ELASTOHYDRODYNAMIC LUBRICATION AND SURFACE FATIGUE MODELLING OF SPUR GEARS OVER THE MESHING CYCLE

by

Sergey Khaustov

Thesis submitted in candidature for the degree of

Doctor of Philosophy at Cardiff University

Tribology and Contact Mechanics Research Group

Institute of Mechanics and Advanced Material

School of Engineering

Cardiff University

July 2016

DECLARATION

This work has not been submitted in substance for any other degree or award at this or any other university or place of learning, nor is being submitted concurrently in candidature for any degree or other award.

Signed …… Sergey Khaustov Date ……

STATEMENT 1

This thesis is being submitted in partial fulfillment of the requirements for the degree of PhD

Signed …… Sergey Khaustov Date ……

STATEMENT 2

This thesis is the result of my own independent work/investigation, except where otherwise stated. Other sources are acknowledged by explicit references. The views expressed are my own.

Signed …… Sergey Khaustov Date ……

STATEMENT 3

I hereby give consent for my thesis, if accepted, to be available online in the University's Open Access repository and for inter-library loan, and for the title and summary to be made available to outside organisations.

Signed …… Sergey Khaustov Date ……

Abstract

This thesis presents a modern method to evaluate spur gears based on the transient elastohydrodynamic lubrication (EHL) emulation of the full meshing cycle, evaluating elastic stresses in the gear flanks, collecting the stress history and applying stress and strain-life methods to calculate fatigue parameters and cumulative fatigue damage, i.e. predicting the fatigue life taking measured surface roughness into account.

The EHL model is formulated as the coupled system of the hydrodynamic Reynolds equation and the elastic deflection equation. These are solved simultaneously including the transient effect by incorporating the squeeze film term of the Reynolds equation with a Crank-Nicolson discretization of time. The finite difference discretisation of the elastic deflection equation utilises the differential form first formulated at Cardiff to allow coupling of the equations. The Reynolds equation can be discretised either by a finite difference or by a finite element method. The coupled system is solved simultaneously either by a narrow bandwidth Gaussian elimination or a Gauss-Seidel iterative method.

The elastic stresses due to the superimposed discrete values of the EHL pressure and shear stress at the EHL mesh nodes are evaluated by carrying out the necessary convolution of the stresses by a Fast Fourier Transform method. The weighting functions required have been calculated analytically. The stresses are obtained on the EHL solution mesh and are interpolated to meshes fixed in the pinion and the gear flanks. They are then sorted and stored efficiently to enable fatigue life prediction algorithms to be applied.

A detailed description of the EHL and the stress evaluation models are provided as well as a brief description of some fatigue life theories and calculations. The results of the complete analysis are provided for test gears obtained from the NASA Glenn laboratory fatigue tests and the Newcastle University Design Unit micro-pitting investigation. The analyses were carried out for real operating conditions from gear testing under extreme conditions. The surface roughness profiles used were real measured profiles taken from the test gears after initial running-in. The simulations reported are therefore as realistic as can be achieved and represent the true mixed lubrication conditions occurring in heavily loaded gears. The research also shows the importance of precise alignment of the roughness profiles in these conditions.

Acknowledgments

This project was made possible by funding from the U.S. Army Research Laboratory, Glenn Research Center, Cleveland, Ohio to whom I am very grateful, especially Dr. T. L. Krantz and Dr. S. P. Berkebile. I acknowledge the contribution of The Design Unit, Newcastle University for providing the gears to measure for the full gear meshing cycle emulation.

I am indebted to my research advisors, Prof. H. P. Evans and Dr. A. J. Clarke for their time, effort and invaluable support through the duration of this project and for providing me the opportunity to pursue a PhD.

Thanks must be directed to Dr. K. Sharif for helping me to get to grips with the software provided and some helpful advises and Dr. I. Weeks for sharing some experimental data. I also thank my colleagues in the Cardiff Tribology and Contact Mechanics Group, in particular: Mr. M.F.M. Al-Mayali, Mr. N.S.T. Almuramady, Mr. O. Savencu, Dr. R. Gay, Mr. G.J. Dennis, Dr. M. Bryant, Dr. A. Al-Hamood for creating a friendly environment in the office and Prof. F.M. Borodich for productive discussions and many words of wisdom.

Finally, I would like to thank my family and friends for their help and support.

Contents

Declaration I
Abstract II
Acknowledgments III
Contents IV
Nomenclature VII

Chapter 1 – Introduction and literature review

1.1 Introduction 1-1
1.2 History of Gearing 1-1
1.3 History of Tribology 1-4
1.4 History of Elastohydrodynamic Lubrication (EHL) 1-8
1.5 Development of a full numerical solution in line contact 1-14
1.6 Roughness effect 1-16
1.7 Spur Gear geometry and operational conditions 1-16
1.8 Aims of the current work 1-17
1.9 Thesis structure 1-18

Chapter 2 – Introduction and basic theory

2.1 Introduction 2-1
2.2 Geometry and Kinematics of spur gear contact 2-2
2.3. The Elastic deformation equation 2-12
2.4. Load conditions over meshing cycle of spur gear contact 2-13
2.5. The viscosity equation 2-15
2.6. The density equation 2-17
2.7. The Hydrodynamic Reynolds equation 2-18
2.8. The Eyring Rheological model 2-24

Chapter 3 – Numerical theory for the Finite EHL model

3.1 Introduction 3-1
3.2 Computational mesh 3-1
3.3. Numerical formulation of the elastic film thickness equation 3-4
3.4. Numerical formulation of Reynolds hydrodynamic equation 3-14
3.5. Finite central difference formulation 3-15
3.6. Finite element formulation 3-17
3.7. Matrix formulation and the solving process 3-23

Chapter 4 – Elastic stress evaluation and fatigue life expectancy analysis

4.1 Introduction 4-1
4.2. Elastic stress evaluation 4-1
4.3. Verifying the accuracy of the elastic stress calculations 4-13
4.4. Storing stress history 4-28
4.5. Fatigue of Materials: definition and brief history 4-32
4.6. Fatigue of Materials: basic terms and nomenclature 4-33
4.7. Multiaxial elastic stress-life methods 4-40
4.8. Multiaxial strain-life approach 4-42
4.9. Summary 4-44

Chapter 5 – The results of the complete analysis of the gear meshing cycle. The NASA Glenn case

5.1 Introduction 5-1
5.2. NASA test set up parameters 5-2
5.3. EHL and stress analysis solution approach 5-5

Chapter 6 – The results of the complete analysis of the gear meshing cycle. The Design Unit case

6.1 Introduction _____ 6-1
6.2. The test set up parameters _____ 6-2
6.3. Results of the EHL smooth contact analysis _____ 6-4
6.4. Surface roughness measurements and treatment _____ 6-5
6.5. Results of the rough surface contact analysis _____ 6-13

Chapter 7 – Conclusions and Recommendations for future work

7.1 Conclusions _____ 7-1
7.2. Recommendations for future work _____ 7-2

Appendix A – MatLab code

Nomenclature

a	Hertz contact half-width	m
c_1, c_2, c_3, c_4	Weighting functions for directional elastic stress evaluation	
E'	Equivalent elastic modulus	Pa
E_1, E_2	Young's modulus of the driving and driven gears respectively	Pa
f	Weighting function for elastic deflection calculation	
h	Clearance between surfaces / Film thickness	m
h_u	Total separation between two involutes/parabolas with no load applied	m
n_1, n_2	Numbers of teeth of the driving and driven gears respectively	
N_1, N_2, N_3	Lagrange interpolation functions	
p	Pressure	Pa
p_0	Maximum Hertz pressure	Pa
Q	Mass flowrate	$kg \cdot s^{-1}$
q	Tangential shear stress at flank surface	Pa
Q_{CF}	Couette flow	$kg \cdot s^{-1} \cdot m^{-1}$
Q_{PF}	Poiseuille flow	$kg \cdot s^{-1} \cdot m^{-1}$
R	Radius of relative curvature	m
R_1, R_2	Radii of curvature of surfaces	m
r_{b1}, r_{b2}	Base radii of the pinion and the gear respectively	m
r_{t1}, r_{t2}	Tip radii of the pinion and the gear respectively	m

s	Instantaneous position of the contact on the line of action with origin at the pitch point	m
S	Non-Newtonian flow factor	
s_{c1}, s_{c2}	Boundary of the single tooth contact zone	m
s_f, s_l	Coordinates s of the first and last contact of meshing cycle	m
u_1, u_2	Velocity of surfaces relative to contact	$m \cdot s^{-1}$
v_1, v_2	Poisson's ratio	
w'	Load per unit length of contact line	$N \cdot m^{-1}$
x	Axis orientated parallel to the entrainment direction	
z	Axis orientated perpendicular to the entrainment direction, common normal direction	
z_1, z_2	Distance directed into the tooth material of the driving and driven gears respectively	m
α	Pressure-viscosity coefficient	$m^2 \cdot N^{-1}$
γ	Involute roll-angle	rad
$\dot{\gamma}$	Shear strain rate	s^{-1}
γ_ρ	Compressibility constant of lubricant density formula	Pa^{-1}
Δ	Mesh spacing	m
ε	Lubricant thermal expansion coefficient	K^{-1}
η	Lubricant viscosity	$Pa \cdot s$
η_0	Viscosity at reference temperature and pressure	$Pa \cdot s$
η_{eff}	Effective viscosity	$Pa \cdot s$
θ	Temperature	K
θ_0	Reference or ambient temperature	K
Λ	Elastic deflection	m
λ	Mean shear stress for lubricant rheological model	
λ_ρ	Compressibility constant of lubricant density formula	Pa^{-1}

μ	Coefficient of friction	
ρ	Density	$kg \cdot m^{-3}$
ρ_0	Density at reference pressure	$kg \cdot m^{-3}$
σ	Flow factor	$m \cdot s$
Σ	Pressure gradient for lubricant rheological model	
σ_r, σ_θ	Direct stress components in polar coordinate system	Pa
σ_x, σ_z	Direct stress components in Cartesian coordinate system	Pa
τ	Shear stress	Pa
τ_0	Eyring shear stress	Pa
τ_m	Shear stress on lubricant midplane	Pa
$\tau_{r\theta}$	Directional shear stress in polar coordinate system	Pa
τ_{xz}	Directional shear stress in Cartesian coordinate system	Pa
ϕ	Roughness	m
ψ	Pressure angle	rad
ω_1, ω_2	Angular velocities of the driving and driven gears respectively	$rad \cdot s^{-1}$

Chapter 1

Introduction and literature review

1.1. Introduction

This Chapter provides a brief history of gear design and Tribology, placing emphasis on transient elastohydrodynamic lubrication (EHL) numerical models of spur gear contact. Due to the thesis structure, the overview of stress evaluation theory and its development are presented in Section 4.2 as an essential part of the mathematical formulation of the technique. The evolution of fatigue methods is exploited as a justification of the choice of the fatigue life prediction procedures in Chapter 4.

The thesis aims and structure are provided in the last section of this chapter for the reader's convenience and to help them navigate through the document.

1.2. History of Gearing

The use of gear technology can be traced back to fourth century b.c. in Greece and China, although there is some indirect evidence that they existed even earlier than that. There are relatively few publications on the history of gears, partly because the gears were overlooked for centuries due to unavailability of powerful engines, and therefore, the early development of the subject is not properly documented. The fact that the history of modern gears only extends over only two centuries and is tied together with the development of new manufacturing methods and the Industrial revolution is the other reason.

The ancient history of the gears is related to clocks, watches and mechanical calendar computational devices. Price (1959 and 1974) provides some examples of ancient gear based mechanisms. Field and Write (1985), inspired by the arrival of Early Byzantine gearing to the Science Museum, London, summarised the early years of history of gearing. The most comprehensive document describing the evolution of the cogwheel contact was written by Dudley (1969) and a comprehensive summary is provided in NASA-RP-1152 and NASA contribution to gearing in AVRADCOM 82-C-16 by Coy.

Whilst, the fundamental principles of gearing have not changed and ancient cog-wheels served the same purposes, they have very little in common with modern gears. Initially the shape of teeth was either not important or just flat, for example in power transmissions like the Vitruvius' watermills dated back to 40 b.c. in Figure 1.1, or a precision of contact was achieved by manual adjustment of each pair of gears in contact as for the Antikythera mechanism.

Figure 1.1 – Transmission of the Nether Alderley Mill dates from the 12th century as an example of the Vitruvius' watermills design

The history of modern gears started in 1525 when Albrecht Durer discovered the epicycloidal shape, but it was not until 1694 that Phillip de la Hire suggested the involute shape for gear flanks. However, the revolution in gear design dates from 1754, the year in which Leonard Euler formulated the conjugate action law, which allows a steady speed ratio to be maintained. Since then the fundamental design of gear geometry has not been changed significantly. During the Industrial Revolution, gear design and manufacturing rapidly evolved and was properly documented. The most important development over that period is the formulation of the gear tooth stress evaluation method by Lewis (1893). This method is still widely used with some limitations. For example, BS ISO 6336-3:2006 'Calculation of load capacity of spur and helical gears. Calculation of tooth bending strength' states a modified Lewis formula, which only differs by the number of correction factors incorporated into the formula. Finally, Grant (1899) published a comprehensive manual for gear design and manufacturing.

The Technological Revolution set new standards and requirements for gear transmissions. Precision, durability and noise became the crucial limiting factors. Some problems were resolved by the invention of new types of gearing, novel materials and a significant improvement of the manufacturing processes, which is not the focus of interest of this work. For conventional types of gears, such as spur, helical and bevel gears, the traditional design approach cannot meet the demand for new areas of application, for instance in the aerospace industry, characterised by high transmitted speeds and loads, and the critical importance of weight. The robust solution to the problem can only be achieved by the use of the finite-element analysis, which takes into account all major parameters, such as roughness, lubricant, load variation etc.

1.3. History of Tribology

Etymologically the word 'tribology' originated from the Greek root '*τριβ*-', tribo-, translating as 'rubbing' and the suffix -*λογία*', -logia, meaning 'study of'. Jost (1966) introduced the term '*tribology*' in the report on the financial losses to the UK economy due to friction and wear. The current definition of the word 'tribology', according to the Oxford dictionary, is 'The study of friction, wear, lubrication, and the design of bearings; the science of interacting surfaces in relative motion'.

Tribology is multidisciplinary in nature, and includes mechanical engineering, especially machine elements such as gears, journal and roller bearings, materials science with research into wear resistance, surface technology with surface topography analysis and coatings, and the chemistry of lubricants and additives.

As stated above, tribology as a science was established in 1966, but the problems it covers were known from the dawn of the human civilization. The chronicler of tribology, Dowson (1998), published a comprehensive history of Tribology from the beginning of humanity to that time. The evidence of the first very basic tribological studies goes back to the Paleolithic period. According to Furon (1963) and Forbes (1967), humanlike creatures such as Pithecanthropus, Sinanthropus, Atlanthropus and the later Heidelburg, Swanscombe, Fontechevade and Neanderthtal man were familiar with fire, stone tools and weapons, which is direct evidence of tribological skills from that period. Percussion of flint stones and friction of wood were recognised as a cause of frictional heating. There are some serious gaps in the historical data over Mesolithic and Neolithic periods, but Singer et al. (1954) mentioned that stone and wood bearings for door-posts and hand-held bearings of antler, bone and stone for drills were invented and developed over that period of time. In contrast to these simple tools, the Early Civilisations invented wheeled vehicles and potter's wheels that used bearings and were lubricated either by water or by bitumen. The first recorded tribologist, painted in an Egyptian fresco of transporting the statue of Ti dated back to 2400 b.c., poured lubricant in front of a sledge. Due to a poor quality of that painting, another fresco dated back to 1880 b.c. and shown in Figure 1.2 is used as an example of lubricant application.

The Greek and Roman period brought us new technologies and materials as well as scientific research into tribological matters. Herodotus (484-425 b.c.) described ancient methods of producing bitumen and a lighter oil from petroleum. Aristotle (384-322 b.c.), in his *Questiones Mechanicae*, recognised the force of friction and its relation to the shape of the objects in contact. The use of rolling elements to reduce friction becomes a common practice as well as an utilisation of bronze.

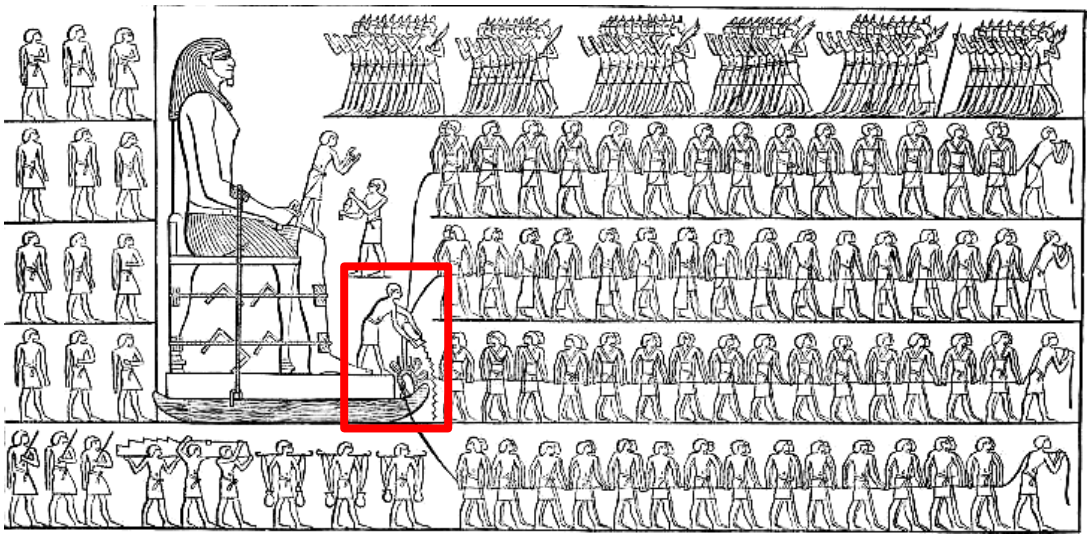


Figure 1.2 – Transporting an Egyptian colossus,
Tomb of Tehuti-Hepter, El-Bersheh, 1880 b.c.

The Medieval Ages are characterised by an evolutionary development of lubricants and structural materials. Vegetable oils and animal fats were used as lubricants. Iron and brass become popular bearing materials. The source of the power over that time was water and wind, which were employed, for example, in mills. The quintessential example of the design level is the invention of the mechanical escapement clock that replaced an hourglass and a water clock.

Over the Renaissance period, while Columbus was on his way to the undiscovered land of America, da Vinci was conducting his studies of friction. The sketches of his experiments from the *Codex Atlanticus* and the *Codex Arundel* are shown in Figure 1.3. He formulated the first two laws of friction, namely:

1. The force of friction is directly proportional to the applied force;
2. The force of friction is independent of the apparent area of contact.

This is the first recorded scientific study of friction.

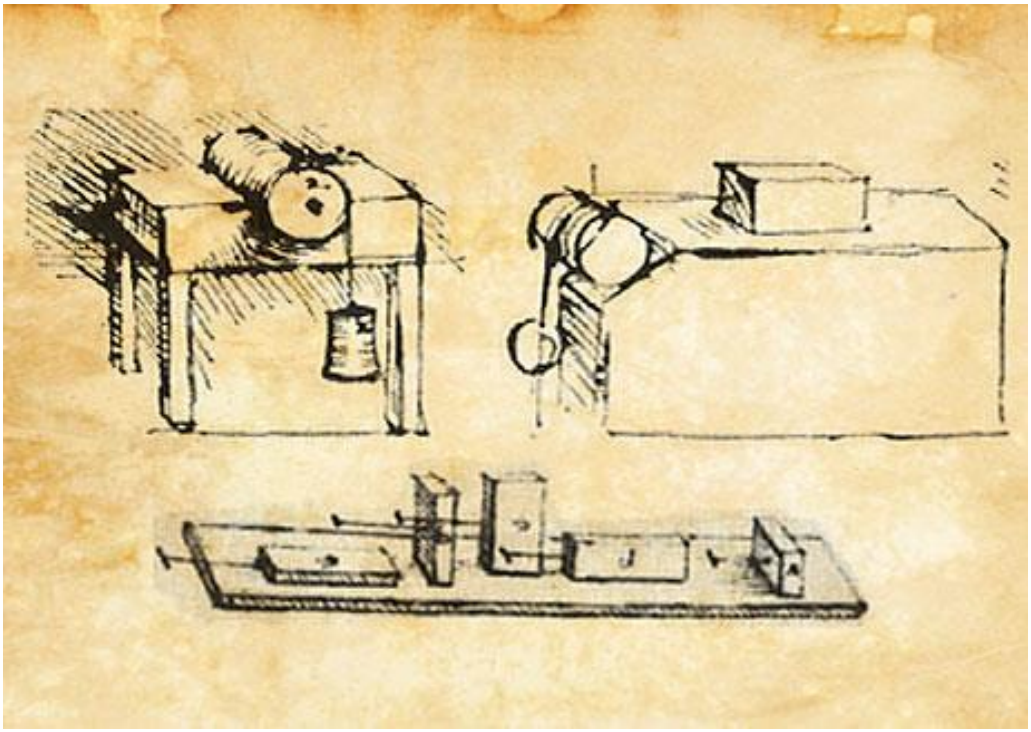


Figure 1.3 – Leonardo da Vinci's studies of friction, the *Codex Atlanticus*

In the *Codex Madrid I* da Vinci continued his study of friction and proposed some rolling-element bearing designs including an early version of a 'cage' to prevent contact between the balls. He considered some lubrication systems and tribological aspects of gears and screw-jacks. However, an intensive description of mining technology in *De re Metallica* by Agricola (1494-1555) demonstrates that there were no significant developments in tribological applications since the Greek and Roman ages.

Due to the increasing demand for mechanical power during the seventeenth century and, consequently, the urge to decrease wear of moving machine elements, led to a considerable development in the science of tribology by Hooke, Newton and Amontons. Hooke (1635-1703) studied hardness of materials and rolling friction, which resulted in a series of concepts on bearing design, seals and lubrication. In 1737 Desaguliers published the first table of friction stating values that can be converted into coefficients of friction with a respect to the material of the surfaces in contact and lubrication. In Berlin, the famous mathematician Euler (1707-1783) defined the force required to move a weight up a slope of inclination to the horizontal and introduced the term 'coefficient of friction'. He also showed the distinction between kinetic and static friction, followed by a recognition of differences between sliding and rolling friction stated by Leibnitz (1646-1716). In 1706 Newton (1642-1757) conducted a study of viscous flow. Section IX of Book II of his *Principia Mathematica Philosophiae Naturalis* (1687), which covers the 'Circular Motion of Fluids', opens with a statement now described as Newton's law of viscous flow. The '*defectus lubricitatis*' in his words is nowadays known as *viscosity*.

The rapid change of technology during the First Industrial revolution induced scientific research in general, including tribological studies. Charles Augustin de Coulomb defined friction formulas based on his own experimental studies. The basics of the fluid mechanics were formulated by Euler, Bernoulli, Poiseuille etc. Newton's law of flowing fluids was elaborated by Claude-Louis Navier by including the viscous terms. At the same time Stokes defined the basics of viscous flow. Subsequently, the equations formulated in this way became known as the Navier-Strokes equations. The first distillation of mineral-oil based lubricant was conducted in Prague in 1812 and the first graphite-containing lubricant was patented in the United Kindom in 1835.

According to Mang, Bobzin and Bartels (2011), during the Second Industrial revolution the cornerstone of the tribological study was laid by numerous scientists, including: studies of rolling friction by Reynolds (1875); Hertz's analysis of contact between elastic materials (1881); Petrov's studies on unloaded journal bearings (1883); the development of viscometers by Engler, Saybolt and Redwood (1884-1886); the hydrodynamic theory of Tower and Reynolds (1865); friction measurements on journal bearings by Stribeck (1902); the analytical solution of Reynolds equation by Sommerfeld (1904); further investigations on Stribeck's results

and defining the major source of dry friction by Gumbel (1914-1925); Langmuir's studies of thin surface films (1917); continuous improvements of the journal bearing by Tower, Kingsbury, Michell and Rayleigh (1915-1925).

In the beginning of the twentieth century, tribological studies branched out according to the subject of interest. The research aims of this thesis are mainly associated with the field of Elastohydrodynamic Lubrication (EHL), therefore, only the timeline of major development in this particular area is covered in the next section.

1.4. History of Elastohydrodynamic Lubrication (EHL)

According to the Encyclopedia of tribology (2013), Elastohydrodynamic Lubrication (EHL) is a mode of fluid-film lubrication in which hydrodynamic action is significantly enhanced by surface elastic deformation and lubricant viscosity increases due to high pressure. A comprehensive discussion of the EHL history was published by Zhu and Wang (2011) and reprinted with minor alterations in the Encyclopedia of tribology (2013), which complemented reviews by Dowson and Ehret (1999), Gohar (2001) and Spikes (2006).

Although, EHL was established as a discipline and separated from the other tribological studies in 1930s, the first idea that lubricants do not just fill the gaps between surfaces in contact caused by roughness, but separates them was suggested by Rennie (1829). The presence of fluid films in journal bearings was discovered at the same time by Petrov (1883) of the Moscow Politechnical Society and Tower (1883) of the Institution of Mechanical Engineers. Both experiments shown a substantial pressure in the oil films. Petrov concentrated on the development of the friction function formula, but Tower investigated the fluid film in a variety of different lubricated contacts, which led a second report, published in 1885, presenting the variation of pressure in a journal bearing. Tower's research provided the experimental basis and conceptual stimulus for the milestone theoretical lubrication analysis of a journal bearing conducted by Reynolds (1886). The Reynolds equation that governs fluid film lubrication was derived and the approximate solutions shown to be in good agreement with Tower's results. The Reynolds equation has since become a cornerstone of hydrodynamic lubrication theory. However, the results were only

obtained for conformal contact. In a conformal contact, the surfaces in contact have a very similar shape; therefore, the area of contact is comparable to the dimension of the mechanical elements. A non-conformal contact involves surfaces whose form is very different, thus, the contact area is small either in both principal dimensions, and is either an elliptic contact, or in the one dimension case, a line contact. Examples of conformal and non-conformal contact from Hamrock and Anderson (1983) are shown in Figure 1.4.

The first analytical solution for frictionless dry non-conformal contact of smooth surfaces was developed by Hertz (1881), which was combined with the results of the experiments and explained the Newton's rings phenomena. Martin (1916) solved the Reynolds equation for line contact in spur gears applying following assumptions: a pair of gear teeth can be approximated as two parallel cylinders in contact; the elastic deflection of the gear flanks can be neglected, i.e. rigid; the lubricant can be considered incompressible, i.e. isoviscous and, therefore, Newtonian. However, the predicted film thickness was extremely small, in some cases even smaller than the average roughness that can be achieved by manufacturing processes, although, gears operated successfully and showed only moderate traces of surface damage or wear.

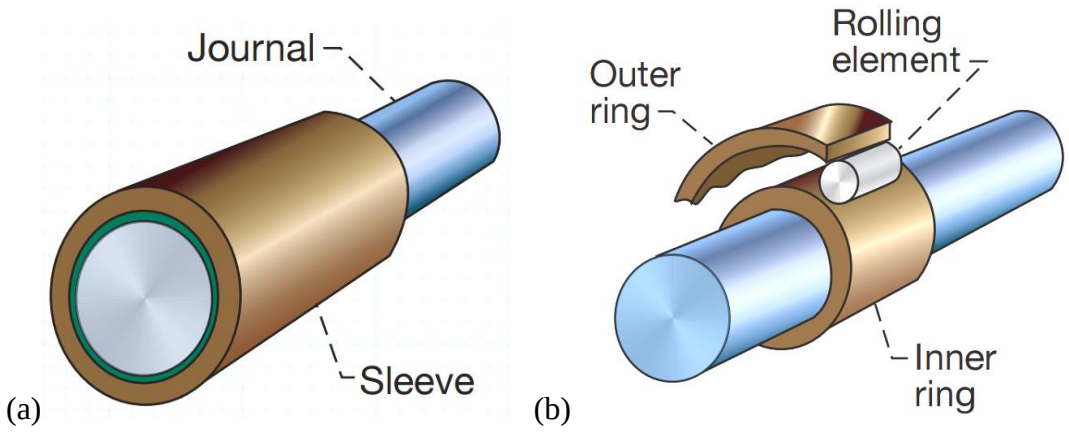


Figure 1.4 – Examples of (a) conformal and (b) non-conformal contact, Hamrock and Anderson (1983)

Starting from 1930s, there were some attempts to develop a robust EHL analysis that takes into account either the localised elastic deformation (Peppler, 1938, Meldahl, 1941, etc.) or the real behaviour of a lubricant (Gatcombe, 1945, Blok, 1950, etc.). Grubin and Vinogradova (1949), based on Ertel's (1939) experimental results, derived a theory, which incorporated the effect of the elastic deformation and the viscous response to the pressure simultaneously. The dramatic life of Alexander Mikhailovich Mohrenstein-Ertel and his contribution to Tribology was presented by Popova and Popov (2014). Two assumptions were made to overcome an unavailability of sufficient computing resources:

1. The shape of the elastic deformation is not affected by the presence of the lubricant and it is identical to the dry contact conditions, i.e. Hertz (1881) contact theory.
2. The hydrodynamic pressure tends to infinity at the inlet boundary of the Hertzian contact.

The pressure-viscosity relation was defined by the Barus (1893) equation, where the viscosity depends on the pressure according to

$$\eta = \eta_0 \, e^{\alpha p} \tag{1.1}$$

Based on those assumptions, Grubin numerically integrated a simplified Reynolds equation at the inlet zone to the Hertzian contact area. This was done in terms of the reduced pressure, q, where for the Barus viscosity formula (1.1)

$$q = \frac{1}{\alpha}\left(1 - e^{-\alpha p}\right) \text{ and } \; p = -\frac{1}{\alpha}\ln\left(1 - \alpha q\right) \tag{1.2}$$

And the Reynolds equation for the reduced pressure in the isoviscous form

$$\frac{d}{dx}\left(h^3\frac{dq}{dx}\right)=12\bar{u}\,\eta_0\frac{dh}{dx}\text{, i.e. }\frac{dq}{dx}=12\bar{u}\,\eta_0\frac{h-h^*}{h^3} \qquad (1.3)$$

When q reaches a value of $1/\alpha$ the corresponding pressure becomes infinite.

Grubin numerically integrated equation (1.3) for $q = 0$ at $x = -\infty$ and obtained the value of q at $x/a = 1$ for a series of values of constant of integration h^*. A curve fitting to these data allowed the contradiction $q = 1/\alpha$ at $x/a = 1$ to be stated and this was developed to obtain the film thickness formula. The predicted values of clearance between the surfaces in contact were over an order of magnitude greater than the ones calculated according to the Martin's theory and were plausible for limited practical applications.

Petrusevich (1951) produced the sets of results for different speed, the same load by combining the full Reynolds piezoviscous equation and the Grubin inlet formula. The resultant pressure distribution and film thickness are presented in Figure 1.5. (Note that entrainment is in the negative axis direction here)

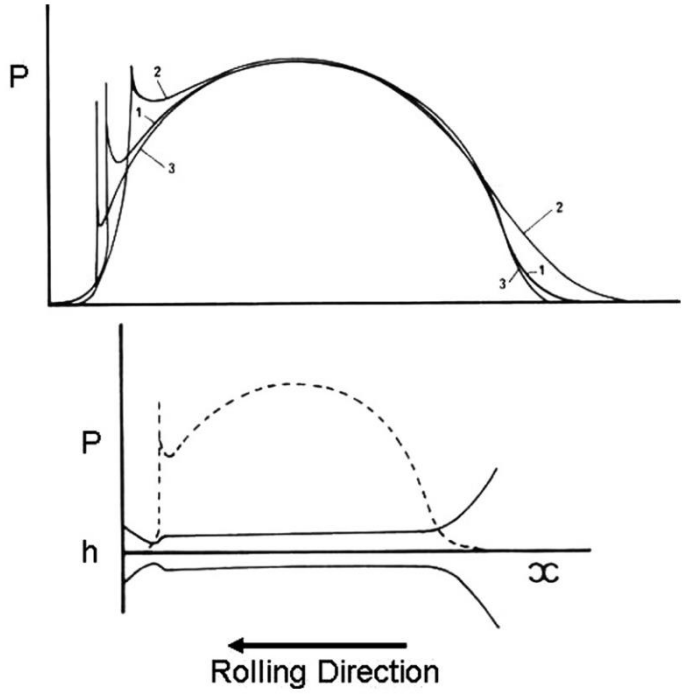


Figure 1.5 – Graphs of resultant pressure distributions (upper) for three line-contact Petrusevich cases and film thickness (lower), Petrusevich (1951)

It is apparent that the magnitudes of the pressure distribution are similar to the Hertzian contact results for the larger part of the contact length of all three cases with a significant discrepancy at the inlet and outlet zones. The later studies confirmed the characteristic build-up of pressure with a sharp spike, named the 'Petrusevich Spike', corresponding to the beginning of film constriction at the outlet area. Considering the insufficient statistical sampling size, he formulated an analytical relation between the film thickness and the entrainment velocity, which describes the nature of the process quite accurately.

The Petrusevich research was shortly followed by a new formulation and algorithm proposed by Dowson and Higginson (1959), which allowed converged solutions to be obtained for a larger variety of operating conditions in a small amount of iterative steps. This was an innovative development that solved the Reynolds equation in an inverse way where the film thickness was obtained for a given pressure distribution. The pressure distribution thus led to two different film thickness curves, one a solution to the elastic equation and one a solution to the Reynolds equation. The method proceeded to assume an initial approximate pressure distribution (a modified Hertzian pressure) and then adjusted it so as to obtain agreement between the two film thickness curves. The first solutions were obtained with some manual intervention by the authors in the iterative process. This was then developed to be an automatic algorithm by Dowson, Higginson and Whitaker (1962).

A curve fit to a set of results was used to produce a formula for predicting line contact EHL minimum film thickness by Dowson and Higginson (1961). Another formula was also presented by Dowson and Higginson (1966). The predicted values of film thickness were highly dependent on the rolling speed and almost unresponsive to the load conditions. This showed a good agreement with numerous experimental studies of line contact such as Crook (1961, 1963) and Dyson et al. (1966) using the capacitance technique, and Sibley and Orcutt (1961) employing the X-ray transmission method. The book 'Elastohydrodynamic lubrication' by Dowson and Higginson (1966) has been considered the foundation of the smooth line-contact EHL theory. Later on, by application of thin-film transducers onto a test disc the EHL pressure distribution was measured and the existence of the Petrusevich spike was confirmed by Kannel (1966) and Hamilton and Moore (1971). Further contributions to the

prediction of minimum film thickness were made by Moes (1966), Theyse (1966), Archard (1968), Greenwood (1969), Johnson (1970), Moes and Bosma (1972), etc.

At the same time, some experimental measurements of the heavy loaded point contact formed by two crossed cylinders were conducted by Archard and Kirk (1961) using an electrical capacitance technique. Shortly after that the optical interferometry method was developed, which allowed observing of the film thickness distribution rather than an average or central film thickness, or an estimate of minimum film thickness provided by the electrical capacitance and x-ray methods respectively. Zhu and Wang (2011) refer to Gohar and Cameron (1963) as the original developers of the optical interferometry method, however, Kirk (1962) and Archard and Kirk (1963) used white light to study the influence of the materials on the film produces by brining Perspex and glass cylinders into a contact. The first effective use of optical interferometry was presented by Cameron and Gohar (1966), which led to a rapid growth in popularity of this technique. It was further developed and modified by Foord et al. (1969-1970), Gohar and Cameron (1967), Wedeven (1970). More recently Johnson, Wayet and Spikes (1991) made a significant advance by combining an optical spacer layer, introduced by Westlake and Cameron (1967), and a spectrometer. In this way modern interferometers can measure film thickness down to one nm overcoming an obvious physical limit due to the wavelength.

The first simplified analytical solution was obtained only about two decades after Grubin's line contact elucidation by Archard and Cowking (1966) and Cheng (1970). The implementation of the numerical analysis for point contact by Ranger et al. (1973) followed more than two decades later and this delay was principally due to the computational resources available to researchers.

The later period of the EHL history can be characterised as a rapid development and branching of the EHL studies, therefore, only the subjects, relative to this research will be reviewed in the next sections.

1.5. Development of a full numerical solution in line contact

The EHL problem is formulated as a system of two governing equations: the elastic deflection, which maintains the load balance, and Reynolds equation that describes the pressure distribution and film thickness, equations. The latter one incorporates density and viscosity that are heavily dependent on pressure. This system exhibits strong nonlinear behaviour predominantly due to the elastic deformation and the rise of viscosity at high pressure. Therefore, it causes severe numerical difficulties and makes a traditional direct integration method, for example, Ranger et al. (1973), Hamrock and Dowson (1976a, 1976b, 1977a, 1977b), Evans and Snidle (1981) ineffective, especially under heavy loading conditions. Over the years four major numerical techniques to find the EHL solution were developed.

The *Inverse Solution* was first put together by Dowson and Higginson (1959) for the line contact and Evans and Snidle (1981) and Hou et al. (1987) for the point contact problem. Initially, the pressure distribution is assumed; therefore, the two film thickness distributions are obtained by integration of the Reynolds equation and by solving the elastic deflection equation. Then those two functions are compared and the difference is used to modify the initial pressure shape. The calculations loop until the difference between the hydrodynamic and elastic deflection functions complies with the convergence criteria. However, this procedure involves some manual adjustment manipulations based on professional expertise and knowledge and it is quite unstable for light loads.

The *Newton-Raphson Iterative Procedure* was first applied by Rohde and Oh (1975) for line contact problems and Oh and Rhode (1977) for point contacts. Some formulation improvements were performed by Okamura (1982). The method involves a simultaneous solution of the elastic deflection and Reynolds equations, which are coupled in the solution scheme. It only takes a few iterations to converge, but it requires a good initial guess of the pressure distribution and it is compulsory to obtain a solution from the full matrix, as all the mesh nodes are included in the elastic equation. Houpert and Hamrock (1986) introduced a non-uniform computational mesh to decrease computational time and memory demand, which was improved by Hsiao et al. (1998) as an automatic meshing-remeshing algorithm was employed for the point contact problem.

The *Coupled Differential Deflection Method* originated from Okamura's (1982) first order full coupled formulation, presented by Houpert and Hamrock (1986) and further developed by Elcoate (1996) and Hughes et al. (2000) for the line contact problem and by Holmes et al. (2003) for point contacts. A different formulation of the elastic deflection equation was suggested by Evans and Hughes (2000). The method allows to use a limited bandwidth matrix in solving the coupled EHL governing equations simultaneously and to increase stability under heavy loading conditions. The method was extended by Elcoate et al. (1999, 2001) to include real roughness and to make appropriate adjustments in the mixed lubrication regime.

The *Improved Direct Iterative Techniques* include the *Multigrid* (MG) *method, Semi-System Approach* and the *Progressive Mesh Densification* (PMD) *Procedure*. MG and PMD methods use a set of grids with different spacing to remove an error between the initial value and final solution as an alternative relaxation process or to generate a better initial candidate solution for the finer mesh respectively. Unfortunately, they are not applicable for rough surface contacts. The main contribution to MG method was made by Lubrecht et al. (1986), Lubrecht (1987), Venner (1991), Ai (1993) and Venner and Lubrecht (2000). The PDM technique was developed by Hu and Zhu (2000) and Zhu (2007). The Semi-system method was developed by Ai (1993) and basically modifies the Gauss-Seidel (Dowson and Higginson, 1966) relaxation process by considering the Couette entraining flow term of the Reynods equation as a function of unknown nodal pressures, on the contrary to the Gauss-Seidel, which only takes into account the contribution of the Poiseuille flow. It allows ensuring stability even under extremely severe conditions. PDM method was employed by Zhu and Hu (1999) and Hu and Zhu (2000).

1.6. Roughness effect

In reality, there are no ideally smooth surfaces and most of the time roughness is comparable to the predicted film thickness or even greater, therefore, it must affect the behaviour of the EHL contact. Engineering observations proved that the contact surfaces of the heavily loaded mechanisms show no traces of damage even considering that their roughness is much higher than the minimum film thickness predicted for smooth contact conditions. Early attempts to obtain a numerical solution revealed some major complications such as high sensitivity of the micro deformations to thermal and non-Newtonian effect as well as a necessity of very high resolution.

There are two groups of rough surface analyses: stochastic and deterministic. At the beginning, Lee and Cheng (1973) and Cheng and Dyson (1978) implemented a stochastic model, but, Patir and Cheng (1978a), Majumbar and Hamrock (1982), Prakash and Czichos (1983), Zhu and Cheng (1988), Sadeghi and Sui (1989) used the average flow EHL model of Pair and Cheng (1978b). They show tendencies to increase film thickness due to higher roughness or lower speeds of the surfaces in contact.

Poon and Sayles (1994) and Ai and Cheng (1994) pioneered the inclusion of the real roughness into an EHL analysis and many followed them. It became apparent that in rough EHL problems the film thickness is very thin and metal on metal contact at some asperities is quite common. That type of regime is called mixed lubrication and some of the first results for real roughness in that type of condition were obtained by Tao et al. (2003) and Holmes et al. (2003a, 2003b).

1.7. Spur Gear geometry and operational conditions

The operational conditions and the undeformed geometry of involute spur gears vary through meshing cycle. In 1980 Wang and Cheng submitted the report, NASA CR 3241 ‘Thermal Elastohydrodynamic Lubrication of Spur Gears’ that included a method of evaluation of film thickness along the path of contact based on the Grubin approach. Larson (1997) presented a full transient analysis of the meshing cycle, which was enhanced by thermoanalysis and more realistic load variation by Wang et al. (2004). Davis (2005) developed a comprehensive TEHL numerical solution incorporating the non-simplified involute shape and real roughness of the gear flanks.

1.8. Aims of the current work

The aims of this thesis were to produce a robust tool to emulate the full meshing cycle of spur gears evaluating and storing elastic stress history in the way appropriate for applying a fatigue-life prediction tool. Another goal was to provide a comprehensive description of the model and input/output structure as well as increasing the reliability of the algorithm.

The goals stated above can be broken down into the following steps

1. Develop a full meshing cycle EHL contact analysis for spur gears in extreme mixed lubrication conditions utilising and enhancing existing software components where available. This requires:
    (a) Developing a method to include contact occurrences within a full meshing cycle analysis
    (b) Increasing robustness of the code in contact cases
    (c) Investigation of the importance of accurate surface roughness alignment
    (d) Considering methods to accelerate the software performance

2. Development of the analysis to allow stress evaluation for the whole meshing cycle and production of the gear stress history information

3. Application of fatigue tools using the stress history

4. Packaging the gear analysis software for non-expert use. This is anticipated to include a consideration of
    (a) Interface control
    (b) Appropriate tools for dealing with job failures
    (c) Development of recovery/debug logs for non-expert and remote use
    (d) Removal of non-essential code options and optimisation of the software

5. Providing tools for the graphic representation of the results at each stage of the analysis.

The analyses were carried out for real operating conditions from gear testing under extreme conditions. The surface roughness profiles used were real measured profiles taken from the test gears after initial running-in. The simulations reported are therefore as realistic as can be achieved and represent the true mixed lubrication conditions occurring in heavily loaded gears. The research also shows the importance of precise alignment of the roughness profiles in these conditions.

1.9. Thesis structure

The thesis includes the necessary summary of the theoretical methods used as well as a comprehensive description of their practical implementation and the results of the analysis. It is divided into the following chapters:

Chapter 2 provides information about geometry, kinematics and load conditions of spur gears through the meshing cycle, EHL governing equations and formulae describing non-Newtonian behaviour and mechanical properties of the lubricant.

Chapter 3 describes all implemented options of the full EHL transient numerical analysis of spur gear meshing cycle, including the calculations of the mesh spacing, the finite element and the finite element formulation of the Reynolds equation, Evans and Hughes (2000) differential method of deflection evaluation, structure of the matrix equation and explanation of the solving processes.

Chapter 4 explains the elastic stress evaluation techniques with verification of the calculations, the collection of the stress history through the meshing cycle, some multiaxial fatigue-life prediction theories as well as a brief history of the fatigue studies and the structure of the output files.

Chapters 5 and 6 present the results of the EHL and fatigue-life analyses of the NASA Glenn spur gear fatigue test and the Design Unit, Newcastle University experimental micropitting study. The preparation of the input data is described and a comparison of those two cases is provided.

The conclusions are drawn and some ideas for the future development are considered in *Chapters 7*.

Chapter 2

Introduction and basic theory

2.1. Introduction

This chapter describes the initial software for rough surface non-Newtonian transient line contact elastohydrodynamic lubrication (EHL) analysis for spur gears (Davis 2005) as well as the basic theory used and main limitations and key features of this algorithm.

The solution method used in the software was first developed by Elcoate et al. (1997) and it can be described as a simultaneous solution of Reynolds hydrodynamic equation and the elastic deflection equation formulated as a differential equation. The latter was formulated in this form by Evans and Hughes (2000). Both equations are functions of pressure and separation, where the elastic deflection equation sets the geometry of the contact and the hydrodynamic equation describes the behaviour of the lubricant that separates the surfaces. Further development was performed by Holmes (2000) and Sharif et al. (2001). The Eyring (1936) shear thinning rheological model was used to incorporate the non-Newtonian effect into the formulation of the problem.

The software for applying these techniques to the gear meshing cycle for spur gears was created by Davies (2005). The coupled method as further developed by Holmes, Evans, Snidle (2005) was used to carry out full meshing cycle non-Newtonian EHL analysis with real roughness of both tooth surfaces. The program provides three meshing options: three node quadratic or two node linear finite elements, or central finite difference, and two solvers that are based either on the Gaussian elimination method or on the Gauss-Seidel iterative method.

2.2. Geometry and Kinematics of spur gear contact

The EHL analysis of the lubricated contact of spur gears is a complex problem due to constantly changing radii of curvature and tangential velocities of the contacting surfaces, and the changing loading conditions through the meshing cycle. This is because the contact point progresses continuously over the involute gear tooth flanks and the number of teeth pairs engaged in contact varies through the meshing cycle. It is necessary to deduce a parameter that links all the operating conditions together. According to Kennedy's theorem and the basic concept of the involute spur gears, the contact point between two mating gear flanks advances uniformly along the straight line *BD*, shown in Figure 2.1, called the line of action or the pressure line. It is tangential to the base circles of both gears and inclined at the pressure angle, ψ, to the normal to the line of centres, A_1A_2. Setting the line of action as a coordinate axis with an origin at its intersection with the line of centres called the pitch point, *P*, coordinate *s* is defined and is positive in the direction of motion of the contact point.

Traditionally, subscript *1* is used to define the driving gear and subscript *2* denotes the driven gear. The subscript letters refer to the nomenclature of the gears: *b* – base circle; *t* – tip (addendum) of the teeth; *r* – root (dedendum) of the teeth; no letter subscript – pitch line or circle.

The formulation of Reynolds hydrodynamic equation implies that any geometry of a contact can be reduced to simple contact of a plane and a parabolic surface. Since a plane is flat, the parabola must incorporate the curvatures of the surfaces in the contact. For example, the contact of two rollers can be considered as a contact of two parabolas, which represent the curvatures of the upper and the lower rollers. The gap between the parabola and the common tangent at the contact position, $x=0$, can be formulated as:

$$z(x) = \frac{x^2}{2R} \tag{2.1}$$

So the total separation between two parabolas in contact at the origin can be determined as:

$$h_u(x) = z_1 + z_2 = \frac{x^2}{2}\left(\frac{1}{R_1} + \frac{1}{R_2}\right) \tag{2.2}$$

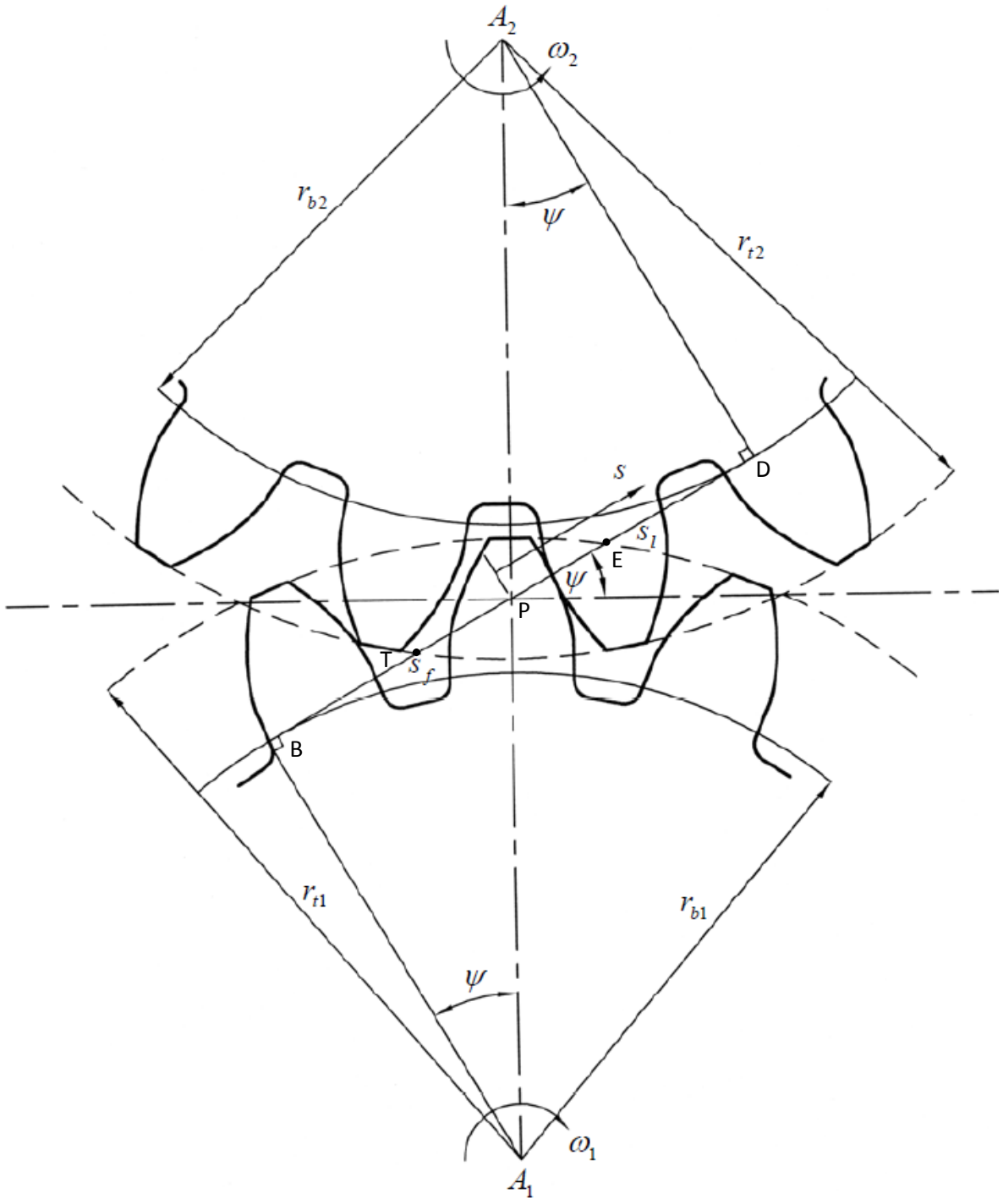


Figure 2.1 – Line of action of meshing gears (Davies, 2005)

As discussed earlier, any contact can be reduced to the contact of the plane and parabola. So the total separation or the undeformed geometry will have a form of equation (2.1), where R is the radius of relative curvature and can be determinate from the comparison of equations (2.1) and (2.2):

$$\frac{1}{R} = \left(\frac{1}{R_1} + \frac{1}{R_2} \right) \tag{2.3}$$

There are two levels of digitization of the undeformed geometry of the spur gear contact that can be used: (a) the equivalent rollers method (Martin, 1916) and (b) involute profile (Euler, 1754).

The equivalent roller method (fig. 2.2) is based on the assumption that the curvature change within the contact is insignificant due to the diminutive contact length. Hence the radii of the equivalent rollers at the specified position of the meshing cycle can be determined as:

$$\begin{aligned} R_1 &= r_{b1} \tan \psi + s \\ R_2 &= r_{b2} \tan \psi - s \end{aligned} \tag{2.4}$$

where r_{b1}, r_{b2} are base radii of the pinion and the gear respectively; ψ is the pressure angle; s is an instantaneous position of the contact on the line of action with the origin at the pitch point. These are the expressions for R_1 and R_2 because the fundamental form of the involute curve means that the centres of the radii of curvature at the contact point are always at the points B and D.

The expressions for R_1 and R_2 in equations (2.4) can be substituted into equation (2.3) to define undeformed geometry of the contacting teeth at any position in the meshing cycle according to the value of parameter s.

The second method is more complicated and more time consuming, but it increases the precision of the undeformed geometry calculations significantly by evaluating the gap between two involutes at any given value of s.

Litvin and Fuentes summarized and published a robust method to evaluate coordinates of the involute profile in 2004.

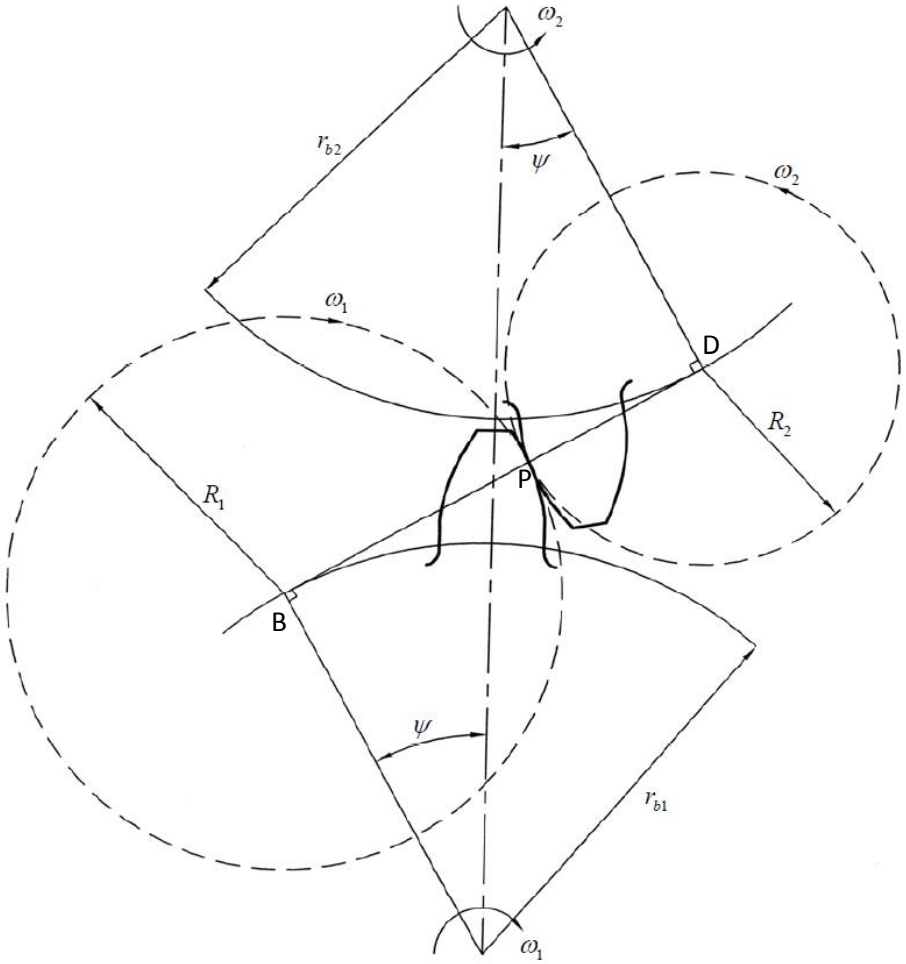


Figure 2.2 – Method of equivalent rollers (Davies, 2005)

Two branches of an involute curve are plotted in Figure 2.3. They are generated by point M_o of the straight line that rolls over the base circle radius, r_b, both clockwise and counterclockwise. Each branch represents its respective side of the tooth. The analytical representation of an involute curve is based on the following consideration:

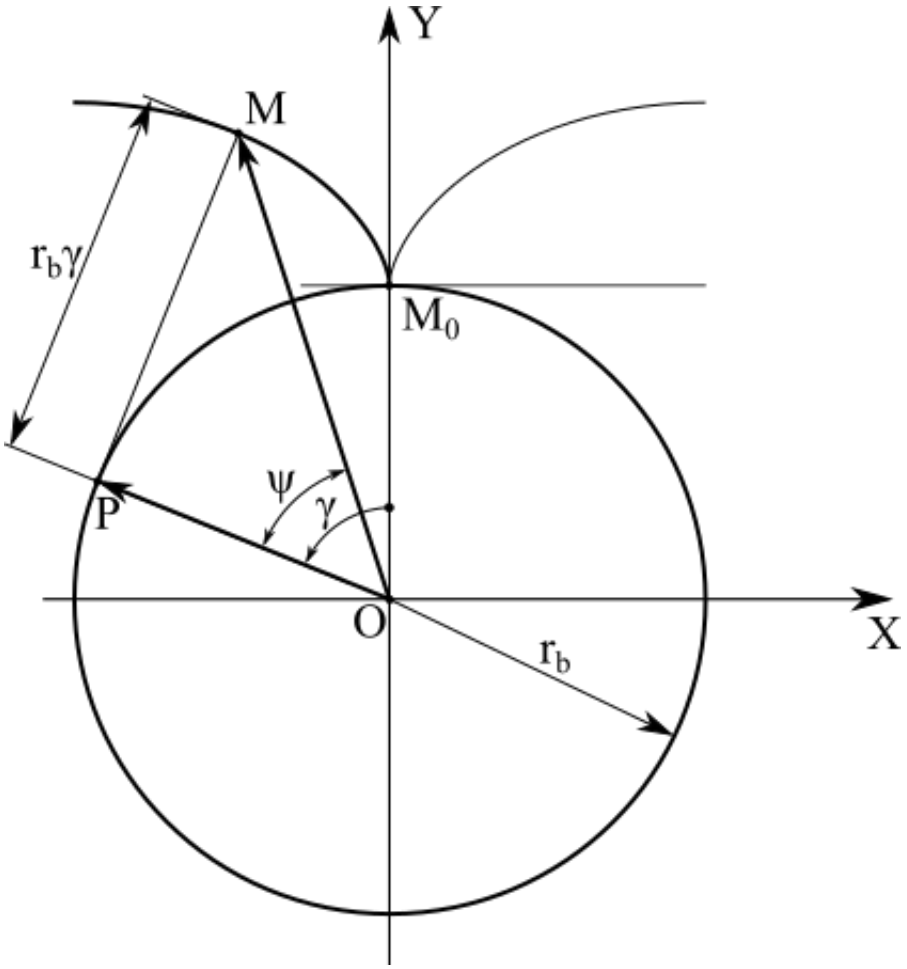


Figure 2.3 – For derivation of the equation of an involute curve

A current point M of the involute curve is determined by the vector equation

$$\overrightarrow{OM} = \overrightarrow{OP} + \overrightarrow{PM} \tag{2.5}$$

where

$$\overrightarrow{OP} = r_b[sin\gamma \quad cos\gamma]^T$$

$$\overrightarrow{PM} = PM[-cos\gamma \quad sin\gamma]^T$$

i. Due to rolling without sliding, we have

$$PM = \widehat{M_oP} = r_b\gamma \tag{2.6}$$

Here γ (rad) is the angle of rotation of rolling motion.

ii. Equations (2.5) and (2.6) give

$$\begin{bmatrix} X \\ Y \end{bmatrix} = r_b \begin{bmatrix} sin\gamma & -cos\gamma \\ cos\gamma & sin\gamma \end{bmatrix} \begin{bmatrix} 1 \\ \gamma \end{bmatrix} \tag{2.7}$$

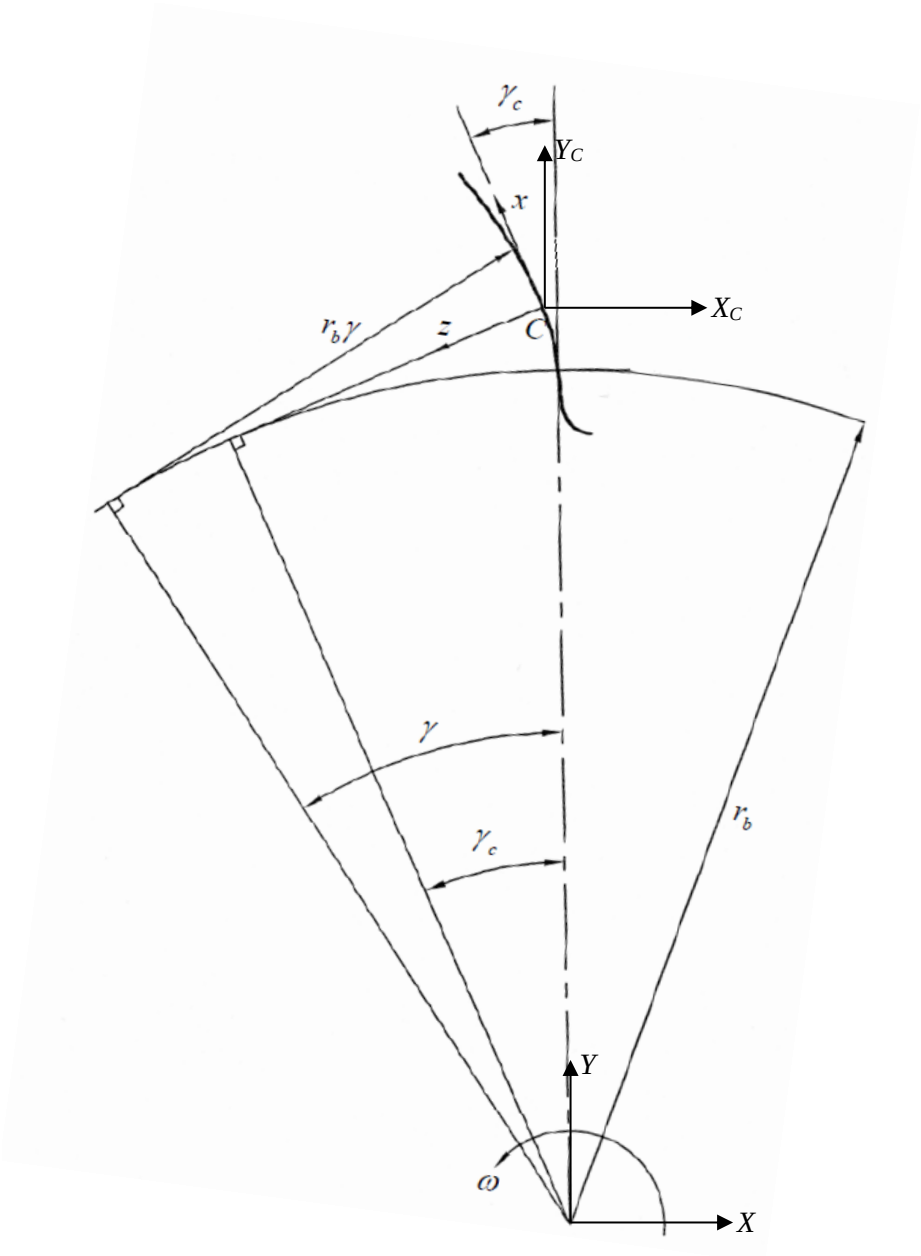


Figure 2.4 – For Coordinate transformation

Equation (2.7) is valid for the coordinate system with an origin at the centre of the gear, but the EHL problem is formulated around the contact point. In order to obtain coordinates of the point of interest in a new coordinate system with an origin at the contact point:

$$\begin{bmatrix} X \\ Y \end{bmatrix} - \begin{bmatrix} X_C \\ Y_C \end{bmatrix} = r_b \left(\begin{bmatrix} sin\gamma & -cos\gamma \\ cos\gamma & sin\gamma \end{bmatrix} \begin{bmatrix} 1 \\ \gamma \end{bmatrix} - \begin{bmatrix} sin\gamma_C & -cos\gamma_C \\ cos\gamma_C & sin\gamma_C \end{bmatrix} \begin{bmatrix} 1 \\ \gamma_C \end{bmatrix} \right)$$
$$= r_b \begin{bmatrix} sin\gamma - \gamma cos\gamma - sin\gamma_C + \gamma_C cos\gamma_C \\ cos\gamma + \gamma sin\gamma - cos\gamma_C - \gamma_C sin\gamma_C \end{bmatrix} \quad (2.8)$$

The EHL axis *x* is parallel to the tangent plane at the contact point and is positive in the direction of rotation and axis *z* is normal to that plane and is positive towards the centre of the tooth. It means that the current coordinate system must be rotated by γ_C, as it is shown in Figure 2.4. It can be achieved by multiplying equation (2.8) by the rotation matrix.

$$\begin{bmatrix} z \\ x \end{bmatrix} = r_b \begin{bmatrix} cos\gamma_C & -sin\gamma_C \\ sin\gamma_C & cos\gamma_C \end{bmatrix} \begin{bmatrix} sin\gamma - \gamma cos\gamma - sin\gamma_C + \gamma_C cos\gamma_C \\ cos\gamma + \gamma sin\gamma - cos\gamma_C - \gamma_C sin\gamma_C \end{bmatrix} \quad (2.9)$$

The coordinate axes z_1 and z_2 are directed into the tooth material of the driving and driven gears respectively, so they are opposite vectors with tails at the contact point of the mated teeth. The coordinate axis *x* must be common for both teeth; therefore, the sign of *x* in the equation (2.9) is inverted. Finally, the coordinates of the involute profiles are:

$$\begin{bmatrix} z_2 \\ x \end{bmatrix} = r_b \begin{bmatrix} cos\gamma_{C2} & -sin\gamma_{C2} \\ sin\gamma_{C2} & cos\gamma_{C2} \end{bmatrix} \begin{bmatrix} sin\gamma_2 - \gamma_2 cos\gamma_2 - sin\gamma_{C2} + \gamma_{C2} cos\gamma_{C2} \\ cos\gamma_2 + \gamma_2 sin\gamma_2 - cos\gamma_{C2} - \gamma_{C2} sin\gamma_{C2} \end{bmatrix} \quad (2.10)$$
$$\begin{bmatrix} z_1 \\ x \end{bmatrix} = r_b \begin{bmatrix} cos\gamma_{C1} & -sin\gamma_{C1} \\ -sin\gamma_{C1} & -cos\gamma_{C1} \end{bmatrix} \begin{bmatrix} sin\gamma_1 - \gamma_1 cos\gamma_1 - sin\gamma_{C1} + \gamma_{C1} cos\gamma_{C1} \\ cos\gamma_1 + \gamma_1 sin\gamma_1 - cos\gamma_{C1} - \gamma_{C1} sin\gamma_{C1} \end{bmatrix}$$

The contact point translates along the line of action and over the mating tooth flanks as the meshing cycle progresses. The distance from the point of tangency the base circle and the line of action, *P*, (Figure 2.3) and the contact point, *M*, can be obtained from the equations (2.4), but, due to the rolling without sliding, it must be equal to the arc length of the involute from the base circle to the contact point, equation (2.6).

Therefore, the reference angles of the contact point in the coordinate system of the driving, γ_{C1}, and driven, γ_{C2}, gears are:

$$\gamma_{C1} = \frac{r_{b1}\tan\varphi + s}{r_{b1}} \qquad (2.11)$$

$$\gamma_{C2} = \frac{r_{b2}\tan\varphi - s}{r_{b2}}$$

The reference angles at the meshing point, γ_1 or γ_2, can be determined from the x part of equations (2.10), which allows z_1 and z_2 to be evaluated. The total separation between two involutes is the undeformed geometry of the contact:

$$h_u(x) = z_1 + z_2 \qquad (2.12)$$

The motion of the contact point along the line of action is shown in Figure 2.5. The gear revolves about the centre of rotation is point A with constant angular velocity $\underline{\omega}_1 = \omega_1 \underline{\hat{k}}$. Therefore, the velocity of point C is the cross product of $\underline{\omega}_1$ and $\underline{A_1C}$, which can be represented as a vector sum of $\underline{A_1B}$ and $\underline{BC}$. Thus

$$v_C = \underline{\omega}_1 \times \underline{A_1 C} = \underline{\omega}_1 \times \left(\underline{A_1 B} + \underline{BC}\right)$$

The length of the vector $\underline{A_1B}$ is equal to the radius of the base circle, r_{b1}, and it is collinear to the unit vector $\underline{\hat{i}}$. $\underline{BC}$ is collinear to the unit vector $\underline{\hat{j}}$ and equal to R_1 as in equation (2.4). Hence

$$v_C = \omega_1 \underline{\hat{k}} \times \left(r_{b1}\underline{\hat{i}} + R_1 \underline{\hat{j}}\right) = \omega_1 r_{b1}\underline{\hat{j}} - \omega_1 R_1 \underline{\hat{i}} \qquad (2.13)$$

Given that the gear ratio of the meshing involute spur gears is constant, the point of contact travels along the line of action at constant velocity from equation (2.13):

$$v_{contact} = r_{b1}\omega_1 = r_{b2}\omega_2$$

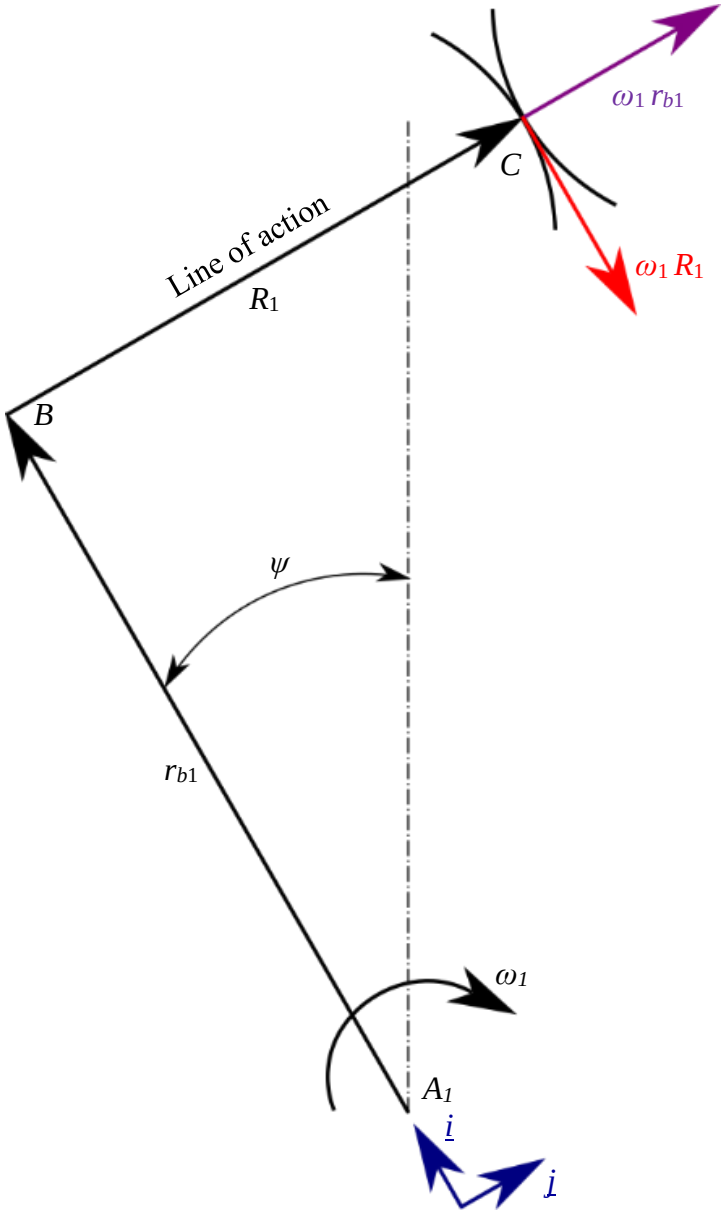


Figure 2.5 – Kinematics of the contact

The tangential velocity of the mating tooth flanks at the contact point can be obtained by substitution of equations (2.4) to equation (2.13):

$$u_1 = \left(r_{b1} \tan\psi + s\right)\omega_1 \qquad (2.14)$$

$$u_2 = \left(r_{b2} \tan\psi - s\right)\omega_2$$

As shown above, the contact geometry and operating conditions are formulated in terms of position of the contact point on the line of action relative to the pitch point. It is necessary to specify the boundary values of s, which are the point where the mating gear flanks first make contact, s_f, and last make contact, s_l, as well as the highest and the lowest points of a single tooth contact, s_{c1} and s_{c2}. The latter two are important for the simplified loading conditions, covered in Section 4, and as a reference for the analysis results processing.

According to Figure 2.1, the gear teeth first make the contact at the point of intersection, T, of the addendum circle of the driven gear, radius r_{t2}, and the line of action, BD. Bearing in mind that the origin is set at the pitch point, P, the coordinate s_f can be determined as the length of PT. The line of action is tangential to the base circle at D, therefore, A_2D is normal to PD and TB and can be defined:

$$PD = r_{b2} \tan\psi$$

$$TD = \sqrt{r_{t2}^2 - r_{b2}^2}$$

$$s_f = PD - TD = r_{b2} \tan\psi - \sqrt{r_{t2}^2 - r_{b2}^2} \tag{2.15}$$

Correspondingly:

$$PB = r_{b1} \tan\psi$$

$$EB = \sqrt{r_{t1}^2 - r_{b1}^2}$$

$$s_l = EB - PB = \sqrt{r_{t1}^2 - r_{b1}^2} - r_{b1} \tan\psi \tag{2.16}$$

The boundary of the single tooth contact zone can be determined as:

$$s_{c1} = s_l - \frac{2\pi r_{b1}}{n_1} \tag{2.17}$$

$$s_{c2} = s_f + \frac{2\pi r_{b2}}{n_2}$$

Where n_1 and n_2 are the numbers of teeth of the driving and driven gears respectively.

2.3. The Elastic deformation equation

The elastically deformed geometry, $h(x)$, can be represented as a combination of the undeformed geometry, $h_u(x)$, explained in section 2, the constant separation term , C, used to achieve appropriate load conditions and the deflection, $\Lambda(x)$. Those terms can be seen in Figure 2.6 and as follows:

$$h\ (x) = \Lambda\ (x) + h_u\ (x) + C \tag{2.19}$$

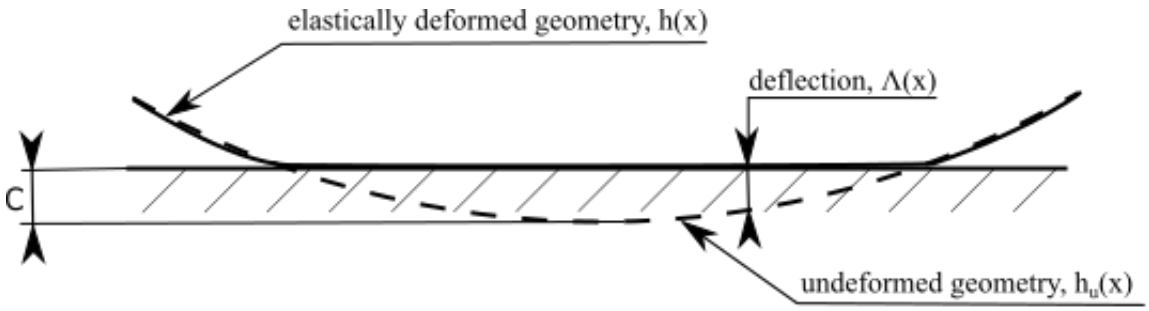


Figure 2.6 – Components of the elastic film thickness

The deflection equation for semi-infinite body was formulated by Johnson (1985):

$$\Lambda(x) = -\frac{4}{\pi E'}\int_a^b p(s)\ln\left|\frac{x-s}{r-s}\right|ds$$

This expression effectively evaluates the deflection at the point of interest x, relative to the point $x=r$, it can be simplified to:

$$\Lambda(x) = -\frac{4}{\pi E'}\int_a^b p(s)\ln|x-s|ds + D \tag{2.20}$$

The surface roughness profile, $\phi(x)$, can be incorporated into equation (2.19), therefore:

$$h\ (x) = \Lambda\ (x) + h_u\ (x) + \phi\ (x) + C \tag{2.21}$$

2.4. Load conditions over meshing cycle of spur gear contact

The variation of the shared tooth load over the meshing cycle can be determined statically and dynamically and depends on the gear flank profile and modifications. The influence of different types of the profile adjustments on the static and dynamic loading conditions are provided, for example, in NASA-TM-101444/AVSCOM-TR-88-C-003. The static shared load function, is easy to define due to its dependence only on the geometry of the contact and the stage of the meshing cycle. Hence, it is widely used for different types of spur gear analysis. The schematic distribution of the load sharing function across the gear flank profile shown in Figure 2.7, where LPSTC and HPSTC stand for the lowest and highest point of single tooth contact respectively.

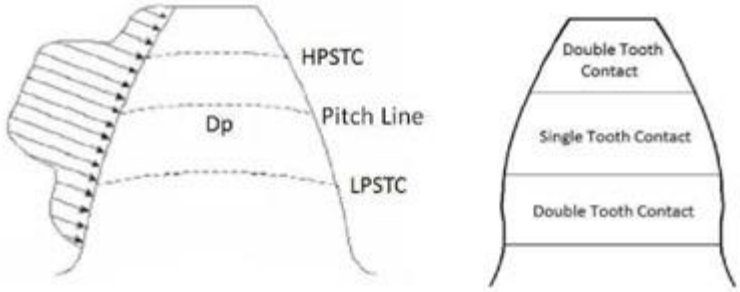


Figure 2.7 – Schematic load distribution and Single and Double tooth contact zones (Imrek, 2009)

Two load sharing functions plotted across of the coordinate on the line of action, s, in Figure 2.8. The load function shown in Figure 2.8(a) does not take into account lubricant behaviour and tooth deflection at the root, therefore, at the double tooth contact (DTC) zones, $[s_f, s_{c1}]$ and $[s_{c2}, s_l]$, the load has the constant magnitude of the half of the maximum constant load at the single tooth contact (STC) area, $[s_{c1}, s_{c2}]$. The second load function is influenced by the tooth deflection and presence of the lubricant, hence, the initial, at s_f, and final, at s_l, values of the meshing cycle are three times lower than the constant maximum load at the STC, , $[s_{c1}, s_{c2}]$ zone and gradually increase over $[s_f, s_{c1}]$ and decrease through $[s_{c2}, s_l]$. At the LPSTC, s_{c1}, the load instantaneously rise from two thirds to unity of the maximum load, and there is the instant drop of the load at HPSTC, s_{c2}. The coordinates s_f, s_{c1}, s_{c2} and s_l are defined by equations 2.15-2.17.

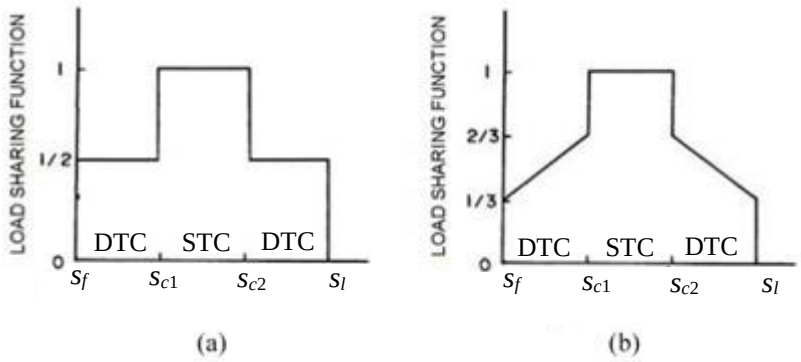


Figure 2.8 – Schematic load distribution and Single, STC, and Double, DTC, tooth contact zones (ASM, 1992)

However, the static load-sharing functions provide a crude approximation of the load variation over the meshing cycle, therefore, the dynamic load function shall be used whenever the measured values are available. One of the dynamic gear tooth load measuring techniques is described in NASA-TM-103281/AVSCOM-TR-90-C-023. The measurements are presented in NASA/TM-2005-213958/ARL-TR-3134 and shown in Figure 2.9.

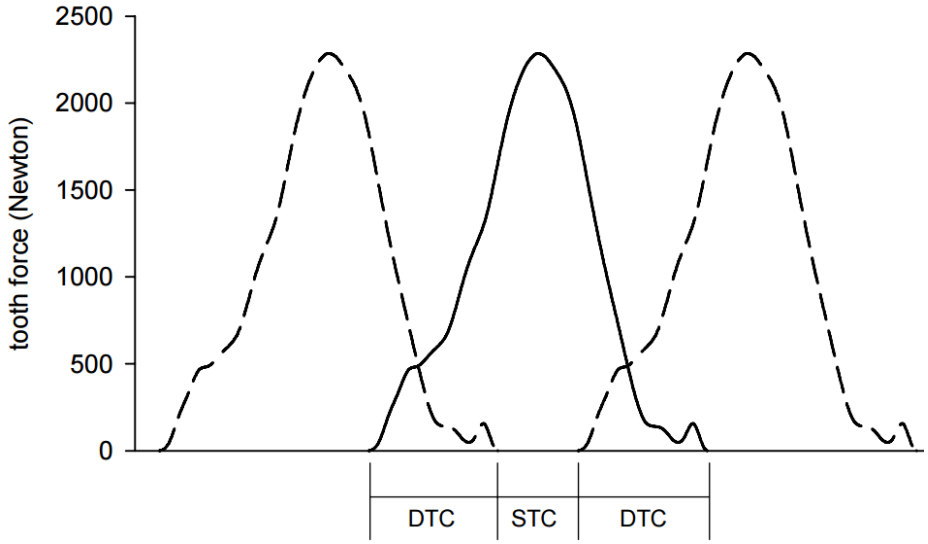


Figure 2.9 – Measured dynamic tooth force. The solid line is the measured data, the dashed lines are replicates of the measured data spaced along the ordinate at the equivalent of one tooth pitch. (NASA/TM-2005-213958/ARL-TR-3134)

2.5. The viscosity equation

Gear tooth contact is non-conformal in that high load is distributed over a very small area. The concentrated contact pressure is high and can easily exceed 1GPa. As experiments show, viscosity is a function of pressure and temperature, however, the response to the variation of the temperature is less significant within the average gear operation temperatures. Viscosity increases dramatically with a rise of pressure so its behaviour must be taken into account. There are two common relationships that are used:

Barus (1893):

$$\eta = \eta_0 \, e^{\alpha p} \tag{2.22}$$

where η_0 is the viscosity at reference (atmospheric) pressure in N-s/m^2 and α is the pressure–viscosity coefficient of units m^2/N;

and Roelands (1966):

$$\eta = \eta_0 \, e^{\alpha^* p} \tag{2.23}$$

where

$$\alpha^* p = [\ln(\eta_0) + 9.67)] \left\{ \left(\frac{\theta - 138}{\theta_0 - 138} \right)^{-S_0} \left[\left(1 + \frac{p}{p_0} \right)^{Z} - 1 \right] \right\}$$

where $p_0 = 1.98 \times 10^8$ Pa, θ_0 is a reference or ambient temperature, θ_0 and θ are in K, η and η_0 are in Pa.s, p is in Pa.

Since both relationships describe the same behaviour the calculated values at low pressures must be similar. It can be done by equating the slopes of ln(η) at p=0:

$$Z = \frac{\alpha}{5.1 \times 10^{-9}(\ln(\eta_0) + 9.67))}$$

$$S_0 = \frac{\beta(\theta_0 - 138)}{\ln(\eta_0) + 9.67}$$

As it can be seen Z and S_0 are constants of oil, independent of temperature and pressure. η_0, β and α are usually given for well known oils.

An example of the variation of the absolute viscosity due to pressure for three different lubricants calculated by equations (2.22) and (2.23) shown in Figure 2.10.

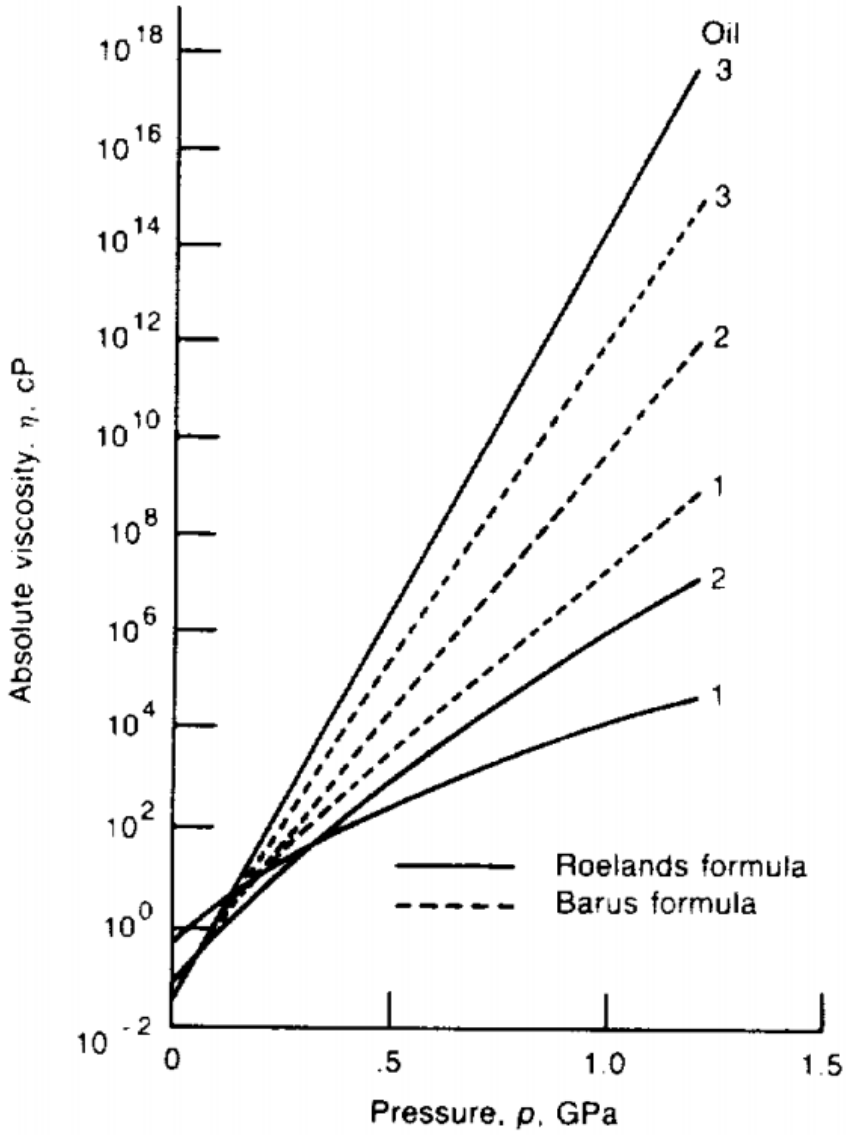


Figure 2.10 – Comparison of absolute viscosity obtained from Barus (dashed) and Roelands (solid) formulas for three different lubricants at 311 K: 1 – synthetic paraffinic oil; 2 – supperrefined naphthetic mineral oil; 3 – synthetic hydrocarbon (NASA RP – 1255 / Hamrock, 1994)

2.6. The density equation

Another liquid parameter that depends on pressure and temperature is density (Fig 1). It is necessary to know its behaviour to accurately determine the film shape. The corresponding variation of density with temperature is considered insignificant. A commonly used relationship describing its behaviour was proposed by Dowson and Higginson (1966):

$$\rho = \rho_0 \left(\frac{1 + \gamma_\rho p}{1 + \lambda_\rho p} \right)$$

where γ_ρ and λ_ρ are the compressibility constants, ρ_0 is the reference density of units kg/m^3. Typical values of γ_ρ and λ_ρ are 2.266 GPa^{-1} and 1.683 GPa^{-1} respectively.

As stated variation of density with temperature is considered insignificant but non-conformal contact leads to high pressures and, probably, significant temperature changes. In this case, it is quite improvident to ignore a temperature impact. Sui and Sadeghi (1991) incorporated a linear temperature influence into Dowson and Higginson relationship so that it became

$$\rho = \rho_0 \left(\frac{1 + \gamma_\rho p}{1 + \lambda_\rho p} \right) [1 - \varepsilon(\theta - \theta_0)]$$

2.7. The Hydrodynamic Reynolds equation

The governing equation of the hydrodynamic behaviour, which relates the pressure and the film thickness of the lubricant is based on two physical principles: the conservation of mass and the Navier-Strokes equation that describes the motion of a viscous fluid. This equation was named after Osborne Reynolds, who was the first scientist to derive it in 1886. Due to complexity of the nature of the problem, he simplified it by introducing following assumptions:

i. Flow is laminar and inertia terms are negligible (Reynolds number is small);
ii. There's no slip at the liquid-solid interfaces;
iii. Shear stress is proportional to shear strain (the lubricant behaves in a Newtonian manner);
iv. The pressure is constant through the thickness of the film, therefore the viscosity and the density do NOT deviate across the lubricant, so they can only alter in the plane of the film;
v. The body forces are negligible;
vi. The dimensions of the contact zone, including the contact length and the radii of the surfaces' curvature, are much greater than the film thickness.

As was mentioned in the section 2, based on the first and sixth assumptions, the geometry of any contact can be reduced to simple contact of a plane and a parabolic surface. A general control volume shown in Figure 2.11 is a subject to force due to the pressure and the shear stress acting on its boundary. The resultant force in the *x* direction is:

$$F_x = \Delta z\left(p - \frac{\Delta x}{2}\frac{\partial p}{\partial x}\right) - \Delta z\left(p + \frac{\Delta x}{2}\frac{\partial p}{\partial x}\right) + \Delta x\left(\tau - \frac{\Delta z}{2}\frac{\partial \tau}{\partial z}\right) - \Delta x\left(\tau - \frac{\Delta z}{2}\frac{\partial \tau}{\partial z}\right)$$

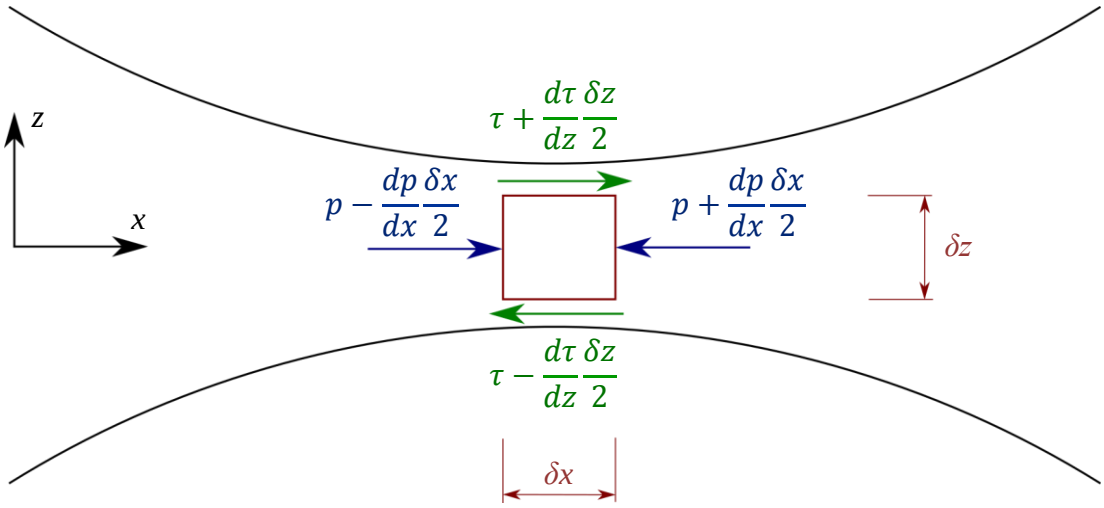


Figure 2.11 – Force equilibrium for a finite liquid control volume

Considering the conservation of the momentum and the absence of the inertial effect in the *x* direction:

$$F_x = \Delta x \Delta z \left(\frac{\partial p}{\partial x} - \frac{\partial \tau}{\partial z} \right) = 0$$

If dimensions of the control element tend to zero:

$$\frac{\partial p}{\partial x} = \frac{\partial \tau}{\partial z} \tag{2.24}$$

For a Newtonian fluid, considering the scale of the contact shear stress and shear strain can be related as

$$\tau = \eta \dot{\gamma} = \eta \left(\frac{\partial u}{\partial z} + \frac{\partial w}{\partial x} \right) = \eta \frac{\partial u}{\partial z} \tag{2.25}$$

as $\frac{\partial u}{\partial z} >> \frac{\partial w}{\partial x}$ from scale considerations.

Substituting equation (2.24) to (2.25) gives

$$\frac{\partial p}{\partial x} = \eta \frac{\partial^2 u}{\partial z^2} \tag{2.26}$$

Since the pressure is assumed to be constant across the film thickness, i.e. does not vary with respect to *z*, equation 2.26 can be integrated twice with respect to z gives the general solution of the velocity profile as:

$$\frac{\partial u}{\partial z} = \frac{z}{\eta}\frac{\partial p}{\partial x} + A$$

$$u = \frac{z^2}{2\eta}\frac{\partial p}{\partial x} + Az + B \tag{2.27}$$

It should be noted that the integration above only applies to the isothermal conditions. Taking into account temperature variation across the film thickness can make the derivation more intricate and include thermal effects.

The particular solution of the velocity profile can be found by applying boundary conditions from the assumption that slip does not occur at the surfaces, which means that the lubricant in contact must travel at the same speed as the surfaces.

$$u = u_1, \text{ at } z = \frac{h}{2} \quad \text{and} \quad u = u_2, \text{ at } z = -\frac{h}{2}$$

So

$$u_1 = \frac{h^2}{8\eta}\frac{\partial p}{\partial x} + \frac{Ah}{2} + B \tag{2.28}$$

$$u_2 = \frac{h^2}{8\eta}\frac{\partial p}{\partial x} - \frac{Ah}{2} + B \tag{2.29}$$

Summation of equations 2.28 and 2.29 gives

$$B = \frac{u_1 + u_2}{2} - \frac{h^2}{8\eta}\frac{\partial p}{\partial x}$$

Subtraction of equation 2.29 from 2.28 gives

$$A = \frac{u_1 - u_2}{h}$$

Substituting A and B at 2.27

$$u = \frac{z^2}{2\eta}\frac{\partial p}{\partial x} + \frac{z(u_1 - u_2)}{h} + \frac{u_1 + u_2}{2} - \frac{h^2}{8\eta}\frac{\partial p}{\partial x}$$

$$\therefore u = \frac{4z^2 - h^2}{8\eta}\frac{\partial p}{\partial x} + \frac{z(u_1 - u_2)}{h} + \frac{u_1 + u_2}{2} \qquad (2.30)$$

It can be noted that the velocity profile incorporates two components:

- The velocity caused by motion of the surfaces, also referred as the Couette flow after M.F.A. Couette (1858 – 1943):

$$u_{CF} = \frac{z(u_1 - u_2)}{h} + \frac{u_1 + u_2}{2}$$

- The velocity caused by the pressure gradient, also referred as the Poiseuille flow after J.L.M. Poiseuille (1797 – 1869):

$$u_{PF} = \frac{4z^2 - h^2}{8\eta}\frac{\partial p}{\partial x}$$

The principle of conservation of mass for the fixed control volume of the fluid system says that the rate of the change of the mass of the fluid system must be identical to the sum of (i) the rate of the change of the mass in the fixed control volume and (ii) the rate at which the mass leaves the fixed control volume.

$$0 = \frac{\partial(\Delta x \rho h)}{\partial t} + \left(Q + \frac{\Delta x}{2}\frac{\partial Q}{\partial x}\right) - \left(Q - \frac{\Delta x}{2}\frac{\partial Q}{\partial x}\right)$$

where Q is the mass flowrate in the x direction per unit transverse width between the solid boundaries.

Which leads to the following when Δx tends to zero:

$$\frac{\partial(\rho h)}{\partial t} + \frac{\partial Q}{\partial x} = 0 \tag{2.31}$$

The value of the mass flowrate, Q, can be obtained by integration as

$$Q = \int_{-\frac{h}{2}}^{\frac{h}{2}} \rho u(z) dz \tag{2.32}$$

Substitution of equation (2.30) to (2.32) gives

$$Q = \int_{-\frac{h}{2}}^{\frac{h}{2}} \rho \left[\frac{4z^2 - h^2}{8\eta}\frac{\partial p}{\partial x} + \frac{z(u_1 - u_2)}{h} + \frac{u_1 + u_2}{2}\right] dz$$

$$= \rho \left[\frac{4z^3 - 3h^2 z}{24\eta}\frac{\partial p}{\partial x} + \frac{z^2(u_1 - u_2)}{2h} + \frac{z(u_1 + u_2)}{2}\right]_{-\frac{h}{2}}^{\frac{h}{2}}$$

$$= \rho \left[\frac{h^3 - 3h^3}{24\eta}\frac{\partial p}{\partial x} + \frac{0 \cdot (u_1 - u_2)}{2h} + \frac{h(u_1 + u_2)}{2}\right]$$

Leading to the result

$$Q = \rho\left[\frac{h(u_1+u_2)}{2} - \frac{h^3}{12\eta}\frac{\partial p}{\partial x}\right] \tag{2.33}$$

Defining a mean velocity, $\bar{u}$:

$$\bar{u} = \frac{u_1+u_2}{2} \tag{2.34}$$

Substitution of equation (2.34) to (2.33) gives

$$Q(x,t) = \rho\bar{u}h - \frac{\rho h^3}{12\eta}\frac{\partial p}{\partial x} \tag{2.35}$$

It can be noted that, as with the velocity profile, the mass flow incorporates the Couette flow and the Poiseuille flow:

$$Q_{CF} = \rho\bar{u}h$$

$$Q_{PF} = -\frac{\rho h^3}{12\eta}\frac{\partial p}{\partial x}$$

Substitution of equation (2.35) to (2.31) gives:

$$\frac{\partial(\rho h)}{\partial t} + \frac{\partial}{\partial x}\left(\rho\bar{u}h - \frac{\rho h^3}{12\eta}\frac{\partial p}{\partial x}\right) = 0$$

Defining a flow factor σ as

$$\sigma = \frac{\rho h^3}{12\eta} \tag{2.36a}$$

The standard form of the Reynolds equation becomes

$$\frac{\partial}{\partial x}\left(\sigma\frac{\partial p}{\partial x}\right) - \frac{\partial}{\partial x}(\rho\bar{u}h) - \frac{\partial(\rho h)}{\partial t} = 0 \tag{2.36b}$$

2.8. The Eyring Rheological model

As it was stated in the section 2.7, the Reynolds equation was derived for fluid that behaves in a Newtonian manner. This is not the case for the lubrication of gears due to the high sliding velocities and the high shear rates. Therefore, the model must be relaxed to include the impact of the Non-Newtonian behaviour. Currently, the most popular shear thinning rheological model was developed by Eyring (1936), however, there are many shear thinning and limiting shear stress rheological models at our disposal. The examples of the latter one are the Bair and Winer (1979); The Lee and Hamrock (1990), Gecim and Winer (1981) etc. In 2005 C.N. Davis presented a comprehensive comparison of the characteristic behaviour of the Non-Newtonian functions based on the rheological models, which are in wide use. Drawing a conclusion of his work, those rheological models have similar formulation and show a good agreement of the calculation results, which differing only in the region of the boundaries and have a different tolerance for numerical singularities.

All rheological theories relate shear stress and shear strain rate. The Eyring shear thinning model was used

$$\dot{\gamma} = \frac{du}{dz} = \frac{\tau_0}{\eta} sinh\left(\frac{\tau}{\tau_0}\right) \tag{2.37}$$

where τ_0 is the Eyring shear stress. The value used in this thesis is 10 MPa. This representative value is commonly used in the field in the absence of accurate friction measurement (Spikes and Jie, 2014).

Conry, Wang and Cusano (1987) integrated equation (2.24) based on the assumption that the pressure is constant through the thickness of the film, so

$$\tau = \tau_m + z\frac{dp}{dx} \tag{2.38}$$

where τ_m is a shear stress in the middle of the film thickness.

Substitution of (2.38) into (2.37) gives:

$$\dot{\gamma} = \frac{du}{dz} = \frac{\tau_0}{\eta} sinh\left(\frac{\tau_m + z\frac{dp}{dx}}{\tau_0}\right)$$

Which can be simplified by introducing:

$$\Sigma = \frac{h}{2\tau_0}\frac{dp}{dx} \tag{2.39a}$$

$$\lambda = \frac{\tau_m}{\tau_0} \tag{2.39b}$$

So

$$\dot{\gamma} = \frac{du}{dz} = \frac{\tau_0}{\eta} sinh\left(\lambda + z\frac{2\Sigma}{h}\right) \tag{2.40}$$

The general velocity profile can be found by integration of equation (2.40) with respect to *z* as

$$\int du = \int \frac{\tau_0}{\eta} sinh\left(\lambda + z\frac{2\Sigma}{h}\right) dz$$

$$u = \frac{\tau_0 h}{2\eta\Sigma} cosh\left(\lambda + z\frac{2\Sigma}{h}\right) + A \tag{2.41}$$

The specific solution can be found by applying the boundary conditions:

$$u = u_1, \text{ at } z = \frac{h}{2} \text{ and } u = u_2, \text{ at } z = -\frac{h}{2}$$

So for the upper surface velocity is

$$u_1 = \frac{\tau_0 h}{2\eta\Sigma} cosh\left(\lambda + \frac{h}{2}\frac{2\Sigma}{h}\right) + A \tag{2.42}$$

And for the lower surface velocity is

$$u_2 = \frac{\tau_0 h}{2\eta\Sigma} cosh\left(\lambda - \frac{h}{2}\frac{2\Sigma}{h}\right) + A \tag{2.43}$$

Isolating A in equation (2.43):

$$A = u_2 - \frac{\tau_0 h}{2\eta\Sigma} cosh\left(\lambda - \frac{h}{2}\frac{2\Sigma}{h}\right)$$

Substituting A into (2.41):

$$u = u_2 + \frac{\tau_0 h}{2\eta\Sigma}\left[cosh\left(\lambda + z\frac{2\Sigma}{h}\right) - cosh(\lambda - \Sigma)\right] \tag{2.44}$$

The instantaneous sliding velocity of the surfaces in the contact can be found as the difference of the velocities at the surfaces:

$$u_s = u_1 - u_2 = \frac{\tau_0 h}{2\eta\Sigma} cosh(\lambda + \Sigma) + A - \frac{\tau_0 h}{2\eta\Sigma} cosh(\lambda - \Sigma) - A$$

$$= \frac{\tau_0 h}{2\eta\Sigma}[cosh(\lambda + \Sigma) - cosh(\lambda - \Sigma)]$$

$$= \frac{\tau_0 h}{2\eta\Sigma}[cosh(\lambda)cosh(\Sigma) + sinh(\lambda)sinh(\Sigma)$$

$$= -cosh(\lambda)cosh(\Sigma) + sinh(\lambda)sinh(\Sigma)]$$

$$\therefore u_s = \frac{\tau_0 h}{\eta\Sigma} sinh(\lambda)sinh(\Sigma) \tag{2.45}$$

Isolating the term involving λ

$$sinh(\lambda) = \frac{\eta u_s}{\tau_0 h}\frac{\Sigma}{sinh(\Sigma)} \tag{2.46}$$

Equation (2.32), which defines the mass flow, was manipulated using integration by parts by Greenwood (2000):

$$Q(x,t) = \int_{-\frac{h}{2}}^{\frac{h}{2}} \rho u(z) dz = \int_{-\frac{h}{2}}^{\frac{h}{2}} \rho u(z) \frac{dz}{dz} dz = [\rho \bar{u} h]_{-\frac{h}{2}}^{\frac{h}{2}} - \int_{-\frac{h}{2}}^{\frac{h}{2}} \rho z \frac{du}{dz} dz$$

i.e.

$$Q(x,t) = \rho \bar{u} h - \int_{-\frac{h}{2}}^{\frac{h}{2}} \rho z \frac{du}{dz} dz \tag{2.47}$$

As explained in the Section 2.7

$$Q_{CF} = \rho \bar{u} h \tag{2.48}$$

$$Q_{PF} = -\rho \int_{-\frac{h}{2}}^{\frac{h}{2}} z \frac{du}{dz} dz \tag{2.49}$$

Substituting (2.39) into the mass flow term due to the Poiseuille flow (2.49) gives

$$Q_{PF} = -\frac{\rho \tau_0}{\eta} \int_{-\frac{h}{2}}^{\frac{h}{2}} z sinh\left(\lambda + z\frac{2\Sigma}{h}\right) dz$$

Integrating by parts

$$\begin{bmatrix} u = z & du = dz \\ dv = sinh\left(\lambda + z\frac{2\Sigma}{h}\right) dz & v = \frac{h}{2\Sigma} cosh\left(\lambda + z\frac{2\Sigma}{h}\right) \end{bmatrix}$$

$$Q_{PF} = -\frac{\rho \tau_0}{\eta} \left[\frac{h}{2\Sigma} \left[z \cdot cosh\left(\lambda + z\frac{2\Sigma}{h}\right) \right]_{-\frac{h}{2}}^{\frac{h}{2}} - \frac{h}{2\Sigma} \int_{-\frac{h}{2}}^{\frac{h}{2}} cosh\left(\lambda + z\frac{2\Sigma}{h}\right) dz \right]$$

$$= -\frac{\rho h \tau_0}{2\eta\Sigma}\left[\left[z \cdot cosh\left(\lambda + z\frac{2\Sigma}{h}\right)\right]_{-\frac{h}{2}}^{\frac{h}{2}} - \frac{h}{2\Sigma}\left[sinh\left(\lambda + z\frac{2\Sigma}{h}\right)\right]_{-\frac{h}{2}}^{\frac{h}{2}}\right]$$

$$= -\frac{\rho h^2 \tau_0}{4\eta\Sigma^2}\left[\Sigma[cosh(\lambda+\Sigma) + cosh(\lambda-\Sigma)] - [sinh(\lambda+\Sigma) - sinh(\lambda-\Sigma)]\right]$$

$$= -\frac{\rho h^2 \tau_0}{4\eta\Sigma^2}[2\Sigma cosh(\lambda)cosh(\Sigma) - 2cosh(\lambda)sinh(\Sigma)]$$

$$\therefore Q_{PF} = -\frac{\rho\tau_0 h^2 cosh(\lambda)}{2\Sigma^2\eta}\left(\Sigma cosh(\Sigma) - sinh(\Sigma)\right) \tag{2.50}$$

Expressing τ_0 from (2.39a)

$$\tau_0 = \frac{h}{2\Sigma}\frac{dp}{dx} \tag{2.51}$$

Substituting (2.51) into (2.50)

$$Q_{PF} = -\frac{\rho h^3 cosh(\lambda)}{4\Sigma^3\eta}\frac{dp}{dx}\left(\Sigma cosh(\Sigma) - sinh(\Sigma)\right)$$

$$\therefore Q_{PF} = -\frac{\rho h^3}{12\eta}\frac{dp}{dx}\left[\frac{3cosh(\lambda)\left(\Sigma cosh(\Sigma) - sinh(\Sigma)\right)}{\Sigma^3}\right] \tag{2.52}$$

Equation (2.36b) has the same form as (2.52), so, the flow factor (2.36a) becomes

$$\sigma = \frac{\rho h^3}{12\eta}S \tag{2.53}$$

Where Non-Newtonian correction to the viscosity is

$$S = \frac{3cosh(\lambda)\left(\Sigma cosh(\Sigma) - sinh(\Sigma)\right)}{\Sigma^3} \tag{2.54}$$

Since

$$cosh^2(\alpha) + sinh^2(\alpha) = 1$$

Substituting (2.46) into (2.54) gives

$$S = \frac{3\left(\Sigma cosh(\Sigma) - sinh(\Sigma)\right)}{\Sigma^3}\sqrt{1 + \left(\frac{\eta u_s}{\tau_0 h}\frac{\Sigma}{sinh(\Sigma)}\right)^2} \tag{2.55}$$

Therefore, an effective viscosity can be found as

$$\eta_{eff} = \frac{\eta}{S}$$

Chapter 3

Numerical theory for the Finite EHL model

3.1. Introduction

The theory summarised in Chapter 2 was used to set up the EHL model. This chapter provides details on the numerical representation of that model, tools used to solve it, major numerical difficulties faced and a discussion of the results. The numerical solution of the EHL problem for spur gears was brought together by Davies (2005). That software supplies a robust implementation of the Non-Newtonian rheological line contact EHL problem to evaluate pressure distribution and film thickness between smooth or rough surfaces for the complete meshing cycle. Unfortunately, there were some limitations: the algorithm became unstable with a reduction of the λ ratio and it did not take into account a permanent change of the initial geometry that will occur through plastic deformation due to the excess pressure.

3.2. Computational Mesh

The load and the geometry of the contact varies over the meshing cycle. It means that the EHL analysis can be only formulated as a transient problem, which is a series of timesteps uniformly spaced by time intervals Δt. For spur gears, the general EHL formulation can be reduced to the one-dimensional line contact analysis (Holmes et al., 2003). The pressure distribution generated by a tooth pair is similar to that formed by two stationary parallel smooth solid cylinders in a dry contact, illustrated in

Figure 3.1, under the normal load. Hertz (1881) showed that contact width can be calculated as:

$$a = \sqrt{\frac{8Rw'}{\pi E'}} \tag{3.1}$$

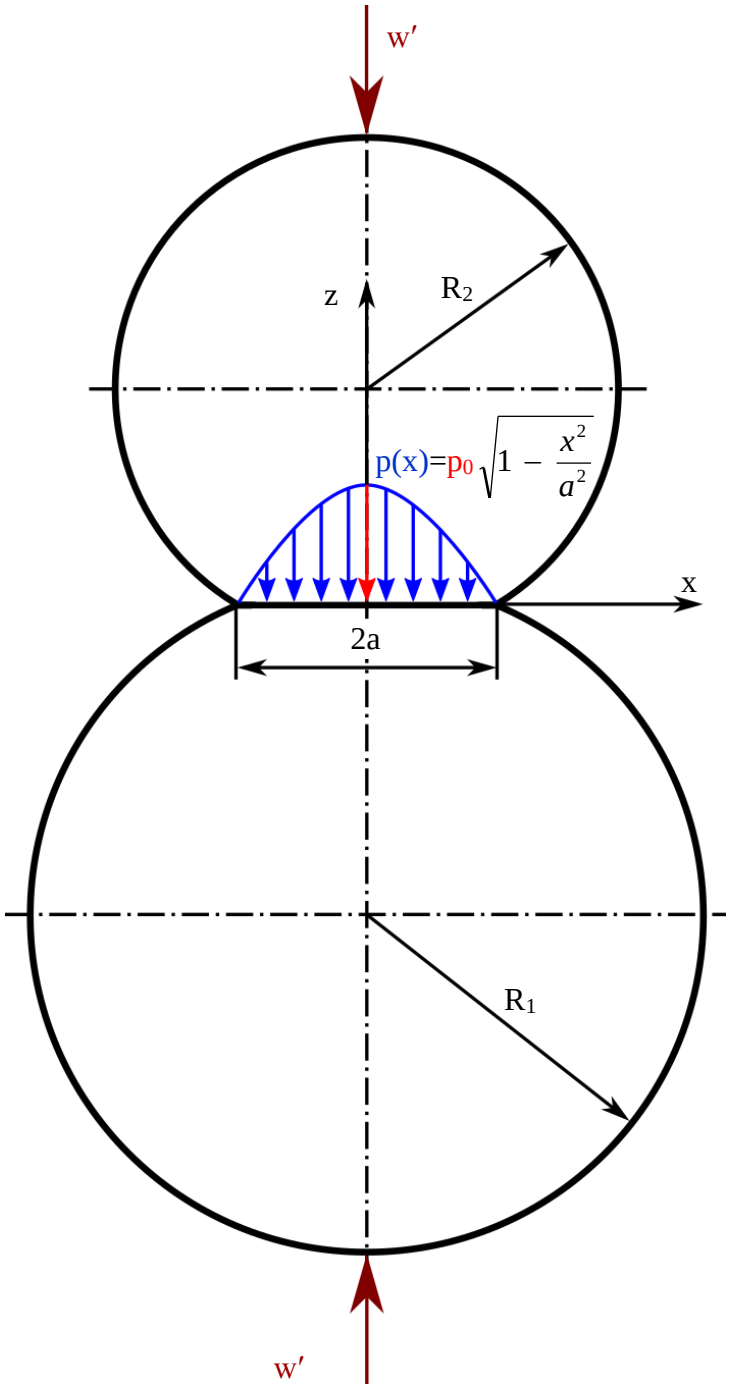


Figure 3.1 – Graphical representation of two stationary parallel smooth solid cylinders in a dry Hertzian contact

Where the equivalent elastic modulus, *E'*, that incorporates elastic properties of both solids in contact is defined as:

$$\frac{2}{E'} = \frac{1-\nu_1^2}{E_1} + \frac{1-\nu_2^2}{E_2}$$

The pressure is distributed in a semi-elliptic manner:

$$p(x) = p_0\sqrt{1 - \frac{x^2}{a^2}}$$

The maximum pressure, p_0, can be found as:

$$p_0 = \frac{2w'}{\pi a} = \sqrt{\frac{E'w'}{2\pi R}}$$

The centre of contact is chosen as the origin and the area of interest is defined in terms of the contact length upstream and downstream of the contact point. In a steady load analysis for fixed radius rollers this results in a time independent geometry and it is usual to use the Hertz contact dimension as a scaling factor in defining the spatial computing mesh. However, due to the variation of load, *w'*, over the meshing cycle the contact length, *a*, is not constant, and using it to scale the mesh would cause an inconsistency of mesh spacing from one timeframe to another. To avoid this, the tooth flanks are meshed with respect to the Hertzian contact length calculated for the maximum load. Therefore, the mesh spacing, Δx, can be defined in terms of the number of nodes per half-Hetzian maximum contact length, a_{max}.

3.3. Numerical formulation of the elastic film thickness equation

As shown in Chapter 2, the elastic film thickness equation is:

$$h(x) = \Lambda(x) + h_u(x) + \phi(x) + C \tag{3.2}$$

and the relative elastic deflection term, $\Lambda(x)$, can be evaluated as (Johnson, 1985):

$$\Lambda(x) = -\frac{4}{\pi E'}\int_a^b p(s)\ln|x-s|ds + D \tag{3.3}$$

The kernel of the integral is singular at the point at which deflection is being calculated, i.e. at x=s. This can be overcome by taking limits on either side of the singular point.

$$\Lambda(x) = -\frac{4}{\pi E'}\lim_{\delta\to 0}\left[\int_a^{x-\delta} p(s)\ln|x-s|ds + \int_{x+\delta}^b p(s)\ln|x-s|ds\right] + D \tag{3.4}$$

Numerically the relative elastic deflection can be expressed using a suitable quadrature:

$$\Lambda(x_i) = \sum_{k=1}^{n} g_{k-i}\, p_k \tag{3.5}$$

where the weighting function $g_{k\text{-}i}$ depends on the type of approximation applied to evaluate pressure between the adjacent mesh nodes, for example, Hamrock and Dowson (1974).

In 2000 Evans and Hughes devised an alternative differential formulation of the relative elastic deflection. It can be seen that the effect of the pressure at the point is more dependent on the slope of the deflection than on its value (Johnson, 1985). The advantage of this method is that pressure has an extremely localised effect on the second derivative of deflection which allows the elastic and fluid differential equations to be solved as a coupled pair. The quadrature formula of the ordinary differential equation obtained by differentiating equation (3.4) twice has weighting function that decays rapidly with increasing distance from the evaluation point as shown below.

Differentiating equation (3.4) with respect to x using the chain rule in the limits of integration gives:

$$-\frac{\pi E'}{4}\frac{d\Lambda(x)}{dx}=\lim_{\delta\to 0}\left[\,p(x-\delta)\ln|\delta|+\int_{a}^{x-\delta}\frac{p(s)}{x-s}ds-p(x+\delta)\ln|\delta|+\int_{x+\delta}^{b}\frac{p(s)}{x-s}ds\right] \qquad (3.6)$$

The pressure terms can be expressed by using Taylor series expansion:

$$p(x\pm\delta)=p(x)+\sum_{n=1}^{\infty}\frac{(\pm\delta)^n}{n!}\frac{d^n p(x)}{dx^n}$$

On taking the limit:

$$\lim_{\delta\to 0}\left[\ln|\delta|\left(\,p(x)-p(x)+\sum_{n=0}^{\infty}\frac{\delta^n}{n!}\frac{d^n p(x)}{dx^n}-\sum_{n=0}^{\infty}\frac{(-\delta)^n}{n!}\frac{d^n p(x)}{dx^n}\right)\right]=0$$

Therefore, equation (3.6) becomes:

$$-\frac{\pi E'}{4}\frac{d\Lambda(x)}{dx}=\lim_{\delta\to 0}\left[\int_{a}^{x-\delta}\frac{p(s)}{x-s}ds+\int_{x+\delta}^{b}\frac{p(s)}{x-s}ds\right]$$

To obtain the second derivative, the process is repeated:

$$-\frac{\pi E'}{4}\frac{d^2\Lambda(x)}{dx^2}=\lim_{\delta\to 0}\left[\frac{p(x-\delta)}{\delta}-\int_{a}^{x-\delta}\frac{p(s)}{(x-s)^2}ds+\frac{p(x+\delta)}{\delta}-\int_{x+\delta}^{b}\frac{p(s)}{(x-s)^2}ds\right] \qquad (3.7)$$

To solve equation (3.7) numerically, a quadrature mesh is applied. Let the mesh spacing be Δ with each interval centred at the point where the relative deflection is calculated as shown in Figure 3.2.

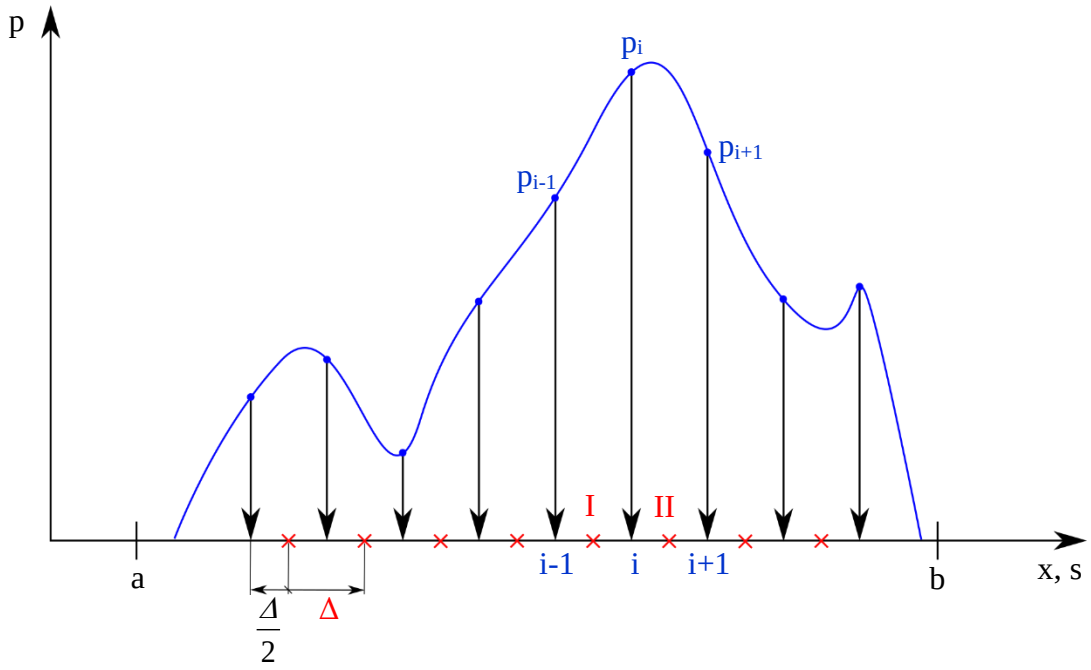


Figure 3.2 – The EHL mesh and the quadrature

The relative deflection is only singular over the domain $\left(-\frac{\Delta}{2},\frac{\Delta}{2}\right)$. Therefore, the equation (3.7) can be written in the form:

$$-\frac{\pi E'}{4}\frac{d^2\Lambda(x)}{dx^2}=-\int_a^{x-\frac{\Delta}{2}}\frac{p(s)}{(x-s)^2}ds-\int_{x+\frac{\Delta}{2}}^{b}\frac{p(s)}{(x-s)^2}ds$$

$$+\lim_{\delta\to 0}\left[\frac{p(x-\delta)}{\delta}-\int_{x-\frac{\Delta}{2}}^{x-\delta}\frac{p(s)}{(x-s)^2}ds+\frac{p(x+\delta)}{\delta}-\int_{x+\delta}^{x+\frac{\Delta}{2}}\frac{p(s)}{(x-s)^2}ds\right] \qquad (3.8)$$

Using the Taylor series expansion and considering the pressure distribution is continuous and differentiable over that range, it can be expressed in terms of $z = s - x$ as:

$$p(x+z)=p(x)+z\frac{dp(x)}{dx}+\frac{z^2}{2}\frac{d^2p(x)}{dx^2} \qquad (3.9)$$

By substitution of the equation (3.9) into the limit term of the equation (3.8) becomes:

$$\lim_{\delta\to 0}\left[\frac{p(x)-\delta\frac{dp(x)}{dx}+\frac{\delta^2}{2}\frac{d^2p(x)}{dx^2}}{\delta}+\frac{p(x)+\delta\frac{dp(x)}{dx}+\frac{\delta^2}{2}\frac{d^2p(x)}{dx^2}}{\delta}\right.$$

$$\left.-\int_{-\frac{\Delta}{2}}^{-\delta}\frac{p(x)+z\frac{dp(x)}{dx}+\frac{z^2}{2}\frac{d^2p(x)}{dx^2}}{z^2}dz-\int_{\delta}^{\frac{\Delta}{2}}\frac{p(x)+z\frac{dp(x)}{dx}+\frac{z^2}{2}\frac{d^2p(x)}{dx^2}}{z^2}dz\right]$$

$$=\lim_{\delta\to 0}\left[\frac{p(x)}{\delta}+\frac{\delta}{2}\frac{d^2p(x)}{dx^2}+\frac{p(x)}{\delta}+\frac{\delta}{2}\frac{d^2p(x)}{dx^2}-p(x)\int_{-\frac{\Delta}{2}}^{-\delta}\frac{dz}{z^2}-\frac{dp(x)}{dx}\int_{-\frac{\Delta}{2}}^{-\delta}\frac{dz}{z}-\frac{1}{2}\frac{d^2p(x)}{dx^2}\int_{-\frac{\Delta}{2}}^{-\delta}dz\right.$$

$$\left.-p(x)\int_{\delta}^{\frac{\Delta}{2}}\frac{dz}{z^2}-\frac{dp(x)}{dx}\int_{\delta}^{\frac{\Delta}{2}}\frac{dz}{z}-\frac{1}{2}\frac{d^2p(x)}{dx^2}\int_{\delta}^{\frac{\Delta}{2}}dz\right]$$

$$=\lim_{\delta\to 0}\left[2\frac{p(x)}{\delta}+\delta\frac{d^2p(x)}{dx^2}-p(x)\left[-\frac{1}{z}\right]_{-\frac{\Delta}{2}}^{-\delta}-\frac{dp(x)}{dx}\left[\ln|z|\right]_{-\frac{\Delta}{2}}^{-\delta}-\frac{1}{2}\frac{d^2p(x)}{dx^2}\left[z\right]_{-\frac{\Delta}{2}}^{-\delta}\right.$$

$$\left.+p(x)\left[-\frac{1}{z}\right]_{\frac{\Delta}{2}}^{\delta}+\frac{dp(x)}{dx}\left[\ln|z|\right]_{\frac{\Delta}{2}}^{\delta}+\frac{1}{2}\frac{d^2p(x)}{dx^2}\left[z\right]_{\frac{\Delta}{2}}^{\delta}\right]$$

$$=\lim_{\delta\to 0}\left[2\frac{p(x)}{\delta}+\delta\frac{d^2p(x)}{dx^2}-p(x)\left(\frac{1}{\delta}-\frac{2}{\Delta}\right)-\frac{dp(x)}{dx}\ln\left(\frac{2\delta}{\Delta}\right)+\frac{1}{2}\frac{d^2p(x)}{dx^2}\left(\delta-\frac{\Delta}{2}\right)\right.$$

$$\left.-p(x)\left(\frac{1}{\delta}-\frac{2}{\Delta}\right)+\frac{dp(x)}{dx}\ln\left(\frac{2\delta}{\Delta}\right)+\frac{1}{2}\frac{d^2p(x)}{dx^2}\left(\delta-\frac{\Delta}{2}\right)\right]$$

$$=\lim_{\delta\to 0}\left[2\frac{p(x)}{\delta}+\delta\frac{d^2p(x)}{dx^2}-2p(x)\left(\frac{1}{\delta}-\frac{2}{\Delta}\right)+\frac{d^2p(x)}{dx^2}\left(\delta-\frac{\Delta}{2}\right)\right]$$

$$=\lim_{\delta\to 0}\left[2\frac{p(x)}{\delta}-2\frac{p(x)}{\delta}+\delta\frac{d^2p(x)}{dx^2}+\delta\frac{d^2p(x)}{dx^2}+\frac{4p(x)}{\Delta}-\frac{\Delta}{2}\frac{d^2p(x)}{dx^2}\right]$$

$$= \frac{d^2 p(x)}{dx^2} \lim_{\delta \to 0} [2\delta] + \frac{4p(x)}{\Delta} - \frac{\Delta}{2} \frac{d^2 p(x)}{dx^2}$$

$$= \frac{4p(x)}{\Delta} - \frac{\Delta}{2} \frac{d^2 p(x)}{dx^2}$$

Therefore, equation (3.9) becomes:

$$\frac{d^2 \Lambda(x)}{dx^2} = \frac{4}{\pi E'} \left[\int_a^{x-\frac{\Delta}{2}} \frac{p(s)}{(x-s)^2} ds + \int_{x+\frac{\Delta}{2}}^{b} \frac{p(s)}{(x-s)^2} ds - \frac{4p(s)}{\Delta} + \frac{\Delta}{2} \frac{d^2 p(s)}{dx^2} \right] \quad (3.10)$$

The integral terms of the equation (3.10) can be evaluated numerically by application of the quadrature defined earlier and shown in Figure 3.2. Let $s–x = z$, then new limits of integration are:

$$\begin{aligned} s &= a & z &= a - x \\ s &= x - \frac{\Delta}{2} & z &= x - \frac{\Delta}{2} - x = -\frac{\Delta}{2} \\ s &= x + \frac{\Delta}{2} & z &= x + \frac{\Delta}{2} - x = \frac{\Delta}{2} \\ s &= b & z &= b - x \end{aligned}$$

So

$$\int_a^{x-\frac{\Delta}{2}} \frac{p(s)}{(x-s)^2} ds + \int_{x+\frac{\Delta}{2}}^{b} \frac{p(s)}{(x-s)^2} ds = \int_{a-x}^{-\frac{\Delta}{2}} \frac{p(z+x)}{(-z)^2} dz + \int_{\frac{\Delta}{2}}^{b-x} \frac{p(z+x)}{(-z)^2} dz$$

Evidently, $z = 0$ is the point where $s = x$. For a signed integer, i, $z = i\Delta$ is the centre of the i^{th} interval from x. Thus

$$\int_a^{x-\frac{\Delta}{2}} \frac{p(s)}{(x-s)^2} ds + \int_{x+\frac{\Delta}{2}}^{b} \frac{p(s)}{(x-s)^2} ds = \sum_{\substack{all\, i \\ i \neq 0}} \int_{(i-0.5)\Delta}^{(i+0.5)\Delta} \frac{p(x+z)}{z^2} dz \quad (3.11)$$

The pressure over domain $[(i-0.5)\Delta,(i+0.5)\Delta]$ can be approximated by a parabola:

$$p(x+z)=A_i+B_i\frac{z}{\Delta}+C_i\left(\frac{z}{\Delta}\right)^2 \tag{3.12}$$

Then

$$\int_{(i-0.5)\Delta}^{(i+0.5)\Delta}\frac{p(x+z)}{z^2}dz=\int_{(i-0.5)\Delta}^{(i+0.5)\Delta}\frac{A_i}{z^2}dz+\int_{(i-0.5)\Delta}^{(i+0.5)\Delta}\frac{B_i}{z\Delta}dz+\int_{(i-0.5)\Delta}^{(i+0.5)\Delta}\frac{C_i}{\Delta^2}dz$$

$$=A_i\left[\frac{1}{z}\right]_{(i+0.5)\Delta}^{(i-0.5)\Delta}+\frac{B_i}{\Delta}\left[\ln|z|\right]_{(i-0.5)\Delta}^{(i+0.5)\Delta}+\frac{C_i}{\Delta^2}\left[z\right]_{(i-0.5)\Delta}^{(i+0.5)\Delta}$$

$$\therefore\int_{(i-0.5)\Delta}^{(i+0.5)\Delta}\frac{p(x+z)}{z^2}dz=\frac{4A_i}{(2i-1)(2i+1)\Delta}+\frac{B_i}{\Delta}\ln\left(\frac{(2i+1)}{(2i-1)}\right)+\frac{C_i}{\Delta} \tag{3.13}$$

Coefficients A_i, B_i and C_i can be evaluated from equation (3.12). As it can be seen in Figure 3.2, the pressure values at the nodes $(i\text{-}1)\Delta$, $i\Delta$ and $(i+1)\Delta$ are p_{i-1}, p_i and p_{i+1} respectively. Therefore

$$\begin{cases}p_{i-1}=A_i+B_i\dfrac{(i-1)\Delta}{\Delta}+C_i\left(\dfrac{(i-1)\Delta}{\Delta}\right)^2;\\ p_i=A_i+B_i\dfrac{i\Delta}{\Delta}+C_i\left(\dfrac{i\Delta}{\Delta}\right)^2;\\ p_{i+1}=A_i+B_i\dfrac{(i+1)\Delta}{\Delta}+C_i\left(\dfrac{(i+1)\Delta}{\Delta}\right)^2.\end{cases}$$

$$\begin{cases}A_i=p_{i-1}-B_i(i-1)-C_i(i-1)^2;\\ p_i=p_{i-1}+B_i i-B_i(i-1)+C_i i^2-C_i(i-1);\\ p_{i+1}=p_{i-1}+B_i(i+1)-B_i(i-1)+C_i(i+1)^2-C_i(i-1)^2.\end{cases}$$

$$\begin{cases}A_i=p_{i-1}-B_i(i-1)-C_i(i-1)^2;\\ B_i=p_i-p_{i-1}-C_i(2i-1);\\ p_{i+1}=p_{i-1}+2B_i+4iC_i.\end{cases}$$

$$\begin{cases} A_i = p_{i-1} - (i-1)\left(p_i - p_{i-1} - C_i(2i-1)\right) - C_i(i-1)^2; \\ B_i = p_i - p_{i-1} - C_i(2i-1); \\ p_{i+1} = p_{i-1} + 2p_i - 2p_{i-1} - C_i(4i-2) + 4iC_i. \end{cases}$$

$$\begin{cases} A_i = ip_{i-1} - (i-1)p_i + C_i\left(i^2 - i\right); \\ B_i = p_i - p_{i-1} - C_i(2i-1); \\ C_i = \dfrac{p_{i-1} - 2p_i + p_{i+1}}{2}. \end{cases}$$

$$\begin{cases} A_i = \dfrac{\left(i^2 + i\right)p_{i-1} - 2\left(i^2 - 1\right)p_i + \left(i^2 - i\right)p_{i+1}}{2}; \\ B_i = \dfrac{-(2i+1)p_{i-1} + 4ip_i - (2i-1)p_{i+1}}{2}; \\ C_i = \dfrac{p_{i-1} - 2p_i + p_{i+1}}{2}. \end{cases} \tag{3.14}$$

Substituting coefficients (3.14) into expression (3.13) gives:

$$\begin{aligned} \int_{(i-0.5)\Delta}^{(i+0.5)\Delta} \frac{p(x+z)}{z^2}\,dz = {} & \frac{2\left(\left(i^2 + i\right)p_{i-1} - 2\left(i^2 - 1\right)p_i + \left(i^2 - i\right)p_{i+1}\right)}{\left(4i^2 - 1\right)\Delta} \\ & - \frac{(2i+1)p_{i-1} - 4ip_i + (2i-1)p_{i+1}}{2\Delta} \ln\left(\frac{(2i+1)}{(2i-1)}\right) \\ & + \frac{p_{i-1} - 2p_i + p_{i+1}}{2\Delta} \end{aligned}$$

Separation of pressure terms gives:

$$\begin{aligned} \int_{(i-0.5)\Delta}^{(i+0.5)\Delta} \frac{p(x+z)}{z^2}\,dz = \frac{1}{2\Delta}\Bigg[& \left(\frac{8i^2 + 4i - 1}{4i^2 - 1} - (2i+1)\ln\left(\frac{(2i+1)}{(2i-1)}\right)\right)p_{i-1} \\ & + \left(\frac{10 - 16i^2}{4i^2 - 1} + 4i\ln\left(\frac{(2i+1)}{(2i-1)}\right)\right)p_i \\ & + \left(\frac{8i^2 - 4i - 1}{4i^2 - 1} - (2i-1)\ln\left(\frac{(2i+1)}{(2i-1)}\right)\right)p_{i+1}\Bigg] \end{aligned} \tag{3.15}$$

Introducing terms:

$$F_i^{(-)} = \frac{1}{2\Delta}\left(\frac{8i^2+4i-1}{4i^2-1} - (2i+1)\ln\left(\frac{(2i+1)}{(2i-1)}\right)\right), i \neq 0 \qquad (3.16a)$$

$$F_i = \frac{1}{2\Delta}\left(\frac{10-16i^2}{4i^2-1} + 4i\ln\left(\frac{(2i+1)}{(2i-1)}\right)\right), i \neq 0 \qquad (3.16b)$$

$$F_i^{(+)} = \frac{1}{2\Delta}\left(\frac{8i^2-4i-1}{4i^2-1} - (2i-1)\ln\left(\frac{(2i+1)}{(2i-1)}\right)\right), i \neq 0 \qquad (3.16c)$$

allows the equation (3.15) to be written as:

$$\int_{(i-0.5)\Delta}^{(i+0.5)\Delta} \frac{p(x+z)}{z^2}\,dz = F_i^{(-)}p_{i-1} + F_i p_{i-1} + F_i^{(+)}p_{i+1} \qquad (3.17)$$

Parameters $F_i^{(-)}$, F_i and $F_i^{(+)}$ are singular at $i = 0$ and the values of these functions at $i = 0$ can be obtained from the limiting process used in developing equation (3.10). This shows that

$$\int_{-\frac{\Delta}{2}}^{\frac{\Delta}{2}} \frac{p(x+z)}{z^2}\,dz = -\frac{4p(x)}{\Delta} + \frac{\Delta}{2}\frac{d^2p(x)}{dx^2} = -\frac{4p_0}{\Delta} + \frac{\Delta}{2}\frac{p_{-1}-2p_0+p_{+1}}{\Delta^2}$$

$$= \frac{p_{-1}-10p_0+p_{+1}}{2\Delta}$$

Therefore

$$F_0^{(-)} = F_0^{(+)} = \frac{1}{2\Delta},\ F_0 = \frac{-5}{\Delta} \qquad (3.18)$$

The relative differential deflection for any general node of the Δ spaced mesh can thus be formulated as:

$$\frac{d^2\Lambda(x_i)}{dx^2} = \frac{4}{\pi E'}\sum_{all\,k} f_{k-i}\,p_k \tag{3.19}$$

where the quadrature weighting function shown in Figure 3.3 is

$$f_{k-i} = F_{i+1}^{(-)} + F_i + F_{i-1}^{(+)} \tag{3.20}$$

This variation of $f_{k\text{-}i}$ with *k-i* is shown in Figure 3.3.

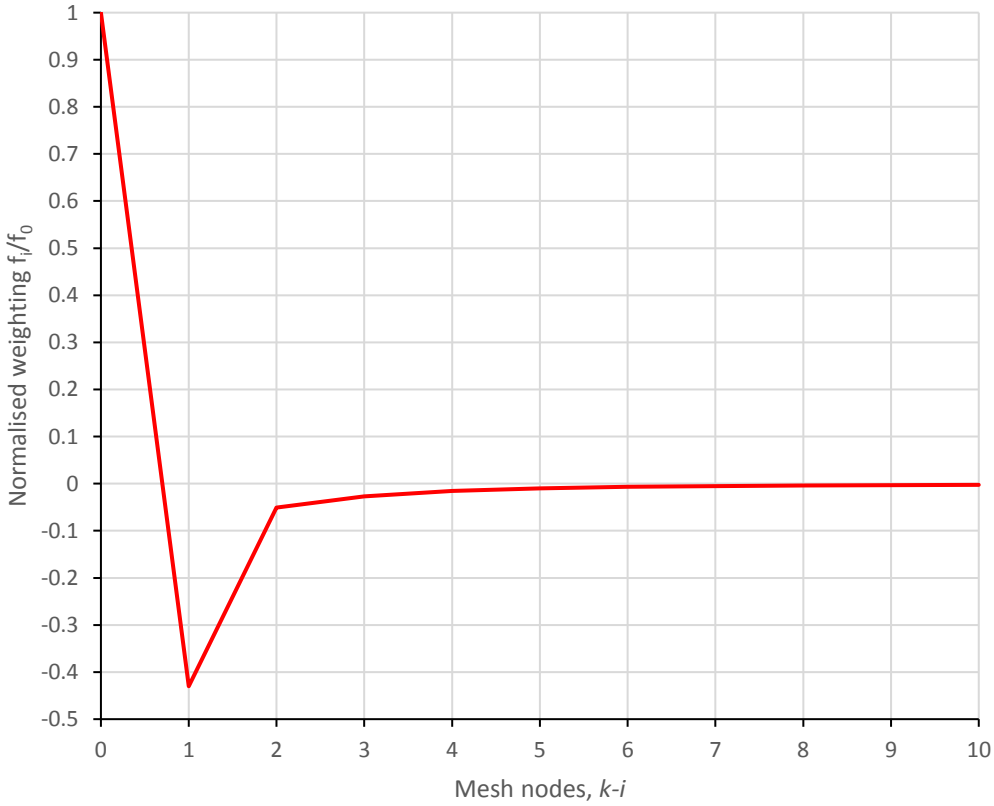


Figure 3.3 – Normalized shape of the weighting function in the neighbourhood of the point of evaluation (*i*=0)

Differentiating equation (3.2) with respect to x twice gives:

$$\frac{d^2h(x_i)}{dx^2} = \frac{d^2\Lambda(x_i)}{dx^2} + \frac{d^2h_u(x_i)}{dx^2} + \frac{d^2\phi(x_i)}{dx^2} \tag{3.21}$$

So that

$$\frac{d^2h(x_i)}{dx^2} = \frac{4}{\pi E'}\sum_{all\,k} f_{k-i}p_k + \frac{d^2h_u(x_i)}{dx^2} + \frac{d^2\phi(x_i)}{dx^2} \tag{3.22}$$

The relative differential deflection can be evaluated using the expression (3.19). The second order differentials of the profile roughness and the underformed geometry profiles can be obtained as follows:

$$\frac{d^2h_u(x_i)}{dx^2} = \frac{h_u(x_{i-1}) - 2h_u(x_i) + h_u(x_{i+1})}{\Delta^2}$$
$$\frac{d^2\phi(x_i)}{dx^2} = \frac{\phi(x_{i-1}) - 2\phi(x_i) + \phi(x_{i+1})}{\Delta^2}$$

If the surfaces in the contact are smooth, the profile roughness term can be omitted. If the simplification of the gear flank profile using the equivalent rollers method (Martin, 1916) described in section 2.2 is used, the second order derivative of the undeformed geometry profile can be reduced to $1/R$.

As can be seen in Figure 3.3, the weighting function determines that the pressure distribution in the neighbourhood of the point of evaluation has a significant influence on the relative deflection. Furthermore, the influence factors decay rapidly as index $k–i$ increases, as can be seen in Figure 3.3, where, for example, the magnitude f_5 is less than 1% of f_0 and also the pressure at the mesh point has an opposite sign. It is therefore possible to reduce the bandwidth of equation (3.22) by taking the influential pressure terms to the left-hand side and evaluating the remaining terms on the right-hand side based on previous cycle pressures.

The equation then becomes:

$$\frac{h(x_{i-1})-2h(x_i)+h(x_{i+1})}{\Delta^2}-\frac{4}{\pi E'}\sum_{|k-i|\le n_c} f_{k-i}p_k$$

$$=\frac{4}{\pi E'}\sum_{|k-i|>n_c} f_{k-i}p_k+\frac{h_u(x_{i-1})-2h_u(x_i)+h_u(x_{i+1})}{\Delta^2}+\frac{\phi(x_{i-1})-2\phi(x_i)+\phi(x_{i+1})}{\Delta^2} \quad (3.23)$$

where n_c is the bandwidth of the main matrix.

3.4. Numerical formulation of Reynolds hydrodynamic equation

As shown in Chapter 2, Reynolds hydrodynamic equation (2.36b) is:

$$\frac{\partial}{\partial x}\left(\sigma\frac{\partial p}{\partial x}\right)-\frac{\partial}{\partial x}\left(\rho\bar{u}h\right)-\frac{\partial(\rho h)}{\partial t}=0 \quad (3.24)$$

where the flow factor, σ, is

$$\sigma=\frac{\rho h^3}{12\eta}S \quad (3.25)$$

The flow factor for the isothermal formulation is a function of the film thickness and pressure, which includes the density and the viscosity terms. If the thermal conditions are taken into the account, the density and the viscosity become the functions of pressure and temperature. To accommodate the Non-Newtonian behaviour the *S* term was introduced. As was stated in the section 2.8, *S* is a function of the sliding velocity and the pressure gradient, as well as film thickness and viscosity.

Note that the flow factor, σ, is a function of pressure and, therefore, is a function of position, *x*, which makes partial differential equation (3.24) non-linear, but it can be linearized by the use of an iterative technique: values of the flow factors are calculated for a current candidate pressure distribution to find a new candidate pressure distribution and film thickness. Then the flow factors are then recalculated for the new

values of the pressure distribution and the process is repeated until the overall solution satisfies the convergence criterion. This process requires some under-relaxation in updating the candidate pressure distribution.

Reynolds hydrodynamic equation can be formulated numerically by using finite difference and finite element methods. According to Elcoate (1996) the finite central difference and finite element models are stable and enable a fast convergence for both the smooth surface analysis and the rough surface analysis.

3.5. Finite central difference formulation

By the central difference formulation for the node *i* of the mesh shown in Figure 3.2, the first term equation (3.24) can be expressed as:

$$\frac{\partial}{\partial x}\left(\sigma \frac{\partial p}{\partial x}\right) = \frac{\sigma_{II} \dfrac{p_{i+1} - p_i}{\Delta} - \sigma_I \dfrac{p_i - p_{i-1}}{\Delta}}{\Delta}$$

$$\therefore \frac{\partial}{\partial x}\left(\sigma \frac{\partial p}{\partial x}\right) = \frac{\sigma_{II} p_{i+1} + \sigma_I p_{i-1} - (\sigma_{II} + \sigma_I) p_i}{\Delta^2} \qquad (3.25)$$

Taking into account that the instantaneous velocity at the centre of the film, $\bar{u}$, is constant for the whole mesh and applying the chain rule, the second term of the equation (3.24) can be found as:

$$\frac{\partial}{\partial x}(\rho \bar{u} h) = \bar{u}\left(\rho_i \frac{\partial h}{\partial x} + h_i \frac{\partial \rho}{\partial x}\right)$$

$$\therefore \frac{\partial}{\partial x}(\rho \bar{u} h) = \bar{u}\left(\rho_i \frac{h_{i+1} - h_{i-1}}{2\Delta} + h_i \frac{\rho_{i+1} - \rho_{i-1}}{2\Delta}\right) \qquad (3.26)$$

Therefore, the equation (3.24) can be written as:

$$\frac{\partial(\rho h)}{\partial t} = \frac{\sigma_{II} p_{i+1} + \sigma_I p_{i-1} - (\sigma_{II} + \sigma_I) p_i}{\Delta^2} - \bar{u}\left(\rho_i \frac{h_{i+1} - h_{i-1}}{2\Delta} + h_i \frac{\rho_{i+1} - \rho_{i-1}}{2\Delta}\right) \quad (3.27)$$

where σ_I and σ_{II} are the values of the flow factor at the midpoints between mesh positions $i - 1$, i and $i + 1$, as shown in Figure 3.2.

The Crank-Nicolson (1947) method is used to evaluate the time dependent term of the equation (3.24). The implicit numerical method is a combination of the explicit forward Euler method and the implicit backward Euler method: it is based on the trapezoidal rule, giving second-order convergence in time. The Crank-Nicolson method is unconditionally stable (Thomas, 1995). Equation (3.27), expressed in this way, becomes:

$$\frac{(\rho h)_i^{k+1} - (\rho h)_i^k}{\delta t} = \frac{1}{2}\left(\left[\frac{\sigma_{II} p_{i+1} + \sigma_I p_{i-1} - (\sigma_{II} + \sigma_I) p_i}{\Delta^2} - \bar{u}\left(\rho_i \frac{h_{i+1} - h_{i-1}}{2\Delta} + h_i \frac{\rho_{i+1} - \rho_{i-1}}{2\Delta}\right)\right]^k + \left[\frac{\sigma_{II} p_{i+1} + \sigma_I p_{i-1} - (\sigma_{II} + \sigma_I) p_i}{\Delta^2} - \bar{u}\left(\rho_i \frac{h_{i+1} - h_{i-1}}{2\Delta} + h_i \frac{\rho_{i+1} - \rho_{i-1}}{2\Delta}\right)\right]^{k+1}\right) \quad (3.28)$$

where k and k+1 are successive timesteps (and not powers).

By moving all k terms to the left-hand side and all $k+1$ terms to the right-hand side:

$$-\left[\frac{\sigma_{II} p_{i+1} + \sigma_I p_{i-1} - (\sigma_{II} + \sigma_I) p_i}{\Delta^2} - \bar{u}\left(\rho_i \frac{h_{i+1} - h_{i-1}}{2\Delta} + h_i \frac{\rho_{i+1} - \rho_{i-1}}{2\Delta}\right) - \frac{2\rho_i h_i}{\delta t}\right]^{k+1}$$

$$= \left[\frac{\sigma_{II} p_{i+1} + \sigma_I p_{i-1} - (\sigma_{II} + \sigma_I) p_i}{\Delta^2} - \bar{u}\left(\rho_i \frac{h_{i+1} - h_{i-1}}{2\Delta} + h_i \frac{\rho_{i+1} - \rho_{i-1}}{2\Delta}\right) + \frac{2\rho_i h_i}{\delta t}\right]^k$$

$$-\left[\frac{\sigma_I}{\Delta^2}p_{i-1}+\frac{\bar{u}\rho_i}{2\Delta}h_{i-1}-\frac{\sigma_I+\sigma_{II}}{\Delta^2}p_i-\left(\frac{\rho_{i+1}-\rho_{i-1}}{\Delta}\bar{u}+\frac{2\rho_i}{\delta t}\right)h_i+\frac{\sigma_{II}}{\Delta^2}p_{i+1}-\frac{\bar{u}\rho_i}{2\Delta}h_{i+1}\right]^{k+1}$$

$$=\left[\frac{\sigma_I}{\Delta^2}p_{i-1}+\frac{\bar{u}\rho_i}{2\Delta}h_{i-1}-\frac{\sigma_I+\sigma_{II}}{\Delta^2}p_i-\left(\frac{\rho_{i+1}-\rho_{i-1}}{\Delta}\bar{u}-\frac{2\rho_i}{\delta t}\right)h_i+\frac{\sigma_{II}}{\Delta^2}p_{i+1}-\frac{\bar{u}\rho_i}{2\Delta}h_{i+1}\right]^{k} \tag{3.29}$$

3.6. Finite element formulation

The finite element method is a numerical technique for solving partial differential equations by discretising these equations in their space dimensions into finite number of suitably shaped elements. These elements are interconnected and subdivided by a finite number of Gauss points. The Galerkin Weighted Residual, GWR, method was used for the Reynolds hydrodynamic equation formulation:

$$\int N_i\frac{d}{dx}\left(\sigma\frac{dp}{dx}\right)dx-\int N_i\frac{d}{dx}\left(\rho\bar{u}h\right)dx-\int N_i\frac{\partial(\rho h)}{\partial t}dx=0 \tag{3.30}$$

where N_i is the shape function that interpolates tabulated independent variable, which is pressure, and a set of dependant variable, such as viscosity, density and film thickness.

The equation (3.30) can be reformulated by applying the chain rule to the Couette flow term and by integrating by parts the Poiseuille term:

$$\left[N_i\sigma\frac{dp}{dx}\right]-\int\frac{dN_i}{dx}\sigma\frac{dp}{dx}dx-\int N_i\left(\rho h\frac{d\bar{u}}{dx}+\bar{u}h\frac{d\rho}{dx}+\rho\bar{u}\frac{dh}{dx}\right)dx$$

$$-\int N_i\frac{\partial(\rho h)}{\partial t}dx=0 \tag{3.31}$$

The quantities $\left[N_i\sigma\frac{dp}{dx}\right]$ are the element boundary conditions. They are equal and opposite at the common nodes of adjacent elements; therefore, they are cancelled out

when the full problem matrix is assembled and can be ignored in the model formulation.

At either end of the mesh, the pressure gradient is zero. At the inlet zone it can be achieved by setting pressure at the first two mesh nodes equal to zero, but the position of the downstream boundary is initially unknown. It can be established automatically during the iterative solving process by setting the negative pressure values at the outlet zone, which correspond to cavitated elements, to zero.

The Crank-Nicolson method is used to evaluate the time dependent term of the equation (3.31):

$$\left[\int N_i\,\rho\,h\,dx+\frac{\Delta_t}{2}\left(\begin{array}{l}\int\frac{dN_i}{d\,x}\sigma\frac{d\,p}{d\,x}\,dx\\+\int N_i\left(\rho\,h\frac{d\,\bar{u}}{d\,x}+\bar{u}\,h\frac{d\,\rho}{d\,x}+\rho\,\bar{u}\,\frac{d\,h}{d\,x}\right)dx\end{array}\right)\right]^{k+1}$$

$$=\left[\int N_i\,\rho\,h\,dx-\frac{\Delta_t}{2}\left(\begin{array}{l}\int\frac{dN_i}{d\,x}\sigma\frac{d\,p}{d\,x}\,dx\\+\int N_i\left(\rho\,h\frac{d\,\bar{u}}{d\,x}+\bar{u}\,h\frac{d\,\rho}{d\,x}+\rho\,\bar{u}\,\frac{d\,h}{d\,x}\right)dx\end{array}\right)\right]^{k} \qquad (3.31)$$

If it is assumed, that the tabulated function varying linearly inside each element, then:

$$f(x)=a_0+a_1x$$

To evaluate two unknown terms a_0 and a_1 we need at least two values of the tabulated function:

$$\begin{cases}f(x_1)=a_0+a_1x_1,\\ f(x_2)=a_0+a_1x_2.\end{cases}\quad\Rightarrow\quad\begin{array}{l}a_0=\frac{x_2}{x_2-x_1}f(x_1)-\frac{x_1}{x_2-x_1}f(x_2)\\ a_1=-\frac{1}{x_2-x_1}f(x_1)+\frac{1}{x_2-x_1}f(x_2)\end{array}$$

Therefore

$$f(x)=\frac{x_2-x}{x_2-x_1}f(x_1)+\frac{x-x_1}{x_2-x_1}f(x_2)=N_1(x)f(x_1)+N_2(x)f(x_2)$$

If the Lagrange interpolation functions are formulated over element length [-0.5dx, 0.5dx], as shown in Figure 3.4, the shape functions become:

$$N_1(x) = \frac{x_2 - x}{x_2 - x_1} = \frac{1 - x}{2} \qquad \text{and} \qquad N_2(x) = \frac{x - x_1}{x_2 - x_1} = \frac{x - 1}{2}$$

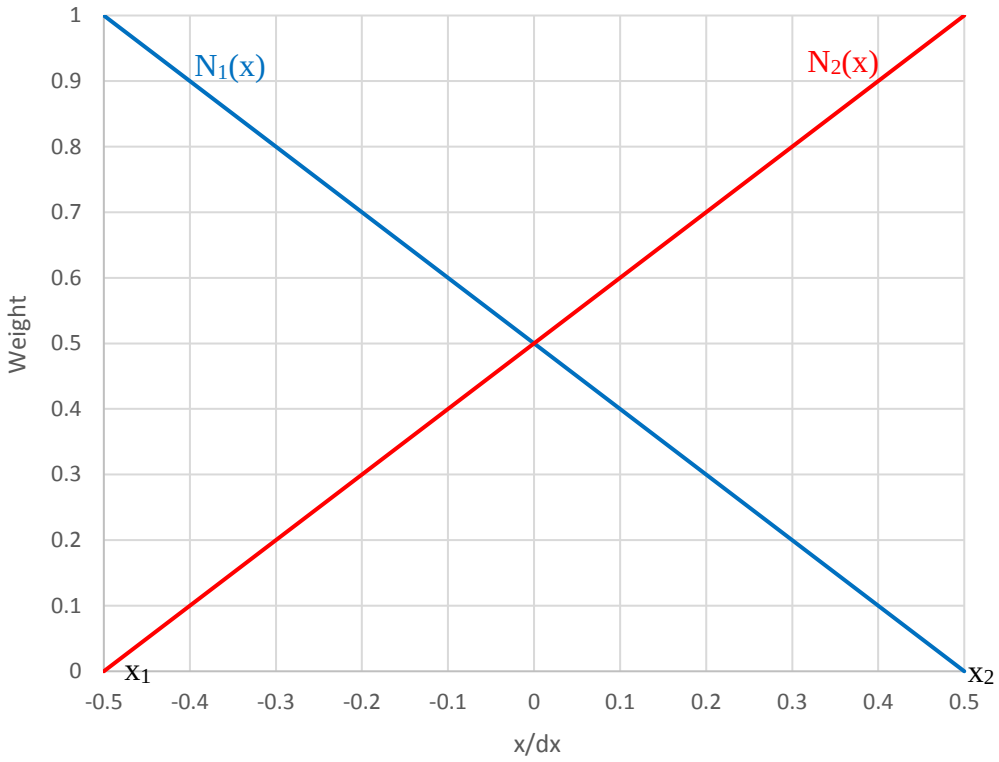


Figure 3.4 – Shape functions of a linear finite element

The first derivatives of the shape functions are:

$$\frac{dN_1(x)}{dx} = \frac{-x}{x_2 - x_1} = -\frac{1}{2} \qquad \text{and} \qquad \frac{dN_2(x)}{dx} = \frac{x}{x_2 - x_1} = \frac{1}{2}$$

In case of a quadratic approximation, the tabulated function is:

$$f(x) = a_0 + a_1 x + a_2 x^2$$

There are three unknown terms a_0, a_1 and a_2, so, a system of three equations is needed:

$$\begin{cases} f\left(x_1\right) = a_0 + a_1 x_1 + a_2 x_1^2 \\ f\left(x_2\right) = a_0 + a_1 x_2 + a_2 x_2^2 \\ f\left(x_3\right) = a_0 + a_1 x_3 + a_2 x_3^2 \end{cases}$$

The Lagrange interpolation functions defined at [-dx, dx], shown in Figure 3.5:

$$N_1\left(x\right) = \frac{\left(x_2 - x\right)\cdot\left(x_3 - x\right)}{\left(x_2 - x_1\right)\cdot\left(x_3 - x_1\right)} = \frac{-x\cdot\left(1-x\right)}{2}$$

$$N_2\left(x\right) = \frac{\left(x_1 - x\right)\cdot\left(x_3 - x\right)}{\left(x_1 - x_2\right)\cdot\left(x_3 - x_2\right)} = 1 - x^2$$

$$N_3\left(x\right) = \frac{\left(x_1 - x\right)\cdot\left(x_2 - x\right)}{\left(x_1 - x_3\right)\cdot\left(x_2 - x_3\right)} = \frac{x\cdot\left(1+x\right)}{2}$$

The first derivatives of the shape functions are:

$$\frac{dN_1\left(x\right)}{\mathrm{dx}} = \frac{-1+2x}{2} \qquad \frac{dN_2\left(x\right)}{\mathrm{dx}} = -2x \qquad \frac{dN_3\left(x\right)}{\mathrm{dx}} = \frac{1+2x}{2}$$

The integral terms of the equation (3.31) can be evaluated on an elemental level by application of the Gauss-Legendre quadrature:

$$\int_{-\frac{el}{2}}^{\frac{el}{2}} f\left(x\right) = \sum_{i=1}^{ngp} c_i \cdot f\left(x_i\right)\cdot el$$

where *el* is the element length; *ngp* is a number of integration Gauss points; x_i and c_i are positions and weights of the quadrature points respectively, shown for two, three and four point quadrature in Table 3.1.

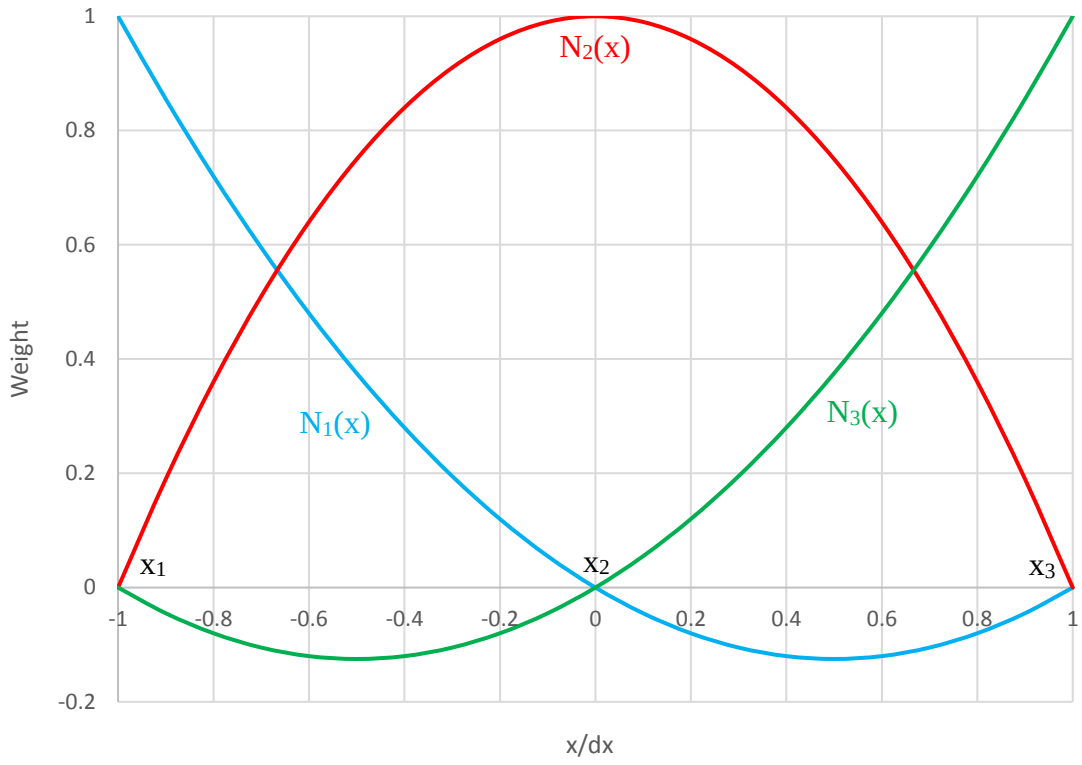


Figure 3.5 – Shape functions of a quadratic finite element

Table 3.1 – Gauss-Legendre integration points

Number of points	Position (x_i)	Weight (c_i)
2	$-\sqrt{\frac{1}{3}}$	1
	$\sqrt{\frac{1}{3}}$	1
3	$-\sqrt{\frac{3}{5}}$	$\frac{5}{9}$
	0	$\frac{8}{9}$
	$\sqrt{\frac{3}{5}}$	$\frac{5}{9}$

Table 3.1 – Gauss-Legendre integration points (continued)

Number of points	Position (x_i)	Weight (c_i)
4	$-\sqrt{\frac{3}{7}+\frac{2}{7}\sqrt{\frac{6}{5}}}$	$\frac{18-\sqrt{30}}{36}$
	$-\sqrt{\frac{3}{7}-\frac{2}{7}\sqrt{\frac{6}{5}}}$	$\frac{18+\sqrt{30}}{36}$
	$\sqrt{\frac{3}{7}-\frac{2}{7}\sqrt{\frac{6}{5}}}$	$\frac{18+\sqrt{30}}{36}$
	$\sqrt{\frac{3}{7}+\frac{2}{7}\sqrt{\frac{6}{5}}}$	$\frac{18-\sqrt{30}}{36}$

The average terms evaluated at the Gauss points have the same format, and they can be summarized in a single equation as

$$\tilde{y} = \sum_{i=1}^{ngp} N_i \cdot y_i \tag{3.32}$$

So, finally, the equation (3.31) can be written as

$$\left[\int_{-\frac{el}{2}}^{\frac{el}{2}} N_i \, \tilde{\rho} \left(\sum_{j=1}^{ep} N_j h_j \right) dx + \frac{\Delta_t}{2} \left(\begin{array}{l} \int_{-\frac{el}{2}}^{\frac{el}{2}} \frac{dN_i}{dx} \tilde{\sigma} \left(\sum_{j=1}^{ep} \frac{dN_j}{dx} p_j \right) dx \\ + \int_{-\frac{el}{2}}^{\frac{el}{2}} N_i \, \tilde{\rho} \, \tilde{\tilde{u}} \left(\sum_{j=1}^{ep} \frac{dN_j}{dx} h_j \right) dx \\ + \int_{-\frac{el}{2}}^{\frac{el}{2}} N_i \left(\tilde{\rho} \frac{d\tilde{\tilde{u}}}{dx} + \tilde{\tilde{u}} \frac{d\tilde{\rho}}{dx} \right) \left(\sum_{j=1}^{ep} N_j h_j \right) dx \end{array} \right) \right]^{k+1}$$

$$= \left[\int N_i \, \tilde{\rho} \, \tilde{h} \, dx - \frac{\Delta_t}{2} \left(\begin{array}{l} \int \frac{dN_i}{dx} \tilde{\sigma} \frac{d\tilde{p}}{dx} dx \\ + \int N_i \left(\tilde{\rho} \, \tilde{h} \frac{d\tilde{\tilde{u}}}{dx} + \tilde{\tilde{u}} \, \tilde{h} \frac{d\tilde{\rho}}{dx} + \tilde{\rho} \, \tilde{\tilde{u}} \frac{d\tilde{h}}{dx} \right) dx \end{array} \right) \right]^{k} \tag{3.33}$$

3.7. Matrix formulation and the solving process

The Reynolds (3.29) or (3.33) and elastic film thickness (3.23) equations are formulated to produce a banded matrix and can be expressed as:

$$\begin{cases} \sum_{k=1}^{n_c} A_k p_k + \sum_{k=1}^{n_c} B_k h_k = R_i \\ \sum_{k=1}^{n_c} C_k p_k + \sum_{k=1}^{n_c} D_k h_k = E_i \end{cases} \tag{3.34}$$

where A_k and B_k are the pressure and film thickness coefficients of the Reynolds equation respectively; R_i incorporates the values from the previous timestep solution; C_k and D_k are the pressure and film thickness coefficients of the elastic deflection equation; E_i determines the contribution of the pressure distribution over the problem formulation bandwidth, n_c, excluding the i mesh node.

Based on the operating load, the undeformed geometry of the contact, the mechanical properties of the flank material and the lubricant, the initial candidate pressure distribution and film thickness are set as for the Hertzian contact, i.e. the parabolic pressure function as in Figure 3.1 and the flat gap between the surfaces in the Hertz contact zone. Therefore, the density and the viscosity at the mesh nodes can be evaluated, subsequently, allowing the coefficients of the Reynolds equation, A_k and B_k, to be obtained and to make the system linear. The coefficients C_k and D_k do not depend on the pressure and the film thickness values.

The converged timestep solution can be obtained using either the Gaussian elimination and back substitution or the iterative Gauss-Seidel methods. The Gaussian elimination algorithm for a banded matrix is relatively simple and a ready-to-use code is available, for example, Press *et al. (1992*). It obtains the candidate solution in two sweeps: the matrix modification and backsubstitution. However, it cannot reformulate the problem to accommodate the occurrence of the dry contact and produces a negative film thickness instead, which must be addressed elsewhere and then, after the reformulation of the problem by setting A_k and B_k to zero, the system of equations (3.34) must be resolved again.

The Gauss-Seidel iterative algorithms, including the black-red one, involve a modification of the system of equations (3.34) by moving all terms except the diagonal ones for the pressure and the film thickness to the right-hand side as

$$\begin{cases} A_i p_i + B_i h_i = R_i - \sum\limits_{\substack{k=1 \\ \text{except i}}}^{n_c} A_k p_k + \sum\limits_{\substack{k=1 \\ \text{except i}}}^{n_c} B_k h_k = \hat{R}_i \\ C_i p_i + D_i h_i = E_i - \sum\limits_{\substack{k=1 \\ \text{except i}}}^{n_c} C_k p_k + \sum\limits_{\substack{k=1 \\ \text{except i}}}^{n_c} D_k h_k = \hat{E}_i \end{cases} \tag{3.35}$$

These equations are solved as a simultaneous pair of equations to give new values for p_i and h_i

$$p_i^{\text{new}} = \frac{D_i \hat{R}_i - B_i \hat{E}_i}{A_i D_i - B_i C_i} \qquad h_i^{\text{new}} = \frac{A_i \hat{E}_i - C_i \hat{R}_i}{A_i D_i - B_i C_i} \tag{3.36}$$

In case of the occurrence of the dry contact, i.e. a negative value of the film thickness, there is no fluid to separate the surfaces, hence, Reynolds equation cannot be applied and the pressure is governed by the elastic deflection formula. The film thickness must be set to zero and the candidate solution is

$$p_i^{\text{new}} = \frac{\hat{E}_i}{C_i} \qquad h_i^{\text{new}} = 0 \tag{3.37}$$

The iterative process is recursive and continues until the mean change of the candidate solution is smaller than the convergence criterion, which is set to 0.5% of the mean value of the solution (p or h) throughout this thesis.

The initial solution values are relaxed towards the new results produced by the elimination or converged by the iterative method and used to update the coefficients of equation (3.34). The process is repeated until the converged solution for the current timestep is obtained. The flow-chart of the EHL analysis is presented in Figure 5.5.

Chapter 4

Elastic stress evaluation and fatigue life expectancy analysis

4.1. Introduction

The results of the EHL analysis include tabulated functions of normal and tangential loads at the specified timesteps of the meshing cycle. This chapter describes the method to evaluate the elastic stresses in the gear flanks in the contact; provides some verification of the accuracy of this method; explains the effective method developed to sort and store the stress history as well as the output file structure; briefly chronicles the timeline of the fatigue study; spells out the nomenclature and the main definitions of the fatigue analysis; tells about the fatigue-life prediction methods used.

4.2. Elastic stress evaluation

The stress analysis of the EHL contact of two rough surfaces is a complex problem, which can only be solved using numerical methods. However, the mathematical formulation is based on the traditional established theory of elasticity.

Chapter 4 of Theory of Elasticity by Timoshenko & Goodier (1951) presents a derivation of the two-dimensional elastic equation in polar coordinates. It shows that the stress components that satisfy the equilibrium equations are:

$$\begin{cases} \sigma_r = \dfrac{1}{r}\dfrac{\partial \phi}{\partial r} + \dfrac{1}{r^2}\dfrac{\partial^2 \phi}{\partial \theta^2} \\ \sigma_\theta = \dfrac{\partial^2 \phi}{\partial r^2} \\ \tau_{r\theta} = -\dfrac{\partial}{\partial r}\left(\dfrac{1}{r}\dfrac{\partial \phi}{\partial \theta}\right) \end{cases} \tag{4.1}$$

where $\phi(r,\theta)$ is the stress function.

Flamant (1892) showed that in case of a concentrated nodal load P in a direction normal to the surface of a semi-infinite body, as shown in Figure 4.1, the stress function is

$$\phi(r,\theta) = A\, r\, \theta\, \sin\, \theta \tag{4.2}$$

where A is an arbitrary constant.

By substitution of the equation (4.2) into the equations (4.1) they become:

$$\begin{cases} \sigma_r = 2A\, \dfrac{\cos\, \theta}{r} \\ \sigma_\theta = 0 \\ \tau_{r\theta} = 0 \end{cases} \tag{4.3}$$

The angle θ varies from $-\dfrac{\pi}{2}$ at the surface to the left of the line of action through $\dfrac{\pi}{2}$ at the surface to the right of the point of application. The system is in a state of equilibrium and there is only one non-zero stress component, so that:

$$-P = \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \sigma_r\, r\, \cos\, \theta\, d\theta \tag{4.4}$$

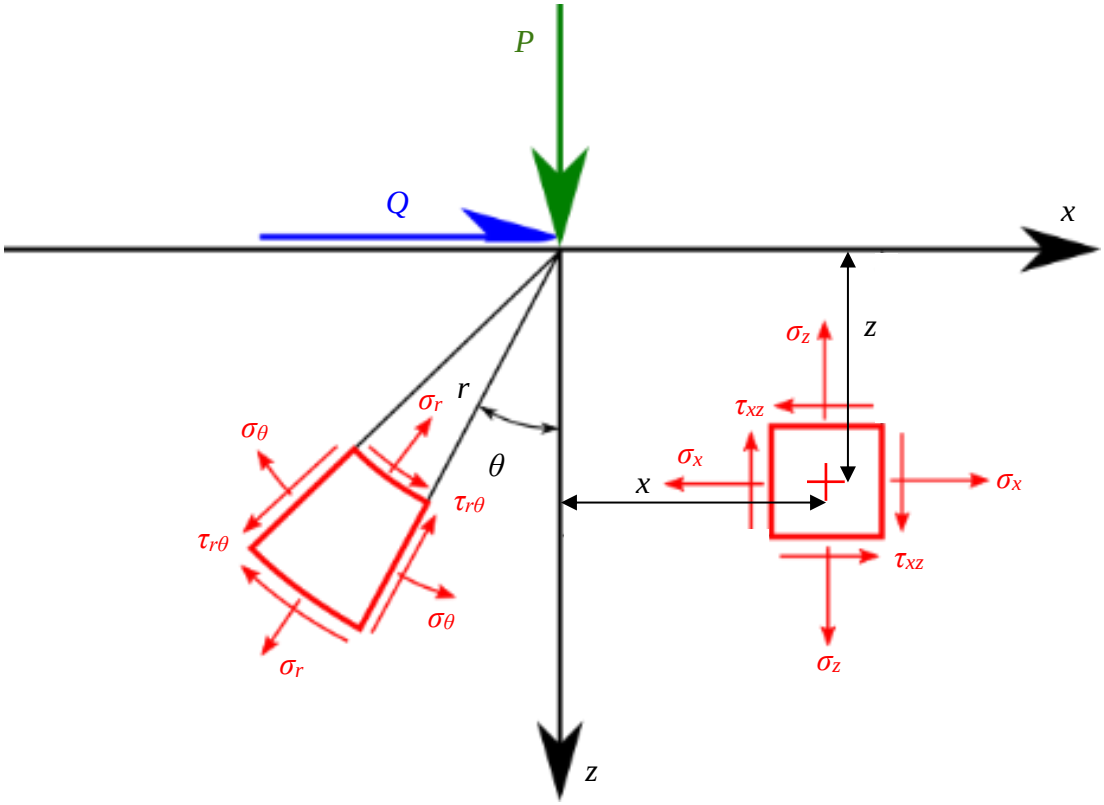


Figure 4.1 – Stress components in Cartesian and polar coordinate systems

The stress distribution is axisymmetric about the line of action of p, and σ_r is defined by equation (4.3), therefore

$$-P = \int_0^{\frac{\pi}{2}} 4A \frac{\cos\theta}{r} r \cos\theta \, d\theta = \int_0^{\frac{\pi}{2}} 4A \cos^2\theta \, d\theta$$

$$= \int_0^{\frac{\pi}{2}} 2A \left(1+\cos 2\theta\right) d\theta = A \cdot \left[2\theta + \sin 2\theta\right]_0^{\frac{\pi}{2}} = A\pi \qquad (4.4)$$

Thus

$$\sigma_r = -\frac{2P}{\pi} \frac{\cos\theta}{r} \qquad (4.5)$$

The directional stress components are obtained from the radial stress distribution

$$\begin{cases} \sigma_x = \sigma_r \sin^2 \theta = -\dfrac{2P}{\pi} \dfrac{\cos \theta \sin^2 \theta}{r} \\ \sigma_z = \sigma_r \cos^2 \theta = -\dfrac{2P}{\pi} \dfrac{\cos^3 \theta}{r} \\ \tau_{xz} = \sigma_r \sin \theta \cos \theta = -\dfrac{2P}{\pi} \dfrac{\cos^2 \theta \sin \theta}{r} \end{cases}$$

As can be seen in Fig. 4.1

$$\sin\theta = \frac{x}{\sqrt{x^2 + z^2}},\ \cos\theta = \frac{z}{\sqrt{x^2 + z^2}} \text{ and } r = \sqrt{x^2 + z^2}$$

So

$$\begin{cases} \sigma_x = -\dfrac{2P}{\pi} \dfrac{x^2 z}{\left(x^2 + z^2\right)^2} \\ \sigma_z = -\dfrac{2P}{\pi} \dfrac{z^3}{\left(x^2 + z^2\right)^2} \\ \tau_{xz} = -\dfrac{2P}{\pi} \dfrac{z^2 x}{\left(x^2 + z^2\right)^2} \end{cases} \tag{4.6}$$

A concentrated tangential load Q is argued by Johnson (1985) to produce a radial stress similar to a concentrated normal load P but revolved by 90°.

$$\begin{cases} \sigma_r = -\dfrac{2Q}{\pi} \dfrac{\cos \theta}{r} \\ \sigma_\theta = 0 \\ \tau_{r\theta} = 0 \end{cases}$$

Since it was revolved by 90°, the directional components are

$$\begin{cases} \sigma_x = -\dfrac{2Q}{\pi}\dfrac{x^3}{\left(x^2+z^2\right)^2} \\ \sigma_z = -\dfrac{2Q}{\pi}\dfrac{z^2x}{\left(x^2+z^2\right)^2} \\ \tau_{xz} = -\dfrac{2Q}{\pi}\dfrac{x^2z}{\left(x^2+z^2\right)^2} \end{cases} \tag{4.7}$$

As presented by Johnson (1985), for example, a general surface load distribution can be approximated by a set of concentrated normal $p(s)$ and tangential $q(s)$ forces on an finite elemental area of width ds, as shown in Figure 4.2, therefore, equations (4.6) and (4.7) can be used to evaluate the stress components at any point of the half-space by replacing x by $(x–s)$ and integrating them over the loaded strip ($–b < x < a$).

$$\begin{cases} \sigma_x(x,z) = -\dfrac{2z}{\pi}\displaystyle\int_{-b}^{a}\frac{p(s)(x-s)^2}{\left((x-s)^2+z^2\right)^2}ds - \frac{2}{\pi}\int_{-b}^{a}\frac{q(s)(x-s)^3}{\left((x-s)^2+z^2\right)^2}ds \\ \sigma_z(x,z) = -\dfrac{2z^3}{\pi}\displaystyle\int_{-b}^{a}\frac{p(s)}{\left((x-s)^2+z^2\right)^2}ds - \frac{2z^2}{\pi}\int_{-b}^{a}\frac{q(s)(x-s)}{\left((x-s)^2+z^2\right)^2}ds \\ \tau_{xz}(x,z) = -\dfrac{2z^2}{\pi}\displaystyle\int_{-b}^{a}\frac{p(s)(x-s)}{\left((x-s)^2+z^2\right)^2}ds - \frac{2z}{\pi}\int_{-b}^{a}\frac{q(s)(x-s)^2}{\left((x-s)^2+z^2\right)^2}ds \end{cases} \tag{4.8}$$

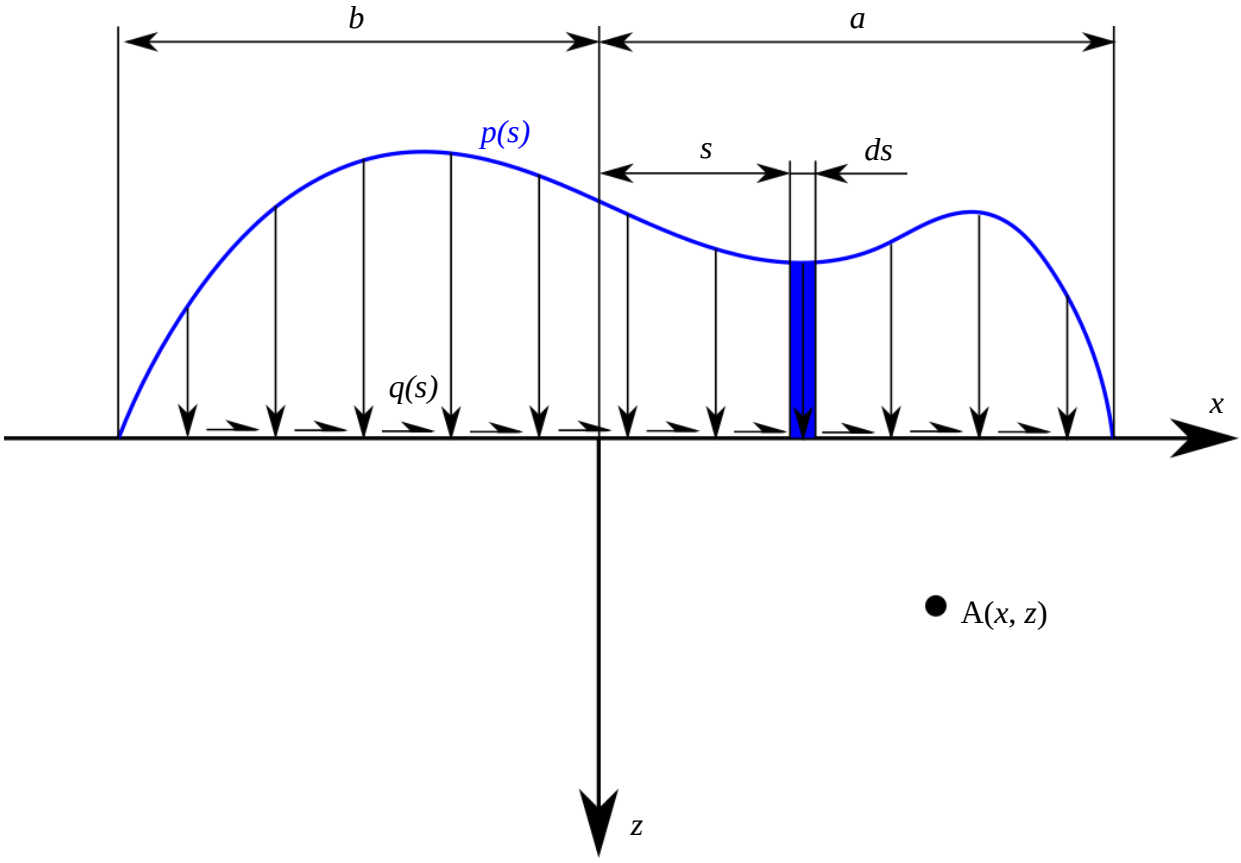


Figure 4.2 – For evaluation of stress components due to a load distributed over interval (–*b*, *a*)

The EHL analysis explained in Chapter 3 provides the tabulated normal and tangential tractions evaluated on the EHL mesh, as shown schematically in Figure 4.3. The load can be approximated to a set of loads uniformly distributed over elemental areas Δ. According to the superposition principle the total stress at any point of the half-space can be evaluated as a sum of stresses at that particular point due to loads over each elemental area Δ.

Consider the stress due to a constant pressure, p_c, and shear stress, q_c, applied to the portion $\left[-\frac{\Delta}{2}, \frac{\Delta}{2}\right]$ of the surface as in Figure 4.3.

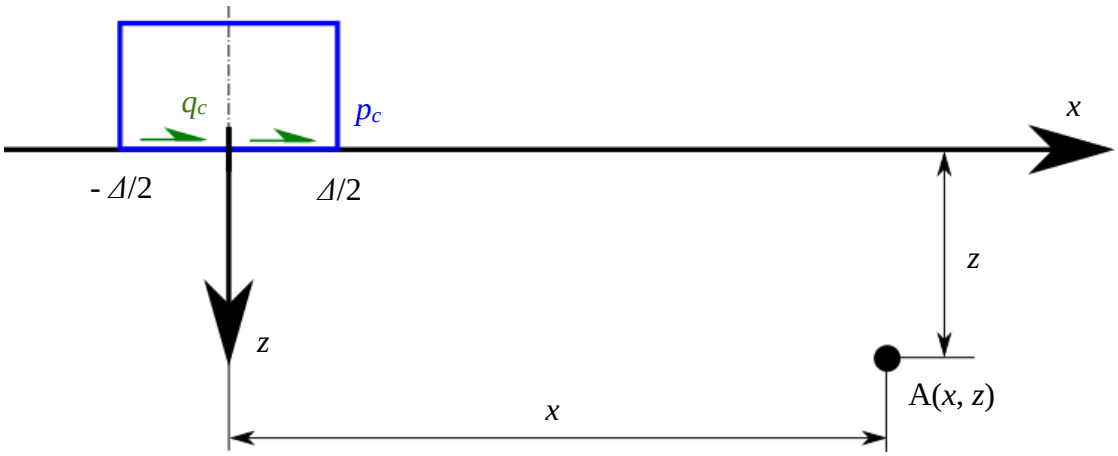


Figure 4.3 – Approximation of loading conditions

Then the resulting stress components at the general point A (x, z) are

$$\begin{cases} \sigma_x(x,z) = -\dfrac{2z}{\pi} p_c \displaystyle\int_{-\Delta/2}^{\Delta/2} \frac{(x-s)^2}{\left((x-s)^2+z^2\right)^2} ds - \frac{2}{\pi} q_c \int_{-\Delta/2}^{\Delta/2} \frac{(x-s)^3}{\left((x-s)^2+z^2\right)^2} ds \\ \sigma_z(x,z) = -\dfrac{2z^3}{\pi} p_c \displaystyle\int_{-\Delta/2}^{\Delta/2} \frac{1}{\left((x-s)^2+z^2\right)^2} ds - \frac{2z^2}{\pi} q_c \int_{-\Delta/2}^{\Delta/2} \frac{(x-s)}{\left((x-s)^2+z^2\right)^2} ds \\ \tau_{xz}(x,z) = -\dfrac{2z^2}{\pi} p_c \displaystyle\int_{-\Delta/2}^{\Delta/2} \frac{(x-s)}{\left((x-s)^2+z^2\right)^2} ds - \frac{2z}{\pi} q_c \int_{-\Delta/2}^{\Delta/2} \frac{(x-s)^2}{\left((x-s)^2+z^2\right)^2} ds \end{cases} \quad (4.9)$$

These expressions give the weighting functions for the element of surface loading at the origin. There are four types of integral to evaluate, which can be denoted c_1, c_1, c_3, c_4.

$$c_1(x,z) = -\frac{2z}{\pi} \int_{-\Delta/2}^{\Delta/2} \frac{(x-s)^2}{\left((x-s)^2+z^2\right)^2} ds \quad (4.10a)$$

$$c_2(x,z) = -\frac{2z^3}{\pi} \int_{-\Delta/2}^{\Delta/2} \frac{1}{\left((x-s)^2+z^2\right)^2} ds \quad (4.10b)$$

$$c_3(x,z) = -\frac{2z^2}{\pi}\int_{-\Delta/2}^{\Delta/2}\frac{(x-s)}{\left((x-s)^2+z^2\right)^2}\,ds \tag{4.10c}$$

$$c_4(x,z) = -\frac{2}{\pi}\int_{-\Delta/2}^{\Delta/2}\frac{(x-s)^3}{\left((x-s)^2+z^2\right)^2}\,ds \tag{4.10d}$$

The stress components at point (x, z) are given by

$$\begin{cases}\sigma_x(x,z) = p_c c_1(x,z) + q_c c_4(x,z)\\ \sigma_z(x,z) = p_c c_2(x,z) + q_c c_3(x,z)\\ \tau_{xz}(x,z) = p_c c_3(x,z) + q_c c_1(x,z)\end{cases} \tag{4.11}$$

By the principle of superimposition, if the load is composed of a series of such constant pressure, constant shear stress blocks, (p_0, q_0), (p_1, q_1), … , (p_{n-1}, q_{n-1}), as in Figure 4.4, the stress at general point A is given by

$$\begin{cases}\sigma_x(x,z) = \sum_{i=0}^{n-1} p_i \cdot c_1(x-x_i,z) + \sum_{i=0}^{n-1} q_i \cdot c_4(x-x_i,z)\\ \sigma_z(x,z) = \sum_{i=0}^{n-1} p_i \cdot c_2(x-x_i,z) + \sum_{i=0}^{n-1} q_i \cdot c_3(x-x_i,z)\\ \tau_{xz}(x,z) = \sum_{i=0}^{n-1} p_i \cdot c_3(x-x_i,z) + \sum_{i=0}^{n-1} q_i \cdot c_1(x-x_i,z)\end{cases} \tag{4.12}$$

where x_i is the coordinate at the centre of the i^{th} loading block.

To evaluate the weighting functions c_1, c_1, c_3 and c_4 it is convenient to let $x = i\Delta$, $z = \lambda\Delta$ and $s = \kappa\Delta$.

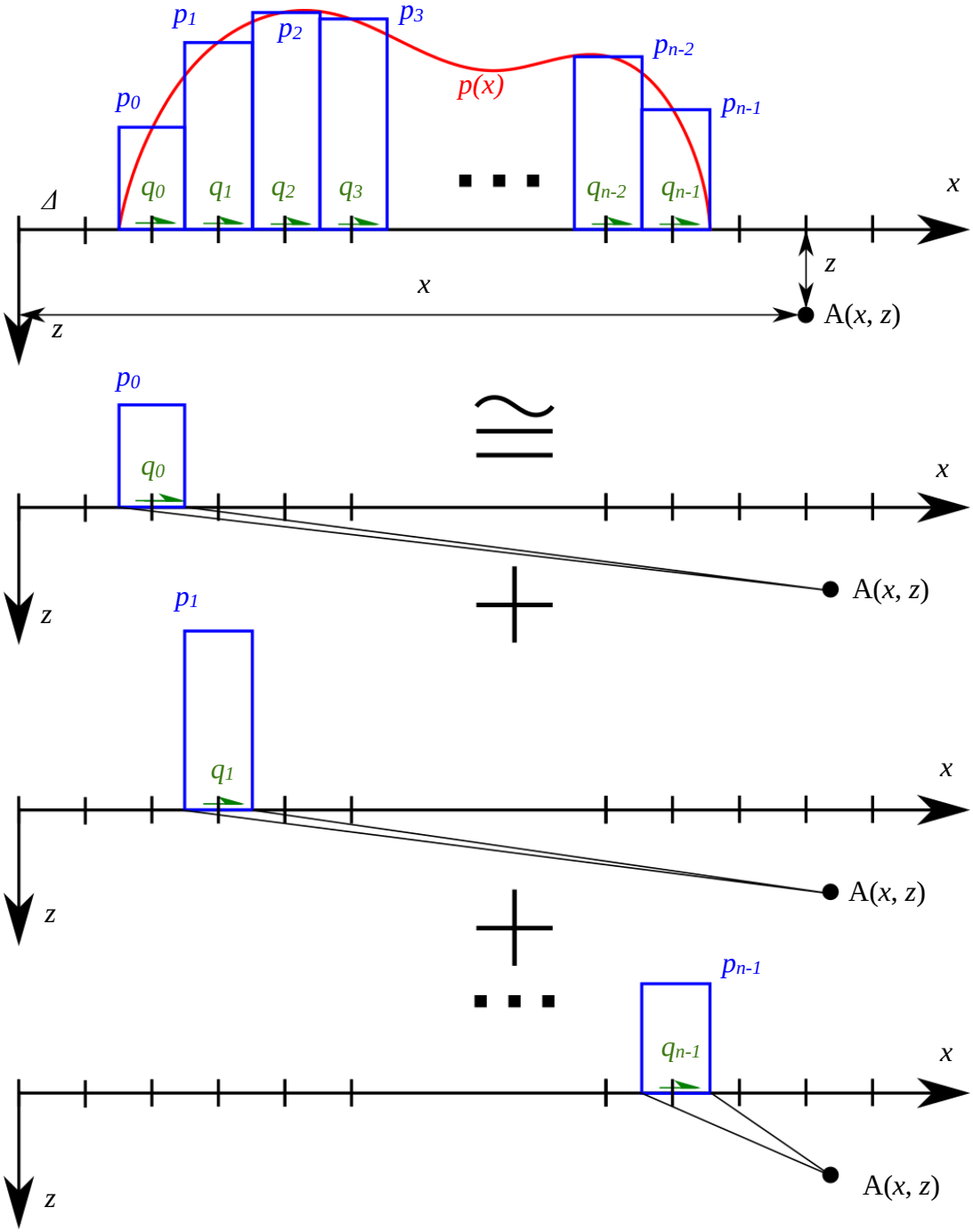


Figure 4.4 – Approximation of loading conditions

Therefore, the weighting functions are

$$c_1(i,\lambda) = -\frac{2\lambda}{\pi}\int_{-0.5}^{0.5}\frac{(i-\kappa)^2}{\left((i-\kappa)^2+\lambda^2\right)^2}\,d\kappa \tag{4.13a}$$

$$c_2(i,\lambda) = -\frac{2\lambda^3}{\pi}\int_{-0.5}^{0.5}\frac{1}{\left((i-\kappa)^2+\lambda^2\right)^2}\,d\kappa \tag{4.13b}$$

$$c_3(i,\lambda) = -\frac{2\lambda^2}{\pi}\int_{-0.5}^{0.5}\frac{(i-\kappa)}{\left((i-\kappa)^2+\lambda^2\right)^2}\,d\kappa \tag{4.13c}$$

$$c_4(i,\lambda) = -\frac{2}{\pi}\int_{-0.5}^{0.5}\frac{(i-\kappa)^3}{\left((i-\kappa)^2+\lambda^2\right)^2}\,d\kappa \tag{4.13d}$$

They can be evaluated by substitution $u = i - \kappa$, and then $d\kappa = -du$. The limits of integration are reversed due to negative sign of du. According to Brychinov, Marichev and Prudnikov (1989), the integrals of the functions (4.13) are

$$c_1(i,\lambda) = -\frac{2}{\pi}\int_{i-0.5}^{i+0.5}\frac{\lambda u^2}{\left(u^2+\lambda^2\right)^2}\,du = -\frac{1}{\pi}\left[-\frac{\lambda u}{u^2+\lambda^2}+\arctan\left(\frac{u}{\lambda}\right)\right]_{i-0.5}^{i+0.5} \tag{4.14a}$$

$$c_2(i,\lambda) = -\frac{2}{\pi}\int_{i-0.5}^{i+0.5}\frac{\lambda^3}{\left(u^2+\lambda^2\right)^2}\,du = -\frac{1}{\pi}\left[\frac{\lambda u}{u^2+\lambda^2}+\arctan\left(\frac{u}{\lambda}\right)\right]_{i-0.5}^{i+0.5} \tag{4.14b}$$

$$c_3(i,\lambda) = -\frac{2}{\pi}\int_{i-0.5}^{i+0.5}\frac{\lambda^2 u}{\left(u^2+\lambda^2\right)^2}\,du = \frac{1}{\pi}\left[\frac{\lambda^2}{u^2+\lambda^2}\right]_{i-0.5}^{i+0.5} \tag{4.14c}$$

$$c_4(i,\lambda) = -\frac{2}{\pi}\int_{i-0.5}^{i+0.5}\frac{u^3}{\left(u^2+\lambda^2\right)^2}\,du = -\frac{1}{\pi}\left[\frac{\lambda^2}{u^2+\lambda^2}+\ln\left(u^2+\lambda^2\right)\right]_{i-0.5}^{i+0.5} \tag{4.14d}$$

As can be seen in Figure 4.5

$$\chi_1^2 = (i-0.5)^2+\lambda^2 \qquad \chi_2^2 = (i+0.5)^2+\lambda^2$$

$$\frac{\pi}{2}-\theta_1 = \arctan\left(\frac{(i-0.5)}{\lambda}\right) \qquad \frac{\pi}{2}-\theta_2 = \arctan\left(\frac{(i+0.5)}{\lambda}\right)$$

$$\theta_1 = \arctan\left(\frac{\lambda}{i-0.5}\right) \qquad \theta_2 = \arctan\left(\frac{\lambda}{i+0.5}\right) \tag{4.15}$$

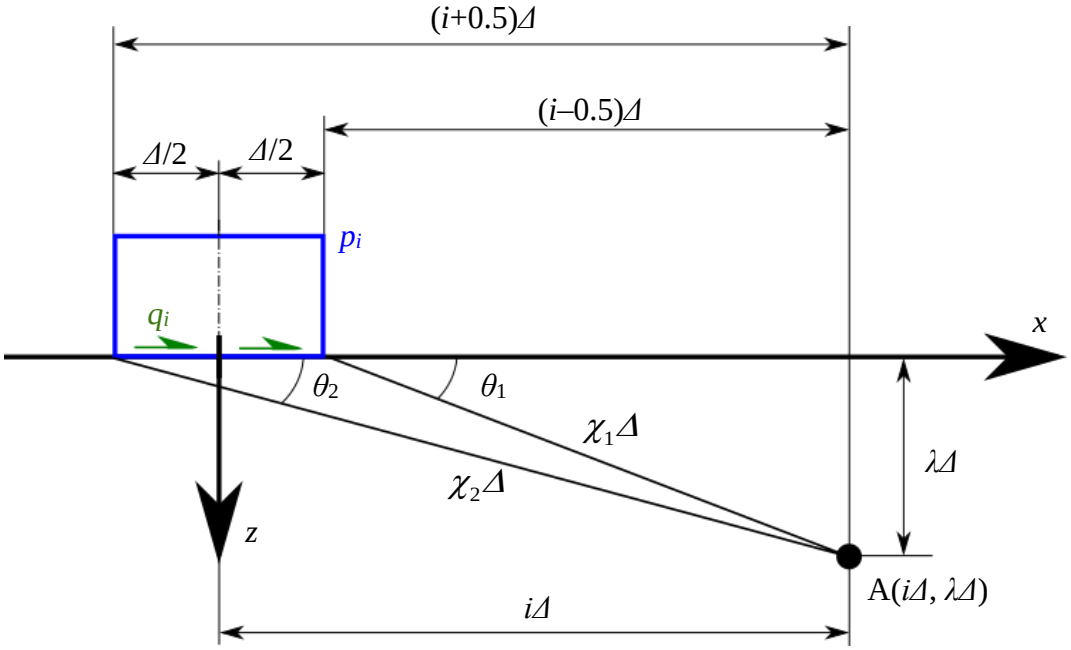


Figure 4.5 – Approximation of loading conditions

Applying limits of integration and substituting equations (4.15)

$$c_1(i,\lambda) = -\frac{1}{\pi}\left[-\frac{\lambda(i+0.5)}{(i+0.5)^2+\lambda^2} + \frac{\lambda(i-0.5)}{(i-0.5)^2+\lambda^2} + \arctan\left(\frac{(i+0.5)}{\lambda}\right) - \arctan\left(\frac{(i-0.5)}{\lambda}\right)\right]$$

$$= -\frac{1}{\pi}\left[\frac{\lambda\left(i^2-\lambda^2-0.25\right)}{\chi_1^2\chi_2^2} + \theta_1 - \theta_2\right] \qquad (4.16a)$$

$$c_2(i,\lambda) = -\frac{1}{\pi}\left[\frac{\lambda(i+0.5)}{(i+0.5)^2+\lambda^2} - \frac{\lambda(i-0.5)}{(i-0.5)^2+\lambda^2} + \arctan\left(\frac{(i+0.5)}{\lambda}\right) - \arctan\left(\frac{(i-0.5)}{\lambda}\right)\right]$$

$$= -\frac{1}{\pi}\left[\frac{\lambda\left(\lambda^2-i^2+0.25\right)}{\chi_1^2\chi_2^2} + \theta_1 - \theta_2\right] \qquad (4.16b)$$

$$c_3(i,\lambda) = \frac{\lambda^2}{\pi}\left[\frac{1}{(i+0.5)^2+\lambda^2} - \frac{1}{(i-0.5)^2+\lambda^2}\right] = -\frac{\lambda^2}{\pi}\left(\frac{1}{\chi_1^2} - \frac{1}{\chi_2^2}\right) \qquad (4.16c)$$

$$c_4(i,\lambda) = -\frac{1}{\pi}\left[\frac{\lambda^2}{(i+0.5)^2+\lambda^2} - \frac{\lambda^2}{(i-0.5)^2+\lambda^2} + \ln\left((i+0.5)^2+\lambda^2\right) - \ln\left((i-0.5)^2+\lambda^2\right)\right]$$

$$= -\frac{1}{\pi}\left[-\lambda^2\left(\frac{1}{\chi_1^2} - \frac{1}{\chi_2^2}\right) + \ln\left(\frac{\chi_1^2}{\chi_2^2}\right)\right] = -\frac{1}{\pi}\left[-\frac{c_3(i,\lambda)}{2} + \ln\left(\frac{\chi_1^2}{\chi_2^2}\right)\right] \qquad (4.16d)$$

The stress components defined in the equation (4.12) can be evaluated by applying a Discrete Convolution and Fast Fourier Transform (DC-FFT) method for contact analysis, developed by Liu, Wang and Liu (2000). The algorithm steps include the following:

Notation: a vector $\{v(i)\}_N$ has N entries and its index , i, varies between 0 and N-1.

1. Calculate the influence coefficients (4.16), $\{c_1(i,\lambda)\}_N$, $\{c_2(i,\lambda)\}_N$, $\{c_3(i,\lambda)\}_N$, $\{c_4(i,\lambda)\}_N$. Note that the influence coefficients are a function of the position at which the stress components are being evaluated, therefore they are constant for the static mesh.
2. The EHL normal and tangential load domain is from 0 through N–1, where surface tractions, $\{p(i)\}_N$ and $\{q(i)\}_N$ are applied. The load vectors must be extended from N to 2N, by zero padding over $i \in [N, 2N\text{-}1]$ and the influence coefficients must be mirrored about index N, where the value must be set to 0, in order to convert linear convolution into the cyclic one (Press *et al. 1992*). Note that c_1 and c_2 are even functions and c_3 and c_4 are odd.
3. Apply FFT of discretely sampled data (Press *et al. 1992*) to $\{c_1(i,\lambda)\}_{2N}$, $\{c_2(i,\lambda)\}_{2N}$, $\{c_3(i,\lambda)\}_{2N}$, $\{c_4(i,\lambda)\}_{2N}$, $\{p(i)\}_{2N}$, $\{q(i)\}_{2N}$ to obtain $\{C_1(\omega,\lambda)\}_{2N}$, $\{C_2(\omega,\lambda)\}_{2N}$, $\{C_3(\omega,\lambda)\}_{2N}$, $\{C_4(\omega,\lambda)\}_{2N}$, $\{P(\omega)\}_{2N}$, $\{Q(\omega)\}_{2N}$
4. Calculate a temporal frequency series $\{\Omega_{xx}(\omega,\lambda)\}_{2N}$, $\{\Omega_{zz}(\omega,\lambda)\}_{2N}$ and $\{T_{xz}(\omega,\lambda)\}_{2N}$ by the element-by-element multiplication of complex numbers based on the equations (4.11)

$$\begin{cases} \Omega_{xx}(\omega,z) = P(\omega)\cdot C_1(\omega,z) + Q(\omega)\cdot C_4(\omega,z) \\ \Omega_{zz}(\omega,z) = P(\omega)\cdot C_2(\omega,z) + Q(\omega)\cdot C_3(\omega,z) \\ T_{xz}(\omega,z) = P(\omega)\cdot C_3(\omega,z) + Q(\omega)\cdot C_1(\omega,z) \end{cases} \tag{4.17}$$

5. Apply IFFT to the temporal frequencies $\{\Omega_{xx}(\omega,\lambda)\}_{2N}$, $\{\Omega_{zz}(\omega,\lambda)\}_{2N}$ and $\{T_{xz}(\omega,\lambda)\}_{2N}$
6. Discard the spoiled terms $i \in [N, 2N\text{-}1]$ and divide the rest of them by 2π to obtain stress components σ_x, σ_z and τ_{xz}

4.3. Verifying the accuracy of the elastic stress calculations

To check the accuracy of the DC-FFT method, the two-dimensional contact of cylindrical bodies, as in Figure 3.1, was selected due to availability of analytical solutions. McEwen (1949) expressed the stress components at a general point (x,z)

$$\begin{cases} \sigma_x(x,z) = -\dfrac{p_0}{a}\left(m\left(1+\dfrac{z^2+n^2}{m^2+n^2}\right) - 2z \right) \\ \sigma_z(x,z) = -\dfrac{p_0 m}{a}\left(1-\dfrac{z^2+n^2}{m^2+n^2}\right) \\ \tau_{xz}(x,z) = -\dfrac{p_0 n}{a}\left(1-\dfrac{m^2-z^2}{m^2+n^2}\right) \end{cases} \tag{4.18}$$

where

$$\begin{aligned} m^2 &= \frac{1}{2}\left[\sqrt{\left(a^2-x^2+z^2\right)^2+4x^2z^2}+\left(a^2-x^2+z^2\right)\right] \\ n^2 &= \frac{1}{2}\left[\sqrt{\left(a^2-x^2+z^2\right)^2+4x^2z^2}-\left(a^2-x^2+z^2\right)\right] \end{aligned} \tag{4.19}$$

The signs of n and m are the same as the signs of x and z respectively.

The results of DC-FFT calculations are presented as two-dimensional filled contour plots in Figures 4.6 through 4.8 with the results of the analytical solution (4.17) superimposed on top as solid black iso-lines. The example illustrated has p_0 = 0.8 MPa.

The numerical errors plotted in Figure 4.9 are of the same format

$$Error \quad s = \frac{|s-S|}{|s|}\cdot 100\ \% \tag{4.20}$$

where s is appropriate stress component evaluated by using analytical equations (4.18) and S is the same stress component calculated by application of the DC-FFT method.

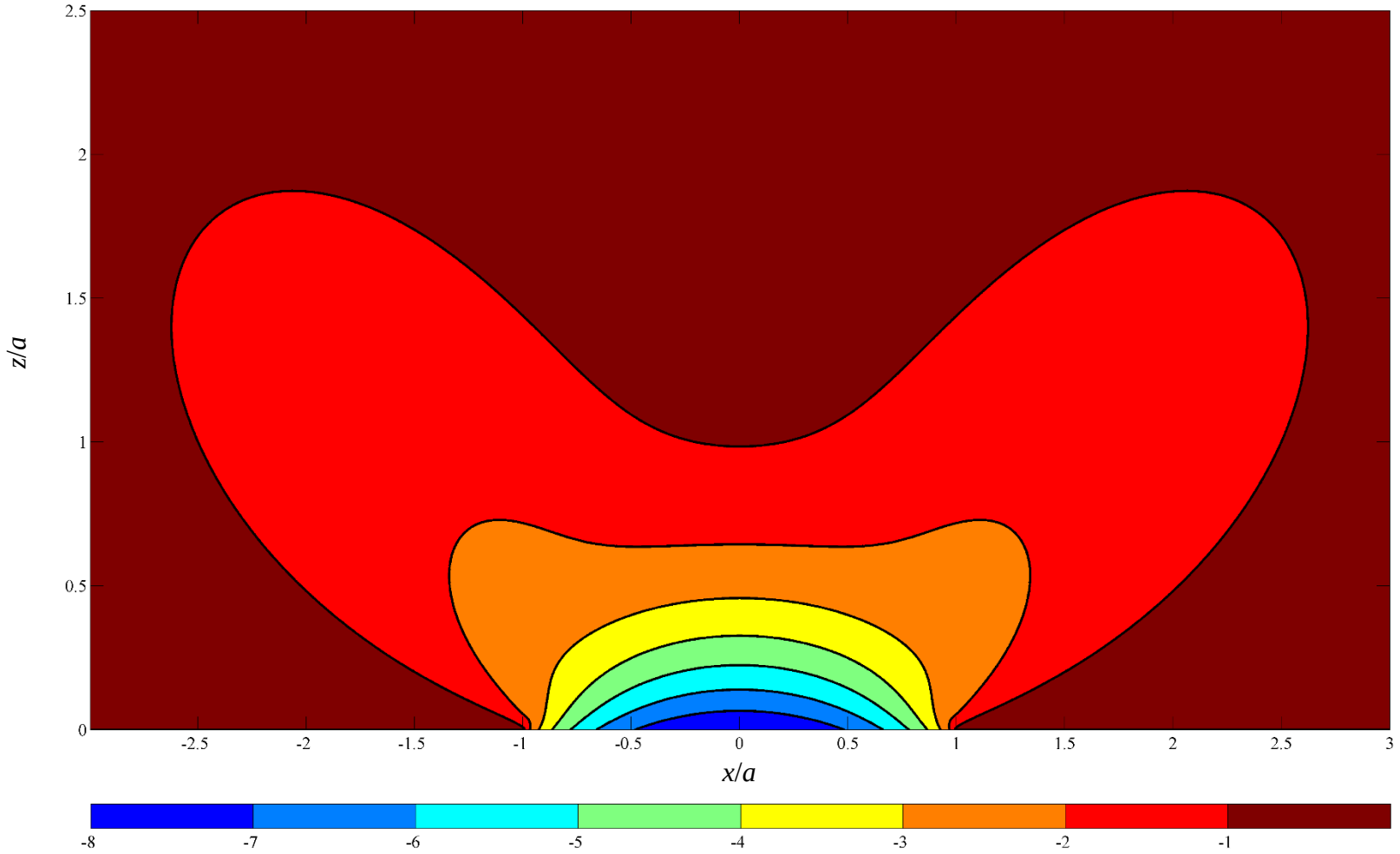


Figure 4.6 – Contact of cylinders: horizontal stress component, contour plot of σ_x / 10^5 Pa

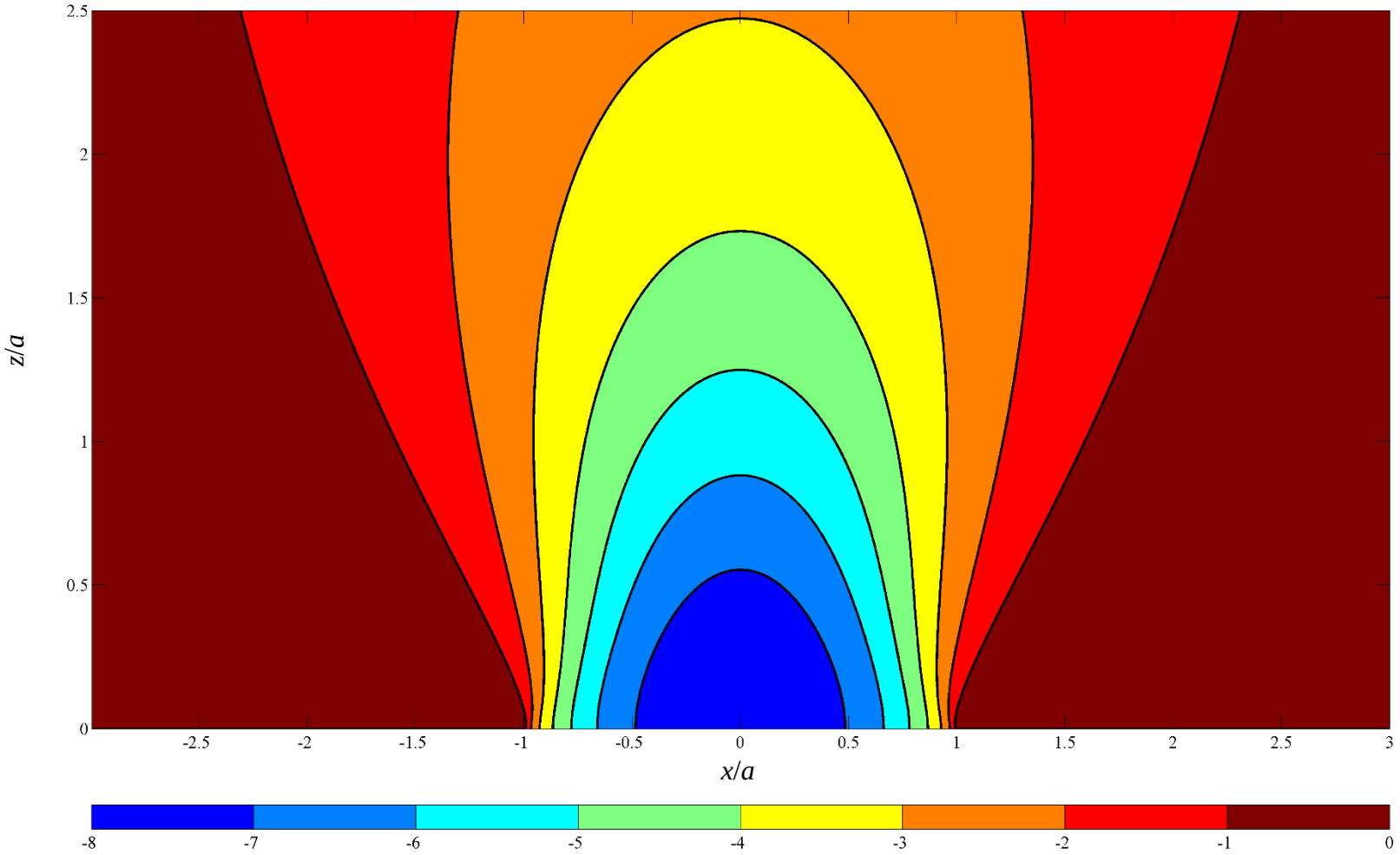


Figure 4.7 – Contact of cylinders: vertical stress component, contour plot of σ_z / 10^5 Pa

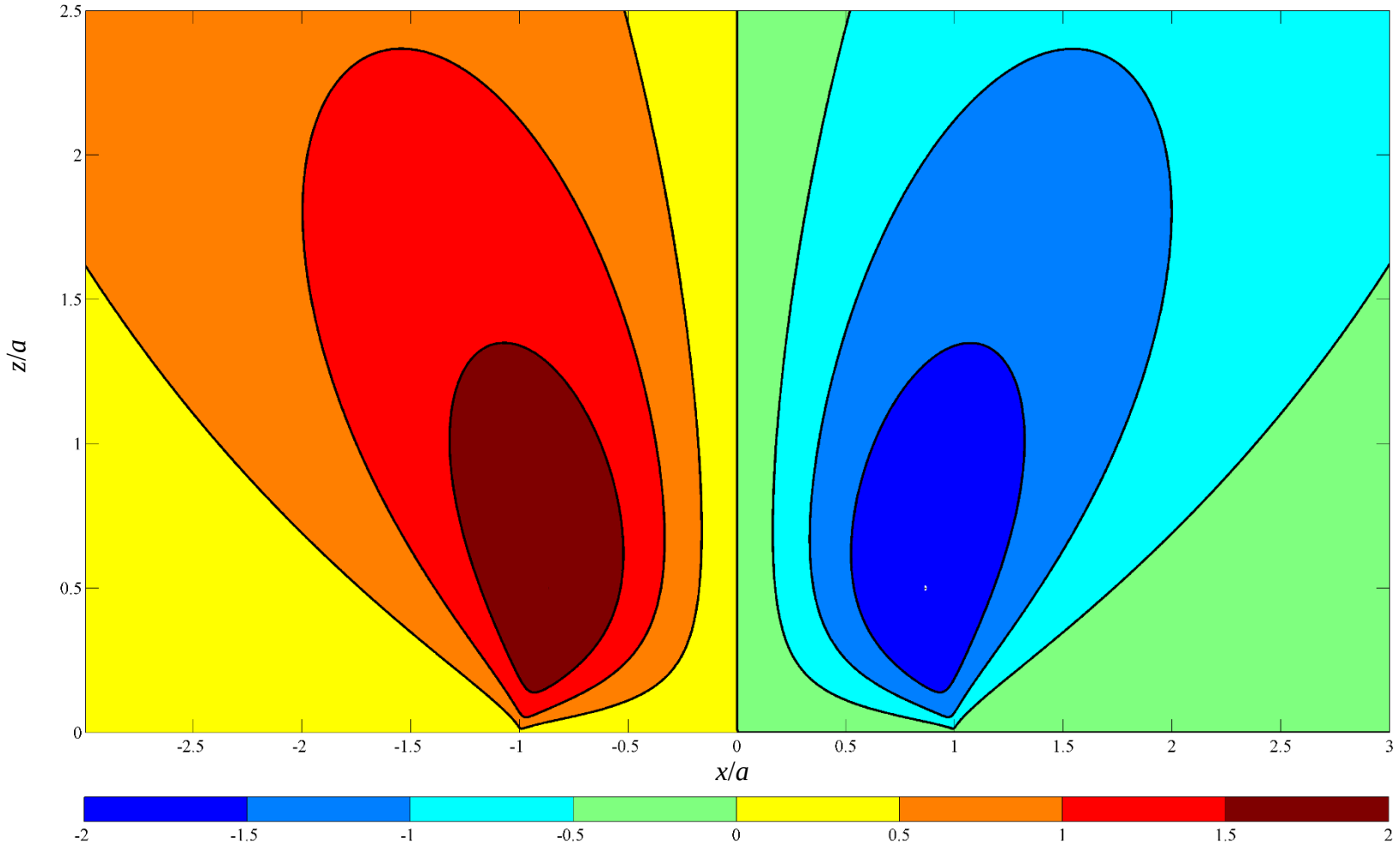


Figure 4.8– Contact of cylinders: shear stress component, contour plot of τ_{xz} / 10^5 Pa

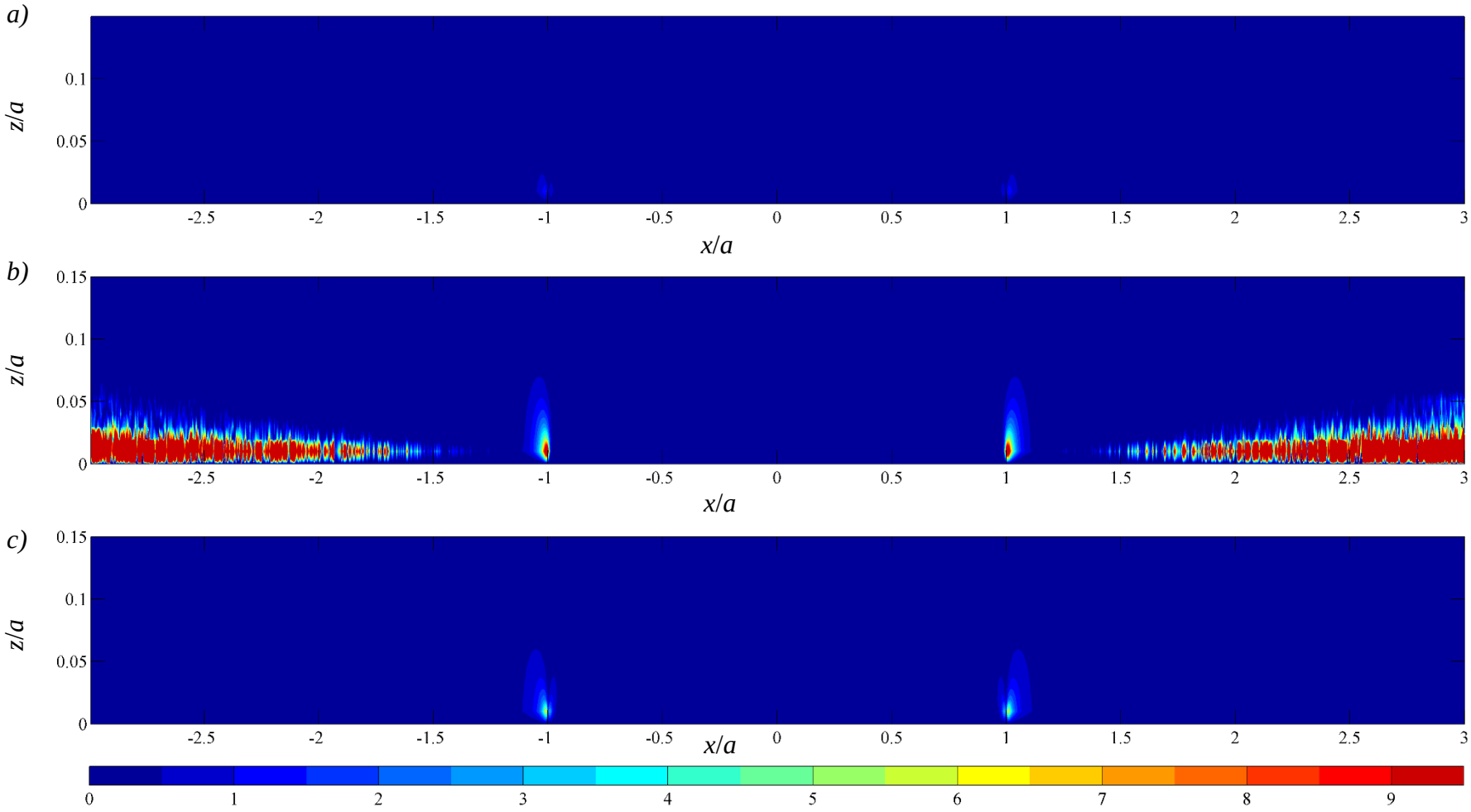


Figure 4.9 – Numerical error (%): (a) σ_x , (b) σ_z, (c) τ_{xz}

The numerical error does not exceed 0.5% for the vast majority of the mesh nodes with no more than 1.4% and 5% difference at the zones characterized by a high pressure gradient and pressure gradient discontinuity at $x = \pm a$ for σ_x and τ_{xz} respectively. The numerical error in evaluating σ_y at the inlet and outlet zones at the close proximity of the surface can reach up to 200% for the length up to three half-Hetzian contact widths and rises rapidly with the increase of distance from the centre of the contact. This is because of the cyclic loading that is actually being applied to the surface and can be reduced by adding zero padding. However, the values of stress components at those zones do not exceed 0.001% of p_0, therefore, these differences are insignificant for evaluating the fatigue damage.

From equations 4.6 and 4.7 it can be seen that stress components σ_x and τ_{xz} in the case of load directed normally to the surface are of the same form as τ_{xz} and σ_z in case of load directed tangentially to the surface respectively. Therefore, if the tangential traction is set in the form of the Hertzian stress distribution

$$q(x) = \begin{cases} 0, & x \in [-\infty; -a] \cup [a; +\infty] \\ q_0 \sqrt{1 - \dfrac{x^2}{a^2}}, & x \in (-a; a) \end{cases} \tag{4.21}$$

the stress components σ_z and τ_{xz} can be evaluated by applying McEwen's (1949) solution

$$\begin{cases} \sigma_z(x,z) = -\dfrac{p_0 n}{a}\left(1 - \dfrac{m^2 - z^2}{m^2 + n^2}\right) \\ \tau_{xz}(x,z) = -\dfrac{p_0}{a}\left(m\left(1 + \dfrac{z^2 + n^2}{m^2 + n^2}\right) - 2z\right) \end{cases} \tag{4.22}$$

The form of σ_x component in equation 4.7 is significantly different to all of the equations 4.6, so that McEwen's (1949) solution cannot be applied. Nevertheless, from equation (4.8) if there is no normal load $\sigma_x = (q \otimes c)(x)$, where $c(x,z)$ is a weighting function, which is defined for constant load across finite element by equation (4.16d) or, according to Johnson (1985)

$$c\ (x,z) = \frac{4\ln\left(\frac{r_1}{r_2}\right) - \cos\ 2\theta_1 + \cos\ 2\theta_2}{2\pi} \tag{4.23}$$

where

$$r_{1,2} = \sqrt{\left(x \mp \frac{\Delta}{2}\right)^2 + z^2} \qquad \arctan\ \theta_{1,2} = \frac{z}{x \mp \frac{\Delta}{2}}$$

Hence, σ_x can be approximated by the summation of the products of the load by weighting function at all mesh nodes.

$$\sigma_x(x,z) \cong \sum_{i=0}^{N}\sum_{j=0}^{N} q_i \cdot c(x_j - i, z) \tag{4.24}$$

where N is the number of mesh intervals.

The stress components due to the tangential load defined by equation (4.21) are plotted in Figures 4.10 through 4.12. The two-dimensional filled contours present the results of the DC-FFT calculations and the superimposed iso-lines show the results of the analytical solution (4.22) for σ_z and τ_{xz} and the numerical integration (4.24) using Johnson's (1985) weighting function (4.23) for σ_x. The numerical errors in Figure 4.13 are calculated by using equation (4.20).

The numerical error can be only evaluated for σ_z and τ_{xz} due to unavailability of an analytical solution for σ_x. It is under 0.5% for most mesh nodes reaching 3.8% and 1.4% for σ_z and τ_{xz} respectively at the nodes characterized by a rapid change of tangential traction.

In the case of σ_x, the results of two numerical calculations are compared, therefore, both result sets incorporate a numerical error. However, the difference of the results does not exceed 0.5% for stresses above 20% of maximum stress values, increasing up to 4% for stresses 10-20% of maximum stress values and reaching 9% at the zone characterized by the low stress values and the change of the sign..

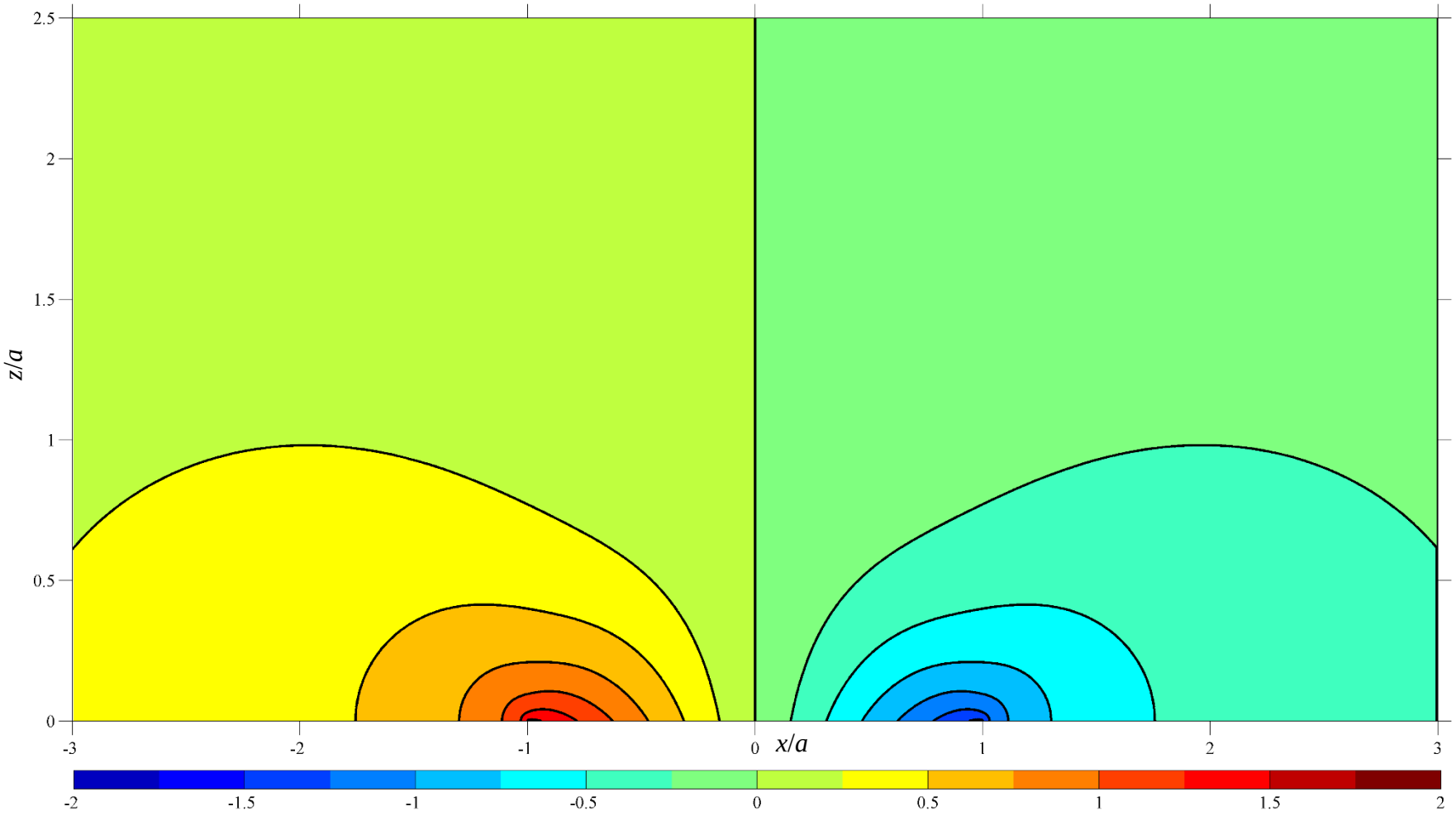


Figure 4.10 – Hertzian distribution of the tangential traction: horizontal stress component, contour plot of σ_x / 10^6 Pa

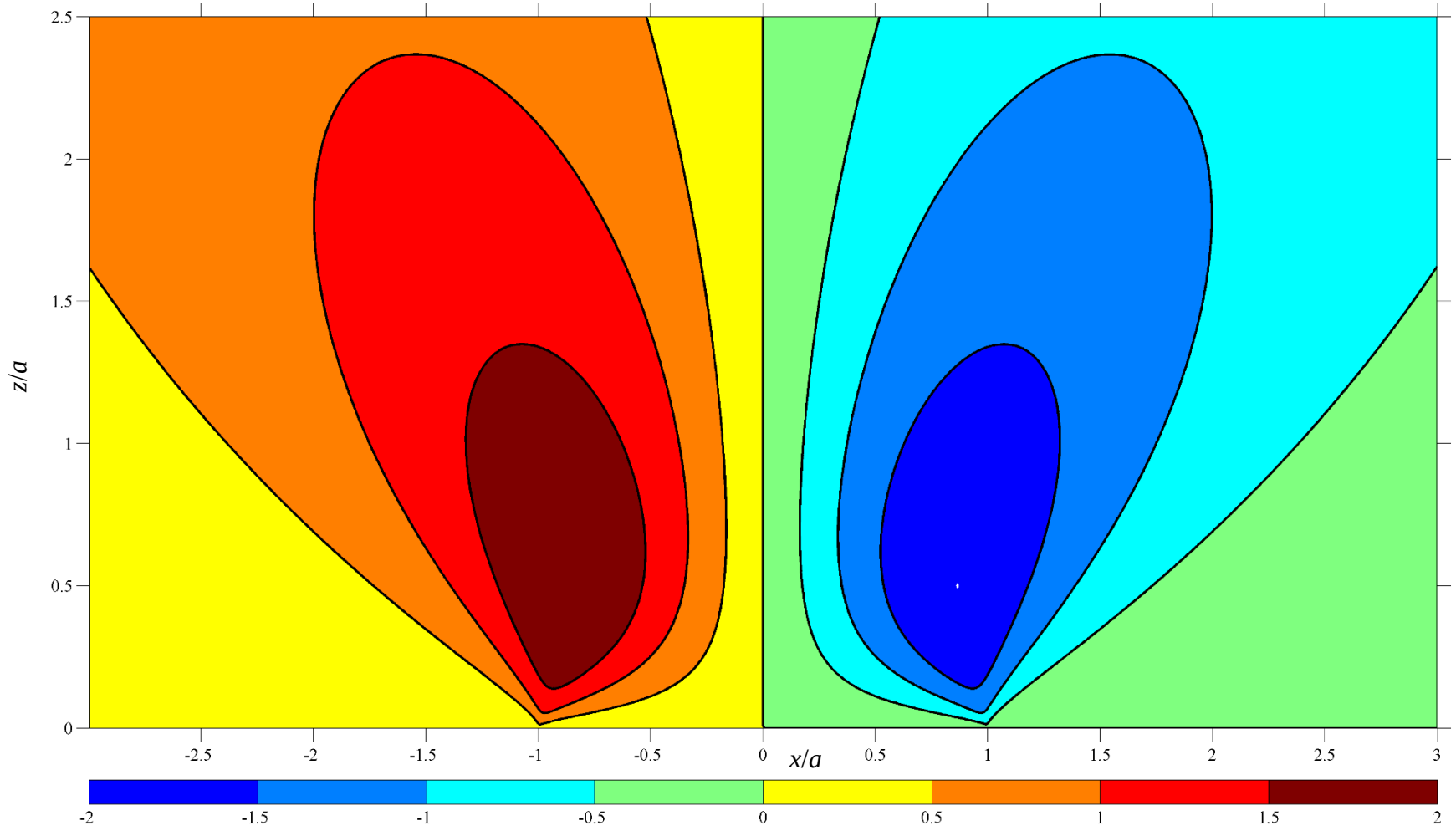


Figure 4.11 – Hertzian distribution of the tangential traction: vertical stress component, contour plot of σ_z / 10^6 Pa

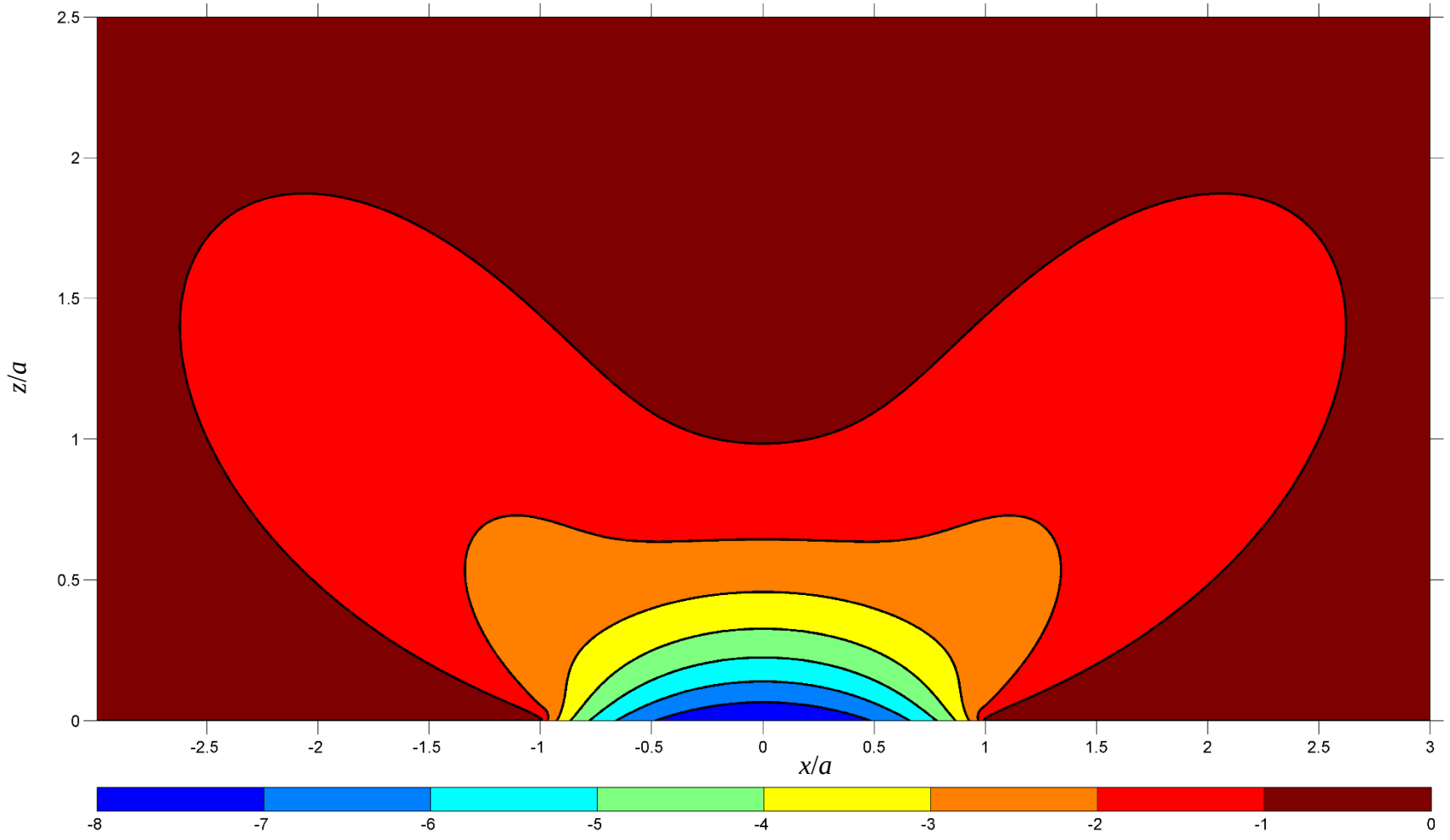


Figure 4.12 – Hertzian distribution of the tangential traction: shear stress, contour plot of $\tau_{xz} / 10^5$ Pa

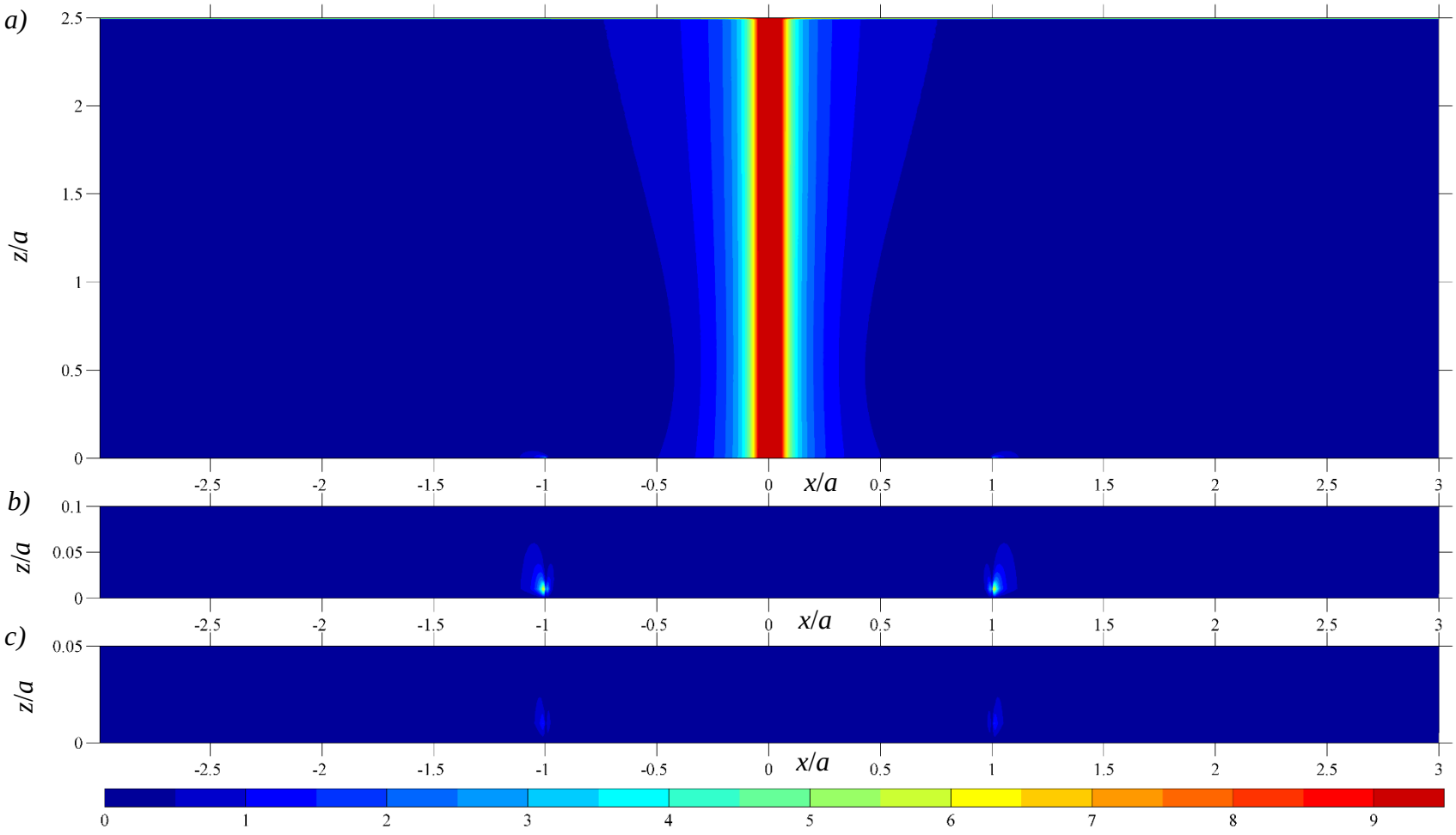


Figure 4.13 – Numerical error (%): (a) σ_x , (b) σ_z, (c) τ_{xz}

Unfortunately, there is no analytical solution for more complex problems, so, it is not possible to evaluate the numerical error, but it is feasible to check the stress patterns along with the pressure and tangential traction distributions with rough surfaces arrangement as shown in Figures 4.14 through 4.16. The normal stresses are primarily driven by normal load, therefore, the σ_x and σ_z evaluated in the EHL mesh were plotted against the EHL pressure distribution shown in blue. It can be seen that peaks of pressure correspond to high stress values at the surface, which gradually reduce further down from the surface and the lower pressure zones cause lower stress level. The magnitude of the stress at the surface agrees with the magnitude of the pressure distribution and the stress pattern can be described as the Hertzian stress distribution as in Figures 4.6 and 4.7 with severe numerical noise at the surface and near subsurface due to the hydrodynamic pressure caused by the contact kinematics and the incorporated roughness. The gap between the surfaces is plotted in green.

The shear stress τ_{xz} at the surface and near subsurface is heavily influenced by the tangential traction but its impact gradually reduces with the increase of the depth where the shear stress is dominated by the normal load. In Figure 4.16 the tangential load is plotted in blue and its pattern shows a good agreement with the shear stress pattern at the surface and near subsurface and further down from the surface it tends to the Hertzian stress distribution as in Figure 4.8.

From the reasoning above, it can be concluded that the DC-FFT method provides a robust tool to evaluate stress components due to the normal and tangential load provided.

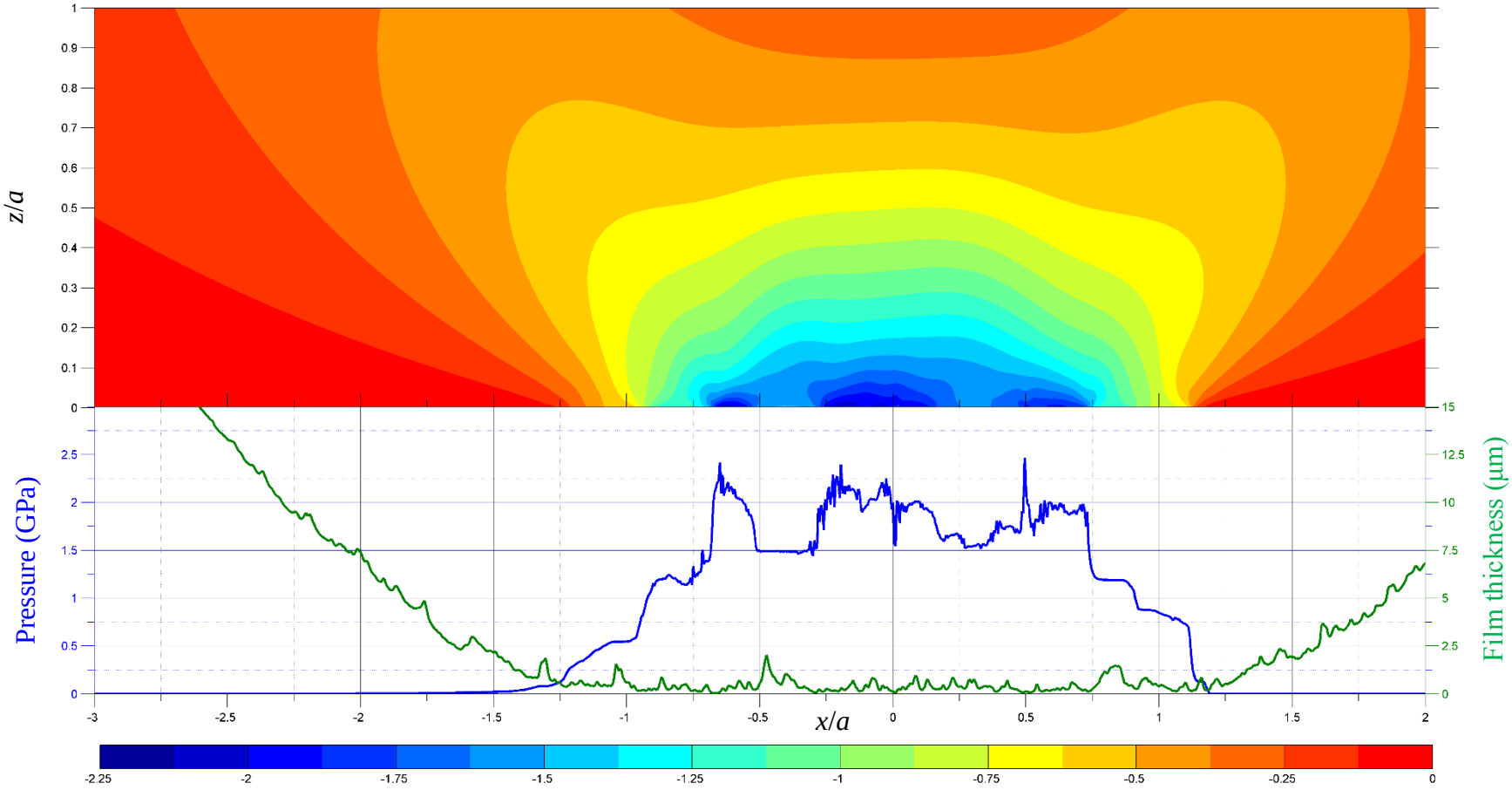


Figure 4.14 – Contact of rough surfaces: contour of horizontal stress component σ_x / GPa plotted against EHL pressure distribution and film thickness

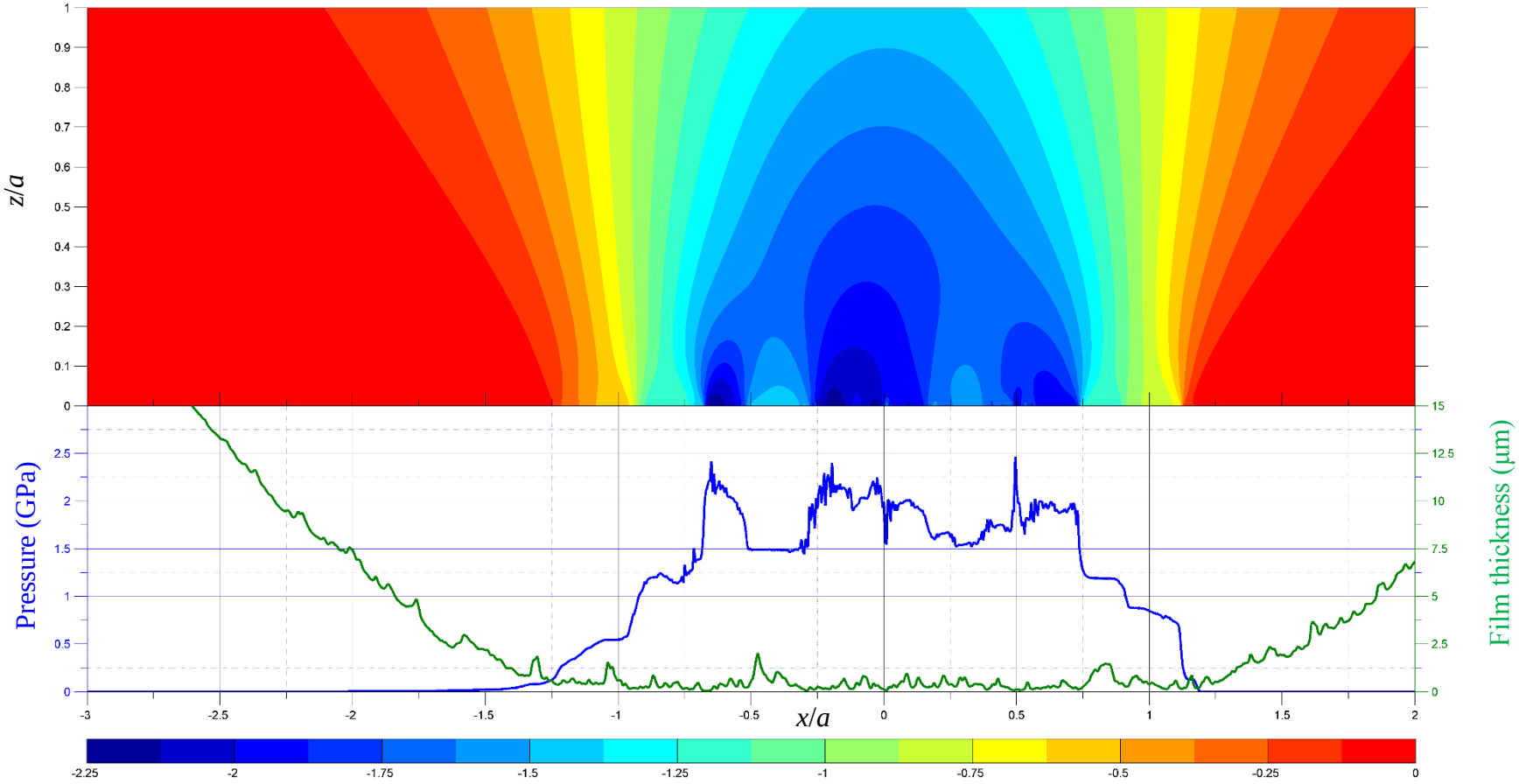


Figure 4.15 – Contact of rough surfaces: contour of horizontal stress component σ_z / GPa plotted against EHL pressure distribution and film thickness

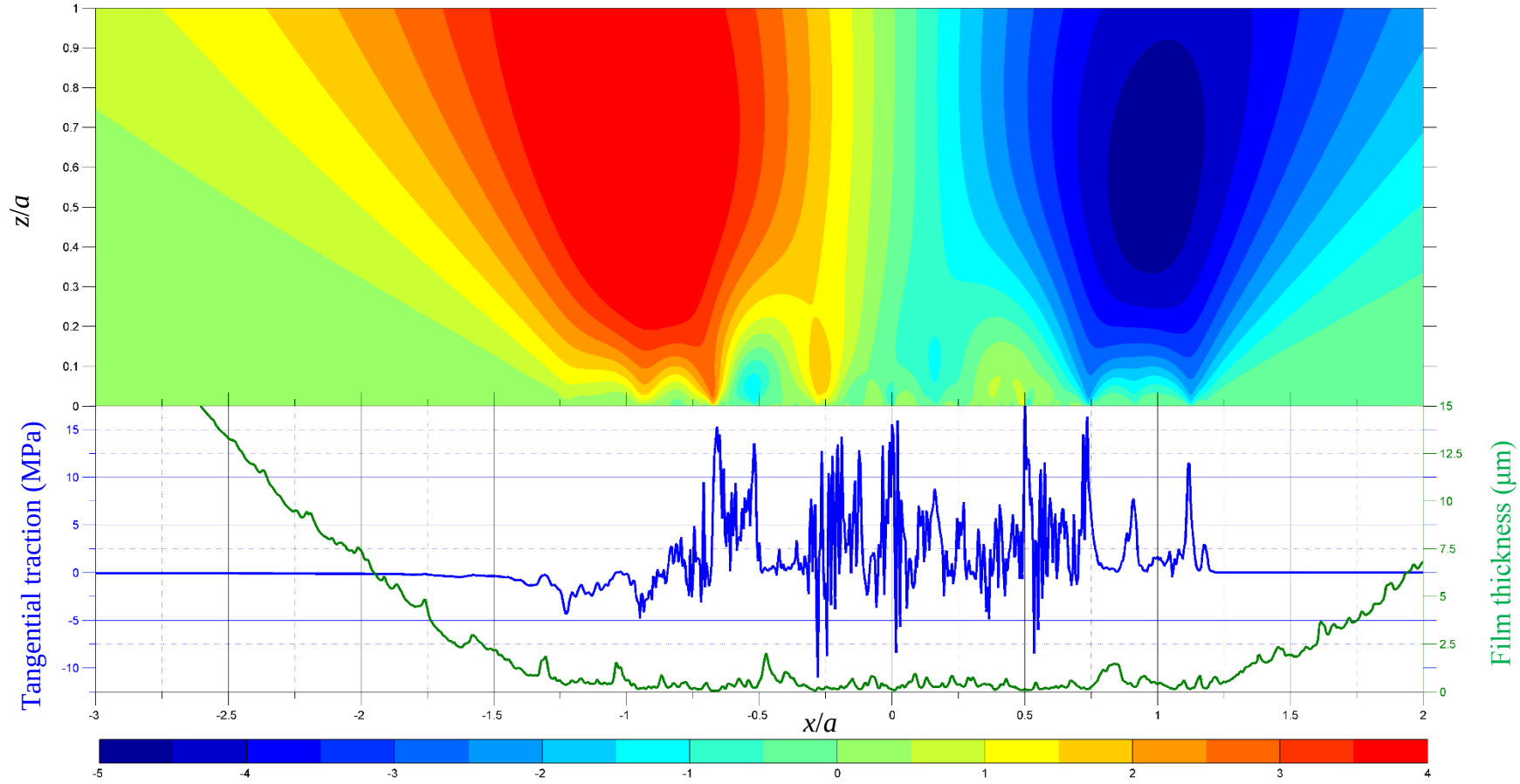


Figure 4.16 – Contact of rough surfaces: contour of horizontal stress component τ_{xz} / 10^8 Pa plotted against EHL tangential traction and film thickness

4.4. Storing stress history

The stress components in the previous section are evaluated in the same EHL mesh, shown in red in Figure 4.16, as for the pressure distribution and the film thickness. This is defined relative to the centre of the contact with its origin at the contact point. Consequently, the EHL mesh moves non-uniformly along the gear flank during the meshing cycle.

For the fatigue analysis the stress history must be collected for each mesh point of the gear flank. Therefore, the stress components must be interpolated to the gear flank mesh. The stress history mesh is set up at the first timestep for which the stress evaluation function is called. It is equally spaced in direction x as illustrated in Figure 4.16 along the gear flank surface. The spacing in the direction z, normal to the surface and directed into the material can be specified in a general way and it is convenient to adopt a non-uniform mesh to resolve rapid change in stress in the near surface material efficiently.

The fatigue analysis is based on the stress history at the point of interest and is not influenced by stresses at any other point, so, it is reasonable to store stress histories as a series of files. Each of these files contains the stress history of the nodes on a vertical line of the grid shown in grey in Figure 4.16 as it progresses through the EHL mesh associated with the centre of the contact zone. The position of the EHL mesh and corresponding flank material is shown for four timesteps in the gear meshing cycle. The file structure was developed to reduce the amount of data to be stored and system requirements for the hardware by excluding repeating patterns and storing them separately as well as sorting data in the most convenient manner to accommodate the fatigue analysis calculations. The file name is the six-digit number, which identifies the column of the material of the gear flank mesh.

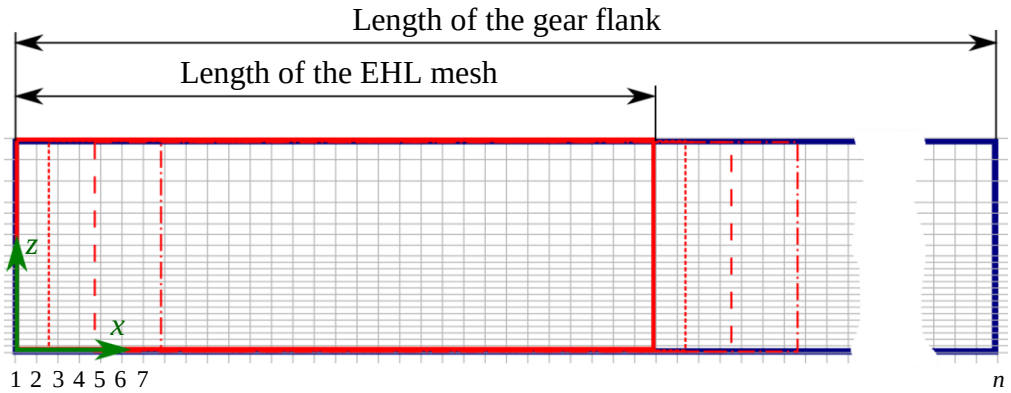


Figure 4.16 – Schematic illustration of the gear flank mesh with the location of the EHL mesh and corresponding flank material shown in red at different timesteps

The structure of these material column stress history files is specified in Table 4.1. The first line contains three numbers which are:

1. The number of stress evaluation nodes normal to the surface, *numZ*;
2. The spacing of the nodes parallel to the surface, Δx, in meters;
3. The number of nodes in dimension a_{max} parallel to the surface, *ninb*.

Those values must not vary within the dataset of the stress history. They are set by the first file of the dataset, then they are only used to ensure that the data is not corrupted.

The subsequent lines give the stress components, σ_{xx}, σ_{zz} and τ_{xz}, evaluated in Pa at the node spacing specified.

An example of the stress history for the column of the gear flank material is plotted at the surface and subsurface at four depths in Figure 4.17. The first timestep that the column entered the EHL mesh is defined as 0. It is approaching the contact area for about 1600 timesteps and leaving it after 2600 timesteps. The stress levels tend to zero asymptotically as the distance of the column from the contact zone increases. The highest values of stresses correspond to the timesteps spent in the contact zone as expected.

Table 4.1 – Structure of files holding stress history

	numZ	*Δx*	*ninb*
Timestep 1	σ_{xx}^{0}	σ_{zz}^{0}	τ_{xz}^{0}
	σ_{xx}^{1}	σ_{zz}^{1}	τ_{xz}^{1}
	…		
	$\sigma_{xx}^{\mathrm{numZ}}$	$\sigma_{zz}^{\mathrm{numZ}}$	$\tau_{xz}^{\mathrm{numZ}}$
Timestep 2	σ_{xx}^{0}	σ_{zz}^{0}	τ_{xz}^{0}
	σ_{xx}^{1}	σ_{zz}^{1}	τ_{xz}^{1}
	…		
	$\sigma_{xx}^{\mathrm{numZ}}$	$\sigma_{zz}^{\mathrm{numZ}}$	$\tau_{xz}^{\mathrm{numZ}}$
…			
Timestep k	σ_{xx}^{0}	σ_{zz}^{0}	τ_{xz}^{0}
	σ_{xx}^{1}	σ_{zz}^{1}	τ_{xz}^{1}
	…		
	$\sigma_{xx}^{\mathrm{numZ}}$	$\sigma_{zz}^{\mathrm{numZ}}$	$\tau_{xz}^{\mathrm{numZ}}$

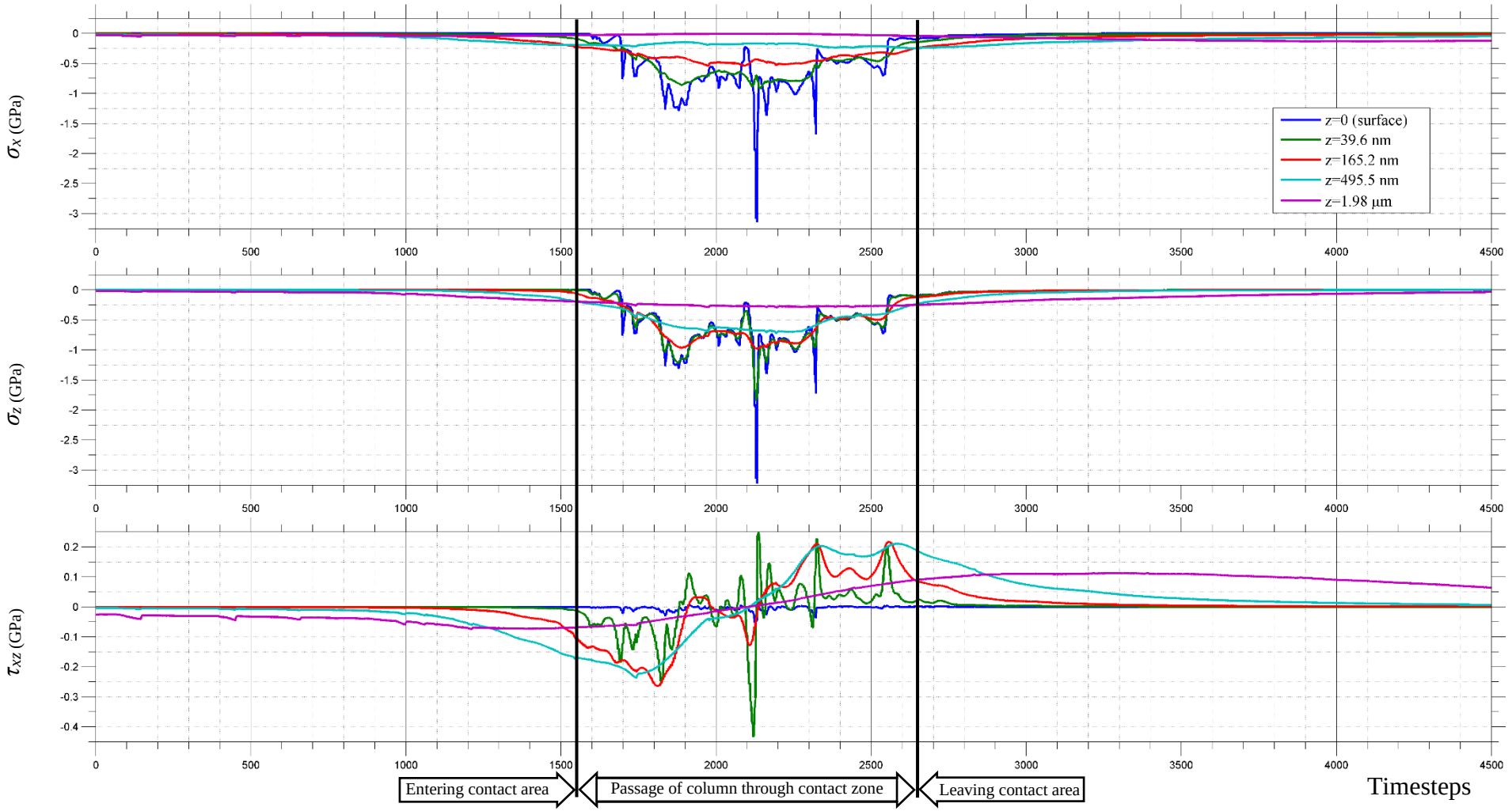


Figure 4.17 – Plotted examples of stress history

4.5. Fatigue of Materials: definition and brief history

The fatigue failure can be defined as the tendency of a material to fracture by means of progressive brittle cracking under repeated alternating or cyclic stresses of an intensity considerably below the normal strength. This definition excludes the phenomenon of static fatigue, which is used to describe stress corrosion cracking.

According to Schultz (1996), the history of fatigue begins with Albert (1837) who published the first fatigue test results known, investigating a failure of the conveyor chains used in the Clausthal mines. It started an international discussion across the UK, Germany and France mainly concentrating on transport accidents, involving trains and horse-drawn coaches. In 1860s and 1870s the topic was dominated by Wohler (1858, 1860, 1863, 1866, 1867, 1870, 1870a). In the UK, Fairbairn (1864) ran systematic fatigue testing. Ewing and Humfrey (1903) published the first metallurgical description of the fatigue process.

From the beginning of the WWI and until the end of the WWII the scientific interest in fatigue grew and researchers from all over the world became interested in the subject. The first books on fatigue were published by Gough (1927) in the UK, by Moore and Kommers (1927) in the USA, by Cazaud (1937) in France and by Serensen (1937) in Russia. The damage accumulation hypotheses for fatigue life prediction under variable loading amplitudes were published by Palmgren (1924), Langer (1937), Serensen (1940) and Miner (1945). But the most important development was a proposal by Basquin (1910) to represent the Wohler curves so as to define the finite life region in the form of a plot of stress against fatigue life on logarithmic axes, which is still in use and is often referred to as an S-N curve. Another important evolution of the fatigue theory of the time was popularisation of Palmgren's (1924) linear damage hypothesis by Miner (1945) which provided a practical method to predict a fatigue life taking accumulated damage into account.

After three fatal crashes of the world's first jetliner, the de Havilland DH 106 Comet, in 1953 and 1954 which caused 93 fatalities, the importance of including fatigue and stress concentration in the structural analysis became apparent. Since then the theory of fatigue has become a vital part of engineering and a variety of different fatigue related topics were investigated.

4.6. Fatigue of Materials: basic terms and nomenclature

The basic factors causing fatigue failure are a maximum tensile stress of sufficiently high value, a large amount of variation or fluctuation in the applied stress, and a sufficiently large number of cycles of the applied stress. Additional factors, such as stress concentration, corrosion, temperature, overload, metal structure, residual and combined stress, can expedite the failure.

The study of cyclic behaviour is based on the total-life approach, which includes stress-life and strain-life, and the damage-tolerant approach, which concentrates on the fracture mechanics and fatigue crack growth. The stress-life approach to fatigue was first introduced in 1860s by Wohler as a concept of an 'endurance limit', which specifies the applied stress amplitude below which a nominally defect-free material is expected to have an infinite fatigue life. This empirical method has found widespread use in fatigue analysis, mostly in applications where low-amplitude cyclic stress induce primarily elastic deformation in a component which is designed for long life, i.e. in the so called high-cycle fatigue (HCF) applications. When considerable plastic deformation occurs during cyclic loading as, for example, a consequence of high stress amplitude or stress concentrations, the fatigue life is significantly shortened. This calls for the so-called low-cycle fatigue (LCF) approach.

The EHL analysis presented in Chapter 3 does not include any procedure to accommodate plastic behaviour of the gear flanks as well as any indicators of the cracks on the surfaces. It narrows down the useful fatigue studies to the HCF approach.

As mentioned previously, the HCF models are empirical, i.e. they are based on statistically processed results of a series of tests. The test methods and the appropriate processing of the results is covered in detail in BS ISO 1143:2010, BS 3518-3:1963, BS ISO 12107:2003 or ASTM International Standards E3, E466-E468, E606, E739, E1012 and E1823. The combination of the results for several specimens are usually provided as a so-called S-N curve, a log-log plot of stresses, S, against the number of cycles to failure, N. Because the S-N fatigue data is normally scattered, an S-N curve represents the average probability of failure. The presentation of results is defined by BS 3518-1:1993 as well as symbols, terms and definitions presented in Figures 4.18-4.22 and Tables 4.2-4.3.

Table 4.2 – Symbols, terms and definitions relating to stress controlled testing

Symbol	Term	Definition
σ, τ, S	Stress	The force applied divided by the original cross-sectional area; tensile stress is considered positive and compressive stress negative
σ_{max} τ_{max}	Maximum stress	The highest algebraic value of stress in the stress cycle
σ_{min} τ_{min}	Minimum stress	The lowest algebraic value of stress in the stress cycle
σ_m τ_m	Mean stress	Half of the algebraic sum of the maximum and minimum stresses
σ_a τ_a	Stress amplitude	Half of the algebraic difference between the maximum and minimum stresses
$\Delta\sigma$ $\Delta\tau$	Range of stresses	The algebraic difference between the maximum and minimum stresses
R	Stress ratio	The algebraic ratio of the minimum stress to the maximum stress in one cycle
n	Number of stress cycles	The number of cycles applied
f	Frequency of cycles	The number of cycles applied per second
N or N_f	Endurance or fatigue life	The number of stress cycles to failure. NOTE This is generally stated as decimal fractions or multiples of 10^6
σ_N τ_N	Fatigue strength at N cycles	The value of the stress amplitude at a stated stress ratio under which the test piece would have a life of at least N cycles with a stated probability. NOTE If no probability is stated 50 % is implied. If no stress ratio is stated a value of – 1 is implied
σ_D τ_D	Fatigue limit	The value of the stress amplitude below which the test piece would be expected to endure an infinite number of stress cycles with a stated probability NOTE Certain materials do not show a fatigue limit. Others only show a fatigue limit in certain environments
σ_{af}	Fatigue limits	For fully reversed bending test
τ_{af}		For fully reversed pure torsion test
σ_u	Ultimate tensile strength	

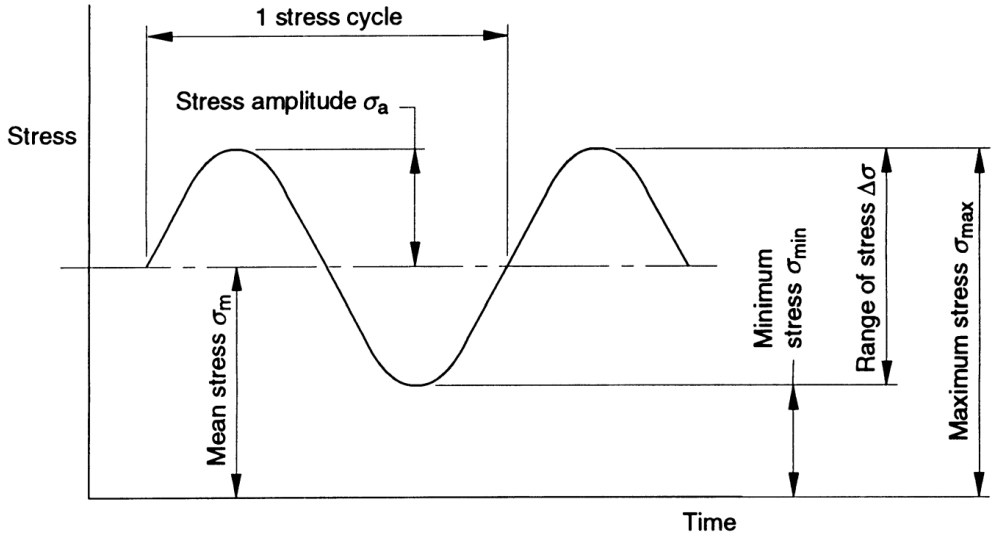


Figure 4.18 – Fatigue stress cycle (BS 3518-1:1993)

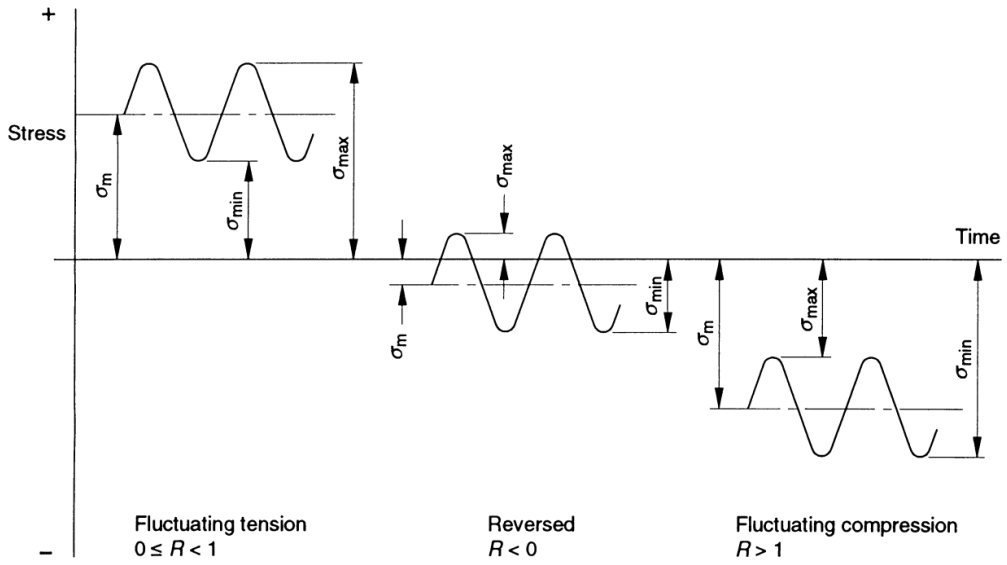


Figure 4.19 – Types of stress cycle with algebraic notation (BS 3518-1:1993)

Table 4.3 – Symbols, terms and definitions relating to strain controlled testing

Symbol	Term	Definition
ε, γ	Strain	The extension of the gauge length divided by the original gauge length. It is taken to be positive when the gauge length increases in length and negative when it contracts
ε_{max}, γ_{max}	Maximum strain	The highest algebraic value of strain in the strain cycle
ε_{min}, γ_{min}	Minimum strain	The lowest algebraic value of strain in the strain cycle
ε_m, γ_m	Mean strain	One-half the algebraic sum of the maximum and minimum strain
$\Delta\varepsilon_t$, $\Delta\gamma_t$	Total strain range	The algebraic difference between the maximum and minimum strain in one strain cycle
$\Delta\varepsilon_p$, $\Delta\gamma_p$	Plastic strain range	The width of the hysteresis loop of stress plotted against strain, determined at the mean stress
$\Delta\varepsilon_e$, $\Delta\gamma_e$	Elastic strain range	The difference between the total strain range and the plastic strain range
ε_a γ_a	Strain amplitude	Half the total strain range
$0.5\Delta\varepsilon_p$, $0.5\Delta\gamma_p$	Plastic strain amplitude	Half the plastic strain range
$0.5\Delta\varepsilon_e$, $0.5\Delta\gamma_e$	Elastic strain amplitude	Half the elastic strain range
$2N_f$	Fatigue life in reversals	The number of reversals, or half cycles, to failure
b	Fatigue strength exponent	The slope of the "elastic" line obtained by plotting the logarithm of the elastic strain amplitude at half-life as the ordinate and the logarithm of the number of reversals to failure as the abscissa, the mean stress being zero as in Figure 4.21
σ'_f, τ'_f	Fatigue strength coefficient	The stress amplitude corresponding to a life of one reversal, obtained by extrapolating the "elastic" line
c	Fatigue ductility exponent	The slope of the "plastic" line obtained by plotting the logarithm of the plastic strain amplitude at half-life as the ordinate and the logarithm of the number of reversals to failure as the abscissa, the mean stress being zero as in Figure 4.21
ε'_f, γ'_f	Fatigue ductility coefficient	The plastic strain amplitude corresponding to a life of one reversal obtained by extrapolating the "plastic" line as
n'	Cyclic hardening exponent	The slope of the line obtained by plotting the logarithm of the stress amplitude at half-life as the ordinate and the logarithm of the half-life plastic strain amplitude as the abscissa, the mean stress being zero as in Figure 4.22
K'	Cyclic hardening coefficient	The stress amplitude corresponding to a plastic strain amplitude of unity; obtained by plotting the logarithm of the stress amplitude at half-life as the ordinate and the logarithm of the half-life plastic strain amplitude as the abscissa, the mean stress being zero as in Figure 4.22

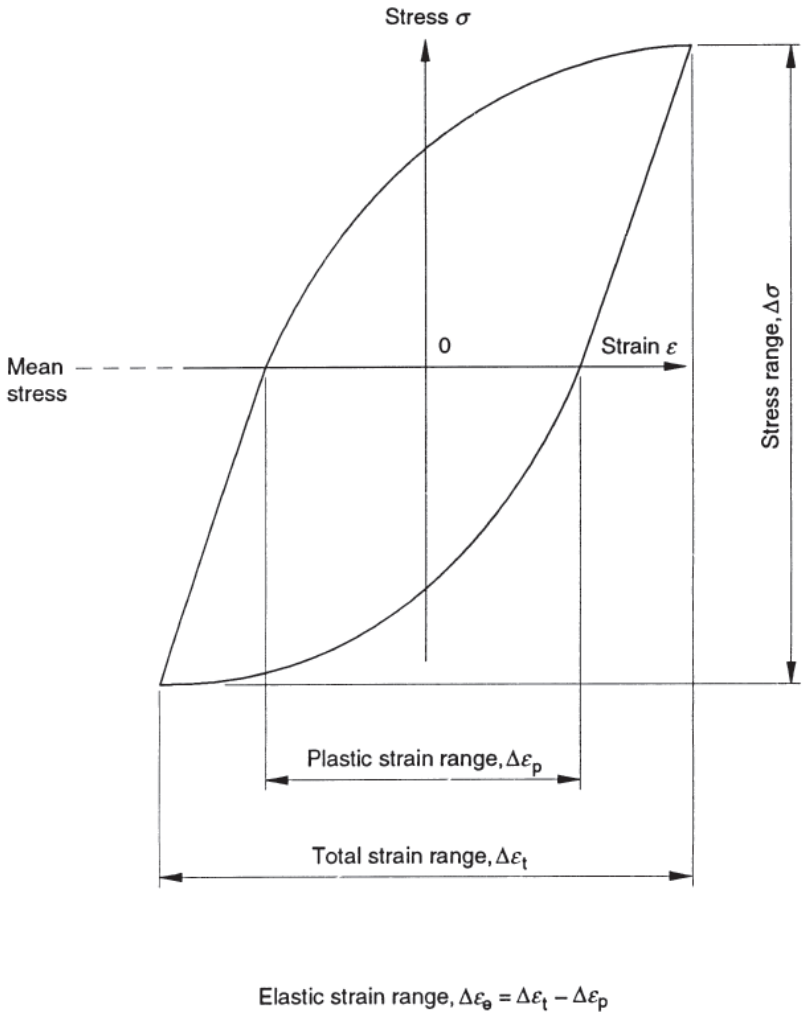


Figure 4.20 – Stress-strain hysteresis loop (BS 3518-1:1993)

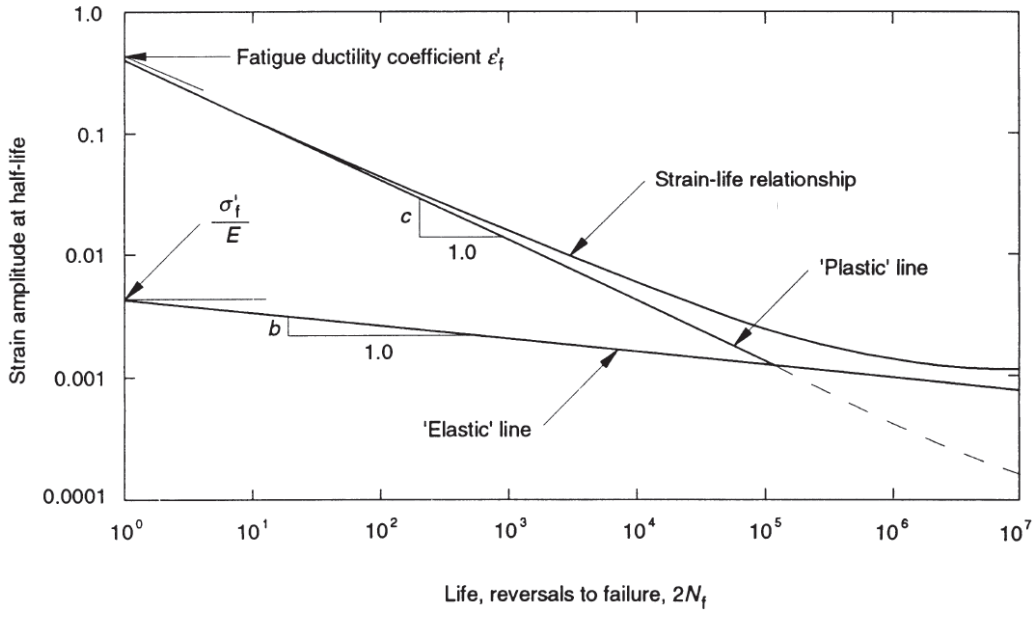


Figure 4.21 – Strain amplitude versus reversals to fatigue (BS 3518-1:1993)

To evaluate fatigue strength and ductility exponents and coefficients, b, c, σ'_f, ε'_f, BS3518-1:1993 directs that the results of fully reversed loading strain controlled tests that comply with BS 7270:2006 are used. An accepted and convenient scheme of analysis involves the plotting of the strain amplitude against the number of reversals to failure on logarithmic scales as in Figure 4.21. The two material constants σ'_f and b are obtained from the intercept and slope of the line fitted through the high cycle part of the data by a regression analysis. This line, which is frequently referred to as the "elastic" line, has the equation:

$$\frac{\Delta\varepsilon_e}{2} = \frac{\sigma'_f}{E}\left(2N_f\right)^b \tag{4.25}$$

In a similar manner, the plastic strain amplitude is plotted against the number of reversals to failure using logarithmic scales as in Figure 4.21. Two further material constants, ε'_f and c, are obtained from the intercept and slope of the line fitted through the data by regression analysis, this time using a regression fit to the low cycle data. This line, which is frequently referred to as the "plastic" line, has the equation:

$$\frac{\Delta\varepsilon_p}{2} = \varepsilon'_f\left(2N_f\right)^c \tag{4.26}$$

In the absence of creep deformations the total strain amplitude consists only of time independent elastic and plastic strain amplitudes. Thus for a material for which straight "elastic" and "plastic" lines can be demonstrated as in Figure 4.21, the strain-life relationship may be represented by an equation of the form:

$$\frac{\Delta\varepsilon_t}{2} = \frac{\sigma'_f}{E}\left(2N_f\right)^b + \varepsilon'_f\left(2N_f\right)^c \tag{4.27}$$

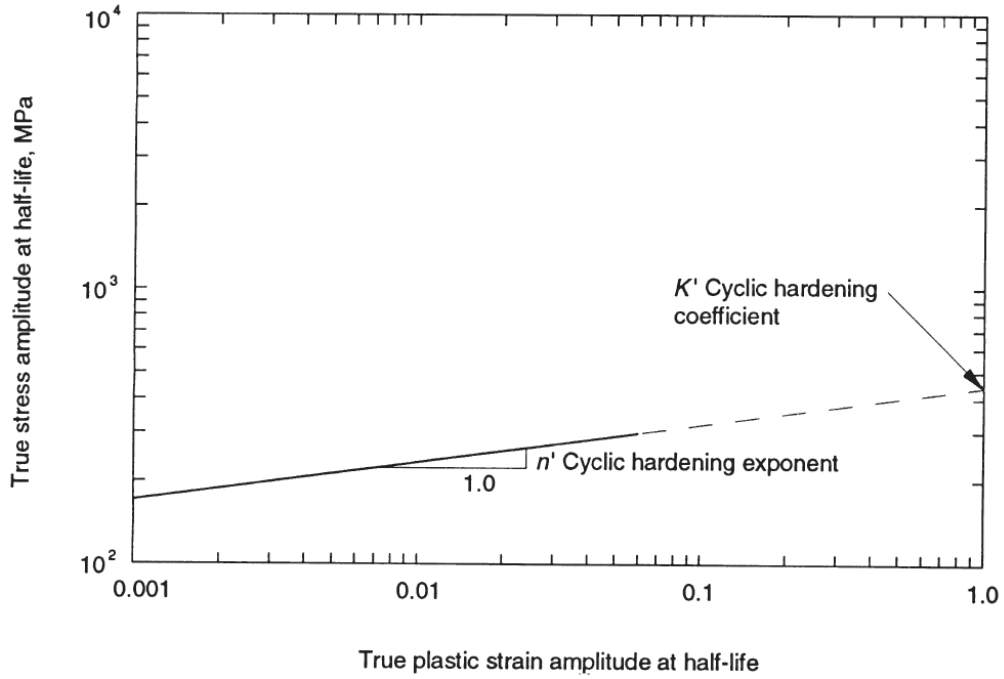


Figure 4.22 – Determination of the cyclic hardening and the cyclic hardening exponent (BS 3518-1:1993)

A cyclic stress-strain curve for a material may be generated from paired values of the stress amplitude at half-life plotted against the total strain amplitude as in Figure 4.22. Cyclic stress-strain behaviour is frequently analysed by separating the total strain amplitude into elastic and plastic strain amplitudes and describing the relationship between cyclic stress amplitude and plastic strain amplitude in the form of a power law. If the values of true stress and true strain are employed, many metals conform to the empirical relationship:

$$\frac{\Delta\sigma}{2} = K'\left(\frac{\Delta\varepsilon_p}{2}\right)^{n'} \tag{4.28}$$

4.7. Multiaxial elastic stress-life methods

A large variety of techniques and methods can be employed to predict the fatigue life of specimens. It is a hot topic nowadays and, for example, the Scopus database offers more than ten thousand papers published for the last five years. The most comprehensive summary of those methods is ASTM's STP1387. However, most companies use self-developed modifications of the existing methods, which are protected by the copyright law. Therefore, it was decided to employ the well-known methods that can be easily modified by adjusting parameters.

The fatigue criteria based on the elastic stress-life methods can be written in a general form:

$$A + \bar{k}B \le \bar{\lambda} \tag{4.29}$$

Therefore, the fatigue parameter can be determined as:

$$FP = \frac{A + \bar{k}B}{\bar{\lambda}} \tag{4.30}$$

The fatigue parameter is calculated for a specific number of life cycles, N_f, and a value below unity means that fatigue failure is unlikely to happen in N_f cycles; if the FP is unity or above, fatigue failure is likely to happen in N_f cycles.

For the Crossland (1956) and Sines (1955, 1959) criteria which are based on the amplitude of second invariant of stress tensor deviator, which corresponds to the von Mises stress, the parameters for equation 4.30 are given in Table 4.4.

Table 4.4 – Parameters of equation 4.29 for Cossland and Sines criteria in formulations referred to by Papadopoulos et al. (1997)

Fatigue method	A	B	$\bar{k}$	$\bar{\lambda}$
Crossland	$\sqrt{J_{2,a}}$	$\sigma_{h,\max}$	$\frac{3\tau_{af}}{\sigma_{af}} - \sqrt{3}$	τ_{af}
Sines	$\sqrt{J_{2,a}}$	$\sigma_{h,m}$	$\frac{3\tau_{af}\left(1+\sigma_{af}/\sigma_u\right)}{\sigma_{af}} - \sqrt{3}$	τ_{af}

where

$$\sigma_h = \frac{\sigma_x + \sigma_y + \sigma_z}{3} \tag{4.31}$$

$$\sqrt{J_{2,a}} = \sqrt{\frac{\left(\sigma_x - \sigma_y\right)^2 + \left(\sigma_y - \sigma_x\right)^2 + \left(\sigma_z - \sigma_x\right)^2}{6} + \tau_{xy}^2 + \tau_{yz}^2 + \tau_{zx}^2} \tag{4.32}$$

It was shown by Papadopoulos et al. (1997), Carpinteri and Spagnoli (2001), Banvillet et al (2003) that the Crossland's criterion provides a better estimation of the fatigue failure.

Another group of fatigue criteria considers load conditions on a so called "critical plane". The critical plane for each particular case can be determined by maximisation of the right-hand side of equation (4.30). The principal normal and shear stress components on a plane for the plane strain case can be determined from the directional stresses on another plane, according to Timoshenko and Goodier (1973) as:

$$\sigma = \sigma_x \cos^2\theta + \sigma_z \sin^2\theta + 2\tau_{xz} \sin\theta\cos\theta \tag{4.33a}$$

$$\tau = \tau_{xz}\left(\cos^2\theta - \sin^2\theta\right) + \left(\sigma_z - \sigma_x\right)\sin\theta\cos\theta \tag{4.33b}$$

The time varying shear stress amplitude for the Dang Van (1973) criterion is used in Ekberg's (1997) formulation

$$\tau_a(t) = \left|\tau(t) - \tau_m\right| \tag{4.34}$$

Parameters of equation 4.29 for the Findley (1953, 1957) criterion modified by Socie (1993), the revision of the McDiarmid (1991,1994) criterion and the Dang Van (1973) criterion are provided in Table 4.5.

Table 4.5 – Parameters of equation 4.29 for Findley, McDiarmid and Dang Van criteria

Fatigue method	A	B	$\bar{k}$	$\bar{\lambda}$
Findley	$\tau_{cp,a}$	$\sigma_{cp,\max}$	$\frac{2-\sigma_{af}/\tau_{af}}{2\sqrt{\sigma_{af}/\tau_{af}-1}}$	$\frac{\sigma_{af}}{2\sqrt{\sigma_{af}/\tau_{af}-1}}$
McDiarmid	$\tau_{cp,a}$	$\sigma_{cp,\max}$	$\frac{\tau_{af}}{2\sigma_u}$	τ_{af}
Dang Van	$\tau_a(t)$	$\sigma_{cp,\max}$	$3\frac{\tau_{af}}{\sigma_{af}}-1.5$	τ_{af}

4.8. Multiaxial strain-life approach

The strain-life prediction methods can be broken down into similar steps. The only significant difference from the numerical point of view is the formula that defines the total damage. Since there is no universally accepted approach and the algorithm similarities it was decided to employ the Fatemi and Socie (1988) criterion in the form of the Bannantine and Socie (1992) revision.

$$\gamma_{cp,a}\left(1+K\frac{\sigma_{cp,\max}}{\sigma'_o}\right)=\frac{\tau'_f}{G}\left(2N_f\right)^b+\gamma'_f\left(2N_f\right)^c \qquad (4.35)$$

where γ_a is the amplitude of the principal shear strain; $\sigma_{\max}$ is the maximum principal tensile stress; b is the fatigue strength exponent; c is the fatigue ductility exponent; τ'_f and γ'_f are fatigue stress and ductility coefficients; G is a shear modulus; σ'_o is the yield stress.

To evaluate the total-life the following steps are taken for each point in the fatigue evaluation area:

1. Read the stress history for the current point.

2. Calculate strain history using Hook's law in compliance format as

$$\varepsilon_x = \frac{\sigma_x - \nu(\sigma_y + \sigma_z)}{E} \tag{4.36a}$$

$$\varepsilon_z = \frac{\sigma_z - \nu(\sigma_x + \sigma_y)}{E} \tag{4.36b}$$

$$\gamma_{xz} = \frac{\tau_{xz}}{G} = \frac{2(1+\nu)}{E}\tau_{xz} \tag{4.36c}$$

3. Consider each potential critical plane orientation in turn using steps 4 to 7 as follows.

4. Obtain normal and shear stress history by equations 4.33 and normal and shear strain histories on the critical plane as

$$\varepsilon = \varepsilon_x \cos^2\theta + \varepsilon_z \sin^2\theta + \gamma_{xz}\sin\theta\cos\theta \tag{4.37a}$$

$$\frac{\gamma}{2} = \frac{\gamma_{xz}}{2}\left(\cos^2\theta - \sin^2\theta\right) + \left(\varepsilon_z - \varepsilon_x\right)\sin\theta\cos\theta \tag{4.37b}$$

5. Count cycles based on the shear strain history according to ASTM International E1049-85(2011)

6. Calculate the fatigue life, N_f, for each cycle identified at step 5 using equation 4.35

7. Calculate the total cumulative damage using the Palmgren-Miner (1924) rule:

$$D = \frac{1}{N_f} = \sum_{i=1}^{\text{all cycles}} \frac{1}{N_{f,i}} \tag{4.38}$$

   for the current candidate critical plane orientation.

8. Repeat steps 3 to 7 for each candidate critical plane and identify the plane that has the largest value of D, D_{max}. This is the critical plane for the point being considered, and D_{max} is the value of the accumulated damage at that point.

9. Repeat from step 1 for all points in the fatigue evaluation area.

10. Plot the value of D_{max} for each point in the fatigue evaluation area.

4.9. Summary

This chapter provides all necessary information to evaluate elastic stresses at the surface and subsurface of the gear flanks, store stress history in the most convenient manner for the fatigue analysis and a brief overview of the fatigue theories and their possible applications.

Chapter 5

The results of the complete analysis of the gear meshing cycle The NASA Glenn case

5.1. Introduction

The method described in Chapters 2 through 4 was implemented in a software package, providing an opportunity to solve a full spur gear meshing cycle, obtain the stress history for the gear flanks in contact and apply the fatigue-life prediction tool. This Chapter describes the structure of the input and output files as well as the step-by-step user manual of the solving process. The NASA Glenn experimental investigation setup presented in report NASA/TM-2005-213956 / ARL-TR-3126 was taken to obtain the results by use of computer modelling. The Chapter provides the solution for each step of the analysis, including smooth and rough surface transient non-Newtonian EHL emulation of the full meshing cycle, plots of the elastic stresses and the fatigue-life evaluation.

The explanation of the results is available, however, their comparison with The Design Unit case and a brief discussion of their differences and the reliability of the method is provided in Chapter 6.

5.2. NASA test set up parameters

The results are presented for the NASA Glenn Research Centre gear fatigue test rig, shown in Figure 5.1 and described in NASA-TN-D-7261.

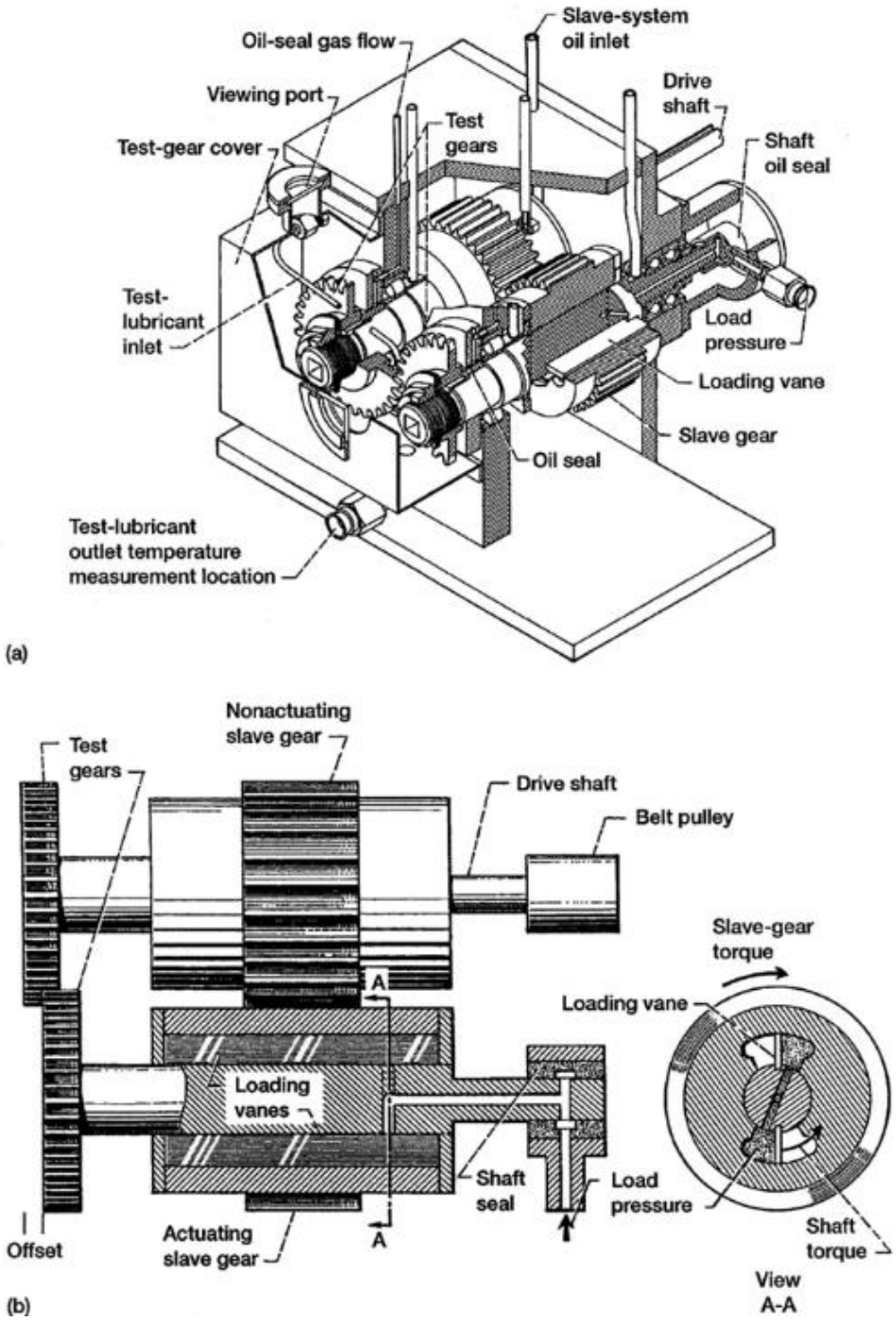


Figure 5.1 – NASA Glenn Research Center gear fatigue test apparatus: (a) cutaway view; (b) schematic view. (Krantz, 2015)

The spur gears mounted on the test rig are identical. Their parameters are provided in Table 5.1. The gears were tested at 10,000 rpm and it was assumed that the gear temperature was equal to the oil outlet temperature, 348 K. The test lubricants used and properties at 348 K are given in Table 5.2. The load conditions were determined from the dynamic tooth force measurements by Krantz (2002), which were carried out according to the technique developed by Rebbechi, Oswald and Townsend (1996). The measured results are plotted in Figure 5.2. The abscissa of this graph is the position of tooth contact line on the line of action in millimetres with the origin at the pitch point. The graph has two ordinate axes which define the magnitude of the load: the left-hand side one states the tooth force in newtons and the right-hand side one specifies the load per unit face width in kN/m. The maximum measured tooth force is 2280 N and it is considered evenly distributed across a 2.79 mm contact line, which results in a pitch-line load of approximately 820 kN/m and a pitch-line maximum Hertz stress of about 2 GPa. This gives a maximum half-Herzian contact dimension of a = 0.264 mm.

Table 5.1 – Spur test gears design parameters

Geometrical parameters		
Module / mm	m	3.175
Number of teeth	n	28
Pressure angle / °	ψ	20
Outside radius / mm	r_t	47.625
Base radius / mm	r_b	41.769
Material properties of AISI 9310		
Young's modulus / GPa	E	206.85
Poisson's ratio	υ	0.3
Coefficient of friction (dry contact)	μ	0.1

Table 5.2 –Lubricant properties (NASA/TM-2005-213956/ARL-TR-3126)

Name	Absolute viscosity at 348 K / N-s/m^2	Pressure-viscosity coefficient / m^2/N
Basestock oil 1	0.014	9.0×10^{-9}
MIL-L-23699	0.018	10.5×10^{-9}
MIL-L-7808J	0.010	7.5×10^{-9}
DOD-L_85734	0.017	11.0×10^{-9}
DRED-2478	0.022	12.0×10^{-9}
Basestock oil 2	0.028	11.0×10^{-9}

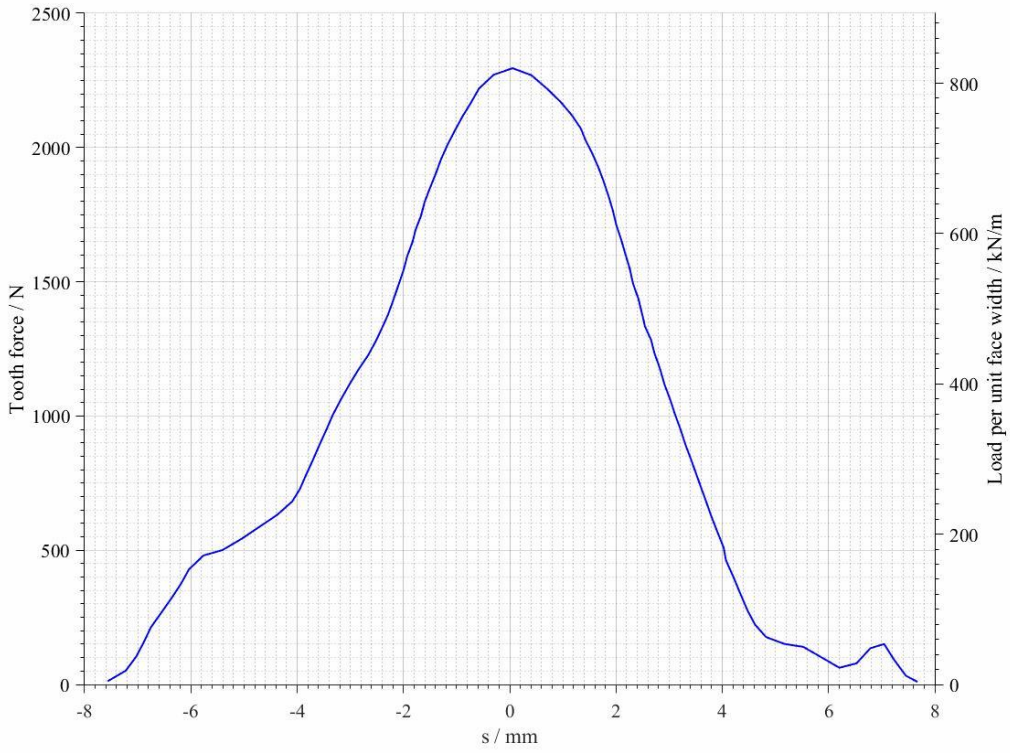


Figure 5.2 – Measured tooth force (Krantz, 2002)

The roughness profiles used were provided by Dr. Krantz on behalf of NASA and ARL. There were two groups of profiles: as manufactured, which were measured before any type of tests or running-in were conducted, and run-in, which were run-in at a load per unit face width of 123 kN/m for one hour. In the test the load is then increased to the static test load of 580 kN/m. However, as it was mentioned before, the maximum measured dynamic load is about 820 kN/m. The run-in gear profiles do not show any traces of damage of any kind. The EHL analysis at the current stage of development does not incorporate any tool which would allow the change of the initial geometry due to stresses above the critical level to be taken into consideration. Therefore, instead of the permanent change of the surface profiles due to the plastic deformation, surface roughness features can drive the pressure at the nodes to be unfeasibly high, which causes unrealistic behaviour of the lubricant. However, if the gears run at a higher load then further plastic deformation will take place due to the higher load and the higher temperatures reached due to increased frictional heating. Hence, it was decided to use the run-in profiles because all plastic processes due to the

contacts with the counterfaces changed the gear tooth profiles over running-in period and any further contacts are predominantly elastic. The profiles used for the EHL analysis were run against each other for a short period of time at the full test load and are shown in Figures 5.3 and 5.4. Figures 5.3(a) and 5.4(a) present the profilometer measurements of the gear flanks taken using a 2 μm radius, conisphere-tipped stylus. Then a least-square curve form removal was applied followed by the application of an ISO-conforming Gaussian roughness filter having a 0.8 mm and 0.25 mm cutoff, shown in Figures 5.3(b), 5.4(b) and 5.3(c), 5.4(c) respectively. Figures 5.3(b) and 5.4(b) still show some traces of waviness, therefore, the roughness profiles shown in Figures 5.3(c) and 5.4(c) were used for the further analysis.

5.3. EHL and stress analysis solution approach

The geometrical parameters, the kinematics and the loading conditions vary as the contact progresses along the line of action over the meshing cycle. Therefore, the EHL analysis is a transient problem, which can be considered as a series of timesteps with appropriate set up parameters linked by the full transient equations.

The size and parameters of the computational EHL mesh are constant and it is scaled to the Hertz dimension at the pitch point, which is covered in Section 3.2. The mesh is worked out at the beginning of the EHL analysis and it is determined by four parameters specified in the input file: the pitch point load , *load_max*, number of nodes in the half-Hertzian width, *ninb*, number of the half-Hertzian dimensions from the centre of the contact upstream and downstream, *upb* and *dob* respectively. Hence, the spacing of the mesh is

$$\Delta x = \frac{a_{\text{pitch point}}}{ninb} \tag{5.1}$$

and the total number of the EHL mesh nodes, *nnode*, is

$$nnode = \left(upb + dob \right) \cdot ninb + 1 \tag{5.2}$$

The algorithm of the solution process is shown in Figure 5.5.

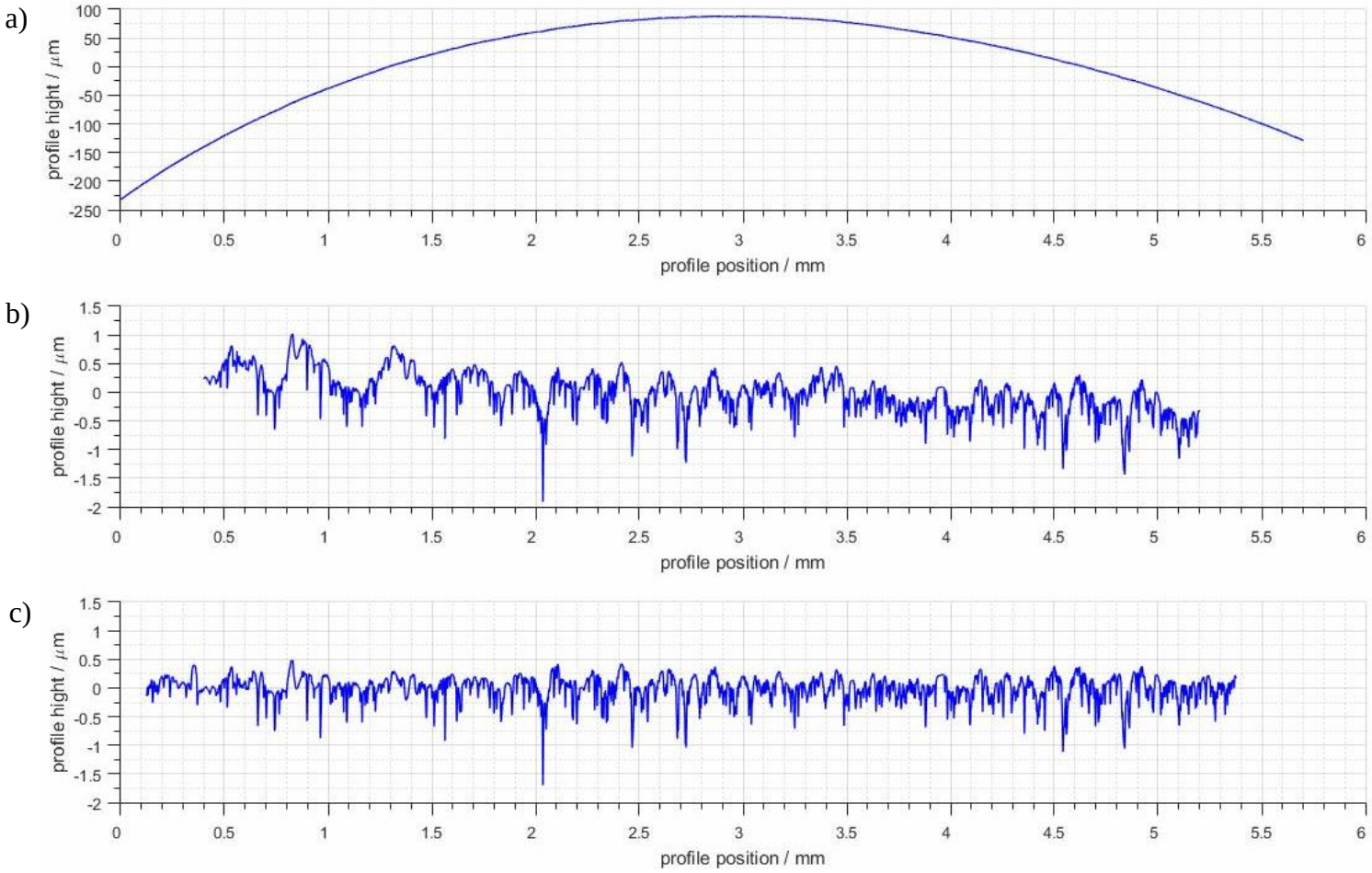


Figure 5.3 – Pinion surface roughness: (a) raw data; (b) form removed and Gaussian filter with 0.8 mm cutoff applied; (c) form removed and Gaussian filter with 0.25 mm cutoff applied

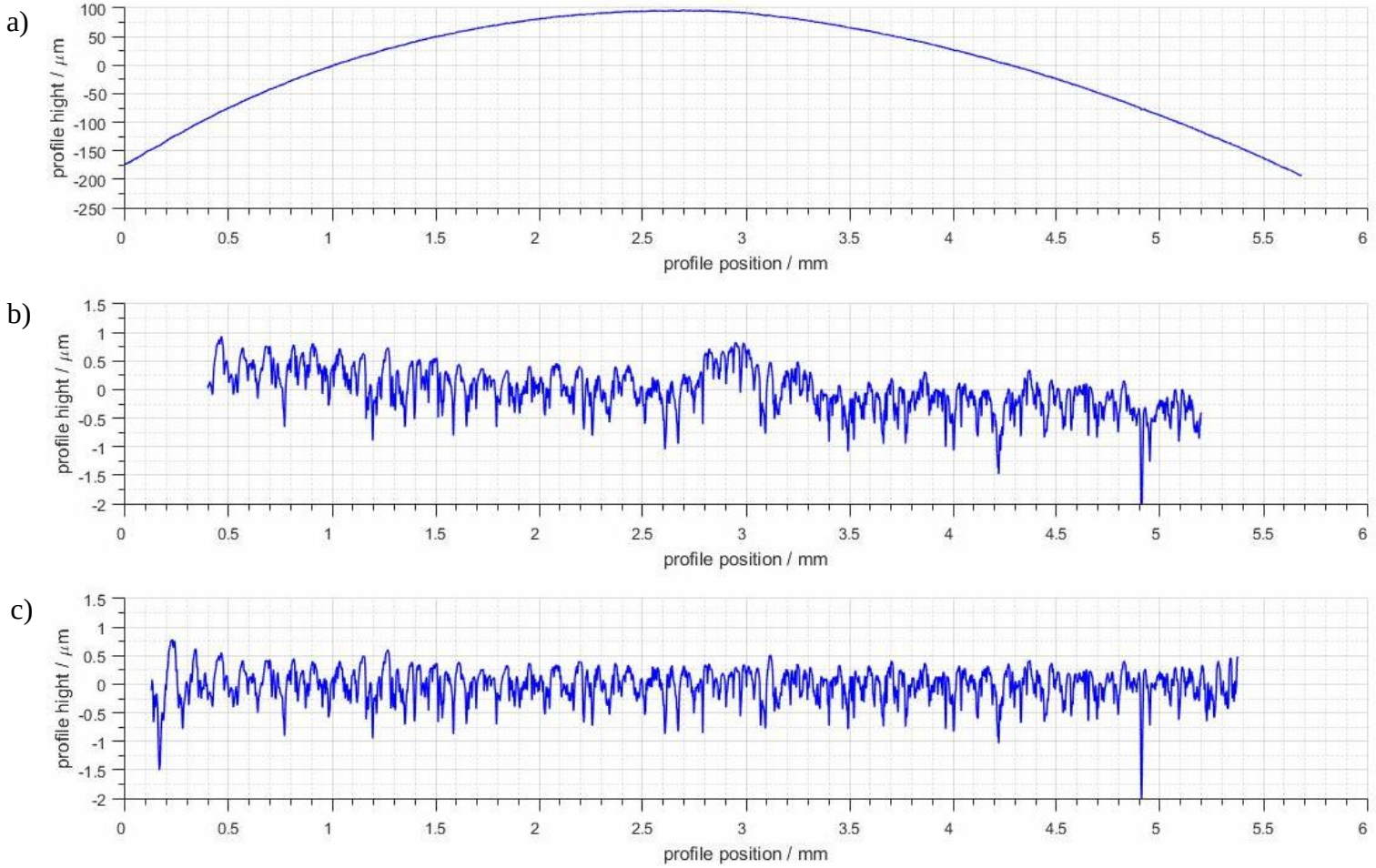


Figure 5.4 – Cog-wheel surface roughness: (a) raw data; (b) form removed and Gaussian filter with 0.8 mm cutoff applied; (c) form removed and Gaussian filter with 0.25 mm cutoff applied

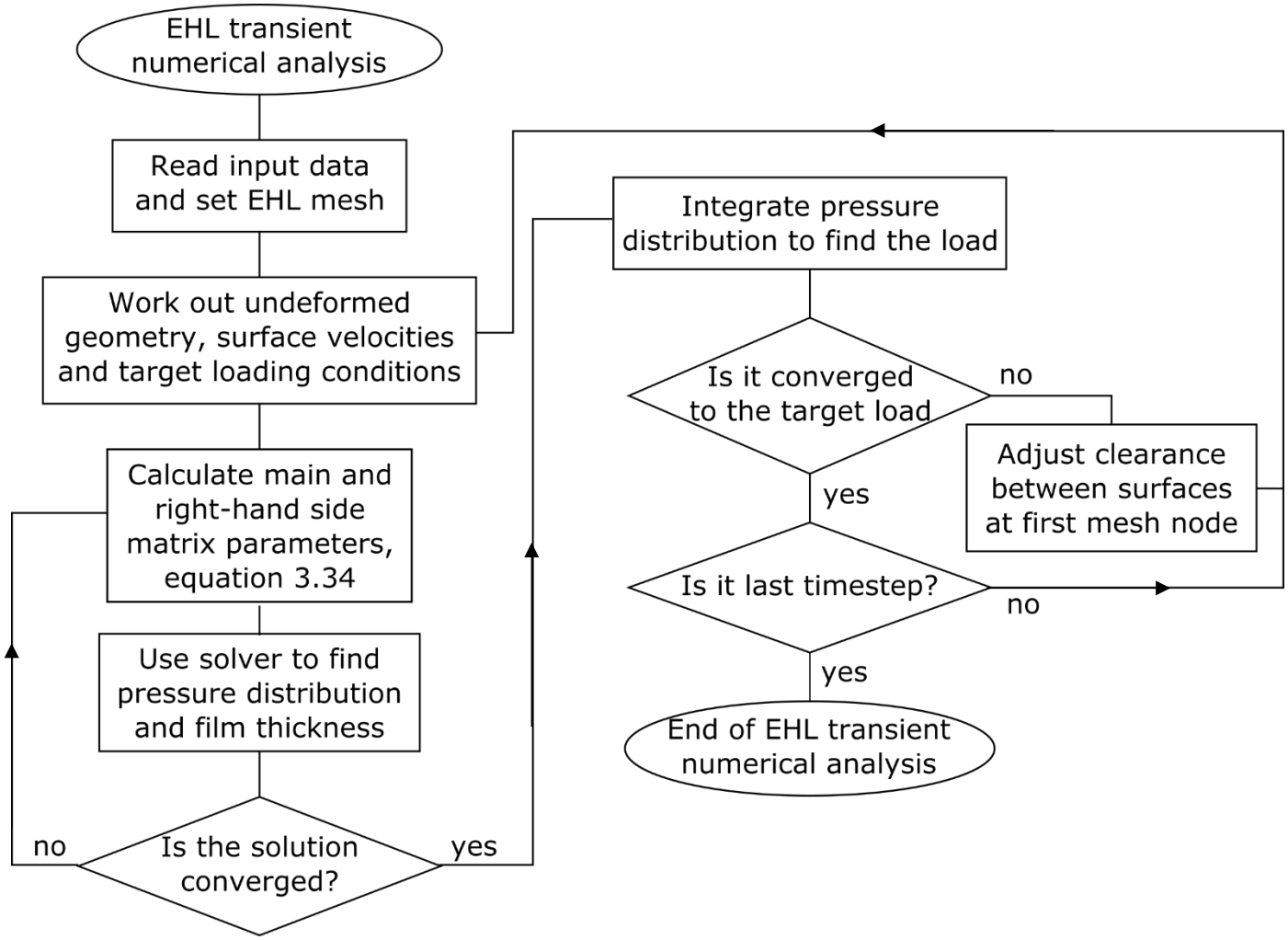


Figure 5.5 – Flowchart of the EHL analysis

The software reads the input data stored in two files: *ipg01.dat* and *load.txt*. The *ipg01.dat* is the main input file, which contains the names of the current job and the surface roughness files, the selection parameters for the formulation and the solution method, the precision of the calculations, the geometric and kinematic parameters of the gear pair, the maximum load and the discretization parameters for the mesh, relaxation factors, mechanical properties of the material and the lubricant. These parameters are used to find the maximum dry contact Hertz dimension for smooth surfaces, a_{max}, which is a scaling factor in the analysis. The file also specifies the timesteps at which results are to be tabulated. File *load.txt* holds the load conditions. The first column contains the coordinate on the path of action, *s*, in meters and the second one has the values of the load in N/m. The file is only used for the initial smooth surfaces analysis.

The undeformed geometry and surface velocities can be calculated by equations (2.9) and (2.14) respectively. The target load is determined by linear interpolation of the appropriate segment of the tabulated target load read from the file. The load conditions can be defined by setting the clearance between the surfaces at any mesh node. Since the first mesh node in the inlet zone experiences the lowest elastic deflection due to the significant distance from the centre of contact, an adjustment of the separation at this point allows faster convergence of the load and reduces potential numerical instabilities. Thus, the clearance between flank surfaces at the first mesh node is set and adjusted through the analysis to obtain the loading conditions for the particular mesh.

The content of the main matrix and the right-hand side matrix is explained in Chapter 3 and can be formulated by the finite difference or the finite element methods. There are two solvers provided, one is based on the Gauss-Seidel method of successive displacement and other on the Gaussian elimination. The Gauss-Seidel method is an iterative technique which allows modification of the formulation of the problem within the convergence process but at the same time might cause some numerical instability and add some noise to the results obtained. The Gaussian elimination is a row reduction algorithm, which obtains the solution by direct algebraic manipulation but all necessary measures to alter the problem formulation, in case of a dry contact of the surfaces at a node for example, must be done separately and can increase the

computing time. The odd row entries of the matrices, which come from the Reynolds equation, must be updated for each candidate pressure distribution due to the non-linearity of the equation, in view of the fact that density and viscosity are functions of pressure. The solution is considered converged when the difference between candidate solutions of previous and current convergence cycles comply with the convergence criteria set in the input file.

The product of the sum of positive values of the converged pressure at all nodes and the mesh spacing gives the load in N/m. It is compared with the target value for that gear mesh position as illustrated in Figure 5.2 and if it does not comply with the convergence criteria an adjustment of the clearance between the surfaces at the fist mesh node is made. The timestep solution finding procedure is repeated until the load is correct. The algorithm must be repeated for each timestep.

In the case of rough surfaces, it is problematic to converge the load due to the interaction of the asperities, because the formulation of the contact can change due to a different disposition of the surfaces, for example from full-film contact to dry one. Hence, the constant change of the clearance between surfaces in order to achieve appropriate load conditions for the rough surface analysis can cause considerable numerical difficulties. Moreover, the software only takes into account elastic deflection of the gear flanks, but the gear teeth are large elastic bodies which deflect at the root when they brought into contact. Therefore, the excess or scarcity of the generated load can be adjusted by an insignificant change of the deflection at the root of the gear teeth.

To overcome the problem described above, the smooth surface case results for the specified lubricant, gears and the target load variation are obtained and the converged time varying clearance values between mating surfaces at the first mesh node are used to set the load conditions for the rough surface analysis. Note that the mesh boundaries and spacing must be identical for both analyses in order to achieve commensurable load conditions. The results of the converged loads and clearances between the mating gear flanks are shown in Figures 5.6 and 5.7 respectively. The abscissae determines the coordinate on the path of action, s, in mm and the ordinates specify the load per unit face width in N/m and the clearance at the first mesh node, h_1, in μm. The

minimum magnitude of the load is set to 100 kN/m because lower values cause some numerical instabilities and make load convergence process very time consuming.

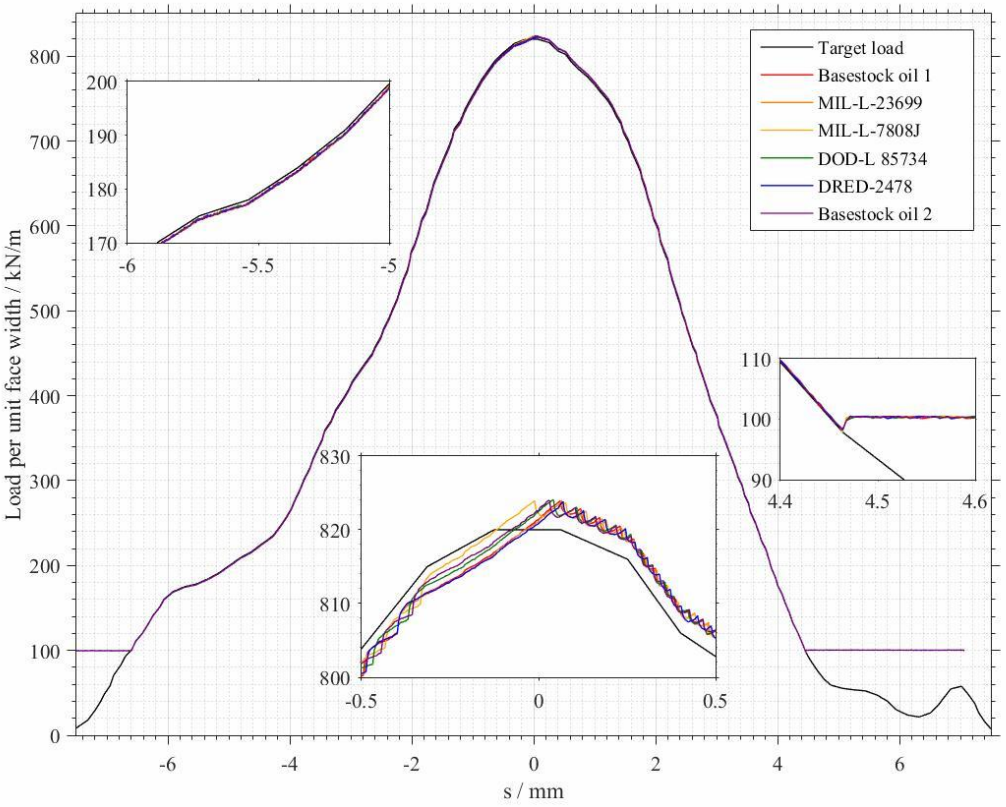


Figure 5.6 – Converged loads for lubricants specified in Table 5.2

The global convergence criteria of the analysis was set as 0.5%.The zoomed-in plot in Figure 5.6 shows the deviation of the converged loads from the target load. The graphs are very close for the lower loads, since the acceptable loads, for example, at 100 kN/m are [99.5; 100.5] kN/m, which can be seen at the end of the meshing cycle. At the pitch point, the load must be [815.9; 824.1] kN/m. Note, that during the monotonic increase of the load, the converged values are always below the target and right after the pitch point, where the load function starts to decrease monotonically the converged values are higher than target ones. The convergence process causes the mild oscillations at the tip, because the clearance between surfaces in contact is adjusted only if the candidate load is not within the allowed range of the target load. The converged loads increase linearly at *s* [-0.4; 0] in Figure 5.6 which corresponds to the constant clearance values in Figure 5.7. There is an immediate response of the clearance between the

surfaces to the change of the load and visible kinks at s = -6.6 and s = 4.5, i.e. the end and the beginning of the load modification respectively and the clearance function is clearly different at the beginning and the end of the meshing cycle.

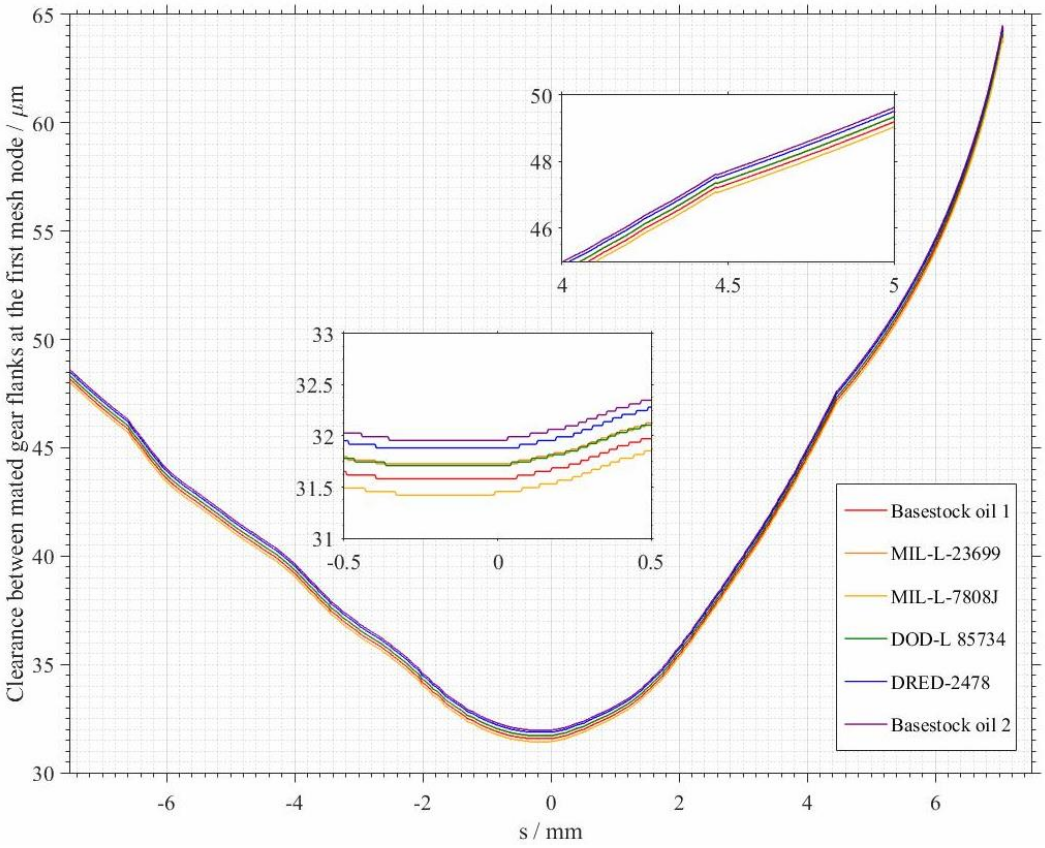


Figure 5.7 – Converged clearance between mated gear flanks at the fist mesh node for lubricants specified in Table 5.2

The results of the rough surface analysis can be grouped into two classes: the EHL analysis results and the stress history. The EHL analysis results are split in two subcategories: general and detailed. The general parameters are saved in ***Job_name*.inf** and the column structure of it is presented in Table 5.3. The minimum film thickness and the maximum pressure as well as film thickness at the centre of contact are plotted in Figure 5.8 for the smooth surface analysis. The radii of curvature and the tangential velocities are plotted in Figure 5.9.

Table 5.3 – Column structure of ***Job_name*.inf**

Column	Parameter	Units
1	Coordinate on the path of action, *s*	mm
2	Film thickness at the contact point	μm
3	Maximum value of the pressure at the EHL mesh	GPa
4	Minimum value of the film thickness at the EHL mesh	μm
5	Radius of curvature of the pinion flank	mm
6	Radius of curvature of the gear flank	mm
7	Radius of relative curvature	mm
8	Tangential velocity of the pinion flank	ms^{-1}
9	Tangential velocity of the gear flank	ms^{-1}
10	Mean tangential velocity of the pinion and the gear flanks	ms^{-1}
11	Integrated load	N/m

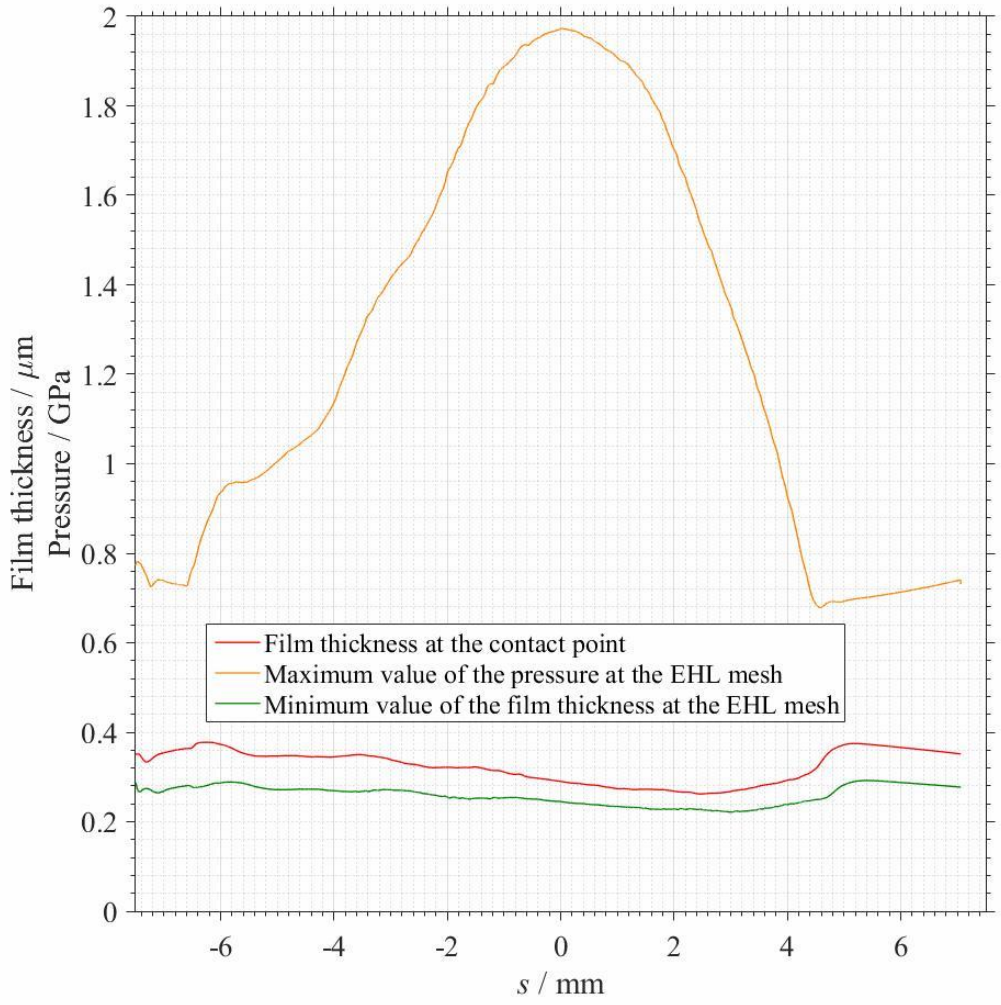


Figure 5.8 – Film thickness at the contact point and minimum value and maximum pressure value over meshing cycle of smooth surfaces

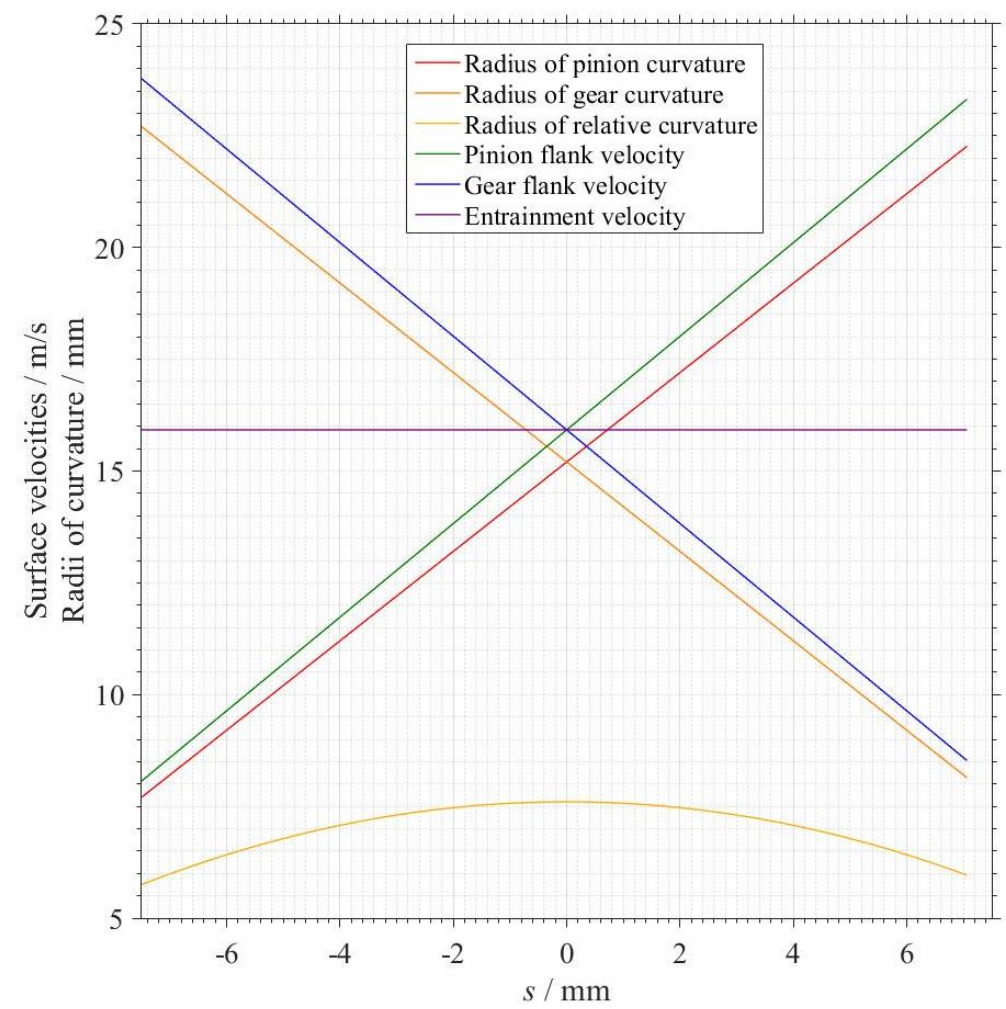


Figure 5.9 – Undeformed geometry and kinematics of the meshing cycle

Figure 5.10 presents the minimum film thickness and the maximum pressure for the EHL mesh as well as film thickness at the centre of contact for the analysis incorporating the rough surfaces. The difference between data plotted in Figures 5.8 and 5.10 is apparent. The difference in film thickness and maximum pressure is driven by the incorporated roughness of the mating gear flanks. The maximum pressure values are more than twice the smooth surface contact ones. The minimum film thickness of the rough surface analysis is of the same order as that of the rough surface, but regularly drops to zero which means that at least one of the mesh nodes the lubricant is unable to separate the mating surfaces. This is particularly apparent at the highest loads that occur between s = -2 and s = +2.

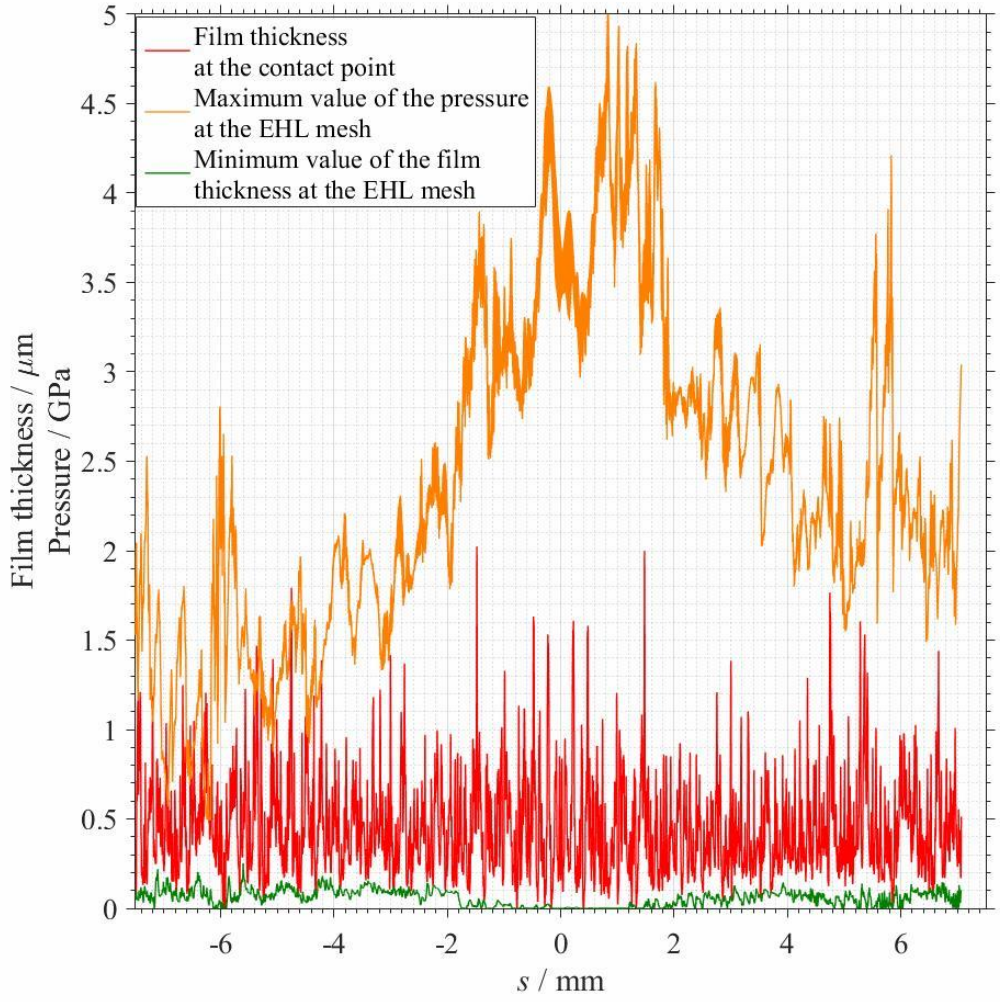


Figure 5.10 – Film thickness at the contact point and minimum value and maximum pressure value over meshing cycle of rough surfaces

The number of nodes in a dry contact for the rough surface contact lubricated by MIL-L-7808J oil over the mesh cycle is plotted in Figure 5.12. The abscissa is the coordinate on the line of action with the origin at the pitch point and the ordinate is the number of mesh nodes at which dry contact is occurring in that timestep. This shows that direct contact of the asperity features happens during the meshing cycle but occurs at relatively few mesh points. It is most prevalent at the higher load area of the meshing cycle.

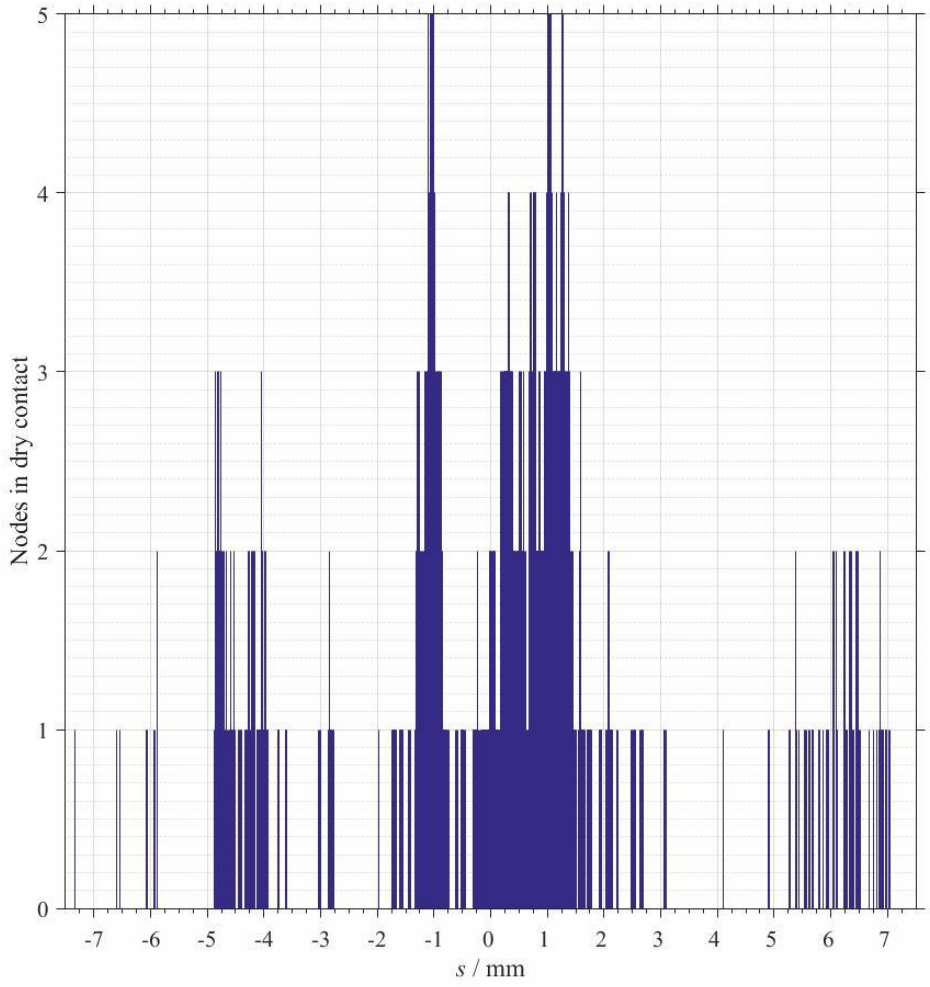


Figure 5.11 – Number of mesh nodes in dry contact at each timestep of the analysis

The load results of the rough surface analysis set by the clearance between the surfaces at the first mesh node for MIL-L-7808J oil over the meshing cycle, shown in Figure 5.7 in amber, are plotted against the converged load of the smooth surface analysis for the same lubricant in Figure 5.12. The abscissa is the coordinate on the line of action with the origin at the pitch point and the ordinate shows the load per unit face width in kN/m. The load exceeds the target values at the close proximity to the pitch point, however, the difference is less than 4% so that no modification of the load set up data is needed. In some cases the interaction of the asperities can cause a significant divergence from the target load, so, the clearance at the first mesh node must be

modified. The value of the change can be determined by trial and error method over several complete analyses.

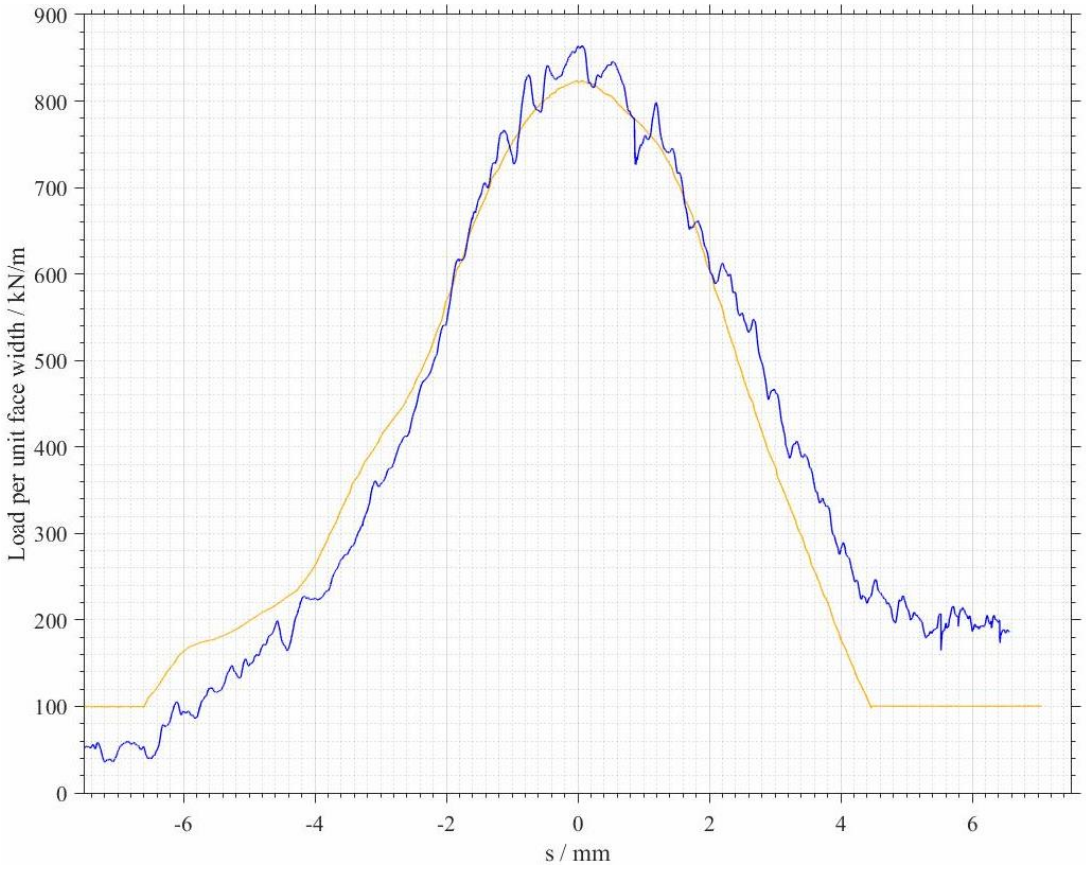


Figure 5.12 – Load over the meshing cycle: smooth (amber) and rough surface contact lubricated by MIL-L-7808J oil (blue)

The difference between the load values up to 600 kN/m is apparent in Figure 5.12. The rough case load is lower than the smooth case one at the monotonically increasing side and greater at the decreasing side. This might be due to differences in the squeeze film load contribution at the lower level surface interaction. The load generated due to the approach velocity may be different for the smooth and the rough surface analysis and if so, the effect would be the opposite when the velocity was separating the surfaces.

The detailed parameters of the EHL analysis are stored in the series of flies, ***.txt**. The file name is a six-digit number, which identifies the timestep number. The column structure of those files is specified in Table 5.4.

Table 5.4 – Column structure of ***.txt**

Column	Parameter	Units
1	x/a	
2	Pressure	GPa
3	Film thickness	μm
4	Offset of the pinion flank	μm
5	Offset of the gear flank	μm
6	Tangential load at the pinion flank surface	Pa
7	Tangential load at the gear flank surface	Pa
8	Indicator of cavitation (-1)/full film (0)/ direct contact (1)	

A universal design of the graphic representation of the results for each timestep defined in Table 5.4 was developed. The data can be automatically processed either for a single timestep or for a series of timesteps by the MATLAB code presented in Appendix A. The results are presented in Figures 5.13 through 5.16. There are three plots sharing the same abscissa, which gives the relative distance from the contact point in fractions of the half-Hertzian contact dimension at the maximum pitch load. The origin is set at the contact point and the axis is oriented in the direction of motion. The upper subplot contains the graphs of pressure distribution in red plotted on the left axis and the film thickness in green plotted on the right axis. The middle plot shows the offsets of the mating gear flanks at a reduced vertical scale and offset for clarity. The lower plot is a three-state indicator, where a filled lower-half position designates cavitated nodes, a filled upper-half designates the dry contact and the mid-line means that the surfaces in the contact are separated by lubricant.

Figure 5.13 shows the very beginning of the meshing cycle. The surfaces are just engaged in the contact. The radius of curvature of the pinion, upper surface offset, is visibly smaller than that of the gear. The area of the contact is small in terms of Hertz dimension, *a*, because the load is low at this timestep integrating to 75 kN/m.

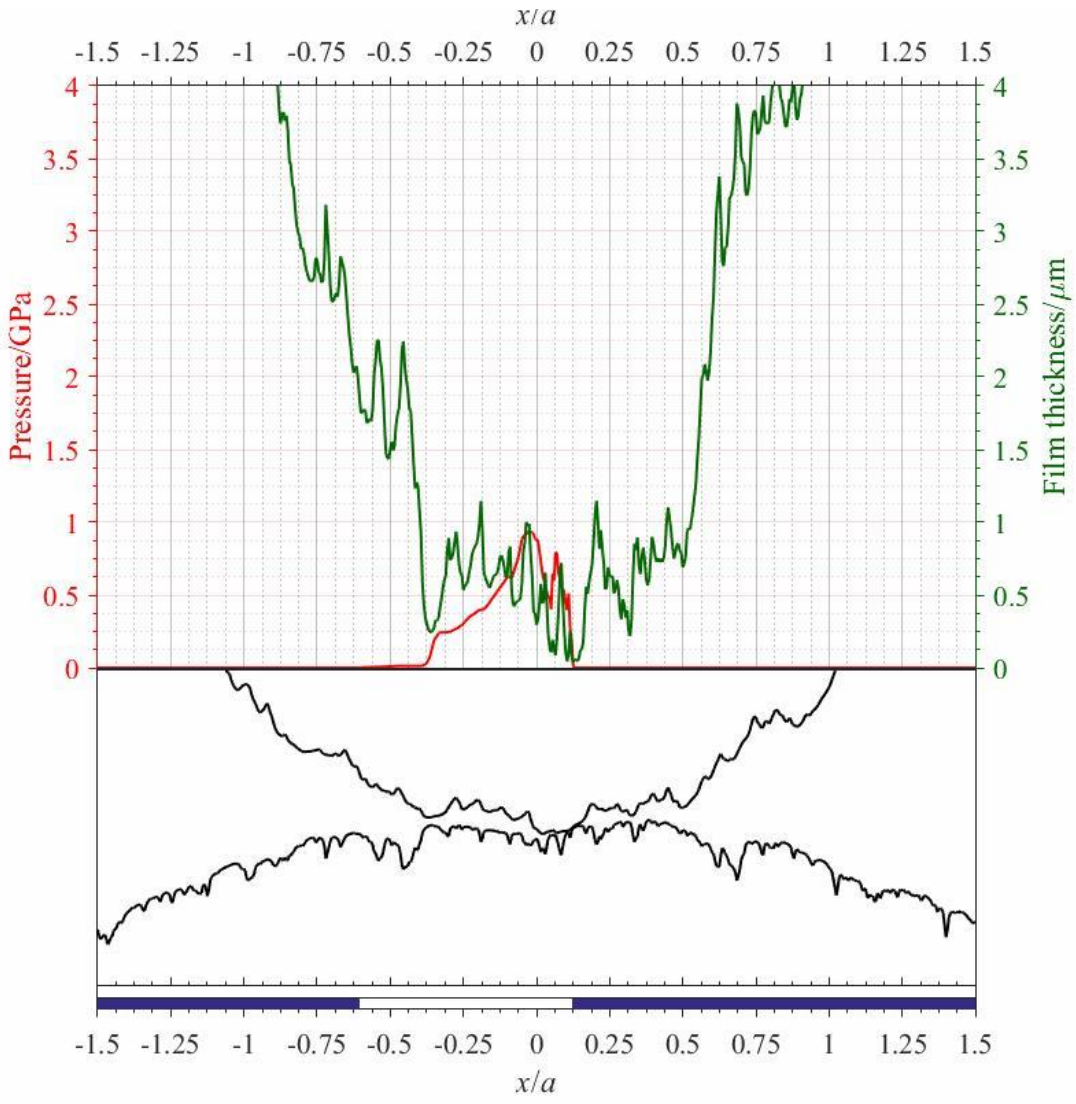


Figure 5.13 – Pressure distribution (red), film thickness (green), offsets of the surfaces (black, factor 4.446) and regime indicator (blue) at s = – 7.6765 mm

Figure 5.14 shows the results of timstep 5535, which corresponds to the coordinate of – 1.001 mm on the line of action with the origin at the pitch point. The radius of curvature of the pinion, upper surface offset, is still visibly smaller than that of the gear, but the difference is not as apparent as in Figure 5.13. The length of the contact has grown significantly. The mating surfaces are in dry contact at mesh points located at x/a values of – 0.645, – 0.340, 0.020 and 0.510. These four nodes split the contact into five local contacts with a significant increase of the hydrodynamic pressure at the outlet zone. There is a local cavitation at x/a = – 0.345. The pressure distribution integrates to 753 kN/m.

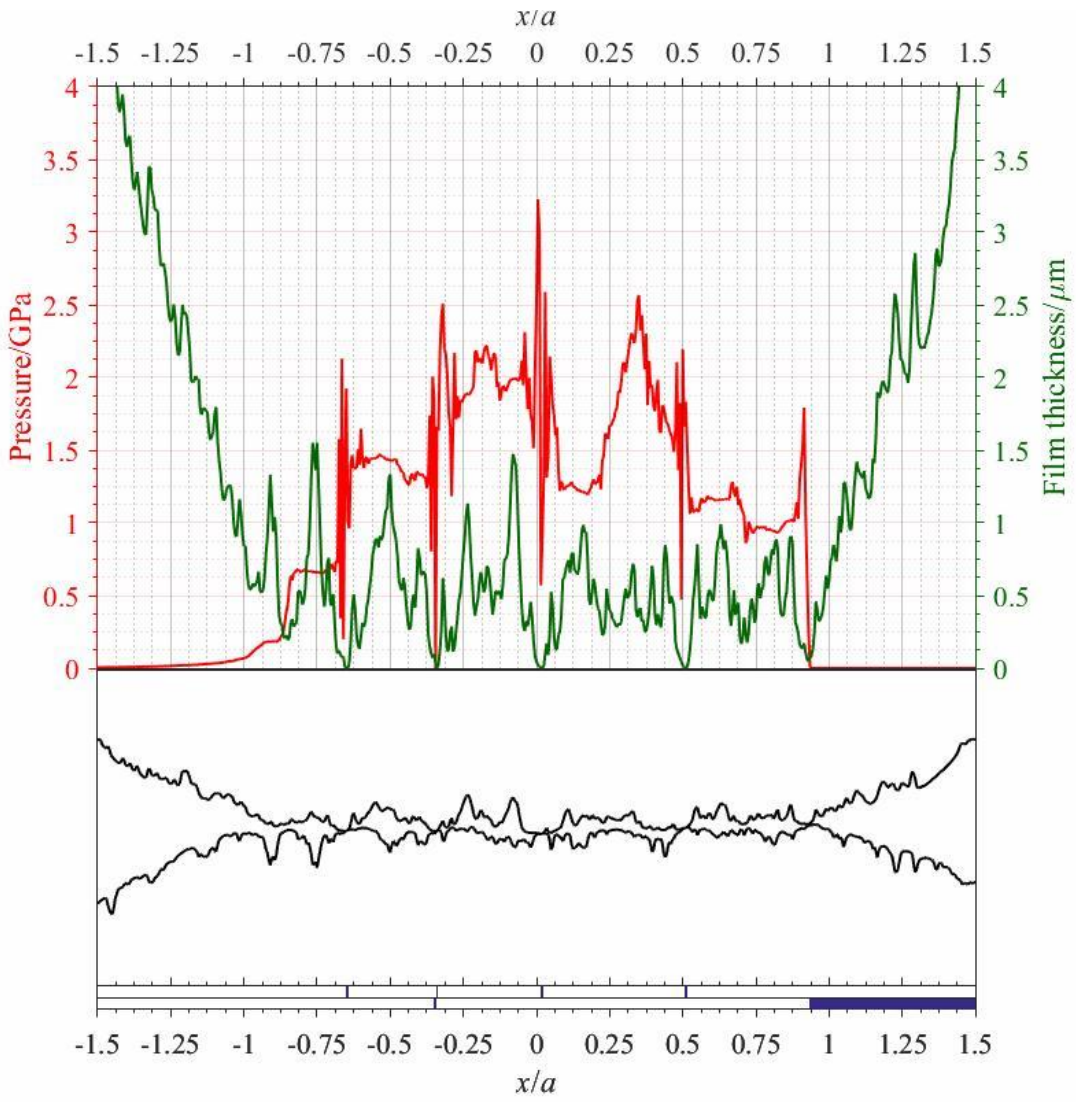


Figure 5.14 – Pressure distribution (red), film thickness (green), offsets of the surfaces (black, factor 4.446) and regime indicator (blue) at s = – 1.001 mm

Figure 5.15 presents the results at the pitch point, timstep 6365. The radii of curvature of the pinion, upper surface offset, and the gear, lower surface offset, are equal due to the identical number of the teeth. The length of the contact has grown even more and is approximately the same as the reference Hertzian dimension, since it is within the range [-1, 1] of x/a. There is no local cavitation and dry contact only occurs at one mesh point at this timestep. The pressure distribution integrates to 823 kN/m.

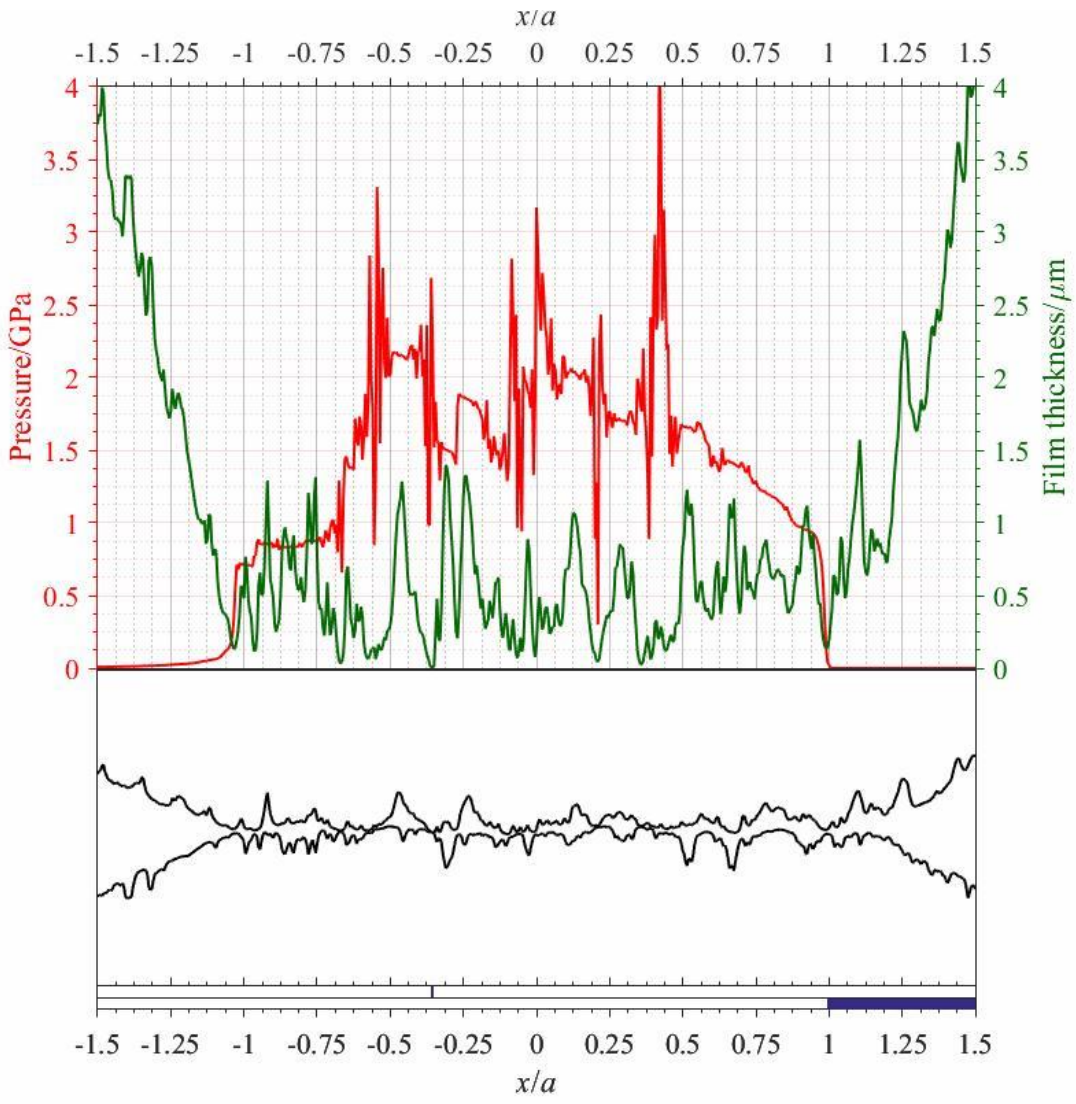


Figure 5.15 – Pressure distribution (red), film thickness (green), offsets of the surfaces (black, factor 4.446) and regime indicator (blue) at the pitch point, $s = 0$

The picture late in the meshing cycle is plotted in Figure 5.15. This is timstep 12224, which corresponds to the coordinate of s = 7.0675 mm. The surface offsets are close to the inverse of the surface offsets in Figure 5.12. The length of the contact has shrunk down and the pressure has been reduced significantly. There is no local cavitation or dry contact. The pressure distribution integrates to 104 kN/m.

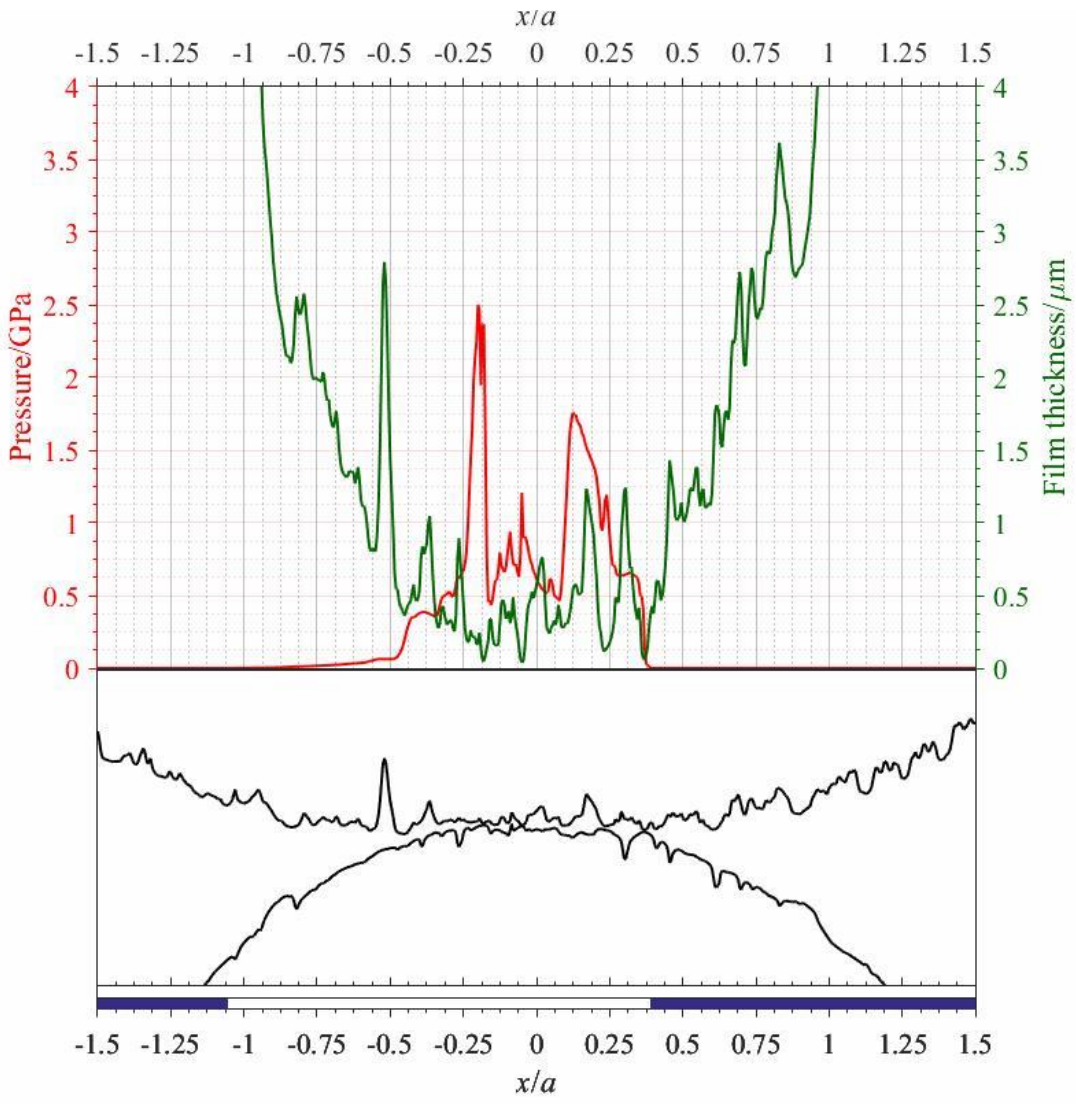


Figure 5.16 – Pressure distribution (red), film thickness (green), offsets of the surfaces (black, factor 4.446) and regime indicator (blue) at *s* = 7.0675

The stress developed in the gear material due to the contact loading is calculated for the EHL mesh. The three stress components are stored in series of files ***.cmp**, where the asterisk is a six-digit timestep ID number. The structure of the file is defined in Table 5.5.

Table 5.5 – Column structure of ***.cmp**

Column	Parameter		Units
1	x/a		
2	z/a		
3	Pinion	σ_{xx}	Pa
4		σ_{zz}	
5		τ_{xz}	
6	Gear	σ_{xx}	
7		σ_{zz}	
8		τ_{xz}	

The stress components based on the pressure and tangential traction distribution when the contact is at the pitch point are plotted in Figure 5.17 for both surfaces. The von Mises and the Principal shear stresses are shown in Figure 5.18. Each of the plots incorporates two subplots, which share abscissa with the origin at the contact point and oriented in the direction of motion. The ordinates originate at the surface and are directed normally into the surface. Both coordinates are normalized in terms of the half-Hertzian dimension at the pitch point. The stress plots generally follow the Hertz stress pattern for two collinear cylinders in contact with a significant noise due to the roughness which has maximum impact at the surface and rapidly reduces with depth into the gear flanks. The maximum values of the stress correspond to the peaks of the pressure as expected.

The stress components are then interpolated from the EHL mesh into the gear flank mesh. The details of that process are covered in Section 4.4. The resulting stress history files for the material are used for further fatigue analysis. Results of these calculations are presented in Figures 5.19 through 5.26 for the EHL analysis carried out using MIL-L-7808J oil. The pinion flank results are shown from the root to the tip and the other way around for the gear tooth profile.

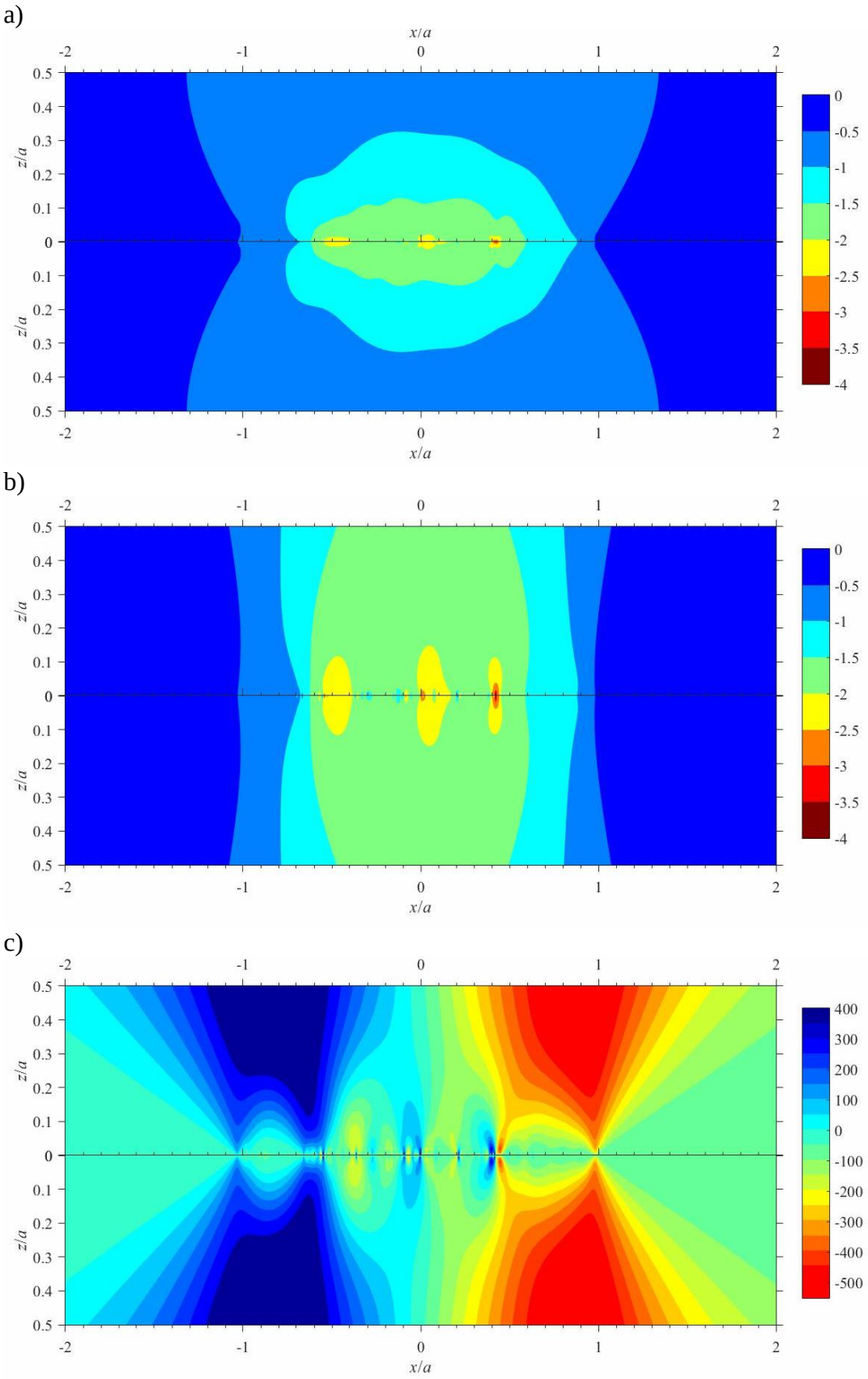


Figure 5.17 – Stress components at the pitch point:

(a) σ_{xx} in GPa, (b) σ_{zz} in GPa, (c) τ_{xz} in MPa

a)

b)

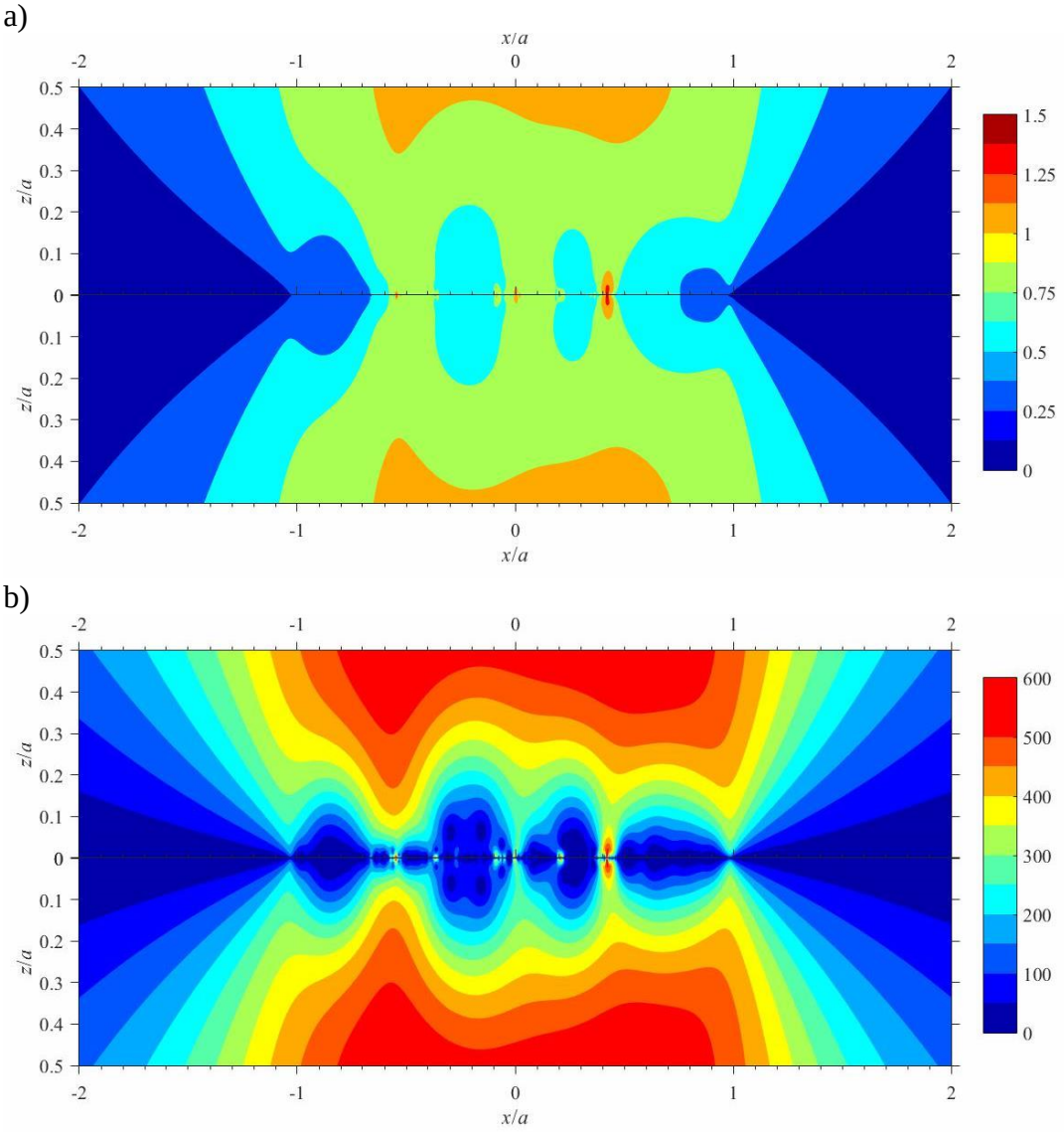
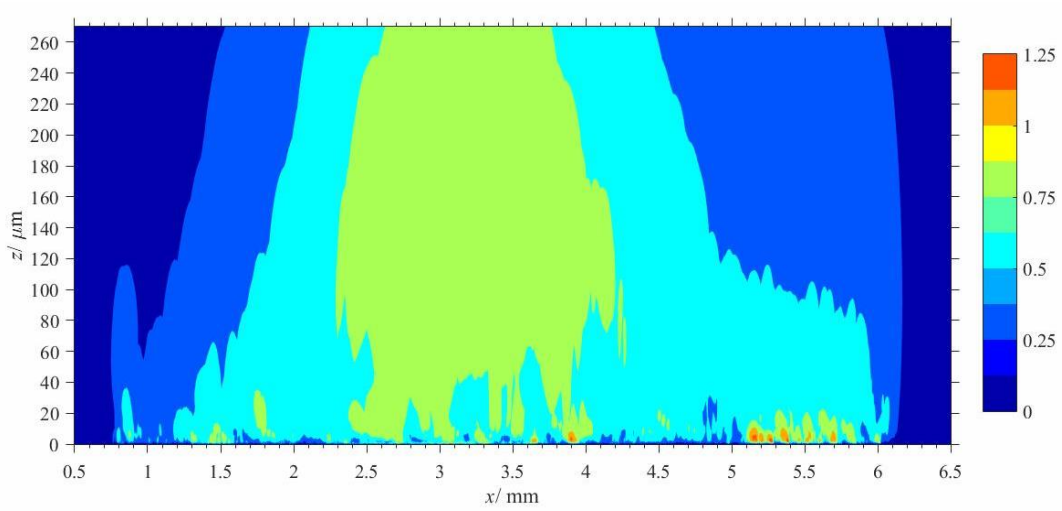
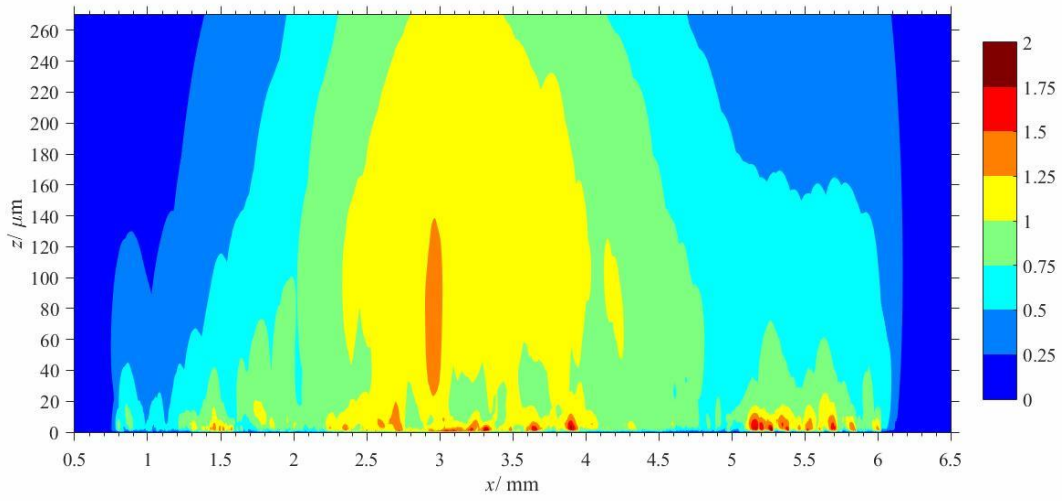


Figure 5.18 – (a) von Mises stress in GPa and
(b) Principal shear stress in MPa at the pitch point

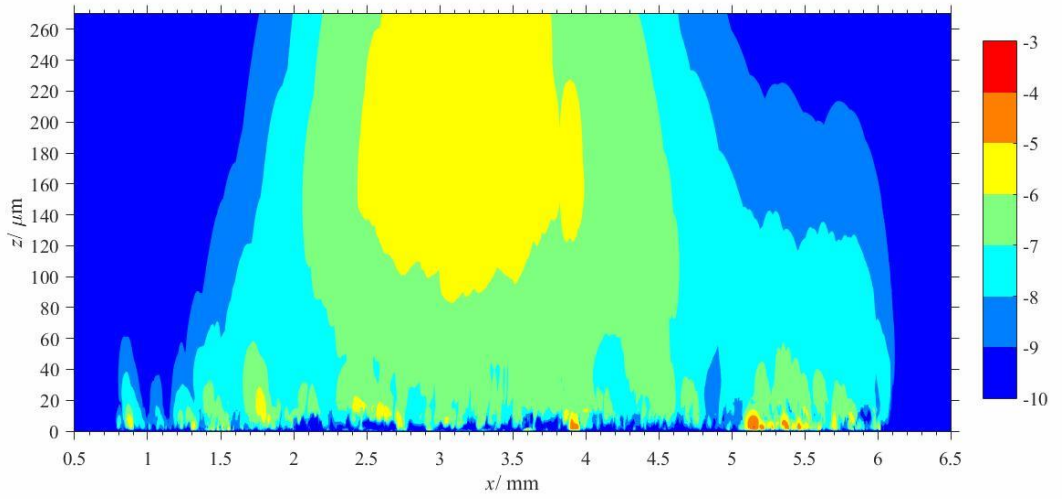
The values of the Dang Van and the Findley fatigue parameters greater than or equal to unity flags out the areas at which, most probably, fatigue processes will take place within 10^7 meshing cycles. The Fatemi and Socie model takes into account the cyclic behaviour of the load during the meshing cycle and the damage value plotted is the reciprocal of the number of probable load cycles to fatigue. To draw a comparison with the first two methods, the values of the probabilistic Fatemi and Socie fatigue life greater than 10^{-7} point out the fatigued nodes, where fatigue can be expected to occur in 10^7 meshing cycles.



a)

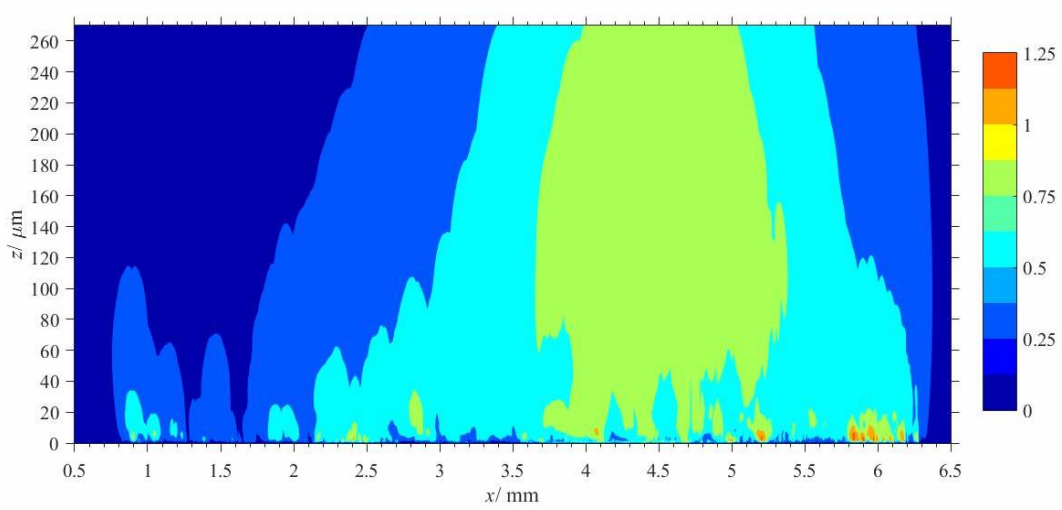


b)

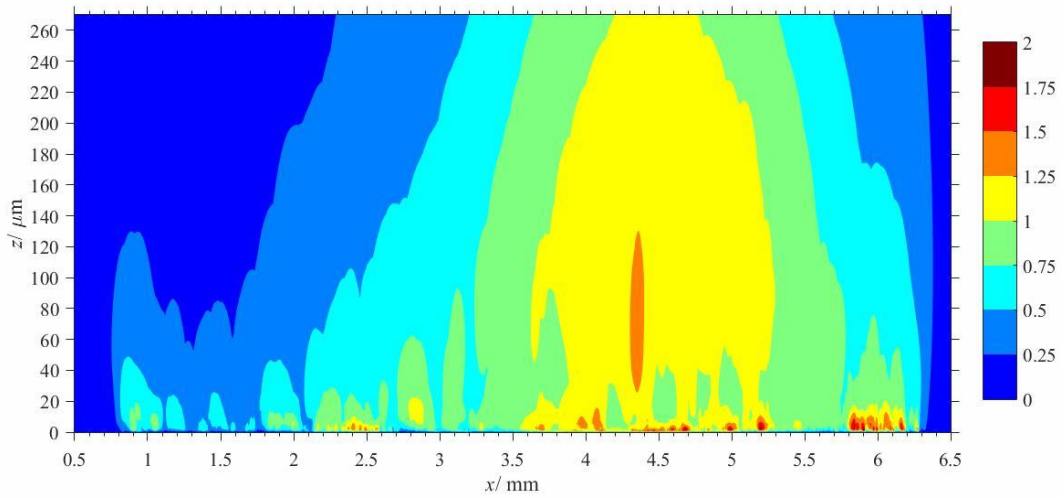


c)

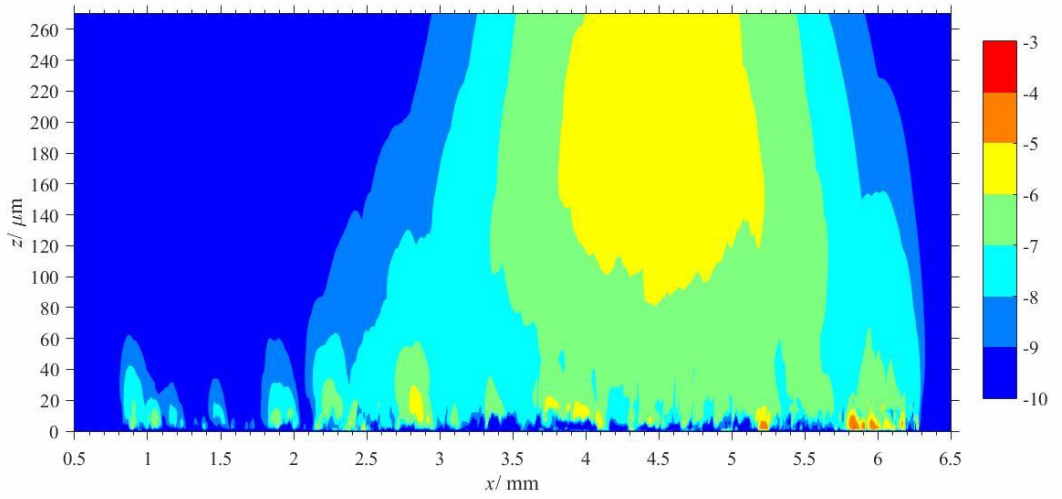
Figure 5.19 – Pinion: (a) Dang Van and (b) Findley fatigue parameters for 10^7 loading cycles and (c) Fatemi and Socie accumulated damage, 10^{-n}, indicating fatigue in 10^n cycles



a)



b)



c)

Figure 5.20 – Gear: (a) Dang Van and (b) Findley fatigue parameters for 10^7 loading cycles and (c) Fatemi and Socie accumulated damage, 10^{-n}, indicating fatigue in 10^n cycles

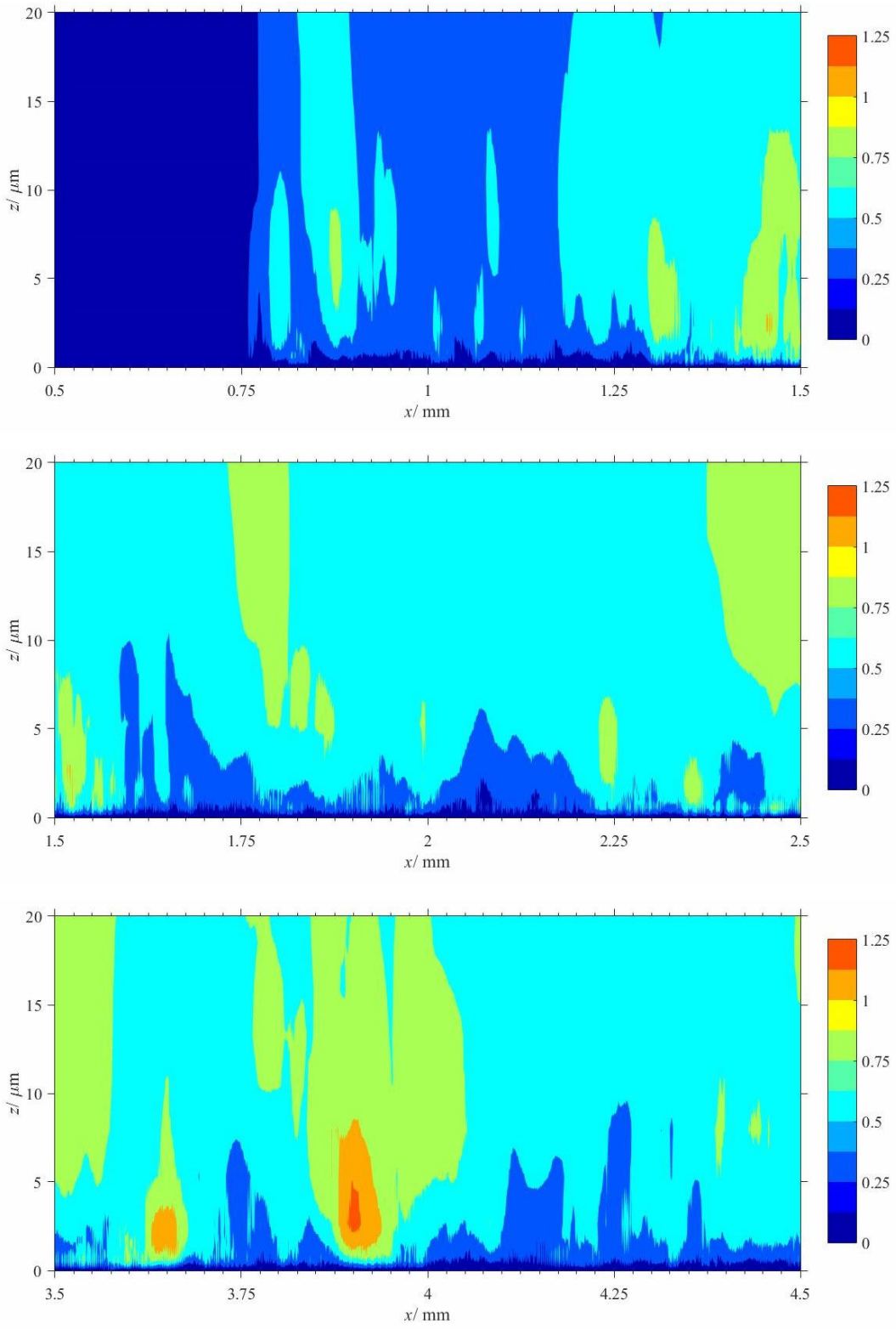


Figure 5.21 – Pinion: Dang Van fatigue parameters at 10^7 loading cycles for the close proximity to the pinion flank surface

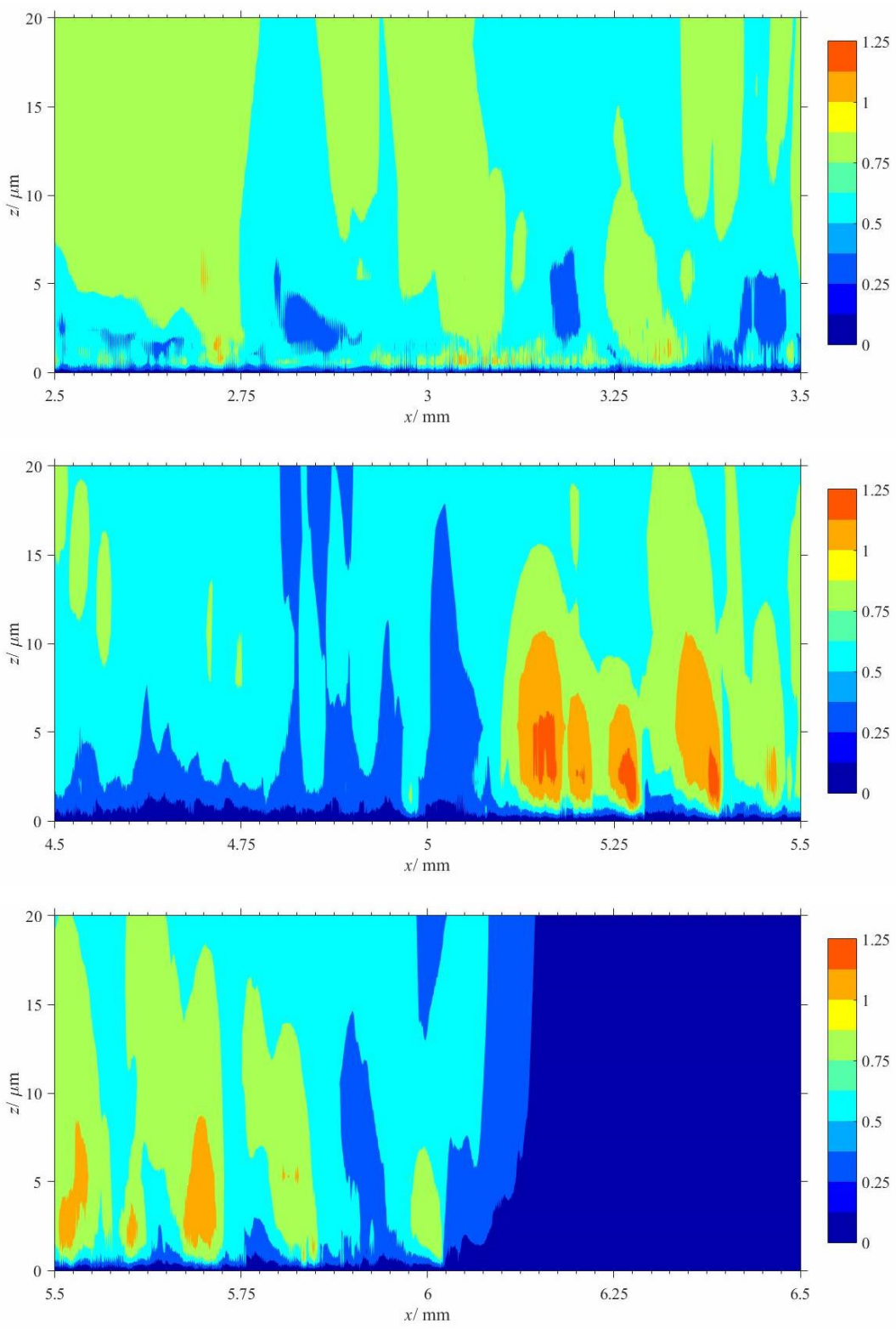


Figure 5.21 – Pinion: Dang Van fatigue parameters for 10^7 loading cycles at the close proximity to the pinion flank surface

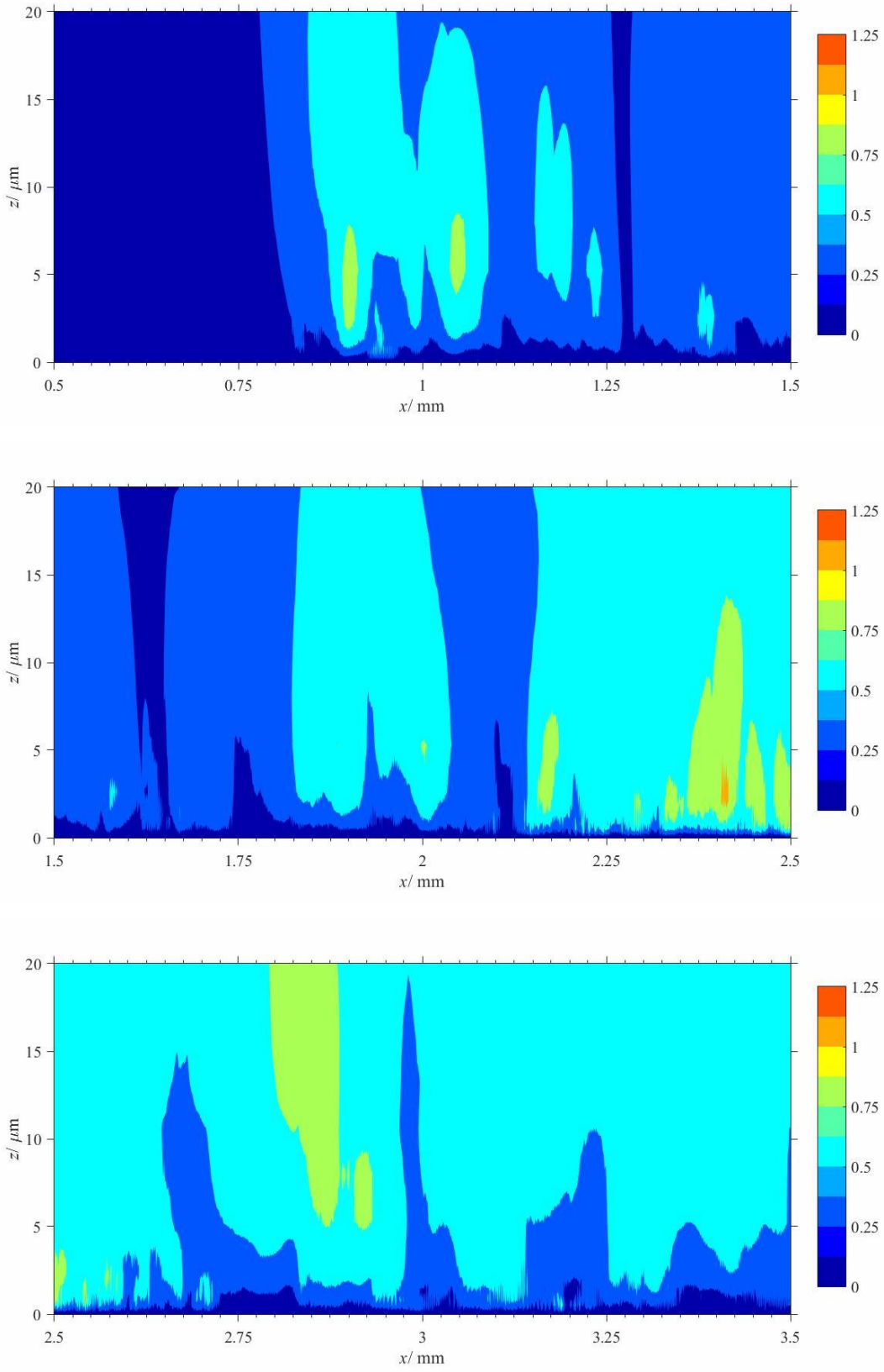


Figure 5.22 – Gear: Dang Van fatigue parameters for 10^7 loading cycles at the close proximity to the pinion flank surface

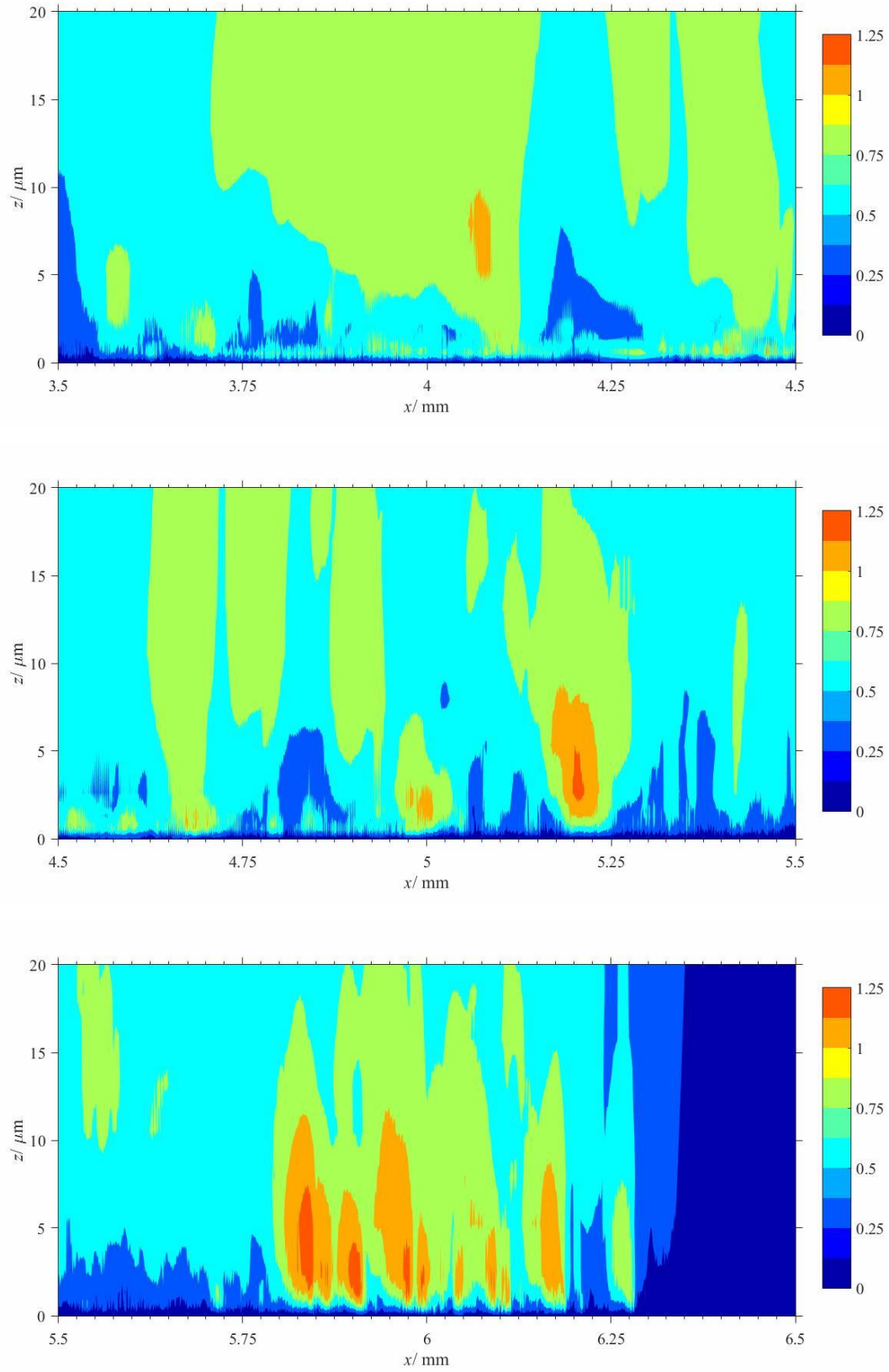


Figure 5.22 – Gear: Dang Van fatigue parameters for 10^7 loading cycles at the close proximity to the pinion flank surface

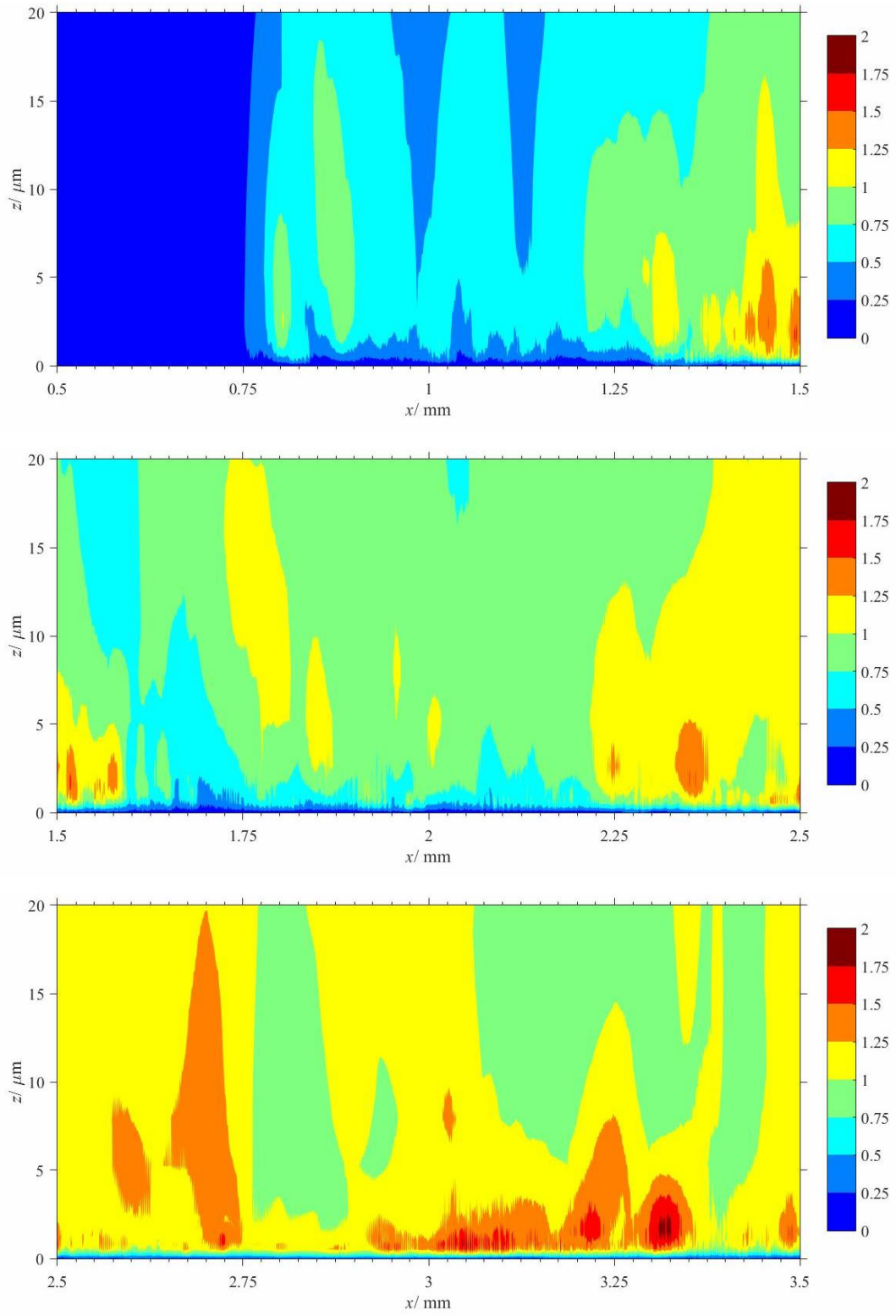


Figure 5.23 – Pinion: Findley fatigue parameters for 10^7 loading cycles at the close proximity to the pinion flank surface

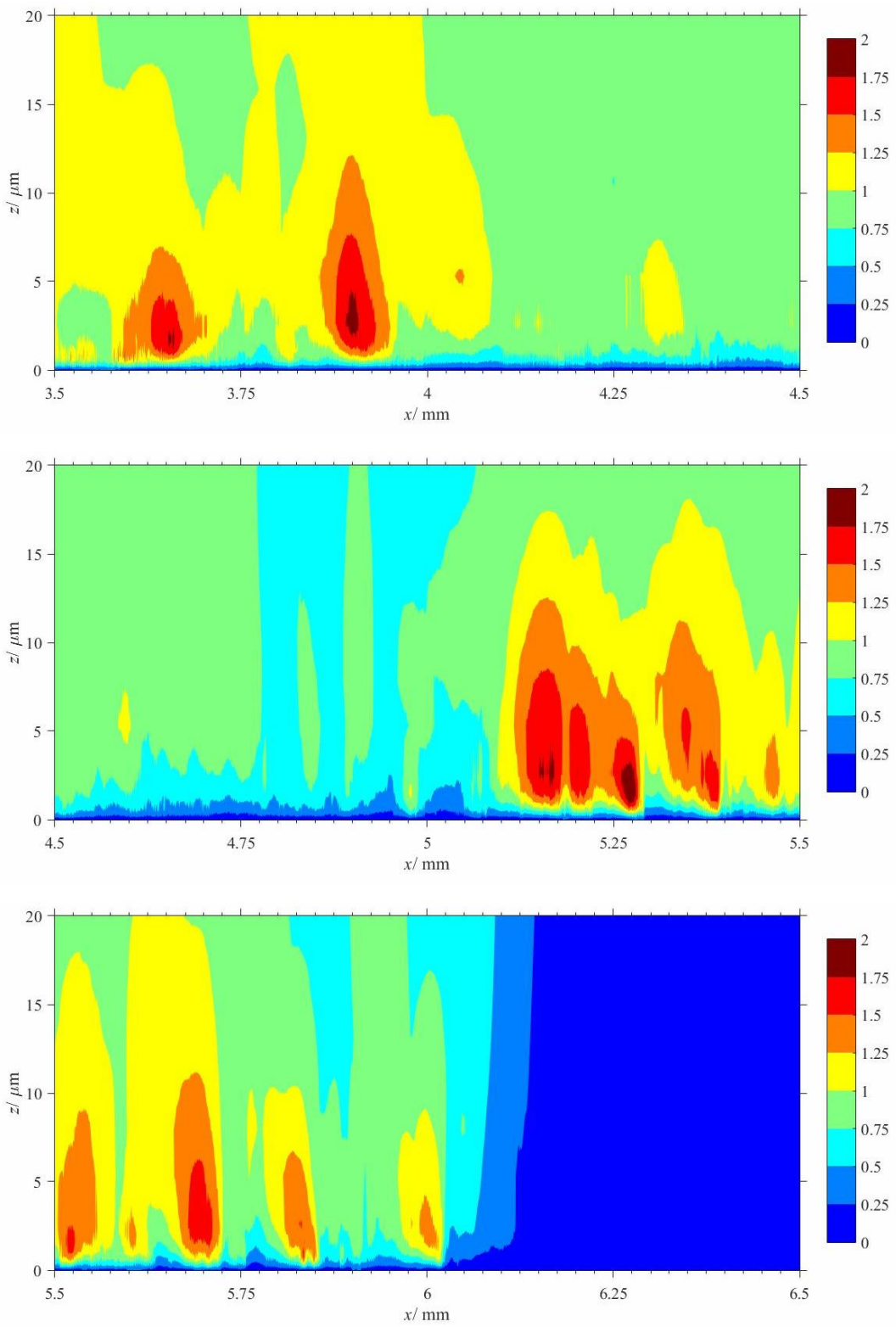


Figure 5.23 – Pinion: Findley fatigue parameters for 10^7 loading cycles at the close proximity to the pinion flank surface

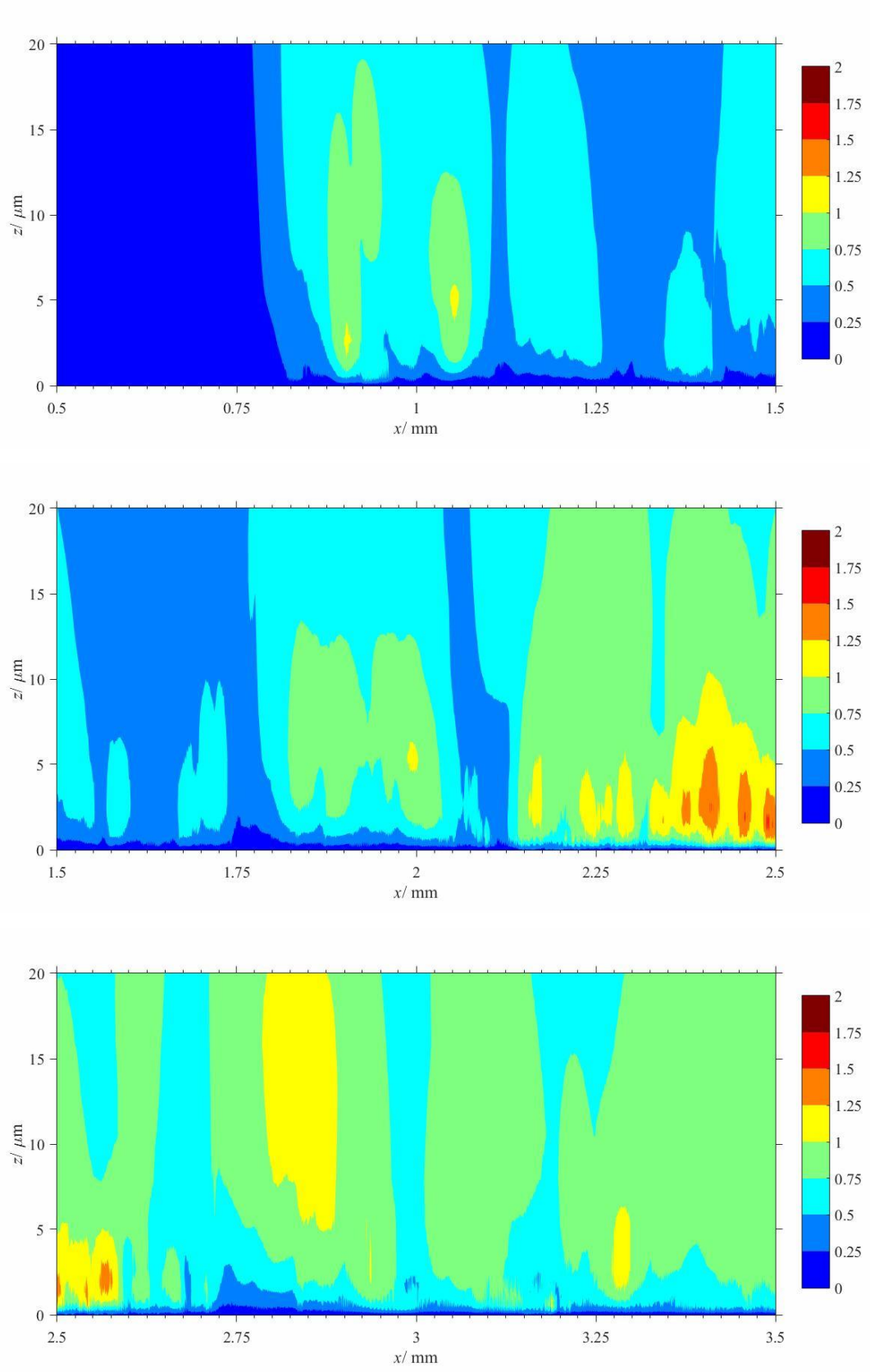


Figure 5.24 – Gear: Findley fatigue parameters for 10^7 loading cycles at the close proximity to the pinion flank surface

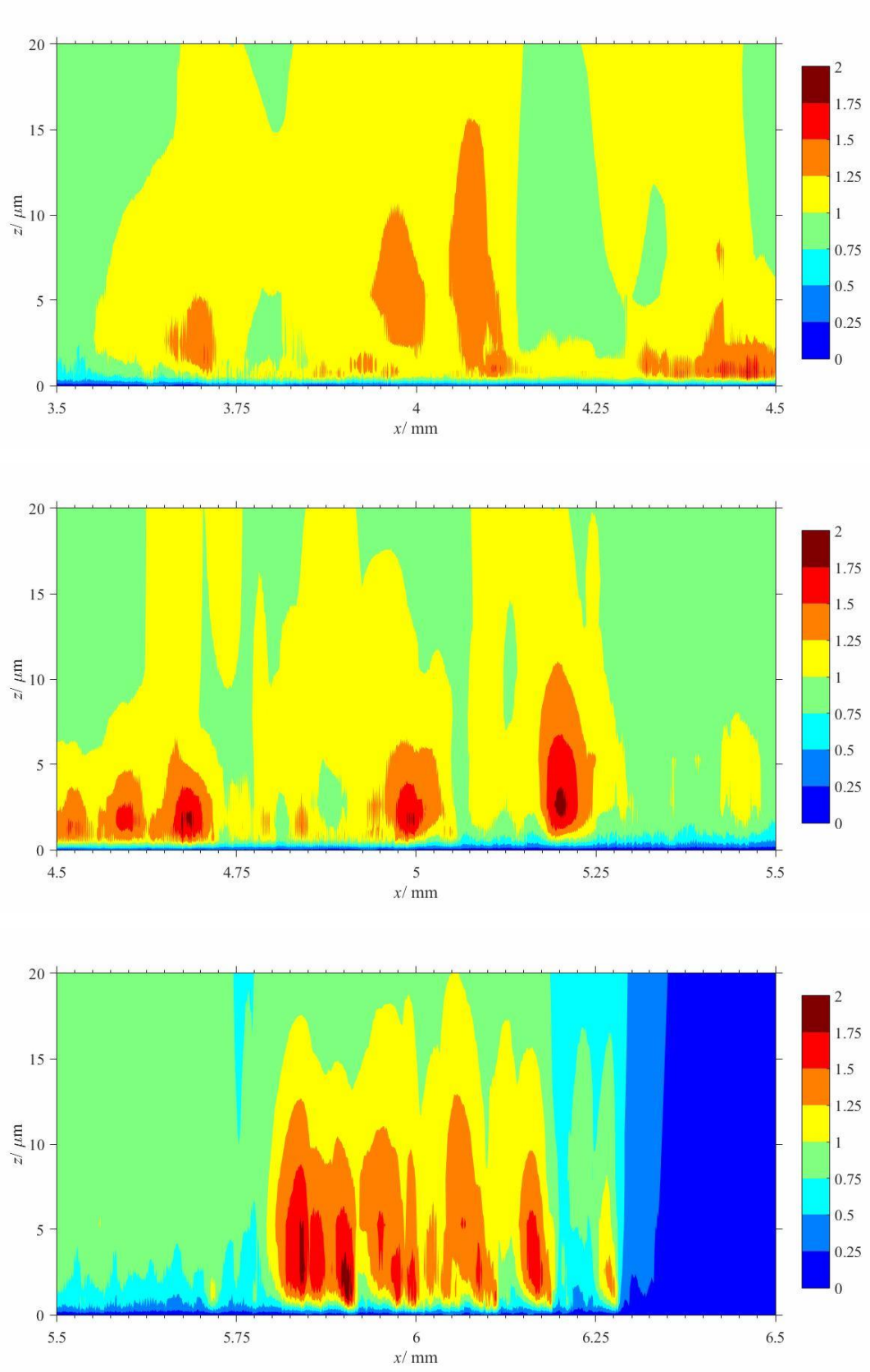


Figure 5.24 – Gear: Findley fatigue parameters at 10^7 loading cycles at the close proximity to the pinion flank surface

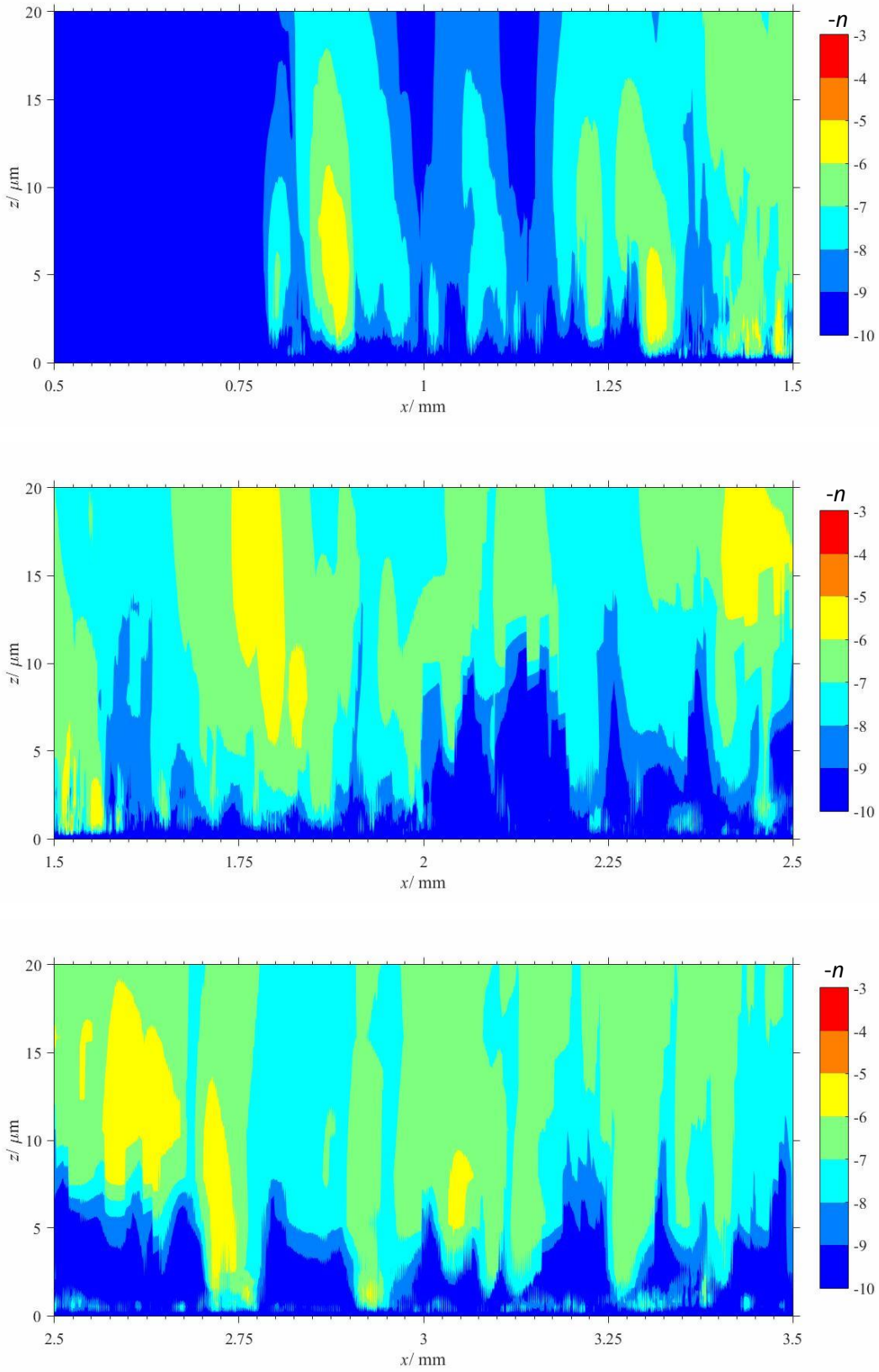


Figure 5.25 – Pinion: Fatemi and Socie accumulated damage, 10^{-n}, at the close proximity to the pinion flank surface indicating fatigue in 10^{n} cycles

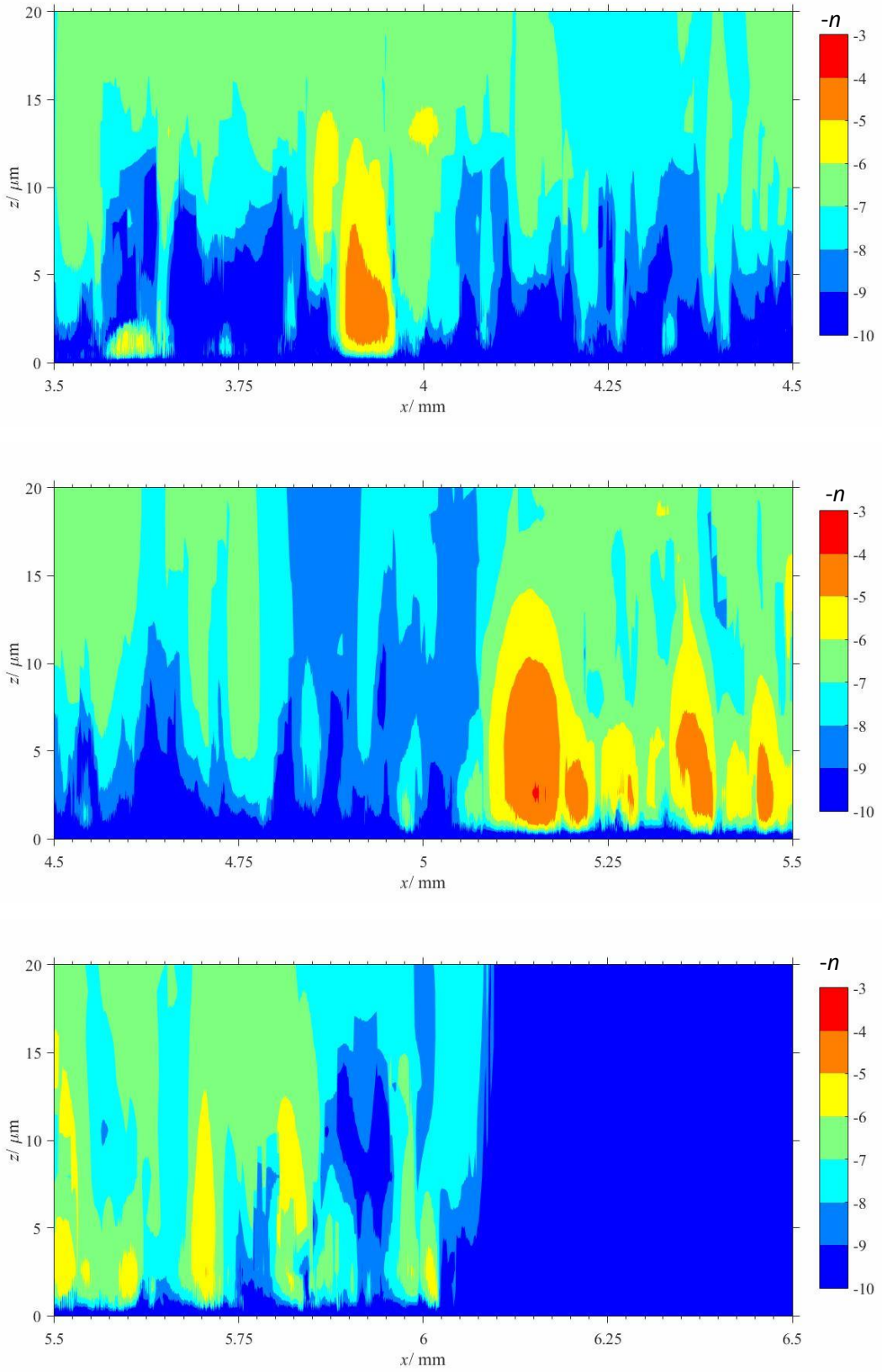


Figure 5.25 – Pinion: Fatemi and Socie accumulated damage, 10^{-n}, at the close proximity to the pinion flank surface indicating fatigue in 10^{n} cycles

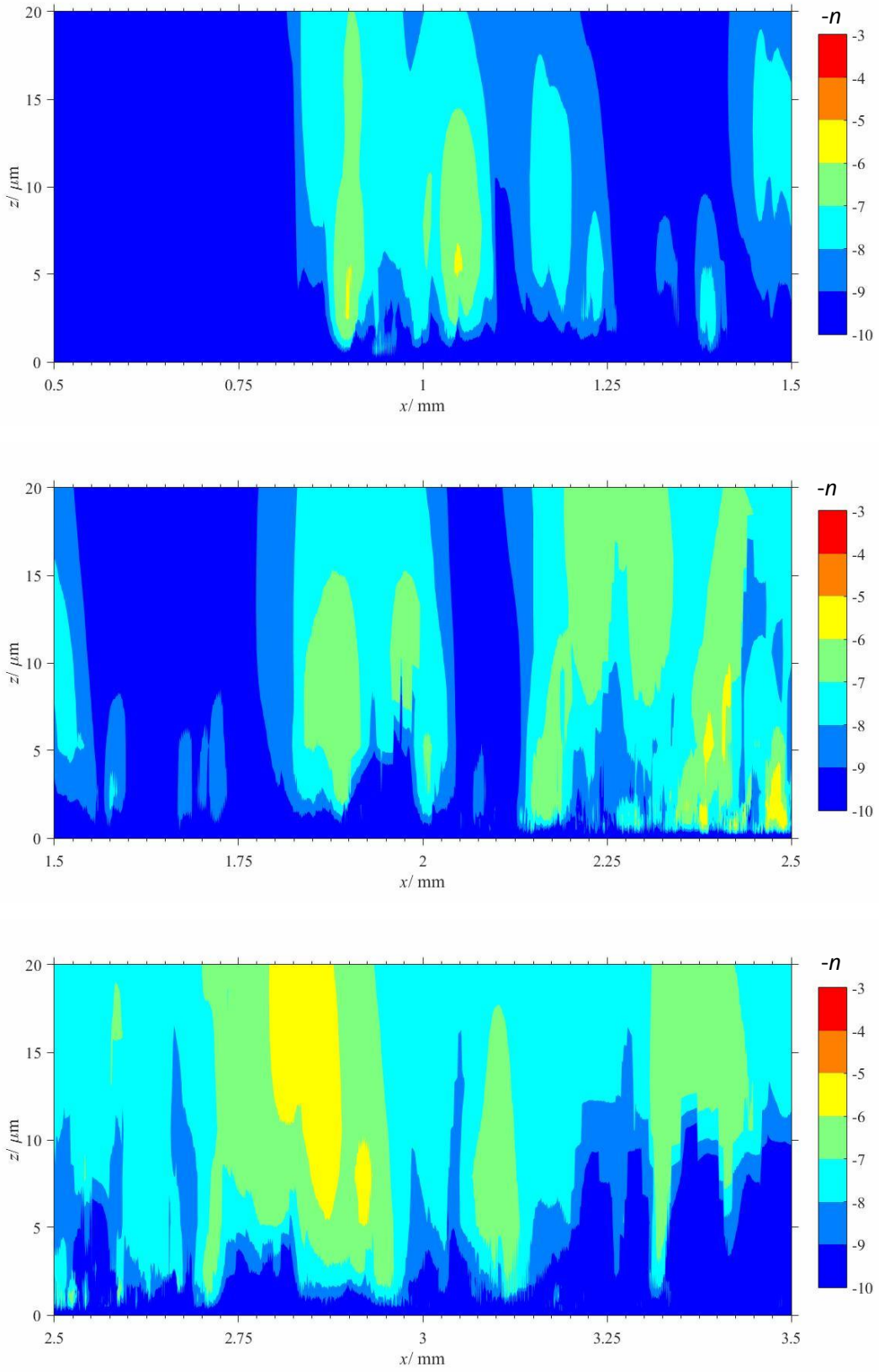


Figure 5.26 – Gear: Fatemi and Socie accumulated damage, 10^{-n}, at the close proximity to the gear flank surface indicating fatigue in 10^{n} cycles

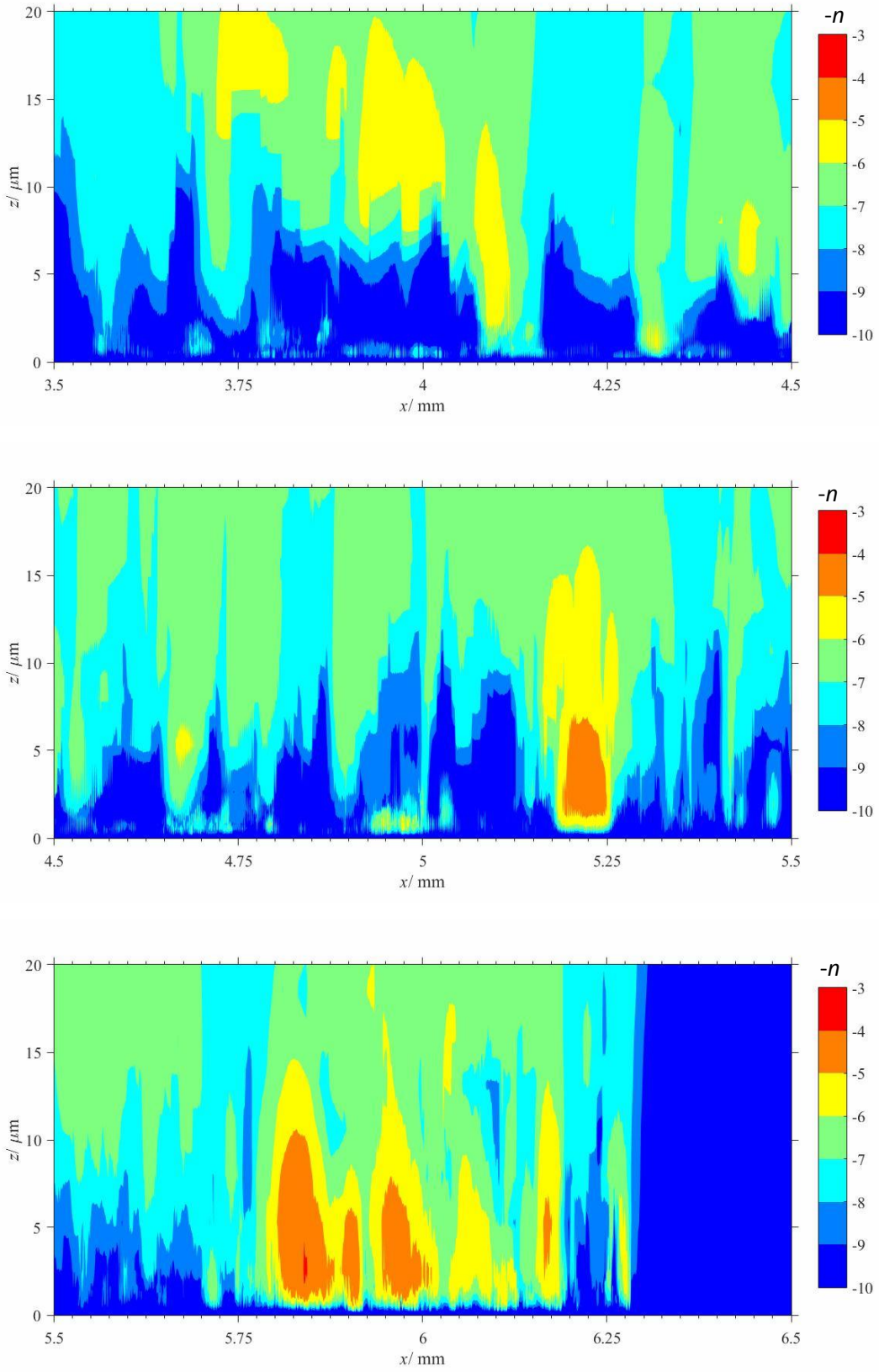


Figure 5.26 – Gear: Fatemi and Socie accumulated damage, 10^{-n}, at the close proximity to the gear flank surface indicating fatigue in 10^{n} cycles

Two processes different by nature can be seen in Figures 5.19 and 5.20. According to Johnson (1985), the maximum principal shear stress at the Hertz contact occurs at $z = 0.78a$ (for the current smooth case it is 206 μm at the pitch point), therefore, the high values of the fatigue parameters shown as solid green and yellow contours for all three fatigue models deep down under the surface are primarily driven by the load and almost unaffected by the surface roughness. It is not surprising, considering the pressure distribution in Figures 5.14 and 5.15 that in this they are not very different from the Hertzian pressure function, with occasional peaks due to the incorporated roughness of the surfaces.

The surface roughness contribution to the fatigue is the most apparent at the positions of high sliding: the tip of the pinion, x [5; 6] mm, and the root of the gear, x [5.75; 6.25] mm. The Dang Van and the Fateni and Socie methods predict a mild wear of the middle part of the pinion and the gear flanks, with one particular fatiguing feature at $x = 3.9$ and $x = 5.2$ at the pinion and gear surfaces respectively. The Findley model predicts a severely more aggressive wear, because it is highly dependent on the shear stress, therefore, it is more responsive to the roughness effect. At the same time, it only considers the worst loading conditions of the meshing cycle, disregarding the stress history pattern, the duration of the peak load, softening and hardening processes. Hence, there is a fatigue feature shown in orange in Figures 5.19b and 5.20b in the close subsurface (z [20; 140] μm, $x = 3$ mm for the pinion and $x = 4.35$ mm for the gear).

The NASA fatigue test results in NASA/TM–2005-213956/ARL-TR-3126 include photos of worn gear teeth for all the experiments replicated numerically in this chapter. The photo of the worn gear flank tip to root for the MIL-L-7808J test case is shown in Figure 5.27. There is a clear similarity between the predicted fatigue life parameters pattern shown in Figures 5.19 and the worn profile in Figure 5.27.

Figure 5.27 – Worn surface of the test gear tip to root
(NASA/TM–2005-213956/ARL-TR-3126)

Chapter 6

The results of the complete analysis of the gear meshing cycle: The Design Unit case

6.1. Introduction

Due to some restrictions and limitations, the analysis produced of the NASA gears presented in Chapter 5 does not allow the roughness treatment procedure to be presented. At the same time, the fatigue-life prediction results shown in Chapter 5 are significantly different to the typical results previously obtained by the Cardiff Tribology group for representative steady state conditions. Therefore, it was necessary to use a different gear test case to explain the crucial steps to make sure that the surface roughness does not have any redundant artefacts and is properly located on the involute profiles so that the EHL analysis resembles the real meshing cycle behaviour. The operating conditions were also chosen so that they differed from the NASA test setup to demonstrate a different perspective on the results.

This Chapter describes the surface roughness treatment approach, presents the EHL and fatigue-life analyses results of test gears provided by the Design Unit, Newcastle University (DU), and discusses the differences between them and the NASA case.

6.2. The test set up parameters

Unfortunately, the profiles provided by NASA were pre-processed and cropped in such a way that the tooth root and tip features are not included. This creates significant uncertainties in aligning the rough profiles on the involute curves. Another complication in that analysis was the inability to obtain roughness of a counter surface of the particular gear flank. This was particularly important due to the equal number of teeth of the pinion and the gear, i.e. each pinion flank is brought into contact with one particular gear flank and that gear flank only. The setup of the test made it impossible to confirm that the pinion and gear profiles provided for the analysis were actually in contact with each other during the test and as such the probability of that being the case is 0.035. This is probably the reason for the numerical difficulties experienced in running the EHL analysis as each meshing flank pair will run in to accommodate each other.

Therefore, it was decided to run a different analysis to include all steps of the gear meshing cycle analysis as well as investigating the influence of the roughness and the operating conditions on the results. The Cardiff Tribology group collaborates closely with the DU studying lubrication of helical gears. A pair of test gears taken from a micropitting test was available for measuring and had no traces of pitting damage. The gear design parameters as well as the lubricant properties are provided in Table 6.1.

Table 6.1 – DU test gears design parameters

Geometrical parameters		Pinion	Gear
Module / mm	m_n	4.500	
	m_s	4.776	
Number of teeth	n	33	34
Pressure angle / °	ψ_n	20.000	
	ψ_s	21.121	
Outside radius / mm	r_t	83.305	85.695
Base radius / mm	r_b	73.512	75.739
Material properties of AMS 6260			
Young's modulus / GPa	E	206.850	
Poisson's ratio	υ	0.300	
Coefficient of friction (dry contact)	μ	0.1	
Lubricant OEP 80			
Absolute viscosity at 350 K / N-s/m^2	η_0	0.0155	
Pressure-viscosity coefficient at 350 K / m^2/N	α	1.55×10^{-8}	

The smooth surface analysis was conducted to obtain the clearance between surfaces at the first mesh node, which determines the load as was explained in Section 5.3 for the NASA gears. The target load function over the meshing cycle was scaled in order to obtain about 1.3 GPa pitch-line maximum Hertz stress. The converged operating load plotted against the target one and the clearance between surfaces at the first mesh node are shown in Figure 6.1 and 6.2 respectively. The half-Hertzian contact dimension at the maximum load is 0.311 mm. Due to numerical difficulties in converging the load for low values, because even small adjustment of the clearance between two surfaces in contact can cause a situation when the surfaces are too far away to maintain the contact, the lowest value of the target load function, which allowed a converged solution to be obtained was determined as 100 kN/m by a trial and error method. The modification of the target load function is an automatic process and the result of it is clear in the shape of the converged load plot, shown in red in Figure 6.1, at the beginning and the end of meshing cycle.

6.3. Results of the EHL smooth contact analysis

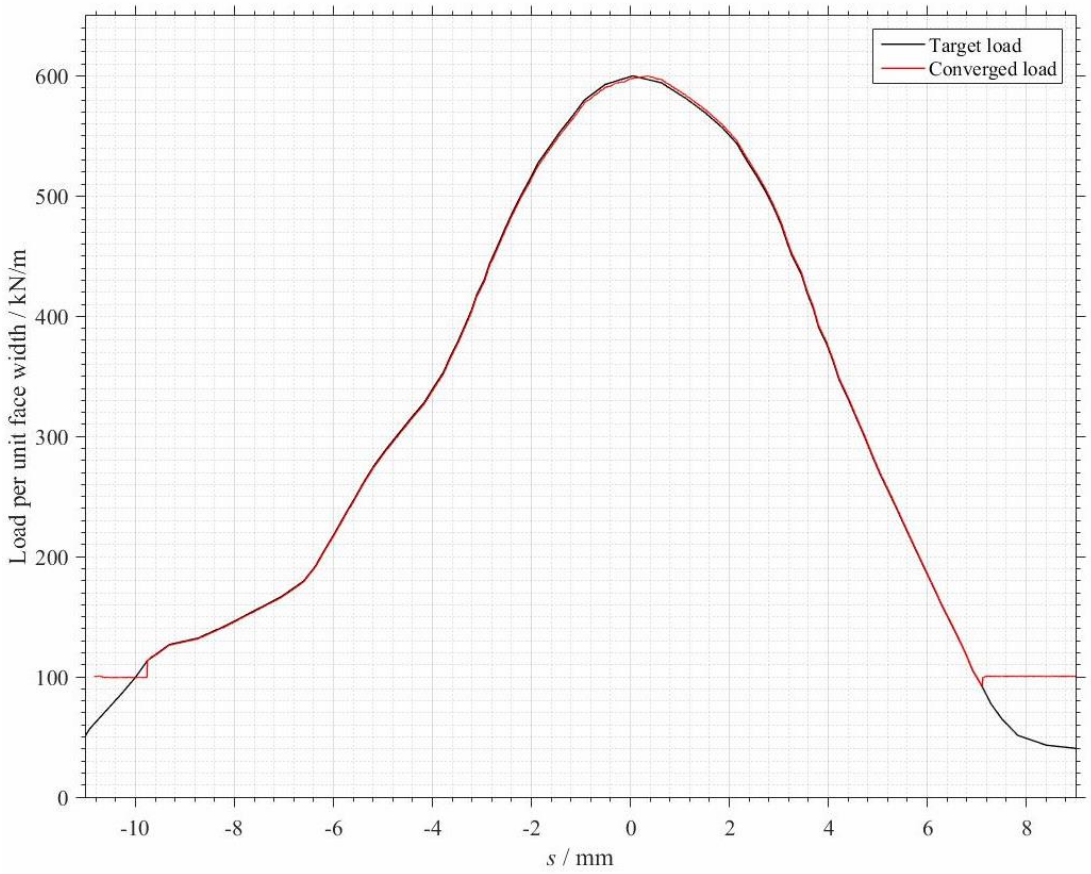


Figure 6.1 – Target (black) and converged (red) loads

The divergence of the converged clearance function shown in red in Figure 6.2 at the beginning and the end of meshing cycle due to the modification of the load function was resolved by manual extrapolation of the data obtained for the non-modified load region. The resultant modified clearance ready to use for the rough surface analysis is plotted in black in Figure 6.2.

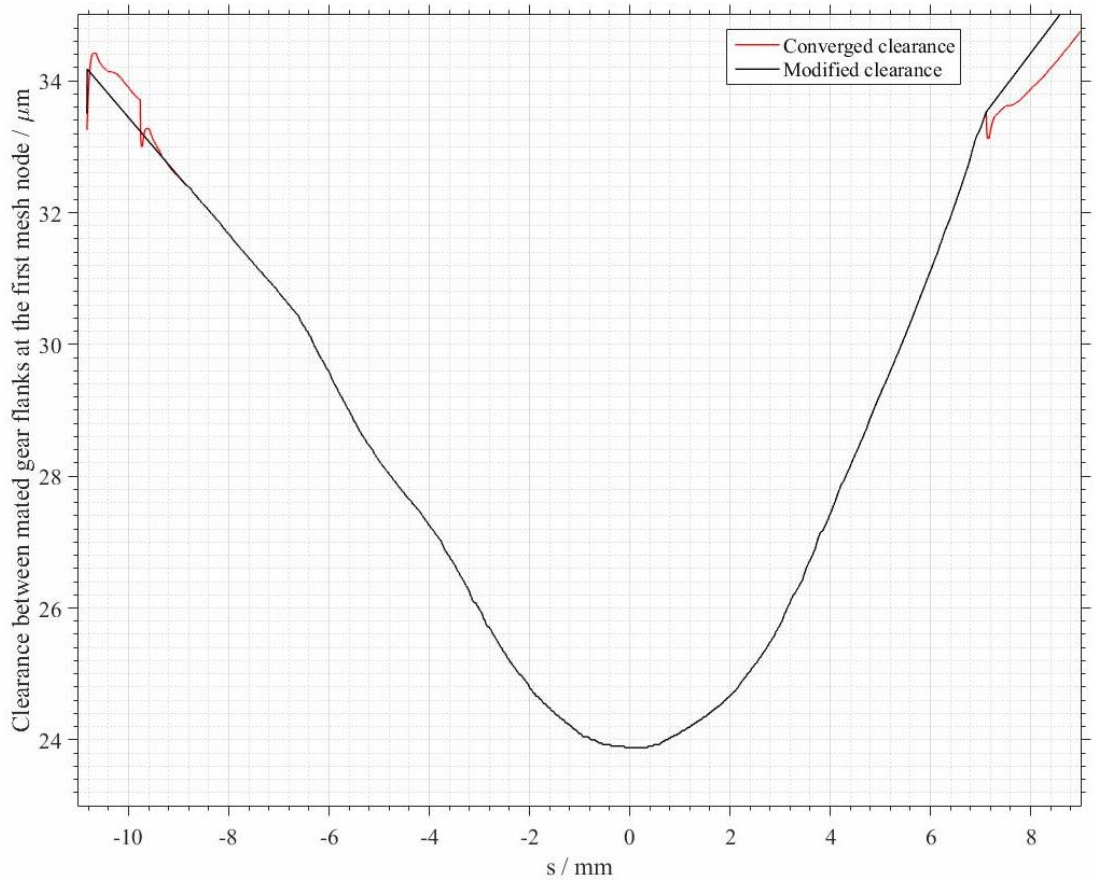


Figure 6.2 – Converged (red) and manually modified for the rough surface analysis (black) clearance between mated gear flanks at the first mesh node

6.4. Surface roughness measurements and treatment

The processing of the roughness profiles differs significantly from the one employed for the NASA case, which used a standard process to remove waviness and filter the data. The helical gears were mounted in a purpose made jig that inclined the gear axis to the vertical by the tooth helix angle so that the tooth surface to be measured was nominally horizontal. The gear was then rotated so that the profilometer could take a root to tip profile measurement within the 1 mm height range of the instrument. The photograph in Figure 6.3 shows the pinion mounted in the jig at the measuring position.

The measuring objective is to take root to tip profiles where the tip of the gear can be clearly identified in the profile as a reference point. If the profilometer stylus moves outside the measuring range during a traverse then the instrument automatically abandons the profile measurement. This makes it difficult to achieve the measuring objective of including the gear tip as the profile must take the stylus contact beyond the gear tip and must end before the stylus loses contact with the surface. The profilometer was fitted with an accessory that can be adjusted to provide a stop support for the stylus beam to solve this difficulty. The accessory prevents the stylus from moving below the measurement range of the instrument causing it to lose contact with the surface if the surface becomes too low during the traverse. So with the accessory installed the height becomes constant when the stylus loses contact with the tip.

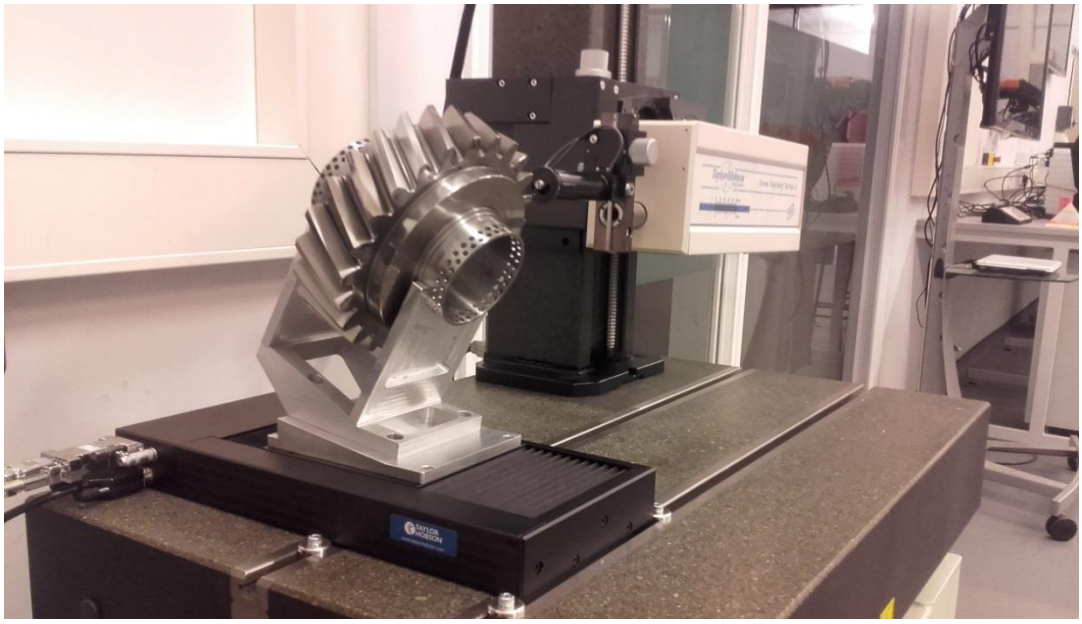


Figure 6.3 – Photograph of gear mounted in measuring jig on profilometer y-stage.

A typical root to tip profile is illustrated in Figure 6.4 which consists of 40,000 measured heights giving the form of the tooth and its roughness measurement. Point T, arrowed, is the tip of the tooth and point C is the highest point of the form. The profile heights are thus measured perpendicular to involute tangent at C. Coordinates of points T and C are obtained by curve fitting the roughness data in the vicinity of the points and this gives an accurate measurement of distance TC. This coordinate information is then used to determine the radius of C from the gear axis and thus locate C on the involute whose base circle radius is known. Tooth tip relief must be accounted

for in this procedure but this is known from the gear manufacturing drawings and standard profile error measurements taken by a Klingelnberg P65 gear checker which confirm the relief magnitude at the tip.

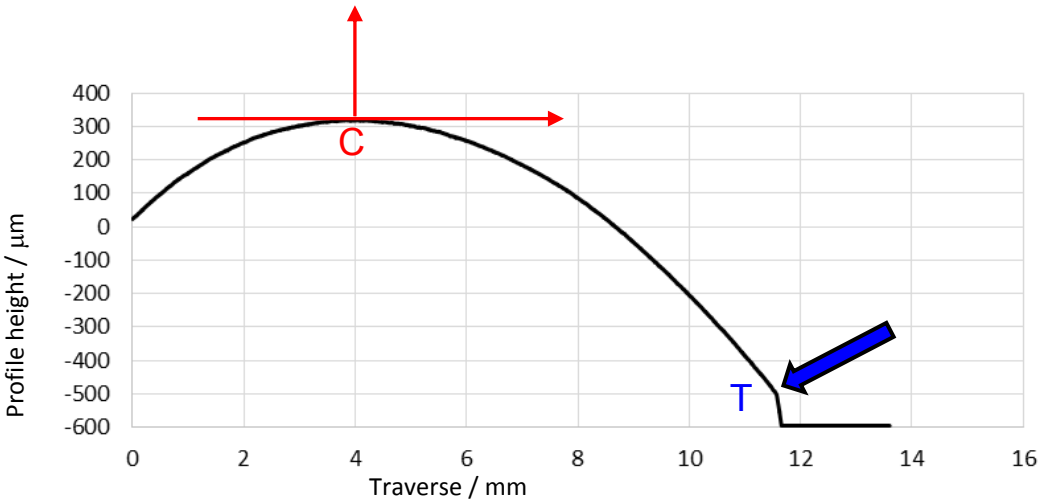


Figure 6.4 – Root to tip gear profile measurement taken from test gear.

The analytic fit to the involute is used to remove the form so that the profile gives the deviation from the involute. An example of this is shown in Figure 6.5 which shows the measured profile.

The profiles shown in Figure 6.5 are expressed in terms of roll angle which can be obtained in terms of the profile traverse coordinate from the involute geometry and the position of point C on that involute. The *SAP* and *STR* positions are defined in terms of roll angle in the gear geometry specification and it is clear that the specified *STR* position corresponds to the location where the tip relief becomes apparent in the involute form removed profile, h_{diff}.

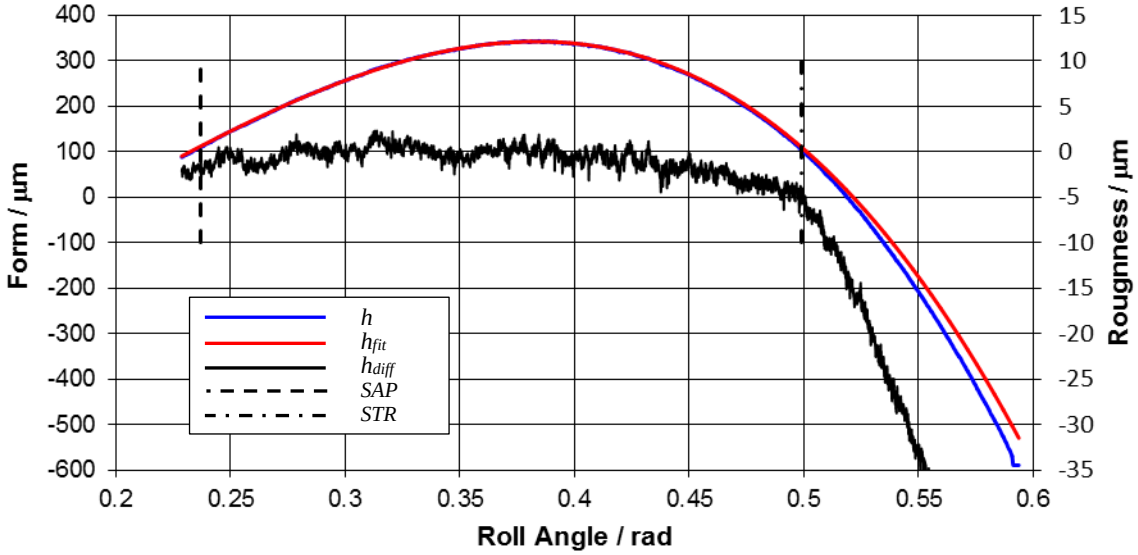


Figure 6.5 – Profile measurement, h, involute fit, h_{fit}, and profile with fit removed, h_{diff}. Also shown are start positions of active profile, *SAP*, and tip relief, *STR*.

Figures 6.6 and 6.7 show the results of profile measurement for the pinion and wheel tooth measurements used for the EHL analysis, respectively, with all plots presented in terms of profile traverse position. Frames (a). show the raw profile data measurements, and frames (b) show the corresponding profiles after the fitted involute form is removed. To obtain the final roughness profile for the EHL analysis two further steps are carried out. Firstly the remaining form of the true involute portion is removed within the TalyMap roughness software using a 2nd or 4th order polynomial. An ISO standard Gaussian filter is then applied to remove the waviness and produce the roughness profile using a standard cut-off of 0.25 mm, which is of the order of the Hertzian contact dimension, a. The resulting roughness profiles, shown in frames (c), are used for the EHL analysis and they do not show any traces of pitting damage.

The roughness of the DU gear flanks after running-in is more aggressive than the NASA values: For the DU pinion roughness profile shown in Figure 6.6 the roughness parameter values of Ra=0.376 µm, and Rq=0.505 µm are obtained, which compare with values of Ra=0.214 µm and Rq=0.279 µm for the the NASA pinion. For the wheel roughness profile the DU gear has Ra=0.424 µm and Rq=0.563 µm while the NASA gear has Ra=0.157 µm and Rq=0.213 µm. The minimum smooth surface film thicknesses are 0.23 µm and 0.48 µm giving minimum lambda ratios of the NASA and the DU cases, assessed over complete meshing cycle, of 0.655 and 0.635 respectively.

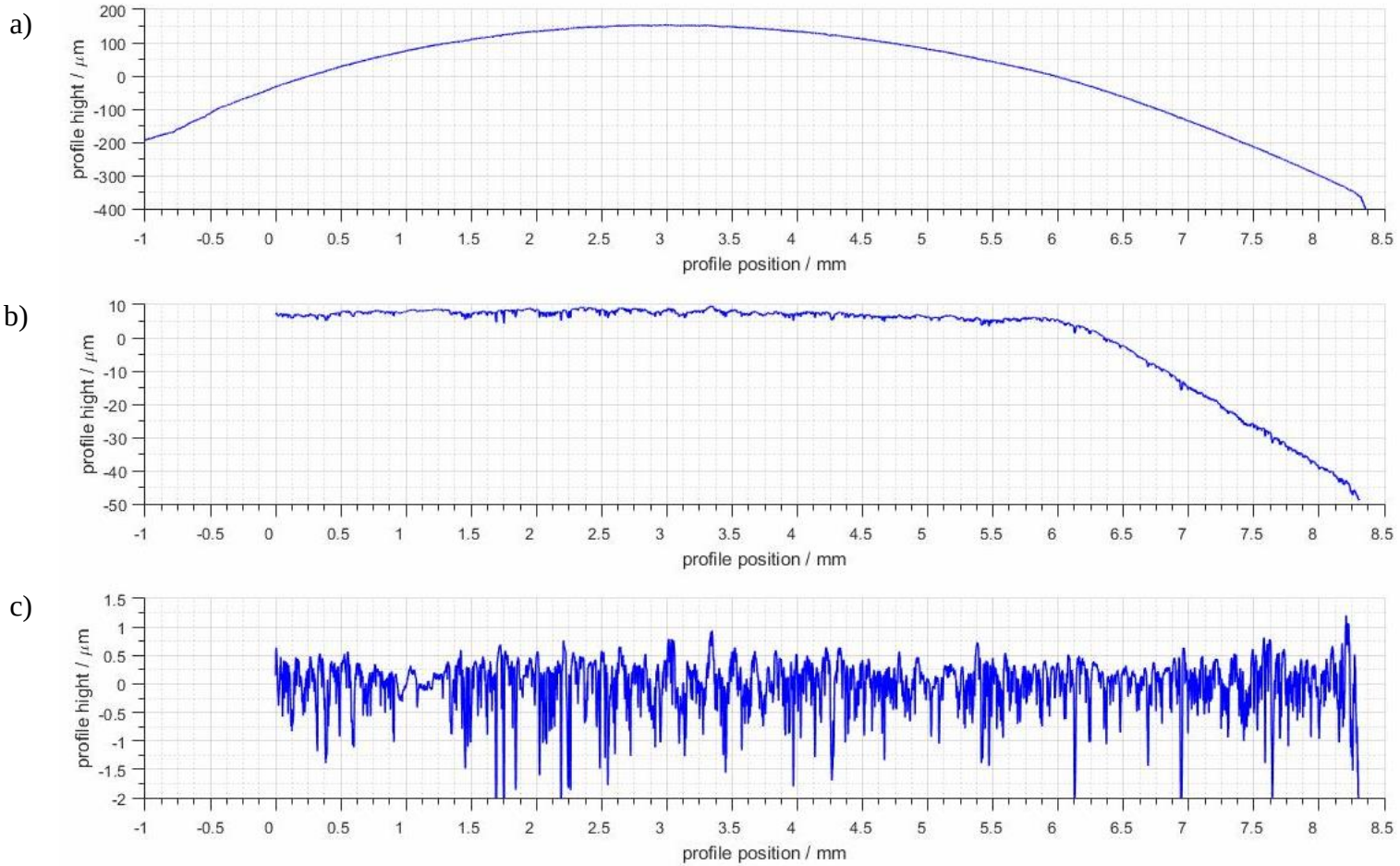


Figure 6.6 – Pinion surface roughness: (a) raw data; (b) form removed;
(c) form removed and Gaussian filter with 0.25 mm cutoff applied

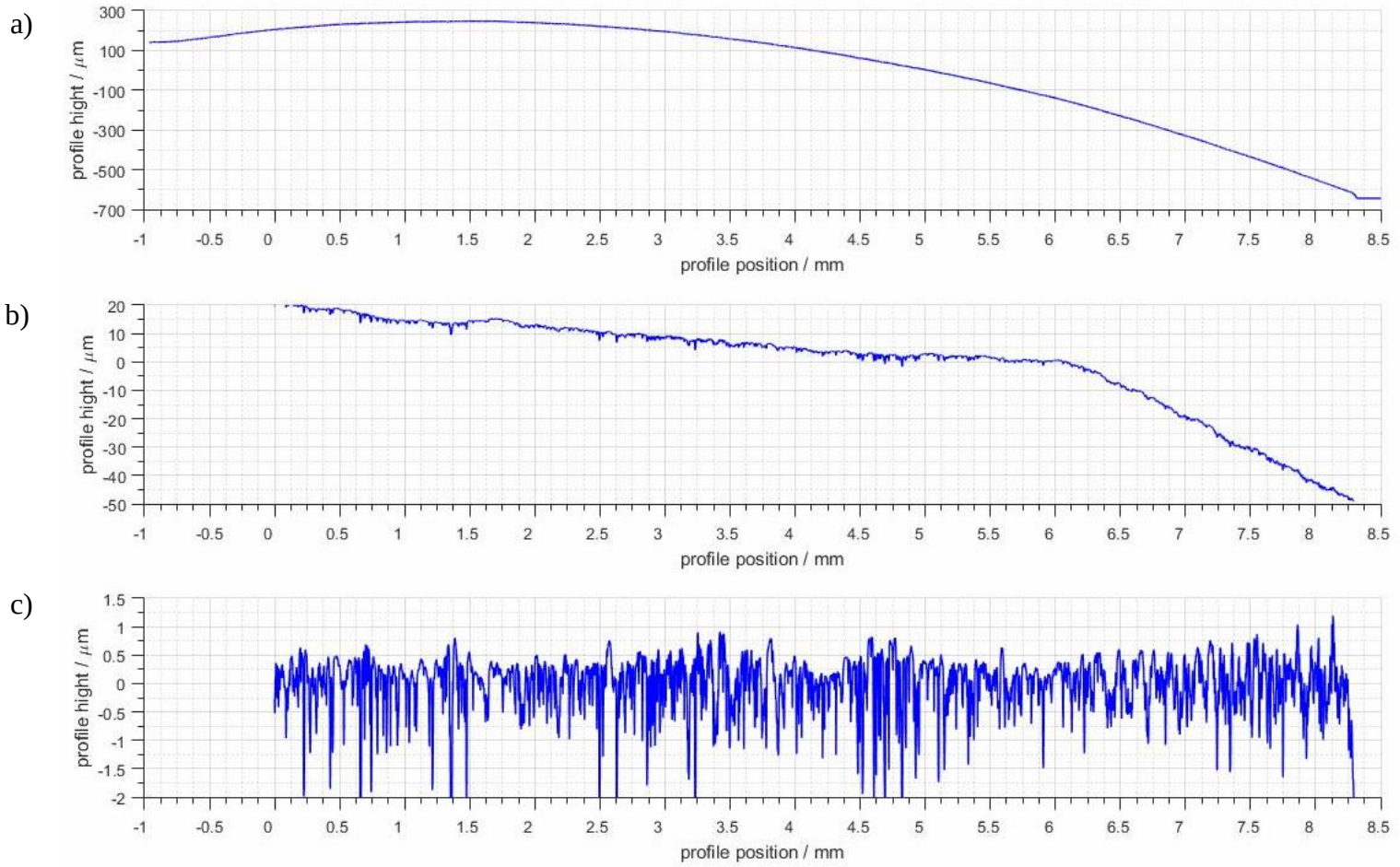


Figure 6.7 – Wheel surface roughness: (a) raw data; (b) form removed;
(c) form removed and Gaussian filter with 0.25 mm cutoff applied

Positioning the roughness profile accurately on the involute for the analysis is possible as the position of point C is known. The surfaces are aligned by specifying a point on the roughness profile and providing the corresponding roll-angle of the involute gear flank at that particular point. The easiest way to find the roll angle is by using equation 2.7. The *y* coordinate is the radius from the centre of the gear to the point of interest:

$$y = r_b \left(\sin \gamma - \gamma \cdot \cos \gamma \right) \tag{6.1}$$

The profiles are set to be in the correct positions at the pitch point where there is no sliding and the accurate relative location of asperities is particularly important.

The graphic representations of Equation 6.1 for the pinion and the gear are plotted in Figure 6.8 as solid blue and red lines respectively. The horizontal dashed lines show the gear tip radii. The abscissa gives the roll angle of the involute in degrees and the ordinate specifies the radius of the circle, which defines the position of the profile from the centre of the gear in mm. The labels provide information about the arc length of the involute from the base circle and the roll angle in mm and degrees respectively. These data are obtained for each integer radius value between the base and the tip circles but can be evaluated for any radius. The values are also provided for the gear tip circles.

The last point of the roughness profile corresponds to the tip, therefore, knowing the length and spacing of the roughness profile and the arc length of the involute at the tip point, the index of the point of interest can be determined as

$$index = N_{rough} - \frac{al_{tip} - al_{pi}}{spacing} \tag{6.2}$$

Where N_{rough} is the size of the rough profile mesh; al_{tip} and al_{pi} are the arc lengths of the involute profile at the tip point and the point of interest respectively; *spacing* is the spacing of the rough profile mesh.

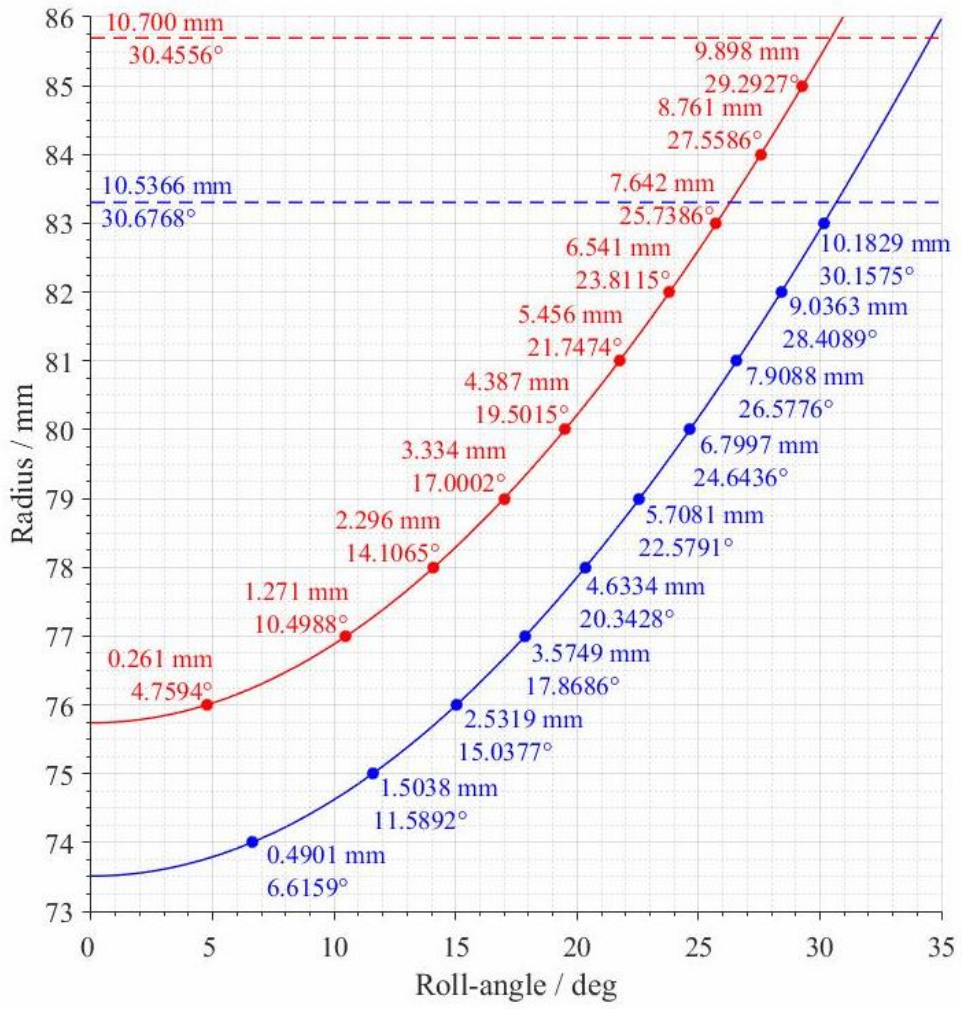


Figure 6.8 – Pinion (blue) and Gear (red) alignment parameters

However, it is almost impossible to identify the exact tip point, which causes some imperfections in aligning profiles. It is not very important at the beginning and the end of the meshing cycle due to the high sliding velocity and low load, but at positions near the pitch point the load is high and the surfaces move with almost equal velocities, i.e. a feature of the pinion flank interacts only with a particular feature of the gear flank.

6.5. Results of the rough surface contact analysis

The EHL analysis result when the contact is at the pitch point is shown in Figure 6.9 where the pressure, film thickness and offset rough surfaces are shown, as before. There are three areas that can cause significant numerical problems if the surfaces are not aligned properly. The high and lengthy asperity at the pinion surface in zone A, shown in red in Figure 6.6, fits perfectly into the deep valley feature of the gear flank at x/a = -0.21. In zone B, shown in amber, two massive asperities barely miss each other and interact with an appropriate valley of the counter surface at x/a = 0.56. In green zone C the high sharp gear surface asperity at x/a = -1 follows the deep valley feature at the pinion surface, but does not interfere with it even later on during the meshing cycle. In this situation even a 5 μm misalignment can bring the features described above into a contact, which causes a rapid rise of pressure beyond the elastic limit and numerical difficulties with converging a candidate solution for the current timestep as a result. Keeping in mind that the gears of interest were previously used for some tests and operated under the similar loads, therefore, all plastic processes took place long before the roughness profiles were measured and the contact can be expected to be purely elastic.

Due to the higher roughness the contact itself is clearly split into four subcontacts, characterized by the low pressure at the inlet zone and high hydrodynamic pressure at the outlet area.

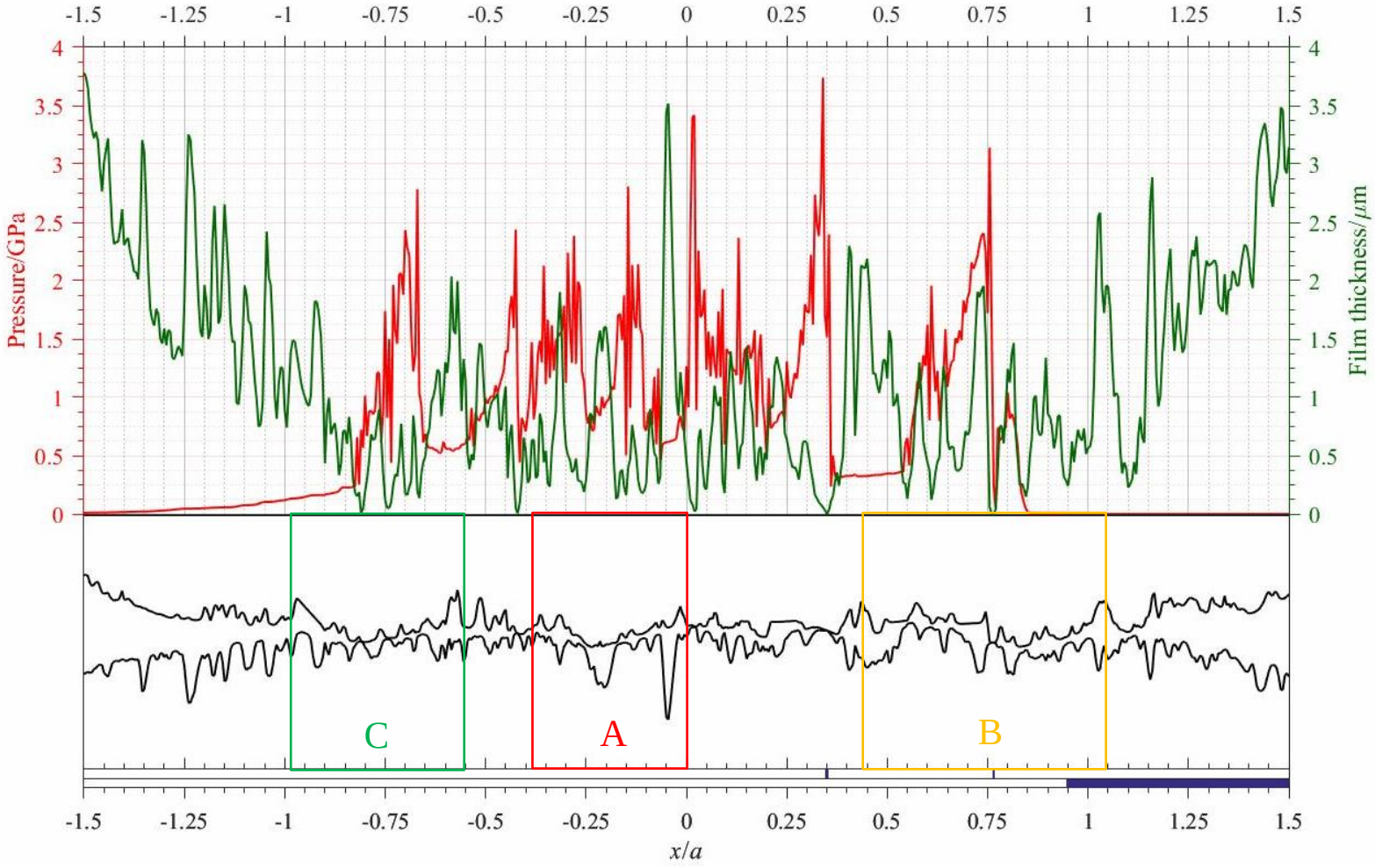


Figure 6.9 – Pressure distribution (red), film thickness (green), offsets of the surfaces (black) and regime indicator (blue) at pitch point

The contact for the DU gear set is more aggressive than the NASA one, which can be seen in Figure 6.10 that shows the number of nodes in a dry contact at each timestep of the analysis and in Figure 6.11 which presents the load variation over the meshing cycle.

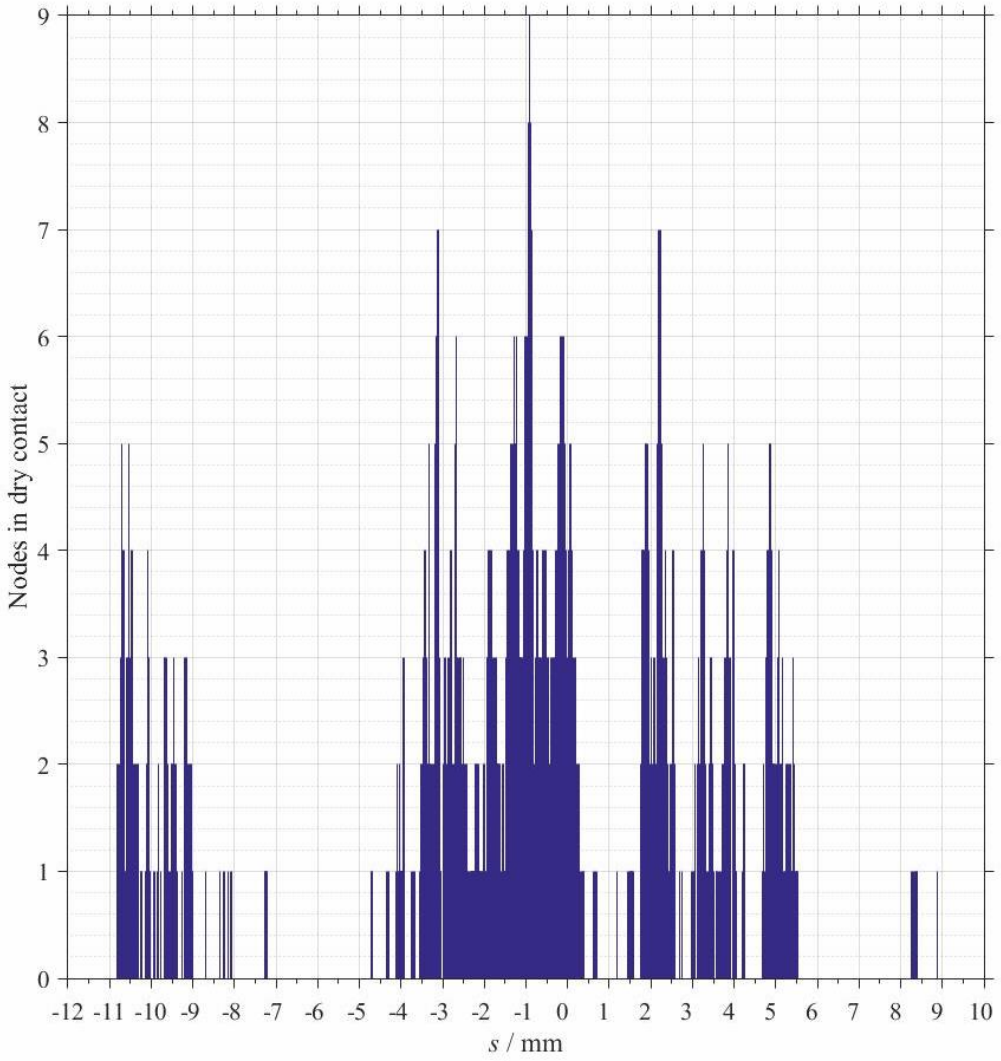


Figure 6.10 – Number of nodes in dry contact at each timestep of the analysis

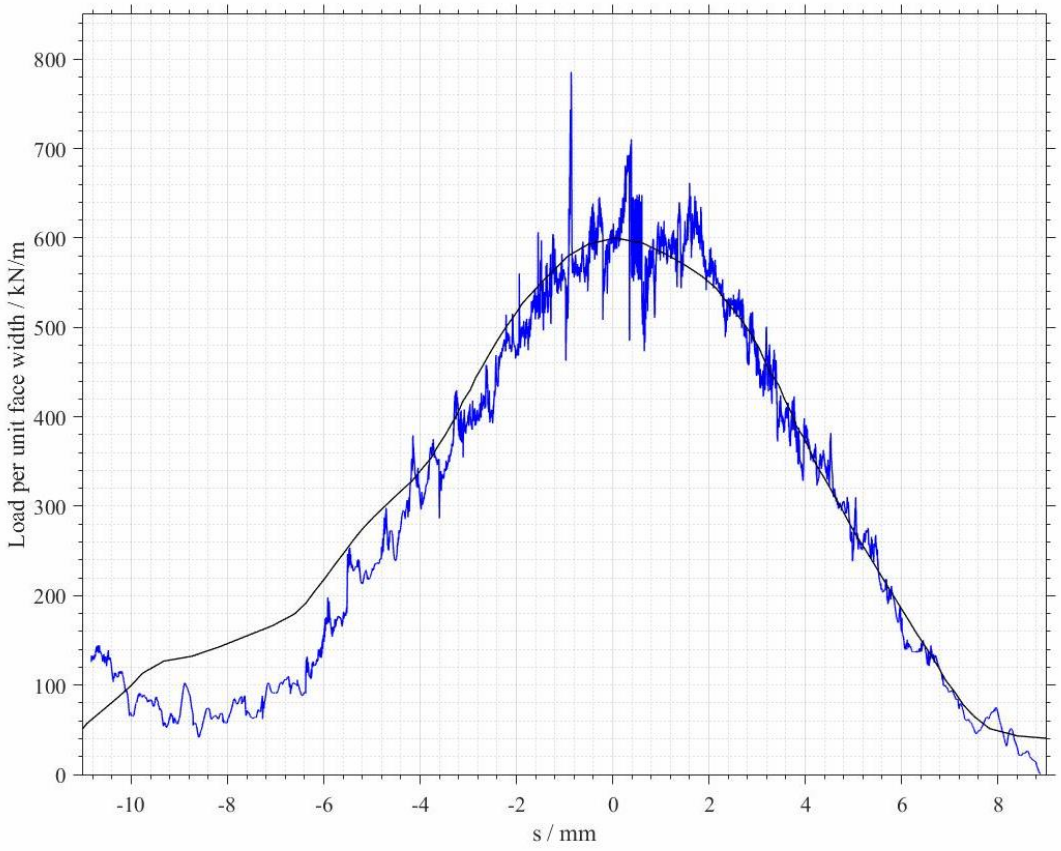


Figure 6.11 – Target (black) and rough surface contact (blue) load over the meshing cycle

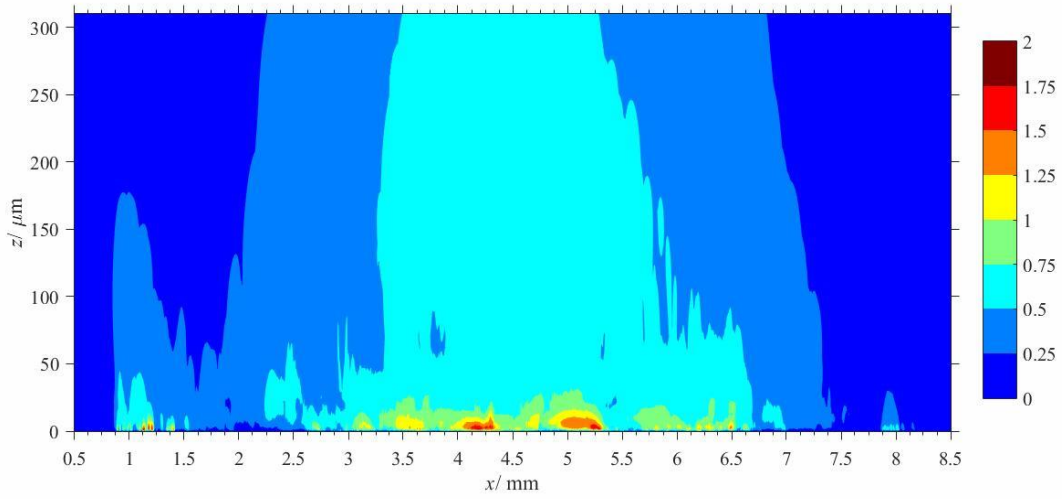
The fatigue analysis results for the two gear flanks are presented in Figures 6.12 to 6.19 adopting the same format and approach as was adopted for the results in Chapter 5. Figures 6.12 and 6.13 show the overall results for the pinion and the gear surfaces for the Dang Van fatigue parameter, the Findley fatigue parameter and the Fatemi and Socie accumulated damage.

The contour colours for the plots reflect the magnitude range of the parameter being plotted so that in comparing with the magnitude range of the corresponding figures, 5.19 and 5.20, for the NASA gears where roughness is lower it must be noted that the colours may represent different values. Comparing the core values at the position of maximum principal shear, the Dang Van fatigue parameters are at similar levels, the Findley fatigue parameter is 25% higher in the NASA gears results, and the Fatemi

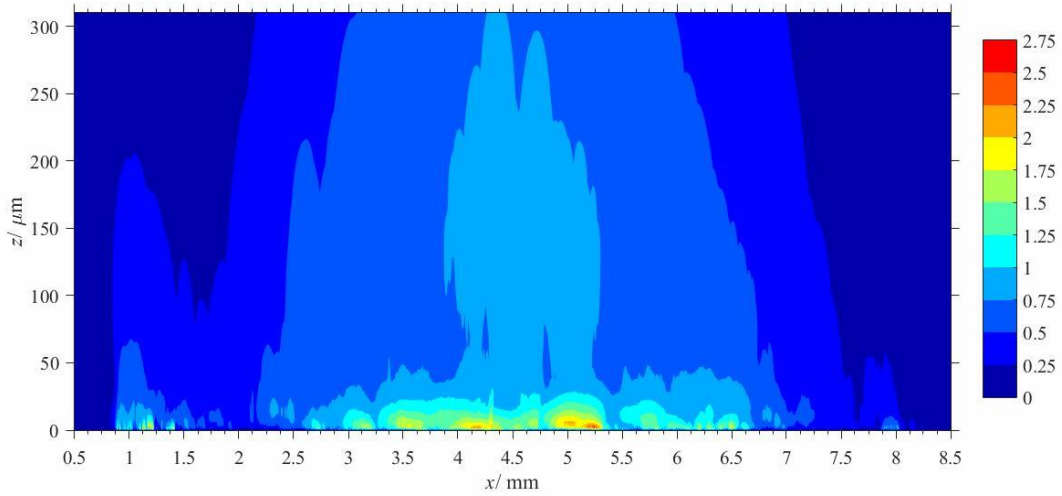
and Socie damage values are of similar order. The peak values occur near the surface and in this case the Dang Van and Findley fatigue parameters are approximately 50% higher for the DU gears. The Fatemi and Socie damage levels are again of similar order.

The near surface difference can be seen more clearly in the more detailed larger scale plots. These confirm the levels of difference given in discussing Figures 6.12 and 6.13, above. It is also apparent that there are larger areas which are subject to the highest plotted values in each case, for the DU gears. In particular the Fatemi and Socie peak damage is higher in the DU gears although in the same order of magnitude range.

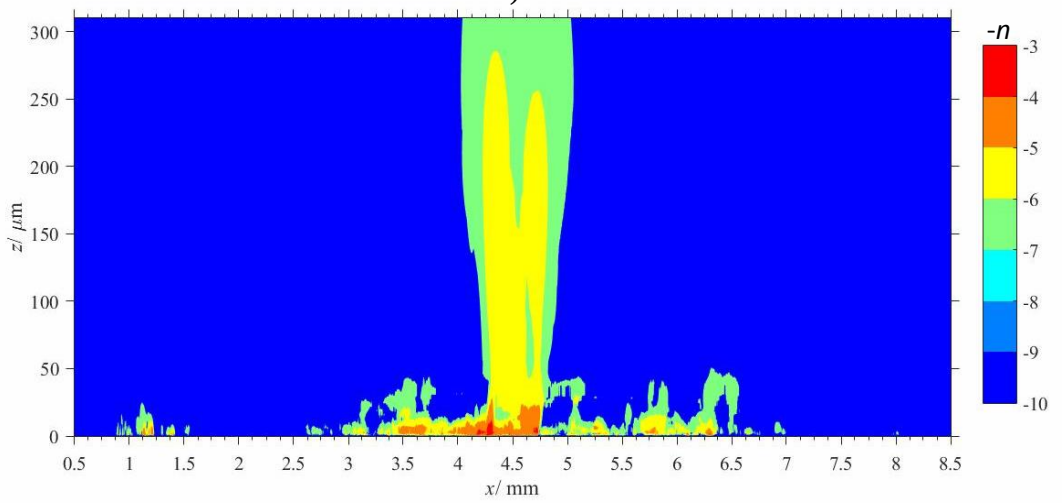
In summary, the core values for the fatigue calculation are higher for the NASA test case which operates at a peak Hertzian pressure of 2 GPa compared to 1.3 GPa for the Design Unit gears. However, the higher surface roughness for the Design Unit gears leads to more aggressive asperity loading and increases the probability of fatigue in the near surface layer in spite of the significantly lower Hertzian contact pressure.



a)

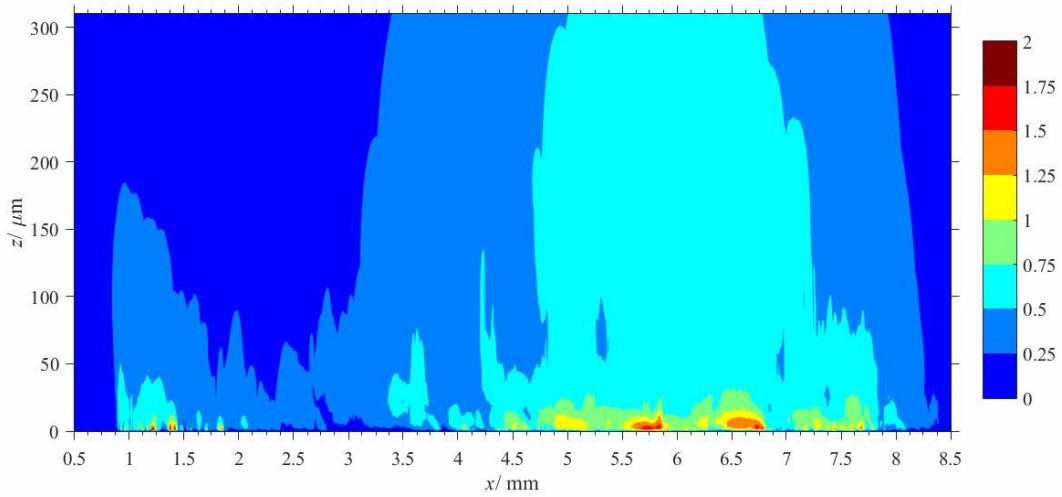


b)

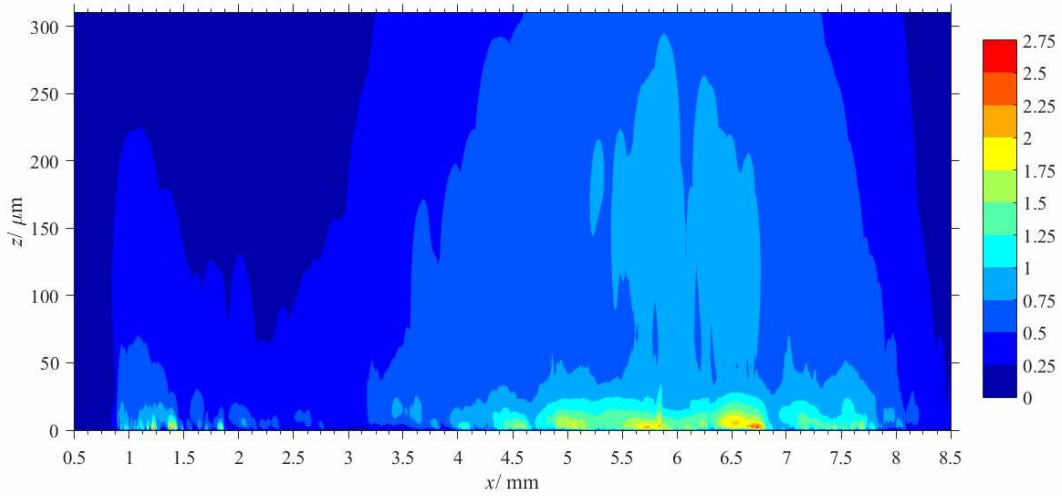


c)

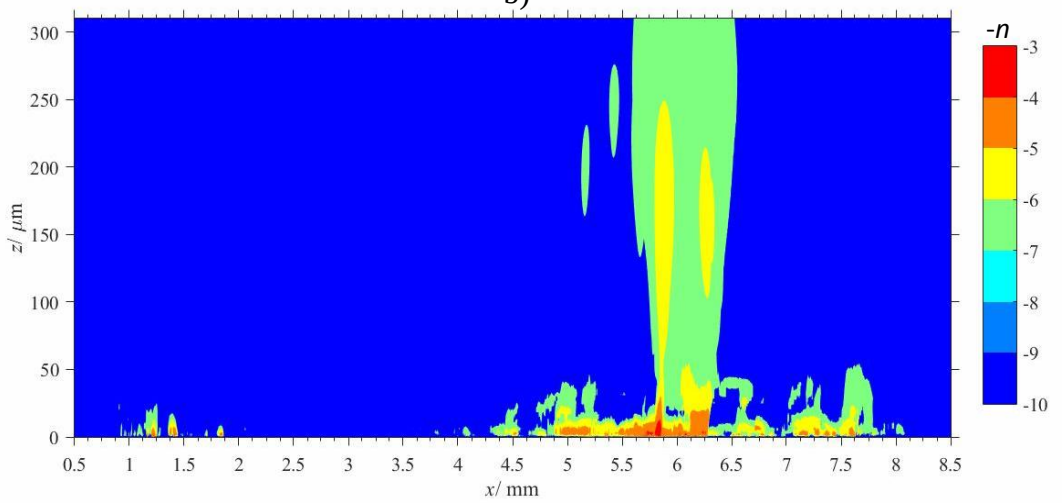
Figure 6.12 – Pinion: (a) Dang Van and (b) Findley fatigue parameters for 10^7 loading cycles and (c) Fatemi and Socie accumulated damage, 10^{-n}, indicating fatigue in 10^n cycles



a)



b)



c)

Figure 6.13 – Gear: (a) Dang Van and (b) Findley fatigue parameters for 10^7 loading cycles and (c) Fatemi and Socie accumulated damage, 10^{-n}, indicating fatigue in 10^n cycles

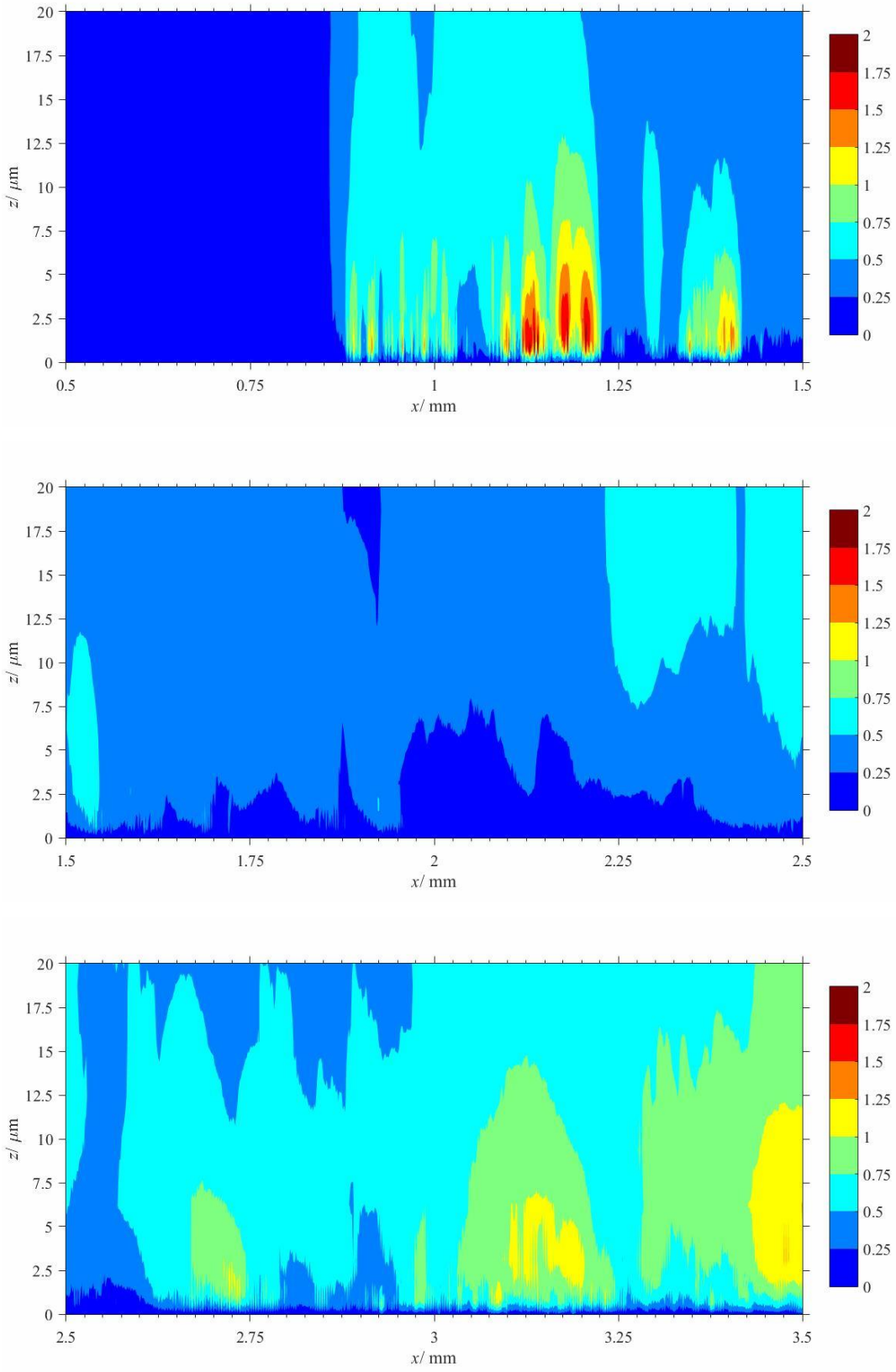


Figure 6.14 – Pinion: Dang Van fatigue parameters for 10^7 loading cycles at the close proximity to the pinion flank surface

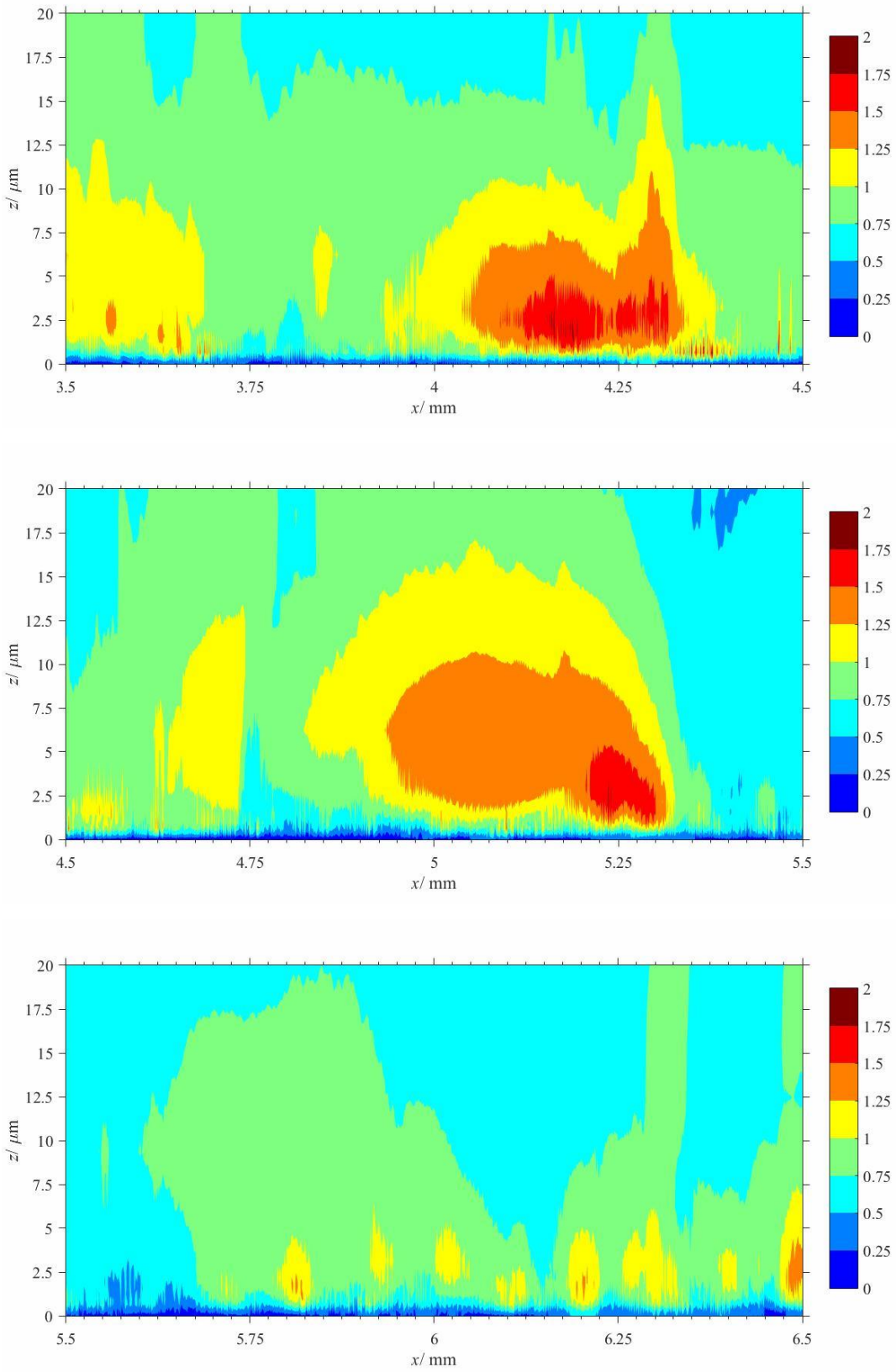


Figure 6.14 – Pinion: Dang Van fatigue parameters for 10^7 loading cycles at the close proximity to the pinion flank surface

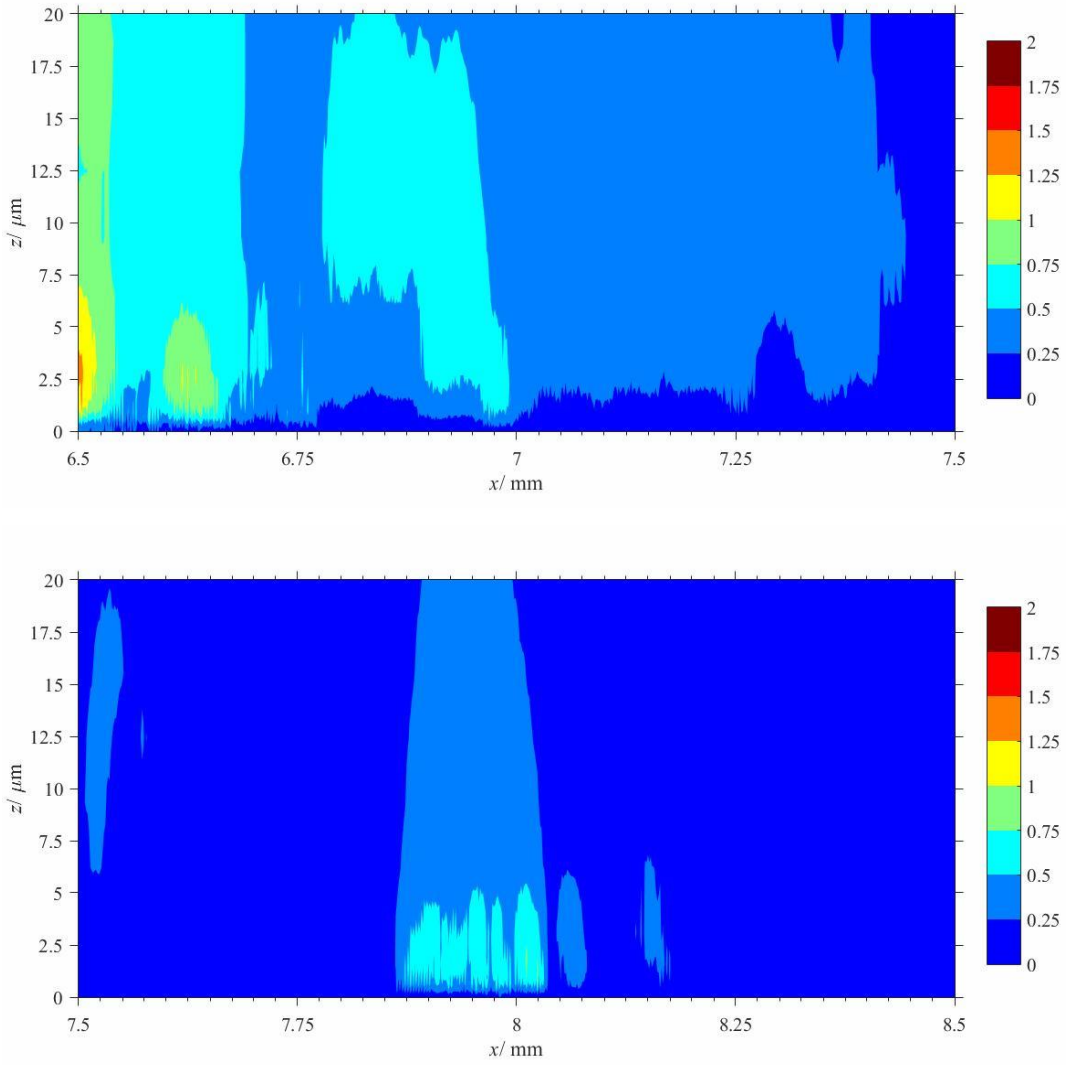


Figure 6.14 – Pinion: Dang Van fatigue parameters for 10^7 loading cycles at the close proximity to the pinion flank surface

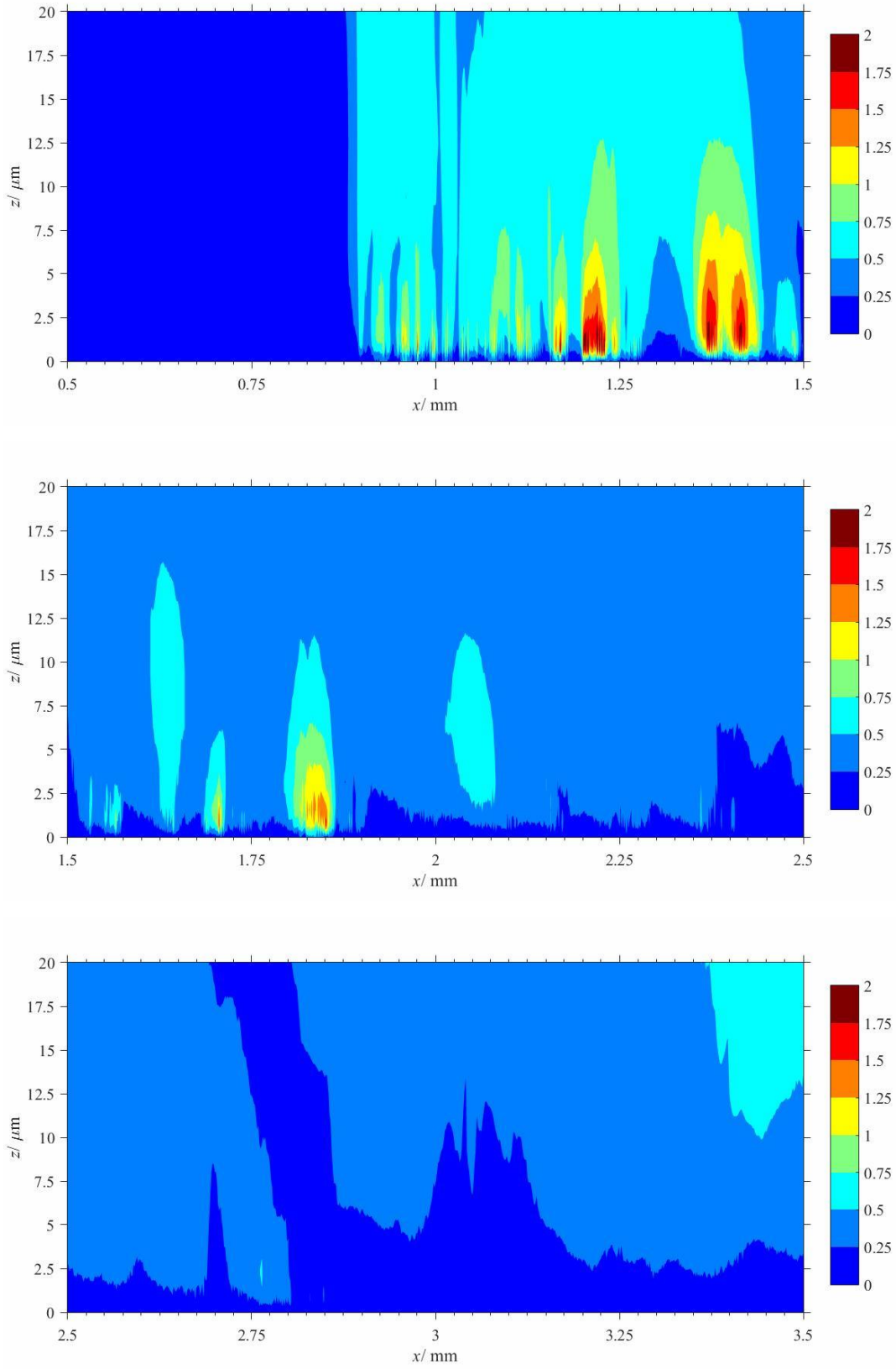


Figure 6.15 – Gear: Dang Van fatigue parameters for 10^7 loading cycles at the close proximity to the pinion flank surface

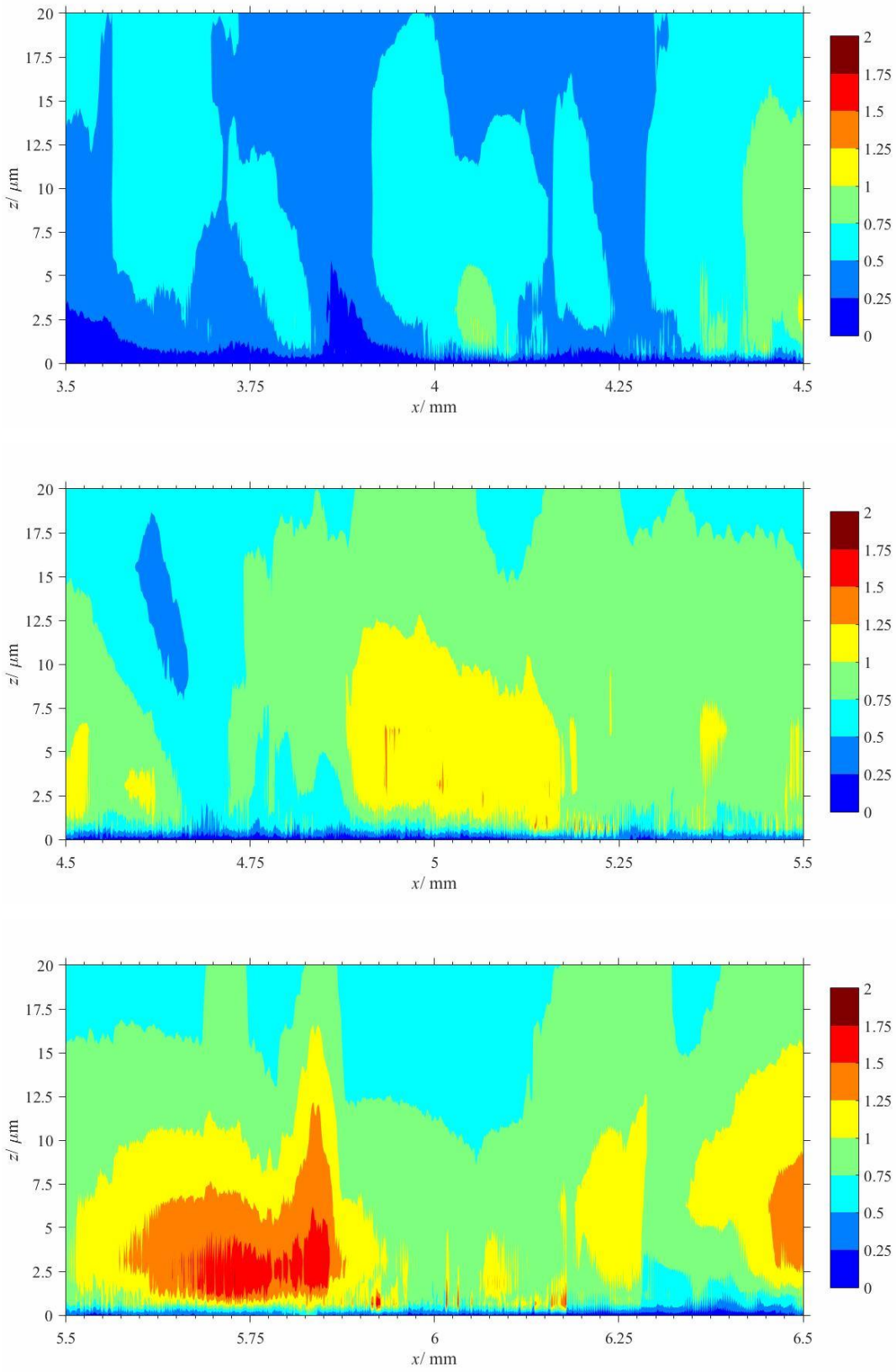


Figure 6.15 – Gear: Dang Van fatigue parameters for 10^7 loading cycles at the close proximity to the pinion flank surface

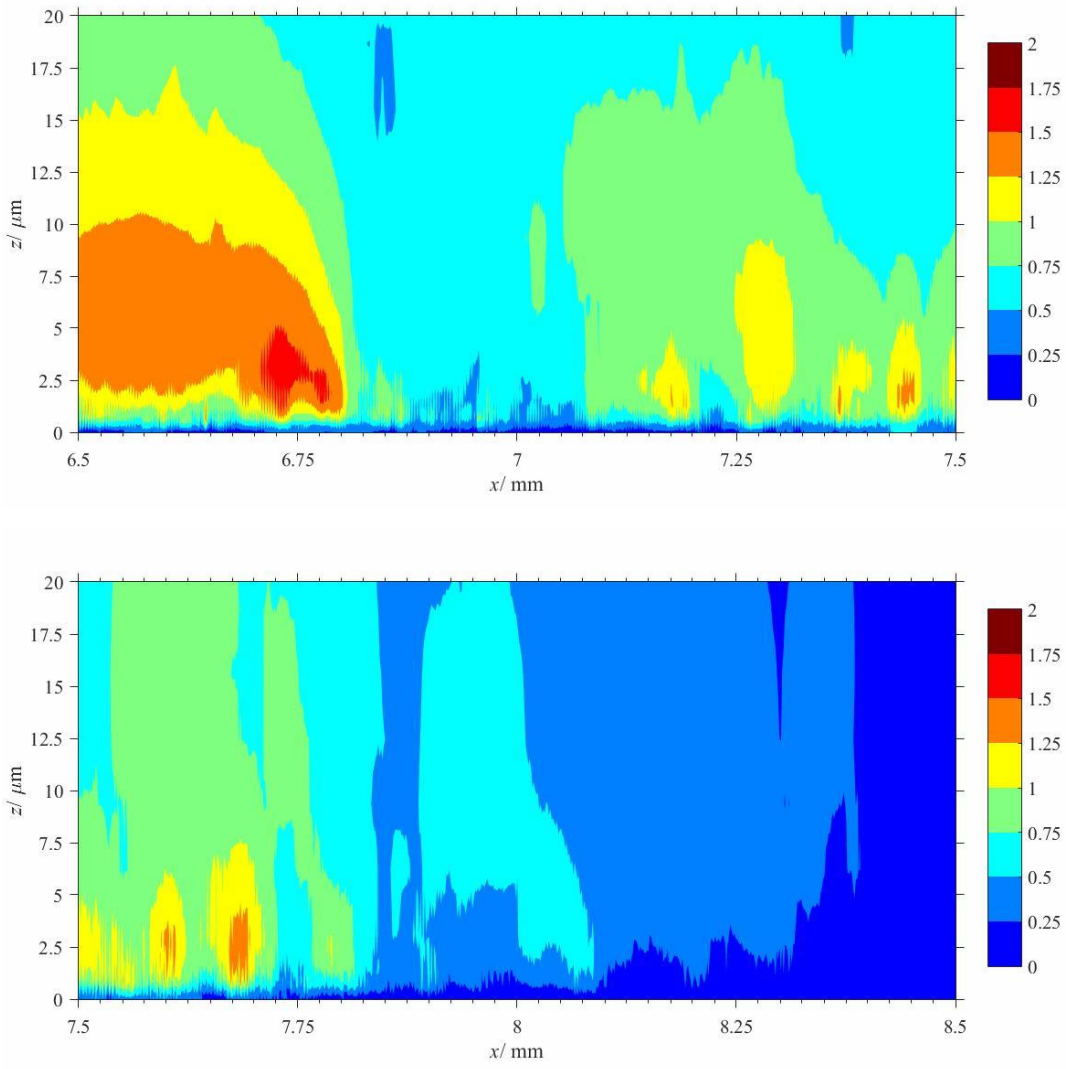


Figure 6.15 – Gear: Dang Van fatigue parameters for 10^7 loading cycles at the close proximity to the pinion flank surface

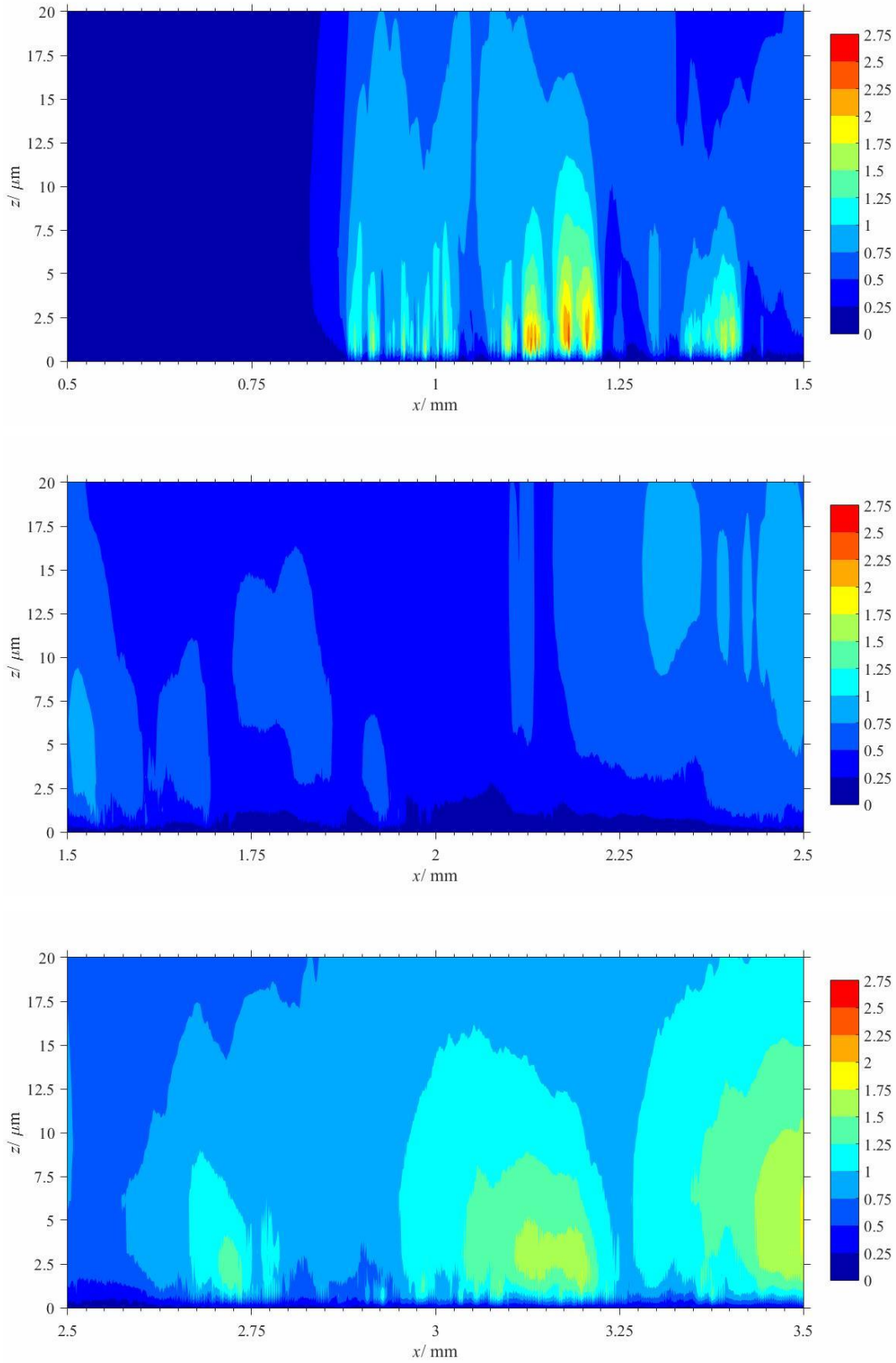


Figure 6.16 – Pinion: Findley fatigue parameters for 10^7 loading cycles at the close proximity to the pinion flank surface

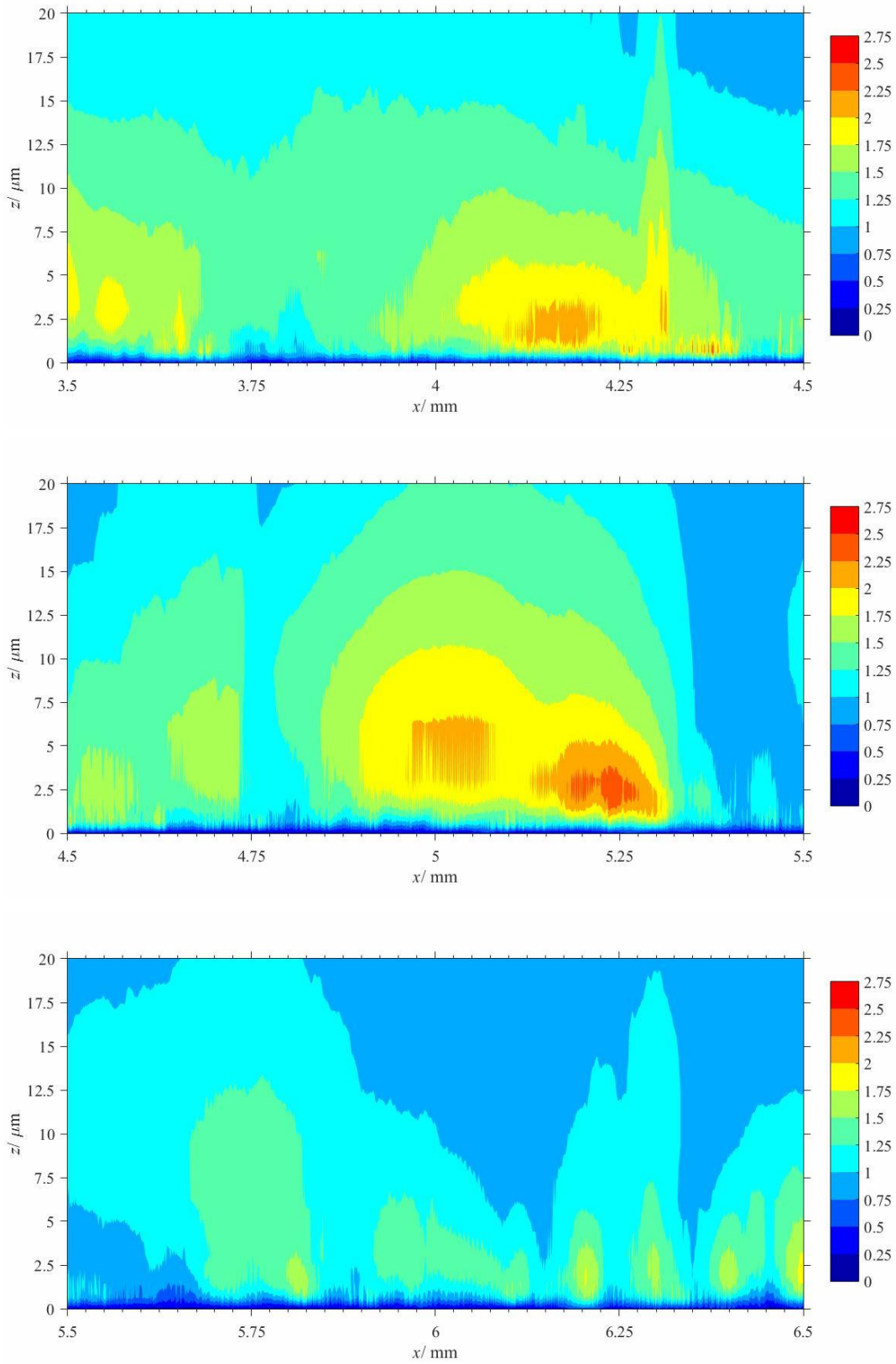


Figure 6.16 – Pinion: Findley fatigue parameters for 10^7 loading cycles at the close proximity to the pinion flank surface

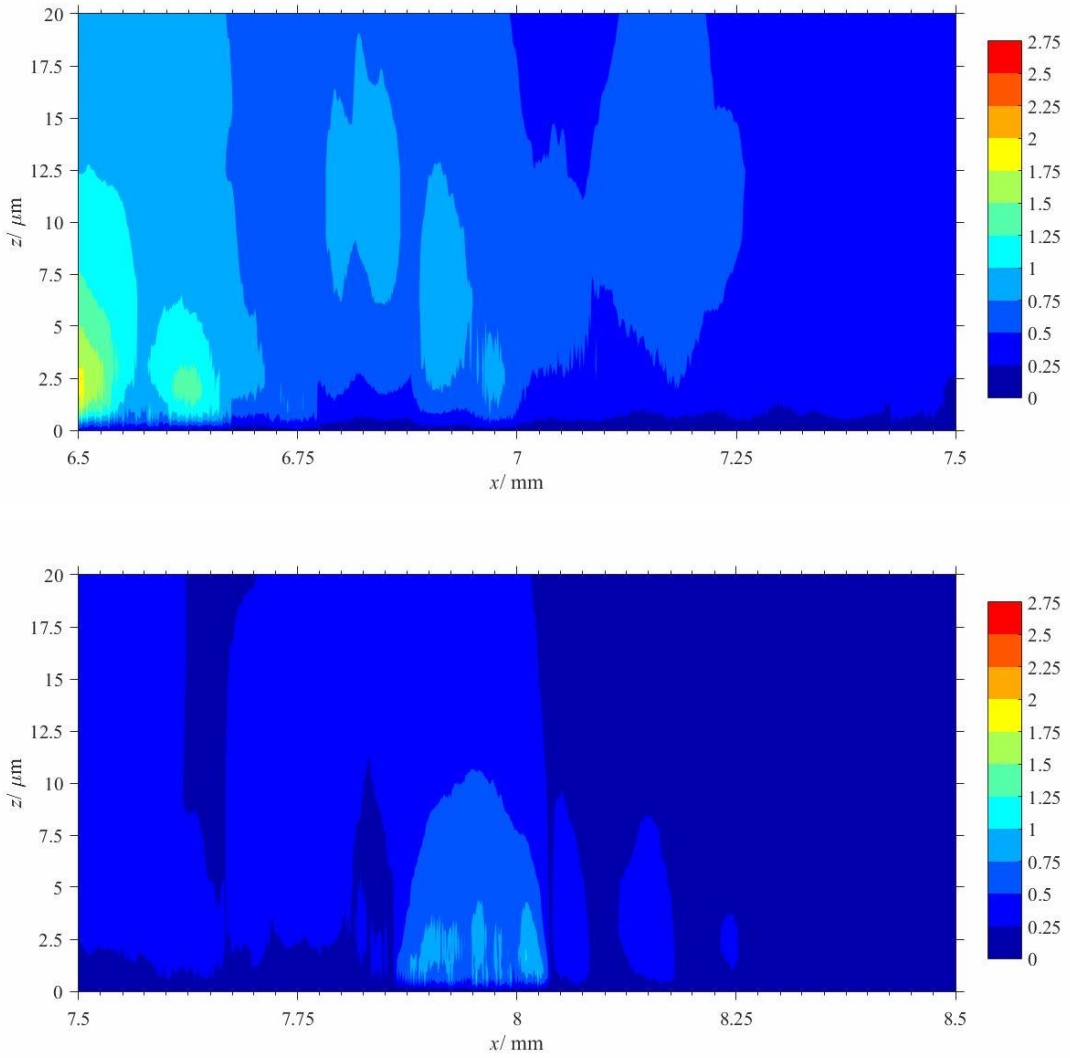


Figure 6.16 – Pinion: Findley fatigue parameters for 10^7 loading cycles at the close proximity to the pinion flank surface

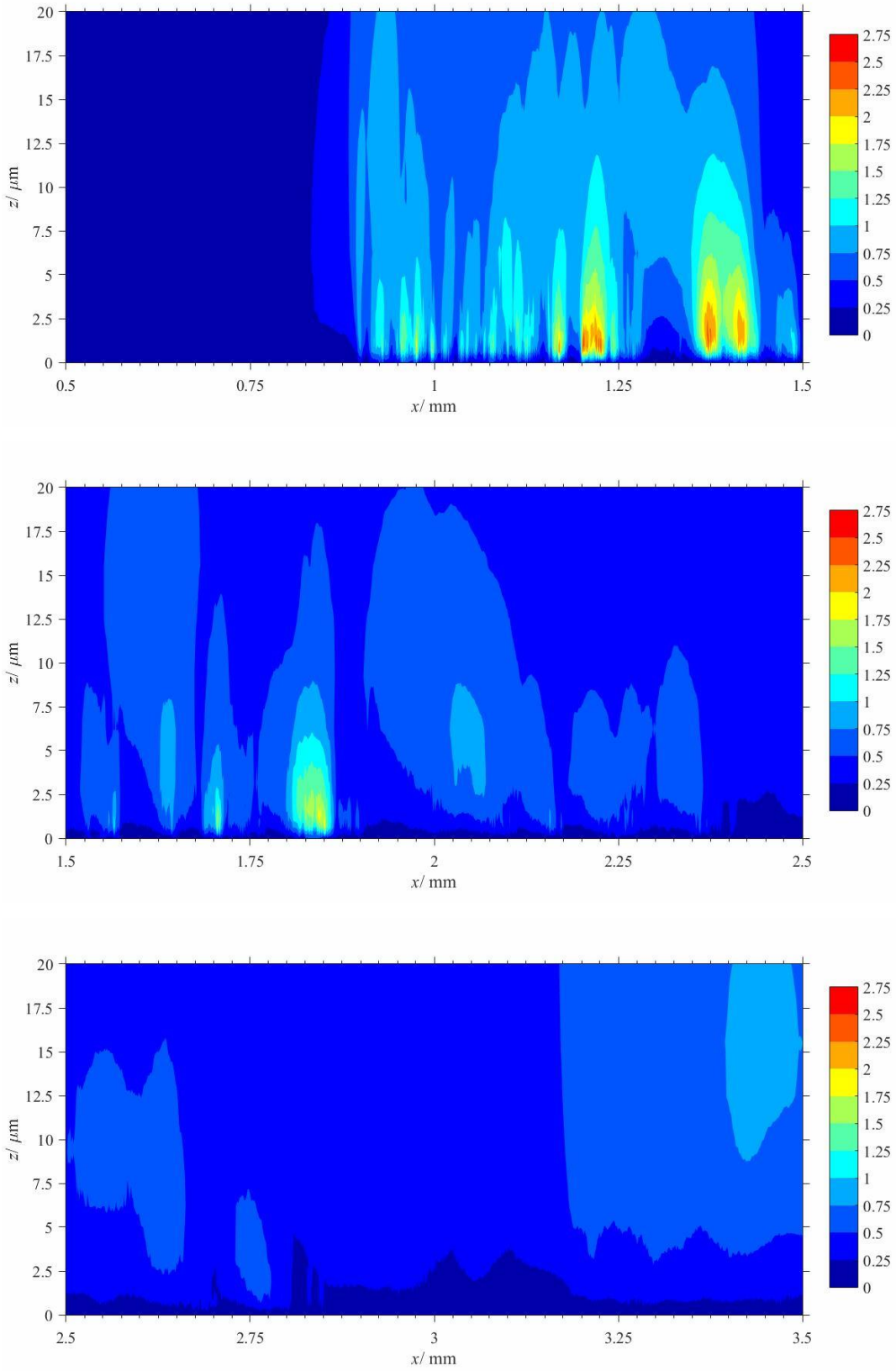


Figure 6.17 – Gear: Findley fatigue parameters for 10^7 loading cycles at the close proximity to the pinion flank surface

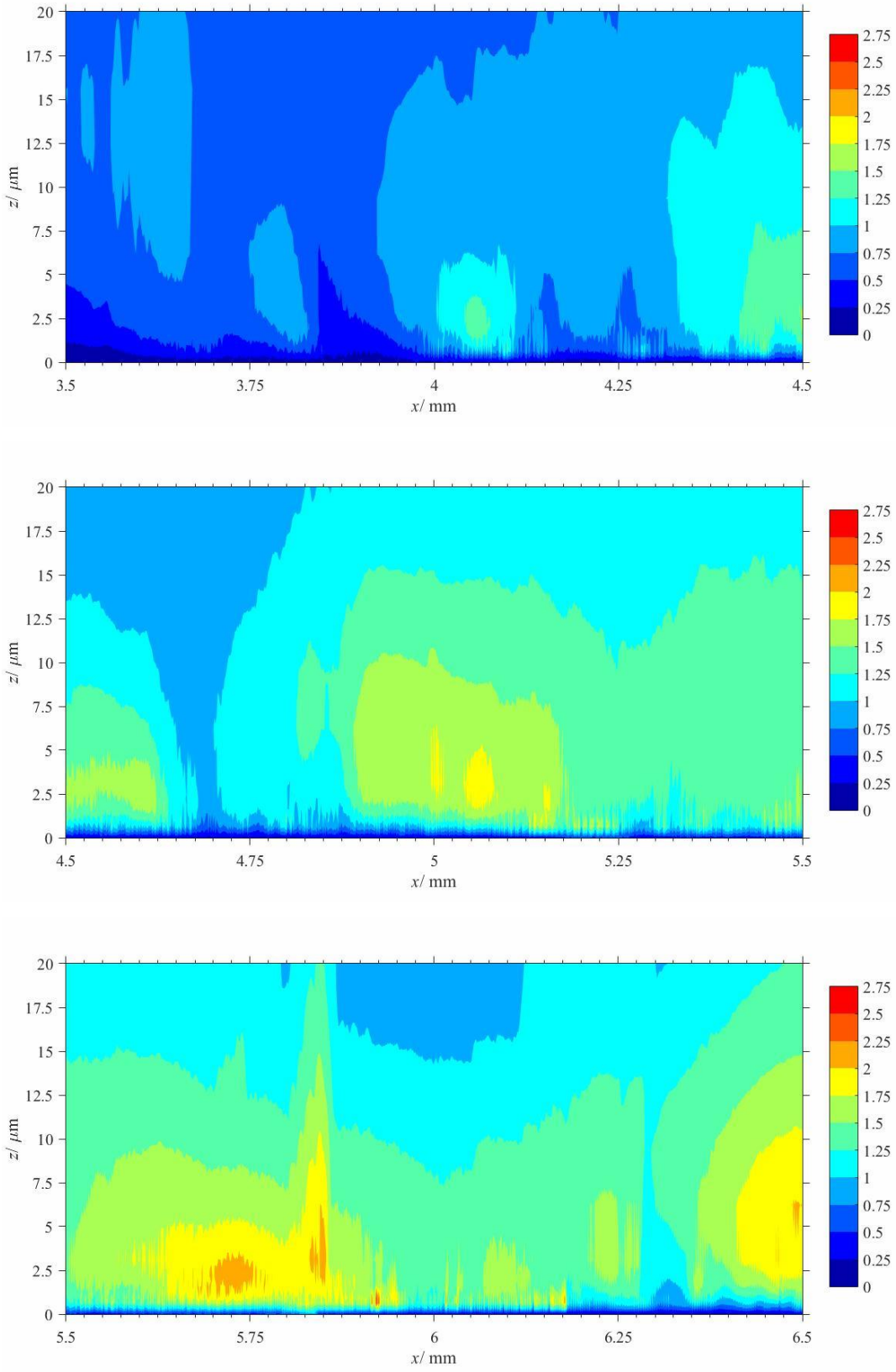


Figure 6.17 – Gear: Findley fatigue parameters for 10^7 loading cycles at the close proximity to the pinion flank surface

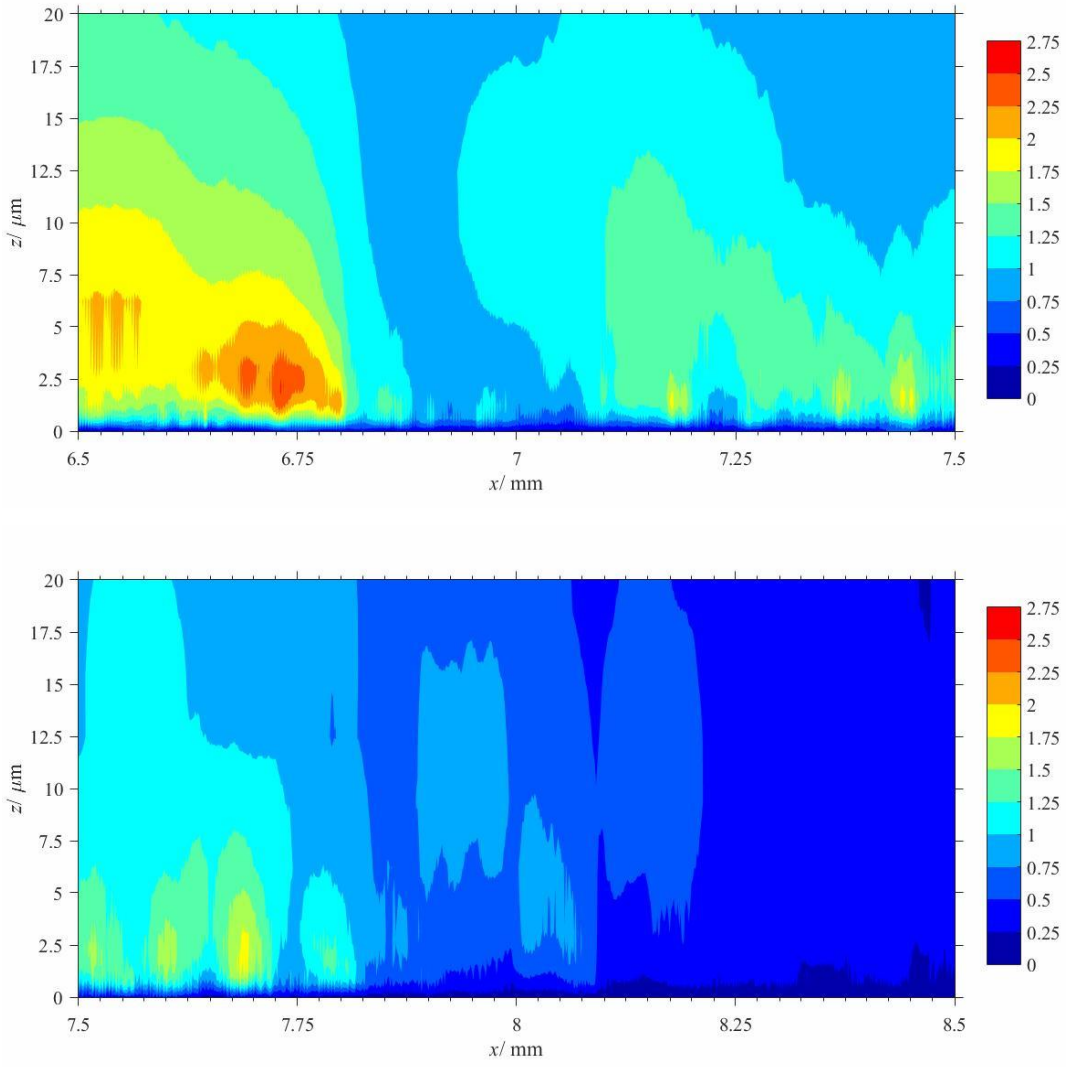


Figure 6.17 – Gear: Findley fatigue parameters for 10^7 loading cycles at the close proximity to the pinion flank surface

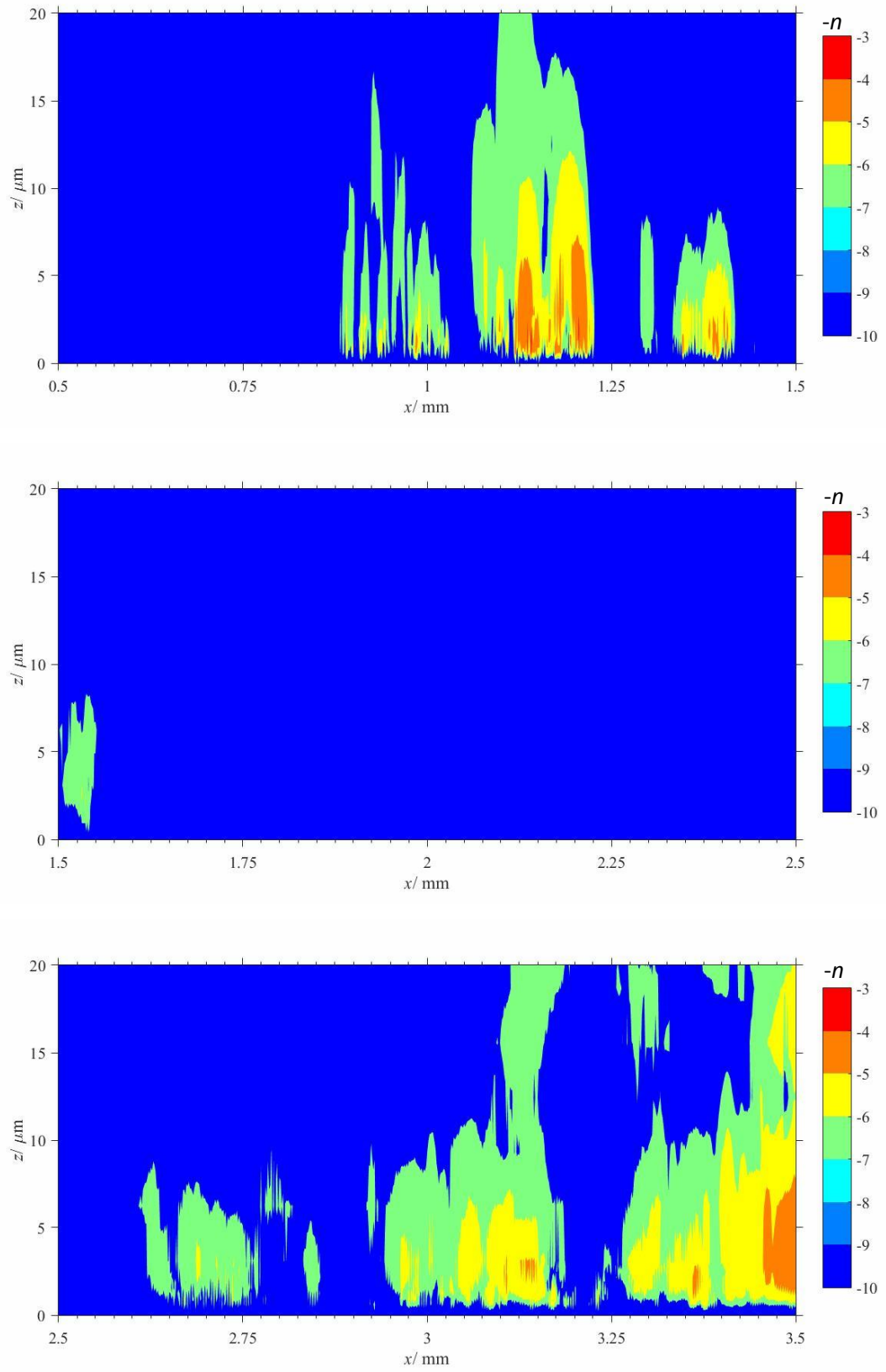


Figure 6.18 – Pinion: Fatemi and Socie accumulated damage, 10^{-n}, at the close proximity to the pinion flank surface indicating fatigue in 10^{n} cycles

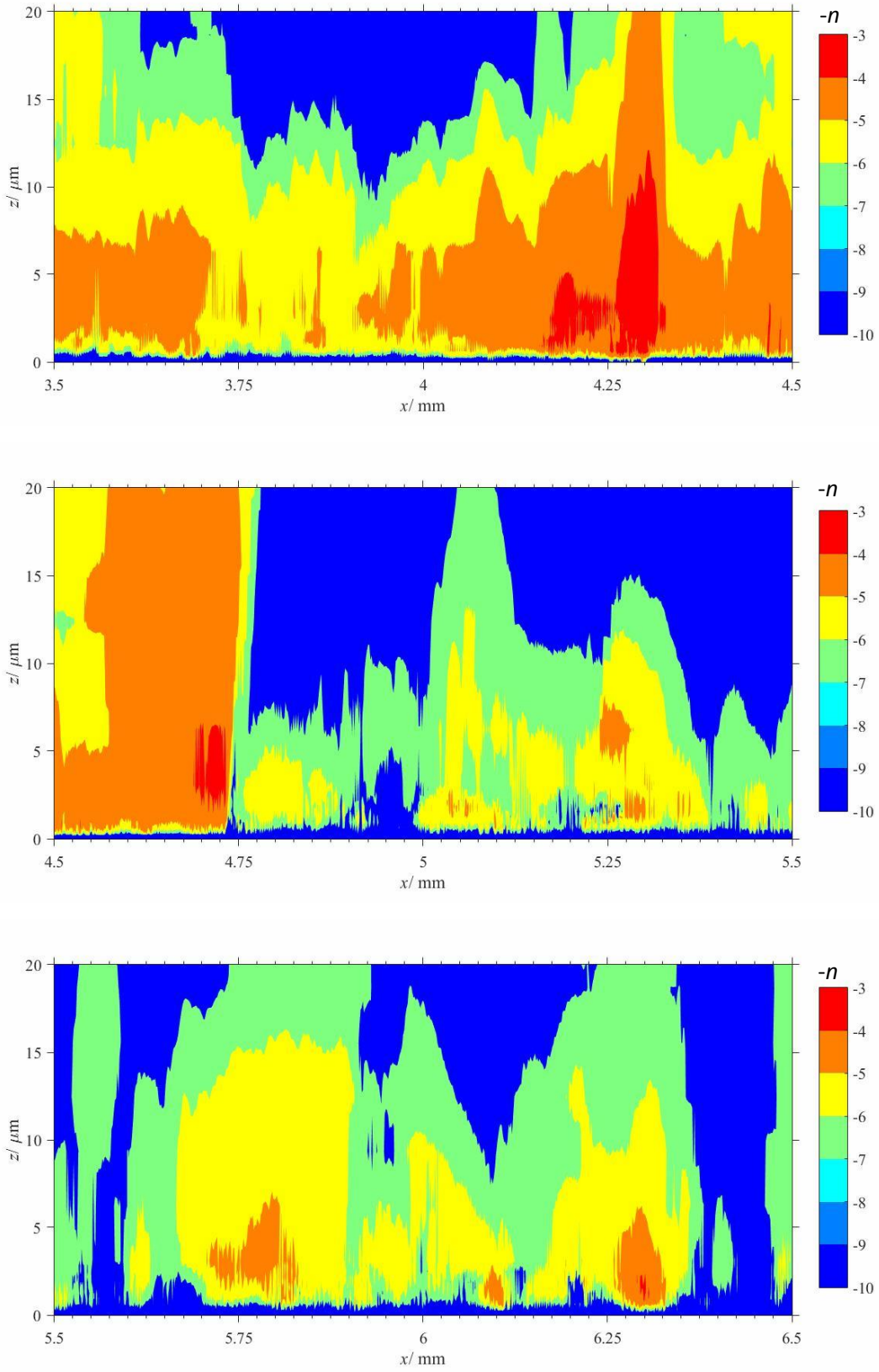


Figure 6.18 – Pinion: Fatemi and Socie accumulated damage, 10^{-n}, at the close proximity to the pinion flank surface indicating fatigue in 10^{n} cycles

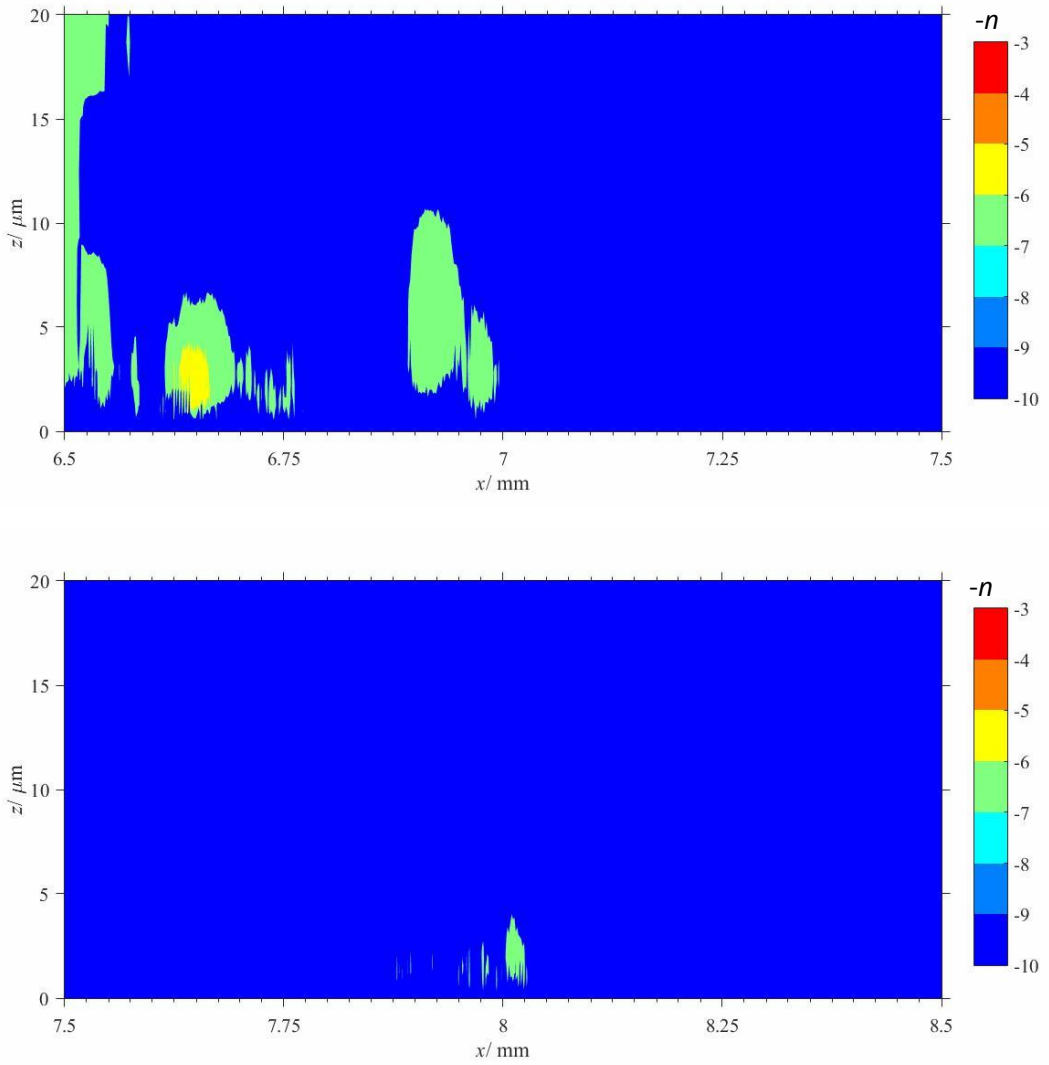


Figure 6.18 – Pinion: Fatemi and Socie accumulated damage, 10^{-n}, at the close proximity to the pinion flank surface indicating fatigue in 10^{n} cycles

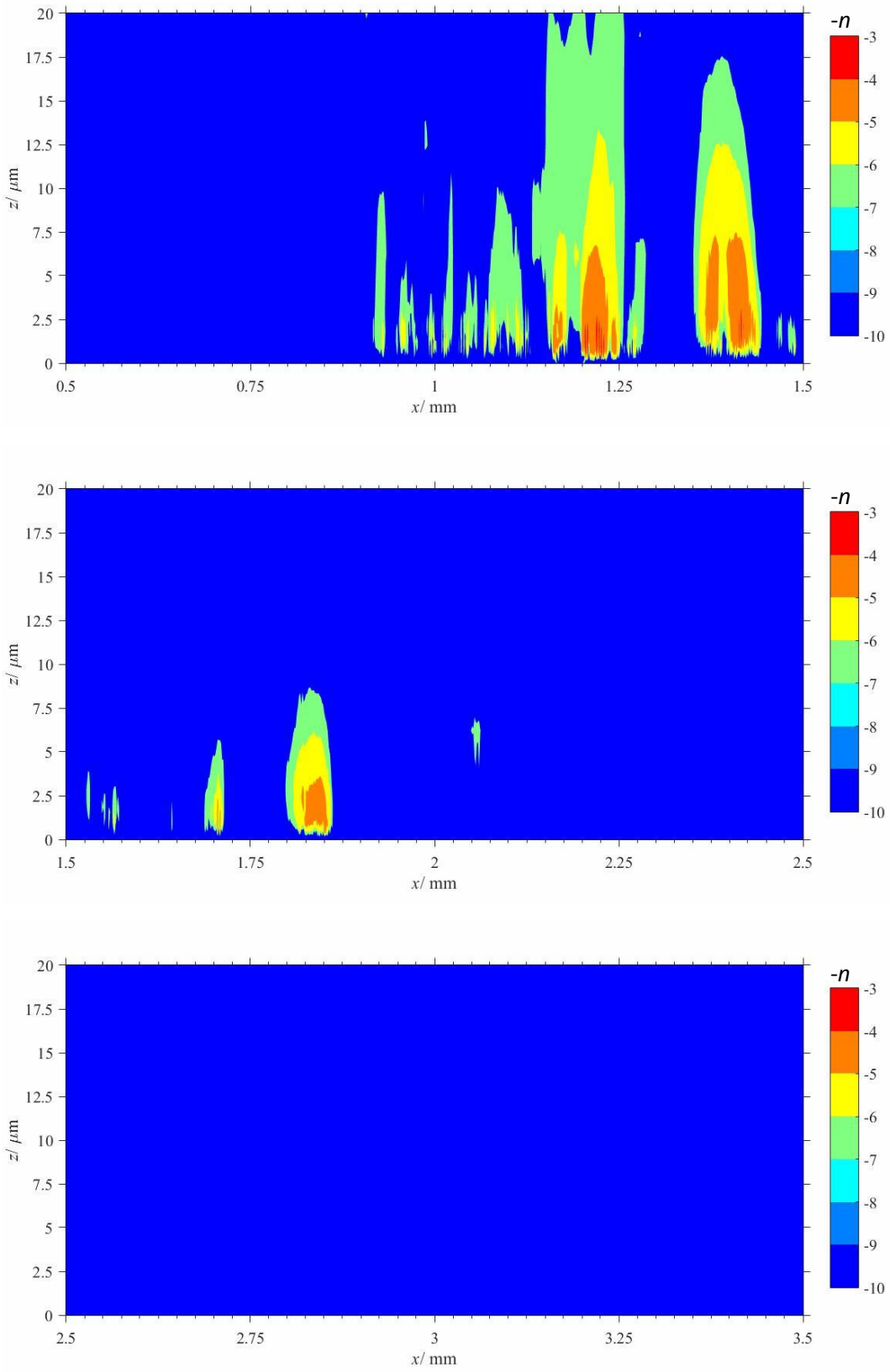


Figure 6.19 – Gear: Fatemi and Socie accumulated damage, 10^{-n}, at the close proximity to the gear flank surface indicating fatigue in 10^{n} cycles

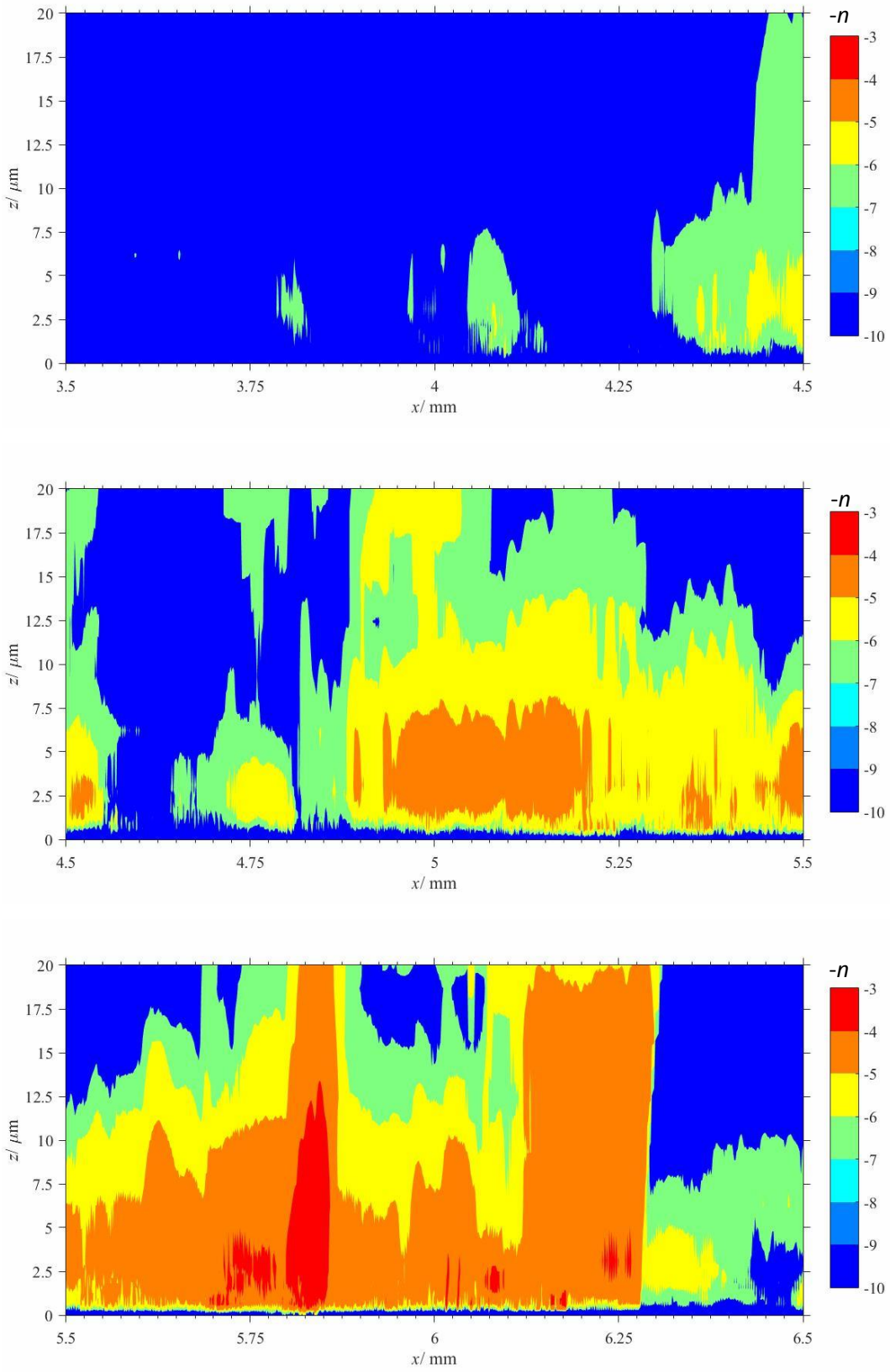


Figure 6.19 – Gear: Fatemi and Socie accumulated damage, 10^{-n}, at the close proximity to the gear flank surface indicating fatigue in 10^{n} cycles

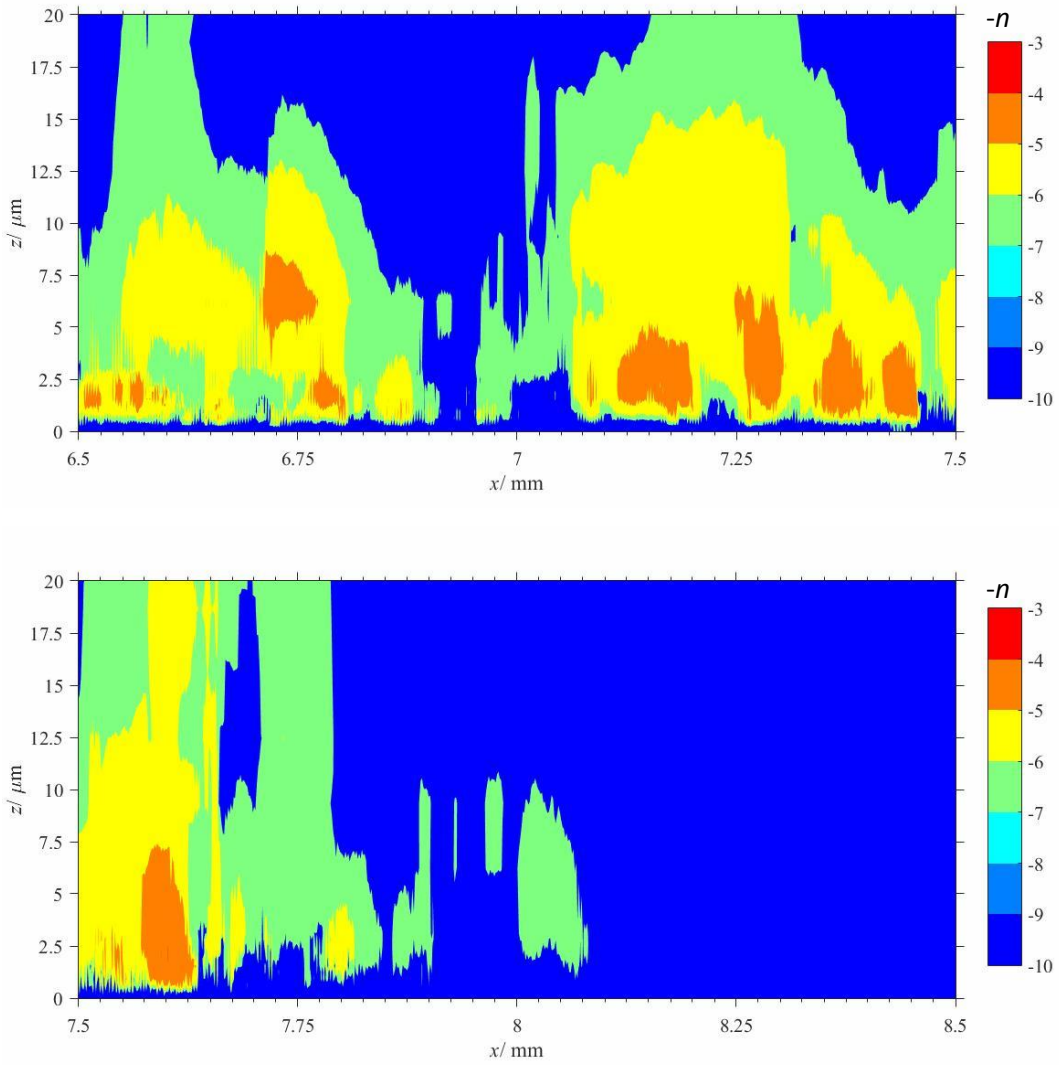


Figure 6.19 – Gear: Fatemi and Socie accumulated damage, 10^{-n}, at the close proximity to the gear flank surface indicating fatigue in 10^{n} cycles

Chapter 7

Conclusions and Recommendations for future work

7.1. Conclusions

The thesis presents a procedure to:

- ✓ Obtain a transient isothermal numerical EHL solution of the spur gear full meshing cycle, employing the Eyring rheological shear-thinning model, to accommodate the non-Newtonian behaviour of the lubricant and taking into account measured surface roughness of the gear flanks and different types of the operating load variation;
- ✓ Calculate elastic stresses at the different stages of the meshing cycle using the Fast Fourier Transform numerical technique to reduce computing time;
- ✓ Sort and store loading stress histories for the material of the gear flanks effectively;
- ✓ Evaluate the predicted fatigue-life of the gear flanks in contact by a number of fatigue theory approaches.

The work provides a comprehensive description of the EHL problem formulation, including the calculations of the geometry and kinematics of the spur gear contact over the meshing cycle, the application of the coupled method for solution of the Reynolds and the elastic deflection equations. The elastic stress evaluation method is explained in detail and verified. In addition some popular multi-axial and cumulative damage fatigue theories in the literature are discussed and implemented.

The aims of the research itemised in Section 1.8 have all been achived in the software developed, and the graphical tools produced have been used extensively in presenting the results at all stages of the work.

The analysis methods have been applied to the extreme conditions adopted in endurance testing of gears and give a detailed insight into the interface conditions experienced by the gear flank material over all of the meshing cycle. They also allow an assessment of fatigue life to be made.

The results of calculation are presented for two case studies showing that the faster speed and smoother NASA Glenn test gears are less likely to fatigue than the slower speed and rougher Design Unit ones as expected. However, the NASA Glenn gears appeared to have to centres of fatigue: at the surface driven by the combined surface roughness and at 0.79 of the half-Hertzian contact dimension as in a dry contact of collinear cylinders, which had not been the case in the previous studies of the Cardiff Tribology group. The EHL pressure distribution is similar to the parabolic Hertz pressure function with moderate oscillations due to the surface roughness. Contact analysis of the Design Unit test gears shows a more common picture: the fatigue is predicted to happen at the surface due to the gear flanks roughness. The EHL analysis results present a mixed lubricated contact, which can be considered as a combination of the simultaneous EHL contacts. Therefore, the fatigue-life is predicted to be lower due to the heavier interaction of the surfaces.

7.2. Recommendations for future work

The method presented in this thesis is based on the previous studies, which were proved correct by a series of tests. However, the combination of those methods is novel and was never investigated experimentally. Therefore, the obvious development of study would be to obtain test data and compare it with the simulation results.

The suggested development of the software would include: (i) elastic-plastic EHL analysis, rather than just elastic one employed now; (ii) improvement of the surface roughness processing and aligning technique; (iii) including a thermal analysis in the EHL solution.

At the current stage, the algorithm cannot accommodate plastic behaviour of the gear flank material, resulting in the unrealistically high pressure values and the unreasonable change of the lubricant properties instead of the permanent change of the surface geometry. Some early attempts to incorporate the plastic deformation treating it as a permanent surface profile change of the same magnitude as the elastic response to the excess pressure were investigated and some results were obtained, but the timeline of the project did not allow this to be developed to the stage of experimental testing of the process..

The importance of accurate alignment of the surface roughness was emphasised in Chapter 6. Currently, the final tuning is done by a trial and error method, which is time consuming and requires a high level of expertise. Therefore, the process requires some simplification and better tools to identify the position of the filtered roughness data to the gear flank.

The accuracy of the measured data is limited by the hardware. The Talysurf probe measures roughness perpendicular to the profile traverse direction rather than perpendicular to the gear flank surface which produces an error that could be corrected.

The thermal analysis would increase the accuracy of the EHL solution and provide an opportunity to include the thermal stress calculations, which may be important in terms of the plastic yielding and crack propagation processes.

Appendix A – MatLAB code

A.1. Plotting the EHL Results

```
% setting directory and file type
folder='E:\Chapter 5\'; ext='.txt';

% timestep sequence to plot
first=9000;step=1;last=11000;

% pressure axis set up in GPa
prmin=0;prstep=.5;prmax=4;
% film thickness axis set up in µm
filmmin=0;filmstep=.5;filmmax=4;
% horizontal axis set up in x/a
xup=-1.5;xstep=.25;xdown=1.5;

% creats a figure 800x800 pixels and white background
fig=figure;
set(fig, 'Position',[1, 1, 800, 800]);
set(fig, 'Color','white');

for n=first:step:last

    % reads data from files
    name=num2str(n,'%06d');
    full_name=strcat(folder,name,ext);
    toplot=dlmread(full_name);
    dx=abs(toplot(2,1)-toplot(1,1));

    % creates three subplots
    sub1=subplot(3,1,1);
    sub2=subplot(3,1,2);
    sub3=subplot(3,1,3);


    % plots pressure and film thickness in subplot one
    [ax,pr,he]=plotyy(toplot(:,1),toplot(:,2),toplot(:,1),toplot(:,3),'Parent',sub1);
```

```
    % formats subplot one
    set(ax,'Layer','top');
    set(ax,'TickDir','out');
    set(ax(1),'YMinorTick','on');
    set(ax(2),'YMinorTick','on');
    set(ax,'XAxisLocation','top');
    set(ax,'XMinorTick','on');
    set(ax(1),'YMinorGrid','on');
    set(ax,'XMinorGrid','on');
    set(ax,'XGrid','on');
    xlabel('\itx\rm/\ita\rm','Parent',sub1);
    set(ax(1),'YGrid','on');
    set(ax(1),'Ylim',[prmin,prmax]);
    set(ax(2),'Ylim',[filmmin,filmmax]);
    set(ax(1),'YTick',[prmin:prstep:prmax]);
    set(ax(2),'YTick',[filmmin:filmstep:filmmax]);
    set(ax,'Xlim',[xup,xdown]);
    set(ax,'XTick',[xup:xstep:xdown]);
    set(ax(1),'YColor','red');
    set(ax(2),'YColor',[0 0.4 0]);
    set(get(ax(1), 'Ylabel'), 'String', 'Pressure/GPa');
    set(get(ax(2), 'Ylabel'), 'String', 'Film thickness/\mum');
    set(pr,'Color','red');
    set(he,'Color',[0 0.4 0]);
    set(pr,'LineWidth',1.5);
    set(he,'LineWidth',1.5);

    % plots offsets of flanks in subplot two

offsets=plot(toplot(:,1),toplot(:,4),toplot(:,1),toplot(:,5),'Parent',sub2,'LineWidth',1.5,'
Color', 'black');

    % formats subplot one
    set(sub2,'XLim',[xup,xdown]);
    set(sub2,'XTick',[xup:xstep:xdown]);
    set(sub2,'YTick',[]);
    set(sub2,'YLim',[min(toplot(:,4))-5,min(toplot(:,4))+5]);
    set(sub2,'TickDir','out');
    set(sub2,'XTickLabel',[]);
    set(sub2,'XTick',[]);

    % plots regime indicator in subplot three
    ind=bar(toplot(:,1),toplot(:,6),'Parent',sub3);


    %f ormats subplot three
    set(sub3,'XTick',[xup:xstep:xdown]);
    set(sub3,'XLim',[xup,xdown]);
    set(sub3,'YLim',[-1, 1]);
    set(sub3,'YTick',[]);
```

```
    set(sub3,'XMinorTick','on');
    set(sub3,'TickDir','out');
    xlabel('\itx\rm/\ita\rm','Parent',sub3);

    % sets fonts
    set(sub1,'FontName','Times New Roman');
    set(sub1,'FontSize',16);
    set(ax(2),'FontName','Times New Roman');
    set(ax(2),'FontSize',16);
    set(sub3,'FontName','Times New Roman');
    set(sub3,'FontSize',16);

    % sets size of subplots
    set(sub3,'Position',[0.1, 0.09, 0.8, 0.02]);
    set(sub2,'Position',[0.1, 0.1, 0.8, 0.298]);
    set(sub1,'Position',[0.1, 0.4, 0.8, 0.53]);

    % rasterises figure and saves it as jpeg in same directory
    % using timestep id as name of file
    frame=getframe(fig);
    im = frame2im(frame);
    [A,map] = rgb2ind(im,256);
    pic_name=strcat(folder,name,'.jpg');
    imwrite(A,map,pic_name,'jpg');

end
```

A.2. Plotting the roll-angles to fit the roughness

```
% Declaration of function with four arguments:
% rbp and rap are base and addendum radii of pinion
% rbp and rap are base and addendum radii of gear
function rollangle(rbp, rap, rbg, rag)

clear xinvp yinvp lnrap  xinvg yinvg lnrag gamma sum;

% calculating involute coordinates
for gamma=1:350000
    gr=degtorad(gamma/10000);
    A=[sin(gr),-cos(gr);cos(gr),sin(gr)]*[1; gr];
    xinvp(gamma)=rbp*A(1,1);
    yinvp(gamma)=rbp*A(2,1);
    xinvg(gamma)=rbg*A(1,1);
    yinvg(gamma)=rbg*A(2,1);
end
```

```
% setting tooth tip horizontal line
lnrap(1:max(size(xinvp)))=rap;
lnrag(1:max(size(xinvg)))=rag;
gamma=1:350000;
gamma=gamma/10000;

% creating a figure 800x800 pixels and white background
fig=figure;
set(fig, 'Position',[1, 1, 800, 800]);
set(fig, 'Color','white');

% plotting involutes and tip lines
sub=subplot(1,1,1);
radp=line(gamma,yinvp,'Parent',sub,'LineWidth',1,'Color', 'blue');
radg=line(gamma,yinvg,'Parent',sub,'LineWidth',1,'Color', 'red');
addp=line(gamma,lnrap,'Parent',sub,'LineWidth',1,'Color', 'blue','LineStyle', '--');
addg=line(gamma,lnrag,'Parent',sub,'LineWidth',1,'Color', 'red','LineStyle', '--');

% formating plot
set(sub,'XLim',[0,35]);
set(sub,'XTick',0:5:35);
set(sub,'XTickLabel',0:5:35);
set(sub,'YLim',[73,86]);
set(sub,'YTick',73:86);
set(sub,'TickDir','out');
set(sub,'YMinorTick','on');
set(sub,'XMinorTick','on');
set(sub,'XMinorGrid','on');
set(sub,'XGrid','on');
set(sub,'YMinorGrid','on');
set(sub,'YGrid','on');
xlabel('Roll-angle / deg','Parent',sub);
ylabel('Gear radius / mm','Parent',sub);

% calculating arc lengths and adding markers and text for pinion
n=ceil(rbp);
while n<=ceil(rap)

   if n~=ceil(rap)
      rtarg=n;
   else
      rtarg=rap;n=rap;
   end
   i=2;arc=0;
   while yinvp(i)<rtarg
      dx=xinvp(i)-xinvp(i-1);dy=yinvp(i)-yinvp(i-1);
      arc=arc+sqrt(dx*dx+dy*dy);
      i=i+1;
   end
```

```
    txt=sprintf('%7.4f mm\n%7.4f',arc,gamma(i-1));
    line(gamma(i-1),n,'Parent',sub,'Marker','o','MarkerEdgeColor', 'blue','
MarkerFaceColor', 'blue','MarkerSize',6,'LineStyle', 'none');
    text(gamma(i),n,[txt,'\circ'],'HorizontalAlignment','left','Color','blue','FontName',
'Times New Roman','FontSize',14)
    n=n+1;
end



% calculating arc lengths and adding markers and text for gear
n=ceil(rbg);
while n<=ceil(rag)

    if n~=ceil(rag)
        rtarg=n;
    else
        rtarg=rag;n=rag;
    end
    i=2;arc=0;
    while yinvg(i)<rtarg
        dx=xinvg(i)-xinvg(i-1);dy=yinvg(i)-yinvg(i-1);
        arc=arc+sqrt(dx*dx+dy*dy);
        i=i+1;
    end
    txt=sprintf('%7.4f mm\n%7.4f',arc,gamma(i-1));
    line(gamma(i-1),n,'Parent',sub,'Marker','o','MarkerEdgeColor',
'red','MarkerFaceColor', 'red','MarkerSize',6,'LineStyle', 'none');

text(gamma(i),n,[txt,'\circ'],'HorizontalAlignment','right','Color','red','FontName','Tim
es New Roman','FontSize',14)
    n=n+1;
end

end
```

References

Ai, X., 1993. *Numerical analyses of elastohydrodynamically lubricated line and point contacts with rough surfaces by using semi-system and multigrid methods*. Ph. D. Thesis, Northwestern University, Evanston.

Ai, X., Cheng, H.S., 1994, A Transient EHL Analysis for Line Contacts with Measures Surface Roughness Using Multigrid Technique", *Trans. ASME, Journal of Tribology*, Vol. 116, No. 3, pp. 549 – 556.

Albert, W. A. J., 1837 Über Treibseile am Harz. *Archive für Mineralogie Geognosie Bergbau und Hüttenkunde*, vol. 10, pp 215-34

Archard, J.F, Cowking, E.W., 1966. A simplified treatment of elastohydrodynamic lubrication theory for a point contact. *Proceedings of the Institution of Mechanical Engineers*, 1965–66; 180, part 3B; *Elastohydrodynamic Lubrication: A symposium arranged by the Lubrication and Wear Group 21st–23rd September 1965.*

Archard, J.F., 1968. Non-dimensional parameters in isothermal theories of elastohydrodynamic lubrication. *Proc. IMechE Sci.*, Vol. 10, Part. 2, pp. 165-167.

Archard, J.F., Kirk, M.T., 1961. Lubrication at Point Contacts. *Proc. R. Soc. Lond. A.*, Vol. 261, pp. 532 – 550.

ASM, 1992, *ASM Handbook* Volume 18*: Friction, Lubrication and Wear Technology*, ASM International, pp. 535-545.

ASTM E1012 *Practice for Verification of Testing Frame and Specimen Alignment Under Tensile and Compressive Axial Force Application*

ASTM E1049, 2011. *Standard Practices for Cycle Counting in Fatigue Analysis*

ASTM E1823 *Terminology Relating to Fatigue and Fracture Testing*

ASTM E3 *Guide for Preparation of Metallographic Specimens*

ASTM E466 *Standard Practice for Conducting Force Controlled Constant Amplitude Axial Fatigue Tests of Metallic Materials*

ASTM E467 *Practice for Verification of Constant Amplitude Dynamic Forces in an Axial Fatigue Testing System*

ASTM E468 *Practice for Presentation of Constant Amplitude Fatigue Test Results for Metallic Materials*

ASTM E606/E606M *Test Method for Strain-Controlled Fatigue Testing*

ASTM E739 *Practice for Statistical Analysis of Linear or Linearized Stress-Life (S-N) and Strain-Life (ε-N) Fatigue Data*

ASTM STP 1122, 1992. *Advances in Fatigue Lifetime Predictive Techniques*

ASTM STP 1191, 1993. *Advances in Multiaxial Fatigue*.

ASTM STP 1387 *Multiaxial Fatigue and Deformation: Testing and Prediction*

AVRADCOM 82-C-16 / **NASA-CP-2210**, 1983. *Advanced Power Transmission Technology*. Geared Power Transmission Technology. Coy, John J., pp. 49-77

Bair, S., Winer, W.O., 1979, A Rheology Model for Elastohydrodynamic Contacts Based on Primary Laboratory Data, *Trans. ASME, Journal of Lubrication Technology*, Vol. 101, pp. 258 – 265.

Bannantine, J. A., Socie, D. F., 1992 A multiaxial fatigue life estimation technique. *Advances in Fatigue Lifetime Predictive Techniques*, **ASTM STP 1122**. [Eds.: M. R. Mitchell and R. W. Landgraf. Philadelphia, ASTM]. pp. 249-275.

Banvillet, A., Palin-Luc, T., Lasserre, S., 2003. A volumetric energy based high cycle multiaxial fatigue criterion. *Int. J. Fatigue* 25, pp. 755-769.

Barus, C., 1893, "Isothermal, Isopiestics and Isometrics in relation to Viscosity", *American Journal of Science, 3rd Series*, Vol. 45, pp. 87 – 96.

Blok, H., 1950, Fundamental Mechanical Aspects of Thin Film Lubrication. *Ann. N.Y. Acad. Sci.*, 53, p. 779–804.

Brychkov, Y.A., Prudnikov, A.P., Marichev, O.I. 1989. *Tables of Indefinite Integrals*, Amsterdam: Gordon and Breach Publishers Ltd.

BS 3518-1:1993. *Methods of fatigue testing. Methods of fatigue testing. Guide to general principles*

BS 3518-3:1963. *Methods of fatigue testing. Direct stress fatigue tests*

BS 7270:2006. *Metallic materials. Constant amplitude strain controlled axial fatigue. Method of test*

BS ISO 1143:2010. *Metallic materials. Rotating bar bending fatigue testing*

BS ISO 12107:2003. *Metallic materials. Fatigue testing. Statistical planning and analysis of data*

BS ISO 6336-3:2006. *Calculation of load capacity of spur and helical gears. Calculation of tooth bending strength.*

Cameron A, Gohar R., 1966. Theoretical and experimental studies of the oil film in lubricated point contact. *Proceedings of the Royal Society of London* A 291. Pp. 520–536.

Carpinteri, A., Spagnoli, A., 2001. Multiaxial high-cycle fatigue criterion for hard metals. *Int. J. Fatigue* 23, pp. 135-145.

Cazaud, R., Persoz, L., 1937. *La Fatigue des Mdtaux*. Dunod Press, Paris

Cheng, H.S., 1970, A numerical solution of the elastohydrodynamic film thickness in an elliptical contact. *Trans. ASME F, J. Lubric. Technol.*, **92**(1), pp. 155-162.

Cheng, H.S., 1999. *Tribology Surveillance*, Orlando, pp. 150–156

Cheng, H.S., Dyson, A., 1978. Elastohydrodynamic Lubrication of Circumferentially Ground Rough Disks. *Trans. ASLE*, vol. 21(1), pp. 25-40.

Conry, T.F., **Wang, S., Cusano, C.A.**, 1987, A Reynolds-Eyring Equation for Elastohydrodynamic Lubrication in Line Contacts, *Trans. ASME, Journal of Tribology*, Vol. 109, pp. 648 – 658.

Crank, J., Nicolson, P. 1947. A practical method for numerical evaluation of solutions of partial differential equations of the heat conduction type. *Proc. Camb. Phil. Soc.* 43 (1) pp. 50–67

Crook, A.W., 1958. The Lubrication of Rollers, I. *Philos. Trans. R. Soc. Lond. A.,* Vol 250, pp. 387 – 409.

Crook, A.W., 1961a. The Lubrication of Rollers, II. Film Thickness with Relation to Viscosity and Speed. *Philos. Trans. R. Soc. Lond. A.*, Vol 254, pp. 223 – 236.

Crook, A.W., 1961b. The Lubrication of Rollers, III. A Theoretical Discussion of Friction and Temperatures in the Oil Film. *Philos. Trans. R. Soc. Lond. A.*, Vol 254, pp. 237 – 258.

Crook, A.W., 1963. The Lubrication of Rollers, IV. Measurements of friction and effective viscousity. *Philos. Trans. R. Soc. Lond. A.*, Vol 255, pp. 281 – 312.

Crossland, B., 1956. Effect of large hydrostatic pressure on the torsional fatigue strength of an alloy steel. *Proc. Int. Conf. on Fatigue of Metals, Institution of Mechanical Engineers*, London, pp. 138-149.

Dang Van, K., 1973. *Sur la résistance a la fatigue des métaux.* These de Doctorat es Sciences, Sci. Techniq. l´Armement, 47, p. 647.

Davies, C.N. 2005. *Effects of non-Newtonian rheology on the line contact elastohydrodynamic lubrication problem*. PhD thesis, Cardiff University.

Department of Education and Science, 1966. *Lubrication (Tribology) Education and research. A report on the Present Position and Industry's Needs.* London: HMSO

Dowson D., Higginson G.R., 1961. New roller bearing lubrication formula. *Engineering* 192, 158–159

Dowson, D. and **Higginson, G.R.**, 1966, *Elastohydrodynamic Lubrication*, Second Edition, Pergammon Press, Oxford.

Dowson, D., 1998. *History of Tribology*. London: Professional Engineering Publishing.

Dowson, D., Ehret, P, 1999. Past, Present and future studies in elastohydrodynamics. *Proc. Instn Mech Engrs*, Vol. 213, Part. J, pp. 317 – 333.

Dowson, D., Higginson, G.R., 1959. A Numerical Solution to the Elastohydrodynamic Problem. *Journal of Mechanical Engineering Science*, Vol. 1,No. 1, pp. 6 – 15.

Dowson, D., Higginson, G.R., 1966, *Elastohydrodynamic Lubrication*, Second Edition, Pergammon Press, Oxford.

Dudley, Darle W., 1969. The Evolution of the Gear Art. *American Gear Manufacturers Association*.

Dyson, A., Naylor, H., Wilson, A.R., 1966. The measurement of oil-film thickness in elastohydrodynamic contacts. *Proc. Instn. Mech. Engrs, 1965-1966, Elastohydrodynamic Lubrication*, Vol. 180(3B), pp. 135-142.

Elcoate, C.D., 1996. *Coupled solution methods for the elastohydrodynamic lubrication problem.* PhD thesis, Cardiff University.

Elcoate, C.D., Evans, H.P., Hughes, T.G. 1997. Fully coupled elastohydrodynamic solution techniques for the analysis of real rough line contacts using finite element and finite difference models. *Tribology Series*, 32, pp. 27-36.

Elcoate, C.D., Evans, H.P., Hughes, T.G. and Snidle, R.W., 2001, "Transient Elastohydrodynamic Analysis of Rough Surfaces Using a Novel Coupled Differential Deflection Method", *Proc. Instn Mech. Engrs, Part J, Journal of Engineering Tribology*, Vol. 215, pp 319 – 337.

Elcoate, C.D., Evans, H.P., Hughes, T.G., 1999. On the coupling of the elastohydrodynamic problem. *Proceedings of the Institution of Mechanical Engineers*, Part C: *Journal of Mechanical Engineering Science*, 212 (4), pp. 307-318.

Encyclopedia of Tribology, 2013. [Eds. Wang,Q.J. , Chung Y.W.], Springer.

Engler, C., 1885. Ein Apparat zur Bestimmung der sogenannten Viskosität der Schmieröle. *Chem.* Z. 9, 189-90

Ertel, A.M., 1939. Hydrodynamic lubrication based on new principles. *Akad. Nauk. SSSR, Prikladnaya Matematika i Mekhanika* 3(2), 41–52

Evans, H.P., Hughes, T.G. 2000. Evaluation of deflection in semi-infinite bodies by a differential method. *Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science*, 214 (4), pp. 563-578.

Evans, H.P., Snidle, R.W., 1981b Inverse solution of Reynolds equation of lubrication under point contact elastohydrodynamic conditions. *J. Lubr. Technol*. 103, 539–546

Ewing, J. A., Humfrey, J. C. W., 1903. The fracture of metals under repeated alternations of stress. *Phil. Trans. Royal Society*, London, Vol. CC, pp. 241 - 250

Eyring, H. 1936. Viscosity, Plasticity, and Diffusion as Examples of Absolute Reaction Rates. *Journal of Chemical Physics*, 4, pp. 283–287.

Fairbairn, W., 1864. Experiments to determine the effect of impact, vibratory action and long continued changes of load on wrought iron girders. *Phil. Trans. R. Soc.* 154, 311-325

Fatemi, A., Socie, D. F., 1988. A critical plane approach to multiaxial fatigue damage including out-of-phase loading. *Fatigue Fract. Engng. Mater. Struct.*, Vol. 11, No. 3, pp. 149-165.

Field, J.V., Write, M.T., 1985. The early history of mathematical gearing *Endeavour, New Series*, 9(4), pp. 198-202.

Findley, W. N., 1953. Combined-stress fatigue strength of 76S-T61 aluminum alloy with superimposed mean stresses and corrections for yielding. [Technical report **NASA TN-2924**].

Findley, W. N., 1957. Fatigue of Metals under Combinations of Stresses, *Transactions of ASME*, vol. 79.

Flamant 1892. Compt. Rendus, *Paris*

Foord, C.A., Hammann, W.C., Cameron, A., 1968. Evaluation of lubricants using optical elasto-hydrodynamics. *ASLE Transactions* 11:31–43.

Forbs, R.J., 1963. The beginnings of technology and man. [*Technology in Western Civilization,* Vol. 1, *The Emergence of Modern Industrial Society. Earliest times to 1900,* Taton, R.], Oxford: Oxford U.P.

Furon, R., 1963. The dawn of science: Prehistoric beginnings. [*Ancient and Medieval Science: From Prehistory to A.D. 1450,* Taton, R.], London: Thames and Hudson.

Gatcombe, E.K., 1945. Lubrication characteristics of involute spur gears — a theoretical investigation. *Transactions of the ASME*, 67, pp. 177–185.

Gecim, B., Winer, W.O., 1980, Lubricant Limiting Shear Stress Effect on EHD Film Thickness, *Trans. ASME, Journal of Lubrication Technology*, Vol. 102, pp. 213 – 221.

Gohar R., 2001. *Elastohydrodynamics*, 2nd edition. Imperial College Press: London.

Gohar, R., Cameron, A., 1963. Optical measurement of oil film thickness under elasto-hydrodynamic lubrication. *Nature*, 200 (4905), pp. 458-459.

Gohar, R., Cameron, A., 1967. The mapping of elastohydrodynamic contacts. *ASLE Transactions*, 10 (3), pp. 215-225.

Gough, H. J., 1924. *The Fatigue of Metals*. London, Scott, Greenwood

Grant, George B., 1980. *A Treatise on Gear Wheels.* Twenty-first ed., Philadelphia Gear Works, Inc., 1980

Greenwood, J.A., 1969. Presentation of elastohydrodynamic film thickness results. *IMechE Sci.*, Vol. 11, Part. 2, pp. 128-132.

Greenwood, J.A., 2000, Two-dimensional Flow of a Non-Newtonian Lubricant, *Proc. Instn Mech. Engrs, Part J, Journal of Engineering Tribology*, Vol. 214, pp. 29 – 41.

Grubin, A.N., Vinogradova, I.E., 1949. Investigation of the Contact of machine Components. *Central Scientific Research Institute for Technology and Mechanical Engineering*, Book No. 30, Moscow [DSIR Translation No. 337]

Gumbel, L., Everling, E., 1925. *Reibung und Schmierung im Maschinenbau*, Krain.

Hamilton, G.M., Moore, S.L., 1971. Deformation and pressure in an elastohydrodynamic contact. *Proc. R. Soc. Lond. A.*, Vol 322

Hamrock, B.J., Dowson, D. 1974 Numerical evaluation of the surface deformation of elastic solids subjected to a Hertzian contact stress. *NASA Report TN D-7774.*

Hamrock B.J, Dowson D., 1976a. Isothermal elastohydrodynamic lubrication of point contacts. Part I. Theoretical formulation. *Transactions of the ASME, Journal of Lubrication Technology*, 98, pp. 223–229.

Hamrock B.J, Dowson D., 1976b. Isothermal elastohydrodynamic lubrication of point contacts. Part II. Ellipticity parameter results. *Transactions of the ASME, Journal of Lubrication Technology*, 98, pp. 375–383.

Hamrock B.J, Dowson D., 1977a. Isothermal elastohydrodynamic lubrication of point contacts. Part III. Fully flooded results, *Transactions of the ASME, Journal of Lubrication Technology*, 99, pp. 264–276.

Hamrock B.J, Dowson D., 1977b. Isothermal elastohydrodynamic lubrication of point contacts. Part IV. Starvation results *Transactions of the ASME, Journal of Lubrication Technology*, 99, pp. 15–23.

Hamrock, B.J. 1994. *Fundamentals of fluid lubrication.* McGraw-Hill.

Hamrock, B.J., Anderson, W.J., 1983. *Rolling-Element Bearing*. **NASA-RP-1105**

Hertz, H., 1881, "Über die Berührung fester Elastischer Körper", [On the Contact of Elastic Solids], *J. reine und angewandte Mathematik*, 92, 156 – 171, [English Translation: **Jones** and **Schott**, Eds., 1896, *Misc. Papers by H. Hertz*, Macmillan, London.].

Hertz, H., 1896. On the contact of elastic solids. *In Miscellaneous Papers by Heinrich Hertz (Eds D. E. Jones and G. A. Schott)*, Macmillan, London)

Holmes, M.J.A 2000. A fully coupled method for solving the Newtonian, steady state, isothermal elastohydrodynamic point contact problem. *Proc. SouthWales Inst. Engrs*.

Holmes, M.J.A., Evans, H.P., Hughes, T.G., Snidle, R.W., 2003, Transient Elastohydrodynamic Point Contact Analysis Using a New Coupled Differential Deflection Method Part 1: Theory and Validation, *Proc. Instn Mech. Engrs, Part J, Journal of Engineering Tribology*, Vol. 217, No. 4, pp. 289 – 304.

Holmes, M.J.A., Evans, H.P., Snidle, R.W., 2003. Analysis of microelastohydrodynamic and mixed lubrication effects in gear tooth contacts. *Proceedings of the ASME Design Engineering Technical Conference*, 4 B, pp. 903-912.

Holmes, M.J.A., Evans, H.P., Snidle, R.W., 2005. Analysis of mixed lubrication effects in simulated gear tooth contacts. *Journal of Tribology*, 127 (1), pp. 61-69.

Hou K.P., Zhu D., Wen S.Z., 1987. An inverse solution to the point contact EHL problem under heavy loads. *J. Tribol.* 109, 432–436

Houpert L.G., Hamrock B.J., 1986. Fast approach for calculating film thicknesses and pressures in elastohydrodynamically lubricated contacts at heavy loads. *J. Tribol*. 108, 411–420

Hsiao, H.-S.S., Hamrock, B.J., Tripp, J.H., 1998. Finite element system approach to EHL of elliptical contacts: Part I-Isothermal circular non-Newtonian formulation. *Journal of Tribology*, 120 (4), pp. 695-704.

Hu, Y.Z., Zhu, D., 2000. A full numerical solution to the mixed lubrication in point contacts. *J. Tribol.* 122, 1–9

Hughes, T.G., Elcoate, C.D., Evans, H.P., 2000. Coupled solution of the elastohydrodynamic line contact problem using a differential deflection method. *J. Mech. Eng. Sci*. 214, 585–598

Imrek, H., 2009, Performance improvement method for Nylon 6 spur gears, *Tribology International*, Vol. 42, pp. 503-510.

Johnson, K.L., 1970. Regimes of elastohydrodynamic lubrication. *J. Mech. Engng. Sci.*, Vol. 12(1), pp. 9-16.

Johnson K.L., 1985. *Contact mechanics.* Cambridge: The Press Syndicate of the University of Cambridge.

Johnston, G.J., Wayte, R., Spikes, H.A., 1991. The measurement and study of very thin lubricant films in concentrated contacts. *Tribology Transactions*, 34, pp. 187–194.

Jost, H.P., 1966. *Jost report.* See Department of Education and Science

Kannel, J.W., Dow, T.A., 1966. *Evaluation of contact stresses between rough-elastic and layered cylinders*. Butterworths

Kirk, M.T., 1962. Hydrodynamic lubrication of 'perspex'. *Nature* 194, pp. 965–966

Krantz, T.L., 2002, *The Influence of Roughness on Gear Surface Fatigue*, Ph.D. Thesis, Case Western Reserve University, Cleveland, Ohio.

Krantz, T.L., 2015, On the Correlation of Specific Film Thickness and Gear Pitting life, ***AGMA-14FTM21***.

Langer, B. F., 1937. Fatigue failure from stress cycles of varying amplitude. *Trans. ASME J. appl. Mech.* 59, A160-A162.

Larsson, R., 1997, Transient Non-Newtonian Elastohydrodynamic Lubrication Analysis of an Involute Spur Gear, *Wear*, Vol. 207, pp. 67 – 73.

Lee, K.M., Cheng, H.S., 1973, Effects of Surface Asperity on Elastohydrodynamic Lubrication", *NASA CR-2195, Washington, D.C*.

Lee, R.T., Hamrock, B.J., 1990, A Circular Non-Newtonian Fluid Model: Part I – Used in Elastohydrodynamic Lubrication, *Trans. ASME, Journal of Tribology*, Vol. 112, pp. 486 – 496.

Lewis, Wilfred, 1893. Investigations of the Strength of Gear Teeth. *Proc. of Engineers Club of Philadelphia*, pp. 16-23.

Litvin F.L., Fuentes, A. 2004. *Gear Geometry and Applied Theory.* Cambridge: The Press Syndicate of the University of Cambridge.

Liu, S., Wang, Q., Liu, G. 2000. A versatile method of discrete convolution and FFT (DC-FFT) for contact analyses. *Wear* 243(1-2), pp. 101-111

Lubrecht, A.A., 1987. *The numerical solution of elastohydrodynamic lubricated line and point contact problems using multigrid techniques*. Ph.D. Thesis, University of Twente, The Netherlands

Lubrecht, A.A., ten Napel, W.E., Bosma, R., 1986. Multigrid, an alternative method for calculating film thickness and pressure profiles in elastohydrodynamically lubricated line contacts. *Journal of Tribology*, 108 (4), pp. 551-556.

Majumdar, B.C., Hamrock, B.J., 1982. Extension of the patir-cheng flow simulation of a rough surface bearing to a compressible lubricant. *Journal of Mechanical Engineering Science*, 24 (4), pp. 209-214.

Mang, T., Bobzin, K., Bartels, T., 2011. *Industrial Tribology: Tribosystems, Friction, Wear and Surface Engineering, Lubrication.* WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

Martin, H.M. 1916. The lubrication of Gear Teeth. *Engineering*, 102, pp. 199-204.

McDiarmid, D. L., 1991. A general criterion for high cycle multiaxial fatigue failure. *Fatigue Fract. Engng. Mater. Struct*., 14, No. 4, pp. 429-453.

McDiarmid, D. L., 1994. A shear stress based critical-plane criterion of multiaxial fatigue failure for design and life prediction. *Fatigue Fract. Engng. Mater. Struct* 17, No. 12, pp. 1475-1484.

McEwen, E. 1949. Stress in elastic cylinders in contact along a generatrix. *Philosophical Magazine* 40

Meldahl, A., 1941, Contribution of Theory of Lubrication of Gears and of Stressing of Lubricated Flanks of Gear Teeth. *Brown Boveri Rev.*, Vol. 28, No. 11, pp. 374-382.

Miner, M. A., 1945. Cumulative damage in fatigue. *Trans. ASME J. appl. Mech*. 12, AI59-A164

Moes, H., 1966. Communications. Symposium on Elastohydrodynamic Lubrication. *Proc. Instn Mech Engrs*, Vol. 108, Part. 3J, pp. 244-245.

Moes, H., Bosma, R., 1972. Film thickness and traction in EHL at point contact. *Proceedings of 1972 IMechE Symposium on Elastohydrodynamic lubrication*, Vol. 108, Part. 3J, pp. 149-152.

Moore, H. F., Kommers, J. B., 1927. *The Fatigue of Metals*. New York, McGraw-Hill

NACA TN-2924, 1953. *Critical plane approaches for multiaxial fatigue damage asessment*. *Advances in Multiaxial Fatigue.* Findley, W. N.

NACA TN-3495, 1955. *Fatigue of materials under combined repeated stresses with superimposed static stresses.* Sines, G.

NASA CR-3241, 1980. *Thermal elastohydrodynamic lubrication of spur gears.* Wang, K. L.; Cheng, H. S.

NASA-RP-1105, 1983. *Rolling-Element Bearing*. Hamrock, B.J., Anderson, W.J.

NASA-RP-1152 / **AVRADCOM 82-C-15**, 1985. *Gearing*. Coy, John J., Townsend, Dennis P., Zaretsky, Erwin V.

NASA-RP-1255, 1991. *Fundamentals of fluid lubrication*. Hamrock, B.J.

NASA-TM-101444 / **AVRADCOM-TR 88-C-003**, 1989. *Dynamic loading of spur gears with linear or parabolic tooth profile modification*. Lin, H.H., Oswald, F.B., Townsend, D.P.

NASA-TM-103281 / **AVRADCOM-TR 90-C-023**, 1991. *Dynamic measurements of gear tooth friction and load*. Rebbechi, B., Oswald, F.B., Townsend, D.P.

NASA-TM-107279 / **ARL-TR-1165**, 1996. *Measurement of gear tooth dynamic friction*. Rebbechi, B., Townsend, D.P., Oswald, F.B.

NASA-TM-2005-213956 / **ARL-TR-3126**, 2005. *An Experimental Investigation of the Influence of the Lubricant Viscosity and Additives on Gear Wear*. Krantz, T., Kahraman, A.

NASA-TM-2005-213958 / **ARL-TR-3134**, 2005. *The Influence of Roughness on Gear Surface Fatigue*. Krantz, T.

NASA-TN-D-7261, 1973. *Pitting fatigue characteristics of AISI M-50 and super nitralloy spur gears*. Townsend, D. P., Chevalier, J. L., Zaretsky, E. V.

Newton, I., 1687. *Philosophiae Naturales Principia Mathematica*. Imprimatur S. Pepys, Reg. Soc. Praeses, 5 Julii 1686. [Translated by Motte, A. 1729, Revised and supplied with historical and explanatory appendix by Cajory, F., edited by Crawford, R.T. (1934), published by the University of California Press, Berkeley and LA (1966)]

Oh, K.P., Rohde, S.M., 1977. Numerical solution of the point contact problem using the finite element method. *Int. J. Numer. Meth. Eng*. 11, 1507–1518

Okamura, H., 1982. A contribution to the numerical analysis of isothermal elastohydrodynamic lubrication. *Proceedings of 9th Leeds-Lyon Symposium on Tribology*, Leeds, pp. 313–320

Palmgren, A., 1924. Die Lebensdauer von Kugellagern. *VDI-Zeitschrift* 68, 339-341

Papadopoulos, I.V., Davoli, P., Gorla, C., Filippini, M., Bernasconi, A., 1997 A comparative study of multiaxial high-cycle fatigue criteria for metals. *Int. J. Fatigue* 19, No. 3, pp. 219-235.

Patir, N., Cheng, H.S., 1978a. Effect of surface roughness orientation on the central film thickness in EHD contacts. *Proceedings of the 5th Leeds-Lyon Symposium on Tribology*, Leeds, pp. 15–21,

Patir, N., Cheng, H.S., 1978b, An Average Model for Determining Effect of Three Dimensional Roughness on Partial Hydrodynamic Lubrication", *Trans. ASME, Journal of Lubrication Technology*, Vol. 100, pp. 12 – 17.

Peppler W., 1938. *Druckübertragung an geschmiert zylindrischen Gleit und Walzflächen*. VDI Verlag: Berlin.

Petrov, N., 1883, "Friction in Machines and the Effect of the Lubricant", *Inzh. Zh., St-Peterb.*, Vol. 1, pp. 71 – 140; Vol. 2, pp. 227 – 279; Vol. 3, pp. 377 – 436; Vol. 4, pp.535 – 64.

Petrusevich, A.I., 1951. Fundamental Conclusions From the Contact-Hydrodynamic Theory of Lubrication, *Izv. Akad. Nauk. SSSR (OTN)*. Vol. 2, pp. 1 -11. [Ministry of Defence Translation No. 293]

Poon, C.Y., Sayles, R.S., 1994, Numerical Contact Model of a Smooth Ball on an Isotropic Rough Surface. *Trans. ASME, Journal of Tribology*, Vol. 116, No. 2, pp. 194 – 201.

Popova, E., Popov, V.L., 2014. On the history of elastohydrodynamics: The dramatic destiny of Alexander Mohrenstein-Ertel and his contribution to the theory and practice of lubrication. *Angew. Math Mech.*, **95**(7), pp.652-663

Prakash, J., Czichos, H., 1983. Influence of surface roughness and its orientation on partial elastohydrodynamic lubrication of rollers. *Journal of lubrication technology*, 105 (4), pp. 591-597.

Press, W.H., Teukolsky, S.A., Vetterling, W.T., Flannery, B.P. 1992. *Numerical recipes in Fortran*. Cambridge: Cambridge U.P..

Price, D.J. de Solla, 1959. On the origin of clockwork, perpetual motion devices and the compass. *United States National Museum Bulletein 218* , pp. 81-112.

Price, D.J. de Solla, 1974. Gears from the Greeks. The Antikythera Mechanism: A Calendar Computer from ca. 80 B. C. *Transactions of the American Philosophical Society New Series* 64 (7), pp. 1–70.

Ranger AP, Ettles CMM, Cameron A., 1973. Solution of the point contact elastohydrodynamic problem. *Proceedings of the Royal Society of London* A 1973; **346**, pp. 227–244

Rebbechi, B., Townsend, D.P., Oswald, F.B., 1996. Measurement of gear tooth dynamic friction. *American Society of Mechanical Engineers*, Design Engineering Division (Publication) DE 88 , pp. 355-363, 1996

Redwood, B., 1885. The Russian Petroleum Industry. *J. Soc. Chem. Ind.*. **4**, 70.

Redwood, B., 1886. On viscosimetry or viscometry. *J. Soc. Chem. Ind.*. **5**, 121-9.

Rennie, G., 1829. Experiments on the Friction and Abrasion of the Surfaces of Solids. *Philos. Trans. R. Soc. Lond.*, Vol. 34, part 1, pp. 143 – 170.

Reynolds, O., 1875. On rolling friction. *Phil. Trans. R. Soc.*, **166**, Pt1, 155.

Reynolds, O., 1886, "On the Theory of Lubrication and its Application to Mr. Beauchamp Tower's Experiments, Including an Experimental Determination of the Viscosity of Olive Oil", *Philos. Trans. R. Soc. Lond.*, Vol. 177, pp.157 – 234.

Roelands,, C.J.A., 1966, *Correlation Aspects of the Viscosity – Temperature – Pressure Relationship of Lubricating Oils*, PhD Thesis, Delft University, Delft [Druk, V.R.B., Groningen.].

Rohde, S.M., Oh, K.P., 1975. A unified treatment of thick and thin film elastohydrodynamic problems by using higher order element methods. *Proc. Roy. Soc. Lond*. A343, 315–331

Sadeghi, F., Sui, P.C., 1989. Thermal elastohydrodynamic lubrication of rolling/sliding contacts. *American Society of Mechanical Engineers*, pp. TRIB45 7p.

Schutz, W., 1996. A history of fatigue. *Engineering Fracture Mechanics* 54: 263–300.

Serensen, S. V., 1937. *The Endurance of Metals and the Design of Machine Components.* M.L.Gos Izd-vo Teckhnich Liter.

Serensen, S. V., 1940. Theory of strength under variable loading. Akademici Nauk Ukrainskoc, SSR. *Nukwoi pratsi Instytuta Budivel noi K. McKhani* [Abstract in Stahl und Eisen 60, 285].

Sharif, K.J., Kong, S., Evans, H.P., Snidle, R.W. 2001. Contact and elastohydrodynamic analysis of worm gears part 1: Theoretical formulation. *Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science*, 215 (7), pp. 817-830.

Sibley, L.B., Orcutt, F.K., 1961. Elasto-hydrodynamic lubrication of rolling-contact surfaces. *ASLE Transactions*, 4 (2), pp. 234-249.

Sines, G., 1955. *Fatigue of materials under combined repeated stresses with superimposed static stresses.* **NACA TN-3495**.

Sines, G., 1959. Behaviour of metals under complex static stresses. *Metal Fatigue*. McGraw Hill, [Eds. Waisman, J.L.], pp. 145-169.

Singer, C., Holmyard, E.J., Hall, A.R. (eds), 1954. *A History of Technology.* Vol I: *From Ancient times to Fall of Ancient Empires.* Oxford: Oxford U.P.

Socie, D. F., 1993. *Critical plane approaches for multiaxial fatigue damage asessment*. *Advances in Multiaxial Fatigue,* **ASTM STP 1191**. [Eds: D. L. Dowell and R. Ellis. Philadelphia, ASTM] pp. 7-36.

Spikes, H.A., 2006. Sixty years of EHL. *Lubrication Science*, **18**, pp. 265-291.

Spikes, H., Jie, Z., 2014, History, origins and prediction of elastohydrodynamic friction, *Tribology Letters*, **56** (1), pp. 1-25.

Sui, P.C., Sadeghi, F., 1991, Non-Newtonian Thermal Elastohydrodynamic Lubrication, *Trans. ASME, Journal of Tribology*, Vol. 113, pp. 390 – 397.

Tao, J., Hughes, T.G., Evans, H.P., Snidle, R.W., Hopkinson, N.A., Talks, M., Starbuck, J.M., 2003, Elastohydrodynamic Lubrication Analysis of Gear Tooth Surfaces From Micropitting Tests. *Trans. ASME, Journal of Tribology*, Vol. 125, No. 2, pp. 267 – 274.

Theyse, F.H.., 1966. Some aspects of the influence of hydrodynamic film formation on the contact between rolling/sliding surfacces. *Wear,* Vol. 9, pp. 41-59.

Thomas, J. W. 1995. Numerical Partial Differential Equations: Finite Difference Methods. *Texts in Applied Mathematics* 22. Berlin, New York: Springer-Verlag, pp 135-137

Timoshenko, S., Goodier, J.N. 1951. Theory of Elasticity, *3rd Edn. New York, London et al.:* McGraw-Hill

Tower, B., 1883, "First Report on friction Experiments (Friction of Lubricated Bearings)", *Proc. Instn Mech. Engrs*, Nov., pp. 632 – 659.

Tower, B., 1885, Second Report on friction Experiments (Experiments on the Oil Pressure in a Bearing), *Proc. Instn Mech. Engrs*, Jan., pp. 50 – 70.

Venner, C.H., 1991. *Multilevel solution of EHL line and point contact problems*. Ph.D Thesis, University of Twente, The Netherlands

Venner, C.H., Lubrecht, A.A., 2000. Multilevel methods in lubrication *Tribology Series*, 37, pp. i-xx+1-374.

Wang, Y., Li, H., Tong, J and **Yang, P.**, 2004, Transient Thermoelastohydrodynamic Lubrication Analysis of an Involute Spur Gear, *Tribology International*, No. 37, pp.773 – 782.

Wedeven, L.D., 1970. *Optical Measurements of Elastohydrodynamics in Rolling-Contact Bearings*. PhD thesis, University of London.

Westlake, F.J., Cameron, A., 1967, Fluid Film Interferometry in Lubrication Studies, *Nature*, Vol. 214, pp. 633 – 634.

Wohler, A., 1858. Bericht Fiber die Versuche, welche auf der körnig. Niederschlesisch-märkischen Eisenbahn mit Apparaten zum Messen der Biegung und Verdrehung von Eisenbahnwagenachsen während der Fahrt angestellt wurden. *Zeitschrift für Bauwesen*, VIII, 641- 652

Wohler, A., 1860. Versuche zur Ermittlung der auf die Eisenbahnwagenachsen einwirkenden Kräfte und die Widerstandsfähigkeit der Wagen-Achsen. *Zeitschrift für Bauwesen*. X, 583-616

Wohler, A., 1863. Ober Versuche zur Ermittlung der Festigkeit von Achsen, welche in den Werkstatten der Niederschlesisch-Märkischen Eisenbahn zu Frankfurt, angestellt sind. *Zeitschrift für Bauwesen*, XII1, 233-258

Wohler, A., 1866. Resultate der in der Central-Werkstatt der Niederschlesisch-M/irkischen Eisenbahn zu Frankfurt angestellten Versuche Fiber die relative Festigkeit von Eisen, Stahl und Kupfer. *Zeitschrift für Bauwesen*, XVI, 67-84

Wohler, A., 1867. Experiments on the strength of metals. *Engineering* 4, 160-161.

Wohler, A., 1870a. Uber die Festigkeits-Versuche mit Eisen und Stahl. *Zeitschrift für Bauwesen*, XX, 73-106

Wohler, A., 1870b. *Über die Festigkeitsversuche mit Eisen und Stahl*. Auf Anordnung des Ministers für Handel und Gewerbe. Berlin, Ernst und Korn

Zhu D., Cheng H.S., 1988. Effect of surface roughness on the point contact EHL. *J. Tribol.* 110, 32–37

Zhu D., Wang Q., 2011. Elastohydrodynamic lubrication: a gateway to interfacial mechanics – review and prospect, *J Tribol.* 133(041001):1–14

Zhu, D., 2007. On some aspects in numerical solution of thin-film and mixed EHL. *J. Eng. Tribol.* 221, pp. 561–579.

Zhu, D., Hu, Y.Z., 1999. The study of transition from full film elastohydrodynamic to mixed and boundary lubrication. The Advancing Frontier of Engineering Tribology, *Proceedings of the 1999 STLE/ASME*

Zhu, D., Wang, Q.J., 2011. Elastohydrodynamic lubrication: A gateway to interfacial mechanics - Review and prospect. *Journal of Tribology*, 133 (4), art. no. 41001