On Heavy-Haul Wheel Damages using Vehicle Dynamics Simulation

Saeed Hossein-Nia

Doctoral Thesis
Stockholm, Sweden
2018
To my smart and beautiful daughter Kiana,

Never forget that I love you!
Abstract

Maintenance cost is one of the important issues in railway heavy-haul operations. In most of the cases, these costs are majorly referring to reprofiling and changing the wheels of the locomotives and the wagons. The main reason of the wheel damages is usually severe wear and/or surface initiated rolling contact fatigue (RCF).

This work tries to enhance and improve the knowledge of the wheel wear and RCF prediction models using dynamic simulations. While most of the contents of this study can be generalised to other operational networks, this study is focused on the locomotives and wagons of the Swedish iron-ore company LKAB. The trains are operating on the approximately 500 km long IORE line from Luleå to Narvik in the north of Sweden and Norway respectively.

Firstly, a literature survey of dynamic modelling of the wagons with various three-piece bogie types is presented. Then, with concentrating on the standard three-piece bogies, parameter studies are carried out to find out what the most important reasons of wheel damages are. Moreover, the long-term stability of wheel profiles of the IORE wagons is analysed. This is done by visualising the wear and RCF evolution on the wheel profiles over 150,000km of simulated running distance.

Most of the calculations for the wagons are repeated for the locomotives. However, traction and braking are also considered in the simulation model and their effects on wheel damages are briefly studied. To improve the accuracy of the wheel damage analysis, a newly developed algorithm called FaStrip is used to solve the tangential contact problem instead of FASTSIM. The damage prediction model developed in the thesis is used to study the effects of increasing axle load, correcting the track gauge, limiting the electro-dynamic braking and using a harder wheel material on the wheel life. Furthermore, a new method is developed to predict the running distance between two consecutive reprofilings due to severe surface initiated fatigue. The method is based on shakedown analysis and laboratory tests.

Most of the research works in wear calculation are limited to two approaches known as wear number and Archard methods. The correlation between these two methods is studied. The possibility of using the relation between the two methods for the wear calculation process is investigated mainly to reduce the calculation time for wheel profile optimisation models.

Keywords: wear, RCF, rolling contact, traction, braking, heavy-haul, FASTSIM, FaStrip
Sammanfattning

Underhållskostnad är en av de viktigaste frågorna för tunga järnvägstransporter. I de flesta fall är dessa kostnader i stort sätt begränsade till svarvning och byte av hjul på lok och vagnar. Den främsta orsaken till hjulskadorna är vanligtvis svårt slitage och/eller ytinitierad rullkontaktutmattning (RCF).

Detta arbete syftar till att utöka och förbättra kunskapen om predikteringsmodellerna för hjulslitage och RCF med hjälp av fordonsdynamiska simuleringar. Medan merparten av innehållet i denna studie kan generaliseras till andra typer av järnvägstrafik, är denna studie inriktad på lok och vagnar hos det svenska järnmalmbolaget LKAB. Tågen trafikerar den ca 500 km långa sträckan från Luleå till Narvik på Malmbanan och Ofotbanan i norra Sverige respektive Norge.

Först presenteras en litteraturstudie med fokus på fordonsdynamisk modellering av vagnar med olika typer av tredelade boggier (three-piece bogies). Därefter utförs parameterstudier för att ta reda på de viktigaste orsakerna till hjulskador för de vanligaste boggierna. Dessutom analyseras den långsiktiga stabiliteten hos hjulprofiler för malmvagnar. Detta görs genom att visualisera utveckling av slitage och RCF på hjulprofiler över 150 000 km simulerad körrörelse.


De flesta av undersökningarna av slitageberäkningar är begränsade till två metoder som kallas ”wear number” och Archard-metoder. Korrelationen mellan dessa två metoder studeras. Möjligheten att använda relationen mellan de två metoderna i slitageberäkningsprocessen undersöks främst för att minska beräkningstiden för optimeringsmodeller för hjulprofiler.

Nyckelord:
slitage, RCF, rullkontakt, dragkraft, bromsning, tung transport, FASTSIM, FaStrip
Preface

This doctoral thesis is the summery of my research work at the Department of Aeronautical and Vehicle Engineering, KTH Royal Institute of Technology in Stockholm starting in January 2012.

I gratefully thank LKAB especially the former LKAB employee and the current PhD student Thomas Nordmark for providing the financial and technical support throughout the research.

This thesis would not have been possible without the guides and assists of my supervisors prof. Sebastian Stichel, Dr. Per-Anders Jönsson and assist. Prof. Carlos Casanueva. Sebastian! You have provided a safe environment for me throughout my studies to fail and learn from. This paved my path to where I am now, and I will always be thankful for it.

For the help with the simulations and technical question with GENSYS, I owe my deepest gratefulness to Ingemar Persson at AB DEsolver.

I want to thank all my current and former colleagues at the rail vehicle unit especially Prof. Mats Berg, the head of the division, for providing me an environment to work on the thesis. Mats! I hope one day (soon!) we celebrate championship of Liverpool F.C. in the Premiere League.

I also wish to thank my wonderful parents for all the moral support and the amazing chances they’ve given me over the years.

My better half Mandana! This would not have been possible without you. Your support and encouragement were in the end what made not only this dissertation but also every other achievement in my life possible.

Saeed Hossein-Nia

Stockholm, Nov 2017
Outline of thesis

The thesis is divided into an OVERVIEW part, including an introduction describing the background of the project and the contents of the thesis. The second part of the thesis is APPENDAD PAPERS consists of the following scientific papers written during the PhD work:

Paper A


Paper B


Paper C


Paper D


Paper E


Paper F


Division of work between authors

Saeed H-Nia performed all the simulations, post processing of the results and writing the papers: A, C, D, E and F.

The following Sections of paper B are written by Saeed H-Nia:

- 2.4 Three-piece bogie
- 4.2 Rolling contact fatigue
- 5.2 Rolling contact on wheels of iron ore wagon
Dr. Per-Anders Jönsson, Mr. Thomas Nordmark, Dr. Matin Sh. Sichani and Dr. Carlos Casanueva supervised Saeed H-Nia during the works where their names are mentioned as co-authors, while Prof. Sebastian Stichel was leading the research and supervision throughout the research.

Publications not included in this thesis


S. H-Nia, C. Casanueva and S. Stichel, Prediction of rolling contact fatigue (RCF) for iron-ore locomotive wheels; comparison of an alternative tangential contact solution with FASTSIM. Proc. 25th IAVSD Conf. 2017, Queensland, Australia.

Thesis contribution

This work tries to enhance and improve the knowledge of wheel wear and RCF prediction models. The main contributions of the present work reported in this thesis are as follows:

• The root cause of the RCF problem in the IORE line is investigated via multibody simulation software and it is shown that track degradation, cant deficiency and wheel/rail friction level play an important role in RCF while track stiffness is not a main concern

• Using the concept of the shakedown diagram, a new model to predict and visualise evolution of RCF on the wheel profiles is proposed. The results have good agreement with field observations.

• The effect of traction and braking on RCF of the locomotive wheels is studied.

• Wheel profile evolution of both iron-ore locomotives and wagons due to wear is calculated and the results are in good agreement with field observations.

• A new methodology to predict and visualise evolution of RCF considering partial-slip condition is developed. The results are well matched with field observations.

• It is shown that increasing the axle load by 2.5t slightly escalates the risk of RCF. However, controlling the track gauge significantly reduces the RCF expansion on wheels.

• A newly developed algorithm to solve the tangential contact problem called FaStrip is used in the RCF prediction tool. It is shown that FASTSIM has more deviations from field observations in the nominal wheel-rail radius vicinity than FaStrip. The differences come from the accumulation of the FASTSIM error in the long-term process.

• A new methodology is developed to predict the running distance between two consecutive reprofilings due to RCF using the shakedown theorem and laboratory test results. The results are compared with a five-years field study and good agreement is achieved.
Table of Contents

1 INTRODUCTION ............................................................................................................. 3
  1.1 Dynamic simulation inputs and features .............................................................. 6
    1.1.1 Vehicle models: Wagon ............................................................................. 6
    1.1.2 Vehicle models: Locomotive................................................................. 8
    1.1.3 Track ................................................................................................. 10
    1.1.4 Wheel and rail profiles ................................................................. 14
    1.1.5 Wheel-rail friction coefficient .................................................. 17
  1.2 Validation of the three-piece bogie model ........................................... 18
    1.2.1 Vehicle-based measurement .............................................. 18
    1.2.2 Track-based measurement .................................................. 20

2 CONTACT MECHANICS ............................................................................................... 25
  2.1 Quasi-identity .............................................................................................. 26
  2.2 Hertzian solution ......................................................................................... 27
  2.3 Slip .............................................................................................................. 29
  2.4 Carter’s theory ............................................................................................ 31
  2.5 Kalker’s linear theory ............................................................................... 33
  2.6 Strip theory .................................................................................................. 34
  2.7 FASTSIM .................................................................................................... 35
  2.8 FaStrip .......................................................................................................... 39

3 WHEEL DAMAGE ........................................................................................................ 45
  3.1 Rolling Contact Fatigue (RCF) ................................................................. 46
    3.1.1 Features of inelastic response of metals ....................................... 47
    3.1.2 Hertzian line contact in full-slip (plane strain) ......................... 49
    3.1.3 Hertzian line contact in partial-slip (plane strain) .................. 57
    3.1.4 Applications of the shakedown diagram .................................. 60
    3.1.5 Beyond the shakedown diagram .............................................. 64
  3.2 Wear .............................................................................................................. 68
3.3 Influence of Wear on RCF ................................................................. 74

4 SUMMARY OF THE PAPERS .................................................................. 79
  4.1 Paper A .................................................................................................. 79
  4.2 Paper B .................................................................................................. 79
  4.3 Paper C .................................................................................................. 80
  4.4 Paper D .................................................................................................. 80
  4.5 Paper E .................................................................................................. 81
  4.6 Paper F .................................................................................................. 81

5 FUTURE RESEARCH DIRECTIONS ......................................................... 83
  5.1 Non-elliptical contact ................................................................. 83
  5.2 Slip control and wheel damage .................................................... 84
  5.3 Active control of wheel damage .................................................... 85

6 REFERENCES .......................................................................................... 87
Part I

OVERVIEW
1 Introduction

In heavy-haul rail transportation, there has always been a demand for increasing the transportation capacity by using faster and longer trains carrying more load. The Swedish Mining company, LKAB, with the 500 km IORE (IORE) line is no exception. Since 2004 the axle load increased from around 25t to 30t and the top speed of the vehicles increased by five km/h. However, this increase in capacity is achieved at the expense of escalating the wheel-rail damages, especially due to rolling contact fatigue (RCF). According to a recent field study at least 75% of the wheels of the locomotives are reprofiled after 50,000km due to excessive surface RCF [1], cf. Figure 1-1.

![Figure 1-1](image)

Figure 1-1, Probability of the wheel reprofiling distance due to RCF [1]

LKAB uses several measurement stations on Malmbanan to detect high vertical impact forces, possibly due to wheel flats, where urgent maintenance actions are needed. Besides, regular inspections for wagon wheels are carried out every 80,000km. These inspections are made daily at the yards and more thoroughly after 80,000km in the Kiruna workshop and they are mainly for detecting cracks on the wheels, and thus a possible need for re-profiling. The average mileage interval
1.1. Dynamic simulation inputs and features

between two consecutive wheel turnings for a wagon wheel is around 250,000 km. The rails are checked by the infrastructure owner Trafikverket and are ground once a year. The maintenance policies have led to a total service life of the wagon wheels of 1,000,000 km. Figure 1-2 shows a picture of wagon’s wheel subjected to RCF, see [2].

![Figure 1-2](image1.png)

Figure 1-2, RCF damage on a wagon wheel treads: (a) damage band in an early stage; and (b) developed damage band [2].

Regular inspections for the locomotive wheels are carried out every 26,000 km and the average life of a locomotive wheel is much less than the wagon’s wheel and it is about 400,000 km. Figure 1-3, shows the typical RCF damage on the surface of the loco wheels in a time interval of five years. As it is seen in the figure, the approximate crack size, position and orientation have not changed during this time.

![Figure 1-3](image2.png)

Figure 1-3, RCF damage on two loco wheels, pictures are taken on two distinct dates, the running distance of each wheel is also mentioned in the figure, the pictures are taken from [1].

The most common RCF bands are the first two bands, between (-0.05 mm to -0.02 mm) and (-0.02 mm to 0). The third band at the flange of the wheel is less common compare to the other ones.
CHAPTER 1: Introduction

With respect to wear damage, the loco wheels are rarely affected by wear, this is usually due to short running distance intervals between two consecutive repolishing. However, wagon wheels are more affected by wear, see Figure 1-4. According to specialists in LKAB around 50% of the wagon wheels repolishing is due to wear damage.

![Graph showing wheel profiles](image)

Figure 1-4, The original and a worn wagon wheel profiles after 150,000km of running distance

Both wear and RCF change the wheel profile from its designed geometry in a way that is not desirable for both railway operators and infrastructure owners. Moreover, wheel profile geometry is a crucial factor for the derailment safety.

With severe wheel flange wear, the flange inclination gets too high and the top of the flange might hit switch blades. Also, too high flange inclination increases the conicity of the wheels. This affects the ride stability and reduces the critical speed. With extreme tread wear, however, problems in turnouts or crossings can arise. The flange wear usually depends on the flexibility of the running gear, the curve radii and the wheel-rail friction level, while the wheel tread wear is more a function of axle load, tread braking and, for the locomotives, traction.

To enhance the service life of wheels, LKAB has decided to change the wheel profiles to obtain a better curving performance for the vehicle, and consequently, reduce the creep forces. Moreover, a new wheel material with greater yield strength is tested to achieve a higher resistance against the initiation of cracks. LKAB is also improving the lubrication strategy by moving from flange lubrication towards top of the rail lubrication to control and reduce the friction level between wheel and rail.

To detect and predict the mechanisms of deterioration and sources of damages, besides many other benefits such as investigating the vehicle-
track interaction and studying the track forces, computer simulation is used in the present work.

This thesis is focused on detecting the sources of the wheel damages and predicting the wear and RCF evolution on the wheel profiles as functions of running distance using vehicle dynamic simulations. The wagon simulation model is built at KTH Rail Vehicle Division by the means of the multibody simulation software GENSYS [3], and the simulation results are validated against measurements while the loco model is built by Bombardier Transportation with the multibody simulation software SIMPACK and has been translated into GENSYS by MiW Rail Technology AB.

Some of the features and inputs of the simulation models are discussed in Chapter 0. In Chapter 2, The necessary backgrounds and theories of contact mechanics are reviewed. These theories are used to provide the basis of the RCF and wear calculation methods which are reviewed and described in Chapter 3. Finally, Chapter 4 and Chapter 5 present summaries of the appended papers and ideas for future research direction, respectively.

1.1 Dynamic simulation inputs and features

As it is mentioned above, simulation models of the IORE wagons and locomotives are used in this study to predict the wheel damages of the IORE line vehicles. In this section, some of the needed inputs and features of such models are discussed.

1.1.1 Vehicle models: Wagon

The present IORE wagons are so called Fanoo wagons running on Amsted Motion Control M976 three-piece bogies with load sensitive frictional damping. Fanoo wagons contain two units. One of the units is called master and the other is called slave unit. The controller of the braking system is attached to the master wagon as shown in Figure 1-5.

![Figure 1-5, A two-unit IORE wagon](image-link)
The characteristics of the IORE wagons are given in Table 1-1.

Table 1-1, IORE wagon characteristics (Fanoo)

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Length of wagons</td>
<td>10.29 (m)</td>
</tr>
<tr>
<td>Distance between centre plates</td>
<td>6.77 (m)</td>
</tr>
<tr>
<td>Total wagon height</td>
<td>3.64 (m)</td>
</tr>
<tr>
<td>Basket width</td>
<td>3.49 (m)</td>
</tr>
<tr>
<td>Weight of empty wagon</td>
<td>21.6 (tons)</td>
</tr>
<tr>
<td>Payload</td>
<td>102 (tons)</td>
</tr>
<tr>
<td>Maximum speed (empty)</td>
<td>70 (km/h)</td>
</tr>
<tr>
<td>Maximum speed (loaded)</td>
<td>60 (km/h)</td>
</tr>
<tr>
<td>Wheel base</td>
<td>1778 (mm)</td>
</tr>
<tr>
<td>Wheel diameter (max)</td>
<td>915 (mm)</td>
</tr>
<tr>
<td>Wheel diameter (min)</td>
<td>857 (mm)</td>
</tr>
<tr>
<td>Weight of wheelset (max)</td>
<td>1341 kg</td>
</tr>
<tr>
<td>Weight bogie incl. wheelsets and braking equipment</td>
<td>4 650 kg</td>
</tr>
</tbody>
</table>

Many authors have used multibody simulations to analyse the dynamic behaviour of wagons with three-piece bogies. Using the multibody simulation software GENSYS, Berghuvud developed simulation models for vehicles running on

- a standard three-piece bogie with frictional contact in the primary suspension;
- a typical frame braced bogie;
- a typical inter-axle linkage bogie with cross-braced design.

He compared the curving performance and calculated lateral contact forces of the mentioned three bogies [4] and [5]. Bogojevic also used GENSYS; however, he focused on the standard three-piece bogie with elastic couplings in the primary suspension. He validated the model by comparing simulation results with on-track data [3]. In [6], Orlova presents a simulation model of the Russian 18-100 three-piece bogies using the software MEDYNA. The fundamentals of all mentioned modelling methods are the same. Primary suspension usually is modelled as an elastic stiffness and damping in parallel unless there is no elastic rubber pad between the axle box and the side frame which should be modelled as friction element. The model of the frictional damping depends on its design and whether it is load dependent or not. Figure 1-6 summarises the entire vehicle parts connections in vertical direction. The vehicle in question is a standard three-piece bogie.
1.1. Dynamic simulation inputs and features

![Diagram of vehicle model]

Figure 1-6, Connection between masses in vertical direction, [3].

For more details about modelling and validation of several types of friction damping in three-piece bogies see [7], [8] and [9].

For this study, a very similar model to Bogojevic [3] is used for dynamic simulations of the IORE wagon wheels.

1.1.2 Vehicle models: Locomotive

Due to the confidentiality agreement, no detailed information about the locomotive model can be given. However, a summary of the vehicle description is presented in this section. The IORE locomotive is a two-unite electric Co-Co vehicle hauling 68 wagons. All the axles of the loco are driven and most of the braking effort is carried out by electrodynamic braking system. The bogie model is a Bombardier Transport FLEXX Power 120H designed for a speed of 80km/h and a tractive effort of 300kN per bogie. Some of the characteristics of the locomotive is presented in Table 1-2.
Table 1-2, IORE locomotive characteristics

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Length of loco</td>
<td>22.89 (m)</td>
</tr>
<tr>
<td>Distance between bogie centres</td>
<td>12.89 (m)</td>
</tr>
<tr>
<td>Height</td>
<td>4.465 (m)</td>
</tr>
<tr>
<td>Width</td>
<td>2.95 (m)</td>
</tr>
<tr>
<td>Weight</td>
<td>180 (tons)</td>
</tr>
<tr>
<td>Maximum speed</td>
<td>80 (km/h)</td>
</tr>
<tr>
<td>Wheel base</td>
<td>1920 (mm)</td>
</tr>
<tr>
<td>Wheel diameter (max)</td>
<td>1250 (mm)</td>
</tr>
<tr>
<td>Wheel diameter (min)</td>
<td>1150 (mm)</td>
</tr>
</tbody>
</table>

To simulate the braking and acceleration a PID controller is used to calculate the required torque on gear boxes. The controller goal is to keep the vehicle speed constant while the resistance/gravitational forces applied at the rear of the vehicle try to push/pull the vehicle. These forces are calculated according to the geometrical characteristics of the track sections, weight and speed of the vehicle. Note that there is no limitation for the tractive forces in the simulation. However, as the drivers are told to not use braking forces above 500kN (for both loco units) unless necessary, the limit of the simulated braking force is set to this value. Figure 1-7 shows a photo of an empty IORE vehicle negotiating a curve in mid of March 2009.

Figure 1-7, IORE vehicle; photo by David Gubler/ Copyright: bahnbilder.ch
1.1. Dynamic simulation inputs and features

1.1.3 Track

*The line geometry*

The length of the IORE railway line between Luleå and Narvik is around 500 km including around 50% of curves with radii below 1000m. About 75% of the curves below 450m radius are located on the Norwegian side. The line map is shown in Figure 1-8.

![Figure 1-8, The IORE line [10]](image)

The details of the line geometry are presented in Table 1-3 including the distribution of the left- and right-hand curves respectively.

<table>
<thead>
<tr>
<th>Curve Radius Interval (m)</th>
<th>Mean Radius (m)</th>
<th>Length (%)</th>
<th>Mean Cant (mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;350</td>
<td>305-305</td>
<td>2.3-2.2</td>
<td>68-68</td>
</tr>
<tr>
<td>350:400</td>
<td>386-368</td>
<td>0.7-0.4</td>
<td>66-61</td>
</tr>
<tr>
<td>400:450</td>
<td>412-409</td>
<td>0.6-0.8</td>
<td>63-66</td>
</tr>
<tr>
<td>450:600</td>
<td>559-547</td>
<td>8.4-7.0</td>
<td>54-51</td>
</tr>
<tr>
<td>600:800</td>
<td>649-648</td>
<td>10.1-8.0</td>
<td>45-46</td>
</tr>
<tr>
<td>800:1000</td>
<td>947-934</td>
<td>5.3-4.0</td>
<td>36-37</td>
</tr>
</tbody>
</table>

Heavy axle load and not strong enough rail fastenings have caused widening of the track gauge on the IORE line. The situation is worse where the curve radii are tight and track forces are high as it is shown in Figure 1-9. The track gauge affects the range of the wheel-rail contact patch locations. Therefore, the size and geometry of the patch changes.
CHAPTER 1: Introduction

Figure 1-9, Histogram of the IORE line track gauge (left) and its relation to curve radius for sections with radii less than 900m (right).

**Track stiffness**

For most track sections on the IORE line concrete sleepers are used. Therefore, in this study a track model representing a concrete sleeper track is used within the simulation. This model comprises of ground, sleeper, rails, stiffness and damping between these bodies. For more on track flexibility characteristics and its validation, see [11]. Figure 1-10 is the track deflection measured by Trafikverket (former Banverket) in winter and summer on wooden and concrete sleepers.

Figure 1-10, Vertical rail displacement as a function of vertical static load measured by Trafikverket (former Banverket) [23]

The frozen track is stiffer than the non-frozen one; consequently, in winter tracks are less flexible and have less deflection under the wagon load.
1.1. Dynamic simulation inputs and features

**Track irregularities**

Any deviation from the designed track geometry is known as track irregularities. They have a profound influence on dynamic wheel-rail contact forces and vibration which may cause severe RCF and noise. There are four types of track irregularities that can be captured by the Swedish track recording vehicle STRIX:

- geometrical track errors in vertical direction;
- geometrical track errors in lateral direction;
- deviation from nominal track gauge;
- deviation from nominal cant.

![Figure 1-11, Four types of track irregularities](image)

The vertical and lateral irregularities are calculated as the mean value of errors of left and right rail and they are often called longitudinal level and line irregularities. Figure 1-11 shows these four track irregularities.

Three classes of track qualities are defined regarding the necessity of maintenance and the applicability for acceptance tests of vehicles:

- QN1 refers to the value which necessitates observing the condition of the track or taking maintenance measures as part of regularly planned maintenance operations;
- QN2 refers to the value which requires short term maintenance action;
- QN3 refers to the value which, if exceeded, leads to the track section being excluded from the acceptance analysis because the track quality encountered is not representative of usual quality standards.
Table 1-4, Permissible standard deviations for different speed intervals in longitudinal and line levels for track classes QN1 and QN2 according to UIC Code 518

<table>
<thead>
<tr>
<th>Vehicle speed intervals (km/h)</th>
<th>Standard deviation for longitudinal level (mm)</th>
<th>Standard deviation for line level (mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>QN1</td>
<td>QN2</td>
</tr>
<tr>
<td>0 &lt; v ≤ 80</td>
<td>2.3</td>
<td>2.6</td>
</tr>
<tr>
<td>80 &lt; v ≤ 120</td>
<td>1.8</td>
<td>2.1</td>
</tr>
<tr>
<td>120 &lt; v ≤ 160</td>
<td>1.4</td>
<td>1.7</td>
</tr>
<tr>
<td>160 &lt; v ≤ 200</td>
<td>1.2</td>
<td>1.5</td>
</tr>
<tr>
<td>200 &lt; v ≤ 300</td>
<td>1.0</td>
<td>1.3</td>
</tr>
</tbody>
</table>

The details of track irregularity classification as function of vehicle speed are presented in Table 1-4. The values are calculated for every 100m of the IORE line and the corresponding statistics for track irregularity measured on IORE line in 2012 are presented in Table 1-5.

Table 1-5: Track quality distribution on IORE line according to UIC 518; Maximum speed of 80 km/h assumed. Measurements from 2012.

<table>
<thead>
<tr>
<th>Track quality classes</th>
<th>Definition of the classes</th>
<th>Distribution of each class</th>
<th>Recommended by UIC 518</th>
</tr>
</thead>
<tbody>
<tr>
<td>QN ≤ QN1</td>
<td>Sections with good track standard</td>
<td>29%</td>
<td>Should be &gt; 50%</td>
</tr>
<tr>
<td>QN1 &lt; QN ≤ QN2</td>
<td>Regularly planned maintenance operations</td>
<td>23%</td>
<td>Should be &lt; 40%</td>
</tr>
<tr>
<td>QN2 &lt; QN ≤ QN3</td>
<td>Short term maintenance action</td>
<td>43%</td>
<td>Should be &lt; 10%</td>
</tr>
<tr>
<td>QN &gt; QN3</td>
<td>Sections to be excluded from the analysis</td>
<td>5%</td>
<td>Should be = 0%</td>
</tr>
</tbody>
</table>

The details of the track quality distribution for 2013 track irregularity measurement with respect to the curve section intervals are presented in Table 1-6 for both left- and right-handed curves. As seen in the tables the track quality becomes worse for tighter curves, probably due to the wheel flange contacts with rails.

Table 1-6, Track quality distribution for left-right hand curve, Measurements from 2013.
1.1. Dynamic simulation inputs and features

<table>
<thead>
<tr>
<th>Curve radius interval (m)</th>
<th>Mean Radius (m)</th>
<th>QN1 (%)</th>
<th>QN2 (%)</th>
<th>QN3 (%)</th>
<th>&gt;QN3 (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;350</td>
<td>305-305</td>
<td>40-33</td>
<td>17-25</td>
<td>43-42</td>
<td>0-0</td>
</tr>
<tr>
<td>350:400</td>
<td>386-368</td>
<td>41-70</td>
<td>11-19</td>
<td>47-11</td>
<td>0-0</td>
</tr>
<tr>
<td>400:450</td>
<td>412-409</td>
<td>74-70</td>
<td>11-8</td>
<td>15-22</td>
<td>0-0</td>
</tr>
<tr>
<td>450:600</td>
<td>559-547</td>
<td>79-82</td>
<td>7-10</td>
<td>13-8</td>
<td>1-0</td>
</tr>
<tr>
<td>600:800</td>
<td>649-648</td>
<td>81-80</td>
<td>10-11</td>
<td>9-9</td>
<td>0-0</td>
</tr>
<tr>
<td>800:1000</td>
<td>947-934</td>
<td>79-71</td>
<td>8-8</td>
<td>13-21</td>
<td>0-0</td>
</tr>
</tbody>
</table>

1.1.4 Wheel and rail profiles

GENSYS pre-calculates six wheel-rail contact geometry functions: contact position, wheel rolling radius, contact angle, wheel lift, lateral curvature, lateral position of the contact point on the wheel and lateral position of the contact point on the rail. The program steps from right to left on the wheel profile and calculates the contact point functions for all relative lateral displacements between wheel and rail.

![Figure 1-12, Comparison between the rail profiles (rail inclination: 1/30)](image)

The wheels of the IORE wagons have the so-called WP4 profile and the wheels of the locomotives have the so-called WPL9 profile. The rail profile distributions, however, are more complicated. On the Norwegian side of the line for the curve sections with radii below 600m the higher rail has the MB1 and the lower rail has the MB4 rail profiles. The MB4 rail profile is the standard UIC60 rail with a slight gauge corner relief.
This moves the contact point slightly towards the field side and leads to a higher rolling radius difference and better steering performance. A comparison between the implemented rail profiles of the line is shown in Figure 1-12. The MB1 rail profile is more ground at the gauge corner to avoid any contact and consequently head checking on the corner of the rails. For the track sections having radii above 600m including the straight track sections, the MB4 rail profile is used for both the high and low rails. On the Swedish side of the line for all curve sections the MB1 and UIC60 rail profiles are used for the high and low rails respectively. On straight lines both rails have either MB4 or UIC60 rail profiles. The distribution of the rail profiles along the entire line is shown in Table 1-7.

Table 1-7, The nominal rail profiles along the IORE line

<table>
<thead>
<tr>
<th></th>
<th>Curve radius &lt;600</th>
<th>Curve radius &gt;600</th>
<th>Straight line</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>High rail</td>
<td>Low rail</td>
<td>High rail</td>
</tr>
<tr>
<td>Norway</td>
<td>MB1_assymetric</td>
<td>MB4</td>
<td>MB4</td>
</tr>
<tr>
<td>Sweden</td>
<td>MB1</td>
<td>UIC60</td>
<td>MB1</td>
</tr>
</tbody>
</table>

Any combination of the mentioned wheel profiles (WP4 and WPL9) and each of the mentioned rail profiles produces a unique wheel-rail geometry function. Here the equivalent conicities for each combination mentioned in Table 1-7 for the nominal track gauge (1435mm), and the average track gauge of the line (1445mm) are compared. As shown in Figure 1-13, at nominal track gauge the highest equivalent conicity is achieved if both rails have the UIC60 rail profile while at the average track gauge of the line (1445mm) using MB4 rail profiles for both rails gives the highest equivalent conicity. Generally, higher equivalent conicity leads to a greater rolling radius difference and a better steering performance in curves. A wheel-rail profile combination with low values of equivalent conicity usually helps the stability of the vehicle at higher speeds.
1.1. Dynamic simulation inputs and features

Figure 1-13, Equivalent conicity of the wheel-rail profile combinations as a function of lateral displacement (mm) of the wheelset.

It is also possible to see the contact positions on the wheel and rail depending on the lateral displacement of the wheelset. As it is seen in Figure 1-14, UIC60 gives the most distributed contact positions on the wheel and rail and the gauge corner relief MB1 rail profile gives the most condensed contact positions. Moreover, both the MB4 and MB1 rail profiles give two-point contacts at the outer wheels in tight curves since there is no material at the corner of the rails. This, on one hand, will divide the load on two different contact points and avoids the concentrated load on one point leading to safer transferring the load through the fastenings and to the ballast, especially when the fastenings are not strong enough or not positioned correctly. Having two-point contacts on the outer wheel, on the other hand, decreases the steering ability. Since the directions of the longitudinal creep forces would be different on the two contact patches the total resulting steering forces decrease. Consequently, higher attack angles are predicted in these cases.
Figure 1-14, Contact point positions on wheel and rail for relative displacements: combinations of WP4 wheel and UIC60, Mb4 and Mb1 rail profiles.

1.1.5 Wheel-rail friction coefficient

The proportionality factor of the normal load and the resistance force in sliding is called the friction coefficient. The dynamic behaviour of the vehicle, among many other things, strongly depends on the coefficient of friction. On one hand, a high friction coefficient is needed for adequate adhesion conditions and better radial steering. On the other hand, higher friction leads to larger creep forces and will increase the
risk of RCF. The coefficient of friction greatly depends on the weather, rail temperature, passing axle number and tribological surface conditions like roughness, hardness etc. It may vary dramatically during a day from morning to night. In [12] the proportionality of the friction level to the rail temperature (among other parameters like humidity of the weather) is experimentally investigated. Usually from very humid to extremely dry weather, the friction coefficient can vary from 0.2 to 0.75. In cold dry climate conditions, as the water content in air reduces significantly, the wheel-rail friction coefficient increases.

1.2 Validation of the three-piece bogie model

The simulation model against the available measured forces and accelerations is validated. There are two types of measurement data available from the IORE line. One is the track-based strain gauge measurement mostly for online monitoring the vehicle performance to detect wheel failures. The other one is the measurement via accelerometers attached to the vehicle. The latter type of measurement has been performed by Interfleet Technology during the summer of 2004 and the winter of 2011.

1.2.1 Vehicle-based measurement

Bogojevic [3], compared the calculated and measured carbody accelerations in vertical and lateral direction for both empty and loaded conditions. Here some of the results are presented, cf. Figure 1-15 and Figure 1-16.

Both the measured and calculated results are filtered with a cut-off frequency of 20 Hz. However, by including the sleeper-passing frequency, the model can be valid up to around 35 Hz, cf. Figure 1-17.
Figure 1-15, Forces at the wheel-rail contact in the vertical (top) lateral (bottom) direction on a 650m curve on the outer wheel: time histories (left) and CDF (right) [13].

Figure 1-16, Measured (top) and simulated (bottom) vertical acceleration of empty wagon above leading bogie on tangent track [3].
1.2. Validation of the three-piece bogie model

Figure 1-17, PSD of the measured and simulated vertical track force on the right wheel of the first wheelset in the leading bogie [14].

1.2.2 Track-based measurement

There are three measurement stations along the IORE line. Two of them are in Sweden near Sävast outside Luleå and Tornehamn, and the third one is in Norway in Haugfjell. The measurement system consists of several strain gauge sensors. These sensors measure the strain in the rail in lateral and vertical direction. The strain is converted to the track forces after proper calibration. Figure 1-18 shows waterproof strain gauge measurement equipment and its sensor locations on the rail at Sävast measurement station. For more information about the details of the measurement station, see [10].

Figure 1-18, Measurement equipment and the sensor locations on the rail, [10].

To compare simulation results and the measurement data from Sävast the actual track geometry information of the section is used as the simulation inputs [15].

The length of the section is around 550m and the curve is left-hand. The rail profiles in the simulation are the profiles measured with the MiniProf equipment before and after grinding together with the nominal UIC60 and MB1 profiles.
Both the worn and nominal WP4 wheel profiles are considered in the simulations. Figure 1-4 shows the difference between the profiles. The characteristics of the wheel profiles are shown in Table 1-8, which are in the range of the measured profile characteristics that passed the measurement station during a year [10].

Table 1-8, Wheel profile characteristics

<table>
<thead>
<tr>
<th></th>
<th>Flange thickness</th>
<th>Flange height</th>
<th>Flange gradient,</th>
</tr>
</thead>
<tbody>
<tr>
<td>New wheel</td>
<td>26.89 (mm)</td>
<td>29.08 (mm)</td>
<td>10.07 (mm)</td>
</tr>
<tr>
<td>Worn wheel</td>
<td>26.78 (mm)</td>
<td>30.42 (mm)</td>
<td>9 (mm)</td>
</tr>
</tbody>
</table>

The minimum, maximum and nominal values of the speed and the load of the vehicles that passed the section for one year are chosen for the simulation input cf. Figure 1-19.

Figure 1-19, The speed of the loaded wagons (left), the vertical load of the loaded wagons (right), [10].

A parameter study is carried out on the wheel-rail friction coefficient that varied between 0.2 and 0.5 to simulate the weather conditions during a year.

As the measurement results contain also high frequent dynamic forces (cut of frequency is 100Hz), the contribution of the high frequent dynamic forces as suggested in [16] is calculated and the results are compared with the measurements.

Figure 1-20 shows the range of all calculated $Q$ forces and Figure 1-21 shows the range of the calculated $Y$ forces in comparison with the measured ones.
1.2. Validation of the three-piece bogie model

Figure 1-20, Measured and calculated vertical forces. a) left wheel, b) right wheel. The blue and red lines show the maximum and minimum of the calculated forces while the background grey diagrams show the distribution of the measured forces in a year [15].

Figure 1-21, Measured and calculated lateral forces. The blue and red lines show the maximum and minimum of the range of the calculated forces while the background diagrams show the distribution of the measured forces in a year.
As can be seen in the figures there is a good agreement between the simulation results and the measurements. The highest deviation occurred at the outer wheel of the leading bogie. The vehicle usually runs with flange contact at these conditions. Usually the flange contact (especially for worn wheels and rails with more conforming contact) violates the half-space assumption in the Hertzian theory which may lead to non-realistic results.
2 Contact Mechanics

The focus of this thesis is to predict and estimate wheel damages. These damages start to initiate at the wheel-rail contact area. Thus, naturally, theories of wheel-rail contact mechanics which is used in the damage calculation process should be reviewed and discussed first.

In solid mechanics, the term “contact mechanics” is the knowledge of mathematically modelling the deformation and stresses arising where surfaces of two solid bodies are brought into contact. Many of engineering problems depends on the theories of contact mechanics. In railway engineering, the multibody simulation (MBS) software use these theories to calculate wheel-rail contact forces. Moreover, the contact theories are used to predict and model several types of wheel-rail damages e.g. RCF, wear etc.

In this chapter, some of the fundamentals of such knowledge is presented. Firstly, a general contact mechanic problem is discussed. Then, the importance of Boussinesq equations and the assumption of the quasi identical materials are addressed. Moreover, the Hertzian solution to the normal contact problem is studied. Before jumping to the existing solutions of the complete rolling contact problem (including the tangential contact problem) the fundamentals of slip, creepage and spin are explained. Later in the chapter, some of the most commonly used theories of rolling contact i.e. Carter, Strip, Kalker’s linear and simplified theories are presented. As it will be discussed, for general 3-D contact, there is no closed solutions for the rolling contact problem and using numerical methods is inevitable. Therefore, at first, the numerical implementation of the Kalker’s simplified theory (FASTSIM) is described and finally a recently published alternative method to FASTSIM is illustrated, i.e. FaStrip.

Briefly, the contact problem is to find the tangential and the normal pressure distribution and size and shape of the contact area for either known deformations or known loads or a combination of these when two bodies of arbitrary surface come into contact at interface. At the
contact surface may exist friction or the contact is frictionless. The material of the bodies in contact can be elastic or inelastic. Finally, the nature of the contact can be non-conformal (small patch) or conformal (large patch). In railway applications, the conformal contact may occur on worn wheel and rail profiles.

2.1 Quasi-identity

Considering the compatibility relation between deformation and strain, the Hooke’s law of stress and strain relation at the surface can be re-written as

\[ u_i(x, y) = \int_C A_{i,z}(x, y)p(x, y)dc, \quad i = x, y, z \]  

(2.1)

Where, \( u_i(x, y) \) is the deformations due to the contact pressure \( p(x, y) \) in \( x, y \), and \( z \) direction. \( A_{i,z} \) is called the influence function which relates the applied pressure in \( z \) direction to the deformations in \( x, y \), and \( z \) direction. The integral is over the total contact area \( C \).

![Figure 2-1, Defined coordinate system [1]](image)

Based on the Boussinesq [18] (and Cerruti [19]) formula, \( A_{i,z}(x, y) \) in Equation (2.1) can be calculated for elastic half-space condition. The influence function depends on material properties and geometry of surfaces and its derivation for point load is reported in [20]. Finally, the displacements in three directions due to normal load are presented in Equations (2.2 - 2.4).

\[ u_x(x, y) = \frac{K}{\pi G} \int_C p(\xi, \eta) \left( \frac{\xi-x}{R} \right) d\xi \, d\eta \]  

(2.2)

\[ u_y(x, y) = \frac{K}{\pi G} \int_C p(\xi, \eta) \left( \frac{\eta-y}{R} \right) d\xi \, d\eta \]  

(2.3)
\[ u_z(x, y) = \frac{1-v}{\pi G} \int_C \left( \frac{p(\xi, \eta)}{R} \right) d\xi \, d\eta \quad (2.4) \]

where, \( K \) is a material and \( R \) is a geometry function respectively. Note that indices 1 and 2 refer to the surfaces in contact, \( G \) is the material shear modulus and \( v \) is the Poisson ratio.

\[
K = \left( \frac{G}{4} \right) \left\{ \frac{1-2\nu_1}{G_1} - \frac{1-2\nu_2}{G_2} \right\} \quad (2.5)
\]

\[
\frac{1}{\bar{G}} = \frac{1}{2} \left( \frac{1}{G_1} + \frac{1}{G_2} \right) \quad (2.6)
\]

\[
R = \sqrt{(\xi - x)^2 + (\eta - y)^2} \quad (2.7)
\]

As it is seen in Equations (2.2 and 2.3), normal load \( p \) causes tangential components \( u_x \) and \( u_y \) which means that the normal and tangential contact problems are coupled. However, if one assumes that the materials of the surfaces are quasi-identical in which

\[
\frac{1-2\nu_1}{G_1} = \frac{1-2\nu_2}{G_2}, \quad (2.8)
\]

the material function \( K \) (i.e. Equation (2.5)) vanishes and the displacements \( u_x \) and \( u_y \) due to the normal load at the surface \( (z = 0) \) will be zero and the normal problem is decoupled from the tangential part. In an analogous way, it can be shown that in presence of tangential traction, where materials are quasi-identical, the displacements in \( x \) and \( y \) direction are not influenced by the normal pressure and the displacement in \( z \) direction is not influenced by tangential tractions [21]. In conclusion, the quasi-identity condition plays a key role in the contact mechanics since it makes it possible to solve the normal and tangential contact problems separately.

### 2.2 Hertzian solution

Some believe that the knowledge of the contact mechanics is started in 1882 when Henrich Hertz a young researcher at the University of Berlin first published his paper “On the contact of elastic solids” [22]. His theory provides a closed-form solution to the normal contact problem.
2.2. Hertzian solution

Nowadays almost all MBS software use Hertz theory to solve normal the contact problem.

Hertz theory is based on the following assumptions:

- displacements and strains are small;
- the bodies in contact are homogeneous, isotropic, and linearly elastic;
- the surfaces are smooth and surface roughness is neglected i.e. frictionless contact;
- dimensions of the contact patch are significantly smaller than the dimensions of the bodies in contact. Moreover, the dimensions of the patch are smaller than the relative radii of the bodies in contact. Therefore, the stresses can then approximately be calculated with the assumption that the contact partners are semi-infinite bodies limited by a straight plane. This is so called half-space assumption;
- the curvature of the bodies in contact are constant;
- the materials of the bodies in contact are quasi-identical.

Although, the assumptions of Hertz theory are generally valid for most of the railway applications, in few cases some of them could be violated. For example, linear elasticity is not valid in case of very high loads (above the elastic shakedown) where some degree of plastic flow could take place. This also affects the size and shape of the patch. Moreover, the half-space assumption is doubtful where wheel and rail have a conformal contact, which can occur in the flange root but also due to wheel profile wear.

Hertz assumed an elliptical pressure distribution over an elliptic contact area with semi-axes $a$ and $b$, see Figure 2-2.

![Figure 2-2, Contact pressure distribution in contact patch [23].](image)

The pressure distribution is
\[ p(x, y) = p_0 \sqrt{1 - \left(\frac{x}{a}\right)^2 - \left(\frac{y}{b}\right)^2}, \]  
\[ p_0 = \frac{3N}{2\pi ab}, \]  

where, \( N \) is the normal force applied to the patch. Moreover, the semi-axes \( a \) and \( b \) can be calculated as

\[ a = m \sqrt{\frac{3N(1-v^2)}{2E(A+B)}} \]  
\[ b = n \sqrt{\frac{3N(1-v^2)}{2E(A+B)}} \]

where, \( A \) and \( B \) are the geometry functions related to the relative longitudinal and lateral curvatures of the bodies in contact respectively.
Knowing the relative curvatures, one can use the tables presented in [23] to find the constants \( m \) and \( n \).

### 2.3 Slip

In rigid body dynamics where the contact patch is as small as a point, if a free wheelset rolls down on a surface by its weight, there will be a constant friction force between the wheelset and the rail opposing the direction of the rolling. Now, if the wheel-rail friction coefficient is higher than the quotient between the normal force in the contact and the friction force (Coulomb’s law), the wheelset stays in the rolling condition, in which the relative velocity between the wheel and rail at the point of contact is zero. However, if the friction coefficient is less than the bound, there will be a sliding motion too. This is called rolling with sliding, where, there exists a relative non-zero velocity between the bodies at the point of contact. Sliding more or less always exists in railway vehicles as the suspension prevents the wheelset from pure rolling and conical shape of the wheel and rail. The direction of the relative velocity is different in traction and braking.

In reality, particles at the contact surface undergo a deformation and they displace relative to each other. If the work done on these particles is less than the recoverable strain energy (elasticity) the particles stick together (adhesion area) and if the work exceeds the limit they slide over each other. This is called micro-slip [24].
2.3. Slip

The slip velocity in the contact patch contains the contributions of the rigid body motions and the elastic deformation, see Equation (2.11).

$$S = C - \left( \frac{\partial u(x,t)}{\partial x} \right) V + \frac{\partial u(x,t)}{\partial t} \quad (2.11)$$

where, $C$ is the relative rolling velocity between the bodies (Creep), $u$ is the displacement of the particles and $V$ is the rolling speed. If we divide $S$ by $V$ in steady state (where, $\frac{\partial u(x,t)}{\partial t} = 0$) and expand the creep term, then the non-dimensional slip velocities in $x$ and $y$ directions are

$$\frac{s_x}{V} = v_x - \phi y - \frac{\partial u_x}{\partial x},$$

$$\frac{s_y}{V} = v_y + \phi x - \frac{\partial u_y}{\partial x} \quad (2.12)$$

$v_x$ and $v_y$ are the creepages in longitudinal and lateral direction respectively, and $\phi$ is the so-called spin. The full derivation of the spin and creepages can be found in [23] and it is briefly presented in [21]. Here the most dominant parts of these terms are discussed.

The lateral creepage depends on the yaw angle $\Psi$ and the lateral velocity of the wheelset $\dot{y}$

$$v_y = \frac{\dot{y}}{V} - \Psi \quad (2.13)$$

The Longitudinal creepage is determined by

- difference of the rolling radii;
- rotation of the wheelset about the $z$ axis;
- braking or acceleration.

Note that the signs of the longitudinal creepages are different on the left and the right wheels (in absence of braking or acceleration)

$$v_x = \frac{-\omega(R_{left}-R_{right})+\dot{\Psi}b+(V-\omega R_{mean})}{V} \quad (2.14)$$

Here, $b$ is half the distance between the contact points on the left and the right wheels, $\omega$ is the angular velocity of the wheelset and $R$ is the radius of the wheel.

Finally, the spin is a function of geometric and kinematic spin
\[ \varphi = \frac{\Psi \cos \gamma}{V} \pm \frac{\sin \gamma}{R_0}, \quad (2.15) \]

where, \( \gamma \) is the contact angle and \( R_0 \) is the actual rolling radius.

### 2.4 Carter’s theory

Carter published his work on running quality of electric locomotives in 1926 [25]. He was looking for a relationship between creepage and creep forces. He reduced the contact problem into a plane strain by considering the contact of two cylinders with parallel axes. Then, he divided the contact patch into two distinct stick and slip zones. In pure longitudinal creepage, the shear stress \( q_x(x) \) increases nonlinearly from the leading edge of the patch until it reaches the boundary \( \mu p(x) \) and it follows it to the trailing edge where \( \mu p(x) = 0 \). Thus, the derivation of the shear stress in the slip zone \( q'_{x} \) is rather straightforward

\[ q'_{x} = \mu p = \mu p_0 \sqrt{1 - \left( \frac{x}{a} \right)^2}, \quad (2.16) \]

where, \( p_0 \) is the maximum normal pressure at the centre of the contact ellipse and \( a \) is the half-length of the contact area. Moreover, Carter proposed that the shear stress distribution in the adhesion area can be obtained by subtracting a similar elliptic traction distribution \( q''_{x} \) from the traction bound \( q'_{x} \).

\[ q''_{x} = -\frac{d}{a} \mu p_0 \sqrt{1 - \left( \frac{x+(a-d)}{d} \right)^2}, \quad (2.17) \]

where, \( d \) is the half-length of the contact area in the stick zone. Consequently, the total shear stress distribution is

\[ q = q'_{x} + q''_{x}. \quad (2.18) \]

The half-length \( d \) of the stick area can be obtained as a function of longitudinal creepage and the length of the contact patch [24], see Equation 2.19.

\[ d = a - \frac{\nu_x E a}{4 \mu p_0 (1-\nu^2)} \quad (2.19) \]
2.4. Carter’s theory

Finally, integration of Equation (2.18) leads to the total shear force as

\[ Q = \mu \left( 1 - \left( \frac{d}{a} \right)^2 \right). \]  \hspace{1cm} (2.20)

The shear stress distributions according to Equations (2.16-2.18) for friction coefficient 0.5 is presented in Figure 2-3.

![Figure 2-3, Shear stress distribution according to Carter’s theory for \( \mu = 0.50 \) and \( Q/P = 0.2 \)](image)

The magnitude of the creep force can be simply derived by integrating from the total shear stress distribution function, i.e. Equation (2.18) and it is presented in Equation (2.21) in non-dimensional form. For Hertzian distribution of the normal pressure for line contact the normalised creep forces are plotted against the normalised longitudinal creepage, see Figure 2-4.

\[ \frac{Q}{\mu N} = f(x) = \begin{cases} 
-\left( 2 + \frac{R}{\mu a} |v_x| \right) \frac{Rv_x}{\mu a}, & \frac{Rv_x}{\mu a} < 1 \\
-\frac{v_x}{|v_x|}, & \frac{Rv_x}{\mu a} \geq 1 
\end{cases} \]  \hspace{1cm} (2.21)

where, \( \frac{1}{R} \) is the relative curvature of the wheel and rail cylinders.


2.5 Kalker’s linear theory

Kalker’s linear theory is based on the fact that in vanishing creepages the relationship between the creep force and creepage is linear [26]. It assumes that where the friction value becomes infinite this linearity remains. It is also based on the boundary condition that the stress at the leading edge is zero. This leads to a theory that the shear stress continues to grow boundlessly from zero at the leading edge. The creep force and spin moment can be obtained by integrating the shear stress over the patch, see Equation (2.22)

\[
F_x = -C_{11} G c^2 v_x, \\
F_y = -C_{22} G c^2 v_y - C_{22} G c^3 \phi, \\
M_y = -C_{23} G c^3 v_y - C_{33} G c^4 \phi, 
\]

where, \(c = \sqrt{ab}\) and \(C_{ij}\) are the so called Kalker’s coefficients. The coefficients are tabulated and can be found in [21]. The coefficients are valid for dry friction which according to Kalker corresponds to a friction coefficient \(\mu = 0.6\). Kalker’s linear theory is valid only in very small creepages where roughly (see [23])

\[
abs(v) + abs(\phi/1000) \leq 0.002. 
\]

Therefore, it is rarely used in railway applications. However, it has a significant role in derivations of other theories in rolling contact.
2.6 **Strip theory**

Haines & Ollerton extended the two-dimensional Carter’s theory into a three-dimensional problem by considering two spheres in contact for the longitudinal creepage case [27]. In this theory, the contact patch is divided into separate strips in the rolling direction and each strip is considered as a stand-alone plane strain problem. Kalker extended the strip theory by considering lateral creepage and limited spin [28]. The longitudinal and lateral stress distributions in the stick zone according to Kalker are

\[ q_x(x, y) = \frac{\mu p_0}{a_0} \left[ \kappa_k \sqrt{a^2(y) - x^2} - \kappa'_k \sqrt{(a(y) - d(y))^2 - (x - d(y))^2} \right], \]

\[ q_y(x, y) = \frac{\mu p_0}{a_0} \left[ \lambda_k \sqrt{a^2(y) - x^2} - \lambda'_k \sqrt{(a(y) - d(y))^2 - (x - d(y))^2} \right], \tag{2.24} \]

where, \( p_0 \) is the maximum Hertzian pressure and half-length of the slip zone and the half-length of the patch in any strip as a function of lateral coordinate are \( d(y) \) and \( a(y) \) respectively. Also,

\[ a(y) = a_0 \sqrt{1 - \left( \frac{y}{p_0} \right)^2}, \tag{2.25} \]

\[ d(y) = \left( \frac{a_0}{1 - \nu'} \right) \frac{\sqrt{\eta^2 + (1 - \psi^2)\xi^2 + \psi \tilde{\eta} \xi}}{(1 - \psi^2)}, \tag{2.26} \]

\[ \kappa_k = \kappa'_k = \frac{\tilde{\xi}}{\xi^2 + \tilde{\eta}^2} \left[ -\tilde{\eta} \psi + \sqrt{\tilde{\eta}^2 + (1 - \psi^2)\tilde{\xi}^2} \right], \tag{2.27} \]

\[ \lambda_k = \frac{\tilde{\eta}}{\xi^2 + \tilde{\eta}^2} \left[ -\tilde{\eta} \psi + \sqrt{\tilde{\eta}^2 + (1 - \psi^2)\tilde{\xi}^2} \right] + \psi, \tag{2.28} \]

\[ \lambda'_k = \frac{\tilde{\eta}}{\xi^2 + \tilde{\eta}^2} \left[ -\tilde{\eta} \psi + \sqrt{\tilde{\eta}^2 + (1 - \psi^2)\tilde{\xi}^2} \right], \tag{2.29} \]

where, \( \nu' \) is the equivalent Poisson’s ratio and \( \tilde{\xi}, \tilde{\eta} \) and \( \psi \) are non-dimensional terms as functions of creepages see Equation (2.30).
\\begin{equation}
\begin{aligned}
\xi &= \frac{-G}{2\mu p_0} \nu_x - \frac{\psi y}{a_0}, \\
\eta &= \frac{-G}{2\mu p_0} (1 - \nu) \nu_y, \\
\psi &= \frac{-G}{2\mu p_0} a_0 \varphi.
\end{aligned}
\end{equation}

(2.30)

The shear stresses in the slip zone depend on whether the strip is in full-slip or partial slip. In the case of partial slip, the shear stresses are

\[ q_x(x, y) = \frac{\mu p_0}{a_0} \frac{\xi}{\sqrt{\eta^2 + (1 - \psi^2)^2 + \xi^2}} \sqrt{a^2(y) - x^2}, \]

(2.31)

\[ q_y(x, y) = \frac{\mu p_0}{a_0} \frac{\eta + \sqrt{\eta^2 + (1 - \psi^2)^2 + \xi^2}}{\sqrt{\eta^2 + \xi^2}} \sqrt{a^2(y) - x^2}, \]

and in the strips in full-slip

\[ q_x(x, y) = \frac{\mu p_0}{a_0} \frac{\xi}{\sqrt{(\eta^2 (1 - \nu)) + \xi^2}} \sqrt{a^2(y) - x^2}, \]

(2.32)

\[ q_y(x, y) = \frac{\mu p_0}{a_0} \frac{\eta + \sqrt{\eta^2 + (1 - \psi^2)^2 + \xi^2}}{\sqrt{(\eta^2 (1 - \nu)) + \xi^2}} \sqrt{a^2(y) - x^2}. \]

Kalker’s strip theory is applicable for a limited range of spin $|\psi| < 1$. The other limitation of this theory is the contact patch size. Its accuracy significantly reduces in large values of semi-axes ratio $\frac{a_0}{b_0}$ of the contact area. Sichani [17] showed that in the case of pure spin and circular contact the relative error in force estimation with respect to CONTACT [29] is up to 40%.

2.7 FASTSIM

Kalker’s complete theory [30], which is implemented in CONTACT, provides a rigorous solution to the elastic half-space rolling contact problem. However, it is computationally demanding and time consuming. Thus, it is barely used in dynamics simulations. Therefore, Kalker introduced his simple and fast numerical algorithm (FASTSIM)
to overcome this problem. FASTSIM is very common in simulation software since it can cover all variety of combinations of creepages and spin, and it is about 1000 times faster than CONTACT. Also, accuracy of FASTSIM in estimation of the contact forces is acceptable. This is since its relative error in a worst-case scenario (combinations of lateral creepage and spin) can rise to 20% [31]. Note that for dynamic simulations the contact forces have decisive role rather than the distribution of the contact stresses.

To begin with the FASTSIM algorithm one should mention Kalker’s simplified theory [32] first. In fact, FASTSIM is the numerical implementation of Kalker’s simplified theory to cope with rolling contact problems containing spin. The simplified method is based on the Winkler bed theory in the normal direction and a similar wire brush model for the tangential direction. In which an elastic foundation is introduced as a mattress (brush) with springs independent of each other in a way that their deformation will not affect their neighbouring springs. Therefore, a linear relation can be maintained between the deformation $u(x, y, z)$ and the stresses $\tau(q_x, q_y, p_z)$ as

$$ u = L\tau, \quad (2.33) $$

where, $L$ is the flexibility parameter and can be compared to the inverse of the elasticity module. However, unlike $E$, it depends on the shape of the bodies and the loading. Since, a closed solution for the vertical direction already exists (Hertz), Equation (2.33) is not used in $z$ direction. However, using the simplified theory in the vertical direction gives a parabolic distribution of the normal pressure unlike the elliptic Hertzian distribution.

In the tangential direction and in full adhesion, $(S_x, S_y) = 0$, integrating from Equation (2.12) holds

$$ u_x = v_x x - \varphi xy + k(y), $$
$$ u_y = v_y x + \varphi \frac{x^2}{2} + l(y), \quad (2.34) $$

where, $k$ and $l$ are integral constants and are functions of $y$. Applying the boundary conditions $q_x(a(y), y) = q_y(a(y), y) = 0$ at the leading edge the shear stress distributions will be
\[ \begin{align*}
q_x &= \frac{1}{L} \left[ v_x(x - a(y)) - \varphi(x - a(y))y \right], \\
q_y &= \frac{1}{L} \left[ v_y(x - a(y)) + \varphi \frac{(x^2 - a^2(y))}{2} y \right].
\end{align*} \tag{2.35} \]

Therefore, the forces in \( x \) and \( y \) directions are
\[ \begin{align*}
F_x &= -\frac{8a^2 b v_x}{3L}, \\
F_y &= -\frac{8a^2 b v_y}{3L} - \frac{\pi a^2 b \varphi}{4L}.
\end{align*} \tag{2.36} \]

Equating the coefficients of longitudinal, lateral and spin creepages from Equation (2.36) by the ones achieved by Kalker’s linear theory, i.e. Equations (2.22), three distinct flexibility parameters are possible to derive
\[ \begin{align*}
L_x &= \frac{8a}{3Gc_{11}}, \quad L_y = \frac{8a}{3Gc_{22}}, \quad L_\varphi = \frac{\pi a^2}{4Gc_{23}}. \tag{2.37} \end{align*} \]

Kalker also introduced a single value of the flexibility parameter \( L \) as a weighted average of the three parameters gives in Equation (2.37).
\[ L = \frac{L_x |v_x| + L_y |v_y| + L_\varphi |\varphi| c}{\sqrt{v_x^2 + v_y^2 + ab \varphi^2}} \tag{2.38} \]

In the linear theory, the traction starts to grow linearly from zero at the leading edge and it suddenly vanishes at the trailing edge. However, in the simplified theory, the traction rises until it reaches the traction bound where the slip takes place and the shear stress continues with the bound until reaching the trailing edge. The traction bound is calculated according to the Coulomb’s law. Kalker realised that using the elliptical traction bound resulting from the Hertzian solution has “grave defects”. Therefore, he tried the parabolic solution from the simplified theory in the normal direction, compare Equations (2.39 and 2.9).
\[ \begin{align*}
N &= \iint p_0' \left\{ 1 - \left( \frac{x}{a} \right)^2 - \left( \frac{y}{b} \right)^2 \right\} \, dx \, dy = \frac{\pi ab}{2} \, p_0, \\
p_0' &= \frac{2N}{\pi ab}.
\end{align*} \tag{2.39} \]
Thus, the parabolic traction bound will be

\[ g = \mu p'(x, y) = \mu \frac{2N}{\pi ab} \left( 1 - \left( \frac{x}{a} \right)^2 - \left( \frac{y}{b} \right)^2 \right). \tag{2.40} \]

Using Equation (2.40), Kalker could improve his simplified theory in mainly two aspects (see [17]):

- more accurate stick-slip boundaries;
- more accurate level of forces.

Kalker justified the use of parabolic traction bound by indicating “Once simplified theory, always simplified theory”.

In the absence of spin the simplified theory provides a theoretical solution. However, in the presence of spin a numerical algorithm is needed. Kalker called this numerical implementation FASTSIM algorithm.

In FASTSIM, first the elliptic contact patch is discretised into longitudinal strips and then each strip is divided into rectangular elements. The shear stress values of each element are given by the Equation (2.41).

\[
\begin{align*}
\bar{q}^{n+1}_x &= q^n_x - \left( \frac{v_x}{L_x} - \frac{q_y}{L_q} \right) dx, \\
\bar{q}^{n+1}_y &= q^n_y - \left( \frac{v_y}{L_y} - \frac{q_x}{L_q} \right) dx,
\end{align*}
\tag{2.41}
\]

where, \( n \) is counter of elements in the contact patch. As it is mentioned, the total shear stress magnitude \( \bar{q}_t = \sqrt{\bar{q}_x^2 + \bar{q}_y^2} \) grows from the leading edge until it reaches the parabolic traction bound (\( g \)) in Equation (2.40). From this point, \((\bar{q}_t \geq g)\), the shear stresses calculated from Equation (2.41) must be reduced proportionally with respect to the magnitude of the traction bound in that element, as shown in Equation (2.42). As an example, the creep-creepage curves for two cases with and without spin are shown in Figure 2-5.

\[ q^{n+1}_x = \frac{\bar{q}^{n+1}_x}{\bar{q}_t} G \tag{2.42} \]
\[ q_y^{n+1} = \frac{q_y^{n+1}}{q_t} \]

Figure 2-5, Comparison of FASTSIM and CONTACT creep-creepage curves for the case of no spin (top) and the case of pure spin (bottom) [21].

2.8 FaStrip

Although the accuracy of FASTSIM in force estimation is satisfactory, its shear stress estimation error compared to CONTACT may rise to 33% at the point of contact and reaches infinity close to the boundaries [33]. This is mainly due to the parabolic traction bound where its higher values of the shear stress in the slip area (compared to the Hertzian elliptic traction bound) compensate the errors in the linear stick zone which results in better estimation of the force (integral of the shear stress). However, many wheel-rail damage analysis models use the surface shear stress as the basis of their methodologies e.g. [34]-[36]. To cope with this issue a novel algorithm is proposed by Sichani [33] at KTH which significantly improves the accuracy of FASTSIM while it
is as fast as FASTSIM. The methodology is based on the properties of Strip theory and FASTSIM algorithm. Therefore, he called the algorithm Fast-Strip or in short FaStrip. In other words, FaStrip is the modified Kalker Strip theory using FASTSIM properties.

As it is mentioned in Section 2.6, the Strip theory has two main drawbacks:

- it is not accurate for \( \frac{a_0}{b_0} \geq 1 \);
- it is not accurate in high spin.

To solve the first issue, Sichani equated the resulting integrals of the shear stresses in the stick area from the Strip theory i.e. Equation (2.24) to the forces from the Kalker’s linear theory Equation (2.22) in vanishing creepages and spin and suggested the following adjustment in the parameters of Equations (2.30):

\[
\begin{align*}
\xi &= \left( \frac{-G}{2 \mu p_0} \nu_x \right) \left( \frac{4(1-\nu)}{\pi^2} c_{11} \right), \\
\eta &= \left( \frac{-G}{2 \mu p_0} (1 - \nu) \nu_y \right) \left( \frac{4}{\pi^2} c_{22} \right), \\
\psi &= \left( \frac{-G}{2 \mu p_0} a_0 \phi \right) \left( \frac{3}{\pi} \frac{b_0}{a_0} c_{23} \right)
\end{align*}
\] (2.43)

Using the shear stress of the stick zone from the Strip theory results in a more accurate non-linear growth of the shear stresses from the leading edge as in CONTACT. Furthermore, the adjustment in Equation (2.43) results in a better slope of the creep-creepage curve for higher values of \( \frac{a_0}{b_0} \). Compare the top and bottom plots of Figure 2-6 to see the improvement of slope of the creep-creepage curve as well as the stick-slip area.
Figure 2-6, Creep-creepage curve (left) and shear stress distribution (right) in pure longitudinal creepage $\nu_x = 0.02$, estimated by the Strip theory (top) and the adjusted Strip theory (bottom) for semi-axes $a_0 = 5\text{mm}$ and $b_0 = 1\text{mm}$ from [33].

In the slip area $\left(x \leq -a(y) + 2d(y)\right)$, see Equations (2.25 and 2.26), the traction bound is elliptic

$$q_t(x,y) = \sqrt{q_x^2 + q_y^2} = \mu p_0 \sqrt{a(y)^2 - x^2}. \quad (2.44)$$

Using the elliptic traction bound results in an accurate estimation of the total shear stress $q_t$, however, it is shown that the parabolic distribution results in more accurate shear stress direction i.e. magnitudes of $q_x$ and $q_y$ in the slip region. Therefore, Sichani used the FASTSIM algorithm to find the stress directions. He also noticed that using the weighted average flexibility parameter of Equation (2.38) results in better estimation of the creep forces in the falling part of the creep-creepage curve in higher values of creepages. Therefore, he adjusted the FASTSIM part in such that for any strip in full-slip, the flexibility parameter of Equation (2.38) is used instead of the ones in Equation (2.37). The final adjustment is due to the error of the creep forces in the
2.8. FaStrip

falling part of the creep-creepage curve. Sichani noticed that his proposed model underestimates the creep forces in higher spin values for smaller semi-axes ratios \( \frac{a_0}{b_0} < 1 \), as shown in Figure 2-7 (left). Therefore, he made the following adjustment in normal force of the FASTSIM part in the presence of spin.

\[
N_f = \begin{cases} 
\left[ 1 + \left( 1 - \frac{a_0}{b_0} \right)^4 \right] N, & \frac{a_0}{b_0} < 1 \\
N & \frac{a_0}{b_0} \geq 1
\end{cases}
\]  

(2.45)

As it is seen in Figure 2-7 (right), the latest change, significantly improved the accuracy of the lateral forces in the case of pure spin.

![Figure 2-7](image)

Figure 2-7, Lateral creep curves of pure spin case for different semi-axes ratios of contact patch estimated by the proposed method using prescribed normal force (left) and modified normal force of the Equation (right) compared to FASTSIM and Contact, from [33].

The equations of the FASTSIM part (direction of the stresses in the slip zone) are the same as the ones presented in Section 2.7 with the mentioned adjustments in this Section. Finally, the FaStrip estimation of the shear stresses in the slip area is

\[
q_x(x, y) = r_x \frac{\mu p_0}{a_0} \sqrt{a(y)^2 - x^2},
q_y(x, y) = r_y \frac{\mu p_0}{a_0} \sqrt{a(y)^2 - x^2},
\]

(2.46)

where, \( r_x = \left( \frac{q_x}{q_{tx}} \right)_{FASTSIM} \) and \( r_y = \left( \frac{q_y}{q_{ty}} \right)_{FASTSIM} \) are the correction factors from the FASTSIM part of the algorithm to improve the shear stress directions in the slip zone.
In conclusion, Equation (2.24), modified by Equation (2.43) for the stick zone, and Equation (2.46) for the slip zone construct the FaStrip algorithm.

It is possible to calculate the elastic contribution, $\left(\frac{\partial u_x}{\partial x}\right)$ and $\left(\frac{\partial u_y}{\partial x}\right)$, of the relative slip velocity, i.e. Equation (2.12), based on the FaStrip stress distributions, see [33]. The shear stress distribution and the relative velocity of CONTACT, FaStrip and FASTSIM, calculated for a tread (left) and a flange (right) contact points are presented in Figure 2-8.

Figure 2-8. Comparison of shear stresses and relative velocity of CONTACT, FaStrip and FASTSIM, calculated for a tread (left) and a flange (right) contact points.
3 Wheel Damage

From the tribology point of view, damage to a solid surface involving progressive loss of material and relocation of material when two surfaces are interacting via a relative motion, is called wear. How a material wears depends not only on the nature of the material but also on other elements of the tribo-system such as geometry of contacting pairs, surface topography, loading, lubrication, and environment. The mechanisms causing such damage usually are complicated and most of the time it is not possible to distinguish one from another. In [37], approximately, 60 terms describing wear behaviour and mechanisms are listed. Some of the most important wheel-rail related mechanisms are listed below.

- **Abrasive wear**: wear caused by rough and hard surfaces sliding on each other or wear caused by hard particles trapped between two surfaces like hard oxide debris;
- **Adhesive wear**: wear caused by shearing of junctions formed between two contacting surfaces, sometimes used as a synonym for dry sliding wear;
- **Chemical wear** (Corrosive wear): wear caused by formation of any oxide or other components on surfaces due to chemical reaction of the surfaces with the environment;
- **Erosive wear**: Wear due to relative motion of contact surfaces while a fluid containing solid particles is between the surfaces;
- **Rolling contact fatigue (RCF)**: caused by cyclic stress variations leading to fatigue of the materials. Generally resulting in the formation of sub-surface and deep-surface cracks, material pitting and spalling.

Kimura [38] studied adhesive wear and RCF and concluded that both phenomena have elemental processes in common. In this study, the term “wear” is used for adhesive wear otherwise the complete term of the mechanism is mentioned. Note that the term “mechanical wear” is also used for abrasion, erosion, and adhesion.
3.1. Rolling Contact Fatigue (RCF)

In this Chapter firstly, RCF mechanisms and prediction models are discussed. Then, some of the dominant theories around wear calculation and wheel profile evolution are presented and finally, the effect of wear on RCF is reviewed.

3.1 Rolling Contact Fatigue (RCF)

It is common among railway engineers and researchers to categorise RCF of the wheels based on the distance of the initiation of the cracks from the surface of the wheel (e.g. see [39]) as:

- surface initiated fatigue;
- sub-surface initiated fatigue (~3-10mm from the surface);
- deep-surface initiated fatigue (~10-25mm from the surface).

The sub-surface and deep surface cracks are usually rare; however, they are the most dangerous ones compared to the surface cracks and may cause derailments. Both phenomena occur where there exists a locally low fatigue resistance of the material such as microscopic manganese sulphide inclusions (sub-surface) and voids (deep-surface). For typical features of fatigue initiated at sub-surface and deep-surface see [40].

The surface initiated cracks, however, are easy to cope with and usually can be treated by repolishing the wheel profiles or even automatically by natural wear [41]. However, frequent repolishing is costly, especially when it interrupts the train operation. Moreover, it may escalate the track degradation caused by high impact forces. Therefore, many researchers have tried to analyse the roots of surface initiated RCF and developed models to predict, visualize and prevent it.

The focus of this section is to review the basis and theories behind such models. We start with the stress analysis both below and at the surface of the wheel-rail contact. Then, the behaviour of the material above the yield limit, plasticity, under the cycling loads will be discussed. Such behaviours include the concept of shakedown, the role of residual stresses and strain hardening. The examples of this section are mainly restricted to 2-D plane strain of Hertzian line contacts with sufficiently small strains. However, a conclusion in point contact (3-D) will be presented.
3.1.1 Features of inelastic response of metals

The stress-strain graph of a metal such as steel from a simple tension and compression test is well-known. The most important parts of this diagram are shown in Figure 3-1.

![Stress-Strain Graph](image)

Figure 3-1, General material response in tension and compression

There are several mathematical theories to model such material behaviour, which each can be useful based on the problem demands. Some of these models are:

- Perfectly plastic: equal yield in tension and compression;
- Isotropic hardening: equal strain hardening in tension and compression;
- Linear kinematic hardening: constant range between yield in tension and compression;
- Non-linear kinematic hardening: contains an elastic response below the yield limit and non-linear hardening rate in plastic region.

Among the few models mentioned here, perfectly plastic models cannot show the “hardening” effect. Isotropic hardening is not able to model the “Bauschinger” effect (specimen deformed in compression and then loaded in tension will start to deform plastically in lower tensile stress). This leads to an ultimate elastic response after a few cycles of loadings. And finally, the linear kinematic hardening is not able to model the cyclic “creep” [42] and leads to a closed cycle of stress-strain relationship. In solid mechanics, creep is the small and slow plastic deformation under a constant stress. If a material undergoes a cyclic creep deformation it will be damaged sooner or later. This accumulative plastic strain in railway engineering is called “ratchetting” or
3.1. Rolling Contact Fatigue (RCF)

incremental collapse [21]. In fact, the surface initiated RCF is caused by the ratchetting phenomenon. This is shown schematically in Figure 3-2. However, not always the material subjected to a loading above the yield limit ratchets. In the first few cycles after the elastic yield limit is reached, the material hardens and in the presence of residual stresses it may end with a totally elastic response. This is the so called “shakedown” behaviour since the material shakes down to a lower state of deformation.

![Figure 3-2, Schematic of ratchetting or incremental collapse](image)

Before entering the shakedown concept in detail, some basics of the two most popular yield criteria (onset of plasticity) should be discussed.

The first one is the Tresca yield criterion. According to Tresca, a material subjected to tension or compression starts to plastically deform where its maximum shear stress (principal shear stress): \( (\tau_1) \) reaches the yield stress in shear \( k \). Equation (3.1) shows this as functions of principal stresses. Note that at yield, the shear stress is half the corresponding normal stress for centric uniaxial loading test, \( k = \frac{\sigma_y}{2} \).

\[
\tau_1 = \frac{1}{2} \max \{|\sigma_1 - \sigma_2|, |\sigma_2 - \sigma_3|, |\sigma_3 - \sigma_1|\} = \frac{\sigma_y}{2} = k, \tag{3.1}
\]

where, \( \sigma_1, \sigma_2 \) and \( \sigma_3 \) are the principal stresses.

The other well-known yield criterion is the Von-Mises yield criterion in which

\[
(\sigma_1 - \sigma_2)^2 + (\sigma_2 - \sigma_3)^2 + (\sigma_3 - \sigma_1)^2 = 2\sigma_y^2 = 6k^2. \tag{3.2}
\]
Tests suggest that the Von-Mises criterion provides a slightly better fit to experiments than Tresca. However, the differences are very small, and it is never more than 15% [21]. For simplicity, the Tresca criterion is used in this study.

### 3.1.2 Hertzian line contact in full-slip (plane strain)

Under a cylinder with a radius $R$ rolling and/or sliding in full-slip condition over an elastic half space (see Figure 3-3), the interface patch will be modelled as a line of contact with the half-length of $a$.

According to Hertz (cf. Section 2.2) the normal pressure is

\[ p(x) = \frac{2P}{\pi a^2} (a^2 - x^2)^{1/2}, \]  

(3.3)

where,

\[ P = \pi a^2 E^*/4R, \quad a^2 = \frac{4PR}{\pi E^*} \quad \text{and} \quad E^* = \frac{E}{2(1-\nu^2)} \]

(3.4)

\[ p_0 = \frac{2P}{\pi a} = 4 \frac{P}{\pi R} = \left( \frac{PE^*}{\pi R} \right)^{1/2} \]

in which $E^*$ is the equivalent modulus of elasticity. The general state of the stress for such case under rolling and sliding with the friction level $\mu$ is given by [24]. The stresses due to the normal pressure $p$ are:
3.1. Rolling Contact Fatigue (RCF)

\[
\begin{align*}
(\sigma_x)_p &= -\frac{p_0}{a} \left\{ m \left( 1 + \frac{z^2 + n^2}{m^2 + n^2} \right) - 2z \right\}, \\
(\sigma_z)_p &= -\frac{p_0}{a} \left\{ m \left( 1 - \frac{z^2 + n^2}{m^2 + n^2} \right) \right\}, \\
(\tau_{xz})_p &= -\frac{p_0}{a} n \left\{ \frac{m^2 - z^2}{m^2 + n^2} \right\}.
\end{align*}
\]

where \((\sigma_x)_p\) and \((\sigma_z)_p\) are the stresses in longitudinal and vertical directions, \((\tau_{xz})_p\) is the shear stress due to the normal force and \(m\) and \(n\) are given by Equation (3.6).

\[
\begin{align*}
m^2 &= \frac{1}{2} \left\{ (a^2 - x^2 + z^2)^2 + 4x^2z^2 \right\}^{1/2} + (a^2 - x^2 + z^2) \\
n^2 &= \frac{1}{2} \left\{ (a^2 - x^2 + z^2)^2 + 4x^2z^2 \right\}^{1/2} - (a^2 - x^2 + z^2)
\end{align*}
\]

In the presence of friction, the stress distributions due to the tangential traction \(q(x)\) will be

\[
\begin{align*}
q(x) &= \frac{2\mu P}{\pi a^2} (a^2 - x^2)^{1/2} \\
(\sigma_x)_q &= \frac{q_0}{a} \left\{ n \left( 2 - \frac{z^2 - m^2}{m^2 + n^2} \right) - 2x \right\} \\
(\sigma_z)_q &= q_0 \left\{ \frac{\tau_{xz}}{p_0} \right\} \\
(\tau_{xz})_q &= q_0 \left\{ \frac{(\sigma_x)}{p_0} \right\}
\end{align*}
\]

Note that the relation between \(q(x)\) and the normal pressure is provided by Ollerton and Haines [27]. The total state of the stresses can be obtained by superposition of Equations (3.5 and 3.7).

\[
\begin{align*}
\sigma_x &= (\sigma_x)_p + (\sigma_x)_q \\
\sigma_z &= (\sigma_z)_p + (\sigma_z)_q \\
\tau_{xz} &= (\tau_{xz})_p + (\tau_{xz})_q
\end{align*}
\]

Finally, the principal shear stress will be
\[ \tau_1 = \frac{1}{2} ((\sigma_x - \sigma_z)^2 + 4\tau_{xz}^2)^{\frac{1}{2}} \]

The contours of normalised principal shear stress \( \tau_1/p_0 \) below and along the surface under rolling with no friction is shown in Figure 3-4. As it can be seen from the figure, the maximum shear stress is located below surface at \( z = -0.78a \) and the line \( x = 0 \) is the symmetry line. As it is briefly mentioned in Section 3.1, this case is safe against the surface initiated fatigue.

![Figure 3-4, Contours of the principal shear stress beneath a rolling (\( \mu = 0 \)) Hertzian line contact.](image)

However, the condition is different under a sliding contact both at the surface and below the surface.

At the surface (\( z = 0 \)), due to the tangential traction \( q(x) \), \( \sigma_x \) reaches its maximum compressive stress \( -2q_0 \) at the leading edge of the contact area (\( x = -a \)) and a maximum tension \( 2q_0 \) at the trailing edge (\( x = a \)). It should be noted that the normal pressure at the surface leads to an equal stress at the surface. Equation (3.10) shows the surface stresses at \( z = 0 \).

\[
\begin{align*}
\sigma_x(x, 0) &= -2\mu x / a \\
\sigma_z(x, 0) &= 0 \\
\tau_{xz}(x, 0) &= -q(x) = \mu p_0 \left(1 - \left( \frac{x}{a} \right)^2 \right)^{1/2} \\
\tau_1(x, 0) &= \mu p_0
\end{align*}
\]
3.1. Rolling Contact Fatigue (RCF)

The surface stresses due to frictional traction \( q(x) \) are shown in Figure 3-5.

![Figure 3-5](image)

Figure 3-5, Surface stresses for sliding contact at the presence of friction cf. Equation (3.10).

Below the surface, the tangential traction \( q(x) \) leads to not only an increase in magnitude of maximum principal shear stress but also it moves it closer to the surface. This strongly depends on the friction values. Figure 3-6 shows the contours of the normalised principal shear stress in two distinct friction values. As the figure depicts, the maximum value of the principal shear stress increases with around 13% from the case with \( \mu = 0.1 \) to the case with \( \mu = 0.3 \). Also, its location moves towards the surface and the trailing edge (\( x = a \)) of the contact patch.

![Figure 3-6](image)

Figure 3-6, Contours of the principal shear stress beneath a rolling Hertzian line contact for \( \mu = 0.1 \) (left) and \( \mu = 0.3 \) (right)

The change of the location of the maximum principal shear stress as a function of friction coefficient is shown in Figure 3-7. As it is seen in the figure, at a friction value \( \mu \approx 0.37 \) the maximum value of \( \tau_1 \) will be
located at the surface of the contact where its magnitude is the same at the entire contact patch with \( \tau_1 = \mu p_0 \) due to the full-slip condition cf. Equation (3.10).

![Figure 3-7](image)

Figure 3-7, Location of the maximum principal shear stress as a function of friction values.

According to Tresca, cf. Equation (3.1), yield occurs wherever the applied load leads to a condition that the magnitudes of \( \tau_1 \) reaches the yield strength in shear \( k \). This can be shown with a line in a dimensionless map where its axis are friction values and normalised vertical load, see Figure 3-8. Such a line indicates that in a certain friction level if the magnitudes of \( p_0 / k \) do not pass the limit, yield does not take place and the response of the material will be elastic. Therefore, this line is called the “elastic limit” [42] or “load limit” [43]. And the corresponding map is called the “shakedown map” or “shakedown diagram”.
3.1. Rolling Contact Fatigue (RCF)

For the loads above the elastic limit, some plastic deformation takes place in the first few cycles and therefore three possible changes may occur:

- residual stresses form below the surface;
- material strain hardens;
- the contact area may expand and may results in conforming contact.

As the conformal contacts violate the half-space assumption, we do not discuss them here. However, the residual stresses and the hardening phenomenon will be briefly reviewed.

After the first few cycles the material is subjected to a combination of residual and contact stresses which together may not reach the yield limit. This means that the system shakes down to a resultant condition which is entirely elastic [44]. The corresponding limit line is the so called “shakedown limit”, calculated using the Von-Mises criterion and tabulated for various friction levels in both line and point contacts in [43]. However, Johnson [24] suggests a simpler approach based on Melan’s theorem.

According to Johnson some assumptions including plane strain and steady-state rolling must be made for considering the effect of residual stresses on the shakedown diagram.

The Tresca criterion holds (cf. Equations (3.1) and (3.9)):
\[ \frac{1}{4} \{ \sigma_x + (\sigma_x)_r - \sigma_z \}^2 + \tau_{xz}^2 \leq k^2 \] (3.11)

where, \((\sigma_x)_r\) is the only remaining residual stress due to Melan’s assumptions. According to Equation (3.11), \(\tau_{xz}\) cannot be higher than the yield limit in shear \(k\) (avoiding infinite plastic strain). However, it is possible to have \(\tau_{zx} = k\), when choosing the residual stress in the way that \((\sigma_x - \sigma_z) = (\sigma_x)_r\). Thus, in analogy to the elastic limit, wherever under or at the surface of a material the shear stress \(\tau_{zx}\) exceeds the yield limit in shear, plastic flow occurs. As no hardening effect is considered, the corresponding shakedown limit is limited to elastic-perfectly plastic materials, see Figure 3-8.

For kinematically hardening material the calculations are more complicated. Ponter [45] introduces a fictitious residual stress \(\sigma_r^*\) which is the combination of back stress (coordinate of the centre of the yield surface) and residual stresses. He re-stated Melan’s theorem with considering the first two mentioned assumptions. Ponter describes that shakedown will occur if there exists \(\sigma_r^*\) in combination with the distribution of stresses in which the yield will not be exceeded. In other words, if at the given load no system of fictitious residual stresses can be found that satisfy the Ponter’s theorem, plastic deformation occurs in every load cycle. This plastic deformation is either in the form of low cycle fatigue with a closed loop of deformation or in the form of ductile fracture with cyclic creep known as ratchetting. The shakedown limit for kinematically hardening material is provided by Johnson in [46]. Moreover, Johnson [47] suggests a graphical approach to obtain the shakedown limit using the Ponter’s theorem. He suggests to re-write Equation (3.11) as

\[ \left( \frac{\sigma_x - \sigma_z}{2p_0} \right)^2 + \left( \frac{\tau_{xz}}{p_0} \right)^2 = \left( \frac{k}{p_0} \right)^2. \] (3.12)

Equation (3.12) is clearly in a form of a circle equation. It is possible to plot Equation (3.12) in a diagram with the horizontal axis \(\left( \frac{\sigma_x - \sigma_z}{2p_0} \right)\) and the vertical axis \(\left( \frac{\tau_{xz}}{p_0} \right)\) in all the depths, \(z\), of the contact patch. Thus, the radius of obtained circles will be \(\left( \frac{k}{p_0} \right)\) which is the inverse of the shakedown limit at given friction values. This is shown in Figure 3-9 at \(\mu = 0.15\) for a few depths. Johnson shows that the smallest circle which

---

CHAPTER 3: Wheel Damage
3.1. Rolling Contact Fatigue (RCF)

circumscribes all the other circles with the radius \( \left( \frac{k}{p_0} \right) \) in any depths from the centre (o) of the stress trajectories shown in Figure 3-9 leads to the elastic limit of the shakedown diagram.

![Diagram showing stress trajectories at \( \mu = 0.15 \).](image)

Figure 3-9, Stress trajectories at \( \mu = 0.15 \) in few depths.

Johnson also shows that for elastic-perfectly plastic materials in the presence of residual stresses \( (\sigma_x)_r \) the centre of the diagram can move horizontally leading to a smaller circle (o’) with the same condition. Finally, for kinematically hardening materials according to Ponter in the presence of \( (\sigma_x)_r^* \) and \( (\tau_{xz})_r^* \), the centre of the circle can move freely and consequently leading to a circle (o’’) with even smaller radius. These are all shown in Figure 3-9 and for convenience some of the shakedown limits for kinematically hardening material are provided in Table 3-1.

Table 3-1, Shakedown limit values for kinematically hardening material.

<table>
<thead>
<tr>
<th>( \mu )</th>
<th>0.00</th>
<th>0.10</th>
<th>0.15</th>
<th>0.20</th>
<th>0.25</th>
<th>0.28</th>
<th>0.30</th>
<th>0.35</th>
<th>0.40</th>
<th>0.50</th>
<th>0.60</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \frac{p_0}{k} )</td>
<td>4.00</td>
<td>3.93</td>
<td>3.85</td>
<td>3.72</td>
<td>3.56</td>
<td>3.42</td>
<td>3.30</td>
<td>2.85</td>
<td>2.50</td>
<td>2.00</td>
<td>1.67</td>
</tr>
</tbody>
</table>

When yield occurs first at the surface of the contact patch (at \( \mu \approx 0.37 \)), the elastic limit and the shakedown limit are equal. In this case, the shakedown limit is equal to \( 1/\mu \) since at the surface \( \tau_1 = \mu p_0 = k \).
3.1.3 Hertzian line contact in partial-slip (plane strain)

In many cases in railway applications full-slip condition is not the case and the contact is in partial-slip. For low friction in which the maximum shear stress locates beneath the surface, the shakedown diagram does not differ drastically between the partial and full-slip conditions. This is since the sub-surface stresses are not influenced significantly by the surface tractions [47]. However, the situation at the surface is quite different and the stress distributions for the partial-slip deviate from the full-slip condition.

The stresses at the surface for the partial-slip case are presented in Section (2.4) for the 2-D line contact. Using Equations (2.18) and (3.9), the relative principal shear stresses for both the contact in partial-slip with traction coefficient \( \frac{Q}{P} = 0.4 \) and the contact in full-slip with friction level \( \mu = 0.5 \) are presented in Figure 3-10. The figure shows that surprisingly, for the case with partial-slip, the maximum principal stress is higher than for the contact in full-slip condition.

![Figure 3-10](image)

**Figure 3-10.** Principal shear stress at the surface: partial-slip with \( \frac{Q}{P} = 0.4 \) vs. full-slip for \( \mu = 0.5 \).

The methodology behind the calculation of the elastic limit is the same as what is mentioned in Section (3.1.2). However, it is not possible to have a general shakedown diagram for all friction levels as in Figure 3-8. This is due to the dependency of the principal stresses on the traction coefficient (Q/P). Thus, for each friction level the elastic limit (and the shakedown limits) should be uniquely calculated as a function of (Q/P). This is shown in Figure 3-11 for \( \mu = 0.5 \). It should be noted that to calculate the elastic limit for a certain friction level, the
longitudinal creepage $v_x$ is increased from zero until the full-slip condition is achieved.

![Shakedown diagram: line contact partial-slip for $\mu = 0.5$.](image)

Figure 3-11, Shakedown diagram: line contact partial-slip for $\mu = 0.5$.

To calculate the shakedown limit considering the residual stresses and the strain hardening of the material, a graphical approach is suggested by Johnson [47] similar to what he has suggested for the full-slip condition. The stress trajectories in analogy to Figure 3-9 are presented in Figure 3-12 (right) for the partial-slip case for again $\mu = 0.5$ and $Q/p_0 = 0.4$. Moreover, the corresponding relative surface shear stress $q/p_0 = \tau_{zx}/p_0$ (cf. Equation (2.18)) is shown in Figure 3-12 (left). Johnson, first calculated the corresponding non-zero stresses at points A, B, and C of Figure 3-12 (left) as:

\[
A(x = -a): \quad \frac{\sigma_x}{\mu p_0} = \frac{2(a-d)}{a},
\]

\[
B(x = 2d - a): \quad \begin{cases} 
\frac{\sigma_x}{\mu p_0} = \frac{2(a-d)}{a} \\
\frac{\tau_{zx}}{\mu p_0} = -2 \left( \frac{a-d}{a} \left( 1 - \frac{a-d}{a} \right) \right)^{\frac{1}{2}} 
\end{cases}
\]

\[
C(x = a): \quad \frac{\sigma_x}{\mu p_0} = 2 \frac{a-d}{a} - 4 \left( \frac{a-d}{a} \right)^{\frac{1}{2}}.
\]
Figure 3-12, Surface shear stress (left) and stress trajectories (right) for partial-slip condition for \( \mu = 0.5 \) and \( \frac{Q}{P} = 0.4 \).

Then, he transferred them into a diagram with horizontal axis \( \frac{\sigma_x - \sigma_z}{2\mu P_0} \) and vertical axis \( \frac{\tau_{xz}}{\mu P_0} \), see Figure 3-12 (right). For a perfectly plastic material considering the effect of residual stresses the centre O moves horizontally to O’, and therefore, from the geometry of the figure and Equation (3.13), the shakedown limit is given by

\[
\frac{p_0^s}{k} = \mu \left( \frac{a-d}{a} \right)^{\frac{1}{2}} \left( 2 - \frac{a-d}{a} \right)^{-1}, \tag{3.14}
\]

and for the kinematically hardening material as the centre O can move freely to O’’ the shakedown limit will be

\[
\frac{p_0^s}{k} = \mu^{-1} \left( \left( \frac{a-d}{a} \right)^{\frac{1}{2}} \left( 2 - \frac{a-d}{a} \right) \right)^{\frac{1}{2}}. \tag{3.15}
\]

Using Equations (3.14) and (3.15) the shakedown limits are plotted in Figure 3-11 for \( \mu = 0.5 \) together with the full-slip shakedown limit for kinematically hardening material. From the figure, it can be concluded that for a fixed traction Q/P plastic flow is more likely to occur when operating under partial-slip condition.

Everything in this chapter so far was based on the calculations for line contact. However, obviously the 3-D point contact is a better approximation for the wheel-rail contact. The calculations for the point contact are much more complicated, and the approximate results presented by Johnson [24] for the full-slip condition show that the
3.1. Rolling Contact Fatigue (RCF)

Stresses are not as severe as for the line contact. The shakedown limit in the absence of friction is at $\frac{P_0^s}{k} = 4.7$. Therefore, the line contact results are a more conservative measure to use. An approximate shakedown diagram for point contact in full-slip is presented in Figure 3-13 (left).

3.1.4 Applications of the shakedown diagram

The Shakedown diagram maybe is the most common tool to detect RCF in railway applications. For a heavy haul-locomotive simulation model, collecting all working points of hundreds of simulations with various operational conditions (vehicle speed, friction level, track geometry etc.), one can construct the shakedown diagram as a density function as in Figure 3-13 (right). From the figure, it can be concluded that the risk of RCF is quite high for the locomotive wheels.

![Figure 3-13, Approximate shakedown diagram for point contact in full-slip (left) and corresponding simulation results for various operational cases after 50,000km simulated running distance (right).](image)

There is another approach to use the shakedown diagram in railway applications. Perhaps, it was first Ekberg who proposed an engineering model for RCF risk assessment based on an approximate shakedown diagram for point contact in full-slip. His model can predict the surface initiated RCF. Such cracks take place when the friction values are quite high (approximately above 0.3). The shakedown limit in this case is an inverse of the traction coefficient, see Section (3.1.2). Moreover, he calculates the horizontal distance of any working point (WP) from this limit as $FI_{surf}$ (Surface Fatigue Index).
\[ F_{I_{surf}} = \frac{Q}{p} - \frac{k}{p_0} = \frac{\sqrt{Q_x^2 + Q_y^2}}{p} - \frac{2\pi a \cdot b \cdot k}{3p}. \] (3.16)

Obviously, positive values of \( F_{I_{surf}} \) represent the ratchetting part of the shakedown diagram resulting in surface initiated fatigue.

All parameters of Equation (3.16) are available as outputs of MBS software. Thus, it is not surprising to see that almost all MBS software use this method to calculate the indication of the surface cracks. As an example, some of the results of two simulation cases of a heavy-haul locomotive operating on a moderate curve section and straight line using the MBS software GENSYS are shown in Table 3-2.

Table 3-2, Simulation results of \( F_{I_{surf}} \) values for two different cases for heavy-haul locomotive (wheel tread)

<table>
<thead>
<tr>
<th>a(m)</th>
<th>b(m)</th>
<th>P(N)</th>
<th>( \mu )</th>
<th>Curve radius (m)</th>
<th>( F_{I_{surf}} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0075</td>
<td>0.0029</td>
<td>142668</td>
<td>0.3700</td>
<td>550</td>
<td>0.2590</td>
</tr>
<tr>
<td>0.0071</td>
<td>0.0039</td>
<td>147886</td>
<td>0.3000</td>
<td>Tangent</td>
<td>-0.0511</td>
</tr>
</tbody>
</table>

As the table shows, for the case operating on a curve section there is a positive \( F_{I_{surf}} \) value and an indication of surface fatigue.

As it is mentioned above, \( F_{I_{surf}} \) as a function of simulation time is usually available as an output of an MBS software as shown schematically in Figure 3-14.

![Figure 3-14, Schematic output of an MBS software for \( F_{I_{surf}} \) values](image-url)
3.1. Rolling Contact Fatigue (RCF)

The area below the $F_{lsurf}$ curve represents the depth of the ratchetting working points and can be used as an index of RCF severity. Figure 3-15, shows the higher risk of RCF in higher friction coefficients. The data are obtained from the simulation results of IORE wagons with three-piece bogies.

![Image of Figure 3-15](image)

Figure 3-15, Area below the $F_{lsurf}$ curve over the maximum possible area at friction coefficient. 0.75 [48]

Moreover, the number of the working points with positive fatigue index over the whole number of the working points can be used as RCF probability. As an example, see Figure 3-16 which shows an inverse relation of the RCF risk to the cant deficiency for relatively tight curves for heavy-haul wagons.

![Image of Figure 3-16](image)

Figure 3-16, RCF probability vs Track Cant [48].

The other very useful application of the shakedown diagram is to visualise the surface crack locations on the wheel profile. It is possible
to find the position of the working points with positive $F_{I_{surf}}$ values on the wheel profile. For Figure 3-17, the wheel profile is discretized into five-millimetre intervals and the points are placed in each corresponding interval. Moreover, the number of the located simulation points subjected to RCF regarding their positions on the wheel profile is shown in the figure. It is obvious that the positions with higher numbers of simulation points subjected to RCF are more at risk regarding surface initiated cracks.

Figure 3-17, (a) Shakedown diagram for one simulation case and (b): the corresponding position of the simulation points with risk of RCF (R=409 m) [49].

One can show the accumulation of the ratchet working points on the wheel profiles after several thousand kilometres of continuous running distance. In this way, the evolution of expansion of surface cracks can be visualised. As an example, see Figure 3-18 to compare the effect of lubrication on evolution of RCF at the surface of an IORE locomotive wheel after 35’000 km of simulated running distance.
3.1. Rolling Contact Fatigue (RCF)

Figure 3-18, Evolution of RCF on the first axle for a locomotive wheel profile (a) lubricated and (b) non-lubricated. The darker the area the more severe the RCF [49].

Another way to visualise the expansion of the cracks on the wheel surface is to show the accumulation of ratchet working points as a function of longitudinal and lateral position of the wheels, see Figure 3-19. In this case, as the shakedown diagram is calculated for the full-slip condition, it is assumed that the entire contact patch is affected by RCF.

Figure 3-19, Accumulated RCF number on a wheel profile after 50,000km running distance; calculated for full-slip condition, the max. value of the colour-bar is set to 300,000.

Spangenberg et al. [50] suggest using the shortest distance from the working point on the shakedown diagram \( F_{I_{surf}} \) instead of the horizontal distance. This is to avoid the diagram limitation in very high traction and low normalised load.

3.1.5 Beyond the shakedown diagram

The shakedown diagram despite its many advantages has some limitations. The \( F_{I_{surf}} \) tool as it is used in MBS software is only valid for full-slip condition; however, as shown in Section (3.1.3) the corresponding shakedown diagram for partial-slip condition is significantly different in higher values of friction level. A comparison is made by Dirks in [51] between the \( F_{I_{surf}} \) values obtained in full-slip condition (distributed elliptically over the patch) and in partial slip condition (using FASTSIM algorithm). In the mentioned study, the full-slip values are called “FI global” and the partial-slip ones are called “FI local”, Equation (3.17).
As described in Section (2.8), the FASTSIM algorithm uses parabolic traction distribution which results in higher stress values. Moreover, as mentioned in Section (3.1.3) higher values of shear stress at the surface are expected in the partial-slip case (also see Figure 3-10). Therefore, it is not surprising to have higher fatigue index values for the local case. Figure 3-20 shows the FIsurf values calculated in a relatively tight curve (radius: 450m) for the IORE locomotive. Here, the maximum FIsurf values in longitudinal direction for each lateral position in the patch have been taken as the FIsurf for that specific lateral position on the wheel profile. The figure also shows the comparison of the obtained local and global values obtained with the FASTSIM algorithm and the FIsurf values obtained with the FaStrip algorithm. As mentioned in Section (2.8), lower values of surface traction (more accurate) result in lower FIsurf values for the FaStrip algorithm, see Figure 3-20.

![Figure 3-20](image)

Figure 3-20, FIsurf values for global, local: FASTSIM and local FaStrip algorithm for a simulation of an IORE locomotive in a curve with 450 m radius and friction level $\mu \approx 0.40$.

The other limitation of the shakedown diagram is that it is only capable of detecting the initiation of the surface cracks. However, it is not able to calculate the size of the cracks nor the time interval between the initiation and full expansion of the cracks at the surface. In the rest of this section some of the approaches to tackle this issue will be discussed.

To begin with the first solution, the results of laboratory tests which are conducted in 2009 by Kabo [52] should be discussed.
3.1. Rolling Contact Fatigue (RCF)

In different operational conditions including high and low friction levels and various contact pressure using roller-rig, twin-disk, and linear-rig, Kabo found a relation between the $F_{I_{surf}}$ values and the number of cycles to failure $N_f$ of the material. Failure is defined by observing the cracks on the surface of the specimen. This relationship is given by Equation (3.18),

$$F_{I_{surf}} = 1.78(N_f)^{-0.25}.$$  \hspace{1cm} (3.18)

If one re-writes Equation (3.18) for $N_f$ in each wheel revolution $i$, it yields Equation (3.19) which sometimes is called the “damage index” or in short $D_i$ model.

$$\left(\frac{1}{N_f}\right)_i = D_i = \left(\frac{F_{I_{surf,i}}}{10}\right)^4$$  \hspace{1cm} (3.19)

Ekberg [53] suggests that the tangential stresses can be used in Equation (3.19) instead of $F_{I_{surf}}$ to consider the partial-slip condition as in Equation (3.20). Note that for this model the maximum value of stresses in each longitudinal strip is considered for each lateral coordinate of the patch.

$$D_{y_i} = \max \left(\frac{SI}{p(x,y)}\right)^4 10,$$  \hspace{1cm} (3.20)

where, $SI$ is called “stress index” which is given in [54] as

$$SI(x,y) = \sqrt{q_x(x,y)^2 + q_y(x,y)^2} - k.$$  \hspace{1cm} (3.21)

To this end all the proposed models are to predict the running distance from initiation of the cracks to the full expansion of them to the surface. To calculate the depth of the cracks, Ekberg [53] suggests the following assumptions (see Figure 3-21):

- A fixed ratio $r$ exists between the mouth width $\rho$ of a crack and its length $L$: $r = \frac{2\rho}{L}$. 

---

66
• The surface length is proportional to the fatigue damage magnitude $D$ so that $\frac{\rho}{\rho_f} = \frac{D}{D_f}$, $\rho_f$ and $D_f$ stand for surface length of the crack and its corresponding damage index at failure.

• The crack propagates with angle $\theta$.

Figure 3-21, Surface crack with surface length $\rho$, length $L$, angle $\theta$ and depth $h$.

Therefore, the depth $h$ will be

$$h = L \cdot \sin \theta = \frac{2\rho}{L} \sin \theta = \frac{2D\rho_f}{rD_f} \sin \theta.$$  \hspace{1cm} (3.22)

Dirks used this idea and with some adjustment developed a model that can predict the crack depth. Instead of Equation (3.20) she assumes a general power law relationship between the magnitude of the stresses and the number of load reversals to failure as

$$\sigma_a = \alpha (N_f)^\beta.$$ \hspace{1cm} (3.23)

If the stresses can be calculated by dynamic simulations, the unknown $N_f$ will be

$$N_f = \alpha' (\sigma_a)^{\beta'}$$ with

$$\alpha' = 9.6 \cdot 10^5 \text{ & } \beta' = -1.06.$$ \hspace{1cm} (3.24)

where, $\alpha, \beta, \alpha'$ and $\beta'$ are unknown fatigue strength coefficients which should be calibrated against measurement data. For the calibration process, crack measurements ($\rho$ and $L$) in a curve on the Dutch railways over a period of five years are used.

However, in another study [56] she claims that one of the calibrated strength parameters is not valid for wheel applications and adjusted again against wheel related measurement data.
3.2. Wear

In a different approach, the authors of [57] and [58], propose to calculate accumulation of the cyclic plastic strains, $\gamma$, until the critical strain, $\gamma_c$ (fracture limit) is reached. The key factor here, is the relation between the applied stress and the resultant plastic flow. For each layer of the material $i$ at the depth $z$ this relation is given by

$$\Delta \gamma^i = C \left( \frac{\tau_{z\alpha}}{k_{eff}^i} - 1 \right),$$

where, $C = 0.00237$ for BS11 rail steel. The effective shear strength, $k_{eff}$, is applied to consider the hardening effect and it is related to its initial values by

$$\frac{k_{eff}}{k_0} = max\{1, \beta \sqrt{a - \exp(-\alpha \gamma)}\},$$

where, $\alpha$ and $\beta$ are the material hardening coefficient and should not be confused with the ones in Equation (3.23).

3.2 Wear

Published research and studies in wear modelling usually apply three approaches:

- field measurements,
- laboratory tests and
- theoretical prediction models.

Most of the field measurements have addressed the effect of lubrication on wear such as [59] and [60].

Archard [61] took the idea of the adhesive wear definition and found that the volume of material removed by wear per sliding distance ($W$) is proportional to the quotient of the pressure ($p$) and the hardness ($H$) of the softer material. The proportionality factor is called the wear coefficient ($k'$)

$$W = k' \frac{p}{H}.$$  

(3.27)

The wear coefficient depends on the governing wear mechanisms. Archard validated his model by determining the wear rates for different material pairs by pin-on-disk tests.
Lim and Ashby [62] have performed substantial amounts of laboratory tests and introduced wear maps where the wear coefficient is plotted as a function of sliding velocity and the nominal pressure (normal load divided by the nominal contact area). The wear map corresponding to medium carbon steel, based largely on pin-on-disc data, is generally divided into two main regions of mechanical and chemical wear, cf. Figure 3-22. The mechanical wear occurs at low sliding velocities where the wear coefficient is more a function of nominal pressure than the velocity. The chemical mechanism occurs at higher sliding velocities (above 1 m/s). The mechanical part contains three regions of mild and severe wear together with a transition area in between. Childs [63] has suggested a wear map for the mechanical wear mechanism where the wear coefficient is a function of the asperities attack angle and the relative strength of the interfaces. The chemical part of the map, however, contains of two regions: mild and severe oxidational wear. As seen in the figure, mild oxidation could even be protective as at a given level of pressure and high sliding speed the wear coefficient drops to low values. This mild oxide material behaves like a lubricant in between the surfaces.

![Figure 3-22, Lim and Ashby wear map for medium carbon steel based largely on pin-and disc data [62].](image-url)
Olofsson and Telliskivi [64] have investigated the evolution of rail profiles of a commuter line track within two years together with performing several laboratory tests with two different testing machines: a two-roller (disc on disc) and a pin-on-disk machine. The tests result in a simplified wear map with the wear coefficients depending on sliding velocity and contact pressure, cf. Figure 3-23. For lubricated conditions, Jendel [65] suggests a different wear map with coefficients almost six times smaller than the ones shown in the figure. These values are claimed to be obtained from several field measurements. Furthermore, Zue suggests a different wear map for greased conditions where the wear coefficients are about 100 times less [66].

Figure 3-23, Wear map for wheel and rail steel; \( H \) is the hardness of the material.

McEven and Harvey [67] using a full-scale wheel-on-rail rig, proposed a linear relation between the wear rate and the dissipated energy per unit distance rolled \( \bar{E} \), per unit area \( A \), adjusted with a constant off-set term, \( K \). The energy dissipation per unit distance area is the creep forces times the creepages added to the moment times the spin in the contact patch. The proportionality factor \( K \) is expected to be function of wheel and rail material. They also predicted two wear regimes, i.e. mild (tread contact) and severe (flange contact) wear.

\[
\text{Wear rate Per unit rolled distance} = k \frac{\bar{E}}{A} + K \tag{3.28}
\]

\[
\bar{E} = F_x \nu_x + F_y \nu_y + M \phi \tag{3.29}
\]
\( \bar{E} \) sometimes is called the wear number. In another study [68] the authors suggested a simple relationship between the material loss and the energy dissipation:

\[
\begin{aligned}
    &\bar{E} < 100N \rightarrow Material\ loss(mm^2) = 0.25 \frac{\bar{E}}{D} \\
    &100 \leq \bar{E} \leq 200 N \rightarrow Material\ loss(mm^2) = \frac{25}{D} \\
    &\bar{E} \geq 200N \rightarrow Material\ loss(mm^2) = \frac{(1.19\bar{E} - 154)}{D}
\end{aligned}
\]  

(3.30)

where, \( D \) is diameter of the wheel in millimetres.

Krause and Poll [69] reviewed several other test results and investigations regarding the proportionality of the wear rate to the longitudinal creepage, normal force and sliding distance. They concluded that the volume of material loss (\( W_V \)) is proportional to the frictional work \( W_R \)

\[
W_V = I_w \cdot W_R 
\]

\[
W_R = \bar{E} \cdot S 
\]

(3.31)

where, \( S \) is the sliding distance. The proportionality factor is \( I_w \) and it depends on the environmental conditions such as the humidity, the material of contacting surfaces and the temperature of the contact patch. The temperature itself is proportional to the frictional power \( P_R \)

\[
P_R = \bar{E} \cdot V 
\]

(3.32)

where \( V \) is the speed. Krause and Poll also concluded that it is difficult to derive a simple mathematical wear law because:

- different parameters are affecting each other. For example, the frictional work affects the surface temperature which changes the tribology of the surfaces and material behaviour.
- different mechanisms are involved in a wear process.

Several authors have investigated and developed wheel wear prediction tools. Some of them are reviewed in this section. Note that the studies which are mentioned focus on uniform wear.

Kalker [70] has developed a method to predict the wheel profile evolution considering that the material loss is proportional to the
3.2. Wear

frictional work. Proportionality factors were gained from field studies. His method was only applied on the tread contact and no flange wear is considered. As he used Hertzian contact and the simplified theory for his calculations it was not possible to predict severe wear on the wheel profile as the wheel gets a hollow shape and the contact becomes more conforming. No comparison between measurements and calculations are mentioned.

Pearce and Sherratt [71] developed their prediction method using the energy approach where the amount of material loss is proportional to the energy dissipation. The values of the wear coefficients are taken from [68]. To validate the model, they compared the development of the equivalent conicity over the running distance from both simulation and measurement; the conclusion was that only a qualitative judgment is possible.

Ward [72] also followed the energy approach. To find the wear coefficient he carried out twin disk testing. The wheel profile is discretized into longitudinal strips and each contact patch is divided into several cells. The numbers of cells are equal for each strip. The wear depth is calculated separately in each of the cells and at the end integrated along the longitudinal strips as shown in Figure 3-24.

![Figure 3-24, Summation of the wear per strip to give total wear depth; t is the duration of the contact and $C_{pw}$ is the position of the centre of the contact point.](image)

Jendel [73] used Archard’s approach instead. His methodology is based on a load collective concept, which determines a set of dynamic time-domain simulations. These simulations reflect the actual rail network for
vehicles in question including for example, track design geometry, track irregularities, rail profiles and the vehicle operating conditions. Hertzian theory is used for the normal contact and the FASTSIM method is applied for the tangential contact problem. The simulation results are compared with measurements and good agreement was observed. A summary of the methodology is presented in Figure 3-25.

Enblom [74] has used the same methodology as Jendel. However, he also included the elastic strain in the sliding velocity assessment, cf. Equation (2.12). He also extended the simulation set with simulation of braking.

Figure 3-25, Methodology of wheel wear prediction developed by Jendel.

An attempt is made to study the relationship between the energy approach and the Archard’s method. A wear ratio has been proposed that maintains a linear relationship between the calibrated wear depth (see Figure 3-25) as calculated by Archard (time consuming) and the calculated energy dissipation (simulation output), see [75].
3.3. Influence of Wear on RCF

Figure 3-26, Wear ratio as a function of lateral wheel position calculated for (a) the case of loaded wagons and (b) the case of empty wagons [75].

Using the obtained wear ratio, it is much faster to calculate the wear depth for long distances rather than using Jendel’s approach with Archard’s theory. However, this comes at the expense of losing accuracy. The volume of the removed material has an error between 10% to 60%. However, the error is mostly at the flange side where lubrication is involved.

3.3 Influence of Wear on RCF

Many attempts have been made to combine wear and RCF effects in the prediction of wheel and rail damage. At a certain level of wear depth, the initiated cracks could be polished away from the surface.

There are generally two approaches to consider the effect of wear on RCF. The first one is to calculate the wear depth and RCF crack length separately and consider the geometrical effect of one to another, see [56]. The method is a combined Jendel methodology for wear and the one which is described in Section (3.1.5) for RCF. The methodology contains a phase in which the model parameters should be tuned against field measurements. The following paragraph is a brief review of the geometrical approach.

After calculating the crack length without considering the wear effect (cf. Equation (3.20)-(3.24)), first, the task is to find a relation between the crack depth, $L$, and its surface width, $\rho$, from measurements, see Figure 3-21 and Equation (3.22). This proportionality can be extended to a general relationship between the wear depth and the RCF crack size. The wear is calculated by Jendel’s methodology (see Figure 3-25). Thus, by subtracting the calculated surface crack length without considering
the effect of wear and the worn off surface crack length the actual value can be obtained.

The other way to consider the wear effect on RCF is to include it into the fatigue damage evaluation. Burstow’s fatigue life model [76] may be the best example of such an approach. In this approach the fatigue life is calculated as a function of wear number cf. Equation (3.29). A schematic Burstow diagram is presented in Figure 3-27.

![Figure 3-27, Rolling contact fatigue function](image)

It is suggested that the turning points of the figure can be calculated by the material properties of the rail/wheel as Equation (3.33).

\[
\bar{E}_i = \frac{\nu_i A}{2 \sqrt{3}} (\sigma_y + \sigma_U), \text{for } i = 1, 2, 3
\]  \hspace{1cm} (3.33)

where, \(\sigma_y\) and \(\sigma_U\) are material yield limit and ultimate tensile strength respectively, and the corresponding suggested creepages \(\nu_i\) are 0.1%, 0.3% and 1% (see also [50]). The suggested values need to be tuned by field measurements for each application otherwise they remain as an approximation.

To visualise RCF on the wheel profile as it is shown in Section (3.1.4), one can use the concept of Burstow’s diagram in both full-slip and partial-slip condition. The local energy dissipation is given by

\[
E(x, y) = q_x(x, y) \cdot (v_x - \phi \cdot y) + q_y(x, y) \cdot (v_y + \phi \cdot x).
\]  \hspace{1cm} (3.34)

Using the dynamic simulation results, counting the number of cycles with RCF risk (“RCF number”) according the shakedown theorem
3.3. Influence of Wear on RCF

(positive SI values cf. Equation (3.21)) gives a good indication of RCF development.

To apply this theory, the stresses, which could be calculated via both the FaStrip and FASTSIM algorithms are distributed over the meshed contact patch. In the elements where the Tresca yield criterion is exceeded, there will be an indication of RCF risk. The number of incidents is counted and scaled based on the contribution of the length of the simulated curves to the total length on the line (according to Jendel’s methodology, see Figure 3-25). Furthermore, to consider the effect of wear, the Burstow fatigue index diagram is used. If the calculated energy values are less than $E_1$, it is assumed that RCF is not affected by wear, and therefore, the corresponding RCF number will be unchanged. However, for energy values between $E_1$ and $E_2$, wear starts to dominate, and thus the RCF number is decreased proportionally, meaning that more wheel cycles are needed to initiate the cracks. Finally, if the calculated energy dissipation is higher than $E_2$, it is assumed that wear is completely governing the situation and there are no cracks to propagate. As an example, Figure 3-28 shows the calculated energy dissipation for all contact points of the inner and outer wheels normalised by $E_1$ midway in a tight curve. In the figure, the $E_1$ threshold level is shown by a transparent surface. In this case, regardless of the propagation assumption, the RCF number would be affected both on the flange and the tread area, since in both areas energy dissipation values exceed the first threshold level. The $E_2$ value is almost three times as high as $E_1$. Therefore, it is not shown for clarity of the figure.

![Figure 3-28](image.png)

Figure 3-28, An example of normalised energy dissipation by the threshold $E_1$ value on the wheel profile with its corresponding corrected RCF number scaled by the length of the curve section.
Surface initiated cracks occur in the entire contact patch. However, as wear does not occur in the stick area the RCF correction factor is only activated in the slip area. Nevertheless, to consider the effect of neighbouring points in uniform calculation, it is assumed that if the total energy dissipation in the contact patch is greater than the global $\overline{E_2}$, all RCF number values in both the slip and stick area will be discarded and equated to zero. Although the shakedown theorem shows a high percentage of RCF damages in the flange area, there are less indications of cracks observed in the field studies. This is due to the crack propagation nature. Since the expansion of the cracks is usually due to the water (or any viscous lubricant) entrapment and hydrostatic pressure, propagation of the cracks rarely takes place in the negative direction of the forces on the wheels, as it happens in the flanges [77].

Figure 3-29 shows the simulated RCF results after 50,000 km simulated running distance. The RCF numbers are corrected by Burstow’s concept. Both the FASTSIM and FaStrip algorithms are used to calculate the surface shear stresses and the results are shown separately.

Worthy to note is that Dirks [55] also used Equation (3.34) as the “Energy Index” or in short EI model instead of the SI model i.e. Equation (3.21) in her crack length prediction model.
4 Summary of the papers

4.1 Paper A

Wheel damage on the Swedish IORE line investigated via multibody simulation

The paper presents some details and a general summary of the three-piece bogie simulation model. Track geometries and irregularities are discussed and some statistics from the IORE line are presented. Moreover, after discussing the rail profiles and wheel-rail friction variations, the theory behind the surface initiated RCF is described via the concept of the shakedown diagram. During a series of parameter studies, it is concluded that RCF mostly occurs in curve sections with radii below 450 m. Improvement of the lubrication policies is suggested especially on the Norwegian part of the line. Moreover, it is shown that track degradation plays a key role in initiating the cracks. Simulation results show that the track stiffness is not the main concern with regard to RCF cracks. The study also confirms the significant impact of the wheel-rail friction coefficient on RCF. Good agreement between the predicted location of cracks and field observations is achieved.

4.2 Paper B

Modelling and simulation of freight wagon with special attention to the prediction of track damage

To develop and validate simulation models of freight wagons is much more difficult than it might seem at first sight. The main reasons are the strong non-smoothness in almost all suspension components and the variation of suspension characteristics that can be found from component measurements. In the paper, several suggestions are given on how to model the typical suspension components in freight wagons. The specific difficulties to validate the models are discussed. The principles on how to use the models to calculate wear, RCF and the
overall track damage are explained. Examples are given on how the models were used for prediction.

4.3 Paper C

*Prediction of RCF and wear evolution of IORE locomotive wheels*

The summary of the project background is discussed in the introduction of the paper. Then, a literature review of previous works on wear and RCF is presented. Moreover, the employed wear and RCF calculation methodologies are described. As locomotives are equipped with a flange lubrication system, both lubricated and non-lubricated conditions are considered in the simulations. To include the traction effects on the wheel damages, total resistance forces for each of the simulation cases are calculated. Both the impacts of lubrication and tractive forces on RCF and wear are described via examples. To consider the effect of wear on RCF, a method is presented. This method is based on the previous works on modifying the calculated RCF index by the corresponding wear number. Good agreement is achieved when comparing the simulated and measured worn profiles. Moreover, comparing the calculated and observed RCF locations, it is concluded that the vehicle lubrication system is, at least partly, not performing as it is expected. The procedure is repeated for two new proposed wheel profiles to choose one for the field tests. These wheel profiles are designed to reduce the number of wheel damages. Finally, one of the profiles is suggested to be better than the other one because of producing milder wear and RCF in the long term.

4.4 Paper D

*Study of the long-term evolution of low-RCF wheel profiles for LKAB IORE wagons*

With a similar approach to Paper B, the long-term stability of wheel profiles of the IORE wagons running in northern Sweden and Norway is analysed. However, the effect of shoe brakes is estimated first by calculating the amount of braking force required for each wheel and then by comparing the simulated worn profile with the measured one. The methodology is first validated by comparing the simulated and measured worn profiles for the actual wheel profiles. Simulated wear and RCF evolution of the current wheel profiles shows good agreement with measurement and observation data; minor differences are mainly due to
the wear at the corner of the field side due to switches and crossings and due to the lack of measured extremely worn rail profiles in the simulations.

4.5 Paper E

Fast wear calculation for wheel profile optimization

The paper is an attempt to establish the relation between the dissipated energy in the contact patch and the wear depth calculated according to Jendel’s wear calculation method. As the energy dissipation is an output of simulation software, it is possible to reduce the wear calculation time significantly by using the proposed energy-wear depth relationship. To achieve the first goal, a so-called wear ratio is introduced. The wear ratio is the quotient between the calculated wear depths using Jendel’s method and the corresponding wear number along the wheel profile. This ratio is used to estimate the worn wheel profile by multiplying this ratio to the calculated wear number. With this technique, the wear simulation time is halved. To further decrease the simulation time, the calculated wear depth obtained by the new method is extrapolated several times to achieve the worn wheel profile corresponding to the desired distance before a new simulation is started. The accuracy of the method is checked by comparing the extrapolated wheel profiles and the reference wheel profile. The methodology seems to be more correct in shorter simulation distances. However, for a wheel profile optimization process when one wants to have control on the shape of the profile after some ten thousand of kilometres the method can give a reasonable result with much less computational effort like for the original method proposed by Jendel.

4.6 Paper F

Wheel life prediction model - An alternative to FASTSIM algorithm for RCF

Like paper C, the IORE locomotive simulation model is used to study wear and RCF on wheels. However, a wheel life prediction model is also proposed to estimate the running distance intervals between two consecutive wheel reprofilings due to surface fatigue damage. Also, instead of only considering the full-slip condition, wheel-rail contacts in partial-slip condition are studied too. As the FASTSIM algorithm could be affected by large errors in stress levels, to calculate the tangential
stresses a newly developed algorithm (FaStrip) is adopted. The results are compared with five years of field studies. The calculations are repeated for other cases than the normal operational case to study the effects of higher axle load, Ed-braking force and the controlled track gauge on wheel damages.

From the analysis of the results, the following conclusions are obtained:

- The wear calculation results show a very good agreement with the measurement data. However, more studies in the wear map data for the heavy axle load is suggested.
- Using full slip condition for RCF calculations has more deviations from the field observations around the wheel-rail nominal radius towards the flange area, compared to RCF analysis using local slip conditions.
- FASTSIM shows more deviations from the field observations in the nominal wheel-rail radius vicinity than FaStrip. The differences come from the accumulation of the FASTSIM error in the long-term process, which is a decisive factor in prediction models, where the length and the depth of the cracks are the purpose of the model.
- When analysing different operational conditions, no major differences are observed regarding the wear of the profiles.
- The case of controlled track gauge has significantly less RCF area compared to the other cases.
- It has been shown that a higher axle load slightly increases the risk of RCF.
- Regarding the prediction of the reprofiling intervals, the comparison between field observations and simulations is very promising.

As an overall conclusion, damage prediction tools such as the described ones have enough sensitivity to allow comparisons within the range of operational conditions of a specific railway system, which is very useful from a vehicle operator point of view, as they allow a much broader analysis than tests and measurements.
5 Future research directions

5.1 Non-elliptical contact

Estimating the wheel life due to RCF and wear greatly depends on the methodologies, assumptions, and their accuracies. In calculations with the goal of long-term evaluations the time of calculations plays a key role. Estimating the surface shear stresses is one of the main concerns in this field. As it is seen in Section 2 most of the methodologies are based on Hertzian theory for the normal contact problem and the FASTSIM algorithm for the tangential contact problem. However, both have shown to be not accurate enough especially in estimation of the shear stresses. Nevertheless, using algorithms with non-elliptical contacts is usually time consuming which makes it almost impossible to be used in long-term evaluations of the wheel damages. In this work (cf. Paper F) an alternative FaStrip algorithm is used instead of FASTSIM which significantly enhanced the accuracy of estimation of the shear stresses. However, still the normal contact problem is solved by Hertzian theory. A recent study at KTH suggests using a fast and accurate non-elliptical approach to tackle the normal contact problem called ANALYN [78]. However, the method has not been used in the large-scale calculations. See Figure 5-1 for comparison of accuracy of different methods as well as the time of calculations.

![Figure 5-1, Shear stress distribution in the wheel–rail contact patch for three different methods [79].](image-url)

\[ \text{Hertz+FASTSIM} \quad \text{ANALYN+FaStrip} \quad \text{CONTACT} \]

\[ \sim 0.02 \text{ seconds} \quad \sim 0.12 \text{ seconds} \quad \sim 20 \text{ seconds} \]
To this end almost everything which has been mentioned is about post processing of the MBS simulation results. As described in Section 2, the simulation software relies on FASTSIM algorithm and Hertzian theory, due to their acceptable accuracies in estimation of the track forces. However, even in estimation of the forces in large spin cases the errors could be unacceptable (more than 20%). Therefore, it is obvious that there is a need to improve the simulation software contact algorithms as well as the post processing tools and methods.

5.2 Slip control and wheel damage

In some locomotives where the adhesion condition is poor, usually due to rail surface roughness, over-slip is applied on the wheels during a brief period to clean the rail surface and enhance the adhesion condition. There is a need to study the effects of over-slip on wear of the wheels. With applied artificial wear on wheels it is possible to wear off the initiated cracks on the wheel surface before they propagate further. Figure 5-2 shows the simulation results on a straight line for dependency of the energy dissipation on the over-slip (dv) in distinct friction levels. Here, the over-slip as introduced by applying an extra torque to the driven loco wheels.

![Figure 5-2](image)

Figure 5-2, Simulation results for the effects of over-slip (dv) on dissipated energy in various friction levels.

Further studies should be directed towards using the slip-control system effectively to minimise the wheel damages in the long run.
5.3 **Active control of wheel damage**

Nowadays, some of the modern locomotives are using active suspension technologies to improve the passenger ride comfort. Recently an attempt has been made at KTH to investigate the effects of active wheel steering on the wear number [80]. The simulation results clearly show that the dissipated energy is significantly reduced by changing the steering conditions, see [60]. In the study, the control goal (i.e. reference signal) is to modify the longitudinal primary suspension characteristics (using actuators) to change the yaw angle of the wheelset according to [81].

![Figure 5-3](image)

**Figure 5-3.** Wear number as function of time for the combination of active/passive vehicle on track

The vehicle model used in this study is a simple two-axle vehicle with one stage suspension system and the simulated track sections are limited to few numbers. Firstly, the vehicle model should be extended to a full complete loco model with traction and braking capacity and the long-term evolution of the wheel profiles should be examined to study the effect of active steering on operational life of the wheels. Secondly, further studies should be conducted to investigate the best possible control goals (other than the wheelset steering) to minimise the wheel damages (wear and RCF).
6 References


[45] A. R. S. Ponter, A general shakedown theorem for elastic-plastic bodies with work hardening, 3rd Int. Conf. on Structural Mechanics in Reactor Tech., London (1976)


