The effect of bidirectional traffic on the microstructure, mechanical properties and rolling contact fatigue defect formation of rail material

Jorik Huisman
Student number: 4368371
Date: 11^{th} of June 2021

ProRail

Company supervisor: Ir. B. Schotsman
Company: ProRail
TU Delft supervisor: Prof. Dr. Ir. J. Sietsma

Materials Science and Engineering

TU Delft

Delft University of Technology

Department name: Materials Science & Engineering
University: TU Delft
Country: The Netherlands

Abstract

Railways are an essential part of the infrastructure in the Netherlands. Before the pandemic, the network of 6500 km track was used by 5400 passenger trains, servicing 1.2 million people and 300 freight trains transporting 100.000 tonnes of cargo [1] every day. Railways play a vital role in the transport of people between cities and suburban areas as well as the transport of cargo, while also being one of the safest and most energy-efficient means of transportation. If the energy consumption is expressed as energy cost per person per kilometre, travel by rail consumes only one-seventh of the energy used for travel by car [2]. Before the pandemic, rail traffic had been increasing at a rate of 5% each year [1]. It is important to keep the railways in proper condition and maximize their lifetime while minimizing maintenance time and costs.

Rails are central to the functioning of railways. The lifetime of a rail is primarily governed by the surface of the railhead, where wear or rolling contact fatigue (RCF) occurs due to wheel contact. [3]. RCF is the most critical failure mechanism. Due to RCF, cracks can initiate at the rail surface. If no action is taken, these cracks can cause spallation of the rail surface, or worse, transverse rail fracture, which may cause a derailment. Furthermore, the replacement of rails affected by RCF is a main cost driver in railway maintenance. It is important to have an understanding of the initiation and propagation of RCF cracks to maximize the lifetime of rails and minimize maintenance costs.

Many studies have been published which study RCF crack formation on rails. Almost all of these studies consider RCF crack formation on tracks with unidirectional traffic. This is to be expected because the majority of railways consist of two (or more) tracks where each track has a dedicated travel direction. However, on routes where demand is low, railways exist that consist of only one track. On these railways, traffic travels in both directions. Little is known in literature on the effect of bidirectional traffic on the formation of RCF cracks. Further investigation is required.

The motivation for this study is the discovery of RCF defects on the railway between Zutphen and Hengelo. This railway is located in one of the less densely populated areas in the Netherlands, where passenger demand is relatively low. The majority of the railway consists of a single track where the traffic is bidirectional. Furthermore, due to the low passenger demand, the majority of train traffic consists of a single type of train. These trains are relatively short and light. As a result, the annual tonnage (the cumulative sum of the total axle loads) is low. Despite the low annual tonnage, a large number of RCF defects have developed on the rails of this track.

The objective of the present work is to examine the effect of the bidirectional traffic on the rail material and the formation of RCF cracks. Two research goals are formulated:

- Understanding how bidirectional traffic changes the effects of wheel-rail contact on the microstructure and mechanical properties of the rail material compared to unidirectional traffic.
- Understanding how RCF cracks initiate and propagate as a result of bidirectional traffic.

To accomplish these research goals, rail specimens are taken from the railroad between Zutphen and Hengelo. The specimens are examined by performing a metallographic investigation. The metallographic investigation is supported by the analysis of the wheel-rail contact, which is simulated with multibody wheel-rail contact simulation software. Based on the results of this investigation, a number of conclusions is drawn.

In longitudinal direction, optical microscopy revealed the absence of plastic strain accumulation in the surface layer of the rail. This plastic strain, which is the result of the tangential surface traction, is always observed on rails subjected to unidirectional traffic and is considered the leading cause of fatigue crack initiation. The absence of this plastic strain can have two explanations: (1) the material (after initial hardening) is not repeatedly plastically deformed, or (2) the material is repeatedly plastically deformed, but does not accumulate plastic strain in a particular direction. The wheel-rail contact simulation showed that the contact stresses exceed the elastic shakedown limit, so plastic deformation is expected. By the analysis of the trajectory of the principal shear stress, it was shown that repeated plastic deformation will occur at a lower stress level for bidirectional traffic than for unidirectional traffic, due to the frequent reversal of tangential surface traction. From the observations on the microstructure and the analysis of the contact stresses, it is concluded that the frequent direction reversal of the tangential surface traction causes cyclic plasticity without the accumulation of strain, detectable with optical microscopy.

Etching of the running surface, optical microscopy and hardness measurements all showed the presence of WEL on the surface of the studied rail specimens. Based on the observed shape and orientation of the WEL on the rail surface, it was concluded that the WEL formed due to slip in the wheel-rail contact patch. However,

the wheel-rail contact simulation showed that, for normal running behaviour, the creep ratio is not high enough to generate the frictional heat flux into the rail required to elevate the material temperature above the austenitization temperature. No indication was found that indicates that bidirectional traffic has had an effect on the formation of WEL.

All cracks on the rail samples are positioned on grinding facet two but a clear initiation mechanism could not be determined. Although WEL was found in the proximity of the RCF defects, no observation was made that WEL actually resulted in crack initiation. The spatial distribution of defects also does not provide insight into the conditions of initiation. There is a real possibility, especially for the larger category C defects, that crack initiation of the defects found on grinding facet two occurred prior to the grinding maintenance in 2017. It is expected that the layer of material from which the crack had initiated was removed by wear or during grinding.

Optical microscopy did not observe any indication of crack initiation by either MnS inclusions or PE ferrite. Fatigue cracks have not been initiated by delamination of heavily strained MnS inclusions and pro-eutectoid ferrite grains as a result of bidirectional traffic. MnS inclusions and PE ferrite grains are not elongated in the direction of the strain. Therefore, the boundaries between MnS inclusion or PE ferrite and the pearlite matrix are not weakened enough to facilitate crack initiation.

Once initiated, the propagation of the RCF cracks is governed by the shear stresses generated in the contact patch. The observed propagation direction is consistent with the simulated direction of maximum shear stress. The presence of fluid in the crack, as indicated by the oxides on the crack faces, likely decreased the crack face friction and assisted crack propagation.

Although developed as a result of bidirectional traffic, the observed crack morphology showed similarities to the squat RCF defect. Analysis of the subsurface crack morphology of a severe squat-like RCF defect showed a high degree of symmetry, which is not observed for squats formed on rails subjected to unidirectional traffic. Furthermore, branching was observed in either longitudinal propagation directions, which for unidirectional traffic is only observed in the traffic direction.

The subsurface crack morphology of a large category C U-crack is analysed. In contrast to squat defects formed by unidirectional traffic, the subsurface crack morphology of the category C U-crack showed to be symmetric. Furthermore, branching was observed in either longitudinal propagation directions, which for unidirectional traffic is only observed in the traffic direction. New insights were generated by considering the implications of bi-directional traffic and the rolling direction reversal on existing theories about the rail squat formation mechanism. It is concluded that bidirectional traffic is the cause of this symmetry.

Further research is required to improve the understanding of the effects of bidirectional traffic and the process of the RCF defect formation in general.

Optical microscopy showed an absence of longitudinal plastic strain accumulation in the surface layer of the rail. Further research is necessary to examine the plastic behaviour of the surface layer. Analysis of the plastic behaviour could provide insight into the relation between bidirectional traffic and crack initiation and could lead to the improvement of plasticity models.

The observed symmetry of the subsurface crack morphology and indicates that the propagation of the crack front in either longitudinal direction is governed by the same crack propagation mechanisms. Analysis of the asperities on the crack faces could further clarify the responsible crack propagation mechanisms and could provide insight into the crack growth rate.

Additional research should be performed on the conditions of WEL formation on the rail surface. Given the simulated creep ratios, it is expected that there must be some other factors at play which enable the generation of the heat flux required for material transformation. Field measurements of the running behaviour of trains could provide new insights into this phenomenon.

Based on the results in this study, no conclusion could be drawn on the crack initiation mechanism. Additional research should be done to further investigate the processes involved in crack initiation. Since the process of RCF defect initiation and propagation develops over time, further research should consider all stages in the life of RCF formation in finer detail. That way there is a higher chance of determining the initiation point with certainty.

Acknowledgement

I would like to express my gratitude to several people, whose help and guidance was essential for the completion of this project. First of all, I would like to thank ir. B. Schotsman for his continued support and daily supervision. His supervision was essential throughout all stages of the project, and his expertise, insights and critical eye provided the needed guidance and direction for the progression of this project. My second thank you goes to prof. dr. ir. J. Sietsma, who provided the opportunity to work on this project and helped to keep the project on the right track.

This project was carried out in collaboration with ProRail. Even though I did not (yet) meet any ProRail colleagues in person due to the pandemic, my thanks go out to all of the members of the track expert team for their contributions. Furthermore, I would like to express my gratitude to several people outside of ProRail who contributed to this project: dr. M Hinsch and ir. L van Rijen of Dekra rail for their help on the wheel-rail contact simulation, Caroline van Loenen of Keolis and Arno Vagevuur of NS for providing the access to the LINT trains. Special thanks go out to Gijs Brefeld of VolkerRail, for providing the rail specimen.

My gratitude goes out to the staff of the TU Delft Materials Science and Engineering staff for providing the training and access to the equipment required to perform the experimental work. Special thanks go out to prof. dr. ir. J van Turnhout for his good company and support.

Last, but certainly not least, I would like to thank my family and girlfriend for their continued support and encouragement. Their much-appreciated feedback helped a great deal in the writing of this report.

Contents

Chapter 1: Introduction

1.1 Importance of maximizing the lifetime of rails

Railways are an essential part of the infrastructure in the Netherlands. Before the pandemic, the network of 6500 km track was used by 5400 passenger trains, servicing 1.2 million people and 300 freight trains transporting 100.000 tonnes of cargo [1] every day. Railways play a vital role in the transport of people between cities and suburban areas as well as the transport of cargo, while also being one of the safest and most energy-efficient means of transportation. If the energy consumption is expressed as energy cost per person per kilometre, travel by rail consumes only one-seventh of the energy used for travel by car [2]. Before the pandemic, rail traffic had been increasing at a rate of 5% each year [1]. It is important to keep the railways in proper condition and maximize their lifetime while minimizing maintenance time and costs.

Rails are central to the functioning of railways. The lifetime of a rail is primarily governed by the surface of the railhead, where wear or rolling contact fatigue (RCF) occur due to contact with train wheels [3]. RCF is the most critical failure mechanism. Due to RCF, cracks can initiate at the rail surface. If no action is taken, these cracks can cause spallation of the rail surface, or worse, transverse rail fracture, which may cause a derailment. Furthermore, the replacement of rails affected by RCF is a main cost driver in railway maintenance. It is important to have an understanding of the initiation and propagation of RCF cracks to maximize the lifetime of rails and minimize maintenance costs.

Many studies have been published which study RCF crack formation on rails. Almost all of these studies consider RCF crack formation on tracks with unidirectional traffic. This is to be expected because the majority of railways consist of two (or more) tracks where each track has a dedicated travel direction. However, on routes where demand is low, railways exist that consist of only one track. On these railways, traffic travels in both directions. Little is known in literature on the effect of bidirectional traffic on the formation of RCF cracks. Further investigation is required.

1.2 The objective of the study

The motivation for this study is the discovery of RCF defects on the railway between Zutphen and Hengelo. This railway is located in one of the less densely populated areas in the Netherlands, where passenger demand is relatively low. Therefore the majority of the railway consists of only one track where the traffic is bidirectional. Furthermore, due to the low passenger demand, the majority of traffic consists of a single type of train. These trains are relatively short and light. Therefore the annual tonnage (the cumulative sum of the total axle loads) is low. Despite the low annual tonnage, a large number of RCF defects have developed on the rails of this track.

The objective of the present work is to examine the effect of the bidirectional traffic on the rail material and the formation of RCF cracks. To study this effect, specimens containing RCF defects are taken from the track and analysed with multiple characterization techniques. The wheel-rail contact is analysed through a multibody wheel-rail contact simulation. The research goals are as follows:

- Understanding how bidirectional traffic changes the effects of wheel-rail contact on the microstructure and mechanical properties of the rail material compared to unidirectional traffic.
- Understanding how RCF cracks initiate and propagate as a result of bidirectional traffic.

1.3 The structure of the report

The structure of the report is as follows. In chapter 2, the literature review is presented. The goal of the literature review is to create an overview of the current understanding of the effects of unidirectional traffic on the rail material and RCF defect formation, to which the bidirectional results of the current study can be compared. This review will cover pearlitic rail steel grades, the mechanics of the wheel-rail contact, the effects of the wheel-rail contact on the rail material and the formation of common RCF defect types. To conclude this chapter the effects of the unconventional track conditions (low annual tonnage, homogeneous rolling stock and bidirectional traffic) on the rail material and RCF formation is presented. Chapter 3 continues by listing the history and characteristics of the rail specimens taken from the studied track. Next, the analysis methods and equipment used for the macro- and micro-characterization are listed. Then, the set-up and input for the multibody wheel-rail contact simulation are presented. Finally, the construction process of a 3D crack

contour model is explained. In chapter 4 the results of the simulations and characterization experiments are presented. The results are discussed in chapter 5. The discussion consists of two parts. The first part discusses the effect of bidirectional traffic on the microstructure and mechanical properties of the rail material. In the second part, the process of RCF crack formation is discussed in the context of the material microstructure and mechanical properties affected by the bidirectional traffic. Finally, in chapter 6 a summary is given of the knowledge acquired for each of the research goals and possibilities for further research will be pointed out.

Chapter 2: Literature Review

In this literature review, the reader is introduced to the subject of rolling contact fatigue (RCF) in pearlitic rail steels. Firstly, the microstructure, mechanical properties and manufacturing process of pearlitic rail steels will be discussed in section 2.1. Special attention is given to the R260Mn rail steel grade. The formation of pearlite and other important phase transitions in the Fe-C system will be explained based on the phase diagram. Secondly, an overview of the stresses to which the rail is subjected during its lifetime is presented in section 2.2. These stresses are the primary cause of the development of RCF defects. The combination of multiple simultaneous loading types during a stress cycle creates a complex stress state, especially near the wheel-rail interface. Thirdly, the effects of the stress cycles on the material properties and the microstructure are analyzed in section 2.3. Factors such as plastic deformation, wear and thermal loading affect the initiation and propagation of RCF. Fourthly, the formation of defects caused by RCF is discussed in section 2.4. Weak sites in the microstructure, which function as nucleation sites for RCF cracks, are identified. An overview of different types of RCF defects is presented. Special attention is given to the "squat" type RCF defect. Theories in literature on the initiation and propagation of squats are presented and discussed. Finally, an overview of the current understanding of the effects of annual tonnage, homogeneous rolling stock and bidirectional traffic RCF is presented in section 2.5.

2.1 Pearlitic rail steels

Since the introduction of steel rails, the focus of rail steel standardization has been on fatigue and wear resistance to meet the increasing requirements for higher axle loads, increased speeds, lower costs and higher safety[4]p2.2. Over the past decades standardized rail steel grades have been developed. The majority of these grades are low-alloy, carbon-manganese steels with a carbon content of 0.5 weight percent (w%) to eutectoid composition (0.76 w%), or even hyper-eutectoid weight percentages [5][6]. The microstructure of these steel grades consists of pearlite with varying fractions of pro-eutectoid (PE) ferrite or cementite [5][7]. Pearlitic grades are used primarily because of the combination of their mechanical properties and their low costs [5]. In this section, the formation mechanisms, microstructure, mechanical properties and manufacturing process of pearlitic steels are discussed.

2.1.1 Phase transformations and microstructure of pearlitic steels

While not being an actual phase in itself, pearlite is one of the most common microstructures found in steels. A pearlitic microstructure is characterised by colonies that contain a lamellar structure of alternating ferrite (α) and cementite (θ) lamellae [8] An example of this structure is shown in figure 2.1. Although there are alloying elements present, pearlite formation can be understood with the aid of the Fe-C equilibrium phase diagram in figure 2.2. This phase diagram is calculated with the Thermo-Calc software [9]. Two types of equilibrium transformation regimes can be distinguished depending on the carbon fraction being smaller (hypo-) or larger (hyper-) than the eutectoid composition [8].

The hypo-eutectoid transformation starts by slowly cooling from point 1_a in figure 2.2. Here the structure is fully austenitic (γ) and has a face-centred cubic (FCC) crystal lattice with all carbon in solid solution. When the A_3 temperature is reached (point 1_b), grains of body centred cubic (BCC) ferrite start to nucleate at the

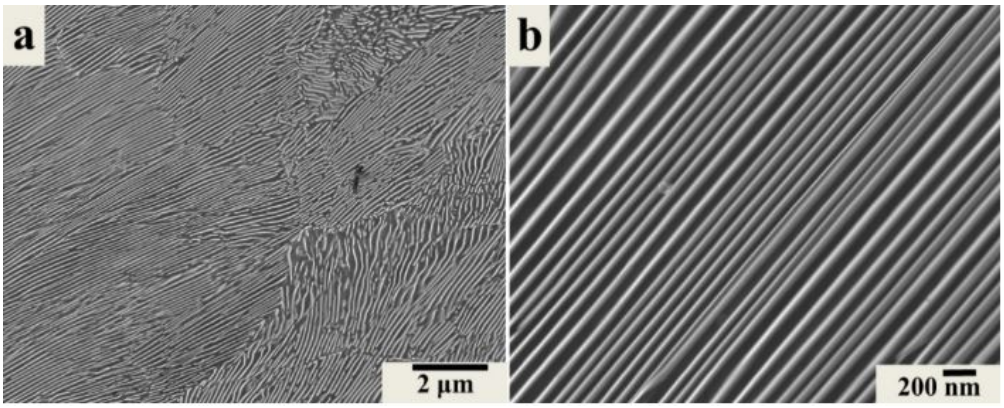


Figure 2.1: *SEM image of a pearlitic structure, showing ferrite in dark grey and cementite in light grey [10]*

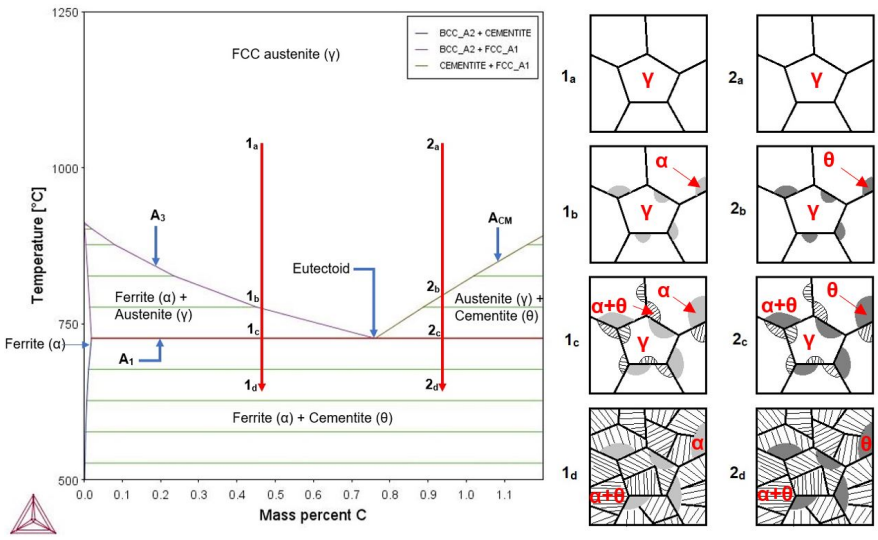


Figure 2.2: *The Fe-C equilibrium phase diagram, as calculated with Thermo-Calc [9]. The two vertical red arrows indicate (1) hypo- and (2) hyper-eutectoid transitions. The drawing on the right schematically represents the microstructure during the transformation from austenite to pearlite.*

austenite grain boundaries, classically named pro-eutectoid ferrite (PE ferrite) [8][11]. The solubility of carbon in ferrite is much lower than in austenite (0.02 w% and 2.06 w% respectively at the A_1 temperature) [8]. PE ferrite grain growth depends on the diffusion of carbon from the PE ferrite to the surrounding austenite. This increases the carbon concentration in the austenite [8]. When the eutectoid temperature A_1 is reached in point 1_c, the remaining austenite, now at eutectoid carbon concentration, simultaneous decomposition into ferrite and cementite occurs ($\gamma \rightarrow \alpha + \theta$) to form pearlite [6][8]. Pearlite colonies originate at the austenite grain boundaries, austenite-ferrite boundaries and/or inclusions. They grow into the austenite grains until all austenite is transformed to pearlite at a point below 1_c [8][12]. Pearlite with a non-eutectoid composition can also be formed if the austenite is quickly cooled and subsequently held at a temperature below the A_1 temperature. Due to the quick undercooling, the nucleation of pro-eutectoid phases is suppressed and the austenite is transformed into a fully pearlitic structure. The temperatures and compositions where this transformation is possible are given by the Hultgren extrapolation of the α/γ and γ/θ phase boundaries [8].

If austenite with a carbon concentration larger than eutectoid concentration is slowly cooled to the A_{cm} temperature (point 2_b in figure 2.2), pro-eutectoid cementite grains form at the austenite grain boundaries [8][11]. Cementite has a higher carbon concentration (6.67 w%) than austenite [13]. Therefore it depletes the remaining austenite in carbon. When the eutectoid temperature is reached (point 2_c), the remaining austenite transforms into pearlite at eutectoid composition.

Two processes cost energy during the formation of pearlite: the transformation of austenite into ferrite and cementite and the creation of ferrite/cementite interfaces [8]. This diffusional transformation is temperature-dependent, specifically on the level of undercooling below the A_1 temperature [8]. The undercooling has two effects: it slows down diffusion and it increases the energetic driving force for the transformation of FCC to BCC to occur [8][13]. This has several consequences for the microstructure. For a small undercooling the driving force is relatively small, but the diffusivity is high. Therefore carbon can diffuse over large distances to form wide lamellae with a larger interlamellar distance, which minimizes the energy needed for ferrite/cementite interfaces. For a larger undercooling carbon can not diffuse as far. However, due to the larger driving force, it is possible to create more ferrite/cementite interface, resulting in a smaller interlamellar spacing [8][13]. Additionally, a larger driving force increases the pearlite colony nucleation rate, which results in a smaller average colony size [8][13].

So far, all described transformations are equilibrium phase transformations, which are governed by diffusion. Diffusivity decreases with temperature. Therefore equilibrium transformations can only occur at sufficiently low cooling rates, where there is enough time for all atoms to diffuse. An example of a low cooling rate the 0.1°Cs^{-1} trajectory in the continuous cooling transformation (CCT) diagram in figure 2.3 [14]. This is not a diagram of rail steel, but they show a behaviour similar to hypo-eutectoid steel. For high cooling rates, for example the 43°Cs^{-1} trajectory in figure 2.3 [14], there is not enough time to complete the transformation

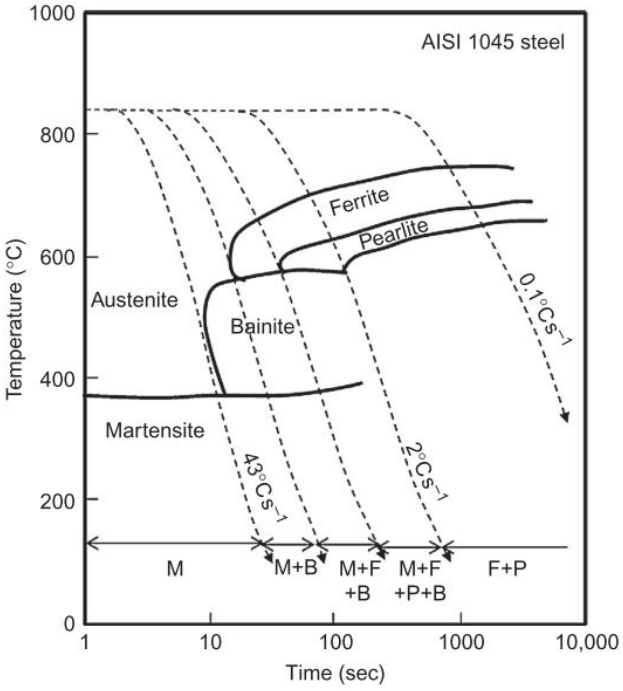


Figure 2.3: *CCT diagram of AISI 1045 steel (0.42-0.5 w% C, 0.6-0.9w% Mn*

through diffusion. However, the driving force for the FCC lattice to transform into a BCC lattice becomes so large that a diffusionless shear transformation occurs. The resulting meta-stable phase is known as martensite [15]p414. This transformation causes a shape change and expands the volume of the transformed region by 3 to 5%, which exerts stress on the surrounding matrix [8][15]p415. Excess carbon is retained in solid solution, which must be accommodated in the BCC lattice. In the FCC austenite, the carbon atoms sit at the octahedral interstices. Austenite has one octahedral interstice per atom. Ferrite has three octahedral sites per unit cell, which are located in the centre of the cell edges. Therefore, BCC contains three sublattices of octahedral interstices; one for each orthogonal cube edge of the BCC unit cell. Only one out of the three octahedral sublattices in BCC coincides with the octahedral interstice lattice in FCC. Therefore, the diffusionless transformation causes all carbon atoms to end up at one of the three octahedral sites (at the z-axis) in BCC. This causes the lattice to strain in one particular direction (the Bain strain), resulting in a body-centred tetragonal (BCT) unit cell [16].

Martensite is known to have a high hardness and yield strength but is also known to be very brittle [16]. The high yield stress of martensite is closely related to the carbon concentration; higher carbon concentrations cause an increase in solid solution strengthening [8]. Martensite is also strengthened by its high dislocation and twin density, which are created during the transformation. The brittleness is caused by a combination of the static stress that is exerted on the matrix due to the supersaturation of carbon and the high defect density created in the transformation [17]. This can lead to brittle fracture at low external loads. In practice, the material fractures before the yield strength is reached [17].
At intermediate cooling rates, the distinction between diffusional and diffusionless transformations is not as clear as for high or low cooling rates; the larger iron atoms undergo a shear transformation, while the smaller carbon atoms are still able to diffuse. It is at this regime where bainite is formed [14][18]. Depending on the cooling rate multiple different phases can be formed at different temperatures during cooling, as indicated in figure 2.3. Bainitic steels are used in some railway applications but have not seen widespread use [5].

2.1.2 Microstructure and mechanical properties of pearlitic rail steel

It is well known that for steels an increased carbon concentration generally increases its strength, but decreases its fracture toughness [19]. For low carbon steels, the mechanical behaviour of the ferritic microstructure governs the mechanical behaviour of the material. Some pearlite does occur as a second phase but has little effect on the properties. The fraction of pearlite increases with increasing carbon content. As the pearlite fraction increases, so does its influence on the mechanical properties of the material [8][20][21].

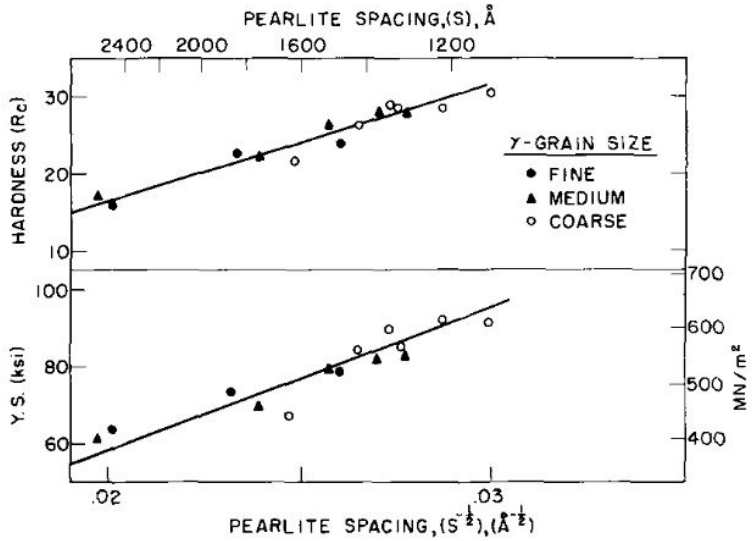


Figure 2.4: *Yield strength and hardness as a function of pearlite interlamellar spacing [21]*

For fully pearlitic steels it is reported that the yield strength and hardness have a strong correlation with the interlamellar spacing; a reduction in the interlamellar spacing results in an increase in the yield strength, see figure 2.4. This relationship obeys the Hall-Petch relationship [8][21][22][23]. A possible explanation for this relation is that for smaller interlamellar spacing, the mean free path for dislocations in the ferrite lamellae is reduced, as the cementite lamellae are effective barriers against dislocation motion [21]. Macroscopic yielding would therefore require new dislocations to be created in the ferrite lamellae at first, as existing dislocations pile up at the ferrite-cementite boundaries. Cementite also starts to deform when work hardening strengthens the ferrite to a similar level [8].

The fracture toughness and ductility of pearlite are controlled by the prior austenite (PA) grain size; an increased austenite grain size decreases the toughness and ductility [21][24][25]. An explanation for this relation is given in [25]: pearlite colonies originate at the PA grain boundaries and grow into one of the two adjacent austenite grains while having an orientational relationship with the other austenite grain. For a large austenite grain size, multiple colonies may originate at the same PA grain boundary and therefore have a similar crystallographic orientation, forming an "orientation unit". For smaller PA grain size, the number of similarly oriented colonies would be limited, resulting in a smaller orientation unit. The size of these orientation units can be considered as the effective parameter for fracture toughness, as a crack can propagate without a change of direction within a single orientation unit. It is reported that the pearlite colony size does not have a significant impact on the fracture toughness [21].

R260Mn grade steel

For the past few decades, rails have been manufactured from low alloyed carbon-manganese steel with medium to near-eutectoid carbon concentrations. On the Dutch railways the steel grades R260Mn, R350HT and R370CrHT are used [26]. The steel names are composed of two principal symbols and optionally two additional symbols [27]. The two principal symbols are a letter and a three-digit number, representing the type of steel (R = rail steel) and the minimum Brinell hardness of the surface. The additional symbols indicate possible important alloying elements or heat treatments. An overview of the chemical compositions of these steels is given in table 2.1.

This research will focus on the R260Mn grade. This grade is the most widely used, as it is used in straight tracks and curves with a radius larger or equal to 2000 m [29]. The other grades are used for the outer rail for curves with radii smaller than 2000 m, because their increased hardness provides the required wear resistance. As indicated by the name, R260Mn steel has a minimum Brinell hardness of 260 HBW measured at the surface of the rail head. Its ultimate tensile strength of 880 MPa is high compared to other types of mass-produced steels such as structural steels or pipe steels, which have tensile strengths between 360 and 550 MPa [30][31]. R260Mn has a minimal fracture toughness of 26 MPa $m^{1/2}$ [28]. This is lower than for structural steels, which have a fracture toughness of approximately 39 MPa $m^{1/2}$ [32]. The carbon composition range for R260Mn

Table 2.1: *Chemical composition of rail steel grades on Dutch railways[26][28]*

Steel grade	Curve radius (m)	% by mass						Tensile strength σ_{uts} (MPa)	Elonga-tion ϵ (%)	Hard-ness of run-ning surface (HBW)
		C	Si	Mn	P (max)	S (max)	Cr			
R260Mn	$>$ 2000	0.53 to 0.77	0.13 to 0.62	1.25 to 1.75	0.025	0.030	$\leq$ 0.15	880	10	260 to 300
R350HT	$<$ 2000	0.70 to 0.82	0.13 to 0.60	0.65 to 1.25	0.025	0.025	$\leq$ 0.15	1175	9	350 to 390
R370CrHT	$<$ 2000	0.68 to 0.84	0.38 to 1.02	0.65 to 1.15	0.025	0.025	0.35 to 0.65	1280	9	370 to 410

grade steel is relatively large, ranging from a moderate ferrite-pearlite steel to a fully pearlitic steel. The roles of the other alloying elements are as follows:

- Mn: manganese shifts the eutectoid composition to a lower carbon concentration. Therefore, the addition of manganese will result in a larger pearlite fraction for a certain carbon concentration [5][8]. It also lowers the A_3 and A_1 temperatures, which reduces the growth rate of pearlite [33]. Additionally, it binds to sulphur to create MnS inclusions [5]. The effect of inclusions will be discussed in section 2.4.1.
- Si: silicon is used as a solid solution strengthener for (pearlitic) ferrite [34].
- Cr: Chromium helps to refine the interlamellar spacing of the pearlite. Additionally, it reduces softening in welds [34]. Due to the low Cr weight fraction, its effect in R260Mn steel is limited.
- S & P: sulphur and phosphor are impurities remaining from the steel production process [35].

2.1.3 Manufacturing process

Similar to most other steels, the preferred manufacturing method for rail steel is continuous casting. Steel strands are created when liquid steel flows through a mould, where it starts to solidify. The strands are cut into billets with a standard length and cross-section and are allowed to cool [4]p4.17. The billets are shaped into the desired rail profile by several rolling passes. Before rolling, the billets are austenitized by reheating them to 1250°C. All rolling passes are performed when the material is in the austenitic phase. After rolling, the rails are transferred to cooling beds to increase the homogeneity of temperature during cooling by controlling the cooling rate [4]p4.18,[36]. Due to the difference in geometry, the rail head and rail foot cool at different rates. This causes the rails to bend in vertical direction. Therefore the rails are straightened by cold-rolling in a roller straightener [4]p4.19,[37].

2.2 Stresses in the rail

During its lifetime, a rail experiences a large number of loading cycles due to the wheel loads exerted on it by passing trains. The combination of the geometry of the wheel-rail interface and the train loads induce stresses in the rail. The shear stresses caused by these loads are the cause for the development of rolling contact fatigue (RCF) defects. An overview of all the possible stresses that can act on a rail is given in figure 2.5. These will be discussed in the following sections.

2.2.1 Track geometry

The geometry of the wheel-rail interface has a strong influence on the stresses that act upon the system, and the way the train is guided along the track [39][40]p17. Instead of the cylindrical shape of normal wheels, train wheels are conical. A wheel-set is mounted on a fixed axle. This causes a centering lateral force when there is a lateral displacement, which guides the train along the rails [40]p17. On the gauge side (see figure 2.6), train wheels have flanges to prevent derailment. The rails have standardized I-shaped profiles and are

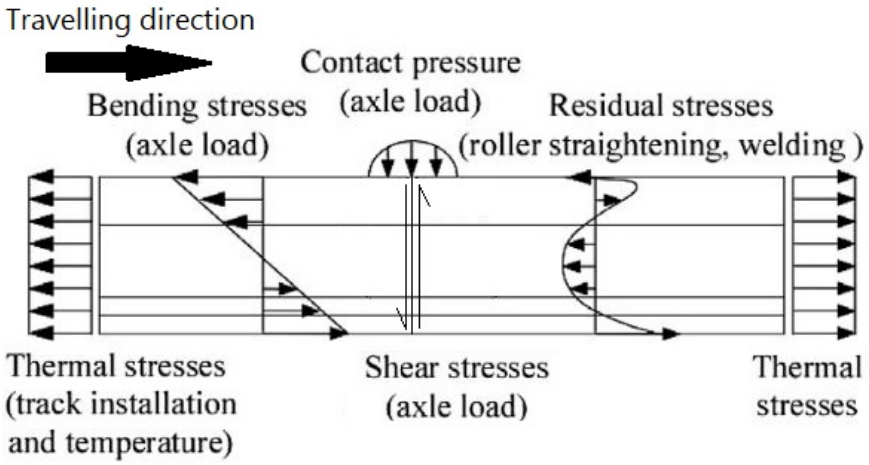


Figure 2.5: *Overview of types of stresses acting on rails. Side-view of the rail. Adapted from [38]*

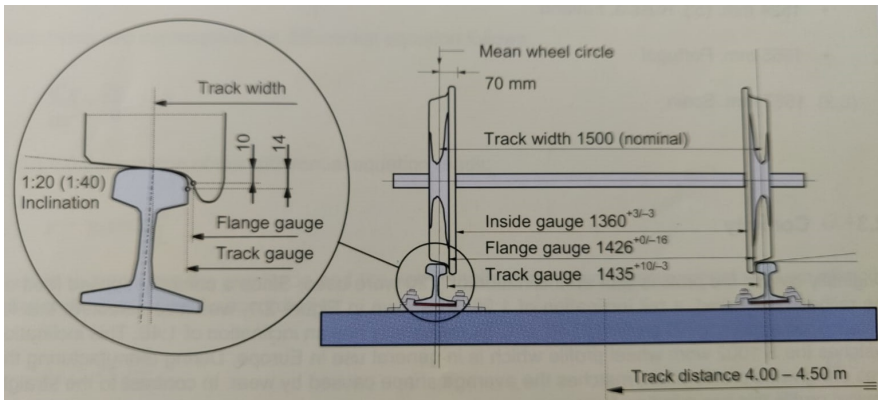


Figure 2.6: *Transverse cross-section of a wheel set on a straight track [40]p17*

connected to the sleepers under a slight inclination, which matches the inclination of the wheel conicity [39]. In straight tracks, contact is expected to occur between the rail head and wheel tread [41].

2.2.2 Stresses due to wheel loading

Bending stresses

A wheel load induces a bending stress envelope in the rail head around the position of the wheel. This envelope is shown in figure 2.7. The beam-on-elastic-foundation theory predicts that directly underneath the wheel a maximal compressive stress is reached [38]. At a distance x_0 (order of magnitude $\tilde{1}$ m) in front of and behind the point of contact, two tensile stress maxima are located. The material in the rail head experiences a cycle of axial stresses when a wheel passes over it: tensile stress, followed by compressive stress, followed by tensile stress. The distance x_0 is a function of the geometry and the material characteristics parameter λ_V, as given by the relation in figure 2.7 [38]. λ_V is calculated with the following relation:

$$\lambda_V = \left(\frac{k_V}{4E \cdot I_{yy}}\right)^{\frac{1}{4}} \tag{2.1}$$

where k_V is the vertical stiffness of the elastic foundation, E the modulus of elasticity and I_{yy} the second area moments of inertia with respect to the horizontal axis.

In practice, train axles are grouped in sets of two, which causes two bending stress envelopes to overlap. This causes the actual bending stress envelope to be a superposition of the two individual wheel envelopes. Depending on the wheel spacing two tensile bending stress maxima can overlap which increases the maximal bending stress experienced by the rail [36][40]p85. The maximum bending stress is reached at the surface, as shown in figure 2.5. Therefore they contribute to the initiation and growth of surface-initiated defects by

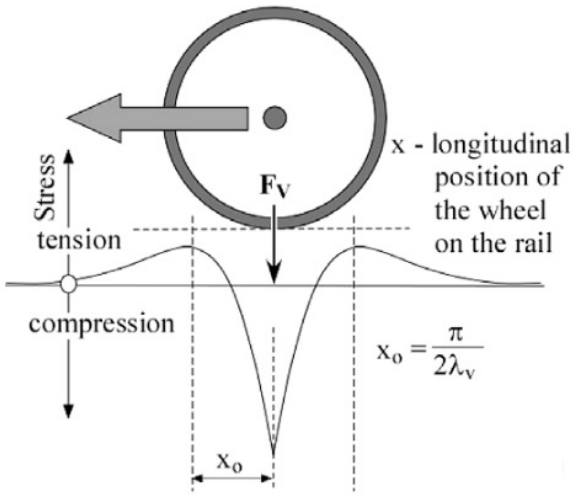


Figure 2.7: *Bending stress in the rail head as function of the wheel position. Adapted from [38]*

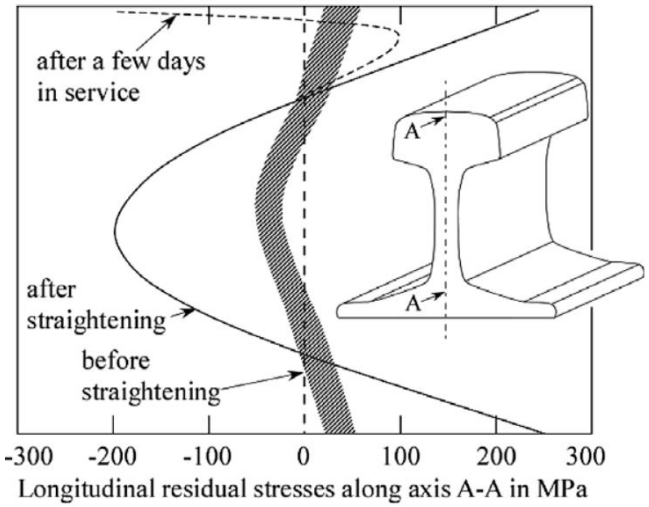


Figure 2.8: *Longitudinal residual stress profile in the center line of the rail before straightening, after straightening and after a few days of service [38]*

mode I loading [36][38]. To calculate the bending stresses, equations can be found in works by Zerbst et al. [38] and Esveld [40]p74.

Vertical shear stresses

At a certain point in the rail, the direction of the train load and the reaction forces from the sleepers induce vertical shear stresses in the rail. The direction of the shear stress changes after the wheel passes over it, as shown in figure 2.5. The rail experiences a large number of cycles of alternating shear stress during its lifetime. These alternating shear stress cycles can contribute to the propagation of surface cracks by mode II loading [38].

2.2.3 Residual stresses

During the manufacturing process, longitudinal residual stresses are created in the rail. As mentioned in section 2.1.3, the rails are shaped by multiple rolling passes while the steel is still austenitic. When the rails are cooled afterwards, internal stresses are created as a result of non-uniform temperature gradients. The internal stresses cause the rails to bend [37]. The subsequent roller-straightening plastically deforms the rail to remove the curvature but simultaneously creates residual stresses in the rail. The residual stresses from roller-straightening are tensile in the rail head and foot and compressive in the rail web [38][37]. After a few train passages the tensile residual stress in the top layer of the rail head changes to compression as a result of the strain-hardening [38], see figure 2.8.

2.2.4 Thermal stresses

Due to the differences between the installation temperature and the continuously changing service temperature thermal stresses develop in rails. Rails are installed with a temperature of 25°C [40]p388,[38]. Therefore rails experience tensile thermal stress for lower service temperatures and compressive thermal stress for higher service temperatures. For continuous straight tracks, the thermal stress σ_T can be calculated with $\sigma_T = \alpha E(T_N - T)$, where α is the thermal expansion coefficient, E Young's modulus, T_N the installation (neutral) temperature and T the service temperature [38].

2.2.5 Contact stresses

The wheel-rail contact patches are the only contact points between the train and the track. All forces that arise from travelling along the track must be transferred through these contact patches. The area of a contact patch has a typical size of 1 cm^2 [41]. Large train loads that are exerted on the rail via a small contact area cause high, complex stresses in the surface layer beneath the contact patch.

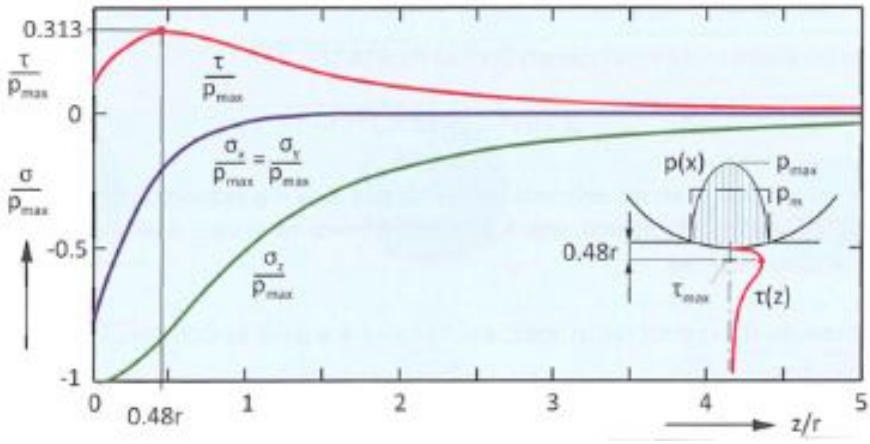


Figure 2.9: *Normalised Hertzian stress distribution under the centre of the contact patch of a ball in contact with a flat plane[42]*

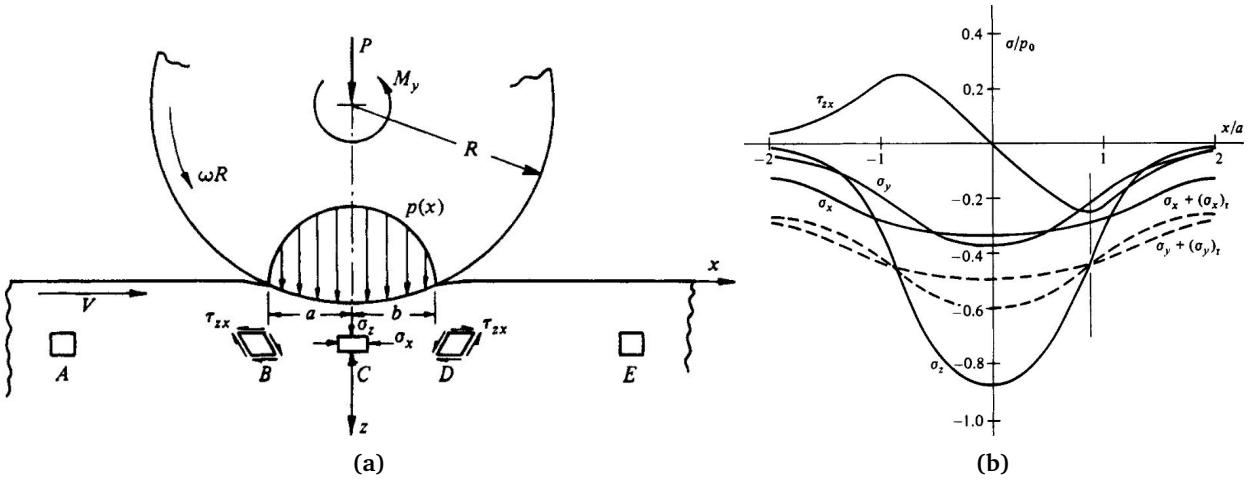


Figure 2.10: *Deformation and stress cycle of an element 0.48a below the surface due to a rolling contact of an elastic cylinder: (a) deformation cycle, (b) stress cycle [43], the broken lines represent the addition of residual stresses. The z-axis is fixed to the wheel and the rail material moves underneath the wheel with a velocity V.*

To understand the mechanics of bodies in rolling contact, the mechanics of bodies in stationary contact need to be understood first. The simplest and well-known model for understanding contact stresses between two bodies in contact is the Hertzian contact model [42][40]. For simplicity, it is first assumed a wheel-rail contact can be approximated by the contact of two parallel cylinders. For this case the Hertzian contact model predicts a line contact and parabolic compressive normal stress distribution $p(x) = p_0(1 - x^2/a^2)^{(1/2)}$ acting on a contact patch with width $2a$. Figure 2.9 shows the stress distribution as a function of distance below the surface at the centre of the contact patch. The normal stress reaches a maximum p_0 on the surface in the centre of the contact patch. Due to the Poisson effect compressive stresses are generated in x- and y-direction. The combination of stresses in x-, y- and z-direction results in shear stress. This shear stress reaches a maximum at $0.48a$ beneath the surface and acts on the set of planes that are oriented at an angle of 45 degrees to the surface normal.

The Hertzian contact model is not an exact description of reality as it makes four assumptions: "*(1) The surfaces are continuous and non-conforming, (2) the strains are small, (3) each solid can be considered as an elastic half-space, (4) the surfaces are frictionless*" [44]. For wheel-rail contact, the first two assumptions are a good approximation of reality. However, assumptions three and four need to be taken into account.

The surfaces are assumed to be frictionless so that only a normal pressure can be transmitted between them. The purpose of this assumption is to eliminate the effect of the tangential stresses in the contacting surfaces that can arise as a result of elastic dissimilarities. Since wheel and rail materials are elastically similar, these tangential stresses are minimal. As a result, the Hertzian solution is a good approximation of the elastic normal

stresses in the rail, despite the presence of friction between the wheel and rail surfaces. The implications of assumption (3) will be further discussed in section 2.3.1.

As a consequence of friction, tangential forces (i.e. traction) can be transmitted between the surfaces of two rolling bodies. Two types of rolling exist depending on the presence of a tangential force: free-rolling, where the tangential force is zero, and tractive rolling, where the tangential force is non-zero [45]. The application of traction influences the stresses in the contacting bodies.

First, we look at the case of free rolling. From the viewpoint of the axle of a freely rolling wheel, the stress distribution in the contacting bodies is Hertzian and fixed in space. From the perspective of a material element in the rail, however, this Hertzian stress distribution moves along in space with the movement of the wheel. As a result of this moving stress distribution, a material element in the rail experiences a stress and deformation cycle. This cycle is presented in figure 2.10. A cycle starts at position A in figure 2.10a, where no influence of the wheel is experienced by the element. When the wheel approaches the element, the shear in the xz-plane increases until a maximum is reached at position B. After this maximum is reached the shear decreases until the centre of the contact patch is exactly above the material element at position C. Here the shear direction reverses until another maximum is reached at position D. When the wheel moves away the stress returns to zero at position E. The shear maxima in the xz-plane are smaller than the maximum shear reached in the centre of the contact patch. However, the reversed cycle of shear is critical for fatigue, which will be discussed in section 2.3.1.

We now consider the case of tractive rolling. On top of the Hertzian contact stresses, the tangential force between the two contacting surfaces creates additional stresses in the contacting bodies. To find a description for these stresses, we first look at the effect of a tangential force on a stationary contact. For this, we follow Johnson's [46] analysis of a stationary contact on the onset of sliding. We consider two stationary surfaces in contact subjected to a normal force P and a tangential force Q, where the traction coefficient $\frac{Q}{P} \leq \mu$. We assume that P and Q are independent and that the normal stress distribution $p(x, y)$ is Hertzian. The bodies in contact will deform in shear as a result of this applied tangential force. If the two bodies as a whole do not slide relative to each other (i.e. stick), there must be at least one point where the contacting surfaces are stationary relative to each other. However, this does not mean that the whole contact area is in stick; slip could occur at other points on the surface. For s stick, the tangential force must be smaller than the limiting friction, thus equation 2.2 must apply.

$$|q(x, y)| \leq \mu |p(x, y)| \tag{2.2}$$

For surfaces in slip condition 2.2 is violated, then $q(x, y)$ and $p(x, y)$ are related by

$$|q(x, y)| = \mu |p(x, y)| \tag{2.3}$$

Additionally, the direction of traction must oppose the slip direction. Thus

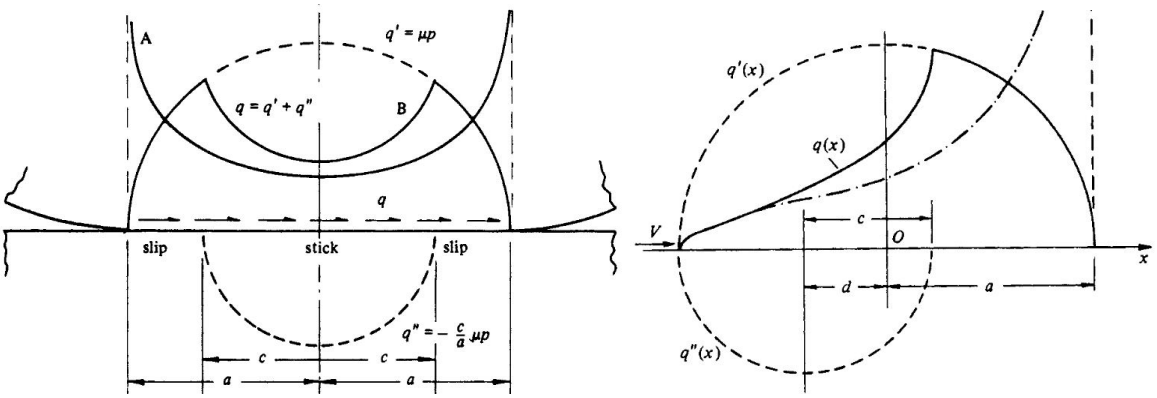
$$\frac{q(x, y)}{|q(x, y)|} = -\frac{\dot{s}(x, y)}{|\dot{s}(x, y)|} \tag{2.4}$$

where $\dot{s}(x, y)$ is the slip speed.

We now assume that the friction between the contacting surfaces is high enough to prevent sliding completely i.e. the complete contact patch is a 'stick' area. As depicted in figure 2.11a, we consider a strip through the centre of the contact patch at $y = 0$ where the surfaces are in contact from $-a \leq x \leq a$. If no sliding would occur over the whole contact area $q(x) \leq \mu p_0 (1 - x^2/a^2)^{(1/2)}$, where $p_0 = 2P/\pi a^2$. At the edge of the contact patch ($x = \pm a$), $q(x)$ goes to zero. This implies that the coefficient of friction at $x = \pm a$ would have to be infinite to remain in stick, which obviously cannot be. Therefore surface near the edge of the contact patch must be able to move relative to each other.

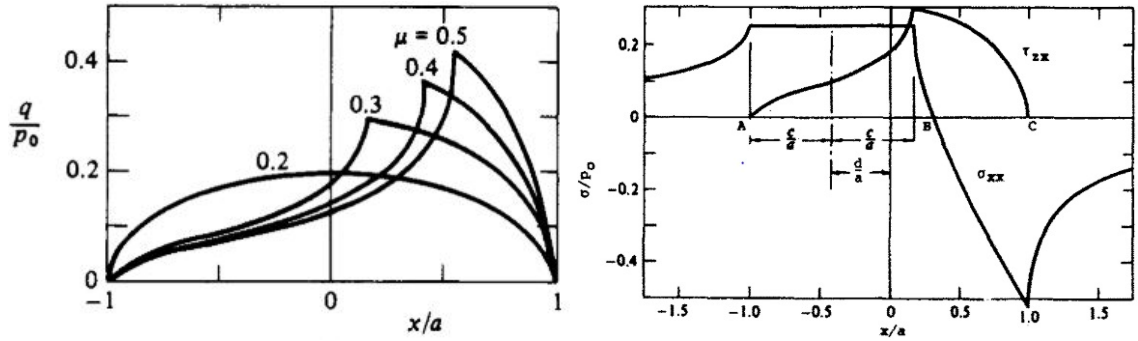
Johnson [45] considers another case where partial slip is allowed. If the tangential force Q is set to its limit μP, such that the surfaces are on the onset of sliding, then the tangential traction is given by

$$q'(x) \leq \mu p_0 (1 - \frac{x^2}{a^2})^{1/2} \tag{2.5}$$



(a) **(b)**

Figure 2.11: *Contact of cylinders with parallel axis (a) stationary contact [46], (b) tractive rolling [45]*



(a) **(b)**

Figure 2.12: *Surface stresses for a rolling line contact with tangential traction $Q_x = 0.2P$ (a) Tangential surface traction for varying μ [43], (b) Surfaces stress σ_x and τ_{zx} for μ = 0.3 [47].*

In this case, the surfaces are in slip except at $x = 0$. Johnson [45] now introduces an additional traction distribution

$$q''(x) = \frac{-c}{a}\mu p_0(1 - \frac{x^2}{c^2})^{1/2} \tag{2.6}$$

over a distance $-c \leq x \leq c$. Johnson [45] shows that by superimposing $q''(x)$ on $q'(x)$ that the resulting surface strains within the area $-c \leq x \leq c$ satisfy the conditions for stick, while in the regions $c \leq |x| \leq a$ the conditions for slip are satisfied. The size of the stick area can now be determined by following expression:

$$\frac{c}{a} = (1 - \frac{Q}{\mu P})^{1/2} \tag{2.7}$$

From this relation, it can be concluded that when P is kept constant, the application of a tangential load Q will lead to two slip areas beginning at the edges of the contact area, which will gradually progress towards the centre of the contact when Q increases.

The next step is to consider how the contact conditions change when the bodies in contact are rolling. Similar to a stationary contact, the contacting surfaces in a rolling contact deform as a result of the normal load distribution. This deformation changes the effective rolling circumference of both bodies. If a torque smaller than the maximum tractive force is applied to one of the bodies, it is found that the circumferential velocities of both bodies are no longer equal, without the occurrence of complete sliding. The difference in circumferential velocities divided by the rolling velocity is called creep. The ratio $(V_1 - V_2)/V_1$, where V_1 and V_2 are the velocities surfaces in contact, is defined as the creep ratio (ξ) [45]. Similar to a stationary contact, this results in the creation of a stick and slip area on the contacting surfaces. However, the position of these areas is different in rolling contact. If it is assumed that for a rolling contact the positions of the stick and slip areas are similar to a stationary contact, then there must be a slip zone at the leading edge of the contact

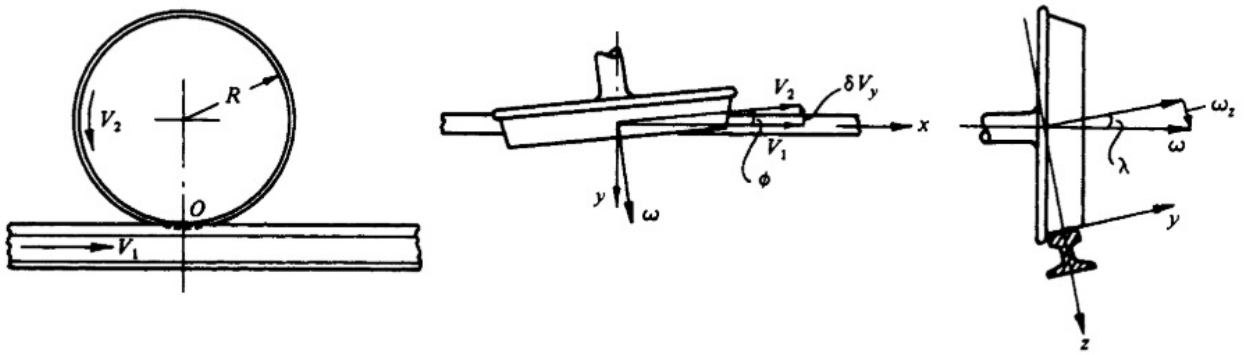


Figure 2.13: *Creep motion of railway wheel [45]*

patch. However, the direction of slip in this zone would be in the same direction as the tangential traction, which violates relation 2.4. This implies that the stick area must border the leading edge of the contact patch, and the slip area borders the trailing edge of the contact patch. The question remains where the boundary between the stick and slip areas is located.

To incorporate the stick area bordering the leading edge of the contact, Johnson [45] adapts the solution for a stationary contact by displacing the tangential traction distribution $q''(x)$ by a distance d such that it is adjacent to the leading edge, as depicted in figure 2.11b. This changes the $q''(x)$ to:

$$q''(x) = \frac{-c}{a}\mu p_0(1 - \frac{(x+d)^2}{c^2})^{1/2} \tag{2.8}$$

Johnson [45] shows that the conditions of stick are satisfied in the stick region and that the creep ratio and the size of the stick area are both related to the applied traction force Q, by the following equations:

$$\frac{d}{a} = 1 - (1 - \frac{Q_x}{\mu P})^{1/2} \tag{2.9}$$

$$\xi_x = \frac{-\mu a}{R}(1 - (1 - \frac{Q_x}{\mu P}))^{1/2} \tag{2.10}$$

From these expressions, it follows that in the absence of traction the complete contact area is in stick and there is no creep in the contact. If the Q is increased such that $d/a > 1$, the complete contact area is sliding. The friction coefficient also affects the size and shape of the stick area in the contact. This effect is illustrated in figure 2.12a. Not only does the size of the stick area increase for a larger friction coefficient, the peak of the shear stress cycle increases as well. Apart from the shear stress cycle, a surface element also experiences a cycle of tension and compression in longitudinal direction as a result of the applied traction. This is illustrated in figure 2.12b.

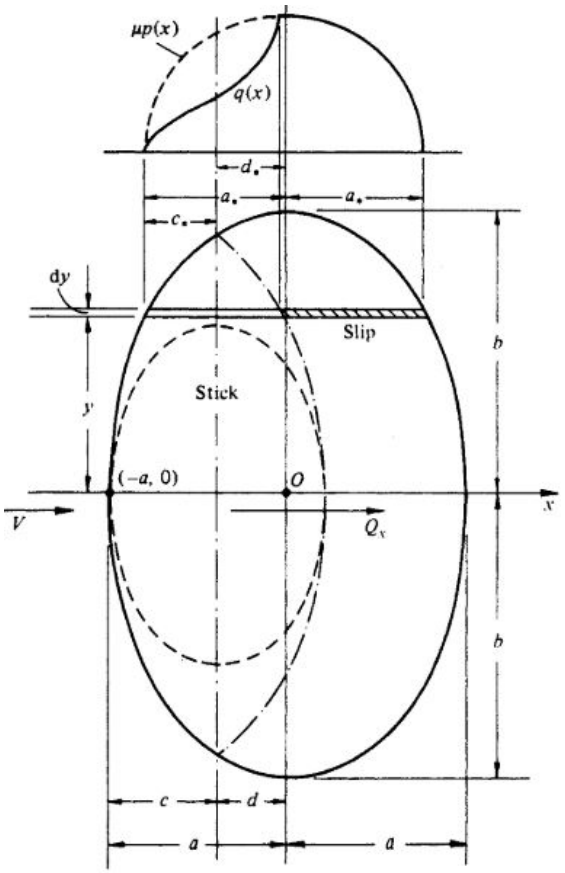


Figure 2.14: *Tractive rolling of an elliptical contact under an longitudinal force Q_x [45]. Broken line: elliptical stick zone [45], chain line: strip theory of Haines and Ollerton [48].*

So far all descriptions of the contact patch were made with the assumption of the contact of two parallel cylinders in two dimensions. However, the wheel-rail contact is a 3D phenomenon, where the contacting surfaces can be subjected to tangential forces in longitudinal and transverse directions, while also being subjected to an angular velocity about the normal axis called spin [45]. Therefore creep can have components in multiple directions. Creep is essential for the guidance of train wheels along rails. Johnson [45] distinguishes three types of creep that occur in wheel-rail contact (see figure 2.13):

a "*Longitudinal creep is caused either by wheels running on different radii as a result of wheel conicity, or by driving/braking forces.*"[45]

$$\xi_x = (V_1 - V_2)/V_1 \tag{2.11}$$

b "*Lateral creep arises if the wheel plane makes an angle ϕ with the rail axis.*"[45]

$$\xi_y = \delta V_y/V_1 = tan(\phi) \tag{2.12}$$

c "*Spin is the result of the difference between the wheel cone angle λ and the axis of rotation.*"[45]

$$\psi = \omega(ab)^{1/2}/V_1 = ((ab)^{1/2}/R)tan(\lambda) \tag{2.13}$$

ξ is the creep ratio, ψ is the spin parameter, V is velocity, a and b are the semi-axes of the contact ellipse, ϕ and λ are angles, R is the radius of the train wheel and where x,y,z are the longitudinal, transverse and vertical direction respectively.

All previous derivations were made with the assumption of a cylinder rolling on an elastic half-space. However, for a real wheel-rail contact patch this is not an accurate approximation, as the contacting surfaces have curvatures in more than one direction. For this three-dimensional case, Hertzian contact theory predicts an elliptical contact patch[42]. To find the distribution of stick and slip areas in an elliptical rolling contact a theory is proposed by Haines and Ollerton [48]. This theory divides the contact ellipse into longitudinal strips. It is assumed that the interaction between the strips can be ignored and each strip can be approximated by a cylinder on an elastic half-plane. Utilising that the creep ratio must be equal for the whole contact area, the boundary between the stick and slip areas can be calculated for each strip. An example is shown in figure 2.14. A symmetrical 'lemon-shaped' stick area borders the leading edge, while the rest of the contact is in slip.

2.3 Effects of wheel-rail contact on rail material

In sections 2.1 and 2.2, the reader was introduced to the properties of rail material and the stresses caused by wheel-rail contact loads. In this section, the effects of the wheel-rail contact on the rail material will be discussed. The effects can be distinguished: plastic deformation, wear and thermomechanical loading. These three aspects all influence the initiation and propagation of rolling contact fatigue (RCF), which is discussed in more detail in section 2.4.

2.3.1 Plastic deformation

In section 2.2.5 the assumptions of the Hertzian contact theory were discussed. Up to this point, the material response is assumed to be perfectly elastic. However, no real material is perfectly elastic. Plastic deformation will occur when the maximum stress first exceeds the material's yield stress (Y in simple tension, k in simple shear), i.e. the onset of yield. Plastic deformation has two effects. (1) Residual stresses will be created in the deformed area. These partly counteract the stresses generated in subsequent cycles, limiting further plastic deformation. (2) Strain-hardening raises the yield strength of the material, which also limits further plastic deformation [50]. During their lifetime, rails experience millions of stress cycle. During the first cycles, some plastic deformation will occur in the rail. After a certain number of cycles with plastic deformation, the residual stresses and strain-hardening may raise the yield stress of the rail material, such that all subsequent loading cycles result in an elastic material response. This process is called elastic shakedown [43].

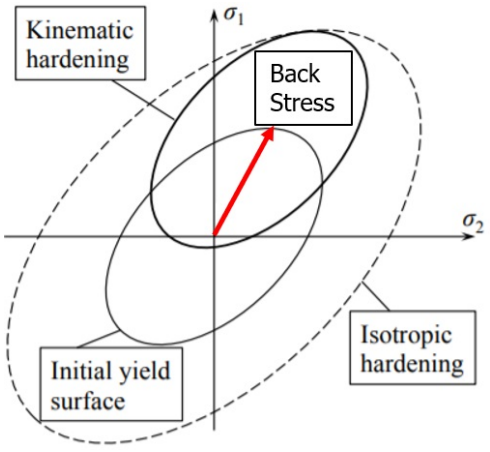


Figure 2.15: *Isotropic and linear kinematic hardening models in principal stress space. Adapted from [49].*

Strain hardening and residual stresses cannot raise the yield strength indefinitely. Beyond a certain threshold, plastic deformation will occur each loading cycle. This threshold is called the elastic shakedown limit. The plastic deformation of a material is governed by the hardening behaviour. The two simplest hardening models are isotropic and linear kinematic hardening [51], as illustrated in figure 2.16a and b. Isotropic hardening assumes that yield stress is raised to the maximum applied stress and all further loads below that stress are in

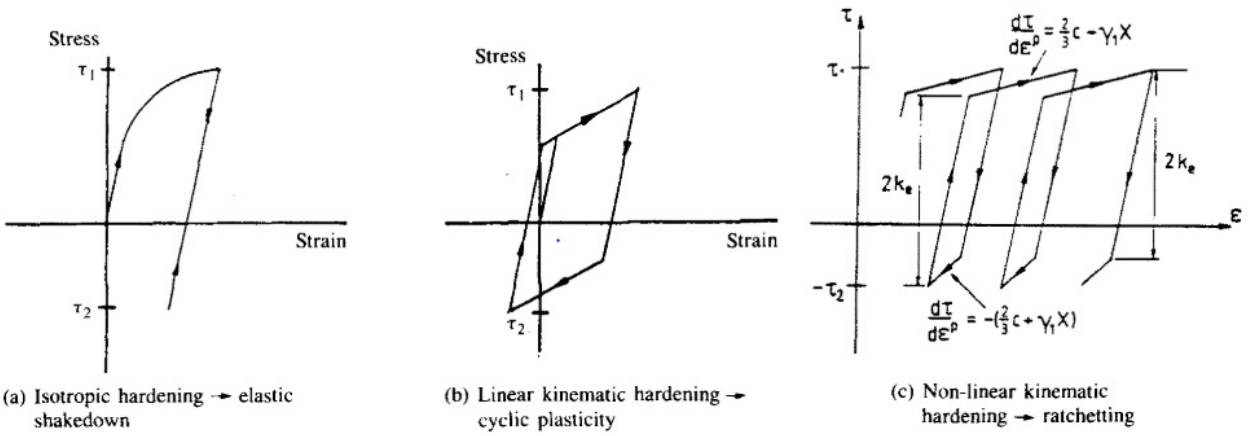


Figure 2.16: *Material response for different hardening models, (a) isotropic, (b) linear kinematic, (c) non-linear kinematic. Adapted from [51] and [52].*

the elastic regime, therefore shaking down in a quarter stress cycle. Linear kinematic hardening assumes a linear relationship between stress and strain in the plastic regime. If the yield stress is raised in the positive loading direction, a linear kinematic hardening material loses strength in the negative loading direction such that the elastic range stays constant at a span twice the yield stress. Therefore it will start yielding earlier if the loading direction is reversed. If the amplitude of a cyclic load is larger than the yield stress, plastic deformation will occur every cycle. The behaviour of isotropic and linear kinematic hardening can also be represented in principal stress space, as shown in figure 2.15. For isotropic hardening, the yield surface increases in size when the stress exceeds the yield stress, but its origin is constant. For kinematic hardening, the yield surface does not change in size once the yield stress is exceeded, but translates in space. The components of the translated yield surface origin is called the back stress [52].

It is observed that the material at the rail surface does accumulate strains in the direction opposite to the rolling direction on tracks with unidirectional traffic. The behaviour of accumulating unidirectional strain during each cycle is called ratcheting [51]. Neither isotropic nor linear kinematic hardening models can predict strain accumulation due to cyclic loading; an isotropic hardening material does not deform plastically after the first cycle, a linear kinematic material shakes down to a closed loop with cyclic plastic deformation [51]. Ratcheting cannot be modelled by either of the two simple hardening laws. To model ratcheting Bower [52] proposes a non-linear kinematic hardening law. The hardening law gives good predictions for the ratcheting rates of rail material when compared to experimental results. It does not reproduce the shape of the stress-strain loop in detail, but this is not critical for the investigation of damages after a large number of cycles. An example of this hardening law is presented in figure 2.16c. A material is subjected to an unsymmetrical cycle of pure shear stress, where the mean shear stress is positive. As a result of the positive mean shear stress, a positive back stress X is created. Whereas linear kinematic behaviour assumes a constant value for the slope of the stress-strain curve $d\tau/d\epsilon^p$, non-linear kinematic hardening assumes that this slope is dependant on the back stress X with the relation

$$\frac{d\tau}{d\epsilon^p} = \frac{2}{3}c - \gamma_1 X \tag{2.14}$$

for a positive strain increment and

$$\frac{d\tau}{d\epsilon^p} = -(\frac{2}{3}c - \gamma_1 X) \tag{2.15}$$

for a negative strain increment, where c is the hardening rate and γ_1 is the ratcheting rate. Due to the positive mean shear stress the back stress is also positive. Therefore the slope $d\tau/d\epsilon^p$ for a positive strain increment is less steep than for a negative strain increment, resulting in the accumulation of strain in the direction of the mean stress for each cycle [52].

As mentioned in section 2.2.5, the maximum stress for free-rolling is located subsurface. However, if a traction force is applied to the bodies in contact, the resulting stresses generated by this traction force shift the point of maximum stress. If the ratio between the normal force and the traction force Q/P exceeds 0.3, the maximum stress, and therefore the onset of yield, will be located at the surface [47][51].

Subsurface flow

For a free rolling wheel the stresses in x-, y- and z-direction are compressive, as shown in figure 2.9. First yield will occur in the point of maximum principal shear stress. If the wheel-rail contact is approximated by a cylinder rolling on a half-space, the maximum principal shear stress τ_1 will occur in the center of contact at a depth of $0.78a$, where τ_1 is given by Johnson [43]:

$$\tau_1 = \frac{1}{4}(\sigma_{xx} + (\sigma_{xx})_r - \sigma_{zz})^2 + \tau_{zx}^2 \leq k^2 \tag{2.16}$$

where subscript r indicates residual stress.

If a normal load is applied that exceeds the elastic limit first yield occurs at C in figure 2.10a. As a result of this load, the material element wants to expand parallel to the surface. However, this expansion is hindered by other elements at the same depth, who experience the same load. This creates residual stresses $(\sigma_{xx})_r$ such that $(\sigma_{xx})_r = \sigma_{zz} - \sigma_{xx}$ and hinders further plastic deformation during subsequent cycles. However, it follows from equation 2.16 that the effect of the reversed shear cycle at points B and D in figure 2.10a cannot be compensated by the creation of residual stresses [43]. Therefore it is the cycle of alternating orthogonal shear stress τ_{zx} that is critical for elastic shakedown. For free rolling the elastic shakedown limit is reached at a normal load of 1.66 times the elastic limit [43]. If the normal load is increased even further, plastic deformation will occur every loading cycle. The shear stress cycle experienced by the subsurface element is symmetric, thus back stress X is equal to zero. Therefore, the non-linear kinematic hardening behaviour reduces to linear kinematic behaviour, hence no unidirectional strain will be accumulated [53].

Surface flow

If the friction coefficient is larger than 0.3, the maximum stress in the surface exceeds the subsurface stress maximum. Therefore yield will occur at the surface first. Contrary to a subsurface element, a surface element does not experience a symmetrical cycle of orthogonal shear stress τ_{zx}. Additionally a surface element experiences a cycle of tension and compression in longitudinal direction σ_{xx}, as presented in figure 2.12b. The combined stress cycles in σ_{xx} and τ_{zx} are critical for the elastic shakedown limit. This is illustrated in figure 2.17, where the trajectory of deviatoric stresses are plotted in stress space. Points A, B and C correspond to the equally named points in figure 2.12b. The yield surface (according to the Tresca yield criterion) is indicated with the chain line. If non-linear kinematic hardening behaviour is assumed, this yield surface is free to translate in stress space and has a constant radius of k. The elastic shakedown limit is reached if the stress trajectory can no longer be contained within the circle of the yield surface [47]. If the elastic shakedown limit is exceeded, plastic deformation will occur each cycle. Due to the unsymmetrical cycle of shear stress, ratcheting will occur. At a critical value of unidirectional strain, ratcheting can cause damage to a material [50]. This will be discussed in section 2.4.1.

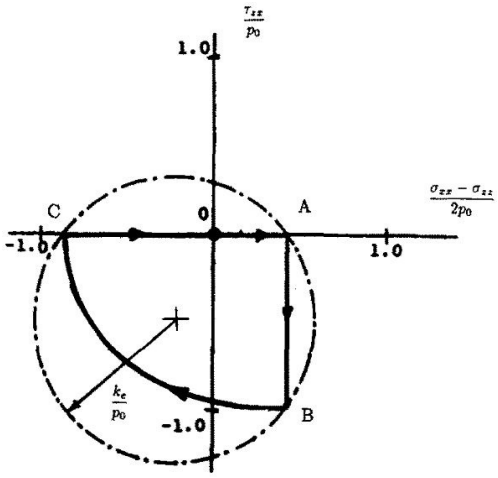


Figure 2.17: *Trajectory in stress space of surface stresses in a line contact for partial slip conditions. The yield surface at shakedown is shown as a chain line [47].*

2.3.2 Wear

Wear is defined as "*the progressive loss of substance from the operating surface of a body occurring as a result of relative motion at the surface.*" *[54]* Depending on the conditions, wear is caused by various mechanisms, such as adhesion, abrasion, oxidation, delamination, etc [55]. Archard and Hirst [56] studied adhesive wear and made a classification into two regimes: mild and severe. They observed that, after an initial running-in period characterized by severe wear, the amount of material lost to wear was dependent on the applied load and the sliding distance as given by

$$W = \frac{KsP}{p_m} \tag{2.17}$$

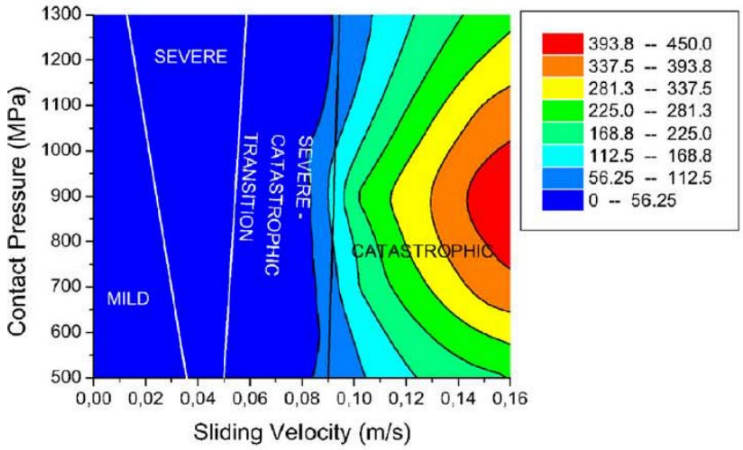


Figure 2.18: *Wear map of eutectoid rail steel [57]*

where W is the worn volume, s is the sliding distance, P is the applied load, p_m is the flow pressure i.e. hardness of the softer material and K is the wear coefficient; a dimensionless constant related to the probability of a wear particle being created. Note that this relation does not include any effects of sliding speed. Later work by Lim and Ashby [55] expanded Archards wear law by plotting the wear rate as a function of the contact pressure and the sliding velocity in wear maps. These wear maps help in predicting the wear regime for a given combination of contact conditions. Lewis and Olofsson [57] applied the approach of Lim and Ashby to wheel and rail material, but instead of plotting the wear rate they used Archards wear coefficient K. The wear map for standard rail steel is shown in figure 2.18.

Although Archards theory is consistent with experimental results, Suh [58] has the criticism that "*It completely ignores the physics and physical metallurgy of metal deformation; many of the assumptions employed in the mathematical derivation are unreasonable and arbitrary; the theory does not provide any insight to the wear of metals under different sliding conditions*"[58]. In his delamination wear theory, Suh [58] explains wear based on dislocation theory, plastic deformation and the fracture of metals. His theory states that the surface layer of the material has a reduced dislocation density due to the image force; a force that arises due to the existence of a stress-free surface. Therefore, after continued sliding, dislocation pile-ups create voids beneath the surface. These voids coalesce to form cracks parallel to the surface which, after they reach a critical length, grow to the surface and detach a piece of material. This process repeats and produces plate-like flakes to detach periodically. Wear test from Tyfour et al. [59] found such plate-like wear particles as described by Suh's delamination wear theory. However, the question remains if the particles detach as plates, or if they are rolled to plates by subsequent rolling cycles.

2.3.3 Thermomechanical loading and the formation of white (and brown) etching layer

As discussed in section 2.2.5, there exist a slip area contact patch as a result of the applied traction. Due to the friction between the sliding surfaces frictional heat is generated, which dissipates into the contacting bodies. This process subjects the material at the surface to a fast thermal cycle, which is believed to be able to alter the microstructure and create a so-called white etching layer (WEL) on the surface. WEL is often linked to the initiation of RCF [60][61][62]. This relation is discussed in section 2.4.1.

Since measurement of temperature during real wheel-rail contact is difficult, if not impossible, temperature calculations are important in the analysis of the rolling contact [60]. In literature analytical [60][63][64] and numerical [61][62][65] methods are presented to calculate the temperature in a wheel-rail contact. The temperature of a rail element starts to increase from the first point in the contact where slip occurs. The maximum temperature, also known as the flash temperature, is reached at the surface close to the trailing edge of the contact [60][61]. Since the frictional heat source moves over the rail surface with a high velocity, the temperature is only affected for a thin surface layer called the thermal penetration depth. This depth is calculated to be 75 to 80 μm [61][60][64]. Knothe and Liebelt [64] give the following expression for the flash temperature in an elliptical contact where complete sliding occurs between smooth surfaces

$$T_{max} = 1.23\sqrt{\frac{a\kappa}{v}}\frac{\alpha\mu v_s p_{max}}{\lambda} \tag{2.18}$$

where T_{max} is maximum surface temperature, a is the contact half-width, κ is thermal diffusivity, v is the travelling speed, α is the ratio between the heat conducted into the rail and the total frictional heat, v_s is the sliding velocity, p_{max} is the maximum normal pressure and λ is the thermal conductivity. The flash temperature is related to the sliding velocity and therefore also to the creep ratio. Naeimi et. al. [61] calculated the flash temperature as a function of the longitudinal creepage using a 3D finite element model. They assumed a friction coefficient of 0.6 and that the material properties are not dependant on the temperature. They found a temperature increase of 311 °C for a creepage of 1.2%, which is the maximum creepage possible without complete sliding. This temperature increase is not sufficient to cause any material transformation. Their approximation by linear extrapolation indicates that temperatures sufficient for material transformation are reached at creepages of 2.5% or higher. Therefore it is unlikely that such temperatures will be reached during normal train operation [64].

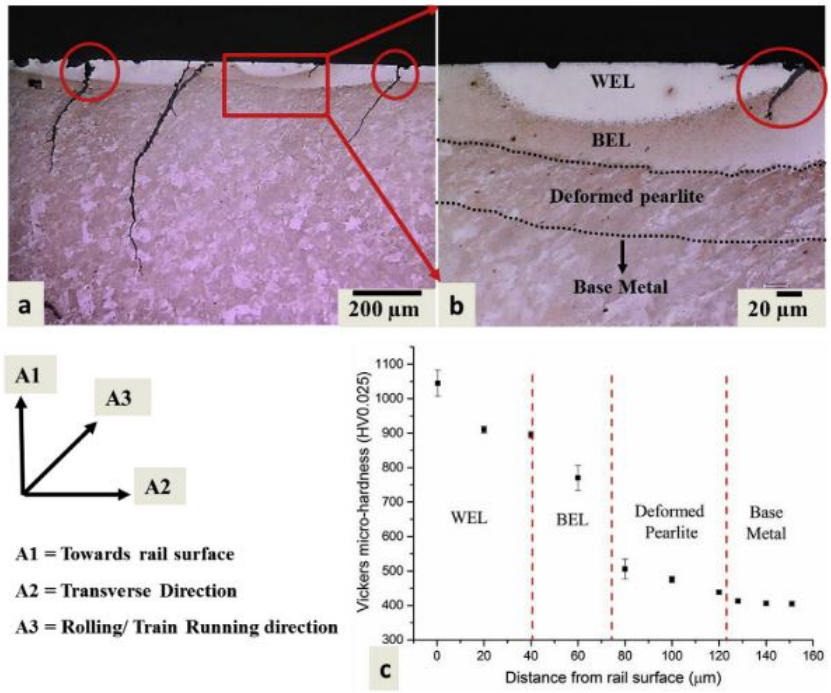


Figure 2.19: *(a,b) Microstructure of WEL and BEL at rail surface, (c) micro-hardness profile perpendicular to surface [66]*

If, for example, due to sudden wheel slip, temperatures are reached above the austenitization temperature, phase transformations can occur. Due to the speed of the train, the thermal cycle caused by the wheel only lasts a millisecond [66] The high cooling rate can cause martensite formation. Phase changes due to thermal loading are believed to play a major role in the formation of the so-called white etching layer (WEL), which is often linked to the initiation of RCF cracks [67][60][62]. WEL is a hard and brittle material that forms at the contact surface during the service life of the rail [6]p15. It owes its name to its white appearance under a microscope after it is etched with Nital [6]p15. WEL typically has a thickness of 30-100 μm [68] and a hardness of 650-1200 HV, which is two or three times higher than the bulk hardness [6]p15,[68]. The formation mechanism of WEL remains a point of debate in literature [66][6]p16. Two possible formation mechanisms are proposed: (1) the formation of martensite due to a rapid heating and cooling cycle caused by frictional heat from wheel-rail contact, (2) dissolution of cementite and grain refinement leading to a carbon-supersaturated nanocrystalline ferrite caused by severe plastic deformation [66][6]p16. Recent characterization studies by Kumar et al. [66] and Wu et al. [6]p62 conclude that the first mechanism is responsible for WEL due to the presence of martensite and a larger fraction of retained austenite than in the bulk material.

Another layer with an altered microstructure which is often linked to the effects of frictional thermal cycles due to wheel-rail contact is brown etching layer (BEL) [69][66][70]. Similar to WEL, BEL also owes its name to its colour observed in optical microscopy when etched with Nital [69]. BEL is most commonly found beneath WEL or it is found as a single layer at the surface [69]. BEL also shows an increased hardness relative to the bulk hardness. Average values between 700 and 800 HV are reported [66][70]. Li et al. [69] measured hardness profile over a BEL with a thickness of approximately 100 μm and found decreased hardness of 540

HV at the interface between the WEL and BEL, whereafter the hardness increased with depth to a maximum of 740 HV. They conclude that BEL is formed through a martensitic transformation after rapid heating and cooling cycles caused by frictional contact heating. The decrease in hardness after the WEL/BEL interface is contributed to the tempering of the formed martensite by continued heating cycles. Kumar et al. [66] present two theories of BEL formation. The first formation mechanism occurs when the peak temperature remains below the A_1 temperature; severe plastic deformation leads to the partial dissolution of the cementite laths. Subsequent heating cycles just above the A_1 temperature forms austenite, which upon cooling is transformed into martensite or is retained as austenite. This leads to a microstructure containing supersaturated ferrite, partially dissolved cementite and martensite/retained austenite islands. The second formation mechanism occurs when the peak temperature does exceed the A_1 temperature from the start. This leads to the formation of a larger amount of austenite. Upon cooling, the austenite also transforms into martensite. Subsequent heating cycles temper the martensite, which leads to carbide precipitation.

2.4 Rolling contact fatigue (RCF) defects in rails

The service life of a rail is controlled by either wear or RCF [71]. Wear is the preferred deterioration mechanism, as it is visible and predictable. In contrast, RCF cracks are unpredictable, difficult to detect and can become a safety hazard if left to propagate unattended [72]. The development of RCF cracks is often divided into three phases: initiation, growth due to contact stress and growth due to other stresses [38][3]. For a safe operation of railways and the planning of maintenance, understanding of initiation and propagation of cracks and the development of RCF defects is required. In this section crack initiation sites, crack growth mechanisms and a common type of RCF defect are discussed. To conclude this section, a short overview of rail grinding maintenance and its effect on the rail surface is presented.

2.4.1 Crack initiation sites

As mentioned in section 2.3.1, contact stresses cause repeated plastic deformation and/or ratcheting in the surface layer of the rail [73][50]. The rail material contains some inhomogeneities which can cause stress concentrations and can act as "weak sites" in the deformed microstructure. Crack initiation will be facilitated by these weak sites [10]p8,[6]p14.

Inclusions

Brittle inclusions, like alumina and silica particles, can greatly reduce the fatigue life [5]. They cause stress concentrations in the matrix, which lead to early fatigue crack initiation [6]p15. However, due to improvements in rail steel making, failure caused by brittle inclusions is rare [5][74]. Ductile inclusions, like the MnS inclusions found in rail steel, do not have such a detrimental effect on fatigue life as brittle inclusions [74]. MnS inclusions typically have a hardness half that of the steel matrix [75]. In the heavily sheared rail surface layer, MnS inclusions become strain-flattened and align along the longitudinal rail axis. This can lead to delamination along the incoherent inclusion boundary [74] or to the initiation of micro-cracks caused by a stress concentration at the strain-sharpened inclusion tip [75].

Pro-eutectoid ferrite - pearlite interface

As discussed in section 2.1.1, hypo-eutectoid steels contain a fraction of PE ferrite at the prior austenite grain boundaries. Multiple studies [71][72][76] have found that fatigue cracks initiate at and grow along the heavily deformed boundaries of PE ferrite and pearlite grains in the surface layer of the rail. Eden et al. [71] performed nano-hardness measurements on deformed and undeformed PE ferrite and pearlite grains in rail steel with approximately 20 volume% PE ferrite. When comparing the increase in hardness between the undeformed and deformed material, they found that the PE ferrite showed a larger average increase in hardness (33%) than the pearlite (8%). It was concluded that strain-partitioning occurred between the PE ferrite and pearlite, causing the PE ferrite to strain-harden at a higher rate than the pearlite. This led the PE ferrite to exhaust its ductility earlier than the pearlite and enabled cracks to initiate.

Garnham and Davis [72][76] performed experiments on the rolling contact fatigue behaviour of rail steels with lower volume percentages (11%, 5.8% and 2.0%) PE ferrite. They found that a decrease in the PE ferrite content caused an increase in the life to RCF crack initiation. For the lowest PE ferrite concentration, they found that crack initiation at the strained PE ferrite-pearlite boundaries was not the dominant initiation mechanism.

White etching layer (WEL)

WEL's are often mentioned in literature as initiation sites for RCF cracks due to their hard and brittle nature [77][78][79][80]. Recently two studies by Kumar et al. [81] and Saxena et al. [82] measured the fracture toughness of WEL on R350HT grade steel to be 21-25 MPa $m^{1/2}$ and 16.0 $\pm$ 1.2 MPa $m^{1/2}$ respectively, which are comparable to martensite. The minimum bulk fracture toughness of R350HT steel is 30 MPa $m^{1/2}$ [28], thus WEL is significantly more brittle than the bulk material. Although no measurements of the fracture toughness of WEL in R260Mn steel exist, a similarly increased brittleness is expected. Saxena et al. [82] estimate a critical defect size in WEL of 5-10 μm, which is significantly smaller (approximately one order of magnitude) than the typical layer thickness. Therefore cracks are easily nucleated from small inclusions or surface irregularities, such as embedded oxide particles [79].

Carroll and Beynon [83] performed an extensive study into the crack morphology of RCF cracks in WEL using a SUROS twin disk machine [84]. They simulated WEL in the laboratory by two techniques: spot-welding and gross sliding two disks over each other using the SUROS machine. The latter technique formed a continuous WEL over the disk. After the test disks were subjected to varying loading cycles, cross-sections of the surface showed two types of crack morphologies, labelled "single cracks" and "gaps" (of 5 to 30 μm wide). Single cracks were observed perpendicular or under an angle to the surface. Gaps were found in different morphologies depending on their formation mechanism. Gaps can form because of the continued wear of an angled single crack, spalling of a piece of white layer between two closely spaced single cracks, plastic deformation of the WEL or the accumulation of wear debris in a crack. Cracks in WEL that continue to grow into the pearlite matrix follow the flow direction of the plastically deformed microstructure.

WEL on rails is not continuous, but consist of separate islands/patches of transformed material [77][83][79]. Apart from transverse crack initiation in WEL, crack initiation is also observed at the surface on the boundary between a WEL patch and the untransformed material [77]. On the disks tested by Carroll and Beynon [83] where the WEL was simulated by spot welding, cracks were also found on the boundary between the WEL and pearlite matrix. They attribute the formation of these cracks to the exhaustion of ductility of the pearlite due to a difference in the amount of plastic deformation accumulated during testing.

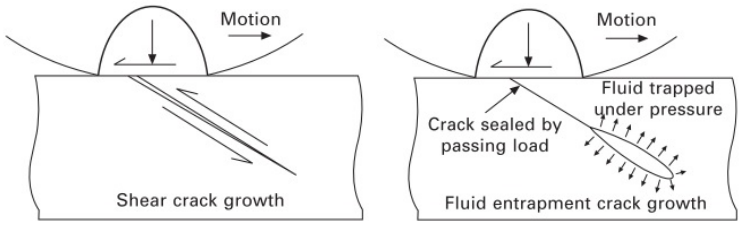
Brown etching layer (BEL)

BEL is also linked to the initiation of RCF cracks. Li et al. [69] observed cracks initiating from the edge of a BEL patch at the interface between BEL and the pearlite matrix, which grew to a depth of approximately 300 μm. The crack grew under an angle to the traffic direction. The angle varied with depth. Other cracks that they observed were transverse cracks through the BEL, oriented perpendicular to the traffic direction. Most were contained within the WEL and BEL, but one was found propagating into the pearlite below. At the BEL-pearlite interface, this crack changed direction and followed the strain direction of the pearlite.

Messaadi and Steenbergen [70] observed two similar types of cracks related to BEL: transverse fracture and initiation at the interface of WEL/BEL and pearlite. Some of the transverse fractures were contained only in the BEL and did not propagate into the pearlite below, nor the WEL above. Cracks were also found to simultaneously form in both the WEL and BEL, which, by subsequent growth, merged into one crack. In contrast to the findings of Li et al. [69], cracks initiated at the interface of WEL/BEL and pearlite were found not to propagate into the pearlite matrix. Instead, they propagated along the WEL-BEL boundary.

2.4.2 Crack propagation

The cracks transition from the first phase to the second phase when they reach a length in the order of tens of μm up to a hundred μm [3]. Wear slows down crack growth in the first phase, as the removal of material from the surface decreases the crack length or removes the crack entirely [38][3]. In the second stage, additional crack growth mechanisms start contributing to crack propagation, which are primarily activated by the contact stresses [38][3]. After this transition cracks usually propagate oriented along the sheared microstructure, under an angle of 10-20 degrees with the surface [38]. Water is believed to be important for crack growth in this stage because defects on rails rarely occur without the presence of water (for example in tunnels) [38][3][85][86]. Bower [87] modelled the influence of fluid on crack growth due to contact stresses and suggested two crack growth mechanisms enabled by the presence of fluid in the crack: crack face lubrication (fig 2.20a) and fluid entrapment fig 2.20b). Crack face lubrication allows shear stresses to reach the crack tip, which would otherwise be restricted by the friction between the crack faces. The shear stresses load the crack tip in mode II. Fluid entrapment is caused by the closing of the crack mouth due to compressive contact forces. The compressive force pressurizes the trapped liquid in the crack. The pressurized fluid pushes against



(a) **(b)**

Figure 2.20: *Fluid assisted crack growth mechanisms: (a) shear crack growth enabled fluid-assisted crack face lubrication, (b) crack pressurization through fluid entrapment [3].*

the crack faces and loads the crack tip in mode I [88]. In addition to the contact stresses, other sources of stress (such as thermal stress, residual stress and bending stress) can also cause mode II loading at the crack tip. However, the contact stresses primarily govern crack propagation [3].

Phase two cracks propagate up to a critical length in the order of millimetres [38][3]. This point marks the start of phase three. The crack changes its growth direction and/or branches. Depending on the stress state in the rail the crack propagates either almost vertically down into the rail, or continues to grow parallel to the rail surface [38]. The first case will ultimately lead to a transverse fracture, the second case to a spallation [38].

2.4.3 Squats

RCF cracks can occur at different locations on the surface of the railhead. Depending on the type of loading and the location of initiation, cracks can develop into different types of defects. The two primary types of defects at the railhead are squats and head-checks (or gauge corner cracking). In this section, only the squat-type defects will be discussed. Descriptions of head-checks can be found in [3][38][89].

Although some debate exists in literature on its exact definition, a squat is commonly characterised by three visual features; (1) a local depression with a dark surface in the shape of two lungs/kidneys, similar to the imprint left behind by someone sitting or squatting on the rail surface (hence, the name squat[90]); (2) a V- or U-shaped crack in the running surface, with its opening in the direction of the gauge corner of the rail; and (3) a local widening of the running band around the dark spots [77][78][85][91]. If squats subjected to repeated critical loads are left to propagate without intervention, they develop cracks that propagate down into the rail which can cause brittle transverse rail fracture [92], resulting in possible derailment. An example of a developed squat is shown in figure 2.21a. Squats do not possess these three previously mentioned visual characteristics at the moment of initiation but develop these features due to repeated wheel-rail contact loading cycles. In literature, different theories exist on the initiation and development of squats.

According to Li et al. [91][92], the formation of squats is related to short wave irregularities in the rail surface, such as indentations, welds and corrugations. Li et al. [92] developed a finite element model to simulate the contact forces resulting from a train wheel rolling over singular depression (similar to the left picture in 2.22) and developed squat. They found that the surface depressions caused an oscillation in the contact force which resulted in multiple maxima of approximately 1.8 times the static load. Due to these peak loads stresses are created that exceed the yield stress of the rail material. The distance between contact force maxima coincided well with the characteristic lengths of the squat, as seen in figure 2.21b. Li et al. [91][92] conclude that this oscillation in the contact force, which is the consequence of the eigencharacteristics of the vehicle-track system, is responsible for squat development.

Deng et al. [85] performed an extensive field monitoring of squat development. The development of one of the monitored squats is shown in figure 2.22. They observed that a squat first develops a small, darkened depression in the running surface of the rail. This depression grows with subsequent loading cycles and alters the shape of the contact patch. This initiates a crack at the gauge side of the depression once the depression has reached a critical size. This crack propagates under a shallow angle into the rail and grows into a U-shape under influence of traction shear force. Deng et al. [85] contribute the formation of the black depression to the collapse of a network of micro-cracks as a result of plastic deformation and strain-hardening. The localised deformation and hardening is believed to be caused by a local oscillation in contact force as a result of surface irregularities, as described by Li et al. [91][92].

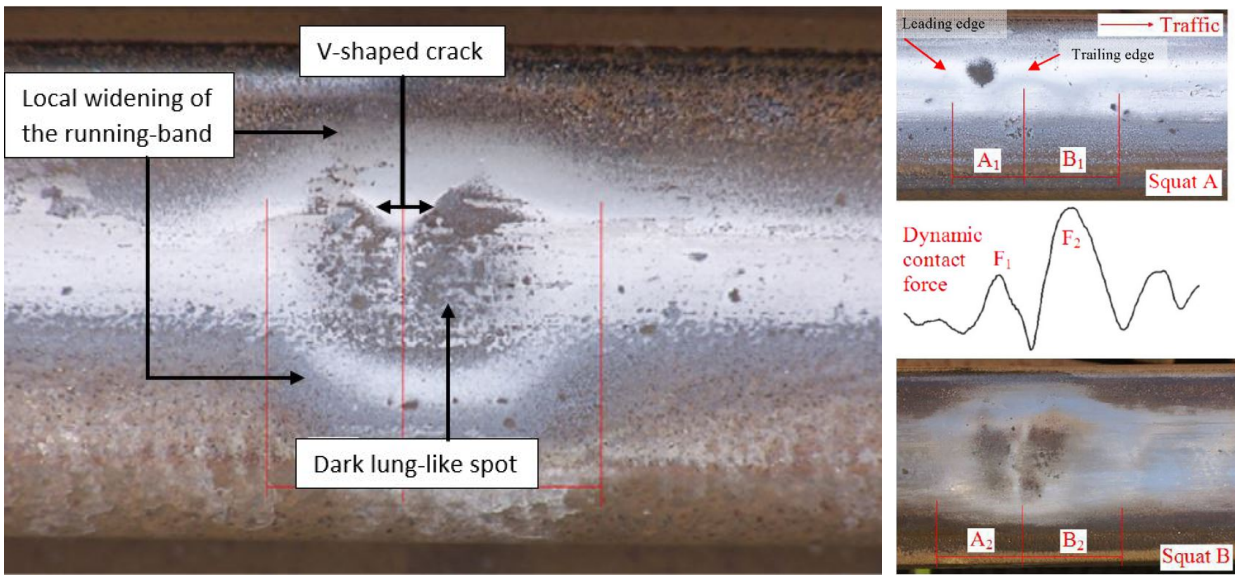


Figure 2.21: *(a) Visual appearance of a squat on the running band. Adapted from [92] (b) Contact force maxima coinciding with squat geometry [91]*

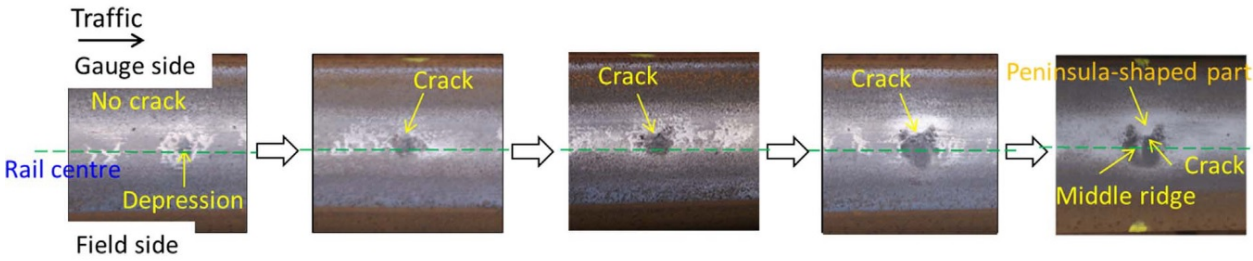


Figure 2.22: *The growth of a squat defect, adapted from [85]. Photos are taken every six months.*

Steenbergen and Dollevoet [77][93] take a more metallurgical approach in their two-part theory on squat initiation and growth. In their paper on initiation [77] they theorise that squats initiate from micro-cracks at the edges of WEL patches on the running surface. The cracks grow as a result of τ_{zy} and τ_{zx} in the direction indicated in figure 2.23a. Multiple micro-cracks in the surface link up and form a large crack as indicated in figure 2.23b. The crack mouth that is visible on the running surface is oriented perpendicular to the principal tensile direction. This crack propagates down into the rail in the direction of the traffic along the texture of the ratcheted surface layer. This crack is called the leading crack as indicated in figure 2.23c. The presence of the leading crack causes a redistribution of the tangential stress τ_{zy}, which can cause a trailing crack to branch off from the leading crack. This trailing crack does not follow the sheared texture of the microstructure but propagates by a brittle failure mechanism in the direction opposite to the traffic. However, not all squats develop a trailing crack. When they are present, they lag in growth relative to the leading crack. The leading crack is therefore the most critical.

In their second paper on squat growth, Steenbergen and Dollevoet [93] observe that the cracks propagate into the railhead under a slight angle until they reach a typical depth of 2-4 mm, from where they continue to propagate parallel to the surface. This depth coincides with the depth where the residual axial stress is expected to change from a compressive to tensile state. The compressive axial stresses in the top layer of the rail are caused by cycles of wheel-rail contact during train operation. The development of cracks reduces the load-bearing capability of the material above the cracks, which causes a vertical settlement of the surface. This vertical settlement (a dent) changes the curvature of the rail surface and therefore causes a redistribution of the load-bearing surface. The settled surface no longer makes contact with the wheels. Therefore corrosive areas can build up on the surface, which causes a dark spot to appear on the surface.

Simon et al. [78] performed a tribological and metallurgical analysis of rails that had been in service. A wave-like pattern on the field side of the running band is observed. The running surface itself is covered with

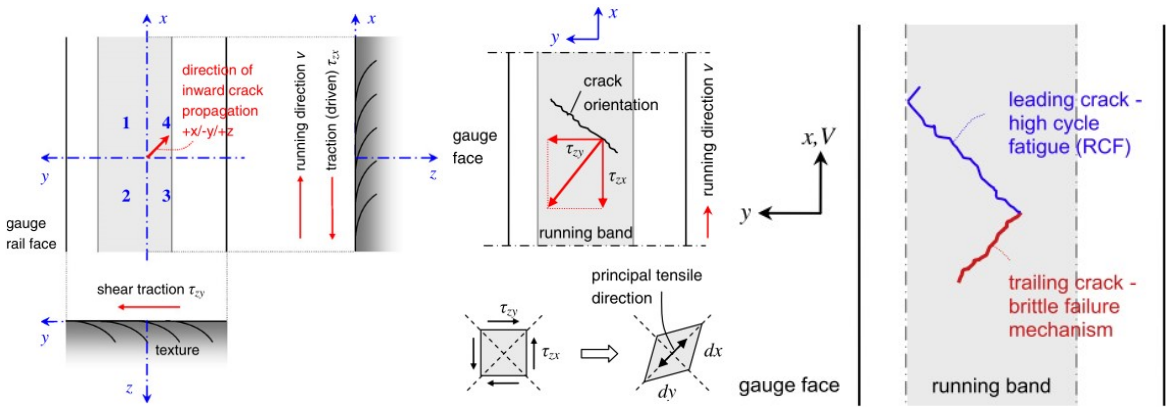


Figure 2.23: *Steenbergen and Dollevoet's [77] theory on squat formation. (a) Crack propagation direction as result of material texture, (b) orientation of cracks in the running surface, (c) differences in propagation mechanisms of the branches.*

a thin WEL, areas, where no wheel-rail contact occurs, are not covered with WEL. Transverse and longitudinal cross-sections of the microstructure beneath the rail surface reveal a textured layer. The microstructure of the surface layer besides the running band is strained in the direction of traffic, but the surface layer of the running band itself is strained opposite the traffic direction. It is theorised that after a certain number of cycles, the mechanical properties between these areas with opposite strain direction become incompatible. Therefore, the border between these areas becomes a weak site where cracks can initiate. Fracture surface analysis revealed a fatigue crack that initiated at the border of the WEL covered and non-WEL covered surface near the field side of the rail.

In his work, Grassie [94] theorises that, besides regular squats, there exists another distinct squat-type defect that has the same three superficial characteristics as a squat, which he calls a "stud". He defines a squat as follows: "*Squats are a surface-initiated RCF defect. These are initiated at the gauge-corner edge of the contact band and are oriented under an angle to the running direction. The crack is initiated in the heavily sheared and plastically deformed surface layer of the rail. The initial crack propagates as a result primarily of fluid entrapment and an angle of about 20°to the rail surface, roughly normal to the crack mouth on the surface, and in the direction of traffic. This crack propagates further into the rail and in most cases develops into a transverse defect*"[90]. Studs differ from squats in several ways. First, studs also develop in areas where there is little unidirectional strain near the surface [94]. Second, studs are always found to have a WEL at the surface near the cracks [94]. Third, studs are found to develop within 10 million gross tonnes (MGT) of traffic, whereas squats develop after 40 to 100 MGT [94]. Finally, studs have never been found to cause transverse fractures. Instead, they continue to propagate parallel to the surface after they have reached a depth of 3-6 mm [86]. Grassie contributes the initiation of studs to the effects of creepage induced thermal damage, due to the strong relation to the presence of WEL [94][86].

These different theories and explanations indicate that squats are not yet fully understood. Squat initiation and propagation seem to be influenced by a wide variety of variables. Different approaches in the analysis method yield different results and conclusions, each with its strengths and shortcomings. Li et al. [91][92] and Deng et al. [85] provide a good description of the growth of a squat once initiated, but lack an explanation for the initiation itself. They also do not consider the effects of subsurface crack propagation to a large extend. The distinction between squats and studs made by Grassie [94][86] does seem to be relevant, as the defects are found to form under different circumstances. However, Grassie does not provide much direct evidence or a mechanism for stud formation. Steenbergen and Dollevoet [77][93] do provide a complete description of both initiation and propagation of squats. However, they do not give a clear explanation on why the trailing crack is sometimes not formed. Simon et al. [78] also provides an extensive explanation for squat initiation and propagation. However, they conclude that the cracks initiate at the field side of the running surface; a conclusion that is not shared by any other study.

2.4.4 Rail maintenance: grinding

Grinding is a maintenance technique that can be used as a corrective action in response to the development of RCF defects. Kalpakjian [95]p727 defines grinding as "*a chip-removal process that uses an individual abrasive grain as the cutting tool*". Chips are created as a result of a combined cutting, ploughing and shearing operation

of the top layer of the surface [95]p729. Due to the heat generated by friction between the abrasive and the workpiece the temperature can rise as a result of grinding. Kalpakjian [95]p730 relates the temperature rise ΔT to the process parameters by

$$\Delta T \propto D^{\frac{1}{4}} d^{\frac{3}{4}} (\frac{V}{v})^{\frac{1}{2}} \tag{2.19}$$

where D is the grinding wheel diameter, d is the cutting depth, V is the wheel speed and v is the workpiece speed. This is an empirical relation and is only valid for a certain range of wheel and workpiece speeds.

Rail grinding is performed by pressing multiple rotating ceramic grinding stones mounted on the underside of a train against the rail surface, while the train travels at a constant speed [7]. It is intended to reset the rail surface to the initial conditions of the rail [96]. To accomplish this, grinding has two objectives. "*Firstly to remove defect initiations at the contact surface of the rail. And secondly to re-profile the railhead to maintain the intended contact points and avoid ratcheting in the gauge corner*"[36]. The resulting rail surface after grinding is shown in figure 2.24. A polygonal surface profile is created by the longitudinal facets due to the grinding of multiple grinding stones. Transverse grinding marks are still present.

Although grinding is intended to counteract RCF, studies have found that rail grinding can create surface conditions that induce the development of RCF defects [97][96][98]. Directly after grinding, the rail surface is covered with a continuous layer of white etching material paired with an increased roughness in longitudinal direction due to the grinding grooves. The microstructure under the WEL is not heavily strained due to grinding [96]. On the softer kinds of rail steel like R260Mn, the subsequent passing of trains after grinding causes the formed WEL to spall off [96]. On harder kinds of rail steel, the WEL formed during grinding is pressed down into the pearlite matrix underneath [97][96]. The pressed down WEL causes stress concentrations in the surface layer of the rail, which result in extensive crack initiation [97][96]. After a few MGT of traffic loading, there are many small cracks in the surface, but no large cracks are present. With increasing MGT the number of short ($\leq 100\mu$m) cracks decreases, but the number and length of the long cracks increased. This suggests that some of the short cracks continued to propagate, while most other cracks are lost to wear [98].

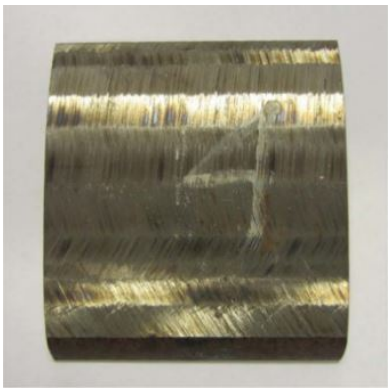
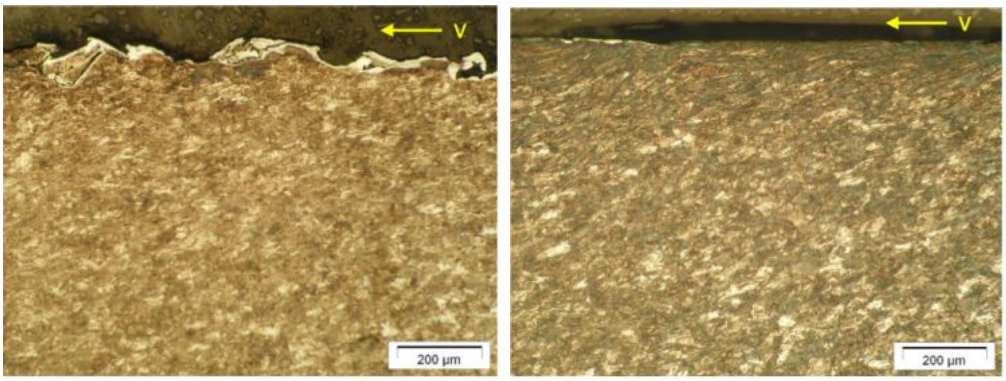


Figure 2.24: *Rail surface directly after grinding [97]*



(a) **(b)**

Figure 2.25: *Longitudinal cross-section of the surface of a R260Mn rail (a) immediately after grinding (b) after grinding and a few days of train operation. Adapted from [96]*

2.5 The influence of annual tonnage, homogeneous rolling stock and bidirectional traffic on RCF defect formation

The theories on squat initiation and propagation discussed in the previous section are all based on the analysis of squats that developed in rails subjected to, unidirectional traffic, heterogeneous rolling stock and an average annual tonnage. As mentioned in chapter 1, the rails analysed in this study are subjected to a low annual tonnage, homogeneous rolling stock and bi-directional traffic. Due to the uncommon nature of these

conditions, only a limited body of literature is available on the subjects. In this section, the effects of each of the unconventional operating conditions on RCF will be discussed separately.

2.5.1 Influence of annual tonnage

RCF defects are generally observed to increase with passed tonnage [90][99][100]. At the Australian rail network only 1% of the squats have been found at lines with less than 10 MGT annual load [99]. Railways with an annual MGT of 30 or more contributed over 45% of squat occurrence. It was concluded that traffic tonnage has a significant effect on squat occurrence. Kondo et al. [100] monitored the RCF defect development on the Japanese Shinkansen lines. They observed that dents developed in the railhead after 40 MGT, cracks developed at 100 MGT and transverse fracture occurred at approximately 350 MGT, see figure 2.26. Grassie [90] mentions similar numbers for the amount of traffic needed for the development of squats, but he states that studs can develop within 10 MGT. The railway line analyzed in this study experiences approximately 2 MGT per year, which is relatively low compared to the traffic loads associated with squat formation. However, the rails have been in use for over twenty years.

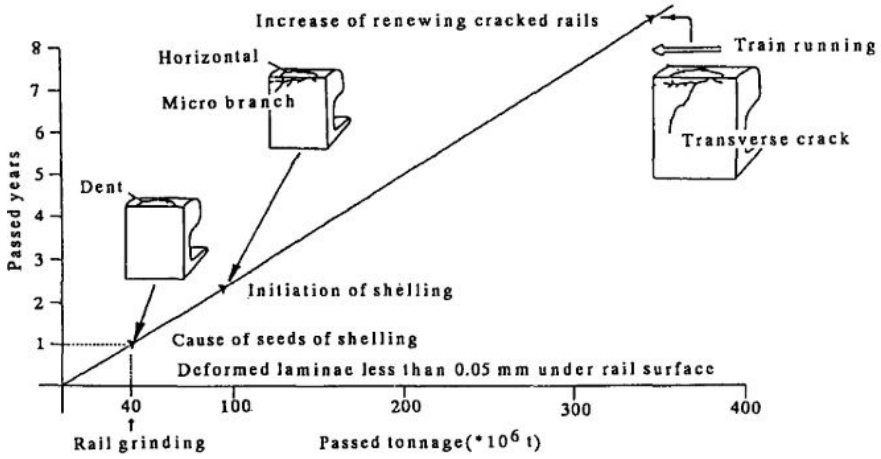


Figure 2.26: *RCF defect development as function of accumulated tonnage over time [100]*

2.5.2 Effect of homogeneous rolling stock

Although no study has been found looking into the specific case of the effects of homogeneous rolling stock (a single train type operating exclusively on a particular railroad) on RCF, some literature does exist that studies the effect of rolling stock on track degradation in a more general way. An example of the effects of rolling stock is discussed by Cannon et al. [101], who noticed frequent transverse fracture that initiated at the gauge corner, which did not commonly occur in passenger or mixed passenger-freight railways. Heavy-haul railways are typically characterized by high axle loads, uniform rolling stock and high annual tonnages. Grassie [102] mentions that a lack of variation in traffic type can enhance the development of quasi-sinusoidal irregularities with wavelengths less than a metre i.e. corrugation. Corrugation is known to accelerate wear and the development of RCF defects due to a periodic increase of dynamic force. Due to the dynamic nature of corrugation, consistency in rolling stock and travel speeds also increases consistency in corrugation wavelength [102]. Steenbergen and de Jong [103] studied the effect of rolling stock on the degradation of track geometry. They found that vertical low-frequency dynamic loads caused by wheel out-of-roundness were especially damaging.

2.5.3 Effect of bi-directional traffic

All previously discussed literature on railway wear, RCF crack initiation and defect growth were based on unidirectional traffic conditions. Unidirectional traffic is the standard on the majority of railway, as one track is used per traffic direction. However, on railways with low traffic intensity, there is no need for dual tracks and single track railways are used. On these single-track railways, the traffic is bidirectional. Some literature exists on the effects of rolling direction reversal on rail degradation, but all studies are based on twin disk tests. No studies of actual in-service rail tracks subjected to bidirectional traffic have been found in literature.

Tyfour and Beynon [104],[105] performed two studies on the effects of rolling direction reversal (RDR) using a twin disk machine; one related to fatigue crack morphology and propagation, the other related to the wear rate and wear mechanism. In the study on the effect of RDR on fatigue crack morphology, Tyfour and Beynon [104] first determined the number of cycles N_{uni} required to reach a predetermined maximum crack depth for a unidirectional rolling test. Then they performed additional tests where the rolling direction of the rail disk was reversed after a predetermined number of cycles $N_{reversal}$ until a crack reached the predetermined maximum crack depth at a total number of cycles N_{tot}. For each additional test two factors were calculated: a reversal factor RF and a life factor LF, where $RF = \frac{N_{reversal}}{N_{uni}}$ and $LF = \frac{N_{tot}}{N_{uni}}$. By varying $N_{reversal}$ in each additional test, Tyfour and Beynon [104] found that $N_{tot} > N_{N_uni}$ for all tests with RDR. A maximum fatigue life factor LF of 1.87 for $RF = 0.267$ (10.000 cycles per reversal). The crack morphology of the initiated fatigue cracks was observed to be affected by rolling direction depending on the reversal factor, as schematically shown in figure 2.27. For the lowest reversal factor attempted in the study it is observed that the crack displays a wavy shape, as depicted in figure 2.27b. It is suggested that the interval between each RDR were not long enough to cause branching or an inversion in the propagation direction. For larger reversal factors, branching and propagation direction inversion is observed.

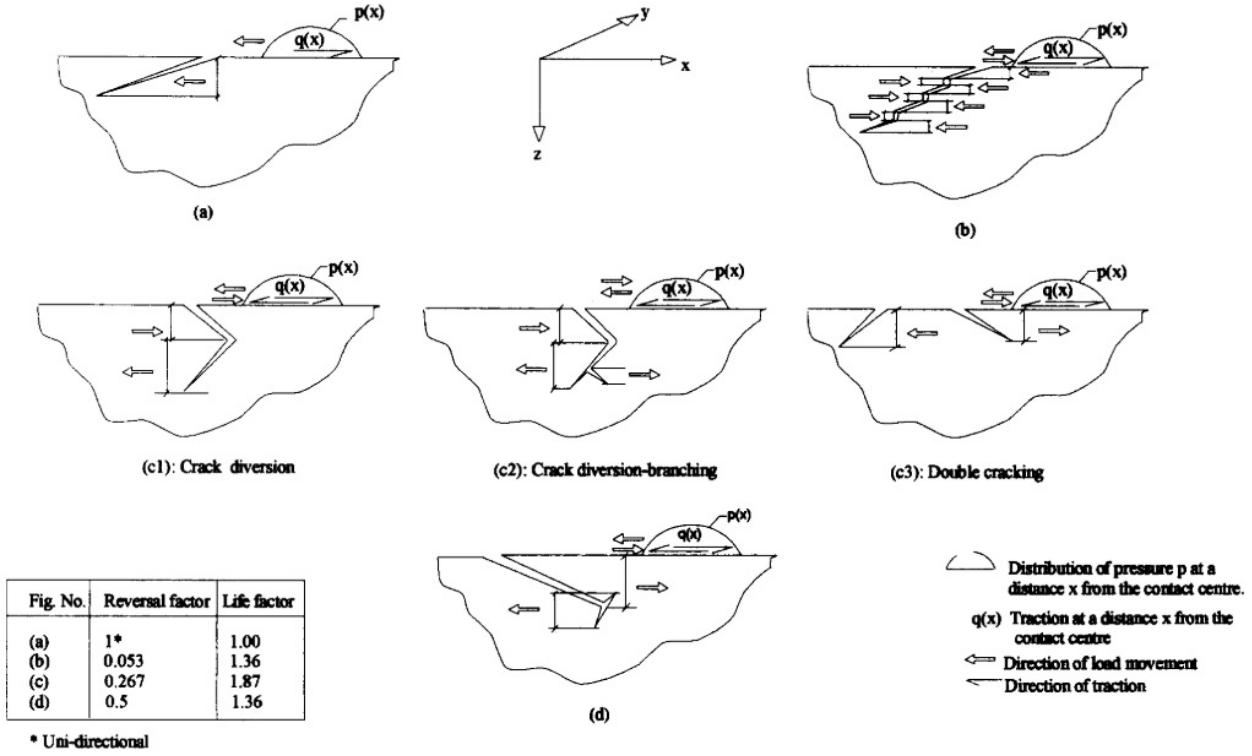


Fig. No.	Reversal factor	Life factor
(a)	1*	1.00
(b)	0.053	1.36
(c)	0.267	1.87
(d)	0.5	1.36

* Uni-directional

Figure 2.27: *Schematic representation of crack morphology for varying different reversal factor RF[104]*

The study by Tyfour and Beynon [105] on the relation between RDR and wear found that RDR caused a decrease in wear rate. First, they performed a unidirectional test. It was observed that for unidirectional rolling the wear rate was determined by the degree of accumulated unidirectional strain due to ratcheting. After an RDR, the wear rate was observed to decrease significantly. This is contributed to a reduction in the amount of accumulated unidirectional ratcheting strain. This behaviour is reflected in the microstructure, as seen in figure 2.28; a decrease in the number of cycles per RDR causes a decrease in the degree of unidirectional ratcheting strain in the surface layer, which in turn causes a reduced wear rate. The minimum wear rate was found for tests with the lowest number of cycles per RDR, which was 2500 cycles.

The minimum number of cycles per RDR attempted in the studies by Tyfour and Beynon [104][105] is large in comparison to the number of cycles experienced by the rail specimen in this study. As a result of the bidirectional traffic rolling direction is reversed after just six loading cycles. As a result not obvious if the relation between the number of cycles per RDR and the wear and fatigue behaviour found by Tyfour and Beynon [104][105] can be extrapolated to the present study. If it is assumed that these relations do hold up for an RDR interval of 6 cycles, two predictions can be made. Firstly, it is expected that the accumulated unidirectional ratcheting strain should be effectively non-existent, as six cycles would not be enough to

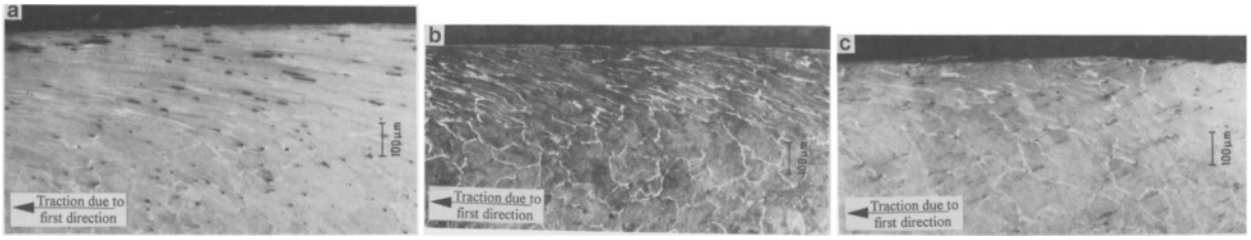


Figure 2.28: *Deformation of microstructure due to multiple RDR's after every (a) 15000 cycles, (b) 10000 cycles and (c) 2500 cycles. A reduction in the degree of accumulated unidirectional ratcheting strain is observed when the rolling direction is reversed more frequently [105].*

accumulate this strain. Secondly, it would be expected that the observed waviness of the crack would reduce, such that its appearance would be similar to a crack caused by unidirectional rolling. However, this crack would propagate slower than a crack caused by unidirectional rolling as only half of the cycles are oriented in the direction that causes crack propagation.

Zeng et al. [106] performed a similar study on the relation of RDR and wear as Tyfour and Beynon, except they changed the rolling direction of the wheel disk instead of the rail disk. They observed a drop in wear rate for the wheel disk after RDR while keeping the rolling direction of the rail disk constant. The decrease in wear rate was observed to be the result of the reversal of the direction of plastic strain accumulation. Li et al. [91][92] predict that the traffic direction does not affect the dynamic oscillation in the normal force on the rail, which is postulated to be responsible for squat formation. However, they do not take into account the effect of rolling direction reversal.

Chapter 3: Materials and Methods

In this chapter, the materials and methods used to study the rail specimens are discussed. First, a description is given of the location, track design, maintenance and traffic of the studied railway line. Secondly, after discussing sample selection, a description of the sample material is given. Thirdly, the sample preparation and examination methods are discussed and the equipment used for examination is listed. Fourthly, the set-up of the multibody wheel-rail contact simulation is discussed. Finally, the construction procedure of the subsurface crack morphology 3D model is listed.

3.1 Materials

In this study, rail specimen from the railway between Zutphen and Hengelo are examined. This railway is located in one of the less densely populated areas of the Netherlands. As a result, traffic demand is relatively low. The railroad consists of a singular track, except at the intermediary train stations. Every 30 minutes a train departs from either end of the line. In table 3.1 the total traffic in February 2020 is shown. The traffic is measured by a Quo Vadis measuring station located in Lochem. The Quo Vadis station registers the number of passing trains and axles. It also measures the weight of passing trains within an error margin of $\pm$3% by measuring the rail deflection due to the weight of a train.

The Quo Vadis data shows that the majority of traffic consists of Alstom Coradia LINT® passenger trains. LINT® trains are constructed with three bogies of two axles each. The four axles on the front and rear bogies are driven, the two axles on the middle bogie roll freely. The wheels of a LINT® have a DIN 5573 profile. Although 98.4% of the total number of trains are the LINT trains, they account only for 89.7% of the total transported tonnage. This is to be expected as freight trains are longer and have a larger average axle load than passenger trains. However, the average axle load for freight trains is not meaningful, as there exists a difference between the loaded and unloaded condition. Loaded freight wagons have an average axle load of 21.7 kN, while unloaded wagons only have an axle load of 6.0 kN. This difference is visible in the difference in the number of tons between opposite directions. Additionally, freight locomotives have an average axle load of 19.8 kN.

At the locations with RCF defects, the track is straight and supported by wooden sleepers. The rails consist of standard grade R260Mn steel and have a UIC 54E1 profile [28]. The rails are installed with a standard gauge of 1435 mm and an inclination of 1:40. The rails were manufactured by the French rolling mill Hayange in 1992. The date of installation is unknown but is most likely not long after the manufacturing date.

During their lifetime the rails received two instances of cyclic grinding maintenance, in Q1 2014 and Q1 2017. According to the cyclic grinding contract at the time of the maintenance, it would suffice to perform cyclic grinding maintenance every 15 to 22.5 mega gross tons (MGT). For the approximately 2 MGT this railway line experiences annually, it would require a maintenance interval of 7.5 to 11.25 years. The actual maintenance interval is significantly shorter than the predicted interval. During each maintenance instance, 0.17 to 0.20 mm of material is removed from the railhead.

Table 3.1: *Traffic Zutphen - Hengelo in Febuary 2020 as measured by the Quo Vadis station in Lochem.*

	# of trains	% of total trains	# of tons	% of total tons	# of axles	% of total axles
Alstom Coradia LINT®	**1922**	**98.4**	**137,057**	**89.7**	**12,456**	**91.9**
Zutphen → Hengelo	961		68,530		6228	
Hengelo → Zutphen	961		68,527		6228	
Freight trains	**30**	**1.5**	**15,614**	**10.2**	**1068**	**7.9**
Zutphen → Hengelo	15		3972		526	
Hengelo → Zutphen	15		11642		542	
Other	**2**	**0.1**	**206**	**0.1**	**24**	**0.2**
Zutphen → Hengelo	1		114		8	
Hengelo → Zutphen	1		92		16	
Monthly total	**1954**		**152,877**		**13,548**	
Estimated annual total	80,624		1,924,142		170,392	

Eddy current and ultrasonic measurements of the rail surface, performed by VolkerRail, indicated the presence of cracks on multiple sections of the track. Based on these measurements, three locations were selected for visual inspection. During these inspections, it was found that the majority of RCF defects were located on the right rail (the right rail being the rail on the right-hand side when travelling in the direction of Hengelo). Based on the visual inspections two locations were selected where multiple RCF defects of different sizes were located within a distance of six meters. The selected sample sections of rail were removed from the track on the night of the 10^{th} of September 2020. At both locations a rail section with a length of six meters was removed and replaced. From the six meter sections, the RCF defect containing parts were removed and cut into manageable pieces. An overview of the obtained sample sections is given in table 3.2.

Table 3.2: *Overview of acquired rail sections. The location is indicated with the travel distance from Zutphen station. All sections are taken from the right rail, as seen from the perspective of a train moving towards Hengelo.*

Section number	Section length (m)	Location (km)
1.1	1	7.25
1.2	1	7.25
1.3	1	7.25
2.0	1.5	5.21

3.2 Macro examination methods

The rail specimen were visually inspected to identify the locations of the defects and characteristic features left behind by wheel-rail contact and grinding maintenance. Each visible defect was given a unique number based on the distance in centimetres from the Zutphen end of the rail section. Transverse profiles of the rail head were measured with a MiniProf [107] profile measuring device. Multiple measurements were made on the removed sample sections as well as the remaining rail in the vicinity of the removed rail sections. The related MiniProf software package was used to calculate the amount of wear relative to the original UIC 54E1 profile.

3.3 Metallographic preparation and examination methods

To create metallographic samples the rail sections were cut into smaller sections with a BekaMak BMSY 320C water-cooled band saw. From these smaller sections, the railhead was separated by band saw as well. Afterwards, the samples were cut to a size suited for embedding with a water-cooled Struers Discotom-6 ceramic circular saw. The samples were embedded in Struers PolyFast hot mounting resin. A maximum heating temperature of 180 °C with a heating and cooling time of 10 and 6 minutes respectively. Sample surfaces were prepared by grinding with SiC grinding paper with a consecutively finer grit, with grit sizes of 80, 160, 320, 800, 1200, 2000 and 4000. Finally, the specimens were polished with 3 and 1 μm diamond polishing fluid and etched with 2% Nital.

A Keyence VHX-5000 digital microscope was used to optically examine the surface and microstructure of the samples. Distances and angles on the micrographs were measured with the Keyence software package. The carbon concentration of the bulk pearlite was measured with a LECO chemical analyzer. For this analysis 5 samples were cut from the rail approximately 1 cm below the running surface. Each of these samples was dimensioned to weigh 0.5 to 1 gram. After cutting, all sample surfaces were ground with Al_2O_3 grinding paper, instead of SiC grinding paper to prevent carbon from being deposited on the sample by the grinding paper. Vickers hardness measurements were performed with a Struers DuraScan. A load of 3 kgf was used.

3.4 Multibody wheel-rail contact simulations

Multibody simulation is a method that simulates the dynamics of a system composed of multiple rigid or elastic bodies. Connections between the bodies can be made through joints, spring-dampers and friction between surfaces. Multibody wheel-rail contact simulation is a tool frequently used to gain insight into the contact phenomenon. The simulation software can calculate parameters such as position, area and shape of the contact patch, as well as forces, stresses and creep. The goal of the simulation performed in the present work is to examine if there exists a relation between the simulated running behaviour and the position of the defects found on the rail surface.

The simulations in the present work are performed with the VAMPIRE® software package. The software constructs physical models of the vehicle and track based on several inputs, such as rail profile, wheel profile, track geometry, track stiffness, vehicle speed and vehicle dynamic behaviour. Before a simulation is executed the software calculates a wheel-rail contact data file. This file contains a table of all possible combinations of contact points between the rail and wheel profile. These calculations are based on the Hertzian contact theory of double-curved surfaces. Therefore the contact patch is assumed to have an elliptical shape. The relation between creep and creep forces is assumed to obey a non-linear creep law. This non-linear creep law uses the pre-calculated Kalker table to compute the creep forces in each time step given the contact parameters. A more complete description of the non-linear creep law is given in appendix A.3.

For the vehicle, model data is required such as dimensions, mass, the moment of inertia, stiffness, etc, for multiple components of the train. The input for these parameters was based on the technical data Alstom LINT® 41/H train. Furthermore, the vehicle model requires the input of transverse wheel profiles. Multiple MiniProf measurements were made on the wheels of three of the actual trains that operate on the railway between Zutphen and Hengelo. From these measurements, two profiles are selected. In total, three wheel profiles are selected:

- DIN 5573 profile in new condition
- DIN 5573 profile with the heaviest wear
- S1002 profile in new condition

LINT train wheels do not have the S1002 profile, but it is the most widely used wheel profile on other trains. Therefore it is used as a benchmark for measured wheel profiles.

To complete the vehicle model an input for the train weight must be set. Train weight can have two settings:

- Crush: 145 kN per axle (at maximum passenger capacity)
- Tare: 106 kN per axle (no passengers)

As input for the track model transverse rail profiles and data on track geometry are required. Three rail profiles are selected:

- UIC 54E1 profile in new condition
- UIC 54E1 profile at km 5.21 measured by MiniProf at the moment of removal from the track
- UIC 54E1 profile at km 5.21 measured by MiniProf measured directly after grinding maintenance in 2017

For the track geometry, two inputs can be selected:

- Perfectly straight track
- Track with irregularities

Track irregularities are small perturbations in the track geometry parameters, such as horizontal and vertical level, gauge offset and cross-level. In reality, railway tracks are never perfectly straight. Therefore track irregularities result in a more realistic track model. The track geometry parameters are taken from the TRACK 110 file in the VAMPIRE® library, which is shown in appendix A.2 figure A.2. These track irregularities are representative of an average railway track with a maximum speed of 120 km/h.

For each unique combination of input parameters, a simulation is run. This resulted in a total of 36 runs, which are indicated in table 3.3. However, just 8 of the 36 input combinations result in a running behaviour that simulates reality. The 24 other runs are theoretical and serve as a reference. 4 of these 8 runs, which were simulated with crush weight, are evaluated in section 4.2. Their combination of input parameters is highlighted in red in table 3.3.

Each simulation run calculated the running behaviour of the train over a distance of 500 m, while travelling at a speed of 120 km/h. The output was recorded for the last 100 m of the run, where the running behaviour had reached a steady state. All selected output parameters were recorded for the two right wheels of the front bogie of the train. The output time step was set to 1 ms and a friction coefficient of 0.32 is assumed between the wheel and rail surfaces. For the last 50 m of each run, an external torque of 0.7 kNm was applied to each wheelset to simulate the tractive torque that would be required to maintain a constant travelling speed of 120 km/h. The first 450 m of the run the train was artificially fixed at travelling speed.

Table 3.3: *Simulation matrix. The four letter label of each run indicates the wheel profile, rail profile, weight (C = crush, T = tare) and track irregularities (S = straight/without irregularities, I = with irregularities) in that order. The labels indicated in red are evaluate in the results.*

Wheel profile	Rail profile						Weight
	New (N)		Worn (W)		Ground (G)		
New DIN 5573 (N)	NNCS	NNCI	NWCS	NWCI	NGCS	NGCI	Crush
	NNTS	NNTI	NWTS	NWTI	NGTS	NGTI	Tare
Worn DIN 5573 (W)	WNCS	WNCI	WWCS	WWCI	WGCS	WGCI	Crush
	WNTS	WNTI	WWTS	WWTI	WGTS	WGTI	Tare
S1002 (S)	SNCS	SNCI	SWCS	SWCI	SGCS	SGCI	Crush
	SNTS	SNTI	SWTS	SWTI	SGTS	SGTI	Tare
Track irregularities	S	I	S	I	S	I	

3.5 3D model of subsurface crack morphology

A 3D model of the crack contour underneath the surface was constructed. For the construction of this model, a 20 cm section of rail containing a large surface crack was milled in an EMCO FB-4 milling machine. The rail section was positioned horizontally on the table of the milling machine, with its longitudinal axis aligned with one of the horizontal movement axes of the table. The vertical position of the table was set to zero when the mill just touched the top of the rail. To expose the shape of the crack underneath the surface, a total of 66 layers of material were milled from the rail. From a milling depth of 0 to 1 mm, layers with a thickness of 0.05 mm were removed during each milling pass. After a depth of 1 mm was reached, the layer thickness was increased to 0.1 mm per milling pass until all material containing cracks was removed. A mill with a diameter of 80 mm was used, such that a layer of the 70 mm wide rail could be removed in one milling pass.

Between each milling pass, the rail surface was photographed to capture the shape of the crack. To ensure that all photos were taken from the same viewpoint, the camera was mounted on a tripod which was carefully positioned in the same location relative to the rail after each milling pass. A scale bar was placed on the milled surface before each photograph was taken.

From the collected images an image stack was created. In this image stack, each image represents a horizontal slice of the three-dimensional crack contour. By stacking the images in the vertical direction the 3D crack contour was reconstructed. First, the images were processed using Microsoft Powerpoint. The images were scaled to compensate for the increasing focal distance by each milling pass. Then the crack contour was manually traced with the pencil tool to highlight it in a colour with saturation and value of 100%. The processed images were imported into MatLab, where the traced crack contours were extracted from the images by filtering out the highlight colour. Then the pixels of the extracted crack contours were plotted as points in 3D space for each of the image stack layers. Finally, the axes were scaled to the right dimensions and a colour gradient was added to better display the depth of each layer.

Chapter 4: Results

In this chapter the results of the investigation of the rail samples are presented. In section 4.1 the characteristics of the rail surface and defects are described based on visual examination and the profile measurements of rails and LINT train wheels are presented. In section 4.2 the results of the wheel-rail contact simulation are presented. Multiple parameters related to wheel-rail contact phenomena are evaluated as a function of the rail contact position discusses. The microstructure and hardness of the rail surface material are examined in section 4.3. Finally the 3D model of the crack contour of a large RCF defect is reconstructed in section 4.4

4.1 Macro examination

4.1.1 Visual examination

Visual examination shows a shiny band on the head of the rail, as presented in figure 4.1. This band, called the running surface, is caused by contact with train wheels. Areas outside the running surface are covered in a brown oxide layer. The running surface has a width of 22 $\pm$ 1 mm. The centre of the running surface is offset by 4 mm to the gauge side with respect to the centreline of the rail. On the running surface patterns of alternating light and dark areas are seen. These patterns are the result of the grinding maintenance performed in 2017 and the wear by wheel-rail contact. Just like a recently ground rail surface, multiple grinding facets can be identified on the running surface. On the running surface in figure 4.1 four grinding facets can be identified based on the remaining grinding patterns. Grinding facets on the areas outside the running surface can not be distinguished as the oxide layer blurs the boundaries between them.

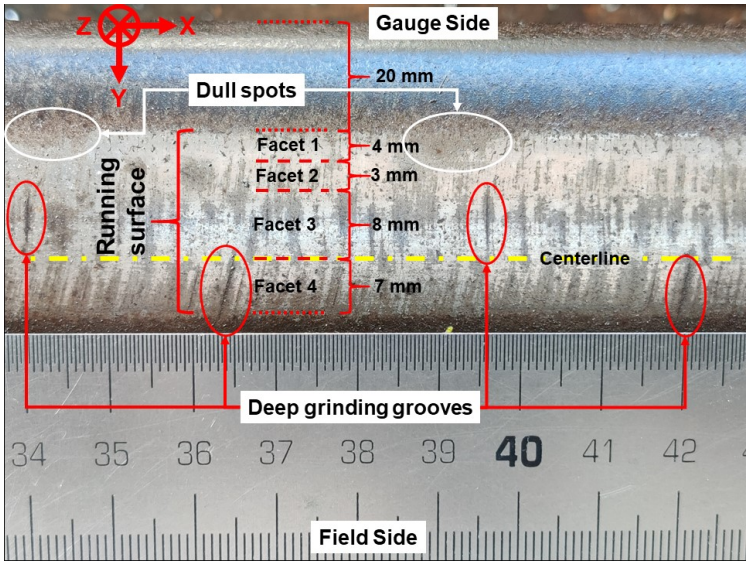


Figure 4.1: *Structure of the running surface.*

In figure 4.1 there are large grooves visible on facets 3 and 4, which are encircled in red. On both facets these grooves repeat periodically with an equal interval of 57 $\pm$ 1 mm. Closer examination of the alternating light and dark areas on the running surface reveals a repeating pattern with an interval equal to that of the large groove interval. Encircled in white are dull spots on facet 1, which also repeat periodically. However, these dull spots repeat with a period of 49 $\pm$ 2 mm.

4.1.2 Surface-breaking defects and wear patterns

Visual examination of the running surface shows defects of different sizes and shapes. Examples of these defects are presented in figure 4.2. Based on their appearance the defects can be classified into three categories, increasing in severity. This classification is based on the classification of Li et al. [92].

(A) crack
(B) crack and dark spot
(C) crack, dark spot & running surface widening

Categories A, B and C can be further classified into two kinds of shapes: U-cracks and slash-cracks. Figure 4.2a shows a category A and B U-cracks. The surface cracks are slightly curved and resemble the bottom of a U-shape. The concave side of the U-cracks is oriented towards the gauge side of the rail. All of the U-cracks found are either completely or partially contained within the second grinding facet. The surface crack tips often coincide with the boundary between grinding facet 1 and 2, whereas the point of the surface crack nearest to the field side often coincided with the boundary between facet 2 and 3. Category B U-cracks show

a darkened spot at the field side of the crack. This dark spot is contained within the second and third grinding facets.

Figure 4.2b shows category A and B slash-cracks. Slash-cracks owe their name to their resemblance to the \ and / symbols. Contrary to the U-shape cracks, slash-cracks are straight and oriented under an angle of 45° relative to the rail axis. Category A slash-cracks are all completely contained within the second grinding facet. The crack tips coincide with the edges of facet 2. Right-hand slash-cracks (/) (as in figure 4.2b) as well as left-hand slash-cracks (\ ) are present on the rail surface. Category B slash-cracks are accompanied by a dark spot at their field side. This dark spot is contained within the second and third grinding facets. No category B left-hand slash-cracks are observed.

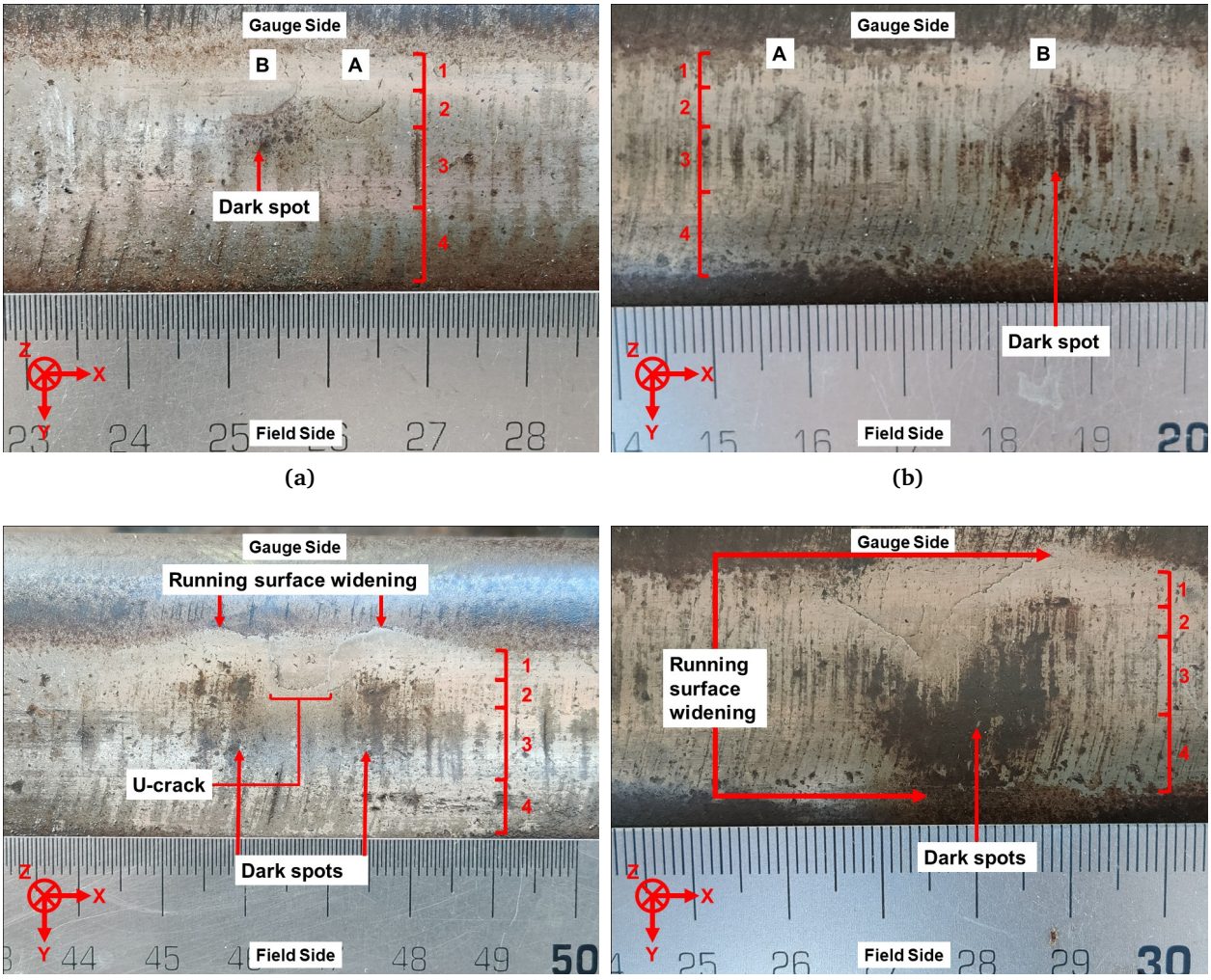


Figure 4.2: *Surface defects (a) category A and B U-cracks, (b) category A and B /-cracks, (c) category C U-crack, (d) category C /-crack.*

Figure 4.2c shows a U-shape category C defect. The bottom horizontal part of the U-shape is slightly curved and falls within the second grinding facet. This part of the U-shape resembles the shape of a category A and B U-crack as shown in figure 4.2a. On both sides the crack sharply turns towards the gauge side. The points at which the crack turns coincide with the boundary between grinding facet 1 and 2. At the edge of the widened running surface the crack turns horizontally and follows the contour of the widened running surface. At the field side of the crack a dark spot is present. This dark spot has a lighter shade directly beneath the horizontal part of the U-crack.

Figure 4.2d shows the other type of category C defect: the V-crack. The two branches of the V-shape make an 90° angle and the point where they meet approximately coincides with the boundary between grinding facet 2 and 3. The right branch is longer than the left branch. At the point where the right branch crosses the

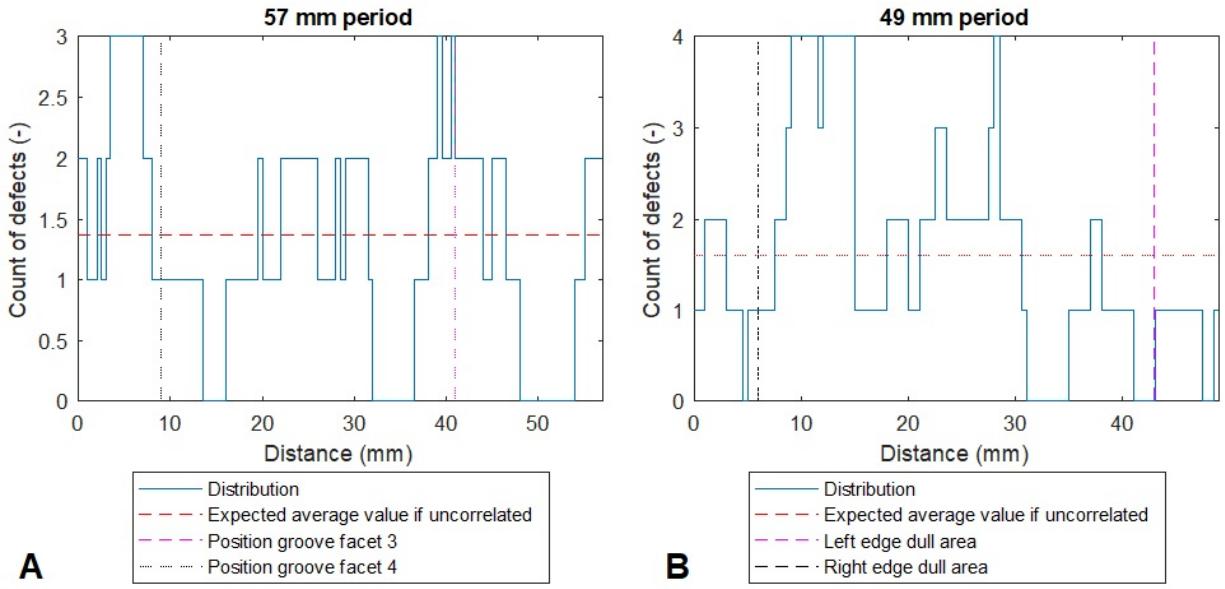


Figure 4.3: *Distribution of defect locations along periodic patterns in running surface (a) 57 mm period (b) 49 mm period.*

boundary between facets 1 and 2 the crack turns by 30°. The right branch ends at the edge of the widened running surface. A dark spot is present at the field side of the V-crack. Similar to the V-crack, the dark spot is larger on the right hand side than on the left.

Other than the defects that are concentrated on facet 2, facet 1 is strewn with small cracks. These cracks can are visible in figures 4.1 and 4.2a. These small cracks do not have a fixed orientation and position on facet 1.

Figure 4.3 shows the distribution of the surface defect position along periodic surface patterns mentioned in section 4.1.1. This section contains a total of 17 surface defects, which span a combined longitudinal length of 78 mm. If the defect positions are uncorrelated to the periodic running surface patterns average value is expected of 1.37 and 1.59 for the 57 and 49 mm periods respectively, as indicated by the dashed red lines. Defect count can only take discrete values, thus if there were no relation and the defects were distributed randomly along the period, it would result in a distribution alternating between a count of 1 and 2. For the 57 mm period this is the case for 69.1% of the period. The maximum count of three also indicates that there is no significant concentration of defects along the period. For the 49 mm period the distribution is similar, with 67.3% of the interval having a count of 1 or 2 defects. Between 9 and 15 mm there is a peak with a count of 4, but this also does not indicate a relation.

4.1.3 Wheel and rail profile measurements

To analyze the geometry of the wheel and rail surfaces, multiple MiniProf [107] profile measurements of transverse wheel and rail profiles are made. Other than serving as input for the multibody wheel-rail contact simulation, which will be discussed in the next section, the transverse profiles give insight into the wear of the rails and wheels. The rail head dimensions of a new UIC 54E1 rail profile are shown in figure 4.4. The surface of the rail head consists of three sections, which have different radii of curvature. The running surface as measured in figure 4.1 is located within the area indicated with B.

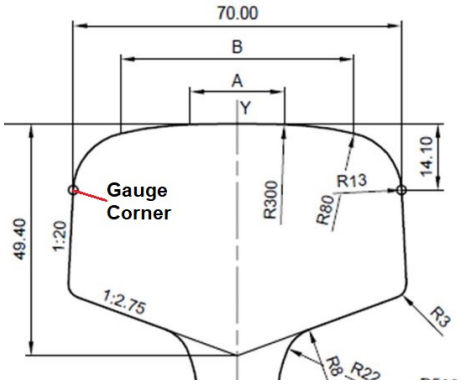


Figure 4.4: *Rail head dimensions of a new UIC 54E1 rail profile. A = 20,024 mm, B = 49,727 mm [28].*

Figure 4.5 shows the three transverse rail profiles that are used as input for the simulation. The rail profiles are aligned such that the R3 corners overlap. The position of these corners does not change due to wear or grinding maintenance. As a result of this alignment, the total volume of removed material due to wear and grinding maintenance is visible. The vertical difference between the highest points on the surface of the profile measured at removal from the

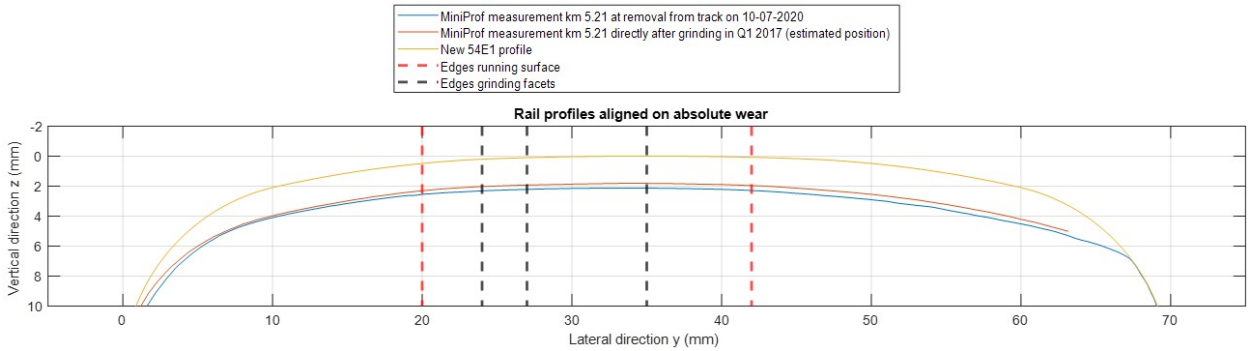


Figure 4.5: *Transverse rail profiles measured with MiniProf [107] aligned on absolute wear. The exact position of the profile measured directly after grinding relative to the other profiles is unknown. The position of this profile in the graph is estimated. $y = 0$ at the gauge corner of the new UIC 54E1 profile.*

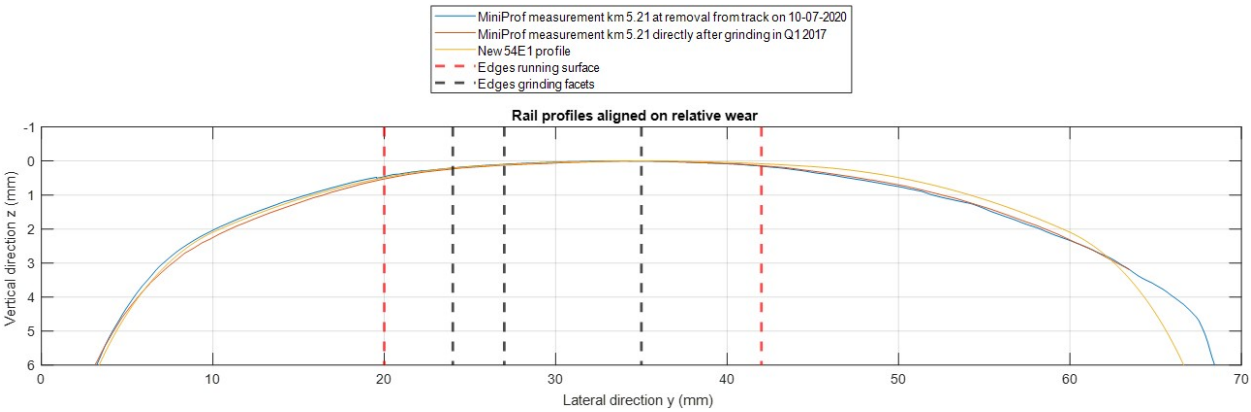


Figure 4.6: *Transverse rail profiles measured with MiniProf aligned on relative wear. The measured profiles are positioned such that they overlap at $y = 0$ and their highest point is at $z = 0$.*

track on 10-07-2020 and the new UIC 54E1 profile is 2.130 mm. Near the gauge corner, the rail is worn in lateral direction due to contact with the wheel flange. At the gauge corner, wear and grinding maintenance has removed a layer with a lateral thickness of 0.776 mm (not shown in the figure). The wear of the profile measured directly after grinding can not be measured exactly, because the measurement does not include the sides of the rail. Therefore its vertical position relative to the new UIC 54E1 profile is unknown. The position shown in 4.5 is estimated based on the average annual wear.

In figure 4.6, the same three transverse rail profiles are aligned such that their relative wear is visible. These relative differences in profile have a larger effect on the running behaviour of the train than the absolute differences, since these affect the contact point between wheel and rail. The relative differences in surface height and curvature are largest on the parts of the rail outside of the running surface. On the running surface itself the relative difference between the measured profiles and the new profile is small; the largest difference in surface height is 0.07 mm at the field side edge of the running surface. On the gauge side of the running surface the maximum difference in height is 0.025 mm.

In figure 4.7, two of the measured transverse wheel profiles are presented. The profiles are aligned such that the wheel flanges overlap, such that the profile difference as a result of wear is visible. The maximum vertical difference between the new and worn profiles is 2.3 mm at y = 750 to 755 mm. Wear caused the surface of the wheel to hollow out. The slope of the tangent is negative for every point of the new rail profile. For the worn profile this is not the case; from y = 760 to 769 mm the tangent is positive. The maximum amount of hollowing wear allowed for safe train operation is 2.5 mm, which is measured as the vertical difference between the lowest point and the highest point on the field side of the wheel. The hollowing wear for the worn profile shown in 4.7 is 1.206 mm.

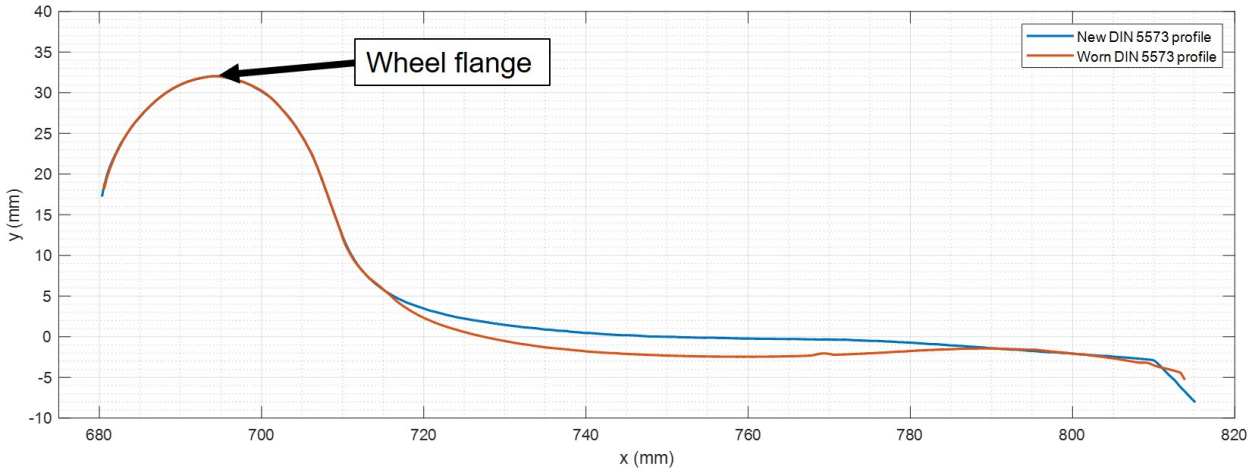


Figure 4.7: *Wheel profiles of LINT® train measured with MiniProf*

4.2 Wheel-rail contact simulation

Multibody wheel-rail contact simulation is a tool frequently used to gain insight into the contact phenomenon. The simulation software can calculate parameters such as position, area and shape of the contact patch, as well as forces, stresses and creep. The goal of the simulation performed in the present work is to examine if there exists a relation between the simulated running behaviour and the position of the defects found on the rail surface. The parameters will be evaluated for a selection of 4 of the 36 simulation runs, as indicated in table 3.3. Paragraphs 4.2.1 to 4.2.3 will only consider the output of wheel 1, as the differences in the output between the two wheels is negligible. For the creep, discussed in 4.2.4, the differences between wheels are significant.

4.2.1 Contact patch position

When a wheelset with conical wheel profiles is laterally displaced from the central position, this displacement is counteracted by the difference in rolling radii of the wheels. This causes a oscillating lateral motion of the wheelset, which is called Klingel movement [40]p20. As a result Klingel movement, the wheel-rail contact patch moves laterally over the wheel and rail surfaces. This is illustrated in figure 4.8. When the lateral displacement is zero, the contact patch on both rail is located at an equal distance from the gauge corner. If the wheelset is displaced in the positive y direction, the contact patch on the right rail moves towards the gauge corner and moves towards the field side on the left rail. The opposite effect occurs for a negative lateral displacement.

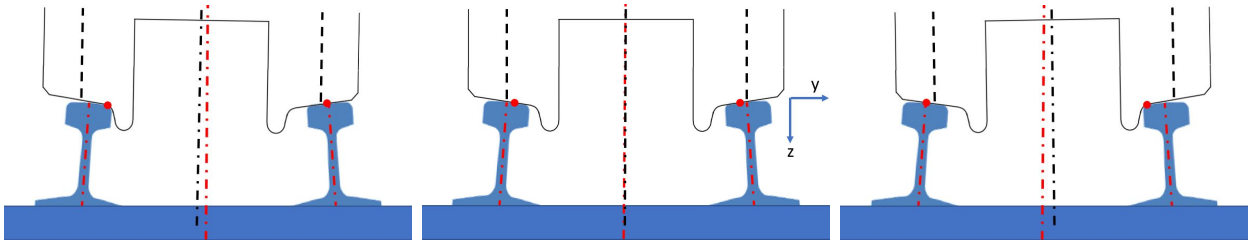


Figure 4.8: *Lateral shift of the contact patch as a result of Klingel movement for (a) negative, (b) zero, (c) positive lateral displacement. The position of the contact patch is indicated by the red dots.*

The effects of Klingel movement can also be seen in the simulation results. Figure 4.9 shows the position of the contact patch centre and the lateral displacement of the wheel relative to the rail as a function of travel distance. When the lateral displacement is positive the contact point shift towards the gauge corner and when it is negative contact shifts towards the field corner. When the different runs are compared against each other, it appears that the wheel profile has the largest influence on the running behaviour. The oscillation frequency of the lateral displacement is lower for the new wheel profile than for the worn wheel profile. This can be

seen from the number of times the lateral displacement crosses the neutral position: 13 and 15 times for the new wheel profile, 24 and 26 times for the worn profile. The rail profile does not have such a significant impact on the running behaviour as the wheel profile. This behaviour can be understood by comparing the differences between individual wheel and rail profiles in figures 4.6 and 4.7. For wheels wear causes a much more significant change in profile than it does for rails. The change in profile alters the curvature of the wheel surface, which influences the running behaviour. For rails the relative differences in surface curvature are much more subtle than for wheels. Therefore the wheel profile affects the running behaviour most.

The simulated contact position on the rail matches reality well; the majority of simulated wheel-rail contact occurs within the boundaries of the running surface (20-42 mm) as measured on the rail specimen in figure 4.1. In figure 4.9 the edges of the measured running surface are indicated by the striped red lines. Within the boundaries of the running surface the rail contact position appears to oscillate between two primary contact bands; one near the gauge side of the running surface, the other near the field side of the running surface. For the simulations with a new DIN 5573 wheel profile, the gauge side contact band is located at a contact position of 21-24, the field side contact band at 33-35 mm. The transitions of contact between these two primary contact bands are sudden: the contact point changes within 0.1 to 0.4 m of distance travelled. For the simulations with a worn wheel profile contact the gauge side contact band is located at a rail contact position of 21-24 and the field side contact band at 39-41 mm. The transition between these two bands is less sudden than for the new wheel profile; for the worn rail profile the contact point shifts between band in 0.4 to 1 m of distance travelled.

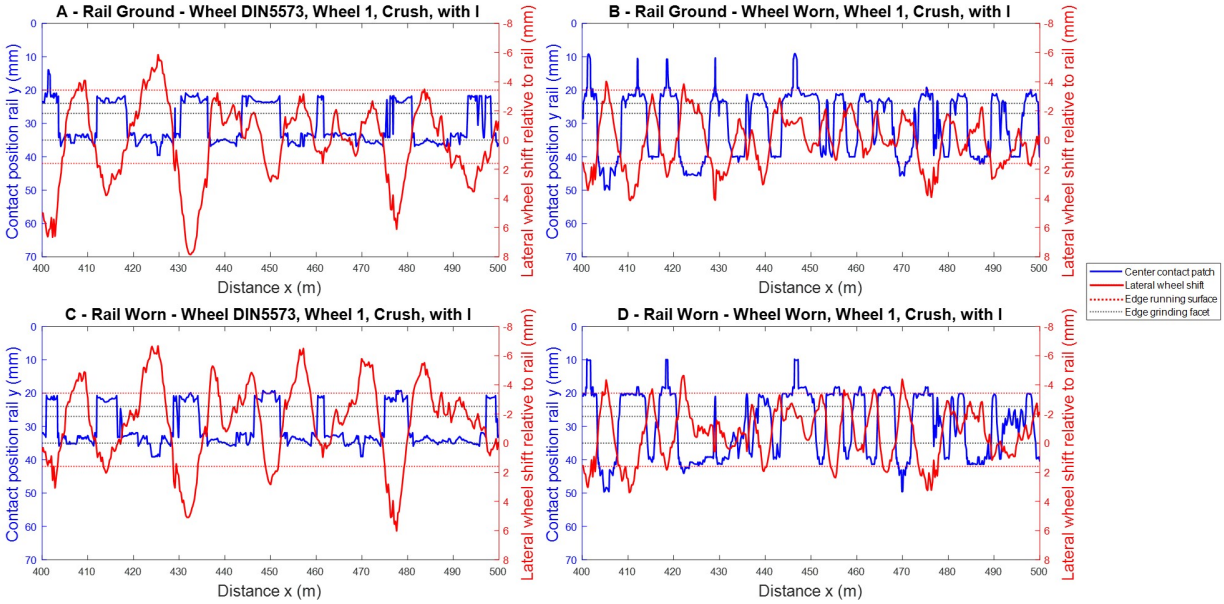


Figure 4.9: *Contact position on right rail and lateral displacement of the wheel relative to the rail as function of travel distance. Positive lateral displacement is for wheelset movement to the right, i.e. towards flange contact on the right rail.*

4.2.2 Area and ellipticity of the contact patch

The VAMPIRE® simulation package assumes a elliptical contact patch. A real contact patch will never be perfectly elliptical, but it is a close approximation. The ellipticity e is the ratio between the longitudinal and lateral axes of the contact ellipse and solely depends on the curvature of the surfaces in contact [42]p108. The curvature in longitudinal direction is constant for both the wheel and rail. Therefore the ellipticity depends primarily on the curvature of the transverse wheel and rail profiles. This effect is visible in figure 4.10. The ellipticity is largest towards the edges of the running surface, where contact occurs between parts of the wheel and rail with a relatively large curvature. The ellipticity is lowest in the centre of the running surface. Here the surfaces in contact have a relatively low curvature and are more conforming. In reality, the contact patch will never be perfectly elliptical, but it is a close approximation.

The area of the contact patch varies as a function of the rail contact position. The area is not only a function of the surface curvature, but also depends on the magnitude of the normal load and modulus of elasticity of the contacting materials [42]p108. From figure 4.10 it can be seen that the area and ellipticity have an

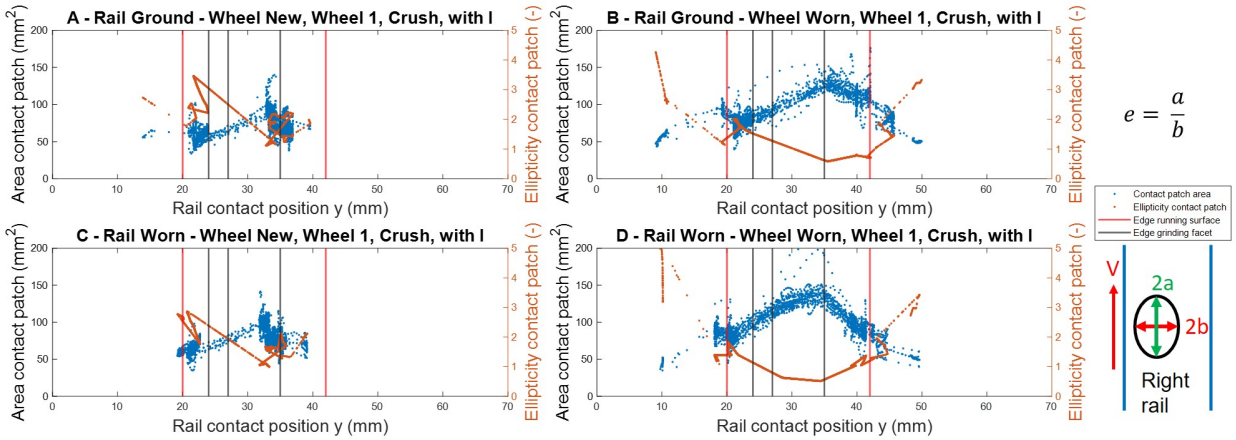


Figure 4.10: *Area and ellipticity of the contact patch as function of rail contact position.*

inverse relation; when the ellipticity is large the area is small and vice versa. Unlike the ellipticity, the area displays a spread. This is caused by fluctuations in the normal load P as a result of the track irregularities, $P = 70 \pm 9$ kN.

In the previous paragraph it is mentioned that contact on the rail concentrates on two primary bands. This can also be seen from the high density of the data points for the contact patch area. the wheel profile has the most significant effect on the running behaviour. For both runs with the new wheel profile, the two primary contact bands are clearly visible due to the density of the blue dots. The distribution into two primary contact bands is the clearest for the new wheel profile.

4.2.3 Normal stress

If the contact patch area and the normal force are known, the maximum normal stress in the rail can be calculated with following equation

$$p_0 = \frac{3}{2}\frac{P}{\pi ab} \tag{4.1}$$

where πab is the area of the elliptic contact patch. In figure 4.11, the maximum normal stress is plotted as a function of contact position on the rail. The inverse relation to the contact stress is clearly visible. The new wheel profile causes a significantly larger normal stress than the worn wheel profile. On average the maximum normal stress is 438 MPa larger higher for a new wheel profile than for a worn rail profile. The rail profile also has a small effect on the magnitude of the normal stress; on average the ground rail profile causes a normal stress which is 108 MPa higher than the worn rail profile. The largest differences in normal stress between wheel and rail profile are found on grinding facet 2; on average the new wheel profile causes a normal stress 503 MPa larger than the worn wheel profile and the ground rail profile causes a normal stress 230 MPa larger than the worn rail profile. The differences in stress for different wheel and rail profiles is the result of wear. The removal of material due to wear causes the wheel and rail profiles to increase their conformity. A higher degree of conformity increases the contact area, which in turn decreases the stress.

4.2.4 Creep

Creep is essential for the guidance of wheels along rails [45]. As a result of creep, forces are created between the contacting surfaces which act tangential to the plane of contact. These creep forces govern how the wheels roll over the rail. The relation between creep and creep force can best be understood by the analogy of walking on a slippery floor. If you push off with your foot on the floor, your body moves forward, but your foot will slide backwards due to the lack of grip. Although they are sliding, the foot exerts a tangential force on the ground, and the ground exerts an equal and opposite force back on the foot. These forces are the creep forces. A creep force that acts on the foot (the wheel) will point in the direction opposite the slip direction, and the creep force that acts on the floor (the rail) points in the same direction as the creep. This relation is given by the following equation

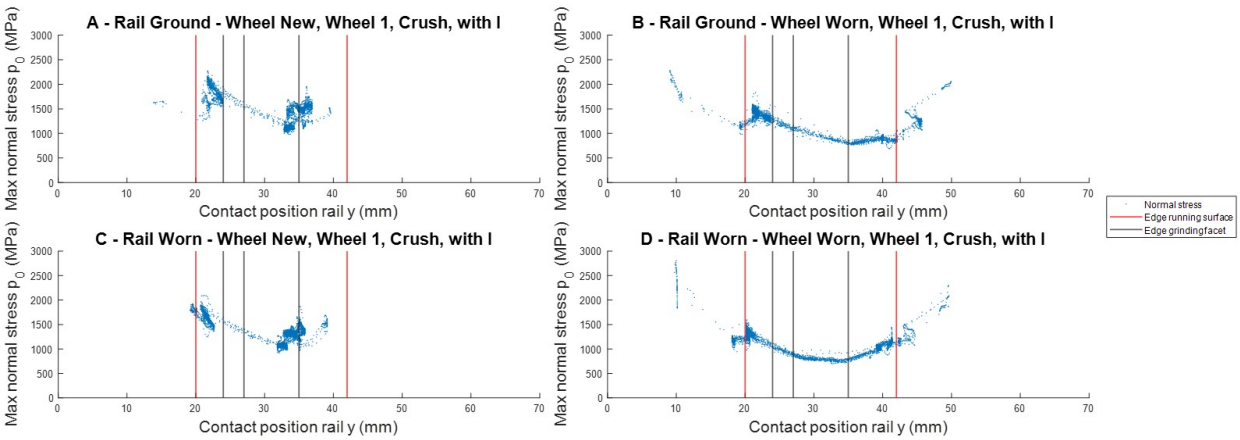


Figure 4.11: *Maximum normal stress p_0 as function of rail contact position.*

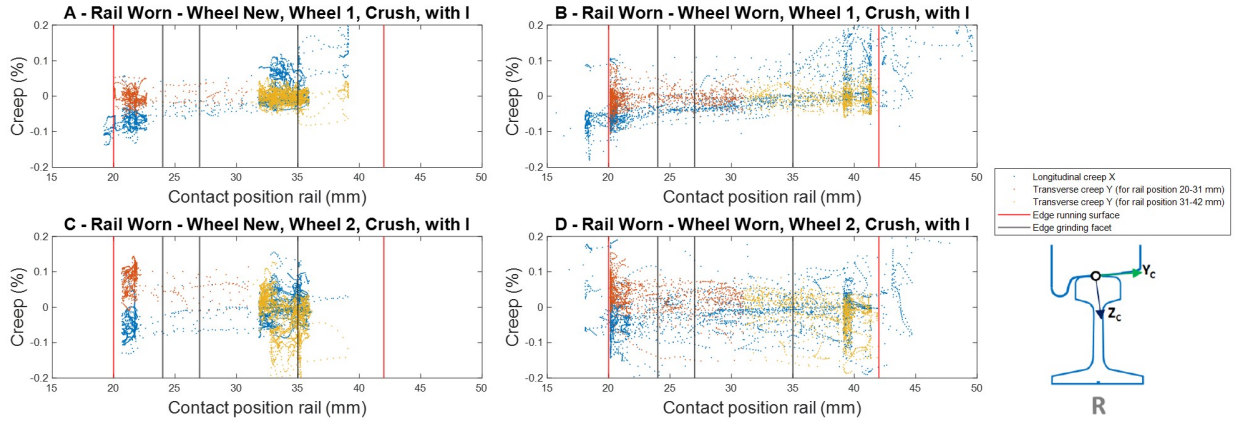


Figure 4.12: *Longitudinal and lateral creep (ξ_x and ξ_y) as function of rail contact position. The lateral creep is divided into two colours to better clarify the results in figure 4.13.*

$$Q_{\xi,w} = -Q_{\xi,r} = -A\xi \tag{4.2}$$

where $Q_{\xi,w}$ and $Q_{\xi,r}$ are the creep forces on the wheel and rail respectively, A is a constant and ξ is the creep ratio. In the remainder of this paragraph the term 'creep force' will refer to $Q_{\xi,r}$.

The VAMPIRE® software calculates the creep in the centre of the contact patch in longitudinal and lateral direction (ξ_x and ξ_y) as function of the contact geometry (see appendix A.3 for the equations). In figure 4.12 the longitudinal and lateral creep are plotted as function of rail contact position. The creepages are expressed in the wheel-rail contact axis system as indicated by the image on the right side of figure 4.12. This axis system is fixed to the centre of the contact patch and the z-axis of this axis system is always perpendicular to the contacting surfaces. The top row shows the results for wheel 1, the bottom row for wheel 2. ξ_x displays the same behaviour for all profile combinations as well as wheel position; ξ_x is negative on the gauge side contact band and positive on the field side contact band. Recalling equation 2.11, ξ_x is negative if $V_1 < V_2$ and positive if $V_1 > V_2$. The behaviour of the longitudinal creep is the result of wheel conicity. This effect is also illustrated in figure 4.8. Whenever a wheel set gains a positive lateral displacement, the effective radius R of the right wheel increases as the contact point shifts towards the wheel flange. Since $V_2 = \omega R$, where ω is the angular velocity of the axle, the velocity of the wheel surface increases due to the increase in R. Therefore ξ_x is negative towards the gauge corner and ξ_x is positive towards the field corner. ξ_x changes sign in the middle of the running surface at rail contact position of 31 mm.

The lateral creep is a function of the attack angle, i.e the angle between the plane of the wheel and the longitudinal rail direction [45], and the lateral velocity of the wheel with respect to the rails. The attack angle changes continuously as a result of the Klingel movement. Unlike ξ_x, ξ_y does not display the same behaviour for all runs; the behaviour is different for wheel 1 and 2. For wheel 1, ξ_y is both positive and negative for all

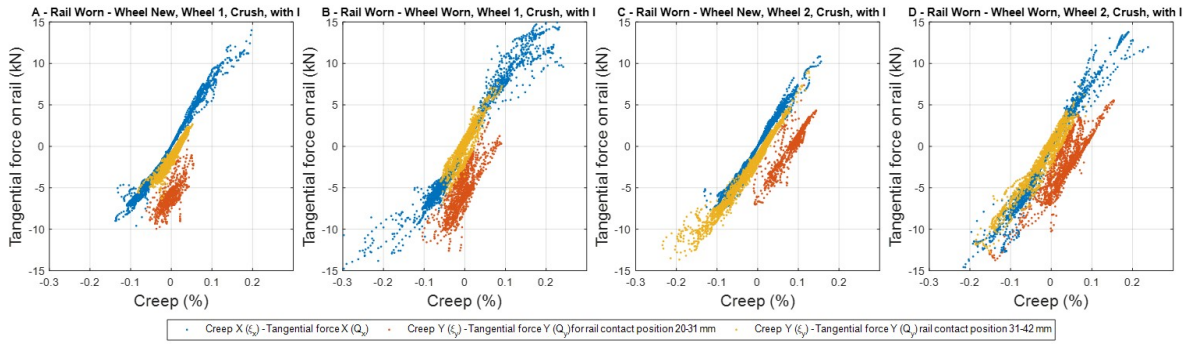


Figure 4.13: *Relation creep - tangential force in longitudinal (X) and lateral (Y) directions. The two colors (orange and yellow) that indicate the lateral creep and creep force are divided in the same way as in figure 4.12*

rail contact positions, resulting in a mean creep close to zero. For wheel two, ξ_y is positive on the gauge side of the rail and negative on the field side of the rail.

As a result of creep, tangential forces are generated in the wheel-rail contact. These forces guide the train along the rails. VAMPIRE® calculates the tangential forces in the wheel-rail contact as function of the area, ellipticity, equivalent curvature and creep in the contact patch. In figure 4.13 the tangential forces acting on the rail are plotted as function of the creep ratio. The tangential force in longitudinal direction Q_x obeys equation 4.2; all blue data points are located on a line with a positive A slope through the origin. The yellow and orange data points indicate the relation between lateral creep ξ_y and tangential force in lateral direction Q_y. The colours indicate the corresponding contact position on the rail: orange indicates contact at the gauge side contact band, yellow indicates contact at the field side contact band. If contact occurs on the field side of the running surface, the simulated creep and creep force obey equation 4.2. However, when contact occurs at the gauge side of the running surface, the creep and creep force do not obey equation 4.2. Instead the points are located on a line with the following equation:

$$Q_y = A\xi + B \tag{4.3}$$

This relation implies that there is another source of lateral tangential force B, which is a function of the contact position on the rail. From figure 4.13 B can be determined; B is the value of Q_y when ξ_y is zero. If the rail contact position lies between 31 and 42 mm, B must be close to zero. For contact positions between 20 and 31 mm, B must must be -5 $\pm$ 2 kN.

4.3 Metallographic examination

In this section the results of the metallographic examination are presented. First, the microstructure of the bulk and the defectless surface are investigated. Secondly, the cross sections of three Category A defects are examined. Then, the results of hardness measurements on the running surface are presented. Finally, the 3D model of the subsurface crack morphology of a category C U-crack is presented.

4.3.1 Bulk microstructure

Figure 4.14 shows a longitudinal section of the microstructure of the R260Mn bulk material. MnS inclusions are present in the material. The inclusions appear as slender rods in the longitudinal cross section, while the transverse cross section shows tiny sphere-like inclusions, see figure 4.15. Therefore the inclusions must have a rod-like shape oriented in the longitudinal direction. The inclusions are formed during the steel making process and formed into the rod shape by hot rolling. The carbon weight fraction of the rail steel is measured to be 0.62 w%, which is within the specified limits in [28].

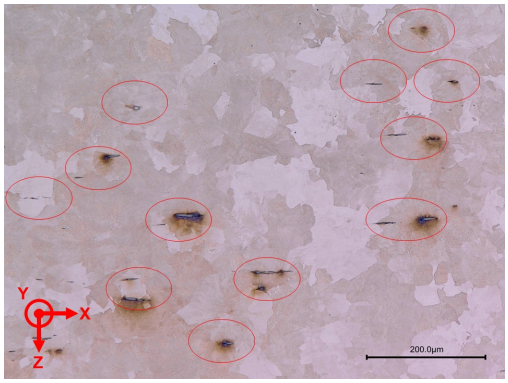


Figure 4.14: *Longitudinal cross section of the fully pearlitic bulk microstructure. MnS inclusions are encircled.*

4.3.2 Surface microstructure of defectless rail section

To get a reference for the microstructure at the surface without defects, a cross section of a defectless part of the rail is evaluated first. Figure 4.15a shows a transverse cross section of the gauge side of the running surface. The orange double-arrows indicate the edge of the running surface and the boundaries between grinding facets. The red boxes show the location of the details in figure 4.15b through d. Figure 4.15b shows microstructure at the surface of facet 1. The pearlite grains at the surface are plastically deformed in the direction of the gauge side. This is the result of the transverse shear stresses induced by wheel-rail contact. This deformation is present along all of the surface of facet 1. A small crack can be seen at the surface which is oriented along the direction of deformation. Although not visible in the figure 4.15b small patches of WEL exist on the surface of facet 1.

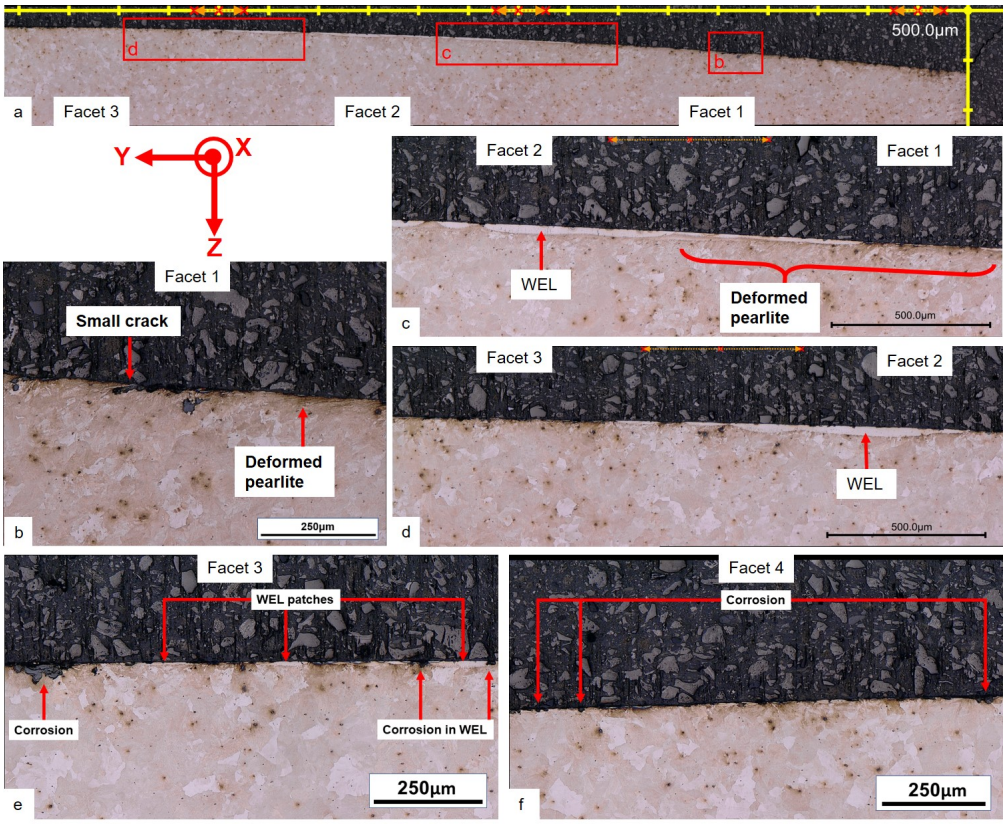
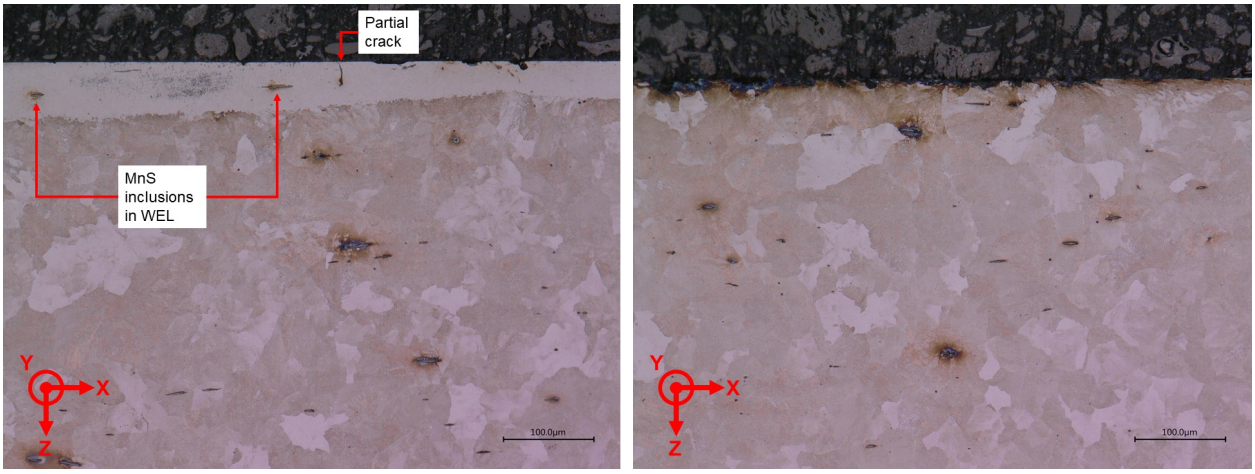


Figure 4.15: *Transverse cross section of the running surface, (a) overview, (b) middle of facet 1, (c) transition between facets 1 and 2, (d) transition between facets 2 and 3, (e) middle of facet 3, (f) middle of facet 4.*

Figure 4.15c shows the boundary between facet 1 and 2. The edge of the WEL coincides with this boundary. The WEL extends in the direction of the field side and covers facet 2 continuously. As indicated in the figure some plastically deformed pearlite is seen underneath the WEL. The thickness of this plastically deformed layer decreases as the thickness of the WEL increases, until no plastic deformation is visible the left of the boundary between facet 1 and 2. Figure 4.15d shows the boundary between facets 2 and 3. The edge of the WEL approximately coincides with the boundary between facet 2 and 3. No plastically deformed layer is visible.

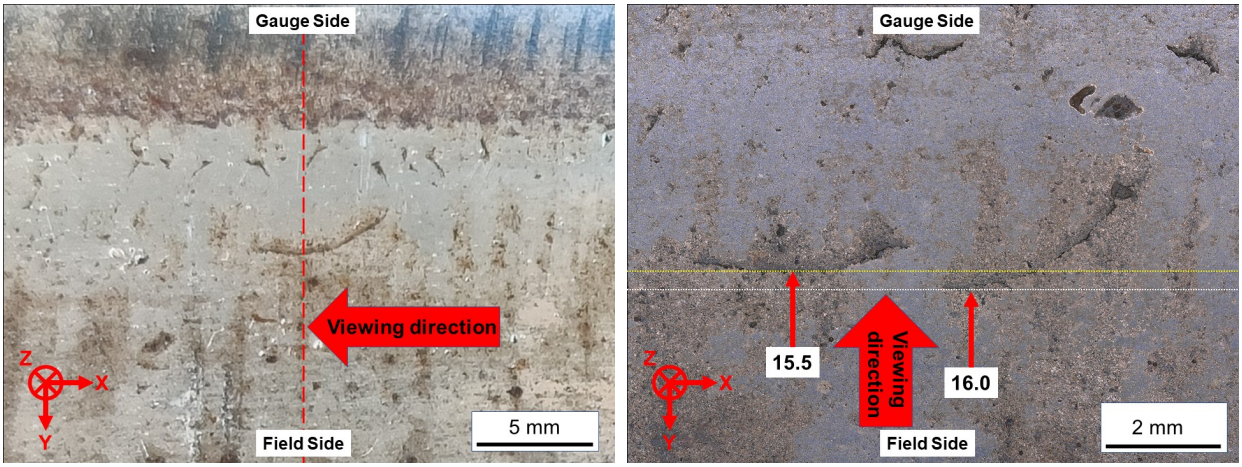
Figure 4.15e shows the surface microstructure in the middle of facet 3. Small non-continuous patches of WEL are visible on the surface. Corrosion has affected the WEL patches as well as the pearlite. Uncorroded WEL surfaces are smooth, the roughness of the corroded parts of the WEL and pearlite appear to have an increased roughness compared to facet 1 and 2. Further towards the field side facet 4 also shows an increased roughness, shown in figure 4.15f. On facet 4 no WEL and deformed pearlite are visible. Approximately 1 mm from the field side edge of the running face the surface roughness increases gradually until the edge of the running surface where the surface is completely covered with an oxide layer.

Two sections of a defectless longitudinal cross section of facet 2 is shown in figure 4.16. Similar to figure 4.15 large parts of facet 2 are covered with WEL. Some MnS inclusions as well as a partial vertical crack are visible



(a) **(b)**

Figure 4.16: *Longitudinal cross section of facet 2 (a) part with WEL, (b) part without WEL.*



(a) **(b)**

Figure 4.17: *Overview of the surfaces of category A defects from rail section 2.0: (a) 11.5 cm and (b) 15.5 cm and 16.0 cm. The arrows and striped lines indicate the viewing direction and the location of the cross section.*

in the WEL. Contrary to the results of Steenbergen and Dollevoet [77] for rails subjected to unidirectional traffic no clear deformation in either of the two possible travel directions is observed, neither for the pearlite underneath the WEL nor the pearlite directly at the surface.

4.3.3 Category A defect

Figure 4.17 shows the surfaces of category A defects 11.5, 15.5 and 16.0. All cracks are contained within the second grinding facet. Some small cracks are present in the first grinding band above the crack. In 4.17a the transverse cross-section is positioned such that it captures the parts of the crack nearest to the field side. The longitudinal cross-sections presented in 4.17b are positioned such that they capture the part of the cracks nearest to the field side.

Transverse cross section

Figure 4.18a shows an overview of the three cracks found in the transverse cross section 11.5 mid as indicated in figure 4.17a. Figure 4.18b shows the subsurface crack propagating from the U crack. It starts at an angle of 30° with the surface, until it reaches a depth of 190 µm. From thereon it continues under an 22° angle with the surface until a depth of 540 µm, where it changes direction again to a 40° angle. The crack tip is located at a distance of 675 µm from the surface and the horizontal distance between the crack tip and mouth is 1440 µm. As expected the crack mouth is located near the edge between facet 2 and 3. Corrosion has occurred

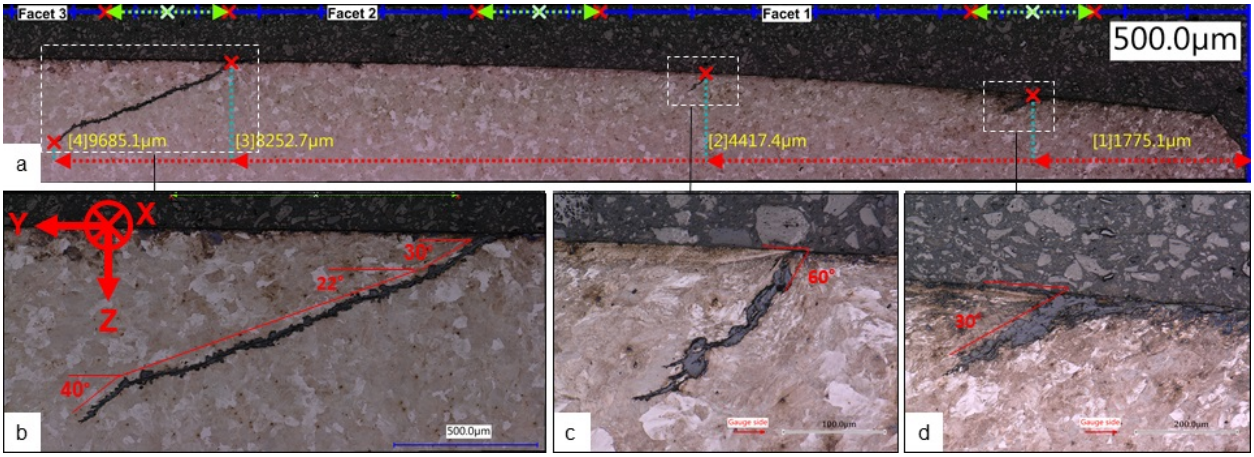


Figure 4.18: *Transverse cross section 11.5 mid from fig 4.17a. Gauge side is to the right. (a) Overview of the first and second grinding facets. The green arrows indicate the transition area between facets. (b) Crack propagating from the U crack. (c) Small crack in facet 1. (d) Crack on the edge of the running surface.*

on both crack flanks, which filled the crack with oxides. Due to the rough interface between the oxide layer and pearlite it appears that the oxidation has transformed individual pearlite colonies, rather than growing uniformly. After the turn from 22° to 40° the thickness oxide layers appears to decline.

In figure 4.18c a small crack in facet 1 is shown. Near the surface this crack grows at an angle of 60° with the surface, but this angle gradually reduces with increasing distance from the surface. This angle does not match the angle of the direction of the deformed pearlite in the surface layer. The crack mouth is located at the edge of a WEL patch on the surface, which is the most likely the point of initiation. Again, the crack is filled with a non-uniform layer of oxide. A small branch can be seen. Figure 4.18d shows the magnification of the crack on the gauge side edge of the running surface. The crack makes an angle with the surface of approximately 30°, but due to the build-up of oxides in the crack it is difficult to determine exactly. This angle does match the direction of plastic deformation. However, to the left of the crack mouth a WEL/BEL can be seen. Therefore it is not clear if this crack is initiated by ratcheting or delamination of the WEL/BEL patch or both.

When figure 4.18 is compared to the defect free cross section in figure 4.15, a difference in surface condition on the first two grinding facets is observed. Where in figure 4.15 only a few small patches of WEL were present on facet 1, in figure 4.18 facet 1 is mostly covered with WEL. The opposite is true for facet 2 as in figure 4.18 no WEL is visible on facet 2.

Longitudinal cross section

The positions of the longitudinal cross sections of defects 15.5 and 16.0 are indicated in figure 4.17b. The cross sections are positioned such that they show the part of the surface crack that is nearest to the field side. Figure 4.19 shows the longitudinal cross section of defect 15.5. Like the surface crack, the shape of subsurface is symmetrical with respect to the transverse direction. When compared to the cracks seen in the transverse cross sections this crack has reached a relatively shallow maximum depth of 220 µm. The longitudinal distance from tip to tip is 3230 µm, which matches the length of the surface crack. Figure 4.19b shows the left crack tip. Near the crack tip the crack is narrow, but the crack width increases in the direction of the crack mouth. The material between the surface and the crack has discoloured more than the material under the crack, especially around the MnS inclusions. Figure 4.19c shows one of the two pieces of material above the crack near the surface breaking part of the cross section. The tip is heavily discoloured and out of focus, which indicates that it lies beyond the plane of the cross section. The material at these tips has a small thickness because the surface crack is located behind them. Therefore the tips are most likely bent backwards during sanding and/or etching.

Figure 4.20 shows the longitudinal cross section of defect 16.0. In this cross section the crack has a longitudinal length of 3000 µm and reaches a maximum depth of 537 µm at the crack tip. The surface crack of defect 16.0, as shown in figure 4.17b, is not positioned symmetrically with respect to the transverse direction. This asymmetry is also visble in the subsurface crack, as only the right crack flank has propagated. At the crack mouth there is a gap in the surface with a width of 54 µm and a depth of 55 µm. This gap is not filled with embedding material. Therefore it must have been filled with material before sample embedding. The missing

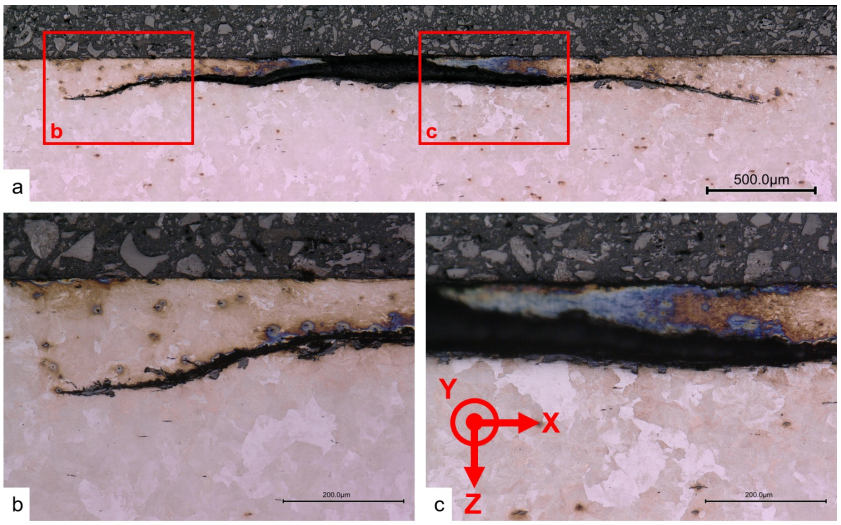


Figure 4.19: *Longitudinal cross section of defect 15,5 (a) overview (b) close-up left crack tip (c) edge of crack mouth.*

material is probably removed during sample preparation. The crack has propagated from the lower right corner of this gap at a 10° angle. After the crack reached a depth of approximately 140 µm it continued propagating with a meandering path as indicated in 4.20a. Similar to all other cracks, the crack width decreases with increasing distance from the crack mouth. Oxidation products have built up in the crack. On the bottom crack surface the oxide is attached to the pearlite beneath, but on the top crack surface the oxide has detached resulting in a rugged surface profile.

In the close up of the left side of the gap in the surface of defect 16,0 in figure 4.20b a small crack can be seen originating from the lower left corner of the gap. The surface to the left of the gap is covered with WEL. The pearlite underneath the WEL directly left of the gap shows some deformation. Further left the WEL increases in thickness and the pearlite underneath it does not show deformation. Figure 4.20c shows the right side of the gap in the surface of defect 16.0. Similar to defect 15.5 the pearlite above the crack shows a larger discolouration than the bulk. The tip of this discoloured area is in focus and therefore is not bent backwards by sample preparation. It is possible that the tip of this part has broken off as the tip shape in figure 4.20c and 4.19c are different.

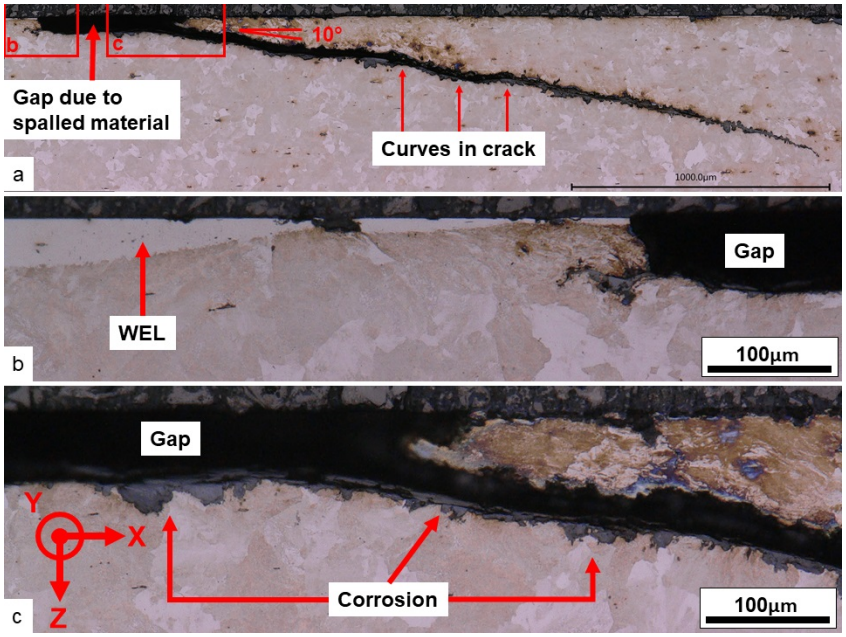


Figure 4.20: *Longitudinal cross section of defect 16,0 (a) overview (b) close-up of left gap corner (c) close-up right gap corner and crack mouth.*

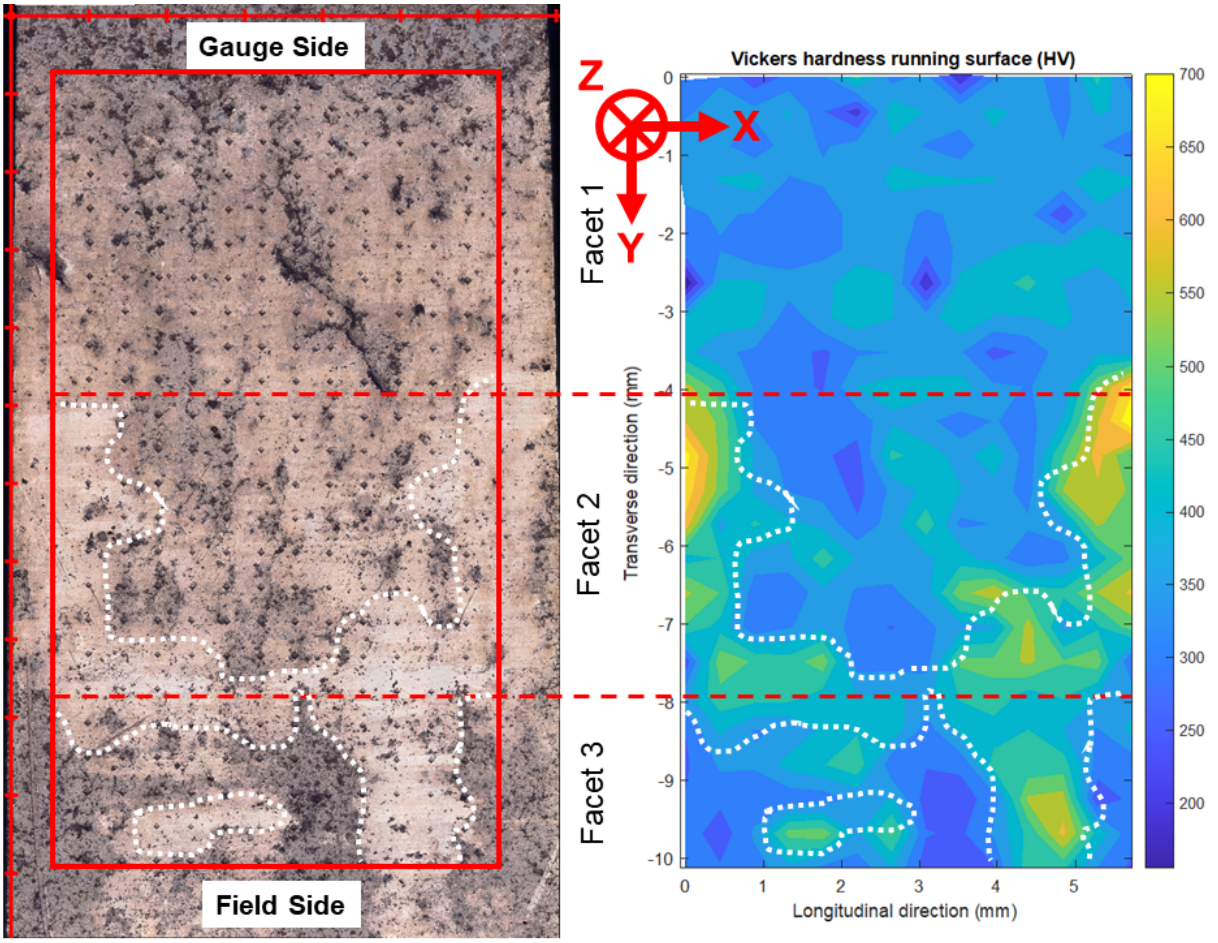


Figure 4.21: *Running surface etched with Nital 2% (left) and its Vickers hardness (HV3) (right). The red rectangle indicates the edge of the measured area. The areas where WEL is observed and measured are indicated with the dotted white line.*

4.3.4 Surface structure and hardness

On the left side of figure 4.21 a close up of the running surface is shown. Prior to the capture of the image, the rail surface was etched with Nital 2%. Etching with Nital reveals the areas on the running surface which are covered in WEL. WEL is known to have a high hardness. To get an estimate of the layer thickness of the white etching material a grid of vickers hardness measurements is made. The results of the hardness measurements are presented in the contour plot on the right of figure 4.21. If the image of the etched surface and the hardness contour plot are compared it is clear that the hardness is higher is areas which are affected most by etching. The WEL appears to be primarily concentrated on facet 2. This is similar to the transverse cross section of a defectless part of the rail presented in figure 4.15. Facet 1 does not contain any large patches of WEL. The average hardness across facet 1 is 372 HV, which is equal to 342 HBW. This is higher than the hardness specified for the centerline of the rail in the norm [28]. Facet 1 contains a crack of approximately 1.5 mm. As mentioned in 4.1.2, facet 1 is strewn with small cracks. A similar crack is found in the transverse cross section of 4.18. Some parts of the surface on facet 3 are grey/black as a result of corrosion. Due to the corroded surface layer these areas have a much lower hardness than the non-corroded parts of the surface.

In figure 4.22 a close-up of the area indicated in the COLOR rectangle is shown. Three areas where the surfaces is covered with WEL are indicated. The surface of the WEL displays a stripe like pattern in longitudinal direction. This pattern is oriented perpendicular to the grinding groove pattern left behind by grinding maintenance.

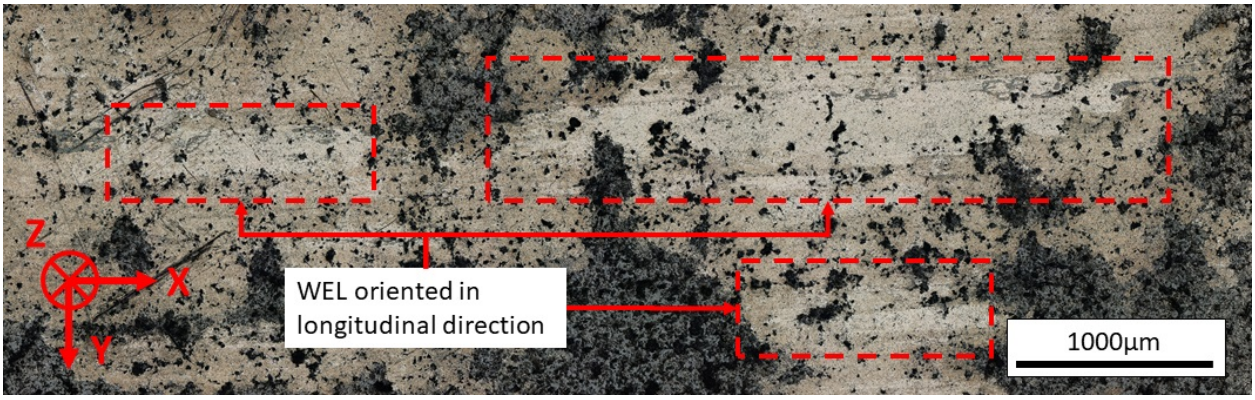


Figure 4.22: *Orientation of WEL patches on the running surface. The running surface is etched with Nital 2%.*

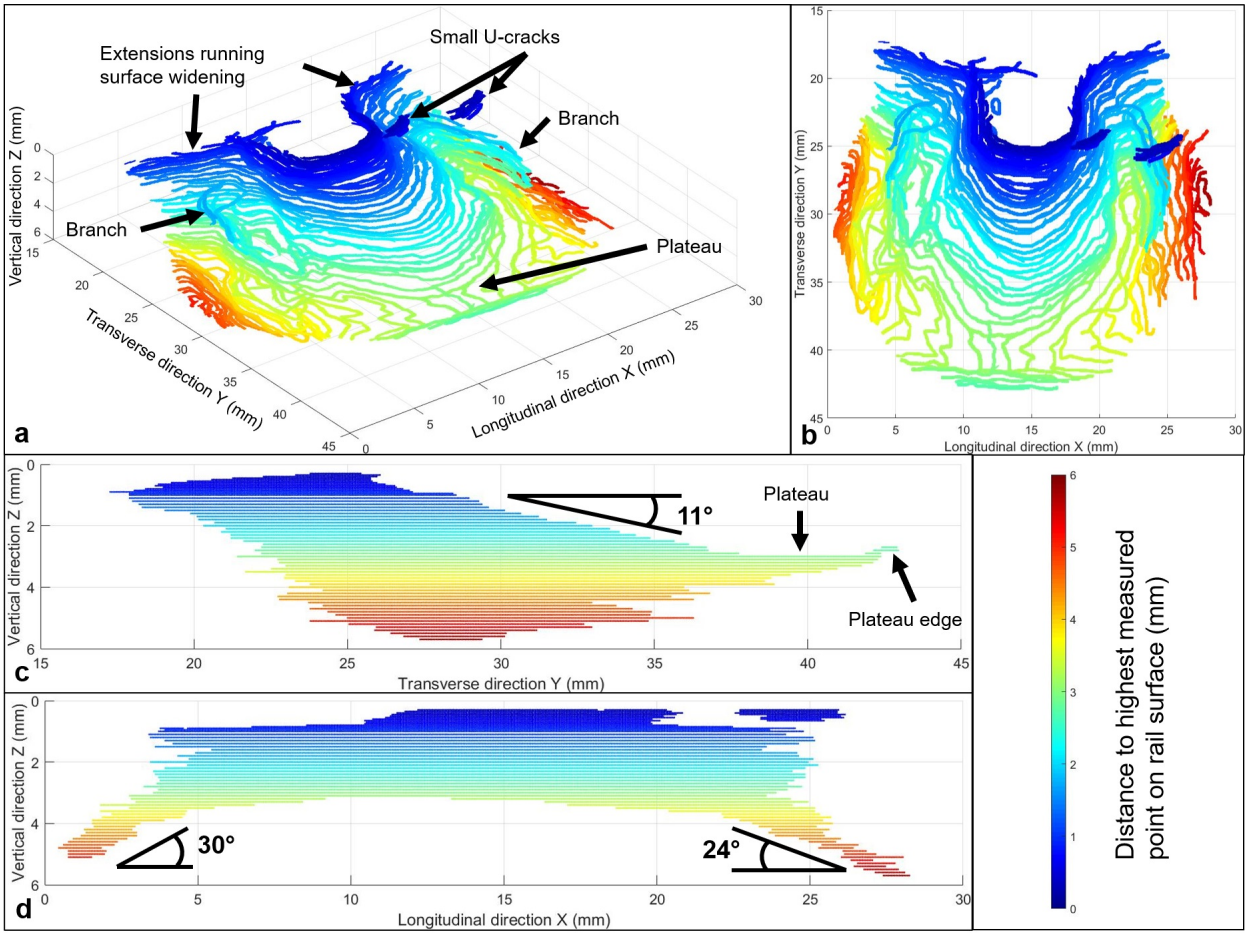


Figure 4.23: *3D model subsurface crack contour C-category defect. (a) isometric view, (b) top view, (c) transverse view, (d) longitudinal view.*

4.4 3D model crack morphology of a category C defect

Figure 4.23 shows the 3D model of the subsurface crack contour of the defect shown in figure 4.2c. The height where the mill first touched the rail surface is set to zero in the vertical direction. In the transverse direction zero is set at the gauge corner, see figure 4.1. The crack contour in the horizontal slices shows certain characteristics depending on the distance below the surface. Depending on those characteristics nine ranges are identified.

0 to 0.25 mm: during the first few milling passes no crack containing material is removed. However, milling showed a depression in the surface around the defect.

0.25 to 0.8 mm: at a depth of 0.25 mm the U-shape crack contour first appears on the milled surface. With increasing depth the U-shape widens in longitudinal direction and the bottom of the U-shape gradually propagates down towards the field side, under an angle of 11°, as indicated by the yellow arrows in figure 4.24. The crack tips of the U-shape extend to the gauge side, as indicated by the blue arrows in figure 4.24. To the side of the category C U-shaped crack, two other category A U-cracks are located. These were not spotted during the visual inspection, due to the dark spots at the surface. These U cracks both reach a maximum depth of 0.65 mm.

0.8 to 1.6 mm: at a depth of 0.8 mm the running surface is completely removed from the rail. In this depth range the crack tips of the U-shape extends outward in longitudinal direction, as indicated by the blue arrows in 4.24. These longitudinal extensions follow the contour of the running face widening, which is indicated in figure 4.2c. The U-shape itself continues to widen in longitudinal direction and the bottom also continues to move towards the field side.

1.6 to 2.6 mm: at a depth of 1.7 mm two crack branches originate from the corners between the U-shape sides and the longitudinal extensions. With increasing depth the tips of these branches extend towards the field side. At a depth of 2.6 mm the branch at the left side connects back to the side of the U-shape. The U-shape itself again widens longitudinally and moves towards the field side.

2.6 to 2.9 mm: at a depth of 2.6 mm the field side edge of the plateau becomes visible in the horizontal cross section. The edge is located at a distance of 43 mm from the gauge corner, which coincides with the field side edge of the running surface. The field side edge of the plateau is located 0.3 mm above the plateau itself, which is clearly visible in figure 4.23c. The plateau edge extends in longitudinal direction.

2.9 to 3.1 mm: at this depth range the transverse propagation direction levels off until it coincides with the edge of the running surface, creating a plateau. At this depth the crack front stops propagating in the transverse direction and splits up into two crack fronts propagation in opposing longitudinal directions.

3.1 to 3.7 mm: the two crack fronts are completely separated at a depth of 3.1 mm. Towards the field side they still contain some kinks related to the U-shape. With increasing depth these kinks are straightened out, until to two parallel crack fronts remain. These crack fronts propagate in opposite longitudinal direction under an angle of 30°.

3.7 to 5.1 mm: the two crack fronts propagate further down into the rail material. With increasing depth they get narrower in lateral direction. The left crack front reaches its maximum depth at 5.1 mm.

5.1 to 5.7 mm: after the left crack front has reached it maximum depth, the right front continues to propagate, until it reaches its maximum depth at 5.7 mm. The final milling pass at a depth of 5.8 mm showed no cracks.

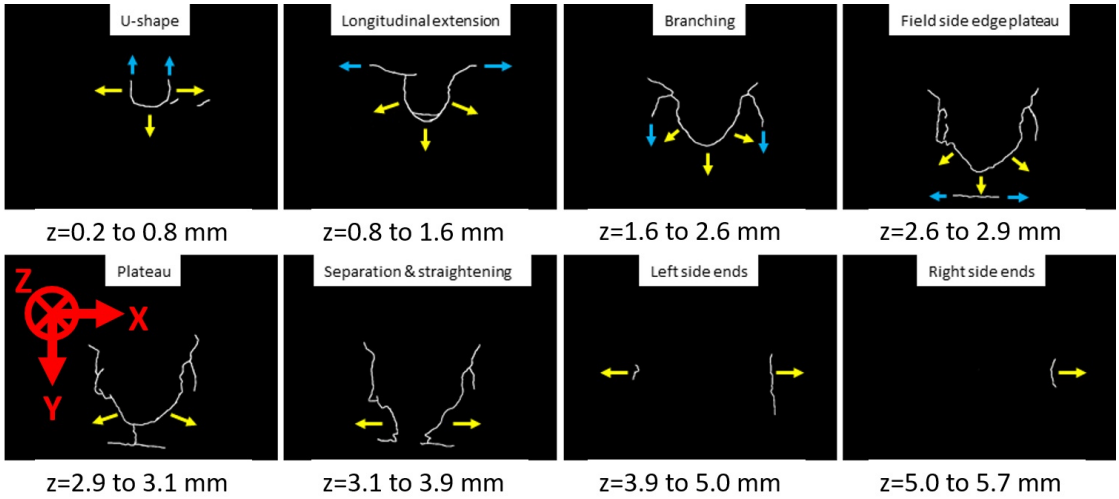


Figure 4.24: *Characteristic shapes of the crack contour in the xy-plane at multiple depth ranges. Yellow arrows indicate crack front propagation direction, blue arrows indicate the crack tip propagation direction and z indicates the distance to the highest measured point on the surface.*

Chapter 5: Discussion

The discussion consists of two parts. The first part discusses the effect of bidirectional traffic on the microstructure and mechanical properties of the rail material. In the second, part the process of RCF crack formation is discussed in the context of the material microstructure and mechanical properties affected by the bidirectional traffic.

5.1 The effects of bi-directional traffic on the microstructure and mechanical properties of the rail material

5.1.1 Longitudinal deformation

In literature, unidirectional plastic deformation opposite to the running direction (ratcheting) has been observed many times [71][72][76][77][78][79][92][96] and is widely considered as one of the primary causes for crack initiation and propagation [71][72][76][77][80]. The development of a ratcheting strain is the result of repeated rolling contact in the same direction [53][59][72]. In all of the above mentioned reports, the rail material was subjected to unidirectional traffic. However, this is not the case in the present study. In the longitudinal cross-section of figure 4.16 it is observed that the surface layer of the rail does not contain unidirectional plastic deformation. Tyfour and Beynon [105] found a similar result studying the effects of rolling direction reversal in a twin disk testing device; the accumulation of unidirectional strain significantly decreased when the rolling direction is reversed at a higher frequency, as seen in figure 2.28. However, these results cannot be directly compared to the results in this study. Tyfour and Beynon [105] ran the experiment for a total of 40.000 cycles, reversing the rolling direction every 2500 cycles. This number of cycles is much larger than the number of cycles per direction reversal in the present work. Each time a train passes by it travels in the opposite direction as the previous train. LINT® trains have six axles, of which four are driven. Therefore the rails experience only six cycles per rolling direction.

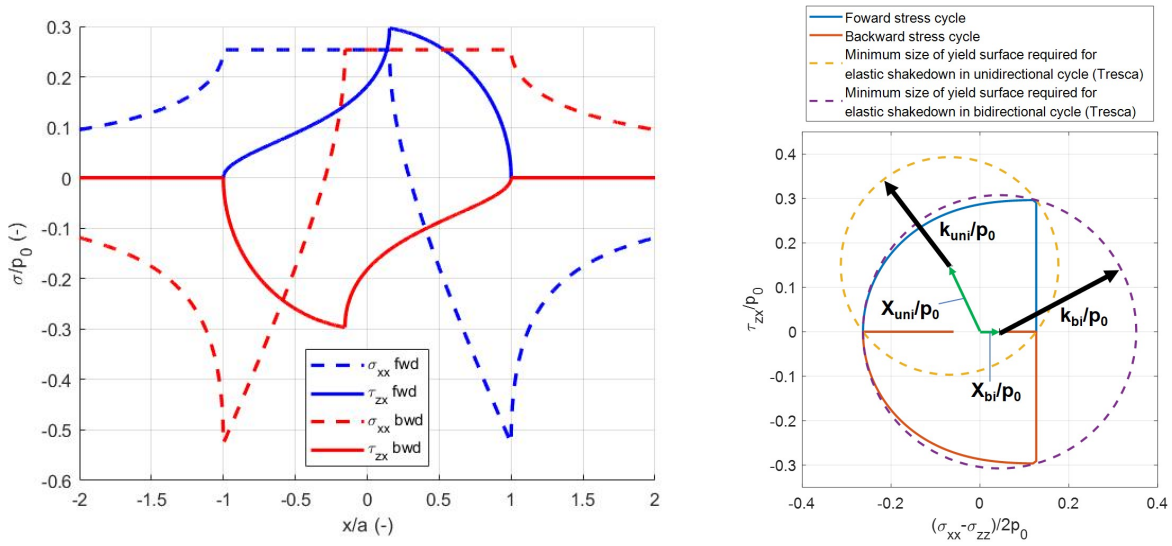
When the rolling direction is reversed, a material element in the rail experiences a reversed cycle of contact stresses in longitudinal direction. We will examine the effect of rolling direction reversal on the plasticity of the rail material in longitudinal direction. We assume that the rail material behaves according to the non-linear kinematic hardening law proposed by Bower [52], which has experimentally been shown to predict longitudinal ratcheting of rail material. First, we will discuss the effect of rolling reversal on the material response to cyclic stress. We will assume that the wheel-rail contact can be approximated by a cylinder rolling on an elastic half-space. Two cases of cyclic stress are considered: low traction (μ<0.3), where the maximum stress occurs subsurface, and high traction (μ>0), where the maximum stress occurs at the surface. Then the contact stresses experienced by the studied rail specimen will be calculated based on the calculated contact parameters of the VAMPIRE® wheel-rail contact simulation. From the calculated stresses the strain response of the material is discussed.

Maximum stress subsurface

When the maximum stress is located beneath the surface, the shakedown limit for is determined by the cycle of orthogonal shear stress τ_{zx} [43]. When the rolling direction is reversed the stress cycle as a function of position, as shown in figure 2.10b, is mirrored in the vertical axis. The cycles of axial stress σ_x, σ_y, and σ_z do not change for a reversed cycle. However, the cycle of orthogonal shear stress τ_{zx} does change. As a result of the reversed cycle, τ_{zx} changes sign at every point. If it is assumed that the rail material obeys Bower's [52] non-linear kinematic hardening law, rolling direction reversal does not change plastic behaviour. The reversed cycle of shear stress remains symmetric with a mean stress of zero. Therefore the material will behave as if it is linearly kinematic. Thus, when the elastic shakedown limit is exceeded, the material will be subjected to cyclic plastic deformation, but no unidirectional strain will be accumulated. This cyclic plastic deformation occurs at a depth of $0.5a$ [43], where a is the radius of the contact patch. The area of the contact patch calculated in the wheel-rail contact simulation lies between 50 and 150 mm^2, so $2 \leq a \leq 3.5$ mm. Therefore, if the elastic shakedown limit is exceeded, subsurface plastic deformation would occur at a depth of 2 to 3.5 mm beneath the surface, depending on the wheel load and contact position.

Maximum stress at surface

At the surface, the effects of rolling direction reversal are more significant than beneath the surface. In figure 5.1a the contact stresses at the surface for two loading cycles in opposite directions are presented. The stresses



(a) **(b)**

Figure 5.1: *(a) Deviatoric surface stresses for a line contact rolling with a tangential traction $Qx = 0.2P$ and $\mu = 0.3$ in forwards (fwd) and backwards (bwd) directions as function of contact position, (b) the trajectory of τ_1 stress space. The radii of the yield surfaces indicate the minimum shear yield strength k required for elastic shakedown. X_{uni} and X_{bi} indicate the required back-stresses for the unidirectional and bidirectional cycles, respectively.*

in the forward cycle are equal to those in figure 2.12b. In the backward cycle, the stress trajectories are mirrored along the vertical axis. Due to the reversed traction direction, the shear stress acts in the opposite direction. For the assumed case of a cylinder rolling on a half-plane, the material can shake down such that all subsequent stress cycles result in an elastic material response, as long as the trajectory principal shear stress τ_1 can be contained within the yield surface. When plotted in figure 5.1a this yield surface has the shape of a circle with a radius of k, where k is the Tresca shear yield strength [47][51]. As result of the non-linear kinematic behaviour, yield surface is free to move in stress space. The vector from the origin to the centre of the yield surface is the back-stress X.

$$\tau_1 = (\frac{1}{4}(\sigma_{xx} - \sigma_{zz})^2 + \tau_{zx}^2)^{\frac{1}{2}} - X \leq k \tag{5.1}$$

To examine how bidirectional traffic affects the shakedown behaviour, the trajectories of τ_1 are plotted in figure 5.1b for a forward and backward stress cycle. The shakedown limit is exceeded if the stress trajectory cannot be contained within a circle with a radius k/p_0. The two circles plotted in figure 5.1b represent the two yield surfaces that just contain a unidirectional and a bidirectional cycle for constant μ and Q_x. From figure 5.1b, it is obvious that for bidirectional rolling a larger yield strength k is required to contain the trajectory of τ_1 within the yield surface than for unidirectional rolling. Therefore the bidirectional elastic shakedown limit will be exceeded before the unidirectional elastic shakedown limit if the material and load are equal.

Example with simulated contact parameters

We will now examine the material response of the rail material studied in the present work. The deviatoric stresses $\sigma_{xx} - \sigma_{zz}$ and τ_{zx} will be calculated based on simulated contact parameters at facet 1 of simulation run NWCI, as indicated in table 3.3. Figure 5.2 shows the calculated

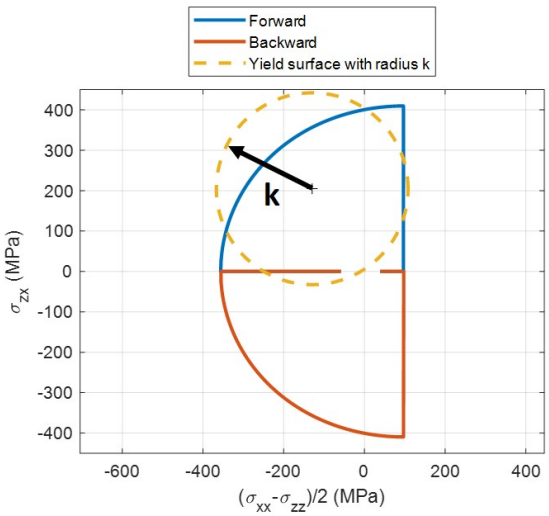


Figure 5.2: *Stress trajectories of forwards and backwards cycle. The P = 70 kN, Q_x = 7.5 kN, μ = 0.32, p_0 = 1650 MPa, a = 4.4 mm and k = 237.5 MPa.*

trajectories of τ_1 for a forward and backward cycle, together with the yield surface. The size of the yield surface is based on the value of the yield stress of R260Mn rail steel reported by Steenbergen [96]. Figure 5.2 shows that trajectory of a unidirectional and a bidirectional cycle both cannot be contained within the yield surface. Therefore, the surface material of the rail will be plastically deformed during each loading cycle. It must be noted that the calculated trajectories of τ_1 as shown in figure 5.2 are based on the contact conditions of the most severely stress part of the running surface, for the wheel-rail profile combination which resulted in the larges contact loads. On other parts of the running surface, and for other wheel-rail profile combinations, the stresses are less severe. However, if wheel-rail contact occurs in completely dry conditions, the friction coefficient μ can be as high as 0.6, increasing the contact stresses.

Plastic deformation due to exceedance of the shakedown limit

When the principal shear stress τ_1 exceeds the yield surface, plastic deformation will occur. A strain hardening material strains until it carries the applied load. Bower en Johnson [47][53] assume that the stresses at the onset of plastic deformation can be approximated by the elastic (i.e. Hertzian) stress field, given that the plastically deformed layer is thin. Furthermore, is is assumed that longitudinal and lateral strain ϵ_x and ϵ_y are prevented by the creation of residual stresses $\sigma_{xx,r}$ en $\sigma_{yy,r}$. As a result, ϵ_{zx} is the only accumulated strain. It is experimentally shown that with these assumptions the accumulation of plastic shear strain in the surface layer of rail steel can be modelled by the non-linear kinematic hardening law [73] under unidirectional rolling conditions.

It is not clear if this model would also be able to predict the plastic behaviour of rail material subjected to bidirectional rolling. When the yield stress is exceeded, the yield surface, as shown in figure 5.2, will translate in stress space. This translation is caused by a change in the back stress X. For unidirectional rolling, the back stress will only change by a small margin as a result of plastic deformation. However, if the rolling direction is reversed after an initial cycle in forward direction, the back stress, especially its τ_1 component, will change significantly more than for unidirectional rolling. Additional research is needed to investigate the role of the back stress on cyclic plasticity without the accumulation of plastic strain.

5.1.2 Transverse deformation

Despite the lack of plastic deformation in the longitudinal direction on the studied rail specimen, unidirectional strain is observed in lateral direction. In the transverse cross-section of a defectless part of the rail surface in figure 4.15b, the microstructure is strained in the direction of the gauge corner. This deformation is not present over the entire cross-section; it is only observed underneath facet 1 (y = 20 to 24 mm). In literature similar deformations are reported [77][78]. The strain indicates that there must have been a tangential contact force in the gauge side direction when wheel-rail contact occurred on facet 1. There are two sources of lateral tangential load in the contact patch, which are schematically represented in figure 5.3. The first source is the lateral component of the normal load in the wheel-rail contact plane [77]. As a result of the curvature of the rail surface and the transverse inclination of installation α, the rail surface is not perfectly horizontal. Instead, the normal of the contact plane z' is tilted by an angle δ relative to the z-axis. Due to this angle δ, the gravity force vector G decomposes into a normal force $P = G\cos(\delta)$ and a tangential force $Q_{g,y} = G\sin(\delta)$. Due to the dependence on curvature of the rail surface, $Q_{g,y}$ is a function of the lateral position of the wheel-rail contact patch. On facet 1, δ is approximately 3.6° : 1.4° from the inclination of installation and 2.2° from the curvature of the rail surface. This results in a $Q_{g,y}$ of -4.4 $\pm$0.6 kN, due to the variance in the gravity force.

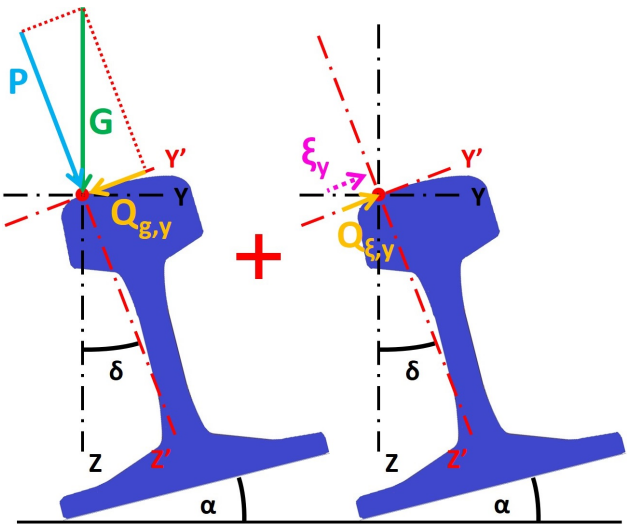


Figure 5.3: *Forces on rail as a result of gravity (left) and creep (right). The transverse inclination of installation α is exaggerated on purpose. ξ_y and $Q_{\xi,y}$ can be oriented in either the positive or negative Y' direction.*

The wheel rail-contact simulation shows that creep is the second source of lateral tangential load in the contact patch. The tangential force on the rail as a result of creep is oriented in the same direction as the creep vector, as described by equation 4.2. However, it is observed that when contact occurs on the gauge side of the running surface (y = 20 to 31 mm), the relation between the lateral creep and tangential force is described by equation 4.3: $Q_{\xi,y} = A\xi_y + B$. From figure 4.13 is determined that B is equal to -5 $\pm$ 2 kN. This result matches the value found for $Q_{g,y}$ at facet 1, which indicates that $B = Q_{g,y}$. Furthermore, the lack of a B term in the relation between the creep and the lateral tangential force for contact positions on field side of the running surface (y = 31 to 42 mm) is consistent with $B = Q_{g,y}$, since δ is smaller on the field side of the running surface due to the curvature of the rail surface. It can be concluded that total lateral tangential force can be expressed as

$$Q_y = Q_{\xi,y} + Q_{g,y} = A\xi_y + \sin(\delta)G \tag{5.2}$$

Figures 4.12 and 4.13 show wheel one and two have a difference creep behaviour on the gauge side of the running surface. For wheel two the lateral creep is primarily positive. As a result the lateral tangential forces $Q_{g,y}$ and $Q_{\xi,y}$ are oriented in opposite directions, which reduces to total lateral tangential force acting on the rail. For wheel one the lateral creep can be both positive and negative. A negative lateral creep creates a force $Q_{\xi,y}$ in the same direction as $Q_{g,y}$. This results in a maximum Q_y of -10 kN. It is therefore expected that the lateral plastic deformation beneath facet 1 is caused by contact with wheel one.

Bidirectional traffic does not influence the plastic deformation in lateral direction. It is shown that both sources of lateral tangential force on the rail are not dependant on the traffic direction. Furthermore, the LINT trains are symmetric in layout of their carriages, which causes the running behaviour to be the same in both directions.

5.1.3 The origin and position of WEL on the running surface

WEL is created as a result of a rapid thermal cycle in the surface layer of the rail. During this cycle, the material is heated above the austenitization temperature and subsequently quenched as the heat is quickly transported away from the surface through conduction and convection [7][68]. Two processes can cause the thermal cycle required for WEL formation: grinding maintenance and wheel slip [7][10][70]. Grinding causes a heat flux into the rail material as a result of the friction between the abrasive and the rail. Immediately after grinding the rail surface is covered with a continuous layer of white etching material [97][96]. Due to the grinding grooves, the surface of the rail has a high roughness. During the initial few gross tons after grinding, the roughness is removed from the surface due to contact with the train wheels. During this process one of two things can happen to the WEL created by grinding: it can either spall off or it is pressed down into the pearlite matrix underneath [97][96]. If the WEL is pressed down into the pearlite matrix it forms transverse WEL bands along the direction of the grinding grooves as shown in figure 5.4. For R260Mn rail steel, it is observed that the WEL created by grinding is removed from the surface by spallation rather than being pressed down into the pearlite matrix [96].

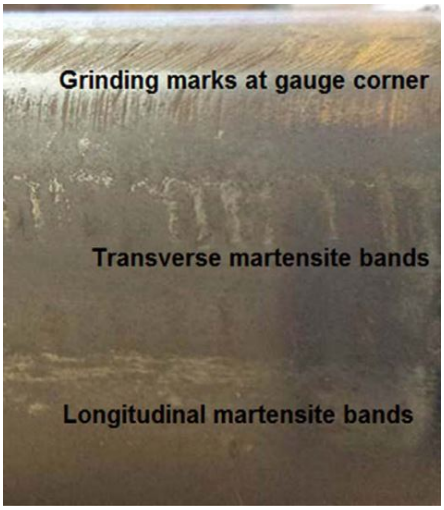


Figure 5.4: *Longitudinal and transverse WEL (martensite) bands on surface of a worn R350HT rail [97]. The area shown has been etched with Nital.*

The heat flux into the rail material caused by wheel slip originates from the frictional heating between the two contacting surfaces [60][63][64]. The primary cause of this slip is the tangential force produced by driving or braking wheels, which results in partial or complete slip in the contact patch [62]. Due to the speed of the train, this frictional heating source moves rapidly over the rail surface in longitudinal direction. If the heat flux into the rail material generated by wheel slip is significant enough to cause material transformation, it is expected that rail material is transformed along a strip in the slip direction. The most significant slip occurs in longitudinal direction as a result of the torque applied to the driven wheels. Therefore areas of WEL created by wheel slip are oriented in longitudinal direction [97].

At first sight, it looks like grinding maintenance is the most dominant influence on the surface appearance; the positions of the grinding facets is visible on the running surface. Within those facets, the remnants of the grinding grooves are visible from the alternating transverse light and dark bands. These bands look very similar to the transverse WEL bands in figure 5.4. However, etching of the running surface shown in figures

4.21 and 4.22 shows that the patches of WEL on the running surface contain a stripe-like pattern oriented in longitudinal direction. This pattern cannot be created by grinding. Therefore the WEL on the running surface must have been formed as a result of wheel slip. This result is consistent with the findings of Steenbergen [96] that WEL created during grinding is removed by wheel contact shortly after the rails are ground.

WEL formation as a result of wheel slip can only occur if the sliding velocity/creep ratio is high enough to reach the austenitization temperature. In literature, multiple models are presented which calculate the surface temperature as a function of the creep ratio [61][62][63][65][67]. The creep ratio is required to reach the austenitization temperature varies substantially between the different models: longitudinal creep ratios of 3% [61] to 26% [62] are reported. A direct comparison between the different models cannot be made, as the models vary in complexity and make different assumptions for parameters such as vehicle speed, coefficient of friction, etc. Despite the variation between models, the calculated creep ratios required for WEL formation are at least 10 times higher than the creep ratio calculated in the VAMPIRE® multibody wheel-rail contact simulation in the present work. Figure 4.12 shows that the maximum creep ratio in longitudinal direction is $\pm$ 0.2%. All models predict that WEL formation cannot occur at such small creep percentages. Therefore some other factors must be at play, which can reduce the coefficient of friction to cause a sliding velocity high enough for the formation of WEL. Two possible candidates could be tree leaves or bird excrement.

It is not expected that bidirectional traffic has an influence on the formation of WEL. WEL formation is only a function of the heat flux into the rail surface, which is not dependant on the slip direction. No mention of rolling direction or slip direction is made in references [7][60][61][62][63][64][65][68].

5.2 The effects of bidirectional traffic on RCF defects

The life of an RCF crack can be divided into three phases: initiation, propagation as a result of contact stresses, and propagation due to bending stresses [3][38]. In the following two paragraphs the effects of bidirectional traffic on the formation of RCF defects are discussed for each of the three phases.

5.2.1 Effect of bidirectional traffic on RCF crack initiation

The microstructure of the rail material contains many inhomogeneities. When the material accumulates plastic strain as a result of wheel contact, these inhomogeneities act as weak sites in the microstructure. The inhomogeneities either lower the strength of the material locally, or induce stress concentrations. Fatigue cracks are prone to initiate at these weak sites in the microstructure. The most common inhomogeneities in rail steels are inclusions [74][75], pro-eutectoid (PE) ferrite - pearlite interfaces, WEL and BEL.

Two types of inclusions exist: ductile and brittle. When the rail material is strained, ductile inclusions elongate in the direction of the strain. Elongation of the ductile inclusions has two effects. Firstly, the tip of the elongated inclusion acts as a notch, which causes a stress concentration in the surrounding pearlite [75]. Secondly, voids can form at the interface between the elongated inclusion and the pearlite matrix, due to interface decoherence. These voids are initiation sites for fatigue cracks [74][75]. Brittle inclusions do not strain when the material is deformed, which causes stress concentration. Furthermore, the brittle inclusion itself can fracture, which creates initiation sites for fatigue cracks [75].

Similar to the crack initiation mechanisms of ductile inclusions, fatigue cracks initiate on a PE ferrite - pearlite interface when the material is heavily strained. The PE ferrite has a lower strength than pearlite. As a consequence, PE ferrite is strained more relative to the pearlite, which results in a relatively larger increase in hardness due to the build-up of dislocations [71][72]. The build-up of dislocations weakens the PE ferrite - pearlite grain boundaries, which facilitates crack initiation [72]. The likelihood of crack initiation along heavily strained PE ferrite grain boundaries depends on the volume fraction of PE ferrite; the larger the PE ferrite volume fraction, the more dominant crack initiation along PE ferrite grain boundaries [72].

Crack initiation in rail material is often linked to WEL [77][78][79][80]. The thermally transformed layers are harder and more brittle than the underlying pearlite [66][82]. Due to their brittleness, they fracture easily. Furthermore, they induce stress concentrations in the underlying pearlite, when they are pressed down into the material due to a wheel load. There are two possible crack initiation mechanisms related to WEL. Firstly, cracks can initiate from transverse brittle fracture of the WEL. The brittle fracture halts at the WEL - pearlite interface, but can continue to propagate into the pearlite due to subsequent wheel loads. Secondly, cracks can initiate at a point where the interface between WEL and pearlite (or BEL) is located at the surface [77][10]. In both cases, the resulting crack propagates along the strain direction of the underlying pearlite.

The fatigue crack initiation mechanisms share a common factor; crack initiation is always paired with a significant accumulation of plastic shear strain in the surface layer of the material. As a result of the bidirectional traffic experienced by the rails in the present work, the longitudinal component of this strain is not present on the rail surface. Due to the lack of strain accumulation, it is expected that ductile inclusions did not play a major role in crack initiation. The experimental results are in line with this expectation. Although MnS inclusions are present throughout the microstructure of the rail, none are observed that are elongated as a result of strain accumulation. PE ferrite likely also did not have a large effect on the initiation of the RCF defects found in the present work. Firstly, the results of optical microscopy do not indicate that PE ferrite is present in the microstructure of the rail. Although the measured carbon concentration of 0.62 w% is lower than the eutectoid composition, a pearlitic microstructure is formed. This is the result of the relatively large fraction of manganese, which shifts the eutectoid point to a lower carbon concentration [5]. Small grains of PE ferrite may exist in the studied rail material, which could not be observed by optical microscopy. However, the effect on fatigue crack initiation would be minimal. Secondly, on the second grinding facet, where all RCF defects are initiated, the material lacks the ratcheting strain required for crack initiation at the PE ferrite - pearlite interface.

On the rail specimen studied in the present work, WEL is observed at multiple locations on the running surface. Despite the lack of accumulated plastic strain, fatigue cracks are observed to initiate as a result of WEL. In figure 4.18c, a crack is found in the surface layer of facet 1, which has initiated at the edge of a WEL. This crack has not propagated far beyond its initiation point, as it has only reached a depth of 150 μm. The cracks in figures 4.18b and 4.19 show no indication of initiation as a result of WEL, since there is no WEL present near the crack mouth. Furthermore, no WEL is found on the surface above the cracks. This stands in contrast to the observations shown in figures 4.15 and 4.21, where it is found that WEL is primarily concentrated on grinding facet two. The lack of WEL at the crack mouths of the category A cracks does not rule out initiation as a result of WEL. It is possible that WEL was present near the crack mouth at the moment of initiation, but has since been removed from the surface as a result of wear.

In the cross-section of figure 4.20, WEL is found near the crack mouth. Unlike other crack mouths, the crack appears to be initiated from a gap in the surface. Before sanding and polishing, this gap must have been filled by rail material, otherwise it would have been filled with embedding material. It is possible that the gap was previously filled with WEL which detached during the sample preparation. On the surface inside the gap, oxides are present. Oxidation can only occur if water was able to reach underneath the WEL while it was still in place. This indicates that there must have existed a crack along the interface between the WEL that was removed during sample preparation and the pearlite below. The crack that has propagated down into the pearlite, must have originated from the crack on the interface between the WEL and pearlite.

5.2.2 The effect of bidirectional traffic on RCF crack propagation and the formation of RCF defects

Once a crack reaches a length in the order of 0.1 mm, the first phase is completed and the crack is considered to be fully initiated [3][38]. At this length, the effects of plastic deformation are no longer the most dominant mechanism governing crack propagation. Instead, the crack now continues to propagate as a result of the contact stresses[3][38]. As discussed in section 2.2.5, the contact subsurface axial contact stresses are all compressive. Fatigue cracks do not propagate as a result of compressive stress. However, the shear stresses generated by the wheel load can cause crack propagation.

For shear-driven crack propagation, it is required that the crack flanks can move relative to each other as depicted in figure 2.20a. However, due to the compressive contact stresses, the crack faces are pressed against each other. Friction between the crack faces hinders their movement. Fluids (primarily rainwater) can enter the crack and reduce the friction between the crack faces, enhancing shear-driven crack propagation [3][38][88]. Furthermore, fluids may be trapped inside the crack and pressurize the crack flanks, inducing tensile stress at the crack tip [3][88]. Fluids have had an effect on crack propagation on the studied rail specimen. The presence of water in the crack caused corrosion, which covered the crack faces with an oxide layer. This oxide layer is observed on all cross-sections of RCF cracks. Generally, this oxide layer is thicker near the crack mouth than near the crack tip, as the crack faces near the crack mouth have been exposed to fluids for a longer period.

Although fluid-assisted shear-driven crack propagation gives an explanation for the mechanism of crack propagation at the crack tip, it not give an explanation for the exact shape, orientation, and position of the RCF defects observed on the rail specimen. The defects observed on the surface of the studied rail specimen possess the three typical characteristics of squat (or stud) defects: (1) a local depression with a dark surface in the shape of two lungs/kidneys, (2) a V- or U-shaped crack in the running surface, with its opening in

the direction of the gauge corner of the rail and (3) a local widening of the running band around the dark spots [77][78][85][91]. In literature, multiple theories are presented which describe the process of squat formation. In the next two paragraphs, two of the most cited theories on squat formation by Steenbergen and Dollevoet [77][93] and Li et al. [92][91] will be discussed. We will examine what the implications of bidirectional traffic are on these theories and compare those to observations in the present work.

According to Li et al. [91][92], squats initiate from an irregularity in the rail surface. This irregularity causes a dynamic oscillation in the normal force on the rail, which causes peak loads. The fixed frequency of this oscillation causes the peak loads to occur at fixed distances from the initial surface irregularity, as indicated in figure 2.21b. The first load peak occurs at the initial irregularity and forms the first dark spot of the squat. The second load peak occurs a fixed distance after the first peak load and creates the second dark spot of the squat. The dark spots are the result of a settlement of the surface, which enables discoloration due to corrosion. Li et al. [91][92] mention that this oscillation in normal force, which is determined by the eigenfrequency of the vehicle-track system, does not depend on the traffic direction. Therefore, the appearance of the squat in a later stage of development will be similar for both travel directions, as long as the traffic direction is not alternated. However, when this theory is applied to the case of bidirectional traffic an inconstancy arises. The first force peak occurs at the location of the initial surface irregularity, irrespective of the traffic direction. As a result, the first lobe of the squat forms at the position of the initial surface irregularity. However, the position of the second force peak does depend on the traffic direction. If the traffic direction is reversed often, it is expected that a second dark spot will simultaneously form on either side of the first dark spot, resulting in a total of three dark spots. This prediction does not match the observations in the current study; no defects are found with more than two dark spots. Li et al. [91][92] do not take into account the cracks that form underneath the surface of a squat and report that cracks may or may not be present during initiation. No observation in the present study indicates that a dark spot without an accompanying crack could initiate a squat defect.

In contrast to the theory proposed by Li et al. [91][92], Steenbergen and Dollevoet [77][93] take on a more metallurgical a fractographic approach in the analysis of squats. According to Steenbergen and Dollevoet [77][93], the formation of a squat starts with the initiation of small cracks from fracture or delamination of WEL patches on the rail surface. These small cracks link up perpendicular to the direction of the principal tensile stress resulting from the tangential shear stresses τ_{zx} and τ_{zy} as depicted in figure 2.23b. The resulting surface-breaking crack propagates down into the rail along the texture of the plastically deformed surface layer. It propagates by fatigue in the direction indicated in figure 2.23a. The crack always propagates in the traffic direction and is therefore deemed the leading crack. Once the leading crack is initiated, a trailing crack may initiate. This trailing crack branches off from the leading crack as a result of brittle fracture. If the postulated squat formation mechanism is applied to the case of bidirectional traffic, one issue arises immediately; the definition of direction and propagation mechanism of the leading and trailing cracks lose their meaning in the context of an alternating traffic direction. As discussed in paragraph 5.1.1 and 5.1.2, bidirectional traffic alters the direction of τ_{zx}, while the direction of τ_{zy} stays constant. As a result, a leading crack can form in either of the two principal tensile stress directions. Furthermore, the longitudinal component of the plastically deformed surface layer does not exist in the case of bidirectional traffic, thus the initiated crack cannot propagate along the deformed texture.

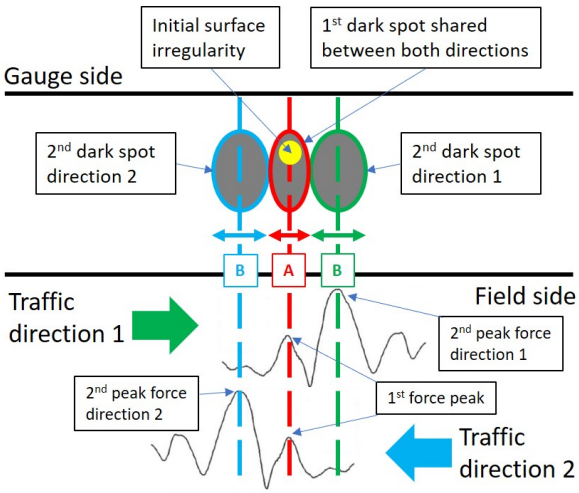
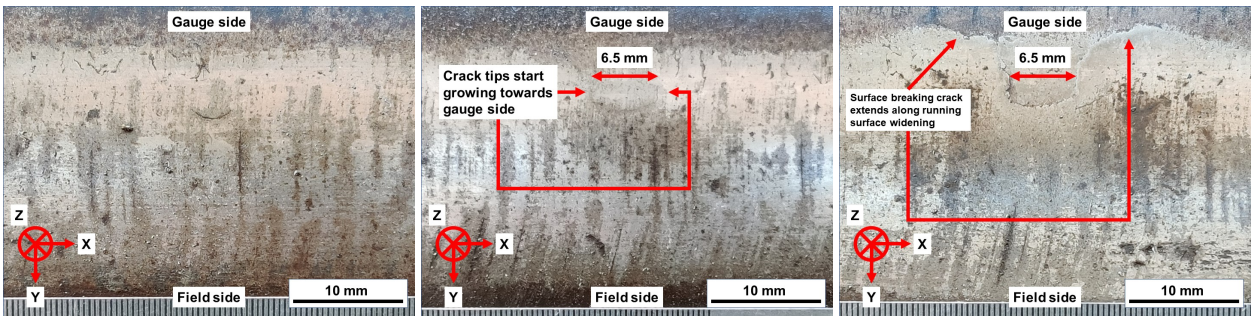


Figure 5.5: *Schematic representation of the effect of bidirectional traffic on the theory on squat formation by Li et al. [91][92].*

Despite the inconsistencies that arise if the theory of Steenbergen and Dollevoet [77][93] is applied to the case of an alternating traffic direction, similarities are observed between the crack morphology predicted by the theory and some of the RCF defects observed in the present study; the slash-cracks observed on the rail surface of the studied rail specimen are oriented at the same 45° angle as predicted by Steenbergen and Dollevoet [77][93]. These slash-cracks are observed in two different orientations (\ and /), which is consistent with the discussed effect of the bidirectional traffic on crack initiation. The possibility still exists



(a) **(b)** **(c)**

Figure 5.6: *The three stages in the development of a U-crack: (a) category A crack, surface breaking U-crack in longitudinal direction, (b) category B crack, after longitudinal length of 6.5 mm is reached the crack tips propagate towards the gauge side, (c) category C crack, crack tips propagated along the widened running surface.*

that a branch initiates from the initial slash-crack. The initiation of a branch turns the slash-crack into a V-shape crack, which is also observed on the rail surface. Due to the alternating traffic direction, the branch will alternate between the role of leading and trailing crack. The propagation mechanism of this branch should therefore alternate between the fatigue propagation mechanism and the brittle failure mechanism. The alternating propagation mechanisms should result in a fracture surface that contains alternating areas with fatigue striations and brittle cleavage.

Besides the slash and V-cracks, another crack shape is observed on the rail surface: U-cracks. Unlike the (bi)linear surface-breaking cracks of slash and V-cracks, the surface-breaking U-cracks are slightly curved. Multiple U-cracks of varying sizes are observed. Category A U-cracks, as shown in figure 5.6a, have a tip-to-tip length between 3 to 5.5 mm. The transverse cross-section in figure 4.18 shows that the subsurface crack front propagates towards the field side of the rail under an angle of 20° with the rail surface. The category A U-crack continues to propagate until a tip-to-tip distance of 6.5 mm is reached. At this critical distance size, the surface-breaking crack stops propagation in the longitudinal direction and instead continues propagating towards the gauge side of the rail, creating the sides of the U-shape as indicated in figure 5.6b. While the sides of the U-crack develop, the subsurface crack front has propagated to a large enough depth to cause a settlement of the rail material above the crack. As a result, the rail surface no longer makes contact with train wheels, which enables a dark spot to form due to corrosion. Due to the surface settlement the curvature of the rail surface changes, which shifts the contact patch to the gauge side of the rail and causes the running surface to widen. Once the tips of the surface-breaking crack reach the edge of the running surface, they stop propagating directly towards the gauge side follow the edge of the widened running surface as shown in figure 5.6c.

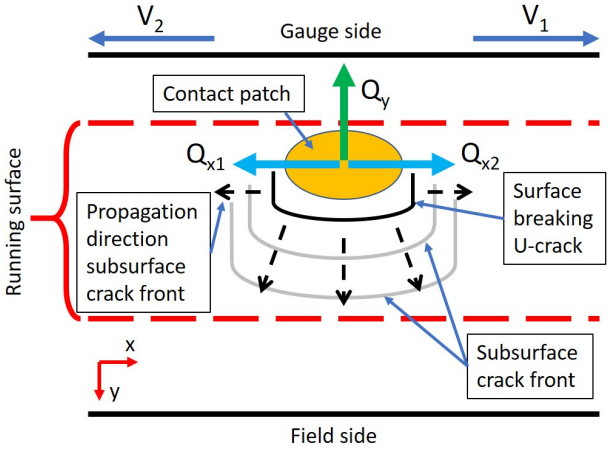


Figure 5.7: *Schematic representation of subsurface propagation of a U-crack*

U-shaped cracks are also observed by Deng et al. [85], who observed that U-shape squats originate from a settlement in the running surface. Since the settled surface does not bear any contact load, a stress concentration is created at the gauge side edge of the settlement. Once the settlement reaches a certain critical size, a crack is initiated as a result of this stress concentration. This initiation mechanism appears to contradict the observed category A U-cracks in the present work, which are found without an accompanying dark spot. However, Deng et al. [85] observed that after grinding maintenance a part of some U-cracks that had initiated from dark spots remained in the surface, while the initial dark spot was completely removed by grinding, leaving just a U-crack behind. Subsequent wheel loads caused the dark spot to reappear at the field side of the U-crack. If this is also the case for the rail specimen in the present study, it would imply that all U-cracks have been initiated before the grinding maintenance in Q1 2017.

Below the surface, the U-crack propagates as a result of the lateral forces on the rail surface that are generated when the contact patch is located on the gauge side of the running surface. This is schematically represented in figure 5.7. When wheel-rail contact occurs on the gauge side of the rail, the wheel-rail contact simulation shows that Q_y is always oriented in the direction of the gauge side and Q_x is oriented opposite to the traffic direction. The bottom centre of the U-crack propagates towards the field side of the rail, opposite the direction of Q_y. The opposition of Q_y and the crack propagation direction is consistent with the shear-driven crack propagation mechanism [3][38]. As shown in figure 4.4, the crack front propagates down into the rail at an inclination of 11° with the rail surface, until a depth of 3 mm is reached. Here the propagation direction levels off and continues to propagate parallel to the rail surface, creating a plateau. The plateau continues to propagate parallel to the rail surface until it the crack front is located directly below the field side edge of the running surface. Other studies [78][108] have found similar crack plateaus at this depth. Simon et al. [78] report that at this depth a transition occurs in the stress state in the rail. At a depth range of 0 to 3 mm, the contact stresses are dominant and control crack propagation. Beyond a depth of 3 mm, the influence of the contact force decreases, and the bending stress becomes the most dominant stress. This is consistent with the Hertzian contact theory [42][44], which predicts a subsurface stress maximum at a depth of $0.5a$, where a is the contact patch radius. The wheel-rail contact simulation showed that the area of the contact patch varies between 50 to 100 mm^2, which locates the maximum subsurface stress at a depth of 2 to 2.8 mm.

From the sides of the surface breaking U-crack two crack fronts propagated down into the rail as well. These crack fronts propagated in opposite longitudinal directions under an angle of 25° with the rail surface. In figure 4.4, two branches are indicated. These appear to have initiated at the same depth as where the plateau is found. Similar subsurface branching is observed in multiple other studies [88][90][94][100]. However, in these studies, subsurface branching was only observed on the leading crack (as defined by [77]). The fact that in the present study subsurface branches are observed on both crack flanks, is an indication that both crack flanks propagated as if they were a leading crack. This would also explain the high degree of symmetry that is observed in 3D model of the subsurface crack morphology in figure 4.23, which is not observed for squats that formed by unidirectional rolling [77][85][90][94].

Chapter 6: Conclusion

6.1 Conclusions

In this study, an investigation of the effect of bidirectional traffic on rail material has been presented. The goal is to answer to research questions: (1) how does bidirectional traffic change the effects of wheel-rail contact on the microstructure and mechanical properties of the rail, and (2) how does bidirectional traffic affect RCF crack initiation and propagation? To answer these questions, rail samples, taken from the railroad between Zutphen and Hengelo, were metallographically examined. The metallographic investigation was supported by the analysis of the wheel-rail contact, which is simulated with the VAMPIRE® multibody wheel-rail contact simulation software. Based on the results, a number of conclusions can be drawn.

In longitudinal direction, optical microscopy revealed the absence of plastic strain accumulation in the surface layer of the rail. This plastic strain, which is the result of the tangential surface traction, is always observed on rails subjected to unidirectional traffic and is considered the leading cause of fatigue crack initiation. The absence of this plastic strain can have two explanations: (1) the material (after initial hardening) is not repeatedly plastically deformed, or (2) the material is repeatedly plastically deformed, but does not accumulate plastic strain in a particular direction. The wheel-rail contact simulation showed that the contact stresses exceed the elastic shakedown limit, so plastic deformation is expected. By the analysis of the trajectory of the principal shear stress, it was shown that repeated plastic deformation will occur at a lower stress level for bidirectional traffic than for unidirectional traffic, due to the frequent reversal of tangential surface traction. From the observations on the microstructure and the analysis of the contact stresses, it is concluded that the frequent direction reversal of the tangential surface traction causes cyclic plasticity without the accumulation of strain, detectable with optical microscopy.

Contrary to the longitudinal direction, plastic strain is detected in lateral direction. Optical microscopy revealed plastic strain in a part of the rail surface layer. The material is deformed in the direction of the gauge side of the rail as a result of tangential surface traction. The wheel-rail contact simulation shows that there are two sources of lateral tangential surface traction: lateral creep and the lateral component of the normal load in the wheel-rail contact plane. The observed deformation is in good agreement with the simulated position and orientation of the peak lateral surface traction. In both traffic directions the lateral surface traction is oriented towards the gauge side of the rail. Therefore it is concluded that lateral deformation is independent of the traffic direction.

Etching of the running surface, optical microscopy and hardness measurements all showed the presence of WEL on the surface of the studied rail specimens. Based on the observed shape and orientation of the WEL on the rail surface, it was concluded that the WEL formed due to slip in the wheel-rail contact patch. However, the wheel-rail contact simulation showed that, for normal running behaviour, the creep ratio is not sufficient to generate the frictional heat the rail above the required austenitization temperature, which is required for WEL formation. No indication was found that indicates that bidirectional traffic has had an effect on the formation of WEL.

All cracks on the rail samples are positioned on grinding facet two but a clear initiation mechanism could not be determined. Although WEL was found in the proximity of the RCF defects, no observation was made that WEL actually resulted in crack initiation. The spatial distribution of defects also does not provide insight into the conditions of initiation. There is a real possibility, especially for the larger category C defects, that crack initiation of the defects found on grinding facet two occurred prior to the grinding maintenance in 2017. It is expected that the layer of material from which the crack had initiated was removed by wear or during grinding.

Optical microscopy did not observe any indication of crack initiation by either MnS inclusions or PE ferrite. Fatigue cracks have not been initiated by delamination of heavily strained MnS inclusions and pro-eutectoid ferrite grains as a result of bidirectional traffic. MnS inclusions and PE ferrite grains are not elongated in the direction of the strain. Therefore, the boundaries between MnS inclusion or PE ferrite and the pearlite matrix are not weakened enough to facilitate crack initiation.

Once initiated, the propagation of the RCF cracks is governed by the shear stresses generated in the contact patch. The observed propagation direction is consistent with the simulated direction of maximum shear stress. The presence of fluid in the crack, as indicated by the oxides on the crack faces, likely decreased the crack face friction and assisted crack propagation.

Although developed as a result of bidirectional traffic, the observed crack morphology showed similarities to the squat RCF defect. New insights were generated by considering the implications of bi-directional traffic and

the rolling direction reversal on existing theories about the rail squat formation mechanism. The theory by Li et al. [91][92] does not provide a complete explanation of squat formation. In the context of bidirectional traffic it would predict a defect shape that is not observed on the studied rail. The definitions of a leading and trailing crack in the theory of Steenbergen and Dollevoet [77][93] lose their meaning in the context of bidirectional traffic.

The subsurface crack morphology of a large category C U-crack is analysed. In contrast to squat defects formed by unidirectional traffic, the subsurface crack morphology of the category C U-crack showed to be symmetric. Furthermore, branching was observed in either longitudinal propagation directions, which for unidirectional traffic is only observed in the traffic direction.

6.2 Recommendations for further research

Further research is required to improve the understanding of the effects of bidirectional traffic and the process of the RCF defect formation in general.

Optical microscopy showed an absence of longitudinal plastic strain accumulation in the surface layer of the rail. Further research is necessary to examine the plastic response of the surface layer. Analysis of the plastic behaviour could provide insight into the relation between bidirectional traffic and crack initiation and could lead to the improvement of plasticity models.

The observed symmetry of the subsurface crack morphology and indicates that the propagation of the crack front in either longitudinal direction is governed by the same crack propagation mechanisms. Analysis of the beach marks on the crack faces could further clarify the responsible crack propagation mechanisms and could provide insight into the crack growth rate.

Additional research should be performed on the conditions of WEL formation on the rail surface. Given the simulated creep ratios, it is expected that there must be some other factors at play which enable the generation of the heat flux required for material transformation. Field measurements of the running behaviour of trains could provide new insights into this phenomenon.

Two kinds of surface breaking crack shapes were observe on the studied rail specimen. Further investigation is required to determine the cause of this observed difference.

Based on the results in this study, no conclusion could be drawn on the crack initiation mechanism. Additional research should be done to further investigate the processes involved in crack initiation. Since the process of RCF defect initiation and propagation develops over time, further research should consider all stages in the life of RCF formation in finer detail. By designing the research in that way, increases the possibility to make a conclusive description of the initiation mechanism.

Bibliography

[1] A Zoeteman et al. “28: Managing the wheel–rail interface: the Dutch experience”. In: *Wheel–rail interface handbook*. Woodhead Publishing Limited, 2009, pp. 792–818. DOI: 10.1533/9781845696788.2.792. URL: http://dx.doi.org/10.1533/9781845696788.2.792.

[2] OPEUS. “Urban rail system energy requirements in Europe”. In: (2017).

[3] D I Fletcher, F J Franklin, and A Kapoor. “9: Rail surface fatigue and wear”. In: *Wheel–rail interface handbook*. Woodhead Publishing Limited, 2009, pp. 280–310. DOI: 10.1533/9781845696788.1.280. URL: http://dx.doi.org/10.1533/9781845696788.1.280.

[4] Alberto Javier Perez-Unzueta. “WEAR RESISTANCE OF PEARLITIC RAIL STEELS”. PhD thesis. University of Leicester, 1992.

[5] J E Garnham and C L Davis. “5: Rail materials”. In: *Wheel–rail interface handbook*. Woodhead Publishing Limited, 2009, pp. 125–171. DOI: 10.1533/9781845696788.1.125. URL: http://dx.doi.org/10.1533/9781845696788.1.125.

[6] Jun Wu. “Microstructure evolution in pearlitic rail steel due to rail / wheel interaction”. PhD thesis. TU Delft, 2018. DOI: 10.4233/uuid.

[7] Jari Boesten. “Study of degradation and damage in a pearlitic steel rail during cyclic grinding maintenance and subsequent wheel-rail contacts .” PhD thesis. TU Delft, 2019.

[8] Sir Harshad Bhadeshia and Sir Robert Honeycombe. “1. Iron and Its Interstitial Solutions”. In: 2017. Chap. 1, pp. 1–22. ISBN: 9780081002704.

[9] J. O. Andersson. *Thermo-Calc*. 2002.

[10] A. Kumar et al. “Microstructural evolution of white and brown etching layers in pearlitic rail steels”. In: *Acta Materialia* 171 (2019), pp. 48–64. ISSN: 13596454. DOI: 10.1016/j.actamat.2019.04.012. URL: https://doi.org/10.1016/j.actamat.2019.04.012.

[11] M. V. Kral. *Proeutectoid ferrite and cementite transformations in steels*. Vol. 1. Woodhead Publishing Limited, 2012, pp. 225–275. ISBN: 9781845699703. DOI: 10.1533/9780857096104.2.225. URL: http://dx.doi.org/10.1533/9780857096104.2.225.

[12] S. E. Offerman et al. “In-situ study of pearlite nucleation and growth during isothermal austenite decomposition in nearly eutectoid steel”. In: *Acta Materialia* 51.13 (2003), pp. 3927–3938. ISSN: 13596454. DOI: 10.1016/S1359-6454(03)00217-9.

[13] George Krauss. “Pearlite, Ferrite, and Cementite”. In: *Steels: Processing, Structure and Performance*. 2015. Chap. 4.

[14] R.E. Smallman and A.H.W. Ngan. “Steel Transformations”. In: *Modern Physical Metallurgy* (2014), pp. 473–498. DOI: 10.1016/b978-0-08-098204-5.00012-2.

[15] Günter Gottstein. *Physical Foundations of Materials Science*. 2004. ISBN: 9783642072710. DOI: 10.1007/978-3-662-09291-0.

[16] Sir Harshad Bhadeshia and Sir Robert Honeycombe. “5. Formation of Martensite”. In: 2017. Chap. 5, pp. 135–177. ISBN: 9780081002704.

[17] V. I. Sarrak and G. A. Filippov. “Brittleness of martensite”. In: 2.3 (1978), pp. 377–389.

[18] Sir Harshad Bhadeshia and Sir Robert Honeycombe. “6. Bainite”. In: 2017. Chap. 6, pp. 179–202. ISBN: 9780081002704.

[19] George Krauss. “High-Carbon Steels - Fully Pearlitic Microstructures and Wire and Rail Applications”. In: *Steels: Processing, Structure and Performance*. 2015. Chap. 15.

[20] George Krauss. “Non-Martensitic strengthening of medium-carbon steels”. In: *Steels: Processing, Structure and Performance*. 2015. Chap. 14.

[21] J. M. Hyzak and I. M. Bernstein. “The role of microstructure on the strength and toughness of fully pearlitic steels”. In: *Metallurgical Transactions A* 7.8 (1976), pp. 1217–1224. ISSN: 15431940. DOI: 10.1007/BF02656606.

[22] K. K. Ray and D. Mondal. “The effect of interlamellar spacing on strength of pearlite in annealed eutectoid and hypoeutectoid plain carbon steels”. In: *Acta Metallurgica Et Materialia* 39.10 (1991), pp. 2201–2208. ISSN: 09567151. DOI: 10.1016/0956-7151(91)90002-I.

[23] C. M. Bae, C. S. Lee, and W. J. Nam. “Effect of carbon content on mechanical properties of fully pearlitic steels”. In: *Materials Science and Technology* 18.11 (2002), pp. 1317–1321. ISSN: 02670836. DOI: 10.1179/026708302225007556.

[24] John J. Lewandowski and Anthony W. Thompson. “Effects of the Prior Austenite Grain Size on the Ductility of Fully Pearlitic Eutectoid Steel.” In: *Metallurgical transactions. A, Physical metallurgy and materials science* 17 A.3 (1986), pp. 461–472. ISSN: 03602133. DOI: 10.1007/BF02643953.

[25] Y. J. Park and I. M. Bernstein. “The process of crack initiation and effective grain size for cleavage fracture in pearlitic eutectoid steel”. In: *Metallurgical Transactions A* 10.11 (1979), pp. 1653–1664. ISSN: 03602133. DOI: 10.1007/BF02811698.

[26] ProRail. “SPC00011”. In: (2019).

[27] (CEN) European Committee for Standardisation. “EN 10027-1: 2005 Designation systems for steels - Part 1: Steel names”. In: *European Standards* 3 (2005), pp. 1–25.

[28] NEN. “NEN-EN 13674-1+A1”. In: 1 (2017).

[29] ProRail. “OVS00056-5.1”. In: (2019).

[30] Corus Strip Products UK. “EN 10025-2:2004, European standard for hot-rolled structural steel”. In: (2006).

[31] Arcelor Mittal. “Material Grade Comparison Chart”. In: 65.Psl 1 (), p. 460. URL: http://ds.arcelormittal.com/repo/Projects%20Oil%20and%20Gas/Material%20Grade%20Comparison%20Chart.pdf.

[32] Fatih Bozkurt and Eva Schmidová. “Fracture toughness evaluation of S355 steel using circumferentially notched round bars”. In: *Periodica Polytechnica Transportation Engineering* 47.2 (2019), pp. 91–95. ISSN: 15873811. DOI: 10.3311/PPtr.11560.

[33] Sir Harshad Bhadeshia and Sir Robert Honeycombe. “4. Solutes that Substitute for Iron”. In: 2017. Chap. 4, pp. 101–134. ISBN: 9780081002704.

[34] F Fau, P. Sécordel, and J. Jaiswal. “Tata Steel Premium Grades to address Wear and Rolling Contact Fatigue”. In: ().

[35] Sir Harshad Bhadeshia and Sir Robert Honeycombe. “1. Iron and Its Interstitial Solutions”. In: 2017. Chap. 1, pp. 1–22. ISBN: 9780081002704.

[36] Bart Schotsman. “Spalling defects in MHH steel rails”. In: October (2017).

[37] W. H. Hodgson. “Residual Stresses in Rail”. In: *Rail Quality and Maintenance for Modern Railway Operation*. 1992, pp. 61–76. ISBN: 9788578110796. DOI: 10.1017/CBO9781107415324.004.

[38] U. Zerbst et al. “Introduction to the damage tolerance behaviour of railway rails - a review”. In: *Engineering Fracture Mechanics* 76.17 (2009), pp. 2563–2601. ISSN: 00137944. DOI: 10.1016/j.engfracmech.2009.09.003. URL: http://dx.doi.org/10.1016/j.engfracmech.2009.09.003.

[39] S Iwnicki, S Björklund, and R Enblom. “3: Wheel–rail contact mechanics”. In: *Wheel–rail interface handbook*. Woodhead Publishing Limited, pp. 58–92. DOI: 10.1533/9781845696788.1.58. URL: http://dx.doi.org/10.1533/9781845696788.1.58.

[40] Coenraad Esveld. “Wheel-Rail Interface”. In: *Modern Railway*. 2001. Chap. 2.

[41] R Lewis and U Olofsson. “2: Basic tribology of the wheel–rail contact”. In: *Wheel–rail interface handbook*. Woodhead Publishing Limited, pp. 34–57. DOI: 10.1533/9781845696788.1.34. URL: http://dx.doi.org/10.1533/9781845696788.1.34.

[42] Anton van Beek. “Load Rating and Rolling Contact Fatigue”. In: *Advanced Engineering Design: Lifetime performance and reliability*. 2015. Chap. 4.

[43] K. L. Johnson. “9: Rolling contact of inelastic bodies”. In: *Contact Mechanics*. 1985. Chap. 9, pp. 284–311. DOI: 10.1017/cbo9781139171731.010.

[44] K. L. Johnson. “4: Normal contact of elastic solids: Hertz theory”. In: *Contact Mechanics*. 1985. Chap. 4.

[45] K. L. Johnson. “8: Rolling contact of elastic bodies”. In: *Contact Mechanics*. 1985. Chap. 8, pp. 242–283. DOI: 10.1017/cbo9781139171731.009.

[46] K. L. Johnson. “7: Tangential loading and sliding contact”. In: *Contact Mechanics*. 1985. Chap. 7. ISBN: 0521347963. DOI: 10.1017/cbo9781139171731.008.

[47] A. F. Bower and K. L. Johnson. “Plastic flow and shakedown of the rail surface in repeated wheel-rail contact”. In: *Wear* 144.1-2 (1991), pp. 1–18. ISSN: 00431648. DOI: 10.1016/0043-1648(91)90003-D.

[48] D. J. Haines and E. Ollerton. “Contact Stress Distributions on Elliptical Contact Surfaces Subjected to Radial and Tangential Forces”. In: *Proceedings of the Institution of Mechanical Engineers* 177.1 (1963), pp. 95–114. ISSN: 0020-3483. DOI: 10.1243/pime{_}proc{_}1963{_}177{_}014{_}02.

[49] Tomasz Wierzbicki. *Isotropic and Kinematich Hardening*. 2021. URL: https://eng.libretexts.org/Bookshelves/Mechanical_Engineering/Structural_Mechanics_(Wierzbicki)/11%3A_Fundamental_Concepts_in_Structural_Plasticity/11.05%3A_Isotropic_and_Kinematic_Hardening.

[50] K. L. Johnson. "The strength of surfaces in rolling contact". In: *Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science* 203.3 (1989), pp. 151–163. ISSN: 20412983. DOI: 10.1243/PIME{_}PROC{_}1989{_}203{_}100{_}02.

[51] K. L. Johnson. "The strength of surfaces in rolling contact". In: *Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science* 203.3 (1989), pp. 151–163. ISSN: 20412983. DOI: 10.1243/PIME{_}PROC{_}1989{_}203{_}100{_}02.

[52] A. F. Bower. "Cyclic hardening properties of hard-drawn copper and rail steel". In: *Journal of the Mechanics and Physics of Solids* 37.4 (1989), pp. 455–470. ISSN: 00225096. DOI: 10.1016/0022-5096(89)90024-0.

[53] A. F. Bower and K. L. Johnson. "Shakedown, Residual Stress and Plastic Flow in Repeated Wheel-Rail Contact". In: *Rail Quality and Maintenance for Modern Railway Operation* (1993), pp. 239–249. DOI: 10.1007/978-94-015-8151-6{_}19.

[54] OECD. *Glossary of terms and definitions in the field of friction, wear and lubrication - tribology*. 1969.

[55] S. C. Lim and M. F. Ashby. "Overview no. 55 Wear-Mechanism maps". In: *Acta Metallurgica* 35.1 (1987), pp. 1–24. ISSN: 00016160. DOI: 10.1016/0001-6160(87)90209-4.

[56] J. F. Archard and W. Hirst. "The wear of metals under unlubricated conditions". In: (1956).

[57] R. Lewis and U. Olofsson. "Mapping rail wear regimes and transitions". In: *Wear* 257.7-8 (2004), pp. 721–729. ISSN: 00431648. DOI: 10.1016/j.wear.2004.03.019.

[58] Nam P. Suh. "The delamination theory of wear". In: *Wear* 25.1 (1973), pp. 111–124. ISSN: 00431648. DOI: 10.1016/0043-1648(73)90125-7.

[59] W R Tyfour, J H Beynon, and A Kapoor. "The steady state wear behaviour of pearlitic rail steel under dry rolling-sliding contact conditions". In: 180 (1995), pp. 79–89.

[60] Martin Ertz and Klaus Knothe. "A comparison of analytical and numerical methods for the calculation of temperatures in wheel/rail contact". In: *Wear* 253.3-4 (2002), pp. 498–508. ISSN: 00431648. DOI: 10.1016/S0043-1648(02)00120-5.

[61] M. Naeimi et al. "Computation of the flash-temperature at the wheel-rail contact using a 3D finite element model and its comparison with analytical methods". In: *CM 2015 - 10th International Conference on Contact Mechanics of Wheel / Rail Systems* (2015).

[62] Meysam Naeimi et al. "Thermomechanical analysis of the wheel-rail contact using a coupled modelling procedure". In: *Tribology International* 117.February 2018 (2018), pp. 250–260. ISSN: 0301679X. DOI: 10.1016/j.triboint.2017.09.010. URL: https://doi.org/10.1016/j.triboint.2017.09.010.

[63] F. D. Fischer, W. Daves, and E. A. Werner. "On the temperature in the wheel-rail rolling contact". In: *Fatigue and Fracture of Engineering Materials and Structures* 26.10 (2003), pp. 999–1006. ISSN: 8756758X. DOI: 10.1046/j.1460-2695.2003.00700.x.

[64] Klaus Knothe and S. Liebelt. "Determination of temperatures for sliding contact with applications for wheel-rail systems". In: *Wear* 189 (1995), pp. 91–99. ISSN: 00431648. DOI: 10.1016/S0043-1648(96)07318-8.

[65] Jay Prakash Srivastava et al. "A numerical study on effects of friction-induced thermal load for rail under varied wheel slip conditions". In: *Simulation* 95.4 (2019), pp. 351–362. ISSN: 17413133. DOI: 10.1177/0037549718782629.

[66] A. Kumar et al. "In situ study on fracture behaviour of white etching layers formed on rails". In: *Acta Materialia* 180 (2019), pp. 60–72. ISSN: 13596454. DOI: 10.1016/j.actamat.2019.08.060. URL: https://doi.org/10.1016/j.actamat.2019.08.060.

[67] Khoa Vo et al. "A tool to estimate the wheel/rail contact and temperature rising under dry, wet, oil conditions". In: (2014).

[68] Rui Pan et al. "The microstructure analysis of white etching layer on treads of rails". In: *Engineering Failure Analysis* 82.May (2017), pp. 39–46. ISSN: 13506307. DOI: 10.1016/j.engfailanal.2017.06.018. URL: http://dx.doi.org/10.1016/j.engfailanal.2017.06.018.

[69] Shaoguang Li et al. ""Brown etching layer": A possible new insight into the crack initiation of rolling contact fatigue in rail steels?" In: *Engineering Failure Analysis* 66 (2016), pp. 8–18. ISSN: 13506307. DOI: 10.1016/j.engfailanal.2016.03.019. URL: http://dx.doi.org/10.1016/j.engfailanal.2016.03.019.

[70] Maha Messaadi and Michaël Steenbergen. "Stratified surface layers on rails". In: *Wear* 414-415.July (2018), pp. 151–162. ISSN: 00431648. DOI: 10.1016/j.wear.2018.07.019.

[71] H. C. Eden, J. E. Garnham, and C. L. Davis. "Influential microstructural changes on rolling contact fatigue crack initiation in pearlitic rail steels". In: *Materials Science and Technology* 21.6 (2005), pp. 623–629. ISSN: 02670836. DOI: 10.1179/174328405X43207.

[72] John E Garnham and Claire L Davis. "The role of deformed rail microstructure on rolling contact fatigue initiation". In: 265 (2008), pp. 1363–1372. DOI: 10.1016/j.wear.2008.02.042.

[73] A F Bower and K L Johnson. "The influence of strain hardening on cumulative plastic deformation in rolling and sliding line contact". In: 37.4 (1989).

[74] J. E. Garnham, R. G. Ding, and C. L. Davis. "Ductile inclusions in rail, subject to compressive rolling-sliding contact". In: *Wear* 269.11-12 (2010), pp. 733–746. ISSN: 00431648. DOI: 10.1016/j.wear.2010.07.010. URL: http://dx.doi.org/10.1016/j.wear.2010.07.010.

[75] C. D. Liu, M. N. Bassim, and S. St Lawrence. "Evaluation of fatigue-crack initiation at inclusions in fully pearlitic steels". In: *Materials Science and Engineering A* 167.1-2 (1993), pp. 107–113. ISSN: 09215093. DOI: 10.1016/0921-5093(93)90343-D.

[76] J E Garnham and C L Davis. "Very early stage rolling contact fatigue crack growth in pearlitic rail steels". In: 271 (2011), pp. 100–112. DOI: 10.1016/j.wear.2010.10.004.

[77] Michaël Steenbergen and Rolf Dollevoet. "On the mechanism of squat formation on train rails – Part I : Origination". In: *International Journal of Fatigue* 47 (2013), pp. 361–372. DOI: 10.1016/j.ijfatigue.2012.04.023.

[78] S Simon et al. "Tribological characterization of rail squat defects". In: *Wear* 297.1-2 (2013), pp. 926–942. ISSN: 0043-1648. DOI: 10.1016/j.wear.2012.11.011. URL: http://dx.doi.org/10.1016/j.wear.2012.11.011.

[79] A. Al-Juboori et al. "Squat formation and the occurrence of two distinct classes of white etching layer on the surface of rail steel". In: *International Journal of Fatigue* 104 (2017), pp. 52–60. ISSN: 01421123. DOI: 10.1016/j.ijfatigue.2017.07.005.

[80] Sarvesh Pal et al. "Metallurgical and physical understanding of rail squat initiation and propagation". In: *Wear* 284-285 (2012), pp. 30–42. ISSN: 00431648. DOI: 10.1016/j.wear.2012.02.013. URL: http://dx.doi.org/10.1016/j.wear.2012.02.013.

[81] Anit Kumar. "Multiscale study of microstructural evolution and damage in rail steels". PhD thesis. TU Delft, 2019. DOI: 10.4233/uuid.

[82] Ashish Kumar Saxena et al. "Micro fracture investigations of white etching layers". In: *Materials and Design* 180 (2019), p. 107892. ISSN: 18734197. DOI: 10.1016/j.matdes.2019.107892. URL: https://doi.org/10.1016/j.matdes.2019.107892.

[83] R. I. Carroll and J. H. Beynon. "Rolling contact fatigue of white etching layer: Part 1. Crack morphology". In: *Wear* 262.9-10 (2007), pp. 1253–1266. ISSN: 00431648. DOI: 10.1016/j.wear.2007.01.003.

[84] David I. Fletcher and John H. Beynon. "Development of a Machine for Closely Controlled Rolling Contact Fatigue and Wear Testing". In: *Journal of Testing and Evaluation* 28.4 (2000), pp. 267–275. ISSN: 00903973. DOI: 10.1520/jte12104j.

[85] Xiangyun Deng et al. "Investigation of the formation of corrugation-induced rail squats based on extensive field monitoring". In: *International Journal of Fatigue* 112.February (2018), pp. 94–105. ISSN: 01421123. DOI: 10.1016/j.ijfatigue.2018.03.002. URL: https://doi.org/10.1016/j.ijfatigue.2018.03.002.

[86] Stuart L Grassie. "Studs and squats: The evolving story". In: 367 (2016), pp. 194–199. DOI: 10.1016/j.wear.2016.03.021.

[87] A. F. Bower. "The influence of crack face friction and trapped fluid on surface initiated rolling contact fatigue cracks". In: *Journal of Tribology* (1988).

[88] P E Bold, M W Brown, and R J Allen. "Shear mode crack growth and rolling contact fatigue". In: *Wear* 144 (1991).

[89] Dylewski Benoît, Bouvier Salima, and Risbet Marion. "Multiscale characterization of head check initiation on rails under rolling contact fatigue : Mechanical and microstructure analysis". In: *Wear* 366-367 (2016), pp. 383–391. ISSN: 0043-1648. DOI: 10.1016/j.wear.2016.06.019. URL: http://dx.doi.org/10.1016/j.wear.2016.06.019.

[90] S. L. Grassie. “Squats and squat-type defects in rails: The understanding to date”. In: *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit* 226.3 (2012), pp. 235–242. ISSN: 09544097. DOI: 10.1177/0954409711422189.

[91] Zili Li et al. “Squat growth-Some observations and the validation of numerical predictions”. In: *Wear* 271.1-2 (2011), pp. 148–157. ISSN: 00431648. DOI: 10.1016/j.wear.2010.10.051. URL: http://dx.doi.org/10.1016/j.wear.2010.10.051.

[92] Zili Li et al. “An investigation into the causes of squats-Correlation analysis and numerical modeling”. In: *Wear* 265.9-10 (2008), pp. 1349–1355. ISSN: 00431648. DOI: 10.1016/j.wear.2008.02.037.

[93] Michaël Steenbergen and Rolf Dollevoet. “On the mechanism of squat formation on train rails – Part II : Growth”. In: 47 (2013), pp. 373–381. DOI: 10.1016/j.ijfatigue.2012.04.019.

[94] S. L. Grassie et al. “Studs: A squat-type defect in rails”. In: *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit* 226.3 (2012), pp. 243–256. ISSN: 09544097. DOI: 10.1177/0954409711421462.

[95] Serope Kalpakjian and Stephen Schmid. “Manufacturing, Engineering and Technology SI 6th Edition”. In: (2013), p. 1216. DOI: 0133128741.

[96] Michaël Steenbergen. “Rolling contact fatigue in relation to rail grinding”. In: *Wear* 356-357 (2016), pp. 110–121. ISSN: 0043-1648. DOI: 10.1016/j.wear.2016.03.015. URL: http://dx.doi.org/10.1016/j.wear.2016.03.015.

[97] Carsten Jørn Rasmussen et al. “Surface crack formation on rails at grinding induced martensite white etching layers”. In: *Wear* 384-385.February (2017), pp. 8–14. ISSN: 00431648. DOI: 10.1016/j.wear.2017.04.014. URL: http://dx.doi.org/10.1016/j.wear.2017.04.014.

[98] V Dikshit. “Investigation of rolling contact fatigue in a head-hardened rail”. In: *Wear* 144 (1991), pp. 89–102.

[99] H. Zhu et al. “Understanding and treatment of squat defects in a railway network”. In: *Wear* 442-443.November 2019 (2020). ISSN: 00431648. DOI: 10.1016/j.wear.2019.203139.

[100] Kunihiro Kondo, Katsunori Yoroizaka, and Yoshihiko Sato. “Cause, increase, diagnosis, countermeasures and elimination of Shinkansen shelling”. In: *Wear* 191.1-2 (1996), pp. 199–203. ISSN: 00431648. DOI: 10.1016/0043-1648(95)06727-2.

[101] D. F. Cannon et al. “Rail defects: An overview”. In: *Fatigue and Fracture of Engineering Materials and Structures* 26.10 (2003), pp. 865–886. ISSN: 8756758X. DOI: 10.1046/j.1460-2695.2003.00693.x.

[102] S. L. Grassie. “Rail corrugation: Characteristics, causes, and treatments”. In: *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit* 223.6 (2009), pp. 581–596. ISSN: 09544097. DOI: 10.1243/09544097JRRT264.

[103] Michaël Steenbergen and Edward De Jong. “Railway track degradation: The contribution of rolling stock”. In: *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit* 230.4 (2015), pp. 1164–1171. ISSN: 20413017. DOI: 10.1177/0954409715585371.

[104] W. R. Tyfour and J. H. Beynon. “The effect of rolling direction reversal on fatigue crack morphology and propagation”. In: *Tribology International* 27.4 (1994), pp. 273–282. ISSN: 0301679X. DOI: 10.1016/0301-679X(94)90007-8.

[105] W. R. Tyfour and J. H. Beynon. “The effect of rolling direction reversal on the wear rate and wear mechanism of pearlitic rail steel”. In: *Tribology International* 27.6 (1994), pp. 401–412. ISSN: 0301679X. DOI: 10.1016/0301-679X(94)90017-5.

[106] Dongfang Zeng et al. “Effect of bi-directional operation on rolling-sliding wear of wheel steel under variable amplitude loading”. In: *Proceedings of the Institution of Mechanical Engineers, Part J: Journal of Engineering Tribology* 230.2 (2016), pp. 186–195. ISSN: 2041305X. DOI: 10.1177/1350650115594955.

[107] GreenwoodEngineering. *MiniProf*. 2020.

[108] Meysam Naeimi et al. “Reconstruction of the rolling contact fatigue cracks in rails using X-ray computed tomography”. In: *NDT and E International* 92.April (2017), pp. 199–212. ISSN: 09638695. DOI: 10.1016/j.ndteint.2017.09.004. URL: https://doi.org/10.1016/j.ndteint.2017.09.004.

Chapter A: Appendix

A.1 Running behaviour all runs

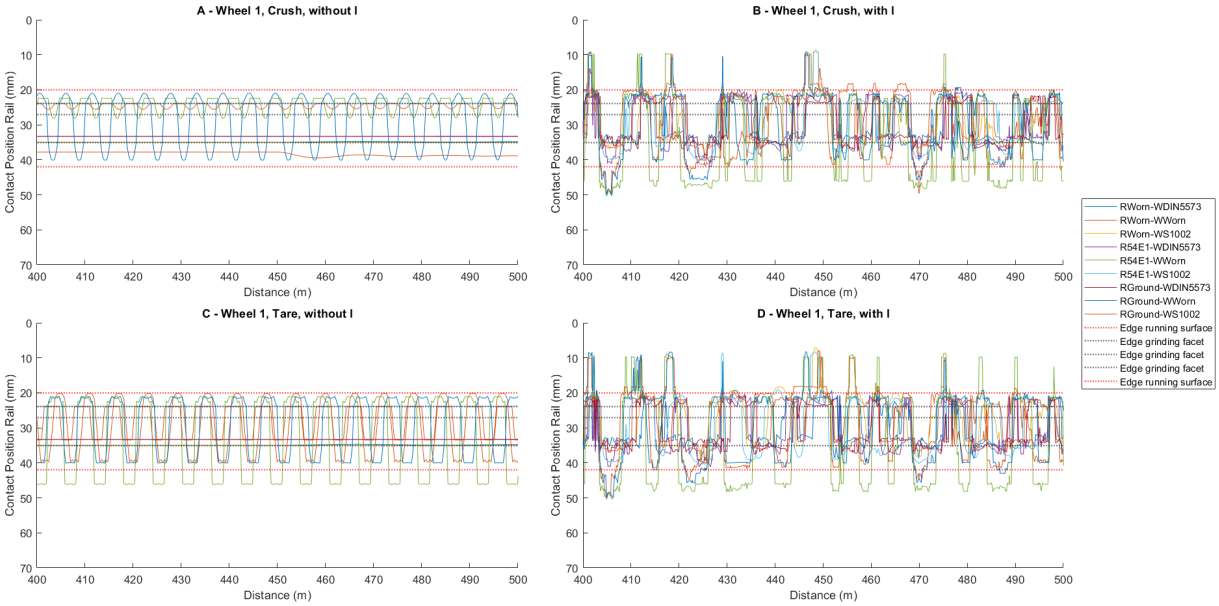


Figure A.1: *Contact position on the rail as a function of travel distance. A contact position of zero corresponds to contact at the gauge corner, 70 corresponds to contact at the field corner.*

Figure A.1 shows the position of the center of the contact patch on the rail plotted against the travel distance for all runs. The lines indicate the trajectories centre of the contact patch along the surface of the rail. Depending on the different combinations of train weight and track irregularities the runs are divided into 4 subfigures. Each subfigure shows the 9 simulation runs, one for each different wheel-rail profile combination. The position of the running surface as measured in figure 4.1 is indicated by the dotted red lines. In a similar fashion the boundaries between grinding facets are indicated by the black dotted lines. For all runs the majority of the contact occurs within the boundaries of the running surface. This indicates that the position of the simulated contact position matches reality well. Figures A.1a and c show the contact position trajectories for the runs where no track irregularities are applied. For these runs the trajectory along the rail is either a straight line or has a distinct cyclic pattern that repeats with a fixed interval. The RWorn-WWorn trajectory in figure A.1a is straight line from 400 to 450 m, but shows some small oscillations starting 450 m. This is caused by the application of the engine torque on the wheels. For all other runs without track irregularities such an effect is not present. This highly regular running behaviour is obviously is not representative the running behaviour of a real train, but it does serve as a benchmark for the runs with track irregularities. The running behaviour of runs with the track irregularities are plotted in figure 2.5b and d. When compared the runs without track irregularities the running behaviour is less regularly periodic. For some runs the contact position moves outside of the boundaries of the running surface. The train weight appears to influence the running behaviour of the train less than the track irregularities, as the contact position trajectories do not change significantly depending on train weight.

A.2 Track irregularities

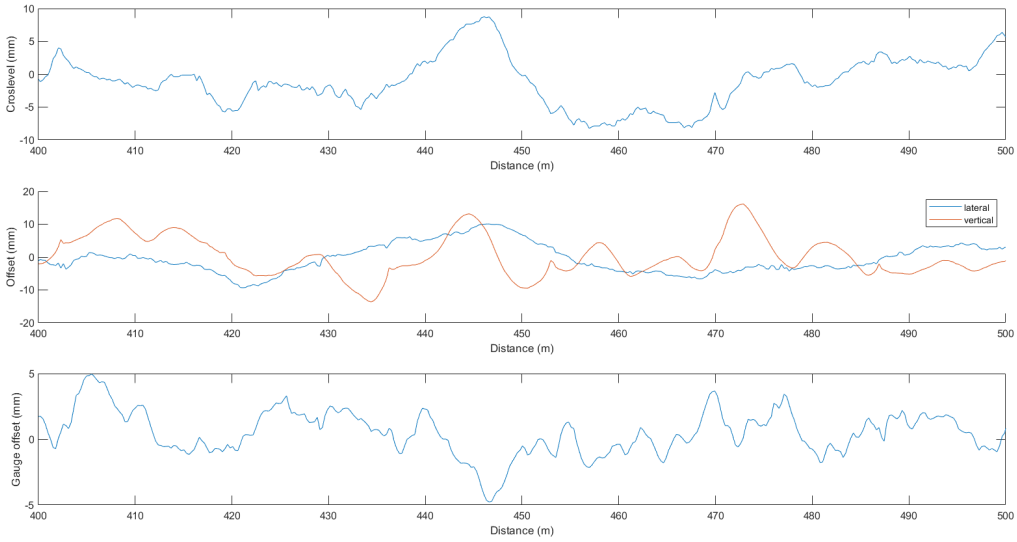


Figure A.2: *Track irregularities T110 from VAMPIRE® library.*

A.3 Calculation of creep VAMPIRE

Calculation of Creepages

At each time step the lateral displacement of each wheel relative to the rail is calculated by the analysis program taking into account rail movements. This value is then used to look up the contact data for each contact patch between wheel and rail. Having obtained the position, rolling radius and contact angle for each contact patch, the creepages are calculated for each wheel rail contact patch as follows.

For longitudinal creepage on the left and right contacts

$$\gamma_{1L} = 1 + \frac{l_{0L}\dot{W}}{V} + \frac{r_L\dot{P}}{V} + \frac{l_{0L}}{R} + \frac{\dot{X}}{V}$$

$$\gamma_{1R} = 1 - \frac{l_{0R}\dot{W}}{V} + \frac{r_R\dot{P}}{V} - \frac{l_{0R}}{R} + \frac{\dot{X}}{V}$$

For lateral creepage on the left and right contacts

$$\gamma_{2L} = \left(\frac{\dot{Y}}{V} + \frac{r_L\dot{P}W}{V}\right) sec(\delta_L)$$

$$\gamma_{2R} = \left(\frac{\dot{Y}}{V} + \frac{r_R\dot{P}W}{V}\right) sec(\delta_R)$$

For spin creepage on the left and right contacts

$$\gamma_{3L} = -\left(\frac{\dot{P}}{V}\right) sin(\delta_L) + \frac{\dot{W}}{V} cos(\delta_L)$$

$$\gamma_{3R} = \left(\frac{\dot{P}}{V}\right) sin(\delta_R) + \frac{\dot{W}}{V} cos(\delta_R)$$

where V is the vehicle speed

X,Y,W,P are displacements of the wheelset in track axes, dot denotes rate of change with time

r_L, r_R are rolling radii on the left and right wheels

R is the curve radius of the track (CURVING program only)

l_{0L}, l_{0r} are the distances from the centre of the wheelset to the left and right contact patch

δ_L, δ_R are the contact angles at the left and right contacts.

Figure A.3: *Calculation Creep*