Adaptive Numerical Methods for Elastohydrodynamic Lubrication

by

Christopher Edward Goodyer

Submitted in accordance with the requirements for the degree of Doctor of Philosophy.

The University of Leeds
School of Computing

May 2001

The candidate confirms that the work submitted is his own and that the appropriate credit has been given where reference has been made to the work of others.

Abstract

Numerical solutions to elastohydrodynamic lubrication problems have been computed for the last half century. Over the past decade multilevel techniques have been successfully applied in several solvers and significant speed-ups achieved. The aim of numerical research in this field is to develop techniques in order to calculate accurate solutions to demanding industrial problems as efficiently as possible.

In this work the numerical solver, previously developed by Nurgat, is examined. Despite being successful in achieving converged results on a single grid, there were some unresolved issues relating to the multigrid performance. These problems are explained and the necessary modifications to the method used are detailed.

There is much current interest in obtaining results to transient elastohydrodynamic lubrication problems. These are examined in detail and the justification for the methods used are discussed. Example results for industrially relevant cases, such as variation of lubricant entrainment, oscillation of the applied load and the presence of surface defects are considered.

In many other fields, adaptation in both space and time is used to increase performance and accuracy. However, these techniques are not currently used for elastohydrodynamic lubrication problems. It is shown that they can be successfully applied and substantial benefits accrued.

A method of variable timestepping has been introduced and results are presented showing that not only is it as accurate as fixed time stepping methods, but that the computational work required to obtain these solutions is significantly reduced. Local error control on each individual timestep is also implemented.

Adaptation of the spatial mesh is also developed. By developing a hierarchy of refined meshes within the multigrid structure it is seen how significantly fewer computational points are used in the most expensive numerical calculations. This, in turn, means that the computational time required is reduced. Different criteria for adaptation are explained and results presented showing the relative levels of accuracy and speed-up achieved.

Acknowledgements

I would like to thank Professor Martin Berzins for all his assistance and guidance in supervising this research. Whilst it was not always straightforward, without his help it would have been nearly impossible. Many thanks are also due to Dr Laurence Scales from Shell Global Solutions for all his encouragement and positive feedback over the course of this project. I also wish to thank Dr Victoria Pennington for her guidance during the first year of the research.

Thanking Roger Fairlie for his help is almost taken for granted. Having both spent many hours wrestling with the minutiae of the code, and trying it make it do what we want, it is just *so* satisfying that it has finally all come together.

I am indebted to my family for all their help and support. To my mother who kept me sane and my brother who kept me amused, thanks.

Finally, I'd like to thank three people without whom I'd never have got this far. Anne Williams, who taught me that maths is fun, Ann Weitzel who convinced me that I could do maths, and Anna Holladay who helped me through the hardest part of this research without allowing me to lose all faith.

This work was funded through an EPSRC CASE studentship with Shell Global Solutions.

Declarations

Some parts of the work presented in this thesis have been published in the following articles:

Goodyer, C.E., Fairlie, R., Berzins, M. and Scales, L.E., "An In-depth investigation of the Multigrid approach for steady and transient EHL problems", in *Thinning Films in Lubrication, Proceedings of the 26th Leeds-Lyon Symposium on Tribology*, ed. D. Dowson *et al.*, Elsevier, 2000.

Goodyer, C.E., Fairlie, R., Berzins, M. and Scales, L.E., "Adaptive Techniques for Elastohydrodynamic Lubrication Solvers", in *Tribology Research: From Model Experiment to Industrial Problem, Proceedings of the 27th Leeds-Lyon Symposium on Tribology*, ed. G. Dalmaz *et al.*, Elsevier, 2001.

Contents

List of Figures

List of Tables

Notation

a	halfwidth of Hertzian contact
a_k	Gaussian quadrature weight
$\mathscr{A}$	maximum amplitude of roughness
c	relaxation factor for the Force Balance Equation
$\underline{e}$	discretisation error
E'	reduced modulus of elasticity (Pa)
f^k	right hand side vector on grid k
F	applied load (N)
$\underline{F}_1$, $\underline{F}_2$, $\underline{F}_3$	Differential algebraic matrices
$\underline{g}$	global error in time integration
G	Hamrock and Dowson non-dimensional material parameter
$\mathscr{G}$	undeformed surface geometry
h	dimensional film thickness (m)
h_{00}	dimensional central offset film thickness (m)
H	non-dimensionalised film thickness
H_{00}	dimensionless central offset film thickness
i, j, k, l	integer indices
`i`, `j`, `k`,	computational array element indices
I, J	coarse mesh integer indices
I_l^k	intergrid transfer operator from level l to level k
$\mathscr{I}$	point intensity in pseudo-interferometry pictures
k	grid level
k	order of temporal method
K	film thickness kernel matrix
L	Moes non-dimensional material parameter
L_X, L_Y	dimensionless radii of surface defect in X and Y directions respectively
$\mathscr{L}$	differential operator
m, m_1, m_2	multilevel multi-integration correction patch widths
m	multigrid iteration number
M	Moes non-dimensional loading parameter
n	timestep number
N	total number of mesh point in computational domain
n_x, n_y	number of mesh point in x and y directions respectively
N_X, N_Y	number of mesh point in X and Y directions respectively
p	dimensional pressure (Pa)

p_0	pressure coefficient in Roelands equation (Pa)
p_h	maximum Hertzian pressure (Pa)
P	non-dimensionalised pressure
r	stepsize change ratio
r^k	residual vector on grid k
r_1	radius of surface defect
$\mathscr{R}$	surface roughness term
R_x, R_y	reduced radii of curvature in X and Y directions respectively
t	dimensional time (s)
t_{min}	time of minimum film thickness in transient calculation (s)
T	non-dimensionalised time
T_0	non-dimensionalised time of initial timestep
$\underline{u}$	general solution vector
u_1, u_2	velocities of contacts 1 and 2 respectively, in the X direction
u_{ref}	sum of velocities at $T=0$
u_s	sum of velocities of contacts
U	Hamrock and Dowson non-dimensional speed parameter
$\underline{U}$	Differential algebraic solution vector
v_1, v_2	velocities of contacts 1 and 2 respectively, in the Y direction
W	Hamrock and Dowson non-dimensional loading parameter
$\mathscr{W}$	dimensionless wavelength of surface defect
x, y, z	dimensional coordinates
X, Y	dimensionless coordinates
x', y', X', Y'	dummy variables in integration
X_d	dimensionless position of surface defect
z_i	viscosity index
α	pressure-viscosity index (Pa^{-1})
$\overline{\alpha}$	dimensionless pressure-viscosity index
ΔT	non-dimensionalised timestep size
ΔT_n	non-dimensionalised timestep size for timestep n
ΔT_{max}	maximum non-dimensionalised timestep size
δx^k	mesh spacing in x direction on grid k
ΔX, ΔY	mesh sizes in X and Y directions respectively
$(\Delta X)^k$, $(\Delta Y)^k$	mesh sizes in X and Y directions respectively on grid k
ε	coefficient in non-dimensionalised Reynolds equation
ζ	coordinate in transformed range for numerical integration
ζ_k	abscissa in transformed range for numerical integration

η_0	viscosity at ambient pressure
η	dimensional viscosity (Pa s)
$\bar{\eta}$	non-dimensionalised viscosity
κ	ellipticity ratio
λ	coefficient in non-dimensionalised Reynolds equation
Λ	dimensionless wavelength for pseudo-interferometry pictures
ν_0	number of coarse grid smoothing cycles
ν_1	number of pre-smoothing cycles
ν_2	number of post-smoothing cycles
ν_3	total number of pre and post-smoothing cycles
ξ	error tolerance
ρ_0	density at ambient pressure
ρ	dimensional density ($\mathrm{kg\,m^{-3}}$)
$\bar{\rho}$	non-dimensionalised density
σ	convergence test parameter
τ^k	force balance correction term on grid k
τ_j^k	relative error tolerance between grids j and k
Φ	Non-dimensionalised load for shock loading case
Ψ	transformed function for numerical integration
Ω	solution domain
Ω_{ad}	adapted region of solution domain
Ω_{sing}	region of solution domain requiring correction of singularity
Ω_e	excitation frequency of variable load
tol $ATOL$ $RTOL$	error tolerances

Chapter 1

Introduction

1.1 Introduction

Elastohydrodynamic Lubrication, hereinafter referred to as EHL, is a topic which is concerned with understanding and modelling lubrication problems in which solid metal surfaces deform under large loads. The problems considered occur most commonly in combustion engines, although the ideas apply to many other regimes.

An engine is comprised of large numbers of individual elements, many of which are in motion relative to each other. Surfaces will therefore be in contact. Elementary mechanics demands that when such a motion is occurring then there will be a frictional force opposing the movement. The friction not only reduces the efficiency of the component, since work must be done to overcome friction, but also increases the wear.

In order to reduce the frictional force, a *lubricant* (oil) is applied between the surfaces. This separates the two contacting surfaces only slightly, but this is enough to stop them impacting upon each other. Friction is reduced to a tenth of the *dry contact* (unlubricated) case, and thus the wear is also dramatically reduced. This situation is called *hydrodynamic lubrication*. This is illustrated in Figure 1.1 with oil flow from left to right.

One particular component of interest is the journal bearings of a car. In this situation a very large pressure is applied over a very small surface area. Once the pressure exceeds about 0.3 GPa (i.e. 3×10^8 Pascals) the contact behaviour moves from being hydrodynamically lubricated to the elastohydrodynamic regime. *Elastohydrodynamic lubrication*

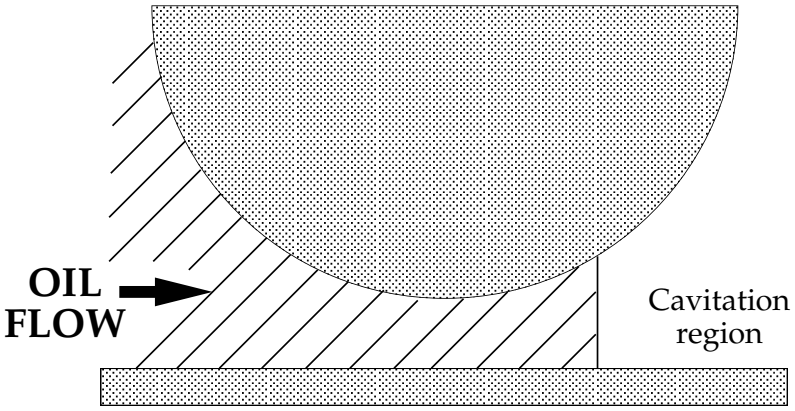


Figure 1.1: Representation of a hydrodynamically lubricated contact

(*EHL*) is different from hydrodynamic problems because here there is actual deformation of the contacting surfaces. This may sound unlikely but pressures in a journal bearing or gear commonly reach up to 3 GPa. Assuming pressure is force over area, this would correspond to three elephants balancing on the end of a pen!

With such a wide range of operating pressures in the contact, it is not difficult to understand that the properties of the lubricant itself will change across the contact. It is, however, of great importance to lubricant manufacturers that the oils being developed are as efficient as possible for the operating conditions for which they are intended. It is therefore necessary for designers of both lubricants and components to obtain performance results for a variety of lubricants in different operating conditions.

The range of scales in EHL problems is great. Applied loads cause pressure distributions across the contact of the order of giga-Pascals, minimum film thickness are in the micrometre range, and lubricant molecules pass through the contact in a hundredth of a second. This illustrates how difficult it is to conduct physical experiments into the behaviour of EHL contacts. That consistent results are achievable at all is a great accomplishment, and a testament to the skills of those people whose experimental work pioneered the techniques now used, described in Section 1.2.

It is now the case that research into EHL problems involves a combination of experiments and numerical simulations. Assuming that accurate computer software (code) is available, then solutions to numerous EHL test cases may be obtained at minimal expense. The more efficient the code, the quicker results may be obtained and used. The development of these numerical techniques is charted in Section 1.3, where emphasis is placed on those techniques with some bearing on this work.

To summarise a century of work by many outstanding engineers in a few pages is somewhat difficult. There are several comprehensive reviews already in the literature,

notably those of Dowson [30], Dowson and Ehret [31] and Jacobson [76]. The following two sections provide the highlights and notable milestones in the study of EHL. In Section 1.4 the rest of this thesis is outlined.

1.2 History of EHL Research

All work in fluid film lubrication can be traced back to the 1880s, when a combination of experiments was followed by a unifying mathematical theory. In 1886 Osborne Reynolds [121] formulated equations derived from the Navier-Stokes equations, to describe the pressure distribution for an applied load on a given geometry, relating the pressure to the speed of the moving surfaces. This work was itself an attempt to explain the results of the experimental work of Beauchamp Tower [137, 138] which was the first to detect high pressures in the lubricant film. This pressure variation was also the conclusion of Nicoli Petrov [116], after he had conducted friction experiments on railway axle bearings, at about the same time.

Despite initial success in the application of the Reynolds Equation to the design of journal bearings, e.g. [105], in 1916 the case of trying to model lubrication in gears caused problems. Martin [100] considered an isoviscous lubricant between two smooth rigid cylinders, representing the gear teeth. A relationship between the operating conditions and the minimum lubricant film thickness was obtained. However, applying known physical values into the formula, a film thickness of 0.01 μm was predicted, which was significantly less than the known surface roughness of gears at around 0.4–0.8 μm. This, therefore, meant that there must have been other factors involved in the near wearless operation of gears than simply the model developed thus far.

This impasse was not resolved quickly. For over 30 years, two possible themes were investigated into what the missing link could be. It was assumed to be either due to the elastic deformation of the solids, or due to the viscosity-pressure characteristics of the lubricant. Work was carried out independently on each of these ideas. Deformation was shown to have some effect, e.g. [103], but the simultaneous calculation of both elasticity and hydrodynamic equations was too complex for a numerical problem at that time. The viscosity-pressure work, e.g. [102] also produced larger film thickness values than those predicted by Martin, but still not nearly enough to obtain numerical results consistent with experiments.

The work of Ertel [39] and Grubin [54] combined the analytical solution of the deformation of a dry contact [67] with the viscosity-pressure effects calculated using the exponential viscosity-pressure relation of Barus [11]. The minimum film thickness for-

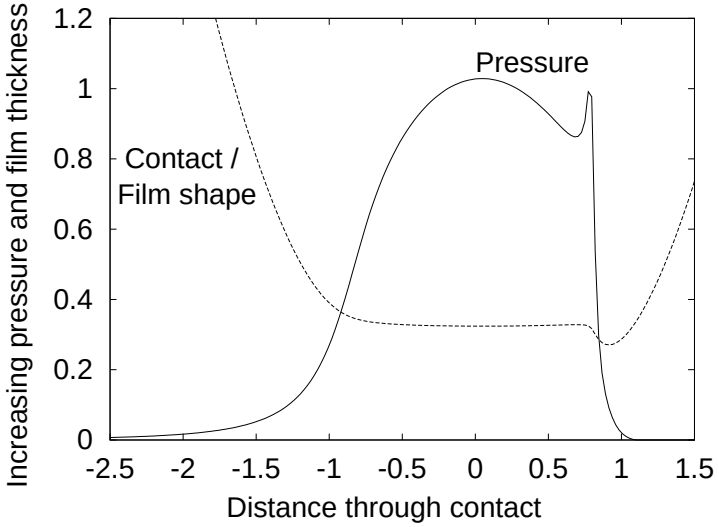


Figure 1.2: Typical solution for pressure and film thickness in an EHL line contact

mula obtained was only valid in a very limited parameter region, but this was the most major advance since the formulation of the Reynolds equation.

The first solution of the full numerical line contact elastohydrodynamic lubrication problem was presented by Petrusevich in 1951 [117]. This work also used the Barus viscosity-pressure relation, although only three solutions were obtained. It was also the first to show what has since become known as the *Petrusevich pressure spike*, itself an important factor in the development of subsurface stresses and on the life of rolling contacts [34]. The Petrusevich spike is shown in Figure 1.2 where the typical features of an EHL contact are shown. Lubricant entrainment is from left to right, and the inlet is assumed to be fully flooded. In the centre of the contact, known as the *contact region*, there is a near Hertzian pressure profile, with the pressure spike towards the outflow. From the end of the contact region the pressure solution is zero in what is known as the *cavitation region*. This is where the lubricant film is no longer contiguous; bubbles of air at (near) ambient pressure are present in the oil film, as can be seen in the experimental results shown in Figure 5.5. In terms of the geometry of the surfaces, it can be clearly seen that in the elastohydrodynamically lubricated case the parabolic shape on the contact has been deformed. The roller has flattened out across the contact area, with a constriction present in the outflow where the pressure drops steeply.

With the computing power available at that time, obtaining numerical solutions to the two dimensional point contact EHL problem was simply not possible due to the added mathematical difficulties involved. In the mid-1960s two attempts had been made using

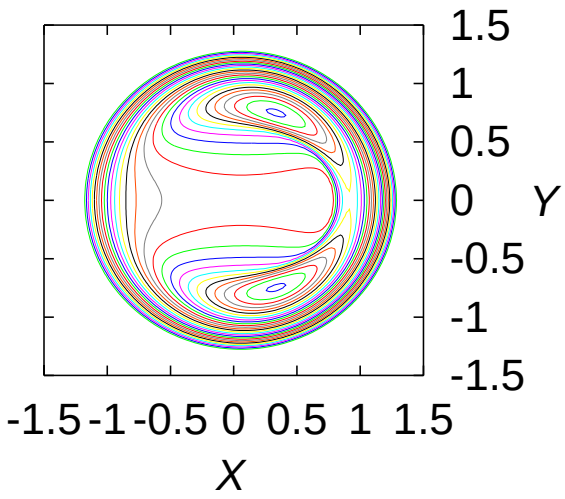


Figure 1.3: Contour plot of a typical solution for film thickness across an EHL circular point contact

the Grubin approach [4, 23], but full solutions did not appear until the 1970s. Ranger *et al.* concluded their 1975 solution paper [120] by indicating that the two limitations on obtaining solutions were the "computer capacity" available, and the "ignorance of the relationship between pressure, viscosity and density". The effect of ellipticity in the point contact was investigated by Hamrock and Dowson [60–64], in which expressions for central and minimum film thickness were developed. These were found to be in reasonable agreement with experimental results. The formulae obtained are given as Equations (2.30) and (2.31) respectively. Work was also conducted at this time on the problem of *starved lubrication* where the inlet region is not fully flooded [63, 120].

Examples of point contact problem solutions will be seen throughout the rest of this work. This early work established that there are several physical features not present in the line contact case. These will be shown in depth in Section 4.3 but most notable is the change in deformation shape from that predicted in the line contact case. In point contacts the minimum film thickness is not found on the centreline but on each of the two *sidelobes* which develop to form a horseshoe shape. This is characteristic of point contact EHL problems. An example solution is shown in Figure 1.3 which shows a contour plot of the central region. It can be seen that away from the centre of the contact the film shape is undeformed from the original parabolic profile. Inside the centre of the contact, the large flat central region is again visible, with the horseshoe replacing the film thickness dip of the line contact case.

Once the ability to solve two dimensional EHL problems had been established, attention returned to the models of the lubricant used. Now-a-days probably the most es-

tablished viscosity model used is that of Roelands [123], a more complicated version of the Barus equation. All models used must only be expected to be accurate within certain parameter ranges, because the number of variables required for modelling real life lubricants is far too large to try to model all fluids with one simple equation. This is especially true when using non-Newtonian lubricant models.

The study of Non-Newtonian fluids is known as *rheology* and "rheology is a difficult subject" [10]. Lubricants are typically non-Newtonian because they often contain long molecular polymer chains or additive suspensions, both of which affect their flow characteristics. Non-Newtonian behaviour of a fluid exists whenever the rate of shear is not proportional to the shear stress, or there are non-zero normal stresses. There are three main viscoelastic effects which govern the behaviour of non-Newtonian fluids, namely shear-thinning, variation of extensional viscosity and fluid memory. Shear-thinning is when the viscosity of the fluid reduces with increasing rate of shear in a steady shear flow [10]. A common example of this is how paint gets easier to stir, the more you stir it. Extensional flows are non-Newtonian whenever the fluid flow is not a shear flow and this is important in such cases as polymer processing. The extensional viscosity may increase or decrease with increasing rate of strain, being called tension-thickening or tension-thinning respectively. This subject is considered in detail by Petrie [115]. Fluid memory is the idea that a fluid's behaviour at any moment is not just related to the conditions it is experiencing at that moment, but also to its previous states. Again, see [10] for examples.

The choice of rheological model to use is very much fluid dependent. Many fluids are Newtonian in their characteristics, whilst the behaviour of others varies dramatically with the operating conditions. The simplest elastic non-Newtonian models include those of Maxwell [101] and Oldroyd [114], the latter of which includes shear thinning. There are many more detailed or specialised viscoelastic models; however there have been very few applied to EHL modelling, besides those detailed below.

Newtonian models have been effective in modelling EHL film thickness, which depends primarily upon conditions in the inlet to the contact. However, they greatly overestimate the friction in a contact because the shear stresses are not accurately predicted in the high pressure/shear flow in the Hertzian zone using these models. The first published difficulties with a Newtonian model came in 1959 when experimental work by Smith [134] reported that the lubricant acted as a plastic solid in the contact area. Use of the Ree-Eyring model to deal with the non-linear relationship between shear stress and shear rate has been developed in many studies, e.g. [9, 40, 81] and still continues today. This work includes using Ree-Eyring models for line contacts, e.g. by Chang *et al.* [25], for circular contacts, e.g. by Kim and Sadeghi [85], and a related model for line contact

cases by Lee and Hamrock [87]. Other models used include the White-Metzner model [153] which Scales [127] successfully compared to experimental friction results, building upon the work of Walters' group at Aberystwyth on journal bearings [122]. This model explicitly included fluid memory effects on macroscopic friction, including their variation with local conditions in the contact (extensional effects were assumed to be negligible in the shear dominated flows considered). The notion that the lubricant changes phase to a solid at very high pressures, leading to a breakdown in film continuity under shear and hence rendering the Reynolds Equation invalid in certain regions, has been explored by Ehret *et al.* [37] in a plug flow model. Ehret's work allowed slip conditions at the boundary, and a good agreement was obtained between experimental and numerical results for the regimes he considered.

Besides the viscosity-pressure relation, another important physical factor is that of thermal effects. At the high pressures existing in EHL contacts there can be significant temperature changes over the surfaces, and through the lubricant film. These effects have been included in some of the rheological studies mentioned above, however they deserve a separate consideration too. The temperature in the inlet region of an EHL contact is very important to the resulting film thickness profile across the rest of the contact [53], and this heat is conducted almost entirely to the contacts [29]. Under sliding conditions, rather than pure rolling, the temperature rise across the contact can be as much as 100° C [5].

The components used in engines and machinery are real surfaces which have not been specially prepared before each use, and therefore are not perfectly smooth. This may not be a great revelation, knowing that the roughness of the gears in Martin's work [100] was known to be larger than the minimum hydrodynamic film thickness predicted, however the ability to model rough contacts is now growing in importance. Applications, such as computer hard drives, are continuously reducing the lubricant film width, and hence the effect of surface asperities is becoming more important in estimating the life of components. In cases where the surface roughness is of at least the same magnitude as the elastohydrodynamic film thickness, then it is well known that the components may operate as though lubricated with a fluid film. This is because the surface roughness will generate pressures great enough to flatten the asperities to leaving only smooth ripples on the surface. This is known as *micro-elastohydrodynamic lubrication*.

Accurately measuring surface roughness is a topic which is both limited and defined by the accuracy of the measuring equipment. Today, it is the effective roughness that is being investigated. Work by Venner *et al.* [141, 143], Chang *et al.* [26, 27], Ai and Cheng [2] and Hooke [71, 72], investigated line contact problems with either ridges or waviness patterns passing through the contact. The effect of different slide to roll ratios

was also investigated. Point contact cases with dents, ridges, and/or waviness patterns, were solved initially in steady state by Lubrecht *et al.* [97], but more importantly for transient analyses by Venner and Lubrecht [142, 144], Ai and Cheng [1] and Ehret *et al.* [38]. Recent work by Venner and collaborators [94, 98, 146] has been investigating the amplitude reduction of waviness in both line and point contacts. As with all simulations of real life phenomena it is important to ensure that the assumptions made in the models are valid when compared to experimental results. Comparisons against the experimental work of Kaneta *et al.* [83, 84] have been done in some of the work mentioned above. However, in the conclusion of his analysis on the validity of the use of Newtonian models in these cases, Greenwood [51] highlights that "there is considerable danger that these beautiful calculations are only of mathematical interest". He does, however, leave open the question of which lubricant model to use to successfully reproduce the experimental results from [82] where dimples not associated with surface defects appeared in the solution, but only when the steel ball was stationary and the glass disc was mobile.

1.3 Numerical Methods

With analytic solutions to EHL problems only possible in extremely limited regions for very basic models, the ability to obtain solutions to these problems numerically became essential. The complexity of problems available for consideration has always been constrained by the computing power available. The early numerical work of Petrusevich [117] led to many different solution methods for the line contact case. However, the highly computationally expensive two dimensional cases, such as will be considered in this work, were not solved until the early 1970s, e.g. [120].

The main area in which there is a choice of numerical methods available is in the solution of the Reynolds Equation. It will be shown later that, for the Newtonian fluids to be considered here, the other governing equations are mathematically relatively simple to solve – even if not always computationally cheap. The Reynolds Equation, however, is a highly non-linear partial differential equation which gives the pressure distribution for a given geometry.

The elastic deformation of the surfaces, by definition, gives the geometry of the contact. This deformation is governed by the pressure distribution over the contact and hence there is a very important counterpoint between the solution of these two equations. One method, which will be the one applied here, is to solve them each in turn, and then iterate the process to obtain converged solutions of both pressure and geometry.

Gauss-Seidel relaxation has been widely used for solving the Reynolds Equation. Ex-

amples can be found for the line contact case in Hamrock and Jacobson [65], and for point contacts in Hamrock and Dowson [60], Chittenden *et al.* [28], and Zhu and Cheng [158]. This method is unstable for highly loaded contacts [140]. Under-relaxation does alleviate this, but reduces the speed of convergence, which can already be slow.

The Newton-Raphson method has been used to solve for both pressure and geometry simultaneously. First described by Okamura [113], this method inverts the Jacobian matrix of all the solution variables to obtain new approximations. This does, however, lead to very expensive calculations, since for a domain of N points, the inversion will require $\mathcal{O}(N^3)$ operations and $\mathcal{O}(N^2)$ computational storage space. Other drawbacks include the near singularity of the Jacobian matrix for very high loads and the difficulties posed by the varying location of the cavitation boundary across the domain [95]. Use is therefore mainly confined to lightly loaded line contact problems, although some advancements have been made since [24, 66, 74, 86, 111, 125].

The inverse methods of Ertel [39] were first applied to EHL line contact problems by Dowson and Higginson in 1959 [32], from which a formula for predicting minimum film thickness was developed. Point contact solutions followed in the 1980s by Evans and Snidle [41, 42]. The method operates by comparing the geometry obtained by solving the elasticity equation for a given pressure distribution, with the geometry obtained by solving the Reynolds equation for the geometry rather than the pressure. The difference between these two results is used to correct the pressure. This method requires solution of a cubic equation for the geometry, where the correct one of the three roots must be selected. This method is not based on strong mathematical principles, but inspired guesswork. It is also only stable in highly loaded situations, meaning another solver must be used in the non-contact region [41, 42]. It does, however, allow solutions to extremely highly loaded situations to be obtained. Computationally, this method approaches $\mathcal{O}(N^3)$ operations and is therefore undesirable for large systems.

Finite differences are not the only numerical approach that can be used. Using finite elements has enabled many demanding, highly non-linear problems to be solved on unstructured grids up to many dimensions, in fields such as computational fluid dynamics, e.g. [90], and solid mechanics, e.g. [160]. Finite element techniques have been applied to EHL problems since the early 1970s. Line contacts were first considered by Taylor and O'Callaghan [136] and Rohde and Oh [124] and point contacts by Oh and Rohde [112]. Whilst these techniques are not widely used to solve EHL problems today, there is still research being done, e.g. [36].

The finite element techniques used so far appear to have been mainly restricted to the use of the Galerkin method, see for example [112, 159]. A disadvantage of this ap-

proach is that it is analogous to using central difference approximations to the convective terms [78], and would require additional artificial viscosity in practice. There are several similar possible extensions to this method which may prove better for solving EHL problems, namely Petrov-Galerkin, streamline diffusion and Galerkin least squares [77]. In these cases the amount of artificial viscosity introduced is related to the resolution of the mesh, and adds extra stability to the scheme. This is particularly important in convection-diffusion problems when there is a high mesh Peclet number [77]. The use of exponential fitting for a coarse mesh to obtain optimal amounts of artificial viscosity is also a possibility, see, for example, [77].

Without going into great detail there are some other methods deserving of mention, which have been employed in solving EHL problems. The effective influence Newton method of Wang [151] uses a small region around each point in the domain to calculate each new pressure solution. In contrast, the homotopy method, used by Schlijper *et al.* [129], uses a full solve including all points in the Jacobian matrix rather than just a well chosen selection. Although the homotopy does lead to very long computing times, it is a very powerful and reliable technique. Preconditioning techniques have been used successfully in variational methods, as proposed by Verstappen [149]. Wavelet preconditioning has also been attempted for line contact solutions by Ford *et al.* [45] but the extension of this technique to two dimensions is currently unclear.

The above methods all have drawbacks, either in their applicability to both highly and lightly loaded cases, or in their computational complexity. Limiting the computational work became a necessity. Lubrecht [95,96] extended the Gauss-Seidel relaxation methods to include multigrid techniques for both line and point contacts. Multigrid techniques are described by Briggs [22] and explained in detail by Brandt [16], Hackbusch [58] and Wesseling [152]. Whilst greatly accelerating convergence of results, the deformation calculation remained $\mathcal{O}(N^2)$.

The multigrid methods of Lubrecht did not attempt to make the Gauss-Seidel scheme applicable to high loads. In 1991, Venner [140] introduced a new relaxation method. This considered the lightly and heavily loaded regions of the domain separately. In the non-contact region the Gauss-Seidel scheme was still applied, and solved in a line sense. The high pressures in the contact region mean that the elastic deformation dominates the problem here, and so Venner proposed the use of a distributive relaxation scheme in this area of the domain.

Venner's method of using different numerical schemes to solve the Reynolds Equation in different parts of the domain has been expanded by several other authors. Nurgat [108, 110] applied a Jacobi line scheme instead of the distributive scheme, in the

contact region. Ehret *et al.* [38] used the same schemes as Venner except a pointwise Gauss-Seidel scheme was applied on the cavitation boundary.

Before leaving the solution methods which have been used for the Reynolds Equation, it would be remiss not to mention a common alternative solve applied in the cavitation region. Usually the pressure here is constrained to be non-negative by setting all calculated negative pressures to be identically zero, the so-called *cavitation condition* imposed using a Christopherson approach [35]. A penalty function approach was proposed by Wu [157] and has been used since in work such as Schlijper *et al.* [129].

It is interesting to note, at this point, that the cavitation region in EHL is modelled in a very different manner to cases of modelling full journal bearings, e.g. [88, 89]. In that field the fluid in the cavitated region must be modelled since it will later become inlet lubricant. Possible approaches are described in Gwynllyw *et al.* [57] and include the modelling of the cavitation region as a continuous film of lubricant, just with a very low viscosity.

The computational work of the deformation calculation is the largest part of the numerical solve. This is particularly important for two dimensional cases. The idea of multilevel multi-integration was developed by Brandt and Lubrecht [17]. This successfully reduces the deformation calculation from $\mathcal{O}(N^2)$ to $\mathcal{O}(N \ln N)$. The first published EHL solutions using this method include Venner [140] and Venner *et al.* [147]. Multilevel multi-integration has recently been compared to Fast Fourier Transforms for analysis of rough surfaces [119] and was found to be faster when maximum or even moderate accuracy was specified. The use of superconvergence [150] for solving this integration could be possible, although there would need to be much work done in order to produce results that were much more efficient that those using multi-integration.

1.4 The Layout of this Thesis

The intention of this work is to continue the development of a fast, efficient numerical solver for EHL problems. It has been built upon the previous work by Nurgat [108] whose work was primarily concerned with a new numerical scheme, used to obtain the solution of the Reynolds Equation. This scheme is accurate enough to be employed here, however other results from [108] suggested that it was possible to improve parts of the solver. These issues will be addressed in the subsequent chapters.

The general problem, described in Section 1.1, will be formulated in a strict mathematical sense in Chapter 2. The governing equations will be described, although derivations of these well known results will not be presented. These equations will then be

non-dimensionalised. The various non-dimensional parameters used to characterise EHL cases will be quantified in relation to the physical characteristics of the real life problem. This chapter is concluded by presenting the steady state discretisation of the non-dimensional equations.

It was explained in Section 1.3 that the use of multilevel techniques had greatly improved the efficiency of EHL solvers since their introduction to the field fifteen years ago. It will be especially necessary for the multigrid work of Chapter 4 to have an understanding of the processes involved, and hence in Chapter 3 the multigrid methods used will be explained. Also in this chapter, the multilevel multi-integration process, employed in this work, is described.

The numerical solver used is described in Chapter 4. The different types of equations to be solved for an EHL solution require different numerical techniques to be used. These schemes are explained, along with how the multilevel techniques of Chapter 3 are applied in the algorithm. Example steady state solutions to a circular point contact EHL problem are shown, and the difficulties posed for numerical solvers are highlighted.

The aforementioned improvements to the solver of Nurgat [108] are also explained in Chapter 4. The scheme developed by Goodyer *et al.* [49] to improve the convergence of Nurgat's solver is described, alongside results showing its effectiveness. This is then followed by details of the changes made to the solver in order that the convergence problems experienced by the code of Nurgat, are now avoided. The chapter is concluded by a comprehensive set of results showing how the multigrid solver which has been developed does now obtain fast, numerically accurate results. The efficiency of the solver has also been shown by comparing single grid results to multigrid results. Finally, results have been shown detailing some of the possible performance increases that may be obtained by using parallelism in the solver.

In Chapter 5 the numerical solver will be extended to transient problems. Again it has been possible to make dramatic improvements on the preliminary attempts of Nurgat [108]. At the start of the chapter the steady state EHL problem of Chapter 2 is reformulated as a transient Differential-Algebraic one, before being discretised. The changes made to the solution method between steady state and transient cases are outlined. A wide selection of examples is then presented, encompassing the range of problems tackled today. These examples will be used to show how the solver is developed further for maximum efficiency and accuracy. Results will be compared to both other numerical, and experimental results which are available.

Variable timestepping for EHL problems has not, to the best of the author's knowledge, been previously attempted. It will be shown in Section 5.7 that it is possible to apply

the same techniques used in ODE problems to the highly non-linear PDE EHL problem. It will be shown that results from variable timestep runs are of similar accuracy to those of fixed step cases, and can be achieved in considerably less computational time, as larger timesteps will be taken during periods of the solve with linear behaviour, and smaller timestep sizes when the non-linear effects are dominating.

Adaptive meshing is an idea which has been used for EHL problems only once [93, 99]. Apparently this was a great success although it has not been attempted by the authors since, whose work has concentrated on the development of multilevel multi-integration. In Chapter 6 this idea is revived and combined with the solver developed thus far, without negating the speed-ups achieved by the use of multilevel multi-integration. Results are presented showing both the accuracy of the achieved results, and the decrease in the computational work. Finally a transient example is presented showing how adaptive meshes can be used in combination with variable timestepping.

This work will be concluded in Chapter 7 where the advances made are summarised. Possible future work is described here, in relation to further developments of both the grid adaptation and solver parallelism, in addition to the solution of more complicated transient EHL cases.

The computational timings given in Chapters 4 to 6 are intended for comparison with each other only. The code has been optimised in the same manner for all results. With the exception of the parallel results of Section 4.6.3 all timings have all been calculated on a single R10000 processor of an Silicon Graphics Origin 2000 shared memory machine, hence some small scale variation is possible on identical runs (see, for example, Table 4.7). The timings have been provided solely to show the performance increases achieved, rather than as a benchmark of optimum performance.

Throughout the production of this work the computational code used to generate the results has been refined and improved to achieve better efficiency, and hence results between separate sections may appear to have different times for the same example, and hence for each section a control case will be given against which the other results should be compared, rather than those in previous chapters.

Chapter 2

Governing Equations

2.1 Introduction

In this chapter the mathematical model governing EHL calculations will be presented. The only results presented in subsequent chapters will be for circular point contact problems, but reference will be made to the extra difficulties by moving from one to two dimensional cases, hence, where relevant, the differences to the equations will be highlighted.

The EHL problem is governed by two main groups of equations; that is those concerned with the physical model of the lubricant used, and those concerned with the EHL problem itself. The latter set is split up as follows:

- The Reynolds Equation. This governs the pressure distribution across the contact, for given geometry and lubricant properties.
- The Film Thickness Equation. For a given pressure distribution across the contact this defines the elastic deformation, and hence new geometry of the surfaces.
- The Force Balance Equation. This is a conservation law ensuring that the applied load across the contact is fixed, at any time.

The applications that EHL calculations are now used in mean that the properties of the actual lubricant can be as important as the physical operating conditions, such as loading. In the development of new oils, the lubricant industry needs to be able to analyse how

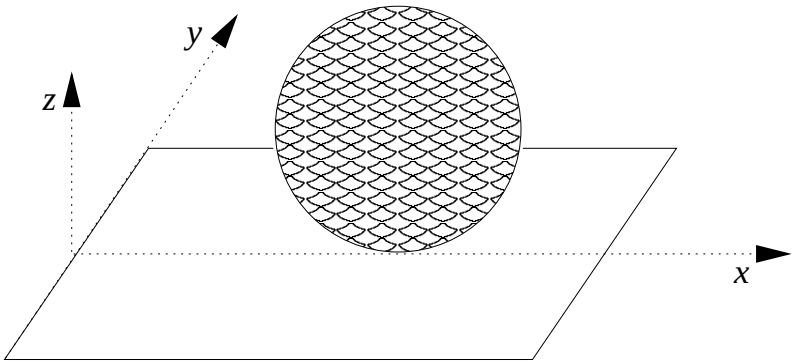


Figure 2.1: Representation of the domain of an EHL problem

the desired properties can be obtained, without manufacturing real lubricants with those properties. The need for accurate rheological models is now very important. Bearing these things in mind, however, much of this detail is not needed for this work. The overall solution properties are similar between the different models and this work is concerned with the actual numerical solution techniques used rather than the real world applicability of the solutions obtained. The techniques developed will be independent of the fluid model employed and hence an isothermal generalised Newtonian model will be used. This, then, only requires equations for the density and viscosity of the lubricant, which are both non-linear in terms of the pressure.

The equations described above will be presented in the first few sections of this chapter. They will be non-dimensionalised, using Hertzian [67] parameters, in Section 2.7. The relationships between these parameters and the non-dimensional ones of Moes [106] and Hamrock and Dowson [60] will also be presented in this section.

Before any numerical results of EHL problems can be calculated, it is necessary to discretise the equations. In Section 2.8 the steady state discretisation schemes used, are presented for the regular mesh, finite difference calculations carried out in Chapter 4. The discretisation for transient problems will be shown in Chapter 5.

2.2 Solution Domain

The circular point contact EHL problem can be considered as that of a spherical bearing on a plane, as shown in Figure 2.1. Coordinate axes directions are taken as shown, with the origin taken to be the point on the plane closest to the centre of the ball. The convention taken throughout this work will be that surface 1 refers to the ball, and surface 2 to the plane. This is only important in cases of *sliding*. This is when the speeds of the two surfaces differ. Lubricant entrainment is taken as parallel to the x-axis in the x-y plane.

The z-direction is one of the primary variables in the calculation, since it represents the geometry of the contact, h. There is assumed to be no variation of lubricant properties in this direction for the cases considered here. This is because the pressure gradient across the film in the z direction can be shown to be of the order of h/l and $h \ll l$, where l is the representative length of the contact [33]. If a viscoelastic model was employed then these gradients could be significant and would also have to be taken into account, hence a modified Reynolds Equation would have to be used.

The contact being represented need not necessarily be a ball and plate, however it is possible to reduce the real geometry to this arrangement. Throughout the rest of this chapter the word *reduced* will be taken to mean 'the quantity obtained when the geometry is transformed to the ball and plane scenario'. This will be applied to both the 'reduced radius' of the ball considered as well as the combined physical properties of the surfaces and loading conditions. Inside the contact area the undeformed geometry of the ball will be taken as parabolic in both x and y directions. Since only circular, rather than elliptical, contacts are being considered then this geometry will be axially symmetric about the z-axis.

2.3 The Reynolds Equation

The Reynolds Equation defines the pressure distribution of an applied load for a given geometry. It was derived from the Navier-Stokes equations by Osborne Reynolds in 1886. It therefore only applies to Newtonian (and generalised Newtonian) lubricants. Using Cartesian coordinates (x, y), and time t it is given by

$$\frac{\partial}{\partial x}\left(\frac{\rho h^3}{12\eta}\frac{\partial p}{\partial x} - \rho h\frac{(u_1+u_2)}{2}\right) + \frac{\partial}{\partial y}\left(\frac{\rho h^3}{12\eta}\frac{\partial p}{\partial y} - \rho h\frac{(v_1+v_2)}{2}\right) - \frac{\partial(\rho h)}{\partial t} = 0, \quad (2.1)$$

where p is the pressure
h is the geometry, or *film thickness*
ρ is the density of the lubricant
η is the viscosity of the lubricant
u_1 and u_2 are the speeds of the two surfaces in x-direction
and v_1 and v_2 are the speeds of the two surfaces in y-direction

Choosing lubricant flow parallel to the x-axis, and assuming no flow velocity variation

across the contact, then Equation (2.1) may be rewriten as

$$\frac{\partial}{\partial x}\left(\frac{\rho h^3}{\eta}\frac{\partial p}{\partial x}\right)+\frac{\partial}{\partial y}\left(\frac{\rho h^3}{\eta}\frac{\partial p}{\partial y}\right)-6u_s\frac{\partial(\rho h)}{\partial x}-12\frac{\partial(\rho h)}{\partial t}=0, \tag{2.2}$$

where the entrainment velocity, u_s is defined by $u_s = u_1 + u_2$.

This equation is referred to as having three different parts. The first two terms involving the second derivatives of pressure are called the *Poiseuille* terms. The *wedge* term is the other spatial derivative whilst the *squeeze* term is the temporal derivative.

The Reynolds Equation has no in-built concept of giving physical solutions for pressure, and hence in much of the region beyond the centre of the contact, it will be satisfied by negative solutions of pressure. In actuality, at the point of the outflow where this occurs, air pockets will have been formed in the lubricant. This can be seen in an interferometry picture such as shown in Figure 5.5. This physical effect means Equation (2.2) must be solved as a free boundary problem.

2.4 The Film Thickness Equation

The film thickness is the separation of the two surfaces in the contact. In elastohydrodynamic cases it is assumed that these surfaces are allowed to deform. This deformation is therefore dependent on the pressure applied; more particularly, the pressure distribution across the whole contact. This, therefore, is very dependent on the type of contact being modelled.

The dominant, underlying shape of the contact will be assumed to be the original parabolic shape of the contact. The presence of a lubricant - even in a hydrodynamic case, with no deformation - will separate the contacts by an extra scalar quantity, referred to as the central offset film thickness, h_{00}. The final term describing the film thickness is the deformation which, in the reduced geometry, is all taken to be in the curved surface rather than the plane.

In one dimension (i.e. the line contact case) the film thickness equation is given by

$$h(x) = h_{00} + \frac{x^2}{2R_x} - \frac{4}{\pi E'}\int_{-\infty}^{\infty} \ln\left|\frac{x-x'}{x_0}\right| p(x')\,\mathrm{d}x', \tag{2.3}$$

whereas in two dimensions, for the point contact case, it is

$$h(x,y) = h_{00} + \frac{x^2}{2R_x} + \frac{y^2}{2R_y} - \frac{2}{\pi E'}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{p(x',y')\,\mathrm{d}x'\mathrm{d}y'}{\sqrt{(x-x')^2+(y-y')^2}}, \tag{2.4}$$

where E' is the reduced elastic modulus of the contact

and R_x and R_y are the reduced radii of curvature in the x and y directions.

In this work only circular contacts will be considered, so $R_x = R_y$.

2.5 The Lubricant Model

In the EHL problem there are very large changes in the pressures in the lubricant over very small distances. It is therefore important to use models of the lubricant which can accurately model this behaviour. When codes, such as the one developed here, are applied to model real life situations then it is often to investigate particular properties of the oils. Much research is currently being done into the use of non-Newtonian models, as outlined in Section 1.2, however, for this work, generalised Newtonian models will suffice.

2.5.1 Density Equation

The density model that will be used throughout this work, except where stated, is that of Dowson and Higginson [33]. This commonly used model takes account of compressibility of the lubricant and is given by

$$\rho(p) = \rho_0 \frac{0.59 \times 10^9 + 1.34p}{0.59 \times 10^9 + p}, \tag{2.5}$$

where ρ_0 is the density at ambient pressure. This relationship is locally pressure dependent in that there is no knowledge of the pressure distribution from the surrounding areas affecting the density at any given point. An alternative expression sometimes seen in the literature is

$$\rho(p) = \rho_0 \left(1 + \frac{0.59 \times 10^{-9} p}{1 + 1.7 \times 10^{-9} p}\right), \tag{2.6}$$

although the two are almost mathematically identical.

2.5.2 Viscosity Equation

The viscosity of the lubricant is very important in EHL contacts. For high pressure cases it is necessary to use an accurate model which applies at pressures of up to at least 1 GPa. The usual Newtonian model used is that of Roelands [123]. This has recently been tested further against experimental results and has shown to be a good comparison up to loads of 400 MPa, however beyond this point the correlation is not always as accurate [52]. It is, however, still much better at higher loads than the previously employed model of

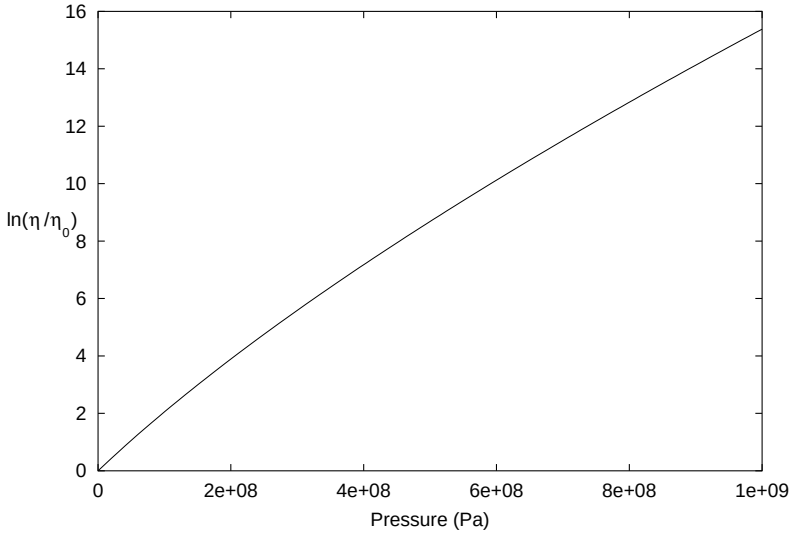


Figure 2.2: Relative viscosity (η/η_0) against pressure

Barus [11] (1893) which only had a linear dependence on pressure inside the exponential term. Using the Roelands model, the viscosity is defined by

$$\eta(p) = \eta_0 \mathrm{e}^{\left\{\frac{\alpha p_0}{z_i}\left[-1+\left(1+\frac{p}{p_0}\right)^{z_i}\right]\right\}}, \tag{2.7}$$

where η_0 is the viscosity at ambient pressure
p_0 is a constant (typically 1.98×10^8)
z_i is the pressure viscosity index, taken as $z_i = 0.68$
and α is the pressure viscosity coefficient given by

$$\alpha = \frac{1}{\eta}\left[\frac{\partial \eta}{\partial p}\right]_{p=0}. \tag{2.8}$$

Once again, it is clear that this relationship depends on the pressure. This is shown in Figure 2.2 where the relative viscosity, η/η_0, is plotted against increasing pressure, up to 1 GPa.

2.6 The Force Balance Equation

The Force Balance Equation is needed to ensure conservation of applied force over the contact. This is because the load applied across the contact must be completely carried by the lubricant film, since it is assumed that the fluid film does not break down to give even

a partially dry contact. For the point contact case this is expressed as

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} p(x,y)\,\mathrm{d}x\mathrm{d}y = F \tag{2.9}$$

where F is the applied load.

2.7 Non-dimensionalisation

For isothermal point contact EHL problems the system of dimensional equations to be solved is defined by equations (2.2), (2.4), (2.5), (2.7), and (2.9). The orders of numbers in these equations vary dramatically, from $\mathscr{O}(10^9)$ for the pressure, down to $\mathscr{O}(10^{-8})$ for the minimum film thickness. To numerically compute solutions, without incurring artifacts from the floating point arithmetic of the computer used, non-dimensionalisation is highly advisable.

In the process of non-dimensionalisation, dimensionless quantities may also be computed. These values are commonly used to characterise the individual case being studied. These relations are explored in Sections 2.7.1 and 2.7.2.

The first parameters to be introduced are the *maximum Hertzian pressure*, and the *Hertzian radius*, which are derived from Hertz's theory for dry contacts [67]. This assumes a pressure distribution of

$$p(x,y) = \begin{cases} p_h\sqrt{1-\frac{x}{a}^2-\frac{y}{a}^2} & \left|x^2+y^2\right| < a. \\ 0 & \text{otherwise.} \end{cases} \tag{2.10}$$

The contact has now been assumed to be circular rather than generally elliptical, i.e. $R_x = R_y$. The maximum Hertzian pressure, p_h is given by

$$p_h = \frac{3F}{2\pi a^2}, \tag{2.11}$$

and the Hertzian radius, a, by

$$a^3 = \frac{3FR_x}{2E'}. \tag{2.12}$$

Using Equations (2.11) and (2.12), along with the density and viscosity at ambient pressure, the EHL system described in Sections 2.3 to 2.6 can be non-dimensionalised. This is done by substituting into the equations the dimensionless variables:

$$X = \frac{x}{a} \qquad Y = \frac{y}{a} \qquad \overline{\rho} = \frac{\rho}{\rho_0} \qquad \overline{\eta} = \frac{\eta}{\eta_0}$$

$$P = \frac{p}{p_h} \qquad H = \frac{hR_x}{a^2} \qquad T = \frac{tu_s(0)}{2a}.$$

The Reynolds Equation (2.2) becomes

$$\frac{\partial}{\partial X}\left(\frac{\overline{\rho}H^3}{\overline{\eta}}\frac{\partial P}{\partial X}\right) + \frac{\partial}{\partial Y}\left(\frac{\overline{\rho}H^3}{\overline{\eta}}\frac{\partial P}{\partial Y}\right) - \lambda\frac{u_s}{u_s(0)}\frac{\partial(\overline{\rho}H)}{\partial X} - \lambda\frac{\partial(\overline{\rho}H)}{\partial T} = 0, \tag{2.13}$$

where the dimensionless parameter λ is given by

$$\lambda = \frac{6\eta_0 R_x^2 u_s(0)}{a^3 p_h}. \tag{2.14}$$

Defining another dimensionless quantity, ε, by

$$\varepsilon = \frac{\overline{\rho}H^3}{\overline{\eta}\lambda}, \tag{2.15}$$

means that the non-dimensional Reynolds Equation, (2.13) can be rewritten as

$$\frac{\partial}{\partial X}\left(\varepsilon\frac{\partial P}{\partial X}\right) + \frac{\partial}{\partial Y}\left(\varepsilon\frac{\partial P}{\partial Y}\right) - \frac{u_s}{u_s(0)}\frac{\partial(\overline{\rho}H)}{\partial X} - \frac{\partial(\overline{\rho}H)}{\partial T} = 0. \tag{2.16}$$

Assuming oil entrainment in the positive X direction, this has boundary conditions given by $P(X{=}-\infty) = P(Y{=}\infty) = P(Y{=}-\infty) = 0$ and a free boundary cavitation condition ensuring $P \geq 0$ inside the domain.

The Film Thickness Equation (2.4) non-dimensionalises to give

$$H(X,Y) = H_{00} + \mathscr{G}(X,Y) + \frac{2}{\pi^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\frac{P(X',Y')\,\mathrm{d}X'\mathrm{d}Y'}{\sqrt{(X-X')^2+(Y-Y')^2}}, \tag{2.17}$$

for given undeformed surface geometry $\mathscr{G}(X,Y)$. Assuming smooth surfaces, as in most of the examples to follow, this is given by

$$\mathscr{G}(X,Y) = \frac{X^2}{2} + \frac{Y^2}{2}. \tag{2.18}$$

Non-dimensionalising the Force Balance equation (2.9) removes explicit dependence on the applied force to give

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} P(X,Y)\,\mathrm{d}X\,\mathrm{d}Y = \frac{2\pi}{3}. \tag{2.19}$$

Finally, the density and viscosity take the form:

$$\overline{\rho}(P) = \frac{0.59 \times 10^9 + 1.34 p_h P}{0.59 \times 10^9 + p_h P}, \tag{2.20}$$

and

$$\overline{\eta}(P) = \mathrm{e}^{\left\{ \frac{\alpha p_0}{z_i} \left[-1 + \left(1 + \frac{p_h P}{p_0} \right)^{z_i} \right] \right\}}. \tag{2.21}$$

2.7.1 Moes Parameters

The EHL problem described thus far, can be characterised by a reduced number of non-dimensional variables which relate to the wide ranges of physical parameters available in experiments. The two common sets used for point contact cases, such as used here, are the *Moes parameters*, M and L, described in this section, and the *Hamrock and Dowson parameters* described in Section 2.7.2. These parameters are defined differently between line and point contact cases, but since only the point contact case will be used in this work, this will be all that is presented. More detailed descriptions of the relationships between these sets are provided, for example, in [140]. The examples presented in the rest of this work will, for those cases using the expressions for density and viscosity described in Section 2.5, always be presented in terms of both the Moes and the Hamrock and Dowson parameters.

There are six physical parameters that are to be reduced: α, E', η_0, R_x, F and u_s. These can be combined, as in Equations (2.11) and (2.12), to get expressions for p_h and a respectively. Also defining

$$\overline{\alpha} = \alpha p_h, \tag{2.22}$$

and recalling Equation (2.14) for λ, the Moes parameters, M and L [106, 107] may then be defined by:

$$\overline{\alpha} = \frac{L}{\pi} \left(\frac{3M}{2} \right)^{\frac{1}{3}}, \tag{2.23}$$

and

$$\lambda = \left(\frac{128\pi^3}{3M^4} \right)^{\frac{1}{3}}. \tag{2.24}$$

2.7.2 Hamrock and Dowson

Dowson and Higginson [33] characterised the line contact problem in terms of three non-dimensional parameters, W, U, and G, for load, speed, and material parameters respectively. In 1976, Hamrock and Dowson [60] similarly defined the following relations for circular point contact problems:

$$G = \alpha E'. \tag{2.25}$$

$$U = \frac{\eta_0 u_s}{2E'R_x} \tag{2.26}$$

$$W = \frac{F}{E'R_x^2} \tag{2.27}$$

These parameters may be related to the Moes parameters using the expressions

$$M = W(2U)^{-\frac{3}{4}}, \tag{2.28}$$

and

$$L = G(2U)^{\frac{1}{4}}. \tag{2.29}$$

Clearly, to relate three parameters to just two will require one of G, U and W, to be chosen as fixed.

Hamrock and Dowson [64] later calculated expressions for central and minimum film thickness in an elliptic contact, in terms of G, U and W. These are

$$H_{cen} = 2.69U^{0.67}G^{0.53}W^{-0.067}\left(1-0.61\mathrm{e}^{-0.73\kappa}\right), \tag{2.30}$$

$$H_{min} = 3.63U^{0.68}G^{0.49}W^{-0.073}\left(1-\mathrm{e}^{-0.68\kappa}\right), \tag{2.31}$$

for ellipticity ratio κ.

2.8 Discretisation

In order to solve the EHL system given by Equations (2.16) to (2.21) by a numerical scheme, it is first necessary to discretise them. In this section these discrete equations will be built up from first principles. This process requires that a set of sample points inside the domain are chosen at which the equations will be satisfied. The more points chosen, the closer the points will be to each other and hence the more accurate the solution. However, as the number of mesh points increases, so does the amount of computational work

required to solve the system.

In this work the domain $X \in [X_a, X_b]$, $Y \in [Y_a, Y_b]$ is represented by a regularly spaced mesh of N_X x N_Y nodes. It is then a simple matter to relate $u_{i,j}$ to being the value of solution variable u at the mesh point (i,j) where $1 \leq i \leq N_X$, $1 \leq j \leq N_Y$. Here the coordinates of point (i,j) are easily calculated by

$$X_i = X_a + (i-1)\Delta X, \tag{2.32}$$

and

$$Y_j = Y_a + (j-1)\Delta Y, \tag{2.33}$$

where ΔX and ΔY are the mesh point spacings in the X and Y directions respectively.

The numerical solution method used in this work to solve the Reynolds Equation (2.16) is that of *finite differences*. Using this technique requires only the solutions at the mesh points, and numerical derivatives are calculated using neighbouring solutions. The simplest method of doing this is by using a *first order backward difference* scheme. This defines the derivative of solution variable u at mesh point (i,j) by

$$\left.\frac{\mathrm{d}u}{\mathrm{d}x}\right|_{(i,j)} \approx \frac{u_{(i,j)} - u_{(i-1,j)}}{\Delta x}, \tag{2.34}$$

i.e. the gradient between (i,j) and its upstream neighbour $(i-1,j)$. Similarly the *first order forward difference* scheme is given by

$$\left.\frac{\mathrm{d}u}{\mathrm{d}x}\right|_{(i,j)} \approx \frac{u_{(i+1,j)} - u_{(i,j)}}{\Delta x}. \tag{2.35}$$

These two equations form the basic building blocks for all finite difference schemes.

There are several schemes commonly employed to solve the Reynolds Equation. The simplest steady state version - used throughout Chapter 4 - is first order, and is defined by using backward differences. Equation (2.16)becomes:

$$\begin{aligned} &\frac{\varepsilon_{i-\frac{1}{2},j}\left(P_{i-1,j} - P_{i,j}\right) + \varepsilon_{i+\frac{1}{2},j}\left(P_{i+1,j} - P_{i,j}\right)}{(\Delta X)^2} \\ +&\frac{\varepsilon_{i,j-\frac{1}{2}}\left(P_{i,j-1} - P_{i,j}\right) + \varepsilon_{i,j+\frac{1}{2}}\left(P_{i,j+1} - P_{i,j}\right)}{(\Delta Y)^2} \\ &-\frac{\rho_{i,j}H_{i,j} - \rho_{i-1,j}H_{i-1,j}}{\Delta X} = 0, \end{aligned} \tag{2.36}$$

where

$$\begin{aligned}
\varepsilon_{i+\frac{1}{2},j} &= \frac{\varepsilon_{i+1,j}+\varepsilon_{i,j}}{2},\\
\varepsilon_{i-\frac{1}{2},j} &= \frac{\varepsilon_{i-1,j}+\varepsilon_{i,j}}{2},\\
\varepsilon_{i,j+\frac{1}{2}} &= \frac{\varepsilon_{i,j+1}+\varepsilon_{i,j}}{2},\\
\text{and} \qquad \varepsilon_{i,j-\frac{1}{2}} &= \frac{\varepsilon_{i,j-1}+\varepsilon_{i,j}}{2}.
\end{aligned} \tag{2.37}$$

The boundary conditions are prescribed with all exterior boundaries having P=0, and the line j=1 being a symmetry condition in the Y direction.

An alternative steady state formulation is the *second order upstream* discretisation used by Venner and Lubrecht, e.g. [145]:

$$\begin{aligned}
&\frac{\varepsilon_{i-\frac{1}{2},j}\left(P_{i-1,j}-P_{i,j}\right)+\varepsilon_{i+\frac{1}{2},j}\left(P_{i+1,j}-P_{i,j}\right)}{(\Delta X)^2}\\
+&\frac{\varepsilon_{i,j-\frac{1}{2}}\left(P_{i,j-1}-P_{i,j}\right)+\varepsilon_{i,j+\frac{1}{2}}\left(P_{i,j+1}-P_{i,j}\right)}{(\Delta Y)^2}\\
&-\frac{3\rho_{i,j}H_{i,j}-4\rho_{i-1,j}H_{i-1,j}+\rho_{i-2,j}H_{i-2,j}}{2\Delta X} \;=\; 0,
\end{aligned} \tag{2.38}$$

This is valid for all points in the domain with i>2. For i=2 the first order scheme of Equation (2.36) is used. The only difference between Equations (2.36) and (2.38) is the order of the derivatives used in the discretisation of the wedge term.

Transient discretisations are undertaken in similar fashions, and will be considered in Chapter 5.

Assuming that the undeformed surfaces have geometry, $\mathscr{G}$, given by Equation (2.18), then the Film Thickness Equation (2.17) is discretised to give:

$$H_{i,j} = H_{00} + \frac{{X_i}^2}{2} + \frac{{Y_j}^2}{2} + \sum_{k=1}^{N_X}\sum_{l=1}^{N_Y} K_{i,j,k,l}P_{k,l} \tag{2.39}$$

where X_i and Y_j are as defined by Equations (2.32) and (2.33) respectively, and K is the *kernel* matrix. The kernel is the analytic solution of the double integral required for the deformation calculation given by:

$$K_{i,j,k,l} = \frac{2}{\pi^2}\left\{\left|X_p\right|\sinh^{-1}\left(\frac{Y_p}{X_p}\right)+\left|Y_p\right|\sinh^{-1}\left(\frac{X_p}{Y_p}\right)\right.$$

$$
\begin{aligned}
&- |X_m| \sinh^{-1}\left(\frac{Y_p}{X_m}\right) - |Y_p| \sinh^{-1}\left(\frac{X_m}{Y_p}\right) \\
&- |X_p| \sinh^{-1}\left(\frac{Y_m}{X_p}\right) - |Y_m| \sinh^{-1}\left(\frac{X_p}{Y_m}\right) \\
&+ |X_m| \sinh^{-1}\left(\frac{Y_m}{X_m}\right) + |Y_m| \sinh^{-1}\left(\frac{X_m}{Y_m}\right) \Bigg\} \qquad (2.40)
\end{aligned}
$$

where

$$
\begin{aligned}
X_p &= X_i - X_k + \frac{\Delta X}{2}, \\
X_m &= X_i - X_k - \frac{\Delta X}{2}, \\
Y_p &= Y_j - Y_l + \frac{\Delta Y}{2}, \\
\text{and} \qquad Y_m &= Y_j - Y_l - \frac{\Delta Y}{2}. \qquad (2.41)
\end{aligned}
$$

The Force Balance Equation (2.19) discretises to give:

$$
\Delta X \Delta Y \sum_{i=1}^{N_X} \sum_{j=1}^{N_Y} P_{i,j} = \frac{2\pi}{3}. \qquad (2.42)
$$

The equations for density and viscosity are pointwise calculations, Equations (2.20) and (2.21) respectively, and are thus simply discretised to give

$$
\overline{\rho}_{i,j} = \frac{0.59 \times 10^9 + 1.34 p_h P_{i,j}}{0.59 \times 10^9 + p_h P_{i,j}}, \qquad (2.43)
$$

and

$$
\overline{\eta}_{i,j} = \mathrm{e}^{\left\{ \frac{\alpha p_0}{z_i} \left[-1 + \left(1 + \frac{p_h P_{i,j}}{p_0} \right)^{z_i} \right] \right\}}. \qquad (2.44)
$$

with all symbols as defined previously.

Chapter 3

Multilevel Techniques

3.1 Introduction

Multilevel techniques are used to provide solutions of the same accuracy significantly faster than could be achieved on just one very fine grid. These techniques have been around since the late 1970s [16], being developed primarily for boundary value problems. The field of multilevel techniques goes far beyond just multigrid. Other areas operating on a multiscale basis include wavelet transforms (such as used in [45]) and multi-resolution, e.g. [15].

Since their introduction, multigrid use has increased dramatically, from such fields as disparate as quantum chemistry and electrostatics [12] to hurricane tracking [46]. Their first application to EHL problems came in 1986 with the work of Lubrecht *et al.* [95]. Since this time they have become generally accepted as an effective method for getting EHL solutions quickly.

The two multilevel techniques to be used in this work are explained in this chapter. First *Multigrid* is explained in Section 3.3. It will be seen in Chapter 4 how this accelerates convergence of the numerical solver for the Reynolds Equation. Section 3.3 includes examples of different types of multigrid cycles, and explains the advantages of using a full multigrid start. Secondly, *Multilevel Multi-Integration* is considered in Section 3.4. This is used to greatly reduce the computational time spent in calculating the surface deformation. Both multigrid and multi-integration are explained in terms of their general

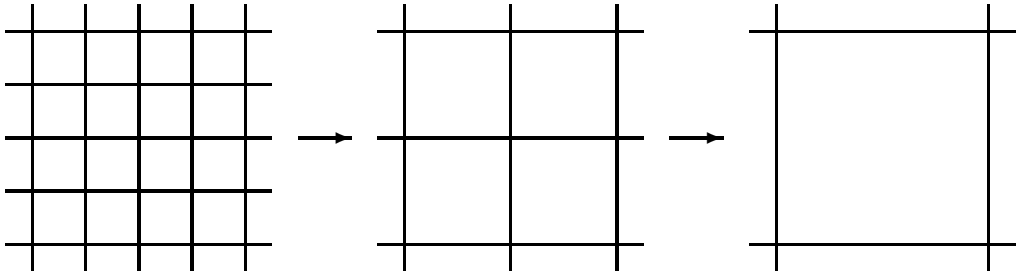


Figure 3.1: Hierarchy of grids

methods and properties although only those aspects of the theory used in later chapters will be explained in detail. There are several common pieces of notation between the two techniques which will be described, first, in Section 3.2.

There are now several detailed books on the application of multilevel methods. For multigrid, the tutorial by Briggs [22] is a good introduction to the subject and has recently been updated to include more advanced techniques. A more detailed look at multigrid is provided by Hackbusch [58], Trottenberg *et al.* [139], and by Wesseling [152]. Both multigrid and multilevel multi-integration for EHL problems are described in detail by Venner and Lubrecht in [145]. This chapter provides a summary of the techniques utilised in the solver described in Chapter 4.

3.2 Multilevel Formulation

In Chapter 2 a discrete system of mathematical equations has been presented. The aim is to obtain the solution of this system as efficiently as possible on a fine mesh. Assuming that this fine mesh is a regular grid and has (2^k+1) x (2^k+1) points, then it can be referred to as *grid k*. A hierarchy of grids with decreasing values of k as shown in Figure 3.1 may then be defined. Grid points on a coarser grid l are separated by a distance $2^{k-l}\delta^k$ where δ^k is the separation of points on grid k, in the appropriate direction. Hence all points on grid $k-1$ will have coincident points on grid k, with the additional points on grid k being the mid-points between coincident points.

Rather than just referring to a solution vector, $\underline{u}$ say, it will be necessary to define a solution vector for each individual grid level. Hence on grid k such a solution will be represented by $\underline{u}^k$. The operators employed for transferring solutions from one grid to another will be explained in Section 3.3.3.

3.3 Multigrid

This work is not intended to cover multigrid techniques in their entirety. There are numerous works devoted to providing a more thorough investigation, e.g. [22,58,152]. However, it is necessary to explain the basic methods being used before continuing to solve the EHL problem. Only those parts used in this solver will be explained here.

The motivation for the use of multigrid techniques is briefly outlined in Section 3.3.1. Although there are various types of multigrid method (see e.g. [58]), the only one of interest in this work is the *Full Approximation Scheme* (FAS). This is because, as has been shown in Chapter 2, EHL problems are inherently very non-linear, and hence the simple *Correction Scheme* cannot be employed. The FAS will be explained in Section 3.3.2.

It will be shown that the method of transfer of both solutions and errors between grids will be important to the usefulness and efficiency of the scheme. The operators used will be presented in Section 3.3.3. The types of multigrid cycle used, as well as the mechanics of the process (the correction scheme), will explained in Section 3.3.4. Finally, the use of the Full Multigrid algorithm to obtain a good initial solution on the finest mesh will be outlined in Section 3.3.5.

3.3.1 Introduction to Multigrid

Consider a simple case where the system to be solved is given by the one-dimensional equation

$$\mathscr{L}\underline{u} = f, \tag{3.1}$$

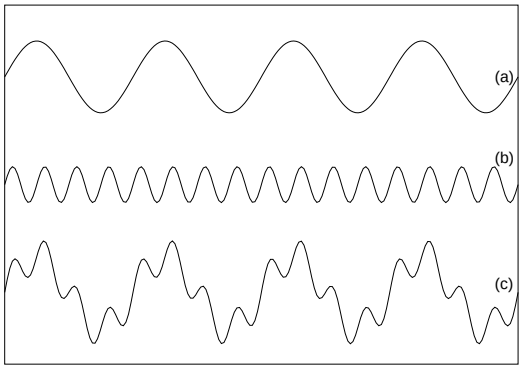
where $\mathscr{L}$ is the differential operator defining the system, and f is the right hand side function. This system needs to be solved in order to obtain a vector $\underline{u}$, an approximation to the discretised numerical solution, on a regular mesh of N points (where $N = 2^k + 1$ for some k), separated by distance δx^k. Hence, this problem can now be represented as

$$\mathscr{L}^k\underline{u}^k = f^k. \tag{3.2}$$

At any particular stage, a solution vector $\underline{\tilde{u}^k}$ will have been calculated, which is an approximate solution to Equation (3.1) with an error, $\underline{e}^k$, such that

$$\underline{u}^k = \underline{\tilde{u}}^k + \underline{e}^k. \tag{3.3}$$

The solution is relaxed iteratively to obtain new solutions which are (hopefully) better than the previous one. The aim of the solution process is to reduce this error to below



(a) Low frequency component (b) High frequency component (c) Two-phase error

Figure 3.2: Example of errors of different frequencies.

some pre-specified tolerance level.

It has been shown (e.g. [22,140] as well as Section 4.6.1.1) that the relaxation schemes used to solve EHL problems are very good at reducing high frequency errors, but very slow at reducing low frequency errors. This idea is illustrated in Figure 3.2 where a high frequency sine wave is imposed on a low frequency one. Standard single grid smoothing techniques would quickly eliminate errors of similar frequencies to the mesh spacing, but the lower frequency error components could be almost unchanged.

Multigrid is a technique to try and combat this problem. Given that the smoother is able to reduce errors of the frequency of the grid size, then lower frequency errors can be reduced by using a (coarser) grid with similar order to that of the error. In the case shown in Figure 3.2 a coarser grid with mesh spacing four times more than that used to reduce the fine grid error, would be appropriate.

3.3.2 Full Approximation Scheme (FAS)

From Equation (3.2) the residual, r^k, of the system can be calculated from

$$r^k = f^k - \mathscr{L}^k \tilde{\underline{u}}^k, \tag{3.4}$$

for an approximation $\tilde{\underline{u}}^k$ to $\underline{u}^k$. Substituting for f^k from Equation (3.2) gives

$$r^k = \mathscr{L}^k \underline{u}^k - \mathscr{L}^k \tilde{\underline{u}}^k. \tag{3.5}$$

The non-linearity of the operator $\mathscr{L}^k$ means this cannot be directly factorised. However, using Equation (3.3) enables us to define the residual as

$$r^k = \mathscr{L}^k(\tilde{\underline{u}}^k + \underline{e}^k) - \mathscr{L}^k\tilde{\underline{u}}^k, \tag{3.6}$$

which can be reordered to give

$$\mathscr{L}^k(\tilde{\underline{u}}^k + \underline{e}^k) = \mathscr{L}^k\tilde{\underline{u}}^k + r^k. \tag{3.7}$$

Consider now a coarser grid, grid j. To represent Equation (3.7) in the same form as (3.2), i.e.

$$\mathscr{L}^j\hat{\underline{u}}^j = \hat{f}^j, \tag{3.8}$$

it is necessary to define $\hat{\underline{u}}^j$ by

$$\hat{\underline{u}}^j = I_k^j\tilde{\underline{u}}^k + \underline{e}^j, \tag{3.9}$$

The term $\hat{f}^j$ in Equation (3.8) is called the *FAS right hand side*, and is given by

$$\hat{f}^j = \mathscr{L}^j(I_k^j\tilde{\underline{u}}^k) + I_k^j r^k, \tag{3.10}$$

and I_k^j is an inter-grid transfer operator from grid k to grid j, to be described in Section 3.3.3.

The solution $\hat{\underline{u}}^j$ to Equation (3.8) can be approximated by $\overline{\underline{u}}^j$ which can then be used to calculate the coarse grid approximation to the error by

$$\underline{\tilde{e}}^j = \overline{\underline{u}}^j - I_k^j\tilde{\underline{u}}^k. \tag{3.11}$$

This is then used to update the fine grid solution in the following manner:

$$\tilde{\underline{u}}^k \leftarrow \tilde{\underline{u}}^k + I_j^k(\overline{\underline{u}}^k - I_k^j\tilde{\underline{u}}^k). \tag{3.12}$$

3.3.3 Grid Transfer

It is now necessary to define the operators for transferring solutions between grids. To move to a coarser grid a *coarsening* or *restriction* operator is needed. The transfer from grid k to grid $k-1$ will be denoted by I_k^{k-1}. Similarly, moving to a finer grid needs a *prolongation* operator, denoted by I_{k-1}^k. These operators will be defined in terms of *stencils* describing how the new pointwise solution is constructed. Efficient multigrid solvers are very reliant on the choice of correct intergrid operators. It is especially important for

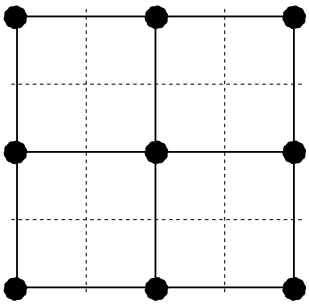


Figure 3.3: Injection coarsening operator, I_k^{k-1}

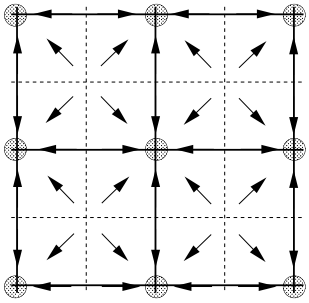


Figure 3.4: Full weighting coarsening operator, I_k^{k-1}

linear problems that the two operators are transposes of each other.

To define a coarsening operator there are two common choices. Either *injection* or *full weighting*. The easiest of these is injection which is simply direct transfer of the solution at coincident points between fine and coarse grids. This stencil is given mathematically, in two dimensions, for the coarse grid points, by

$$I_k^{k-1} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \tag{3.13}$$

and shown in Figure 3.3. Full weighting involves a weighted average of the surrounding fine grid points:

$$I_k^{k-1} = \frac{1}{16} \begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix}, \tag{3.14}$$

illustrated in Figure 3.4. Full weighting can be more desirable because the highest frequency components are not represented on the coarser grids, improving the Coarse Grid Correction process.

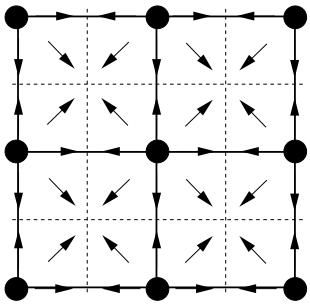


Figure 3.5: Prolongation operator, I_{k-1}^{k}

The prolongation operator to be used linearly interpolates the coarse grid function to the fine mesh, hence at coincident points injection will be used, whilst at other points linear interpolation of either two or four coarse grid points will be employed. This is shown in Figure 3.5 and given by

$$I_{k-1}^{k} = \frac{1}{4}\begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix}. \tag{3.15}$$

This is the transpose of the coarsening operator described by Equation (3.14).

3.3.4 Multigrid Cycles

The multigrid process is the combination of the individual tools described above. Assuming that the same iterative process can be used to solve the coarse grid system as the fine grid system, then the finest grid will be used to smooth the highest frequency errors, and progressively coarser grids used to smooth errors of progressively lower frequencies, before returning to get an updated solution on the finest mesh. The smoothing cycles done before coarsening are called *pre-smooths* and those done after prolongation and correction of the solution are referred to as *post-smooths*.

The simplest multigrid cycle is the V-cycle. An example of this is shown in Figure 3.6 which shows one cycle over four levels of mesh. An initial approximation on the finest grid has ν_1 pre-smooths before being coarsened. This is then repeated until the coarsest mesh is reached where ν_0 smoothing cycles are done. The solution on the next finer mesh is then corrected according to Equation (3.12) before having ν_2 post-smooths. Again this process is repeated until a corrected, smoothed solution is reached on the finest mesh. This V-cycle is known as a V(ν_1,ν_2)-cycle. Typical values for ν_1 and ν_2 are three or less, although ν_0 may be much larger in order to obtain a much better coarse grid representation

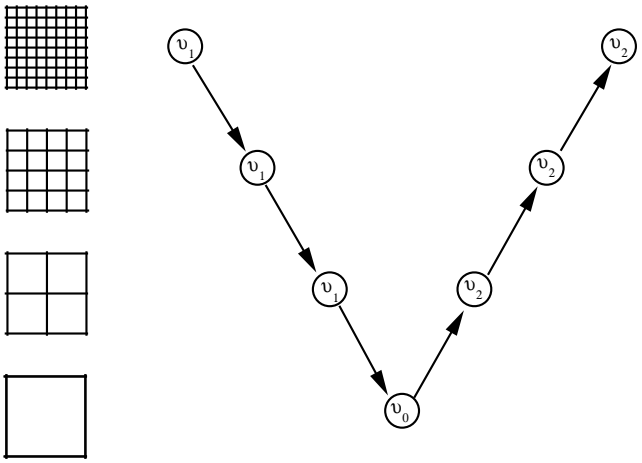


Figure 3.6: A multigrid V-cycle

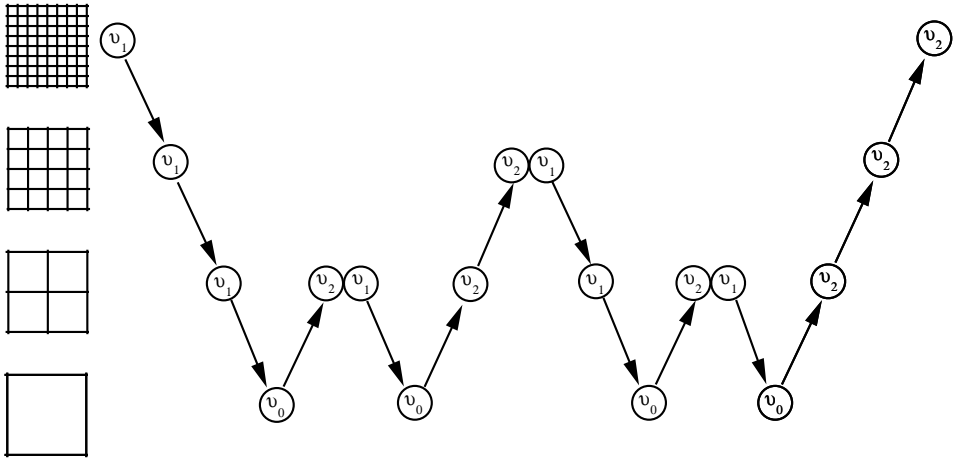


Figure 3.7: A multigrid W-cycle

of the solution.

An alternative multigrid cycle is the W(ν_1,ν_2)-cycle. This is where two coarse grid correction cycles are used to correct the solution on each grid rather than just the one of the V-cycle. An example, again on four levels, is shown in Figure 3.7. The advantage of the W-cycle over the V-cycle is that there are twice as many coarse grid corrections for each level per multigrid cycle.

3.3.5 Full Multigrid

The process of *Full Multigrid* (FMG) is designed to eliminate the large errors which would exist on the fine grid, before it is first used. Solutions, especially for the EHL point contact problem solved here, become more computationally expensive to calculate, the more mesh points there are present in the domain. Hence the ability to use a solution

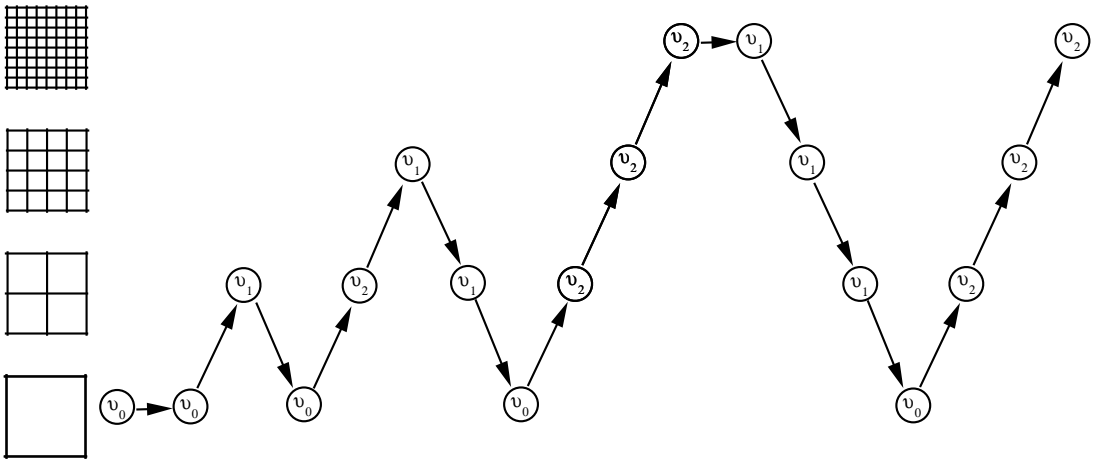


Figure 3.8: Full multigrid, with one V-cycle per level

prolonged from a coarser grid as an initial approximation will be clearly advantageous.

FMG uses the same multigrid techniques and V- or W-cycles as described above in Section 3.3.4, but applied as shown in Figure 3.8. This example demonstrates just one V-cycle per level, but there will usually be several to obtain a reasonably converged solution. At the end of each set of V-cycles this solution is then prolonged up to a new finest grid. For the EHL problem it will be seen in Chapter 4 how beneficial this is compared to simply starting off with the Hertzian approximation, given by Equation (2.10), to the pressure on the finest grid level employed.

3.4 Multilevel Multi-Integration

Multilevel multi-integration is a multiscale technique designed to significantly speed up the evaluation of integrals. In this work it will be employed for calculation of the deformation of the surfaces, given mathematically by the double integral in Equation (2.17). Besides its use in EHL, it also has applications in integral equations, integro-differential equations, elasticity problems and acoustic problems. The process behind the derivation and application of these techniques is explained by Brandt, Lubrecht and Venner in [17, 19, 140, 145] in much greater detail than need be provided here.

In one space dimension, a general example would be to solve the following integral:

$$w(x) = \int_{\Omega} K(x,y)u(y)\,\mathrm{d}y, \tag{3.16}$$

where the domain $\Omega = (a,b)$. The function K is referred to as the *kernel* and its discretised form as the *kernel matrix*. Multi-integration is applicable in situations where the kernel is

a dense rather than a sparse matrix. This means that Equation (3.16) actually represents a full matrix-vector multiplication. Multi-integration for higher dimensional cases than the one dimensional case to be considered here, is applied by using the same algorithms described below, applied separately to each dimension.

This method is only possible where the kernel itself has sufficient smoothness properties. For the EHL point contact problem being considered in this work it is necessary to remember that Equation (2.40) is, in fact, singular around the point (i, j). Therefore K is not smooth around this point, although far enough away it is smooth enough to apply multi-integration techniques. This means that applying multi-integration over the whole domain Ω will not be correct, unless the region around the singularity, Ω_{sing} is corrected afterwards.

Assume now that Equation (3.16) is discretised on a regular mesh of n_x^h points, separation $\delta x^h = (b-a)/(n_x^h - 1)$, for grid level h. The single grid method for calculating this multi-summation at each $x_i^h = a + (i-1)\delta_x^h$, for $i = 1, \ldots, n_x$ is

$$w_i^h = w^h(x_i^h) = \sum_{j=1}^{n_x^h} K_{i,j}^{hh} u_j^h, \tag{3.17}$$

where $K_{i,j}^{hh}$ is the discretised kernel and u_i^h is the approximation to $u^h(x_i)$ Hence for every point in Ω, this calculation is $\mathscr{O}(n_x^h)$ meaning that the evaluation of the summation for the whole domain is $\mathscr{O}\left((n_x^h)^2\right)$.

Outside the influence of any singularity, multi-integration is defined by two stages for the formulation of the coarse grid equation. First, an approximation to w_i^h must be calculated including only summation of coarse grid points (Equations (3.18) to (3.22)). Then the second dimension of the kernel matrix will be included using the knowledge that it displays similar behaviour (Equations (3.23) to (3.26)).

For ease of notation let grid H be grid $h-1$, i.e. one level coarser, with coarse grid indices I and J coincident with fine grid indices i and j respectively. Only two levels will be considered for the formulation of multi-integration, although it will be shown how these techniques can be applied iteratively, similar to the ideas of multigrid.

Consider first the coarse grid kernel, $K_{i,J}^{hH}$, being the fine grid kernel evaluated at coarse grid points. This is therefore, clearly, just the fine grid values injected upwards by $K_{i,J}^{hH} \equiv K_{i,2J-1}^{hh}$. The fine grid kernel can then be approximated by using an interpolation of a high enough order, defined by

$$\tilde{K}_{i,j}^{hh} \equiv \left[I_H^h K_{i,\cdot}^{hH} \right]_j, \tag{3.18}$$

where the dot $(\cdot)$ refers to interpolation in the y direction (the dummy variable in Equation 3.16)), and the final j is the new index.

It is then possible to rewrite Equation (3.17) as

$$w_i^h = \sum_{j=1}^{n_x^h} \tilde{K}_{i,j}^{hh} u_j^h + \sum_{j=1}^{n_x^h} \left(K_{i,j}^{hh} - \tilde{K}_{i,j}^{hh} \right) u_j^h, \tag{3.19}$$

which can be broken down as follows. Only coarse grid points have been used to construct $\tilde{K}_{i,j}^{hh}$ and hence at coincident fine grid points $\left(K_{i,j}^{hh} - \tilde{K}_{i,j}^{hh} \right) \equiv 0$. This just leaves the non-coincident points, which will have an interpolation error of $K_{i,j}^{hh} - \tilde{K}_{i,j}^{hh}$ which is of the order of $K^{(2p)}(\xi)$ where $2p$ is the order of the interpolation, and $K^{(2p)}(\xi)$ is the $2p^{\text{th}}$ derivative of K at some point ξ in Ω. The assumption that the kernel is sufficiently smooth compared to u means that the discretisation error is large enough for the interpolation error to be discounted. Thus, Equation (3.19) becomes

$$w_i^h \approx \tilde{w}_i^h = \sum_{j=1}^{n_x^h} \tilde{K}_{i,j}^{hh} u_j^h = \sum_{j=1}^{n_x^h} \left[I_H^h K_{i,\cdot}^{hH} \right]_j u_j^h = \sum_{J=1}^{n_x^H} K_{i,J}^{hH} \left[\left(I_H^h \right)^{\mathrm{T}} u_\cdot^h \right]_J, \tag{3.20}$$

where the adjoint of the interpolation matrix, known as the *anterpolation matrix* is given by $\left(I_H^h\right)^{\mathrm{T}}$. Hence, defining a coarse grid representation of the fine grid solution by

$$u_J^H \equiv \left[\left(I_H^h \right)^{\mathrm{T}} u_\cdot^h \right]_J, \tag{3.21}$$

the coarse grid integration, Equation (3.20) is reduced to

$$w_i^h \approx \tilde{w}_i^h = \sum_{J=1}^{n_x^H} K_{i,J}^{hH} u_J^H, \tag{3.22}$$

which is an approximation to the fine grid integration without increasing the complexity of the algorithm from $\mathscr{O}(n_x^h)$.

Applying a similar process to the x direction allows this integration to be reduced still further. Similar to Equation (3.18) the fine grid kernel may be approximated in the x direction to $\hat{K}_{i,j}^{hh}$, by

$$\hat{K}_{i,j}^{hh} \equiv \left[I_H^h K_{\cdot,j}^{Hh} \right]_j, \tag{3.23}$$

with all symbols defined as before. Therefore, Equation (3.17) can be rewritten as

$$w_i^h = \left[I_H^h w_\cdot^H\right]_i + \sum_{j=1}^{n_x^h} \left(K_{i,j}^{hh} - \hat{K}_{i,j}^{hh}\right) u_j^h, \tag{3.24}$$

where

$$w_I^H \equiv \sum_{j=1}^{n_x^H} K_{I,j}^{Hh} u_j^h. \tag{3.25}$$

Again, as $K_{I,j}^{Hh}$ is an injection of the coincident fine points onto the coarser mesh, then $\left(K_{i,j}^{hh} - \hat{K}_{i,j}^{hh}\right) \equiv 0$ at these points. Also, if K is sufficiently smooth in the x direction, the interpolation error should be sufficiently small to neglect the summation in Equation (3.24) completely, thus becoming:

$$w_i^h \approx \left[I_H^h w_\cdot^H\right]_i. \tag{3.26}$$

Therefore, provided K is sufficiently smooth in both x and y directions, combining Equations (3.22) and (3.26) gives:

$$w_i^h \approx \left[I_H^h w_\cdot^H\right]_i \approx \left[I_H^h \sum_{J=1}^{n_x^H} K_{\cdot,J}^{hH} u_J^H\right]_i, \tag{3.27}$$

which defines multi-integration for smooth kernels.

If, however, the kernel is singular, and hence not smooth in a region $\Omega_{sing} \subset \Omega$ then a different expression for w_i^h is required. To build this up the coincident and non-coincident points between grids h and H in the x direction will be considered separately.

Considering, first, the coincident points in the x direction, Equation (3.19) can be rewritten as

$$w_i^h = \sum_{j=1}^{n_x^h} \tilde{K}_{i,j}^{hh} u_j^h + \sum_{(i,j)\in\Omega_{sing}} \left(K_{i,j}^{hh} - \tilde{K}_{i,j}^{hh}\right) u_j^h + \sum_{(i,j)\notin\Omega_{sing}} \left(K_{i,j}^{hh} - \tilde{K}_{i,j}^{hh}\right) u_j^h. \tag{3.28}$$

Since only coincident points are being calculated, in the smooth region, as before, the final term's sum can be considered to be zero. Note also that at these points,

$$\sum_{j=1}^{n_x^h} \tilde{K}_{i,j}^{hh} u_j^h = \sum_{J=1}^{n_x^H} \tilde{K}_{I,J}^{HH} u_j^h = w_I^H. \tag{3.29}$$

Hence,

$$w_i^h = w_I^H + \sum_{(i,j)\in\Omega_{sing}} \left(K_{i,j}^{hh} - \tilde{K}_{i,j}^{hh}\right) u_j^h. \tag{3.30}$$

For non-coincident points in the x direction, recalling Equation (3.24) and reducing the smooth part of the domain as before, then the value of w_i^h is given by

$$w_i^h = \left[I_H^h w_\cdot^H\right]_i + \sum_{(i,j)\in\Omega_{sing}} \left(K_{i,j}^{hh} - \tilde{K}_{i,j}^{hh}\right) u_j^h. \tag{3.31}$$

Therefore the fine grid solution of the integration in Equation (3.16) can be approximated using multi-integration by first calculating the coarse grid multi-summation, and then correcting around the singularity, as given by Equation (3.30), before interpolating the coarse grid multi-summation to the non-coincident points on the fine grid, and correcting again, as given by Equation (3.31).

The problem of deciding which points $(i,\ j)$ are in Ω_{sing}, thus requiring correction, is important both in terms of solution accuracy (by having enough points) and optimal efficiency (by not having too many). For each point i in the one dimensional example above, the region requiring correction may be defined by $\Omega_{sing}^i = \{j \in \Omega : |i-j| < m\}$. The choice of the radius, m, is problem dependent. In [17] Brandt and Lubrecht showed that, for the kernel of the EHL line contact problem, $m \sim \ln n$, should be used. For the two dimensional circular point contact EHL case, Equation (2.40) this linear idea per point was extended to give a rectangular shape, hence it is corrected over a $(2m_1+1)\times(2m_2+1)$ rectangle. They found that taking $m_1 \sim \ln n_x$ in the interpolation direction, and $m_2 = 2$ perpendicular to the interpolation, then the algorithm maintains its $\mathcal{O}(N \ln N)$ efficiency. Elliptic problems were considered by Wijnant in [154] where the different directions have different weights, meaning that the area requiring correction has different sizes in the X and Y directions, for ellipticity $\kappa \neq 1$.

Multi-integration need not — and should not — just be restricted to evaluation of the multi-summation on one level coarser. It may be applied recursively, provided that the kernel matrix still has sufficient smoothness properties on the coarser grids. This is illustrated in Figure 3.9 where four levels of grid are used for the multi-integration process. The larger the number of coarser grids which are used, then the greater the possible speed up, although there is no point in going beyond a grid with $\sqrt{N}$ points, where N is the number of the points on the finest level, because this summation is already $\mathcal{O}(N)$, which is the same order as the intergrid operators required to coarsen further.

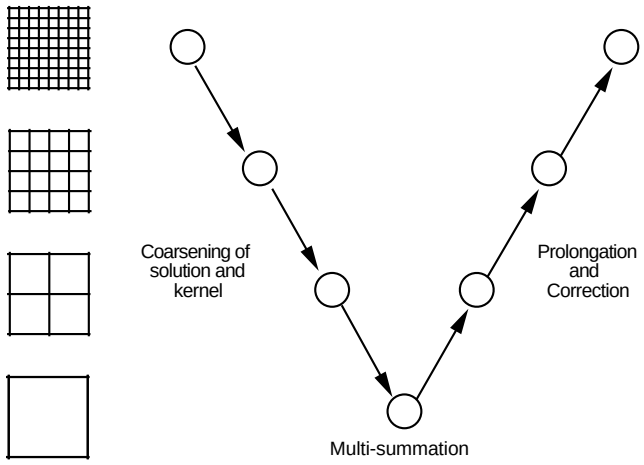


Figure 3.9: Multilevel multi-integration being applied over four levels

Chapter 4

Solving EHL Problems

4.1 Introduction

EHL problems have been calculated numerically since the first results of Petrusevich [117] in 1951. Industry demands results for EHL calculations quickly, to assist in the design and analysis of the performance and wear of components and lubricants. The enormous advances in computing power over the last half century have enabled more and more demanding problems to drive current research. The speed of the code is obviously an important issue. Since first being applied to EHL problems less than fifteen years ago, by Lubrecht *at al.* [95, 96], multigrid has been widely recognised as the way forward for obtaining quick numerical solutions. The multilevel multi-integration technique of Brandt and Lubrecht [17] in 1990, for solving the film thickness equation decreased the solution time still further by reducing the order of the deformation calculation from $\mathcal{O}(N^2)$ to $\mathcal{O}(N \log N)$ where N is the number of mesh points in the computational domain.

The equation system which needs to be solved is that described in Chapter 2. Once discretised on the required mesh, various different techniques need to be employed to obtain an accurate numerical approximation to the solution. These techniques are described in Section 4.2.

An example of a steady state EHL problem solution is presented in Section 4.3. This is, as with the entirety of this work, for a two dimensional circular point contact problem. This example is provided to show the typical solution profiles, and highlight some of the

numerical difficulties in obtaining solutions.

This work aims to build on the numerical techniques already used in EHL. With this in mind, the methods used to solve the equations presented in Chapter 2 are described in Section 4.2. This work has been developed around an earlier attempt by Nurgat [108] which was based on the FDMG Multigrid Software [133]. This solver is outlined in Section 4.4. The multigrid results obtained by Nurgat for steady state problems [108,110] were similar to those he obtained using a single grid scheme. These were, in turn, similar to the published results of Ehret *et al.* [38], Venner [140] and Wang [151]. The solver described in [108] did, however, have a number of deficiencies in the accuracy of the solutions obtained, some of which are highlighted in [110]. In Section 4.5 these problems will be explained in detail before describing the series of improvements employed to avoid them, including those of Goodyer *et al.* described in [49]. The improvement in accuracy is shown using a series of examples.

The performance of the solver is very important, and any alterations to the solution scheme must be justified by an increase in performance, but without a loss of accuracy. These factors are considered in Section 4.6 where the advantages of the multilevel techniques described in Chapter 3 are illustrated. Also considered in Section 4.6.3 are performance benefits which may be obtained by the application of parallel computing to the problem.

4.2 Solution Scheme

The numerical solution of the EHL problems considered here requires the solution of the system of partial differential integro-equations outlined in Chapter 2. The solution scheme for this system needs to be convergent and stable, as well as being as quick as possible. For the lubricant models being used, the solution of Equations (2.43) and (2.44) for the density and viscosity respectively, is a simple pointwise calculation at every node in the mesh. However more complicated algorithms are necessary to obtain accurate converged solutions for the other equations.

The finite difference solution methods used to solve the finite difference form of the Reynolds Equation (2.36) will be explained in Section 4.2.1. The Film Thickness Equation (2.39) will be solved using the multilevel techniques described in Section 3.4, and these will be applied in Section 4.2.2. Finally, the solution scheme for the Force Balance Equation will be discussed in Section 4.2.3.

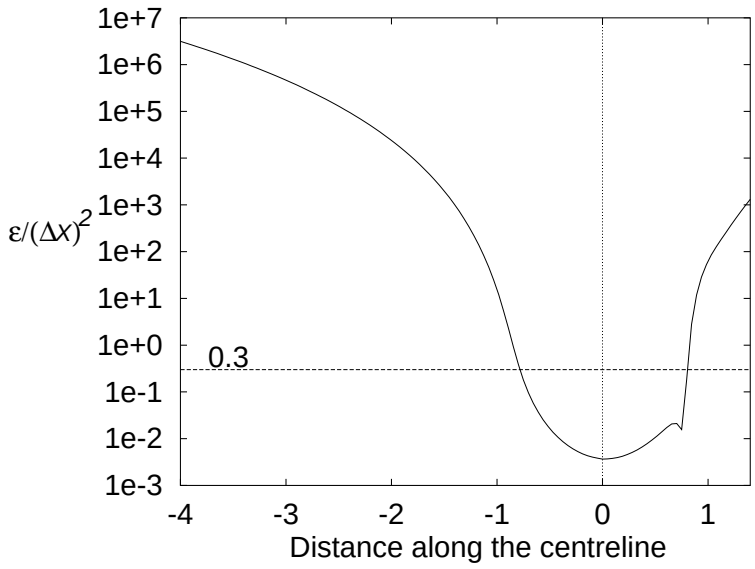


Figure 4.1: Example plot of $\varepsilon/(\Delta X)^2$ used to locate the edge of the contact region

4.2.1 Numerical Solution of the Reynolds Equation

The EHL point contact problem has three very distinct regions of interest. These are the contact region, where the pressure is very high; the non-contact region, where the pressure is very small; and the cavitation region where the Reynolds Equation (2.16) is not valid because a negative, hence non-physical, pressure would be predicted. To solve the discretised Reynolds Equation numerically (for pressure) these three regions need different mathematical schemes to be employed. There are various papers considering these options, such as [38, 108], so only the methods used in this work will be described here.

The first step is to decide on where the boundary between the contact and non-contact region is located. We do this in the manner described by Venner [140]. Here the criterion used is based on the value of ε from Equation (2.15). From Figure 4.1 it can be seen that where $\varepsilon/(\Delta X)^2 = 0.3$ it is a good correlation to the area inside the contact region.

In the non-contact region the Gauss-Seidel line scheme is used. This region has low pressures and hence there is very little deformation of the surface. This means that the previously calculated values for the film thickness are still valid approximations and hence the most recent pressure information can be used for updating the pressure.

In the contact region, the pressures are very large and hence large deformations of the surfaces occur. The wedge term, $\frac{\partial \overline{\rho} H}{\partial X}$, is dominant here and thus there is a very close relationship between the solutions of pressure and film thickness. This means that only current information for film thickness, density and viscosity, should be used in this

area. There are two schemes which have been commonly used: the distributed relaxation scheme, developed by Venner [140], and the Jacobi line relaxation of Nurgat *et al.* [108, 110]. It is the Jacobi line scheme which is used here. Several Newton iterations of the Gauss-Seidel scheme are typically done before updating the contact region.

In both the contact and non-contact regions the smoothing process will take place subject to a relaxation parameter limiting the size of changes made to the solution. These parameters will typically be different between the two regions, and will need controlling to ensure that solutions converge and optimum performance is achieved.

The Reynolds Equation does not hold in the cavitation region. However it is important that this region is not ignored totally because the code may move mesh points "in and out" of the cavitation region as it tries to compute the free boundary position. There are several options available in the treatment of this region. The method used here is to calculate the pressure solution only at pressure-positive points by imposing a *cavitation condition*. This means that for any negative pressures calculated, these values are set to be zero, following the Christopherson approach used by Dowson and Taylor [35].

4.2.2 Numerical Solution of the Film Thickness Equation

The discretised film thickness equation (2.39) to be solved for smooth circular contact cases, is

$$H_{i,j} = H_{00} + \frac{X_i^2}{2} + \frac{Y_j^2}{2} + \sum_{k=1}^{N_X} \sum_{l=1}^{N_Y} K_{i,j,k,l} P_{k,l} \tag{4.1}$$

with the matrix K defined as in Equation (2.40). This has three parts. The quadratic terms, representing the undeformed parabolic geometry, are clearly defined for every point in the mesh. The central offset film thickness, H_{00}, is a scalar displacement variable which will be calculated as described in Section 4.2.3. The third term, the double sum, is the deformation term. Hence for every point in the mesh, the film thickness is based on a multi-summation of all the other points in the mesh.

Calculating the double sum is computationally a very expensive process. For instance, a single deformation calculation on a 257x129 (half) grid would have $257 \times 129 \times 257 \times 257 \approx 2 \times 10^9$ multiplications. Unsurprisingly this leads to very long run times for problems on these fine grids (see Section 4.6.2 for details). This high calculation time was one of the inhibiting factors in the solution of point contact EHL problems.

The smoothness of the kernel matrix allows the use of Multilevel Multi-Integration, as described in Section 3.4. The work of Brandt and Lubrecht [17] for line and point contact kernel matrices was extended to the actual solution of EHL problems in [147], and for

elliptic EHL problems by Wijnant [154]. The implementation used here has been done using the details from Venner and Lubrecht [145]. An optimised version of the multi-summation algorithm was developed by Fairlie [43], and the difference in performance will also be compared in Section 4.6.2.

4.2.3 Numerical Solution of the Force Balance Equation

The Force Balance Equation (2.19) is a conservation law for pressure. This means that any pressure solution from the Reynolds Equation must also satisfy this equation. The difference between the calculated sum of pressures and the non-dimensional 'target' of $\frac{2\pi}{3}$, may be used to relax the central offset film thickness, H_{00} in the manner described by Venner [140]. This solves both the problems of how to include the Force Balance Equation, and how to calculate the correct value of H_{00}.

The relaxation for H_{00} is therefore defined by:

$$H_{00} \leftarrow H_{00} + c\left(\frac{2\pi}{3} - \Delta X \Delta Y \sum_{i=1}^{N_X}\sum_{j=1}^{N_Y} P_{i,j}\right), \tag{4.2}$$

where c is a small relaxation parameter, typically of the order of 0.05, although different values may be chosen on different grid levels; finer grids tend to require smaller relaxation parameters. Venner and Lubrecht [145] show how they reduce this relaxation parameter when using multigrid W-cycles rather than V-cycles.

4.3 Example Solutions

The EHL problem described thus far is characterised by various physical parameters. These represent a particular set of operating conditions, for particular materials in contact with a particular lubricant. These inputs are then combined to give a set of non-dimensional parameters which may be used to characterise the solution, as described in Section 2.7.

In this section, a particular set of input parameters will be represented to show typical solutions for the variables across the domain. The non-dimensional quantities they correspond to are shown in Table 4.1. Employing the Dowson and Higginson density equation (2.20) and the Roelands viscosity equation (2.21) implies a Newtonian model of fluid behaviour.

The solution for the pressure obtained across the domain is shown in Figure 4.2 where, as in all of the examples to follow, unless otherwise stated, lubricant entrainment is from

Parameter	Value
Viscosity index, α	2.1×10^{-8} Pa
Maximum Hertzian pressure, p_h	0.64 GPa
Material parameter, G	4729
Load parameter, W	4.7×10^{-7}
Speed parameter, U	1.0×10^{-11}
Moes parameter, M	50
Moes parameter, L	10

Table 4.1: Non-dimensional parameters for an example EHL solution

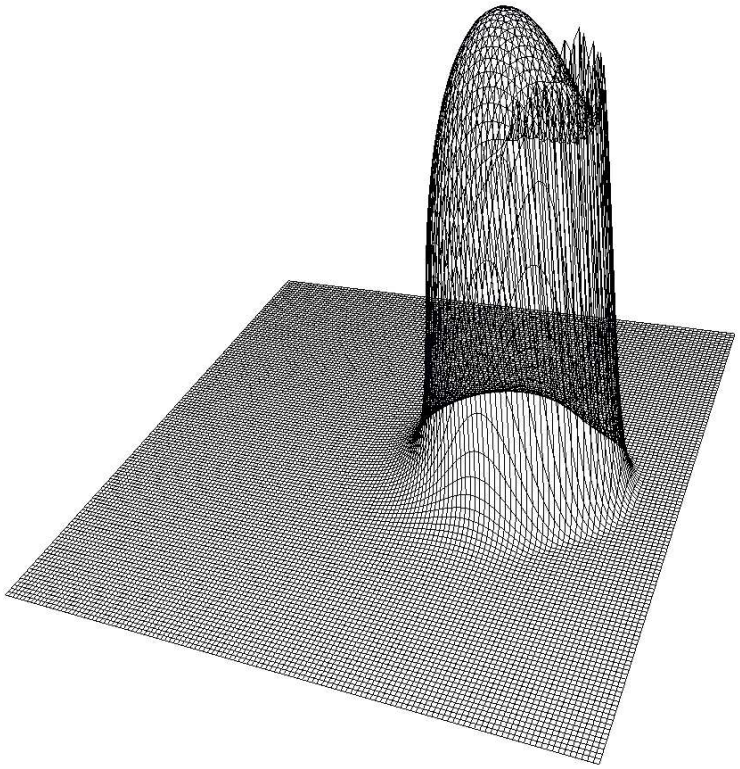


Figure 4.2: Example pressure solution across an EHL point contact. Non-dimensionalised range: $0\rightarrow 1.03$

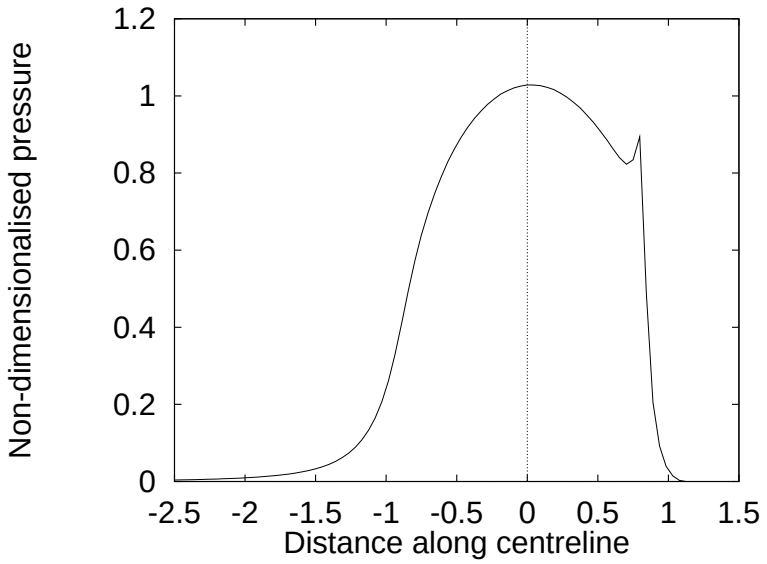


Figure 4.3: Example pressure solution along the centreline of an EHL point contact

left to right, parallel to the X-axis. This shows the three solution areas. Most noticeable is the contact area, in a unit radius circular area about (0,0). The profile along the centreline is shown in Figure 4.3. These pictures also show the Petrusevich pressure ridge/spike. This is a physical feature which is only present for Newtonian lubricants. Although it does not appear smooth in the picture it is a continuous smooth ridge. The saw-tooth behaviour seen is due to the resolution of the grid used to generate the figures.

On the right hand side of Figure 4.3 the cavitation region is clearly visible. This is where the Reynolds equation had calculated a negative solution, representing a non-continuous film of lubricant in the outflow of the contact, and the pressure has been set to zero. The rest of the domain is the non-contact region, where the pressure is low, but non-zero.

The film thickness is shown in Figure 4.4 where there is clear deformation from the undeformed parabolically circular shape. The centreline solution is plotted in Figure 4.5. Notable parts of this solution are the relative flatness of the contact inside the contact area, that the minimum film thickness is not found on either centreline, and that a constriction in film thickness at the outflow occurs after the pressure spike at the end of the contact region.

Solution plots for density and viscosity are shown in Figures 4.6 and 4.7 respectively. These show exactly how much the properties of the lubricant change across the contact. The dependence upon the pressure is very clear with the influence of the pressure ridge being observed in both cases.

Figure 4.4: Example film thickness solution across an EHL point contact. Minimum non-dimensionalised film thickness: 0.14

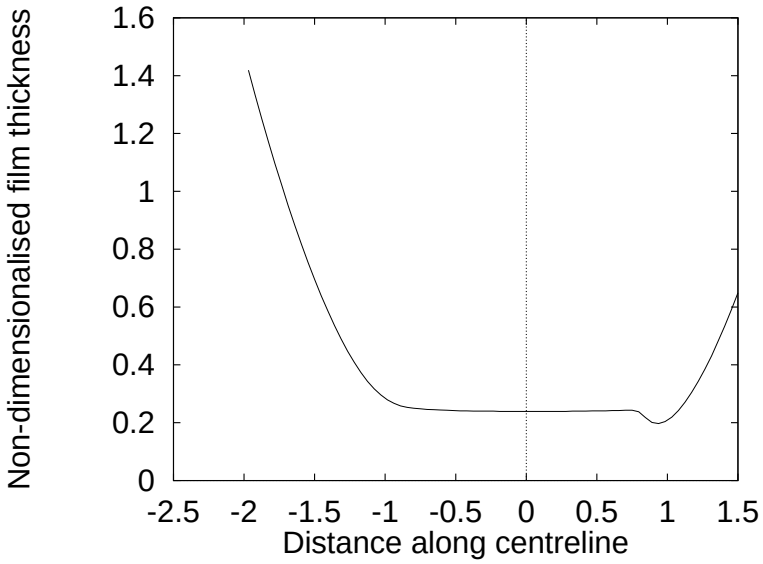


Figure 4.5: Example film thickness solution along the centreline of an EHL point contact

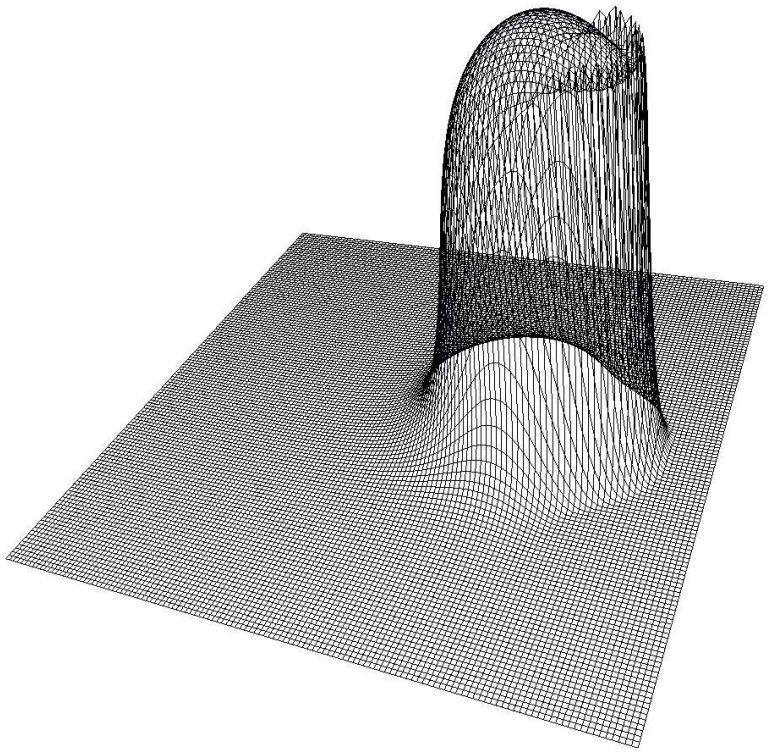


Figure 4.6: Example density solution across an EHL point contact. Non-dimensionalised range $1 \rightarrow 1.18$

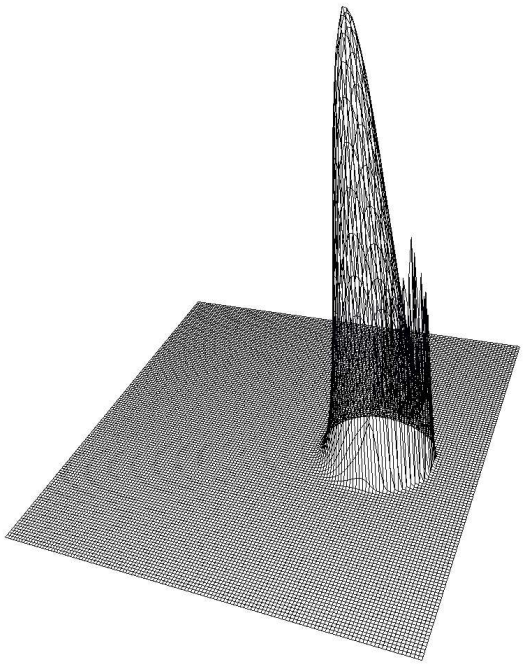


Figure 4.7: Example viscosity solution across an EHL point contact. Non-dimensionalised range: $1 \rightarrow 33500$

4.4 Solution Algorithm

The code used here to solve EHL point contact problems has been developed on that of Nurgat [108] which was itself built on the FDMG software of Shaw [133]. This is a finite difference multigrid code. The choice of V- or W-multigrid cycles is left to the user. What follows is a description of Nurgat's code.

Referring back to the schematics from Chapter 3, Figures 3.6 to 3.8, on each grid level the code will perform one (or more) relaxations. Each smooth consists of one, or more, iterations of a solver for each of the five discrete equations in the system, namely Equations (2.36, 2.39, 2.42, 2.43, and 2.44).

First, the Reynolds Equation (2.36) is solved to give a new solution for pressure. The system to be solved has an inherently full Jacobian matrix, but the most significant contributions come from the terms with pressure derivatives. This, combined with the strong directionality of the problem, allows the update of pressure to be solved using just a tridiagonal Jacobian matrix, for the first order discretisation of Equation (2.36).

Once a new pressure solution has been obtained, the Force Balance Equation (2.42) is relaxed to obtain a new corrected value for H_{00} as shown by Equation (4.2). This is then used in the calculation of the film thickness distribution. Finally, the density and viscosity solutions are updated.

Using the multigrid techniques of Chapter 3, the coarsest grid is used first, before progressing to the finest grid using the Full Multigrid (FMG) technique. After arriving on the finest grid for the first time, the chosen multigrid cycle is employed until a solution of sufficient accuracy is reached. For coarsening on boundaries, injection of the pressure solution was used, as described in Venner [140].

The initial approximation is usually the Hertzian pressure profile, described in Equation (2.10), although it is possible to use a continuation solution from a previously run case. This technique is especially useful in highly loaded cases where good solutions can be computationally expensive to obtain. Another technique employed is rather than having the non-contact region identically zero, as in Equation (2.10), it is set to an arbitrary small value, to ensure it is not mistaken for the cavitation region on the first smooth. This first solve of the Reynolds Equation gives a sensible 'shape' and position for the cavitation boundary.

Iterations	Hcent	Hmin	RMSRES	ΣP	ΔP_m
1	0.1950	0.1110	3.9001E-04	2.0832	2.091E-02
5	0.1928	0.1040	1.7222E-04	2.1144	2.884E-03
10	0.1927	0.1038	1.1407E-04	2.1196	2.693E-03
15	0.1927	0.1038	9.5991E-05	2.1201	2.630E-03
20	0.1927	0.1038	9.1076E-05	2.1202	2.621E-03

Table 4.2: Nurgat *et al.* Multigrid Test Problem One, M=99 & L=16.

Iterations	Hcent	Hmin	RMSRES	ΣP	ΔP_m
1	0.4612	0.3076	1.3773E-02	2.0842	1.377E-02
5	0.4529	0.3057	1.6256E-04	2.0909	8.322E-04
10	0.4526	0.3054	7.3911E-05	2.0904	2.452E-04
15	0.4525	0.3053	4.3010E-05	2.0905	2.251E-04
20	0.4525	0.3053	3.6674E-05	2.0905	2.236E-04
25	0.4525	0.3053	3.6051E-05	2.0905	2.234E-04

Table 4.3: Nurgat *et al.* Multigrid Test Problem Two, M=20 & L=10.

4.5 Improvements to Nurgat's Scheme

4.5.1 Why are Improvements Needed?

The scheme described above is that implemented in [108–110]. It was shown in these works that this does give results for central and minimum film thickness which are similar to the previously published work of Ehret *et al.* [38], Venner [140] and Wang [151]. However it is clear from the results shown in [110] (reproduced here as Tables 4.2 and 4.3) that although the sum of pressures has converged, it has done so to an incorrect value. The value for the sum of the pressures (ΣP), should be $\frac{2\pi}{3}$, which is 2.0943, correct to four decimal places. This result is important because not only is the sum of pressures the Force Balance condition, but it is also used to relax the central offset film thickness which is then used in the film thickness calculation.

The RMSRES column shows the calculated root mean square residual for the Reynolds Equation. This is only for a given film thickness solution, hence an error in the film thickness may not stop these being small. However it can also be seen that these residuals are not falling to machine precision with increased cycles.

These test problems were not especially heavily loaded - Test Problem Two had a maximum Hertzian pressure of only 0.58 GPa. As the loading is increased the 'converged' solution becomes further from the correct sum of pressures, hence further from the true solution. This can be seen by considering the differences between Test Cases One and Two, where the maximum Hertzian pressure in Test Case One is 1.2 GPa. The

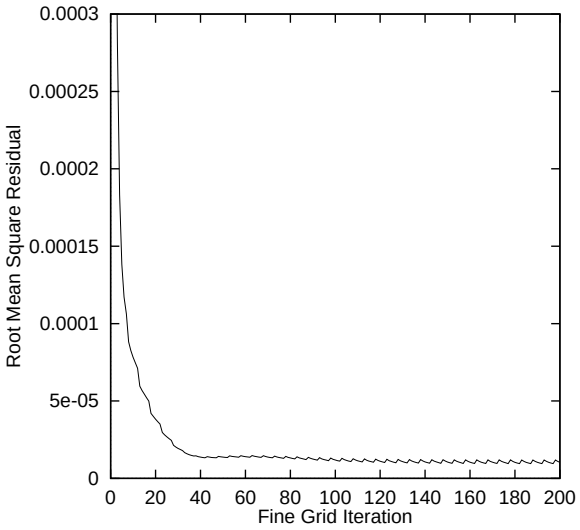


Figure 4.8: Multigrid Stalling exhibited by the code of Nurgat

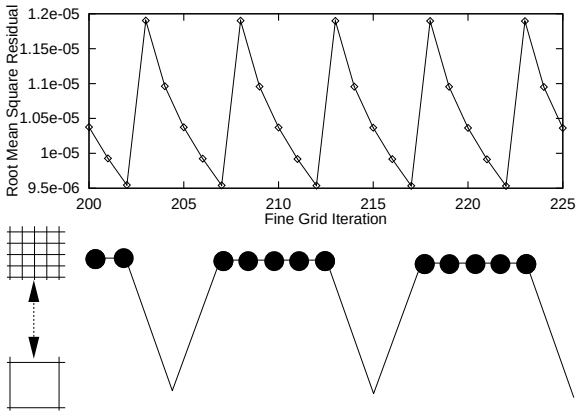


Figure 4.9: Multigrid Stalling Saw-tooth behaviour

convergence rate of the solver appears to slow down until it reaches a *'stalling point'* beyond which further multigrid cycles do not improve the accuracy of the solution. This is illustrated in Figure 4.8 which shows the level of the root mean square of the residuals (on the finest grid) falling steadily until it reaches the stalling point. It is clearly seen that no amount of further multigrid iterations will produce a more accurate solution.

Figure 4.8 does show that once the solution has reached this particular level, the magnitude of the residual exhibits *saw-tooth* behaviour. This is enlarged in Figure 4.9 where an indication of the reason behind this stalling is shown. The problem is that the process of making the coarse grid correction to the fine grid solution is adding errors which are equally balanced by the smoothing carried out on the finest grid. Looking again at Tables 4.2 and 4.3 the final column, ΔP_m highlights this, being as it is the change in pressure solution over a multigrid cycle: the same changes are being made each time.

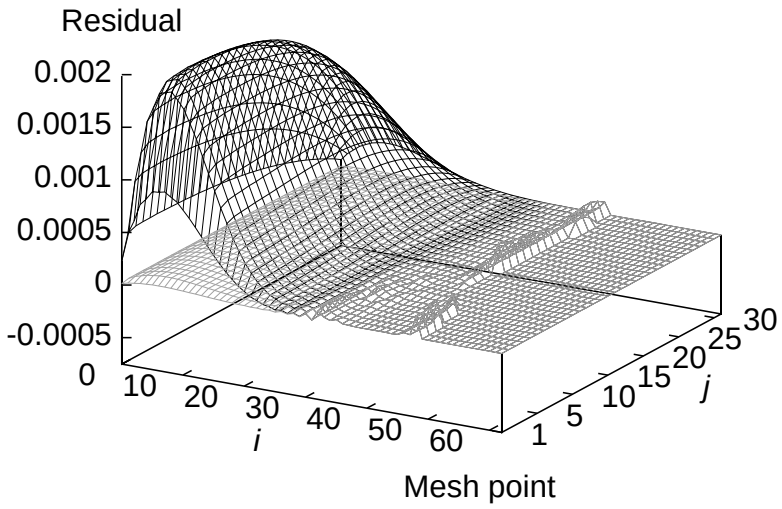


Figure 4.10: Residual levels across the half domain showing errors on the cavitation boundary are not reduced.

4.5.2 Alterations to Solution Scheme

4.5.2.1 The Stalling Problem

The problems described above are not normally associated with multigrid techniques. However, the EHL problem is highly non-linear and requires the FAS method to be used, as described in Section 3.3.2. The main cause of problems is the presence of the cavitation region. This is, as has been previously stated, considered to be a free boundary whose position must be allowed to move. Therefore, the coarse mesh solution may be inherently different to the fine mesh solution on the edge of the cavitation region because the position of the free boundary may move half a coarse mesh cell (one fine cell). This means that when interpolating back, the new solution is introducing an error at this boundary. This is shown in Figure 4.10 which shows the pointwise residual levels across the half domain at two separate stages in the solution process - both after returning from the coarse grid correction (CGC) process. The more prominent, bolder surface shows the early stages of convergence where residual levels across the whole domain are noticeable. The lower, lighter surface shows that most of the error has been smoothed away except that exactly the same error is reappearing on the cavitation boundary. This error is then smoothed away on the fine grid but is reintroduced the next time the CGC is made. This means there is a stage when the errors smoothed away on the finest grid are equally balanced by the errors added to the solution by the CGC process.

It was seen in [108] that the solver of Nurgat did successfully obtained converged solutions when applied only to the finest grid. Attention was therefore drawn to the cor-

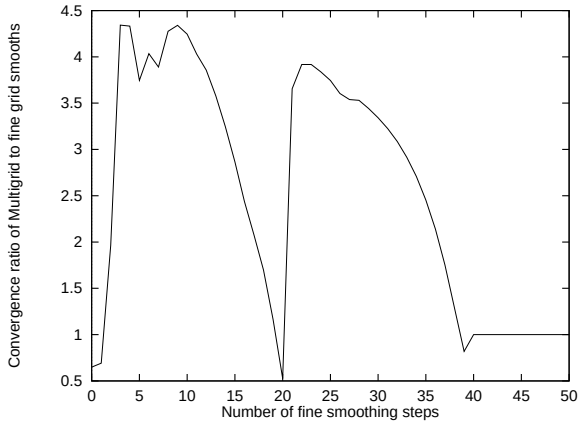


Figure 4.11: Ratio of effective smoothing rates between multigrid cycles and fine grid smooths

rections being made at each level throughout the CGC process. A method for combating this was proposed by Goodyer *et al.* [49]. This made use of both the muligrid scheme and the Nurgat solver.

In the method of [49] multigrid cycles were initially used to reduce the errors in the calculated solution down to the point where stalling occurs. At this point it was proposed that the use of the coarsest grid be dropped, and that multigrid cycles be continued on the rest of the grid levels. This in turn could reach a new (lower) stalling point at which this new coarsest grid was deemed 'too coarse'. How many grids needed to be dropped was problem specific, but to obtain solutions with residuals at the level of machine precision it was not uncommon to reach a point where only the finest grid was used.

An important issue with this method is the choice of when to stop using grids. This cut-off point is determined by a number of factors. This most obvious of these is determined by the ratio between the reduction in residual size from the CGC process to that which would have been expected by smoothing on the fine grid alone. Clearly if this process is not providing an improved solution quicker than would be accomplished without the use of multigrid, then there is something wrong with the multigrid process being used. This ratio is illustrated in Figure 4.11 where it is plotted against the number of fine grid smooths, for a three grid level example. It can be seen that initially multigrid is very effective in reducing the fine grid residual level. However this soon decays to the point where it starts becoming detrimental (20 iterations). Since this ratio has dropped below 1, the coarsest grid is removed from future use. Again the multigrid is seen to be immediately effective, although this usefulness, too, is eventually limited, with the grid being removed just before 40 iterations. The code then proceeds just using the finest grid, hence the ratio is identically 1.

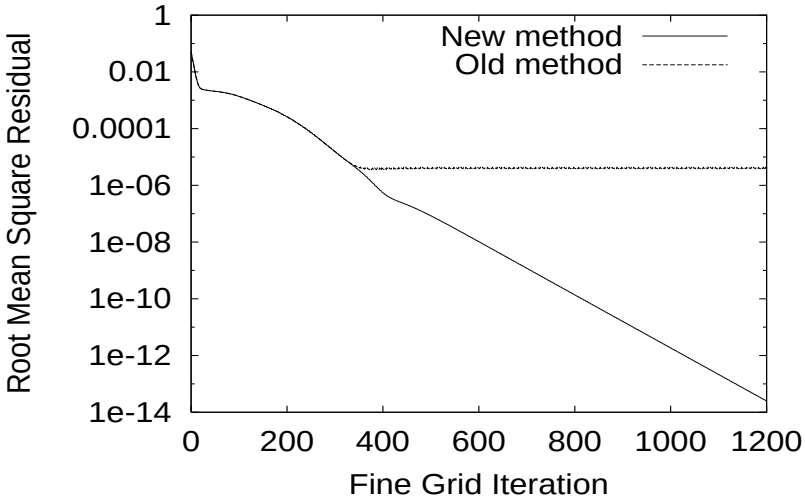


Figure 4.12: Root mean square residual levels in a thermal viscoelastic case with sliding, using the multigrid grid-elimination method of Goodyer *et al.* [49]

These modifications to the multigrid method of [108] meant that numerically more accurate solutions to steady state EHL problems could be obtained. It should be noted that the multi-integration solve is still done over the same number of grid levels irregardless of the removal of coarse grids from the multigrid process.

Results showing the convergence of this code were presented by Goodyer *et al.* in [49]. One of those cases for which the new method was demonstrated was for a thermal viscoelastic example with sliding, reproduced here as Figure 4.12. It is clear that the elimination of the grids has enabled convergence to be continued beyond the previous stalling point, towards machine precision.

This method greatly improved the code of Nurgat, since the dropping of grids lessens the change at the cavitation boundary. However this method does also remove much of the speed-ups possible due to the use of multigrid. Ideally a cavitation boundary treatment was needed without losing the benefits of multigrid.

4.5.2.2 Cavitation Boundary Treatment

The cavitation boundary is a physical constraint on the problem, rather than a mathematical one. Any pressure calculated to be negative is set to be identically zero, as described by Dowson and Taylor [35]. Any pressures at zero are not being allowed to change to their 'true' value, as far as the Reynolds Equation is concerned. The line solver will be calculating a negative solution which will feed back into the positive pressures.

On the finest grid this will not be a problem because the solution will converge to a solution which defines a particular value of the free boundary. However on all coarser

grids this boundary will almost certainly not be in the same position. Over several levels of grids the boundary may move considerably between solutions on different grids.

There are several issues concerning the treatment of this boundary during the transfer of solutions between grids. It is particularly important in the prolongation process because if no change is made near the boundary, then the fine grid boundary is unmoved.

The formulation of the Right Hand Side function for the coarser grids must also ensure that there are no contributions from inside the cavitation region. In fact, this very issue has recently been discussed by Venner and Lubrecht in [145] where they, too, talk about the possibility of "a narrow band around the cavition *(sic)* boundary where [the residuals] remain fixed at a certain level and do not converge because of this switching back and forth between cavitated and non-cavitated". Their solution is as described here – not allowing coarser grids to move a fine grid cavitation boundary, or allowing transfer of information about residuals in the cavitation region to affect the solution in the rest of the domain.

The treatment of the cavitation boundary as only being free to be updated on the finest grid meant that stalling no longer ever occurred. The region from one fine cell before the boundary was only ever updated on the finest grid. This means that these points will never receive multigrid speed ups in convergence, however nor will they ever be wrongly cavitated. Results showing the multigrid convergence will be presented in Section 4.6.1.

4.5.2.3 H_{00} Relaxation

The third major area where improvements were necessary in the code was in the relaxation of H_{00}. It was explained in Section 4.2.3 that the Force Balance Equation was included into the equation system being solved by adding a contribution of the difference between the desired and calculated values for the total pressure.

Once again, when applied on only one (fine) grid, this method converged on the numerically accurate value of H_{00}. However when applied in the multigrid framework of Nurgat [108] this calculated value was different to the single grid method. In this code a single value of H_{00} was calculated which was relaxed on all levels. This is shown in Figure 4.13 where the black dots indicate the levels on which multigrid smooths occur, the circles indicate Force Balance relaxations, and the arrows indicate the transfer of H_{00} between levels.

The problem with this method can be easily demonstrated by considering the ideal example where the solution on the finest grid has zero numerical error, thus having the correct H_{00} value. That the pressure solutions – and hence the sum of the pressures – on coarser grids will be different to those on the finest can be easily seen by considering

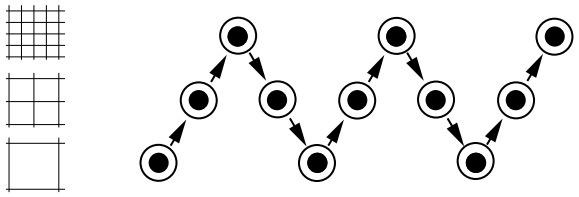


Figure 4.13: H_{00} strategy applied by Nurgat

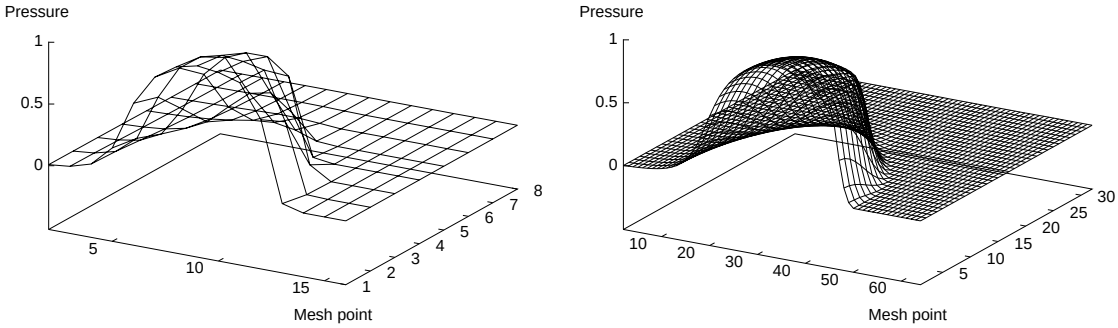
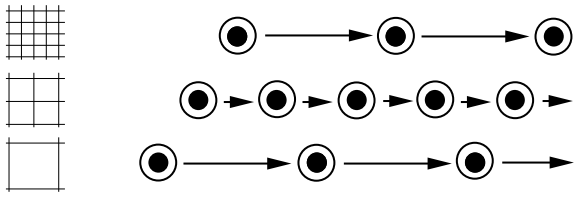


Figure 4.14: Difference in pressure solutions between grid level 3 (17×9 points) and grid 5 (65×33) on a half domain

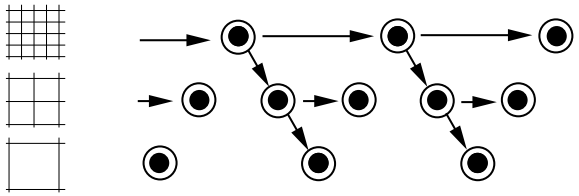
the difference in the level of refinement. This is demonstrated in Figure 4.14 where two half domain solutions for pressure are plotted. It can be seen that on the coarser grid the true shape of the pressure in the contact area is only an approximation to the fine grid. The coarse grid pressure has been calculated using a coarsening operator of the fine grid solution. That this approximation could have the same sum of discretised pressures is unrealistic, since the coarsened solution on a grid is not the same as the true solution on that grid. If the coarsest grid is then used to recalculate H_{00} which is then used on a finer grid, an error will have been introduced into the solution.

Several different strategies to combat this problem were proposed by Goodyer *et al.* [49]. These revolved around the idea that the value of H_{00} on grid k could not be changed by pressure solutions on any grid j with $j<k$. Two of these are demonstrated in Figure 4.15. In Method (a) the value of H_{00} is kept independent between the grids. In Method (b) the fine grid value is transferred onto the coarser grid. However when the solution is being prolonged back up to this grid the H_{00} value used is that previously calculated on the grid. It was shown in [49] that these two methods both produced results of optimal accuracy on the finest grid, rather than those of the method shown in Figure 4.13, however Method (b) had a higher rate of convergence.

Since the publication of [49] further work has been done into the solution of the Force Balance Equation. It has been realised that none of the methods described in either [108]



Method (a) Keep each grids' H_{00} values independent



Method (b) Transfer fine grid H_{00} values down to coarsest, but not propogate back up

Figure 4.15: H_{00} strategies considered by Goodyer *et al.* [49]

nor [49] was accurately satisfying the coarse grid problem. The value of H_{00} was always being relaxed with the aim of having the total sum of non-dimensional pressures being $\frac{2\pi}{3}$ on every grid. However neither of these two previous works had made the necessary correction to the applied load being relaxed upon, based on the difference between the fine and coarsened pressure solutions. Hence, although relaxation on the finest grid, say k, would be given, as in Equation (4.2), by

$$H_{00} \leftarrow H_{00} + c\left(\frac{2\pi}{3} - (\Delta X)^k (\Delta Y)^k \sum_{i=1}^{N_X^k} \sum_{j=1}^{N_Y^k} P_{i,j}^k\right), \tag{4.3}$$

with $(\Delta X)^k$ and $(\Delta Y)^k$ being the mesh spacings in the X and Y directions respectively on grid k, on grid $k-1$ Equation (4.3) should then become

$$\begin{aligned} H_{00} \quad \leftarrow \quad & H_{00} + c\left[\frac{2\pi}{3} - \left((\Delta X)^{k-1} (\Delta Y)^{k-1} \sum_{i=1}^{N_X^{k-1}} \sum_{j=1}^{N_Y^{k-1}} P_{i,j}^{k-1}\right.\right. \\ & \left.\left. + (\Delta X)^k (\Delta Y)^k \sum_{i=1}^{N_X^k} \sum_{j=1}^{N_Y^k} P_{i,j}^k - (\Delta X)^{k-1} (\Delta Y)^{k-1} \sum_{i=1}^{N_X^{k-1} N_Y^{k-1}} \sum_{j=1} \overline{P}_{i,j}^{k-1}\right)\right], \end{aligned} \tag{4.4}$$

where $\overline{P}_{i,j}$ is $P_{i,j}$ after the coarsening procedure, before any pre-smooths have been done. When this process is applied iteratively, a series of corrections to the sum of pressure are

produced. These can be written as

$$\tau^{k-1} = \tau^k + (\Delta X)^k (\Delta Y)^k \sum_{i=1}^{N_X^k} \sum_{j=1}^{N_Y^k} P_{i,j}^k - (\Delta X)^{k-1} (\Delta Y)^{k-1} \sum_{i=1}^{N_X^{k-1}} \sum_{j=1}^{N_Y^{k-1}} \overline{P}_{i,j}^{k-1}, \qquad (4.5)$$

and then may be used to define the correction to the Force Balance Equation on any grid by

$$H_{00} \leftarrow H_{00} + c \left(\frac{2\pi}{3} - (\Delta X)^k (\Delta Y)^k \sum_{i=1}^{N_X^k} \sum_{j=1}^{N_Y^k} P_{i,j}^k + \tau^k \right). \qquad (4.6)$$

Using this method it is not necessary to relax as often as before, and relaxation can simply be done on the coarsest grid used. The correction terms mean that the 'target' sum of pressures will always be $\frac{2\pi}{3}$ on the finest grid.

4.5.3 Summary

In this section several problems with the multigrid method used by Nurgat [108] have been explained. They have been broken down into three distinct parts and the corrections implemented have been explained. Now that the accurate treatment of the relaxation of the Force Balance Equation has been implemented, and with more careful treatment of the cavitation region it has been possible to eliminate the phenomenon of stalling. The success of these modifications will be shown in the next section.

It is important to note, however, that the EHL problem is very sensitive to other problem dependent issues. These include making a good choice of relaxation parameters used in each region for the solution of the Reynolds equation. This is to ensure that convergence occurs as quickly as possible, but that the non-linear solver does not diverge. The choice of domain size is sometimes difficult because the wide range of operating conditions modelled. A balance has to be made between having a sufficiently small domain for highly loaded examples, and having a sufficiently large inlet region. It is also important to make sure that the coarsest grid used is "sufficiently fine" [145] to be able to be useful in accurately representing the solution. The success of these improvements will be seen in the next section.

4.6 Performance of the Code

In considering the performance of the EHL code it is important to characterise how much improvement is being made. It is well known that the multilevel techniques employed

Parameter	Test Case 3	Test Case 4
Viscosity index, α	2.2×10^{-8} Pa^{-1}	2.2×10^{-8} Pa^{-1}
Viscosity at ambient pressure, η_0	0.04 Pa s	0.04 Pa s
Maximum Hertzian pressure, p_h	0.45 G Pa	0.97 G Pa
Material parameter, G	4972	4972
Load parameter, W	1.63×10^{-7}	1.63×10^{-6}
Speed parameter, U	8.18×10^{-12}	8.18×10^{-12}
Moes parameter, M	20	200
Moes parameter, L	10	10

Table 4.4: Non-dimensional parameters for multigrid performance benchmarking

have performance as described in Chapter 3, but in this section it will be shown how these techniques combine within the solver employed here. Especially considering that alterations to the standard multigrid method have been made, the benefits of multigrid (Section 4.6.1) and multilevel multi-integration (Section 4.6.2) will be examined separately.

The EHL point contact problem is being solved on increasingly fine meshes and the computation time increases dramatically with every extra level of finest mesh added. Despite the high computational cost of EHL problems parallel computers do not appear to have been used to reduce the run times. The availability of multiprocessor machines for relatively cheap cost is growing with the introduction of commercially available Linux Beowulf clusters. In Section 4.6.3 parallelism is introduced and explained, with speed-up results presented.

4.6.1 Benefits of Multigrid

The multigrid techniques described in Chapter 3 have been applied to the EHL problem by various authors, e.g. [38,145,154], and over the course of this chapter the implementation of Nurgat [108] has been further refined.

There are various parts to the multigrid solution process that can be quantified. In [145] Venner and Lubrecht have presented a series of benchmarks for their code. Detailed results for residual levels at each iteration, as well as calculated values for central and minimum film thickness, and H_{00}, are provided for an incompressible lubricant using the Barus viscosity-pressure equation. There are also some more limited results provided for the model described in Chapter 2.

A similar investigation of the performance of the code used in this work is provided here for the two test cases shown in Table 4.4. Comparisons will be made to [145] where possible. The computational timings of some selected cases will also be presented to

Iteration	Level 3	Level 4	Level 5	Level 6
1	5.2×10^{-1}	1.1×10^{-1}	5.2×10^{-2}	3.2×10^{-2}
100	1.8×10^{-5}	4.1×10^{-3}	8.8×10^{-3}	3.6×10^{-3}
200	6.8×10^{-9}	2.6×10^{-5}	3.9×10^{-3}	3.0×10^{-3}
300	8.5×10^{-12}	2.6×10^{-7}	1.4×10^{-3}	2.5×10^{-3}
400	1.1×10^{-14}	4.5×10^{-10}	2.0×10^{-4}	2.1×10^{-3}
500	8.5×10^{-16}	2.8×10^{-12}	2.6×10^{-5}	1.5×10^{-3}
600	8.5×10^{-16}	2.1×10^{-14}	3.5×10^{-6}	9.5×10^{-4}

Table 4.5: Root mean square residual levels during single grid convergence of Test Case 3

Iteration	Level 3	Level 4	Level 5	Level 6
1	8.8×10^{0}	2.4×10^{0}	1.1×10^{0}	1.4×10^{2}
100	2.9×10^{-2}	4.5×10^{-3}	9.1×10^{-3}	3.7×10^{-3}
200	2.2×10^{-3}	9.0×10^{-6}	4.0×10^{-3}	3.1×10^{-3}
300	4.6×10^{-4}	4.0×10^{-9}	1.3×10^{-3}	2.6×10^{-3}
400	1.9×10^{-5}	3.3×10^{-12}	1.6×10^{-4}	2.1×10^{-3}
500	6.3×10^{-7}	1.3×10^{-14}	1.9×10^{-5}	1.5×10^{-3}
600	1.9×10^{-8}	3.6×10^{-16}	2.3×10^{-6}	9.7×10^{-4}

Table 4.6: Root mean square residual levels during single grid convergence of Test Case 4

demonstrate the multigrid performance. These cases were all run without use of multilevel multi-integration in order to demonstrate the effectiveness of multigrid alone.

The calculations were carried out on a domain X∈[-4.5,1.5], Y∈[-3.0,3,0]. The relaxation parameters used, except where stated, were 0.4 for the relaxation of the Reynolds Equation, and 0.1 for the Force Balance Equation relaxation of H_{00}.

4.6.1.1 Single Grid Performance

The performance of multigrid must be compared to the performance on only one fine grid. The two test cases were run on single grids between levels 3 (half domain 17×9 points) and level 6 (129×65). Each case has the Hertzian pressure profile, given by Equation (2.10), as the initial approximation. The relaxation parameter of the Force Balance Equation for H_{00}, c in Equation (4.6), was 0.05 on all levels other than Level 3 where it was set to be 0.1 for Test Case 3 and 0.25 for Test Case 4. This value was lower for the coarsest grid on Test Case 4 because the higher loading of the problem makes good coarse solutions initially harder to obtain.

The root mean square of the residual level of the Reynolds Equation is a good measure of the convergence of the solution. In Tables 4.5 and 4.6 these are shown at various points for each of the two test cases. The overall convergence properties are shown graphically

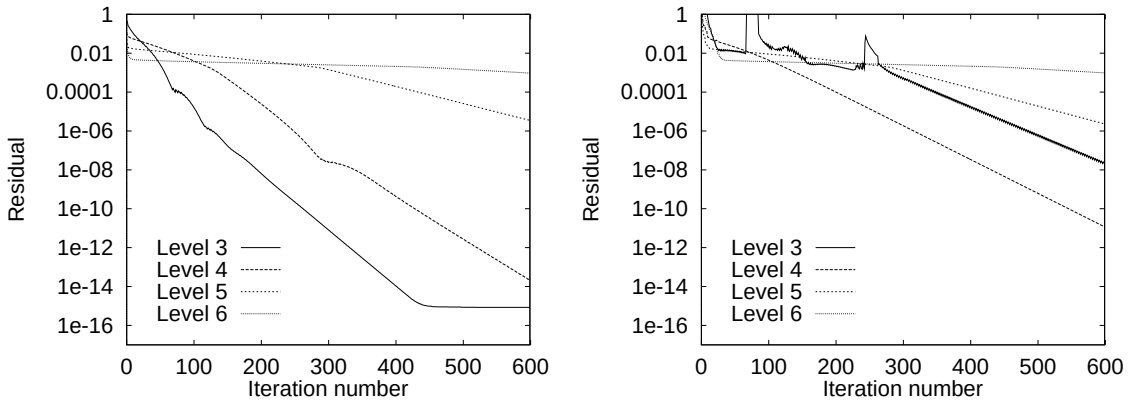


Figure 4.16: Convergence of the residual during single grid convergence of Test Case 3 (left) and 4 on levels 3 to 6

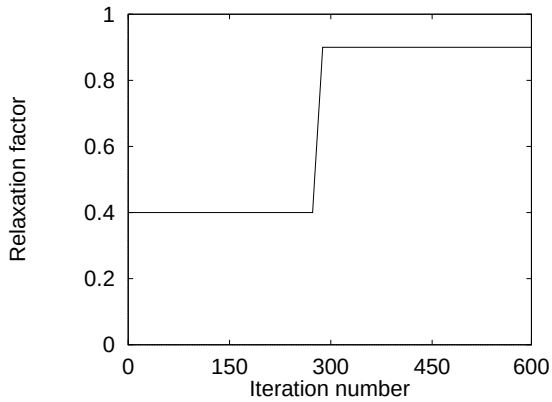


Figure 4.17: Reynolds Equation relaxation factor increasing through convergence, Test Case 4, grid 5

in Figures 4.16.

From these results it can be seen that as the grid being used gets finer the errors in the solution become harder to smooth away. This is consistent with the known properties of the smoothing methods used, and is the reason why multigrid was first applied to EHL problems.

A notable difference between the results presented above and those on page 196 of Venner and Lubrecht [145], is that the convergence rates on the finer grids do improve from their behaviour over the first 100-200 iterations. This is because of another modification to the Reynolds Equation solution scheme introduced by Goodyer *et al.* in [49]. During the later stages of convergence, the solution is requiring smaller changes to be made in each update. This means that the relaxation parameter used can be increased since the solution is now stable. It has been allowed to increase in 0.1 intervals up to a maximum of 0.9. Figure 4.17 shows how the relaxation factor increased on grid 5 for

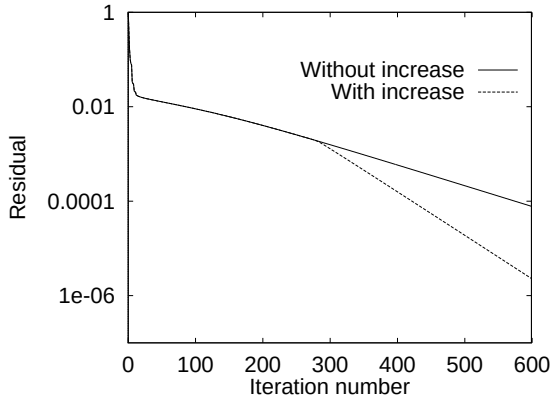


Figure 4.18: Convergence of the residual during single grid convergence of Test Case 4, grid 5, with and without variable relaxation factors

	Level 3	Level 4	Level 5	Level 6
Test Case 3	21.5	132.6	1379.4	16666.5
Test Case 4	21.8	139.1	1417.0	17227.6

Table 4.7: Computational time (s) for 600 smooths on a single fine grid

Test Case 4, and comparative convergence rates with and without increasing this factor are shown in Figure 4.18.

The computational time required to obtain solutions is clearly directly proportional to the number of iterations taken. Table 4.7 shows the computational time to obtain a solution after 600 smooths on these single grids. It is clear that single grid smooths on Level 7 and above would not be feasible in a realistic time, and that the results obtained would also not show any noticeable convergence.

4.6.1.2 Multigrid Performance

The theory that the use of multigrid provides convergence at a far greater rate than would have been previously obtainable on a single fine grid, has been explained in Chapter 3. Here, by comparison with the single grid results provided above, this speed-up will be quantified for the code used in this work.

The root mean square residual levels for the two test cases are shown in Tables 4.8 to 4.11. For each case, the results are shown for both V- and W-cycles. For the W-cycles the technique of halving the relaxation factor of the Force Balance Equation for H_{00}, used in the code of Venner and Lubrecht [145], has been adopted. Calculated values for central and minimum film thickness are also presented in Tables 4.12 to 4.15. Each solution has been calculated with a coarsest level of grid 3 with 17×9 points in the half domain.

These results emphasise exactly how multigrid solutions converge significantly faster

Iteration	Level 5	Level 6	Level 7
1	3.7×10^{-2}	1.2×10^{-2}	3.8×10^{-3}
2	8.9×10^{-3}	4.8×10^{-3}	2.4×10^{-3}
3	2.9×10^{-3}	1.9×10^{-3}	1.1×10^{-3}
4	1.1×10^{-3}	6.7×10^{-4}	4.5×10^{-4}
5	7.1×10^{-4}	3.1×10^{-4}	1.8×10^{-4}
6	5.5×10^{-4}	2.0×10^{-4}	8.5×10^{-5}
7	4.1×10^{-4}	1.5×10^{-4}	5.7×10^{-5}
8	3.2×10^{-4}	1.1×10^{-4}	4.3×10^{-5}
9	2.4×10^{-4}	8.5×10^{-5}	3.3×10^{-5}
10	1.8×10^{-4}	6.8×10^{-5}	2.6×10^{-5}
20	1.3×10^{-5}	6.0×10^{-6}	4.9×10^{-6}
30	7.0×10^{-7}	2.8×10^{-7}	6.3×10^{-7}
40	2.7×10^{-8}	2.1×10^{-8}	8.6×10^{-8}

Table 4.8: Root mean square residual levels during multigrid V(3,1)-cycle convergence of Test Case 3

Iteration	Level 5	Level 6	Level 7
1	3.0×10^{-2}	1.1×10^{-2}	4.4×10^{-3}
2	8.1×10^{-3}	3.4×10^{-3}	1.7×10^{-3}
3	2.0×10^{-3}	8.5×10^{-4}	6.3×10^{-4}
4	8.7×10^{-4}	3.0×10^{-4}	2.8×10^{-4}
5	5.9×10^{-4}	1.5×10^{-4}	6.1×10^{-5}
6	4.4×10^{-4}	1.1×10^{-4}	3.1×10^{-5}
7	3.6×10^{-4}	1.0×10^{-4}	3.1×10^{-5}
8	2.9×10^{-4}	1.0×10^{-4}	2.9×10^{-5}
9	2.4×10^{-4}	1.7×10^{-4}	3.2×10^{-5}
10	2.1×10^{-4}	1.3×10^{-4}	3.0×10^{-5}
20	1.7×10^{-5}	3.1×10^{-6}	1.7×10^{-6}
30	1.2×10^{-6}	2.8×10^{-7}	2.5×10^{-7}
40	5.1×10^{-8}	3.0×10^{-8}	4.1×10^{-8}

Table 4.9: Root mean square residual levels during multigrid W(3,1)-cycle convergence of Test Case 3

Iteration	Level 5	Level 6	Level 7†
1	7.3×10^{-2}	2.1×10^{-2}	9.9×10^{-3}
2	1.5×10^{-2}	5.3×10^{-3}	2.6×10^{-3}
3	5.1×10^{-3}	2.5×10^{-3}	1.2×10^{-3}
4	7.7×10^{-3}	1.3×10^{-3}	5.9×10^{-4}
5	1.8×10^{-3}	7.4×10^{-4}	3.1×10^{-4}
6	1.3×10^{-3}	4.8×10^{-4}	1.8×10^{-4}
7	9.5×10^{-4}	3.3×10^{-4}	1.3×10^{-4}
8	6.9×10^{-4}	2.4×10^{-4}	1.1×10^{-4}
9	5.0×10^{-4}	1.7×10^{-4}	9.5×10^{-5}
10	3.6×10^{-4}	1.3×10^{-4}	8.2×10^{-5}
20	1.1×10^{-5}	6.3×10^{-6}	1.5×10^{-5}
30	9.8×10^{-7}	6.4×10^{-7}	3.5×10^{-6}
40	6.5×10^{-8}	8.6×10^{-8}	8.4×10^{-7}

† H_{00} relaxation parameter, c=0.05.

Table 4.10: Root mean square residual levels during multigrid V(2,1)-cycle convergence of Test Case 4

Iteration	Level 5	Level 6	Level 7
1	1.0×10^{0}	3.0×10^{-1}	1.5×10^{-2}
2	3.6×10^{-2}	4.8×10^{-3}	2.1×10^{-3}
3	6.0×10^{-3}	1.7×10^{-3}	7.1×10^{-4}
4	2.7×10^{-3}	7.4×10^{-4}	8.1×10^{-4}
5	1.7×10^{-3}	3.7×10^{-4}	3.2×10^{-4}
6	1.3×10^{-3}	1.9×10^{-4}	1.4×10^{-4}
7	9.3×10^{-4}	1.3×10^{-4}	4.7×10^{-5}
8	7.2×10^{-4}	1.0×10^{-4}	2.0×10^{-5}
9	5.5×10^{-4}	9.8×10^{-5}	1.6×10^{-5}
10	4.3×10^{-4}	9.8×10^{-5}	1.3×10^{-5}
20	1.8×10^{-5}	5.9×10^{-6}	6.8×10^{-7}
30	1.4×10^{-6}	8.7×10^{-7}	1.3×10^{-7}
40	1.3×10^{-7}	1.2×10^{-8}	3.0×10^{-8}

Table 4.11: Root mean square residual levels during multigrid W(2,1)-cycle convergence of Test Case 4

Iteration	Level 5	Level 6	Level 7
1	0.380	0.383	0.386
2	0.287	0.283	0.280
3	0.297	0.293	0.286
4	0.297	0.289	0.285
5	0.301	0.295	0.289
6	0.289	0.283	0.277
7	0.291	0.285	0.278
8	0.295	0.288	0.283
9	0.298	0.291	0.286
10	0.300	0.294	0.289
20	0.308	0.302	0.297
30	0.308	0.303	0.298
40	0.308	0.303	0.298

Table 4.12: Minimum film thickness during multigrid V(3,1)-cycle convergence of Test Case 3

Iteration	Level 5	Level 6	Level 7
1	0.419	0.428	0.429
2	0.438	0.415	0.408
3	0.467	0.441	0.430
4	0.454	0.437	0.444
5	0.452	0.441	0.424
6	0.448	0.440	0.434
7	0.456	0.444	0.430
8	0.457	0.445	0.437
9	0.459	0.446	0.437
10	0.460	0.447	0.439
20	0.461	0.449	0.441
30	0.461	0.449	0.441
40	0.461	0.449	0.441

Table 4.13: Central film thickness during multigrid V(3,1)-cycle convergence of Test Case 3

Iteration	Level 5	Level 6	Level 7
2	0.170	0.00685	0.0203
4	0.0131	0.0122	0.0371
6	-0.00366	0.0329	0.0382
8	0.0169	0.0324	0.0383
10	0.0244	0.0315	0.0387
20	0.0343	0.0397	0.0405
30	0.0367	0.0400	0.0406
40	0.0368	0.0400	0.0406

Table 4.14: Minimum film thickness during multigrid W(2,1)-cycle convergence of Test Case 4

Iteration	Level 5	Level 6	Level 7
2	0.339	0.0339	0.0996
4	0.154	0.0806	0.0914
6	0.112	0.0860	0.0886
8	0.0976	0.0865	0.0866
10	0.0925	0.0870	0.0869
20	0.0899	0.0898	0.0871
30	0.0900	0.0898	0.0871
40	0.0900	0.0898	0.0871

Table 4.15: Central film thickness during multigrid W(2,1)-cycle convergence of Test Case 4

	Level 5	Level 6	Level 7
Test Case 3 - V(3,1)-cycles	471.5	5161.5	71663.5
Test Case 3 - W(3,1)-cycles	556.4	5836.2	78020.1
Test Case 4 - V(2,1)-cycles	368.8	3929.0	54495.6
Test Case 4 - W(2,1)-cycles	443.2	4511.6	58972.9

Table 4.16: Computational time (s) for 40 multigrid cycles

than single grid solutions. Each multigrid iteration has either three or four fine grid smooths, and there are obviously many more less computationally expensive smooths on coarser grids, but the reduction in computational time on introducing multigrid is palpable by comparing Tables 4.7 and 4.16.

Both V- and W-cycles appear to obtain similar convergence results. The W-cycles do accelerate past an initial slow-down approaching 10 multigrid iterations. This perceived slow-down is most likely associated with the continued convergence of the Force Balance Equation rather than a lack of convergence of the Reynolds Equation.

The film thickness results are comparable to those presented in Venner and Lubrecht [145]. The accuracy of the solutions obtained can be gauged by comparing the results obtained on each grid, to those on the finest mesh. Note that on the finer meshes more accurate results are obtained with significantly less iterations than to get similarly converged results on coarser meshes. For example in Table 4.14 it can be seen that the minimum film thickness is within 10% of the final converged value after only 4 iterations on grid 7 but takes between 10 and 20 W-cycles on levels 5 and 6.

4.6.1.3 Full Multigrid Performance

The multigrid process accelerates the elimination of errors of different frequencies. The initial approximation used on the finest mesh in the above results was the Hertzian pressure distribution of Equation (2.10). Whilst this is not a bad estimate of the shape there will be very large errors in it compared to the true solution. Using the process of Full Multigrid (FMG), described in Section 3.3.5, it is possible to use the coarser grids to obtain a much better first estimate on the finest grid. The use of FMG cannot increase the convergence rate once the fine grid has been reached, but does give the solution a headstart.

The process of FMG is constrained by the accuracy of the interpolation operator to each freshly encountered grid. In moving from grid $k-1$ to grid k, there is very little point in obtaining a solution that has converged to round off on grid $k-1$. This is because the difference between discretisation errors on the grids, and the order of interpolation are

Iteration	0 cycles	1 cycle	2 cycles	4 cycles	40 cycles
1	7.3×10^{-2}	2.5×10^{-3}	2.4×10^{-3}	2.4×10^{-3}	2.4×10^{-3}
2	1.5×10^{-2}	6.0×10^{-4}	5.1×10^{-4}	4.5×10^{-4}	4.5×10^{-4}
3	5.5×10^{-3}	2.5×10^{-4}	2.1×10^{-4}	1.8×10^{-4}	1.8×10^{-4}
4	2.8×10^{-3}	1.5×10^{-4}	1.4×10^{-4}	1.4×10^{-4}	1.4×10^{-4}
5	2.0×10^{-3}	1.3×10^{-4}	1.5×10^{-4}	1.5×10^{-4}	1.3×10^{-4}
10	3.9×10^{-4}	2.0×10^{-5}	2.2×10^{-5}	2.2×10^{-5}	2.0×10^{-5}
40	8.8×10^{-8}	1.8×10^{-8}	1.4×10^{-8}	1.1×10^{-8}	1.0×10^{-8}

Table 4.17: Residual level with finest grid level 5 with varying numbers of V-cycles in an FMG start for Test Case 4

Iteration	0 cycles	1 cycle	2 cycles	4 cycles	40 cycles
1	2.7×10^{-2}	6.9×10^{-4}	6.3×10^{-4}	6.0×10^{-4}	6.0×10^{-4}
2	5.6×10^{-3}	2.1×10^{-4}	1.5×10^{-4}	1.2×10^{-4}	1.2×10^{-4}
3	2.5×10^{-3}	9.2×10^{-5}	6.8×10^{-5}	6.1×10^{-5}	6.3×10^{-5}
4	1.3×10^{-3}	5.6×10^{-5}	5.2×10^{-5}	5.2×10^{-5}	5.6×10^{-5}
5	7.4×10^{-4}	5.1×10^{-5}	5.0×10^{-5}	5.3×10^{-5}	5.5×10^{-5}
10	1.5×10^{-4}	9.0×10^{-6}	9.0×10^{-6}	9.9×10^{-6}	1.1×10^{-5}
40	7.2×10^{-8}	7.8×10^{-9}	6.9×10^{-9}	6.9×10^{-9}	6.6×10^{-8}

Table 4.18: Residual level with finest grid level 6 with varying numbers of V-cycles in an FMG start for Test Case 4

unlikely to come close to maintaining this accuracy, and hence much of the computational effort done would have been wasted.

To quantify the benefits of the FMG start it is therefore necessary to consider the reduction in errors due to the use of the algorithm. The highly loaded Test Case 4 has been used to obtain a series of results using different numbers of multigrid cycles in the FMG period. These are shown in Tables 4.17, 4.18 and 4.19 for finest level 5, 6 and 7 respectively.

The results show that the use of an FMG start does significantly improve the first approximation on the finest grid, as can be seen by looking at the residual levels after just one iteration. Increasing the number of cycles in the FMG start does produce better results, although beyond two cycles these changes are minimal.

4.6.2 Benefits of Multilevel Multi-Integration

The mathematical basis behind the validity of multilevel multi-integration was first outlined by Brandt and Lubrecht in [17], and has been described here in Section 3.4. Results showing the effectiveness of their algorithm have been presented by themselves and their co-workers, for different cases in works such as [17, 140, 145, 154]. In this section the

Iteration	0 cycles	1 cycle	2 cycles	4 cycles	40 cycles
1	1.3×10^{-2}	1.7×10^{-4}	1.7×10^{-4}	1.6×10^{-4}	1.8×10^{-4}
2	3.1×10^{-3}	6.4×10^{-5}	6.2×10^{-5}	5.1×10^{-5}	1.0×10^{-4}
3	1.2×10^{-3}	5.6×10^{-5}	4.9×10^{-5}	5.4×10^{-5}	6.6×10^{-4}
4	7.2×10^{-4}	4.8×10^{-5}	5.1×10^{-5}	3.6×10^{-5}	4.3×10^{-5}
5	3.0×10^{-4}	5.4×10^{-5}	4.5×10^{-5}	4.3×10^{-5}	4.1×10^{-5}
10	7.5×10^{-5}	1.4×10^{-5}	9.1×10^{-6}	9.2×10^{-6}	8.8×10^{-6}
40	2.2×10^{-5}	2.9×10^{-7}	1.7×10^{-7}	2.2×10^{-7}	2.2×10^{-7}

H_{00} relaxation parameter, c=0.025.

Table 4.19: Residual level with finest grid level 7 with varying numbers of V-cycles in an FMG start for Test Case 4

		Coarsest grid used in multi-integration			
		Level 3	Level 4	Level 5	Level 6
Level	5	1.9×10^{-4}	7.4×10^{-5}	-	-
of	6	1.8×10^{-4}	7.5×10^{-5}	3.7×10^{-5}	-
solution	7	1.8×10^{-4}	9.5×10^{-5}	4.2×10^{-5}	1.9×10^{-5}

Table 4.20: Multilevel Multi-integration maximum errors for a single film thickness calculation

validity and efficiency of the implementation used in this work will be shown.

The multi-summation of Equation (4.1) has a strictly defined value for every point on the mesh. Use of multi-integration calculates an approximation to this. It is therefore necessary to compare the results obtained using the multilevel approach to calculate the error in the solution. This will be done in two instances. First, results for a single film thickness calculation will be shown. Secondly, a full EHL problem will be solved to convergence, to investigate any error propagation effects. These investigations will be done for the multi-integration algorithm used in this work, implemented as in Venner and Lubrecht [145].

Results for the single solve are presented in Tables 4.20 and 4.21. For each case the maximum and the root mean square error, compared to the multi-summation case, has been calculated. The rows refer to which grid level the final solution is on, whilst the

		Coarsest grid used in multi-integration			
		Level 3	Level 4	Level 5	Level 6
Level	5	3.2×10^{-5}	1.2×10^{-5}	-	-
of	6	3.5×10^{-5}	1.8×10^{-5}	7.2×10^{-6}	-
solution	7	3.7×10^{-5}	2.1×10^{-5}	1.0×10^{-5}	3.8×10^{-6}

Table 4.21: Multilevel Multi-integration root mean square errors for a single film thickness calculation

		Coarsest grid used in multi-integration				
		Level 3	Level 4	Level 5	Level 6	Level 7
Level	5	0.14	0.22	2.29	-	-
of	6	0.35	0.50	2.30	28.5	-
solution	7	1.14	1.26	3.21	28.2	397

Table 4.22: Multilevel Multi-integration computational times (s) for a single film thickness calculation

		Maximum Error	RMS Error
Level	5	2.2×10^{-4}	4.0×10^{-5}
of	6	1.5×10^{-3}	1.4×10^{-4}
solution	7	1.0×10^{-1}	8.9×10^{-3}

Table 4.23: Multilevel Multi-integration errors for 20 multigrid V(3,1) cycles

columns refer to the coarsest level used in the multi-integration.

The corresponding computation times are shown in Table 4.22. Since the calculation time using multi-integration is so short, repeated calculations (100 iterations on levels 3 to 5, 10 on level 6 and 5 on level 7) have been done and the average times shown. The speed up due to the use of multi-integration is dramatic.

To investigate the error propagation due to multilevel multi-integration, Test Case 3 was reinvestigated using the coarsest grid (level 3) possible. A FMG start comprising of 3 multigrid V-cycles was used, followed by 20 multigrid V(3,1) cycles. The maximum and root mean square errors are shown in Table 4.23, and the timings are shown in Table 4.24. It can be seen that the use of multi-integration does not lead to a significant net increase in the error in the solution compared to using the multi-summation method. The errors that are present could be smoothed away if desired by a few full multi-summation solves. It can be seen, however, that for grid 7 two full multi-summations would take longer than the 20 multigrid cycles previously employed when using multi-integration.

Before closing this section on the performance of the film thickness calculation, it is interesting to note how much of a difference the optimisation of the code can make to the computation time. The optimisation being considered here is not that of the compiler

		With multi-integration	Without multi-integration
Level	5	58	254
of	6	141	2680
solution	7	666	36684

Table 4.24: Multilevel Multi-integration computational times (s) for for 20 multigrid V(3,1) cycles

Grid Level	Unoptimised code	Optimised code
5	2.29	0.67
6	28.5	10.3
7	397	162

Table 4.25: Comparison of computation times for multi-summation using different data structures

(which does make a very large difference) but in how the code is written down, ordered and how memory allocation and retrieval is managed. A highly optimised version of the multi-summation was developed by Fairlie [43] which shows speed-ups of up to three times for identical results. These differences are shown in Table 4.25. The main differences between the two implementations are the data structures involved. In the method of Fairlie only one dimensional arrays are used, e.g. the array of pressures is not represented as `P(i,j)` but as `P(k)` where

$$\mathtt{k} = (\mathtt{j} - 1) \times N_X + \mathtt{i}, \tag{4.7}$$

using the FORTRAN array element ordering convention. These are accessed quicker than the two dimensional arrays in the other version. Implementation of these ideas into the code used in this work could yield further improvements in performance.

4.6.3 A Preliminary Investigation into the Benefits of Parallelism

Parallel computing is the utilisation of more than one processor to perform a task. In terms of a numerical solver, such as in this work, the computational effort required for the calculations needed is shared between the processors available. Issues which affect the performance of parallel solvers will not be described in detail here, but the two which are important are the ideas of minimising communication time, and load balancing.

The chosen method for parallelisation is the Message Passing Interface (MPI) which allows portability of the code between dedicated parallel machines and networks of workstations. There will be one processor chosen as the *master* which is in charge of all *serial* (non-parallel) work. This processor will then communicate the necessary information to the other (*slave*) processors, when parallel work is required.

The most computationally expensive part of the numerical calculation, as previously indicated, is the deformation calculation. Without the use of multilevel multi-integration the order of this calculation far exceeds that of the other parts of the solver. This, com-

Grid level	Percentage time evaluating film thickness
3	26
4	54
5	84
6	95
7	94

Table 4.26: Example percentages of time spent in evaluating the film thickness on each grid level

bined with the nature of the required calculation, makes the film thickness calculation the obvious target for parallelisation. Examples of the relative percentages of time spent in the film thickness calculation on each grid level are shown in Table 4.26 for a sequential solve.

To parallelise this calculation efficiently, reducing the amount of information being communicated between processors is very important. By reconsideration of Equation (4.1):

$$H_{i,j} = \mathscr{G}(X,Y) + H_{00} + \sum_{k=1}^{N_x} \sum_{l=1}^{N_y} K_{i,j,k,l} P^h_{i,j}, \tag{4.8}$$

it is clear that the undeformed geometry, $\mathscr{G}$, and the multi-summation kernel, K, are predefined quantities that do not require communicating at every solve, only at initialisation. The only variables between calls are the calculated values of the central offset film thickness, H_{00}, the current grid (both of which are scalars), and the pressure solution across the whole domain P^h.

After the relevant information has been communicated to the grid, the work is partitioned between the processors. For regular Cartesian grids, such as are being employed here, this is relatively simple. The work has been allocated on a purely line based scheme. Rows of constant j are allocated to the processors in ascending order. This does usually lead to a small imbalance in the workload on some processors; for example if 65 rows are allocated to eight processors, then one processor will have an extra line of film thickness to calculate, but in situations where the ratio of processors to rows is low, this is not a significant problem.

Once the individual processors have all the relevant information to perform all their allotted film thickness calculations, they then proceed on their own, until they have finished. The results are then communicated back to the nominated master processor. Once all the processors have returned their results, the master then continues with the serial

Number of processors	Solution Time (s)	Speed-up
1	168.4	1.00
2	105.5	1.60
4	75.2	2.24
8	59.3	2.84

Table 4.27: Parallel performance on Grid 5

Number of processors	Solution Time (s)	Speed-up
1	2131.7	1.00
2	1166.3	1.83
4	684.6	3.11
8	439.0	4.86

Table 4.28: Parallel performance on Grid 6

solve until the next film thickness solution is required.

The advantage of parallelism is only evident if the time spent in communicating the information between processors is more than offset by the reduction in time due to the reduced scale calculation. The parallelism has been implemented in such a way that it is grid independent, hence it is even parallelising the film thickness calculation on the coarsest grid. In the results that follow, times for 10 multigrid cycles are shown. The results for finer grids, therefore, include the results from the coarser grids too. The timings, presented in Tables 4.27 to 4.29 were performed without the use of multi-integration. It can be seen that, as expected, the parallelism is most effective on the finer grids. With only the film thickness calculation parallelised speed-ups of the same order as the number of processors used would be unrealistic. These results can be visualised in Figure 4.19 where the increasing effectiveness of parallelism on finer grids is very evident.

4.7 Conclusion

The use of multigrid techniques to solve steady state EHL problems is now well established. This thesis is built on the EHL solver developed by Nurgat [108]. Whilst this work

Number of processors	Solution Time (s)	Speed-up
1	31511	1.00
2	16311	1.93
4	9876	3.19
8	4850	6.50

Table 4.29: Parallel performance on Grid 7

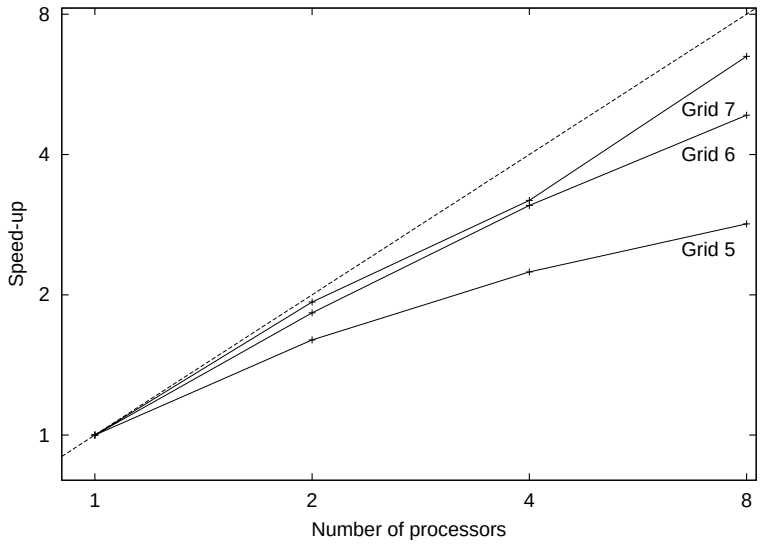


Figure 4.19: Speed-up comparison between various grids for parallel example

did appear to obtain good solutions, there were some deficiencies in some of the methods used. Some of these problems had been previously indicated in works such as [109] although no solutions had been proposed.

The main intention of this chapter was to examine these difficulties in detail, and to demonstrate that these problems can be initially overcome, and then cured. After explaining the numerical methods used for solving each of the equations, example results for steady state case were presented.

The problems in the methods of [108] have been explained in detail, and solutions proposed. First the technique of eliminating the coarsest grid from the multigrid solve, as presented in [49], was described, with results showing that numerically accurate results could now be obtained. Further adaptations to the methods used were then shown to ensure that the desired multigrid performance can be obtained without any grid elimination. The validity of the new solution methods has been shown by use of examples considering not just the physical solution, but also the estimated errors in that solution.

The overall performance of the code has been considered. Results have been presented showing the relative performance increases due to the use of both multigrid and multilevel multi-integration.

The idea of parallelism of the code has been introduced and has been shown to be very beneficial in speeding up the calculation of the deformation equation on fine grids. Further work is possible here into combining the use of parallelism with multi-integration. It is, however, unlikely that the method of parallelism will be attempted in this fashion because multi-integration has meant that the multigrid solver has become a much larger

percentage of the total work required. Parallel multigrid is now starting to be investigated, e.g. [139], and these techniques should prove transferable to EHL modelling.

Chapter 5

Solving Transient EHL problems

5.1 Introduction

Many EHL problems of practical interest are transient. The contacts being modelled are constantly rotating at speed, and there is a flow of oil between them the whole time. The results presented thus far have only been concerned with the steady state case, where none of the physical parameters governing the solution are changing. However, in many applications it is important to model the reaction of the system to changing conditions. For example, as the teeth of a gear engage, they experience large changes in the applied load over a short time period, affecting the wear. Similarly, the starting and stopping of the head in a hard disc drive is becoming more important to model as technological advances mean that the distance between the two surfaces is continuously being decreased for increased storage. Another very important topic for EHL research is the effect of surface roughness. When a surface is not completely smooth the asperities will progress through the contact with the rotation of the surfaces, causing very different surface deformations to those occuring for smooth surfaces.

Clearly, all these problems need to be solved transiently, because simply solving a series of steady state problems, such as was done by Lubrecht *et al.* in [97], will not include the squeeze effect in the film, where the non-linear behaviour in time becomes most important. Much research is currently being done into the long term behaviour of roughness in contacts, and how the amplitude of the asperities is reduced, e.g. [70–73, 75,

94, 146].

In this chapter the transient EHL problem will be considered. First, in Section 5.2, the transient form of the problem to be solved will be stated, before being discretised in Section 5.3. In Section 5.4 this system will be reformulated as a differential-algebraic one. The solution methods to be used will be described in Section 5.5. These methods will then be employed to show solutions for a series of very different example problems in Section 5.6. In this section, the methods used will be further developed to arrive at a more robust and efficient solver. The results of these examples will be presented in terms of solution variables or residual (error) levels, depending on the motivation behind the considered problem.

Variable timestepping has not, to the best of the author's knowledge, been previously employed for EHL problems. When computing accurate solutions on fine meshes, using a fixed timestep size could lead to many more steps being taken than necessary. Transient case solutions may not be changing greatly over individual timesteps, and so to restrict the timestep unduly may well waste computational resources. In Section 5.7 variable timestepping is introduced in terms of the differential-algebraic nature of the EHL problem, outlined in Section 5.4. Results are shown emphasising both the increase in performance and the quality of the solutions obtained. This section includes the work of Goodyer *et al.* [50] where much of the method was first presented.

5.2 Equations

The governing equations for transient EHL problems are those presented in Chapter 2 as Equations (2.16 - 2.21). Only the Reynolds Equation (2.16) has a transient term, and hence this is the only equation requiring any modification from the steady state case.

There are two changes to be made from the steady state case solved in Chapter 4. The first is to scale the contribution from the wedge term, $\frac{\partial(\overline{\rho}H)}{\partial X}$, of the Reynolds Equation. In order to do this, a reference speed, u_{ref}, is calculated, related to the original speeds of the two contact surfaces, where

$$u_{ref} = u_1(T{=}0) + u_2(T{=}0). \tag{5.1}$$

This is then used to scale u_s, the new sum of the roller speeds, where

$$u_s(T) = u_1(T) + u_2(T). \tag{5.2}$$

The other change is the introduction of the *squeeze* term, $\frac{\partial(\overline{\rho}H)}{\partial T}$, into the equation. The

equation to be solved is therefore

$$\frac{\partial}{\partial X}\left(\varepsilon\frac{\partial P}{\partial X}\right)+\frac{\partial}{\partial Y}\left(\varepsilon\frac{\partial P}{\partial Y}\right)-\left(\frac{u_s(T)}{u_{ref}}\right)\frac{\partial(\overline{\rho}H)}{\partial X}-\frac{\partial(\overline{\rho}H)}{\partial T}=0, \tag{5.3}$$

with all other symbols as defined in Chapter 2.

The other non-dimensional equations to be solved have no transient terms explicitly, but for completeness they are given by

$$H(X,Y) = H_{00}+\mathscr{G}(X,Y)+\frac{2}{\pi^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\frac{P(X',Y')\,\mathrm{d}X'\mathrm{d}Y'}{\sqrt{(X-X')^2+(Y-Y')^2}}, \tag{5.4}$$

$$\frac{2\pi}{3} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}P(X,Y)\,\mathrm{d}X\,\mathrm{d}Y, \tag{5.5}$$

$$\overline{\rho}(P) = \frac{0.59\times10^9+1.34p_hP}{0.59\times10^9+p_hP}, \tag{5.6}$$

$$\text{and}\qquad \overline{\eta}(P) = \mathrm{e}^{\left\{\frac{\alpha p_0}{z_i}\left[-1+\left(1+\frac{p_hP}{p_0}\right)^{z_i}\right]\right\}}, \tag{5.7}$$

with all parameters and functions as defined in Chapter 2. It will be shown in the examples of Section 5.6 that for variable load cases the left hand side of Equation (5.5) requires modification, and for surface roughness cases $\mathscr{G}(X,Y)$ will become $\mathscr{G}(X,Y,T)$.

5.3 Discretisation of the Transient Problem

The finite difference discretisation of the steady state problem was explained in Section 2.8. This is the starting basis for the transient discretisations used. Again there is a choice for the order of discretisation method, but now in both space and time. The first order scheme uses a similar discretisation of the temporal (squeeze) term as was used for the wedge term. This extension to Equation (2.36) is therefore given by

$$\begin{aligned}&\frac{\varepsilon_{i-\frac{1}{2},j}\left(P^n_{i-1,j}-P^n_{i,j}\right)+\varepsilon_{i+\frac{1}{2},j}\left(P^n_{i+1,j}-P^n_{i,j}\right)}{(\Delta X)^2}\\ +&\frac{\varepsilon_{i,j-\frac{1}{2}}\left(P^n_{i,j-1}-P^n_{i,j}\right)+\varepsilon_{i,j+\frac{1}{2}}\left(P^n_{i,j+1}-P^n_{i,j}\right)}{(\Delta Y)^2}\\ &-\frac{u_s(T)}{u_{ref}}\frac{\rho^n_{i,j}H^n_{i,j}-\rho^n_{i-1,j}H^n_{i-1,j}}{\Delta X}\\ &-\frac{\rho^n_{i,j}H^n_{i,j}-\rho^{n-1}_{i,j}H^{n-1}_{i,j}}{\Delta T} = 0,\end{aligned} \tag{5.8}$$

where n is the current timestep, and

$$\begin{aligned}
\varepsilon_{i+\frac{1}{2},j} &= \frac{\varepsilon^n_{i+1,j}+\varepsilon^n_{i,j}}{2},\\
\varepsilon_{i-\frac{1}{2},j} &= \frac{\varepsilon^n_{i-1,j}+\varepsilon^n_{i,j}}{2},\\
\varepsilon_{i,j+\frac{1}{2}} &= \frac{\varepsilon^n_{i,j+1}+\varepsilon^n_{i,j}}{2},\\
\text{and} \qquad \varepsilon_{i,j-\frac{1}{2}} &= \frac{\varepsilon^n_{i,j-1}+\varepsilon^n_{i,j}}{2}.
\end{aligned} \tag{5.9}$$

The boundary conditions are as previously, in Chapter 2, with all exterior boundaries having P=0, and the line j=1 being a symmetry condition in the Y direction.

The second order upstream discretisation of Equation (2.38) may also be used in the transient case:

$$\begin{aligned}
&\frac{\varepsilon_{i-\frac{1}{2},j}\left(P_{i-1,j}-P_{i,j}\right)+\varepsilon_{i+\frac{1}{2},j}\left(P_{i+1,j}-P_{i,j}\right)}{(\Delta X)^2}\\
&+\frac{\varepsilon_{i,j-\frac{1}{2}}\left(P_{i,j-1}-P_{i,j}\right)+\varepsilon_{i,j+\frac{1}{2}}\left(P_{i,j+1}-P_{i,j}\right)}{(\Delta Y)^2}\\
&-\frac{u_s(T)}{u_{ref}}\frac{3\rho^n_{i,j}H^n_{i,j}-4\rho^n_{i-1,j}H^n_{i-1,j}+\rho^n_{i-2,j}H^n_{i-2,j}}{2\Delta X}\\
&-\frac{3\rho^n_{i,j}H^n_{i,j}-4\rho^{n-1}_{i,j}H^{n-1}_{i,j}+\rho^{n-2}_{i,j}H^{n-2}_{i,j}}{2\Delta T} \;=\; 0,
\end{aligned} \tag{5.10}$$

with all symbols as defined previously. This discretisation is that used by Venner and Lubrecht [145]. Their motivation in developing this scheme was to minimise the total discretisation error in the characteristic X-T direction, since they claim that the leading term of the truncation error vanishes, provided both wedge and squeeze terms have coefficient one.

A similar alternative discretisation, used by Wijnant [154], is that of the *narrow upstream second order* scheme. This uses a combination of both temporal and spatial directions in the formulation of the wedge and squeeze terms. The discretisation also differs depending on which part of the solution domain is being calculated. This method has additional directions for which the leading truncation error term vanishes, for $\Delta T = \Delta X$ and $\Delta T = 2\Delta X$. This method again requires the leading coefficients of both the wedge and squeeze terms to be one. In the examples which follow this will not always be the case, because the coefficient of the wedge term depends upon the rolling speed of the contact, which will vary in, for example, the reversal example of Section 5.6.2. This discretisation

will not, therefore, be considered further here.

For all of the above discretisations, in order to solve the reversal example introduced in Section 5.6.2, once the entrainment direction has been reversed, the direction of the discretisation of the wedge term also needs to be changed. For example, when using the first order discretisation of Equation (5.8) the discretisation becomes:

$$\begin{aligned}
&\frac{\varepsilon_{i-\frac{1}{2},j}\left(P^n_{i-1,j}-P^n_{i,j}\right)+\varepsilon_{i+\frac{1}{2},j}\left(P^n_{i+1,j}-P^n_{i,j}\right)}{(\Delta X)^2}\\
+&\frac{\varepsilon_{i,j-\frac{1}{2}}\left(P^n_{i,j-1}-P^n_{i,j}\right)+\varepsilon_{i,j+\frac{1}{2}}\left(P^n_{i,j+1}-P^n_{i,j}\right)}{(\Delta Y)^2}\\
&-\frac{u_s(T)}{u_{ref}}\frac{\rho^n_{i+1,j}H^n_{i+1,j}-\rho^n_{i,j}H^n_{i,j}}{\Delta X}\\
&-\frac{\rho^n_{i,j}H^n_{i,j}-\rho^{n-1}_{i,j}H^{n-1}_{i,j}}{\Delta T} \;=\; 0. \qquad (5.11)
\end{aligned}$$

The discretisation of Equations (5.4 - 5.7) are as given in Chapter 2, namely

$$H_{i,j} = H_{00}+\mathscr{G}_{i,j}+\sum_{k=1}^{N_X}\sum_{l=1}^{N_Y}K_{i,j,k,l}P_{k,l}, \qquad (5.12)$$

$$\frac{2\pi}{3} = \Delta X\Delta Y\sum_{i=1}^{N_X}\sum_{j=1}^{N_Y}P_{i,j}, \qquad (5.13)$$

$$\overline{\rho}_{i,j} = \frac{0.59\times10^9+1.34p_hP_{i,j}}{0.59\times10^9+p_hP_{i,j}}, \qquad (5.14)$$

$$\text{and}\quad \overline{\eta}_{i,j} = \mathrm{e}^{\left\{\frac{\alpha p_0}{z}\left[-1+\left(1+\frac{p_hP_{i,j}}{p_0}\right)^z\right]\right\}}, \qquad (5.15)$$

with all symbols as defined previously.

5.4 Differential-Algebraic Formulation of the Transient Problem

The EHL system defined in Section 5.2 once discretised, as in Section 5.3, can be represented by a system of Differential-Algebraic Equations (DAEs) [20]. There are several reasons why this representation of the problem will be useful, and these will be explained below.

Defining the film thickness vector across the whole domain, $\underline{H}$, and its multiple by the

density vector $\underline{\overline{\rho}}$ in a pointwise manner by

$$\begin{aligned} [\underline{H}]_k &= H_{i,j} \quad \text{for} \quad k = i + (j-1) \times N_x \\ \left[\underline{\overline{\rho}H}\right]_k &= \overline{\rho}_{i,j} H_{i,j} \qquad\quad i = 1, \ldots, N_x \\ &\qquad\qquad\quad\; j = 1, \ldots, N_y \end{aligned} \tag{5.16}$$

and the pressure vector, $\underline{P}$, likewise, then the Reynolds Equation (5.3) is given by

$$\underline{F}_1(\underline{P}, \underline{\overline{\rho}H}) - \left[\underline{\overline{\rho}H}\right]' = 0, \tag{5.17}$$

the film thickness equation, (5.4), by

$$\underline{F}_2(\underline{P}, \underline{H}) = 0, \tag{5.18}$$

and the density equation, (5.6), by

$$\underline{F}_3(\underline{P}, \underline{\overline{\rho}}) = 0, \tag{5.19}$$

where $'$ denotes differentiation in time. These can then be combined to define a DAE system for $\underline{U}^{\mathrm{T}} = (\underline{P}^{\mathrm{T}}, \underline{H}^{\mathrm{T}}, \underline{\overline{\rho}}^{\mathrm{T}})$, by:

$$\underline{F}(\underline{U}, \underline{U}', t) = 0. \tag{5.20}$$

The classification of a DAE system is given by the *index* which can give an indication of the numerical difficulties that could occur when being solved. The index of a DAE is defined as the minimum number of times that all, or part of (5.20) must be differentiated with respect to T in order to determine U' as a continuous function of U and T [20].

ODEs are, by definition, index zero, and both index zero and index one systems are generally easier to solve than those with higher index [20]. In the EHL case the DAE aspect arises from the lack of a temporal derivative of pressure in Equation (5.17). Issues affecting the differences between solving ODE and DAE systems are discussed in detail in Brenan *et al.* [20], in particular why classical solution methods, such as backward differentiation formulae, cannot be used for all DAE systems.

To find the index of the EHL DAE system given by Equation (5.20), it would be necessary to first differentiate Equation (5.18) with respect to T to get

$$\frac{\partial \underline{F}_2}{\partial \underline{P}} \frac{\mathrm{d}\underline{P}}{\mathrm{d}T} + \frac{\partial \underline{F}_2}{\partial \underline{H}} \frac{\mathrm{d}\underline{H}}{\mathrm{d}T} = 0. \tag{5.21}$$

Hence, if $\left(\frac{\partial \underline{F}_2}{\partial \underline{H}}\right)^{-1}$ exists, this gives an expression for $\frac{\mathrm{d}H}{\mathrm{d}T}$. Similarly, differentiating Equation (5.19) gives

$$\frac{\partial \underline{F}_3}{\partial \underline{P}}\frac{\mathrm{d}\underline{P}}{\mathrm{d}T} + \frac{\partial \underline{F}_3}{\partial \underline{\overline{\rho}}}\frac{\mathrm{d}\underline{\overline{\rho}}}{\mathrm{d}T} = 0, \tag{5.22}$$

which gives an expression for $\frac{\mathrm{d}\overline{\rho}}{\mathrm{d}T}$ provided $\frac{\partial \underline{F}_3}{\partial \underline{\overline{\rho}}}$ is non-singular.

Noting that Equation (5.17) may also be written as

$$\frac{\mathrm{d}\underline{\overline{\rho}H}}{\mathrm{d}T} = \underline{H}\frac{\mathrm{d}\underline{\overline{\rho}}}{\mathrm{d}T} + \underline{\overline{\rho}}\frac{\mathrm{d}\underline{H}}{\mathrm{d}T} = \underline{F}_1(\underline{P}, \underline{\overline{\rho}H}), \tag{5.23}$$

then expressions for P', H' and $\overline{\rho}'$ have been obtained in one differentiation step. Hence, the index of the transient EHL problem is one, provided both the aforementioned inverses exist.

The calculation of the density, $\overline{\rho}$, is entirely local, leading to a diagonal Jacobian matrix. This is therefore non-singular provided that none of the diagonal entries are zero. These entries are given by differentiation of Equation (5.6) with respect to $P_{i,j}$, by

$$\frac{\partial \overline{\rho}_{i,j}}{\partial P_{i,j}} = \frac{\partial}{\partial P_{i,j}}\left(\frac{0.59\times 10^9 + 1.34 p_h P_{i,j}}{0.59\times 10^9 + p_h P_{i,j}}\right) \tag{5.24}$$

$$= \frac{0.2006\times 10^9 p_h}{\left(0.59\times 10^9 + p_h P_{i,j}\right)^2}. \tag{5.25}$$

The physical constraints on the problem mean that the pressure, P, can never fall below zero. Hence this matrix is invertible.

The invertibility of $\frac{\partial H}{\partial P}$ is much harder to establish. This is because the kernel matrix K is dense, and hence simple analysis is not possible. It is assumed that the inverse is non-singular because it is computed when solving using the Newton-Raphson iteration method [113], however this remains mathematically unproved. During the course of this work a computational investigation was conducted into the properties of this matrix. It was seen that the matrices tested were definitely non-singular with eigenvalues less than one, but definitely non-zero. Unfortunately, due to the very large size of the matrices involved, it was only possible to investigate cases up to grid size 33×33, which remains far short of the fine mesh cases of especial interest.

The development of methods for solving DAE problems of the form $F(t, y, y') = 0$ has received much attention. Most of this work has been done on solving index zero and index one problems, including codes such as DASSL [20], SPRINT [14] and LSODI [68].

These all use strategies based on the Backward Differentiation Formulae of Gear [47], of which the Backward Euler Method is the simplest example.

When solving DAE systems there are added difficulties for numerical solvers at solution discontinuities, as discussed by Petzold [118]. This is especially true in relation to using temporal derivatives over the discontinuity to select a new timestep size. The solver may believe that the first solution after the discontinuity is still satisfying the convergence test, possibly increasing the timestep size still further. In fact a different problem entirely is being solved and the previous timestep should be retaken. This situation must be borne in mind when variable timestepping for EHL solvers will be introduced in Section 5.7.

5.5 Transient Solution Method

The solution method for transient problems differs slightly from that for steady state cases because of the addition of the squeeze term to the problem. Physically this causes interesting effects, especially in the way the contacts deform. For example, in the reversal example below (Section 5.6.2) at the point of reversal there would be no surface rotation. A steady state solve would then predict zero film thickness, representing the surfaces impacting upon each other, however the squeeze effects ensure that this never happens.

The starting point for any transient problem is to solve a steady state case for the initial timestep ($T = T_0$). This is done using the multigrid solver developed in Chapter 4, and is the equivalent of running in the physical components until all start-up anomalies have been ironed out. Inclusion of temporal derivatives here should lead to a near identical solution. Small differences will be due to the temporal derivatives being extended to the end of the previous multigrid cycle, rather than the current solution, hence a very small measure of the convergence will be affecting the solution obtained.

Once the initial $T = T_0$ solution has been obtained, a timestep, size ΔT, is taken. In the method of Nurgat [108] the initial solution at the next timestep was taken to be the converged solution from the end of the previous step. In Section 5.6.2.1 it will be seen how this can be improved upon.

The discretised system, Equations (5.8 - 5.15), will then be solved using exactly the same multigrid solver as was developed in Chapter 4. The multigrid cycles are then continued until a suitably converged solution is obtained. Only then may the next timestep be taken.

The efficiency of the time integration method will depend upon how accurate the initial solution is on each step, and how much work must be done to get the solution accurate enough. On an individual timestep, the code must be able to decide if the solution has

converged sufficiently by considering whether further work is likely to provide significant improvements to the solution. To accomplish this, a strategy, such as the Shampine convergence test [130], must be used. In this test, the iteration cycle per timestep, m, with a solution $\underline{H}^{(m)}(t_n)$, is continued until

$$\frac{\sigma}{1-\sigma}\left\|\underline{H}^{(m+1)}(t_n)-\underline{H}^{(m)}(t_n)\right\| < 0.33tol, \tag{5.26}$$

where tol is an error tolerance for the iteration, $\|\cdot\|$ is a suitable norm (usually the root mean square), and σ is an estimate of the rate of convergence, defined by

$$\sigma = \left(\frac{\left\|\underline{H}^{(m+1)}(t_n)-\underline{H}^{(m)}(t_n)\right\|}{\left\|\underline{H}^{(1)}(t_n)-\underline{H}^{(0)}(t_n)\right\|}\right)^{\frac{1}{m}}. \tag{5.27}$$

This cycle therefore relates the newly calculated solution, $\underline{H}^{(m+1)}(t_n)$, to that of the initially predicted solution $\underline{H}^{(0)}(t_n)$. It is a measure of the relative change in solution over the most recent multigrid iteration against that on the first iteration of the current timestep.

Selecting the variable which will be used in the above convergence test is very important. There is a choice between testing for the convergence of H or P. This has been illustrated above for H because this is the algebraic variable present in the transient squeeze term. It is this which is dominant in transient calculations. The difference in the number of iterations required for convergence, using the same tolerances, is shown in the example of reversal, below, in Section 5.6.2.2. The accuracy of the solutions obtained is also compared. The choice of tolerance of the time integration, tol, is also very important in ensuring the overall accuracy of the solutions obtained. It will also be seen how having too large a tolerance can lead to anomalous behaviour, and so a balance must be struck between performance and accuracy.

5.6 Examples

In the field of EHL it is not a trivial matter to experimentally measure the film thickness across a contact. In transient cases, where physical parameters are changing over very short time scales, it can be nearly impossible to obtain results directly corresponding to the conditions that can occur in real systems. Optical interferometry is one commonly used technique that is employed in EHL test rigs. This works by directing a beam of light at a glass disc and photographing the light reflected off it. However, this is an expensive and time consuming process. There may also be purely experimental problems where the

frequency of the light used corresponds to interference fringes of visible light [128]. Very high speed cameras are needed to capture these results.

The ability to successfully apply a numerical solver to physical, real world problems is paramount to it being a worthwhile exercise. In the examples which follow, the results obtained will be compared to either known experimental results or those from other numerical solvers whose results have been validated elsewhere.

The first example, shown in Section 5.6.1, is that of a pseudo-steady state case. Here, no physical parameters are changing and the results at every timestep should be identical to those from the corresponding (initial) steady state case. This enables the numerical stability of the solver to be tested in a seemingly trivial case, allowing various numerical factors to be considered.

The second example is that of *reversal of entrainment*, solved in Section 5.6.2. This case corresponds to the oil entrainment slowing down and reversing direction. In real life this happens with a cam and follower, or in non-involute gears. In addition to the results for the solution variables being compared to both the experimental and numerical results of Scales *et al.* [126, 128], using prediction to reduce the initial error at each timestep will be introduced here. Also in this section the justification will be shown for the choice of testing for the convergence of the film thickness rather than the pressure in the Shampine convergence test of Equation (5.26).

A further example which involves variation of a physical parameter governing the system is given in Section 5.6.3. This particular case involves the sinusoidal variation of the applied load, and is an approximation to the kind of vibrational effects which frequently occur in components. The example chosen from Wijnant [154] is one for which oscillations at varying frequencies occur. It will be seen that the oscillation frequency governs how significant effects from the squeeze term will be.

The final set of examples, presented in Section 5.6.4, demonstrate various kinds of surface defects. These are presented as geometrically perfect deviations from the undeformed parabolic shape of the contact, since accurate modelling of a surface requires extremely fine meshes to try to represent the true geometry. The ‘real’ roughness exists at a much lower frequency than the short wavelengths of the measured roughness, and hence representations as will be used, are valid. The transient examples of roughness provided will be those of a ridge and a dent on the curved surface.

Parameter	Value
Viscosity index	2.1×10^{-8} Pa^{-1}
Maximum Hertzian pressure	0.45 GPa
Material parameter, G	2961
Load parameter, W	6.58×10^{-7}
Speed parameter, U	1.47×10^{-11}
Moes parameter, M	52.2
Moes parameter, L	6.9

Table 5.1: Non-dimensionalised parameters for reversal example

5.6.1 Pseudo-steady State

The solver initially employed for transient EHL cases was the first attempt of Nurgat [108]. It soon became apparent that this alone was not suitable for these problems. Various numerical problems were discovered which required important remedies. These are discussed here through the use of a pseudo-steady state example.

The pseudo-steady state problem is designed to show that, provided the physical conditions defining the problem do not change, then the solution is also unaffected. The example to be used in this section is the initial steady state solution to the problem described below in Section 5.6.2. The non-dimensional quantities are given in Table 5.1. The inaccuracies with the method, however, apply to all transient problems being solved, not just any particular case.

Two particular parts of the solution process were causing changes in the numerical solution between timesteps. These were numerical instabilities due to small changes in the solution, and too small a system being considered in the Jacobian matrix. The problems involved with having the timestep size set too small will also be explained.

Considering first the problem of numerical instabilities, Figure 5.1 shows how the root mean square (RMS) residual of the pressure solution on the finest grid falls until reaching unit round off level (1×10^{-16}) before growing again. This example was run without multigrid and hence the convergence rate is slower than is the case if multigrid is used. This allowed the problems with the transient method to be isolated away from any multilevel effects. The first timestep is taken at 1000 iterations and subsequent ones taken every 25 fine grid iterations. It can be seen that on the first few timesteps there is no impediment to continuing to reduce the RMS residual level, however the spikey pattern which can be observed beyond 2000 iterations reveals there is a problem. Each spike is caused at the time the timestep is taken, and the errors are then smoothed away until the next timestep is taken. Since the solution is not returning to round-off, a progressively worse solution is obtained on each step. The actual residual distribution across the computational domain is

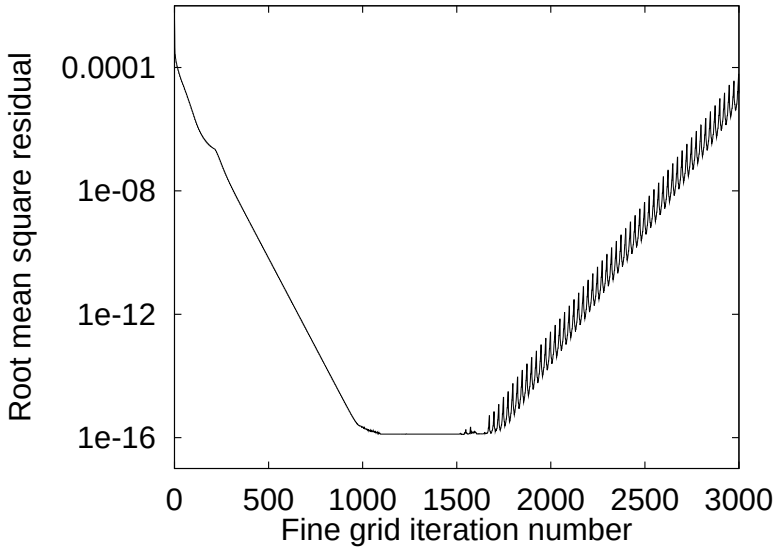


Figure 5.1: Root mean square residual level falling, then growing in a pseudo-steady state case

shown in Figure 5.2 which shows very small residual levels. However when a timestep is taken these are magnified to cause the instabilities shown in Figure 5.1. By introducing a minimum limit on the size of changes allowed in the Jacobi region of the pressure update it is then possible to maintain the root mean square residual level at round off, as shown in Figure 5.3.

In cases with variation in the operating conditions, if the timestep size is chosen to be too small then this would, as above, be introducing steep temporal gradients into the solution. This gives rise to the same kind of behaviour as above, although often on a much larger scale because the residual level will not usually be down at round off when the time step is taken. This leads to another reason for choosing small values of the under-relaxation parameters of the Reynolds Equation, especially in the contact region. This will therefore be coupled with a reduction in the speed of convergence. This minimum level will need to be borne in mind in Section 5.7 when timestep sizes will be allowed to change by the code rather than being predefined by the user.

The second, and possibly most crucial, change to the method in [108] was the expansion of the Jacobian system used in the calculation of the line solve for the Reynolds Equation. For the first order finite difference scheme, the method previously used by Nurgat [108], was always to calculate just a tri-diagonal Jacobian matrix. This had always given accurate solutions quickly for the steady-state case. However for transient cases it appears that this approximation is not satisfactory. Extension of the matrix of derivatives to be a penta-diagonal system does eliminate this problem. This also increases the robust-

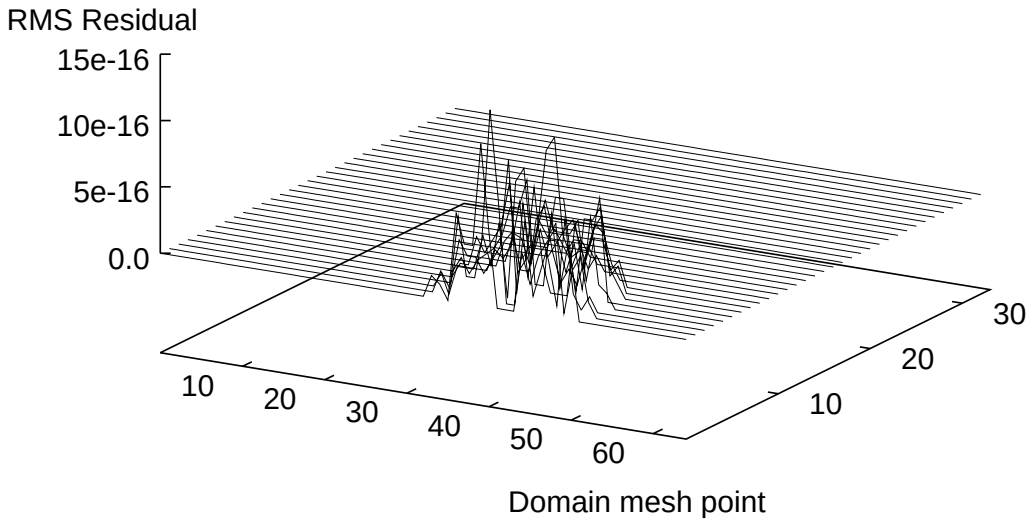


Figure 5.2: Residual level across domain at 1500 iterations in a pseudo-steady state case

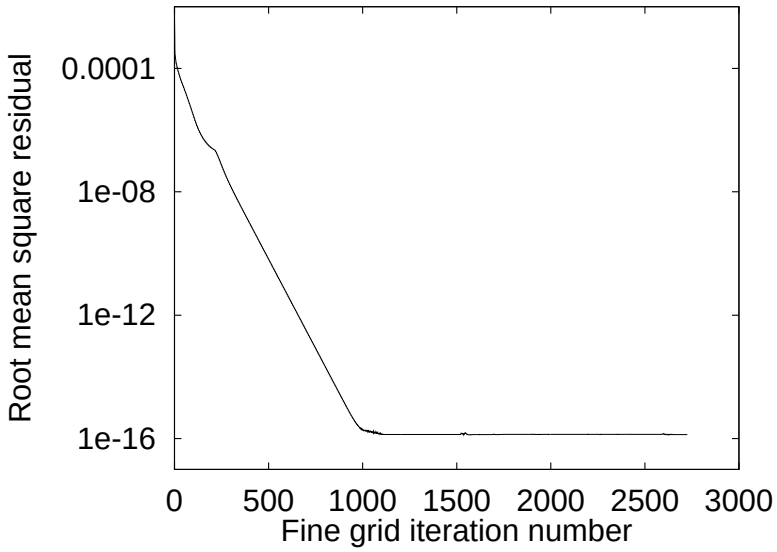


Figure 5.3: Root mean square residual level falling, and remaining at unit round-off in a pseudo-steady state case

Parameter	Value
Radius of ball	0.0127 m
Applied load	15 N
Initial velocity	0.05 ms^{-1}
Time of reversal	0.1 s
Lubricant viscosity	525 cP at 30° C

Table 5.2: Parameters used to define reversal example

ness of the solver for steady state cases and is consistent with its required use for applying a second order discretisation scheme, such as Equation (5.10).

Implementing the general scheme of Nurgat's methods, with the above modifications, solutions to pseudo-steady state problems can be obtained easily, and more demanding (and useful) transient problems can be tackled with confidence.

5.6.2 Reversal of Entrainment

The example of reversal of entrainment is typical of one being used in the design of components for industry. It is most commonly associated with the motion of a cam and follower, or a non-involute gear, with no change in the applied load. Starting from a steady state solution, the oil entrainment velocity is linearly decreased from, u_{ref} ms^{-1}, through 0 ms^{-1}, the *point of reversal*, until reaching typically $-u_{ref}$ ms^{-1}.

As the point of reversal is reached, a saucer of highly viscous oil forms in the centre of the contact, seemingly trapped between closures at both ends. This then proceeds across the domain towards the new outflow before the deformation of the surface takes up its characteristic horseshoe shape at the reverse side of the contact. This has been shown experimentally using a steel ball on a glass plate in an optical interferometer to model a cam and follower in both [128] (reproduced here in Figure 5.5) and [48]. Numerical results have been shown for both the line contact case, e.g. [69], and for the point contact case in [126, 128]. A line contact case for the complete cam and tappet problem has recently been published [104].

It is almost impossible to accurately measure all the physical conditions, time scales, and lubricant properties in an experimental test rig in order that a corresponding numerical case may be run. It is also non-trivial to map the resultant interferometry pictures obtained to real values of lubricant film thickness. In [128] an attempt was made to marry the experimental results to the numerical ones for known physical conditions. The results presented below will use the same conditions, to facilitate comparison against both sets. They are shown in Table 5.2. The lubricant model used is still Newtonian, although

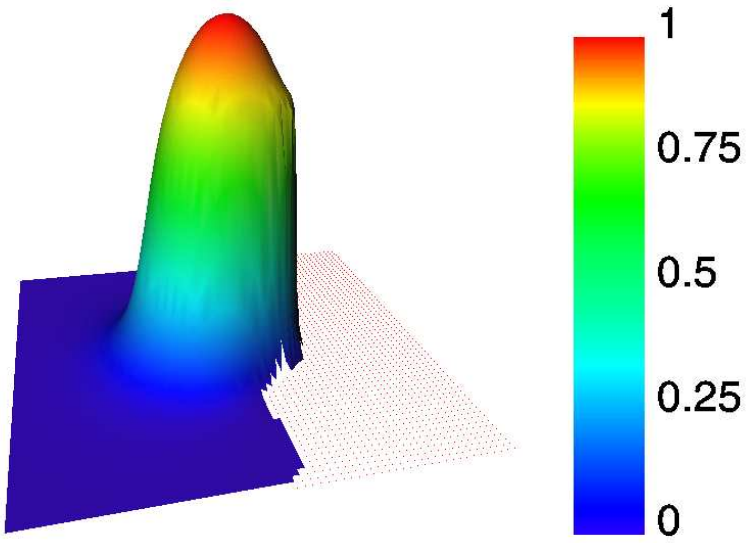


Figure 5.4: Non-dimensional pressure solution at $t = 0$ *s* for reversal example

there is a slight modification to the viscosity-pressure relation, the details of which are described in [128]. The corresponding values for the non-dimensional parameters are as given previously in Table 5.1. Using this modified version of the viscosity-pressure relation the Petrusevich spike is not as evident as in previous examples, although the pressure along the centreline is disturbed from its smooth shape at the point where the spike would have been. The pressure solution at $t = 0$ is shown in Figure 5.4, where the cavitation region is shown by the dots on the right hand side of the picture.

Results for the problem were generated on both 65×65 and 129×129 mesh point computational domains. The experimental interferometry film thickness pictures from the experiments of Scales *et al.* [128] are shown in Figure 5.5. Images are shown at the initial solution, at three later times, all after the point of reversal, and finally at the opposing rolling velocity to the initial conditions. Colour plots of the numerical solver developed here are shown in Figure 5.6 with the key for the colour plots shown in Figure 5.7, although this colour mapping does not apply to the interferometry results. It can be clearly seen that the expected saucer of higher film thickness does progress from right to left through the contact just after reversal.

For industrial applications there are certain values which are of particular interest to the user. These monitoring values usually include both the central and minimum film thickness as well as the central offset film thickness, H_{00}. For the reversal problem calculated above, these values are shown in Figures 5.8 and 5.9, alongside the numerical results calculated using the homotopy method of [126]. It is clear that there is very good agreement between the two sets of results, but the computation time was significantly

Reproduced by permission of Shell Global Solutions

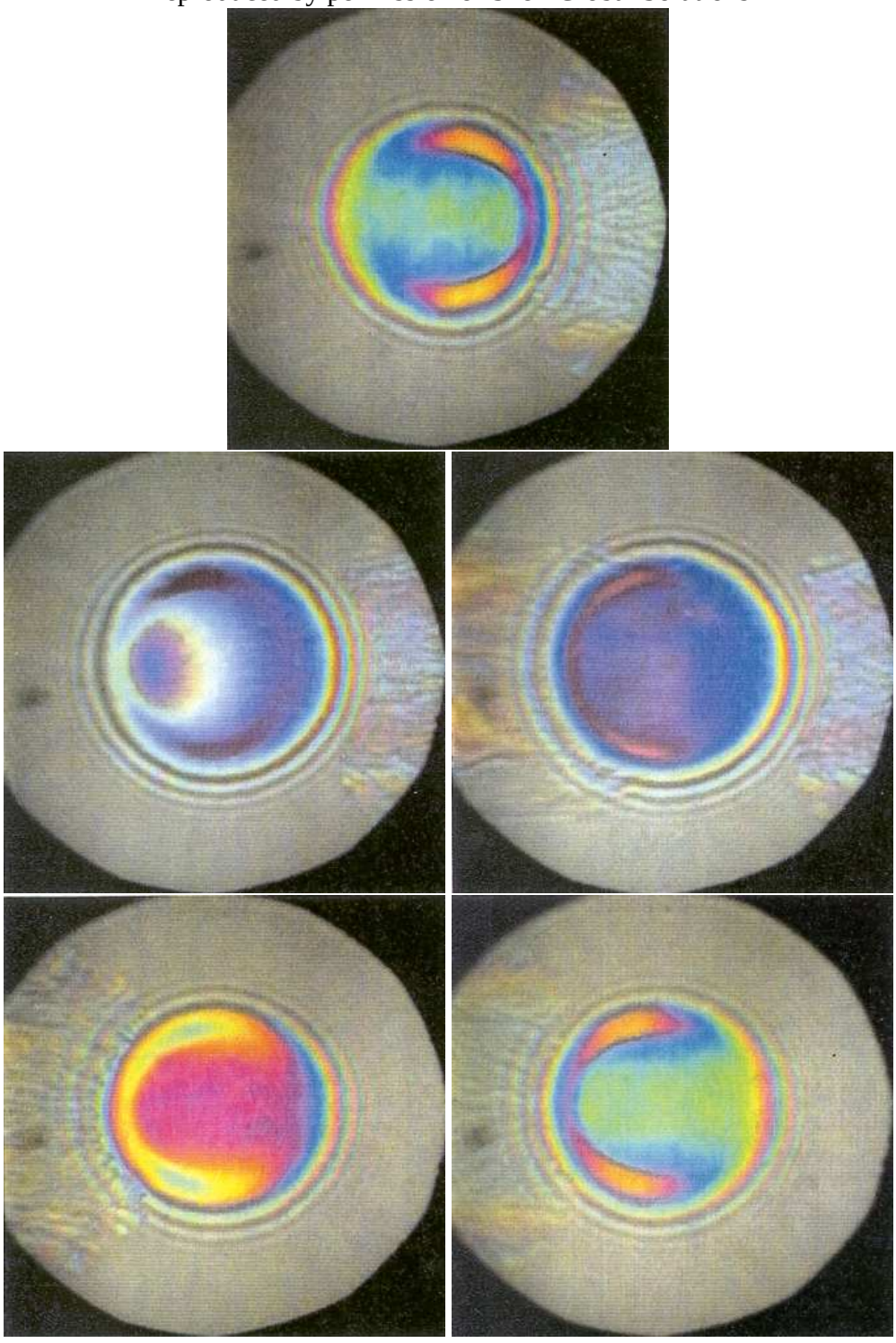


Figure 5.5: Reversal of entrainment experimental film thickness interferometry results [Scales, Rycroft, Horswill, Williamson] [128]

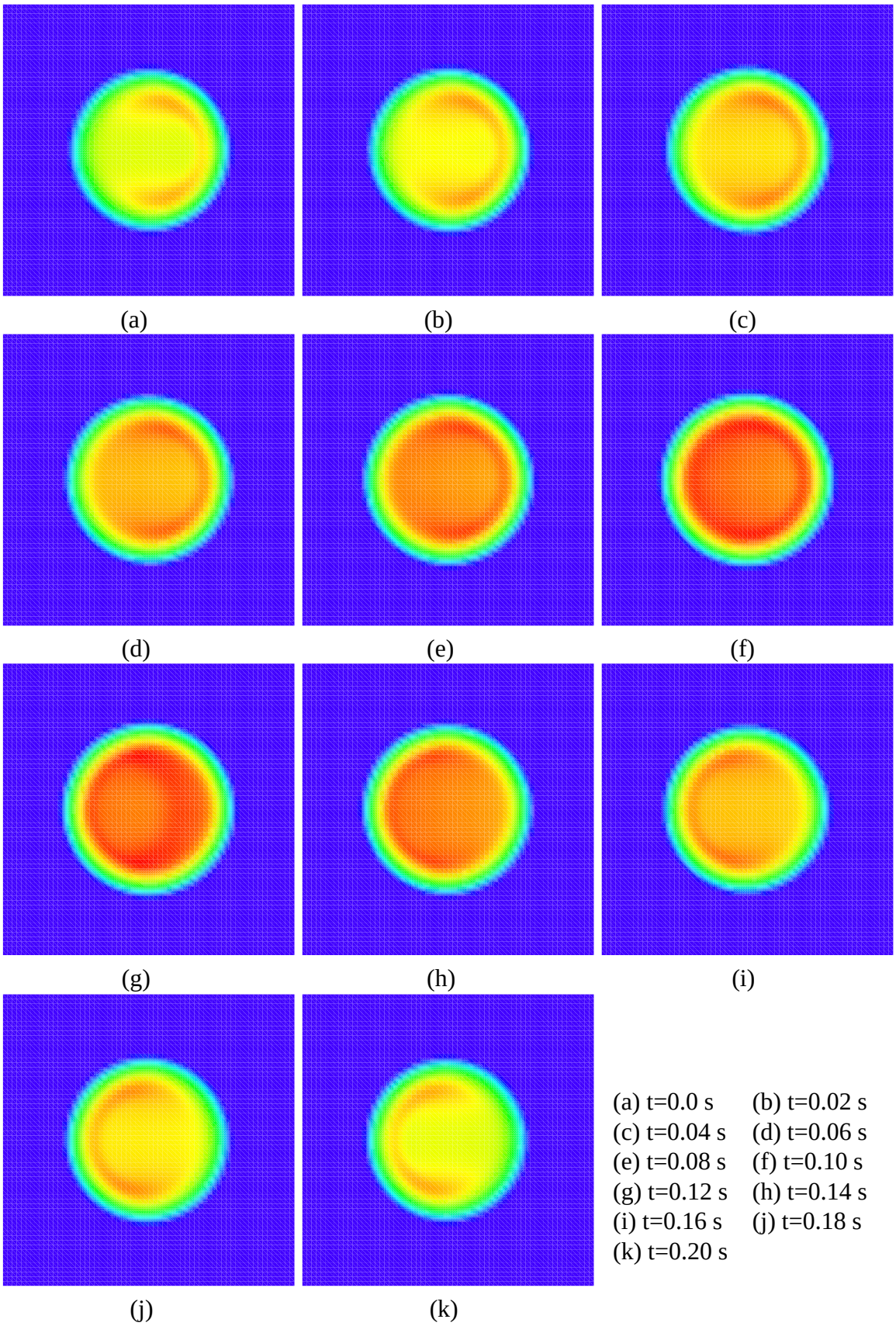


Figure 5.6: Film thickness numerical results for reversal of entrainment

Figure 5.7: Colour key for film thickness plots in Figure 5.6

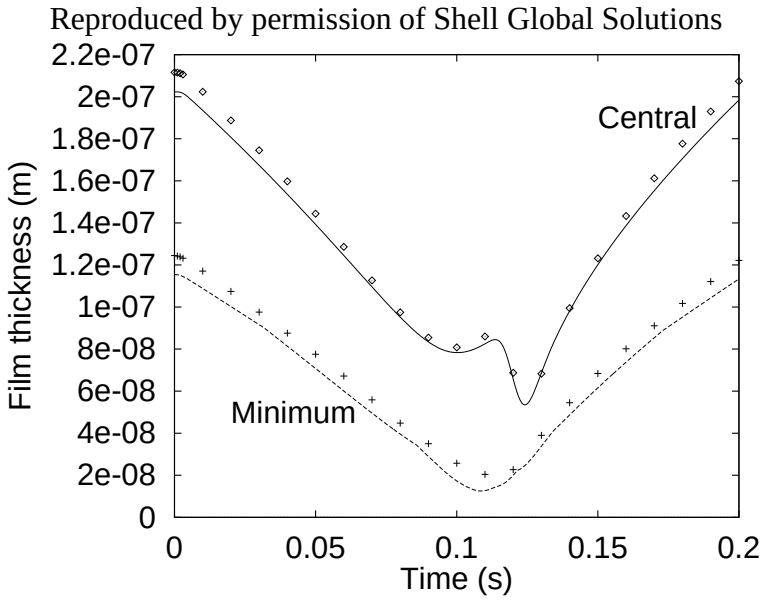


Figure 5.8: Minimum and central film thickness plots compared to previous numerical results of Scales (shown as points)

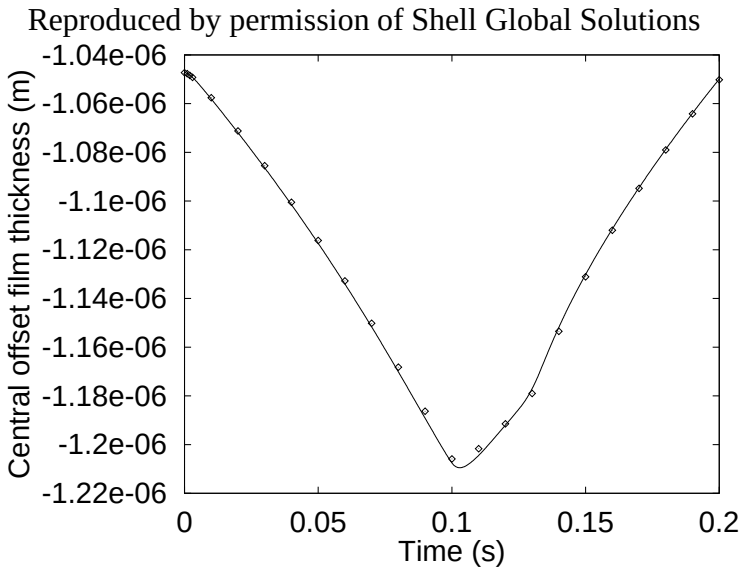


Figure 5.9: Central offset film thickness plot compared to previous numerical results of Scales (shown as points)

shorter (hours rather than days) using the method described here.

5.6.2.1 Prediction

The performance of the solver is governed by how quickly a suitably accurate solution can be obtained. The benefits of the use of multilevel techniques have already been shown in Chapter 4, but for transient problems there are other factors that can be varied. Theoretically, if the solver starts a timestep with a solution that is already within the allowed error tolerances, then no further work should be necessary. Whilst it is necessary to perform at least two multigrid cycles to establish this, as shown in Section 5.5, the better the initial approximation on the step, the greater the likelihood of a converged solution being produced quickly.

Historically, the problems associated with the use of coarser grids in this code, outlined in Section 4.5, made the use of the multigrid F-cycle of Brandt, as described in [140], impractical. In this method at the start of every new timestep the previous fine grid solution is restricted to the coarsest grid without any smoothing operations being undertaken, before the FMG algorithm is applied again.

Instead the method employed here is to use linear interpolation of the solutions on the previous two timesteps as an initial guess for the new fine grid solution. This is done for the pressure and the film thickness as follows (for the pressure):

$$P_{i,j}^{n+1} = P_{i,j}^{n} + \frac{\Delta T_{n+1}}{\Delta T_n}\left(P_{i,j}^{n} - P_{i,j}^{n-1}\right), \tag{5.28}$$

with H_{00} similarly predicted. The density and viscosity are pointwise calculations and are not computationally expensive. Initial approximations for these are therefore calculated using the predicted pressure.

The advantage of using this prediction can be seen in Figure 5.10 which shows how the use of prediction reduces the level of the root mean square residual at the start of each timestep. It can be seen that for much of the solution period the initial solutions using prediction are over two orders of magnitude better than those without. The reduction in work is seen by the reduction of the total number of fine grid smooths, showing a reduction in the number of multigrid cycles required to obtain a converged solution.

5.6.2.2 Convergence

In Section 5.4 the Shampine convergence test, Equation (5.26), was explained. In this test the convergence of the solution on an individual timestep is estimated by comparing the

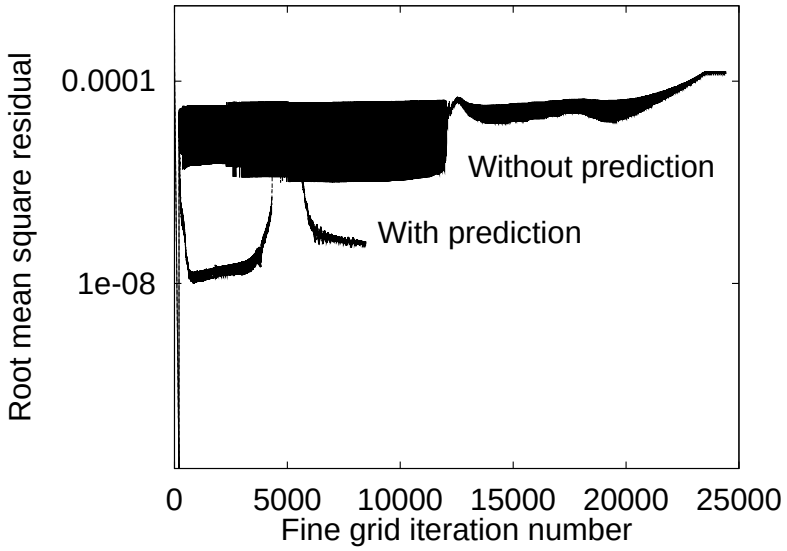


Figure 5.10: Residual level during reversal example, with and without prediction

current approximation obtained to the predicted solution for that step from the end of the previous step. In an example, such as that of reversal, the first time a timestep is taken the predicted solution will be the same as the solution at the end of the previous step. This may mean that convergence is not recognised straight away, since the correct temporal derivatives are not set up. The same situation will occur at any discontinuity in the physical behaviour of the example. This situation is discussed in relation to DAE problems by Petzold [118]. Apart from this situation, this convergence test should highlight non-linear behaviour in the solution.

In Figures 5.11 to 5.13 the number of multigrid cycles required per timestep are shown over the course of reversal. Each figure has two curves, showing the effects of testing for convergence of pressure or film thickness. This required number of iterations increases when approaching reversal as the non-linear effects dominate. Once the linearity of the solution change has re-established its prominence, then the required number of iterations per step falls again.

It was stated that the choice between estimating the convergence of $\underline{H}$ or $\underline{P}$ was quite important. It is, however, necessary to justify any choice in terms of the same overall solution being obtained using fewer iterations, as well as the known properties of the system. Three test cases are shown here. In Figures 5.11 and 5.12 prediction has not been used, with the value of tol in Equation (5.26) being 0.3 and 0.03 respectively. Figure 5.13 was calculated using prediction with tol=0.03. These three figures show the number of timesteps taken with either $\underline{H}$ or $\underline{P}$ being the test variable. It is clear that in all cases testing for pressure demands fewer iterations for the same tolerance. However, the three

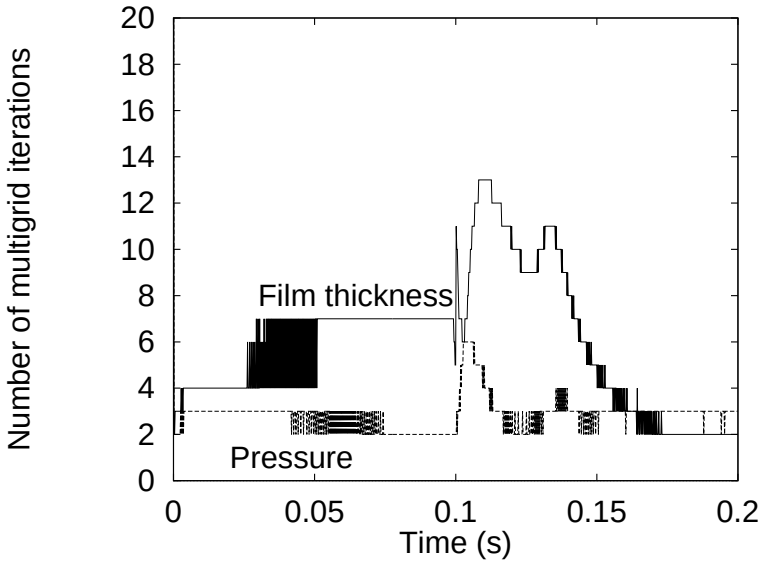


Figure 5.11: Reversal: Number of multigrid iterations per timestep when testing for pressure or film thickness in the Shampine convergence test with *tol*=0.3, without using prediction

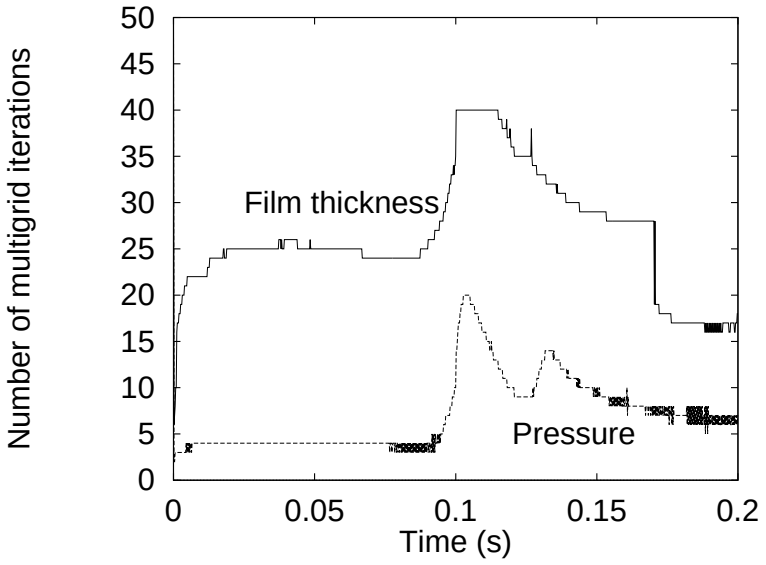


Figure 5.12: Reversal: Number of multigrid iterations per timestep when testing for pressure or film thickness in the Shampine convergence test with *tol*=0.03, without using prediction

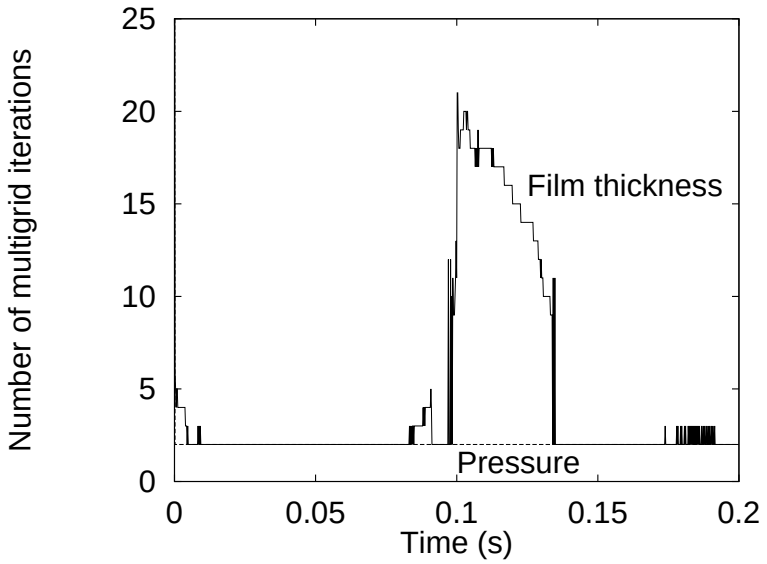


Figure 5.13: Reversal: Number of multigrid iterations per timestep when testing for pressure or film thickness in the Shampine convergence test with tol=0.03, using prediction

companion graphs, Figures 5.14 to 5.16, show the differences in the calculated solution values for central, minimum and central offset film thickness, and it can be seen that there is a significant difference in accuracy between the two methods. Clearly the film thickness test where more iterations have been done, will be more accurate. In the case of the higher tolerance prediction-less case this is about a 20% difference in the minimum film thickness at the point of reversal. Comparing the cases for tol=0.03 shows the advantages of prediction once again. Not only are fewer cycles required, but the solutions obtained are also more accurate, as can be seen away from the point of reversal in Figure 5.16 where the differences between the solutions are very small compared to those in Figure 5.15.

Overall these results confirm that the use of the algebraic variable, $\underline{H}$, is preferable. They do, however, indicate that care must be taken when choosing the tolerance for the required test, since too few iterations can lead to inaccuracies entering the solution.

5.6.3 Sinusoidally Varying Loads

The applied load across a contact is not always constant. Typical situations include gears engaging and disengaging, and vibrations of rolling element bearings. This vibrational behaviour was examined in detail by Wijnant [154] including an investigation of the case of a sinusoidally varying load. Beyond the first cycle periodic behaviour was observed. The example which follows is the same as the one presented in Section 5.5.2 of [154].

The applied load, F, is actually included in the solution of the EHL problem in the left

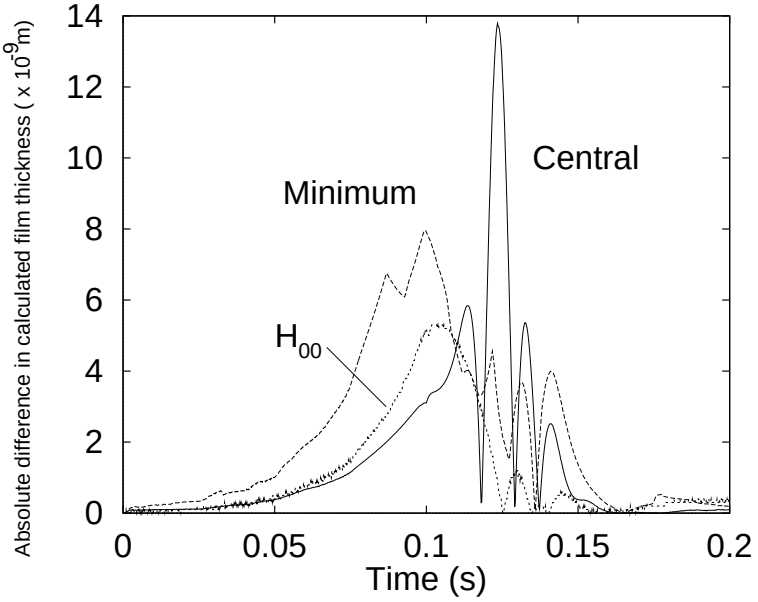


Figure 5.14: Reversal: Difference between calculated solutions of key variables, between testing for pressure or film thickness in the Shampine convergence test with *tol*=0.3, without using prediction

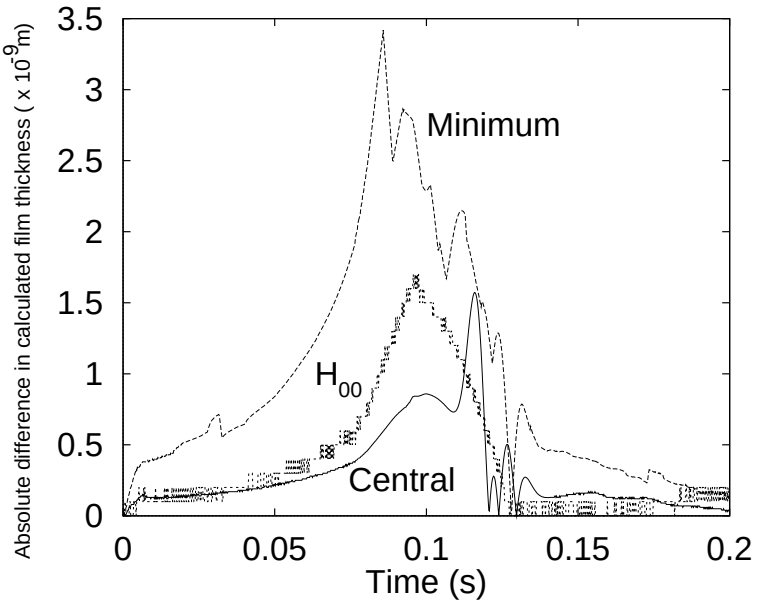


Figure 5.15: Reversal: Difference between calculated solutions of key variables, between testing for pressure or film thickness in the Shampine convergence test with *tol*=0.03, without using prediction

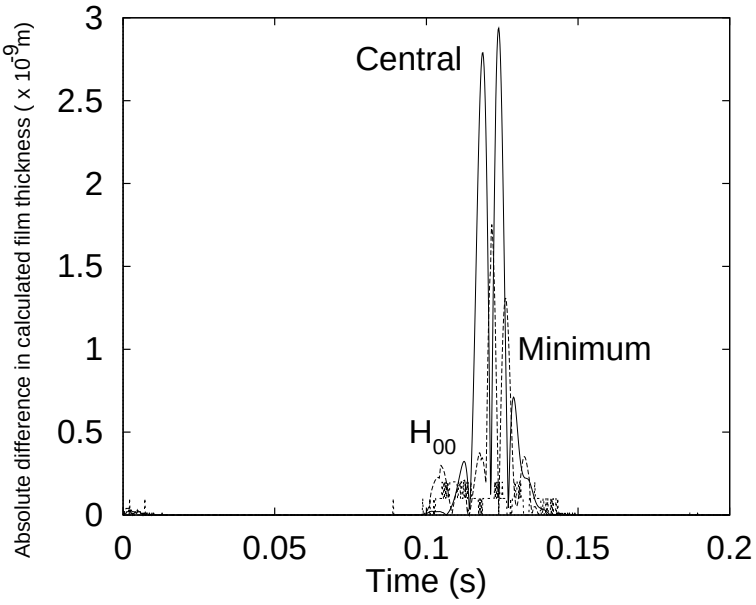


Figure 5.16: Reversal: Difference between calculated solutions of key variables, between testing for pressure or film thickness in the Shampine convergence test with *tol*=0.03, using prediction

hand side of Equation (5.5). However, this has already been non-dimensionalised from view. The results of Wijnant, and hence here, were calculated by instead using

$$\frac{2\pi}{3}\left(1+\mathscr{A}\sin(\Omega_e T)\right)=\Delta X\Delta Y\sum_{i=1}^{N_x}\sum_{j=1}^{N_y}P_{i,j}, \tag{5.29}$$

where $\mathscr{A}$ is the relative amplitude of the vibrations, and Ω_e is the excitation frequency. Values for the parameters used to solve the problem are shown in Table 5.3.

Three cases were investigated. The first, with $\Omega_e=\frac{\pi}{10}$, was a very slow oscillation,

Parameter	Value
Moes parameter, M	100
Moes parameter, L	5
Viscosity index	2.1×10^{-8} Pa^{-1}
Maximum Hertzian pressure	0.40 GPa
Material parameter, G	4729
Load parameter, W	1.18×10^{-7}
Speed parameter, U	6.25×10^{-13}
Oscillation amplitude, $\mathscr{A}$	0.1
Excitation frequency, Ω_e	$\frac{\pi}{10}$, π, 2π

Table 5.3: Parameters used for sinusoidally varying load transient example, after Wijnant

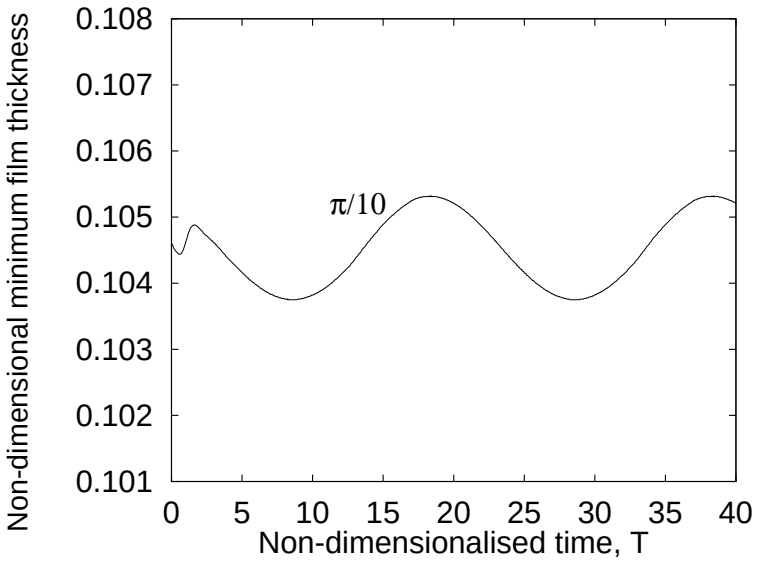


Figure 5.17: Central film thickness during sinusoidal load oscillation with $\Omega_e = \frac{\pi}{10}$

which was not considered by Wijnant. It was seen that with this period there were very few transient effects with an almost steady state solution being obtained at each timestep. Central and minimum film thickness results are plotted in Figures 5.17 and 5.18 respectively. The solution for the minimum film thickness does not appear as smooth as in the cases to follow. This is because the location of the minimum undergoes large positional changes. This movement is shown in Figure 5.19 where it can be seen moving over a range of X and Y points as the loading changes. A selection of results for the film thickness profiles in and around the contact area are shown in Figure 5.20. Although not a lot of change is discernible between them, it can be seen that with the changing loading the contact circle grows and diminishes in size, for example decreasing in radius along the second row and increasing along the third.

The pictures in Figure 5.20 have been generated by use of the pseudo-interferometry technique used by Lubrecht, Venner, and Wijnant. For this, the calculated film thickness is used to calculate pointwise intensities of the image using

$$\mathscr{I}(X,Y) = 0.5 + 0.5\cos\left(\frac{2\pi H(X,Y)}{\Lambda}\right), \tag{5.30}$$

for dimensionless wavelength Λ. This allows easy comparison between numerical results and experimental interferometry pictures. Throughout this section $\Lambda = 0.06$, as in [154], will be used.

The cases $\Omega_e = \pi$ and $\Omega_e = 2\pi$ both have significant squeeze effects in the film profile. A series of pseudo-interference plots of the film thickness for $\Omega_e = 2\pi$ is shown in Fig-

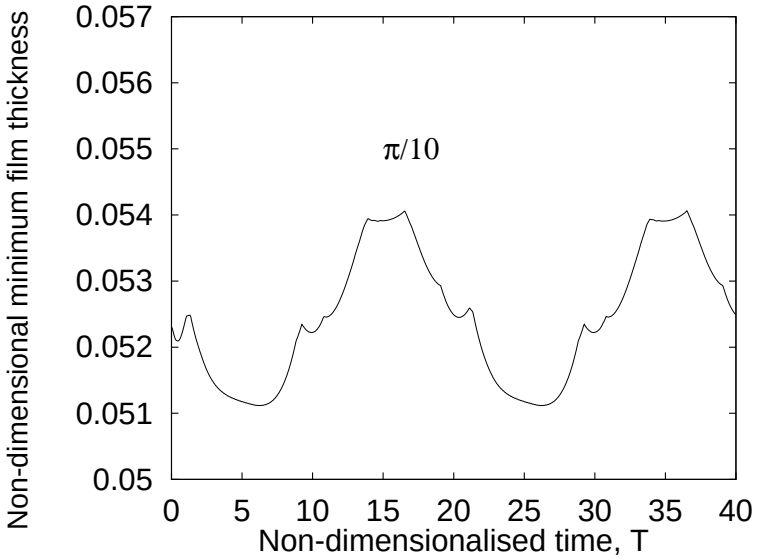


Figure 5.18: Minimum film thickness during sinusoidal load oscillation with $\Omega_e = \frac{\pi}{10}$

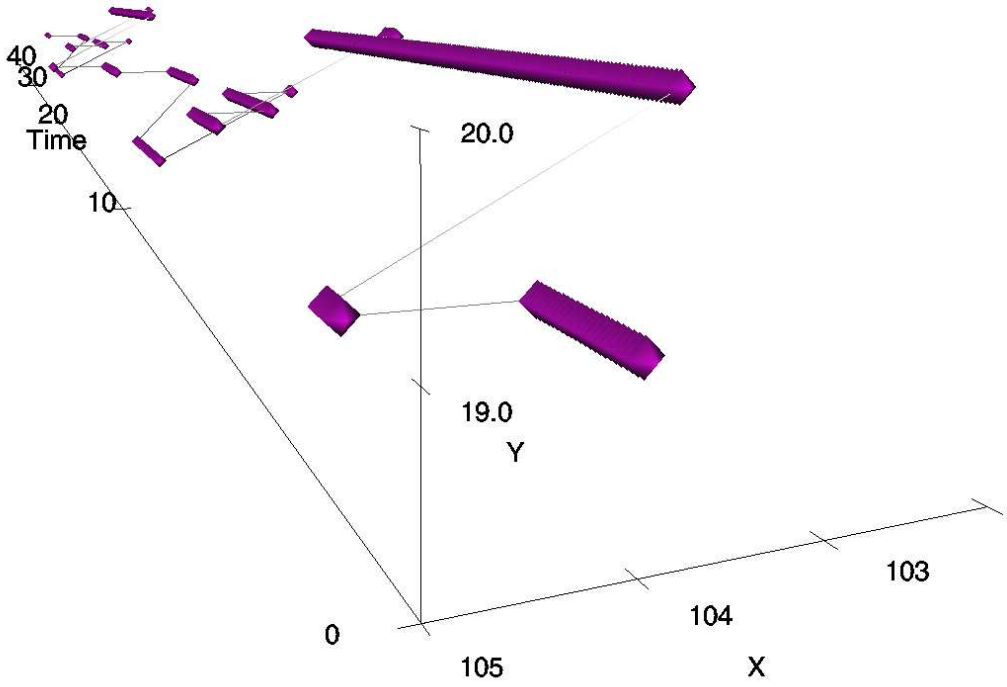


Figure 5.19: Location of minimum film thickness during sinusoidal load oscillation with $\Omega_e = \frac{\pi}{10}$

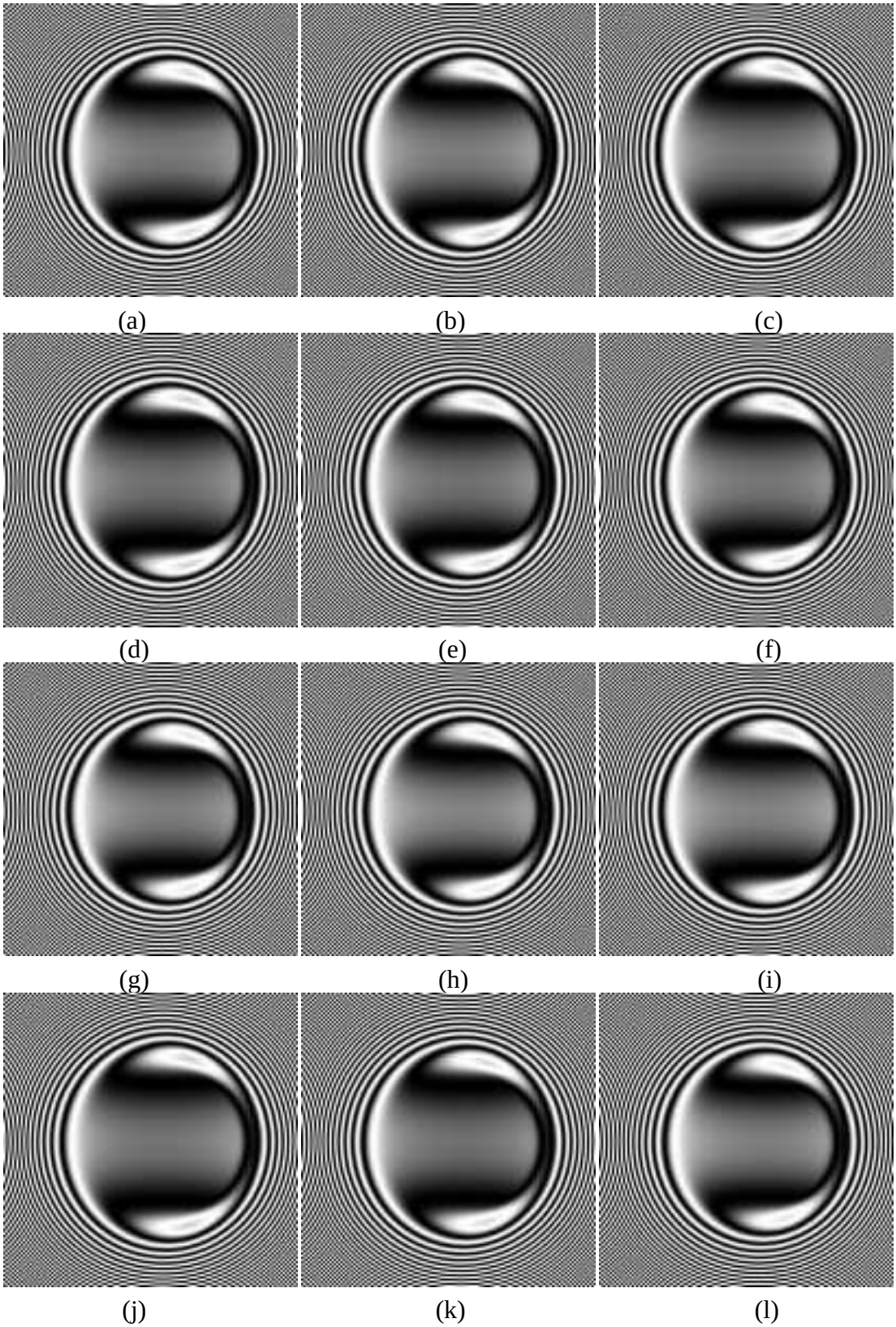


Figure 5.20: Pseudo-interferometry film thickness plots for sinusoidal load oscillation with $\Omega_e = \frac{\pi}{10}$ at times T = (a) 0.00, (b) 0.25, (c) 0.50, (d) 0.75, (e) 1.00, (f) 1.25, (g) 1.50, (h) 1.75, (i) 2.00, (j) 2.50, (k) 3.00 and (l) 3.50.

ure 5.21 where the oscillatory behaviour can be seen beyond $T = 1$. This is also observed in the behaviour of both the central and minimum film thickness, shown in Figures 5.22 and 5.23. In this case the movement of the minimum is smooth enough not to display the behaviour shown in Figure 5.18. It can be seen from Figure 5.24 that the difference in behaviour between the first and subsequent cycles does not change the profile of H_{00} in any way. A slight difference in amplitude of the H_{00} oscillation between the different frequencies can, however, be observed.

For completeness, centreline pressure solutions are shown in Figure 5.25 where the steady state pressure solution is included as a guide.

5.6.4 Surface Features

The study of surface features on EHL contacts enables more realistic surface geometries to be modelled. Even small deviations from a regular smooth parabolic contact can result in very different behaviour across the contact. The addition of geometrically regular asperities simulates roughness in a manner which can be easily implemented, and actually approximates the true surface roughness of real contacts, rather than the microscopic surface roughness.

The problem of surface roughness is currently receiving large amounts of attention. This is split between numerical calculations of amplitude reduction, primarily for line contacts, e.g. [51,70–73,75,94,146], and experimental results such as those of Guangteng *et al.* [55,56]. One of the cases considered in [55] is that of a sputtered surface, i.e. one with a regular pattern of conical deformations spread over it. A similar case is shown in Figure 5.26, where three dimensional profiles of pressure and surface geometry are plotted. Of particular note are the large pressure spikes of almost twice the height of the smooth surface case, corresponding to the small deformations of the contact. It can also be seen from the surface profile that on each of the bumps a separate EHL problem is evident, with individual sidelobes of minimum film thickness.

In this section two examples of transient EHL surface roughness problems will be solved. These are both based on examples taken from the literature, namely the transverse ridge considered by Venner and Lubrecht [142] and a circular dent passing through the domain, proposed by Ai and Cheng [1]. The surface deformations will be included into the mathematical model of Section 5.2 by modifying the undeformed geometry in Equation (5.4) from $\mathscr{G}(X,Y) = \frac{X^2}{2} + \frac{Y^2}{2}$.

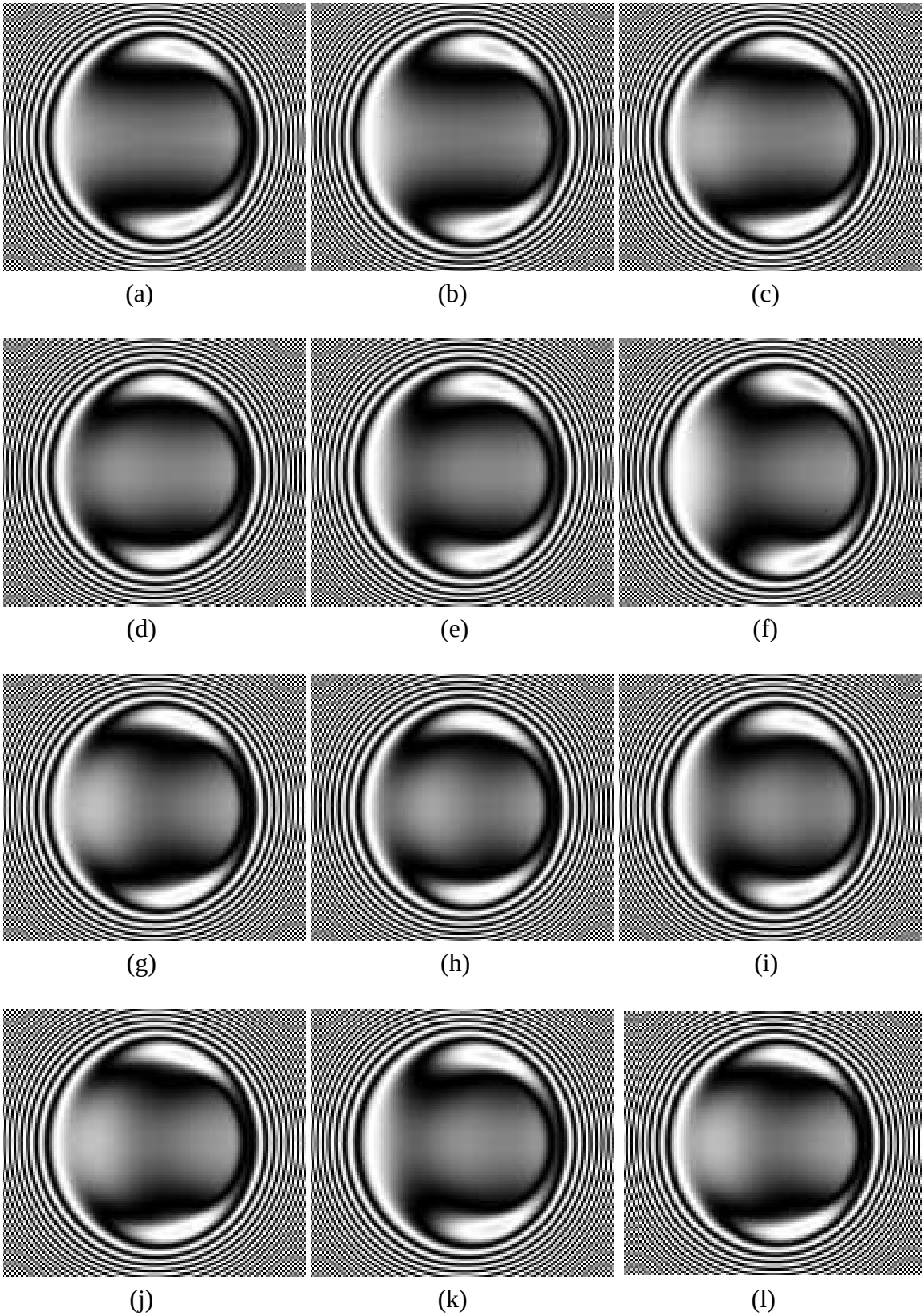


Figure 5.21: Pseudo-interferometry film thickness plots for sinusoidal load oscillation with $\Omega_e = 2\pi$ at times T = (a) 0.00, (b) 0.25, (c) 0.50, (d) 0.75, (e) 1.00, (f) 1.25, (g) 1.50, (h) 1.75, (i) 2.00, (j) 2.50, (k) 3.00 and (l) 3.50.

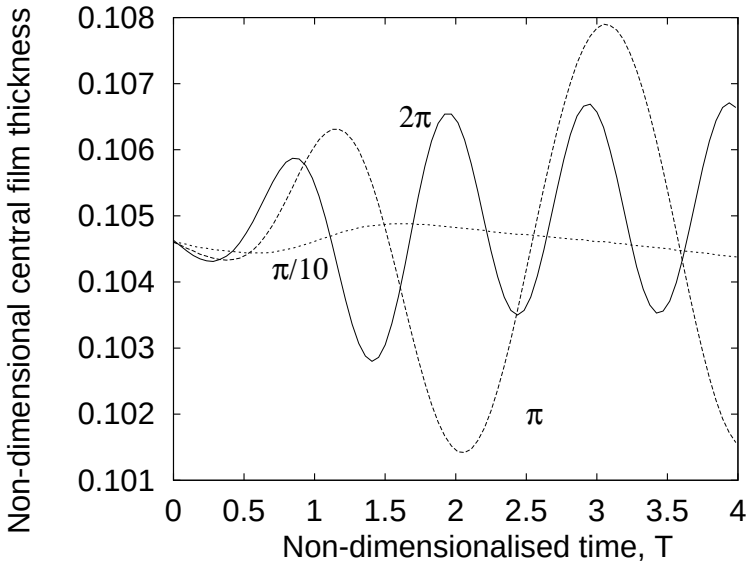


Figure 5.22: Central film thickness during sinusoidal load oscillation with $\Omega_e = \pi$ and $\Omega_e = 2\pi$

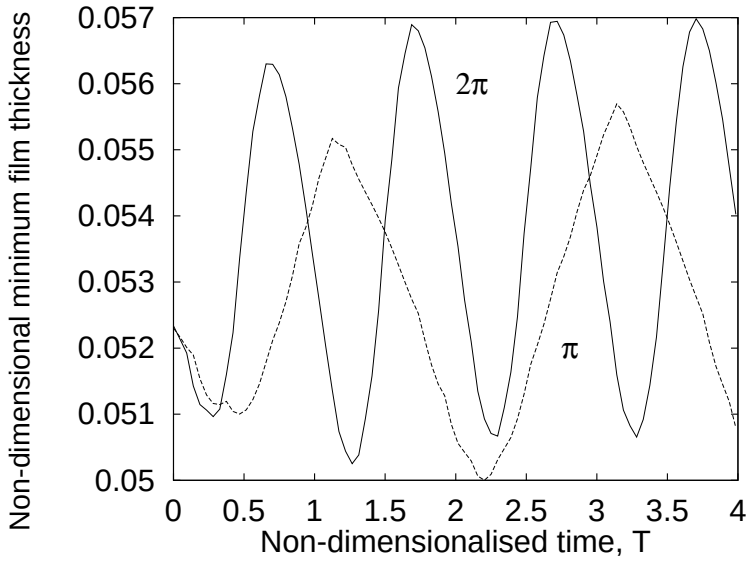


Figure 5.23: Minimum film thickness during sinusoidal load oscillation with $\Omega_e = \pi$ and $\Omega_e = 2\pi$

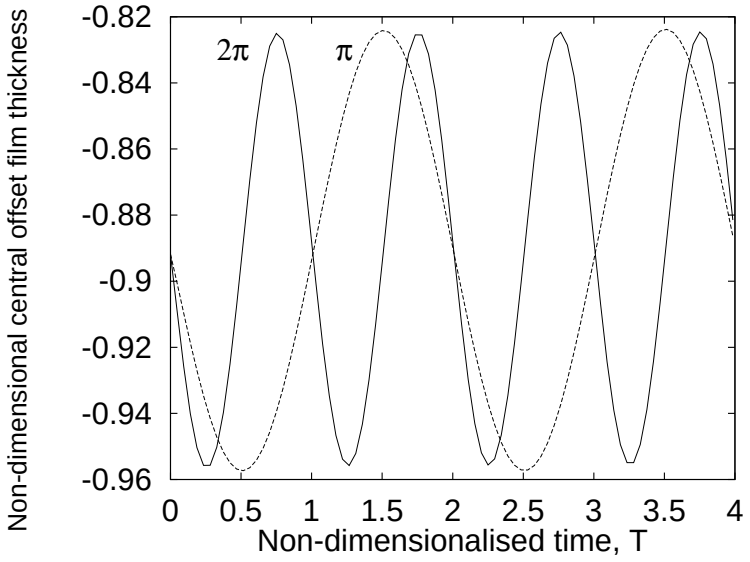


Figure 5.24: H_{00} variation with sinusoidally varying loads for $\Omega_e = \pi$, $\Omega_e = 2\pi$ and $\Omega_e = \frac{\pi}{10}$

5.6.4.1 Ridge Tracking

Following the experimental work of Kaneta et al. [83], Venner and Lubrecht [142] solved this transient example of a transverse ridge proceeding from left to right through the domain. They also considered the effects of the slide to roll ratio.

The undeformed geometry is given by

$$\mathscr{G}(X,Y) = \frac{X^2}{2} + \frac{Y^2}{2} - \mathscr{A} \times 10^{-10\left(\frac{X-X_d}{\mathscr{W}}\right)^2} \cos\left(2\pi \frac{X-X_d}{\mathscr{W}}\right), \tag{5.31}$$

where $\mathscr{A}$ is the dimensionless amplitude of the ridge,
$\mathscr{W}$ is the dimensionless wavelength of the ridge,
and $X_d(T)$ is the dimensionless position of the ridge given by

$$X_d(T) = X_d(0) + \frac{2u_2 T}{u_s}. \tag{5.32}$$

The parameters used in the computation of this problem are shown in Table 5.4. They correspond to a ridge of height 0.2 μm and a width of 0.07 mm, approximately what was used by Kaneta *et al.*, and precisely what was used by Venner and Lubrecht.

The results are shown in Figure 5.27 for calculated film thickness and pressure. The film thickness plots are presented, as in Section 5.6.3, using pseudo-interference graphs. This is to aid the comparison between the numerical results and the monochrome interfer-

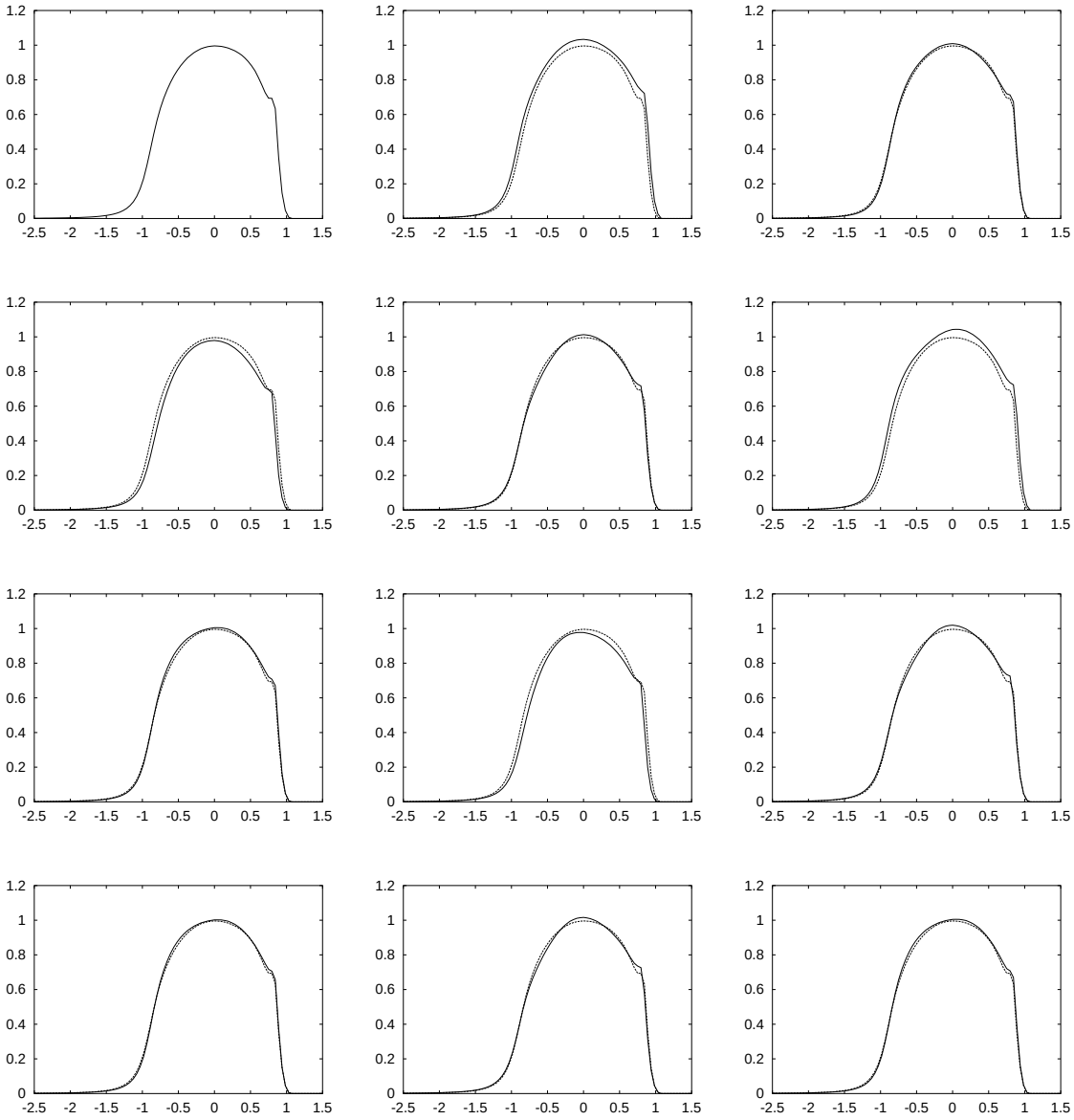
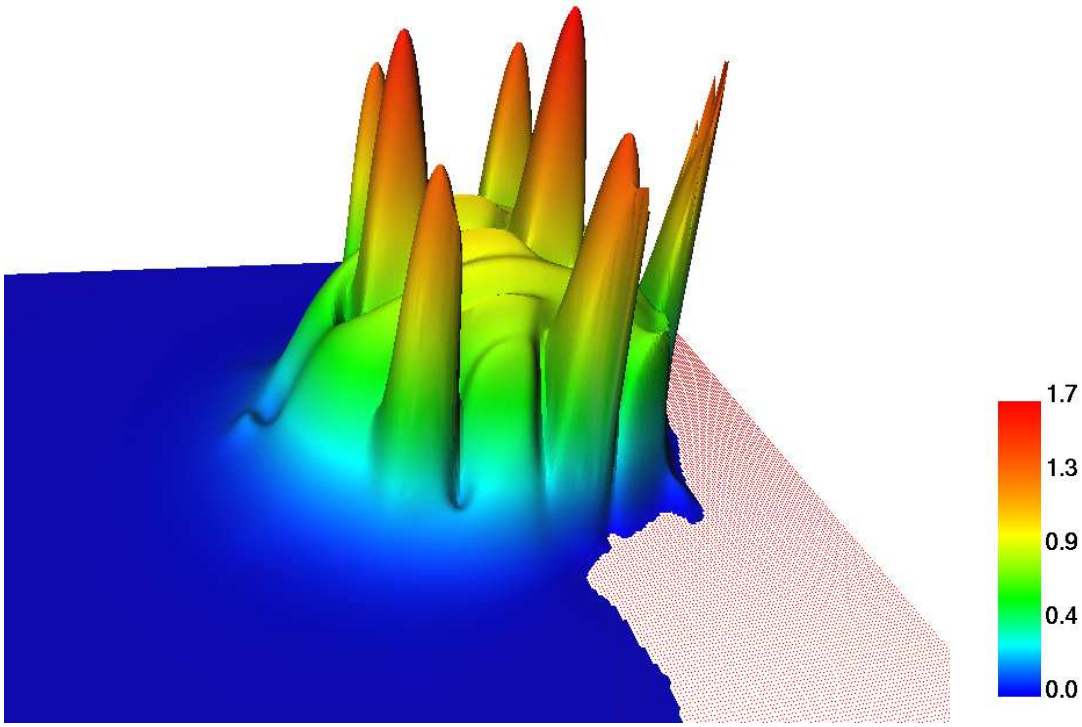
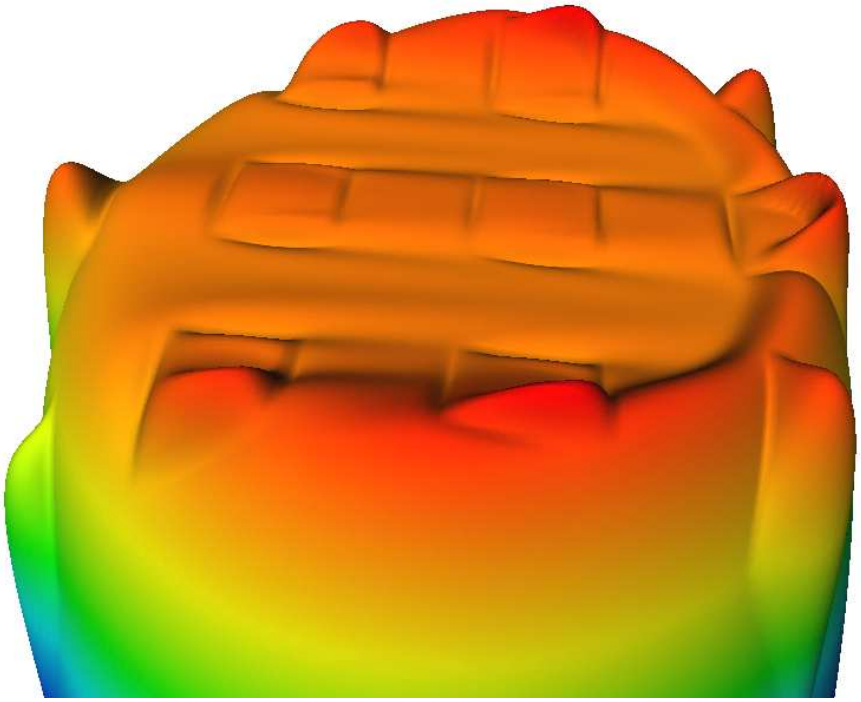


Figure 5.25: Centreline pressure solutions for a sinusoidally varying load with $\Omega_e = 2\pi$ at times T = 0.00, 0.25, 0.50, 0.75, 1.00, 1.25, 1.50, 1.75, 2.00, 2.50, 3.00 and 3.50, with the dotted line showing the T = 0.00 solution for each case



Non-dimensional pressure profile



Surface geometry profile

Figure 5.26: Pressure and surface geometry profiles for a sputtering example

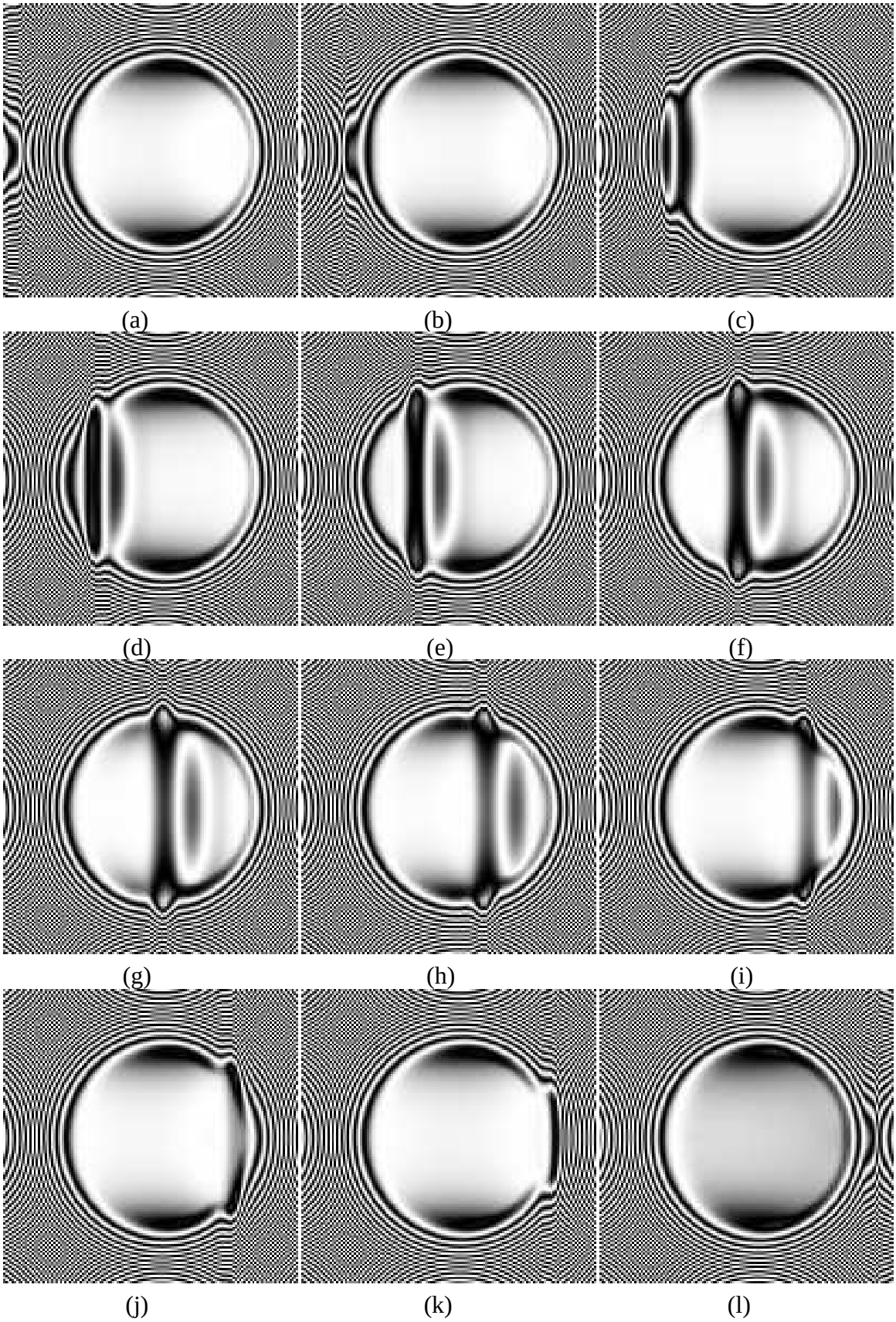


Figure 5.27: Pseudo-interferometry film thickness plots for transverse ridge example with ridge position, X_d = (a) -1.50, (b) -1.25, (c) -1.00, (d) -0.75, (e) -0.50, (f) -0.25, (g) 0.00, (h) 0.25, (i) 0.50, (j) 0.75, (k) 1.00, (l) 1.25

Parameter	Value
E'	1.17×10^{11} Pa
α	2.2×10^{-8} Pa^{-1}
η_0	1.22 Pa s
u_s	0.0215 m s^{-1}
R_x	0.0127 m
a	1.84×10^{-4} m
p_h	0.54×10^{9} Pa
Moes parameter, M	233
Moes parameter, L	5.42
Material parameter, G	2.62×10^{3}
Speed parameter, U	8.8×10^{-12}
Load parameter, W	2.0×10^{-6}
Ridge amplitude, $\mathscr{A}$	0.075
Ridge wavelength, $\mathscr{W}$	0.7

Table 5.4: Parameters used for transverse ridge example, after Venner and Lubrecht

ometry pictures of Kaneta. The dimensionless wavelength has been chosen to be $\Lambda = 0.05$ for the interferometry formula of Equation (5.30), again for comparison with [142].

Overall, it has been seen that the results presented are visually almost identical to those of Venner and Lubrecht. These, in turn, had shown good comparison for Kaneta's experimental results, except locally to the ridge where some differences occurred. These differences were probably the result of incomplete modelling of the lubricant as an isothermal Newtonian model was used [142].

5.6.4.2 Dent Tracking

This is similar to the ridge example, although rather than the asperity being convex, it is concave. Also the deformation of the surface is finite in all directions, with its direction of travel being along the centreline of the contact. The undeformed geometry of the contact is given by

$$\mathscr{G}(X,Y) = \frac{X^2}{2} + \frac{Y^2}{2} + \mathscr{R} \tag{5.33}$$

where the roughness, $\mathscr{R}$ defined by

$$\mathscr{R} = \begin{cases} \mathscr{A}\,[1+cos(\pi r_1)] & r_1 \leq 1 \\ 0 & \text{elsewhere} \end{cases} \tag{5.34}$$

Parameter	Value
α	2.0×10^{-8} Pa^{-1}
η_0	4.0 Pa s
p_h	2.02×10^{9} Pa
Moes parameter, M	578
Moes parameter, L	13.3
Material parameter, G	4616
Speed parameter, U	3.47×10^{-11}
Load parameter, W	1.39×10^{-5}
Ridge amplitude, $\mathscr{A}$	0.02
Radius of dent in X direction, L_X	0.30
Radius of dent in Y direction, L_X	0.30

Table 5.5: Parameters used for transverse ridge example, after Ai and Cheng

where

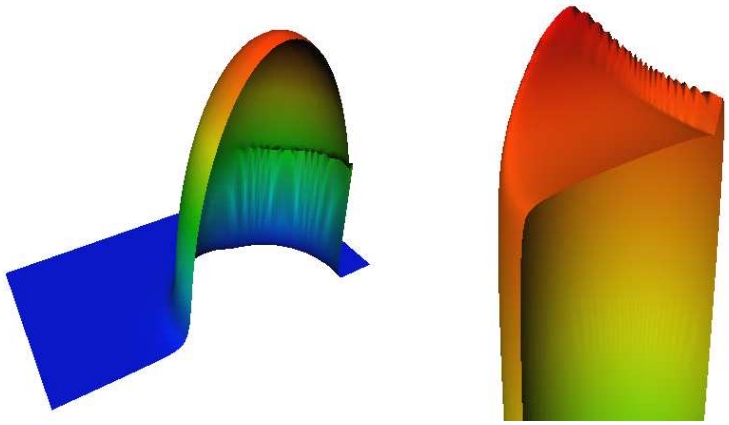
$$r_1 = \sqrt{\left(\frac{X - X_d}{L_X}\right)^2 + \left(\frac{Y}{L_Y}\right)^2} \tag{5.35}$$

where $\mathscr{A}$ is the dimensionless amplitude of the ridge,
L_X and L_Y are the dimensionless radii of the dent in the
X and Y directions respectively,
and $X_d(T)$ is the dimensionless position of the dent given by Equation (5.32).

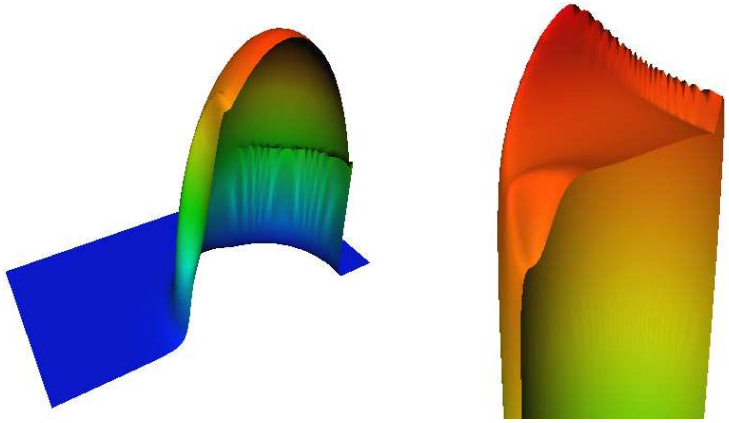
Using the parameters of Table 5.5 the example is similar to the one considered by Ai and Cheng [1]. This is a very heavily loaded case with a conical deformation travelling along the centreline. The amplitude of the deformation shown here is an order of magnitude larger than that considered in [1]. This provides more noticeable changes in the solutions obtained. In addition the lubricant model used is that described in Chapter 2 rather than the modified Barus equation used in [1]. The computation domain used was $X\in$[-2.5,1.5], $Y\in$[-2.0:2.0], with multigrid being applied between the finest level of 129×65 points and the coarsest of 17×9 for the half domain. The dent was positioned at X=-2.5 at T=0.

Three dimensional solution profiles for pressure and film thickness are shown in Figures 5.28 and 5.29. Only half of the domain has been shown to aid the visualisation of the dent's influence on the solution. The centreline solutions for non-dimensionalised pressure are shown in Figure 5.30. It can be seen that both before and after the dent is inside the contact region, there is almost no change in the solutions. However as the dent passes through the high pressure region just that very small surface deformation causes a significant change in the solution profile.

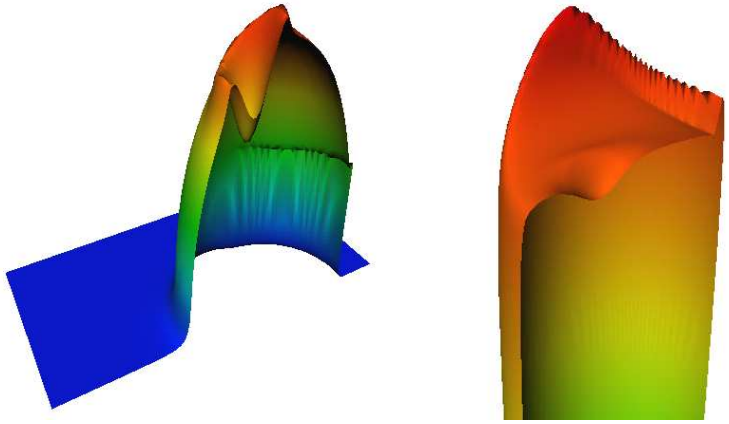
These results compare well against those of Ai and Cheng for the pure rolling case in-



(a) X_d = -1.5 (i) Pressure (ii) Film thickness

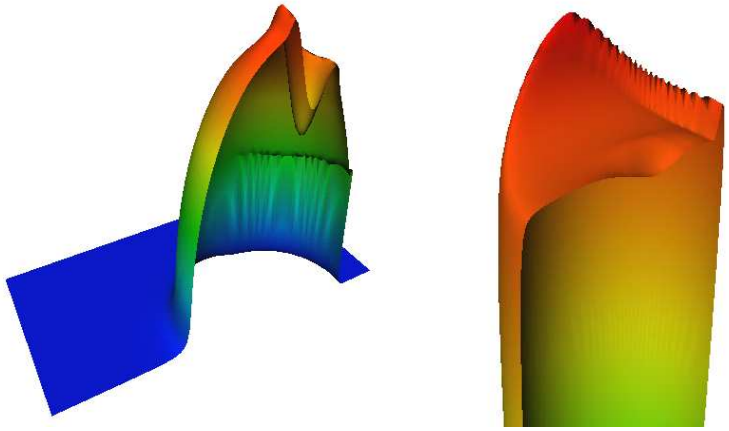


(b) X_d = -1.0 (i) Pressure (ii) Film thickness

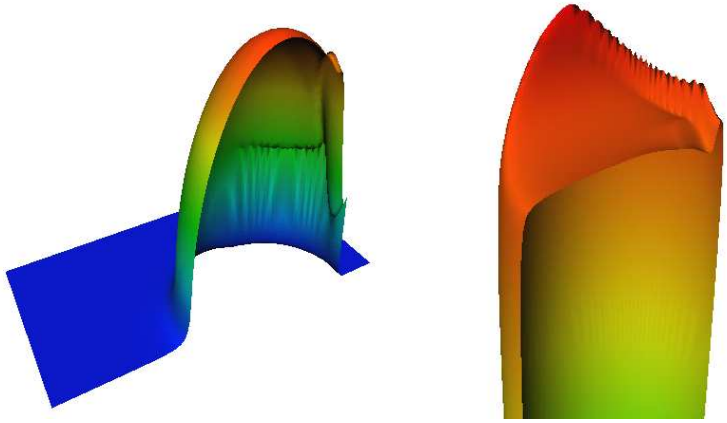


(c) X_d = -0.5 (i) Pressure (ii) Film thickness

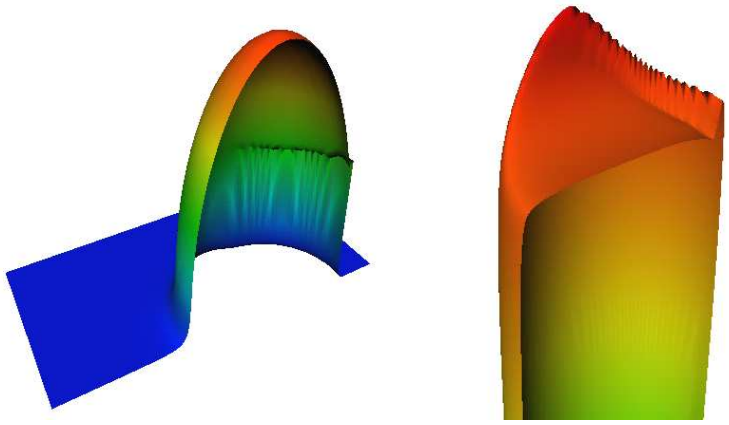
Figure 5.28: Pressure and film thickness profiles for dent example with dent position, X_d = (a) -1.50, (b) -1.00, (c) -0.50



(d) X_d = 0.0 (i) Pressure (ii) Film thickness



(e) X_d = 0.5 (i) Pressure (ii) Film thickness



(f) X_d = 1.0 (i) Pressure (ii) Film thickness

Figure 5.29: Pressure and film thickness profiles for dent example with dent position, X_d = (d) 0.00, (e) 0.50, (f) 1.00

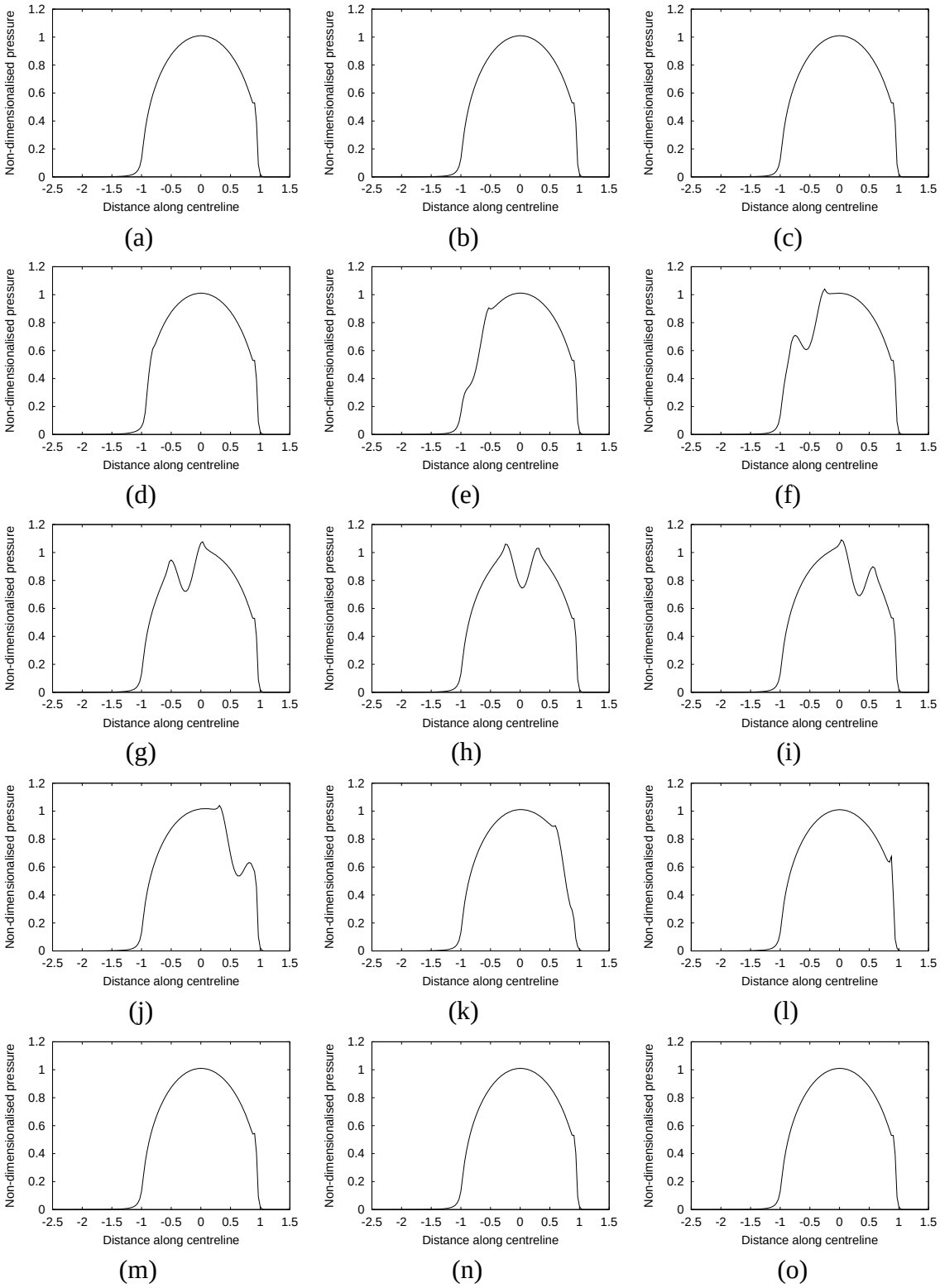


Figure 5.30: Centreline pressure solutions for dent example with dent position, X_d = (a) -2.50, (b) -2.00, (c) -1.50, (d) -1.25, (e) -1.00, (f) -0.75, (g) -0.50, (h) -0.25, (i) 0.00, (j) 0.25, (k) 0.50, (l) 0.75, (m) 1.00, (n) 1.25, (o) 1.50

vestigated here. Similar profiles for both pressure and film thickness have been presented, despite the differences in fluid model used.

5.7 Variable Timestepping

5.7.1 Introduction

Industry is now driving for solutions to be calculated to more realistic EHL problems, in a fast and robust manner. For transient cases this varies from variable loading and contact speeds, through complicated rheological behaviour involving fluid memory, to solutions incorporating surface features. This requires that the solvers being used must be designed to be extensible to meshes of several thousand points in each direction, and, importantly, that the solutions can be obtained sufficiently quickly.

The ability to reduce the work needed for individual problems is thus paramount. The mathematical and numerical analytic techniques behind many existing solvers combine traditional finite difference meshes with innovative multilevel methods, as has been shown in the preceding chapters.

In many other application areas in which the rate of change of the solution does not remain constant with time, it has proved beneficial to vary the timestep to control the error in the solution [20]. The methods, thus, have obvious applications to EHL calculations. In Section 5.7.2 this will be examined, and a method for variable timestepping proposed. Examples of variable timestepping in action will be presented in Section 5.7.3. These will show the potential speed-up over fixed timestep cases, whilst also highlighting the accuracy of the solutions obtained. The example of shock loading in this section demonstrates a powerful application of variable timestepping where fixed step solves would not be cost effective.

5.7.2 Changing Timestep Size

For transient numerical calculations the choice of correct timestep size is critical. If the timestep is too large then important physical features may be missed should they have a smaller timescale than the stepsize. Also, the calculated result may have larger local errors than are desirable for an accurate solution some timesteps later. Equally choosing a very small stepsize, may, at best, lead to a large amount of computational work for very small changes in the solution; at worst, result in solutions diverging, for example due to the magnification of temporal gradients. This is due, not to the stability of the problem, but the convergence properties of the non-linear solver outlined in Chapter 4, and highlighted

in Section 5.6.1. For example, experiments have shown that should the timestep become very small then any corrections made may amplify, rather than reduce, the errors in the solution unless very small under-relaxation parameters are used.

In EHL solutions ΔT has always been chosen to be fixed. Whilst for early transient solutions it was chosen to be larger than ΔX, the choice ΔT=ΔX has been pioneered by Venner and collaborators. This was introduced to minimise the total discretisation error for the Standard Upstream Second Order discretisation scheme, Equation (5.10), which they employ. Wijnant [154] additionally proposed the use of ΔT=$2\Delta X$ for the Narrow Upstream Second Order scheme.

The optimal choice of timestep is governed by successfully relating the spatial error of the solution, with the time error. It is well established in the ODE literature, e.g. [13,131], that controlling the local (temporal) error per step, so that the spatial error dominates, provides efficient, reliable algorithms. This approach, therefore, requires estimates of both components of the error.

Let the continuous equation system, defined by Equations (5.3) and (5.4), have an exact solution $\underline{u}(t)$, and the discretised equation system, defined by Equation (5.20), have exact solution $\underline{U}(t)$. If, at time t, the numerical approximation to the solution of the system is $\underline{\tilde{U}}(t)$, then the total error, $E(t)$, is defined by

$$\begin{aligned} \underline{E}(t) &= \underline{u}(t) - \underline{\tilde{U}}(t) \\ &= (\underline{u}(t) - \underline{U}(t)) + (\underline{U}(t) - \underline{\tilde{U}}(t)) \\ &= \underline{e}(t) + \underline{g}(t), \end{aligned} \tag{5.36}$$

where $\underline{e}(t) = \underline{u} - \underline{U}$ represents the spatial discretisation error, and $\underline{g}(t) = \underline{U} - \underline{\tilde{U}}$ is the global error in the time integration. Given that a solution has been discretised in space to a particular degree of accuracy, $\underline{e}(t)$, it is not worthwhile solving the transient part to a much higher degree of accuracy, but equally this transient error $\underline{g}(t)$ must not degrade the spatial accuracy.

The strategy employed here is similar to that described in [20] which is used in DASSL, a package designed to solve both index zero and index one DAE systems.

As outlined in Section 5.4, there is a free choice as to whether it is the errors in $\underline{P}$ or in $\underline{H}$ which are controlled. Intuitively, because the system given by (5.17) is solved for $\underline{P}$, and (5.18) is solved for $\underline{H}$ using this $\underline{P}$, it seems sensible to control the errors in $\underline{P}$. This was the approach suggested in [108]. However, it is the area inside the contact region where the most change is taking place, and this is dominated by the wedge and squeeze terms in (5.3). This depends upon the film thickness, $\underline{H}$, which is also the algebraic

variable of the system. Experiments have confirmed that controlling these errors requires significantly less work per timestep and less timesteps are required. The error tests will therefore be formulated for the variable $\underline{H}$. However, note that if $\underline{P}$ is chosen instead, the only points to be considered for the error tests are non-cavitation ones.

Once the Shampine convergence test has been satisfied, a local error calculation is undertaken to establish a new timestep size. Since the EHL problem is a non-linear DAE system, the LU decomposition of the system is not available, and hence the approaches described in [118] are not used. Instead, the local truncation error will be used to estimate the local error over the step. Defining the local truncation error for $\underline{P}$, $\underline{leP}$, as in [131, page 355], by:

$$\underline{leP} = \frac{1}{2}\left(P_n - P_n^{pred}\right), \tag{5.37}$$

and $\underline{leH}$ similarly, then the equations for these errors, in the same form as Equation (5.4.9) in [20], are

$$\begin{bmatrix} -1-\Delta T\frac{\partial \underline{F}_1}{\partial \overline{\rho}H} & -\Delta T\frac{\partial \underline{F}_1}{\partial P} \\ -\Delta T & \Delta T\,K \end{bmatrix}\begin{bmatrix} \underline{leH} \\ \underline{leP} \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ 0 & 0 \end{bmatrix}\frac{1}{2}\begin{bmatrix} H_n - H_n^{pred} \\ P_n - P_n^{pred} \end{bmatrix}. \tag{5.38}$$

This gives us a relationship between the local truncation errors in $\underline{H}$ and $\underline{P}$:

$$\underline{leH} = K\,\underline{leP}, \tag{5.39}$$

where K is the film thickness integration kernel matrix. It is possible to rewrite the first equation of (5.38) as the standard estimate for the local truncation error:

$$-\Delta T\left(\frac{\underline{leH}}{\Delta T} + \frac{\partial \underline{F}_1}{\partial \overline{\rho}H}\underline{leH} + \frac{\partial \underline{F}_1}{\partial P}\underline{leP}\right) = -\frac{\left(\underline{H}_n - \underline{H}_n^{pred}\right)}{2}. \tag{5.40}$$

Since these Jacobians are never calculated, consider Taylor's Theorem for two variables:

$$\underline{F}_1\left(\overline{\rho}\underline{H} + \underline{leH}, \underline{P} + \underline{leP}\right) \approx \underline{F}_1\left(\overline{\rho}\underline{H}, \underline{P}\right) + \frac{\partial \underline{F}_1}{\partial \overline{\rho}H}\underline{leH} + \frac{\partial \underline{F}_1}{\partial P}\underline{leP} + h.o.t. \tag{5.41}$$

Assuming that the residual on the timestep, $\underline{F}_1\left(\overline{\rho}\underline{H} + \underline{leH}, \underline{P} + \underline{leP}\right)$ is zero, substitution into Equation (5.40) gives the following equation for the local errors $\underline{leH}$ and $\underline{leP}$:

$$\underline{F}_1\left(\underline{P}_n + \underline{leP}, \overline{\underline{\rho}H}_n + \underline{leH}\right) - \frac{\overline{\underline{\rho}H} + \underline{leH} - \overline{\underline{\rho}H}_n}{\Delta T} = \frac{1}{2\Delta T}\left(H_n - H_n^{pred}\right). \tag{5.42}$$

Defining $\tilde{\underline{P}} \equiv \underline{P}_n + \underline{leP}$ and $\tilde{\underline{H}}$ similarly, then the equation for the local error (5.42) may

then be rewritten in the same form as Equation (5.17) with a different right hand side:

$$\underline{F}_1\left(\underline{\tilde{P}}, \underline{\overline{\rho}\tilde{H}}_n\right) - [\underline{\overline{\rho}\tilde{H}}_n]' = \frac{\left(H_n - H_n^{pred}\right)}{2\Delta T}. \tag{5.43}$$

This equation may then be solved for $\underline{\tilde{H}}$ using the standard EHL multigrid algorithm with right hand side $\frac{1}{2\Delta T}\left(H_n - H_n^{pred}\right)$. Therefore, in summary, to estimate the local error on a timestep, after a sufficiently converged solution has been obtained, two or three more V-cycles are carried out to obtain solutions, $\underline{\tilde{P}}$ and $\underline{\tilde{H}}$, to the local error problem.

Once these new solutions are calculated, an estimate of the total local error in $\underline{H}$, may be defined as

$$\|\underline{leH}\|_\omega = \left\|\underline{\tilde{H}}_{n+1} - \underline{H}_{n+1}\right\|_\omega, \tag{5.44}$$

where $\|\cdot\|_\omega$ is a weighted root mean square L_2-norm, as used in DASSL [20] defined by

$$\|\underline{H}\|_\omega = \sqrt{\frac{1}{N_x N_y}\sum_{i=1}^{N_x}\sum_{j=1}^{N_y}\left(\frac{H_{i,j}}{\omega_{i,j}}\right)^2}, \tag{5.45}$$

with weights, $\omega_{i,j}$ defined by

$$\omega_{i,j} = ATOL + H_{i,j}^{(0)} RTOL \tag{5.46}$$

which are themselves given in terms of the predicted solution at that mesh point on that timestep, $H_{i,j}^{(0)}$, and the absolute and relative error tolerances, $ATOL$ and $RTOL$ respectively. These tolerances have been chosen to be dynamically defined by

$$ATOL = \frac{1}{10}\sqrt{\frac{1}{N_x^C N_y^C}\sum_{I=1}^{N_x^C}\sum_{J=1}^{N_y^C}\left[\tilde{H}_{i,j} - \tilde{H}_{I,J}\right]^2} \tag{5.47}$$

and

$$RTOL = ATOL, \tag{5.48}$$

for fine mesh points (i,j) with coincident coarse points (I,J).

Once the local error has been established, it is then necessary to use this information to calculate the most desirable stepsize for the next timestep. The method chosen here is that of Shampine and Gordon [132], where any change to the stepsize is governed by the value of r in

$$r = (2\|\underline{leH}\|_\omega)^{\frac{-1}{k+1}}, \tag{5.49}$$

with k being the order of the method ($k = 1$ for the Backward Euler Method). The method of [132] suggests that the new stepsize should be given by

$$\Delta T_{n+1} = r\Delta T_n, \tag{5.50}$$

subject to some limitations described below.

These tests now mean that the code itself relates future timestep sizes to the magnitude of the local error. If the error is small, e.g. $r > 1.5$ in Equation (5.49), then the stepsize may be increased for the next timestep. If the error is 'too large', $r < 0.9$, then the stepsize is reduced, either for the following step, or, if the current step is considered to have failed, the current timestep may be retaken with a new stepsize. There is also a 'comfort region' in between these extremes where the stepsize is left unchanged.

Limits are also imposed on when, and by how much, the timestep size may change. In some codes, as is used here, it is never allowed to change up or down by more than a factor of 2, which is a standard approach in ODE initial value problem solvers. This helps, both in terms of keeping temporal derivatives of similar scales, and in keeping a check on what changes are allowed. A safety factor – usually of just one timestep – prevents the stepsize increasing too rapidly. The size is, however, allowed to reduce as often as necessary to capture features in the solution. Maximum and minimum timesteps may be specified by the user before runtime. These allow controls to be placed on the code to stop the stepsize diminishing towards nothing, for example, if it is failing at some point, and to impose physical constraints to the individual problem being solved: e.g. if the problem is being solved on T=[0.0, 1.0] then there would be no point in a ΔT_{max} of 0.5, but 0.05 could be acceptable.

5.7.3 Examples

Any modification to the numerical methods used to solve a problem must both enable accurate solutions to be obtained, and ensure that any additional computational overhead is minimal. In the case of variable timestepping this is broadened to solutions of similar accuracy as the best fixed timestep results, in a reduced computational time. It will also enable some problems which were previously unrealistic in a fixed ΔT=ΔX sense, to be solved.

In this section three examples the use of variable timestepping will be presented. In Section 5.7.3.1 the reversal of entrainment example will be tackled again, and both accuracies and computational timings compared. This example is useful because of the two linear periods of change at the start and at the end, with the nonlinear behaviour occurring

in the middle. This idea is also seen in Section 5.7.3.2 where a surface feature proceeds through the contact, and hence both before and after the feature influences the solution inside the contact region, the expected behaviour is almost a steady state, whereas large transient effects occur as it passes through the centre of the contact.

The third example, presented in Section 5.7.3.3 is that of *shock loading*. This is an industrially relevant example modelling the changes in conditions as gears interlock, and then separate again. Here the timestep will be expected to be small over the very short timescale of loading (or unloading) and then increase as the new steady state conditions are reached.

5.7.3.1 Reversal

A clear example of the benefits of using variable timesteps can be seen in the case of reversal of entrainment. This test case is that used in [128] and presented above in Section 5.6.2. The most interesting, and most non-linear part of this example is the formation at the point of reversal (0.1 s) of a saucer of viscous fluid which then proceeds across the domain (towards the new outflow) before the deformation pattern re-adopts its characteristic horseshoe shape. All physical parameters are as given previously in Table 5.2, with the non-dimensionalisation as in Table 5.1. The use of variable timestepping should pick out the transient effects around reversal with smaller timesteps than those preceding and following reversal, where larger timesteps may be more appropriate.

Precise direct comparison of accuracies between transient calculations using different timestep sizes is not trivial. In this section the results obtained will be compared for certain notable variables, namely central and minimum film thickness, and the central offset film thickness. This will be done between fixed and variable timestepping runs on two different finest grid levels.

The general accuracy of the approaches can be compared by visualising the central and minimum film thickness results throughout the solve. This is done in Figure 5.31 for a level 5 finest mesh with accuracy tolerance tol in Equation (5.26) set to be 0.3, in Figure 5.32 with tol=0.03 and in Figure 5.33 for a level 6 finest mesh with tol=0.3. In each graph the lines indicate the fixed timestep solution, and the points indicate the results at each timestep of the variable timestepping case. All three cases show that for the variable timestepping cases, the timesteps are much closer together around reversal than in the linear sections. It is clear that these results are very close to the fixed step results.

A more detailed examination of these differences is presented in Table 5.6. This shows the values of the central and minimum film thickness, and the central offset film thick-

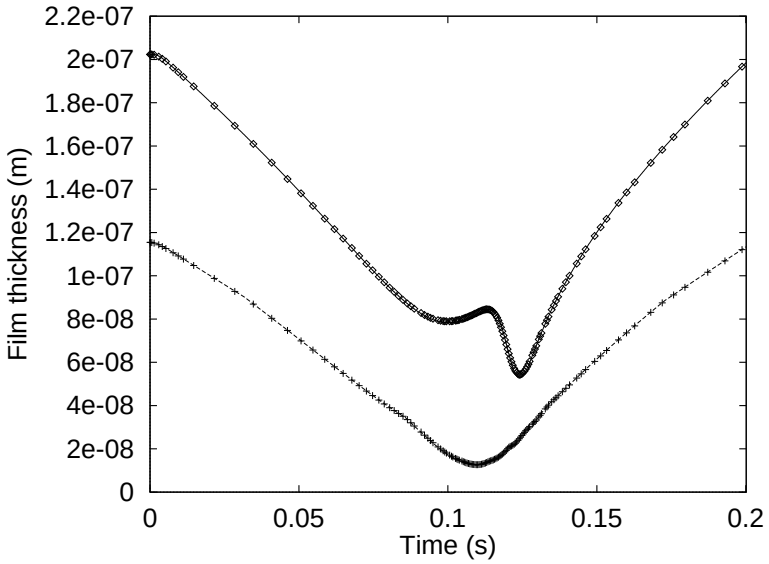


Figure 5.31: Central and minimum film thickness during reversal with, and without variable timestepping on grid 5, tol=0.3, where lines show the fixed step case and points are the variable timestepping results

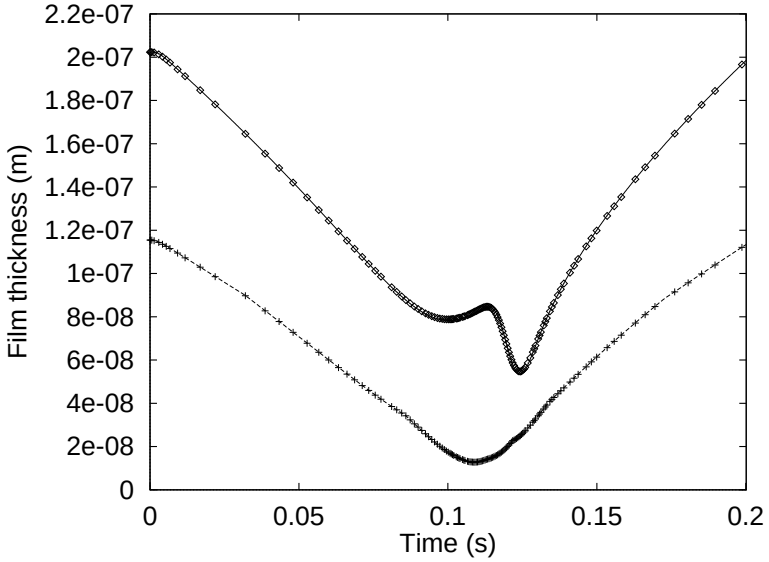


Figure 5.32: Central and minimum film thickness during reversal with, and without variable timestepping on grid 5, tol=0.03, where lines show the fixed step case and points are the variable timestepping results

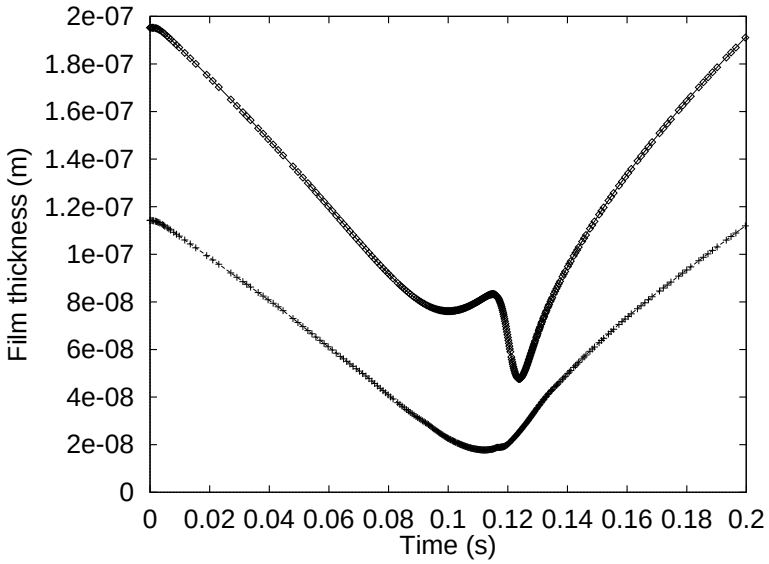


Figure 5.33: Central and minimum film thickness during reversal with, and without variable timestepping on grid 6, where lines show the fixed step case and points are the variable timestepping results

Grid level	Convergence factor, *tol*	Timestep scheme	At reversal, t=0.1 s		
			Central	Minimum	H_{00}
65×65	0.3	Fixed	7.83×10^{-8}	1.74×10^{-8}	-1.208×10^{-6}
65×65	0.3	Variable	7.90×10^{-8}	1.76×10^{-8}	-1.207×10^{-6}
65×65	0.03	Fixed	7.83×10^{-8}	1.74×10^{-8}	-1.208×10^{-6}
65×65	0.03	Variable	7.88×10^{-8}	1.75×10^{-8}	-1.207×10^{-6}
129×129	0.3	Fixed	7.59×10^{-8}	2.26×10^{-8}	-1.207×10^{-6}
129×129	0.3	Variable	7.61×10^{-8}	2.26×10^{-8}	-1.207×10^{-6}
			At t=t_{min}		
			Central	Minimum	H_{00}
65×65	0.3	Fixed	8.19×10^{-8}	1.25×10^{-8}	-1.205×10^{-6}
65×65	0.3	Variable	8.26×10^{-8}	1.27×10^{-8}	-1.205×10^{-6}
65×65	0.03	Fixed	8.16×10^{-8}	1.25×10^{-8}	-1.206×10^{-6}
65×65	0.03	Variable	8.21×10^{-8}	1.28×10^{-8}	-1.206×10^{-6}
129×129	0.3	Fixed	8.14×10^{-8}	1.77×10^{-8}	-1.203×10^{-6}
129×129	0.3	Variable	8.17×10^{-8}	1.78×10^{-8}	-1.202×10^{-6}

Table 5.6: Comparison between test parameters at reversal and t_{min}

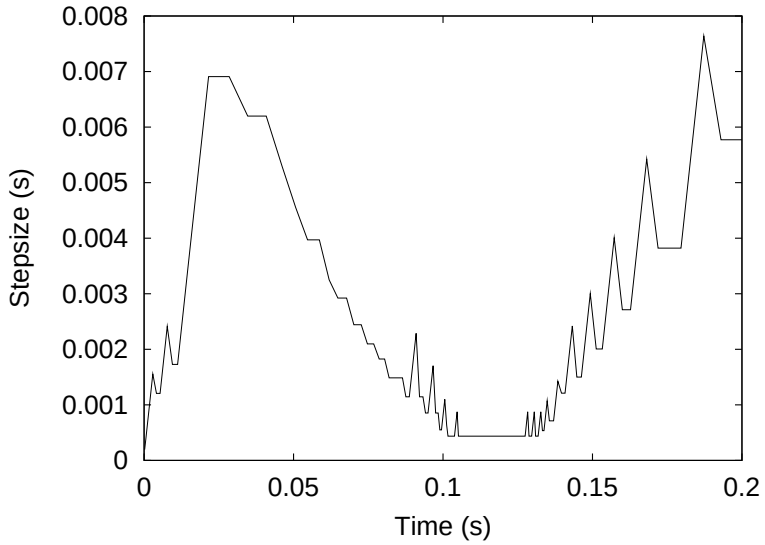


Figure 5.34: Timestep sizes during reversal using variable timestepping on grid 5, tol=0.3

ness, H_{00}, at two reference times. These have been chosen to be in the area of greatest importance and are the values at the point of reversal, t=0.1 s, and at t=t_{min}, the time of minimum calculated film thickness, around 0.11 s. At these times the minimum film thickness is an order of magnitude less than the initial steady state. It can be seen that there is excellent agreement between the results of fixed and variable timestepping cases.

The actual timestep sizes can be seen in Figures 5.34 to 5.36. These show how the timestep rises from the initial value of ΔT=ΔX to reach a maximum value before falling to reach a constant value just after reversal. This minimum ΔT is still larger than the initial stepsize in all three cases.

The values of the calculated stepsize change ratio, r, are shown in Figures 5.37 to 5.39. They show exactly how 'good' the current timestep size is considered to be. The flat period at reversal coincides with the smallest timesteps and the desire of the local error estimation to increase the stepsize afterwards is clearly visible.

It was explained above that there can be advantages in not changing the stepsize too often. Choosing the new stepsize based on an *a priori* error test cannot guarantee that the new stepsize will be valid for more than one timestep. Thus having a large enough range of values for r in Equation (5.49) where the stepsize remains unchanged is important. The size of this region also governs how often the stepsize can change. Setting it too narrow can result in the stepsize being successively increased and decreased. This 'chattering' effect, well known in the ODE community, may cause instabilities in the solution. This region is considered, for example, by Shampine [132] and Hairer *et al.* [59]. It is the range of values calculated for r in the error test, for which no change in stepsize should be

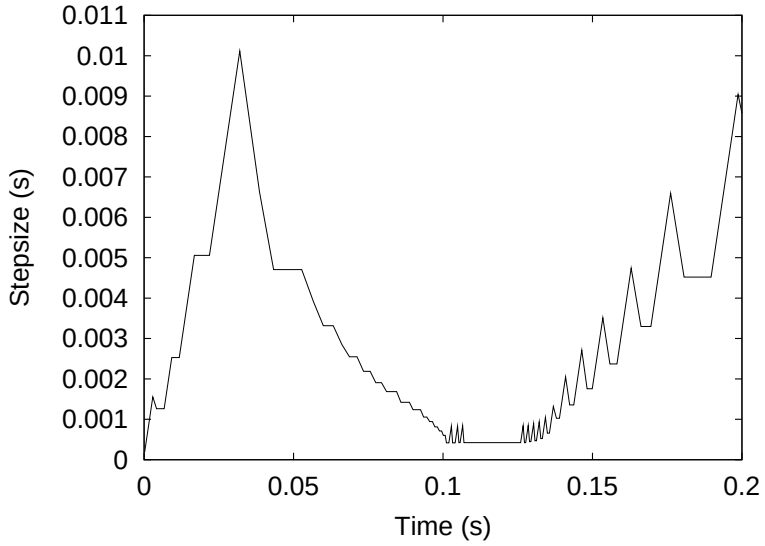


Figure 5.35: Timestep sizes during reversal using variable timestepping on grid 5, tol=0.03

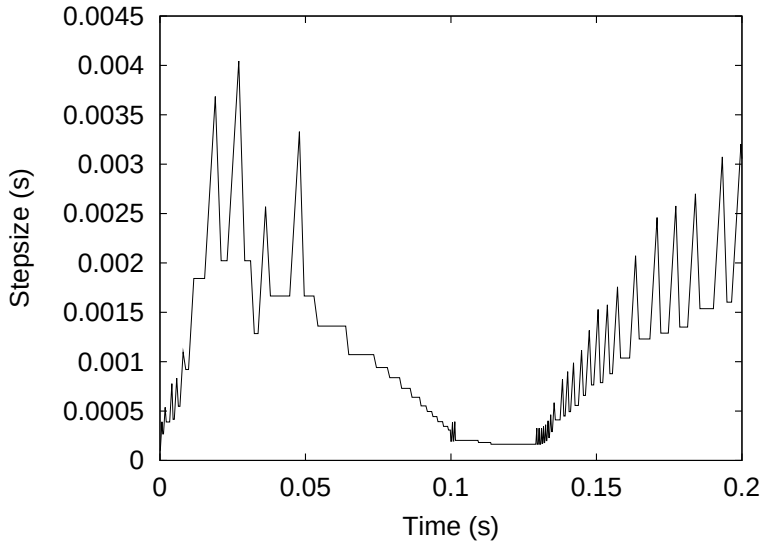


Figure 5.36: Timestep sizes during reversal using variable timestepping on grid 6, tol=0.3

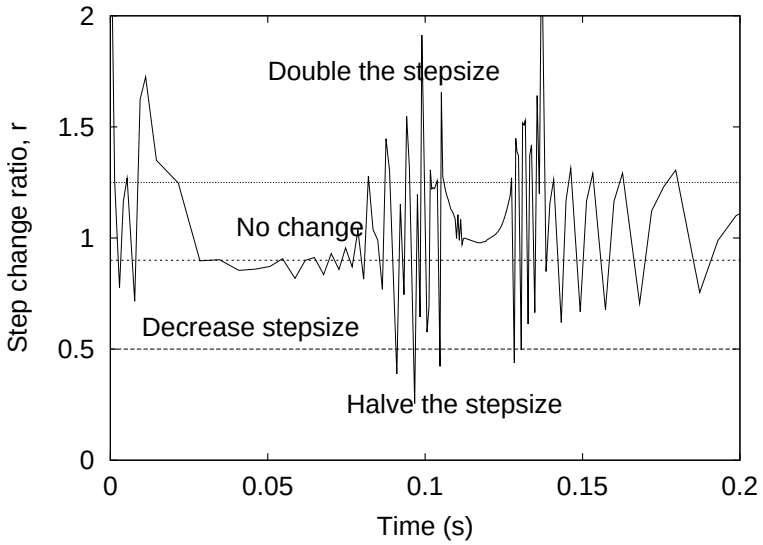


Figure 5.37: Stepsize change ratio, *r*, during reversal using variable timestepping on grid 5, tol=0.3

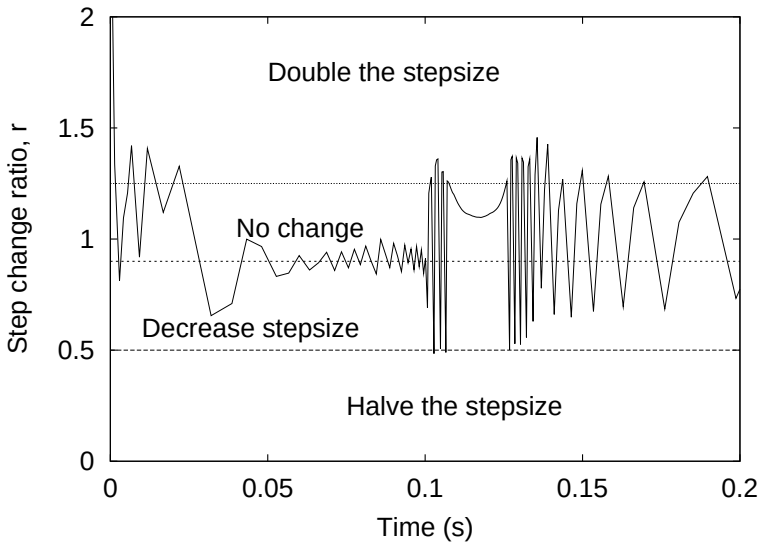


Figure 5.38: Stepsize change ratio, *r*, during reversal using variable timestepping on grid 5, tol=0.3

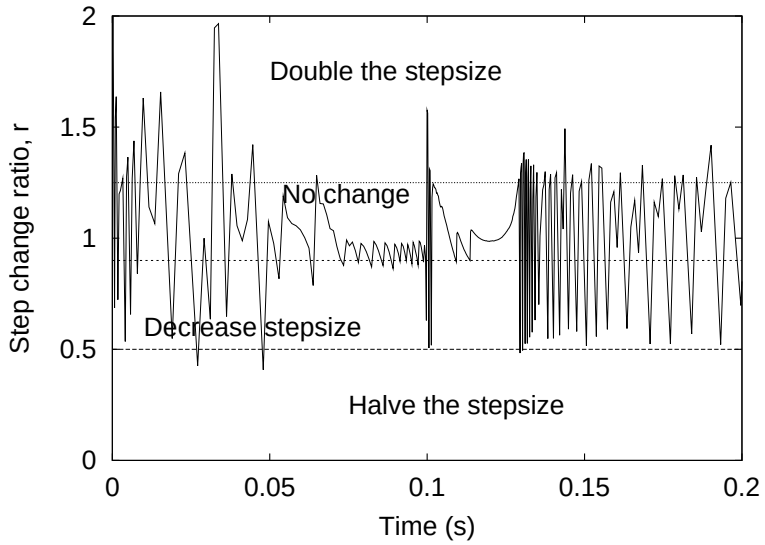


Figure 5.39: Stepsize change ratio, r, during reversal using variable timestepping on grid 6, tol=0.3

Grid Dimension	Convergence tolerance	Fixed / Variable	Time taken (s)	Iterations required
65×65	0.3	Fixed	1493	2164
65×65	0.3	Variable	1116	1126
65×65	0.03	Fixed	2998	4356
65×65	0.03	Variable	2440	3019
129×129	0.3	Fixed	11090	4181
129×129	0.3	Variable	5374	951

Table 5.7: Computational performance comparison between fixed and variable timestepping codes

made. An example of reduction in chatter is shown by comparing Figures 5.39 and 5.40 where the level at which r increases the timestep value has been increased from 1.25 to 1.5.

The use of variable timestepping may require more iterations to reach the same level of approximation for the solution at individual timesteps. However, the important factor is not that more cycles may be needed per timestep, but that over a complete run less are taken. It is also possible to limit the number of iterations per timestep if the convergence of individual steps is failing to satisfy the convergence test of Equation (5.26), quickly enough. This may occur at solution discontinuities.

The balance between variable timestepping taking extra iterations per timestep, and fixed timestepping taking more timesteps overall is quantified in Table 5.7 where the six cases above are considered. Both the computational times and the iterations required,

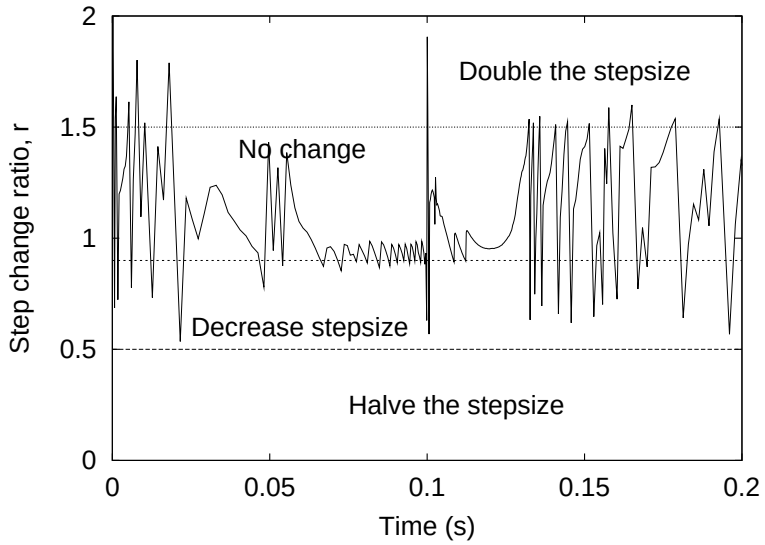


Figure 5.40: Stepsize change ratio, r, during reversal using variable timestepping on grid 6, tol=0.3, with a larger threshold for increasing timestep size

excluding those for evaluating the local temporal error, are compared. It is seen that variable timestepping is especially beneficial on the finer grid where the computational time is reduced by over 50% and the total number of iterations by a factor of four.

Overall, it is clear that the use of variable timestepping can produce significant savings in computational work whilst producing results of similar accuracy.

5.7.3.2 Ridge Tracking

Variable timesteps are of great value for distinguishing the change in solutions and capturing features of particular importance. A good example of this is the overrolling of a transverse ridge, considered by Venner and Lubrecht [143] and previously here in Section 5.6.4.1. The parameters are, again, as given in Table 5.4 with the ridge entering from the left hand side of the domain, with $X_d(0)$=-2.5, and then progressing through the domain. Initially the ridge is outside the area of influence on the solution, and the timestep is expected to be larger than the optimal timestep as the ridge passes through the contact region. Once the ridge has entered the cavitation region no further transient effects should be present, and hence the solutions should return to the steady state conditions, and the timestep size should increase.

The physical solutions obtained were as shown in Section 5.6.4.1 and hence the only part of interest is the timestep information. In Figure 5.41 the timestep size is shown. As expected it does rise to a maximum which then falls as the ridge enters the contact area at T=1.5. This remains fairly constant until the ridge enters the cavitation region just after

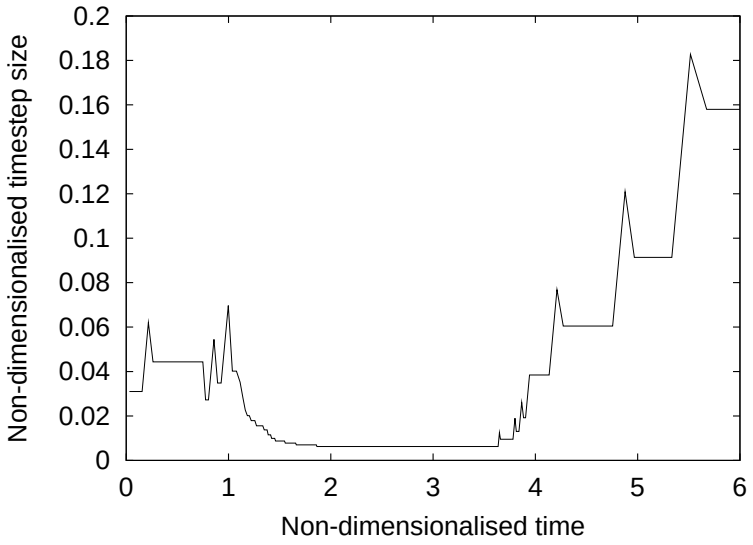


Figure 5.41: Timestep sizes during overrolling of a transverse ridge using variable timestepping on grid 6, tol=0.3

T=3.5. Once out of the contact the solution returns to steady state conditions and the timestep size increases dramatically.

Consideration of the stepsize change parameter, r, in Figure 5.42 reveals how satisfactory the selected timestep size is, as the ridge passes through the contact area. The desire to increase the stepsize both initially and after overrolling is contrasted against the reduction in timestep size between T=1 and T=2.

5.7.3.3 Shock Loading

The example of shock loading models the kind of change to the physical conditions experienced by the teeth on gears during meshing. Two kinds will be demonstrated in the section: shock loading and shock unloading. Rather than the variable loading examples shown in Section 5.6.3, the load changes are of much greater amplitude, and the timescale is very short.

The examples to be considered here both start from the initial conditions of the reversal example, shown in Tables 5.1 and 5.2. The loading will take place by increasing the maximum Hertzian pressure to double its initial value for loading, or by halving it for the unloading case. This is incorporated into the solve by modifying the target sum of pressures in Equation (5.5) to be

$$\frac{2\pi}{3}\Phi = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} P(X,Y)\,\mathrm{d}X\,\mathrm{d}Y, \tag{5.51}$$

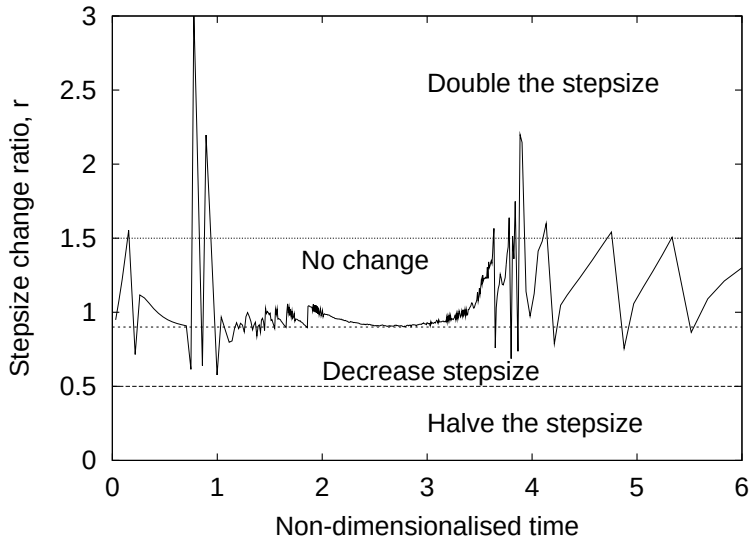


Figure 5.42: Stepsize change ratio, r, during overrolling of a transverse ridge using variable timestepping on grid 5, tol=0.3

where the variable load, Φ, is given by

$$\Phi = \left\{ \begin{array}{ll} 1+450t & t \leq 0.02\text{s} \\ 10 & t > 0.02\text{s} \end{array} \right\} \quad \text{for increasing loading} \tag{5.52}$$

$$\Phi = \left\{ \begin{array}{ll} 1-45t & t \leq 0.02\text{s} \\ \frac{1}{10} & t > 0.02\text{s} \end{array} \right\} \quad \text{for decreasing loading.} \tag{5.53}$$

Considering first the physical results. The central and minimum film thickness results are as shown in Figures 5.43 and 5.44 for increasing and decreasing the load, respectively. Of particular interest is the initial unexpected behaviour as the loading changes: for the shock loading case the film thickness rises across the contact before attaining its new significantly lower value, whereas for the unloading the reverse behaviour is less evident, although can be observed in the central value.

Pressure solutions for the two cases are very different. The centreline pressure is shown at selected timesteps in Figure 5.45, with the arrow indicating the direction of increasing time of solution plots. The increase in both the size of the contact area and maximum pressure is very evident.

The unloading example has a very different behaviour, as shown in Figure 5.46, because the problem has dropped out of the EHL regime to just being hydrodynamically lubricated instead. This means that the pressure spike has completely disappeared, and the deformation is now minimal, with no sidelobes present at all.

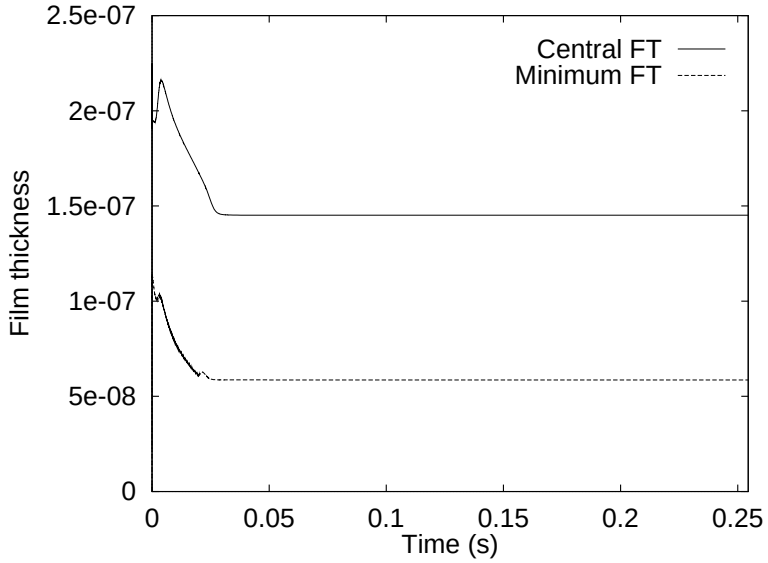


Figure 5.43: Central and minimum film thickness during shock loading from 0.45 to 0.90 GPa in 0.02 s

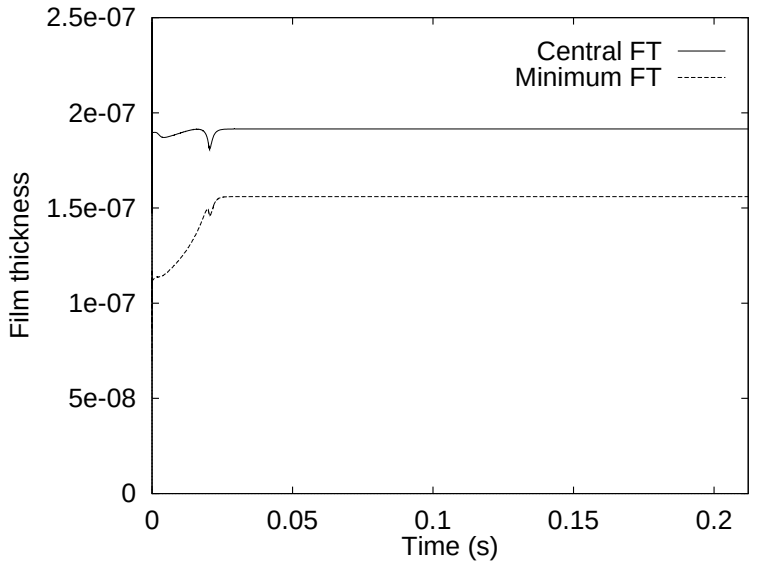


Figure 5.44: Central and minimum film thickness during shock unloading from 0.45 to 0.22 GPa in 0.02 s

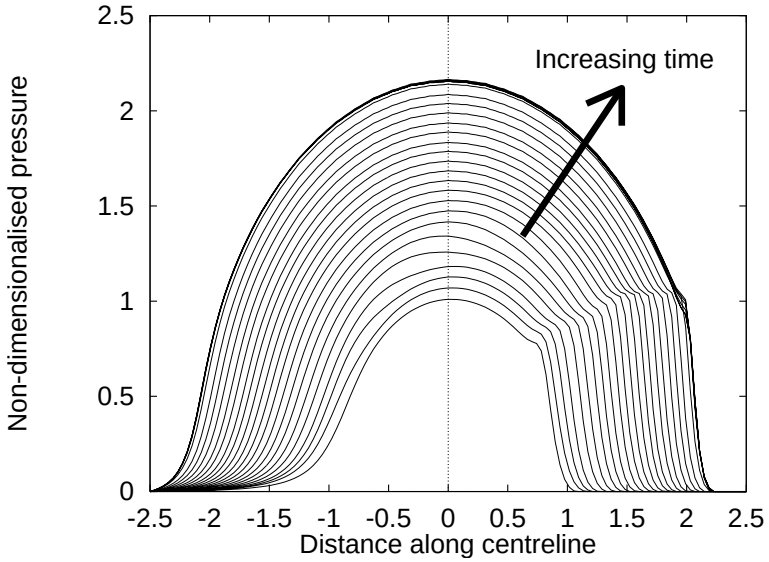


Figure 5.45: Centreline non-dimensional pressure during shock loading from 0.45 to 0.90 GPa in 0.02 s

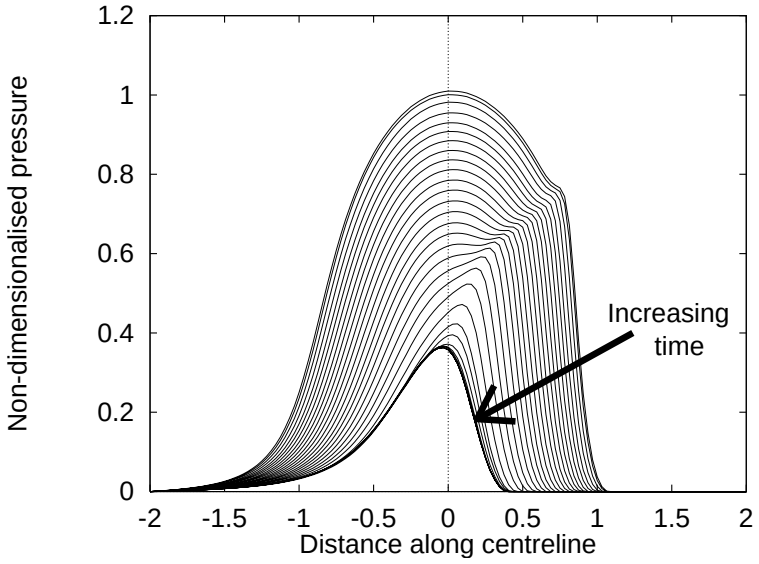


Figure 5.46: Centreline non-dimensional pressure during shock unloading from 0.45 to 0.22 GPa in 0.02 s

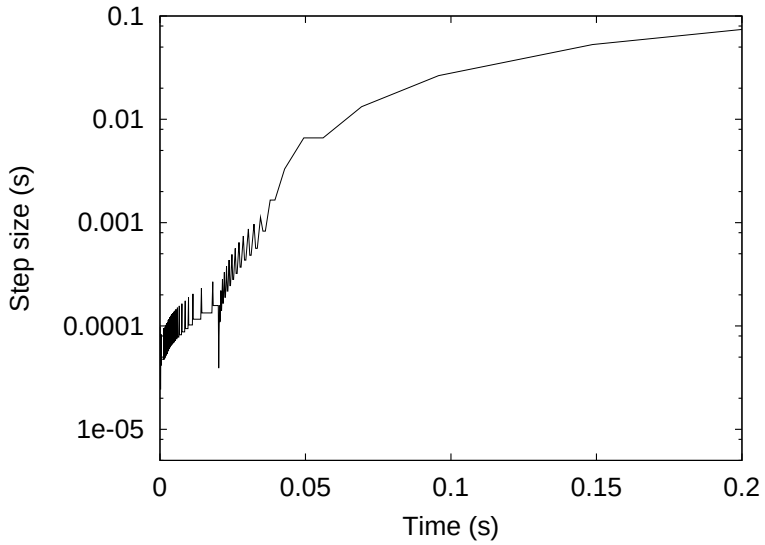


Figure 5.47: Timestep sizes selected during shock loading from 0.45 to 0.22 GPa in 0.02 s

The timestep size should reflect the amount of solution change, hence beyond 0.02 s when the system has reached the final steady state solution the timestep should be much larger than in the initial stages when the greatest change is taking place. For the shock loading example the timestep size is shown in Figure 5.47 with the stepsize change ratio, r, shown in Figure 5.48. It can be clearly seen that the timestep does initially start out much smaller than the final steady state value.

The timestep sizes and the step change ratios for the shock unloading examples are shown in Figures 5.49 and 5.50. Similar behaviour to the shock loading case can be observed, although with a significantly less smooth increase in the timestep size. It is also noticeable that the position of the minimum of the dip in film thickness at the end of unloading, shown in Figure 5.44, coincides with the minimum timestep size selected.

5.8 Conclusion

In this chapter it has been shown that the solution methods used in Chapter 4 can be developed to solve transient circular point contact EHL problems. The numerical methods employed make use of the differential-algebraic nature of the equations to be solved. These include the use of the Shampine convergence test, for individual timesteps, to estimate when convergence has been obtained based on a predicted solution. The choice of which variable should be monitored for error tests was also discussed.

Several example solutions were presented for very different physical cases. The pseudo-steady state case was presented to demonstrate the overall stability of the solver.

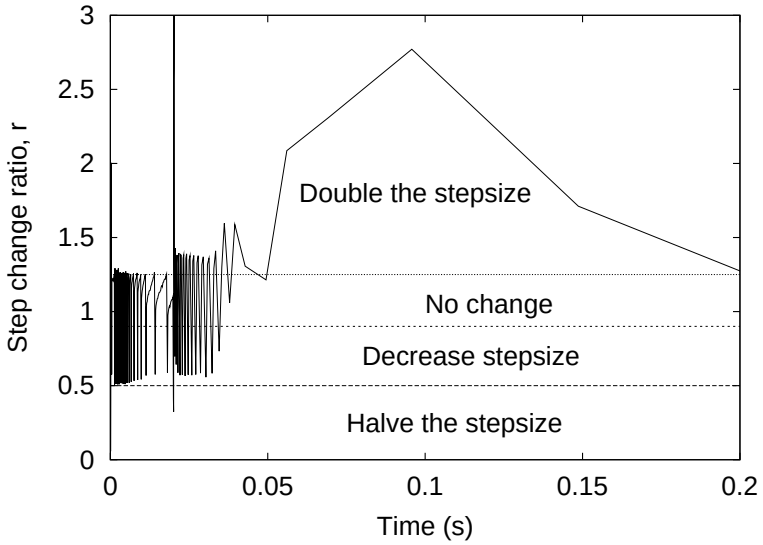


Figure 5.48: Stepsize change ratio, *r*, values during shock loading from 0.45 to 0.22 GPa in 0.02 s

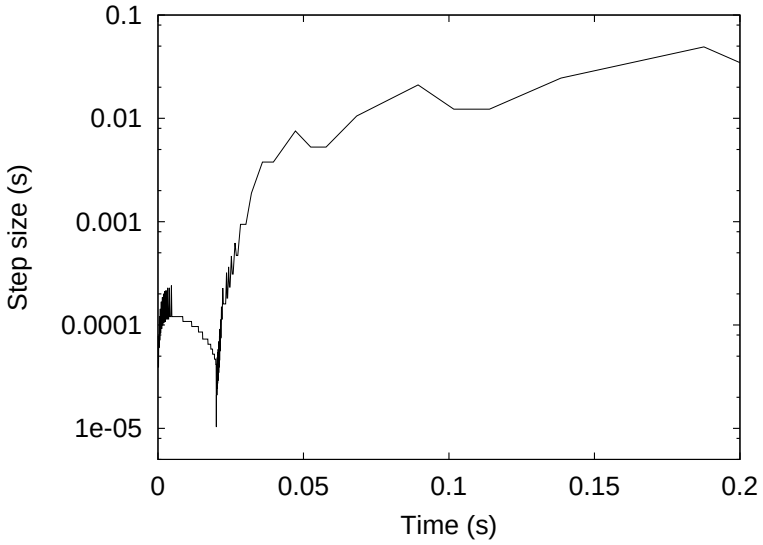


Figure 5.49: Timestep sizes selected during shock unloading from 0.45 to 0.22 GPa in 0.02 s

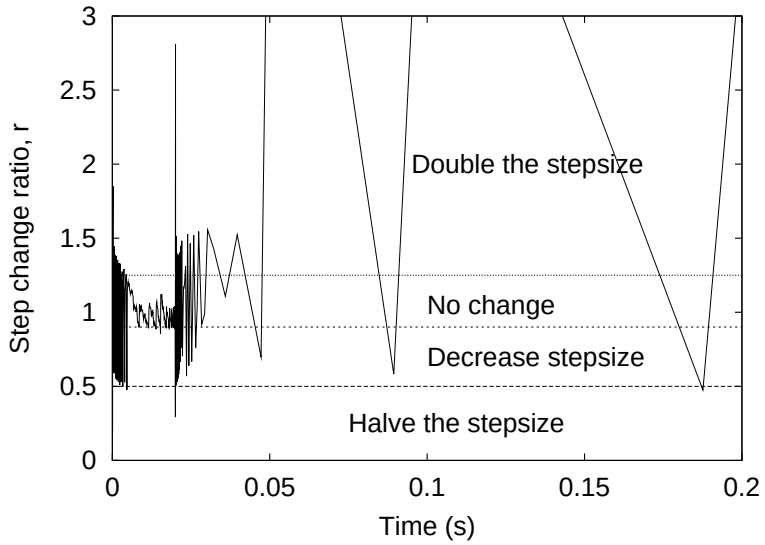


Figure 5.50: Stepsize change ratio, r, values during shock unloading from 0.45 to 0.22 GPa in 0.02 s

The example of reversal of entrainment was presented. Results were successfully compared to both the experimental and numerical results of Scales *et al.* [126, 128]. It was shown how the use of prediction for the initial solution on a timestep significantly reduces the residual level and hence the amount of work required to obtain a solution on that step. It was also seen how the accuracy of the solutions obtained relate to the tolerances chosen for the Shampine convergence test.

Another case examined was that of sinusoidally varying loads. The results were compared to those of Wijnant [154] and again a very close similarity was observed. The frequency of the oscillations was also decreased, well beyond those considered by Wijnant, and it was seen how the non-linear effects due to the squeeze term were no longer visible.

Problems involving surface features are very important in the continuing drive to model increasingly realistic contacts. In this chapter an example, for a single timestep, of the pressure and film thickness results was shown for a sputtering pattern applied to one of the contacts. Two other cases were solved transiently, with significant effects from the squeeze term of the Reynolds Equation. The case of overrolling of a transverse ridge was successfully compared to the results of Venner and Lubrecht [143]. The second example was a variation on the case of a circular dent passing through the domain, considered by Ai and Cheng [1].

To improve the computational efficiency of multigrid finite difference EHL codes, a method of variable timestepping has been presented based on the differential algebraic

nature of the system. Variable timestepping has been shown, by experiments, to substantially reduce the required work whilst maintaining the same level of solution accuracy. The overhead in calculating new stepsizes is small, relative to the increase in performance. Changing the stepsize away from ΔT, within predefined limits, has been seen to pose no problems for the solver.

Examples considered were the case of reversal, the overrolling of a transverse ridge, and the case of shock loading and unloading of the contact. These examples all showed that the convergence tests did identify regions of greatest non-linear solution change by reducing the timestep in these periods, and later increasing it again once the linearity of the solution had been re-established. In the example of shock unloading it was also seen how the code was able to still function without difficulties despite dropping out of the EHL regime into the hydrodynamic lubrication region.

Chapter 6

Grid Adaptation

6.1 Introduction

The drive towards solutions of problems incorporating surface features and, ultimately, true roughness means that the solvers being used must be designed to be extensible to meshes of several thousand points in each dimension. However, increasing the resolution of the domain means that the amount of work required to obtain solutions also grows. The need to obtain solutions sufficiently quickly means that new techniques must be investigated.

Grid adaptation is widely used in many areas of computing for obtaining fast, accurate solutions to many numerical problems. This is especially true for finite element methods. Error estimates can be derived in order to decide the areas in need of refinement or de-refinement [3, 148]. The style of adaptation chosen will depend upon the nature of the problem. For example, it is possible to track a shock moving through a domain by keeping a fine mesh around the front, through the use of node movement, called *r-refinement*, with the number of mesh points remaining fixed [8]. Conversely, *h-refinement* is the addition of more mesh points in regions of interest [135, 148]. Finally *p-refinement* changes the order of the approximating polynomial solution used with the higher order polynomials giving a local increase in accuracy [6].

The use of adaptive meshing in EHL, though, is relatively unknown. Early work was done by Lubrecht and co-workers [93, 99] into the use of adaptive grids, however this

appears to have been an isolated experiment. In [50] Goodyer *et al.* presented preliminary results for the use of adaptive meshing and considerable work has been done since that time to improve upon the methods used in terms of both accuracy and performance.

In this chapter the ideas behind re-meshing are to be explored. It will be shown that without major changes to the solution method, grid adaptation can maintain the accuracy of the results whilst obtaining significant improvements in the performance of the code.

The numerical techniques to be used for the mesh adaptation are discussed in Section 6.2. The results of [50] were predominantly based on refinement of the film thickness solve, however with the large advances made in the multilevel multi-integration algorithm used, the deformation calculation has become less important to the overall solution time.

Choice of where to adapt is very important. There are three possible methods for choosing this, and these will be considered in Section 6.3. The first is by apportioning parts of the domain for certain levels of refinement based on problem specific knowledge. The second is based on solution properties, be they actual solution values, derivatives or some other type of monitor function. The third is by applying some sort of error test.

Results will be shown in Section 6.4 for various different methods of adaptation for a steady state case. They will be presented in terms of both solution accuracy and computational time. Similarly, in Section 6.5 a transient case will be examined.

6.2 Theory

The addition of more fine grid points means that the resolution of the solution can be increased. However, it may not be necessary to use a fine grid in regions where the solution does not change greatly. The intention of adaptive meshing is to focus the computational work by placing mesh points in the areas of the domain where they are most required.

In cases where solution discontinuities or very steep gradients exist, the solution at these points must be updated differently from smooth parts of the solution, otherwise smearing of the numerical discontinuity will not allow accurate resolution of physical features. These ideas were used by Harten to produce a multilevel approximation strategy, called multiresolution (see [15] for 2D work). In Harten's work, in order to obtain the solution over the entire fine mesh, the only points stored are those required to interpolate the calculated solution using defined smoothness properties. This reduces both the computational work required in evaluating the new solution, and the storage of said solution.

It is straightforward to extend Harten's work to multigrid. Multigrid grid adaptation has been undertaken for many years, e.g. [7], including some early work with EHL

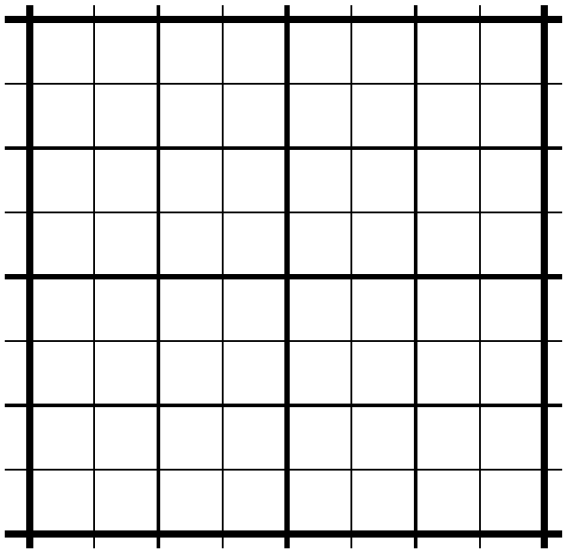


Figure 6.1: Example of a three level multigrid mesh

solvers [93, 99]. Goodyer *et al.* [50] applied the idea of smoothness to the calculation of the film thickness deformation calculation to generate a full fine grid representation of the solution. However, since the publication of [50], the increased efficiency of the multilevel multi-integration algorithm used has reduced the potential performance increase due to grid adaptation in the film thickness calculation.

With increasing numbers of points in the domain the linear algebra system required to be solved for each line gets larger. This can then dominate the calculation time. The method being applied in this chapter is to only use fine grids in selected regions of the domain. This idea is explained in detail in Trottenberg *et al.* [139] and is summarised below.

Considering a three level multigrid formulation as shown in Figure 6.1 then an extra level of refinement may be added over only part of the domain if so desired. This can be seen in Figure 6.2 where the bottom right hand of the four coarsest squares has been refined. It is possible to just solve for a new finest grid solution using just the points inside the shaded area. The points on the boundary with the unadapted region, namely those marked $\bullet$ and $\circ$ will be treated as Dirichlet boundary conditions. Those outside the shaded region will not be included in the solve. The difference between the two types of boundary points is the accuracy with which they have been obtained. Those marked $\bullet$ are a direct prolongation of the coincident coarser grid point, whilst the hanging nodes, marked by $\circ$, are only obtained by some interpolation of surrounding coarse grid points.

This method fits inside the multigrid framework as follows. Solutions from the finest grid are only generated inside the shaded region, say Ω_{ad}. It is clearly only valid to

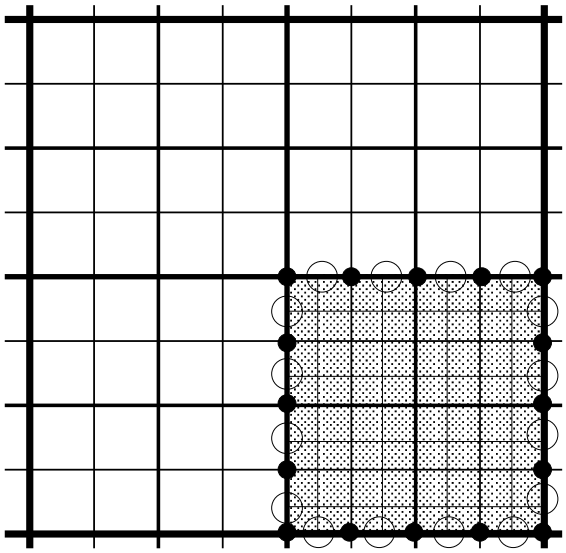


Figure 6.2: Example of a multigrid mesh with one quadrant refined further

calculate coarse grid corrections inside this region. Similarly the calculation of the right hand side functions, previously shown in Equation (3.10), now becomes

$$\hat{f}^j = \begin{cases} \mathscr{L}^j(I_k^j \underline{\tilde{u}}^k) + I_k^j r^k & (i,j) \in \Omega_{ad} \\ 0 & \text{elsewhere} \end{cases} \tag{6.1}$$

with all symbols as defined in Chapter 3. This means that only the points marked $\diamond$ in Figure 6.3 will have knowledge of finer multigrid meshes.

This procedure may be applied iteratively to produce a hierarchy of increasingly refined multigrid meshes. An example of this is shown in Figure 6.4 which shows five different grid levels with different adaptation domains on each. It is however important to ensure that refinement can only take place on an area which was refined on the next coarsest level.

The solution of the film thickness equation will still be performed on the entire unrefined fine mesh. A prolonged version of the pressure on the next coarsest grid will be used to generate the pressure solution on points not in the new adaptive mesh. Updating the film thickness at points which are not used may seem like an unnecessary overhead, however the use of multi-integration, in its present form, means that this overhead is not large. In future work it should be possible to include the adaptive mesh multilevel multi-integration techniques described by Brandt and Venner [18] to increase efficiency still further.

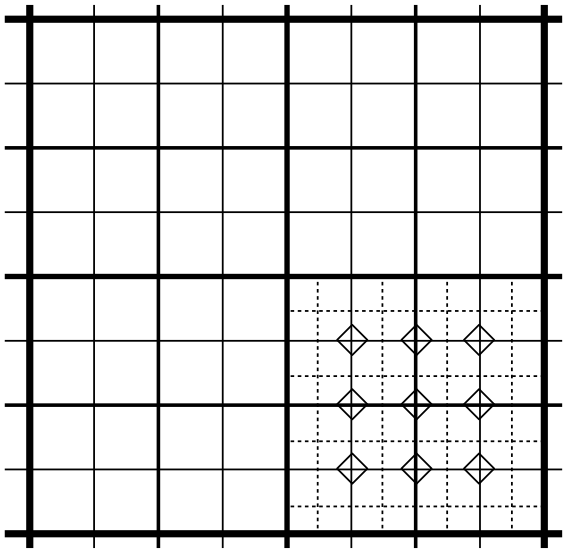


Figure 6.3: Example of the next coarsest multigrid mesh to that shown in Figure 6.2

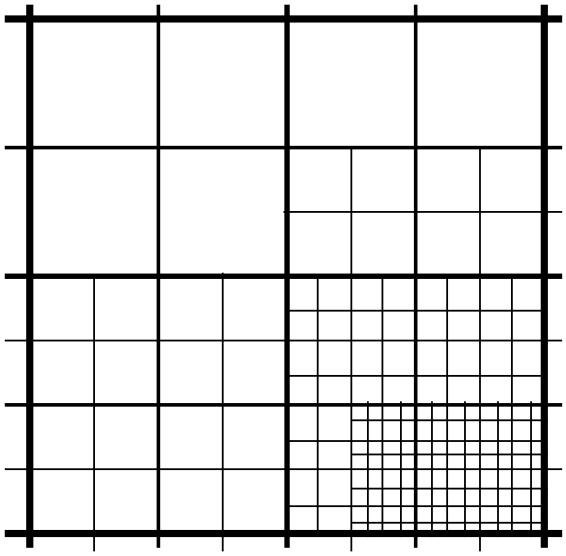


Figure 6.4: Example of many levels of adaptation applied to a hierarchy of grids

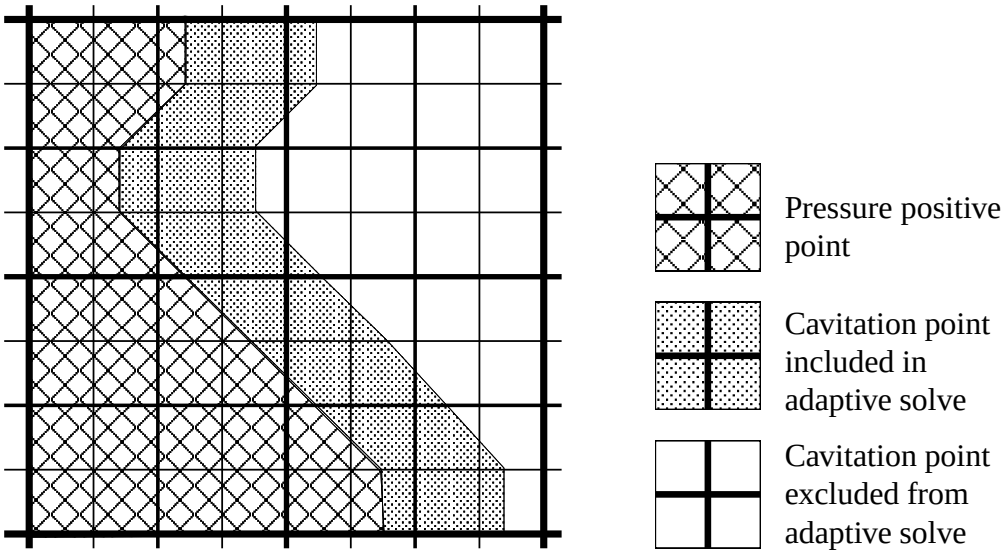


Figure 6.5: Example showing adaptation around cavitated free boundary

6.3 Monitoring Where to Adapt

Solutions of the EHL system are characterised by three regions of the domain: the contact region, where the pressure is high; the non-contact region, where the pressure is low; and the cavitation region, where the pressure is assumed to be identically zero. In deciding where to adapt the differences between these regions will be of great importance.

The most obvious choice of region in which to adapt the solution is the cavitation region. Here, there is no reason to spend CPU cycles calculating a pressure solution which will come out negative and then have to be set to zero. By deciding that only those points of positive pressure are in the linear solve automatically reduces the size of every calculation. There must, however, be at least one, preferably two points in the cavitation region included to allow the free boundary to adjust its position. An example of adaptation over the free boundary is shown in Figure 6.5 where the cross-hatched region indicates the area of positive pressure, and the dotted region includes the extra points included into each line solve.

In the rest of the domain there is a lot more freedom about how to choose where to adapt. Deciding the correct refinement criteria is very important to obtaining optimal performance from the use of grid adaptation, however even some crude assumptions will be shown to be very beneficial. There are three possible types of scheme available for deciding on refinement levels; namely arbitrary geometrical decisions, adaptation based on a monitor function, or refinement based on an error test.

Effective use of the arbitrary geometrical decision scheme relies on *a priori* knowl-

edge of how the solution behaves. In the EHL case this is well defined, since it is known that the high pressure contact region is found inside the unit disc centred at (0,0), that the cavitation region is almost entirely the X-positive region not in the contact region, and the non-contact region is the rest of the X-negative domain. It is therefore sensible to surmise that refinement in the area of high pressure and great deformation is advisable. Using this method of adaptation, the decisions on where to adapt are all taken before the code is started.

The adaptation can also be made more automatic. This is often done through use of a *monitor function* which will take account of the properties of the solution before deciding which regions need adapting. This idea is already commonly used in EHL problems to decide which numerical scheme should be used at individual points, as shown in Section 4.2.1. The monitor function may be as simple as "On grid 6 adapt if $P_{i,j}<1\times10^{-3}$". Equally, it may be chosen to monitor derivatives of solution variables to identify regions of greatest change. This method again requires certain criteria to be decided before the run is started, but the actual regions adapted will move as the solution proceeds.

Fully automatic refinement can only come through use of an accurate error control function. Much work has been done into grid refinement in many different applications. It is, however, the case that at present little or no analysis has been done into the error control issues of the EHL problem. In the work of Lubrecht [93] the method of Bai and Brandt [7] was used to decide where to adapt. This uses the quantity τ_k^{k-1} known as the *(k,k-1) relative truncation error* which is the quantity that is to be added to the right hand side of the coarse grid problem, and is thus a measure of the extent to which the local introduction of the finer grid has influenced the global solution [139]. The convergence test given in [139] is to refine whenever

$$(\Delta X)^k \tau_k^{k-1} < \xi \tag{6.2}$$

for a chosen tolerance ξ.

In Section 6.4 a steady state example will be solved using each of these three methods, and their relative effectiveness and applicability judged.

6.4 Adaptation Example - Steady State

To test the effectiveness of the grid adaptation, the test case shown in Table 6.1 was chosen. This was solved on the half-domain $X\in$[-4.5,1.5], $Y\in$[0,3] using finest meshes between level 5 and level 8. The accuracy of the solutions were monitored to ensure that

Parameter	Test Case 3
Viscosity index, α	2.2×10^{-8} Pa^{-1}
Viscosity at ambient pressure, η_0	0.04 Pa s
Maximum Hertzian pressure, p_h	0.45 G Pa
Material parameter, G	4972
Load parameter, W	1.63×10^{-7}
Speed parameter, U	8.18×10^{-12}
Moes parameter, M	20
Moes parameter, L	10

Table 6.1: Non-dimensional parameters for steady state adaptation example

potential speed-ups were not detrimental to the quality of the solution obtained.

Measuring the accuracy of the results can be done in several different ways. Comparisons can be made between the values of the different 'notable' variables, for example the central and minimum film thickness and the height of the pressure spike. The accuracy desired is that the solution of an adapted grid on level *k* should be significantly closer to the results obtained on an unadapted level *k* than those obtained on an unadapted level *k*-1 mesh, or coarser.

Comparison between grid levels does, however, rely upon the numerical solution to the equations being similar. Around the pressure spike this need not be the case. The effect of grid spacing for line contact cases has been investigated for many years, and results up to level 10 were shown using adaptation in the work of Lubrecht [93] and Breukink [21]. Recent work done by Fairlie [44] into the resolution of the pressure spike in the line contact case is reproduced here in Figures 6.6 to 6.8. In Figure 6.6 the pressure distribution across the whole domain is shown. It can the seen the curves are almost coincidental apart from around the pressure spike. This area is shown in detail in Figure 6.7 where the addition of many orders of magnitude more points has now captured the pressure spike completely and appears to have achieved a converged solution. Whilst the accurate capturing of such a feature is an obvious use for adaptation, it can be seen from Figure 6.8 that even the small differences in spike height have caused much greater variation in the film thickness profile.

The accuracy of the whole solution domain can also be compared, although the limitations mentioned above must be remembered. The points at which comparisons are to be taken is again important to the level of accuracy obtained. If all points on the finest level are to be considered then it is the case that the results at all points outside the adapted region have not had a fine grid solution solved on them. There would be the errors in the accuracy of the interpolation routines used to generate the non-coincident intermediary

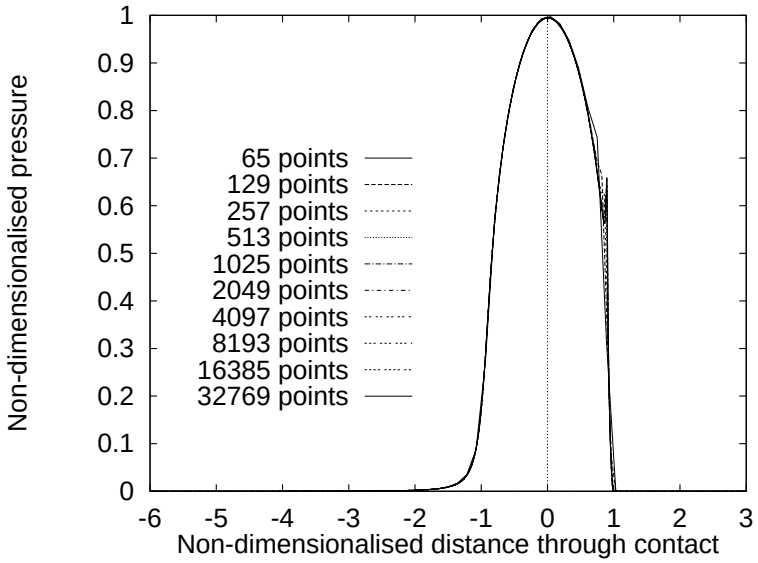


Figure 6.6: Non-dimensional pressure plot of a line contact problem on levels 5 to 14 [Fairlie]

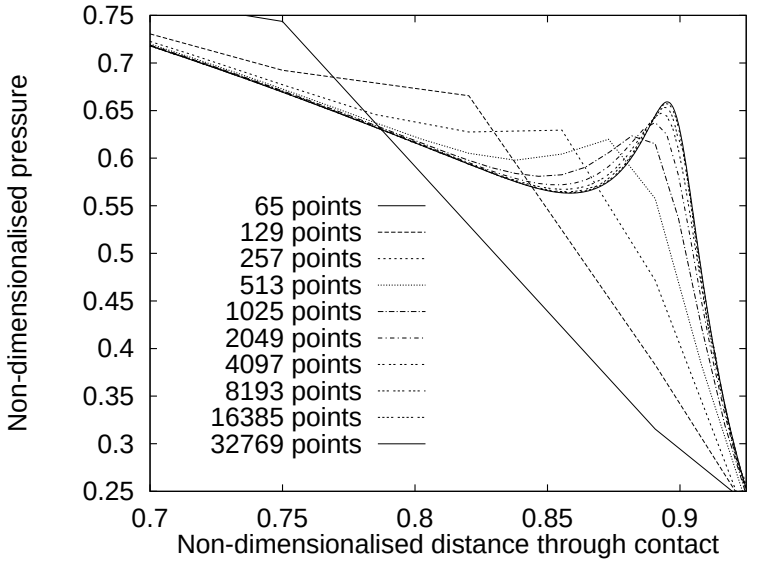


Figure 6.7: Non-dimensional pressure plot around spike in a line contact problem on levels 5 to 14 [Fairlie]

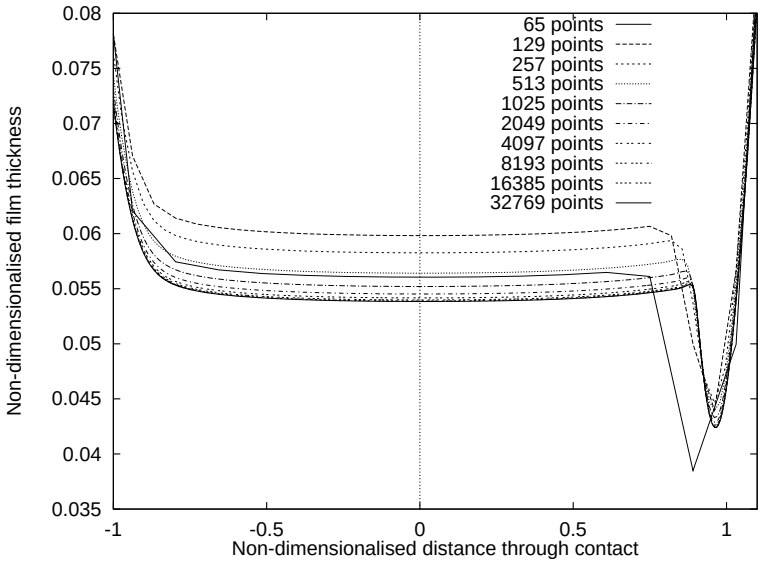


Figure 6.8: Non-dimensional film thickness plot of a line contact problem on levels 5 to 14 [Fairlie]

points, coupled with the lack of relaxations on all non-adapted points. It therefore seems sensible to only calculate the accuracy at points which were solved upon at their finest level, hence accuracy will be compared on a mesh similar to the one shown in Figure 6.4, which itself was never explicitly used in any calculation.

The method chosen for comparing the accuracies of the methods is by calculating the L_2-norm of the differences between solutions on adapted and unadapted grids. This is given by

$$\left\| u_{ref} - u_{ad} \right\|_2 = \int_{Y_{min}}^{Y_{max}} \int_{X_{min}}^{X_{max}} \left(u_{ref} - u_{ad} \right)^2 \mathrm{d}X \mathrm{d}Y \tag{6.3}$$

where u_{ref} is the reference solution against which the adapted solution, u_{ad}, is compared. The integrals have been calculated using three point Gauss quadrature applied in both X and Y directions. To solve the integral on each square pseudo-element of the mesh, a substitution is undertaken to calculate on points $\zeta \in$[-1,1], before using the approximation

$$\int_{-1}^{1} \Psi(\zeta) \mathrm{d}\zeta \approx \sum_{k=1}^{3} a_k \Psi(\zeta_k) \tag{6.4}$$

with abscissas, ζ_k, and weights, a_k as given in Table 6.2, with the solution values, Ψ being given by linear interpolation of the mesh point solution values. By applying this equation in both X and Y directions, a nine-point formula defines the numerical integral. For adapted grids, with prolongation of the solution as defined by the linear interpolation of Equation (3.15), this quadrature will preserve the validity of norm, regardless of the

k	Abscissa	Weight
1	-0.774597	5/9
2	0.0	8/9
3	0.774597	5/9

Table 6.2: Abscissa and weights for calculating three point Gauss quadrature on a domain $\zeta \in$[-1,1].

number of further sub-divisions of this element that are made.

The grid adaptation schemes used in each of the tests are shown in Table 6.3, where each section refers to the finest grid used in the multigrid cycle. For each grid there are examples of each of the three types of adaptation described in Section 6.3. In addition to retaining consistency between the tests on the same finest grid level, there is a strong element of correlation between different levels, e.g. cases 6.3 and 6.4 with cases 7.4 and 7.5.

The errors, as described above, are measured by two different methods. First, the values for non-dimensional central and minimum film thickness, and centreline non-dimensional pressure spike height for each test case are shown in Table 6.4. By considering each of the three adaptation methods in turn, a good intuition to the veracity of the results may be gauged. The geometrical adaptation schemes, shown first for each grid level, all preserve the three test parameters in the neighbourhood of the fully fine results for that grid level, and all are a significant distance from those of other grid levels. Similar results are also obtained for the pressure dependent grid adaptation schemes. However, for the automatic grid refinement schemes the results are not as close for larger values of ξ in Equation (6.2) as the grid gets finer. Whilst on its own the trade off between performance and accuracy is expected to reduce the quality of results, the final entries for each grid level exhibit starkly different solutions.

On a related note, in the test descriptions of Table 6.3 it was noted that the pressure dependent style of test would find the first pressure point above the chosen level and then refine all the way to the cavitation boundary. If the refinement had only been done in regions with pressure greater than this level, and hence the area between the contact region and the cavitation region had not been refined, then the results for the less stringent tests given in Table 6.5 would be as given in Table 6.6. These show that beyond only the finest level being adapted there is a marked drop of accuracy. This is most likely to do with the narrowness region between contact region and cavitation region, and the effects of the multigrid moving the cavitation boundary on too coarse a grid level to be accurate for a fine grid solution.

The second set of accuracy results are those of the L_2-norm for both the pressure and film thickness results, again performed on the non-dimensional quantities. Comparisons

Grid 5

5.0	No adaptation, fully converged
5.1	Solution after 10 multigrid cycles

Grid 6

6.0	No adaptation, fully converged
6.1	Solution after 10 multigrid cycles
6.2	Including adaption of cavitation region on all levels
6.3	Cavitation, and grid 6 for $X < -2.44$, $Y > 1.50$ de-refined
6.4	As 6.3, with grid 5 for $X < -3.57$, $Y > 2.27$ de-refined
6.5	Refined along line from first $P > 1\times10^{-2}$ on grid 6
6.6	As 6.5, with grid 5 from $P > 2\times10^{-3}$
6.7	Refined along lines using Equation (6.2), with $\xi = 1\times10^{-6}$
6.8	As 6.7, with $\xi = 1\times10^{-5}$
6.9	As 6.7, with $\xi = 1\times10^{-4}$
6.10	As 6.7, with $\xi = 1\times10^{-3}$

Grid 7

7.0	No adaptation, fully converged
7.1	Solution after 10 multigrid cycles
7.2	Including adaption of cavitation region on all levels
7.3	Cavitation, and grid 7 for $X < -1.50$, $Y > 1.17$ de-refined
7.4	As 7.3, with grid 6 for $X < -2.44$, $Y > 1.50$ de-refined
7.5	As 7.4, with grid 5 for $X < -3.57$, $Y > 2.27$ de-refined
7.6	Refined along line from first $P > 5\times10^{-2}$ on grid 7
7.7	As 7.6, with grid 6 from $P > 1\times10^{-2}$
7.8	As 7.7, with grid 5 from $P > 2\times10^{-3}$
7.9	Refined along lines using Equation (6.2), with $\xi = 1\times10^{-7}$
7.10	As 7.9, with $\xi = 1\times10^{-6}$
7.11	As 7.9, with $\xi = 1\times10^{-5}$
7.12	As 7.9, with $\xi = 1\times10^{-4}$

Grid 8

8.0	No adaptation, fully converged
8.1	Solution after 10 multigrid cycles
8.2	Including adaption of cavitation region on all levels
8.3	Cavitation, and grid 8 for $X < -2.27$, $Y > 0.95$ de-refined
8.4	As 8.3, with grid 7 for $X < -3.00$, $Y > 1.50$ de-refined
8.5	As 8.4, with grid 6 for $X < -3.56$, $Y > 2.06$ de-refined
8.6	As 8.5, with grid 5 for $X < -4.13$, $Y > 2.44$ de-refined
8.7	Refined along line from first $P > 1\times10^{-1}$ on grid 8
8.8	As 8.7, with grid 7 from $P > 5\times10^{-2}$
8.9	As 8.8, with grid 6 from $P > 1\times10^{-2}$
8.10	As 8.9, with grid 5 from $P > 2\times10^{-3}$
8.11	Refined along lines using Equation (6.2), with $\xi = 1\times10^{-8}$
8.12	As 8.11, with $\xi = 1\times10^{-7}$
8.13	As 8.11, with $\xi = 1\times10^{-6}$
8.14	As 8.11, with $\xi = 1\times10^{-5}$

Table 6.3: Adaptation schemes used for the steady state grid adaptivity tests

Test	Central film thickness	Minimum film thickness	Centreline pressure spike height
5.0	0.4615	0.3083	1.1373
5.1	0.4616	0.3090	1.1410
6.0	0.4488	0.3026	1.1679
6.1	0.4490	0.3024	1.1737
6.2	0.4490	0.3025	1.1728
6.3	0.4491	0.3028	1.1719
6.4	0.4492	0.3028	1.1717
6.5	0.4490	0.3027	1.1675
6.6	0.4491	0.3028	1.1673
6.7	0.4489	0.3026	1.1683
6.8	0.4489	0.3026	1.1683
6.9	0.4506	0.3043	1.1661
6.10	0.3991	0.2096	1.0851
7.0	0.4411	0.2978	1.2435
7.1	0.4412	0.2968	1.2477
7.2	0.4414	0.2972	1.2464
7.3	0.4419	0.2977	1.2473
7.4	0.4421	0.2980	1.2469
7.5	0.4421	0.2980	1.2467
7.6	0.4418	0.2976	1.2464
7.7	0.4423	0.2979	1.2470
7.8	0.4423	0.2980	1.2470
7.9	0.4414	0.2975	1.2454
7.10	0.4424	0.2982	1.2477
7.11	0.4184	0.2893	1.1856
7.12	0.4184	0.2921	1.1231
8.0	0.4369	0.2954	1.3062
8.1	0.4370	0.2953	1.3250
8.2	0.4370	0.2955	1.3057
8.3	0.4371	0.2960	1.3100
8.4	0.4371	0.2962	1.3111
8.5	0.4372	0.2962	1.3117
8.6	0.4372	0.2962	1.3120
8.7	0.4375	0.2961	1.3122
8.8	0.4380	0.2978	1.3178
8.9	0.4382	0.2969	1.3196
8.10	0.4382	0.2970	1.3200
8.11	0.4370	0.2954	1.3167
8.12	0.4274	0.2961	1.3206
8.13	0.4264	0.2940	1.2833
8.14	0.4215	0.2931	1.0974

Table 6.4: Values of central and minimum film thickness, and centreline pressure spike height for grid adaptation tests, in non-dimensional units

Grid 6

6.11	Refined wherever $P > 1\times10^{-2}$ on grid 6
6.12	As 6.11, with grid 5 from $P > 2\times10^{-3}$

Grid 7

7.13	Refined wherever $P > 1\times10^{-2}$ on grid 7
7.14	As 7.13, with grid 6 from $P > 2\times10^{-3}$
7.15	As 7.14, with grid 5 from $P > 5\times10^{-4}$

Grid 8

8.15	Refined wherever $P > 1\times10^{-2}$ on grid 8
8.16	As 8.15, with grid 7 from $P > 2\times10^{-3}$
8.17	As 8.16, with grid 6 from $P > 5\times10^{-4}$
8.18	As 8.17, with grid 5 from $P > 1\times10^{-6}$

Table 6.5: Grid adaptation schemes used for pressure dependent adaptivity without refinement to the cavitation boundary

Test	Central film thickness	Minimum film thickness	Centreline pressure spike height
6.0	0.4488	0.3026	1.1679
6.11	0.4490	0.3027	1.1675
6.12	0.3491	0.2235	1.4918
7.0	0.4411	0.2978	1.2435
7.13	0.4414	0.2960	1.2452
7.14	0.4304	0.2762	1.3437
7.15	0.3506	0.1499	1.5739
8.0	0.4369	0.2954	1.3062
8.15	0.4363	0.2928	1.2998
8.16	0.4348	0.2917	1.3594
8.17	0.4725	0.2690	1.4276
8.18	0.3848	0.1777	1.6698

Table 6.6: Values of central and minimum film thickness, and centreline pressure spike height for pressure dependent adaptation tests, in non-dimensional units, without refinement to the cavitation boundary

Pressure

	Compared against test			
	6.0	6.1	5.0	5.1
6.1	2.11×10^{-5}	-	9.04×10^{-3}	7.70×10^{-3}
6.2	2.07×10^{-5}	2.78×10^{-7}	9.02×10^{-3}	7.68×10^{-3}
6.3	1.26×10^{-5}	1.73×10^{-6}	8.91×10^{-3}	7.59×10^{-3}
6.4	1.07×10^{-5}	2.73×10^{-6}	8.88×10^{-3}	7.56×10^{-3}
6.5	1.76×10^{-5}	3.27×10^{-5}	8.46×10^{-3}	7.22×10^{-3}
6.6	3.89×10^{-5}	4.04×10^{-5}	8.41×10^{-3}	7.18×10^{-3}
6.7	1.53×10^{-7}	2.13×10^{-5}	8.57×10^{-3}	6.61×10^{-3}
6.8	1.53×10^{-7}	2.13×10^{-5}	8.57×10^{-3}	7.31×10^{-3}
6.9	1.03×10^{-4}	1.97×10^{-4}	7.72×10^{-3}	7.31×10^{-3}
6.10	1.53×10^{-2}	1.54×10^{-2}	1.64×10^{-2}	1.52×10^{-2}

Film thickness

	Compared against test			
	6.0	6.1	5.0	5.1
6.1	3.64×10^{-5}	-	4.88×10^{-1}	4.86×10^{-1}
6.2	3.18×10^{-5}	1.72×10^{-7}	4.88×10^{-1}	4.86×10^{-1}
6.3	8.62×10^{-6}	1.09×10^{-5}	4.88×10^{-1}	4.86×10^{-1}
6.4	5.04×10^{-6}	1.69×10^{-5}	4.88×10^{-1}	4.86×10^{-1}
6.5	5.90×10^{-6}	7.09×10^{-5}	4.87×10^{-1}	4.86×10^{-1}
6.6	1.38×10^{-5}	9.41×10^{-5}	4.87×10^{-1}	4.86×10^{-1}
6.7	5.11×10^{-9}	3.69×10^{-5}	4.87×10^{-1}	4.86×10^{-1}
6.8	5.11×10^{-9}	3.69×10^{-5}	4.87×10^{-1}	4.86×10^{-1}
6.9	2.50×10^{-4}	4.69×10^{-5}	4.87×10^{-1}	4.86×10^{-1}
6.10	1.46×10^{-2}	1.69×10^{-5}	5.08×10^{-1}	5.04×10^{-1}

Table 6.7: L_2-norms of differences in non-dimensionalised pressure and film thickness between adapted and unadapted cases on grid 6

have been made against the fully converged results, i.e. those of tests 5.0, 6.0, 7.0 and 8.0, and against those of the similarly converged unadapted cases of tests 5.1, 6.1, 7.1 and 8.1. These results are presented in Tables 6.7 to 6.10. Again, by considering the three different adaptation methods, the overall usefulness of the methods may be evaluated. It is important, first, to note that even the unadapted cases, i.e. 6.1, 7.1 and 8.1, have a marked difference in terms of the film thickness L_2-norm, when compared to coarser grids. This translates to up to six orders of magnitude difference. It can be seen in all the cases that these differences will never be significantly reduced. Comparison between the results of the fully converged and the ten iteration results are also seen to be comparable, showing that generally the results achieved after these few cycles are not radically different from those with near-zero numerical residuals.

Pressure

	Compared against test					
	7.0	7.1	6.0	6.1	5.0	5.1
7.1	1.01×10^{-4}	-	4.70×10^{-3}	4.35×10^{-3}	1.84×10^{-2}	1.66×10^{-2}
7.2	1.19×10^{-4}	5.50×10^{-6}	4.57×10^{-3}	4.23×10^{-3}	1.82×10^{-2}	1.64×10^{-2}
7.3	1.48×10^{-4}	1.91×10^{-5}	4.43×10^{-3}	4.10×10^{-3}	1.80×10^{-2}	1.62×10^{-2}
7.4	1.68×10^{-4}	3.30×10^{-5}	4.35×10^{-3}	4.03×10^{-3}	1.79×10^{-2}	1.61×10^{-2}
7.5	1.75×10^{-4}	3.83×10^{-5}	4.32×10^{-3}	4.00×10^{-3}	1.78×10^{-2}	1.61×10^{-2}
7.6	1.71×10^{-4}	3.75×10^{-5}	4.30×10^{-3}	3.99×10^{-3}	1.78×10^{-2}	1.60×10^{-2}
7.7	2.23×10^{-4}	7.28×10^{-5}	4.16×10^{-3}	3.86×10^{-3}	1.76×10^{-2}	1.59×10^{-2}
7.8	2.44×10^{-4}	8.82×10^{-5}	4.12×10^{-3}	3.82×10^{-3}	1.75×10^{-2}	1.58×10^{-2}
7.9	1.22×10^{-4}	1.30×10^{-5}	4.45×10^{-3}	4.13×10^{-3}	1.80×10^{-2}	1.62×10^{-2}
7.10	2.12×10^{-4}	6.47×10^{-5}	4.19×10^{-3}	3.88×10^{-3}	1.76×10^{-2}	1.59×10^{-2}
7.11	2.81×10^{-3}	3.26×10^{-3}	8.11×10^{-3}	7.72×10^{-3}	2.01×10^{-2}	1.92×10^{-2}
7.12	6.51×10^{-3}	7.00×10^{-3}	7.51×10^{-3}	7.28×10^{-3}	1.49×10^{-2}	1.32×10^{-2}

Film thickness

	Compared against test					
	7.0	7.1	6.0	6.1	5.0	5.1
7.1	1.18×10^{-5}	-	2.43×10^{-1}	2.42×10^{-1}	4.59×10^{-1}	4.57×10^{-1}
7.2	2.07×10^{-6}	1.69×10^{-5}	2.43×10^{-1}	2.42×10^{-1}	4.59×10^{-1}	4.56×10^{-1}
7.3	2.62×10^{-6}	6.75×10^{-5}	2.42×10^{-1}	2.42×10^{-1}	4.58×10^{-1}	4.56×10^{-1}
7.4	5.92×10^{-5}	1.18×10^{-4}	2.42×10^{-1}	2.42×10^{-1}	4.58×10^{-1}	4.56×10^{-1}
7.5	7.27×10^{-5}	1.37×10^{-4}	2.42×10^{-1}	2.42×10^{-1}	4.58×10^{-1}	4.56×10^{-1}
7.6	4.92×10^{-5}	1.04×10^{-4}	2.42×10^{-1}	2.42×10^{-1}	4.58×10^{-1}	4.56×10^{-1}
7.7	1.15×10^{-4}	1.95×10^{-4}	2.42×10^{-1}	2.42×10^{-1}	4.58×10^{-1}	4.56×10^{-1}
7.8	1.45×10^{-4}	2.33×10^{-4}	2.42×10^{-1}	2.42×10^{-1}	4.58×10^{-1}	4.56×10^{-1}
7.9	9.82×10^{-6}	3.95×10^{-5}	2.42×10^{-1}	2.42×10^{-1}	4.59×10^{-1}	4.57×10^{-1}
7.10	1.07×10^{-4}	1.84×10^{-4}	2.42×10^{-1}	2.42×10^{-1}	4.58×10^{-1}	4.56×10^{-1}
7.11	3.40×10^{-3}	3.05×10^{-3}	2.48×10^{-1}	2.47×10^{-1}	4.68×10^{-1}	4.65×10^{-1}
7.12	2.11×10^{-3}	1.86×10^{-3}	2.46×10^{-1}	2.46×10^{-1}	4.66×10^{-1}	4.62×10^{-1}

Table 6.8: L_2-norms of differences in non-dimensionalised pressure and film thickness between adapted and unadapted cases on grid 7

Pressure

	Compared against test			
	8.0	8.1	7.0	7.1
8.1	1.20×10^{-5}	-	2.50×10^{-3}	2.78×10^{-3}
8.2	4.09×10^{-6}	6.37×10^{-6}	2.49×10^{-3}	2.78×10^{-3}
8.3	6.70×10^{-6}	8.55×10^{-6}	2.36×10^{-3}	2.64×10^{-3}
8.4	1.05×10^{-5}	1.34×10^{-5}	2.31×10^{-3}	2.58×10^{-3}
8.5	1.27×10^{-5}	1.63×10^{-5}	2.28×10^{-3}	2.56×10^{-3}
8.6	1.44×10^{-5}	1.85×10^{-5}	2.27×10^{-3}	2.54×10^{-3}
8.7	1.44×10^{-5}	1.40×10^{-5}	2.29×10^{-3}	2.56×10^{-3}
8.8	4.27×10^{-5}	5.33×10^{-5}	2.09×10^{-3}	2.35×10^{-3}
8.9	6.56×10^{-5}	7.91×10^{-5}	2.02×10^{-3}	2.27×10^{-3}
8.10	7.35×10^{-5}	8.80×10^{-5}	1.99×10^{-3}	2.24×10^{-3}
8.11	9.93×10^{-6}	1.54×10^{-5}	2.54×10^{-3}	2.82×10^{-3}
8.12	2.19×10^{-5}	1.13×10^{-5}	2.54×10^{-3}	2.83×10^{-3}
8.13	1.53×10^{-3}	1.38×10^{-3}	4.09×10^{-3}	4.66×10^{-3}
8.14	4.63×10^{-3}	4.15×10^{-3}	4.74×10^{-3}	5.36×10^{-3}

	Compared against test			
	6.0	6.1	5.0	5.1
8.1	1.02×10^{-2}	9.70×10^{-3}	2.47×10^{-2}	2.26×10^{-2}
8.2	1.01×10^{-2}	9.62×10^{-3}	2.45×10^{-2}	2.24×10^{-2}
8.3	9.91×10^{-3}	9.43×10^{-3}	2.43×10^{-2}	2.22×10^{-2}
8.4	9.83×10^{-3}	9.36×10^{-3}	2.42×10^{-2}	2.22×10^{-2}
8.5	9.79×10^{-3}	9.32×10^{-3}	2.41×10^{-2}	2.21×10^{-2}
8.6	9.76×10^{-3}	9.30×10^{-3}	2.41×10^{-2}	2.21×10^{-2}
8.7	9.81×10^{-3}	9.34×10^{-3}	2.42×10^{-2}	2.21×10^{-2}
8.8	9.49×10^{-3}	9.04×10^{-3}	2.38×10^{-2}	2.18×10^{-2}
8.9	9.37×10^{-3}	8.92×10^{-3}	2.37×10^{-2}	2.17×10^{-2}
8.10	9.33×10^{-3}	8.88×10^{-3}	2.36×10^{-2}	2.16×10^{-2}
8.11	1.02×10^{-2}	9.72×10^{-3}	2.47×10^{-2}	2.26×10^{-2}
8.12	1.02×10^{-2}	9.70×10^{-3}	2.46×10^{-2}	2.25×10^{-2}
8.13	1.11×10^{-2}	1.06×10^{-2}	2.41×10^{-2}	2.20×10^{-2}
8.14	8.94×10^{-3}	8.60×10^{-3}	1.89×10^{-2}	1.71×10^{-2}

Table 6.9: L_2-norms of differences in non-dimensionalised pressure between adapted and unadapted cases on grid 8

Film thickness

	Compared against test			
	8.0	8.1	7.0	7.1
8.1	2.16×10^{-6}	-	1.21×10^{-1}	1.21×10^{-1}
8.2	3.04×10^{-7}	8.91×10^{-7}	1.21×10^{-1}	1.21×10^{-1}
8.3	8.69×10^{-4}	1.84×10^{-5}	1.21×10^{-1}	1.21×10^{-1}
8.4	1.98×10^{-5}	3.37×10^{-5}	1.21×10^{-1}	1.21×10^{-1}
8.5	2.66×10^{-5}	4.25×10^{-5}	1.21×10^{-1}	1.21×10^{-1}
8.6	3.15×10^{-5}	4.86×10^{-5}	1.21×10^{-1}	1.21×10^{-1}
8.7	2.26×10^{-5}	3.71×10^{-5}	1.21×10^{-1}	1.21×10^{-1}
8.8	1.16×10^{-4}	1.47×10^{-4}	1.21×10^{-1}	1.21×10^{-1}
8.9	1.76×10^{-4}	2.14×10^{-4}	1.21×10^{-1}	1.21×10^{-1}
8.10	1.97×10^{-4}	2.37×10^{-4}	1.21×10^{-1}	1.21×10^{-1}
8.11	2.04×10^{-6}	3.10×10^{-8}	1.21×10^{-1}	1.21×10^{-1}
8.12	6.63×10^{-6}	2.24×10^{-6}	1.21×10^{-1}	1.21×10^{-1}
8.13	4.32×10^{-4}	3.86×10^{-4}	1.21×10^{-1}	1.22×10^{-1}
8.14	1.24×10^{-3}	1.17×10^{-3}	1.23×10^{-1}	1.23×10^{-1}

	Compared against test			
	6.0	6.1	5.0	5.1
8.1	2.28×10^{-1}	2.28×10^{-1}	4.21×10^{-1}	4.19×10^{-1}
8.2	2.28×10^{-1}	2.28×10^{-1}	4.21×10^{-1}	4.19×10^{-1}
8.3	2.28×10^{-1}	2.28×10^{-1}	4.21×10^{-1}	4.18×10^{-1}
8.4	2.28×10^{-1}	2.28×10^{-1}	4.20×10^{-1}	4.18×10^{-1}
8.5	2.28×10^{-1}	2.28×10^{-1}	4.20×10^{-1}	4.18×10^{-1}
8.6	2.28×10^{-1}	2.27×10^{-1}	4.20×10^{-1}	4.18×10^{-1}
8.7	2.28×10^{-1}	2.27×10^{-1}	4.20×10^{-1}	4.18×10^{-1}
8.8	2.28×10^{-1}	2.27×10^{-1}	4.20×10^{-1}	4.18×10^{-1}
8.9	2.27×10^{-1}	2.27×10^{-1}	4.19×10^{-1}	4.17×10^{-1}
8.10	2.27×10^{-1}	2.27×10^{-1}	4.19×10^{-1}	4.17×10^{-1}
8.11	2.28×10^{-1}	2.27×10^{-1}	4.21×10^{-1}	4.19×10^{-1}
8.12	2.28×10^{-1}	2.28×10^{-1}	4.21×10^{-1}	4.19×10^{-1}
8.13	2.30×10^{-1}	2.29×10^{-1}	4.24×10^{-1}	4.21×10^{-1}
8.14	2.32×10^{-1}	2.31×10^{-1}	4.27×10^{-1}	4.24×10^{-1}

Table 6.10: L_2-norms of differences in non-dimensionalised film thickness between adapted and unadapted cases on grid 8

Case	L_2 pressure norm	L_2 film thickness norm
6.0	9.98×10^{-3}	2.28×10^{-1}
6.1	9.51×10^{-3}	2.28×10^{-1}
6.2	9.49×10^{-3}	2.28×10^{-1}
6.3	9.62×10^{-3}	2.28×10^{-1}
6.4	9.66×10^{-3}	2.28×10^{-1}
6.5	1.01×10^{-2}	2.29×10^{-1}
6.6	1.02×10^{-2}	2.29×10^{-1}
6.7	1.00×10^{-2}	2.28×10^{-1}
6.8	1.00×10^{-2}	2.28×10^{-1}
6.9	1.12×10^{-2}	2.30×10^{-1}
6.10	2.98×10^{-2}	2.31×10^{-1}
7.0	2.41×10^{-3}	1.21×10^{-1}
7.1	2.68×10^{-3}	1.21×10^{-1}
7.2	2.82×10^{-3}	1.21×10^{-1}
7.3	3.99×10^{-3}	1.21×10^{-1}
7.4	3.10×10^{-3}	1.21×10^{-1}
7.5	3.13×10^{-3}	1.21×10^{-1}
7.6	3.14×10^{-3}	1.21×10^{-1}
7.7	3.35×10^{-3}	1.21×10^{-1}
7.8	3.43×10^{-3}	1.21×10^{-1}
7.9	2.91×10^{-3}	1.21×10^{-1}
7.10	3.32×10^{-3}	1.21×10^{-1}
7.11	3.11×10^{-3}	1.22×10^{-1}
7.12	8.17×10^{-3}	1.21×10^{-1}

Table 6.11: Solution accuracy for non-dimensionalised pressure and film thickness comparing the unadapted grid 8 against both adapted and unadapted cases on grids 6 and 7

The geometrical adaptation results are seen to be similar on all grids for both pressure and film thickness. The accuracy deteriorates as more of the mesh is de-refined, and this has the greatest effect on the film thickness, e.g. cases 7.1 to 7.5. The pressure based refinement has similar properties and is of similar accuracy for all the tests. Finally, the automatic grid adaptation cases show similar behaviour to that suggested by the tracked variables shown in Table 6.4. As the tolerance ξ increases the quality of the results decreases sharply. Of particular interest, though, are those results for the highest factors, i.e. cases 6.7, 6.8 and 7.9, which show film thickness results significantly better than the unadapted cases.

It is also worthwhile to consider the accuracy of the adapted grid results against those on finer grids. In Table 6.11 the L_2-norms of the solutions for both non-dimensionalised pressure and film thickness are shown for the tests on grids 6 and 7 against the unadapted

solution on grid 8. It is seen that, as expected, the solutions on the adapted grids have similar accuracy to those on the unadapted grids with many more mesh points.

Finally, in Table 6.12, computational timings are shown for each case. These show that solutions on adapted grids can be computed significantly faster than in the unadapted case, often up to around half the time. This means that results of significantly greater accuracy than those on grid k, say, can be computed on grid k+1 in roughly only twice the time, rather than the factor of four previously.

6.5 Adaptation Example - Transient

In this section an example is shown which combines all the methods developed in this work. With the adaptive meshing still at an early stage of development, and without any adaptation of the film thickness solve, it would be unrealistic to expect substantial savings in computational speed overall. This is especially true considering that an unadapted solve is done on the first multigrid cycle per timestep to ensure that only accurate information is being used in deciding the location for regions to be adapted.

The example chosen was the reversal example shown in Section 5.6.2 with oil entrainment initially from left to right before slowing down and reversing direction. This has been solved, as in Section 5.7.3.1, using variable timestepping. This also means that all of the computational benefits developed in Chapter 5, such as prediction of the solution at the next timestep, and convergence testing on each timestep, are utilised.

The adaptation schemes used are shown in Table 6.13 where it can be see that the pressure based adaptation scheme has been chosen. Both the plots of central and minimum film thickness were indistinguishable from each other between adaptation schemes, and, as such, there is little point in reproducing them here. Instead the accuracy can again be measured by comparing the central and minimum film thickness results, as well as the central offset film thickness, at certain reference times of the calculation. As in Table 5.6 these have been chosen to be at the point of reversal, and at the time of minimum film thickness. These results are shown in Table 6.14. It can be seen that there is almost no difference between any of the schemes in any of these key variables. This suggests that they all have the same properties in terms of accuracy.

The computational efficiency is shown in Table 6.15. It can be seen that generally the adaptive meshing cases are slightly quicker, although not by very much. The largest variation between the cases is for the number of timesteps required. Coupling this with the fact that the accuracy of the different methods is similar then it is clear that any errors being introduced by the adaptive meshing are more than compensated for by the balancing

Case	Time for 10 iterations (s)	Saving on unadapted case (s)	Percentage time saved
6.1	31.4	-	-
6.2	30.0	1.3	4.2
6.3	21.1	10.3	32.8
6.4	19.5	11.8	37.7
6.5	21.6	9.7	31.1
6.6	20.5	10.8	34.6
6.7	30.4	1.0	3.1
6.8	30.0	1.3	4.1
6.9	21.7	9.6	30.7
6.10	18.8	12.6	40.1
7.1	117.2	-	-
7.2	108.0	9.2	7.8
7.3	65.5	51.7	44.1
7.4	57.3	59.8	51.1
7.5	56.0	61.2	52.2
7.6	65.6	51.6	44.0
7.7	59.5	57.7	49.2
7.8	58.7	58.4	49.9
7.9	110.3	6.9	5.9
7.10	80.9	36.3	31.0
7.11	72.5	44.9	38.3
7.12	63.3	53.9	46.0
8.1	482	-	-
8.2	445	37	8.0
8.3	267	215	44.0
8.4	241	241	49.6
8.5	233	249	51.3
8.6	237	245	51.3
8.7	268	213	47.8
8.8	220	262	54.2
8.9	214	268	53.7
8.10	218	264	55.6
8.11	450	32	6.6
8.12	276	206	44.3
8.13	241	241	46.6
8.14	236	246	48.8

Table 6.12: Computational timings for 10 multigrid V-cycles for adaptation test cases

Case	Adaptation scheme
6.1	No grid adaptation, $tol = 0.3$
6.2	As 6.1, with cavitation region adapted
6.3	As 6.2, with refinement along line from first $P > 2\times10^{-2}$ on grid 6
6.4	As 6.3, with refinement from $P > 5\times10^{-3}$ on grid 5
6.5	No grid adaptation, $tol = 0.03$
6.6	As 6.5, with cavitation region adapted
6.7	As 6.6, with refinement along line from first $P > 2\times10^{-2}$ on grid 6

Table 6.13: Grid adaptation schemes for transient reversal example

Test case	At reversal, t=0.1 s		
	Central	Minimum	H_{00}
6.1	7.613×10^{-8}	2.263×10^{-8}	-1.207×10^{-6}
6.2	7.613×10^{-8}	2.264×10^{-8}	-1.207×10^{-6}
6.3	7.613×10^{-8}	2.264×10^{-8}	-1.207×10^{-6}
6.4	7.613×10^{-8}	2.264×10^{-8}	-1.207×10^{-6}
6.5	7.613×10^{-8}	2.264×10^{-8}	-1.207×10^{-6}
6.6	7.613×10^{-8}	2.264×10^{-8}	-1.207×10^{-6}
6.7	7.627×10^{-8}	2.267×10^{-8}	-1.207×10^{-6}
	At t=t_{min}		
	Central	Minimum	H_{00}
6.1	8.192×10^{-8}	1.779×10^{-8}	-1.202×10^{-6}
6.2	8.192×10^{-8}	1.779×10^{-8}	-1.202×10^{-6}
6.3	8.192×10^{-8}	1.779×10^{-8}	-1.202×10^{-6}
6.4	8.192×10^{-8}	1.779×10^{-8}	-1.202×10^{-6}
6.5	8.191×10^{-8}	1.779×10^{-8}	-1.202×10^{-6}
6.6	8.192×10^{-8}	1.779×10^{-8}	-1.202×10^{-6}
6.7	8.207×10^{-8}	1.782×10^{-8}	-1.202×10^{-6}

Table 6.14: Comparison between test parameters at reversal and t_{min}

Case	Time taken (s)	Timesteps taken	Iterations required
6.1	6088	378	863
6.2	5834	367	841
6.3	5944	375	859
6.4	6075	383	875
6.5	5853	361	846
6.6	5849	361	872
6.7	6136	366	1114

Table 6.15: Computational comparisons for adaptive meshing reversal cases

of these errors against the temporal errors in the choice of new timestep size.

In summary it has been seen that grid adaptation can be successfully incorporated into the EHL solver, and be combined with variable timestepping. Overall, though, it is clear that more work needs to be done to investigate transient adaptive meshing cases. If finer grids were used for the transient problems then it is likely that the speed-up per iteration will improve but it is not clear whether this would be cancelled out with the need for more timesteps to be taken. Perhaps a more comprehensive adaptation criterion is needed to combine both adaptation in space and time.

6.6 Conclusion

In this chapter the idea of adaptive meshing has been introduced and examined from the starting point of the EHL solver explained in previous chapters.

Refinement of grids has been done by refining the number of points used in the pressure solve. This has allowed the order of the linear algebra systems being solved to be significantly decreased. Solutions for film thickness have still been done on the full grid to enable easy re-use of the multilevel multi-integration algorithm.

Both steady state and transient examples have been presented showing the benefits of adaptation in terms of increasing computational performance against the accuracy of the solutions attained. For the steady state case, the grids have been adapted using three different methods. The first method was based on deciding which regions of the domain would be solved before the code was started. This method was shown to accurately solve the system to the same level of accuracy as the unadapted case, whilst reducing the time taken by up to 50%. It does, however, require *a priori* knowledge of which area of the solution domain will most benefit from grid adaptation.

The second method of grid adaptation was solution based. In the areas of highest pressure the Reynolds equation was solved, but outside these areas it was assumed that the solution was smooth enough to use a linearly interpolated values without affecting solution accuracy. A succession of pressure values were used on different grid levels to define a set of multigrid ‘patches’ which would be adapted. Again significant speed-ups were achieved for accuracies of solution. This method is preferable over the previous one because there is no need to know where the regions of high pressure are before the run is started, and hence is a more automatic method.

The third criteria used to adapt the grid was that of an error test based on the relative truncation error, τ_k^{k-1}, which is a measure of the similarity between pressure solutions on adjacent grids. This test is automatic in that only the regions of the grid demanding

greater accuracy are refined. It was shown that the chosen tolerance is very important beacuse having too little an area adapted resulted in significantly larger errors, however for correctly chosen tolerances gave accurate results up to around 40% faster. This error test needs more work before it can be used confidently for this application. Another automatic error test may well be preferable.

Adaptation has been seen to be very beneficial in reducing the required work to solve a steady state EHL problem. It is clear that refining the multi-integration method to include adaptive meshes will reap even greater benefits. Some work on this method has already been published [18] and will be an obvious future extension to this work, although this implementation will be non-trivial. Extension to transient cases has not seen similar reductions in computational time, although there has been no decrease in accuracy of results because the variable timestepping appears to have reduced the timestep size to compensate for the additional errors from the adapted meshes. Further work is clearly needed here.

Chapter 7

Conclusion

7.1 Summary

In this work the numerical solution of point contact EHL problems has been investigated. Although numerical solutions to these computationally demanding problems have been computed for a quarter of a century, modern mathematical techniques have increased the speed with which solutions can be obtained. The use of multilevel techniques have been particularly important in these advances, and these were summarised in Chapter 3.

The computational code used in this work was built on that of Nurgat. Whilst this had been able to calculate accurate solutions to single grid problems, some of the elements of the multigrid processes used still required work. In Chapter 4 these deficiencies were examined in detail. It was seen that residuals around the cavitation boundary were not falling away with continued multigrid iterations, but being reintroduced when making coarse grid corrections to the fine grid solution. First a method to avoid this stalling of residuals was proposed, and shown to be effective. Later, an accurate treatment of the free boundary was introduced inside the multigrid correction process. When this was then coupled with the correct treatment of the force balance equation in the multigrid framework, asymptotic convergence on all grids was achieved as desired.

In Chapter 4 the use of parallelism was also considered. Whilst only done for the multi-summation part of the deformation calculation for the film thickness, this did show some encouraging results. However, with the introduction of multilevel multi-integration

into the code the potential time savings from this method of parallelisation were significantly reduced. The computational timings for the use of multi-integration in this code were also shown in this chapter. The benefits of writing optimised FORTRAN were emphasised by comparing two different single processor implementations of the multi-summation process. Since conducting the multi-integration timing experiments, these optimisation techniques have been incorporated into the code and timing results in later chapters all include their benefits.

Transient EHL problems are of particular interest to industry as the behaviour of both components and lubricant in changing physical conditions is very important. The first attempt by Nurgat has been redeveloped in Chapter 5. Examples were presented showing that solution inaccuracies could grow if the temporal derivatives were not adequately controlled. These were then followed by a wide variety of test problems, namely modelling variable entrainment velocity, sinusoidal loading and examples of surface roughness. Results were compared against those of other numerical solvers and corresponding experimental results.

Error control on a timestep was considered, and experimental results were shown emphasising the advantages of choosing to test for convergence of the algebraic variable for film thickness present in the time dependent squeeze term of the Reynolds Equation. The reduction in residual levels at the start of each timestep using prediction from previous timestep solutions, was also demonstrated.

Variable timestepping was introduced using the differential algebraic properties of the system of equations. It was shown, by a series of experiments, that the chosen error test was able to identify the periods of solution with non-linear behaviour, and hence small timesteps were taken in these times. One of the examples in this section was the modelling of shock loading and unloading, where the system undergoes a large change in loading in a very short period of time.

Grid adaptation was explained in Chapter 6. Selection of areas of the mesh where solutions are relatively smooth allows a reduced domain to be considered for the solution of the equations on the finest mesh. By applying increasingly large regions on decreasing levels of refinement in the multigrid hierarchy the computational effort on the finest meshes can be directed only to these areas of the solution domain where it would be most beneficial. Three different schemes for deciding where to refine were evaluated for a steady state example. The computational time was typically reduced by a half, without a significant drop in accuracy, and solutions were still better by the same order of accuracy than those on any coarser mesh. A transient example was also shown, however whilst the results were of the same accuracy as the unadapted cases, the increase in performance was

small, since more timesteps were taken as the spatial and temporal errors were balanced.

7.2 Future Work

The code developed in this work has been shown to be capable of solving many different problems that are typical of the cases being studied in industry. The work is part of a continuing collaboration with Shell Global Solutions. Parts of the code developed during this research are already being used in their lubrication software.

There are three main directions for future work. First, demanding physical problems can now be tackled with confidence. Secondly, it has been seen that there are powerful numerical tools available, such as grid adaptation, which require further exploration and deeper understanding with regard to their application in EHL modelling. Finally, the development of an all inclusive environment for solving EHL is proposed. Each of these will be outlined in slightly more detail, and it will be clear that there is a significant overlap between them.

The problems in this work have been restricted to a very basic lubricant model. The first step away from this must come by using more realistic and applicable viscoelastic models. In addition to a different steady state solution, these will have very different transient behaviour. The solution of thermal problems is also of great importance to accurately measuring the contact conditions. With differing slide to roll ratios there will be different temperature profiles on each of the two contacts and the lubricant has another profile, too.

In Chapter 5 problems involving surface roughness were tackled. These were done by assuming an overall smoothness to the asperities. Real surfaces have far greater roughness than those represented here, and this is as varied as the measuring equipment is precise. To solve EHL problems for a measured roughness is not a technically difficult inclusion to the problem, however the resolution of the details are far greater than anything attempted in this work. To be able to solve these problems accurately, the finest mesh used must be the equivalent of about 10000×10000 points. With the use of adaptation this finely meshed region can be reduced to only those points in the contact region. Such fine meshes will, however, require a much sleeker data structure to be implemented. At present the storage is allocated for the full computational domain, even if it is not used in the adaptive solve. Exactly how easy this will be to implement is an issue requiring deep thought, especially due to the deformation calculation kernel matrix requiring solution on the entire domain for every point.

It was also explained in Chapter 6 that the grid adaptation undertaken thus far has

only been centred on reducing the order of the linear algebra system used in solving the Reynolds Equation. The next critical step in reducing computation times still further is to use adaptive meshes for the multi-integration [18]. Once this has been successfully incorporated then the transient case should be reconsidered in order to achieve the substantial speed-ups seen in the steady state case.

The large systems of equations to be solved, combined with the extra computational time increases which come with increasing the finest mesh size, e.g. increasing the fineness of the coarsest meshes used in both the multigrid and multi-integration algorithms, are examples of why effective use of parallelism should be very useful. The work of Chapter 4 has now been superseded by the use of multi-integration, but the ideas are still important. Successful partitioning of the domain to balance loads and eliminate the serial calculations will be very important. Work on parallel multigrid has already been done, e.g. [91, 92, 139], and the issues of how to best partition the domain for optimum load balancing is still unclear for adaptive meshes. Clearly the added difficulties of multi-integration will confuse the issue still further.

A *Problem Solving Environment* (PSE) is a technique for changing the nature of the way calculations are done. Codes, such as the one in this work, are usually designed, run, the results post-processed and visualised before editing the input parameters and re-running. A PSE, on the other hand, combines these steps to allow visualisation of results as the solution proceeds, and allows changes to be made to the problem being solved without recompiling or starting the solution from the beginning again. PSEs either take the form of purpose built environments such as SCIRun [79], or as additions to more widely available applications such as IRIS Explorer [156]. These have already been used to solve PDE problems, e.g. [80, 155]. The ability to write the solver in a modular form means that elements of the solver can be switched mid-solve, for example between lubricant models or roughness patterns. Computational parameters, such as adaptivity controls, iterations required, and mesh levels, will also be freely adjustable. Such an environment would allow ease of use to an end-user, and would facilitate fast interactive development of the code, whilst also enabling additional insight to be gained into the inherent difficulties of the problem being solved.

Bibliography

[1] Ai X. and Cheng H.S. 'The influence of moving dent on point EHL contact.' *STLE Tribology Transactions*, Vol 37, pp. 323–335, 1994.

[2] Ai X., Cheng H.S. and Zheng L. 'A transient model for micro-elastohydrodynamic lubrication with three dimensional irregularities.' *Trans. ASME, Journal of Tribology*, Vol 115, pp. 102–110, 1993.

[3] Ainsworth M. and Oden J.T. *A Posterori Error Estimation in Finite Element Analysis*. 2000.

[4] Archard J.F. and Cowking E.W. 'Elastohydrodynamic lubrication at point contacts.' *Proceedings of the Institute of Mechanical Engineers*, Vol 180, pp. 47–56, 1966.

[5] Ausherman V.K., Nagaraj H.S., Sanborn D.M. and Winer W.O. 'Infrared temperature mapping in elastohydrodynamic lubrication.' *Trans. ASME, Journal of Lubrication Technology*, Vol 98(2), pp. 236–243, 1976.

[6] Babuska I. and Suri M. 'The p and h-p versions of the finite element method, basic principles and properties.' *SIAM Review*, Vol 36(4), pp. 578–632, 1994.

[7] Bai D. and Brandt A. 'Local mesh refinement multigrid techniques.' *SIAM Journal on Scientific and Statistical Computing*, Vol 8(2), pp. 109–134, 1987.

[8] Baines M.J. *Moving Finite Elements*. Clarendon Press, 1994.

[9] Bair S. and Winer W.O. 'A rheological model for elastohydrodynamic contacts based on primary laboratory data.' *Trans. ASME, Journal of Lubrication Technology*, Vol 101, pp. 258–265, 1979.

[10] Barnes H.A., Hutton J.F. and Walters K. *An Introduction to Rheology*. Elsevier, 1989.

[11] Barus C. 'Isothermals, isopiestics and isometrics relative to viscosity.' *American Journal of Science*, Vol 45, pp. 87–96, 1893.

[12] Beck T. 'Multiscale techniques for electrostatics and eigenvalue problems in real space.' In: G. Hummer and L.R. Pratt (editors), *Simulation and Theory of Electrostatic Interactions in Solution*. AIP Press, New York, 1999.

[13] Berzins M. 'Temporal error control for convection-dominated equations in two space dimensions.' *SIAM Journal of Scientific Computation*, Vol 16(3), pp. 558–580, 1995.

[14] Berzins M., Dew P.M. and Furzeland R.M. 'Developing software for time-dependent problems using the method of lines and differential algebraic integrators.' *Applied Numerical Mathematics*, Vol 5, pp. 375–397, 1989.

[15] Bihari B.L. and Harten A. 'Multiresolution schemes for the numerical solution of 2-D conservation laws I.' *SIAM Journal on Scientific Computing*, Vol 18(2), pp. 315–354, 1997.

[16] Brandt A. 'Multi-level adaptive solutions to boundary value problems.' *Mathematics of Computation*, Vol 31, pp. 333–390, 1977.

[17] Brandt A. and Lubrecht A.A. 'Multilevel matrix multiplication and fast solution of integral equations.' *Journal of Computational Physics*, Vol 90(2), pp. 348–370, 1990.

[18] Brandt A. and Venner C.H. 'Multilevel evaluation of integral transforms on adaptive grids.' Technical Report WI/GC-5, Weizmann Institute of Science, 1996.

[19] Brandt A. and Venner C.H. 'Multilevel evaluation of integral transforms with asymptotically smooth kernels.' *SIAM Journal of Scientific Computing*, Vol 19(2), pp. 468–492, 1998.

[20] Brenan K.E., Campbell S.L. and Petzold L.R. *Numerical Solution of Initial-Value problems in Differential-Algebraic Equations*. SIAM, 1996.

[21] Breukink G.A.C. *Het oplossen van het Elastohydrodynamische smeringsprobleem gebruikmakend van multigrid*. Master's Thesis, University of Twente, Endschende, The Netherlands, 1986. (In Dutch).

[22] Briggs W.L. *A Multigrid Tutorial, second edition*. SIAM, 2000. ISBN 0-89871-462-1.

[23] Cameron A. and Gohar R. 'Theoretical and experimental studies of the oil in lubricated point contact.' *Proceedings of the Royal Society of London*, Vol A291, pp. 520–536, 1966.

[24] Chang L., Conry T.F. and Cusano C. 'An efficient, robust, multi-level computational algorithm for elastohydrodynamic lubrication.' *ASME Journal of Tribology*, Vol 111, pp. 193–199, 1989.

[25] Chang L., Cusano C. and Conry T.F. 'Effects of lubricant rheology and kinematic conditions on micro-elastohydrodynamic lubrication.' *Trans. ASME, Journal of Tribology*, Vol 111, pp. 344–351, 1989.

[26] Chang L. and Webster M.N. 'A study of elastohydrodynamic lubrication of rough surfaces.' *Trans. ASME, Journal of Tribology*, Vol 113, pp. 110–115, 1991.

[27] Chang L., Webster M.N. and Jackson A. 'On the pressure rippling and roughness deformation in elastohydrodynamic lubrication of rough surfaces.' *Trans. ASME, Journal of Tribology*, Vol 115, pp. 439–444, 1993.

[28] Chittenden R.J., Dowson D., Dunn J.F. and Taylor C.M. '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.' *Proceedings of the Royal Society of London*, Vol A397, pp. 245–269, 1985.

[29] Dowson D. 'Elastohydrodynamic lubrication: an introduction and a review of theoretical studies.' *Proceedings of the Institute of Mechanical Engineers*, Vol 180(3B), pp. 7–16, 1966.

[30] Dowson D. 'Elastohydrodynamic and micro-elastohydrodynamic lubrication.' *WEAR*, Vol 190, pp. 125–138, 1995.

[31] Dowson D. and Ehret P. 'Past, present and future studies in elastohydrodynamics.' *Proceedings of the Institution of Mechanical Engineers Part J, Journal of Engineering Tribology*, Vol 213, pp. 317–333, 1999.

[32] Dowson D. and Higginson G.R. 'A numerical solution of the elastohydrodynamic problem.' *Journal of Mechanical Engineering Science*, Vol 1(1), pp. 6–15, 1959.

[33] Dowson D. and Higginson G.R. *Elasto-hydrodynamic Lubrication, The Fundamentals of Roller and Gear Lubrication*. Permagon Press, Oxford, Great Britain, 1966.

[34] Dowson D., Higginson G.R. and Whitaker A.V. 'Stress distribution in lubricated rolling contacts.' In: *Proceedings of IMechE Symposium on Fatigue in Rolling Contact*, paper 6.66, pp. 66–75. 1963.

[35] Dowson D. and Taylor C.M. 'Cavitation in bearings.' *Annual Review of Fluid Mechanics*, Vol 11, pp. 35–66, 1979.

[36] Durany J., García G. and Vázquez. 'Numerical computation of free boundary problems in elastohydrodynamic lubrication.' *Applied Mathematical Modelling*, Vol 20(2), pp. 104–113, 1996.

[37] Ehret P., Dowson D. and Taylor C.M. 'On lubricant transport conditions in elastohydrodynamic conjunctions.' *Proceedings of the Royal Society of London*, Vol A454, pp. 763–787, 1998.

[38] Ehret P., Dowson D., Taylor C.M. and Wang D. 'Analysis of isothermal elastohydrodynamic point contacts lubricated by Newtonian fluids using multigrid methods.' *Proceedings of the Institute of Mechanical Engineers Part C, Journal of Mechanical Engineering Science*, Vol 211(7), pp. 493–508, 1997.

[39] Ertel A.M. 'Hydrodynamic lubrication based on new principles.' *Akad. Nauk SSSR Prikadnaya Mathematica i Mekhanika*, Vol 3(2), pp. 41–52, 1939.

[40] Evans C.R. and Johnson K.L. 'The rheological properties of elastohydrodynamic lubricants.' *Proceedings of the Institute of Mechanical Engineers*, Vol 200, pp. 303–312, 1984.

[41] Evans H.P. and Snidle R.W. 'Inverse solution of the Reynolds' equation of lubrication under point-contact elastohydrodynamic conditions.' *ASME Journal of Tribology*, Vol 103, pp. 539–546, 1981.

[42] Evans H.P. and Snidle R.W. 'The elastohydrodynamic lubrication of point contacts under high loads.' *Proceedings of the Royal Society of London*, Vol A382, pp. 183–199, 1982.

[43] Fairlie R. 'Optimised FDMG implementation of multi-summation for EHL problems.' Private communication, 2000.

[44] Fairlie R. 'Investigation into the effects of grid resolution on the convergence of EHL line contact problems.' Private communication, 2001.

[45] Ford J., Chen K. and Scales L.E. 'A new wavelet transform preconditioner for iterative solution of elastohydrodynamic lubrication problems.', 1999. Report, Chester College, England.

[46] Fulton S.R. 'A comparison of multilevel adaptive methods for hurricane track prediction.' *Electronic Transactions on Numerical Analysis*, Vol 6, pp. 120–132, 1997.

[47] Gear C.W. 'The simultaneous numerical solution of differential-algebraic equations.' *IEEE Transactions on Circuit Theory*, Vol CT-18, pp. 89–95, 1971.

[48] Glovnea R.P. and Spikes H.A. 'The influence of cam-follower motion on elastohydrodynamic film thickness.' In: G. Dalmaz *et al.* (editor), *Tribology Research: From Model Experiment to Industrial Problem, Proceedings of the 27th Leeds-Lyon Symposium on Tribology*. Elsevier, to appear.

[49] Goodyer C.E., Fairlie R., Berzins M. and Scales L.E. 'An in-depth investigation of the multigrid approach for steady and transient EHL problems.' In: D. Dowson *et al.* (editor), *Thinning films and Tribological Interfaces, Proceedings of the 26th Leeds-Lyon Symposium on Tribology*, pp. 95–102. Elsevier, 2000.

[50] Goodyer C.E., Fairlie R., Berzins M. and Scales L.E. 'Adaptive techniques for elastohydrodynamic lubrication solvers.' In: G. Dalmaz *et al.* (editor), *Tribology Research: From Model Experiment to Industrial Problem, Proceedings of the 27th Leeds-Lyon Symposium on Tribology*, pp. 709–719. Elsevier, 2001.

[51] Greenwood J.A. 'Transverse roughness in elastohydrodynamic lubrication.' *Proceedings of the Institution of Mechanical Engineers Part J, Journal of Engineering Tribology*, Vol 213, pp. 383–396, 1999.

[52] Greenwood J.A. 'Non-Newtonian lubrication.' In: G. Dalmaz *et al.* (editor), *Tribology Research: From Model Experiment to Industrial Problem, Proceedings of the 27th Leeds-Lyon Symposium on Tribology*. Elsevier, to appear.

[53] Greenwood J.A. and Kauzlarich J.J. 'Inlet shear heating in elastohydrodynamic lubrication.' *Trans. ASME, Journal of Lubrication Technology*, Vol 95(4), pp. 401–416, 1973.

[54] Grubin A.N. and Vinogradova I.E. *Investigation of the contact of machine components*. Central Scientific Research Institute for Technology and Mechanical Engineering, Book No. 30, Moscow, 1949. DSIR translation No. 337.

[55] Guangteng G., Cann P.M., Olver A.V. and Spikes H.A. 'An experimental study of film thickness between rough surfaces in EHD contacts.' *Tribology International*, Vol 33, pp. 183–189, 2000.

[56] Guangteng G., Cann P.M., Olver A.V. and Spikes H.A. 'Lubricant film thickness in rough surface, mixed elastohydrodynamic contact.' *Trans. ASME, Journal of Tribology*, Vol 122, pp. 65–76, 2000.

[57] Gwynllyw D.R., Davies A.R. and Phillips T.N. 'On the effects of a piezoviscous lubricant on the dynamics of a journal bearing.' *Journal of Rheology*, Vol 40, pp. 1239–1266, 1996.

[58] Hackbusch W. *Multi-grid methods and applications*. Springer-Verlag, 1985.

[59] Hairer E., Nørsett S.P. and Wanner G. *Solving Ordinary Differential Equations I: Nonstiff Problems*. Springer-Verlag, 1980.

[60] Hamrock B.J. and Dowson D. 'Isothermal elastohydrodynamic lubrication of point contacts: Part I - theoretical formulation.' *Trans. ASME, Journal of Lubrication Technology*, Vol 98, pp. 223–229, 1976.

[61] Hamrock B.J. and Dowson D. 'Isothermal elastohydrodynamic lubrication of point contacts: Part II - ellipticity parameter results.' *Trans. ASME, Journal of Lubrication Technology*, Vol 98, pp. 375–383, 1976.

[62] Hamrock B.J. and Dowson D. 'Isothermal elastohydrodynamic lubrication of point contacts: Part III - fully flooded results.' *Trans. ASME, Journal of Lubrication Technology*, Vol 99, pp. 264–276, 1977.

[63] Hamrock B.J. and Dowson D. 'Isothermal elastohydrodynamic lubrication of point contacts: Part IV - starvation results.' *Trans. ASME, Journal of Lubrication Technology*, Vol 99(1), pp. 15–23, 1977.

[64] Hamrock B.J. and Dowson D. *Ball Bearing Lubrication. The Elastohydrodynamics of elliptical contacts*. John Wiley, New York, 1981.

[65] Hamrock B.J. and Jacobson B.O. 'Elastohydrodynamic lubrication of line contacts.' *ASLE Transactions*, Vol 27(4), pp. 275–287, 1984.

[66] Hamrock B.L., Lee R.T. and Pan P. 'Pressure spikes in elastohydrodynamically lubricated conjunctions.' *ASME Journal of Tribology*, Vol 110, pp. 279–284, 1988.

[67] Hertz H. 'The contact of elastic solids.' *Journal für die reine angew. Math.*, Vol 92, pp. 156–171, 1881.

[68] Hindmarsh A.C. 'LSODE and LSODI, two new initial value ordinary differential equation solvers.' *ACM-SIGNUM Newsletters*, Vol 15, pp. 10–11, 1980.

[69] Hooke C.J. 'The minimum film thickness in line contacts during reversal of entrainment.' *Trans. ASME, Journal of Tribology*, Vol 115, pp. 191–199, 1993.

[70] Hooke C.J. 'The behaviour of heavily loaded line contacts with transverse roughness.' *Proceedings of the Institution of Mechanical Engineers Part C, Journal of Mechanical Engineering Science*, Vol 213, pp. 309–320, 1999.

[71] Hooke C.J. 'The behaviour of low-amplitude surface roughness under line contacts.' *Proceedings of the Institution of Mechanical Engineers Part J, Journal of Engineering Tribology*, Vol 213(4), pp. 275–286, 1999.

[72] Hooke C.J. 'The behaviour of low amplitude surface roughness under line contacts: non-Newtonian fluids.' *Proceedings of the Institution of Mechanical Engineers Part J, Journal of Engineering Tribology*, Vol 214(3), pp. 253–265, 2000.

[73] Hooke C.J. and Venner C.H. 'Surface roughness attenuation in line and point contacts.' *Proceedings of the Institution of Mechanical Engineers Part J, Journal of Engineering Tribology*, Vol 214(5), pp. 439–444, 2000.

[74] Houpert L.G. and Hamrock B.J. 'A fast approach for calculating film thicknesses and pressures in elastohydrodynamically lubricated contacts at high loads.' *ASME Journal of Tribology*, Vol 108, pp. 411–420, 1986.

[75] Jacob B., Lugt P.M., Dumont M.L., Tripp J.H. and Venner C.H. 'Amplitude reduction of waviness in elastohydrodynamic lubrication using an Eyring fluid model.' *Proceedings of the Institution of Mechanical Engineers Part J, Journal of Engineering Tribology*, Vol 214(4), pp. 343–350, 2000.

[76] Jacobson B. 'Thin film lubrication of real surfaces.' *Tribology International*, Vol 33, pp. 205–210, 2000.

[77] Johnson C. *Numerical solution of partial differential equations by the finite element method*. Cambridge University Press, 1987.

[78] Johnson C. 'Finite element methods for flow problems.' In: H. Deconinck and T. Barth (editors), *AGARD Report 787: Special Course on Unstructured Grid Methods for Advection Dominated Flows*, pp. 1–1–1–47. NATO, 1992.

[79] Johnson C., Parker S., Hansen C., Kindlmann G. and Livnat Y. 'Interactive simulation and visualization.' *IEEE Computer*, Vol 32(12), pp. 59–65, 1999.

[80] Johnson C.R., Berzins M., Zhukov L. and Coffey R. 'SCIRun: Application to atmospheric dispersion problems using unstructured meshes.' In: M.J. Baines (editor), *Numerical Methods for Fluid Mechanics VI*, pp. 111–122. ICFD '98, Oxford, 1998.

[81] Johnson K.L. and Tevaarwerk. 'Shear behaviour of elastohydrodynamic oil films.' *Proceedings of the Royal Society of London*, Vol A356(12), pp. 215–236, 1977.

[82] Kaneta M., Nishikawa H., Kameishi K., Sakai T. and Ohno N. 'Effect of elastic moduli of contact surfaces in elastohydrodynamic lubrication.' *Trans. ASME, Journal of Tribology*, Vol 114, pp. 75–80, 1992.

[83] Kaneta M., Sakai T. and Nishikawa H. 'Optical interferometric observations of the effects of a bump on point contact EHL.' *Trans. ASME, Journal of Tribology*, Vol 114, pp. 779–784, 1992.

[84] Kaneta M., Sakai T. and Nishikawa H. 'Effects of surface roughness on point contact EHL.' *STLE Tribology Transactions*, Vol 36(4), pp. 605–612, 1993.

[85] Kim K.H. and Sadeghi F. 'Non-Newtonian elastohydrodynamic lubrication of point contact.' *Trans. ASME, Journal of Tribology*, Vol 113(4), pp. 703–711, 1991.

[86] Kostreva M.M. 'Elasto-hydrodynamic lubrication: A nonlinear complementarity problem.' *International Journal of Numerical Methods in Fluids*, Vol 4, pp. 377–397, 1984.

[87] Lee R.T. and Hamrock B.J. 'A circular non-Newtonian model: Part I – used in elastohydrodynamic lubrication.' *Trans. ASME, Journal of Tribology*, Vol 112, pp. 486–496, 1990.

[88] Lemaître F. and Berker A. 'Non-newtonian cavitation analysis in journal bearings.' *Journal of non-Newtonian Fluid Mechanics*, Vol 59, pp. 31–48, 1995.

[89] Li X.K., Rh. G.D., Davies A.R. and Phillips T.N. 'On the influence of lubricant properties on the dynamics of two-dimensional journal bearings.' *Journal of non-Newtonian Fluid Mechanics*, Vol 93, pp. 29–59, 2000.

[90] Löhner R. *Applied computational fluid dynamics techniques : an introduction based on finite element methods*. Wiley, 2001.

[91] Lötstedt P. 'Parallel adaptive multigrid solution of compressible flow problems.' Technical Report Report 202, Department of Scientific Computing, Uppsala University, 1998.

[92] Lou J.Z. and Ferraro R. 'A parallel incompressible flow solver package with a parallel multigrid elliptic kernel.' *Journal of Computational Physics*, Vol 125, pp. 225–243, 1995.

[93] Lubrecht A.A. *Numerical solution of the EHL line and point contact problem using multigrid techniques*. Ph.D. Thesis, University of Twente, Endschende, The Netherlands, 1987. ISBN 90-9001583-3.

[94] Lubrecht A.A., Graille D., Venner C.H. and Greenwood J.A. 'Waviness amplitude reduction in EHL line contacts under rolling-sliding.' *Trans. ASME, Journal of Tribology*, Vol 120, pp. 705–709, 1998.

[95] Lubrecht A.A., ten Napel W.E. and Bosma R. 'Multigrid, an alternative method of calculating film thicknesses and pressure profiles in elastohydrodynamically lubricated line contacts.' *Trans. ASME, Journal of Tribology*, Vol 108(4), pp. 551–556, 1986.

[96] Lubrecht A.A., ten Napel W.E. and Bosma R. 'Multigrid, an alternative method of solution for two-dimensional elastohydrodynamically lubricated point contact calculations.' *Trans. ASME, Journal of Tribology*, Vol 109, pp. 437–443, 1987.

[97] Lubrecht A.A., ten Napel W.E. and Bosma R. 'The influence of longitudinal and transverse roughness on the elastohydrodynamic lubrication of circular contacts.' *Trans. ASME, Journal of Tribology*, Vol 110(3), pp. 421–426, 1988.

[98] Lubrecht A.A. and Venner C.H. 'Elastohydrodynamic lubrication of rough surfaces.' *Proceedings of the Institution of Mechanical Engineers Part J, Journal of Engineering Tribology*, Vol 213, pp. 397–403, 1999.

[99] Lubrecht A.A., Venner C.H., ten Napel W.E. and Bosma R. 'Film thickness calculations in elastohydrodynamically lubricated circular contacts, using a multigrid method.' *Trans. ASME, Journal of Tribology*, Vol 110, pp. 503–507, 1988.

[100] Martin H.M. 'Lubrication of gear teeth.' *Engineering (London)*, Vol 102, p. 199, 1916.

[101] Maxwell J.C. 'On the dynamical theory of gases.' *Encyclopaedia Britannica*, 1867.

[102] McEwen E. 'The effect of variation of viscosity with pressure on the load carrying capacity of oil films between gear teeth.' *Journal of the Institute of Petroleum*, Vol 38, p. 646, 1952.

[103] Meldahl A. 'Contribution to the theory of lubrication of gears and of the stressing of the lubricated flanks of gear teeth.' *Brown Boveri Review*, Vol 28(11), p. 374, 1941.

[104] Messé S. and Lubrecht A.A. 'Transient elastohydrodynamic analysis of an overhead cam/tappet contact.' *Proceedings of the Institution of Mechanical Engineers Part J, Journal of Engineering Tribology*, Vol 214(5), pp. 415–425, 2000.

[105] Michell A.G. *Improvements in thrust and like bearings*, 1905. British patent No. 875.

[106] Moes H. 'Discussion on a paper by D. Dowson.' *Proceedings of the Institute of Mechanical Engineers*, Vol 180, pp. 244–245, 1965.

[107] Moes H. 'Optimum similarity analysis with applications to elastohydrodynamic lubrication.', 1990. Internal memorandum, University of Twente, Endschende, The Netherlands.

[108] Nurgat E. *Numerical Methods in Lubrication Modelling*. Ph.D. Thesis, University of Leeds, Leeds, England, 1997.

[109] Nurgat E. and Berzins M. 'Multigrid methods for EHL problems.' In: N.D. Melson, T.A. Manteuffel, S.F. McCormick and C.C. Douglas (editors), *Seventh Copper Mountain Conference on Multigrid Methods*, Vol Part 2, pp. 623–636. National Aeronautics and Space Administration, Washington, D.C., 1996.

[110] Nurgat E., Berzins M. and Scales L.E. 'Solving EHL problems using iterative, multigrid and homotopy methods.' *Trans. ASME, Journal of Tribology*, Vol 121(1), pp. 28–34, 1999.

[111] Oh K.P. 'The numerical solution of dynamically loaded elastohydrodynamic contact as a nonlinear complementarity problem.' *Trans. ASME, Journal of Tribology*, Vol 106, pp. 88–95, 1985.

[112] Oh K.P. and Rohde S.M. 'Numerical solution of the point contact problem using the finite element method.' *International Journal for Numerical Methods in Engineering*, Vol 11, pp. 1507–1518, 1977.

[113] Okamura H. 'A contribution to the numerical analysis of isothermal elastohydrodynamic lubrication.' In: *Proceedings of the 9^{th} Leeds-Lyon Symposium on Tribology*, pp. 313–320. Leeds, England, 1982.

[114] Oldroyd J.G. 'On the formulations of rheological equations of state.' *Proceedings of the Royal Society*, Vol A200, pp. 523–541, 1950.

[115] Petrie C.J.S. *Elongational Flows*. Pitman, 1979.

[116] Petrov N.P. 'Friction in machines and the effect of the lubricant.' *Inzherrernii Zhurnal St. Petersburg*, Vol 1, 2, 3, pp. 71–140, 227–229, 377–436, 1883.

[117] Petrusevich A.I. 'Fundamental conclusions from the contact-hydrodynamic theory of lubrication.' *Izv. Akad. Nauk. SSSR (OTN)*, Vol 2, p. 209, 1951.

[118] Petzold L. 'Differential/Algebraic Equations are not ODE's.' *SIAM Journal of Scientific and Statistical Computation*, Vol 3(3), pp. 367–384, 1982.

[119] Polonsky I.A. and Keer L.M. 'Fast methods for solving rough contact problems: a comparative study.' *Trans. ASME, Journal of Tribology*, Vol 122, pp. 36–41, 2000.

[120] Ranger A.P., Ettles C.M.M. and Cameron A. 'The solution of the point contact elasto-hydrodynamic problem.' *Proceedings of the Royal Society of London*, Vol A346, pp. 227–244, 1975.

[121] Reynolds O. 'On the theory of lubrication and its application to Mr Beauchamp Tower's experiments, including an experimental determination of the viscosity of olive oil.' *Philosophical Transactions of the Royal Society of London*, Vol 177, pp. 157–234, 1886.

[122] Roberts G. and Walters K. 'On viscoelastic effects in journal bearing lubrication.' *Rheology Acta*, Vol 31, pp. 55–62, 1992.

[123] Roelands C.J.A. *Correlational Aspects of the viscosity-temperature-pressure relationship of lubricating oils*. Ph.D. Thesis, Technische Hogeschool Delft, The Netherlands, 1966.

[124] Rohde S.M. and Oh K.P. 'A unified treatment of thick and thin film elastohydrodynamic problems using higher order element methods.' *Proceedings of the Royal Society of London*, Vol A343, pp. 315–331, 1975.

[125] Sadeghi F. and Sui P.C. 'Compressible elastohydrodynamic lubrication of rough surfaces.' *ASME Journal of Tribology*, Vol 111, pp. 56–62, 1989.

[126] Scales L.E. 'Minimum and central film thickness results during reversal from SAE paper SP-1182.', 1998. Private communication.

[127] Scales L.E. 'Quantifying the rheological basis of traction fluid performance.' In: *Proceedings of the SAE International Fuels and Lubricants Meeting, Toronto, Canada*. Society of Automotive Engineers, 1999.

[128] Scales L.E., Rycroft J.E., Horswill N.R. and Williamson B.P. 'Simulation and observation of transient effects in elastohydrodynamic lubrication, SP-1182.' In: *SAE International Fuels and Lubricants Meeting, Dearborn, Michigan*, pp. 23–34. 1996.

[129] Schlijper A.G., Scales L.E. and Rycroft J.E. 'Current tools and techniques for EHL modelling.' *Tribology International*, Vol 29(8), pp. 669–673, 1996.

[130] Shampine L.F. 'Implementation of implicit formulas for the solution of ODEs.' *SIAM Journal of Scientific and Statistical Computation*, Vol 1(1), pp. 103–118, 1980.

[131] Shampine L.F. *Numerical Solution of Ordinary Differential Equations*. Chapman and Hall, 1994.

[132] Shampine L.F. and Gordon M.K. *Computer Solution of Ordinary Differential Equations : the Initial Value Problem*. W. H. Freeman and Co., 1975.

[133] Shaw G.J. *FDMG Multigrid Software, Version 3*.

[134] Smith C.F. 'Lubricant behaviour in concentrated contact systems – the castor oil-steel system.' *Wear*, Vol 2(4), p. 250, 1959.

[135] Speares W. and Berzins M. 'A 3d unstructured mesh adaptation algorithm for time-dependent shock-dominated problems.' *International Journal for Numerical Methods in Fluids*, Vol 25(1), pp. 81–104, 1997.

[136] Taylor C.R. and O'Callaghan J.F. 'A numerical solution of the elastohydrodynamic lubrication problem using finite elements.' *Journal of Mechanical Engineering Science*, Vol 14, pp. 229–237, 1972.

[137] Tower B. 'First report on friction experiments (friction of lubricated bearings).' *Proceedings of the Institute of Mechanical Engineers (London)*, pp. 632–659, 1883.

[138] Tower B. 'Second report on friction experiments (experiments on the oil pressure in a bearing).' *Proceedings of the Institute of Mechanical Engineers (London)*, pp. 58–70, 1885.

[139] Trottenberg U., Oosterlee C. and Schüller A. *Multigrid*. Academic Press, 2001. ISBN 0-12-701070-X.

[140] Venner C.H. *Multilevel Solution of the EHL Line and Point Contact Problems*. Ph.D. Thesis, University of Twente, Endschende, The Netherlands, 1991. ISBN 90-9003974-0.

[141] Venner C.H., Lubrecht, A. A. and ten Napel W.E. 'Numerical simulation of the overrolling of a surface feature in an EHL line contact.' *Trans. ASME, Journal of Tribology*, Vol 113, pp. 777–783, 1991.

[142] Venner C.H. and Lubrecht A.A. 'Numerical simulation of a transverse ridge in a circular EHL contact under rolling/sliding.' *Trans. ASME, Journal of Tribology*, Vol 116, pp. 751–761, 1994.

[143] Venner C.H. and Lubrecht A.A. 'Transient analysis of surface features in an EHL line contact in the case of sliding.' *Trans. ASME, Journal of Tribology*, Vol 116, pp. 186–193, 1994.

[144] Venner C.H. and Lubrecht A.A. 'Numerical analysis of the influence of waviness on the film thickness of a circular EHL contact.' *Trans. ASME, Journal of Tribology*, Vol 118, pp. 153–161, 1996.

[145] Venner C.H. and Lubrecht A.A. *Multilevel Methods in Lubrication*. Elsevier, 2000. ISBN 0-444-50503-2.

[146] Venner C.H. and Morales-Espejel G.E. 'Amplitude reduction of small-amplitude waviness in transient elastohydrodynamically lubricated line contacts.' *Proceedings of the Institution of Mechanical Engineers Part J, Journal of Engineering Tribology*, Vol 213(6), pp. 487–504, 1999.

[147] Venner C.H., ten Napel W.E. and Bosma R. 'Advanced multilevel solution of the EHL line contact problem.' *Trans. ASME, Journal of Tribology*, Vol 112, pp. 426–431, 1990.

[148] Verfürth R. *Review of Posteriori Error Estimation*. Wiley, 1996.

[149] Verstappen R. 'A simple numerical algorithm for elastohydrodynamic lubrication, based on a dynamic variation principle.' *Journal of Computational Physics*, Vol 97, pp. 460–488, 1991.

[150] Wahlbin L. *Superconvergence in Galerkin Finite Element Methods*. Springer-Verlag, 1995.

[151] Wang D. *Elastohydrodynamic lubrication of point contacts for layers of soft solids and for monolithic hard materials in the transient bouncing ball problem*. Ph.D. Thesis, Department of Mechanical Engineering, University of Leeds, England, 1994.

[152] Wesseling P. *An introduction to multigrid methods*. Wiley, 1992. ISBN 0-471-93083-0.

[153] White J.L. and Metzner A.B. 'Development of constitutive equations for polymeric melts and solutions.' *Journal of Applied Polymeric Science*, Vol 7, pp. 1867–1889, 1963.

[154] Wijnant Y. *Contact Dynamics in the Field of Elastohydrodynamic Lubrication.*. Ph.D. Thesis, University of Twente, Endschende, The Netherlands, 1998. ISBN 90-3651223-9.

[155] Wright H., Brodlie K.W. and David T. 'Navigating high-dimensional spaces to support design steering.' In: *VIS 2000*, pp. 291–296. IEEE, 2000.

[156] Wright H. and Walton J.P.R.B. 'HyperScribe: A data management facility for the dataflow visualisation pipeline.' Technical Report IETR/4, NAG, 1996.

[157] Wu S.R. ‘A penalty formulation and numerical approximation of the Reynolds-Hertz problem of elastohydrodynamic lubrication.’ *Journal of Engineering Science*, Vol 24(6), pp. 1001–1013, 1986.

[158] Zhu D. and Cheng H.S. ‘Effect of surface roughness on the point contact EHL.’ *Trans. ASME, Journal of Tribology*, Vol 110, pp. 32–37, 1988.

[159] Zienkiewicz O.C. and Morgan K. *Finite Elements and Approximation*. John Wiley and Sons, 1983.

[160] Zienkiewicz O.C. and Taylor R.L. *The Finite Element Method*. Butterworth-Heinemann, 2000.