Preface

The present thesis has been prepared at the Department of Mathematical Modelling (IMM), Technical University of Denmark and at the Development Office – Danish State Railway (DSB) as a partial accomplishment of the requirements for the Ph.D.-degree. The work has been carried out in the period from September 1995 to September 1998 and was financially supported by DSB and STVF – The Danish Research Council.

During all three years the work has been surveyed by a group consisting of Jørgen Tornhøj Christensen, ScanRail Consult, Per Grove Thomsen, IMM and Peter Leth Christiansen, IMM who continuously have contributed with many relevant comments regarding the outline of the ongoing work.

The work will be evaluated by a panel of external examiners: John Hansen, FAM, Technical University of Denmark, Jean Pierre Pascal, INRETS and Klaus Knothe, TU-Berlin. I appreciate that they have agreed to devote some of their sparse time to my work.

As a part of my work I have stayed two times three months at Institut für Luft- und Raumfahrt (ILR), TU-Berlin under the supervision of Professor Klaus Knothe. I am most thankful to Professor Knothe who has served as a great inspiration and introduced me to the world of contact mechanics with much enthusiasm and a never failing trust in my somewhat different approach to the subject. A further thanks to all the employees at the ILR for making my stays interesting and pleasant. Especially I would like to thank André Theiler with whom I have had many interesting and alternative discussions of topics more or less related to contact mechanics.

The last three years I have learned much about daily life railway problems thanks to my colleagues at DSB. This insight is perhaps not conspicuous in the present work, but I believe that this more practical approach to wheel/rail contact has provided me with a more complete conception of contact mechanics. Even though my work may seem far away from the problems my colleagues deal with every day, they have always shown a genuine interest in my work, which I appreciate very much.

The first two years at IMM I shared office with Carsten Nordstrøm Jensen – a pleasant experience with the exchange of many interesting views upon the daily life of a Ph.D.-student.

My work has been supervised by Jørgen Jessen, DSB and Hans True, DTU. Jørgen Jessen has provided me with some of his overwhelming knowledge of practical wheel/rail problems, which has given me the opportunity to relate my theory to the real world. I owe Jørgen much thank for this very inspiring insight.

Hans True is perhaps the main responsible for me ending up in contact mechanics. During all three years he has always inspired and encouraged me. Using his many contacts in Denmark and abroad he has thrown me into the railway world with an apparently blind faith in my abilities which has given me many interesting experiences and added an extra dimension to my Ph.D.-work. For that I am most thankful.

Finally I thank Jane Hvolbæk Larsen for the support and encouragement during my work.

Lyngby, August 1998

Jakob Birkedal Nielsen

Summary

Today many simulation routines concerning railway dynamics employ rather primitive contact models which are not necessarily suited for the specific wheel/rail contact problem. The objective of the present thesis is to derive a more flexible contact model which can be applied on a variety of contact problems.

When it comes to the modelling of the wheel/rail contact it is always a compromise between computational speed and accuracy. Many numerical methods provide a very good accuracy, but since most railway simulations necessitates the evaluation of many consecutive contact situations the relative slow computational speed is extremely critical. To avoid this problem the present model is based on an analytical approach.

The model derived in the thesis is a two-dimensional contact model based on elastic half spaces. It is demonstrated that the solution to a three-dimensional contact problem with no spin has many similarities with the two-dimensional solution. Thus, the results obtained with the present model can qualitatively be extended to the three-dimensional contact problem.

The thesis is divided into two parts: one containing the derivation of the contact model and one containing examples of application. The model is applied on four different types of contact problems which cannot be treated with the most common contact models:

- contact between corrugated surfaces
- contact with velocity dependent friction coefficient
- contact between rough surfaces
- non-steady contact

The calculations demonstrate with much clearness that the solution to the contact problem is very sensitive to the choice of contact model. This illustrates how crucial it is to employ an adequate contact model in a given simulation routine in order to obtain a realistic result. If the assumptions of the contact model do not fulfill the actual contact situation the result can be most erroneous and thus misleading.

Resumé

Mange simulations-programmer, der behandler jernbane-dynamiske problemer, anvender relativt primitive kontakt-modeller, der ikke nødvendigvis er velegnede til det pågældende hjul/skinne kontakt-problem. Formålet med denne afhandling er at udlede en mere fleksible kontakt-model, som kan anvendes på en lang række kontakt-problemer.

Når det gælder modelleringen af hjul/skinne kontakt sker der altid en afvejning mellem regne-hastighed og præcision. Mange numeriske metoder regner med stor præcision, men da de fleste jernbane-dynamiske simulationer kræver at mange på hinanden følgende kontakt-problemer bliver løst, er den relativt lave regne-hastighed meget kritisk. For at undgå dette problem bygger nærværende model på en analytisk metode.

Den model, der udledes i afhandlingen, er en to-dimensional kontakt-model baseret på teorien for elastiske halv-rum. Det påvises at løsningen til et tre-dimensionalt kontakt-problem uden spin i vid udstrækning er lig den to-dimensionale løsning. Således kan de resultater, der er opnået med

nærværende model, blive udvidet så de kvalitativt også gælder for tredimensional kontakt.

Afhandlingen er opdelt i to hovedafsnit: en del der omhandler udledningen af kontakt-modellen, og en del hvor eksempler på anvendelse af modellen bliver gennemgået. Modellen er anvendt på fire forskellige typer af kontaktproblemer, der ikke kan behandles med de sædvanlige kontakt-modeller:

- kontakt mellem riflede overflader
- kontakt med hastighedsafhængig friktions-koefficient
- kontakt mellem ru overflader
- ikke-stationær kontakt

Beregningerne viser med al ønskelig tydelighed at løsningen til et givet kontakt-problem er meget følsomt med hensyn til valget af løsnings-model. Dette illustrerer, at det er meget vigtigt at anvende en passende kontaktmodel i et simulations-program for at opnå realistiske resultater. Hvis forudsætningerne for kontakt-modellen ikke opfylder det faktiske kontaktproblem, kan resultatet blive yderst fejlagtigt og dermed vildledende.

Nomenclature

Lower Case Letters

a	half the contact length	[m]
a^*	half the length of the stick zone	[m]
b	half the contact width	[m]
d	distance	[m]
g_{ij}	influence functions	[m/N]
h	distance between undeformed bodies	[m]
k	wave number	[1/m]
p	normal pressure	[N/m^2]
q	tangential stress	[N/m^2]
r_{a_0}	ratio a_0^*/a_0	[$\cdot$]
r_τ	ratio τ^*/τ	[$\cdot$]
s	slip	[$\cdot$]
t	time	[s]
u	displacement	[m]

x	local longitudinal position	[m]
y	local lateral position	[m]

Upper Case Letters

A_n	matrix coefficients	[m^{2n}]
A_n^{-1}	matrix coefficients	[m^{2n}]
A_n^*	matrix coefficients	[m^{2n}]
$\mathcal{A}$	asperity	
B_n	polynomial coefficients	[m^{1-n}]
B_n^*	polynomial coefficients	[m^{1-n}]
C_{11}	Kalker's creep coefficient	[$\cdot$]
C_{22}	Kalker's creep coefficient	[$\cdot$]
C_{23}	Kalker's creep coefficient	[$\cdot$]
E	modulus of elasticity	[N/m^2]
K	wear coefficient	[m^2/N]
L	wave length	[m]
N	normal force (2D,3D)	([N/m],[N])
R	equivalent radius	[m]
R_1,R_2	radii of bodies	[m]
S	contact patch	
S_{slip}	slip zone	
S_{stick}	stick zone	
T	tangential force (2D, 3D)	([N/m],[N])
V	velocity	[m/s]

W	wear	[m]
X	longitudinal position	[m]
Y	lateral position	[m]
Z	vertical position	[m]
Z_1,Z_2	shape of undeformed bodies	[m]

Greek Letters

β_n	polynomial coefficients	[m^{-n}]
β_n^*	polynomial coefficients	[m^{-n}]
δ	penetration	[m]
Δ	shift in position of contact point	[m]
ζ	longitudinal position	[m]
η	lateral position	[m]
λ_n	geometric moment	[m^{1-n}]
μ	friction coefficient	[$\cdot$]
μ_k	kinematic friction coefficient	[$\cdot$]
ν	Poisson ratio	[$\cdot$]
ξ	creepage	[$\cdot$]
τ	time to pass half the contact patch	[s]
τ^*	time to pass half the stick zone	[s]
φ	spin	[1/m]
ϕ	phase	[degs]
Φ	spin coefficient	[$\cdot$]
ω	angular velocity	[1/s]
Ω	angular velocity	[1/s]

Indices

$*$	stick zone
0	time invariant
A	cosine term
B	sine term
F	filter
lin	linear
m	mode number
mod	modified
r	rough
R	Remington filter
SHE	Shen, Hedrick, Elkins
x	longitudinal
y	lateral
z	vertical

Notation

$\widehat{a}$	amplitude
$\tilde{N}$	normalised
$\boldsymbol{s}$	vector, matrix
$\{\}$	vector
$\{\}^T$	transposed vector
$[\,]$	matrix
$[\,]^{-1}$	inverted matrix

Contents

Chapter 1

Introduction

The theory of contact mechanics plays an important role in the description of a large variety of engineering problems e.g. roller bearings, gear wheels or the rolling deformation of bodies. In railway dynamics the contact between wheel and rail is a crucial property: it is via the forces transmitted through the contact patch that vibrations and wear are generated. Thus, it is important for the simulations of the railway dynamics to be able to make a very accurate description of the rolling contact of wheel and rail.

The first known literature on rolling contact is a surprisingly perceptive paper by Reynolds [61], who formulated some basic ideas concerning the behaviour of iron cylinders rolling on rubber surfaces. The paper contains no calculations and only few experiments and yet Reynolds' conception of fundamental contact mechanical properties is very close to what subsequently has been verified with experiments and calculations.

By the formulation of the half space theory in the 1880's a mathematical foundation for the theory of contact mechanics was introduced. The main contributors from this period are Boussinesq [4], Cerruti [11] and Hertz [27] all treating the normal contact problem. In the 1920's the first papers on two-dimensional tangential contact problem was published by Carter [9] and Fromm [18]. The two-dimensional tangential contact problem was then continuously improved until 1958 where Johnson as the first treated the three-dimensional tangential contact problem [31].

The main contributor to modern rolling contact mechanics is beyond doubt J.J. Kalker who was the first to apply modern numerical methods to contact problems. The amount of work made by Kalker is impressive and he has published a multitude of papers on all sorts of contact problems. A survey of his principles can be found in his book from 1990 [38]. There is no doubt that Kalker's theory applied on certain contact problems yields a very high degree of accuracy. The drawback of the numerical approach which Kalker apply is that the computation times are rather high and thus not very well suited for dynamical simulations.

As Kalker's theory is known to be exact, people with less insight in the field of contact mechanics sometimes apply it uncritically also when the assumptions of the theory are not fulfilled for the given contact problem. This results in simulations where important properties of the contact problem are disregarded, which in worst case can lead to qualitatively wrong results.

The objective of the present work is to derive a flexible contact theory which can be applied on contact problems not covered by the assumptions in Kalker's theory or in other common contact models. The derived model is then applied on a number of basic contact problems in order to investigate how sensitive the solutions are with respect to the properties which often are neglected in more primitive contact models.

The contact model derived in the present work is two-dimensional and based on elastic half spaces. It is obviously a weakness of the model that it only is two-dimensional, but it is demonstrated that the two-dimensional solution is qualitatively similar to the one for a three-dimensional contact problem with no spin. Thus, the results obtained with the present model can be considered also to be an indicator of the behaviour of a three-dimensional contact problem.

1.1 Outline of the Thesis

The thesis is divided into two parts. In Chapter 2–3 the theory of contact mechanics is introduced and the fundamental problems are described. In Chapter 4–7 the theory derived in the first part of the thesis is applied on a variety of contact problems. The contents of each chapter is in brief:

Chapter 2. Contact Mechanics: This serves as a general introduction to contact mechanics. Some principle problems of contact mechanics are introduced and a variety of contact models are described. Finally a new approach to the three-dimensional contact problem based on the two-dimensional solution is presented.

Chapter 3. A Polynomial Approach: The mathematical foundation of the new contact model is introduced. The objective of the approach is to transform the constitutive equation into an algebraic equation which makes it possible to calculate the stresses from a given displacement.

Chapter 4. Corrugation: The example of a cylinder rolling on a corrugated surface is investigated using the new contact model. The normal contact problem and the tangential contact problem are solved. Furthermore the wear in the contact patch is investigated in order to predict the evolution of the corrugation. Finally the application of contact filters is considered.

Chapter 5. Velocity Dependent Friction Coefficient: The tangential contact problem is investigated for the case where the friction coefficient is velocity dependent. The influence on the tangential stress distribution and the influence on the outline of the creep curve are examined.

Chapter 6. Rough Surfaces: The problem of contact between rough bodies is treated. The influence of the geometry of a roughness asperity is analysed for both the normal contact problem and for the tangential contact problem. Furthermore the cross influence between adjacent contact patches is investigated.

Chapter 7. Non-Steady Two-Dimensional Contact: Here the two-dimensional non-steady tangential contact problem is investigated. The oscillations of the tangential stress distribution are analysed and a relation between tangential force and the creepage is derived.

The thesis is concluded by Chapter 8 where a discussion and conclusion of the obtained results are carried out.

Chapter 2

Contact Mechanics

The objective of the present chapter is to outline some of the basic aspects of contact mechanics. Different approaches to solve contact problems will be treated with the emphasis on the half space method. Finally a number of solutions to the normal contact problem and to the tangential contact problem are described.

2.1 The Basic Problems

The subject of contact mechanics covers a very large variety of problems concerning the interaction between deformable bodies. In order to make the topic more clear it is divided into smaller groups depending on either the material of the bodies (e.g. elastic, plastic or viscoelastic contact) or how the bodies in contact interact (e.g. normal contact, tangential contact, rolling contact, impact etc.).

In the present work only the case of elastic contact is investigated i.e. it is assumed that no plastic deformation of the bodies takes place and that there is no time-dependent behaviour in the materials relationship between stress and strain. An introduction to inelastic contact can be found in Hill's book [28] and a more thorough investigation of viscoelastic rolling contact can be found in the work of Wang [69].

If it furthermore is assumed that the bodies in contact are quasi identical i.e. have the same material properties, then the tangential stress does not influence the normal pressure distribution. This implies that the normal contact problem can be solved independently of the tangential contact problem. As both the wheel and the rail are made of steel, the wheel/rail interaction is a typical quasi identical contact. If the bodies in contact have different material properties an iterative method must be applied e.g. the Panagiotopoulos process [55].

It is presupposed that the wheel and rail interact in a rolling motion. The bodies must thus always be in contact which implies that no impact occurs. For some special cases of wheel/rail contact this assumption is not valid e.g. heavily corrugated rails or sudden flange contact. Still the normal load is usually sufficiently large to ensure a continuous contact between wheel and rail, and so the no-impact assumption is in general reasonable.

2.2 Calculation Methods

There exists numerous different methods to solve contact problems. The choice of approach is not at all easy because it often is a choice between computational speed and accuracy. The obvious way to solve a contact problem is to utilize a finite element method (FEM). Dividing the bodies into a finite number of elements and assume the displacements and stresses in each element to be of a simple form e.g. constant or linear and then by putting the elements together by means of some compability relations [71], the stresses and strains in the bodies are found [48]. The FEM can handle very complicated geometries with a very high degree of accuracy but is much to slow to be utilized in connections where numerous consecutive contact situations must be evaluated.

A much more appropriate approach to the investigation of wheel/rail contact is the half space method. It assumes that the overall contact problem can be solved just by analysing the contact patch itself. This leads of course to major simplifications of the solution as the problem then is reduced by one dimension.

2.3 The Half Space Approach

For wheel/rail contact as well as many other applications the characteristic dimensions of the bodies in contact are much larger than the size of the contact patch. When this is the case the contact stresses do not depend on the shape of the bodies distant from the contact area, and so the bodies may be approximated by semi-infinite bodies having plane surfaces i.e. half spaces.

2.3.1 Elastic Half Space Theory

In order to apply the elastic half space theory it is necessary that some basic properties of the contact are fulfilled:

i. The characteristic sizes of the bodies in contact are large compared to the size of the contact patch

ii. the bodies are smooth

iii. strain and stresses are small

iv. The bodies are fully elastic

v. The bodies are homogeneous and isotropic

The restrictions (i) and (ii) ensure that the bodies can be considered as half spaces. Provided that the strains and stresses are small, the small strain theory can be applied [64]. The assumption of small strains and stresses is crucial because the derivation of the constitutive equations is based on the principle of superposition, which is only valid if the strains and stresses are linear.

The aim is now to derive a relation between stresses and displacements i.e. a constitutive equation. This is done by considering the influence function $\boldsymbol{g}(X, Y, \zeta, \eta)$, which should be interpreted as the displacement at the point

(X, Y) when a point load is acting at (ζ, η). Because the stresses and strains can be superposed this leads to the constitutive equation

$$\boldsymbol{u}(X, Y) = \int\int \boldsymbol{g}(X, Y, \zeta, \eta)\, \boldsymbol{q}(\zeta, \eta)\, d\zeta d\eta \tag{2.1}$$

where $\boldsymbol{u}(X, Y)$ is the displacement vector and $\boldsymbol{q}(X, Y)$ is the stress vector. The influence functions for the half space theory were found by Cerruti [11] and Boussinesq [4] and are listed in Appendix C.

The question is of course whether the half space theory can be applied for the case of wheel/rail contact. In general the restrictions mentioned above are fulfilled, but for some special cases the half space approximation may be too primitive:

1. Flange contact: as the width of a flange is of magnitude 30 mm and the characteristic size of the contact patch is approximately 10 mm, assumption (i) may be violated.
2. Heavily corrugated rails: investigations have shown that the surface material at the top of a corrugation asperity is harder than the surface material in a corrugation trough [3]. This implies that assumption (v) is not fulfilled.
3. Roughness: when the contact patch due to roughness of the wheel and rail are divided into numerous small patches, the stresses may be so large that the material will undergo a plastic deformation. In this case the assumptions (ii), (iii) and (iv) are violated.

Despite the above listed cases it must be emphasized that the half space approximation considering accuracy and computation time by far is the best approach for most investigations of wheel/rail contact.

2.3.2 The Normal Contact Problem

The normal contact problem is an overall term for the case of two elastic bodies pressed together under a normal force. The aim is to find the characteristic properties of the contact such as displacements, pressure distribution, penetration and size and shape of the resulting contact patch.

Consider two elastic bodies where the shape of the undeformed bodies are given by the functions $Z_1(X, Y)$ and $Z_2(X, Y)$. The separation between them is then defined as

$$h(X, Y) = Z_1(X, Y) - Z_2(X, Y) \tag{2.2}$$

As the bodies approach one another the first contact will take place at the point (X_0, Y_0) which is the global minimum of $h(X, Y)$. This point is in the following referred to as the contact point.

When the bodies are pressed together they will deform around the contact point and a contact patch S is created. The shape and size of the contact patch depend on the geometry of the bodies, the normal load and the characteristic material constants.

If several local minima of the function $h(X, Y)$ lie inside S the contact is said to be multiple. In the case of multiple contact, S can be either coherent or divided into more separate contact zones. The distinction between single point contact and multiple point contact is important for many applications. Another way to categorize a contact situation is by distinguishing between conforming and non-conforming contact. If the bodies have dissimilar profiles in the vicinity of the contact point the contact is non-conforming

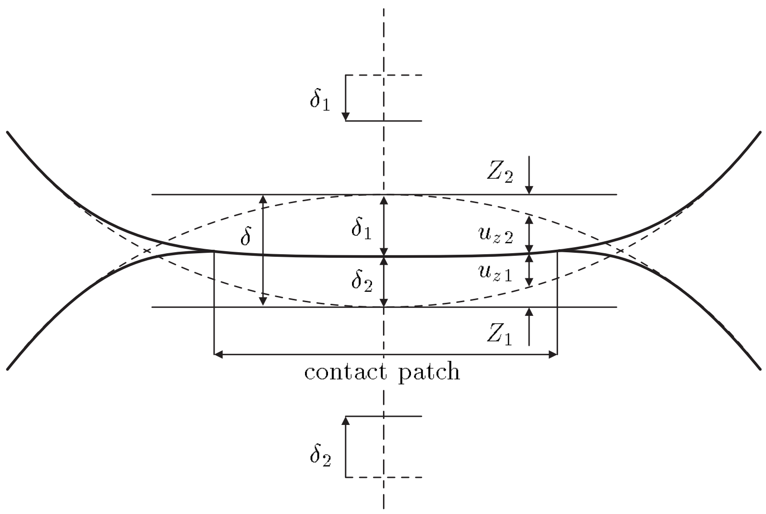

Figure 2.1. The normal deformation of elastic bodies in contact.

whereas bodies which fit almost together without deformation - i.e. a concave body and a convex body - are said to be conforming.

Now let $u_z(X, Y) = u_{z1}(X, Y) + u_{z2}(X, Y)$ be the vertical displacement of material as the bodies are deformed. Then

$$u_z(X, Y) + h(X, Y) = \delta \ , \ (X, Y) \in S \tag{2.3}$$

$$u_z(X, Y) + h(X, Y) > \delta \ , \ (X, Y) \notin S \tag{2.4}$$

where $\delta = \delta_1 + \delta_2$ is the penetration (see Figure 2.1). The relation between the vertical displacements and the normal pressure distribution is found from the constitutive equation of Cerruti-Boussinesq [32]:

$$u_z(X, Y) = \frac{2(1-\nu^2)}{\pi E} \int\!\!\int_S \frac{p(\zeta, \eta)}{\sqrt{(X-\zeta)^2 + (Y-\eta)^2}} \, d\zeta \, d\eta \tag{2.5}$$

where E is the modulus of elasticity and ν is the Poisson ratio. The contact is assumed to be frictionless or quasi identical i.e. the vertical displacements do not depend on the tangential stress. It is assumed that the normal pressure always is zero outside the contact patch so

$$p(X,Y) \neq 0 \quad \Leftrightarrow \quad (X,Y) \in S \tag{2.6}$$

$$p(X,Y) = 0 \quad \Leftrightarrow \quad (X,Y) \notin S \tag{2.7}$$

The last equation that is necessary to solve the normal contact problem arises from the fact that the normal force is equal to the normal pressure distribution integrated over the contact patch, i.e.

$$N = \iint_S p(X,Y)\, dX\, dY \tag{2.8}$$

By this the set of equations necessary to solve the normal contact problem is established. The complexity of the normal contact problem is closely related to the various types of equations. With one inequality and two integral equations, the normal contact problem is very difficult to solve and for many applications a numerical approach is the only way to obtain a solution to the normal contact problem.

2.3.3 The Tangential Contact Problem

Now consider two elastic bodies in contact. If a torque is applied to one of the bodies a tangential force will be transmitted to the other body due to the friction in the contact patch and the bodies will roll over each other. The tangential contact problem consists in finding the tangential stress distribution, the tangential displacements and the relative velocity in the contact patch.

Let the motion of the two bodies be defined with respect to a reference point which coincides with the contact point, and define the linear velocity of the bodies

$$\boldsymbol{V_1}(t) = \begin{Bmatrix} V_{x1} \\ V_{y1} \\ V_{z1} \end{Bmatrix} \tag{2.9}$$

$$\boldsymbol{V_2}(t) = \begin{Bmatrix} V_{x2} \\ V_{y2} \\ V_{z2} \end{Bmatrix} \tag{2.10}$$

and the angular velocity

$$\boldsymbol{\Omega_1}(t) = \begin{Bmatrix} \Omega_{x1} \\ \Omega_{y1} \\ \Omega_{z1} \end{Bmatrix} \tag{2.11}$$

$$\boldsymbol{\Omega_2}(t) = \begin{Bmatrix} \Omega_{x2} \\ \Omega_{y2} \\ \Omega_{z2} \end{Bmatrix} \tag{2.12}$$

then the relative velocity of the rigid bodies in the contact patch is given as

$$\boldsymbol{V}(t) = \begin{Bmatrix} V_x \\ V_y \\ \Omega_z \end{Bmatrix} = \begin{Bmatrix} V_{x1} - V_{x2} \\ V_{y1} - V_{y2} \\ \Omega_{z1} - \Omega_{z2} \end{Bmatrix} \tag{2.13}$$

Provided the bodies remain in contact, the linear vertical velocity is always zero and so the mean velocity

$$V_m = \frac{1}{2}|\boldsymbol{V_1} + \boldsymbol{V_2}| \tag{2.14}$$

is always parallel to the contact plane. The creepage is then defined as the relative velocity of the rigid bodies normalised with the mean velocity

$$\boldsymbol{\xi}(t) = \left\{ \begin{array}{c} \xi_x \\ \xi_y \\ \varphi \end{array} \right\} = \frac{1}{V_m} \left\{ \begin{array}{c} V_x \\ V_y \\ \Omega_z \end{array} \right\} \tag{2.15}$$

where ξ_x is the longitudinal creepage, ξ_y is the lateral creepage and φ is the spin.

Now introduce a new coordinate system (x, y, z) which moves along with the contact patch and let $\boldsymbol{u}^T(x, y, t) = \{u_x, u_y\}$ be the displacement in the contact plane. Defining the slip $\boldsymbol{s}^T(x, y, t) = \{s_x, s_y\}$ as the local, relative velocity in the contact patch normalised with V_m then the kinematic constraints read

$$s_x(x, y, t) = \xi_x(t) - \varphi(t)y + \frac{\partial}{\partial x}u_x(x, y, t) - \frac{1}{V_m}\frac{\partial}{\partial t}u_x(x, y, t) \tag{2.16}$$

$$s_y(x, y, t) = \xi_y(t) + \varphi(t)x + \frac{\partial}{\partial x}u_y(x, y, t) - \frac{1}{V_m}\frac{\partial}{\partial t}u_y(x, y, t) \tag{2.17}$$

A detailed derivation of the kinematic constraints can be found in [38].

In tangential contact mechanics there is often distinguished between stationary and non-steady rolling contact. If the term with $\partial \boldsymbol{u}(x, y, t)/\partial t$ is negligible the tangential contact problem is said to be stationary and the kinematic constraints are then reduced to

$$s_x(x, y) = \xi_x - \varphi y + \frac{\partial}{\partial x}u_x(x, y) \tag{2.18}$$

$$s_y(x, y) = \xi_y + \varphi x + \frac{\partial}{\partial x}u_y(x, y) \tag{2.19}$$

The omitting of the time derivative of $\boldsymbol{u}(x, y, t)$ leads to major simplifications in solving the tangential contact problem and for that reason most contact theories assume stationary contact.

The contact patch is divided into a stick zone S_{stick} and a slip zone S_{slip}:

$$(x,y) \in S_{stick} \Leftrightarrow \left\{ \begin{array}{rcl} \boldsymbol{s}(x,y,t) & = & \mathbf{0} \\ |\boldsymbol{q}(x,y,t)| & < & \mu p(x,y,t) \end{array} \right. \tag{2.20}$$

$$(x,y) \in S_{slip} \Leftrightarrow \left\{ \begin{array}{rcl} \boldsymbol{s}(x,y,t) & \neq & \mathbf{0} \\ |\boldsymbol{q}(x,y,t)| & = & \mu p(x,y,t) \end{array} \right. \tag{2.21}$$

where $\boldsymbol{q}^T(x,y,t) = \{q_x, q_y\}$ is the tangential stress distribution and μ is the friction coefficient according to the friction law of Coulomb. It is noticed that the direction of the tangential stress always is opposite the direction of the slip, i.e.

$$\left\{ \begin{array}{rcl} \boldsymbol{s}(x,y,t) \cdot \boldsymbol{q}(x,y,t) & < & 0 \\ \boldsymbol{s}(x,y,t) \cdot \boldsymbol{q}^T(x,y,t) & = & 0 \end{array} \right\} \;,\quad (x,y) \in S_{slip} \tag{2.22}$$

As the displacements tend towards zero as the distance from the contact patch increases, it follows from the kinematic constraints that

$$\lim_{(x,y)\to\infty} s_x(x,y,t) = \xi_x(t) - \varphi(t)y \tag{2.23}$$

$$\lim_{(x,y)\to\infty} s_y(x,y,t) = \xi_y(t) + \varphi(t)x \tag{2.24}$$

The relation between the tangential stress distribution and the tangential displacements is established from the constitutive equation of Cerruti-Boussinesq:

$$\boldsymbol{u}(x,y,t) \iint_S \boldsymbol{g}(x,y,\zeta,\eta)\boldsymbol{q}(\zeta,\eta,t)\, d\zeta\, d\eta \tag{2.25}$$

where $\boldsymbol{g}(x,y,\zeta,\eta,t)$ is the influence matrix with the coefficients

$$g_{11}(x,y,\zeta,\eta) = \frac{2(1+\nu)}{\pi E}\left[\frac{(1-\nu)}{\rho} + \frac{\nu(x-\zeta)^2}{\rho^3}\right] \tag{2.26}$$

$$g_{12}(x,y,\zeta,\eta) = \frac{2(1+\nu)}{\pi E}\left[\frac{\nu(x-\zeta)(y-\eta)}{\rho^3}\right] \tag{2.27}$$

$$g_{21}(x,y,\zeta,\eta) = g_{12}(x,y,\zeta,\eta) \tag{2.28}$$

$$g_{22}(x,y,\zeta,\eta) = \frac{2(1+\nu)}{\pi E}\left[\frac{(1-\nu)}{\rho} + \frac{\nu(y-\eta)^2}{\rho^3}\right] \tag{2.29}$$

$$\rho = \sqrt{(x-\zeta)^2 + (y-\eta)^2} \tag{2.30}$$

It is assumed that the bodies are quasi identical so that the normal pressure does not influence the tangential displacements.

The last equation that is necessary in order to solve the tangential contact problem states that the tangential force $\boldsymbol{T}^T(t) = \{T_x, T_y\}$ is equal to the tangential stress distribution integrated over the entire contact patch:

$$\boldsymbol{T}(t) = \iint_S \boldsymbol{q}(x,y,t)\,dx\,dy \tag{2.31}$$

As the area of the contact patch is finite, a moment $M_z(t)$ acting about the normal to the contact plane is generated:

$$M_z(t) = \iint_S \left[q_y(x,y,t)x - q_x(x,y,t)y\right]\,dx\,dy \tag{2.32}$$

In most applications the size of the contact patch is however so small compared to other characteristic sizes that the influence of the moment in the contact patch can be neglected.

The solution to the tangential contact problem is not unique: for a given creepage an infinity of tangential stress distributions fulfill the equations derived in this section. The physical explanation to this apparently non-physical behaviour is that only the solution that minimizes the tangential force is stable: all other solutions are unstable and will only occur in a transition phase.

2.4 Solutions to the Two-Dimensional Contact Problem

In the theory of contact mechanics a problem is often referred to as two-dimensional or three-dimensional. The three-dimensional contact problem handles real bodies whereas the two-dimensional contact problem is reduced by one degree of freedom so the bodies in contact have a characteristic cross section which is constant in the direction perpendicular to the direction of motion. The most often investigated two-dimensional contact problem is the one of two infinite cylinders in rolling contact. The omitting of one dimension is naturally strictly speaking non-physical. It is however a case of interest as the simplification makes it possible to solve certain contact problems analytically. Furthermore the qualitative behaviour of a two-dimensional contact problem is in many cases equivalent to the one of a three-dimensional case as will be demonstrated in section 2.6.

2.4.1 The Hertz Solution

When two infinite cylinders with the radii R_1 and R_2 are pressed together under the normal load per unit length N, they deform around the contact line and a contact strip is created. As the problem is two-dimensional only a cross section of the cylinders is considered and so the contact line is transformed into a contact point and the contact strip into a line. For historical reasons this is however still referred to as the contact patch. Assuming that the radii of the cylinders are much larger than the size of contact patch, the shape of the cylinders in the vicinity of the contact point can be approximated by the second order Taylor expansions

$$Z_1(x) \simeq \frac{1}{2R_1}x^2 \tag{2.33}$$

$$Z_2(x) \simeq -\frac{1}{2R_2}x^2 \tag{2.34}$$

Let the length of the contact patch be $2a_0$, then the vertical displacement of material in the contact patch is found from the constitutive equation which for the two-dimensional case reads:

$$u_z(x) = -\frac{4(1-\nu^2)}{\pi E}\int_{-a_0}^{a_0} p(\zeta)\ln(x-\zeta)\, d\zeta + C_1 \tag{2.35}$$

where the constant C_1 depends on the choice of datum of the displacements. This implies that the equations (2.3)-(2.4) can only be solved apart from a constant. To avoid this unknown constant the two equations are differentiated with respect to x and so C_1 disappears and the problem can be solved. Unfortunately the penetration δ is also independent of x and will thus be removed by the differentiation. Thus, it is impossible to find the penetration for the two-dimensional contact problem. The constitutive equation derived with respect to x reads

$$\frac{du(x)}{dx} = -\frac{4(1-\nu^2)}{\pi E}\int_{-a_0}^{a_0} \frac{p(\zeta)}{x-\zeta}\, d\zeta \tag{2.36}$$

and so

$$\frac{4(1-\nu^2)}{\pi E}\int_{-a_0}^{a_0} \frac{p(\zeta)}{x-\zeta}\, d\zeta = \frac{d}{dx}\left[Z_1(x) - Z_2(x)\right] \quad , \quad -a_0 < x < a_0 \tag{2.37}$$

Introducing the equivalent radius R as

$$\frac{1}{R} = \frac{1}{R_1} + \frac{1}{R_2} \tag{2.38}$$

equation (2.37) is reduced to

$$\frac{4(1-\nu^2)}{\pi E}\int_{-a_0}^{a_0} \frac{p(\zeta)}{x-\zeta}\, d\zeta = \frac{x}{R} \quad , \quad -a_0 < x < a_0 \tag{2.39}$$

The normal contact problem was first solved in 1882 by Hertz [27] who found that the normal pressure distribution is elliptic with

$$p(x) = \frac{p_0}{a_0}\sqrt{a_0^2 - x^2} \quad , \quad -a_0 < x < a_0 \tag{2.40}$$

$$p_0 = \sqrt{\frac{NE}{2(1-\nu^2)\pi R}} \tag{2.41}$$

$$a_0 = \sqrt{\frac{8(1-\nu^2)RN}{\pi E}} \tag{2.42}$$

It should be noticed that the shape of the two surfaces are represented linearly in equation (2.37). This implies that geometric properties can be moved from one body to another and so the case of two cylinders pressed together can always be transformed into the case of a cylinder with the equivalent radius R pressed into a level surface (a cylinder with radius equal infinity).

2.4.2 The Carter Solution

Now apply an axial torque to one of the cylinders from the previous section. Due to the friction in the contact patch a tangential force will be transmitted between the bodies and the cylinders will roll over each other. Let the relative global velocity of the two cylinders be non zero i.e.

$$\xi = \frac{1}{V_m}\left(\omega_1 R_1 - \omega_2 R_2\right) \tag{2.43}$$

where ω_1 and ω_2 are the angular velocities of the cylinders and V_m is the mean velocity

$$V_m = \frac{1}{2}\left|\omega_1 R_1 + \omega_2 R_2\right| \tag{2.44}$$

The kinematic constraint then gives a relation between the slip $s(x,t)$ and the derivatives of the tangential displacement $u_x(x,t)$

$$s(x,t) = \xi(t) + \frac{\partial u_x(x,t)}{\partial x} - \frac{1}{V_m}\frac{\partial u_x(x,t)}{\partial t} \tag{2.45}$$

where the constitutive equation provides the relation between the tangential stress distribution and the tangential displacement of the material in the contact patch:

$$u_x(x,t) = -\frac{4(1-\nu^2)}{\pi E}\int_{-a_0}^{a_0} q(\zeta)\ln(x-\zeta)\,d\zeta + C_2(t) \tag{2.46}$$

Analogue to the constitutive equation for the normal contact problem $C_2(t)$ depends on the choice of datum of the displacements. This implies that $\partial u(x,t)/\partial t$ can only be found apart from an unknown function, and it is thus impossible to solve the two-dimensional tangential contact problem for the non-steady case.

The stationary tangential contact problem for a Hertzian normal pressure distribution was solved by Carter [9] in 1926 and by Fromm [18] in 1927. Whereas Fromm succeeded in solving the problem of two-dimensional elasticity without the use of a half space approximation, Carter regarded the cylinders as elastic half spaces. They both found that the tangential stress distribution can be calculated as the sum of two ellipses. A new coordinate system where the transformation between the old coordinates and the new coordinates is given as

$$x^* = x + a_0 - a_0^* \tag{2.47}$$

is introduced so that one of the ellipses has its centre in $O(x)$ and the other in $O(x^*)$ as indicated in Figure (2.2).

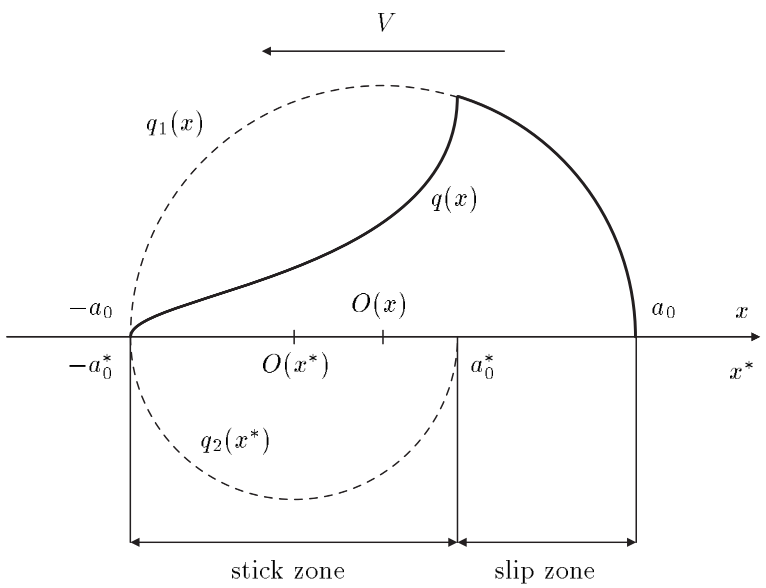

Figure 2.2. The Carter solution for the two-dimensional tangential contact problem.

The tangential stress distribution $q(x)$ is then

$$q(x) = q_1(x) + q_2(x^*) \tag{2.48}$$

$$q_1(x) = \begin{cases} \frac{\mu p_0}{a_0}\sqrt{a_0^2 - x^2} & , \quad -a_0 < x < a_0 \\ 0 & , \quad \text{otherwise} \end{cases} \tag{2.49}$$

$$q_2(x^*) = \begin{cases} -\frac{\mu p_0}{a_0}\sqrt{a_0^{*2} - x^{*2}} & , \quad -a_0^* < x^* < a_0^* \\ 0 & , \quad \text{otherwise} \end{cases} \tag{2.50}$$

where the size of a_0^* depends on the size of the contact length and the magnitude of the creepage

$$a_0^* = a_0 + \frac{R}{\mu}\xi_0 \ \ , \ \ -\mu a_0/R \leq \xi_0 \leq 0 \tag{2.51}$$

By integrating the tangential stress distribution over the entire contact patch the tangential force is found, and so the creepage and the tangential force are given as functions of a_0^*:

$$\xi_0 = \frac{\mu}{R}\left(a_0^* - a_0\right) \tag{2.52}$$

$$T_0 = \frac{\mu N}{a_0^2}\left(a_0^2 - a_0^{*2}\right) \tag{2.53}$$

The above expressions are only valid for $a_0^* \leq a_0$ which yields the critical creepage

$$\xi_c = -\frac{\mu a_0}{R} \tag{2.54}$$

This is exactly the value for which $T = \mu N$ i.e. where the tangential force is equivalent to the tangential force according to the friction law of Coulomb. Consequently the cases where $|\xi_0| \geq |\xi_c|$ are referred to as complete sliding.

The classic way to evaluate a tangential contact problem is via a creep curve, where the tangential force is plotted as a function of the creepage. The creep curve for the Carter solution is shown in Figure 2.3 where the creepage is normalised with ξ_c. It is seen that when the size of the creepage is small the tangential force is below the Coulomb value μN, whereas it for a certain size of the creepage will reach the saturated regime where complete sliding occurs and the tangential force will then be equal to the Coulomb value.

If the tangential stress distribution is inserted into the constitutive equation and into the kinematic constraint the local relative velocity between the

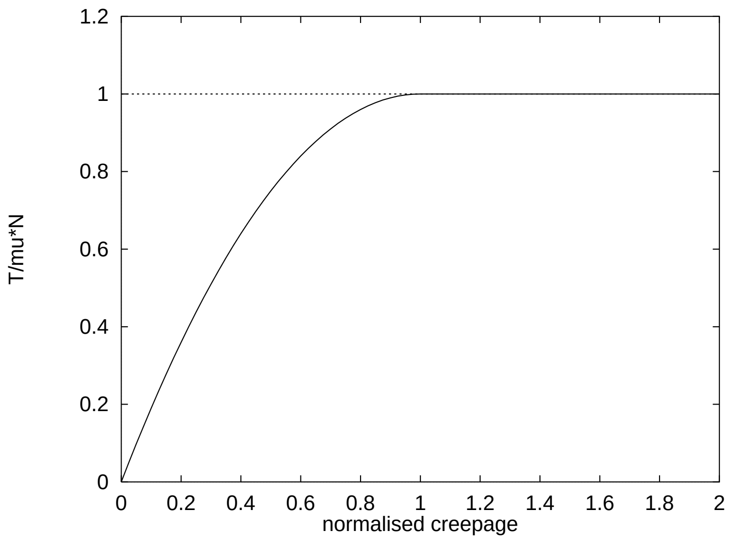

Figure 2.3. Creep curve for the Carter solution.

bodies is found to be

$$s(x) = \begin{cases} -\frac{\mu}{R}\left(\sqrt{x^2 - a_0^2} - \sqrt{x^{*2} - a_0^{*2}}\right) & , \quad x \le -a_0 \\ 0 & , \quad -a_0 < x < 2a_0^* - a_0 \\ -\frac{\mu}{R}\sqrt{x^{*2} - a_0^{*2}} & , \quad 2a_0^* - a_0 \le x < a_0 \\ \frac{\mu}{R}\left(\sqrt{x^2 - a_0^2} - \sqrt{x^{*2} - a_0^{*2}}\right) & , \quad a_0 \le x \end{cases} \tag{2.55}$$

This implies that the contact patch is divided into a stick zone with the length $2a_0^*$ at the leading edge and a slip zone at the trailing edge (Figure 2.1 and Figure 2.4):

$$S_{stick} = \{x \in S \mid -a_0 < x < 2a_0^* - a_0\} \tag{2.56}$$

$$S_{slip} = \{x \in S \mid 2a_0^* - a_0 \le x < a_0\} \tag{2.57}$$

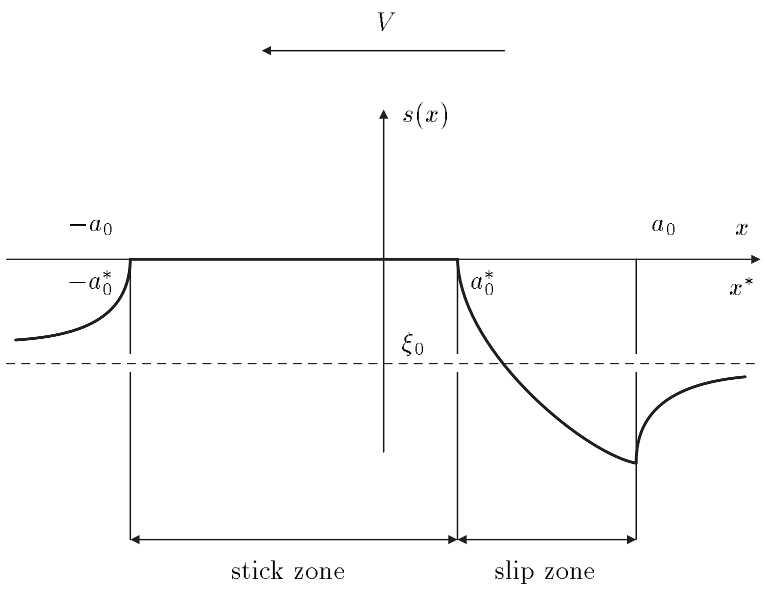

Figure 2.4. Local relative velocity for the Carter solution.

It is furthermore seen that

$$\lim_{x \to \pm\infty} s(x) = \xi_0 \tag{2.58}$$

which states the obvious property that the local relative velocity between the cylinders at a position far away from the contact patch is equal to the global relative velocity.

It can be shown that the magnitude of the creepage only depends on the position of the small ellipse and that the tangential force only is related to the size of the small ellipse. This implies that there is no correlation between creepage and tangential force: any position and size of the small ellipse satisfies the equations for the tangential contact problem as long as it lies inside the big ellipse. As stated in section 2.3.3 this is due to

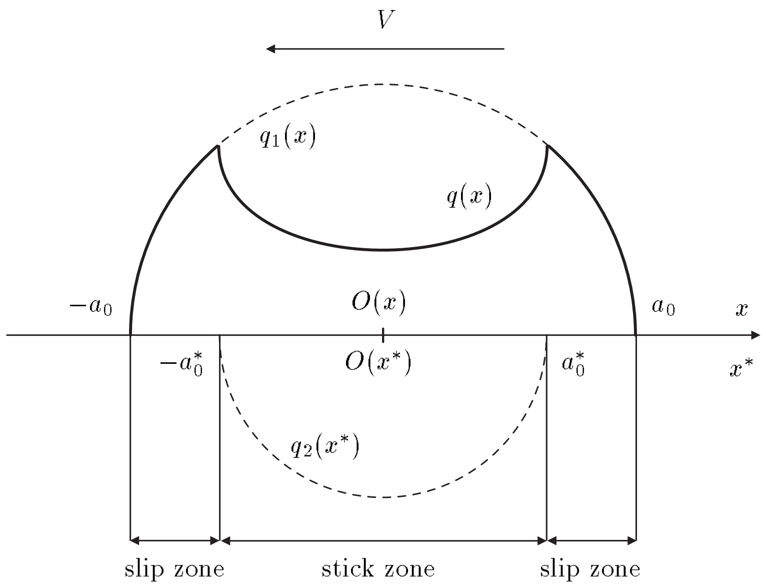

Figure 2.5. The Cattaneo solution to the two-dimensional tangential contact problem.

the fact that only the solution which for a given creepage minimizes the tangential force is a stable solution to the tangential problem. In Appendix A it is shown that the Carter solution is the only stable solution to the two-dimensional stationary contact problem.

The Cattaneo solution is a typical example of an instable solution to the two-dimensional contact problem. Consider two cylinders pressed together and then shifted in the tangential direction under the application of a sudden tangential force. Cattaneo [10] then found that the tangential stress distribution when the rolling motion starts can be described as the sum of two ellipses with the same origo but different widths (see Figure 2.5). As indicated in Appendix A this leads to a situation with zero creepage which is obvious because the rolling is about to start. The tangential force is however different from zero and thus bigger than the lowest possible force which is zero as in the Carter solution. Thus, the Cattaneo solution is unstable and the tangential stress distribution and the creepage will evolve until the Carter solution is reached. The transition from the Cattaneo solution to the Carter solution has been investigated by Kalker [36].

2.5 Solutions to the Three-Dimensional Contact Problem

In the previous section some analytical solutions to the two-dimensional contact problem were introduced. Regarding three-dimensional problems the solutions are much more complicated and cannot be found analytically. Instead it is necessary to apply other methods to deal with the problem. In the following sections some different approaches are outlined. The presentation serves as an introduction to the various methods, and only a few examples are mentioned in each category. A more complete listing of solutions to the three-dimensional contact problem can be found in [38].

2.5.1 Analytical Solutions

As in the two-dimensional case there also exists a three-dimensional Hertz solution to the normal contact problem. If the bodies in contact in the vicinity of the contact point are non-conforming and can be approximated by second order polynomials the contact patch will be elliptic and the normal pressure distribution will also be elliptic

$$S = \left\{ (x,y) \in \mathbb{R} \;\middle|\; \left[1 - (x/a_0)^2 - (y/b_0)^2\right] > 0 \right\} \tag{2.59}$$

$$p(x,y) = p_0 \sqrt{1 - \left(\frac{x}{a_0}\right)^2 - \left(\frac{y}{b_0}\right)^2} \tag{2.60}$$

The expressions for the semi axes a_0 and b_0 of the contact ellipse and for the maximum normal pressure p_0 are very complex. The Hertz solution for the three-dimensional normal contact problem is nevertheless by far the most convenient contact model because the elliptic properties make it possible to describe the normal contact with very few characteristic parameters. The derivation of the three-dimensional Hertz solution can be found in [32].

Regarding a solution to the stationary tangential contact problem, Kalker has made an analytical approach. With the assumption that the normal contact is Hertzian and that the stick zone covers the entire contact patch, Kalker solved the constitutive equations by assuming the tangential stress distribution to be on a polynomial form [34]. In principle it is possible to apply an arbitrary high order of the polynomial approximation, but as the complexity of the calculations explodes with the number of polynomial coefficients, Kalker restricted himself to a 5th order approximation. With this he found a linear relation between the tangential force and the creepage

$$T_{x,lin} = -\frac{E}{2(1+\nu)}\, a_0 b_0 C_{11} \xi_x \tag{2.61}$$

$$T_{y,lin} = -\frac{E}{2(1+\nu)}\, a_0 b_0 \left(C_{22}\xi_y + \sqrt{a_0 b_0} C_{23} \varphi\right) \tag{2.62}$$

where C_{11}, C_{22} and C_{23} are the Kalker creep coefficients which depend on a_0 and b_0. The creep coefficients are listed in Appendix B.

2.5.2 Heuristic Solutions

If the two bodies are conforming or the contact is multiple, the Hertzian theory is no longer valid. Because the theory of Hertz is the foundation of many methods for the tangential contact problem, it is of major interest to transform a non-Hertzian contact into an equivalent Hertzian contact, i.e. to define an elliptic contact which has the same properties as the actual contact situation. One example of establishing an equivalent contact patch for a multiple contact was made by Pascal and Sauvage [56]. They calculated the Hertzian contact for each minimum of $h(X, Y)$ in the contact patch and then established the properties of one resulting elliptic contact patch by a weighted sum of the characteristic properties of the small contact ellipses.

Concerning the tangential contact problem, the disadvantage of Kalker's linear theory presented in the previous section is that it is derived for the case where the stick zone covers the entire contact patch, i.e. for infinite small creepages. A direct consequence of this assumption is that the tangential force does not reach a saturated regime and thus violates the friction law of Coulomb. To compensate for this non-physical behaviour Shen, Hedrick and Elkins (SHE) have developed a heuristic modification of Kalker's linear theory so that the friction law of Coulomb is fulfilled when the creepage is large.

Let $T_{lin} = \sqrt{T_{x,lin}^2 + T_{y,lin}^2}$ and define the size of the tangential force according to SHE as

$$T_{SHE} = \begin{cases} \mu N \left[\left(\frac{T_{lin}}{\mu N}\right) - \frac{1}{3}\left(\frac{T_{lin}}{\mu N}\right)^2 + \frac{1}{27}\left(\frac{T_{lin}}{\mu N}\right)^3 \right] & , \quad T_{lin} \leq 3\mu N \\ \mu N & , \quad T_{lin} > 3\mu N \end{cases} \tag{2.63}$$

then the lateral and the longitudinal force components are

$$T_{x,SHE} = T_{x,lin}\frac{T_{SHE}}{T_{lin}} \tag{2.64}$$

$$T_{y,SHE} = T_{y,lin}\frac{T_{SHE}}{T_{lin}} \tag{2.65}$$

The theory of SHE is based on the theory of Vermeulen-Johnson [66] but is still a pure mathematical modification of the linear theory, made in order to obtain a saturated regime. This implies that SHE is a macroscopic model which can only be applied to establish a creep curve: it is absolutely inadequate to use for an evaluation of what happens inside the contact patch such as finding the slip and the stress or determining the location of the stick zone and of the slip zone.

The range of validity of SHE is limited. Whereas the theory provides a very good approximation for pure creepage it is much to inaccurate when the spin is large. SHE is however a very often used approximation as it is analytical and thus easy to implement and very fast. A modified version of SHE was introduced by Zhang [70] in order to ameliorate the accuracy for large spin.

Because SHE is based on the linear theory of Kalker it presupposes a Hertzian contact patch and a stationary contact.

2.5.3 Numerical Solutions

Because the equations of the half space theory in general are impossible to solve analytically, a natural approach would be to use numerical methods based on discretisation of the contact patch. The idea is to divide the contact patch into many small surface elements and then assume the stresses and displacements to be constant within each element. In this way the integral equations from the half space theory are transformed into matrix equations which usually can be solved quite easily. The discretisation method can be applied for both the normal contact problem and the tangential contact problem, but an often used simplification is to assume the contact to be Hertzian i.e. the contact patch to be elliptic and then only utilize the discretisation method for the tangential contact problem.

Divide the contact patch into n surface elements and assume the displacements and stresses inside each element to be constant and acting in the centre of the contact element. The constitutive equation can then be written as

$$\{\boldsymbol{u}\} = [\boldsymbol{G}]\{\boldsymbol{q}\} \tag{2.66}$$

$$\{\boldsymbol{u}\}^T = \{\boldsymbol{u}_1^T, \boldsymbol{u}_2^T, \ldots, \boldsymbol{u}_n^T\} \tag{2.67}$$

$$\{\boldsymbol{q}\}^T = \{\boldsymbol{q}_1^T, \boldsymbol{q}_2^T, \ldots, \boldsymbol{q}_n^T\} \tag{2.68}$$

$$[\boldsymbol{G}] = \begin{bmatrix} \boldsymbol{G}_{11} & \cdots & \boldsymbol{G}_{1n} \\ \vdots & \ddots & \vdots \\ \boldsymbol{G}_{n1} & \cdots & \boldsymbol{G}_{nn} \end{bmatrix} \tag{2.69}$$

$$\boldsymbol{G}_{ij} = \int\!\!\int_{S_j} \boldsymbol{g}(x_i, y_i, \zeta, \eta)\, d\zeta d\eta \tag{2.70}$$

where S_j is the jth surface element and (x_i, y_i) is the centre of the ith surface element. The vector $\boldsymbol{u}_j^T = \{u_{xj}, u_{yj}\}$ is the displacement in the jth element and similarly $\boldsymbol{q}_j^T = \{q_{xj}, q_{yj}\}$ is the tangential stress in the jth element.

Defining the creepage vector

$$\{\boldsymbol{\xi}\}^T = \{\boldsymbol{\xi}_1^T, \boldsymbol{\xi}_2^T, \ldots, \boldsymbol{\xi}_n^T\} \tag{2.71}$$

$$\boldsymbol{\xi}_j = \left\{ \begin{array}{c} \xi_x - \varphi y_j \\ \xi_y + \varphi x_j \end{array} \right\} \tag{2.72}$$

and the slip vector

$$\{\boldsymbol{s}\}^T = \{\boldsymbol{s}_1^T, \boldsymbol{s}_2^T, \ldots, \boldsymbol{s}_n^T\} \tag{2.73}$$

where $\boldsymbol{s}_j^T = \{s_{xj}, s_{yj}\}$ then the kinematic constraint for the stationary tangential contact problem reads

$$\{\boldsymbol{s}\} = \{\boldsymbol{\xi}\} + \frac{d}{dx}\{\boldsymbol{u}\} \tag{2.74}$$

The simplest way to solve this matrix equation is to assume that the stick zone covers the entire contact patch apart from an infinitely narrow strip at the trailing edge where a singularity occurs. This set-up is equivalent to Kalker's linear theory.

The boundary conditions then yield that the tangential stress is zero at the leading edge of the contact patch and that the slip is different from zero at the trailing edge. Now let the surface elements at the leading edge be numbered from 1 to nl and the surface elements at the trailing edge be the ones from $(n - nl)$ to n, then

$$\boldsymbol{s}_j = \mathbf{0}\ ,\ j = 1, 2, \ldots, (n - nl) \tag{2.75}$$

and the matrix equation (2.74) can be rewritten as

$$\frac{d}{dx}\left[\begin{array}{ccc} \boldsymbol{G}_{11} & \cdots & \boldsymbol{G}_{1,n-nl} \\ \vdots & \ddots & \vdots \\ \boldsymbol{G}_{n-nl,1} & \cdots & \boldsymbol{G}_{n-nl,n-nl} \end{array} \right] \left\{ \begin{array}{c} \boldsymbol{q}_{nl+1} \\ \vdots \\ \boldsymbol{q}_n \end{array} \right\} = \left\{ \begin{array}{c} \boldsymbol{\xi}_1 \\ \vdots \\ \boldsymbol{\xi}_{n-nl} \end{array} \right\} \tag{2.76}$$

which is solved with respect to the stresses. Finally the tangential force is found as

$$\boldsymbol{T} = \sum_{j=nl+1}^{n} \left[\boldsymbol{q}_j \iint_{S_j} d\zeta d\eta \right] \tag{2.77}$$

The accuracy of the method depends obviously on the discretisation: the more surface elements the better is the approximation. Kalker compares in [34] the creep coefficients C_{11}, C_{22} and C_{23} calculated with the discretisation method with the ones from his linear theory. It turns out that with a discretisation of $[11 \times 8]$ elements the error varies from $5 - 10\%$.

The above example demonstrates the fundamental principles of the discretisation method, but it is not very convenient for practical use since it is equivalent to the linear theory which is analytical and thus faster and more precise. The real application of discretisation methods is for cases where both a stick zone and a slip zone exist inside the contact patch. This results of course in an augmentation of unknowns in equation (2.74) as the assumption from equation (2.75) is no longer valid. Even when the boundary conditions (2.20)-(2.22) are introduced the equation system (2.74) is still unconstraint. This implies that an infinity of solutions exists just as in the two-dimensional contact problem. To overcome this problem the potential energy must be minimized which is done with a variational principle. A thorough description of this approach can be found in [38].

One of the most common numerical routines to solve contact problems is CONTACT by Kalker [67], a program which is based on discretisation of the contact patch and covers a broad range of contact problems. CONTACT is often referred to as the exact theory. This is of course an exaggeration as it is a numerical method, but it is beyond doubt that the routine is a very powerful tool to investigate half space contacts. The drawback of CONTACT is the computation time needed to solve a single contact problem. As the contact patch must be very fine discretized and a number of iterations is needed to solve the variational problem, CONTACT is not well suited for investigations which demand many consecutive contact calculations as railway dynamic simulations or wear calculations.

2.5.4 Tables

A way to avoid the long computation times for the numerical solutions is to use tables. This approach is based on the numerical solutions, but instead of solving the contact problem every time, many different contact situations are solved once and for all and then listed in a large table. To find the tangential force for a given contact situation it is thus a matter of interpolation in the table.

The advantage of tables is that the interpolation is much faster than computing the numerical solution. Of course this demands a large storing capacity, but that is not a problem with the computers of today. A much more sophisticated problem is how to define a given contact situation: it craves for much generality but yet as few entries as possible. It is crucial to utilize appropriate normalisations of the characteristic values in order to incorporate as many constants as possible. Typical entries of a contact table would be the semi axes of the contact ellipse and the creepage components.

Due to the limited number of entries, the use of tables is only appropriate for Hertzian contact: the information necessary to determine a non-elliptic contact patch would demand too many entries. For the same causes the use of tables is restricted to stationary contact. USETAB is an example of a contact table. It was made by Kalker [39] and is based on CONTACT. To give an impression of the magnitude of the contact tables, it is noted that the storage capacity needed for USETAB is 4.5 MB. When comparing the computation speed USETAB is 15.000 times faster than CONTACT.

For some applications the interpolation procedure may cause some numerical problems. If the interpolation in the table is linear the values are represented as piecewise linear functions and are thus non-smooth. If these values are used for e.g. numerical integration the non-smoothness may result in convergence problems. This can naturally be prevented by a higher order of interpolation, but this increases the complexity of the interpolation procedure and only moves the non-smoothness of the variable one level up.

2.6 A Three-Dimensional Solution Based on Two-Dimensional Contact Theory

The previous sections clearly indicate that where the two-dimensional contact problem for certain presumptions can be solved analytically, the three-dimensional contact problem in general must be approached by employing numerical methods. As the set of equations defining the three-dimensional contact problem in many ways is similar to the one from the two-dimensional problem, it is obvious to try to derive a three-dimensional solution by modifying the solution to the two-dimensional contact problem.

2.6.1 The Strip Theory

The concept of employing a two-dimensional approach to solve a three-dimensional contact problem was first introduced by Haines and Ollerton in 1963 [24] for the case of an elliptic contact patch subjected to a longitudinal traction. By dividing the contact patch into narrow strips parallel to the rolling direction, they argued that each strip would have a tangential stress distribution equal to the Carter solution. In the following the basic equations of the strip theory will be derived.

Let the contact patch be an ellipse with the semi-axes a_0 and b_0 then the normal pressure distribution according to Hertz is given as

$$p(x,y) = p_0\sqrt{1-\left(\frac{x}{a_0}\right)^2-\left(\frac{y}{b_0}\right)^2} \tag{2.78}$$

The length of the contact strip located at the lateral position $y = y_0$ is then $2a(y_0)$ where $a(y) = a_0\sqrt{1-(y/b_0)^2}$. With this definition the Carter solution for a contact strip is formulated as

$$q(x,y) = \mu p(x,y) - q_2(x,y) \tag{2.79}$$

$$q_1(x,y) = \frac{\mu p_0}{a_0}\sqrt{[a(y)]^2 - x^2} \tag{2.80}$$

$$q_2(x,y) = \frac{\mu p_0}{a_0}\sqrt{[a^*(y)]^2 - x^{*2}} \tag{2.81}$$

where $x^* = [x - a(y) + a^*(y)]$ and $a^*(y)$ is half the length of the stick zone (see Figure 2.6). The two-dimensional constitutive equation applied on a given strip then yields that

$$\frac{\partial u_x(u,y)}{\partial x} = \frac{4(1-\nu^2)}{\pi E}\frac{\mu p_0}{a}[a(y) - a^*(y)] \tag{2.82}$$

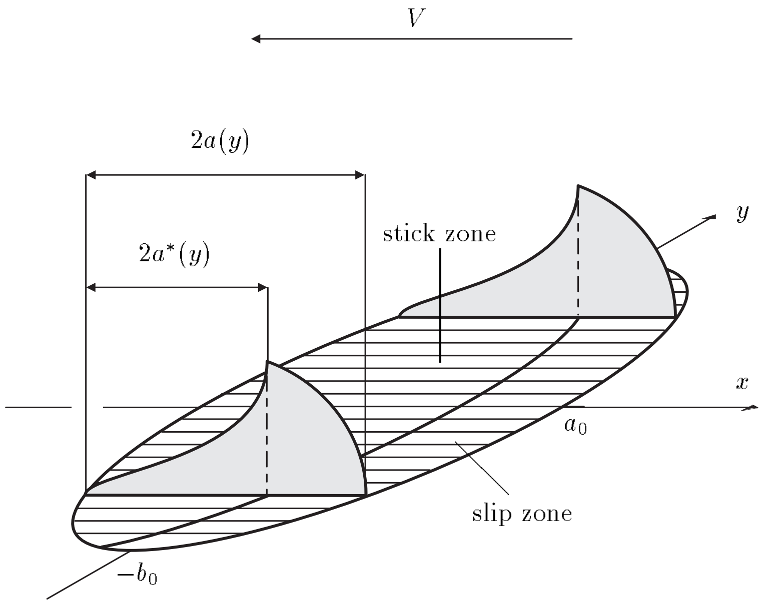

Figure 2.6. Basic concept of the strip theory.

which inserted into the two-dimensional kinematic constraint provides a relation between $a^*(y)$ and the creepage:

$$a^*(y) = a_0 \left(\frac{E}{4(1-\nu^2)\mu p_0}\xi_x + \sqrt{1-\left(\frac{y}{b_0}\right)^2} \right) \tag{2.83}$$

If complete sliding occurs in a strip then $a^*(y) = 0$. This implies that the stick zone, which is symmetric around the x-axis, is limited to the interval $y \in [-b^*; b^*]$ where

$$b^* = b_0\sqrt{1-\tilde{\xi}_x^2} \tag{2.84}$$

and where the normalised creepage $\tilde{\xi}_x$ is defined as

$$\tilde{\xi}_x = \frac{E}{4(1-\nu^2)\mu p_0}\xi_x \tag{2.85}$$

This definition further implies that complete sliding for the entire contact patch occurs when $\tilde{\xi}_x = -1$.

With $a^*(y)$ defined, the tangential stress distribution is given by the equations (2.79)–(2.81). Integrating the stress distribution over the entire contact patch the tangential force is found as

$$T_x = \mu N - \int_{-b^*}^{b^*} \int_{-a^*(y)}^{a^*(y)} \frac{\mu p_0}{a_0} \sqrt{\left[a^*(y)\right]^2 - x^{*2}}\, dx^*\, dy \tag{2.86}$$

which can be solved to

$$T_x = \mu N \left[1 - \left(1 + \frac{1}{2}\tilde{\xi}_x^2\right)\sqrt{1 - \tilde{\xi}_x^2} - \frac{3}{2}\tilde{\xi}_x \arcsin\left(\sqrt{1 - \tilde{\xi}_x^2}\right)\right] \tag{2.87}$$

where $-1 \leq \tilde{\xi} \leq 0$. If $\tilde{\xi}_x < -1$ complete sliding occurs and the tangential force is equal the Coulomb value $T_x = \mu N$. By this the tangential stress distribution is found plus a relation between tangential force and creepage is established and so the three-dimensional contact problem is solved for the case of pure longitudinal traction.

2.6.2 A Modified Strip Theory

The problem with the strip theory is that it is based on the two-dimensional constitutive equation and thus neglects the cross influence between the stress distributions in the contact strips. The consequence of this simplification is obvious when the initial slope of the creep curve is investigated. Kalker's linear theory which can be considered to be exact when $\xi \to 0$ results in the initial slope

$$\left.\frac{dT_x}{d\xi_x}\right|_{\xi_x=0} = -\frac{E}{2(1+\nu)} a_0 b_0 C_{11} \tag{2.88}$$

which indicates that the initial slope depends on C_{11} and thus on the ratio a_0/b_0. As the strip theory provides the slope

$$\left.\frac{dT_x}{d\xi_x}\right|_{\xi_x=0} = -\frac{E}{2(1-\nu^2)}a_0 b_0 \frac{\pi^2}{4} \tag{2.89}$$

which does not depend on the ratio a_0/b_0 , the strip theory is in principle only valid for one shape of the contact patch namely when $C_{11} = \pi^2/4$ i.e. for cases where $b_0 \gg a_0$. It is however possible to modify the strip theory so it is valid for all shapes of the contact patch namely by introducing the modified creepage

$$\tilde{\xi}_{x,mod} = \frac{4(1-\nu)C_{11}}{\pi^2}\tilde{\xi}_x \tag{2.90}$$

When this new creepage is inserted into equation (2.87) the initial slope of the creep curve will be equal to the one obtained by Kalker's linear theory. In Figure 2.7 the creep curve for the strip theory with the modified creepage ($\diamond$) is compared with creep curves obtained with CONTACT (solid line) and SHE (dashed line). For the stationary contact with Hertzian normal pressure CONTACT can be considered as an exact solution where SHE is merely an approximation. It is seen that the strip theory is very close to the result from CONTACT and more accurate than SHE. It must be emphasized that the accuracy of the modified strip theory does not depend on the shape of the contact ellipse. So it can be concluded that for the case of pure longitudinal traction the modified strip theory provides an analytical solution which is more accurate than SHE.

Another problem with the original strip theory is, that it only considers tangential problems with longitudinal creepage. Kalker has made an approach to the strip theory where also the lateral creepage and the spin are taken into account [35], but this leads to very complicated calculations based on numerical approximations, and so the concept of a three-dimensional theory

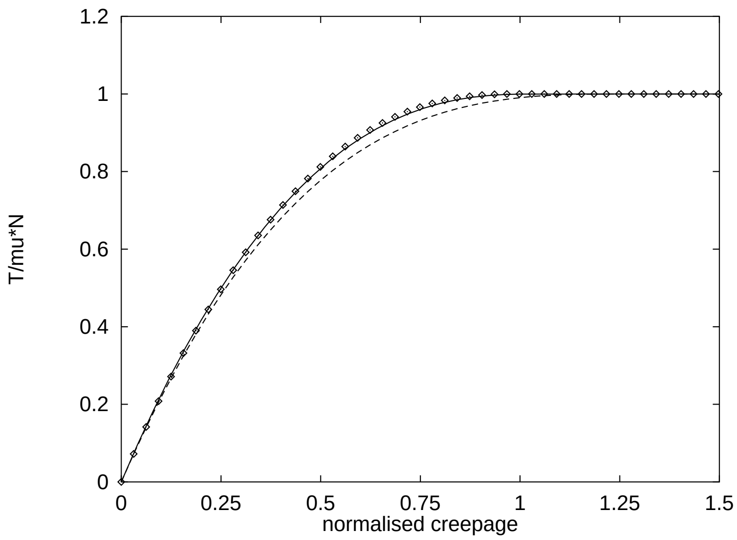

Figure 2.7. Longitudinal tangential force when $\xi_y = 0$. Solid line: CONTACT. Dashed line: SHE. $\diamond$: modified strip theory.

based on the simplicity of the two-dimensional solution vanishes. The inclusion of lateral creepage can be made fairly easily by dividing the contact problem into two separate problems: one for the longitudinal traction and one for the lateral traction. Applying the strip theory on both problems the following expressions are derived:

$$T_x = \mu N \frac{p_{0x}}{p_0} \left[1 - \left(1 + \frac{1}{2}\tilde{\xi}^2_{x,mod} \right) \sqrt{1 - \tilde{\xi}^2_{x,mod}} - \frac{3}{2}\tilde{\xi}_{x,mod} \arcsin\left(\sqrt{1 - \tilde{\xi}^2_{x,mod}} \right) \right] \tag{2.91}$$

$$T_y = \mu N \frac{p_{0y}}{p_0} \left[1 - \left(1 + \frac{1}{2}\tilde{\xi}^2_{y,mod} \right) \sqrt{1 - \tilde{\xi}^2_{y,mod}} - \frac{3}{2}\tilde{\xi}_{y,mod} \arcsin\left(\sqrt{1 - \tilde{\xi}^2_{y,mod}} \right) \right] \tag{2.92}$$

where the modified creepages are defined from Kalker's linear theory as

$$\tilde{\xi}_{x,mod} = \frac{EC_{11}}{\pi^2(1+\nu)\mu p_{0x}}\xi_x \tag{2.93}$$

$$\tilde{\xi}_{y,mod} = \frac{EC_{22}}{\pi^2(1+\nu)\mu p_{0y}}\xi_y \tag{2.94}$$

in order to compensate for the neglected cross influence between the contact strips. The unknown stresses p_{0x} and p_{0y} are found from the constraint that

$$p_0^2 = p_{0x}^2 + p_{0y}^2 \tag{2.95}$$

plus the assumption that the ratio between the initial slope of the creep curve for T_x and the initial slope for the creep curve for T_y are identical with the ratio of the slopes given by Kalker's linear theory:

$$\frac{p_{0x}}{p_{0y}} = \frac{C_{11}\xi_x}{C_{22}\xi_y} \tag{2.96}$$

The resulting creep curves are shown in Figure 2.8–2.9 where they are compared with the creep curves due to CONTACT and with the creep curves due to SHE.

Again it is noticed that the modified strip theory gives a better approximation than SHE both when it comes to the total traction and for each traction component. It is further seen that the creep curve for the total traction is identical with the creep curve for the case of pure tangential traction (see Figure 2.8). Thus, the accuracy of the method does not decrease when a lateral traction component is introduced.

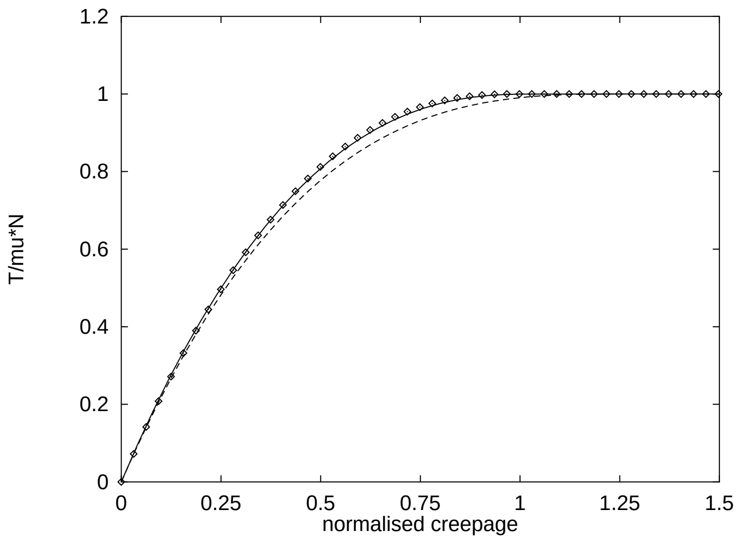

Figure 2.8. Total tangential force when $\xi_x = \xi_y$. Solid line: CONTACT. Dashed line: SHE. $\diamond$: modified strip theory.

The illustrated tangential problem is calculated for the case $\xi_x = \xi_y$ but the discrepancies are qualitatively the same for other combinations of ξ_x and ξ_y, so it must be concluded that for general no-spin contact situations the modified strip theory is a better approximation than SHE. A further advantage of the strip theory is that it calculates the stick zone which is not the case with SHE. It is obviously an approximation but comparisons with CONTACT indicate that there is a fairly good agreement between the very accurate calculated stick zone of CONTACT and the one found by the strip theory.

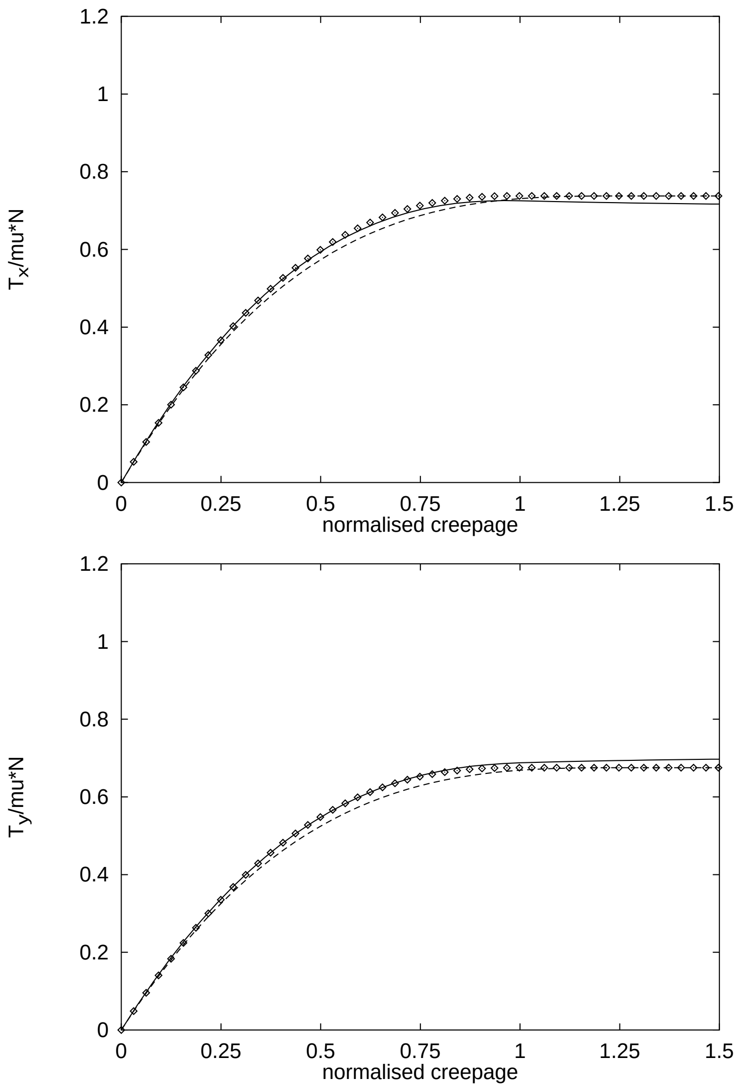

Figure 2.9. Components of the tangential force when $\xi_x = \xi_y$.
Top: Longitudinal tangential force. Bottom: Lateral tangential force.
Solid line: CONTACT. Dashed line: SHE. $\diamond$: modified strip theory.

2.6.3 The Inclusion of Spin

The main problem when deriving a three-dimensional contact theory based on the two-dimensional solution arises when the spin must be included. Where the longitudinal and the lateral creepages are very much comparable with the creepage from the two-dimensional problem, the spin leads to a qualitatively different contact situation. When the spin is included in the contact problem the direction of the relative velocity varies inside the contact patch. As the direction of the tangential stress in the slip zone is known to be opposite the direction of the slip this will lead to very complicated stress distributions.

From the kinematic constraints it is seen, that a pseudo creepage vector can be defined as $\{\xi_x - y\varphi, \xi_y + x\varphi\}^T$. The spin pole is then defined as the position where the pseudo creepage is zero:

$$(x, y) = \left(-\frac{\xi_y}{\varphi}, \frac{\xi_x}{\varphi}\right) \tag{2.97}$$

Now introduce the spin coefficient Φ as

$$\Phi = \frac{\sqrt{a_0^2 + b_0^2}}{\sqrt{\left(\frac{\xi_y}{\varphi}\right)^2 + \left(\frac{\xi_x}{\varphi}\right)^2}} \tag{2.98}$$

then the spin pole is located inside the contact patch if $\Phi > 1$. In Figure 2.10 the global relative velocity inside the contact patch is shown for the cases where (A): $\Phi = 0.5$ and (B): $\Phi = 2$. It is seen how the direction of the relative velocity changes inside the the contact patch when Φ is large. For that reason the strip theory is only applicable when $\Phi \ll 1$ where the direction of the relative velocity does not change considerably within the contact patch.

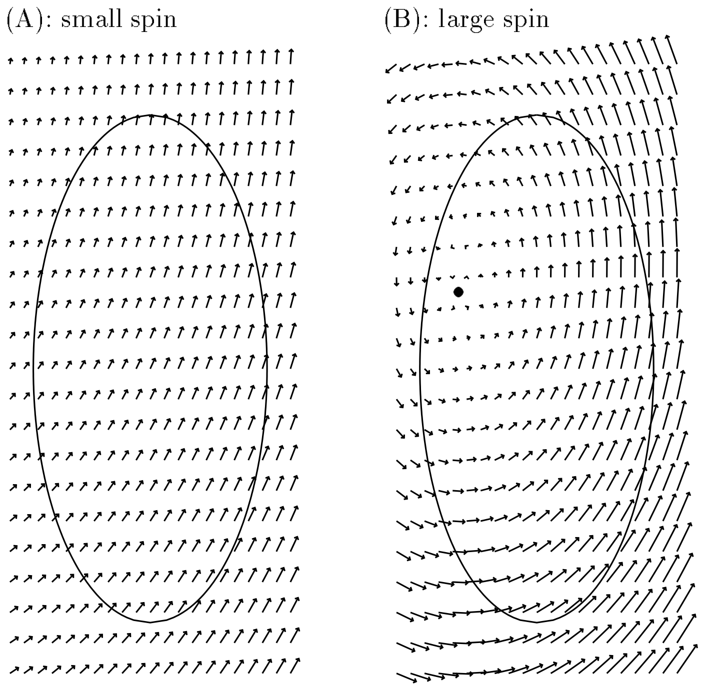

Figure 2.10. Global relative velocity in the contact patch. (A): small spin ($\Phi = 0.5$). (B): large spin ($\Phi = 2$. •: spin pole).

It is possible to compensate for relative small spins by using Kalker's linear theory. This can be done by making the transformation

$$C_{22}\xi_y \quad \rightarrow \quad C_{22}\xi_y + \sqrt{a_0 b_0} C_{23}\varphi \tag{2.99}$$

which ensures that the initial slope of the creep curve agrees with Kalker's linear theory also for cases where the spin is included. It must however be emphasized that this modification only holds when the spin is small and it does not imply a decaying behaviour of the creep curve which Kalker has shown is the effect of large spin [38].

The above derivation of a three-dimensional theory indicates that the qualitatively behaviour of the three-dimensional contact problem is closely related to an equivalent two-dimensional contact problem, when no spin occurs. This fact is very useful as two-dimensional contact problems thus provide a good understanding of the behaviour of three-dimensional problems. For this reason the remaining chapters are devoted to two-dimensional contact theory.

Chapter 3

A Polynomial Approach

In the present chapter a new approach to the two-dimensional contact problem is introduced. The objective of the method is to transform the constitutive equation which is defined as an integral equation into an algebraic equation. By this simplification it is possible to solve a variety of contact problems which do not fulfill the assumptions of the Hertz theory and of the Carter theory. The fundamental principles of the method will be explained and some simple examples of the application of the method are demonstrated.

3.1 An Extended Two-Dimensional Solution

The Hertz solution and the Carter solution are the classic approaches to the two-dimensional contact problem. They have many advantages but are only valid under certain assumptions:

1. the bodies in contact can be approximated with second order polynomials in the vicinity of the contact point.
2. the friction law of Coulomb is valid.
3. the bodies in contact are smooth.
4. the contact is stationary.

Now the question arises: how can a two-dimensional contact problem be solved if those restrictions are not fulfilled? One way to overcome this problem is to employ a boundary element method where the contact patch is divided into a finite number of surface elements. The program VISCON developed by Wang [68] is an examples of a boundary element method for two-dimensional contact. It is, however, very difficult to gain an understanding of the nature of a contact problem by using numerical methods. The influence of the various parameters can only be investigated by calculating numerous examples and then make some conclusions regarding the behaviour of the problem with respect to the different parameters. This is a time consuming approach, which may imply that certain properties of the problem are disregarded.

Instead of treating the contact problem numerically it is sought to derive an analytical method to solve the contact problem. The advantage of analytical methods is that the solution is fast and exact. Furthermore an analytical solution facilitates the investigation of how the various parameters influence the contact problem.

3.2 The Basic Idea

The main obstacle in solving a contact problem based on elastic half spaces is the constitutive equation. If the stresses are given the displacements can be calculated quite easily, but it is very difficult to derive an expression for the stresses related to a predefined relative displacement. So the aim of the polynomial approach is to transform this integral equation into an algebraic equation for an arbitrary choice of stress distribution or displacement gradient. When this is done both the normal contact problem and the tangential contact problem can be expressed as a system of algebraic equations and are thus easily solved.

3.2.1 Integral Transformation

The relation between the stress and the displacement gradient is given by the constitutive equation, which is formulated as the integral equation

$$\frac{du_z(x)}{dx} = -\frac{4(1-\nu^2)}{\pi E} \int_{-a}^{a} \frac{p(\zeta)}{(x-\zeta)}\, d\zeta \tag{3.1}$$

Now assume that the stress distribution can be approximated by a polynomial form:

$$p(x) = \frac{\sum_{n=0}^{N} B_n x^n}{\sqrt{a^2 - x^2}} \tag{3.2}$$

where N in the following is referred to as the degree of the polynomial form. As the size of the contact patch is finite and the order of the polynomial can be chosen to an arbitrary high degree, this is a quite reasonable assumption. The scope of the present section is to demonstrate that if the stress distribution can be expressed as a sum of polynomial forms like the one in equation (3.2) then the displacement gradient in the contact patch will be a sum of polynomials.

Introduce the polynomial

$$\sum_{n=0}^{N} B_n x^n$$

where the boundary conditions

$$\sum_{n=0}^{N} B_n(-a)^n = \sum_{n=0}^{N} B_n(a)^n = 0 \tag{3.3}$$

are fulfilled. It can then be shown (see Appendix D) that the following relation exists:

$$\int_{-a}^{a} \frac{\sum_{n=0}^{N} B_n \zeta^n}{(x-\zeta)\sqrt{a^2-\zeta^2}} d\zeta = \sum_{m=0}^{N-1} \beta_m x^m + I_{-1} \sum_{n=0}^{N} B_n x^n \tag{3.4}$$

$$I_{-1} = \begin{cases} 0 & , \quad |x| \leq a \\ \frac{\text{sign}(x)\pi}{\sqrt{x^2-a^2}} & , \quad |x| > a \end{cases} \tag{3.5}$$

The connection between the B_n's and the β_m's is described by the matrix equation

$$\{\beta_0\} = [A]\{B_1\} \tag{3.6}$$

$$\{\beta_0\}^T = \{\beta_0, \beta_1, \beta_2, \ldots, \beta_{N-1}\} \tag{3.7}$$

$$\{B_1\}^T = \{B_1, B_2, B_3, \ldots, B_N\} \tag{3.8}$$

$$[A] = \begin{bmatrix} A_0 & 0 & A_1 & 0 & A_2 & \dots & 0 \\ 0 & A_0 & 0 & A_1 & 0 & \dots & A_{N/2-1} \\ 0 & 0 & A_0 & 0 & A_1 & \dots & 0 \\ 0 & 0 & 0 & A_0 & 0 & \dots & A_{N/2-2} \\ 0 & 0 & 0 & 0 & A_0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & 0 & \dots & A_0 \end{bmatrix} \tag{3.9}$$

where the coefficients A_k are defined as

$$A_k = -\pi \frac{(2k)!}{k!} \left(\frac{a}{2}\right)^{2k} \quad , \quad k = 0, 1, 2, \dots \tag{3.10}$$

Because the matrix $[A]$ always is nonsingular there exists a unique solution to the inverse problem

$$\{B_1\} = [A]^{-1}\{\beta_0\} \tag{3.11}$$

i.e. that if the β_n's are given, then the B_n's can be determined from equation (3.11). The matrix $[A]^{-1}$ is

$$[A]^{-1} = \begin{bmatrix} A_0^{-1} & 0 & A_1^{-1} & 0 & A_2^{-1} & \dots & 0 \\ 0 & A_0^{-1} & 0 & A_1^{-1} & 0 & \dots & A_{N/2-1}^{-1} \\ 0 & 0 & A_0^{-1} & 0 & A_1^{-1} & \dots & 0 \\ 0 & 0 & 0 & A_0^{-1} & 0 & \dots & A_{N/2-2}^{-1} \\ 0 & 0 & 0 & 0 & A_0^{-1} & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & 0 & \dots & A_0^{-1} \end{bmatrix} \tag{3.12}$$

where the coefficients A_k^{-1} are defined as

$$A_k^{-1} = \frac{1}{\pi} \frac{(2k)!}{(2k-1)k!} \left(\frac{a}{2}\right)^{2k} \quad , \quad k = 0, 1, 2, \dots \tag{3.13}$$

Now the question is whether the above solution is unique. It has been demonstrated that if the stress distribution has the polynomial form as in equation (3.2) then the displacement gradient always will be a polynomial. But is the inverse problem also unique: will a polynomial displacement gradient always imply a stress distribution on a polynomial form?

As $\beta_{N-1} = -\pi B_N$ it will always be true that if the stress distribution is a polynomial form of a finite degree, then the displacement gradient will be a polynomial of one degree lower than the stress distribution. Now let the displacement gradient be a polynomial of the finite degree $(N-1)$ and assume that the stress distribution is

$$p(x) = f(x) \sum_{m=0}^{M} B_m x^m \tag{3.14}$$

It is already known that if $f(x) = 1/\sqrt{a^2 - x^2}$ and $M = N$ then the stress distribution (3.2) is a solution to the inverse problem. If there exists another solution then it can be written as

$$p(x) = \frac{g(x) \sum_{m=0}^{M} B_m x^m}{\sqrt{a^2 - x^2}} \tag{3.15}$$

Now assume that $g(x)$ is infinitely smooth for $-a < x < a$ and thus can be expressed as a Taylor expansion of order M_g. This implies that $(M + M_g) = (N - 1)$ i.e. that M_g is finite, which ensures that if the displacement gradient is a polynomial of finite degree, then the stress distribution is also a polynomial form of a finite degree. This is exactly the class of stress distributions for which the equations (3.6)-(3.13) are valid and for which a unique solution to the inverse problem exists. With this argumentation the basic idea of the polynomial approach is demonstrated:

$$p(x) = \frac{\sum_{n=0}^{N} B_n x^n}{\sqrt{a^2 - x^2}} \quad \Leftrightarrow \quad \frac{du_z(x)}{dx} = \sum_{m=0}^{N-1} \beta_m x^m \quad , \quad |x| \leq a \tag{3.16}$$

The polynomial approach is based on comparison of polynomial coefficients. As the stress distributions or the displacement gradients can be defined in several coordinate systems it is important to be able to change the base of a certain polynomial. This is done by applying the binomial formulae which states that if

$$\sum_{n=0}^{N} G_n (x+d)^n = \sum_{n=0}^{N} \gamma_n x^n \tag{3.17}$$

then there is a linear relation between the G_n's and the γ_n's

$$\{\gamma_0\} = [D]\{G_0\} \tag{3.18}$$

where

$$\{\gamma_0\}^T = \{\gamma_0, \gamma_1, \gamma_2, \ldots, \gamma_N\} \tag{3.19}$$

$$\{G_0\}^T = \{G_0, G_1, G_2, \ldots, G_N\} \tag{3.20}$$

and where a matrix element in $[D]$ at the position (i, j) is found as

$$D_{ij} = \begin{cases} \begin{pmatrix} j \\ i \end{pmatrix} d^{j-i} & , \quad j \geq i \\ \\ 0 & , \quad j < i \end{cases} \tag{3.21}$$

The matrix $[D]$ is nonsingular for all values of d and can thus be inverted. The elements in $[D]^{-1}$ are

$$D_{ij}^{-1} = (-1)^{j-i} D_{ij} \tag{3.22}$$

The last necessary integral transformation is related to the integral of the stress distribution i.e. the evaluation of the force. Let the force be equal the stress distribution integrated over the entire contact patch:

$$N_{force} = \int_{-a}^{a} \frac{\sum_{n=0}^{N} B_n \zeta^n}{\sqrt{a^2 - \zeta^2}} \, d\zeta \tag{3.23}$$

then the force is found as

$$N_{force} = -\{A_0\}^T \{B_0\} \tag{3.24}$$

$$\{A_0\}^T = \{A_0, 0, A_1, 0, \ldots, A_{N/2-1}, 0\} \tag{3.25}$$

$$\{B_0\}^T = \{B_0, B_1, B_2, B_3, \ldots, B_{N-1}, B_N\} \tag{3.26}$$

With the above derived transformations of integrals into algebraic equations it is possible to turn the entire contact problem into a set of algebraic equations, which facilitates the derivation of a solution as it will be demonstrated in the next section.

3.3 Application of the Theory

The advantage of the polynomial approach is that the displacement gradients are polynomials with known coefficients. By comparing the coefficients a few complicated polynomial equations are split into many very simple equations which are easily solved. In the next sections two examples of the application of the polynomial approach are briefly described: one for a normal contact problem and one for a tangential contact problem.

3.3.1 Two-Dimensional Non-Hertzian Contact

The Hertz solution to the normal contact problem is only valid if the bodies in contact can be approximated by second order polynomials. Now assume that it is necessary to employ a higher order approximation of the separation

$h = Z(X)$ in order to achieve a satisfying accuracy:

$$Z(X) = \sum_{n=0}^{N} Z_n X^n \tag{3.27}$$

This expression is inserted into the constitutive equation:

$$\frac{d}{dX}\left[\sum_{n=0}^{N} Z_n X^n\right] = \frac{4(1-\nu^2)}{\pi E}\int_{-a}^{a}\frac{p(\zeta)}{x-\zeta}\,d\zeta \tag{3.28}$$

According to the polynomial approach this implies that the normal pressure distribution will be a polynomial form of the degree N. Thus the contact problem contains $(N+3)$ unknowns: the $(N+1)$ coefficients from the polynomial form, the size of half the contact patch, a, plus the position of the contact patch in the global coordinate system - the centre of the contact patch is not necessarily located at the origo of the coordinate system in which $Z(X)$ is defined.

The separation between the bodies – $Z(x)$ – defined in the local coordinate system is calculated with the aid of the binomial matrix (equation (3.18)) where d is the position of the centre of the contact patch defined in the global coordinate system. Thus, $Z(x)$ is given as the polynomial

$$Z(x) = \sum_{n=0}^{N} \bar{Z}_n x^n \tag{3.29}$$

where $x = X - d$. The unknowns a and d plus the coefficients of the polynomial form are now found from the matrix equation (3.11) where $\beta_n = (n+1)\bar{Z}_{n+1}$, from the boundary conditions (equation (3.3)) and from the relation for the normal force (equation (3.24)). As the coefficients for the polynomial form are represented linearly in the equation system they can be substituted directly. This implies that the entire normal contact problem is reduced to solving two nonlinear equations with the unknowns a and d.

For some applications it is possible to derive analytical expressions for a and d whereas it in other cases is necessary to apply a numerical method. With the values from the Hertzian solution as initial guess, an iterative method will converge in very few steps. Compared with boundary element methods where an iterative process must be applied within each contact element the present approach is of course much faster. A further advantage is that the problem is solved without using variational methods which are very time consuming. The existence of many different values of a and d as solutions to the nonlinear equations is of course a problem, but provided the contact patch is coherent the iterative process will converge quickly towards the real a- and d-values. The other solutions to the equations yield non-physical stress distributions where the normal pressure in some areas is negative.

3.3.2 The Tangential Problem for a Two-Dimensional Non-Hertzian Contact

The Carter solution to the tangential contact problem is based on a Hertzian normal pressure distribution. If the normal pressure distribution is non-Hertzian the Carter solution is no longer valid and the polynomial approach must then be applied. Let the normal pressure distribution be

$$p(x) = \frac{\pi E}{4(1-\nu^2)} \frac{\sum_{n=0}^{N} B_n x^n}{\sqrt{a^2 - x^2}} \tag{3.30}$$

and assume the tangential stress distribution to be the sum of two polynomial forms

$$q(x) = q_1(x) + q_2(x^*) \tag{3.31}$$

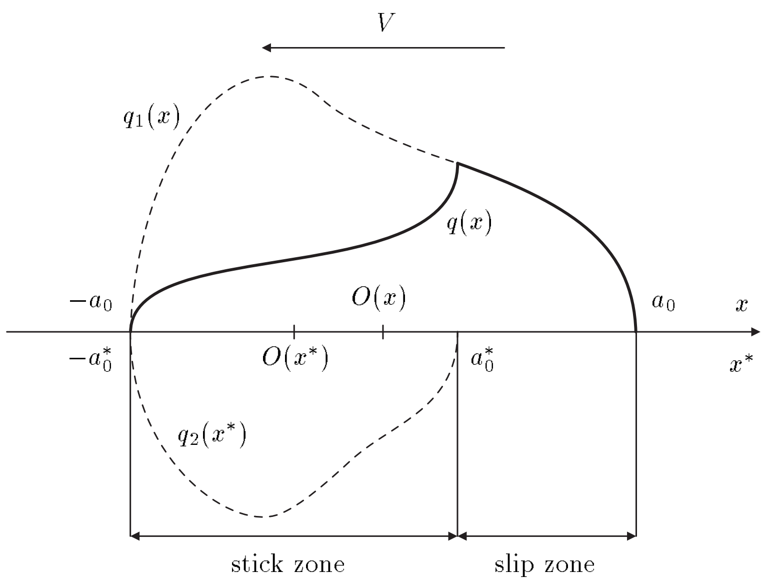

Figure 3.1. The tangential stress distribution according to the polynomial approach.

$$q_1(x) = \begin{cases} \frac{\mu\pi E}{4(1-\nu^2)} \frac{\sum_{n=0}^{N} B_n x^n}{\sqrt{a^2 - x^2}} & , \quad -a < x < a \\ 0 & , \quad \text{otherwise} \end{cases} \tag{3.32}$$

$$q_2(x^*) = \begin{cases} \frac{\mu\pi E}{4(1-\nu^2)} \frac{\sum_{m=0}^{M} B_m x^{*m}}{\sqrt{a^{*2} - x^{*2}}} & , \quad -a^* < x^* < a^* \\ 0 & , \quad \text{otherwise} \end{cases} \tag{3.33}$$

This is equivalent to the Carter solution but with the ellipses replaced by polynomial forms (see Figure 3.1). Thus, the contact patch is still assumed to be divided into a stick zone and a slip zone where the stick zone is located at the leading edge of the contact patch. The displacement gradients arising from the two contributions to the tangential stress distribution are found from the polynomial approach as:

$$\frac{du_1(x)}{dx} = -\{x^0\}^T[A]\{B_1\} \ , \ \ -a < x < a \tag{3.34}$$

$$\frac{du_2(x^*)}{dx^*} = -\{x^{*0}\}^T[A^*]\{B_1^*\} \ , \ \ -a^* < x^* < a^* \tag{3.35}$$

with the notation

$$\{x^0\}^T = \{1, x, x^2, \ldots, x^{N-1}\} \tag{3.36}$$

$$\{x^{*0}\}^T = \{1, x^*, x^{*2}, \ldots, x^{*M-1}\} \tag{3.37}$$

and where the elements in the matrix $[A^*]$ are equivalent to those from the matrix $[A]$ (see equation (3.9)) just with a replaced by a^*. With this inserted into the kinematic constraint where $s(x) = 0$ as $x \in S_{stick}$ a relation between the polynomial forms and the creepage is established

$$\xi_0 = \mu\left(\{x^0\}^T[A]\{B_1\} + \{x^{*0}\}^T[A^*]\{B_1^*\}\right) \tag{3.38}$$

Changing the base of $q_2(x^*)$ this equation is rewritten as

$$\xi_0 = \mu\left(\{x^0\}^T[A]\{B_1\} + \{x^0\}^T[D][A^*]\{B_1^*\}\right) \tag{3.39}$$

where $d = a - a^*$ in the matrix $[D]$ (see equation (3.18)). It follows immediately that the degree of the two polynomials are the same i.e. $N = M$. As the creepage is independent of the position x a new matrix equation is derived

$$0 = [A_s]\{B_2\} + [D_s][A_s^*]\{B_2^*\} \ \ \Rightarrow \tag{3.40}$$

$$\{B_2^*\} = -[A_s^*]^{-1}[D_s]^{-1}[A_s]\{B_2\} \tag{3.41}$$

where the indices s denotes that the matrices are sub-matrices where the first row and the first column are removed. Provided the normal contact problem is already solved and thus a and the B_n's are known, the tangential contact problem consists of $(N + 2)$ unknowns: the $(N + 1)$ polynomial coefficients B_n^* plus the size of the stick zone a^*. The matrix equation (3.41) yields $(N - 1)$ equations and the boundary conditions

$$\sum_{n=0}^{N} B_n^*(-a^{*n}) = \sum_{n=0}^{N} B_n^* a^{*n} = 0 \tag{3.42}$$

provides another two equations. The last equation necessary to solve the contact problem is derived from equation (3.39) and reads

$$\xi_0 = \mu \left(\{A_0\}^T \{B_1\} + \{I\}^T [D][A^*]\{B_1^*\} \right) \tag{3.43}$$

where $\{I\}$ is the unity vector $\{I\}^T = \{1, 0, 0, \ldots, 0\}$. Equivalent to the normal contact problem the polynomial coefficients can be substituted directly so that one nonlinear equation with the unknown a^* remains. The solution to this equation can sometimes be found analytically or else with the aid of an iterative process which converges in very few steps if the Carter value a_0^* is utilized as initial guess. When a^* is found the tangential force is equal to the sum of the contributions from $q_1(x)$ and $q_2(x^*)$:

$$T = -\frac{\mu \pi E}{4(1-\nu^2)} \left(\{A_0\}^T \{B_0\} + \{A_0^*\}^T \{B_0^*\} \right) \tag{3.44}$$

The distribution of the slip is found by inserting the displacement gradients into the kinematic constraint:

$$s(x) = \begin{cases} \pi\mu \left(\frac{\sum_{n=0}^{N} B_n x^n}{\sqrt{x^2 - a^2}} + \frac{\sum_{n=0}^{N} B_n^* x^{*n}}{\sqrt{x^{*2} - a^{*2}}} \right) & , \quad -a \leq x \\ 0 & , \quad -a < x < 2a^* - a \\ -\pi\mu \frac{\sum_{n=0}^{N} B_n^* x^{*n}}{\sqrt{x^{*2} - a^{*2}}} & , \quad 2a^* - a < x < a \\ -\pi\mu \left(\frac{\sum_{n=0}^{N} B_n x^n}{\sqrt{x^2 - a^2}} + \frac{\sum_{n=0}^{N} B_n^* x^{*n}}{\sqrt{x^{*2} - a^{*2}}} \right) & , \quad a \geq x \end{cases} \tag{3.45}$$

The above solution to the tangential contact problem is only valid if the contact patch is coherent and divided into a stick zone at the leading edge and a slip zone at the trailing edge. If the contact patch is divided into several stick zones then the term

$$\frac{\operatorname{sign}(x^*)\pi}{\sqrt{x^{*2} - a^{*2}}} \tag{3.46}$$

in equation (3.4) must be taken into account, because it will influence the displacement gradient in the adjacent stick zone. This implies that the displacement gradients no longer can be evaluated as polynomials of a finite degree. As the polynomial approach is based on comparison of polynomial coefficients this will cause the model to break down. It is however possible to make a Taylor approximation of the term (3.46) and then utilize the polynomial approach as will be done in Chapter 6.

Some contact theories assume that the stick zone covers the entire contact patch i.e. $a^* = a$. With this assumption the number of unknowns in the system of equations is reduced by one and the problem is thus short of one degree of freedom. It is for that reason necessary to neglect on of the boundary conditions. Normally it is assumed that $q(-a) = 0$ and so the restriction

$$\sum_{n=0}^{N} B_n^* a^{*n} = 0 \tag{3.47}$$

must be ignored. This implies that a singularity occurs at the trailing edge at the contact patch i.e.

$$\lim_{x \to a} q(x) = \infty \tag{3.48}$$

which of course is a violation of the restriction that $|q(x)| \leq \mu p(x)$ unless the friction coefficient is infinite. For this reason contact theories where the stick zone covers the entire contact patch is usually referred to as contact with infinite friction.

Chapter 4

Corrugation

The Hertz solution and the Carter solution are not valid for heavily corrugated surfaces. The objective of the present chapter is to demonstrate how the polynomial approach can be utilized to investigate corrugation phenomena.

4.1 Introduction to Corrugation

A big problem in railway traffic is the corrugation of the rails. Corrugation appears as short-wave ripples across the surface of the rail which generate noise and cause discomfort for the passengers. The presence of corrugation also augment the dynamical load on the tracks which increases the wear of the rails and the running gear. A high dynamical load also implies a faster deterioration of the track. The corrugation evolves locally without any apparent reason and develops amplitudes of magnitudes up to 100 μm depending on the wave lengths.

It is a common theory that the corrugation can be generated by several different mechanisms. As the characteristic wave length apparently depends on the wear mechanism it is convenient to divide the rail corrugation into several classes depending on the wave length [2] or the wear mechanism [19]. In the latter work by Grassie and Kalousek the wave length fixing mechanisms are divided into six different groups. It is stated that the wear mechanisms are known for all the groups except the one denoted as short pitch corrugation. This type of corrugation has a typical wave length which lies within the range from $0.03 - 0.1$ m, and is recognized by characteristic shiny patches on the rail heads, where each patch indicates a trough in the corrugation pattern.

Because the wear mechanism which causes short pitch corrugation is unknown, the only way to treat the corrugation problem is to grind the rails, which is a very time and money consuming process. It is thus of great interest for the railway companies to understand the nature of the corrugation: how does it evolve and what determines the characteristic wave length and the growth rate of the corrugation? With the answers to these questions in hand it will be possible to take measures to prevent the corrugation and especially to develop a maintenance strategy that minimizes the costs related to the grinding of the rails without worsening the general quality of the track surface.

As the corrugation evolves over thousands of train passages it is very difficult to make experiments in order to investigate the evolution of the corrugation.

However it is possible to analyse the existing corrugation on a given rail and then from the knowledge of the traffic on the site draw some general conclusions. Such field experiments only make sense if the traffic on the rails is homogeneous, i.e. on closed systems such as metro lines [1], because a large diversity in the rolling stock complicates the establishing of a relation between the wave length of the corrugation and the characteristics of the traffic on the site.

A much more powerful approach is to employ numerical simulations in the pursuit of achieving an understanding of the corrugation mechanism. Much theoretical work has been carried out over the years in order to explain the phenomenon of corrugation, e.g. [12], [14] and [62]. An overview of the different approaches are listed in a state of the art review by Knothe and Grassie [41].

Most of the theoretical models focus on the dynamics of wheel and rail to explain and describe the formation of corrugation, but make more or less primitive approximations of the contact mechanics. Surely the dynamics of the rail and wheel must be taken into account in a complete model, but in the present chapter it will be shown that a more accurate model of the contact mechanics itself influences the corrugation heavily and must be included in the ordinary simulation programs to yield more realistic results.

4.2 An Infinite Cylinder Rolling on a Corrugated Surface

The corrugation model derived in the present chapter is a simplified model only focusing on the contact mechanics of wheel/rail interaction. The wheel and rail are thus described as bodies with no mass and thus with no eigenmodes or eigenfrequencies. The system is furthermore considered to be two-dimensional, and so the problem of a wheel rolling on a corrugated rail is transformed into the case of an infinite cylinder rolling on a corrugated surface. These simplifications will evidently give rise to some discrepancies between the results of the simulations and what is observed for three-dimensional contact situations, but as demonstrated in section 2.6 there are many similarities between a three-dimensional contact problem without spin and a two-dimensional problem which ensures that the behaviour of the two-dimensional model in those cases is qualitatively equivalent to the behaviour of the three-dimensional model.

4.2.1 The Physical Model

To investigate the influence of pure contact mechanics on the evolution of corrugation, the two-dimensional case of a cylinder rolling on a surface is examined (see Figure 4.1), where it is assumed that the level of the surface at any time can be described by a series of harmonic functions

$$Z_2(X) = Z_0 + \sum_{m=1}^{M} \left[Z_{A,m} \cos(k_m X) + Z_{B,m} \sin(k_m X) \right] \tag{4.1}$$

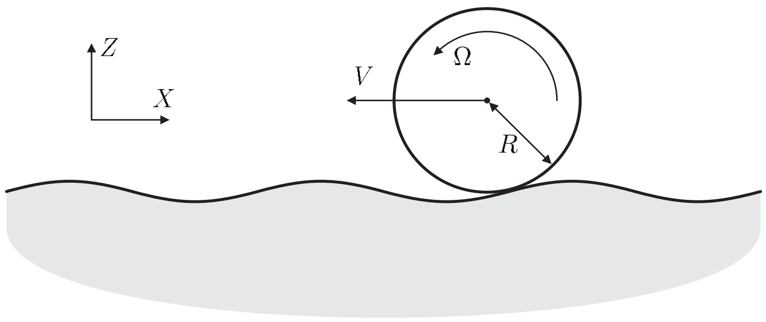

Figure 4.1. An infinite cylinder rolling on a corrugated surface.

To simplify the model, it is assumed that the entire wear is laid upon the surface, i.e. the cross section of the cylinder always will be circular with the constant radius R. This presumption is reasonable as the model is intended to simulate many wheels running on the same surface.

If the local curvature of the surface is larger than the curvature of the cylinder a two point contact occurs when the cylinder is located in a trough of the corrugation pattern. Consequently a sudden shift in the location of the contact point will occur and the rolling motion is replaced by impacts between the cylinder and the surface. This is however only the case when the surface is very heavily corrugated and thus of minor interest for the investigation of the formation of corrugation. Cases where impacts between the bodies are taken into account can be found in [32]. Thus it is assumed that the curvature of the cylinder always is much larger than the curvature of the surface i.e.

$$k_m^2 R \sqrt{Z_{A,m}^2 + Z_{B,m}^2} \ll 1 \quad , \quad m = 1, 2, \ldots, M \tag{4.2}$$

The above restriction also ensures that multiple contact is avoided which is a further simplification of the contact problem. Much work has been carried out to treat the case of multiple contact points, e.g. [56] and [57], where the configuration of an equivalent contact patch is derived from the actual contact patches. These approaches may be applicable for general simulations of the dynamics of a body rolling on a corrugated surface, but they are not well-suited for investigations of the contact mechanics. The following sections will demonstrate that the behaviour of the problem is very strongly related to what happens inside the contact patch, which implies that heuristic approaches provide poor accuracy and may even cause important properties of the contact problem to be disregarded.

4.2.2 The Normal Contact Problem

Assume that the cylinder rolls along the surface with the constant velocity V and with the constant angular velocity Ω and choose the origo of the global coordinate system such that the vertical projection of the cylinder axis has the position $X = 0$ to the time $t = 0$. Now define the shape of the two bodies in a moving coordinate system with origo at the vertical projection of the cylinder axis on Z_2:

$$Z_1(\bar{x}) = R - \sqrt{R^2 - \bar{x}^2} \tag{4.3}$$

$$Z_2(\bar{x}) = \sum_{m=1}^{M} \left[Z_{A,m} \cos\left[k_m(Vt + \bar{x})\right] + Z_{B,m} \sin\left[k_m(Vt + \bar{x})\right]\right] \tag{4.4}$$

Introduce the wave length L_m and the local angular velocity ω_m as

$$L_m = \frac{2\pi}{k_m} \tag{4.5}$$

$$\omega_m = V k_m \tag{4.6}$$

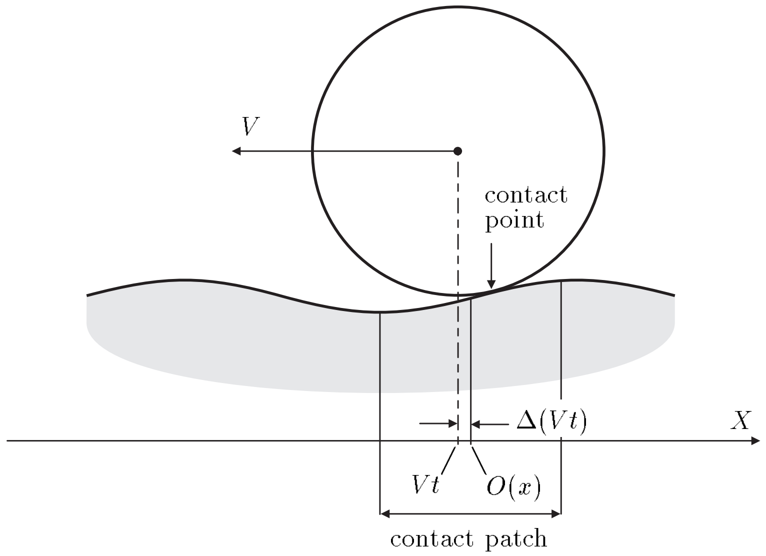

Figure 4.2. Local coordinate systems.

and let the parameter Δ be the longitudinal shift in the location of the centre of the contact patch i.e.

$$\bar{x} \in S \quad \Leftrightarrow \quad -a < \bar{x} - \Delta < a \tag{4.7}$$

where it is assumed that $|\Delta| \ll a$. The definition of Δ provides that if $\Delta > 0$ then the centre of the contact patch is located behind the cylinder axis with respect to the rolling direction. The magnitude of Δ depends on the shape of the bodies in the vicinity of the contact point and thus on the position of the cylinder on the surface.

Now define a new coordinate system where $x = \bar{x} - \Delta$ (see Figure 4.2). With this notation the bodies in contact are described in a coordinate system with origo in the centre of the contact patch

$$Z_1(x) = R - \sqrt{R^2 - (x+\Delta)^2} \tag{4.8}$$

$$Z_2(x) = \sum_{m=1}^{M} \left[Z_{A,m} \cos\left(\omega_m t + k_m x + k_m \Delta\right) + Z_{B,m} \sin\left(\omega_m t + k_m x + k_m \Delta\right) \right] \tag{4.9}$$

Making a Taylor expansion of the shape of the bodies, the separation $Z(x) = Z_1(x) - Z_2(x)$ can be written as

$$\begin{aligned} Z(x) = R &\left(\frac{1}{2}\left(\frac{x+\Delta}{R}\right)^2 + \frac{1}{8}\left(\frac{x+\Delta}{R}\right)^4 + \dots \right) - \\ &\sum_{m=1}^{M} \Bigg[Z_{A,m} \cos(\omega_m t + k_m \Delta) \left(1 - \frac{(k_m x)^2}{2!} + \frac{(k_m x)^4}{4!} - \dots \right) + \\ &\quad Z_{A,m} \sin(\omega_m t + k_m \Delta) \left(-k_m x + \frac{(k_m x)^3}{3!} - \dots \right) + \\ &\quad Z_{B,m} \sin(\omega_m t + k_m \Delta) \left(1 - \frac{(k_m x)^2}{2!} + \frac{(k_m x)^4}{4!} - \dots \right) + \\ &\quad Z_{B,m} \cos(\omega_m t + k_m \Delta) \left(k_m x - \frac{(k_m x)^3}{3!} + \dots \right) \Bigg] \end{aligned} \tag{4.10}$$

As in the Hertzian case it is presupposed that the size of the contact patch is much smaller than the radius of the cylinder, which justifies a second order Taylor approximation of $Z_1(x)$.

It is obvious that $k_m x$ is a critical quantity for the contact problem. Because x has the same magnitude as a the contact situation is described with the characteristic parameter $k_m a$ which is equal to $2\pi a/L_m$. The ratio between the size of the contact patch and the wave length of the corrugation is thus very important for the contact problem. When $2\pi a/L_m \to 0$ the contact problem can be approximated with the Hertzian solution whereas larger ratios implies that more terms in the polynomial approximation of the surface must be included in order to obtain a satisfactory accuracy.

In the present investigation the complete Taylor expansion is used in the calculations. At first this is strictly speaking impossible as the degree of the expansion is infinite, but it turns out that the infinite series eventually can be substituted by known functions. It is noticed that the application of the infinite series does not provide any problems concerning the magnitude of the coefficients as they are of the size $(k_m x)^n/n!$ which tends towards zero as n tends towards infinity.

By differentiating $Z(x)$ with respect to x and inserting the result into the constitutive equation this yields that

$$\frac{dZ(x)}{dx} = \sum_{n=0}^{\infty} \beta_n x^n = \frac{4(1-\nu^2)}{\pi E} \int_{-a}^{a} \frac{p(\zeta)}{x-\zeta}\, d\zeta \tag{4.11}$$

where the coefficients β_n are

$$\beta_0 = \frac{\Delta}{R} + \sum_{m=1}^{M} k_m \left[Z_{A,m} \sin(\omega_m t + k_m \Delta) - \right.$$
$$\left. Z_{B,m} \cos(\omega_m t + k_m \Delta) \right] \tag{4.12}$$

$$\beta_1 = \frac{1}{R} + \sum_{m=1}^{M} k_m^2 \left[Z_{A,m} \cos(\omega_m t + k_m \Delta) + \right.$$
$$\left. Z_{B,m} \sin(\omega_m t + k_m \Delta) \right] \tag{4.13}$$

$$\beta_2 = -\sum_{m=1}^{M} \frac{k_m^3}{2!} \left[Z_{A,m} \sin(\omega_m t + k_m \Delta) - \right.$$
$$\left. Z_{B,m} \cos(\omega_m t + k_m \Delta) \right] \tag{4.14}$$

$$\beta_3 = -\sum_{m=1}^{M} \frac{k_m^4}{3!} \left[Z_{A,m} \cos(\omega_m t + k_m \Delta) + \right.$$
$$\left. Z_{B,m} \sin(\omega_m t + k_m \Delta)\right] \quad (4.15)$$

$$\vdots$$

$$\beta_{2n} = (-1)^n \sum_{m=1}^{M} \frac{k_m^{2n+1}}{2n!} \left[Z_{A,m} \sin(\omega_m t + k_m \Delta) - \right.$$
$$\left. Z_{B,m} \cos(\omega_m t + k_m \Delta)\right] \quad (4.16)$$

$$\beta_{2n+1} = (-1)^n \sum_{m=1}^{M} \frac{k_m^{2n+2}}{(2n+1)!} \left[Z_{A,m} \cos(\omega_m t + k_m \Delta) + \right.$$
$$\left. Z_{B,m} \sin(\omega_m t + k_m \Delta)\right] \quad (4.17)$$

With the aid of the polynomial approach it is now possible to express the normal pressure distribution as the polynomial form

$$p(x) = \frac{\pi E}{4(1-\nu^2)} \frac{\sum_{n=0}^{\infty} B_n x^n}{\sqrt{a^2 - x^2}} \quad (4.18)$$

with the coefficients

$$B_1 = -\frac{\Delta}{\pi R} + \sum_{m=1}^{M} k_m \left[Z_{A,m} \sin(\omega_m t + k_m \Delta) - \right.$$
$$\left. Z_{B,m} \cos(\omega_m t + k_m \Delta)\right] \sum_{j=0}^{\infty} (-1)^j A_j^{-1} \frac{k_m^{2j}}{(2j)!} \quad (4.19)$$

$$B_2 = -\frac{1}{\pi R} + \sum_{m=1}^{M} k_m \left[Z_{A,m} \cos(\omega_m t + k_m \Delta) + \right.$$
$$\left. Z_{B,m} \sin(\omega_m t + k_m \Delta)\right] \sum_{j=0}^{\infty} (-1)^j A_j^{-1} \frac{k_m^{2j+1}}{(2j+1)!} \quad (4.20)$$

$$B_3 = -\sum_{m=1}^{M} k_m \left[Z_{A,m} \sin(\omega_m t + k_m \Delta) - \right.$$
$$\left. Z_{B,m} \cos(\omega_m t + k_m \Delta)\right] \sum_{j=0}^{\infty} (-1)^j A_j^{-1} \frac{k_m^{2j+2}}{(2j+2)!} \tag{4.21}$$
$$B_4 = -\sum_{m=1}^{M} k_m \left[Z_{A,m} \cos(\omega_m t + k_m \Delta) + \right.$$
$$\left. Z_{B,m} \sin(\omega_m t + k_m \Delta)\right] \sum_{j=0}^{\infty} (-1)^j A_j^{-1} \frac{k_m^{2j+3}}{(2j+3)!} \tag{4.22}$$
$$\vdots$$
$$B_{2n-1} = (-1)^{n+1} \sum_{m=1}^{M} k_m \left[Z_{A,m} \sin(\omega_m t + k_m \Delta) - \right.$$
$$\left. Z_{B,m} \cos(\omega_m t + k_m \Delta)\right] \sum_{j=0}^{\infty} (-1)^j A_j^{-1} \frac{k_m^{2j+2n-2}}{(2j+2n-2)!} \tag{4.23}$$
$$B_{2n} = (-1)^{n+1} \sum_{m=1}^{M} k_m \left[Z_{A,m} \cos(\omega_m t + k_m \Delta) + \right.$$
$$\left. Z_{B,m} \sin(\omega_m t + k_m \Delta)\right] \sum_{j=0}^{\infty} (-1)^j A_j^{-1} \frac{k_m^{2j+2n-1}}{(2j+2n-1)!} \tag{4.24}$$

The restriction that $p(-a) = p(a) = 0$ yields the expression

$$B_0 = -\sum_{n=1}^{\infty} B_{2n} a^{2n} \tag{4.25}$$

which can be calculated to

$$B_0 = \frac{a^2}{\pi R} - \left(\frac{a}{\pi}\right)^2 \sum_{m=1}^{M} k_m \left[Z_{A,m} \cos(\omega_m t + k_m \Delta) + \right.$$
$$\left. Z_{B,m} \sin(\omega_m t + k_m \Delta)\right] \sum_{j=0}^{\infty} (-1)^j A_j \frac{k_m^{2j+1}}{(2j+1)!} \tag{4.26}$$

and so the coefficients for the polynomial form of the normal pressure distribution are derived. It is noticed that they only depend on the two unknown quantities Δ and a. It is not possible to bring the coefficients on a closed form but the magnitude of the contribution from each polynomial coefficient is found to be

$$B_n a^{n-1} = \mathcal{O}\left(\frac{(k_m a)^n}{n!}\right) \tag{4.27}$$

Let $p_n(x)$ be defined as an approximation of the real normal pressure distribution $p(x)$ where the infinite series is truncated so it is a polynomial of degree n, then a measure of the relative error introduced by the truncation is defined as

$$\lim_{x \to a}\left[\frac{p(x) - p_n(x)}{p(x) - p_{n+1}(x)}\right] \simeq \frac{B_{n+1} a^n}{B_n a^{n-1}} = \frac{k_m a}{n+1} \tag{4.28}$$

A result of equation (4.28) is that a large a/L_m ratio demands a high degree of the Taylor expansion in order to reduce the error introduced by the truncation. If only a second order approximation is employed the solution is naturally equal to the Hertz solution. A rule of thumb says that the Hertzian theory can be applied if $a/L_m < 1/20$. With this ratio inserted into the error estimation from equation (4.28) the relative error is calculated to be about 10 %. Thus the polynomial approach provides a powerful tool to estimate the magnitude of the error introduced by applying a truncated solution to a given contact problem. It is further noticed that

$$\lim_{n \to \infty} B_n a^n = 0 \tag{4.29}$$

which ensures that the expression for the normal pressure distribution converges as n tends towards infinity.

To obtain a solution to the normal contact problem it only remains to derive some expressions for the unknowns a and Δ. The boundary conditions provide the expression

$$\sum_{n=0}^{\infty} B_{2n+1} a^{2n} = 0 \tag{4.30}$$

With the B_n's inserted, this equation can be rewritten as

$$0 = \Delta + \sum_{m=1}^{M} k_m R \mathrm{J}_0(k_m a) \left[Z_{A,m} \sin(\omega_m t + k_m \Delta) - Z_{B,m} \cos(\omega_m t + k_m \Delta) \right] \tag{4.31}$$

where J_0 is the Bessel function of the first kind of order 0. The second necessary equation is established from the expression for the normal force

$$N = \int_{-a}^{a} p(x)\, dx = -\frac{\pi E}{4(1-\nu^2)} \{A_0\}^T \{B_0\} \tag{4.32}$$

Inserting the polynomial coefficients in the matrix equation a simple expression for the normal force N is derived

$$N = \frac{\pi E}{4(1-\nu^2)} \left[\frac{a^2}{2R} + \sum_{m=1}^{M} k_m a \mathrm{J}_1(k_m a) \left[Z_{A,m} \cos(\omega_m t + k_m \Delta) + Z_{B,m} \sin(\omega_m t + k_m \Delta) \right] \right] \tag{4.33}$$

The two equations (4.31) and (4.33) can be solved numerically with the Hertzian values $a = a_0$ and $\Delta = 0$ as initial guess. A much more convenient way to solve the problem is to apply Taylor expansions of the two expressions about the Hertzian values. The latter approach leads to the closed form expressions

$$\Delta = -\sum_{m=1}^{M} k_m R \mathrm{J}_0(k_m a_0) \left[Z_{A,m} \sin(\omega_m t) - Z_{B,m} \cos(\omega_m t) \right] \tag{4.34}$$

$$a = a_0 - \sum_{m=1}^{M} k_m R \mathrm{J}_1(k_m a_0) \left[Z_{A,m} \cos(\omega_m t) + Z_{B,m} \sin(\omega_m t) \right] \tag{4.35}$$

which are excellent approximations as long as the amplitudes of a and Δ are small. Due to the restriction from equation (4.2) this will always be the case in the present model. Making the substitution $\omega_m t = k_m X$ the shift in the location of the centre of the contact patch and the size of half the contact patch are expressed in the global coordinate system as

$$\Delta(X) = -\sum_{m=1}^{M} k_m R\mathrm{J}_0(k_m a_0)\,[Z_{A,m}\sin(k_m X) - Z_{B,m}\cos(k_m X)] \tag{4.36}$$

$$a(X) = a_0 - \sum_{m=1}^{M} k_m R\mathrm{J}_1(k_m a_0)\,[Z_{A,m}\cos(k_m X) + Z_{B,m}\sin(k_m X)] \tag{4.37}$$

i.e. that when the cylinder axis is located at the position X then the values for Δ and a are given by the equations derived above. It is in the expression for $a(X)$ assumed that the normal force is constant. If the normal force oscillates and thus can be put on the form

$$N = N_0 + \sum_{m=1}^{M_N} [N_{A,m}\cos(k_{N,m}X) + N_{B,m}\sin(k_{N,m}X)] \tag{4.38}$$

then the expression for half the size of the contact patch is rewritten as

$$a = a_0 + \frac{a_0}{2N_0}\sum_{m=1}^{M_N} [N_{A,m}\cos(k_{N,m}X) + N_{B,m}\sin(k_{N,m}X)] - \sum_{m=1}^{M} k_m R\mathrm{J}_1(k_m a_0)\,[Z_{A,m}\cos(k_m X) + Z_{B,m}\sin(k_m X)] \tag{4.39}$$

This expression is like the others restricted to small amplitudes of a which also implies that the amplitude of the normal force is small. The expression for Δ remains unchanged as the oscillating normal force in this context is a secondary effect. In the following analysis the normal force is presupposed to be constant. The inclusion of an oscillating normal force can easily be made but it does not result in any changes in the qualitative behaviour of the contact problem.

By this the normal contact problem for a cylinder rolling on a corrugated surface is solved. As the characteristic parameters a and Δ are expressed on closed forms it is possible to carry out some basic analysis concerning the contact problem. It is important to notice, that the position of the contact point and the centre of the contact patch are not identical. Let Δ_c be the shift in the location of the contact point with respect to the axis of the cylinder. The contact point is then found as the position for which the tangents of the cylinder and the surface have the same slope, which can be done by solving the equation

$$\frac{\Delta_c}{\sqrt{R^2 - \Delta_c^2}} = -\sum_{m=1}^{M} k_m \left[Z_{A,m} \cos[k_m(Vt + \Delta_c)] - Z_{B,m} \sin[k_m(Vt + \Delta_c)] \right] \qquad (4.40)$$

with respect to Δ_c. Applying a Taylor expansion this leads to the solution

$$\Delta_c = -\sum_{m=1}^{M} Rk_m \left[Z_{A,m} \sin(k_m X) - Z_{B,m} \cos(k_m X) \right] \qquad (4.41)$$

which is not equivalent to Δ. Considering a corrugated surface with only one characteristic wave length, it is noticed that

$$\Delta = \mathrm{J}_0(a_0 k)\Delta_c \qquad (4.42)$$

The size of the difference between Δ and Δ_c is thus depending on the relative size of half the contact patch. This is quite an evident fact as Δ_c is a pure geometrical property, independent of the deformation in the contact patch, whereas Δ is related to both geometry and deformation. From the

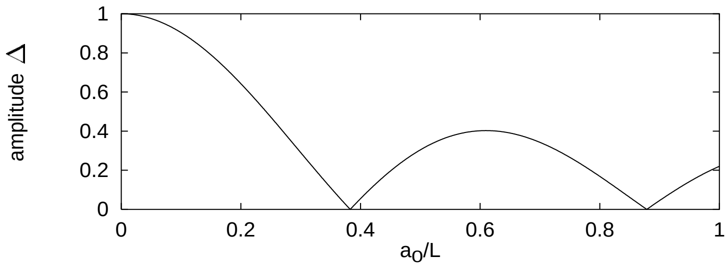

Figure 4.3. The amplitude of the location of the centre of the contact patch normalised with $kR\sqrt{Z_A^2 + Z_B^2}$.

expression (4.42) it is seen that the contact point can be located on one side of the cylinder axis whereas the centre of the contact patch is on the other side of the axis. This is due to the asymmetry of the surface: when the cylinder is on a spike or in a trough, the surface is symmetric around the axis which implies that the contact point and the centre of the contact patch are coincident. For all other positions of the cylinder - provided the surface is corrugated - the two points will not be identical. The amplitude of Δ normalised with $kR\sqrt{Z_A^2 + Z_B^2}$ is plotted in Figure 4.3.

If the normal force is constant then the size of the contact patch depends strongly on $a_0 k$ and thus on the ratio a_0/L. The amplitude of a – referred to as $\widehat{a}$ – normalised with $kR\sqrt{Z_A^2 + Z_B^2}$ is thus equal to the Bessel function of first kind order 1 which is shown in Figure 4.4. It is seen that for certain values of a_0/L the size of the contact patch will be constant even though the shape of the surface varies. Similarly a resonance effect will occur for the a_0/L values for which $\widehat{a}$ has a local maximum. The global maximum of the amplitude is obtained for $a_0/L \simeq 0.293$ and is found to be $\widehat{a} \simeq 0.582kR\sqrt{Z_A^2 + Z_B^2}$.

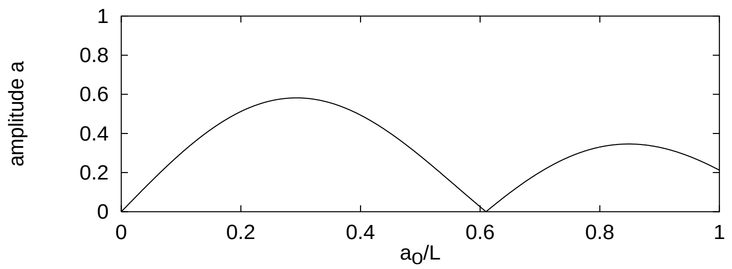

Figure 4.4. Amplitude of half the size of the contact patch normalised with $kR\sqrt{Z_A^2 + Z_B^2}$.

In Figure 4.5 the normal pressure distribution for different positions of the cylinder is shown for a corrugation with one distinct wave number k where $k^2 R\sqrt{Z_A^2 + Z_B^2} = 0.4$ and $a_0/L = 0.45$. The calculations are made for a constant normal force and are compared with the Hertzian solution for the cylinder rolling on a level surface. It is clearly seen that the normal pressure distribution due to the asymmetries at the vicinity of the contact point is asymmetric itself. This feature turns out to be extremely important for the solution to the tangential contact problem as demonstrated in the following section.

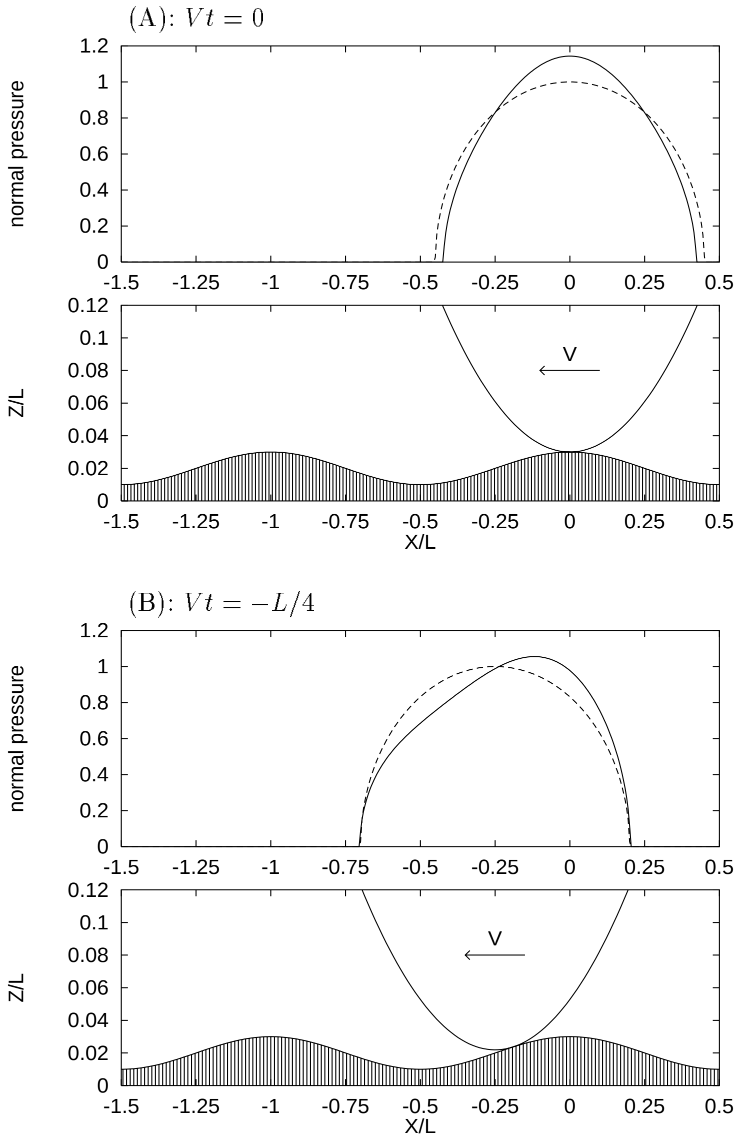

Figure 4.5. (A)-(B) Normal pressure distribution. Top: stress distribution calculated with the polynomial approach (solid line) and calculated with the Hertzian theory (dashed line). Bottom: position of the cylinder.

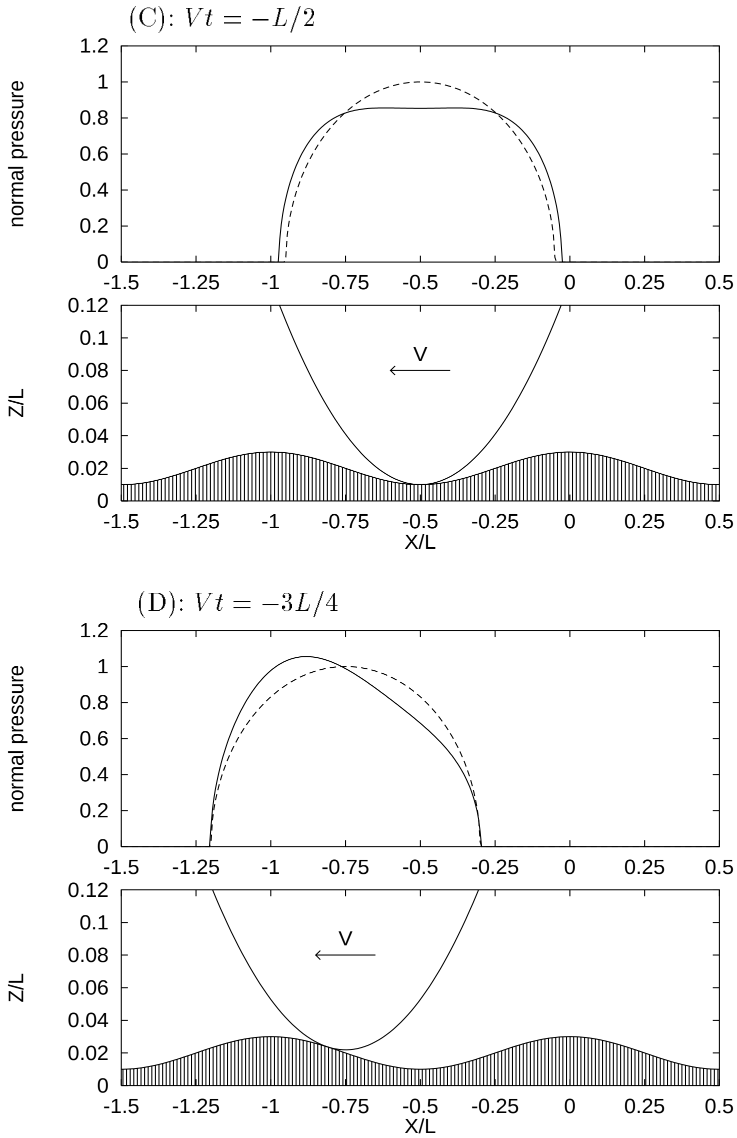

Figure 4.5. (C)-(D) Normal pressure distribution. Top: stress distribution calculated with the polynomial approach (solid line) and calculated with the Hertzian theory (dashed line). Bottom: position of the cylinder.

4.2.3 The Tangential Contact Problem

In the previous section the normal contact problem was solved for the case of an infinite cylinder rolling on a corrugated surface. The objective of the present section is to derive a solution to the equivalent stationary tangential contact problem. It is naturally a dubious simplification to presuppose the contact to be stationary, especially for the case of short pitch corrugation where the wave lengths are so small that the displacement gradients inside the contact patch certainly are time dependent. The tangential problem is however treated as a stationary contact as this provides a good understanding of the behaviour of the problem. The case of non-steady contact will be treated in Chapter 7.

As indicated in section 3.3.2 the tangential stress distribution is assumed to be described as the sum of the two polynomial forms $q_1(x)$ and $q_2(x^*)$, where $q_1(x) = \mu p(x)$ and where $q_2(x^*)$ is unknown. From the normal contact problem it is known that the contribution from $q_1(x)$ to the displacement gradient inside the contact patch is

$$\frac{du_{x1}(x)}{dx} = \sum_{n=0}^{\infty} \mu\beta_n x^n \quad \Rightarrow$$

$$\frac{du_{x1}(x)}{dx} = \frac{\mu(x+\Delta)}{R} + \sum_{m=0}^{M} \mu k_m \left[Z_{A,m} \sin(\omega_m t + k_m x + k_m\Delta) - \right.$$
$$\left. Z_{B,m} \cos(\omega_m t + k_m x + k_m\Delta) \right] \qquad (4.43)$$

With this expression inserted into the kinematic constraint

$$0 = \xi_0 + \frac{du_{x1}(x)}{dx} + \frac{du_{x2}(x^*)}{dx} \qquad (4.44)$$

the contribution from $q_2(x^*)$ is calculated as

$$\frac{\partial u_{x2}(x,t)}{\partial x} = -\frac{\mu x}{R} - \xi_0 + \sum_{m=0}^{M} \mu k_m \left[Z_{A,m} \cos(k_m x + \omega_m t) + Z_{B,m} \sin(k_m x + \omega_m t)\right] \tag{4.45}$$

which in the x^* coordinate system is

$$\frac{\partial u_{x2}(x^*,t)}{\partial x} = -\xi_0 - \frac{\mu(x^* + \Delta^*)}{R} + \sum_{m=0}^{M} \mu k_m \left[Z_{A,m} \cos[k_m(x^* + \Delta^*) + \omega_m t] + Z_{B,m} \sin[k_m(x^* + \Delta^*) + \omega_m t]\right] \tag{4.46}$$

where

$$\Delta^* = \Delta - a + a^* \tag{4.47}$$

Thus the displacement gradient can be described by an infinite polynomial in x^*

$$\frac{du_{x2}(x)}{dx} = \sum_{n=0}^{\infty} \beta_n^* x^{*n} \tag{4.48}$$

Because $du_{x2}(x)/dx$ is a polynomial, the polynomial approach can be applied, i.e. $q_2(x^*)$ can be expressed as a polynomial form:

$$q_2(x^*) = \frac{\mu \pi E}{4(1-\nu^2)} \frac{\sum_{n=0}^{\infty} B_n^* x^{*n}}{\sqrt{a^{*2} - x^{*2}}} \tag{4.49}$$

Equivalent to the calculation of the B_n's in the normal contact problem, the coefficients to the polynomial form $q_2(x^*)$ are found as

$$B_1^* = -\frac{\xi_0}{\pi\mu} + \frac{\Delta^*}{\pi R} - \sum_{m=1}^{M} k_m \left[Z_{A,m} \sin(\omega_m t + k_m \Delta^*) - Z_{B,m} \cos(\omega_m t + k_m \Delta^*)\right] \sum_{j=0}^{\infty} (-1)^j A_j^{*-1} \frac{k_m^{2j}}{(2j)!} \tag{4.50}$$

$$B_2^* = \frac{1}{\pi R} - \sum_{m=1}^{M} k_m \left[Z_{A,m} \cos(\omega_m t + k_m \Delta^*) + \right.$$
$$\left. Z_{B,m} \sin(\omega_m t + k_m \Delta^*) \right] \sum_{j=0}^{\infty} (-1)^j A_j^{*-1} \frac{k_m^{2j+1}}{(2j+1)!} \tag{4.51}$$

$$B_3^* = \sum_{m=1}^{M} k_m \left[Z_{A,m} \sin(\omega_m t + k_m \Delta^*) - \right.$$
$$\left. Z_{B,m} \cos(\omega_m t + k_m \Delta^*) \right] \sum_{j=0}^{\infty} (-1)^j A_j^{*-1} \frac{k_m^{2j+2}}{(2j+2)!} \tag{4.52}$$

$$B_4^* = \sum_{m=1}^{M} k_m \left[Z_{A,m} \cos(\omega_m t + k_m \Delta^*) + \right.$$
$$\left. Z_{B,m} \sin(\omega_m t + k_m \Delta^*) \right] \sum_{j=0}^{\infty} (-1)^j A_j^{*-1} \frac{k_m^{2j+3}}{(2j+3)!} \tag{4.53}$$

$$\vdots$$

$$B_{2n-1}^* = (-1)^n \sum_{m=1}^{M} k_m \left[Z_{A,m} \sin(\omega_m t + k_m \Delta^*) - \right.$$
$$\left. Z_{B,m} \cos(\omega_m t + k_m \Delta^*) \right] \sum_{j=0}^{\infty} (-1)^j A_j^{*-1} \frac{k_m^{2j+2n-2}}{(2j+2n-2)!} \tag{4.54}$$

$$B_{2n}^* = (-1)^n \sum_{m=1}^{M} k_m \left[Z_{A,m} \cos(\omega_m t + k_m \Delta^*) + \right.$$
$$\left. Z_{B,m} \sin(\omega_m t + k_m \Delta^*) \right] \sum_{j=0}^{\infty} (-1)^j A_j^{*-1} \frac{k_m^{2j+2n-1}}{(2j+2n-1)!} \tag{4.55}$$

The coefficient B_0^* is derived from one of the boundary conditions

$$B_0^* = -\frac{a^{*2}}{\pi R} + \left(\frac{a^*}{\pi} \right)^2 \sum_{m=1}^{M} k_m \left[Z_{A,m} \cos(\omega_m t + k_m \Delta^*) + \right.$$
$$\left. Z_{B,m} \sin(\omega_m t + k_m \Delta^*) \right] \sum_{j=0}^{\infty} (-1)^j A_j^* \frac{k_m^{2j+1}}{(2j+1)!} \tag{4.56}$$

while the other boundary condition yields that

$$\xi_0 = \frac{\mu\Delta^*}{R} + \sum_{m=0}^{M} \mu k_m \mathrm{J}_0(k_m a^*) \left[Z_{A,m} \sin(\omega_m t + k_m \Delta^*) - Z_{B,m} \cos(\omega_m t + k_m \Delta^*) \right] \tag{4.57}$$

There are obvious similarities between the β_n's and the β_n^*'s and thus also between the B_n's and the B_n^*'s. Consequently the considerations from the previous section regarding the convergence of the coefficients B_n and the error introduced by truncating the infinite series are also valid for the B_n^*'s in the tangential contact problem.

The only unknown quantity in the above equation is half the size of the stick zone a^*. It is not possible to find an exact analytical closed form expression for a^*, but with numerical methods where the a_0^* from the Carter solution is used as initial guess, a^* can be found in just a few iterations. Another approach is to assume the amplitude of a^* to be small and then employ a Taylor expansion of the equation as in the normal contact problem. With this method the size of the stick zone is found as

$$a^* = a_0^* - \Delta(X) + a(X) - a_0 - \sum_{m=0}^{M} k_m R \mathrm{J}_0(k_m a_0^*) \left[Z_{A,m} \sin[\omega_m t + k_m(a_0^* - a_0)] - Z_{B,m} \cos[\omega_m t + k_m(a_0^* - a_0)] \right] \tag{4.58}$$

Making the substitution $\omega_m t = k_m X$ and inserting the expressions for $\Delta(X)$ and $a(X)$ where the normal load is assumed to be constant, the size of half the stick zone depending on the cylinders position on the surface is expressed as

$$a^*(X) = a_0^* + \sum_{m=0}^{M} \left[a_{A,m}^* \cos(k_m X) + a_{B,m}^* \sin(k_m X) \right] \tag{4.59}$$

$$\left\{ \begin{array}{c} a^*_{A,m} \\ a^*_{B,m} \end{array} \right\} = \left[\begin{array}{cc} a^*_{1,m} & -a^*_{2,m} \\ a^*_{2,m} & a^*_{1,m} \end{array} \right] \left\{ \begin{array}{c} Z_{A,m} \\ Z_{B,m} \end{array} \right\} \tag{4.60}$$

where the matrix coefficients are

$$a^*_{1,m} = k_m R \left[-\mathrm{J}_1(k_m a_0) - \mathrm{J}_0(k_m a^*_0) \sin\left[k_m(a^*_0 - a_0)\right]\right] \tag{4.61}$$

$$a^*_{2,m} = k_m R \left[\mathrm{J}_0(k_m a_0) - \mathrm{J}_0(k_m a^*_0) \cos\left[k_m(a^*_0 - a_0)\right]\right] \tag{4.62}$$

Now let the surface be harmonic with only one characteristic wave length L i.e. one characteristic wave number k. It is seen from equation (4.59) that the size of the stick zone oscillates with the same wave length as the wave length of the corrugation. Let $\widehat{a}^*$ denote the amplitude of a^* and let ϕ_{a^*} be the phase of a^* with respect to the surface then

$$\widehat{a}^* = \sqrt{(a^*_1)^2 + (a^*_2)^2}\sqrt{Z_A^2 + Z_B^2} \tag{4.63}$$

$$\phi_{a^*} = \arctan\left(\frac{a^*_2}{a^*_1}\right) \tag{4.64}$$

With the use of the polynomial approach it is thus possible to establish analytical expressions for the amplitude and the phase of the oscillating size of half the stick zone. It is seen that these properties can be normalised so they only depend on the ratio r_{a_0} defined as

$$r_{a_0} = \frac{a^*_0}{a_0} \tag{4.65}$$

and the relative size of the contact patch (a_0/L). In Figure 4.6 the amplitude (equation (4.63)) and the phase lag (equation (4.64)) are plotted for various values of r_{a_0} and a_0/L. It is seen that ϕ_{a^*} is quite sensible to the size of the r_{a_0} ratio. When $r_{a_0} = 0.5$ i.e. when the stick zone covers half the contact patch, a sudden shift in the phase of a^* occurs.

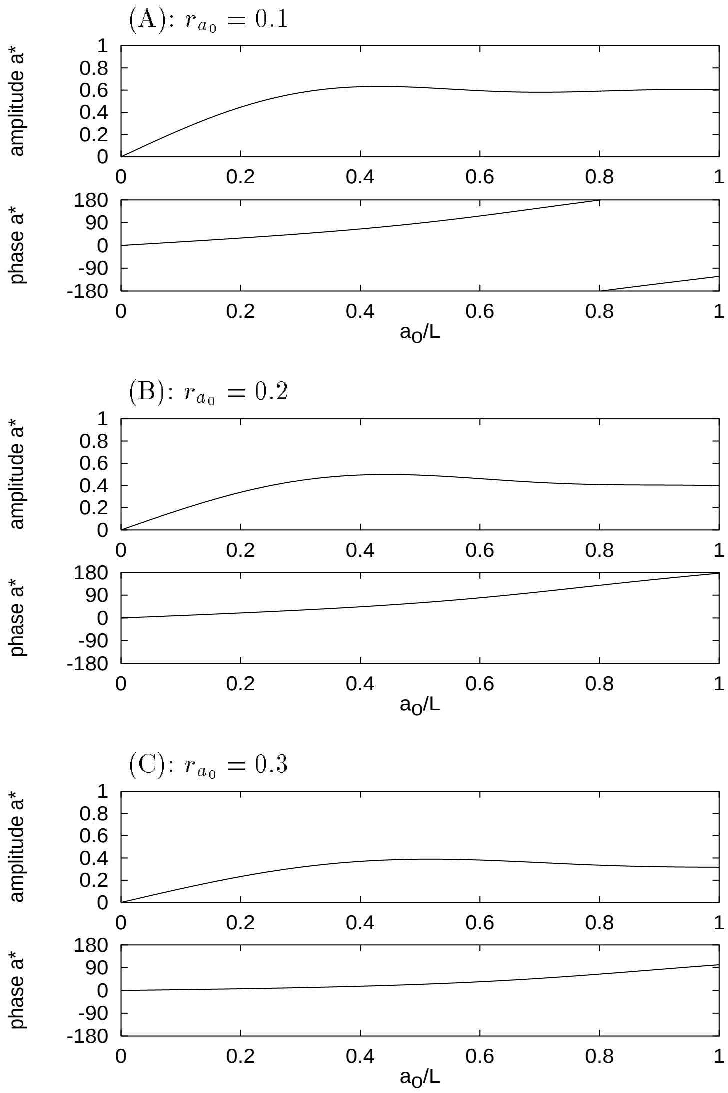

Figure 4.6. (A)-(C) The size of half the stick zone. Top: amplitude normalised with $kR\sqrt{Z_A^2 + Z_B^2}$. Bottom: phase with respect to $Z_2(X)$.

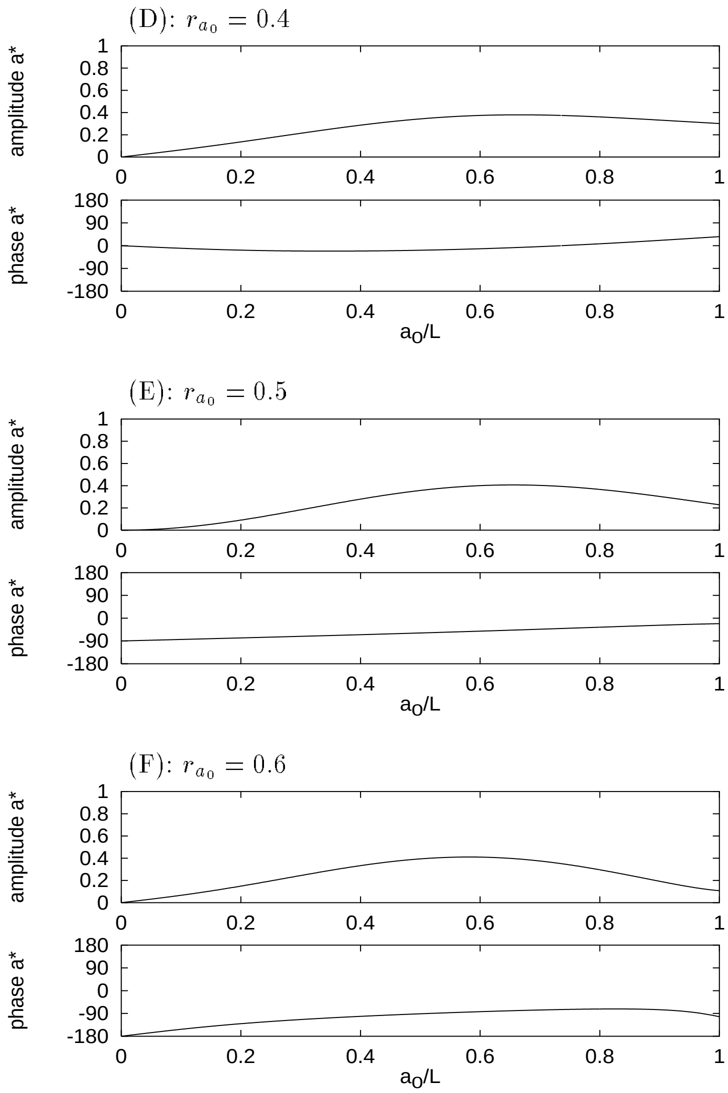

Figure 4.6. (D)-(F) The size of half the stick zone. Top: amplitude normalised with $kR\sqrt{Z_A^2 + Z_B^2}$. Bottom: phase with respect to $Z_2(X)$.

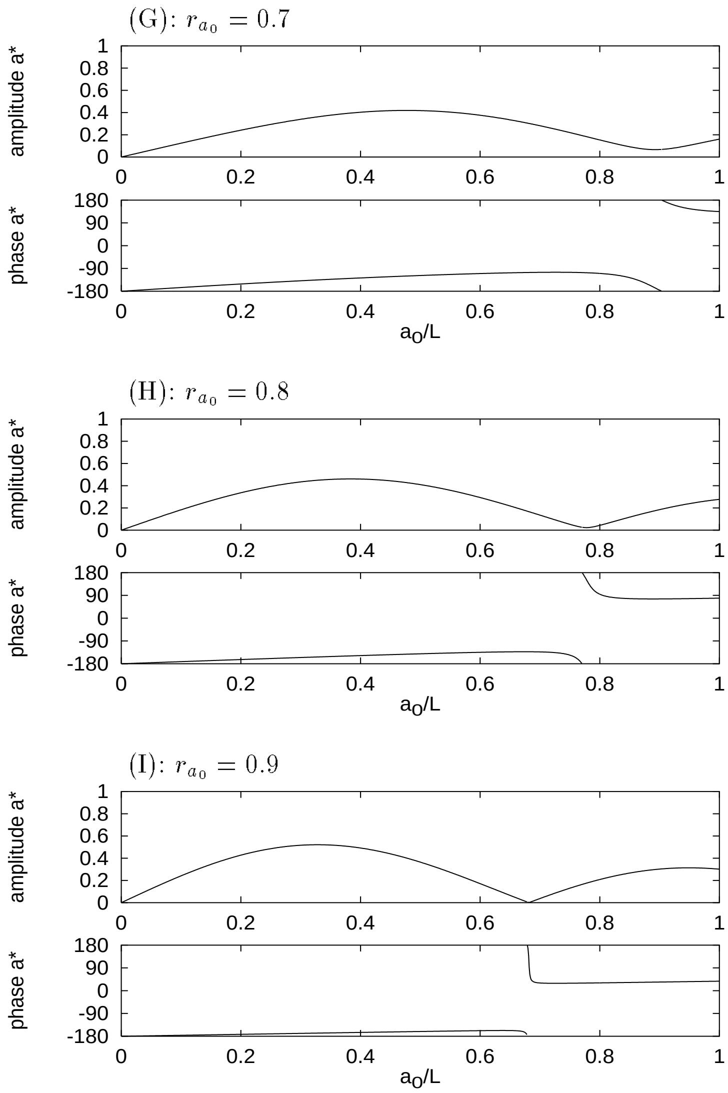

Figure 4.6. (G)-(I) The size of half the stick zone. Top: amplitude normalised with $kR\sqrt{Z_A^2 + Z_B^2}$. Bottom: phase with respect to $Z_2(X)$.

Like for the case of the normal pressure distribution it is not possible to derive a closed form expression for the tangential stress distribution, but an arbitrary good approximation is found by employing a finite number of the B_n^* coefficients. In Figure 4.7 the tangential stress distribution calculated with the polynomial approach is compared with the Carter solution for the cylinder rolling on a level surface. As the tangential stress distribution in the slip zone is closely related to the normal pressure distribution due to the friction law of Coulomb, an asymmetry equivalent to the one from the normal contact problem also appears in the tangential contact problem.

It is very important to notice that an oscillating behaviour takes place inside the contact patch both with respect to the size of the stick zone and with respect to the magnitude of the local tangential stress. This property is important when it comes to wear calculations as the wear only takes place in the slip zone and depends on the magnitude of the tangential stress. A thorough investigation of wear phenomena will be carried out in the next section.

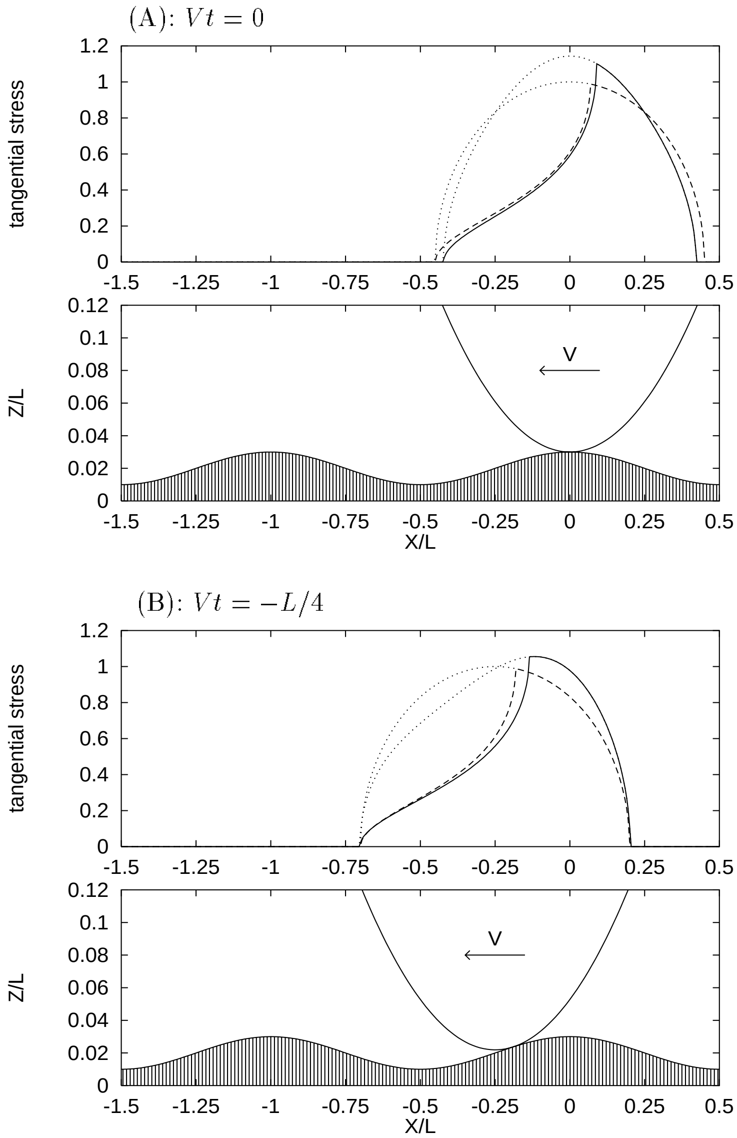

Figure 4.7. (A)-(B) Tangential stress distribution. Top: stress distribution calculated with the polynomial approach (solid line) and calculated with the Carter theory (dashed line). Bottom: position of the cylinder.

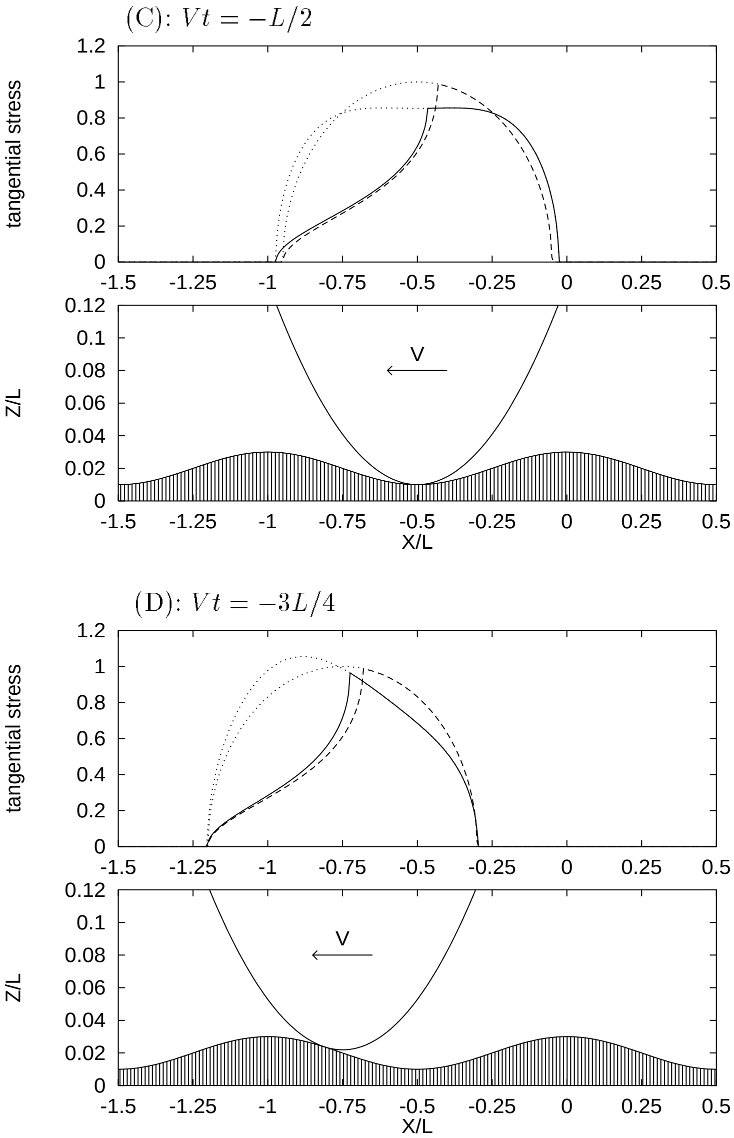

Figure 4.7. (C)-(D) Tangential stress distribution. Top: stress distribution calculated with the polynomial approach (solid line) and calculated with the Carter theory (dashed line). Bottom: position of the cylinder.

The tangential force is found by integrating the tangential stress distribution over the entire contact patch. It is evident that the contribution from $q_1(x)$ is μN i.e. it is only necessary to integrate $q_2(x^*)$ in order to find the tangential force. By applying the equation

$$T_2 = -\frac{\mu \pi E}{4(1-\nu^2)} \{A_0^*\}^T \{B_0^*\} \tag{4.66}$$

the tangential force is found to be $T = \mu N + T_2$:

$$T(X) = \mu N - \frac{\pi \mu E}{4(1-\nu^2)} \left\{ \frac{a^{*2}}{2R} + \sum_{m=0}^{M} k_m a^* \mathrm{J}_1(k_m a^*) \cdot \right.$$
$$\left. [Z_{A,m} \cos[k_m(X+\Delta^*)] + Z_{B,m} \sin[k_m(X+\Delta^*)]] \right\} \tag{4.67}$$

When $a(X)$, $a^*(X)$ and $\Delta(X)$ are known the tangential force can be calculated in a straightforward way. If the amplitudes of these quantities are small it is however possible to make a simplification without any appreciable loss of accuracy. The tangential force is then expressed as

$$T(X) = T_0 + \sum_{m=0}^{M} [T_{A,m} \cos(k_m X) + T_{B,m} \sin(k_m X)] \tag{4.68}$$

$$\begin{Bmatrix} T_{A,m} \\ T_{B,m} \end{Bmatrix} = \begin{bmatrix} T_{1,m} & -T_{2,m} \\ T_{2,m} & T_{1,m} \end{bmatrix} \begin{Bmatrix} Z_{A,m} \\ Z_{B,m} \end{Bmatrix} \tag{4.69}$$

where T_0 is the Carter value and the matrix coefficients are

$$T_{1,m} = \frac{\pi \mu E}{4(1-\nu^2)} k_m a_0^* \left[\frac{-a_{1,m}^*}{k_m R} - \mathrm{J}_1(k_m a_0^*) \cos[k_m(a_0^* - a_0)] \right] \tag{4.70}$$

$$T_{2,m} = \frac{\pi \mu E}{4(1-\nu^2)} k_m a_0^* \left[\frac{-a_{2,m}^*}{k_m R} + \mathrm{J}_1(k_m a_0^*) \sin[k_m(a_0^* - a_0)] \right] \tag{4.71}$$

Assuming that the level of the surface is harmonic with the characteristic wave length L, this implies that the amplitude $\widehat{T}$ and the phase ϕ_T of the tangential force are

$$\widehat{T} = \sqrt{(T_1)^2 + (T_2)^2}\sqrt{Z_A^2 + Z_B^2} \tag{4.72}$$

$$\phi_T = \arctan\left(\frac{T_2}{T_1}\right) \tag{4.73}$$

These values are plotted in Figure 4.8 for different values of r_{a_0} and a_0/L. It is seen that when r_{a_0} is small both the amplitude and the phase are almost constant when a_0/L also is small. When the r_{a_0} ratio grows $\widehat{T}$ is more sensible to the value of a_0/L. The reason for this behaviour lies in the fact that the size of the stick zone compared with the wave length of the corrugation is important for the behaviour of the tangential contact problem.

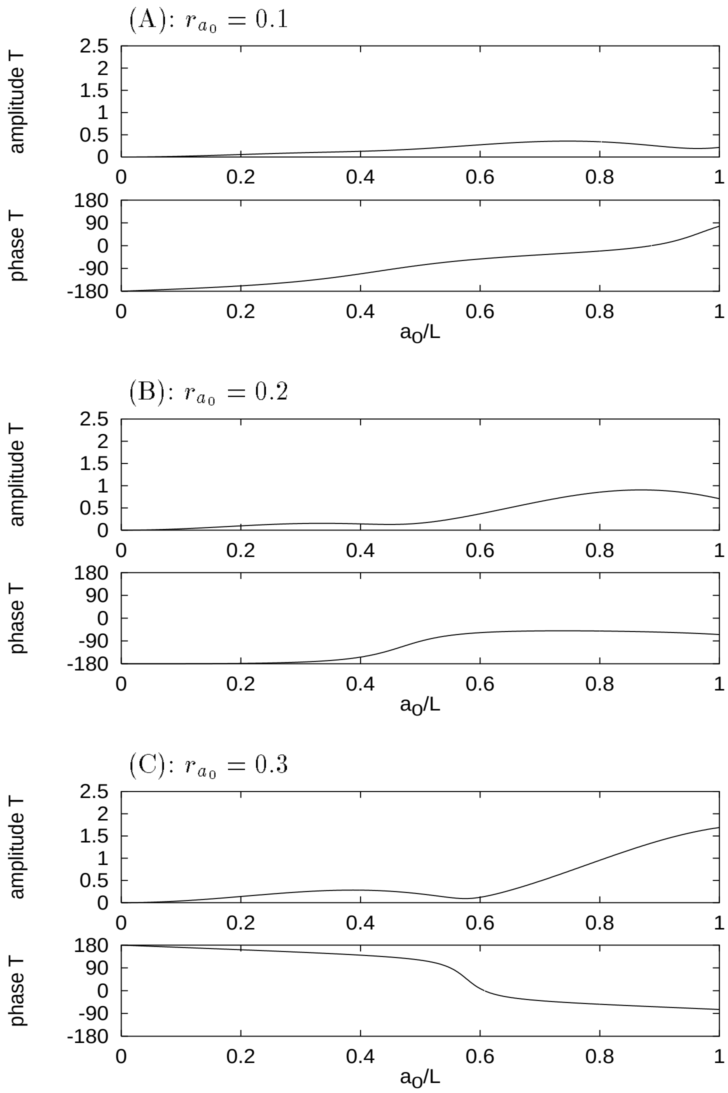

Figure 4.8. (A)-(C) The tangential force. Top: amplitude normalised with $\sqrt{Z_A^2 + Z_B^2}\,\mu\pi E/(4(1-\nu^2))$. Bottom: phase with respect to $Z_2(X)$.

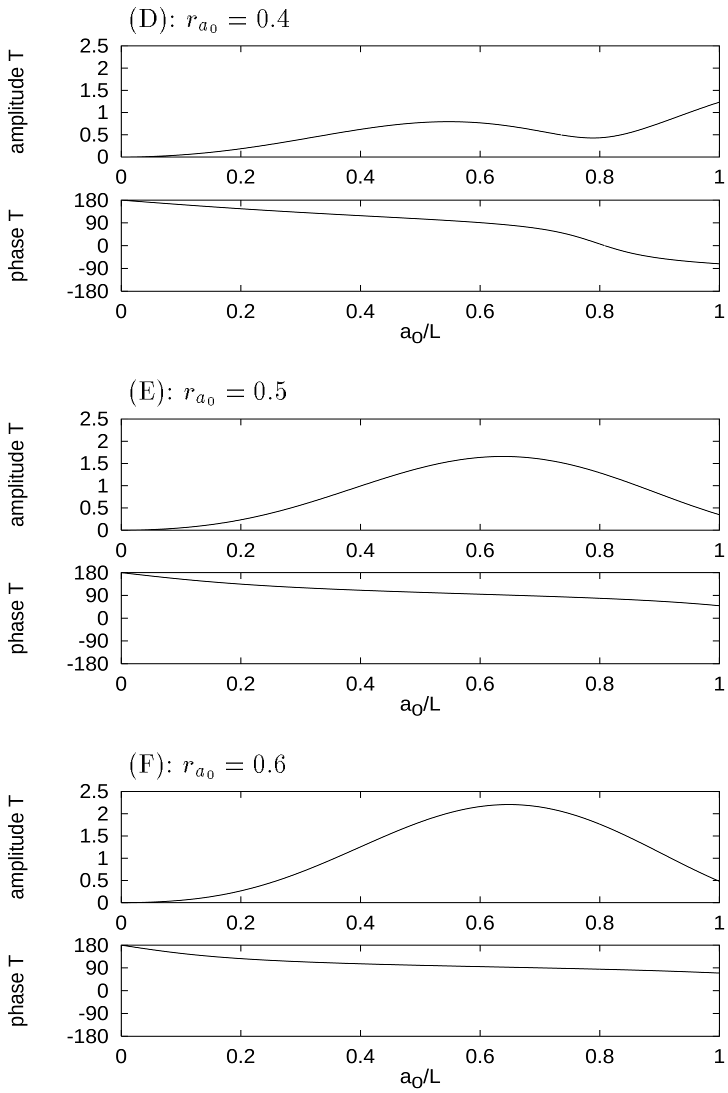

Figure 4.8. (D)-(F) The tangential force. Top: amplitude normalised with $\sqrt{Z_A^2 + Z_B^2}\,\mu\pi E/(4(1-\nu^2))$. Bottom: phase with respect to $Z_2(X)$.

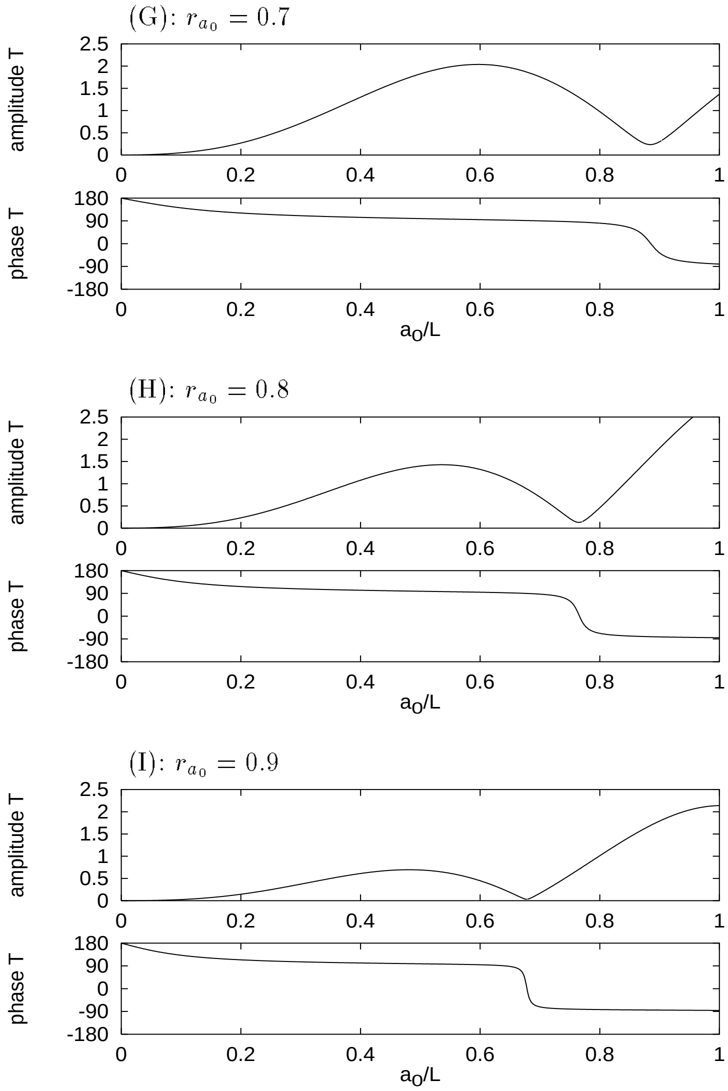

Figure 4.8. (G)-(I) The tangential force. Top: amplitude normalised with $\sqrt{Z_A^2 + Z_B^2}\,\mu\pi E/(4(1-\nu^2))$. Bottom: phase with respect to $Z_2(X)$.

4.3 Wear

Corrugation of rails is a result of a wear process which takes place as the wheel rolls over the rail. Many different hypotheses have been developed to explain the wear mechanism. In general wear mechanisms are divided into two groups: one containing local and sudden wear and one including wear which develops over a long period of time. The local wear is a result of isolated abnormalities in the wheel/rail interaction such as wheel flats caused by sudden blocking of the wheel or shelling due to high local temperatures [45]. Whereas this type of wear is caused by isolated events the long time wear is related to the consecutive rolling motion and occurs without non-typical events such as wheel blocking or major irregularities of the track. This implies that the long time wear mechanism will cause the wheel and the rail to be worn even if the wheel is perfectly circular and smooth and the rail is completely level and smooth.

The reason for long time wear can either be rolling contact fatigue, plastic deformation or material removal in the contact patch. Rolling contact fatigue is a crack generating mechanism caused by the characteristic sequence of tensile and shear stresses under a rolling load. The rolling contact fatigue is certainly a long time effect, but as there is no apparent relation between rolling contact fatigue and cracks on one side and corrugation of wheel and rail on the other side the topic will not be treated in this work. A survey of rolling contact fatigue can be found in [16].

Plastic deformation is a result of large normal pressures in the contact patch. It is generally accepted that plastic deformation is of minor interest in the investigation of corrugation phenomena [13]. Initially the surface of a newly ground rail may undergo plastic deformations in a very thin layer at the surface, but this will introduce residual stresses which remain after the normal pressure on the material is relieved. As the number of wheel passages increases, the material will harden as the residual stresses are built up and eventually the surface of the rail will be sufficiently hard to sustain the normal pressures without any plastic deformation. This process is often referred to as a shakedown mechanism [6].

This leaves the removal of material due to frictional work as the major wear mechanism when it comes to the formation of corrugation. The most common theory for frictional wear is developed by Frederic [17] and states that the wear is proportional to the frictional work in the contact patch. When a wheel rolls over a rail under the influence of a certain torque, tangential forces are transmitted between the bodies in the contact patch. As the relative velocity in a part of the contact patch – the slip zone – is different from zero, frictional work is created. Defining the wear $W(X)$ as the material removed from the surface after one passage of the wheel then

$$W(X) = \left| KV \int_S q(X,t)s(X,t)\, dt \right| \tag{4.74}$$

where K is a material depending constant. For steel $K \simeq 7.58 \cdot 10^{-6}\ \mathrm{m}^2/\mathrm{N}$.

Many models where calculation of the wheel/rail wear is carried out have been published e.g. [29], [30] and [65]. In most models the wear hypothesis is simplified such that the integral in equation (4.74) is replaced by the tangential force multiplied with the creepage:

$$W(X) = |KT(X)\xi| \tag{4.75}$$

This linearised version of equation (4.74) is however much too primitive for the case of a cylinder rolling on a corrugated surface. In the previous section it was shown that even when the amplitude of the tangential force is very small, the size of the stick zone and the local tangential stress oscillate with quite a large amplitude (see e.g. Figure 4.6 (B) and Figure 4.8 (B). In those cases the wear calculated with equation (4.75) is almost constant whereas the actual wear derived from equation (4.74) will oscillate. It is thus necessary to solve the integral in equation (4.74) in order to obtain a satisfactory evaluation of the wear.

4.3.1 Calculating the Wear

To derive a solution to the integral (4.74) it is convenient to find an expression for the amount of time an arbitrary point on the surface is located in the slip zone. Consider a point X_0 on the corrugated surface. This point will be exactly at the leading edge of the contact zone at the time t_0 for which

$$Vt_0 + \Delta(Vt_0) - a(Vt_0) = X_0 \tag{4.76}$$

Introducing t_* as the time when X_0 is located at the limit between the stick zone and the slip zone and t_1 as the time where X_0 is located at the trailing edge of the contact patch (see Figure 4.9) then

$$Vt_* + \Delta(Vt_*) + 2a^*(Vt_*) - a(Vt_*) = X_0 \tag{4.77}$$

$$Vt_1 + \Delta(Vt_1) + a(Vt_1) = X_0 \tag{4.78}$$

Now let τ be half the time it takes for X_0 to travel through the contact patch and similarly let τ^* be half the time it takes for the same point to travel through the stick zone then

$$\tau(X_0) = \frac{1}{2}(t_1 - t_0) \tag{4.79}$$

$$\tau^*(X_0) = \frac{1}{2}(t_* - t_0) \tag{4.80}$$

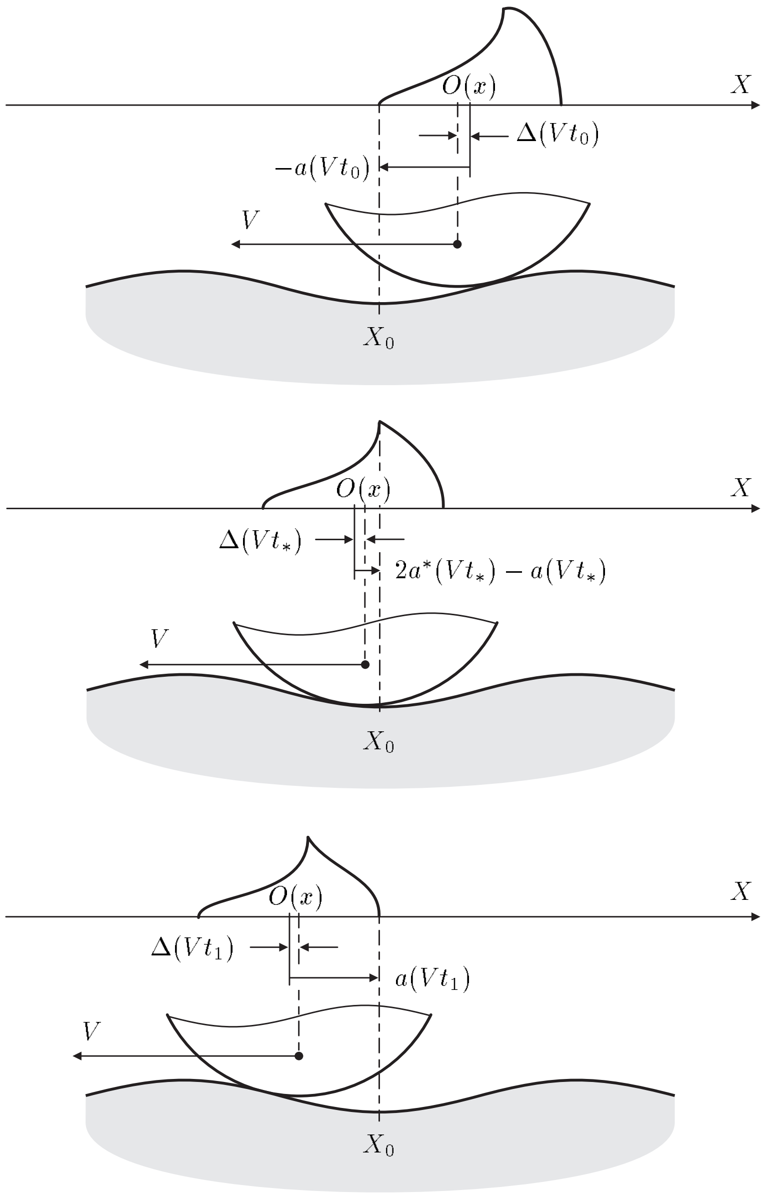

Figure 4.9. The contact patch passing a point on the surface.

Assuming the amplitudes of a, Δ and a^* to be small it is possible to make a first order Taylor expansion of τ and τ^* which gives that

$$\tau(X_0) = -\frac{a_0}{V} + \sum_{m=0}^{M} \left[\tau_{A,m}\cos(k_m X_0) + \tau_{B,m}\sin(k_m X_0)\right] \tag{4.81}$$

$$\left\{\begin{array}{c} \tau_{A,m} \\ \tau_{B,m} \end{array}\right\} = \left[\begin{array}{cc} \tau_{1,m} & 0 \\ 0 & \tau_{1,m} \end{array}\right] \left\{\begin{array}{c} Z_{A,m} \\ Z_{B,m} \end{array}\right\} \tag{4.82}$$

where

$$\tau_{1,m} = \frac{k_m R}{V}\left[\mathrm{J}_0(k_m a_0)\sin(k_m a_0) + \mathrm{J}_1(k_m a_0)\cos(k_m a_0)\right] \tag{4.83}$$

Similarly $\tau^*(X_0)$ is found as

$$\tau^*(X_0) = -\frac{a_0^*}{V} + \sum_{m=0}^{M} \left[\tau^*_{A,m}\cos(k_m X_0) + \tau^*_{B,m}\sin(k_m X_0)\right] \tag{4.84}$$

$$\left\{\begin{array}{c} \tau^*_{A,m} \\ \tau^*_{B,m} \end{array}\right\} = \left[\begin{array}{cc} \tau^*_{1,m} & -\tau^*_{2,m} \\ \tau^*_{2,m} & \tau^*_{1,m} \end{array}\right] \left\{\begin{array}{c} Z_{A,m} \\ Z_{B,m} \end{array}\right\} \tag{4.85}$$

where

$$\begin{aligned} \tau^*_{1,m} = \frac{k_m R}{2V} & \left[\mathrm{J}_0(k_m a_0)\left[3\sin\left[k_m(2a_0^* - a_0)\right] - \sin\left[k_m a_0\right]\right] + \right. \\ & \mathrm{J}_1(k_m a_0)\left[\cos\left[k_m(2a_0^* - a_0)\right] + \cos[k_m a_0]\right] - \\ & \left. 2\mathrm{J}_0(k_m a_0^*)\sin\left[k_m a_0^*\right]\right] \qquad (4.86) \\ \tau^*_{2,m} = -\frac{k_m R}{2V} & \left[\mathrm{J}_0(k_m a_0)\left[3\cos\left[k_m(2a_0^* - a_0)\right] - \cos[k_m a_0]\right]\right] - \\ & \mathrm{J}_1(k_m a_0)\left[\sin\left[k_m(2a_0^* - a_0)\right] - \sin[k_m a_0]\right] - \\ & 2\mathrm{J}_0(k_m a_0^*)\cos\left[k_m a_0^*\right]] \qquad (4.87) \end{aligned}$$

There are obviously strong similarities between a and τ and between a^* and τ^*. The a and a^* represent the size of the contact patch and the size of the stick zone in space-domain whereby τ and τ^* are the equivalent properties in time-domain. Even though there evidently are points of resemblance between the quantities in time-domain and in space-domain it is relevant to focus on some qualitative differences.

Whereas τ as well as a always will have the same phase as the corrugated surface, the representation of the stick zone given by a^* and τ^* has a phase shift with respect to the corrugation pattern. It is however worth noticing that the phases of τ^* and a^* are not equivalent. This property is a consequence of the fact that τ^* depends on the position of the leading edge of the contact patch at the time t_0 and the position of the limit between stick zone and slip zone at the time t_*. Because the size of the stick zone oscillates in time and τ^* depends on two different times the relation between a^* and τ^* is quite complicated.

In the following calculations it is demonstrated that the ratio $\tau^*(X)/\tau(X)$ is important for the properties of the wear. Due to the above stated discrepancies between the description in time-domain and in space-domain it must for that reason be emphasized that

$$\frac{\tau^*(X)}{\tau(X)} \neq \frac{a^*(X)}{a(X)} \tag{4.88}$$

neither with respect to the phase nor with respect to the amplitude.

When the expressions for τ and τ^* are derived it is possible to solve the integral in equation (4.74). From the previous sections it follows that the tangential stress distribution and the slip in the slip zone can be expressed as polynomial forms and so the wear is expressed as

$$W(X) = \left| KV \int_{2\tau^*(X)-\tau(X)}^{\tau(X)} q(X,t)s(X,t)\,dt \right| \tag{4.89}$$

$$q(X,t) = \frac{\mu \pi E}{4(1-\nu^2)} \frac{\sum_{n=0}^{\infty} B_n(Vt)[X - Vt - \Delta(X)]^n}{\sqrt{a^2(Vt) - [X - Vt - \Delta(Vt)]^2}} \tag{4.90}$$

$$s(X,t) = \mu \pi \frac{\sum_{m=0}^{\infty} B_n^*(Vt)[X - Vt - \Delta^*(Vt)]^n}{\sqrt{a^{*2}(Vt) - [X - Vt - \Delta^*(Vt)]^2}} \tag{4.91}$$

which can be calculated to

$$W(X) = \left| KVN \frac{R}{a_0^2} \left(\frac{\mu \pi B_0(X)}{a(X)} \right)^2 \tau(X) \left[1 - r_\tau(X) - r_\tau^2(X) + r_\tau^3(X) \right] \right| \tag{4.92}$$

$$r_\tau(X) = \frac{\tau^*(X)}{\tau(X)} \tag{4.93}$$

Presupposing that the amplitudes of a, a^* and Δ are small the expression for the wear can be approximated by

$$W(X) = W_0 + \sum_{m=0}^{M} \left[W_{A,m} \cos(k_m X_0) + W_{B,m} \sin(k_m X_0) \right] \tag{4.94}$$

$$\left\{ \begin{array}{c} W_{A,m} \\ W_{B,m} \end{array} \right\} = \left[\begin{array}{cc} W_{1,m} & -W_{2,m} \\ W_{2,m} & W_{1,m} \end{array} \right] \left\{ \begin{array}{c} Z_{A,m} \\ Z_{B,m} \end{array} \right\} \tag{4.95}$$

where

$$W_0 = K\mu^2 N \frac{a_0}{R} \left[1 - r_{a_0} - r_{a_0}^2 + r_{a_0}^3 \right] \tag{4.96}$$

$$\begin{aligned} W_{1,m} = K\mu^2 N \frac{k_m V}{R} \Big[\tau_1^* \left[1 + 2r_{a_0} - 3r_{a_0}^2 \right] - \\ \tau_1 \left[1 + r_{a_0}^2 - 2r_{a_0}^3 \right] \Big] + \\ 2K\mu^2 N \left[1 - r_{a_0} - r_{a_0}^2 + r_{a_0}^3 \right] \cdot \\ \left[\mathrm{J}_1(k_m a_0) - \widehat{B}_0 \right] \end{aligned} \tag{4.97}$$

$$W_{2,m} = K\mu^2 N \frac{k_m V}{R} \; \tau_2^* \left[1 + 2r_{a_0} - 3r_{a_0}^3 \right] \tag{4.98}$$

The term W_0 represents the global wear of the surface and is thus uninteresting when it comes to the investigation of corrugation. Naturally it is of importance to know how the geometry of a cross section of the rail changes as the number of train passages increases, because a modified shape of the cross section influences on the contact geometry, but this effect is neglected in the present investigation. It is however noted that W_0 and thus the global wear does not depend on the initial corrugation pattern and is therefore equal to the expression for the wear arising from the Carter solution.

4.4 Evolution of the Corrugation

When the cylinder rolls over the surface a certain amount of material is removed from the surface. The height of the material removed at a given position on the surface is denoted the wear. The wear in itself is of limited interest as it is small compared to the characteristic size of the corrugation on the surface. As the number of cylinder passages increases the shape of the surface will slowly undergo an evolution from the initial corrugation pattern to another surface geometry. This evolution of the corrugation is crucial in wheel/rail wear because an adequate maintenance strategy depends on how fast the corrugation grows. The present section treats the subject of numerous consecutive cylinder passages and thus the evolution of an initial corrugation.

From equation (4.94) it is seen that one distinct wave length of the corrugation results in exactly the same wave length in the wear. The amplitudes and phases are different, but the wave lengths are identical. This property implies that there is no cross influence between the different wave lengths and so the problem of a corrugated surface with several characteristic wave lengths can be solved by calculating the wear for each wave length and then adding these results, provided that the component W_0 is included only once. Thus, to simplify the calculations it is assumed that the surface is harmonic with one distinct wave length.

With the wear defined as the height of the material removed as the cylinder rolls over the surface, it is obvious that the shape of the surface after one passage of the cylinder is equal the old surface minus the wear:

$$Z_2^{\{1\}}(X) = Z_2^{\{0\}}(X) - W^{\{0\}}(X) \tag{4.99}$$

where the high indices refer to the number of cylinder passages. As the initial surface and the wear are given by

$$Z_2^{\{0\}}(X) = Z_A^{\{0\}} \cos(kX) + Z_B^{\{0\}} \sin(kX) \tag{4.100}$$

$$W^{\{0\}}(X) = W_A^{\{0\}} \cos(kX) + W_B^{\{0\}} \sin(kX) \tag{4.101}$$

the shape of the new surface is

$$\begin{aligned} Z_2^{\{1\}}(X) &= \left(Z_A^{\{0\}} - W_A^{\{0\}}\right) \cos(kX) + \left(Z_B^{\{0\}} - W_B^{\{0\}}\right) \sin(kX) \quad \Rightarrow \\ Z_2^{\{1\}}(X) &= Z_A^{\{1\}} \cos(kX) + Z_B^{\{1\}} \sin(kX) \end{aligned} \tag{4.102}$$

In the previous section it was found that W_A and W_B depend linearly on Z_A and Z_B with the coefficients W_1 and W_2. Because W_1 and W_2 do not depend on the amplitude of the surface they are constant even when the shape of the surface changes as the number of cylinder passages increases. Due to this fact it is possible to find the surface coefficients after one passage of the cylinder via the discrete mapping

$$\left\{ \begin{array}{c} Z_A^{\{n+1\}} \\ Z_B^{\{n+1\}} \end{array} \right\} = \left[\begin{array}{cc} 1 - W_1 & W_2 \\ -W_2 & 1 - W_1 \end{array} \right] \left\{ \begin{array}{c} Z_A^{\{n\}} \\ Z_B^{\{n\}} \end{array} \right\} \tag{4.103}$$

With the above discrete mapping it is possible to calculate the shape of the surface after each passage of the cylinder in a straight forward analytical way. Thus time consuming integrations or space stepping methods are replaced by algebraic expressions, which makes the present approach extremely fast. As the method includes the oscillating behaviour inside the contact patch in contradiction to many other approaches it is also able to monitor effects that more primitive methods disregard. It must be emphasized that all the calculations presuppose that the amplitude of the corrugation is small compared to the wave length and the radius of the cylinder, but this is always fulfilled when it comes to the evolution of a very slight corrugation.

4.4.1 Stability of the Corrugation

The fact that the corrugation can be calculated by analytical closed forms makes it possible to derive some qualitative properties concerning the evolution of an initial corrugation. From the discrete mapping (4.103) it is found that the amplitude of the surface after one passage of the cylinder can be described by the former surface configuration as

$$\sqrt{\left(Z_A^{\{n+1\}}\right)^2 + \left(Z_B^{\{n+1\}}\right)^2} = \sqrt{1-2W_1}\sqrt{\left(Z_A^{\{n\}}\right)^2 + \left(Z_B^{\{n\}}\right)^2} \tag{4.104}$$

which can be generalised to

$$\sqrt{\left(Z_A^{\{n+1\}}\right)^2 + \left(Z_B^{\{n+1\}}\right)^2} = (1-2W_1)^{\frac{n+1}{2}}\sqrt{\left(Z_A^{\{0\}}\right)^2 + \left(Z_B^{\{0\}}\right)^2} \tag{4.105}$$

This means that the growth rate of the corrugation only depends on the absolute value of $(1-2W_1)$: if this term is smaller than 1 any initial amplitude will be levelled out whereas the amplitude grows exponentially if

$(1-2W_1) > 1$. As $|W_1| \ll 1$ this criteria of stability can be formulated as:

$$\begin{aligned} W_1 > 0 &\quad\Leftrightarrow\quad \text{the corrugation is levelled out} \\ W_1 = 0 &\quad\Leftrightarrow\quad \text{the corrugation is constant} \\ W_1 < 0 &\quad\Leftrightarrow\quad \text{the corrugation is amplified} \end{aligned} \tag{4.106}$$

It is important to notice that the wear component W_2 does not influence the growth rate of the corrugation even though it contributes to the magnitude of the wear amplitude. This illustrates the fact that the magnitude of the wear alone is not important for the evolution of the corrugation: it is the magnitude and the phase of the wear which is important. If the phase of the wear is such that it has its maximum on a corrugation spike the corrugation levels out whereas a maximum wear located at the position of a corrugation trough causes the corrugation to evolve very fast.

4.4.2 Amplifying and Levelling Zones

From the stability criteria (4.106) it is seen that W_1 is crucial for the evolution of the corrugation: this quantity determines the stability and the growth rate of the corrugation. A typical outline of W_1 is given in Figure 4.10. For a given r_{a_0}-value the outline of W_1 will always be qualitatively equivalent to the one monitored in Figure 4.10 with only one critical L/a_0 ratio for which $(1-2W_1) = 0$. This is the limit of stability i.e. surface irregularities with L/a_0 ratios smaller than this value are levelled out while the corrugation is amplified if the L/a_0 value is above the critical value. Furthermore it is seen that $(1-2W_1)$ tends towards 1^+ for long wave lengths. This states that if the wave length of the corrugation is large compared with a_0 then the amplitude of the corrugation is unaffected by the wear.

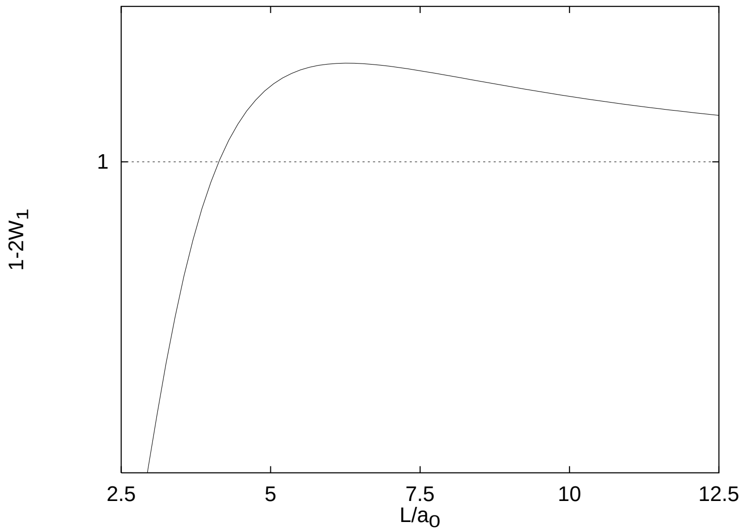

Figure 4.10. The wear coefficient $(1 - 2W_1)$ for $r_{a_0} = 0.6$

Because W_1 only depends on the relative size of the contact patch and the relative size of the stick zone it is possible to monitor the critical wave length in a $(a_0^*/a_0, L/a_0)$-diagram as done in Figure 4.11. The line indicates the $(a_0^*/a_0, L/a_0)$-values for which $W_1 = 0$. For all the combinations of a_0^*/a_0 and L/a_0 lying above this line $W_1 < 0$, which result in an amplification of the corresponding corrugation. Similarly values of a_0^*/a_0 and L/a_0 located below the line imply that the corrugation is levelled out.

In Figure 4.12 the qualitative evolution of a corrugated surface is shown. The initial surface corrugation consists of two wave lengths L_1 and L_2 where L_1 lies in the levelling zone and L_2 in the amplifying zone. The amplitude

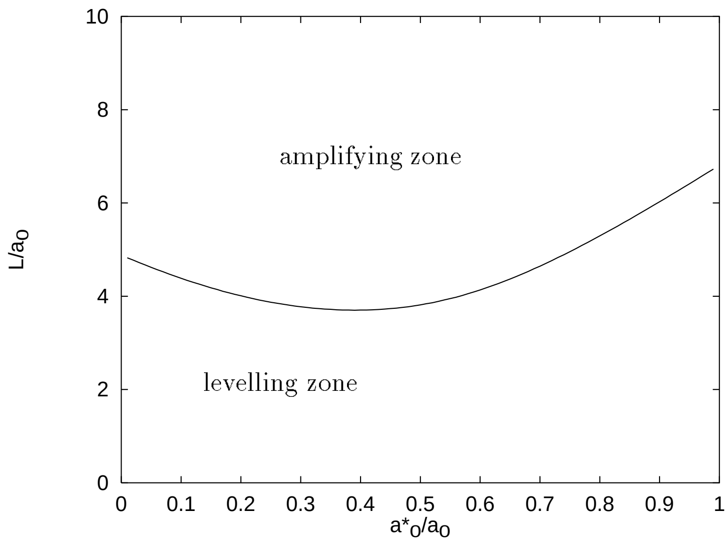

Figure 4.11. The smallest possible wave length for the corrugation

related to L_1 is initially 10 times bigger than the amplitude for the contribution from L_2:

$$Z_2^{\{0\}}(X) = Z_{L_1}^{\{0\}}(X) + Z_{L_2}^{\{0\}}(X) \tag{4.107}$$

$$Z_{L_1}^{\{0\}}(X) = 10\left[Z_A^{\{0\}}\cos(k_1X) + Z_B^{\{0\}}\sin(k_1X)\right] \tag{4.108}$$

$$Z_{L_2}^{\{0\}}(X) = \left[Z_A^{\{0\}}\cos(k_2X) + Z_B^{\{0\}}\sin(k_2X)\right] \tag{4.109}$$

i.e. $Z_{L_1}^{\{0\}}(X)$ is dominating the initial surface. As the corrugation evolves the wear will cause $Z_{L_1}^{\{0\}}(X)$ to be levelled out whereas $Z_{L_2}^{\{0\}}(X)$ is becoming more and more dominant. After a while the initial surface corrugation has completely vanished and is replaced by a corrugation pattern with another distinct wave length.

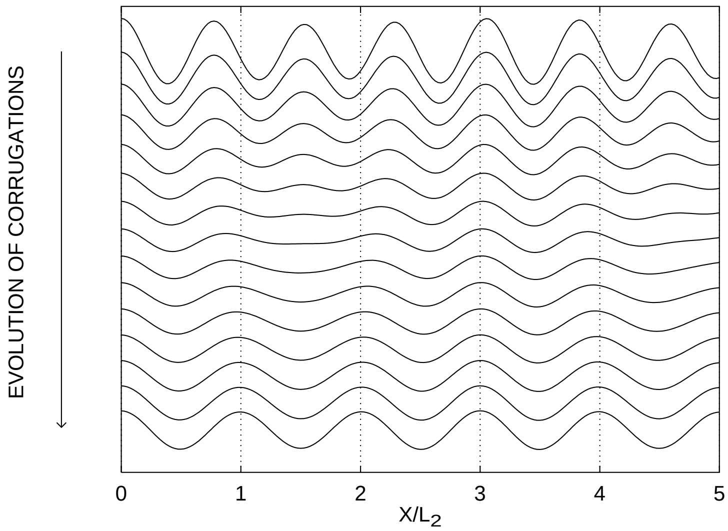

Figure 4.12. The qualitative evolution of corrugation

4.4.3 Characteristic Wave Length

It was in the previous section demonstrated how certain wave lengths will be amplified and other wave lengths will be levelled out. Now the question is whether one distinct wave length will evolve from an arbitrary surface configuration. In Figure 4.10 it was demonstrated that if a_0^*/a_0 is constant then $(1-2W_1)$ has a maximum for a certain L/a_0-value. At first this maximum does not seem to be very dramatic, but because the wear rate is given as $(1-2W_1)^{\frac{n+1}{2}}$ a very distinct peak in the frequency spectra will grow up as the number of cylinder passages increases. So this relative wave length will be dominating the corrugation, which explains why a certain corrugation pattern usually evolves with one and only one distinct wave length. This

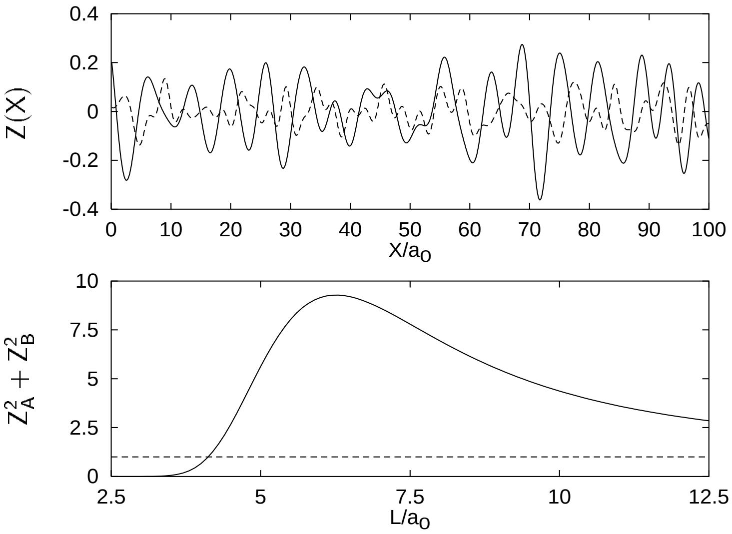

Figure 4.13. The evolution of surface regularities represented by white noise. Top: Level of the surface. Bottom: Spectrum of the surface. Dashed line: initial surface. Solid line: surface after a number of cylinder passages.

effect is seen in Figure 4.13 where the initial surface corrugation is represented by white noise. After a number of cylinder passages a corrugation pattern with one dominating wave length has evolved. The wear mechanism operates thus as a filter on the initial surface corrugation.

The filtering effect is a crucial feature for the evolution of the corrugation. In general the initial amplitudes of the surface irregularities are not important for the evolution of the corrugation. As the growth of the corrugation is exponential, the wave length of a surface component is far more important than the amplitude. In practice all wave lengths are to some extend represented on the surface of a rail, and so a certain wave length will emerge

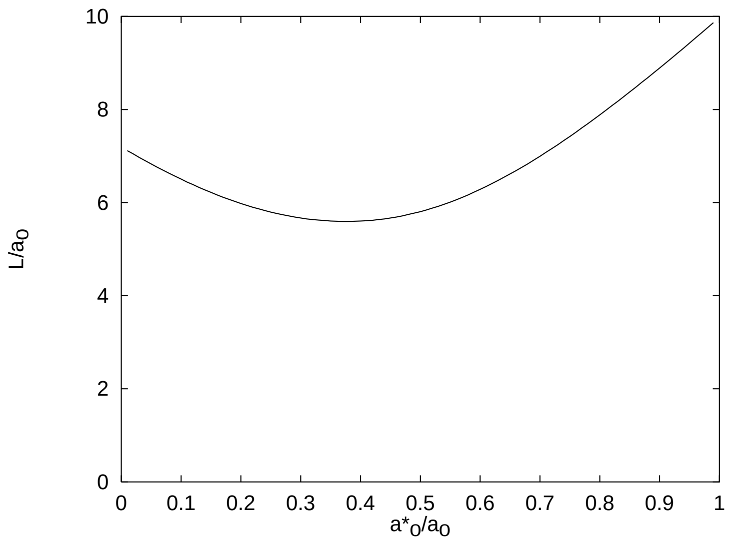

Figure 4.14. The characteristic wave lengths for which the corrugation is most likely to develop

without any apparent reason even though it is not very strongly represented in the initial wheel/rail system.

The characteristic wave length depends on the ratio a_0^*/a_0 and thus on the magnitude of the creepage. So if the creepage changes, the characteristic wave length of the corrugation will change. In Figure 4.14 the characteristic wave length is shown for different r_{a_0}-values. The line indicates the parameter combinations for which $(1-2W_1)$ has its local maximum. It is seen that the L/a_0-ratio for the characteristic wave length approximately lies in the range from 5-10. In wheel/rail contact the typical size of a_0 is somewhere between 5 mm and 10 mm, which thus provides a characteristic wave length in the interval 0.025 – 0.1 m. This fits very well with the observed wave lengths for short pitch corrugation [19].

The above analysis is made for the case where a_0^* and a_0 are constant for all the passages of the cylinder. In reality this assumption is only valid for closed line systems where the traffic is homogeneous. On railway lines with a large diversity in the traffic the size of the contact patch and the magnitude of the creepage depends on the type of the rolling stock. As the principle of superposition is valid the resulting wear rate after n passages is calculated as

$$\sqrt{\left(Z_A^{\{n\}}\right)^2 + \left(Z_B^{\{n\}}\right)^2} =$$

$$\left(\prod_{j=1}^{n}\left[1 - 2W_1\left(\frac{a_0^{\{j\}}}{L}, \frac{a_0^{*\{j\}}}{a_0^{\{j\}}}\right)\right]\right)^{\frac{1}{2}} \sqrt{\left(Z_A^{\{0\}}\right)^2 + \left(Z_B^{\{0\}}\right)^2} \quad (4.110)$$

where $a_0^{\{j\}}$ and $a_0^{*\{j\}}$ refer to the contact parameters for the j'th passage. Because the qualitative behaviour of $(1-2W_1)$ is unique, the resulting wear rate for a situation with many different types of rolling stock will in general have a similar behaviour i.e. an amplifying zone and a levelling zone plus one distinct wave length.

4.4.4 Moving Corrugation

A result of the changing phase of the wear is that the corrugation has a tendency to move along the surface. If the maximum wear is located on the downhill side of the corrugation spikes the corrugation pattern will move in the opposite direction of the cylinder whereas a maximum wear on the uphill side results in a corrugation pattern moving in the same direction as the cylinder. When the maximum of the wear is located in the corrugation

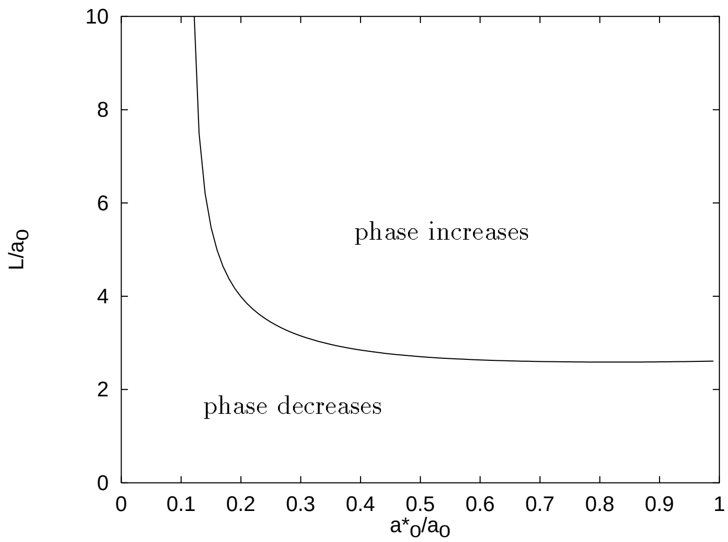

Figure 4.15. The limit between increasing and decreasing phase of the corrugation.

troughs or at the corrugation spikes the corrugation pattern will remain at the same position as the initial corrugation.

The phase of the surface is represented by Z_B/Z_A. From the discrete mapping in equation (4.103) a relation for the evolution of the phase is derived as

$$\frac{Z_B^{\{n+1\}}}{Z_A^{\{n+1\}}} = \frac{Z_B^{\{n\}}}{Z_A^{\{n\}}} - W_2 \left[1 + \left(\frac{Z_B^{\{n\}}}{Z_A^{\{n\}}} \right)^2 \right] \tag{4.111}$$

This expressions demonstrates that it is the sign of W_2 which is determining

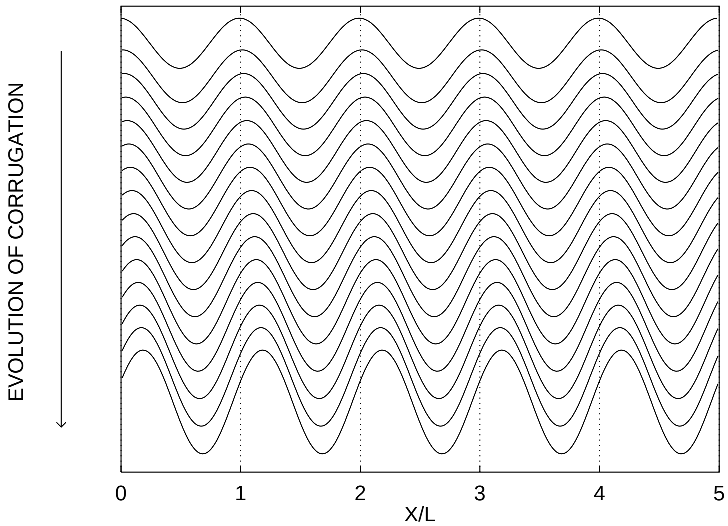

Figure 4.16. Corrugation moving along the surface.

the evolution of the phase of the corrugation pattern:

$$\begin{aligned} W_2 > 0 \quad &\Leftrightarrow \quad \text{the phase decreases} \\ W_2 = 0 \quad &\Leftrightarrow \quad \text{the phase is constant} \\ W_2 < 0 \quad &\Leftrightarrow \quad \text{the phase increases} \end{aligned} \tag{4.112}$$

In Figure 4.15 the values for which $W_2 = 0$ are plotted. For the values situated to the left or below of this line the phase of the corrugation will be decreasing while the values above or to the right of the line results in an increasing phase. It must be noticed that while the expression for the growth rate is exponential (equation (4.105)) the expression for the change of phase is additive. This implies that the phenomenon of moving corrugation often is suppressed by the growth of the amplitude. In the regions where W_1 is close to zero and thus the growth rate is very small the change of phase, however, will be quite distinct. An example of a moving corrugation is shown in Figure 4.16.

4.5 Contact Filters

The previous sections demonstrate that the size of the contact patch and the size of the stick zone relative to the wave length of the corrugation are crucial parameters in order to determine the qualitative behaviour of the wear. The fact that the size of the contact patch acts like a mechanical filter on the surface irregularities is quite obvious: the finite size of the contact patch implies that the cylinder is not able to sense irregularities with relative small wave lengths. Many contact theories try to take this property into account by applying a filter on the surface irregularities. One of the most frequently used filters is suggested by Remington [60] who has introduced a contact filter for a rectangular contact patch:

$$F_R = \frac{L}{2\pi a_0} \sin\left(\frac{2\pi a_0}{L}\right) \tag{4.113}$$

This filter is not directly applicable for all problems as it for certain a_0/L-values is negative. This implies e.g. for a wear problem that material is added to the surface instead of being removed as the cylinder passes by. To avoid this apparent non-physical behaviour Hempelmann [26] has introduced a modified Remington contact filter where negative values do not occur:

$$F_{RM} = \frac{1}{1 + 5.32\left(\frac{a_0}{L}\right)^2 + 6\left(\frac{a_0}{L}\right)^4 - 1.984\left(\frac{a_0}{L}\right)^6} \tag{4.114}$$

The two contact filters are monitored in Figure 4.17.

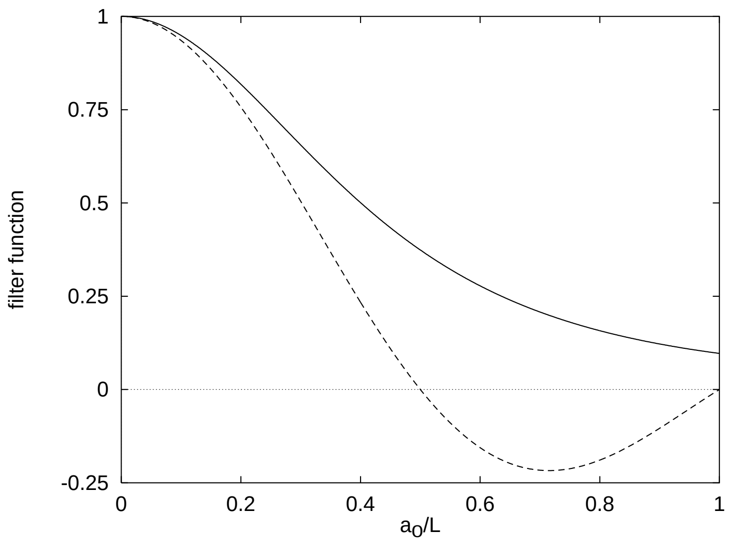

Figure 4.17. Contact filters. Dashed line: the Remington filter F_R. Solid line: the modified Remington filter F_{RM}.

4.5.1 The Remington Filter Compared with The Polynomial Approach

The problem with the Remington filter both in its original version and in the modified version is that it only depends on the ratio a_0/L. This is a critical simplification as the results from the previous sections demonstrate that also the ratio $r_{a_0} = a_0^*/a_0$ influences the characteristic behaviour of the problem. The qualitative discrepancies between solutions obtained by applying the Remington filter and the solutions derived with the polynomial approach can be illustrated by considering a modified version of the Carter solution.

Consider the problem of a cylinder with the radius R_0 rolling on the corrugated surface

$$Z_2(X) = Z_A \cos(kX) + Z_B \sin(kX) \tag{4.115}$$

When the local curvature of the surface is taken into account this yields that the equivalent radius of the contact problem can be approximated by

$$R_{eqv}(X) = R - (Rk)^2 \left[Z_A \cos(kX) + Z_B \sin(kX)\right] \tag{4.116}$$

Now apply the Remington filter to the surface irregularities and insert the equivalent radius into the Hertz solution and into the Carter solution, then the following characteristic parameters for the contact problem are derived

$$a_F(X) = a_0 - \frac{1}{2}k^2 a_0 R \left[Z_A \cos(kX) + Z_B \sin(kX)\right] F_R(ka_0) \tag{4.117}$$

$$\begin{aligned} a_F^*(X) = a_0^* + \frac{1}{2}k^2 a_0 R \left(1 - 2r_{a_0}\right) \cdot \\ \left[Z_A \cos(kX) + Z_B \sin(kX)\right] F_R(ka_0) \end{aligned} \tag{4.118}$$

$$\begin{aligned} T_F(X) = \mu N \left[1 - r_{a_0}^2 + k^2 R r_{a_0} \left(1 - r_{a_0}\right) \cdot \right. \\ \left. \left[Z_A \cos(kX) + Z_B \sin(kX)\right] F_R(ka_0)\right] \end{aligned} \tag{4.119}$$

The first thing to notice is that the above derived modified version of the Carter solution still results in a contact patch with a centre which is exactly located on the vertical projection of the cylinder axis i.e. $\Delta(X) = 0 \;\forall\; X$. It is possible to correct for this error by assuming that the centre of the contact patch is located at the contact point, but as demonstrated in section 4.2.2 this assumption is also erroneous.

When it comes to the size of the contact patch the solution with the filter is quite acceptable. The phase with respect to the surface is constant for both the filter solution and for the polynomial approach. For the amplitudes of the contact patch the calculations yield that the relative error committed by applying the Remington filter is

$$1 - \frac{\widehat{a}_F}{\widehat{a}} = \frac{\pi^2}{6}\left(\frac{a_0}{L}\right)^2 + \mathcal{O}\left(\left(\frac{a_0}{L}\right)^4\right) \tag{4.120}$$

It is seen that for small a_0/L ratios the filter solution provides a good accuracy, but the error grows quadraticly in a_0/L and so an a_0/L ratio of 0.25 results in an error of 10 %. In general it must however be concluded that the filter approach is satisfactory for the calculation of a, which is not surprising as the contact filter is a geometrical filter suited for a stationary contact situation i.e. for the normal contact problem.

For the tangential contact problem where the size of the stick zone enters into the solution as an important parameter, the accuracy of the filter solution decreases considerably. The contact filter is not able to handle the phase shift introduced by the oscillating behaviour of the contact patch. As a result both a_F^* and T_F have constant phase lag with respect to the corrugated surface. This is in sharp contrast to the solutions found by the polynomial approach illustrated by Figure 4.6 and Figure 4.8.

The amplitudes of a^* and a_F^* are shown in Figure 4.18. It is evident that there are very large discrepancies between the results obtained with the filter approach and the results from the polynomial approximation. First of all it must be noticed that the filter solution results in a symmetry around $r_{a_0} = 0.5$ i.e. that

$$\widehat{a}_F^*(r_{a_0}) = \widehat{a}_F^*(1 - r_{a_0}) \tag{4.121}$$

$$\widehat{a}_F^*(0.5) = 0 \quad \forall\ \frac{a_0}{L} \tag{4.122}$$

This symmetric behaviour is far from the results obtained with the polynomial approach. Another important result is that the contact filter yields

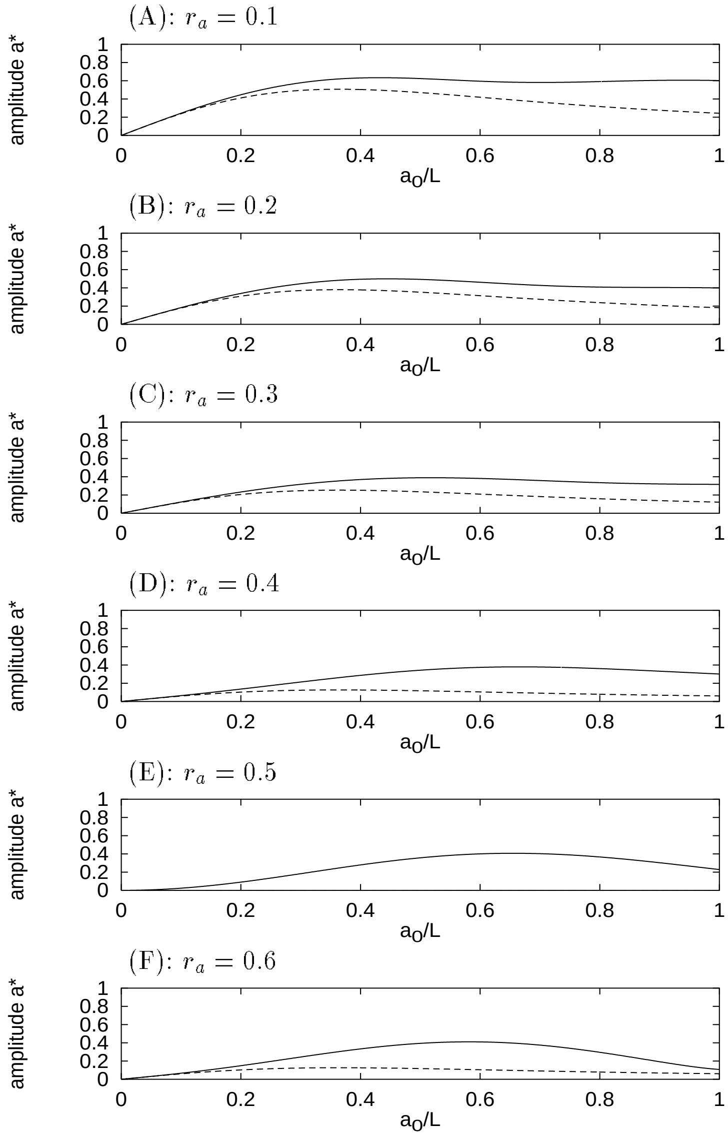

Figure 4.18. (A)-(F): Amplitude of half the size of the stick zone calculated with the polynomial approach (solid line) and calculated with the filter approach (dashed line).

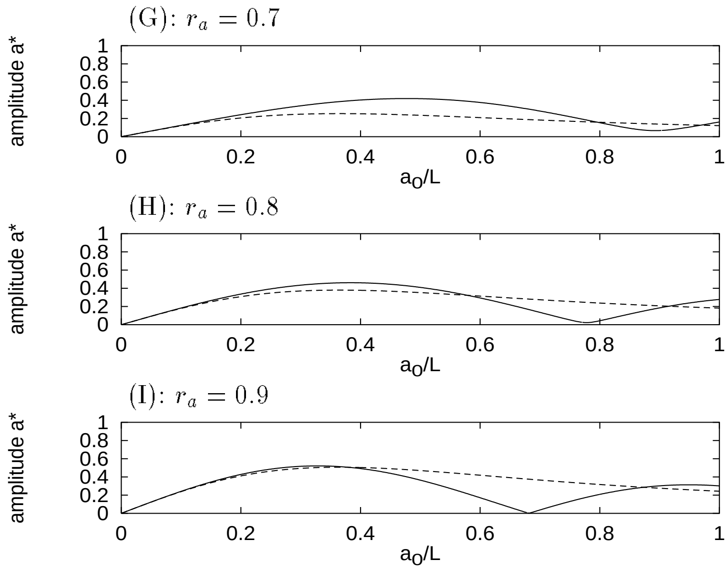

Figure 4.18. (G)-(I): Amplitude of half the size of the stick zone calculated with the polynomial approach (solid line) and calculated with the filter approach (dashed line).

the largest error for r_{a_0} values in the vicinity of 0.5. This is not at all an extreme contact situation and yet the error is considerable in the entire a_0/L interval. It must thus be concluded that the validity of the filter approach is very questionable for the calculation of a^* because the filter is indifferent with respect to r_{a_0}.

The same qualitative discrepancies are found for the amplitude of T (see Figure 4.19). The filter solution also provides an amplitude of T which is symmetric around $r_{a_0} = 0.5$ whereas the polynomial approach do not have this property. The error introduced by utilizing a contact filter for the calculation of $\widehat{T}$ is very large even for relative small values of a_0/L, which illustrates the inadequacy of the Remington contact filter for the tangential contact problem. Because the filter only depends on the a_0/L ratio the error introduced by the filter solution is quite large both with respect to the amplitudes and the phases of the quantities of the tangential contact problem.

The lack of accuracy of the filter solution results in qualitatively false results when they are applied to a wear calculation. In the previous section it was demonstrated how important the phase of the wear is for the evolution of a corrugation pattern. Because the wear is related to the tangential force and the size of the stick zone, errors in the calculation of these quantities will naturally be transmitted into the wear calculations

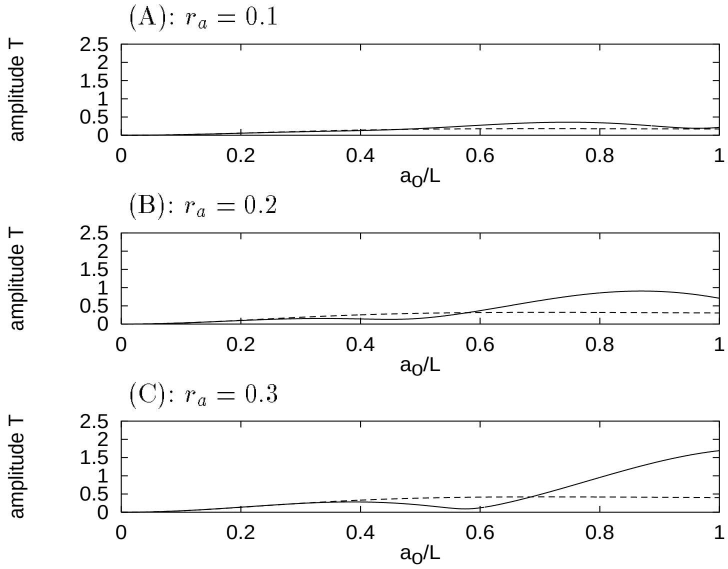

Figure 4.19. (A)-(C): Amplitude of the tangential force calculated with the polynomial approach (solid line) and calculated with the filter approach (dashed line).

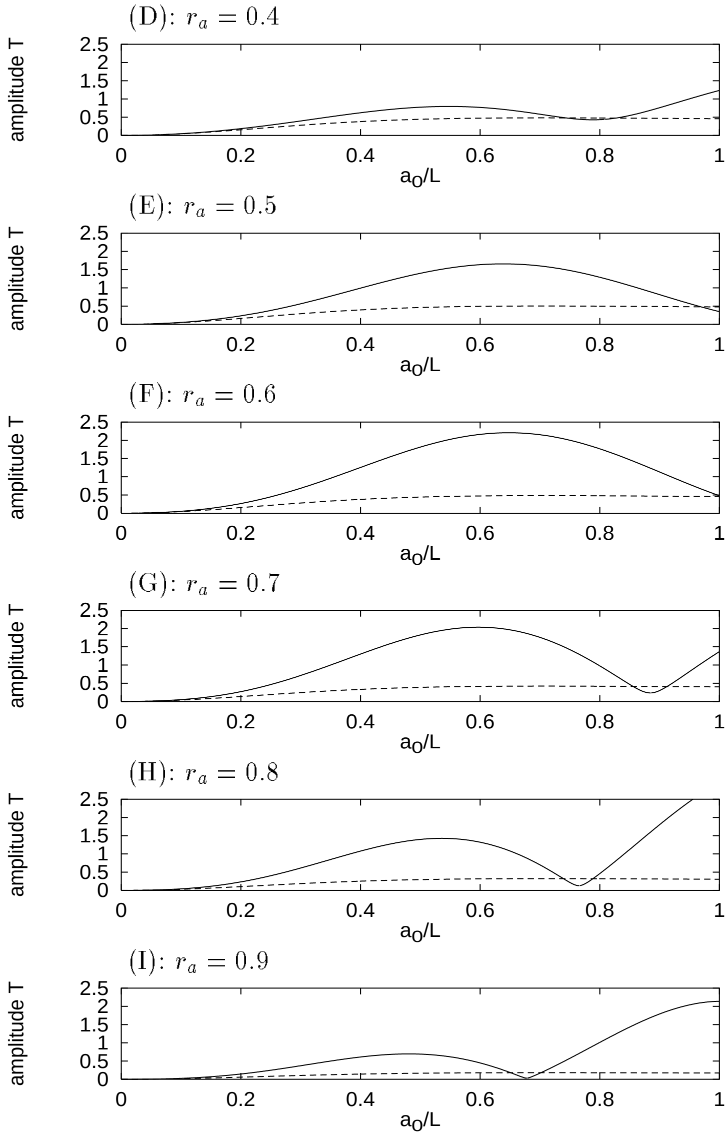

Figure 4.19. (D)-(I): Amplitude of the tangential force calculated with the polynomial approach (solid line) and calculated with the filter approach (dashed line).

An often used approach for wear calculations based on the Carter solution is to assume that the wear can be considered to be concentrated in the centre of the slip zone i.e. that if the cylinder axis is located at the position X_0 then the resulting wear for this instant contact situation is laid on the point $X_0 + a^*$. With the wear approximation from equation (4.75) this provides the following expression for the wear:

$$W_F(X) = |KT_F(X - a^*)\xi| \quad \Rightarrow$$

$$W_F(X) = W_0 + W_{F,A}\cos(kX) + W_{F,B}\sin(kX) \tag{4.123}$$

$$\left\{ \begin{array}{c} W_{F,A} \\ W_{F,B} \end{array} \right\} = \left[\begin{array}{cc} W_{F,1} & -W_{F,2} \\ W_{F,2} & W_{F,1} \end{array} \right] \left\{ \begin{array}{c} Z_A \\ Z_B \end{array} \right\} \tag{4.124}$$

where

$$W_{F,1} = -K\mu^2 Nk^2 a_0 r_{a_0}\left(1 - 2r_{a_0} + r_{a_0}^2\right)\cos(ka_0^*)F_R \tag{4.125}$$

$$W_{F,2} = -K\mu^2 Nk^2 a_0 r_{a_0}\left(1 - 2r_{a_0} + r_{a_0}^2\right)\sin(ka_0^*)F_R \tag{4.126}$$

It is seen that by assuming the wear to be concentrated in the centre of the slip zone a phase lag with respect to the initial corrugation is introduced. However, as demonstrated in the previous calculations the value of a_F^* and thus the location of the centre of the slip zone is erroneous. A consequence is that also the estimation of the phase of the wear is wrong. This property is illustrated in Figure 4.20 where the tangential force and the wear are shown as functions of the position on the corrugated surface.

The evaluation of the tangential force according to the filter solution is only slightly different from the solution obtained with the polynomial approach. This picture is however grossly disturbed when it comes to the evaluation of the wear. The discrepancy between the two amplitudes remains small, but

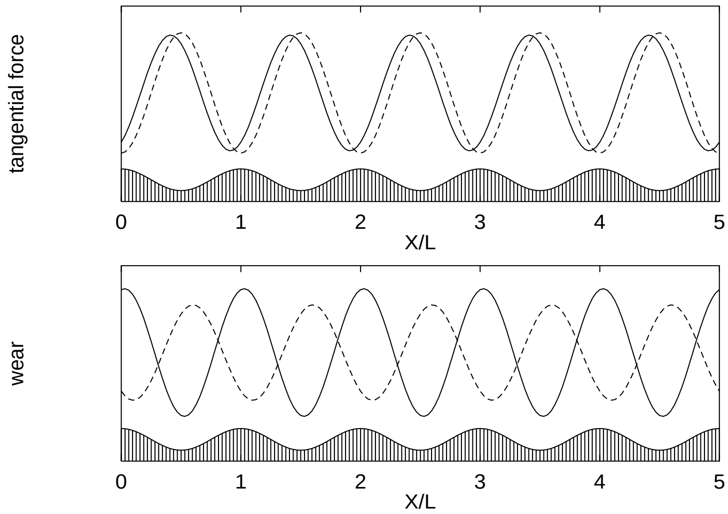

Figure 4.20. Comparison between polynomial approach (solid line) and filter solution (dashed line). Contact configuration: $L/a_0 = 3$ and $r_{a_0} = 0.3$.
Top: the tangential force acting on a given position on the surface.
Bottom: the wear acting on a given position on the surface.

the phase lags are completely different. In the present example the maximum wear according to the filter solution is located in the troughs of the corrugation pattern whereas the polynomial approach yields a maximum wear at the corrugation spikes. A consequence of the different phase lags is that a wear calculation made with the filter approach for the present case results in a growing corrugation while the polynomial approach demonstrates that the corrugation will be levelled out. The reason for this behaviour is related to the asymmetry of the stress distribution, which implies that the tangential stress distribution calculated with the polynomial approach has a tendency to move towards the top of the corrugation spike compared with the Carter solution (see Figure 4.7). Thus, when the locations of the

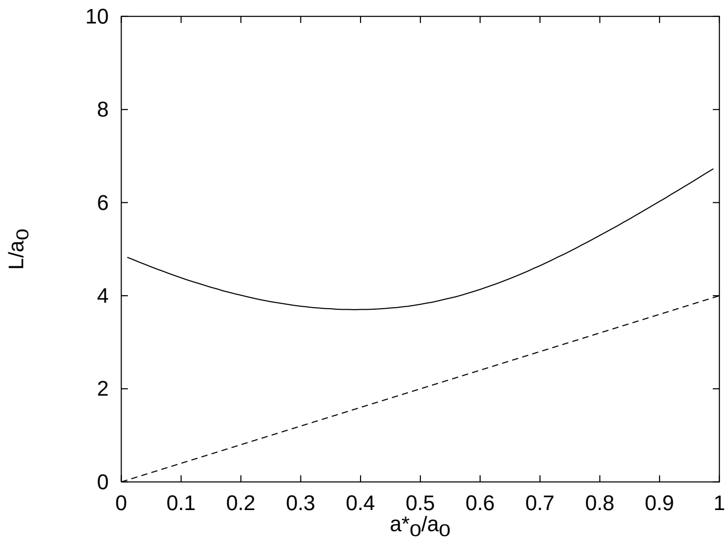

Figure 4.21. The limit between amplifying zone and levelling zone. Solid line: polynomial approach. Dashed line: solution with contact filter.

contact patches for the two solutions are almost identical the phases of the two calculated tangential forces do not differ very much, but the internal asymmetry of the tangential stress distribution in the polynomial approach leads to a considerable shift of the phase when it comes to the wear.

The stability of the corrugation calculated with the filter approximation is determined by the quantity $(1-2W_{F,1})$. Like in the case with the polynomial approach the limit of stability is given by the values for which $W_{F,1} = 0$, which implies that $\cos(ka_0^*) = 0$. Thus, the limit between the amplifying zone and the levelling zone is defined as

$$\frac{a_0}{L} = \frac{1}{4r_{a_0}} \tag{4.127}$$

It is noticed that this value actually does not depend on the applied contact filter. The differences in the location of the limit between the amplifying zone and the levelling zone for the two approaches are demonstrated in Figure 4.21. It is clearly seen that the two solutions differ qualitatively. Consequently the filter approach will predict a growth in the corrugation where the more advanced polynomial approach demonstrates that the given corrugation actually will be levelled out.

It must thus be concluded that for cases where the L/a_0 ratio is small the Remington filter is only suited for the normal contact problem. When it comes to the tangential contact problem the omitting of the size of the stick zone leads to qualitatively false results. If a contact filter is utilized for the tangential contact problem, it is therefore important that the size of the contact patch is included in the filter function in order to obtain a satisfactory wear model.

Chapter 5

Velocity Dependent Friction Coefficient

The objective of the present chapter is to investigate how a velocity dependent friction coefficient influences the tangential contact problem. Two types of friction coefficients will be treated: a friction coefficient defined as a step function with one static value and one kinematic value plus the case of a friction coefficient defined as a smooth function of the local relative velocity. As characteristic quantities for the tangential contact problem the outline of the creep curve and the shape of the tangential stress distribution are examined.

5.1 Friction Function

When an object slides over a surface a tangential force will be transmitted between the two bodies due to friction in the contact patch. A fundamental problem is now how the magnitude of this tangential force can be calculated. The classic friction law of Coulomb states that there exists a linear relation between normal force and tangential force given as

$$T = \mu N \tag{5.1}$$

where μ is the friction coefficient. Originally μ was presupposed to be a material dependent constant, but today it is accepted that this assumption for many contact situation is much too primitive.

When it comes to rolling friction the relation between the tangential force and the normal force is given by the creep curve. The shape of the creep curve found by experimental observations, however, often differs considerably from the shape derived by theoretical calculations. A common deviation is that where the calculated creep curve enters a saturated regime with constant tangential force for complete sliding, experiments indicate that the tangential force reaches a maximum and then decays as the creepage is increased [46]. This qualitative shape of the creep curve is important for railway dynamics, as a decreasing creep curve introduces a negative damping in the system, which may cause the system to loose its stability.

It is a common theory that the high frequency noise which occurs as a railway vehicle runs through a curve is closely related to the decaying creep curve [25]. Due to the negative damping of the system, the tangential force will start to oscillate, resulting in rapid variations in the location of the limit between the stick zone and the slip zone. The resulting curve-shrieking is therefore often referred to as a stick-slip phenomenon. The equivalence in the case of sliding friction is the friction oscillator where very interesting nonlinear dynamics can be found as a result of the slip-stick-behaviour (see e.g. [7] and [58]).

One way to explain a decaying creep curve is that the classic friction law of Coulomb is not valid for the given contact situations. Most contact models are based on the friction law of Coulomb where the friction coefficient is assumed to be constant but experiments indicate that the friction coefficient depends on the sliding velocity (e.g. [43] and [54]). The question is now how a velocity dependent friction coefficient influences the tangential contact problem.

Assume that the friction coefficient depends on the relative velocity between the bodies in contact. With the further assumption that the macroscopic friction law of Coulomb can be applied on the microscopic case i.e. that $q(x) = \mu p(x)$, then μ is depending on the local relative velocity between the bodies i.e. the slip:

$$\mu = f(s) \tag{5.2}$$

The function f is in the following referred to as the friction function.

The present work will not go into investigations of the outline of $f(s)$ but just assume it to be predefined. Work on the determination of the friction function for various contact situations can be found in e.g. [5], [44] and [59]. In the following sections the tangential contact problem will be solved for various configurations of $\mu = f(s)$.

5.2 A Friction Coefficient Defined as a Step Function

A simple outline of a velocity dependent friction function is obtained by assuming the friction coefficient to be a step function. In 1989 Ohyama [53] suggested a model describing the two-dimensional contact problem where the friction coefficient has one static value and one kinematic value, i.e.

$$\mu(s) = \begin{cases} \mu_0 & , \quad s = 0 \\ \mu_k & , \quad s \neq 0 \end{cases} \tag{5.3}$$

The solution found by Ohyama states that the tangential stress distribution in the stick zone is identical to the Carter solution for $\mu = \mu_0$ whereas the tangential stress distribution in the slip zone is equivalent to the Carter solution with $\mu = \mu_k$. The tangential stress distribution according to Ohyama is shown in Figure 5.1 for the case where $\mu_k < \mu_0$ and for $\mu_k > \mu_0$.

In Figure 5.2 the creep curves corresponding to the stress distributions monitored in Figure 5.1 are shown. The creep curve for the Ohyama solution where $\mu_k < \mu_0$ has a distinct maximum and then decays until it has reached the regime of complete sliding where $T = \mu_k N$. This behaviour fits to some extend with experimental data [54], only the decaying shape for the Ohyama solution is restricted to a limited range of the creepage whereas the experiments indicate that the slope of the creep curve remains negative also when complete sliding occurs.

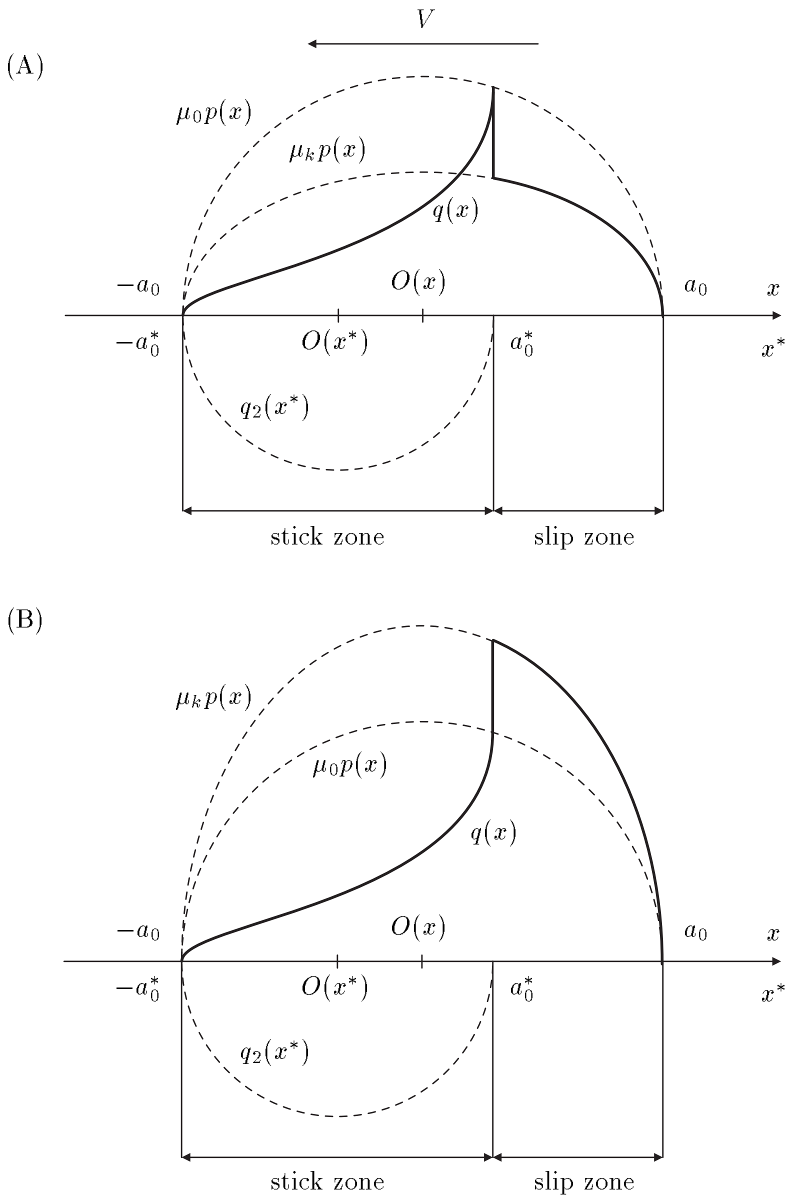

Figure 5.1. Tangential stress distributions according to Ohyama [53]. Top (A): $\mu_k = 0.6\mu_0$. Bottom (B): $\mu_k = 1.4\mu_0$.

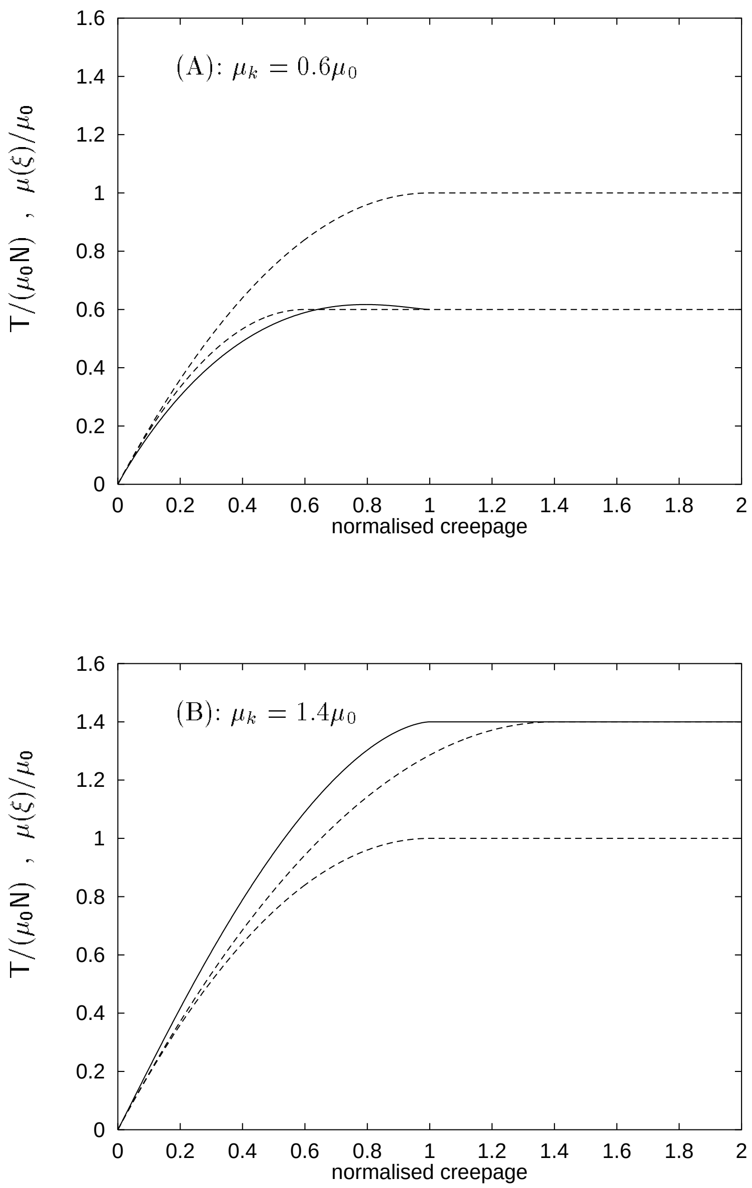

Figure 5.2. Creep curves according to Ohyama [53]. Solid line: Ohyama solution. Dashed lines: Carter solution for $\mu = \mu_0$ and $\mu = \mu_k$.

The problem with the Ohyama solution is that when $\mu_k < \mu_0$ it is not a minimum solution to the tangential contact problem. If the kinematic friction coefficient is smaller than the static friction coefficient the solution is therefore not valid i.e. the tangential stress distribution from Figure 5.1 (A) and the creep curve from Figure 5.2 (A) are wrong.

In order to solve the tangential contact problem when $\mu_k < \mu_0$ the polynomial approach is employed with the conditions that

$$|q(x)| < \mu_0 p(x) \quad , \quad -a < x < 2a^* - a \tag{5.4}$$

$$q(x) = \mu_0 p(x) \quad , \quad x = 2a^* - a \tag{5.5}$$

$$q(x) = \mu_k p(x) \quad , \quad 2a^* - a < x < a \tag{5.6}$$

It can then easily be shown that the solution to this tangential contact problem actually is equivalent to the Carter solution for $\mu = \mu_k$ with the only modification that a Kronecker peak occurs at the limit between stick zone and slip zone (see Figure 5.3). Because the integral of this peak is zero, it will not contribute to the tangential force and so the creep curve for the case when the kinematic friction coefficient is smaller than the static friction coefficient is equal to the creep curve for the Carter solution where $\mu = \mu_k$.

Consequently a friction function defined as a step function introduces no decay in the creep curve, and so the stability of the rolling contact problem in this case apparently is unaffected by the varying friction coefficient. It must thus be concluded that classic slip-stick oscillations will not occur in rolling contact when the friction coefficient is defined as a step-function.

Another problem of the Ohyama solution is related to the calculation of the creepage which is inaccurate. An obvious indicator of this fact is seen

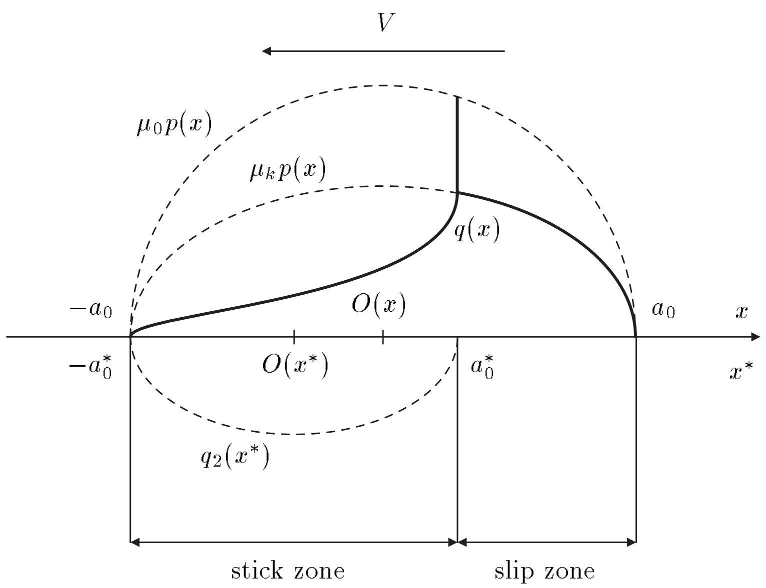

Figure 5.3. Actual tangential stress distribution with one static friction coefficient and one kinematic friction coefficient when $\mu_k < \mu$.

from the creep curves in Figure 5.2. When the friction coefficient is defined as a function lying in the interval from μ_0 to μ_k the resulting creep curve should lie between the equivalent creep curves for the Carter solution. This is clearly not the case for the creep curves calculated with the Ohyama solution. Like the actual creep curve for $\mu_k < \mu_0$ is identical with the Carter creep curve for $\mu = \mu_k$ the equivalent result is obtained when $\mu_k > \mu_0$. Thus, the creep curve related to the tangential stress distribution indicated in Figure 5.1 (B) is actually equivalent to the upper Carter creep curve in Figure 5.2 (B).

5.3 A Friction Coefficient Defined as a Smooth Function

Now assume that the friction coefficient is a smooth function of the slip. The problem of solving the tangential contact problem then arises from the fact that the slip and thus the friction coefficient varies over the slip zone. As the tangential stress distribution in the slip zone is defined as $q(x) = \mu(s)p(x)$ the fact that $s = s(x)$ leads to a quite complicated tangential stress distribution.

Because the friction coefficient is assumed to be a smooth function of $s(x)$ it is possible to write it as a polynomial in x:

$$\mu(x) = \sum_{m=0}^{M} \bar{\mu}_m x^m \quad , \quad x \in S_{slip} \tag{5.7}$$

With the usual assumption of one stick zone and one slip zone the tangential stress distribution can be expressed as

$$q(x) = q_1(x) + q_2(x^*) \tag{5.8}$$

where $q_1(x) = \mu(x)p(x)$ in the slip zone. The normal pressure distribution is in the present calculations defined to be Hertzian, but the calculations can be made for an arbitrary normal pressure distribution as long as it is expressed as a polynomial form. The contribution from $q_1(x)$ in the slip zone is thus

$$\begin{aligned} q_1(x) &= \frac{p_0}{a_0}\sqrt{a_0^2 - x^2} \sum_{m=0}^{M} \bar{\mu}_m x^m \quad \Rightarrow \\ q_1(x) &= \frac{\pi E}{4(1-\nu^2)} \frac{\sum_{m=0}^{M+2} B_m x^m}{\sqrt{a_0^2 - x^2}} \end{aligned} \tag{5.9}$$

where the coefficients are given as

$$B_0 = \frac{a_0^2}{\pi R}\bar{\mu}_0 \tag{5.10}$$

$$B_1 = \frac{a_0^2}{\pi R}\bar{\mu}_1 \tag{5.11}$$

$$B_2 = \frac{1}{\pi R}\left(\bar{\mu}_2 a_0^2 - \bar{\mu}_0\right) \tag{5.12}$$

$$\vdots$$

$$B_M = \frac{1}{\pi R}\left(\bar{\mu}_M a_0^2 - \bar{\mu}_{M-2}\right) \tag{5.13}$$

$$B_{M+1} = -\frac{1}{\pi R}\bar{\mu}_{M-1} \tag{5.14}$$

$$B_{M+2} = -\frac{1}{\pi R}\bar{\mu}_M \tag{5.15}$$

With $q_1(x)$ expressed as a polynomial form $q_2(x^*)$ can be found using the polynomial approach.

The problem is however that $s(x)$ and thus $\mu(x)$ depend on both $q_1(x)$ and $q_2(x^*)$, which implies that $\mu(x)$ must be recalculated by which new expressions for $q_1(x)$ and $q_2(x^*)$ are found. This iterative process converges however after a few steps and so the tangential contact problem for a velocity dependent friction coefficient is solved.

To illustrate how the solution to the tangential contact problem changes when the friction coefficient is assumed to be velocity dependent, two examples are investigated. Let the friction coefficient be defined as

$$\mu(s) = \frac{\mu_0 - \mu_k}{1 + |\kappa s|} + \mu_k \tag{5.16}$$

then one example is calculated with $\mu_k = 1.4\mu_0$ whereas the other example is for the case where $\mu_k = 0.6\mu_0$. The coefficient κ determines the size of the initial slope of the friction function and can be chosen arbitrarily. With the above definition of the friction function the kinematic friction coefficient μ_k is given as

$$\mu_k = \lim_{s \to \infty} \mu(s) \tag{5.17}$$

In Figure 5.4 the friction functions and the creep curves are plotted for the two cases. It is noticed that the initial slopes of the creep curves are unaffected by the velocity dependent friction coefficient and is thus equal to the slope of the creep curves for the Carter solutions. As the magnitude of the creepage increases the tangential force tends towards the value $\mu(\xi)N$. It must be emphasized that the shown creep curves are not equal to the creep curve for a Carter solution where the friction coefficient is defined as $\mu = \mu(\xi)$. Because the creepage is the global relative velocity whereas the slip is the local relative velocity the two solutions will only be identical when $\xi = 0$ or when $\xi \to \infty$.

For the case where $\mu_k < \mu_0$ (see Figure 5.4 (A)) the creep curve has a maximum and then decays. The location of the maximum depends on the ratio μ_k/μ_0 and the initial slope of the friction function given by κ. Larger initial slopes will imply that the maximum occurs for smaller values of the creepage, which also is the case if the ratio μ_k/μ_0 decreases. In all cases the maximum occurs before complete sliding takes place.

Examples of the tangential stress distribution for the two cases are shown in Figure 5.5. It is seen that the magnitude of the tangential stress always lies between the Carter solution for respectively $\mu = \mu_0$ and $\mu = \mu_k$. At the limit between stick zone and slip zone the tangential stress distribution has a vertical tangent just like in the Carter solution. Crossing this limit the tangential stress distribution will continue to grow for a while when $\mu_k > \mu_0$ whereas it decays immediately when $\mu_k < \mu_0$.

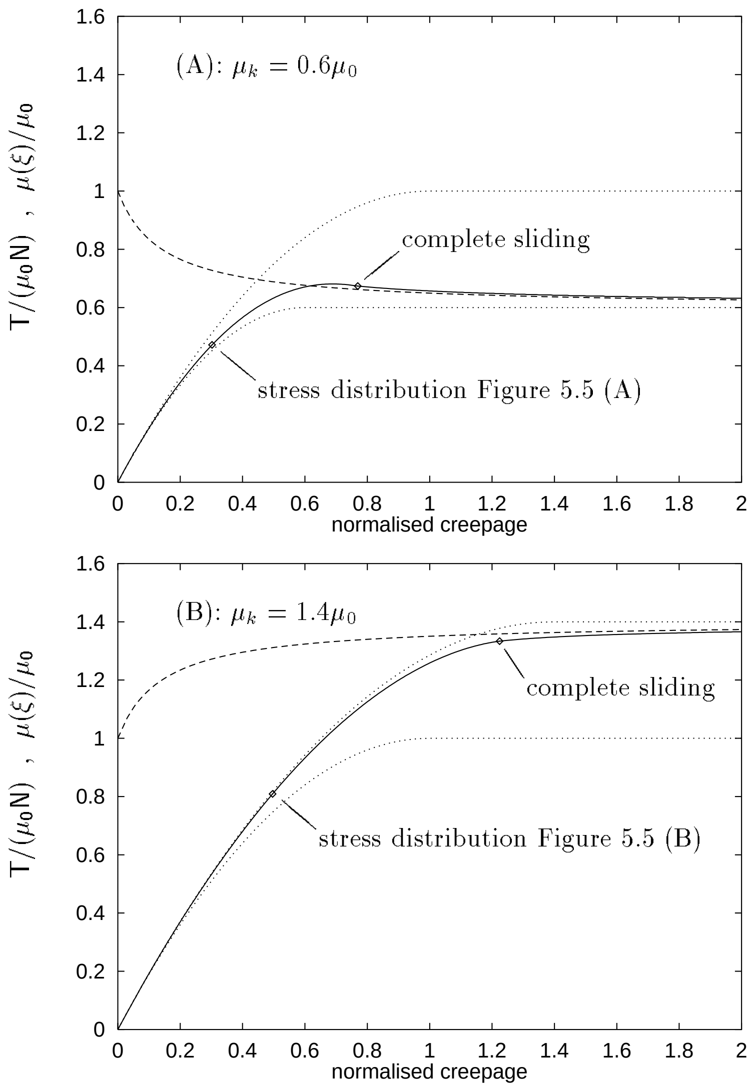

Figure 5.4. Creep curves when the friction coefficient is velocity dependent. Solid line: creep curve. Dashed line: $\mu(\xi)/\mu_0$. Dotted lines: creep curve for the Carter solution with $\mu = \mu_0$ and $\mu = \mu_k$.

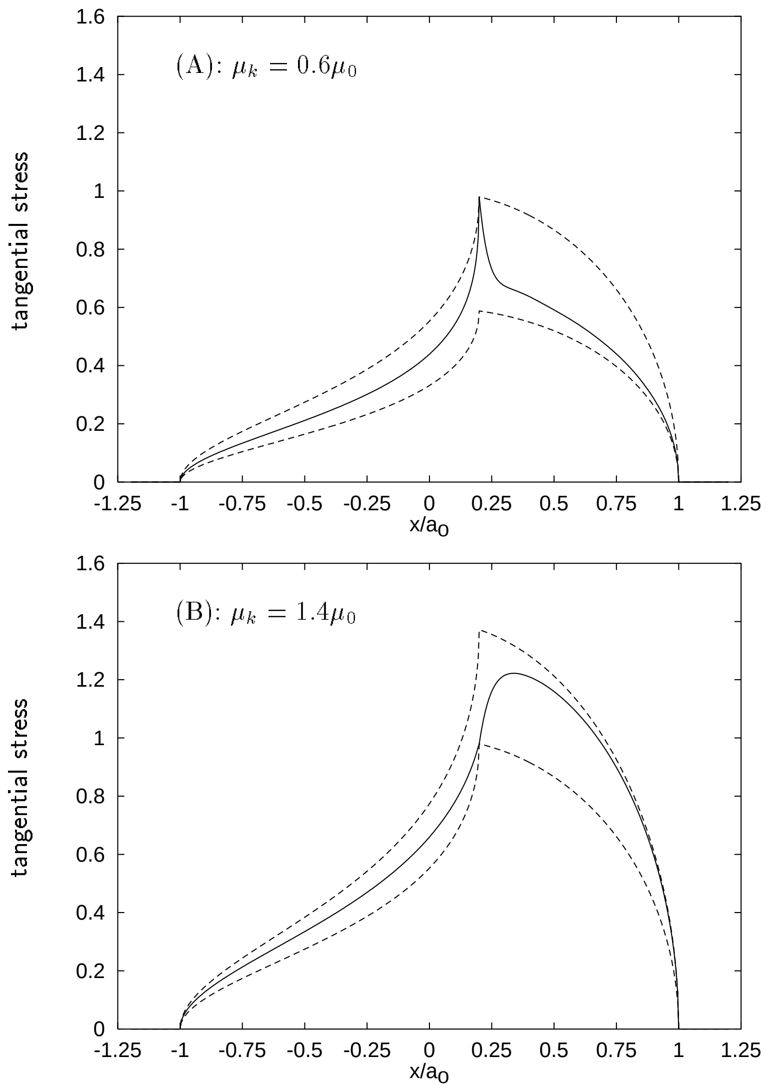

Figure 5.5. Tangential stress distributions when the friction coefficient is velocity dependent. Solid line: tangential stress. Dashed lines: tangential stress for the Carter solution with $\mu = \mu_0$ and $\mu = \mu_k$.

As the initial slope of the friction function increases the tangential stress distribution in the slip zone will approach the Carter solution for $\mu = \mu_k$ more rapidly. When κ and thus the slope tends towards infinity the tangential stress distribution will have the outline due to Ohyama for $\mu_k > \mu_0$ whereas the tangential stress distribution for the case $\mu_k < \mu_0$ will be equivalent to the one indicated in Figure 5.3.

It has in the present chapter been demonstrated that a decaying friction function results in a decaying creep curve. As an isolated result this is not at all unexpected. The power of the calculations lies however in the fact that when the friction function is known then it is possible to determine for what size of the creepage the maximum of the creep curve is located. This is a critical value of the creepage for which instability of the system is likely to occur.

Chapter 6

Rough Surfaces

In the contact problems investigated so far it has been assumed that the bodies are smooth. The present chapter serves as an introduction to the topic of contact problems involving rough surfaces. First the normal contact problem is solved for one isolated roughness asperity and afterwards the cross influence between isolated contact patches is investigated. Finally the tangential problem for rough surfaces in contact will be investigated briefly.

6.1 Contact between Rough Surfaces

To make the assumption that the bodies in contact are smooth is naturally a simplification. Any surface, manufactured or worn, will always have local asperities and troughs - it is just a matter of scaling. Consequently an extended model of wheel/rail interaction must also include the case where wheel and rail are considered to be rough. Whereas numerous papers on normal contact between rough surfaces have been published, e.g. [8] and [21], the amount of work on rolling contact between bodies with rough surfaces is quite limited [42].

When two rough spheres are pressed together a contact patch is created equivalently to the contact of smooth spheres. Two characteristic properties make the rough problem differ from the smooth problem. First of all the presence of asperities will result in an incoherent contact patch. Let $2a_r$ be the area of the contact patch for the rough surface contact and let $2a_0$ be the contact area for the equivalent Hertzian contact then the relative size of the contact area a_r/a_0 is an important parameter for the rough contact. A second effect related to the asperities is that the normal pressure distribution $p_r(x)$ locally will reach values much higher than the Hertzian value. In Figure 6.1 two examples of the two-dimensional normal contact problem for rough bodies are shown. The results are calculated by Knothe and Theiler [42] and are compared with the equivalent Hertzian solution, where the roughness of the bodies is omitted.

As indicated in Figure 6.1, the solution to the normal contact problem strongly depends on the size of the roughness wave length. Thus, one of the basic problems of rough contact is: what is the minimum wave length of the roughness which should be included in the model? As a given surface in theory can be represented by arbitrary small wave lengths it is necessary to define a cut-off wave length L_r which indicates the smallest wave length in the representation of the surface. It can be demonstrated [42] that a_r/a_0 depends heavily on L_r: smaller cut-off wave lengths result in smaller a_r/a_0-values.

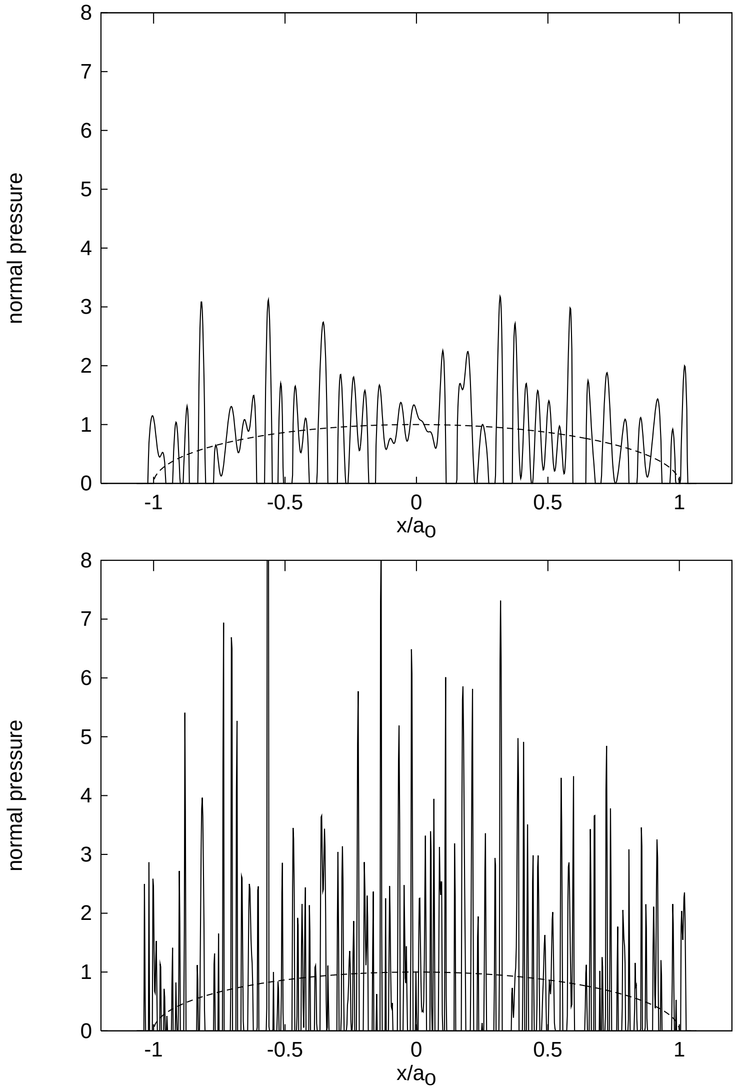

Figure 6.1. Normal pressure distribution for a two-dimensional contact between rough surfaces. Top: $a_0/L_r = 23.4$. Bottom: $a_0/L_r = 69.7$. Solid line: the actual normal pressure. Dashed line: the equivalent Hertzian normal pressure. The calculations are made by Knothe and Theiler [42].

From Figure 6.1 it is furthermore seen that the maximum of the normal pressure distribution increases as the cut-off wave length decreases. When L_r is very small and thus the maximum normal pressure is very large the utilization of the half space approach is questionable because the assumptions of fully elastic material and small strains are violated. Similarly very small a_r/a_0 values will indicate that the size of the local contact patch is of the same magnitude as the characteristic size of the local roughness asperity. In these cases the half space approach is no longer valid and other methods must be applied. In the present chapter it is however assumed that the contact between the local asperities can be evaluated as a half space contact i.e. the cut-off wave length L_r is sufficiently large.

6.2 Characterization of Roughness

Calculations including the contact between rough surfaces are often based on roughness measurements of real surfaces. Such measurements result in a large amount of data which makes the calculation of the contact mechanics very difficult. Thus, it is convenient to be able to describe a given rough surface with few parameters which can be used as input in a contact model. In the previous section it was demonstrated that the cut-off wave length is a critical parameter for the contact situation, but it is an open question what other properties are relevant for the contact situation. In other words: how is it possible to describe a rough surface with few parameters and still be sure that the description is unique in a contact mechanical context?

Much work has been carried out in this field without any decisive conclusions being made. A list of different methods to characterize a rough surface can be found in the book by Thomas [63]. Greenwood and Williamson [20] advocate for a description based on mean values and standard deviations of the asperity-height and asperity-curvature and then express the surface as a Gaussian distribution. This approach is quite erroneous as the height distribution in reality is not symmetric. Due to the wear of the surface the asperities of the surface are flattened, i.e. the curvatures of the asperities are smaller than the curvatures of the troughs. It is of course possible to correct for this property by applying an asymmetric representation of the data, but the basic problem remains: will surfaces with the same mean values and standard deviations for the asperity-height and asperity-curvature yield the same contact mechanical properties? The remainder of the present chapter is devoted to the investigations of different surface properties in the search for crucial surface parameters.

6.3 The Normal Contact Problem for one Asperity

In order to make the problem as simple as possible the first investigations consider only one surface asperity $Z(\bar{x})$ in contact with a nominally flat and smooth surface. The local coordinate system is defined such that the contact point is located in $\bar{x} = 0$. From section 2.4.1 it is known that the geometrical properties for two bodies in contact can be moved from one body to the other without loss of accuracy i.e. the case of two rough surfaces can be transformed into the case of one rough surface and one smooth surface in contact. Now let the roughness asperity be described as

$$Z(\bar{x}) = \lambda_0 + \frac{1}{2}\lambda_2\bar{x}^2 + \frac{1}{6}\lambda_3\bar{x}^3 + \frac{1}{24}\lambda_4\bar{x}^4 \tag{6.1}$$

which implies that the geometric moments of the asperity are given as

$$\lambda_i = \left. \frac{d^{(i)} Z(\bar{x})}{d\bar{x}^{(i)}} \right|_{\bar{x}=0} \qquad i = 0, 2, 3, 4 \tag{6.2}$$

where λ_1 is not present due to the choice of origo and where λ_2 is the curvature, λ_3 is the skewness and λ_4 is the flatness of the asperity. The values of the λ's are chosen such that the asperity is convex in the vicinity of the contact point, which ensures that the contact patch is coherent and that there is only one contact point. It is further assumed that the fourth order approximation of the asperity provides an satisfactory accuracy i.e. that higher order terms of the series expansion can be neglected.

Employing the polynomial approach this leads to a contact patch with the centre located the distance Δ from $O(\bar{x})$ and with half the contact width a. Introducing the new coordinate system $x = \bar{x} - \Delta$ the normal pressure distribution over the contact patch is

$$p(x) = \frac{\pi E}{4(1-\nu^2)} \frac{\sum_{n=0}^{4} B_n x^n}{\sqrt{a^2 - x^2}} \tag{6.3}$$

$$B_0 = \frac{a^2}{\pi} \left[\lambda_2 + \lambda_3 \Delta + \frac{1}{2} \lambda_4 \Delta^2 + \frac{1}{12} \lambda_4 a^2 \right] \tag{6.4}$$

$$B_1 = \frac{1}{\pi} \left[\frac{1}{4} a^2 \left(\lambda_3 + \lambda_4 \Delta \right) - \Delta \left(\lambda_2 + \frac{1}{2} \lambda_3 \Delta + \frac{1}{6} \lambda_4 \Delta^2 \right) \right] \tag{6.5}$$

$$B_2 = \frac{1}{\pi} \left[\frac{1}{12} a^2 \lambda_4 - \lambda_2 - \lambda_3 \Delta - \frac{1}{2} \lambda_4 \Delta^2 \right] \tag{6.6}$$

$$B_3 = -\frac{1}{2\pi} \left[\lambda_3 + \lambda_4 \Delta \right] \tag{6.7}$$

$$B_4 = -\frac{1}{6\pi} \lambda_4 \tag{6.8}$$

where the two unknowns a and Δ are derived from the boundary conditions plus the constraint that the integral of the normal pressure distribution is equal to the normal force:

$$0 = \Delta\left(\lambda_2 + \frac{1}{2}\lambda_3\Delta + \frac{1}{6}\lambda_4\Delta^2\right) + \frac{1}{4}a^2\left(\lambda_3 + \lambda_4\Delta\right) \tag{6.9}$$

$$N = \frac{\pi E}{8\left(1-\nu^2\right)}\left\{\left(\lambda_2 + \lambda_3\Delta + \frac{1}{2}\lambda_4\Delta^2\right)a^2 + \frac{1}{8}\lambda_4 a^4\right\} \tag{6.10}$$

The values for a and Δ for various combinations of the λ's are given in Table 6.1 where $\tilde{N}$ is the normal force normalised with $(\pi E)/8(1-\nu^2)$.

$\lambda_4 = 0$	$\lambda_3 \neq 0$	$\lambda_2 > 0$	$a = \left(\frac{2}{3}\frac{\lambda_2}{\lambda_3}\left(2\cos\left(\Theta\right) - 1\right)\right)^{\frac{1}{2}}$ $\Theta = \frac{1}{3}\arccos\left(\frac{27}{8}\frac{\tilde{N}^2\lambda_3^4}{\lambda_2^6} - 1\right)$ $\Delta = \frac{1}{\lambda_3}\left(\frac{\tilde{N}}{a^2} - \lambda_2\right)$ restriction: $\frac{27}{16}\frac{\tilde{N}^2\lambda_3^4}{\lambda_2^6} < 1$
		$\lambda_2 = 0$	asperity is not convex
	$\lambda_3 = 0$	$\lambda_2 > 0$	$a = \left(\frac{\tilde{N}}{\lambda_2}\right)^{\frac{1}{2}}$ $\Delta = 0$
		$\lambda_2 = 0$	asperity is not convex

Table 6.1. (A): $\lambda_4 = 0$. Characteristics of the contact patch for various asperities.

$\lambda_4 \neq 0$	$\lambda_3 \neq 0$	$\lambda_2 > 0$	if $\lambda_3 + \lambda_4 \Delta = 0$: $a = \left(2\frac{\lambda_2}{\lambda_4}\left(\sqrt{1 + 2\frac{\lambda_4}{\lambda_2^2}\tilde{N}} - 1\right)\right)^{\frac{1}{2}}$ $\Delta = -\frac{\lambda_3}{\lambda_4}$ restriction: $\lambda_2 = \frac{\lambda_3^2}{3\lambda_4}$
			if $\lambda_3 + \lambda_4 \Delta \neq 0$: (no analytical solution) $a = \left(\frac{-4\Delta}{\lambda_3 + \lambda_4 \Delta}\left(\lambda_2 + \frac{1}{2}\lambda_3 \Delta + \frac{1}{6}\lambda_4 \Delta^2\right)\right)^{\frac{1}{2}}$ $\tilde{N} = \left(\lambda_2 + \lambda_3 \Delta + \frac{1}{2}\lambda_4 \Delta^2\right) a^2 + \frac{1}{8}\lambda_4 a^4$
		$\lambda_2 = 0$	asperity is not convex
	$\lambda_3 = 0$	$\lambda_2 > 0$	$a = \left(4\frac{\lambda_2}{\lambda_4}\left(\sqrt{1 + \frac{1}{2}\frac{\lambda_4}{\lambda_2^2}\tilde{N}} - 1\right)\right)^{\frac{1}{2}}$ $\Delta = 0$ restriction: $\frac{1}{2}\frac{\lambda_4}{\lambda_2^2}\tilde{N} > -1$
		$\lambda_2 = 0$	$a = \left(8\frac{\tilde{N}}{\lambda_4}\right)^{\frac{1}{4}}$ $\Delta = 0$ restriction: $\lambda_4 > 0$

Table 6.1. (B): $\lambda_4 \neq 0$. Characteristics of the contact patch for various asperities.

All the combinations except for one can be solved analytically. Only when it comes to an asperity where $\lambda_2, \lambda_3, \lambda_4 \neq 0$ the solution must be found by employing an approximation method e.g. series expansions or numerical iterations. The restrictions indicated in the table provides the parameter combinations for which the assumptions of coherent contact patch and single contact point are fulfilled.

The results from Table 6.1 indicate that the relation between the size of half the contact patch and the λ's is strongly nonlinear. A consequence of this nonlinearity is that the principle of superposition is not valid. In Figure 6.2 four different asperities are shown: $\mathcal{A}_i\ ,\ i = 1, 2, 3, 4$. The configuration of the asperities implies that each asperity can be expressed as a combination of the others

$$\mathcal{A}_1 - \mathcal{A}_2 - \mathcal{A}_3 + \mathcal{A}_4 = 0 \tag{6.11}$$

i.e. there is a linear dependency between the four asperities. It is seen that the discrepancies between the geometry of the asperities are minor in the vicinity of the contact point.

Now the four different asperities are pressed down on a smooth, level surface under the application of the normal load $\tilde{N}$. This way a contact patch with a normal pressure distribution is generated (see equation (6.3)). The pressure distributions for the asperities are shown in Figure 6.3. It is evident that the principle of superposition does not hold. The linear dependency of the geometry of the asperities cannot be retrieved in their normal pressure distributions, i.e. it is not possible to create the normal pressure for one asperity by combining the stress distributions for the other asperities. This implication demonstrates clearly that a statistical representation of the asperities do not provide a unique characterization of the surface in a contact mechanical context. A surface consisting of n asperities of type $\mathcal{A}_1$ and n

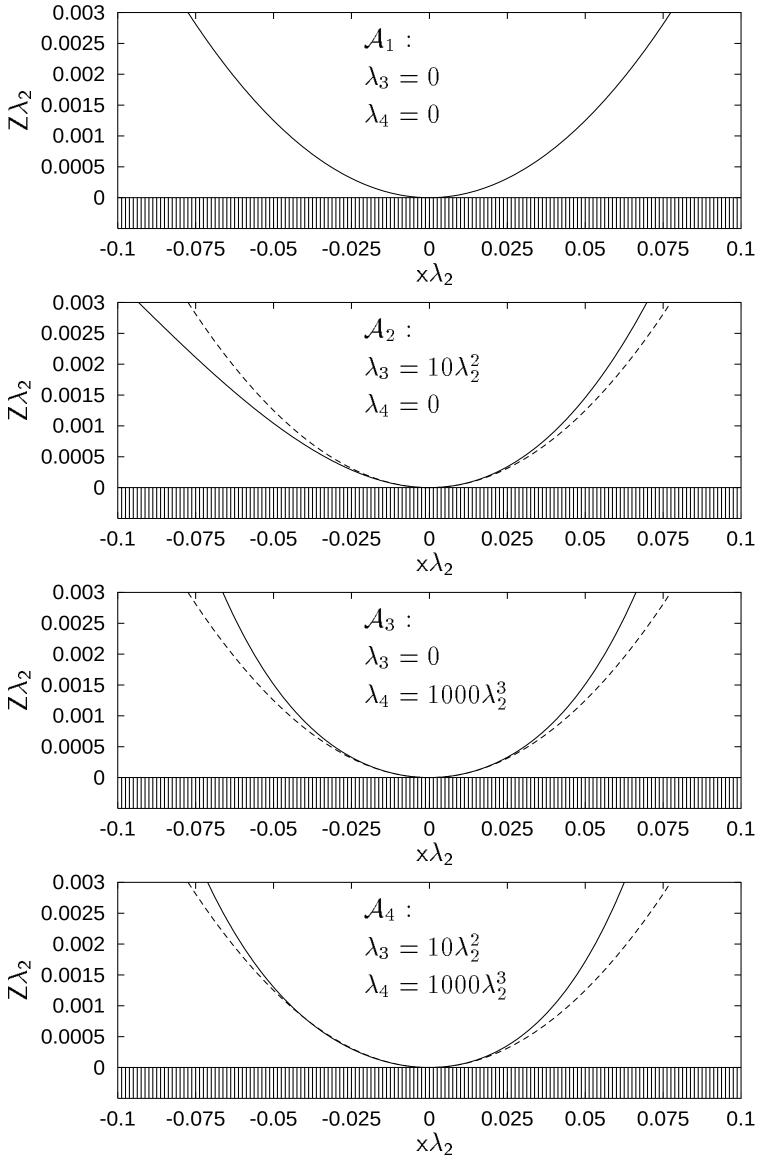

Figure 6.2. Different asperities. The dashed line indicates $\mathcal{A}_1$.

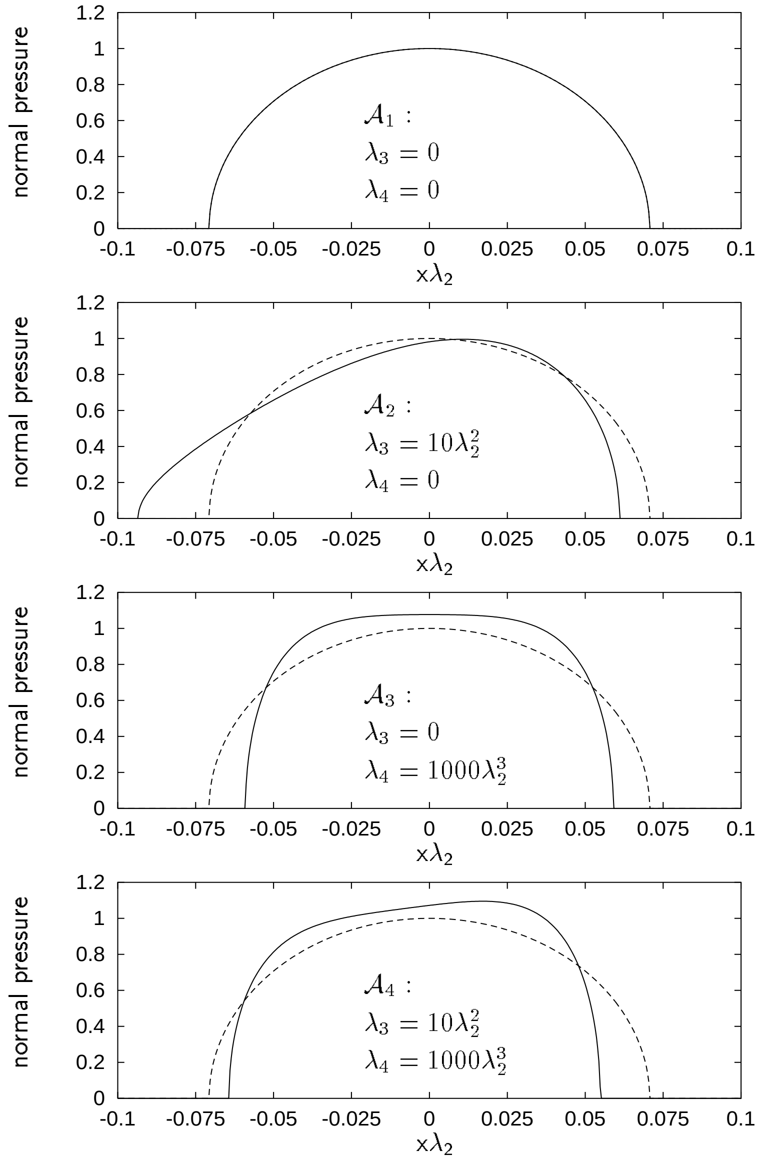

Figure 6.3. Normal pressure distribution. The dashed line indicates $\mathcal{A}_1$.

asperities of type $\mathcal{A}_4$ is statistical equal to a surface with n asperities of type $\mathcal{A}_2$ and n asperities of type $\mathcal{A}_3$, but the resulting size of half the contact patch a_r or the magnitude of the normal pressure p_r are not identical when the two surfaces are pressed down on a level surface.

6.4 Cross Influence between Adjacent Contact Patches

In the previous section only isolated asperities were examined. As a rough surface consists of many adjacent asperities it is of interest to know how the normal contact problem for a given asperity is influenced by the normal contact problem of the adjacent asperities. The cross influence between more contact patches arises from the relative displacement of material due to the deformation in the contact patch. Where the normal pressure is zero outside the contact patch, the relative displacements will in principle affect the entire surface of the bodies in contact. In order to simplify the investigations only the case of two neighbouring contact patches is analysed, but the methodology can easily be applied for more contact patches.

Consider the surface $Z(X)$ being pressed down on a level surface. Provided that $Z(X)$ has two minima inside the potential region of contact, there will be two contact points. If the normal load furthermore is sufficiently small or if the trough between the local minima is sufficiently deep, a contact situation with two separate contact patches occurs as indicated in Figure 6.4. Now assume that $Z(X)$ can be expressed as a polynomial in the vicinity of both contact points, then the polynomial approach states that the pressure distributions will be polynomial forms:

$$p_1(x_1) = \frac{\pi E}{4(1-\nu^2)} \frac{\sum_{m=0}^{M} B_m x_1^m}{\sqrt{a_1^2 - x_1^2}} \tag{6.12}$$

$$p_2(x_2) = \frac{\pi E}{4(1-\nu^2)} \frac{\sum_{m=0}^{M} G_m x_2^m}{\sqrt{a_2^2 - x_2^2}} \tag{6.13}$$

where a_1 and a_2 refer to half the size of the two contact patches and where the local coordinate systems x_1 and x_2 are defined with respect to the global coordinate system as

$$x_1 = X - d_1 - \Delta_1 \tag{6.14}$$

$$x_2 = X - d_2 - \Delta_2 \tag{6.15}$$

as indicated on Figure 6.4. With the aid of the polynomial approach the displacement gradients of the contact patches are found to be

$$\frac{du_{z1}(x_1)}{dx_1} = \sum_{m=0}^{M-1} \beta_m x_1^m + I_{-1,1} \sum_{m=0}^{M} B_m x_1^m \tag{6.16}$$

$$\frac{du_{z2}(x_2)}{dx_2} = \sum_{m=0}^{M-1} \gamma_m x_2^m + I_{-1,2} \sum_{m=0}^{M} G_m x_2^m \tag{6.17}$$

$$I_{-1,1} = \begin{cases} 0 & , \quad |x_1| \leq a_1 \\ \frac{\mathrm{sign}(x_1)\pi}{\sqrt{x_1^2 - a_1^2}} & , \quad |x_1| > a_1 \end{cases} \tag{6.18}$$

$$I_{-1,2} = \begin{cases} 0 & , \quad |x_2| \leq a_2 \\ \frac{\mathrm{sign}(x_2)\pi}{\sqrt{x_2^2 - a_2^2}} & , \quad |x_2| > a_2 \end{cases} \tag{6.19}$$

where the β_m's and the γ_m's are derived from the B_m's and the G_m's, respectively, applying the polynomial approach, where a in the matrix $[A]$ is substituted with a_1 and a_2. The normal contact problem for the two

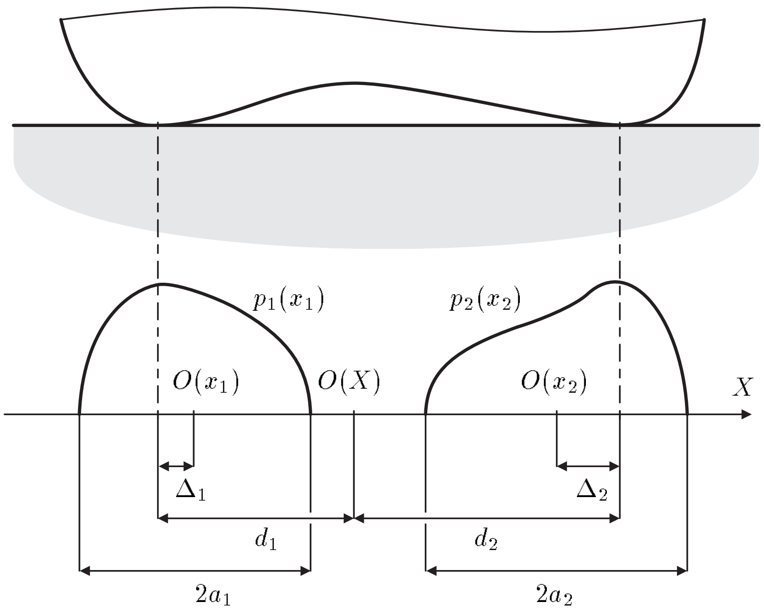

Figure 6.4. Local coordinate systems for adjacent contact patches.

contact patches is then given as

$$\frac{dZ(X)}{dX} = \sum_{m=0}^{M-1} [\beta_m x_1^m + \gamma_m x_2^m] + I_{-1,2} \sum_{m=0}^{M} G_m x_2^m \quad , \quad |x_1| \leq a_1 \tag{6.20}$$

$$\frac{dZ(X)}{dX} = \sum_{m=0}^{M-1} [\gamma_m x_2^m + \beta_m x_1^m] + I_{-1,1} \sum_{m=0}^{M} B_m x_1^m \quad , \quad |x_2| \leq a_2 \tag{6.21}$$

It is not possible to solve this system of equations in a straightforward way due to the square roots in $I_{-1,1}$ and $I_{-1,2}$. One way to overcome this problem is by making a Taylor approximation of $I_{-1,1}$ and $I_{-1,2}$ with the point of evolution in the centre of the adjacent contact patch i.e. at $x_1 = 0$ for $I_{-1,2}$ and at $x_2 = 0$ for $I_{-1,1}$. Now the equations (6.20) and (6.21) are reduced to pure polynomial equations, which implies that the unknown B_m's and G_m's can be found by comparing the polynomial coefficients.

The remaining unknowns a_1, a_2, d_1, d_2, B_0 and G_0 can be found from the boundary conditions:

$$\sum_{m=0}^{M} B_m(-a_1)^m = \sum_{m=0}^{M} B_m a_1^m = 0 \tag{6.22}$$

$$\sum_{m=0}^{M} G_m(-a_2)^m = \sum_{m=0}^{M} G_m a_2^m = 0 \tag{6.23}$$

and the relation between the normal pressure and the normal force:

$$\tilde{N}_1 = -2\{A_{0,1}\}^T\{B_0\} \tag{6.24}$$

$$\tilde{N}_2 = -2\{A_{0,2}\}^T\{G_0\} \tag{6.25}$$

where the notation $\{A_{0,1}\}$ denotes the vector $\{A_0\}$ with $a = a_1$. The only problem remaining is to define $\tilde{N}_1$ and $\tilde{N}_2$ which refer to the normal force acting on each contact patch and where $\tilde{N} = \tilde{N}_1 + \tilde{N}_2$. Unfortunately this problem is unsolvable because it is impossible to calculate the penetration in the two-dimensional contact problem (see section 2.4.1). It is thus necessary to make some sort of estimate of how the normal force is distributed between the two asperities. Several approaches to this problem have been formulated (e.g. [57]) and the qualitatively behaviour of the cross influence between the contact patches is not critical with respect to the distribution of the normal forces. Even though the contact problem for two adjacent contact patches now is formulated as a set of algebraic equations it is in general not possible to find an analytical solution because the quantities a_1, a_2, d_1 and d_2 are represented in the equations in a nonlinear way. The problem can however be solved with an iterative method in just a few steps if the Hertzian values are utilized as initial guess.

After having derived a method to solve the normal contact problem for two interfering contact patches, it is of interest to investigate how the cross influence between the contact patches affect the entire contact problem i.e. to

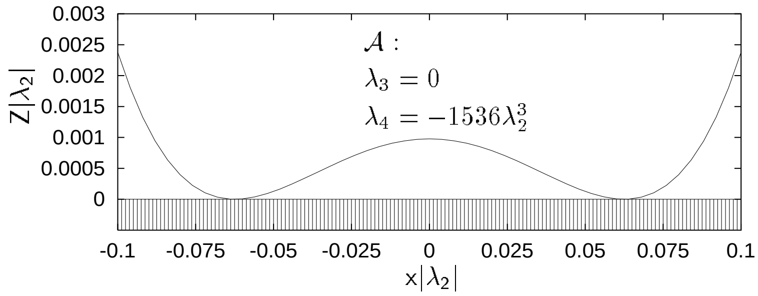

Figure 6.5. The asperity $\mathcal{A}$ providing two distinct contact patches.

compare the result for the calculations where the cross influence is included with an approach without cross influence. Consider a roughness asperity $\mathcal{A}$ where $\lambda_3 = 0$ and $\lambda_4 = -1536\lambda_2^3$ (see Figure 6.5). The asperity has two minima at $d_{1,2} = \pm 1/(16\lambda_2)$. If the normal force is small this will result in two distinct contact patches. Assuming that there is no cross influence between the contact problems, the theory from the previous section can be applied on each minimum. Introduce the transformations $\bar{x}_1 = X - d_1$ and $\bar{x}_2 = X - d_2$ then the two local asperities

$$Z_1(\bar{x}) = \frac{1}{2}\tilde{\lambda}_2\bar{x}^2 + \frac{1}{6}\tilde{\lambda}_3\bar{x}^3 + \frac{1}{24}\tilde{\lambda}_4\bar{x}^4 \qquad (6.26)$$

$$Z_2(\bar{x}) = \frac{1}{2}\tilde{\lambda}_2\bar{x}^2 - \frac{1}{6}\tilde{\lambda}_3\bar{x}^3 + \frac{1}{24}\tilde{\lambda}_4\bar{x}^4 \qquad (6.27)$$

$$\tilde{\lambda}_2 = -2\lambda_2 \ , \ \tilde{\lambda}_3 = 96\lambda_2^2 \ , \ \tilde{\lambda}_4 = -1536\lambda_2^3 \qquad (6.28)$$

can be investigated with the theory from the previous section.

The normal pressure distributions for the two approaches are demonstrated in Figure 6.6 for various normal forces. Not surprisingly, the stress distributions are very much alike when the normal force is small, i.e. when the relative distance between the two contact patches is large. Because the relative displacements decrease rapidly as the distance from the contact patch

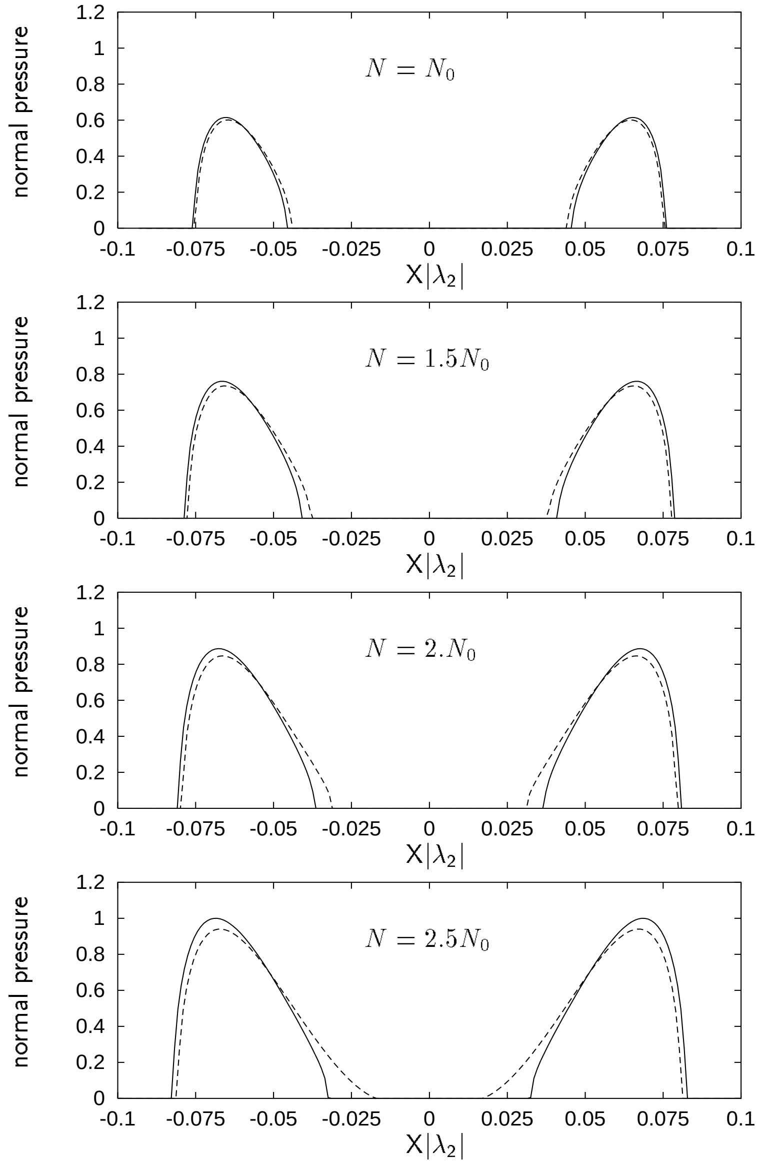

Figure 6.6. Normal pressure distribution for adjacent contact patches. Solid line: with cross influence. Dashed line: without cross influence

grows, there will hardly be any cross influence in these cases. As the normal load increases two things happen: the size of the contact patches becomes larger and the location of the centre of each contact patch represented by Δ_1 and Δ_2 is shifted more towards the position $X = 0$. Both properties result in a shorter distance between the contact patches and thus in a larger cross influence. The result is quite evident: for large normal loads the discrepancies between the two solutions are rather big. A consequence of the importance of the cross influence is, that a description of a rough surface also must include the distance between the asperities if it is sought to make a description of the surface which is unique in a contact mechanical context. Thus, mean-line roughness parameters as root-mean-square or centre-line average ([63]) are not adequate as isolated descriptors of a rough surface, since they cannot evaluate the distance between the roughness asperities.

Due to the symmetry of $\mathcal{A}$ it is obvious that the investigated contact patches are symmetric around $X = 0$. The qualitative result of the above comparison between calculations with cross influence and calculations without cross influence is however not a result of the symmetry. The demonstrated properties will also occur if the local asperities are not symmetric or if the normal loads acting on them are not identical.

6.5 Tangential Contact of Rough Surfaces

When the normal contact problem is solved for the case of rough bodies in contact, it should be possible to solve the tangential contact problem in a similar way, just like it has been done in the previous chapters.

The solution of the tangential contact problem for rough surfaces poses however some conceptual problems. If both bodies in contact are rough, the slope of the asperities will cause the normal force between two asperity sides to affect the tangential contact problem [15]. Consequently it is not possible to superpose the roughness onto one of the bodies such that the problem is transformed into the case of a rough body in contact with a perfectly smooth body. Assuming that the wave lengths of the asperities are much larger than their amplitudes, this effect can however be disregarded.

Another problem is that as the local asperity travels through the Hertzian contact patch, the local normal force varies swiftly, which implies that the contact problem actually is non-steady. The following investigations consider the tangential contact as a sum of local, stationary contact problems, i.e. that the tangential contact problem is solved for each local contact patch in a straightforward way. Important properties may thus be ignored, but the investigation will yet serve to provide an understanding of the basic properties of tangential contact for rough surfaces.

6.5.1 Tangential Contact for one Asperity

Assume that the contact patch is divided into a stick zone at the leading edge and a slip zone at the trailing edge, then the tangential stress distribution is defined as

$$q(x) = q_1(x) + q_2(x^*) \tag{6.29}$$

with the kinematic constraint

$$0 = \xi + \frac{du_{x1}(x)}{dx} + \frac{du_{x2}(x^*)}{dx} \tag{6.30}$$

As always $q_1(x) = \mu p(x)$.

Consider the asperities investigated in section 6.3, where the normal pressure distribution had the form

$$p(x) = \frac{\pi E}{4(1-\nu^2)} \frac{\sum_{n=0}^{4} B_n x^n}{\sqrt{a^2 - x^2}} \tag{6.31}$$

Employing the polynomial approach this leads to an expression for $q_2(x^*)$

$$q_2(x^*) = \frac{\mu \pi E}{4(1-\nu^2)} \frac{\sum_{n=0}^{4} B_n^* x^{*n}}{\sqrt{a^{*2} - x^{*2}}} \tag{6.32}$$

where the coefficients are given as

$$B_0^* = -\frac{a^{*2}}{\pi}\left[\lambda_2 + \lambda_3 \Delta^* + \frac{1}{2}\lambda_4 \Delta^{*2} + \frac{1}{12}\lambda_4 a^{*2}\right] \tag{6.33}$$

$$B_1^* = -\frac{1}{\pi}\left[\frac{\xi}{\mu} + \frac{1}{4}a^{*2}\left(\lambda_3 + \lambda_4 \Delta^*\right) - \Delta^*\left(\lambda_2 + \frac{1}{2}\lambda_3 \Delta^* + \frac{1}{6}\lambda_4 \Delta^{*2}\right)\right] \tag{6.34}$$

$$B_2^* = -\frac{1}{\pi}\left[\frac{1}{12}a^{*2}\lambda_4 - \lambda_2 - \lambda_3 \Delta^* - \frac{1}{2}\lambda_4 \Delta^{*2}\right] \tag{6.35}$$

$$B_3^* = \frac{1}{2\pi}\left[\lambda_3 + \lambda_4 \Delta^*\right] \tag{6.36}$$

$$B_4^* = \frac{1}{6\pi}\lambda_4 \tag{6.37}$$

and where $\Delta^* = \Delta - a + a^*$. The unknowns of the tangential contact problem are now a^* and either the creepage ξ or the tangential force T. They can be found by employing the boundary conditions and the relation between the tangential stress and the tangential force, which yields the equations

$$\xi = \mu \Delta^*\left(\lambda_2 + \frac{1}{2}\lambda_3 \Delta^* + \frac{1}{6}\lambda_4 \Delta^{*2}\right) + \frac{1}{4}\mu a^{*2}\left(\lambda_3 + \lambda_4 \Delta^*\right) \tag{6.38}$$

$$T = \mu N - \frac{\pi \mu E}{8\left(1-\nu^2\right)}\left\{\left(\lambda_2 + \lambda_3 \Delta^* + \frac{1}{2}\lambda_4 \Delta^{*2}\right) a^{*2} + \frac{1}{8}\lambda_4 a^{*4}\right\} \tag{6.39}$$

These two equations can for certain values of the λ's be solved analytically just like in the case of the equivalent normal contact problems. If an analytical solution cannot be found, an iterative process converges in just a few steps if the Carter solution is applied as initial guess.

In Figure 6.7 the tangential stress distributions for the asperities $\mathcal{A}_1, \ldots, \mathcal{A}_4$ are monitored for $T = 0.75\mu N$. It is clearly seen that the size of the stick zone varies with the geometry of the asperity. A result of the varying a^* is, that the creepage is not the same for the four contact situations. In Figure 6.8 the equivalent creep curves are plotted, where the $\diamond$ indicates the contact situations from Figure 6.7. Two distinct properties must be noticed when it comes to the creep curves: the magnitude of the critical creepage and the initial slope of the creep curve depend on the geometry of the asperities.

The critical creepage ξ_c for which complete sliding occurs can be derived from equation (6.38) by setting $a^* = 0$, which yields the expression

$$\xi_c = (\Delta - a)\left(\lambda_2 + \frac{1}{2}\lambda_3(\Delta - a) + \frac{1}{6}\lambda_4(\Delta - a)^2\right) \tag{6.40}$$

The slope of the creep curve is found by differentiating equation (6.38) and equation (6.39) with respect to a^*, divide the two quantities and then let a^* approach a. This calculation leads to the expression

$$\lim_{\xi \to 0^-}\left[\frac{dT}{d\xi}\right] = -\frac{\pi E}{8(1-\nu^2)}\left(2a + \mathcal{O}(\lambda_5)\right) \tag{6.41}$$

The consequence of this result is very interesting. If the asperity in the vicinity of the contact point can be approximated with a fourth order polynomial then the initial slope of the creep curve will always be proportional to the size of the contact patch.

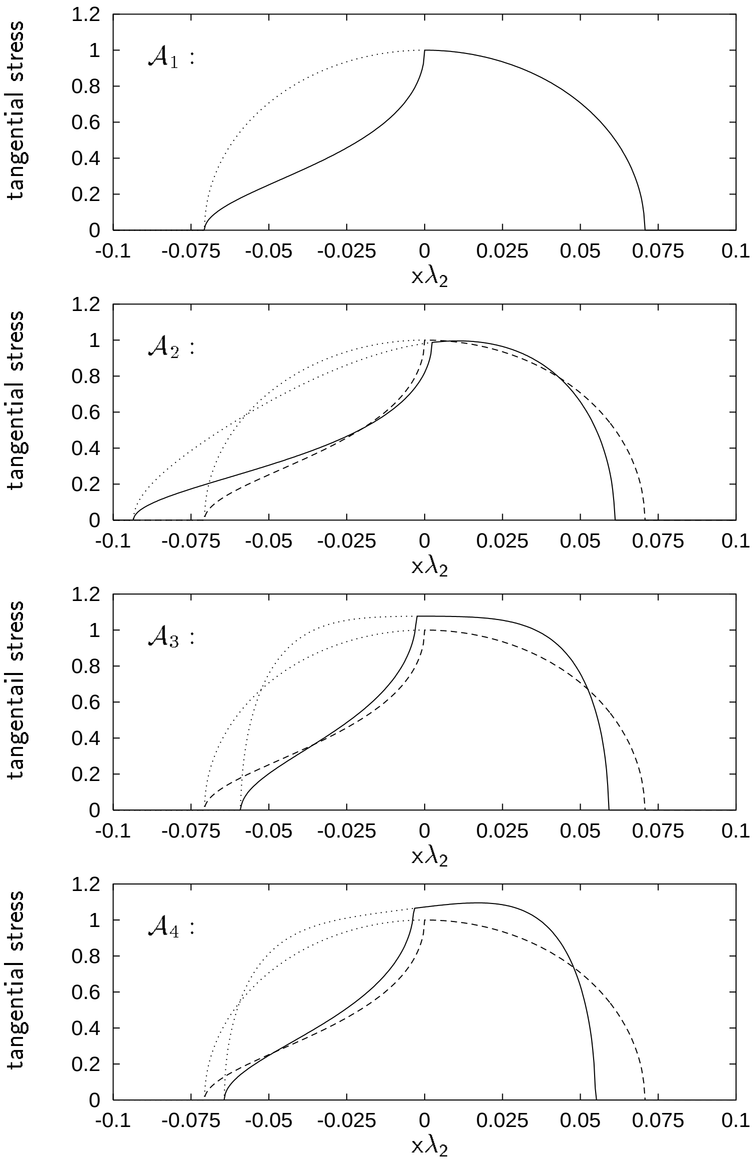

Figure 6.7. Tangential stress distribution for $T = 0.75\mu N$. The dashed line indicates $\mathcal{A}_1$.

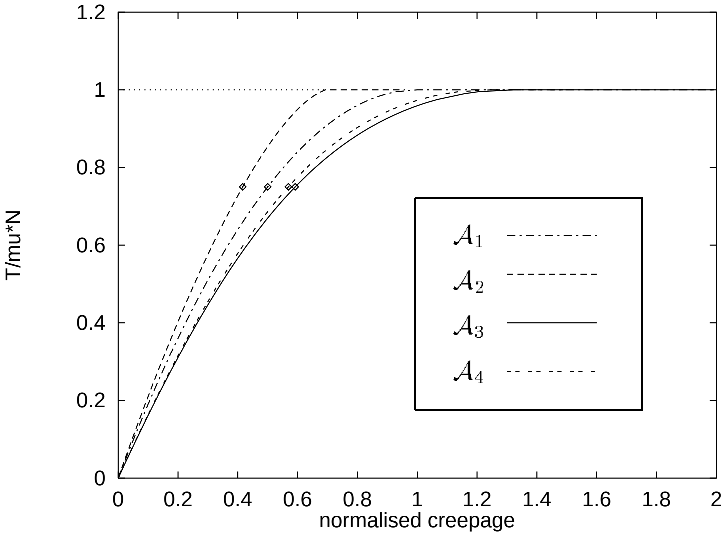

Figure 6.8. Creep curves for the four different asperities. The $\diamond$ indicates the contact situations shown in Figure 6.7.

Assuming that the tangential contact problem is unaffected by the cross influence between adjacent contact patches when the creepage tends towards zero, the result can be generalized. With no cross influence between the contact patches, the principle of superposition is valid, and so the initial slope of the creep curve for a rough contact with arbitrary many local contact patches will be proportional to the global size of the contact patch as long as the asperities can be approximated by fourth order polynomials. This relation has been suggested by Knothe and Theiler [42] as a result of numerical calculations based on measured rough surfaces and is thus actually proved. It must therefore be emphasized that the size of the contact patch is a crucial parameter also when it comes to the tangential contact problem for rough surfaces.

6.5.2 The Tangential Contact Problem with Cross Influence

Like in the case with the normal contact problem, adjacent contact patches will influence each other when it comes to the tangential contact problem. The tangential problem with the cross influence between adjacent contact patches can be solved by applying an approach equivalent to the one from the normal contact problem.

Consider two contact patches with the normal pressure distributions $p_1(x_1)$ and $p_2(x_2)$ defined as in equation (6.12) and equation (6.13), then the tangential stress distributions over the contact patches are assumed to have the form

$$q_1(x_1) = \mu p_1(x_1) + q_{1,2}(x_1^*) \tag{6.42}$$

$$q_2(x_1) = \mu p_2(x_2) + q_{2,2}(x_2^*) \tag{6.43}$$

$$q_{1,2}(x_1^*) = \frac{\mu\pi E}{4(1-\nu^2)} \frac{\sum_{m=0}^{M} B_m^* x_1^{*m}}{\sqrt{a_1^{*2} - x_1^{*2}}} \tag{6.44}$$

$$q_{2,2}(x_2^*) = \frac{\mu\pi E}{4(1-\nu^2)} \frac{\sum_{m=0}^{M} G_m^* x_2^{*m}}{\sqrt{a_2^{*2} - x_2^{*2}}} \tag{6.45}$$

where the local coordinates x_1^* and x_2^* are defined as

$$x_1^* = x_1 + a_1 - a_1^* \tag{6.46}$$

$$x_2^* = x_2 + a_2 - a_2^* \tag{6.47}$$

Each of the two contact patches are thus divided into a stick zone at the leading edge with half the size a_1^* and a_2^*, respectively, and a slip zone at the trailing edge.

The gradients of the relative displacements can now be found using the polynomial approach, which provides the kinematic constraints for the two stick zones

$$\xi = \sum_{m=0}^{M-1} [\beta_m x_1^m + \gamma_m x_2^m + \beta_m^* x_1^{*m} + \gamma_m^* x_2^{*m}] + \sum_{m=0}^{M} \left[I_{-1,2} G_m x_2^m + I_{-1,2}^* G_m^* x_2^{*m}\right] \ , \ \ |x_1^*| \le a_1^* \qquad (6.48)$$

$$\xi = \sum_{m=0}^{M-1} [\beta_m x_1^m + \gamma_m x_2^m + \beta_m^* x_1^{*m} + \gamma_m^* x_2^{*m}] + \sum_{m=0}^{M} \left[I_{-1,1} B_m x_1^m + I_{-1,1}^* B_m^* x_1^{*m}\right] \ , \ \ |x_2^*| \le a_2^* \qquad (6.49)$$

where the β_m^*'s and the γ_m^*'s are derived from the B_m^*'s and the G_m^*'s, respectively, applying the polynomial approach, where a in the matrix $[A]$ is substituted with a_1^* and a_2^*. Finally $I_{-1,1}^*$ and $I_{-1,2}^*$ are defined as

$$I_{-1,1}^* = \begin{cases} 0 & , \ \ |x_1^*| \le a_1^* \\ \frac{\mathrm{sign}(x_1^*)\pi}{\sqrt{x_1^{*2} - a_1^{*2}}} & , \ \ |x_1^*| > a_1^* \end{cases} \qquad (6.50)$$

$$I_{-1,2}^* = \begin{cases} 0 & , \ \ |x_2^*| \le a_2^* \\ \frac{\mathrm{sign}(x_2^*)\pi}{\sqrt{x_2^{*2} - a_2^{*2}}} & , \ \ |x_2^*| > a_2^* \end{cases} \qquad (6.51)$$

The procedure is now exactly the same as for the normal contact problem. $I_{-1,1}$, $I_{-1,2}$, $I_{-1,1}^*$ and $I_{-1,2}^*$ are transformed into polynomials and so the unknowns B_0^*, G_0^*, a_1^*, a_2^* plus either the local tangential force or the local creepage can be found from the boundary conditions and the condition, that the tangential force is equal the integral of the tangential stress distribution.

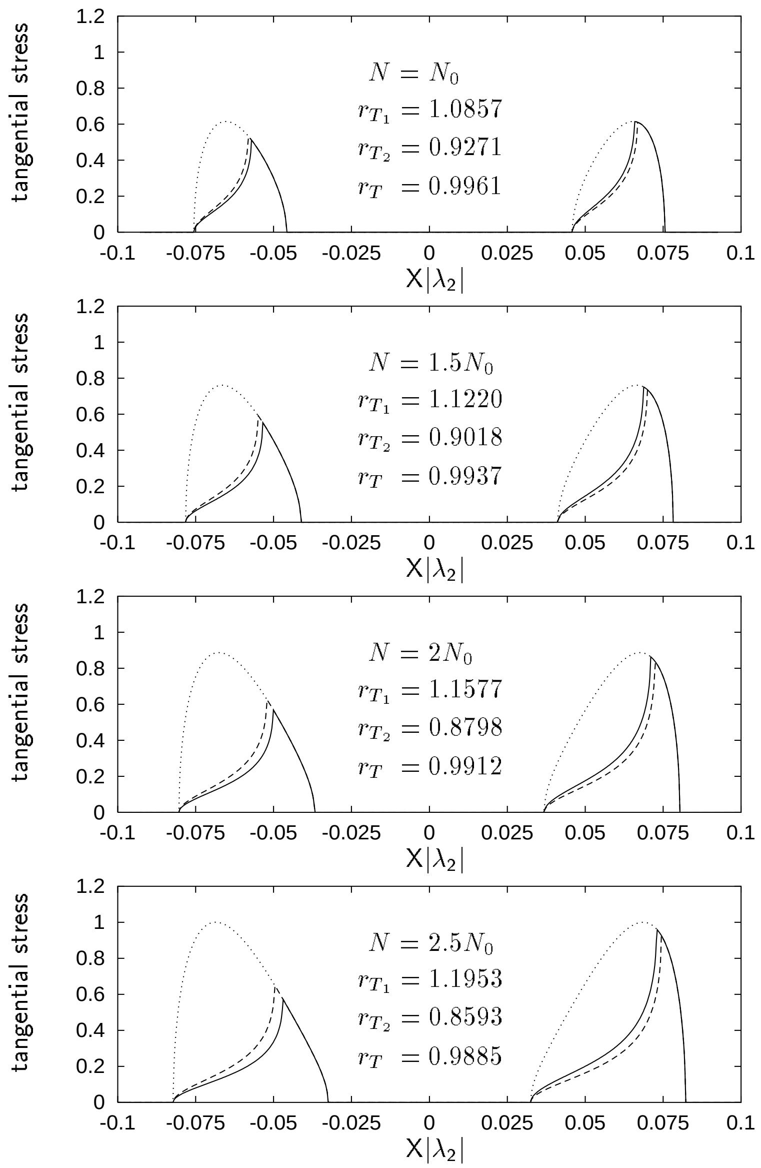

Figure 6.9. Tangential stress distribution for adjacent contact patches. Solid line: with cross influence. Dashed line: without cross influence.

As the contact problem is assumed to be stationary, the local creepages acting on the two contact patches must be identical. In Figure 6.9 the tangential stress distribution is calculated for the normal contact problem from section 6.4 (see Figure 6.6). The normal pressure distribution which is utilized as foundation for the calculations of the tangential contact problem is the one where the cross influence is incorporated. The four contact situations are calculated for the constant creepage $\xi = \xi_0$, which implies that the tangential force varies as the normal force is increased. Due to the nonlinearity of the problem, the ratio between normal force and tangential force is not constant.

In Figure 6.9 three different ratios are listed. The ratios are defined as

$$r_{T_1} = \frac{T_1 \text{ without cross influence}}{T_1 \text{ with cross influence}} \tag{6.52}$$

$$r_{T_2} = \frac{T_2 \text{ without cross influence}}{T_2 \text{ with cross influence}} \tag{6.53}$$

$$r_{T} = \frac{T \text{ without cross influence}}{T \text{ with cross influence}} \tag{6.54}$$

where T_1 refers to the first contact patch, T_2 refers to the second contact patch and T is the entire tangential force. The calculated ratios indicate that the entire tangential force is almost unaffected by the cross influence. The cross influence implies that the first contact patch will have a larger stick zone and a smaller resulting tangential force whereas a larger tangential force is transmitted through the second contact patch because the stick zone diminishes.

	$N = N_0$	$N = 1.5N_0$	$N = 2N_0$	$N = 2.5N_0$
$\xi = 0.5\xi_0$	$r_{T_1} = 1.1108$ $r_{T_2} = 0.9087$ $r_T = 0.9966$	$r_{T_1} = 1.1505$ $r_{T_2} = 0.8836$ $r_T = 0.9952$	$r_{T_1} = 1.1903$ $r_{T_2} = 0.8613$ $r_T = 0.9935$	$r_{T_1} = 1.2392$ $r_{T_2} = 0.8402$ $r_T = 0.9916$
$\xi = \xi_0$	$r_{T_1} = 1.0857$ $r_{T_2} = 0.9271$ $r_T = 0.9961$	$r_{T_1} = 1.1220$ $r_{T_2} = 0.9018$ $r_T = 0.9937$	$r_{T_1} = 1.1577$ $r_{T_2} = 0.8798$ $r_T = 0.9912$	$r_{T_1} = 1.1953$ $r_{T_2} = 0.8593$ $r_T = 0.9885$
$\xi = 1.5\xi_0$	$r_{T_1} = 1.0630$ $r_{T_2} = 0.9497$ $r_T = 0.9986$	$r_{T_1} = 1.0968$ $r_{T_2} = 0.9212$ $r_T = 0.9950$	$r_{T_1} = 1.1297$ $r_{T_2} = 0.8987$ $r_T = 0.9918$	$r_{T_1} = 1.1639$ $r_{T_2} = 0.8781$ $r_T = 0.9884$
$\xi = 2\xi_0$	$r_{T_1} = 1.0481$ $r_{T_2} = 0.9774$ $r_T = 1.0067$	$r_{T_1} = 1.0733$ $r_{T_2} = 0.9460$ $r_T = 1.0006$	$r_{T_1} = 1.1040$ $r_{T_2} = 0.9212$ $r_T = 0.9960$	$r_{T_1} = 1.1356$ $r_{T_2} = 0.8995$ $r_T = 0.9918$

Table 6.2. Ratios between tangential forces calculated without cross influence and with cross influence for various values of the creepage and the normal force.

In Table 6.2 the three ratios are calculated for various values of the creepage and the normal force, where the reference creepage ξ_0 is the one which is used for the contact situations shown in Figure 6.9. Three trends are apparent:

1. Increasing normal force resulting in shorter distance between the contact patches implies that the second contact patch contributes more to the total tangential force.

2. Smaller creepage implies that the second contact patch contributes more to the total tangential force.
3. The total tangential force is almost unaffected by the cross influence between the contact patches.

The above calculations indicate that the tangential contact problem is less affected by the cross influence than the normal contact problem. The two problems are solved for exactly the same configuration in Figure 6.6 and Figure 6.9. Considering the size of the contact patch as a crucial quantity for the normal contact problem and the total tangential force as the important quantity for the tangential contact problem it is evident that the cross influence is not as important for the tangential contact problem as it is for the normal contact problem. It must however be emphasized that when the distance between the contact patches is small compared to the characteristic size of the contact patches, both the normal contact problem and the tangential contact problem will be affected by the cross influence.

Chapter 7

Non-Steady Two-Dimensional Contact

In the present chapter an approximative method to investigate non-steady contact for the two-dimensional tangential contact problem is introduced. The objective of the method is to find the tangential stress distribution and establish a relation between the creepage and the tangential force. It is demonstrated that exact solutions for the tangential stress distribution and the tangential force can be derived, whereas the relation between the creepage and the tangential force depends on an approximation of the displacements at a reference point. To illustrate the application of the theory a contact problem with an oscillating tangential force is examined.

7.1 Non-Steady Contact

In the previous chapters the tangential contact problem is considered to be stationary, i.e. it is assumed that the quantities of the contact problem changes very slowly implying that the time derivative of the relative displacements can be neglected. Thus, it is a precondition for stationary contact that the characteristic wave length of the contact problem is much larger than the size of the contact patch. This assumption is much too primitive for a variety of problems of wheel/rail contact such as corrugation problems or contact involving rough surfaces. For that reason it is necessary to be able to incorporate non-stationarity in the contact model in order to extend the field of application.

The case of three-dimensional non-steady contact has been investigated thoroughly by Groß-Thebing [22] and [23], who extended Kalker's theory based on discretization of the contact patch [38] to a non-steady application. By introducing a linearization a non-steady theory equivalent to Kalker's linear theory is obtained, where the creep coefficients now are frequency dependent. In the latter work of Groß-Thebing the approach is modified so the linearization is made with respect to an arbitrary reference state of the nonlinear stationary contact model. The problem with Groß-Thebing's contact model is that it is based on linearizations of the basic equations within each discretized contact element. This approach is therefore only adequate for contact situations where the amplitudes of the oscillating quantities are small. When this is not the case a very fine discretization of the contact patch much be applied, which augment the computation time considerably.

In Chapter 2 it was demonstrated that it is impossible to solve the two-dimensional non-steady contact problem because the relative displacement depends on the choice of datum. It is however possible to make some qualitative investigations of the problem. As in many other fields of rolling contact mechanics Kalker has made some of the primary work in this field [36] and [37]. The first paper [36] is an analytical description of the case where the stick zone is assumed to cover the entire contact patch, i.e. a contact with infinite friction, whereas the second paper [37] introduces a numerical approach to non-steady rolling contact. In the present chapter an analytical approach to two-dimensional, non-steady rolling contact where the contact patch is divided into one stick zone and one slip zone will be introduced.

7.2 Deriving a Non-Steady Theory

In the stationary contact theory it is assumed that the relative displacement $u_x = u_x(x)$ where the local coordinate x is given by the global coordinate X as $x = X - V_m t$. The kinematic constraint for a particle in the contact patch then reads

$$-V_m s(x) = -V_m \xi + \frac{du_x(x)}{dt} \tag{7.1}$$

where V_m is the mean velocity [33]. With the definition of x this leads to the expression

$$-V_m s(x) = -V_m \xi + \frac{du_x(x)}{dx}\frac{dx}{dt} \quad \Rightarrow$$

$$s(x) = \xi + \frac{du_x(x)}{dx} \tag{7.2}$$

This is the kinematic constraint which is the foundation of all stationary contact theories.

If the relative displacement now also depends on the time i.e. $u_x = u_x(x,t)$ then the kinematic constraint is calculated to

$$-V_m s(x,t) = -V_m \xi(t) + \frac{\partial u_x(x,t)}{\partial x}\frac{dx}{dt} + \frac{\partial u_x(x,t)}{\partial t} \quad \Rightarrow$$

$$s(x,t) = \xi(t) + \frac{\partial u_x(x)}{\partial x} - \frac{1}{V_m}\frac{\partial u_x(x,t)}{\partial t} \tag{7.3}$$

With the definition of the slip being zero in the stick zone this yields that

$$0 = \xi(t) + \frac{\partial u_x(x,t)}{\partial x} - \frac{1}{V_m}\frac{\partial u_x(x,t)}{\partial t} \quad , \quad x \in S_{stick} \tag{7.4}$$

In order to find an expression for the time dependent creepage it is necessary to be able to calculate the derivatives of $u_x(x,t)$. Due to the outline of the kinematic constraint it is assumed that $u_x(x,t)$ can be expressed as

$$u_x(x,t) = U_x(x + V_m t) - \xi_0 x - U_x(x_0 + V_m t) + \xi_0 x_0 + u_x(x_0,t) \tag{7.5}$$

where the reference coordinate $x_0 \in S_{stick}$ is defined as

$$x_0 = X - V_m t + f(t) \tag{7.6}$$

The derivatives of the relative displacement are then

$$\frac{\partial u_x(x,t)}{\partial x} = U_x'(x + V_m t) - \xi_0 \tag{7.7}$$

$$\frac{\partial u_x(x,t)}{\partial t} = V_m U_x'(x + V_m t) - U_x'(x_0 + V_m t) f'(t) + \left(\xi_0 + \frac{\partial u_x(x_0,t)}{\partial x_0}\right)(f'(t) - V_m) + \frac{\partial u_x(x_0,t)}{\partial t} \tag{7.8}$$

which inserted into the kinematic constraint yields that

$$\xi(t) = -\frac{1}{V_m} U_x'(x_0 + V_m t) f'(t) + \frac{1}{V_m}\left(\xi_0 + \frac{\partial u_x(x_0,t)}{\partial x_0}\right) f'(t) - \frac{\partial u_x(x_0,t)}{\partial x_0} + \frac{1}{V_m}\frac{\partial u_x(x_0,t)}{\partial t} \tag{7.9}$$

In order to simplify this expression it is assumed that $f'(t)$ and $U_x'(x_0+V_m t)$ are small i.e. that their product is small compared to the other quantities and consequently can be neglected. This assumption is reasonable if the position of x_0 is located in a small interval of the contact patch. As x_0 always is located in the stick zone it further follows that the term

$$\left(\xi_0 + \frac{\partial u_x(x_0,t)}{\partial x_0}\right) f'(t)$$

also can be neglected. With these simplifications the expression for the creepage is reduced to

$$\xi(t) = -\frac{\partial u_x(x_0,t)}{\partial x_0} + \frac{1}{V_m}\frac{\partial u_x(x_0,t)}{\partial t} \quad , \quad x_0 \in S_{stick} \tag{7.10}$$

Since $u_x(x_0,t)$ depends on the choice of datum the derivative $\partial u_x(x_0,t)/\partial t$ is an unknown function and so the problem cannot be solved unless an appropriate approximation for $u_x(x_0,t)$ is applied.

In order to find an expression for $u_x(x_0,t)$, Johnson has suggested an approach where the three-dimensional case of a rectangular contact patch with uniform tangential stress distribution is considered [32]. Denoting the side lengths of the rectangle as $2a$ and $2b$ where $2a$ is the length parallel to the rolling direction, the value of $u_x(x_0,t)$ is approximated with the displacement at the centre of the rectangle when the uniform tangential stress distribution $q(x,y,t) = T/2a$ is acting on the contact patch. T is in this context the tangential force per unit length. In principle the assumption of complete sticking leads to a singularity in the tangential stress distribution at the trailing edge of the contact patch. To avoid this phenomenon the Johnson approach is slightly modified, so the tangential stress distribution is expressed as the sum of two uniform stress distributions where one is defined over the entire contact patch and the other is defined over the stick zone:

$$q(x,y,t) = q_1(x,t) + q_2(x^*,t) \tag{7.11}$$

$$q_1(x,t) = \frac{\mu N}{2a} \quad , \quad -a < x < a \tag{7.12}$$

$$q_2(x^*,t) = -\frac{\mu N - T}{2a^*} \quad , \quad -a^* < x^* < a^* \tag{7.13}$$

Applying the three-dimensional constitutive equation (see Appendix C) on this tangential stress distribution under the assumption that $b \gg a$, an expression for $u_x(x_0,t)$ is found to be

$$u_x(x_0,t) = \frac{4(1-\nu^2)}{\pi E}\left\{\mu N\left[g(x_0,a) - g(x_0^*,a^*)\right] + Tg(x_0^*,a^*)\right\} \tag{7.14}$$

$$g(x_0,a) = \frac{1}{1-\nu} + \ln(2b) + \frac{1}{2}\frac{x_0}{a}\ln\left(\frac{a-x_0}{a+x_0}\right) - \frac{1}{2}\ln\left(a^2 - x_0^2\right) \tag{7.15}$$

where x_0 is defined to be the centre of the stick zone i.e. $x_0 = a^* - a$ and $x_0^* = 0$.

The above approximation of $u_x(x_0,t)$ is quite primitive and is only valid when the size of the slip zone is very small i.e. when $a_0^* \to a_0$. This constraint implies that the theory is equivalent to the stationary linear model where the stick zone always covers the entire contact patch. The simplified expression for the tangential stress distribution results in values for the stationary creepage which are different from the stationary value of the Carter solution when the stick zone does not cover the entire contact patch.

A better approximation of $u_x(x_0,t)$ can be achieved by considering the same rectangular contact patch with the side lengths $2a$ and $2b$, but where the tangential stress distribution in a cross section of this contact patch is equal to the Carter stress distribution:

$$q(x,y,t) = q_1(x,t) + q_2(x^*,t) \tag{7.16}$$

$$q_1(x,t) = \frac{\mu p_0}{a}\sqrt{a^2 - x^2} \ , \quad -a < x < a \tag{7.17}$$

$$q_2(x^*,t) = -\frac{\mu p_0}{a}\sqrt{a^{*2} - x^{*2}} \ , \quad -a^* < x^* < a^* \tag{7.18}$$

Again the constitutive equation provides an expression for $u_x(x_0,t)$:

$$u_x(x_0,t) = \frac{4(1-\nu^2)}{\pi E}\left[\mu N\left[g(x_0,a) - g(x_0^*,a^*)\right] + Tg(x_0^*,a^*)\right] \tag{7.19}$$

$$g(x_0,a) = \frac{1+\nu}{2(1-\nu)} + \ln(4b) - \ln(a) - \left(\frac{x_0}{a}\right)^2 \tag{7.20}$$

where x_0 still is defined as the centre of the stick zone.

The problem with the above calculations of $u_x(x_0,t)$ is that the approximation to an infinite cylinder rolling on a surface implies that $b \to \infty$ which also causes $u_x(x_0,t)$ to tend towards infinity. Instead it is only assumed that $b \gg a$ and then consider the result to be an indicator of the qualitative behaviour of the non-steady problem.

When all quantities except for the tangential force and the creepage are constant in time the non-steady contact problem is in principle solved by inserting the derivatives of $u_x(x_0,t)$ which can be found from equation (7.14) or from equation (7.19) into the kinematic constraint from equation (7.10). However, it is sought to derive a theory which can be applied on a wide range of non-steady contact problems where other quantities such as the size of the contact patch, the size of the stick zone or the velocity also vary in time. So it remains to establish a set of equations which describes how these other oscillating quantities are related. To do that it is necessary to solve the entire contact problem i.e. to find the tangential stress distribution, the size of the stick zone and the size of the tangential force.

In order to illustrate how the non-steady contact problem may be solved, a very simple example of non-steady contact is treated. It is obviously possible to employ the procedure on much more complicated problems, but the simple problem is chosen in order to demonstrate the basic concept of the method.

Consider a cylinder rolling on a smooth surface with constant velocity and constant normal force but with the oscillating tangential force:

$$T(t) = T_0 + \sum_{m=0}^{M} \left[T_{A,m} \cos(\omega_m t) + T_{B,m} \sin(\omega_m t) \right] \tag{7.21}$$

It is noticed that the assumption of constant normal force implies that the size of the contact patch is constant in time i.e. $a = a_0$. The objective of the calculations is now to establish an expression for the creepage on the form

$$\xi(t) = \xi_0 + \sum_{m=0}^{M} \left[\xi_{A,m} \cos(\omega_m t) + \xi_{B,m} \sin(\omega_m t) \right] \tag{7.22}$$

To do that the point of departure is taken in the gradient of the relative displacement

$$\partial u_x(x,t)/\partial x = U_x'(x + V_m t) - \xi_0 \tag{7.23}$$

As both the tangential force and the creepage are harmonic functions in time it is assumed that the function $U_x(x + V_m t)$ also is harmonic i.e.

$$\begin{aligned} U_x(x + V_m t) = \sum_{m=0}^{M} \mu \, [U_{A,m} \sin[k_m(x + V_m t)] - \\ U_{B,m} \cos[k_m(x + V_m t)]] \end{aligned} \tag{7.24}$$

In order to find the tangential stress distribution and the size of the stick zone an approach similar to the one for the corrugated surface must be applied. Assuming that the contact patch is divided into one stick zone and one slip zone, the gradient of the relative displacement in the stick zone is given as

$$\frac{\partial u_x(x,t)}{\partial x} = \frac{\partial u_{x1}(x,t)}{\partial x} + \frac{\partial u_{x2}(x,t)}{\partial x} \tag{7.25}$$

where $\partial u_{x1}(x,t)/\partial x = x\mu/R$ as the normal stress distribution is Hertzian. The equation (7.23) then indicates that

$$\frac{\partial u_{x2}(x,t)}{\partial x} = -\frac{\mu x}{R} - \xi_0 + \sum_{m=0}^{M} \mu k_m \left[U_{A,m}\cos(k_m x + \omega_m t) + U_{B,m}\sin(k_m x + \omega_m t)\right] \tag{7.26}$$

where $\omega_m = k_m V$. Introducing the coordinate transformation $x^* = x + a - a^*$ the displacement gradient is rewritten as

$$\frac{\partial u_{x2}(x^*,t)}{\partial x} = -\xi_0 - \frac{\mu(x^* + \Delta^*)}{R} + \sum_{m=0}^{M} \mu k_m \left[U_{A,m}\cos[k_m(x^* + \Delta^*) + \omega_m t] + U_{B,m}\sin[k_m(x^* + \Delta^*) + \omega_m t]\right] \tag{7.27}$$

where $\Delta^* = a^* - a$. Employing the polynomial approach, a procedure equivalent to the example of the corrugated surface (see section 4.2.3) yields the tangential stress distribution

$$q(x,t) = \mu p(x,t) + q_2(x^*,t) \tag{7.28}$$

where

$$q_2(x^*,t) = \frac{\mu\pi E}{4(1-\nu^2)} \frac{\sum_{n=0}^{\infty} B_n^* x^{*n}}{\sqrt{a^{*2} - x^{*2}}} \tag{7.29}$$

$$B_0^* = -\frac{a^{*2}}{\pi R} + \left(\frac{a^*}{\pi}\right)^2 \sum_{m=1}^{M} k_m \left[U_{A,m}\cos(k_m\Delta^* + \omega_m t) + \right.$$

$$\left. U_{B,m}\sin(k_m\Delta^* + \omega_m t)\right] \sum_{j=0}^{\infty} (-1)^j A_j^* \frac{k_m^{2j+1}}{(2j+1)!} \quad (7.30)$$

$$B_1^* = -\frac{\xi_0}{\pi\mu} + \frac{\Delta^*}{\pi R} - \sum_{m=1}^{M} k_m \left[U_{A,m}\sin(k_m\Delta^* + \omega_m t) - \right.$$

$$\left. U_{B,m}\cos(k_m\Delta^* + \omega_m t)\right] \sum_{j=0}^{\infty} (-1)^j A_j^{*-1} \frac{k_m^{2j}}{(2j)!} \quad (7.31)$$

$$B_2^* = \frac{1}{\pi R} - \sum_{m=1}^{M} k_m \left[U_{A,m}\cos(k_m\Delta^* + \omega_m t) + \right.$$

$$\left. U_{B,m}\sin(k_m\Delta^* + \omega_m t)\right] \sum_{j=0}^{\infty} (-1)^j A_j^{*-1} \frac{k_m^{2j+1}}{(2j+1)!} \quad (7.32)$$

$$\vdots$$

$$B_{2n-1}^* = (-1)^n \sum_{m=1}^{M} k_m \left[U_{A,m}\sin(k_m\Delta^* + \omega_m t) - \right.$$

$$\left. U_{B,m}\cos(k_m\Delta^* + \omega_m t)\right] \sum_{j=0}^{\infty} (-1)^j A_j^{*-1} \frac{k_m^{2j+2n-2}}{(2j+2n-2)!} \quad (7.33)$$

$$B_{2n}^* = (-1)^n \sum_{m=1}^{M} k_m \left[U_{A,m}\cos(k_m\Delta^* + \omega_m t) + \right.$$

$$\left. U_{B,m}\sin(k_m\Delta^* + \omega_m t)\right] \sum_{j=0}^{\infty} (-1)^j A_j^{*-1} \frac{k_m^{2j+2n-1}}{(2j+2n-1)!} \quad (7.34)$$

The boundary conditions provide the restriction that

$$\xi_0 = \frac{\mu\Delta^*}{R} + \sum_{m=0}^{M} \mu k_m \mathrm{J}_0(k_m a^*) \left[U_{A,m}\sin(k_m\Delta^* + \omega_m t) - \right.$$

$$\left. U_{B,m}\cos(k_m\Delta^* + \omega_m t)\right] \quad (7.35)$$

from which an expression for half the size of the stick zone a^* can be found depending on the coefficients $U_{A,m}$ and $U_{B,m}$.

Integrating the tangential stress distribution over the contact patch, the tangential force is found to be

$$T(t) = \mu N - \frac{\mu \pi E}{4(1-\nu^2)} \left\{ \frac{a^{*2}}{2R} + \sum_{m=0}^{M} k_m a^* \mathrm{J}_1(k_m a^*) \cdot \right.$$
$$\left. [U_{A,m} \cos(k_m \Delta^* + \omega_m t) + U_{B,m} \sin(k_m \Delta^* + \omega_m t)] \right\} \quad (7.36)$$

Comparing the coefficients from this expression with the ones from equation (7.21) the unknowns $U_{A,m}$ and $U_{B,m}$ are determined.

In order to employ the size of the stick zone and the magnitude of the tangential force in further calculations it is convenient to derive closed form expressions for these two quantities. Just like in the case of the corrugated surface a series expansion is applied under the assumption that $k^2 R\sqrt{U_A^2 + U_B^2} \ll 1$, which leads to the expressions:

$$a^*(t) = a_0^* + \sum_{m=0}^{M} \left[a_{A,m}^* \cos(\omega_m t) + a_{B,m}^* \sin(\omega_m t) \right] \quad (7.37)$$

$$T(t) = T_0 + \sum_{m=0}^{M} \left[T_{A,m} \cos(\omega_m t) + T_{B,m} \sin(\omega_m t) \right] \quad (7.38)$$

where the coefficients with indices A and B are found from the usual matrix equation

$$\left\{ \begin{array}{c} \left\{a_{A,m}^*, T_{A,m}\right\}^T \\ \\ \left\{a_{B,m}^*, T_{B,m}\right\}^T \end{array} \right\} = \left[\begin{array}{cc} \left\{a_{1,m}^*, T_{1,m}\right\}^T & -\left\{a_{2,m}^*, T_{2,m}\right\}^T \\ \\ \left\{a_{2,m}^*, T_{2,m}\right\}^T & \left\{a_{1,m}^*, T_{1,m}\right\}^T \end{array} \right] \left\{ \begin{array}{c} U_A \\ \\ U_B \end{array} \right\} \quad (7.39)$$

with the matrix coefficients

$$a^*_{1,m} = -k_m R \mathrm{J}_0(k_m a^*_0)\sin[k_m(a^*_0 - a_0)] \tag{7.40}$$

$$a^*_{2,m} = -k_m R \mathrm{J}_0(k_m a^*_0)\cos[k_m(a^*_0 - a_0)] \tag{7.41}$$

$$T_{1,m} = \frac{\mu\pi E}{4(1-\nu^2)} k_m a^*_0 \left[\frac{-a^*_{1,m}}{k_m R} - \mathrm{J}_1(k_m a^*_0)\cos[k_m(a^*_0 - a_0)]\right] \tag{7.42}$$

$$T_{2,m} = \frac{\mu\pi E}{4(1-\nu^2)} k_m a^*_0 \left[\frac{-a^*_{2,m}}{k_m R} + \mathrm{J}_1(k_m a^*_0)\sin[k_m(a^*_0 - a_0)]\right] \tag{7.43}$$

In Figure 7.1 the tangential stress distribution is shown for the case of one distinct wave length where $a_0/L = 1$ i.e. a contact situation where the oscillations in time are so fast, that the time derivative of the relative displacement must be included in the contact model. The importance of including the term $\partial u_x(x,t)/\partial t$ is clearly demonstrated in the figure, where the non-steady solution (solid line) is compared with the stationary Carter solution (dashed line). The discrepancies between the two solutions are significant and even though the amplitude of a^* indicated by the location of the limit between stick zone and slip zone is minor, the tangential stress distribution varies considerably inside the stick zone.

At the extreme a contact situation where a^* is constant but the tangential stress distribution still oscillates can occur. This is the case for configurations where $\mathrm{J}_0(ka^*) = 0$. Thus, the common conception of oscillating tangential forces as a slip-stick phenomenon where the sizes of the stick zone and the slip zone oscillate does not provide a complete picture of the non-steady contact problem. It is possible to have a situation where the tangential force and thus the tangential stress distribution and the slip oscillates even though the position of the limit between stick zone and slip zone is kept constant.

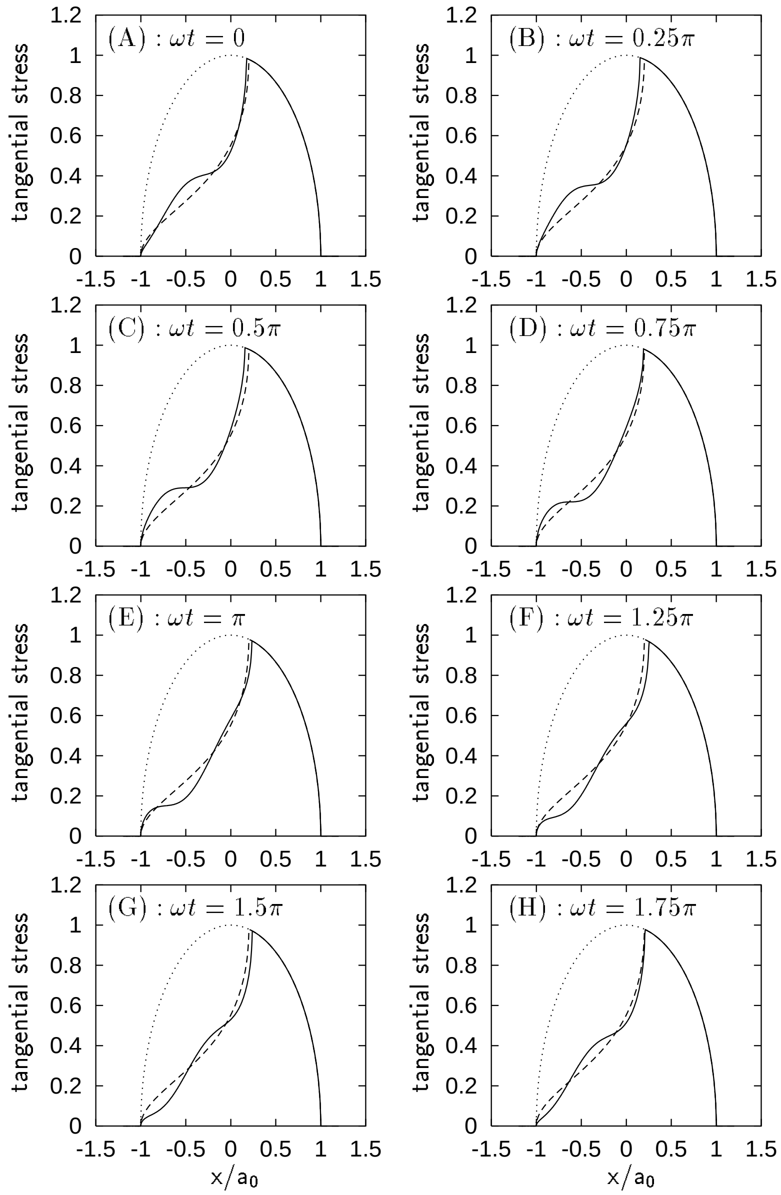

Figure 7.1. (A)-(H): Tangential stress distribution for the case $a_0/L = 1$. Solid line: non-steady solution. Dashed line: Carter solution.

With the expressions for $a^*(t)$ and $T(t)$ derived it is possible to establish a relation between the tangential force and the creepage. Inserting the approximation for $u_x(x_0,t)$ into the kinematic constraint in equation (7.10) this leads to the result that

$$\xi(t) = \xi_0 + \sum_{m=)}^{M} \left[\xi_{A,m}\cos(\omega_m t) + \xi_{B,m}\sin(\omega_m t)\right] \tag{7.44}$$

where

$$\left\{ \begin{array}{c} \xi_{A,m} \\ \xi_{B,m} \end{array} \right\} = \left[\begin{array}{cc} \xi_{1,m} & -\xi_{2,m} \\ \xi_{2,m} & \xi_{1,m} \end{array} \right] \left\{ \begin{array}{c} U_A \\ U_B \end{array} \right\} \tag{7.45}$$

The coefficients $\xi_{1,m}$ and $\xi_{2,m}$ obviously depend on which approximation of $u_x(x_0,t)$ there is utilized. The approach based on the modified version of Johnson's stress distribution results in the coefficients

$$\xi_0 = 0 \tag{7.46}$$

$$\xi_{1,m} = \mu(k_m a_0)^2 \left(\mathrm{J}_0(k_m a_0)\left[\frac{1}{1-\nu} + \ln\left(\frac{2b}{a_0}\right)\right] - \mathrm{J}_0(k_m a_0) \right) \tag{7.47}$$

$$\xi_{2,m} = \mu(k_m a_0)^2 \left(\mathrm{J}_1(k_m a_0)\left[\frac{1}{1-\nu} + \ln\left(\frac{2b}{a_0}\right)\right] - \frac{1}{k_m a_0}\mathrm{J}_0(k_m a_0) \right) \tag{7.48}$$

provided that $a_0^* = a_0$. The fact that the size of the stick zone oscillates even when $a_0^* = a$ is due to fact that the location of the slip zone changes between the leading edge and the trailing edge of the contact patch. Thus, a value of a^* greater than a is interpreted as a situation where a slip zone with half the length $(a^* - a)$ is located at the leading edge of the contact patch.

The approximation of $u_x(x_0, t)$ based on the tangential stress distribution similar to the Carter solution (equation (7.16)) yields that

$$\xi_0 = \frac{\mu a_0}{R}\left(r_{a_0} - 1\right) \tag{7.49}$$

$$\xi_{1,m} = \frac{\mu}{R}\left(a^*_{1,m} + k_m a_0 r_{a_0} a^*_{2,m}\right) + \frac{4(1-\nu^2)}{\pi E} k_m T_{2,m}\left[\frac{1+\nu}{2(1-\nu)} + \ln\left(\frac{4b}{a^*_0}\right)\right] \tag{7.50}$$

$$\xi_{2,m} = \frac{\mu}{R}\left(a^*_{2,m} - k_m a_0 r_{a_0} a^*_{1,m}\right) - \frac{4(1-\nu^2)}{\pi E} k_m T_{1,m}\left[\frac{1+\nu}{2(1-\nu)} + \ln\left(\frac{4b}{a^*_0}\right)\right] \tag{7.51}$$

It is seen that for $r_{a_0} = 1$ the solution for the approach based on the modified Johnson stress distribution is almost identical to the solution where $u_x(x_0, t)$ is found from a Carter stress distribution. This must evidently be the case as the Johnson stress distribution can be considered as a linearization of the Carter solution for the case where $r_{a_0} = 1$ with the only difference that the first approximation is made for a uniform stress distribution whereas the latter approximation is derived from an elliptic stress distribution.

By this the non-steady contact problem is solved. The obvious similarities between the above derived expressions for the tangential stress distribution and for the tangential force and the same quantities for the case of a cylinder rolling on a corrugated surface (see section 4.2.3) should be noticed. Replacing $U_{A,m}$ with $Z_{A,m}$ and $U_{B,m}$ with $Z_{B,m}$ the stress distribution and the tangential force are identical with the ones from the corrugated surface. Consequently the solution to the stationary tangential problem for the case of a cylinder rolling on a corrugated surface which was derived in section 4.2.3 is also a solution to the non-steady contact problem when it comes to the tangential force and the tangential stress distribution. The omitting of the time-derivative in the kinematic constraint only affects the size of the creepage, which then is oscillating instead of being constant.

Another interesting feature is that the quantity $u_x(x_0, t)$ does not influence the tangential stress distribution or the tangential force. The approximation of $u_x(x_0, t)$ does only affect the creepage, and so the derived solutions for $q(x, t)$ and $T(t)$ are exact, even though the two-dimensional non-steady contact problem by definition is unsolvable.

To analyse the relation between the creepage and the tangential force the ratio $T(t)/\xi(t)$ is investigated. In order to simplify the calculations it is assumed that $U_x(x + V_m t)$ only has one distinct wave length i.e.

$$U_x(x + V_m t) = \mu k \left[U_A \cos[k(x + V_m t)] + U_B \sin[k(x + V_m t)]\right] \qquad (7.52)$$

The ratio between the amplitudes of $T(t)$ and $\xi(t)$ then reads

$$\frac{\widehat{T}}{\widehat{\xi}} = \left(\frac{T_1^2 + T_2^2}{\xi_1^2 + \xi_2^2}\right)^{\frac{1}{2}} \qquad (7.53)$$

and the difference in phase is given as

$$\phi_T - \phi_\xi = \arctan\left(\frac{T_2}{T_1}\right) - \arctan\left(\frac{\xi_2}{\xi_1}\right) \qquad (7.54)$$

It is seen that both the ratio of the amplitudes $\widehat{T}/\widehat{\xi}$ and the difference in phase $(\phi_T - \phi_\xi)$ depend on the two ratios a_0/L and $r_{a_0} = a_0^*/a_0$ i.e. on the relative size of the contact patch and the magnitude of the reference creepage or the reference tangential force. The coefficients $U_{A,m}$ and $U_{B,m}$ do not influence the ratio of the amplitudes or the difference in phase, a property which is quite obvious as no restrictions concerning these coefficients are given in the derivation of the theory.

Groß-Thebing has derived an analytical expression for the ratio $\widehat{T}/\widehat{\xi}$ for the Johnson approach where the stick zone covers the entire contact patch [23]. Defining the tangential force and the creepage as the complex quantities

$$T(t) = \widehat{T} e^{i\omega t} \tag{7.55}$$

$$\xi(t) = \widehat{\xi} e^{i\omega t} \tag{7.56}$$

where $i = \sqrt{-1}$ and $\widehat{T}$ and $\widehat{\xi}$ are complex constants, Groß-Thebing demonstrates that

$$\frac{\widehat{T}}{\widehat{\xi}} = \frac{\pi E a}{4(1-\nu^2)} \left[\frac{1}{\frac{\mathrm{J}_0(ka)}{\mathrm{J}_0(ka)+i\mathrm{J}_1(ka)} + ika\left[\frac{1}{1-\nu} + \ln\left(\frac{2b}{a}\right)\right]} \right] \tag{7.57}$$

This solution is shown in Figure 7.2 (dashed line) where it is compared with the approach based on the modified Johnson stress distribution (solid line). The size of the pseudo contact width is set to be $b = 20a_0$. It is seen that the two solutions are identical when $a_0/L \rightarrow 0$ which implies that the wave length is very large compared with the size of the stick zone i.e. the influence of the non-steady term in the kinematic constraint vanishes. Furthermore the two solutions are identical when $J_0(ka_0) = 0$ which exactly are the cases where the size of the stick zone is constant in time and thus covers the entire contact patch. The discrepancies between the two solutions for all other values of a_0/L arises from the fact that Groß-Thebing's solution has a singularity at the trailing edge of the contact patch as a result of the complete stick assumption, whereas the approach based on the modified Johnson stress distribution has incorporated a small slip zone at the trailing edge in order to avoid this singularity.

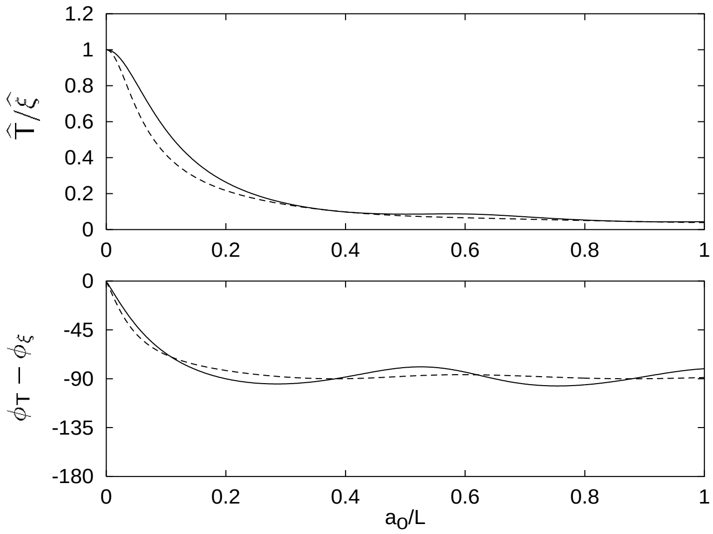

Figure 7.2. Non-steady contact for the case $r_{a_0} = 1$. Dashed line: Groß-Thebing's solution (7.57). Solid line: approach based on the modified Johnson stress distribution. Top: ratio between the amplitude of the tangential force and the amplitude of the creepage normalised with $a_0^* \pi E/(4(1-\nu^2))$. Bottom: difference between the phase of the tangential force and the phase of the creepage.

In general the Groß-Thebing solution is quite accurate, also when it comes to contact situations where $r_{a_0} \neq 1$. In Figure 7.3 the ratio of the amplitudes and the difference in phase according to the solution based on the Carter stress distribution are compared with the same values given by Groß-Thebing's approach where a_0 is substituted by a_0^*. The size of the pseudo contact width is set to be $b = 20a_0$. It is seen that the Groß-Thebing theory actually provides a very accurate result also when the stick zone does not cover the entire contact patch. This is caused by the fact that the qualitative behaviour of the problem does not differ significantly when r_{a_0} changes e.g. the outline of $\widehat{T}/\widehat{\xi}$ and $(\phi_T - \phi_\xi)$ for $r_{a_0} = 0.1$ is not very different from the outline of the same quantities when $r_{a_0} = 0.8$ (see Figure 7.3 (A) and Figure 7.3 (H)). Thus, the non-steady behaviour of the contact problem is not very sensitive to the stationary size of the stick zone and the slip zone.

In the above example a non-steady contact problem with oscillating tangential force is solved employing a simple series expansion of $T(t)$ and $a^*(t)$. This approach is used in order to demonstrate the application of the derived non-steady theory. It is obviously possible to solve the problem deriving more sophisticated closed form expressions for $T(t)$ and $a^*(t)$, just as the problem also can be solved if other quantities such as the normal force, the curvature or the material properties oscillate as the cylinder rolls along the surface. In all cases the expressions for the tangential stress distribution and the tangential force are exact whereas the creepage is found using an approximation of the value $u_x(x_0, t)$. The only constraints necessary to solve the non-steady problem are that $U_x = U_x(x + V_m t)$ and that the contact patch is divided into one stick zone and one slip zone.

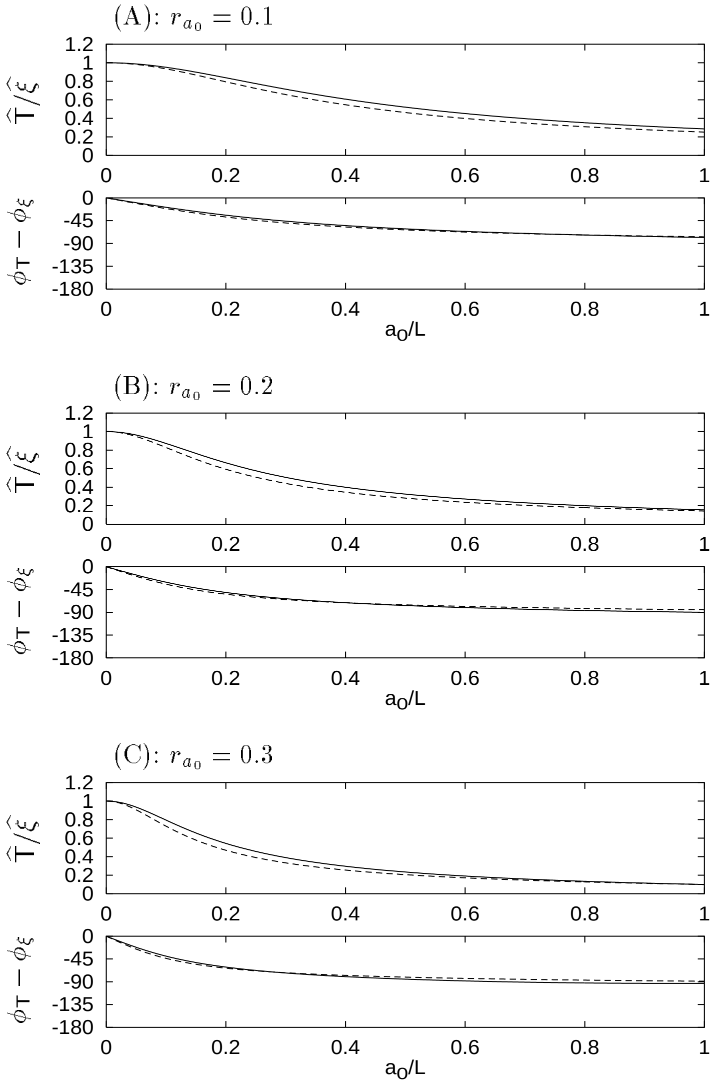

Figure 7.3. (A)-(C) Non-steady contact. Dashed line: Groß-Thebing's solution (7.57). Solid line: approach with Carter's stress distribution. Top: ratio between the amplitude of the tangential force and the amplitude of the creepage normalised with $a_0^* \pi E/(4(1-\nu^2))$. Bottom: difference between the phase of the tangential force and the phase of the creepage.

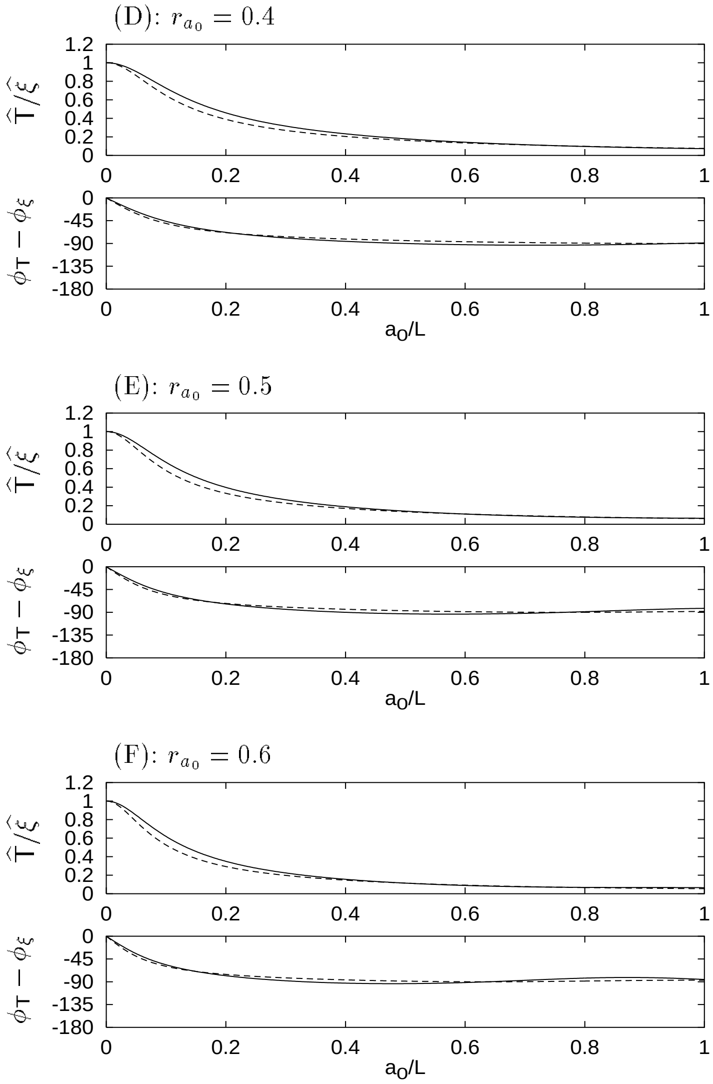

Figure 7.3. (D)-(F) Non-steady contact. Dashed line: Groß-Thebing's solution (7.57). Solid line: approach with Carter's stress distribution. Top: ratio between the amplitude of the tangential force and the amplitude of the creepage normalised with $a_0^* \pi E/(4(1-\nu^2))$. Bottom: difference between the phase of the tangential force and the phase of the creepage.

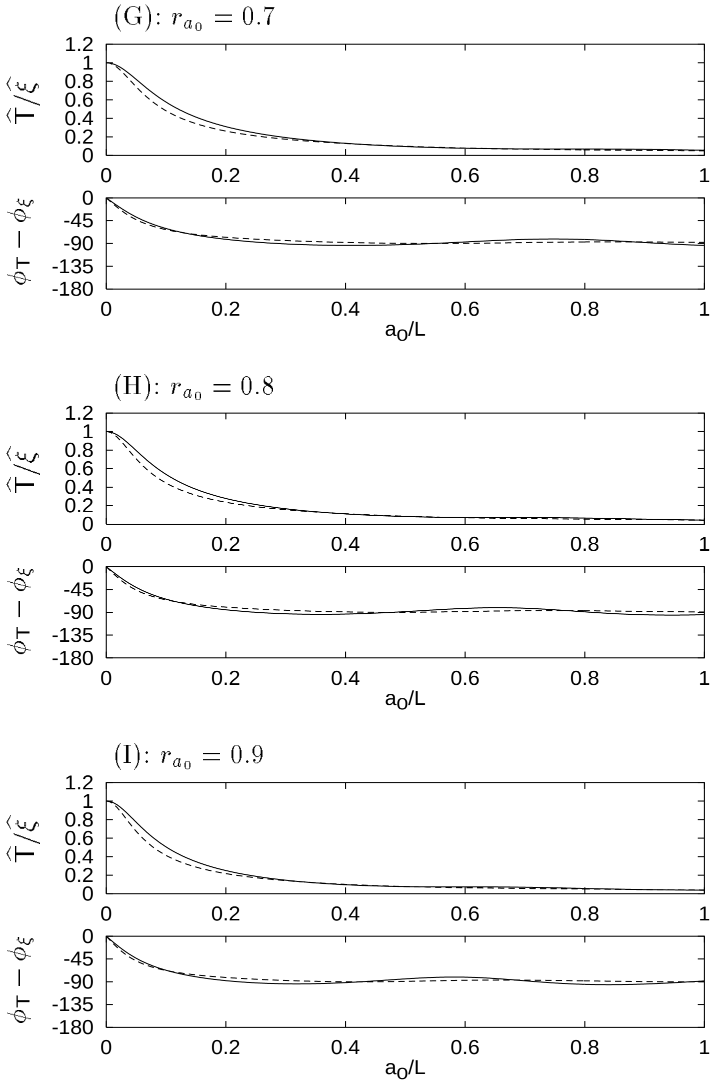

Figure 7.3. (G)-(I) Non-steady contact. Dashed line: Groß-Thebing's solution (7.57). Solid line: approach with Carter's stress distribution. Top: ratio between the amplitude of the tangential force and the amplitude of the creepage normalised with $a_0^* \pi E/(4(1-\nu^2))$. Bottom: difference between the phase of the tangential force and the phase of the creepage.

Chapter 8

Conclusion

In the present work a two-dimensional contact model based on half space approximations has been formulated. The fundamental problem of the half space approach to a contact problem arises from the constitutive equation which provides the relation between the stress distribution $q(x)$ and the gradient of the relative displacement $du_x(x)/dx$:

$$\frac{du_x(x)}{dx} = \frac{4(1-\nu^2)}{\pi E} \int_{-a}^{a} \frac{q(\zeta)}{x-\zeta}\, d\zeta$$

When the stress distribution is known this equation does not cause any problems, but when the stress distribution has to be found from a given gradient of the relative displacement, the constitutive equation provides major difficulties.

It has in the present work been shown that

$$p(x) = \frac{\sum_{n=0}^{N} B_n x^n}{\sqrt{a^2 - x^2}} \quad \Leftrightarrow \quad \frac{du_z(x)}{dx} = \sum_{m=0}^{N-1} \beta_m x^m \quad , \quad |x| \leq a$$

where the coefficients B_n and β_m are linearly dependent. Thus, the constitutive equation originally formulated as an integral equation is reduced to an algebraic equation. Since this equation consists of polynomials the actual calculation is reduced to the comparison of polynomial coefficients, which simplifies the solution of the contact problem significantly. Because the constitutive equation for the normal pressure distribution and the gradient of the relative normal displacement is equivalent to the constitutive equation for the tangential stress distribution and gradient of the relative tangential displacement, both the normal contact problem and the tangential contact problem can be solved with the derived model.

The model has been applied on four different contact problems which cannot be solved by more primitive contact models which are nevertheless the most common in simulations investigating wheel/rail contact:

Contact between corrugated surfaces: It has been demonstrated that the case of a cylinder rolling on a corrugated surface is very sensitive to the choice of contact model. The presence of corrugation implies that the stress distribution becomes asymmetric, a property which is decisive for the evolution of the corrugation. This effect is disregarded if the tangential stress distribution is assumed to be equal to the one for the Carter solution.

Contact with velocity dependent friction coefficient: The derived model has been applied on the case where the friction coefficient depends on the local relative velocity between the surfaces in contact. The tangential stress distribution and the creep curve were calculated. If the friction coefficient is defined as a step function with one static value and one kinematic value, the outline of the creep curve is identical with the creep curve for the Carter solution where the friction coefficient is defined as the kinematic friction coefficient. When the friction coefficient is defined as a decaying function of the local relative velocity, the creep curve will have a maximum and then decay as the creepage increases. The location of this maximum can be determined.

Contact between rough surfaces: Employing the new contact model it can be demonstrated that the combination of curvature, skewness and flatness of a roughness spike is important for the normal contact problem, thus statistical representations of these contact properties for a given surface may be misleading in a contact mechanical sense as the contact situation is not unique for a given statistical representation. It is further demonstrated that the initial slope of the creep curve is proportional to the size of the actual contact patch. Finally it has been shown that the normal contact problem is more sensitive to the cross influence between adjacent contact patches than the tangential contact problem.

Non-steady contact: An approximative method for the case of non-steady contact has been derived. It is shown that an expression for the time dependent creepage depends on an approximation of the displacement at a reference point whereas the expressions for the tangential stress distribution and the tangential force are exact. Calculations show that the ratio between the amplitude of the creepage and the amplitude of the tangential force with appreciable accuracy can be found by a linear approach to the contact problem.

8.1 Further Investigations

It is obviously possible to solve combinations of the different contact problems with the derived model e.g. the case of a corrugated surface with a velocity dependent friction coefficient or the non-steady contact of rough surfaces. Due to the generality of the derived model the application can also be extended to many other cases of contact mechanics which are not covered by more conventional contact models. An example of such an application is the case where the material properties - e.g. the modulus of elasticity or the coefficient of friction - depends on the temperature in the contact patch. Due to the frictional work the temperature is not constant over the contact patch. If the temperature field is known and an expression for the material properties dependency of the temperature also is known, then both the normal contact problem and the tangential contact problem can be solved employing a method equivalent to the one described for the case of the velocity dependent friction coefficient.

The main weakness of the new contact model is evidently that it is a two-dimensional model. It has been demonstrated that the two-dimensional model can be extended also to cover three-dimensional contact situations with no spin by integrating the two-dimensional solution over an elliptic contact patch. An obvious continuation of the present work is to perform this integration for the four different contact problems in order to obtain a three-dimensional model for each case. A much harder task is to include the effect of spin in the contact model. It is obviously possible to introduce a correction factor depending on the spin which makes the creep curve resemble the numerically calculated creep curve (e.g. by CONTACT). However, the introduction of a correction factor implies that the strictly physical background of the derived expressions is eliminated. A much better way to include the spin in the model is by redefining the location of the strips in the strip theory. Thus, instead of having linear strips parallel to the rolling direction, curved strips parallel to the pseudo creepage vector must be introduced. This results in more complicated calculations but it is an interesting approach to the three-dimensional problem with spin which may provide very accurate results.

Appendix A

Minimum of the Two-Dimensional Solution

The following calculations will demonstrate that the Carter solution to the two-dimensional tangential problem is a minimum solution.

Consider a tangential stress distribution defined as the sum of two ellipses

$$q(x) = q_1(x) + q_2(x^*) \tag{A.1}$$

$$q_1(x) = \frac{\mu p_0}{a_0}\sqrt{a_0^2 - x^2} \tag{A.2}$$

$$q_2(x^*) = p_0^*\sqrt{a_0^{*2} - x^{*2}} \tag{A.3}$$

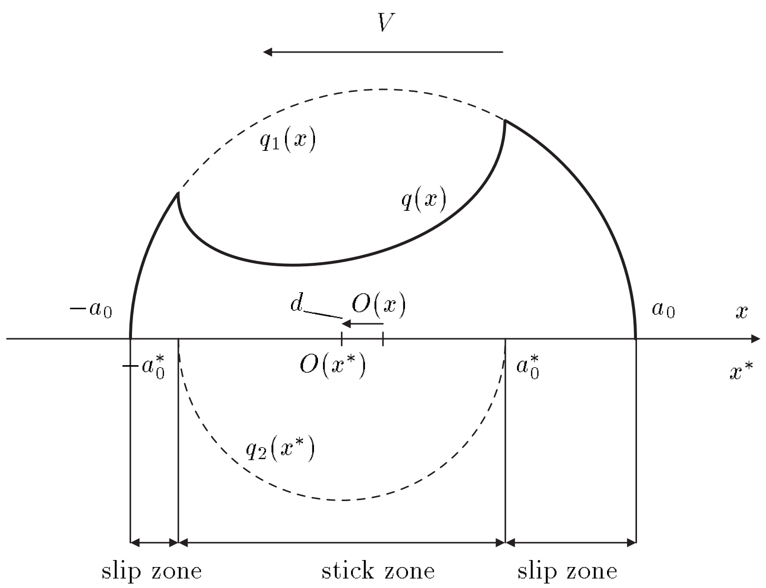

Figure A.1. Tangential stress distribution.

$$x^* = x + d \tag{A.4}$$

Then the gradient of the relative lateral displacement due to Cerruti-Boussinesq is:

$$\frac{du_x(x)}{dx} = -\frac{4(1-\nu^2)}{\pi E}\left[\frac{\pi\mu p_0}{a_0}x + \pi p_0^*(x+d)\right] \quad , \quad -a_0^* < x + d < a_0^* \tag{A.5}$$

which inserted into the kinematic constraint yields that

$$\xi_0 = \frac{4(1-\nu^2)}{\pi E}\left[\frac{\pi\mu p_0}{a_0}x + \pi p_0^*(x+d)\right] \tag{A.6}$$

As ξ_0 is independent of x this implies that

$$p_0^* = -\frac{\mu p_0}{a_0} \tag{A.7}$$

and so

$$\xi_0 = -\frac{4(1-\nu^2)}{E}\frac{\mu p_0}{a_0}d \tag{A.8}$$

The resulting tangential force is found by integrating $q(x)$

$$T = \frac{1}{2}\frac{\pi \mu p_0}{a_0}\left[a_0^2 - a_0^{*2}\right] \tag{A.9}$$

This implies that the minimum tangential force is found by maximizing a_0^{*2}.

The only restriction for a_0^* and d is that

$$a_0^* + |d| \leq a_0 \tag{A.10}$$

i.e. the ellipse $q_2(x^*)$ always lies inside the ellipse $q_1(x)$, and so the maximum value of a_0^* is

$$a_{0,max}^* = a_0 - |d| \tag{A.11}$$

and consequently

$$d = a_0 - a_{0,max}^* \tag{A.12}$$

which is equivalent to the Carter solution. This demonstrates that under the assumption that the tangential stress distribution can be expressed as the sum of two ellipses, the Carter solution is the only minimal solution.

Appendix B

Kalker's Creep Coefficients

The coefficients are listed for a Poisson ratio $\nu = 0.25$.

a_0 is the semi axis of the contact ellipse in the rolling direction.

b_0 is the semi axis of the contact ellipse perpendicular to the rolling direction.

The derivation of the creep coefficients can be found in [34].

a_0/b_0	C_{11}	C_{22}	C_{23}
0.1	3.31	2.52	0.473
0.2	3.37	2.63	0.603
0.3	3.44	2.75	0.715
0.4	3.53	2.88	0.823
0.5	3.62	3.01	0.929
0.6	3.72	3.14	1.03
0.7	3.81	3.28	1.14
0.8	3.91	3.41	1.25
0.9	4.01	3.54	1.36

b_0/a_0	C_{11}	C_{22}	C_{23}
1.0	4.12	3.67	1.47
0.9	4.22	3.81	1.59
0.8	4.36	3.99	1.75
0.7	4.54	4.21	1.95
0.6	4.78	4.50	2.23
0.5	5.10	4.90	2.62
0.4	5.57	5.48	3.24
0.3	6.34	6.40	4.32
0.2	7.78	8.14	6.63
0.1	11.7	12.8	14.6

Appendix C

Constitutive Equations

The below listed constitutive equations for the loaded half space are derived for quasi identical bodies. A derivation can be found in [47].

Two-dimensional case:

$$u_z(x) = -\frac{4(1-\nu^2)}{\pi E} \int p(\zeta) \ln |x-\zeta| \, d\zeta + C_1 \tag{C.1}$$

$$u_x(x) = -\frac{4(1-\nu^2)}{\pi E} \int q(\zeta) \ln |x-\zeta| \, d\zeta + C_2 \tag{C.2}$$

$$\frac{du_z(x)}{dx} = -\frac{4(1-\nu^2)}{\pi E} \int \frac{p(\zeta)}{x-\zeta} \, d\zeta \tag{C.3}$$

$$\frac{du_x(x)}{dx} = -\frac{4(1-\nu^2)}{\pi E} \int \frac{q(\zeta)}{x-\zeta} \, d\zeta \tag{C.4}$$

Three-dimensional case:

$$u_z(x,y) = \frac{2(1-\nu^2)}{\pi E} \iint \frac{p(\zeta,\eta)}{\rho}\, d\zeta d\eta \tag{C.5}$$

$$u_x(x,y) = \iint \{g_{11}(x,y,\zeta,\eta)\, q_x(\zeta,\eta) + g_{12}(x,y,\zeta,\eta)\, q_y(\zeta,\eta)\}\, d\zeta d\eta \tag{C.6}$$

$$u_y(x,y) = \iint \{g_{21}(x,y,\zeta,\eta)\, q_x(\zeta,\eta) + g_{22}(x,y,\zeta,\eta)\, q_y(\zeta,\eta)\}\, d\zeta d\eta \tag{C.7}$$

$$g_{11}(x,y,\zeta,\eta) = \frac{2(1+\nu)}{\pi E}\left[\frac{(1-\nu)}{\rho} + \frac{\nu(x-\zeta)^2}{\rho^3}\right] \tag{C.8}$$

$$g_{12}(x,y,\zeta,\eta) = \frac{2(1+\nu)}{\pi E}\left[\frac{\nu(x-\zeta)(y-\eta)}{\rho^3}\right] \tag{C.9}$$

$$g_{21}(x,y,\zeta,\eta) = \frac{2(1+\nu)}{\pi E}\left[\frac{\nu(x-\zeta)(y-\eta)}{\rho^3}\right] \tag{C.10}$$

$$g_{22}(x,y,\zeta,\eta) = \frac{2(1+\nu)}{\pi E}\left[\frac{(1-\nu)}{\rho} + \frac{\nu(y-\eta)^2}{\rho^3}\right] \tag{C.11}$$

$$\rho = \sqrt{(x-\zeta)^2 + (y-\eta)^2} \tag{C.12}$$

Appendix D

Transformation of the Constitutive Equation

The aim of the present investigation is to derive a general solution to the integral

$$I = \int_{-a}^{a} \frac{\sum_{n=0}^{N} B_n \zeta^n}{(x-\zeta)\sqrt{a^2-\zeta^2}}\, d\zeta \tag{D.1}$$

Introducing the transformation $\eta = x - \zeta$ then the polynomial is rewritten as

$$\sum_{n=0}^{N} B_n \zeta^n = \sum_{n=0}^{N} B_n \sum_{i=0}^{n} (-1)^i \left(\begin{array}{c} n \\ i \end{array} \right) x^{n-i} \eta^i \tag{D.2}$$

and so the integral (D.1) is transformed into

$$I = \sum_{n=0}^{N} B_n \sum_{i=0}^{n} (-1)^i \left(\begin{array}{c} n \\ i \end{array} \right) x^{n-i} \int_{x+a}^{x-a} \frac{\eta^{i-1}}{\sqrt{-\eta^2 + 2x\eta + a^2 - x^2}}\, d\eta \tag{D.3}$$

Now let I_{i-1} denote the integral

$$I_{i-1} = \int_{x+a}^{x-a} \frac{\eta^{i-1}}{\sqrt{-\eta^2 + 2x\eta + a^2 - x^2}}\, d\eta \tag{D.4}$$

then the value of I_i is given by the recursive formulae

$$I_i = \left(\frac{1}{i} - 1\right)\left(x^2 - a^2\right) I_{i-2} + \left(2 - \frac{1}{i}\right) x I_{i-1} \quad , \quad i = 1, 2, \dots \tag{D.5}$$

$$I_{-1} = \begin{cases} 0 & , \quad |x| \le |a| \\ \frac{\text{sign}(x)\pi}{\sqrt{x^2 - a^2}} & , \quad |x| > |a| \end{cases} \tag{D.6}$$

$$I_0 = -\pi \tag{D.7}$$

This implies that if $|x| \le |a|$ then I_{i-1} can be written as a polynomial in x

$$I_{i-1} = \sum_{j=1}^{i} \alpha_{ij} x^{j-1} \tag{D.8}$$

where the coefficients are given as

$$\alpha_{ij} = \begin{cases} -\pi \frac{(i-1)!}{(j-1)!\left[\left(\frac{1}{2}(i-j)\right)!\right]^2} \left(\frac{a}{2}\right)^{i-j} & , \quad \frac{i-j}{2} \in \mathbb{N}_0 \\ 0 & , \quad \text{otherwise} \end{cases} \tag{D.9}$$

Inserted into equation (D.3) this yields the new equation

$$I = \sum_{n=0}^{N} B_n \sum_{i=1}^{n} (-1)^i \begin{pmatrix} n \\ i \end{pmatrix} x^{n-i} \sum_{j=1}^{i} \alpha_{ij} x^{j-1} \tag{D.10}$$

i.e. a polynomial in x with the coefficients β_m:

$$I = \sum_{m=0}^{N-1} \beta_m x^m \tag{D.11}$$

Comparing the coefficients of the polynomials in equation (D.10) and equation (D.11) this implies that

$$j = m - n + i + 1 \tag{D.12}$$

which also yields the restriction that $(n-m-1)/2 \in \mathbb{N}_0$. It is now possible to derive a relation between the B_n's and the β_m's

$$\beta_m = \sum_{n=m+1}^{N} B_n \sum_{i=n-m}^{n} (-1)^i \binom{n}{i} x^{n-i} \alpha_{i,m-n+i+1} \tag{D.13}$$

which can be reduced to

$$\beta_m = \sum_{n=m+1}^{N} B_n A_k \quad , \quad k \in \mathbb{N}_0 \tag{D.14}$$

$$A_k = -\pi \frac{(2k)!}{k!} \left(\frac{a}{2}\right)^{2k} \tag{D.15}$$

$$k = \frac{n-m-1}{2} \tag{D.16}$$

This demonstrates that there exists a linear relation between the B_n's and the β_m's only depending on a and $n-m$. The above derivation is made with the assumption that $I_{-1} = 0$ i.e. for the case where $|x| \leq |a|$. When I_{-1} is included the general solution to the integral (D.1) is given as

$$\int_{-a}^{a} \frac{\sum_{n=0}^{N} B_n \zeta^n}{(x-\zeta)\sqrt{a^2-\zeta^2}} \, d\zeta = \sum_{m=0}^{N-1} \beta_m x^m + I_{-1} \sum_{n=0}^{N} B_n \tag{D.17}$$

Bibliography

[1] Ahlbeck, D.R. and Daniels, L.E. (1991): *Investigations of Rail Corrugations on the Baltimore Metro.* 2nd Mini Conference on Contact Mechanics and Wear of Wheel/Rail Systems, Budapest, pp. 197–210.

[2] Alias, J. (1986): *Characteristics of wave formation in rails.* Rail International, 17(11), pp. 17 – 23.

[3] Baumann, G., Grohmann, H.D. and Knothe, K. (1996): *Wirkungsketten bei der Ausbildung kurzwelliger Riffeln auf Schienenlaufflächen.* Eisenbahntechnische Rundschau, 45, pp. 792 – 798.

[4] Boussinesq, J. (1885): *Applications des Potential à l'etude de l'èquilibre er du mouvement des solides èlastiques.* Gauthier-Villars, Paris.

[5] Bowden, F.P. and Tabor, D. (1950): *The Friction and Lubrication of Solids.* Part 1, Clarendon Press, Oxford.

[6] Bower, A.F. and Johnson, K.L. (1991): *Plastic Flow and Shakedown of the Rail Surface in Repeated Wheel-Rail Contact.* 3rd International Conference on Contact Mechanics and Wear of Rail/Wheel Systems, Cambridge July 22–26, pp. 1–18.

[7] Brommundt, E. (1991): *Ein Reibschwinger mit Selberregung ohne fallende Reibkennlinie.* Zeitschrift für Angewandte Mathematik und Mechanik, 75(11), pp. 811-820.

[8] Carneiro Esteves,A. (1987): *Resolution du Contact Elastique entre Deux Corps Rugueux.* Thesis no. 87 ISAL 0047, L'Institut National de Sciences Appliqueés de Lyon, France.

[9] Carter, F.W. (1926): *On the action of a locomotive driving wheel.* Proceedings of the Royal Society of London, A112, pp. 151–157.

[10] Cattaneo, C. (1938): *Sul contatto di due corpi elastici: distribuzione locale degli sforzi.* Academia Nazionale Lincei, Rendiconti, Ser. 6, Vol. XXVII, pp. 342–348, 434–436, 474–478.

[11] Cerruti, V. (1882): *Academia dei Lincei, Roma.* Mem. fis. mat.

[12] Clark, R.A.; Dean, P.A.; Elkins, J.A. and Newton, S.G. (1982): *An investigation into the dynamic effects of railway vehicles running on corrugated rails.* Journal of Mechanical Engineering Society, 24, pp. 65 – 76.

[13] Clayton, P. and Allery, M.B.P. (1982): *Metallurgical aspects of surface damage problems in rails.* Canadian Metallurgical Quarterly, 21(1), pp. 31–46.

[14] Dahlberg, T.; Fenander, A.; Igeland, A and Nielsen, J. (1991): *Railway vehicle on randomly profiled track.* Chalmers University of Technology, Solid Mechanics Report F150, Göteborg, Sweden.

[15] DANKOWICZ, H. (1997): *Dynamical Friction Modelling. submitted to* ASME – Journal of Tribology.

[16] EKBERG, A. and BJARNEHED, H. (1995): *Rolling Contact Fatigue of Wheel/Rail Systems – A Literature Survey.* Report F182. Chalmers University of Technology, Division of Solid Mechanics, Göteborg, Sweden.

[17] FREDERIC, C.O. (1986): *A Rail Corrugation Theory.* 2nd International Conference on Contact Mechanics and Wear of Rail/Wheel Systems, Kingston, Rhode Island, pp. 181–211.

[18] FROMM, H. (1927): *Berechnung des Schlüpfes beim Rollen deformierbaren Scheiben.* ZAMP, 7, pp. 27–58.

[19] GRASSIE, S.L. and KALOUSEK, J. (1993): *Rail Corrugation: Characteristics, Causes and Treatments.* Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit, 207, pp. 57–68.

[20] GREENWOOD, J.A. and WILLIAMSON, J.P.B. (1966): *Contact of Nominally Flat Surfaces.* Proceedings of the Royal Society London, A295, pp. 300 –319.

[21] GREENWOOD, J.A. and TRIPP, J.H. (1967): *The Elastic Contact of Rough Spheres.* Trans. ASME, series E, Journal of Applied Mechanics, 34. pp. 300 –319.

[22] GROSS-THEBING, A. (1989): *Frequency Dependent Creep Coefficients for Three-Dimensional Rolling Contact Problems.* Vehicle System Dynamics, 18, pp. 359–374.

[23] Gross-Thebing, A. (1993): *Lineare Modellierung des instationären Rollkontaktes von Rad und Schiene.* VDI - Fortschrittsbericht, 199, Reihe 12, VDI Verlag Düsseldorf.

[24] Haines, D.J. and Ollerton, E. (1963): *Contact stress distributions on elliptical contact surfaces subjected to radial and tangential forces.* Proceedings, institution of Mechanical Engineers, 177, pp. 95–114.

[25] Hecht, M. (1995): *Kurvenkreischen – Ursachen und Gegenmassnahmen.* Schweizer Eisenbahn-Revue, 3, pp.103–108.

[26] Hempelmann, K. (1995): *Short Pitch Corrugation on railway rails - A linear model for prediction.* VDI - Fortschrittsbericht, 231, Reihe 12, VDI Verlag Düsseldorf.

[27] Hertz, H. (1882): *Über die Berührung fester elastischer Körper.* Journal für die reine und angewandte Mathematik, 92, pp. 156–171.

[28] Hill, R. (1950): *Theory of Plasticity.* University Press, Oxford.

[29] Igeland, A. (1997): *Dynamic Train/Track Interaction: Simulation of Railhead Corrugation Growth under a Moving Bogie Using Mathematical Models Combined with Full-Scale Measurements.* Ph.D.-Thesis,Chalmers University of Technology, Division of Solid Mechanics, Göteborg, Sweden.

[30] Ilias, H. (1996): *Nichtlineare Wechselwirkungen von Radsatz und Gleis beim Überrollen von Profilstörungen.* VDI - Fortschrittsbericht, 297, Reihe 12, VDI Verlag Düsseldorf.

[31] Johnson, K.L. (1958): *The effect of spin upon the rolling motion of an elastic sphere upon a plane.* Journal of Applied Mechanics, 25, pp. 332–338.

[32] Johnson, K.L. (1985): *Contact Mechanics.* Cambridge University Press.

[33] Kalker, J.J. (1966): *Rolling with Slip and Spin in the Presence of Dry Friction.* Wear, 9, pp. 20–38.

[34] Kalker, J.J. (1967): *On the rolling contact of two elastic bodies in the presence of dry friction.* Doctoral Dissertation, Technical University of Delft.

[35] Kalker, J.J. (1967): *A Strip Theory for Rolling with Slip and Spin.* Proceedings, Koninklijke Nederlandse Akademie van Wetenschappen, Amsterdam, B70, pp. 10–62.

[36] Kalker, J.J. (1970): *Transient Phenomena in Two Elastic Cylinders Rolling Over Each Other With Dry Friction.* Journal of Applied Mechanics, September 1970, pp.677–688.

[37] Kalker, J.J. (1971): *A Minimum Principle for the Law of Dry Friction, With Application to Elastic Cylinders in Rolling Contact. Part 1–2.* Journal of Applied Mechanics, 38, pp. 875–887.

[38] Kalker, J.J. (1990): *Three-Dimensional Elastic Bodies in Rolling Contact.* Solid Mechanics and its Applications, Vol.2, Kluwer Academic Publishers, Dordrecht.

[39] Kalker, J.J. (1996): *Book of Tables for the Hertzian Creep-Force.* 2nd Mini Conference on Contact Mechanics and Wear of Wheel/Rail Systems, Budapest, pp. 11–20.

[40] Knothe, K. and Gross-Thebing, A. (1986): *Derivation of Frequency Dependent Creep Coefficient Based on an Elastic Half-Space Model.* Vehicle System Dynamics, 15, pp. 133–153.

[41] Knothe, K. and Grassie, S.L. (1993): *Modelling of Railway Track and Vehicle/Track Interaction at High Frequencies*. Vehicle System Dynamics, 22(3–4), pp. 209–262.

[42] Knothe, K. and Theiler, A. (1996): *Normal and Tangential Contact Problems with Rough Surfaces*. 2nd Mini Conference on Contact Mechanics and Wear of Wheel/Rail Systems, Budapest, pp. 34–43.

[43] Kraft, K. (1967): *Der Einfluß der Fahrgeschwindigkeit auf dem Haftwert zwischen Rad und Schiene*. Archiv für Eisenbahntechnik, 22, pp. 58–78.

[44] Kragelski, I.V. (1965): *Friction and Wear*. Butterworth, Washington D.C.

[45] Kumagai, N. *et. al.* (1991): *Factors of Wheel Flats Occurence and Preventive Measures*. 3rd International Conference on Contact Mechanics and Wear of Rail/Wheel Systems, Cambridge July 22–26, pp. 277–287.

[46] Lang, W. and Roth, G. (1993): *Optimale Kraftschlußausnutzung bei Hochleistungs-Schienenfahrzeugen*. Eisenbahn-technische Rundschau, 42, pp. 61–66.

[47] Love, A.E.H. (1952): *A Treatise on the Mathematical Theory of Elasticity*. 4th Edition, Cambridge University Press.

[48] Nackenhorst, U. (1993): *Berechnung von Rollkontaktproblemen mit der Finite Element Methode*. Zamm, 73, pp. 363–366.

[49] Nielsen, J.B. and Theiler, A. (1996): *Tangential Contact Problem with Friction Coefficients Depending on Sliding Velocity*. 2nd Mini Conference on Contact Mechanics and Wear of Wheel/Rail Systems, Budapest, pp. 44–51.

[50] Nielsen, J.B. (1997): *A Nonlinear Wear Model.* ASME Rail Transportation, RTD-Vol. 13, pp. 7–12.

[51] Nielsen, J.B. (1997): *Evolution of Rail Corrugation Predicted with a Nonlinear Wear Model.* Submitted to Journal of Sound and Vibration.

[52] Nielsen, J.B. (1998): *Computational Methods in Contact Mechanics.* Proceedings of the IV International Congress of Computational Mechanics, Buenos Aires.

[53] Ohyama, T. (1989): *Some Basic Studies on the Influence of Surface Contamination on Adhesion Force Between Wheel and Rail at Higher Speeds.* Quarterly Report, Railway Technical Research Institute, 30(3), pp. 127–135.

[54] Ohyama,T. (1991): *Tribological Studies on Adhesion Phenomena Between Wheel and Rail at High Speeds.* Wear, 144, pp 263–275.

[55] Panagiotopoulos, P.D. (1975): *A nonlinear programming approach to the unilateral contact and friction-boundary value problem in the theory of elasticity.* Ingenieur-Archiv, 44, pp. 421–432.

[56] Pascal, J.P. and Sauvage, G. (1991): *New Method for Reducing the Multi-contact Wheel/Rail Problem to one Equivalent Rigid Contact Patch.* Proceedings from the 12th IAVSD-Symposium, Lyon, pp. 475–489.

[57] Piotrowsky, J. (1982): *A Theory of Wheelset Forces for Two Point Contact Between Wheel and Rail.* Vehicle System Dynamics, 11, pp. 69–87.

[58] Popp, K. (1991): *The Chaotic Motions of a Frictional Oscillator with Self- and External-excitation.* Zeitschrift für Angewandte Mathematik und Mechanik, 71(4).

[59] RABIBOWICZ:65 (1965): *Friction and Wear of Materials.* Wiley, New York.

[60] REMINGTON, P.J. (1976): *Wheel-Rail-Noise Part I-IV.* Journal of Sound and Vibration, 46(3), pp. 359–451.

[61] REYNOLDS, O. (1875): *On rolling friction.* Philosophical Transactions of the Royal Society, 166, pp. 155–174.

[62] RIPKE, B. (1995): *Hochfrequente Gleismodellierung un Simulation der Fahrzeug-Gleis-Dynamik unter Verwendung einer nichtlinearen Kontaktmechanik.* VDI - Fortschrittsbericht, 249, Reihe 12, VDI Verlag Düsseldorf.

[63] THOMAS, T.R. (1982): *Rough Surfaces.* Longman, London and New York.

[64] TIMOSHENKO, S. and GOODIER, J.N. (1951): *Theory of Elasticity,* 3rd Edition, McGraw–Hill, Tokyo.

[65] VALDIVIA, A. (1988): *Die Wechselwirkung zwischen hochfrequenter Rad-Schiene-Dynamik und ungleichförmigem Schienenverschleiß - ein lineares Modell.* VDI - Fortschrittsbericht, 93, Reihe 12, VDI Verlag Düsseldorf.

[66] VERMEULEN, P.J. and JOHNSON, K.L. (1964): *Contact of Nonspherical Bodies Transmitting Tangential Forces.* Journal of Applied Mechanics, 31, pp. 338–340.

[67] VOLLEBREGT, E.A.H. (1992): *Users Manual Con93.* Dept. of Math. and Comp. Science, Delft University of Technology.

[68] WANG, G. (1988): *Theorie und numerische Behandlung des allgemeinen rollenden Kontaktes zweier viscoelastischer Walzen.* VDI - Fortschrittsbericht, 165, Reihe 1, VDI Verlag Düsseldorf.

[69] WANG, G. (1991): *Rollkontakt zweier viskoelasticher Walzen mit Coulombscher Reibung*. Thesis, Technische Universität Berlin, Institut für Luft- und Raumfahrt.

[70] ZHANG, J. (1996): *Dynamisches Bogenlaufverhalten mit Stochastischen Gleislagefehlern*. VDI - Fortschrittsbericht, 304, Reihe 12, VDI Verlag Düsseldorf.

[71] ZIENKIEWICZ, O. (1988): *The Finite Element Method in Engineering Science*. 4th Ed., McGraw-Hill.

Ph. D. theses from IMM

1. **Larsen, Rasmus.** (1994). *Estimation of visual motion in image sequences.* *xiv* + 143 pp.
2. **Rygard, Jens Moberg.** (1994). *Design and optimization of flexible manufacturing systems.* *xiii* + 232 pp.
3. **Lassen, Niels Christian Krieger.** (1994). *Automated determination of crystal orientations from electron backscattering patterns.* *xv* + 136 pp.
4. **Melgaard, Henrik.** (1994). *Identification of physical models.* *xvii* + 246 pp.
5. **Wang, Chunyan.** (1994). *Stochastic differential equations and a biological system.* *xxii* + 153 pp.
6. **Nielsen, Allan Aasbjerg.** (1994). *Analysis of regularly and irregularly sampled spatial, multivariate, and multi-temporal data.* *xxiv* + 213 pp.
7. **Ersbøll, Annette Kjær.** (1994). *On the spatial and temporal correlations in experimentation with agricultural applications.* *xviii* + 345 pp.
8. **Møller, Dorte.** (1994). *Methods for analysis and design of heterogeneous telecommunication networks.* Volume 1-2, *xxxviii* + 282 pp., 283-569 pp.
9. **Jensen, Jens Christian.** (1995). *Teoretiske og eksperimentelle dynamiske undersøgelser af jernbanekøretøjer.* ATV Erhvervsforskerprojekt EF 435. *viii* + 174 pp.

10. **Kuhlmann, Lionel.** (1995). *On automatic visual inspection of reflective surfaces.* ATV Erhvervsforskerprojekt EF 385. Volume 1, *xviii* + 220 pp., (Volume 2, *vi* + 54 pp., fortrolig).

11. **Lazarides, Nikolaos.** (1995). *Nonlinearity in superconductivity and Josephson Junctions.* *iv* + 154 pp.

12. **Rostgaard, Morten.** (1995). *Modelling, estimation and control of fast sampled dynamical systems.* *xiv* + 348 pp.

13. **Schultz, Nette.** (1995). *Segmentation and classification of biological objects.* *xiv* + 194 pp.

14. **Jørgensen, Michael Finn.** (1995). *Nonlinear Hamiltonian systems.* *xiv* + 120 pp.

15. **Balle, Susanne M.** (1995). *Distributed-memory matrix computations.* *iii* + 101 pp.

16. **Kohl, Niklas.** (1995). *Exact methods for time constrained routing and related scheduling problems.* *xviii* + 234 pp.

17. **Rogon, Thomas.** (1995). *Porous media: Analysis, reconstruction and percolation.* *xiv* + 165 pp.

18. **Andersen, Allan Theodor.** (1995). *Modelling of packet traffic with matrix analytic methods.* *xvi* + 242 pp.

19. **Hesthaven, Jan.** (1995). *Numerical studies of unsteady coherent structures and transport in two-dimensional flows.* Risø-R-835(EN) 203 pp.

20. **Slivsgaard, Eva Charlotte.** (l995). *On the interaction between wheels and rails in railway dynamics.* *viii* + 196 pp.

21. **Hartelius, Karsten.** (1996). *Analysis of irregularly distributed points.* *xvi* + 260 pp.

22. **Hansen, Anca Daniela.** (1996). *Predictive control and identification - Applications to steering dynamics.* *xviii* + 307 pp.

23. **Sadegh, Payman.** (1996). *Experiment design and optimization in complex systems.* *xiv* + 162 pp.

24. **Skands, Ulrik.** (1996). *Quantitative methods for the analysis of electron microscope images.* *xvi* + 198 pp.

25. **Bro-Nielsen, Morten.** (1996). *Medical image registration and surgery simulation.* *xxvii* + 274 pp.

26. **Bendtsen, Claus.** (1996). *Parallel numerical algorithms for the solution of systems of ordinary differential equations.* *viii* + 79 pp.

27. **Lauritsen, Morten Bach.** (1997). *Delta-domain predictive control and identification for control.* *xxii* + 292 pp.

28. **Bischoff, Svend.** (1997). *Modelling colliding-pulse mode-locked semiconductor lasers.* *xxii* + 217 pp.

29. **Arnbjerg-Nielsen, Karsten.** (1997). *Statistical analysis of urban hydrology with special emphasis on rainfall modelling.* Institut for Miljøteknik, DTU. *xiv* + 161 pp.

30. **Jacobsen, Judith L.** (1997). *Dynamic modelling of processes in rivers affected by precipitation runoff.* *xix* + 213 pp.

31. **Sommer, Helle Mølgaard.** (1997). *Variability in microbiological degradation experiments - Analysis and case study.* *xiv* + 211 pp.

32. **Ma, Xin.** (1997). *Adaptive extremum control and wind turbine control.* *xix* + 293 pp.

33. **Rasmussen, Kim Ørskov.** (1997). *Nonlinear and stochastic dynamics of coherent structures.* x + 215 pp.

34. **Hansen, Lars Henrik.** (1997). *Stochastic modelling of central heating systems.* *xxii* + 301 pp.

35. **Jørgensen, Claus.** (1997). *Driftsoptimering på kraftvarmesystemer.* 290 pp.

36. **Stauning, Ole.** (1997). *Automatic validation of numerical solutions.* *viii* + 116 pp.

37. **Pedersen, Morten With.** (1997). *Optimization of recurrent neural networks for time series modeling.* x + 322 pp.

38. **Thorsen, Rune.** (1997). *Restoration of hand function in tetraplegics using myoelectrically controlled functional electrical stimulation of the controlling muscle.* x + 154 pp. + Appendix.

39. **Rosholm, Anders.** (1997). *Statistical methods for segmentation and classification of images.* *xvi* + 183 pp.

40. **Petersen, Kim Tilgaard.** (1997). *Estimation of speech quality in telecommunication systems.* x + 259 pp.

41. **Jensen, Carsten Nordstrøm.** (1997). *Nonlinear systems with discrete and continuous elements.* 205 pp.

42. **Hansen, Peter S.K.** (1997). *Signal subspace methods for speech enhancement.* x + 214 pp.

43. **Nielsen, Ole Møller.** (1998). *Wavelets in scientific computing. xiv* + 232 pp.

44. **Kjems, Ulrik.** (1998). *Bayesian signal processing and interpretation of brain scans. iv* + 125 pp.

45. **Hansen, Michael Pilegaard.** (1998). *Metaheuristics for multiple objective combinatorial optimization. x* + 163 pp.

46. **Riis, Søren Kamaric.** (1998). *Hidden markov models and neural networks for speech recognition. x* + 223 pp.

47. **Mørch, Niels Jacob Sand.** (1998). *A multivariate approach to functional neuro medeling. xvi* + 147 pp.

48. **Frydendal, Ib.** (1998.) *Quality inspection of sugar beets using vision. iv* + 97 pp. + app.

49. **Lundin, Lars Kristian.** (1998). *Parallel computation of rotating flows. viii* + 106 pp.

50. **Borges, Pedro.** (1998). *Problems with multiple criteria heuristic approaches.*

51. **Nielsen, Jakob Birkedal.** (1998). *New developments in the theory of wheel/rail contact mechanics. xviii* + 223 pp.