Contact Dynamics in the field of Elastohydrodynamic Lubrication

Y.H. Wijnant

This research project was supported by SKF Engineering & Research Centre B.V., Nieuwegein, the Netherlands, and was carried out at the University of Twente, Enschede, the Netherlands.

CIP-GEGEVENS KONINKLIJKE BIBLIOTHEEK, DEN HAAG

Wijnant, Ysbrand Hans

Contact Dynamics in the field of Elastohydrodynamic Lubrication / Ysbrand Hans Wijnant
PhD-thesis, University of Twente, Enschede, the Netherlands
ISBN: 90-36512239

Subject headings: elastohydrodynamic lubrication
multigrid/multilevel methods

CONTACT DYNAMICS IN THE FIELD OF ELASTOHYDRODYNAMIC LUBRICATION

PROEFSCHRIFT

ter verkrijging van
de graad van doctor aan de Universiteit Twente,
op gezag van de rector magnificus,
prof.dr. F.A. van Vught,
volgens besluit van het College voor Promoties
in het openbaar te verdedigen
op vrijdag 30 oktober 1998 te 15.00 uur

door

Ysbrand Hans Wijnant

geboren op 7 januari 1970

te 's-Hertogenbosch

DIT PROEFSCHRIFT IS GOEDGEKEURD DOOR:

PROMOTOR: PROF.IR. A.W.J. DE GEE
PROMOTOR: PROF.DR.IR. H. TIJDEMAN
ASSISTENT PROMOTOR: DR.IR. C.H. VENNER

Voor Lieke

Abstract

The quality of materials and lubricants which are used in rolling element bearings have increased significantly over the past decades. Consequently, for certain applications in the household and car industry, the load capacity and the "lifetime" of the bearing are guaranteed and other selection criteria, like the vibrational behaviour of the bearing, become important. The vibrational behaviour is directly related to the noise production of the application, since the bearing is an important link in the transmission path of vibrations from the shaft to the bearing housing. Noise production and its prevention are expected to become more important in the near future, as governmental regulations and customer demands become more restrictive.

Essential for the formulation of guidelines for reducing the noise production of the bearing and the transmission of vibrations through the bearing, is a thorough knowledge of the dynamics of the structural elements, e.g. the rolling elements, outer and inner ring. An additional prerequisite is a detailed understanding of the behaviour of the individual, elastohydrodynamically lubricated contacts that connect these elements. Such contacts and their relation to structural vibrations are the topics of this thesis.

A mathematical model is presented which describes the lubricated contact between a rolling element and the raceway under varying loads and free vibrations. This model is used as a tool to obtain the main dynamical characteristics of the contact: its stiffness and damping. In rolling element bearings, the elastic "footprint" is elliptically shaped and such contacts have therefore been included. Since the amount of lubricant used in bearings continues to decrease, so-called starved lubrication is accounted for too.

Analytical solutions of the problem are available in only a few, simplified cases. Numerical techniques have therefore been applied. The results for circular contacts form the basis of the discussion as to how the various parameters affect the film thickness, pressure, stiffness and damping. From this basis, we will study the effect of ellipticity and starved lubrication. Experimental results will also be presented. The good, qualitative agreement between theory and experiment shows that the model accurately describes the various phenomena observed in the contact region.

The model thus proves to be a valuable tool for understanding the behaviour of the contact. Aided by a dynamical analysis of the structural elements, this will enable us to predict and control the dynamic behaviour of the bearing.

Samenvatting

Naarmate, voor bepaalde toepassingen in de witgoed en auto-industrie, de levensduur en belastbaarheid van kogellagers voldoende is gewaarborgd, wordt de kwaliteit ervan in toenemende mate bepaald door het trillingsgedrag van het lager. Dit gedrag is direct gekoppeld aan de geluidsproductie van de toepassing waarin het lager wordt gebruikt. Immers, de overdracht van trillingen van de as naar de omkasting loopt via het lager. De geluidsproductie en de reduktie ervan worden steeds belangrijker naarmate de klant, maar ook de wet, striktere eisen stelt.

Essentieël voor het formuleren van richtlijnen voor reductie van de geluidsproductie en transmissie van trillingen door het lager, is een gedegen kennis van de dynamica van de structurele elementen, zoals de kogels, de buiten- en de binnenring. Tevens is een gedetailleerd begrip vereist van het gedrag van de individuele, zogenaamde, elastohydrodynamische contacten, die deze elementen verbinden. Dergelijke contacten en hun relatie tot trillingen van het lager zijn de onderwerpen van dit proefschrift.

Een mathematisch model wordt gepresenteerd dat het gesmeerde contact beschrijft tussen een enkele kogel en de "raceway". Dit model wordt gebruikt om de belangrijkste dynamische eigenschappen van het contact te bepalen: stijfheid en demping. Aangezien in kogellagers, het contactgebied tussen kogel en "raceway" elliptisch van vorm is, beschrijft het model ook dit soort contacten. Omdat vaak een gereduceerde toevoer van smeermiddel wordt toegepast, is tevens zogenaamde schrale smering gemodelleerd.

Analytische oplossingen voor het probleem zijn slechts in enkele, vereenvoudigde gevallen bekend. Numerieke methoden zijn daarom toegepast. De resultaten voor het cirkelvormig contact vormen de basis voor de discussie betreffende de filmdikte, druk, stijfheid en demping. Vanuit deze basis wordt de invloed van ellipticiteit en schrale smering bestudeerd. Ook experimentele resultaten worden beschreven. De goede, kwalitatieve overeenkomst tussen theorie en experiment laat zien dat het model de verschillende fenomenen die in het contactgebied worden waargenomen, nauwkeurig beschrijft.

Het model vormt, dientengevolge, een waardevol gereedschap waarmee het gedrag van het contact kan worden gesimuleerd. Met de beschreven resultaten en een analyse van het dynamisch gedrag van de strukturele elementen, kan het trillingsgedrag van kogellagers worden voorspeld en beïnvloed.

Contents

Nomenclature

a	Hertzian contact length	$a = (3fR/E')^{1/3}(2\kappa\mathcal{E}/\pi)^{1/3}$
A	force variation amplitude	
$\mathtt{A}$	force variation amplitude	
$\mathcal{A}$	amplitude of oscillation	
b	Hertzian contact width	$b = a/\kappa$
B	amplitude of oscillation	
c	Hertzian approach	$c = (a^2/(2R))(\mathcal{K}/\mathcal{E})$
c_Δ	Relaxation factor	
$\mathtt{C}$	constant of integration	
$\mathcal{C}$	constant of integration	
d	damping constant	[Ns/m]
D	damping constant	$D = dcu_s/(2af)$
E	modulus of elasticity, variable dimensionless dissipated energy	
E'	reduced modulus of elasticity	$\frac{2}{E'} = \frac{1-\nu_1^2}{E_1} + \frac{1-\nu_2^2}{E_2}$
$ERR(k, k-1)$	difference norm	
$\mathcal{E}$	elliptic integral (second kind)	
f	nominal load, right hand side	
f_f	final load	
f_i	initial load	
f_n	frequency	
$f(z)$	function	
F	variable, dimensionless load, force	
g_2	variable	$g_2 = 0$
g_3	variable	$g_3 = -1/800$
h	film thickness	
h_c	central film thickness	
h_l	lubricant layer	
$h_{l;inlet}$	inlet lubricant layer	
h_m	minimum film thickness	
h_{oil}	constant inlet lubricant layer	

h_t	timestep	
h_T	timestep	
h_x	meshsize in the x-direction	
h_X	meshsize in the X-direction	
h_y	meshsize in the y-direction	
h_Y	meshsize in the Y-direction	
$h.o.t$	higher order terms	
H	dimensionless film thickness	$H = h/c$
H_c	dimensionless central film thickness	$H_c = h_c/c$
H_{cff}	central fully flooded film thickness	
H_l	dimensionless h_l	$H_l = h_l/c$
$H_{l;inlet}$	dimensionless $h_{l;inlet}$	$H_{l;inlet} = h_{l;inlet}/c$
H_m	dimensionless minimum film thickness	$H_m = h_m/c$
H_{oil}	dimensionless h_{oil}	$H_{oil} = h_{oil}/c$
I_H^h	interpolation operator	
$\mathbb{I}_H^h$	higher order interpolation operator	
I_h^H	restriction operator	
k	Hertzian contact parameter	
K	kernel, linearized stiffness	
$\mathcal{K}$	elliptic integral (first kind)	
L	dimensionless lubricant parameter	$L = \alpha E'(\eta_0 u_s/(E' R_x))^{1/4}$
	differential operator	
L^h	difference operator	
m	mass	$m \equiv m_1$
m_i	mass of solid i	
M	dimensionless load parameter	$M = (f/(E' R_x^2))$
	number of cycles	$\cdot (E' R_x/(\eta_0 u_s))^{3/4})$
n	normal vector, number of unknowns	
n_t	component of n in the t-direction	
n_x	component of n in the x-direction	
n_y	component of n in the y-direction	
p	pressure	
p_h	maximum Hertzian pressure	$p_h = 3f/(2\pi ab)$
p_r	constant (Roelands)	$p_r = 1.96\ 10^8$
p_s	specific friction loss	
$p(L)$	function	
P	dimensionless pressure	$P = p/p_h$
P_s	dimensionless specific friction loss	
$\wp$	Weierstrass elliptic function	
q	integer	

$q(L)$	function	
r^h	residual	
R	reduced radius of curvature	$R^{-1} = R_x^{-1} + R_y^{-1}$
R_{x1}	radius of c. in the x-direction of solid 1	
R_{x2}	radius of c. in the x-direction of solid 2	
R_{y1}	radius of c. in the y-direction of solid 1	
R_{y2}	radius of c. in the y-direction of solid 2	
R_x	reduced radius of curvature in x-direction	$R_x^{-1} = R_{x1}^{-1} + R_{x2}^{-1}$
R_y	reduced radius of curvature in y-direction	$R_y^{-1} = R_{y1}^{-1} + R_{y2}^{-1}$
s	coordinate, variable, order of discretization	$s = (-3Z_0^2\zeta - Z_0)/20$
S	domain, contact area	
$\mathcal{S}$	shape factor	$\mathcal{S} = (\mathcal{E} - \kappa^2\mathcal{K})/(\mathcal{K} - \kappa^2\mathcal{K})$
t	time	
T	dimensionless time	$T = tu_s/(2a)$
T_f	time associated with final load	
$\mathtt{T}$	dimensionless time	$\mathtt{T} = t\omega$
$\mathtt{T}_e$	dimensionless period of oscillation	$\mathtt{T} \equiv 2\pi$
$\mathcal{T}$	dimensionless time	$\mathcal{T} = t(f/(mc))^{1/2}$
$\mathcal{T}_n$	dimensionless period of oscillation	$\mathcal{T}_n \approx 5.13$
u	function	
u_i	velocity of solid i	
u_s	sum velocity	$u_s = u_1 + u_2$
$u(y)$	function	
v	error function	
	elastic deformation	
w	integral transform	
W	work	
WU	work unit	
W_V	work associated with one V-cycle	
W_{FMG}	work associated with FMG algorithm	
x	coordinate in direction of flow	
$x_{a,b}$	domain boundaries	
x_c	position of the meniscus	
$\dot{x}_c$	velocity of the meniscus	
X	dimensionless coordinate	$X = x/a$
$X_{a,b}$	dimensionless domain boundaries	
X_m	variable	$X_m = X_i - X_k - h_X/2$
X_p	variable	$X_p = X_i - X_k + h_X/2$
y	coordinate perpendicular to x	
$y_{a,b}$	domain boundaries	
Y	dimensionless coordinate	$Y = y/b$

Y_m	variable	$Y_m = (Y_i - Y_k - h_Y/2)/\kappa$
Y_p	variable	$Y_p = (Y_i - Y_k + h_Y/2)/\kappa$
$Y_{a,b}$	dimensionless domain boundaries	
z	pressure viscosity index (Roelands),variable	$z = p_r\alpha/(\ln(\eta_0) + 9.67)$
Z	variable	$Z = \Delta^{1/2}$
Z_0	constant	$(5/2)^{1/3}$
α	pressure-viscosity coefficient	
$\bar{\alpha}$	dimensionless viscosity parameter	$\bar{\alpha} = \alpha p_h$
γ	constant	$\gamma^2 = -20$
Γ	$\mathcal{T}/\gamma$	
δ	mutual approach	
δ_∞	mutual approach at equilibrium	
δ_0	initial mutual approach	
$\dot{\delta}_0$	initial approaching velocity	
Δ	dimensionless mutual approach	$\Delta = \delta/c$
Δ_∞	Δ at equilibrium	$\Delta_\infty = \delta_\infty/c$
Δ_0	dimensionless initial mutual approach	$\Delta_0 = \delta_0/c$
$\dot{\Delta}_0$	dimensionless initial velocity	$\dot{\Delta}_0 = \dot{\delta}_0/c$
Δ_s	scaled dimensionless approach	
ι	imaginary unit	$\iota = \sqrt{-1}$
ϵ	coefficient	$\epsilon = \bar{\rho}H^3/(\bar{\eta}\lambda)$
ϕ	phase shift	
η	viscosity	
η_0	viscosity at ambient pressure	
$\bar{\eta}$	dimensionless viscosity	$\bar{\eta} = \eta/\eta_0$
θ	fractional film content	$\theta = h_l/h$
θ_c	θ *in* the starved region *closest* to the cavitation boundary	
κ	ellipticity ratio	$\kappa = a/b$
λ	dimensionless speed parameter,factor	$\lambda = (6u_s\eta_0 a)(c^2 p_h)$
Λ	dimensionless wavelength, wavelength	
μ	asymptotic reduction factor	
ν_i	Poisson's ratio solid i, number of pre- and post relaxations	
τ	variable,discretization error	$\tau = (z - Z_0)^{-1}$
ζ	variable	$\zeta = (Z - Z_0)^{-1}$
σ	variable	$\sigma = (-3Z_0^2\tau - Z_0)/20$
ρ	density	
ρ_0	density at ambient pressure	
$\bar{\rho}$	dimensionless density	$\bar{\rho} = \rho/\rho_0$
ω	angular frequency	

ω_r	frequency	
Ω_n	dimensionless natural frequency	$\Omega_n^2 = 4fa^2/(mu_s^2c)$
Ω_e	dimensionless excitation frequency	$\Omega_e = 2a\omega/u_s$
∇	gradient operator	
$\tilde{}$	approximation	
$\bar{}$	updated approximation	
$\hat{}$	associated with FAS	

Chapter 1

Introduction

In tribology, one studies the design, friction, wear and lubrication of interacting surfaces in relative motion. Perhaps unexpectedly, this is not a small field of research. In fact, it includes phenomena that occur in everyday life, from trying to walk on a slippery bathroom floor, but also "high-tech" applications like the bearings used in spacecraft. The understanding of the underlying mechanisms, requires a combination of knowledge from fields as varied as physics, chemistry, materials science, mechanical engineering and mathematics and makes tribology all the more interesting. In this thesis, we will restrict ourselves to a small, but important, tribological phenomenon: elastohydrodynamic lubrication.

The term elastohydrodynamically lubricated contacts, or EHL contacts, is used to describe the situation in which two solids are pressed against each other and the lubricant, present in the gap between the solids, prevents the two surfaces (the asperities) from touching. In addition, the contact pressures are so large that the elastic deformation of (one of) the solids, is of the order of the thickness of the lubricant film or larger. EHL contacts can be found in, e.g. gears and rolling element bearings, the latter being illustrated in Figure 1.1. In fact, lubrication and elastic deformation are what makes these mechanical devices work the way they do.

In recent years, the introduction of high quality materials, new manufacturing methods and lubricants has significantly increased the load capacity and lifetime of rolling element bearings. At present, bearings are expected to survive contact pressures up to 3 GPa and more, for an almost indefinite period of time. "Indefinite," in the sense that they outlive the application they are in. If the load increases, the thickness of the lubricant film separating the surfaces, decreases and if it becomes too small, asperties will touch. It is evident that, in such cases, the life expectancy of the bearing drops significantly and the surface topography becomes important. Also with respect

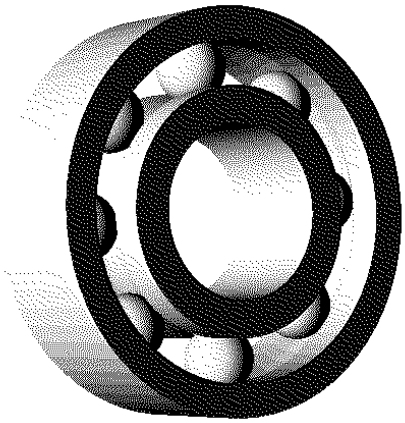


Figure 1.1: A rolling element bearing (source: J. Wensing).

to low friction, for which rolling element bearings were designed in the first place, both experimental and theoretical research is required.

Since in this thesis we will mainly restrict ourselves to theory, in the subsequent paragraphs we will discuss the developments in theoretical work only.

1.1 EHL and numerical solution

The theory of lubrication started over a century ago, when Reynolds described the flow of Newtonian fluids in narrow gaps, see [60]. At approximately the same time, Hertz [30] was the first to describe the elastic deformation of two, non-conforming solids in contact. It is interesting to notice that it took over 50 years before Ertel [20] and Grubin [24] combined both effects into what is now known as EHL. Since then, however, considerable progress has been made in theory as well as experiment.

Petrusevich [59] was the first to present numerical solutions that satisfied both the Reynolds equation and the equation describing the elastic deformation. Dowson and Higginson [18] presented numerical solutions for a wide range of the parameters involved and combined these solutions to the first film thickness formula, see [17].

In the last decades, a number of techniques have been proposed to solve the set of equations describing EHL contacts. For instance, direct methods have been proposed, in which the pressure distribution is determined for a given film thickness. As a next step, an iteration is performed, by means of Gauss-Seidel iteration or by Newton-Raphson algorithms, to satisfy the

Reynolds and the elasticity equation, simultaneously, see e.g. Hamrock and Dowson [28], Chittenden, Dowson, Dunn and Taylor [15, 16] and Okamura [56]. An inverse method, which includes the method of Dowson and Higginson, has also been proposed. This method was applied to circular contacts by Evans and Snidle [21]. The multigrid method, among others developed by Brandt [10], was applied to the EHL equations by Lubrecht [47]. Although, compared to all methods given above, multigrid significantly increased the rate of convergence, the evaluation of the elastic deformation integral still dominated calculation times. Hence, the use of the algorithms for practical studies was long hampered by CPU time. In addition, these methods became unstable at higher loads. Only by including the multilevel integration technique, as developed by Brandt and Lubrecht [12], and through, so-called, distributive relaxation techniques, was Venner [66] able to present an algorithm for both the line and point contact problem that was stable for high loads. The complexity of the algorithm was as low as $O(n \ln n)$, where n is the number of unknowns; a complexity way below the complexity of conventional algorithms. The combination of methods even allowed transient problems to be addressed. As an example, Figure 1.2 shows some film thickness contour plots, as calculated by Venner [66], which clearly shows the "power of the method" (263, 169 nodal points were used in the computation). The numbers M and L, as they appear in the plot, are the (dimensionless) load number and piezoviscosity number and will be discussed in Chapter 3. ΔH denotes the change in the values of the film thickness between consecutive contour lines.

Using these algorithms, and by means of fast computer hardware, the tribologist is now able to study the (transient) effects of surface topography, such as waviness, dents and bumps, see [69]. In addition, the algorithms allow for more complex rheological models to be included in the analysis, as well as analysis of temperature effects [9] and even starved or parched lubrication [14].

1.2 Vibrational behaviour

A separate development, which brings us to the topic of the present thesis, is the behaviour of the contact in relation to vibrations of the bearing, bearing noise and machine noise. The interest in this behaviour is not only imposed by customer demands, but also by governmental regulations. As load capacity and the lifetime of rolling element bearings are more and more guaranteed, their quality is increasingly determined by their acoustical or vibrational performance.

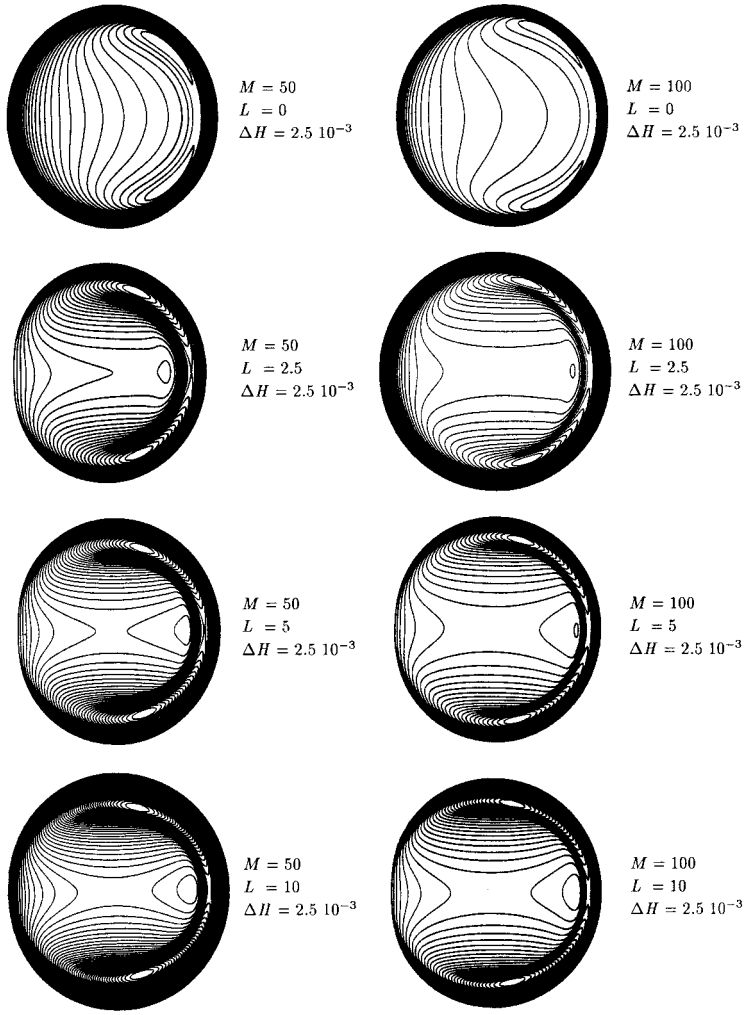


Figure 1.2: Film thickness contour plots calculated by Venner.

The determination of how vibrations are transmitted from the shaft to the bearing housing requires solving a system of equations, that govern the dynamics of the structural components in the bearing and those of the EHL contacts between these components. In addition, the dynamical behaviour of the bearing is affected by the dynamical behaviour of the application, and the equations describing the application should thus be included as well. Since it is already difficult to obtain the solution for a single EHL contact operating under dynamic loads, including all contacts would be a huge task. Not only would calculation times be beyond acceptable bounds, but, more importantly, the large number of variables in such a model would make a pa-

rameter analysis simply impossible. This would thus prevent the deduction of general guidelines to reduce noise.

One may circumvent solving all equations simultaneously, if the solution of a subset of the equations, which in our case will concern the behaviour of the EHL contact, is known in sufficient detail. Then, these solutions can be included in a bearing model, which effectively reduces the overall complexity. Obviously, such an approximation is only valid if the cross-coupling between the individual subsets of equations is small.

Two examples (modes) of how a bearing responds to vibrations, or the way by which vibrations are transmitted through the bearing, are shown in Figure 1.3. The deflections are greatly exaggerated for clarity. In fact, Figure 1.3 shows the results obtained by means of a finite element model of the bearing, which includes approximate relations describing the lubricated contacts, see Wensing [71, 72].

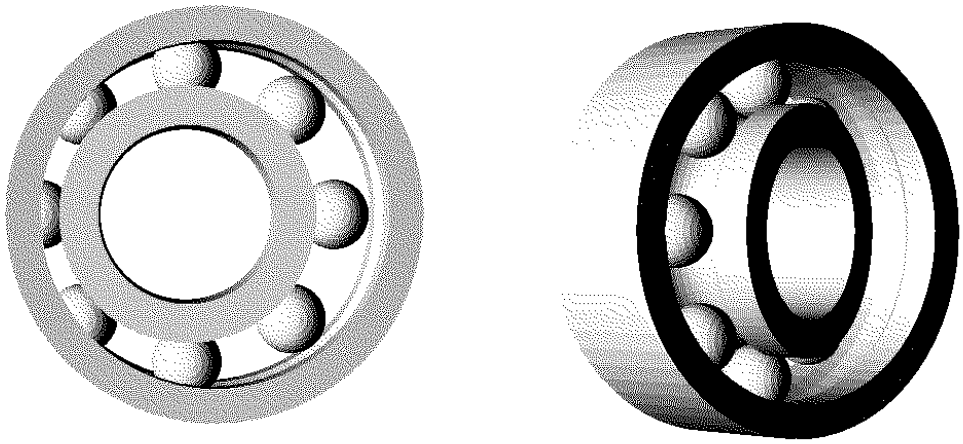


Figure 1.3: Two vibrating modes (source: J.Wensing).

The results obtained with these models provide valuable and accurate information regarding the dynamic response of the bearing. With the aid of additional information, obtained from experiments, these models can be used as a tool to reduce the noise production in all kinds of machines and equipment.

1.3 Problem definition

In this thesis, the main question we would like to address is: how do EHL contacts behave under dynamic loads or vibrations? An additional question is whether this behaviour can be described by simple, approximated relations. As a reference case, since it is the limiting case of high loads and vanishing lubricant supply, we start with the behaviour of a dry contact. Subsequently, the attention focuses on how lubrication affects the stiffness and the effects of the parameters involved. Can the stiffness be approximated by a simple relation? If the structural elements vibrate, how does this affect the film thickness and the pressure distribution? How do the different (dimensionless) variables affect the damping of the circular contact? Furthermore, we would like to understand in what way elliptically shaped contacts differ, with respect to stiffness and damping, from a circular contact. (An additional question is how ellipticity can be included most effectively in the algorithm for the circular contact.) If lubrication is insufficient to fully fill the gap, how does this affect the solution and should this effect be included in the analysis? Obviously, a final question to be answered is also, how realistic the simulated phenomena are.

1.4 Outline

The outline of this thesis is as follows. In Chapter 2, the dry contact situation is discussed. This discussion serves several purposes. Firstly, it forms an introduction to the topic and provides a reference for the interpretation and explanation of the results in the more complex lubricated case. Secondly, as the dry contact is the asymptotic limit of the lubricated contact at high loads and for vanishing lubricant supply, it has a practical significance of its own. Appropriate scales are derived, by which the equations are made dimensionless and analytical solutions for both sinusoidally varying loads and free vibrations are presented.

In Chapter 3, the theory associated with the lubricated case is discussed. The model accounts for ellipticity as well as starved lubrication. The dimensionless equations are presented and the theory of starved lubrication is discussed in some detail.

Chapter 4 focuses on the numerical methods used to solve the relevant equations. Firstly, the discretized equations are given, which approximate the equations given in the previous chapter. Secondly, some essential elements of the algorithm for circular contacts are discussed. Subsequently, we direct our attention to the modifications in the circular contact algorithm, which

are required as a result of ellipticity and starved lubrication.

In the next chapter, the circular contact problem is considered. As a reference, it shows the steady state pressure and film thickness distribution as a function of the parameters involved. From this steady state solution, the stiffness of the EHL circular contact is derived. In the remaining part of the chapter, emphasis is placed upon the transient solution for both free vibrations and sinusoidally varying loads. It is shown how the film thickness and pressure distribution change in time, and how the results relate to the solution for a dry contact. From the solutions for sinusoidally varying loads, the damping induced by the lubricant is calculated. Next, a curve-fit function is presented that approximates the calculated values, as a function of the operating conditions.

Subsequently, in Chapter 6, the effect of ellipticity is discussed. The steady state solution is presented for a number of values of the parameters involved. From this solution, the stiffness of the elliptical contact is deduced. The transient solution is used to reveal the effect of ellipticity on damping.

Chapter 7 is about starved lubrication. It is shown how the solution changes if the lubricant supply decreases, and that the phenomena, derived from continuity relations, indeed show up in the numerically obtained solutions. From the steady state solution, the effect of starvation on the stiffness is discussed, and it will be shown how damping is affected by starvation.

The concluding chapter presents a comparison between results obtained from experiments on a ball on disc apparatus, conducted by R. Larsson and P. Eriksson at Luleå University of Technology, Sweden, and results obtained with the mathematical model, see also [76]. It shows that the phenomena described in this thesis can indeed be observed in actual experiments.

Chapter 2

Theory: Dry contact

In this chapter the dynamic behaviour of two elastic solids in contact is discussed, where it is assumed that no lubricant is present in the gap between the solids. The Hertzian steady state solution will be discussed first. Subsequently, the quasi-static solution for sinusoidally varying loads, as well as the solution for which inertia is included are discussed.

2.1 Dry contact model

Figure 2.1 shows a rolling element running on a raceway. At this stage, it is assumed that there is no lubricant film separating the surfaces of the bodies, leaving the discussion of the lubricated problem for the next chapter.

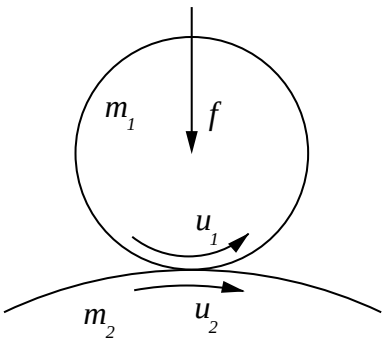


Figure 2.1: Rolling element on raceway.

Due to the applied load f both solids will deform elastically, forming an elastic footprint or contact area. If the contact area dimensions are small compared to the radii of curvature involved, locally, the *undeformed* surfaces

may well be approximated by paraboloids. (From Pythagoras' rule, the circular arc given in Figure 2.2 is described by $z = R_x - R_x\sqrt{1-(x/R_x)^2}$. A Taylor series around $x = 0$, yields $z \approx x^2/R_x$.) These approximate relations will be used to describe the undeformed surfaces.

Since the surfaces can be approximated by paraboloids, the shape of the gap between the undeformed solids, abbreviated to as the gap, can also be expressed by paraboloids. The reduced radii of curvature of the gap, R_x and R_y, are defined according to:

$$R_x^{-1} = R_{x1}^{-1} + R_{x2}^{-1}, \tag{2.1}$$

$$R_y^{-1} = R_{y1}^{-1} + R_{y2}^{-1}, \tag{2.2}$$

where R_{x1} and R_{x2} are the radii of curvature in the x-direction of, respectively, solid 1 and solid 2. R_{y1} and R_{y2} denote the radii of curvature in y-direction, see Figure 2.2

In addition, it is assumed that the x- and y-coordinates are aligned to the principal directions of the gap, i.e. the cross-term xy cancels in its description. The directions of the surface velocities u_1 and u_2 are assumed to coincide and to be aligned to the x-direction. The y-coordinate is taken perpendicular to the x-coordinate.

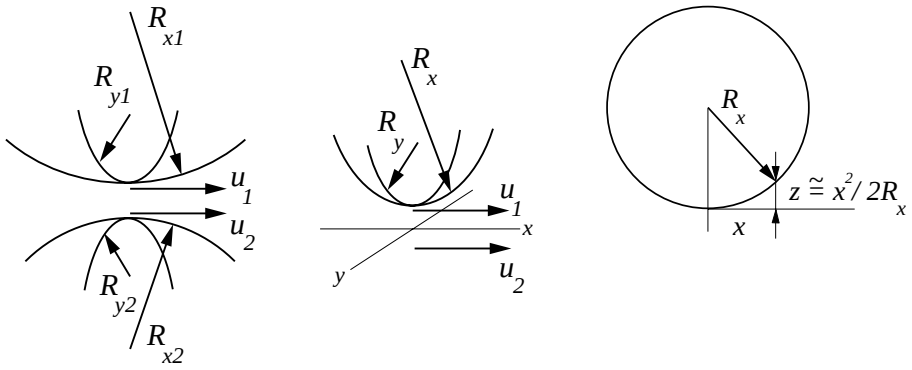


Figure 2.2: Definition of the reduced radii of curvature and approximation of the undeformed surface by a paraboloid.

If the elastic deformations are small, it is justified to apply linear elastic theory. Furthermore, if the contact length and width are small compared to the size of the contacting bodies (in all three dimensions), the elastic deformation of the solids is approximated accurately by the deformation of two semi-infinite bodies. These assumptions restrict the validity of the solution to concentrated contacts.

The elastic deformation of the solids can be calculated in the following way. Neglecting tangential stresses on the surfaces, the normal displacement $v(x, y)$ of the surface of a semi-infinite solid, due to a point load f in the origin is:

$$v(x,y) = \frac{1-\nu^2}{\pi E}\frac{f}{\sqrt{x^2+y^2}}, \tag{2.3}$$

where E is Young's modulus or the modulus of elasticity and ν denotes Poisson's ratio. This solution has been obtained by assuming that the elastic deformation is zero at an infinite distance from the load. The reader is referred to Love [45] and Johnson [33] for a detailed discussion on the subject.

The deformation due to a distributed normal pressure $p(x, y)$ on the surface can subsequently be obtained by integration:

$$v(x,y) = \frac{1-\nu^2}{\pi E}\iint_S \frac{p(x',y')dx'dy'}{\sqrt{(x-x')^2+(y-y')^2}}, \tag{2.4}$$

where S is the contact area.

The actual gap $h(x, y)$ between the two solids is now obtained by adding the elastic deformation of both solids to the parabolic approximation of the undeformed gap:

$$h(x,y) = -\delta + \frac{x^2}{2R_x} + \frac{y^2}{2R_y} + \frac{2}{\pi E'}\iint_S \frac{p\,(x',y')\,dx'dy'}{\sqrt{(x-x')^2+(y-y')^2}}. \tag{2.5}$$

Here, δ is the mutual distance of approach, also referred to as the mutual approach, of two remote points in the solids. The term remote is used to indicate that the elastic deformation is negligible at these points. Furthermore E' is the *reduced* modulus of elasticity:

$$\frac{2}{E'} = \frac{(1-\nu_1^2)}{E_1} + \frac{(1-\nu_2^2)}{E_2}, \tag{2.6}$$

where E_1 and E_2 denote the elasticity modulus of solids 1 and 2, respectively, and ν_1 and ν_2 denote Poisson's ratio for solids 1 and 2.

Since the surfaces do not penetrate, the gap is restricted to positive values only. Besides, only positive contact pressures exist if the contact is non-adhesive, as is assumed here. In addition, because a gap larger than zero implies that the pressure is also zero, and a pressure larger than zero causes the gap to close, the following complementarity condition holds:

$$\begin{aligned} h(x,y)\,p(x,y) = 0, \text{ with} \\ h(x,y) \geq 0 \text{ and } p(x,y) \geq 0. \end{aligned} \tag{2.7}$$

Given the approach δ, Equations 2.7 and 2.5 uniquely determine the pressure inside the contact area as well as the gap outside it. The determination of the approach δ itself requires an additional condition. For instance, it is usually required that the integral of the pressure over the contact area equals the applied load. However, δ may also follow from contact dynamics, as will be discussed in subsequent sections.

For smooth parabolic surfaces, the pressure in the contact region and the gap outside can be solved analytically. This solution was already presented by Hertz [30] over a century ago and, since it will be used extensively in the present work, will be discussed in the next section. For general, non-smooth, surfaces the solution can efficiently be obtained numerically using the multi-level algorithm presented by Lubrecht and Ioannides [48].

2.2 Hertzian solution

Hertz [30] was the first to present the solution to the dry contact problem. He derived the deformation of the surfaces, the resulting mutual approach of two remote points in the solids, as well as the displacements, strains and stresses. Hertz showed that the shape of the contact area is elliptical. The length of the elastic footprint on the line $y = 0$, denoted by $2a$, and the width of the footprint on $x = 0$, denoted by $2b$, can be expressed as:

$$a = \left(\frac{3fR}{E'}\right)^{1/3} \left(\frac{2\kappa\mathcal{E}}{\pi}\right)^{1/3} \tag{2.8}$$

and

$$b = a/\kappa, \tag{2.9}$$

where f is the applied load, $R = \left(R_x^{-1} + R_y^{-1}\right)^{-1}$ is the reduced radius of curvature, $\mathcal{E} \equiv \mathcal{E}(1-\kappa^2)$ is the complete elliptic integral of the second kind[1] and κ denotes the ellipticity.

The ellipticity κ only depends on the ratio between the reduced radii of curvature in the x- and y-directions, and is given implicitly by the following expression:

$$\frac{R_x}{R_y} = \kappa^2 \frac{\mathcal{K} - \mathcal{E}}{\mathcal{E} - \kappa^2\mathcal{K}}, \tag{2.10}$$

[1] The complete elliptic integrals of the first and second kind are defined by, respectively, $\mathcal{K}(m) = \int_0^{\pi/2} \frac{1}{\sqrt{1-m^2\sin^2(\psi)}} d\psi$ and $\mathcal{E}(m) = \int_0^{\pi/2} \sqrt{1-m^2\sin^2(\psi)}\, d\psi$, where m is the parameter of the elliptic integral. Note that $\mathcal{K}(0) = \mathcal{E}(0) = \pi/2$.

where $\mathcal{K} \equiv \mathcal{K}(1-\kappa^2)$ is the complete elliptic integral of the first kind[1]. Analytical curve-fits for κ are available, see Moes [51], but nowadays κ can easily be solved numerically, directly from Equation 2.10. Figure 2.3 shows κ as a function of R_x/R_y.

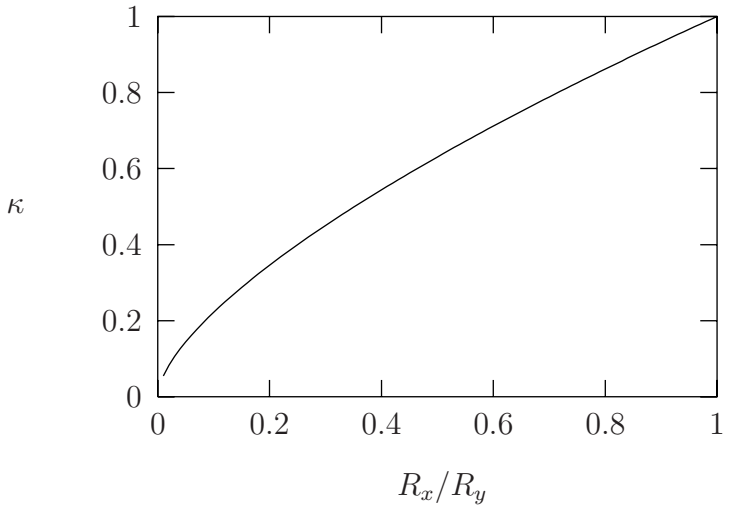


Figure 2.3: Ellipticity κ as a function of R_x/R_y

Hertz also showed that the pressure in the contact area is semi-ellipsoidal and the maximum normal pressure p_h is given by:

$$p_h = \frac{3f}{2\pi ab}. \tag{2.11}$$

The mutual approach c of two remote points in the solids was shown to be:

$$c = (f/k)^{2/3} = \frac{a^2}{2R}\frac{\mathcal{K}}{\mathcal{E}}, \tag{2.12}$$

with

$$k = \frac{2}{3}E'\sqrt{2R}\sqrt{\frac{\pi^2\mathcal{E}}{(2\kappa)^2\mathcal{K}^3}}. \tag{2.13}$$

c has been introduced to denote the mutual approach at the nominal load f, which for time-dependent problems is generally different from δ.

Equation 2.12 defines the so-called Hertzian spring, i.e. for any approach δ, the contact force can be obtained from $f = k\delta^{3/2}$. The corresponding stiffness equals $\frac{3}{2}k\delta^{1/2}$, showing that the Hertzian spring is a stiffening spring. The flexibility of the Hertzian contact, i.e. the inverse of the stiffness, is given by $\frac{2}{3}k^{-2/3}f^{-1/3} = \frac{2}{3}k^{-1}\delta^{-1/2}$.

For the sake of completeness, it should be mentioned that Hertz also presented the solution for two impacting solids. He assumed that no external load was applied, which allowed him to determine the contact pressures, the mutual approach and the total contact time as a function of the impacting velocity. Since we are interested in cases for which an external load is applied, Hertz' transient result will not be discussed here. Instead, we will present the more general equation of motion.

2.3 Contact dynamics

If the mass of the raceway is large compared to the mass of the rolling element, we can assume it to be a solid of infinite mass ($m_2 \rightarrow \infty$). From Newton's second law, the equation describing the relative motion of two remote points in the solids becomes:

$$m\frac{d^2\delta}{dt^2} + \iint_S p(x,y)dxdy = f(t), \tag{2.14}$$

where $m \equiv m_1$ is the mass of the rolling element and $f(t)$ is the applied load. Henceforth, $f(t)$ will denote a force that may vary in time, whereas f will be used to denote the average load. Equation 2.14, being a second order differential equation, requires two initial conditions, e.g. one for the initial mutual approach and one for the initial approaching velocity:

$$\begin{aligned} &\delta(t=0) = \delta_0 \text{ and} \\ &\frac{d\delta}{dt}(t=0) = \dot{\delta}_0. \end{aligned} \tag{2.15}$$

In dry contact situations, the force which results from a given approach δ (the second term in Equation 2.14), is simply the Hertzian spring $k\delta^{3/2}$. The equation of motion then reduces to:

$$m\frac{d^2\delta}{dt^2} + k\delta^{3/2} = f(t). \tag{2.16}$$

For negative δ the solids separate and the contact force is obviously zero. Hence, Equation 2.16 only holds for positive δ. For negative δ, it is valid provided the second term is set to zero.

The solution of this equation for arbitrary initial conditions and loads is easily obtained numerically. However, we will restrict ourselves to two particular sets of solutions, in view of what will be needed for the analysis of the lubricated contact. The first solution is the quasi-static solution for sinusoidal perturbations of the load. The second solution includes inertia and assumes a constant load.

2.3.1 Sinusoidally varying loads

If the load is perturbed sinusoidally, i.e. if $f(t) = f \cdot (1 + \mathtt{A}\sin(\omega t))$, and there is an instantaneous adaptation to the applied load, the approach δ directly follows from Equation 2.16, i.e. the solution is quasi-static. However, to facilitate a comparison between the lubricated case and the present dry contact case, the problem will be expressed in a time dependent form.

Before examining the solution, Equation 2.16 will be made dimensionless to reduce the number of independent variables. Introducing:

$$\begin{aligned} \Delta &= \delta/c, \\ \mathtt{T} &= t\omega, \end{aligned} \tag{2.17}$$

where c is the Hertzian approach at the nominal load f, see Equation 2.12, the equation reduces to:

$$\Delta^{3/2} = 1 + \mathtt{A}\sin(\mathtt{T}). \tag{2.18}$$

Note that Equation 2.18 does not hold for $\mathtt{A} > 1$. This is evident because in this case the load becomes negative, resulting in tensile stresses at the contacting surfaces. This already contradicts the Hertzian assumptions, since the solids separate whenever negative pressures occur as they are assumed to be non-adhesive.

Multiplying both sides of Equation 2.18 by $\dot{\Delta}$ and integrating with respect to $\mathtt{T}$ yields:

$$\frac{2}{5}\Delta^{5/2} = \Delta + \frac{1}{5}\left(1 + \mathtt{A}\sin\mathtt{T}\right)^{2/3}\left(-3 + 2\mathtt{A}\sin\mathtt{T}\right) + \mathtt{C}, \tag{2.19}$$

where $\mathtt{C}$ is a constant of integration. It represents the amount of deviation from equilibrium and ranges from $\mathtt{C} = 0$ for $\mathtt{A} = 0$, to $\mathtt{C} = 2^{2/3}$ for $\mathtt{A} = 1$. The individual terms in this equation may be interpreted as the potential energy due to the elastic deformation, the work done by the nominal load and the work done by the perturbation load, respectively. Obviously, such an energy equation could be derived only since, per definition, the elasticity model does not include any dissipative forces.

Figure 2.4 shows $\Delta(\mathtt{T})$ for $\mathtt{A} = 0.1$, 0.5 and 1.0. In addition, it illustrates how the solution can simply be constructed from the relation between the mutual approach and the applied load, i.e. $\Delta = F^{2/3}$, where F is the dimensionless load. Due to the non-linearity, the asymmetry of the solution around the equilibrium solution $\Delta = 1$ clearly shows. Note that in terms of the dimensionless variables, the frequency ω vanishes from the set of independent variables. Thus, the "period of oscillation" maps onto $\mathtt{T}_e = 2\pi$.

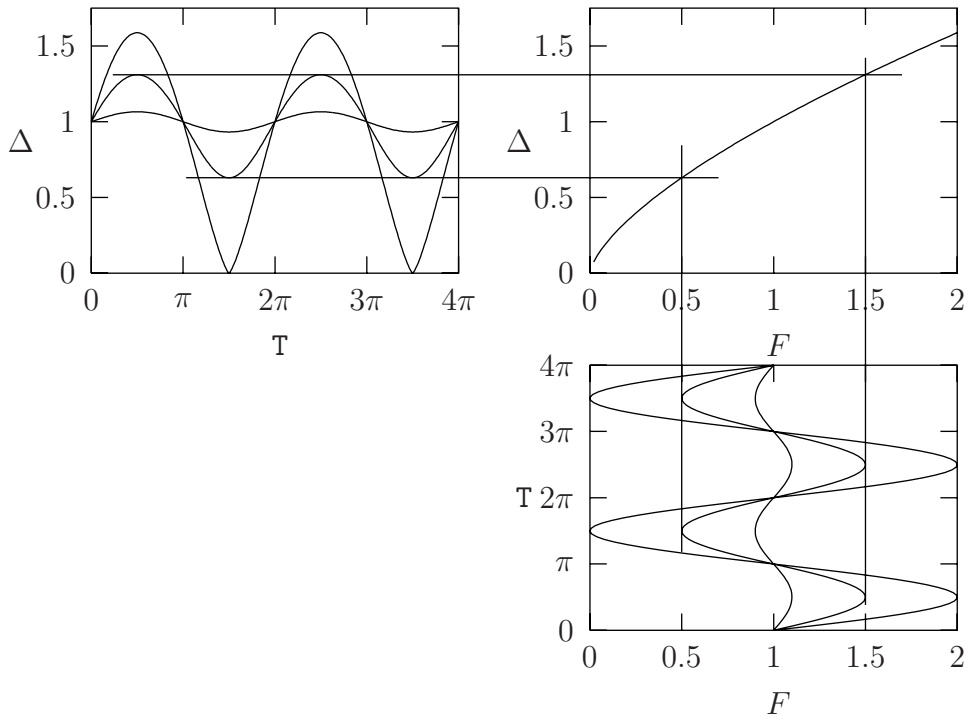


Figure 2.4: Response $\Delta(\mathtt{T})$ for sinusoidally varying loads for $\mathtt{A} = 0.1$, 0.5 and 1.0.

2.3.2 Inertia

A second model can be obtained by introducing forces due to the inertia of the rolling element. Assuming a constant load, substitution of the following variables makes the independent variables in the equation of motion vanish, i.e. substituting

$$\begin{aligned} \Delta &= \delta/c \text{ and} \\ \mathcal{T} &= t\sqrt{\frac{f}{mc}}, \end{aligned} \tag{2.20}$$

the equation of motion reduces to:

$$\frac{d^2\Delta}{d\mathcal{T}^2} + \Delta^{3/2} = 1. \tag{2.21}$$

In this case, initial conditions are required, which in terms of the new variables read:

$$\Delta(\mathcal{T} = 0) = \delta_0/c = \Delta_0 \text{ and}$$

$$\frac{d\Delta}{d\mathcal{T}}(\mathcal{T}=0)=\dot{\delta}_0\sqrt{\frac{m}{fc}}=\dot{\Delta}_0. \tag{2.22}$$

Just as for the equation for sinusoidally varying loads, Equation 2.21 can be integrated once, yielding the energy equation:

$$\frac{1}{2}\dot{\Delta}^2+\frac{2}{5}\Delta^{5/2}=\Delta+\mathcal{C}, \tag{2.23}$$

where $\dot{\Delta}$ denotes differentation with respect to $\mathcal{T}$ and $\mathcal{C}$ is a constant of integration. $\mathcal{C}$ ranges from $\mathcal{C}=0$ for oscillations, for which at some instant Δ becomes zero but otherwise remains positive, to $\mathcal{C}=-3/5$ for the steady state solution $\Delta=1$. The individual terms may now be interpreted as, respectively, the kinetic energy, the potential energy due to the elastic deformation and the work done by the unit force.

Analytical solutions of Equation 2.21, subject to the initial conditions given in Equation 2.22, can only be obtained in some special cases. One such solution is the linearized solution for small oscillations. A second solution is the one for which $\mathcal{C}=0$.

For small initial deviations from the equilibrium solution $\Delta=1$, Equation 2.21 can be linearized and the solution is readily obtained as:

$$\Delta=1+\mathcal{A}\cos\left(\sqrt{3/2}\mathcal{T}+\phi\right), \tag{2.24}$$

where $\mathcal{A}$ is the amplitude of the oscillation and ϕ denotes a phase shift. Both the amplitude and phase shift can be determined from the initial conditions. Thus, for small oscillations, the period of oscillation is $\mathcal{T}_n=2\pi/\sqrt{3/2}\approx 5.130$.

If $\mathcal{C}=0$, Δ will become zero at some time, otherwise it remains positive. This solution can be given in terms of the Weierstrass elliptic function $\wp(z;g_2,g_3)$, where $g_2=0$ and $g_3=-1/800$:

$$\Delta=400\wp\left(\mathcal{T}-\iota\sqrt{20}\mathcal{C}_2;0,-1/800\right)^2, \tag{2.25}$$

where ι is the imaginary unit and $\mathcal{C}_2$ is a constant of integration and represents a phase shift, see Appendix A. Between the zeroes of this function there are alternating regions where the solution is bounded and regions where it tends to infinity. Since, obviously, the bounded solution is the appropriate one, the period of oscillation $\mathcal{T}_n$ can be shown to be

$$\mathcal{T}_n=2\int_{\frac{1}{10}\sqrt[3]{\frac{5}{2}}}^{\infty}\left(4\sigma^3+1/800\right)^{-1/2}d\sigma\approx 5.383. \tag{2.26}$$

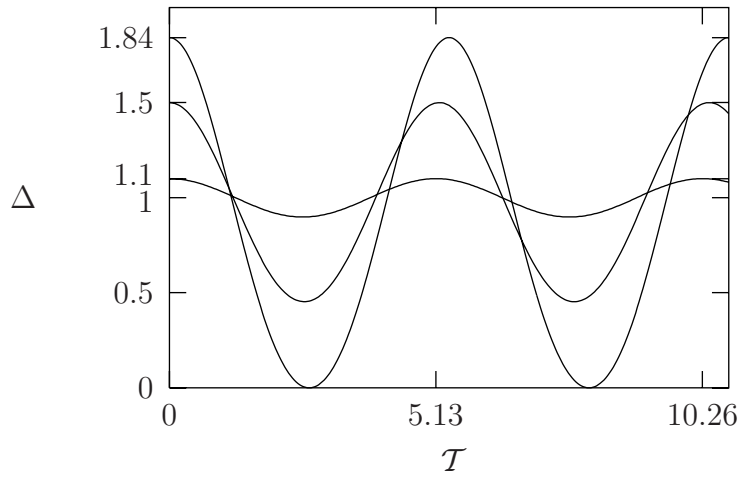


Figure 2.5: $\Delta(\mathcal{T})$ for $\Delta_0 = 1.1$, 1.5 and $(5/2)^{2/3}$ and $\dot{\Delta}_0 = 0$.

The maximum mutual approach is $\Delta_{max} = (5/2)^{2/3} \approx 1.842$.
The solution for arbitrary values of $\mathcal{C}$ and $\mathcal{C}_2$ can easily be solved numerically. Figure 2.5 shows Δ for different values of Δ_0 where $\dot{\Delta}_0 = 0$. One observes that the asymmetry is less profound than in the case of sinusoidal perturbations of the load.

From the linearized, the numerical and the Weierstrass solution, it can be seen that the period of oscillation only slightly increases with $\mathcal{C}$, i.e. it is virtually independent of the amplitude of the oscillation. In fact, the difference in natural frequency between the linearized and the Weierstrass solution is only about 4.7%. Hence, one may conclude that the non-linearity in the amplitude of the oscillation is only small. This result will be shown to be a useful reference in the discussion on lubricated contacts.

Chapter 3

Theory: Lubricated contact

In this chapter the model, as presented in the previous chapter, is extended to lubricated contacts. The additional equation describing the lubricant flow is given, as well as the cavitation condition and two empirical relations concerning the viscosity and density of the lubricant. Furthermore, using appropriate scales, dimensionless equations are presented.

A further extension that will be described is starved lubrication, i.e. when lubrication is insufficient to fully fill the gap between the contacting bodies. For that purpose, the equation describing the lubricant flow is modified and so-called continuity conditions will be discussed. The chapter ends by presenting the modified dimensionless equations.

3.1 EHL theory

The equations describing the fully flooded EHL contact are, among others, the Reynolds equation describing the flow in the gap, the equation for the gap or film thickness and the equation of motion. Each equation is discussed below.

Reynolds equation

In EHL contacts, the thickness of the lubricant film between the contacting surfaces is generally small compared to the characteristic dimensions of the contact area. Moreover, the forces due to inertia in the lubricant are negligibly small compared to the viscous forces. In this case, the Navier Stokes equations, describing the flow of a Newtonian fluid, reduce to a single equation that relates the pressure in the gap to its geometry. This equation is known as the Reynolds equation, see Reynolds [60], named after the person

who first derived it. If the velocities of both surfaces are in the x-direction only and fully flooded conditions apply, the Reynolds equation reads:

$$\frac{\partial}{\partial x}\left(\frac{\rho h^3}{\eta}\frac{\partial p}{\partial x}\right)+\frac{\partial}{\partial y}\left(\frac{\rho h^3}{\eta}\frac{\partial p}{\partial y}\right)=6u_s\frac{\partial \rho h}{\partial x}+6\rho h\frac{\partial u_s}{\partial x}+12\frac{\partial \rho h}{\partial t}, \tag{3.1}$$

where $p = p(x, y, t)$ is the hydrostatic pressure above ambient pressure and $h = h(x, y, t)$ is the gap between the contacting surfaces. The viscosity and the density of the lubricant are denoted by $\eta = \eta(x, y, t)$ and $\rho = \rho(x, y, t)$, respectively. Finally, the sum velocity u_s equals the sum of the velocities of the two surfaces, i.e. $u_s = u_1 + u_2$.

In the present study, it is assumed that u_s is constant. Hence, the so-called stretch effect (the second term on the right hand side of Equation 3.1) disappears. The remaining terms on the right hand side are generally referred to as the wedge term and the squeeze term according to the mechanism of pressure generation they represent. Flow induced by the wedge term is also known as Couette flow. Pressure induced flow, as on the left hand side of the equation, is usually termed Poiseuille flow.

The Reynolds equation requires conditions on the boundary enclosing the spatial domain. Since p is the pressure above ambient pressure, the pressure on the boundary of the domain is set equal to zero:

$$p\,(x, y, t) = 0; \quad \forall (x, y, t) \in \partial S, \tag{3.2}$$

where ∂S denotes the boundary of the spatial domain.

Cavitation

If the pressure in the lubricant becomes smaller than the vapour pressure, the lubricant cavitates. In the cavitated region, the pressure in the lubricant remains constant and equal to the vapour pressure, when the void that is formed is filled by the lubricant vapour. The lubricant pressure equals the ambient pressure if the surrounding medium occupies the void. This effect is not accounted for in the Reynolds equation and, without any further provisions, the Reynolds equation predicts negative pressures. Since the vapour pressure and the ambient pressure are generally small compared to the pressure occurring in the contact, it is justified to assume that the lubricant cavitates at zero pressure. Hence, cavitation is taken into account by restricting the pressure to positive values only, i.e.:

$$p\,(x, y, t) \geq 0; \quad \forall (x, y, t) \in S, \tag{3.3}$$

where S is the domain.

Film thickness equation

If the undeformed surfaces are approximated by paraboloids and the elastic deformation of the surfaces is accounted for, the gap between the surfaces can be described by:

$$h(x,y,t) = -\delta(t) + \frac{x^2}{2R_x} + \frac{y^2}{2R_y} + \frac{2}{\pi E'} \iint_S \frac{p\left(x',y',t\right) dx'dy'}{\sqrt{\left(x-x'\right)^2 + \left(y-y'\right)^2}}, \quad (3.4)$$

as was given in Equation 2.5.

Viscosity/Compressibility

The viscosity and the density of the lubricant are generally pressure dependent. Particularly at the high pressures occurring in EHL contacts, this dependence cannot be neglected.

A simple relation that is frequently used to describe the increase of the viscosity with pressure is the so-called Barus relation, see Barus [5]:

$$\eta = \eta_0 e^{\alpha p}, \quad (3.5)$$

where η_0 is the viscosity at ambient pressure and α is the pressure-viscosity coefficient. Experience has shown that, for most lubricants, the Barus relation overestimates the viscosity. Roelands [61] proposed a relation which is valid up to pressures of approximately 1 GPa. For isothermal conditions Roelands' equation reduces to:

$$\eta = \eta_0 \exp\left(\ln\left(\eta_0\right) + 9.67\right)\left(-1 + \left(1 + p/p_r\right)^z\right), \quad (3.6)$$

where $p_r = 1.96\,10^8$ and z is the pressure-viscosity index. The index z ranges from approximately 0.1 to 1.5 for various lubricants, see [61]. In the present work Roelands' relation will be used with $z = 0.67$.

From:

$$\alpha \equiv \frac{1}{\eta}\left(\frac{d\eta}{dp}\right)_{p=0}, \quad (3.7)$$

α, p_r and z are related according to:

$$z = \frac{p_r \alpha}{\ln\left(\eta_0\right) + 9.67}. \quad (3.8)$$

The density of most mineral oils increases with pressure. Often this effect is small and can be neglected. However, at the large pressures that occur in

EHL contacts, this is no longer true. An empirical relation that describes this increase was proposed by Dowson and Higginson [17]:

$$\rho = \rho_0 \frac{0.59\,10^9 + 1.34p}{0.59\,10^9 + p}, \tag{3.9}$$

where ρ_0 is the density at ambient pressure. According to Equation 3.9, the density can increase with approximately 30%.

Equation of motion

If the mutual approach δ is given, the Reynolds equation and the film thickness equation form a well posed problem, i.e. there is one unique solution for the pressure and film thickness. However, generally δ is unknown too. For example for steady state conditions, our particular interest is in the value of δ for which the contact force, i.e. the integral of the pressure over the contact area, equals the applied load f. In this case, the equation of motion reduces to the equation:

$$\iint_S p(x,y)dxdy = f. \tag{3.10}$$

In the transient case, the mutual approach is governed by the general equation of motion as is given in Section 2.3. Assuming sinusoidal perturbations of the load and including inertia terms, this equation reads:

$$m\frac{d^2\delta}{dt^2} + \iint_S p\,(x,y,t)\,dxdy = f \cdot (1 + A\sin(\omega t)). \tag{3.11}$$

As for the dry contact case, this equation will not be used in its complete form, i.e., either the inertia term or A is set to zero.

3.1.1 Dimensionless equations

The number of independent parameters presented in the above equations can be greatly reduced by the introduction of similarity groups. These groups are preferably dimensionless but this is not essential. A powerful tool to obtain similarity groups for a given set of equations is the so-called *optimum similarity analysis*, as presented by Moes [50]. This method ensures that the number of independent parameters is minimal.

In the present study, the similarity groups are based on the Hertzian solution. This leads to a number of independent similarity groups that equals the minimum number obtained by the optimum similarity analysis. These specific groups are preferred because the value of most of the variables will

be around unity, which minimizes the loss of accuracy in numerical computations. Moreover, as will become clear, it facilitates the interpretation of the solution.

Hence, after substitution of the following variables:

$$\begin{aligned} X &= x/a, \quad & Y &= y/b, \\ H &= h/c, \quad & \Delta &= \delta/c, \\ P &= p/p_h, \quad & T &= tu_s/\left(2a\right), \\ \bar{\rho} &= \rho/\rho_0, \quad & \bar{\eta} &= \eta/\eta_0, \end{aligned}$$

the Reynolds equation reduces to:

$$\frac{\partial}{\partial X}\left(\frac{\bar{\rho}H^3}{\bar{\eta}\lambda}\frac{\partial P}{\partial X}\right)+\kappa^2\frac{\partial}{\partial Y}\left(\frac{\bar{\rho}H^3}{\bar{\eta}\lambda}\frac{\partial P}{\partial Y}\right)=\frac{\partial \bar{\rho}H}{\partial X}+\frac{\partial \bar{\rho}H}{\partial T}, \tag{3.12}$$

where λ is defined by:

$$\lambda=\frac{6u_s\eta_0 a}{c^2 p_h}. \tag{3.13}$$

The boundary condition is $P(X,Y,T)=0;\ \forall(X,Y,T)\in\partial S$, where ∂S now denotes the enclosure of the dimensionless spatial domain S. The cavitation condition poses the condition $P(X,Y,T)\geq 0;\ \forall(X,Y,T)\in S$.

The dimensionless film thickness, in terms of the given variables, reads:

$$H(X,Y,T)=-\Delta(T)+\mathcal{S}X^2+\left(1-\mathcal{S}\right)Y^2+\frac{1}{\pi\mathcal{K}}\iint_S\frac{P\left(X',Y',T\right)dX'dY'}{\sqrt{\kappa^2\left(X-X'\right)^2+\left(Y-Y'\right)^2}}, \tag{3.14}$$

where $\mathcal{S}=\mathcal{S}(\kappa)$ is a shape factor which only depends on the geometry of the undeformed solids. It is defined according to:

$$\mathcal{S}\left(\kappa\right)=\frac{\mathcal{E}-\kappa^2\mathcal{K}}{\mathcal{K}-\kappa^2\mathcal{K}}, \tag{3.15}$$

where $\lim_{\kappa\rightarrow 1}\mathcal{S}(\kappa)=1/2$. The reader is referred to Section 2.2 for the definition of the complete elliptic integrals $\mathcal{K}$ and $\mathcal{E}$. Figure 3.1 shows the shape factor $\mathcal{S}(\kappa)$ for $0<\kappa\leq 1$. Values of the shape factor for $\kappa>1$ can be obtained from its value for $\kappa<1$ using $\mathcal{S}(\kappa)=1-\mathcal{S}(1/\kappa)$. The function $0.5\kappa^{3/11}$ has been included in Figure 3.1. It approximates $\mathcal{S}(\kappa)$ within 0.5%, for $0.1<\kappa\leq 1$. Due to the small values of $\mathcal{S}(\kappa)$ for $\kappa<1$, the percentage becomes less favourable.

The Barus relation describing the dependence of viscosity on pressure reduces to:

$$\bar{\eta}=e^{\bar{\alpha}P}, \tag{3.16}$$

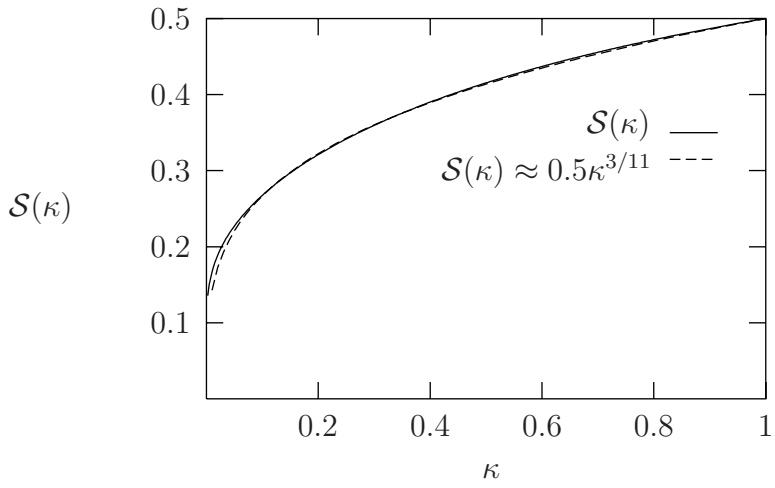


Figure 3.1: Shape factor $\mathcal{S}(\kappa)$.

where $\bar{\alpha}$ is defined as:

$$\bar{\alpha} = \alpha p_h. \tag{3.17}$$

Roelands' equation reduces to:

$$\bar{\eta} = \exp\left\{\left(\frac{\alpha p_r}{z}\right)\left(-1 + (1 + (P p_h/p_r))^z\right)\right\}, \tag{3.18}$$

whereas the Dowson and Higginson relation, in terms of the dimensionless variables, reads:

$$\bar{\rho} = \frac{0.59\,10^9 + 1.34 P p_h}{0.59\,10^9 + P p_h}. \tag{3.19}$$

Finally, Equation 3.10 reduces to:

$$\frac{3}{2\pi}\int_S P(X,Y)dXdY = 1, \tag{3.20}$$

whereas the dimensionless equation of motion reads:

$$\frac{1}{\Omega_n^2}\frac{d^2\Delta}{dT^2} + \frac{3}{2\pi}\int_S P(X,Y,T)dXdY = 1 + A\sin\left(\Omega_e T\right). \tag{3.21}$$

In this equation Ω_n is the dimensionless natural frequency defined as:

$$\Omega_n^2 = \frac{4fa^2}{mu_s^2 c} = \frac{8fR}{mu_s^2}\frac{\mathcal{E}}{\mathcal{K}}, \tag{3.22}$$

and Ω_e is the dimensionless excitation frequency defined as:

$$\Omega_e = \frac{2a\omega}{u_s}, \tag{3.23}$$

where the reader is reminded that f denotes the nominal load.
From the variables introduced, it can be concluded that the steady state solution for fully flooded conditions only depends on the parameters λ, $\bar{\alpha}$ and κ. The solution for sinusoidally varying loads depends also on the frequency Ω_e and the amplitude A. For free vibrations, the solution further depends on the frequency Ω_n and the initial conditions Δ_0 and $\dot{\Delta}_0$.

The set of parameters λ, α and κ are not the only set that is independent and minimal in the sense as described earlier. In fact, there are an infinite number of sets which are independent and minimal, provided they are unique combinations of λ, α and κ. One such set has been derived by Moes [50], and this set is often used within the tribology community. The set is based on the optimum similarity analysis. Although all calculations are performed using the parameters based on the Hertzian solution, this set will be used to identify the specific cases.

Moes' parameters are the load parameter M and lubricant parameter L. They are defined according to:

$$M = \frac{f}{E' R_x^2} \left(\frac{E' R_x}{\eta_0 u_s} \right)^{3/4} \tag{3.24}$$

and

$$L = \alpha E' \left(\frac{\eta_0 u_s}{E' R_x} \right)^{1/4} . \tag{3.25}$$

The parameters λ and α are related to M, L and κ according to:

$$\lambda = \left(\frac{128\pi^3}{3M^4} \right)^{1/3} \left(\frac{16\pi \left(\mathcal{E} - \kappa^2 \mathcal{K} \right)^5}{\kappa^4 \left(1 - \kappa^2 \right) \mathcal{K}^6} \right)^{1/3} , \tag{3.26}$$

and

$$\bar{\alpha} = \frac{L}{\pi} \left(\frac{3M}{2} \right)^{1/3} \left(\frac{\pi^2 \kappa \left(1 - \kappa^2 \right)^2}{16 \left(\mathcal{E} - \kappa^2 \mathcal{K} \right)^2} \right)^{1/3} . \tag{3.27}$$

It is noted that for $\kappa = 1$, one should take the principal values of these expressions ($\lim_{\kappa \to 1}$). In that case, the factors containing the elliptic integrals converge to 1.

3.2 Starved lubrication

The Reynolds equation presented in the previous section is only valid for fully flooded conditions, which apply when the lubricant completely fills the inlet region of the domain. In real applications, however, the amount of lubricant

in the inlet is often insufficient to establish these fully flooded conditions. The lubricant only partly fills the gap, whereas the remaining part is occupied by vapour or the surrounding medium. As a result, the pressure equals the ambient pressure at that location. In addition to these global scale features, the lubricant may also cavitate locally, when surface features like dents or bumps are present. A general term indicating these phenomena is starved lubrication.

Since, under the assumptions made in Section 3.1, the pressure in the cavitated area equals the ambient pressure, one may consider cavitation and starvation to be different words representing the same thing.

For starved lubricated contacts, one may thus distinguish a starved region and a pressurized region. For the two-dimensional flow considered here, the starved region surrounds the pressurized region. (Obviously, the starved region may enclose a number of pressurized regions, which in their turn may enclose starved regions.) This is shown in Figure 3.2. The cross section $A-A$ is also included in Figure 3.2.

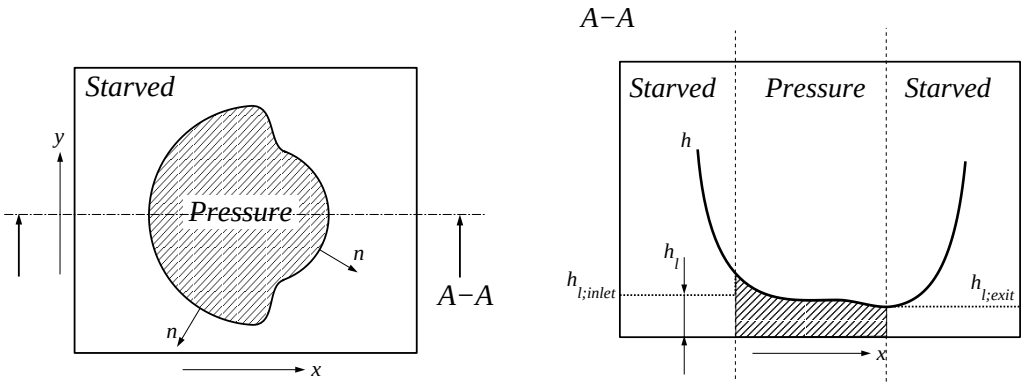


Figure 3.2: The pressurized and starved regions, and the cross section.

In the pressurized region the lubricant fully fills the gap and the Reynolds equation can be used to determine the pressure in the film. Note that the position of the boundary between the pressurized region and the starved region, the so-called meniscus, is unknown. Hence, this boundary is a free boundary, which may even change in time, and its location is part of the solution to be obtained. Compared to fully flooded conditions, this adds to the complexity of the problem.

In recent years, a number of authors have analysed starved lubrication. The first reports on cavitation were the contributions of Stieber [63], Jakobsson and Floberg [32] and Olsson [65]. Jakobsson and Floberg introduced the

concept of a fractional film content and derived continuity relations which hold on the meniscus. Olsson extended the condition posed by Jakobsson and Floberg to the transient case. The derivation of these continuity relations will be given in a subsequent section. An algorithm for the automatic determination of the meniscus in slider bearings, which is basically the algorithm that is adopted here, was presented by Elrod [19]. However, numerical results for steady state EHL contacts were only recently reported in Chevalier [14].

3.2.1 Modified Reynolds equation

Although the cavitation condition, Equation 3.3, is sufficient for simulating a single cavitation boundary, it cannot simulate pressure generation from a starved region. This requires the introduction of a fractional film content θ. The fractional film content is defined as the ratio between the height of the lubricant layer or oil film, denoted by h_l, and the total gap h at that location:

$$\theta = \frac{h_l}{h}, \tag{3.28}$$

see Figure 3.2. It is noted that, contrary as the Figure might suggest, the lubricant may be distributed unevenly across the gap.

Since the lubricant is confined within the gap, the value of θ is within the range:

$$0 < \theta\,(x, y, t) \leq 1; \quad \forall (x, y, t) \in S, \tag{3.29}$$

where $\theta < 1$ in the starved regions and $\theta \equiv 1$ in the pressurized regions.

It is subsequently assumed that, in the starved region, the lubricant is continuously distributed across the film and thus can be modelled as a continuum. However, since the lubricant only partly fills the gap, the density of the continuum has to be adjusted accordingly. The density of the vapour or surrounding medium is neglected and the density of the continuum in the starved region is set equal to $\theta\rho$, where ρ is the density of the lubricant. Since obviously the pressure equals the cavitation pressure, the sole mechanism for transport of lubricant in the starved region is Couette flow.

Based on this concept, the Reynolds equation can be modified to:

$$\frac{\partial}{\partial x}\left(\frac{\rho h^3}{\eta}\frac{\partial p}{\partial x}\right) + \frac{\partial}{\partial y}\left(\frac{\rho h^3}{\eta}\frac{\partial p}{\partial y}\right) = 6u_s\frac{\partial \theta\rho h}{\partial x} + 6\theta\rho h\frac{\partial u_s}{\partial x} + 12\frac{\partial \theta\rho h}{\partial t}. \tag{3.30}$$

This modified Reynolds equation is valid in the pressurized regions, as well as in the starved regions and, indeed, it reduces to the classical equation if $\theta = 1$.

To obtain a unique solution, the following complementarity condition must be fulfilled:

$$\begin{aligned} p(x,y,t)\,(1-\theta(x,y,t)) &= 0, \quad \text{with} \\ p(x,y,t) \geq 0 \;\text{ and }\; 0 &< \theta(x,y,t) \leq 1, \end{aligned} \tag{3.31}$$

since a point is either in a pressurized region ($p > 0$ and $\theta = 1$) or in a starved region ($p = 0$ and $\theta < 1$).

The modified Reynolds equation and the complementarity condition ensure continuity of flow across the boundary between starved regions and pressurized regions. This will be discussed further in Section 3.2.2.

The boundary conditions for the modified Reynolds equation differ from the boundary conditions for the classical Reynolds equation. Instead of specifying the pressure, one now needs to specify the fractional film content θ on the boundary of the domain. Note that because, in the starved region, the modified Reynolds equation is a hyperbolic differential equation, boundary conditions may only be specified at the inlet of the domain.

In the present study, the value of the fractional film content on the boundary will be derived from the "height" of the inlet lubricant layer, $h_l = h_{l;inlet}$. Furthermore, it will be assumed that $h_{l;inlet}$ is constant and equal to h_{oil}, i.e. it is assumed to be independent of the spatial coordinates as well as independent of time. The fractional film content at the inlet of the domain then follows from:

$$\theta\,(x_a,y,t) = \frac{h_{oil}}{h(x_a,y,t)}, \tag{3.32}$$

where x_a is the position of the inlet boundary, see Figure 3.2. Note that $h(x_a,y,t)$ depends on the pressure distribution, and hence the boundary condition is a function of the solution itself.

3.2.2 Continuity relations

As was mentioned before, a continuity relation must hold on the boundary between the starved region and the pressurized region. This relation is known as the Jakobsson-Floberg-Olsson (JFO) relation, see [32] and [23].

For steady state conditions, the JFO relation reduces to the so-called Reynolds cavitation boundary condition at the outlet meniscus. This condition is ascribed to Reynolds although he did not state it explicity. In the sequel of this section, the JFO relation will be extended to two-dimensional problems. For reasons of simplicity, the continuity relation for the one dimensional problem will be derived first. Contrary to Olsson, its derivation will be based on Green's theorem, which allows a more straightforward extension to the two-dimensional, time dependent situation.

One dimensional flow

If the radius of curvature R_y is large compared to R_x, the modified Reynolds equation becomes independent of y and reduces to:

$$\frac{\partial}{\partial x}\left(\frac{\rho h^3}{\eta}\frac{\partial p}{\partial x}\right) - 6u_s\frac{\partial \theta\rho h}{\partial x} - 12\frac{\partial \theta\rho h}{\partial t} = 0. \tag{3.33}$$

This can be rewritten in terms of the divergence of a vector field V as

$$\nabla \cdot V = 0, \tag{3.34}$$

where

$$V = (\frac{\rho h^3}{\eta}\frac{\partial p}{\partial x} - 6u_s\theta\rho h\ ,\ -12\theta\rho h), \tag{3.35}$$

and $\nabla \equiv (\partial/\partial x, \partial/\partial t)$. Then, also the integral $\iint_S \nabla \cdot V dS$ must vanish on an arbitrary domain S in (x,t)-space.

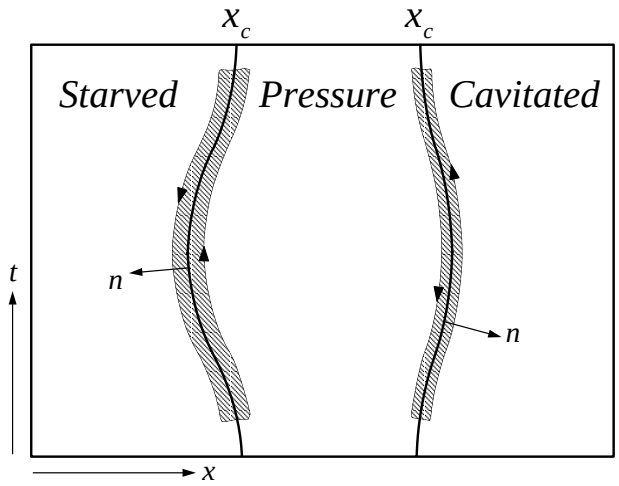


Figure 3.3: The regions S, as indicated by the shaded areas, on and across the boundary between the starved and pressurized region (left) and the pressurized and starved region (right). Please note the t-direction.

Applying Green's theorem[1] on a region S near and across the boundary between the starved region and the pressurized region or the pressurized

[1]Green's theorem reads:

$$\iint_S \nabla \cdot V dx dt = \oint_{\partial S} V \cdot n ds, \tag{3.36}$$

where ∂S is the contour, n is the exterior unit normal on ∂S and s is the coordinate which runs along the contour counter clockwise.

region and the starved region, see Figure 3.3, and letting the contours ∂S collapse onto the boundary, one obtains:

$$-\frac{h^2}{\eta}\frac{\partial p}{\partial x}n_x + 6u_s\left(1-\theta_c\right)n_x + 12\left(1-\theta_c\right)n_t = 0, \tag{3.37}$$

where n_x and n_t denote the components of the unit normal vector n *on* the boundary. Although this choice is arbitrary, we will let n point in the direction of the starved region, see Figure 3.3. In Equation 3.37, θ_c is the value of the fractional film content which is *in* the starved region and *nearest* to the boundary. Note that in Equation 3.37, the term $(1-\theta_c)$ is actually the difference between the fractional film content in the fully flooded region, $\theta \equiv 1$, and the fractional film content in the starved region, $\theta_c < 1$. The relation shows that if a pressure gradient exists on the meniscus, a jump in θ must occur! From $n_t/n_x = -\dot{x}_c$, where $\dot{x}_c$ denotes the velocity of the meniscus, we can rewrite Equation 3.37 as:

$$-\frac{h^2}{\eta}\frac{\partial p}{\partial x} + \left(6u_s - 12\dot{x}_c\right)\left(1-\theta_c\right) = 0. \tag{3.38}$$

In Figure 3.3, the inlet- and the outlet meniscus have both been denoted by x_c, to show that the continuity relation is valid on both boundaries.

Steady state

In the steady state situation, Equation 3.38 reduces to:

$$-\frac{h^2}{\eta}\frac{\partial p}{\partial x} + 6u_s\left(1-\theta_c\right) = 0. \tag{3.39}$$

Near the exit of a pressurized region, see Figure 3.3, it is obvious that the pressure gradient must be negative or zero, since on the left hand side of the boundary the pressure is larger than the vapour pressure on the right hand side of the boundary. If $\partial p/\partial x$ were negative, θ_c would have to be larger than unity in the cavitated region, for Equation 3.39 to be true. This violates Equation 3.29 and, as a result, the pressure gradient must vanish at the exit of the pressurized region. The fractional film content is a continuous function.

This is not true at the entrance of the pressurized region, as is explained in the frame below. Thus the pressure gradient as well as the discontinuity do exist at any inlet to a pressurized region, which is illustrated in Figure 3.4.

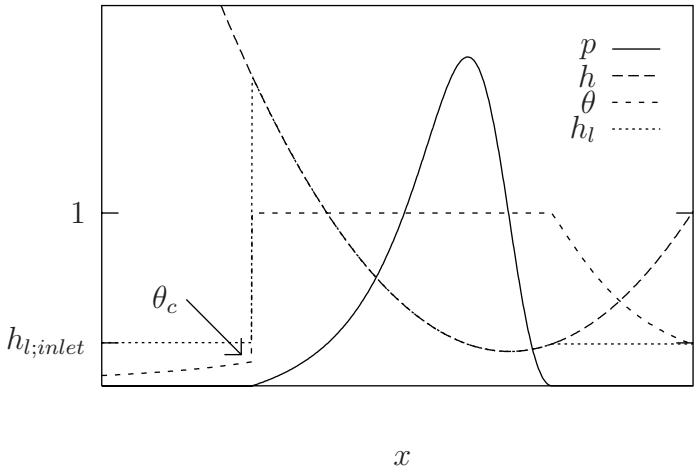


Figure 3.4: Pressure, film thickness, fractional film content and $h_{l;exit}$.

On the discontinuity of the fractional film content and pressure gradient.

In this frame, we will explain why the pressure gradient and the discontinuity of θ must exist at the inlet of the pressurized region. For this purpose, it is shown that, at the first point at which the pressure gradient vanishes, the pressure itself is maximal. This contradicts the statement that, by definition, the pressure is minimal in the starved region and hence, a jump in the fractional film content and a discontinuity in the pressure gradient must occur.

At the outlet meniscus, it was shown that the pressure gradient vanishes, therefore only Couette flow remains at the exit. For one dimensional flow, the flow at that position thus equals the total flow and is equal to $\rho h_{l;exit}$, where $h_{l;exit}$ denotes the film thickness at the outlet boundary. As a result, the pressure gradient vanishes only at locations where $\rho h = \rho h_{l;exit}$. Assuming the density to be constant, this implies that $\partial p/\partial x = 0$ if $h = h_{l;exit}$. This is illustrated in Figure 3.4. Note that in this case, $h_{l;exit} = h_{l;inlet}$. This is obvious since incompressible and one-dimensional flow is assumed.

Furthermore, for concentrated contacts, the gap tends to infinity for $x \to \pm\infty$. Since cavitation can only occur when the gap diverges, there must at least be one additional point in the pressure zone for which $h = h_{l;exit}$. In fact, h must cross the line $h = h_{l;exit}$, in an even number of points for it to go to infinity for $x \to \pm\infty$. This is illustrated in Figure 3.4 where the gap h is shown to cross the line 2

times.

As a result, since the pressure is minimal at the exit, the pressure must attain a maximum at the first point at which the pressure gradient is zero, and there can be no point in front of this point which is within the pressurized region, for which the pressure gradient vanishes. Thus, the pressure gradient on the inlet boundary of the pressurized region is non-zero.

As a consequence, the pressure gradient and the discontinuity in θ must occur at the inlet meniscus between the starved region and the pressurized region.

Transient

From Equation 3.38, one can observe that a similar analysis applies for the transient case. However, one now needs to include the velocity $\dot{x}_c$ of either the inlet or the outlet meniscus. Based on the arguments given in the paragraph above, the pressure gradient vanishes and θ is continuous if $(u_s/2 - \dot{x}_c)$ is positive, i.e. regardless of whether the meniscus is at the entrance or at the exit of the pressurized region. However, if $(u_s/2 - \dot{x}_c)$ becomes negative, i.e. if $\dot{x}_c > u_s/2$, the argument ceases to hold and a pressure gradient as well as the discontinuity in θ will show. Simply stated, the JFO relation states that whenever lubricant enters the pressurized region, a jump occurs. Whenever lubricant leaves the pressurized region, no discontinuity is observed. This topic will be further discussed in Chapter 7.

Two-dimensional flow

The continuity relation for the two-dimensional situation can be derived in a similar way. In this case, the vector field V, as given in Equation 3.34, is given by:

$$V = (\frac{\rho h^3}{\eta}\frac{\partial p}{\partial x} - 6u_s\theta\rho h \ , \ \frac{\rho h^3}{\eta}\frac{\partial p}{\partial y} \ , \ -12\theta\rho h), \tag{3.40}$$

and ∇ is defined according to $\nabla \equiv (\partial/\partial x, \partial/\partial y, \partial/\partial t)$. Gauss' theorem, i.e. the three-dimensional version of Green's theorem, can now be applied on a volume S that is near and across the boundary in the three-dimensional (x, y, t)-space, where the volume integral now reduces to a surface integral. By letting this surface collapse onto the boundary, the following continuity relation results:

$$-\frac{h^2}{\eta}\left(\frac{\partial p}{\partial x}n_x + \frac{\partial p}{\partial y}n_y\right) + 6u_s(1-\theta_c)n_x + 12(1-\theta_c)n_t = 0, \tag{3.41}$$

where n denotes the unit normal vector on the boundary surface (pointing towards the starved region), n_x is its component in the x-direction, n_y in the y-direction and n_t in the t-direction. Apparently, the continuity relation in the two-dimensional situation is similar to the continuity relation for the one dimensional case, except that it is valid in the direction normal to the boundary. The JFO relation for two-dimensional flows is discussed in more detail in Chapter 7.

Dimensionless equations

Adding the fractional film content adds only one extra variable to the number of variables that was introduced in Section 3.1.1. Since θ is already dimensionless, it can easily be included in the set of dimensionless variables. Substitution into the modified Reynolds equation gives:

$$\frac{\partial}{\partial X}\left(\frac{\bar{\rho}H^3}{\bar{\eta}\lambda}\frac{\partial P}{\partial X}\right)+\kappa^2\frac{\partial}{\partial Y}\left(\frac{\bar{\rho}H^3}{\bar{\eta}\lambda}\frac{\partial P}{\partial Y}\right)=\frac{\partial\theta\bar{\rho}H}{\partial X}+\frac{\partial\theta\bar{\rho}H}{\partial T}. \tag{3.42}$$

The dimensionless boundary condition for θ reads:

$$\theta(X_a,Y,T)=\frac{H_{l;inlet}(X_a,Y,T)}{H(X_a,Y,T)}, \tag{3.43}$$

where $H_{l;inlet}=h_{l;inlet}/c$ and X_a denotes the position of the inlet boundary of the dimensionless domain. As was stated earlier, $H_{l;inlet}$ may be a function of X,Y as well as T, however, it is assumed to be constant and equal to $H_{oil}=h_{oil}/c$. The introduction of θ does not affect the other equations.

Hence, starvation influences both the steady state and the transient solution and adds the dimensionless thickness of the inlet lubricant layer or inlet oil film H_{oil} to the set of parameters. In particular, its value influences the stiffness of the contact as well as the damping.

Chapter 4

Multilevel methods

This chapter discusses the methods that were used to solve the equations given in the previous chapter. It especially focuses on the effect of ellipticity and starved lubrication. First, the discrete equations will be presented. Secondly, the methods are described that were used to solve the system of equations; the multigrid method and the multilevel multi-integration method.

4.1 Introduction

Only in a few, simplified, asymptotic cases can solutions to the EHL problem be obtained analytically. In general however, a numerical approach is required. In the last few decades a number of methods have been proposed to numerically solve the EHL equations. However, due to stability and computing time problems, most algorithms were not really suited for realistic, often highly loaded conditions. This situation has changed dramatically since the introduction and further development of multigrid techniques by Lubrecht [48] and Venner [66]. Lubrecht was the first to apply the multigrid technique to the steady state line and point contact problem. Compared to the algorithms in use at that time, Lubrecht's algorithm was very efficient. Unfortunately, it was not stable at high loads and also the $O(n)$ efficiency, where n is the number of unknowns, was not obtained as the evaluation of the elastic deformation integral required $O(n^2)$ operations. To reduce the time needed to evaluate this deformation integral, Brandt and Lubrecht [12] developed the so-called multilevel multi-integration.

Venner [66] incorporated this technique into algorithms for the steady state line and point contact problem, as well as a time dependent line contact problem. He also introduced distributive relaxation in EHL, which solved the stability problems at higher loads. The resulting algorithms were of

$O(n \ln n)$ complexity and this efficiency enabled even transient problems to be addressed.

Venner's algorithm has formed the starting point in the present work. However, to facilitate accurate determination of damping, also in elliptically shaped contacts, further development was needed.

Accuracy

For steady state problems, a first order discretization provides sufficiently accurate results. However for the transient problem, second order discretization is indispensable. Due to the shear flow dominance in the contact region, the steady state Reynolds equation reduces to:

$$-\frac{\partial \bar{\rho} H}{\partial X} - \frac{\partial \bar{\rho} H}{\partial T} \approx 0, \tag{4.1}$$

forcing $\bar{\rho} H = \bar{\rho} H(X - T)$. From computational fluid dynamics, it is well known that a first order discretization results in what is known as numerical damping. In the present study, using a first order discretization for the equation above would result in the calculated damping being larger than the actual value of the damping for the continuum problem. Therefore, second order discretization is essential.

Many second order schemes are available for the discretization of the wedge/squeeze term, e.g. see Yavneh [77]. In the present study two discretization schemes have been used. For simplicity, the standard upstream discretization ($SU2$) was used for the starved lubricated problem. The so-called narrow upstream discretization ($NU2$) was used for fully flooded conditions.

The choice for the $NU2$ discretization was motivated by the fact that the discretization error for any characteristic component is smaller than the error using $SU2$ discretization. In addition, a time step $h_T = 2h_X$ can be used for small Ω_n or Ω_e without loss of accuracy, see Appendix B.

Ellipticity

A straightforward extension of Venner's circular contact algorithm to elliptically shaped contacts is described in Nijenbanning [54]. In Nijenbanning's approach, contrary to the scaling of the y-coordinate with the Hertzian contact width b ($Y = y/b$), y was scaled with a, the Hertzian contact size in the x-direction ($Y_N = y/a$). (The dimensionless coordinates given in Nijenbanning will be denoted by Y_N and X_N.) For ellipticity ratios around $\kappa = 1$, this approach is quite efficient and virtually no modifications need to

be applied to the algorithm. However, for small ellipticity ratio's, i.e. if the contact width is large compared to the contact length, the domain becomes very large in the Y_N-direction compared to its size in the X_N-direction.

Thus, if the mesh size h_{X_N} in the X_N-direction equals the mesh size h_{Y_N} in the Y_N-direction, the number of unknowns in Y_N-direction, n_{Y_N}, becomes much larger than the number of unknowns in X_N-direction, n_{X_N}. Apart from the fact that such a large number is not necessary (the discretization error is mainly due to the small number of unknowns in X_N-direction), the coarsest possible grid will still contain a large number of points. This has a significant adverse effect on the performance of a multigrid algorithm.

One way to overcome this problem is to increase the mesh size h_{Y_N}, such that n_{Y_N} is of the same order of magnitude as n_{X_N}. However, this requires a modification of the multilevel multi-integration technique, as is described at the end of the chapter.

Since a modification is required anyway, a more elegant approach is to scale the y-coordinate with the Hertzian contact width b, i.e. $Y = y/b$. This was indeed applied in the equations given in Chapter 3. Then, in terms of the dimensionless coordinates, the actual elliptical contact region maps onto a circle and a mesh size $h_X = h_Y$ can be used again.

Starved lubrication

For starved lubrication, Elrod [19] presented an algorithm that automatically determines the position of the meniscus between the starved and pressurized region. Chevalier [14] adopted Elrod's algorithm for EHL and presented steady state solutions for the EHL circular contact obtained using multilevel methods. However, this algorithm used only a first order discretization and also "line-relaxation" was not implemented.

The extensions of Chevalier's algorithm to second order accuracy, time dependence and line-relaxation, will be discussed below.

The outline of the chapter is as follows. First, the discrete equations are introduced. Subsequently, the multigrid method is discussed briefly with its application to fully flooded and starved EHL. The technique used for the fast evaluation of the elastic deformation integral, the multilevel multi-integration, will be discussed separately in Section 4.5.

4.2 Discrete equations

The equations describing the EHL contact problem are discretized on a rectangular domain $X_a \leq X \leq X_b$ and $Y_a \leq Y \leq Y_b$. The domain is covered

by a mesh of points whose distance from one to another is the mesh size h_X and h_Y in the X- and Y-directions, respectively. In the present analysis, $h_X = h_Y$. The mesh size in the time direction or the time step is denoted by h_T.

The different variables, like pressure and film thickness, now reside on this mesh and discrete equations are derived that approximate the equations involved and relate the variables on the mesh. For the discretization of the Reynolds equation, a standard second order discretization for the Poiseuille term and a $NU2$ discretization for the wedge/squeeze term are used. For starved lubricated problems, $SU2$ discretization was used. In the discrete film thickness equation as well as the force balance equation and the equation of motion, the pressure is assumed to be constant on $X - h_X/2 \leq X < X + h_X/2$ and on $Y - h_Y/2 \leq Y < Y + h_Y/2$. For free vibrations, the equation of motion is discretized using the Newmark scheme. In this section the choices for the specific discretizations are elaborated upon.

Reynolds equation

On writing

$$\epsilon_{i,j} \equiv \frac{\bar{\rho}_{i,j} H_{i,j}^3}{\bar{\eta}_{i,j} \lambda}, \tag{4.2}$$

where i and j are the spatial grid indices and taking

$$\begin{aligned}
\epsilon_{i+1/2,j} &\equiv (\epsilon_{i,j} + \epsilon_{i+1,j})/2, \\
\epsilon_{i-1/2,j} &\equiv (\epsilon_{i,j} + \epsilon_{i-1,j})/2, \\
\epsilon_{i,j+1/2} &\equiv (\epsilon_{i,j} + \epsilon_{i,j+1})/2, \\
\epsilon_{i,j-1/2} &\equiv (\epsilon_{i,j} + \epsilon_{i,j-1})/2,
\end{aligned} \tag{4.3}$$

the $SU2$ discretization of the Reynolds equation, used for starved lubrication, reads:

$$\begin{aligned}
h_X^{-2} \left(\epsilon_{i+1/2,j} \left(P_{i+1,j} - P_{i,j} \right) - \epsilon_{i-1/2,j} \left(P_{i,j} - P_{i-1,j} \right) \right) \\
+ \kappa^2 h_Y^{-2} \left(\epsilon_{i,j+1/2} \left(P_{i,j+1} - P_{i,j} \right) - \epsilon_{i,j-1/2} \left(P_{i,j} - P_{i,j-1} \right) \right) \\
- h_X^{-1} \left(1.5\, \theta \bar{\rho} H_{i,j} - 2.0\, \theta \bar{\rho} H_{i-1,j} + 0.5\, \theta \bar{\rho} H_{i-2,j} \right) \\
- h_T^{-1} \left(1.5\, \theta \bar{\rho} H_{i,j} - 2.0\, \theta \bar{\rho} H_{i,j,k-1} + 0.5\, \theta \bar{\rho} H_{i,j,k-2} \right) = 0.
\end{aligned} \tag{4.4}$$

The terms without a time index denote the term at time index k, e.g. $P_{i,j} = P_{i,j,k}$. Furthermore, terms like $\theta \bar{\rho} H_{i,j,k}$ have been used to denote the term $\theta_{i,j,k} \bar{\rho}_{i,j,k} H_{i,j,k}$.

The $NU2$ discretization, used for fully flooded conditions, has two forms depending on the ratio h_X/h_Y. For $h_X \leq h_T$ and time index k, it reads:

$$\begin{aligned} h_X^{-2}\left(\epsilon_{i+1/2,j}\left(P_{i+1,j}-P_{i,j}\right)-\epsilon_{i-1/2,j}\left(P_{i,j}-P_{i-1,j}\right)\right) \\ +\kappa^2 h_Y^{-2}\left(\epsilon_{i,j+1/2}\left(P_{i,j+1}-P_{i,j}\right)-\epsilon_{i,j-1/2}\left(P_{i,j}-P_{i,j-1}\right)\right) \\ -\left(h_X^{-1}-h_T^{-1}\right)\left(1.5\ \bar{\rho}H_{i,j}-2.0\ \bar{\rho}H_{i-1,j}+0.5\ \bar{\rho}H_{i-2,j}\right) \\ -h_T^{-1}\left(1.5\ \bar{\rho}H_{i,j}-2.0\ \bar{\rho}H_{i-1,j,k-1}+0.5\ \bar{\rho}H_{i-2,j,k-2}\right)=0. \end{aligned} \tag{4.5}$$

For $h_X > h_T$ it reads:

$$\begin{aligned} h_X^{-2}\left(\epsilon_{i+1/2,j}\left(P_{i+1,j}-P_{i,j}\right)-\epsilon_{i-1/2,j}\left(P_{i,j}-P_{i-1,j}\right)\right) \\ +\kappa^2 h_Y^{-2}\left(\epsilon_{i,j+1/2}\left(P_{i,j+1}-P_{i,j}\right)-\epsilon_{i,j-1/2}\left(P_{i,j}-P_{i,j-1}\right)\right) \\ -h_X^{-1}\left(1.5\ \bar{\rho}H_{i,j}-2.0\ \bar{\rho}H_{i-1,j,k-1}+0.5\ \bar{\rho}H_{i-2,j,k-2}\right) \\ -\left(h_T^{-1}-h_X^{-1}\right)\left(1.5\ \bar{\rho}H_{i,j}-2.0\ \bar{\rho}H_{i,j,k-1}+0.5\ \bar{\rho}H_{i,j,k-2}\right)=0. \end{aligned} \tag{4.6}$$

If a single mesh and a constant time step were used, obviously, the $NU2$ discretization would require only one of the two equations given above; however, in the multigrid algorithm which is explained below, the solution is calculated on different grids and different grids are used to update the solution on the finest grid. Consequently, both forms of the discretization are used in the algorithm.

Film thickness equation

Assuming the pressure to be constant on $X_r - h_X/2 \leq X < X_r + h_X/2$ and on $Y_s - h_Y/2 \leq Y < Y_s + h_Y/2$, the equation for the film thickness becomes:

$$H_{i,j} = -\Delta + \mathcal{S}X_i^2 + (1-\mathcal{S})Y_j^2 + \sum_{r=0}^{n_X}\sum_{s=0}^{n_Y} K_{irjs}P_{r,s}. \tag{4.7}$$

The elastic deformation integral thus reduces to a multi-summation; for all $(n_X+1)(n_Y+1)$ film thicknesses $H_{i,j}$, one needs to carry out a multi-summation over $(n_X+1)(n_Y+1)$ points.

The expression of the discrete kernel K_{irjs} is given in Appendix C. Its value depends on the mesh size, the ellipticity κ and on the absolute distance between X_i and X_r and Y_j and Y_s, i.e. it depends on $|r-i|$ and $|s-j|$. As a result, K_{irjs} can be expressed as a two-dimensional matrix.

Equation of motion / Equations of state

For harmonic excitation, the second order derivative in the equation of motion is set to zero and the discrete force balance equation at time index k is simply:

$$\frac{3}{2\pi} h_X h_Y \sum_{i=0}^{n_X} \sum_{j=0}^{n_Y} P_{i,j,k} = 1 + A \sin(\Omega_e k h_T) \tag{4.8}$$

Note that, by setting $A = 0$, this equation is also valid for steady state conditions.

For free vibrations, where the load is assumed to be constant, the equation is discretized according to the so-called Newmark integration scheme, see [53]. This implicit second order scheme is widely used in finite element calculations. The derivation of the scheme can also be found in Bathe [6].

One of our aims is to find the influence of lubrication on damping. Hence, in order to present quantitative results, the amplitude decay of an oscillatory component due to discretization errors must be small. Since for linear second order differential equations the Newmark scheme shows *no* amplitude decay, this scheme was adopted. It is noted however, that the Newmark scheme does exhibit some period elongation.

The Newmark scheme is derived by taking the average of the second order derivative at times kh_T and $(k-1)h_T$, i.e.:

$$\Delta_k = \Delta_{k-1} + h_T \dot{\Delta}_{k-1} + \frac{1}{2} h_T^2 \left(\frac{1}{2} \ddot{\Delta}_{k-1} + \frac{1}{2} \ddot{\Delta}_k \right) \tag{4.9}$$

and

$$\dot{\Delta}_k = \dot{\Delta}_{k-1} + h_T \left(\frac{1}{2} \ddot{\Delta}_{k-1} + \frac{1}{2} \ddot{\Delta}_k \right). \tag{4.10}$$

For the equation of motion, substitution and rearranging of terms yields:

$$\Delta_k + E \frac{3}{2\pi} h_X h_Y \sum_{i=0}^{n_X} \sum_{j=0}^{n_Y} P_{i,j,k} = F, \tag{4.11}$$

where

$$E = \frac{1}{4} \Omega_n^2 h_T^2 \tag{4.12}$$

and

$$F = \frac{1}{4} \Omega_n^2 h_T^2 + \Delta_{k-1} + h_T \dot{\Delta}_{k-1} + \frac{1}{4} h_T^2 \ddot{\Delta}_{k-1}. \tag{4.13}$$

The derivation of the discrete forms of Roeland's equation and the equation of Dowson and Higginson is straightforward and hence will not be given here explicitly.

4.3 Multigrid

The general idea behind multigrid methods, quoting Brandt [10], is to find algorithms for which the following "golden rule" applies:

> *The amount of computational work should be proportional to the amount of real physical changes in the computed system.*

This implies that if a process stalls despite the work that is invested, there is bound to be a faster way to solve the problem. Multigrid methods seek for these ways by trying to solve components of the solution on the scale on which they appear.

A discrete approximation of the solution of a non-linear differential equation needs to be solved by means of some iterative solution process, so-called relaxation, since the coefficients of the differential- or integral-operator depend on the solution itself. Conventional relaxation schemes, like Gauss-Seidel relaxation, are known to efficiently reduce high-frequency error components in discrete approximations of the solution of elliptic differential equations. (High-frequency components are those components whose wavelength is of the order of the mesh size of the grid.) On the other hand, if the computational domain is large compared to the mesh size, most relaxation schemes hardly reduce the low-frequency components. Hence, after a number of relaxation sweeps, the error consists of low-frequency components only and error reduction stalls.

The stalling of this process can be circumvented by approximating the now smooth error on a coarser grid. The mesh size of the coarse grid is usually taken twice as large as the mesh size of the fine grid, which is referred to as standard coarsening. Subsequently, so-called "coarse grid equations" are derived and solved and its solution is used to update the solution on the fine grid. Obviously, the coarse grid equations need to be solved by an iterative procedure as well and a similar procedure can be applied to these equations. In fact, the process can be repeated recursively until a grid is reached where the mesh size is of the order of the domain and a fixed number of relaxations is sufficient to completely solve the coarse grid equations. This recursive procedure is called a "cycle". If 1 coarse grid correction per level is performed, the cycle is called a V-cycle, its name being derived from the shape of the picture, illustrating the cycle in Figure 4.1. From the Figure, it can be seen that the W-cycle denotes a cycle for which two coarse grid corrections per level are used.

Although multigrid cycles can be applied to any initial approximation, an initial approximation on the fine grid can be obtained by interpolation of a solution from a coarser grid which has an error below the discretization error

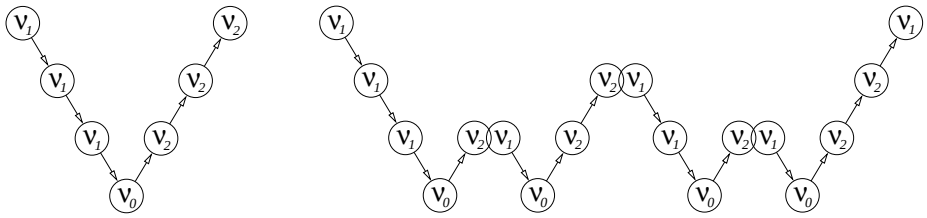


Figure 4.1: $V(\nu_1, \nu_2)$- and $W(\nu_1, \nu_2)$-cycle.

on that grid. Subsequently, this fine grid approximation is updated using the multigrid cycles described above. The preference for such a strategy is that the number of cycles, needed to obtain a converged solution, is constant and independent of the number of unknowns. In the "full multigrid"-algorithm (FMG), this strategy is applied recursively. The FMG algorithm is illustrated in Figure 4.2.

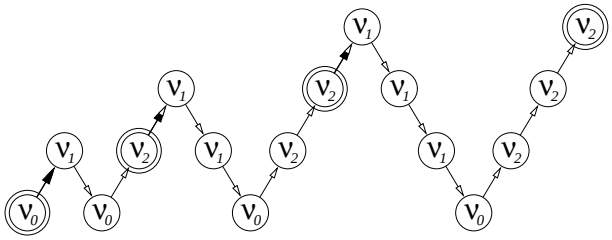


Figure 4.2: The FMG algorithm with $1V(\nu_1, \nu_2)$-cycle per grid.

The FMG algorithm has the prospect of solving the problem below the discretization error in an amount of work that is proportional to the number of unknowns n, i.e. its complexity is $O(n)$. Provided the individual components of the algorithm are at most of this complexity $O(n)$, the FMG algorithm is optimal (a further reduction is impossible as a plain evaluation of the solution already requires $O(n)$ operations).

For an introduction to the multigrid method the reader is referred to Briggs [11]. Elaborate discussions can be found in Brandt [10], Stüben and Trottenberg [64], and Joppich [35]. A résumé of the method is given in Appendix D.

4.3.1 Relaxation

The cornerstone of the multigrid procedure is relaxation. Obviously, the relaxation procedure should be stable, but, for the multigrid method to work, it should also effectively reduce high-frequency error components. As the point and elliptical contact problems are two-dimensional problems, it should do so in both the X and Y-directions, so coarsening can be applied in both directions.

A prerequisite to obtain a fast and stable multigrid algorithm is to understand the behaviour of the equations in the calculational domain. As the convergence of the solution is dominated by the convergence of the Reynolds equation, we only need to focus on this equation.

In the Reynolds equation, as stated by Venner [66], two types of behaviour can be recognized. One is the behaviour of the equations for low pressures, which will be discussed in the next section. The other type is the behaviour in the high pressure zone, which is discussed later.

Low pressures

For low pressures, the elastic deformation of the surfaces is small and, consequently, the differential aspect of the equations dominate the behaviour of the equation, i.e. the Reynolds equation reduces to the anisotropic Poisson equation with varying coefficients:

$$\frac{\partial}{\partial X}\left(\epsilon\frac{\partial P}{\partial X}\right)+\kappa^2\frac{\partial}{\partial Y}\left(\epsilon\frac{\partial P}{\partial Y}\right)=F(X,Y), \tag{4.14}$$

where $F(X,Y)$ denotes the right-hand side function. For this equation, which is an equation of elliptic type, one-point Gauss-Seidel relaxation is stable. However, as will be explained shortly, for small values of κ, error components with high-frequency in the Y-direction are hardly reduced. Hence, coarsening in this direction would not be possible.

Fortunately, the loss of coupling can be restored using so-called line-relaxation. Instead of applying single, consecutive changes, in line-relaxation all changes on a line perpendicular to the direction of weak coupling, are solved and applied simultaneously. Heuristically, the distributive relaxation relies on the property that the coupling between *many*, almost uncoupled variables is much stronger than the coupling between the *individual* variables. Contrary to the single changes, line-relaxation requires solving a system of equations. The solution of this system of equations can, for instance, be obtained using decimation or Gaussian elimination with partial pivoting, see Atkinson [3].

The error reduction or smoothing behaviour of relaxation schemes, such as the one-point relaxation and line-relaxation, is clearly illustrated by means of local mode analysis. In the frame below, the error reduction for the anisotropic Poisson equation is studied for both the Gauss-Seidel one-point and line-relaxation. (For convenience, the variable u will be used for the pressure P, x for X, y for Y and f will be used to denote the right-hand side function.)

Local mode analysis for the anisoptropic Poisson equation.

Consider the anisotropic Poisson equation:

$$\frac{\partial^2 u}{\partial x^2} + \kappa^2 \frac{\partial^2 u}{\partial y^2} = f. \tag{4.15}$$

The central second order discretization of this equation relates the unknown $u_{i,j}$ to its neighbours according to:

$$\frac{u^h_{i-1,j} - 2u^h_{i,j} + u^h_{i+1,j}}{h_x^2} + \kappa^2 \frac{u^h_{i,j-1} - 2u^h_{i,j} + u^h_{i,j+1}}{h_y^2} = f^h_{i,j}, \tag{4.16}$$

where $u^h_{i,j} = u(x_i, y_j)$ and $h_x = h_y$ denotes the mesh size. (In the following discussion, $h_x = h_y \equiv h$.) It is assumed that the boundary conditions are either periodic or sufficiently far away and, as a result, the influence of the boundary on the analysis may be neglected.

Let $\tilde{u}^h$ be some approximation to u^h. Applying one-point Gauss-Seidel relaxation in lexicographic order, the unknowns are scanned one by one and $\tilde{u}^h_{i,j}$ is replaced by a new value $\bar{u}^h_{i,j}$, such that Equation 4.16 is satisfied. That is, $\bar{u}^h_{i,j}$ satisfies:

$$\frac{\bar{u}^h_{i-1,j} - 2\bar{u}^h_{i,j} + \tilde{u}^h_{i+1,j}}{h^2} + \kappa^2 \frac{\bar{u}^h_{i,j-1} - 2\bar{u}^h_{i,j} + \tilde{u}^h_{i,j+1}}{h^2} = f^h_{i,j}. \tag{4.17}$$

Since after one complete relaxation sweep, $\bar{u}^h_{i,j}$ does *not* satisfy Equation 4.16, further sweeps are required to improve it.

Substitution of the error $\tilde{v}^h_{i,j} \equiv u^h_{i,j} - \tilde{u}^h_{i,j}$ *before* the relaxation sweep and $\bar{v}^h_{i,j} \equiv u^h_{i,j} - \bar{u}^h_{i,j}$ *after* the sweep, yields:

$$\bar{v}^h_{i-1,j} - 2\bar{v}^h_{i,j} + \tilde{v}^h_{i+1,j} + \kappa^2 \left(\bar{v}^h_{i,j-1} - 2\bar{v}^h_{i,j} + \tilde{v}^h_{i,j+1} \right) = 0, \tag{4.18}$$

which shows that the error after relaxation is some average of the errors before and after the relaxation.

A Fourier expansion of the error yields:

$$v_{i,j}^h = \sum\sum A(\theta_1, \theta_2) e^{\iota(\theta_1 i + \theta_2 j)}, \tag{4.19}$$

where $\iota = \sqrt{-1}$. For each component (θ_1, θ_2), substitution provides a relation between the amplitude of the Fourier component $A(\theta_1, \theta_2) = \bar{A}(\theta_1, \theta_2)$ *after* the relaxation sweep and the amplitude $A(\theta_1, \theta_2) = \tilde{A}(\theta_1, \theta_2)$ *before* the sweep:

$$\mu(\theta_1, \theta_2) = \left| \frac{\bar{A}(\theta_1, \theta_2)}{\tilde{A}(\theta_1, \theta_2)} \right| = \left| \frac{e^{\iota\theta_1} + \kappa^2 e^{\iota\theta_2}}{2 + 2\kappa^2 - e^{-\iota\theta_1} - \kappa^2 e^{-\iota\theta_2}} \right|, \tag{4.20}$$

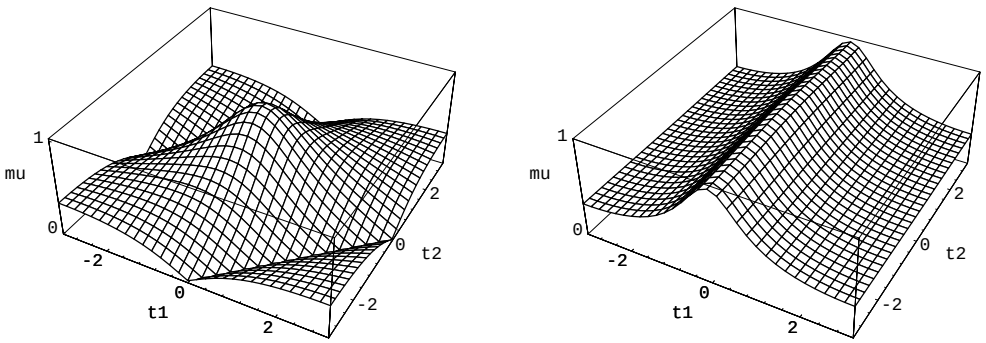


Figure 4.3: Amplification factor $\mu(\theta_1, \theta_2)$ for one-point Gauss-Seidel relaxation; $\kappa = 1$ (left) and $\kappa = 0.1$ (right).

Figure 4.3 shows the amplification factor for one-point Gauss-Seidel relaxation for $\kappa = 1$ and $\kappa = 0.1$ and indeed, for $\kappa = 1$, the Figure shows that Gauss-Seidel reduces high-frequency components ($\pi/2 \leq |\theta_{1,2}| \leq \pi$), whereas low-frequency components ($0 < |\theta_{1,2}| < \pi/2$) are hardly affected. This smoothing behaviour is reflected in the so-called smoothing rate. It is defined as the largest factor by which the high-frequency error components, i.e. those components which can not be represented on a coarser grid, are reduced. The smoothing rate for $\kappa = 1$ is 0.5.

For $\kappa = 0.1$, the figure shows that, although the relaxation efficiently reduces high-frequency components in the x-direction, ($\pi/2 \leq |\theta_1| \leq \pi$), high-frequency components in the y-direction, ($\pi/2 \leq |\theta_2| \leq \pi$), remain unaffected. As a result, the smoothing rate in this case increases up to 0.98. The slow convergence of these components is due to the loss of coupling between the variables in the y-direction. (For small κ, the variable $u_{i,j}^h$ is related to $u_{i-1,j}^h$ and $u_{i+1,j}^h$ only.)

So-called line-relaxation does not show bad error reduction for high-frequency components in the y-direction, provided the line is perpendicular to the direction of weak coupling, i.e. in this case the x-direction. For line-relaxation in the x-direction, instead of single changes, new approximations $\bar{u}^h$ are applied simultaneously on each line of constant y. The new approximations thus follow from:

$$\frac{\bar{u}^h_{i-1,j} - 2\bar{u}^h_{i,j} + \bar{u}^h_{i+1,j}}{h^2} + \kappa^2 \frac{\bar{u}^h_{i,j-1} - 2\bar{u}^h_{i,j} + \tilde{u}^h_{i,j+1}}{h^2} = f^h_{i,j}, \qquad (4.21)$$

and the amplification factor is given by:

$$\mu(\theta_1, \theta_2) = \left| \frac{\kappa^2 e^{\iota\theta_2}}{2 + 2\kappa^2 - e^{-\iota\theta_1} - \kappa^2 e^{-\iota\theta_2} - e^{\iota\theta_1}} \right|. \qquad (4.22)$$

The amplification factors for Gauss-Seidel line-relaxation for $\kappa = 1$ and $\kappa = 0.1$ are shown in Figure 4.4.

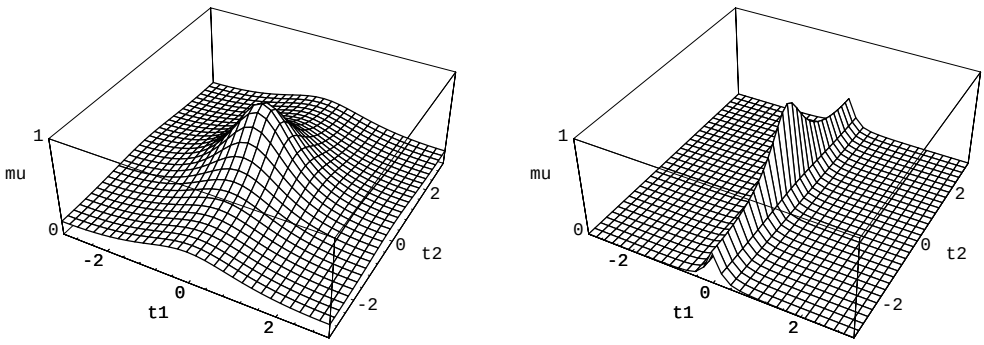


Figure 4.4: Amplification factor $\mu(\theta_1, \theta_2)$ for line-relaxation; $\kappa = 1$ (left) and $\kappa = 0.1$ (right).

The smoothing rate for this type of relaxation is 0.447, irrespective of κ, provided $\kappa \leq 1$. For $\kappa = 1$, the smoothing rate is thus smaller than for the one-point Gauss-Seidel relaxation. (It must be noted that there is a price to pay for the smaller smoothing rate, i.e. one needs to solve a system of equations.) More importantly however, the amplification factor for $\kappa = 0.1$ is such that high-frequency components in the y-direction are greatly reduced. In fact, the Figure shows that actually *all* error components, except those around $\theta_1 = 0$, are solved. Although this may come as a surprise, note that all related unknowns are updated simultaneously and as a result so do all error components.

Finally it is noted that for $\kappa > 1$, instead of x-line relaxation, y-line relaxation should be applied.

From the analysis above it is concluded that, although one-point relaxation suffices for circular contacts, line-relaxation is essential to obtain an efficient multigrid algorithm for small values of κ. The implementation details for Gauss-Seidel relaxation for the Reynolds equation are discussed in Section 4.4.

With respect to the line-relaxation, a problem arises on how these pressure changes should be solved simultaneously, with the new values of the film thickness. Since the film thickness involves a summation of these changes, this leads to a full matrix and the solution of this set of equations would require a solver which is equivalent to a solver for the EHL line contact problem. Fortunately, the equations need not be solved exactly to obtain full efficiency. A reduction of the error by one or two orders of magnitude proves to be sufficient and one only needs to take into account the three largest terms in the summation. The full matrix then reduces to a banded matrix, which can be solved using some standard matrix solver, e.g. using decimation or Gaussian elimination. Recalculation of the elastic deformation is not required after the relaxation of a single line. An update after the relaxation of all lines suffices.

High pressures

For high loads, as will be shown in subsequent chapters, the film thickness H becomes very small and when the exponential increase of viscosity with pressure is included, the viscosity is large as well. Consequently, $\epsilon = \bar{\rho}H^3/(\bar{\eta}\lambda)$, as it appears in Equation 3.12, becomes very small and the steady state Reynolds equation reduces to:

$$\frac{\partial \bar{\rho}H}{\partial X} \approx 0. \tag{4.23}$$

The Reynolds equation thus prescribes that for high pressures $\bar{\rho}H = \bar{\rho}H(X)$ is constant and independent of Y. Hence, in contrast to being a differential equation, the Reynolds equation reduces to an integral equation for the unknown pressure, i.e. one needs to solve the pressure for a given film thickness, which involves an integral of the product of pressure and the elastic deformation kernel, see 3.14.

Gauss-Seidel relaxation is known to be unstable for this type of integral equation, see Brandt and Lubrecht [12]. The instability can easily be shown as follows. First, it is noted that a pressure change $\delta_{r,s}$ in some point causes the elastic deformation at a distance $d = \sqrt{(X_i - X'_r)^2 + (Y_j - Y'_s)^2}$ from that point, to change as $\delta_{r,s}/d$. Hence, the elastic deformation in a specific point is only affected by pressure changes close to that point. In the process

of updating the pressures so as to satisfy the Reynolds equation, the effect of such a pressure change on the elastic deformation, and consequently on the film thickness, is local.

The change in the elastic deformation as a result of the pressure changes in a *complete* relaxation sweep, should also remain local. Otherwise relaxation, which is a local process, cannot correct for these accumulated changes. Since a constant pressure change δ on a line in all points up to a distance d, causes a change in the elastic deformation which is proportional to $\delta \log d$, the change in the film thickness indeed grows without bounds for increasing d and as a result the scheme is unstable.

The solution to the stability problem requires the introduction of a different type of relaxation; so-called distributive relaxation. In this type of relaxation, instead of changing one unknown, a number of unknowns is updated according to some predefined distribution.

For second order, uniform, distributive relaxation, five unknown pressures are updated simultaneously according to:

$$\begin{aligned}
\bar{P}_{i,j+1} &= \tilde{P}_{i,j+1} - \delta_{i,j}/4, \\
\bar{P}_{i-1,j} &= \tilde{P}_{i-1,j} - \delta_{i,j}/4, \\
\bar{P}_{i,j} &= \tilde{P}_{i,j} + \delta_{i,j}, \\
\bar{P}_{i+1,j} &= \tilde{P}_{i+1,j} - \delta_{i,j}/4, \\
\bar{P}_{i,j-1} &= \tilde{P}_{i,j-1} - \delta_{i,j}/4,
\end{aligned} \tag{4.24}$$

where $\bar{P}$ denotes the new update and $\tilde{P}$ the old approximation. In so-called stencil notation, the distribution thus reads:

$$\begin{array}{ccc}
 & -\delta_{i,j}/4 & \\
-\delta_{i,j}/4 & \delta_{i,j} & -\delta_{i,j}/4 \\
 & -\delta_{i,j}/4 &
\end{array} .$$

After one relaxation sweep, the new pressure $\bar{P}_{i,j}$ thus relates to the old pressure $\tilde{P}_{i,j}$ as:

$$\bar{P}_{i,j} = \tilde{P}_{i,j} + \delta_{i,j} - \left(\delta_{i+1,j} + \delta_{i-1,j} + \delta_{i,j+1} + \delta_{i,j-1}\right)/4, \tag{4.25}$$

If updated "a là" Jacobi, the elastic deformation, and as a result the film thickness $H_{i,j}$, thus changes according to:

$$\bar{H}_{i,j} = \tilde{H}_{i,j} + \sum_{r=1}^{n_X-1} \sum_{s=1}^{n_Y-1} \Delta K_{irjs} \delta_{r,s}, \tag{4.26}$$

where

$$\Delta K_{irjs} = K_{irjs} - \left(K_{ir+1js} + K_{ir-1js} + K_{irjs+1} + K_{irjs-1}\right)/4. \tag{4.27}$$

Because ΔK decays much faster with d than K itself, the distributive relaxation forces the changes in the film to be local, i.e. a constant pressure change δ in all points upto a distance d, now induces a change δ/d^2 in the elastic deformation. For circular contacts a stable relaxation is thus obtained.

For wide elliptical contacts ($\kappa \ll 1$), the uniform distributive relaxation is stable, only if line-relaxation is applied, i.e. the uniform distributive relaxation is unstable if one-point changes are applied. This instability is due to the amplification of error components which have a high frequency in the Y-direction and is apparentely a result of forcing changes to the unknowns according to a distribution which does not reflect the distribution of the integral operator. A possible remedy for the instability is a distribution which does reflect the distribution in the integral operator; the non-uniform distributive relaxation given by:

$$\begin{array}{ccc} & -\kappa\delta_{i,j}/(2+2\kappa) & \\ -\delta_{i,j}/(2+2\kappa) & \delta_{i,j} & -\delta_{i,j}/(2+2\kappa) \\ & -\kappa\delta_{i,j}/(2+2\kappa) & \end{array} .$$

That the non-uniform distribution is stable might be verified by examining the behaviour of the relaxation for vanishing κ, i.e. for the line contact problem. In that case, the distribution reduces to the second order distributive relaxation, $\delta_{i,j}/2 \;\; \delta_{i,j} \;\; \delta_{i,j}/2$, that already proved to be stable for line contacts, see Venner [66]. In addition to the stability, the relaxation has reasonable smoothing properties in both directions.

An additional problem, again with respect to the multigrid solver, is the loss of coupling in the Y-direction. Since in the high pressure zone $\bar{\rho}H$ was shown to be independent of Y, the coupling of the variables in the Y-direction is lost. Hence, as was observed for the Gauss-Seidel relaxation at small values of κ, the relaxation becomes increasingly ineffective in reducing high-frequency components in this direction. Consequently, after a number of relaxations, the error is not smooth in the Y-direction and coarsening in this direction would not be possible.

A solution to this problem was already given earlier, in the discussion on the loss of coupling due to small values of κ. That is to say, also in the high

pressure zone, the relaxation should be applied in a line-relaxation manner and the line should be aligned to the X-direction.

The implementation details of Jacobi distributive relaxation is elaborated upon in Section 4.4.

The combination of the distributive and line-relaxation thus provides a stable scheme in the high pressure zone, which has good smoothing properties in both directions.

Cavitation and the equation of motion

Cavitation is simply taken into account by setting the pressure to zero, whenever negative pressures are encountered in the relaxation scheme.

Furthermore, since the relaxation of the Reynolds equations dominates the convergence of the solver, the relaxation of the force balance equation or equation of motion (the residual of the equation decreases by changing the value of Δ), only needs to be performed after a fixed number of relaxations of the Reynolds equation. The unknown derivative of the equation with respect to Δ, which is required in the Newton-Raphson iteration, can be approximated by a constant, i.e. a fraction of the residual is added or substracted to Δ, according to the sign of the residual.

The details with respect to cavitation and the force balance equation and the equation of motion are given in Section 4.4.

Starved lubrication

In the previous chapter, it was explained that in a starved lubricated contact, one can distinguish a starved region enclosing one or more pressurized regions. In fact, it might be possible that, in its turn, a pressurized region encloses a starved region. The position of the boundary between these regions, referred to as the meniscus, is unknown and in time dependent situations, this position may even change in time. Hence, an algorithm must be able to find this position automatically.

Chevalier [14] first adopted Elrod's algorithm for EHL and presented solutions for the EHL circular contact at steady state conditions. Chevalier used the modified Reynolds equation, given in Chapter 3, which is valid in both the starved and pressurized region, see also Bayada [8].

In the pressurized zone, similar problems arise as those encountered in the fully flooded situation. That is to say, in the low pressure region, the Gauss-Seidel line-relaxation is needed, whereas Jacobi distributive line-relaxation is required in the high pressure zone.

The modified Reynolds equation determines the fractional film content in the starved region and, since line-relaxation is applied anyway, it is very convenient and straightforward to include the changes $\delta_{i,j}$ for $\theta_{i,j}$, in the system of equations for the changes $\delta_{i,j}$ for $P_{i,j}$.

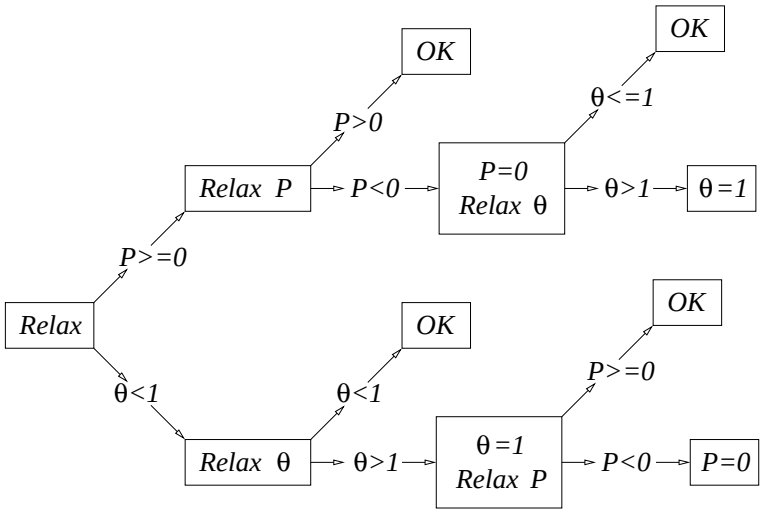


Figure 4.5: Relaxation scheme for starved lubricated contacts.

An additional step is the application of the complementarity condition immediately after the change δ has been applied to either $\tilde{\theta}$ or $\tilde{P}$, i.e. if $\bar{\theta} > 1$ then $\bar{\theta} = 1$ and if $\bar{P} < 0$ then $\bar{P} = 0$ and, if a point changes its state, a verification whether the change was legitimate. The reason for such a scheme is as follows. One can imagine that a change of state from a pressurized point to a starved point, or vice versa, can affect the states of other points in the vicinity of that point. In fact, a single change of state can induce a complete cascade of state-changes which may cause the relaxation scheme to swap between different solutions (or even become unstable). This cascade is prevented by an immediate relaxation on the new variable. That is to say, if a new approximation $\bar{\theta}_{i,j} > 1$, $\bar{\theta}_{i,j} = 1$ and a new update $\bar{P}_{i,j}$ is determined by means of a one-point Gauss-Seidel relaxation. If this new approximation $\bar{P}_{i,j} < 0$, $\bar{P}_{i,j}$ is set to zero. On the other hand, if after the first relaxation, the approximation $\bar{P}_{i,j} < 0$, $\bar{P}_{i,j} = 0$ and a new approximation $\bar{\theta}_{i,j}$ is determined by a one-point Gauss-Seidel change. If this new update $\bar{\theta}_{i,j} > 1$, $\bar{\theta}$ is set to 1. The relaxation scheme thus established is illustrated in Figure 4.5.

4.4 Implementation details

In this section, details with respect to the relaxation and the implementation of the multigrid method to the EHL equations are given.

4.4.1 Relaxation

As stated in the previous section, for fully flooded conditions two types of relaxation are needed, depending on whether a point is in the high pressure zone or in the low pressure zone. Actually, as was stated in Venner [66], the value of ϵ/h_X^2 determines whether Gauss-Seidel line-relaxation or Jacobi distributive relaxation should be applied; Gauss-Seidel relaxation is sufficient if $\epsilon/h_X^2 > 0.3$ and Jacobi distributive relaxation is needed when $\epsilon/h_X^2 \leq 0.3$.

The two discretization forms in the $NU2$ discretization require two different relaxation schemes depending on the ratio h_X/h_T. Hence, two relaxation scheme are needed for both the Gauss-Seidel line-relaxation and the Jacobi distributive relaxation.

If $h_X \leq h_T$, the Gauss-Seidel changes $\delta_{i,j}$ for each line j on time step k thus follow from the system of equations:

$$\begin{aligned}
h_X^{-2}\Big\{\epsilon_{i+1/2,j}\Big((\tilde{P}_{i+1,j}+\delta_{i+1,j})-(\tilde{P}_{i,j}+\delta_{i,j})\Big)\\
-\epsilon_{i-1/2,j}\Big((\tilde{P}_{i,j}+\delta_{i,j})-(\tilde{P}_{i-1,j}+\delta_{i-1,j})\Big)\Big\}\\
+\kappa^2 h_Y^{-2}\Big\{\epsilon_{i,j+1/2}\Big(\tilde{P}_{i,j+1}-(\tilde{P}_{i,j}+\delta_{i,j})\Big)\\
-\epsilon_{i,j-1/2}\Big((\tilde{P}_{i,j}+\delta_{i,j})-\bar{P}_{i,j-1}\Big)\Big\}\\
-\Big(h_X^{-1}-h_T^{-1}\Big)\Big(1.5\ \bar{\rho}\bar{H}_{i,j}-2.0\ \bar{\rho}\bar{H}_{i-1,j}+0.5\ \bar{\rho}\bar{H}_{i-2,j}\Big)\\
-h_T^{-1}\Big(1.5\ \bar{\rho}\bar{H}_{i,j}-2.0\ \bar{\rho}H_{i-1,j,k-1}+0.5\ \bar{\rho}H_{i-2,j,k-2}\Big)=0, \qquad (4.28)
\end{aligned}$$

where $\tilde{P}$ denotes the old and $\bar{P}$ and $\bar{H}$ the new approximations. It is noted that $H_{i-1,j,k-1}$ and $H_{i-2,j,k-2}$ are, respectively, the film thickness on the previous time step and the time step before and thus do not have the bar as superscript.

For $h_X > h_T$, the Gauss-Seidel changes follow from:

$$\begin{aligned}
h_X^{-2}\Big\{&\epsilon_{i+1/2,j}\Big((\tilde{P}_{i+1,j}+\delta_{i+1,j})-(\tilde{P}_{i,j}+\delta_{i,j})\Big)\\
&-\epsilon_{i-1/2,j}\Big((\tilde{P}_{i,j}+\delta_{i,j})-(\tilde{P}_{i-1,j}+\delta_{i-1,j})\Big)\Big\}\\
+\kappa^2 h_Y^{-2}\Big\{&\epsilon_{i,j+1/2}\Big(\tilde{P}_{i,j+1}-(\tilde{P}_{i,j}+\delta_{i,j})\Big)\\
&-\epsilon_{i,j-1/2}\Big((\tilde{P}_{i,j}+\delta_{i,j})-\bar{P}_{i,j-1}\Big)\Big\}\\
-h_X^{-1}&\Big(1.5\ \bar{\rho}\bar{H}_{i,j}-2.0\ \bar{\rho}\bar{H}_{i-1,j,k-1}+0.5\ \bar{\rho}\bar{H}_{i-2,j,k-2}\Big)\\
-\Big(h_T^{-1}-h_X^{-1}\Big)&\Big(1.5\ \bar{\rho}\bar{H}_{i,j}-2.0\ \bar{\rho}H_{i,j,k-1}+0.5\ \bar{\rho}H_{i,j,k-2}\Big)=0. \qquad (4.29)
\end{aligned}$$

As stated earlier, only the three largest contributions to the elastic deformation need to be included. Hence, the new approximation of the film thickness $\bar{H}$ is related to the old approximation $\tilde{H}$ by:

$$\bar{H}_{i,j} = \tilde{H}_{i,j} + \sum_{r=i-1}^{i+1} K_{irjs}\delta_{r,s}. \qquad (4.30)$$

Because the Reynolds equation is non-linear, under-relaxation should be applied, i.e. only a fraction of the calculated changes δ are added to the variables. For the Gauss-Seidel relaxation, these factors range from 0.8 for small values of M down to 0.4, if M is large.

The Jacobi distributive changes for $h_X \geq h_T$, applied when $\epsilon/h_X^2 \leq 0.3$, follow from the system of equations:

$$\begin{aligned}
h_X^{-2}\Big\{&\epsilon_{i+1/2,j}\Big((\tilde{P}_{i+1,j}+\delta_{i+1,j}-(\delta_{i,j}+\delta_{i+2,j})/4)\\
&-(\tilde{P}_{i,j}+\delta_{i,j}-(\delta_{i-1,j}+\delta_{i+1,j})/4)\Big)\\
&-\epsilon_{i-1/2,j}\Big((\tilde{P}_{i,j}+\delta_{i,j}-(\delta_{i-1,j}+\delta_{i+1,j})/4)\\
&-(\tilde{P}_{i-1,j}+\delta_{i-1,j}-(\delta_{i-2,j}+\delta_{i,j})/4)\Big)\Big\}\\
+\kappa^2 h_Y^{-2}\Big\{&\epsilon_{i,j+1/2}\Big((\tilde{P}_{i,j+1}-\delta_{i,j}/4)\\
&-(\tilde{P}_{i,j}+\delta_{i,j}-(\delta_{i-1,j}+\delta_{i+1,j})/4)\Big)\\
&-\epsilon_{i,j-1/2}\Big((\tilde{P}_{i,j}+\delta_{i,j}-(\delta_{i-1,j}+\delta_{i+1,j})/4)\\
&-(\bar{P}_{i,j-1}-\delta_{i,j}/4)\Big)\Big\}\\
-\Big(h_X^{-1}-h_T^{-1}\Big)&\Big(1.5\ \bar{\rho}\bar{H}_{i,j}-2.0\ \bar{\rho}\bar{H}_{i-1,j}+0.5\ \bar{\rho}\bar{H}_{i-2,j}\Big)\\
-h_T^{-1}&\Big(1.5\ \bar{\rho}\bar{H}_{i,j}-2.0\ \bar{\rho}H_{i-1,j,k-1}+0.5\ \bar{\rho}H_{i-2,j,k-2}\Big)=0. \qquad (4.31)
\end{aligned}$$

If $h_X < h_T$, the distributive changes follow from:

$$\begin{aligned}
h_X^{-2} \Big\{ & \epsilon_{i+1/2,j} \Big((\tilde{P}_{i+1,j} + \delta_{i+1,j} - (\delta_{i,j} + \delta_{i+2,j})/4) \\
& - (\tilde{P}_{i,j} + \delta_{i,j} - (\delta_{i-1,j} + \delta_{i+1,j})/4) \Big) \\
& - \epsilon_{i-1/2,j} \Big((\tilde{P}_{i,j} + \delta_{i,j} - (\delta_{i-1,j} + \delta_{i+1,j})/4) \\
& - (\tilde{P}_{i-1,j} + \delta_{i-1,j} - (\delta_{i-2,j} + \delta_{i,j})/4) \Big) \Big\} \\
+ \kappa^2 h_Y^{-2} \Big\{ & \epsilon_{i,j+1/2} \Big((\tilde{P}_{i,j+1} - \delta_{i,j}/4) \\
& - (\tilde{P}_{i,j} + \delta_{i,j} - (\delta_{i-1,j} + \delta_{i+1,j})/4) \Big) \\
& - \epsilon_{i,j-1/2} \Big((\tilde{P}_{i,j} + \delta_{i,j} - (\delta_{i-1,j} + \delta_{i+1,j})/4) \\
& - (\bar{P}_{i,j-1} - \delta_{i,j}/4) \Big) \Big\} \\
- h_X^{-1} & \left(1.5\ \bar{\rho}\bar{H}_{i,j} - 2.0\ \bar{\rho}\bar{H}_{i-1,j,k-1} + 0.5\ \bar{\rho}\bar{H}_{i-2,j,k-2} \right) \\
- \left(h_T^{-1} - h_X^{-1} \right) & \left(1.5\ \bar{\rho}\bar{H}_{i,j} - 2.0\ \bar{\rho}H_{i,j,k-1} + 0.5\ \bar{\rho}H_{i,j,k-2} \right) = 0.
\end{aligned} \tag{4.32}$$

For all distributive changes, the new update of the film thickness $\bar{H}$ follows from:

$$\bar{H}_{i,j} = \tilde{H}_{i,j} + \sum_{r=i-1}^{i+1} \Delta K_{irjs} \delta_{r,s}, \tag{4.33}$$

with ΔK_{irjs} defined in Equation 4.27. Again, note that only the three largest contributions have been included. Relaxation factors for the Jacobi distributive relaxation range from 0.6 down to 0.2 for large values of M.

The value of the mutual approach Δ is determined by the force balance equation or the equation of motion. As stated earlier, since the derivative of the equations with respect to Δ is unknown, one can only replace the derivative by a constant value c_Δ.

In the steady state situation, the new value $\bar{\Delta}$ is thus calculated from the old value $\tilde{\Delta}$, by:

$$\bar{\Delta} = \tilde{\Delta} + c_\Delta \left(\frac{2\pi}{3} - h_X h_Y \sum_{i=0}^{n_X} \sum_{j=0}^{n_Y} P_{i,j} \right). \tag{4.34}$$

For sinusoidally varying loads, the relaxation changes to:

$$\bar{\Delta} = \tilde{\Delta} + c_\Delta \left(\frac{2\pi}{3} \left(1 + A \sin(\Omega_e k h_T) \right) - h_X h_Y \sum_{i=0}^{n_X} \sum_{j=0}^{n_Y} P_{i,j,k} \right). \tag{4.35}$$

If the equation of motion is included, the value is updated according to:

$$\bar{\Delta} = \tilde{\Delta} + c_\Delta \left(F - (\Delta_k + E \frac{3}{2\pi} h_X h_Y \sum_{i=0}^{n_X} \sum_{j=0}^{n_Y} P_{i,j,k}) \right), \tag{4.36}$$

where E and F are defined in Equations 4.12 and 4.13. c_Δ ranges from 0.2 down to 0.05 for the finer grids, see also Section 4.4.2.

Since the convergence of the Reynolds equation dominates the convergence of the system of equations, the force balance equation/equation of motion only needs to be relaxed after a number of relaxations, say 30, of the Reynolds equation.

Starved lubrication

As stated above, for starved lubricated conditions the Reynolds equation was discretized using the $SU2$ discretization. The Gauss-Seidel changes on the line j thus follow from:

$$\begin{aligned}
h_X^{-2}\Big\{&\epsilon_{i+1/2,j}\Big((\tilde{P}_{i+1,j}+\delta_{i+1,j})-(\tilde{P}_{i,j}+\delta_{i,j})\Big)\\
&-\epsilon_{i-1/2,j}\Big((\tilde{P}_{i,j}+\delta_{i,j})-(\tilde{P}_{i-1,j}+\delta_{i-1,j})\Big)\Big\}\\
+\kappa^2 h_Y^{-2}\Big\{&\epsilon_{i,j+1/2}\Big(\tilde{P}_{i,j+1}-(\tilde{P}_{i,j}+\delta_{i,j})\Big)\\
&-\epsilon_{i,j-1/2}\Big((\tilde{P}_{i,j}+\delta_{i,j})-\bar{P}_{i,j-1}\Big)\Big\}\\
-h_X^{-1}&\Big(1.5\;\theta\bar{\rho}\bar{H}_{i,j}-2.0\;\theta\bar{\rho}\bar{H}_{i-1,j}+0.5\;\theta\bar{\rho}\bar{H}_{i-2,j}\Big)\\
-h_T^{-1}&\Big(1.5\;\theta\bar{\rho}\bar{H}_{i,j}-2.0\;\theta\bar{\rho}H_{i,j,k-1}+0.5\;\theta\bar{\rho}H_{i,j,k-2}\Big)=0,
\end{aligned}\tag{4.37}$$

where $\tilde{P}$ denotes the old and $\bar{P}$ the new approximations. $\bar{H}$ is the new approximation for H and is given in Equation 4.30.

The distributive changes follow from the system of equations:

$$\begin{aligned}
h_X^{-2}\Big\{&\epsilon_{i+1/2,j}\Big((\tilde{P}_{i+1,j}+\delta_{i+1,j}-(\delta_{i,j}+\delta_{i+2,j})/4)\\
&-(\tilde{P}_{i,j}+\delta_{i,j}-(\delta_{i-1,j}+\delta_{i+1,j})/4)\Big)\\
&-\epsilon_{i-1/2,j}\Big((\tilde{P}_{i,j}+\delta_{i,j}-(\delta_{i-1,j}+\delta_{i+1,j})/4)\\
&-(\tilde{P}_{i-1,j}+\delta_{i-1,j}-(\delta_{i-2,j}+\delta_{i,j})/4)\Big)\Big\}\\
+\kappa^2 h_Y^{-2}\Big\{&\epsilon_{i,j+1/2}\Big((\tilde{P}_{i,j+1}-\delta_{i,j}/4)\\
&-(\tilde{P}_{i,j}+\delta_{i,j}-(\delta_{i-1,j}+\delta_{i+1,j})/4)\Big)\\
&-\epsilon_{i,j-1/2}\Big((\tilde{P}_{i,j}+\delta_{i,j}-(\delta_{i-1,j}+\delta_{i+1,j})/4)\\
&-(\bar{P}_{i,j-1}-\delta_{i,j}/4)\Big)\Big\}\\
-h_X^{-1}&\Big(1.5\;\theta\bar{\rho}\bar{H}_{i,j}-2.0\;\theta\bar{\rho}\bar{H}_{i-1,j}+0.5\;\theta\bar{\rho}\bar{H}_{i-2,j}\Big)\\
-h_T^{-1}&\Big(1.5\;\theta\bar{\rho}\bar{H}_{i,j}-2.0\;\theta\bar{\rho}H_{i,j,k-1}+0.5\;\theta\bar{\rho}H_{i,j,k-2}\Big)=0,
\end{aligned}\tag{4.38}$$

where $\bar{H}$ is given in Equation 4.33.

The Gauss-Seidel changes for θ, included in the system of equations for the changes for P, follow from:

$$\begin{aligned} h_X^{-2}\left(\epsilon_{i+1/2,j}\left(\tilde{P}_{i+1,j}-\tilde{P}_{i,j}\right)-\epsilon_{i-1/2,j}\left(\tilde{P}_{i,j}-\tilde{P}_{i-1,j}\right)\right) \\ +\kappa^2 h_Y^{-2}\left(\epsilon_{i,j+1/2}\left(\tilde{P}_{i,j+1}-\tilde{P}_{i,j}\right)-\epsilon_{i,j-1/2}\left(\tilde{P}_{i,j}-\tilde{P}_{i,j-1}\right)\right) \\ -h_X^{-1}\left(1.5\,(\tilde{\theta}_{i,j}+\delta_{i,j})\bar{\rho}_{i,j}\bar{H}_{i,j}-2.0\,(\tilde{\theta}_{i-1,j}+\delta_{i-1,j})\bar{\rho}_{i-1,j}\bar{H}_{i-1,j}\right. \\ \left.+0.5\,(\tilde{\theta}_{i-2,j}+\delta_{i-2,j})\bar{\rho}_{i-2,j}\bar{H}_{i-2,j}\right) \\ -h_T^{-1}\left(1.5\,(\tilde{\theta}_{i,j}+\delta_{i,j})\bar{\rho}_{i,j}\bar{H}_{i,j}-2.0\,\theta\bar{\rho}H_{i,j,k-1}+0.5\,\theta\bar{\rho}H_{i,j,k-2}\right)=0. \end{aligned} \tag{4.39}$$

4.4.2 Additional insights

With respect to the multigrid solver, some remarks must be made.

First of all, the FMG algorithm should start at a sufficiently fine level. Otherwise, the discrete approximation is an inaccurate approximation of the continuum solution. Especially for high loads, this inaccuracy leads to negative film thicknesses and, as a result, the relaxation scheme is unstable.

A convenient initial approximation for the pressure is the dry contact pressure for the initial value of the mutual approach Δ, i.e. $P(X,Y) = \sqrt{\Delta - X^2 - Y^2}$. The initial approximation for the film thickness follows from the elastic deformation associated with the initial pressure.

If, after a number of relaxations on the coarsest grid, the error is below the discretization error, the solution is interpolated to the next finer grid. For this purpose, a fourth order interpolation suffices.

Subsequently, the solution on the next finer grid is updated by means of the coarse grid correction cycles. Since the Reynolds equation is non-linear, the Full Approximation Scheme (FAS) should be applied to all equations involved, i.e. the Reynolds equations, the film thickness equation and the force balance equation/equation of motion, see Appendix D. This implies that the right-hand side of Equations 4.28, 4.29, 4.31, 4.32, 4.38 and 4.39, should be replaced by the FAS right-hand terms. Also, the FAS right-hand side should be added to the right-hand side of Equations 4.30 and 4.33. For Equations 4.34, 4.35 and 4.36, the terms $2\pi/3$, $2\pi/3 \cdot (1 + A\sin(\Omega_e T))$ and F should be replaced by the appropriate FAS right-hand side.

With respect to the relaxation of the force balance equation and the equation of motion, it is noted that a change in Δ only introduces global changes in the solution. Therefore, since the relaxation of the force balance

or equation of motion affects low-frequency components in the solution, it should be performed on the coarsest grid only.

For fully flooded conditions, the Reynolds equation is not valid in the cavitation zone. Hence, injection should be applied in the residual transfer from the fine to the coarse grid. For starved lubricated contacts however, full weighting may be applied, as the modified Reynolds equation is valid in both the starved and the pressurized region.

Injection should be used in the restriction operator of the solution to the coarse grid at both the cavitation boundary in fully flooded situations and at the position of the meniscus in starved lubricated conditions.

With respect to the interpolation of the solution to the next finer grid in the FMG algorithm, higher order interpolation should only be applied if all points involved in the interpolation are either starved of fully flooded. Linear interpolation can be used otherwise.

Finally, it is noted that generally 3 or 4 $V(2,1)$ cycles per level are sufficient to converge below the discretization error.

4.5 Multilevel multi-integration

Obviously, the complexity of the FMG algorithm is only $O(n)$, if the individual components of the algorithm are, at most, of the same complexity. Unfortunately, it is this assumption that does not hold for the EHL equations; plain evaluation of the deformation integral (the last term in Equation 3.4) is an $O(n^2)$ process.

Similar to the multigrid technique, a reduction of this complexity may be achieved by the so-called multilevel multi-integration method (MLMI), developed by Brandt and Lubrecht [12]. The reduction is achieved by exploiting smoothness properties of the elastic deformation kernel. Although the kernel is singular near $(X', Y') = (X, Y)$, sufficiently far away from this point the kernel is smooth. Hence, in both the X- and X'-direction, as well as in the Y- and Y'-direction, the kernel in "odd" points may thus be interpolated from its values in the "even" points. In doing so, Brandt and Lubrecht managed to reduce the evaluation time to $O(n \ln n)$. The deviation from the optimal $O(n)$ complexity is due to the $O(\ln n)$ corrections for the large interpolation errors near the singularity.

Résumé

The elastic deformation integral is an example of so-called integral transforms. In terms of the notation of Brandt and Lubrecht, the two-dimensional

integral transform reads:

$$w(x,y) = \iint_S K(x,x',y,y')u\,(x',y')\,dx'dy', \quad (x,x',y,y') \in S. \tag{4.40}$$

In this relation, $K(x,x',y,y')$ denotes the kernel of the integral transform and $u(x',y')$ is a given function.

According to Equation 3.14, the elastic deformation kernel, in the notation of Equation 4.40, is:

$$K(x,x',y,y') = \frac{1}{\pi\mathcal{K}} \frac{1}{\sqrt{\kappa^2\,(x-x')^2 + (y-y')^2}}. \tag{4.41}$$

As an illustration, Figure 4.6 shows the circular kernel ($\kappa = 1$) and an example of an elliptic kernel ($\kappa = 0.1$), as a function of $|x-x'|$ and $|y-y'|$. Obviously, because of the singularity, the function is truncated in the Figure.

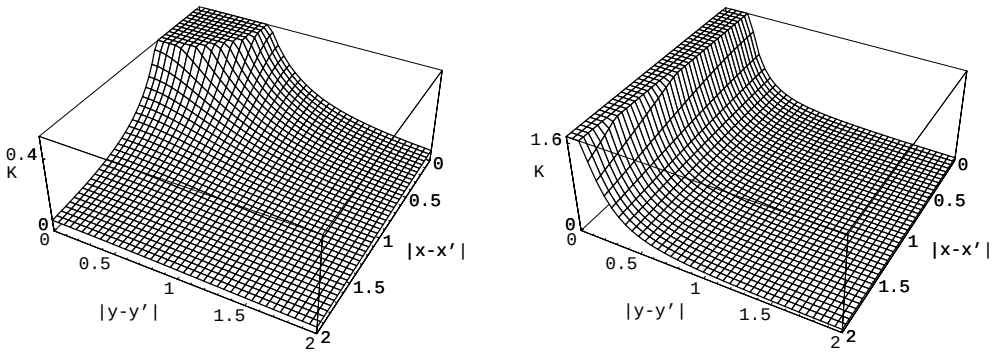


Figure 4.6: The circular kernel ($\kappa = 1$, left) and an example of an elliptic kernel ($\kappa = 0.1$, right).

The discrete analogue of Equation 4.40 is

$$w_{i,j} = \sum_{r=0}^{n_x} \sum_{s=0}^{n_y} K_{irjs} u_{r,s}, \tag{4.42}$$

which shows that the evaluation of all elastic deformations is an $O(n^2)$ process. For large number of unknowns, these $O(n^2)$ operations will eventually dominate the total number of operations and computing times will increase accordingly.

A reduction of the complexity can be achieved by exploiting the fact that, since a numerical approach is adopted anyway, the discrete deformation does not need to be evaluated exactly. An approximation of the deformation

is sufficiently accurate if the difference between the approximation and the exact solution is dominated by the discretization error. Indeed, under this restriction, multilevel multi-integration will evaluate the integral in $O(n \ln n)$ operations.

From this point, it is assumed that the reader is familiar with MLMI for the circular contact kernel, i.e. for $\kappa = 1$. For those unfamiliar with the method, a detailed discussion can be found in Appendix E. In this section, apart from a brief résumé, we will focus on the modifications required for small values of κ.

As explained in the introduction, MLMI exploits smoothness properties of the discrete elastic kernel by interpolating the discrete kernel in "odd" (fine grid) points, from its values in the "even" (coarse grid) points. Note that the interpolation is performed in the x-, x'-, y- and y'-directions.

To illustrate the method, consider the one-dimensional multi-summation:

$$w_i^h = \sum_j K_{i,j}^{hh} u_j^h, \tag{4.43}$$

for which the discrete kernel $K_{i,j}^{hh}$ is smooth with respect to the mesh size h and non-singular. In addition, let $\tilde{K}_{i,j}^{hh}$ be a coarse grid approximation of the discrete kernel $K_{i,j}^{hh}$, defined by:

$$\tilde{K}_{i,j}^{hh} \equiv \left[I\!I_H^h K_{i,\cdot}^{hH} \right]_j, \tag{4.44}$$

where $I\!I_H^h$ is the interpolation matrix. The index on which $I\!I_H^h$ works is denoted by a dot, i.e. in this case the interpolation works on the index j. If the kernel is sufficiently smooth, the summation can now be approximated by:

$$\begin{aligned} w_i^h \approx \tilde{w}_i^h &\equiv \sum_j \tilde{K}_{i,j}^{hh} u_j^h = \sum_j \left[I\!I_H^h K_{i,\cdot}^{hH} \right]_j u_j^h \\ &= \sum_J K_{i,J}^{hH} \left[(I\!I_H^h)^T u_\cdot^h \right]_J \equiv \sum_J K_{i,J}^{hH} u_J^H. \end{aligned} \tag{4.45}$$

The operator $(I\!I_H^h)^T$ is generally referred to as anterpolation, since it is the adjoint (or transpose) of the interpolation matrix. Anterpolation thus maps u_j^h onto a coarse grid variable u_J^H (J denotes the grid index on the coarser grid whose mesh size is H). Moreover, rewriting the summation in this form, reduces the fine grid summation to a summation over coarse grid points only. Apart from the operations needed for the anterpolation, the number of operations is thus reduced by a factor 2. Note that the work associated with the anterpolation is only $O(n)$ and thus does not add to the complexity of the method.

Assuming that the kernel has similar smoothness properties in the x-direction, which is true for the elastic deformation kernel, the kernel may be interpolated in this direction as well. Hence, defining

$$\hat{K}^{hh}_{i,j} = \left[I^h_H K^{Hh}_{\cdot,j} \right]_i , \tag{4.46}$$

the summation may also be approximated by:

$$w^h_i \approx \left[I^h_H w^H_\cdot \right]_i , \tag{4.47}$$

where w^H_I is the summation evaluated in the coarse grid points. Also in this case, the number of operations is reduced by a factor 2.

The two steps may be combined in such a way that both the summation and the evaluation of the integral transform extend over the coarse grid points only. The integral transform in the fine grid points is then calculated by means of interpolation of the values in the coarse grid points. Neglecting the intergrid transfers, the amount of work is thus reduced by a factor 4.

This may not seem much of an improvement. However, with respect to the coarser grid, the kernel may still be smooth and an additional coarsening procedure can be applied. In fact, the coarsening may be applied recursively, until a grid is reached for which the number of points is only $O(\sqrt{n})$. On this grid, plain summation requires $O(n)$ operations only. Consequently, an optimal method is obtained, since the anterpolation and interpolation also require $O(n)$ operations.

The method is illustrated in Figure 4.7, where the index k is used to denote the grid number.

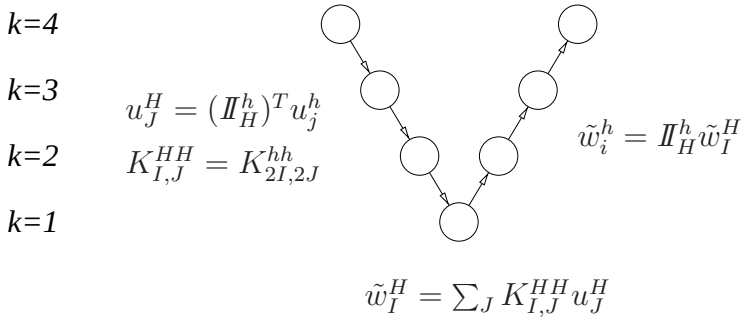


Figure 4.7: The multilevel multi-integration cycle.

Unfortunately, the elastic deformation kernel is singular in $(x', y') = (x, y)$. Only sufficiently far away from the singularity is the kernel again smooth; a property referred to as singular-smooth. MLMI may be applied to singular-smooth kernels, provided one corrects for the large errors introduced by the

interpolation from points near the singularity. Brandt and Lubrecht showed that, for the one-dimensional kernel $\ln|x - x'|$, the $m = 3 + 2\ln n$ points, closest to the singularity, need to be corrected. The complexity of the method for the singular smooth kernel is thus $O(n \ln n)$, which, for large n, is very close to the $O(n)$ complexity.

For two-dimensional kernels, the method described above can easily be extended by alternately coarsening in the x- and x'- directions and the y- and y'- directions. Then, the anterpolation and interpolation procedures are essentially the same as in the one-dimensional case and, as a result, so is the complexity of the method. Furthermore, despite the additional storage required for the "half coarsened grids," it is very effective and easy to extend to higher dimensions.

It must be realized that, for two-dimensional kernels, also the correction patch is two-dimensional. Hence, a correction size in the direction of interpolation as well as perpendicular to the direction of interpolation is required. Based on the correction size for the one-dimensional kernel, Brandt and Lubrecht concluded and verified that the number of corrections in the direction of interpolation should equal $m_1 = 3 + 0.5 \ln n$ and the number of corrections perpendicular to the direction of interpolation should be $m_2 = 2$, for the additional error to be below the discretization error.

Ellipticity

Essentially, MLMI remains unchanged when ellipticity is included. That is, the values of the kernel in the fine grid points may be approximated by the values of the kernel in the coarse grid points. However, as Figure 4.6 already suggests, the smoothness of the kernel in the x-direction is now different from the smoothness in the y-direction. (Note that, like the y- and y'-directions, the properties of the kernel with respect to the x'-coordinate are the same as those with respect to the x-coordinate. In this section, we will therefore assume $x' = 0$ and $y' = 0$.) Ellipticity thus only affects the size of the correction patch.

The intention of the next analysis is not to give a rigorous proof of the actual size of the correction patch. Instead, our goal will be to examine how it should be changed, compared to the correction patch for the circular contact, in order to obtain errors which are small compared to the discretization error. The comparison of the Hertzian elastic deformation with the discrete deformation obtained using MLMI, will verify whether the size of the correction patch satisfies this restriction.

It is known from standard numerical analysis that, for p^{th} order interpolation, the p^{th} order derivative of the interpolated function is a measure for

the interpolation error, see Atkinson [3]. Since a function is termed smooth if the p^{th} order derivative is sufficiently small, smooth functions have small interpolation errors.

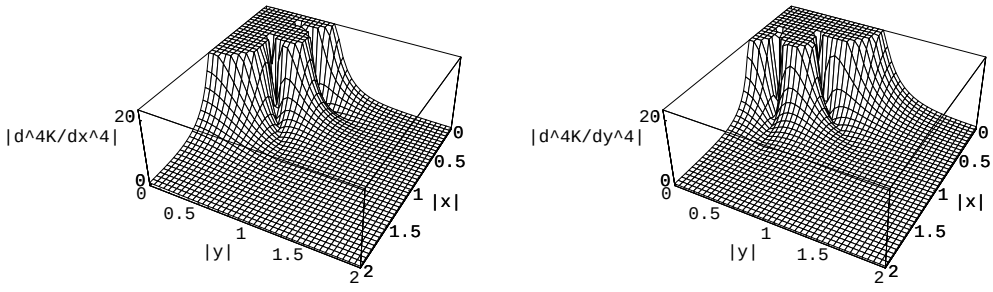


Figure 4.8: $|\partial^4 K/\partial x^4|$ (left) and $|\partial^4 K/\partial y^4|$ (right) for the circular contact kernel ($\kappa = 1.0$).

Figure 4.8 shows the absolute value of the 4^{th} order derivative of the circular kernel with respect to the x-direction ($|\partial^4 K/\partial x^4|$) as well as the y-direction ($|\partial^4 K/\partial y^4|$), for $x' = 0$ and $y' = 0$. From the Figure it can clearly be seen that in both pictures values are large in an approximately circular shaped patch. Thus, instead of the narrow rectangle proposed by Brandt and Lubrecht, the Figure suggests that the shape of the correction patch should be approximately circular. Apparently, the total error introduced by the interpolation is so small, that only a limited number of these largest errors require correction.

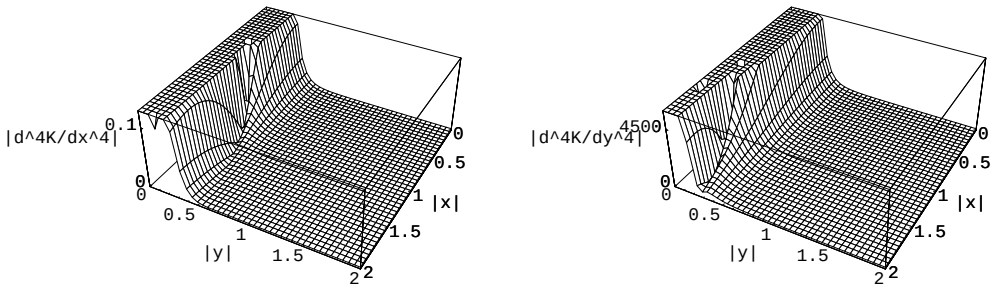


Figure 4.9: $|\partial^4 K/\partial x^4|$ (left) and $|\partial^4 K/\partial y^4|$ (right) for the elliptical contact kernel ($\kappa = 0.1$).

For $\kappa = 0.1$, Figure 4.9 shows the absolute value of the 4^{th} order derivative in the x- and the y-directions. One observes immediately that, for both directions, the largest values occur in a narrow strip close to $y = 0$. For the derivative in the x-direction however, the values are much smaller than those observed for the circular kernel. Instead, the values are much larger for the derivative in the y-direction.

Hence, the two correction sizes m_1 and m_2, which were sufficient for the circular kernel, no longer suffice. In fact, four correction sizes need to be defined. A correction size m_1^x will be used to denote the number of points of the correction patch in the direction of interpolation, if interpolation is in the x-direction. The number of points perpendicular to the direction of interpolation is denoted by m_2^x. Similarly, m_1^y will denote the number of points in the direction of interpolation, if it is in the y-direction, whereas m_2^y denotes the number in the direction perpendicular to it.

Compared to the circular kernel, the most significant contributions to the error are thus introduced by interpolation in the y-direction. In particular, compared to the correction size in the direction of interpolation (m_1^y), the Figures show that perpendicular to this direction (m_2^y), the correction size should be much larger. Evidently, interpolation from values which include the singularity should be avoided. This restricts the reduction of the value of m_1^x.

An obvious choice would thus be to make m_2^y proportional to $1/\kappa$ and, accordingly, m_2^x proportional to κ.

Verification

The actual values of the correction sizes have been determined from a numerical experiment. For this purpose, consider the elastic deformation integral:

$$w(x,y) = \frac{1}{\pi\mathcal{K}} \iint_S \frac{u\left(x',y'\right) dx'dy'}{\sqrt{\kappa^2\left(x-x'\right)^2 + \left(y-y'\right)^2}}. \tag{4.48}$$

If $u(x,y) = \sqrt{1-x^2-y^2}$, which is similar to the Hertzian pressure distribution, the solution in the unit circle $x^2 + y^2 < 1$, is known analytically:

$$w(x,y) = 1 - \mathcal{S}x^2 - (1-\mathcal{S})y^2. \tag{4.49}$$

In this equation, $\mathcal{S}$ is the shape function defined in Equation 3.15.

To monitor the error in the multilevel multi-integration, the error norm E_k^l is defined according to:

$$E_k^l = h_x h_y \sum_{i=0}^{n_x} \sum_{j=0}^{n_y} |w_{i,j}^{k,l} - w(x_i, y_j)|, \tag{4.50}$$

where $w_{i,j}^{k,l}$ is the discrete approximation on level l, for which the multi-summation has been carried out on level k. Hence, E_k^l is the average absolute error of the integral and, for $k = l$, the norm reduces to the L_1 norm of the discretization error. Then, if $E_k^l \approx E_l^l$ at a level k for which the number of points is $O(\sqrt{n})$, the error introduced by MLMI is small compared to the discretization error and the summation would cost $O(n)$ operations.

Based on this condition, the following correction sizes were obtained: $m_1^x = m_1^y = p + 4$, $m_2^x = 2\kappa + 4$ and $m_2^y = 2/\kappa + 4$, where p is the order of interpolation. The values have been verified for values of κ down to 0.0125 and up to 20.

Tables 4.1 and 4.2, which show the E_k^l-norm for $\kappa = 0.63$, confirm the correctness of the correction sizes. The interpolation orders are, respectively, 6 and 8. The calculational domain is $-2 \leq x \leq 2$ and $-2 \leq y \leq 2$. The coarsest level, $l = 1$, contains $(4 + 1)(4 + 1)$ grid points, level $l = 2$ contains $(8 + 1)(8 + 1)$ points, etc. Hence, the mesh sizes h_x and h_y decrease by a factor 2 for each level.

l	$k = l$	$k = l - 1$	$k = l - 2$	$k = l - 3$	$k = l - 4$	$k = l - 5$	$k = l - 6$
1	$2.87\ 10^{-1}$						
2	$1.58\ 10^{-1}$	$1.58\ 10^{-1}$					
3	$4.66\ 10^{-2}$	$4.66\ 10^{-2}$	$4.66\ 10^{-2\star}$				
4	$1.55\ 10^{-2}$	$1.55\ 10^{-2}$	$1.55\ 10^{-2\star}$	$1.55\ 10^{-2}$			
5	$3.05\ 10^{-3}$	$3.04\ 10^{-3}$	$3.04\ 10^{-3}$	$3.04\ 10^{-3\star}$	$3.04\ 10^{-3}$		
6	$9.46\ 10^{-4}$	$9.38\ 10^{-4}$	$9.28\ 10^{-4}$	$9.25\ 10^{-4\star}$	$9.25\ 10^{-4}$	$9.25\ 10^{-4}$	
7	$2.46\ 10^{-4}$	$2.42\ 10^{-4}$	$2.34\ 10^{-4}$	$2.23\ 10^{-4}$	$2.20\ 10^{-4\star}$	$2.20\ 10^{-4}$	$2.21\ 10^{-4}$
8	$8.92\ 10^{-5}$	$8.68\ 10^{-5}$	$8.21\ 10^{-5}$	$7.41\ 10^{-5}$	$6.35\ 10^{-5\star}$	$6.09\ 10^{-5}$	$6.09\ 10^{-5}$

Table 4.1: E_k^l-norm for $\kappa = 0.63$ ($R_x/R_y = 0.5$) using 6^{th} order transfer.

l	$k = l$	$k = l - 1$	$k = l - 2$	$k = l - 3$	$k = l - 4$	$k = l - 5$	$k = l - 6$
1	$2.87\ 10^{-1}$						
2	$1.58\ 10^{-1}$	$1.58\ 10^{-1}$					
3	$4.66\ 10^{-2}$	$4.66\ 10^{-2}$	$4.66\ 10^{-2\star}$				
4	$1.55\ 10^{-2}$	$1.55\ 10^{-2}$	$1.55\ 10^{-2\star}$	$1.55\ 10^{-2}$			
5	$3.05\ 10^{-3}$	$3.06\ 10^{-3}$	$3.06\ 10^{-3}$	$3.06\ 10^{-3\star}$	$3.05\ 10^{-3}$		
6	$9.46\ 10^{-4}$	$9.47\ 10^{-4}$	$9.48\ 10^{-4}$	$9.48\ 10^{-4\star}$	$9.48\ 10^{-4}$	$9.46\ 10^{-4}$	
7	$2.46\ 10^{-4}$	$2.47\ 10^{-4}$	$2.47\ 10^{-4}$	$2.48\ 10^{-4}$	$2.48\ 10^{-4\star}$	$2.48\ 10^{-4}$	$2.46\ 10^{-4}$
8	$8.92\ 10^{-5}$	$8.94\ 10^{-5}$	$8.98\ 10^{-5}$	$9.05\ 10^{-5}$	$9.13\ 10^{-5\star}$	$9.15\ 10^{-5}$	$9.15\ 10^{-5}$

Table 4.2: E_k^l-norm for $\kappa = 0.63$ ($R_x/R_y = 0.5$) using 8^{th} order transfer.

These Tables indeed confirm that the additional error introduced by MLMI,

on a level which has approximately $O(\sqrt{n})$ points (indicated by the stars in the Table) is small compared to the discretization error in L_1 norm. In addition, it shows that 6^{th} order interpolation suffices.

To show that also for smaller values of κ, the correction sizes are correct, the E_k^l-norm for $\kappa = 0.22$ and $\kappa = 0.055$ is shown in Tables 4.3 and 4.4. Indeed, also in this case, 6^{th} order interpolation is sufficient.

l	$k = l$	$k = l-1$	$k = l-2$	$k = l-3$	$k = l-4$	$k = l-5$	$k = l-6$
1	$3.01\ 10^{-1}$						
2	$1.53\ 10^{-1}$	$1.53\ 10^{-1}$					
3	$4.47\ 10^{-2}$	$4.47\ 10^{-2}$	$4.47\ 10^{-2\star}$				
4	$1.50\ 10^{-2}$	$1.50\ 10^{-2}$	$1.50\ 10^{-2\star}$	$1.50\ 10^{-2}$			
5	$2.94\ 10^{-3}$	$2.92\ 10^{-3}$	$2.92\ 10^{-3}$	$2.92\ 10^{-3\star}$	$2.92\ 10^{-3}$		
6	$9.19\ 10^{-4}$	$9.06\ 10^{-4}$	$8.88\ 10^{-4}$	$8.84\ 10^{-4\star}$	$8.84\ 10^{-4}$	$8.84\ 10^{-4}$	
7	$2.40\ 10^{-4}$	$2.33\ 10^{-4}$	$2.18\ 10^{-4}$	$2.00\ 10^{-4}$	$1.95\ 10^{-4\star}$	$1.95\ 10^{-4}$	$1.96\ 10^{-4}$
8	$8.82\ 10^{-5}$	$8.41\ 10^{-5}$	$7.57\ 10^{-5}$	$6.10\ 10^{-5}$	$4.26\ 10^{-5\star}$	$3.79\ 10^{-5}$	$3.78\ 10^{-5}$

Table 4.3: E_k^l-norm for $\kappa = 0.22$ ($R_x/R_y = 0.1$) using 6^{th} order transfer.

l	$k = l$	$k = l-1$	$k = l-2$	$k = l-3$	$k = l-4$	$k = l-5$	$k = l-6$
1	$3.18\ 10^{-1}$						
2	$1.46\ 10^{-1}$	$1.46\ 10^{-1}$					
3	$4.09\ 10^{-2}$	$4.09\ 10^{-2}$	$4.09\ 10^{-2\star}$				
4	$1.41\ 10^{-2}$	$1.41\ 10^{-2}$	$1.41\ 10^{-2\star}$	$1.41\ 10^{-2}$			
5	$2.74\ 10^{-3}$	$2.72\ 10^{-3}$	$2.72\ 10^{-3}$	$2.72\ 10^{-3\star}$	$2.72\ 10^{-3}$		
6	$8.70\ 10^{-4}$	$8.44\ 10^{-4}$	$8.25\ 10^{-4}$	$8.17\ 10^{-4\star}$	$8.17\ 10^{-4}$	$8.18\ 10^{-4}$	
7	$2.33\ 10^{-4}$	$2.05\ 10^{-4}$	$1.75\ 10^{-4}$	$1.55\ 10^{-4}$	$1.49\ 10^{-4\star}$	$1.49\ 10^{-4}$	$1.50\ 10^{-4}$
8	$8.62\ 10^{-5}$	$6.71\ 10^{-5}$	$3.88\ 10^{-5}$	$4.98\ 10^{-5}$	$6.35\ 10^{-5\star}$	$6.87\ 10^{-5}$	$6.88\ 10^{-5}$

Table 4.4: E_k^l-norm for $\kappa = 0.055$ ($R_x/R_y = 0.01$) using 6^{th} order transfer.

It has thus been established that, for interpolation in the y-direction, the correction patch becomes very large in the x-direction. Unfortunately, for very small values of κ, these corrections may start to dominate the total number of operations. It might be worthwhile to examine whether the evaluation of these corrections could be done by a similar MLMI procedure. It has in fact been found that the correction kernel, i.e. the kernel minus the interpolated kernel, is smooth in the x- and the x'-direction and thus may be interpolated from its values on coarser grids. Since the correction patch does not extend over all points in the domain (for $\kappa = 0.05$, $m_2^y = 40$ and the grids for which the method is designed to work contains about $257x257$ or $513x513$ points),

the implementation might however be cumbersome. In addition, it is not obvious whether the number of corrections is large enough for MLMI to be more efficient than plain summation.

Chapter 5

EHL Circular contact

The objective of the present work is to study the behaviour of EHL contacts for cases where the structural elements vibrate as a result of force variations (applied or inertia induced). This chapter shows the influence of vibrations on the film thickness and the pressure distribution of the *circular* contact. Also the flexibility (stiffness) and damping of the contact will be discussed.

From the steady state solution, the flexibility of the contact is derived. It is computed for a wide range of values of the governing parameters. An analytical curve-fit function is presented that accurately predicts the flexibility.

To study damping, clearly the time dependent problem must be considered. The characteristic behaviour of time dependent solutions, for the case of vibrations resulting from the inertia of the rolling element, is discussed. It is shown that small oscillations are indeed damped by the viscous forces in the lubricant film.

To quantify this damping, simulations for the case of sinusoidally varying loads are performed. Damping values for a wide parameter range are obtained and a simple analytical curve-fit function is presented that approximates the calculated values.

5.1 Steady state solution

This section focuses on the stiffness of the EHL circular contact model. First, we will discuss the film thickness and pressure distribution and how they change as the parameters change. Next, the accuracy of the presented solutions is estimated.

5.1.1 Pressure and film thickness

Figures 5.1 and 5.2 show typical examples of the pressure distribution and the film thickness distribution in an EHL circular contact. The Moes dimensionless parameters for the case presented are $M = 100$ and $L = 5$. The parameters based on the Hertzian solution are $\lambda = 2.37\ 10^{-3}$ and $\bar{\alpha} = 8.46$.

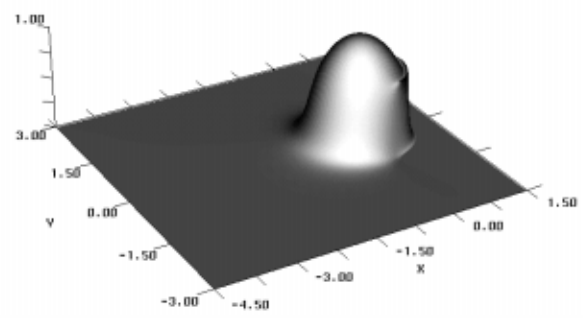


Figure 5.1: Pressure distribution $P(X, Y)$ for $M = 100$ and $L = 5$.

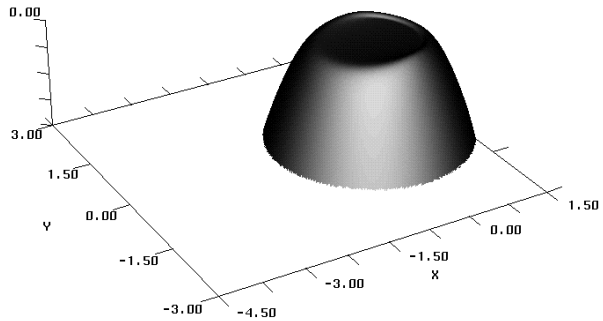


Figure 5.2: Film thickness distribution $H(X, Y)$, plotted upside-down, for $M = 100$ and $L = 5$.

The pressure in the contact closely resembles the Hertzian semi-ellipsoidal distribution. It deviates from this distribution in the inlet region and in the outlet region of the contact. In the inlet region, instead of the discontinuity in the pressure gradient, occurring for dry contacts, now a smooth pressure distribution is obtained. In the outlet, an additional local pressure maximum occurs. This local maximum is known as the "pressure spike," first discovered by Petrusevich [59], and has been the subject of many discussions, see

Kostreva [41], Hamrock, Lee and Houpert [26], Lubrecht, Breukink, Moes, ten Napel and Bosma [46], Baumann, Von Frey and Haller [7], Hamrock, Pan and Lee [27] and Venner [66]. The spike is a consequence of the exponential increase of the viscosity with pressure. Its height depends on the density and consequently on the density pressure relation that is used.

Figure 5.2 shows that the film thickness is approximately constant in the high pressure zone. The minimum film thickness is generally found in the so-called side lobes.

Figure 5.3 shows the film thickness of Figure 5.2, by means of an interference plot[1]. Interference plots will be used extensively in this thesis, not only because the film thickness distribution is observed more clearly (the horseshoe shaped constriction is clearly visible), but also because these interference plots can actually be observed experimentally, see Chapter 8.

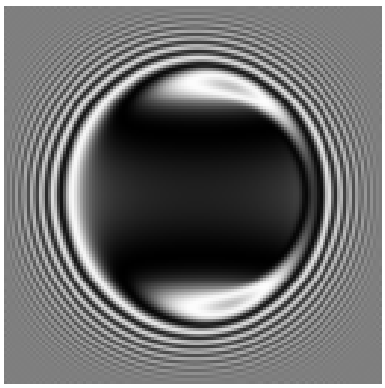


Figure 5.3: Interference plot of film thickness for $M = 100$ and $L = 5$. Flow is from left to right.

The influence of the load parameter M on the film thickness and pressure distribution is illustrated in Figure 5.4. The figure shows the film thickness and corresponding pressure along the centre line ($Y = 0$) and in a direction normal to the centre line ($X = 0$), for different values of M. The value of the lubricant number L is 5.

[1]In the interference plot the film thickness is shown by means of the intensity I. It is defined according to:

$$I(X, Y, T) = 0.5 + 0.5 \cos\left(\frac{2\pi H(X, Y, T)}{\Lambda}\right), \tag{5.1}$$

where Λ is the dimensionless wavelength. The grey tones in the interference plot range from white for $I = 1$ to black for $I = 0$. Hence for every other fringe, the film thickness differs by an amount which is equal to the wavelength Λ.

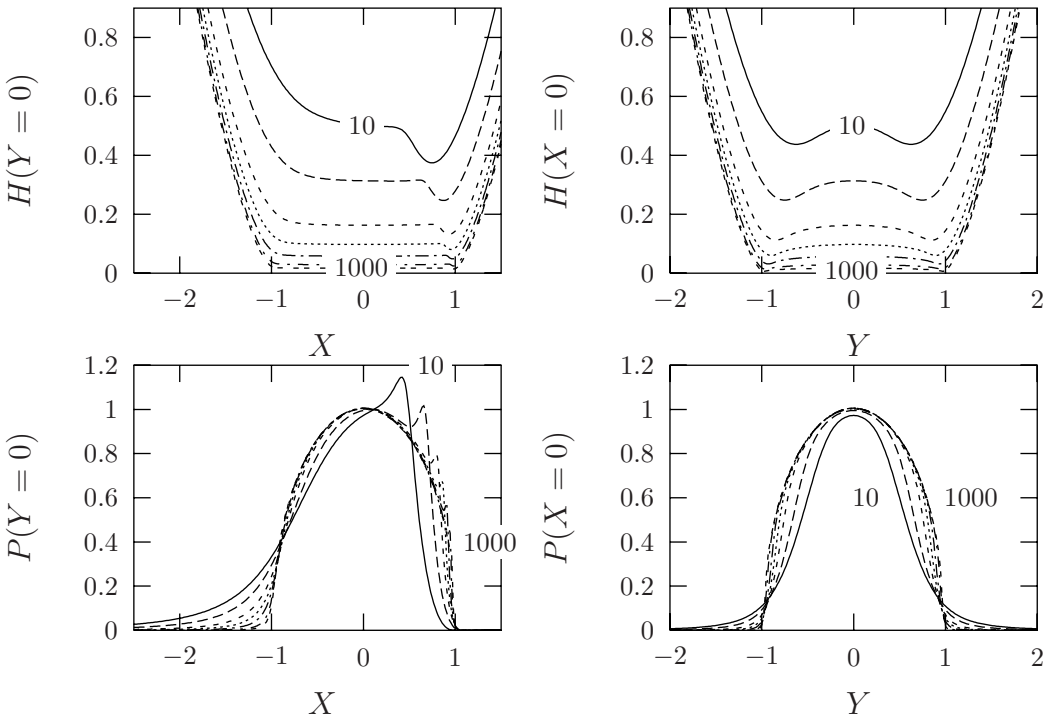


Figure 5.4: Film thickness and pressure along the centre line (left) and perpendicular to the centre line (right) for $L = 5$ and $M = 10$, 20, 50, 100, 200, 500 and 1000.

As one would expect, with increasing load (M), the film thickness decreases. Also, the film thickness becomes more uniform within the high pressure zone, and the pressure distribution resembles the Hertzian distribution more and more.

Figure 5.5 illustrates the influence of the value of the lubricant parameter L on the film thickness and pressure distribution. The value of the load parameter M is 100. The Figure shows that the film thickness increases with increasing values of L. This illustrates the beneficial effect of pressure dependent viscosity for film formation. In addition, it shows that the film thickness is indeed minimal in the side lobes. The variation of the pressure profile appears to be much smaller as the variation observed in Figure 5.4 for varying M. The overall shape is still roughly Hertzian, however, with increasing L, the pressure rise is steeper in the inlet region. Note that, although difficult to distinguish because of the small differences, the Petrusevich spike for this particular value of M only occurs for L values larger than approximately 5. These observed differences can be explained as follows.

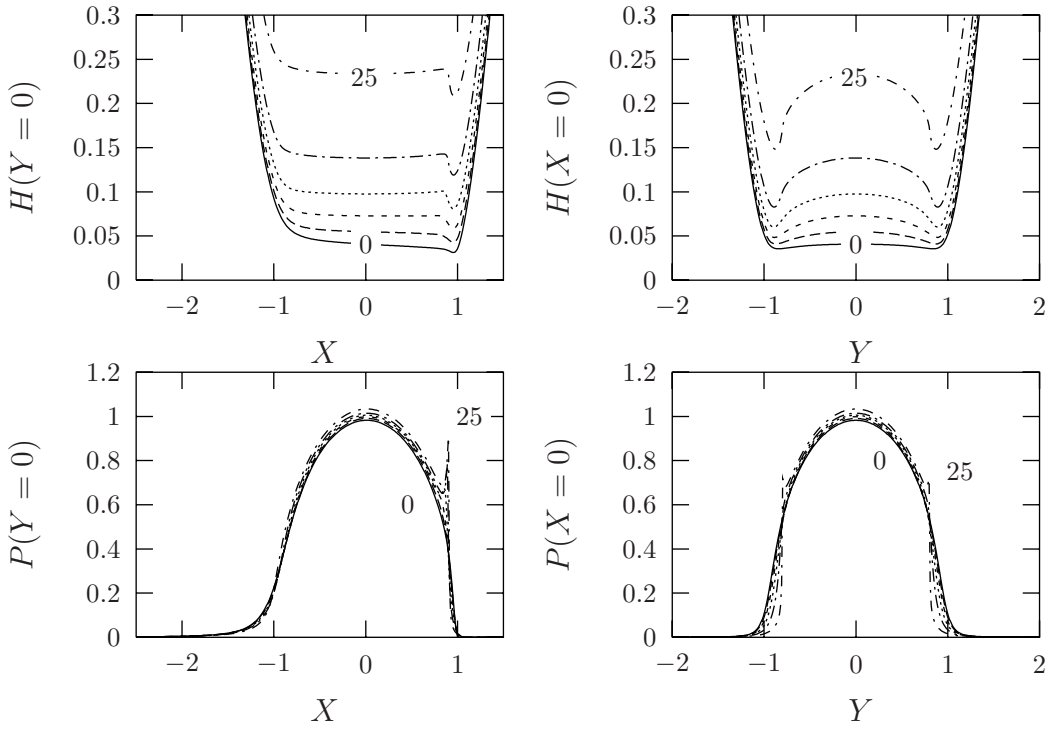


Figure 5.5: Film thickness and pressure along the centre line (left) and perpendicular to the centre line (right) for $M = 100$ and $L = 0$, 1, 2.5, 5, 10, and 25.

For larger values of L, the viscosity increases with pressure at a steeper rate. Since the Poiseuille term decreases with increasing viscosity, it will be less at a certain pressure, if L is larger. Later in this chapter, Poiseuille flow will be shown to be negligibly small in the high pressure zone. Hence, for larger values of L, it will vanish at lower pressures. This means that, with increasing values of L, an increasing amount of lubricant is forced to go through, rather than flowing around the contact, thus increasing the gap between the two surfaces.

In the past, emphasis was placed on the prediction of the minimum and central film thickness in the contact. For example, based on a series of numerical results, similar to those presented above, curve-fit formulae were derived, e.g. see Hamrock and Dowson [29], Venner [66], and Nijenbanning, Venner and Moes [54]. Here, we are interested in a quantity important from the viewpoint of dynamic behaviour, i.e. the stiffness of the contact.

5.1.2 Flexibility and stiffness

The stiffness is defined as $\partial f/\partial \delta_\infty$, where f is the applied load and δ_∞ denotes the mutual approach at equilibrium. To obtain the stiffness from the model in terms of the dimensionless quantities, we thus should study the variation of M with varying Δ_∞, for a given L. However, in view of what will be discussed in the time-dependent case, we will adopt the inverse approach, i.e. given the parameters M and L, we calculate Δ_∞. The relation of Δ_∞ versus M defines the flexibility of the contact.

For example, the calculated dimensionless approach for $M = 100$ and $L = 5$ is given by $\Delta_\infty = 0.897$. This means that the approach is 0.897 times the Hertzian approach and indicates that the rolling element and raceway are pushed slightly apart compared to the positions they would have if no lubricant were present.

Figure 5.6 displays the calculated values of the equilibrium approach Δ_∞ for various values of M and L.

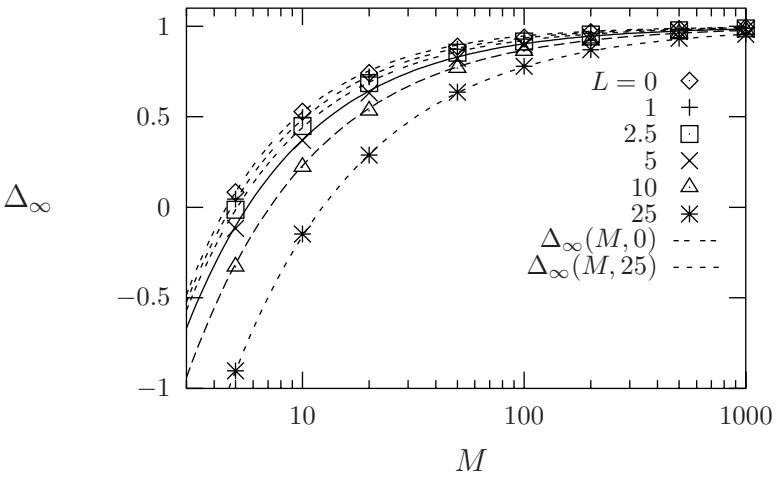


Figure 5.6: Δ_∞ as a function of M for $L = 0$, 1, 2.5, 5, 10 and 25 and the approximated function.

The Figure shows that, for a given L, Δ_∞ increases with increasing M, eventually approaching $\Delta_\infty = 1$. This is the natural limit. After all, with increasing M, the film thickness decreases and becomes more uniform. Thus, the contact approximates the Hertzian dry contact more and more, for which $\Delta_\infty \equiv 1$, see Chapter 2. Hence, for large M, the flexibility of the surfaces determines the flexibility of the contact.

For small M, particularly if L is large, the situation is completely different. For these cases, negative values of Δ_∞ are observed, which implies that

the film thickness at the centre of the contact exceeds the Hertzian deformation. Hence, the flexibility of the contact will be determined much more by the flexibility of the lubricant film than by the flexibility of the elastic bodies.

For engineering purposes, a curve-fit function of the flexibility is constructed that closely approximates the calculated values. Using the fact that, for high values of M, $\Delta_\infty \to 1$, one can assume that Δ_∞ is of the form $1 - pM^q$. Fitting the parameters p and q for different values of L, the following curve-fit function was derived:

$$\begin{aligned} \Delta_\infty(M,L) &= 1 - p(L)M^{q(L)} \quad \text{where} \\ p(L) &= \left((4-0.2L)^7 + (3.5+0.1L)^7\right)^{1/7} \\ q(L) &= -\left(0.6 + 0.6\,(L+3)^{-1/2}\right) \end{aligned} \tag{5.2}$$

The function is included in Figure 5.6.

From the definition of Δ, the actual approach δ_∞ follows from:

$$\delta_\infty = \Delta_\infty(M(f,\ldots),L(\ldots))\left(\frac{9f^2}{8E'^2R}\right)^{1/3}, \tag{5.3}$$

with $\Delta_\infty(M,L)$ according to Equation 5.2. If, on the other hand, the approach δ is given, the load f is obtained implicitly from Equation 5.3. The stiffness of the EHL contact subsequently follows from (numerical) differentiation with respect to δ.

Example

In Figure 5.7, the stiffness of the EHL circular contact, for some typical values of the variables involved, is given and a comparison is made with the stiffness of the Hertzian contact ($\frac{3}{2}k\delta^{1/2}$). A dry contact has no stiffness for negative δ, but the stiffness of the lubricated contact is still significant. For larger values of δ, the stiffness converges to the stiffness of the Hertzian contact.

5.1.3 Numerical accuracy

For most of the computations presented in this section, the size of the computational domain is $X_a = -4.5$, $X_b = 1.5$ and $Y_a = -Y_b = 3.0$. However, larger domains were used for small M to avoid the boundary conditions, specifically the conditions at the inlet, to affect the solution.

The finest mesh, i.e. the mesh at which the solution was evaluated, consisted of 513x513 nodal points. The coarsest mesh employed in the FMG

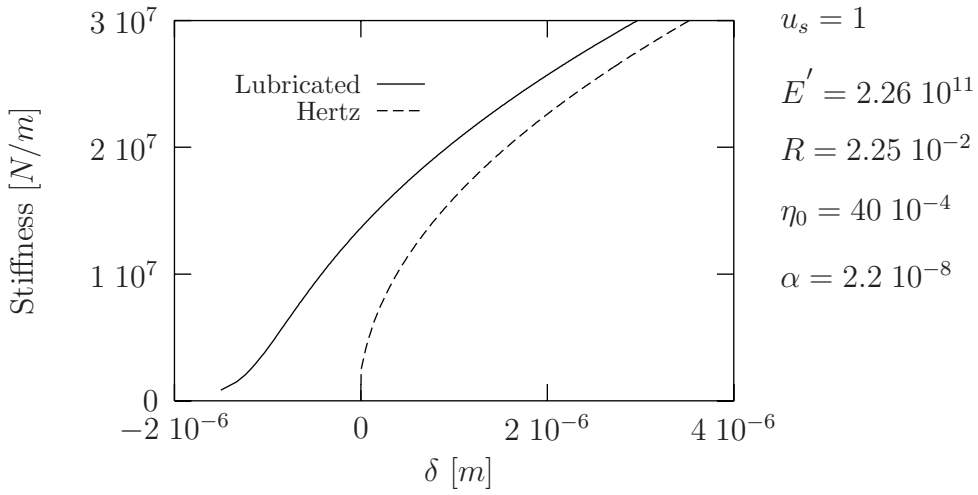


Figure 5.7: Stiffness of the EHL contact compared to the Hertzian stiffness. The variables are in SI units.

algorithm, contained 33x33 nodal points. (For compatibility with the previous chapter, we will denote the mesh consisting of 5x5 points by level 1, the mesh consisting of 9x9 points by level 2, etc. The finest level is thus level 8.)

$3V(2,1)$ cycles and 30 relaxations on the coarsest grid are sufficient for the solution to converge below the discretization error. This can be verified in the first column of Table 5.1, which shows the value of Δ_∞^h obtained on the different grids, for $M = 100$ and $L = 5$. Clearly, with decreasing mesh size, the value rapidly converges to a limiting value. It can thus be inferred that the mesh size is small compared to the wavelength of the Fourier components of the solution.

level #	Δ_∞^h	$\Delta_\infty^h - \Delta_\infty^H$	$\frac{\Delta_\infty^h - \Delta_\infty^H}{\Delta_\infty^H - \Delta_\infty^{2H}}$
4	0.943		
5	0.9086	$-3.44\ 10^{-2}$	
6	0.8996	$-9.00\ 10^{-3}$	0.26
7	0.89726	$-2.34\ 10^{-3}$	0.26
8	0.89668	$-5.80\ 10^{-4}$	0.25

Table 5.1: Convergence test for $M = 100$ and $L = 5$.

The second order convergence with decreasing mesh size in the spatial coordinates is illustrated by the second and third columns in Table 5.1. In the Table, $\Delta_\infty^h - \Delta_\infty^H$ denotes the difference between the value at level l and the

value obtained at level $l-1$. The third column thus shows that this difference decreases roughly by a factor of 4, each time the mesh size is halved. This confirms the second order convergence.

From the value for level 8, it may therefore be concluded that the difference between this value and the value for the continuum problem is below 1%. For general values of M and L, and especially for large L values, the accuracy is of the order of 2%. This is well below the accuracy of the model.

5.2 Time dependent solution

To study damping, clearly, the solution of the time-dependent EHL model is needed. First, by means of numerical experiments using free vibrations, it is shown that the lubricant damps any small oscillation due to an initial deviation from equilibrium. The effects of the vibrations on the film thickness and pressure distribution are discussed.

Next, simulations for sinusoidally varying loads are used to show that the damping is indeed viscous and thus can be modelled as a viscous damper. Damping values are obtained as a function of the various parameters involved. Finally, an estimate of the numerical accuracy of the time-dependent solutions is given.

5.2.1 Free vibrations

As initial conditions, one can either choose $\Delta_0 \neq \Delta_\infty$ and/or take $\dot{\Delta}_0 \neq 0$. As a result of these conditions, the rolling element will start an oscillatory motion that resembles the motion observed earlier for Hertzian contacts, see Section 2.3. However, this time the amplitude of the oscillation will decrease and eventually become zero, as a result of the viscous losses in the lubricant. The pressure and film thickness thus converge to the steady state solution. In this section, the influence of the parameters involved, i.e. the frequency Ω_n, the initial deviation Δ_0 and the parameters M and L, will be studied.

Varying Ω_n

To show the effect of the frequency Ω_n on the mutual approach, the film thickness and the pressure, simulations were performed for $\Omega_n = 2.56$, 5.13 and 10.26. (The motivation for choosing these specific values is given below.) The Moes parameters are $M = 100$ and $L = 5$, whereas Δ_0 was chosen 15% less than the equilibrium approach Δ_∞, i.e., $\Delta_0 = 0.85\Delta_\infty = 0.76$. The initial approaching velocity $\dot{\Delta}_0$ was set to zero.

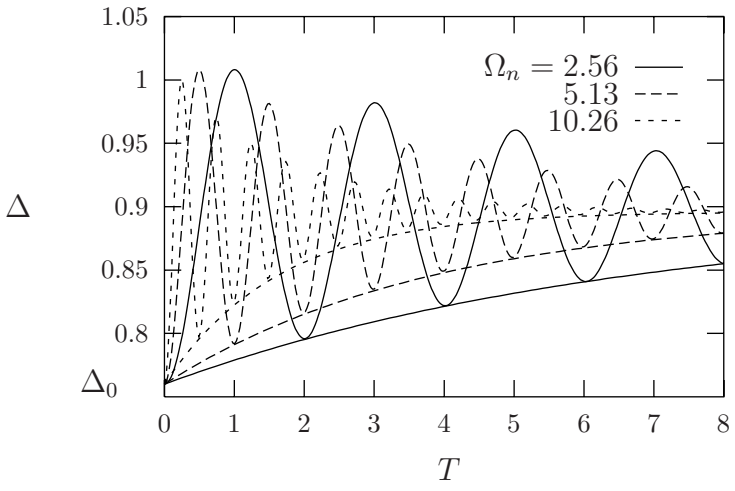


Figure 5.8: Mutual approach Δ versus time T for $M = 100$ and $L = 5$ and envelope, connecting the minima. The frequency is $\Omega_n = 2.56$, 5.13 or 10.26.

Figure 5.8 shows the time history of the mutual approach Δ, from which it can be observed that the amplitude of the oscillatory motion decreases in time. (The included envelope, connecting all minima is discussed below.) After 8 simulated time units, the amplitude for $\Omega_n = 10.26$ has nearly converged to Δ_∞. The Figure shows that the amplitude reduction per oscillation is approximately the same for the three presented solutions, i.e. the value of Δ after one oscillation for $\Omega_n = 2.56$ is approximately the same as the value after one oscillation for $\Omega_n = 5.13$ and $\Omega_n = 10.26$.

The frequency of the oscillation can be explained by the relation between the time scale $\mathcal{T}$, introduced in Chapter 2 for the Hertzian dry contact, and the time scale T, used in the simulation. These time scales are related according to:

$$\mathcal{T} = \Omega_n T. \tag{5.4}$$

It is recalled here that, for Hertzian contacts, the period of oscillation for small perturbations is given by $\mathcal{T}_n \approx 5.13$, see Section 2.3. Thus, the period of oscillation for Hertzian contacts in terms of the time scale T is $T_n = 5.13/\Omega_n$. This means that, provided the influence of the lubricant on the frequency is small (a statement which can be verified later), the frequency in terms of T should be about equal to $f_n = \Omega_n/5.13$. Consequently, for the frequencies of the solutions presented in Figure 5.8 for $\Omega_n = 2.56$, 5.13 and 10.26, the corresponding values of f_n should be 0.5, 1 and 2, respectively. This is indeed the case and thus justifies the assumption that the influence of the lubricant on the frequency can be neglected.

The envelope curves, included in Figure 5.8, are exponential functions. Exponential functions are characteristic for linear viscous dampers (in a linear viscous damper, the viscous force is proportional to the velocity). Hence, one may expect that the viscous forces, or the damping induced by the lubricant, can be accurately described by a linear viscous damper. That this is indeed the case, will further be discussed in the section on sinusoidally varying loads.

The changes in the film thickness are shown in Figures 5.9, 5.11 and 5.13 by means of interference plots. Figures 5.10, 5.12, and 5.14 show the associated pressure and film thickness along the centre line $Y = 0$. It is noted that the time increment between the presented interference plots is enlarged after the first 9 images.

Although the undeformed surfaces are smooth, the interference plots clearly show film thickness variations in the contact region. These variations are crescent and are most clearly observed in Figure 5.13. They have a particular wavelength and propagate through the high pressure zone at the dimensionless speed of unity, i.e. it takes two time units for the film thickness modulation to pass through the contact. The propagation speed is independent of Ω_n. These phenomena can be explained as follows.

Because $\Delta_0 \leq \Delta_\infty$, the rolling element is initially lifted slightly from the raceway. After release, the element starts to accelerate towards the raceway. Since the film thickness is small and the viscosity is large, Poiseuille flow is virtually absent in the high pressure zone and, as a result, the central film thickness hardly decreases, despite the increase in the mutual approach. Instead, the high pressure zone expands to compensate for the increased "load". Thus, as long as the rolling element approaches the raceway, squeeze motion occurs only in the region near the circumference of the Hertzian contact circle and the film thickness only decreases at that location. However, note that as a result of the pressure generated by the squeeze motion, the reduction of the film thickness is smaller than it would be if the load was increased quasi-statically. A larger film thickness thus enters the high pressure zone. As the motion of the rolling element reverses, by a similar mechanism, the film thickness is reduced in the inlet and film thickness modulations are induced. This process repeats periodically, with decreasing amplitude, until it converges to the steady state solution.

As is shown in the Figures, the film thickness changes induced at the inlet are propagated through the high pressure zone. This immediately follows from the Reynolds equation. Since in the high pressure zone the Poiseuille term vanishes, the Reynolds equation reduces to the advection equation:

$$-\frac{\partial(\bar{\rho}H)}{\partial X} - \frac{\partial(\bar{\rho}H)}{\partial T} \approx 0, \tag{5.5}$$

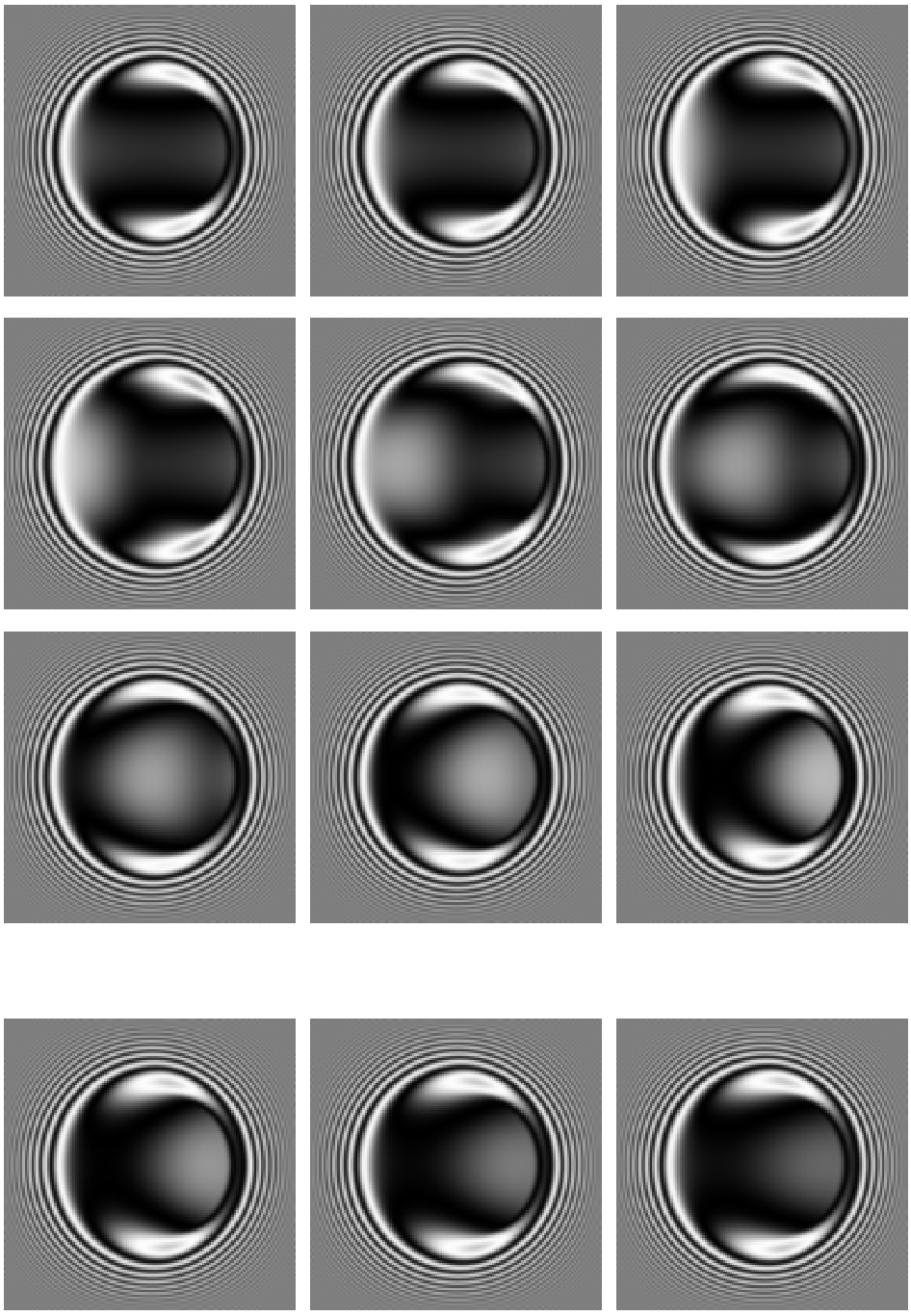


Figure 5.9: Interference plots of film thickness for $M = 100$, $L = 5$ and $\Omega_n = 2.56$ at times $T = 0$, 0.25, 0.5, 0.75, 1.0, 1.25, 1.5, 1.75, 2.0, 4.0, 6.0 and 8.0. The dimensionless wavelength is $\Lambda = 0.06$.

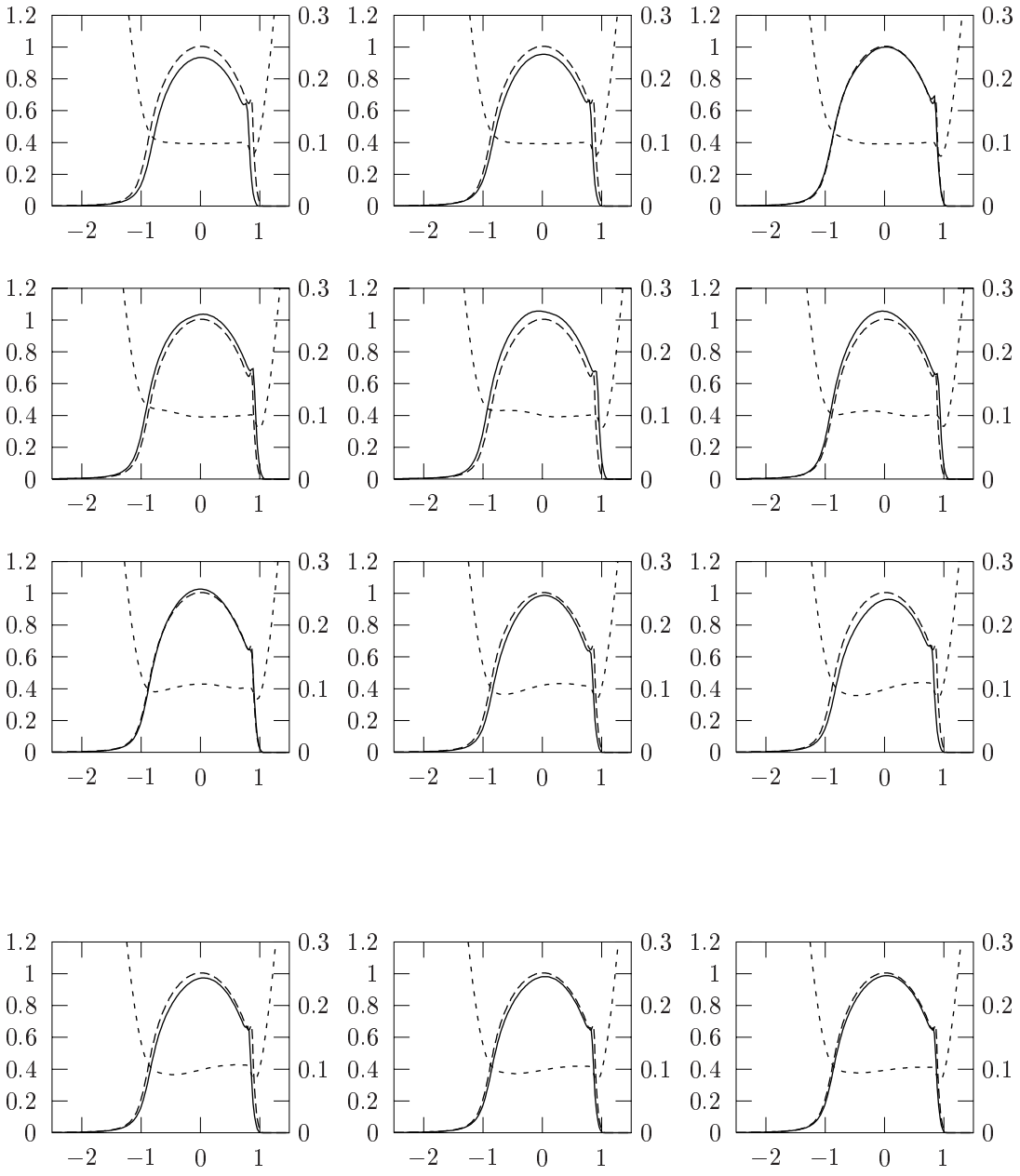


Figure 5.10: Centre line pressure (left labels) and film thickness (right labels) for $M = 100$, $L = 5$ and $\Omega_n = 2.56$ at times $T = 0$, 0.25, 0.5, 0.75, 1.0, 1.25, 1.5, 1.75, 2.0, 4.0, 6.0 and 8.0. The centre line pressure at equilibrium is included for comparison by the long-dashed line.

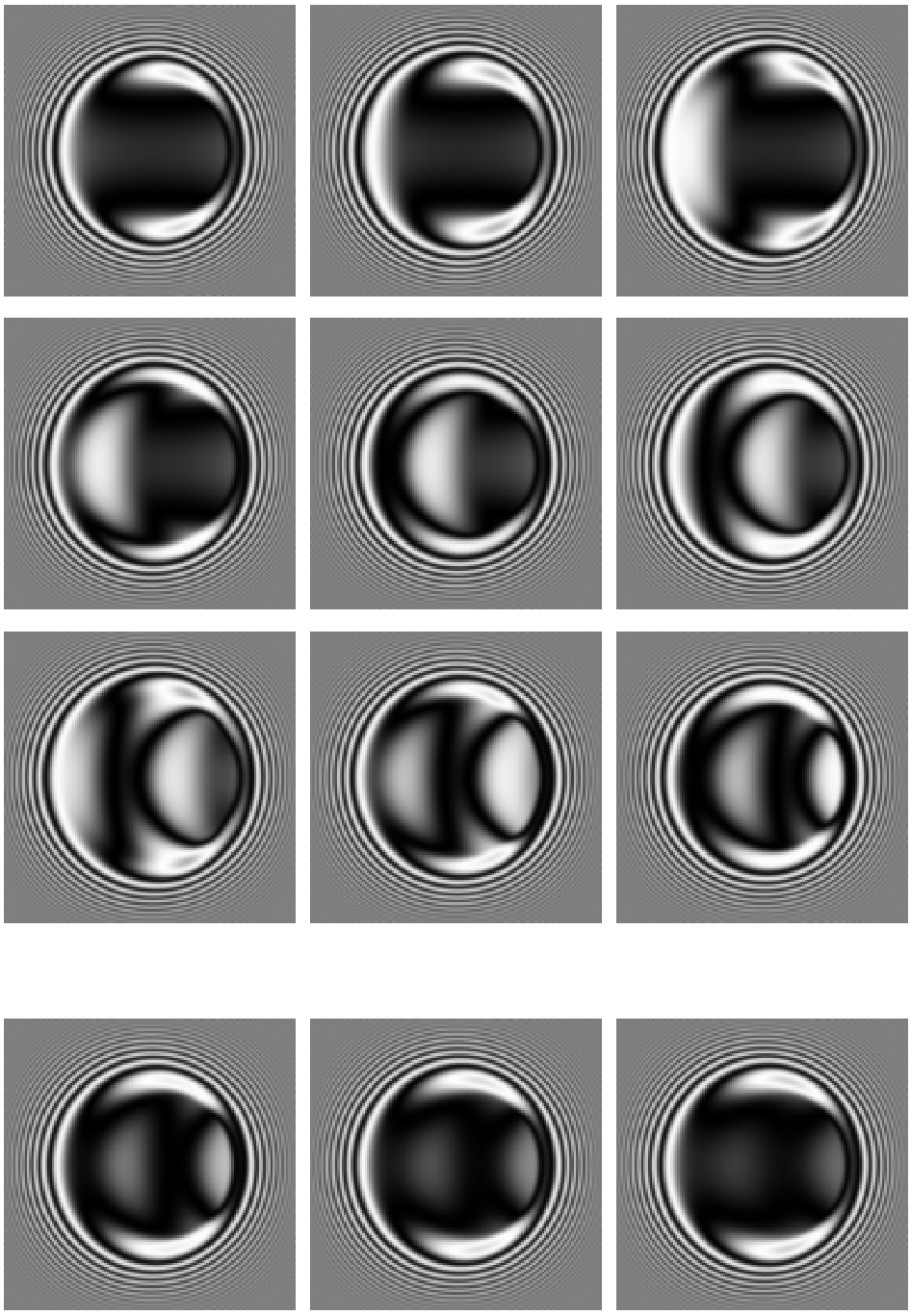


Figure 5.11: Interference plots of film thickness for $M = 100$, $L = 5$ and $\Omega_n = 5.13$ at times $T = 0$, 0.25, 0.5, 0.75, 1.0, 1.25, 1.5, 1.75, 2.0, 4.0, 6.0 and 8.0. The dimensionless wavelength is $\Lambda = 0.06$.

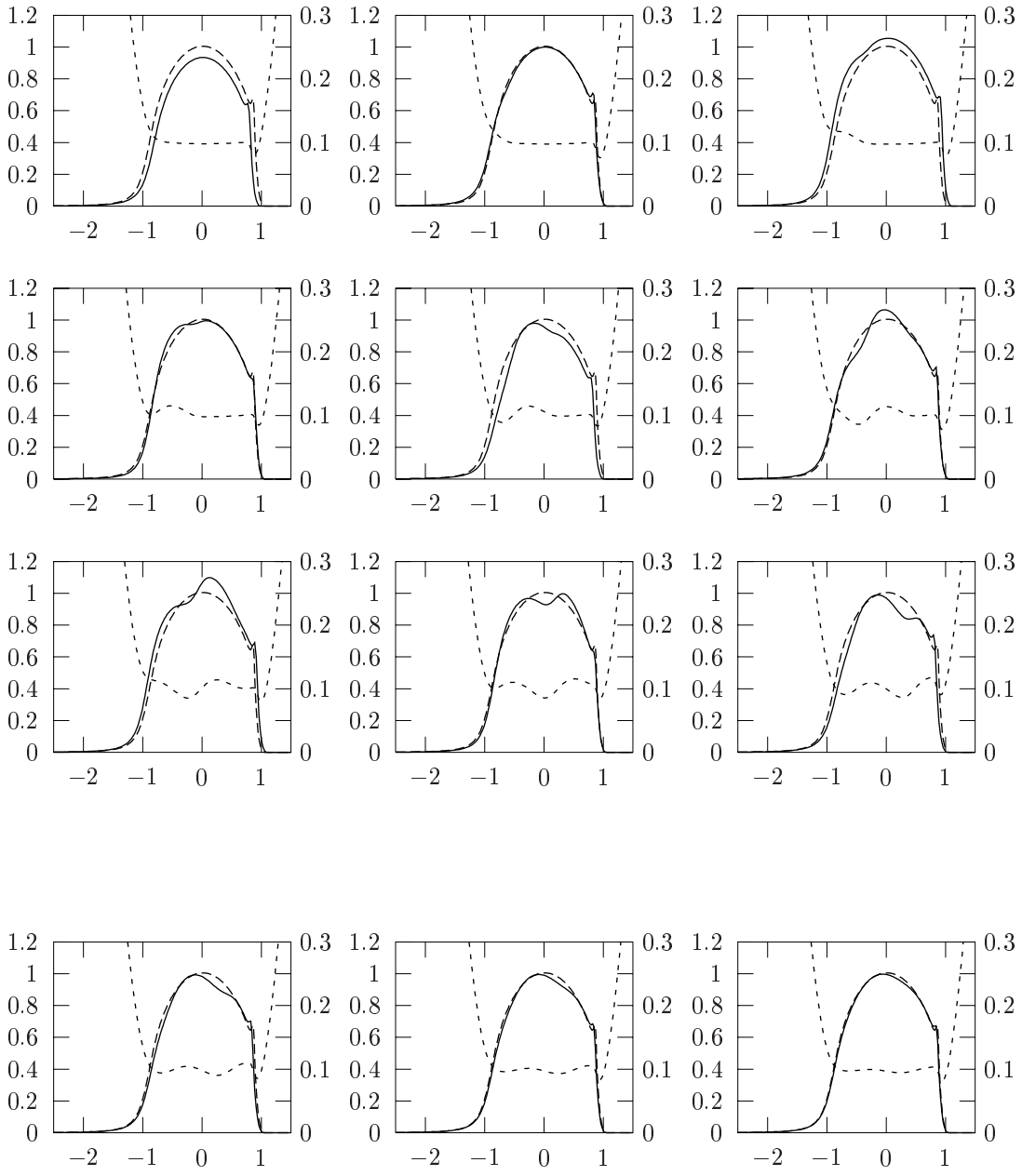


Figure 5.12: Centre line pressure (left labels) and film thickness (right labels) for $M = 100$, $L = 5$ and $\Omega_n = 5.13$ at times $T = 0$, 0.25, 0.5, 0.75, 1.0, 1.25, 1.5, 1.75, 2.0, 4.0, 6.0 and 8.0. The centre line pressure at equilibrium is included as a reference by the long-dashed line.

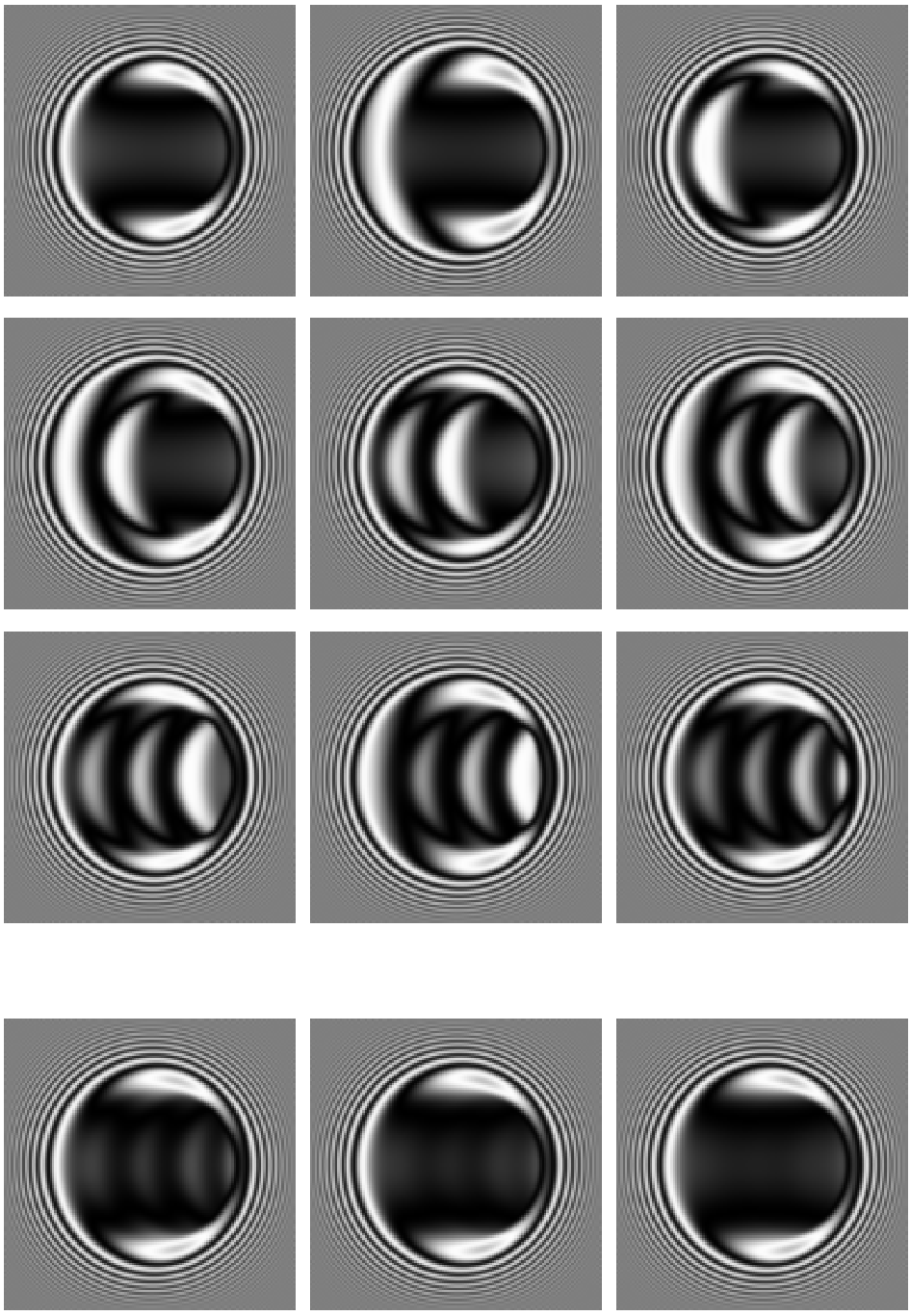


Figure 5.13: Interference plots of film thickness for $M = 100$, $L = 5$ and $\Omega_n = 10.26$ at times $T = 0$, 0.25, 0.5, 0.75, 1.0, 1.25, 1.5, 1.75, 2.0, 4.0, 6.0 and 8.0. The dimensionless wavelength is $\Lambda = 0.06$.

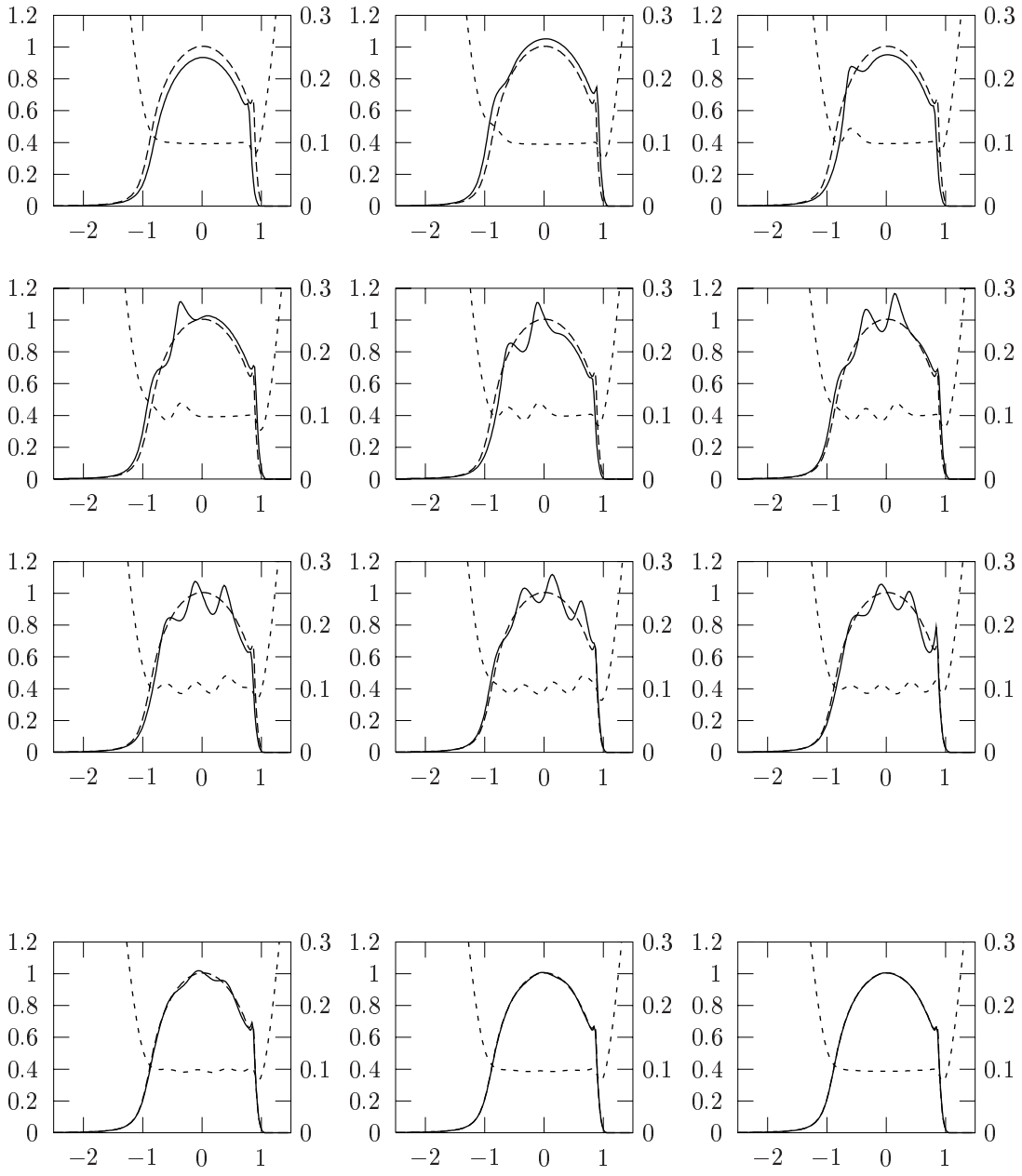


Figure 5.14: Centre line pressure (left labels) and film thickness (right labels) for $M = 100$, $L = 5$ and $\Omega_n = 10.26$ at times $T = 0$, 0.25, 0.5, 0.75, 1.0, 1.25, 1.5, 1.75, 2.0, 4.0, 6.0 and 8.0. The centre line pressure at equilibrium is included as a reference by the long-dashed line.

which prescribes $\bar{\rho}H \approx \bar{\rho}H(X-T)$. Because the lubricant's compressibility is limited, eventually this results in $H \approx H(X-T)$ and any film thickness modulation, induced at the inlet of the high pressure zone, propagates through the contact at the dimensionless speed of unity. In real terms, this is the average velocity of the surfaces, also referred to as the entrainment velocity. As a result, the propagation speed is independent of Ω_n.

Note that the wavelength of the film modulations is larger for smaller values of the frequency Ω_n. This is evident since the frequency with which these film variations are induced is equal to $\Omega_n/5.13$, whereas the propagation speed is fixed. This implies that the wavelength of the film variations equals $5.13/\Omega_n$, which is exactly what is observed in the Figures.

It is interesting to note that the changes in the film thickness increase in the X-direction, most clearly shown in Figure 5.14 for $\Omega_n = 10.26$. This is due to the fact that the amplitude of oscillation at the beginning of the simulation is larger than the amplitudes induced at a later time. Accordingly, the film thickness modulations are thus also larger at the beginning of the simulation. These larger film modulations have thus propagated through the contact further downstream than the film modulations induced at a later time.

The propagation mechanism has been observed earlier, both experimentally and numerically, in studies on surface features like dents, bumps and waviness, see [13, 57, 66, 52, 2]. For instance, in studies on the effect of dents and bumps, some components in the solution must travel at the speed at which the dent or bump moves through the contact. Particularly, in cases where the contacting surfaces move at different speeds (so-called slip), the solution contains the aforementioned component as well as the characteristic component $H \approx H(X-T)$. Note that, since the undeformed surfaces are smooth, the present model is independent of the slip ratio and no such components exist.

The propagation phenomenon is further illustrated in Figure 5.15, which displays the central and minimum film thickness versus time for $M = 100$, $L = 5$ and $\Omega_n = 5.13$. In the initial stage, the central film thickness is approximately constant and only at $T \approx 1$, the time at which the initial variation reaches $X = 0$, is the first change in the film thickness observed. Note again that the increase of the mutual approach, at the beginning of the simulation, leads to an increase in the film thickness. As stated before, this is explained by the pressure, induced by the squeeze motion, which causes the film thickness to be larger than the quasi-static film thickness associated with the increased load.

Also note that the minimum film thickness is immediatly affected by the squeeze motion. Since the minimum film is located in the side lobes, it

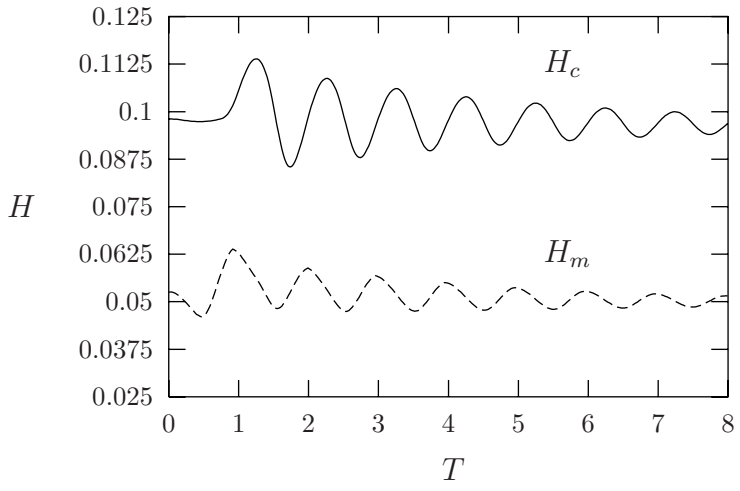


Figure 5.15: Central- and minimum film thickness versus T for $M = 100$, $L = 5$ and $\Omega_n = 5.13$

responds immediately to changes in Δ. As observed in the Figure, the film thickness oscillations gradually decay and would, for larger simulation times, vanish completely.

The pressure distribution on $Y = 0$, is shown in Figures 5.10, 5.12 and 5.14. The Figures show that the contact force, i.e. the integral of the pressure distribution, increases in the beginning of the simulation to compensate for the inertia forces. At $T = 0.5T_n$ the contact force reaches its maximum value, at $T = T_n$ it reaches its minimum value again, after which it starts to oscillate and converges towards the applied load. The induced pressure changes in the inlet are "fixed" to the film thickness changes and propagate through the contact at the dimensionless speed of unity.

It is obvious that, as long as no additional parameters are introduced, one may use different dimensionless variables in which the solution is expressed. Specifically, it is interesting to look at the equations which result when using the time scale $\mathcal{T}$, introduced in Chapter 2 for Hertzian contacts.

The Reynolds equation and the equation of motion in terms of the timescale $\mathcal{T}$ read, respectively:

$$\frac{\partial}{\partial X}\left(\frac{\bar{\rho}H^3}{\bar{\eta}\lambda}\frac{\partial P}{\partial X}\right) + \kappa^2\frac{\partial}{\partial Y}\left(\frac{\bar{\rho}H^3}{\bar{\eta}\lambda}\frac{\partial P}{\partial Y}\right) = \frac{\partial \bar{\rho}H}{\partial X} + \Omega_n\frac{\partial \bar{\rho}H}{\partial \mathcal{T}}, \tag{5.6}$$

and

$$\frac{d^2\Delta}{d\mathcal{T}^2} + \frac{3}{2\pi}\int_S P(X,Y,T)dXdY = 1. \tag{5.7}$$

As a result, for all values of Ω_n the period of oscillation would approximately equal 5.13 (the period of oscillation for Hertzian contacts), since it was shown that lubrication does not affect the period. Neglecting the Poiseuille term in the high pressure zone, the Reynolds equation reduces to:

$$\frac{\partial \bar{\rho} H}{\partial X} + \Omega_n \frac{\partial \bar{\rho} H}{\partial \mathcal{T}} = 0. \tag{5.8}$$

Since the compressibility of the lubricant was shown to be limited, the solution in the high pressure zone, in terms of the "Hertzian" time scale, is thus:

$$H \approx H\left(X - \mathcal{T}/\Omega_n\right). \tag{5.9}$$

Hence, $1/\Omega_n$ would be the dimensionless speed at which film thickness modulations travel through the high pressure zone. For large values of Ω_n, the film thickness modulations would thus propagate very slowly through the contact and the wavelength of the induced modulations is small. The wavelength for small values of Ω_n is very large and a nearly quasi-static solution would be observed.

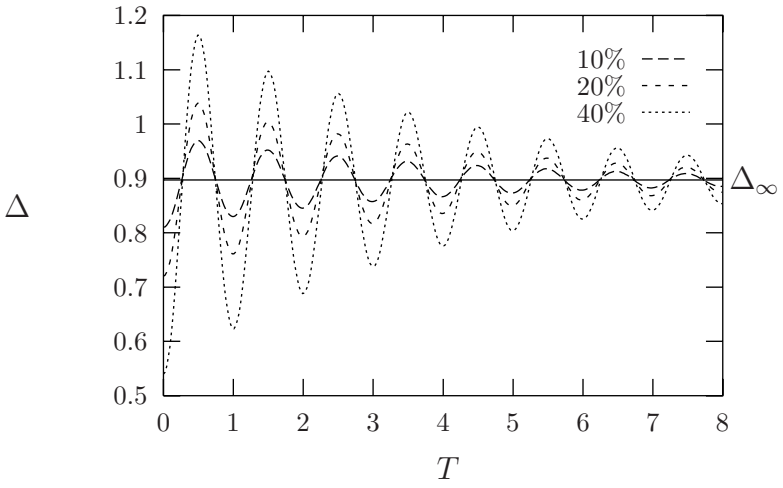


Figure 5.16: Response Δ for a 10%, 20% and 40% deviation from equilibrium, i.e. $\Delta_0 = 0.8$, 0.71 and 0.54, respectively. The line $\Delta = \Delta_\infty$ is shown as a reference.

Varying Δ_0

The effect of the initial approach Δ_0 on the response Δ is shown in Figure 5.16. The parameters used are $M = 100$, $L = 5$ and $\Omega_n = 5.13$. The

initial mutual approach Δ_0 is 10, 20 and 40% smaller than the equilibrium approach Δ_∞.

For Hertzian contacts, see Chapter 2, it was shown that the response due to an initial deviation from equilibrium is asymmetric with respect to Δ_∞, being more apparent for larger deviations from the equilibrium approach.

The asymmetry is also observed for the lubricated case. A close examination of Figure 5.16 shows that, the absolute value of the deviation from the equilibrium approach is larger at a minimum value, than it is at the previous maximum value, whereas for a linear system, the deviation would be less than the deviation at the previous maximum. There is no indication that the effect of the lubricant introduces an additional non-linearity, e.g. through period elongation.

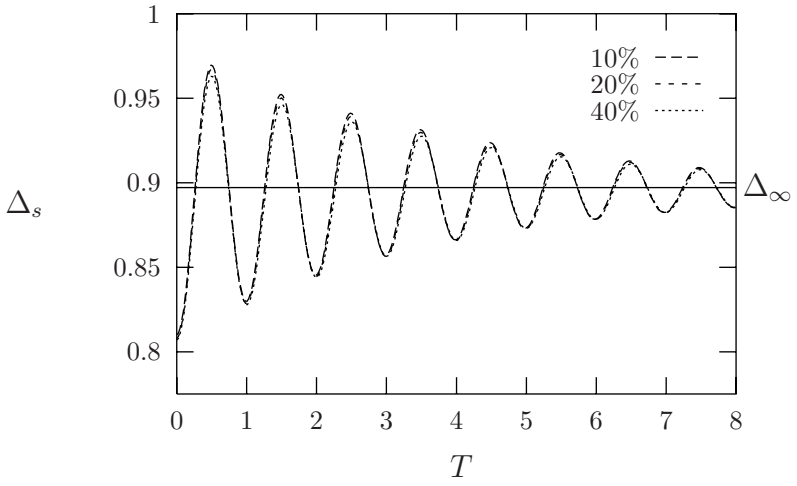


Figure 5.17: Response Δ for a 10% deviation and the scaled response Δ_s onto Δ for a 20% and 40% deviation from equilibrium. The line $\Delta = \Delta_\infty$ is shown as a reference.

To verify that the response is linear in the deviation from equilibrium, Figure 5.17 shows the solutions given in Figure 5.16 scaled onto the 10% solution, i.e. for the 20% solution, $\Delta_s = \Delta_\infty + 0.5(\Delta - \Delta_\infty)$ and for the 40% solution, $\Delta_s = \Delta_\infty + 0.25(\Delta - \Delta_\infty)$. One observes that the difference between the scaled solutions is practically negligible and one may thus conclude that the response is linear in the amplitude of the oscillation for values up to 40% of the equilibrium approach.

Varying M and L

Finally, we discuss the effect of the load parameter M and the lubricant parameter L on the response Δ, the film thickness and the pressure. Specifically, the effect on Δ illustrates the effect of the parameters on the damping of the contact, which will be discussed in Section 5.2.3.

For different values of M, $\Delta(T)$ is shown in Figure 5.18. The remaining parameters are $L = 5$, $\Omega_n = 5.13$ and $\Delta_0 = 0.85\Delta_\infty$. As was shown in Section 5.1.2, Δ_∞ is smaller for smaller values of M. This implies that the actual difference between the initial mutual approach and the mutual approach at equilibrium is larger for larger M. For the present purpose this is, however, irrelevant.

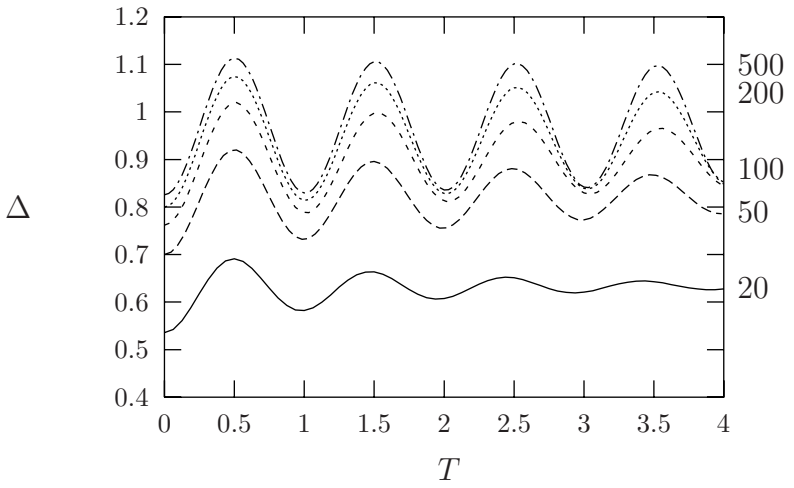


Figure 5.18: Δ versus T for $M = 20$, 50, 100, 200 and 500 and $L = 5$

For all values of M, one observes a similar behaviour of the mutual approach as shown earlier for $M = 100$ and $L = 5$, i.e. Δ oscillates around the equilibrium position and the oscillations are damped for all values of M. However, with the help of the logarithmic decrement of the oscillation, one can show that the amplitude decay is larger for smaller values of M. The viscous losses in the lubricant thus become less with increasing values of M. This observation will be discussed in more detail in Section 5.2.3. In that Section, also a quantification of the damping will be given.

Figure 5.19 shows $\Delta(T)$ for different values of L at $M = 100$ and similar conditions for Ω_n and Δ_0 as before. Although less apparent than that found for higher values of M, one observes that the amplitude decay is smaller for larger values of L. Since the viscosity increases with pressure at a higher rate for larger L, this seems contradictory. However, it is noted that energy

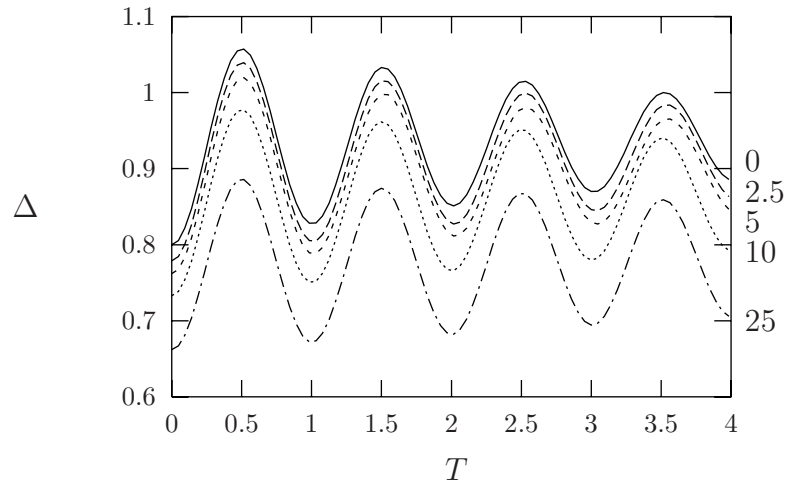


Figure 5.19: Δ versus T for $L = 0$, 2.5, 5 and 10 and $M = 100$

dissipation can only occur in regions where pressure induced flow occurs and, for larger values of L, these regions become smaller. In fact, it is the high viscosity which prevents pressure induced flow and, as a result, no energy is dissipated. As stated earlier, the damping induced by the lubricant will be further discussed in Section 5.2.3.

The film thickness and pressure distribution for the different values of M and L are quite similar to those presented earlier for $M = 100$ and $L = 5$. Both the film thickness and the pressure are perturbations around the steady state solution and converge towards the steady state solution.

Note however the formation of Petrusevich-like spikes near the inlet of the contact, shown in Figure 5.20, which can be observed for larger values of L and Ω_n. The spikes are a result of the squeeze motion in the inlet of the contact region and subsequently move through the contact practically undisturbed. Also note that the constriction of the film thickness, *just* before the spike, resembles the constrictions that can be observed in steady state situations just *after* the spike. As far as pressure generation is concerned, this illustrates that there is no difference between squeeze motion and entraining motion, as also the Reynolds equation states; variations of $\bar{\rho}H$ in the X-direction induce pressure changes similar to variations of $\bar{\rho}H$ in the T-direction.

Petrusevich-like spikes in studies on pure squeeze motion have been observed earlier by a number of researchers, see e.g. Larsson and Höglund [43], Larsson and Lundberg [44], Safa and Gohar [62] and Peiran and Shizhu [58]. Hence, the present observations are a mere combination of results reported

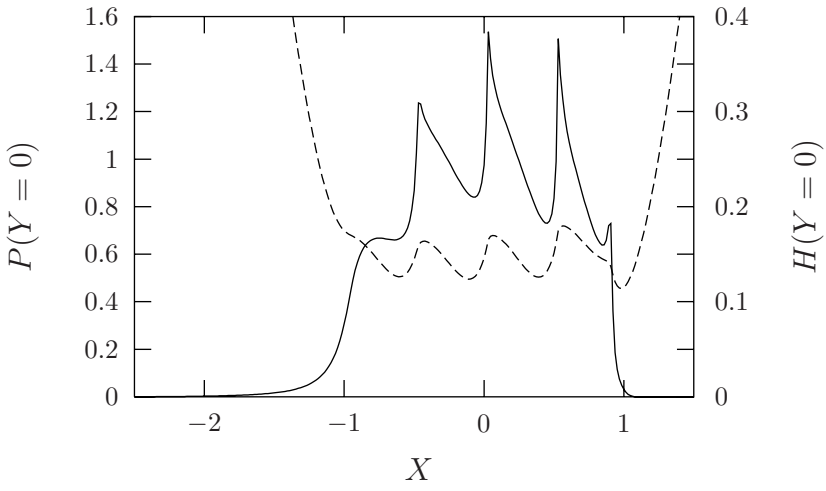


Figure 5.20: Centre line pressure and film thickness at $T = 1.75$ for $M = 100$, $L = 10$ and $\Omega_n = 10.26$. The initial conditions are $\Delta_0 = 0.7$ and $\dot{\Delta}_0 = 0$.

in the studies mentioned and the steady-state solutions discussed earlier.

5.2.2 Sinusoidally varying loads

By means of simulations for free vibrations, we have explained the basic phenomena occurring in the contact. The results showed clearly that the oscillations are damped and that this damping depends on the governing parameters. From the decrease of amplitude as a function of time, this damping could be quantified. A better way to do so, is to determine the damping by means of the simulation for sinusoidally varying loads. Then, the solution is periodic, which makes it easier to deduce damping values.

The simulation for sinusoidally varying loads is quite similar to the one of free vibrations. Contrary to free vibrations no initial conditions need to be specified since there is an instantaneous response to the load. Apart from a transient effect in the initial stages of the simulation, the solution is periodic.

An additional motivation for using sinusoidally varying loads, is to show that film thickness modulations occur in all situations, whenever a variable is varied at a certain frequency. For instance, film thickness modulations, similar to the ones shown earlier, were observed by Kaneta [36] in an experiment in which u_s was varied periodically. The modulations can thus be explained by the mechanisms described here and the proposed wall slip might not be necessary to explain the observations. (The present model is independent of the slip ratio and the difference between the pure rolling and pure slip case in [36] only affects the frequency, which is responsible for the difference in

wavelength of the induced film thickness modulations.)

Response Δ

Figure 5.21 shows the mutual approach Δ obtained for $M = 100$, $L = 5$, $A = 0.1$ and $\Omega_e = \pi$, $\Omega_e = 2\pi$ and $\Omega_e = 4\pi$. One observes a periodic solution for which the frequency is $f_e = \Omega_e/(2\pi)$. Furthermore, the Figure shows that the transient effects, which are associated with the film thickness changes in the time span $0 < T \leq 2$, are not seen in the mutual approach Δ.

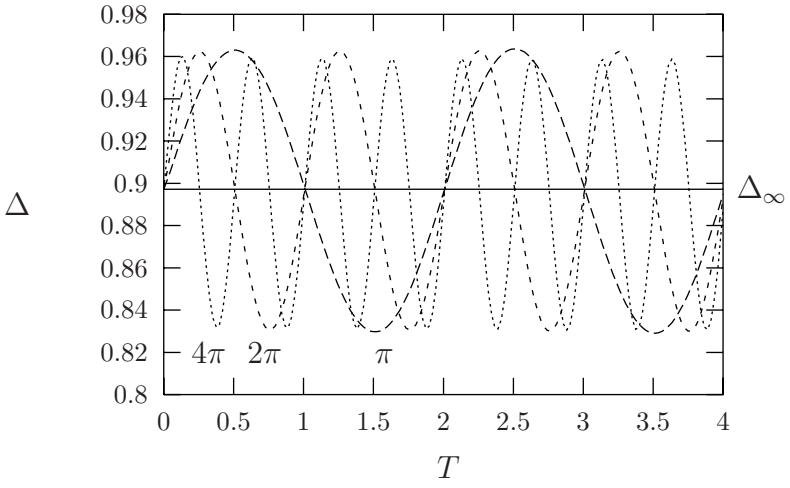


Figure 5.21: Response Δ versus T for $M = 100$, $L = 5$ and $\Omega_e = \pi$, 2π and 4π.

It is interesting to point out that, similar to the free vibrations, the time scales for the Hertzian contacts, $\mathtt{T}$, and the time scale T used in the simulation, are related according to:

$$\mathtt{T} = \Omega_e T. \tag{5.10}$$

Film thickness and pressure

In the simulations for sinusoidally varying loads, similar film thickness and pressure distributions are observed as in the simulation of free vibrations. However, instead of the decay, in this case, the solution becomes periodic. The periodicity starts after the initially induced film thickness changes have propagated through the contact, i.e. for $T \geq 2$.

Since the frequency at which the film changes are induced equals $\Omega_e/(2\pi)$, the wavelength of the induced film thickness variations is now $2\pi/\Omega_e$, see Figure 5.22. The wavelength for sinusoidally varying loads can thus be compared

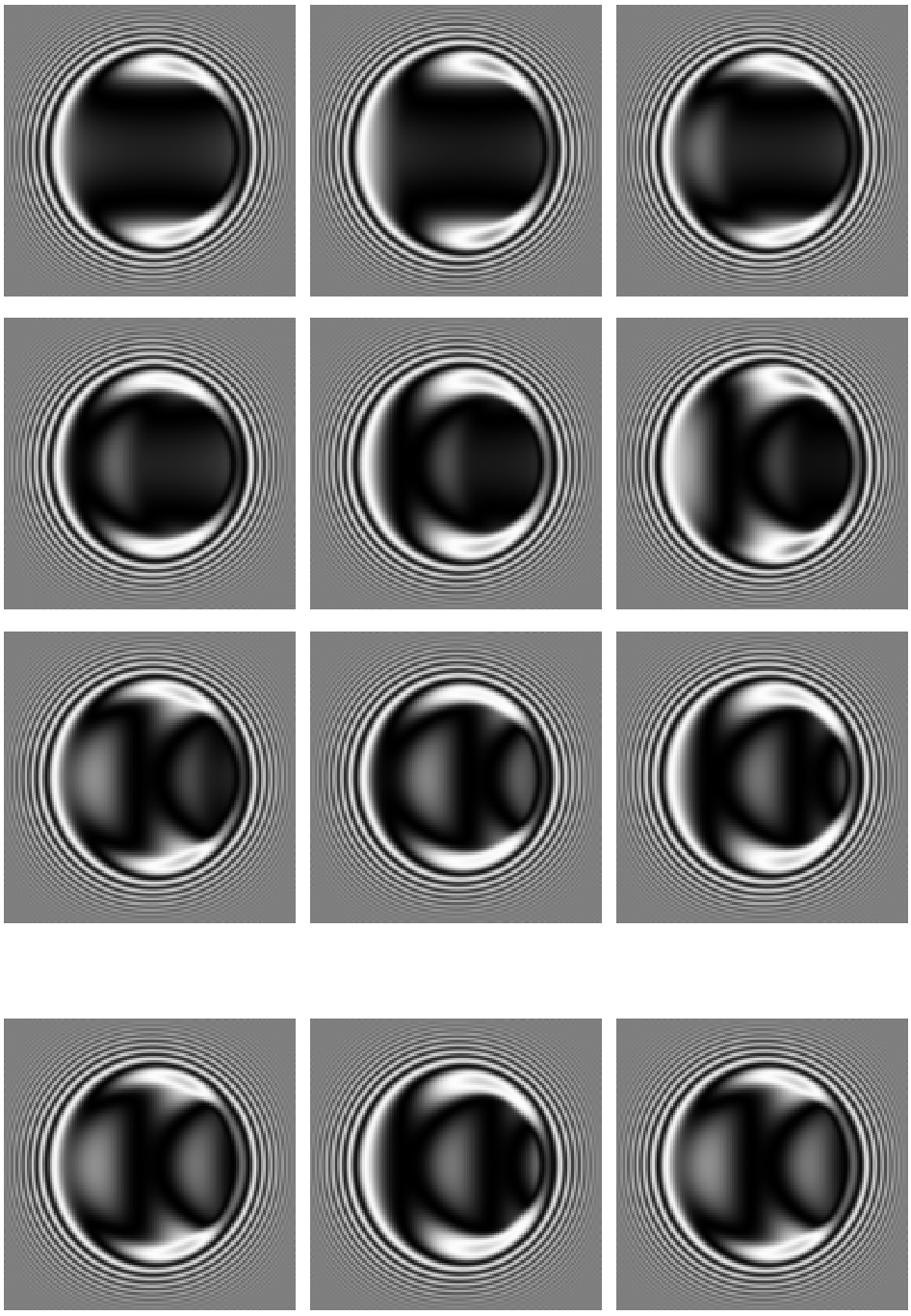


Figure 5.22: Interference plots of film thickness for $\Omega_e = 2\pi$ at times $T = 0$, 0.25, 0.5, 0.75, 1.0, 1.25, 1.5, 1.75, 2.0, 2.5, 3.0 and 3.5. The dimensionless wavelength is $\Lambda = 0.06$.

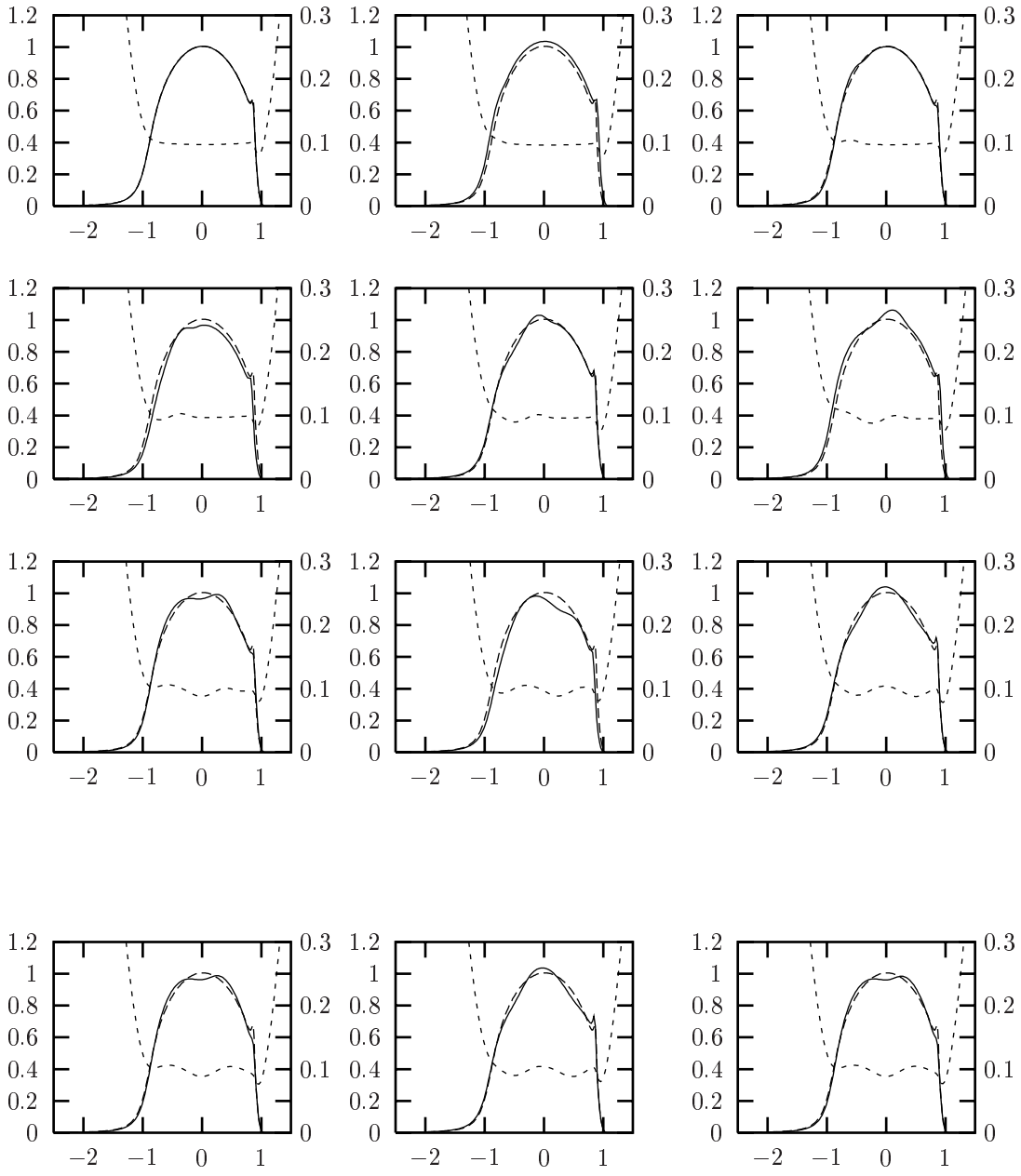


Figure 5.23: Centre line pressure (left labels) and film thickness (right labels) for $M = 100$, $L = 5$ and $\Omega_e = 2\pi$ at times $T = 0$, 0.25, 0.5, 0.75, 1.0, 1.25, 1.5, 1.75, 2.0, 2.5, 3.0 and 3.5. The centre line pressure at equilibrium is included as a reference by the long-dashed line.

to the wavelength in the simulations of free vibrations, if the frequency Ω_e is related to Ω_n according to $\Omega_n \approx (5.13/2\pi)\Omega_e$.

With respect to the amplitude of the observed pressure variations in Figure 5.23, it is noted that, due to the large stiffness of the contact, an amplitude of 10% from the average load has a smaller effect than a 10% variation from the equilibrium approach. This is indeed shown in Figure 5.23 in which snapshots of the centre line pressures and film thickness are given at the same times as those given in Figure 5.22. Hence, the perturbation of the pressure from the steady state solution is smaller than that observed in the case of free vibrations, see Figure 5.11.

5.2.3 Damping

As was shown before, the amplitude decay in the simulation of free vibrations decreases for larger values of M and L. In this section, we will present damping values for a variety of M and L values that support this statement.

The damping values for each value of M and L will be calculated from the so-called hysteresis loop. In a hysteresis loop, the contact force is plotted as a function of the mutual approach Δ. (The dimensionless contact force is defined as $(3/(2\pi)) \iint_S PdXdY$.) As an example, the loop obtained for $A = 0.1$, $M = 100$, $L = 5$ and $\Omega_e = 2\pi$ is shown in Figure 5.24.

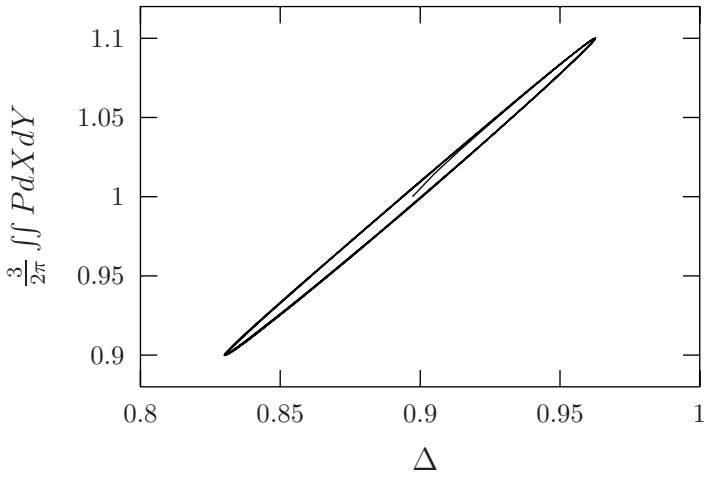


Figure 5.24: Hysteresis loop for $M = 100$, $L = 5$ and $\Omega_e = 2\pi$

Obviously, the area enclosed by the hysteresis loop equals the work per cycle done by the load. It thus equals the energy dissipated by the viscous forces present in the lubricant.

Using a linear viscous damper, the viscous forces are assumed to behave like $d\dot{\delta}$, where d is the damping constant. In terms of the dimensionless variables, the damping constant, denoted by D, is defined as:

$$D = d\frac{cu_s}{2af}. \tag{5.11}$$

For a linear viscous damper, the hysteresis loop is an ellipse, see the frame below. Apart from the initial stages in the simulation, the hysteresis loop shown in Figure 5.24 is indeed an ellipse, which supports the conclusion, given in Section 5.2.1, that modelling the damping of the EHL circular contact by a linear viscous damper is an accurate approximation.

On the shape of the hysteresis loop.

For harmonic excitation, the excitation force provides the energy that is dissipated in the system due to the viscous losses in the lubricant. Assuming a linear spring-damper model, such that the viscous force is equal to $D\dot{\Delta}$, the dimensionless dissipated energy per period equals:

$$E = -\int_0^{T_e} D\dot{\Delta}d\Delta = -\int_0^{2\pi/\Omega_e} D\dot{\Delta}^2 dT. \tag{5.12}$$

Assuming a response $\Delta = B\cos(\Omega_e T)$, the dissipated energy equals:

$$E = -\pi D\Omega_e B^2. \tag{5.13}$$

The total force F for the single degree of freedom is:

$$F = -D\Omega_e B\sin(\Omega_e T) + KB\cos(\Omega_e T), \tag{5.14}$$

where the contact stiffness K is assumed to be linear around the equilibrium position. Then, the force is related to the approach Δ according to:

$$F = K\Delta \pm D\Omega_e\sqrt{B^2 - \Delta^2}. \tag{5.15}$$

Equation 5.15 is the equation of an ellipse which is approximately alligned with the line $F = K\Delta$.

When the hysteresis loop is known by simulation or measurement, one can obtain an equivalent damping constant D for which the area enclosed by the elliptic hysteresis loop equals the area enclosed by the simulated or measured loop.

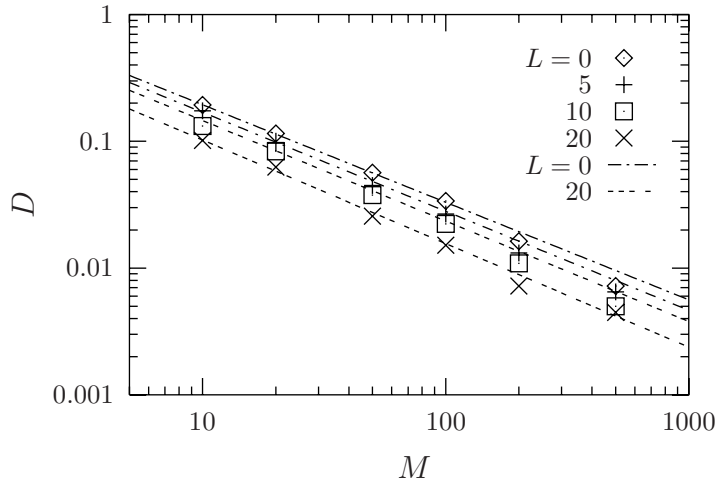


Figure 5.25: The damping constant D and curve-fit function by the drawn lines, as a function of M for different values of L. The frequency is $\Omega_e = 2\pi$.

The damping constant D was determined for a variety of values of M and L at $\Omega_e = 2\pi$ and is shown in Figure 5.25. As was already observed for free vibrations, the damping decreases with increasing M as well as with increasing L.

For engineering practice, a curve-fit function for the damping was obtained. Assuming the form $D = f(L)M^{g(L)}$, the following approximating function appears to be quite accurate:

$$\begin{aligned} D(M,L) &= f(L)M^{g(L)} \text{ where} \\ f(L) &= 1.14 - 0.0234L \\ g(L) &= -0.769 - 0.0024L \end{aligned} \tag{5.16}$$

The function for $L = 0$, 5, 10 and 20, is included in Figure 5.25 by the dashed lines.

5.2.4 Specific friction loss

The damping constant D is a global number. It does not reveal the phenomena in the solution, that account for these losses. This information can however be deduced from the specific friction loss p_s, i.e., the fricition loss per unit area. The dimensionless specific friction loss, P_s, is defined according to:

$$P_s = p_s \frac{4\pi^2 (2R)^3 \eta_0}{9 f^2 \kappa^2} \left(\frac{\mathcal{E}}{\mathcal{K}}\right)^3. \tag{5.17}$$

Expressed in terms of the (earlier) introduced dimensionless parameters, the expression for the friction loss is:

$$P_s = \frac{H^3}{12\bar{\eta}} \left\{ \left(\frac{\partial P}{\partial X} \right)^2 + \kappa^2 \left(\frac{\partial P}{\partial Y} \right)^2 \right\}, \tag{5.18}$$

see Moes [51]. From the specific friction loss, the total dissipated energy per cycle is:

$$E = \int_0^{2\pi/\Omega_e} \iint_S P_s dX dY dT, \tag{5.19}$$

where the term $\iint_S P_s dX dY$ equals the energy per time unit, dissipated in the lubricant.

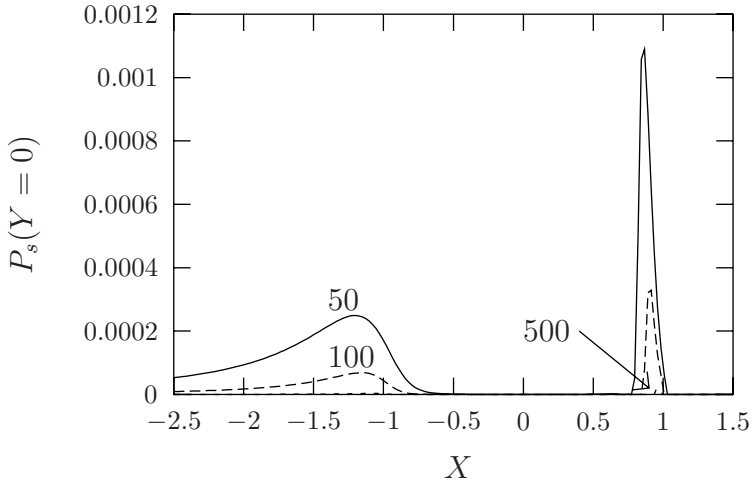


Figure 5.26: Specific friction loss at $Y = 0$ for $L = 5$, $M = 50$, 100 and 500.

Figure 5.26 shows the specific friction loss for steady state conditions at $Y = 0$ for $M = 50$, 100 and 500 at $L = 5$. It can be observed that P_s is largest in the low pressure zone, just in front of the Hertzian contact circle. Also at the location of which the Petrusevich spike, the specific friction loss is significant. However, for all cases shown, P_s completely vanishes in the centre of the contact. The picture further shows that P_s decreases with M. For $M = 500$, it even becomes negligible small. For steady state conditions, the dimensionless energy that is needed (and lost) in order to keep the rolling element rolling, thus decreases with M.

In time-dependent calculations, the specific friction loss is almost identical to the specific friction loss in the steady state situation. (Since it can hardly be distinguished from the steady state distribution, it is not shown.) The

small differences occur at the locations where the friction loss for steady state conditions was already large, i.e. just before the Hertzian circle and near the spike.

The (small) difference between the steady state and the time-dependent distribution equals the additional energy dissipation that causes the oscillation of the rolling element to be damped. Therefore, it can be concluded that the additional energy dissipation is small compared to the energy which is already dissipated in the steady state situation. In addition, this illustrates the decrease of the damping observed for larger values of M.

Varying Ω_e

The influence of the parameter Ω_e on the damping coefficient is shown in Figure 5.27. Again, the parameters are $M = 100$ and $L = 5$. It can be observed that, except for the increase for small Ω_e, the damping is approximately constant, i.e. about the same amount of energy is dissipated per cycle. For small amplitudes, one would expect the damping *constant* to be constant and independent of Ω_e. Only near the transition from the squeeze dominated motion to the entraining dominated motion at $\Omega_e \approx 2\pi$, see Section 5.3, one might observe some deviation from this constant value.

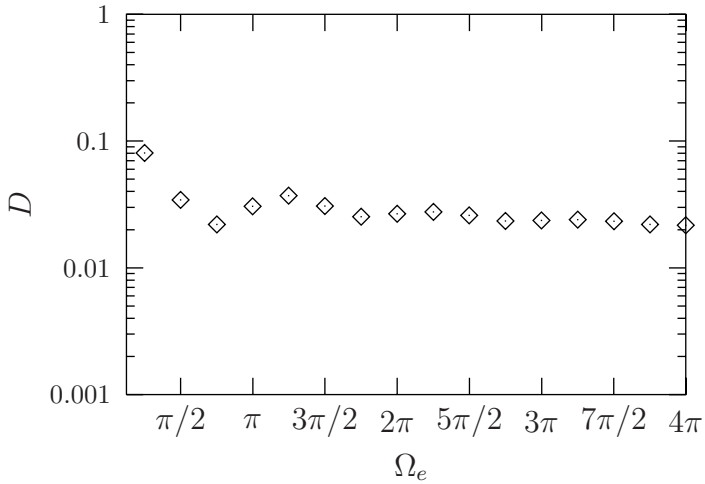


Figure 5.27: Damping constant D as a function of Ω_e for $M = 100$ and $L = 5$

It is interesting to note the inconsistency of approximating the viscous forces in the lubricant by a linear viscous damper for $\Omega_e \to 0$. As stated before, for small values of Ω_e the solution closely resembles the quasi-static solution, given a particular value of the mutual approach Δ. Since for any quasi-static solution the velocity is zero, adding a linear viscous damper having

an arbitrary damping constant, does not affect its solution. It must thus be concluded that the limit for vanishing Ω_e does not exist.

This inconsistency is reflected by the (numerical) problems that arise in determining the solution for small Ω_e. Firstly, since time span that is needed to simulate one hysteresis loop is at least the period of oscillation $2\pi/\Omega_e$, the simulation times are large for small Ω_e and it is very difficult to preserve the accuracy of the solution for these long simulation times. Secondly, and more importantly, the two branches of the hysteresis loop virtually coincide, which makes it impossible to determine damping constants within reasonable accuracy.

On the other hand, for very large values of Ω_e, it should be noted that, although one hysteresis loop is obtained within a short simulation time, the accuracy of the solution becomes smaller because the discretization error in the equation of motion becomes larger. Moreover, one still needs to simulate at least two time units before the film thickness modulations, induced in the initial stages of the simulation, have propagated through the contact and the solution becomes stationary. Reducing the time step is not an option here, because this would lead to the same excessive computation times and associated problems as encountered for small Ω_e. Moreover, because of the reduction of the time step, discretization errors in the discrete Reynolds equation increase, see Appendix B. The problems for small Ω_e can however be solved using a pure squeeze model.

Numerical accuracy

To determine the accuracy of the solutions presented so far, in this section we will present the solutions obtained on different mesh sizes. Contrary to the steady state solution, the presented time-dependent solutions have been obtained using a mesh with 257x257 nodal points. At the expense of a small increase in the discretization error, the coarser mesh has been adopted to reduce simulation times.

So-called F-cycles were used in the simulation, see Venner [66], where $3V(2,1)$ cycles per time step proved to be sufficient to converge below the incremental discretization error. In some cases however, $4V(2,1)$ cycles were needed, specifically for large M combined with large L values. This is due to the under-relaxation which is needed to stablize the numerical process. Underrelaxation slows down the convergence rate per cycle and, as a result, one additional cycle is necessary to converge below discretization error.

To show that the presented solutions indeed converge quickly, Figure 5.28 shows Δ obtained in the simulation of free vibrations, at different mesh sizes. Let $h = h_X = h_Y = h_T = 6/256$ denote the finest mesh size in the simulation,

then Figure 5.28 shows the approach Δ versus time on grids $4h$, $2h$ and h.

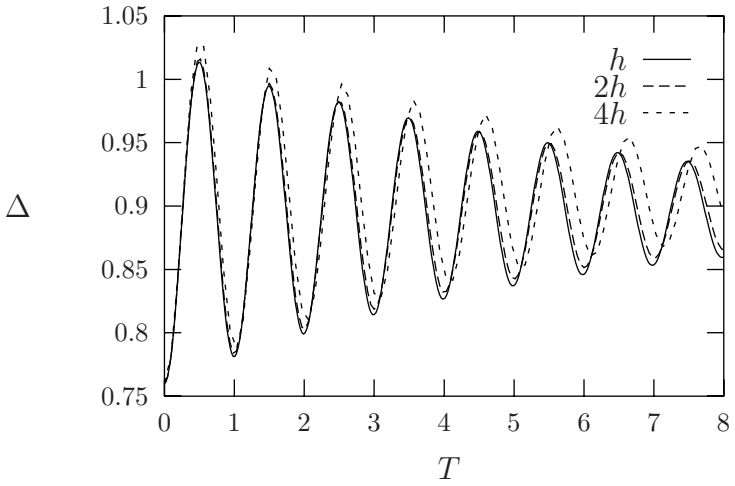


Figure 5.28: Δ versus T obtained on grid h, $2h$ and $4h$.

First, it is observed that, since Δ_∞ is slightly larger for coarser grids (see Table 5.1), the initial deviation from the equilibrium approach, $\Delta_\infty - \Delta_0$, is larger on the coarser grids. Secondly, the solution will converge to an equilibrium approach Δ_∞ which matches the value for that specific grid. For coarser grids it thus converges to a larger value. Both effects are related to the spatial accuracy.

The rate of decline of the amplitude of the oscillation is approximately the same for all solutions. Hence, the incremental discretization error must be small compared to the discretization error in the spatial coordinates. This indicates that the incremental discretization error is $O(h_T^3)$, which is less than the spatial discretization error. Consequently, the accuracy of the solution, presented in Section 5.2.1 for a given time T, is indeed $O(h_T^2)$ and the solution is estimated to be accurate to within 2%.

5.3 Concluding remarks

In this chapter, the stiffness and damping of the EHL circular contact were determined numerically. It was shown that vibrations of the structural elements induce film thickness modulations and associated pressure changes in the contact. In addition, it was shown that free vibrations of the rolling element are damped by the lubricant. Specifically, the obtained curve-fit functions for the stiffness and damping may be used in a structural dynamics analysis or acoustical analysis of rolling element bearings.

For instance, as reported by Wensing [71], finite element models can be used to describe the structural dynamics of the raceways and rollers. The behaviour of the EHL contacts are then introduced by a spring and damper, for which the stiffness and damping are those given in, respectively, Equations 5.2 and 5.16.

It is interesting to note that the frequencies Ω_n or Ω_e bring together two reasearch fields that exist in tribology. For instance, if $\Omega_n \gg 5.13$ or $\Omega_e \gg 2\pi$, one can conclude beforehand that squeeze motion dominates. In that case, a pure squeeze model yields an accurate approximation. Pure squeeze models are adopted in studies of the film thickness between a ball impacting and rebouncing on a lubricated surface. References are given at the end of Section 5.2.1.

On the other hand, if $\Omega_n \ll 5.13$ or $\Omega_e \ll 2\pi$, i.e. when the surface speeds are much larger than the speed at which the bodies approach each other, the steady state model that is used in standard EHL theory is sufficient. Note that this does not imply that the mutual approach may not change in time. Its value may still be governed by the equation of motion, however, the pressure and the film thickness equal the quasi-static solution given a particular value of the mutual approach Δ.

Chapter 6

Elliptical contacts

In this chapter, the solutions for elliptical contacts are presented. They were obtained with the algorithm described in Chapter 4. First, we discuss the effect of ellipticity on the film thickness and the pressure distribution. The stiffness of the elliptical contact, derived from the steady state solution, will be presented. From the simulations with sinusoidally varying loads, the effect of ellipticity on damping is deduced.

6.1 Steady state solution

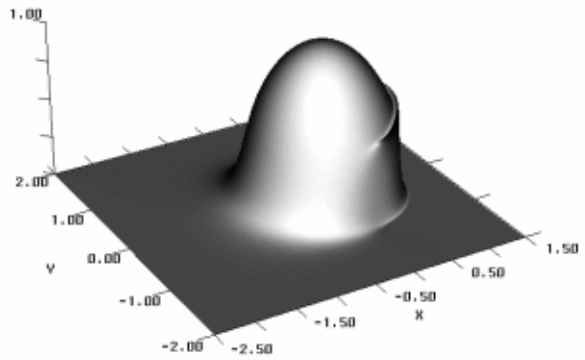


Figure 6.1: Pressure distribution $P(X, Y)$ for $R_x/R_y = 0.05$, $M = 500$ and $L = 5$.

As an illustration, Figures 6.1 and 6.2 show typical examples of the pressure and film thickness distribution of elliptical contacts, in terms of the variables used in the simulation. For this particular case, the ratio of curvatures

$R_x/R_y = 0.05$, i.e. $\kappa \approx 0.14$ and the actual contact width is thus approximately 7 times as large as the contact length. The Moes dimensionless parameters are $M = 500$ and $L = 5$.

The Figures clearly show that the actual elliptical contact region maps onto a circle. The same characteristic phenomena are observed as for circular contacts, i.e. the pressure distribution resembles the Hertzian pressure distribution except for the smooth increase in the inlet and the (start of a) Petrusevich-like spike near the exit. A closer examination reveals that the pressure distribution resembles the pressure distribution for circular contacts, corresponding to a smaller value of M, see Figure 5.1. This will be elaborated upon later.

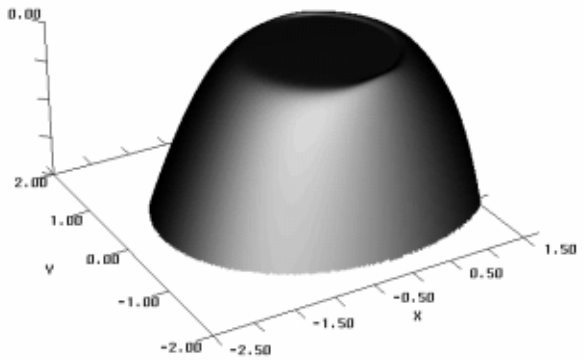


Figure 6.2: Film thickness distribution $H(X, Y)$ for $R_x/R_y = 0.05$, $M = 500$ and $L = 5$.

The film thickness, which is plotted in Figure 6.2, again confirms the mapping of the elliptical contact region onto a circle. In addition, the Figure clearly shows the difference in *dimensionless* radii of curvature in the two directions. Note that the difference in radii, as observed in the Figure, is a result of the applied scaling and is the exact opposite of the difference in the actual radii of curvature. The Figure also shows the constriction near the outlet of the contact.

6.1.1 Varying R_x/R_y

To further illustrate the effect of ellipticity on film thickness and pressure, Figures 6.3 and 6.4 show interference plots of the film thickness, the pressure and the film thickness on $Y = 0$ and $X = 0$, for different values of the curvature ratio R_x/R_y. Again, the values of the Moes dimensionless parameters are $M = 500$ and $L = 5$.

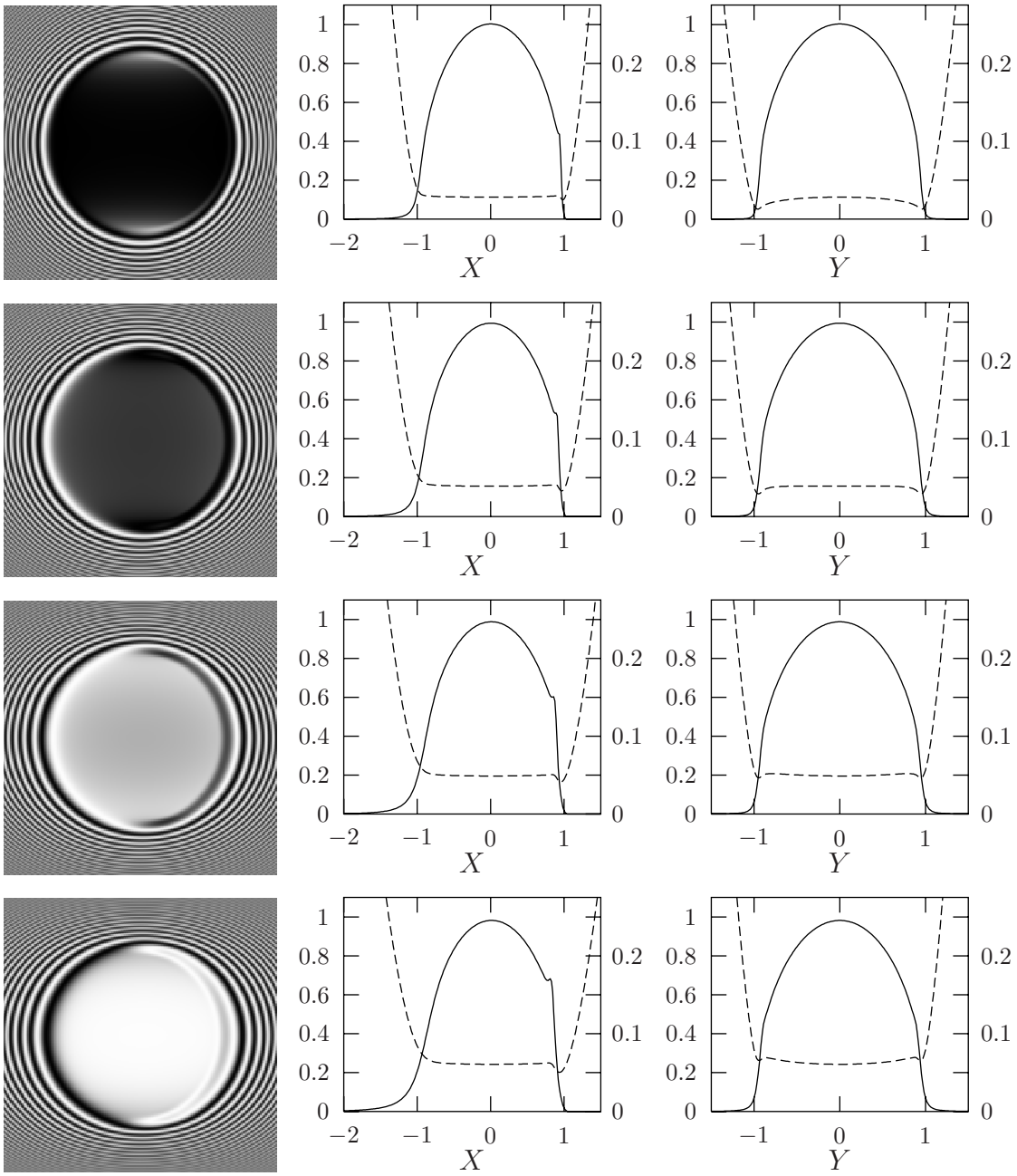


Figure 6.3: Film thickness interference plots, $P(X,0)$, $P(0,Y)$ (solid line, left labels), $H(X,0)$ and $H(0,Y)$ (dashed line, right labels). The Moes parameters are $M = 500$, $L = 5$ and, from top to bottom, $R_x/R_y = 1$, 0.4, 0.2, 0.1 .

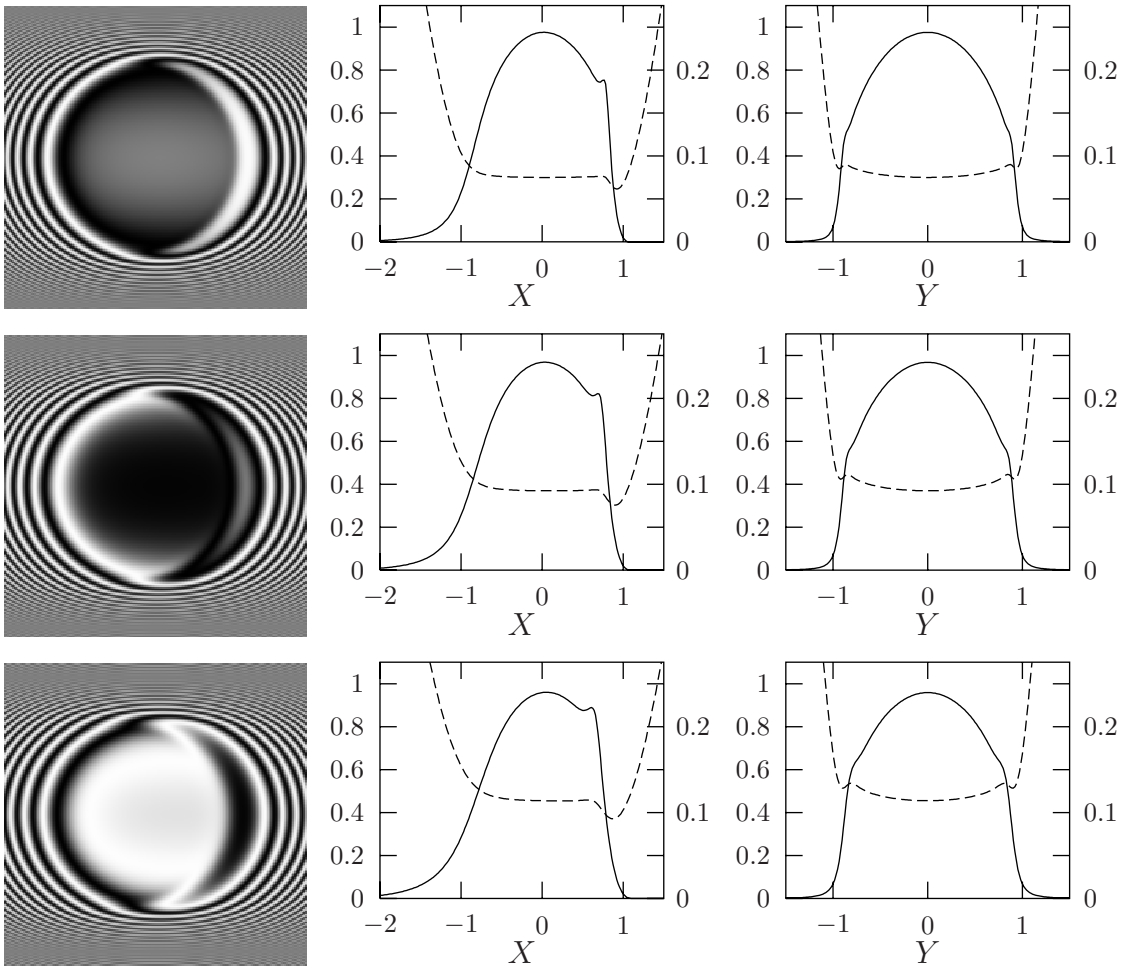


Figure 6.4: Film thickness interference plots, $P(X,0)$, $P(0,Y)$ (solid line, left labels), $H(X,0)$ and $H(0,Y)$ (dashed line, right labels). The Moes parameters are $M = 500$, $L = 5$ and, from top to bottom, $R_x/R_y = 0.05$, 0.025, 0.0125.

As was already described in Nijenbanning [54], one observes clearly that with decreasing values of R_x/R_y and constant values of M and L, the overall film thickness increases. Furthermore, the position of the minimum film thickness shifts from the side lobes to the exit on the centre line of the contact. In addition, as is shown in the Figure, a local minimum in the film thickness forms in the centre of the contact as R_x/R_y decreases.

These phenomena can be explained as follows. By fixing M, the value of the applied load is also fixed. Further analysis shows that if the ratio R_x/R_y decreases, the maximum Hertzian pressure decreases. As a result, the

film thickness increases and the pressure distribution resembles the circular contact distribution at a smaller Hertzian pressure and thus a smaller M value.

The local minimum film thickness, observed in the centre of the contact, is due to Dowson and Higginson's compressibility relation, which is included in the analysis. In this relation, the increase in density is limited to about 30%. Hence, as long as the pressure is sufficiently large, this will also be the density increase in the high pressure zone. Now, as the maximum Hertzian pressure decreases, the pressures in the contact reach a level where the changes in the density vary in the contact. The density increase will be highest in the centre of the contact and the film thickness will decrease to compensate for the density increase. This corresponds with the fact that the Reynolds equation reduces to $\bar{\rho}H = C(Y)$ with $C(Y)$ being constant and dependent on Y only, see also Nijenbanning [54].

6.1.2 Varying M and L

Figure 6.5 and 6.6 show, respectively, pressure contour plots and interference film thickness plots for different values of M and L at $R_x/R_y = 0.05$. The motivation for choosing the particular M values will be explained below.

The pressure distribution in Figure 6.5 closely resembles the pressure distribution for circular contacts. Indeed, as for circular contacts, the pressure distribution converges to the Hertzian pressure distribution for larger values of M, whereas, for smaller values of M, the pressure build-up starts further upstream. The pressure spike can be seen only for larger L values. The effect of the ellipticity only shows in the pressure distribution for smaller values of M; the pressure spikes extend further to the sides of the contact, especially for $L = 10$.

Compared to the circular contact, the film thickness constriction, seen at the outlet, is more pronounced for the elliptical contact and extends further to the sides of the contact. Otherwise, the film thickness distribution resembles the distribution for circular contacts. The reader is referred to Kapitza [40], Kweh [42] and Nijenbanning [54] for additional information about the film thickness and the pressure in elliptical contacts.

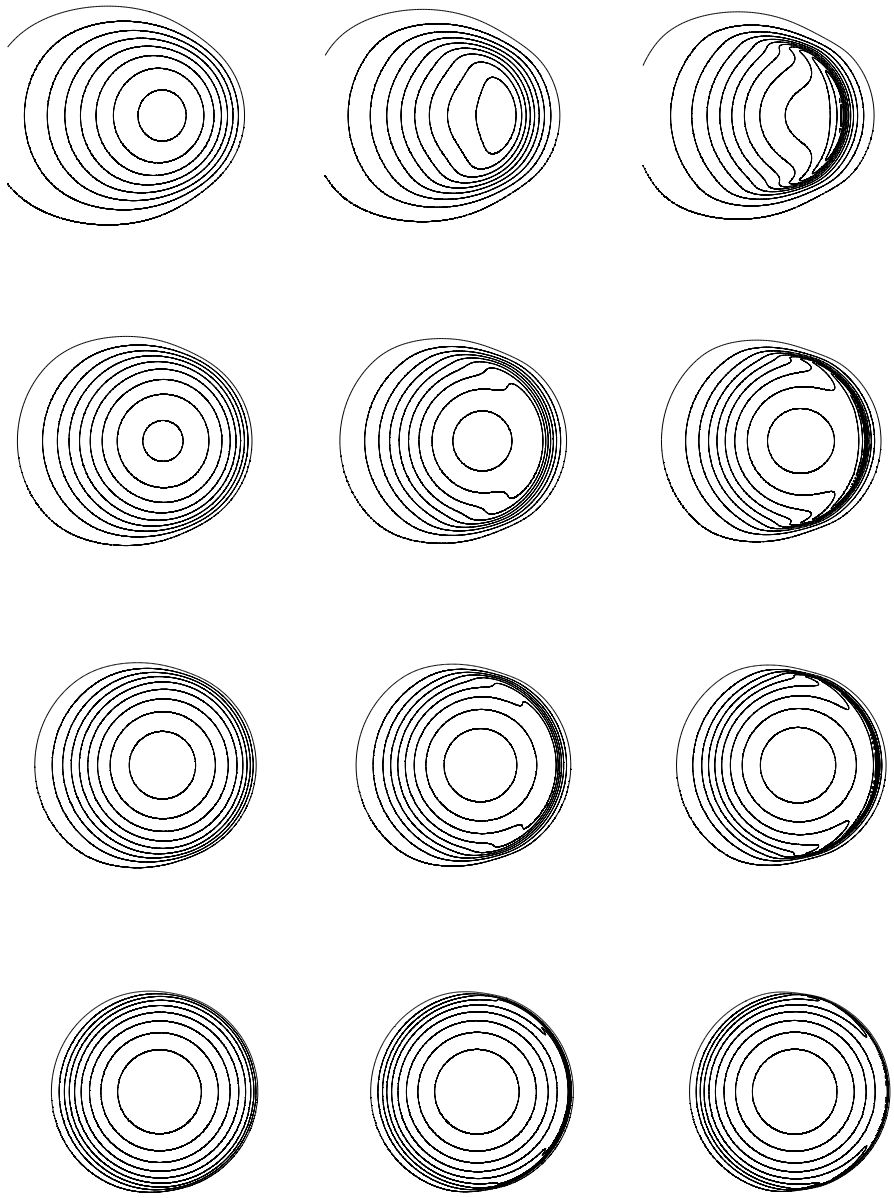


Figure 6.5: Pressure contour plots for $R_x/R_y = 0.05$, M is, from top to bottom, 89.4, 223.6, 447.2, and 2236 and L is, from left to right, 0, 5 and 10. The pressure increment between consecutive lines is 0.1.

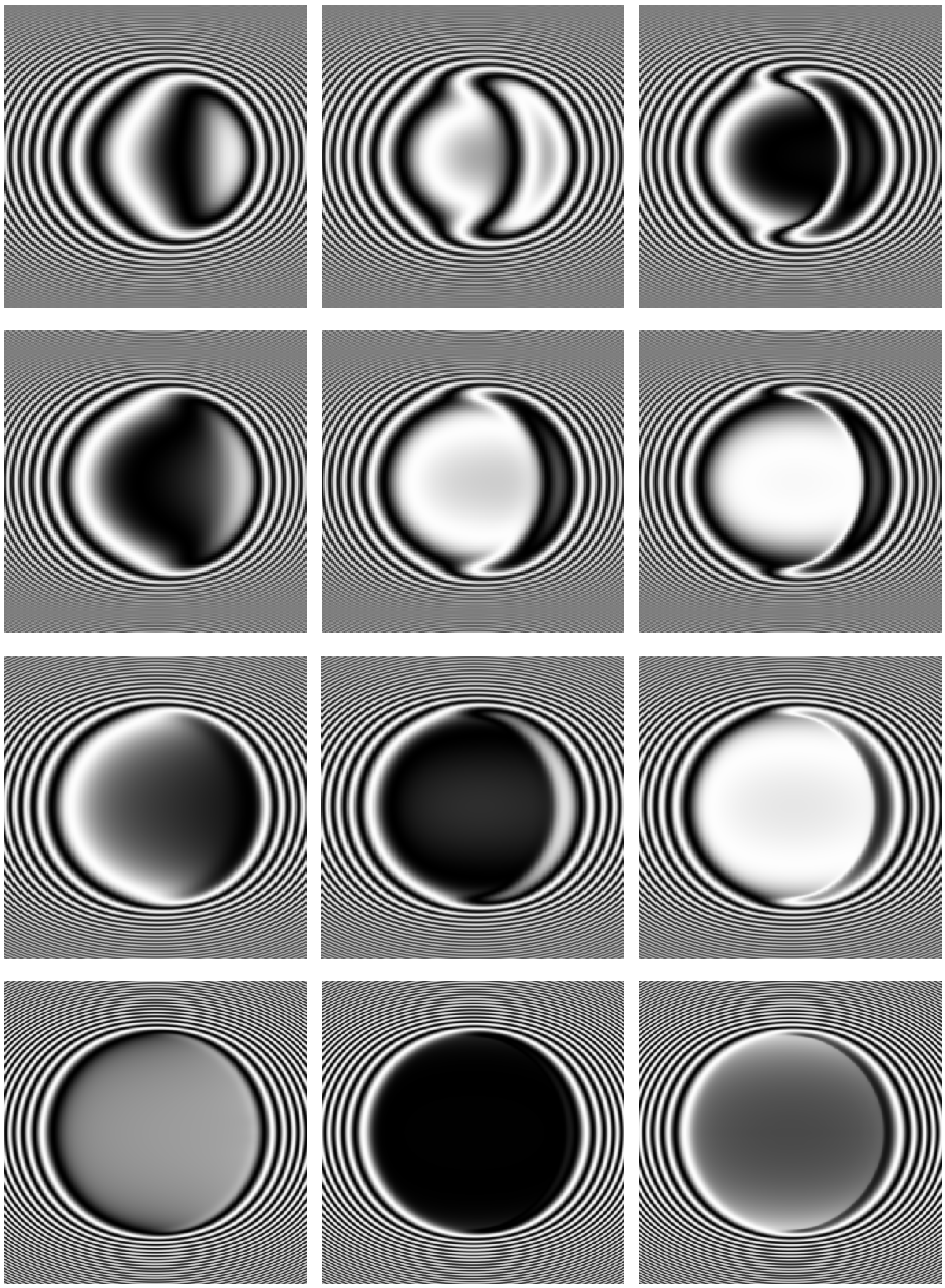


Figure 6.6: Film thickness interference plots for $R_x/R_y = 0.05$, M is, from top to bottom, 89.4, 223.6, 447.2 and 2236 and L is, from left to right, 0, 5 and 10.

6.1.3 Flexibility and stiffness

The effect of the ellipticity ratio on the mutual approach Δ_∞ is shown in Figure 6.7. The curve-fit functions for the circular contact, see Equation 5.2, for two L values are included in the Figure and indicated by Δ^c_∞.

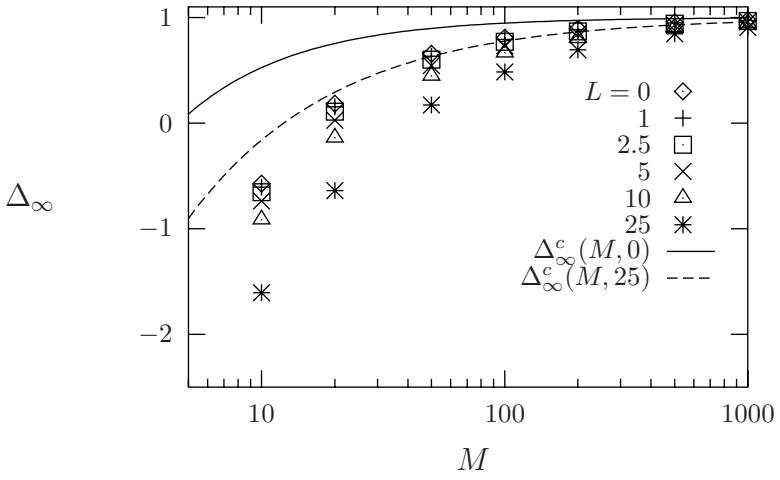


Figure 6.7: Δ_∞ as a function of M for different values of L. $R_x/R_y = 0.1$.

As was observed earlier for the circular contact, for a given L, Δ_∞ increases with increasing M, approaching the Hertzian value $\Delta_\infty = 1$. Hence, for large M, the flexibility of the surfaces determines the flexibility of the contact. For small M, and particularly if L is large, the situation is completely different. In these cases, the values for the elliptical contacts are smaller than those calculated for circular contacts and negative, implying that the film thickness at the centre of the contact exceeds the Hertzian deformation. Clearly, the flexibility of the lubricant film plays a more important role as it did for circular contacts.

From a similarity analysis it follows that the solution in the rigid, isoviscous case only depends on N, defined as $N = (R_x/R_y)^{1/2}M$, and on a factor, depending on R_x/R_y only, see Moes and Bosma [49]. The factor is relatively independent of the ratio R_x/R_y when it is small. Hence, at least for the rigid, isoviscous asymptote, the mutual approach, as a function of N, should be independent of R_x/R_y.

Figures 6.8 and 6.9, showing the mutual approach Δ_∞ as a function of N for $R_x/R_y = 0.2$ and 0.05, confirm that the values of Δ_∞ closely resemble the values of the curve-fit function derived for circular contacts. Also for larger values of L and N, the scaling proves to be quite accurate. As a result, the curve-fit function for circular contacts given in Equation 5.2 can also be

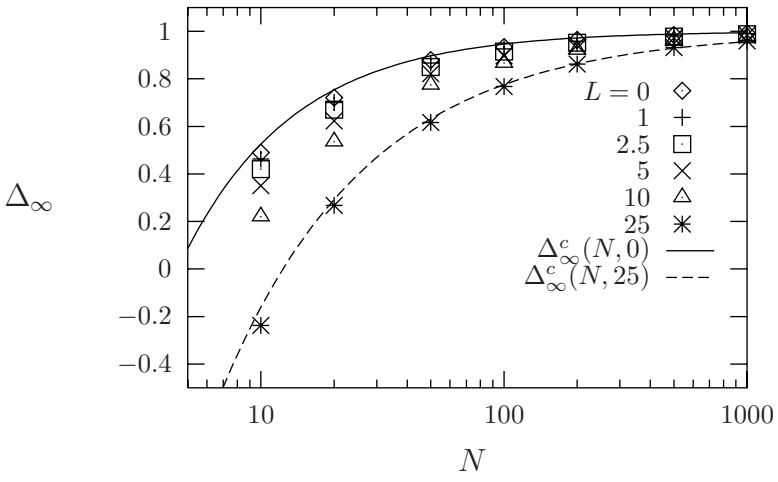


Figure 6.8: Δ_∞ as a function of N for different values of L. $R_x/R_y = 0.2$,

applied for elliptical contacts, provided that M is replaced by N. The mutual approach Δ_∞ may thus be described by:

$$\begin{aligned} \Delta_\infty(N, L) &= 1 - p(L)N^{q(L)} \quad \text{where} \\ p(L) &= \left((4 - 0.2L)^7 + (3.5 + 0.1L)^7\right)^{1/7} \\ q(L) &= -\left(0.6 + 0.6\,(L + 3)^{-1/2}\right) \end{aligned} \tag{6.1}$$

The introduction of N motivated the choice for the particular values of M which were used in Figures 6.5 and 6.6, i.e. for $R_x/R_y = 0.05$; they are equivalent to $N = 20$, 50, 100 and 500.

6.2 Time dependent solution

6.2.1 Film thickness and pressure

For free vibrations as well as for sinusoidally varying loads, the solution, including the mutual approach, the pressure and the film thickness, is similar to the solution for circular contacts. For free vibrations, the rolling element's oscillatory motion, as a result of an initial disturbance, is damped by the lubricant and the solution will eventually converge to the steady state solution. For sinusoidally varying loads, a periodical solution is obtained.

As the rolling element oscillates, it introduces film thickness modulations at the inlet of the contact, which propagate through the contact with the dimensionless speed of unity. Also, for elliptical contacts the wavelength of

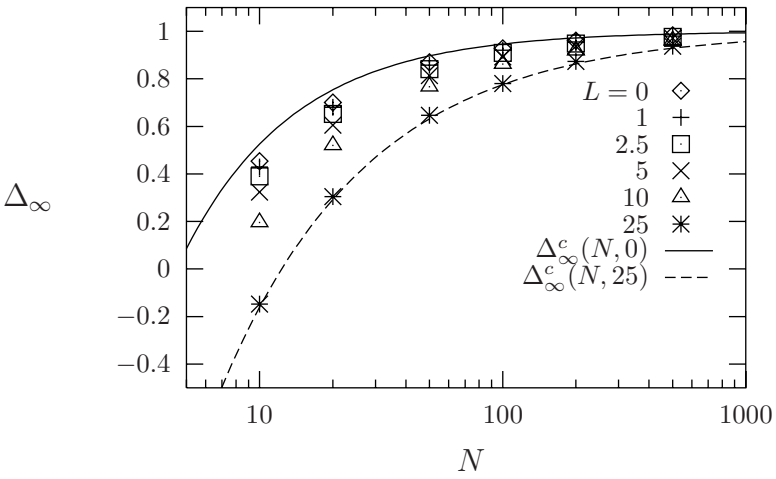


Figure 6.9: Δ_∞ as a function of N for different values of L. $R_x/R_y = 0.05$,

these modulations is approximately equal to $5.13/\Omega_n$ for free vibrations and $2\pi/\Omega_e$ for sinusoidally varying loads.

As an illustration and confirmation of the propagation mechanism and wavelength of the induced film modulations, Figure 6.10 shows interference plots of the film thickness at different instants in time in the simulation for $N = 500$, $L = 20$ and $R_x/R_y = 0.05$. The associated pressure contours are given in Figure 6.11. For clarity, the centre line film thickness and pressure are plotted in Figure 6.12. The amplitude of the oscillation is $A = 0.2$. The excitation frequency is $\Omega_e = 2\pi$ and, indeed, for this particular case, the wavelength of the film modulations is 1.

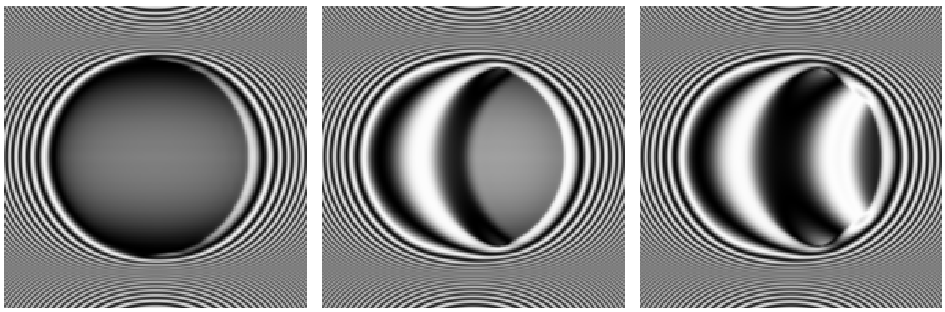


Figure 6.10: Interference plots of the film thickness at $T = 0.0$, 1.0 and 2.0, for $M = 500$, $L = 20$, $A = 0.2$ and $\Omega_e = 2\pi$. $R_x/R_y = 0.05$.

Here, we end the discussion of the film thickness and pressure and focus our attention to the damping induced by elliptical contacts.

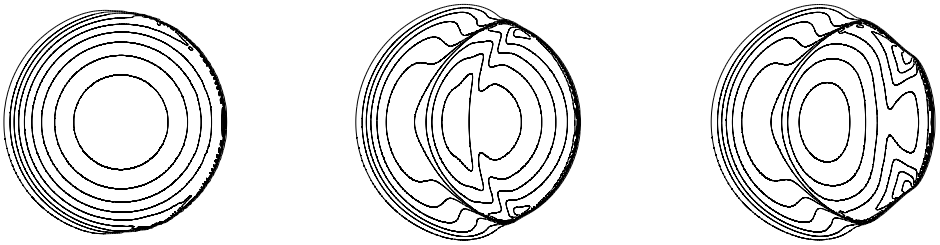


Figure 6.11: Pressure contour plots at $T = 0.0$, 1.0 and 2.0, for $M = 500$, $L = 20$, $A = 0.2$ and $\Omega_e = 2\pi$. $R_x/R_y = 0.05$.

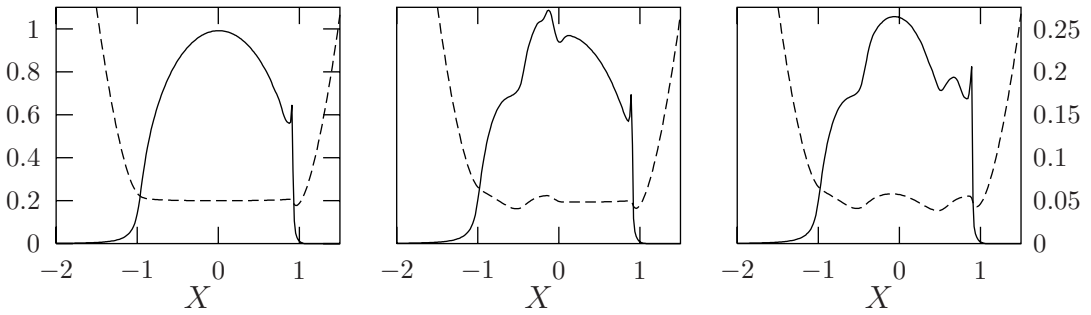


Figure 6.12: Centre line pressure (solid lines, left labels) and film thickness at $T = 0.0$, 1.0 and 2.0, for $M = 500$, $L = 20$, $A = 0.2$ and $\Omega_e = 2\pi$. $R_x/R_y = 0.05$.

6.2.2 Damping

From the hysteresis loop, values for the damping constant D were obtained for a variety of N and L values at $R_x/R_y = 0.05$. The results are collected in Figure 6.13. The curve-fit function of Equation 5.16, as derived for circular contacts, is included in the Figure. This function is denoted by D_c.

As is shown in the plot, despite the use of N, the damping values for the elliptical contact are slightly larger than those obtained for circular contacts. The general trend, however, i.e. less damping for higher values of N and higher values of L, remains the same. Apparently, viscous flow does vanish, but it does not vanish as rapidly towards the sides of the contact, as happens in the equivalent circular contact.

The curve-fit function of Equation 5.16, derived for the circular contact, needs some minor adjustments to fit the values for the elliptical contacts, i.e.

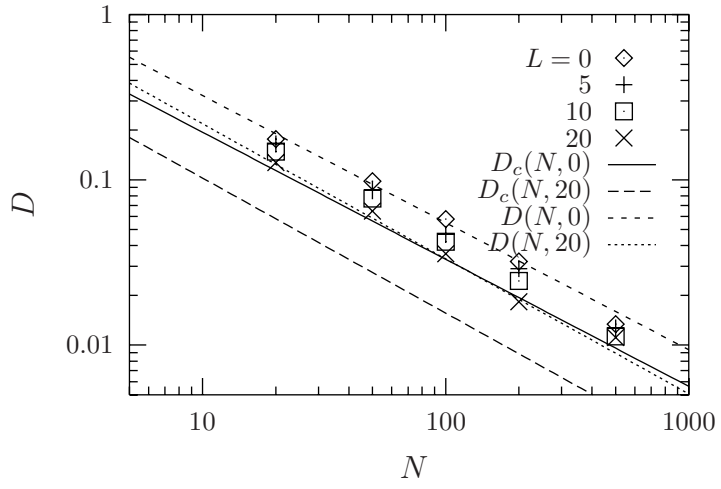


Figure 6.13: Damping constant D versus N for $R_x/R_y = 0.05$ and various values of L.

the number 1.14 in the function $f(L) = 1.14 - 0.0234L$, must increase up to 1.9 for $R_x/R_y = 0.05$. The curve-fit function of the damping constant D for $R_x/R_y = 0.05$, also included in Figure 6.13, thus reads:

$$\begin{aligned} D(N,L) &= f(L)N^{g(L)} \text{ where} \\ f(L) &= 1.9 - 0.0234L \\ g(L) &= -0.769 - 0.0024L \end{aligned} \tag{6.2}$$

6.3 Numerical accuracy

As for circular contacts, most solutions presented in this chapter were calculated on $-4.5 \leq X \leq 1.5$ and $3 \leq Y \leq 3$. Depending on the load number M, this size was adjusted so that the inlet boundary did not influence the solution. The finest mesh used in the steady state calculations contained 513x513 nodal points. The transient solutions were obtained using 257x257 points. The coarsest mesh employed in the FMG algorithm, consisted of 33x33 points. Again, $3V(2,1)$ cycles and 30 relaxations on the coarsest grid are enough, for the steady state solution to converge below the discretization error.

The second order covergence is confirmed in Table 6.1, which shows Δ_∞^h calculated using different mesh sizes h, the difference $\Delta_\infty^h - \Delta_\infty^H$ and the ratio $(\Delta_\infty^h - \Delta_\infty^H)/(\Delta_\infty^H - \Delta_\infty^{2H})$. From the Table, it can be concluded that the value

of the mutual approach is well below 1%.

level #	Δ_∞^h	$\Delta_\infty^h - \Delta_\infty^H$	$\frac{\Delta_\infty^h - \Delta_\infty^H}{\Delta_\infty^H - \Delta_\infty^{2H}}$
4	0.915		
5	0.9069	$-7.92\ 10^{-3}$	
6	0.90495	$-1.98\ 10^{-3}$	0.25
7	0.90450	$-4.50\ 10^{-4}$	0.23
8	0.90438	$-1.22\ 10^{-4}$	0.27

Table 6.1: Convergence test for Δ_∞. $M = 500$, $L = 5$ and $R_x/R_y = 0.05$.

The convergence of the central (H_c) and minimum film thickness (H_m) is confirmed in Table 6.2.

level #	H_c	H_m
4	$6.12\ 10^{-2}$	$5.23\ 10^{-2}$
5	$7.17\ 10^{-2}$	$6.03\ 10^{-2}$
6	$7.43\ 10^{-2}$	$6.15\ 10^{-2}$
7	$7.49\ 10^{-2}$	$6.16\ 10^{-2}$
8	$7.51\ 10^{-2}$	$6.17\ 10^{-2}$

Table 6.2: Convergence test for H_c and H_m. $M = 500$, $L = 5$ and $R_x/R_y = 0.05$.

The discussion of the accuracy of the transient solution is similar to the discussion given in Section 5.2.4. The accuracy is estimated to be below 4% for all cases considered.

6.4 Concluding remarks

The solutions presented in this chapter show that the pressure and film thickness for elliptical contacts can very effectively be solved by applying the scaling based on the Hertzian parameters given in Chapter 3. In addition, it may be concluded that the results, given for the circular contacts in Chapter 5 apply also for elliptical contacts, provided the parameter M is replaced by N.

Chapter 7

Starved lubrication

In the previous chapters, it was assumed that the gap between the the rolling element and the raceway was completely filled by the lubricant. In real applications, however, the amount of lubricant is generally insufficient to realize these fully flooded conditions. As a result, one can distinguish a pressurized region, where the amount of lubricant is sufficient to fully fill the gap, and a starved region, where part of the gap is filled with the surrounding medium or the oil vapour. In the pressurized region, the classical Reynolds equation accurately predicts the flow in the gap. In the starved region, one may assume that the lubricant is transported downstream by shear flow only.

As stated in Chapter 3, the thickness of the inlet lubricant layer, $H_{l;inlet}$, was used to describe starvation. Assuming that the inlet lubricant layer is constant ($H_{l;inlet} = H_{oil}$), the steady state starved lubricated problem is a three parameter problem (M, L and H_{oil}). With sinusoidally varying loads, the solution also depends on the frequency Ω_e and the amplitude A of the varying load. For free vibrations there are three additional parameters, i.e. the frequency Ω_n, the initial mutual approach Δ_0 and the initial approaching velocity $\dot{\Delta}_0$.

7.1 Steady state solution

Figures 7.1, 7.2 and 7.3 show, respectively, the steady state pressure, the fractional film content and the film thickness in the starved lubricated contact, obtained with the algorithm described in Chapter 4. The parameters used are $M = 100$, $L = 5$ and $H_{oil} = 9.78\ 10^{-2}$. This specific value of H_{oil} equals the value of the central film thickness in the fully flooded situation, which will be denoted by H_{cff}. As was shown by Chevalier [14], in this case starvation is significant but not severe or "parched".

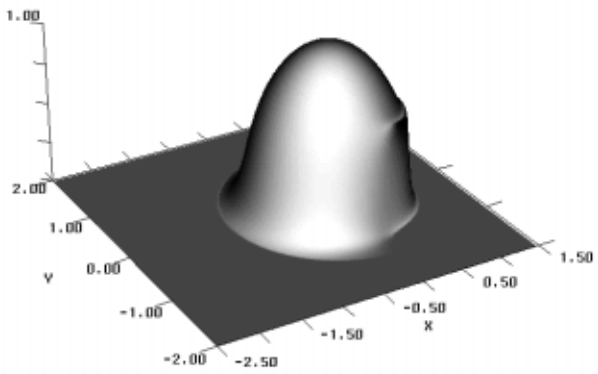


Figure 7.1: The pressure distribution $P(X, Y)$ for $M = 100$, $L = 5$ and $H_{oil} = H_{cff}$.

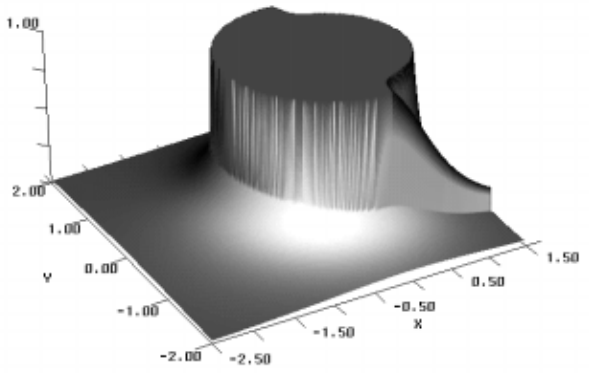


Figure 7.2: The fractional film content distribution $\theta(X, Y)$ for $M = 100$, $L = 5$ and $H_{oil} = H_{cff}$.

As for the fully flooded case, the pressure resembles the Hertzian pressure profile, except for the smooth increase just in front of the Hertzian circle and the Petrusevich-like spike near the outlet. However, under starved conditions, the pressure build-up only starts once the amount of lubricant is sufficient to fill the gap, i.e. when $\theta = 1$. This leads to the free boundary, close to the Hertzian contact circle. The jump in the pressure gradient at the inlet meniscus is also shown in the Figure. This jump will be illustrated more clearly in subsequent Figures.

At this stage, it is interesting to see that the actual shape of the inlet meniscus is a circular arc, concentric to the Hertzian circle. Hence, the

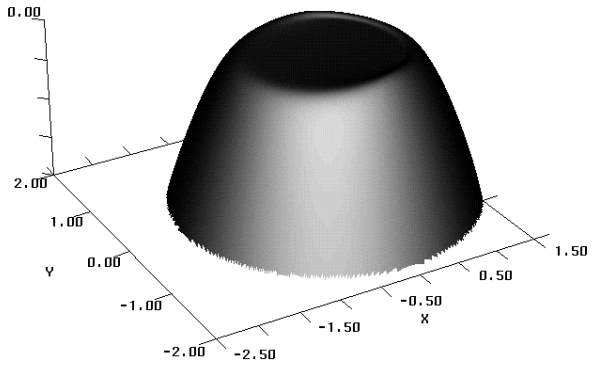


Figure 7.3: Film thickness $H(X, Y)$ for $M = 100$, $L = 5$ and $H_{oil} = H_{cff}$.

assumption that the inlet meniscus is a straight line aligned to the Y-axis, as was assumed in the fully flooded situation, is realistic only if the meniscus is sufficiently far away from the contact.

In the fractional film content distribution, shown in Figure 7.2, one clearly observes a part of the domain, for which the value of the fractional film content is unity, which corresponds to the region of positive pressures in Figure 7.1. This confirms that the solution complies with the complementarity condition. That is to say, the region in which $P > 0$ exactly matches the region in which $\theta = 1$. In addition, as expressed by the JFO relation, the discontinuity in θ clearly shows. (As a reminder, the JFO relation for steady state conditions states that, a discontinuity exists only at the inlet of the pressurized region.) It is observed that at the outlet meniscus, the fractional film content is a smooth function. The latter observation will again be illustrated more clearly in subsequent Figures, see Figure 7.7.

Obviously, since less lubricant is available to the contact, the film thickness values are smaller than in the fully flooded case. In this case, the film thickness distribution shown in Figure 7.3, resembles the fully flooded distribution. However, this is not true in general, as will be discussed later on.

In Figure 7.4 the lubricant layer or lubricant profile $H_l = \theta H$ is shown for the conditions stated before. This may seem a peculiar quantity to depict, but, as will be shown, it is very useful to explain phenomena in starved contacts. That is to say, in the starved region, H_l represents the amount of oil in the gap, as if it were adhered onto a single surface (as may indeed be true in the inlet). In the pressurized region, it is simply the film thickness. Thus, if the lubricant completely adheres to the raceway, one would observe the

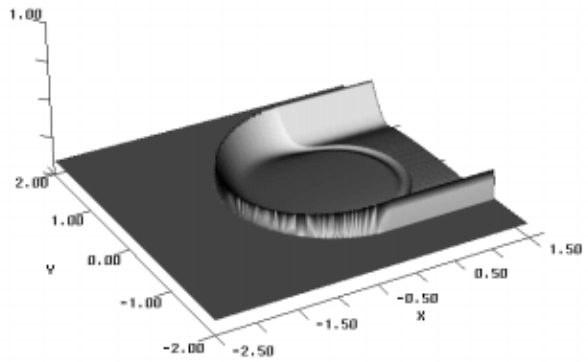


Figure 7.4: Lubricant distribution $H_l(X, Y)$ for $M = 100$, $L = 5$ and $H_{oil} = H_{cff}$.

lubricant film of Figure 7.4, as if the observer is looking through a transparent rolling element. (Note that, as indicated before, the actual distribution of the lubricant across the gap is irrelevant.)

It can be seen that up to the inlet meniscus $H_l = H_{oil}$. This means that, in the starved region, the lubricant is transported downstream undisturbed. In addition, it illustrates the bow-wave at the inlet meniscus; a phenomenon that one can actually observe in experiments. (The bow-wave might be less steep because of the surface tension of the lubricant.) The side leakage and the outlet distribution, which result from the rolling element's passage, are clearly visible; almost all the lubricant is pushed aside, forming the two trailing bands downstream. Apparently, only a fraction of the available lubricant is able to enter the narrow gap between the surfaces in the high pressure zone. Also note that the two bands separate the wake from that part of the lubricant, which remains unaffected by the passage of the rolling element. As found in the fully flooded situation, the Figure shows the film thickness constriction near the outlet of the contact.

7.1.1 Varying H_{oil}

To illustrate how the thickness of the inlet lubricant layer, H_{oil}, influences the solution, Figures 7.5 and 7.6 show, respectively, the pressure and the film thickness along the centre line $Y = 0$, for values of H_{oil} ranging from 4 times H_{cff} to 0.25 times H_{cff}. The lubricant layer, H_l, is included in Figure 7.6. The associated fractional film content θ is shown in Figure 7.7.

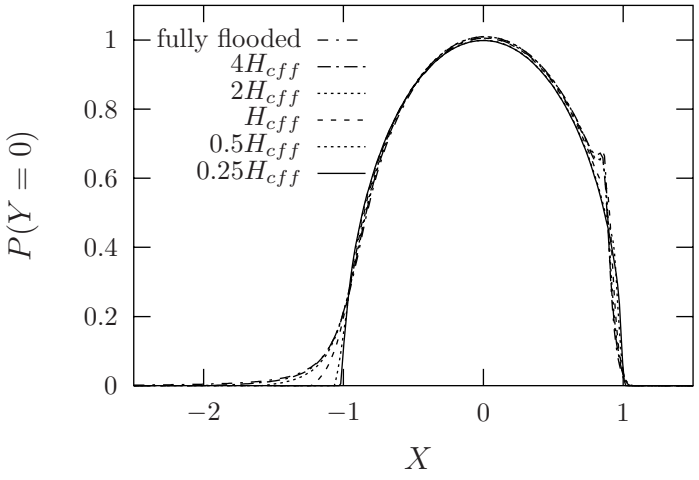


Figure 7.5: Pressure $P(X, Y = 0)$ for fully flooded conditions and various values of H_{oil}. The Moes parameters are $M = 100$ and $L = 5$.

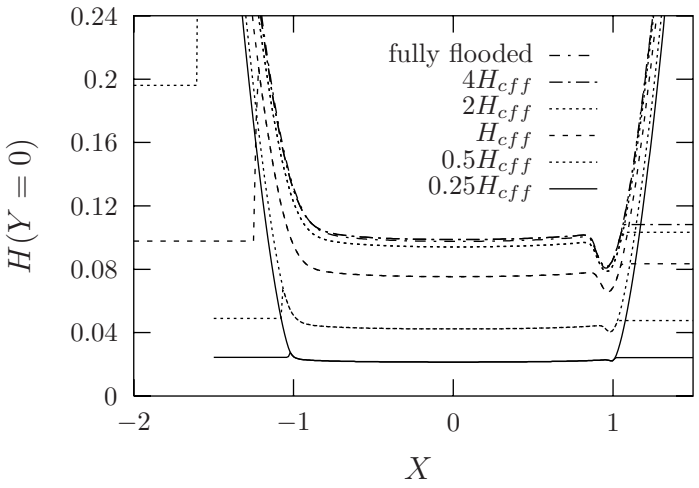


Figure 7.6: Film thickness $H(X, Y = 0)$ and lubricant profile $H_l(X, Y = 0)$ for fully flooded conditions and various values of H_{oil}. The Moes parameters are $M = 100$ and $L = 5$.

In starved lubricated contacts, two asymptotes can be distinguished. As could be expected, the solution converges to the fully flooded solution for increasing values of H_{oil}, whereas for vanishing values of H_{oil}, the solution converges to the dry contact solution.

It is interesting to see that, if H_{oil} is only 4 times H_{cff}, the difference between the solutions for starved lubrication and the fully flooded case already becomes very small. Apparently, only a very limited amount of lubricant is required to obtain fully flooded conditions. Actually, this is one of the main reasons why rolling element bearings operate as reliably as they do.

For vanishing values of H_{oil}, the pressure converges to the Hertzian pressure, which can be verified in Figure 7.5. The associated film thickness would therefore converge to zero. However, before the point of vanishing film thicknesses, as can be observed in Figure 7.6, the film thickness converges to H_{oil}. Apparentely, if practically no lubricant is available, all of it is used to separate the surfaces. This conclusion is also supported by the vanishing difference between the inlet oil film H_{oil}, and the value of H_l at the outlet of the contact. This confirms the statement made by Chevalier [14], that starved lubrication is indeed very efficient.

With respect to the film thickness distribution, it should be noted that the constriction near the outlet almost disappears, making the film thickness more flat. The slight curvature that remains must be ascribed to the increased density at higher pressures, included in the analysis.

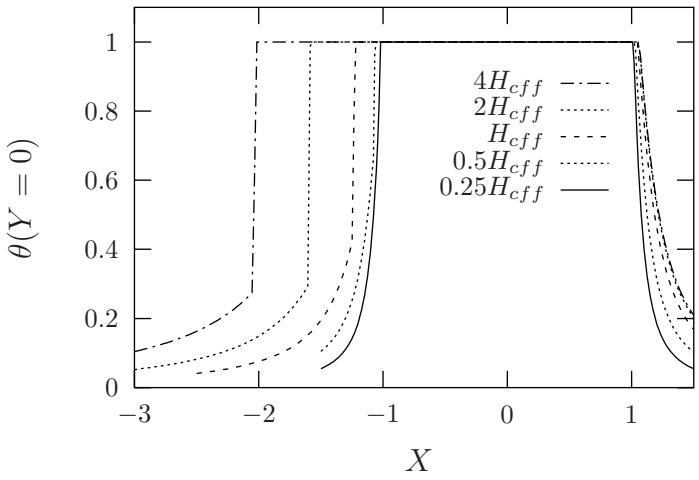


Figure 7.7: Fractional film content $\theta(X, Y = 0)$ for various values of H_{oil}. The Moes parameters are $M = 100$ and $L = 5$.

The fractional film content of Figure 7.7 and the lubricant distribution in

Figure 7.6, reveal that the discontinuity and the associated bow-wave become less for decreasing values of H_{oil}. In addition, Figure 7.7 shows that the discontinuity in the fractional film content vanishes at the outlet meniscus.

The reader is referred to the thesis of Chevalier [14] and the references therein for a more detailed discussion on all the topics mentioned above, as well as the presentation of numerous solutions at different M and L values. In addition, a very satisfactory agreement between numerical and experimental results was found in this reference. In the present study, the focus will be upon the stiffness and the time dependent solution, as discussed in subsequent sections.

7.1.2 Stiffness

The effect of starvation on the mutual approach Δ_∞, presented in Figure 7.8, shows that, for values of H_{oil} significantly larger than H_{cff}, Δ_∞ is close to the fully flooded approach ($\Delta_\infty = 0.897$). With decreasing values of H_{oil}, Δ_∞ slowly increases. Effects of starvation are only significant for $H_{oil}/H_{cff} < 1$. Below this value, Δ_∞ linearly increases to the Hertzian approach.

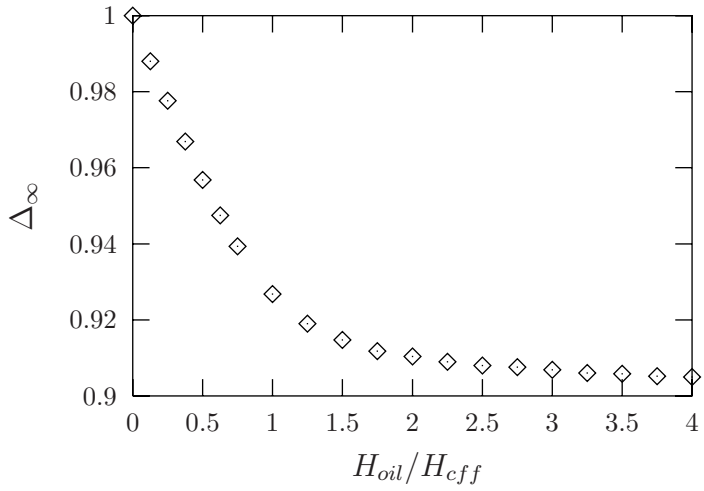


Figure 7.8: Δ_∞ as a function of H_{oil}/H_{cff} for $M = 100$ and $L = 5$.

These results support the statement that the curve-fit function of the flexibility, presented in Chapter 5, is quite accurate down to lubricant layers which are equal to the central fully flooded film thickness. If starvation is more severe, the flexibility, and hence the stiffness, is more Hertzian.

7.2 Time dependent solution

In this section, the time dependent solution is discussed. For the sake of brevity, we will restrict ourselves to the case of sinusoidally varying loads. In particular, it will be shown how the meniscus changes its shape with time. From the solution, it will be concluded that starvation leads to smaller values of the damping compared to the fully flooded situation.

The simulation starts from the steady state solution, as given in the previous section. Subsequently, the pressure and the film thickness as well as the fractional film content and the lubricant layer distribution are monitored as the force varies sinusoidally.

Figure 7.9 shows interference plots of the lubricant profile H_l at different instants in the simulation. The meniscus is indicated in the Figure by the solid line (see the arrow in the first frame). This line encloses the pressurized region. In the region enclosed by the line, the Figure thus shows the usual interference plot of the film thickness. In the outer region, it shows the lubricant layer. Figure 7.10 shows the associated pressure, the film thickness and lubricant film along the centre line ($Y = 0$). For clarity reasons, Figure 7.11 shows Δ versus T. In the Figure, letters have been included at certain points, marking the instants corresponding to those given in Figure 7.9 and Figure 7.10.

Similarly as in the fully flooded situation, the central film thickness hardly decreases in the initial stages of the simulation, despite the increase of the load. To compensate for the increased load, the contact area increases and the film becomes thinner only just outside the Hertzian contact circle, as can be observed most clearly in Figure 7.10. At this location, squeeze motion is induced. Thus, the reduction in film thickness with increasing load is less than would be observed if the load was increased quasi-statically. As in the fully flooded situation, film thickness modulations are thus induced in the inlet of the contact. They subsequently propagate downstream with the dimensionless speed of unity. Since the period of oscillation in terms of $\mathtt{T}$ was shown to map onto 2π, the period of oscillation in terms of T is $2\pi/\Omega_e$. Because the velocity with which the film thickness modulations propagate is unity, the wavelength of the modulations is $2\pi/\Omega_e$ and is 1 for this particular case.

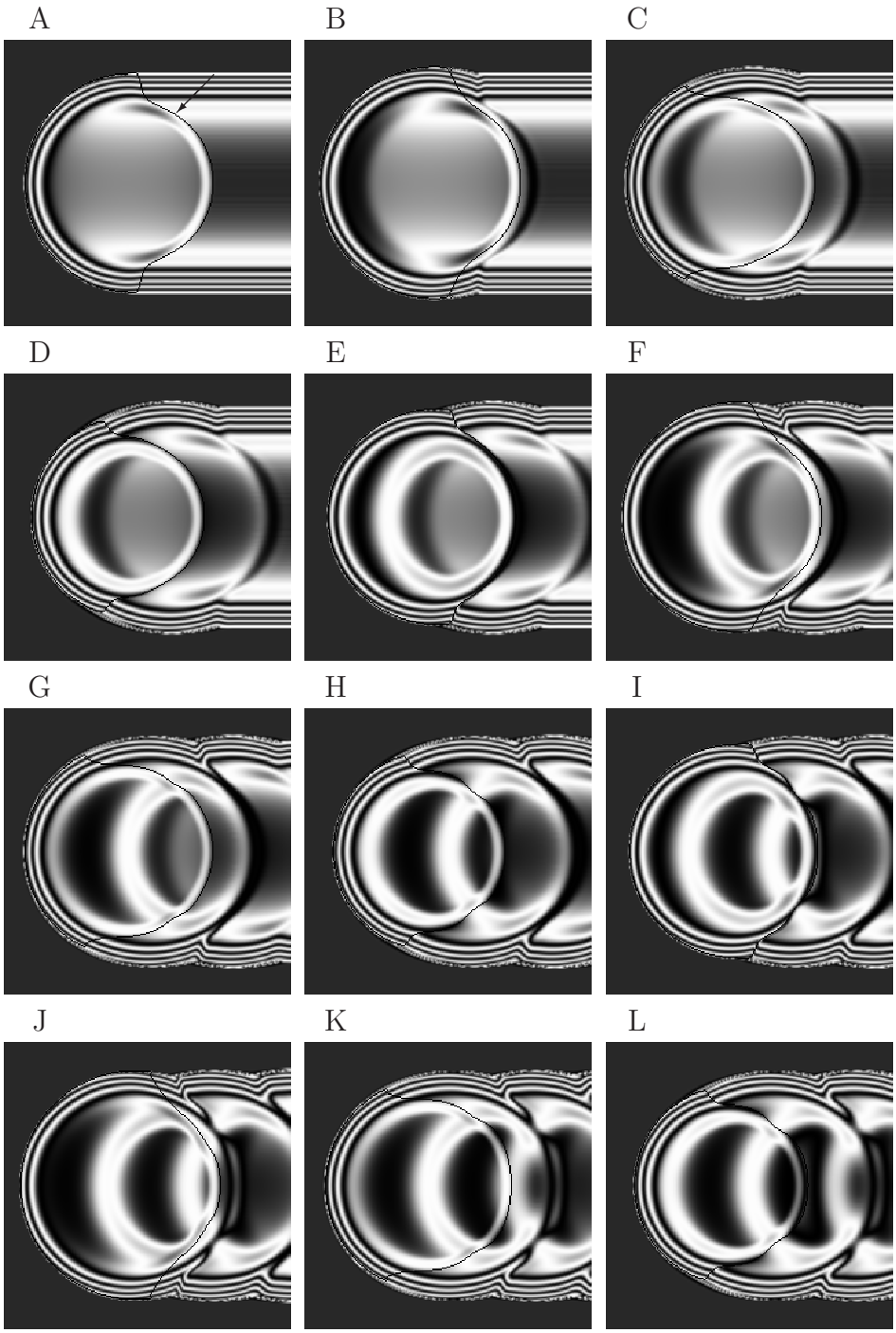


Figure 7.9: Lubricant distribution for $M = 100$, $L = 5$, $A = 0.3$ and $\Omega_e = 2\pi$. The time between consecutive frames is 0.25. The dimensionless wavelength is $\Lambda = 0.06$.

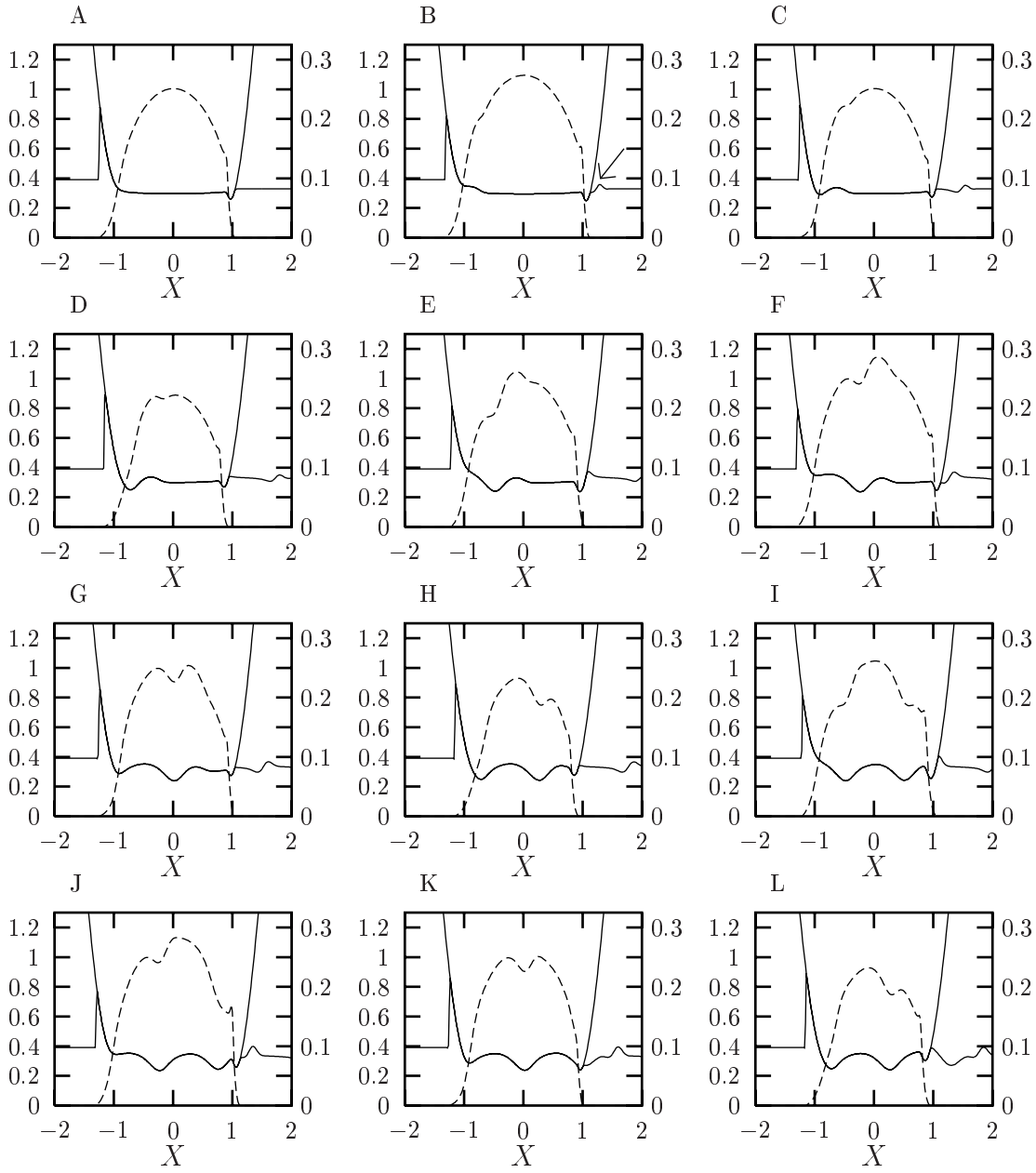


Figure 7.10: Centre line pressure (dashed lines, left labels), lubricant film and film thickness (solid lines, right labels) for $M = 100$, $L = 5$ and $\Omega_e = 2\pi$. The time between consecutive frames is 0.25.

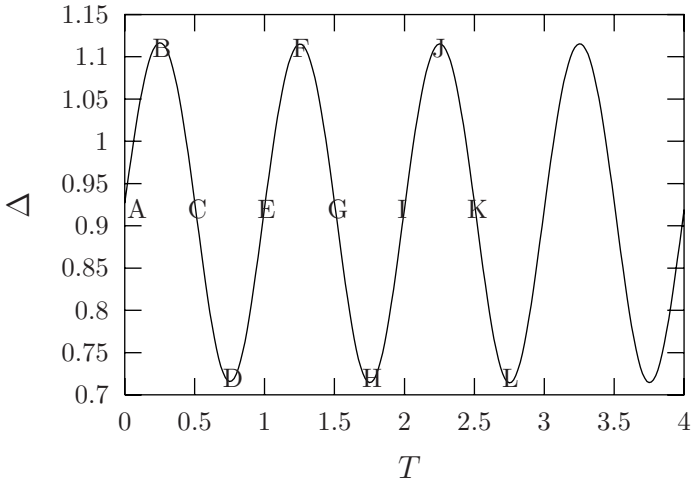


Figure 7.11: Δ as a function of T.

In the fully flooded case, the inlet was filled with oil everywhere. In the present case, one can clearly see that the size of the pressurized region depends on the load, i.e. it expands if the load increases and reduces in size as the load decreases. Hence, as the load increases, the lubricant that results from the squeeze motion, added to the lubricant fraction near the meniscus, is enough to fill the gap and the meniscus moves into the (initially) starved region. Figure 7.10 clearly shows that the squeeze motion at $T = 0$ has introduced an uprise of lubricant at the outlet of the contact, which can be observed clearly at the outlet at $T = 0.25$ (indicated by the arrow in picture B). The forming of the uprise is further illustrated in Figure 7.12. It shows the lubricant layer and film thickness at $T = 1.75$, $T = 2.00$ and $T = 2.25$. These instants correspond to the situation at the minimum, the average and the maximum load. It is noted that the uprise of lubricant is formed at the instant at which the approaching velocity is a maximum and not, as one might believe, at the maximum load.

In the same manner as the propagation of the film thickness modulations, lubricant layer modulations formed at the outlet of the contact propagate at the dimensionless speed of unity. This follows directly from the assumption that lubricant in the starved region is transported by means of Couette flow only. That is to say, in the starved region, the modified Reynolds equation reduces to:

$$-\frac{\partial \theta H}{\partial X} - \frac{\partial \theta H}{\partial T} = 0. \tag{7.1}$$

Consequently, $H_l = \theta H = \theta H(X - T)$ and, indeed, the lubricant layer modulations propagate at a dimensionless velocity of 1.

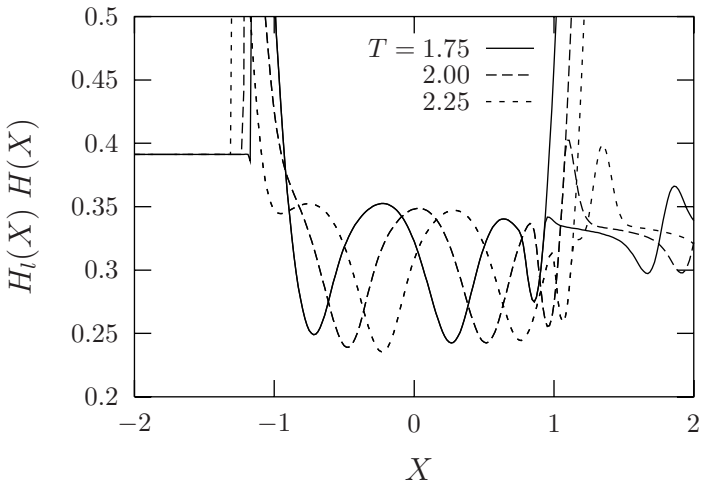


Figure 7.12: Lubricant layer and film thickness for $T = 1.75$, 2.00 and 2.25

It is noted that the transportation of lubricant inside the wake and, in fact, in the entire starved region is governed by the same mechanism as by which lubricant is transported within the high pressure zone. As stated earlier, in the high pressure zone Poiseuille flow is virtually absent and the Reynolds equation reduces to $-\partial(\bar{\rho}H)/\partial X - \partial(\bar{\rho}H)/\partial T = 0$. Hence, $\bar{\rho}H = \bar{\rho}H(X-T)$ which, by neglecting the effect of pressure on $\bar{\rho}$, validates the given statement.

From Figure 7.10, it can be inferred that the lubricant layer modulations at $Y = 0$ are transported downstream at a higher velocity than the velocity of the outlet meniscus; a conclusion which follows directly from the JFO relation. That is to say, if the velocity of the meniscus were larger, a bow-wave would form near the outlet, which should have shown up in the wake at later times. It may thus be concluded that at $T = 0$, at which for sinusoidally varying loads the approaching velocity is a maximum, the velocity of the meniscus in the X-direction at $Y = 0$, was smaller than unity.

Near the two lubricant ridges however, the normal velocity of the meniscus has exceeded the normal component of the outflow velocity. This can be concluded from the solution in Figure 7.9 at $T = 0.25$, by the somewhat ragged circular arc showing the discontinuity.

At $T = 0.25$, the load has reached its maximum value and squeeze motion becomes zero. Thus, after this instant, the pressurized region becomes smaller. Note that if the load becomes smaller, near the sides the inlet meniscus ceases to be concentric to the Hertzian contact circle. At this location, no lubricant is available; it was squeezed out earlier and deposited further away.

It is shown most clearly in the interference plot at $T = 0.75$, that the lubricant profile in the wake resembles a "footprint" of the film thickness at the maximum load. Since Poisseuille flow is virtually absent in the high pressure zone and, by definition, in the starved region, only a very small "boundary layer" exists, for the Poisseuille term to be of any influence. From the foot-print, it can be inferred that this influence must be very small.

Another interesting point is the lubricant layer distribution in the wake at times $T > 1.75$. From this time on, the film thickness modulations, which were induced at the start of the simulation, start to affect the lubricant profile at the outlet. Obviously, at locations where the film thickness is larger, the outflow is larger and, subsequently, so is the uprise in the lubricant profile. Hence, also crescent-shaped film thickness modulations appear in the lubricant profile in the wake. At $T = 2.75$, one clearly observes the lubricant profile which is seen periodically at all later times.

Finally, it is noted that if slip is to be included, i.e. if the surface speeds of the contacting solids are different from the average speed, the actual disctribution of the lubricant on the two surfaces becomes important. If all the lubricant adheres to one surface, it will, obviously, be transported with the velocity of that particular surface. This differs from the velocity with which it would be transported if all the lubricant adheres to the other surface, clearly illustrating that the model cannot be valid. However, it does remain valid as long as the lubricant is distributed evenly between the two contacting solids. For pure rolling, the model is valid regardless of the distribution.

7.3 Damping

To see whether, and if so when, starved lubrication affects damping, Figure 7.13 shows the dimensionless damping constant D for four values of the ratio H_{oil}/H_{cff}. For fully flooded lubrication, the value of the damping constant for $M = 100$ and $L = 5$ is 0.027 (see Chapter 5).

Although further investigation is required, the values shown in Figure 7.13 indicate that starvation, when it becomes important in the actual application, does have a significant effect on damping. Unlike the limited influence of starvation on stiffness, compared to the fully flooded situation the value of D decreases already by a factor of about 4 for $H_{oil} = H_{cff}$.

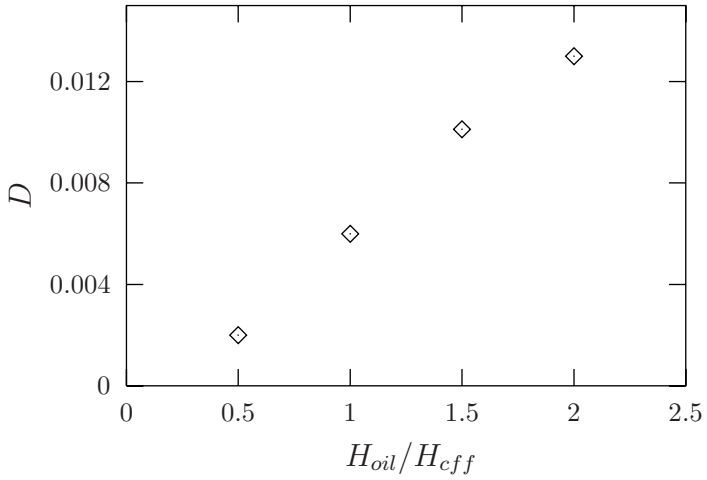


Figure 7.13: Damping constant D for four values of H_{oil}/H_{cff}.

7.4 Concluding remarks

From this chapter, it may be concluded that, if the thickness of the lubricant layer available to the contact decreases to values which are of the order of the central fully flooded film thickness, starvation should be included in the analysis. Fortunately however, the formula for the stiffness, as presented in Chapter 5, is fairly accurate for starved lubricated contacts down to $H_{oil} \approx H_{cff}$. Damping values calculated under conditions of full flooding, decrease (approximately in a linear fashion) as the amount of lubricant available is reduced. In that case, other sources of damping might dominate the damping in the bearing.

Especially in rolling element bearings, the assumption of a constant inlet film H_{oil} is not very realistic. In fact, the wake which forms at the outlet of the contact will be the inlet distribution for the next rolling element. (Since the time between two consecutive passages is very small, surface tension may be neglected and also the replenishment of the track.) The inclusion of a perturbed inlet oil layer in the model is beyond the scope of this thesis, but is an interesting topic for future research.

Chapter 8

Experimental verification

In this chapter, we compare results from experiments that were carried out on a ball and disc apparatus with results obtained with the EHL circular contact model. In the experiment, the applied load was rapidly increased. This results in an oscillatory motion of the contacting bodies. Modulations in the film thickness which result from these oscillations are clearly visible. After the numerical contact model was tailored to this experiment, a very satisfactory agreement was found. The experiments were carried out by R. Larsson and P. Eriksson at Luleå University of Technology, Luleå, Sweden.

8.1 Experimental setup

The experiments were performed on a so-called ball and disc apparatus for interferometric measurements of lubricant film thickness. The apparatus is schematically sketched in Figure 8.1.

A polished steel ball with a diameter of 50 mm and mounted on a shaft, is pressed against a Ø104 mm circular glass disc. The glass disc is mounted in a bearing and is free to rotate around its centre. The ball shaft is driven by an electric motor. As the ball rotates in contact with the glass disc, the disc also starts to rotate. Driven by the friction in the contact, a pure rolling condition occurs. The lubricant is applied in a small reservoir, in which the ball on its shaft is mounted. Fully flooded lubrication conditions occur since the rotation of the ball brings ample lubricant from the reservoir to the contact.

The contact area is illuminated by white light, filtered by means of an interference filter of 20 nm band width, to a wavelength of 577.5 nm. The glass disc is made semi-transparent by means of a thin chromium layer and interference occurs between light reflected by the chromium layer and light

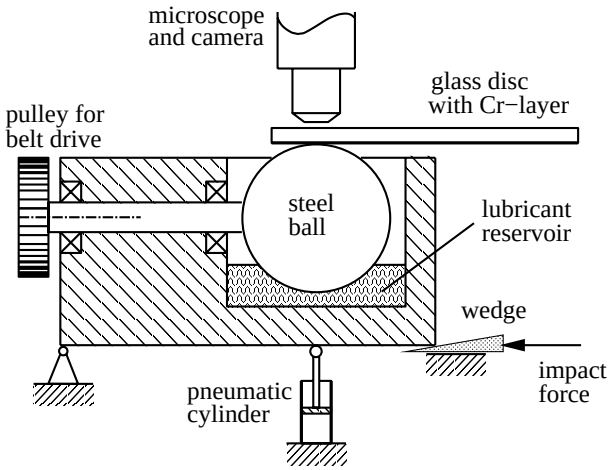


Figure 8.1: Experimental apparatus.

reflected by the ball surface. A pattern of bright and dark fringes can be observed in the contact. Each fringe corresponds to a level of more or less constant film thickness and the difference in film thickness between two neighbouring bright (or dark) fringes is $\Lambda/2n$, where Λ is the wavelength and n is the refractive index of the lubricant. Bright fringes occur at film thicknesses given by:

$$h = \frac{\Lambda}{2n}\left(q - \frac{1.2}{2\pi}\right) \qquad q = 1, 2, 3, \ldots \tag{8.1}$$

For more details about optical interferometry applied to EHL, reference is made to, e.g. Foord, Wedeven, Westlake and Cameron [22].

An initial static load is applied by a pneumatic cylinder mounted between the ball shaft holder and the base of the apparatus. A sudden change of the load is obtained by impacting a steel wedge into the space between the base and the ball holder. The load will then be increased within a few milliseconds.

The contact is magnified by a microscope and the time history recorded with a B/W high-speed video system. The system includes a high-speed video camera, an intensifier unit and a processor unit. The recordings are transferred to a PC for permanent storage after each session. The recording speed in this case was 4500 frames/s.

The lubricant used in this investigation is a fully formulated polyalphaolefin oil of viscosity grade VG46. The lubricant temperature was held constant at room temperature (23°C). The lubricant refractive index at atmospheric pressure and room temperature is $n_0 = 1.460$. The speed of the rotating ball shaft was 141 r.p.m., which corresponds to a surface velocity of 0.37 m/s.

Parameter	Value	Dimension
R	$12.5\,10^{-3}$	m
E'	$1.17\,10^{11}$	Pa
η_0	$92\,10^{-3}$	Pa s
α	$1.8\,10^{-8}$	Pa^{-1}
z	0.484	-
u_s	0.74	m/s
f_i	45	N
f_f	165	N

Table 8.1: Parameters and their values in the experiment.

The initial load of 45 N was almost stepwise increased to 165 N during the experiment. These loads correspond to maximum Hertzian pressures of 0.35 and 0.54 GPa respectively. The experimental data are summarised in Table 8.1.

8.2 Results

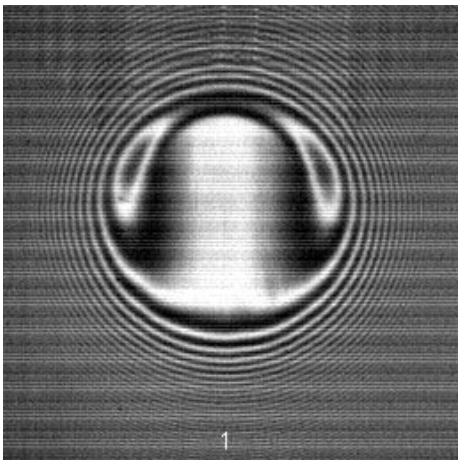


Figure 8.2: Interferogram at the initial load (t=0). In all subsequent pictures, the direction of flow is from bottom to top.

Figure 8.2 displays the measured interferogram of the film thickness at the initial load and a surface velocity of 0.37 m/s. All the characteristics of an EHL contact can be seen, i.e. a horseshoe-shaped film thickness distribution

with a plateau of roughly constant value in the contact centre, and the overall minimum value at the side lobes. For this particular case the measured central film thickness was $h_c = 0.34\ \mu$ m, and the minimum film thickness was $h_m = 0.15\ \mu$ m. The influence of pressure on the refractive index has been taken into account.

Figure 8.3 shows snapshots of the film thickness for times $t > 0$, i.e. after the sudden increase of the load. The time between two consecutive frames is 0.22 ms. The increase of the load quickly leads to a larger contact region. The contact size in the fifth frame already corresponds to the size of the Hertzian contact circle, associated with the final load. However, because of the greater stiffness of the film in the central region, see Section 5.2.1, the main changes in the film thickness due to squeeze motion occur just outside the initial Hertzian contact circle. The changes thus induced on the inlet side are propagated through the high viscosity region with the average surface speed. As stated in the previous chapters, this characteristic behaviour has been observed earlier, experimentally as well as theoretically, for surface features like dents, bumps and waviness, see Wedeven and Cusano [70], Kaneta [36], Kaneta, Sakai and Nishikawa [37] and [38] and Venner and Lubrecht [67] and [68].

It is noted that the film thickness does not immediately reach the new steady state value associated with the higher load. Due to inertia effects, the ball and disc start to carry out an oscillatory motion. In the same way as before, the changes in film thickness induced by this motion occur *near* the Hertzian contact circle corresponding with the final load. The periodic changes in film thickness induced in the inlet propagate through the high viscosity region, giving rise to the pattern of waves with a specific wavelength. Due to the energy dissipation in the lubricant, the amplitude of these waves decreases with time. Transient effects could not be distinguished after about $0.022s$, i.e. after 100 frames. Evidently, from that moment on the steady state situation was reached. From this relatively large period, it can be concluded that damping in the experimental apparatus together with the viscous damping in the lubricant must be small.

8.3 Theoretical study

The EHL contact model, as presented in Chapter 3, needs some adjustments, before it can be applied to the experiment just described. The conditions of the experiment are modelled by assuming force balance at $T = 0$. Subsequently, we assume that the load increases linearly with time, in a short interval after which, at T_f, it reaches its final value. A motivation for this

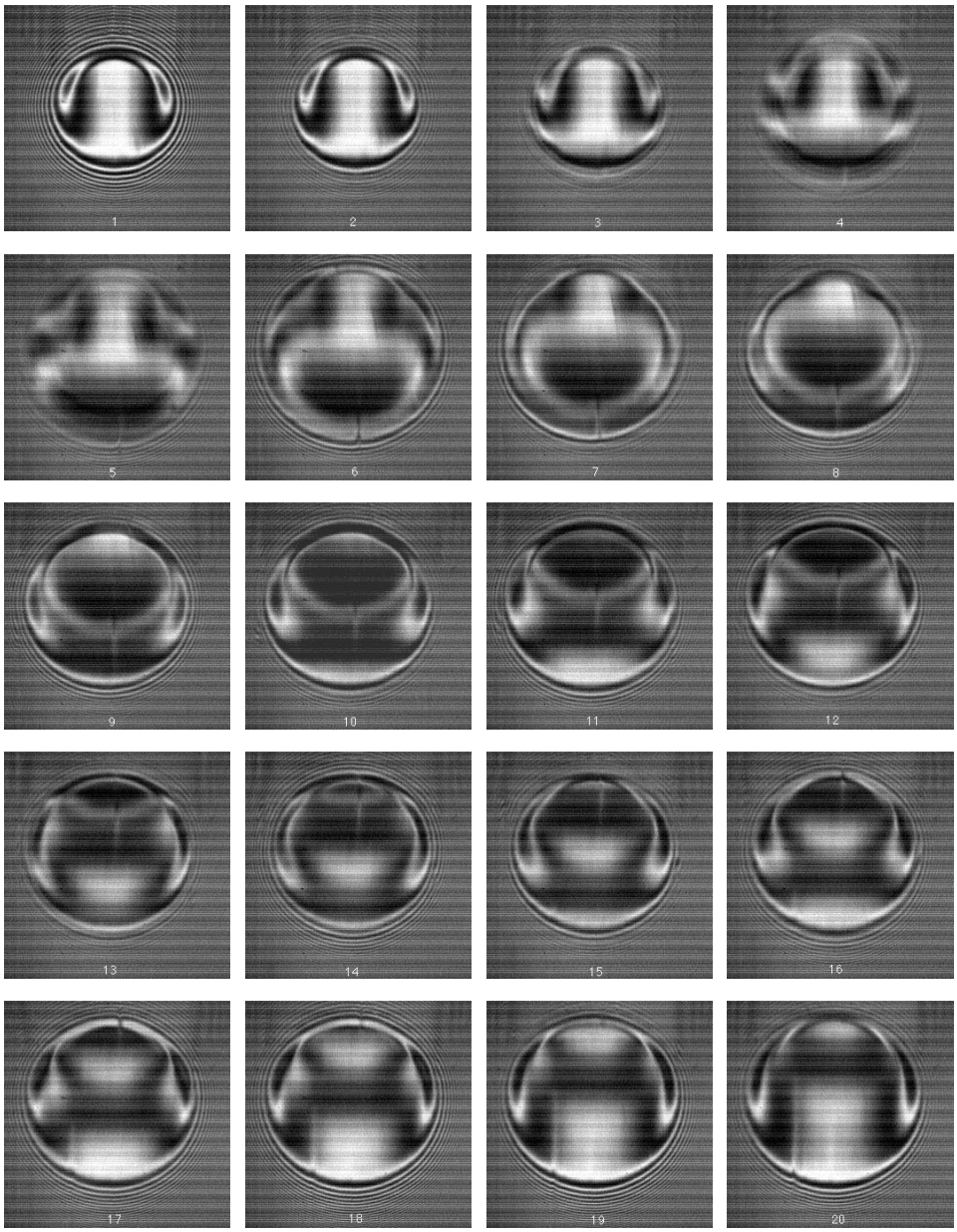


Figure 8.3: Interferogram of film thickness for $0 \leq t \leq 4.22ms$ (left to right/top to bottom). The time increment between consecutive pictures is $0.22ms$.

choice is given in Section 8.3.1.

Based on the Hertzian parameters associated with the initial load, the dimensionless equation of motion thus reads:

$$\frac{1}{\Omega_n^2}\frac{d^2\Delta}{dT^2} + \frac{3}{2\pi}\int_S P(X,Y,T)\,dX\,dY = \mathcal{F}(T), \tag{8.2}$$

where $\mathcal{F}(T)$ is the dimensionless load, given by:

$$\begin{array}{ll} \mathcal{F} = 1 & T = 0; \\ \mathcal{F} = 1 + \frac{(f_f - f_i)}{f_i}\frac{T}{T_f} & 0 < T \leq T_f; \\ \mathcal{F} = \frac{f_f}{f_i} & T > T_f. \end{array} \tag{8.3}$$

In this equation, f_i is the initial load and f_f denotes the final load.

The determination of the values Ω_n and T_f, is given in the next section.

8.3.1 Numerical solution

The discrete equations were solved in the domain ($-5.5 \leq X \leq 2.5$, $-4.0 \leq Y \leq 4.0$) on a uniform grid with mesh size $dX = dY = 0.03125$, i.e using 257x257 grid points. The steady state situation was calculated on a two times finer mesh.

The different variables were scaled by means of the Hertzian contact parameter values corresponding to the initial load. For the initial load, the region of large pressures thus roughly covers the domain $X^2 + Y^2 \leq 1$. Because the load increases, the region of large pressures grows and, after a very short time, is equal to Hertzian contact circle corresponding to the final load, i.e. to roughly $X^2 + Y^2 \leq (f_f/f_i)^{2/3}$. This increase must be anticipated in the choice of the calculational domain. Its boundaries must stay sufficiently far away from the high pressure region to avoid numerical starvation effects.

The values of the dimensionless parameters used in the numerical simulation are listed in Table 8.2. All values were computed directly from the data collected in Table 8.1, except for the value of the dimensionless frequency Ω_n and T_f.

In the model, a raceway of infinite mass is assumed. Hence, the only "vibrating mass" is the ball. In the experimental setup, the situation is more complex and, instead of Equation 8.2, one needs a model describing the dynamical behaviour of the entire experimental rig. This can be circumvented however, by the introduction of an *effective* mass, as is explained below.

In Chapter 5 it was shown that the wavelength of the film thickness modulations which propagate through the high pressure zone, is approximately

Parameter	Value	Dimension
a_i	$0.243\,10^{-3}$	m
p_{h_i}	$0.363\,10^{9}$	Pa
c_i	$2.37\,10^{-6}$	m
λ	$48.8\,10^{-3}$	-
$\bar{\alpha}$	6.53	-
Ω	2.4	-
M	58.1	-
L	4.6	-
$\mathcal{F}_i$	1	-
$\mathcal{F}_f$	3.67	-
T_f	1.69	-

Table 8.2: Hertzian and dimensionless parameters used in the simulation

$5.13/\Omega_n$. From this wavelength, the frequency at which the ball and glass disc vibrate can easily be deduced without measuring the actual value of the frequency in the dry contact case. In the measurements, the dimensionless wavelength, in terms of the parameters based on the *final* load, is approximately 1.1. In terms of the parameters based on the *initial* load, as is used in the simulation, this gives $\Omega_n = 2.4$. The associated effective mass is readily calculated as 1.43 kg. As could be expected, the effective mass is larger than the mass of the rolling element (0.51 kg), because it includes parts of the mass of the glass disc and the structural supports.

By lack of more detailed information, it is assumed that the load increases linearly in a short time interval. An assessment of this interval, which is difficult to determine experimentally, can be made with the help of the measured interferograms. As can be observed from Figure 8.3, the contact size only increases in the first five or six interferograms. After this time frame, the contact size starts to oscillate. Hence, it can be assumed that the load increases within the first five frames, i.e. within 1.11ms, and remains constant afterwards. This corresponds with $T_f = 1.7$.

8.3.2 Results

To facilitate a direct comparison with the experimental results the calculated film thickness will be presented in *pseudo* interferograms, i.e. an interferogram generated from the computed film thickness, based on the same wavelength of the light used in the experiments. Figure 8.4 presents the computed

interferogram for the film thickness at $t = 0$, together with the (dimensionless) pressure and film thickness along the centre line of the contact. The calculated central film thickness equals $h_c = 0.32\ \mu$ m, which differs only 5.9% from the experimental value. The calculated minimal film thickness was $h_m = 0.18\ \mu$ m, resulting in a much larger difference of approximately 20% from the measured value.

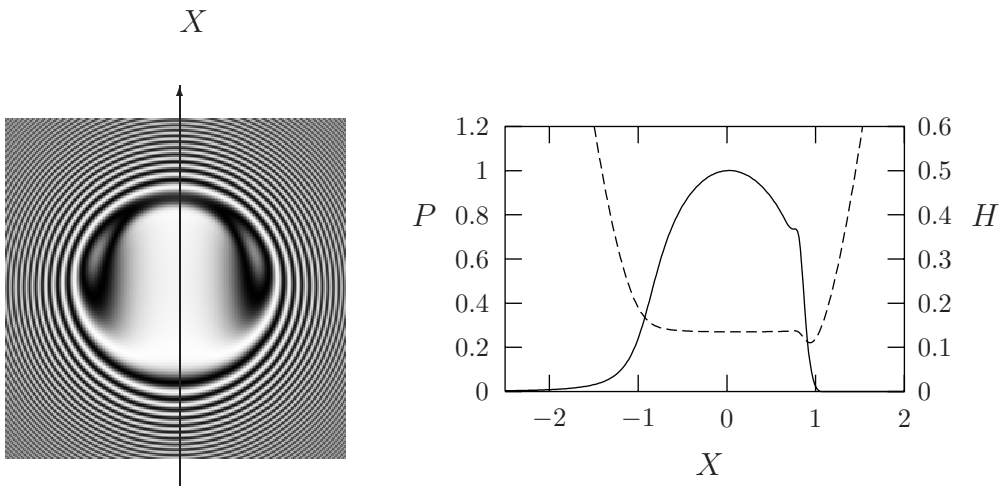


Figure 8.4: Steady state interferogram of the film thickness (left) and pressure and film thickness along the centre line ($Y = 0$).

Figure 8.5 illustrates the changes in the solution after the load has been changed. This figure shows snapshots of film thickness and pressure at the same instants as those in the experiments. From a comparison between Figures 8.3 and 8.5, it can be seen that a qualitatively good agreement with the theory and experiment exists, i.e. a fast growth of the contact size immediately after the load is increased and, subsequently, the film thickness modulations that travel through the contact. In both cases, these modulations gradually decay until the steady state solution is reached.

Since, with respect to the film thickness, the most interesting phenomena occur within the first twenty frames, we have limited the simulations to these frames only.

More results of the numerical simulation are given in Figures 8.6 and 8.7. These figures display, respectively, the approach of ball and the glass disc and the central and minimum film thickness as a function of time. In Figure 8.6 one observes that the approach δ quickly rises to approximately the steady state value associated with the final load. From that moment on, it oscillates around this value with hardly any reduction in amplitude in the considered

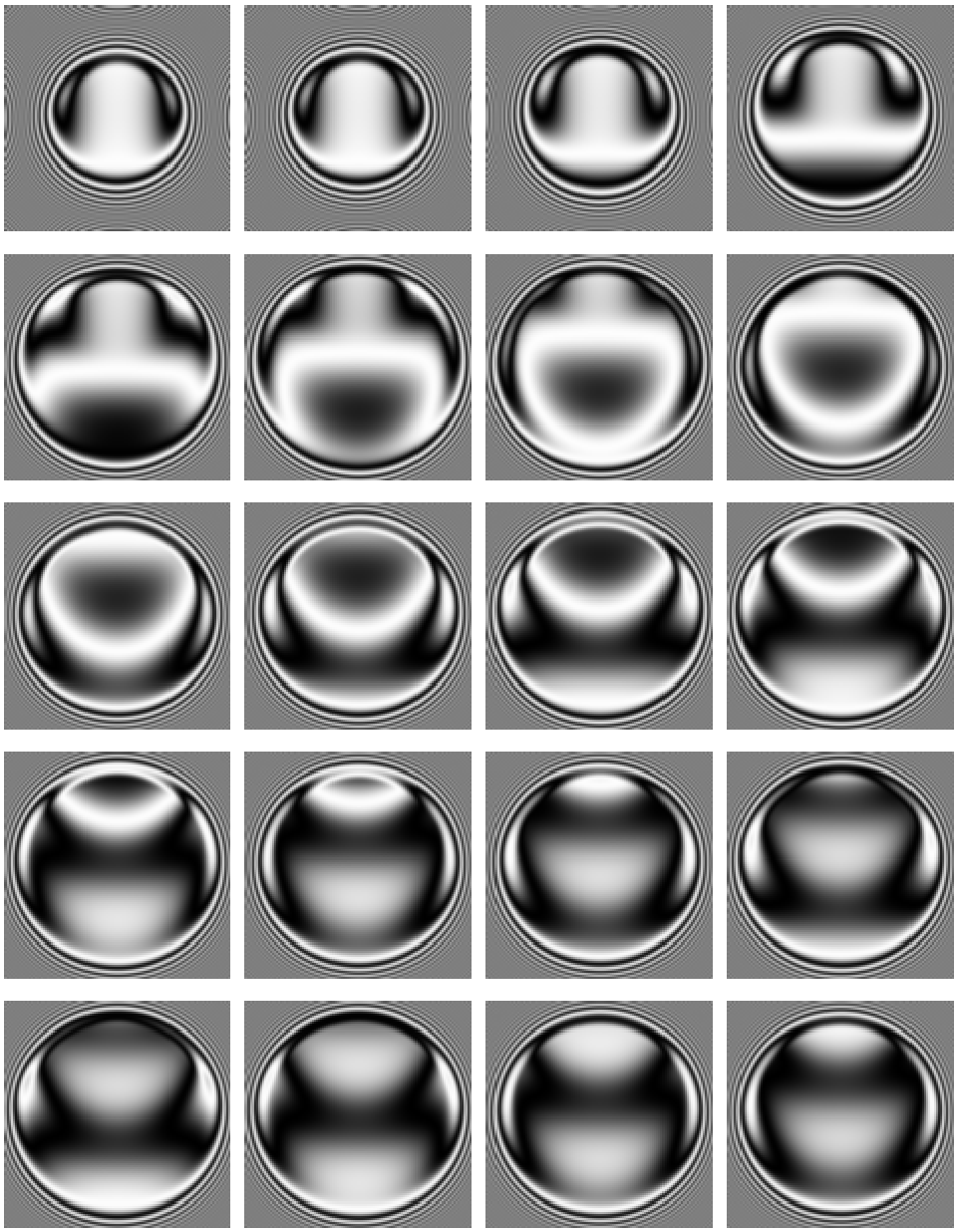


Figure 8.5: Computed pseudo interferograms of film thickness for $0 \leq t \leq 4.22\ ms$ (left to right/top to bottom). The time increment between consecutive pictures is $0.22\ ms$.

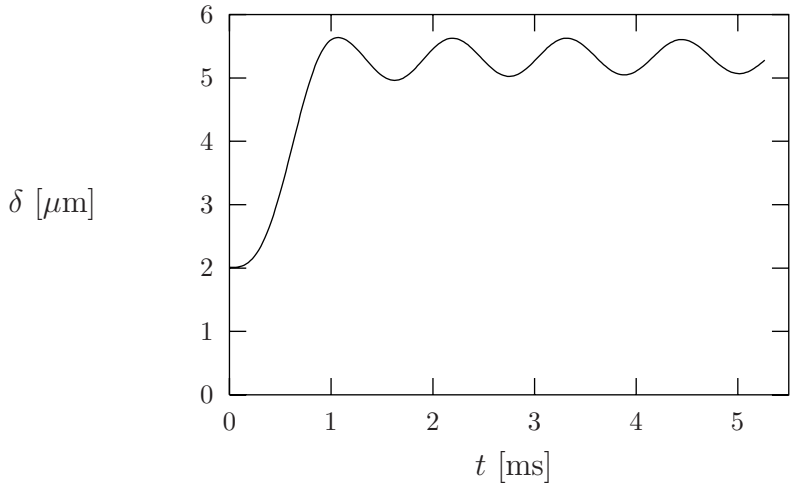


Figure 8.6: Approach δ as a function of time.

time frame. Energy dissipation in the lubricant must thus be small.

Figure 8.7 confirms the earlier observations, see Section 5.2.1, that the central film thickness increases in the initial stage of the experiment, although the load also increases rapidly during this period.

8.4 Concluding remarks

In this chapter, emphasis was placed upon the effect of the vibrations on film thickness and pressure. We circumvented the need to model the entire dynamic interaction between the ball and glass disc (including supports and other structural elements of the test apparatus) by introducing the concept of an *effective* mass. This choice is justified by the good agreement between the experimental and the theoretical results.

On the other hand, the results clearly indicate the need to include both squeeze and entraining motion in the analysis of vibrational behaviour on film thickness and pressure. As stated earlier, this can be determined in advance from the value of Ω_n, see Chapter 5.

Finally, no quantitative comparison was made between the experimental and calculated data. This has been left out since, in the experimental setup, the verification of fringe order is lost after the sudden impact. Furthermore, with respect to the determination of Ω_n, it is recognized that it would be more appropriate to measure the frequency in a dry contact situation and derive Ω_n from that measurement.

With respect to the study of noise and vibration of bearings, it is noted

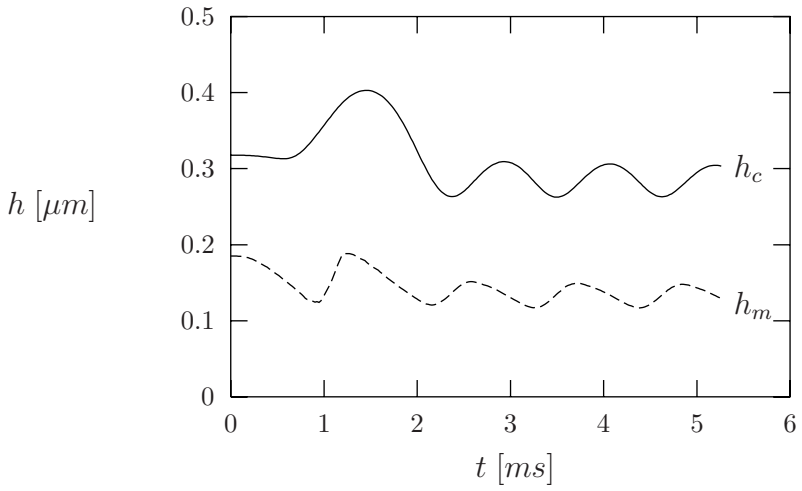


Figure 8.7: Central- (h_c) and minimum film thickness (h_m) as a function of time.

that it is extremely difficult to investigate the influence of the individual EHL contacts from measurements on a full bearing. Therefore, it would seem more efficient to derive simplified but accurate models for the individual EHL contacts, which may be based on both theory and experiment, and subsequently incorporate these models in dynamic models of the bearing. This subsequently enables the study of the vibrational behaviour and noise production of the bearing. The obtained good qualitative agreement between theory and experiment indicates that the model can indeed be used for this purpose and the present results must be looked upon as an initial step in understanding the influence of the lubricant on vibrations of rolling element bearings, which is important with respect to sound production and vibration control.

Recommendations for future research

The present research deals with the stiffness and damping of a single EHL contact as a function of the operating conditions, the contact geometry and lubricant supply. The presented relations are ready to be included in structural dynamics models of rolling element bearings, see Wensing [73]. With such models, general guidelines can be developed concerning the control of the vibrational behaviour and noise emission of bearings.

It is recommended that "ball on disc" experiments are carried out to validate more quantitatively the theoretical results, see Larsson and Lundberg [44], Nishikawa, Handa, Teshima, Matsuda and Kaneta [55] and Guangteng, Cann, Spikes and Olver [25]. The results may also be verified from "full bearing" measurements as described in Wensing [73] and Zeillinger [78].

It would be interesting to study, both experimentally and theoretically, whether vibrations from a preceeding rolling element will affect the vibrations of the next element via the disturbed lubricant layer in the wake. In the model, this has already been taking into account by taking the outlet lubricant layer as the boundary condition for the inlet lubricant layer. This would verify whether this coupling, which is neglected in structural dynamic models, should be included.

The presented results all apply to situations where the sum speed is non-zero. Thus, the asymptotic situation of "pure squeeze" motion ($\Omega_{n,e} \rightarrow \infty$) has been excluded. Obviously, this situation can not be analysed efficiently with the scaled equations given in the thesis and appropriate scales are thus required.

Finally, we restricted ourselves to an isothermal model, assuming Newtonian behaviour of the lubricant. For pure rolling, one may expect the predictions to be accurate. However, if sliding occurs, both non-Newtonian lubricant behaviour and thermal effects may become important, see Bos [9], Johnson and Tevaarwerk [34], Bair and Winer [4] and Jacobson [31].

Acknowledgements

I gratefully acknowledge all people who have in one way or another contributed to this work.

Especially, I would like to thank Kees Venner for his continuous support, stimulation and enthusiasm during the last four years.

I also thank Roland Larsson and Patrik Eriksson from Luleå University, Luleå, Sweden, for carrying out the experiments.

Thanks also to my promotors Ton de Gee and Henk Tijdeman for their valuable contributions to the thesis.

I would like to thank Marco Beltman, my room mate, for our discussions and valuable suggestions and Douwe Dijkstra and Hans Moes for their interest and help in finding the Weierstrass solution.

This research project was supported by SKF Engineering & Research Centre B.V., Nieuwegein, the Netherlands. From SKF, I thank Marie-Laure Dumont, Stathis Ioannides, Piet Lugt, Gerrit van Nijen, John Tripp and Jeroen Wensing, for their cooperation, support and interest.

Furthermore, I would like to thank the staff and colleagues of the Tribology Group, Faculty of Mechanical Engineering, University of Twente, for their continuous support and a pleasant working atmosphere: Ton de Gee, Johan Ligterink, Hans Moes, Wijtze ten Napel, Dik Schipper, Kees Venner, Laurens de Boer, Willy Kerver, Walter Lette, Erik de Vries, Jan Bos, Rob Cuperus, Edwin Gelinck, Rudi ter Haar, Qiang Liu, Harald Lubbinge, Henk Metselaar, Elmer Mulder, Daniel van Odyck, Matthijn de Rooij, Jan-Willem Sloetjes, Ronald van der Stegen, Harm Visscher en Andre Westeneng. Thanks also to all the students in the group for their interest. In particular, I thank Mark Aalbers for working with me on starved lubrication.

I am indebted to Debbie Vrieze and Annemarie Teunissen for their administrative support and I wish to thank Katrina Emmett and Andrew Stepek for their help concerning the English language.

I also would like to thank my parents, family and friends for their interest and encouragement.

Finally, I thank Lieke for her love, support and endurance.

Appendix A

Weierstrass solution

The energy equation for dry contact oscillations, induced by inertia forces, is given by Equation 2.23:

$$\frac{1}{2}\left(\frac{d\Delta}{d\mathcal{T}}\right)^2 + \frac{2}{5}\Delta^{5/2} = \Delta + \mathcal{C}. \tag{A.1}$$

If $\mathcal{C} = 0$ and introducing $\Delta = Z^2$, this equation reduces to:

$$\left(\frac{dZ}{d\mathcal{T}}\right)^2 + \frac{1}{5}Z^3 - \frac{1}{2} = 0. \tag{A.2}$$

Introducing $\mathcal{T} = \gamma\Gamma$, where $\gamma^2 = -20$, the differential equation becomes:

$$\left(\frac{dZ}{d\Gamma}\right)^2 - 4Z^3 + 10 = 0. \tag{A.3}$$

This equation is solved by the Weierstrass elliptic function $\wp(\Gamma; g_2, g_3)$ with invariants $g_2 = 0$ and $g_3 = 10$:

$$Z(\Gamma) = \wp(\Gamma + \mathcal{C}_2; 0, 10), \tag{A.4}$$

see Abramowitz and Stegun [1] and Whittaker and Watson [74]. $\mathcal{C}_2$ is a constant of integration which represents a phase shift. Note that, since $\mathcal{T}$ is real, Γ and $\mathcal{C}_2$ must be imaginary. Finally, using one of the properties of homogeneity: $\wp(\lambda\mathcal{T}; \lambda^{-4}g_2, \lambda^{-6}g_3) = \lambda^{-2}\wp(\mathcal{T}; g_2, g_3)$ and backsubstitution of $\Delta = Z^2$, yields:

$$\Delta = 400\wp\left(\mathcal{T} + \iota\sqrt{20}\mathcal{C}_2; 0, -1/800\right)^2, \tag{A.5}$$

where $\iota = \sqrt{-1}$ and the argument becomes real again. The Weierstrass solution is shown in Figure A.1 for two full periods. Note that Equation A.5

shows regions for which the solution goes to infinity and regions for which it is bounded. Obviously, the bounded solution is the appropriate one (the solution between $\mathcal{T}_n$ and $2\mathcal{T}_n$ ($4\mathcal{T}_n$ and $5\mathcal{T}_n$) given in Figure A.1).

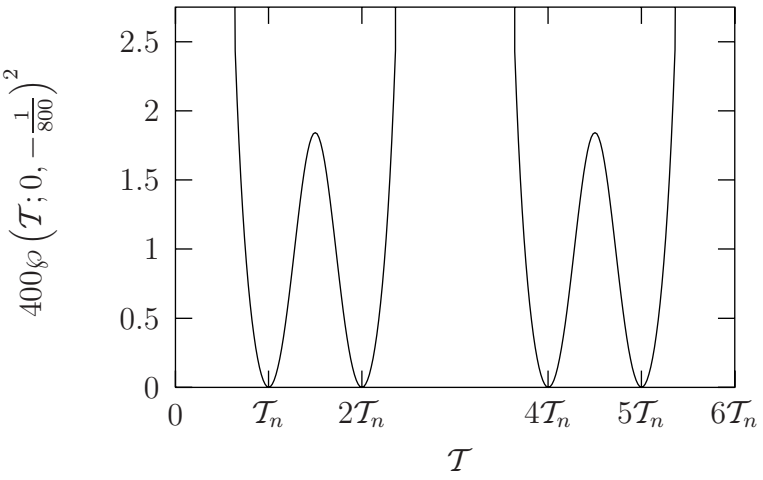


Figure A.1: The Weierstrass solution $400\wp(\mathcal{T};0,-1/800)^2$

The period of oscillation for the bounded solution, $\mathcal{T}_n$, can be determined as follows. As a starting point, we invert Equation A.2:

$$\frac{d\mathcal{T}}{dZ} = \left(-\frac{1}{5}Z^3 + \frac{1}{2}\right)^{-1/2}. \tag{A.6}$$

Integration yields:

$$\mathcal{T} = \int_{Z_0}^{Z} \left(-\frac{1}{5}z^3 + \frac{1}{2}\right)^{-1/2} dz = \int_{Z_0}^{Z} \{f(z)\}^{-1/2}\, dz, \tag{A.7}$$

where $f(z) = -z^3/5 + 1/2$. For our purpose of determining $\mathcal{T}_n$, we take $f(Z_0) = 0$, i.e. $Z_0 = (5/2)^{1/3}$. Note that for the mutual approach Δ, the lower bound is $\Delta_0 = Z_0^2 = (5/2)^{2/3}$, which is the maximum mutual approach of rolling element and raceway. Since the Weierstrass solution is symmetric about Z_0, half the period of oscillation can be obtained by taking the upper bound of the integral equal to $Z = \sqrt{\Delta} = 0$.

Since $f(Z_0) = 0$, Taylor's expansion for $f(z)$ around Z_0 reads:

$$f(z) = -\frac{3}{5}Z_0^2\,(z - Z_0) - \frac{3}{5}Z_0\,(z - Z_0)^2 - \frac{1}{5}\,(z - Z_0)^3\,. \tag{A.8}$$

On writing $(z - Z_0)^{-1} = \tau$ and $(Z - Z_0)^{-1} = \xi$, we have

$$\mathcal{T} = \int_{\xi}^{\infty} \left(-\frac{3}{5}Z_0^2\tau^3 - \frac{3}{5}Z_0\tau^2 - \frac{1}{5}\tau\right)^{-1/2} d\tau. \tag{A.9}$$

To remove the second term in the cubic involved, write:

$$\tau = -\frac{1}{3Z_0^2}\left(20\sigma + Z_0\right) \text{ and } \xi = -\frac{1}{3Z_0^2}\left(20s + Z_0\right), \tag{A.10}$$

and we get

$$\mathcal{T} = \int_s^\infty \left(4\sigma^3 - g_2\sigma - g_3\right)^{-1/2} d\sigma, \tag{A.11}$$

where $g_2 = 0$ and $g_3 = -1/800$, as was obtained before. This expression is the integral formula for the Weierstrass elliptic function with invariants g_2 and g_3 and thus $s = \wp(\mathcal{T}; g_2, g_3)$.

From the Taylor expansion and the variables s and ξ introduced, one can deduce that:

$$Z = Z_0 + \frac{1}{4}f'(Z_0)\left\{\wp(\mathcal{T}; g_2, g_3) - \frac{1}{24}f''(Z_0)\right\}^{-1}, \tag{A.12}$$

where the prime indicates differentation with respect to z. Taking $Z = 0$, which is the upper bound needed in the integral expression of Equation A.7, this provides the lower bound $s = \wp(\frac{1}{2}\mathcal{T}_n; g_2, g_3) = \frac{1}{10}\sqrt[3]{\frac{5}{2}}$, needed in expression Equation A.11. Hence, conform Equation 2.26, the period of oscillation is:

$$\mathcal{T}_n = 2\int_{\frac{1}{10}\sqrt[3]{\frac{5}{2}}}^\infty \left(4\sigma^3 + 1/800\right)^{-1/2} d\sigma \approx 5.383. \tag{A.13}$$

Appendix B

Advection operator

Using a first order discretization, especially in the high pressure zone, it is known that film thickness variations as observed in Chapter 5 decay fast a result of the discretization; not as a result of physical phenomena. Hence, for time dependent problems, great care must be taken to obtain accurate results and this is the topic of this Appendix.

The general form of what is known in computational fluid dynamics as the advection operator, acting on an unknown function u, is:

$$Lu = a\frac{\partial u}{\partial x} + b\frac{\partial u}{\partial y}. \tag{B.1}$$

This function is discussed here since, if the Poiseuille term vanishes, the Reynolds equation reduces to:

$$-\frac{\partial \bar{\rho}H}{\partial X} - \frac{\partial \bar{\rho}H}{\partial T} = 0, \tag{B.2}$$

which is a special form of the advection operator. (By inspection it follows that Equation B.2 reduces to Equation B.1 by setting $a = b = -1$, $\bar{\rho}H = u$, $X = x$ and $T = y$. We may thus restrict ourselves to the discussion to cases where $a = b = 1$.)

The solution of the homogeneous advection equation is given by:

$$u = \sum_{r=0}^{\infty} A_r e^{i\omega_r(y-x)}, \tag{B.3}$$

where ω_r is the frequency of oscillation in the "cross-stream"-direction and A_r is the amplitude. The amplitude and frequency follow from the initial conditions and the component $e^{i\omega(y-x)}$ is a so-called characteristic component.

Many stable second order schemes exist for the discretization of the advection operator. One such discretization is the so-called standard upstream discretization. It is defined by:

$$\begin{aligned} L^h u^h = & h_x^{-1} \left(1.5u_{i,j} - 2.0u_{i-1,j} + 0.5u_{i-2,j}\right) \\ & + h_y^{-1} \left(1.5u_{i,j} - 2.0u_{i-1,j-1} + 0.5u_{i-2,j-2}\right), \end{aligned} \tag{B.4}$$

where h_x and h_y denote the mesh sizes and i and j are the indices in the x- and y-directions. Neglecting higher order terms, the truncation error τ^{h_x,h_y}, i.e. the difference between the operator and the discrete approximation of the operator, is given by:

$$\tau^{h_x,h_y} = -\frac{h_x^2}{3}\frac{\partial^3 u}{\partial x^3} - \frac{h_y^2}{3}\frac{\partial^3 u}{\partial y^3} + \ldots, \tag{B.5}$$

where $\ldots$ denote higher order terms. For brevity, τ^{h_x,h_y} will henceforth be referred to as τ. For a single characteristic component $e^{i\omega(y-x)}$, the absolute value of the error can be shown to be:

$$|\tau| = \left|\frac{1}{3}h_x^2\omega^3\left(1 - \frac{h_y^2}{h_x^2}\right)\right| + \ldots \tag{B.6}$$

Apparently, the frequency ω strongly affects the discretization error. Furthermore, it is noted that the error will be dominated by higher order terms for $h_y = h_x$. Equation B.6 can now be used to choose the mesh size in the y-direction, when the mesh size in the x-direction is prescribed. As an illustration, Figure B.1 shows $3\tau/(h_x^2\omega^3)$ as a function of h_y/h_x. The picture shows that for the truncation error to be small, the mesh size h_y should equal h_x. Although h_y may be chosen smaller, values of h_y significantly larger than h_x lead to large truncation errors.

An alternative scheme is the so-called narrow upstream discretization. The narrow upstream discretization depends on the ratio h_y/h_x. For $h_x^{-1} \geq h_y^{-1}$ it reads:

$$\begin{aligned} L^h u^h = & \left(h_x^{-1} - h_y^{-1}\right)\left(1.5u_{i,j} - 2.0u_{i-1,j} + 0.5u_{i-2,j}\right) \\ & + \left(h_y^{-1}\right)\left(1.5u_{i,j} - 2.0u_{i-1,j-1} + 0.5u_{i-2,j-2}\right), \end{aligned} \tag{B.7}$$

whereas for $h_x^{-1} < h_y^{-1}$ it reads:

$$\begin{aligned} L^h u^h = & \left(h_x^{-1}\right)\left(1.5u_{i,j} - 2.0u_{i-1,j-1} + 0.5u_{i-2,j-2}\right) \\ & + \left(h_y^{-1} - h_x^{-1}\right)\left(1.5u_{i,j} - 2.0u_{i,j-1} + 0.5u_{i,j-2}\right). \end{aligned} \tag{B.8}$$

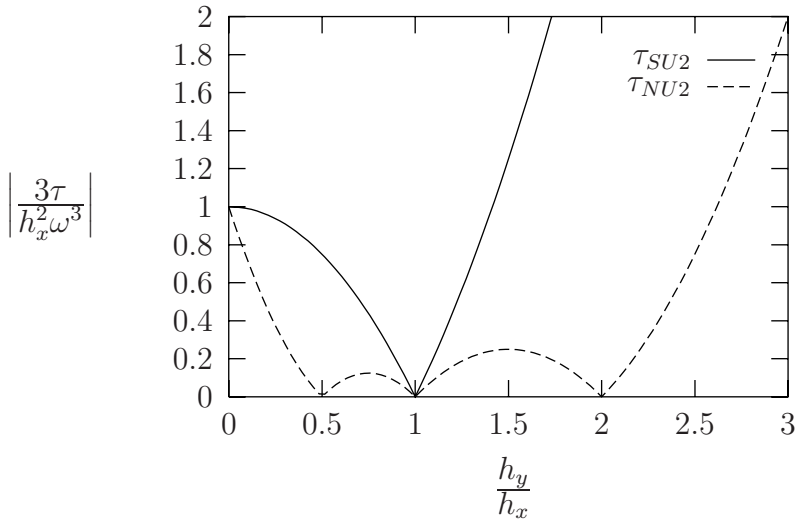


Figure B.1: Discretization error for SU2 and NU2 discretization.

Using a Taylor expansion, the discretization error for $h_x^{-1} \geq h_y^{-1}$ is:

$$\tau = -\frac{h_x^2}{3}\left(\frac{\partial^3 u}{\partial x^3} + 3\frac{\partial^3 u}{\partial x^2 \partial y} + 3\frac{h_y}{h_x}\frac{\partial^3 u}{\partial x \partial y^2} + \frac{h_y^2}{h_x^2}\frac{\partial^3 u}{\partial y^3}\right). \tag{B.9}$$

Otherwise, it is equal to:

$$\tau = -\frac{h_x^2}{3}\left(\frac{\partial^3 u}{\partial x^3} + 3\frac{h_y}{h_x}\frac{\partial^3 u}{\partial x^2 \partial y} + 3\frac{h_y^2}{h_x^2}\frac{\partial^3 u}{\partial x \partial y^2} + \frac{h_y^2}{h_x^2}\frac{\partial^3 u}{\partial y^3}\right). \tag{B.10}$$

Substitution of a characteristic component gives:

$$|\tau| = \left|\frac{1}{3}\omega^3 h_x^2\left(2 - \frac{h_y}{h_x}\right)\left(1 - \frac{h_y}{h_x}\right)\right|, \tag{B.11}$$

if $h_x^{-1} \geq h_y^{-1}$, whereas otherwise the error becomes:

$$|\tau| = \left|\frac{1}{3}\omega^3 h_x^2\left(2\frac{h_y}{h_x} - 1\right)\left(1 - \frac{h_y}{h_x}\right)\right|. \tag{B.12}$$

The error for the narrow upstream discretization is included in Figure B.1. One observes that the error is generally smaller than the error for the standard upstream discretization. Besides it vanishes not only at $h_y = h_x$, but also at $h_y = 2h_x$ and $h_y = 0.5h_x$. Given the mesh size h_x, the optimum choice for h_y would be $h_y = 2h_x$, as it is the largest mesh size for which the error is dominated by higher order terms. As a result, since the work associated

with both discretizations is equal, the narrow upstream discretization is to be preferred.

When the Poiseuille term vanishes in the Reynolds equation, which indeed occurs in the high pressure zone, one may conclude that the narrow upstream discretization is more accurate and justifies the choice for this type of discretization. The time step, i.e. the mesh size h_y in the discussion above, is chosen to equal h_x. Although the optimum choice would be $2h_x$, the choice is motivated by the smaller error that results from the discretization of the equation of motion.

An additional advantage of the narrow upstream discretization with respect to the multigrid solver is the accurate representation of the characteristic components on coarser levels. This especially shows if the time step is taken to equal twice the spatial mesh size of the finest level. Then, the characteristic component is represented with an accuracy that is dominated by higher order terms, not only on the finest grid, but also on two coarser grids. For $h_y = h_x$, this only holds for characteristic components on one level coarser.

Appendix C

Discrete elastic deformation

The film thickness equation for general elliptical contacts is given by:

$$H(X,Y,T) = -\Delta(T) + \mathcal{S}X^2 + (1-\mathcal{S})\,Y^2 + \frac{1}{\pi\mathcal{K}} \iint_S \frac{P\left(X',Y',T\right) dX'dY'}{\sqrt{\kappa^2\left(X-X'\right)^2 + \left(Y-Y'\right)^2}}, \tag{C.1}$$

The elastic deformation, i.e. the last term in the equation above, is discretized by approximating the pressure by a piecewise constant function on a uniform grid with mesh size h and value $P_{k,l}$ in $X'_k - h/2 \leq X' \leq X'_k + h/2$ and $Y'_l - h/2 \leq Y' \leq Y'_l + h/2$. The discrete deformation in (X_i, Y_j) can then be written as:

$$\sum_{k=0}^{n_X} \sum_{l=0}^{n_Y} K_{ikjl} P_{k,l}, \tag{C.2}$$

where K_{ikjl} follows from:

$$K_{ikjl} = \frac{1}{\pi\mathcal{K}} \int_{Y_l - h/2}^{Y_l + h/2} \int_{X_k - h/2}^{X_k + h/2} \frac{dX'dY'}{\sqrt{\kappa^2\left(X_i - X'\right)^2 + \left(Y_j - Y'\right)^2}} \tag{C.3}$$

The discrete kernel K_{ikjl} can be calculated analytically:

$$\begin{aligned} K_{ikjl} \;=\; & \frac{1}{\pi\mathcal{K}} \Bigg\{ |X_p| \operatorname{arcsinh}\left(\frac{Y_p}{X_p}\right) + |Y_p| \operatorname{arcsinh}\left(\frac{X_p}{Y_p}\right) \\ & - |X_m| \operatorname{arcsinh}\left(\frac{Y_p}{X_m}\right) - |Y_p| \operatorname{arcsinh}\left(\frac{X_m}{Y_p}\right) \\ & - |X_p| \operatorname{arcsinh}\left(\frac{Y_m}{X_p}\right) - |Y_m| \operatorname{arcsinh}\left(\frac{X_p}{Y_m}\right) \\ & + |X_m| \operatorname{arcsinh}\left(\frac{Y_m}{X_m}\right) + |Y_m| \operatorname{arcsinh}\left(\frac{X_m}{Y_m}\right) \Bigg\} \end{aligned} \tag{C.4}$$

where

$$\begin{aligned} X_p &= X_i - X_k + h_X/2, \quad X_m = X_i - X_k - h_X/2, \\ Y_p &= (Y_j - Y_l + h_Y/2)/\kappa, \quad Y_m = (Y_j - Y_l - h_Y/2)/\kappa. \end{aligned} \tag{C.5}$$

Appendix D

Multigrid

In this appendix a résumé of the most important features of the multigrid technique is given. It discusses the behaviour of standard relaxation schemes. Subsequently, it shows how this behaviour can be exploited to reduce computation time. The so-called Correction Scheme for linear and the Full Approximation Scheme for non-linear operators is discussed. In addition, it discusses the Full Multigrid Algorithm.

D.1 Relaxation

To illustrate the multigrid procedure, consider the following equation:

$$L(u) = f, \tag{D.1}$$

where L is a differential operator, u is the unknown solution and f is the right-hand side function. A discrete approximation u^h of u should then satisfy the approximate equation:

$$L^h(u^h) = f^h, \tag{D.2}$$

where L^h and f^h are discrete approximations of L and f, respectively. The superscript h denotes the mesh size.

Since the ultimate goal is to find the solution of the continuum problem $L(u) = f$, our objective is to obtain an approximation $\bar{u}^h$ to u^h for which the algebraic error $\|u^h - \bar{u}^h\|$ is smaller than the discretization error $\|u - u^h\|$. A more accurate approximation to u^h would only imply additional work without increasing the accuracy; the discretization error dominates the difference between u and $\bar{u}^h$.

Starting from an approximation $\tilde{u}^h$, the error in $\tilde{u}^h$ can be reduced by relaxations. Gauss-Seidel relaxation efficiently reduces high-frequency error

components in $\tilde{u}^h$. Low-frequency components however are hardly affected and hence the asymptotic speed of convergence is low.

From a local mode analysis, one can deduce the asymptotic reduction factor $\bar{\mu}$, defined according to:

$$\bar{\mu} = max\left\{|\mu(\theta_1,\theta_2)|\,;\pi/2 \leq (\theta_1,\theta_2) \leq \pi\right\}, \tag{D.3}$$

where $\mu(\theta_1,\theta_2)$ is the amplification factor. This factor equals the ratio between the absolute value of the amplitude of a Fourier component (θ_1,θ_2) *after* a relaxation sweep and the absolute value of the same component *before* the sweep. It tells us how well high-frequency components, which can not be represented on a coarser grid, are reduced by a single relaxation sweep. Hence, high-frequency error components are at least reduced by a factor $\bar{\mu}^\nu$, where ν is the total number of relaxations on the grid.

D.2 Correction Scheme (CS)

After a number of relaxation sweeps, due to the fast convergence of the high-frequency components, the error becomes a smooth function. Instead of continuing the relaxation process on the fine grid, the error is now represented on a coarser grid by means of restriction. Once the error is solved on the coarse grid, it can subsequently be used to update the solution on the fine grid. This process forms the basis for the so-called Correction Scheme (CS) which is discussed below.

Consider again the discrete equation:

$$L^h(u^h) = f^h \tag{D.4}$$

and let $\tilde{u}^h$ be the approximation obtained after a few relaxations on the fine grid h.

A measure for the accuracy of the approximation $\tilde{u}^h$ is given by the residual r^h, defined according to:

$$r^h = f^h - L^h(\tilde{u}^h). \tag{D.5}$$

Using $L^h(u^h) = f^h$, and assuming L^h to be linear, it follows that:

$$r^h = L^h(u^h) - L^h(\tilde{u}^h) = L^h(u^h - \tilde{u}^h). \tag{D.6}$$

Substitution of the error, defined according to $v^h = u^h - \tilde{u}^h$, yields:

$$L^h(v^h) = r^h. \tag{D.7}$$

Since, after a number of relaxations, v^h is smooth compared to the mesh size h, it can be approximated on a coarser grid H, where usually $H = 2h$. This serves two purposes. First, the computation time decreases since the number of unknowns is reduced by a factor 2^d, where d is the dimension of the problem. Secondly, some of the low-frequency errors become high-frequency errors compared to the mesh size H and, as a result, they can efficiently be reduced by relaxation. The coarse grid approximation v^H to v^h then follows from:

$$L^H(v^H) = I_h^H r^h, \tag{D.8}$$

where L^H is a representation of L^h on the coarse grid and I_h^H is a restriction operator. It is usually sufficient to take L^H to be the fine grid operator evaluated on the coarse grid, but other representations do exist. Once a sufficient approximation $\bar{v}^H$ to v^H is obtained, it is used to update the solution on the fine grid according to:

$$\bar{u}^h = \tilde{u}^h + I_H^h \bar{v}^H, \tag{D.9}$$

where I_H^h is an interpolation operator. High-frequency components introduced by the interpolation are eliminated by one or two additional relaxations on the fine grid.

The restriction and interpolation operators will be discussed in a subsequent section.

D.3 Full Approximation Scheme (FAS)

The CS can only be applied for linear operators L^h. For non-linear operators, the Full Approximation Scheme (FAS) is required.

FAS is derived from the equation for the residual:

$$r^h = L^h(u^h) - L^h(\tilde{u}^h). \tag{D.10}$$

Substitution of the error $v^h = u^h - \tilde{u}^h$ and rearranging yields:

$$L^h(\tilde{u}^h + v^h) = L^h(\tilde{u}^h) + r^h. \tag{D.11}$$

Since v^h is again smooth compared to the mesh size h and the non-linearity of L^h prevents the reduction to Equation D.8, the "full" equation is represented on the coarse grid:

$$L^H(\hat{u}^H) = \hat{f}^H, \tag{D.12}$$

where $\hat{u}^H$ is the coarse grid variable:

$$\hat{u}^H \equiv I_h^H(\tilde{u}^h + v^h) = I_h^H \tilde{u}^h + v^H, \tag{D.13}$$

and $\hat{f}^H$ is the FAS right-hand side, defined according to:

$$\hat{f}^H \equiv L^H(I_h^H \tilde{u}^h) + I_h^H r^h. \tag{D.14}$$

From a sufficiently accurate approximation $\bar{\hat{u}}^H$ of the coarse grid variable $\hat{u}^H$, a coarse grid approximation $\bar{v}^H$ of v^h follows from:

$$\bar{v}^H = \bar{\hat{u}}^H - I_h^H \tilde{u}^h, \tag{D.15}$$

which is subsequently used to update the fine grid solution according to:

$$\bar{u}^h = \tilde{u}^h + I_H^h \bar{v}^H. \tag{D.16}$$

Note that, if applied to linear operators, FAS reduces to CS.

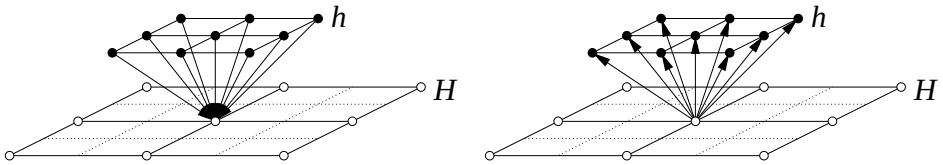
D.4 Restriction and interpolation operators

The restriction operator I_h^H transfers a function, e.g. the residual, to the coarser grid. If standard coarsening is applied, the simplest form of restriction is so-called injection. Injection transfers the value of the function on the fine grid to the corresponding coarse grid point. A more accurate transfer is obtained using full weighting. With full weighting, the value in the coarse grid point is a weighted average of function values in points adjacent to the corresponding fine grid point. Full weighting has the advantage that the components $|\theta_{1,2}| = \pi$, i.e. the highest frequency which can be represented on the grid, vanish in the coarse grid representation, whereas injection projects these components onto $\theta_{1,2} = 0$. The coarse grid correction is therefore less accurate for injection than it is in the case of full weighting.

The interpolation or prolongation operator I_H^h interpolates the function on the coarse grid to the fine grid. Linear interpolation is usually sufficient. Full weighting and linear interpolation are illustrated in Figure D.4 for two-dimensional problems.

D.5 Coarse grid correction cycle

The coarse grid equations can usually be solved by means of the same relaxation procedure as was derived for the original set of equations. Hence, as for the fine grid equation, the convergence of the coarse grid equation slows down due to the poor reduction of the low-frequency error components and, after a number of relaxations, the error is smooth with respect to the coarser mesh. It can thus be approximated accurately on an even coarser grid.



$$I_h^H = \frac{1}{16}\begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix} \qquad I_H^h = \frac{1}{4}\begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix}$$

Figure D.1: Restriction operator I_h^H and prolongation operator I_H^h.

Indeed, the procedure can be repeated recursively until a grid is reached for which the solution can be obtained within only a limited number of relaxations. This recursive procedure is known as the coarse grid correction cycle or $V(\nu_1, \nu_2)$-cycle, where ν_1 and ν_2 denote, respectively, the number of pre- and post relaxations. One refers to $W(\nu_1, \nu_2)$-cycles if, on each grid, 2 coarse grid correction cycles are used to update the solution on that particular grid. The $V(\nu_1, \nu_2)$- and $W(\nu_1, \nu_2)$-cycles are illustrated in Figure D.2. ν_0 indicates the number of relaxations needed to solve the equations on the coarsest grid.

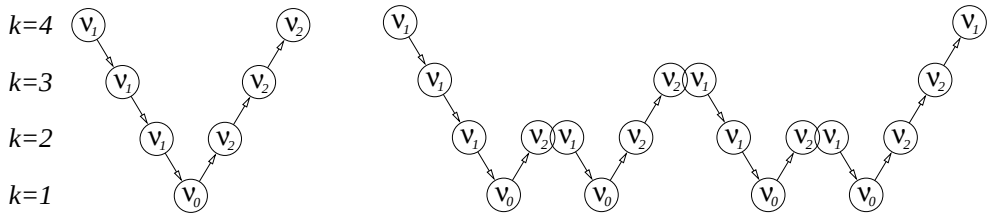


Figure D.2: $V(\nu_1, \nu_2)$- and $W(\nu_1, \nu_2)$-cycle.

The amount of work, required to obtain a solution using the cycles described above, can be given in terms of so-called work units (WU). One WU equals the amount of work needed for one complete relaxation sweep on the finest grid. If n denotes the number of unknowns on the finest grid, one WU thus requires $O(n)$ operations.

Neglecting the work associated with the intergrid transfers I_h^H and I_H^h, which only require $O(n)$ operations, as well as the ν_0 relaxations on the

coarsest grid, the work associated with one V-cycle, W_V, adds up to:

$$W_V = (\nu_1 + \nu_2)\, WU \left(1 + 2^{-d} + 2^{-2d} + \ldots\right) \leq \frac{\nu_1 + \nu_2}{(1 - 2^{-d})} WU. \tag{D.17}$$

If it is assumed that, by the coarse grid correction cycle, all low-frequency error components are solved exactly and that there is no interaction between low- and high-frequencies in the grid transfers, then the error reduction of one cycle equals the reduction of high-frequency components on the finest grid.

Hence, starting from an arbitrary approximation, for which the error is $O(1)$, the number of cycles M needed to solve the problem below the discretization error follows from:

$$\left(\bar{\mu}^{\nu_1+\nu_2}\right)^M \leq h^s, \tag{D.18}$$

where s is the order of the discretization. Since $h = O(1/n)$, it follows that:

$$M \geq \frac{s \ln(n)}{\ln\left(1/\bar{\mu}^{\nu 1+\nu 2}\right)}. \tag{D.19}$$

Hence, $O(\ln n)$ cycles are needed to converge below the discretization error and, since the work needed for one cycle is $O(n)$, the total amount of work equals $O(n \ln n)$. The amount of work needed to converge below the discretization error using relaxation only is $sn^2 \ln(n)$, showing that already by applying correction cycles only, a significant gain in computing time can be achieved.

D.6 Full Multigrid (FMG)

As explained in the previous section, the poor initial approximation dictates that $O(\ln n)$ coarse grid correction cycles are required to converge below the discretization error. The number of cycles can be reduced to $O(1)$ by improving the initial approximation.

In FMG, the initial approximation on the fine grid follows from interpolation of the solution on the next coarser grid, which has previously been calculated by a similar FMG algorithm. The recursive algorithm, thus established, is depicted in Figure D.3.

Since the error in the solution on the coarser grid is now $O(H^s)$, the number of cycles M, follows from the condition:

$$\left(\bar{\mu}^{\nu_1+\nu_2}\right)^M H^s \leq h^s. \tag{D.20}$$

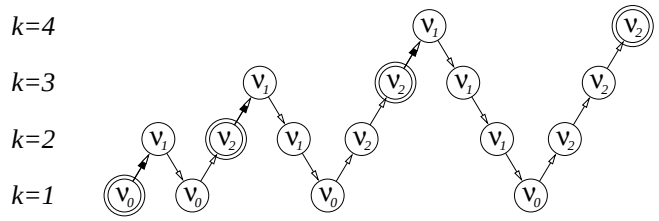


Figure D.3: The FMG algorithm with $1V(\nu_1, \nu_2)$-cycle per refinement.

For standard coarsening, this yields for M:

$$M \geq \frac{s \ln(2)}{\ln\left(1/\bar{\mu}^{\nu_1+\nu_2}\right)}. \tag{D.21}$$

Hence, only $O(1)$ cycles are needed to obtain a converged solution, provided the interpolation is at least of order s, but preferably of higher order.

The total amount of work that is needed in the FMG algorithm equals:

$$W_{FMG} = MW_V\left(1 + 2^{-d} + 2^{-2d} + \ldots\right) \leq \frac{1}{(1-2^{-d})} MW_V. \tag{D.22}$$

Substitution of Equation D.17 yields:

$$W_{FMG} = \frac{\nu_1 + \nu_2}{(1-2^{-d})^2} MWU. \tag{D.23}$$

Consequently, since 1 WU is $O(n)$, the total amount of work is $O(n)$ and the algorithm is optimal.

D.7 Convergence

Since a sequence of grids is used in the FMG algorithm, one can easily check whether the solution has converged below the level of the discretization error by comparing converged solutions on different grids. These solutions are indicated by the double circles in Figure D.3. In addition, the order of convergence can easily be verified.

If the converged solution on grid k is denoted by $\bar{u}^k$, then a difference norm $ERR(k, k-1)$ can be defined according to:

$$ERR(k, k-1) = H^d \sum \left| \bar{u}^{k-1} - I_h^H \bar{u}^k \right|, \tag{D.24}$$

where H denotes the mesh size on grid $k-1$, d is the dimension of the problem and I_h^H denotes injection. The difference norm consists of four different components. First, it contains the discretization errors on level k and level $k-1$, i.e. $h^d \sum \left| u - \bar{u}^h \right|$ and $H^d \sum \left| u - \bar{u}^H \right|$. In addition, it contains the algebraic errors on level k and $k-1$, i.e. $h^d \sum \left| \bar{u}^h - u^h \right|$ and $H^d \sum \left| \bar{u}^H - u^H \right|$. If the algebraic errors on both grids are small compared to the discretization error, the difference norm is dominated by the difference between the discretization errors on the grids. Then, the difference norm remains unchanged if the algebraic error is reduced by means of one additional coarse grid correction cycle and the solutions must have converged to the level of the discretization error.

The order of discretization is verified by monitoring the ratio between the difference norms $ERR(k, k-1)$ and $ERR(k-1, k-2)$. From a similar analysis, this ratio should equal $(2)^{-s}$ for standard coarsening.

Appendix E

Multilevel multi-integration

In this appendix, the multilevel multi-integration technique is discussed in detail. The discussion of results will be restricted to the circular contact kernel. The elliptical kernel is discussed in the main text.

E.1 Introduction

The elastic deformation integral is an example of so-called integral transforms. Integral transforms are integrals of the type:

$$w(x) = \int_S K(x,y)u\,(y)\,dy, \quad (x,y) \in S \subseteq I\!R^d, \tag{E.1}$$

where $K(x,y)$ denotes the kernel of the integral transform and $u(y)$ is a given function. Obviously, the evaluation of a discrete approximation of the integral transform requires a complete $O(n)$ summation, where n is the number of grid points on which u is defined. Consequently, the evaluation of the integral in all the n grid points x_i requires $O(n^2)$ operations and this would also be the complexity in the FMG algorithm in which evaluating $w(x)$ appears as a subtask, i.e. when u is to be solved. The computing time would thus be dominated by the matrix multiplication. Multilevel integration aims at the reduction of the $O(n^2)$ complexity.

E.2 Discretization

A discretization of Equation E.1 is obtained by approximating the function $u(y)$ on the interval (y_j^h, y_{j+1}^h) by a piecewise polynominal function $\hat{u}^h$ of degree $2s-1$, where $y_j^h = y_a + jh$ denote the gridpoints on which $u(y_j^h)$ is given. Let the grid points $x_i^h = x_a + ih$ coincide with y_i^h, $u_j^h = u^h(y_j^h)$ be the

function values in y_j^h and $w_i^h = w(x_i^h)$ the integral transform evaluated in x_i^h. The discretization is then obtained by piecewise integration using $\hat{u}^h$ instead of u^h, which yields:

$$w_i^h \equiv \int_S K(x,y)\bar{u}^h(y)dy = \sum_j K_{i,j}^{hh} u_j^h. \tag{E.2}$$

The error introduced by the discretization is $h^{2s}u^{2s}(\zeta)$, where $u^{2s}(\zeta)$ denotes the $2s^{th}$ derivative of u in some point $\zeta \in S$.

The influence coefficients $K_{i,j}^{hh}$ can often be evaluated analytically. The coefficients for the elastic deformation are given in Appendix C. In terms of the elastic deformation, $K_{i,j}^{hh}$ represents the deformation in x_i^h due to a unit pressure on $y_a + (j-1/2)h \leq y \leq y_a + (j+1/2)h$.

E.3 Coarse grid integration

As in multigrid, in order to obtain a fast algorithm one needs to look for smooth functions which may be solved or evaluated on coarser grids. In the elastic deformation integral, if the discrete kernel $K_{i,j}^{hh}$ is smooth compared to the grid size h, the kernel may well be approximated by an interpolation of its values on a coarser grid. Since the elastic kernel is singular, it is never smooth around the singularity. However, it is smooth a sufficient distance away from the singularity. For simplicity, the general idea of the algorithm is discussed for non-singular kernels first. The extension to singular kernels follows in a subsequent section.

E.3.1 Smooth kernels

The coarse grid integral transform is obtained in two steps. First, let $\tilde{K}_{i,j}^{hh}$ be a coarse grid approximation of the discrete kernel $K_{i,j}^{hh}$, defined by:

$$\tilde{K}_{i,j}^{hh} \equiv \left[I\!I_H^h K_{i,\cdot}^{hH} \right]_j, \tag{E.3}$$

where $I\!I_H^h$ is the interpolation matrix. The index on which $I\!I_H^h$ works is denoted by a dot. Hence, in this case, the dot corresponds to an interpolation with respect to the y-direction. The new index appears after the final bracket. The coarse grid kernel $K_{i,J}^{hH}$ is simply the fine grid kernel, evaluated in the coarse grid points. Mathematically, the coarse grid kernel is thus obtained by injection, i.e. $K_{i,J}^{hH} \equiv K_{i,2J}^{hh}$, where J is the coarse grid index for standard coarsening.

Using the interpolated kernel, the integral can be written as:

$$w_i^h = \sum_j \tilde{K}_{i,j}^{hh} u_j^h + \sum_j \left(K_{i,j}^{hh} - \tilde{K}_{i,j}^{hh} \right) u_j^h. \tag{E.4}$$

Since $\tilde{K}_{i,j}^{hh}$ is an interpolation of $K_{i,j}^{hh}$ itself, using only coarse grid points, $K_{i,j}^{hh} - \tilde{K}_{i,j}^{hh}$ vanishes in even points (the points which coincide with coarse grid points). In odd points, the interpolation error $K_{i,j}^{hh} - \tilde{K}_{i,j}^{hh}$ is $h^{2p} K^{2p}(\xi)$, where $2p$ is the interpolation order and $K^{2p}(\xi)$ is the $2p^{th}$ derivative of K in some point $\xi \in S$. Whenever the kernel is smooth compared to u, i.e. if $h^{2p} K^{2p}(\xi)$ is sufficiently small compared to the discretization error $h^{2s} u^{2s}(\zeta)$, the interpolation error may be neglected and the integral can be approximated by:

$$w_i^h \approx \tilde{w}_i^h \equiv \sum_j \tilde{K}_{i,j}^{hh} u_j^h = \sum_j \left[I\!I_H^h K_{i,\cdot}^{hH} \right]_j u_j^h = \sum_J K_{i,J}^{hH} \left[(I\!I_H^h)^T u_\cdot^h \right]_J, \tag{E.5}$$

where $(I\!I_H^h)^T$ is the adjoint (or transpose) of the interpolation matrix, generally referred to as *anterpolation*. Consequently, defining

$$u_J^H \equiv \left[(I\!I_H^h)^T u_\cdot^h \right]_J, \tag{E.6}$$

the fine grid multi-integration can be reduced to a coarse grid integration:

$$\tilde{w}_i^h = \sum_J K_{i,J}^{hH} u_J^H. \tag{E.7}$$

It is noted that, since the interpolation is local, the work associated with the intergrid transfer $(I\!I_H^h)^T$ is only $O(n)$. Therefore, it does not increase the complexity of the algorithm.

Since the smoothness of the kernel $K(x, y)$ is very often the same for both the x- and y-direction, a similar procedure can be applied in the x-direction. Let $\hat{K}_{i,j}^{hh}$ be an approximation to $K_{i,j}^{hh}$ defined according to:

$$\hat{K}_{i,j}^{hh} = \left[I\!I_H^h K_{\cdot,j}^{Hh} \right]_i, \tag{E.8}$$

where again the coarse grid kernel $K_{I,j}^{Hh}$ is the fine grid kernel evaluated in the coarse grid points. Then, w_i^h can be written as:

$$w_i^h = \left[I\!I_H^h w_\cdot^H \right]_i + \sum_j \left(K_{i,j}^{hh} - \hat{K}_{i,j}^{hh} \right) u_j^h, \tag{E.9}$$

where

$$w_I^H \equiv \sum_j K_{I,j}^{H,h} u_j^h. \tag{E.10}$$

Since $\hat{K}_{i,J}^{hH}$ is an interpolation of $K_{i,J}^{hH}$ itself, $K_{i,J}^{hH} - \hat{K}_{i,J}^{hH}$ vanishes in even points and in odd points it is only $O(h^{2p}|K^{2p}|)$. Hence, if the kernel is sufficiently smooth, the interpolation error may be neglected and the integral reduces to:

$$w_i^h \approx \left[I\!I_H^h w_\cdot^H \right]_i . \tag{E.11}$$

The two steps may be combined such that both the summation and the evaluation of the integral transform extends over the even points only. The integral transform in the odd points can subsequently be obtained by interpolation from its value in the even points.

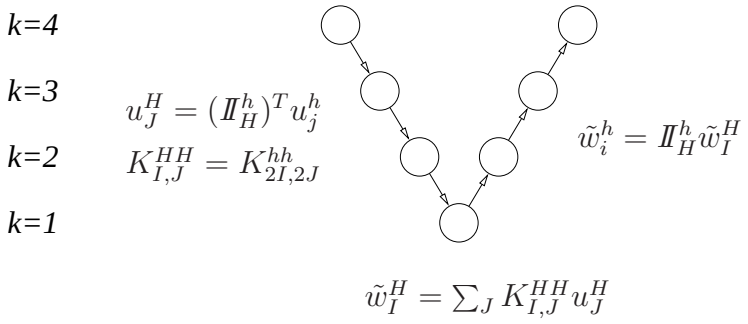


Figure E.1: The multilevel multi-integration cycle.

Obviously, if the kernel is still smooth with respect to the coarse grid, even coarser grids may be used. Indeed, the process can be applied recursively until a grid is obtained for which the number of points is $O(\sqrt{n})$. Then, plain evaluation on this grid only requires $O(n)$ operations and an optimum reduction is achieved. The multilevel algorithm thus obtained is illustrated in Figure E.1.

With respect to the interpolation $I\!I$, it is noted that, since the kernel is well defined outside the domain, central interpolation can, and for smallest errors preferably should, be used near the boundaries of the domain. The order of interpolation should be at least equal to the discretization order but is generally 6^{th} or 8^{th} order.

E.3.2 Singular smooth kernel

In the section above, the kernel was assumed to be smooth in the entire domain. The elastic deformation kernel $K(x, y) = 1/|x - y|$ however, is singular and as a result it is never smooth around $x = y$. On the other hand, for $\|x - y\| \gg h$, or in the discrete summation $\|i - j\| \gg 1$, the kernel is again smooth. (m will be used to denote the discrete distance from which

the kernel is again smooth.) The delimiters $\|$ have been used to indicate that for the two-dimensional elastic deformation kernel, m has two independent components.

Kernels of this type are referred to as singular smooth kernels. The multilevel multi-integration may still be applied for singular smooth kernels, provided one corrects for the large interpolation errors that are introduced near the singularity. Obviously, this can only work if the region which needs to be corrected is sufficiently small.

Consider first the integral transform in the even points, i.e. the points which coincide with the coarse grid points. For these points, Equation E.4 can be replaced by the equality:

$$w_i^h = \sum_j \tilde{K}_{i,j}^{hh} u_j^h + \sum_{\|i-j\|\leq m} (K_{i,j}^{hh} - \tilde{K}_{i,j}^{hh}) u_j^h + \sum_{\|i-j\|>m} (K_{i,j}^{hh} - \tilde{K}_{i,j}^{hh}) u_j^h. \quad \text{(E.12)}$$

Since for $\|i-j\| > m$ the discrete kernel is again smooth, the last term in Equation E.12 may be neglected. (Note that, by definition, $K_{i,j}^{hh} - \tilde{K}_{i,j}^{hh}$ vanishes in the even points.) Reduction of the integral to a coarse grid summation, as described previously for the smooth kernels, yields the approximation:

$$w_i^h \approx \tilde{w}_i^h + \sum_{\|i-j\|\leq m} (K_{i,j}^{hh} - \tilde{K}_{i,j}^{hh}) u_j^h. \quad \text{(E.13)}$$

For the odd points, Equation E.9 can be written as:

$$w_i^h = \left[I\!I_H^h w_\cdot^H \right]_i + \sum_{\|i-j\|\leq m} (K_{i,j}^{hh} - \hat{K}_{i,j}^{hh}) u_j^h + \sum_{\|i-j\|>m} (K_{i,j}^{hh} - \hat{K}_{i,j}^{hh}) u_j^h, \quad \text{(E.14)}$$

which, since for $\|i-j\| \gg m$ the term $K_{i,j}^{hh} - \hat{K}_{i,j}^{hh}$ can be neglected, reduces to:

$$w_i^h = \left[I\!I_H^h w_\cdot^H \right]_i + \sum_{\|i-j\|\leq m} (K_{i,j}^{hh} - \hat{K}_{i,j}^{hh}) u_j^h. \quad \text{(E.15)}$$

Combination of the two individual steps implies that firstly, the integral transform is evaluated on the coarse grid. The values are subsequently corrected according to Equation E.13 and injected in to the fine grid. Secondly, the values in the odd points are obtained by interpolation and the correction according to Equation E.15.

The problem that remains is the determination of the correction size m. For the one-dimensional kernel $\ln(X - X')$, Brandt and Lubrecht [12] showed that $m \sim \ln n$. Hence, the total complexity of the algorithm is $O(n \ln n)$.

E.4 Higher dimensions

For kernels of higher dimension, the algorithm described above can be applied in each direction separately. This ensures that the complexity of the algorithm remains at most $O(n \ln n)$, since the individual components are of this complexity.

For the two-dimensional kernel $1/r$, the reader must realize that the region which needs to be corrected is a two-dimensional patch. Hence, for each coarsening direction, one needs to determine the size of the correction region in the direction of interpolation, m_1, as well as the size perpendicular to this direction, m_2.

For the circular kernel, since the kernel has similar smoothness properties in both the X- and Y-directions, Brandt and Lubrecht showed that m_1 and m_2 are independent of the direction of interpolation. In addition, they showed that the correction size that was determined for the one-dimensional kernel may also be applied for the correction size in the direction of interpolation of the two-dimensional kernel. $m_1 \sim \ln n$ and $m_2 = 2$ proved to be sufficient for interpolation errors which are small compared to the discretization error.

Below, the algorithm is illustrated for the circular kernel. For elliptical kernels, the reader is referred to the main text.

E.5 Circular kernel

Consider the elastic deformation integral:

$$w(x,y) = \frac{2}{\pi^2} \iint_S \frac{u\,(x',y')\,dx'dy'}{\sqrt{(x-x')^2 + (y-y')^2}}. \tag{E.16}$$

For $u(x,y) = \sqrt{1-x^2-y^2}$, the solution in the unit circle $x^2 + y^2 < 1$, is readily obtained as:

$$w(x,y) = 1 - \frac{x^2}{2} - \frac{y^2}{2}. \tag{E.17}$$

To monitor the error in the multilevel multi-integration, an error norm E_k^l is defined according to:

$$E_k^l = h_x h_y \sum_{i=0}^{n_x} \sum_{j=0}^{n_y} |w_{i,j}^{k,l} - w(x_i, y_j)|, \tag{E.18}$$

where $w_{i,j}^{k,l}$ is the discrete transform on level l, where, by the multilevel multi-integration algorithm, the multi-summation has been carried out on level k.

Hence, E_k^l is the average absolute error of the integrals and for $k = l$ the error norm reduces to the L_1 norm of the discretization error. Then, if $E_k^l \approx E_l^l$ at level k for which the number of points is $O(\sqrt{n})$, the error introduced by the multilevel multi-integration is small compared to the discretization error and the total amount of work needed for the multi-summation is $O(n \ln n)$.

Table E.1 shows E_k^l for the discrete transform on a domain $-2 \leq x \leq 2$ and $-2 \leq y \leq 2$ using 6^{th} order interpolation. 8^{th} order interpolation was used in Table E.2. The coarsest level $l = 1$, contains $(4+1)(4+1)$ gridpoints, level $l = 2$ contains $(8 + 1)(8 + 1)$ points, etc. Hence, the mesh sizes h_x and h_y decrease by factor 2 for each level. Furthermore, the correction size in the direction of interpolation in both the X- and Y-directions is, conform Brandt and Lubrecht, $m_1 = 3 + \ln(1/h)$. The correction size perpendicular to the direction of interpolation is $m_2 = 2$.

l	$k = l$	$k = l - 1$	$k = l - 2$	$k = l - 3$	$k = l - 4$	$k = l - 5$
1	$2.86\ 10^{-1}$					
2	$1.58\ 10^{-1}$	$1.58\ 10^{-1}$				
3	$4.68\ 10^{-2}$	$4.69\ 10^{-2}$	$4.70\ 10^{-2}$			
4	$1.56\ 10^{-2}$	$1.53\ 10^{-2}$	$1.53\ 10^{-2}$	$1.54\ 10^{-2}$		
5	$3.07\ 10^{-3}$	$3.00\ 10^{-3}$	$2.95\ 10^{-3}$	$3.02\ 10^{-3}$	$3.13\ 10^{-3}$	
6	$9.50\ 10^{-4}$	$9.19\ 10^{-4}$	$8.72\ 10^{-4}$	$8.41\ 10^{-4}$	$9.19\ 10^{-4}$	$1.03\ 10^{-2}$
7	$2.47\ 10^{-4}$	$2.39\ 10^{-4}$	$2.26\ 10^{-4}$	$2.06\ 10^{-4}$	$1.20\ 10^{-4}$	$3.35\ 10^{-4}$
8	$8.93\ 10^{-5}$	$8.53\ 10^{-5}$	$7.77\ 10^{-5}$	$6.41\ 10^{-5}$	$4.55\ 10^{-5}$	$5.88\ 10^{-5}$

Table E.1: E_k^l-norm for $\kappa = 1.0$ using 6^{th} order transfer.

Table E.2 shows E_k^l using 8^{th} order interpolation.

l	$k = l$	$k = l - 1$	$k = l - 2$	$k = l - 3$	$k = l - 4$	$k = l - 5$
1	$2.86\ 10^{-1}$					
2	$1.58\ 10^{-1}$	$1.58\ 10^{-1}$				
3	$4.68\ 10^{-2}$	$4.69\ 10^{-2}$	$4.71\ 10^{-2}$			
4	$1.56\ 10^{-2}$	$1.66\ 10^{-2}$	$1.70\ 10^{-2}$	$1.73\ 10^{-2}$		
5	$3.07\ 10^{-3}$	$3.14\ 10^{-3}$	$3.19\ 10^{-3}$	$3.17\ 10^{-3}$	$3.22\ 10^{-3}$	
6	$9.50\ 10^{-4}$	$9.67\ 10^{-4}$	$9.91\ 10^{-4}$	$1.00\ 10^{-3}$	$9.92\ 10^{-4}$	$1.04\ 10^{-3}$
7	$2.47\ 10^{-4}$	$2.48\ 10^{-4}$	$2.49\ 10^{-4}$	$2.49\ 10^{-4}$	$2.41\ 10^{-4}$	$2.40\ 10^{-4}$
8	$8.93\ 10^{-5}$	$8.98\ 10^{-5}$	$9.06\ 10^{-5}$	$9.18\ 10^{-5}$	$9.19\ 10^{-5}$	$8.45\ 10^{-5}$

Table E.2: E_k^l-norm for $\kappa = 1.0$ using 8^{th} order transfer.

The Tables show that, indeed, the algorithm introduces interpolation errors which are small compared to the discretization error. Specifically, Table E.1 shows that 6^{th} order interpolation is already sufficient to achieve this goal. The increased accuracy by the 8^{th} order interpolation, and the associated additional work, is thus superfluous.

Bibliography

[1] Abramowitz, M and Stegun, I.A., (1965), "Handbook of Mathematical Functions," New York.

[2] Ai, X., and Cheng, H.S., (1993), "The Influence of Moving Dent on Point EHL Contacts," presented at the 48th Annual Meeting in Calgary, STLE preprint 93-AM-3C-2.

[3] Atkinson, K. E., (1989), "An Introduction to Numerical Analysis (2nd ed.)," John Wiley & Sons, United States, ISBN 0-471-62489-6.

[4] Bair, S. and Winer, W.O., (1979), "A Rheological Model for Elastohydrodynamic Contacts Based Primary Laboratory Data," *Journal of Lubrication Technology*, **101**, 258-265.

[5] Barus, C., (1893), "Isothermals, isopietics and isometrics relative to viscosity," *Am. J. of Science*, **45**, 87-96.

[6] Bathe, K.J., (1982), "Finite Element Procedures in Engineering Analysis," Prentence-Hall Inc.

[7] Baumann, H., Von Frey, D., und Haller, R., (1988), "Druck und Temperaturvertielungen in EHD-kontakten," *Tribologie und Schmierungstechniek*, 35 Jahrgang, **2**, 84-96.

[8] Bayada, G., Chambat, M., El Alaoui, M., (1990), "Variational Forumlations and Finite Element Algorithms for Cavitation Problems," *ASME JOT*, **112**, 398-403.

[9] Bos, J., (1995), "Frictional Heating of Tribological Contacts," PhD Thesis, University of Twente, Enschede, The Netherlands, ISBN 90-9008920-9.

[10] Brandt, A., (1984), "Multigrid Techniques: 1984 Guide with Applications to Fluid Dynamics," available as G.M.D.-Studien No. **85**, from G.M.D.-F1T, Postfach 1240, D-5205, St. Augustin 1, Germany.

[11] Briggs, W.L., (1987), "A Multigrid Tutorial," SIAM, Philadelphia, Pennsylvania, ISBN 0-89871-221-1.

[12] Brandt, A., and Lubrecht, A.A., (1990), "Multilevel Matrix Multiplication and Fast Solution of Integral Equations," *J. of Comp. Phys.*, **90**, **2**, 348-370.

[13] Chang, L., Webster, M.N., and Jackson, A., (1992), "On the Pressure Rippling in Elastohydrodynamic Lubrication of Rough surfaces," *presented at the 1992 ASME/ASLE Tribology conference*, San Diego.

[14] Chevalier, F., (1996), "Modélisation des Conditions d'Alimentation dans les Contacts Élastohydrodynamiques Ponctuels," PhD Thesis, L'Institut National des Sciences Appliquées de Lyon, France.

[15] Chittenden, R.J., Dowson, D., Dunn, J.F., and Taylor, C.M., (1985), "A Theoretical Analysis of the Isothermal Elastohydrodynamic Lubrication of Concentrated Contacts I. Direction of Lubricant Entrainment Coincident with the Major Axis of the Hertzian Contact Ellipse," *Proc. R. Soc. London.*, **A 397**, pp. 245-269.

[16] Chittenden, R.J., Dowson, D., Dunn, J.F., and Taylor, C.M., (1985), "A Theoretical Analysis of the Isothermal Elastohydrodynamic Lubrication of Concentrated Contacts II. General case, with Lubricant Entrainment along Either Principal Axis of the Hertzian Contact Ellipse or at some Intermediate Angle," *Proc. R. Soc. London.*, **A 397**, pp. 271-294.

[17] Dowson, D. and Higginson, G.R., (1966), "Elasto-hydrodynamic Lubrication, The Fundamentals of Roller and Gear Lubrication," Pergamon Press, Oxford, Great Britain.

[18] Dowson, D. and Higginson, G.R., (1959), "A Numerical Solution to the Elastohydrodynamic Problem," *J. Mech. Eng. Sci.*, **1**, 1, pp. 6-15.

[19] Elrod, H.G., (1981), "A Cavitation Algorithm," *ASME J. of Lubrication Technology*, **Vol 103**, **No 3**, 350-354.

[20] Ertel, A.M., (1939), "Hydrodynamic Lubrication Based on New Principles," Akad. Nauj SSSR Prikadnaya Mathematica i Mekhanika, 3, 2, pp. 41-52.

[21] Evans, H.P., and Snidle, R.W., (1981), "Inverse Solution of Reynolds' Equation of Lubrication under Point-contact Elastohydrodynamic Conditions," *ASME JOT*, **103**, pp. 539-546.

[22] Foord, C.A., Wedeven, L.D., Westlake, F.J. and Cameron, A., (1969-1970), "Optical elastohydrodynamics," *Proceedings Inst. Mech. Eng.*, 184 Pt 1 No 28, pp. 487-505.

[23] Floberg, L., (1974), "Cavitation Boundary Conditions with Regard to the Number of Streamers and Tensile Strength of the Liquid," *Cavitation and Related Phenomena in Lubrication, Proceedings of the 1st Leeds-Lyon Symposium on Tribology*, Leeds, pp 31-35.

[24] Grubin, A.N., (1949), "Fundamentals of the Hydrodynamic Theory of Lubrication of Heavily Loaded Cylindrical Surfaces," Central Scientific Research Institute for Technology and Mechanical Engineering, book no. 30, Moscow, D.S.I.R. translation, pp. 115-166.

[25] Guangteng, G., Cann, P.M., Spikes, H.A., and Olver, A.V., (1998), "Mapping of Surface Features in the Thin Film Lubrication Regime," presented at the 25th Leeds-Lyon Symposium on Tribology, 8-11 sept 1998, Lyon.

[26] Hamrock, B.J., Lee, R.T., and Houpert, L.G., (1986), "Parametric Study of Performance in Elastohydrodynamic Lubricated Line Contacts," *Proc. 1986 Leeds-Lyon symposium on Tribology*, Leeds, 199-206.

[27] Hamrock, B.J., Ping Pan, and Rong-Tsong Lee, (1988), "Pressure Spikes in Elastohydrodynamically Lubricated Conjunctions," *ASME JOT*, **110**, 279-284.

[28] Hamrock, B.J., and Dowson, D., (1976), "Isothermal Elastohydrodynamic Lubrication of Point Contacts, part 1-Theoretical Formulation" *ASME JOT*, **98**, pp. 223-229.

[29] Hamrock, B.J., and Dowson, D., (1978), "Elastohydrodynamic Lubrication of Elliptical Contacts for Materials of Low Elastic Modulus 1 - Fully Flooded Conjunction," *ASME JOT*, **100**, 236-245.

[30] Hertz, H., (1881), "On the Contact of Elastic Solids," *J. Reine Angew. Math.*, **92**, 292-302.

[31] Jacobson, B.O., (1991), "Rheology and Elastohydrodynamic Lubrication," *Tribology Series 19*, ISBN 0-444-88146-8.

[32] Jakobsson, B. and Floberg, L., (1957), "The Finite Journal Bearing, Considering Vaporization," *Trans. Chalmers Univ. Tech. Göteborg*, 190.

[33] Johnson, K.L., (1985), "Contact Mechanics," Cambridge University Press.

[34] Johnson, K.L and Tevaarwerk, J.L., (1977), "Shear Behaviour of Elastohydrodynamic Oil Films," Series A, *Proceedings of the Royal Society of London*, **356**, 215-236.

[35] Joppich, W., (1996), "Grundlagen der Mehrgittermethode. Eine Einführung in Standardverfahren," Skriptum zur Vorlesung, Algorithmen I und II, Fachhochschule Köln, Abteilung Gummersbach.

[36] Kaneta, M., (1992), "Effects of Surface Roughness in Elastohydrodynamic Lubrication," *JSME*, III, 35, 4, pp. 535-546.

[37] Kaneta, M., Sakai, T., and Nishikawa, H., (1992), "Optical Interferometric Observations of the Effects of a Bump on Point Contact EHL," *ASME JOT*, 114, pp. 779-784.

[38] Kaneta, M., Sakai, T., and Nishikawa, H., (1993), "Effects of Surface Roughness on Point Contact EHL," *STLE Tribology Transactions*, 36, 4, pp. 605-612.

[39] Kaneta, M., Nishikawa, H., Kameishi, K., (1990), "Observation of Wall Slip in Elastohydrodynamic Lubrication," *ASME JOT*, **112**, 447-452.

[40] Kapitza, P.L., (1955), "Hydrodynamic Theory of Lubrication During Rolling," *ZH.Tekh. Fiz*, **25**, 747-162.

[41] Kostreva, M.M., (1984), "Pressure Stability Considerations in Elasto Hydrodynamic Lubrication Models," *ASME JOT*, **108**, 551-556.

[42] Kweh C.C., Evans, H.P., Snidle, R.W., (1889), "Micro-elastohydrodynamic Lubrication of an Elliptical Contact with Transverse and Three-dimensional Sinusoidal Roughness," *ASME JOT*, **111**, 577-584.

[43] Larsson, R., and Höglund, E., (1995) "Numerical Simulation of a Ball Impacting and Rebounding a Lubricated Surface," *ASME JOT*, **117**, 94-102

[44] Larsson, R., and Lundberg, J., (1995) "Study of Lubricated Impact using Optical Interferometry," *Wear*, **190**, 184-189.

[45] Love, A.E.H., (1944), "A Treatise on the Mathematical Theory of Elasticity," (4^{th} edition), Dover Publications, New York.

[46] Lubrecht, A.A., Breukink, G.A.C., Moes, H., ten Napel, W.E., and Bosma, R., (1987), "Solving Reynold's Equation for EHL Line Contacts by Application of a Multigrid Method," *Proc. Leeds-Lyon symposium on Tribology*, Leeds, 175-182.

[47] Lubrecht, A.A., (1987) "The Numerical Solution of the Elastohydrodynamically Lubricated Line- and Point Contact Problem using Multigrid Techniques," PhD. Thesis, University of Twente, Enschede, ISBN 90-9001583-3.

[48] Lubrecht, A.A., Ioannides, E., (1989), "A Fast Solution of the Dry Contact Problem and the Associated Sub-surface Stress Field, using Multilevel Techniques," *ASME JOT*, **113**, 128-133.

[49] Moes, H., Bosma, R., (1972), "Film Thickness and Traction in EHL at Point Contact," *Inst. Mech. Eng.*, **38**, 149-152.

[50] Moes, H., (1992), "Optimum Similarity Analysis with Applications to Elastohydrodynamic Lubrication," *Wear*, **159**, 57-66.

[51] Moes, H., (1995), "An Advanced Course of Full Film Lubrication," lecture notes code 115531, University of Twente, P.O.Box 217, 7500 AE Enschede, the Netherlands.

[52] Morales Espejel, G.E., (1993), "Elastohydrodynamic Lubrication of Smooth and Rough surfaces," PhD Thesis, University of Cambridge, Department of Engineering.

[53] Newmark, N.M., (1959), "A Method of Computation for Structural Dynamics," *A.S.C.E. J. of Engineering Mechanics Division*, **85**, 76-94.

[54] Nijenbanning, G., Venner, C.H., and Moes, H., (1994), "Film Thickness in Elastohydrodynamically Lubricated Elliptic Contacts," *Wear*, **176**, 217-229.

[55] Nishikawa, H., Handa, K., Teshima, K., Matsuda, K., Kaneta, M., (1995), "Behavior of EHL Films in Cyclic Squeeze Motion," *JSME International Journal*, Series C, **38**, 2, 577-585.

[56] Okamura, H., (1982), "A Contribution to the Numerical Analysis of Isothermal Elastohydrodynamic Lubrication," *Proc. 9th Leeds-Lyon Symp. on Tribology*, Leeds, pp. 313-320.

[57] Osborn, K.F., and Sadeghi, F., (1992), "Time Dependent Line EHD Lubrication Using the Multigrid/Multilevel Technique," *ASME Journal of Tribology*, **114**, 68-74.

[58] Peiran, Y., and Shizhu, W., (1989), "Pure Squeeze Action in an Isothermal Elastohydrodynamically Lubricated Spherical Conjunction. Part 1. Theory and Dynamic Load Results," *Wear*, **142**, 1-16.

[59] Petrusevich, A.I., (1951), "Fundamental Conclusions from the Contact-hydrodynamic Theory of Lubrication," *Izv. Akad. Nauk. SSSR (OTN)*, **2**, 209.

[60] Reynolds, O., (1886), "On the Theory of Lubrication and its Application to Mr. Beachamp Tower's Experiments, Including an Experimental Determination of the Viscosity of Olive Oil," *Phil. Trans. R. Soc.*, **177**, 157-234.

[61] Roelands, C.J.A., (1966), "Correlation Aspects of the Viscosity-Temperature-Pressure Relationship of Lubricating Oils," PhD Thesis, Technische Hogeschool Delft, The Netherlands.

[62] Safa, M.M.A., and Gohar, R., (1986), "Pressure Distribution Under a Ball Impacting a Thin Lubricant Layer," *ASME JOT*, **108**, 372-376.

[63] Stieber, W., (1933), "Das Swimmlager. Hydrodynamische Theorie des Gleitlagers," Berlin.

[64] Stüben, K., and Trottenberg, U., (1984) "Multigrid Methods: Fundamental Algorithms, Model Problem Analysis and Applications," available as G.M.D.-Studien No. **96**, from G.M.D.-F1T, Postfach 1240, D-5205, St. Augustin 1, Germany.

[65] Olsson, K., (1965), "Cavitation in Dynamically Loaded Bearings," *Trans. Chalmers Univ. Tech. Goteborg*, 308.

[66] Venner, C.H., "Multilevel Solution of the EHL Line and Point Contact Problems," PhD Thesis, University of Twente, Enschede, The Netherlands, ISBN 90-9003974-0.

[67] Venner, C.H., and Lubrecht, A.A., (1994), "Numerical Simulation of Transverse Ridge in a Circular EHL Contact Under Rolling/Sliding," *ASME JOT*, 116, pp. 751-761.

[68] Venner, C.H., and Lubrecht, A.A., (1995), "Numerical Analysis of Influence of Waviness on the Film Thickness of a Circular EHL Contact," *ASME JOT*, 118, pp. 153-161.

[69] Venner C.H. and Lubrecht, A.A, (1998), "Amplitude Reduction of Non-Isotropic Harmonic Patterns in Circular EHL Contacts, under Pure Rolling," *Proc. 25th Leeds-Lyon Symposium on Tribology*, Lyon, to appear.

[70] Wedeven, L.D., and Cusano, C., (1979), "Elastohydrodynamic Film Thickness Measurements of Artificially Produced Surface Dents and Grooves," *ASLE Transactions*, 22, pp. 369-381.

[71] Wensing, J.A., and van Nijen, G.C., (1996), "2-Dimensional Computational Model for Vibration Analysis of Waviness in Rolling Bearing Applications," *Proceedings 6th International Conference on Vibrations in Rotating Machinery, IMECHE MEP*, London, pp. 371-382, ISBN 1-86058-009-2.

[72] Wensing, J.A., (1998), "Dynamic Behaviour of Ball Bearings on Vibration Test Spindles," *Proceedings of the 16th International Modal Analysis Conference*, Santa Barbara, California, Vol I, pp.788-794, Society for Experimental Mechanics Inc., Bethel, USA, ISBN 0-912053-59-3.

[73] Wensing, J.A., (1998), "On the Dynamics of Ball Bearings," PhD Thesis, University of Twente, Enschede, The Netherlands.

[74] Whittaker, E.T., and Watson, G.N., (1962), "A Course of Modern Analysis," (4^{th} edition), Cambridge University Press.

[75] Wijnant, Y.H., and Venner, C.H., () "Analysis of an EHL circular contact incorporating rolling element vibration," *Proc. 23th Leeds-Lyon Symp. on Tribology*, Leeds, pp. 445-456.

[76] Wijnant, Y.H., Venner, C.H., Larsson, R., Eriksson, P., (1999), "Effects of Structural Vibrations on the Film Thickness in an EHL Circular Contact," *ASME JOT*, 121, accepted for publication.

[77] Yavneh, I., (1991), "Multigrid Techniques for Incompressible Flows," Ph.D. Thesis, The Weizmann Institute of Science, Rehovot, Israel.

[78] Zeillinger, R., (1995), "Zum Dämpfungsvermögen von Wälzlagern und Wälzlagerverbindungen," Dissertation, Technischen Universität Wien, Fakultät für Maschinenbau, Wien, Österreich.