

THESIS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY IN SOLID AND
STRUCTURAL MECHANICS

Squat defects and rolling contact fatigue clusters

Numerical investigations of rail and wheel deterioration mechanisms

ROBIN ANDERSSON

Department of Industrial and Materials Science
CHALMERS UNIVERSITY OF TECHNOLOGY

Gothenburg, Sweden 2018

Squat defects and rolling contact fatigue clusters
Numerical investigations of rail and wheel deterioration mechanisms
ROBIN ANDERSSON
ISBN 978-91-7597-710-2

© ROBIN ANDERSSON, 2018

Doktorsavhandlingar vid Chalmers tekniska högskola
Ny serie nr. 4391
ISSN 0346-718X
Department of Industrial and Materials Science
Chalmers University of Technology
SE-412 96 Gothenburg
Sweden
Telephone: +46 (0)31-772 1000

Chalmers Reproservice
Gothenburg, Sweden 2018

Squat defects and rolling contact fatigue clusters
Numerical investigations of rail and wheel deterioration mechanisms
ROBIN ANDERSSON
Department of Industrial and Materials Science
Chalmers University of Technology

ABSTRACT

Squat defects, a type of localised rolling contact fatigue damage appearing on rail surfaces with rail break as an ultimate consequence, have concerned infrastructure managers for the last couple of decades. In recent years similar types of defects—so-called studs—that are visually resembling squats, have started to appear. In contrast to conventional rolling contact fatigue of rails, these defects are associated with a thin surface layer of brittle material—a “white etching layer”. The wheel counterpart of squats/studs are called “rolling contact fatigue clusters”. Despite significant research efforts, the exact initiation mechanisms of the defects are still unknown and it is difficult to relate the occurrence of squats/studs and rolling contact fatigue clusters to specific operational scenarios.

The current work aims to deepen the understanding of squat/stud and rolling contact fatigue cluster initiation by comparing and ranking predicted damage from various potential causes of initiation under different operational scenarios. Special emphasis is put on local surface irregularities. These are studied using dynamic vehicle–track interaction simulations to evaluate the impact of e.g. irregularity size, vehicle velocity, wheel–rail friction conditions and position relative to a sleeper. It is seen that surface irregularities might cause substantial fatigue impact. Rolling contact fatigue initiation connected to operational scenarios of specific interest are studied more in detail by mapping dynamic contact stresses from simulations of vehicle–track interaction to finite element models for subsequent stress analyses and ranking of operational scenarios via ratchetting response and low cycle fatigue impact. Among the results, it is seen that larger irregularities and higher wheel–rail friction promote higher fatigue impact. In order to study the influence of irregularity geometry when macroscopic cracks are present, dynamic contact stresses are mapped onto finite element models of a cracked rail head. The severity is assessed using an equivalent stress intensity factor, which is seen to increase with the size of the irregularity. This effect holds also for clusters of irregularities. It is furthermore seen that even a shallow irregularity can make a substantial impact.

The influence of white etching layers is investigated by simulating thermally induced phase transformations occurring in spots on rail and wheel surfaces, subjected to subsequent mechanical loading. The influence of axle load and wheel–rail friction is investigated with respect to fatigue impact. It is seen that the axle load has a rather low influence whereas an increased frictional loading increases the fatigue impact considerably.

Keywords: Squats, studs, rolling contact fatigue clusters, rolling contact fatigue, rail surface irregularities, dynamic vehicle–track interaction, fracture mechanics

PREFACE

The following work was carried out between April 2013 and May 2018 at the Department of Applied Mechanics and the Department of Industrial and Materials Science at Chalmers University of Technology. It was conducted as a part of the project MU31 – “Squats in rails and RCF clusters on wheels” within the Centre of Excellence CHARMEC (CHalmers Railway MECHANics, www.charmec.chalmers.se). The project was partly financed within the European Horizon 2020 Joint Technology Initiative Shift2Rail through contract no. 730841. Some of the computations were performed on resources at Chalmers Centre for Computational Science and Engineering (C3SE) provided by the Swedish National Infrastructure for Computing (SNIC).

ACKNOWLEDGEMENTS

First of all, I would like to thank my main supervisor Elena Kabo and co-supervisors Anders Ekberg, Fredrik Larsson and Peter Torstensson for all their support, encouragement and commitment during these years. It has been very instructive and also fun to work together.

I would furthermore like thank my co-author Johan Ahlström. I also feel gratitude towards members of the project’s reference group as well as participants in the Swedish–Austrian CHARMEC workshops, who have contributed to fruitful discussions. Moreover, I would like to thank friends and colleagues at the divisions of Material and Computational Mechanics and Dynamics for a nice working environment.

Last, but definitely not least, I would like to thank my family for their support and patience throughout the years.

Gothenburg, May 2018
Robin Andersson

THESIS

This thesis consists of an extended summary and the following appended papers:

- Paper A** R. Andersson, P. T. Torstensson, E. Kabo, and F. Larsson. The influence of rail surface irregularities on contact forces and local stresses. *Vehicle System Dynamics* **53.1** (2015), 68–87. DOI: 10.1080/00423114.2014.982890
- Paper B** R. Andersson, P. T. Torstensson, E. Kabo, and F. Larsson. An efficient approach to the analysis of rail surface irregularities accounting for dynamic train–track interaction and inelastic deformations. *Vehicle System Dynamics* **53.11** (2015), 1667–1685. DOI: 10.1080/00423114.2015.1081701
- Paper C** R. Andersson, P. T. Torstensson, E. Kabo, F. Larsson, and A. Ekberg. Integrated analysis of dynamic vehicle–track interaction and plasticity induced damage in the presence of squat defects. *Wear* **366-367** (2016), 139–145. DOI: 10.1016/j.wear.2016.05.014
- Paper D** R. Andersson, F. Larsson, and E. Kabo. Evaluation of stress intensity factors under multiaxial and compressive conditions using low order displacement or stress field fitting. *Engineering Fracture Mechanics* **189** (2018), 204–220. DOI: 10.1016/j.engfracmech.2017.11.015
- Paper E** R. Andersson, J. Ahlström, E. Kabo, F. Larsson, and A. Ekberg. Numerical investigation of crack initiation in rails and wheels in the presence of white etching layers. *Submitted for international publication* (2018)
- Paper F** R. Andersson. Assessment of rolling contact fatigue cracks in the vicinity of rail surface irregularities. *Submitted for international publication* (2018)

Papers A–E were prepared in collaboration with the co-authors. The author of this thesis was responsible for the major progress of the work, i.e. took part in planning the papers, developed the numerical models, carried out the simulations, interpreted the results and wrote the main part of the papers. The dynamic vehicle–track interaction software used for parts of the analyses in Papers A–C and F was developed by Peter Torstensson.

CONTENTS

Abstract	i
Preface	iii
Acknowledgements	iii
Thesis	v
Contents	vii
I Extended Summary	1
1 Introduction	1
1.1 Background	1
1.2 Aim and scope of research	3
1.3 Outline	4
2 Fundamentals of squat/RCF cluster formation	4
3 Overview of potential damage mechanisms	7
3.1 Rolling contact fatigue damage	7
3.2 Dynamic loads	9
3.3 Thermal damage	9
4 Modelling of wheel–rail contact	11
4.1 Basic concepts of rolling contact	11
4.2 Analytical solutions	13
4.2.1 Normal contact – Hertz solution	13
4.2.2 Tangential contact – Carter solution	14
4.3 Kalker’s complete and simplified theories	16
4.4 Finite element method	17
5 Simulation of dynamic vehicle–track interaction	18
5.1 Method of moving Green’s functions	18
5.2 Transient finite element analysis	22
5.3 Multibody simulation	23
6 Analysis of rail and wheel damage	23
6.1 Stress–strain analysis	23
6.2 Numerical approaches to RCF and fracture mechanics	26
6.2.1 Maximum contact stresses and forces	26

6.2.2	Surface fatigue index	27
6.2.3	$T\gamma$ approach	27
6.2.4	Equivalent stress/strain	28
6.2.5	Jiang–Sehitoglu fatigue parameter	29
6.2.6	Fracture mechanics	30
6.3	Thermal loading and damage	34
7	Summary of appended papers	35
7.1	Paper A: The influence of rail surface irregularities on contact forces and local stresses	35
7.2	Paper B: An efficient approach to the analysis of rail surface irregularities accounting for dynamic train–track interaction and inelastic deformations	38
7.3	Paper C: Integrated analysis of dynamic vehicle–track interaction and plasticity induced damage in the presence of squat defects	38
7.4	Paper D: Evaluation of stress intensity factors under multiaxial and compressive conditions using low order displacement or stress field fitting	39
7.5	Paper E: Numerical investigation of crack initiation in rails and wheels in the presence of white etching layers	40
7.6	Paper F: Assessment of rolling contact fatigue cracks in the vicinity of rail surface irregularities	42
8	Conclusions, main results and future work	43
	References	46

Part I

Extended Summary

1 Introduction

1.1 Background

Since the 19th century, railway traffic has been an important mode of transportation for passengers and freight. Its ability to transport large volumes efficiently is of increasing importance, not least from an environmental perspective. One reason for the efficiency is that rolling resistance decreases with an increase of the stiffness of the contacting bodies [49]. Wheels and rails are made of steel, which causes the rolling resistance to be relatively low. The tyres of road vehicles are on the other hand produced of rubber, which is considerably more flexible than steel and thus provides a higher rolling resistance in contact with asphalt. Johnson [49] has estimated the rolling resistance of a road tyre to be approximately 100 times higher than the rolling resistance of a railway wheel.

According to the Hertzian theory of contact [41], the maximum contact pressure between wheel and rail increases with the material's stiffness [49]. This means that the benefit of using a stiff wheel and rail material, such as steel, comes with the drawback of high contact pressures. Consider, for example, a wheelset carrying an axle load of 22.5 tonnes. Half the load, i.e. 11.25 tonnes, will be transferred to the rail through a contact patch of a size similar to the area of a thumb nail, cf. methods in [48] for estimations. Additional loads in the form of friction forces are also present in wheel–rail contacts, especially when the wheel is braked, accelerated or negotiating curves. The resulting stress magnitudes are of a scale rarely seen within engineering.

The high stresses in the wheel–rail contact may cause different types of deterioration, such as wear and plastic deformation. Another frequently occurring deterioration phenomenon is the initiation and growth of cracks under rolling contact loading. This phenomenon, which occurs on both wheels and rails, is referred to as Rolling Contact Fatigue (RCF) and is usually caused by ratchetting and/or low cycle fatigue [33]. This thesis treats two types of defects that are associated to RCF: *squats*, that appear on rails, and *RCF clusters*, appearing on wheel treads.

According to the UIC Code of rail defects [77], squats can be defined as “[...] a widening and a localised depression of the rail/wheel contact band, accompanied by a dark spot containing cracks with a circular arc or V shape”, cf. Figure 1.1. If no actions are taken and the squat cracks are free to grow, they might cause a rail break [40]. This is of course something that infrastructure managers want to avoid and is also the reason for the efforts that have been made to understand the root causes of squat formation.



Figure 1.1: A typical squat-like defect found in Stockholm, Sweden. Two distinct lobes are clearly visible. Also note the V-shaped surface breaking crack and the slight widening of the running band.

RCF clusters on wheels can be defined as a “[...] localised damage effect which can occur in the centre of the wheel tread” [29], cf. Figure 1.2. These are prone to crack propagation, with the potential of detaching wheel tread material. This leads to higher loads, which in turn promote further damage on the wheel (and also on the running gear and track components) and a higher rate of deterioration. This is something that vehicle owners want to avoid, which motivates a research interest in understanding root causes and initiation mechanisms.

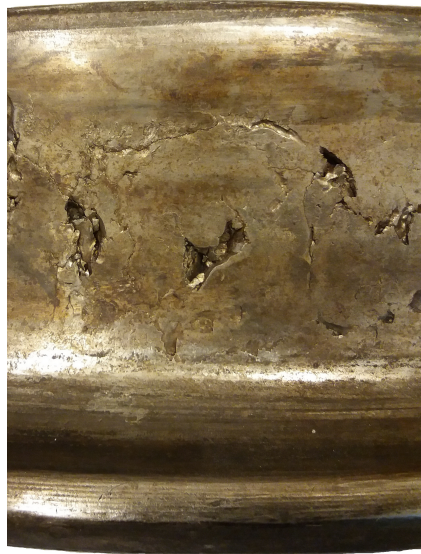


Figure 1.2: *Typical Rolling Contact Fatigue (RCF) cluster on a wheel tread. Adopted from **Paper E**.*

Conventional RCF is, as mentioned above, considered to be responsible for initiating squats in rails and RCF clusters on wheels. In addition, another potential initiator has been linked to rail and wheel defects, namely brittle so-called White Etching Layers (WELs). At least some of these are thought to be formed due to thermally induced material phase transformation. Rail defects believed to be triggered by the occurrence of WELs are known as “studs” [40]. These defects visually resemble squats, cf. Figure 1.1. In order to find effective mitigating actions to squats, studs and RCF clusters it is important to understand similarities and differences between thermally induced defects and the conventional RCF counterparts, as will be further discussed.

1.2 Aim and scope of research

In order to develop rational countermeasures to prevent the development of squats and RCF clusters, their root causes must be known and understood. Even though some main features have been identified, it is still difficult to explain why initiation occurs in some situations but not in other. The overall aim of this thesis is to forward the understanding of squats and RCF clusters through numerical simulations. As discussed in Chapter 3, such an approach should preferably be able to account for several physical phenomena, for example:

- high-frequency wheel–rail interaction, especially when a rail surface irregularity is impacted by a wheel
- nonlinear material response due to plastic deformation in the vicinity of a squat/RCF cluster
- high contact stresses
- material phase transformations due to thermal loading

To this end, an in-house program is used to simulate dynamic vehicle–track interaction, while finite element simulations are used to evaluate the resulting material response. All the aforementioned phenomena are considered in **Papers A–F**.

There are, of course, some restrictions on the subjects covered in this thesis. Firstly, the focus is on numerical simulations. Experiments and field investigations employed to contrast simulation results are taken from the literature. Furthermore, wear is not explicitly modelled, even though its influence might, in some cases, be significant. Finally, since the initiation phase is of most practical interest, the main focus of the studies is on the relatively early stages of squat/RCF cluster development.

1.3 Outline

The remainder of this extended summary is outlined as follows. Chapter 2 contains a general survey of squats and RCF clusters. A more detailed literature review regarding specific initiation causes is presented in Chapter 3. Chapters 4–6 are devoted to strategies for studying squats/studs and RCF clusters using computer simulations. To facilitate the discussion, the topic is divided into three chapters: modelling of wheel–rail contact, simulation of vehicle–track interaction and analysis of rail and wheel damage. Chapter 7 contains summaries of the appended papers. Conclusions and suggestions for future work are found in Chapter 8.

2 Fundamentals of squat/RCF cluster formation

The history of squats dates back to the 1950s, when black spots with embedded cracks started to appear on the rail surface of Japanese railway lines [42]. The problem was given further attention after the Shinkansen high speed line was opened in 1964, as the black spots began causing “Shinkansen shelling”, which in some cases lead to rail breaks [61]. Some hypotheses regarding the root cause involved surface initiated RCF, effects related to wheel torque (all wheels were driven), wheel flats and entrapment of e.g. stones between the wheel and the rail. It was furthermore observed that the Shinkansen shelling did not appear in dry tunnels [61]. Eventually, it was possible to attenuate the problem on the Shinkansen line by rail grinding and rail profile modification [61].

At almost the same time, similar defects started to appear in Great Britain, where they were named “squats”. This subject caught the interest of researchers from the mid 1970s. It could be concluded that surface initiated RCF is the physical mechanism at work when squats are formed [38]. These type of cracks tend to grow a few millimetres at a shallow angle to the rail surface until a critical depth is reached. Then the crack typically branches, either upwards causing spalling or downwards possibly resulting in a rail break [19]. Grassie et al. [38, 40] have mentioned several key findings from the early squat research. Some of these observations include:

- Squats appear where the driving traction is high.
- Squats are not found in tunnels.
- Squats tend to appear on straight lines and in mild curves.
- Squats tend not to appear on heavy haul lines.

The two first items are generally expected in the case of surface initiated RCF. High traction will result in large amounts of plastic deformation that will promote RCF crack initiation. Lubrication, here in the form of rain water, will then facilitate RCF crack growth once cracks are formed [16]. The two latter items are reflecting that excessive surface tractions will result in significant wear that will grind away initiated RCF cracks.

Some confusion occurred when it was eventually realised that a type of defect, visually resembling a squat, did not share the typical signs of surface initiated RCF. These defects lacked the amount of plastic deformation typically related to surface initiated RCF and seemed to originate from a thin and brittle material layer known as a White Etching Layer (WEL) [40]. The name originates from the light appearance of these layers after being etched with a mixture of nitric acid and ethanol, see e.g. [92]. Even though different hypotheses have been presented, at least some of the WELs seem to consist of martensite, originating from thermal loading [36]. Other than the presence of WELs, it has been observed that this “new” type of defect develops significantly faster than the classic squat. According to Grassie et al. [40] some 10 MGT of traffic is needed for initiation. This is approximately one tenth compared to conventional squat initiation. On the other hand, it was also noted that these new cracks did not seem to be as prone to cause rail breaks as squats. Grassie et al. [40] have proposed that these new defects should be called “studs” to distinguish them from the conventional squats.

To summarise, two different, but visually similar, rail surface defects exist: squats and studs. Squats stem from surface initiated RCF cracks which might cause transverse rail failure or shelling. Stud cracks, on the other hand, initiate in the vicinity of surface layers of martensite and can cause shelling, but do not seem to cause rail breaks.

RCF clusters on wheels have historically been given less attention than squats and studs. One possible reason is that squats have been considered more serious since they might cause rail breaks and subsequent derailments. Nonetheless, RCF clusters, followed

by shelling of the wheel tread, might lead to high dynamic impact loads which promote further damage to the wheel, but also to the running gear, track components etc. One key characteristic of RCF clusters is that they initiate locally on the wheel tread [29] with extension around the circumference as the cluster expands. This is in contrast to other types of wheel RCF that often initiate around the entire wheel tread.

There are similarities between squats/studs appearing on rail surfaces and RCF clusters appearing on wheel treads. For example, they both involve locally initiated surface defects and subsequent crack growth. Where they differ is mainly in how crack propagation and (final) fracture proceeds. Knowledge gained from squat/stud research can therefore be applied also on RCF clusters to a large extent. This chapter, and most of the work in the appended papers, mainly treats squats and studs, but as discussed, initiation mechanisms of squats/studs should largely be valid also for RCF clusters. It should however be kept in mind that there are differences between the stress states in rails and wheels, e.g.:

- When the rail is loaded it will deflect and thus introduce bending stresses which are not present in wheels.
- Residual stresses after manufacturing differ between rails and wheels [17, 63].
- The wheels are regularly subjected to thermal loads stemming from tread braking. As a consequence, tensile residual stresses can occur in the surface [33]. Rails are on the other hand susceptible to the surrounding temperature as a cold climate might introduce tensile stresses due to constrained thermal deformation.
- For a given position along a rail, the running band is essentially the same over time. This is not the case for a wheel tread that will have different contact positions depending on e.g. curve and switch negotiations.

These effects will all affect crack initiation and propagation. The remaining part of this thesis will focus on squats/studs, but highlight similarities and differences to RCF clusters when needed.

In order to deepen the knowledge about the formation of squats/studs and RCF clusters, and to form a basis for effective mitigation strategies, the overall initiation mechanisms presented in this chapter must be related to more specific damage related studies. This is the subject of the next chapter.

3 Overview of potential damage mechanisms

A general overview on squats and studs can be found in the works by Grassie et al. [38–40]. Further, general experiences from squats in Australia and the Netherlands have been described in Daniel et al. [27] and Zoeteman et al. [104] respectively.

This chapter provides a literature survey regarding the mechanisms behind squat/stud formation. To facilitate the overview, the discussion is treated under three different subsections: RCF damage, dynamic loads and thermal damage. The headlines are chosen to reflect the topics of the appended papers rather than making a stringent categorisation of possible crack initiation and growth mechanisms. There are no rigid borders between the subjects. Thermal damage can for example introduce a surface irregularity that causes higher dynamic loads, which in turn could initiate additional RCF damage.

3.1 Rolling contact fatigue damage

Rolling contact fatigue differs in several ways from fatigue stemming from other types of loading. As an example, Ekberg and Kabo [33] mention that the loading conditions are multiaxial with out of phase stress components. RCF cracks are furthermore subjected to compressive loading during most parts of the load cycle, whereby mode II and III loading are typically the dominating modes of deformation¹. It has however been difficult to obtain stable mode II crack growth in laboratory experiments as the cracks tend to arrest or branch into mode I cracks [15]. Researchers have therefore put a lot of effort into understanding how RCF cracks can propagate due to the significant shear loads despite crack closure. Bower [16] used simulations to investigate three different possibilities, all related to the presence of a fluid:

- Lubrication of the crack faces, which lowers the crack face friction and thus facilitates mode II propagation.
- The fluid fills the crack and exerts a pressure on the crack faces, enabling both mode I and II loading.
- The fluid is entering the crack and gets entrapped and forced towards the crack tip due to crack mouth closure. This results in both mode I and II loading.

Bold et al. [15] used biaxial testing to produce a cycle of mode I loading followed by mode II loading, resembling the conditions of the aforementioned entrapment model, and obtained crack growth on the plane of the maximum shear stress range. Wong et al. [95] examined under which conditions a RCF crack deviates from coplanar growth and branches. They found that both the “mode I effective stress intensity range” (see [95] for definition) as well as the interaction of the mode I and II load cycles play important roles. Kaneta et al. [59] later used twin-disc experiments and presented a photomicrograph showing a crack network where water has reached the tips of approximately 1 mm long

¹See Section 6.2.6 for a definition of stress intensity factors and modes of deformation.

cracks. They furthermore managed to initiate defects similar to squats in the same experiment. To the best knowledge of the author, this is the only time experimentally induced squat initiation has been reported in the literature.

The aforementioned studies have provided significant insight in the growth of RCF cracks in general, including squats. There are also studies focusing more on squats in particular. One of these were performed by Clayton and Allery [24]. They distinguished between two different initiation categories for a squat-like defect: initiation from small surface indentations and initiation from fine cracks appearing at the gauge corner side of the rail surface. In a later research review, Clayton [23] described squats as surface initiated RCF defects and mentioned indentations caused by WELs as possible initiation triggers. Kondo et al. [61] described the Japanese experience of Shinkansen shelling and suggested four different initiation mechanisms (which all relate to RCF in one way or another): 1) Surface initiated RCF 2) Effect of torque since all wheels are driven 3) Dynamic loads from wheel flats 4) Other, such as rail inclusions or dents caused by entrapped ballast stones etc. Simon et al. [84] performed tribological and metallurgical examinations of a number of squats and presented a possible initiation mechanism. The underlying idea is that the rail surface can be divided into longitudinal contact strips. These are subjected to different loading conditions, leading to different amounts of plastic deformation, which in turn introduce different types of microstructures (as well as WELs). As the material becomes more and more dissimilar between the contact strips, the borders tend to be prone to crack initiation, which in turn might initiate a squat. Jörg et al. [50] introduced the “Squat staircase concept”, which aims at describing squat formation in a qualitative way. The idea is that certain parameters linked to squat formation (for example contact conditions, wear, traction and system stiffness), which would not cause a problem taken individually, can be combined in such a way that a squat initiation threshold is reached. Rail surface irregularities, for example, appear frequently on rail surfaces, but are not a guarantee for squat initiation. However, if combined with traction and some other unfavourable parameters, squat initiation could appear according to the theory. This approach could also explain the seeming randomness in squat formation. The question that remains is then to identify which conditions that are necessary for squats to initiate and relate them to operational scenarios.

So far, the squat formation process has been discussed up until the point of crack initiation, but the subject of existing squat cracks has also been studied in several publications. Bogdański et al. [14] used Finite Element (FE) analysis to investigate, among other things, how bulk stresses and liquid entrapment affect Stress Intensity Factors (SIFs). More studies followed in Bogdański et al. [13] and Bogdański and Brown [12], where, e.g., crack face friction and branching were examined. Farjoo et al. [35] used finite element analysis to investigate the influence of, e.g., bending stresses on SIFs. The importance of water entrapment was also considered. Farjoo et al. [34] have furthermore showed that SIFs increase with softer track foundation, which could explain the occurrence of squats close to areas with poor ballast (where the rail consequently deflects more).

3.2 Dynamic loads

Dynamic loads, caused by e.g. rail surface irregularities, can cause wheel–rail contact force magnitudes considerable higher than the corresponding static load. Since these load peaks will relate to more severe RCF damage, it is of great interest to study dynamic phenomena in relation to squats. This is typically done either by measurements or simulations. The former method allows for online squat identification, as done by Molodova et al. [68, 69], who used axle box acceleration data to identify typical squat frequencies. Axle box acceleration data in relation to squats have also been presented by Kaewunruen et al. [54].

Computer simulations are of course well suited to analyse dynamic wheel–rail contact forces, either in the context of squats, or rail surface irregularities that might develop into squats. Li et al. [66] developed a three dimensional transient FE² model and simulated the contact force response as a wheel traversed rail surface irregularities. The distance between peak values of the contact forces were related to visible rail damage of field samples. A squat growth hypothesis was presented and eventually supported by field observations, cf. Li et al. [64]. The FE model was later validated by Molodova et al. [67] by comparing simulated results to corresponding results provided from axle box acceleration measurements. Li et al. [65] have also used the model to investigate how the influence of bolt pretension of fishplates (used in insulation joints) influences squat initiation. Later, Zhao et al. have used similar models to investigate the influence of small rail surface irregularities [99], small areas of low friction [100] and fastening modelling in the context of surface irregularities [98]. Furthermore, Zhao et al. [96] have used field examinations in combination with the aforementioned FE model as well as a two-dimensional model to study local rolling contact fatigue defects in railway wheels. These defects resemble RCF clusters and it is suggested in the study that indentations on the wheel tread are essential for RCF cluster formation.

Naeimi et al. [72] used transient FE simulations together with a multiaxial fatigue criterion to identify crack initiation angles. These were later compared to a field sample and some agreement was found. Deng et al. [28] used a similar modelling technique to study different squat geometries with respect to fatigue life as well as number of cycles to ratchetting failure. In contrast to the aforementioned studies, which mainly relate to crack initiation, Zhao et al. [101] also used explicit FEM to compute stress intensity factors of a cracked rail head during a wheel passage.

3.3 Thermal damage

As touched upon previously, so-called White Etching Layers (WELs) consist of a hard and brittle material phase. Since the brittle material may shatter, it is believed that WELs are prone to crack initiation. It should however be noted that WELs may also remain intact. WELs are commonly believed to contain martensite. This material phase has a larger volume than pearlite [1]. Martensite formation may therefore cause tensile

²See Section 5.2 for a background to transient finite element analysis.

residual stresses in the vicinity of the martensite, which promotes RCF initiation.

As pointed out by e.g. Vargolici et al. [92], there are two common ways of explaining WEL formation—mechanical or thermal origin. The kind of damage associated to studs seem to lack significant plastic deformation, therefore the latter initiation mechanism will be the main focus here. The idea is then that heat is causing a temperature increase of the steel, which, if severe enough, will cause austenization. If this state is followed by rapid cooling, another phase transformation will occur and a thin layer of martensite will form. For a general discussion on WEL in relation to RCF, see e.g. Carroll and Beynon [20, 21].

In addition to studies regarding the WEL phenomena itself, several publications treat WEL in relation to squats and studs³. Pal et al. [73] have performed metallurgical investigations of squats found in field and also suggested that a WEL can act as an initiator of conventional squats [74]. Steenbergen and Dollevoet [86, 87] have described a possible initiation and growth mechanism where the initiation takes place in a WEL. Even though grinding is often used to remove squat and stud defects, Steenbergen [88] points out that it can also be harmful, especially in the case of heat treated grades, in the sense that a WEL might be introduced by grinding. Scott et al. [81] simulated traction control systems and showed that the temperature increase caused by adhesion recovery after wheel slip can cause martensite formation in some situations. Naeimi et al. [71] used a coupled thermomechanical three-dimensional FE analysis to study frictional heating appearing in the wheel–rail contact as well as the resulting temperature increase and stress–strain response. It was found that the austenization temperature was reached at high creepage levels (26% in the simulation). Bernsteiner et al. [11] have used results from computer simulations and test rig experiments to show that austenization and subsequent WEL formation on rail surfaces are possible during some circumstances. Fletcher [36] used the boundary element method to study a stud crack in the presence of a mechanical contact load in combination with a thermal load and computed, among other things, stress intensity factors. The study was followed up by Fletcher and Sanusi [37], where the influence of volume expansion due to phase transformation from pearlite to martensite was investigated. It was shown that the formation of a WEL might accelerate crack growth due to unfavourable stress distributions.

A related area, which is mainly important in relation to RCF clusters, but has bearing on studs, is that of martensite formed during wheel flat formation⁴. This topic was investigated by Jergéus [44, 46] via two-dimensional FE simulations that included phase transformation. Residual stresses as well as resulting martensite depths were evaluated. The results were later compared to full scale field experiments followed by sectioning and metallurgical examination, see Jergéus et al. [45] and Ahlström and Karlsson [2].

³Note that not all authors apply this terminology. The name “squat” figures also in the context of WELs.

⁴In the case of studs, the corresponding phenomenon is a “wheel burn” mark on the rail head.

4 Modelling of wheel–rail contact

The ability to accurately resolve wheel–rail contact stresses is of great importance in a rolling contact fatigue assessment, especially in cases related to squats and RCF clusters. To this end, the analysis of a wheel rolling along a rail requires the application of a rolling contact theory able to predict the size of the contact patch as well as the magnitude and distribution of the normal contact pressure. However, most often it is not sufficient to consider the normal contact pressure alone. Tangential shear tractions, stemming from frictional slip inside the contact patch, often have a substantial contribution to the stress state. The resultants to the tangential tractions are known as creep forces and occur during, e.g., curving, acceleration and braking of the rail vehicle. The creep forces are important for the dynamic behaviour of the vehicle and depend on certain rolling characteristics described below. The contact stress distribution will also determine the stress field inside the wheel and the rail. This stress field will govern propagation of surface initiated cracks.

The following chapter begins with the introduction of a few basic concepts. It is followed by some analytical solutions to the rolling contact problem. Although there are limitations to these theories, they provide an inexpensive way to compute contact stresses, and increase the overall insight into the subject. The chapter continues with more general theories that require numerical solutions. Rolling contact analysis is a wide research topic in itself and is only covered briefly here. See e.g. Johnson [48] and Kalker [57] for thorough discussions on the subject.

4.1 Basic concepts of rolling contact

In this section, basic concepts of rolling contact are illustrated by considering a wheel (here denoted body 1) with radius R rolling along a tangent track (here denoted body 2). The bodies are initially considered to be rigid. The translational velocity is denoted v and the angular velocity Ω , see Figure 4.1. If $v = R\Omega$, there is no relative velocity between the wheel and the rail and the corresponding points on the wheel and rail are said to “stick” to each other. This situation is known as free rolling. The other possibility, where $v \neq R\Omega$ is known as “slip”, which means that there is relative motion between a contacting point on the wheel and the corresponding point on the rail. In the extreme case, $\Omega = 0$ while $v \neq 0$ and the wheel slides along the rail. For a more detailed introduction to rolling contact, see e.g. [82].

If the deformation of the contacting bodies is taken into account, the contact point will be replaced by a contact area. This allows for an intermediate state between stick and full slip, where there is stick in part of the contact patch and slip in the remaining part. The ability to characterise the situation thus requires that the motion of particles inside the contact patch is resolved, see Figure 4.2. The velocity of a point is then divided into two parts; one stemming from rigid body motion (v_1, v_2) and one due to deformation (\dot{u}_1, \dot{u}_2). The so-called slip velocity \dot{S} is then introduced as the velocity difference between

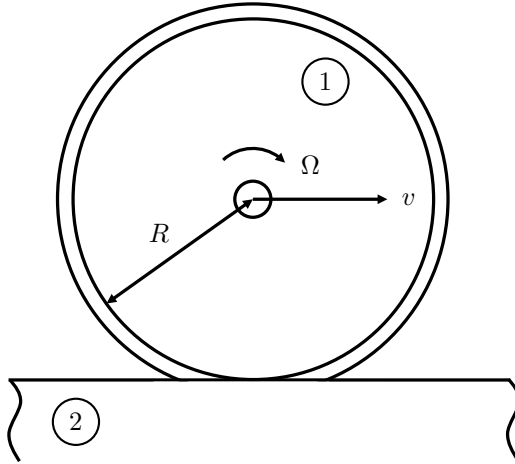


Figure 4.1: A wheel, denoted body 1, with radius R is rolling along a rail, denoted body 2, at a translational velocity v and an angular velocity Ω .

two, originally opposite, points [48, 57, 82]¹:

$$\dot{S} = v_1 + \dot{u}_1 - (v_2 + \dot{u}_2) \quad (4.1)$$

where $\langle \dot{\bullet} \rangle$ denotes time derivative. By rearranging the terms and introducing the longitudinal creepage² as

$$\xi = \frac{v_1 - v_2}{v} \quad (4.2)$$

The slip velocity can be expressed as

$$\dot{S} = \xi v + \dot{u}_1 - \dot{u}_2 \quad (4.3)$$

By using the Coulomb model as a constitutive model of the friction, it is possible to formally define areas of stick and slip respectively [57, 82]:

$$\text{Stick region: } \dot{S} = 0, |q(x)| \leq p(x) \mu \quad (4.4)$$

$$\text{Slip region: } \dot{S} \neq 0, q(x) = -\mu p(x) \frac{\dot{S}}{|\dot{S}|} \quad (4.5)$$

where $p(x)$ is the wheel–rail contact pressure, $q(x)$ denotes the interfacial shear traction and μ is the friction coefficient.

¹It is sometimes convenient to use an Eulerian description of motion, where the reference frame is tracing the contact patch and the material is considered to flow into the contact, see, e.g., [48, 57, 82]. In that case \dot{u} is replaced by the material time derivative.

²The spin creepage, occurring from a difference in angular velocity in the plane of contact, is not considered here. For details, see e.g. [48, 57].

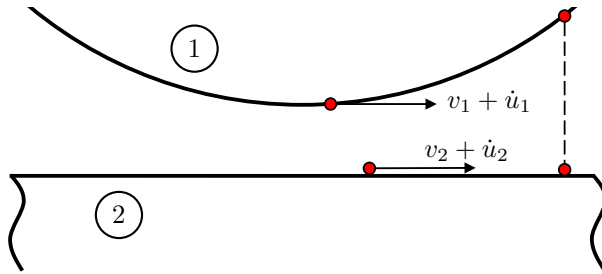


Figure 4.2: *Relative velocity between two, originally aligned, particles in rolling-sliding contact.*

4.2 Analytical solutions

4.2.1 Normal contact – Hertz solution

In 1882 Hertz [41] presented a theory on elastic bodies in contact. The theory, which is widely used within contact mechanics, is based on the following assumptions [48, 82]:

- Isotropic, linear elastic material.
- Small strains.
- The topology of the contacting surfaces close to the contact patch can be approximately described by quadratic functions.
- The bodies in contact can be approximated as elastic half-spaces, which implies that the contact area is small compared to the contacting bodies. This excludes the analysis of bodies where the contacting surfaces fit each other well and thus produce a large contact area, so-called conformal contact.
- Frictionless contact surfaces. This requirement is not relevant in the case of quasi-identical materials in contact³.

If these requirements are fulfilled, the contact patch will become elliptical with semi-axes a and b [48], see Figure 4.3. The magnitudes of a and b depend, in a nonlinear fashion, on the loading, the elastic properties and the radii of curvature of the contacting bodies.

If N denotes the resulting normal contact force, the maximum contact pressure p_0 is given by

$$p_0 = \frac{3N}{2\pi ab} \quad (4.6)$$

The pressure distribution over the contact patch is then obtained as

$$p(x, y) = p_0 \sqrt{1 - \left(\frac{x - \bar{x}}{a}\right)^2 - \left(\frac{y - \bar{y}}{b}\right)^2} \quad (4.7)$$

³This means that the shear modulus G and Poisson's ratio ν of bodies 1 and 2 fulfill $\frac{G_1}{1-2\nu_1} = \frac{G_2}{1-2\nu_2}$ [82].

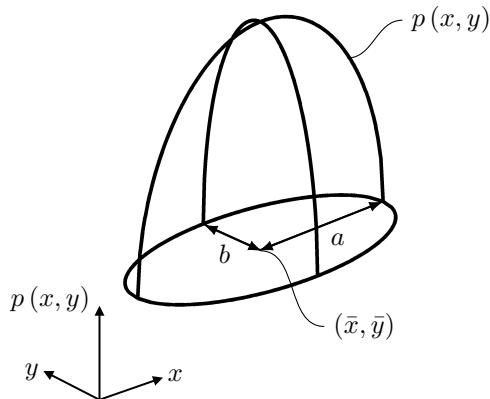


Figure 4.3: *Schematic illustration of a Hertzian pressure distribution.*

where x and y are coordinates along the surface and (\bar{x}, \bar{y}) is the center of the contact patch [48], see Figure 4.3.

The aforementioned limitations are often violated to some extent in the contact between wheel and rail. Plasticity, just to take one example, will bound the magnitude of the contact pressure and influence the size of the contact patch. Nevertheless, the closed form solution offers many benefits, not least in a vehicle–track interaction analysis where computational speed is an important issue.

4.2.2 Tangential contact – Carter solution

One of the pioneers within the area of rolling contact was Carter, who studied tractive rolling of cylinders, i.e. line contact [22]. The corresponding pressure distribution is obtained by letting $b \rightarrow \infty$ in Equation (4.7). For convenience, x is furthermore assumed to originate from the center of the contact patch, resulting in the relation $p(x) = p_0 \sqrt{1 - \left(\frac{x}{a}\right)^2}$. Carter showed that, in the case of Coulomb friction, the shear traction distribution can be obtained as a superposition of two parts, see e.g. [48]:

$$q(x) = q'(x) + q''(x) \quad (4.8)$$

where

$$q'(x) = \mu p(x) = \mu p_0 \sqrt{1 - \left(\frac{x}{a}\right)^2}, \quad -a \leq x \leq a \quad (4.9)$$

$$q''(x) = -\frac{c}{a} \mu p_0 \sqrt{1 - \frac{(x-d)^2}{c^2}}, \quad -(c-d) \leq x \leq a \quad (4.10)$$

Here, c is the half-length of the stick zone and d is the distance from the center of the stick zone to the center of the contact patch, see Figure 4.4. The maximum creep force is obtained during full slip conditions, i.e. when $q''(x)$ vanishes and $q(x) = \mu p(x)$.

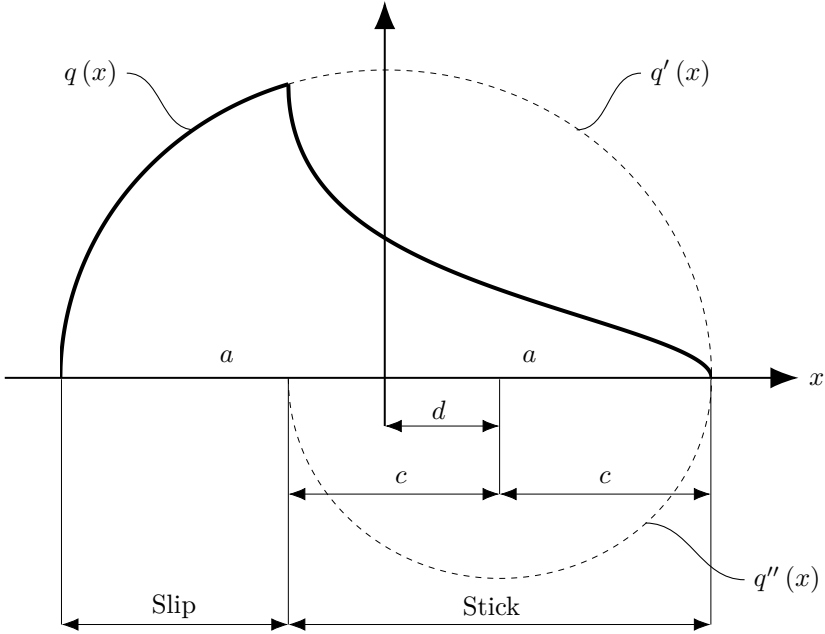


Figure 4.4: *Sketch of the shear stress distribution in tractive rolling, according to the Carter solution. Here, a denotes half-length of the contact patch, c is the half-length of the stick zone and d is the distance from the center of the contact to the midpoint of the stick zone.*

By integrating the shear traction it is possible to obtain the resulting creep force Q as:

$$Q = \int_{-a}^a q(x) dx \quad (4.11)$$

It can be shown, see e.g. [48], that the creepage is related to the creep force Q as:

$$\xi = -\frac{\mu a}{R} \left(1 - \sqrt{1 - \frac{Q}{\mu N}} \right) \quad (4.12)$$

where N is the resulting normal force, $\frac{1}{R} = \frac{1}{R_1} + \frac{1}{R_2}$ with R_i being the radius of body i . When the creepage becomes sufficiently high, the contact patch will reach a state of full slip and the creep force saturates at $Q = \mu N$. The principal appearance is shown in a so-called creep curve as presented in Figure 4.5. The curve becomes horizontal during full slip conditions, which is a consequence of the Coulomb friction model assumption. A real world creep curve might look somewhat different, e.g. due to the influence of so-called third body layers. To some extent it is in practice possible to control the slope of the creep curve inside the slip zone by the introduction of so-called friction modifiers [63, 89].

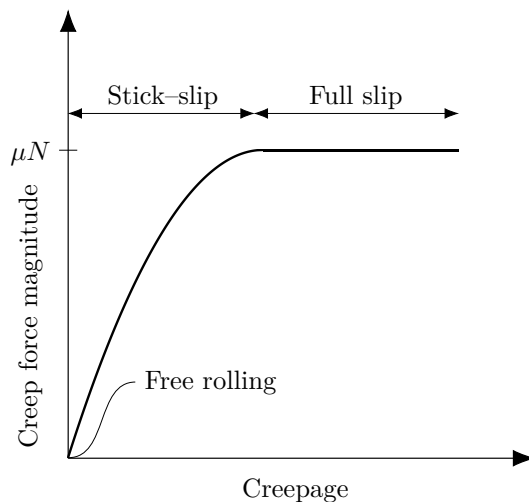


Figure 4.5: *Schematic relation between creepage and creep force magnitude according to the Carter theory.*

4.3 Kalker’s complete and simplified theories

As mentioned above, the Carter solution is restricted to two-dimensional situations. Several approximate theories of three-dimensional rolling contact (also including lateral and spin creepage) have been presented throughout the years, see e.g. Kalker [58]. A milestone occurred in 1979 when Kalker presented what has become known as his complete theory⁴ of rolling contact [56, 57]. It is here presumed that the contacting bodies can be considered as linear elastic half-spaces [82]. The theory was implemented by use of the Boundary Element Method⁵ (BEM) in the computer software CONTACT [93], which is able to compute the pressure and traction distributions between general (to the point of the underlying assumptions) bodies in rolling contact. It should be noted that some of the original limitations, such as restriction to non-conformal contact, can be circumvented [94].

Kalker has also presented a simplified theory, on which the algorithm FASTSIM [55] is based. The main difference, as compared to the complete theory, is that the displacement of one surface point is considered to depend only on the surface tractions in that particular point [58]. That is not presumed in the complete theory, where the surface displacement in one point depends on the surface tractions in all points. FASTSIM is considered to be fast [58] and is often used multibody simulation software⁶ [82].

⁴Other names such as Kalker’s variational theory and Kalker’s exact theory occur in the literature [82].

⁵See Section 6.1 for a description of the Boundary Element Method.

⁶See Section 5.3 for a discussion on such software.

4.4 Finite element method

The most general way to solve the rolling contact problem is probably provided by the Finite Element Method (FEM), which is a method to solve partial differential equations in an approximate manner, see e.g. [25, 103]. FEM makes it possible to relax presumptions such as elastic half-spaces, linear elasticity, isotropy etc. The limiting factor is thus rather the computational cost than assumptions related to the underlying theory.

The generality of the FE approach makes it suitable as a reference to assess the validity of the classical contact theories mentioned previously. One such comparison was done by Telliskivi and Olofsson [91] who computed the maximum contact pressure as well as contact patch size using FEM (including an elasto-plastic constitutive model), CONTACT and Hertz theory. It was found that the maximum pressure became quite similar between the three methods when contact occurred on top of the rail. The agreement was however poor during flange contact, partly because the half-space assumption is violated as the contact becomes conformal. Zhao and Li [102] utilised an explicit FE analysis to solve the transient problem of a cylindrical wheel in frictional rolling on top of a rail. The position of the contact patch in combination with a linear elastic material model allowed for a comparison with CONTACT and Hertz theory. Good agreement in terms of size of the contact patch as well as contact stresses were found. In a succeeding study, the model was enhanced with an isotropic hardening plasticity model [97]. It was shown that in the case of plastic flow, the contact patch size increases and becomes egg shaped rather than elliptic, while the maximum contact pressure decreases. It was also shown that it is not possible to decouple the normal and tangential problems under such conditions, as is commonly done otherwise.

Within solid mechanics, a so-called Lagrangian description of the motion is usually employed. This means that specific particles are tracked during motion. In an FE setting, the FE mesh is thus connected to the body of interest and will distort during deformation, which might cause problems when large deformations occur. Further, if a wheel is rolling a long distance along the rail, the full circumference of the wheel must be discretized by a dense mesh in order to resolve the contact stresses. This is computationally expensive. As an alternative, the so-called Arbitrary Lagrangian-Eulerian description can be employed. It is then possible to use the undeformed configuration following from a rigid body motion as a reference configuration. In many applications that configuration will furthermore be more or less constant in time, which means that the mesh refinement can be concentrated to the area of contact, which decreases the computational cost significantly. See e.g. [31] for a thorough description of the subject.

5 Simulation of dynamic vehicle–track interaction

Tractive and braking forces are transferred to the rail vehicle through the relatively small contact patch between the wheel and the rail. Contact forces will affect the vehicle dynamics, but vehicle dynamics will also affect the contact forces. In many situations this condition makes it necessary to account for the dynamic vehicle–track interaction in order to get detailed information on the contact stresses, which in turn is needed for fatigue and strength assessments. Different approaches to vehicle–track dynamics are therefore discussed in this chapter. Special emphasis is put on the method of moving Green’s functions since it is utilised in several of the appended papers. The corresponding section has previously been presented in the licentiate thesis by Andersson [9], but is reiterated here for clarity.

5.1 Method of moving Green’s functions

The convolution integral method, also known as the Duhamel integral method, is used to compute the dynamic response of linear systems subjected to general load histories. The concept is demonstrated in the subsequent example, which follows [26], where also a rather extensive background is provided.

For illustration, a system with one translational degree of freedom (DOF), $u(t)$, describing the position of a mass m attached to a linear spring with stiffness k is considered. The system is subjected to an impulse acting during the time t_d :

$$I = \int_0^{t_d} F(t) dt \quad (5.1)$$

where t is the time and $F(t)$ is the force. If the mass is initially at rest, the equation of motion and its initial conditions can be stated as

$$m\ddot{u} + ku = \begin{cases} F(t), & 0 \leq t \leq t_d \\ 0, & t > t_d \end{cases} \quad (5.2)$$

$$u(0) = \dot{u}(0) = 0 \quad (5.3)$$

where \dot{u} and \ddot{u} are the first and second time derivatives of the displacement, respectively. Integrating Equation (5.2) over the interval $0 \leq t \leq t_d$ gives the following relation

$$m\dot{u}(t_d) + ku_{\text{avg}}t_d = I \quad (5.4)$$

where u_{avg} is the the average displacement. Consider now the solution for $t_d \rightarrow 0$ whereby

$$m\dot{u}(0^+) = I \quad (5.5)$$

which yields the following initial conditions for interval $t > t_d$ of Equation (5.2)

$$\begin{cases} \dot{u}(0^+) = \frac{I}{m} \\ u(0^+) = 0 \end{cases} \quad (5.6)$$

Solving the resulting differential equation for $t > 0$ gives

$$u(t) = \frac{I}{m\omega_n} \sin(\omega_n t) \quad (5.7)$$

where $\omega_n = \sqrt{\frac{k}{m}}$. The unit impulse response, also known as a so-called Green's function, denoted $G(t)$, is obtained by putting $I = 1$:

$$G(t) = \frac{1}{m\omega_n} \sin(\omega_n t) \quad (5.8)$$

Equation (5.8) describes how a suspended mass will oscillate due to a unit impulse acting over a short time. This result can be exploited when the mass is subjected to an entire load history. The impulse due to a load $F(\tau)$ acting during the small time interval $d\tau$ can be expressed as

$$dI = F(\tau) d\tau \quad (5.9)$$

Combining Equations (5.7)–(5.9) gives the displacement increment at time point t due to the load applied at time τ :

$$du(t) = \frac{dI}{m\omega_n} \sin(\omega_n(t - \tau)) = F(\tau) G(t - \tau) d\tau \quad (5.10)$$

The total displacement response at an arbitrary time point t is obtained by summing the contributions of all impulse responses from the previous time steps:

$$u(t) = \int_0^t F(\tau) G(t - \tau) d\tau \quad (5.11)$$

In this example, the Green's function $G(t - \tau)$ was derived for the case of an undamped 1-DOF system. However, Equation (5.11) represents a general (as long as the superposition principle holds) expression for how the dynamic response can be obtained. For more complicated systems, like a track model consisting of beam elements resting on discrete springs and dampers, the Green's functions are not derived analytically. Instead Frequency Response Functions (FRFs), i.e. functions describing the response in one point of the dynamic system due to a harmonic excitation at another point, are computed in a standard manner using the system matrices of an FE model. The Green's functions can then be obtained by computing the inverse Fourier transform of the FRFs [76].

In this thesis, Green's functions are used to evaluate the dynamic response beneath a wheel moving along a track. To this end, the concept of moving Green's functions is introduced [76]. For simplicity a homogeneous rail is henceforth considered although the concept can easily be extended to involve discrete sleeper support. A general Green's

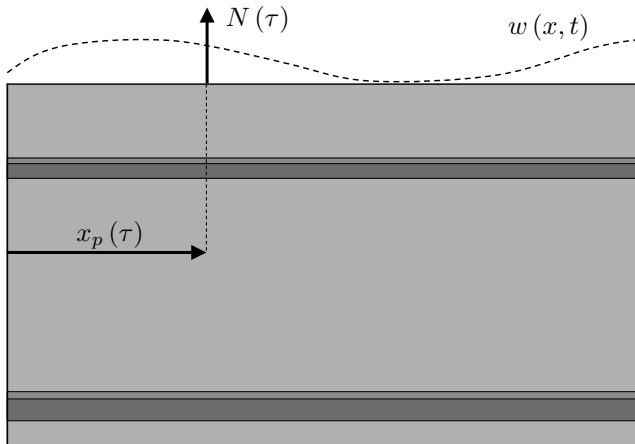


Figure 5.1: The load $N(\tau)$ acting at coordinate x_p at time τ causes the displacement $w(x, t)$ for $t > \tau$.

function, which is allowed to vary both in time and along the longitudinal coordinate x , is established: $G = G(x, t)$. The wheel's position is described by $x_p(t)$ and the wheel–rail normal contact force is denoted $N(t)$. The rail displacement $w(x, t)$ due to an applied contact force $N(t)$, see Figure 5.1, is calculated as

$$w(x, t) = \int_0^t N(\tau) G(x - x_p(\tau), t - \tau) d\tau \quad (5.12)$$

As touched upon previously, the displacement of the rail beneath the wheel is typically of primary interest. Therefore it is sufficient to consider the dynamic response at the wheel's current position. This is also the point of excitation. If the wheel is restricted to move only at constant velocity v , the load position at time $t = \tau$ becomes $x_p(\tau) = x_0 + v\tau$, where x_0 is the initial position. Hence, by use of Equation (5.12) the displacement beneath the wheel (indicated by $\tilde{\cdot}$) can be expressed as

$$w(x_0 + vt, t) = \tilde{w}(t) = \int_0^t \tilde{N}(\tau) G(v(t - \tau), t - \tau) d\tau \quad (5.13)$$

If the Green's function G , the load history applied to the wheel $\tilde{N}(\tau)$, as well as the time difference $t - \tau$ is known, it is possible to evaluate the integral. Note that for a given dynamic system and constant velocity v , the Green's function is a function of time alone. The moving Green's function $\tilde{G}_v(t - \tau)$ is now introduced and Equation (5.13) becomes

$$\tilde{w}(t) = \int_0^t \tilde{N}(\tau) \tilde{G}_v(t - \tau) d\tau \quad (5.14)$$

The moving Green's function $\tilde{G}_v(t)$ is obtained by considering position $v(t - \tau)$ of $G(x, t)$ at time $t - \tau$. Figure 5.2 schematically illustrates how a moving Green's function $\tilde{G}_v(t)$ (indicated by the thick solid line) is obtained from a generic Green's function $G = G(x, t)$.

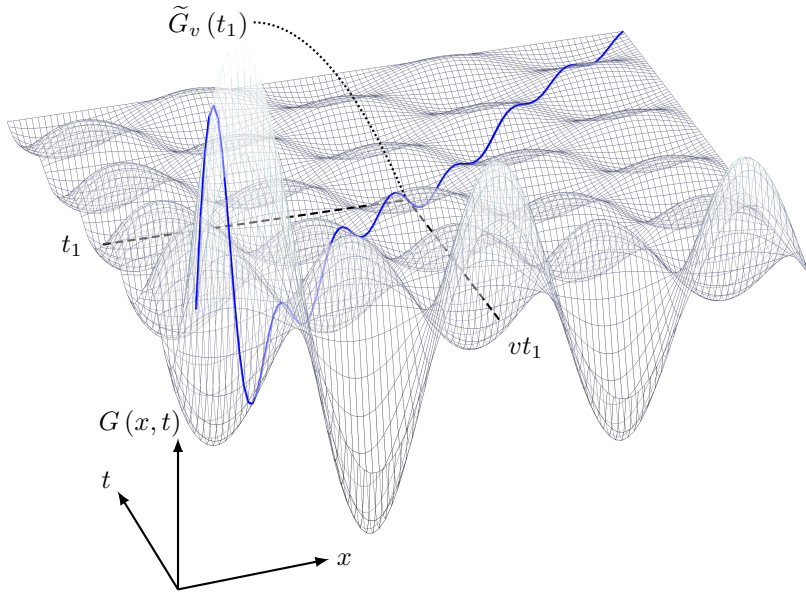


Figure 5.2: The surface indicates a generic Green's function $G = G(x, t)$ from which the moving Green's function $\tilde{G}_v(t)$ (marked by the thick solid line) of a point travelling with a constant velocity v is constructed.

The method of moving Green's functions is implemented in an in-house software for dynamic vehicle–track interaction—RAVEN, see **Paper B**. Note that the method is not restricted to solve the dynamic response of a track. This is utilised in most of the simulations in the appended papers, where Green's functions are also used to compute wheel displacements. A summary of benefits and drawbacks of the described modelling technique are:

- + Pre-calculated Green's functions enable fast simulations.
- + High-frequency dynamic behaviour can be accounted for without adding significant computational costs.
- Only the response of linear systems can be analysed, which means that non-linear characteristics of e.g. ballast (or suspension if modelling the vehicle) are not captured. In contrast, the cause of excitation (in this case the wheel–rail contact force) can be due to non-linear phenomena (in this case the non-linear wheel–rail contact stiffness).
- The trajectory of the contact point needs to be prescribed, which makes the modelling technique inappropriate for investigations of low-frequency lateral vehicle dynamics (e.g. dynamic stability analyses).

5.2 Transient finite element analysis

One great benefit of using a finite element approach is the possibility to directly incorporate nonlinearities stemming from e.g. frictional contact and plasticity. As mentioned previously, this makes it possible to relax assumptions related e.g. to the contact model. By considering the full transient response it is possible to analyse the full problem of rolling contact. In many situations it is then possible to write the resulting equations of motion on the following form, cf. [25]:

$$\mathbf{M}\ddot{\mathbf{a}} + \mathbf{C}\dot{\mathbf{a}} + \mathbf{K}\mathbf{a} = \mathbf{f} \quad (5.15)$$

where \mathbf{M} is the mass matrix, \mathbf{C} is the damping matrix, \mathbf{K} is the stiffness matrix, \mathbf{f} is the load vector, \mathbf{a} are the nodal displacements and $\langle \bullet \rangle$ denotes time derivatives. One common and quite general method to solve Equation (5.15) is to reformulate the time derivatives in terms of finite differences and then use direct numerical integration. Two main classes of integration schemes exist, namely *explicit* and *implicit* methods [25]. The former expresses the displacements of the next time step ($n + 1$) solely in terms of quantities related to previous time steps:

$$\mathbf{a}_{n+1} = f(\mathbf{a}_n, \dot{\mathbf{a}}_n, \ddot{\mathbf{a}}_n, \mathbf{a}_{n-1}, \dots) \quad (5.16)$$

Implicit methods, on the other hand, express the displacements of time step $n + 1$ in terms of quantities related to previous time steps as well as the next ($n + 1$):

$$\mathbf{a}_{n+1} = f(\dot{\mathbf{a}}_{n+1}, \ddot{\mathbf{a}}_{n+1}, \mathbf{a}_n, \dot{\mathbf{a}}_n, \ddot{\mathbf{a}}_n, \dots) \quad (5.17)$$

There are two important differences between explicit and implicit solution methods: explicit schemes are only conditionally stable, which means that a too large time step will cause numerical instabilities. Implicit schemes, on the other hand, are unconditionally stable, allowing for larger time steps. However, the computational cost of taking one time step is lower for an explicit method than for an implicit. If the high frequency dynamics is of importance, as is often the case during e.g. a wheel passage over a surface irregularity, it is often efficient to use an explicit solver since many time steps are required even for a short travel distances. Note that although the principles above are illustrated using a problem from solid mechanics, the integration schemes are applicable also to other engineering disciplines.

Some benefits and drawbacks of transient FEM are:

- + Dynamics, contact modelling, stress evaluation etc. can be incorporated into a single analysis.
- + No formal need for approximate theories, such as the elastic half-space assumption used in Hertzian contact theory.
- The overall solution is computationally expensive.

5.3 Multibody simulation

Dynamic vehicle–track interaction is commonly analysed using so-called multibody simulation software. Well-known titles in this genre include SIMPACK, GENYSYS, VAMPIRE, ADAMS/Rail and NUCARS [43]. The vehicle is typically considered to consist of rigid bodies (although some software allow for flexible bodies as well) connected via flexible elements resembling suspensions, dampers etc. The resulting equations of motion of the system are then formulated automatically by the software. As mentioned before, an accurate wheel–rail model is, in general, nonlinear, which adds additional constraints to the system of equations. Even though many different analysis options are commonly available, the most relevant in the context of this thesis is numerical time integration of the full nonlinear system in order to resolve the wheel–rail contact stresses. It is however important to note that it is necessary for the software to include high-frequency dynamics if a small rail surface irregularity (e.g. a dimple as in **Paper A–C** and **Paper F**) is to be analysed. For a thorough discussion on multibody simulations, see e.g. Iwnicki (ed.) [43].

6 Analysis of rail and wheel damage

The ultimate goal of dynamic vehicle–track interaction simulations in general and wheel–rail contact stress analyses in particular, is often to relate the corresponding results to some sort of deterioration assessment. Within railway engineering two such mechanisms are of particular interest—rolling contact fatigue and wear. In this thesis, the emphasis is on the former. The RCF impact is usually quantified through either some crack initiation criterion or by means of fracture mechanics. Some of the available alternatives are treated in this chapter.

6.1 Stress–strain analysis

Before any fatigue (or fracture) assessment criterion can be used, it is commonly necessary to know the state of stress and strain within the body of interest. In the field of railway mechanics, essentially three methods are used to obtain these—analytical methods, the Boundary Element Method (BEM) and the Finite Element Method (FEM).

Analytical methods, to begin with, are often restricted to linear elastic isotropic materials and only valid under special conditions. This does, however, not prevent them from being useful in many situations. The most common example is probably Hertzian contact theory, from which it is possible to evaluate contact stresses that can be used to derive the stress fields in the contacting bodies [48]. No analytical methods are however used for stress or strain evaluation of the analysed bodies in the appended papers.

As an alternative, BEM approaches might be used. These are based on the existence of a fundamental solution that fulfils the boundary value problem for a unit point load [25]. For linear problems, more complicated load cases can then be obtained as combinations

and scalings of unit loads. The fundamental solution must be known beforehand, which often puts restriction on the generality of the solution. On the other hand it is possible to incorporate e.g. stress singularities in the fundamental solution, which is convenient within fracture mechanics, see Section 6.2.6. Once the fundamental solution is known, BEM only requires the discretization of the body's surface and not the entire volume. This often leads to a drastic decrease of the computational cost, as a three-dimensional problem can be treated as a two-dimensional etc. [25]. BEM is a rather popular method within railway mechanics where the computer implementations of Kalker's complete theory [57] is probably the most famous example. An implementation of this theory is used in **Paper A–Paper C** and **Paper F**.

Various aspects of FEM have been treated in Sections 4.4 and 5.2. Its usefulness in solving general solid mechanics problems has already been mentioned. However, applying conventional FEM to cracked bodies is problematic for several reasons; it can be cumbersome to achieve a good mesh quality in the vicinity of the crack, which makes it laborious to change crack geometry or analyse crack growth. Another disadvantage is that finite stresses will be obtained at crack tips, even though the theory of elasticity predicts a stress singularity, see Section 6.2.6. These problems have largely been resolved with methods such as the eXtended Finite Element Method (XFEM), see Belytschko and Black [10] and Moës et al. [70].

XFEM constitutes a large research topic in itself and will not be covered in detail here. The method will instead be briefly illustrated via an example following Moës et al. [70]. To this end a cracked isotropic linear elastic two-dimensional body is considered. The crack does not coincide with element boundaries, but is embedded inside the relevant elements, see Figure 6.1. In the example there are two elements completely split by the crack, one element containing the crack tip and nine elements not containing any crack at all. The FE approximation of the displacement field $\mathbf{u}^h(x, y)$ belonging to the latter nine elements follows the standard format:

$$\mathbf{u}^h(x, y) = \mathbf{N}(x, y) \mathbf{a} \quad (6.1)$$

where $\mathbf{N}(x, y)$ are shape functions and \mathbf{a} are the nodal displacements.

The conventional FE formulation above does not hold for the elements cut by the crack. Consider for example a case where the lower boundary of Figure 6.1 is clamped and a load in the y -direction is applied to the upper boundary. The vertical displacement of the upper crack face must be larger than on the lower. A crack splitting an element will in other words introduce a discontinuity in the displacement field. This can be accounted for by enriching the standard FE approximation of the cracked elements with an additional term:

$$\mathbf{u}^h(x, y) = \mathbf{N}(x, y) \mathbf{a} + \mathbf{N}(x, y) H(x, y) \mathbf{b} \quad (6.2)$$

where \mathbf{b} are additional degrees of freedom and the Heaviside function $H(x, y)$ is defined as:

$$H(x, y) = \begin{cases} 1, & y > 0 \\ -1, & y < 0 \end{cases} \quad (6.3)$$

The singularity occurring inside the crack tip element can be captured by adding the following enrichment to the standard FE approximation:

$$\mathbf{u}^h(x, y) = \mathbf{N}(x, y) \mathbf{a} + \sum_{i=1}^4 \mathbf{N}(x, y) F_i(r, \theta) \mathbf{c}_i \quad (6.4)$$

where (r, θ) are polar coordinates originating from the crack tip, see Figure 6.3, \mathbf{c}_i are additional (enriched) nodal degrees of freedom and $F_i(r, \theta)$ contain asymptotic crack tip functions, cf. [70]:

$$F_1(r, \theta) = \sqrt{r} \sin\left(\frac{\theta}{2}\right) \quad (6.5)$$

$$F_2(r, \theta) = \sqrt{r} \cos\left(\frac{\theta}{2}\right) \quad (6.6)$$

$$F_3(r, \theta) = \sqrt{r} \sin\left(\frac{\theta}{2}\right) \sin(\theta) \quad (6.7)$$

$$F_4(r, \theta) = \sqrt{r} \cos\left(\frac{\theta}{2}\right) \sin(\theta) \quad (6.8)$$

For the sake of simplicity, the example above contains several simplifications. It is, for example, not necessary to use the same shape functions for the standard FE approximation and the enrichment functions, the crack tip enrichment is not restricted to only the element containing the crack tip and the crack does not have to be aligned with the global coordinate system. Nevertheless the example illustrates the overall idea with XFEM, which is to allow modelling of cracks not following the element boundaries. The crack topology is instead provided only implicitly. This allows for easy modifications of crack geometries compared to conventional FE procedures.

XFEM is used in **Paper D** and **Paper F**.

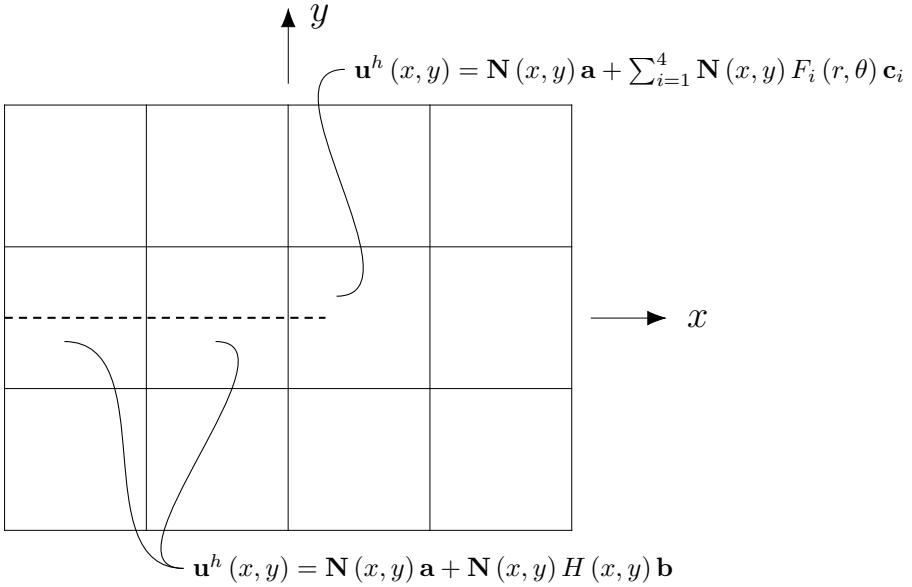


Figure 6.1: *FE mesh with a crack embedded within three elements and the corresponding XFEM approximation.*

6.2 Numerical approaches to RCF and fracture mechanics

6.2.1 Maximum contact stresses and forces

A straight forward approach for evaluating the risk of possible RCF damage is to extract maximum wheel–rail contact stresses and the corresponding resulting forces from a dynamic vehicle–track analysis. In a two-dimensional case, this results in four RCF impact candidates: maximum normal pressure $p(x)$, shear traction $q(x)$, normal force N and creep force Q . The overhead cost related to an assessment of these quantities is low since they are typically directly available as an output from multibody simulation tools. Some benefits and drawbacks are listed below:

- + The influence of local surface irregularities can be accounted for if a sufficiently accurate contact model is used.
- No detailed material response is taken into consideration.
- The RCF impact measures are rough with only implicit links to RCF initiation mechanisms.

This approach is used in **Paper A**.

6.2.2 Surface fatigue index

The surface fatigue index, FI_{surf} , was developed as an engineering tool, to be used in e.g. multibody dynamic simulations, for fast assessment of the risk for surface initiated RCF [32]. The idea is to parametrize the loading in terms of the normal force N and the traction coefficient, defined as the creep force to the normal force ratio: $f = Q/N$. Under the assumptions of Hertzian contact and full slip, it is possible to relate the stress state induced by the loading to the one needed for ratchetting to take place. This is quantified by dimensionless surface fatigue index, FI_{surf} [32]:

$$FI_{\text{surf}} = f - \frac{2\pi abk}{3N} \quad (6.9)$$

where a and b are semiaxes of the elliptic contact patch and k is the (cyclic) yield stress in pure shear. Some benefits and drawbacks of the model are listed below.

- + Low computational cost, which allows for online implementation in simulations of dynamic vehicle–track interaction.
- + Clear relation to a RCF initiation mechanism.
- Discrete surface irregularities cannot be accounted for since the assumption of Hertzian contact is not fulfilled.
- Only full slip conditions are considered.
- No detailed material response is taken into consideration.

This RCF impact model is not used in the appended papers since it was not developed to account for discrete surface irregularities.

6.2.3 $T\gamma$ approach

The so-called wear number $T\gamma$, which is a measure of the dissipated energy inside the contact patch, is defined as the product between in-plane creep forces and their corresponding creepages:

$$T\gamma = T_x\gamma_x + T_y\gamma_y \quad (6.10)$$

where T_x and T_y are longitudinal and lateral creep forces, while γ_x and γ_y are the longitudinal and lateral creepages respectively¹. By introducing a damage function calibrated towards field observations, it has been possible to relate the wear number not only to wear, but also to RCF damage [18]. Some benefits and drawbacks are listed below.

- + Low computational cost that allows for online implementation in simulations of dynamic vehicle–track interaction.
- + The competitive mechanism between wear and RCF is taken into account.

¹Note that this notation is not consistent with the one introduced in Chapter 4, where the longitudinal creep force and creepage are denoted Q and ξ respectively. It is nevertheless kept to be in accordance with the literature.

- The damage function must be calibrated towards the conditions under which it should be used.

This RCF impact model is not used in the appended papers since calibration of the damage function would be needed to account for surface irregularities.

6.2.4 Equivalent stress/strain

As touched upon previously, one of the mechanisms behind surface initiated RCF is ratchetting, i.e. unrestrained accumulation of strain under cyclic loading. This initiation mode is particularly relevant if the directions of the surface tractions do not vary over time [33]. By employing a suitable material model, able to capture nonlinear hardening, it is possible to study this phenomenon in an FE analysis. The state of stress and strain is however multiaxial and for a convenient analysis it is therefore necessary to introduce single scalar values to represent the stress and strain. Since metal plasticity is considered, the influence of hydrostatic stresses are negligible. To this end, the equivalent von Mises strain is introduced as a scalar representation of the strain:

$$\bar{\epsilon} = \frac{\sqrt{2}}{3} \sqrt{(\epsilon_{xx} - \epsilon_{yy})^2 + (\epsilon_{yy} - \epsilon_{zz})^2 + (\epsilon_{zz} - \epsilon_{xx})^2 + 6(\epsilon_{xy}^2 + \epsilon_{yz}^2 + \epsilon_{xz}^2)} \quad (6.11)$$

where ϵ_{ij} is a component of the total strain tensor. Similarly, the equivalent von Mises stress is introduced as a scalar representation of the stress:

$$\bar{\sigma} = \frac{1}{\sqrt{2}} \sqrt{(\sigma_{xx} - \sigma_{yy})^2 + (\sigma_{yy} - \sigma_{zz})^2 + (\sigma_{zz} - \sigma_{xx})^2 + 6(\sigma_{xy}^2 + \sigma_{yz}^2 + \sigma_{xz}^2)} \quad (6.12)$$

where σ_{ij} contains the components of the stress tensor. By plotting the von Mises stress against the strain it is possible to graphically indicate whether or not a state of shakedown is obtained. Some benefits and drawbacks are listed below.

- + This method can capture potential RCF initiation due to ratchetting.
- + The approach is rather general and suitable also in cases of local surface irregularities.
- Detailed stress and strain fields, typically obtained at a high computational cost, must be provided.
- An advanced material model is required to predict accurate evolution of local stresses and strains.
- The criterion is local whereas, in reality, fatigue is a “semi-local” phenomenon since steep stress gradients might initiate a crack which is soon halted.

This approach is used in **Papers B** and **C**.

6.2.5 Jiang–Sehitoglu fatigue parameter

If the surface tractions change directions over time, the net accumulation of strain will become low. Instead of ratchetting, low cycle fatigue will then become the dominating mode of RCF initiation in terms of damage accumulation [33]. This effect can be captured using the Jiang–Sehitoglu low cycle fatigue criterion [47]. The idea is to project stresses and strains onto a material plane, see Figure 6.2, in order to obtain the corresponding normal stress σ , normal strain ϵ , shear stress τ and engineering shear strain γ . The Jiang–Sehitoglu fatigue parameter can then be evaluated for a particular plane i according to:

$$FP_i = \left\langle \frac{\Delta\epsilon}{2} \sigma_{\max} \right\rangle + c_J \Delta\gamma \Delta\tau \quad (6.13)$$

Where

$$\sigma_{\max} = \max_{t \in L} [\sigma(t)] \quad (6.14)$$

$$\Delta\epsilon = \max_{t \in L} [\epsilon(t)] - \min_{t \in L} [\epsilon(t)] \quad (6.15)$$

$$\Delta\tau = \max_{t, \bar{t} \in L} [|\tau(t) - \tau(\bar{t})|] \quad (6.16)$$

$$\Delta\gamma = \max_{t, \bar{t} \in L} [|\gamma(t) - \gamma(\bar{t})|] \quad (6.17)$$

c_J is a material parameter, t and \bar{t} are instants in time related to the current load cycle L . Further, τ and γ are the vectorial shear stress and engineering shear strain of the studied plane and $\langle \bullet \rangle = 1/2 (|\bullet| + \bullet)$. Note that the $\langle \bullet \rangle$ notation is not present in [47] but occurs in different forms in the literature, cf. [53, 79]. The material plane giving the highest fatigue parameter is typically taken as representative for the fatigue damage:

$$FP = \max \left[\left\langle \frac{\Delta\epsilon}{2} \sigma_{\max} \right\rangle + c_J \Delta\gamma \Delta\tau \right] \quad (6.18)$$

where the maximization should be carried out over all material plane orientations. FP can then be related to the fatigue life N_f in the following manner [47]:

$$(FP - FP_0)^m N_f = C \quad (6.19)$$

where FP_0 , m and C are three additional material parameters. Some benefits and drawbacks are listed below.

- + The approach explicitly accounts for the influence of plastic deformation on fatigue damage.
- + The approach is rather general and can handle the influence of local surface irregularities.
- Detailed stress and strain fields, typically obtained at a high computational cost, must be provided.

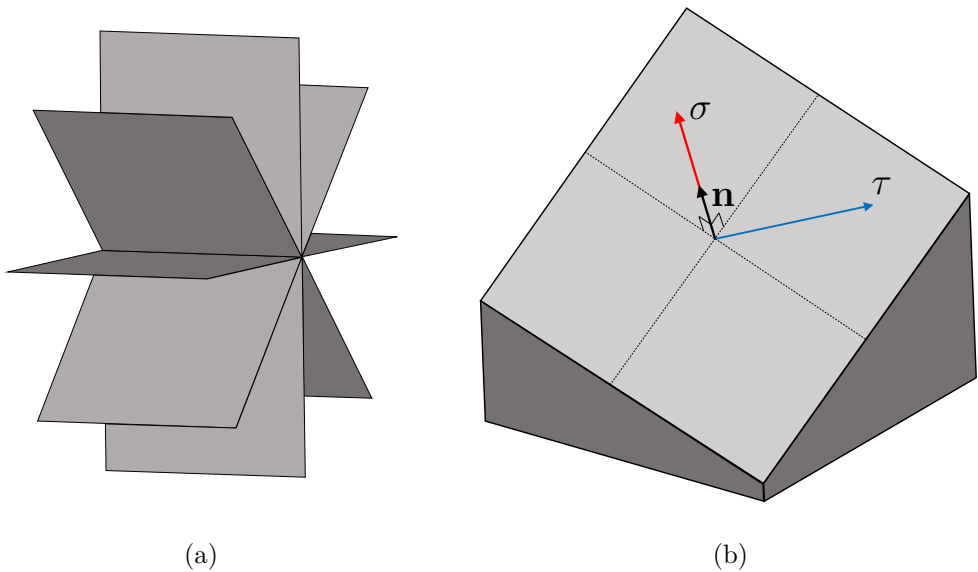


Figure 6.2: (a) Example of different material plane orientations in a point. Unlike in the figure, the material planes are rotated about all axes in an analysis. (b) Normal stress σ and shear stress τ acting on a material plane defined by the unit normal vector \mathbf{n} .

- It is computationally expensive to calculate the fatigue parameter for all material planes and positions of interest.
- The criterion is local whereas, in reality, fatigue is a “semi-local” phenomenon since steep stress gradients might initiate a crack which is soon halted.

FP is used as a measure of RCF impact in **Paper B**, **Paper C** and **Paper E**.

6.2.6 Fracture mechanics

For a linear elastic material response, stress magnitudes go to infinity in the vicinity of a sharp crack tip. It thus appears that the stress magnitudes are not useful quantifiers for the severity of a loaded crack. A more suitable choice is to use so-called Stress Intensity Factors (SIFs) defined as [62]:

$$\begin{aligned}
 K_{\text{I}} &= \lim_{r \rightarrow 0} \sigma_{yy}(r, \theta = 0) \sqrt{2\pi r} \\
 K_{\text{II}} &= \lim_{r \rightarrow 0} \sigma_{xy}(r, \theta = 0) \sqrt{2\pi r} \\
 K_{\text{III}} &= \lim_{r \rightarrow 0} \sigma_{yz}(r, \theta = 0) \sqrt{2\pi r}
 \end{aligned} \tag{6.20}$$

where r and θ are polar coordinates originating from the crack tip, see Figure 6.3. K_{I} , K_{II} and K_{III} are related to crack opening, in-plane shear and out-of-plane shear,

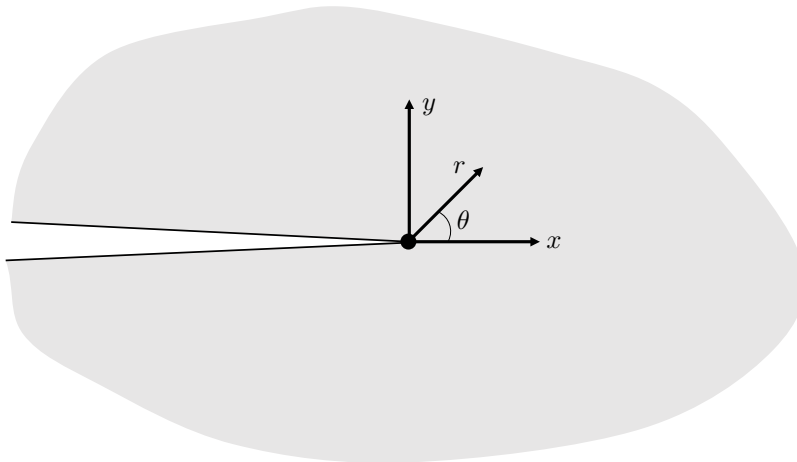


Figure 6.3: A crack with two coordinate systems originating from the crack tip—the polar r, θ -system and the Cartesian x, y -system (with z pointing out of the plane) respectively.

respectively, see Figure 6.4. Note that SIFs are only relevant under (predominantly) linear elastic conditions. This is commonly referred to as Linear Elastic Fracture Mechanics (LEFM) [30].

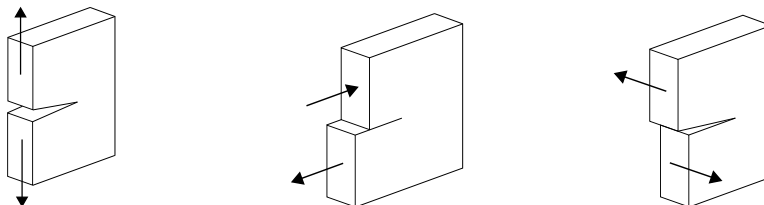


Figure 6.4: Opening mode (mode I), in plane shear mode (mode II) and out of plane shear mode (mode III).

The SIFs can be expressed as functions of the crack length a and nominal stress σ_0 [30]:

$$K = g\sigma_0\sqrt{\pi a} \quad (6.21)$$

where g is a function of the geometry and load characteristics. Values of g are available in handbooks for a wide range of problems. Fracture is expected when K reaches a critical value K_c , known as the fracture toughness [30]. $K \geq K_c$ might thus be used as a fracture criterion. It should be noted that fracture tends to appear in mode I and K_I is therefore commonly of most concern. This is however often not the case for RCF cracks, see Section 3.1.

SIFs are useful, not only for prediction of brittle failure, but also in the assessment of fatigue crack growth. If the SIF range ΔK , defined as $\Delta K = K_{\max} - K_{\min}$ over a load

cycle, is greater than a threshold value ΔK_{th} , crack growth is expected. Several attempts have been made to formulate crack growth models. One of the first models, which is still the most common, is often referred to as Paris law. It can be expressed as [30, 75]:

$$\frac{da}{dN} = C (\Delta K)^m \quad (6.22)$$

where $\frac{da}{dN}$ is the crack growth per load cycle, i.e. the crack growth rate, while C and m are material parameters. RCF cracks are associated with mixed mode nonproportional loading and crack closure, which are outside the validity of the original Paris law [33]. Nevertheless, this does not prevent that other SIF based propagation criteria are employed, or that the SIFs are used as overall fatigue impact factors.

In complicated general situations, such as a railway wheel traversing a cracked rail surface, it is not possible to find analytical SIF solutions and they must hence be estimated. One attractive way of achieving this is to use FEM or XFEM. Throughout the last decades, significant efforts have been devoted to how this can be done and several methods exist, see e.g. [62]. In the following the focus will be restricted to two methods.

The first, and probably most common approach, available in several commercial FE codes, is based on the J -integral, defined as [78, 90]:

$$J = \int_{\Gamma} \left(W dy - T_i \frac{\partial u_i}{\partial x} d\Gamma \right) \quad (6.23)$$

where W is the strain energy density, u_i the displacements, T_i is the traction vector and Γ is an arbitrary contour that should be taken counterclockwise (i.e. start on the lower crack face and end on the upper) as in Figure 6.5. If the material is linear elastic, J can be related to the SIFs [62]. For implementation purposes, it is often convenient to rewrite the classical J -integral of Equation (6.23) into a domain integral, cf. [83]. If the crack faces are traction free and the material is elastic (not restricted to linear elasticity), it can be shown that J , pertaining to the energy release rate, becomes path independent [62]. As the path independence makes it possible to avoid the singularity close to the crack tip, the mesh can be made rather coarse. Note however that RCF problems are usually associated with crack closure and the assumption of traction free crack faces is thus violated.

The second SIF evaluation approach brought up here is more straight forward than the J -integral method. It starts off from LEFM theory and the asymptotic functions describing the displacement fields close to the crack tip. Presuming plane strain conditions, traction free crack surfaces and neglecting higher order terms, it is possible to express the displacements in the x -, y - and z -direction as [90]:

$$\begin{aligned} u_x &= \frac{(1+\nu)}{4\pi E} \sqrt{2\pi r} \left((5-8\nu) \cos \frac{\theta}{2} - \cos \frac{3\theta}{2} \right) K_{\text{I}} + \frac{(1+\nu)}{4\pi E} \sqrt{2\pi r} \left((9-8\nu) \sin \frac{\theta}{2} + \sin \frac{3\theta}{2} \right) K_{\text{II}} \\ u_y &= \frac{(1+\nu)}{4\pi E} \sqrt{2\pi r} \left((7-8\nu) \sin \frac{\theta}{2} - \sin \frac{3\theta}{2} \right) K_{\text{I}} - \frac{(1+\nu)}{4\pi E} \sqrt{2\pi r} \left((3-8\nu) \cos \frac{\theta}{2} + \cos \frac{3\theta}{2} \right) K_{\text{II}} \\ u_z &= \frac{2(1+\nu)}{\pi E} \sqrt{2\pi r} \sin \frac{\theta}{2} K_{\text{III}} \end{aligned} \quad (6.24)$$

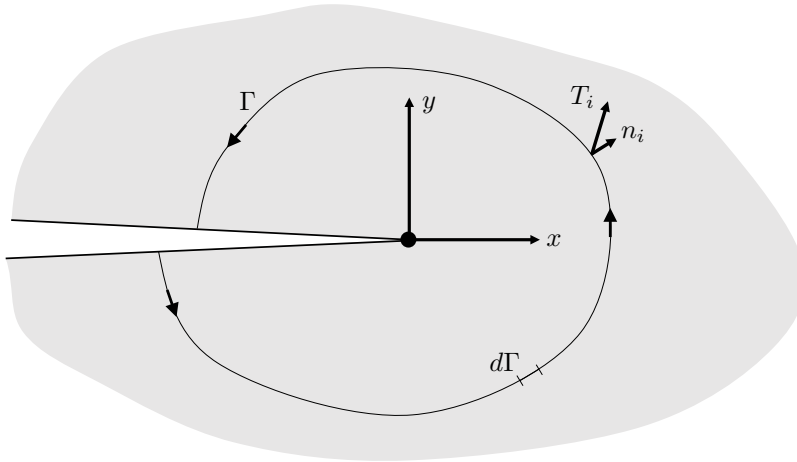


Figure 6.5: *Schematic illustration of a contour Γ around a crack tip. n_i is an outward pointing normal vector and $T_i = \sigma_{ij}n_j$ is the traction vector.*

where E is the Young's modulus, ν Poisson's ratio, u_i the displacement in direction i of the coordinate system illustrated in Figure 6.3. Consequently, elastic displacement fields in the vicinity of a crack tip can be derived from SIFs and vice versa. FEM/XFEM can be used to obtain approximate displacements for a number of discrete points close to the crack tip. Combined with Equation (6.24), an overdetermined system of equations is formed and the SIFs can be obtained via the method of least squares. More involved methods to extract SIFs, via both experiments and FEM, by the method of least squares have been studied in e.g. [51, 52, 80].

To clarify the method described above, a small schematic example is given. Consider a situation of pure mode I loading (i.e. $K_{II} = K_{III} = 0$) with $E=210$ GPa and $\nu=0.3$. For simplicity, the only displacements considered are those on the upper crack face (see Figure 6.3), i.e. where $\theta = \pi$. It is seen from Equation (6.24) that the only non-zero displacement component is u_y :

$$u_y = 8 \frac{1 - \nu^2}{4\pi E} \sqrt{2\pi r} K_I \quad (6.25)$$

Now assume that a FEM/XFEM solution, denoted u_y^h , exist for five discrete points $r_1 \dots r_5$ located on the upper crack face. Applying Equation (6.25) yields five equations where K_I is the only unknown:

$$\underbrace{\begin{bmatrix} u_{y1}^h \\ \vdots \\ u_{y5}^h \end{bmatrix}}_{\mathbf{A}} \approx 8 \frac{1 - \nu^2}{4\pi E} \underbrace{\begin{bmatrix} \sqrt{2\pi r_1} \\ \vdots \\ \sqrt{2\pi r_5} \end{bmatrix}}_{\mathbf{Y}} K_I \quad (6.26)$$

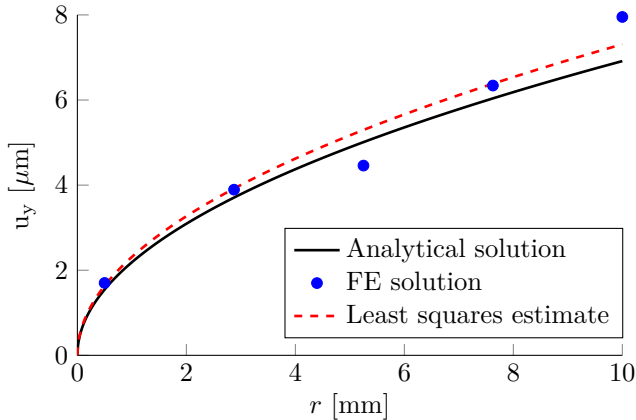


Figure 6.6: Schematic illustration of the asymptotic displacement field close to a crack tip under mode I loading. The solid line represents the exact solution, the circles are sampling points of the FE solution and the dashed line is a least squares fit of the FE results.

The least squares approximation is then obtained as

$$K_I = \left(\mathbf{A}^T \mathbf{A} \right)^{-1} \mathbf{A}^T \mathbf{Y} \quad (6.27)$$

The procedure is illustrated schematically in Figure 6.6. The difference between the exact solution and the least squares fit of the FE solution reflects the error of the estimated K_I compared to the exact. There are of course many delicate questions related to the method, e.g. how crack closure can be handled and sampling points selected. This and several other issues are discussed in **Paper D**, where the approach is applied to RCF cracks (thus treating mixed mode loading and crack closure) and results are compared to J -integral evaluations.

6.3 Thermal loading and damage

The mechanical properties of steel depend to a large extent on the phase² content, which in turn is a function of the temperature history, see e.g. [85]. This is an important topic within rail and wheel manufacturing. The main concern in the context of this thesis is the risk of wheel and rail damage in the form of studs on rails or RCF clusters on wheels due to phase transformations. As touched upon in Section 3.3, the idea is that frictional heating due to sliding in the wheel–rail contact might cause a temperature increase of the pearlite beyond the austenization temperature. When the heat source is removed, the relatively high conductivity of the steel will effectively cool the hot area and cause a rapid temperature decrease, possibly leading to martensite formation and the creation of a white etching layer.

²A material domain that has a microstructure (or composition) distinguishable from other domains of the material [85].

The mechanical properties, such as the yield stress, of the material are typically both phase and temperature dependent and will hence change during the process described above. The residual stresses, that affect damage formation, will thus be related to the mechanical and thermal loading in a highly nonlinear manner. This makes FE analysis a suitable tool for the corresponding stress evaluation, see e.g. Jergéus [46] and Ahlström [1].

The constitutive model needed to describe the phase transformation can be made highly complex to involve different physical phenomena. **Paper E** focusses on rapid temperature changes and some physical phenomena are therefore neglected. As an example, transformation from austenite into martensite is of great importance, while transformation from austenite into e.g. pearlite or bainite is of less interest. This limits the demands on the constitutive model to the inclusion of austenization and martensite transformation. To this end, a simplified approach based on the work by Jergéus [44, 46] is used.

7 Summary of appended papers

The following chapter contains summaries of the appended papers. The inclusion of overall topics are shown in Table 7.1.

Table 7.1: Overall content of the appended papers.

Paper	A	B	C	D	E	F
Dynamic vehicle–track interaction	✓	✓	✓			✓
RCF initiation		✓	✓		✓	
Fracture mechanics				✓		✓
Thermal damage					✓	

7.1 Paper A: The influence of rail surface irregularities on contact forces and local stresses

The initiation of squats has been linked to the occurrence of rail surface irregularities. This paper investigates the influence of rail surface irregularities in the form of dimples, but also some cases of corrugation, on the overall RCF impact, quantified by maximum wheel–rail contact stresses and forces. To this end, an in-house code for dynamic vehicle–track interaction, RAVEN, is utilised. The overall idea, summarised in Figure 7.1 is to compute the displacements of the wheel and track via the computationally efficient method of moving Green’s functions, see Section 5.1. The wheel and rail displacements are then used as input to a contact model based on an implementation of Kalker’s complete theory, see Section 4.3. Output in the form of the resulting normal contact force is used as input to the calculation of wheel and rail displacements of the next time step. Contact stresses

and the resulting contact forces are saved for post-processing.

Altogether, 486 operational scenarios are analysed in a parametric study involving the size of surface dimples, position in relation to sleepers, friction coefficient, velocity and creepage. It is shown that even relatively shallow surface irregularities might cause a substantial RCF impact. How the different RCF impact factors are affected by the different input parameters is presented and discussed. As an example, the influence on RCF impact of dimple semi-length a and depth h are investigated, see Figure 7.2. In addition, three different types of corrugation are studied in combination with different magnitudes of velocity, friction coefficient and creepage. It is found that the evaluated RCF impact from rail surface dimples is larger than for the studied cases of corrugation.

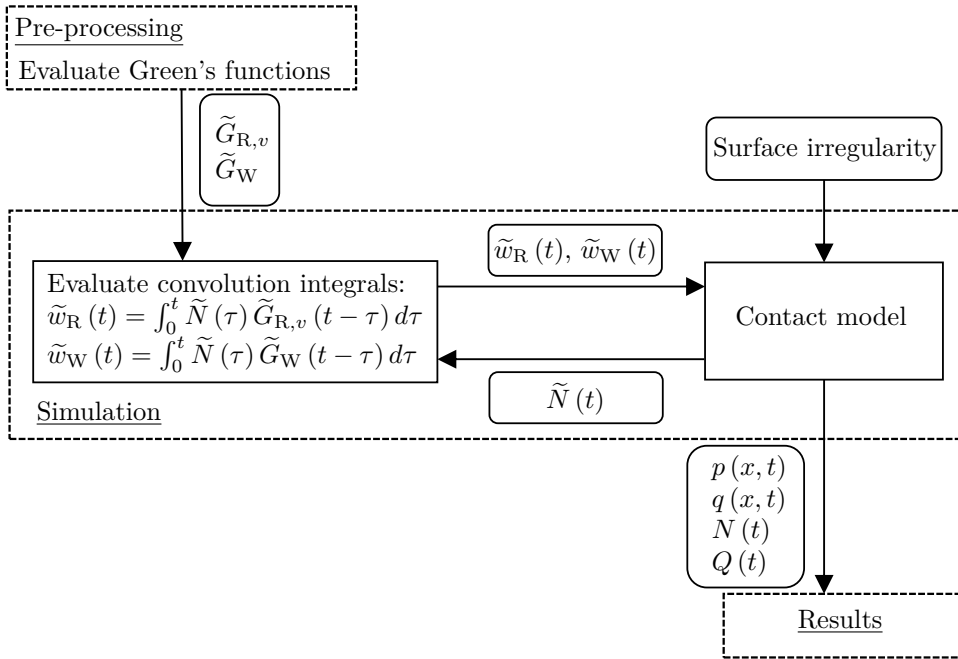


Figure 7.1: Flow chart describing the in-house vehicle-track interaction code RAVEN. $\tilde{G}_{R,v}$ and \tilde{G}_W are Green's functions for rail and wheel, $\tilde{w}_R(t)$ and $\tilde{w}_W(t)$ are displacements of rail and wheel at the current position, and $\tilde{N}(t)$ is the resulting normal contact force in the current position. The output consists of time histories of contact pressure $p(x,t)$ and shear traction $q(x,t)$ and corresponding normal and shear forces $N(t)$ and $Q(t)$ respectively. Surface irregularities, e.g. dimples and corrugation patterns, are introduced as input to the contact model.

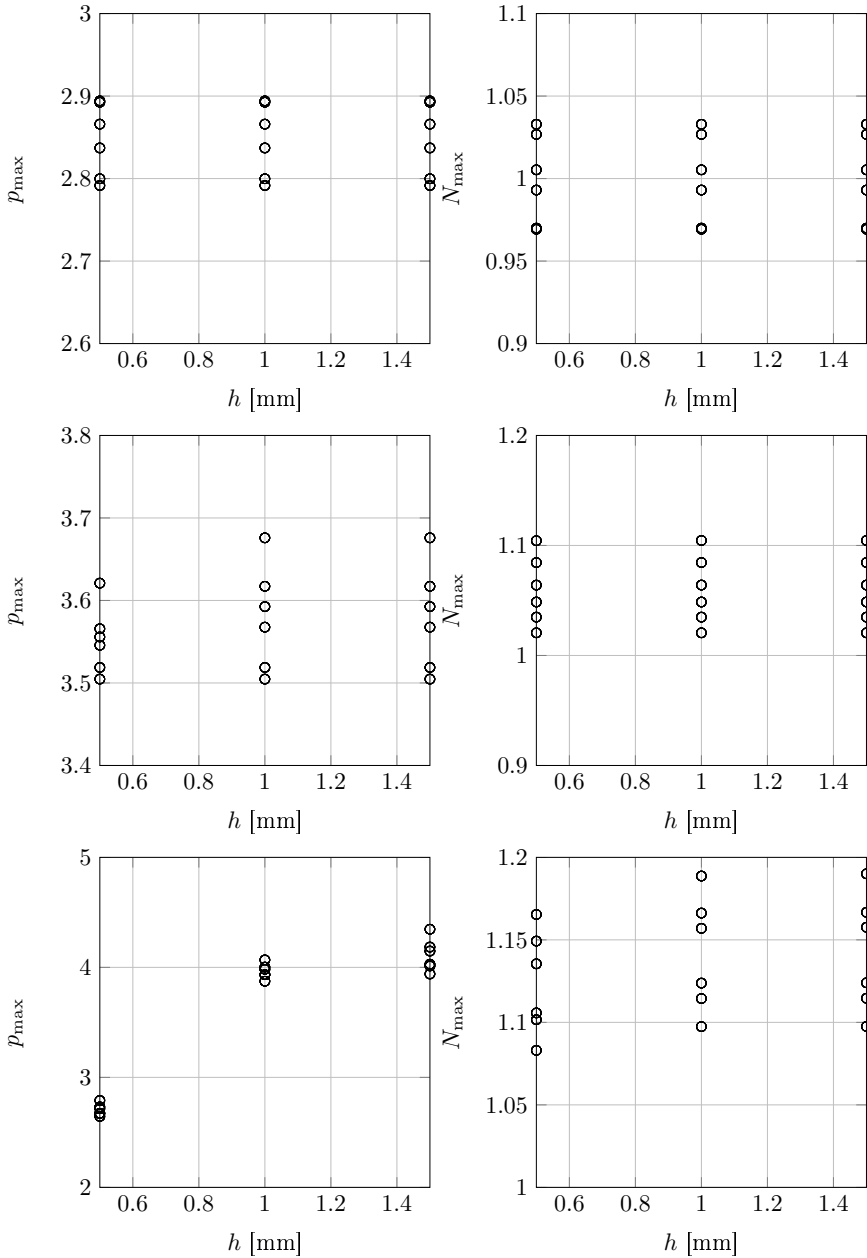


Figure 7.2: Two RCF impact measures, maximum contact pressure p_{\max} and the maximum contact force N_{\max} , plotted as functions of dimple depth h . Three different dimple semi-lengths are considered. From above: $a=4, 6$ and 8 mm. Adopted from **Paper A**.

7.2 Paper B: An efficient approach to the analysis of rail surface irregularities accounting for dynamic train–track interaction and inelastic deformations

This paper continues the work in **Paper A** by investigating how the computed contact stress histories can be used in combination with finite element analysis in order to evaluate the RCF impact. The FE analyses involve nonlinear mixed kinematic and isotropic hardening models fitted against cyclic experimental data of a rail material. The RCF impact is quantified using two different RCF impact measures: the accumulated von Mises strain, see Section 6.2.4, and the Jiang–Sehitoglu fatigue parameter, see Section 6.2.5. These measures reflect the two overall RCF initiation mechanisms: ratchetting and low cycle fatigue, respectively. The procedure is finally demonstrated through numerical examples involving rail surface irregularities. The method presented in this paper constitutes the foundation of the analyses in **Paper C**.

7.3 Paper C: Integrated analysis of dynamic vehicle–track interaction and plasticity induced damage in the presence of squat defects

A number of parameter combinations identified in **Paper A** are analysed in accordance with the procedure proposed in **Paper B**. This includes a small, medium sized and large dimple, lower and higher friction coefficient as well as a double dimple representing a cluster of defects. The effect of a limited region of low friction, resulting in a state of full slip and subsequent recovery, is also examined. It is shown that a higher friction coefficient (and thereby higher wheel–rail shear tractions) has a significant impact on the RCF impact, see Figure 7.3. It is furthermore evident that the RCF impact increases with dimple size and that defect clusters have a detrimental effect. It is finally shown that the impact of a low friction section on a rail is comparable with that of a small surface dimple.

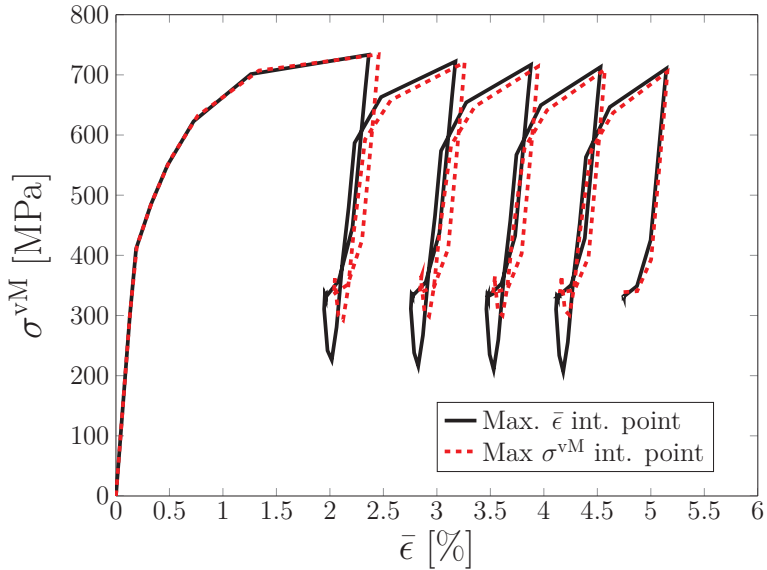


Figure 7.3: *Stress–strain response in the vicinity of a dimple for a case with relatively high friction loading. Results from two integration points are considered: the point subjected to the highest von Mises stress of all load cycles and the point with the largest accumulated von Mises strain after the final load cycle. Adopted from **Paper C**.*

7.4 Paper D: Evaluation of stress intensity factors under multiaxial and compressive conditions using low order displacement or stress field fitting

There are several complicating factors related to the evaluation of Stress Intensity Factors (SIFs) of RCF cracks. First of all, the loading is typically multiaxial, which means that mode I–III SIFs are all activated. Secondly, the loading is compressive during part of a wheel passage and special care must be taken to avoid non-physical negative mode I SIFs. In this study it is proposed that SIFs can be accurately evaluated for RCF cracks by least squares fitting of a low order displacement ansatz towards an XFEM solution, see Sections 6.1 and 6.2.6. The accuracy is demonstrated by comparing the proposed approach to handbook solutions and J -integral results for pure mode I, II and III loading. Special emphasis is put on compressive conditions, see Figure 7.4. The ability to separate between mode I and II SIFs as well as convergence with mesh size are also demonstrated. The procedure is finally applied to evaluate SIFs for an inclined half-circular RCF crack embedded in a 3D rail head. The procedure presented in this paper constitutes the foundation of the analyses in **Paper F**.

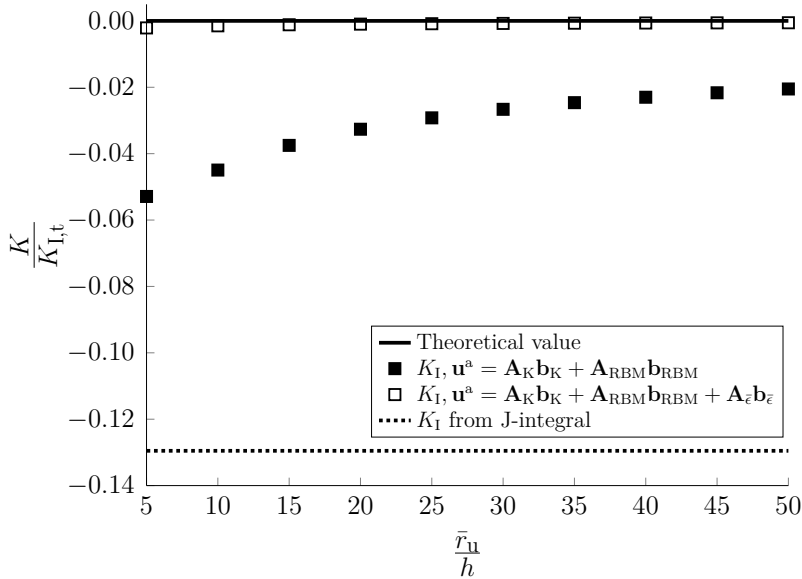


Figure 7.4: A case of compressive mode I loading is considered. The vertical axis shows a normalised mode I SIF. The horizontal axis shows a cut-off distance \bar{r}_u , used in the SIF evaluation scheme, expressed in terms of element lengths. Two different ansatzes are considered. The three-term ansatz can account also for constant strain fields and more accurate results are thus obtained. Theoretical values and J-integral results are shown for reference. Adopted from **Paper D**.

7.5 Paper E: Numerical investigation of crack initiation in rails and wheels in the presence of white etching layers

Stud defects on rail crowns and RCF clusters on wheel treads have been linked to the presence of white etching layers (WELs). This study utilises three-dimensional FE analyses to study how the axle load (12.5/25 tonnes), traction coefficient (0/0.4) and WEL thickness (500 μm on rails and 1000 μm on wheels) affect the propensity for crack initiation. By prescribing a temperature history it is possible to model the resulting material phase transformation and capture the phase dependent elasto-plastic material response.

The phase transformation algorithm consists of the following steps:

- Pearlite, austenite and martensite are represented by their respective volume fractions: v_p , v_a , v_m , which can vary between 0 and 1 and should add up to 1 in each integration point. When the analysis starts, it is presumed that $v_p = 1$.
- If the temperature reaches the austenization temperature T_{aust} during heating, the

volume fraction of austenite instantly becomes 1, i.e. $v_a = 1$.

- The volume fraction martensite obtained during a temperature decrease is modelled using an empirical relation known as the Koistinen–Marburger equation [60]:

$$v_m = 1 - \exp(-0.011 (T_{Ms} - T)), \quad T \leq T_{Ms} \quad (7.1)$$

where T is the current temperature and T_{Ms} is a material parameter.

To illustrate the behaviour of the model $T_{aust}=850$ °C and $T_{Ms}=270$ °C are adopted. The resulting phase transformations are presented in Figure 7.5 (a) for heating and (b) cooling respectively.

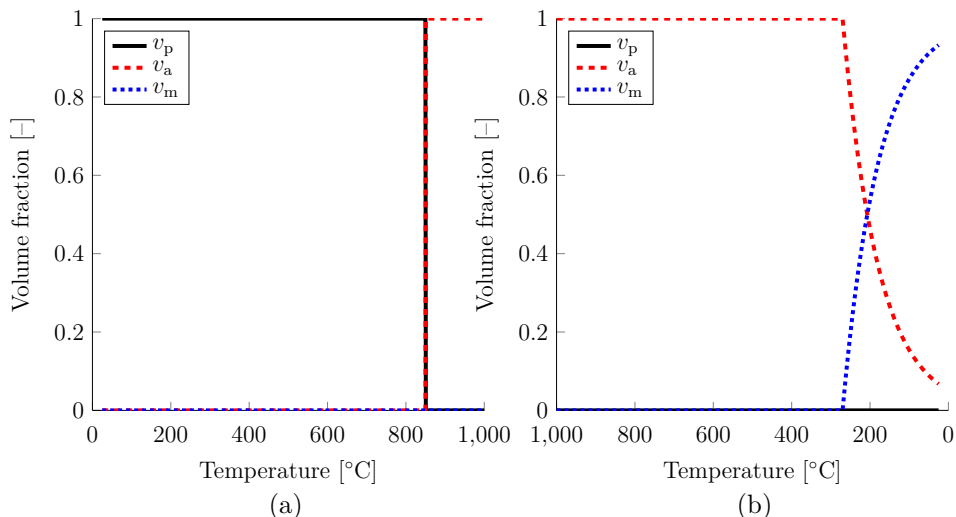


Figure 7.5: *Illustration of phase transformation during a temperature increase (a) and temperature decrease (b). v_p , v_a and v_m denote the volume fraction of pearlite, austenite and martensite respectively. Adopted from **Paper E**.*

Subsequent mechanical load cases are included by simulations featuring moving Hertzian loads. Fatigue damage is quantified by the Jiang–Sehitoglu fatigue parameter, see Section 6.2.5 and Figure 7.6. The study shows that the axle load has a minor impact on wheel and rail damage. An increased frictional load does on the other hand increase the damage significantly. Although the WEL is modelled to be twice as thick on wheels as compared to rails, similar trends are observed. Finally, the effect of eccentric wheel passages are considered. It is seen that the predicted fatigue damage becomes more than twice as high compared to a centric wheel passage for the cases studied. This might help explaining why stud cracks often grow from the gauge corner side of the rail surface.

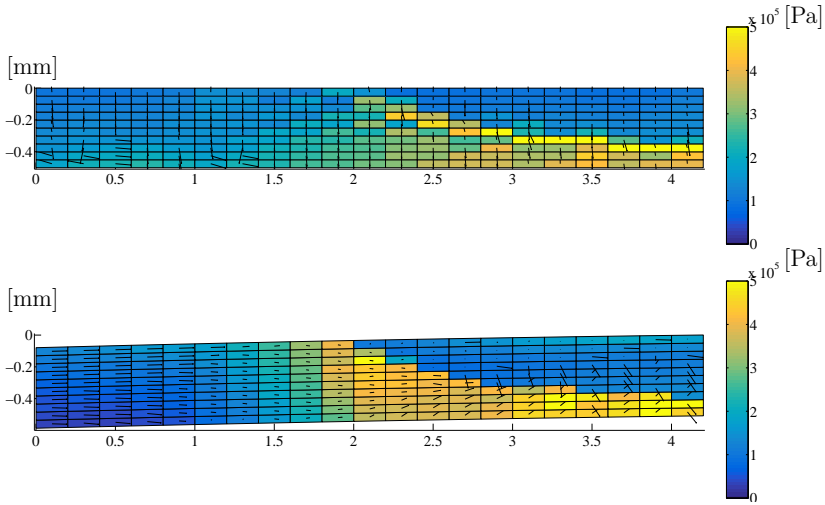


Figure 7.6: *Distribution of the Jiang–Sehitoglu fatigue parameter FP in the vicinity of a WEL in a rail. Axle load 25 tonnes and no frictional load. The upper and lower plots refer to evaluation planes along and perpendicular to the rolling direction, respectively. Identified critical material planes are indicated by projected normal vectors. The FE discretization is indicated by black lines. Adopted from **Paper E**.*

7.6 Paper F: Assessment of rolling contact fatigue cracks in the vicinity of rail surface irregularities

As been discussed in **Papers A–C**, rail surface irregularities in the form of dimples generate contact pressure peaks as well as dynamic impact loads. These effects might in turn promote further deterioration of the rail material. This study considers cracks in the vicinity of surface dimples and assess their severity with the help of equivalent SIF ranges. To this end, three-dimensional dimples as well as realistic wheel and rail profiles are considered in a dynamic vehicle–track interaction analysis similar to those conducted in **Papers A–C**. The resulting contact pressures are then mapped onto a three-dimensional FE model representing a rail section with embedded XFEM cracks. SIFs are evaluated in a post-processing step with help of the method described in **Paper D**. An equivalent SIF range is employed to characterise the multiaxial loading of the crack. It is shown that the equivalent SIF range for a crack increases $\sim 10\text{--}20\%$ in the presence of a small dimple and $\sim 30\text{--}40\%$ for a large dimple as compared to the situation of a smooth rail surface, see Figure 7.7. It is found that also dimples as shallow as 0.125 mm might increase the equivalent SIF range by $\sim 15\%$. Double dimples are furthermore examined to investigate possible cluster effects. It is seen that the trends are similar to those obtained for single dimples. Equivalent SIF ranges are, in other words, expected to increase with dimple size.

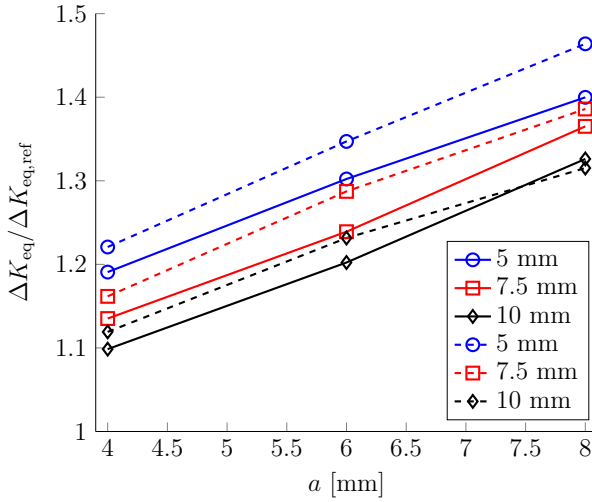


Figure 7.7: Influence of dimple semi-length a on equivalent SIF range ΔK_{eq} . Single and double dimple results are indicated by solid and dashed lines, respectively. Crack lengths are indicated in the legend. Note that the equivalent SIF ranges are scaled with values corresponding to a smooth rail surface, $\Delta K_{\text{eq,ref}}$. Adopted from **Paper F**.

8 Conclusions, main results and future work

This chapter summarises the main results of the appended papers and provide suggestions for future work. In general, the current work has contributed in the development of numerical procedures to facilitate the analysis of squat/RCF cluster initiation and subsequent crack analysis. This involves:

- Methods for mapping contact stresses from dynamic simulations onto finite element models able to capture nonlinear hardening plasticity (**Paper B**).
- Subsequent RCF assessments based on accumulated strain and the Jiang–Sehitoglu fatigue parameter (**Paper B**).
- A procedure for evaluation of stress intensity factors, with particular focus on RCF cracks, using results from XFEM/FEM analyses in commercial software (**Paper D**). Special emphasis is given to crack closure.
- A simulation procedure for mapping contact stresses from dynamic simulations onto an XFEM model featuring a cracked rail head, allowing for subsequent SIF evaluation (**Paper F**).

There are also several practical implications of the work in the appended papers. These involve:

- Surface irregularities in the form of dimples can cause a substantial impact on the propensity for RCF initiation (**Paper A**).

- The RCF impact due to surface dimples might be larger than the impact caused by even severe corrugation (**Paper A**).
- The depth of a dimple becomes increasingly important for longer dimples both with regard to RCF impact and the loading of macroscopic cracks (**Paper A** and **Paper F**).
- Compared to a smooth rail surface, also shallow dimples might have a substantial impact on RCF and crack loading of macroscopic cracks (**Paper A** and **Paper F**).
- The RCF impact increases with friction also when dimples are present (**Paper C**).
- The RCF impact typically increases with the size (i.e. length in combination with depth) of the dimple (**Paper C**). The same conclusion can be drawn also for the severity of macroscopic cracks in the vicinity of a dimple (**Paper F**).
- Clusters of dimples can potentially lead to a significant increase of the RCF impact (**Paper C**).
- Recovery from full slip, caused by a temporary decrease of the friction coefficient, might give rise to RCF impact in the same order as a small surface dimple (**Paper C**).
- The axle load has no major influence on crack initiation in the vicinity of a WEL as long as the friction is negligible (**Paper E**).
- The traction coefficient has a significant influence on the crack initiation in the vicinity of a WEL (**Paper E**).
- The overall trends regarding fatigue impact in the vicinity of a WEL are the same regardless if a 500 μm or 1000 μm thick WEL is considered (**Paper E**).
- Offset (with respect to the WEL position) wheel passages result in high fatigue impact and might explain crack initiation at the gauge corner side of the rail surface (**Paper E**).
- The severity of cracks in the vicinity of double dimples (representing a cluster) increases with the size (i.e. length in combination with depth) of the dimple (**Paper F**). This is similar to the trends obtained for single dimples.

These results imply that surface irregularities should be kept under surveillance. When dimples reach a threshold size, they are detrimental and the rail (or wheel) should be re-profiled. It can here be noted that dimples may deform plastically or wear off. Consequently, a key issue is whether dimples will be removed by such mechanisms or by grinding before substantial cracking has occurred. It seems important to keep surface shear tractions under control, both regarding crack initiation in the presence of dimples and WELs. This can be achieved either by changing how trains are operated or by modifying friction conditions.

It should be noted that in the presence of squats/RCF clusters, several different underlying factors may be responsible for the observable wheel–rail damage. To capture

the effects from all of these, several different physical phenomena need to be accounted for. These include dynamic vehicle–track interaction, wheel–rail contact, nonlinear (and possibly anisotropic) material response, progressive deterioration of the material, crack initiation, crack propagation etc. Preferably, all these effects should be accounted for in an integrated manner and studied for a variety of operational scenarios. Each scenario should in turn consist of a large number of train passages. However, such an approach would hardly work in practice due to high computational costs etc. A less expensive procedure is to study some of the effects individually, or in combination, as done in this thesis. An alternative, and suggestion for future work, would be to study some of the effects more in detail using e.g. finite element analysis and then produce meta-models of the response. If, for example, the Jiang–Sehitoglu fatigue parameter is evaluated for a large number of surface irregularities, friction conditions etc., it would be possible to fit a meta-model to the corresponding results. Unlike the underlying finite element simulations, the meta-model would be fast and could be incorporated directly into a vehicle–track interaction simulation, allowing for extensive parametric studies. This type of analyses would be suitable to identify potentially problematic areas along a railway line.

In contrast to the broad but not so detailed analyses mentioned above, it could also be fruitful to combine, e.g., XFEM with a suitable crack propagation criterion in order to simulate realistic crack growth given certain operational scenarios. This would however require a criterion able to capture the specific characteristics of RCF cracks.

Another topic that is possible to improve as future work is the analyses related to material phase transformation in relation to stud/RCF cluster simulations. More involved phase transformation models, including evolution laws for state variables etc., could be introduced. Further, it could be important to also model the source of the frictional heating. In that way it would be possible to start off with an operational scenario, predict the temperature increase in the wheel and the rail, obtain a size of the white etching layers and finally make a crack initiation assessment. It would, in other words, be possible to directly relate the risk of crack initiation to how the trains are operated.

Last, but not least, it is always necessary to relate the conclusions drawn in numerical simulations to experimental results and field observations. In that way it could become more clear which RCF impact measures that are the most relevant. This is, however, often easier said than done as it is complicated and time consuming to perform full-scale fatigue experiments. Further, the test conditions are typically not possible to control in detail. Therefore it is probably more fruitful to rely on test rig and laboratory tests in order to achieve well controlled experiments.

References

- [1] J. Ahlström. Residual stresses generated by repeated local heating events — Modelling of possible mechanisms for crack initiation. *Wear* **366**.SI (2016), 180–187. ISSN: 0043-1648. DOI: 10.1016/j.wear.2016.05.029.
- [2] J. Ahlström and B. Karlsson. Microstructural evaluation and interpretation of the mechanically and thermally affected zone under railway wheel flats. *Wear* **232**.1 (1999), 1–14. ISSN: 0043-1648. DOI: 10.1016/S0043-1648(99)00166-0.
- [3] R. Andersson. Assessment of rolling contact fatigue cracks in the vicinity of rail surface irregularities. *Submitted for international publication* (2018).
- [4] R. Andersson et al. Numerical investigation of crack initiation in rails and wheels in the presence of white etching layers. *Submitted for international publication* (2018).
- [5] R. Andersson, F. Larsson, and E. Kabo. Evaluation of stress intensity factors under multiaxial and compressive conditions using low order displacement or stress field fitting. *Engineering Fracture Mechanics* **189** (2018), 204–220. DOI: 10.1016/j.engfracmech.2017.11.015.
- [6] R. Andersson et al. An efficient approach to the analysis of rail surface irregularities accounting for dynamic train–track interaction and inelastic deformations. *Vehicle System Dynamics* **53**.11 (2015), 1667–1685. DOI: 10.1080/00423114.2015.1081701.
- [7] R. Andersson et al. The influence of rail surface irregularities on contact forces and local stresses. *Vehicle System Dynamics* **53**.1 (2015), 68–87. DOI: 10.1080/00423114.2014.982890.
- [8] R. Andersson et al. Integrated analysis of dynamic vehicle–track interaction and plasticity induced damage in the presence of squat defects. *Wear* **366–367** (2016), 139–145. DOI: 10.1016/j.wear.2016.05.014.
- [9] R. Andersson. “Surface defects in rails — Potential influence of operational parameters on squat initiation”. Licentiate thesis. Gothenburg, Sweden: Chalmers University of Technology, 2015. 79 pp.
- [10] T. Belytschko and T. Black. Elastic crack growth in finite elements with minimal remeshing. *International Journal for Numerical Methods in Engineering* **45**.5 (1999), 601–620. ISSN: 0029-5981.
- [11] C. Bernsteiner et al. Development of white etching layers on rails: simulations and experiments. *Wear* **366** (2016), 116–122. ISSN: 0043-1648. DOI: 10.1016/j.wear.2016.03.028.
- [12] S. Bogdański and M. W. Brown. Modelling the three-dimensional behaviour of shallow rolling contact fatigue cracks in rails. *Wear* **253**.1-2 (2002), 17–25. DOI: 10.1016/S0043-1648(02)00078-9.
- [13] S. Bogdański, M. Olzak, and J. Stupnicki. Numerical modelling of a 3D rail RCF ‘squat’-type crack under operating load. *Fatigue and Fracture of Engineering Materials and Structures* **21**.8 (1998), 923–935.
- [14] S. Bogdański, M. Olzak, and J. Stupnicki. Numerical stress analysis of rail rolling contact fatigue cracks. *Wear* **191**.1-2 (1996), 14–24.

- [15] P. E. Bold, M. W. Brown, and R. J. Allen. Shear mode crack growth and rolling contact fatigue. *Wear* **144.1** (1991), 307–317. ISSN: 0043-1648. DOI: 10.1016/0043-1648(91)90022-M.
- [16] A. F. Bower. Influence of crack face friction and trapped fluid on surface initiated rolling contact fatigue cracks. *Journal of Tribology* **110.4** (1988), 704–711. ISSN: 0742-4787. DOI: 10.1115/1.3261717.
- [17] F. Brunel et al. Prediction of the initial residual stresses in railway wheels induced by manufacturing. *Journal of Thermal Stresses* **36.1** (2013), 37–55. ISSN: 0149-5739. DOI: 10.1080/01495739.2012.720542.
- [18] M. C. Burstow. *Whole Life Rail Model Application and Development for RSSB – Continued development of an RCF damage parameter*. Report. Rail Safety and Standards Board, 2004.
- [19] D. F. Cannon et al. Rail defects: An overview. *Fatigue and Fracture of Engineering Materials and Structures* **26.10** (2003), 865–886. DOI: 10.1046/j.1460-2695.2003.00693.x.
- [20] R. I. Carroll and J. H. Beynon. Rolling contact fatigue of white etching layer: Part 1. Crack morphology. *Wear* **262.9-10** (2007), 1253–1266. ISSN: 0043-1648. DOI: 10.1016/j.wear.2007.01.003.
- [21] R. I. Carroll and J. H. Beynon. Rolling contact fatigue of white etching layer. Part 2. Numerical results. *Wear* **262.9-10** (2007), 1267–1273. ISSN: 0043-1648. DOI: 10.1016/j.wear.2007.01.002.
- [22] F. W. Carter. On the Action of a Locomotive Driving Wheel. *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character* **112.760** (1926), 151–157. ISSN: 0950-1207.
- [23] P. Clayton. Tribological aspects of wheel-rail contact: a review of recent experimental research. *Wear* **191.1** (1996), 170–183. ISSN: 0043-1648. DOI: 10.1016/0043-1648(95)06651-9.
- [24] P. Clayton and M. B. P. Allery. Metallurgical aspects of surface damage problems in rails. *Canadian Metallurgical Quarterly* **21.1** (1982), 31–46.
- [25] R. D. Cook et al. *Concepts and applications of finite element analysis*. 4th ed. New York: Wiley, 2002. 719 pp. ISBN: 0471356050.
- [26] R. R. Craig and A. J. Kurdila. *Fundamentals of Structural Dynamics*. 2nd ed. Hoboken, New Jersey: John Wiley & Sons, Inc., 2006. 728 pp. ISBN: 978-0-471-43044-5.
- [27] W. J. T. Daniel, S. Pal, and M. Farjoo. Rail squats: Progress in understanding the Australian experience. *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit* **227.5** (2013), 481–492. DOI: 10.1177/0954409713500950.
- [28] X. Deng et al. “Residual fatigue life evaluation of rail at squats seeds using 3D explicit finite element analysis”. *International Conference on Ageing of Materials & Structures*. Proceedings of the International Conference on Ageing of Materials & Structures. 2014.
- [29] R. Deuce. *Wheel tread damage - an elementary guide*. Technical report. 2007.

- [30] N. E. Dowling. *Mechanical behavior of materials: engineering methods for deformation, fracture and fatigue*. 3d ed. Upper Saddle River, N.J: Pearson Prentice Hall, 2007. 912 pp. ISBN: 0132256096.
- [31] A. Draganis. “Numerical simulation of thermomechanically coupled transient rolling contact - An arbitrary Lagrangian-Eulerian approach”. PhD thesis. Gothenburg, Sweden: Chalmers University of Technology, 2014. 116 pp.
- [32] A. Ekberg, E. Kabo, and H. Andersson. An engineering model for prediction of rolling contact fatigue of railway wheels. *Fatigue and Fracture of Engineering Materials and Structures* **25**.10 (2002), 899–909. DOI: 10.1046/j.1460-2695.2002.00535.x.
- [33] A. Ekberg and E. Kabo. Fatigue of railway wheels and rails under rolling contact and thermal loading—an overview. *Wear* **258**.7 (2005), 1288–1300. ISSN: 0043-1648. DOI: 10.1016/j.wear.2004.03.039.
- [34] M. Farjoo, W. Daniel, and P. A. Meehan. Modelling a squat form crack on a rail laid on an elastic foundation. *Engineering Fracture Mechanics* **85** (2012), 47–58. DOI: 10.1016/j.engfracmech.2012.02.004.
- [35] M. Farjoo et al. Stress intensity factors around a 3D squat form crack and prediction of crack growth direction considering water entrapment and elastic foundation. *Engineering Fracture Mechanics* **94** (2012), 37–55. DOI: 10.1016/j.engfracmech.2012.07.011.
- [36] D. I. Fletcher. Numerical simulation of near surface rail cracks subject to thermal contact stress. *Wear* **314**.1-2 (2014), 96–103. DOI: 10.1016/j.wear.2013.11.021.
- [37] D. I. Fletcher and S. H. Sanusi. The potential for suppressing rail defect growth through tailoring rail thermo-mechanical properties. *Wear* **366** (2016), 401–406. ISSN: 0043-1648. DOI: 10.1016/j.wear.2016.06.022.
- [38] S. L. Grassie. Squats and squat-type defects in rails: The understanding to date. *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit* **226**.3 (2012), 235–242. DOI: 10.1177/0954409711422189.
- [39] S. L. Grassie. Studs and squats: The evolving story. *Wear* **366** (2016), 194–199. ISSN: 0043-1648. DOI: 10.1016/j.wear.2016.03.021.
- [40] S. L. Grassie et al. Studs: A squat-type defect in rails. *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit* **226**.3 (2012), 243–256. DOI: 10.1177/0954409711421462.
- [41] H. Hertz. Über die Berührung fester elastischer Körper. *Journal für die reine und angewandte Mathematik* **92** (1882), 156–171.
- [42] M. Ishida. Rolling contact fatigue (RCF) defects of rails in Japanese railways and its mitigation strategies. *Electronic Journal of Structural Engineering* **13**.1 (2013), 67–74.
- [43] S. Iwnicki, ed. *Handbook of railway vehicle dynamics*. 1st ed. Boca Raton: CRC/Taylor & Francis, 2006. 526 pp. ISBN: 0849333210.
- [44] J. Jergéus. Martensite formation and residual stresses around railway wheel flats. *Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science* **212**.1 (1998), 69–79. ISSN: 0954-4062. DOI: 10.1243/0954406981521051.

- [45] J. Jergéus et al. Full-scale railway wheel flat experiments. *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit* **213.1** (1999), 1–13. ISSN: 0954-4097. DOI: 10.1243/0954409991530985.
- [46] J. Jergéus. “Railway Wheel Flats. Martensite Formation, Residual Stresses, and Crack Propagation”. PhD Thesis. Gothenburg, Sweden: Chalmers University of Technology, 1998. 115 pp.
- [47] Y. Jiang and H. Sehitoglu. A model for rolling contact failure. *Wear* **224.1** (1999), 38–49.
- [48] K. L. Johnson. *Contact Mechanics*. New York: Cambridge University Press, 1987. 452 pp. ISBN: 0521347963.
- [49] K. L. Johnson. The Strength of Surfaces in Rolling Contact. *Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science* **203.3** (1989), 151–163. ISSN: 0954-4062. DOI: 10.1243/PIME_PROC_1989_203_100_02.
- [50] A. Jörg et al. “The Squat condition of rail materials - A novel approach to Squat prevention”. *CM2015 – 10th International Conference on Contact Mechanics of Wheel / Rail Systems*.
- [51] S. H. Ju. Simulating stress intensity factors for anisotropic materials by the least-squares method. *International Journal of Fracture* **81.3** (1996), 283–297.
- [52] S. H. Ju. Simulating three-dimensional stress intensity factors by the least-squares method. *International Journal for Numerical Methods in Engineering* **43.8** (1998), 1437–1451.
- [53] E. Kabo and A. Ekberg. Material defects in rolling contact fatigue of railway wheels - The influence of defect size. *Wear* **258.7-8** (2005), 1194–1200. DOI: 10.1016/j.wear.2004.03.070.
- [54] S. Kaewunruen, M. Ishida, and S. Marich. Dynamic Wheel–Rail Interaction Over Rail Squat Defects. *Acoustics Australia* **43.1** (2015), 97–107. ISSN: 0814-6039. DOI: 10.1007/s40857-014-0001-4.
- [55] J. J. Kalker. A Fast Algorithm for the Simplified Theory of Rolling Contact. *Vehicle System Dynamics* **11.1** (1982), 1–13. ISSN: 0042-3114. DOI: 10.1080/00423118208968684.
- [56] J. J. Kalker. The computation of three-dimensional rolling contact with dry friction. *International Journal for Numerical Methods in Engineering* **14.9** (1979), 1293–1307. ISSN: 0029-5981. DOI: 10.1002/nme.1620140904.
- [57] J. J. Kalker. *Three-dimensional elastic bodies in rolling contact*. Vol. 2. Dordrecht: Kluwer, 1990. 314 pp. ISBN: 9780792307129.
- [58] J. J. Kalker. Wheel-rail rolling contact theory. *Wear* **144.1** (1991), 243–261. ISSN: 0043-1648. DOI: 10.1016/0043-1648(91)90018-P.
- [59] M. Kaneta et al. A possible mechanism for rail dark spot defects. *Journal of Tribology* **120.2** (1998), 304–309.
- [60] D. P. Koistinen and R. E. Marburger. A general equation prescribing the extent of the austenite-martensite transformation in pure iron-carbon alloys and plain carbon steels. *Acta Metallurgica* **7.1** (1959), 59–60. ISSN: 0001-6160. DOI: 10.1016/0001-6160(59)90170-1.

- [61] K. Kondo, K. Yoroizaka, and Y. Sato. Cause, increase, diagnosis, countermeasures and elimination of Shinkansen shelling. *Wear* **191.1** (1996), 199–203. ISSN: 0043-1648. DOI: 10.1016/0043-1648(95)06727-2.
- [62] M. Kuna. *Finite Elements in Fracture Mechanics: Theory - Numerics - Applications*. Vol. 201. Dordrecht: Springer Netherlands, 2013. 447 pp. ISBN: 9789400766792.
- [63] R. Lewis and U. Olofsson, eds. *Wheel–rail interface handbook*. Cambridge: Woodhead Publishing, 2009. 842 pp. ISBN: 9781439801468.
- [64] Z. Li et al. Squat growth—Some observations and the validation of numerical predictions. *Wear* **271.1-2** (2011), 148–157. DOI: 10.1016/j.wear.2010.10.051.
- [65] Z. Li et al. Differential wear and plastic deformation as causes of squat at track local stiffness change combined with other track short defects. *Vehicle System Dynamics* **46.S1** (2008), 237–246. DOI: 10.1080/00423110801935855.
- [66] Z. Li et al. An investigation into the causes of squats—Correlation analysis and numerical modeling. *Wear* **265.9-10** (2008), 1349–1355. DOI: 10.1016/j.wear.2008.02.037.
- [67] M. Molodova et al. Validation of a finite element model for axle box acceleration at squats in the high frequency range. *Computers and Structures* **141** (2014), 84–93. DOI: 10.1016/j.compstruc.2014.05.005.
- [68] M. Molodova et al. Automatic detection of squats in railway infrastructure. *IEEE Transactions on Intelligent Transportation Systems* **15.5** (2014), 1980–1990. DOI: 10.1109/TITS.2014.2307955.
- [69] M. Molodova et al. Parametric study of axle box acceleration at squats. *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit* **229.8** (2015), 841–851. ISSN: 0954-4097. DOI: 10.1177/0954409714523583.
- [70] N. Moës, J. Dolbow, and T. Belytschko. A finite element method for crack growth without remeshing. *International Journal for Numerical Methods in Engineering* **46.1** (1999), 131–150.
- [71] M. Naeimi et al. Thermomechanical analysis of the wheel-rail contact using a coupled modelling procedure. *Tribology International* **117** (2018), 250–260. DOI: 10.1016/j.triboint.2017.09.010.
- [72] M. Naeimi, Z. Li, and R. Dollevoet. Nucleation of squat cracks in rail, calculation of crack initiation angles in three dimensions. *Journal of Physics: Conference Series* **628.1** (2015). ISSN: 1742-6588. DOI: 10.1088/1742-6596/628/1/012043.
- [73] S. Pal et al. Metallurgical and physical understanding of rail squat initiation and propagation. *Wear* **284-285** (2012), 30–42. DOI: 10.1016/j.wear.2012.02.013.
- [74] S. Pal, W. J. T. Daniel, and M. Farjoo. Early stages of rail squat formation and the role of a white etching layer. *International Journal of Fatigue* **52** (2013), 144–156. ISSN: 0142-1123. DOI: 10.1016/j.ijfatigue.2013.02.016.
- [75] P. C. Paris, M. P. Gomez, and W. E. Anderson. A rational analytic theory of fatigue. *The trend in engineering* **13.1** (1961), 9–14.
- [76] A. Pieringer. “Time-domain modelling of high-frequency wheel/rail interaction”. PhD thesis. Gothenburg, Sweden: Chalmers University of Technology, 2011. 107 pp.
- [77] *Rail defects*. 4th ed. UIC Code 712. Paris: International Union of Railways, 2002. 107 pp. ISBN: 2-7461-0341-9.

- [78] J. R. Rice. A path independent integral and approximate analysis of strain concentration by notches and cracks. *Journal of Applied Mechanics* **35.2** (1968), 379–386.
- [79] J. W. Ringsberg. Life prediction of rolling contact fatigue crack initiation. *International Journal of Fatigue* **23.7** (2001), 575–586. DOI: 10.1016/S0142-1123(01)00024-X.
- [80] R. J. Sanford and J. W. Dally. A general method for determining mixed-mode stress intensity factors from isochromatic fringe patterns. *Engineering Fracture Mechanics* **11.4** (1979), 621–633. ISSN: 0013-7944.
- [81] D. Scott, D. I. Fletcher, and B. J. Cardwell. Simulation study of thermally initiated rail defects. *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit* **228.2** (2014), 113–127. ISSN: 0954-4097. DOI: 10.1177/0954409712465697.
- [82] M. Shahzamanian Sichani. “On Efficient Modelling of Wheel-Rail Contact in Vehicle Dynamics Simulation”. PhD thesis. Stockholm, Sweden: KTH Royal Institute of Technology, 2016. 97 pp.
- [83] C. F. Shih, B. Moran, and T. Nakamura. Energy release rate along a three-dimensional crack front in a thermally stressed body. *International Journal of Fracture* **30.2** (1986), 79–102. ISSN: 0376-9429.
- [84] S. Simon et al. Tribological characterization of rail squat defects. *Wear* **297.1-2** (2013), 926–942. DOI: 10.1016/j.wear.2012.11.011.
- [85] W. F. Smith and J. Hashemi. *Foundations of materials science and engineering*. 4th ed. Boston: McGraw-Hill Higher Education, 2006. 1032 pp. ISBN: 0071256903.
- [86] M. Steenbergen and R. Dollevoet. On the mechanism of squat formation on train rails - Part I: Origination. *International Journal of Fatigue* **47** (2013), 361–372. DOI: 10.1016/j.ijfatigue.2012.04.023.
- [87] M. Steenbergen and R. Dollevoet. On the mechanism of squat formation on train rails - Part II: Growth. *International Journal of Fatigue* **47** (2013), 373–381. DOI: 10.1016/j.ijfatigue.2012.04.019.
- [88] M. Steenbergen. Rolling contact fatigue in relation to rail grinding. *Wear* **356-357** (2016), 110–121. ISSN: 0043-1648. DOI: 10.1016/j.wear.2016.03.015.
- [89] R. Stock et al. Material concepts for top of rail friction management — Classification, characterisation and application. *Wear* **366** (2016), 225–232. ISSN: 0043-1648. DOI: 10.1016/j.wear.2016.05.028.
- [90] C. T. Sun and Z. H. Jin. *Fracture mechanics*. Waltham, MA: Academic Press, 2012. 337 pp. ISBN: 9780123850010.
- [91] T. Telliskivi and U. Olofsson. Contact mechanics analysis of measured wheel-rail profiles using the finite element method. *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit* **215.2** (2001), 65–72. ISSN: 0954-4097. DOI: 10.1243/0954409011531404.
- [92] O. Vargolici et al. Influence of the initial surface state of bodies in contact on the formation of white etching layers under dry sliding conditions. *Wear* **366** (2016), 209–216. ISSN: 0043-1648. DOI: 10.1016/j.wear.2016.06.023.

- [93] E. A. H. Vollebregt. *User guide for CONTACT, Rolling and sliding contact with friction*. Technical report TR09-03, version 17.1. VORtech BV. 2017. URL: www.kalkersoftware.org.
- [94] E. A. H. Vollebregt and G. Segal. Solving conformal wheel-rail rolling contact problems. *Vehicle System Dynamics* **52.1** (2014), 455–468. ISSN: 0042-3114. DOI: 10.1080/00423114.2014.906634.
- [95] S. L. Wong et al. A branch criterion for shallow angled rolling contact fatigue cracks in rails. *Wear* **191.1-2** (1996), 45–53.
- [96] X. Zhao et al. Local rolling contact fatigue and indentations on high-speed railway wheels: Observations and numerical simulations. *International Journal of Fatigue* **103** (2017), 5–16. ISSN: 0142-1123. DOI: 10.1016/j.ijfatigue.2017.05.014.
- [97] X. Zhao and Z. Li. A three-dimensional finite element solution of frictional wheel-rail rolling contact in elasto-plasticity. *Proceedings of the Institution of Mechanical Engineers, Part J: Journal of Engineering Tribology* **229.1** (2015), 86–100. DOI: 10.1177/1350650114543717.
- [98] X. Zhao, Z. Li, and R. Dollevoet. Influence of the fastening modeling on the vehicle-track interaction at singular rail surface defects. *Journal of Computational and Nonlinear Dynamics* **9.3** (2014). DOI: 10.1115/1.4025895.
- [99] X. Zhao, Z. Li, and R. Dollevoet. The vertical and the longitudinal dynamic responses of the vehicle-track system to squat-type short wavelength irregularity. *Vehicle System Dynamics* **51.12** (2013), 1918–1937. DOI: 10.1080/00423114.2013.847466.
- [100] X. Zhao et al. A study on high-speed rolling contact between a wheel and a contaminated rail. *Vehicle System Dynamics* **52.10** (2014), 1270–1287. DOI: 10.1080/00423114.2014.934845.
- [101] X. Zhao et al. A study on dynamic stress intensity factors of rail cracks at high speeds by a 3D explicit finite element model of rolling contact. *Wear* **366** (2016), 60–70. ISSN: 0043-1648. DOI: 10.1016/j.wear.2016.06.001.
- [102] X. Zhao and Z. Li. The solution of frictional wheel-rail rolling contact with a 3D transient finite element model: Validation and error analysis. *Wear* **271.1** (2011), 444–452. ISSN: 0043-1648. DOI: 10.1016/j.wear.2010.10.007.
- [103] O. C. Zienkiewicz and R. L. Taylor. *Finite Element Method: Vol. 1: The Basis*. 5th ed. San Diego: Elsevier Science & Technology Books, 2000. 689 pp. ISBN: 9780750650496.
- [104] A. Zoeteman, R. Dollevoet, and Z. Li. Dutch research results on wheel/rail interface management: 2001-2013 and beyond. *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit* **228.6** (2014), 642–651. DOI: 10.1177/0954409714524379.