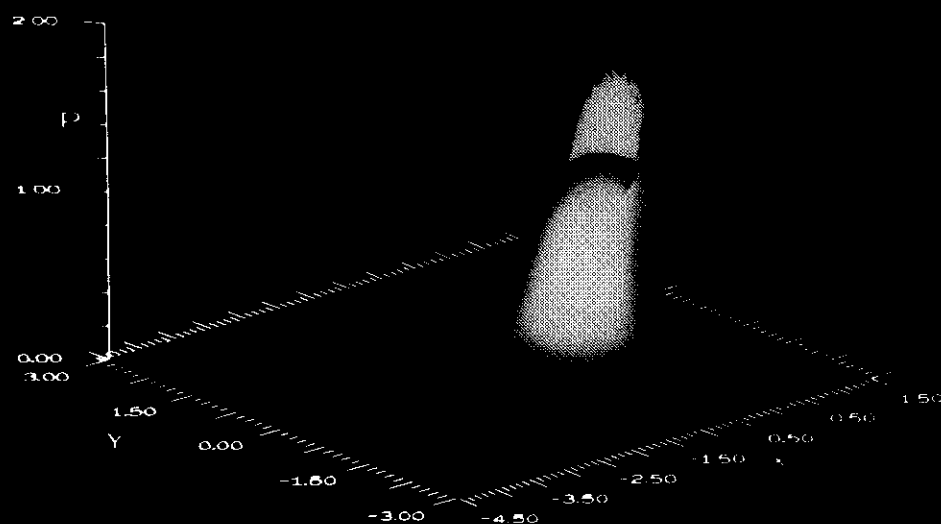


# MULTILEVEL SOLUTION OF THE EHL LINE AND POINT CONTACT PROBLEMS



C.H. VENNER

\_\_\_\_\_

MULTILEVEL SOLUTION OF THE EHL LINE  
AND POINT CONTACT PROBLEMS

C.H. VENNER

CIP-DATA KONINKLIJKE BIBLIOTHEEK, DEN HAAG

Venner, Cornelis Henricus

Multilevel solution of the EHL line and point contact problems /

Cornelis Henricus Venner. - [S.l. : s.n.]. - I11.

Thesis Enschede. - With ref. - With summary in Dutch.

ISBN 90-9003974-0

Subject headings: elastohydrodynamic lubrication/multigrid.

Copyright © 1991 by C.H. Venner, Enschede

# STELLINGEN

behorende bij het proefschrift:

## 'MULTILEVEL SOLUTION OF THE EHL LINE AND POINT CONTACT PROBLEMS'

1. De oorzaak van de instabiliteit van eenvoudige Gauss-Seidel relaxatie toegepast bij het oplossen van elastohydrodynamische smeringsproblemen is dat de (locale) drukveranderingen resulteren in te grote cumulatieve veranderingen van de integralen die de elastische deformatie beschrijven.

Dit proefschrift hoofdstuk 4 en 5

2. Bij de analyse van de invloed van de microgeometrie op de drukverdeling en smeerfilmdikte in elastohydrodynamisch gesmeerde contacten kunnen de veranderingen van deze microgeometrie ten gevolge van de elastische deformatie niet verwaarloosd worden.

Dit proefschrift hoofdstuk 7

3. Extrapolatie van de effecten van oppervlakteoneffenheden op drukverdeling en smeerfilmdikte in een elastohydrodynamisch gesmeerd contact berekend onder stationaire condities, naar situaties waarbij de oneffenheid door het contact beweegt, geeft misleidende resultaten.

Dit proefschrift hoofdstuk 8

4. Door de structuur van de Reynoldsvergelijking is voor veel twee-dimensionale smeringsproblemen, ook wanneer elastische deformatie geen rol speelt, het gebruik van lijnrelaxatie bij de berekening van de drukverdeling in de smeerfilm een noodzakelijke voorwaarde voor een efficient multigrid-algoritme.

5. Juist bij het gebruik van multigrid technieken dient men zich te realiseren dat een zeer nauwkeurige oplossing van de modelvergelijkingen niet per definitie ook een accurate beschrijving van de fysische werkelijkheid is.

6. Wanneer een universiteit onder druk van overheidsbezuinigingen streeft naar een toename van de hoeveelheid door het bedrijfsleven betaald onderzoek ondermijnt zij haar bestaansrecht als dit ontaardt in het vervangen van fundamenteel onderzoek door "marktgericht kluswerk".
7. Het verdient aanbeveling om, in navolging van de waarschuwing op de verpakking van tabaksartikelen, de wervingsadvertenties voor promovendi te voorzien van de tekst:

### Waarschuwing

Dit onderzoek kan schadelijke bijwerkingen hebben  
voor uw sociale contacten en relatie(s).

8. Aangezien *beleid*, *strategie* en *plan* een samenhang tussen maatregelen gericht op het bereiken van een herkenbaar en realiseerbaar doel suggereren, lijkt het verstandig om, ter voorkoming van verwarring, tot nader order deze woorden niet te gebruiken in combinatie met *milieu* en *media*.
9. Een noodzakelijke voorwaarde voor een promotie is de aanvaarding van het risico dat de grootste ontdekkingen misschien pas komen nadat het proefschrift gedrukt is.
10. Door de sterke verontreiniging van bodem en grondwater wordt de struisvogelpoliticus met uitsterven bedreigd.

Enschede, 22 februari 1991

C.H. Venner

---

MULTILEVEL SOLUTION OF THE EHL LINE  
AND POINT CONTACT PROBLEMS

PROEFSCHRIFT

Ter verkrijging van de graad  
van doctor aan de Universiteit Twente,  
op gezag van de rector magnificus,  
prof. dr. ir. J.H.A. de Smit  
volgens besluit van het College van Dekanen  
in het openbaar te verdedigen  
op vrijdag 22 februari 1991 te 16.00 uur

door

Cornelis Henricus VENNER

geboren op 11 januari 1963

te Eindhoven

DIT PROEFSCHRIFT IS GOEDGEKEURD DOOR:

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



# Samenvatting

Het onderwerp van dit proefschrift is elasto-hydrodynamische smering (EHL). Dit is het type smering dat zich bijvoorbeeld voor doet in kogellagers en tussen tandwielen. Meer in detail gaat dit proefschrift over het numeriek oplossen van het elasto-hydrodynamisch gesmeerde lijn en punt contact probleem. Na een korte introductie van het onderwerp wordt een historisch overzicht gegeven van de belangrijkste ontwikkelingen op het werkgebied tot aan de presentatie van de eerste numerieke oplossingen van een EHL contact probleem in 1951. Dit historisch overzicht wordt gevolgd door een overzicht van algoritmes voor het numeriek oplossen van EHL problemen die sindsdien zijn gepresenteerd met een beschrijving van de voor en nadelen van de diverse algoritmes.

Dit overzicht heeft als conclusie dat, ondanks de grote vooruitgang die in de loop der jaren is geboekt, de meeste algoritmes belangrijke beperkingen hebben op het gebied van stabiliteit en/of rekentijd. Door deze beperkingen is het tot dusver niet mogelijk geweest om een aantal problemen die voor de praktijk van groot belang zijn op te lossen. Met name om te kunnen bepalen wat de gevolgen zijn van oppervlakte oneffenheden voor de drukverdeling en de smeerfilmdikte in praktische belastingssituaties zijn snellere en meer stabiele algoritmes nodig.

De ontwikkeling van zulke algoritmes is het onderwerp van het eerste deel van dit proefschrift. Na een beschrijving van een fysisch mathematisch model voor elasto-hydrodynamisch gesmeerde contacten volgt een uitleg van een tweetal zogenaamde multigrid technieken. Deze multigrid technieken maken het in principe mogelijk om de modelvergelijkingen snel en nauwkeurig op te lossen. Vervolgens worden de stabiliteits problemen geanalyseerd en opgelost. Dit resulteert in een algoritme voor zowel het lijn als het puntcontact probleem waarmee, zoals ook aangetoond wordt, zelfs op een mini computer de drukverdeling en filmdikte in het contact met een grote nauwkeurigheid snel kunnen worden berekend.

Het tweede deel van dit proefschrift is toepassingsgericht. Hierin worden onder andere het "standaard" (ideaal gladde oppervlakken) lijn en punt contact probleem uitgebreid beschouwd. Deze problemen zijn opgelost voor een grote variëteit aan belastingcondities, zelfs voor extreem hoge belastingen. Vervolgens wordt aangetoond dat met de ontwikkelde algoritmes bovengenoemde studies naar de effecten van oppervlakte oneffenheden op drukverdeling en filmdikte voor realistische belastingcondities uitgevoerd kunnen worden. Het effect van verschillende soorten oneffenheden wordt besproken voor zowel lijn als puntcontact situaties. Daarbij komen ook dynamische aspecten aan de orde.

De studie naar de invloed van oppervlakte oneffenheden heeft geleid tot enkele interessante nieuwe inzichten zowel met betrekking tot de smering van dit soort contacten in het algemeen als met betrekking tot het falen van een EHL contact ten gevolge van oppervlaktevermoeiing. Als afsluiting van het proefschrift wordt een aantal aanbevelingen gedaan voor toekomstig onderzoek.



# Abstract

This thesis deals with the subject of elastohydrodynamic lubrication (EHL). This is the type of lubrication that for example occurs in roller bearings and between gear wheels. In particular, this work is directed towards the numerical solution of the elastohydrodynamically lubricated line and point contact problems. This thesis starts with an introduction to the subject followed by a brief historical review going from the earliest achievements in the field of fluid film lubrication to the first numerical solutions of an EHL contact problem that were presented in 1951. Subsequently the different algorithms for the numerical solution of EHL contact problems that have been presented since are reviewed and their advantages and drawbacks are discussed.

From this review it is concluded that, in spite of the considerable progress that has been made over the years, most of these algorithms have significant computing time and/or stability limitations. As a consequence a number of problems of great practical interest could not be simulated. For example to accurately study the effects of surface features on the pressure profile and film thickness for practical loading conditions, of essential interest for surface fatigue life predictions, requires faster and more stable solvers.

The development of such solvers for both the line and point contact problems is the subject of the first part of this thesis. Following the explanation of a physical mathematical model describing the lubrication of EHL contacts, first the numerical techniques that basically enable the fast solution of the model equations are outlined. These techniques are usually referred to as “multigrid” or more in general “multilevel” techniques. Subsequently the stability problems are analysed and solved. This results in solvers that enable the solution of the EHL line and point contact problem using large numbers of gridpoints on a small capacity computer.

The second part of this thesis directs attention towards applications. For both the line as well as the point contact situation the “standard” EHL problems, i.e. assuming perfectly smooth surfaces, are solved for a wide range of load conditions, including extremely high loads. Furthermore, it is shown that the newly developed algorithms meet the aforementioned demands by applying them to more complex (realistic) EHL line and point contact situations. In particular, the effects of several surface features upon pressure profile and film thickness for practical (high) loading conditions are investigated for both line and point contact situations. Also transient effects are discussed.

The study of surface features has led to some interesting new insights both in general as well as with respect to the surface fatigue life of EHL contacts. Finally, this thesis is concluded with some recommendations for future research.



# Contents

Samenvatting	v
Abstract	vii
Contents	ix
Nomenclature	xv
<b>1 Introduction</b>	<b>1</b>
1.1 Elastohydrodynamic Lubrication	3
1.2 EHL, a retrospective view	5
1.3 Numerical solution methods	7
1.3.1 Direct methods	8
1.3.1.1 Gauss-Seidel relaxation	8
1.3.1.2 Newton-Raphson system approach	8
1.3.1.3 Multigrid techniques	10
1.3.2 Inverse methods	13
1.4 Variational method	14
1.5 Experimental verification	14
1.6 Outline of this thesis	16
<b>2 EHL Theory</b>	<b>23</b>
2.1 The Reynolds equation	23
2.1.1 Conservation of momentum	24
2.1.2 Conservation of mass	26
2.1.3 Viscosity-pressure	27
2.1.4 Density-pressure	28
2.1.5 Cavitation	28
2.2 The film thickness equation	29
2.2.1 Line contact	29
2.2.2 Point contact	31
2.3 The force balance equation	32
2.4 Dimensionless equations	33

2.4.1	Line contact . . . . .	34
2.4.2	Circular contact . . . . .	35
<b>3</b>	<b>Multilevel techniques</b>	<b>39</b>
3.1	Multilevel solution . . . . .	40
3.1.1	Relaxation . . . . .	40
3.1.2	Correction scheme . . . . .	43
3.1.3	Full approximation scheme . . . . .	44
3.1.4	Intergrid operators . . . . .	45
3.1.5	Coarse grid correction cycle . . . . .	47
3.1.6	Performance . . . . .	49
3.1.7	Full multigrid . . . . .	49
3.1.8	Convergence . . . . .	51
3.1.9	Time dependent problems, the $F$ cycle. . . . .	52
3.1.10	Conclusion . . . . .	54
3.2	Multilevel multi-integration . . . . .	54
3.2.1	Discretization . . . . .	55
3.2.2	General description . . . . .	55
3.2.2.1	Notation . . . . .	56
3.2.2.2	Coarse grid integration . . . . .	56
3.2.3	Coarse grid integration cycle . . . . .	59
3.2.4	Singular smooth kernel . . . . .	59
3.2.5	One dimensional example . . . . .	61
3.2.6	Two dimensional example . . . . .	64
3.2.7	Additional insights . . . . .	67
<b>4</b>	<b>Numerical solution of the line contact problem</b>	<b>71</b>
4.1	Discretization . . . . .	71
4.2	Relaxation . . . . .	73
4.2.1	Linearized problem . . . . .	74
4.2.2	Varying coefficient . . . . .	78
4.2.3	The line contact problem . . . . .	79
4.3	Application of multilevel techniques . . . . .	81
4.3.1	The Reynolds equation . . . . .	81
4.3.2	The film thickness equation . . . . .	81
4.3.3	The force balance equation . . . . .	82
4.3.4	Coarse grid correction cycle . . . . .	83
4.3.5	Full multigrid . . . . .	84
4.3.6	Multilevel multi-integration . . . . .	85
4.4	Example . . . . .	86
4.4.1	Calculational details . . . . .	86
4.4.2	Solution . . . . .	86
4.4.3	Convergence . . . . .	88

---

4.5	Computing times . . . . .	92
4.6	Conclusion . . . . .	94
<b>5</b>	<b>Numerical solution of the circular contact problem</b>	<b>97</b>
5.1	Discretization . . . . .	97
5.2	Relaxation . . . . .	99
5.2.1	Linearized problem . . . . .	100
5.3	Varying coefficients . . . . .	109
5.4	The circular contact problem . . . . .	110
5.5	Application of multilevel techniques . . . . .	111
5.5.1	Full multigrid . . . . .	111
5.5.2	Multilevel multi-integration . . . . .	112
5.6	Example . . . . .	113
5.6.1	Computational details . . . . .	113
5.6.2	Solution . . . . .	113
5.6.3	Convergence . . . . .	116
5.7	Computing times . . . . .	117
5.8	Conclusion . . . . .	119
<b>6</b>	<b>Line contact results: Smooth surfaces</b>	<b>121</b>
6.1	Solutions . . . . .	121
6.1.1	Computational details . . . . .	122
6.1.2	Varying $M$ . . . . .	122
6.1.3	Varying $L$ . . . . .	124
6.1.4	High loads . . . . .	126
6.2	Minimum film thickness . . . . .	127
6.2.1	Asymptotic solutions . . . . .	127
6.2.2	Film thickness formulas . . . . .	128
6.2.3	Film thickness diagram . . . . .	129
6.2.4	Modified film thickness formula . . . . .	130
6.2.5	Comparison of different film thickness formulas . . . . .	131
6.3	The pressure spike . . . . .	136
6.3.1	Onset of the spike . . . . .	137
6.3.2	Spike height . . . . .	138
6.3.3	Comparison of spike height results . . . . .	143
6.4	Conclusion . . . . .	144
<b>7</b>	<b>Line contact results: Surface features</b>	<b>147</b>
7.1	Fatigue life . . . . .	148
7.2	"Large scale" features . . . . .	149
7.2.1	The film thickness equation . . . . .	149
7.2.2	Computational details . . . . .	149
7.2.3	Indentations . . . . .	150

7.2.3.1	Varying wavelength and amplitude . . . . .	152
7.2.3.2	Cavitation . . . . .	154
7.2.3.3	Varying $p_h$ . . . . .	155
7.2.3.4	Varying location . . . . .	158
7.2.3.5	Implications of a dent for the fatigue life . . . . .	160
7.2.4	Bumps . . . . .	161
7.2.4.1	Varying wavelength and amplitude . . . . .	162
7.2.4.2	Varying $p_h$ . . . . .	164
7.2.4.3	Micro EHL . . . . .	166
7.2.4.4	Varying location . . . . .	167
7.2.4.5	Implications of a bump for fatigue life . . . . .	169
7.2.5	Waviness . . . . .	169
7.2.5.1	Varying wavelength and amplitude . . . . .	170
7.2.5.2	Varying $p_h$ . . . . .	172
7.2.5.3	Varying location . . . . .	174
7.3	Surface roughness . . . . .	175
7.3.1	Calculational details . . . . .	177
7.3.2	Variation of speed . . . . .	179
7.3.3	Variation of $p_h$ . . . . .	184
7.4	Conclusion . . . . .	186
<b>8</b>	<b>Simulation of the overrolling of a surface feature</b>	<b>189</b>
8.1	Equations . . . . .	190
8.2	Numerical solution . . . . .	191
8.3	The overrolling of an indentation . . . . .	192
8.3.1	Simulation results . . . . .	193
8.3.2	Sub-surface Stresses . . . . .	199
8.3.3	Condition monitoring . . . . .	201
8.4	The overrolling of a bump . . . . .	202
8.5	The overrolling of waviness . . . . .	208
8.6	Additional insights: sliding . . . . .	212
8.7	Conclusion . . . . .	227
<b>9</b>	<b>Circular contact results: Smooth surfaces</b>	<b>229</b>
9.1	Solutions . . . . .	229
9.1.1	Calculational details . . . . .	230
9.1.2	Varying $M$ and $L$ . . . . .	230
9.2	Film thickness . . . . .	242
9.2.1	Asymptotic solutions . . . . .	243
9.2.2	Film thickness diagrams . . . . .	245
9.2.3	Film thickness formula . . . . .	247
9.3	Conclusion . . . . .	248

---



<b>10 Circular contact results: Surface features</b>	<b>251</b>
10.1 Computational details . . . . .	252
10.2 Smooth surface solution . . . . .	253
10.3 Bumps . . . . .	257
10.4 Indentations . . . . .	262
10.5 Waviness . . . . .	270
10.5.1 Longitudinal . . . . .	270
10.5.2 Transversal . . . . .	276
10.6 Conclusion . . . . .	284
<b>Recommendations for future research</b>	<b>287</b>
<b>Acknowledgement</b>	<b>291</b>
<b>A Dimensionless parameters</b>	<b>293</b>
A.1 Moes and Hertzian dimensionless parameters . . . . .	293
A.1.1 Line contact . . . . .	293
A.1.2 Circular contact . . . . .	296
A.2 Different sets of dimensionless parameters . . . . .	299
A.2.1 Line contact . . . . .	299
A.2.2 Circular contact . . . . .	300
<b>B Elastic deformation</b>	<b>303</b>
B.1 Line contact . . . . .	303
B.2 Circular contact . . . . .	304
<b>C Smoothing rate analysis linearized EHL 1-d problem</b>	<b>307</b>
<b>D Decimation</b>	<b>313</b>
D.1 General description . . . . .	313
D.2 Application to the linearized EHL 2-d problems . . . . .	315



# Nomenclature

$a$	radius Hertzian contact circle (2d)	$a = ((3wR_x)/(2E'))^{1/3}$
$A$	amplitude	
$\mathcal{A}$	dimensionless amplitude	$\mathcal{A} = AR_x/a^2$
$b$	halfwidth Hertzian contact (1d)	$b = \sqrt{(8wR)/(\pi E')}$
$d$	elastic deformation	
$E$	elasticity modulus	
$E'$	reduced modulus of elasticity	$2/E' = (1 - \nu_1^2)/E_1 + (1 - \nu_2^2)/E_2$
$F$	external load (2d)	
$\mathbf{F}$	external force field	
$G$	material parameter	$G = \alpha E'$
$h$	film thickness	
$\bar{h}$	dimensionless film thickness	
	(1d)	$\bar{h} = h/R(2U)^{-1/2}$
	(2d)	$\bar{h} = h/R_x(2U)^{-1/2}$
$H$	dimensionless film thickness	
	(1d)	$H = hR/b^2$
	(2d)	$H = hR_x/a^2$
$h_0$	constant	
$h_{00}$	constant	
$\bar{h}_{00}$	integration constant	
$H_{00}$	integration constant	
$H_{min}$	dimensionless minimum film thickness	(Moes)
	(1d)	$H_{min} = h_{min}/R(2U)^{-1/2}$
	(2d)	$H_{min} = h_{min}/R_x(2U)^{-1/2}$
$H'_m$	dimensionless minimum film thickness	
	(Dowson & Higginson) (1d)	$H'_m = h_{min}/R$
	(Hamrock & Dowson) (2d)	$H'_m = h_{min}/R$
$\Delta H$	distance between lines of constant $H$	
$\mathcal{I}$	integral	
$K$	kernel elastic deformation	
$K_{ij}^{hh}$	discrete approximation of $K$ (1d)	
$K_r^h$	$K_{ij}^{hh}$ for $ i - j  = r$	

$K_{ijkl}^{hhhh}$	discrete approximation of $K$ (2d)	
$\Delta K_{ij}^{hh}$	second order difference of $K_{ij}^{hh}$ (1d)	
$\Delta K_{ijkl}^{hhhh}$	second order difference of $K_{ijkl}^{hhhh}$	
$L$	dimensionless materials parameter	(Moes)
	(1d)	$L = G(2U)^{1/4}$
	(2d)	$L = G(2U)^{1/4}$
$M$	dimensionless load parameter (Moes)	
	(1d)	$M = W(2U)^{-1/2}$
	(2d)	$M = W(2U)^{-3/4}$
$n$	number of nodes on grid	
$\mathbf{n}$	outward normal on surface	
$p$	pressure	
$\bar{p}$	dimensionless pressure	
	(1d)	$\bar{p} = p/E' (2U)^{-1/4}$
	(2d)	$\bar{p} = p/E' (2U)^{-1/4}$
$p_h$	maximum Hertzian pressure	
	(1d)	$p_h = (2w)/(\pi b)$
	(2d)	$p_h = (3F)/(2\pi a^2)$
$P$	dimensionless pressure	$P = p/p_h$
$P_s$	dimensionless spike height (1d)	
$q$	load per unit width	
$r$	radius	
$R$	reduced radius of curvature (1d)	$R^{-1} = R_1^{-1} + R_2^{-1}$
$R_x$	reduced radius of curvature in $x$ direction (2d)	$R_x^{-1} = R_{1x}^{-1} + R_{2x}^{-1}$
$R_y$	reduced radius of curvature in $y$ direction (2d)	$R_y^{-1} = R_{1y}^{-1} + R_{2y}^{-1}$
$S$	surface	
$t$	time	
$T$	dimensionless time	
	(1d)	$T = (u_s t)/(2b)$
	(2d)	$T = (u_s t)/(2a)$
$u$	velocity	
$u_s$	sum velocity	$u_s = u_1 + u_2$
$U$	dimensionless speed parameter	
	(1d)	$U = (\eta_0 u_s)/(2E'R)$
	(2d)	$U = (\eta_0 u_s)/(2E'R_x)$
$\mathbf{U}$	velocity vector	
$(u, v, w)$	velocity components	
$v$	elastic displacement	
$V$	volume in flow	
$w$	wavelength surface feature	
$w$	external load per unit width (1d)	

$W$	dimensionless load parameter (1d) (2d)	$W = w/(E' R)$ $W = w/(E' R_x^2)$
$\mathcal{W}$	dimensionless wavelength surface feature (1d) (2d)	$\mathcal{W} = w/b$ $\mathcal{W} = w/a$
$x$	coordinate	
$\bar{x}$	dimensionless coordinate (1d) (2d)	$\bar{x} = x/R (2U)^{-1/4}$ $\bar{x} = x/R_x (2U)^{-1/4}$
$X$	dimensionless coordinate (1d) (2d)	$X = x/b$ $X = x/a$
$y$	coordinate	
$\bar{y}$	dimensionless coordinate	$\bar{y} = y/R_x (2U)^{-1/4}$
$Y$	dimensionless coordinate	$Y = y/a$
$z$	coordinate	
$z'$	coordinate	
$z$	pressure viscosity parameter (Roelands)	
$\alpha$	pressure viscosity index	
$\bar{\alpha}$	dimensionless parameter	$\bar{\alpha} = \alpha p_h$
$\Delta_x$	mesh size	
$\Delta_t$	time step	
$\epsilon$	coefficient	
$\bar{\Phi}$	dimensionless mass flow per unit width	(1d)
$\bar{\Phi}_{def}$	dimensionless mass flow defect per unit width	(1d)
$\Phi_x, \Phi_y$	pressure flow factors (Patir and Cheng)	
$\Phi_s$	shear flow factor (Patir and Cheng)	
$\eta$	viscosity	
$\eta_o$	viscosity at ambient pressure	
$\bar{\eta}$	dimensionless viscosity	$\bar{\eta} = \eta/\eta_o$
$\lambda$	dimensionless speed parameter (1d) (2d)	$\lambda = (6\eta_o u_s R^2)/(b^3 p_h)$ $\lambda = (6\eta_o u_s R_x^2)/(a^3 p_h)$
$\lambda$	Lamé constant	
$\mu$	Lamé constant	
$\nu$	Poisson's ratio	
$\Omega$	domain	
$\partial\Omega$	boundary of domain	
$\rho$	density	

$\rho_o$	density at ambient pressure	
$\bar{\rho}$	dimensionless density	$\bar{\rho} = \rho/\rho_o$

**Multilevel symbols**

$ERR[h, H]$	convergence norm	
$f$	right hand side	
$d$	dimension of the problem	
$\underline{f}$	right hand side vector	
$I_H^h$	coarse to fine grid interpolation	
$I_H^h$	coarse to fine grid interpolation	
$I_h^H$	fine to coarse grid restriction	
$k$	grid index	
$\mathcal{L}$	differential operator	
$L$	matrix operator approximating $\mathcal{L}$	
$u$	solution	
$\underline{u}$	solution vector of discretized problem	
$\underline{r}$	residual vector	
$\underline{v}$	error vector	
$\mu$	amplitude reduction factor	$\mu =  A/\bar{A} $
$\bar{\mu}$	asymptotic smoothing rate	$\bar{\mu} \equiv \max \mu(\theta)$ for $\pi/2 \leq  \theta  \leq \pi$
$\theta$	error frequency on grid (1d)	
$\bar{\theta}$	error frequency vector (2d)	

**sub-, superscripts**

1,2	body 1,2
x,y	x,y direction
$m$ , or $min$	minimum
$max$	maximum
$i$	grid index (1d)
$i, j$	grid index (2d)
$k, l$	grid index (2d)
$k$	index denoting grid
$h$	grid, fine grid
$H$	coarse grid
$\sim$	approximation
$\wedge$	FAS coarse grid variable
$-$	approximation

Some symbols have more than one meaning. However, generally at the first use in a chapter or section the appropriate meaning is explicitly stated.

# Chapter 1

## Introduction

Friction and wear are two closely related phenomena that play an important role in everyday life. For example, imagine that friction wouldn't exist. That would have a tremendous impact. The slightest force would put objects in eternal motion, and many everyday activities such as simply walking from one spot to another or turning a page of a book, e.g. this thesis, would be very complicated if not impossible. Obviously, it would be hard to live without friction at all. Nevertheless, in many situations friction is a nuisance and should be reduced to a minimum, see for example figure 1.1.

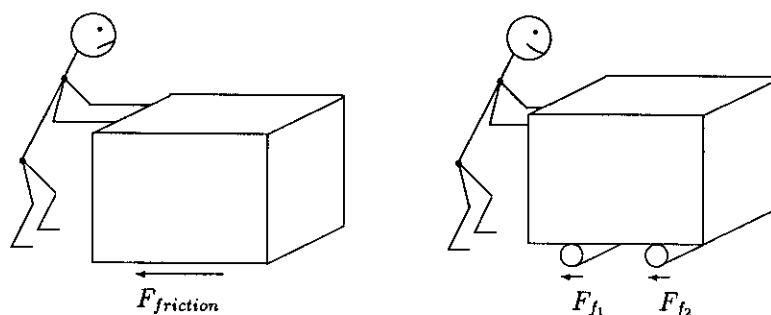


FIGURE 1.1: *An example of a situation where a minimum friction is desired.*  
 $F_{f1} + F_{f2} \ll F_{friction}$

In particular in the design and the operation of technical equipment it is often important to minimize the power loss caused by the friction in the various contacts between machine parts in relative motion, e.g. in bearings and transmissions, in order to obtain a maximum efficiency. Furthermore, wear of the surfaces of these contacting parts should be controlled at a minimum level and if possible avoided at all since uncontrolled wear may lead to an unexpected breakdown with sometimes expensive and maybe even dangerous consequences.

An effective way to reduce friction and wear is lubrication. For example the lubrication of the main bearings of an engine with oil. The geometry of the bearing and the crankshaft and their relative motion enable the formation of a thin oil film separating the bearing from the shaft. If the thickness of this film is large enough to fully separate the two surfaces all shear resulting from the different surface velocities will take place in the lubricant film. Hence, wear will be nearly absent and the coefficient of friction will be small. Typically it is  $O(0.01)$  whereas the coefficient of friction in a dry contact situation is in general  $O(0.1)$ . This type of lubrication that is characterized by the fact that the lubricant film separating the contacting elements is generated by the motion and the geometry of their surfaces is commonly referred to as *Hydrodynamic Lubrication*. If the lubricant film fully separates the surfaces and carries the entire contact load it is also referred to as *full film lubrication*.

In many situations the shape of the lubricant film is exclusively determined by the geometry of the running surfaces. However, in the case of concentrated contacts such as the contacts between gear teeth and between roller (ball) and raceway in rolling element bearings displayed in figure 1.2, this is no longer true.

Because of the high pressures in the contact the shape of the lubricant film is strongly effected by the elastic deformation of the contacting elements. The lubrication of such contacts, commonly referred to as *Elasto-Hydrodynamic Lubrication (EHL)*, is the subject of this thesis. In particular this thesis is directed to the development of a numerical algorithm enabling a fast calculation of the pressure profile and the shape of the lubricant film in elastohydrodynamically lubricated contacts. The algorithm is subsequently employed in the analysis of various contact situations. The presented results are of particular interest with respect to surface fatigue life predictions of EHL contacts.

Before going into detail the remainder of this chapter provides some background with respect to the research contained in this thesis. The next section presents a more detailed introduction to EHL and explains the different types of EHL contacts

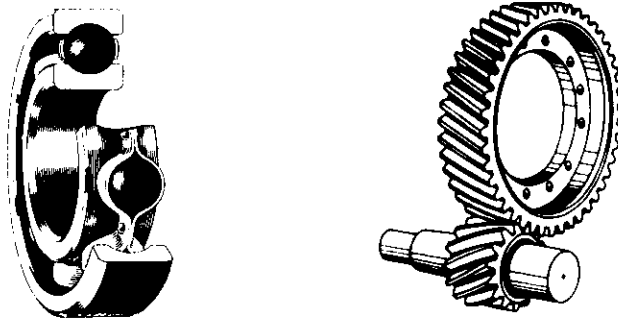


FIGURE 1.2: Two examples of lubricated concentrated contacts.



that are generally distinguished. This introduction is followed by a brief historical review going from the earliest achievements in the field of fluid film lubrication via some milestones in the history of EHL to the first numerical solutions of an EHL contact problem that were presented in 1951. Following this pioneering work many algorithms for the numerical solution of EHL problems have been presented over the last few decades. Section 1.3 presents a survey and discusses characteristics, advantages and drawbacks of the most widely applied algorithms. Finally section 1.5 briefly discusses some experimental techniques that enable verification of the results of numerical calculations and this chapter is concluded with an outline of the research contained in the following chapters.

## 1.1 Elastohydrodynamic Lubrication

Elastohydrodynamic lubrication is defined as the type of hydrodynamic lubrication in which the interaction between the elastic deformation of the contacting elements and the fluid film formation can not be neglected.

Two examples of EHL contacts were already shown in figure 1.2. In these situations, i.e. in contraformal contacts between steel surfaces, the maximum contact pressure can range up to 3.0 GPa. Even in everyday life these contacts play an important role. For example the successful operation of a car depends on at least 200 EHL contacts.

The elastic deformation is not the only characteristic element in EHL studies. The fluid film formation in such contacts is also strongly affected by the lubricant behavior. It is well known that, for most lubricants, the viscosity increases rapidly with increasing pressure and obviously at the high pressures mentioned above this effect can not be neglected. On the contrary, the viscosity-pressure relation and in general the lubricant rheology has become an essential element in the study of EHL contacts.

The EHL theory is not restricted to the aforementioned highly loaded contacts between steel surfaces. It applies to all situations where the stiffness of one or both of the running surfaces is small compared to the pressure in the lubricant film. A typical example is the contact between a steel rod and a rubber seal.

When studying EHL contacts it is not necessary to consider the often rather complex geometry of the contacting machine elements. Since the film thickness and the contact width are generally small compared to the local radius of curvature of the running surfaces, the geometry of the surfaces in the contact area locally can be accurately approximated by paraboloids. This approximation allows a further simplification of the contact geometry. As will be shown, it can be reduced to the contact between a paraboloid and a flat surface.

In general two types of problems are distinguished: the line contact problem and the point contact problem. In the case of the line contact problem the contacting elements are assumed to be infinitely long in one of the principal directions. In fact,

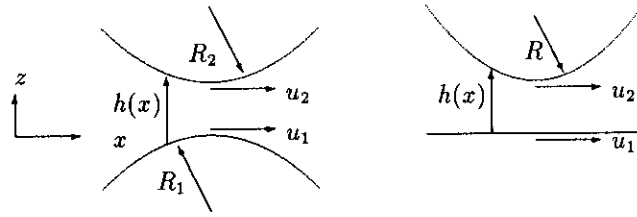


FIGURE 1.3: *The EHL line contact and the reduced geometry employed in the theoretical analysis;*

$R$  = reduced radius of curvature:  $R^{-1} = R_1^{-1} + R_2^{-1}$ .

$h$  = gap width.

the radius of curvature of the paraboloids approximating the surfaces in this direction is infinitely large. In the unloaded dry contact situation the surfaces touch along a straight line. If a load is applied a strip shaped contact region is formed because of the elastic deformations. Figure 1.3 shows the most widely used approximation of the line contact situation: Two parabolically shaped surfaces with local radii of curvature  $R_1$  and  $R_2$  moving with velocities  $u_1$  and  $u_2$  respectively. Figure 1.3 also shows the reduced geometry generally used in the theoretical analysis of this type of contact.

Figure 1.4 shows the most widely used approximation of the more general point contact situation: Two parabolically shaped surfaces with local radii of curvature  $R_{1x}$  and  $R_{2x}$  in the  $x$  direction and  $R_{1y}$  and  $R_{2y}$  in the  $y$  direction. The  $x$  direction is chosen to coincide with the direction of the surface velocities  $u_1$  and  $u_2$ . Figure 1.4 also shows the reduction of this geometry to the contact between a single paraboloid with radii of curvature  $R_x$  and  $R_y$  and a flat surface.  $R_x$  and  $R_y$  are generally referred to as the reduced radii of curvature in  $x$  and  $y$  direction respectively.

In the dry contact situation, i.e. in the absence of a lubricant, both surfaces nominally touch in a point in the unloaded situation whereas the shape of the contact region when a load is applied depends on the ratio of the reduced radii of curvature in  $x$  and  $y$  direction. In general it is an ellipse, and therefore this type of contact is also referred to as an elliptic contact. Examples of point contacts are the contact between a ball and the inner or the outer raceway in a ball bearing and the contact between two crowned disks. A special case of an elliptic contact is the circular contact which occurs if the reduced radii of curvature in both principal directions are equal.

Both line and point contact problem will be addressed in this thesis. With respect to the point contact problem only the circular contact situation is studied,

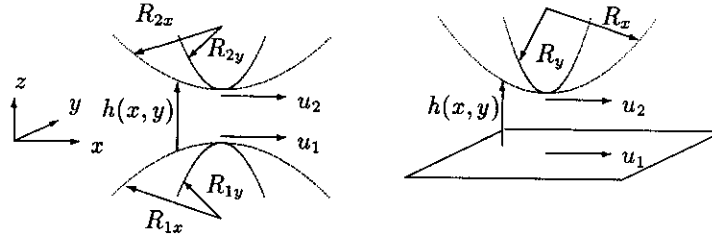


FIGURE 1.4: The EHL point contact and the reduced geometry used in the theoretical analysis;

$$R_x = \text{reduced radius of curvature in } x \text{ direction: } R_x^{-1} = R_{1x}^{-1} + R_{2x}^{-1}$$

$$R_y = \text{reduced radius of curvature in } y \text{ direction: } R_y^{-1} = R_{1y}^{-1} + R_{2y}^{-1}$$

$$h = \text{gap width}$$

because it contains all characteristic elements of 2-D EHL and this work is mainly directed towards the development of algorithms for the solution of the problems.

## 1.2 EHL, a retrospective view

Nowadays it is widely accepted that Beauchamp Tower's [T1] (1883) experimental investigation of friction in lubricated journal bearings, leading to the accidental discovery of substantial pressures in the oil film, laid the foundation for the understanding of fluid film lubrication.

Tower's experimental work provided the stimulus for Reynolds' theoretical analysis [R1] (1886). In his paper Reynolds derives the basic differential equation of fluid film lubrication, nowadays well known as "The Reynolds Equation". This equation relates the pressure in the lubricant film to its geometry and the velocities of the moving surfaces. Reynolds not only derived this differential equation but also presented its approximate solution for restricted situations. Furthermore, he compared his theoretical predictions with Tower's experimental results. Hence, Reynolds was the first who provided the engineers with a mathematical tool for the design of bearings. One of the earliest examples of the application of Reynolds' equation in bearing design was the tilting pad thrust bearing, patented by Michell [M1,M2] (Australia, 1905) and Kingsbury [K1](US, 1910).

Following the successful application of Reynolds' theory to journal and thrust bearings, Martin [M3] (1916) and Gümbel [G1] (1916) focussed the attention on the more complicated problem of gear lubrication. They recognized that the conjunction between the gear teeth could be accurately represented by two cylinders and

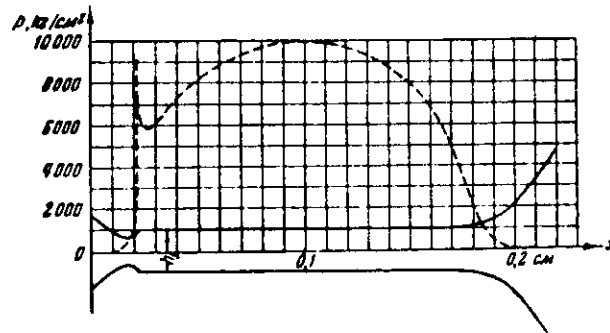


FIGURE 1.5: Pressure profile ( $P$ ) and film thickness ( $H$ ) calculated by Petrusevich [P1].

solved Reynolds' equation for this geometry. Assuming the cylinders to be rigid and the lubricant isoviscous, they deduced a relationship between the minimum film thickness and the operating conditions. However, the film thicknesses predicted by this formula were very small compared to the known surface roughness of gears. Hence, it was concluded that the successful operation of gears almost without wear as observed in practice could not be ascribed to the fluid film action.

It took some 30 years before this contradiction between theory and practical observations was solved and two effects beneficial to fluid film formation in non-conformal contacts were recognized: the elastic deformation of the surfaces due to the high contact pressures and the increase of viscosity with increasing pressure. Ertel [E1] (1939) and Grubin [G2] (1949) were the first to include both effects in the theoretical analysis. From the solution of Reynolds' equation in the inlet region, with the elastic deformations according to the dry contact theory of Hertz [H1] and the linear relation between the logarithm of the viscosity and the pressure proposed by Barus [B1] (1893), they derived an approximate expression for the lubricant film thickness in the center of the contact. This formula yielded film thicknesses which were more than one order of magnitude larger than Martin's predictions for similar conditions. Moreover, it predicted film thickness values that were much larger than the surface roughness in many practical contact situations. Hence, it was confirmed that also in non-conformal contacts separation of the surfaces by a lubricant film can be obtained. Although the formula derived by Ertel and Grubin was, as a result of the different approximations, only valid in a limited parameter regime, their work provided the basis for today's EHL theory.

Shortly afterwards Petrusevich [P1] (1951) presented the first "numerical" solutions of the line contact problem. These solutions simultaneously satisfied both Reynolds' equation and the elasticity equation in a number of discrete positions throughout the conjunction. In addition, the exponential Barus viscosity pressure

relation was used. The results revealed details about the pressure distribution and film shape in the contact region and contain all characteristic features of medium to highly loaded EHL solutions (see figure 1.5): a more or less semi-elliptic pressure distribution, a uniform film in the contact region and a dip in the film thickness preceded by a very local pressure maximum near the outlet. This second maximum in the pressure distribution became the subject of discussions for many years to come and is nowadays generally referred to as “the Petrusevich spike” or “the pressure spike”.

### 1.3 Numerical solution methods

Following the pioneering work of Petrusevich, and stimulated by the introduction of the digital computer and its increasing availability to scientists, many algorithms for the numerical solution of EHL contact problems have been presented over the last few decades. In general the applied mathematical model consists of three equations:

- The Reynolds equation, describing the pressure generation in the lubricant film for a given film thickness.
- The elasticity equation or film thickness equation for the computation of the elastic deformations and the film thickness for a given pressure distribution.
- The force balance equation demanding that the integral over the pressure in the lubricant film balances the externally applied contact load.

The solutions for the pressure ( $P$ ) and the film thickness ( $H$ ) must simultaneously satisfy all three equations. Furthermore, the solution is subject to the condition that all pressures should be larger than, or equal to the vapor pressure of the lubricant. This condition, the so called cavitation condition, is imposed because, for physical reasons, fluids can not be subjected to pressures below this vapor pressure. The fluid will cavitate and the pressure will equal the ambient pressure. Since this vapor pressure is generally already small compared to the atmospheric pressure it certainly will be much smaller than the pressures in an EHL contact and therefore it can be assumed to be zero. Because of this cavitation condition the exit boundary becomes a free boundary.

The algorithms used to solve the set of equations resulting from discretization of the aforementioned constituting equations, are classified into two categories. The direct methods versus the inverse methods. This division is entirely based upon the different way in which Reynolds' equation is treated in the solution process. In both methods the elasticity equation is used only to calculate the elastic deformation for a given pressure distribution. In addition, in both methods the same procedure is followed to solve the force balance equation. Because the latter equation can be solved in a very cheap and straightforward manner, the specific details with respect to this equation are disregarded for the moment.

### 1.3.1 Direct methods

Although the specific iterative procedure applied varies for the different algorithms, they have in common that the Reynolds equation is solved for the pressure given the film thickness. This seems to be the most natural way to solve the system of equations. Indeed, Reynolds' equation was originally derived to obtain the pressure distribution in a conjunction in which the film shape is a priori known.

#### 1.3.1.1 Gauss-Seidel relaxation

One of the most straightforward direct algorithms to solve the set of discrete equations consists of a simple Gauss-Seidel type of relaxation to solve the pressure distribution from Reynolds' equation. After each or after a number of iterations the elastic deformation is recalculated. This sequence is repeated until a pressure profile and film shape have been obtained that simultaneously satisfy all three equations within a desired accuracy. Such an algorithm was used for the solution of the point contact problem by among others Hamrock and Dowson [H2], Chittenden et al. [C1,C2] and Zhu and Cheng [Z1]. Furthermore, Hamrock and Jacobson [H3] applied it to the line contact problem.

This approach has several disadvantages. Asymptotic convergence of the process is generally slow and if  $n$  is the number of points on the grid used, then at least  $O(n)$  relaxations are required to obtain a converged solution. Since the calculation of the elastic deformation in all grid points requires  $O(n^2)$  operations the complexity of the algorithm is at least  $O(n^3)$ . This obviously leads to excessive computing times for large  $n$ , which are unfortunately needed to accurately solve the point contact problem. Moreover, the relaxation process appears to be unstable for highly loaded situations. Even for moderate loads rather large underrelaxation factors were needed to stabilize the relaxation process. Although underrelaxation may stabilize the scheme it unfortunately also reduces the speed of convergence.

Advantages of the method are the straightforward implementation of the cavitation condition and the small storage capacity required.

#### 1.3.1.2 Newton-Raphson system approach

The Newton-Raphson algorithm is based on the linearization of the system of equations around some approximate solution. In this "working point" the Jacobian matrix, i.e. the matrix consisting of the derivatives of all discrete equations with respect to all variables, is evaluated, inverted and used to calculate a new approximation. This process is repeated until a solution simultaneously satisfying all equations within the desired accuracy, is obtained. The basic algorithm is described in detail by Okamura [O1] and with some variations it is still widely used nowadays, see for example Hamrock et al. [H4] and Sadeghi and Sui [S1]. One of the advantages of this algorithm is that in general, convergence close to the solution is fast and that, if it converges, only a few iterations are required. With respect to EHL problems

---

however, it has some severe drawbacks, as was for example pointed out by Lubrecht [L1]:

- Because of the elastic deformations, the Jacobian matrix to be inverted during each iteration, is a full matrix. Since this inversion requires  $O(n^3)$  operations the overall complexity of the algorithm is also  $O(n^3)$ . Hence, accurate solution of point contact problems with large  $n$  will lead to excessive computing times.
- On top of that both the total computer storage needed for the elements of the Jacobian matrix as well as the total computing time needed for their calculation are  $O(n^2)$ .
- The Jacobian matrix becomes almost singular with increasing load, making the solution of highly loaded situations rather cumbersome.
- Because of the simultaneous solution of all pressure updates in each iteration, the implementation of the cavitation condition is rather troublesome. In the case of line contact situations these problems can be avoided simply by scaling the coordinate in the direction of flow, e.g. [B2]. The scaling introduces the location of the exit boundary as a separate unknown to be solved for in the iteration process. Since this location is only a single unknown the scaling doesn't result in a complication of the equations. However, in the two-dimensional case the location of the exit boundary varies over the domain. In that case such a scaling introduces not only a number of unknowns to describe its location but also the derivatives of the exit boundary with respect to both coordinates in the equations. This seriously complicates the problem. Therefore, extension of this approach to the point contact problem seems impossible. The only way out is treating the Reynolds equation with the cavitation condition as a complementarity problem as proposed by Oh [O2] and Kostreva [K2]. That approach however, is rather time consuming and tedious to implement.

Most likely because of these drawbacks the majority of the papers dealing with the application of the global Newton-Raphson method to EHL problems is dedicated to the line contact problem and only a limited number of successful applications of this method to point contact problems have been reported so far.

Over the years improvements to the aforementioned technique have been reported: Houpert and Hamrock [H5] found a way to overcome the load limitation, and recently Chang et al. [C3] reduced the complexity to  $O(n^2)$  by truncating the Jacobian matrix to a tridiagonal matrix. They discarded the terms that reflect the global relation between the film thickness and the pressure caused by the elastic deformation integrals. The truncation allows a more efficient updating of the pressures. However, since these terms become increasingly important for higher loads, it will only increase the difficulties to obtain converged solutions for highly loaded situations.

### 1.3.1.3 Multigrid techniques

The high complexity of the algorithms discussed above, seriously limits the number of nodes that can be used in the analysis, and thus the accuracy that can be obtained. This problem was partly overcome by using local grid refinement techniques and large mainframe or super computers. The benefit of faster hardware however, is limited. The gain in computing time that can be obtained by means of a faster computer is only a constant factor. Consequently, with an algorithm of  $O(n^3)$  complexity, the increase in number of nodes it allows is only proportional to the cubic root of this constant. For example, to solve a point contact problem with half of the original mesh size in the same time, requires a computer that is 64 times faster. Moreover, if a Newton-Raphson algorithm is used it should also have 16 times the storage capacity. Therefore, Lubrecht [L1,L2] shifted the attention to the development of faster software, i.e. developing an alternative solution method of lower complexity.

The presented algorithm was based on Gauss-Seidel relaxation of the Reynolds equation whereas convergence of the process was accelerated using so-called Multigrid techniques. A brief outline of these techniques is presented below. For an introduction to the subject the reader is referred to Briggs [B3]; the application of these techniques to a variety of problems is extensively described by Brandt [B4].

The concept of multilevel (Multigrid) fast solvers is based upon a certain understanding of the convergence behaviour of iterative processes such as simple Gauss-Seidel relaxation. In many situations it can be shown that error components with a wavelength of the order of the mesh size, i.e. high frequency components, are efficiently reduced by the relaxation process.

On the other hand, error components with wavelengths much larger than the mesh size are hardly reduced and consequently, after a few iterations convergence slows down and the asymptotic convergence rate becomes very small. Hence, after a few relaxations the error in the approximation is smooth compared to the mesh size. The remaining error, in fact, can be accurately represented and solved on a coarser grid. Therefore, instead of continuing the relaxation process when, after a few relaxations, convergence slows down, one switches over to a coarser grid for the solution of the error. Once an accurate approximation to the error is obtained on this grid it is used to correct the solution on the fine grid.

Solving the error on a coarser grid generally means solving a similar problem as the original one. So, the same iterative procedure can be applied. However, compared to the computational cost of iterations on the fine grid, the amount of work needed for the solution of the error on the coarse grid is much smaller. The first reason for this reduction is that the number of nodes on the coarse grid is smaller and consequently one iteration on this grid requires less operations. The second reason is that, because of the larger ratio between wavelength and mesh size, the iterative process on the coarse grid converges faster and a given error reduction requires less iterations.

---



The same reasoning applies to the iterative solution of the error on the coarse grid. If the number of nodes on this grid is still relatively large the convergence will slow down again after a few iterations. The remaining error on this grid will be smooth and can be accurately approximated and solved on an even coarser grid. This process can be repeated until a grid is reached on which the problem can be solved in only a few iterations. The result is used to correct the solution of the problem on the next finer grid and so on until the finest grid is reached again. Hence, to solve the problem on some grid, a set of coarser grids is used. On each grid only a few iterations are carried out and only on the coarsest grid the problem is really solved. Usually the number of nodes on the coarsest grid is very small so this requires only a few iterations.

The total number of operations needed to solve a problem up to the level of the discretization error with the aforementioned procedure is generally equivalent to the amount of work of 10 to 25 iterations on the finest grid, independent of the number of nodes on that grid, e.g see [B4].

The complexity of Lubrecht's algorithm was, because of the evaluation of the elastic deformation integrals, still  $O(n^2)$ . Nevertheless, the large reduction in computing time obtained with the Multigrid techniques, enabled him to solve both stationary line and point contact problems with a relatively large number of nodes on a small capacity computer. As a result, extensive parametric studies for both problems could be performed. For the line contact problem this parameter study resulted in an accurate formula to predict the minimum film thickness as a function of the operating conditions. Also the effects of longitudinal and transversal harmonic furrows and ridges and the effects of an isotropic bump on both pressure distribution and film thickness in a circular contact situation were investigated. A characteristic example of a pressure distribution and the associated film thickness in a circular contact calculated by Lubrecht is presented in figure 1.6: A semi-ellipsoid pressure distribution in the central region of the contact. In the inlet region the pressure smoothly builds up to this semi-ellipsoid whereas in the outlet region the pressure is zero. Instead of a single spike the cavitated region is preceded by a spiked region in the pressure profile. The film thickness displays another characteristic element of EHL point contact solutions, i.e. the formation of so called side lobes. The film thickness in these side lobes can be much smaller than the film thickness in the central region of the contact.

Lubrecht's algorithm was not suited to solve the problem for highly loaded situations. As mentioned above, the simple Gauss-Seidel relaxation process is unstable for high loads and, with increasing load, underrelaxation and local relaxation with local film thickness updates were needed to stabilize the relaxation process. In that way, solutions up to a maximum Hertzian pressure of  $\pm 2.0$  GPa were obtained.

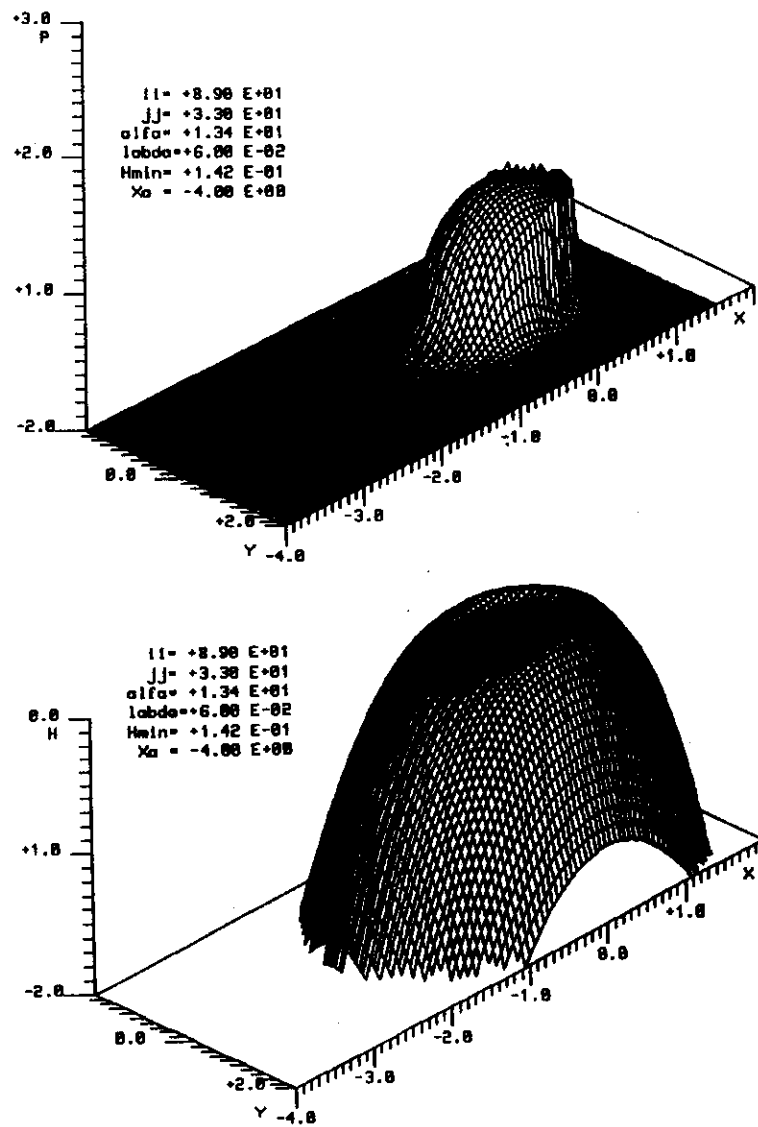


FIGURE 1.6: Pressure profile (P) and film thickness (H) in a circular contact calculated by Lubrecht [L2].

### 1.3.2 Inverse methods

As an alternative, Reynolds' equation can be regarded as an equation for solving the film thickness given a certain pressure profile. This so-called inverse solution technique was introduced by Ertel [E1] in 1939. Driven by stability problems encountered when applying a direct iterative algorithm, Dowson and Higginson [D1] (1959) were the first to develop an algorithm for the numerical solution of the EHL line contact problem based on the inverse solution of Reynolds' equation.

The inverse algorithms are based on the fact that the problem is only solved if the film thickness calculated with the Reynolds' equation, hereafter referred to as the hydrodynamic film thickness, equals the film thickness calculated from the elasticity equation. This leads to the following iterative scheme: For a given approximation of the pressure profile both the hydrodynamic film thickness and the film thickness based on the elastic deflections are calculated. Subsequently, the differences between those two results are used to adjust the pressure profile. This sequence is repeated until the hydrodynamic film thickness deviates less than some prescribed convergence criterion from the elastic deflection film thickness.

Dowson and Higginson [D1] solved the line contact problem for a variety of operating conditions and from the results they derived a formula predicting the minimum film thickness as a function of the operating conditions that is still widely used. Following their work, the inverse approach was extended to point contact problems by Evans and Snidle [E2,E3].

The inverse method has the following disadvantages:

- Solving Reynolds' equation for the film thickness means solving a cubic - equation, with basically three solutions. Since only one of those solutions will be the appropriate one, some problems associated with the selection of the proper root have to be solved.
- Contrary to the direct iterative solvers, the inverse method is only stable for highly loaded situations. Therefore, Evans and Snidle [E2,E3] only used the inverse method in the Hertzian contact region of the conjunction and applied a direct solver in the inlet region.
- The relation used to calculate the pressure changes, given a certain deviation of the hydrodynamic film thickness from the elastic film thickness, is based on experience and insight and its mathematical or physical foundation is not well understood.
- Because the film thickness equation is an integral equation over  $P$  the film thickness is relatively insensitive to local changes in the pressure. Therefore, solution of the problem in the case of rough surfaces may be difficult.

- The complexity of the approach is most likely close to  $O(n^3)$  leading to excessive computing times for large  $n$ .
- The solution process is very sensitive to the accuracy of the first approximation.

An advantage of the inverse method is that implementation of the cavitation condition is, similar to the Gauss-Seidel direct solver, very straightforward and storage requirements are small.

The combined direct-inverse technique of Evans and Snidle was further improved by Kweh et al. [K3]. Circular contact solutions with maximum Hertzian pressures ranging from 1.0 to 4.0 GPa are presented. The same approach was successfully used to investigate the effects of sinusoidal roughness in elliptical contact situations by Kweh et al. [K4] and Barragan de Ling et al. [B5]. Furthermore, a basically similar algorithm was presented by Seabra and Berthe [S2,S3].

In spite of the aforementioned disadvantages this combined inverse-direct approach is the only one so far that has enabled accurate solution of the point contact problem for maximum Hertzian pressures up to 3.3 GPa which is the practical limit in the case of roller bearings. Beyond this value gross plastic deformation of the raceways starts to occur.

## 1.4 Variational method

An alternative approach leading to a solver for the EHL line contact problem was proposed by Verstappen [V1]. This approach is based on variational principles. The resulting algorithm uses an iteration in time to maintain stability with increasing load and has some interesting elements in it. For example the introduction of a preconditioner as a result of which the matrix to be inverted at each time step is no longer a full matrix. It is a pentadiagonal matrix instead. This preconditioning is the more interesting since it is claimed that it can be applied in any solver of the line contact problem based on the inversion of the full matrix. In particular, it should enable a reduction of the complexity of the aforementioned global Newton-Raphson schemes to approximately  $O(n^2)$ .

## 1.5 Experimental verification

The theoretical investigations resulted in the conclusion that even in the case of contacts between non conforming surfaces, complete separation of the running surfaces can be obtained. Furthermore, the dependence of the film shape, pressure distribution and other important properties on the operating conditions was studied. Such results of course, need experimental verification. Because of the high pressures in the contact and the small film thickness, that is not an easy task. Several techniques based on different physical principles have been developed over the years. Most of

---

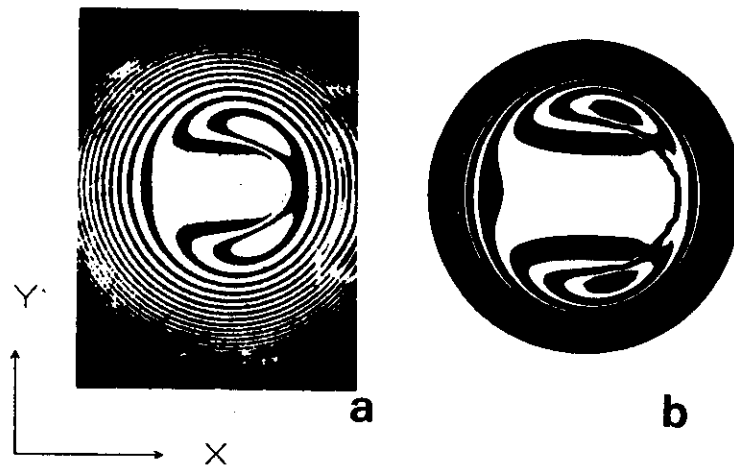


FIGURE 1.7: (a) Interferogram of a circular contact measured by Gohar [G4].  
(b) Film thickness contour plot calculated by Lubrecht [L2] for identical conditions.

these techniques can only be used to measure some average film thickness. However, with the two techniques discussed below also details about the film shape in the contact region can be revealed. The most popular one is the optical interferometry. This technique was for example used by Kirk [K5], Gohar and Cameron [G3], Foord et al. [F1], Koye and Winer [K6], Gohar [G4] and it is still widely used nowadays, for example see Kweh et al. [K3].

As was demonstrated by Lubrecht [L2], see figure 1.7, the optically measured film thickness profiles show good qualitative agreement with the results of numerical calculations. Both results display the characteristic horseshoe shaped film thickness in the central region with the aforementioned side lobes where the overall minimum values occur. In addition, Kweh et al. [K3] reported also a good quantitative agreement between the measured central film thickness and the central film thickness calculated numerically for the same lubricating conditions.

Optical interferometry is limited to film thickness measurements. Information regarding the pressure distribution can not be obtained directly in this way. For that purpose, Safa et al. [S4,S5] and Baumann et al. [B6] performed measurements using micro transducers, vacuum deposited on one of the running surfaces. Depending on the type of transducer, this technique allows measurement of the pressure, the film thickness or the temperature in the conjunction. Because of the high pressures and the high shear rates in the contact these measurements are very difficult to perform and, so far, only results for relatively lightly loaded line contact situations have been presented, see for example figure 1.8. Note that these measurements confirm the occurrence of a second local maximum in the pressure profile, i.e. the "mysterious" spike. However, the transducers are generally not small enough to accurately measure its height.

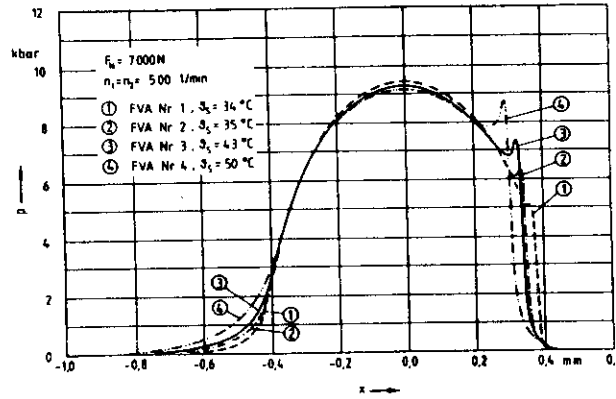


FIGURE 1.8: Pressure profiles measured by Baumann et al. [B6] using micro-transducers.

## 1.6 Outline of this thesis

The accurate prediction of the effect of surface roughness on the fatigue life of EHL contacts in practically relevant situations requires algorithms that allow solution of the line and point contact problems with a large number of nodes in a reasonable computing time, also for highly loaded situations. To give an estimate of what such studies require: for the point contact problem the algorithm must allow solution of the problem with roughly  $500 \times 500$  nodal points.

From the previous sections it is clear that none of the algorithms presented so far is able to accomplish this task in limited times, not even on high speed computers. One might argue that, viewing today's rapid development of faster and faster computers, the next generation of super computers may be fast enough. As was explained in section 1.3.1.3, the high complexity of the algorithms severely limits the increase in number of nodes that a faster computer will allow. In addition, computing time on super computers is generally quite expensive. Therefore, the alternative of developing a solver of lower complexity is adopted in this thesis.

In principle multigrid techniques do have the prospect to solve problems, with the above mentioned number of nodes, in limited computing times on small capacity computers. Therefore, these techniques were adopted as a starting point for further research. The results of this research are presented in the following chapters.

This thesis can be divided into two parts. In the first part, i.e. the chapters 2 to 5 emphasis is on the theory of EHL and the necessary mathematics needed for understanding of the fast numerical solution of the EHL line and point contact problems. Chapter 2 explains the mathematical model describing the isothermal EHL line and point contact problems. The different equations are derived and, for convenience,

they are reworked to a dimensionless form. Next, chapter 3 explains the two multilevel techniques on which fast solvers for both problems will be based. First the "standard" multilevel techniques to accelerate the convergence of a relaxation process, i.e. the procedure briefly outlined in section 1.3.1.3, are explained in detail. Secondly, a recently developed multilevel algorithm *multilevel multi-integration* for the fast numerical evaluation of multi-integrals such as the elastic deformation integrals is explained. In particular, it is demonstrated that, using this algorithm, the computing time needed for the calculation of the elastic deformation in both the line as well as the point contact situation can be reduced from  $O(n^2)$  to  $O(n \ln n)$ . Given a suitable relaxation scheme these two multilevel techniques in principle enable the solution of the entire problem (line or point contact) in  $O(n \ln n)$  operations if  $n$  is the number of nodes on the grid.

With respect to the relaxation scheme, an efficient multilevel solver requires a relaxation process that effectively reduces high frequency error components. For lightly to moderately loaded line and point contact situations this is no problem. As was shown by Lubrecht [L2] a simple Gauss-Seidel relaxation scheme serves well. Unfortunately this scheme, as many other algorithms developed over the past few decades, is unstable for the highly loaded conditions that occur in practical situations. Hence, to obtain solvers of the aforementioned low complexity first alternative relaxation schemes must be developed that remain stable in extreme loading conditions and, in addition provide good error smoothing. This subject is addressed in the chapters 4 and 5.

First chapter 4 describes the step by step development of such a relaxation scheme for the line contact problem with the help of a series of linearized model problems. Subsequently, it is explained how to merge this relaxation process with the multilevel techniques into a fast solver for the problem. The chapter is concluded with the presentation of the results for an example load situation obtained with the solver. In particular, it is demonstrated that, because of its low complexity ( $O(n \ln n)$ ), the algorithm enables solution of the problem with a large number of nodes, e.g.  $O(100.000)$ , on a mini-computer.

Next chapter 5 describes a similar approach leading to a fast solver for the circular contact problem. Again the alternative relaxation scheme is developed with the help of the analysis of a series of linearized model problems. Similar to what is discussed in chapter 4, it is explained how to combine this relaxation process with both the multilevel solution as well as the multilevel multi-integration algorithm into a solver of complexity  $O(n \ln n)$ . By means of an example the low complexity of the solver is demonstrated and particularly it is shown that it is indeed possible to solve the point contact problem with the number of nodes mentioned at the start of this section in reasonable times on a mini-computer.

Having developed "a set of tools" the second part of this thesis directs attention towards applications. The chapters 6 to 10 present the results of the application of the solvers for the line and point contact problem to a variety of EHL contact situations. First the results obtained for line contact problems are presented. In

particular, chapter 6 gives the results of a parametric study of the so-called standard line contact problem, i.e. it discusses the situation in which the surfaces are perfectly smooth. Secondly, chapter 7 addresses a more complex situation, i.e. it studies the effect of surface features. Assuming steady state conditions the effects of some large scale features such as indentations, bumps, and waviness on the pressure profile and film shape are studied for conditions derived from an actual bearing application. In addition to these results for large scale features the effects of a small scale feature, i.e. surface roughness, are studied and pressure profiles and film shapes calculated using the input of an actually measured surface trace are presented.

The analysis of the effects of surface features on the pressure profile and the film shape introduces the need for a time dependent approach since generally the surface feature moves through the contact. This subject is addressed in chapter 8. The solver for the line contact problem described in chapter 4 is extended to transient situations and the results of the simulation of the overrolling of an indentation, a bump and waviness are presented and discussed. These transient studies have resulted in some interesting new insights and demonstrate the importance of such transient studies.

Following these line contact results next the circular contact problem is addressed. First chapter 9 presents the results of the application of the solver developed for this problem to a parametric study of the idealized situation in which the surfaces are perfectly smooth. Secondly chapter 10 presents the results obtained in some more complex situations, i.e. it shows the effects of some surface features on both pressure profile and film thickness.

As a result of covering both the numerical mathematical side, i.e. the theory needed for the understanding of the development of fast solvers for the EHL line and point contact problem, as well as the engineering side, i.e. the application of the solvers to situations of practical interest, this thesis has become quite an extensive work.

However, in spite of all results and answers presented in the chapters 4 to 10 there are still many subjects to be studied, questions that remain unanswered, and new questions to be answered. Therefore this thesis is concluded with a brief outline of some interesting topics for future research.

---



## References

- [B1] Barus, C., 1893, "Isothermals, isopiestic and isometrics relative to viscosity," *Am. J. of Science*, **45**, 87-96.
- [B2] Bisset, E.J., and Glander, D.W., 1988, "A highly accurate approach that resolves the pressure spike of elastohydrodynamic lubrication," *ASME JOT*, **110**, 241-246.
- [B3] Briggs, W.L., 1987, "A Multigrid Tutorial," SIAM, Philadelphia, Pennsylvania, ISBN 0-89871-221-1
- [B4] 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, W. Germany.
- [B5] Barragan de Ling, FdM, Evans, H.P., and Snidle, R.W., 1989, "Micro-elastohydrodynamic lubrication of circumferentially finished rollers: The influence of temperature and roughness," *ASME JOT*, **11**, 730-736.
- [B6] Baumann, H., Von Frey, D., and Haller, R., 1988, "Druck und Temperaturverteilungen in EHD-Kontakten," *Tribologie und Schmierungstechnik*, **35** Jahrgang, **2**, 84-96.
- [C1] 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. Lond.*, **A 397**, 245-269.
- [C2] 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. Lond.*, **A 397**, 271-294.
- [C3] Chang, L., Conry, T.F., and Cusano, C., 1989, "An efficient, robust multi-level computational algorithm for elastohydrodynamic lubrication," *ASME JOT*, **111**, 193-199.
- [D1] Dowson, D., and Higginson, G.R., 1959, "A numerical solution to the elastohydrodynamic problem," *J. of Mech. Eng. Science*, **1**, **1**, 6-15.
- [D2] Dowson, D., and Higginson, G.R., 1966, "Elasto-hydrodynamic lubrication, the fundamentals of roller and gear lubrication," Pergamon Press, Oxford, Great Britain.
- [E1] Ertel, A.M., 1939, "Hydrodynamic lubrication based on new principles," *Akad. Nauk SSSR Prikadnaya Matematika i Mekhanika*, **3**, **2**, 41-52
- [E2] Evans, H.P., and Snidle, R.W., 1981, "Inverse solution of Reynolds' equation of lubrication under point-contact elastohydrodynamic conditions," *ASME JOT*, **103**, 539-546.

- [E3] **Evans, H.P., and Snidle, R.W.**, 1982, "The elastohydrodynamic lubrication of point contacts at heavy loads," *Proc. R. Soc. Lond., A* **382**, 183-199.
  - [F1] **Foord, C.A., Wedeven, L.D., Westlake, F.J., and Cameron, A.**, 1969, "Optical elastohydrodynamics," *Proc. Instn Mech. Engrs*, **184**(1), 487-505.
  - [G1] **Gümbel, L.**, 1916, "Über geschmierte Arbeitsräder," *Z. ges. Turbinenwesen*, **13**, 357.
  - [G2] **Grubin, A.N., and Vinogradova, I.E.**, 1949, "Investigation of the contact of machine components," Kh, F. Ketova (ed.) *Central Scientific Research Institute for Technology and Mechanical Engineering* (Moscow), Book No. 30, (DSIR translation No. 337)
  - [G3] **Gohar, R., and Cameron, A.**, 1963, "Optical measurement of oil film thickness under ehd lubrication," *Nature*, **200**, 458-459.
  - [G4] **Gohar, R.**, "Oil film thickness and rolling friction in elastohydrodynamic point contact," *ASME JOT*, **98**, 223-229.
  - [H1] **Hertz, H.**, 1881, "Über die Berührung fester elastischer Körper," *J. für die reine und angew. Math.*, **92**, 156-171.
  - [H2] **Hamrock, B.J., and Dowson, D.**, 1976, "Isothermal elastohydrodynamic lubrication of point contacts, part 1-Theoretical Formulation," *ASME JOT*, **98**, 223-229.
  - [H3] **Hamrock, B.J., and Jacobson, B.O.**, 1984, "Elastohydrodynamic lubrication of line contacts," *ASLE transactions*, **27**, 4, 275-287.
  - [H4] **Hamrock, B.J., Lee, R.T., and Pan, P.**, 1988, "Pressure spikes in elastohydrodynamically lubricated conjunctions," *ASME JOT*, **110**, 279-284.
  - [H5] **Houpert, L.G., and Hamrock, B.J.**, 1986, "A fast approach for calculating film thicknesses and pressures in elastohydrodynamically lubricated contacts at high loads," *ASME JOT*, **108**, 411-420.
  - [K1] **Kingsbury, A.**, 1910, "Thrust bearings," US Patent No. 947242.
  - [K2] **Kostreva, M., M.**, 1984, "Elasto-hydrodynamic Lubrication: A nonlinear complementarity problem," *Int. Journ. Num. Meth. in Fluids*, **4**, 377-397.
  - [K3] **Kweh, C.C., Evans, H.P., and Snidle, R.W.**, 1989, "Elastohydrodynamic lubrication of heavily loaded circular contacts," *Proc. Instn. Mech. Engrs*, **203**, 133-148.
  - [K4] **Kweh, C.C., Evans, H.P., and Snidle, R.W.**, 1989, "Micro-elastohydrodynamic lubrication of an elliptical contact with transverse and three-dimensional sinusoidal roughness," *ASME JOT*, **111**, 577-583.
  - [K5] **Kirk, M.T.**, 1962, "Hydrodynamic lubrication of perspex," *Nature*, London, **194**, (4832), 965-966.
-

- [K6] Koye, K., and Winer, W.O., 1981, "An experimental evaluation of the Hamrock and Dowson minimum film thickness equation for fully flooded E.H.D. point contacts," *ASME JOT*, **103**, 284-290.
- [L1] Lubrecht, A.A., ten Napel, W.E., and Bosma, R., 1986, "Multigrid, an alternative method for calculating film thickness and pressure profiles in elastohydrodynamically lubricated line contacts," *ASME JOT*, **108**, 551-556.
- [L2] 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.
- [M1] Michell, A.G.M., 1905, "Improvements in thrust and like bearings," British Patent No. 875.
- [M2] Michell, A.G.M., 1905, "The lubrication of plane surfaces," *Z. Math. Phys.*, **52**, 2 123-137.
- [M3] Martin, H.M., 1916, "Lubrication of gear teeth," *Engineering* (London), **102**, 199.
- [O1] Okamura, H., 1982, "A contribution to the numerical analysis of isothermal elastohydrodynamic lubrication," *Proc. 9th Leeds-Lyon Symp. on Tribology, Leeds*, 313-320.
- [O2] Oh, K.P., 1985, "The numerical solution of dynamically loaded elastohydrodynamic contact as a nonlinear complementarity problem," *ASME JOT*, **106**, 88-95.
- [P1] Petrusевич, A.I., 1951, "Fundamental conclusions from the contact-hydrodynamic theory of lubrication," *Izv. Akad. Nauk. SSSR (OTN)*, **2**, 209.
- [R1] Reynolds, O., 1886, "On the theory of lubrication and its application to Mr Beauchamp Tower's experiments, including an experimental determination of the viscosity of olive oil," *Phil. Trans. R. Soc.*, **177**, 157-234.
- [S1] Sadeghi, F., and Sui, P.C., "Compressible elastohydrodynamic lubrication of rough surfaces," *ASME JOT*, **111**, 56-62.
- [S2] Seabra, J., and Berthe, D., 1989, "Elastohydrodynamic point contacts part 1: Formulation and numerical solution," *Wear*, **130**, 301-318.
- [S3] Seabra, J., and Berthe, D., 1989, "Elastohydrodynamic point contacts part 2: Influence of surface speeds, surface waviness and load on the contact behaviour," *Wear*, **130**, 319-335.
- [S4] Safa, M.M.A., 1982, "Elastohydrodynamic studies using thin film transducers," PhD thesis, Dep. of Electrical Engineering, Imperial College of Science and Technology, London.

- [S5] **Safa, M.M.A., Anderson, J.C., and Leather, J.A.**, 1982/1983, "Transducers for pressure, temperature and oil film thickness measurement in bearings," *Sensors and Actuators*, **3**, 119-128.
  - [T1] **Tower, B.**, 1883, "First report on friction experiments (friction of lubricated bearings)," *Proc. Instn mech. Engrs*, Nov. 1883, 632-659.
  - [V1] **Verstappen, R.W.C.P.**, 1989, "Elastohydrodynamic lubrication: a dynamic variation method," *PhD Thesis*, University of Twente, Enschede, The Netherlands, ISBN-90-9003079-4.
  - [Z1] **Zhu, D., and Cheng, H.S.**, 1988, "Effect of surface roughness on the point contact EHL," *ASME JOT*, **110**, 32-37.
-

## Chapter 2

### EHL Theory

In this chapter a mathematical model describing the isothermal EHL contact situation, i.e. the lubricant flow past deformable boundaries, will be discussed. As described before, the model consists of three equations: The Reynolds equation that relates the pressure in the lubricant film to the geometry of the gap and the velocities of the running surfaces. The film thickness equation for the computation of the elastic distortion of the surfaces caused by the pressures in the film and the force balance equation demanding that the integral over the pressure balances the externally applied contact load.

Because of the high pressures in the lubricant film the variation of lubricant properties such as the viscosity and the density with pressure must be accounted for. Hence, the model is completed with some empirical relations describing the variation of the viscosity and the density with pressure.

#### 2.1 The Reynolds equation

Consider the situation displayed in figure 2.1: The flow between two moving surfaces  $z = z_1(x, y, t)$  and  $z = z_2(x, y, t)$  with velocity vectors  $\mathbf{U}_1 = (u_1, v_1, w_1)$  and  $\mathbf{U}_2 = (u_2, v_2, w_2)$  respectively.

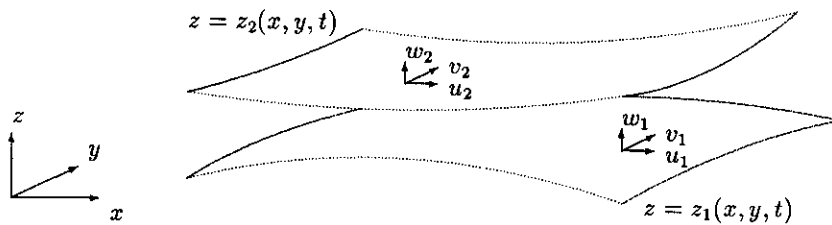


FIGURE 2.1: *fluid flow between two moving surfaces.*

Assuming a narrow gap and a Newtonian lubricant behaviour, the velocities in the flow can be solved from the Navier-Stokes equations and the boundary conditions. Substitution of the results in the equation of continuity, i.e. demanding conservation of mass, yields an equation for the pressure in the film. This equation was derived by Reynolds [R1] and is generally referred to as the Reynolds equation.

### 2.1.1 Conservation of momentum

For a Newtonian fluid the general equation describing the conservation of momentum in the flow is the Navier-Stokes equation, e.g. Milne Thomson [M1]:

$$\rho \frac{D\mathbf{U}}{Dt} = \rho \mathbf{F} - \nabla p - \frac{2}{3} \nabla [\eta \nabla \cdot \mathbf{U}] + 2 [\nabla \cdot (\eta \nabla)] \mathbf{U} + \nabla \times [\eta (\nabla \times \mathbf{U})] \quad (2.1)$$

where:  $p$  = hydrostatic pressure  
 $\eta$  = viscosity  
 $\mathbf{U}$  = velocity vector  
 $\rho$  = density  
 $\mathbf{F}$  = external force field

Assuming the external force field and the mass-inertia terms to be negligible compared to the viscous terms (creeping flow) this equation reduces to:

$$\nabla p = -\frac{2}{3} \nabla [\eta \nabla \cdot \mathbf{U}] + 2 [\nabla \cdot (\eta \nabla)] \mathbf{U} + \nabla \times [\eta (\nabla \times \mathbf{U})] \quad (2.2)$$

In a cartesian coordinate system the equations for the  $x$ ,  $y$ , and  $z$  direction read:

$$\begin{aligned} \frac{\partial p}{\partial x} = & \frac{\partial}{\partial x} \left( \frac{4}{3} \eta \frac{\partial u}{\partial x} - \frac{2}{3} \eta \left( \frac{\partial v}{\partial y} + \frac{\partial w}{\partial z} \right) \right) + \\ & \frac{\partial}{\partial y} \left( \eta \frac{\partial u}{\partial y} + \eta \frac{\partial v}{\partial x} \right) + \\ & \frac{\partial}{\partial z} \left( \eta \frac{\partial u}{\partial z} + \eta \frac{\partial w}{\partial x} \right) \end{aligned} \quad (2.3)$$

$$\begin{aligned} \frac{\partial p}{\partial y} = & \frac{\partial}{\partial x} \left( \eta \frac{\partial v}{\partial x} + \eta \frac{\partial u}{\partial y} \right) + \\ & \frac{\partial}{\partial y} \left( \frac{4}{3} \eta \frac{\partial v}{\partial y} - \frac{2}{3} \eta \left( \frac{\partial u}{\partial x} + \frac{\partial w}{\partial z} \right) \right) + \\ & \frac{\partial}{\partial z} \left( \eta \frac{\partial v}{\partial z} + \eta \frac{\partial w}{\partial y} \right) \end{aligned} \quad (2.4)$$

$$\frac{\partial p}{\partial z} = \frac{\partial}{\partial x} \left( \eta \frac{\partial w}{\partial x} + \eta \frac{\partial u}{\partial z} \right) +$$


---

$$\begin{aligned} & \frac{\partial}{\partial y} \left( \eta \frac{\partial w}{\partial y} + \eta \frac{\partial v}{\partial z} \right) + \\ & \frac{\partial}{\partial z} \left( \frac{4}{3} \eta \frac{\partial w}{\partial z} - \frac{2}{3} \eta \left( \frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} \right) \right) \end{aligned} \quad (2.5)$$

where  $u$ ,  $v$  and  $w$  are the fluid velocities in  $x$ ,  $y$ , and  $z$  direction respectively.

In most lubrication situations the gap is narrow, i.e. if  $h$  is the characteristic size in  $z$  direction and  $L$  is the characteristic size of the domain in both  $x$  and  $y$  directions then  $\epsilon = h/L \ll 1$ . Hence, all derivatives with respect to the  $x$  and  $y$  direction will be much smaller than their equivalents with respect to the  $z$  direction. This narrow gap assumption leads to a significant simplification of the equations. For a detailed analysis the reader is referred to Reynolds [R1] and Langlois [L1]. If only the terms of lowest order in  $\epsilon$  are considered the equations (2.3) to (2.5) reduce to:

$$\frac{\partial p}{\partial x} = \frac{\partial}{\partial z} \left( \eta \frac{\partial u}{\partial z} \right) \quad (2.6)$$

$$\frac{\partial p}{\partial y} = \frac{\partial}{\partial z} \left( \eta \frac{\partial v}{\partial z} \right) \quad (2.7)$$

$$\frac{\partial p}{\partial z} = 0 \quad (2.8)$$

with the boundary conditions:  $U|_{z=z_1} = U_1$   
 $U|_{z=z_2} = U_2$

These boundary conditions are based on the assumption that no slip occurs at the surfaces. Note that, as a result of the narrow gap assumption, the pressure is independent of  $z$ . Consequently, since the fluid flow is assumed to be isothermal, the viscosity will not vary over the height of the gap. Integration of equations (2.6) and (2.7) using the boundary conditions at  $z = z_1$  and  $z = z_2$  renders the following expression for the velocities  $u$  and  $v$ :

$$u = \frac{1}{2\eta} \frac{\partial p}{\partial x} (z^2 - (z_1 + z_2)z + z_1 z_2) + \frac{(u_2 - u_1)}{(z_2 - z_1)} (z - z_1) + u_1 \quad (2.9)$$

$$v = \frac{1}{2\eta} \frac{\partial p}{\partial y} (z^2 - (z_1 + z_2)z + z_1 z_2) + \frac{(v_2 - v_1)}{(z_2 - z_1)} (z - z_1) + v_1 \quad (2.10)$$

Introduction of  $z' = z - z_1$ , i.e.  $0 \leq z' \leq h$ , where  $h = z_2 - z_1$  is the gap width gives:

$$u = \frac{1}{2\eta} \frac{\partial p}{\partial x} (z'^2 - h z') + (u_2 - u_1) \frac{z'}{h} + u_1 \quad (2.11)$$

$$v = \frac{1}{2\eta} \frac{\partial p}{\partial y} (z'^2 - h z') + (v_2 - v_1) \frac{z'}{h} + v_1 \quad (2.12)$$

### 2.1.2 Conservation of mass

The equation for the conservation of mass in the flow is the equation of continuity. In absence of a fluid source the mass balance for a volume  $V$  enclosed by the surface  $S$  entirely in the flow reads, e.g. Batchelor [B1]:

$$\frac{d}{dt} \int_V \rho dV = - \int_S \rho \mathbf{U} \cdot \mathbf{n} dS \quad (2.13)$$

where:  $S$  = surface of volume  
 $\mathbf{n}$  = outwards normal on surface

Considering the volume in the gap shown in figure 2.2 equation (2.13) will be satisfied if:

$$\frac{\partial}{\partial t}(\rho h) + \frac{\partial}{\partial x} \left( \int_0^h (\rho u) dz' \right) + \frac{\partial}{\partial y} \left( \int_0^h (\rho v) dz' \right) = 0 \quad (2.14)$$

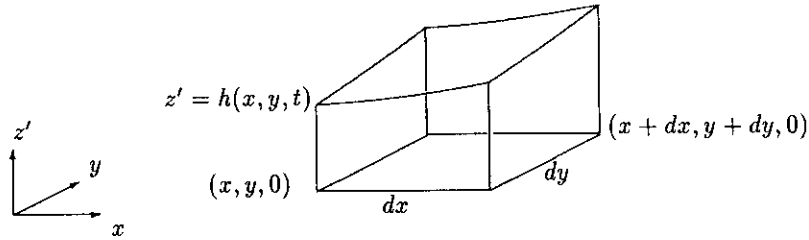


FIGURE 2.2: Control volume in the gap.

Note that in this equation the time rate of change of the volume resulting from the vertical velocity of the running surfaces is accounted for, although the vertical velocities in the flow are neglected because of the narrow gap assumption.

Substitution of  $u$  and  $v$  from equations (2.11) and (2.12) in (2.14) yields the Reynolds equation:

$$\frac{\partial}{\partial x} \left\{ \frac{\rho h^3}{12\eta} \frac{\partial p}{\partial x} - \rho h(u_1 + u_2)/2 \right\} + \frac{\partial}{\partial y} \left\{ \frac{\rho h^3}{12\eta} \frac{\partial p}{\partial y} - \rho h(v_1 + v_2)/2 \right\} - \frac{\partial(\rho h)}{\partial t} = 0 \quad (2.15)$$

In this work it will be assumed that the projections of both surface velocities on a plane  $z = \text{constant}$  are parallel and that the  $x$  axis is chosen to coincide with this direction. Hence,  $v_1 = v_2 = 0$  and equation (2.15) can be written as:



$$\frac{\partial}{\partial x} \left( \frac{\rho h^3}{\eta} \frac{\partial p}{\partial x} \right) + \frac{\partial}{\partial y} \left( \frac{\rho h^3}{\eta} \frac{\partial p}{\partial y} \right) = \underbrace{6u_s \frac{\partial(\rho h)}{\partial x}}_{\text{wedge}} + \underbrace{6\rho h \frac{\partial u_s}{\partial x}}_{\text{stretch}} + \underbrace{12 \frac{\partial(\rho h)}{\partial t}}_{\text{squeeze}} \quad (2.16)$$

where  $u_s = u_1 + u_2$  denotes the sum of the velocities of the running surfaces. The three terms in the right hand side represent three different effects that account for the pressure generation in the lubricant film and are commonly referred to as the *wedge* effect, the *stretch* effect and the *squeeze* effect. In this study  $u_s$  is assumed to be constant, i.e. the stretch effect is not taken into account.

The pressure is subject to the boundary condition that it equals the ambient pressure at the edges of the domain. This pressure is generally defined as zero, i.e. the pressure solved for is in fact the pressure rise from the ambient level. Hence, if the domain is given by:  $\{(x, y) \in \mathbb{R}^2 | x_a \leq x \leq x_b \wedge -y_a \leq y \leq y_a\}$  the boundary conditions are:

$$\begin{aligned} p(x_a, y) &= p(x_b, y) = 0 \\ p(x, -y_a) &= p(x, y_a) = 0 \end{aligned}$$

### 2.1.3 Viscosity-pressure

The increase of viscosity with increasing pressure is one of the two dominant effects accounting for the fluid film formation in EHL contacts. One of the most widely used viscosity pressure relations is the exponential Barus equation, e.g. Barus [B2]:

$$\eta(p) = \eta_0 e^{(\alpha p)} \quad (2.17)$$

where:  $\eta_0$  = the viscosity at ambient pressure

$\alpha$  = the pressure viscosity coefficient, typically  $\alpha \approx 2 \cdot 10^{-8} \text{ Pa}^{-1}$

The advantage of Barus' equation is its simplicity. However, it is accurate for relatively low pressures only. The predicted viscosities for pressures larger than approximately 0.1 GPa are too high, e.g. see Lubrecht [L1].

A more accurate viscosity pressure relation was proposed by Roelands [R2]. The equation is accurate for pressures up to 1 GPa. In S.I. units Roelands' equation reads:

$$\eta(p) = \eta_0 \exp \left\{ (\ln(\eta_0) + 9.67) \left( -1 + \left( 1 + \frac{p}{p_0} \right)^z \right) \right\} \quad (2.18)$$

with:  $z$  = pressure viscosity index, typically  $0.5 \leq z \leq 0.7$ .

In this work  $z=0.68$  will be used

$p_0$  = a constant:  $p_0 = 1.98 \cdot 10^8 \text{ Pa}$

By defining:

$$\alpha = \frac{1}{\eta} \left( \frac{\partial \eta}{\partial p} \right)_{p=0}$$

the Roelands equation can be written as:

$$\eta(p) = \eta_0 \exp \left\{ \frac{\alpha p_0}{z} \left[ -1 + \left( 1 + \frac{p}{p_0} \right)^z \right] \right\} \quad (2.19)$$

$\alpha$ ,  $p_0$  and  $z$  are mutually dependent:

$$\frac{\alpha p_0}{z} = (\ln(\eta_0) + 9.67)$$

#### 2.1.4 Density-pressure

In the analysis of gas lubricated bearings the compressibility of the lubricant is obviously an effect that can not be neglected. If the lubricant is a mineral oil however, the variation of the density with the pressure is usually negligible. In the case of EHL contacts this assumption does not hold. Because of the high pressures the compressibility of the oil must be taken into account. The following relation proposed by Dowson and Higginson [D2] is used throughout this work:

$$\rho(p) = \rho_0 \frac{0.59 \cdot 10^9 + 1.34 p}{0.59 \cdot 10^9 + p} \quad (2.20)$$

where:  $\rho_0$  = density at ambient pressure

This relation limits the compressibility of the lubricant to approximately 30 %. Consequently, the effect of the compressibility of the lubricant on the fluid film formation will be much smaller than for example the effect of the elastic deformation of the surfaces. An alternative relation between density and pressure will be discussed in the recommendations for future research.

#### 2.1.5 Cavitation

Since the lubricant is assumed to be a fluid, pressures lower than the vapour pressure are physically impossible. The fluid will cavitate and the pressure remains constant. Cavitation will for example occur in the outlet region of the conjunction where the gap is widening. This effect is not accounted for in the Reynolds equation presented above. On the contrary, in such regions the equation will allow the pressure to decrease without limit and may predict large negative pressures. The occurrence of cavitation is therefore accounted for separately. Since in most situations the vapour pressure of the lubricant is small compared to the ambient pressure it certainly will be much smaller than the pressures in the contact and therefore the condition is

---

imposed that the pressure should be larger than or equal to zero. As a result of this condition the problem to be solved is a complementarity problem and the outlet boundary of the pressurized region becomes a free boundary.

## 2.2 The film thickness equation

The Reynolds equation can be used to solve the pressure in the gap when its geometry is known. However, because of the elastic deformation of the surfaces the gap width depends on the pressure. In order to solve the problem a relation between pressure and film thickness is required. The film thickness equations for both the line and the point contact problem are presented in this section. With respect to the calculation of the elastic deformations the contacting bodies are considered as semi-infinite elastic bodies. This assumption will be justified as long as the elastic deformations are small compared to the radii of curvature of the undeformed surfaces. Furthermore, it is assumed that the material is isotropic and that the deformations are small enough to justify the application of linear elasticity theory.

### 2.2.1 Line contact

The undeformed surfaces are approximated by parabola's. Hence, the equation for the film thickness reads:

$$h(x) = h_0 + \frac{x^2}{2R} + d(x) \quad (2.21)$$

where:  $R$  = reduced radius of curvature:  $R^{-1} = R_1^{-1} + R_2^{-1}$

$d(x)$  = elastic deformation

$h_0$  = a constant

The elastic deformation  $d(x)$  is obtained by summation of the deformation of both running surfaces caused by the pressure profile. These deflections are calculated in the following manner: First the elastic deformation resulting from the line loading of an elastic half-space is calculated, see figure 2.3.

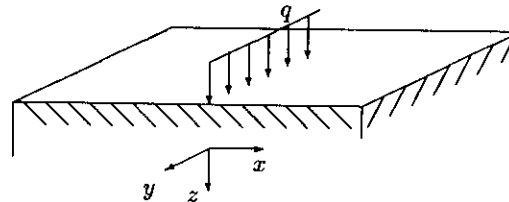


FIGURE 2.3: The line loading of an elastic half space

A line load of intensity  $q$  per unit width distributed along the  $y$  axis acts on the surface  $z = 0$  in the direction normal to this surface. Under plane strain conditions, i.e. no deformation in  $y$  direction, the displacement  $v(x)$  in  $z$  direction of a point of the surface resulting from this load is:

$$v(x) = -\frac{2(1-\nu^2)q}{\pi E} \ln \left| \frac{x}{x_0} \right| \quad (2.22)$$

where:  $q$  = load per unit width  
 $x_0$  = distance where  $v = 0$   
 $E$  = Young's modulus  
 $\nu$  = Poisson's ratio

For details with respect to the analysis the reader is referred to Johnson [J1]. A similar analysis for plane stress conditions was presented by Timoshenko and Goodier [T1].

From equation (2.22) the elastic deformation caused by a distributed load  $p(x)$  can be obtained by integration:

$$v(x) = -\frac{2(1-\nu^2)}{\pi E} \int_{-\infty}^{\infty} \ln \left| \frac{(x-x')}{x_0} \right| p(x') dx' + C \quad (2.23)$$

Hence, the expression for the elastic deformation  $d(x)$  in equation (2.21) reads:

$$d(x) = -\left\{ \frac{2(1-\nu_1^2)}{\pi E_1} + \frac{2(1-\nu_2^2)}{\pi E_2} \right\} \int_{-\infty}^{\infty} \ln \left| \frac{(x-x')}{x_0} \right| p(x') dx' + C \quad (2.24)$$

Note the subscripts indicating that the elastic properties of the two contacting elements are not necessarily the same. With respect to the singularity in the integral it is noted that the so called Cauchy principal value must be taken.

Introduction of the reduced elastic modulus:

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

leads to the following equation for the film thickness:

$$h(x) = h_{00} + \frac{x^2}{2R} - \frac{4}{\pi E'_{-\infty}} \int_{-\infty}^{\infty} \ln \left| \frac{(x-x')}{x_0} \right| p(x') dx' \quad (2.25)$$

where  $h_{00}$  is a constant.

---

### 2.2.2 Point contact

Similar to the line contact situation, see section 2.2.1, the undeformed surfaces are approximated by paraboloids. The expression for the gap width reads:

$$h(x) = h_0 + \frac{x^2}{2R_x} + \frac{y^2}{2R_y} + d(x, y) \quad (2.26)$$

where:  $R_x$  = reduced radius of curvature in  $x$  direction:

$$R_x^{-1} = R_{1x}^{-1} + R_{2x}^{-1}$$

$R_y$  = reduced radius of curvature in  $y$  direction:

$$R_y^{-1} = R_{1y}^{-1} + R_{2y}^{-1}$$

$d(x, y)$  = elastic deformation

The analysis leading to an expression for the elastic deformations is similar to the analysis presented in the previous section for the line contact problem. First the displacements are calculated for a point load acting on an elastic half space, see figure 2.4:

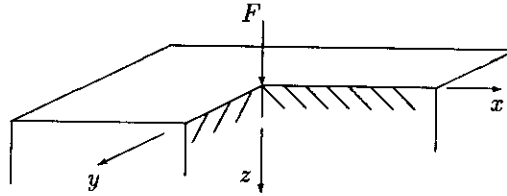


FIGURE 2.4: The point loading of an elastic half space

The displacement  $v(x, y, z)$  of a point  $(x, y, z)$  produced by a concentrated point force  $F$  acting normally to the surface  $z = 0$  at the origin is, e.g. Love [L2]:

$$v(x, y, z) = \frac{F}{4\pi\mu} \frac{z^2}{r^3} + \frac{(\lambda + 2\mu)F}{4\pi\mu(\lambda + \mu)} \frac{1}{r} \quad (2.27)$$

with:  $r = \sqrt{x^2 + y^2 + z^2}$

$\mu, \lambda$  = Lamé constants

The Lamé constants  $\lambda$  and  $\mu$  are related to Poisson's ratio  $\nu$  and Young's modulus  $E$  according to:

$$\lambda = \frac{E\nu}{(1+\nu)(1-2\nu)} \quad \mu = \frac{E}{(1+\nu)} \quad (2.28)$$

Hence, the displacement in  $z$  direction of a point at the surface is given by:

$$v(x, y) = \frac{(1 - \nu^2)F}{\pi E} \frac{1}{\sqrt{(x^2 + y^2)}} \quad (2.29)$$

The elastic deformation of the surface of the elastic half-space caused by a distributed normal load  $p(x, y)$  is obtained by integration:

$$v(x, y) = \frac{(1 - \nu^2)}{\pi E} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{p(x', y') dx' dy'}{\sqrt{(x - x')^2 + (y - y')^2}} + C \quad (2.30)$$

As in the line contact situation, the elastic term in the film thickness equation is obtained by summation of the elastic deformations of the individual surfaces. Consequently the film thickness equation reads:

$$h(x, y) = h_{00} + \frac{x^2}{2R_x} + \frac{y^2}{2R_y} + \frac{2}{\pi E'} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{p(x', y') dx' dy'}{\sqrt{(x - x')^2 + (y - y')^2}} \quad (2.31)$$

### 2.3 The force balance equation

The entire contact load exerted on the contacting elements is carried by the lubricant film. Hence, the integral over the pressure in the film must equal the applied load. This condition is generally referred to as the force balance equation and in the case of the line contact problem it reads:

$$\int_{-\infty}^{\infty} p(x) dx = w \quad (2.32)$$

where:  $w$  = external load per unit width

The equivalent equation for the point contact problem is:

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} p(x, y) dx dy = F \quad (2.33)$$

where:  $F$  = external load

The force balance equation determines the value of the integration constants in the film thickness equations (2.21) and (2.31).

## 2.4 Dimensionless equations

In both line and point contact situations the number of parameters can be significantly reduced by the introduction of suitable dimensionless groups and variables. Different dimensionless groups have been proposed over the years. For specific details with respect to these groups and the relations between them the reader is referred to appendix A. The dimensionless equations presented in this paragraph are mainly based on the Hertzian dry contact parameters. Hertz's theory, see [H1], gives the pressure profile, the geometry of the contact region, and the elastic deformation of the contacting elements in the case of a loaded contact between two parabolically shaped elastic bodies. For the line contact situation the Hertzian pressure profile is given by:

$$p(x) = \begin{cases} p_h \sqrt{1 - (x/b)^2} & \text{if } |x| < b \\ 0 & \text{otherwise} \end{cases} \quad (2.34)$$

$p_h$  is the maximum pressure, generally referred to as the *maximum Hertzian pressure*:

$$p_h = \frac{2w}{\pi b} \quad (2.35)$$

and  $b$  denotes the half width of the region where the bodies are in contact:

$$b = \sqrt{\frac{8wR}{\pi E'}} \quad (2.36)$$

where  $w$  is the external contact load per unit width,  $R$  is the reduced radius of curvature and  $E'$  is the reduced elastic modulus. In the case of a circular contact the Hertzian pressure profile is given by:

$$p(x, y) = \begin{cases} p_h \sqrt{1 - (x/a)^2 - (y/a)^2} & \text{if } |x^2 + y^2| < a^2 \\ 0 & \text{otherwise} \end{cases} \quad (2.37)$$

where  $p_h$  is the maximum pressure:

$$p_h = \frac{3F}{2\pi a^2} \quad (2.38)$$

and  $a$  denotes the radius of the contact circle:

$$a^3 = \frac{3FR_x}{2E'} \quad (2.39)$$

$F$  is the external contact load,  $R_x$  the reduced radius of curvature in  $x$  direction (circular contact  $R_x = R_y$ ) and  $E'$  is the reduced elastic modulus of the contacting bodies.

### 2.4.1 Line contact

If both contacting elements are assumed to be infinitely long in  $y$  direction and the stretch effect is neglected the Reynolds equation (2.16) reads:

$$\frac{\partial}{\partial x} \left( \frac{\rho h^3}{\eta} \frac{\partial p}{\partial x} \right) - 6u_s \frac{\partial(\rho h)}{\partial x} - 12 \frac{\partial(\rho h)}{\partial t} = 0 \quad (2.40)$$

with the cavitation condition  $p \geq 0$  in the domain and  $p = 0$  at the boundaries.

After substitution of the following dimensionless variables:

$$\begin{aligned} \bar{\rho} &= \rho/\rho_0 \\ \bar{\eta} &= \eta/\eta_0 \\ X &= x/b \\ P &= p/p_h \\ H &= hR/b^2 \\ T &= tu_s/(2b) \end{aligned}$$

with  $p_h$  and  $b$  given by (2.35) and (2.36) respectively, this equation reads:

$$\frac{\partial}{\partial X} \left( \frac{\bar{\rho} H^3}{\bar{\eta}} \frac{\partial P}{\partial X} \right) - \lambda \frac{\partial(\bar{\rho} H)}{\partial X} - \lambda \frac{\partial(\bar{\rho} H)}{\partial T} = 0 \quad (2.41)$$

with the cavitation condition  $P \geq 0$  in the domain and  $P = 0$  at the boundaries. The dimensionless parameter  $\lambda$  is given by:

$$\lambda = \frac{6\eta_0 u_s R^2}{b^3 p_h}$$

The dimensionless viscosity  $\bar{\eta}$  is defined according to:

$$\bar{\eta}(P) = \exp(\bar{\alpha} P) \quad (2.42)$$

with  $\bar{\alpha} = \alpha p_h$ , if the Barus equation is used and if Roelands' equation is applied:

$$\bar{\eta}(P) = \exp \left\{ \left( \frac{\alpha p_0}{z} \right) \left( -1 + \left( 1 + \frac{P p_h}{p_0} \right)^z \right) \right\} \quad (2.43)$$

Furthermore, the dimensionless density  $\bar{\rho}$  is unity for an incompressible lubricant and if the lubricant is assumed to be compressible  $\bar{\rho}$  is given by:

$$\bar{\rho}(P) = \frac{0.59 \cdot 10^9 + 1.34 P p_h}{0.59 \cdot 10^9 + P p_h} \quad (2.44)$$

The dimensionless film thickness equation reads:

$$H(X) = H_{00} + \frac{X^2}{2} - \frac{1}{\pi} \int_{-\infty}^{\infty} \ln |X - X'| P(X') dX' \quad (2.45)$$



where  $H_{00}$  is a constant.

The corresponding dimensionless force balance equation is given by:

$$-\infty \int^{\infty} P(X) dX - \frac{\pi}{2} = 0 \quad (2.46)$$

From the dimensionless equations describing the EHL line contact problem it can be observed that, if the lubricant is assumed to be incompressible and if Barus' equation is used the problem reduces to a *two* parameter problem. All solutions can be characterized in terms of the two parameters  $\bar{\alpha}$  and  $\lambda$  (or alternatively the Moes non-dimensional parameters  $M$  and  $L$ , see appendix A). If Roelands' equation is applied and the lubricant is assumed to behave compressible the results can still be presented in terms of these two parameters provided  $\alpha$  and either  $\eta_0$  or  $z$  are given. In this thesis line contact calculations with Roelands equation for a compressible lubricant will be carried out for given values of  $\alpha$  and  $z$ . Unless explicitly stated otherwise,  $\alpha = 1.7 \cdot 10^{-8}$  and  $z = 0.68$  will be used.

### 2.4.2 Circular contact

For the two dimensional circular contact problem the Reynolds equation (2.16) reads:

$$\frac{\partial}{\partial x} \left( \frac{\rho h^3}{\eta} \frac{\partial p}{\partial x} \right) + \frac{\partial}{\partial y} \left( \frac{\rho h^3}{\eta} \frac{\partial p}{\partial y} \right) - 6u_s \frac{\partial(\rho h)}{\partial x} - 12 \frac{\partial(\rho h)}{\partial t} = 0 \quad (2.47)$$

with the cavitation condition  $p \geq 0$  in the domain and  $p = 0$  at the boundaries.

Substitution of the dimensionless variables:

$$\begin{aligned} \bar{\rho} &= \rho/\rho_0 \\ \bar{\eta} &= \eta/\eta_0 \\ X &= x/a \\ Y &= y/a \\ P &= p/p_h \\ H &= hR/a^2 \\ T &= tu_s/(2a) \end{aligned}$$

with  $p_h$  and  $a$  given by (2.38) and (2.39) results in the following dimensionless Reynolds equation:

$$\frac{\partial}{\partial X} \left( \frac{\bar{\rho} H^3}{\bar{\eta}} \frac{\partial P}{\partial X} \right) + \frac{\partial}{\partial Y} \left( \frac{\bar{\rho} H^3}{\bar{\eta}} \frac{\partial P}{\partial Y} \right) - \lambda \frac{\partial(\bar{\rho} H)}{\partial X} - \lambda \frac{\partial(\bar{\rho} H)}{\partial T} = 0 \quad (2.48)$$

with the cavitation condition  $P \geq 0$  in the domain and  $P = 0$  at the boundaries.

The dimensionless parameter  $\lambda$  is given by:

$$\lambda = \frac{6\eta_0 u_s R_x^2}{a^3 p_h}$$

As for the line contact problem the dimensionless viscosity  $\bar{\eta}$  is defined by equation (2.42) if Barus' equation is used and by equation (2.43) if Roelands' equation is applied. Similarly the dimensionless density  $\bar{\rho}$  is unity for an incompressible lubricant and given by equation (2.44) if the lubricant is assumed to behave compressible. Substitution of the same dimensionless parameters in the film thickness equation gives:

$$H(X, Y) = H_{00} + \frac{X^2}{2} + \frac{Y^2}{2} + \frac{2}{\pi^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{P(X', Y') dX' dY'}{\sqrt{(X - X')^2 + (Y - Y')^2}} \quad (2.49)$$

where  $H_{00}$  is a constant.

Finally the corresponding dimensionless force balance equation reads:

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} P(X, Y) dX dY - \frac{2\pi}{3} = 0 \quad (2.50)$$

From the dimensionless equations it follows that, if the lubricant is assumed to be incompressible and if Barus' equation is used, the problem is a two parameter problem. All solutions can be characterized in terms of  $\bar{\alpha}$  and  $\lambda$  (or the Moes dimensionless point contact parameters  $M$  and  $L$ , see appendix A). An approach that still applies if Roelands' equation is used and the lubricant is assumed to behave compressible however, in that case  $\alpha$  and either  $\eta_0$  or  $z$  must be given. In this work calculations with Roelands' equation and a compressible lubricant are performed for given values  $\alpha$  and  $z$ . Fortunately for mineral oils these two parameters do not vary too much. Unless explicitly stated otherwise,  $\alpha = 1.7 \cdot 10^{-8}$  and  $z = 0.68$  will be used.

---

## References

- [B1] **Batchelor, G.K.**, 1967, "An introduction to fluid dynamics," Cambridge University Press.
- [B2] **Barus, C.**, 1893, "Isothermals, isopiestic and isometrics relative to viscosity," *Am. J. of Science*, **45**, 87-96.
- [D2] **Dowson, D., and Higginson, G.R.**, 1966, "Elasto-hydrodynamic lubrication, The fundamentals of roller and gear lubrication," Pergamon Press, Oxford, Great Britain.
- [H1] **Hertz, H.**, 1881, "Über die Berührung fester elastischer Körper," *J. für die reine und angew. Math.*, **92**, 156-171.
- [J1] **Johnson, K.L.**, 1985, "Contact mechanics," Cambridge University Press.
- [L1] **Langlois, W.E.**, 1964, "Slow viscous flow," The Macmillan Company, New York.
- [L2] **Love, A.E.H.**, 1944, "A treatise on the mathematical theory of elasticity," 4<sup>th</sup> edition, Dover publications, New York.
- [M1] **Milne-Thomson, C.B.E.**, 1968, "Theoretical hydrodynamics," 5<sup>th</sup> edition, Macmillan & Co Ltd., London.
- [R1] **Reynolds, O.**, 1886, "On the theory of lubrication and its application to Mr Beauchamp Tower's experiments, including an experimental determination of the viscosity of olive oil," *Phil. Trans. R. Soc.*, **177**, 157-234.
- [R2] **Roelands, C.J.A.**, (1966), "Correlational aspects of the viscosity-temperature-pressure relationship of lubricating oils," Ph.D. Thesis, Technische Hogeschool Delft, The Netherlands, (V.R.B., Groningen, The Netherlands.)
- [T1] **Timoshenko, S.P., and Goodier, J.N.**, 1982, "Theory of elasticity," 3<sup>rd</sup> edition, McGraw-Hill, New York.



## Chapter 3

### Multilevel techniques

Only in some restricted line contact situations can the pressure profile in the film be solved analytically, for example if the bodies are assumed to be rigid and the variation of the viscosity and the density with pressure is disregarded, e.g. Martin [M1] and Gümbel [G1]. However, in general a numerical approach is required. The equations are discretized on a grid covering the contact domain. Subsequently the resulting set of discrete equations is solved using some iterative procedure. A brief discussion of the variety of algorithms for the numerical solution of EHL problems developed in the last few decades, has already been presented in chapter 1. The complexity of most of the solvers is relatively high, i.e.  $O(n^2)$  or  $O(n^3)$ . Consequently the study of features requiring large  $n$  such as the study of the effects of surface roughness on the pressure profile in the film and the consequences for bearing fatigue life becomes rather awkward, if not impossible.

Restricting ourselves to local relaxation processes this high complexity is caused by two effects. Firstly there is slow convergence of the relaxation process for large numbers of nodes. This slowness can be effectively overcome by means of “standard” multilevel techniques, i.e. employing a set of coarser grids to accelerate convergence. These techniques are outlined in detail in section 3.1. An algorithm for the solution of both EHL line and point contact based on these techniques was developed by Lubrecht [L1].

The second reason for the high complexity is the computation of the integrals describing the elastic deformation. Section 3.2 presents a recently developed multilevel technique *multilevel multi-integration* for the fast evaluation of such integrals. This technique has been developed by Brandt and Lubrecht [B1,L2] and as a result the computing time needed for the calculation of the elastic deformation can be reduced from  $O(n^2)$  to  $O(n \ln n)$ , where  $n$  is the number of nodes on the grid.

Together these techniques basically enable the solution of both EHL line and point contact problems in  $O(n \ln n)$  operations. This chapter gives a general description of both techniques. The details regarding the application to the solution of the EHL line and point contact problem are discussed in chapter 4 and 5 respectively.

### 3.1 Multilevel solution

Ever since the development of the first multilevel solvers for elliptical partial differential equations, multilevel techniques have been developed for the solution of a wide variety of scientific problems. For an overview of recent developments in the field the reader is referred to [M2]. The present section explains the basic theory and techniques for the fast solution of an elliptical partial differential problem. All aspects necessary for the understanding of the development of a solver for EHL problems are covered. For additional information on the multilevel solution of partial differential equations the reader is referred to Brandt [B2]. The reader is further referred to the introduction to multilevel solution techniques written by Briggs [B3].

#### 3.1.1 Relaxation

Consider the following problem:

$$\mathcal{L}u = f \quad \text{on } \Omega \quad (3.1)$$

with appropriate boundary conditions on  $\partial\Omega$ .

where:  $\mathcal{L}$  = differential operator  
 $u$  = solution  
 $f$  = right hand side function

Discretization of the problem on a uniform grid characterized by its meshsize  $h$  gives:

$$L^h \underline{u}^h = \underline{f}^h \quad \text{on } \Omega^h \quad (3.2)$$

where:  $L^h$  = discretized differential operator  
 $\underline{u}$  = vector consisting of the solution in all grid points  
 $\underline{f}$  = vector approximating the right hand side function in all points

Assume that this discrete problem is solved using some relaxation process. Starting with a first approximation  $\underline{u}_0^h$  of the solution a new approximation of the solution  $\underline{u}_1^h$  is calculated according to the rule characteristic for the relaxation process under consideration. This sequence is repeated until an approximation of the solution has been obtained that satisfies some prescribed tolerance.

It is characteristic for many relaxation processes that the asymptotic speed of convergence is small. Typically the error reduction per relaxation sweep is  $O(1-h^p)$  where  $p$  is usually 2 or 3. Consequently, for small  $h$  (large number of nodes  $n$ ), many relaxation sweeps are needed to solve the problem up to some fixed accuracy.

Using Fourier analysis techniques it can be shown for simple cases that this slow convergence is due to the fact that, whereas high frequency components in the error are efficiently reduced by the relaxation, it hardly effects low frequency components.

This behaviour is illustrated in the following example, the solution of the 2-D Poisson problem with a simple one-point Gauss-Seidel relaxation.

### Error smoothing: Example

The equation to be solved reads:

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = f \quad (3.3)$$

with some suitable boundary conditions. The usual second order discretization of this equation on a uniform grid with mesh size  $h$  results in the following equation for each interior grid point  $(x_l, y_m)$ :

$$h^{-2}(u_{l-1,m} + u_{l+1,m} + u_{l,m-1} + u_{l,m+1} - 4u_{l,m}) = f_{l,m} \quad (3.4)$$

The approximation of the solution before relaxation is denoted by  $\tilde{u}_{l,m}$ . One relaxation sweep consists of visiting all sites  $(l, m)$  in lexicographic order replacing  $\tilde{u}_{l,m}$  by  $\bar{u}_{l,m}$  to satisfy equation (3.4) in the particular gridpoint. It is characteristic of Gauss-Seidel relaxation that the changes applied in the previously visited sites are taken into account while relaxing a gridpoint. Hence,  $\bar{u}_{l,m}$  is solved from:

$$h^{-2}(\tilde{u}_{l-1,m} + \tilde{u}_{l+1,m} + \bar{u}_{l,m-1} + \bar{u}_{l,m+1} - 4\bar{u}_{l,m}) = f_{l,m} \quad (3.5)$$

The error at site  $(l, m)$  before relaxation is given by:

$$\tilde{v}_{l,m} = u_{l,m} - \tilde{u}_{l,m} \quad (3.6)$$

whereas the error after relaxation at this site is defined as:

$$\bar{v}_{l,m} = u_{l,m} - \bar{u}_{l,m} \quad (3.7)$$

Substitution of  $\tilde{u}$  and  $\bar{u}$  from equations (3.6) and (3.7) in equation (3.5) results in the following equation relating the error after relaxation to the error before relaxation:

$$\bar{v}_{l-1,m} + \bar{v}_{l+1,m} + \bar{v}_{l,m-1} + \bar{v}_{l,m+1} - 4\bar{v}_{l,m} = 0 \quad (3.8)$$

This equation shows that the error in  $(l, m)$  after relaxation is a weighted average of neighboring values of both  $\tilde{v}$  and  $\bar{v}$ . Hence, if the error before relaxation is smooth the error after relaxation must be much smoother.

This smoothing effect can be quantified by means of a so-called "local mode analysis". Because of the local nature of the relaxation process, i.e. the effect of changes applied at a specific site on points several meshsizes away decays exponentially with the distance, the analysis can be restricted to the interior of the domain. The grid is regarded to be embedded in a rectangular domain allowing expansion of both  $\tilde{v}$  and  $\bar{v}$  into Fourier series:

$$\tilde{v}_{l,m} = \sum A(\boldsymbol{\theta}) e^{i(\theta_1 l + \theta_2 m)} \quad (3.9)$$

$$\bar{v}_{l,m} = \sum \bar{A}(\theta) e^{i(\theta_1 l + \theta_2 m)} \quad (3.10)$$

where  $\theta = (\theta_1, \theta_2)$ .

The summations are over a subset of the square  $|\theta| = \max(|\theta_1|, |\theta_2|) \leq \pi$ .

Since the relaxation maps a specific Fourier component on itself, it is sufficient to consider the effect of the relaxation on only one component. From equation (3.8) the following relation between the amplitude of a component before and after relaxation can be derived:

$$(e^{i\theta_1} + e^{i\theta_2}) A(\theta_1, \theta_2) + (e^{-i\theta_1} + e^{-i\theta_2} - 4) \bar{A}(\theta_1, \theta_2) = 0 \quad (3.11)$$

Consequently the *amplification factor* of the  $\theta$  component due to one relaxation sweep is:

$$\mu(\theta) = \left| \frac{\bar{A}(\theta_1, \theta_2)}{A(\theta_1, \theta_2)} \right| = \left| \frac{e^{i\theta_1} + e^{i\theta_2}}{4 - e^{-i\theta_1} - e^{-i\theta_2}} \right| \quad (3.12)$$

This amplification factor as a function of  $(\theta_1, \theta_2)$  is shown in figure 3.1.

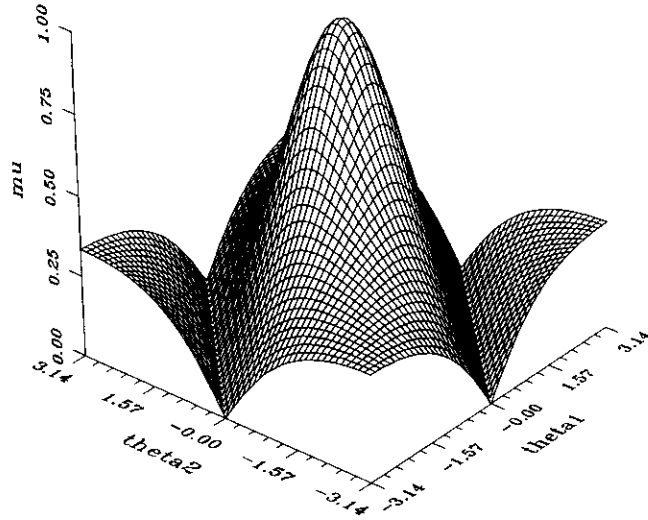


FIGURE 3.1:  $\mu(\theta_1, \theta_2)$  for Gauss-Seidel relaxation on 2-D Poisson problem.

From figure 3.1 it is obvious that, since  $\mu(\theta) \rightarrow 1$  as  $\theta \rightarrow 0$ , the relaxation hardly reduces low frequency components. As a result asymptotic convergence rates are small. In reality, i.e. taking into account the boundary conditions, the asymptotic reduction per relaxation of these components is  $(1 - O(h^2))$ . High frequency components on the other hand are very effectively reduced as is illustrated by the *asymptotic smoothing rate*. This asymptotic smoothing rate  $\bar{\mu}$  is defined as the maximum amplification factor for those error components that can not be represented on a (coarser) grid with, for example, twice the mesh size:



$$\bar{\mu} = \max_{\pi/2 \leq \theta \leq \pi} \mu(\theta_1, \theta_2) \quad (3.13)$$

For the situation considered here, i.e. one point Gauss-Seidel relaxation applied to the solution of the 2-D Poisson problem with the usual 5-point discretization  $\bar{\mu} = 0.5$ . Hence, after only three relaxations the high frequency components are reduced by almost an order of magnitude. Obviously after a few relaxations the error in the solution will mainly consist of components with a wavelength large compared to the mesh size  $h$ .

This completes an example of a local mode analysis. In general such an analysis, also referred to as “smoothing rate analysis”, provides the basis for a multilevel algorithm for the numerical solution of a partial differential problem.

The convergence behaviour shown in the example above is characteristic for many relaxation processes. After a few relaxations the error in the approximation will be smooth compared to mesh size of the grid. Such a smooth error can be approximated accurately and solved on a coarse grid. This is the basic idea behind the multilevel solution technique. Instead of continuing the relaxation process when, after a few relaxations, convergence slows down, one switches over to a coarser grid for the solution of the error. Once an accurate approximation to the error is obtained on this grid it is used to correct the solution on the fine grid.

Two different schemes are discussed. The most straightforward is the so called *Correction Scheme (CS)*. However, this scheme, explained in section 3.1.2, is restricted to linear problems. For non linear problems, e.g. the EHL problem, the *Full Approximation Scheme (FAS)* must be used. This scheme is described in section 3.1.3.

### 3.1.2 Correction scheme

Consider again the solution of the problem:

$$L^h \underline{u}^h = \underline{f}^h \quad \text{on} \quad \Omega^h \quad (3.14)$$

After some relaxations an approximation  $\tilde{\underline{u}}^h$  to the solution  $\underline{u}^h$  has been obtained. The error in this solution is defined as:

$$\underline{v}^h \equiv \underline{u}^h - \tilde{\underline{u}}^h \quad (3.15)$$

Using  $\tilde{\underline{u}}^h$  residuals can be calculated according to:

$$\underline{r}^h = \underline{f}^h - L^h \tilde{\underline{u}}^h \quad (3.16)$$

By definition  $L^h \underline{u}^h = \underline{f}^h$ . Hence, this equation can be written as:

$$\underline{r}^h = L^h \underline{u}^h - L^h \tilde{\underline{u}}^h \quad (3.17)$$

Assuming  $L^h$  to be a linear operator this equation reduces to:

$$L^h(\underline{u}^h - \tilde{\underline{u}}^h) = \underline{r}^h \quad (3.18)$$

Substitution of  $\underline{v}^h$  from equation (3.15) results in an equation for the error:

$$L^h \underline{v}^h = \underline{r}^h \quad (3.19)$$

As was stated in the previous section, after some relaxations, the error  $\underline{v}^h$  will be smooth compared to the mesh size. Hence, it can be represented and solved on a coarser grid. For reasons of simplicity only one coarse grid with mesh size  $H=2h$  is assumed for the moment. A coarse grid approximation  $\underline{v}^H$  to  $\underline{v}^h$  is solved from equation (3.19) on grid  $H$ :

$$L^H \underline{v}^H = I_h^H \underline{r}^h \quad (3.20)$$

$L^H$  is the representation of the matrix operator  $L^h$  on the coarse grid whereas  $I_h^H$  is a restriction operator from the fine grid to the coarse grid. After an approximation  $\tilde{\underline{v}}^H$  to  $\underline{v}^H$  has been calculated it is used to correct the approximation  $\tilde{\underline{u}}^h$  on the fine grid according to:

$$\tilde{\underline{u}}^h = \tilde{\underline{u}}^h + I_H^h \tilde{\underline{v}}^H \quad (3.21)$$

where  $I_H^h$  is an interpolation operator. Because the corrections calculated on the coarse grid are smooth, a linear interpolation (bilinear interpolation if the problem is two dimensional), generally provides sufficient accuracy.

### 3.1.3 Full approximation scheme

If  $L^h$  is a non linear operator the correction scheme is not applicable since the step leading from equation (3.17) to equation (3.18) is not allowed. Therefore, an alternative coarse grid equation for the solution of the error must be derived. This equation is again obtained from equation (3.17):

$$\underline{r}^h = L^h \underline{u}^h - L^h \tilde{\underline{u}}^h \quad (3.22)$$

Substitution of  $\underline{u}^h = \tilde{\underline{u}}^h + \underline{v}^h$  from equation (3.15) gives

$$\underline{r}^h = L^h(\tilde{\underline{u}}^h + \underline{v}^h) - L^h \tilde{\underline{u}}^h \quad (3.23)$$

This equation can be written as:

$$L^h(\tilde{\underline{u}}^h + \underline{v}^h) = L^h \tilde{\underline{u}}^h + \underline{r}^h \quad (3.24)$$

Equation (3.24) is usually referred to as the *FAS* increment equation and will be used to obtain an approximation of the error on the coarse grid. Defining a coarse grid variable:

$$\hat{u}^H = I_h^H \hat{u}^h + v^H \quad (3.25)$$

the coarse grid representation of equation (3.24) reads:

$$L^H \hat{u}^H = \hat{f}^H \quad (3.26)$$

where  $\hat{f}^H$  is referred to as the *FAS* right hand side:

$$\hat{f}^H = L^H(I_h^H \hat{u}^h) + I_h^H r^h \quad (3.27)$$

From an approximation  $\bar{u}^H$  to the solution  $\hat{u}^H$  of equation (3.26) a coarse grid approximation  $\tilde{v}^H$  of the error  $v^h$  can be calculated according to:

$$\tilde{v}^H = \bar{u}^H - I_h^H \hat{u}^h \quad (3.28)$$

This approximation is used to improve the fine grid solution according to:

$$\bar{u}^h = \tilde{u}^h + I_h^h(\bar{u}^H - I_h^H \hat{u}^h) \quad (3.29)$$

### 3.1.4 Intergrid operators

Two different inter-grid operators have been used in the two preceding sections, i.e. a restriction operator  $I_h^H$  and an interpolation operator  $I_h^h$ . If  $v^h$  is a fine grid function then  $I_h^H v^h$  is a coarse grid function. Hence, if  $n_f$  is the number of nodes on the fine grid and  $n_c$  is the number of coarse grid points the restriction operator  $I_h^H$  can be written as a  $n_c \times n_f$  matrix. The most straightforward restriction operator is denoted as *injection*. The value of  $I_h^H v^h$  in a coarse grid point is simply the value of  $v^h$  in the fine grid point coinciding with this coarse grid site. Alternatively, the value  $I_h^H v^h$  in a coarse grid point can be obtained by taking a weighted average of the values of  $v^h$  in the coinciding fine grid point and some sites adjacent to it. Figure 3.2 illustrates the restriction from a fine to a coarse grid for both operators in the one dimensional situation. The stencil of *injection* is given by:

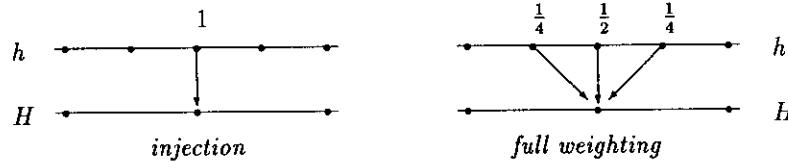
$$I_h^H = [0, 1, 0] \quad (3.30)$$

whereas the stencil of full weighting reads:

$$I_h^H = \frac{1}{4}[1, 2, 1] \quad (3.31)$$

In the two dimensional situation the stencil of  $I_h^H$  reads:

$$I_h^H = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad (3.32)$$

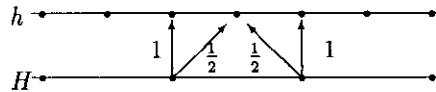
FIGURE 3.2: Restriction operators for  $d = 1$ 

in the case of *injection* whereas for *full weighting* it reads:

$$I_h^H = \frac{1}{16} \begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix} \quad (3.33)$$

The second inter-grid operator is the interpolation from the coarse to the fine grid denoted by the symbol  $I_H^h$ . If  $v^H$  is a coarse grid function then  $I_H^h v^H$  is a fine grid function. Assuming  $n_f$  fine grid nodes and  $n_c$  coarse grid points, the interpolation operator  $I_H^h$  can be written as a  $n_f \times n_c$  matrix. If  $v^H$  is a coarse grid function, then  $I_H^h v^H$  is a fine grid function obtained by multi-polynomial interpolation of some specified order, e.g. linear interpolation in the case of a one dimensional problem, see figure 3.3. Let the subscript  $i$  denote a point of the fine grid and  $I$  a coarse grid point. Assuming that the even fine grid points coincide with the coarse grid points, i.e.  $i = 2I$ , the value of  $[I_H^h v^H]_i$  is defined as:

$$[I_H^h v^H]_i \equiv \begin{cases} v_I^H & i = 2I \\ (v_I^H + v_{I+1}^H)/2 & i = 2I + 1 \end{cases} \quad (3.34)$$

FIGURE 3.3: Linear interpolation  $d = 1$

### 3.1.5 Coarse grid correction cycle

If the number of nodes on the fine grid is large, the number of nodes on a coarse grid with  $H = 2h$  may still be relatively large and after a few relaxations convergence on the coarse grid will slow down again. Obviously, the same reasoning as presented in the preceding section applies to the solution of the error on this grid. After a few relaxations the error on this grid will be smooth and can be described accurately on a grid with mesh size  $2H$ . This process of coarsening can be repeated recursively until a grid is reached where the number of nodes is so small compared to the original fine grid that the equations can be solved exactly or almost exactly in only a few operations.

To describe the coarse grid correction in the case of a sequence of grids the superscripts  $h$  and  $H$  are replaced by the superscript  $k$  indicating the level. The coarsest grid is referred to as level 1. In general, the mesh size on level  $k$  is given by  $2^{1-k}h^1$ , where  $h^1$  is the mesh size on the coarsest grid. Other coarse to fine grid mesh size ratio's are possible but in general the ratio 2 is very convenient and will be assumed throughout this thesis. The problem to be solved on a given level  $k$  is written as:

$$L^k u^k = f^k \quad (3.35)$$

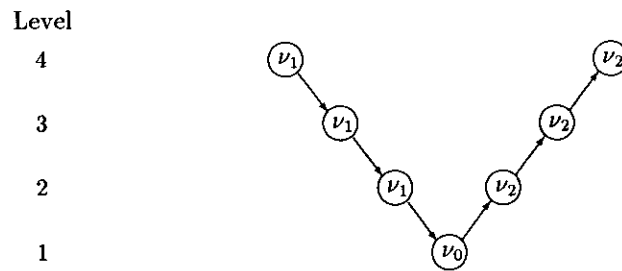
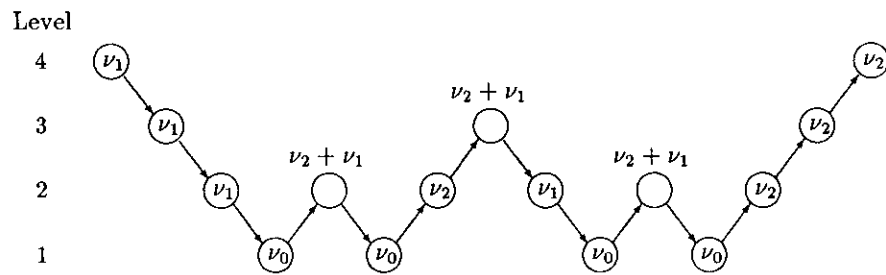
The recursive description of the coarse grid correction for level  $k$  then reads:

- If  $k = 1$ , i.e. the coarsest grid, solve the problem exactly or carry out  $\nu_0$  relaxations to solve the problem nearly exactly.
- if  $k > 1$ 
  - Perform  $\nu_1$  relaxation sweeps on (3.35), resulting in a new approximation  $\bar{u}^k$ .
  - Calculate the coarse grid right hand side, e.g. if the correction scheme is employed calculate  $I_k^{k-1} r^k$ .
  - Perform  $\gamma$  coarse grid correction cycles on the level  $k - 1$  equation.
  - Correct the solution on level  $k$  using the results of level  $k - 1$  e.g. if the correction scheme is used:

$$\bar{\bar{u}}^k = \bar{u}^k + I_{k-1}^k u^{k-1} \quad (3.36)$$

- Finally, perform  $\nu_2$  additional relaxation sweeps on (3.35) starting with  $\bar{\bar{u}}^k$  yielding the final new approximation to  $u^k$ .

The sweep counts  $\nu_1$  and  $\nu_2$  are usually either 0, 1 or 2 whereas  $\nu_0$  is in general larger, e.g.  $\nu_0 = O(\nu_1 + \nu_2)$ . The coarse grid correction cycle for  $\gamma = 1$  is generally referred to as the  $V(\nu_1, \nu_2)$ -cycle. For the situation of 4 levels (grids), this cycle is displayed in figure 3.4. Figure 3.5 shows the coarse grid correction cycle with  $\gamma = 2$  for the same number of levels. This cycle is generally referred to as the  $W(\nu_1, \nu_2)$  cycle.

FIGURE 3.4:  $V(\nu_1, \nu_2)$  cycle.FIGURE 3.5:  $W(\nu_1, \nu_2)$  cycle.

### 3.1.6 Performance

An estimate of the error reduction that is possible with a coarse grid correction cycle can be obtained from the smoothing rate analysis. Assuming that the low frequency components of the error are solved on the coarser grids, the error reduction is determined mainly by the efficiency of the relaxation process to reduce the high frequency components. Hence the reduction factor per cycle is:

$$\bar{\mu}^{(\nu_1+\nu_2)} \quad (3.37)$$

where  $\bar{\mu}$  is the *asymptotic smoothing rate* defined as the maximum amplification factor for the error components that can not be described on the coarse grid:

$$\bar{\mu} = \max_{\pi/2 \leq |\theta| \leq \pi} \mu(\theta_1, \theta_2) \quad (3.38)$$

For example, the asymptotic smoothing rate for Gauss-Seidel relaxation with lexicographic ordering on the 2-D Poisson problem (see section 3.1.1) is 0.5. Hence, a correction cycle with  $\nu_1 = 2$  and  $\nu_2 = 1$  gives an error reduction of almost one order of magnitude. Note that this reduction is independent of the number of nodes on the grid.

Next the amount of work a coarse grid correction cycle costs is investigated. For that purpose the *Work Unit* (WU) is introduced. One work unit is defined as the amount of work equivalent to one relaxation on the finest grid. Let  $WUV$  denote the amount of work of one  $V(\nu_1, \nu_2)$  cycle. If the intergrid transfers and the larger number of relaxations,  $\nu_0$ , on the coarsest grid are neglected it follows from the theory of geometrical series that:

$$WUV \leq \frac{(\nu_1 + \nu_2)}{(1 - 2^{-d})} WU \quad (3.39)$$

where  $d$  is the dimension of the problem. Equation (3.39) shows that, if one relaxation requires  $O(n)$  operations a  $V$  cycle also costs  $O(n)$  operations.

For example for the model problem ( $d = 2$ ) a  $V(2, 1)$  cycle adds up to the amount of work of approximately 4 relaxation sweeps on the finest grid. Three relaxations are actually performed on this grid ( $\nu_1 + \nu_2 = 3$ ) so all the work done on the coarser grids equals the equivalent of only one fine grid relaxation. Consequently, in an amount of work equivalent to four relaxations on the finest grid the error is reduced by an order of magnitude, independent of  $n$ . The power of the multilevel approach for large  $n$  is clearly demonstrated by the fact that, if no coarser grids are used, a similar reduction requires  $O(n)$  relaxations, i.e.  $O(n)$  WU.

### 3.1.7 Full multigrid

The solution of the system of discrete equations is an approximation to the solution of the continuous differential problem with an accuracy limited by the discretization

error. In fact one is only interested in the solution of the differential problem. Consequently, any approximation that differs by less than the discretization error from the exact solution of the discretized problem, is as good an approximation to the solution of the continuous differential problem as the exact solution of the discretized problem itself. An approximation that satisfies this criterion will be denoted as a *converged* solution.

Because the discretization error is usually  $O(h^p)$  it decreases with decreasing mesh size and, starting with an arbitrary first approximation,  $O(\ln n)$  coarse grid correction cycles are needed to obtain an approximation with an error smaller than the discretization error. However, if a converged solution of the discrete problem on a coarser grid is used as a first approximation, only  $O(1)$  cycles are required.

Since coarser grids are used to accelerate convergence anyway, they might as well be used to generate an accurate first approximation. This first approximation is subsequently improved by coarse grid correction cycles. Applying this idea recursively results in the solution process that is generally referred to as the *Full Multigrid* (*FMG*) algorithm. For the situation of 4 levels using one  $V(\nu_1, \nu_2)$  cycle per level the *FMG* algorithm is shown schematically in figure 3.6.

The double circles indicate the *converged* solutions at each level.  $I_H^h$  denotes the interpolation of the converged solution of a coarse grid to the next finer grid to serve as a first approximation. In many situations the same multi-linear interpolation can be applied as is used for the interpolation of the corrections in the coarse grid correction cycle. However, in some situations it should be of a higher order. Therefore it is represented by a different symbol, i.e.  $I_H^h$  instead of  $I_H^h$ . In the next section the subject of convergence is treated in more detail. In general a *FMG* algorithm with one or two coarse grid correction cycles per level is sufficient to solve the

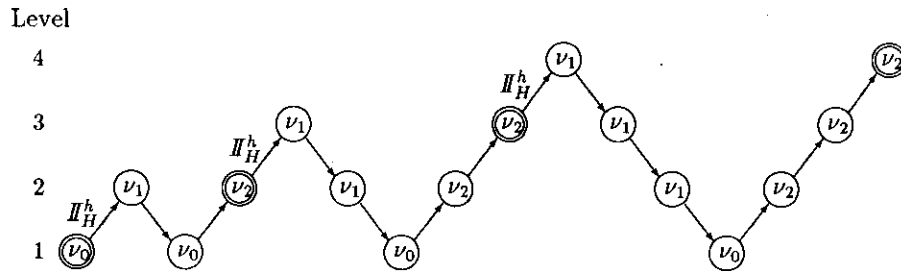


FIGURE 3.6: *FMG* –  $1V(\nu_1, \nu_2)$



problem within discretization error. Using the *FMG* algorithm the total amount of work units needed to solve the problem becomes independent of  $n$ . Neglecting the intergrid transfers, the total amount of work in the case of one  $V(\nu_1, \nu_2)$  cycle per level is approximately:

$$\frac{(\nu_1 + \nu_2)}{(1 - 2^{-d})^2} WU \quad (3.40)$$

Hence, in the case of the example presented in section 3.1.1, such an algorithm provides a solution of the problem in a computing time equivalent to the computing time of approximately 6 relaxations on the finest grid. In fact, only 3 relaxations are actually performed on this grid.

### 3.1.8 Convergence

An additional advantage of the *FMG* algorithm is that, employing the sequence of solutions computed on the different grids, it is easy to check the order of the discretization error and to test if the solutions have indeed converged below the level of the discretization error. For that purpose the following difference norm is calculated for all pairs of converged solutions (the solutions marked by a double circle in figure 3.6)  $\tilde{u}^k$  and  $\tilde{u}^{k-1}$ :

$$ERR(k, k-1) = H^d \sum |\tilde{u}^{k-1} - I_h^H \tilde{u}^k| \quad (3.41)$$

where the summation takes place over all coarse grid points.  $H$  denotes the mesh size on level  $k-1$  whereas  $h$  is the mesh size on level  $k$ ,  $H=2h$ .  $d$  is the dimension of the problem. The multiplication factor  $H^d$  is used to obtain a number that is independent of the level.

$ERR(k, k-1)$  will be made up of several parts:

- The discretization error on level  $k$ , i.e. the difference between the exact solution  $\underline{u}^k$  of the discretized problem on this particular grid and  $\underline{u}$ , the solution of the continuous differential problem represented on the same grid:

$$h^d \sum |\underline{u} - \underline{u}^k| \quad (3.42)$$

- Similarly, the discretization error on level  $k-1$ :

$$H^d \sum |\underline{u} - \underline{u}^{k-1}| \quad (3.43)$$

- The algebraic error on level  $k$ , that is, the error in the approximation  $\tilde{u}^k$  to the exact solution  $\underline{u}^k$  of the discrete problem.

$$h^d \sum |\tilde{u}^k - \underline{u}^k| \quad (3.44)$$

- The algebraic error on the coarsest of the two grids, level  $k - 1$ , similarly defined as:

$$H^d \sum |\tilde{u}^{k-1} - u^{k-1}| \quad (3.45)$$

By calculating solutions on both grids using one additional  $V$  or  $W$  cycle, both solutions ( $\tilde{u}^h$  and  $\tilde{u}^H$ ) can be obtained with algebraic errors being much smaller than the error in the original solutions (generally smaller by an order of magnitude as can be checked using the residuals). If the difference between the value of the norm (3.41) for the original solutions and its value for the solutions with smaller algebraic errors is small, both will be dominated by the discretization error. Obviously, the discretization error is the same in both situations, hence, the algebraic error must be small compared to the discretization error and the solutions are said to have converged below truncation error of the discretized equations.

The norm defined in equation (3.41) also serves as a check of the order of convergence of the discrete approximation to the continuous solution of the differential problem. For that purpose simply calculate the ratios  $ERR(k, k-1)/ERR(k-1, k-2)$ . With increasing  $k$ , i.e. decreasing mesh size, this ratio converges to  $2^{-p}$  where  $p$  is the approximation order obtained in practice. It may deviate from the theoretically expected value because of singularities or boundary effects.

### 3.1.9 Time dependent problems, the $F$ cycle.

The previous sections outlined the multilevel techniques developed for the fast numerical solution of problems governed by elliptical partial differential equations. Using a suitable relaxation scheme these techniques for example enable a fast solution of the stationary EHL line and point contact problem. However, one of the interesting subjects in the study of EHL contact situations is the analysis of the effects of surface features on pressure profile and film shape in the contact region and such an analysis requires a transient simulation since in general the surface feature moves. This section describes one of the multilevel techniques developed by Brandt and co-workers for the efficient solution of such a transient problem.

Consider the following discretized transient problem, where  $L^{h,k}$  is a differential operator discretized on a grid with meshsize  $h$ , time step  $k$ ,  $u^{h,k}$  is the solution to be calculated and  $f^{h,k}$  is a right-hand side function:

$$L^{h,k} u^{h,k} = f^{h,k} \quad (3.46)$$

Using the solution at the previous time step  $k-1$  as an approximation to the solution at the current time step  $k$ , residuals can be calculated according to:

$$r^{h,k} = f^{h,k} - L^{h,k} u^{h,k-1} \quad (3.47)$$


---

Defining:

$$\delta^{h,k} = \underline{u}^{h,k} - \underline{u}^{h,k-1}, \quad (3.48)$$

equation (3.46) can be written as:

$$L^{h,k}(\underline{u}^{h,k-1} + \delta^{h,k}) = L^{h,k} \underline{u}^{h,k-1} + \underline{r}^{h,k} \quad (3.49)$$

In general,  $\delta^{h,k}$  contains all frequencies the grid can represent, though in fact it is very often dominated by low frequencies. Using the normal *FAS* or *CS* coarse grid correction cycle  $O(\ln(n))$  cycles are needed to solve the problem to the level of the truncation error, with  $n$  being the number of nodes in spatial direction. If the first approximation were accurate up to the level of the truncation error on the coarser grid only  $O(1)$  cycles would be needed. To obtain a first approximation of this accuracy is the purpose of the *F* cycle.

Characteristic for the *F* cycle is that equation (3.49) is solved on the coarse grid first. Hence solving  $\hat{\underline{u}}^{H,k}$  from:

$$L^{H,k}(\hat{\underline{u}}^{H,k}) = L^{H,k}(I_h^H \underline{u}^{h,k-1}) + I_h^H \underline{r}^{h,k} \quad (3.50)$$

an approximation  $\tilde{\underline{u}}^{h,k}$  is calculated according to:

$$\tilde{\underline{u}}^{h,k} = \underline{u}^{h,k-1} + I_H^h(\hat{\underline{u}}^{H,k} - I_h^H \underline{u}^{h,k-1}) \quad (3.51)$$

Subsequently, normal *FAS* or *CS* coarse grid correction cycles can be used to reduce the error to the level of the truncation error. The same approach can be applied to the solution of the coarse grid problem, i.e. equation (3.50). Hence, the algorithm is defined recursively. For the case where one  $V(\nu_1, \nu_2)$  cycle is used, the solution process per time step is as depicted in figure 3.7, in which 4 grids are used.

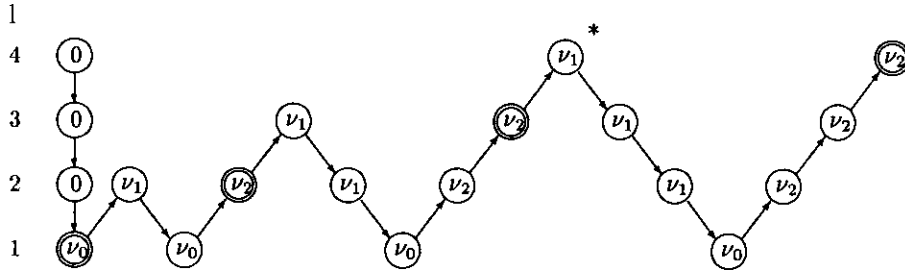


FIGURE 3.7:  $F - V(\nu_1, \nu_2)$  cycle

The figures in the circles denote the number of relaxation sweeps carried out on the grid. Note that in the first coarsening sequence no (0, zero) relaxations are performed. On the coarsest grid the equations should be solved almost exactly. Hence,  $\nu_0$  is in general larger than  $\nu_1$  and  $\nu_2$  which are both usually  $O(1)$ . The double circles represent converged solutions. In this figure the first approximation to  $\underline{u}^{h,k}$  is marked by an \*. This approximate solution is subsequently improved by coarse grid correction cycles.

### 3.1.10 Conclusion

The multilevel techniques described in this section will be used to deal with the slowness of convergence of the relaxation process applied to the solution of the EHL line and point contact problem and enable a fast calculation of the pressure and film thickness as a function of place and time. Attention is now directed towards the integral part of the problem. Because of the elastic deformation a suitable relaxation process when combined with the multilevel solution techniques will still result in a solver of a relatively high complexity, i.e.  $O(n^2)$ . The computing time will be dominated by the computation of the integrals describing the elastic deformation. The next section presents a fast algorithm for the evaluation of such integrals. The algorithm enables computation of the integrals in  $O(n \ln n)$  operations without loss of accuracy.

## 3.2 Multilevel multi-integration

In a general formulation the subject of study is *multi-integration*, i.e. given a function  $u$ , and a kernel  $K$  on a domain  $\Omega$  and requested the integral  $w$  where:

$$w(x) = \int_{\Omega} K(x, y)u(y)dy, \quad x \in \Omega \subseteq \mathbb{R}^d \quad (3.52)$$

To calculate  $w(x)$  at a certain location requires an integration over the entire domain. The discrete analogue of multi-integration is the multiplication of a vector by a dense matrix having certain smoothness properties. This numerical task arises in many physical and mathematical problems, as for example: integro-differential equations, integral equations, elasticity problems, coulombic molecular interactions and other many-body long range interactions.

If equation (3.52) is discretized on a uniform grid with mesh size  $h$  and  $n = O(h^{-d})$  nodes, the evaluation of every integral requires  $O(n)$  operations. Hence, a total of  $O(n^2)$  operations are needed for the calculation of  $w(x)$  in each of  $n$  gridpoints. Since in many situations a large number of nodes is essential, this often leads to rather excessive computing times, particularly if in the course of the iterative solution of an integral or integro-differential problem, e.g. the EHL problem, the integrals need to be evaluated more than once. Under these conditions the multi-integration quite often consumes the major part of the total computing time. For

example, the reader is referred to the multilevel algorithms for the solution of the EHL line- and point contact problem developed by Lubrecht [L1]. The application of the multilevel solution techniques resulted in a significant reduction of the computing time needed to solve the problems compared to the time required if conventional methods are used. However, the complexity of the algorithms was still  $O(n^2)$  because of the calculation of the elastic deformation integrals.

Recent developments enable a reduction of this complexity. Brandt and Lubrecht [B1] developed a multilevel algorithm for the fast numerical evaluation of such integrals. It was shown that for sufficiently smooth kernels the complexity can be reduced from  $O(n^2)$  to  $O(n)$ , whereas for potential type kernels such as  $\ln|x-y|$  and  $|x-y|^{-1}$  a reduction to  $O(n \ln n)$  can be obtained without loss of accuracy. This reduction is obtained by performing part of the integration on coarser grids. Exploiting the smoothness properties of the kernel  $K$ , this can be done in such a way that the error introduced by the coarse grid integration remains small compared to the original discretization error on grid  $h$ . In the following subsections first an outline is given of the multilevel multi-integration algorithm. Next, the results of its application to a one and two dimensional sample problem with the aforementioned kernels are presented.

### 3.2.1 Discretization

Define  $x_i^h = x_0 + ih$  to be equidistant gridpoints in  $\Omega$ , where  $x \in \mathbb{R}^d$  and  $i = (i_1, i_2, \dots, i_d)$  and let  $u_i^h = u^h(x_i^h)$  and  $w_i^h = w^h(x_i^h)$  denote the approximations to  $u$  and  $w$  on this grid. If the function  $u$  is approximated by a piecewise polynomial function  $\hat{u}^h$  of degree  $2s-1$  satisfying  $\hat{u}(x_j^h) = u_j^h$  equation (3.52) can be discretized as:

$$w_i^h \equiv \int_{\Omega} K(x, y) \hat{u}^h(y) dy = h^d \sum_j K_{i,j}^{hh} u_j^h \quad (3.53)$$

The coefficients  $K_{i,j}^{hh}$  can often be calculated analytically. In particular, the discretization of the elastic deformation integrals for both the EHL line and point contact situation, is outlined in appendix B. Provided  $u$  is sufficiently smooth the discretization error in equation (3.53) is  $O(h^{2s}|u|_{2s})$ , where  $|u|_{2s}$  is an upper bound for the  $2s$ -order derivatives of  $u$ .

### 3.2.2 General description

Before dealing with the specific situation of reducing the complexity of multi-integration when the kernel has a singularity (singular smooth), the general idea of the algorithm is described first. Section 3.2.2.2 outlines how to replace the original fine grid multi-summation by a similar summation on a coarser grid in such a way that the additional error remains small compared to the discretization error. For reasons

of simplicity only one coarser grid with mesh size  $H = 2h$  will be used. The extension of the algorithm to situations where the kernel has a singularity is discussed in section 3.2.4.

### 3.2.2.1 Notation

The running index on the coarse grid will be denoted by a capital letter.  $x_J^H = x_0 + JH$  represents a coarse grid point whereas  $x_j^h = x_0 + jh$  is a point in the fine grid. The coarse grid points are chosen to coincide with the even fine grid points satisfying  $x_J^H = x_{2J}^h$ . Similarly  $u_J^H = u^H(x_J^H)$  denotes the value of the coarse grid function  $u^H$  in the coarse grid point  $x_J^H$ . Furthermore, the symbol  $\mathcal{I}_H^h$  will denote an interpolation operator from the coarse to the fine grid of some specified order. The symbol  $\mathcal{I}_H^h$  will be used instead of  $I_H^h$  to indicate that the interpolation will often be of higher order. The index on which the interpolation works will, if necessary, be denoted by a dot. For example  $\mathcal{I}_H^h K_{i,\cdot}^{hH}$  represents, for each  $j$  a fine grid function obtained by interpolation from the coarse grid function  $K_{i,J}^{hH}$ . The value of the result in the fine grid point  $x_j^h$  will be denoted by  $[\mathcal{I}_H^h K_{i,\cdot}^{hH}]_j$ .  $K_{i,j}^{hH}$  denotes a discrete kernel whose first index is in the fine grid and the second in the coarse grid, that approximates  $K(x_i^h, x_J^H)$ .

### 3.2.2.2 Coarse grid integration

The coarse grid equation is obtained from the fine grid equation in two steps. First an approximation  $\tilde{w}_i^h$  to  $w_i^h$  is derived which only involves summations over coarse grid points. Let  $\tilde{K}_{i,j}^{hh}$  be an approximation to  $K_{i,j}^{hh}$  defined as:

$$K_{i,j}^{hh} \approx \tilde{K}_{i,j}^{hh} \equiv [\mathcal{I}_H^h K_{i,\cdot}^{hH}]_j \quad (3.54)$$

where the interpolation  $\mathcal{I}_H^h$  is of sufficiently high order. The coarse grid function  $K_{i,\cdot}^{hH}$  is obtained by "injection" from the fine grid function  $K_{i,\cdot}^{hh}$ ; i.e.,  $K_{i,J}^{hH} \equiv K_{i,2J}^{hh}$ . Using  $\tilde{K}_{i,j}^{hh}$  equation (3.53) can be written as:

$$w_i^h = h^d \sum_j \tilde{K}_{i,j}^{hh} u_j^h + h^d \sum_j (K_{i,j}^{hh} - \tilde{K}_{i,j}^{hh}) u_j^h \quad (3.55)$$

Since  $\tilde{K}_{i,j}^{hh}$  is an interpolation of  $K_{i,j}^{hh}$  itself using only coarse grid points, the operator  $(K_{i,j}^{hh} - \tilde{K}_{i,j}^{hh})$  is given by:

$$(K_{i,j}^{hh} - \tilde{K}_{i,j}^{hh}) = \begin{cases} 0 & j = 2J \\ O(h^{2p} K^{(2p)}(\xi)) & \text{otherwise} \end{cases} \quad (3.56)$$

where  $2p$  is the interpolation order and  $K^{(2p)}(\xi)$  is a  $2p^{th}$  derivative of  $K$  at some intermediate point  $\xi$ . Whenever  $K$  is smooth compared to  $u$  that is, if  $|h^{2p} K^{(2p)}| \ll |h^{2s} u^{(2s)}|$ , the second term in equation (3.55) will be small compared to the discretization error and  $w_i^h$  can be approximated accurately by  $\tilde{w}_i^h$  defined as:

$$\tilde{w}_i^h \equiv h^d \sum_j \tilde{K}_{i,j}^{hh} u_j^h = h^d \sum_j [\mathbb{I}_H^h K_{i,j}^{hH}]_j u_j^h = h^d \sum_j K_{i,j}^{hH} [(\mathbb{I}_H^h)^T u^h]_j \quad (3.57)$$

where  $(\mathbb{I}_H^h)^T$  is the *adjoint* of  $\mathbb{I}_H^h$ . If  $n_f$  denotes the number of nodes on the fine grid and  $n_c$  the number of coarse grid points, then  $\mathbb{I}_H^h$  can be written as a  $n_f \times n_c$  matrix and its transpose  $(\mathbb{I}_H^h)^T$  is a  $n_c \times n_f$  matrix. Hence,  $(\mathbb{I}_H^h)^T$  describes a fine-to-coarse transfer operator. For example if  $\mathbb{I}_H^h$  denotes multi-linear interpolation, its adjoint is, apart from a constant factor  $H/h$ , the well known “full weighting” operator, extensively used in multigrid algorithms.

Defining:

$$u^H \equiv 2^{-d} (\mathbb{I}_H^h)^T u^h \quad (3.58)$$

equation (3.57) can be written as:

$$\tilde{w}_i^h = H^d \sum_j K_{i,j}^{hH} u_j^H \quad (3.59)$$

Note that the evaluation of  $\tilde{w}_i^h$  involves a summation over the coarse grid only.

The second step in the derivation of a coarse grid equation is based on the fact that very often  $K(x, y)$  shows similar behaviour with respect to both variables. Introducing  $\hat{K}_{i,j}^{hh}$  being an approximation to  $K_{i,j}^{hh}$ , which is defined as:

$$\hat{K}_{i,j}^{hh} \equiv [\hat{\mathbb{I}}_H^h K_{i,j}^{hH}]_i \quad (3.60)$$

equation (3.53) can be written as:

$$w_i^h = [\hat{\mathbb{I}}_H^h w^H]_i + h^d \sum_j (K_{i,j}^{hh} - \hat{K}_{i,j}^{hh}) u_j^h \quad (3.61)$$

where:

$$w_I^H \equiv w_{2I}^h = h^d \sum_j K_{I,j}^{Hh} u_j^h = h^d \sum_j K_{2I,j}^{hh} u_j^h \quad (3.62)$$

$K_{I,j}^{Hh}$  is obtained from  $K_{i,j}^{hh}$  by injection, i.e.  $K_{I,j}^{Hh} \equiv K_{2I,j}^{hh}$ . Similar to  $\tilde{K}_{i,j}^{hh}$ ,  $\hat{K}_{i,j}^{hh}$  is an interpolation of  $K_{i,j}^{hh}$  itself, using only the points coinciding with coarse grid points, i.e. the grid points  $i = 2I$ . In particular, if  $K(x, y)$  is symmetric  $\hat{\mathbb{I}}_H^h = \mathbb{I}_H^h$  can be used. Assuming  $\mathbb{I}_H^h$  to be an interpolation of the order  $2p$  the operator  $(K_{i,j}^{hh} - \hat{K}_{i,j}^{hh})$  is given by:

$$(K_{i,j}^{hh} - \hat{K}_{i,j}^{hh}) = \begin{cases} 0 & i = 2I \\ O(h^{2p} K^{(2p)}(\xi)) & \text{otherwise} \end{cases} \quad (3.63)$$

If  $K$  is sufficiently smooth with respect to the  $x$  direction, the second term in equation (3.61) can be neglected and  $w_i^h$  can be approximated by:

$$w_i^h \approx [\hat{\mathcal{I}}_H^h w^H]_i \quad (3.64)$$

with:

$$w_I^H \equiv w_{2I}^h \quad (3.65)$$

Hence, if  $K$  is smooth with respect to the  $x$  direction it is sufficient to calculate the integrals in the points  $i = 2I$  only, using interpolation to obtain the values in the intermediate points.

Combining the results of both steps:

- If  $K$  is sufficiently smooth with respect to the  $y$  direction  $w_i^h$  can be accurately approximated by  $\tilde{w}_i^h$  obtained by summation over the coarse grid points,  $j = 2J$ , only.
- If  $K$  is sufficiently smooth with respect to the  $x$  direction it is sufficient to calculate  $w_i^h$  in the points  $i = 2I$  only, using interpolation to obtain the values in the other points.

leads to:

$$w_i^h \approx [\hat{\mathcal{I}}_H^h \tilde{w}^h]_i \quad (3.66)$$

with:

$$w_I^H \equiv \tilde{w}_{2I}^h = H^d \sum_J K_{I,J}^{HH} u_J^H \quad (3.67)$$

where  $K_{I,J}^{HH}$  is “injected” from  $K_{I,J}^{hH}$ , i.e.  $K_{I,J}^{HH} \equiv K_{2I,J}^{hH} = K_{2I,2J}^{hh}$ .

Hence, the original multi-summation (3.53) can be replaced by:

- The transfer of  $u^h$  to the coarse grid giving  $u^H$  according to equation (3.58).
  - The “calculation” of the coarse grid kernel:  $K_{I,J}^{HH} = K_{2I,2J}^{hh}$ .
  - Multi-summation on the coarse grid (3.67) to obtain  $w_I^H$ .
  - Injection of the coarse grid  $w_I^H$  to the coinciding fine grid points  $i = 2I$ , followed by an interpolation to obtain the values in the intermediate points,  $i = 2I + 1$ .
-



### 3.2.3 Coarse grid integration cycle

The procedure of replacing the multi-summation by a fine-to-coarse transfer, a coarse grid multi-summation and an interpolation described in the previous section can be applied once more to the coarse grid multi-summation and repeated recursively until a grid is reached with a number of gridpoints proportional to  $\sqrt{n}$ . On the latter grid the multi-summation is actually performed, requiring  $O(n)$  operations. The use of even coarser grids would not reduce the overall complexity since both the transfer of  $u^h$  from the finest grid to the next coarser grid, as well as the interpolation of the integrals from this coarser grid to the finest grid, already require  $O(n)$  operations.

Consequently, similar to the coarse grid correction cycle described in section 3.1.5, a sequence of coarser grids is used to obtain the approximations to the integrals on a given grid. The different grids are again referred to as levels and numbered, starting with the coarsest grid as level 1, the next finer grid being level 2, etc. For the situation where 3 coarser grids are employed, the multilevel calculation of the integrals on level 4 is schematically displayed in figure 3.8.

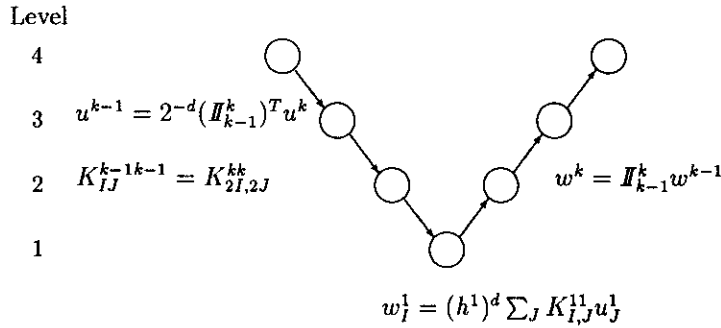


FIGURE 3.8: Multilevel multi-integration using 3 coarser grids.

### 3.2.4 Singular smooth kernel

In section 3.2.2.2 the kernel  $K$  was assumed to be much smoother than the function  $u$  over the entire domain of integration. However, the kernels in the elastic deformation integrals, i.e.  $K(x, y) = \ln |x - y|$  for the line contact problem and  $K(x, y) = |x - y|^{-1}$  in the point contact situation, are what is referred to as “singular smooth”, that is they have a singularity. In the vicinity of the singularity the  $2p^{\text{th}}$  differences of  $K$  will not be small. Hence, the second terms of both equation (3.55) and equation (3.61) will not be negligible. However, far from the singularity ( $\|y - x\| \gg h$  or

$\|j - i\| \gg 1$  for the discrete case), the  $2p^{th}$  differences of  $K$  will again be small. This knowledge of  $K$  will be used to derive corrections for the integrals such that again the error introduced by the coarse grid integration remains small compared to the discretization error. Consider first the calculation of the integrals in the even fine grid points Equation (3.55) can be written as:

$$w_i^h = w_I^H + h^d \sum_{\|j-i\| \leq m} (K_{i,j}^{hh} - \hat{K}_{i,j}^{hh}) u_j^h + h^d \sum_{\|j-i\| > m} (K_{i,j}^{hh} - \hat{K}_{i,j}^{hh}) u_j^h \quad (3.68)$$

where in the one-dimensional situation  $\|i - j\| = |i - j|$ . In situations with two or more dimensions the meaning of this norm is more complicated. It depends on the direction of interpolation as will be discussed in section 3.2.6. For the moment it is assumed that  $m$  is determined in such a way that the last term in equation (3.68) can be neglected. Consequently  $w_i^h$  is approximated by:

$$\tilde{w}_i^h = w_I^H + h^d \sum_{\|j-i\| \leq m} (K_{i,j}^{hh} - \hat{K}_{i,j}^{hh}) u_j^h \quad (3.69)$$

with:

$$w_I^H = H^d \sum_j K_{I,j}^{HH} u_j^H \quad (3.70)$$

Hence, in case of a singular smooth kernel, the multi-summation is again performed on the coarse grid. However, when injecting the results to the fine grid the contribution coming from the region in the vicinity of the singularity is corrected.

A similar procedure applies to the calculation of the integrals in the points  $i = 2I + 1$ . Assuming that  $K$  has similar smoothness properties in  $x$  and  $y$  direction, and that therefore identical interpolation operators are used in (3.54) and (3.60) equation (3.61) is written as:

$$w_i^h = [\mathcal{I}_H^h w^H]_i + h^d \sum_{\|j-i\| \leq m} (K_{i,j}^{hh} - \hat{K}_{i,j}^{hh}) u_j^h + h^d \sum_{\|j-i\| > m} (K_{i,j}^{hh} - \hat{K}_{i,j}^{hh}) u_j^h \quad (3.71)$$

The last term is again neglected and  $w_i^h$  is approximated by:

$$w_i^h \approx [\mathcal{I}_H^h w^H]_i + h^d \sum_{\|j-i\| \leq m} (K_{i,j}^{hh} - \hat{K}_{i,j}^{hh}) u_j^h \quad (3.72)$$

with:

$$w_I^H \equiv \tilde{w}_{2I}^h \quad (3.73)$$

Summarizing, compared to the situation of a smooth kernel, only the step from the coarse grid to the fine grid to obtain the integrals changes. First, as in the case of a smooth kernel described in section 3.2.2.2, coarse grid integrals are calculated. These

integrals are corrected and injected to the fine grid (3.69). Secondly, the corrected integrals in the points  $i = 2l$  are interpolated to the points  $i$  that are not part of the coarse grid and corrected again (3.72).

The remaining problem is to find a value of  $m$  for which the last terms in (3.68) and (3.71) can be neglected. For a detailed analysis the reader is referred to Brandt and Lubrecht [B1]. For example, for the logarithmic kernel in the one dimensional situation, the condition that the coarse grid integration error should be smaller than the fine grid discretization error gives:

$$m = 0.7pn^{s/p} - 1 \quad (3.74)$$

Combining this equation with the requirement of a minimum number of operations per fine grid point gives:

$$m \propto \ln(n) \quad (3.75)$$

Hence, the complexity of multi-integration in case of a singular smooth kernel can be reduced from  $O(n^2)$  to  $O(n \ln(n))$ .

### 3.2.5 One dimensional example

As an example consider the following multi-integral where the exact solution can be calculated analytically:

$$w(x) = \int_{-1}^1 \ln|x-y|(1-y^2)dy, \quad x \in [-1, 1] \quad (3.76)$$

Discretizing the equation on a uniform grid with mesh size  $h$ , as described in appendix B, gives:

$$w_i^h = \sum_j K_{i,j}^{hh} u_j^h \quad (3.77)$$

with:

$$u_j^h = (1 - y_j^2) \quad (3.78)$$

and:

$$\begin{aligned} K_{ij}^{hh} = & (i - j + \frac{1}{2})h(\ln(|i - j + \frac{1}{2}|h) - 1) - \\ & (i - j - \frac{1}{2})h(\ln(|i - j - \frac{1}{2}|h) - 1) \end{aligned} \quad (3.79)$$

Equations (3.69) and (3.72) were used for the fast integration with  $m = 3 + 2\ln(n)$ . The coarsest grid  $l = 1$  consisted of 8+1 points, the next finer grid 16+1 etc. As much as 16 levels ( $n = 262144+1$ ) have been used. To monitor the error in the multilevel multi-integration the following error norm is defined:

$$ERR_k^l = \frac{1}{n+1} \sum_{i=0}^n |w_i^{k,l} - w(x_i)|, \quad (3.80)$$

$ERR_k^l$  is the average absolute error of the integrals on level  $l$  when the actual multi-summation is carried out on level  $k$ , ( $k \leq l$ ). Note that  $ERR_k^l$  is by definition the  $L_1$  norm of the discretization error on grid  $l$ .  $w_i^{k,l}$  is the approximation to  $w_i$  on the finest grid obtained through equations (3.69) and (3.72). Table 3.1 gives the average errors (3.80) when sixth order transfers are used. The use of non-central operators near the boundaries was avoided as described in [B1]. The stencil of the 6<sup>th</sup> order restriction reads:

$$(\mathcal{H}_H^h)^T = \frac{1}{16}[-1, 0, 9, 16, 9, 0, -1] \quad (3.81)$$

The results marked by an asterisk indicate the calculations where the actual multi-summation was performed on a grid with approximately  $\sqrt{n}$  nodes. When one allows the additional error introduced by the coarse grid integration to be as large as the discretization error on the finest grid, the sixth order transfers give good results for  $2l - k \leq 17$ . This result is the same as that found by Brandt and Lubrecht [B1] using different  $K_{i,j}^{hh}$ . Hence, using sixth order transfers, up to  $l=11$  the actual summation can be carried out on a grid with approximately  $\sqrt{n}$  points. The computing times for the calculations presented in table 3.1 are shown in table 3.2.

The first column illustrates the  $O(n^2)$  complexity if no coarser grids are used whereas the sequence of starred results shows a complexity  $O(n \ln n)$ . Comparing the computing times for these results with the computing time needed if  $k = l$  clearly illustrates the power of the algorithm.

The result of similar calculations performed with tenth order transfers are presented in table 3.3. The computing times for the calculations presented in this table are shown in table 3.4. Using tenth order transfers at least up to level 16, that is  $n = 262144 + 1$ , the actual summation can be carried out on a grid with  $O(\sqrt{n})$  points without loss of accuracy. Without the use of the multilevel algorithm, the multi-summation on this level would require some 27 hours of computing time. As a result of the techniques presented here this time is reduced to some 50 seconds, that is by a factor of 2000.

l	k=l	k=l-1	k=l-2	k=l-3	k=l-4	k=l-5	k=l-6	k=l-7
5	9.21e-5	9.17e-5	9.10e-5	8.49e-5				
6	2.29e-5	2.28e-5	2.26e-5	2.23e-5	* 2.19e-5	2.32e-5		
7	5.69e-6	5.64e-6	5.53e-6	5.34e-6	5.01e-6	4.73e-6	5.38e-6	
8	1.42e-6	1.41e-6	1.37e-6	1.32e-6	1.21e-6	* 1.04e-6	9.24e-7	9.35e-7
9	3.55e-7	3.50e-7	3.41e-7	3.23e-7	2.87e-7	2.23e-7	1.30e-7	8.48e-8
10	8.80e-8	8.63e-8	8.17e-8	7.26e-8	5.45e-8	3.54e-8	* 6.76e-8	1.40e-7
11	2.21e-8		2.00e-8	1.71e-8	1.13e-8	1.12e-8	2.77e-8	6.28e-8
12	≈5.5e-9			3.88e-9	2.27e-9	4.07e-9	1.05e-8	* 2.40e-8
13	≈1.4e-9				9.78e-10	2.64e-9		

Table 3.1: Average error  $ERR_k^l$  logarithmic kernel using  $6^{th}$  order transfers

l	k=l	k=l-1	k=l-2	k=l-3	k=l-4	k=l-5	k=l-6	k=l-7
5	0.02	0.02	0.02	0.01	0.01			
6	0.09	0.04	0.03	0.03	* 0.03	0.03		
7	0.35	0.12	0.07	0.06	0.06	0.06	0.06	
8	1.38	0.43	0.19	0.14	0.13	* 0.14	0.14	0.14
9	5.51	1.54	0.58	0.38	0.30	0.30	0.31	0.31
10	22.80	5.48	1.90	0.93	0.77	0.65	* 0.57	0.60
11	95.80		6.33	2.57	1.58	1.46	1.34	1.36
12	≈384			7.84	4.17	3.11	2.90	* 2.94
13	≈1536				10.60	6.74		

Table 3.2: Computing time in seconds for table 3.1 (HP 9000/845s)

l	k=l	k=l-4	k=l-5	k=l-6	k=l-7	k=l-8	k=l-9	k=l-10
5	9.21e-5	9.27e-5						
6	2.29e-5	* 2.29e-5	2.18e-5					
7	5.69e-6	5.67e-6	5.69e-6	6.30e-6				
8	1.42e-6	1.42e-6	* 1.42e-6	1.42e-6	1.44e-6			
9	3.55e-7	3.54e-7	3.53e-7	3.53e-7	3.58e-7	3.58e-7		
10	8.86e-8	8.83e-8	8.80e-8	* 8.74e-8	8.69e-8	8.90e-8	8.93e-8	
11	2.21e-8	2.21e-8	2.20e-8	2.18e-8	2.17e-8	2.15e-8	2.17e-8	2.17e-8
12	≈5.5e-9	5.52e-9	5.50e-9	5.47e-9	* 5.42e-9	5.33e-9	5.28e-9	1.03e-8
13	≈1.4e-9	1.38e-9	1.37e-9	1.35e-9	1.32e-9	1.27e-9	1.17e-9	1.03e-9
14	≈3.5e-10		3.42e-10	3.39e-10	3.32e-10	* 3.17e-10	2.90e-10	2.42e-10
15	≈9.0e-11			8.47e-11	8.32e-11	8.00e-11	7.26e-11	5.74e-11
16	≈2.5e-11				1.88e-11	1.68e-11	* 1.22e-11	8.28e-12

Table 3.3: Average error  $ERR_k^l$  logarithmic kernel using  $10^{th}$  order transfers

l	k=l	k=l-4	k=l-5	k=l-6	k=l-7	k=l-8	k=l-9	k=l-10
5	0.03	0.02						
6	0.09	* 0.04	0.03					
7	0.35	0.06	0.06	0.07				
8	1.38	0.13	* 0.14	0.14	0.14			
9	5.51	0.30	0.29	0.30	0.34	0.24		
10	22.80	0.64	0.59	* 0.60	0.61	0.62	0.609	0.60
11	95.80	1.52	1.33	1.29	1.34	1.29	1.29	1.29
12	≈380	3.91	3.01	2.83	* 2.81	2.78	2.87	2.88
13	≈1.5e3	10.21	6.70	5.77	5.62	5.57	5.91	5.93
14	≈6.0e3		16.90	13.03	12.06	* 12.30	13.25	13.46
15	≈2.4e4			29.60	25.90	25.10	25.10	24.80
16	≈9.6e4				54.90	51.67	* 50.52	50.56

Table 3.4: Computing time in seconds for table 3.3 (HP 9000/845s)

### 3.2.6 Two dimensional example

The algorithm developed for the one dimensional problem presented above can easily be extended to two dimensions by applying the coarsening(s) procedure alternately in the  $x$  and  $y$  directions. In that way, the transfer of the function  $u$  and the kernel  $K$  and the correction of the integrals  $w$  are essentially the same as for the one dimensional situation. This approach may seem awkward at first sight but it has some major advantages amongst which is the simplicity of generalization to even higher dimensions without the need for complicated transfer operators. Furthermore, it ensures that the total work will continue to be  $O(n \ln n)$ , since all parts of the algorithm are at most of this complexity. Its disadvantage is the additional storage (50 %) required for the semi-coarsened grids.

The following multi-integration problem, where the integral can be calculated analytically, is studied:

$$w(x, y) = \int_{\Omega} K(x, y, x', y') u(x', y') dx' dy' \quad (3.82)$$

with:

$$K(x, x', y, y') = [(x - x')^2 + (y - y')^2]^{-1/2}$$

$$u(x', y') = \begin{cases} (1 - x'^2 - y'^2)^{1/2} & x'^2 + y'^2 \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

$$\Omega = \{(x, y) \in \mathbb{R}^2 | x^2 + y^2 \leq 1\}$$

The second order ( $s = 1$ ) discretization of equation (3.82) described in appendix B will be used. The coarsest grid employed in the calculations consisted of  $(4+1) \times (4+1)$

l	k=l	k=l-1	k=l-2	k=l-3	k=l-4	k=l-5
2	2.31e-1	2.32e-1				
3	7.68e-2	7.63e-2	* 7.67e-2			
4	1.51e-2	1.50e-2	1.49e-2	1.53e-2		
5	4.69e-3	4.58e-3	4.43e-3	* 4.25e-3	4.73e-3	
6	$\approx 1.2e-3$		1.10e-3	9.85e-4	9.30e-4	1.59e-3
7	$\approx 3.0e-4$			3.63e-4	* 3.06e-4	3.43e-4
8	$\approx 7.5e-5$				4.32e-5	1.07e-4
9	$\approx 1.8e-5$					* 5.38e-5

Table 3.5: Average error  $ERR_k^l$  in the two-dimensional singular-smooth multilevel multi-integration using  $6^{th}$  order transfers

nodes and the finest grid  $(1024+1) \times (1024+1)$ . The corrections of the integrals according to equations (3.69) and (3.72) have been carried out over rectangles of  $(2m_1 + 1) \times (2m_2 + 1)$  points around the singularity, where  $m_1$  denotes the half-width of the correction region in the direction of interpolation and  $m_2$  is the half-width in the perpendicular direction. The calculations have been performed using  $m_1 = 3 + 0.5 \ln(n)$  and  $m_2 = 2$ . Table 3.5 gives the average error  $ERR_k^l$  using  $6^{th}$  order transfers. The first column of table 3.5 confirms that the discretization is of  $2^{nd}$  order. The results marked by an asterisk indicate the approximations to the integrals on the finest grid obtained when the actual multi-summation is performed on a grid with approximately  $\sqrt{n}$  points. Allowing the additional error resulting from the coarse grid integration to be as large as the discretization error the integration can be carried out on a grid with  $\sqrt{n}$  points up to at least level 8 ( $n \approx 260.000$ ).

Table 3.6 gives the computing time for the calculations presented in table 3.5. The first column illustrates the  $O(n^2)$  complexity if no coarse grids are used. Each time the mesh size is decreased by a factor of 2 the computing time increases by a factor of 16. The computing times for the results marked by an asterisk demonstrate that the multilevel techniques reduce the complexity of multi-integration to approximately  $O(n)$ . ( $\ln(n)$  increases very slowly with  $n$ ). As a result, particularly for large  $n$ , large reductions in computing time are obtained.

The average error and the computing times when using  $8^{th}$  order transfer operators are presented in tables 3.7 and 3.8 respectively. In this case the integrals can be calculated on a grid with  $\sqrt{n}$  points without loss of accuracy up to 9 levels ( $n \approx$  one million). Without the multilevel techniques this computation would require some 36 days whereas with the multilevel algorithm it requires some 14 minutes. The use of this large number of nodes may seem exaggerated for a two-dimensional problem but for a three dimensional problem a grid with only 100 points in each direction already gives a million points.

l	k=l	k=l-1	k=l-2	k=l-3	k=l-4	k=l-5
2	0.12	0.12				
3	0.59	0.29	* 0.28			
4	4.53	1.13	0.82	0.81		
5	47.75	6.08	3.07	* 2.58	2.54	
6	768.80		13.41	10.05	9.61	9.64
7	$\approx 1.2e4$			41.63	* 38.59	38.34
8	$\approx 2.0e5$				153.06	151.87
9	$\approx 3.1e6$					* 628.03

Table 3.6: Computing time in seconds for table 3.5 (HP 9000/845s).

l	k=l	k=l-1	k=l-2	k=l-3	k=l-4	k=l-5
2	2.31e-1	2.31e-1				
3	7.68e-2	7.92e-2	* 8.02e-2			
4	1.51e-2	1.52e-2	1.52e-2	1.51e-2		
5	4.69e-3	4.72e-3	4.77e-3	* 4.77e-3	4.71e-3	
6	$\approx 1.2e-3$		1.25e-3	1.27e-3	1.26e-3	1.25e-3
7	$\approx 3.0e-4$			4.26e-4	* 4.05e-4	3.54e-4
8	$\approx 7.5e-5$				7.71e-5	5.95e-5
9	$\approx 1.8e-5$					* 1.28e-5

Table 3.7: Average error  $ERR_k^l$  in the two-dimensional singular-smooth multilevel multi-integration using  $8^{\text{th}}$  order transfers

l	k=l	k=l-1	k=l-2	k=l-3	k=l-4	k=l-5
2	0.12	0.28				
3	0.59	0.60	* 0.57			
4	4.53	1.87	1.40	1.39		
5	47.75	8.62	4.16	* 3.72	3.73	
6	$\approx 760$		17.65	13.49	12.95	12.92
7	$\approx 1.2e4$			54.74	* 49.31	49.30
8	$\approx 2.0e5$				202.12	195.95
9	$\approx 3.1e6$					* 835.68

Table 3.8: Computing time in seconds for table 3.7 (HP 9000/845s).



### 3.2.7 Additional insights

As mentioned above, in many situations the multi-summation is performed in the course of the iterative solution of the function  $u$  occurring in the integral. For example  $u$  is solved from an integral equation in terms of  $u$  and  $w$ , see [B1], or from a partial differential equation in  $u$  and  $w$ , e.g. EHL problems. In these situations the multilevel solution techniques and the multilevel multi-integration algorithm can be combined very efficiently.

Assuming that the problem is solved using some relaxation process, the techniques described in section 3.1 can be used to accelerate convergence of the relaxation process, whereas the multilevel-multi-integration algorithm can be used for the fast evaluation of the integrals. In particular for the case of a one dimensional problem the two techniques can be combined in a very straightforward manner. The multilevel integration can use the very same coarse grids used by the solver. Additional storage is only required for the coarse grid  $K_{JJ}^{HH}$  and for the correction factors needed if the kernel is singular smooth. The multilevel solution and multi-integration can also be combined easily in the case of a two dimensional problem. However, the situation with respect to storage is less favourable because of the semi-coarsened grids used by the multi-integration that are generally not used by the solver.

The resulting complexity of a solver combining the multilevel solution techniques described in section 3.1 with the multilevel multi-integration algorithm will be  $O(n)$  for a smooth kernel and  $O(n \ln n)$  if the kernel is singular smooth.

If the multi-integration is performed in the course of the iterative solution of the function  $u$  an even larger reduction of the total amount of work of the solver can be obtained by having it resort to full order multi-integration as seldom as possible. In fact, only one such multi-integration on the finest grid, plus some much less expensive ones (on coarser grids and/or using lower accuracy) is all that is desired.

The proposed procedure to obtain a solution with some prescribed accuracy on the finest grid is an *FMG* algorithm as follows: First the equations are solved to a similar accuracy on a coarser grid, say with meshsize  $2h$ . (This is done by a similar procedure - so the algorithm is defined recursively.) The grid  $2h$  solution is then interpolated to grid  $h$  to serve as the first approximation there. As a second step, the first approximation thus obtained, denoted by  $u^0$ , is used for the calculation of the integrals by multi-integration. In the following coarse grid correction cycles to improve this first approximation on grid  $h$  no full order multi-integration is needed anymore. Instead, whenever needed (e.g., after each relaxation sweep or coarse grid correction), a multi-integration is done on the solution *increment*, i.e., the difference between the current approximation of the solution and  $u^0$ , and the integrals thus calculated are added to those of  $u^0$ . The incremental multi-integration can be of lower order, i.e. it can be performed using smaller  $p$  and  $m$  and consequently it requires less operations than the first multilevel multi-integration, see [B1].

Besides this, the calculation of the coarse grid corrections for the finest grid involves multi-summations on the coarser grids. The necessary accuracy of these sum-

mations is even lower, since in fact only a crude correction function is calculated. In summary, the entire *FMG* solution process requires only one full multi-summation on the finest grid, plus a similar one on each of the coarser grids.



## References

- [B1] **Brandt, A., and Lubrecht, A.A.**, 1990, "Multilevel matrix multiplication and fast solution of integral equations," *J. of Comp. Phys.*, **90**, **2**, 348-370.
- [B2] **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, W. Germany.
- [B3] **Briggs, W.L.**, 1987, "A Multigrid Tutorial," SIAM, Philadelphia, Pennsylvania, ISBN 0-89871-221-1
- [G1] **Gümbel, L.**, 1916, "Über geschmierte Arbeitsräder," *Z. ges. Turbinenwesen*, **13**, 357.
- [L1] **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.
- [L2] **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*, to appear.
- [M1] **Martin, H.M.**, 1916, "Lubrication of gear teeth," *Engineering* (London), **102**, 199.
- [M2] **McCormick, S.**, 1988, "Multigrid methods: theory, applications, and supercomputing," ed. by S.F. McCormick, New York, M. Dekker, ISBN 0-89871-214-9.

---

## Chapter 4

# Numerical solution of the line contact problem

From the description of the multilevel solution techniques in chapter 3 it is obvious that an efficient multilevel solver for the line contact problem requires a relaxation process that effectively reduces high frequency error components. In addition, to enable simulation of practically relevant situations, i.e. maximum pressures in excess of 2 GPa, the relaxation process should be very stable. The simple one point Gauss-Seidel relaxation process applied by Hamrock and Jacobson [H1] and Lubrecht [L1] does not meet these demands. For moderately loaded conditions the process serves well, although with increasing load underrelaxation and local relaxation with local film thickness updates are needed to stabilize the process. However, the process is not stable for high loads. With the analysis of a linearized version of the problem the nature of this unstable behaviour is explained (section 4.2). Subsequently, an alternative relaxation scheme is proposed that provides good error smoothing and stability for all load conditions, including high loads and section 4.3 explains how to combine this relaxation process and the multilevel techniques, described in chapter 3, into a (fast) multilevel solver for the line contact problem. Finally, as an example, the results obtained with this solver for a specific load situation are studied in detail. Convergence to the solution of the continuous differential problem is demonstrated as well as its low complexity.

### 4.1 Discretization

The governing equations are discretized on a uniform grid with mesh size  $h$  covering the calculational domain  $X_a \leq X \leq X_b$ . For the moment a stationary situation is assumed (the transient problem is addressed in chapter 8). Hence, the dimensionless Reynolds equation (2.41) reads:

$$\frac{d}{dX} \left( \epsilon \frac{dP}{dX} \right) - \frac{d(\bar{\rho}H)}{dX} = 0 \quad (4.1)$$

with the cavitation condition:  $P \geq 0$

and the boundary conditions:

$$P(X_a) = P(X_b) = 0$$

$\epsilon$  is given by:

$$\epsilon = \frac{\bar{\rho}H^3}{\bar{\eta}\lambda}$$

where

$$\lambda = \frac{6\eta_0 u_s R^2}{b^3 p_h}$$

In case of an incompressible lubricant the dimensionless density is given by  $\bar{\rho} = 1$ , otherwise it is calculated from the Dowson and Higginson equation:

$$\bar{\rho}(P) = \frac{0.59 \cdot 10^9 + 1.34 P p_h}{0.59 \cdot 10^9 + P p_h} \quad (4.2)$$

For an isoviscous lubricant the dimensionless viscosity  $\bar{\eta}$  is equal to one whereas for a viscous lubricant  $\bar{\eta}$  is calculated either from Barus' equation:

$$\bar{\eta}(P) = \exp(\bar{\alpha}P) \quad (4.3)$$

or from Roelands' equation:

$$\bar{\eta}(P) = \exp \left[ \frac{\alpha p_0}{z} \left( -1 + \left( 1 + \frac{p_h}{p_0} P \right)^z \right) \right] \quad (4.4)$$

Unless explicitly stated otherwise the Roelands equation will be used throughout this thesis and the lubricant is assumed to be compressible.

Using second order central discretization for the first term of (4.1) and first order upstream discretization of the wedge term, the discretized Reynolds equation at site  $i$ ,  $X = X_a + ih$ , reads:

$$h^{-2}(\epsilon_{i-\frac{1}{2}}P_{i-1} - (\epsilon_{i-\frac{1}{2}} + \epsilon_{i+\frac{1}{2}})P_i + \epsilon_{i+\frac{1}{2}}P_{i+1}) - h^{-1}(\bar{\rho}_i H_i - \bar{\rho}_{i-1} H_{i-1}) = 0 \quad (4.5)$$

with the cavitation condition:  $P_i \geq 0$ .  $\epsilon_{i-\frac{1}{2}}$  and  $\epsilon_{i+\frac{1}{2}}$  denote the value of  $\epsilon$  at the *intermediate* locations  $X = X_a + (i-1/2)h$  and  $X = X_a + (i+1/2)h$  respectively. They are calculated from values of  $\epsilon$  in the sites  $i$  and  $i-1$  or  $i$  and  $i+1$ . For example:

$$\epsilon_{i-1/2} \equiv (\epsilon_i + \epsilon_{i-1})/2$$

Alternatively an harmonic average can be used:

$$\epsilon_{i-1/2}^{-1} \equiv (\epsilon_i^{-1} + \epsilon_{i-1}^{-1})/2$$

$\epsilon_i$  is simply defined as:

$$\epsilon_i = \frac{\bar{\rho}(P_i) H_i^3}{\bar{\eta}(P_i) \lambda} \quad (4.6)$$

The equation for the dimensionless film thickness reads:

$$H(X) = H_{00} + \frac{X^2}{2} - \frac{1}{\pi} \int_{-\infty}^{\infty} \ln |X - X'| P(X') dX' \quad (4.7)$$

Discretization of the elastic deformation integral as described in appendix B gives the following equation for  $H_i$ :

$$H_i = H_{00} + \frac{X_i^2}{2} - \frac{1}{\pi} \sum_{j=0}^n K_{ij}^{hh} P_j \quad (4.8)$$

where

$$K_{ij}^{hh} = (i - j + \frac{1}{2})h(\ln(|i - j + \frac{1}{2}|h) - 1) \\ - (i - j - \frac{1}{2})h(\ln(|i - j - \frac{1}{2}|h) - 1)$$

The dimensionless force balance equation, stating that the integral over the pressure balances the externally applied contact load, is discretized according to:

$$h \sum_{j=0}^{n-1} \frac{(P_j + P_{j+1})}{2} - \frac{\pi}{2} = 0 \quad (4.9)$$

This global condition determines the value of the constant  $H_{00}$  in equation (4.8).

## 4.2 Relaxation

One of the problems encountered in the simultaneous numerical solution of the equations (4.5) and (4.8) together with the condition (4.9) is that the coefficient  $\epsilon$  in equation (4.5) varies several orders of magnitude over the calculational domain. In both the inlet and the outlet region  $\epsilon \gg 1$  since in those regions  $\bar{\eta}$  and  $\bar{\rho}$  are close to one and  $H^3$  will be large. For example, taking a quite common location of the inlet boundary  $X_a = -4$ ,  $\epsilon(X = X_a)$  will be  $O(500)$ . On the contrary, in the contact region  $\bar{\eta}$  is large,  $O(10)$  to  $O(10^{15})$  or even larger,  $\bar{\rho}$  will never exceed 1.33 and  $H^3$  is small;  $O(10^{-3})$  to  $O(10^{-9})$ . The parameter  $\lambda$  varies from  $O(1)$  for very

lightly loaded situations to  $O(10^{-4})$  for high loads. Consequently,  $\epsilon(X=0)$  can be as small as  $10^{-20}$ . As a result the character of the problem changes. In the regions where  $\epsilon$  is large the problem behaves as a differential problem whereas in the contact region the integral aspect of the problem dominates. An additional complication is that, with increasing load, the "boundary layer", i.e. the region  $X \approx -1$  where  $\epsilon$  changes from a relatively large to a very small value, narrows. In particular for highly loaded situations the pressure gradient in this region is large. The rapidly increasing pressure results, due to the exponential relation between viscosity and pressure, in an even larger increase of the viscosity and as a result the value of  $\epsilon$  drops several orders of magnitude in a region of only a few times the mesh size on the grid. The reverse occurs in the outlet region, i.e.  $X \approx 1$ , where the steeply decreasing pressure results in an increase of  $\epsilon$  of several orders of magnitude over only a few gridpoints.

The relaxation process used in a multilevel solver should be a good and stable error smoother over the entire domain, i.e. for both large and small values of  $\epsilon$ . With respect to the solution of highly loaded conditions (high pressure, large  $\bar{\alpha}$  resulting in large  $\bar{\eta}$ , small  $H$ ) the behaviour for extremely small  $\epsilon$  is of particular interest. A local mode analysis as presented in chapter 3 provides a useful tool for the analysis of the convergence behaviour of a relaxation process. However, carrying out such an analysis for the complete problem with local and global boundary conditions, cavitation condition and non-linearity is far too complex, if not impossible. Therefore the problem is linearized and a new relaxation process for the line contact problem will be developed based on the results obtained from the local mode analysis of the linearized problem. The model problem is chosen in such a way that it is characteristic for the local behaviour of the full problem.

#### 4.2.1 Linearized problem

The following linear approximation of the EHL line contact problem (equation (4.1)) is studied:

$$L(P) = \epsilon \frac{d^2 P}{dX^2} - \frac{dH}{dX} = 0 \quad X_a \leq X \leq X_b \quad (4.10)$$

With the boundary conditions  $P(X_a) = P(X_b) = 0$  and  $H$  given by equation (4.7).  $\epsilon$  is assumed to be a constant. Any value of the constant  $H_{00}$  can be chosen and the force balance equation plays no role. Furthermore the cavitation condition is not taken into account, i.e. negative values of  $P$  are allowed. Discretization of (4.10) on a uniform grid with mesh size  $h$  gives:

$$L_i^h(P_i) = \epsilon(P_{i-1} - 2P_i + P_{i+1})/h^2 - (H_i - H_{i-1})/h = 0 \quad (4.11)$$

with

$$H_i = H_{00} + \frac{X_i^2}{2} - \frac{1}{\pi} \sum_{j=0}^n K_{ij}^{hh} P_j \quad (4.12)$$



Since the matrix  $K_{ij}^{hh}$  is independent of  $P$  the problem is linear in  $P$  and a smoothing rate analysis can be carried out for different types of relaxation. Apart from the limiting situations  $\epsilon \rightarrow \infty$  and  $\epsilon = 0$  the smoothing rate of a specific relaxation will depend on  $\epsilon/h^2$ . If  $\epsilon/h^2 \gg 1$  the first term of equation (4.11) dominates and the problem behaves like the one-dimensional Poisson problem. In that case a simple Gauss-Seidel relaxation has good smoothing and stability properties. This Gauss-Seidel relaxation can be described as follows. Given an approximation  $\tilde{P}_i$  in each grid point and an approximation  $\tilde{H}_i$  to  $H_i$  calculated from equation (4.12), the gridpoints are visited in lexicographic order and in each gridpoint a new approximation  $\bar{P}_i$  is computed according to:

$$\bar{P}_i = \tilde{P}_i + \left( \frac{\partial L_i^h}{\partial P_i} \right)^{-1} r_i \quad (4.13)$$

where  $r_i$  is the dynamic residual defined as:

$$r_i = (\tilde{H}_i - \tilde{H}_{i-1})/h - \epsilon(\bar{P}_{i-1} - 2\tilde{P}_i + \tilde{P}_{i+1})/h^2 \quad (4.14)$$

and

$$\frac{\partial L_i^h}{\partial P_i} = -2\epsilon/h^2 + \frac{1}{\pi}(K_{i,i}^{hh} - K_{i-1,i}^{hh})/h \quad (4.15)$$

The new approximation  $\bar{P}_i$  is subsequently used to recalculate or update the elastic deformation integrals and to compute a new approximation  $\bar{H}_i$  to  $H_i$  from equation (4.12). Hence, starting with an approximation to  $P_i$  in all gridpoints and the associated approximation to  $H_i$ , one relaxation consists of one sweep to compute new pressures according to equation (4.13) followed by a multi-integration to compute the elastic deformation and film thicknesses associated with this new approximation to the pressure.

This type of relaxation scheme in which all integrals are recalculated or updated simultaneously and only once per relaxation sweep is preferred since then all integrals needed in one relaxation sweep can be computed by one (multilevel) multi-integration. The smoothing factor  $\mu$  as a function of  $\theta$  for large  $\epsilon/h^2$  is depicted in figure 4.1. Since  $\mu(\theta) \leq 1$  for all  $\theta$  the scheme is stable. Moreover it efficiently reduces high frequency error components since the asymptotic smoothing rate  $\bar{\mu}$  is 0.45. Consequently the scheme is very well suited to serve in a multilevel solver for the problem.

Unfortunately, with decreasing  $\epsilon/h^2$  the performance of the scheme deteriorates and below a specific value of  $\epsilon/h^2$  low frequency error components are amplified and the relaxation process diverges. In the above argument it was assumed that the integrals are evaluated only once per relaxation sweep. The amplification of low frequency error components is caused by the fact that the changes in the discrete integrals, resulting from changing all pressures, are large. Hence, some improvement

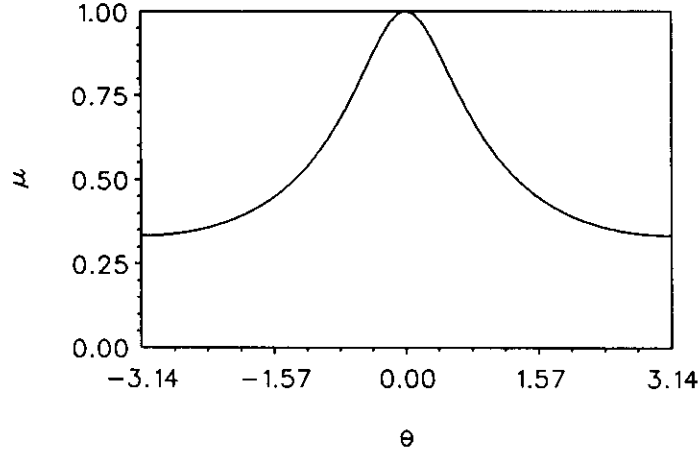


FIGURE 4.1: Amplitude reduction factor  $\mu(\theta)$  for Gauss-Seidel relaxation on linearized problem,  $\epsilon/h^2 \gg 1$

can be obtained by updating the integrals in the vicinity of the point where the pressure is changed during the relaxation.

Alternatively the stability range in terms of  $\epsilon/h^2$  can be extended to smaller values by means of underrelaxation. However, even combining both underrelaxation and local recalculation of integrals does not provide stability for the small values of  $\epsilon/h^2$  that will occur in the contact region at high loads.

To find a stable relaxation scheme with good smoothing properties for small  $\epsilon$  it is important to ensure that the process remains effectively *local* i.e., that relaxing at a point  $x_i$  introduces only small changes to the discrete integral  $H_j$  and in particular to the discrete derivative  $(H_j - H_{j-1})/h$  at points  $x_j$  far away from  $x_i$ ; otherwise each such integral would accumulate too many significant changes in a relaxation sweep. This can be achieved using a suitable kind of *distributed relaxation*.

In general, a distributed relaxation of the order  $r$  is a relaxation where each set of simultaneous changes is an  $r^{\text{th}}$ -order difference of a local function. (e.g., a multiple of a discrete delta function). Instead of changing the approximation in only one point, changes are also applied at one or more adjacent sites. For example, in the case of a first order distributive relaxation two unknowns are changed at a time:  $P_i \leftarrow P_i + \delta_i$  and  $P_{i-1} \leftarrow P_{i-1} - \delta_i$ , where  $\delta_i$  is calculated in such a way that, after applying these changes, the equation to be solved, e.g. equation (4.11), is satisfied at  $x_i$ . Such a relaxation can ensure that the changes in the integrals remain local since the changes in the integral at  $x_j$  caused by  $r^{\text{th}}$ -order distributed changes at  $x_i$  behave like  $\partial K^r(x_j, x_i)/\partial x^r$ , which decays like  $|x_i - x_j|^{-r}$ . Moreover, the changes in the derivative  $dH/dX$  occurring in the equation will behave like  $\partial K^{r+1}(x_j, x_i)/\partial x^{r+1}$ , which decays even faster.

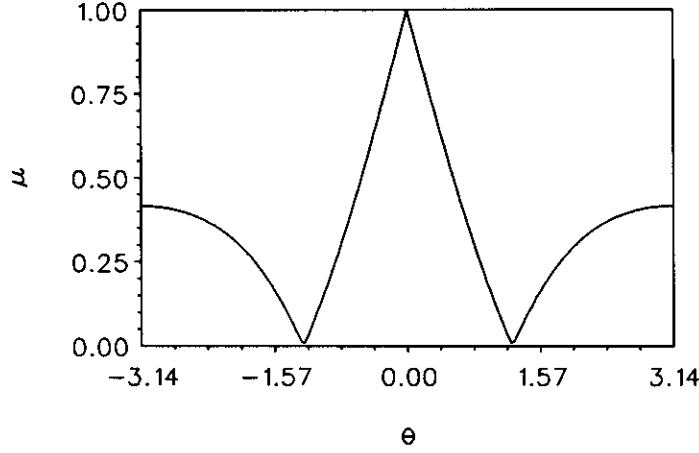


FIGURE 4.2: Amplitude reduction factor  $\mu(\theta)$  for Jacobi dipole relaxation on the linearized problem,  $\epsilon = 0$

The distributive relaxation scheme can be used either as a simultaneous displacement scheme, (new values of  $P$  and  $H$  replacing the old values at the end of a sweep), in which case it is referred to as a *distributive Jacobi* relaxation; or in successive displacements (the changed values being immediately used in relaxing subsequent equations), in which case it is called a *distributive Gauss-Seidel* relaxation.

For the linear model problem the effect of a first order distributed Jacobi relaxation was studied. This relaxation can be described as follows: starting with an approximation  $\tilde{P}_i$  in each gridpoint and an approximation  $\tilde{H}_i$  to  $H_i$  calculated from equation (4.12) changes  $\delta_i$  are computed from:

$$\epsilon([\tilde{P}_{i-1} - \delta_i] - 2([\tilde{P}_i + \delta_i]) + \tilde{P}_{i+1})/h^2 - (\tilde{H}_i - \tilde{H}_{i-1})/h = 0 \quad (4.16)$$

Subsequently all grid points are visited in lexicographic order adding  $\delta_i$  to the approximation in gridpoint  $i$  and subtracting  $\delta_i$  from the approximation in gridpoint  $i - 1$ . Because a change of the same magnitude but with opposite sign is applied at two adjacent gridpoints this relaxation is also referred to as a *dipole* relaxation. Hence, as a result of these changes, the new approximation to the pressure at site  $i$  is given by:

$$\bar{P}_i = \tilde{P}_i + \delta_i - \delta_{i+1} \quad (4.17)$$

The relaxation sweep is completed with the recalculation or update of all film thicknesses. Appendix C presents the calculation of the smoothing rate of this first order distributed relaxation for the model problem with  $\epsilon = 0$ . The result of this calculation, i.e. the smoothing factor  $\mu(\theta)$  is depicted in figure 4.2.

Obviously this relaxation process is stable since  $\bar{\mu}(\theta) \leq 1$  for all  $\theta$ . Moreover, it efficiently reduces high frequency error components. The asymptotic smoothing rate  $\bar{\mu} = 0.4$ . An even better reduction of high frequency components can be obtained if an underrelaxation factor of  $2/3$  is used, in which case  $\bar{\mu} = 0.2$ .

The relaxation process is equally effective for small values of  $\epsilon$ . However, with increasing  $\epsilon/h^2$  the performance deteriorates and the process gradually becomes unstable. This problem can be overcome by using some underrelaxation. Using an underrelaxation factor of  $0.6$  the process is stable even for large values of  $\epsilon/h^2$  with an asymptotic smoothing rate of  $0.7$ . Consequently for large values of  $\epsilon/h^2$  the Jacobi dipole relaxation is outranked by the simple Gauss-Seidel relaxation.

From the results presented above it is concluded that a stable and efficient multi-level solver for equation (4.11) regardless of the value of  $\epsilon$  can be obtained if the two relaxation schemes are combined. However, to ensure overlap in stability regions, at least for the Jacobi relaxation an underrelaxation factor of  $0.6$  should be used.

- On grids where  $\epsilon/h^2$  is larger than a prescribed limit the simple Gauss-Seidel relaxation is used.
- On grids where  $\epsilon/h^2$  is smaller than the prescribed limit, generally the coarsest grids, the Jacobi dipole relaxation is employed.

The optimal value of the cross-over point can be determined from smoothing rate calculations. If no underrelaxation is used for the Gauss-Seidel changes it is approximately  $\epsilon/h^2 = 1.0$  provided  $h$  is sufficiently small. On very coarse grids it is somewhat larger.

However, if an underrelaxation of factor  $0.6$  is used for the distributive changes the specific value of the cross-over point chosen in the calculations does not affect the overall performance too much as long as it is chosen sufficiently large. For example, using the (more conservative) switch criterion  $\epsilon/h = 0.01$  (instead of  $\epsilon/h^2$ ) also serves well. The only consequence is that the Jacobi dipole relaxation is applied on grids where also the simple Gauss-Seidel relaxation is stable. As a result the asymptotic error reduction per coarse grid correction cycle will be somewhat smaller.

#### 4.2.2 Varying coefficient

The next step towards a relaxation process for the solution of the EHL line contact problem is to consider equation (4.1) and, instead of assuming  $\epsilon$  to be a constant (as was done to obtain equation (4.10)),  $\epsilon$  is given as function of the spatial coordinate  $X$ . Hence,  $\epsilon$  varies over the calculational domain and the problem to be solved reads:

$$L(P) = \frac{d}{dX} \left( \epsilon \frac{dP}{dX} \right) - \frac{dH}{dX} = 0 \quad (4.18)$$

with the boundary conditions  $P(X_a) = P(X_b) = 0$  and  $H$  again given by equation (4.7). For the present purpose  $\epsilon(X)$  was chosen as:

$$\epsilon(X) = \begin{cases} X^6 & \text{if } |X| > 1 \\ 0 & \text{otherwise} \end{cases} \quad (4.19)$$

a function that resembles the  $\epsilon$  occurring in equation (4.1). The discontinuity at  $|X| = 1$  was chosen as a *worst case* simulation of the aforementioned boundary layer. Cavitation is not taken into account and again  $H_{00}$  and the force balance equation play no role. The equation is discretized according to:

$$\begin{aligned} h^{-2}(\epsilon_{i-1/2}P_{i-1} - (\epsilon_{i-1/2} + \epsilon_{i+1/2})P_i + \epsilon_{i+1/2}P_{i+1}) \\ - h^{-1}(H_i - H_{i-1}) = 0 \end{aligned} \quad (4.20)$$

where  $\epsilon_i = \epsilon(X_i)$ . In the multilevel algorithm described in 4.2.1, the relaxation type used on a specific grid depended on the value of  $\epsilon/h^2$  (or  $\epsilon/h$ ) on that grid. Since relaxation is, by its nature, a local process, both relaxation types can be combined on one grid. Thus, in regions of the domain where the coefficients are large, simple Gauss Seidel changes are applied, whereas in regions where the coefficients are small Jacobi dipole changes are calculated and applied after the sweep is completed. Using this hybrid relaxation type in a multilevel solver for equation (4.20) with  $\epsilon(X)$  given by (4.19), a reduction of the residuals of one order of magnitude per coarse grid correction cycle ( $V(2,2)$  or  $W(2,2)$  cycle) was easily obtained.

### 4.2.3 The line contact problem

The only remaining step towards the full non linear problem is to replace the dependence of  $\epsilon$  on  $X$  by a dependence on  $H$  and  $P$ . In that case the changes  $\delta_i$  to satisfy equation (4.5) can not be solved exactly. Instead a local Newton linearization is used. Assuming an approximate solution  $\tilde{P}_i$  with  $\tilde{H}_i$  given by equation (4.8) in regions of large  $\epsilon/h^2$  new approximations to  $P_i$  are calculated according to:

$$\bar{P}_i = \tilde{P}_i + \left( \frac{\partial L_i^h}{\partial P_i} \right)^{-1} r_i \quad (4.21)$$

where  $r_i$  is the dynamic residual of the discrete Reynolds equation at site  $i$ :

$$\begin{aligned} r_i = -(\epsilon_{i-1/2}\tilde{P}_{i-1} - (\epsilon_{i-1/2} + \epsilon_{i+1/2})\tilde{P}_i + \epsilon_{i+1/2}\tilde{P}_{i+1})/h^2 + \\ (\tilde{\rho}_i\tilde{H}_i - \tilde{\rho}_{i-1}\tilde{H}_{i-1})/h \end{aligned} \quad (4.22)$$

and  $(\partial L_i^h / \partial P_i)$  is approximated by

$$\frac{\partial L_i^h}{\partial P_i} \approx -(\epsilon_{i-1/2} + \epsilon_{i+1/2})/h^2 + \frac{1}{\pi}(\tilde{\rho}_i K_{i,i}^{hh} - \tilde{\rho}_{i-1} K_{i-1,i}^{hh})/h \quad (4.23)$$

The dipole changes  $\delta_i$  to be added to  $\tilde{P}_i$  and subtracted from  $\tilde{P}_{i-1}$  in regions of small  $\epsilon/h^2$  are calculated from:

$$\delta_i = \left( \frac{\partial L_i^h}{\partial P_i} - \frac{\partial L_i^h}{\partial P_{i-1}} \right)^{-1} r_i \quad (4.24)$$

with:

$$r_i = -(\epsilon_{i-1/2} \tilde{P}_{i-1} - (\epsilon_{i-1/2} + \epsilon_{i+1/2}) \tilde{P}_i + \epsilon_{i+1/2} \tilde{P}_{i+1})/h^2 + (\tilde{\rho}_i \tilde{H}_i - \tilde{\rho}_{i-1} \tilde{H}_{i-1})/h \quad (4.25)$$

and, since the kernel is symmetric ( $K_{i,i-1} = K_{i-1,i}$ ):

$$\frac{\partial L_i^h}{\partial P_i} - \frac{\partial L_i^h}{\partial P_{i-1}} \approx (-2\epsilon_{i-1/2} - \epsilon_{i+1/2})/h^2 + \frac{2}{\pi} (\tilde{\rho}_i K_{i,i}^{hh} - \tilde{\rho}_{i-1} K_{i,i-1}^{hh})/h \quad (4.26)$$

Note the small, but essential, difference between the residual  $r_i$  defined by equation (4.22) and (4.25). At first sight these equations seem the same, however, equation (4.22) already incorporates the changes applied in the previous gridpoint as is characteristic for Gauss-Seidel relaxation. On the contrary, equation (4.25) is entirely based on the old approximation.

The cavitation condition can be easily taken into account by simply replacing a negative  $\tilde{P}_i$  by zero. To obtain a solution that also satisfies the force balance equation the integration constant  $H_{00}$  is adjusted once every  $s$  relaxations in the following way:

$$H_{00} \leftarrow H_{00} - c \left( \frac{\pi}{2} - h \sum_{j=0}^{n-1} 0.5(\tilde{P}_j + \tilde{P}_{j+1}) \right) \quad (4.27)$$

where  $c$  is a suitably chosen constant and  $\tilde{P}_j$  denotes the current approximation to  $P_j$ . Basically this correction is of the same type as the correction of  $P_i$  according to equation (4.21), with the constant  $c$  replacing the derivative. The value of  $H_{00}$  depends on the values of governing parameters such as  $\lambda$  and  $\bar{\alpha}$ . With increasing load however it approaches a limiting value that can be calculated from the Hertzian dry contact theory:

$$H_{00} = -\frac{1}{4} - \frac{1}{2} \ln(2) \quad (4.28)$$

The number of relaxations,  $s$ , after which the force balance equation is relaxed, must be large enough and the multiplication factor  $c$  must be sufficiently small to avoid unstable oscillations. On the other hand, if  $c$  is chosen too small or  $s$  too large the changes of the value of  $H_{00}$  are too small and the relaxation of the force balance condition significantly slows down the overall convergence.

The hybrid relaxation process described above provides a stable solver for the line contact problem up to high loads. Because of the non-linearity some underrelaxation is required. The necessary underrelaxation factor depends on the load conditions and

varies from 0.1 to 0.6 (the aforementioned upper limit) for the Jacobi dipole changes and from 0.5 to 1.0 for the Gauss-Seidel changes. Furthermore, in particular for high loads, the mesh size on the grid should be sufficiently small. This matter is discussed in more detail in section 4.3.5. Since this relaxation scheme has good error smoothing properties it provides a sound basis for a multilevel solver of the problem.

### 4.3 Application of multilevel techniques

Having developed a stable relaxation process with good error smoothing properties this section directs the attention towards its implementation in a *FMG* algorithm. First some specific details with respect to the application of the multilevel solution techniques to Reynolds' equation, the film thickness equation and the force balance equation are discussed. Since the problem is not linear the correction scheme does not apply and the Full Approximation Scheme must be used. This applies to all three equations. Secondly the coarse grid correction cycle is described followed by the implementation of this coarse grid correction cycle into a *FMG* algorithm. Finally the implementation of the multilevel multi-integration algorithm for the fast computation of the elastic deformation integrals is addressed. To distinguish between the different grids used in the calculations, the superscript  $k$  is again used to indicate the level. Level 1 denotes the coarsest grid.

#### 4.3.1 The Reynolds equation

With respect to Reynolds' equation it should be noted that the problem to be solved is in fact a complementarity problem. This equation is not valid in the entire domain: either the fluid obeys Reynolds' equation or the fluid cavitates and the pressure is zero. The cavitation condition can be implemented in the multilevel process in a straightforward manner as described by Brandt and Cryer [B2]. Near the cavitation boundary both for the transfer of the residuals as well as for the transfer of the solution to the coarse grid *injection* should be used.

#### 4.3.2 The film thickness equation

The film thickness equation also deserves special attention. For a given pressure profile this equation allows an exact calculation of the associated  $H_i$ . Hence, provided a relaxation sweep ends with the update or recalculation of all  $H_i$ , the residuals of this equation are zero before going to the coarser grid. It is important to realize that this does not imply that the *FAS* right hand side of this equation on the coarse grid is zero. Recalling the definition of the *FAS* right hand side:

$$\hat{f}^H = L^H(I_h^H \tilde{u}^h) + I_h^H r^h \quad (4.29)$$

If the residuals are zero this equation reduces to:

$$\hat{f}^H = L^H(I_h^H \hat{u}^h) \quad (4.30)$$

Hence, in the case of multiple grids, the film thickness  $H_i$  on a specific level  $k$  is solved from:

$$H_i^k - H_{00} - \frac{X_i^2}{2} + \frac{1}{\pi} \sum_j K_{ij}^{kk} P_j = \hat{f}_i^k \quad (4.31)$$

where  $\hat{f}_i^k = 0$  on the finest grid used in the correction cycle, whereas on all coarser grids:

$$\hat{f}_i^k = [I_{k+1}^k \tilde{H}^{k+1}]_i - H_{00} - \frac{X_i^2}{2} + \frac{1}{\pi} \sum_j K_{ij}^{kk} [I_{k+1}^k \tilde{P}^{k+1}]_j \quad (4.32)$$

The treatment of the film thickness equation outlined here is essential to obtain an efficient multilevel algorithm for the solution of the problem.

### 4.3.3 The force balance equation

If only a single grid is used the treatment of this equation is quite straightforward. After a number of relaxations the value of  $H_{00}$  is changed in the direction that drives the residual of the force balance equation to zero. In case of multiple grids one might suggest to do something similar, that is to change the value of the constant from time to time on the finest grid. However this is not the appropriate way. The force balance equation is a global condition, i.e. it has a large global effect on the solution. The reader is reminded that the purpose of relaxation in a multilevel solver is to smooth the error. Relaxation of such a condition on the finest grid may frustrate this error smoothing process. In fact such a global condition should not be treated at all on the fine grid, see [B1, page 64]. All one should do is to transfer the residual of the condition to serve as the right-hand side for a similar condition on the next coarser grid. This process is repeated all the way down to the coarsest grid. On this grid the condition should be solved. Together with the *FAS* right hand sides of the Reynolds equation and the film thickness equation, the *FAS* right hand side for the force balance equation (4.9) is calculated. As a result the force balance condition on level  $k$  reads:

$$h \sum_{j=0}^{n-1} 0.5(P_j^k + P_{j+1}^k) = g^k \quad (4.33)$$

On the finest grid employed in the correction cycle  $g^k = \pi/2$  whereas on all coarser grids  $g^k$  is calculated from:

$$g^k = h^k \sum_{j=0}^{n_c-1} 0.5([I_{k+1}^k \tilde{P}^{k+1}]_j + [I_{k+1}^k \tilde{P}^{k+1}]_{j+1}) + \quad (4.34)$$



$$g^{k+1} - h^{k+1} \sum_{j=0}^{n_f-1} 0.5(\tilde{P}_j^k + \tilde{P}_{j+1}^k)$$

where  $h^k$  and  $h^{k+1}$  denote the mesh size at level  $k$  and level  $k+1$ . Similarly  $n_f$  is the number of nodes on level  $k+1$  and  $n_c$  denotes the number of nodes level  $k$ . Note that the last two terms in this equation are the residual of the same global condition on the next finer grid. On the coarsest grid,  $k=1$ , the equation is relaxed, i.e.  $H_{00}$  is changed according to:

$$H_{00} \leftarrow H_{00} - c(g^1 - h^1 \sum_{j=0}^{n_1-1} 0.5(\tilde{P}_j^1 + \tilde{P}_{j+1}^1)) \quad (4.35)$$

where  $\tilde{P}^1$  denotes the current approximation to the solution of the level 1 equations. As the treatment of the film thickness equation outlined in the previous section the approach described here is essential to obtain an efficient multilevel solver.

#### 4.3.4 Coarse grid correction cycle

The recursive description of the coarse grid correction cycle reads:

- If  $k=1$ , i.e. the coarsest grid:
  - perform  $\nu_0$  relaxations to solve the problem nearly exactly.
  - every  $s$  ( $\nu_0 > s$ ) relaxations the integration constant  $H_{00}$  is changed according to equation (4.35).
- if  $k > 1$ 
  - Perform  $\nu_1$  relaxations on the level  $k$  equations to obtain an approximation  $\tilde{P}^k$ . One relaxation sweep consists of:
    - \* The calculation of a new approximation to the pressure profile from the discretized Reynolds equation using the hybrid relaxation scheme.
    - \* A multi-integration to compute the elastic deformation and the associated new approximation  $\tilde{H}^k$  to  $H^k$ .
  - Transfer the solution ( $\tilde{P}^k$  and  $\tilde{H}^k$ ) to the coarse grid to serve as a first approximation on this grid and calculate the FAS coarse grid right hand sides for all equations.
  - Perform  $\gamma$  coarse grid correction cycles on the level  $k-1$  equation resulting in  $\tilde{P}^{k-1}$ .
  - Correct the solution on level  $k$  according to equation (3.29):

$$\bar{P}^k \leftarrow \tilde{P}^k + I_{k-1}^k(\tilde{P}^{k-1} - I_k^{k-1}\tilde{P}^k) \quad (4.36)$$

Since the film thickness equation can be solved exactly at any moment only the pressures are corrected. The corrected pressure profile is subsequently used to calculate the associated approximation to the film thickness:  $\bar{H}^k$ .

- Finally, Perform  $\nu_2$  additional relaxation sweeps on the level  $k$  equation.

#### 4.3.5 Full multigrid

The coarse grid correction cycle described in the previous section is implemented in a Full Multigrid algorithm as described in section 3.1.7. To obtain a first approximation to the solution on level  $k$  the converged solution of the preceding coarse grid is used. This approximation is subsequently improved with coarse grid correction cycles. As a first approximation on the coarsest grid, level 1, the dimensionless Hertzian dry contact pressure profile is used:

$$P_i = \begin{cases} \sqrt{1 - X_i^2} & \text{if } |X_i| < 1 \\ 0 & \text{otherwise} \end{cases} \quad (4.37)$$

To interpolate the solution of a coarser grid to the next finer grid to serve as a first approximation there a second order accurate interpolation is recommended. This first approximation is subsequently improved with coarse grid correction cycles.

For low load situations a *FMG* algorithm with two  $V(2,1)$  cycles per level is sufficient to obtain a converged solution. However, with increasing load the solution of the force balance equation becomes more difficult and convergence slows down. This can easily be overcome by using the more robust  $W$  cycle. In general a  $W(2,1)$  cycle gives an error reduction of one order of magnitude, even for highly loaded conditions and large numbers of nodes, and a *FMG* algorithm with two  $W(2,1)$  cycles per refinement provides a solution that has converged below the level of the discretization error. For a particular load situation this is demonstrated in section 4.4.

With respect to the solution of highly loaded situations it is important to separate the two functions of the coarser grids. With increasing load the discretization error on a given grid increases. If the grid is rather coarse the solution of the discrete problem on this grid is a very poor approximation to the solution of the differential problem. Moreover, the solution is also a bad first approximation to the solution on the next finer grid. Characteristic for such situations is the occurrence of rather large ( $O(1)$ ) negative film thicknesses. In combination with the non-linearity and the cavitation condition these negative values of  $H$  frustrate the convergence of the relaxation process. Hence, the first grid employed in the *FMG* algorithm must have a sufficiently small mesh size. The above argument however does not apply to the coarsest grids used in the coarse grid correction cycle. In the correction cycle the coarser grids are used to calculate a smooth correction to the fine grid pressure profile and upon convergence the coarse grid variable approximates the restriction

---

of the fine grid solution to this grid. Consequently in the coarse grid correction cycle the aforementioned problem does not occur. Summarizing, for highly loaded conditions the *FMG* process starts with one or two coarse grid correction cycles on a sufficiently dense grid (in general  $O(100)$  nodes) and in fact the two or three coarsest grids are skipped. However, although the last mentioned grids are not used in the process of generating a first approximation, they are used in the coarse grid correction cycle. Subsequently the normal *FMG* procedure is followed, i.e. the solution is used as a first approximation to the solution on the next finer grid and improved with some coarse grid correction cycles etc.

#### 4.3.6 Multilevel multi-integration

Each relaxation consists of the calculation of a new approximation to the pressure profile followed by the computation of the associated film thicknesses. The elastic deformation integrals needed in the latter computation require a multi-integration that can be performed in  $O(n \ln n)$  operations using the techniques outlined in the previous chapter. Hence, the summations:

$$-\frac{1}{\pi} \sum_{j=0}^n K_{ij}^{hh} P_j \quad (4.38)$$

are transferred to coarser grids, calculated, and interpolated back (using the correction procedure outlined in section 3.2.4) to the fine grid. To obtain the film thickness, the value of

$$H_{00} + \frac{X_i^2}{2} \quad (4.39)$$

is added to the fine grid integral in each point. If, using the advanced techniques described in section 3.2.7, only updates of the integrals are calculated, the summations:

$$-\frac{1}{\pi} \sum_{j=0}^n K_{ij}^{hh} \delta P_j \quad (4.40)$$

are transferred to coarser grids, calculated, interpolated back to the fine grid (using the correction procedure), and added to the fine grid summations that were calculated upon arrival at the specific grid for the first time:

$$-\frac{1}{\pi} \sum_{j=0}^n K_{ij}^{hh} P_j^0 \quad (4.41)$$

where  $P^0$  denotes the first approximation to pressure profile and  $\delta P = P - P^0$ , i.e. the difference between the current approximation and the first approximation.

Obviously calculating a new approximation to the pressure profile requires  $O(n)$  operations. The following update or recalculation of the integrals requires  $O(n \ln n)$  operations. Hence, one relaxation on the finest grid, i.e. one  $WU$ , represents  $O(n \ln n)$  operations.

## 4.4 Example

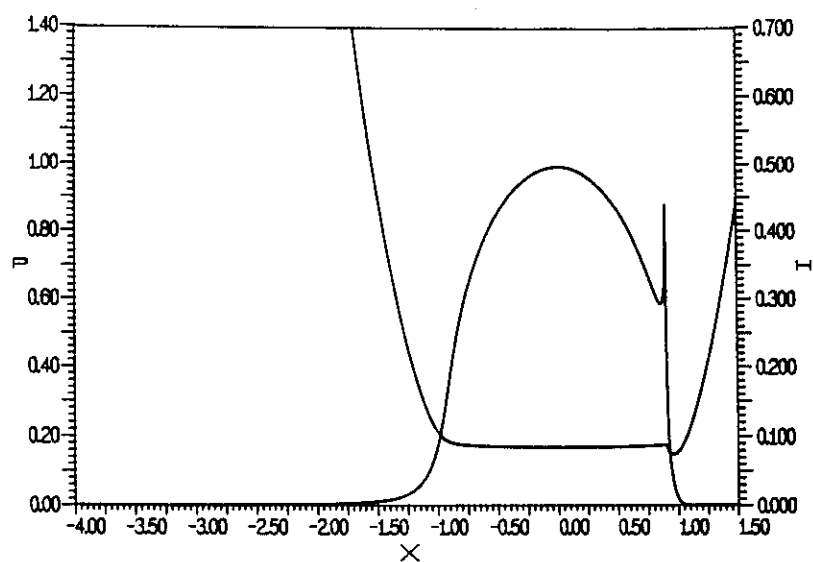
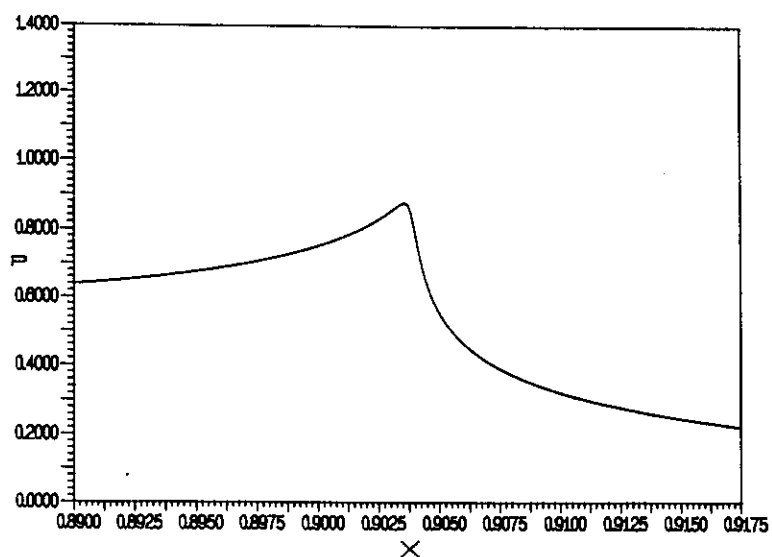
As an example a specific load situation is studied in detail. The values of the governing parameters are:  $M = 20$  and  $L = 10$  in terms of the Moes parameters or  $\bar{\alpha} = 17.84$  and  $\lambda = 9.25e - 3$  in terms of the parameters occurring in the equations or  $W = 8.94e - 5$ ,  $U = 1.0e - 11$  and  $G = 4730$  in terms of the Dowson and Higginson dimensionless parameters.

### 4.4.1 Computational details

The calculational domain extended from  $X_a = -4$  to  $X_b = 1.5$ . The solution has been calculated using a *FMG* algorithm with two  $W(2,1)$  cycles per refinement, each cycle giving an error reduction of approximately one order of magnitude. The coarsest grid employed in the calculations consisted of 15 nodes and as much as 14 levels have been used ( $n \approx 115,000$ ). The solution has been calculated using the Roelands relation and the lubricant was assumed to behave compressible according to the Dowson and Higginson equation.

### 4.4.2 Solution

Figure 4.3 presents the calculated pressure profile and the associated film shape for the situation studied. The solution contains all characteristic elements of medium to highly loaded EHL solutions. Starting at the inlet ( $X = -4$ ) the pressure gradually builds up to a nearly semi-elliptical profile in the Hertzian contact region  $-1 \leq X \leq 1$ . The cavitated outlet region  $X \geq 1$  is preceded by a pressure spike. This spike, caused by the exponential relation between the viscosity and the pressure, is discussed in detail in chapter 6. In particular the height of this spike has been the subject of many discussions. For the conditions considered here it can be shown that it is finite and smooth provided the local nodal density of the grid is sufficiently large. This is illustrated in figure 4.4 which gives an enlargement of the pressure profile in the spike region.

FIGURE 4.3: Pressure profile and film thickness for  $M = 20$  and  $L = 10$ FIGURE 4.4: Pressure profile in the spike region for  $M = 20$  and  $L = 10$

### 4.4.3 Convergence

Convergence to the level of the discretization error will be checked using the procedure outlined in section 3.1.8. The solution has been recalculated using three  $W(2,1)$  cycles. The values of  $ERR(k, k-1)$  calculated for the different levels for both 2 as well as 3 cycles are presented in table 4.1.

One  $W(2,1)$  cycle gives an error reduction of approximately one order of magnitude. Hence, the error in the solution calculated with three  $W(2,1)$  cycles is much smaller than the error if two  $W$  cycles are used. However, the value of  $ERR(k, k-1)$  calculated in both situations hardly differs. Consequently, it is dominated by the discretization error. Hence, two  $W(2,1)$  cycles are sufficient to obtain a solution that has converged below the level of the discretization error.

From table 4.1 it can be observed that, with increasing  $n$ , the ratio of  $ERR(k, k-1)/ERR(k-1, k-2)$  approximates a factor 1/2 indicating a first order convergence to the solution of the continuous differential equation.

The order of convergence can also be checked by monitoring the value of a characteristic parameter such as the minimum film thickness or the pressure spike height as a function of the number of nodes. For the situation considered here these results are presented in table 4.2 and figure 4.5. Both the (dimensionless) minimum film thickness as well as the (dimensionless) spike height converge first order to a limiting value with decreasing mesh size.

This first order convergence is exactly what one would expect since a first order upstream discretization of the wedge term in Reynolds' equation was used. Only in very lightly loaded situations where this term is negligible compared to the Poiseuille term, a second order convergence can be obtained, see for example Lubrecht [L1].

k	2 $W(2,1)$	3 $W(2,1)$
6	1.24e-3	1.24e-3
7	6.64e-4	6.59e-4
8	3.74e-4	3.76e-4
9	1.99e-4	1.97e-4
10	1.06e-4	1.06e-4
11	5.57e-5	5.35e-5
12	2.77e-5	2.72e-5
13	1.36e-5	1.36e-5
14	6.78e-6	6.78e-6

Table 4.1:  $ERR(k, k-1)$  using 2  $W$  and 3  $W$  cycles

level	$n+1$	$H_m$	$P_s$
6	449	0.07502	0.677
7	897	0.07439	0.710
8	1793	0.07404	0.758
9	3585	0.07385	0.787
10	7169	0.07375	0.825
11	14337	0.07370	0.850
12	28673	0.07367	0.867
13	57345	0.07366	0.875
14	114689	0.07365	0.879

Table 4.2: Convergence of spike height and minimum film thickness with increasing number of nodes for  $M = 20$  and  $L = 10$ .

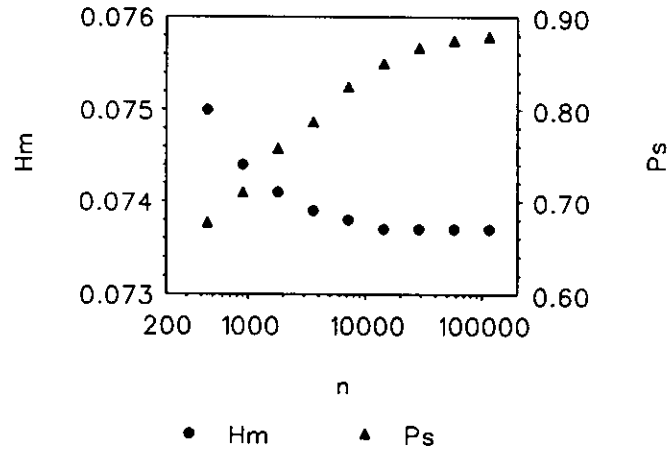


FIGURE 4.5: Minimum film thickness and pressure spike height as a function of the number of nodes  $n$  for  $M = 20$  and  $L = 10$ .

An alternative way to check convergence is to calculate the mass flow defect as a function of the number of nodes. Define the dimensionless mass flow per unit width  $\bar{\Phi}$  as:

$$\bar{\Phi} = -\frac{\bar{\rho}H^3}{\bar{\eta}} \frac{dP}{dX} + \lambda\bar{\rho}H \quad (4.42)$$

Using a second order central discretization for the pressure gradient the dimensionless mass flow per unit width at node  $i$  is given by:

$$\bar{\Phi}_i = -0.5 \frac{\bar{\rho}_i H_i^3}{\bar{\eta}_i} (P_{i+1} - P_{i-1})/h + \lambda\bar{\rho}_i H_i \quad (4.43)$$

Note that this expression is not valid in the cavitated region.

Table 4.3 gives the maximum, minimum, and average value of  $\bar{\Phi}_i$  as a function of the number of nodes. Also the mass flow defect in percent of the average mass flow:

$$\bar{\Phi}_{def} = 100 \frac{(\max(\bar{\Phi}_i) - \min(\bar{\Phi}_i))}{\text{av}(\bar{\Phi}_i)} \quad (4.44)$$

is displayed. Figure 4.6 displays the maximum, minimum and average mass flow as a function of the number of nodes. All three show a first order convergence to the same value and consequently the mass flux defect defined according to (4.44) converges to zero, see figure 4.7.

level	n+1	$\max(\bar{\Phi}_i)$	$\min(\bar{\Phi}_i)$	$\text{av}(\bar{\Phi}_i)$	$\bar{\Phi}_{def}$
6	449	9.969e-4	7.573e-4	9.008e-4	26.60
7	897	9.747e-4	8.535e-4	9.256e-4	13.09
8	1793	9.627e-4	9.016e-4	9.377e-4	6.53
9	3585	9.559e-4	9.255e-4	9.436e-4	3.22
10	7169	9.528e-4	9.376e-4	9.466e-4	1.61
11	14337	9.511e-4	9.435e-4	9.480e-4	0.81
12	28673	9.503e-4	9.461e-4	9.487e-4	0.44
13	57345	9.499e-4	9.477e-4	9.491e-4	0.23
14	114689	9.497e-4	9.486e-4	9.493e-4	0.12

Table 4.3: Convergence of the maximum, minimum and average dimensionless mass flow per unit width and mass flux defect with increasing number of nodes for  $M = 20$  and  $L = 10$ .



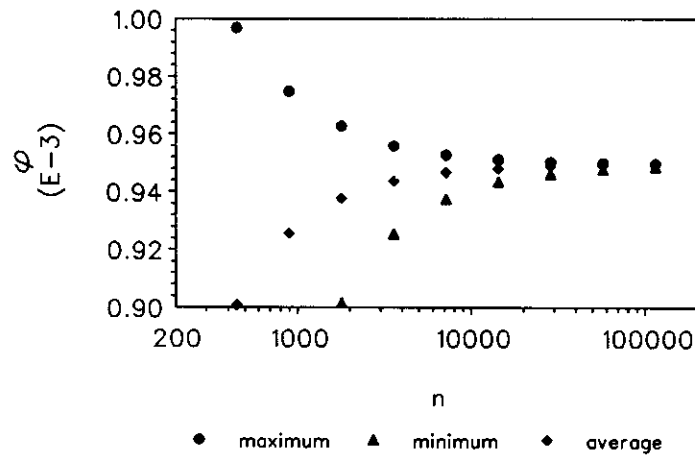


FIGURE 4.6: Maximum, minimum and average dimensionless mass flow per unit width as a function of the number of nodes  $n$  for  $M = 20$  and  $L = 10$ .

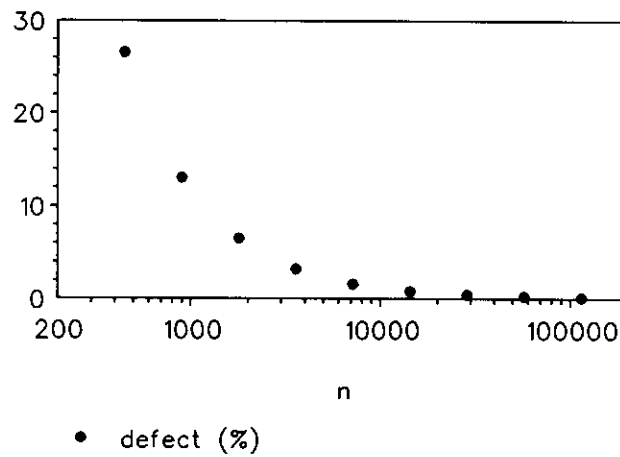


FIGURE 4.7: Mass flux defect in percent defined according to equation (4.44) as a function of the number of nodes  $n$  for  $M = 20$  and  $L = 10$ .

## 4.5 Computing times

To conclude this chapter the complexity of the solver is discussed. Table 4.4 presents the computing time on an HP 9000/845s computer as a function of the number of nodes  $n$  when using a *FMG* algorithm with 15 nodes on the coarsest grid if two  $V(2,1)$  or two  $W(2,1)$  cycles are used per refinement. The computing times in the first column are representative for lightly loaded conditions whereas the second column displays the computing time generally required to obtain a converged solution in a medium to highly loaded situation.

The required computing time approximately doubles with increasing number of nodes indicating a complexity close to  $O(n)$ . The larger increase going from level 12 to level 13 is caused by the change from 6<sup>th</sup> order to 10<sup>th</sup> order transfer operators in the multilevel calculation of the elastic deformation integrals, in order to keep the additional error introduced by the multilevel calculation small compared to the discretization error.

The data presented in table 4.4 is plotted in figure 4.8. For reasons of comparison also a line  $c n$  is drawn. Since one work unit represents  $O(n \ln n)$  operations, a  $V$  cycle and consequently a *FMG* algorithm with  $V$  cycles requires  $O(n \ln n)$  operations. Since  $\ln(n)$  increases very slowly with  $n$  the computing time results run parallel to the  $c n$  curve. With respect to the  $W$  cycles the situation is a bit more complicated. One would expect the ratio between the computing time needed for a *FMG* algorithm with 2  $W$  cycles and the computing time required by a *FMG* algorithm with 2  $V$  cycles to be constant. This ratio is displayed in figure 4.9. Obviously it increases with increasing number of nodes. The reason is that a  $W$  cycle in the case of a one dimensional problem requires  $O(\ln n)$  work units as can be shown with an analysis similar to the one for a  $V$  cycle presented in section 3.1.6. Hence, the complexity of the scheme when  $W$  cycles are used is  $O(n(\ln n)^2)$ .

level	$n+1$	2 $V(2,1)$ cycles	2 $W(2,1)$ cycles
5	225	2	6.2
6	449	4	14.8
7	897	7	34.5
8	1793	14	1:19
9	3585	29	2:58
10	7169	58	6:38
11	14337	1:57	14:56
12	28673	4:04	32:00
13	57345	9:50	1:17:43
14	114689	20:04	2:41:00

Table 4.4: Computing time in hours:minutes:seconds as a function of the number of nodes. (HP 9000/845s computer)

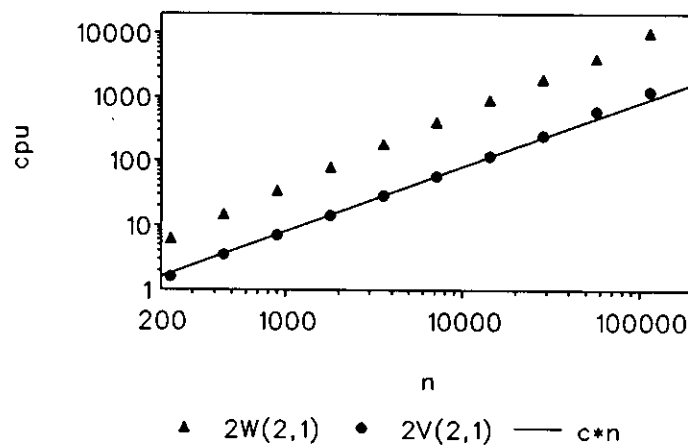


FIGURE 4.8: Computing time on an HP 9000/845s computer as a function of the number of nodes for the FMG algorithm with 2  $V(2,1)$  and 2  $W(2,1)$  cycles per refinement.

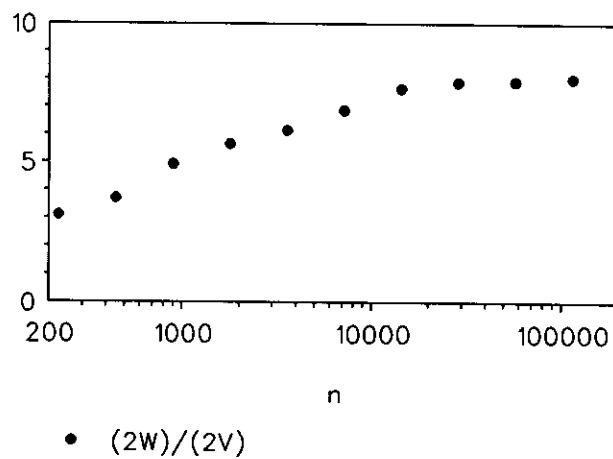


FIGURE 4.9: Ratio of the computing time used by the FMG with  $2W(2,1)$  and FMG with  $2V(2,1)$  cycles per refinement as a function of  $n$ .

## 4.6 Conclusion

Based on the analysis of a linearized model problem an alternative relaxation process for the EHL line contact problem has been developed. Incorporating this relaxation process in a *FMG* algorithm together with the use of multilevel multi-integration for the calculation of the elastic deformation integrals results in a solver with a complexity close to  $O(n)$ . This low complexity enables accurate solution of the problem using a large number of nodes on a small capacity computer. Moreover, as will be demonstrated in chapter 6 the algorithm is stable up to very high loads.

---

## References

- [B1] **Brandt, A.**, 1984, "Multigrid Techniques: 1984 Guide with applications to fluid dynamics," available as G.M.D.-Studie No. 85, from G.M.D.-F1T, Postfach 1240, D-5205, St. Augustin 1, W. Germany.
- [B2] **Brandt, A., and Cryer, C.W.**, 1983 "Multigrid algorithms for the solution of linear complementarity problems arising from free boundary problems," *SIAM J. Sci. Stat. Comput.*, **4**, 4, 655-684.
- [H1] **Hamrock, B.J., and Jacobson, B.O.**, 1984, "Elasto-hydrodynamic lubrication of line contacts," *ASLE trans*, **27**, 275-287.
- [L1] **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.
- [L2] **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*, to appear.



## Chapter 5

# Numerical solution of the circular contact problem

This chapter describes the development of a fast multilevel solver for EHL circular contact problems. The first step towards such a solver is to develop a relaxation process that gives good error smoothing and stability. As for the line contact problem the simple Gauss-Seidel relaxation that was for example applied by Chittenden et al. [C1,C2], Hamrock and Dowson [H1], Lubrecht [L1] and Zhu and Cheng [Z1] does not provide stability for highly loaded conditions. Again a linearized problem, characteristic for the local behaviour of the full problem, is studied to reveal the nature of this unstable behaviour. Based on the results of this study an alternative relaxation scheme providing stability and good error smoothing also for extreme conditions, is developed. This scheme is subsequently merged with the multilevel techniques described in chapter 3 into a solver for the circular contact problem of  $O(n \ln n)$  complexity. This complexity as well as convergence to the solution of the continuous integro-differential problem are demonstrated using a specific load situation.

### 5.1 Discretization

The equations are discretized on a rectangular uniform grid extended over the domain  $\{(X, Y) \in \mathbb{R}^2 | X_a \leq X \leq X_b, -Y_a \leq Y \leq Y_a\}$ , where  $X$  is the dimensionless coordinate in the direction of flow. Assuming a circular contact and steady state conditions the dimensionless Reynolds equation (2.48) reads:

$$\frac{\partial}{\partial X} \left( \epsilon \frac{\partial P}{\partial X} \right) + \frac{\partial}{\partial Y} \left( \epsilon \frac{\partial P}{\partial Y} \right) - \frac{\partial(\bar{\rho}H)}{\partial X} = 0 \quad (5.1)$$

with the cavitation condition  $P \geq 0$  and the boundary conditions:

$$P(X_a, Y) = P(X_b, Y) = 0$$

and

$$P(X, Y_a) = P(X, -Y_a) = 0.$$

$\epsilon$  is given by:

$$\epsilon = \frac{\bar{\rho} H^3}{\bar{\eta} \lambda}$$

where

$$\lambda = \frac{6\eta_0 u_x R_x^2}{a^3 p_h}$$

As in the line contact situation the dimensionless density is unity for an incompressible lubricant. In the case of a compressible lubricant the non dimensional expression of the Dowson and Higginson equation is used, see section 4.1. Similarly for an iso-viscous lubricant the dimensionless viscosity is one, otherwise it is computed from Roelands' equation (4.4)

Using a second order accurate central discretization for the first two terms of equation (5.1) and a first order upstream discretization of the wedge term leads to the following equation to be satisfied at each non boundary site  $(i, j)$ ,  $(X_a + ih, -Y_a + jh)$ :

$$\begin{aligned} h^{-2} & (\epsilon_{i-1/2,j}(P_{i-1,j} - P_{i,j}) + \epsilon_{i+1/2,j}(P_{i+1,j} - P_{i,j}) + \\ & \epsilon_{i,j-1/2}(P_{i,j-1} - P_{i,j}) + \epsilon_{i,j+1/2}(P_{i,j+1} - P_{i,j})) + \\ & - h^{-1}(\bar{\rho}_i H_i - \bar{\rho}_{i-1} H_{i-1}) = 0 \end{aligned} \quad (5.2)$$

with the cavitation condition:  $P_{i,j} \geq 0$ .

$\epsilon_{i-1/2,j}$ ,  $\epsilon_{i+1/2,j}$ ,  $\epsilon_{i,j-1/2}$  and  $\epsilon_{i,j+1/2}$  denote the value of  $\epsilon$  at the *intermediate* locations:

$$\begin{aligned} (X, Y) &= (X_a + (i - 1/2)h, -Y_a + jh), \\ &= (X_a + (i + 1/2)h, -Y_a + jh), \\ &= (X_a + ih, -Y_a + (j - 1/2)h), \text{ and} \\ &= (X_a + ih, -Y_a + (j + 1/2)h) \end{aligned}$$

respectively.

They are computed in an analogous way as described for the line contact problem in section 4.1, for example:

$$\epsilon_{i-1/2,j} \equiv (\epsilon_{i,j} + \epsilon_{i-1,j})/2$$

$\epsilon_{i,j}$  is defined as:

$$\epsilon_{i,j} = \frac{\bar{\rho}(P_{i,j}) H_{i,j}^3}{\bar{\eta}(P_{i,j}) \lambda} \quad (5.3)$$



The dimensionless film thickness equation for the circular contact (equation 2.49) reads:

$$H(X, Y) = H_{00} + \frac{X^2}{2} + \frac{Y^2}{2} + \frac{2}{\pi^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{P(X', Y') dX' dY'}{\sqrt{(X - X')^2 + (Y - Y')^2}} \quad (5.4)$$

Discretizing the elastic deformation integral as described in appendix B gives:

$$H_{i,j} = H_{00} + \frac{X_i^2}{2} + \frac{Y_j^2}{2} + \frac{2}{\pi^2} \sum_{k=0}^{n_x} \sum_{l=0}^{n_y} K_{ikjl}^{hhhh} P_{kl} \quad (5.5)$$

where  $(n_x + 1)$  and  $(n_y + 1)$  denote the number of nodes on the grid in  $X$  and  $Y$  direction respectively. Note that with respect to the elastic deformation integrals the general notation  $i = (i_1, \dots, i_d)$  used in chapter 3 to describe the multilevel multi-integration algorithm is abandoned here. The reason for this change is the combination of the integral equation with Reynolds' equation.

Finally the force balance equation determining the value of the integration constant  $H_{00}$  is discretized according to:

$$h^2 \sum_{i=1}^{n_x-1} \sum_{j=1}^{n_y-1} P_{i,j} - \frac{2\pi}{3} = 0 \quad (5.6)$$

## 5.2 Relaxation

The problems encountered in the simultaneous numerical solution of the discrete equations (5.2) and (5.5) with the global condition (5.6) are very similar to the problems one had to overcome in case of the line contact problem. The coefficient  $\epsilon$  varies several orders of magnitude over the calculational domain. Globally in the Hertzian dry contact region, i.e.  $(X^2 + Y^2) \leq 1$ , the integral aspect of the problem dominates, whereas in the remaining part of the domain the problem behaves like a differential problem and also the intermediate region, the *boundary layer*, narrows with increasing load. However, with respect to the multilevel solution of the problem, an additional complication shows up with increasing load. This complication is related to the fact that the problem is two-dimensional. In the region of small  $\epsilon$  equation (5.1) reduces to:

$$\frac{\partial(\bar{\rho}H)}{\partial X} \approx 0 \quad (5.7)$$

being a relation in  $X$  direction only. Consequently, when discretized, there is no direct coupling via the pressure between adjacent gridpoints in  $Y$  direction. These

points are only indirectly, and very weakly coupled via the elastic deformation integrals. This problem, i.e. the loss of coupling in  $Y$  direction was already observed by Lubrecht [L1] and is addressed in more detail in 5.2.1.

The approach to obtain a relaxation process that will also give good error smoothing for extreme conditions is basically the same as described in chapter 4 for the line contact problem. First a linear model problem is studied. This model problem is chosen in such a way that it is characteristic for the local behaviour of the full problem. Subsequently, using the results of the analysis of this linear problem, step by step, a relaxation process for the full problem is developed.

### 5.2.1 Linearized problem

The following linear approximation of the EHL circular contact problem, i.e. equation (5.1), is studied:

$$L(P) = \epsilon \left( \frac{\partial^2 P}{\partial X^2} + \frac{\partial^2 P}{\partial Y^2} \right) - \frac{\partial H}{\partial X} = 0 \quad (5.8)$$

on the domain  $\{(X, Y) \in \mathbb{R}^2 | X_a \leq X \leq X_b, -Y_a \leq Y \leq Y_b\}$  and the condition that  $P = 0$  on the boundaries.  $H$  is given by equation (5.5) and  $\epsilon$  is assumed to be a constant. In addition, cavitation is disregarded and the force balance equation plays no role. Discretization of equation (5.8) on a uniform grid with mesh size  $h$  gives:

$$\begin{aligned} L_{i,j}^h(P_{i,j}) = & \epsilon(P_{i-1,j} + P_{i+1,j} + P_{i,j-1} + P_{i,j+1} - 4P_{i,j})/h^2 \\ & -(H_{i,j} - H_{i-1,j})/h = 0 \end{aligned} \quad (5.9)$$

with:

$$H_{i,j} = H_{00} + \frac{X_i^2}{2} + \frac{Y_j^2}{2} + \frac{2}{\pi^2} \sum_{k=0}^{n_x} \sum_{l=0}^{n_y} K_{ikjl}^{hhhh} P_{k,l} \quad (5.10)$$

The smoothing rate of a specific relaxation process applied to the solution of equation (5.9) with  $H_{i,j}$  given by (5.10) depends on the ratio  $\epsilon/h^2$ . For large values of this ratio the following simple one point lexicographic Gauss-Seidel relaxation provides good error smoothing and stability.

Given an approximation  $\tilde{P}_{i,j}$  in each grid point and the associated approximation  $\tilde{H}_{i,j}$  to  $H_{i,j}$  calculated from equation (5.10), the gridpoints are visited in lexicographic order and at each gridpoint a new approximation  $\bar{P}_{i,j}$  is computed according to:

$$\bar{P}_{i,j} = \tilde{P}_{i,j} + \left( \frac{\partial L_{i,j}^h}{\partial P_{i,j}} \right)^{-1} r_{i,j} \quad (5.11)$$

where

$$r_{i,j} = (\tilde{H}_{i,j} - \tilde{H}_{i-1,j})/h - \epsilon(\bar{P}_{i-1,j} + \tilde{P}_{i+1,j} + \tilde{P}_{i,j+1} + \bar{P}_{i,j-1} - 4\tilde{P}_{i,j})/h^2 \quad (5.12)$$

and

$$\frac{\partial L_{i,j}^h}{\partial P_{i,j}} = -4\epsilon/h^2 - \frac{2}{\pi^2}(K_{i,i,j,j}^{hhhh} - K_{i,i-1,j,j}^{hhhh})/h \quad (5.13)$$

The new approximation  $\bar{P}_{i,j}$  is subsequently used to recalculate or update the elastic deformation integrals and to compute a new approximation  $\tilde{H}_{i,j}$  to  $H_{i,j}$  from equation (5.10).

Summarizing, given an approximation  $\tilde{P}_{i,j}$  to  $P_{i,j}$  in all gridpoints and the associated approximation  $\tilde{H}_{i,j}$  to  $H_{i,j}$  one relaxation consists of one sweep to compute a new approximation to  $P_{i,j}$  according to equation (5.11) followed by a multi-integration to obtain the associated new approximation to  $H_{i,j}$ .

The preference for such a relaxation scheme, in which all integrals are recalculated or updated simultaneously only once per relaxation sweep, has already been explained. Such a scheme enables an efficient combination with the multilevel multi-integration algorithm explained in section 3.2.

It is no surprise that this relaxation gives good error smoothing for large  $\epsilon/h^2$ . Under these conditions the discrete  $\partial H/\partial X$  term in equation (5.9) is small compared to the first term and the equation approximates the 2-D Poisson equation. For this problem with the usual 5-point discretization the asymptotic smoothing rate of one-point lexicographic Gauss-Seidel relaxation is 0.5 as was demonstrated in section 3.1.1.

With decreasing  $\epsilon/h^2$  two problems occur. Firstly, the relaxation gradually becomes unstable. Below approximately  $\epsilon/h^2 = 0.5$  low frequency error components are amplified and the relaxation process diverges. This unstable behaviour is caused by the accumulation of changes in the elastic deformation integrals during one relaxation sweep. Consequently, updating the integrals during the relaxation sweep and/or applying underrelaxation alleviates the stability problems to a certain extent. However, even a combination of both measures is not sufficient to stabilize the relaxation process for the small values of the  $\epsilon/h^2$  that will definitely occur in high load point contact situations. Secondly, because of the loss of coupling in  $Y$  direction, the relaxation becomes increasingly ineffective in reducing high frequency error components in  $Y$  direction.

A rigorous solution to the first problem, i.e. the stability problem, is provided by a distributed relaxation, since such a relaxation scheme can ensure that the changes of the integrals resulting from changing all pressures remain effectively local, e.g. see section 4.2.1. As a first guess one might suggest the application of the same distributed relaxation that was effective for the line contact problem, i.e. the Jacobi dipole relaxation.

The smoothing factor as a function of the error frequency  $(\theta_1, \theta_2)$  for this first order distributed Jacobi relaxation applied to the solution of the linearized problem

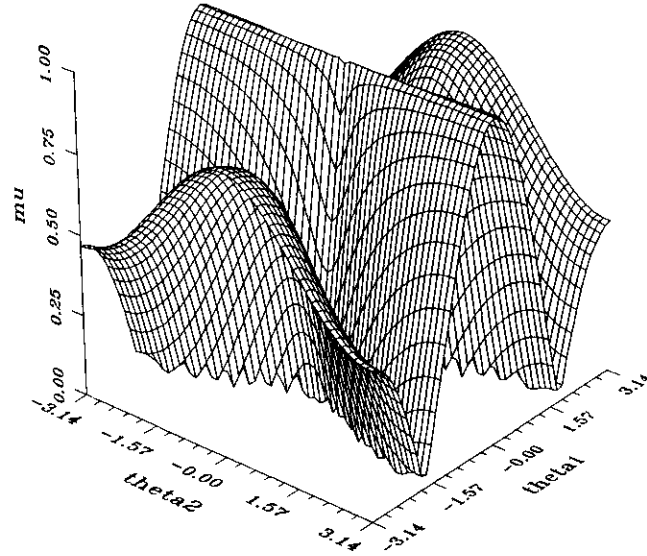


FIGURE 5.1: Amplitude reduction factor  $\mu(\theta_1, \theta_2)$  for Jacobi dipole relaxation on the linearized model problem with  $\epsilon = 0$ .

with  $\epsilon = 0$  is displayed in figure 5.1. From this figure it is obvious that the relaxation is stable since  $\mu(\theta) \leq 1$  for all  $\theta$ . However, with respect to the use of this relaxation in a multilevel solver figure 5.1 reveals two problematic areas, i.e. regions associated with high frequency error components where  $\mu(\theta) \approx 1$ .

Firstly, the relaxation is not effective in reducing the error components that are high frequency in  $X$  direction and low frequency in  $Y$  direction, i.e. the components for which  $(|\theta_1| \approx \pi, |\theta_2| \approx 0)$ . This problem can be overcome easily by means of underrelaxation. This can be seen from figure 5.2, showing the smoothing factor for the same relaxation process if an underrelaxation factor of 0.7 is used. The aforementioned high frequency components are effectively reduced now.

The second problem is that the relaxation is ineffective in reducing low frequency components in  $X$  direction, regardless of the frequency in  $Y$  direction, i.e. the components  $(\theta_1 \approx 0, \theta_2)$ . Comparing figure 5.1 and figure 5.2 shows that underrelaxation does not solve this problem. In fact, the bad reduction of these components is a more fundamental problem. As mentioned before when discussing the simple Gauss-Seidel relaxation it is caused by the fact that for  $\epsilon = 0$  (and by approximation also for small  $\epsilon$ ) equation (5.8) reduces to:

$$\frac{\partial H}{\partial X} = 0 \quad (5.14)$$

Consequently the direct coupling via  $P$  between neighbouring points in  $Y$  direction vanishes. Only the weak and indirect coupling via the elastic deformation integrals remains. Because of this very weak coupling any local relaxation (one point or

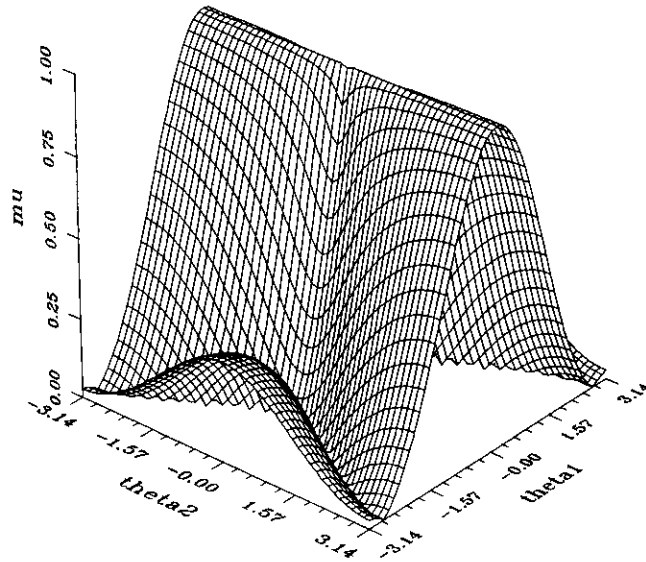


FIGURE 5.2: Amplitude reduction factor  $\mu(\theta_1, \theta_2)$  for Jacobi dipole relaxation on the linearized model problem with  $\epsilon = 0$  and underrelaxation factor 0.7.

distributive, simultaneous or successive displacement) where the changes for  $P$  are calculated from equation (5.9) scanning the gridpoints one by one in some specified order will be inefficient in reducing the aforementioned components. Consequently, if such a relaxation scheme is applied it will result in a tendency of the solutions to oscillate with respect to the  $Y$  direction. For the complete circular contact problem this tendency was indeed observed by Lubrecht, see [L1, page 78], and it is exactly what is reflected in figure 5.1 and figure 5.2. The solution of this problem of weak coupling is a necessity in order to obtain an efficient multilevel solver.

In general there are two possible ways to solve this problem of weak coupling. The first alternative is generally referred to as *semi-coarsening*. The standard choice of a coarser grid having twice the mesh size of the finer grid is entirely based on the fact that the relaxation is effective in reducing high frequency components in *both* spatial directions. In this particular situation one should conclude that this choice can not lead to a fast solver. The relaxation is only effective in reducing error components that are high frequency in  $X$  direction. Consequently after a few relaxations the error will be smooth with respect to the  $X$  direction only and can be represented on a grid that is coarse with respect to this direction only. Subsequently, all techniques discussed in chapter 3 can be applied. For more information on this so-called semi-coarsening the reader is referred to Brandt [B1]. Disadvantages of semi-coarsening are that, assuming a two-dimensional problem, almost twice as much storage is required and that the necessary computing time for a coarse grid correction cycle approximately doubles compared to the situation of coarsening in both directions.

The second, more elegant, solution to the (multilevel) problems related to weak coupling is to apply a line relaxation as is proposed by Brandt, see [B1, page 11]. Instead of visiting the gridpoints one by one in some order, e.g. lexicographic order, and solve the discrete equation at each gridpoint, the discrete equations on a line of points are solved simultaneously. The lines to be solved should be lines in the direction of the strong coupling, i.e. in the situation studied lines in the  $X$  direction (lines of constant  $Y$ ).

At this stage it can be concluded that a relaxation process enabling efficient multilevel solution of the linearized model problem for all  $\epsilon/h^2$  and in particular for  $\epsilon = 0$ , should contain both elements discussed above, i.e. it should combine distributive changes with elements of line relaxation. Assuming for the moment the limiting situation that  $\epsilon = 0$ , such a scheme is derived below. First the equations are written as a system of two equations with two unknowns,  $H$  and  $P$ :

$$H_{i,j} - w_{i,j} = f_{i,j} \quad (5.15)$$

$$(H_{i,j} - H_{i-1,j})/h = 0 \quad (5.16)$$

where  $w_{i,j}$  denotes the discretized elastic deformation integral. Suppose that  $H_{i,j}$  is given and satisfies the second equation. In that case the problem reduces to solving  $P$  from the first equation. Hence, the problem reduces to the basic equation describing the elastic dry contact situation, i.e. see Lubrecht [L2]. This problem can be solved by means of the following second order distributive relaxation scheme: Given an approximation  $\bar{P}_{i,j}$  to  $P_{i,j}$  and the associated integral values  $\tilde{w}_{i,j}$ , all grid points are visited in lexicographic order and changes  $\delta_{i,j}$  to be applied at sites  $(i,j)$  and the adjacent sites  $(i \pm 1, j)$ , and  $(i, j \pm 1)$  according to the following distribution:

$$\frac{1}{4} \delta_{i,j} \begin{bmatrix} 0 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 0 \end{bmatrix} \quad (5.17)$$

are calculated according to:

$$\delta_{i,j} = \frac{f_{i,j} + \tilde{w}_{i,j} - H_{i,j}}{-\frac{2}{\pi^2} \Delta K_{i,j}^{hhhh}} \quad (5.18)$$

where:

$$\Delta K_{i,j}^{hhhh} = K_{i,j}^{hhhh} - (K_{i-1,j}^{hhhh} + K_{i+1,j}^{hhhh} + K_{i,j-1}^{hhhh} + K_{i,j+1}^{hhhh})/4 \quad (5.19)$$

The changes are applied simultaneously after the sweep is completed. As a result of the distributive changes the new approximation  $\bar{P}_{i,j}$  to  $P_{i,j}$  is given by:

$$\bar{P}_{i,j} = \tilde{P}_{i,j} + \delta_{i,j} - (\delta_{i-1,j} + \delta_{i+1,j} + \delta_{i,j-1} + \delta_{i,j+1})/4 \quad (5.20)$$

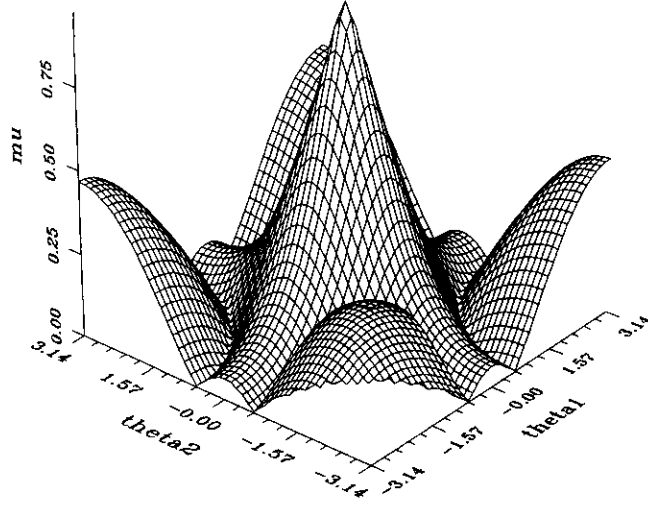


FIGURE 5.3: Amplitude reduction factor  $\mu(\theta_1, \theta_2)$  for Jacobi second order distributed relaxation solving  $P$  from equation 5.15

The relaxation sweep is completed by a multi-summation to obtain all  $\bar{w}_{i,j}$ , i.e. the values of the integrals associated with the new approximation to  $P$ . Because of the distributed changes the accumulated changes in the integrals during one sweep are small and the relaxation process is stable. Moreover, it is very effective in reducing high frequency error components in both directions. This is illustrated in figure 5.3. The asymptotic smoothing rate  $\bar{\mu}$  equals 0.45. Consequently, together with the multilevel techniques described in chapter 3, it provides a fast and stable solver for the problem as is for example clearly demonstrated by Lubrecht [L2].

Hence, given an approximation to the film thickness, a stable and efficient solver for the pressure is available. The next step is to solve  $H$ . For that purpose we consider the second equation of the system, i.e. equation (5.16). This equation should be satisfied everywhere except at the boundaries of the domain. At the boundaries, in particular at the inlet boundary, it does not apply, and the values of  $H$  can be calculated straightforwardly from equation (5.15). Given the values of  $H$  at this boundary associated with an approximate pressure profile, equation (5.16) can be solved in one sweep setting  $\bar{H}_{i,j}$  equal to  $\bar{H}_{0,j}$ .

Obviously, the aforementioned distributive relaxation for solving  $P$  when  $H$  is given, can be combined directly with the solution of  $H$  from equation (5.16). This leads to the following relaxation scheme for the model problem with  $\epsilon = 0$ : Given an approximation  $\bar{P}_{i,j}$  and the associated integral values  $\bar{w}_{i,j}$

- Compute the values of  $\bar{H}$  on the boundary  $X = X_a$ , i.e.  $\bar{H}_{0,j}$ , from equation (5.15).

- Solve  $\bar{H}$  in all interior points from equation (5.16), i.e. set  $\bar{H}_{i,j}$  equal to  $\bar{H}_{0,j}$ .
- Calculate a new approximation  $\bar{P}_{i,j}$  using the distributed relaxation process.
- (Multilevel) Multi-integration to recalculate or update the integrals to obtain  $\bar{w}_{i,j}$ .

With some simple numerical tests it can be shown that this relaxation scheme efficiently reduces all high frequency components. However, due to the interaction with the boundaries when solving  $H$  from equation (5.16), a quantification of the smoothing effect is complex since the relaxation does not map a specific Fourier component onto itself. Nevertheless, it can be shown that the overall performance is reasonably well approximated if the changes of the values of  $H$  are disregarded in the analysis, i.e. by an analysis based on equation (5.15) only. In that case, the predicted asymptotic smoothing rate is 0.45. Consequently a  $W(2,1)$  or  $V(2,1)$  cycle should give an error reduction of one order of magnitude. This reduction is also easily obtained for the complete problem.

This completes the description of a stable relaxation scheme that efficiently reduces high frequency components for the model problem with  $\epsilon = 0$ . Note that the scheme, although in a very simple way, contains both necessary elements mentioned before. This is obvious as far as the distributive changes needed to limit the accumulated changes of the integral values are concerned. However, the scheme also contains the line relaxation element in the form of the simultaneous solution of equation (5.16) in all points on a line  $j$  by setting the value of  $H_{i,j}$  equal to the first value on that line, i.e.  $H_{0,j}$ .

As a next step the aforementioned scheme is extended to a full distributive line relaxation scheme, that is, to visit all lines of constant  $Y$  and to solve at each line simultaneously a new approximation  $\bar{H}_{i,j}$  to  $H_{i,j}$  and all changes  $\delta_{i,j}$  ( $1 \leq i \leq n_x - 1$ ) to be applied distributively from:

$$\bar{H}_{i,j} - \frac{2}{\pi^2} \sum_{k=1}^{n_x-1} \Delta K_{ik,jj} \delta_{k,j} = f_{i,j} + \bar{w}_{i,j} \quad (5.21)$$

$$(\bar{H}_{i,j} - \bar{H}_{i-1,j})/h = 0 \quad (5.22)$$

After all interior lines  $j$  have been visited the changes are applied according to:

$$\bar{P}_{i,j} = \bar{P}_{i,j} + \delta_{i,j} - (\delta_{i-1,j} + \delta_{i+1,j} + \delta_{i,j-1} + \delta_{i,j+1})/4 \quad (5.23)$$

and the new values of the pressures  $\bar{P}_{i,j}$  are used to update or recalculate the integrals. This distributive line relaxation scheme is even more effective in reducing high frequency error components (disregarding the changes of  $H$ :  $\bar{\mu} = 0.2$ ). The only complication is how to solve all changes together with the new values of  $H_{i,j}$  on one line simultaneously. In fact, this requires the solution of a (full) system of



$O(\sqrt{n})$  discrete equations, where  $n$  is the total number of nodes on the grid. However, to obtain the full line relaxation efficiency, the equations need not be solved exactly. In general it is sufficient if the error on that line is reduced by say one or two orders of magnitude. Therefore, since  $\Delta K_{ikjj}^{hhhh}$  decreases very fast with increasing distance  $|i - k|$ , it is sufficient to take into account only the three largest terms of the summation and to solve all  $\delta_{i,j}$  and  $\bar{H}_{i,j}$  from:

$$\bar{H}_{i,j} - \frac{2}{\pi^2} (\Delta K_{iijj}^{hhhh} \delta_{i,j} + \Delta K_{ii-1jj}^{hhhh} \delta_{i-1,j} + \Delta K_{ii+1jj}^{hhhh} \delta_{i+1,j}) = f_{i,j} + w_{i,j} \quad (5.24)$$

$$(\bar{H}_{i,j} - \bar{H}_{i-1,j})/h = 0 \quad (5.25)$$

There are various alternatives enabling a fast solution of all  $\bar{H}_{i,j}$  and  $\delta_{i,j}$  from this reduced system. An effective way is for example by means of decimation. This process and its implementation for the linear model problem considered here are explained in detail in appendix D.

The line relaxation process described above was implemented in a multilevel solver for the model problem with  $\epsilon = 0$  and an error reduction of an order of magnitude per  $V(2,1)$  or  $W(2,1)$  cycle was easily obtained.

Having developed a stable relaxation scheme giving good error smoothing for the extreme situation that  $\epsilon = 0$ , attention is now directed towards the situation where  $\epsilon \neq 0$ . The distributive line relaxation scheme can be easily extended to this situation. In that case the changes  $\delta_{i,j}$  to be applied distributively and the new values  $\bar{H}_{i,j}$  on one line  $j$  must be simultaneously solved from:

$$\bar{H}_{i,j} - \frac{2}{\pi^2} \sum_{k=1}^{n_x-1} \Delta K_{ikjj}^{hhhh} \delta_{k,j} = f_{i,j} + \tilde{w}_{i,j} \quad (5.26)$$

$$\begin{aligned} \epsilon/h^2 \{ & (\tilde{P}_{i-1,j} - \delta_{i-2,j}/4 + \delta_{i-1,j} - \delta_{i,j}/4) + \\ & (\tilde{P}_{i+1,j} - \delta_{i,j}/4 + \delta_{i+1,j} - \delta_{i+2,j}/4) - \\ & 4(\tilde{P}_{i,j} - \delta_{i-1,j}/4 + \delta_{i,j} - \delta_{i+1,j}/4) + \\ & (\tilde{P}_{i,j+1} - \delta_{i,j}/4) + (\tilde{P}_{i,j-1} - \delta_{i,j}/4) \} - \\ & (\bar{H}_{i,j} - \bar{H}_{i-1,j})/h = 0 \end{aligned} \quad (5.27)$$

After all interior lines  $j$  have been visited the changes  $\delta_{i,j}$  are applied distributively and the new values of the pressures, see equation 5.23, are used to update or recalculate the integrals to obtain  $\tilde{w}_{i,j}$ . Again it is not necessary to solve the system exactly to obtain the full line relaxation efficiency and it is sufficient to solve a reduced version of this set of equations. For further details regarding the solution of this system of equations the reader is referred to appendix D. The distributed line relaxation scheme presented here is very well suited for the solution of the model problem for

small  $\epsilon/h^2$ . Also for larger values of  $\epsilon/h^2$  it rather effectively reduces high frequency error components and even in the limiting case of very large  $\epsilon/h^2$ , i.e. the two dimensional Poisson problem, the asymptotic smoothing rate  $\bar{\mu}$  is still 0.6. However, for these large values of  $\epsilon/h^2$  the present relaxation process is outranked by the one point Gauss-Seidel relaxation described at the start of this section, which has an asymptotic smoothing rate of 0.5. An even better alternative for large values of  $\epsilon/h^2$  is a Gauss-Seidel line relaxation with an asymptotic smoothing rate of 0.4. The Gauss-Seidel line relaxation can be described as follows: For each line of constant  $Y$  ( $j$ ) changes  $\delta_{i,j}$  and a new approximation  $\bar{H}_{i,j}$  to  $H_{i,j}$  are solved simultaneously from:

$$\bar{H}_{i,j} - \frac{2}{\pi^2} \sum_{k=1}^{n_x-1} K_{ik,jj} \delta_{k,j} = f_{i,j} + \tilde{w}_{i,j} \quad (5.28)$$

$$\begin{aligned} & \epsilon/h^2 \{ (\tilde{P}_{i-1,j} + \delta_{i-1,j}) + (\tilde{P}_{i+1,j} + \delta_{i+1,j}) - \\ & \quad 4(\tilde{P}_{i,j} + \delta_{i,j}) + \tilde{P}_{i,j+1} + \tilde{P}_{i,j-1} \} - \\ & (\bar{H}_{i,j} - \bar{H}_{i-1,j})/h = 0 \end{aligned} \quad (5.29)$$

When all  $\delta_{i,j}$  for a line  $j$  are solved they are applied immediately:

$$\bar{P}_{i,j} = \tilde{P}_{i,j} + \delta_{i,j} \quad (5.30)$$

These new values of  $P$  on line  $j$  are subsequently used when treating the next line as is characteristic for a successive displacement scheme. In order to obtain the full efficiency it is not necessary to take into account all terms of the summation. Since  $K_{ik,jj}^{hhhh}$  decreases with distance as  $|i-k|^{-1}$  it is, as in the procedure described above, sufficient to take into account only three terms.

As was found for the simple (one point) Gauss-Seidel relaxation the Gauss-Seidel line relaxation, although it does not suffer from the loss of coupling, is not stable for small values of  $\epsilon/h^2$ . Below  $\epsilon/h^2 \approx 0.3$  low frequency error components are amplified and the process diverges.

By combining both the Gauss-Seidel line relaxation and the Jacobi distributed line relaxation an efficient multilevel algorithm for the solution of the model problem can be obtained if:

- On grids where  $\epsilon/h^2 \geq 0.3$  the Gauss-Seidel line relaxation is used.
- On grids where  $\epsilon/h^2 < 0.3$  the Jacobi distributive line relaxation is used.

With a multilevel solver for the linearized problem considered here based on these two relaxations an error reduction of one order of magnitude per  $V(2,1)$  or  $W(2,1)$  cycle was obtained regardless of the value of  $\epsilon$ .

### 5.3 Varying coefficients

The next step leading to a solver for the circular contact problem is to consider equation (5.1) and allow the coefficient  $\epsilon$  to vary over the grid, i.e. to solve the problem:

$$L(P) = \frac{\partial}{\partial X} \left( \epsilon \frac{\partial P}{\partial X} \right) + \frac{\partial}{\partial Y} \left( \epsilon \frac{\partial P}{\partial Y} \right) - \frac{\partial H}{\partial X} = 0 \quad (5.31)$$

on the domain  $\{(X, Y) \in \mathbb{R}^2 | X_a \leq X \leq X_b, -Y_a \leq Y \leq Y_a\}$  with the condition that  $P = 0$  on the boundaries and  $H$  is given by equation (5.5) for different functions  $\epsilon(X, Y)$ . Cavitation is disregarded and the force balance equation plays no role. To approximate the full circular contact problem  $\epsilon(X, Y)$  is chosen as:

$$\epsilon(X, Y) = \begin{cases} (X^2 + Y^2)^3/8 & \text{if } X^2 + Y^2 > 1 \\ 0 & \text{otherwise} \end{cases} \quad (5.32)$$

Equation (5.31) is discretized according to:

$$\begin{aligned} h^{-2} & (\epsilon_{i-1/2,j}(P_{i-1,j} - P_{i,j}) + \epsilon_{i+1/2,j}(P_{i+1,j} - P_{i,j}) + \\ & \epsilon_{i,j-1/2}(P_{i,j-1} - P_{i,j}) + \epsilon_{i,j+1/2}(P_{i,j+1} - P_{i,j})) \\ & - h^{-1}(\bar{\rho}_i H_i - \bar{\rho}_{i-1} H_{i-1}) = 0 \end{aligned} \quad (5.33)$$

$\epsilon_{i-1/2,j}$ ,  $\epsilon_{i+1/2,j}$ ,  $\epsilon_{i,j-1/2}$  and  $\epsilon_{i,j+1/2}$  denote the value of  $\epsilon(X, Y)$  at the *intermediate* locations:

$$\begin{aligned} (X, Y) &= (X_a + (i - 1/2)h, -Y_a + jh), \\ &= (X_a + (i + 1/2)h, -Y_a + jh), \\ &= (X_a + ih, -Y_a + (j - 1/2)h), \text{ and} \\ &= (X_a + ih, -Y_a + (j + 1/2)h) \end{aligned}$$

respectively. In the multilevel algorithm described in section 5.2.1 the relaxation type used on a specific grid depends on the value of  $\epsilon/h^2$ . As mentioned before, relaxation is, by its nature, a local process. Hence different relaxations can be combined on one grid. For the specific problem considered here the two line relaxation processes can be combined in the following way: For a given line of constant  $Y$  the changes  $\delta_{i,j}$  and the new approximations to  $\bar{H}_{i,j}$  are solved from a system of equations, i.e. two equations per gridpoint  $i$ . If at least one of  $\epsilon_{i\pm 1/2,j\pm 1/2}/h^2$  exceeds the value of the crossover point the two equations for this grid point are:

$$\bar{H}_{i,j} - \frac{2}{\pi^2} \sum_{k=1}^{n_x-1} K_{ik,jj} \delta_{k,j} = f_{i,j} + \tilde{w}_{i,j} \quad (5.34)$$

$$\begin{aligned} h^{-2} \{ & (\epsilon_{i-1/2,j}(\tilde{P}_{i-1,j} + \delta_{i-1,j}) + \\ & \epsilon_{i+1/2,j}(\tilde{P}_{i+1,j} + \delta_{i+1,j}) - \\ & (\epsilon_{i,j-1/2} + \epsilon_{i,j+1/2} + \epsilon_{i+1/2,j} + \epsilon_{i-1/2,j}) (\tilde{P}_{i,j} + \delta_{i,j}) + \\ & \epsilon_{i,j+1/2}\tilde{P}_{i,j+1} + \epsilon_{i,j-1/2}\tilde{P}_{i,j-1} \} - \\ (\bar{H}_{i,j} - \bar{H}_{i-1,j})/h & = 0 \end{aligned} \quad (5.35)$$

Otherwise, i.e. if all  $\epsilon_{i\pm 1/2,j\pm 1/2}/h^2$  are smaller than the value of the crossover point, the two equations for this grid point are:

$$\bar{H}_{i,j} - \frac{2}{\pi^2} \sum_{k=1}^{n_x-1} \Delta K_{ikjj}^{hhhh} \delta_{k,j} = f_{i,j} + \tilde{w}_{i,j} \quad (5.36)$$

$$\begin{aligned} h^{-2} \{ & \epsilon_{i-1/2,j}(\tilde{P}_{i-1,j} - \delta_{i-2,j}/4 + \delta_{i-1,j} - \delta_{i,j}/4) + \\ & \epsilon_{i+1/2,j}(\tilde{P}_{i+1,j} - \delta_{i,j}/4 + \delta_{i+1,j} - \delta_{i+2,j}/4) + \\ & \epsilon_{i,j+1/2}(\tilde{P}_{i,j+1} - \delta_{i,j}/4) + \epsilon_{i,j-1/2}(\tilde{P}_{i,j-1} - \delta_{i,j}/4) - \\ & (\epsilon_{i,j-1/2} + \epsilon_{i,j+1/2} + \epsilon_{i+1/2,j} + \epsilon_{i-1/2,j}) \\ & (\tilde{P}_{i,j} - \delta_{i-1,j}/4 + \delta_{i,j} - \delta_{i+1,j}/4) \} - \\ (\bar{H}_{i,j} - \bar{H}_{i-1,j})/h & = 0 \end{aligned} \quad (5.37)$$

In both situations it is sufficient to solve a reduced version of the system of equations, i.e. see appendix D. Once all  $\delta_{i,j}$  and  $\bar{H}_{i,j}$  for the specific line are solved  $\delta_{i,j}$  is added to  $\tilde{P}_{i,j}$  or, when the sweep is completed, to  $\bar{P}_{i,j}$  and partly to its four neighbours depending on the equations solved for the gridpoint  $(i,j)$ .

Using this hybrid relaxation scheme in a multilevel solver for equation (5.33) with  $\epsilon(X,Y)$  given by (5.32), a reduction of the error by almost one order of magnitude (a factor of 8) per coarse grid correction cycle was obtained.

## 5.4 The circular contact problem

In the previous section the coefficient  $\epsilon$  was given as a function of the spatial coordinates  $X$  and  $Y$ . The next step towards a relaxation scheme for the full non-linear circular contact problem is to give  $\epsilon$  as a function of  $H$  and  $P$  instead. This can be done straightforwardly since the description of the relaxation scheme in the previous

section was entirely given in terms of  $\epsilon$ , i.e. instead of using equation (5.32),  $\epsilon_{i,j}$  is calculated according to:

$$\epsilon_{i,j} = \frac{\bar{\rho}(\tilde{P}_{i,j})\tilde{H}_{i,j}^3}{\bar{\eta}(\tilde{P}_{i,j})\lambda}$$

In addition,  $\partial H/\partial X$  is replaced by  $\partial(\bar{\rho}H)/\partial X$  and when applying the pressure changes the cavitation condition is imposed, i.e. negative values of  $P$  are not allowed. The last step is to add the force balance equation to the system. Once every  $s$  relaxations the integration constant  $H_{00}$  is adjusted according to:

$$H_{00} \leftarrow H_{00} - c\left(\frac{2}{3}\frac{\pi}{h^2} - \sum_{i=1}^{n_x-1} \sum_{j=1}^{n_y-1} \tilde{P}_{i,j}\right) \quad (5.38)$$

where  $\tilde{P}_{i,j}$  is the current approximation to the pressure profile. As described for the line contact case in section 4.2.3, the number of relaxations after which the force balance equation is relaxed,  $s$ , must be large enough and the multiplication factor  $c$  must be sufficiently small to avoid unstable oscillations. On the other hand,  $c$  and  $s$  should also be chosen in such a way that the overall convergence is not slowed down too much. The reader is reminded that in the case multiple grids are used the force balance equation is only relaxed on the coarsest grid.

This concludes the description of a relaxation scheme that provides a stable solver for the circular contact problem. Because of the non-linearity underrelaxation is needed. The necessary underrelaxation depends on the load conditions and as experience has shown varies from 0.3 to 1.0 for the distributed changes and from 0.6 to 1.0 for the single changes. In addition, in particular for high loads, see section 4.2.3 and 4.3.5, the mesh size on the grid should be sufficiently small. This subject is discussed in more detail in section 5.5.1.

## 5.5 Application of multilevel techniques

This section deals with the implementation of the relaxation process described above and the multilevel multi-integration algorithm described in chapter 3 into a fast multilevel solver for the circular contact problem. With respect to the treatment of the different equations, the calculation of the *FAS* coarse grid right hand sides, and the coarse grid correction cycle the reader is referred to the description for the line contact case given in section 4.3. Basically the same approach applies to the circular contact problem, hence, the description is not repeated here.

### 5.5.1 Full multigrid

As a first approximation to the solution on the coarsest grid the Hertzian dry contact pressure profile is used. In terms of the dimensionless parameters used in this thesis

this pressure profile is given by:

$$P_{i,j} = \begin{cases} \sqrt{1 - X_i^2 - Y_j^2} & \text{if } |X_i^2 + Y_j^2| < 1 \\ 0 & \text{otherwise} \end{cases} \quad (5.39)$$

To interpolate the solution of a coarse grid to the next finer grid to serve as a first approximation a bi-cubic interpolation is recommended. Subsequently, two  $V(2,1)$  or  $W(2,1)$  cycles are sufficient to obtain a solution that has converged below the level of the truncation error. The  $V(2,1)$  cycles are applied for lightly loaded situations. In the case of high loads the changes of the integration constant related to the force balance equation significantly reduce the reduction factor that can be obtained with a  $V$  cycle. In those situations the more robust  $W$  cycles are used. Another important observation with respect to the solution of highly loaded situations is again to separate the two functions of the coarser grids, i.e. the error correction and the solution approximation, in the *FMG* process as was done in the line contact case, see section 4.3.5.

### 5.5.2 Multilevel multi-integration

The implementation of the multilevel multi-integration algorithm described in chapter 3 in a *FMG* algorithm solving the circular contact problem gives a minor complication. Because in the multilevel calculation of the integrals the coarsening is performed in two steps it requires the introduction of a number of *half-coarsened* grids. These grids are only used for the multilevel multi-integration and are skipped in the solution process. Apart from this complication the implementation does not differ from the approach for the line contact problem described in section 4.3.6. At the first visit on a certain level the summations:

$$\frac{2}{\pi^2} \sum_{k=0}^{n_x} \sum_{l=0}^{n_y} K_{ikjl}^{hhhh} P_{kl} \quad (5.40)$$

are transferred to coarser grids, calculated, and interpolated back (using the correction procedure) to the fine grid. To obtain the film thickness the value of

$$H_{00} + \frac{X_i^2}{2} + \frac{Y_j^2}{2} \quad (5.41)$$

is added to the summation in each point. If the advanced techniques described in section 3.2.7 are used, i.e. if at the following occasions multi-integration on this grid is required only updates for the integrals are calculated, the summations:

$$\frac{2}{\pi^2} \sum_{k=1}^{n_x-1} \sum_{l=1}^{n_y-1} K_{ikjl}^{hhhh} \delta P_{kl} \quad (5.42)$$


---

are transferred to coarser grids, calculated, interpolated back to the fine grid (using the correction procedure), and added to the fine grid summations that were calculated upon arrival at the specific grid for the first time (see also section 4.3.6):

$$\frac{2}{\pi^2} \sum_{k=0}^{n_x} \sum_{l=0}^{n_y} K_{ikjl}^{hhhh} \delta P_{kl}^0 \quad (5.43)$$

## 5.6 Example

To check convergence to the solution of the continuous integro-differential problem and to check the complexity of the algorithm the results obtained for a specific load situation are studied in detail. The values of the Moes dimensionless point contact parameters for this load situation are:  $M = 20$  and  $L = 10$ . This coincides with  $\bar{\alpha} = 9.89$  and  $\lambda = 0.2$ . Upon assuming again  $\alpha = 1.7 \cdot 10^{-8}$  the maximum Hertzian pressure for this load situation is 0.6 GPa. Values of the Hamrock and Dowson non dimensional point contact parameters describing the same load situation are:  $W_{hd} = 1.89 \cdot 10^{-7}$ ,  $U = 1.0 \cdot 10^{-11}$ , and  $G = 4728$ .

### 5.6.1 Calculational details

The solution has been calculated on a uniform grid covering the domain  $\{(X, Y) \in \mathbb{R}^2 \mid -4.5 \leq X \leq 1.5, -3 \leq Y \leq 3\}$ . The *FMG* algorithm described above was used with two  $W(2,1)$  cycles per refinement, each cycle giving an error reduction of almost one order of magnitude. The coarsest grid used in the *FMG* algorithm, level 1, consisted of  $(8+1) \times (8+1)$  nodes and the finest grid consisted of  $(512+1) \times (512+1)$  nodes. Hence, the solution has been calculated using as much as 263.169 nodal points. The calculation has been performed using Roelands equation and assuming a compressible lubricant.

### 5.6.2 Solution

The figures 5.4 and 5.5 show the calculated pressure profile and the associated film shape. It may be obvious that, with the number of nodes used in the present calculations, the usual presentation of the results in a “wire” figure such as figure 1.6, is impossible (it becomes totally black). Therefore, the results are presented using colorgraphics. The different shades represent different angles between the outward normal on the profile and the viewer.

These figures show all characteristic elements of medium to highly loaded EHL circular contact solutions. Firstly, the pressure profile resembles the Hertzian dry contact semi-ellipsoid pressure profile. In the inlet region the pressure gradually builds up to the semi-ellipsoid. The cavitated region is preceded by the three dimensional version of the pressure spike. Note that, when compared to the one dimensional situation this “spike” is a shield of high pressure values “wrapped around”

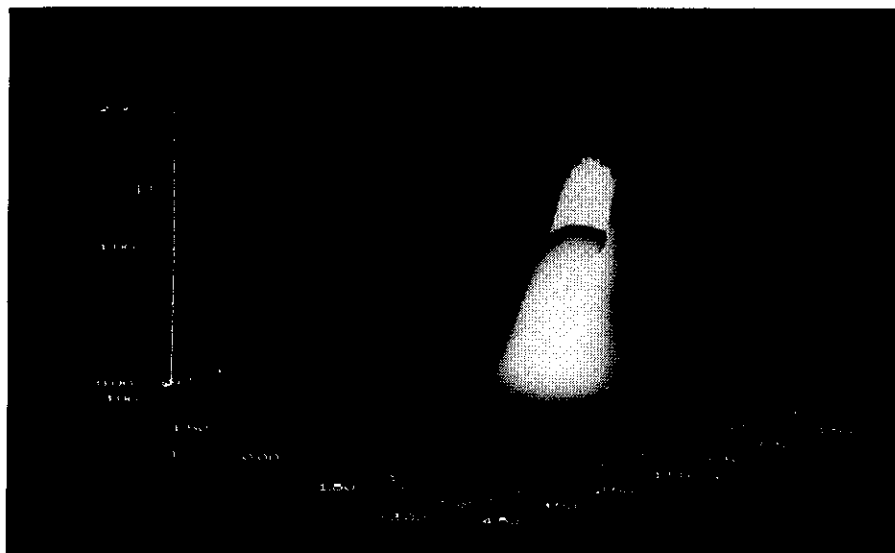


FIGURE 5.4:  $M = 20$ ,  $L = 10$ , circular contact: pressure profile.

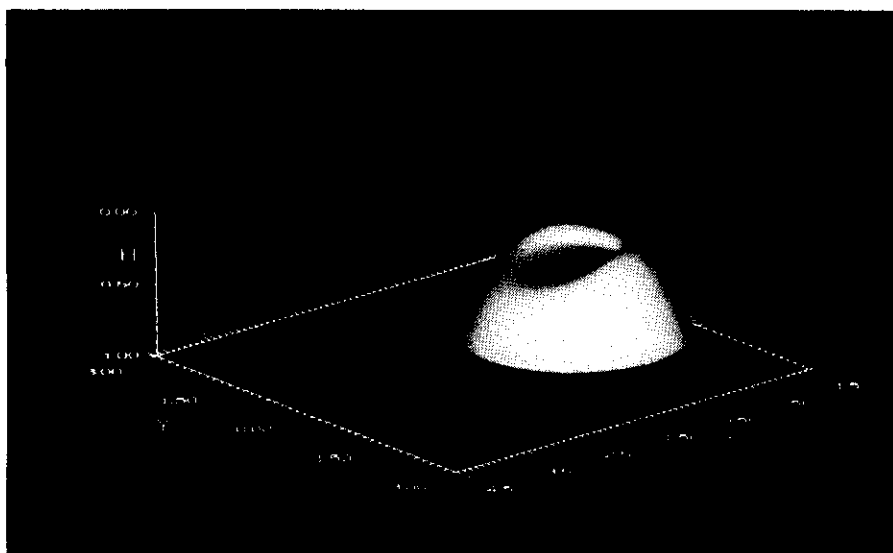


FIGURE 5.5:  $M = 20$ ,  $L = 10$ , circular contact: film thickness.



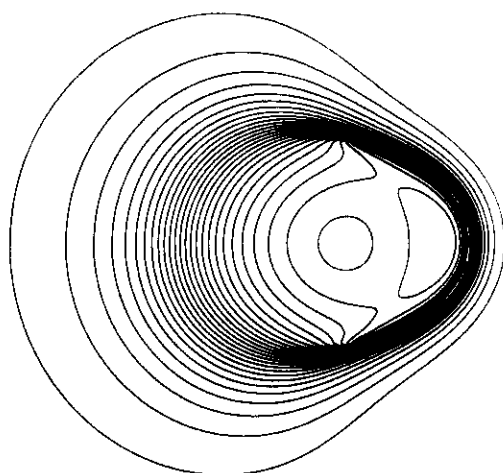


FIGURE 5.6: *Contour line plot of the pressure profile presented in figure 5.4.*

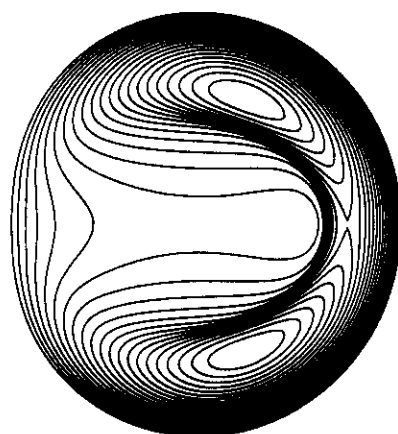


FIGURE 5.7: *Contour line plot of the film thickness presented in figure 5.5.*

the Hertzian contact region. Secondly, the film thickness graph clearly displays the formation of the so called side-lobes or of a horseshoe shaped region. The overall minimum film thickness occurs at these side lobes and its value deviates significantly from the minimum film thickness found on the centerline of the contact.

An alternative way of presenting the calculational results is by means of contour line plots, i.e. drawing lines along which the film thickness or the pressure is constant. Such graphs for the pressure and the film thickness are presented in figures 5.6 and 5.7. In figure 5.6 the pressure spike region can be recognized easily. It is the dark region of large gradients. Also the fact that the spike is in fact a shield wrapped around a more or less semi-ellipsoid is clearly visualised. The film thickness contour graph (figure 5.7) clearly displays the horseshoe shaped region, i.e. the side lobes. Note that the region of large gradients in the film thickness in figure 5.7 coincides with the region of large gradients in the pressure in figure 5.6.

### 5.6.3 Convergence

To check if the solution obtained with the *FMG* algorithm using  $2 W(2,1)$  cycles per level has converged below the level of the discretization error the procedure outlined in section 3.1.8 is used and the solution has been recalculated using an additional  $W(2,1)$  cycle. The calculated values of  $ERR[h, H]$  for the different levels for both two as well as three  $W(2,1)$  cycles are displayed in table 5.1. The error in the solution obtained using three cycles is about 8 times smaller than the error in the solution obtained using two cycles. Hence, as the calculated values of  $ERR[h, H]$  for both situations hardly differ, it is concluded that  $2 W(2,1)$  cycles are sufficient to obtain a converged solution.

Table 5.1 shows that the ratio between the subsequent values of  $ERR[h, H]$  calculated on the different levels approximates a factor of 2. This was also found for the line contact problem, see section 4.4, and it is exactly what could be expected because of the first order upstream discretization of the wedge term in Reynolds' equation. Only in very lightly loaded situations where the wedge term is small compared to the Poiseuille terms second order convergence can be observed, see also Lubrecht [L1].

k	$n_x + 1$	$n_y + 1$	$2 W(2,1)$	$3 W(2,1)$
2	17	17	3.898	3.890
3	33	33	1.024	1.032
4	65	65	$3.192 \cdot 10^{-1}$	$3.203 \cdot 10^{-1}$
5	129	129	$1.216 \cdot 10^{-1}$	$1.215 \cdot 10^{-1}$
6	257	257	$5.486 \cdot 10^{-2}$	$5.463 \cdot 10^{-2}$
7	513	513	$2.713 \cdot 10^{-2}$	$2.717 \cdot 10^{-2}$

Table 5.1:  $ERR(k, k-1)$  using  $2 W(2,1)$  and  $3 W(2,1)$  cycles

k	$n_x + 1$	$n_y + 1$	$H_m$	$H_c$	$H_c/H_m$
1	9	9	$9.813 \cdot 10^{-2}$	5.045	51.42
2	17	17	$3.242 \cdot 10^{-1}$	1.178	3.633
3	33	33	$3.587 \cdot 10^{-1}$	$4.568 \cdot 10^{-1}$	1.274
4	65	65	$3.552 \cdot 10^{-1}$	$4.889 \cdot 10^{-1}$	1.376
5	129	129	$3.492 \cdot 10^{-1}$	$5.023 \cdot 10^{-1}$	1.438
6	257	257	$3.446 \cdot 10^{-1}$	$4.983 \cdot 10^{-1}$	1.446
7	513	513	$3.421 \cdot 10^{-1}$	$4.950 \cdot 10^{-1}$	1.447

Table 5.2: Minimum and central film thickness as a function of the level.

From a sufficiently small mesh size onwards this first order convergence also shows if the value of a characteristic parameter such as the minimum film thickness or the central film thickness is monitored as a function of the mesh size (level). For example, table 5.2 gives the value of the minimum and central film thickness as a function of the mesh size for the load situation considered here. Table 5.2 also displays the ratio between the minimum and central film thickness.

When comparing the convergence behaviour displayed in table 5.2 to the results for the line contact problem presented in chapter 4 the first order convergence is less pronounced. This is obviously caused by the additional dimension of the point contact problem as a result of which solutions with the same number of nodes as in the line case are in fact less accurate.

## 5.7 Computing times

To conclude this example the computing time required to obtain a solution is studied. Table 5.3 presents computing times used by the *FMG* algorithm with two  $V(2,1)$  and two  $W(2,1)$  cycles per refinement as a function of the number of nodes. The first column is representative for lightly loaded situations whereas the second column is representative for moderate to highly loaded situations. The data displayed in table 5.3 is also shown in figure 5.8. For reasons of comparison in this figure also a line representing a computing time of  $O(n)$  and a line representing a computing time of  $O(n^2)$  are drawn. Obviously the complexity of the algorithm is very close to  $O(n)$ . Halving the mesh sizes, i.e. increasing the number of nodes by a factor of 4, the computing time increases by a factor of 4. The fact that it is actually  $O(n \ln n)$  can not even be recognized.

In addition to these results figure 5.9 displays the ratio of the computing time needed by the *FMG* algorithm with 2  $W(2,1)$  cycles and 2  $V(2,1)$  cycles respectively. Neglecting the larger number of relaxations on the coarsest grid and the intergrid transfers the ratio between the number of Work Units required by a *FMG* algorithm with  $V(\nu_1, \nu_2)$  cycles and the number of Work Units in the case of  $W(\nu_1, \nu_2)$  cycles should be 1.5 for a 2 dimensional problem.

k	$n_x + 1$	$n_y + 1$	2 V(2,1)	2 W(2,1)
2	17	17	9.5	9.5
3	33	33	22.1	40.3
4	65	65	1:23	2:44
5	129	129	4:51	10:40
6	257	257	17:35	41:27
7	513	513	1:07:53	2:41:25

Table 5.3: Computing time as a function of the number of nodes for the FMG algorithm with two V(2,1) cycles and two W(2,1) cycles.

This can be shown with the type of analysis presented in section 3.1.6, i.e using the theory of geometrical series, see also [B2]. Figure 5.9 shows that the ratio obtained with the presented algorithm exceeds this value and amounts to approximately 2.5. This deviation from the theoretically expected value can be ascribed to the additional work included to solve the force balance equation (larger number of relaxations performed on the coarsest grid) and the relatively high cost of the multi-integrations on the first few levels.

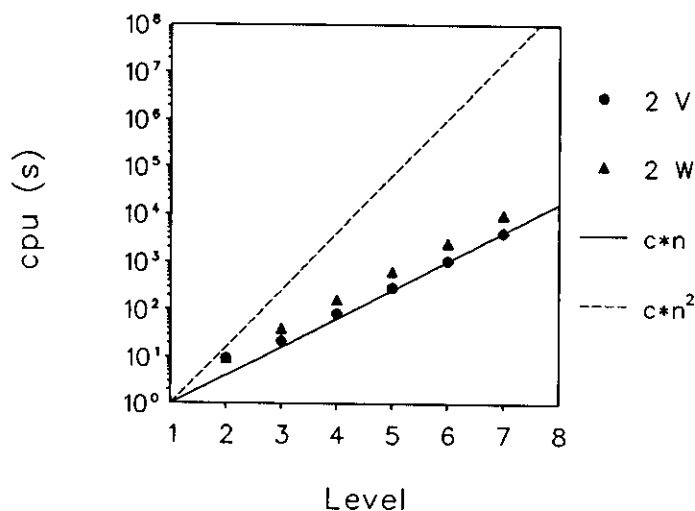


FIGURE 5.8: Computing time on a HP 9000/845s computer for the FMG algorithm with two V(2,1) cycles and two W(2,1) cycles.

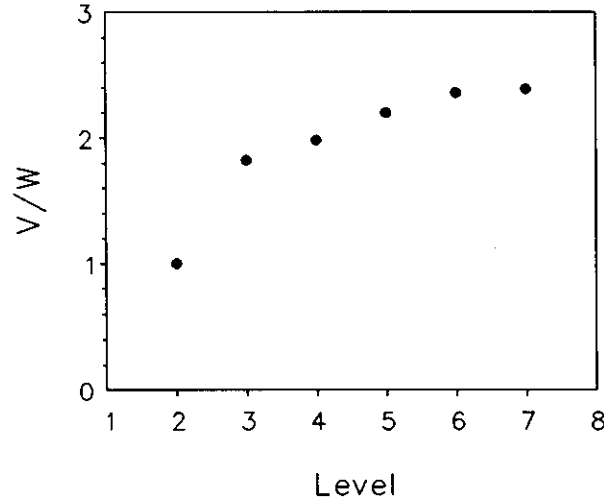


FIGURE 5.9: Ratio of the computing time used by the FMG algorithm with 2  $W$  and 2  $V$  cycles per refinement as a function of the level.

## 5.8 Conclusion

Using distributive relaxation techniques together with the multilevel solution and multi-integration techniques described in chapter 3 an algorithm for the fast solution of the EHL circular contact problem has been developed. It is demonstrated that the algorithm is of very low complexity, i.e.  $O(n \ln n)$  with  $n$  being the number of nodes on the grid. Consequently, it allows solution of the problem with a large number of nodes, e.g. 250.000 on a mini-computer. Moreover, as will be shown in chapter 9 the algorithm is very stable.

The presented algorithm does not only apply to the circular contact situation to which this thesis is restricted. It can also be used to solve elliptical contact situations.

## References

- [B1] **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, W. Germany.
  - [B2] **Briggs, W.L.**, 1987, "A Multigrid Tutorial," SIAM, Philadelphia, Pennsylvania, ISBN 0-89871-221-1
  - [C1] **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. Lond.*, A **397**, 245-269.
  - [C2] **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. Lond.*, A **397**, 271-294.
  - [H1] **Hamrock, B.J., and Dowson, D.**, 1976, "Isothermal elastohydrodynamic lubrication of point contacts, part 1-Theoretical Formulation," *ASME JOT*, **98**, 223-229.
  - [L1] **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.
  - [L2] **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*, to appear.
  - [Z1] **Zhu, D., and Cheng, H.S.**, 1988, "Effect of surface roughness on the point contact EHL," *ASME JOT*, **110**, 32-37.
-

## Chapter 6

### Line contact results: Smooth surfaces

To explore the possibilities of the algorithm described in chapter 4 it has been applied to a parametric study of the “standard” EHL line contact problem, i.e. the situation in which the surfaces are perfectly smooth. In section 4.4 it has already been demonstrated that the low complexity of the algorithm enables the solution of this problem using a large number of nodes which is a necessity in order to study surface roughness effects and point contact problems. The results presented in this chapter show that the algorithm also meets a second requirement for such studies, i.e. it is very stable.

The problem is solved for a variety of load conditions, including extremely high loads. Two subjects are discussed in detail. First the minimum gap width as a function of the governing parameters is studied. Based on the results of the calculations the formula for the minimum film thickness as presented by Lubrecht et al. [L1] has been improved. The modified formula is presented and its predictions are compared with predictions of formulas presented by several other authors. Next the so-called pressure spike is discussed. The occurrence and the height of this second maximum in the pressure profile and its relation with the lubricant behaviour assumed in the calculations are addressed in section 6.3.

#### 6.1 Solutions

With varying values of the dimensionless parameters  $M$  and  $L$  different pressure profiles and film shapes are found. Figure 6.1 gives a global indication of the changes in the pressure profile and film shape with varying  $M$  and  $L$ . These variations are studied in more detail in the following sections.

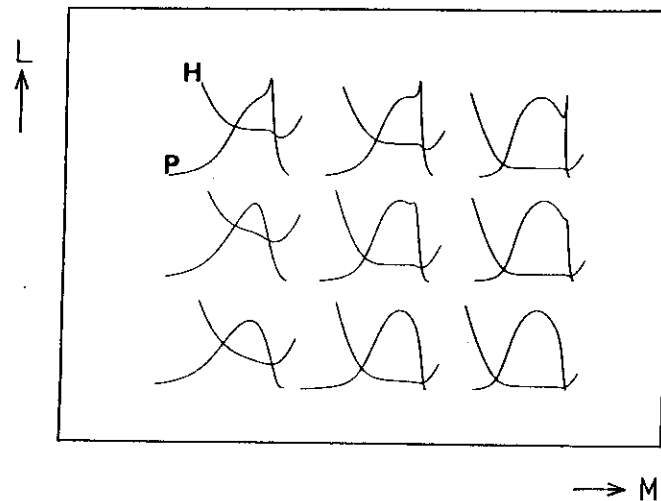


FIGURE 6.1: *Characteristic pressure distributions and film shapes.*

### 6.1.1 Computational details

All solutions presented here have been calculated using Roelands' equation and assuming a compressible lubricant. The *FMG* algorithm described in chapter 4 was used with two  $W(2,1)$  cycles per refinement. The coarsest grid employed in the calculations consisted of 15 nodal points. In most situations the calculational domain extended from  $X_a = -4$  to  $X_b = 1.5$ . However, in some lowly loaded situations a larger inlet region was needed in order to avoid "numerically starved lubrication". Convergence of the solutions was checked as described in section 4.4.3.

### 6.1.2 Varying $M$

Figure 6.2 displays the calculated (dimensionless) pressure profiles for six values of  $M$  and constant  $L$  ( $L = 10$ ). The associated (dimensionless) film shapes are presented in figure 6.3. With increasing load the pressure profile approximates the semi-elliptical Hertzian pressure and deviates from this profile only in the inlet region of pressure generation and in the outlet region where the pressure spike occurs. Furthermore, the width of the spike region and the inlet region narrows down. Here it is recalled that these regions are the boundary layers between the parts of the domain where the problem behaves as a differential problem and the region where the integral aspect dominates. All solutions have been calculated using 28673 nodal points. The resulting local nodal density in the spike region is not always sufficient for the spike height to have converged; in particular for  $M=50$  and  $M=100$  its height is underestimated.



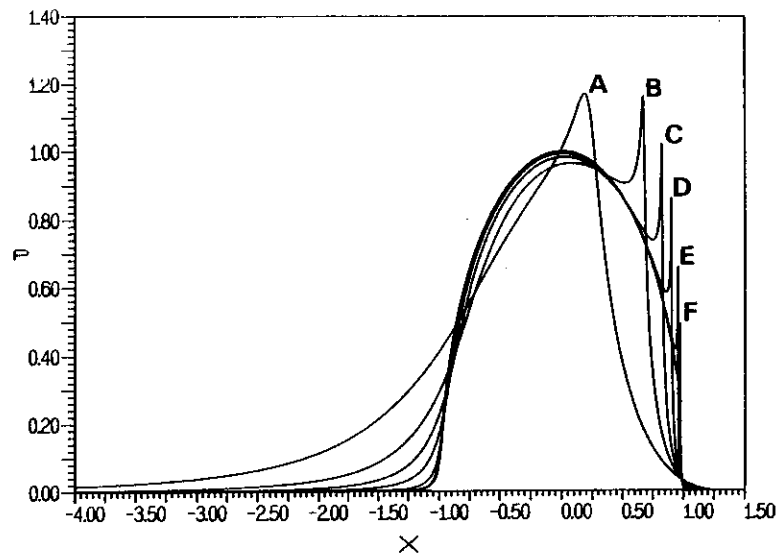


FIGURE 6.2: Pressure profile for  $M=2$  (A), 5 (B), 10 (C), 20 (D), 50 (E) and 100 (F), and  $L = 10$ .

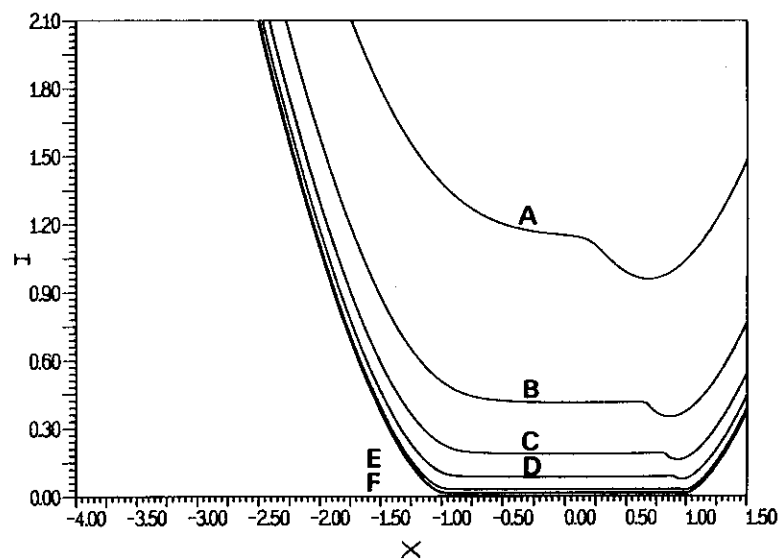


FIGURE 6.3: Film shapes associated with the pressure profiles of figure 6.2.

The results presented in figure 6.3 show that the film thickness and particularly its minimum value, decreases monotonously with increasing  $M$  (load). Meanwhile the effect of the elastic deformation increases as can be observed from the increasing width of the region of nearly uniform film thickness, i.e. the part of the domain where the integral aspect of the problem dominates ( $d(\bar{p}H)/dX \approx 0$ ). This region increases until it covers nearly the entire Hertzian dry contact region  $-1 \leq X \leq 1$ . Assuming  $\alpha = 1.7 \cdot 10^{-8}$  the maximum Hertzian pressure for the situations presented in figure 6.1 and 6.2 varies from 0.3 to 2.35 GPa.

### 6.1.3 Varying $L$

Figure 6.4 displays the calculated (dimensionless) pressure profiles for five values of  $L$  and constant  $M$  ( $M = 20$ ). The associated (dimensionless) film shapes are presented in figure 6.5. The pressure profiles presented in figure 6.4 are difficult to distinguish. The main differences between the various pressure profiles occur in the outlet region as is illustrated by the enlargement of the five pressure profiles in this region displayed in figure 6.6. With increasing  $L$  a spike gradually develops. This subject is discussed in detail in section 6.3. Again the solutions have been calculated using 12 levels ( $n + 1 = 28673$ ). The resulting nodal density in the spike region is not large enough to show convergence of the spike for the  $L = 25$  solution. Obviously, with increasing  $L$  the film thickness increases, illustrating the beneficial effect of a viscosity increasing with pressure on fluid film formation. Assuming  $\alpha = 1.7 \cdot 10^{-8}$  the solution for  $M = 20$  and  $L = 25$  represents a maximum Hertzian pressure of 2.6 GPa.

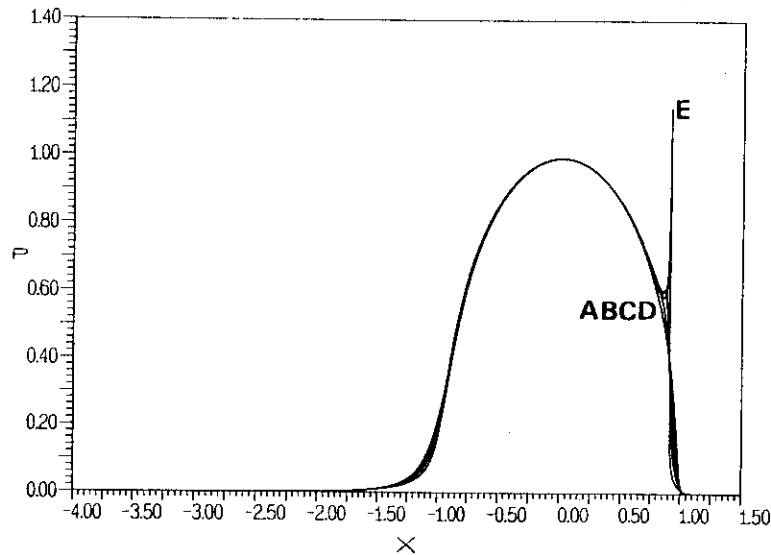


FIGURE 6.4: Pressure profile for  $M=20$ ,  $L=0$  (A), 2.5 (B), 5 (C), 10 (D), 25 (E).

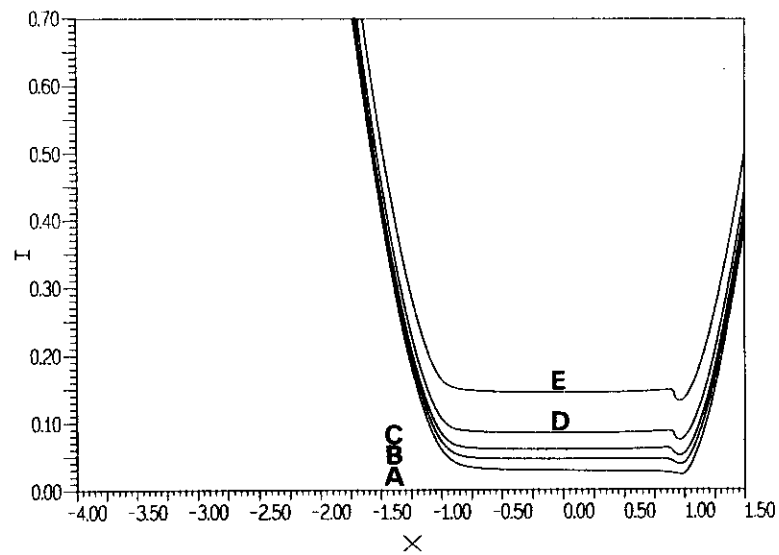


FIGURE 6.5: *Film shapes associated with the the pressure profiles presented in figure 6.4.*

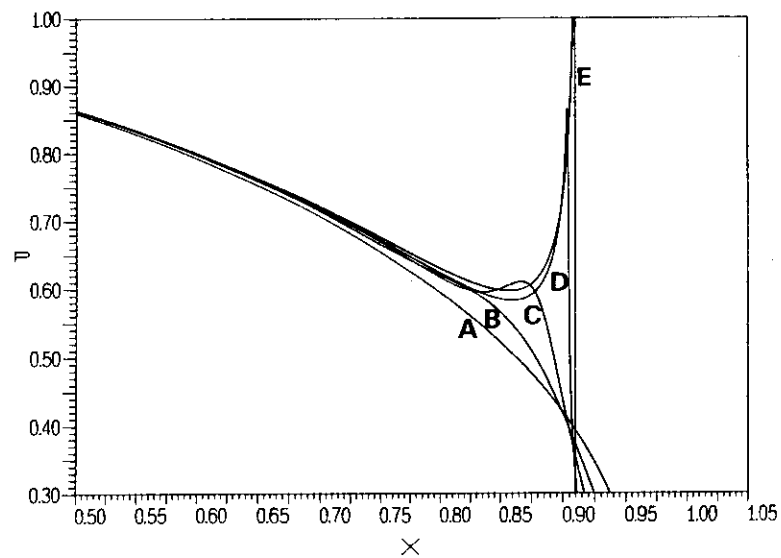


FIGURE 6.6: *Enlargement of the pressure profiles presented in figure 6.4 in the outlet region.*

### 6.1.4 High loads

Here it is recalled that one of the purposes of the present study was to develop an algorithm that enables the analysis of the effects of surface features, e.g. indentations, bumps, waviness, surface roughness, on the pressure profile and film thickness. This kind of information is of particular interest for surface fatigue life predictions of concentrated contacts. The study of these effects for realistic loading situations, i.e. maximum pressures ranging up to 3.0 GPa, requires an extremely stable algorithm since these surface features put an additional demand on the stability of the algorithm, that is they cause small regions of extremely small coefficient values. To test out if the algorithm presented in chapter 4 provides the necessary stability to carry out the aforementioned studies, the smooth surface problem was solved for some extreme loading conditions, for example  $M = 100$  and  $L = 25$ .

This load situation is equivalent to  $W = 2.8 \cdot 10^{-3}$ ,  $U = 3.9 \cdot 10^{-10}$  and  $G = 4730$  in terms of the Dowson and Higginson dimensionless parameters and with  $\bar{\alpha} = 99.74$  and  $\lambda = 3.70 \cdot 10^{-4}$ . With  $\alpha = 1.7 \cdot 10^{-8}$ , the maximum Hertzian dry contact pressure for this situation is 5.9 GPa. Figure 6.7 presents the pressure profile and film shape calculated using 28673 nodal points.

Once more it is emphasized that solutions of the smooth surface EHL line contact problem for such extremely high loads in itself are of little practical relevance for many reasons. Among things, the viscosities are unrealistically high. Moreover, at such high pressures the lubricant will most likely behave as a solid instead of a fluid. Furthermore, in the case of steel surfaces beyond 3.3 GPa gross plastic deformation

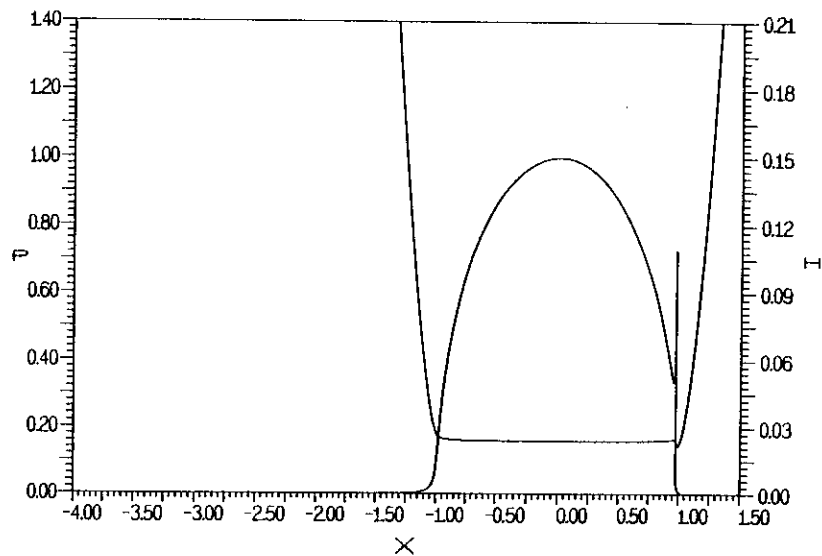


FIGURE 6.7: Pressure profile and film thickness for  $M = 100$ ,  $L = 25$ .

of the raceways will occur which is also not included in the model. However, the fact that the smooth surface problem can be solved for such high loads, (extremely small coefficient values in the entire contact region) gives confidence that the algorithm will also enable the study of the aforementioned more complex EHL situations (extremely small coefficients locally).

## 6.2 Minimum film thickness

From the first theoretical studies of fluid film lubrication onwards one of the main subjects of interest has been the prediction of the minimum film thickness in the conjunction. An interest that is easily explained from a designers point of view. In general, concentrated contacts should run with minimum amount of friction and wear. Hence, it is important to design the contact in such a way that, under the given operating conditions, the surfaces will be fully separated by the lubricant film. It is generally assumed that the surfaces will be fully separated if the minimum gap width calculated for the smooth surface situation is sufficiently large compared to a sum of the roughnesses of the opposing surfaces. However, in particular in EHL contact situations this criterion is very crude. The elastic deformation of the roughness or more in general the deformation of the microgeometry and its effect on the minimum film thickness is not accounted for. This deformation has not necessarily a positive effect on the minimum film thickness as one might expect at first glance. The effect will strongly depend on the geometry of the roughness or surface feature and its position in the contact region as will be shown in chapter 7.

### 6.2.1 Asymptotic solutions

With respect to the minimum film thickness in the conjunction three asymptotic situations are generally distinguished. First the solution for rigid surfaces and an isoviscous, incompressible lubricant, as calculated by Martin [M1] and Gumbel [G1] is described. Expressed in the Moes dimensionless parameters this asymptotic solution reads:

$$H_{min} = 2.45 M^{-1} \quad (6.1)$$

Secondly the elastic-isoviscous asymptote is given, i.e. elastic surfaces and an isoviscous, incompressible lubricant. According to Moes [M2] this asymptote can be written as:

$$H_{min} = 2.05 M^{-1/5} \quad (6.2)$$

This equation merges with the Herrebrugh solution [H1] and is of particular interest for so called *soft* EHL contacts, e.g. the contact between a steel shaft and a rubber

seal or the contact in human joints where, although the pressures in the lubricant film are relatively small, the film shape is dominated by the elastic deformations. Finally there is the Ertel [E1] and Grubin [G2] solution. In fact Ertel and Grubin calculated the film thickness in the center of the contact region but assuming the minimum film thickness to be 3/4 times this central film thickness the asymptote reads, e.g. Terril [T1]:

$$H_{min} = 0.99 M^{-1/8} L^{3/4} \quad (6.3)$$

This formula especially applies to highly loaded contacts between steel surfaces lubricated with a mineral oil (exponential viscosity pressure relation).

### 6.2.2 Film thickness formulas

For design purposes, several researchers have presented formulas predicting the minimum gap width, or a dimensionless minimum film thickness parameter, as a function of the governing parameters. To the author's knowledge Dowson and Higginson [D1] were the first to present such a formula:

$$H'_m = 1.60 G^{0.6} U^{0.7} W^{-0.13} \quad (6.4)$$

where  $H'_m$ ,  $G$ ,  $U$ , and  $W$  are dimensionless parameters defined according to:

$$\begin{aligned} H'_m &= h_{min}/R \\ G &= \alpha E' \\ W &= w/(E' R) \\ 2U &= (\eta_0 u_s)/(E' R) \end{aligned}$$

This set of parameters, although widely used by researchers in the field of EHL is not used throughout this thesis. As was explained in chapter 2 the Moes dimensionless parameters are used because they allow presentation of all film thickness results in one diagram  $H_{min}(M, L)$ . Strictly this only applies if the Barus' equation is used and the lubricant is assumed to be incompressible. When using Roelands' equation and assuming a compressible lubricant such a diagram is restricted to some fixed values of the lubricant parameters  $\alpha$  and  $z$ . The Moes dimensionless parameters  $H_{min}$ ,  $M$  and  $L$  are related to the Dowson and Higginson parameters according to:

$$\begin{aligned} H_{min} &= H'_m (2U)^{-1/2} \\ M &= W (2U)^{-1/2} \\ L &= G (2U)^{1/4} \end{aligned}$$

Equation (6.4) is based on a curve-fit of the results obtained from numerical calculations. Formulas obtained in a similar way were presented by Hamrock and Jacobson [H2]:

$$H'_m = 3.07 G^{0.57} U^{0.71} W^{-0.11} \quad (6.5)$$

and recently by Pan and Hamrock [P2]:

$$H'_m = 1.714 G^{0.568} U^{0.694} W^{-0.128} \quad (6.6)$$

In their paper the constant is given as 1.174. However, to obtain the predictions presented in the paper its value should be 1.714.

A disadvantage of the formulas (6.4) to (6.6) is that they are only accurate within a limited range of values of the governing parameters. In particular for small  $L$  values the predictions are rather inaccurate. A more accurate formula can be obtained if the asymptotic solutions are taken into account. Such a formula has been presented by Lubrecht et al. [L1]. Using two parameters  $s$  and  $r$ , the three asymptotic solutions are merged into one formula predicting the minimum film thickness over the entire parameter range. The parameters  $s$  and  $r$  take care of a smooth transition in the region between the asymptotes. The formula reads:

$$H_{min} = \left[ \underbrace{\{(0.99 M^{-1/8} L^{3/4})^r\}}_{\text{Ertel, Grubin}} + \underbrace{\{(2.05 M^{-1/5})^r\}^{s/r}}_{\text{Elastic/Isoviscous}} + \underbrace{\{(2.45 M^{-1})^s\}^{1/s}}_{\text{Rigid/Isoviscous}} \right]^{1/s} \quad (6.7)$$

where:

$$s = 4 - \exp(-L/2) - \exp(-2/M) \quad (6.8)$$

and:

$$r = \exp\{1 - 4/(L + 5)\} \quad (6.9)$$

In equation (6.7) the aforementioned asymptotic solutions are indicated separately. At first sight equation (6.7) may seem complex compared to the equations (6.4) through (6.6). However, a closer look reveals that the structure is quite straightforward and allows easy programming on a pocket or personal computer.

### 6.2.3 Film thickness diagram

Figure 6.8 presents a plot of the dimensionless minimum film thickness parameter  $H_{min}$  as a function of  $M$  and  $L$ . The drawn lines represent the predictions obtained from equation (6.7) with  $s$  and  $r$  from (6.8) and (6.9) respectively. The aforementioned asymptotic solutions are indicated separately. The markers indicate the presently calculated values. The results rather closely match the predictions of equation (6.7) in the *asymptotic* regions. However, in the intermediate region, in particular for large  $L$ , significant deviations occur between the predicted values and the calculational results. The dashed lines in figure 6.8 are lines of constant  $\bar{\alpha}$  which are drawn using equation (6.7). Given a value of  $\alpha$  these curves represent lines along which the maximum Hertzian pressure  $p_h$  is invariant.

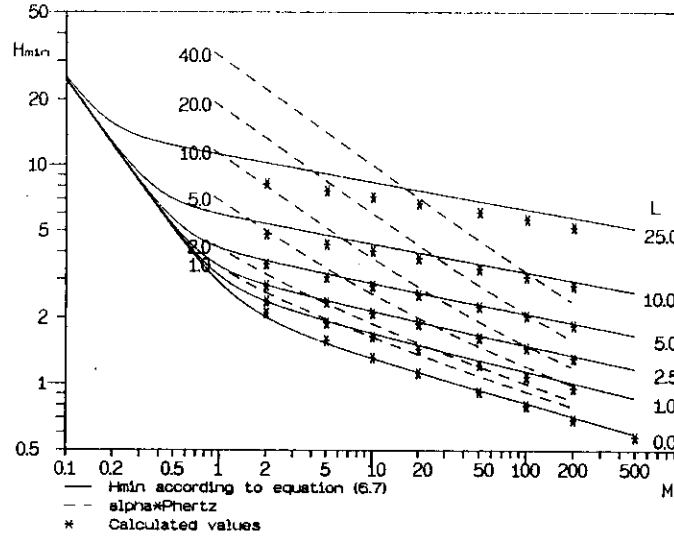


FIGURE 6.8: Calculated values of the dimensionless minimum film thickness parameter  $H_{min}$  (\*) as a function of  $M$  and  $L$ , the drawn lines give the predicted values according to equation (6.7). The dashed lines represent lines of constant  $\bar{\alpha}$

#### 6.2.4 Modified film thickness formula

Based on the results of the present numerical calculations Moes [M3] improved the minimum film thickness formula (6.7). The modified equation reads:

$$H_{min} = \left[ \left\{ (0.99 M^{-1/8} L^{3/4} t)^r + (2.05 M^{-1/5})^r \right\}^{s/r} + (2.45 M^{-1})^s \right]^{1/s} \quad (6.10)$$

where  $r$ ,  $s$  and  $t$  are given by:

$$r = \exp\{1 - 3/(L + 4)\} \quad (6.11)$$

$$s = 3 - \exp(-1/(2M)) \quad (6.12)$$

and:

$$t = 1 - \exp\left\{-3.5 \frac{M^{1/8}}{L^{1/4}}\right\} \quad (6.13)$$

In figure 6.9 the predictions of equation (6.10) with  $r$ ,  $s$  and  $t$  given by (6.11), (6.12) and (6.13) respectively, are compared with the calculational results. In particular for the larger  $L$  values, i.e.  $L \geq 10$ , the predictions agree much closer with the results of the numerical calculations.



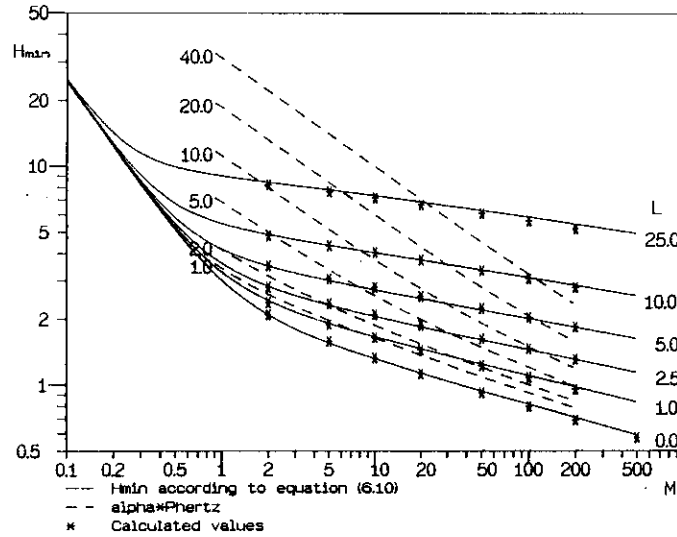


FIGURE 6.9: Calculated values of the dimensionless minimum film thickness parameter  $H_{min}$  (\*) as a function of  $M$  and  $L$ , the drawn lines give the predicted values according to equation (6.10). The dashed lines represent lines of constant  $\bar{\alpha}$

### 6.2.5 Comparison of different film thickness formulas

In the preceding sections different minimum film thickness formulas were reviewed. In this section the values predicted by these formulas are compared with the results of numerical calculations for three load situations. The first situation is more or less representative for a highly loaded contact between steel surfaces, i.e.  $M = 200$  and  $L = 10$  or in terms of the Dowson and Higginson parameters  $U = 1.0 \cdot 10^{-11}$ ,  $G = 4730$ , and  $W = 8.94 \cdot 10^{-4}$ . The values of  $\bar{\alpha}$  and  $\lambda$  for this case are :  $\bar{\alpha} = 56.42$ ,  $\lambda = 9.25 \cdot 10^{-5}$ . With  $\alpha = 1.7 \cdot 10^{-8}$  the maximum Hertzian pressure is 3.3 GPa. Figure 6.10 displays the calculated dimensionless pressure profile and film shape for this load situation. Table 6.1 gives the value of the Moes dimensionless minimum film thickness parameter as a function of the number of nodes. Furthermore, table 6.2 gives the value  $H_{min}$  predicted by the different formulas and the difference between these predictions and the calculational result. For the calculation of this difference the value of  $H_{min}$  calculated on level 12 was taken as a reference.

It appears that in this particular situation all equations give an accurate prediction. This is no surprise since the formulas (6.4) to (6.6) were mainly derived from calculational results for  $L \approx 10$  and also the assumptions that led to the Ertel and Grubin asymptotic solution, equation (6.3), are characteristic for the present load situation.

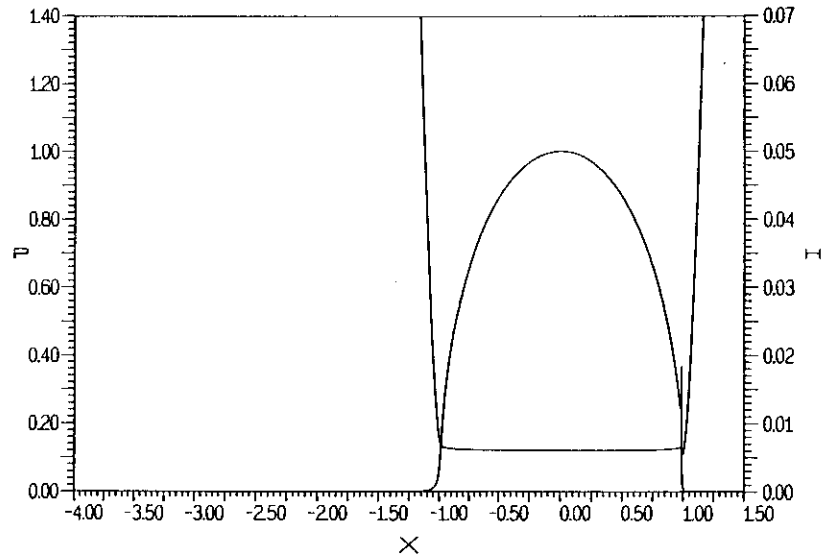


FIGURE 6.10: Calculated pressure profile and film shape for  $M=200$  and  $L=10$ .

level	$n+1$	$H_{min}$
6	449	2.982
7	897	2.903
8	1793	2.873
9	3585	2.845
10	7169	2.827
11	14337	2.820
12	28673	2.816

Table 6.1: Convergence of  $H_{min}$  as a function of the number of nodes for  $M = 200$  and  $L = 10$ .

source:	$H_{min}$	Difference (%)
Ertel, Grubin	2.90	3.0
Dowson and Higginson	2.85	1.4
Hamrock and Jacobson	2.86	1.8
Pan and Hamrock	2.67	4.9
Lubrecht et al.	2.95	5.0
Moes	2.87	2.1

Table 6.2: Comparison of  $H_{min}$  predicted by different formulas with calculated result for  $M = 200$  and  $L = 10$ .

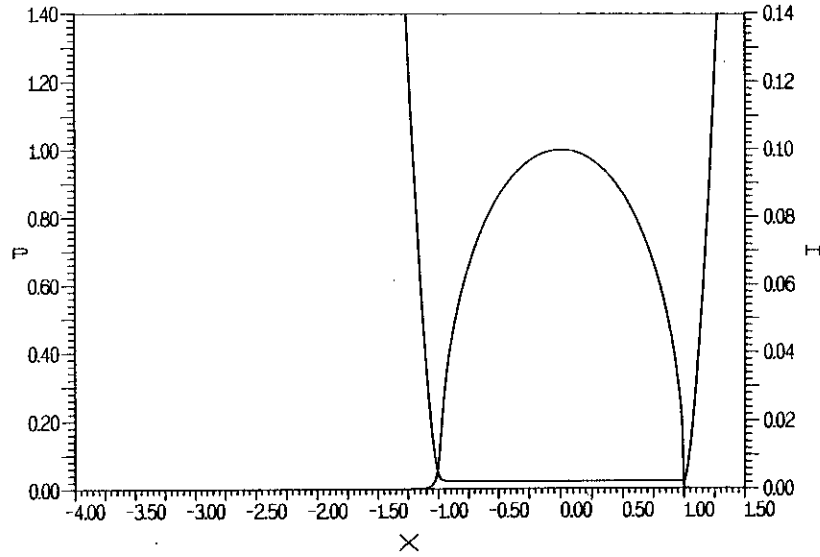


FIGURE 6.11: Calculated pressure profile and film shape for  $M = 200$  and  $L = 1$ .

The second situation is representative for soft EHL contacts, i.e. seals etc. Again  $M = 200$  but this time with  $L = 1$ . Using the Dowson and Higginson parameters this load situation can be described by :  $U = 1.0 \cdot 10^{-11}$ ,  $G = 473$  and  $W = 8.94 \cdot 10^{-4}$ . The values of  $\bar{\alpha}$  and  $\lambda$  for this case are  $\bar{\alpha} = 5.64$  and  $\lambda = 9.25 \cdot 10^{-5}$ . Hence, assuming  $\alpha = 1.7 \cdot 10^{-8}$  the maximum Hertzian pressure is only 0.3 GPa. The calculated pressure profile and film thickness are plotted in figure 6.11. Note that, in spite of the relatively low maximum Hertzian pressure, the film shape is dominated by the elastic deformations (large region of nearly constant film thickness). This is characteristic for soft EHL contacts.

The calculated value of  $H_{min}$  as a function of  $n$  is given in table 6.3. Furthermore, table 6.4 shows the predictions of the different formulas and the differences between these predictions and the numerically calculated result.

level	n+1	$H_{min}$
6	449	1.049
7	897	1.023
8	1793	0.994
9	3585	0.978
10	7169	0.967
11	14337	0.961
12	28673	0.958

Table 6.3: Convergence of  $H_{min}$  as a function of the number of nodes for  $M = 200$  and  $L = 1$ .

source:	$H_{min}$	Difference (%)
Dowson and Higginson	0.716	25.3
Hamrock and Jacobson	0.770	19.6
Pan and Hamrock	0.723	24.5
Lubrecht et al.	1.010	5.4
Moes	0.978	2.1

Table 6.4: Comparison of  $H_{min}$  predicted by different formulas with calculated result for  $M = 200$  and  $L = 1$ .

Obviously the predictions obtained from equations (6.4) to (6.6) are rather inaccurate. The predicted minimum film thickness is much too small. Fortunately, from a designers point of view, the minimum film thickness is underestimated. With respect to Dowson and Higginson's formula this observation was also made by Lubrecht [L1]. This inaccuracy is a consequence of the fact that these formulas are based mainly on a curve fit of numerical results for relatively high  $L$  values. The predictions from equation (6.7) and (6.10) are much more accurate. However, the close agreement between the predictions of these two formula and the numerical calculations was already shown.

The inaccuracy in the predictions from equations (6.4) to (6.6) is even larger if both  $L$  and  $M$  are small, i.e. close to the "rigid isoviscous" asymptote as is demonstrated by the following example:  $M = 1$  and  $L = 1$ . In terms of the Dowson and Higginson parameters this load situation can be described by  $W = 4.5 \cdot 10^{-6}$ ,  $U = 1.0 \cdot 10^{-11}$ , and  $G = 473$ . The values of  $\bar{\alpha}$  and  $\lambda$  for this situation are:  $\bar{\alpha} = 0.40$  and  $\lambda = 3.70$ . Figure 6.12 displays the calculated dimensionless pressure profile and the associated dimensionless film shape. The value of  $H_{min}$  as a function of  $n$  is presented in table 6.5.

In this case it is important that the solution is calculated using a sufficiently large inlet region, otherwise the computed film thickness will be too small, a numerical artefact that is referred to as "numerically starved lubrication" or "numerical starvation". The pressure profile and film shape presented in figure 6.12 have been calculated using  $X_a = -32$ . Figure 6.12 shows that this load situation can hardly be characterized as EHL. The elastic deformations are very small and the film shape closely approximates its undeformed parabolical shape.

The predictions of the different formulas for this load situation are given in table 6.6. In addition, table 6.6 gives the percentual differences between the predictions and the numerically calculated result. Once more it appears that the values predicted by equations (6.4) to (6.6) are inaccurate and, together with the previously considered load situation, the results demonstrate the limited validity of these film thickness formulas.

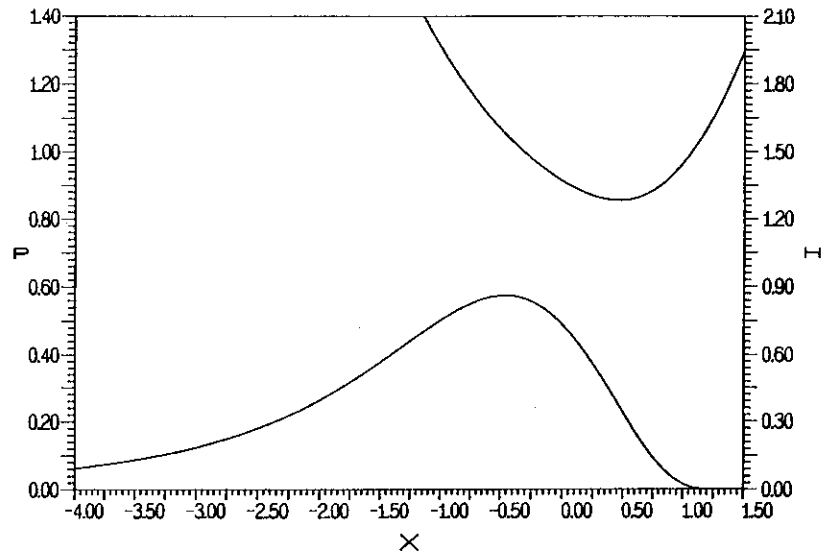


FIGURE 6.12: Calculated pressure profile and film shape for  $M = 1$  and  $L = 1$ .

level	n+1	$H_{min}$
6	449	3.400
7	897	3.333
8	1793	3.300
9	3585	3.282
10	7169	3.275
11	14337	3.270
12	28673	3.267

Table 6.5: Convergence of  $H_{min}$  as a function of the number of nodes for  $M = 1$  and  $L = 1$ .

source:	$H_{min}$	Difference (%)
Dowson and Higginson	1.43	56.2
Hamrock and Jacobson	1.38	57.8
Pan and Hamrock	1.42	56.5
Lubrecht et al.	3.10	5.1
Moes	3.29	0.7

Table 6.6: Comparison of  $H_{min}$  predicted by different formulas with calculated result for  $M = 1$  and  $L = 1$ .

### 6.3 The pressure spike

One of the most intriguing features in the calculated line contact pressure profiles is the occurrence of a very local maximum close to the outlet. Petrusevich [P1] was the first who showed this second local maximum and consequently it is nowadays generally referred to as “the Petrusevich spike” or simply “the pressure spike”.

Ever since, the spike has been the subject of discussion, see for example [K1, H3, L1, B1, H4] and references therein. Two reasons can be given for the interest in the pressure spike. Firstly the academical interest in the phenomenon. The second reason given in literature arises from surface fatigue life considerations. A very small region of high pressures, such as the spike, can act as a stress raiser, i.e. it can lead to stress concentrations in the contacting elastic solids located close to the surface. Consequently one can argue that the spike may have a significant effect on the surface fatigue life of the concentrated contact, e.g. see [H5], and as such requires investigation, the more so since measurements performed with micro transducers [S1, B1] have demonstrated that the spike is not a numerical artefact but that a second local maximum really occurs. However, as will be shown in chapter 7, the possible effect of the spike on the fatigue life of the contact will be small, if not negligible, compared to the effect of surface features, such as indentations and surface roughness, on life. Consequently, from the viewpoint of fatigue life analysis, the study of the effects of surface features on the pressure profile is much more important than the investigation of the pressure spike. Nevertheless, the “academical” interest in the spike remains. An additional reason for studying the spike here is that “capturing” the spike requires a large number of nodal points. As such the study provides another testcase for the stability of the algorithm.

Up till now the discussion with respect to the pressure spike has been focussed mainly on its height in relation with the values of the governing parameters and the lubricant behaviour assumed in the calculations. In particular the question if the spike is a singularity for some values of the operating conditions has been addressed. On the one hand Kostreva [K1] and Lubrecht [L1] argue that the spike might be a logarithmic singularity for some specific situations. On the other hand Hamrock et al. [H3] and Pan et al. [P2] implicitly suggest the opposite by presenting formulas giving the spike height as a function of the governing parameters. Furthermore, Bisset and Glander [B2] explicitly claim *the* spike to be finite and smooth provided it is studied at a sufficiently small length scale. A dubious claim, since it is supported by only a few calculational results.

Notwithstanding this dispute of the spike, it is commonly accepted that it is caused by the exponential relation (Barus or Roelands) between viscosity and pressure. Furthermore, it is agreed upon that for the same conditions the calculated spike height is larger if the Barus viscosity pressure relation is used instead of the Roelands equation. Similarly, higher spikes are calculated if the lubricant is assumed to behave incompressible, see [H4], [L2].

---

In this section the occurrence and the height of the pressure spike in the stationary isothermal EHL line contact situation is investigated. The parameter range where a second local maximum in the pressure profile can be expected is mapped for a compressible and an incompressible lubricant for both the Roelands equation as well as the Barus relation. Furthermore, using the Roelands equation the variation of the spike height with the governing parameters is investigated for both a compressible and an incompressible lubricant.

### 6.3.1 Onset of the spike

One of the differences between the pressure profiles obtained for different values of  $M$  and  $L$  is the number of extremes (locations where  $dP/dX = 0$ ), e.g. see figure 6.1. This observation can be used to define a parameter characterizing the onset of the pressure spike. This parameter,  $L_o$ , is defined as that value of  $L$ , for a certain value of the load parameter  $M$ , for which the pressure distribution has two extremes and below which only one extreme is found (the maximum located near the center of the contact). For  $L > L_o$  three extremes occur; two local maxima, i.e. one in the center of the contact and another, the spike, near the outlet, with a local minimum in between. For example figure 6.13 gives details of the pressure distribution in the outlet region for a specific value of  $M$  ( $M \approx 5$ ) and three values of  $L$ :  $L_A < L_o$ ,  $L_B \approx L_o$  and  $L_C > L_o$  showing the evolution of the second local maximum. For  $L$  slightly larger than  $L_o$  this maximum is only a small bump. However, with increasing  $L$  it gradually changes into a real spike, see also figure 6.6.

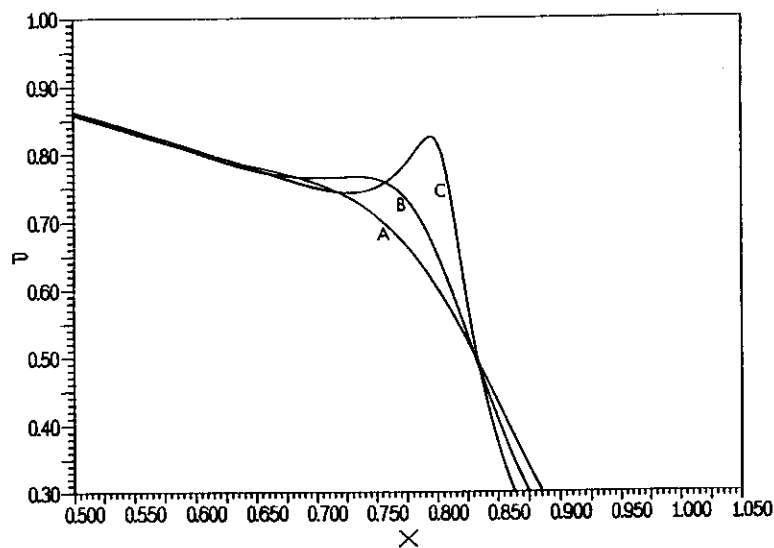


FIGURE 6.13: Onset of pressure spike

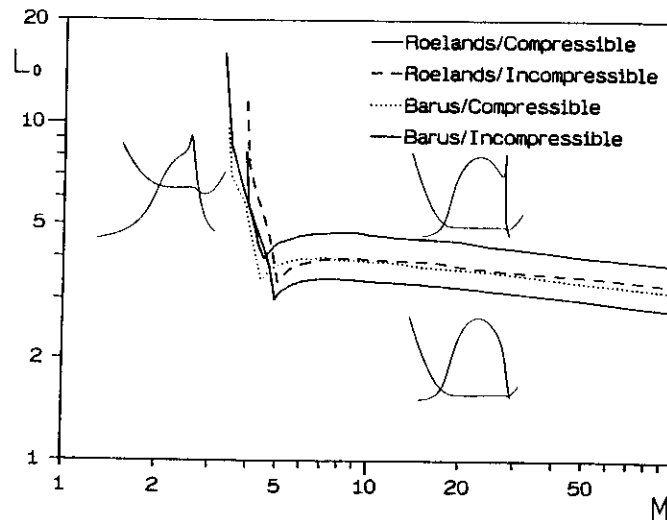


FIGURE 6.14:  $L_o$  as function of  $M$  and  $L$ , and indirectly the zones with different types of pressure profiles

The value of the onset parameter  $L_o$  for a given value of  $M$  depends on the viscosity-pressure relation used in the calculations. Furthermore, it is different for a compressible and an incompressible lubricant. Figure 6.14 gives  $L_o$  as a function of the load parameter  $M$  for different situations of lubricant behaviour, i.e. Barus versus Roelands and incompressible versus compressible. The characteristic pressure distributions prevailing in the regions separated by these curves are also shown. Note that in the region left of the almost vertical part of the curve the pressure profile always has only one extreme. Figure 6.14 shows that the influence of the presumed behaviour of the lubricant is rather small. For large values of  $M$  the curves run nearly parallel and approximately  $L_o \propto M^{-0.09}$ .

From a practical point of view, it may be interesting to plot these curves in the dimensionless minimum film thickness diagram. Consequently, it can not only serve to predict the minimum film thickness but it also provides information on the type of pressure distribution. See for example figure 6.15 where the curve for the Roelands/compressible situation is presented in the Moes plot.

### 6.3.2 Spike height

In the previous section it was shown that there is more to say about the pressure spike than simply concluding that it does occur in the case of an exponential relation between viscosity and pressure and is absent under isoviscous conditions. In particular, it was shown that the spike does not appear suddenly but that it evolves gradually. Having mapped the parameter range where a spike occurs the next step



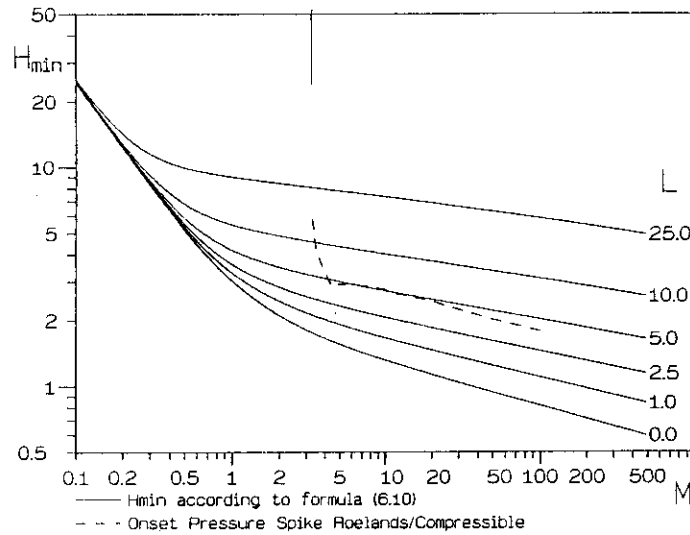


FIGURE 6.15: Onset of pressure spike in the film thickness plot

is to study its height as a function of  $M$  and  $L$  within this range. In particular the variation of the dimensionless spike height with increasing  $M$  is of interest since it has been subject of discussions in the past. One of the complications in such a study is that, given a value of  $L$ , the spike narrows with increasing  $M$ . Consequently, at least locally, increasingly dense grids are required to capture it.

The newly developed algorithm enables the solution of the line contact problem using large nodal densities. As a result in many situations, for example the load condition considered in section 4.4, convergence of the spike height to a finite value can be shown.

The variation of the spike height with the governing parameters observed for two situations is considered, i.e. using the Roelands equation and assuming both an incompressible as well as a compressible lubricant. Figure 6.16 displays the variation of the spike height in the case of a compressible lubricant. Only situations where convergence of the spike height with increasing number of nodes could be shown within 14 levels ( $n + 1 = 114.689$ ) are presented. From the viewpoint of testing out the algorithm this number of nodes was considered sufficiently large. To obtain spike height results beyond the parameter range presented in figure 6.16 local grid refinement techniques are recommended. These techniques are straightforward to apply in a multilevel solver and much more efficient when the only purpose of study is a local feature such as the spike.

Figure 6.16 shows that, within the parameter range presented, the dimensionless spike height decreases with increasing load. This is illustrated in figure 6.17, figure 6.18, and table 6.7. First figure 6.17 presents the pressure profiles for three different values of  $M$  and  $L = 12$ .

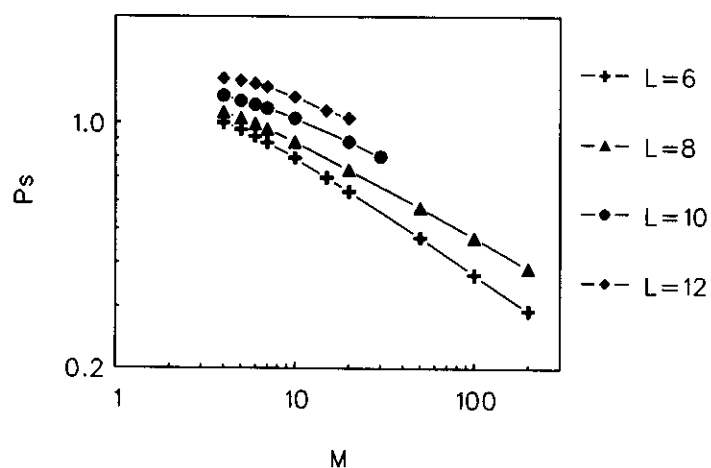


FIGURE 6.16: Variation of the dimensionless spike height with varying  $M$  and  $L$  when Roelands equation is used and the lubricant is assumed to be compressible.

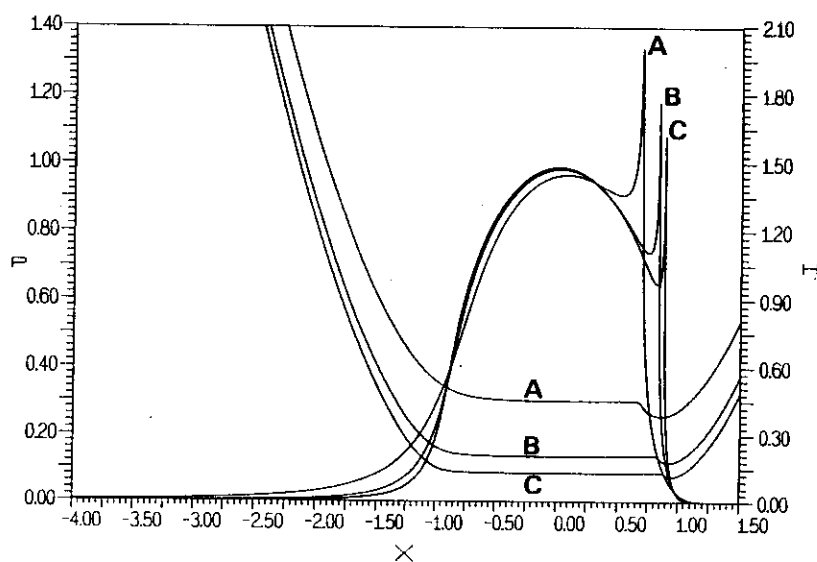


FIGURE 6.17: Pressure profile for  $M=5$  (A), 10 (B), 15 (C), and  $L=12$ .

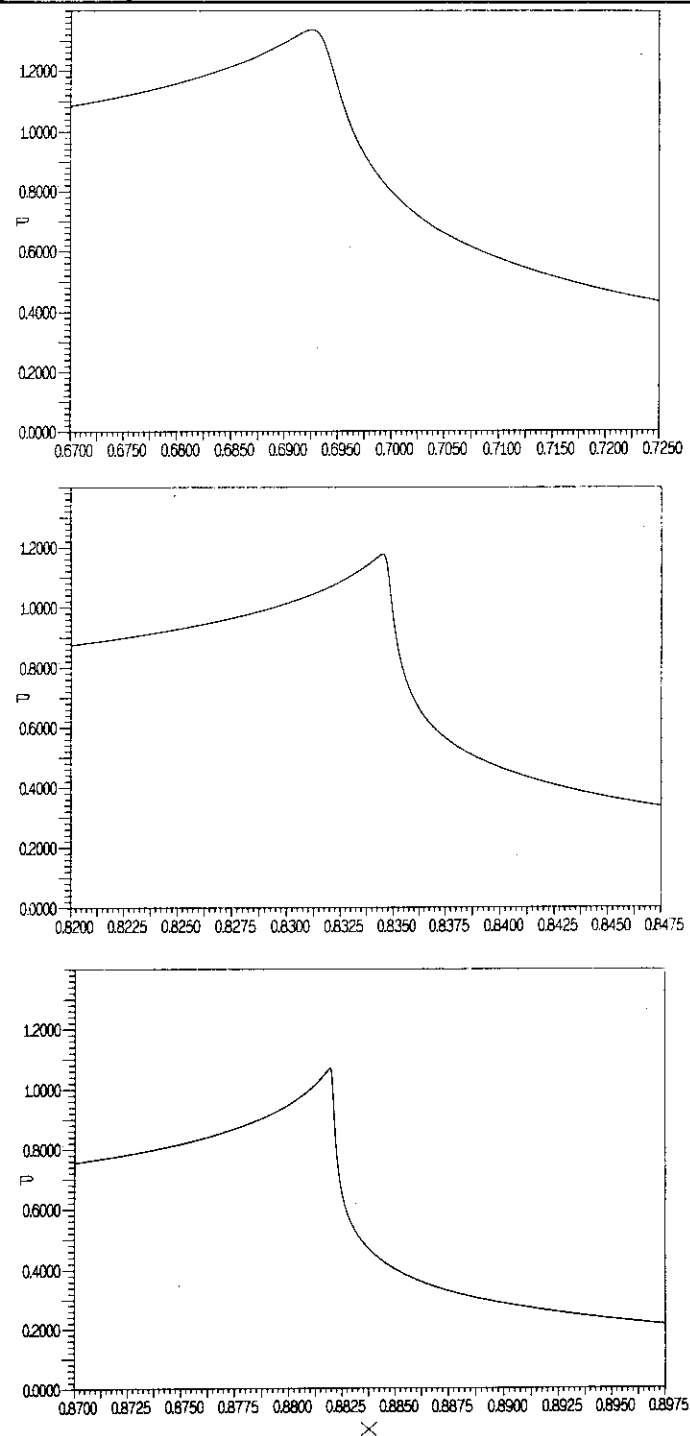


FIGURE 6.18: Enlargements of the pressure profiles presented in figure 6.15 in the spike region .

level	n+1	$M = 5$	$M = 10$	$M = 15$
6	449	0.936	0.915	0.793
7	897	0.964	0.976	0.817
8	1793	1.191	1.024	0.898
9	3585	1.248	1.077	0.938
10	7169	1.288	1.127	0.979
11	14337	1.312	1.150	1.024
12	28673	1.325	1.164	1.047
13	57345	1.332	1.173	1.062
14	114689	1.336	1.177	1.070

Table 6.7: Convergence of the dimensionless spike height as a function of the number of nodes  $n$  for three values of  $M$  and  $L = 12$

Secondly figure 6.18 gives enlargements of the three pressure profiles in the spike region and finally table 6.7 gives the (dimensionless) spike height as a function of the number of nodes.

From figure 6.16 one might be tempted to draw the general conclusion that the spike height decreases with increasing load and, as for example Pan and Hamrock [P2] derive a formula for general use expressing this tendency. However, such a conclusion is doubtful as is shown in figure 6.19 where the calculated dimensionless spike height for an incompressible lubricant is presented.

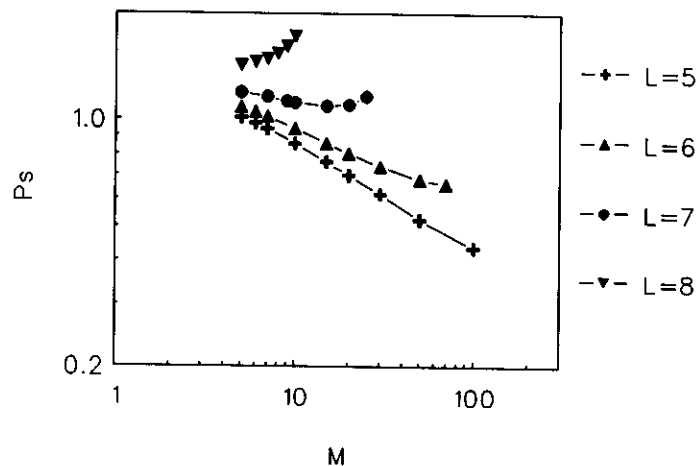


FIGURE 6.19: Variation of the dimensionless spike height with  $M$  and  $L$  when Roelands equation is used and the lubricant is assumed to be incompressible.

In that case the non-dimensional spike height,  $P_s$ , also decreases with  $M$  but only in a limited range of values of  $M$  and  $L$ . Beyond this range  $P_s$  increases rather steeply with  $M$  for a fixed value of  $L$ . The reader is reminded that the Dowson and Higginson relation limits the compressibility of the lubricant to some 33 percent. Hence, for larger values of  $L$  than presented in figure 6.16 a similar increase of the dimensionless spike height with increasing  $M$  might be found.

### 6.3.3 Comparison of spike height results

In figure 6.20 a comparison is made between some present results and spike heights for similar conditions presented by Hamrock et al. [H3], and Pan et al. [P2]. In these papers some situations with fixed  $U$  and  $G$  and varying  $W$  are presented. In terms of the Moes parameters this means  $L$  is fixed whereas  $M$  is varied. The presented results apply to  $U = 1.0 \cdot 10^{-11}$  and  $G = 5007.2$ , hence,  $L = 10.59$  and were obtained using Roelands' viscosity pressure equation and assuming a compressible lubricant. Since  $L$  is fixed a curve similar to the ones presented in figure 6.16 can be expected.

For low loads the results of the different methods match very well whereas the differences increase with increasing  $M$ . Part of the difference may be accounted for by different values of  $\alpha$  and  $z$  used in the calculations with Roelands' equation. However, the rather large differences between the results presented here and in [P2], from the results presented in [H3] are probably caused by numerical inaccuracy of the spike height data presented in [H3]. As mentioned before, with increasing  $M$  the spike narrows and the local nodal density required to describe it accurately increases. Using only 32 nodes in the region  $0.9 < X < 1.0$  to describe the spike, see [H3], has resulted in an underestimation of its height, especially for large  $M$ .

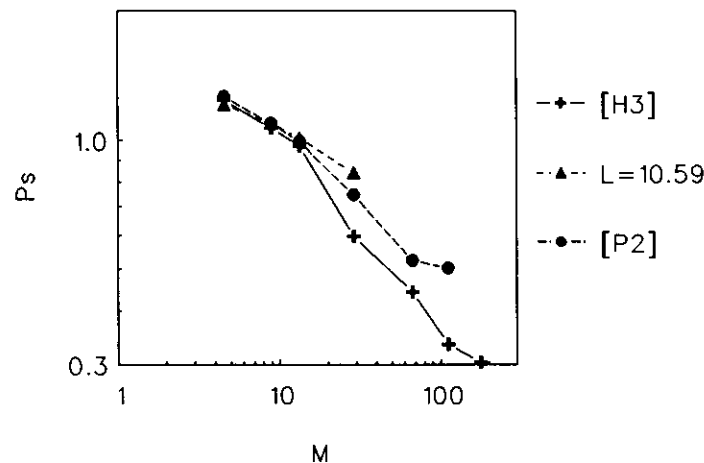


FIGURE 6.20: Comparison of spike height results.

## 6.4 Conclusion

The newly developed algorithm enables accurate solution of the smooth surface EHL line contact problem under all conditions. A parametric study has resulted in a more accurate formula for the prediction of the minimum film thickness. Considering the same loading conditions the results compare favourably with results presented by other authors. With respect to the pressure spike it is demonstrated that it does not occur for all possible combinations of parameter values. The variation of the spike height with varying values of the governing parameters was studied. The results amply demonstrate that the utmost care should be taken in using spike height formulas presented over the years. They should not be used outside the range of parameter values they were based upon. For many values of the governing parameters it was shown that the spike is not a singularity but that its height converges to a finite value with decreasing mesh size. However, the presented results do not justify generalisation of this conclusion to the entire parameter domain.

Finally, the demonstrated stability of the algorithm for highly loaded conditions gives confidence that it can be used successfully in the study of the effect of surface features on the pressure distribution and the film shape. The results of the application of the algorithm to these more complex contact situations will be presented in the following two chapters.

---

## References

- [B1] **Baumann, H., Von Frey, D., and Haller, R.**, 1988, "Druck und temperaturverteilungen in EHD-kontakten," *Tribologie und schmierungstechnik*, 35 Jahrgang, 2, 84-96.
- [B2] **Bisset, E.J., and Glander, D.W.**, 1988, "A highly accurate approach that resolves the pressure spike of elastohydrodynamic lubrication," *ASME JOT*, 110, 241-246.
- [D1] **Dowson, D., and Higginson, G.R.**, 1966, "Elasto-hydrodynamic lubrication, The fundamentals of roller and gear lubrication," Pergamon Press, Oxford, Great Britain.
- [E1] **Ertel, A.M.**, 1939, "Hydrodynamic lubrication based on new principles," *Akad. Nauk SSSR Prikladnaya Matematika i Mekhanika*, 3,2, 41-52
- [G1] **Gümbel, L.**, 1916, "Über geschmierte Arbeitsräder," *Z. ges. Turbinenwesen*, 13, 357.
- [G2] **Grubin, A.N., and Vinogradova, I.E.**, 1949, "Investigation of the contact of machine components," Kh, F. Ketova (ed.) *Central Scientific Research Institute for Technology and Mechanical Engineering* (Moscow), Book No. 30, (DSIR translation No. 337)
- [H1] **Herrebrugh, K.**, 1968, "Solving the incompressible and isothermal problem in elastohydrodynamic lubrication through an integral equation," *ASME JOT*, 90, 262-270.
- [H2] **Hamrock, B.J., and Jacobson, B.O.**, 1984, "Elasto-hydrodynamic lubrication of line contacts," *ASLE trans*, 27, 275-287.
- [H3] **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.
- [H4] **Hamrock, B.J., Ping Pan, and Rong-Tsong Lee**, 1988, "Pressure spikes in elastohydrodynamically lubricated conjunctions," *ASME JOT*, 110, 279-284.
- [H5] **Houpert, L., Ioannides, E., Kuypers, J.C., and Tripp, J.**, 1987, "The effect of the EHD pressure spike on rolling bearing fatigue" *ASME JOT*, 109, 444-451
- [K1] **Kostreva, M.M.**, 1984, "Pressure stability considerations in elasto hydrodynamic lubrication models," *ASME JOT*, 108, 551-556.
- [L1] **Lubrecht, A.A., Breukink, G.A.C., Moes, H., ten Napel, W.E., and Bosma, R.**, 1987, "Solving Reynolds' equation for EHL line contacts by application of a multigrid method," *Proc. 13th Leeds-Lyon Symp. on Tribology, Leeds*, 175-182.
- [L2] **Lubrecht, A.A., Ten Napel, W.E., and Bosma, R.**, 1986, "Multigrid, An alternative method for calculating film thickness and pressure profiles in elastohydrodynamically lubricated line contacts," *ASME JOT*, 108, 551-556.

- [M1] **Martin, H.M.**, 1916, "Lubrication of gear teeth", *Engineering* (London), **102**, 199.
  - [M2] **Moes, H.**, 1985, "On survey diagrams," internal report, Laboratory of Tribology, University of Twente, Enschede, The Netherlands.
  - [M3] **Moes, H.**, 1989, private communications.
  - [P1] **Petrusevich, A.I.**, 1951, "Fundamental conclusions from the contact-hydrodynamic theory of lubrication", *Izv. Akad. Nauk. SSSR (OTN)*, **2**, 209
  - [P2] **Ping Pan, and Hamrock, B.J.**, 1989, "Simple formulas for performance parameters in elastohydrodynamically lubricated line contacts," *ASME JOT*, **111**, 246-251.
  - [S1] **Safa, M.M.A.**, 1982, "Elastohydrodynamic studies using thin film transducers," PhD thesis, Dep. of Electrical Engineering, Imperial College of Science and Technology, London.
  - [T1] **Terril, R.M.**, 1983, "On Grubin's formula in elastohydrodynamic lubrication theory," *Wear*, **92**, 67-78.
-



## Chapter 7

### Line contact results: Surface features

In the previous chapter the pressure generation and fluid film formation in an EHL line contact for perfectly smooth surfaces was discussed. Hence, the analysis was restricted to the macrogeometry of the contact. However, in reality technical surfaces are not perfectly smooth and the pressure generation and lubricant film shape will also be affected by the microgeometry of the contacting surfaces.

For this reason in the present chapter the effects of four different microgeometrical features on the pressure generation and the fluid film formation in an EHL line contact under steady state conditions are investigated. The results of a similar study using a time-dependent approach will be presented in chapter 8. The motivation for studying the effects of surface features arises from sub-surface fatigue life considerations as is explained in section 7.1.

The features are distinguished with respect to their characteristic length scale. In section 7.2 indentations, bumps and waviness are discussed, i.e. features with a characteristic length scale of the order of 0.1 to 1.0 times the size of the half width of the Hertzian contact region, for different load conditions. In particular, the maximum pressure and minimum film thickness in dependence of the wavelength and amplitude of the feature are investigated. Similar theoretical studies have been carried out before, see for example Chow and Cheng [C1], Cheng [C2], Goglia et al. [G1,G2], Lee and Hamrock [L1] and references therein. However, their results were generally restricted to relatively lightly loaded situations.

Subsequently, in section 7.3 a surface feature with a much smaller characteristic length scale is treated, i.e. surface roughness. Typically the characteristic length scale of roughness is  $O(0.1)$  to  $O(0.01)$  times the Hertzian contact size. In this section pressure profiles and film shapes obtained using an actually measured surface trace in the calculations are presented and discussed.

Throughout this chapter it is assumed that the Reynolds equation remains valid, i.e. in all situations considered the characteristic size of the feature in the direction of the flow (its wavelength) is large compared to its characteristic amplitude, typically

$w/A \leq 2 \cdot 10^{-2}$  if  $w$  is the wavelength and  $A$  the amplitude. If this ratio is not small the vertical velocities in the flow resulting from the features can not be neglected and the Stokes equations should be used.

## 7.1 Fatigue life

It has been stated several times that the study of the effects of surface features such as bumps, indentations, waviness and surface roughness is of specific interest for the prediction of the surface fatigue life of rolling element bearings.

This phenomenon may occur when part of the contacting surfaces is damaged, either by incorrect manufacturing, handling or by the overrolling of debris in the oil. The resulting surface features will disturb the pressure in the lubricant film and consequently the stresses in the sub-surface are affected. These surface imperfections can be viewed as local stress raisers, leading to imperfection-related fatigue failure and a substantial reduction in service life.

So far, the relation between surface features and the fatigue life has mainly been investigated experimentally, for example by means of endurance tests, running bearings in batches under specific operating conditions until they fail. By their nature such tests are time-consuming, costly and destructive. However, until recently, these tests were practically the only alternative. The lack of a proper theory relating the geometry of the surface feature to fatigue life, the time consuming numerical solution techniques, and the high price of powerful computers made the subject hard to access theoretically.

Recent developments in the field of theoretical research have significantly changed this situation. Firstly, the New Life Theory presented by Ioannides and Harris [I1] enables a detailed study of the relation between surface features, the associated stresses and fatigue life. Secondly, the introduction of the multilevel multi-integration techniques, enables a fast calculation of the sub-surface stress integrals, i.e. see Lubrecht et al. [L2], and finally the development of fast and stable algorithms enables the computation of the necessary input for such stress calculations, i.e. the pressure profile in the contact. In particular, Lubrecht [L2] developed multilevel solvers based on distributive relaxation schemes to compute the pressure distribution, assuming a dry contact situation. These pressure profiles are more or less *worst case* results for fatigue life studies, since the lubricant film is neglected.

The main motivation for the study presented in this chapter is to provide input for surface fatigue life calculations taking into account the effect of the lubricant film. However, the subject of fatigue life itself will be addressed only briefly. For a detailed study of this subject, where some of the results presented here served as input, reference is made to Lubrecht et al. [L3].

---

## 7.2 “Large scale” features

First the effect of three “large scale” surface features on the pressure profile and film thickness is studied. The study of surface features requires a modification of the film thickness equation. This modification is discussed in section 7.2.1. The features will be studied for three load conditions. These conditions and some calculational details are explained in section 7.2.2. Subsequently the results for the indentations, bumps and waviness are presented and discussed in sections 7.2.3 through 7.2.5.

### 7.2.1 The film thickness equation

In order to account for surface features the dimensionless film thickness equation, i.e. equation (2.45) is extended to:

$$H(X, T) = H_{00}(T) + \frac{X^2}{2} - \mathcal{R}(X, T) - \frac{1}{\pi} \int_{-\infty}^{\infty} P(Y, T) \ln |X - Y| dY \quad (7.1)$$

In general  $\mathcal{R}$  is a function of both space  $X$  and time  $T$  that describes all features on both surfaces that result in a deviation of the undeformed gap from its usual parabolical shape, i.e. surface roughness, surface waviness, bumps, indentations, etc. However, in this chapter a steady state situation is assumed. Hence,  $\mathcal{R}$  represents a feature located on the stationary surface and is a function of  $X$  only.

### 7.2.2 Calculational details

As in this chapter the input conditions are derived from an actual bearing application, the characterization of the load situations using non-dimensional parameters used in the previous chapters is partly abandoned in the current study. Nevertheless, for convenience the results are still presented in terms of the dimensionless pressure  $P$  and film thickness  $H$ .

All parameters, except for the reduced radius of curvature, are fixed. The invariant parameters and their values are listed in table 7.1. To study the effect of load for a given feature the following approach is followed. The half width of the Hertzian contact region  $b$  is fixed and different values of the reduced radius of curvature are considered, resulting in maximum Hertzian pressures of 1, 2, and 3 GPa. This approach was chosen to single out the effect of the larger pressure and it ensures that a specific feature has the same scale compared to the contact size in all three load situations. The values of the reduced radius of curvature, the resulting maximum Hertzian pressures, the values of the Moes dimensionless parameters  $M$  and  $L$  as well as the dimensionless parameters  $\bar{\alpha}$  and  $\lambda$  for the load situations created in this way are displayed in table 7.2. In addition, table 7.2 gives the calculated dimensionless minimum film thickness under smooth surface conditions in micrometers ( $\mu m$ ).

Parameter	Value	Dimension
$E'$	$2.26 \cdot 10^{11}$	$[Pa]$
$\alpha$	$2.2 \cdot 10^{-8}$	$[Pa^{-1}]$
$z$	0.68	
$\eta_0$	$40 \cdot 10^{-3}$	$[Pa \cdot s]$
$u_s$	1.94	$[ms^{-1}]$
$b$	$5.0 \cdot 10^{-4}$	$[m]$

Table 7.1: Parameter values used in the calculations.

$R [m]$	$p_h [Pa]$	$M$	$L$	$\bar{\alpha}$	$\lambda$	$h_{min} [\mu m]$
$2.83 \cdot 10^{-2}$	$1.0 \cdot 10^9$	35.02	9.32	22.00	$3.0 \cdot 10^{-3}$	0.337
$1.41 \cdot 10^{-2}$	$2.0 \cdot 10^9$	100	11.08	44.20	$3.7 \cdot 10^{-4}$	0.232
$9.42 \cdot 10^{-3}$	$3.0 \cdot 10^9$	181.95	12.26	66.03	$1.1 \cdot 10^{-4}$	0.190

Table 7.2: Different loading conditions.

### 7.2.3 Indentations

First the effect of an indentation on the pressure profile and the film thickness is studied. The reason for this choice is that, over the years, experimental investigations, e.g. see Sayles and Macpherson [S1] and Lorösch [L4], have demonstrated that indentations, regardless whether they stem from handling damage, overrolling of hard or ductile particles (contaminants in the lubricant) cause a significant reduction of the surface fatigue life of the rolling element in bearings. The following dent geometry is assumed, e.g. see figure 7.1:

$$\mathcal{R}(X) = \mathcal{A} 10^{-10 \left( \frac{x - X_d}{\mathcal{W}} \right)^2} \cos \left( 2\pi \frac{X - X_d}{\mathcal{W}} \right) \quad (7.2)$$

where:  $\mathcal{A}$  = dimensionless amplitude ( $\mathcal{A} \leq 0$  for a dent):  $\mathcal{A} = AR/b^2$   
 $\mathcal{W}$  = dimensionless wavelength of the dent:  $\mathcal{W} = w/b$   
 $X_d$  = dimensionless position of the centre of the dent:  $X_d = x_d/b$

This specific shape was preferred over a simple harmonic dent because it has no sharp corners at the edges and it emulates a physical dent more realistically because it contains the bulged-out shoulders resulting from the indentation of the material. Since this chapter is restricted to steady state conditions the location of the dent denoted by  $X_d$ , is fixed.

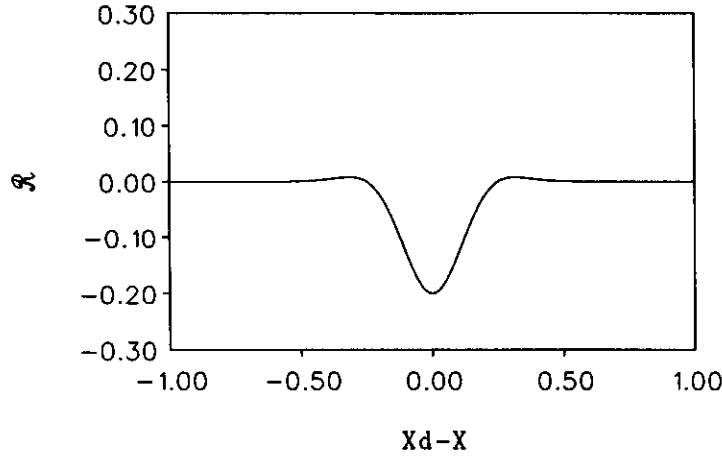


FIGURE 7.1: *Geometry of indentation,  $\mathcal{A} = -0.2$ ,  $\mathcal{W} = 1.0$ .*

As an example figure 7.2 presents the (dimensionless) pressure profile and (dimensionless) film thickness for the 2 GPa loading condition in the case of an indentation with an amplitude of  $2 \mu\text{m}$  and a wavelength of  $0.5 \text{ mm}$  ( $\mathcal{W} = 1$  and  $\mathcal{A} = -0.11$ ) located at the center of the contact ( $X_d = 0$ ). The pressure profile differs from the well known smooth surface pressure profile in the pressure rises at the trailing and leading edge of the indentation with a pressure drop in between, i.e. at the center of the dent. The pressure rises at the leading and trailing edge are equal, i.e. apart from the inlet and outlet region the pressure profile is symmetrical.

The solution has been calculated using 5633 nodal points on a domain extending from  $X_a = -4$  to  $X_b = 1.5$ . The resulting nodal density is not large enough to accurately represent the pressure spike. Consequently the height of the spike is underestimated but the exact height is not our aim here.

Notwithstanding the obvious deviation of the pressure profile from the smooth surface pressure, the film shape is only slightly affected by the presence of the indentation. As the film shape in absence of any surface feature it is nearly uniform in the contact region. This is no surprise since under the given conditions of a relatively high maximum Hertzian pressure ( $\bar{\alpha}=44$ ) the Reynolds equation in the contact region reduces to:

$$\frac{d(\bar{\rho}H)}{dX} \approx 0 \quad (7.3)$$

Dowson and Higginson's density pressure equation limits the compressibility of the lubricant to about 33 % hence, unless the amplitude of the indentation is relatively large (large pressure drop at the center), it will hardly show up in the film shape.

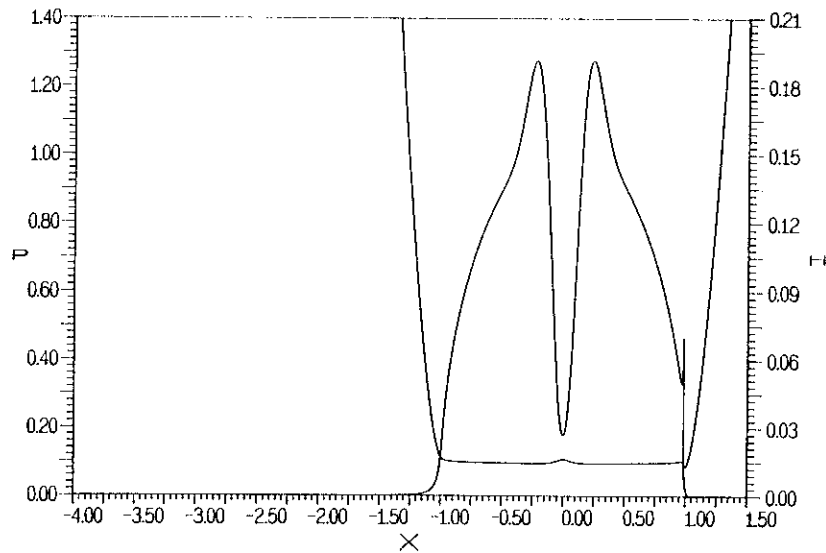


FIGURE 7.2: Pressure profile and film shape for the 2 GPa loading condition in the case of a dent with a wavelength of 0.5 mm and an amplitude of  $2\text{ }\mu\text{m}$  at  $X_d = 0$ .

### 7.2.3.1 Varying wavelength and amplitude

For a given load condition, the pressure rise at the trailing and leading edge of the indentation depends on its wavelength and amplitude. Figure 7.3 presents the ratio of the maximum pressure in the case of an indented surface to the maximum pressure under smooth surface conditions as a function of the wavelength for four different amplitudes, i.e. 1, 2, 3 and  $4\text{ }\mu\text{m}$  ( $A = -0.056, -0.11, -0.17$ , and  $-0.23$ ), for the 2 GPa load condition. All results apply to an indentation located at the center of the contact. From figure 7.3 it can be observed that the relative maximum pressure is inversely proportional to the wavelength. In addition, it is proportional to the amplitude.

A similar graph for the minimum film thickness relative to the minimum film thickness under smooth surface conditions is presented in figure 7.4. This figure shows that the indentation causes a decrease of the minimum film thickness. However, even though the amplitude of the dent is up to 20 times larger than the smooth surface minimum film thickness, in all situations the effect of the indentation on the minimum gap width is small ( $\leq 4\%$ ).

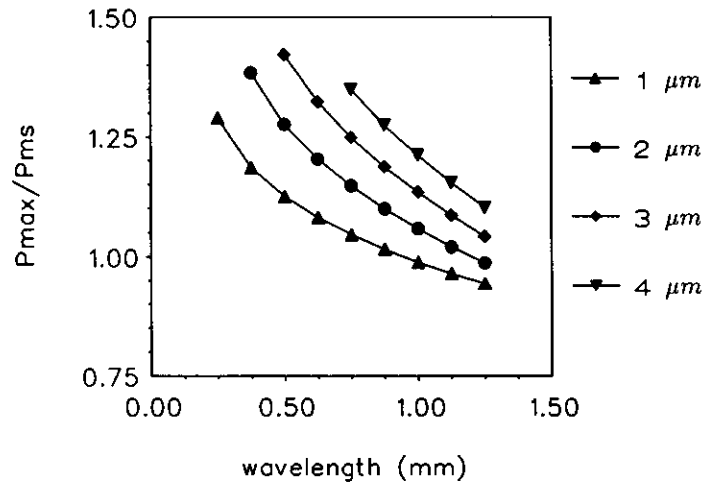


FIGURE 7.3: Relative maximum pressure as a function of the wavelength for the 2 GPa load situation and amplitudes of 1, 2, 3, and 4  $\mu m$ .

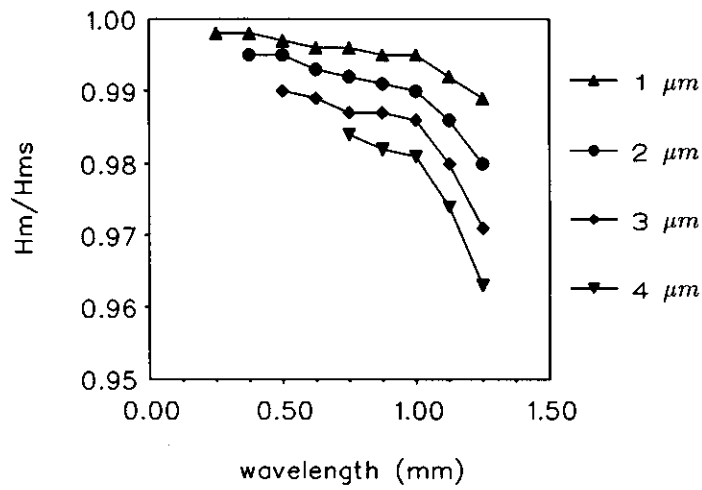


FIGURE 7.4: Relative minimum film thickness as a function of the wavelength for the 2 GPa load situation and amplitudes of 1, 2, 3, and 4  $\mu m$ .

### 7.2.3.2 Cavitation

Note that the results presented in figure 7.3 and 7.4 are limited to rather shallow dents. For a given wavelength, the pressure drop at the center of the dent increases with increasing amplitude. Gradually both sides of the indentation start to act as separate contacts, see for example figure 7.5. This figure presents the dimensionless pressure profile and film shape for the same loading conditions (2 GPa) with an amplitude of  $3\text{ }\mu\text{m}$  and wavelength  $0.5\text{ mm}$  ( $A = -0.17$  and  $W = 1$ ). The pressure at the dent center is *almost* zero. The pressure profile and film shape at both the left and right side of the indentation show all characteristics of the normal smooth surface solution, i.e. a nearly uniform film which narrows at the end where this narrow region is preceded by a pressure spike.

For larger amplitudes, the fluid will actually cavitate at the dent location and both sides of the contact will really become separate contacts, see also Goglia [G1]. This feature limits the validity of the results that can be obtained with the present model since this model is based upon the assumption that the contact is fully flooded throughout the entire pressurized region. If cavitation occurs at the dent location the present model will predict pressure generation downstream of this site from the moment the gap narrows again regardless whether it is fully flooded at that point or not. Hence, to solve situations where cavitation occurs at the center of the dent requires an extension of the *model* allowing positive pressures only if the gap is fully flooded. This can for example be done in the way described by Elrod [E1].

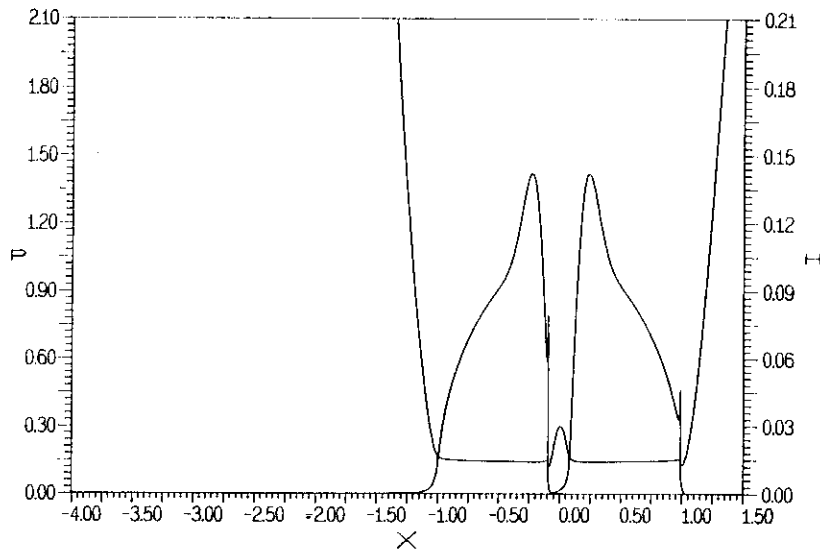


FIGURE 7.5: Pressure profile and film film thickness for the 2 GPa load conditions in the case of an indentation with an amplitude of  $3\text{ }\mu\text{m}$  and a wavelength of  $0.5\text{ mm}$  located at  $X_d = 0$ .



### 7.2.3.3 Varying $p_h$

Having studied the effect of the wavelength and amplitude of the indentation on the maximum pressure and minimum film thickness for a given maximum Hertzian pressure, next the effect of the maximum Hertzian pressure itself is studied. For this purpose the three different load situations described earlier are considered. The results presented in this section apply to an indentation with an amplitude of  $2\ \mu\text{m}$ . This is equivalent to a dimensionless amplitude  $\mathcal{A} = -0.23, -0.11$  and  $-0.075$  for the 1, 2, and 3 GPa load situation respectively.

Figure 7.6 presents the maximum pressure at the shoulder of the indentation relative to the maximum pressure in the case of smooth surfaces as a function of the wavelength of the indentation for the three different load situations. In the 2 and 3 GPa load situations this is also the overall maximum pressure. This is not true for the 1 GPa load condition. In the case of a wavelength of  $0.5\ \text{mm}$  the pressure spike preceding the pressure drop at the indentation exceeds the maximum pressure at the shoulders of the dent. In that case the maximum at the shoulder is not the overall maximum. However, it is expected that the maximum pressure at the shoulder of the dent determines the effect of the dent on the fatigue life. Therefore, and for reasons of comparison with the other load situations, this value is displayed.

The relative maximum pressure increase is largest for the load situation of 1 GPa and smallest for the 3 GPa situation. This is due to the fact that in the 1 GPa situation the elastic deformation is smallest, in other words, the amplitude of the

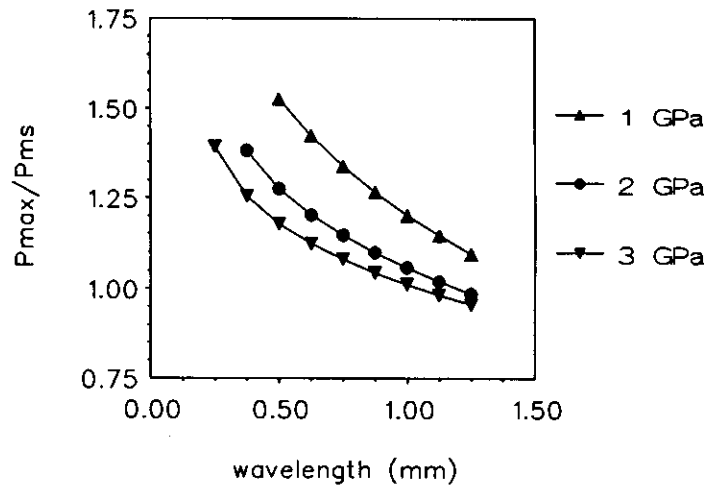


FIGURE 7.6: Relative maximum pressure as function of the wavelength of the indentation for three different values of  $p_h$ . The amplitude of the dent is fixed at  $2\ \mu\text{m}$ .

indentation relative to the maximum elastic deflection is largest resulting in a larger relative pressure increase.

This only applies to the relative increase. Reworking the data to absolute values, see figure 7.7, reveals that the absolute increase of the maximum pressure is roughly independent of  $p_h$ . This implies that the local pressure generated by the indentation is independent of the global pressure (and thus independent of the load). For wavelengths of the order of  $2b$  this is no longer true. In those situations the dent starts to modify the *macro* pressure distribution and is no longer local. However, with increasing wavelength ( $\mathcal{W} \gg 2$ ) these effects disappear again and gradually the smooth surface solution returns.

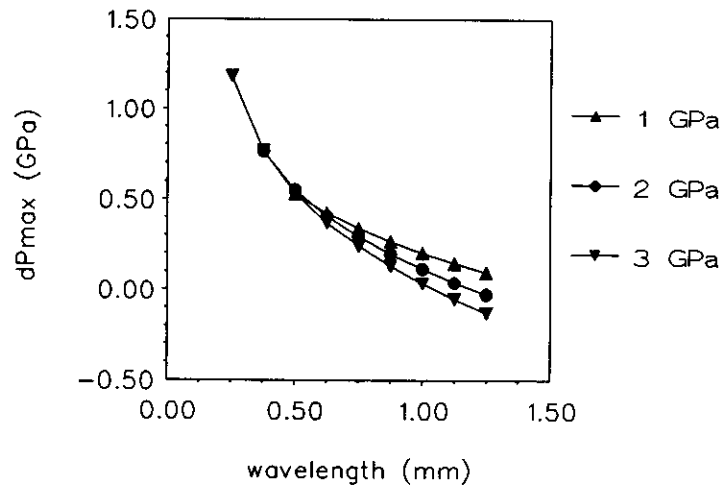


FIGURE 7.7: Absolute maximum pressure rise as a function of the wavelength of the indentation for three values of  $p_h$ . The amplitude of the dent is fixed at  $2 \mu\text{m}$ .

The values of the minimum film thickness relative to the minimum film thickness in the case of smooth surfaces for the situations presented in figure 7.6 are displayed in figure 7.8. This figure shows that the indentation causes a reduction of the minimum film thickness that increases with the wavelength. Furthermore, for a given wavelength the reduction of the relative minimum film thickness increases with decreasing maximum Hertzian pressure. This is explained by the fact that with decreasing  $p_h$  the amplitude of the indentation relative to the elastic deformation increases.

In all situations presented for the 2 and 3 GPa load conditions the minimum film thickness is found at the end of the contact region, i.e. at  $X \approx 1$  and the reduction of the minimum film thickness is quite small. This does not apply to the results of the 1 GPa load condition. For the larger wavelengths the minimum film thickness also occurs at the end of the contact region and the deviation from the smooth surface minimum film thickness is relatively small. However, for the two smallest values of the wavelength of the indentation the minimum film thickness occurs on the inlet edge of the indentation and its value deviates much more from the value calculated in the case of smooth surfaces. In these situations the pressure at the dent location is close to zero and the contact in fact behaves as two separate contacts.

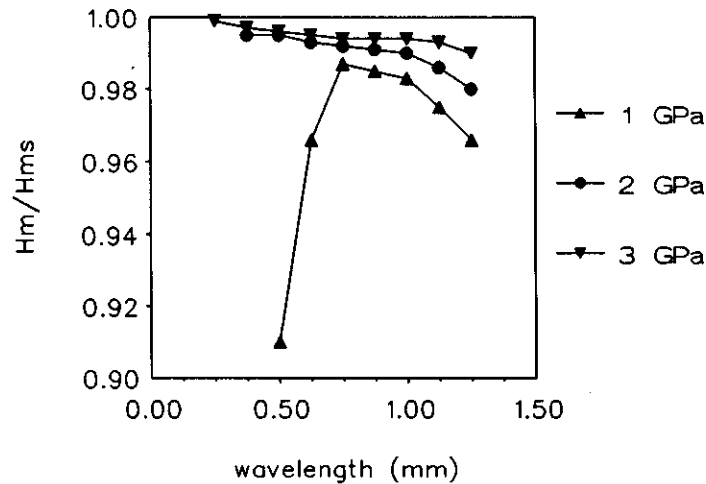


FIGURE 7.8: Relative minimum film thickness values associated with the results presented in figure 7.6.

#### 7.2.3.4 Varying location

So far the location of the indentation was fixed. However, the pressure rise at the edges of the dent and the change of the minimum film thickness resulting from the presence of the indentation will depend on its position in the contact region. The effect of the position is studied here. It is emphasized that still a steady state situation is assumed, i.e. the velocity of the indented surface is zero and only the opposing smooth surface moves (simple sliding).

As an example figure 7.9 shows the calculated dimensionless pressure profile and film shape for the 3 GPa load situation in the case of an indentation with a wavelength of 0.5 mm and an amplitude of 2  $\mu\text{m}$  ( $A = -0.075$  and  $W = 1.0$ ) located at  $X_d = -0.5$ .

Obviously the pressure profile is no longer symmetric with respect to the dent location. The largest pressure occurs on the edge of the indentation that is nearest to the center of the contact. The maximum pressure still exceeds the maximum pressure for smooth surfaces but it differs from the maximum pressure found for the same situation with the dent located in the center of the contact.

The relative maximum pressure at the dent shoulder as a function of the location of the indentation in the contact for the 1, 2, and 3 GPa load conditions listed in table 7.2 is displayed in figure 7.10. The calculations have been carried out assuming a dent with a fixed amplitude of 2  $\mu\text{m}$  and a wavelength of 0.5 mm.

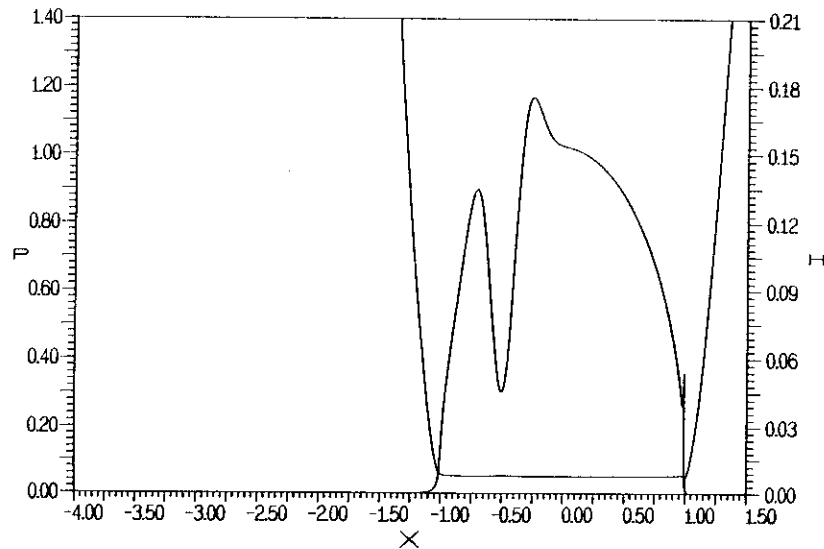


FIGURE 7.9: Pressure profile and film shape for the 3 GPa load situation with indented surface. The wavelength of the dent is 0.5 mm whereas its amplitude is 2  $\mu\text{m}$ . The dent is located at  $X_d = -0.5$ .

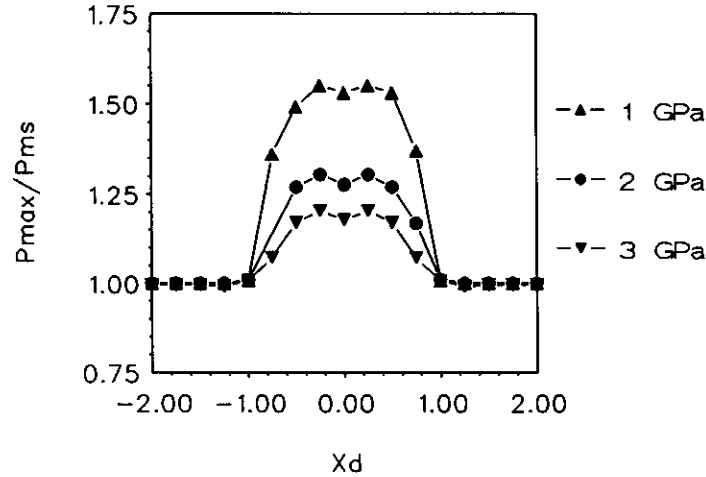


FIGURE 7.10: Relative maximum pressure as a function of the position of the dent in the case of an indentation with an amplitude of  $2 \mu\text{m}$  and a wavelength of  $0.5 \text{ mm}$  for three values of  $p_h$ .

In particular for the 1 GPa load situation the value displayed is not always the overall maximum pressure, as was already explained in section 7.2.3.3.

If the dent is located far away from the center of the contact region the pressure profile and film thickness are almost identical to the smooth surface pressure profile and film shape. The largest maximum pressures are found if the dent is located just beside the center of the contact. The maximum pressure always occurs on the edge of the indentation that is closest to the center of the contact. Hence, if  $X_d < 0$ , the maximum pressure is found on the edge of the dent at the outlet side and if  $X_d > 0$  it occurs on the edge at the inlet side. Note that in fact the maximum pressure is a function of  $|X_d|$ , i.e. it is symmetric with respect to the position of the dent in the contact region.

The associated results for the relative minimum film thickness are presented in figure 7.11. This figure shows that the effect of the indentation on the minimum film thickness depends significantly on its position in the contact region. With respect to the relatively large reduction of the minimum film thickness for the 1 GPa results it should be noted that for all  $-1 \leq X_d \leq 1$  the overall minimum film thickness occurred at the inlet edge of the indentation, i.e. the large pressure drop at the dent location divides the entire contact in two almost separate contacts, whereas in all other situations it was found at the end of the contact region.

Nevertheless, figure 7.11 clearly shows that the largest effect on the minimum film thickness occurs if the dent is located in the inlet region of the contact, i.e. if  $X_d \approx -1$ . This is no surprise since it is well known from the numerous calculations

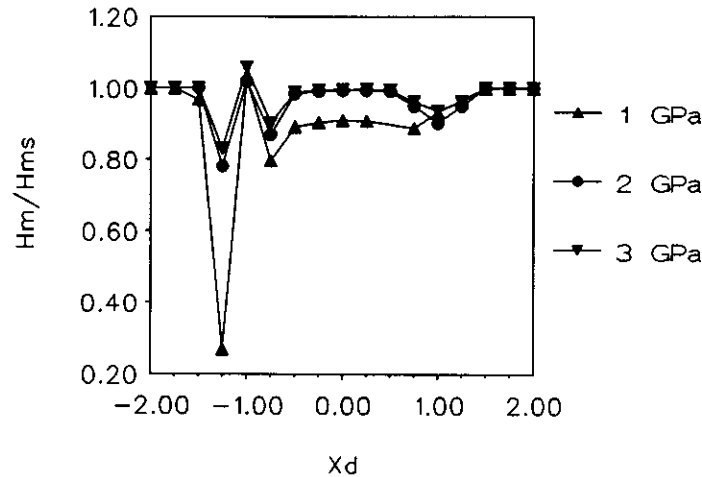


FIGURE 7.11: *Relative minimum film thickness values associated with the results presented in figure 7.10*

for smooth surface situations that the pressure generation in the inlet region determines the minimum film thickness. Any microgeometrical feature located in this region changes the pressure generation and therefore it may significantly affect the minimum gap width.

#### 7.2.3.5 Implications of a dent for the fatigue life

From the results presented in the previous sections it can be concluded that the indentation will result in stress concentrations below both edges of the indentation. These stresses will exceed the maximum sub-surface stresses under smooth surface conditions and consequently the indentation may result in a reduction of the fatigue life of the contact. With respect to the depth below the surface it can be expected that, with decreasing wavelength, the stress concentration will be located closer to the surface.

To include the effect of the location of the dent one could use a semi-transient approach, i.e. calculate the maximum stresses over all positions of the indentation in the contact region, e.g. see [L3]. Such a study gives similar results. The maximum value of this quantity will be found below the edges of the indentation. Moreover, its value will be equal for both sides.

Hence, it is concluded that, as far as the theoretical predictions are concerned, if the contact fails due to surface fatigue, spalling will initiate with equal possibility below either one of the edges of the indentation. As this is not in accordance with the experimentally observed failure location, this subject will be addressed once more in chapter 8.

### 7.2.4 Bumps

The results presented in this section apply to the reverse of the feature discussed in section 7.2.3, i.e. a bump instead of an indentation. The bump geometry used is the same as the dent geometry described by equation (7.2) with an amplitude of the opposite sign. This specific shape was chosen because it has no sharp corners or discontinuities at the edges.

As an example figure 7.12 presents the calculated (dimensionless) pressure profile and (dimensionless) film shape for the 2 GPa load condition in the case of a bump with an amplitude of  $2 \mu\text{m}$  and a wavelength of  $0.5 \text{ mm}$  ( $\mathcal{A} = 0.11$  and  $\mathcal{W} = 1.0$ ) located in the center of the contact ( $X_d=0$ ).

Comparing the pressure profile displayed in figure 7.12 with the smooth surface pressure profile shows that the bump causes a significant increase of the maximum pressure in the contact region. With respect to the film shape it is interesting to see that, although the bump amplitude is 9 times larger than the smooth surface minimum film thickness (see table 7.2), the film shape hardly deviates from the smooth surface film shape. This is, similar to the situation for an indentation discussed in the previous section, explained by the relatively large value of  $\bar{\alpha}$  and the accompanying degeneration of the Reynolds equation in the center region to  $d(\bar{\rho}H)/dX \approx 0$ .

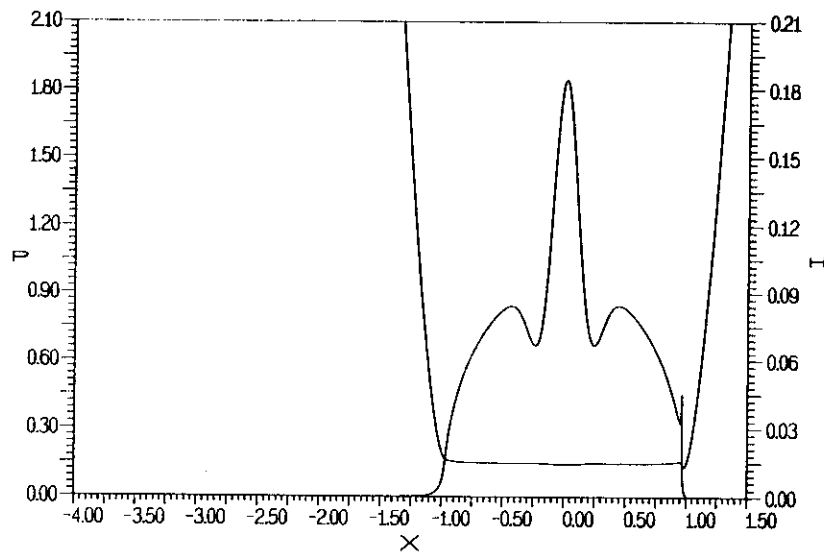


FIGURE 7.12: Pressure profile and film shape for the 2 GPa load condition in the case of a bump with a wavelength of  $0.5 \text{ mm}$  and an amplitude of  $2 \mu\text{m}$  at  $X_d = 0$ .

### 7.2.4.1 Varying wavelength and amplitude

First the effect of the wavelength of the bump on the minimum film thickness is studied in the 2 GPa load situation for different values of the amplitude. Figure 7.13 presents the maximum pressure relative to the smooth surface maximum pressure as a function of the wavelength for amplitude values of 1, 2, 3, and 4  $\mu\text{m}$ . ( $\mathcal{A} = 0.056, 0.11, 0.17, \text{ and } 0.23$ ).

Figure 7.13 shows that the relative pressure rise at the bump is inversely proportional to the wavelength. In addition, it is proportional to the amplitude of the bump as follows from figure 7.13 where the results are presented as a function of the amplitude. Hence the increase of the maximum pressure is approximately proportional to the ratio  $\mathcal{A}/W$ . With respect to the large increases of the maximum pressure found for small wavelengths and large amplitudes it should be noted that these results are not realistic. In reality the pressure at the bump can not rise beyond the plasticity limit of the material ( $\approx 3.3 \text{ GPa}$  in the case of steel surfaces).

The minimum film thickness relative to the minimum film thickness under smooth surface conditions associated with the results presented in figure 7.13 is displayed in figure 7.15. This figure shows that the presence of a bump results in an increase of the minimum film thickness. However, even if the bump amplitude is large compared to the smooth surface minimum film thickness, the increase is still small.

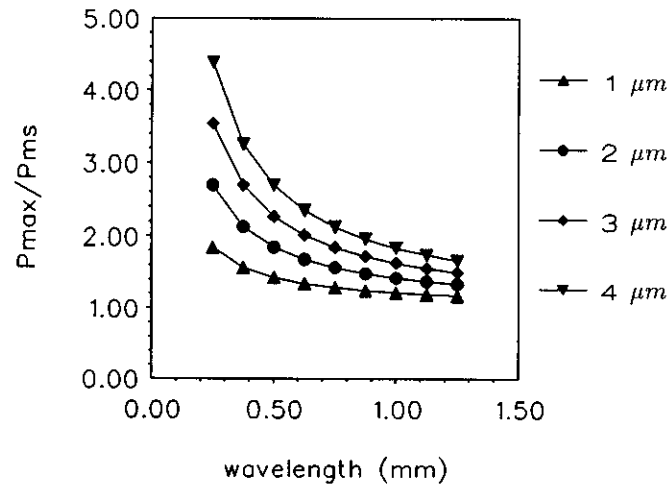


FIGURE 7.13: Relative maximum pressure as a function of the wavelength for the 2 GPa load situation and amplitudes of 1, 2, 3, and 4  $\mu\text{m}$ .



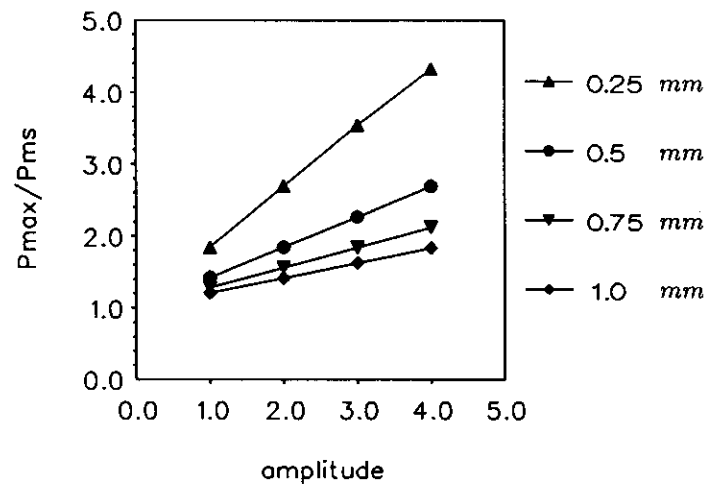


FIGURE 7.14: Relative maximum pressure as a function of the amplitude for different wavelengths.

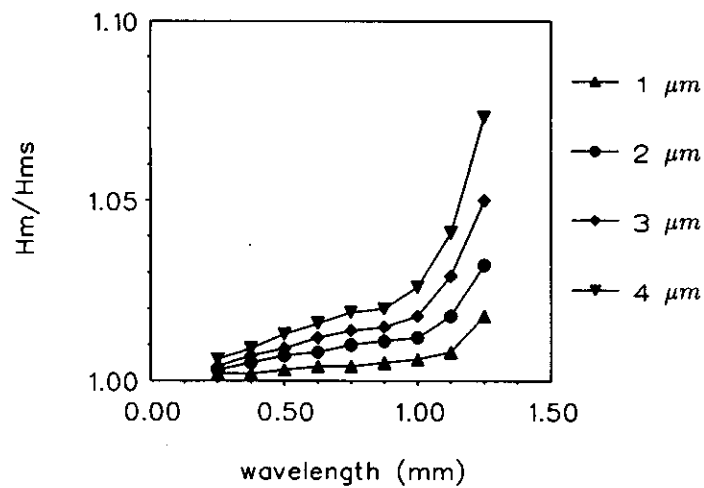


FIGURE 7.15: Relative minimum film thickness as a function of the wavelength for the 2 GPa load situation and amplitudes of 1, 2, 3, and 4 μm.

#### 7.2.4.2 Varying $p_h$

In a similar way as done for an indentation (section 7.2.3.3) for a specific bump the dependence of the maximum pressure in the contact and the minimum film thickness on the maximum Hertzian pressure has been investigated. Figure 7.16 presents the relative maximum pressure at the bump center as a function of the wavelength for the aforementioned load situations of 1, 2, and 3 GPa. In all situations presented here this is also the overall maximum pressure. The relative maximum pressure is largest for the 1 GPa load situation, because in this situation the amplitude of the bump is largest compared with the maximum deformation. Reworking the data presented in figure 7.16 to the absolute increase of the maximum pressure in GPa gives figure 7.17. Obviously the absolute pressure rise caused by the bump is approximately independent of the maximum Hertzian pressure. As was found for an indentation this implies that the local pressure generated by the bump is independent of the global pressure (and thus independent of the load). Again for wavelengths of the order of  $2b$  deviations occur because the bump starts to modify the *macro* pressure distribution and is no longer local. However, with increasing wavelength ( $\mathcal{W} \gg 2$ ) these effects disappear again and gradually the smooth surface solution returns.

Figure 7.18 presents the relative minimum film thickness values associated with the results presented in figure 7.16 and 7.17. In all three load situations the bump results in a relatively small increase of the minimum film thickness.

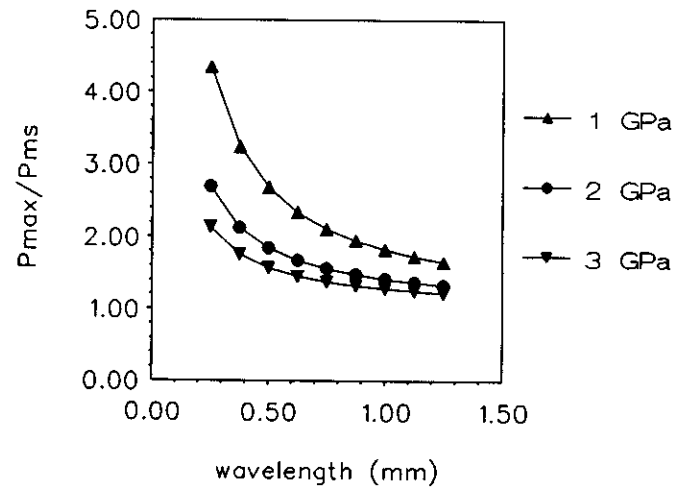


FIGURE 7.16: Relative maximum pressure as a function of the wavelength for three load situations. The bump amplitude is fixed at  $2 \mu\text{m}$ .

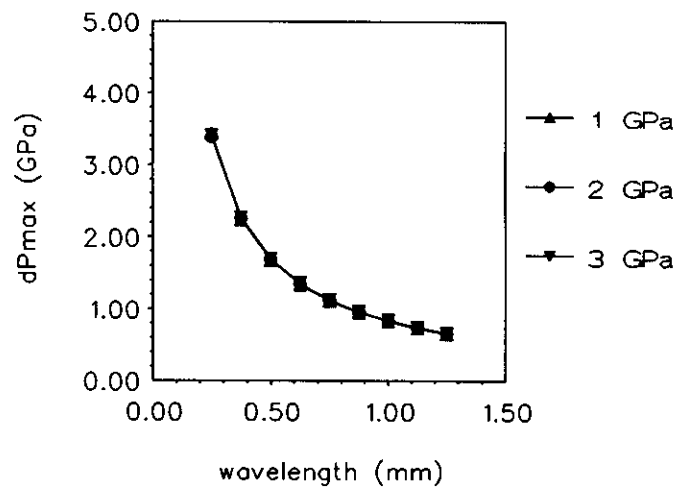


FIGURE 7.17: Absolute maximum pressure rise as a function of the wavelength for three load situations. The bump amplitude is fixed at  $2\text{ }\mu\text{m}$ .

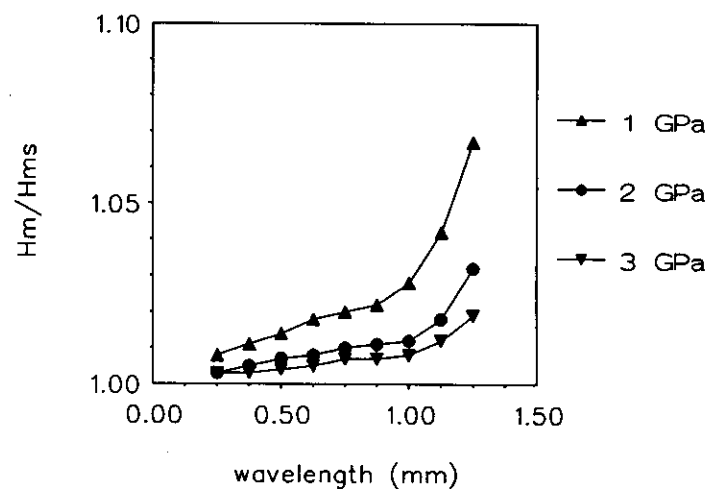


FIGURE 7.18: Relative minimum film thickness as a function of the wavelength for three load situations. The bump amplitude is fixed at  $2\text{ }\mu\text{m}$ .

For a given wavelength the effect of the bump on the relative minimum film thickness increases with decreasing maximum Hertzian pressure. As for an indentation, see section 7.2.3.3, this is explained by the increasing relative size of the indentation compared to the elastic deformation. However, in all situations considered the relative increase of the minimum film thickness is quite small ( $\leq 8\%$ ). Hence, notwithstanding the large size of the amplitude of the bump compared to the smooth surface minimum film thickness ( $A \approx 10 h_{min}$ ) the minimum film thickness in the contact is not significantly affected by the presence of the bump.

#### 7.2.4.3 Micro EHL

For a given wavelength with increasing amplitude the bump gradually starts to act as a separate EHL contact. This behaviour is generally referred to as micro EHL, e.g. see Cheng [C2], Goglia et al. [G1], and Lee and Hamrock [L1]. Figure 7.19 displays the dimensionless pressure profile and film shape for the 1 GPa load condition in the case of a bump with an amplitude of  $4\text{ }\mu\text{m}$  and a wavelength of  $0.5\text{ mm}$  ( $A = 0.46$  and  $\mathcal{W} = 1.0$ ). Because of the large amplitude of the bump the pressure profile and film shape in the three parts of the contact, i.e. the part on the inlet side of the bump, the part on the outlet side of the bump and the bump itself show all characteristics of the normal smooth surface pressure profile and film shape.

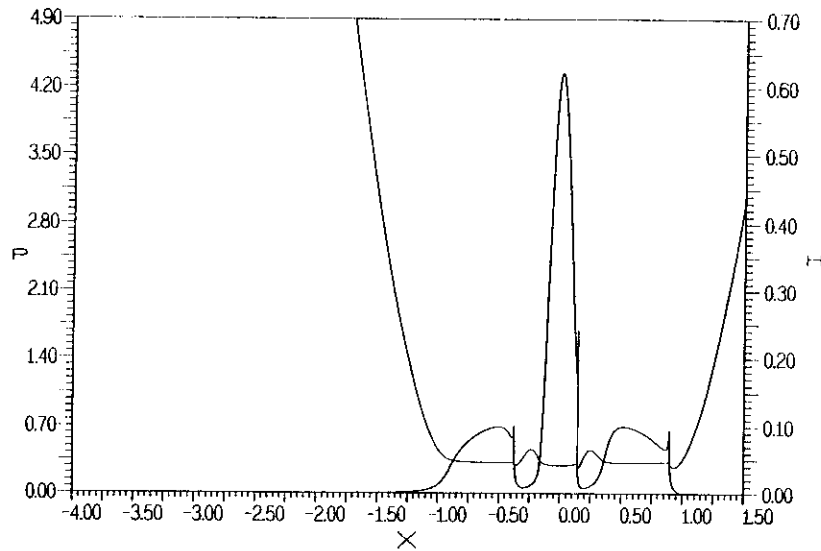


FIGURE 7.19: Pressure profile and film shape for the 1 GPa load situation in the case of a bump with an amplitude of  $4\text{ }\mu\text{m}$  and a wavelength of  $0.5\text{ mm}$  located at  $X_d = 0$ .

#### 7.2.4.4 Varying location

As in the case of an indentation, see section 7.2.3.4, the effect of a bump on the maximum pressure and minimum film thickness depends on the location of the bump in the contact region. As an example figure 7.20 displays the dimensionless pressure profile and film shape for the 3 GPa load condition in the case of a bump with an amplitude of  $2\ \mu\text{m}$  and a wavelength of  $0.5\ \text{mm}$  located at  $X_d = -0.5$  ( $\mathcal{A} = 0.075$  and  $\mathcal{W} = 1.0$ ).

Figure 7.21 presents the maximum pressure at the bump relative to the maximum Hertzian pressure as a function of the location in the contact for the three load situations considered here. The results apply to a bump with an amplitude of  $2\ \mu\text{m}$  and a wavelength of  $0.5\ \text{mm}$ .

With respect to the results presented in figure 7.21 it should be noted that the value displayed is not always the overall maximum pressure in the contact. In particular, in the case of the 1 GPa load situation with the bump located at  $X_d = 1$  the pressure spike exceeds the maximum pressure at the bump center.

The largest values of the relative maximum pressure are found if the bump is located in the center of the contact. Note that, similar to what was found for an indentation, the relative maximum pressure at the bump center is in fact a function of  $|X_d|$ . Figure 7.22 presents the associated relative minimum film thickness as a function of  $X_d$ . The largest changes occur again if the bump is located in the inlet region.

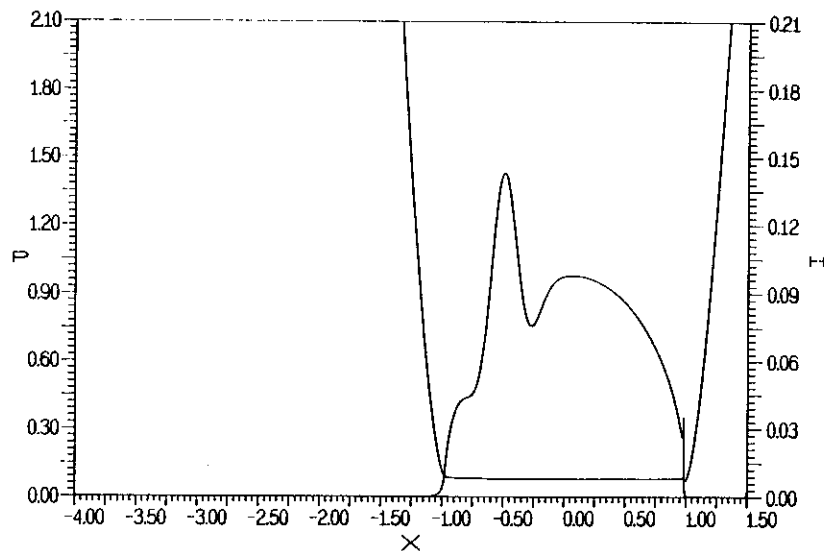


FIGURE 7.20: Pressure profile and film shape for the 3 GPa load condition in the case of a bump with an amplitude of  $2\ \mu\text{m}$  and a wavelength of  $0.5\ \text{mm}$  at  $X_d = -0.5$ .

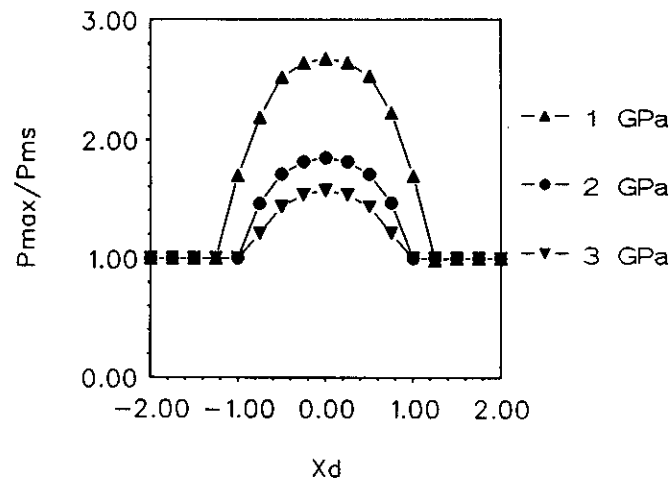


FIGURE 7.21: *Relative maximum pressure as a function of the position for a bump with an amplitude of 2 micrometer and a wavelength of 0.5 mm and three values of  $p_h$ .*

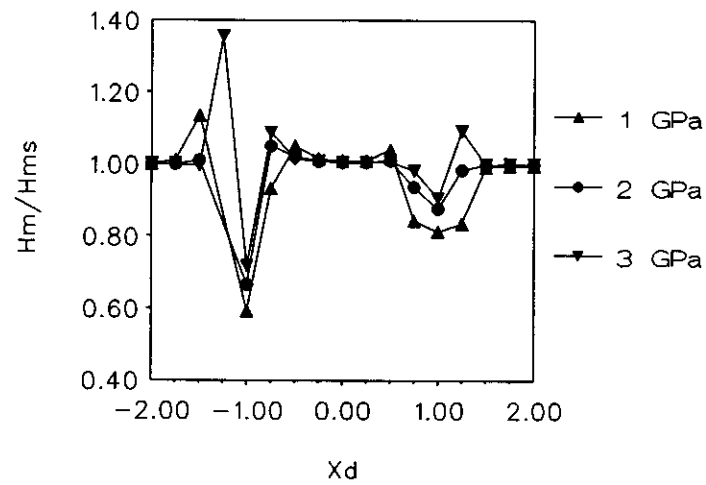


FIGURE 7.22: *Relative minimum film thickness results associated with the results presented in figure 7.21.*

#### 7.2.4.5 Implications of a bump for fatigue life

The previous sections have shown that a bump causes a significant rise of the maximum pressure in the contact. The absolute increase of the maximum pressure is larger if the wavelength is smaller and it is roughly proportional to the slope of the bump and independent of the maximum Hertzian pressure unless the wavelength of the bump is too large and the bump is no longer a local feature. With respect to the sub-surface stresses, the bump will result in a stress concentration below the surface with a value exceeding the maximum sub-surface stress in the case of perfectly smooth surfaces. This bump related stress concentration will be located closer to the surface if the wavelength is smaller and the value is roughly proportional to the ratio  $\mathcal{A}/\mathcal{W}$  unless the wavelength of the bump is too large.

#### 7.2.5 Waviness

In this section the influence of what is generally referred to as waviness on the pressure generation and fluid film formation is investigated for the same load situations as considered in the previous sections. Characteristic for waviness is that the deviation from the smooth surface parabolical shape is periodic with a wavelength of the order of magnitude of half the width of the Hertzian contact region and that it extends over the entire contact domain. Compared to the two features discussed above waviness is one step closer to real surface roughness albeit that the characteristic length scale of surface roughness is much smaller. The following harmonic waviness is considered:

$$\mathcal{R}(X) = \mathcal{A} \cos \left( \frac{2\pi(X - X_d)}{\mathcal{W}} \right) \quad (7.4)$$

where  $\mathcal{W}$  denotes the dimensionless wavelength and  $\mathcal{A}$  is the dimensionless amplitude. As in the preceding sections the steady state situation of simple sliding is assumed, i.e. only the smooth surface moves.

Figure 7.23 presents the dimensionless pressure profile and film shape for the 2 GPa load situation in the case the waviness has an amplitude of  $1 \mu\text{m}$  and a wavelength of  $0.25 \text{ mm}$  ( $\mathcal{W}=0.5$  and  $\mathcal{A}=0.056$ ). The presented results apply to  $X_d = 0$ .

From this figure it can be concluded that the waviness causes a periodic variation on the smooth surface pressure profile with a wavelength that is the same as the wavelength of the feature itself, see also [G2]. Notice the Micro EHL features occurring at the “asperities” or “ridges” on the inlet and outlet side of the Hertzian contact region. In spite of the relatively small amplitude the fluid almost cavitates at these locations. Because of the relatively small elastic deformation in both the inlet and outlet region the cavitation limitation discussed in section 7.2.3.2 applies even more to the calculations with waviness and consequently with the current model only shallow waves can be simulated. This once more stresses the need for an extension of the model to allow for pressure generation following a cavitated region.

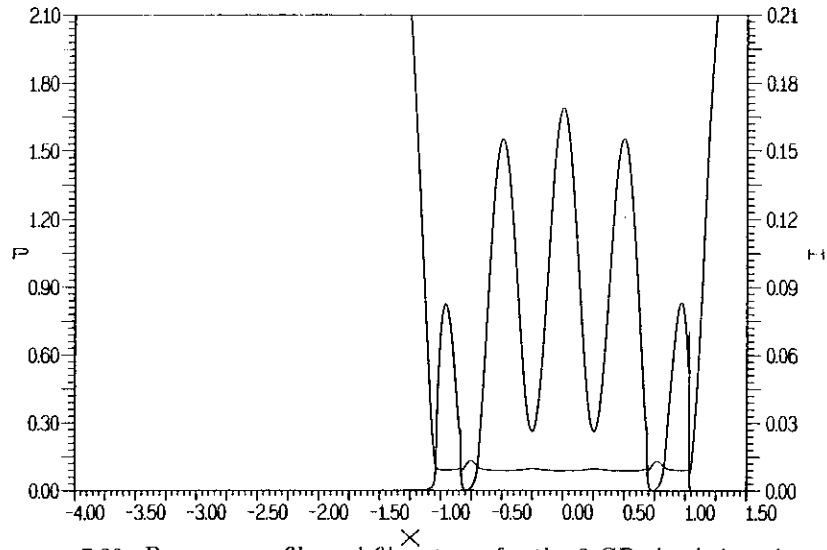


FIGURE 7.23: Pressure profile and film shape for the 2 GPa load situation with a harmonic waviness of amplitude  $1 \mu\text{m}$  and wavelength  $0.25 \text{ mm}$ .

With respect to the film thickness, figure 7.23 shows that, as was observed earlier for indentations and bumps, in the central region the surface feature is flattened and hardly shows up in the film shape.

#### 7.2.5.1 Varying wavelength and amplitude

First the influence of the amplitude and wavelength on the maximum pressure and minimum film thickness are investigated for given load conditions and position of the waviness with respect to the center of the contact. Figure 7.24 displays the relative maximum pressure as a function of the wavelength for two values of the amplitude and  $X_d = 0$ . The presented results apply to the 2 GPa load condition. As was found for indentations and bumps the relative maximum pressure is inversely proportional to the wavelength.

The calculated minimum film thickness relative to the smooth surface minimum film thickness for the situations presented in figure 7.24 is displayed in figure 7.25. This figure shows that, compared to the situation of a bump or a dent considered in the previous section, the effect of the waviness on the minimum film thickness is quite large. This is explained by the fact that the minimum film thickness is mainly determined by the inlet region of pressure generation. A bump and a dent are basically local features that, particularly if the wavelength is small, hardly affect the inlet geometry. On the contrary waviness, being a global feature, also for small wavelengths affects the inlet geometry and the pressure generation in this region. In particular, depending on the wavelength either a ridge or a valley is located in the inlet. This explains the larger effect on the minimum film thickness.



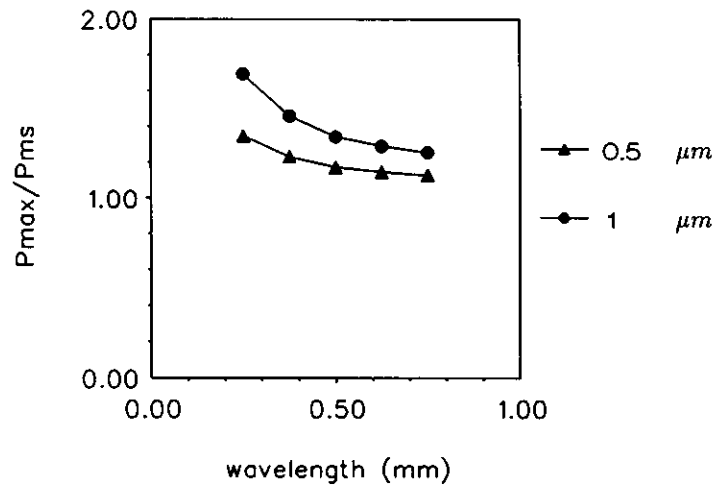


FIGURE 7.24: Relative maximum pressure as a function of the wavelength for amplitudes of 0.5 and 1  $\mu\text{m}$  for given load conditions (2 GPa).

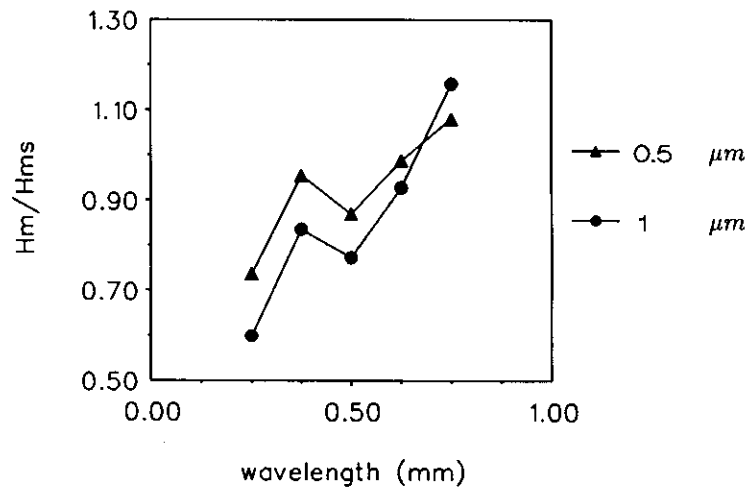


FIGURE 7.25: Relative minimum film thickness as a function of the wavelength for amplitudes of 0.5 and 1  $\mu\text{m}$  for given load conditions.

### 7.2.5.2 Varying $p_h$

Having investigated the effects of the wavelength and amplitude of the feature for a given load condition, the effect of the Hertzian pressure is studied next. The amplitude is fixed at  $0.5 \mu\text{m}$  and all results apply to  $X_d=0$ . Figure 7.26 displays the relative maximum pressure as a function of the wavelength for the 1, 2 and 3 GPa load conditions. This figure shows that the relative maximum pressure is largest for the 1 GPa load condition and smallest for the 3 GPa load condition. Reworking the data presented in figure 7.26 to the absolute increase of the maximum pressure caused by the feature gives figure 7.27. This figure shows that also in the case of an harmonic waviness the absolute increase of the maximum pressure is roughly independent of the maximum Hertzian pressure.

To conclude this section figure 7.28 presents the relative minimum film thickness data associated with the results presented in the figures 7.26 and 7.27. This figure shows that, as was already outlined in the previous section, because of its global character the waviness has a rather large effect on the minimum film thickness. In addition, this figure shows that, as was found for an indentation and a bump, for a given wavelength and amplitude the relative change of the minimum film thickness increases with decreasing maximum Hertzian pressure.

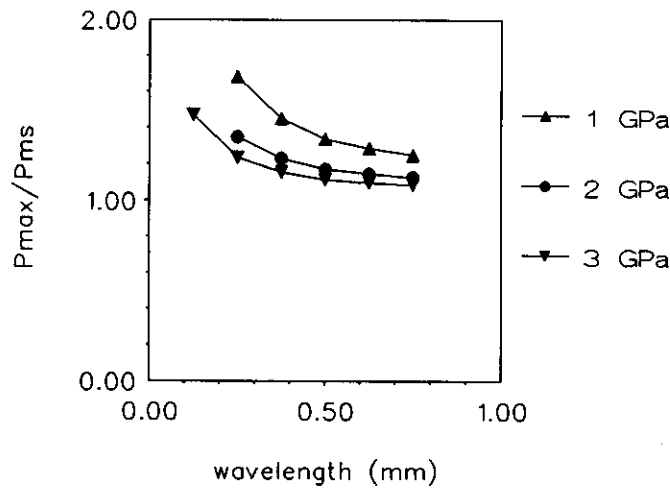


FIGURE 7.26: Relative maximum pressure as a function of the wavelength for three different values of  $p_h$ . The amplitude is fixed at  $0.5 \mu\text{m}$  and  $X_d = 0$ .

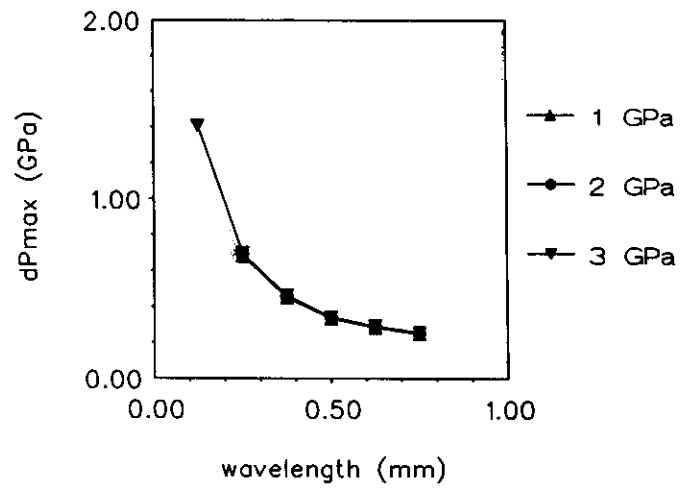


FIGURE 7.27: Absolute increase of maximum pressure as a function of the wavelength for three values of  $p_h$ . The amplitude is fixed at  $0.5 \mu\text{m}$  and  $X_d = 0$ .

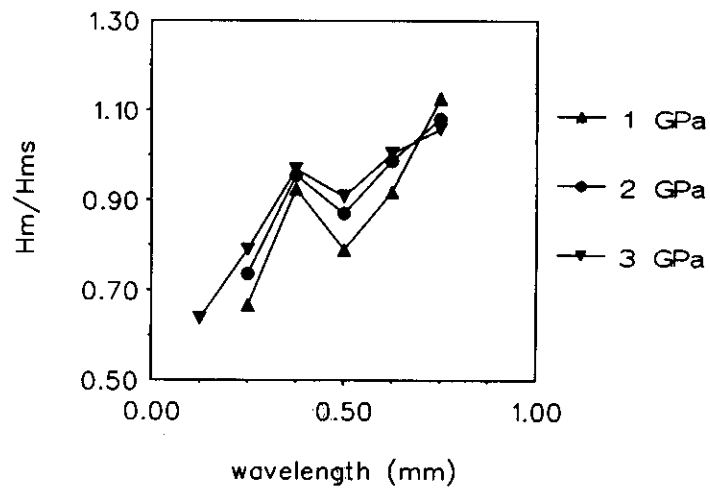


FIGURE 7.28: Relative minimum film thickness values associated with the results presented in figure 7.26.

### 7.2.5.3 Varying location

The results presented so far (section 7.2.5.1 and 7.2.5.2) applied to a waviness with a fixed position relative to the center of the contact, i.e.  $X_d = 0$ . In the present section the effect of this position on the relative maximum pressure and minimum film thickness is investigated. Obviously the results will be periodic in  $X_d$  with a wavelength equal to the wavelength of the feature. Assuming a waviness with a wavelength of  $0.25\text{ mm}$  and an amplitude of  $0.5\text{ }\mu\text{m}$  figure 7.29 presents the relative maximum pressure as a function of  $X_d$ . Obviously for this situation the maximum pressure is relatively insensitive to the position of the wave with respect to the center of the contact.

This does not apply to the minimum film thickness as can be seen from figure 7.30 presenting the relative minimum film thickness values associated with the results presented in figure 7.29. Depending on the position of the feature relative to the center in this specific situation either a decrease or an increase in film thickness is found. The two extreme values occur for  $X_d = 0$  and  $X_d = 0.25$  ( $X_d = 0.5W$ ). In the first situation the minimum film thickness is decreased by some 20 percent, whereas for  $X_d = 0.25$  an increase is observed that adds up to 20 percent for the 2 and 3 GPa load situation. Because the waviness extends over the entire domain a shift of the relative position with respect to the center of the contact implies a change of the inlet geometry.

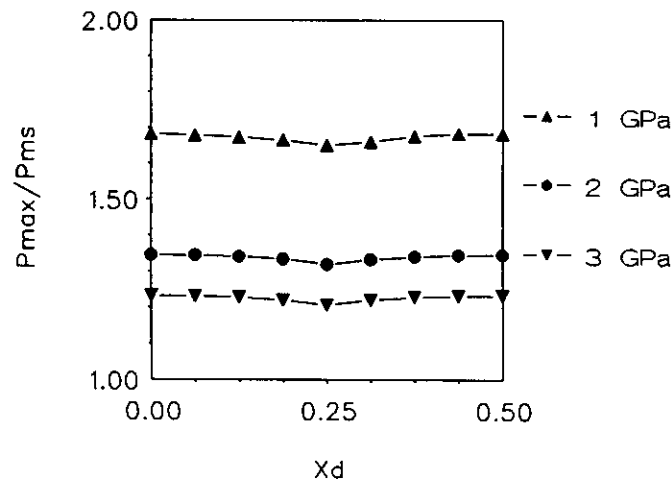


FIGURE 7.29: Relative maximum pressure as a function of  $X_d$  for three values of  $p_h$ . The wavelength is fixed at  $0.25\text{ mm}$  and the amplitude is fixed at  $0.5\text{ }\mu\text{m}$ .

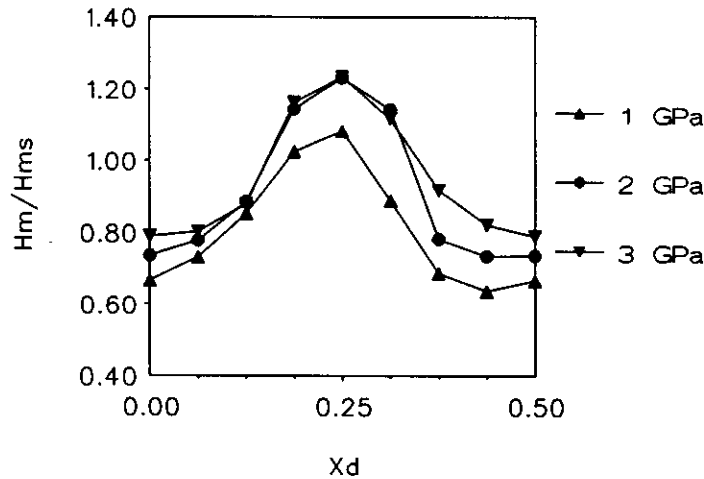


FIGURE 7.30: *Relative minimum film thickness values associated with the results presented in figure 7.29.*

Since the inlet region determines the minimum film thickness this explains the relatively strong influence of the relative position of the waviness on the minimum film thickness.

### 7.3 Surface roughness

Over the last few decades a large number of papers has been published considering the influence of surface roughness on pressure generation and fluid film formation in lubricated contacts. For a general review of work on this topic the reader is for example referred to Elrod [E2]. The majority of the papers is restricted to hydrodynamic lubrication. In addition they are generally limited to so called *Reynolds roughness*. This type of roughness has small “slopes” (the variations in  $z$  direction are small compared to the variations in  $x$  and  $y$  direction) so the underlying assumptions of Reynolds’ equation are not violated and this equation can be used to describe the fluid flow. As mentioned in the introduction to this chapter this assumption is also made in the present work. When the slopes of the surface feature or the roughness profile are too large, the Stokes equation must be used to properly describe the fluid flow. Consequently, that type of roughness is called *Stokes roughness*. To the authors knowledge only a few papers dealing with the latter type of roughness have been presented, e.g. Sun and Chen [S2] and Phan Thien [P1].

The different studies of Reynolds roughness have in common that, based on certain properties of the roughness height distribution, an averaged Reynolds equa-

tion is derived, e.g. Tzeng and Saibel [T1], Christensen [C3], Elrod [E3], Patir and Cheng [P2], Phan-Thien [P3], Tønder [T2], and Chengwei and Linqing [C4]. In particular the so called *flow factor method* has become quite popular. This method was introduced by Patir and Cheng [P2] and further work along this line was carried out by Elrod [E4], Teale and Lebeck [T3], and Tripp [T4]. An average Reynolds equation is introduced with flow factors incorporating the effect of the fluid flow past rough surfaces. For example, the equation for the case of an incompressible lubricant presented by Patir and Cheng [P2] reads:

$$\frac{\partial}{\partial x} \left( \Phi_x \frac{h^3}{12\eta} \frac{\partial \bar{p}}{\partial x} \right) + \frac{\partial}{\partial y} \left( \Phi_y \frac{h^3}{12\eta} \frac{\partial \bar{p}}{\partial y} \right) = \frac{(u_1 + u_2)}{2} \frac{\partial \bar{h}_T}{\partial x} + \frac{(u_1 - u_2)}{2} \sigma \frac{\partial \Phi_s}{\partial x} + \frac{\partial \bar{h}_T}{\partial t} \quad (7.5)$$

where:  $\bar{p}$  = mean hydrodynamic pressure  
 $h$  = nominal film thickness  
 $h_T$  = local film thickness:  $h + \delta_1 + \delta_2$   
 $\delta_1, \delta_2$  = roughness surface 1,2  
 $\bar{h}_T$  = average film thickness  
 $\Phi_x, \Phi_y$  = pressure flow factors  
 $\Phi_s$  = shear flow factor

The flow factors,  $\Phi_x$ ,  $\Phi_y$ , and  $\Phi_s$ , are determined numerically simulating the flow around the rough boundaries. They are generally given as a function of the ratio of film thickness and standard deviation of the roughness height distribution and an orientation parameter  $\gamma$ , where  $\gamma = 1$ ,  $\gamma = 0$ , and  $\gamma = \infty$  represent isotropic, transverse and longitudinal roughness respectively. Although originally developed for hydrodynamic lubrication this flow factor method has been widely applied in the study of surface roughness effects in an EHL line contact situation. The reader is for example referred to Patir and Cheng [P4], Tripp and Hamrock [T5], Sadeghi and Sui [S2], and Dong Zhu et al. [D1].

Notwithstanding its popularity the flow factor method has some serious limitations with respect to EHL problems. The first is that the elastic deformation of the asperities is not accounted for in the analysis. The flow factors used in the calculations are usually the factors presented by Patir and Cheng [P2] which are based on a hydrodynamic analysis only. Hence, by applying the flow factor method to EHL problems only the effect of the elastic deformation on the global geometry is taken into account and the surface roughness is in fact assumed to be rigid. Based on the results presented in the previous section for large scale features it can be expected that this assumption may not be valid and that the elastic deformation of the roughness asperities may have a significant effect on the results. This will apply particularly to the line contact situation, where, as a result of the one dimensional approximation, the fluid can not flow around an asperity.

The second disadvantage is the orientation parameter  $\gamma$ . This parameter does not uniquely define the orientation of the roughness. In particular,  $\gamma = 1$  is not

by definition equivalent with an isotropic roughness profile. Another disadvantage arises when the purpose of study is to investigate the effect of the roughness on the surface fatigue life of the contact. Different roughness profiles may have the same statistical properties. The flow factor method will give the same “averaged” pressure profile and film shape whereas in reality the profiles may very well have a completely different effect on the surface fatigue life.

In this thesis an alternative approach is proposed. The low complexity of the solver developed for the line contact problem enables an accurate description of the surface geometry. Therefore, instead of using some averaged equation, a deterministic description of the roughness can be used. In particular, the calculations can be performed with data obtained from surface roughness measurements as input. This section presents some preliminary results obtained in this way. First, section 7.3.1 discusses some calculational details. Subsequently the results of calculations for various load conditions are presented.

### 7.3.1 Calculational details

For a given surface roughness profile first the effect of the velocity is investigated for a maximum Hertzian pressure of 2 GPa. As a next step the effect of the Hertzian pressure is investigated for a given velocity. As was done in the analysis of large scale features the half width of the Hertzian contact region is fixed at 0.5 mm and, to obtain another value of the maximum Hertzian pressure, the reduced radius of curvature is varied. Apart from the velocity and reduced radius of curvature no parameters are varied. The values of the other parameters are the same as used in the study of the large scale effects, e.g. see table 7.1.

First for a specific value of the reduced radius of curvature coinciding with a maximum Hertzian pressure of 2 GPa ( $R = 1.41 \cdot 10^{-2}$ ) the effect of the roughness is investigated for four values of the velocity. The values of the Moes non dimensional parameters and the values of the dimensionless parameters  $\bar{\alpha}$  and  $\lambda$  describing the four load conditions created in this way are listed in table 7.3.

In addition table 7.3 gives the calculated minimum film thickness, and the calculated average film thickness in the Hertzian contact region,  $-1 \leq X \leq 1$ , assuming perfectly smooth surfaces:

$u_s$ [m/s]	$M$	$L$	$\bar{\alpha}$	$\lambda$	$h_{min}$ [ $\mu m$ ]	$h_{av}$ [ $\mu m$ ]
8	49.2	15.7	44.0	$1.53 \cdot 10^{-3}$	0.647	0.728
4	69.7	13.2	44.0	$7.63 \cdot 10^{-4}$	0.393	0.443
2	98.5	11.1	44.0	$3.82 \cdot 10^{-4}$	0.239	0.270
1	139.2	9.35	44.0	$1.91 \cdot 10^{-4}$	0.146	0.165

Table 7.3: Different loading conditions,  $p_h = 2$  GPa, varying  $u_s$ .

$R$ [m]	$p_h$ [Pa]	$M$	$L$	$\bar{\alpha}$	$\lambda$	$h_{min}$ [ $\mu m$ ]	$h_{av}$ [ $\mu m$ ]
$9.42 \cdot 10^{-3}$	$3.0 \cdot 10^9$	255.4	10.4	66.0	$5.68 \cdot 10^{-5}$	0.119	0.133

Table 7.4: 3 GPa load condition,  $u_s = 1$  m/s.

$$h_{av} = \frac{1}{2b} \int_{-b}^b h \, dx \quad (7.6)$$

Secondly, for a given value of the surface velocity, i.e.  $u_s = 1$  m/s, the reduced radius of curvature is changed to obtain a maximum Hertzian pressure of 3 GPa. The values of the different dimensionless parameters for this load situation are displayed in table 7.4. In addition table 7.4 gives the value of the minimum and average film thickness calculated assuming perfectly smooth surfaces.

The calculations are performed using one specific roughness profile. This measured profile is displayed in figure 7.31. The calculated value of the cla roughness,  $R_a$ , for this profile is  $3.68 \cdot 10^{-2} \mu m$  and the calculated rms roughness,  $R_\sigma$ , is  $4.56 \cdot 10^{-2} \mu m$ . The profile presented in figure 7.31 is obtained from a stylus measurement and consists of 879 height values whereas the trace length is about 1.13 mm. ( $\approx 2.25b$ )

Since the calculational domain is taken from  $X_a = -4$  to  $X_b = 1.5$  the profile can cover only part of the domain and it is extended with zero height values on both sides. The first non zero height value is taken at  $X = -1.45$ . The preliminary calculations presented here have been carried out using 1409 nodal points on the entire domain. To determine the value of the roughness height at a calculational point simple linear interpolation from the measured profile is used.

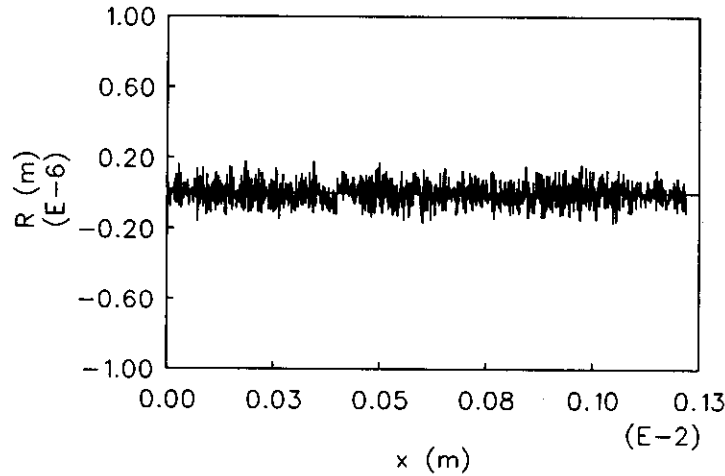


FIGURE 7.31: Roughness profile used in calculations.



### 7.3.2 Variation of speed

Figure 7.32 presents the calculated (dimensionless) pressure profile for the 2 GPa load condition and a surface velocity of 8 m/s. This figure shows that, the roughness causes a high frequency oscillation (high frequency relative to  $b$ ) on the smooth surface pressure profile. Note that the pressure spike can no longer be distinguished from the variations caused by the roughness. The associated film shape is presented in figure 7.33. On a global scale the film shape is very much alike the usual smooth surface film shape. The roughness only causes a high frequency rippling.

Similar results for a much smaller velocity, e.g.  $u_s = 1$  m/s, are presented in figure 7.34 and 7.35. Comparing figure 7.34 with figure 7.32 shows that at the larger velocity the amplitude of the oscillations is smaller. This is explained by the film thickness which is larger for the higher velocity. The lubricant film has a “damping” effect on the pressure oscillations and as a result they increase with decreasing velocity (film thickness). At this point the reader is referred to section 7.1 where pressure profiles calculated assuming a dry contact situation were referred to as *worst case* results.

Comparing the film shape presented in figure 7.33 with figure 7.35 shows that for a smaller velocity the variations on the film thickness are noticably smaller. Since both figures are plotted using the same scale this indicates an increasing elastic deformation of the roughness profile. This subject will be addressed in more detail below. First the calculated minimum and average film thickness as a function of the velocity for the rough surface situation are compared with the results obtained assuming perfectly smooth surfaces, see table 7.5.

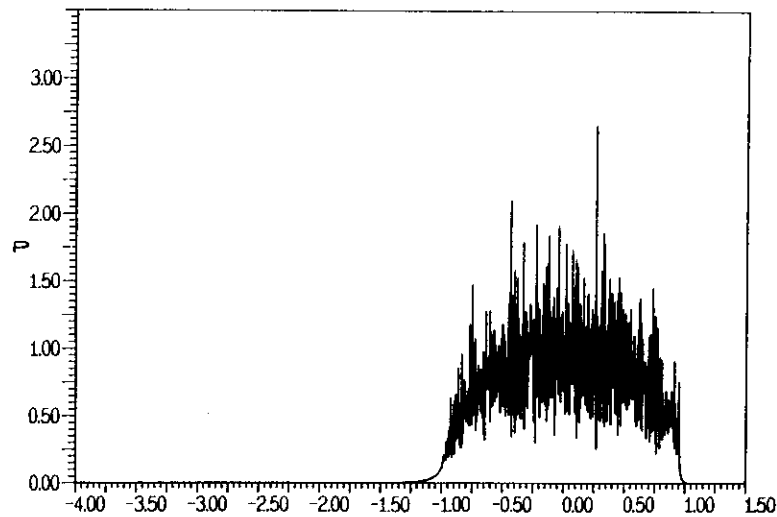


FIGURE 7.32: Pressure profile calculated with rough surface for a maximum Hertzian pressure of 2 GPa and surface velocity of 8 m/s.

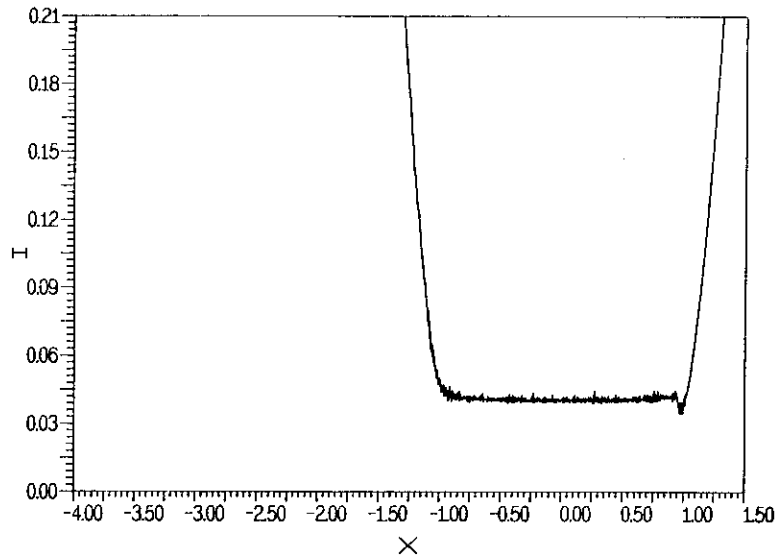


FIGURE 7.33: Film shape associated with the pressure profile presented in figure 7.32.

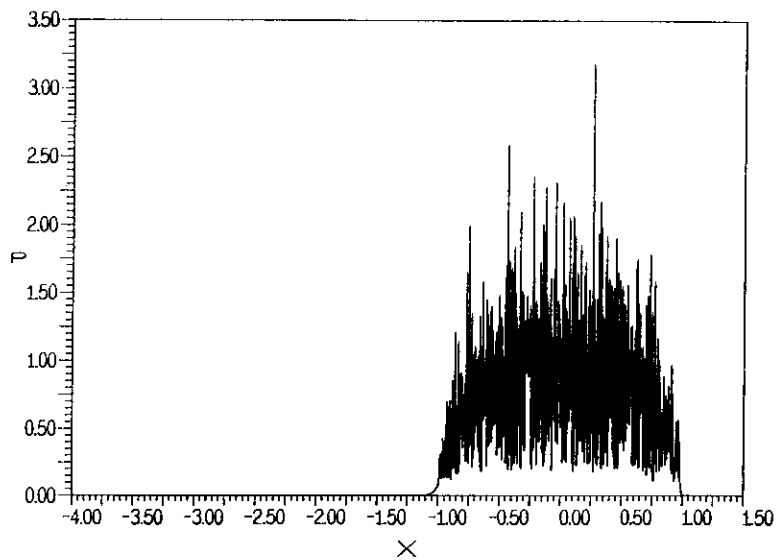


FIGURE 7.34: Pressure profile calculated with rough surface for a maximum Hertzian pressure of 2 GPa and surface velocity of 1 m/s.

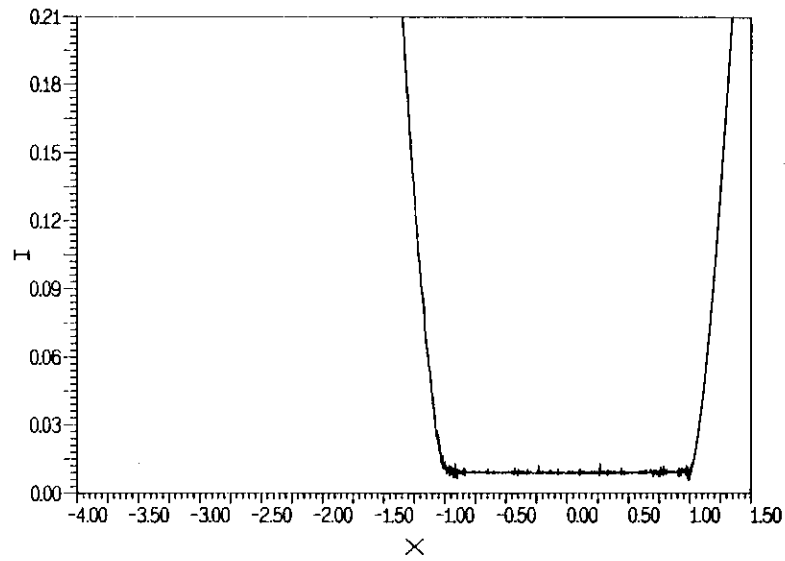


FIGURE 7.35: *Film shape associated with the pressure profile presented in figure 7.34.*

$u_s$ [m/s]	$h_{min}$ [ $\mu m$ ]	$h_{min(smooth)}$ [ $\mu m$ ]	$h_{av}$ [ $\mu m$ ]	$h_{av(smooth)}$ [ $\mu m$ ]
8	0.612	0.647	0.727	0.728
4	0.349	0.393	0.443	0.443
2	0.193	0.239	0.270	0.270
1	0.099	0.146	0.164	0.165

Table 7.5: *Comparison of the calculated minimum and average film thickness as a function of surface speed for  $p_h=2$  GPa in the case of a rough surface with the values obtained assuming smooth surfaces.*

Comparing the values obtained in the present (rough) case with the values for smooth surfaces shows that, in particular for the smaller velocities, the minimum film thickness deviates significantly from the smooth surface value. The reason can be seen from figure 7.35. In these situations the minimum film thickness occurs at one of the asperities near the end of the Hertzian contact region.

A more interesting result is that for all situations the average film thickness hardly deviates from the value found assuming smooth surfaces. This observation contradicts the results generally obtained with the flow factor method. For the situation considered here, i.e. transverse roughness, this method would predict an increase of the average film thickness due to the roughness. This increase is larger for smaller values of a parameter  $h/\sigma$ . For the 1 m/s surface velocity condition presented here the value of this parameter based on the average film thickness is approximately 3. At such a value the flow factor method would already predict a noticeable increase, e.g. [P4,D1], whereas no effect whatsoever is observed here.

To the author's opinion this is explained by the elastic deformation of the roughness profile. This elastic deformation is significant and can not be neglected as will be demonstrated. Figure 7.36 presents an artificial film thickness for the 1 m/s conditions in the contact region. This film thickness was obtained by superimposing the undeformed roughness profile on the calculated smooth surface film thickness. Hence, this figure gives an indication of what the film shape in the contact region would be if the asperities were assumed to be rigid as is done in the flow factor method. Figure 7.37 presents the actually calculated film thickness in the contact region for the same conditions, i.e. figure 7.37 is an enlargement of the film shape presented in figure 7.35. Comparing the two figures shows the effect of elastic deformation of the roughness, i.e. the amplitude of the film thickness variations in figure 7.37 is much smaller than the amplitude of the variations displayed in figure 7.36.

To quantify the deformation of the roughness profile table 7.6 displays the mean value, the average deviation from this value and the standard deviation of the artificial film thickness (figure 7.36) and the calculated film thickness (figure 7.37) in the Hertzian contact region:

$$h_{av} = \frac{1}{nn} \sum_i h(x_i) \quad |x_i| \leq b \quad (7.7)$$

$$R_a = \frac{1}{nn} \sum_i |h(x_i) - h_{av}| \quad |x_i| \leq b \quad (7.8)$$

$$R_\sigma^2 = \frac{1}{nn - 1} \sum_i (h(x_i) - h_{av})^2 \quad |x_i| \leq b \quad (7.9)$$

$nn$  denotes the number of gridpoints in the region  $|x| \leq b$ . This table shows that the mean values are almost the same. However, the values of  $R_a$  and  $R_\sigma$  differ significantly.

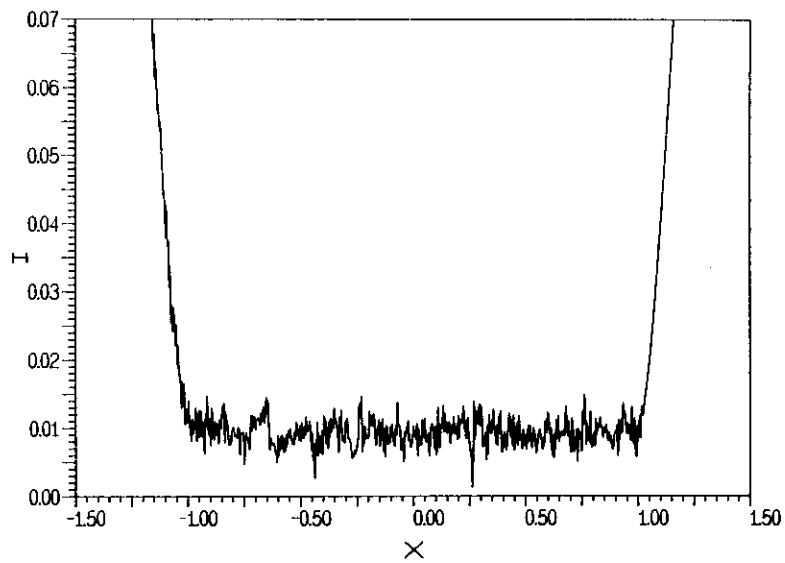


FIGURE 7.36: Film thickness in the contact region obtained by superimposing the undeformed roughness on the smooth surface film thickness.  $u_s = 1$  m/s and  $p_h = 2$  GPa.

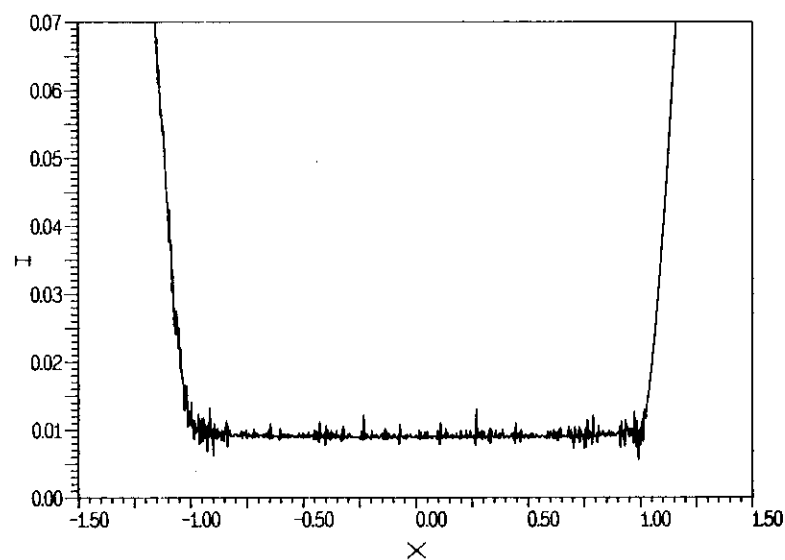


FIGURE 7.37: Actually calculated film shape in the contact region for  $u_s = 1$  m/s and  $p_h = 2$  GPa.

	$h_{av} [\mu m]$	$R_a [\mu m]$	$R_\sigma [\mu m]$
$h$ (artificial)	0.167	0.0286	0.0359
$h$ (calculated)	0.165	0.0101	0.0151

Table 7.6: Comparison of the characteristic parameters for artificially created film thickness and calculated film thickness.

For the real solution of the film thickness they are smaller, roughly by a factor of 2.5, which obviously indicates a significant elastic deformation of the roughness profile.

This finding indicates that, when studying the effects of surface roughness in stationary EHL line contacts, neglecting the deformation of the asperities can not be justified. Moreover, as can be seen by comparing the film shape for a velocity of 8 m/s and the film shape calculated for a velocity of 1 m/s, the elastic deformation of the profile increases with decreasing film thickness (velocity) and the error made by neglecting this deformation will be larger. This applies in particular to the line contact problem considered here, where, as a result of the 1 D approximation, the fluid can not flow around an asperity.

Summarizing, from the results presented in this section the conclusion can be drawn that the increase of the central or average film thickness in the case of a stationary EHL line contact with transverse roughness predicted by the flow factor method may be artefacts resulting from neglecting the deformation of the asperities.

### 7.3.3 Variation of $p_h$

To visualise the effect of the Hertzian pressure on the results, the problem is solved once more for a surface velocity of 1 m/s but this time with another value of the reduced radius of curvature resulting in a maximum Hertzian pressure of 3 GPa, e.g. see table 7.4. Figure 7.38 presents the calculated pressure profile. Comparing this figure with the results for the same velocity but for a smaller Hertzian pressure, e.g. figure 7.34 clearly shows that in the 3 GPa load situation the pressure variations are smaller. This can be explained by the smaller relative height of the roughness compared to the elastic deformation. A similar effect was found in the study of bumps, indentations and waviness. The associated film thickness for this load situation is presented in figure 7.39.

The calculated value of the minimum film thickness for this situation is  $0.069 \mu m$  whereas the average film thickness is  $0.130 \mu m$ . The minimum film thickness is smaller than the value calculated assuming smooth surfaces, e.g. see table 7.4. As was found for the same velocity but with a maximum Hertzian pressure of 2 GPa discussed in the previous section it occurs at one of the asperities. Furthermore, the calculated average film thickness hardly deviates from the value found when assuming perfectly smooth surfaces ( $\leq 2\%$ ). For the conditions considered here the flow factor method would again predict an increase of the average film thickness. The fact that such an increase is not observed once more indicates the importance of

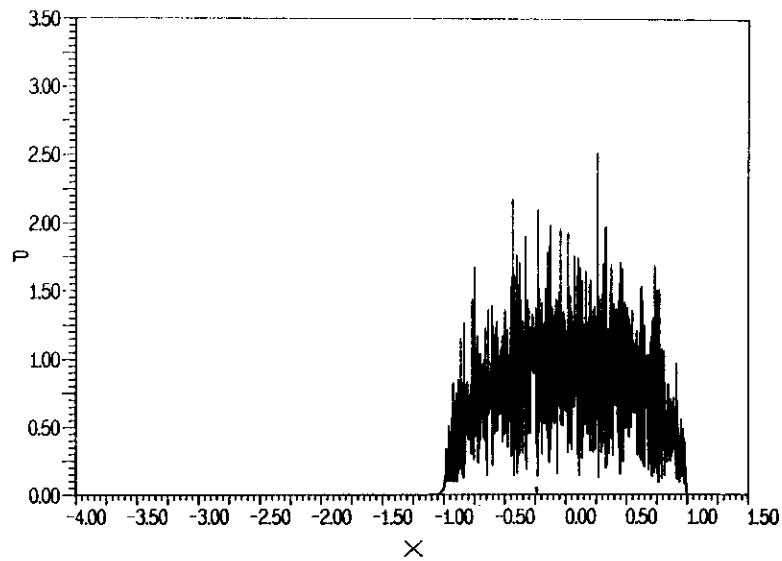


FIGURE 7.38: *Calculated pressure profile for  $p_h=3$  GPa,  $u_s=1$  m/s in the case of a rough surface.*

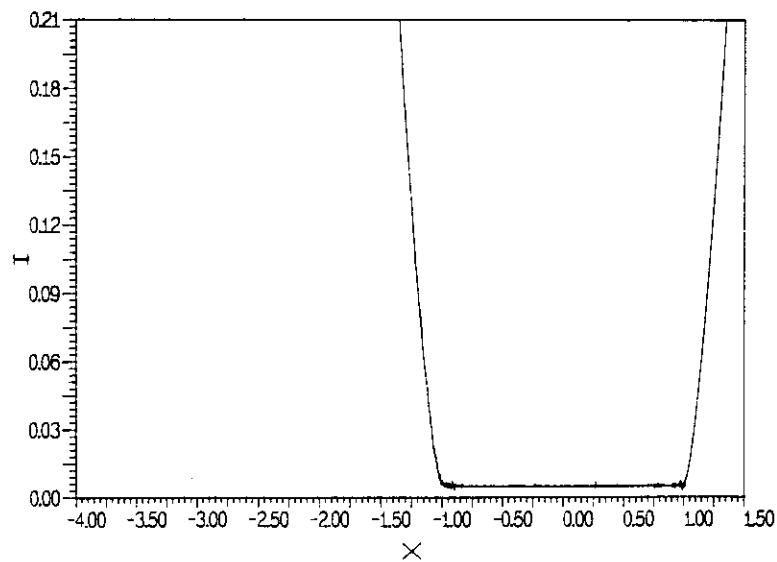


FIGURE 7.39: *Film shape associated with the pressure profile presented in figure 7.38.*

taking the elastic deformation of the asperities into account when studying surface roughness effects in EHL line contact situations.

## 7.4 Conclusion

The algorithm developed for the solution of the EHL line contact problem was used to study the effect of a number of surface features on the pressure generation and the film thickness. With respect to the large scale surface features such as bumps, indentations and waviness it was demonstrated that, when located in the Hertzian contact region, the surface feature mainly affects the maximum pressure in the contact whereas the minimum film thickness changes only slightly. This can be ascribed to the large elastic deformations. However, if located in the inlet region where the elastic deformations are small, the feature may result in a significant change of the minimum film thickness. Furthermore, it was shown that the absolute increase of the maximum pressure caused by the feature is roughly independent of the maximum Hertzian pressure.

The study of these features revealed a shortcoming of the model used in this thesis. In particular only relatively shallow dents and waviness can be simulated with the model as presented in chapter 2. The main reason is that it does not correctly model pressure generation from a cavitated region.

Furthermore, the algorithm was used to solve the pressure profile and film shape using the input of actually measured surface roughness. Although the results were limited to a single roughness profile they clearly indicated the importance of incorporating the elastic deformation of the asperities, when studying surface roughness effects in EHL.

---



## References

- [C1] **Chow, L.S.H., and Cheng, H.S.**, (1976), "Pressure perturbation in EHD contacts due to an ellipsoidal asperity," *ASME JOT*, **98**, 8-15.
- [C2] **Cheng, H.S.**, (1977) "On some aspects of microelastohydrodynamic lubrication," *Proc. Leeds Lyon conference on Tribology 1977*, 71-79.
- [C3] **Christensen, H.**, (1970), "Stochastic models for hydrodynamic lubrication of rough surfaces," *Proc. Instn Mech Engrs*, **184**, 1, 1013-1026.
- [C4] **Chengwei, W., and Linqing, Z.**, (1989), "An average Reynolds equation for partial film lubrication with a contact factor," *ASME JOT*, **111**, 188-191.
- [D1] **Dong Zhu, Cheng, H.S., and Hamrock, B.J.**, (1990), "Effect of surface roughness on pressure spike and film constriction in elastohydrodynamically lubricated line contacts," *ASLE Transactions*, **33**, 2, 267-273.
- [E1] **Elrod, H.G.**, (1981), "A cavitation algorithm," *ASME JOT*, **103**, 350-354.
- [E2] **Elrod, H.G.**, (1977), "A review of theories for the fluid dynamic effects of roughness on laminar lubricating films," *Proceedings Leeds Lyon conference on Tribology, Lyon*
- [E3] **Elrod, H.G.**, (1973), "Thin-film lubrication theory for Newtonian fluids with surfaces possessing striated roughness or grooving," *ASME JOT*, **95**, 484-489.
- [E4] **Elrod, H.G.**, (1979), "A general theory for laminar lubrication with Reynolds roughness," *ASME JOT*, **101**, 8-14.
- [G1] **Goglia, P.R., Cusano, C., and Conry, T.F.**, (1984), "The effects of irregularities on the elastohydrodynamic lubrication of sliding line contacts, Part I: single irregularities," *ASME JOT*, **106**, 104-112.
- [G2] **Goglia, P.R., Cusano, C., and Conry, T.F.**, (1984), "The effects of irregularities on the elastohydrodynamic lubrication of sliding line contacts, Part II: wavy surfaces," *ASME JOT*, **106**, 113-119.
- [I1] **Ioannides, E., and Harris, T.A.**, (1985), "A new fatigue life model for rolling bearings," *ASME JOT*, **107**, 367-378.
- [L1] **Lee, R.T., and Hamrock, B.J.**, (1990), "A circular non-Newtonian fluid model: part II - used in microelastohydrodynamic lubrication," *ASME JOT*, **112**, 497-505.
- [L2] **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*, to appear.

- [L3] **Lubrecht, A.A., Venner, C.H., Lane, S., Jacobsen, B., and Ioannides, E.,** (1990), "Surface damage - comparison of theoretical and experimental endurance lives of rolling bearings," *Proceedings of the 1990 Japan International Tribology Conference, Nagoya, Japan*, **1**, 185-190.
  - [L4] **Lorösch, H.K.,** (1985), "Research on longer life for rolling element bearings," *Lubrication Engineering*, **41**, 37-43.
  - [P1] **Phan-Thien, N.,** (1981), "On the effects of the Reynolds and Stokes surface roughnesses in a two-dimensional slider bearing," *Proc. R. Soc. London.*, **A 377**, 349-362.
  - [P2] **Patir, N., and Cheng, H.S.,** (1978), "An average flow model for determining effects of three-dimensional roughness on partial hydrodynamic lubrication," *ASME JOT*, **100**, 12-17.
  - [P3] **Phan-Thien, N.,** (1982), "On the mean Reynolds equation in the presence of homogeneous random surface roughness," *ASME JAM*, **49**, 476-480.
  - [P4] **Patir, N., and Cheng, H.S.,** (1978), "Effect of surface roughness orientation on the central film thickness in EHD contacts," *Proceedings 1978 Leeds Lyon conference on Tribology*, 15-21.
  - [S1] **Sayles, R., and Mcpherson, P.B.,** (1982), "Influence of wear debris on rolling contact fatigue," *Rolling contact fatigue testing of bearing steels*, ASTM STP 771, 1982, J.J. C. Hoo (Ed), 255-274.
  - [S2] **Sun, D.C., and Chen, K.K.,** (1977), "First effects of Stokes roughness on hydrodynamic lubrication technology," *ASME JOT*, **99**, 2-9.
  - [S2] **Sadeghi, F., and Sui, P.C.,** (1989), "Compressible elastohydrodynamic lubrication of rough surfaces," *ASME JOT*, **111**, 56-62.
  - [T1] **Tzeng, S.T., and Saibel, E.,** (1967), "Surface roughness effect on slider bearing lubrication," *ASLE trans.*, **10**, 334-338.
  - [T2] **Tønder, K.,** (1984), "The hydrodynamic lubrication of rough surfaces based on a new perturbation approach," *ASME JOT*, **106**, 440-447.
  - [T3] **Teale, J.L., and Lebeck, A.O.,** (1980), "An evaluation of the average flow model for surface roughness effects in lubrication," *ASME JOT*, **102**, 360-367.
  - [T4] **Tripp, J.H.,** (1983), "Surface roughness effects in hydrodynamic lubrication: The flow factor method," *ASME JOT*, **105**, 458-465.
  - [T5] **Tripp, J.H., and Hamrock, B.J.,** (1985), "Surface roughness effects in elastohydrodynamic contacts," *Proc. 1984 Leeds Lyon Symposium on Tribology*, 30-39.
-

## Chapter 8

# Simulation of the overrolling of a surface feature

The analysis of surface features, as described in chapter 7, introduces the need for a time-dependent approach since both surfaces generally move. This problem will be addressed in the present chapter. An accurate transient analysis requires the solution of the model equations at a large number of time steps. As the computing time needed for each time step is approximately equal to the time required for the solution of the steady state problem, it is obvious that an accurate transient simulation requires a fast algorithm for the solution of the pressure and film shape in the contact. Moreover, to simulate situations of practical importance, that is with a maximum Hertzian pressure of some 2.0 GPa, the algorithm should also be stable.

As was outlined in chapter 1 most algorithms for the solution of the steady state line and point contact problem presented over the years are of relatively high complexity. When extended to study transient situations the accuracy that can be obtained is rather limited even when high speed computers are used. Hence, the transient results presented so far, have been obtained using a relatively small number of nodes in spatial direction(s) and only a few time steps, e.g. Oh [O1], Lee and Hamrock [L1] and Chang et al. [C1]. Furthermore, most likely because of stability problems, only relatively lightly loaded situations have been studied. To avoid the aforementioned computational complexities, many papers deal only with the steady state problem.

In chapter 4 a fast and stable algorithm for the computation of the pressure distribution and film shape in EHL line contact situations was presented. This algorithm has been extended to solve transient problems, and some results with respect to the overrolling of an indentation, a bump and a specific waviness are presented in the present chapter. The low complexity of the algorithm enabled a simulation using a relatively large number of nodes in spatial direction combined with a small time step.

The results of the simulations and some implications for the fatigue life of the contact are discussed. These results show that extrapolation of results from a stationary analysis to practical conditions where the surface feature is moving can be misleading. Furthermore, the relation between the simulation results and the detection of surface defects using vibration and acoustic diagnostic techniques in the condition monitoring of bearings is discussed.

## 8.1 Equations

The one dimensional Reynolds equation written in a dimensionless way, i.e. equation (2.41), taking into account transient effects (the squeeze term) reads:

$$\frac{\partial}{\partial X} \left( \epsilon \frac{\partial P}{\partial X} \right) - \frac{\partial(\bar{\rho}H)}{\partial X} - \frac{\partial(\bar{\rho}H)}{\partial T} = 0 \quad (8.1)$$

where the cavitation condition  $P \geq 0$  should be satisfied at all times in the entire domain, and the boundary conditions are  $P(X_a) = P(X_b) = 0 \quad \forall T$ .

$\epsilon$  is given by:

$$\epsilon = \frac{\bar{\rho}H^3}{\eta\lambda}$$

where

$$\lambda = \frac{6\eta_0 u_s R^2}{b^3 p_h}$$

The dimensionless lubricant density  $\bar{\rho}$  is assumed to depend on the pressure according to the Dowson and Higginson relation and the Roelands viscosity pressure relation is used. The dimensionless film thickness equation, accounting for a moving surface feature reads:

$$H(X, T) = H_{00}(T) + \frac{X^2}{2} + \mathcal{R}(X, T) - \frac{1}{\pi} \int_{-\infty}^{\infty} P(X', T) \ln |X - X'| dX' \quad (8.2)$$

where:  $H_{00}(T)$  = the integration constant determined by the force balance condition.

$\mathcal{R}(X, T)$  = geometry of the undeformed surface feature

In the most general situation  $\mathcal{R}(X, T)$  consists of all features on both surfaces that result in a deviation of the undeformed gap from its usual parabolical shape, i.e. surface roughness, surface waviness, bumps, indentations, etc.

For reasons of simplicity only large scale features such as a dent, a bump, and waviness are considered here. In addition the feature is assumed to be located on the surface moving with velocity  $u_2$ . For example if the position of a dent or a bump at  $t = 0$  is given by  $x_s$  then its location at time  $t$  is given by:  $x_d = x_s + u_2 t$ . Hence, in terms of the dimensionless variables its position at time  $T$  is given by:

$$X_d = X_s + 2 \frac{u_2}{u_s} T$$

At all times the solution is subject to the condition of force balance, i.e. the integral over the pressure equals the externally applied contact load. Expressed in the dimensionless variables this condition reads:

$$\int_{-\infty}^{\infty} P(X, T) dX - \frac{\pi}{2} = 0 \quad \forall T \quad (8.3)$$

The equations (8.1)-(8.3) are discretized on a uniform grid with mesh size  $\Delta_x$  extending over a domain  $X_a \leq X \leq X_b$ . Using second order central discretization for the Poiseuille-term, first order upstream discretization for the wedge term, and first order backward discretization of the squeeze term, the approximation of Reynolds equation at time step  $k$  in node  $i$  reads:

$$\Delta_x^{-2}(\epsilon_{i-\frac{1}{2},k}P_{i-1,k} - (\epsilon_{i-\frac{1}{2},k} + \epsilon_{i+\frac{1}{2},k})P_{i,k} + \epsilon_{i+\frac{1}{2},k}P_{i+1,k}) - \Delta_x^{-1}(\bar{\rho}_{i,k}H_{i,k} - \bar{\rho}_{i-1,k}H_{i-1,k}) - \Delta_t^{-1}(\bar{\rho}_{i,k}H_{i,k} - \bar{\rho}_{i,k-1}H_{i,k-1}) = 0 \quad (8.4)$$

with the cavitation condition  $P_{i,k} \geq 0$ .

The discretized film thickness equation reads:

$$H_{i,k} = H_{0,k} + \frac{X_i^2}{2} + \mathcal{R}(X_i, T_k) - \frac{1}{\pi} \sum_{j=0}^n K_{ij}^{\Delta_x \Delta_x} P_{j,k} \quad (8.5)$$

where

$$K_{ij}^{\Delta_x \Delta_x} = (i - j + \frac{1}{2})\Delta_x(\ln(|i - j + \frac{1}{2}|\Delta_x) - 1) - (i - j - \frac{1}{2})\Delta_x(\ln(|i - j - \frac{1}{2}|\Delta_x) - 1)$$

The dimensionless force balance equation reads after discretization:

$$\Delta_x \sum_{j=1}^{n-1} \frac{(P_{j,k} + P_{j+1,k})}{2} - \frac{\pi}{2} = 0 \quad \forall k \quad (8.6)$$

## 8.2 Numerical solution

Assume for the moment the overrolling of a feature with a limited length, e.g. a bump or an indentation. The situation in the case of waviness is somewhat different and will be discussed later. Before the leading edge of the bump or dent reaches the boundary of the calculational domain, the squeeze term in the Reynolds equation will be zero in all nodes on the grid. Hence, film thickness and pressure profile are identical to the stationary smooth surface solution. An algorithm allowing a fast and accurate calculation of this solution was described in detail in chapter 4.

From the moment the feature enters the domain the squeeze term will no longer be negligible. At each time step the pressure and film thickness are to be solved

from equations (8.4) to (8.6) using the solution of the previous time step for the evaluation of the squeeze term. When the trailing edge of the feature passes the exit boundary of the domain the steady state solution slowly returns.

In the case of waviness the feature is present in the entire domain also at the start of the simulation. Hence, the starting solution is not the smooth surface solution. In this thesis the simulation is started with the stationary (simple sliding) solution, i.e. the solution calculated for zero velocity of the wavy surface. Subsequently at each time step the pressure and film thickness are solved from equations (8.4) to (8.6). As will be shown in section 8.5, the use of the stationary solution at the start of the simulation causes some numerical “running in” effects. However, after a short time the pressure profile and film shape will be periodic in time with a wavelength that equals the spatial wavelength of the waviness since pure rolling is assumed.

The relaxation process solving the stationary problem is also suitable for the solution of the equations at each time step, provided the squeeze term is properly accounted for in the algorithm. Convergence of the process can be accelerated using the coarser grids. Hence, given a first approximation of the solution at a specific time step the equations can be solved by repeating coarse grid correction cycles until the desired accuracy has been obtained. The required number of cycles depends of course on the accuracy of the first approximation.

The most straightforward approach is to use the pressure profile of the previous time step as a first approximation to the solution on the current time step. Such an algorithm was, for example, employed by Woods et al. [W1] when solving the dynamically loaded journal bearing problem. Hence, at each time step the coarser grids are only used to accelerate convergence of the relaxation process on the finest grid. The disadvantage of this technique is that the error in the first approximation at the finest grid contains all frequency components the grid can represent and at each time step the starting residual is rather large.

A more accurate first approximation, and consequently a reduction in the number of coarse grid correction cycles needed per time step, can be obtained if the coarser grids are employed in a way rather similar to the Full Multi Grid process for stationary problems. This alternative, the so-called *F* cycle, was developed by Brandt and co-workers and has been described in section 3.1.9.

The calculational results presented in this chapter apply to the 2 GPa conditions described in chapter 7, e.g. see table 7.2.

### 8.3 The overrolling of an indentation

The same dent geometry as used in the previous chapter is assumed, e.g. see figure 7.1. Consequently  $\mathcal{R}(X, T)$  is given by:

$$\mathcal{R}(X, T) = \mathcal{A} 10^{-10 \left( \frac{X - X_d}{W} \right)^2} \cos \left( 2\pi \frac{X - X_d}{W} \right) \quad (8.7)$$


---

where:  $\mathcal{A}$  = dimensionless amplitude of the dent  
 $\mathcal{W}$  = dimensionless wavelength of the dent  
 $X_d$  = dimensionless position of the dent center at time  $T$ .

Assuming pure rolling conditions,  $X_d$  is defined according to:

$$X_d = X_s + T \quad (8.8)$$

$X_s$  denotes the position of the dent at  $T = 0$ . Its value is chosen such that the dent is outside the calculational domain at the start of the simulation, i.e.  $X_s < X_a - \mathcal{W}$ .

### 8.3.1 Simulation results

This section presents and discusses the results of the simulation of the overrolling of an indentation with a wavelength of  $0.5 \text{ mm}$  ( $\mathcal{W} = 1.0$ ) and an amplitude of  $2 \text{ }\mu\text{m}$  ( $\mathcal{A} = -0.11$ ). The solution has been calculated using 208 time steps ( $\Delta_t = 0.03125$ ) with 1409 nodes in spatial direction, i.e.  $\Delta_t = 8 \times \Delta_x$ . This choice of a time step that equals an integer number of times the mesh size in space is necessary to accurately monitor the variation of the pressure and film thickness at a specific location on one of the (moving) surfaces as a function of time.

Figure 8.1 shows the steady state solution at  $T = 0$  whereas figures 8.2 to 8.4 show the solutions at the times when  $X_d = -0.5, 0.0$ , and  $0.5$ , respectively. For reasons of comparison figure 8.5 shows the solution for pressure and film thickness under steady state (simple sliding) conditions with the dent located in the center of the contact, i.e. at  $X = 0$ .

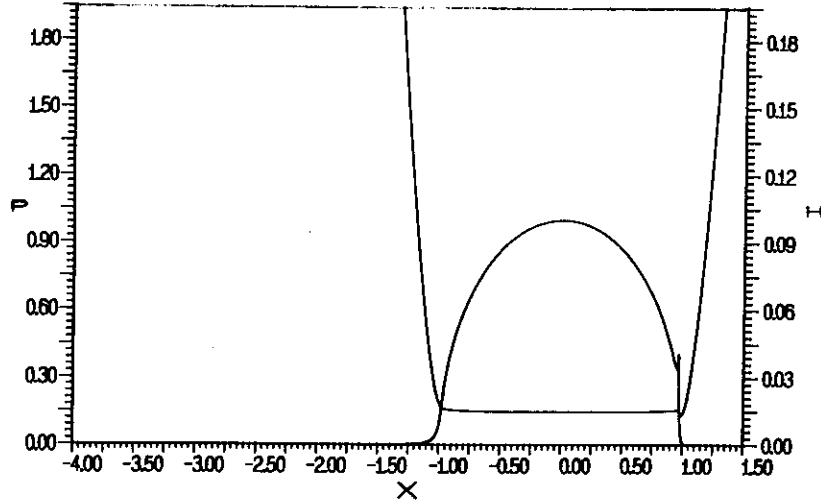


FIGURE 8.1: Solution  $M=100$ ,  $L=11.08$ , Pressure and film thickness as a function of  $X$ . Stationary solution.

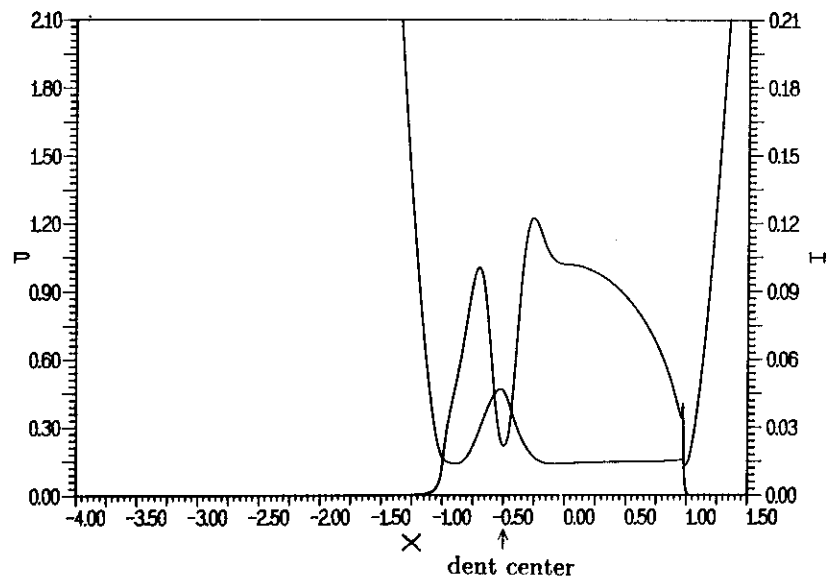


FIGURE 8.2: Solution  $M=100$ ,  $L=11.08$ , Pressure and film thickness as a function of  $X$ . Transient solution in the case of pure rolling with the dent at location:  $X_d = -0.5$ .

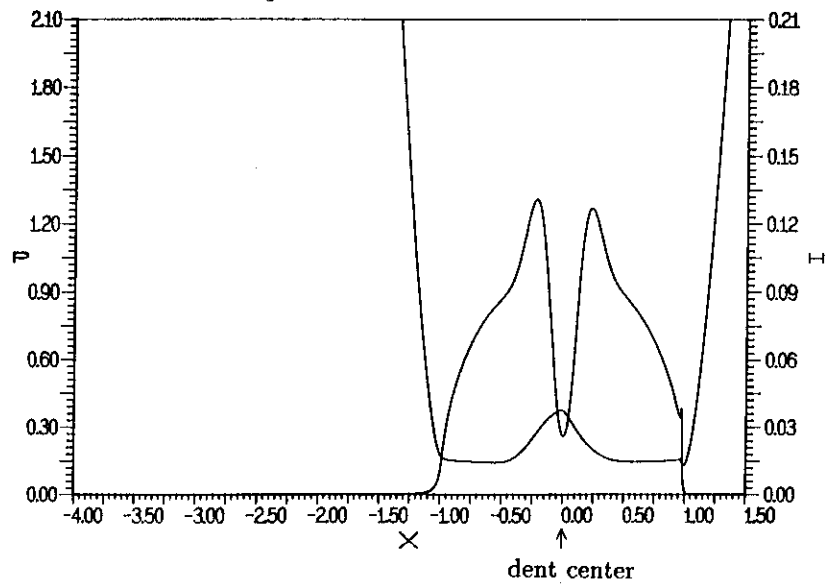


FIGURE 8.3: Solution  $M=100$ ,  $L=11.08$ , Pressure and film thickness as a function of  $X$ . Transient solution in the case of pure rolling with the dent at location:  $X_d = 0.0$ .



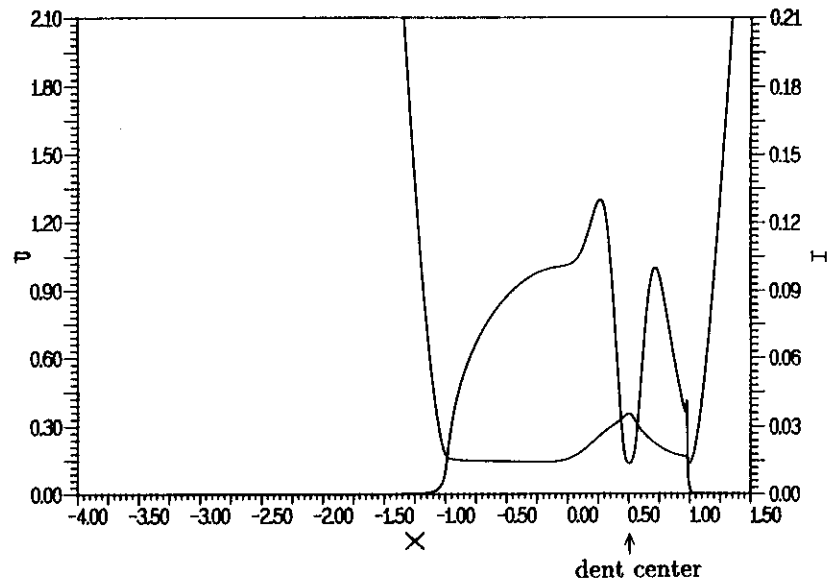


FIGURE 8.4: Solution  $M=100$ ,  $L=11.08$ , Pressure and film thickness as a function of  $X$ . Transient solution in the case of pure rolling with the dent at location:  $X_d = 0.5$ .

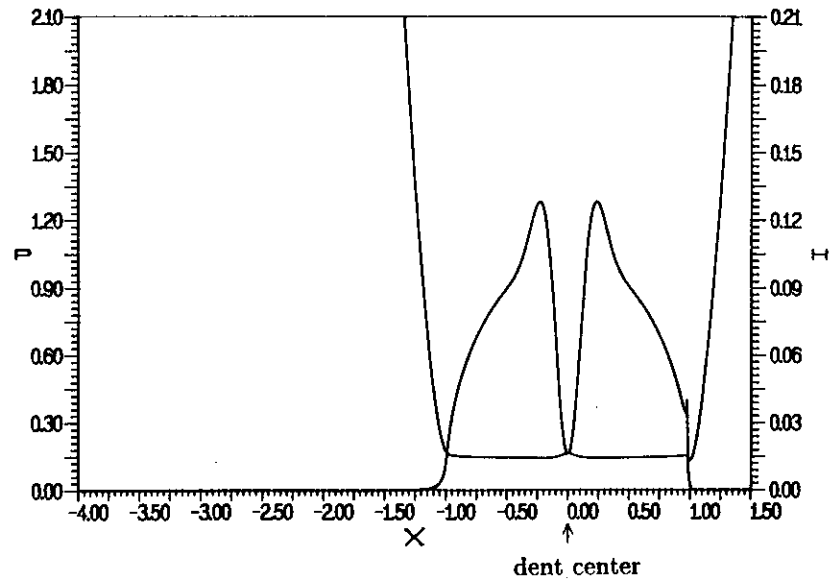


FIGURE 8.5: Stationary solution  $M=100$ ,  $L=11.08$  with the dent at  $X_d = 0.0$ , pressure and film thickness as a function of  $X$ .

Comparing this steady state solution (figure 8.5) with the transient solution (figure 8.3) illustrates the effect of the squeeze term. The most obvious difference is that in the steady state (simple sliding) situation with the dent located in the center the film thickness is nearly uniform, i.e. the dent is flattened whereas in the transient (pure rolling) situation when the dent is at this location it clearly shows up in the film thickness.

The flattening of the dent in the steady state situation is explained by the fact that the Reynolds equation in the contact region reduces to:

$$\frac{d(\bar{\rho}H)}{dX} \approx 0 \quad (8.9)$$

Since Dowson and Higginson's equation limits the compressibility of the lubricant to about 33 %, the result is a nearly uniform film thickness in the Hertzian contact region that rather closely resembles the smooth surface solution, e.g see also chapter 7.

The much smaller change of the dent geometry in the transient situation illustrated in figure 8.3 can be explained using a similar argument as given above. In the transient situation the Reynolds equation in the contact region reduces to:

$$-\frac{\partial(\bar{\rho}H)}{\partial X} - \frac{\partial(\bar{\rho}H)}{\partial T} \approx 0 \quad (8.10)$$

This equation prescribes that by approximation  $(\bar{\rho}H)$  and, since  $\bar{\rho}$  is limited, at high loads eventually  $H$ , will be a function of  $(X - T)$  only. Hence, in the limit it describes a film thickness change moving a dimensionless distance one per dimensionless unit of time. In the pure rolling condition assumed here, this equals the velocity of the feature.

For the full problem this implies that for sufficiently high loads there is a tendency to preserve the geometry of the feature during the time it moves through the Hertzian contact region. This tendency can indeed be recognized in the results presented in the figures 8.2 through 8.4.

Nevertheless, this still does not explain the smaller change in the geometry of the feature compared to the steady state situation. However, note that equation (8.10) only states that the squeeze and wedge term should be equal but of opposite sign. Because of the moving feature the squeeze term is non-zero. Hence, to satisfy equation (8.10) the wedge term must change accordingly resulting in a non-uniform film thickness.

The behavior described above can also be explained in physical terms. As a result of the high viscosities, pressure induced flow is almost absent in the contact region. Hence, shear flow dominates and the additional amount of fluid trapped in the dent upon its entrance of this region is just shifted through. However, as can be seen from the figures 8.2 through 8.4, small geometrical changes of the dent do occur during its motion through the contact region, i.e. the slopes become smaller, since the pressure gradients slowly spread the lubricant, in spite of the high viscosities.

In addition to this observation with respect to the film thickness another interesting effect shows up when comparing the transient and stationary results. The pressure profile displayed in figure 8.5 is almost symmetrical, the pressure rise at the trailing edge of the indentation equals the pressure rise at the leading edge. As can be seen from figure 8.3, in the transient situation the pressure rise at the trailing edge is larger whereas the pressure rise at the leading edge of the dent is smaller than it is in the steady state situation. This difference is obviously caused by the squeeze term.

An alternative way of presenting the results of the simulation is to monitor the pressure or film thickness at a certain point on one of the surfaces during its motion through the contact. As an example, the pressure variations experienced by two points at equal distance from the center of the dent are presented in figure 8.6. The curve on the left is for the point on the leading edge whereas the curve on the right is for the trailing edge. Both points are at a relatively large distance from the center of the dent,  $|X_d - X| = 0.5$ . Note that these are exactly the curves that a transducer on the surface at these locations would measure. Because they are relatively far removed from the center of the dent, both points experience a pressure signal over time that is equal to what any point would experience under smooth surface conditions.

Similarly, figure 8.7 shows the pressure as a function of time for  $|X_d - X| = 0.1875$ , i.e. located closer to the center of the dent. In that case the maximum pressure experienced by the point at the trailing edge is obviously larger than the maximum pressure observed by the point at equal distance from the center of the dent on the leading edge. To visualize this difference we define:

$$\mathcal{I}(X_d - X) = \int_0^{T_{end}} P(X_d - X, T) dT \quad (8.11)$$

to indicate the total "force" experienced by a point on the indented surface during overrolling. This integral is constant for points far removed from the dent. For pure rolling conditions this constant is  $\pi/2$ . For points in the vicinity of the dent the integral will deviate from this value. The value of the integral as a function of the distance from the center of the dent is presented in figure 8.8.

This graph displays an interesting feature. Far from the center of the dent, the integral for a point on the leading edge equals the value of the integral for a point on the trailing edge. However, over the entire dent the integral for points on the trailing edge is larger than the integral for points on the leading edge. The maximum difference between the two curves is some 8 % of the smooth surface value. Similar calculations have been carried out for smaller amplitudes giving the same overall result. However, in the case of a smaller amplitude the maximum difference between the two curves is also smaller. Hence, the maximum difference increases with increasing amplitude.

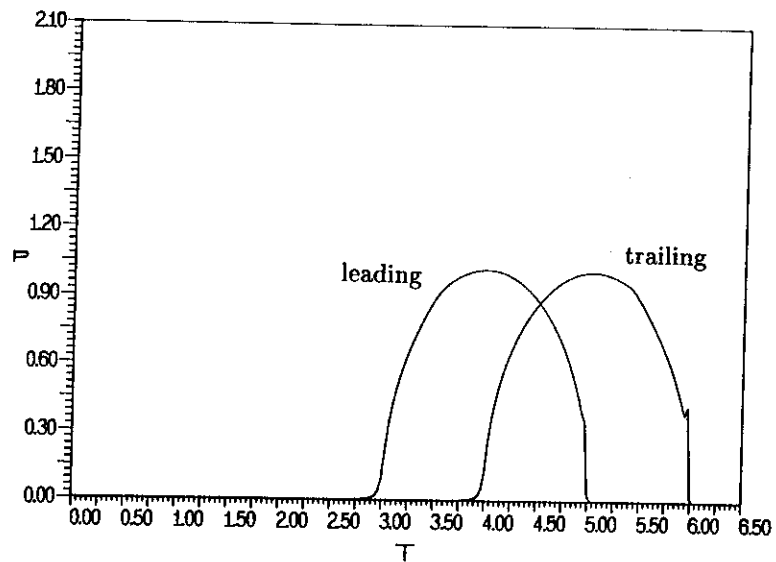


FIGURE 8.6: Pressure as a function of time in the case of pure rolling for two points at location  $X_d - X = -0.5$ ,  $X_d - X = 0.5$  on the indented surface.

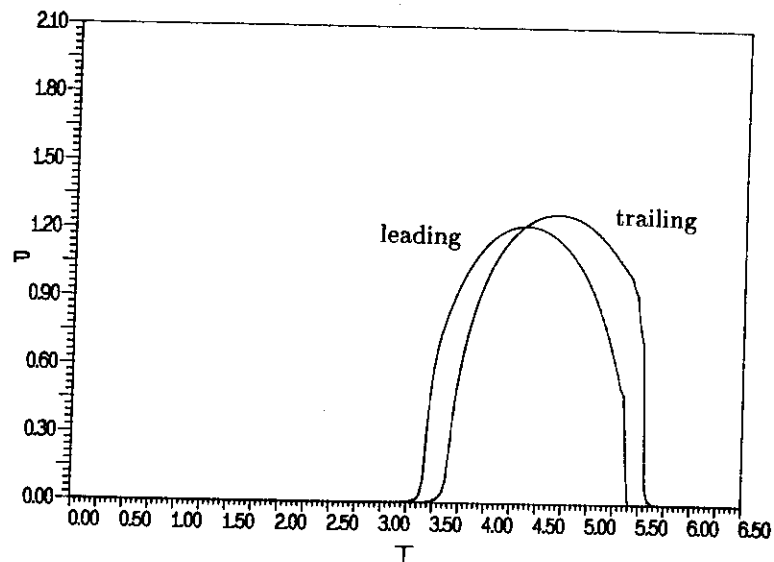


FIGURE 8.7: Pressure as a function of time in the case of pure rolling for two points at location  $X_d - X = -0.1875$  and  $X_d - X = 0.1875$  on the indented surface.

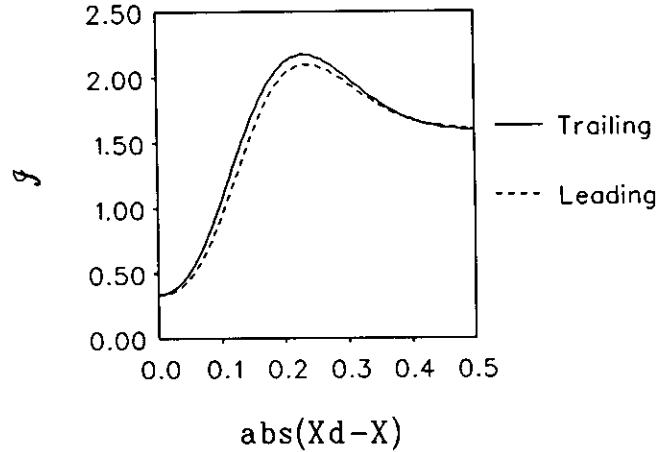


FIGURE 8.8: Value of the integral  $\mathcal{I}$  as a function of the distance from the center of the dent  $|X_d - X|$  on the indented surface (pure rolling).

These results indicate that the maximum sub-surface stress experienced by points below the leading edge will be less than the maximum sub-surface stress experienced by points below the trailing edge; this phenomenon is analyzed in detail in the next section.

### 8.3.2 Sub-surface Stresses

For the specific situation considered here, i.e. the overrolling of an indentation, the implications for the sub-surface stresses have been studied by Lubrecht et al. [L2]. The results of this study are presented in detail in the aforementioned paper and also in [V1]. However, since these results demonstrate the importance of the type of investigation carried out in this chapter they are repeated here.

In section 7.1 it was emphasized that a theoretical model to describe the relations between surface features and fatigue life is of great value since the experimental investigations are very time-consuming, energy-consuming and costly. The theoretical model used is described in detail in [L2]. For this model to be accurate, the time-dependent behaviour of the pressure in the lubricant film is essential, resulting in a time-consuming series of stress calculations (5 cpu hours on a SGI 240).

As an example, the directional preference of fatigue initiation with respect to an indentation is discussed. Generally, a spall is created after (below the trailing edge of) the indentation. Using stationary lubricated or dynamic dry contact calculations, no preference in direction can be found. However, when the time-dependent lubricated calculations are performed an asymmetry in the pressure profile is ob-

served (see figures 8.3, 8.7 & 8.8). This difference is reflected in the sub-surface stresses, which are larger (and much closer to the surface) below the trailing edge of the indentation (see figure 8.9).

Actually, the quantity displayed is not a stress, but a risk-related stress, incorporating the maximum shear stress over any angle, the hydrostatic pressure and the fatigue limit of the material (see [I1]). For convenience we will briefly refer to it as a "stress".

This stress graph was obtained by computing the sub-surface stresses resulting from the pressure distribution at each time step. For this purpose the indented surface is monitored while the pressure profile is sweeping over it, and at each location in the material the maximum stress is recorded with respect to the subsequent time steps. It is obvious that far removed from the dent the iso-stress contours should become straight horizontal lines, only close to the dent will these contours be modified. Note that the stress concentration below the trailing edge of the dent is larger and extends much closer to the surface. As is explained in [L2] the stresses can be converted to a (fatigue) risk integral. The risk integral over the left part of figure 8.9 is almost twice as high as the integral over the right part. Since the difference in stress is concentrated in a small region only (say 10%) the risk of spalling will be much higher for the trailing edge of the dent, explaining the experimentally observed preference. In real applications the depth of the indentations will be larger, and by extrapolation of the theoretical results obtained from shallower dents the asymmetry will be even more pronounced.

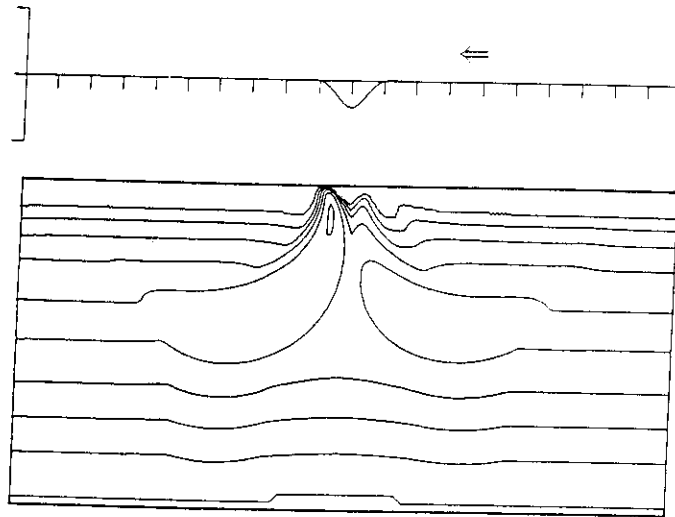


FIGURE 8.9: Maximum sub-surface stresses below a dent during overrolling calculated by Lubrecht [L1] (pure rolling). Conditions as in figure 8.1. The arrow denotes the direction of overrolling.

### 8.3.3 Condition monitoring

As mentioned above, surface indentations generally result in a significant reduction of the fatigue life of an EHL component. Because the unexpected breakdown of one of the bearings in a machine may have dangerous and/or expensive consequences, the detection of surface defects, i.e. indentations or small spalls, is an important topic in the condition monitoring of rolling bearings. One of the techniques employed is based on the analysis of the vibration or acoustic signals. To a certain extent a simulation as presented above can provide some theoretical support for the research on this subject, in particular with respect to the relation between the indentation and its contribution to the aforementioned signals. For that purpose the pressure as a function of time is monitored at a certain location fixed in space. Figure 8.10 displays the pressure in the center of the contact as a function of time for the conditions and the surface feature considered in this section. Before the leading edge of the dent reaches the specific location, and after the trailing edge passes the location the pressure equals the steady state value. The passage of the indentation causes a pressure variation. The magnitude and the frequency content of this variation will be characteristic for its contribution to the vibrations or acoustic emission in the bearing under the full film lubricated conditions assumed in this thesis.

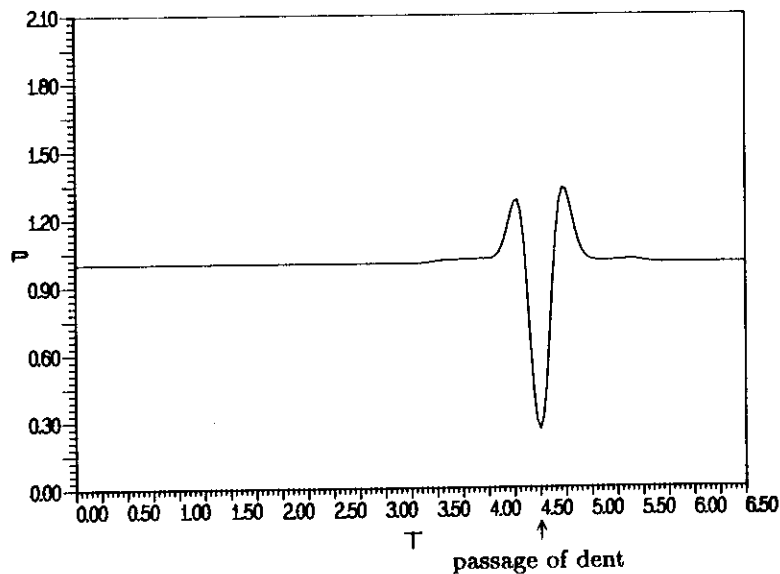


FIGURE 8.10: Pressure in the center of the contact ( $X = 0$ ) as a function of  $T$  in the case of the overrolling of a dent (pure rolling).

## 8.4 The overrolling of a bump

This section presents the results of the numerical simulation of the overrolling of a bump for the same load conditions as considered for a dent in section 8.3, i.e. the maximum Hertzian pressure is 2.0 GPa. The bump geometry is the same as the dent geometry used in the previous section, apart from the opposite sign of the amplitude and the results apply to a wavelength of 0.5 mm and an amplitude of 2  $\mu\text{m}$ . As in the previous section, the simulation has been carried out using 208 time steps with 1409 nodal points in spatial direction. At the start of the simulation ( $T = 0.0$ ) the bump is located at  $X = -4.5$ , i.e. the solution equals the smooth surface steady state solution. For this stationary solution at  $T = 0.0$  the reader is referred to figure 8.1. The figures 8.11 to 8.13 show the solutions at the times when the bump is located at  $X_d = -0.5, 0.0, 0.5$ , respectively. For reasons of comparison figure 8.14 repeats the steady state solution with the bump located at  $X_d = 0.0$  presented earlier. (figure 7.12).

Comparing the steady state solution with the bump located in the center of the contact (figure 8.14) with the transient solution at the moment the bump is at this position (figure 8.12) again shows the effect of the squeeze term. The differences in the pressure profiles are very small. In the transient situation the pressure profile in the vicinity of the bump is slightly asymmetric, i.e. the pressures at the leading edge of the bump are somewhat larger and at the trailing edge they are somewhat smaller than in the steady state situation. These differences can indeed be ascribed to the squeeze term. For example, for points on the leading edge of the bump  $\partial(\bar{p}H)/\partial T$  is smaller than zero which leads to an additional pressure generation compared to a stationary situation. Similarly on the trailing edge the squeeze term is positive which reduces the pressure generation compared to the steady state situation. Notwithstanding the small differences between the pressure profiles the differences in the film shape are more obvious. In the steady state solution the bump was hardly visible in the film shape because of the large elastic deformation. However, in the transient situation the bump clearly shows up in the film shape.



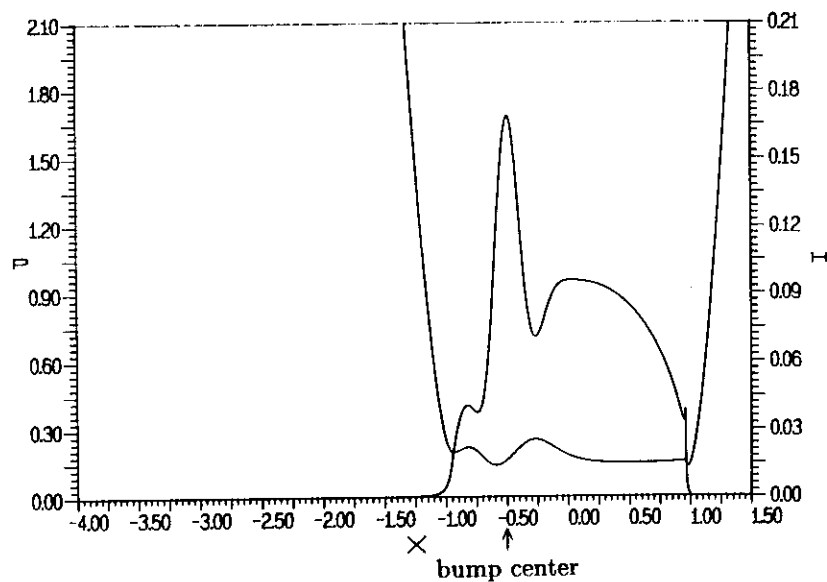


FIGURE 8.11: Solution  $M = 100, L = 11.08$ . Pressure and film thickness as a function of  $X$ . Transient solution in the case of pure rolling with the bump at location:  $X_d = -0.5$ .

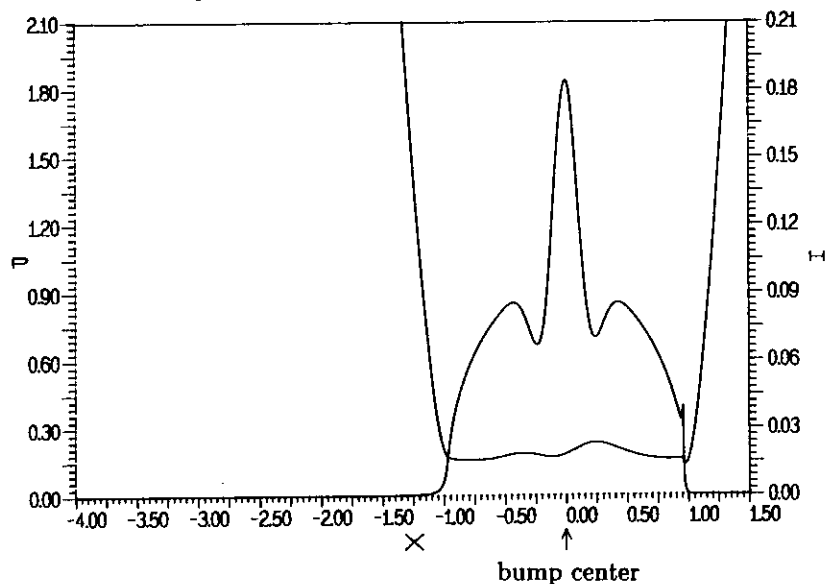


FIGURE 8.12: Solution  $M = 100, L = 11.08$ . Pressure and film thickness as a function of  $X$ . Transient solution in the case of pure rolling with the bump at location:  $X_d = 0.0$ .

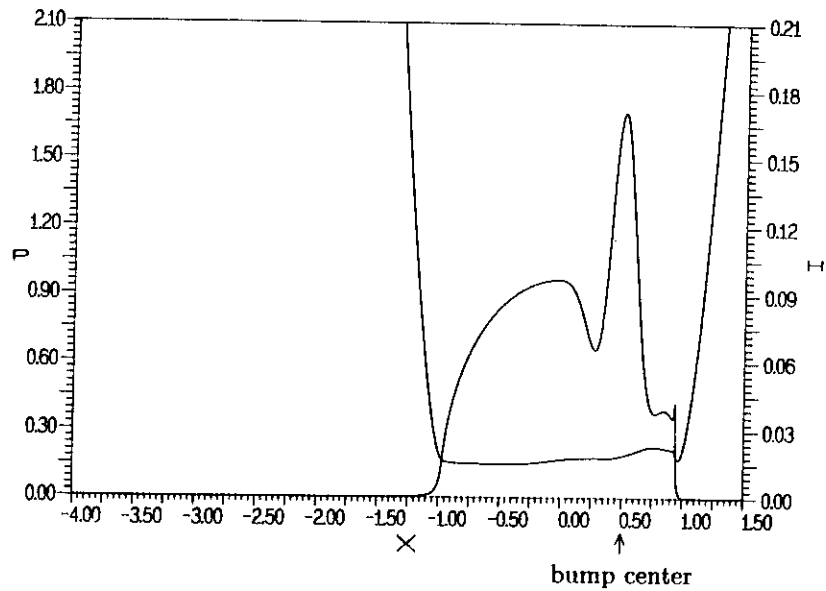


FIGURE 8.13: Solution  $M = 100, L = 11.08$ . Pressure and film thickness as a function of  $X$ . Transient solution in the case of pure rolling with the bump at location:  $X_d = 0.5$ .

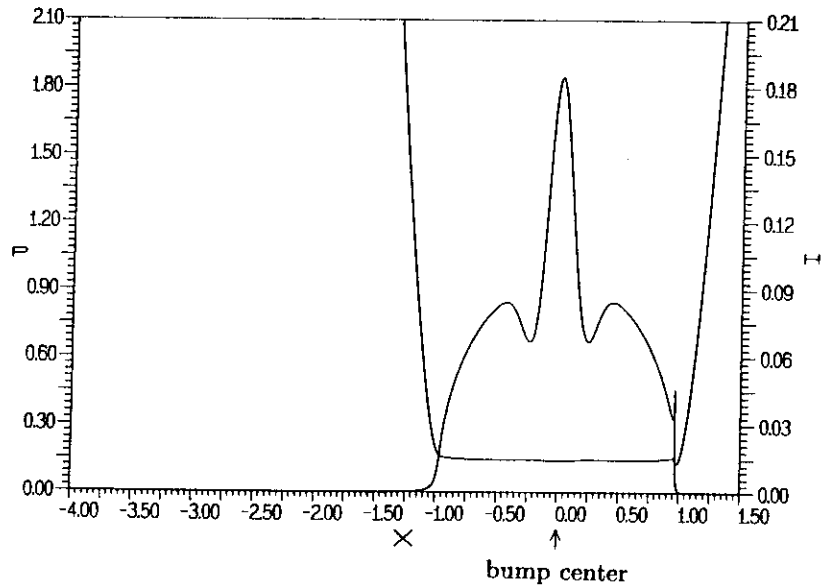


FIGURE 8.14: Stationary solution  $M = 100, L = 11.08$  with the bump at  $X_d = 0.0$ . Pressure and film thickness as a function of  $X$ .

In addition to these pressure profiles and film shapes as a function of  $X$  at different times, the pressure signal experienced by a fixed point of the surface with the bump is monitored as a function of time. For example, figure 8.15 presents the pressure as a function of time for two points on the trailing and leading edge of the bump at equal distance from its center. Since the distance from the center of the bump is relatively large,  $|X_d - X| = 0.5$ , both curves are nearly equal to the variations any point of the surface would experience if the surfaces perfectly were smooth, see also figure 8.6. However, for points located closer to the center of the bump the pressure variation experienced by a point on the leading shoulder differs from the variation experienced by a point at the same distance from the center of the bump on the trailing shoulder. For example, see figure 8.16, where the pressure as a function of time experienced by the two points for which  $|X_d - X| = 0.1875$ .

The differences between the pressure variations experienced by points on the trailing edge and the leading edge are again visualized with the integral of the pressure over time, see equation 8.11. The value of this integral,  $\mathcal{I}$ , as a function of the distance from the center of the bump is presented in figure 8.17. The largest value of the integral is found at the center of the bump. For large distances from the center of the bump the values for a point on the leading and trailing edge are equal and approximate the aforementioned constant for smooth surfaces and pure rolling conditions, i.e.  $\pi/2$ . However, in the intermediate region higher values are found for points on the leading edge. As mentioned before this difference is caused by the squeeze term causing an additional pressure generation since at this edge  $\partial(\bar{p}H)/\partial T$  is negative.

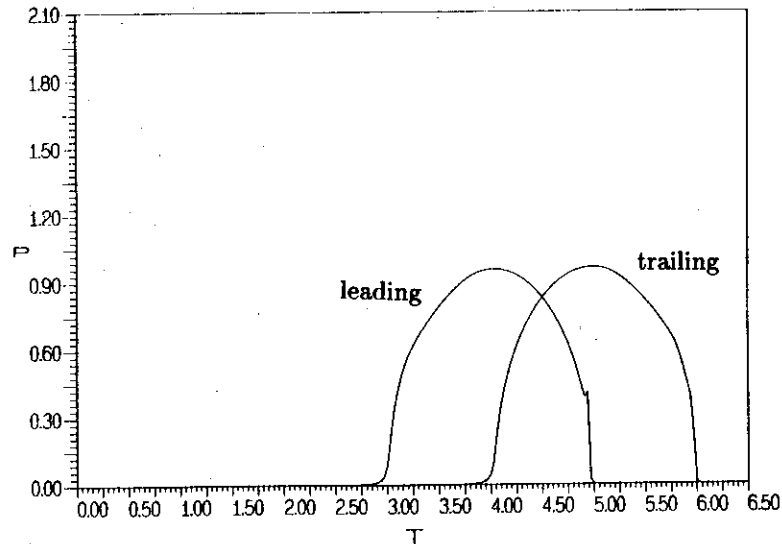


FIGURE 8.15: Pressure as a function of time for two points at location  $X_d - X = -0.5$  and  $X_d - X = 0.5$  on the bumped surface (pure rolling).

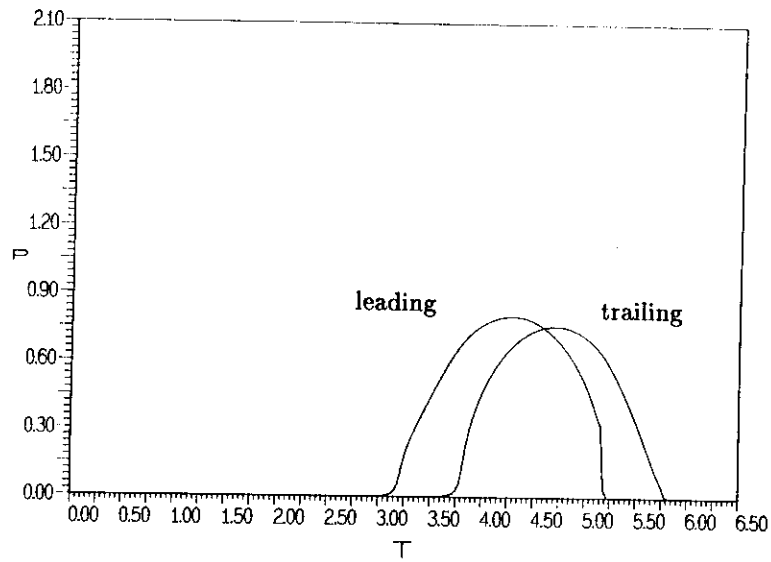


FIGURE 8.16: Pressure as a function of time for two points at location  $X_d - X = -0.1875$  and  $X_d - X = 0.1875$  on the bumped surface (pure rolling).

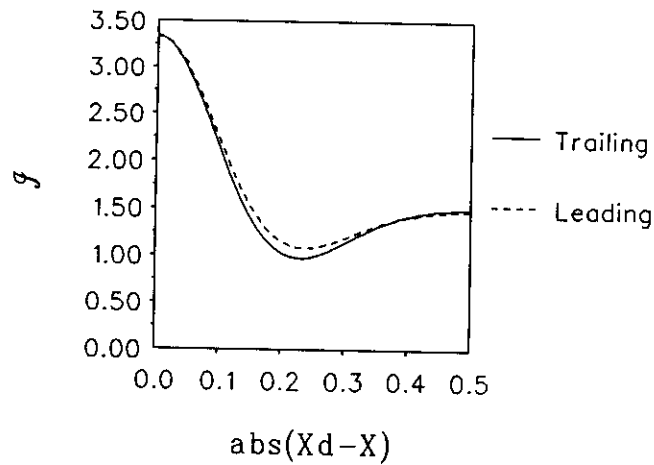


FIGURE 8.17: value of the integral  $I$  as a function of the distance from the center of the bump  $|X_d - X|$  on the surface with the bump (pure rolling).

Contrary to the situation for a dent the consequences of a bump for the fatigue life of the contact have not been studied because bumps don't happen frequently in practice. However, if they were studied the transient results would probably not lead to additional insights compared to a steady state analysis. The largest values of the integral are found at the center of the bump. It can be expected that below the center of the bump the largest sub-surface stresses will be found and spalling will most likely initiate at this location.

To conclude this section the pressure and film thickness at a fixed location in the contact region are monitored as a function of time. As was explained before the study of such "time signals" might be of considerable interest in the investigations of the source of excitation of vibrations and/or acoustic waves, which are used for condition monitoring purposes. For example figure 8.18 presents the pressure and film thickness in the center of the contact as a function of time. Before the trailing edge of the bump reaches the center of the contact the pressure and film thickness are equal to their smooth surface values. The passage of the bump leads to a significant increase in the pressure and, surprisingly, also to an increase of the film thickness. After the trailing edge of the bump passes the center of the contact the pressure and film thickness slowly return to their respective smooth surface values.

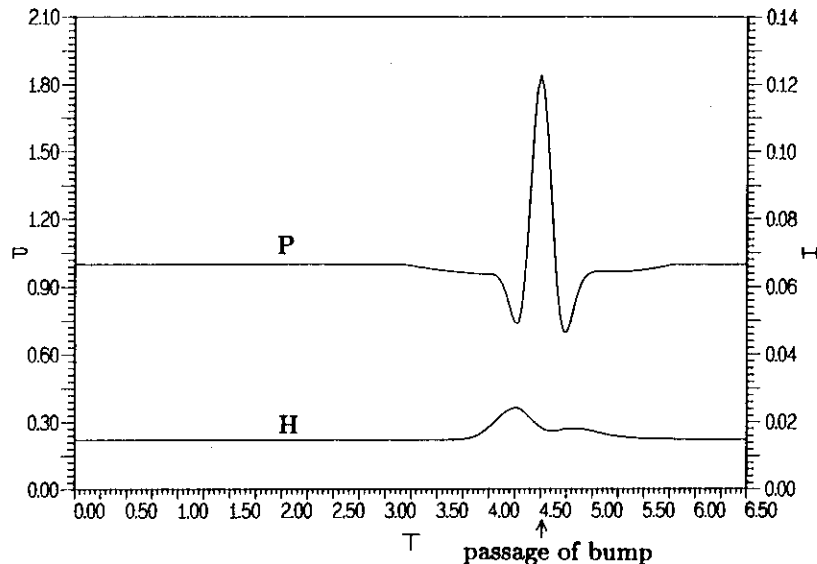


FIGURE 8.18:  $M = 100$ ,  $L = 11.08$ . Pressure in the center of the contact ( $X = 0$ ) as a function of  $T$  in the case the overrolling of a bump (pure rolling).

## 8.5 The overrolling of waviness

This section presents the results of the simulation of the overrolling of an harmonic waviness:

$$\mathcal{R}(X) = \mathcal{A} \cos \left( \frac{2\pi(X - X_d)}{\mathcal{W}} \right) \quad (8.12)$$

where  $\mathcal{W}$  denotes the dimensionless wavelength and  $\mathcal{A}$  is again the dimensionless amplitude. In the transient situation considered here  $X_d$  is defined according to:

$$X_d = X_s + 2 \frac{u_2}{u_s} T \quad (8.13)$$

The results presented in this section apply to a wavelength of 0.25 mm and an amplitude of 0.5  $\mu\text{m}$ . The load conditions are the same as assumed in the previous sections, i.e. the maximum Hertzian pressure is 2.0 GPa. Again pure rolling conditions are assumed, i.e.  $u_2/u_s = 0.5$ . Consequently, the pressure profile and film thickness are expected to be periodic with a wavelength in time that equals the spatial wavelength of the waviness. At  $T = 0$ , the start of the simulation, the solution is assumed to be the stationary solution with  $X_d = 0.0$ . The simulation has been carried out using 1409 nodes in spatial direction and a 256 time steps with  $\Delta t = 0.015625$  to accurately describe the variations in time. Hence, the simulation starts at  $T = 0$  and ends at  $T = 4.0$ .

Because the simulation is started with the stationary solution the results will only display the periodicity in time after a certain “running-in” time. As will be shown later, from  $T = 2.0$  onwards the solution is fully periodic in time with the expected wavelength. Hence, the “running in” time equals the time a point needs to pass the entire Hertzian dry contact region  $-1 \leq X \leq 1$ .

The (dimensionless) pressure and the (dimensionless) film thickness at the start of the simulation are presented in figure 8.19. This figure displays all characteristic features discussed in the previous chapter. Because of the relatively high  $\bar{\alpha}$  the waviness hardly shows up in the film shape. However, it clearly affects the pressure profile.

Figure 8.20 displays the transient solution for  $X_d = 3.0$ , i.e. at a time past the running-in time with the waviness in the same position relative to the center of the contact as at the start of the simulation. In addition figure 8.21 displays the transient (dimensionless) pressure and film shape for  $X_d = 3.25$ , i.e. at the time the waviness is in a position shifted half of its wavelength compared to the situation of figure 8.20.

Again the differences between the solutions presented in figure 8.19 and 8.20 clearly demonstrate the effect of the squeeze term. As in the case of a bump, the differences between the two pressure profiles are small. However, significant differences are found between the film thickness results.

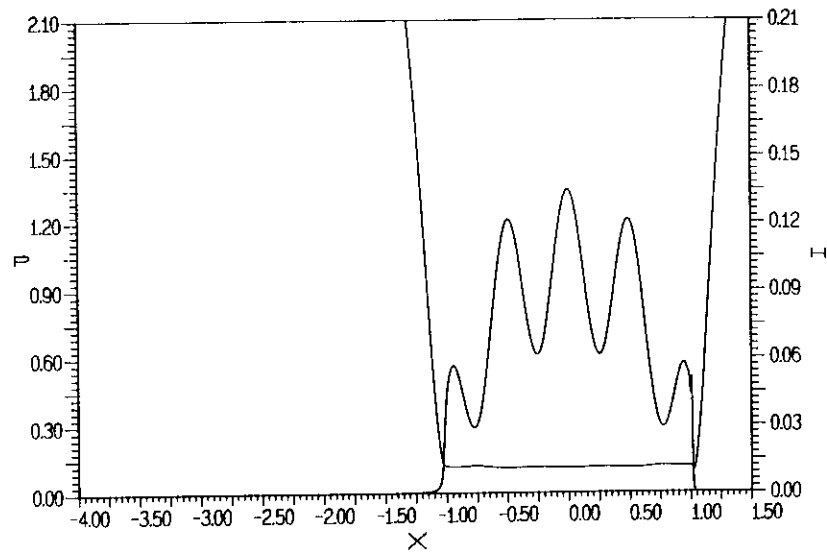


FIGURE 8.19: Solution  $M = 100$ ,  $L = 11.08$ . Pressure and film thickness as a function of  $X$ , stationary solution with waviness,  $X_d = 0.0$ .

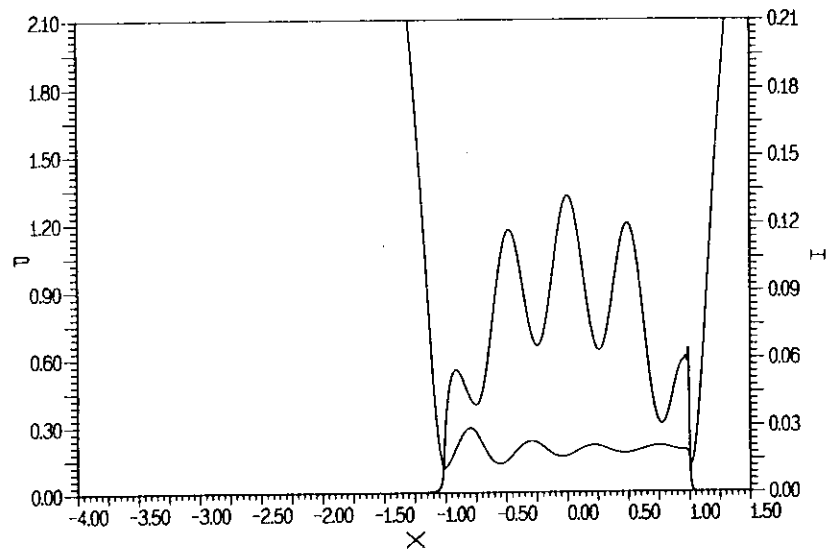


FIGURE 8.20: Solution  $M = 100$ ,  $L = 11.08$ . Pressure and film thickness as a function of  $X$ . Transient solution in the case of pure rolling when  $X_d = 3.0$ .

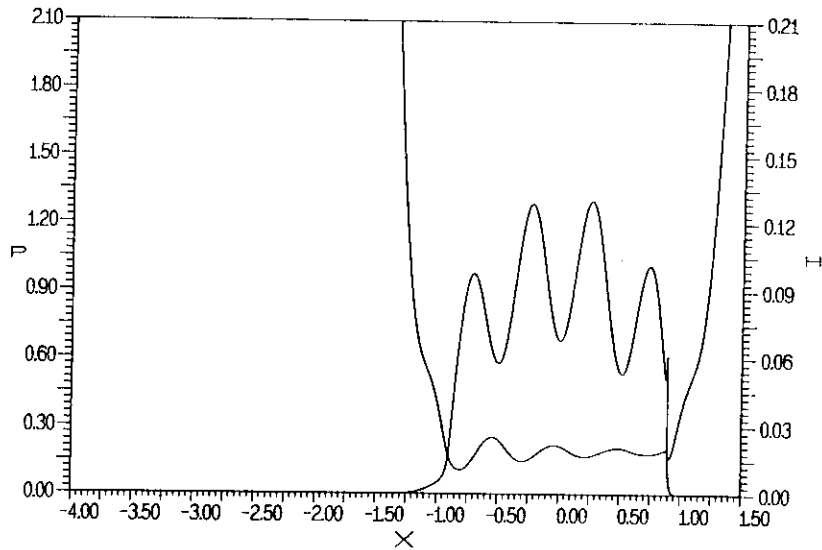


FIGURE 8.21: Solution  $M = 100$ ,  $L = 11.08$ . Pressure and film thickness as a function of  $X$ . Transient solution in the case of pure rolling when  $X_d = 3.25$ .

In the transient situation the geometry of the surface feature is much less affected by the elastic deformation, i.e. the original waviness profile is much more preserved.

As was mentioned earlier both pressure and film thickness should be periodic in time with a wavelength of 0.5. This of course applies to the pressure and film thickness at all locations in the contact region. To demonstrate this the figures 8.22 to 8.24 present pressure and film thickness as a function of time for  $X = -0.5$ ,  $X = 0.0$ , and  $X = 0.5$  respectively.

It appears that, indeed, after some "running in" time the pressure and film thickness are periodic in time. This running in time depends on the location in the contact. The pressure and film thickness in the center of the contact are periodic with respect to time from  $T = 1.0$  onwards whereas the pressure and film thickness at  $X = 0.5$  are only periodic from  $T = 1.5$  onwards. Similarly periodicity will be displayed from  $T = 2.0$  onwards for  $X = 1.0$ . For points further downstream than  $X \approx 1.0$  the pressure is zero at all times. Hence, from  $T = 2.0$  onwards the entire solution will be periodic. All three figures clearly show that the wavelength of the oscillation in time equals the dimensionless wavelength of the waviness, i.e. 0.5, as was expected because of the pure rolling conditions presumed. Note that the amplitude decreases from  $X = -1.0$  to  $X = 1.0$ . This can also be observed from the figures 8.20 through 8.22. The same reasoning as presented in the previous section applies, i.e. in spite of the high viscosities the pressure gradients slowly flatten the surface feature during its motion through the contact region.



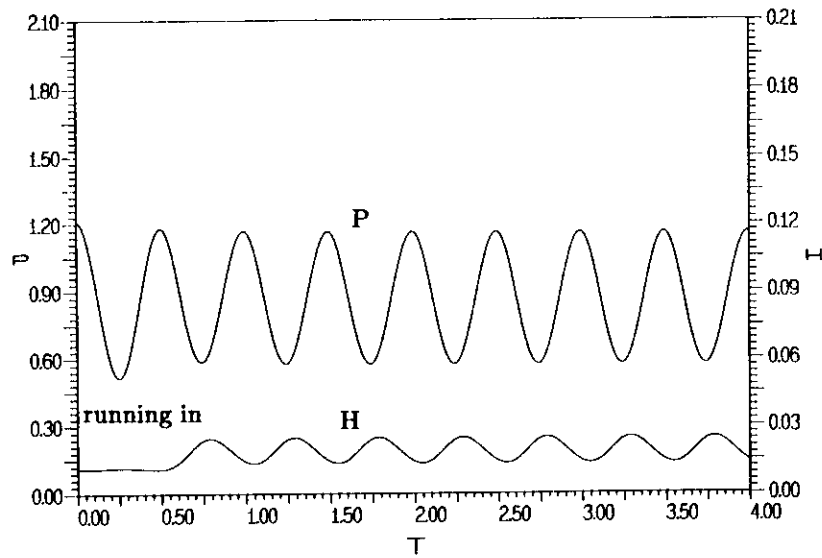


FIGURE 8.22:  $M = 100$ ,  $L = 11.08$ . Overrolling of waviness. Pressure and film thickness at  $X = -0.5$  as a function of time (pure rolling).

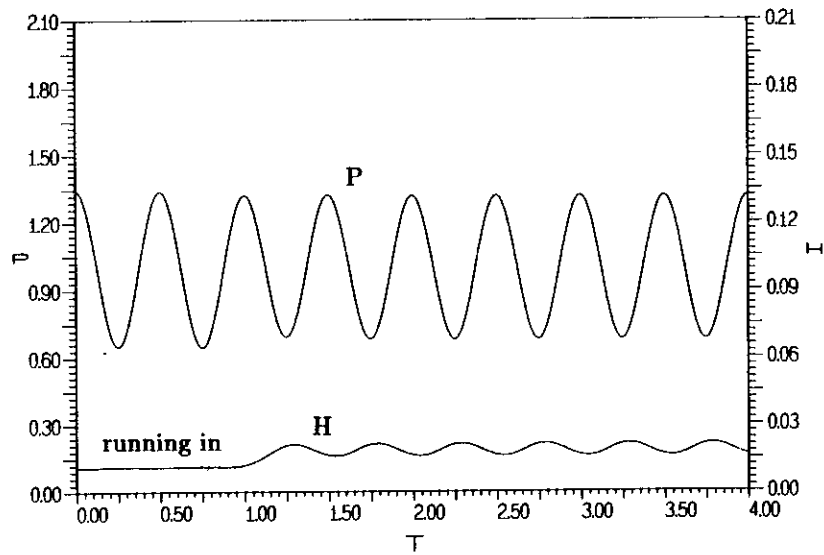


FIGURE 8.23:  $M = 100$ ,  $L = 11.08$ . Overrolling of waviness. Pressure and film thickness at  $X = 0.0$  as a function of time (pure rolling).

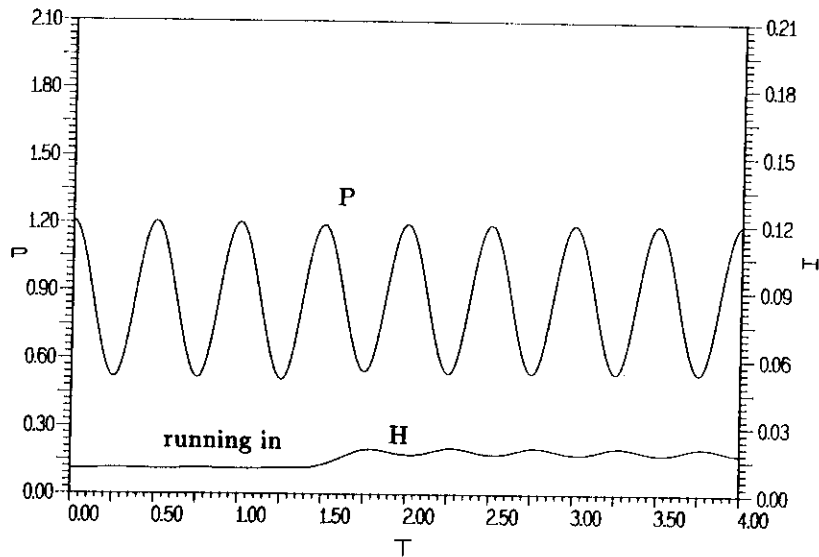


FIGURE 8.24:  $M = 100$ ,  $L = 11.08$ . Overrolling of waviness. Pressure and film thickness at  $X = 0.5$  as a function of time (pure rolling).

## 8.6 Additional insights: sliding

So far the simulations were restricted to pure rolling conditions. The presented results demonstrate that there are significant differences between the effects of a feature on the pressure profile and film shape predicted by a stationary and transient analysis respectively.

This seems to apply even more when both surfaces are not moving at the same velocity, i.e. in the case of slip, as can already be seen from the transient Reynolds' equation. The reader is reminded that in the Hertzian contact region where the coefficients ( $\epsilon$ ) are small this equation in the transient situation reduces to:

$$-\frac{\partial(\bar{\rho}H)}{\partial X} - \frac{\partial(\bar{\rho}H)}{\partial T} \approx 0 \quad (8.14)$$

As mentioned before this equation prescribes that  $H \approx H(X - T)$ , or in other words, it states that the change in the film thickness  $H$  caused by a feature travels with the dimensionless speed of unity (in terms of real velocities it moves with the average velocity:  $0.5 u_s$ ) in the region of small coefficients (the Hertzian dry contact region). This is a very interesting observation since the equation does not state anything about the velocity of the feature itself. Hence, the velocity with which the change in film thickness induced by the feature is propagated through the Hertzian contact region is independent of the velocity of the feature itself.

For example, consider the overrolling of a bump or an indentation. If the velocity of the bumped or indented surface is smaller than the average velocity of the surfaces

the change of the film thickness initiated upon entering the high pressure zone will move faster through this region than the feature itself. Hence, monitoring the film thickness in a fixed location in the Hertzian contact region as a function of time it will already deviate from the steady state value before the bump or dent is even at this location. A similar observation applies to the situation in which the surface with the feature moves at a higher speed than the average velocity. The change in the film thickness caused by the feature moves at a smaller speed through the region of small coefficients than the feature itself, i.e. when the feature leaves the contact region the change in the film thickness it induced is still present. The situation for the overrolling of waviness in the case of slip occurs is even more involved. In that case the pressure profile and film shape will be periodic in time, however, for points in the Hertzian contact region the pressure will be periodic with a wavelength determined by the velocity of the waviness itself ( $P \approx P(X - 2 u_2/u_s T)$ ), and the film thickness with a dimensionless wavelength that fits an integer number of times in one ( $H \approx H(X - T)$ ).

So far, only the equations are discussed. The next step is to demonstrate that indeed such an effect is observed when performing transient simulations under sliding conditions. This section presents the results of the “overrolling” of an indentation with a wavelength of  $0.5 \text{ mm}$  and an amplitude of  $2 \text{ }\mu\text{m}$  located at the surface moving with velocity  $u_2$ . The load conditions are the same as considered in the previous sections, i.e. the maximum Hertzian pressure equals  $2 \text{ GPa}$ . Two situations are considered. In the first situation the feature moves at a velocity smaller than the average velocity, and in the second situation it moves with a velocity exceeding the average velocity. To clearly visualize the aforementioned phenomenon, the first set of results applies to the situation in which  $u_2$  is significantly smaller than  $u_s$ :  $u_2/u_s = 0.25$ . Hence, at time  $T$  the position of the dent is given by:

$$X_d = X_s + 0.5 T \quad (8.15)$$

Similar to the results presented previously the simulation has been carried out using a uniform grid with 1409 nodes covering the domain  $-4 \leq X \leq 1.5$ . The simulation starts at  $T = 0$  with the dent at  $X_s = -4.5$  and a timestep  $\Delta t = 0.03125$  was used. At the start of the simulation the dent is entirely outside the domain. Consequently the solution equals the steady-state smooth surface solution, e.g. see figure 8.1.

Figure 8.25 to 8.31 present the (dimensionless) pressure profiles and film shapes at the times the dent is at different positions in the contact region.

In the situation presented in figure 8.25 the dent has just entered the Hertzian region. This figure displays the initiation of the phenomenon described above. The top of the dent can be recognized clearly at  $X = -0.75$ . The dent itself is symmetric with respect to  $X_d$ , yet, the change in the film thickness caused by the indentation is not. Figure 8.26 shows the dimensionless pressure profile and film thickness when the dent is at  $X_d = -0.5$ , that is 16 time steps later. Now the effect is even more pronounced. In the time it took the dent to move from  $X_d = -0.75$  to  $X_d = -0.5$  the

disturbance in film thickness has travelled a larger distance whereas the disturbance in the pressure profile is fixed at the dent position.

The following figures show how the asynchronism between the pressure variation and the film thickness variation evolves in time. Gradually the dent itself flattens in a way similar to the stationary results presented in chapter 7. See for example figure 8.28. This figure displays the pressure and film thickness at the time the dent is exactly in the center of the contact. At this time the change in the film thickness the dent created upon its entrance already leaves the Hertzian contact region. The difference with the pressure and film shape for pure rolling conditions when the change in film thickness, the dent and the change in pressure travel at the same speed are obvious, e.g. see figure 8.3. In the case of pure rolling conditions a film thickness change observed at a certain time, can be directly coupled to the passage of the dent. It is clear that in the case of sliding, contrary to the pure rolling situation, changes in the film thickness are observed that can not directly be coupled with the dent itself anymore. In fact, they appear to lead a life of their own.

The following figures 8.29 to 8.31 show how the initial change in the film thickness caused by the dent moves out of the Hertzian contact region before the dent does and how the dent itself gradually regains its original shape upon leaving the Hertzian contact dry region itself.

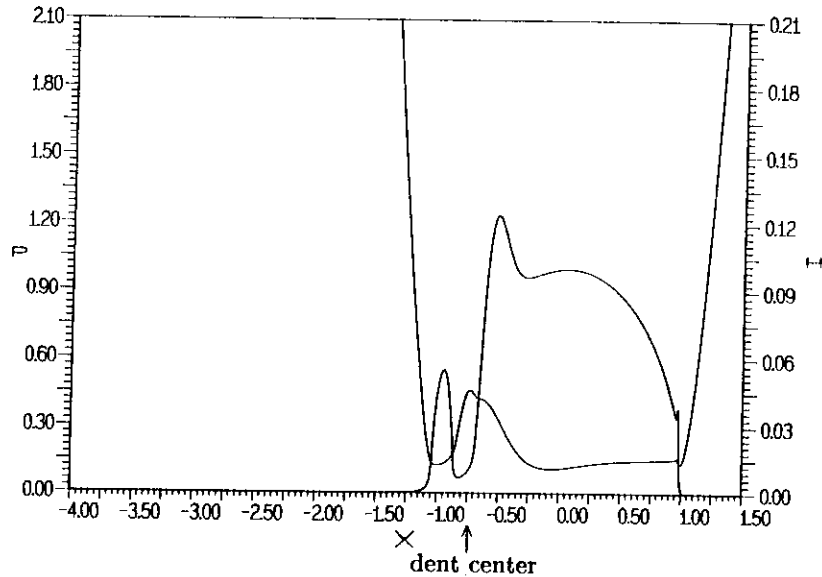


FIGURE 8.25: Solution  $M = 100$ ,  $L = 11.08$ . Pressure and film thickness as a function of  $X$ . Transient solution in the case of slip when the dent is at  $X_d = -0.75$  ( $u_2/u_s = 0.25$ ).

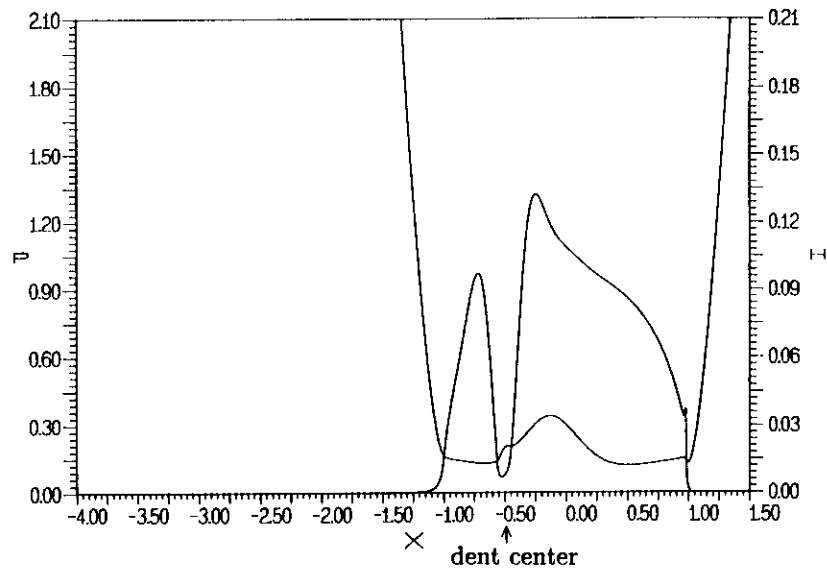


FIGURE 8.26: Solution  $M = 100$ ,  $L = 11.08$ . Pressure and film thickness as a function of  $X$ . Transient solution in the case of slip when the dent is at  $X_d = -0.5$  ( $u_2/u_s = 0.25$ ).

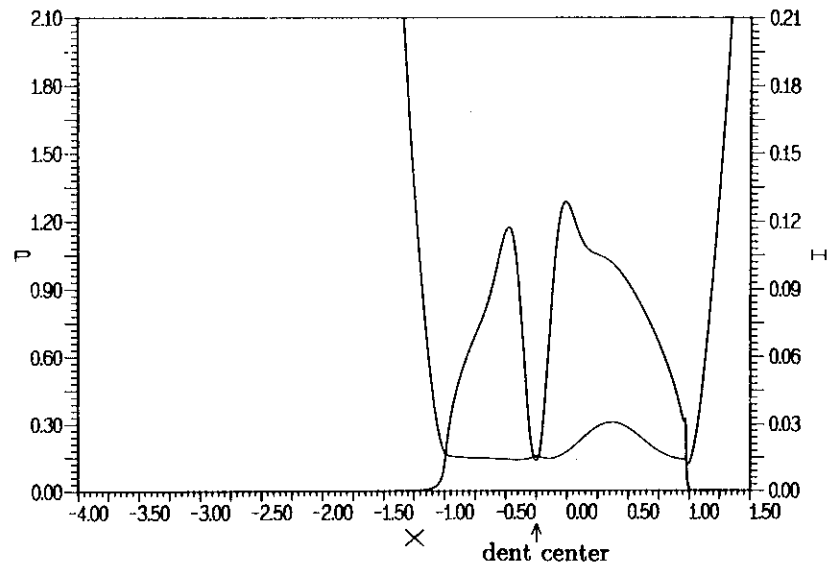


FIGURE 8.27: Solution  $M = 100$ ,  $L = 11.08$ . Pressure and film thickness as a function of  $X$ . Transient solution in the case of slip when the dent is at  $X_d = -0.25$  ( $u_2/u_s = 0.25$ ).

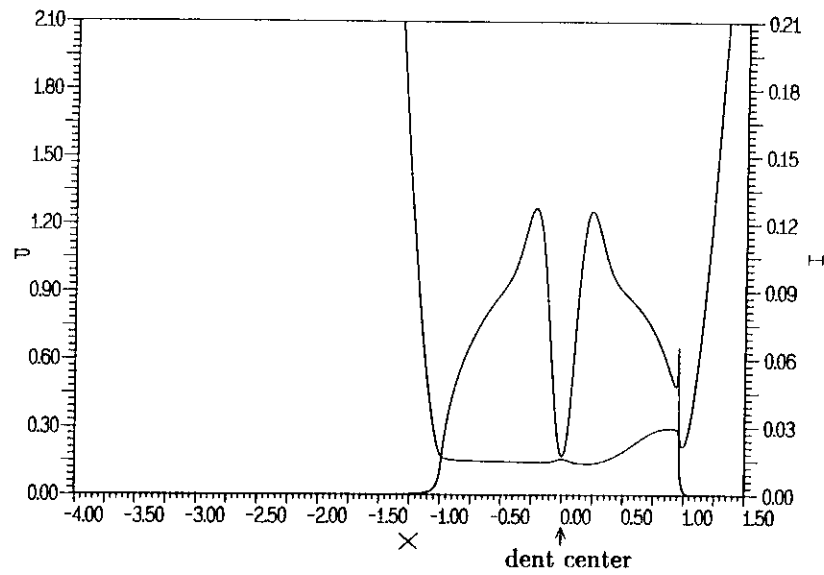


FIGURE 8.28: Solution  $M = 100$ ,  $L = 11.08$ . Pressure and film thickness as a function of  $X$ . Transient solution in the case of slip when the dent is at  $X_d = 0.0$  ( $u_2/u_s = 0.25$ ).

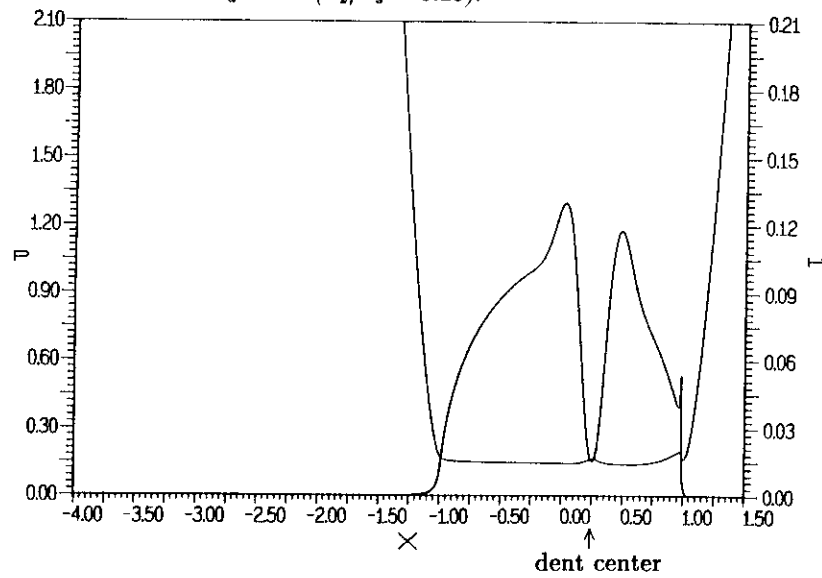


FIGURE 8.29: Solution  $M = 100$ ,  $L = 11.08$ . Pressure and film thickness as a function of  $X$ . Transient solution in the case of slip when the dent is at  $X_d = 0.25$  ( $u_2/u_s = 0.25$ ).

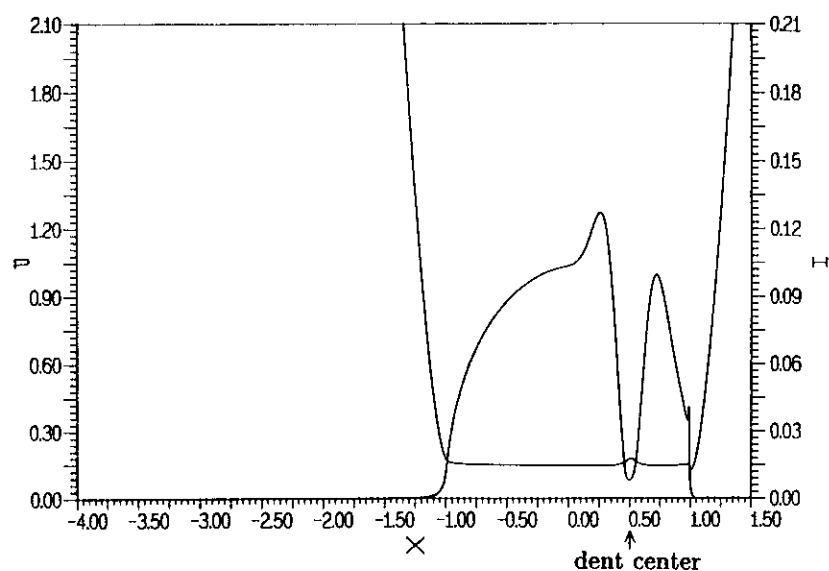


FIGURE 8.30: Solution  $M = 100$ ,  $L = 11.08$ . Pressure and film thickness as a function of  $X$ . Transient solution in the case of slip when the dent is at  $X_d = 0.5$  ( $u_2/u_s = 0.25$ ).

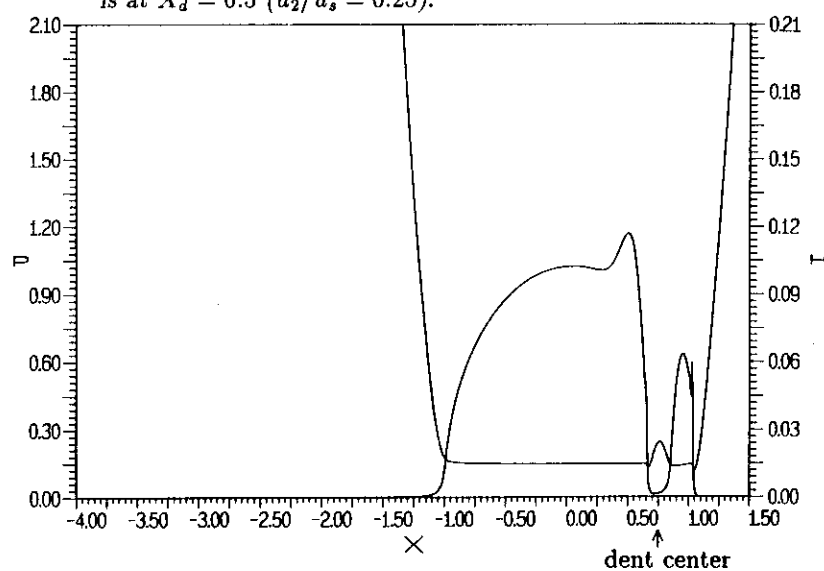


FIGURE 8.31: Solution  $M = 100$ ,  $L = 11.08$ . Pressure and film thickness as a function of  $X$ . Transient solution in the case of slip when the dent is at  $X_d = 0.75$  ( $u_2/u_s = 0.25$ ).

In addition to the pressure profiles and film shapes as a function of the spatial coordinate  $X$  the pressure and film thickness as a function of time have been monitored for different positions in the Hertzian contact region. Figure 8.32 through 8.34 present pressure and film thickness as a function of time for  $X = -0.5$ ,  $X = 0.0$ , and  $X = 0.5$ . All three figures clearly illustrate the asynchronism between the pressure variation and the film thickness variation. A change in the film thickness is already observed before the dent reaches the respective position. Since the dent induced change in the film thickness moves faster than the dent itself, the time interval between the moment the film thickness variation is observed at the specific location and the moment the pressure variation is experienced increases when going further downstream.

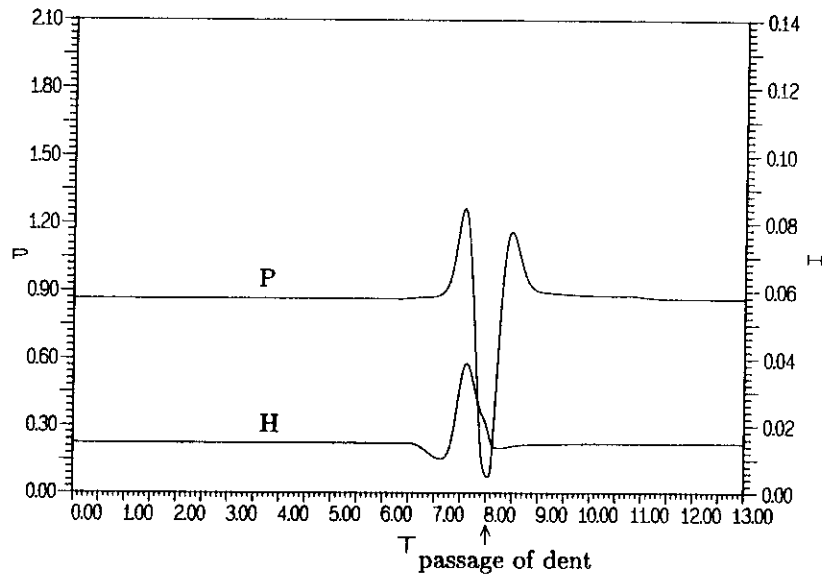


FIGURE 8.32:  $M = 100$ ,  $L = 11.08$ . Overrolling of an indentation. Pressure and film thickness as a function of  $T$  at  $X = -0.5$  in the case of the overrolling of a dent ( $u_2/u_s = 0.25$ ).



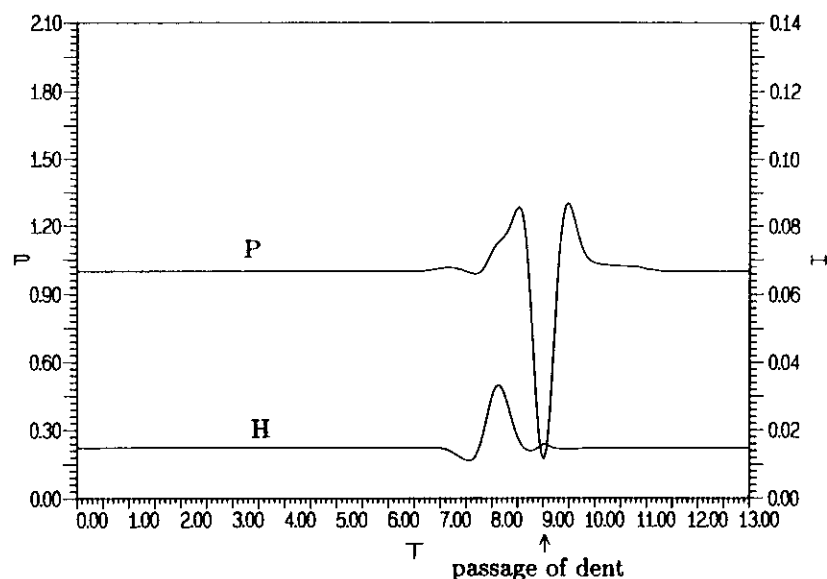


FIGURE 8.33:  $M = 100$ ,  $L = 11.08$ . Overrolling of an indentation. Pressure and film thickness as a function of  $T$  at the center of the contact, i.e.  $X = 0.0$  ( $u_2/u_s = 0.25$ ).

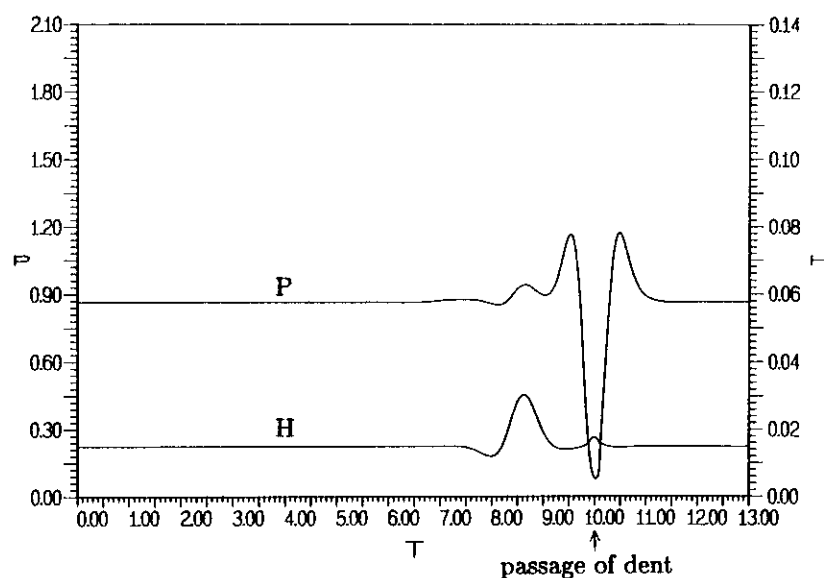


FIGURE 8.34:  $M = 100$ ,  $L = 11.08$ . Overrolling of an indentation. Pressure and film thickness as a function of  $T$  at  $X = 0.5$  ( $u_2/u_s = 0.25$ ).

Secondly the reverse situation is considered, i.e. the same load conditions, the same calculational domain and number of nodes, and the same dent, however, the indented surface moves with a velocity that is larger than the average velocity. In particular, the presented results apply to  $u_2/u_s = 0.75$ . Consequently, the position of the dent at time  $T$  is given by:

$$X_d = X_s + 1.5 T \quad (8.16)$$

At  $T = 0$  the dent is located outside the calculational domain, i.e. at  $X_s = -4.5$ . The simulation has been carried out using 288 time steps with  $\Delta t = 0.015625$ . This different time step was chosen for reasons of comparison, i.e. to ensure that the dent is at  $X = 0.0$  (the center) at an exact timestep. Based on the equation and the results presented earlier it is expected that in this case the dent travels faster through the Hertzian contact region than the change in the film thickness it initiated upon entering this region. Hence, at a specific location well in the Hertzian contact region the largest changes in the film thickness will only be observed at a time when the dent has already passed this position. This effect is indeed observed in the simulation results which are presented in the figures 8.35 to 8.43.

Figure 8.35 displays the (dimensionless) pressure and film shape as a function of  $X$  shortly after the dent has entered the Hertzian dry contact region, i.e. when  $X_d = -0.75$ . From the figure it is clear that already at this time the asynchronism between the passage of the dent with the pressure drop and the film thickness change exists. Figure 8.36 shows the solution 16 time steps later. Since the dent moves faster than the average velocity the dent itself has traveled a larger distance in the associated time interval than the change in the film thickness it caused upon its entrance in the Hertzian dry contact region.

Figure 8.37 displays the film thickness and pressure as a function of  $X$  at the time the dent reaches the center of the contact. The asynchronism between the pressure and film thickness change is fully developed and contrary to the results for pure rolling conditions with the dent at this position, the dent itself hardly shows up in the film thickness.

The following figures 8.38 to 8.40 show how the process proceeds in time. In particular figure 8.39 displays how the dent gradually regains its original shape when moving through the region of decreasing pressure. Finally figure 8.40 shows the pressure profile and film thickness as a function of  $X$  when the dent leaves the Hertzian contact region, i.e. when  $X_d = 1.0$ . Although the indentation has almost left, the changes in the film shape it induced upon its entrance of the Hertzian contact region are still present in the solution.

To demonstrate this effect more clearly the film thickness and pressure are monitored as a function of time at a specific location in the Hertzian contact region. Figure 8.41 to 8.43 show the pressure and film thickness as a function of time for  $X = -0.5$ ,  $X = 0.0$ , and  $X = 0.5$ . These figures again clearly show the asynchronism between the passage of the dent (the pressure drop) and the film thickness change.

In particular, downstream in the contact region the changes in the film thickness caused by the indentation are observed only when the dent has already left the contact region.

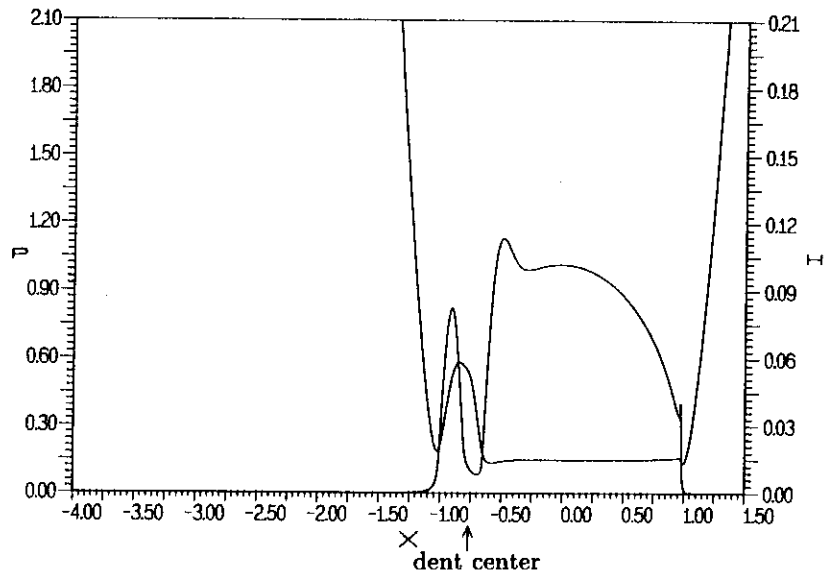


FIGURE 8.35: Solution  $M = 100$ ,  $L = 11.08$ . Pressure and film thickness as a function of  $X$ . Transient solution in the case of slip when the dent is at  $X_d = -0.75$  ( $u_2/u_s = 0.75$ ).

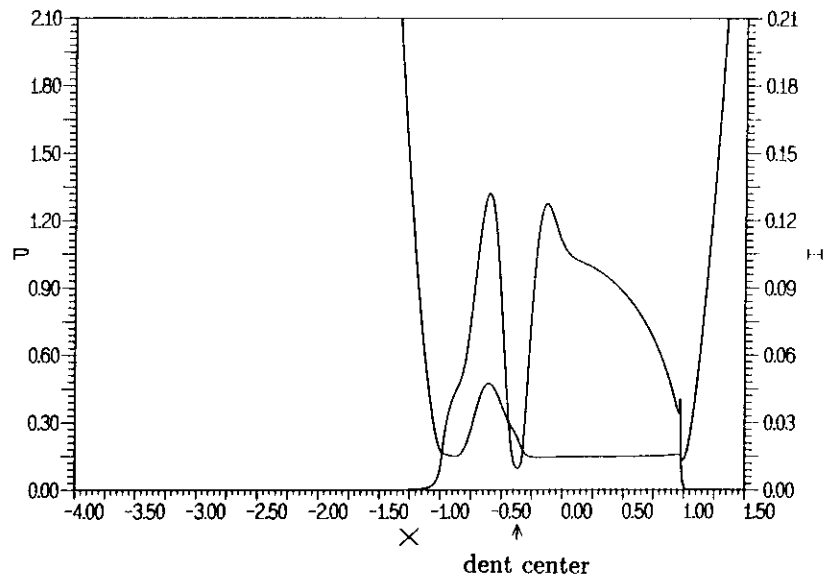


FIGURE 8.36: Solution  $M = 100$ ,  $L = 11.08$ . Pressure and film thickness as a function of  $X$ . Transient solution in the case of slip when the dent is at  $X_d = -0.375$  ( $u_2/u_s = 0.75$ ).

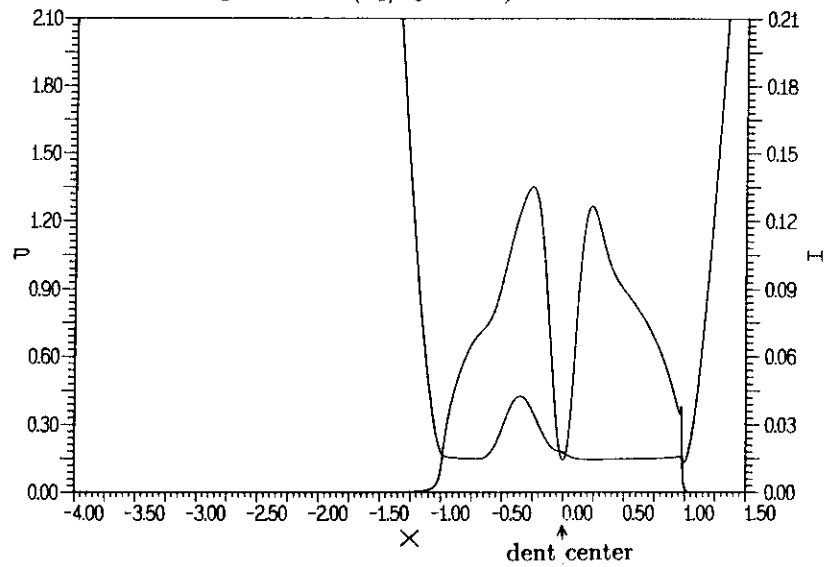


FIGURE 8.37: Solution  $M = 100$ ,  $L = 11.08$ . Pressure and film thickness as a function of  $X$ . Transient solution in the case of slip when the dent is at  $X_d = 0.0$  ( $u_2/u_s = 0.75$ ).

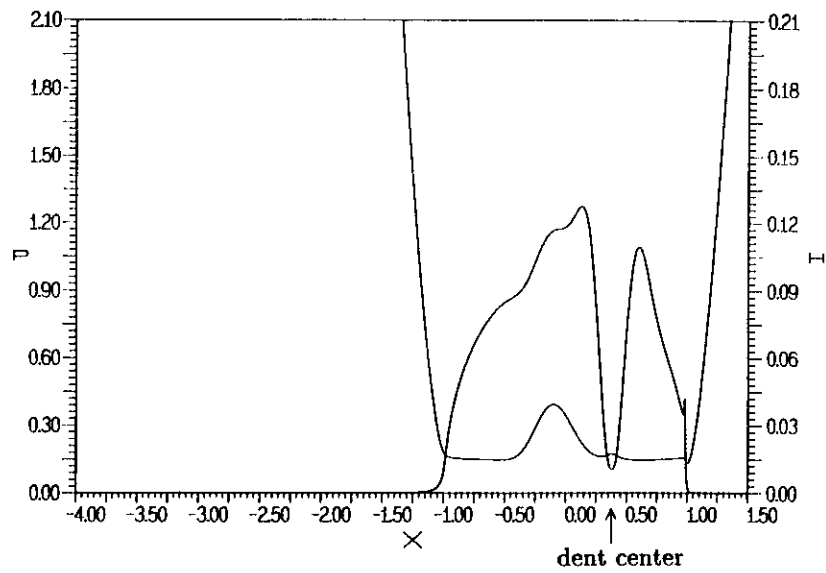


FIGURE 8.38: Solution  $M = 100$ ,  $L = 11.08$ . Pressure and film thickness as a function of  $X$ . Transient solution in the case of slip when the dent is at  $X_d = 0.375$  ( $u_2/u_s = 0.75$ ).

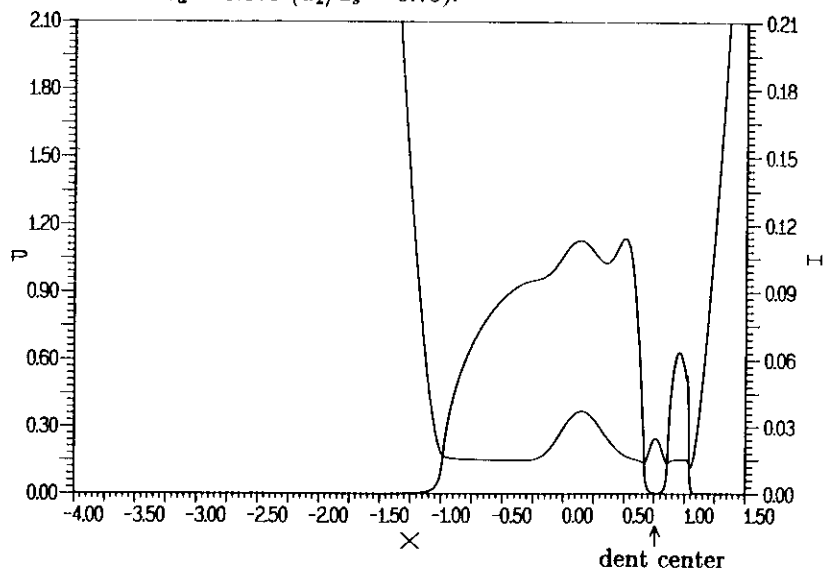


FIGURE 8.39: Solution  $M = 100$ ,  $L = 11.08$ . Pressure and film thickness as a function of  $X$ . Transient solution in the case of slip when the dent is at  $X_d = 0.75$  ( $u_2/u_s = 0.75$ ).

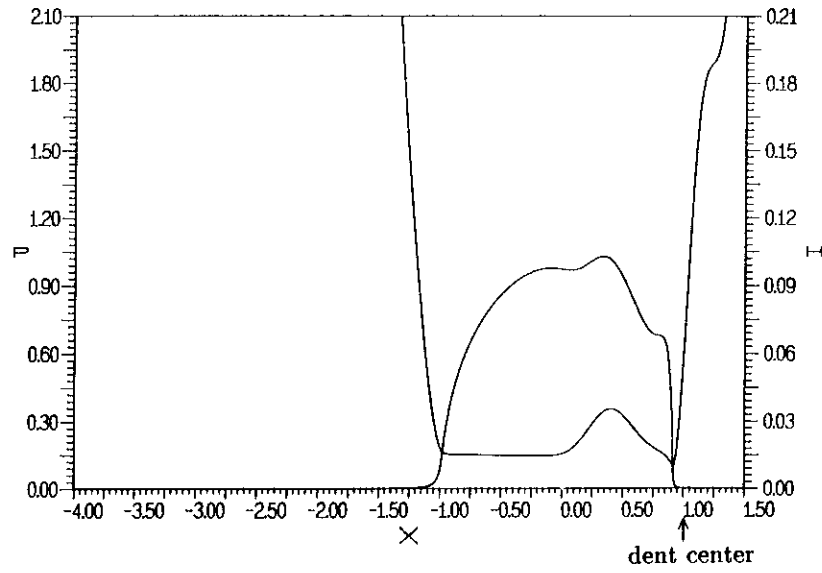


FIGURE 8.40: Solution  $M = 100$ ,  $L = 11.08$ . Pressure and film thickness as a function of  $X$ . Transient solution in the case of slip when the dent is at  $X_d = 1.0$  ( $u_2/u_s = 0.75$ ).

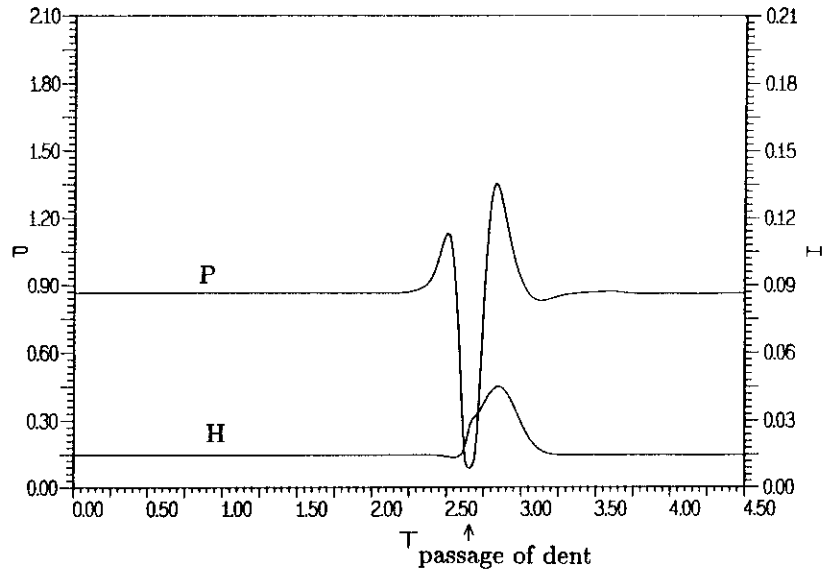


FIGURE 8.41:  $M = 100$ ,  $L = 11.08$ . Overrolling of an indentation. Pressure and film thickness at  $X = -0.5$  as a function of  $T$  ( $u_2/u_s = 0.75$ ).

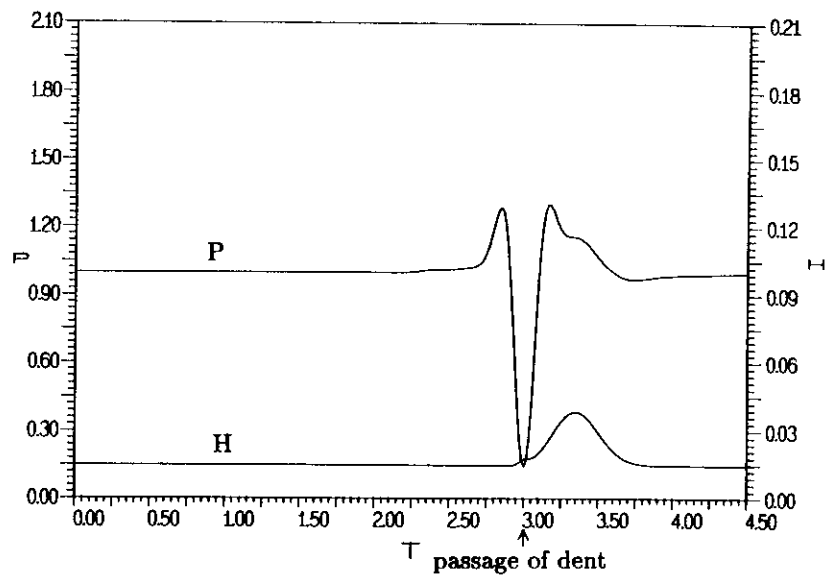


FIGURE 8.42:  $M = 100$ ,  $L = 11.08$ . Overrolling of an indentation. Pressure and film thickness at the center of the contact, i.e.  $X = 0.0$ , as a function of  $T$  ( $u_2/u_s = 0.75$ ).

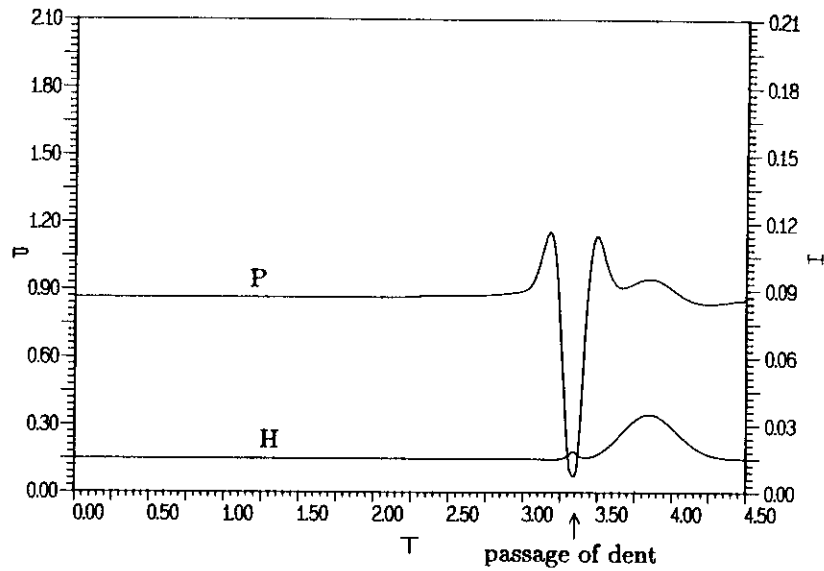


FIGURE 8.43:  $M = 100$ ,  $L = 11.08$ . Overrolling of an indentation. Pressure and film thickness at  $X = 0.5$  as a function of  $T$  ( $u_2/u_s = 0.75$ ).

This section started with the observation that, according to the Reynolds equation, in moderately to highly loaded EHL contacts the changes in the film thickness caused by a feature located on a non-stationary surface travel through the Hertzian contact region with the average velocity of the surfaces. This is obvious for pure rolling conditions but it also applies in the case of slip. Consequently, if slip occurs, from the moment the feature enters the Hertzian contact region changes in the film thickness occur that travel faster or slower than the feature itself. Hence, monitoring the film thickness as a function of time at a certain location the major change in the film thickness already occurs before the dent reaches the location or it occurs when the dent has already passed. The occurrence of this phenomenon was predicted by simply looking at the equations and confirmed by simulation results for the overrolling of a dent in the case of sliding. To the authors knowledge this phenomenon has not been reported before in theoretical studies. Since it is obviously predicted by the equations, the most likely explanation that it has not been observed before is that most transient simulations presented so far have been carried out using only a few (relatively large) timesteps.

According to the author the “physical” explanation of the feature lies in the conservation of mass in the flow. For example, consider the overrolling of a dent. Upon entering the Hertzian contact region the total amount of mass in the film in this region increases because of the fluid trapped in the dent. At moderately to highly loaded conditions, pressure driven flow is virtually impossible in this region. In the model considered here the reason is the high viscosity of the fluid. However, the same reasoning applies in case of solidification of the fluid at high pressures. The only alternative for the system to move the trapped fluid through the Hertzian contact region without violating mass conservation is by means of Couette flow and thus by means of film thickness variations.

The phenomenon has some serious consequences. It shows that, when studying the effect of a surface feature in EHL using a model based on the Reynolds equation it is wrong to assume that, if the feature is located on the surface moving with a non-zero velocity  $u_2$ , the film thickness will be a function of  $x - u_2 t$ . This assumption is obviously true in case of negligible elastic deformation, i.e. in the case of hydrodynamic lubrication. However, in EHL contact situations it only applies in the case of pure rolling or for very lightly loaded situations.

Over the last few decades several researchers have presented modified Reynolds equations to account for the effect of surface features. These equations are quite often based on the aforementioned assumption. Consequently care should be taken when applying such Reynolds equations to EHL contact situations.

Having recognized the phenomenon in the model and confirmed its occurrence by means of numerical simulation it would be interesting to know if it really occurs in practical contact situations. At first sight one might argue that this is unlikely since it is caused by the high viscosities which are not realistic. However, as was explained above, the phenomenon can occur in any situation where pressure induced flow almost vanishes, for example in case of solidification of the lubricant. With

---



respect to roller bearing applications the chances of observing the effect are probably negligible. In these situations the slip is generally small whereas, due to short time a feature spends in the contact region, the phenomenon can only be recognized clearly if the slip is sufficiently large. However, it may be of importance in the contact between gear wheels.

It would be interesting to check the occurrence of the phenomenon by means of experiments. In particular, if it really occurs it should be possible to show it with an optical interferometry rig extended with a high speed camera. In this respect the experimental results presented by Kaneta and Cameron [K1] already a decade ago are interesting. Running an artificially "rough" steel ball against a smooth glass disk they observed a deformation of the asperities that strongly depended on the sliding conditions. In particular, in the case the "rough" ball was moving and the disk was stationary (pure sliding) they found micro constrictions at the inlet side of the asperities whereas the macro constriction occurred at the outlet. To the author's opinion this effect and the results presented in this chapter are related and further research along these lines is recommended.

## 8.7 Conclusion

For the same operating conditions, the overrolling of an indentation, a bump and a harmonic waviness were studied in detail for pure rolling conditions. The results of these simulations demonstrated that transient studies are of great importance and that extrapolation of results from steady state conditions to practical conditions where a surface feature is moving may be misleading.

This particularly applies to the film thickness results. It has been shown that contrary to the steady state situation where the geometry of a feature located in the contact region is significantly changed, in the transient situation assuming pure rolling, once the feature is in the Hertzian contact region it tends to move through while undergoing only small geometrical changes.

In addition, the results presented for the overrolling of an indentation led to an interesting conclusion for the fatigue life of the contact. From the results it was concluded that spalling will most likely occur near the trailing edge of the indentation. The very same preference that has been found in experiments.

The transient analysis also led to the discovery of an interesting property of the Reynolds equation when applied to EHL contact situations. It was shown that the changes in the film thickness caused by a moving surface feature travel through the Hertzian contact region at the average speed of the surfaces regardless of the speed at which the surface feature moves itself. This observation was confirmed by the results of transient simulations of the overrolling of an indentation in case of slip. The simulation results showed that the occurrence of a film thickness change at a certain point in time can no longer directly be associated with the passage of the feature. This phenomenon provides an interesting subject for future research.

## References

- [C1] Chang, L., Cusano, C., and Conry, T.F., 1989, "Effects of lubricant rheology and kinematic conditions on micro-elastohydrodynamic lubrication," *ASME JOT*, **111**, 344-351.
  - [D1] Dowson, D., and Higginson, G.R., 1966, "Elasto-hydrodynamic lubrication, the fundamentals of roller and gear lubrication," Pergamon Press, Oxford, Great Britain.
  - [H1] Hamrock, B.J., and Dowson, D., 1976, "Isothermal elastohydrodynamic lubrication of point contacts, Part 1 - theoretical formulation," *ASME JOLT*, **98**, 223-229.
  - [I1] Ioannides, E., Jacobson, B.O., Tripp, J.H., 1989, "Prediction of rolling bearing life under practical operating conditions," Proc. 15th Leeds-Lyon Symposium on Tribology, 181-187.
  - [K1] Kaneta, M., and Cameron, A., (1980), "Effects of asperities in elastohydrodynamic lubrication," *ASME JOT*, **102**, 374-389.
  - [L1] Lee, R.T., and Hamrock, B.J., 1989, "A circular non-Newtonian fluid model: Part 2 - Used in microelastohydrodynamic lubrication," Presented at the 1989 *ASME/STLE* tribology conference, Fort Lauderdale, U.S.A. *ASME* Preprint 89-Trib-10.
  - [L2] Lubrecht, A.A., Venner, C.H., Lane, S., Jacobson, B., Ioannides, E., 1990, "Surface damage - comparison of theoretical and experimental endurance lives of rolling bearings," *Proceedings of the 1990 Japan International Tribology Conference, Nagoya, Japan*, **1**, 185-190.
  - [O1] Oh, K.P., 1985, "The numerical solution of dynamically loaded elastohydrodynamic contact as a nonlinear complementarity problem," *ASME JOT*, **106**, 88-95.
  - [P1] Patir, N., and Cheng, H.S., 1978, "An average flow model for determining effects of three-dimensional roughness on partial hydrodynamic lubrication," *ASME JOT*, **100**, 12-17.
  - [T1] Tripp, J.H., and Hamrock, B.J., 1984, "Surface roughness effects in elastohydrodynamic contacts," Proc. 11th Leeds-Lyon Symposium on Tribology, 30-39.
  - [V1] Venner, C.H., Lubrecht, A.A., and ten Napel, W.E., (1990) "Numerical simulation of the overrolling of a surface feature in an EHL line contact," Submitted for publication in *ASME JOT*
  - [W1] Woods, C.M., and Brewe, D.E., 1989 "The solution of the Elrod algorithm for a dynamically loaded journal bearing using multigrid techniques," *ASME JOT*, **111**, 302-308.
-

## Chapter 9

# Circular contact results: Smooth surfaces

The previous chapter concluded the results for the EHL line contact problem presented in this thesis. Attention is now directed towards the EHL circular contact problem. To start with the present chapter considers the situation in which the surfaces are perfectly smooth. With the solver described in chapter 5 this problem has been solved for a wide range of load conditions.

First section 9.1 discusses the variation of the pressure profiles and the associated film shapes with the variation of the governing parameters. Subsequently, the dependence of some characteristic film thicknesses on the operating conditions is investigated in section 9.2. In particular, diagrams are presented giving the minimum, central and average film thickness as a function of the governing parameters. From the calculational results a formula has been derived that accurately predicts the central film thickness in the entire parameter range considered.

In section 5.6 it was already demonstrated that the algorithm is fast. The results presented in this chapter demonstrate that the algorithm also meets a second requirement necessary for the study of more complex (non-smooth) EHL circular contact situations, i.e. it is very stable. Results of the application of the algorithm to circular contacts with various types of surface features will be presented in chapter 10.

### 9.1 Solutions

First the variation of the pressure profile and film shape with variation of the operating conditions is investigated. The results are presented in terms of the Moes dimensionless point contact parameters  $M$  and  $L$ . However, for reasons of comparison also values of the Hamrock and Dowson dimensionless parameters, characterizing the operating conditions are given.

### 9.1.1 Computational details

The majority of the presented solutions have been calculated on a domain  $\{(X, Y) \in \mathbb{R}^2 | -4.5 \leq X \leq 1.5, -3 \leq Y \leq 3\}$ . However, in very lightly loaded situations a larger calculational domain has been used. For these conditions the use of a large domain is essential. If the domain, and particularly if the inlet region, is not chosen large enough it results in an underestimation of the film thickness. This numerical artefact is referred to as *numerical starvation*, e.g. see Lubrecht [L1] and also Kweh et al. [K1], and was also mentioned in the previous discussion of the line contact problem (chapter 6). In heavily loaded situations a smaller calculational domain than mentioned above is sufficient to obtain an accurate estimate of the film shape, e.g. the choice of  $X_a = -2.5$ ,  $X_b = 1.5$  and  $Y_a = 2$  serves well.

Because the surfaces are assumed to be perfectly smooth, the solutions for the pressure and the associated film shapes must be symmetric with respect to the centerline of the contact, i.e.  $Y = 0$ . Until now, similar point contact calculations have been performed using this symmetry condition as an explicit boundary condition. This enables solution of the problem on only half of the domain, e.g. see Lubrecht [L1]. For two reasons this approach has not been followed here. Firstly, using the full domain is much more convenient in the calculation of the elastic deformation integrals with the multilevel multi-integration algorithm. Secondly, and even more important, it would limit future studies of the effect of surface features to symmetric features only. Hence, the problem is solved on the entire domain with only the usual Dirichlet condition  $P = 0$  on all four boundaries of the calculational domain. Consequently, in the present application (calculational domain and undeformed gap symmetric with respect to  $Y$ ), the symmetry in the solutions should appear naturally.

### 9.1.2 Varying $M$ and $L$

Figure 9.1 presents the pressure profile for a relatively lightly loaded situation, i.e.  $M = 10$ , and  $L = 10$ . This corresponds with the following values of the Hamrock and Dowson parameters:  $W = 9.46 \cdot 10^{-8}$ ,  $G = 4729$ , and  $U = 1.0 \cdot 10^{-11}$ . The values of the parameters  $\bar{\alpha}$  and  $\lambda$  are:  $\bar{\alpha} = 7.85$ ,  $\lambda = 0.51$ . Hence, for  $\alpha = 1.7 \cdot 10^{-8}$  the maximum Hertzian pressure for this load situation is only 0.46 GPa. The solution has been calculated using 263.169 nodes on the entire domain which extended from  $X_a = -9$  to  $X_b = 3$  in  $X$  direction and from  $-Y_a$  to  $Y_a$  in  $Y$  direction with  $Y_a = 6$ . This large domain was used to avoid numerical starvation. The associated film shape is shown in figure 9.2. Note the reversion of the vertical axis.

In spite of the relatively small load the pressure profile and film shape show the characteristic EHL features. The pressure profile more or less approximates a semi-ellipsoid and because of the low load in the inlet region the pressure smoothly builds up to this semi-ellipsoid. Furthermore, the pressure shows the three dimensional equivalent of the pressure spike just ahead of the cavitated region. Obviously the

local nodal density in the spike region is not large enough to smoothly describe the pressure profile in the spike region and consequently it looks ragged. The film shape shows the well known horse-shoe shaped region where in this particular low load situation the side lobes where the minimum film thickness occurs are relatively large. This can be seen from figure 9.2 but is more obvious from a contour plot of the film thickness, e.g. see figure 9.13.

To demonstrate the effect of load, the pressure profiles and the associated film shapes for varying  $M$  and constant  $L$  ( $L = 10$ ) are displayed in the figures 9.3 to 9.8. In addition to this variation of  $M$  the figures 9.9 to 9.12 give the pressure profile and the associated film shapes for  $M = 200$  and  $L = 0$  and  $L = 5$  respectively. Together with figure 9.5 and figure 9.6 the latter figures display the characteristic changes in the solutions with increasing dependence of the viscosity on the pressure. Finally figure 9.13 gives the film thickness contour plots for a large number of parameter values among which also the situations presented in the figures 9.1 to 9.12.

The variations in the solutions with increasing  $M$  for a given  $L$  resemble the tendencies discussed for the line contact problem in section 6.1.2. With increasing  $M$  the pressure profile approximates the Hertzian dry contact pressure and both the region of pressure generation (inlet) and the pressure spike narrow down. In addition, as for the line contact problem, the minimum film thickness decreases with increasing  $M$  (load).

As mentioned above, characteristic for the present circular contact situation is the formation of the horseshoe shaped region in the film thickness, i.e. the sidelobes. These sidelobes can be recognized clearly in the film thickness plots and it can be seen that with increasing  $M$  they become smaller and “move” in the outward direction until for large  $M$  the minimum film thickness is practically found at the side of the contact i.e. at  $X = 0, Y = \pm 1$ . This can be seen from the colourgraphs but is even more obvious from the film thickness contour plots of figure 9.13. This decrease in size of the sidelobes forms a complication with respect to the accurate estimation of the minimum film thickness. With increasing  $M$  this becomes more difficult, as was also observed by Kweh et al. [K1] and will be discussed in more detail in section 9.2.2.

Comparison of the pressure profiles presented in the figures 9.9, 9.11 and 9.5 shows the influence of the fact that the viscosity increases with pressure on the solution. Firstly, the pressure profile for the isoviscous situation ( $L = 0$ ) presented in figure 9.9 clearly has no spike and for this relatively high value of  $M$  it closely approximates the Hertzian semi-ellipsoid pressure profile. With increasing  $L$  a pressure spike gradually develops in an analogous way as was found for the line contact problem (see section 6.1.3). This is illustrated by figure 9.11 ( $L = 5$ ) where “the spike” shows up in the form of a ridge that can still be described accurately with the present nodal density  $((n_x + 1) \times (n_y + 1) = 263.169)$ . Subsequently, with increasing  $L$  its height increases and it becomes a real “spike”, e.g. see figure 9.5 where it is fully developed. In that situation the local nodal density of the applied grid is obviously not large enough to describe the spike accurately and consequently it looks ragged.

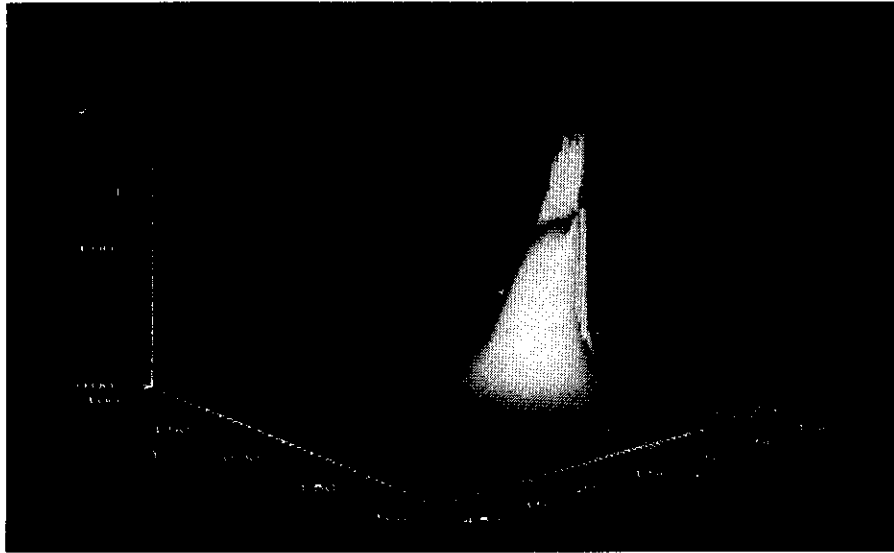


FIGURE 9.1: Pressure profile  $M = 10$  and  $L = 10$  ( $W = 9.46 \cdot 10^{-8}$ ,  $U = 1.0 \cdot 10^{-11}$ ,  $G = 4729$ , or  $\bar{\alpha} = 7.85$  and  $\lambda = 0.51$ ).

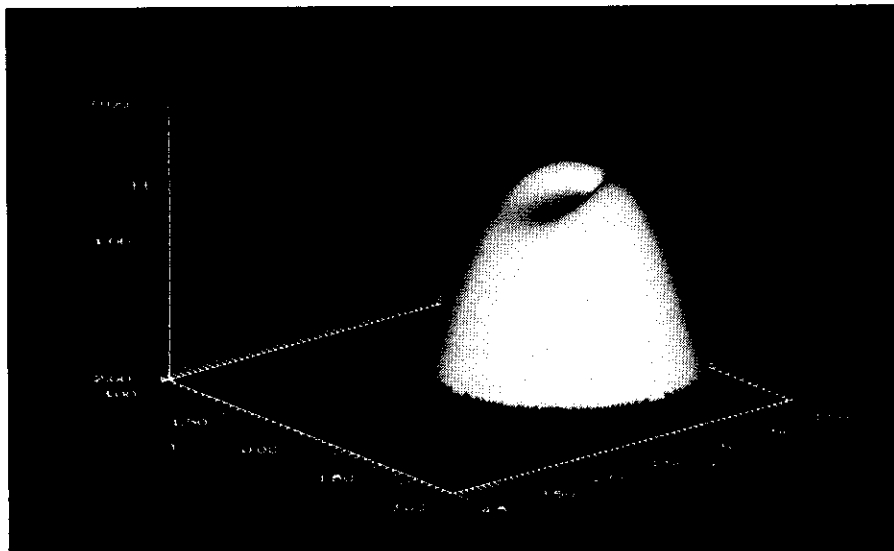


FIGURE 9.2: Film shape associated with the pressure profile presented in figure 9.1, i.e.  $M = 10$  and  $L = 10$ .



FIGURE 9.3: *Pressure profile*  $M = 50$  and  $L = 10$  ( $W = 4.73 \cdot 10^{-7}$ ,  $U = 1.0 \cdot 10^{-11}$ ,  $G = 4729$ , or  $\bar{\alpha} = 13.42$  and  $\lambda = 5.96 \cdot 10^{-2}$ ).

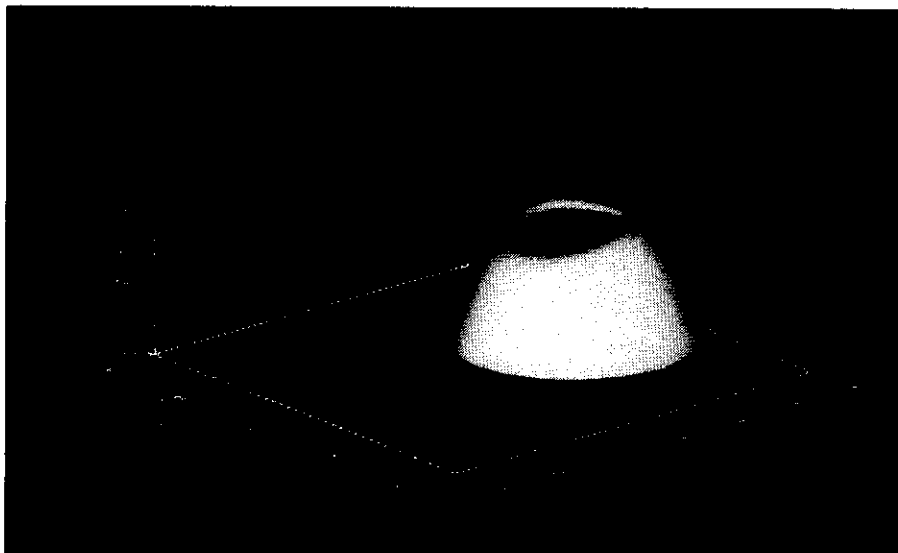


FIGURE 9.4: *Film shape associated with the pressure profile presented in figure 9.3, i.e.  $M = 50$  and  $L = 10$ .*

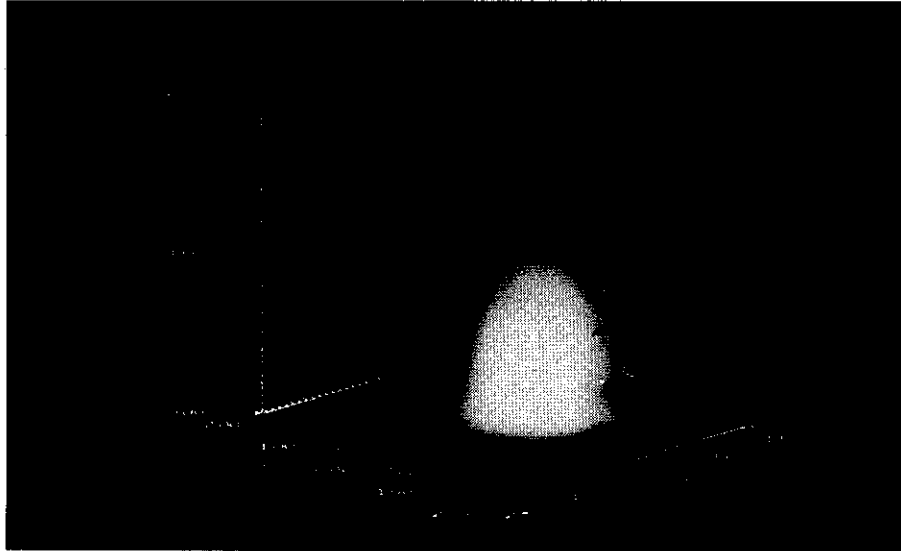


FIGURE 9.5: Pressure profile  $M = 200$  and  $L = 10$  ( $W = 1.89 \cdot 10^{-6}$ ,  $U = 1.0 \cdot 10^{-11}$ ,  $G = 4729$ , or  $\bar{\alpha} = 21.31$  and  $\lambda = 9.39 \cdot 10^{-3}$ ).

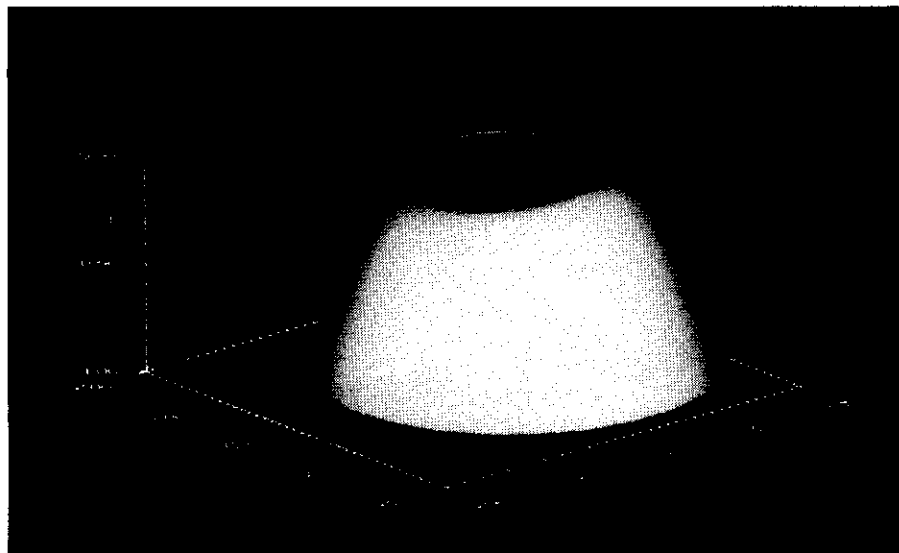


FIGURE 9.6: Film shape associated with the pressure profile presented in figure 9.5, i.e.  $M = 200$  and  $L = 10$ .



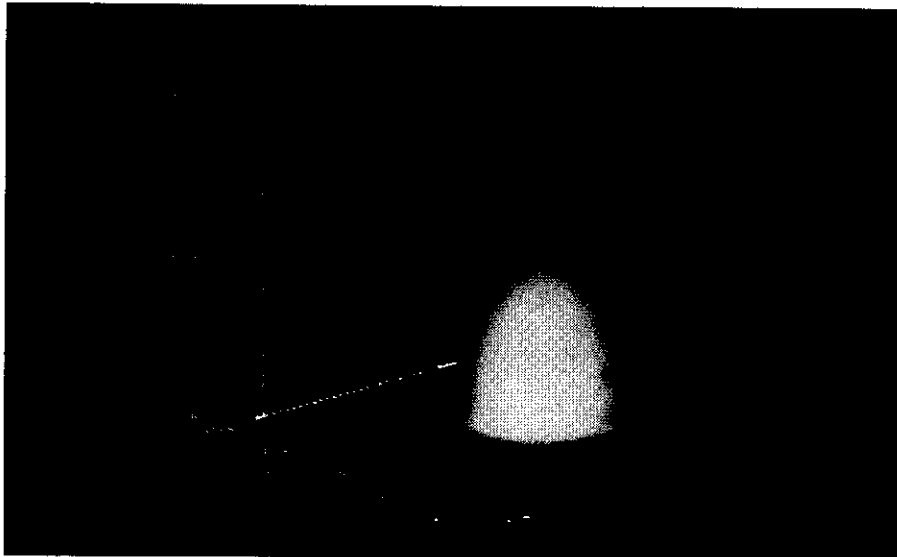


FIGURE 9.7: Pressure profile  $M = 1000$  and  $L = 10$  ( $W = 9.46 \cdot 10^{-6}$ ,  $U = 1.0 \cdot 10^{-11}$ ,  $G = 4729$ , or  $\bar{\alpha} = 36.44$  and  $\lambda = 1.10 \cdot 10^{-3}$ ).

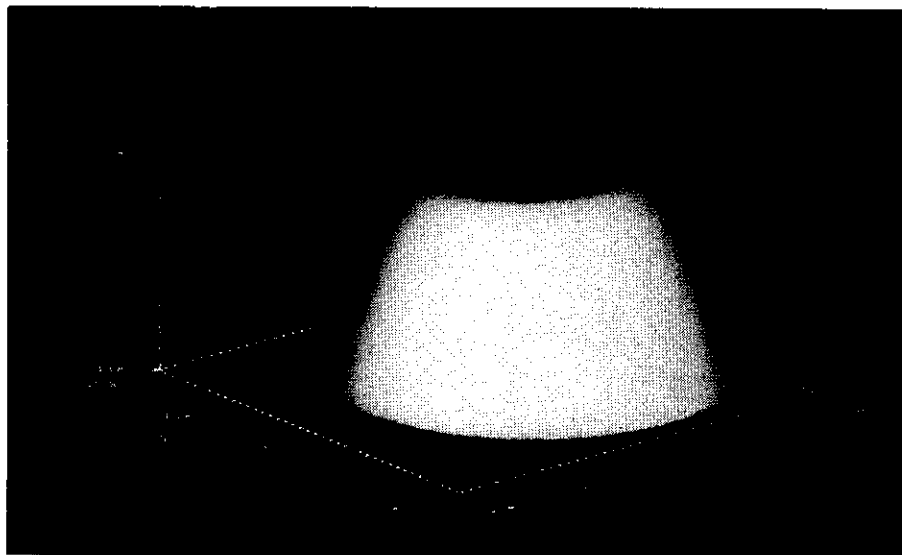


FIGURE 9.8: Film shape associated with the pressure profile presented in figure 9.7., i.e.  $M = 1000$  and  $L = 10$ .

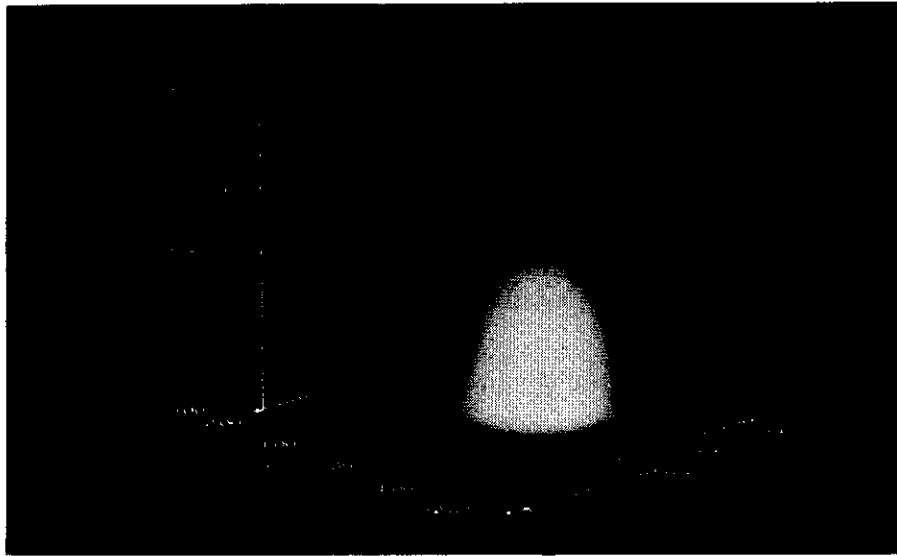


FIGURE 9.9: Pressure profile  $M = 200$  and  $L = 0$  ( $W = 1.89 \cdot 10^{-6}$ ,  $U = 1.0 \cdot 10^{-11}$ ,  $G = 0$ , or  $\bar{\alpha} = 0$  and  $\lambda = 9.39 \cdot 10^{-3}$ ).

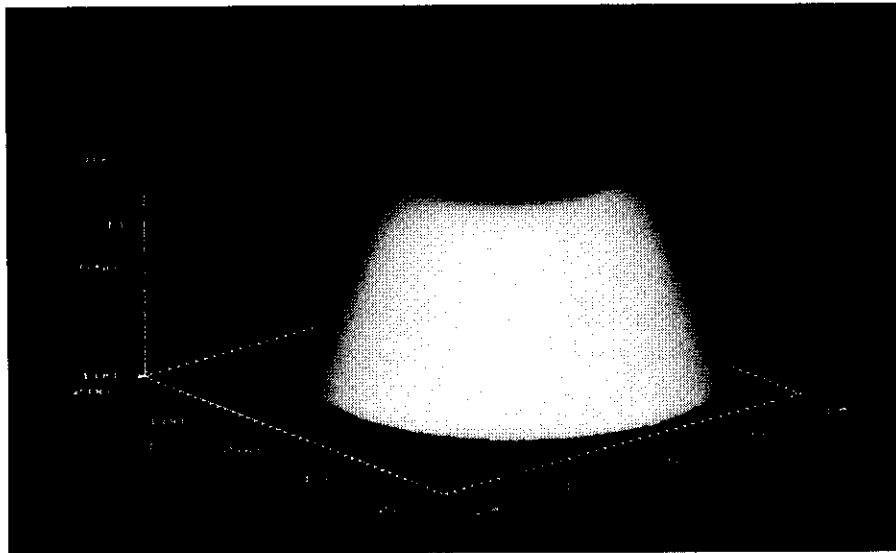


FIGURE 9.10: Film shape associated with the pressure profile presented in figure 9.9, i.e.  $M = 200$  and  $L = 0$ .

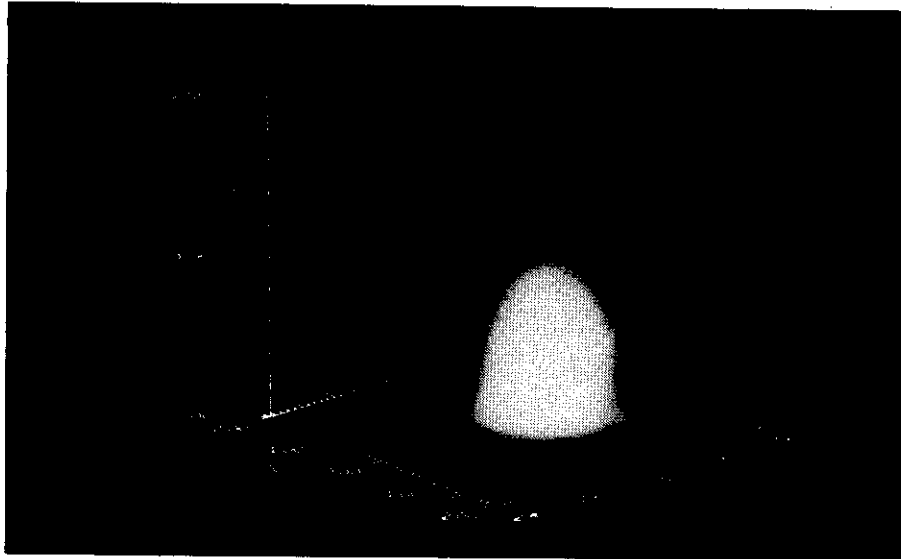


FIGURE 9.11: Pressure profile  $M = 200$  and  $L = 5$  ( $W = 1.89 \cdot 10^{-6}$ ,  $U = 1.0 \cdot 10^{-11}$ ,  $G = 2364$ , or  $\bar{\alpha} = 10.65$  and  $\lambda = 9.39 \cdot 10^{-3}$ ).

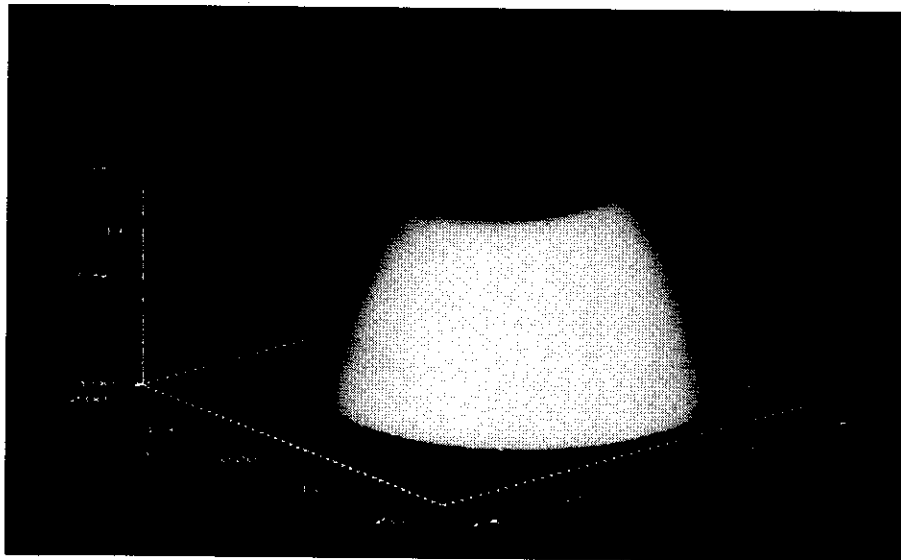


FIGURE 9.12: Film shape associated with the pressure profile presented in figure 9.11, i.e.  $M = 200$  and  $L = 5$ .

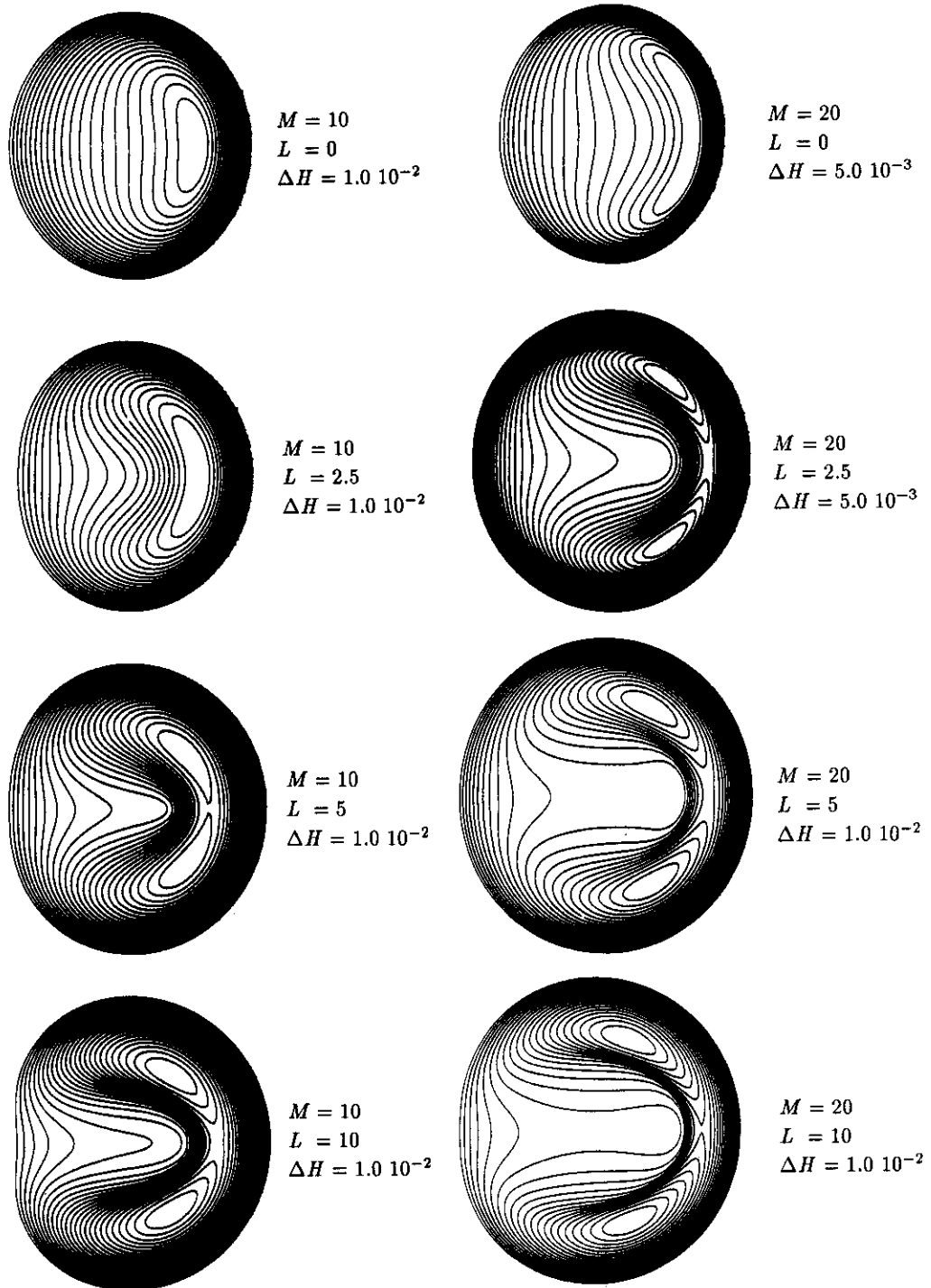
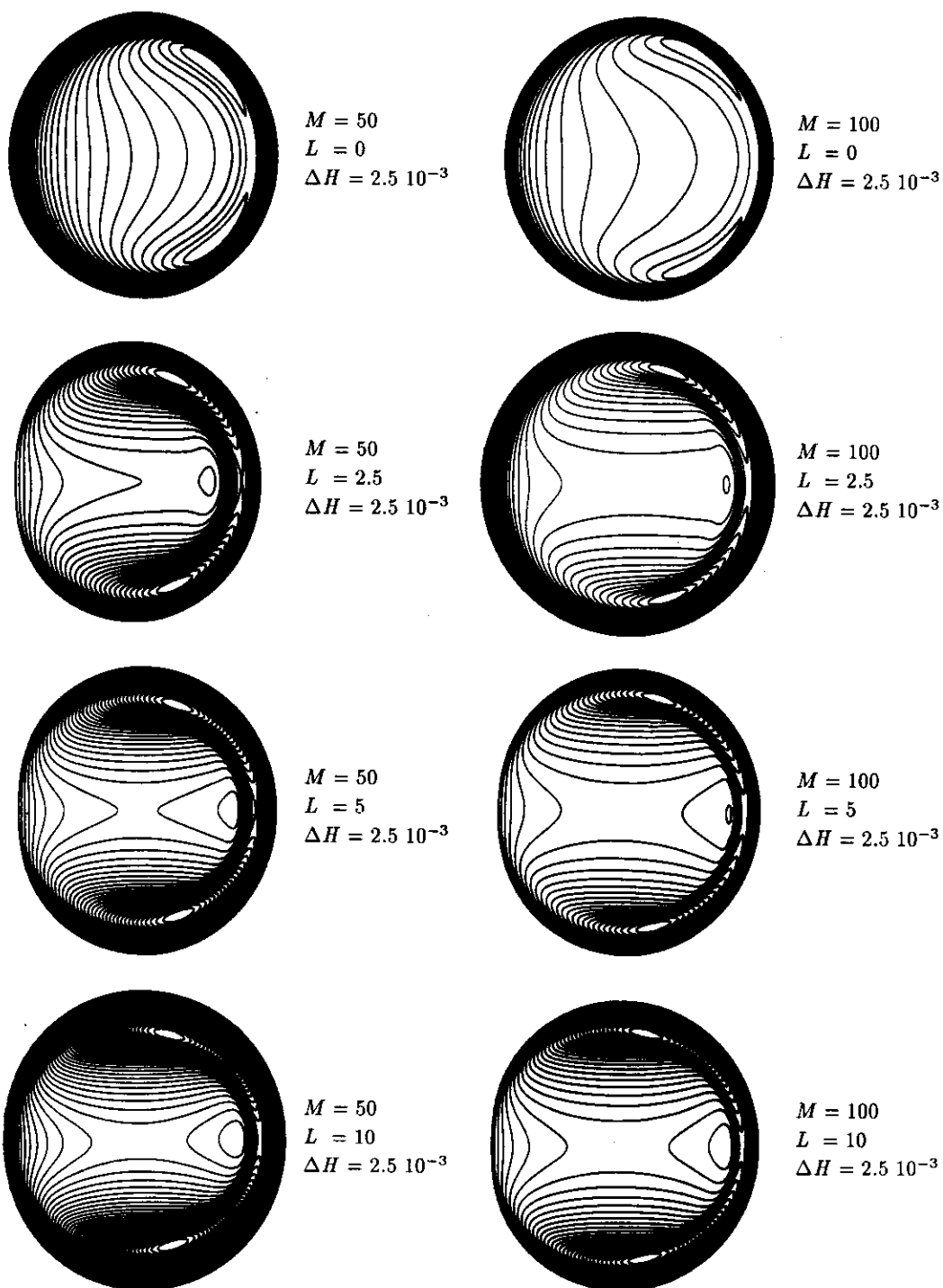
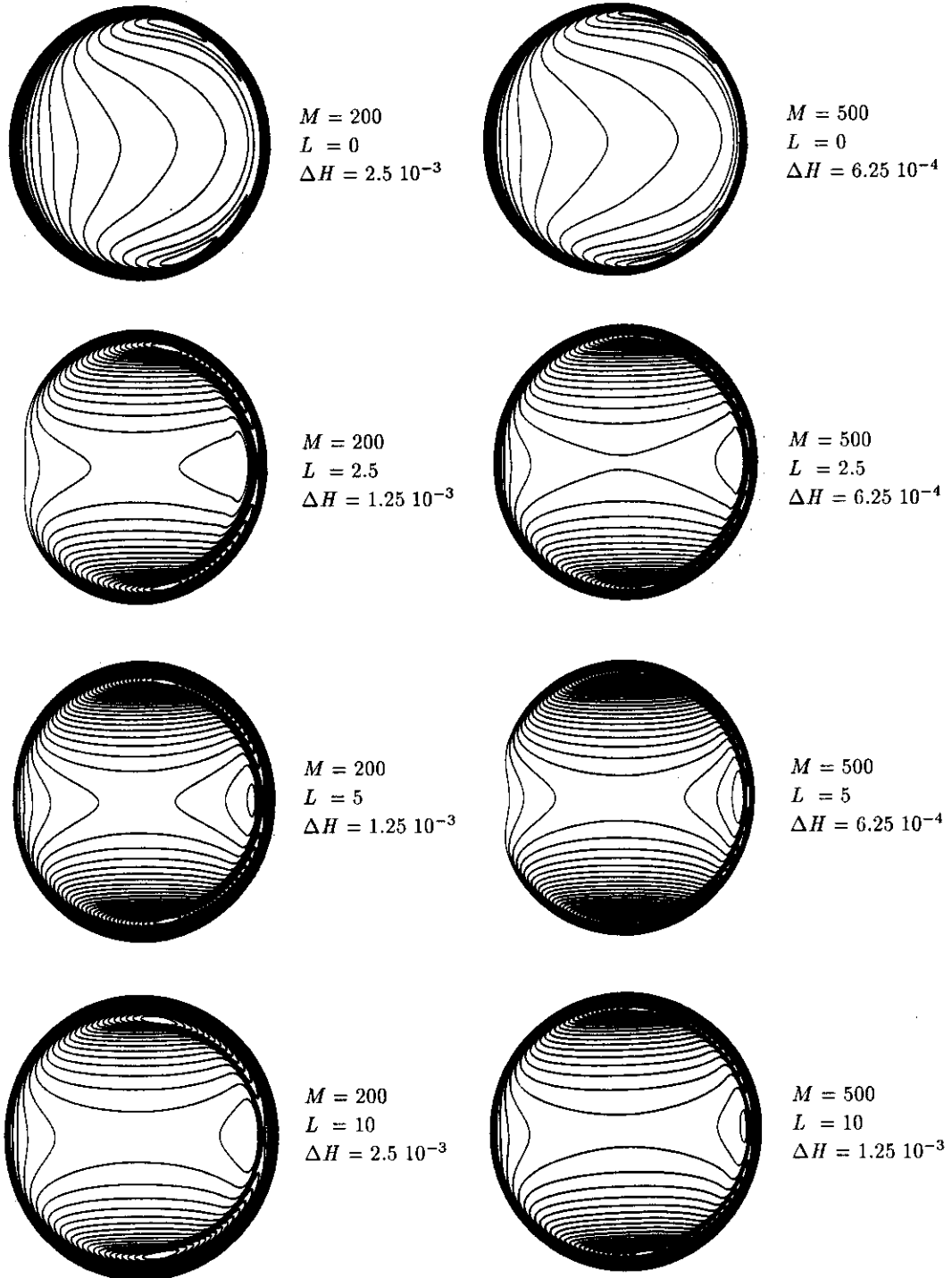
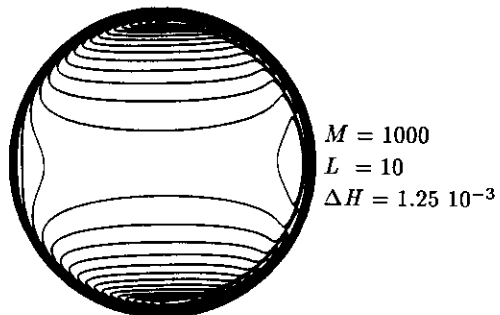
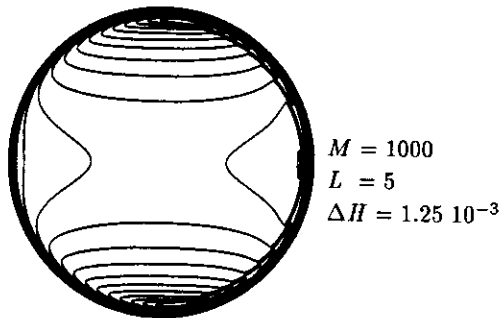
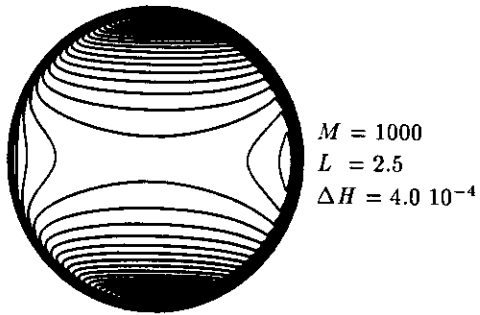
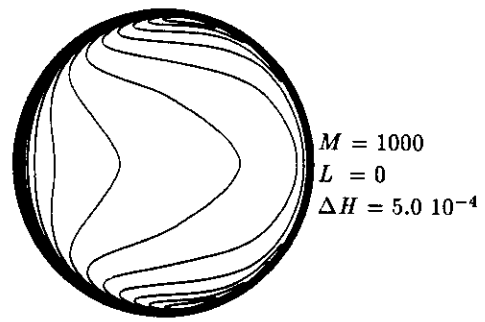


FIGURE 9.13: Film thickness contour plots. Notice the different values of  $\Delta H$ .

FIGURE 9.13: *Continued.*

FIGURE 9.13: *Continued.*

FIGURE 9.13: *Concluded.*

Another effect related to the increasing pressure dependence of the viscosity can be seen from the film thickness contour plots of figure 9.13. For a given value of  $L$  with increasing  $M$  the value of  $\bar{\alpha}$  and in fact the maximum Hertzian pressure increases. Consequently, the viscosity is higher and the flow in the contact region gradually approximates a pure shear flow. This follows from the fact that the film thickness contour plots become almost symmetric with respect to the Y-axis. This effect is particularly visible in the contour plots for  $M = 500$ ,  $L = 10$  and  $M = 1000$ ,  $L = 10$ . With  $\alpha = 1.7 \cdot 10^{-8}$  the latter solution represents a maximum Hertzian pressure of 2.14 GPa.

## 9.2 Film thickness

For the line contact problem discussed in chapter 6 the natural choice for a film thickness parameter to be used in overview graphs was the minimum film thickness. In the case of point contact problems the situation is more complex. Because of the additional dimension the fluid can flow around a region of small film thickness and, as a result, the ratio between the central film thickness (the film thickness at the location where both  $\partial P/\partial X = 0$  and  $\partial P/\partial Y = 0$ ) and the minimum film thickness is not necessarily a constant. This has indeed been observed over the past few years, e.g. see Lubrecht [L1] and Kweh et al. [K1] and the value of this ratio,  $H_c/H_m$ , as a function of  $M$  and  $L$  obtained from the present calculations is displayed in table 9.1.

This table clearly shows that the ratio is not a constant and consequently the minimum and central film thickness differ in their dependence on the load conditions. Hence, the question arises which parameter can be used best to present an overview of the results: The minimum film thickness, the central film thickness or maybe an average film thickness. In this work graphs of all three will be presented. Before doing so, section 9.2.1 discusses some asymptotic solutions for the central and the minimum film thickness that have been presented so far. Subsequently in section 9.2.2 the film thickness diagrams are given and in section 9.2.3 a formula is presented predicting the central film thickness as a function of the operating conditions.

		$M$								
		3	5	10	20	50	100	200	500	1000
$L$	0	1.3	1.3	1.3	1.3	1.4	1.5	1.6	1.7	1.9
	1	1.3	1.3	1.3	1.4	1.5	1.7	1.9	2.3	2.7
	2.5	1.3	1.3	1.3	1.5	1.7	1.9	2.2	2.6	3.1
	5	1.3	1.3	1.3	1.5	1.7	1.9	2.2	2.7	3.1
	10	1.3	1.2	1.3	1.5	1.6	1.8	2.0	2.4	2.8
	25	1.3	1.2	1.2	1.3	1.5	1.6	1.8	2.1	2.5

Table 9.1: Ratio  $H_c/H_m$  as a function of  $M$  and  $L$ .



### 9.2.1 Asymptotic solutions

For the case of the line contact problem the minimum or central film thickness could be solved analytically in some simplified situations, e.g. Martin [M1] and Gümbel [G1], Ertel [E1] and Grubin [G2], Moes [M2], which resulted in the well known asymptotic solutions. In the point contact situation the additional dimension makes an analytical solution of the problem, even in the restricted situations, complicated, if not impossible, and only a few asymptotic solutions for minimum or central film thickness are known to the author.

Firstly, the rigid isoviscous asymptote is discussed, i.e. the equivalent of the Martin and Gümbel solution for the line contact problem. A straightforward dimensional analysis shows that this asymptote can be written as:

$$H_{min} = C M^{-2} \quad (9.1)$$

The remaining problem is to determine the value of the constant. In a more general sense, i.e. for elliptical contacts, this subject has among others been addressed by Kapitza [K2] and also by Brewe et al. [B1]. For a circular contact Kapitza's analysis gives  $C = 28.4$  whereas according to Brewe et al.:  $C = 35.0$ . The relatively large difference between these two values can be ascribed to the different types of analysis used. In particular, Kapitza's prediction is based on half Sommerfeld solutions, i.e. disregarding cavitation.

The author solved this specific asymptotic problem numerically. For the circular contact  $C = 35.5$  was obtained, see [V1]. This value was already reported by Lubrecht [L1]. Comparing this result with the value presented by Brewe et al. it can be stated that this asymptotic solution is established rather accurately. Hence, it is justified to adopt the following formula for the rigid isoviscous asymptote for the circular contact:

$$H_{min} = 35.5 M^{-2} \quad (9.2)$$

In this asymptotic situation in which elastic deformation is absent the minimum film thickness equals 0.75 times the central film thickness. Consequently, the dimensionless central film thickness is given by:

$$H_{cen} = 47.3 M^{-2} \quad (9.3)$$

The second asymptotic situation considered is referred to as *elastic-isoviscous*. This asymptote is of particular interest for situations in which at least one of the surfaces has a small stiffness, e.g. in the case of seals. In terms of the Moes dimensionless parameters these situations are characterized by:  $L = 0$  and large  $M$ .

Contrary to the situation described by (9.2) and (9.3) now the elastic deformation is of great importance for the fluid film formation. All solutions for large  $M$  and  $L \approx 0$  clearly show this elastic deformation for example in the large region of nearly uniform film thickness and the formation of the side lobes, e.g. see figure 9.13.

Consequently, to describe this asymptotic situation in fact both a formula for the minimum as well as a formula for the central film thickness are required.

A solution for long elliptical contacts with the major axis perpendicular to the direction of rolling was derived by Moes [M2]:

$$H_{min} = 2.34 M^{-2/15} \quad (9.4)$$

However, because the analysis leading to this equation neglects the side flow it will probably overestimate the film thickness when applied to the circular contact situation. For that situation Moes proposed:

$$H_{min} = 1.31 M^{-2/15} \quad (9.5)$$

By means of the analytical solution of Reynolds' equation on the centerline of the contact under the assumption of a circular symmetric pressure profile and the elastic deformation given by Hertz [H2], the author derived the following expression for the central film thickness, see [V1]:

$$H_{cen} = 0.97 M^{-2/15} \quad (9.6)$$

Via a comparison of the predictions of this equation with numerically calculated values Lubrecht [L1] showed that the predicted slope of  $-2/15$  was fairly accurate but that the constant was too small.

The three formulas (9.4) to (9.6) are entirely based on analytical arguments. Alternatively Hamrock and Dowson [H1] curve fitted the results of some numerically calculated solutions and presented formulas for both the central and the minimum film thickness in the elastic isoviscous situation. In terms of the Moes dimensionless parameters their equation for the minimum film thickness reads:

$$H_{min} = 2.035 M^{-0.21} \quad (9.7)$$

whereas their equation for dimensionless central film thickness reads:

$$H_{cen} = 5.29 M^{-0.22} \quad (9.8)$$

Note that there is only a slight difference in the way in which the minimum film thickness and the central film thickness depend on the load ( $M$ ). This doesn't seem to agree with results of numerical calculations presented by Kweh et al. [K1] and Lubrecht [L1] who found that the minimum film thickness decreases much faster with increasing  $M$  than the central film thickness does. This subject will be addressed in more detail in section 9.2.2.

For the third asymptotic situation considered in chapter 6 for the line contact problem, i.e. the Grubin solution, the author derived the following equation:

$$H_{cen} = 1.02 M^{-1/12} L^{3/4} \quad (9.9)$$

This formula was obtained solving Reynolds' equation on the centerline of the contact assuming a circular symmetric pressure profile and the elastic deformation given by Hertz [H2], see [V1]. Comparison of the predictions of this formula with calculational results showed that it was relatively inaccurate.

### 9.2.2 Film thickness diagrams

Figure 9.14 displays the presently calculated values of  $H_{min}$  as a function of  $M$  and  $L$ . The most extreme situation represented in this diagram is  $M = 1000$  and  $L = 25$ . Alternatively this case is described by  $\bar{\alpha} = 91.09$  and  $\lambda = 1.10 \cdot 10^{-3}$ . Values of the Hamrock and Dowson parameters describing these conditions are:  $W = 9.46 \cdot 10^{-6}$ ,  $U = 7.81 \cdot 10^{-10}$ , and  $G = 4729$ . With  $\alpha = 1.7 \cdot 10^{-8}$  the maximum Hertzian pressure for this load situation is 5.35 GPa. The reader is reminded that, for the same reasons as mentioned for the line contact problem in chapter 6 (irrealistically high viscosities, lubricant solidification and gross plastic deformation), such high loads are of little practical relevance and the result is only displayed to demonstrate the stability of the algorithm.

In this figure the drawn line on the left represents the rigid isoviscous asymptote, i.e. equation (9.2). In addition the elastic isoviscous asymptote according to Hamrock and Dowson [H1] (equation (9.7)) is plotted. The figure shows that this latter asymptotic solution compares favourably with the numerically calculated values although for large  $M$  ( $M = 1000$ ) the calculated value falls somewhat below the prediction of equation (9.7). Also for  $L > 0$  and large  $M$  the dimensionless minimum film thickness shows a tendency to decrease with a larger slope than equation (9.7) predicts. This probably is a numerical effect caused by the fact that for high values of  $M$  the sidelobes become very small, see figure 9.13. As a result at least locally a large nodal density is required to obtain an accurate estimate of its value and location. This only applies to the minimum film thickness. An accurate estimate of the central film thickness does not require such large nodal densities.

This observation probably explains why the values of  $H_{min}$  for  $M \geq 100$  and small  $L$  presented by Lubrecht [L1] are much smaller than those obtained from the present calculations for the same conditions whereas the value of the central film thickness is about the same. This need for a large local nodal density to accurately represent the side-lobes in these situations also becomes evident from comparing the film thickness contour plots presented in [L1, page 132-135] with those for the same conditions presented in figure 9.13.

Figure 9.15 shows the presently calculated values of the dimensionless central film thickness,  $H_{cen}$ , as a function of  $M$  and  $L$ . The drawn lines indicate the predictions of this film thickness according to a formula presented in section 9.2.3. Clearly the calculational results and the predictions of this formula match quite accurately. Finally, figure 9.16 presents the calculated values of  $H_{av}$ , the dimensionless average film thickness, as a function of  $M$  and  $L$ , where the average is taken over the Hertzian contact region ( $X^2 + Y^2 \leq 1$ ).

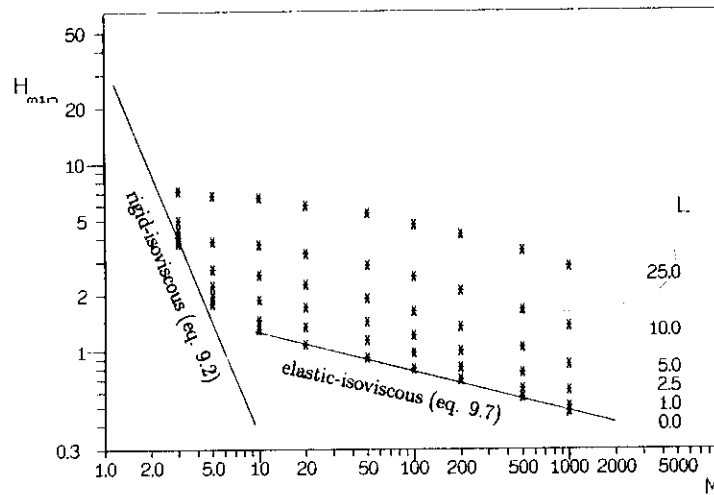


FIGURE 9.14: Calculated values of the dimensionless minimum film thickness parameter  $H_{min}$  (\*) as a function of  $M$  and  $L$ , the drawn lines give the predictions of equation (9.2) and (9.7) respectively.

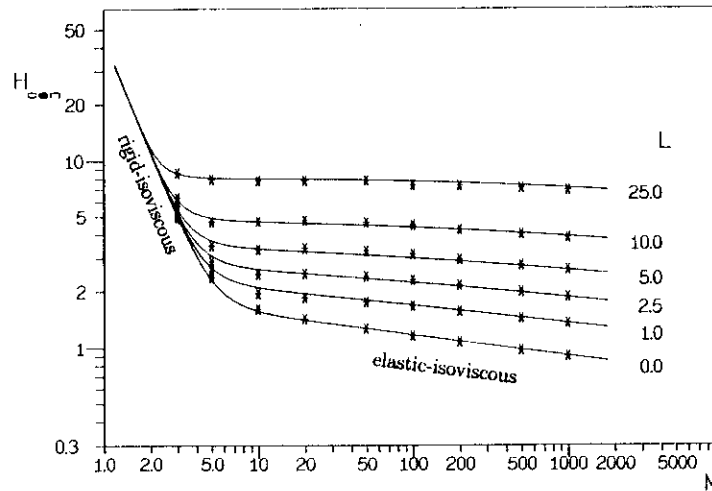


FIGURE 9.15: Calculated values of the dimensionless central film thickness parameter  $H_{cen}$  (\*) as a function of  $M$  and  $L$ , the drawn lines give the predictions of equation (9.12) presented in section 9.2.3.

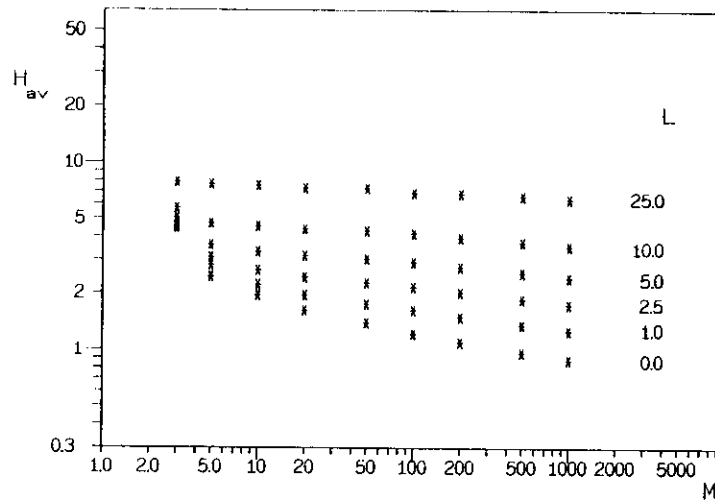


FIGURE 9.16: Calculated values of the dimensionless average film thickness parameter  $H_{av}$  (\*) as a function of  $M$  and  $L$ .

### 9.2.3 Film thickness formula

From the calculational results the following asymptote for the central film thickness in the elastic isoviscous regime was derived:

$$H_{cen} = 1.96 M^{-1/9} \quad (9.10)$$

The predictions of this formula accurately approach the calculated values for large  $M$  and  $L = 0$ . Note that the slope  $-1/9$  is less steep than the slope of the Hamrock and Dowson minimum film thickness asymptote (equation (9.7)) for these conditions. This clearly shows that the minimum film thickness decreases much faster with load than the central film thickness does.

For the same situation (large  $M$ ) but with  $L > 0$  the following equation quite accurately predicts the central film thickness:

$$H_{cen} = 1.70 M^{-1/9} L^{3/4} \quad (9.11)$$

Using three parameters  $r$ ,  $s$  and  $t$  to take care of a smooth transition in the intermediate region the equations (9.3), (9.10) and (9.11) can be merged into one expression that predicts the dimensionless central film thickness over the entire parameter range. This formula was developed by Moes [M2] and reads:

$$H_{cen} = \left[ \left\{ (1.70 M^{-1/9} L^{3/4} t)^r + (1.96 M^{-1/9})^r \right\}^{s/r} + (47.3 M^{-2})^s \right]^{1/s} \quad (9.12)$$

where  $r$ ,  $s$  and  $t$  are given by:

$$r = \exp\{1 - 6/(L + 8)\} \quad (9.13)$$

$$s = 12 - 10 \exp(M^{-2}) \quad (9.14)$$

and:

$$t = 1 - \exp\{-0.9 \frac{M^{1/6}}{L^{1/6}}\} \quad (9.15)$$

Note that this formula has exactly the same structure as the minimum film thickness formula (equation (6.10)) derived for the line contact problem. The straightforward structure allows easy programming on a personal or pocket computer.

The values for the dimensionless central film thickness predicted by equation (9.12) are indicated by the drawn lines in figure 9.15. When compared with the calculated values it shows that this formula accurately predicts the central film thickness in the entire parameter regime considered in the present study.

### 9.3 Conclusion

The algorithm developed for the solution of the EHL circular contact problem that was described in chapter 5 has been applied to the situation where the surfaces are perfectly smooth. The variation of the pressure profile and particularly the variation of the film thickness was studied over a wide range of parameter values. This parametric study has resulted in an accurate formula for the prediction of the central film thickness.

Although the subject was discussed only briefly, the pressure profiles presented in this chapter showed that, as was found earlier for the line contact, the (three dimensional equivalent of) the pressure spike, only occurs in part of the parameter regime. In addition, for a given  $M$  it develops gradually from a small ridge into a real "spike" when the dependence of the viscosity on the pressure increases.

This chapter was restricted to the idealized situation of perfectly smooth surfaces. In the next chapter some more complex EHL circular contact situations will be considered.

## References

- [B1] **Brewer, D., Hamrock, B.J., and Taylor, C.M.**, 1979, "Effect of geometry on hydrodynamic film thickness," *ASME JOT*, **101**, 231-239.
- [E1] **Ertel, A.M.**, 1939, "Hydrodynamic lubrication based on new principles," *Akad. Nauk SSSR Prikladnaya Matematika i Mekhanika*, **3**, **2**, 41-52
- [G1] **Gümbel, L.**, 1916, "Über geschmierte Arbeitsräder," *Z. ges. Turbinenwesen*, **13**, 357.
- [G2] **Grubin, A.N., and Vinogradova, I.E.**, 1949, "Investigation of the contact of machine components," Kh, F. Ketova (ed.) *Central Scientific Research Institute for Technology and Mechanical Engineering* (Moscow), Book No. 30, (DSIR translation No. 337)
- [H2] **Hertz, H.**, 1881, "Über die Berührung fester elastischer Körper," *J. für die reine und angew. Math.*, **92**, 156-171.
- [H1] **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.
- [K1] **Kweh, C.C., Evans, H.P., and Snidle, R.W.**, 1989, "Elastohydrodynamic lubrication of heavily loaded circular contacts," *Proc. Instn. Mech. Engrs.*, **203**, 133-148.
- [K2] **Kapitza, P.L.**, 1955, "Hydrodynamic theory of lubrication during rolling," *Zh. Tekh. Fiz.*, **25**, **2**, 747.
- [L1] **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.
- [M1] **Martin, H.M.**, 1916, "Lubrication of gear teeth," *Engineering* (London), **102**, 199.
- [M2] **Moes, H.**, (1986) internal memo on asymptotical film thickness calculations, University of Twente, Enschede, The Netherlands.
- [V1] **Venner, C.H.**, 1987, M. Sc. thesis, University of Twente, Enschede, the Netherlands (in dutch).





## Chapter 10

### Circular contact results: Surface features

So far the circular contact results were restricted to the situation of perfectly smooth surfaces, i.e. to the macro geometry of the contact only. However, this is a rather crude approximation of the reality in which the surfaces have a specific micro-geometry that will also affect the pressure profile and film shape.

This chapter considers the effect of some microgeometrical surface features on the pressure profile and film thickness in an EHL circular contact. As was explained in chapter 7, one of the reasons for carrying out such studies is that they can provide input, i.e. the pressure profile, to sub-surface stress calculations which are essential to obtain insight in the relation between surface imperfections and the fatigue life of the contact.

However, for point contacts there is another important reason to carry out such studies. In the case of a line contact the entire flow must pass the feature. As a result, in the steady state situation, the feature is almost completely flattened and the film shape is hardly affected. For point contact problems the situation may be quite different. Because of the additional dimension the fluid in principle can flow around a feature and consequently the feature may (locally) have a much more significant effect on the film thickness.

As in the line contact the microgeometrical features can be distinguished with respect to their characteristic length scale. On the one hand there are the features with a relatively large characteristic length scale, such as indentations, bumps and waviness and on the other hand there is surface roughness.

The influence of "large scale" microgeometrical features on the lubrication of circular contacts has already received quite some interest over the last few decades, particularly from an experimental point of view. Using optical interferometry several authors have measured film shapes in the case of artificially created surface imperfections e.g. [J1,W1,W2,K1,C1,C2]. Also from a theoretical point of view this subject has been addressed although, unlike the situation for the line contact problem, the number of papers dealing with the effect of microgeometrical features in point con-

tacts is rather limited, e.g. Lubrecht [L1,L2], Barragan de Ling et al. [B1], Kweh et al. [K2] and also Seabra and Berthe [S1]. In particular, Lubrecht [L1] presents an extensive series of results showing the effect of surface features upon the pressure profile and film shape. Among the features he considered are harmonic waviness (transversal and longitudinal), harmonic furrows and ridges (transversal and longitudinal) and an isotropic harmonic bump. However, his results are restricted to a relatively lightly loaded situation ( $p_h \approx 0.5$  GPa).

With respect to surface roughness the situation is less favourable. Only a few papers dealing with this subject are known to the author. The reader is for example referred to Bush and Skinner [B2], Zhu and Cheng [Z1] and to Jeng and Hamrock [J2]. From a point of view of surface fatigue life, unfortunately, all three studies have in common that an averaged approach is used, i.e. a perturbation approach and the flow factor method respectively. Furthermore, the results have been obtained using a rather limited number of nodes and are restricted to relatively lightly loaded conditions.

The main reason for the limited number of papers dealing with the aforementioned topics from a theoretical (computational) point of view is probably the computational complexity of the problem. Firstly, already in the case of perfectly smooth surfaces the accurate solution of the pressure and film shape requires a relatively large number of nodes. This situation is even worse if a local feature has to be included. It will be clear that for these kind of studies a fast algorithm is a prerequisite as was explained in chapter 1. Secondly, the study of surface features puts a significant strain on the stability of the algorithm since locally very extreme conditions may occur.

In chapter 9 it was demonstrated that the algorithm for the numerical solution of the circular contact presented in this thesis is very stable and allows solution of the problem to high loads with large nodal densities. In the present chapter the algorithm is tested on some situations with "large scale" surface features, i.e. indentations, bumps and longitudinal as well as transverse waviness. The presented results apply to one specific load condition and are restricted to steady state situations. In addition it is assumed that the Reynolds equation remains valid, i.e. only relatively shallow features will be studied. With respect to indentations and waviness the validity of Reynolds' equation is not the only reason for this restriction. As discussed in chapter 7 for the line contact problem, the validity of the results for these features obtainable with the present model is limited because of the way it accounts for cavitation, e.g. see also section 7.2.3.2.

## 10.1 Computational details

The load conditions are representative for an actual bearing application. All parameters are fixed and the conditions are such that the maximum Hertzian pressure is 2.0 GPa. Table 10.1 gives the values of the different parameters and the result-

Parameter	Value	Dimension
$E'$	$2.26 \cdot 10^{11}$	$[Pa]$
$\alpha$	$2.2 \cdot 10^{-8}$	$[Pa^{-1}]$
$z$	0.68	
$\eta_0$	$40 \cdot 10^{-3}$	$[Pa\ s]$
$u_s$	3.5	$[ms^{-1}]$
$R$	$1.80 \cdot 10^{-2}$	$[m]$
$a$	$5.0 \cdot 10^{-4}$	$[m]$
$p_h$	$2.0 \cdot 10^9$	$[Pa]$
$M$	1007.6	
$L$	12.05	
$\bar{\alpha}$	44.0	
$\lambda$	$1.09 \cdot 10^{-3}$	

Table 10.1: Parameter values used in the calculations.

ing values of the Moes dimensionless parameters as well as the values of  $\bar{\alpha}$  and  $\lambda$ . The values of the Hamrock and Dowson dimensionless parameters for this loading situation are:  $W = 1.43 \cdot 10^{-5}$ ,  $U = 1.72 \cdot 10^{-11}$  and  $G = 4972$ .

In section 10.2 the smooth surface pressure profile and film shape is discussed. Subsequently in the following sections of this chapter it is investigated how the pressure profile and film shape for this load situation are affected by some surface features. To take into account a surface feature the equation for the dimensionless film thickness, i.e. equation (2.49), is extended to:

$$H(X, Y) = H_{00} + \frac{X^2}{2} + \frac{Y^2}{2} - \mathcal{R}(X, Y) + \frac{2}{\pi^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{P(X', Y') dX' dY'}{\sqrt{(X - X')^2 + (Y - Y')^2}} \quad (10.1)$$

where  $\mathcal{R}(X, Y)$  denotes the undeformed geometry of the surface feature.

## 10.2 Smooth surface solution

The contour plots of the (dimensionless) pressure profile and the (dimensionless) film shape for this load condition assuming perfectly smooth surfaces are depicted in figure 10.1 and 10.2. In addition figure 10.3 displays the pressure and film thickness at the centerline of the contact ( $Y = 0$ ). Furthermore, figure 10.4 shows a cross-section of the pressure profile and film shape in the perpendicular direction, i.e. pressure and film thickness as a function of  $Y$  on the line  $X = 0$ .

The solution has been calculated using 263.169 nodal points on a grid covering the domain  $\{(X,Y) \in \mathbb{R}^2 | -2.5 \leq X \leq 1.5, -2 \leq Y \leq 2\}$ . Because of the relatively high value of  $\bar{\alpha}$  ( $\bar{\alpha}=44$ ) the pressure profile closely approximates the Hertzian dry contact (semi-ellipsoid) pressure profile. This is illustrated by figure 10.3 and 10.4. In the contour plot this is reflected by the fact that the lines of constant pressure show up as circles. The pressure profile deviates from the Hertzian semi-ellipsoid in the inlet region and just before the outlet region where the pressure spike occurs.

Figure 10.2, the contour plot of the film thickness, shows that the side lobes, obviously visible in figure 10.4, are rather small. This is characteristic for such a large value of  $M$  as considered here. In addition the region of nearly uniform film thickness covers almost the entire Hertzian dry contact circle. The calculated minimum film thickness for this situation is  $0.14 \mu m$  whereas the calculated central film thickness amounts to  $0.40 \mu m$

---

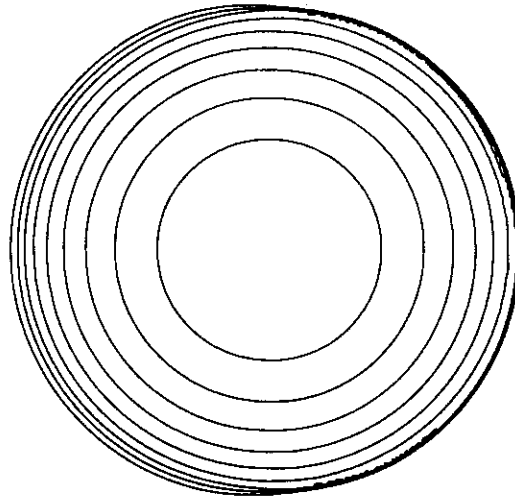


FIGURE 10.1:  $M = 1007.6$ ,  $L = 12.05$ . Pressure profile contour plot in the case of smooth surfaces,  $\Delta P = 0.1$ .

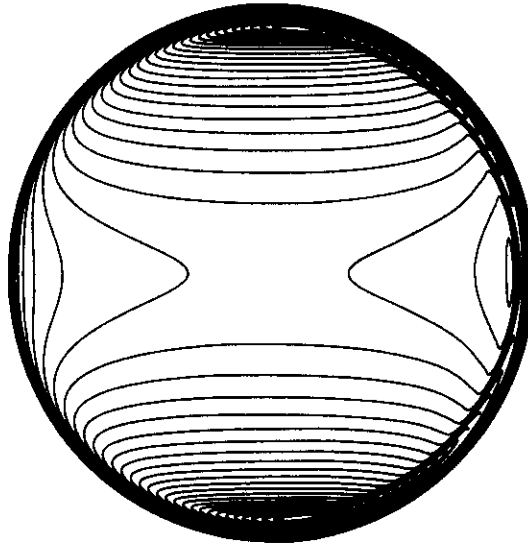


FIGURE 10.2:  $M = 1007.6$ ,  $L = 12.05$ . Film thickness contour plot in the case of smooth surfaces,  $\Delta H = 7.5 \cdot 10^{-4}$ .

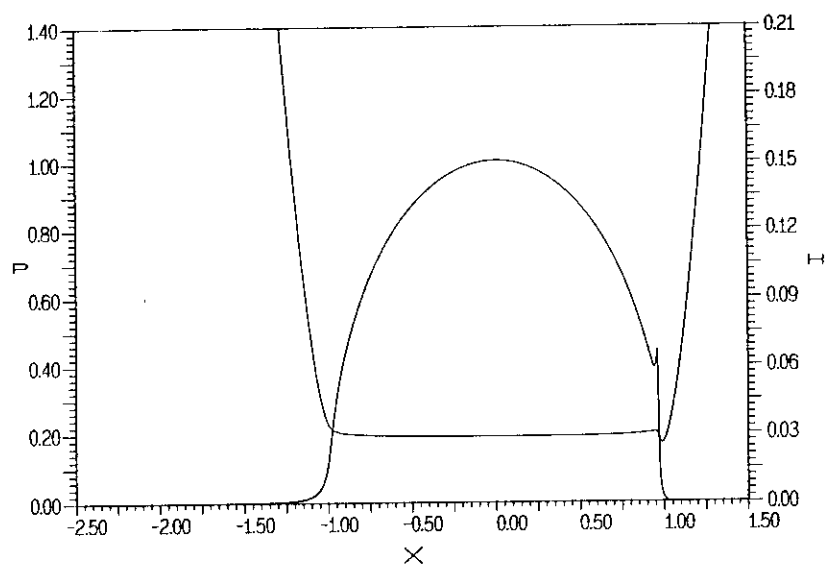


FIGURE 10.3:  $M = 1007.6$ ,  $L = 12.05$ . Pressure,  $P$ , and film thickness,  $H$ , as a function of  $X$  at the centerline  $Y = 0$  in the case of smooth surfaces.

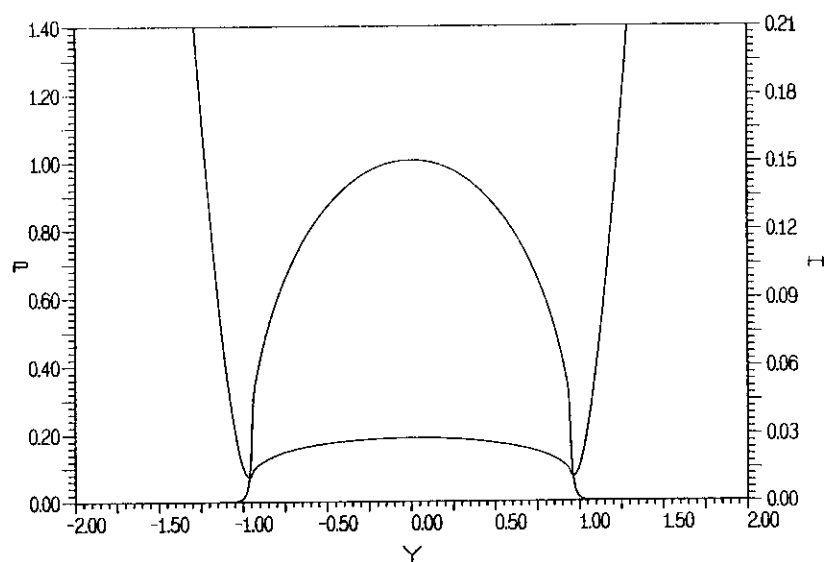


FIGURE 10.4:  $M = 1007.6$ ,  $L = 12.05$ . Pressure,  $P$ , and film thickness,  $H$ , as a function of  $Y$  at the line  $X = 0$  in the case of smooth surfaces.

### 10.3 Bumps

First the effect of a bump located at the center of the contact is studied. The geometry of the bump is the rotational symmetric equivalent of the bump geometry considered in chapter 7:

$$\mathcal{R}(X, Y) = \mathcal{A} 10^{-10\left(\frac{X^2+Y^2}{w^2}\right)} \cos\left(2\pi \frac{\sqrt{X^2+Y^2}}{\mathcal{W}}\right) \quad (10.2)$$

where:  $\mathcal{A}$  = dimensionless amplitude ( $\mathcal{A} > 0$  for a bump):  $\mathcal{A} = AR/a^2$   
 $\mathcal{W}$  = dimensionless wavelength of the bump:  $\mathcal{W} = w/a$

As an example figure 10.5 and figure 10.6 present the contour plots of the pressure and film shape respectively for the case of a bump with a wavelength of  $0.5 \text{ mm}$  and an amplitude of  $1 \text{ }\mu\text{m}$  ( $\mathcal{W} = 1.0$  and  $\mathcal{A} = 0.142$ ) located at the center of the contact. In addition figure 10.7 displays the pressure and film thickness at the centerline of the contact as a function of  $X$  and figure 10.8 shows the pressure and film thickness as a function of  $Y$  on the line  $X = 0$ . The solution has been calculated using the same number of nodes and domain as the smooth surface solution presented in figure 10.1 and 10.2.

The figures 10.5, 10.7 and 10.8 show that the bump causes a local increase of the pressure in the center of the contact. By comparing figure 10.7 with figure 10.3 it is obvious that the maximum pressure exceeds the maximum pressure in the case of smooth surfaces. In fact, disregarding the inlet and outlet regions the pressure profile in the present circular contact situation is the circular symmetric equivalent of the characteristic line contact pressure profile in the case of a bump, as was shown in figure 7.12.

Contrary to the pressure profile the film thickness is only slightly affected by the presence of the bump. In spite of the fact that the amplitude of the bump is more than twice the central film thickness, it hardly shows up in the film shape. As was found for the line contact problem, see section 7.2.4, the bump is almost completely flattened. This obviously indicates that, for the present conditions, in the major part of the Hertzian contact circle the flow is dominated by shear flow ( $\partial(\rho H)/\partial X \approx 0$ ). Hence, because of the high viscosities, the fluid simply can not flow around the bump.

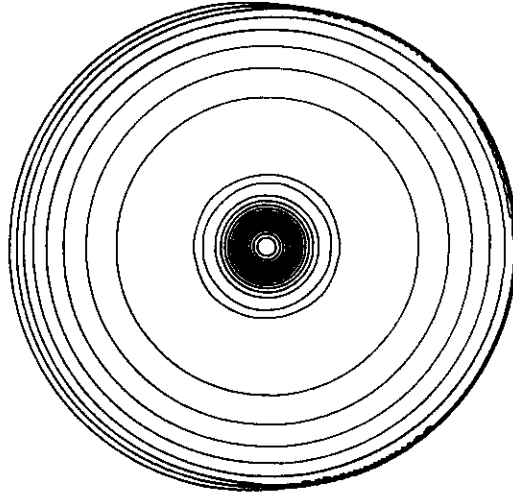


FIGURE 10.5:  $M = 1007.6$ ,  $L = 12.05$ . Pressure profile contour plot in the case of a bump with a wavelength of  $0.5 \text{ mm}$  and an amplitude of  $1 \mu\text{m}$ ,  $\Delta P = 0.1$ .

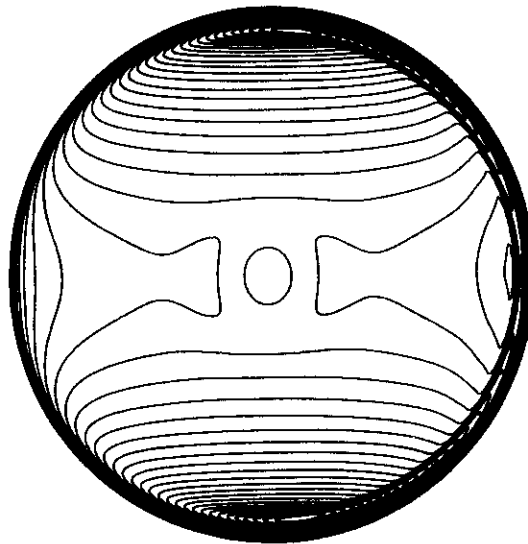


FIGURE 10.6:  $M = 1007.6$ ,  $L = 12.05$ . Film thickness contour plot in the case of a bump with a wavelength of  $0.5 \text{ mm}$  and an amplitude of  $1 \mu\text{m}$ ,  $\Delta H = 7.5 \cdot 10^{-4}$ .



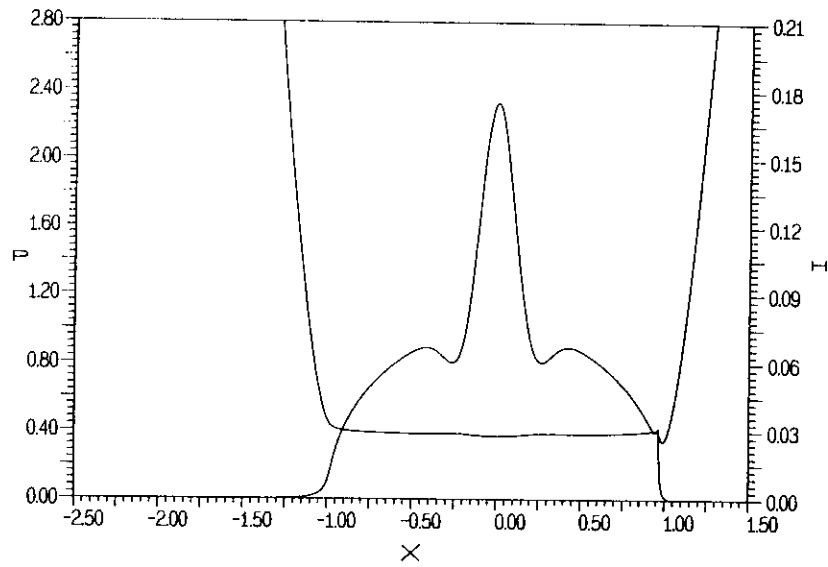


FIGURE 10.7:  $M = 1007.6$ ,  $L = 12.05$ . Pressure  $P$  and film thickness  $H$  as a function of  $X$  at the centerline  $Y = 0$  in the case of a bump with a wavelength of  $0.5 \text{ mm}$  and an amplitude of  $1 \mu\text{m}$ .

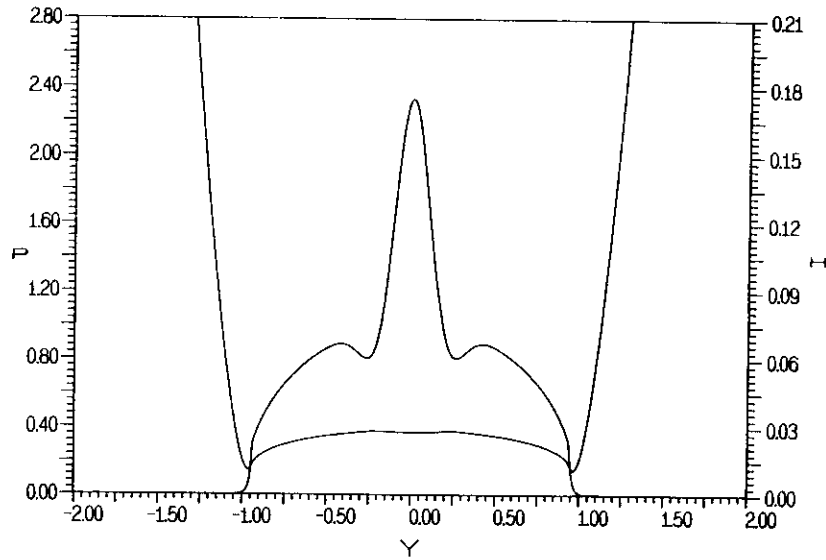


FIGURE 10.8:  $M = 1007.6$ ,  $L = 12.05$ . Pressure  $P$  and film thickness  $H$  as a function of  $Y$  at the line  $X = 0$  in the case of a bump with a wavelength of  $0.5 \text{ mm}$  and an amplitude of  $1 \mu\text{m}$ .

Of course the effect of the bump on the pressure profile and the film shape depends on its amplitude and wavelength. For a fixed amplitude of  $1\text{ }\mu\text{m}$  figure 10.9 shows the relative maximum pressure as a function of the wavelength. Note that the behaviour displayed in this figure is basically the same as the behaviour that was observed for the line contact problem in section 7.2.4: The relative maximum pressure is inversely proportional to the wavelength. With respect to the relatively large increase of the maximum pressure for small wavelengths it should be noted that these results are of limited practical relevance. In reality the pressure at the bump can not rise beyond the plasticity limit of the material.

Next the figures 10.10 and 10.11 show the results for the associated minimum and central film thicknesses. Figure 10.10 shows that for small wavelengths the effect on the film thickness is negligible whereas for the larger wavelengths the bump causes a slight increase of the minimum film thickness. This is explained by the fact that the minimum film thickness, as mentioned before for the line contact problem, is determined in the inlet region. In the case of a relatively small wavelength the bump does not affect the geometry in this region and consequently the effect of the bump on the minimum film thickness is very small.

Hence, in these situations the bump induces only local changes. However, with increasing wavelength the bump gradually starts to effect the global contact geometry, in particular the geometry in the inlet region. Consequently, in these situations the minimum film thickness is affected. For large wavelengths ( $W \gg 2$ ) these effects disappear again and the smooth surface solution slowly returns. Note that the dependence of the minimum film thickness on the wavelength is roughly the same

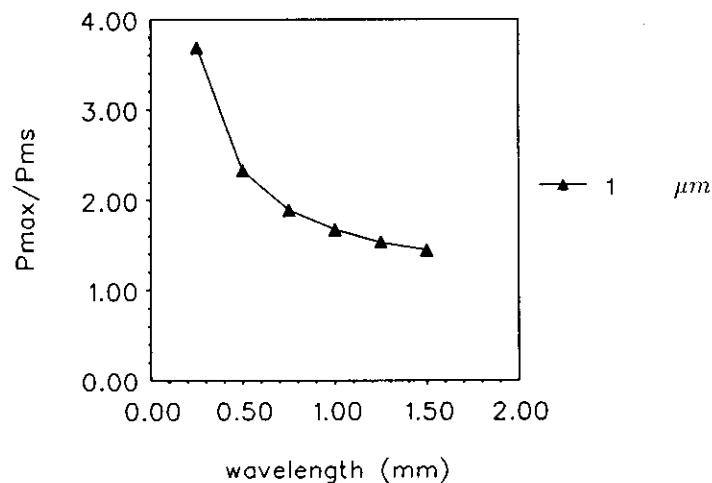


FIGURE 10.9:  $M = 1007.6$ ,  $L = 12.05$ . Relative maximum pressure as a function of wavelength in the case of a bump with an amplitude of  $1\text{ }\mu\text{m}$ .

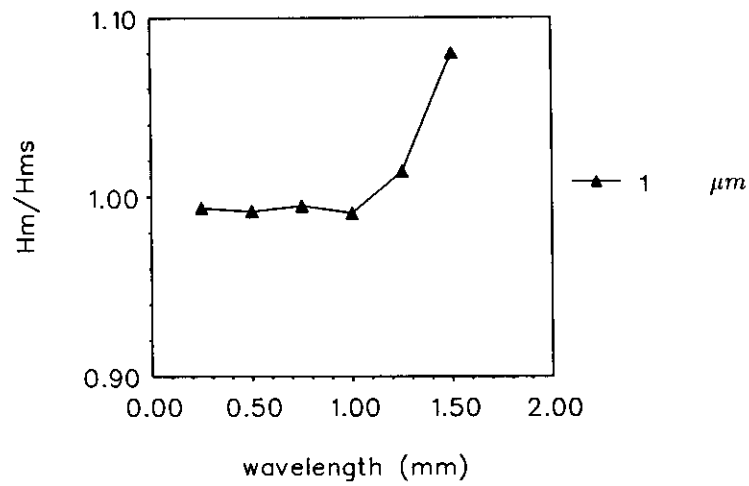


FIGURE 10.10:  $M = 1007.6$ ,  $L = 12.05$ . Relative minimum film thickness as a function of wavelength in the case of a bump with an amplitude of  $1 \mu m$ .

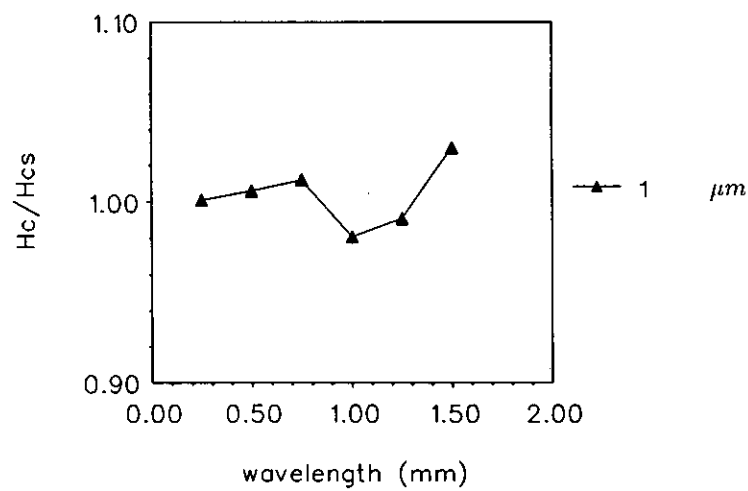


FIGURE 10.11:  $M = 1007.6$ ,  $L = 12.05$ . Relative central film thickness as a function of wavelength in the case of a bump with an amplitude of  $1 \mu m$ .

as the behaviour observed in the case of the line contact problem, see figure 7.13. Figure 10.11 shows that the effect of the bump on the central film thickness is even smaller than the effect on the minimum film thickness. For the situations considered here the changes are less than four percent.

The reader is reminded that the results presented in this section are restricted to a bump located in the center of the contact, i.e. located in the high pressure region. If the bump is located in the inlet or outlet region or if it is located near the sides of the contact the flattening effect may be much smaller and consequently it may have a larger effect on the film thickness.

## 10.4 Indentations

Secondly the effect of an indentation located at the center of the contact on the pressure profile and film shape will be investigated. The geometry of the indentation assumed in the present section is simply the geometry of the bump considered in the previous section apart from the opposite sign of the amplitude.

First, to give a general impression, the situation in the case of a very shallow indentation is considered. Figure 10.12 and figure 10.13 show the pressure contour plot and the contour plot of the associated film thickness for the present load conditions in the case of an indentation with a wavelength of 0.5 mm and an amplitude of 0.5  $\mu\text{m}$  ( $\mathcal{W} = 1.0$  and  $\mathcal{A} = -0.0712$ ). In addition to these contour plots figure 10.14 gives the pressure and film thickness on the centerline of the contact,  $Y = 0$ , as a function of  $X$  and figure 10.15 shows the pressure and film thickness on the line  $X = 0$  as a function of  $Y$ .

Figure 10.12, 10.14 and 10.15 show that the changes in the pressure profile caused by the dent are basically the rotational symmetric equivalent of the changes observed in the case of the line contact, i.e. see section 7.2.4. In particular, note the similarity between the pressure and the film thickness at the centerline, i.e. figure 10.14, and a typical line contact pressure profile and film shape in the case of a bump, e.g. see figure 7.2. In the case of a steady state line contact a dent located in the center of the contact causes an equal pressure rise at both shoulders of the indentation with a pressure drop at its center. In the present situation, of course, the shoulder is a circle and in fact a rotational symmetric pressure ridge is formed.

The film thickness contour plot, i.e. figure 10.13, and the film thickness on the lines  $Y = 0$  (figure 10.14) and  $X = 0$  (figure 10.15) respectively show that the film shape is almost the same as in the smooth surface situation, i.e. the dent hardly shows up. This indicates that in almost the entire Hertzian contact circle the flow is dominated by shear flow ( $\partial(\rho H)/\partial X \approx 0$ ). Hence, in spite of the pressure drop at the center of the dent, even locally, the viscosities are too high to allow pressure induced flow in this region.

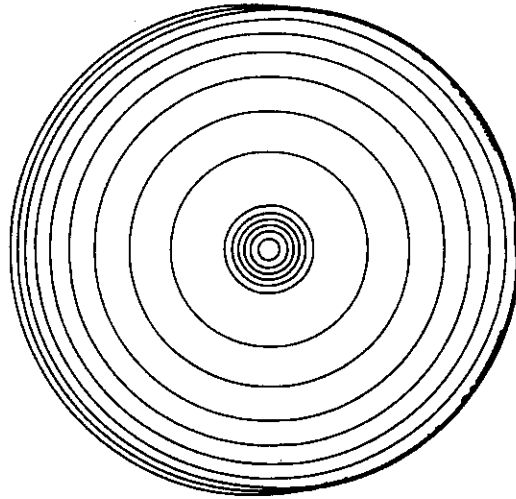


FIGURE 10.12:  $M = 1007.6$ ,  $L = 12.05$ . Pressure profile contour plot in the case of a dent with a wavelength of  $0.5 \text{ mm}$  and an amplitude of  $0.5 \text{ }\mu\text{m}$ ,  $\Delta P = 0.1$ .

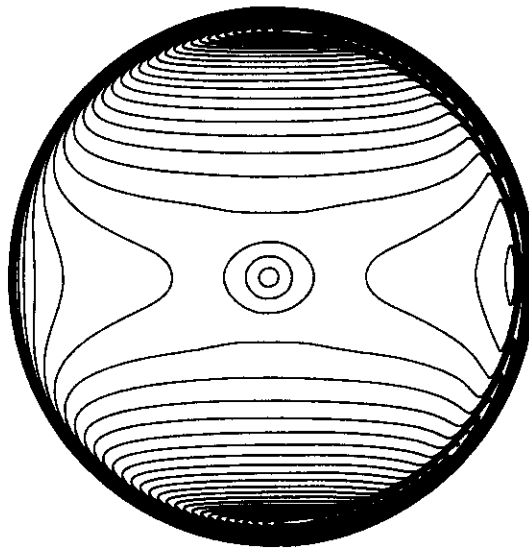


FIGURE 10.13:  $M = 1007.6$ ,  $L = 12.05$ . Film thickness contour plot in the case of a dent with a wavelength of  $0.5 \text{ mm}$  and an amplitude of  $0.5 \text{ }\mu\text{m}$ ,  $\Delta H = 7.5 \cdot 10^{-4}$ .

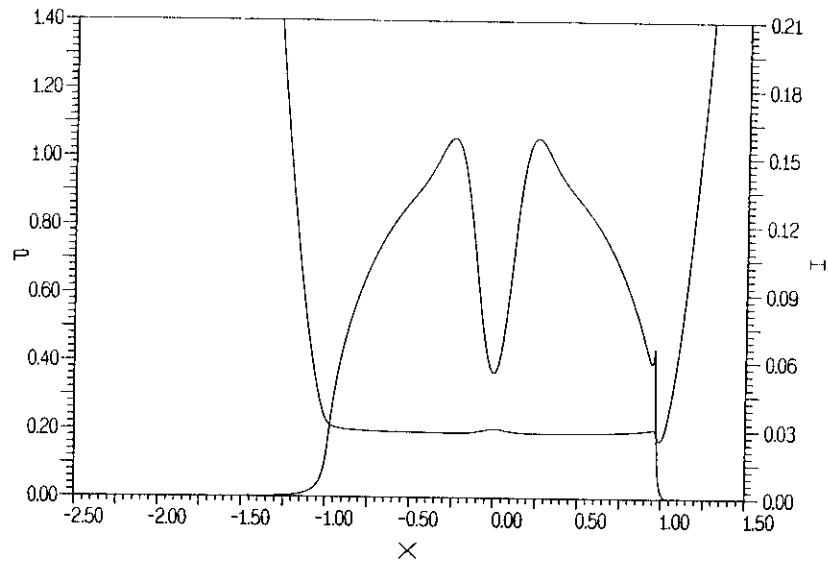


FIGURE 10.14:  $M = 1007.6$ ,  $L = 12.05$ . Pressure  $P$  and film thickness  $H$  as a function of  $X$  at the centerline  $Y = 0$  in the case of a dent with a wavelength of  $0.5 \text{ mm}$  and an amplitude of  $0.5 \mu\text{m}$ .

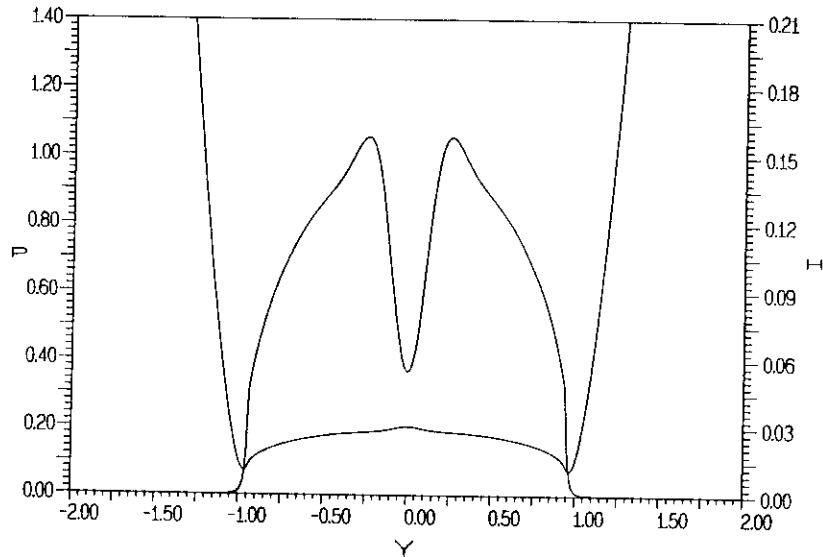


FIGURE 10.15:  $M = 1007.6$ ,  $L = 12.05$ . Pressure  $P$  and film thickness  $H$  as a function of  $Y$  at the line  $X = 0$  in the case of a dent with a wavelength of  $0.5 \text{ mm}$  and an amplitude of  $0.5 \mu\text{m}$ .

In the situation considered above the effect of the dent on the pressure profile is essentially restricted to the center of the contact. However, this situation can change rather drastically as is illustrated below. Figure 10.16 and figure 10.17 show the contour plots of the calculated pressure profile and the associated film thickness in the case of an indentation with an amplitude of  $1\ \mu\text{m}$  and a wavelength of  $0.5\ \text{mm}$  ( $W = 1.0$  and  $A = -0.142$ ). In addition figure 10.18 shows the pressure profile and the film thickness as a function of  $X$  on the centerline of the contact,  $Y = 0$ . Furthermore, figure 10.19 shows the pressure and film thickness as a function of  $Y$  on the line  $X = 0$ . From both the pressure contour plot as well as from the film thickness contour plot it is obvious that, the changes induced by the dent are no longer restricted to the central region only. On the contrary, over a specific width in  $Y$  direction significant changes in the pressure profile and film thickness occur in the entire wake of the dent. In fact, downstream of the dent a furrow in the film thickness occurs that extends to the edge of the Hertzian contact circle. In figure 10.18 this is reflected by the film thickness being uniform at different values on the inlet and outlet side respectively.

Notice the micro EHL features in the film thickness displayed in figure 10.18, i.e. a nearly uniform film thickness with a decrease at the end on both sides. Furthermore, contrary to the situation considered above in this case the pressure in the vicinity of the dent is not circular symmetric. Near the centerline of the contact the pressures in the wake of the dent are higher than the pressure at the same distance upstream of the dent center. This can be seen quite clearly in figure 10.18.

These effects may seem strange at first sight. However, to the authors opinion they are explained by the large pressure drop at the center of the dent as a result of which locally (close to the center of the dent) pressure induced flow is enabled. Since it is a two dimensional problem there is some fluid flow, directed to the center, where the pressure is smallest. This is visualized in figure 10.20.

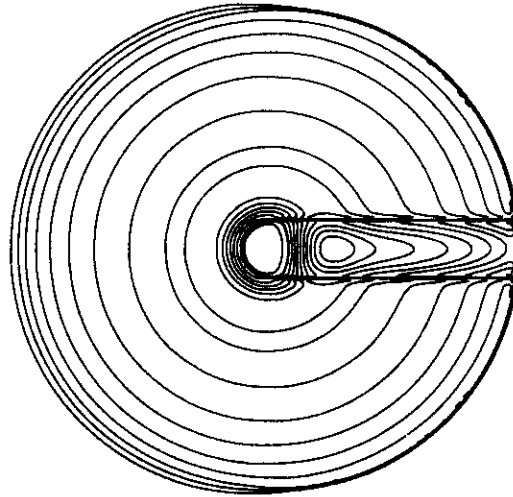


FIGURE 10.16:  $M = 1007.6$ ,  $L = 12.05$ . Pressure profile contour plot in the case of a dent with a wavelength of  $0.5 \text{ mm}$  and an amplitude of  $1 \mu\text{m}$ ,  $\Delta P = 0.1$ .

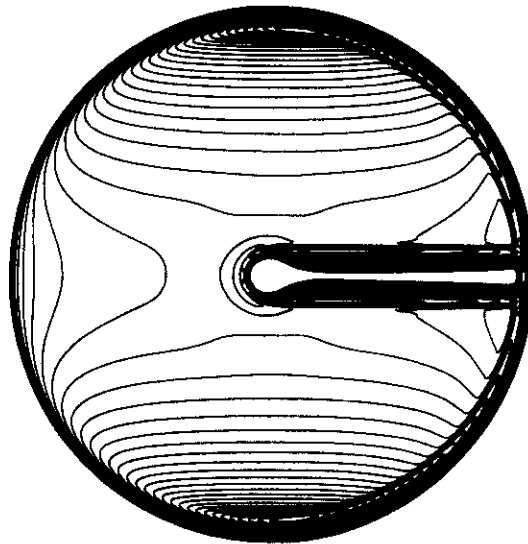


FIGURE 10.17:  $M = 1007.6$ ,  $L = 12.05$ . Film thickness contour plot in the case of a dent with a wavelength of  $0.5 \text{ mm}$  and an amplitude of  $1 \mu\text{m}$ ,  $\Delta H = 7.5 \cdot 10^{-4}$ .



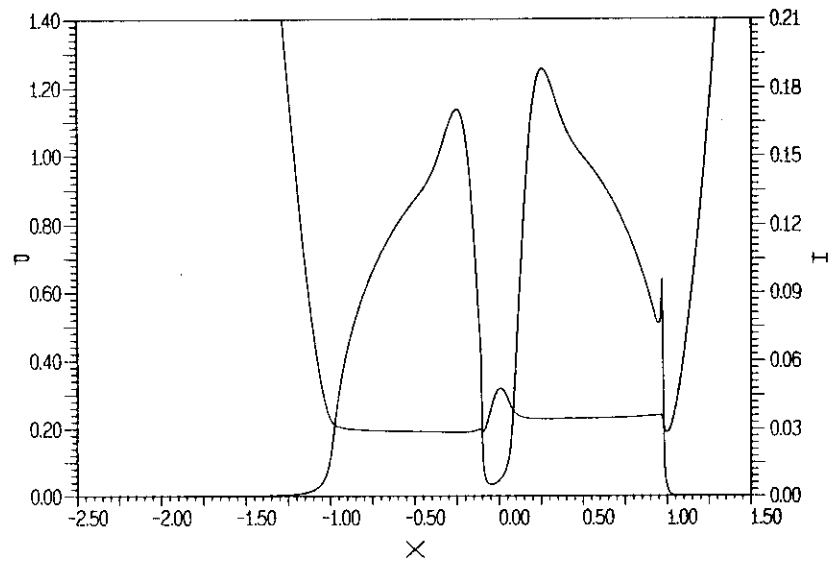


FIGURE 10.18:  $M = 1007.6$ ,  $L = 12.05$ . Pressure,  $P$ , and film thickness,  $H$ , as a function of  $X$  at the centerline  $Y = 0$  in the case of a dent with a wavelength of  $0.5 \text{ mm}$  and an amplitude of  $1.0 \mu\text{m}$ .

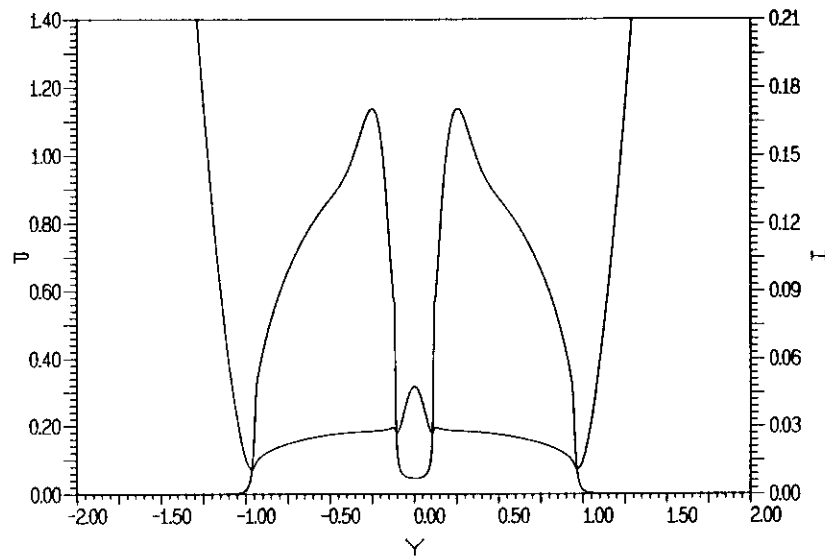


FIGURE 10.19:  $M = 1007.6$ ,  $L = 12.05$ . Pressure,  $P$ , and film thickness,  $H$ , as a function of  $Y$  at the line  $X = 0$  in the case of a dent with a wavelength of  $0.5 \text{ mm}$  and an amplitude of  $1.0 \mu\text{m}$ .

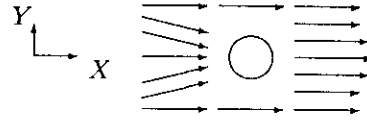


FIGURE 10.20: Visualization of flow near  $(X, Y) = (0, 0)$ , the center of the contact.

Because of mass conservation the “additional fluid” will have to leave the center again. Since the viscosity increases exponentially with increasing pressure and in the entire outer region the pressure is high this can only be done by means of unidirectional shear flow again, i.e. in the region downstream of the dent the Reynolds equation again reduces to  $\partial(\rho H)/\partial X \approx 0$ . However, as a result of the aforementioned “inflow” effect the mass flow in  $X$  direction in the region close to the centerline and on the centerline itself downstream of the dent is larger than upstream of the dent. Consequently downstream the film thickness will be uniform at a higher level. Restricted to the centerline only this is exactly what is displayed in figure 10.18.

Another fact that confirms the occurrence of pressure induced flow directed towards the dent center are the micro EHL features visible in figure 10.19 (the film thickness on the line  $X = 0$  as a function of  $Y$ ), i.e. the decrease of the film thickness on both sides of the dent.

Having described the changes in the pressure profile and film thickness induced by the indentation in two specific situations, the influence of the wavelength on these effects is studied next. First figure 10.21 displays the relative maximum pressure on the dent shoulder as a function of the wavelength for both an amplitude of  $0.5 \mu\text{m}$  and an amplitude of  $1 \mu\text{m}$ . This figure shows that, similar to what was found in the case of the line contact problem, the pressure increase caused by the indentation is approximately inversely proportional to the wavelength of the dent.

The associated minimum film thickness results presented in figure 10.22 show that the indentation results in a slight decrease of the minimum film thickness. In particular for the smallest wavelengths,  $W \leq 1.0$ , the effect is negligible ( $\leq 2\%$ ) which is also about the accuracy in the minimum film thickness result. For larger wavelengths the decrease is larger because in these situations the dent is no longer local. In these situations it changes the geometry in the inlet region and consequently it affects the minimum film thickness. However, for larger wavelengths ( $W \gg 2$ ), these effects disappear again and the smooth surface solution slowly returns.

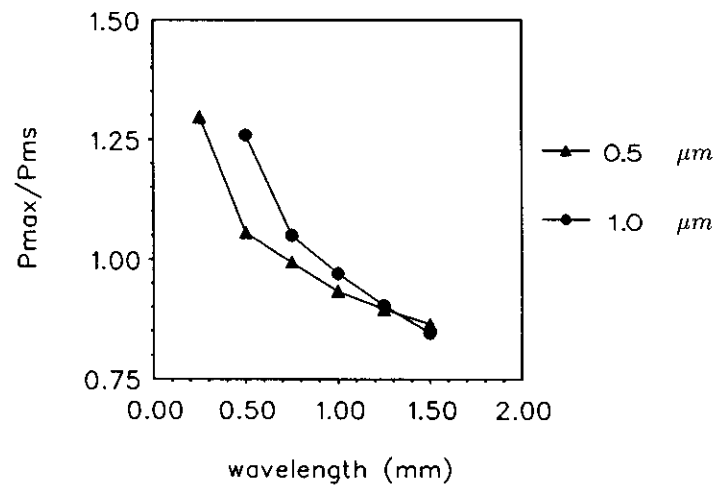


FIGURE 10.21:  $M = 1007.6$ ,  $L = 12.05$ . Relative maximum pressure as a function of wavelength in the case of a dent with an amplitude of 0.5 and 1.0  $\mu m$  respectively.

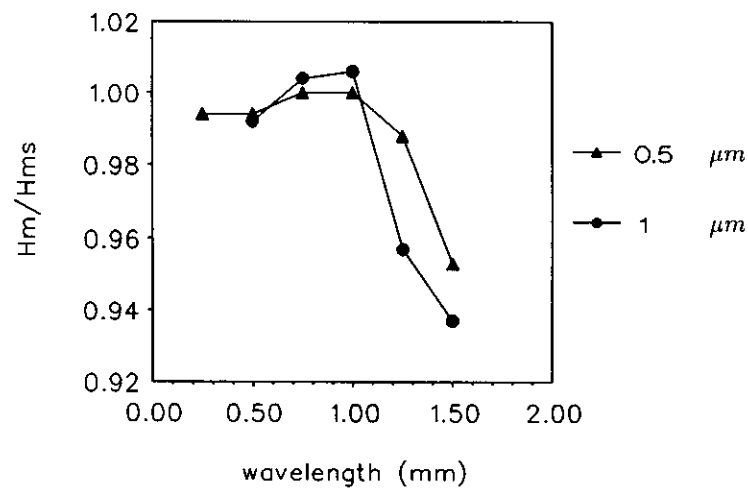


FIGURE 10.22:  $M = 1007.6$ ,  $L = 12.05$ . Relative minimum film thickness as a function of wavelength in the case of a dent with an amplitude of 0.5 and 1.0  $\mu m$  respectively.

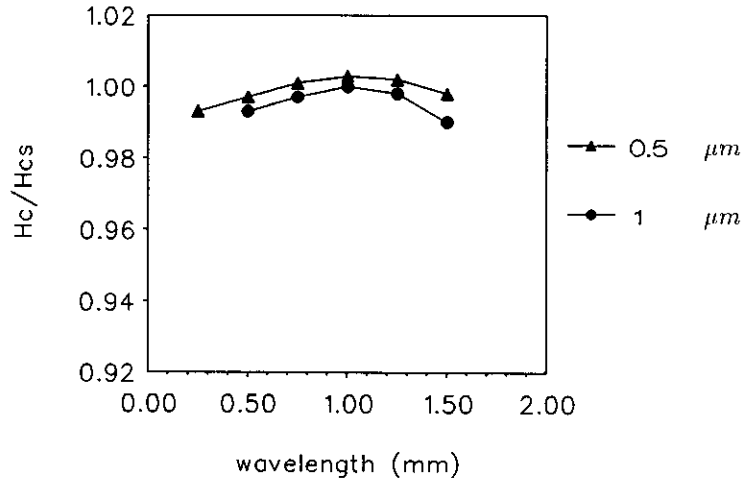


FIGURE 10.23:  $M = 1007.6$ ,  $L = 12.05$ . Relative central film thickness as a function of wavelength in the case of a dent with an amplitude of  $0.5$  and  $1.0 \mu m$ .

Figure 10.23 displays the central film thickness relative to the smooth surface value as a function of the wavelength of the dent. From this figure it can be seen that the changes in the central film thickness induced by the dent are even smaller than the changes in the minimum film thickness discussed above. In all situations considered the change is less than 1 %.

## 10.5 Waviness

This section considers the effect of harmonic waviness on the pressure profile and film shape. Two situations are considered, i.e. longitudinal and transverse waviness.

### 10.5.1 Longitudinal

In the case of longitudinal waviness  $\mathcal{R}$  is given by:

$$\mathcal{R}(X, Y) = \mathcal{A} \cos\left(\frac{2\pi Y}{\mathcal{W}}\right) \quad (10.3)$$

As an example figure 10.24 and 10.25 show the contour plots of the pressure profile and film shape for the present loading conditions in the case of a wavelength of  $0.25 \text{ mm}$  and an amplitude of  $0.25 \mu m$  ( $\mathcal{W} = 0.5$  and  $\mathcal{A} = 0.356$ ). In addition to those contour plots figure 10.26 and figure 10.27 show the pressure profile and film

shape as a function of  $X$  on the line  $Y = 0$  (the centerline) and as a function of  $Y$  on the line  $X = 0$  respectively.

From these figures it can be seen that the waviness results in an oscillation on the pressure profile with respect to the  $Y$  direction, see figure 10.27. Furthermore, comparing the pressure on the centerline of the contact in the present situation, figure 10.26, with the pressure on this line in the case of smooth surfaces, figure 10.3, it is obvious that the maximum pressure is increased. For small amplitudes this effect will be very small. However, with increasing amplitude, the ridge in the center of the contact gradually carries an increasing part of the load. In fact, depending on the wavelength, with increasing amplitude the contact gradually changes to a set of nearly independent elliptical contacts with the major axis in the direction of rolling where the ridge in the center carries the largest part of the external load. This tendency to carry the same load in a smaller region leads to an increased maximum pressure.

From the contour plot it can be seen that the effect of the longitudinal waviness on the global film shape is quite considerable, particularly, when compared to the previously considered situations of a *small* dent and a bump: For example, in section 10.3 it was shown that a bump in the center of the contact, is nearly completely flattened. This was explained by the fact that in the Hertzian contact region unidirectional shear flow dominates because of the high pressures. However, in fact the situation is a bit more complex as will be explained below. In the regions of high pressure (viscosities) the two dimensional Reynolds equation reduces to  $\partial(\bar{\rho}H)/\partial X \approx 0$ . The Dowson and Higginson relation limits the variations of the density, hence in the high viscosity region  $H$  tends to be a function of  $X$  only, the value of which is determined upstream. The important observation is that this applies only with respect to the  $X$  direction. As a result, if at the entrance of the high viscosity region the film thickness shows significant variations in  $Y$  direction, these variations tend to remain present throughout the entire high viscosity region. In this case of longitudinal waviness significant variations of  $H$  in  $Y$  direction obviously occur in the low pressure region. Hence, although the feature is flattened in  $X$  direction, see figure 10.26, since it is already present at the entrance of the high pressure region it tends to remain present with respect to the  $Y$  direction, see figure 10.27.

At this point the reader is reminded of the results presented for a larger dent, i.e. the tunneling effect downstream of the dent (figure 10.17). In fact what applies globally in the present situation of longitudinal waviness in that case applies locally, i.e. downstream of the dent: The tendency to preserve a variation of  $H$  in  $Y$  direction.

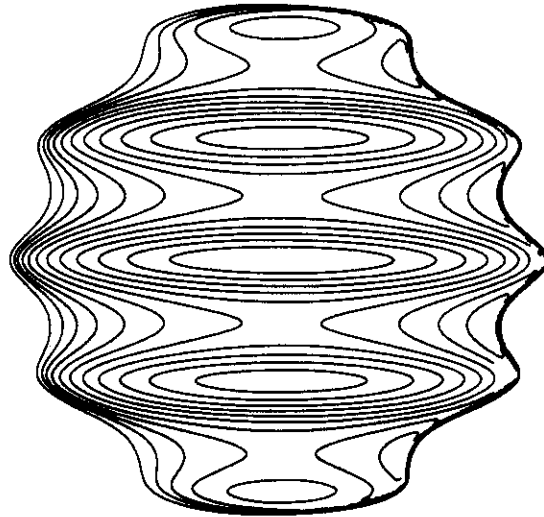


FIGURE 10.24:  $M = 1007.6$ ,  $L = 12.05$ . Pressure profile contour plot in the case of a longitudinal waviness with a wavelength of  $0.25\text{ mm}$  and an amplitude of  $0.25\text{ }\mu\text{m}$ ,  $\Delta P = 0.1$ .

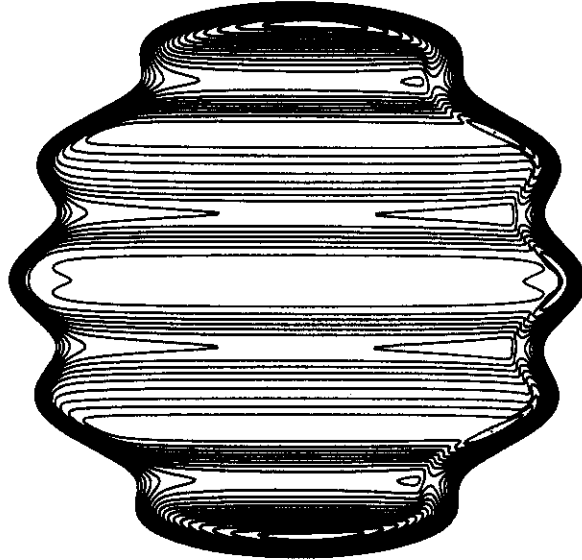


FIGURE 10.25:  $M = 1007.6$ ,  $L = 12.05$ . Film thickness contour plot in the case of longitudinal waviness with a wavelength of  $0.25\text{ mm}$  and an amplitude of  $0.25\text{ }\mu\text{m}$ ,  $\Delta H = 2.5 \cdot 10^{-3}$ .

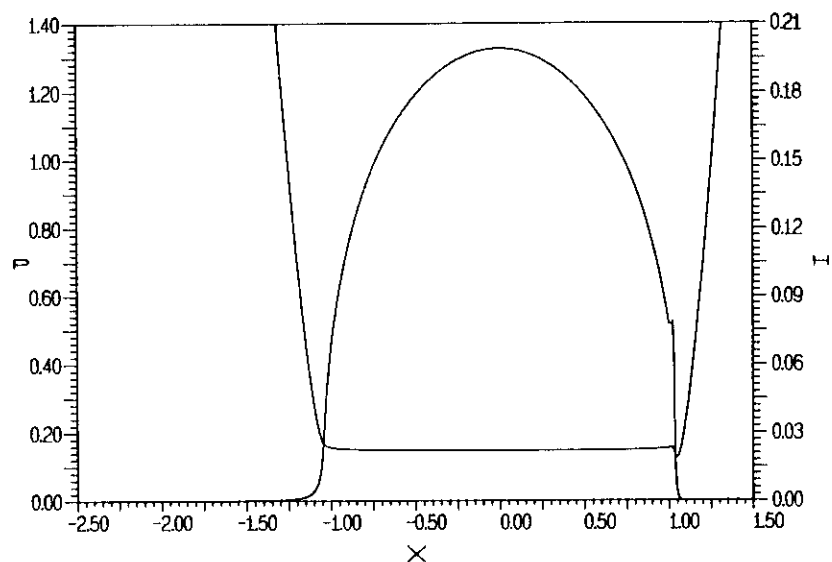


FIGURE 10.26:  $M = 1007.6$ ,  $L = 12.05$ . Pressure and film thickness on the center-line  $Y = 0$  as a function of  $X$  in the case of longitudinal waviness with a wavelength of  $0.25 \text{ mm}$  and an amplitude of  $0.25 \text{ } \mu\text{m}$ .

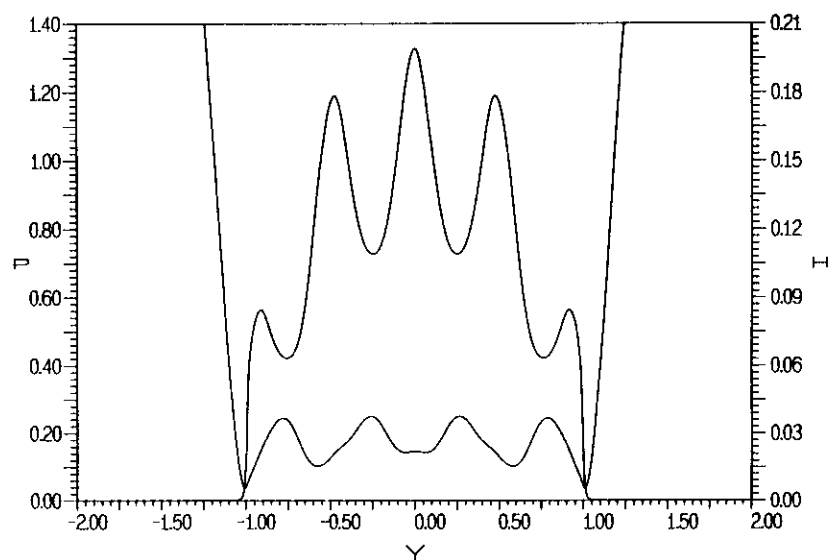


FIGURE 10.27:  $M = 1007.6$ ,  $L = 12.05$ . Pressure and film thickness on the line  $X = 0$  as a function of  $Y$  in the case of longitudinal waviness with a wavelength of  $0.25 \text{ mm}$  and an amplitude of  $0.25 \text{ } \mu\text{m}$ .

In addition to this global difference from the smooth surface solution also specific values such as the minimum and central film thickness differ significantly from the values in the case of smooth surfaces. Comparing figure 10.4 and figure 10.27 it is clear that in the present situation the minimum film thickness is about 50 % smaller. Considering the small value of the amplitude this effect is very large compared to the effects of a bump and indentation considered in the previous sections.

This large (decreasing) influence on the film thickness is explained by the fact that the waviness significantly affects the geometry in the center part of the inlet region. The pressure generation in this region determines the minimum and central film thickness. In the present case the central ridge causes an additional side leakage in this region which reduces the pressure generation compared to the smooth surface situation. Consequently the entire film thickness will be at a lower level and in particular, the minimum and central values are smaller than in the smooth surface situation. It can be expected that this effect will be stronger for larger ratio's  $A/W$ .

For the minimum film thickness there is an additional reason which in the present situation leads to a decrease compared to its smooth surface value. For the wavelength considered here a ridge is located at  $Y = \pm 1$ , i.e. in the region where the minimum film thickness occurs. Because on the one hand, in this region the pressures and consequently the elastic deformations are small and on the other hand, even if the pressures would have been large the ridge would be preserved this decreases the minimum film thickness.

Next the effect of the wavelength of the waviness on the characteristic results such as the maximum pressure, the minimum film thickness and the central film thickness is investigated. For the same value of the amplitude as considered above, i.e.  $0.25 \mu m$ , the maximum pressure relative to the smooth surface maximum pressure is displayed in figure 10.28. This figure shows that the maximum pressure increases with decreasing wavelength. For a given wavelength with increasing amplitude or, as considered here, for a given amplitude with decreasing wavelength the central ridge (only part of the smooth surface contact circle) carries an increasing part of the load and since the total dimensionless load is invariant this leads to an increasing maximum pressure when compared to the smooth surface situation. Although for different conditions, basically the same behaviour was observed by Lubrecht [L1].

Figure 10.29 and 10.30 show the associated film thickness results. Figure 10.30 shows that the central film thickness decreases with decreasing wavelength. This behaviour is explained by the aforementioned increasingly poor lubrication conditions (side leakage). In principle the same reasoning applies to the minimum film thickness although "superimposed" on this the tendency to decrease there is the effect of, depending on the wavelength, a ridge or valley being located at  $Y = \pm 1$ .



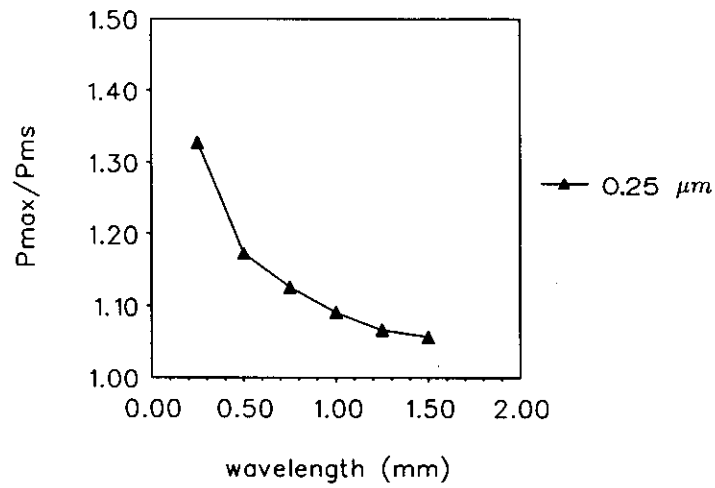


FIGURE 10.28:  $M = 1007.6$ ,  $L = 12.05$ . *Relative maximum pressure as a function of the wavelength in the case of longitudinal waviness with an amplitude of  $0.25 \mu\text{m}$  respectively.*

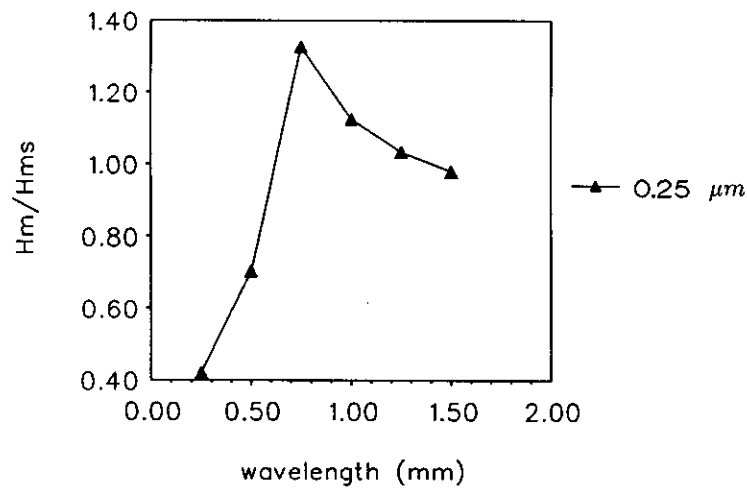


FIGURE 10.29:  $M = 1007.6$ ,  $L = 12.05$ . *Relative minimum film thickness as a function of the wavelength in the case of longitudinal waviness with an amplitude of  $0.25 \mu\text{m}$ .*

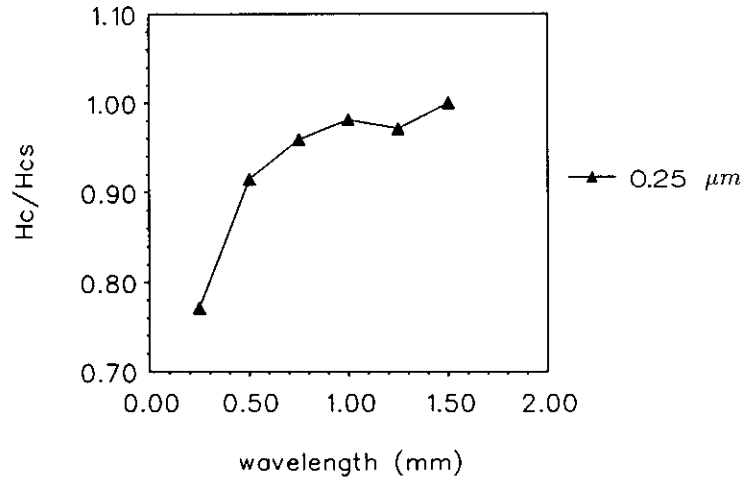


FIGURE 10.30:  $M = 1007.6$ ,  $L = 12.05$ . Relative central film thickness as a function of the wavelength in the case of longitudinal waviness with an amplitude of  $0.25 \mu m$ .

### 10.5.2 Transversal

In the case of transverse waviness  $\mathcal{R}$  is given by:

$$\mathcal{R}(X, Y) = \mathcal{A} \cos\left(\frac{2\pi X}{W}\right) \quad (10.4)$$

Hence, this kind of roughness has ridges perpendicular to the direction of rolling. As an example figure 10.31 and 10.32 show the contour plots of the calculated (dimensionless) pressure and the associated (dimensionless) film thickness for the present loading conditions in the case of a transverse waviness with a wavelength of  $0.375 mm$  and an amplitude of  $0.25 \mu m$  ( $W = 0.75$  and  $\mathcal{A} = 0.356$ ). In addition, figure 10.33 gives the pressure and film thickness as a function of  $X$  on the centerline of the contact ( $Y = 0$ ) and figure 10.34 shows a cross-section in the opposite direction, i.e. the pressure and film shape as a function of  $Y$  on the line  $X = 0$ .

From figure 10.31 and the cross-sections of the pressure profile presented in the figures 10.33 and figure 10.34 it can be concluded that, as is obvious, transverse waviness results in pressure oscillations with respect to the  $X$  direction. In addition, comparing the pressure curves in the figures 10.33 and 10.34 with the pressure on the same lines in the case of smooth surfaces, figure 10.3 and 10.4, shows that transverse waviness results in an increased maximum pressure.

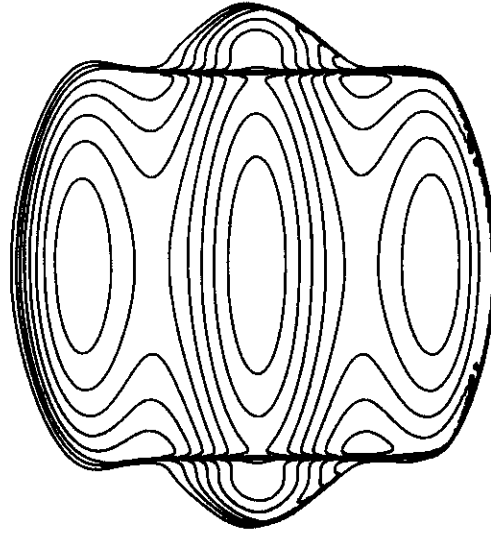


FIGURE 10.31:  $M = 1007.6$ ,  $L = 12.05$ . Pressure profile contour plot in the case of a transverse waviness with a wavelength of  $0.375\text{ mm}$  and an amplitude of  $0.25\text{ }\mu\text{m}$ ,  $\Delta P = 0.1$ .

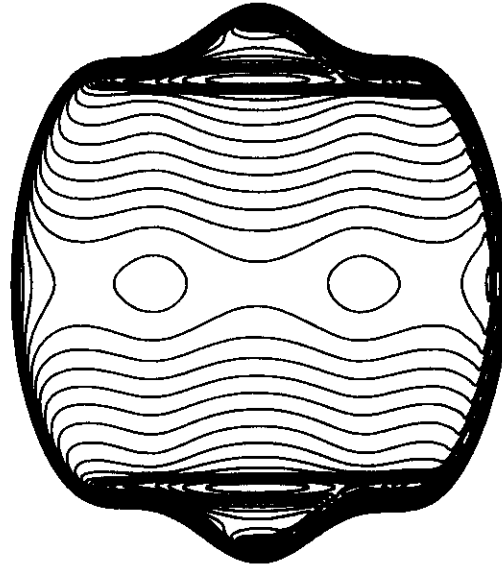


FIGURE 10.32:  $M = 1007.6$ ,  $L = 12.05$ . Film thickness contour plot in the case of a transverse waviness with a wavelength of  $0.375\text{ mm}$  and an amplitude of  $0.25\text{ }\mu\text{m}$ ,  $\Delta H = 7.5 \cdot 10^{-4}$ .

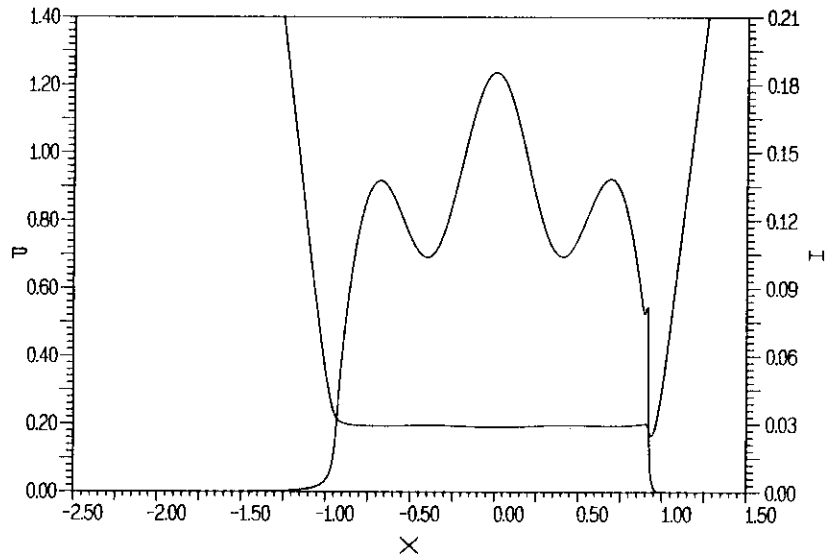


FIGURE 10.33:  $M = 1007.6$ ,  $L = 12.05$ . Pressure and film thickness on the centerline  $Y = 0$  as a function of  $X$  in the case of transverse waviness with a wavelength of  $0.375 \text{ mm}$  and an amplitude of  $0.25 \mu\text{m}$ .

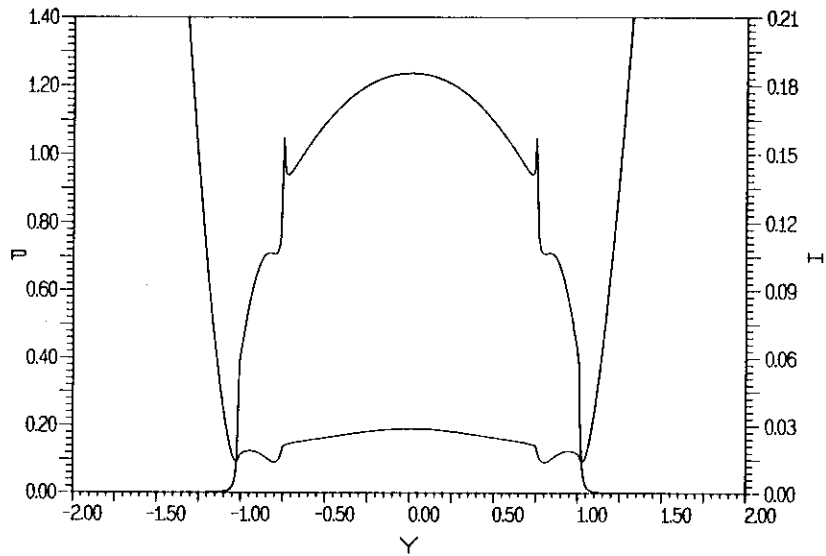


FIGURE 10.34:  $M = 1007.6$ ,  $L = 12.05$ . Pressure and film thickness on the line  $X = 0$  as a function of  $Y$  in the case of transverse waviness with a wavelength of  $0.375 \text{ mm}$  and an amplitude of  $0.25 \mu\text{m}$ .

With respect to the film thickness, figure 10.32 and figure 10.33 show that in the central region the waviness is nearly completely flattened. As far as the  $X$  direction is concerned this is obvious because of the high viscosities. In  $Y$  direction this applies only over a limited distance, i.e. the width of the part of the first ridge where the viscosities are high. The edges of this region are “felt” in the entire flow downstream of this ridge. This is clearly another demonstration of the characteristic 2-d EHL effects already mentioned several times before, i.e. the fact that Reynolds’ equation reduces to  $\partial(\bar{\rho}H)/\partial X \approx 0$ , which is a condition with respect to the  $X$  direction only.

From the figures 10.31 to 10.34 some other interesting features can be observed. In the smooth surface situation side leakage mainly occurs in the inlet region as was for example shown by Kweh et al. [K2], and to a small extend on the edge of the Hertzian contact circle. In the present situation the transverse ridges and furrows enhance the flow in  $Y$  direction near the sides of the contact. This results in some micro-EHL features in the pressure profile and film thickness in the region  $Y = \pm 1$ . This is particularly clear from figure 10.34, i.e. the pressure spikes on both sides and the nip in the film thickness. The very same effect can also be observed in the results for transverse waviness presented by Lubrecht [L1, figure 10.33].

With respect to specific values, comparing figure 10.34 with figure 10.4 it can be seen that the transverse waviness slightly enhances the film thickness. This can be explained by the fact that the transverse waviness affects the global inlet geometry in such a way that the pressure generation is enhanced.

The aforementioned effects are much stronger if the wavelength is decreased. This can be observed from the following figures. Figure 10.35 and figure 10.36 show the contour plot of the pressure profile and film thickness in the case of a waviness with the same amplitude of  $0.25 \mu\text{m}$  ( $\mathcal{A} = 0.356$ ) but with a smaller wavelength of  $0.25 \text{ mm}$  ( $\mathcal{W} = 0.5$ ). In addition figure 10.37 and 10.38 show the pressure and film thickness as a function of  $X$  at the centerline of the contact ( $Y = 0$ ) and as a function of  $Y$  at the line  $X = 0$  respectively. In this case the features described above are more pronounced. For example, in this case both the first as well as the second ridge are felt over a specific width in  $Y$  direction throughout the entire downstream region.

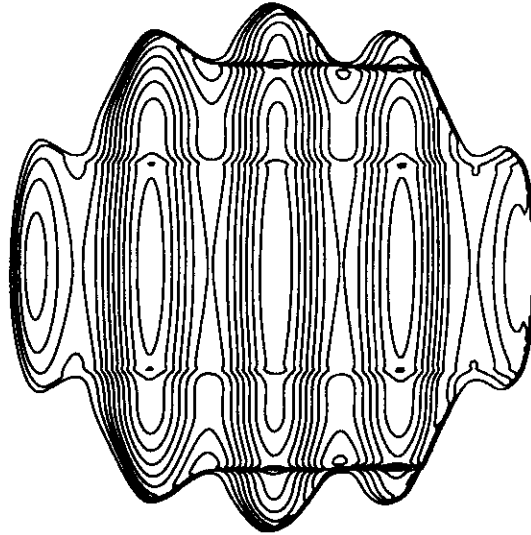


FIGURE 10.35:  $M = 1007.6$ ,  $L = 12.05$ . Pressure profile contour plot in the case of a transverse waviness with a wavelength of  $0.25\text{ mm}$  and an amplitude of  $0.25\text{ }\mu\text{m}$ ,  $\Delta P = 0.1$ .

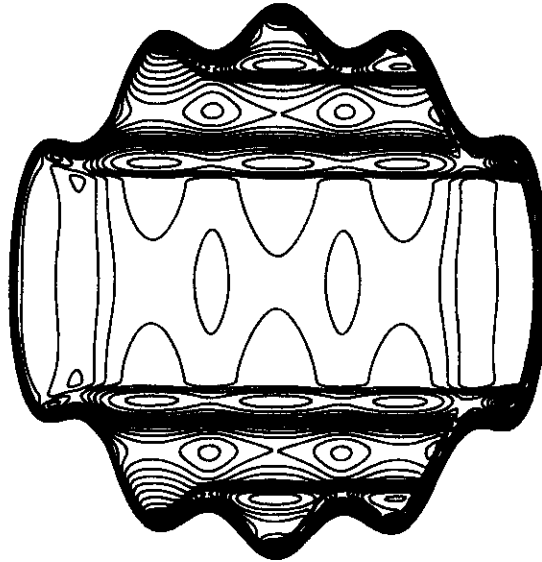


FIGURE 10.36:  $M = 1007.6$ ,  $L = 12.05$ . Film thickness contour plot in the case of a transverse waviness with a wavelength of  $0.25\text{ mm}$  and an amplitude of  $0.25\text{ }\mu\text{m}$ ,  $\Delta H = 7.5 \cdot 10^{-4}$ .

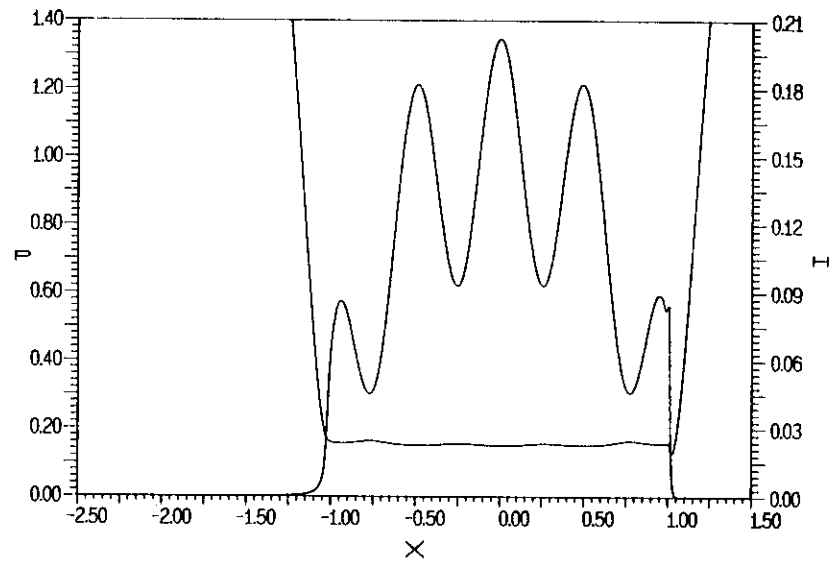


FIGURE 10.37:  $M = 1007.6$ ,  $L = 12.05$ . Pressure and film thickness on the centerline  $Y = 0$  as a function of  $X$  in the case of transverse waviness with a wavelength of  $0.25 \text{ mm}$  and an amplitude of  $0.25 \mu\text{m}$ .

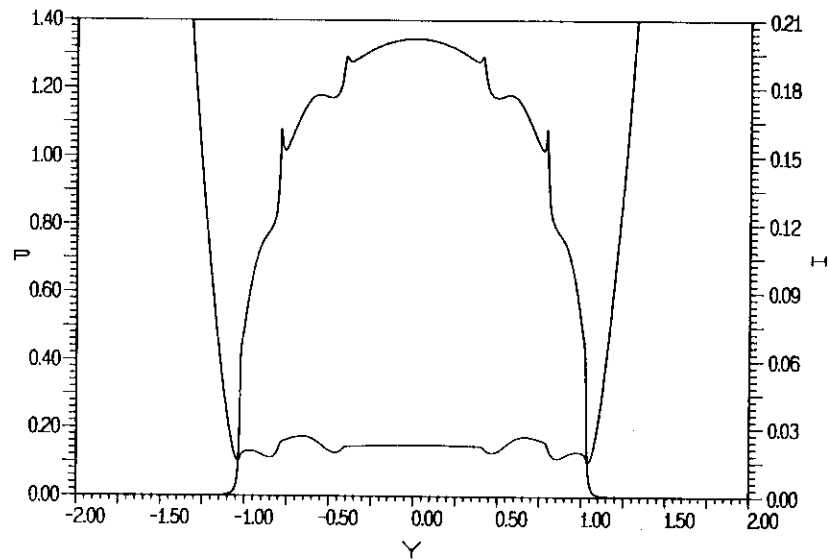


FIGURE 10.38:  $M = 1007.6$ ,  $L = 12.05$ . Pressure and film thickness on the line  $X = 0$  as a function of  $Y$  in the case of transverse waviness with a wavelength of  $0.25 \text{ mm}$  and an amplitude of  $0.25 \mu\text{m}$ .

To conclude this section the effect of the wavelength of the waviness on the maximum pressure, the minimum film thickness and the central film thickness has been investigated. First figure 10.39 shows the maximum pressure relative to the smooth surface situation as a function of the wavelength in the case of a fixed amplitude of  $0.25 \mu\text{m}$ . This figure shows that the maximum pressure is roughly inversely proportional to the wavelength. With increasing wavelength the effect gradually decreases and eventually the smooth surface solution will return.

Figure 10.40 shows that the waviness results in an increased minimum film thickness. For the larger wavelengths, when there is only a single ridge in the Hertzian contact region, this can be explained by the fact that the contact in fact has changed from a circular contact to an elliptic contact with the major axis perpendicular to the direction of rolling. This situation is much more favourable for the pressure generation in the inlet and consequently the film thickness is increased when compared to the smooth surface situation. For the smaller wavelengths basically the same tendency is observed. However with increasing number of ridges in the contact region "superimposed" on the tendency of the minimum film thickness to increase there is an effect of a ridge or valley being located in the inlet region.

Finally figure 10.41 shows the central film thickness relative to the smooth surface value as a function of the wavelength. For the larger wavelengths also the aforementioned increasing tendency can be observed. However, for the small wavelengths the result seem to depend strongly on the number of ridges in the contact region.

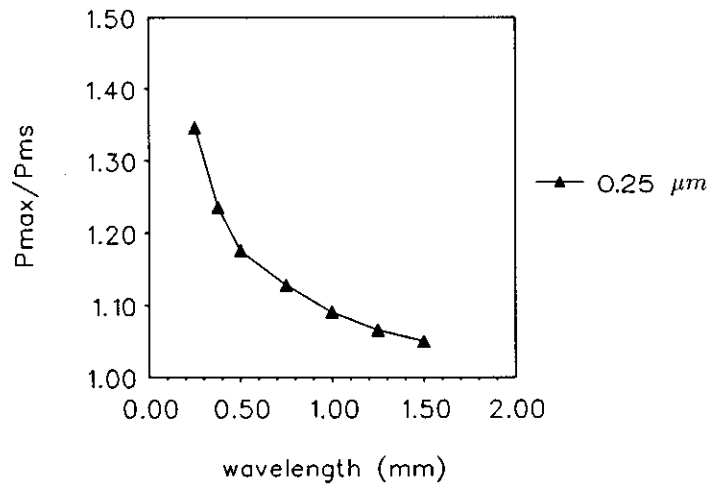


FIGURE 10.39:  $M = 1007.6$ ,  $L = 12.05$ . Relative maximum pressure as a function of the wavelength in the case of transverse waviness with an amplitude of  $0.25 \mu\text{m}$ .



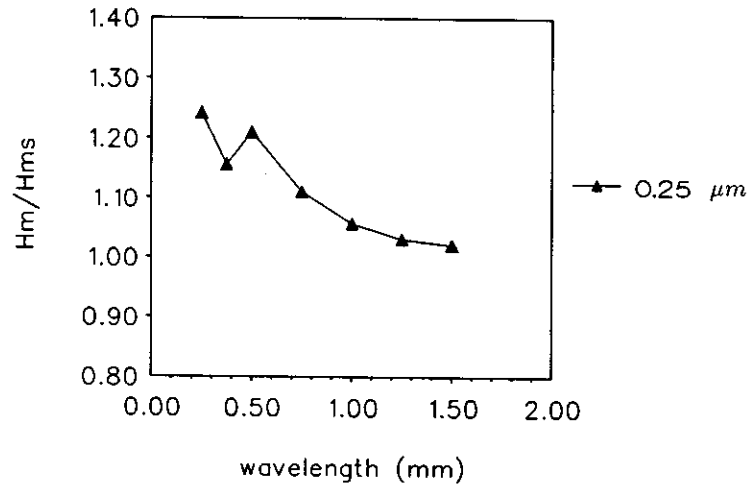


FIGURE 10.40:  $M = 1007.6$ ,  $L = 12.05$ . Relative minimum film thickness as a function of the wavelength in the case of transverse waviness with an amplitude of  $0.25 \mu m$ .

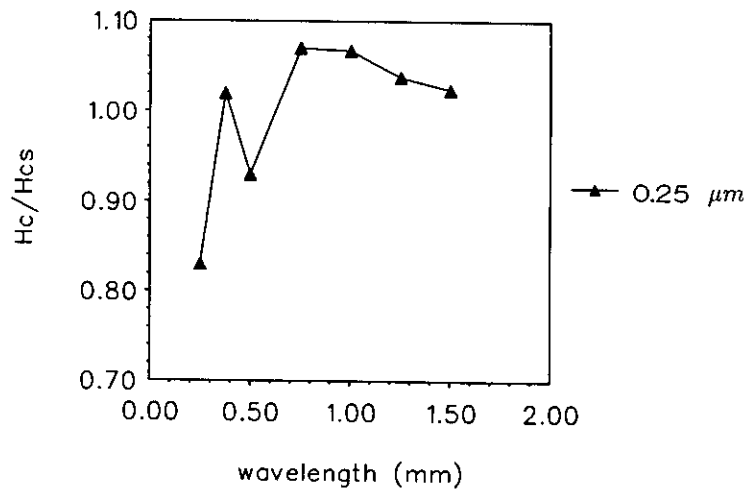


FIGURE 10.41:  $M = 1007.6$ ,  $L = 12.05$ . Relative central film thickness as a function of the wavelength in the case of transverse waviness with an amplitude of  $0.25 \mu m$ .

## 10.6 Conclusion

The algorithm developed for the solution of the steady state EHL circular contact problem was tested out to a more complex point contact situation, i.e. to study the effects of some large scale surface features on the pressure profile and film thickness. For a specific load situation with a maximum Hertzian pressure of 2 GPa the effect of a bump, an indentation, as well as the effects of harmonic longitudinal and transverse waviness have been studied.

Although the study was not as extensive as the study of surface features presented in chapter 7 for the line contact problem the results showed that the additional dimension in the point contact situation leads to some interesting effects. These effects are characteristic for the physical mathematical model of the EHL contact situation used in this thesis, i.e. a Newtonian lubricant behaviour with a viscosity that increases exponentially with pressure. For high loads in the Hertzian contact region Reynolds' equation reduces to  $\partial(\rho H)/\partial X \approx 0$ . This results in a tendency to flatten a feature as was also found for the line contact problem. However, in the point contact problem this only applies with respect to the  $X$  direction, i.e. the direction of rolling. It would be interesting to see to what extent the same effects will be found if non-Newtonian lubricant behaviour is accounted for.

---

## References

- [B1] Barragan de Ling, FdM, Evans, H.P., and Snidle, R.W., 1989, "Micro-elastohydrodynamic lubrication of circumferentially finished rollers: The influence of temperature and roughness," *ASME JOT*, **11**, 730-736.
- [B2] Bush, A.W., and Skinner, P.H., (1982), "Surface roughness effects in point contact elastohydrodynamic lubrication," *WEAR*, **83**, 285-301.
- [C1] Cusano, C., and Wedeven, L.D., 1980, "Elastohydrodynamic film thickness measurements of artificially produced nonsmooth surfaces," *ASLE transactions*, **24**, 1, 1-14.
- [C2] Cusano, C., and Wedeven, L.D., 1982, "The effects of artificially-produced defects on the film thickness distribution in sliding EHD point contacts," *ASME JOT*, **104**, 365-375.
- [J1] Jackson, A., and Cameron, A., 1974, "An interferometric study of the EHL of rough surfaces," *ASLE transactions*, **19**, 1, 50-60.
- [J2] Jeng, Yeau-Ren and Hamrock, B.J., 1988, "The effect of surface roughness on elastohydrodynamically lubricated point contact," *ASLE transactions*, **30**, 4, 531-538.
- [K1] Kaneta, M., and Cameron, A., (1980), "Effects of asperities in elastohydrodynamic lubrication," *ASME JOT*, **102**, 374-389.
- [K2] Kweh, C.C., Evans, H.P., and Snidle, R.W., 1989, "Micro-elastohydrodynamic lubrication of an elliptical contact with transverse and three-dimensional sinusoidal roughness," *ASME JOT*, **111**, 577-583.
- [S1] Seabra, J., and Berthe, D., 1989, "Elastohydrodynamic point contacts part 2: Influence of surface speeds, surface waviness and load on the contact behaviour," *Wear*, **130**, 319-335.
- [L1] 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.
- [L2] Lubrecht, A.A., ten Napel, W.E., and Bosma, R., 1988, "The influence of longitudinal and transverse roughness on the elastohydrodynamic lubrication of circular contacts," *ASME JOT*, **110**, 421-426.
- [W1] Wedeven, L.D., 1977, "Influence of debris dent on EHD lubrication," *ASLE transactions*, **21**, 1, 41-52.
- [W2] Wedeven, L.D., and Cusano, C., 1978, "Elastohydrodynamic film thickness measurements of artificially produced surface dents and grooves," *ASLE transactions*, **22**, 4, 369-381.

- [Z1] Zhu, D., and Cheng, H.S., 1988, "Effect of surface roughness on the point contact EHL," *ASME JOT*, 110, 32-37.
-

## Recommendations for future research

The stability and the low complexity of the algorithms for the solution of the line and point contact problem introduced in this work offer great possibilities for the simulation of EHL contact situations. In that respect the subjects addressed in the chapters 6 through 10 are only a few items of an extensive list.

With respect to the situation of perfectly smooth surfaces an extension of the algorithm developed for the circular contact to elliptic contact situations is recommended.

With respect to surface features extension of the point contact algorithm to transient situations to simulate the overrolling of a feature is needed.

✓ Further subjects for future research in this direction are: The investigation of the influence of different features than the ones considered in the present study. A more extensive study of the effects of surface roughness by performing calculations for different roughness profiles. The extension of the study of roughness effects to point contacts and eventually extending it to transient situations for both line and point contact. However, such further studies, would greatly benefit from an extension of the model to properly account for pressure generation from a cavitated region.

An interesting topic for future study related to these time-dependent surface-feature or rough-surface calculations is the inclusion of a solid-to-solid force which would enable simulation of contacts operating in the “mixed lubrication” regime.

The Dowson and Higginson’s density pressure relation used in this work limits the compressibility of the lubricant. Therefore the use of a more advanced model is recommended. Such a model is for example proposed by Jacobson and Vinet [J1].

The current model is not suited for prediction of the friction in the contact. For that purpose including non-Newtonian lubricant behaviour is essential. Restricted to the line contact problem this subject can be investigated to a limited extent with the present model, see for example the study carried out by Hulselmans [H1] using the algorithm presented in this thesis. In the case of point contacts the situation is

much more complex and in principle the study of non-Newtonian lubricant behaviour requires the development of an alternative physical mathematical model.

Another interesting topic is the inclusion of the energy equation since it is known from literature that the heat generation in the inlet zone in the case of sliding may result in a considerable decrease of the film thickness compared to the isothermal situation considered in this thesis. Furthermore, taking into account thermal effects is essential to obtain reasonable predictions of the friction in EHL contacts.

---

## References

- [J1] **Jacobson, B.O., and Vinet, P.**, 1987, "A model for the influence of pressure on the bulk modulus and the influence of temperature on the solidification pressure for liquid lubricants," *ASME JOT*, **109**, 709-714.
- [H1] **Hulselmans, R.**, 1990, "Elastohydrodynamisch gesmeerd lijncontact: onderzoek naar de effecten van niet Newtons vloeistofgedrag op de drukverdeling en filmdikte in het contact," *Masters thesis* (in dutch), University of Twente, Enschede, The Netherlands.

---



## Acknowledgement

The author is indebted to many people who in one way or another contributed to the present work.

I would like to thank the members of the tribology group of the faculty of Mechanical Engineering of the University of Twente for their cooperation and support over the past four years and for creating a pleasant atmosphere to work: prof. A.W.J. de Gee, ir W.E. ten Napel, ir H. Moes, ir P.H. Vroegop, ir D.J. Ligtink, ing. H. van de Meene, mr. L.J. de Boer and mrs. D.B. Vrieze Zimmerman van Woesik. Special thanks to my fellow PhD. students and friends: ir P.B.Y. ten Hoeve for his guidance through the complex world of "C" and computer graphics, ir P.M. Iugt and ir R. Meerwijk for their continuous interest and stimulating enthusiasm. I'd also like to thank ir F. Janssen and the students that worked on their M.Sc. thesis with me, for their support and stimulation: ir J.F.M. Velthuis, ir R. Hulselmans, and all other students of the tribology group. In addition, I'd like to thank prof. ir R. Bosma for introducing me into tribology.

I would like to thank the members of the "promotie-begeleidings" committee: prof. ir A. W. J. de Gee, prof. dr. ir H. Tjeldeman, prof. dr. ir L. van Wijngaarden, ir W.E. ten Napel, dr. ir D.J. Schipper (TUD) for their suggestions and helpful discussions.

From the faculty of Mathematics I'd like to thank prof. dr. ir E.W.C. van Groesen, drs. F.P.H. van Beckum, and also dr. ir R.W.C.P. Verstappen for their interest in the work.

Special thanks to dr. ir A.A. Lubrecht for teaching me multigrid and multilevel multi-integration. In particular I'd like to thank him for laying the foundation for the present work and for his continuous support, stimulation, and enthusiasm.

I would like to thank prof. dr. A. Brandt of the Weizmann Institute of Science for educating me into multigrid ("if you can not solve a problem there is an easier problem that you can not solve. Your first task is to find that problem.") and for all the helpful discussions and suggestions. I'd also like to thank his coworkers and students, particularly I. Yavneh, V. Mikulinsky and D. Sidilkover for their interest in the work and the many interesting discussions.

This research was sponsored by the SKF Engineering & Research Centre B.V., Nieuwegein, the Netherlands. From this center I'd like to thank prof. dr. E. Ioannides, prof. dr. B.O. Jacobsen, dr. J.H. Tripp, and dr.ir A.A. Lubrecht for their interest in the work and their helpful suggestions with respect to situations of practical relevance.

Part of the research was carried out during visits to professor Brandt at the Weizmann institute of Science, Rehovot, Israel and the University of Colorado at Denver, Denver, USA. I'd like to thank the faculty of Mechanical Engineering of the University of Twente, and the dutch foundations KIVI, and NWO for sponsoring these visits.

Finally I'd like to apologize to fellow users of the computer system for the inconvenience caused by my extensive use of computing time and disk quota and I'd like to thank the "super-user" ing. R. Klumpert, for maintaining the peace on the system.



## Appendix A

### Dimensionless parameters

#### A.1 Moes and Hertzian dimensionless parameters

This section presents the dimensionless equations of the steady state isothermal EHL line and circular contact in terms of the Moes [M1] and Hertzian dimensionless parameters.

##### A.1.1 Line contact

In the stationary situation, using the Barus viscosity pressure relation the governing equations are:

Reynolds equation:

$$\frac{d}{dx} \left( \frac{\rho h^3}{\eta} \frac{dp}{dx} \right) - 6u_s \frac{d(\rho h)}{dx} = 0 \quad (\text{A.1})$$

with the cavitation condition  $p \geq 0$  in the domain and  $p = 0$  at the boundaries. Furthermore  $\eta$  is given by:

$$\eta = \eta_0 \exp(\alpha p) \quad (\text{A.2})$$

and, since an incompressible lubricant is presumed in this section,  $\rho = \rho_0$ , the density at ambient pressure.

The equation for the film thickness:

$$h(x) = h_{00} + \frac{x^2}{2R} - \frac{4}{\pi E'} \int_{-\infty}^{\infty} \ln |x - x'| p(x') dx' \quad (\text{A.3})$$

where  $h_{00}$  is a constant.

and the force balance condition:

$$\int_{-\infty}^{\infty} p(x) dx = w \quad (\text{A.4})$$

where:  $w$  = external load per unit width

The minimum film thickness, as an example of a result, is defined according to:

$$h_{\min} = \min (h) \quad (\text{A.5})$$

From the equations (A.1) to (A.5) it can be concluded that the minimum film thickness is a function of  $\eta_0$ ,  $\alpha$ ,  $u_s$ ,  $E'$ ,  $R$ , and  $w$ . Using a dimensional analysis entitled *optimum similarity analysis*, see Moes [M2], this set of six can be reduced significantly by introducing the following non-dimensional variables and groups:

$$\begin{aligned} \bar{x} &= \frac{x}{R} \left( \frac{E' R}{\eta_0 u_s} \right)^{1/4} & \bar{p} &= \frac{p}{E'} \left( \frac{E' R}{\eta_0 u_s} \right)^{1/4} \\ \bar{h} &= \frac{h}{R} \left( \frac{E' R}{\eta_0 u_s} \right)^{1/2} & M &= \frac{w}{E' R} \left( \frac{E' R}{\eta_0 u_s} \right)^{1/2} \\ L &= \alpha E' \left( \frac{E' R}{\eta_0 u_s} \right)^{-1/4} & H_{\min} &= \frac{h}{R} \left( \frac{E' R}{\eta_0 u_s} \right)^{1/2} \\ \bar{\eta} &= \frac{\eta}{\eta_0} & \bar{\rho} &= \frac{\rho}{\rho_0} \end{aligned}$$

Substitution of these groups and variables in the equations (A.1) to (A.5) using  $\bar{p} \equiv 1$  yields:

$$\frac{d}{d\bar{x}} \left( \frac{\bar{h}^3}{\bar{\eta}} \frac{d\bar{p}}{d\bar{x}} \right) - 6 \frac{d\bar{h}}{d\bar{x}} = 0 \quad (\text{A.6})$$

with the cavitation condition  $\bar{p} \geq 0$  in the domain and  $\bar{p} = 0$  at the boundaries.

$$\bar{\eta} = \exp(L \bar{p}) \quad (\text{A.7})$$

$$\bar{h}(\bar{x}) = \bar{h}_{00} + \frac{\bar{x}^2}{2} - \frac{4}{\pi} \int_{-\infty}^{\infty} \ln |\bar{x} - \bar{x}'| \bar{p}(\bar{x}') d\bar{x}' \quad (\text{A.8})$$

where  $\bar{h}_{00}$  is a constant.

$$\int_{-\infty}^{\infty} \bar{p}(\bar{x}) d\bar{x} = M \quad (\text{A.9})$$

$$H_{\min} = \min (\bar{h}) \quad (\text{A.10})$$

The equations expressed in the new variables show that from the original six parameters only two independent groups remain i.e.  $M$ , and  $L$ . Hence, the substitution shows that the problem can be reduced to a two parameter problem. This allows representation of the results in a very convenient way, particularly for practical purposes (designers). For example, each characteristic parameter, e.g. the dimensionless minimum film thickness,  $H_{min}$ , is a function of two parameters only and all possible solutions can be presented in one chart, i.e. a diagram giving  $H_{min}$  as a function of  $M$  and  $L$ .

Since it is our intention to solve the problem numerically it is convenient to normalize the equations, i.e. to ensure that all variables will be of order 1 to avoid unnecessary loss of accuracy due to round-off errors. This can be done straightforwardly. With increasing load the pressure in the lubricant film will approximate the well known Hertzian dry contact pressure. Therefore it seems reasonable to scale the EHL variables onto the Hertzian parameters:

$$\begin{aligned} b &= \text{half the contact width} \\ p_h &= \text{the maximum Hertzian pressure} \\ h_h &= \text{a characteristic deformation } h_h = b^2/R \end{aligned}$$

where

$$p_h = \frac{2w}{\pi b} \quad (\text{A.11})$$

and

$$b = \sqrt{\frac{8wR}{\pi E'}} \quad (\text{A.12})$$

Defining:

$$\begin{aligned} X &= x/b \\ P &= p/p_h \\ H &= h/h_h \\ \bar{\rho} &= \rho/\rho_0 \\ \bar{\eta} &= \eta/\eta_0 \end{aligned}$$

then yields the dimensionless equations as used in the present work:

$$\frac{d}{dX} \left( \frac{H^3}{\bar{\eta}} \frac{dP}{dX} \right) - \lambda \frac{dH}{dX} = 0 \quad (\text{A.13})$$

with the cavitation condition  $P \geq 0$  in the domain and  $P = 0$  at the boundaries.

$$\bar{\eta} = \exp(\bar{\alpha} P) \quad (\text{A.14})$$

$$H(X) = H_{00} + \frac{X^2}{2} - \frac{1}{\pi} \int_{-\infty}^{\infty} \ln |(X - X')| P(X') dX' \quad (\text{A.15})$$

where  $H_{00}$  is a constant.

$$\int_{-\infty}^{\infty} P(X) dX = \frac{\pi}{2} \quad (\text{A.16})$$

$$H_m = \min (H) \quad (\text{A.17})$$

$\bar{\alpha}$  and  $\lambda$  are the two governing parameters of the scaled dimensionless problem:

$$\bar{\alpha} = \alpha p_h \quad (\text{A.18})$$

$$\lambda = \frac{6\eta_0 u_s R^2}{b^3 p_h} \quad (\text{A.19})$$

and are related to the Moes parameters  $M$  and  $L$  as follows:

$$\bar{\alpha} = L \sqrt{\frac{M}{2\pi}} \quad (\text{A.20})$$

$$\lambda = \frac{3\pi^2}{8 M^2} \quad (\text{A.21})$$

Furthermore,  $H_m$ , the minimum film thickness in terms of the scaled dimensionless problem is related to the Moes dimensionless film thickness parameter  $H_{min}$  according to:

$$H_{min} = \frac{8}{\pi} M H_m \quad (\text{A.22})$$

Note that, obviously, the scaling does not introduce additional parameters, it still remains a two-parameter problem.

### A.1.2 Circular contact

With some minor changes the approach outlined before for the line contact problem also applies to the circular contact problem. When the Barus equation is used in the case of an incompressible lubricant the equations for the isothermal steady state situation are:

Reynolds:

$$\frac{\partial}{\partial x} \left( \frac{\rho h^3}{\eta} \frac{\partial p}{\partial x} \right) + \frac{\partial}{\partial y} \left( \frac{\rho h^3}{\eta} \frac{\partial p}{\partial y} \right) - 6u_s \frac{\partial(\rho h)}{\partial x} = 0 \quad (\text{A.23})$$


---

with the cavitation condition  $p \geq 0$  in the domain and  $p = 0$  at the boundaries. Furthermore,  $\eta$  is given by:

$$\eta = \eta_0 \exp(\alpha p) \quad (\text{A.24})$$

and  $\rho = \rho_0$  since an incompressible lubricant is presumed.

The equation for the film thickness:

$$h(x, y) = h_{00} + \frac{x^2}{2R_x} + \frac{y^2}{2R_y} + \frac{2}{\pi E'} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{p(x', y') dx' dy'}{\sqrt{(x - x')^2 + (y - y')^2}} \quad (\text{A.25})$$

with the integration constant  $h_{00}$ .

The force balance equation:

$$-\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} p(x, y) dx dy = F \quad (\text{A.26})$$

Finally, the obvious definition for the minimum film thickness:

$$h_{\min} = \min(h) \quad (\text{A.27})$$

Hence, in this formulation the minimum film thickness  $h_{\min}$  is a function of the six parameters  $\eta_0$ ,  $\alpha$ ,  $u_s$ ,  $F$ ,  $E'$ , and  $R_x$ . From the dimensional analysis according to Moes, see [M2], the following dimensionless groups and variables to describe the problem can be obtained:

$$\begin{aligned} \bar{x} &= \frac{x}{R_x} \left( \frac{E' R_x}{\eta_0 u_s} \right)^{1/4} & \bar{y} &= \frac{y}{R_x} \left( \frac{E' R_x}{\eta_0 u_s} \right)^{1/4} \\ \bar{p} &= \frac{p}{E'} \left( \frac{E' R_x}{\eta_0 u_s} \right)^{1/4} & \bar{h} &= \frac{h}{R_x} \left( \frac{E' R_x}{\eta_0 u_s} \right)^{1/2} \\ L &= \alpha E' \left( \frac{\eta_0 u_s}{E' R_x} \right)^{1/4} & M &= \frac{F}{E' R_x^2} \left( \frac{E' R_x}{\eta_0 u_s} \right)^{3/4} \\ \bar{\eta} &= \frac{\eta}{\eta_0} & \bar{\rho} &= \frac{\rho}{\rho_0} \end{aligned}$$

Rewriting the equations in these new variables, using  $\bar{\rho} = 1$ , gives:

$$\frac{\partial}{\partial \bar{x}} \left( \frac{\bar{h}^3}{\bar{\eta}} \frac{\partial \bar{p}}{\partial \bar{x}} \right) + \frac{\partial}{\partial \bar{y}} \left( \frac{\bar{h}^3}{\bar{\eta}} \frac{\partial \bar{p}}{\partial \bar{y}} \right) - 6 \frac{\partial \bar{h}}{\partial \bar{x}} = 0 \quad (\text{A.28})$$

$$\bar{\eta} = \exp(L \bar{p}) \quad (\text{A.29})$$

$$\bar{h}(\bar{x}, \bar{y}) = \bar{h}_{00} + \frac{\bar{x}^2}{2} + \frac{\bar{y}^2}{2} + \frac{2}{\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{p(\bar{x}', \bar{y}') d\bar{x}' d\bar{y}'}{\sqrt{(\bar{x} - \bar{x}')^2 + (\bar{y} - \bar{y}')^2}} \quad (\text{A.30})$$

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \bar{p}(\bar{x}, \bar{y}) d\bar{x} d\bar{y} = M \quad (\text{A.31})$$

$$H_{min} = \min(\bar{h}) \quad (\text{A.32})$$

Hence, from the equations expressed in terms of these new variables it appears that there are only two independent parameters, i.e. a load parameter  $M$ , and a materials parameter  $L$ . Consequently, the dimensionless minimum film thickness  $H_{min}$  is a function of  $M$  and  $L$  only, allowing the presentation of all possible load situations in one diagram  $H_{min}(M, L)$  which is particularly convenient from a designers point of view.

However this formulation of the problem is not used for numerical solution. For that purpose the equations are normalized using the Hertzian dry contact parameters:

$$\begin{aligned} a &= \text{half the contact width} \\ p_h &= \text{the maximum Hertzian pressure} \\ h_h &= \text{the maximum Hertzian elastic deformation } h_h = a^2/R_x \end{aligned}$$

where

$$p_h = \frac{3F}{2\pi a^2} \quad (\text{A.33})$$

and

$$a = \left( \frac{3FR_x}{2E'} \right)^{1/3} \quad (\text{A.34})$$

Defining:

$$\begin{aligned} X &= x/a \\ Y &= y/a \\ P &= p/p_h \\ H &= h/h_h \\ \bar{\rho} &= \rho/\rho_0 \\ \bar{\eta} &= \eta/\eta_0 \end{aligned}$$

then yields the dimensionless equations as used in the present work:

$$\frac{\partial}{\partial X} \left( \frac{H^3}{\bar{\eta}} \frac{\partial P}{\partial X} \right) + \frac{\partial}{\partial Y} \left( \frac{H^3}{\bar{\eta}} \frac{\partial P}{\partial Y} \right) - \lambda \frac{\partial H}{\partial X} = 0 \quad (\text{A.35})$$



$$\bar{\eta} = \exp(\bar{\alpha} P) \quad (\text{A.36})$$

$$H(X, Y) = H_{00} + \frac{X^2}{2} + \frac{Y^2}{2} + \frac{2}{\pi^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{P(X', Y') dX' dY'}{\sqrt{(X - X')^2 + (Y - Y')^2}} \quad (\text{A.37})$$

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} P(X, Y) dX dY = \frac{2\pi}{3} \quad (\text{A.38})$$

$$H_m = \min(H) \quad (\text{A.39})$$

This normalization gives an alternative set of two independent parameters, e.g.  $\bar{\alpha}$  and  $\lambda$ . These parameters are defined according to:

$$\bar{\alpha} = \alpha p_h \quad (\text{A.40})$$

and

$$\lambda = \frac{6\eta_0 u_s R_x^2}{a^3 p_h} \quad (\text{A.41})$$

and are related to  $M$  and  $L$  as follows:

$$\bar{\alpha} = \frac{L}{\pi} \left( \frac{3M}{2} \right)^{1/3} \quad (\text{A.42})$$

$$\lambda = \left( \frac{128 \pi^3}{3 M^4} \right)^{1/3} \quad (\text{A.43})$$

Furthermore,  $H_m$ , the minimum film thickness in terms of the scaled variables, is related to the Moes dimensionless minimum film thickness parameter  $H_{min}$  according to:

$$H_{min} = H_m \sqrt{\frac{6\pi}{\lambda}} \quad (\text{A.44})$$

## A.2 Different sets of dimensionless parameters

### A.2.1 Line contact

Dowson and Higginson [D1] introduced a set of three parameters to characterize each load situation:

$$W = \frac{w}{E'R} \quad (\text{A.45})$$

$$G = \alpha E' \quad (\text{A.46})$$

and:

$$U = \frac{1}{2} \frac{\eta_0 u_s}{E' R} \quad (\text{A.47})$$

with a dimensionless film thickness parameter defined according to:

$$H'_m = \frac{h_{min}}{R} \quad (\text{A.48})$$

Apart from the fact that the use of three parameters does not allow presentation of the results in simple diagrams, one of the parameters is redundant. Nevertheless, these parameters are widely used. Hence, for reasons of comparison in particular load situations, in addition to the values of  $M$  and  $L$  characterizing the case, also values of  $W$ ,  $U$ , and  $G$  describing the situation are given in this thesis. Since one of the latter three is redundant, it can be fixed without loss of generality. In this work generally  $U$  will be chosen  $1.0 \cdot 10^{-11}$ . Furthermore, the relation between  $M$  and  $L$  and  $W$  and  $G$  is given by:

$$M = W (2U)^{-1/2} \quad (\text{A.49})$$

$$L = G (2U)^{1/4} \quad (\text{A.50})$$

whereas the relation between the Dowson and Higginson dimensionless minimum film thickness parameter  $H'_m$  and  $H_{min}$  reads:

$$H_{min} = H'_m (2U)^{-1/2} \quad (\text{A.51})$$

### A.2.2 Circular contact

Hamrock and Dowson [H1] use the following set of three parameters to describe a circular contact load condition:

$$W = \frac{w}{E' R_x^2} \quad (\text{A.52})$$

$$G = \alpha E' \quad (\text{A.53})$$

and:

$$U = \frac{1}{2} \frac{\eta_0 u_s}{E' R} \quad (\text{A.54})$$

with a dimensionless minimum film thickness defined as:

$$H_m^{hd} = \frac{h_{min}}{R_x} \quad (A.55)$$

As in the line contact case the use of three parameters to characterize a load situation does not allow presentation of the results in simple diagrams. Moreover since one of the parameters is redundant this set is not used in the present work. Nevertheless, for reasons of comparison with results presented by other authors in particular load situations, in addition to the values of  $M$  and  $L$  characterizing the case, also values of  $W$ ,  $U$ , and  $G$  describing the situation are given in this thesis. Since one of the latter three is redundant, it can be fixed without loss of generality and in this work generally  $U$  is chosen  $1.0 \cdot 10^{-11}$ . The relation between the other two, i.e.  $W$  and  $G$ , and  $M$  and  $L$  is given by:

$$M = W (2U)^{-3/4} \quad (A.56)$$

$$L = G (2U)^{1/4} \quad (A.57)$$

whereas their dimensionless film thickness is related to the Moes dimensionless film thickness parameter according to:

$$H_{min} = H_m^{hd} (2U)^{-1/2} \quad (A.58)$$

## References

- [D1] Dowson, D., and Higginson, G.R., 1966, "Elasto-hydrodynamic lubrication, The fundamentals of roller and gear lubrication", Pergamon Press, Oxford, Great Britain.
  - [H1] Hamrock, B.J. and Dowson, D., 1977, "Isothermal elastohydrodynamic lubrication of point contacts, part III, fully flooded results," *ASME JOT*, **99**, 264-276.
  - [L1] 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.
  - [M1] Moes, H., 1965, "Discussion on a paper by D. Dowson," *Proc. Inst. Mech. Engrs*, **180**, 244-245.
  - [M2] Moes, H., 1990, "Optimum similarity analysis with applications to elastohydrodynamic lubrication," Internal Memo, University of Twente, Enschede, the Netherlands.
-

## Appendix B

### Elastic deformation

#### B.1 Line contact

In the line contact situation the (dimensionless) elastic deformation of the surfaces associated with a pressure profile  $P(X)$  is given by:

$$d(X) = -\frac{1}{\pi} \int_{X_a}^{X_b} \ln |X_i - X'| P(X') dX' \quad (\text{B.1})$$

Approximating the pressure profile by a piecewise constant function with value  $P_j = P(X_j)$  in the region  $X_j - h/2 \leq X' \leq X_j + h/2$  on a uniform grid with mesh size  $h$  the elastic deformation in grid point  $i$  ( $X_i = X_0 + i h$ ) can be written as:

$$d(X_i) = -\frac{1}{\pi} \sum_{j=0}^{j=n} K_{ij}^{hh} P_j \quad (\text{B.2})$$

where the coefficients  $K_{ij}^{hh}$  are defined by:

$$K_{ij}^{hh} = \int_{X_j-h/2}^{X_j+h/2} \ln |X_i - X'| dX' \quad (\text{B.3})$$

These coefficients can be calculated analytically and the result depends on the value of  $|X_i - X'|$  in the integral. Firstly, assume  $(X_i - X') > 0$  for  $X_j - h/2 \leq X' \leq X_j + h/2$ . In that case integration of equation (B.3) yields:

$$K_{ij}^{hh} = (X_i - X_j + h/2)(\ln(X_i - X_j + h/2) - 1) - (X_i - X_j - h/2)(\ln(X_i - X_j - h/2) - 1) \quad (\text{B.4})$$

Similarly, assuming  $(X_i - X') < 0$  for  $X_j - h/2 \leq X' \leq X_j + h/2$  the integration gives:

$$K_{ij}^{hh} = (X_j - X_i + h/2)(\ln(X_j - X_i + h/2) - 1) - (X_j - X_i - h/2)(\ln(X_j - X_i - h/2) - 1) \quad (\text{B.5})$$

Equation B.4 and equation B.5 can be written into one expression valid for all  $X_j \neq X_i$ :

$$K_{ij}^{hh} = \frac{(X_i - X_j + h/2)(\ln |X_i - X_j + h/2| - 1) - (X_i - X_j - h/2)(\ln |X_i - X_j - h/2| - 1)}{2} \quad (\text{B.6})$$

Finally, assume  $(X_i - X')$  changes sign, i.e.  $X_i = X_j$ . In that case the integral is calculated in two steps taking the limit for  $\epsilon \rightarrow 0$  of:

$$K_{ii}^{hh} = \frac{(h/2)(\ln(h/2) - 1) - (\epsilon)(\ln(\epsilon) - 1) + (h/2)(\ln(\epsilon) - 1) - (-h/2)(\ln(h/2) - 1)}{2} \quad (\text{B.7})$$

Hence:

$$K_{ii}^{hh} = (h/2)(\ln(h/2) - 1) + (h/2)(\ln(h/2) - 1) \quad (\text{B.8})$$

which is exactly also what equation (B.6) gives for  $X_j = X_i$ . Summarizing:  $K_{ij}^{hh}$  is given by:

$$K_{ij}^{hh} = \frac{(X_i - X_j + h/2)(\ln |X_i - X_j + h/2| - 1) - (X_i - X_j - h/2)(\ln |X_i - X_j - h/2| - 1)}{2} \quad (\text{B.9})$$

## B.2 Circular contact

In the circular contact situation the elastic deformation is given by:

$$d(X, Y) = \frac{2}{\pi^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{P(X', Y') dX' dY'}{\sqrt{(X - X')^2 + (Y - Y')^2}} \quad (\text{B.10})$$

This integral is discretized in the same way as the discretization of the elastic deformation integral for the line contact problem, i.e. the pressure is approximated by a piecewise constant function on a uniform grid with mesh size  $h$  and value  $P_{k,l}$  in the region  $\{(X, Y) \in \mathbb{R}^2 | X_k - h/2 \leq X \leq X_k + h/2 \wedge Y_l - h/2 \leq Y \leq Y_l + h/2\}$ . Consequently, the elastic deformation in grid point  $(i, j)$  can be written as:

$$d_{i,j} = d(X_i, Y_j) = \frac{2}{\pi^2} \sum_{k=0}^{n_x} \sum_{l=0}^{n_y} K_{ikjl}^{hhhh} \quad (\text{B.11})$$

where the coefficients  $K_{ikjl}^{hhhh}$  are given by:

$$K_{ikjl}^{hhhh} \equiv \int_{X_k-h/2}^{X_k+h/2} \int_{Y_l-h/2}^{Y_l+h/2} \frac{P(X', Y') dX' dY'}{\sqrt{(X_i - X')^2 + (Y_j - Y')^2}} \quad (\text{B.12})$$

and can be calculated analytically. The integration yields nine different results for the cases that  $X_i < X_k$ ,  $X_i > X_k$ ,  $X_i = X_k$  and  $Y_j < Y_l$ ,  $Y_j > Y_l$ ,  $Y_j = Y_l$  respectively. The nine equations can be combined into one expression:

$$K_{ikjl}^{hhhh} = |X_p| \ln \left( \frac{Y_p + \sqrt{X_p^2 + Y_p^2}}{Y_m + \sqrt{X_p^2 + Y_m^2}} \right) + |Y_m| \ln \left( \frac{X_m + \sqrt{Y_m^2 + X_m^2}}{X_p + \sqrt{Y_m^2 + X_p^2}} \right) + \\ |X_m| \ln \left( \frac{Y_m + \sqrt{X_m^2 + Y_m^2}}{Y_p + \sqrt{X_m^2 + Y_p^2}} \right) + |Y_p| \ln \left( \frac{X_p + \sqrt{Y_p^2 + X_p^2}}{X_m + \sqrt{Y_p^2 + X_m^2}} \right)$$

where:

$$X_p = X_i - X_k + h/2, \quad X_m = X_i - X_k - h/2 \\ Y_p = Y_j - Y_l + h/2, \quad Y_m = Y_j - Y_l - h/2$$

In a shorter, more general notation, introducing  $q = (i, j)$  and  $r = (k, l)$  a compact expression can be given for the elastic deformation:

$$d_q = \frac{2}{\pi^2} \sum_r K_{qr}^{hh} P_r \quad (\text{B.13})$$

The latter notation is used in the description of the multilevel multi-integration for the fast calculation of such summations, see chapter 3.





## Appendix C

### Smoothing rate analysis linearized EHL 1-d problem

As an example of a more complicated smoothing rate analysis consider the linearized EHL line contact model problem with  $\epsilon = 0$ .

$$\frac{dH}{dX} = 0 \quad (\text{C.1})$$

Substituting the following definition of  $H$ :

$$H = \int K(X, X') P(X') dX' \quad (\text{C.2})$$

gives:

$$\frac{d}{dX} \int K(X, X') P(X') dX' = 0 \quad (\text{C.3})$$

Note that in the film thickness equation, i.e. equation (C.2), the part coming from the undeformed surface, i.e.  $H_{00} + X^2/2$ , is left out. This part is of no importance with respect to the present analysis.

Our intention is to solve the discretized version of this equation using a Jacobi first order distributive relaxation on  $P$ . The smoothing behaviour of this distributive relaxation can be analysed using the approach outlined by Brandt [B1, page 42]. Define the ghost function  $\Phi$  according to:

$$P = -\frac{d\Phi}{dX} \quad (\text{C.4})$$

Substitution in equation (C.3) gives:

$$-\frac{d}{dX} \int K(X, X') \Phi_{X'}(X') dX' = 0 \quad (\text{C.5})$$

Using the fact that for the present kernel ( $K(X, X') = \ln |X - X'|$ ) holds:  $K_X = -K_{X'}$ , and applying integration by parts, this equation can be reworked to:

$$\int \frac{d^2 K(X, X')}{dX'^2} \Phi(X') dX' = 0 \quad (\text{C.6})$$

Solving  $P$  from the discrete equivalent of equation (C.3) using a Jacobi first order distributive relaxation on  $P$  corresponds with solving  $\Phi$  from the discrete equivalent of equation (C.6) with a simple (non distributive) Jacobi relaxation. Hence, the smoothing rate is the same in both situations. However, in the latter situation this smoothing rate can be calculated straightforwardly, whereas in the first situation, due to the distribution, the computation is more complex. Therefore, the smoothing rate is calculated using the problem in terms of the ghost function  $\Phi$ .

Discretization of equation (C.6) gives:

$$h^2 \sum \frac{\Delta K_{kl}^{hh}}{h^2} \Phi_k = 0 \quad (\text{C.7})$$

where  $\Delta K_{kl}^{hh}$  denotes the second order difference of  $K_{kl}^{hh}$ :

$$\Delta K_{kl}^{hh} = K_{kl-1}^{hh} - 2 K_{kl}^{hh} + K_{kl+1}^{hh} \quad (\text{C.8})$$

The simple jacobi relaxation on  $\Phi$  can be described as follows. Given an approximation  $\hat{\Phi}_k$  a new approximation  $\bar{\Phi}_k$  is calculated from:

$$\sum_{k \neq l} \Delta K_{kl}^{hh} \bar{\Phi}_l + \Delta K_{kk}^{hh} \bar{\Phi}_k = 0 \quad (\text{C.9})$$

Since the kernel  $K$  is symmetric define:

$$\Delta K_r^h = \Delta K_{ij}^{hh} \quad \text{for } |i - j| = r \quad (\text{C.10})$$

and:

$$\Delta K_r^h = K_{r-1}^h - 2 K_r^h + K_{r+1}^h \quad (\text{C.11})$$

Subsequently, assuming the domain to be infinitely large, equation (C.9) can be written as:

$$\sum_{l=1}^{\infty} \Delta K_r^h (\bar{\Phi}_{k+l} + \bar{\Phi}_{k-l}) + \Delta K_0^h \bar{\Phi}_k = 0 \quad (\text{C.12})$$

The error at site  $k$  before relaxation is given by:

$$\tilde{v}_k = \Phi_k - \bar{\Phi}_k \quad (\text{C.13})$$

whereas the error after relaxation at this site is defined as:

$$\bar{v}_k = \Phi_k - \bar{\Phi}_k \quad (\text{C.14})$$

Substitution of  $\tilde{v}$  and  $\bar{v}$  from equations (C.13) and (C.14) in equation (C.12) yields the following equation relating the error after relaxation to the error before relaxation:

$$\sum_{l=1}^{\infty} \Delta K_r^h (\tilde{v}_{k+l} + \tilde{v}_{k-l}) + \Delta K_0^h \bar{v}_k = 0 \quad (\text{C.15})$$

Expanding both  $\tilde{v}$  and  $\bar{v}$  into Fourier series:

$$\tilde{v}_k = \sum A(\theta) e^{i\theta k} \quad (\text{C.16})$$

$$\bar{v}_k = \sum \bar{A}(\theta) e^{i\theta k} \quad (\text{C.17})$$

and using the fact that the relaxation maps a specific Fourier component on itself gives the following expression for the relation between the amplitude of a component before and after relaxation:

$$A \sum_{l=1}^{\infty} \Delta K_r^h 2 \cos(\theta l) + \bar{A} \Delta K_0^h = 0 \quad (\text{C.18})$$

Consequently the *amplification factor* of the  $\theta$  component due to one relaxation sweep is:

$$\mu(\theta) = \left| \frac{\bar{A}(\theta)}{A(\theta)} \right| = \left| \frac{1}{\Delta K_0^h} \sum_{l=1}^{\infty} \Delta K_r^h 2 \cos(\theta l) \right| \quad (\text{C.19})$$

This amplification factor as a function of  $(\theta)$  is shown below (see also figure 4.2) and the asymptotic smoothingrate  $\bar{\mu}$  is 0.4.

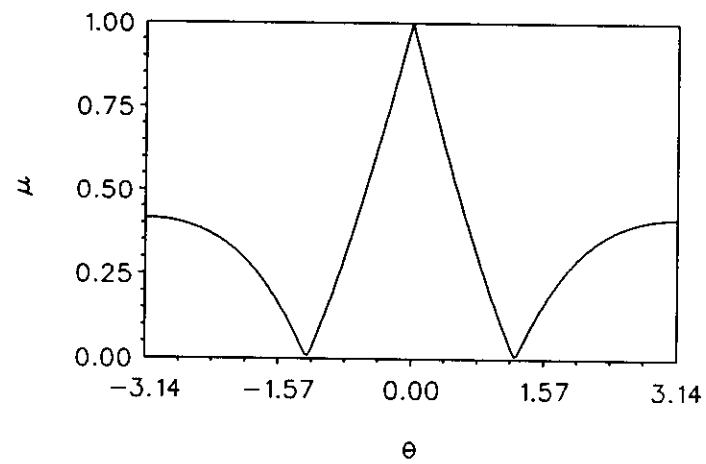


FIGURE C.1: Amplitude reduction factor  $\mu(\theta)$  for Jacobi dipole relaxation on the linearized problem,  $\epsilon = 0$

## References

- [B1] **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, W. Germany.

---

## Appendix D

### Decimation

#### D.1 General description

Consider the following system of discrete equations:

$$a_i v_{i-1} + b_i v_i + c_i v_{i+1} = g_i \quad (\text{D.1})$$

for  $1 \leq i \leq n-1$  with  $v_0$  and  $v_n$  given. Either  $a_i$ ,  $b_i$ , and  $c_i$  are coefficients and  $v_i$  is a scalar unknown or  $a_i$ ,  $b_i$  and  $c_i$  are matrices and  $v_i$  is a vector of unknowns. From this equation an expression for  $v_i$  in terms of its nearest neighbours  $i+1$  and  $i-1$  can be derived:

$$v_i = b_i^{-1}(g_i - a_i v_{i-1} - c_i v_{i+1}) \quad (\text{D.2})$$

where  $b_i^{-1}$  denotes the inverse of  $b_i$ .

Using equation (D.2) to express  $v_{i-1}$  and  $v_{i+1}$  in terms of their neighbouring points and substitution of the result in equation (D.1) gives the following equation relating  $v_i$  to  $v_{i-2}$  and  $v_{i+2}$ :

$$A_i v_{i-2} + B_i v_i + C_i v_{i+2} = G_i \quad (\text{D.3})$$

where:

$$A_i = -a_i b_{i-1}^{-1} a_{i-1} \quad (\text{D.4})$$

$$B_i = -a_i b_{i-1}^{-1} c_{i-1} + b_i - c_i b_{i+1}^{-1} a_{i+1} \quad (\text{D.5})$$

$$C_i = -c_i b_{i+1}^{-1} c_{i+1} \quad (\text{D.6})$$

and

$$G_i = g_i - a_i b_{i-1}^{-1} g_{i-1} - c_i b_{i+1}^{-1} g_{i+1} \quad (\text{D.7})$$

Equation (D.3) can be viewed as the representation of the original problem on a coarser grid, i.e. a system of equations with only half the number of unknowns:

$$\bar{A}_I V_{I-1} + \bar{B}_I V_I + \bar{C}_I V_{I+1} = \bar{G}_I \quad (\text{D.8})$$

where  $V_I = v_{2I}$ ,  $\bar{A}_I = A_{2I}$ ,  $\bar{B}_I = B_{2I}$ ,  $\bar{C}_I = C_{2I}$ , and  $\bar{G}_I = G_{2I}$ . Assume this reduced system of equations is solved, then the solution of the original system of equations follows in two steps. First for all  $i = 2I$ :

$$v_i = V_I \quad (\text{D.9})$$

Next the values of  $v_i$  in the points  $i = 2I + 1$  can be computed from equation (D.2):

$$v_i = b_i^{-1}(g_i - a_i v_{i-1} - c_i v_{i+1}) \quad \text{for } i = 2I + 1 \quad (\text{D.10})$$

Summarizing, the solution of the system of equations (D.1) is obtained in two steps. First a reduced system is constructed according to (D.8). Subsequently, this reduced system is solved. Finally, the solution of the original system follows from (D.9) and (D.10).

The remaining task is to solve the reduced system of equations. Obviously it has exactly the same structure. Hence, the coarsening procedure can be applied once more. In fact, it can be applied recursively until a system of equations is obtained that can be solved in only a few operations.

This completes the general description of the decimation process. The following sections give some details with respect to the application of the process in the line relaxation schemes discussed in chapter 5.



## D.2 Application to the linearized EHL 2-d problems

First consider the system of equations (5.26) and (5.27) to be solved on each line (constant  $j$ ) in the case of the distributed line relaxation applied to the solution of the linearized EHL 2-d problem with constant  $\epsilon$ :

$$\bar{H}_{i,j} - \frac{2}{\pi^2} \sum_{k=1}^{n_s-1} \Delta K_{ikjj}^{hhhh} \delta_{k,j} = f_{i,j} + \tilde{w}_{i,j} \quad (D.11)$$

$$\begin{aligned} \epsilon/h^2 \{ & (\tilde{P}_{i-1,j} - \delta_{i-2,j}/4 + \delta_{i-1,j} - \delta_{i,j}/4) + \\ & (\tilde{P}_{i+1,j} - \delta_{i,j}/4 + \delta_{i+1,j} - \delta_{i+2,j}/4) - \\ & 4 (\tilde{P}_{i,j} - \delta_{i-1,j}/4 + \delta_{i,j} - \delta_{i+1,j}/4) + \\ & (\tilde{P}_{i,j+1} - \delta_{i,j}/4) + (\tilde{P}_{i,j-1} - \delta_{i,j}/4) \} - \\ & (\bar{H}_{i,j} - \bar{H}_{i-1,j})/h = 0 \end{aligned} \quad (D.12)$$

To obtain the line relaxation efficiency it is not necessary to solve the system exactly. Therefore, the following approximate system is solved instead:

$$\bar{H}_{i,j} - \frac{2}{\pi^2} (\Delta K_{iijj}^{hhhh} \delta_{i,j} + \Delta K_{i-1jj}^{hhhh} \delta_{i-1,j} + \Delta K_{i+1jj}^{hhhh} \delta_{i+1,j}) = f_{i,j} + \tilde{w}_{i,j} \quad (D.13)$$

$$\begin{aligned} \epsilon/h^2 \{ & (\tilde{P}_{i-1,j} + \delta_{i-1,j} - \delta_{i,j}/4) + \\ & (\tilde{P}_{i+1,j} - \delta_{i,j}/4 + \delta_{i+1,j}) - \\ & 4 (\tilde{P}_{i,j} - \delta_{i-1,j}/4 + \delta_{i,j} - \delta_{i+1,j}/4) + \\ & (\tilde{P}_{i,j+1} - \delta_{i,j}/4) + (\tilde{P}_{i,j-1} - \delta_{i,j}/4) \} - \\ & (\bar{H}_{i,j} - \bar{H}_{i-1,j})/h = 0 \end{aligned} \quad (D.14)$$

This latter system of equations can easily be written in the form of equation (D.1) by defining:

$$v_i = \begin{pmatrix} \delta_{i,j} \\ \bar{H}_{i,j} \end{pmatrix} \quad g_i = \begin{pmatrix} f_{i,j} + \tilde{w}_{i,j} \\ -L^h(\tilde{P}_{i,j}) \end{pmatrix} \quad (D.15)$$

where  $L^h \tilde{P}_{i,j}$  is defined as:

$$L^h(\tilde{P}_{i,j}) = \epsilon/h^2 (\tilde{P}_{i-1,j} + \tilde{P}_{i+1,j} + \tilde{P}_{i,j+1} + \tilde{P}_{i,j-1} - 4 \tilde{P}_{i,j}) \quad (D.16)$$

$$a_i = \begin{pmatrix} -\frac{2}{\pi^2} \Delta K_{i-1jj}^{hhhh} & 0 \\ 2 \epsilon/h^2 & 1/h \end{pmatrix} \quad b_i = \begin{pmatrix} -\frac{2}{\pi^2} \Delta K_{iijj}^{hhhh} & 1 \\ -5 \epsilon/h^2 & -1/h \end{pmatrix} \quad (D.17)$$

$$c_i = \begin{pmatrix} -\frac{2}{\pi^2} \Delta K_{ii+1jj}^{hhhh} & 0 \\ 2 \epsilon/h^2 & 0 \end{pmatrix} \quad (D.18)$$

Subsequently, the procedure outlined in the previous section can be applied to its solution.

The approach to be followed in the case of the Gauss-Seidel line relaxation is basically similar. Instead of (5.28) and (5.29) the following approximate system is solved:

$$\bar{H}_{i,j} - \frac{2}{\pi^2} (K_{ii+jj}^{hhhh} \delta_{i,j} + K_{ii-1jj}^{hhhh} \delta_{i-1,j} + K_{ii+1jj}^{hhhh} \delta_{i+1,j}) = f_{i,j} + \tilde{w}_{i,j} \quad (D.19)$$

$$\begin{aligned} \epsilon/h^2 \{ & (\tilde{P}_{i-1,j} + \delta_{i-1,j}) + (\tilde{P}_{i+1,j} + \delta_{i+1,j}) - \\ & 4(\tilde{P}_{i,j} + \delta_{i,j}) + \tilde{P}_{i,j+1} + \tilde{P}_{i,j-1} \} - \\ (\bar{H}_{i,j} - \bar{H}_{i-1,j})/h & = 0 \end{aligned} \quad (D.20)$$

This system can be written in the form of (D.1) by defining:

$$v_i = \begin{pmatrix} \delta_{i,j} \\ \bar{H}_{i,j} \end{pmatrix} \quad g_i = \begin{pmatrix} f_{i,j} + \tilde{w}_{i,j} \\ -L^h(\tilde{P}_{i,j}) \end{pmatrix} \quad (D.21)$$

where  $L^h(\tilde{P}_{i,j})$  is defined as:

$$L^h(\tilde{P}_{i,j}) = \epsilon/h^2 (\tilde{P}_{i-1,j} + \tilde{P}_{i+1,j} + \tilde{P}_{i,j+1} + \tilde{P}_{i,j-1} - 4 \tilde{P}_{i,j}) \quad (D.22)$$

$$a_i = \begin{pmatrix} -\frac{2}{\pi^2} K_{ii-1jj}^{hhhh} & 0 \\ \epsilon/h^2 & 1/h \end{pmatrix} \quad b_i = \begin{pmatrix} -\frac{2}{\pi^2} K_{ii+jj}^{hhhh} & 1 \\ -4 \epsilon/h^2 & -1/h \end{pmatrix} \quad (D.23)$$

$$c_i = \begin{pmatrix} -\frac{2}{\pi^2} K_{ii+1jj}^{hhhh} & 0 \\ \epsilon/h^2 & 0 \end{pmatrix} \quad (D.24)$$

To conclude this section the approximate systems of equations of one line solved in the case of varying coefficients are considered. This requires only some minor modifications when compared to the approach outlined above. In the two situations considered above the matrices  $a_i$ ,  $b_i$  and  $c_i$  are independent of  $i$ . This is no longer true in the present case. If at least one of  $\epsilon_{i\pm 1/2,j\pm 1/2}/h^2$  exceeds the value of the crossover point, the approximate system of equations for this grid point is:

$$\bar{H}_{i,j} - \frac{2}{\pi^2} (K_{iijj}^{hhhh} \delta_{i,j} + K_{ii-1jj}^{hhhh} \delta_{i-1,j} + K_{ii+1jj}^{hhhh} \delta_{i+1,j}) = f_{i,j} + \tilde{w}_{i,j} \quad (D.25)$$

$$\begin{aligned} h^{-2} \{ & (\epsilon_{i-1/2,j} (\tilde{P}_{i-1,j} + \delta_{i-1,j}) + \\ & \epsilon_{i+1/2,j} (\tilde{P}_{i+1,j} + \delta_{i+1,j}) - \\ & (\epsilon_{i,j-1/2} + \epsilon_{i,j+1/2} + \epsilon_{i+1/2,j} + \epsilon_{i-1/2,j}) (\tilde{P}_{i,j} + \delta_{i,j}) + \\ & \epsilon_{i,j+1/2} \tilde{P}_{i,j+1} + \epsilon_{i,j-1/2} \tilde{P}_{i,j-1} \} - \\ (\bar{H}_{i,j} - \bar{H}_{i-1,j})/h & = 0 \end{aligned} \quad (D.26)$$

Hence,  $a_i$ ,  $b_i$ , and  $c_i$  are given by:

$$a_i = \begin{pmatrix} -\frac{2}{\pi^2} K_{ii-1jj}^{hhhh} & 0 \\ \epsilon_{i-1/2,j}/h^2 & 1/h \end{pmatrix} \quad b_i = \begin{pmatrix} -\frac{2}{\pi^2} K_{iijj}^{hhhh} & 1 \\ -\sum \epsilon_{kl} & -1/h \end{pmatrix} \quad (D.27)$$

$$c_i = \begin{pmatrix} -\frac{2}{\pi^2} K_{ii+1jj}^{hhhh} & 0 \\ \epsilon_{i+1/2,j}/h^2 & 0 \end{pmatrix} \quad (D.28)$$

where  $\sum \epsilon_{kl} = (\epsilon_{i-1/2,j} + \epsilon_{i+1/2,j} + \epsilon_{i,j+1/2} + \epsilon_{i,j-1/2})$ .

On the other hand, if all  $\epsilon_{i\pm 1/2,j\pm 1/2}/h^2$  are smaller than the value of the crossover point the approximate system of equations is:

$$\bar{H}_{i,j} - \frac{2}{\pi^2} (\Delta K_{iijj}^{hhhh} \delta_{i,j} + \Delta K_{ii-1jj}^{hhhh} \delta_{i-1,j} + \Delta K_{ii+1jj}^{hhhh} \delta_{i+1,j}) = f_{i,j} + \tilde{w}_{i,j} \quad (D.29)$$

$$\begin{aligned} h^{-2} \{ & \epsilon_{i-1/2,j} (\tilde{P}_{i-1,j} + \delta_{i-1,j} - \delta_{i,j}/4) + \\ & \epsilon_{i+1/2,j} (\tilde{P}_{i+1,j} - \delta_{i,j}/4 + \delta_{i+1,j}) + \\ & \epsilon_{i,j+1/2} (\tilde{P}_{i,j+1} - \delta_{i,j}/4) + \epsilon_{i,j-1/2} (\tilde{P}_{i,j-1} - \delta_{i,j}/4) - \\ & (\epsilon_{i,j-1/2} + \epsilon_{i,j+1/2} + \epsilon_{i+1/2,j} + \epsilon_{i-1/2,j}) \\ & (\tilde{P}_{i,j} - \delta_{i-1,j}/4 + \delta_{i,j} - \delta_{i+1,j}/4) \} - \\ (\bar{H}_{i,j} - \bar{H}_{i-1,j})/h & = 0 \end{aligned} \quad (D.30)$$

and the matrices  $a_i$ ,  $b_i$ , and  $c_i$  for this point are defined according to:

$$a_i = \begin{pmatrix} -\frac{2}{\pi^2} \Delta K_{ii-1jj}^{hhhh} & 0 \\ (\epsilon_{i-1/2,j} + \sum \epsilon_{kl}/4)/h^2 & 1/h \end{pmatrix} \quad (D.31)$$

$$b_i = \begin{pmatrix} -\frac{2}{\pi^2} \Delta K_{iijj}^{hhhh} & 1 \\ -5(\sum \epsilon_{kl})/(4h^2) & -1/h \end{pmatrix} \quad (D.32)$$

$$c_i = \begin{pmatrix} -\frac{2}{\pi^2} \Delta K_{i+1,j}^{hhhh} & 0 \\ (\epsilon_{i+1/2,j} + \sum \epsilon_{kl}/4)/h^2 & 0 \end{pmatrix} \quad (\text{D.33})$$

where again  $\sum \epsilon_{kl} = (\epsilon_{i-1/2,j} + \epsilon_{i+1/2,j} + \epsilon_{i,j+1/2} + \epsilon_{i,j-1/2})$ .

---