EFFECTS OF NON-NEWTONIAN RHEOLOGY ON THE LINE CONTACT ELASTOHYDRODYNAMIC LUBRICATION PROBLEM

by

C. N. Davies

Thesis submitted in candidature for the degree of
Doctor of Philosophy at Cardiff University

Institute of Theoretical, Applied and Computational Mechanics
School of Engineering
Cardiff University

June 2005

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

Abstract

The Bair and Winer, Gecim and Winer, Circular and Cardiff limiting shear stress models are compared and contrasted under different load, viscosity, limiting shear stress, surface velocities, Young's modulus, relative radius of curvature and temperature. Isothermal comparisons show that limiting shear stress rheology reduces the pressure gradients in the centre of the contact. Also a localised increase in film thickness in the inlet and a reduction near the exit of the contact occur. The bulk oil temperature significantly influences the limiting shear stress features developed. As a result a full thermal approach was adopted that includes the heat generated, by slip, at the fluid-surface interface. Regions where limiting shear stress and temperature effects dominate are identified under optical interferometry operating conditions.

The heat generated during contact was included in the thermal approach. Sinusoidal roughness features promoted contact and the temperature responses were compared using different coefficients of friction. Contact develops higher temperatures than those generated in full film regions, with much greater temperature rises for higher coefficients of friction.

Thermal and isothermal gear meshes are simulated using equivalent roller and involute geometries with rough and smooth surfaces. Roughness data taken from gear teeth are linked to the numerical involute used in the simulation, allowing roughness features to be used in the correct part of the gear mesh. Smooth and rough surface results are obtained for three different meshes.

'Running in' was investigated by measuring the amount of surface modification and contact developed between two rotating disks using 'in situ' Talysurf profiles and contact resistance measurements respectively. 'Running in' occurred quickly, with temperature effects reducing the films prolonging the contact period. Comparisons using the software developed incorporating actual roughness profiles, bulk disk and bulk oil temperatures, gave good agreement between the numerical and experimental approaches.

Acknowledgements

The fact that this thesis has been attempted let alone completed and written is testament to the hard work and support of many people over the past 25 years!!

The work could not have been attempted without the financial support of EPSRC (Grant GR/L90996/01) and the British Gear Association. I would also like to thank the University of Cardiff and the head of the Engineering department, Prof. H. Thomas, for making available the facilities in which the work has been conducted.

I am indebted to my supervisors, Prof. H.P. Evans and Prof. R.W. Snidle, for their guidance, support and help throughout the duration of the work. I would also like to thank them both for giving me the opportunity to pursue a PhD.

I would also like to thank all of my colleagues in the Tribology and Gearing group, in particular: Dr. K. Sharif for allowing me to adapt his thermal approach for this work as well as providing many hours of entertainment and engaging in many discussions. Dr. M.J.A. Holmes, who helped me get to grips with the software provided and offered many words of wisdom. Mr. M. Alanou for introducing me to the experimental rig he designed and showing me how it all worked. Mr. A. Dhulipalah for his help in providing the SEM pictures presented. Finally, Mr. G. Puiglise, who provided much needed help during the running of the experiments detailed in this work.

Thanks must also be directed towards Mr. A. Coughlin who provided a large amount of the electronics used in the experimental work. Mr. A. Griffiths and his workshop staff, in particular Bob and Erik, helped a great deal in dealing with all of the mechanical work that was necessary in re-designing the experimental rig.

Finally, but by no means least, I would like to thank my family and friends, particularly: My parents, Paul and Sue, for their help and support whilst writing the thesis. Bres and Jugga and my team mates on the American football team for their entertainment during the time away from work. Finally, my girlfriend Ellie, who has somehow tolerated me during the highs and the lows of this work.

Contents

Chapter 3 - Non-Newtonian Behaviour

Chapter 4 - Numerical Methods

Chapter 5 - Steady State Isothermal Results

Chapter 6 - Thermal Non-Newtonian Results

Chapter 7 - Surface Contact Thermal Solution Results

Chapter 8 - EHL Simulation of the Gear Meshing Cycle

Chapter 9 - Experimental Investigations into 'Running In'

Chapter 10 - Conclusions and Recommendations

Appendix A

Nomenclature

a	Half Hertz dimension
c	Specific heat capacity
E_1, E_2	Young's modulus of surfaces
E'	Equivalent elastic modulus
h	Film thickness
h_0	Minimum film thickness
k	Thermal conductivity
p	Pressure
p_0	Maximum Hertz pressure
Q	Flow
Q_c	Couette flow
Q_p	Poiseuille flow
R	Relative radius of curvature
R_1, R_2	Radius of curvature of surfaces
r_b, r_{b1}, r_{b2}	Base radius of gear
r_{t1}, r_{t2}	Tip radius of gear
t	Time
u	Velocity
u_1, u_2	Velocity of surfaces
w_1, w_2	Rotational speed of gear
x	Axis orientated parallel to entrainment direction
z	Axis orientated perpendicular to the entrainment direction
z_1, z_2	Distance in z direction to parabola that defines surface
α	Viscosity pressure coefficient
∇	Deflection
η	Viscosity
η_0	Viscosity at reference temperature and pressure
ρ	Density

ρ_0	Density at reference pressure
θ	Temperature
θ_0	Reference temperature
ε	Thermal expansion coefficient
μ	Coefficient of friction
τ	Shear stress
τ_0	Eyring shear stress
τ_l	Limiting shear stress
λ	Non-dimensional mean shear stress for limiting shear stress
Σ	Non-dimensional pressure gradient for limiting shear stress
ν_1, ν_2	Poisson's ratio of surfaces

All other symbols are defined in the text

Chapter 1

Background and Thesis Aims

1.0 Introduction

This introductory chapter presents the field of Tribology as well as the specialised area of elastohydrodynamic lubrication (EHL). A brief discussion of some of the major discoveries and theories in the field is presented before introducing some of the concepts of fluid rheology in more detail. Finally the thesis layout and aims are covered.

1.1 Tribology and Elastohydrodynamic Lubrication

Tribology as a science and a word was introduced in a Department of Education and Science Report (1966). It is derived from the Greek word *Tribos* which can be translated to mean rubbing. The Department of Education and Science defined Tribology as:
'... The science and technology of interacting surfaces in relative motion and the practices related thereto.'

Although Tribology was given its name and definition in 1966 the principles it is related to have been used since the dawn of man. Dowson (1979) provided an interesting insight into the development and early use of aspects of Tribology. Application of the ideas and concepts encompassed by Tribology can be traced as far back as the first existence of man in the Palaeolithic period some 200,000 years ago. Early man discovered that heat was generated as a result of rubbing two pieces of wood together; this heat is generated by friction and is arguably the first evidence of the practical use of any aspect of Tribolgy.

Other tribological ideas followed through the development of man and the dawn of the first civilisations. Bearings were the first tribological systems, used nearly 7000 years ago, where antler, bone, stone and wood were used in simple rotary drills and door hinges in the Neolithic period. Lubricants and the idea of lubrication were first evident on potter's wheels as far back as 3250 BC. Solutions to the problem of wear were also developed. Cart wheels dated back to 2750 BC have been found with a metal band encasing the outside of the wooden wheel or with copper nails embedded into it.

These components and concepts that form part of Tribology were understood very early in man's existence. However, a major group of engineering components in use today that rely on Tribology to function are gears, which were not developed until much later. The idea behind gears was presented by Archimedes in 250 BC. The first evidence of the use of gears was found to date back to 87 BC where they were found in the Antikythera mechanism which was excavated from a sunken ship in 1900. The mechanism is named after the Greek island near which it was found and contains a complex arrangement of epicyclic differential gearing.

Elastohydrodynamic lubrication, EHL, is a part of the field of Tribology. Its study has developed since the middle of the twentieth century and is concerned with the consequences of applying a large load over a small non-conformal area. This produces very thin films with large pressures that affect the viscosity and density of the fluid as well as producing significant elastic deflection of both surfaces. Typical examples of EHL are in large reduction gear boxes, heavily loaded ball bearings and many other heavily loaded non-conformal contacts.

The problems considered in this thesis are mainly associated with so called line contacts. This is a specific type of contact where the curvature in the entrainment (rolling) direction dictates the problem and the contact produced is a long thin band. Other forms of contacts are point and elliptical contacts, these have a curvature perpendicular to the entrainment direction that affects the shape of the contact produced. Typical line contact problems are gear meshes and roller bearing contacts, and examples of point and elliptical contacts are spherical ball bearing contacts as well as thrust bearings.

1.2 Early Developments

There have been many investigations into the various concepts involved in Tribology and EHL. The major advances as well as the influences on the current work are summarised in the following sections.

Dowson (1979) remarked that the understanding behind the way that lubricants act has changed over the last century. Initially it was believed that lubricants filled the voids in rough surfaces, modifying the topography, allowing them to slide past one another more easily. These beliefs remained in force until Rennie (1829) proposed the idea that lubricants separate the two surfaces.

It was also noted by Dowson (1979) that almost all research before the 19th Century involved experimental investigations into friction in bearings and gear drive trains. Friction was seen as the most important quantity in fluid film lubrication. Towards the end of the 19th Century investigations began into the reliability of fluid films, the work being driven by the necessity to improve the efficiency and reliability of railroad carriages.

Beauchamp Tower was appointed by the Institution of Mechanical Engineers in 1882 to undertake experiments into friction in lubricated contacts. These experiments formed the catalyst for a new era of research, leading to more detailed investigation of the fluid films developed in a variety of different lubricated contacts. Almost by accident Tower (1883) found that the fluid films present in journal bearings produce significant pressures. Continuing his work Tower (1885) presented a second report on friction that included experimental work showing the variation of pressure in a journal bearing.

It should be noted that at the same time that Tower was discovering that fluid film pressures were generated in journal bearings, the Russian engineer Nikolai Petrov (1883) was developing equations of friction in journal bearings. Petrov conducted over 600 experiments with different oils and under different conditions and presented evidence that a fluid film existed between the surfaces. The work of Petrov and Tower is seen as the

beginning of an era of investigating the presence of a fluid film and its features rather than the more limited investigations into friction.

Tower's work influenced many people, but one of the most important works influenced by his discoveries was that by Osborne Reynolds. Reynolds (1886), outlined a general differential equation that governs fluid film lubrication. The paper also produced approximate solutions to the equation and comparisons with Tower's experimental investigations.

1.3 Theoretical Approaches

As mentioned above, very little work was undertaken before the end of the 19th Century on the theoretical aspects of EHL contacts. One exception, to this generalisation, was the analytical work presented by Hertz (1881) who, influenced by the appearance of Newton's rings between lenses in contact, set about investigating the influence that elastic deflection would have upon solid bodies in contact. His initial work on elastic deflection of loaded contacts forms the starting point of almost all theoretical approaches to EHL problems.

Initial analytical work on hydrodynamic theory as applied to non-conformal contacts did not allow for the effect of pressure on viscosity and also neglected elastic deflection of the surfaces. The most important example of these early investigations was presented by Martin (1916), using a parabolic approximation for the shape of the involute used to generate gear teeth. Martin obtained a film thicknesses formula by solving Reynolds equation for rigid surfaces and an incompressible, isoviscous fluid. It was apparent from the film thickness values predicted that they were unrealistic for most of the situations encountered at the time as the film thicknesses were much smaller than the average roughness values of the machined components such as gears. This in spite of the fact that gears were known to operate successfully with little sign of surface damage or wear.

Grubin and Vinogradova (1949) were the first to publish work including the effect of pressure on the viscosity of the fluid as well as the influence on the elastic deflection of

the surfaces. This approximate solution technique assumed a Hertzian pressure distribution and an approximation for the deflection, producing film thicknesses that were over an order of magnitude greater than that determined from Martin's theory.

A few years later, Petrusevich (1951) produced three complete solutions that confirmed the work published by Grubin and Vinogradova. This was later followed by Dowson and Higginson (1959) using an inverse solution to the Reynolds equation, allowing solutions to be achieved for a greater range of operating conditions. Although Dowson and Higginson had developed a method to solve the EHL problem, solutions were still difficult to achieve. Many different approaches were subsequently developed to try to overcome the problems of numerical convergence and stability.

Finite element techniques were used in place of the finite difference approaches used for lubrication problems when applied to the hydrodynamic pad bearing problem by Reddi (1969,1970). Taylor and O'Callaghan (1972) were the first to apply the finite element technique to EHL problems.

Newton-Raphson iterative solution techniques were first introduced to solve the non-linear equations by Rohde and Oh (1975). Their approach used the finite element formulation. The Newton-Rhapson technique was also adopted by Okamura (1983) combined with Euler's difference method. Although Okamura assumed an incompressible fluid, it was the first approach to couple the elastic deflection equation and hydrodynamic Reynolds equation. Houpert and Hamrock (1986) also used the iterative approach and were the first to use a fully coupled elastic and hydrodynamic equation that included lubricant compressibility.

The Newton-Raphson technique was suitable for the line contact problem, however the elastic deflection equation in the point contact problem is such that the resulting matrix is 'full'. Ehret et al. (1997) used an effective Newton-Rhapson approach which removed the problem of the 'full' matrix. Techniques were presented by Lubrecht et al. (1986) that solved the EHL problem using the multigrid method presented by Brandt (1977). The multigrid approach is a more effective method of solving the EHL problem as it does not require as many iterative solution sweeps as the Newton-Rhapson approach. More

recently this method has been adapted to solve a wide range of problems by Venner and Lubrecht (2000).

1.4 Experimental Film Thickness Investigation

It has already been seen that Reynolds' paper led to a substantial amount of interest in simulations of EHL contacts. Experimental investigations into the film thickness developed did not immediately follow the paper by Reynolds. In fact, it was not until after the publication of the numerical results of Dowson and Higginson (1959) that there was significant interest from experimental researchers.

The majority of early experimental apparatus was based around a two disk rig such as that developed by Merritt (1935). He developed a test rig using two disks that can be used to simulate gear tooth contact behaviour. This experimental approach was first used to investigate the friction generated in gear contacts for different oils and materials, it has since been widely adopted for other EHL contact investigations and is still in current use.

Lane and Hughes (1952) outlined an approach where the film formation in gear contacts could be assessed using electrical resistance measurements. This approach was taken further by two other authors. Crook (1961a) used a four disk machine and attached a chromium electrode to one of the disks, which was made of glass, the capacitance generated between the electrode and a mating steel disk was recorded as it passed through the contact which enabled the shape of the oil film to be determined. A different approach to measuring the film thickness was used by Sibley et al. (1960), using X-rays to determine the film thickness and shape generated. These publications confirmed the film shape and features that had been highlighted by Dowson and Higginson (1959), including the restriction present at the exit.

These approaches were followed by Furey (1961) who measured the instantaneous and average electrical resistance generated between the surfaces. This allowed Furey to infer the amount of metallic contact that occurred under different operating conditions. The

results showed that contact does not occur for much of the time and is an erratic event under EHL conditions with well prepared, smooth surfaces.

Another method of investigating film thickness is attributed to Archard and Kirk, who used a cross cylinder machine to generate a point contact and measured the electrical capacitance between the contacts [Archard and Kirk, 1961]. In subsequent papers, Kirk (1962) and Archard and Kirk (1963), they developed their approach to studying the effect that different materials had on the film produced by using cylinders of Perspex and glass. The most important development presented in these papers was the use of white light to measure the shape and size of the contact by optical interferometry. They noticed that the fringes produced moved, indicating increasing film thickness, as the speed of rotation was increased. This approach was further developed by Cameron and Gohar (1966) who presented details of the variation in film thickness for point contact problems. This was the first really effective use of optical interferometry which subsequently became widely adopted to investigate film thickness and shape in a variety of contacts.

Conventional optical interferometry using white light could only be used to measure film thicknesses above about 60nm. Different light sources were adopted by different authors to try and measure smaller film thicknesses, for example monochromatic light was used by Gohar and Cameron (1967) and duochromatic light by Wedeven (1970). A most significant development was made by Johnston, Wayte and Spikes (1991), who combined an optical spacer layer, first used by Westlake and Cameron (1967), with a spectrometer. This ultra-thin film interferometry approach now allows measurements of film thickness to be made down to 1nm.

The ultra-thin film interferometry approach was taken a stage further using colour imaging techniques by Cann, Spikes and Hutchinson (1996). This method allows the film thickness to be determined by examining the interference colour at a point which is then used to produce film thickness maps. This approach is called the Spacer Layer Imaging Method, abbreviated to SLIM.

Most of the experimental work on EHL had concentrated on the way in which the film varies through the contact, not much work had been undertaken on the pressure response that was developed in the contacts since that presented by Tower (1885) in relation to

journal bearings. Kannel, Bell and Allen (1964) used thin manganin strips to infer an accurate pressure distribution within EHL contacts and produced similar distributions to those seen in the numerical work of Dowson and Higginson.

1.5 Thermal Considerations

Early numerical work overestimated the friction that was being measured in EHL contacts. This difference was thought to be due to the increase in temperature that would occur due to the shear heating of the oil. Thus many authors incorporated thermal effects into their numerical approaches in an effort to explain this behaviour.

Most initial work considered the inlet heating effect using a simplified energy equation, for example Crook (1961b). Greenwood and Kauzlarich (1973) investigated the inlet heating effect in greater detail and concluded that the inlet shear heating was responsible for the reduced film thickness seen in experimental results when compared to numerical simulations at high speed. Full thermal solutions were restricted to Newtonian behaviour until 1991 when Sui and Sadeghi (1991) and Salehizadeh and Saka (1991) provided thermal solutions to the non-Newtonian Eyring rheological model. Chang and Farnum (1992) were the first to apply the full energy equation to the solid surfaces. Until this point all investigations had used simple integration of the surface heat flux. This is based on the (correct) assumption that the dominant heat transfer mode is conduction perpendicular to the surface so that the conduction results given by Carslaw and Jaeger (1959) may be utilised. This method becomes impractical in transient situations such as that considered by Chang and Farnum (1992).

Many experimental researchers have investigated surface temperatures but less interest has been directed toward the temperature generated within the oil. Turchina, Sandborn and Winer (1974) introduced an infrared radiation technique to measure the temperature in EHL point contacts. This was refined by Ausherman et al. (1976) where temperature rises up to 100°C were shown. Recent studies such as that by Yagi, Kyogoku and Nakahara (2004) using this technique have been able to measure the temperature profiles across and along the oil film.

1.6 Surface Roughness

The effect that the surface features generated from manufacturing processes have on the fluid film developed has increased in importance in recent studies. The influence that these features have on the pressure generated by the fluid as well as the resulting deflection of the features themselves has attracted much interest. Initially, great difficulties occurred when trying to incorporate surface roughness into an EHL analysis. As a simplification in early work, sinusoidal functions were used to aid the understanding of the response to such perturbations. The inclusion of such roughness features led to a new form of EHL, termed micro-EHL, where due to the asperity tips, the film becomes locally very thin resulting in large pressure ripples. In some cases, the asperity tips can be seen to behave as individual 'micro' EHL contacts with their own pressure spike and film thickness minimum.

One of the first approaches to consider an asperity within an EHL contact was that of Lee and Cheng (1973), where a transverse asperity was run against a smooth surface for the conditions of pure rolling as well as moderate slide roll ratios. This was followed by Patir and Cheng (1978), who developed an average flow model and investigated the effect of the orientation of the roughness, finding that the minimum film thickness increases as the roughness is changed from longitudinal through isotropic to transverse.

Lubrecht, ten Napel and Bosma (1988) used the multigrid, or multilevel, method to produce full solutions to longitudinal and transverse roughness. The roughness features used were single sinusoidal features as well as wavy surfaces. They found that the deformation is larger for transverse roughness and also concluded that the results by Patir and Cheng (1978) overestimated the effects due to the roughness orientation.

Kweh et al. (1992) used two sinusoids of different period and amplitude to simulate real roughness behaviour. Short wavelength, low amplitude features represented the roughness variation which were combined with longer wavelength, larger amplitude features to represent the waviness. It was found that the roughness features underwent less proportional deformation than the waviness feature.

As the numerical solution schemes progressed much more interest was directed towards the incorporation of real roughness. The first inclusion of real roughness into an EHL problem was made by Poon and Sayles (1994) and also Ai and Cheng (1994). Poon and Sayles (1994) considered a plasticity model and did not consider the influence of the fluid film.

With the trend to include real surfaces, the film thicknesses tended to become very small and in some cases metal to metal contact was predicted, this leads to the regime of mixed lubrication. Mixed lubrication is a situation in which a portion of the applied load is carried by metal to metal contact events and the remainder is carried by a lubricating film. Tao et al. (2003) and Holmes et al. (2003a,2003b) were the first to solve real rough surfaces that involve metal to metal contact within a consistent mass conserving numerical approach. Hu and Zhu (2000) have also presented a mass conserving approach, but the accuracy of their approach has recently been brought into question by several communications in the technical press, for example by Venner (2005).

It is relatively simple to use a function to describe a feature within EHL numerical situations. It is a little more complicated to generate such features on real components, termed artificial asperities, that can be used in experimental tests. The first use of artificial asperities was by Jackson and Cameron (1976), where features were created on a steel ball using a lathe. Other features soon followed, such as the work by Wedeven and Cusano (1979), investigating the effect that bumps and dents had on the fluid film.

Repeatable features were rather difficult to achieve using lathes and other machinery. An approach that allowed multiple features, which are reasonably similar to one another, to be created on a surface was presented by Kaneta and Cameron (1980). This technique allows individual features as well as multiple features to be created by sputtering Chromium onto a steel ball through a foil mask. Kaneta and Cameron used this approach to investigate the effects that multiple bumps have on circular point contacts.

This artificial asperity approach developed by Kaneta and Cameron has more recently been used by Kaneta, Sakai and Nishikawa (1992, 1993) to investigate the effects that bumps have on circular point contacts.

1.7 Gear Meshing Cycle

The meshing cycle of involute spur gear teeth is complex. This EHL problem is difficult to solve due to the constant changes in the load, relative radius of curvature and surface speeds. Early work did not consider transient effects and solved the problem of the gear meshing cycle as a series of steady state solutions. Wang and Cheng (1981a, 1981b) used a Grubin analysis to study the change in film thickness along the path of contact. Larsson (1997) provided a full transient analysis of the meshing cycle of involute spur gear teeth illustrating the change in film thickness and pressure. More recently, Wang et al. (2004) introduced temperature effects into the full transient approach and used a more appropriate load variation than that used by Larsson (1997). To the author's knowledge, all studies to date have used a simplified cylindrical geometry to represent the change in radius of curvature of the gear teeth.

1.8 Non-Newtonian Rheology

Most of the numerical work before 1970 was based on a Newtonian fluid. Smith (1959) presented the idea that the lubricant in the high pressure regions shears like a plastic solid. This conclusion, along with the evidence that the thermal Newtonian results tended to over estimate the friction developed within the contacts, led to different formulations that would allow a reduction in viscosity in the high pressure region.

1.8.1 Rheological Models

The first rheological model that was utilised was that of Johnson and Tevaarwerk (1977), using a nonlinear Maxwell model. The non-linear viscous term used was that attributed to Eyring (1936), which was widely adopted as it is based on an activation energy concept. The combination of the two terms yields the following relationship:

$$\dot{\gamma} = \frac{1}{G}\frac{\partial \tau}{\partial t} + \frac{\tau_0}{\eta}\sinh\left(\frac{\tau}{\tau_0}\right)$$

The strain rate shown above consists of two parts, the first term is the elastic strain rate and the second term is the viscous strain rate. The elastic effects in flows with low Deborah numbers are found to be negligible. The Deborah number, D, is a ratio of the relaxation time of the fluid over its time of passage through the contact:

$$D = \frac{\eta \bar{u}}{Ga}$$

The Deborah number for the majority of cases considered in this work varies between 1.1×10^{-8} and 3.6×10^{-7}, hence the elastic term is ignored giving the Eyring rheological model as:

$$\dot{\gamma} = \frac{\tau_0}{\eta}\sinh\left(\frac{\tau}{\tau_0}\right) \tag{1.1}$$

Although this model is still widely used today, it does not describe the friction responses of EHL contacts very well. The model simply reduces the rate of increase of the shear stress, making it non-linear, once the shear stress exceeds τ_0. An alternative view is that the shear stress is constrained by a limiting value. The first model based on this idea was presented by Bair and Winer (1979), shown in equation (1.2). This was based on experimental work, but was difficult to implement in numerical work because of the necessity to incorporate two different equations depending on the shear stress sign.

$$\dot{\gamma} = \begin{cases} -\frac{\tau_l}{\eta}\ln\left(1-\frac{\tau}{\tau_l}\right) & \tau > 0 \\ \frac{\tau_l}{\eta}\ln\left(1+\frac{\tau}{\tau_l}\right) & \tau < 0 \end{cases} \tag{1.2}$$

This complication was resolved by the model developed by Gecim and Winer (1980) where the amended viscous term is shown in equation (1.3). This model is easier to incorporate into the numerical problems, but it is more aggressive in the transition from viscous to plastic behaviour.

$$\dot{\gamma} = \frac{\tau_l}{\eta}\tanh^{-1}\left(\frac{\tau}{\tau_l}\right) \tag{1.3}$$

Equation (1.3) is also difficult to incorporate into the numerical EHL problem. To overcome this difficulty another rheological model was proposed by Lee and Hamrock (1990), this model is called the 'circular' model and is shown in Equation (1.4).

$$\dot{\gamma} = \frac{\tau_l}{\eta} \frac{\frac{\tau}{\tau_l}}{\sqrt{1 - \frac{\tau^2}{\tau_l^2}}} \tag{1.4}$$

This model was based on previous work by Iivonen and Hamrock in 1989, published in Iivonen and Hamrock (1991). They presented the first work using a full numerical solution technique incorporating a non-Newtonian limiting shear stress rheological model.

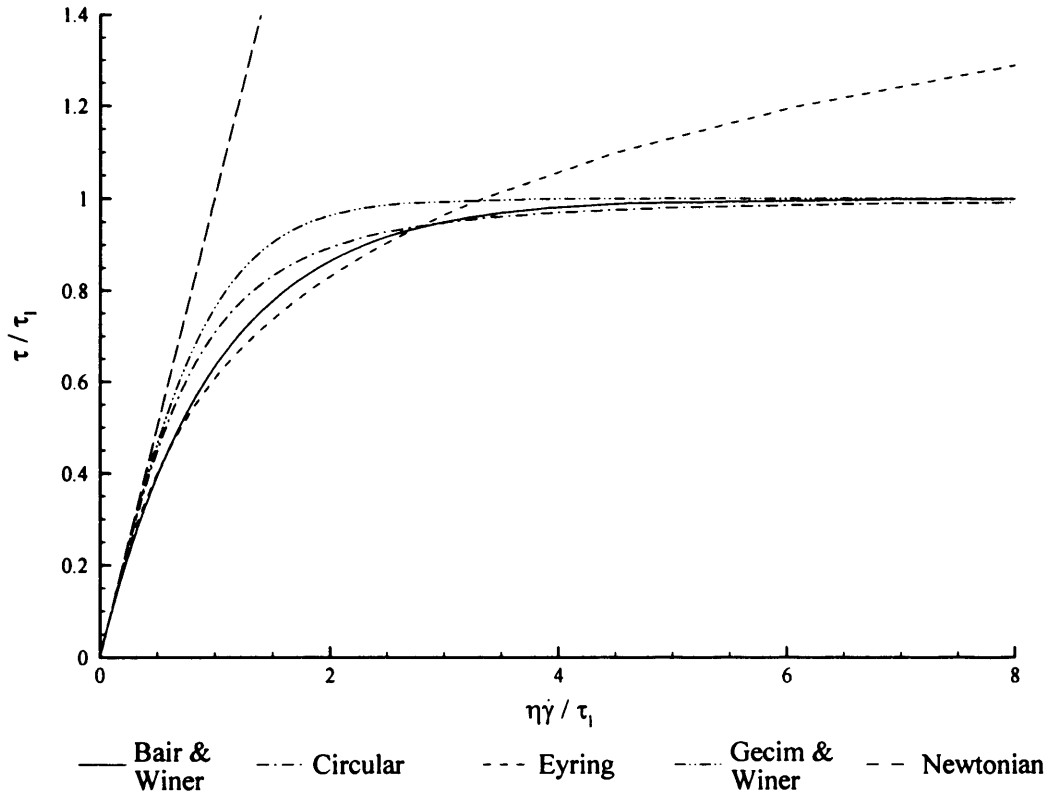

Figure 1.1 – Comparison of different rheological models

These four models, along with the Newtonian response, are shown in Figure 1.1. It can be seen that the circular model approaches the limiting shear stress value in a more progressive manner than the Gecim and Winer model. It is clear that the Gecim and

Winer model is more aggressive as it changes slope very rapidly. The shear thinning Eyring model is very gentle in comparison and does not reach a limiting value.

Using the circular model as a starting point, a general model was developed that would exhibit different limiting shear stress responses simply by altering one particular value, this is termed the Cardiff model and is presented in equation (1.5) below:

$$\dot{\gamma} = \frac{\tau_l}{\eta} \frac{\frac{\tau}{\tau_l}}{\left(1 - \frac{\tau^2}{\tau_l^2}\right)^m} \tag{1.5}$$

The value of m can be changed to give a variety of different responses. The value taken for m must be within the range $0 < m < 1$. It can be seen that for $m = 0.5$ the Cardiff model gives the same response as the Circular model.

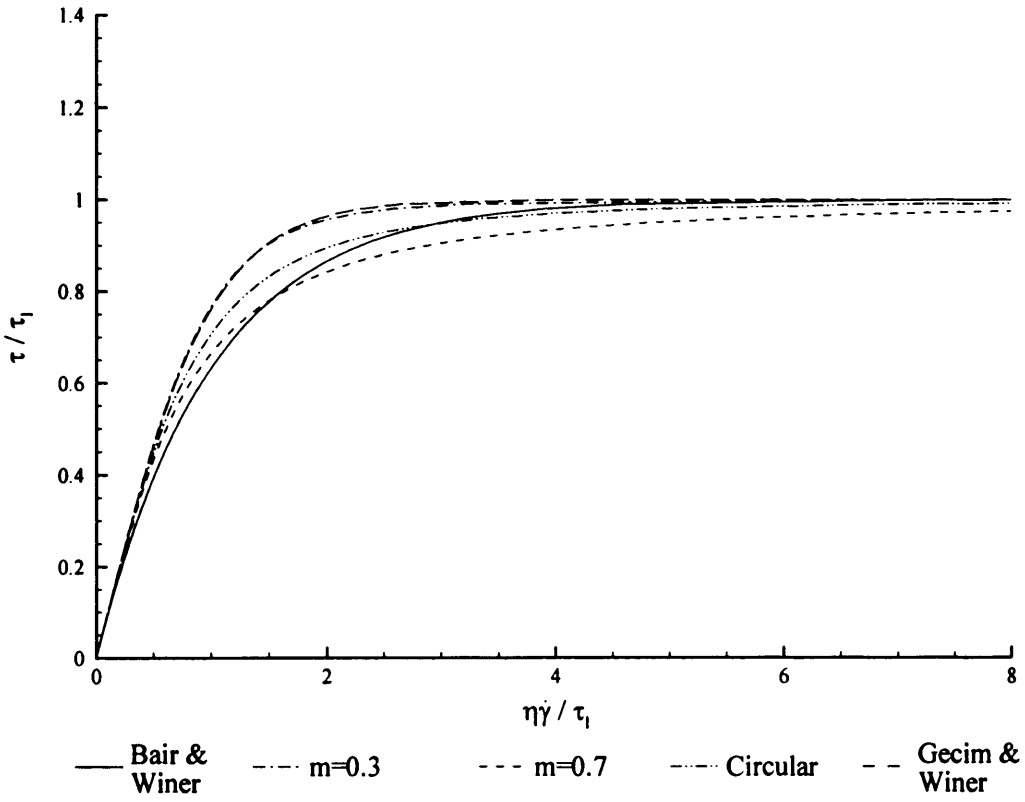

Figure 1.2 – Comparison of Cardiff model to Bair and Winer and Gecim and Winer rheological models

Figure 1.2 shows the effects that different values of m have on the response of the rheological model compared to the Bair and Winer and Gecim and Winer models. Taking lower values of m gives a sharper, more aggressive, transition from viscous to plastic behaviour, whereas larger values give a much more gradual transition. The $m = 0.3$ model gives a very similar response to the Gecim and Winer model, where both generate very aggressive transitions from viscous to plastic behaviour. The $m = 0.7$ model develops a more gradual transition from viscous to plastic behaviour. The word aggressive is used to describe the transition of the models from viscous to plastic behaviour. More aggressive is used to describe a sharper transition and less aggressive to denote a gradual transition.

1.8.2 Theoretical Results

Conry, Wang and Cusano (1987) presented a modified line contact Reynolds equation that took account of non-Newtonian effects based on the Eyring type model developed by Johnson and Tavaarwerk (1977). They also presented results for the Eyring model and presented numerical results that showed a reduction in film thickness when the slide roll ratio is increased.

Lee and Hamrock (1990) showed that for a limiting shear stress the pressure and film thickness experienced large changes as the slide roll ratio is increased, with the effects being far more pronounced than for the Eyring model. The work of Lee and Hamrock showed that the film thickness has a tendency to increase in the inlet to the contact which is accompanied by an increase in the pressure gradient. In the outlet region, the pressure gradient reduces as does the film thickness. This phenomenon was shown to a greater degree in Shieh and Hamrock (1991).

The influence of temperature on the Circular non-Newtonian model was studied by Hsiao and Hamrock (1992), it was shown that the film shape and pressure variation did not change with an increase in the slide roll ratio. The effects that the limiting shear stress had on the film and pressure were removed, essentially the same response was given as that obtained with the Eyring model.

Until very recently very little investigation has been undertaken into the models proposed by the experimental work of Bair and Winer (1979) and Gecim and Winer (1980). Ehret, Dowson and Taylor (1998) used the Bair and Winer rheological model to develop film thickness and pressure responses. Their work compared the limiting shear stress response to the Eyring shear thinning model. The results shown for the isothermal limiting shear stress model gave similar trends to those presented by Shieh and Hamrock (1991).

Very little work has been undertaken to estimate what happens once the limiting shear stress is reached within the film. Jacobson and Hamrock (1984) discussed differing areas within the contact in which slip would occur and formulated different equations for each area. This is the only work to date that has covered slip occurring when the shear stress exceeds a nominally limiting value. Other authors [Ehret, Dowson and Taylor, 1998] have considered a slip parameter and developed formulations based on slip occurring as a function of pressure.

As the interest into limiting shear stress rheology grew, there was a corresponding increase in the importance placed upon the variation of the limiting shear stress under different conditions. Gecim and Winer (1980) presented a relationship for the limiting shear stress indicating that it has an initial value, τ_{l0}, and is then seen to increase, at the rate α_τ, with pressure:

$$\tau_l = \tau_{l0} + \alpha_\tau p$$

Evans and Johnson (1986) conducted a series of experimental investigations into oil properties and it can be seen that the same trend is seen as that presented above. More recently, Larsson et al. (2000) confirmed this relationship in an investigation into the shear behaviour of a number of lubricants.

1.9 Failure Modes of EHL Contacts

All the work mentioned thus far was undertaken in an attempt to understand the fluid films that are developed within EHL contacts. Components that rely on EHL to operate

also suffer failures, some of which are due to design and external factors and others that are related to the breakdown in the EHL mechanism.

Two of the least understood and problematic forms of failure in the EHL contacts occurring in gears and other lubricated contacts are micro-pitting and scuffing. The two failures are very different in the way in which they arise and are discussed in the following sections.

1.9.1 Scuffing

Scuffing is a dramatic failure which can be catastrophic, occurring due to the breakdown of the lubricant film that results in cold welding of the two surfaces, generally occurring early on in the component's life. The cold welding of the surfaces gives rise to a torn, dragged or furrowed appearance to the surfaces. This can be clearly seen in Figure 1.3, where a SEM image of a scuffed region is shown.

Figure 1.3 – SEM image of the scuffed region of a test disk

The scuffing region in figure 1.3 occurs at the edge of the contact region, this is generally seen in disk tests as the edge of the contact region is where the films are thinnest. The area to the left of the scuffed region has not come into contact with another surface and thus is as manufactured. The surface to the right of the scuffed region has run against another surface and has rounded asperities due to the interaction with another surface.

The surfaces rip material off one another giving rise to a dramatic increase in friction and temperature as well as an increase in the noise and vibration in the system. Within gears, scuffing is usually seen at the tip and root of gear teeth where sliding is high. It can be recognised as grooves of varying depth and width which are aligned in the sliding direction.

Scuffing only occurs in systems where the velocity of the two surfaces yields a relative sliding velocity. In systems where there is no relative sliding the components fail mainly due to pitting. When scuffing is initiated the surfaces change rapidly creating more heat, which can lead to further scuffing, possibly resulting in a catastrophic failure. Scuffing leaves hard, white and etch resistant phases separated from the bulk material by a tempered layer.

There are many different hypotheses regarding the conditions that bring about scuffing, however, there is currently no mechanism that is widely adopted and accepted. The various models have related the temperature as well as the frictional power intensity to the onset of scuffing. A components' resistance to scuffing can be improved by modifying the surfaces, making them less rough, and also by the addition of coatings. Different oils can be used that have EP (Extreme-Pressure) additives which help prevent the occurrence of scuffing.

1.9.2 Running In

One of the methods used to reduce the occurrence of scuffing and to improve general performance, that has been investigated widely by tribologists, is to pre-run the surfaces, termed 'running in'. 'Running in' is used to improve the surface topography of the

components that are running against one another, resulting in reduced friction between the components. Run in components can be seen to carry an increased load at the same operating conditions compared to the performance before running in, as shown by Ostvik and Christensen (1968).

'Running in' is usually performed at lower loads and sometimes lower speeds, which are then slowly built up to the operating conditions in stages. During the process of running in the peaks of the asperities are removed, metallurgical processes such as strain hardening occur and oxide layers form that help protect the surfaces. The approach used for running in is based on best engineering practice rather than application of a definitive equation.

More recently, 'running in' has been conducted with different oil to that with which the parts are designed to run with. This allows more conformal surfaces to be generated that provide greater resistance to scuffing [Paliwal, 1987].

1.9.3 Pitting

Pitting is a fatigue effect and does not necessarily occur early in the component's life, unlike scuffing. Pits occur at high points on the surface which are subjected to high stresses. These points are sheared due to the interaction with other asperities creating residual stresses as a result of the shearing which tends to promote cracks which propagate thus removing the high points in an erosive fashion.

Pitting tends to stop after a period of time, as the removal of material re-distributes the load leading to reduced surface stresses. However, in some cases large pits are produced, termed destructive pitting. Destructive pitting tends to primarily occur in the dedendum section of gear teeth due to overload conditions, which do not arrest like normal pitting. The edges of pits can break and form large irregular shapes that can adjoin one another. This type of surface damage is called spalling and occurs due to high contact stresses. Spalling and destructive pitting can be prevented and stopped by increasing the hardness of the surface and by reducing the operating load. It should be noted that pitting can also

occur in the region of impurities or inclusions that are weaker than the surrounding material.

Of greater concern of late is the occurrence of micro-pits. Unlike normal pits which are of the order of 0.3mm or more in diameter, micro-pits are of the scale of the asperities and their diameters tend to measure of the order of microns. Micro-pitting is believed to occur due to high loads and temperatures that create thin films. They are not visible to the naked eye, like normal pitting, and tend to be seen as a grey staining or frosting appearance on the surface. Micro-pitting occurs in the same manner as pitting and is also a fatigue effect, where the asperities themselves tend to have cracks form in them which later form a pit. Micro-pitting can sometimes occur and then be seen to disappear due to the polishing of the surfaces. In future work, it is hoped that the detailed study of the EHL of rough surfaces will help generate a greater understanding of micro-pitting.

1.10 Thesis Aims and Layout

The aims of the thesis were to incorporate limiting shear stress rheology into an existing approach to solve the line contact EHL problem. This was found to require the inclusion of the full energy equation which was adapted to include the heating generated by slip. This led to an approach that allows the heat generated when the two surfaces contact to be determined.

Other aims were to adapt the software to allow gear meshing cycles to be simulated including the actual roughness of the gear teeth. Finally, experimental work was undertaken to investigate the process of running in through monitoring the surface modification and the amount of metallic contact that occurred.

The work undertaken is divided between the following chapters:

The basic equations that define the EHL line contact problem as well as the elastic deflection, hydrodynamic and energy equations are presented in *Chapter 2.* Included are

the equations that describe the behaviour of the gear mesh and the relative geometry of the teeth during the meshing cycle.

An approach is presented in *Chapter 3* that allows the non-Newtonian rheology to be incorporated into the hydrodynamic equation. The equations introduced relate to the Eyring shear stress model as well as the limiting shear stress models, also introduced is the concept of slip.

The equations presented in Chapter 3 are investigated in *Chapter 4* to determine how the functions controlling the non-linear rheological behaviour change as the operating conditions vary. The approaches that are used to solve the equations that define the limiting shear stress models are detailed as well as the approaches that are used to solve the energy equation and the coupled hydrodynamic and elastic deflection equations.

Comparisons of the limiting shear stress models to the Eyring shear thinning model are made in *Chapter 5*. The comparisons are made under isothermal conditions and the operating parameters are adjusted to determine the sensitivity of the models to the different parameters.

The limiting shear stress and Eyring models are compared under thermal conditions in *Chapter 6*. The isothermal and thermal results are compared before a brief investigation of the effect that speed and slide roll ratio have under operating conditions that relate to optical interferometry experiments.

The approach used to calculate the full energy equations is refined as the problem of surface contact is considered in *Chapter 7*. Results are also presented into the influence that the coefficient of friction has on the pressure, film thickness and temperature near contact events.

Chapter 8 presents the manner in which the gear meshes are simulated, including the approach that is used to include the actual roughness of the gear teeth into the simulation. Smooth results are presented for three different gear meshes and rough surface results are attained for two of the meshes.

The experimental work undertaken is detailed in *Chapter 9*. The modifications necessary to acquire the contact resistance and to measure the surfaces of the disks are presented. The experimental approach is also outlined along with the development of the computer data acquisition system. Results are presented for two experiments, one involving a rough disk running against a smooth disk, and another that covers two rough disks running against one another. The results also include numerical simulations of the conditions that allow a quantitative comparison of the results.

The conclusions that can be drawn from the work and the recommendations for future work are presented in *Chapter 10*.

Chapter 2

Basic Theory

2.0 Introduction

This section covers the mathematical equations that specify the line contact EHL problem. A characteristic feature of EHL is that a high load is applied over a small contact area, this produces pressures that are high enough to cause significant elastic deflection of both surfaces as well as to change the oil's properties as it flows through the contact. Two equations, the elastic deflection equation and the Reynolds hydrodynamic equation, specify the problem and have to be solved simultaneously. The two equations link pressure and film thickness, the hydrodynamic equation links pressure with film thickness with respect to the fluid and the elastic deflection equation is concerned with the relationship for the bounding surfaces.

Design aspects usually dictate the load, oil properties, rotational speed, finish of contacting surfaces, geometry and mechanical properties, and this information is essential to construct the line contact problem. Due to the nature of the Reynolds equation, all contacts can be simplified so that the curvature is taken up by one surface, essentially the problem then reduces to the contact between a flat plane and a parabolic surface.

2.1 Contact Geometry

As mentioned previously the complicated geometry of each surface can be simplified to the contact between a parabola and a flat plate. The parabola incorporates the curvature of both surfaces and is two dimensional for the line contact problem and three dimensional for the point and elliptical contact problems.

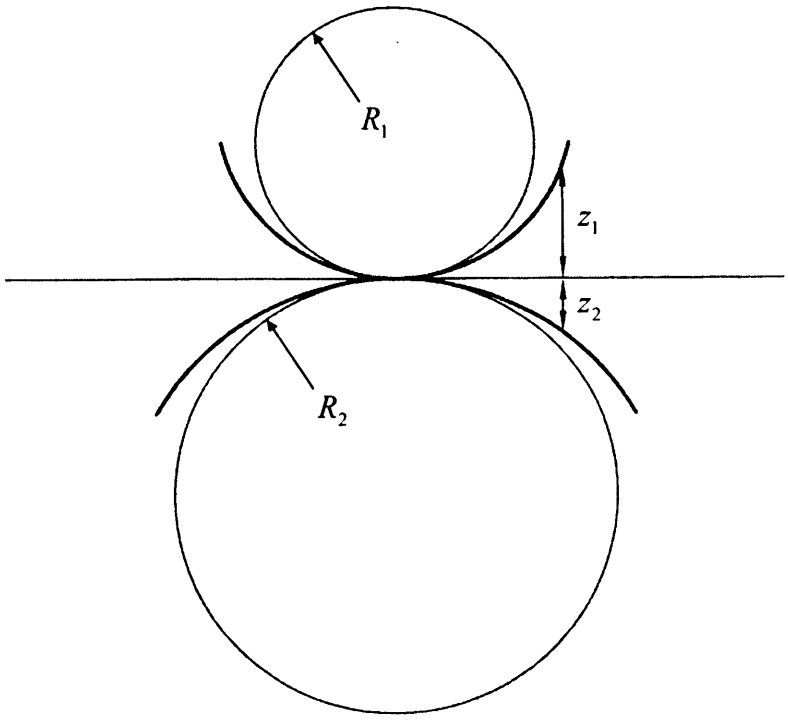

Figure 2.1 – Two cylindrical rollers and respective parabolas

Figure 2.1 shows the contact between two cylindrical rollers as well as parabolas that represent the curvature of each roller. The separation of each surface from the contact plane is given by:

$$z_1 = \frac{x^2}{2R_1} \qquad \text{and} \qquad z_2 = \frac{x^2}{2R_2}$$

The total separation or undeformed geometry, $h_u(x)$, is given by:

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

Using a parabola and a flat plane shown in Figure 2.2 the undeformed geometry can be determined as:

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

Where R is the radius of relative curvature determined from:

$$\frac{1}{R} = \frac{1}{R_1} + \frac{1}{R_2} \tag{2.2}$$

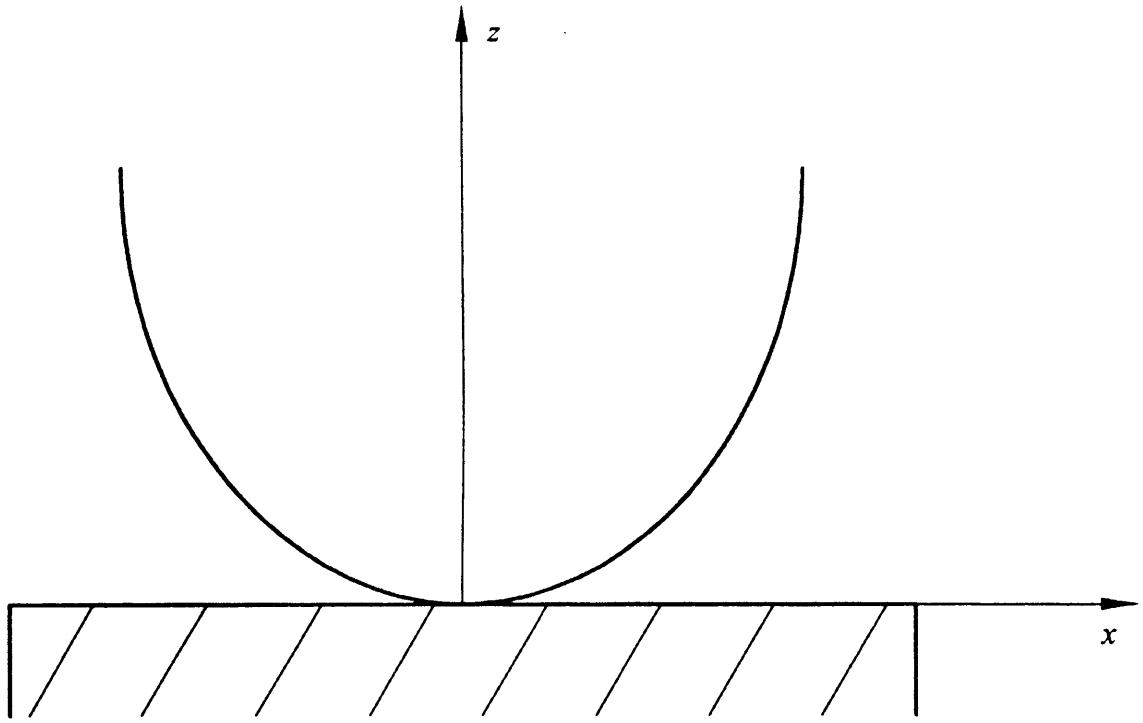

Figure 2.2 – Simplified contact of parabola and a flat plate

2.2 Gear Contacts

The lubricated contact behaviour of gear teeth as they transmit load is an example of an EHL situation and is the main application focus for the current work. For spur gears it can be modelled as a line contact problem. A detailed analysis of such contacts also presents a significant challenge, as the contact geometry and operating conditions are constantly changing through the meshing cycle. The contact geometry of gears is difficult to determine for two reasons. Firstly, the gear teeth are generated from an involute profile and have a constantly changing radius of curvature, and secondly, the contact point is constantly moving along the tooth. Before the geometry can be determined the way in which gears contact must be understood.

The contact between two meshing involute gear teeth progresses along a straight line that is tangential to the base circles of both gears as shown in Figure 2.3. The angle of inclination of this line, termed the line of action, to the line perpendicular to the line of centres A_1A_2 is the pressure angle ψ (which is constant for teeth of involute form). Dimension s is the distance of the instantaneous point of contact from the pitch point and

is a measure of how far along the line of action the contact occurs. It is zero at the pitch point and positive in the direction of motion of the gear teeth.

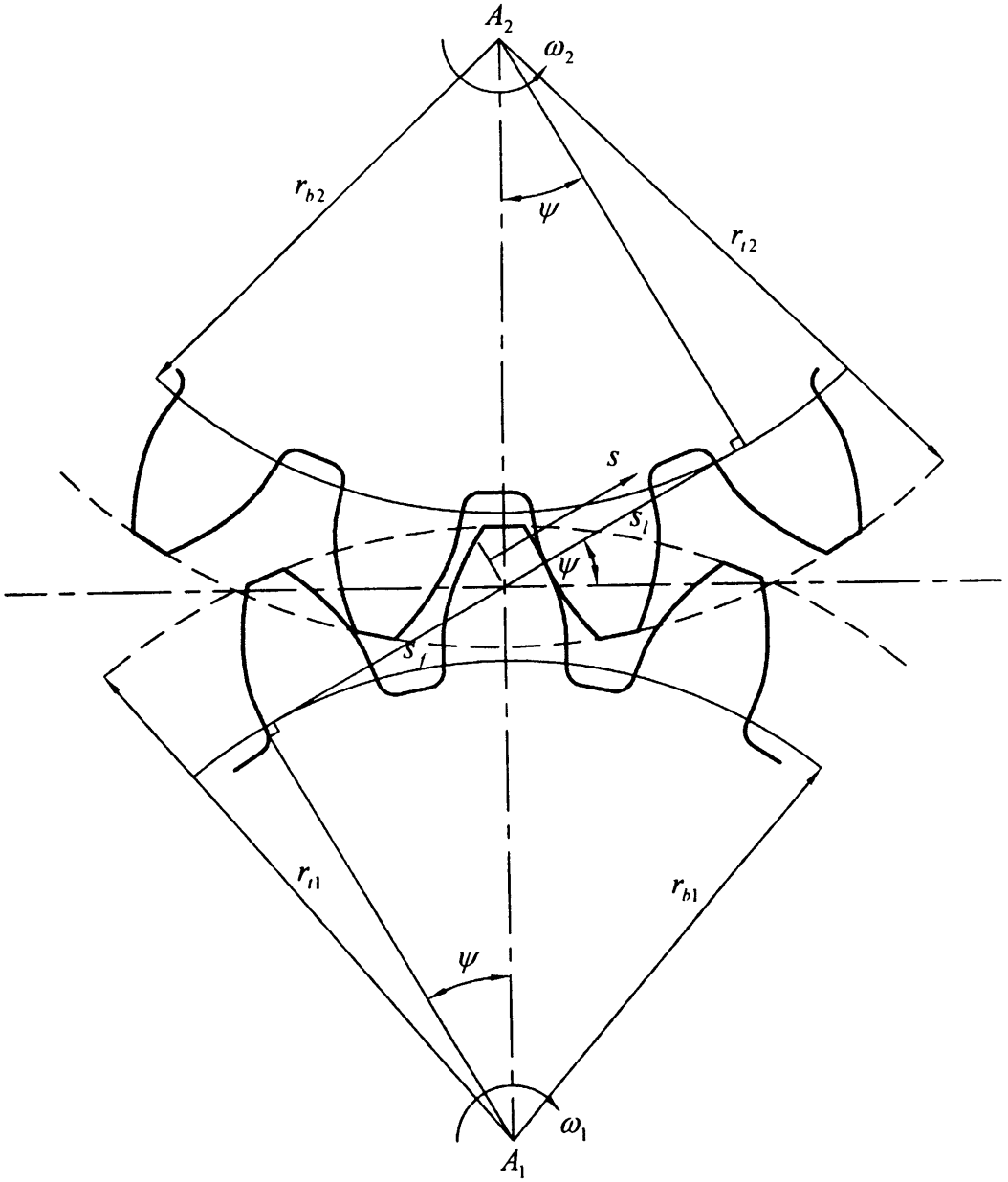

Figure 2.3 – Contact path of meshing gears

There are two sets of subscripts used in figure 2.3, subscript 1 refers to the driving gear, which is the lower component, and subscript 2 is used to define the driven gear, the upper component. The other subscript is used to differentiate between the base circles, subscript b, and the tip radius, subscript t.

The first and last contact points can be determined from the base radius, tip radius and pressure angle. These individual parts are outlined in Figure 2.4 along with the line of action. The lengths PB and TB are distances along the line of action in the positive s direction defined by:

$$PB = r_{b2} \tan\psi \qquad \text{and} \qquad TB = \sqrt{r_{t2}^{\,2} - r_{b2}^{\,2}}$$

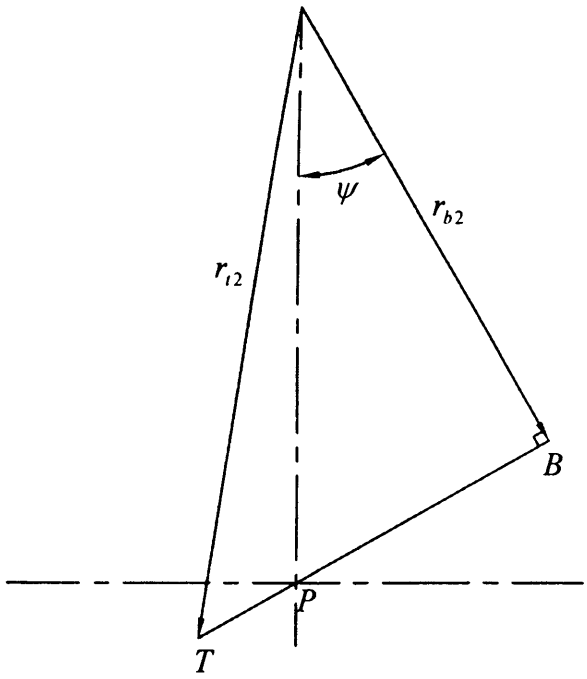

Figure 2.4 – Simplified geometry for driven gear

The first contact point, s_f, is determined from these lengths as:

$$s_f = PB - TB$$

$$s_f = r_{b2} \tan\psi - \sqrt{r_{t2}^{\,2} - r_{b2}^{\,2}}$$

Similarly the last contact point, s_l, is determined from:

$$s_l = \sqrt{r_{t1}^{\,2} - r_{b1}^{\,2}} - r_{b1} \tan\psi$$

The point of contact moves along the line of action between s_f and s_l at constant velocity in a similar manner to a belt in a pulley system. This velocity at which the contact point moves, $v_{contact}$, is determined by:

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

2.2.1 Equivalent Rollers

Consider the meshing of two gear teeth illustrated in Figure 2.5.

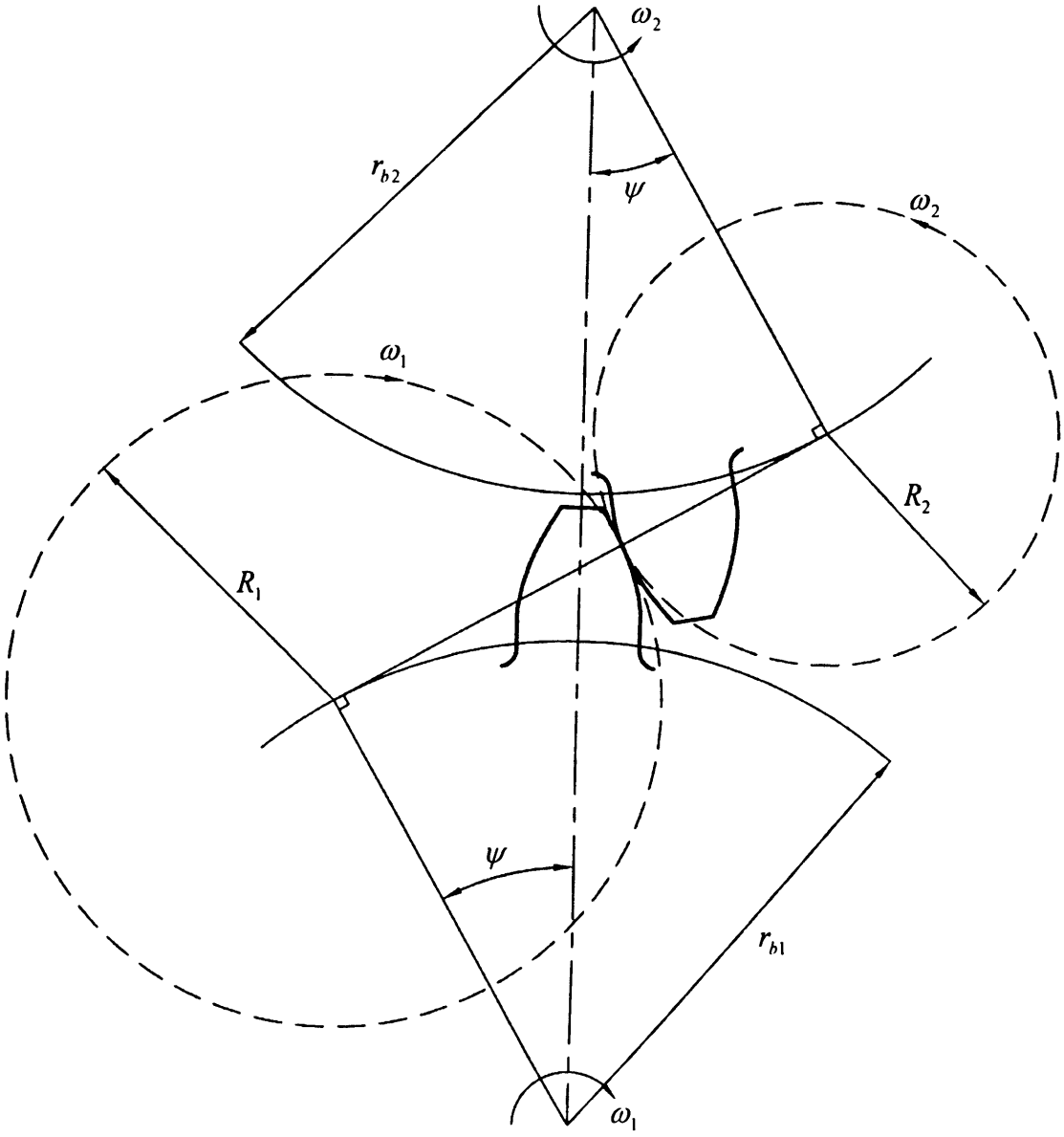

Figure 2.5 – Simplified contact geometry of involute teeth

The gear teeth shown in figure 2.5 are generated so that the tooth profile is of involute form. The radius of curvature of any point on the involute is the distance from that point to a tangential point on the base circle. Since the contact dimension is small the curvature change within the contact due to the involute shape should be very small. Therefore a reasonable assumption is to use the radius of curvature of the point at the centre of the contact allowing the radius of two equivalent rollers, R_1 and R_2, to be determined as:

$$R_1 = r_{b1} \tan\psi + s$$

$$R_2 = r_{b2} \tan\psi - s$$

These can be used in equation (2.1) and (2.2) to determine the undeformed geometry. This approach was first introduced by Martin (1916).

2.2.2 Involute Profile

An involute is generated by a particle on the end of a string that is unwound from a circle, the circle for gear teeth is the base circle. The involute profile of a gear tooth can be determined on an axis set at the origin, O, of the involute with coordinates X_g and Y_g tangential and perpendicular to the base circle respectively, which are shown in Figure 2.6. γ is the angle through which the string has been unwound for the point C on the involute profile, this gives BC as $r_b\gamma$. The coordinates of C are:

$$X_g = -AB\sin\gamma + BC\cos\gamma$$

$$Y_g = -AO + AB\cos\gamma + BC\sin\gamma$$

Substituting for the lengths OA, AB and BC:

$$X_g = (-r_b)\sin\gamma + (r_b\gamma)\cos\gamma$$

$$Y_g = (-r_b) + (r_b)\cos\gamma + (r_b\gamma)\sin\gamma$$

Which simplifies to:

$$X_g = r_b\gamma\cos\gamma - r_b\sin\gamma \qquad (2.3)$$

$$Y_g = r_b(\cos\gamma - 1) + r_b\gamma\sin\gamma \qquad (2.4)$$

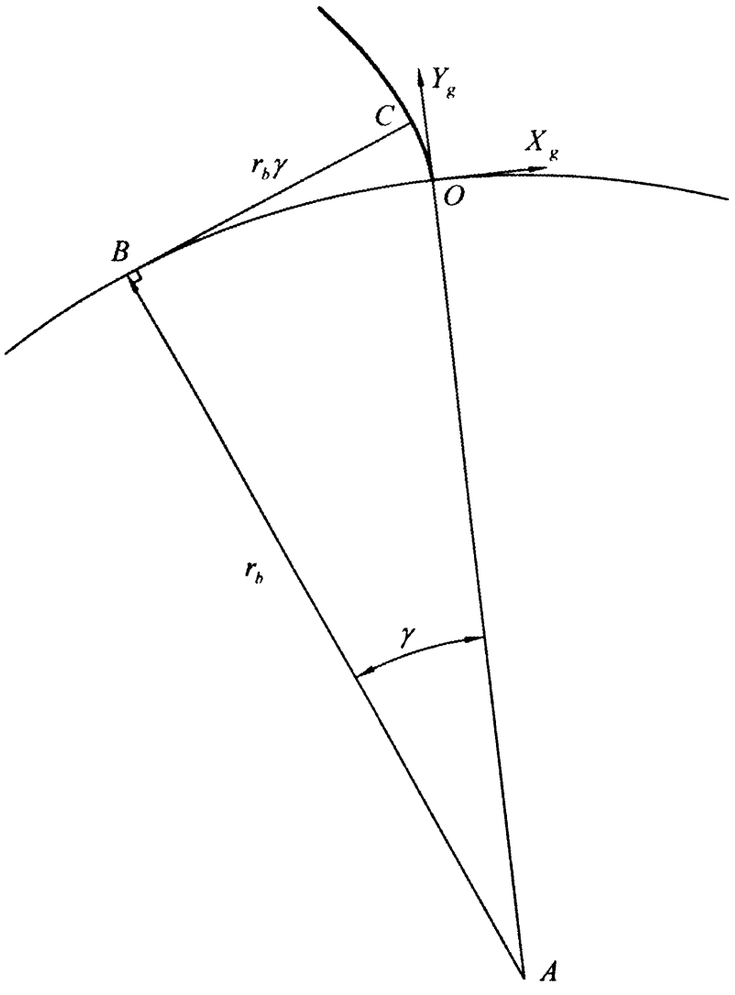

Figure 2.6 – Orientation of axis used to describe involute

The axes for the EHL problem are orientated parallel and perpendicular to the tangent plane at the contact point, the x axis is positive in the direction of rotation and the z axis is positive towards the centre of the tooth. Figure 2.7 illustrates the two sets of axes on the driven gear. γ is the angle through which the string has been unwound for a point of interest and γ_c defines the angle unwound to the contact point C, which is the angle between the two axis sets. Defining the point at which contact is occurring on the axis set at the origin of the involute as X_{gc} and Y_{gc}, the two axis sets can be related by:

$$x = (Y_g - Y_{gc})\cos\gamma_c - (X_g - X_{gc})\sin\gamma_c$$

$$z = -(Y_g - Y_{gc})\sin\gamma_c - (X_g - X_{gc})\cos\gamma_c$$

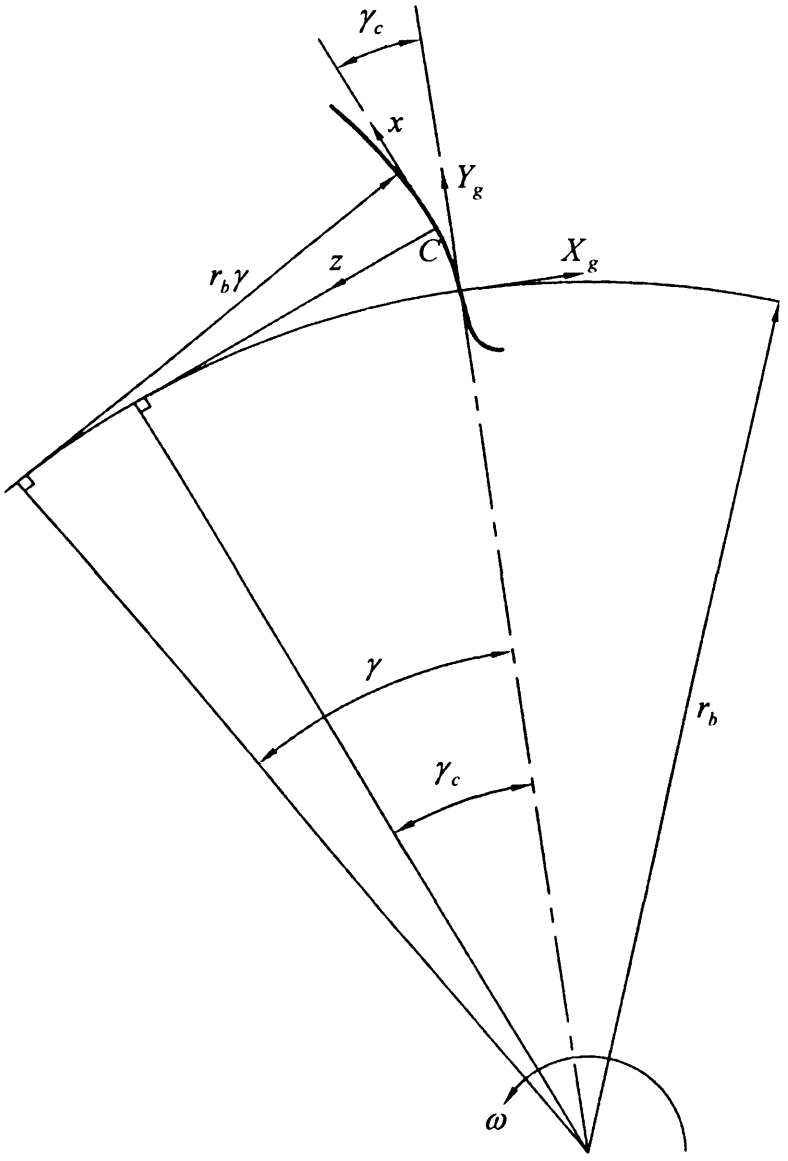

Figure 2.7 – Relation between the EHL and involute axes

Substituting equations (2.3) and (2.4) into the above gives:

$$\begin{aligned} x = {} & r_b \cos\gamma_c(\cos\gamma + \gamma\sin\gamma - \cos\gamma_c - \gamma_c \sin\gamma_c) \\ & + r_b \sin\gamma_c(\sin\gamma - \gamma\cos\gamma - \sin\gamma_c + \gamma_c \cos\gamma_c) \end{aligned} \tag{2.5}$$

$$\begin{aligned} z = {} & r_b \sin\gamma_c(\cos\gamma_c + \gamma_c \sin\gamma_c - \cos\gamma - \gamma\sin\gamma) \\ & + r_b \cos\gamma_c(\sin\gamma - \gamma\cos\gamma - \sin\gamma_c + \gamma_c \cos\gamma_c) \end{aligned} \tag{2.6}$$

Re-expressing equations (2.5) and (2.6) with the relevant subscripts for the driven and driving gears gives:

$$\begin{aligned} x = {} & r_{b2}\cos\gamma_{c2}(\cos\gamma_2 + \gamma_2\sin\gamma_2 - \cos\gamma_{c2} - \gamma_{c2}\sin\gamma_{c2}) \\ & + r_{b2}\sin\gamma_{c2}(\sin\gamma_2 - \gamma_2\cos\gamma_2 - \sin\gamma_{c2} + \gamma_{c2}\cos\gamma_{c2}) \end{aligned} \tag{2.7}$$

$$\begin{aligned} z_2 = {} & r_{b2}\sin\gamma_{c2}(\cos\gamma_{c2} + \gamma_{c2}\sin\gamma_{c2} - \cos\gamma_2 - \gamma_2\sin\gamma_2) \\ & + r_{b2}\cos\gamma_{c2}(\sin\gamma_2 - \gamma_2\cos\gamma_2 - \sin\gamma_{c2} + \gamma_{c2}\cos\gamma_{c2}) \end{aligned} \tag{2.8}$$

Equations (2.7) and (2.8) are only applicable to the driven gear. Similar equations can be established for the driving gear, where the meshing tooth is orientated opposite to the driven tooth portrayed in figure 2.7. The change in orientation of the tooth has no effect on the x axis, but due to the change in orientation of the gear tooth, it is positive towards the root. It does however cause the positive z axis to change direction from that shown in figure 2.7, so that the z axis is still positive toward the centre of the tooth. Equations corresponding to equations (2.7) and (2.8) are obtained as:

$$\begin{aligned} x = {} & r_{b1}\cos\gamma_{c1}(\cos\gamma_{c1} + \gamma_{c1}\sin\gamma_{c1} - \cos\gamma_1 - \gamma_1\sin\gamma_1) \\ & + r_{b1}\sin\gamma_{c1}(\gamma_1\cos\gamma_1 - \sin\gamma_1 - \gamma_{c1}\cos\gamma_{c1} + \sin\gamma_{c1}) \end{aligned} \tag{2.9}$$

$$\begin{aligned} z_1 = {} & r_{b1}\sin\gamma_{c1}(\cos\gamma_{c1} + \gamma_{c1}\sin\gamma_{c1} - \cos\gamma_1 - \gamma_1\sin\gamma_1) \\ & + r_{b1}\cos\gamma_{c1}(\sin\gamma_1 - \gamma_1\cos\gamma_1 - \sin\gamma_{c1} + \gamma_{c1}\cos\gamma_{c1}) \end{aligned} \tag{2.10}$$

It should be noted that x has no subscript, this is due to the fact that x is common to both teeth. As the meshing cycle progresses γ_{c1} and γ_{c2} change, but at each point they can be determined from the gear geometry as:

$$\gamma_{c1} = \frac{r_{b1}\tan\psi + s}{r_{b1}} \qquad \text{and} \qquad \gamma_{c2} = \frac{r_{b2}\tan\psi - s}{r_{b2}}$$

The reference angles γ_{c1} and γ_{c2} can be determined from the above equations for any value of s. These reference angles are used in equations (2.7) and (2.9) with x, which is known for each mesh point, allowing the unknown angles γ_1 and γ_2 to be determined. The angle γ_1 and the reference angle γ_{c1} are used to determine z_1 from equation (2.8), similarly γ_2 and γ_{c2} are used in equation (2.10) to determine z_2. The two distances z_1 and z_2 can be combined to give the undeformed geometry as:

$$h_u(x) = z_1 + z_2$$

2.3 Hertzian Contact Theory

The pressure distribution generated within an EHL contact is similar to that formed by the equivalent dry contact. The dry contact analysis is based on two smooth, cylindrical, stationary solids being pressed together. This was first solved by Hertz (1881) and is used as a starting point for EHL studies.

The theory considers two rollers and simplifies the contact using the radius of relative curvature, R, defined by equation (2.2). The load applied is the load per unit length (in the y direction) w' and the resulting contact area measures $2a$ in the x direction. These are shown in Figure 2.8.

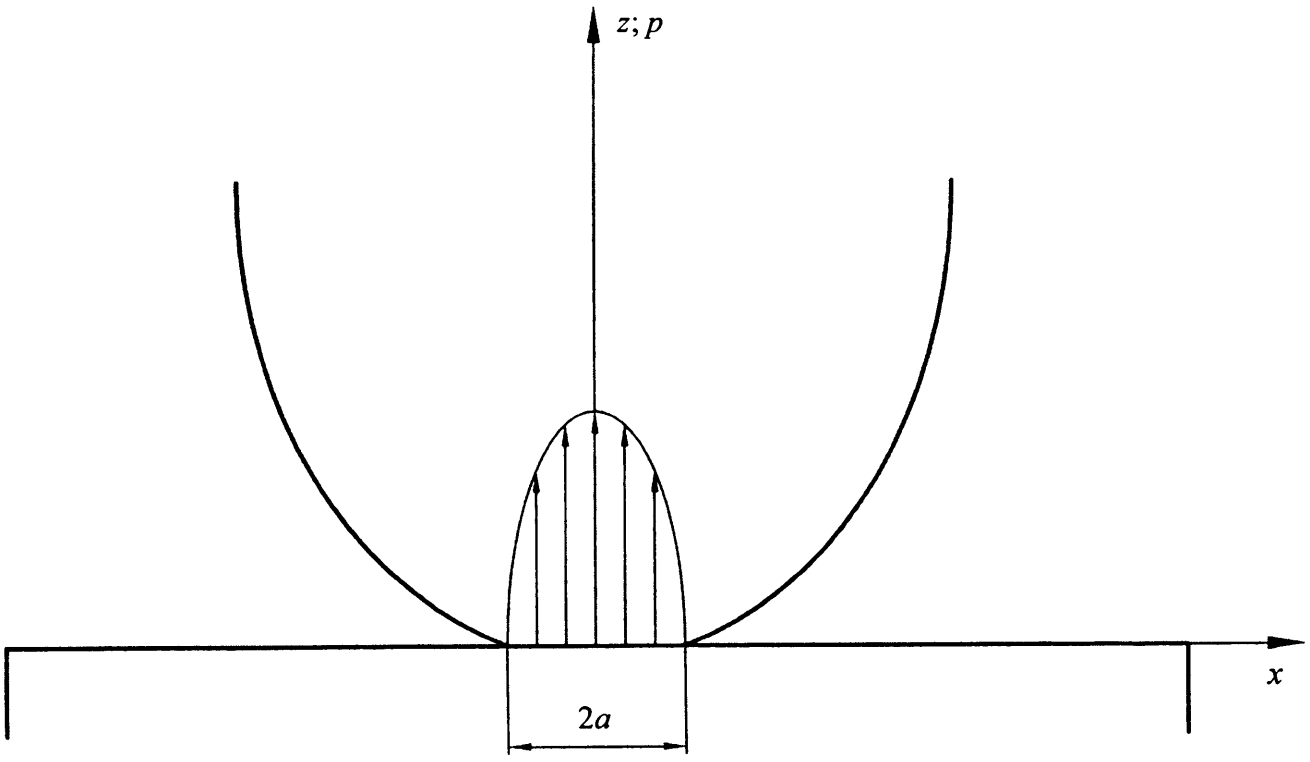

Figure 2.8 – Contact between parabola and flat rigid plate showing Hertzian pressure and contact area

The contact dimension a is given by:

$$a = \sqrt{\frac{8}{\pi}\frac{Rw'}{E'}} \tag{2.11}$$

Where the reduced elastic modulus, E', is taken from:

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

The pressure distribution that is developed is semi-elliptic, as shown in figure 2.8, and is given by:

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

The maximum pressure p_0 is given by:

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

Substituting equation (2.11) into the above gives:

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

2.4 Elastic Film Thickness Equation

The elastic film thickness that is linked with the hydrodynamic equation consists of the undeformed geometry, $h_u(x)$, the deflection, $\Lambda(x)$, and a separation term, h_s. These constituent terms are shown in Figure 2.9 and can be linked by the following equation:

$$h(x) = \Lambda(x) + h_u(x) + h_s \qquad (2.12)$$

The undeformed geometry can be determined from equation (2.1), however, for the case of gear teeth it is slightly more complex and the manner in which it is determined was covered in section 2.2.2. The separation term is used to offset the undeformed geometry, it is a constant that is used to determine the required load (and will be negative for heavily loaded cases). The deflection term is more difficult to determine than the previous two quantities and the approach is outlined as follows.

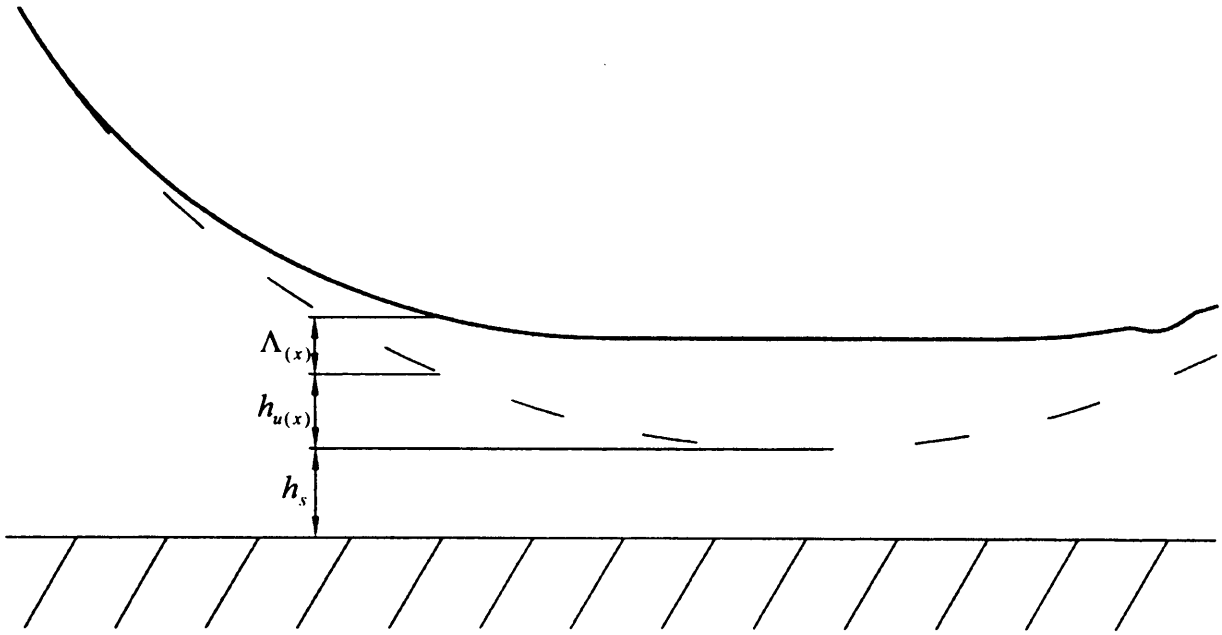

Figure 2.9 – Constituent parts of the elastic film thickness

The deflection solution method that is implemented is the differential deflection method as it reduces the influence of points further away from the point of interest, essentially requiring fewer points in the solution scheme, and making the solution quicker and less complex. The differential deflection method that is used is based around the second differential of deflection and was first presented by Evans and Hughes (2000).

The deflection equation used is based on the deflection of a semi-infinite body as given by Johnson (1985):

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

This gives the deflection of the point of interest, x, relative to the point where $x = r$ and can be simplified to the following form:

$$\Lambda(x) = -\frac{4}{\pi E'}\int_a^b p(s)\ln|x-s|ds + Const \tag{2.13}$$

The major complication in determining the deflection is that the convolution function $\ln|x-s|$ is singular about the point at which deflection is being calculated. This can be overcome by taking the limits of the function on either side of the singularity:

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

Differentiating with respect to x:

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

Taylor's expansion can be used to approximate the two terms that are not in the form of an integral as:

$$\frac{d\Lambda(x)}{dx} = -\frac{4}{\pi E'}\lim_{\delta\to 0}\left\{ \begin{array}{l} \ln(\delta)\left[p(x) - \delta\frac{dp(x)}{dx} + \frac{\delta^2}{2}\frac{d^2p(x)}{dx^2}..... \right] + \int_a^{x-\delta} \frac{p(s)}{x-s}ds \\ -\ln(\delta)\left[p(x) + \delta\frac{dp(x)}{dx} - \frac{\delta^2}{2}\frac{d^2p(x)}{dx^2}..... \right] + \int_{x+\delta}^{b} \frac{p(s)}{x-s}ds \end{array} \right\}$$

Applying the limit:

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

Differentiating again to give the second derivative:

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

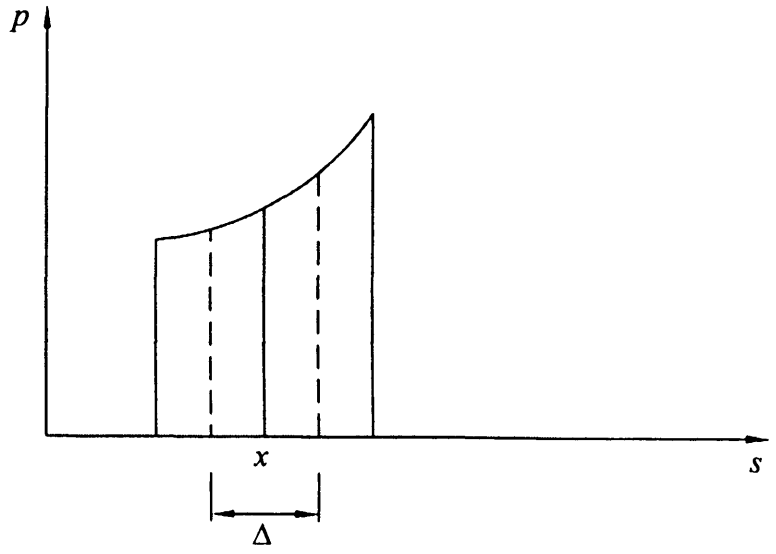

Figure 2.10 – Variation of pressure at node points with quadrature mesh centred at point of interest

Introducing a quadrature mesh, with a spacing Δ, and centring the mesh at the point of interest, as shown in Figure 2.10, the following can be established:

$$\frac{d^2\Lambda(x)}{dx^2} = -\frac{4}{\pi E'}\left\{\begin{array}{c} -\int_a^{x-\frac{\Delta}{2}} \frac{p(s)}{(x-s)^2}ds - \int_{x+\frac{\Delta}{2}}^{b} \frac{p(s)}{(x-s)^2}ds \\ +\lim_{\delta\to 0}\left[\frac{p(x-\delta)}{\delta} - \int_{x-\frac{\Delta}{2}}^{x-\delta} \frac{p(s)}{(x-s)^2}ds + \frac{p(x+\delta)}{\delta} - \int_{x+\delta}^{x+\frac{\Delta}{2}} \frac{p(s)}{(x-s)^2}ds\right] \end{array}\right\} \quad (2.14)$$

Introducing a new variable, $z = s - x$, and assuming the pressure over the range $-\frac{\Delta}{2} < z < \frac{\Delta}{2}$ can be evaluated by:

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

The limit in equation (2.14) can be evaluated giving:

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

Using the quadrature in figure 2.10 it can be seen that $z = 0$ is the point where $s = x$. Introducing i as a positive or negative integer then $z = i\Delta$ is the centre of the ith interval. This allows the integrals to be simplified giving the following form of the second order differential deflection:

$$\frac{d^2\Lambda(x)}{dx^2} = \frac{4}{\pi E'}\left[\sum_{\substack{all\ i \\ i\neq 0}} \int_{(i-0.5)\Delta}^{(i+0.5)\Delta} \frac{p(x+z)}{z^2}dz - \frac{4p(x)}{\Delta} + \frac{\Delta}{2}\frac{d^2p(x)}{dx^2}\right] \quad (2.15)$$

If the point of interest is outside the limits a and b, then equation (2.15) does not need to be used as there is no singularity in the integral. Instead this only requires the evaluation of equation (2.13) based on the quadrature used in figure 2.10.

In the case of rough surfaces, the composite surface roughness profile, $\phi(x)$, can be incorporated into equation (2.12) so that it becomes:

$$h(x) = \nabla(x) + h_u(x) + \phi(x) + h_s$$

Considering a parabolic undeformed geometry, equation (2.1) can be substituted into the above equation giving:

$$h(x) = \nabla(x) + \frac{x^2}{2R} + \phi(x) + h_s$$

This can be differentiated with respect to x twice to give:

$$\frac{d^2h(x)}{dx^2} = \frac{d^2\nabla(x)}{dx^2} + \frac{1}{R} + \frac{d^2\phi(x)}{dx^2} \tag{2.16}$$

It was shown by Evans and Hughes (2000) that the differential deflection equation, equation (2.15), can be represented in the form:

$$\frac{d^2\nabla(x_i)}{dx^2} = \sum_{all\ k} f_{k-i} p_k$$

Substituting in equation (2.16) gives:

$$\frac{d^2h(x_i)}{dx^2} = \sum_{all\ k} f_{k-i} p_k + \frac{1}{R} + \frac{d^2\phi(x)}{dx^2} \tag{2.17}$$

Equation (2.17) gives the second differential of deflection and it can be used in a central difference scheme to determine the elastic deflection relationship between film thickness and pressure in the computational mesh. This is covered in more detail in section 4.5.

2.5 Viscosity Pressure Relationship

Within the contact the pressure increases to the order 1GPa and the viscosity of the oil increases by several orders of magnitude. There are two relationships that are commonly used to take account of this behaviour. They are the Barus (1893) relationship:

$$\eta = \eta_0 e^{\alpha p}$$

and the isothermal Roelands (1966) relationship:

$$\eta = \eta_0 e^{\ln(\eta_0)+9.67\left[\left(1+5.1\times10^{-9}p\right)^{Z}-1\right]}$$

To ensure that the two equations give similar behaviour at low pressure, the parameter Z can be related to the viscosity pressure coefficient, α. This can be achieved by equating the slopes of $\ln(\eta)$ at $p = 0$ which gives:

$$Z = \frac{\alpha}{5.1\times10^{-9}\left[\ln(\eta_0)+9.67\right]}$$

The coefficients α and η_0 are properties of any given oil and vary with temperature. Z is the Roelands pressure viscosity parameter and is generally 0.5 for typical mineral oils. This can be seen in Figure 2.11, where its variation with temperature is also shown. The values for η_0 and α at different temperatures are taken from the formulation proposed by Wu, Klaus and Duda (1989).

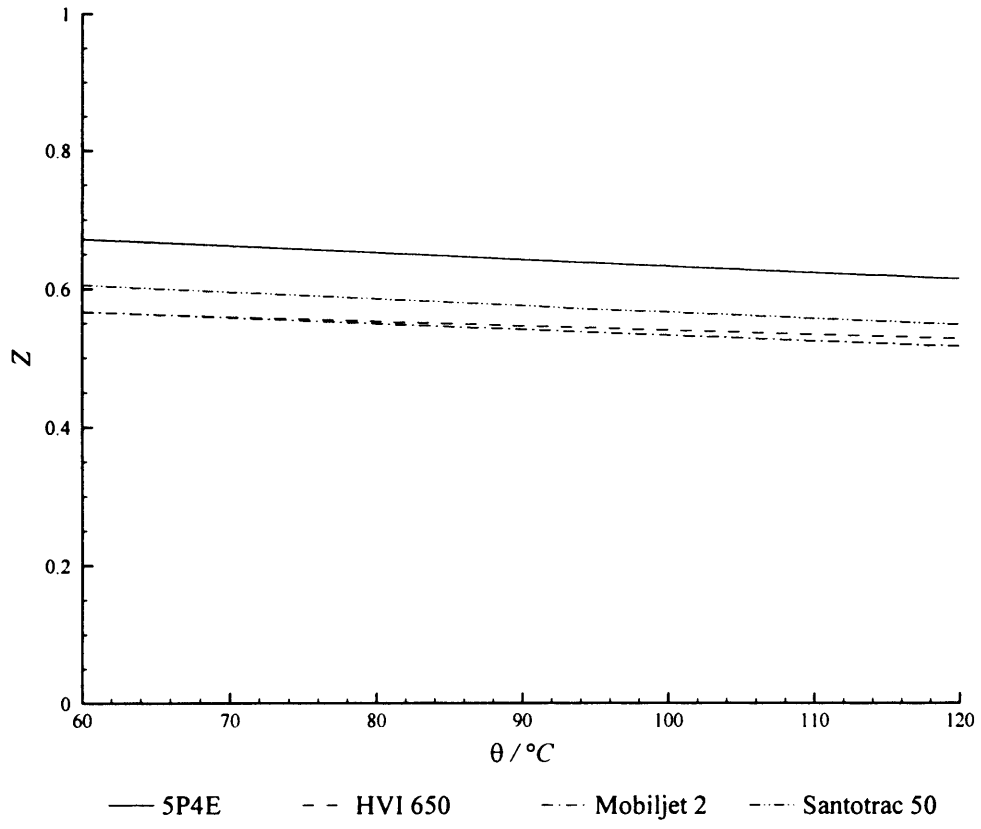

Figure 2.11 – Z variation with temperature

The effect of temperature variation through the contact can be taken account of in the thermal Roelands (1966) relationship:

$$\eta = \eta_0 e^{\ln(\eta_0)+9.67\left\{\left(1+5.1\times10^{-9}p\right)^{Z}\left(\frac{\theta_0-138}{\theta-138}\right)^{S_0}-1\right\}}$$

S_0 is varied to give the response with varying temperature. Table 2.1 shows the values that are necessary to obtain agreement with the results illustrated in figure 2.11.

Table 2.1 – Values of S_0 required to achieve temperature agreement

Oil	S_0
5P4E	1.47
HVI 650	1.29
Mobiljet 2	1.12
Santotrac 50	1.19

The Barus viscosity relationship is only accurate up to pressures up to about 1GPa, above this pressure it tends to over estimate the viscosity. It has been widely used to generate accurately the film forming effects of EHL because the film is formed at low pressures in the inlet. The differences in viscosities predicted by these formulas can be illustrated by considering Mobiljet 2 at 100°C, where $\eta_0 = 0.0048$ and $\alpha = 11.1\text{GPa}^{-1}$. The Barus formula gives viscosities of 1.23Pas, 317.62Pas and 81704.20Pas at pressures of 0.5GPa, 1GPa and 1.5GPa respectively. The isothermal Roelands formula gives viscosities of 0.23Pas, 2.93Pas and 23.05Pas at pressues of 0.5GPa, 1GPa and 1.5GPa respectively. The two relationships are inherently similar at low pressures but there is a marked difference at higher pressures, varying by up to two orders of magnitude at 1.5GPa.

This work uses the Roelands viscosity relationship as it generally gives much better agreement with measurement at the higher pressures that are seen in EHL contacts, although new concerns as to its high pressure accuracy have been recently articulated by Bair (2004).

2.6 Density Pressure Relationship

In a similar manner to the viscosity, the density of the fluid changes due to the pressure increase experienced within the contact. The change in density is much less pronounced as shown in Figure 2.12.

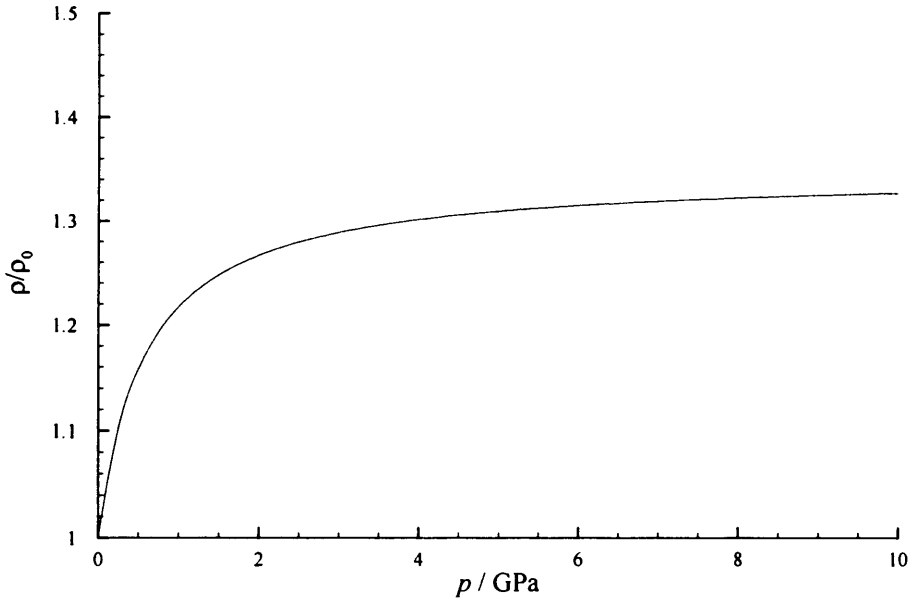

Figure 2.12 – Density variation with pressure

The relationship that is commonly used to describe this behaviour was proposed by Dowson and Higginson (1966) as:

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

The constants γ_ρ and λ_ρ are the compressibility constants for the oil used, typical values are 2.266 GPa^{-1} and 1.683 GPa^{-1} respectively. This equation can be adapted to take account of varying temperature:

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

2.7 The Thermal Equation

The temperature of the fluid and the solid surfaces vary in both the x and z directions. These temperatures are needed so that their influence on viscosity and density can be determined. The required energy equation is obtained by balancing the convection, conduction and heat generation within control volumes in the fluid and solid. The main assumptions made are that the pressure does not vary in the z direction, there is no velocity in the z direction and that the heat generation is represented by the heat generated at the centre of the control volume. The thermal energy balance for the fluid can be written as:

$$k\left(\frac{d^2\theta}{dx^2}+\frac{d^2\theta}{dz^2}\right)-\rho cu\frac{d\theta}{dx}+\varepsilon\left(u\frac{dp}{dx}+\frac{dp}{dt}\right)\theta-\rho c\frac{d\theta}{dt}+\tau\dot{\gamma}=0 \qquad (2.18)$$

Within the solids no heat is generated, the shear heating effect, $\tau\dot{\gamma}$, and the compressibility terms, ε, can be neglected giving:

$$k\left(\frac{d^2\theta}{dx^2}+\frac{d^2\theta}{dz^2}\right)-\rho cu\frac{d\theta}{dx}-\rho c\frac{d\theta}{dt}=0 \qquad (2.19)$$

Equations (2.18) and (2.19) can be applied within the fluid and solid respectively, however care must be taken at the fluid-solid interface. The surface temperatures are present in both equations which requires the equations to be linked in some way. This can be achieved by considering the continuity of flux, $\dot{q}$, between the solid and the fluid:

$$\dot{q}=-k\left.\frac{d\theta}{dz}\right|_{fluid}=-k\left.\frac{d\theta}{dz}\right|_{solid}$$

When contact occurs there is no fluid film, so the fluid equation cannot be used and two surfaces must be linked by considering the continuity of flux. In addition there is some heat generation due to the sliding of the two surfaces and this needs to be quantified. The heat generation changes the flux that flows between the two surfaces, which ensures that the flux flowing out of one surface is different from that flowing into the other. Considering both fluxes as positive in the positive z direction then the two fluxes can be related by:

$$\dot{q}_1-\dot{q}_2=heat\ gen$$

The heat generation does not occur over a volume, it is generated on a plane and is given by:

$$heat\ gen = \tau \times u_s$$

Evaluating the shear stress using the product of the coefficient of friction, μ, and the pressure, this allows the two fluxes to be linked by:

$$\dot{q}_1 - \dot{q}_2 = p\mu u_s$$

Since the product of the pressure and the coefficient of friction is always positive the absolute value of the velocity component must be taken to ensure that the heat generation is positive. This gives the equation that links the two surface fluxes to be written in the form:

$$\dot{q}_1 - \dot{q}_2 = p\mu|u_1 - u_2| \tag{2.20}$$

2.8 The Hydrodynamic Reynolds Equation

The full Navier Stokes equation can be simplified to a general equation for the lubricant flow that is considered in this work. Reynolds (1886) first showed this simplified differential equation and it is based on the following assumptions:

1) The body forces can be neglected
2) The pressure, hence viscosity and density, are constant through the thickness of the film
3) No slip occurs at the boundary edges
4) The lubricant behaves in a Newtonian manner
5) The flow is considered laminar
6) The inertia of the fluid is negligible

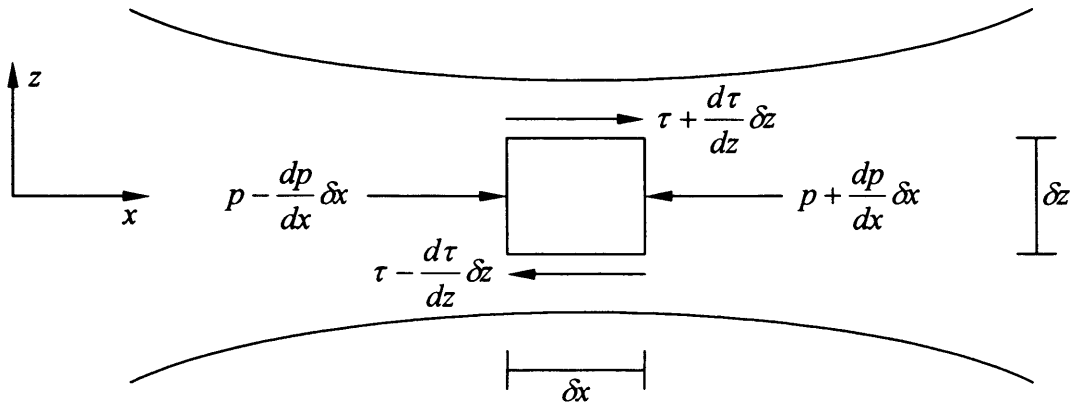

Figure 2.13 – Illustration of the forces that act on a small elemental volume

Considering the elemental volume shown in Figure 2.13, it can be shown using a force balance that:

$$\frac{d\tau}{dz} = \frac{dp}{dx} \tag{2.21}$$

Assumption 2 states that the pressure does not vary through the film, allowing integration of equation (2.21) to achieve:

$$\tau = \tau_m + z\frac{dp}{dx} \tag{2.22}$$

Since the fluid is Newtonian, the shear stress is related to the shear strain by:

$$\tau = \eta\dot{\gamma} = \eta\frac{du}{dz}$$

Substituting into (2.21) gives:

$$\frac{d}{dz}\left(\eta\frac{du}{dz}\right) = \frac{dp}{dx}$$

Integrating this term twice the velocity profile is obtained as:

$$u = \frac{1}{2\eta}\frac{dp}{dx}z^2 + \frac{A}{\eta}z + B \tag{2.23}$$

If there is no slip in the contact then the lubricant in contact with the surfaces must travel at the same speed as the surfaces. Denoting the upper surface velocity with suffix 1, the

lower with suffix 2 and defining the thickness of the lubricant to be h the following boundary conditions must be satisfied:

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

The boundary conditions allow the constants of integration to be determined as:

$$A = \eta\left(\frac{u_1 - u_2}{h}\right) \qquad\qquad B = \frac{u_1 + u_2}{2} - \frac{1}{8\eta}\frac{dp}{dx}h^2$$

Substituting into equation (2.23) yields:

$$u = \frac{u_1 + u_2}{2} + \frac{(u_1 - u_2)z}{h} - \frac{1}{2\eta}\frac{dp}{dx}\left[\left(\frac{h}{2}\right)^2 - z^2\right]$$

Equation (2.23) can be used to develop the mass flow through the contact Q.

$$Q = \int_{-\frac{h}{2}}^{\frac{h}{2}} \rho u.dz = \rho \int_{-\frac{h}{2}}^{\frac{h}{2}} \left\{\frac{1}{2\eta}\frac{dp}{dx}z^2 + \frac{Az}{\eta} + B\right\}.dz$$

Integration yields:

$$Q = \rho\left[\frac{1}{6\eta}\frac{dp}{dx}z^3 + \frac{Az^2}{2\eta} + Bz\right]_{-\frac{h}{2}}^{\frac{h}{2}} = \rho\left[\frac{1}{24\eta}\frac{dp}{dx}h^3 + Bh\right]$$

Substituting for B gives:

$$Q = \rho\left[\left(\frac{u_1 + u_2}{2}\right)h - \frac{1}{12\eta}\frac{dp}{dx}h^3\right] \tag{2.24}$$

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

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

Equation (2.24) becomes:

$$Q = \rho\bar{u}h - \frac{\rho h^3}{12\eta}\frac{dp}{dx} \tag{2.25}$$

It is worth noting that the total flow, Q, comprises of two parts, the couette flow, Q_c, that is related to the entrainment of the fluid and the Poiseuille flow, Q_p, that is the pressure driven flow.

$$Q_c = \rho \bar{u} h \qquad\qquad Q_p = -\frac{\rho h^3}{12\eta}\frac{dp}{dx}$$

Conserving mass in a control volume that spans the gap between the surfaces requires:

$$\frac{dQ}{dx} + \frac{d}{dt}(\rho h) = 0 \tag{2.26}$$

Substituting (2.25) into (2.26) gives:

$$\frac{d}{dx}\left(\frac{\rho h^3}{12\eta}\frac{dp}{dx}\right) - \frac{d}{dx}(\rho \bar{u} h) - \frac{d}{dt}(\rho h) = 0$$

This is the standard form of Reynolds equation. It can be simplified using a flow factor σ:

$$\frac{d}{dx}\left(\sigma \frac{dp}{dx}\right) - \frac{d}{dx}(\rho \bar{u} h) - \frac{d}{dt}(\rho h) = 0 \tag{2.27}$$

Where: $\sigma = \dfrac{\rho h^3}{12\eta}$

2.9 The Load Equation

The pressure that is generated within the contact is dependent on the separation of the surfaces h_s. The pressure that is generated dictates the load that can be supported by the contact.

The load per unit length in the y direction is defined as:

$$w' = \int p(x)dx \tag{2.28}$$

Where the integral is applied over the pressurised region.

In most problems the load is known and equation (2.28) is used to determine the separation between the surfaces. Practically, the separation is adjusted in the numerical solution scheme to allow the required load to be determined. Reducing the separation (or making it more negative) increases the pressure and hence the load and vice versa.

2.10 Friction Force

To determine the overall friction, F, in the contact the shear stress at the surface, τ_s, is obtained and combined with the pressure giving the following equation:

$$F = \int \tau_s \cos\xi + p \sin\xi \, dx$$

In this calculation the slope of the surface is taken into account through the angle, ξ, as the pressure outside the parallel film region can make a contribution to the total friction. This effect is ignored in determining the load as it is trivial in that instance.

Chapter 3

Non-Newtonian Behaviour

3.0 Introduction

This chapter builds upon the hydrodynamic equation outlined in section 2.8, where the fluid used was Newtonian, and considers the effects that non-Newtonian rheological models have as well as introducing various other equations that are necessary to solve the non-Newtonian problem. The rheological models considered include that attributed to Eyring (1936) and a number of different limiting shear stress models. The phenomena that occur as a result of limiting shear stress behaviour are also covered.

Equations (2.21) and (2.22) are still valid but the assumption that the fluid is Newtonian is no longer valid and the remaining equations have to be re-defined for the different rheological models presented.

3.1 Eyring Rheological Model

The Eyring shear thinning rheological model relates shear stress to shear strain rate by:

$$\dot{\gamma} = \frac{du}{dz} = \frac{\tau_0}{\eta}\sinh\left(\frac{\tau}{\tau_0}\right) \tag{3.1}$$

Substituting (2.22) into (3.1) gives:

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

This can be simplified to:

$$\frac{du}{dz} = \frac{\tau_0}{\eta}\sinh\left(\lambda_E + \frac{2\Sigma_E}{h}z\right) \tag{3.2}$$

Where: $\lambda_E = \frac{\tau_m}{\tau_0}$ and $\Sigma_E = \frac{h}{2\tau_0}\frac{dp}{dx}$

Integration of (3.2) allows the velocity of the fluid to be determined as:

$$u = \frac{\tau_0 h}{2\Sigma_E \eta}\cosh\left(\lambda_E + \frac{2\Sigma_E}{h}z\right) + const$$

Using the same boundary conditions as section 2.8 where:

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

Applying the lower surface boundary condition the velocity profile can be determined as:

$$u = u_2 + \frac{\tau_0 h}{2\Sigma_E \eta}\left[\cosh\left(\lambda_E + \frac{2\Sigma_E}{h}z\right) - \cosh(\lambda_E - \Sigma_E)\right]$$

Applying the upper surface boundary conditions the sliding velocity becomes:

$$u_s = \frac{\tau_0 h}{\Sigma_E \eta}\sinh(\lambda_E)\sinh(\Sigma_E)$$

Isolating λ:

$$\sinh(\lambda_E) = \frac{\eta u_s}{\tau_0 h}\frac{\Sigma_E}{\sinh(\Sigma_E)} \tag{3.3}$$

The equation for the mass flow was defined in section 2.8. This can be manipulated using integration by parts after Greenwood (2000):

$$Q = \int_{-\frac{h}{2}}^{\frac{h}{2}} \rho u dz = \int_{-\frac{h}{2}}^{\frac{h}{2}} \rho u \frac{dz}{dz} dz = \rho \bar{u} h - \int_{-\frac{h}{2}}^{\frac{h}{2}} \rho \frac{du}{dz} z dz \tag{3.4}$$

The first term is the Couette flow, thus evaluation of the integral will yield the Poiseuille flow as:

$$Q_p = -\frac{\rho\tau_0}{\eta}\int_{-\frac{h}{2}}^{\frac{h}{2}} z\sinh\left(\lambda_E + \frac{2\Sigma_E}{h}z\right)dz$$

$$= -\frac{\rho h\tau_0}{4\Sigma_E{}^2}\left[2\Sigma_E z\cosh\left(\lambda_E + \frac{2\Sigma_E}{h}z\right) - h\sinh\left(\lambda_E + \frac{2\Sigma_E}{h}z\right)\right]_{-\frac{h}{2}}^{\frac{h}{2}}$$

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

Substituting for τ_0 in this equation and factorising:

$$Q_p = -\frac{\rho h^3}{12\eta}\frac{dp}{dx}\left\{\frac{3[\Sigma_E\cosh(\Sigma_E)-\sinh(\Sigma_E)]}{\Sigma_E{}^3}\cosh(\lambda_E)\right\}$$

To attain the same result as equation (2.27) the flow factor must be redefined:

$$\sigma = \frac{\rho h^3}{12\eta}S \tag{3.5}$$

Where: $S = \dfrac{3[\Sigma_E\cosh(\Sigma_E)-\sinh(\Sigma_E)]}{\Sigma_E{}^3}\cosh(\lambda_E)$

Using equation (3.3) and hyperbolic identities S becomes:

$$S = \frac{3[\Sigma_E\cosh(\Sigma_E)-\sinh(\Sigma_E)]}{\Sigma_E{}^3}\sqrt{1+\left[\frac{\eta u_s}{\tau_0 h}\frac{\Sigma_E}{\sinh(\Sigma_E)}\right]^2} \tag{3.6}$$

S is the non-Newtonian correction to the viscosity, allowing an effective viscosity, η_{eff}, to be defined as:

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

This approach is that first presented by Conry, Wang and Cusano (1987). The concept of a modified flow factor, as seen in equation (3.5), is used for all other non-Newtonian

rheological models so that the Reynolds equation is taken to be as given in equation (2.27) for all rheological models.

3.2 Bair and Winer Rheological Model

This section looks at developing equations for the Bair and Winer rheological model which relates the strain rate to the shear stress by the relationship:

$$\dot{\gamma} = \frac{du}{dz} = \begin{cases} -\frac{\tau_l}{\eta}\ln\left(1-\frac{\tau}{\tau_l}\right) & \tau > 0 \\ \frac{\tau_l}{\eta}\ln\left(1+\frac{\tau}{\tau_l}\right) & \tau < 0 \end{cases} \tag{3.7}$$

From equation (2.22):

$$\frac{\tau}{\tau_l} = \lambda + \frac{2\Sigma}{h}z \tag{3.8}$$

Where: $\lambda = \frac{\tau_m}{\tau_l}$ and $\Sigma = \frac{h}{2\tau_l}\frac{dp}{dx}$

Provided $|\lambda| > |\Sigma|$ the sign of the shear stress does not change through the film and only one expression from equation (3.7) needs to be evaluated. This section is only concerned with examining λ and Σ as positive values and $\lambda \geq \Sigma$. The other combinations are covered in appendix A where general expressions are developed. Considering the combination of λ and Σ outlined above and substituting equation (3.8) into the first of equations (3.7):

$$\frac{du}{dz} = -\frac{\tau_l}{\eta}\ln\left(1-\lambda-\frac{2\Sigma}{h}z\right) \tag{3.9}$$

The velocity of the fluid can be obtained by integration of equation (3.9) as:

$$u = \frac{\tau_l h}{2\Sigma\eta}\left[\left(1-\lambda-\frac{2\Sigma}{h}z\right)\ln\left(1-\lambda-\frac{2\Sigma}{h}z\right)-\left(1-\lambda-\frac{2\Sigma}{h}z\right)\right]+const$$

The velocity variation across the film is obtained by applying the lower surface velocity as a boundary condition as:

$$u = u_2 + \frac{\tau_l h}{2\Sigma\eta}\left[\left(1-\lambda-\frac{2\Sigma}{h}z\right)\ln\left(1-\lambda-\frac{2\Sigma}{h}z\right)-(1-\lambda+\Sigma)\ln(1-\lambda+\Sigma)+\Sigma\left(\frac{2}{h}z+1\right)\right]$$

Applying the other boundary condition gives:

$$u_s = \frac{\tau_l h}{2\Sigma\eta}\left[(1-\lambda-\Sigma)\ln(1-\lambda-\Sigma)-(1-\lambda+\Sigma)\ln(1-\lambda+\Sigma)+2\Sigma\right] \qquad (3.10)$$

Equation (3.10) can be compared to the general expression for the sliding velocity taken from appendix A:

$$u_s = \frac{\tau_l h}{2\Sigma\eta}\left[(1-|\lambda+\Sigma|)\ln(1-|\lambda+\Sigma|)-(1-|\lambda-\Sigma|)\ln(1-|\lambda-\Sigma|)+|\lambda+\Sigma|-|\lambda-\Sigma|\right] \quad (3.11)$$

Equation (3.11) is valid for all combinations of λ and Σ. Similarly a general expression for the velocity across the film can be determined from appendix A as:

$$u = u_2 + \frac{\tau_l h}{2\Sigma\eta}\left[\begin{array}{c}\left(1-\left|\lambda+\frac{2\Sigma}{h}z\right|\right)\ln\left(1-\left|\lambda+\frac{2\Sigma}{h}z\right|\right)+\left|\lambda+\frac{2\Sigma}{h}z\right|-|\lambda-\Sigma| \\ -(1-|\lambda-\Sigma|)\ln(1-|\lambda-\Sigma|)\end{array}\right]$$

The Poiseuille flow can be determined from:

$$Q_p = -\int_{-\frac{h}{2}}^{\frac{h}{2}}\rho\frac{du}{dz}z.dz$$

$$= -\frac{\rho\tau_l}{\eta}\int_{-\frac{h}{2}}^{\frac{h}{2}}z\ln\left(1-\lambda-\frac{2\Sigma}{h}z\right).dz$$

$$= -\frac{\rho\tau_l h^2}{8\Sigma^2\eta}\left\{\begin{array}{l}\left(1-2\lambda+\lambda^2-\frac{4\Sigma^2}{h^2}z^2\right)\ln\left(1-\lambda-\frac{2\Sigma}{h}z\right) \\ -\frac{1}{2}\left[3-6\lambda+3\lambda^2-\frac{4\Sigma(1-\lambda)}{h}z-\frac{4\Sigma^2}{h^2}z^2\right]\end{array}\right\}_{-\frac{h}{2}}^{\frac{h}{2}} \qquad (3.12)$$

Applying the limits of integration to equation (3.12)

$$Q_p = -\frac{\rho\tau_l h^2}{8\Sigma^2\eta}\left[\left(1-2\lambda+\lambda^2-\Sigma^2\right)\ln\left(\frac{1-\lambda-\Sigma}{1-\lambda+\Sigma}\right)+2\Sigma(1-\lambda)\right]$$

The S term can be extracted from this as:

$$S = \frac{3}{4\Sigma^3}\left[\left(1-2\lambda+\lambda^2-\Sigma^2\right)\ln\left(\frac{1-\lambda-\Sigma}{1-\lambda+\Sigma}\right)+2\Sigma(1-\lambda)\right]$$

From appendix A this can be expressed in general form as:

$$S = \frac{3}{4|\Sigma|^3}\left[\left(1-2|\lambda|+\lambda^2-\Sigma^2\right)\ln\left(\frac{1-|\lambda|-|\Sigma|}{1-|\lambda|+|\Sigma|}\right)+2|\Sigma|\left(1-|\lambda|\right)\right] \tag{3.13}$$

Equation (3.13) is valid provided $|\lambda| \geq |\Sigma|$. If this is not the case then, from appendix A, the S term becomes:

$$S = \frac{3}{4|\Sigma|^3}\left[\begin{array}{c}\left(1-2|\lambda|+\lambda^2-\Sigma^2\right)\ln\left(1-|\lambda|-|\Sigma|\right)\\ +\left(1+2|\lambda|+\lambda^2-\Sigma^2\right)\ln\left(1+|\lambda|-|\Sigma|\right)+\Sigma^2+2|\Sigma|-\lambda^2\end{array}\right] \tag{3.14}$$

3.3 Lee and Hamrock's Circular Rheological Model

Lee and Hamrock (1990) proposed the shear strain to shear stress relationship as:

$$\dot{\gamma} = \frac{du}{dz} = \frac{\tau_l}{\eta}\frac{\frac{\tau}{\tau_l}}{\sqrt{1-\frac{\tau^2}{\tau_l^2}}}$$

This can be expressed using equation (3.8) as:

$$\frac{du}{dz} = \frac{\tau_l}{\eta}\frac{\lambda+\frac{2\Sigma}{h}z}{\sqrt{1-\left(\lambda+\frac{2\Sigma}{h}z\right)^2}}$$

Integration gives the fluid velocity:

$$u = -\frac{\tau_l h}{2\Sigma\eta}\sqrt{1-\left(\lambda+\frac{2\Sigma}{h}z\right)^2} + const$$

Applying the lower surface limit yields the velocity profile:

$$u = u_2 - \frac{\tau_l h}{2\eta\Sigma}\left[\sqrt{1-\left(\lambda+\frac{2\Sigma}{h}z\right)^2} - \sqrt{1-(\lambda-\Sigma)^2}\right]$$

Including the upper surface limit the sliding velocity can be obtained as:

$$u_s = \frac{\tau_l h}{2\eta\Sigma}\left[\sqrt{1-(\lambda-\Sigma)^2} - \sqrt{1-(\lambda+\Sigma)^2}\right] \tag{3.15}$$

The Poiseuille flow can be determined from:

$$Q_p = -\frac{\rho\tau_l}{\eta}\int_{-\frac{h}{2}}^{\frac{h}{2}} \frac{\left(\lambda+\frac{2\Sigma}{h}z\right)z}{\sqrt{1-\left(\lambda+\frac{2\Sigma}{h}z\right)^2}}.dz$$

$$Q_p = -\frac{\rho\tau_l h^2}{8\Sigma^2\eta}\left[\left(\lambda-\frac{2\Sigma}{h}z\right)\sqrt{1-\left(\lambda+\frac{2\Sigma}{h}z\right)^2} + \sin^{-1}\left(\lambda+\frac{2\Sigma}{h}z\right)\right]_{-\frac{h}{2}}^{\frac{h}{2}}$$

$$Q_p = -\frac{\rho\tau_l h^2}{8\Sigma^2\eta}\left[\begin{array}{c}(\lambda-\Sigma)\sqrt{1-(\lambda+\Sigma)^2} - (\lambda+\Sigma)\sqrt{1-(\lambda-\Sigma)^2} \\ +\sin^{-1}(\lambda+\Sigma) - \sin^{-1}(\lambda-\Sigma)\end{array}\right]$$

Manipulating this yields S as:

$$S = \frac{3}{4\Sigma^3}\left[\begin{array}{c}(\lambda-\Sigma)\sqrt{1-(\lambda+\Sigma)^2} - (\lambda+\Sigma)\sqrt{1-(\lambda-\Sigma)^2} \\ +\sin^{-1}(\lambda+\Sigma) - \sin^{-1}(\lambda-\Sigma)\end{array}\right] \tag{3.16}$$

3.4 Gecim and Winer Rheological Model

The shear strain is defined by:

$$\dot{\gamma} = \frac{du}{dz} = \frac{\tau_l}{\eta}\tanh^{-1}\left(\frac{\tau}{\tau_l}\right) \tag{3.17}$$

Re-arranging equation (3.17) and substituting in equation (3.8) the strain rate becomes:

$$\frac{du}{dz} = \frac{\tau_l}{2\eta} \ln\left(\frac{1+\lambda+\frac{2\Sigma}{h}z}{1-\lambda-\frac{2\Sigma}{h}z}\right)$$

Integrating this gives the velocity as:

$$u = \frac{\tau_l h}{4\Sigma\eta}\left[\left(1+\lambda+\frac{2\Sigma}{h}z\right)\ln\left(1+\lambda+\frac{2\Sigma}{h}z\right)+\left(1-\lambda-\frac{2\Sigma}{h}z\right)\ln\left(1-\lambda-\frac{2\Sigma}{h}z\right)-2\right]+const$$

Applying the lower limit gives the velocity profile as:

$$u = u_2 + \frac{\tau_l h}{4\Sigma\eta}\left[\begin{array}{c}\left(1+\lambda+\frac{2\Sigma}{h}z\right)\ln\left(1+\lambda+\frac{2\Sigma}{h}z\right)+\left(1-\lambda-\frac{2\Sigma}{h}z\right)\ln\left(1-\lambda-\frac{2\Sigma}{h}z\right)\\ -(1-\lambda+\Sigma)\ln(1-\lambda+\Sigma)-(1+\lambda-\Sigma)\ln(1+\lambda-\Sigma)\end{array}\right]$$

The sliding velocity can then be obtained as:

$$u_s = \frac{\tau_l h}{4\Sigma\eta}\left[\begin{array}{l}(1+\lambda+\Sigma)\ln(1+\lambda+\Sigma)+(1-\lambda-\Sigma)\ln(1-\lambda-\Sigma)\\ -(1-\lambda+\Sigma)\ln(1-\lambda+\Sigma)-(1+\lambda-\Sigma)\ln(1+\lambda-\Sigma)\end{array}\right] \qquad (3.18)$$

Which can be factorised to:

$$u_s = \frac{\tau_l h}{4\Sigma\eta}\left[\ln\left(\frac{1-\lambda^2-2\lambda\Sigma-\Sigma^2}{1-\lambda^2+2\lambda\Sigma-\Sigma^2}\right)+\lambda\ln\left(\frac{1-\lambda^2+2\Sigma+\Sigma^2}{1-\lambda^2-2\Sigma+\Sigma^2}\right)+\Sigma\ln\left(\frac{1+2\lambda+\lambda^2-\Sigma^2}{1-2\lambda+\lambda^2-\Sigma^2}\right)\right]$$

The Poisuielle flow is determined from equation (3.4) as:

$$Q_p = -\frac{\rho\tau_l}{2\eta}\int_{-\frac{h}{2}}^{\frac{h}{2}} z\ln\left(\frac{1+\lambda+\frac{2\Sigma}{h}z}{1-\lambda-\frac{2\Sigma}{h}z}\right).dz$$

Integrating gives:

$$Q_p = -\frac{\rho\tau_l h^2}{8\Sigma^2\eta}\left\{\begin{array}{l}\frac{\left(\frac{2\Sigma}{h}z\right)^2-\lambda^2-1}{2}\left[\ln\left(1+\lambda+\frac{2\Sigma}{h}z\right)-\ln\left(1-\lambda-\frac{2\Sigma}{h}z\right)\right]\\ -\lambda\left[\ln\left(1+\lambda+\frac{2\Sigma}{h}z\right)+\ln\left(1-\lambda-\frac{2\Sigma}{h}z\right)\right]+\frac{2\Sigma}{h}z+3\lambda\end{array}\right\}_{-\frac{h}{2}}^{\frac{h}{2}}$$

This gives the Poisuielle flow as:

$$Q_p = -\frac{\rho\tau_l h^2}{8\Sigma^2\eta}\left\{\begin{array}{l}\dfrac{\Sigma^2-\lambda^2-1}{2}\left[\ln(1+\lambda+\Sigma)+\ln(1-\lambda+\Sigma)-\ln(1+\lambda-\Sigma)-\ln(1-\lambda-\Sigma)\right] \\ +\lambda\left[\ln(1-\lambda+\Sigma)+\ln(1+\lambda-\Sigma)-\ln(1+\lambda+\Sigma)-\ln(1-\lambda-\Sigma)\right]+2\Sigma\end{array}\right\}$$

The corresponding S term is:

$$S = \frac{3}{4\Sigma^3}\left[\frac{\Sigma^2-\lambda^2-1}{2}\ln\left(\frac{1-\lambda^2+2\Sigma+\Sigma^2}{1-\lambda^2-2\Sigma+\Sigma^2}\right)+\lambda\ln\left(\frac{1-\lambda^2+2\lambda\Sigma-\Sigma^2}{1-\lambda^2-2\lambda\Sigma-\Sigma^2}\right)+2\Sigma\right] \quad (3.19)$$

3.5 Cardiff Rheological Model

The shear strain rate is determined from:

$$\dot{\gamma} = \frac{du}{dz} = \frac{\tau_l}{\eta}\frac{\dfrac{\tau}{\tau_l}}{\left(1-\dfrac{\tau^2}{\tau_l^2}\right)^m} \quad (3.20)$$

Where $0 < m < 1$

Substituting equation (3.8) into equation (3.20) and integrating gives:

$$u = \frac{\tau_l h}{4\Sigma\eta(m-1)}\left[1-\left(\lambda+\frac{2\Sigma}{h}z\right)^2\right]^{1-m} + const$$

Applying the limits gives the velocity profile as:

$$u = u_2 + \frac{\tau_l h}{4(m-1)\Sigma\eta}\left\{\left[1-\left(\lambda+\frac{2\Sigma}{h}z\right)^2\right]^{1-m} - \left[1-(\lambda-\Sigma)^2\right]^{1-m}\right\}$$

The sliding velocity is thus:

$$u_s = \frac{\tau_l h}{4(m-1)\Sigma\eta}\left\{\left[1-(\lambda+\Sigma)^2\right]^{1-m} - \left[1-(\lambda-\Sigma)^2\right]^{1-m}\right\} \quad (3.21)$$

This is a general expression that can be evaluated for all values of m between 0 and 1. It should be noted that $m = 0.5$ is Lee and Hamrock's circular model and substitution of this value into equation (3.21) gives the same result as equation (3.15).

The S term is more difficult to determine, in fact there is no general expression that satisfies all values of m. The integration can be done numerically however. The Pousueille flow can be determined from equation (3.4) and equation (3.20) as:

$$Q_p = -\frac{\rho\tau_l}{\eta}\int_{-\frac{h}{2}}^{\frac{h}{2}} \frac{\frac{\tau}{\tau_l}}{\left(1-\frac{\tau^2}{\tau_l^2}\right)^m} z dz$$

Comparing the Poisueille flow from the standard form of Reynolds equation:

$$\frac{\rho h^3}{12\eta} S\frac{dp}{dx} = \frac{\rho\tau_l}{\eta}\int_{-\frac{h}{2}}^{\frac{h}{2}} \frac{\frac{\tau}{\tau_l}}{\left(1-\frac{\tau^2}{\tau_l^2}\right)^m} z dz$$

Re-arranging S can be determined as:

$$S = \frac{6}{\Sigma h^2}\int_{-\frac{h}{2}}^{\frac{h}{2}} \frac{\frac{\tau}{\tau_l}}{\left(1-\frac{\tau^2}{\tau_l^2}\right)^m} z dz \qquad (3.22)$$

The best way of evaluating the integral in equation (3.22) was examined. The most successful method in relation to accuracy and time was to use Gauss integration. In high sliding situations the function changes very dramatically near the surfaces and as a result it was found best to split the whole film into the 3 smaller zones listed below:

1: $-0.46h \le z \le 0.46h$

2: $-0.5h \le z < -0.46h$

3: $0.46h < z \le 0.5h$

The first zone was evaluated using a small number of points as the function was almost linear in this region for all values of λ and Σ. Since the function changes dramatically in the other two zones they needed a considerable increase in the number of points for the integration. This reduces the time taken to perform the integration whilst maintaining a

good level of accuracy. The number of points to be used for the various zones was determined by comparing the integration for the $m = 0.5$ Cardiff model with the equations developed for the Circular model.

The number of points chosen for the zones was 6 for zone 1 and 16 for each of zones 2 and 3, giving a maximum percentage error of 0.99%. The variation of the error over the range of values can be seen in Figure 3.1 where it can be seen that the maximum error occurs at the limits of λ and Σ. The majority of the area has an error of less than 0.1%.

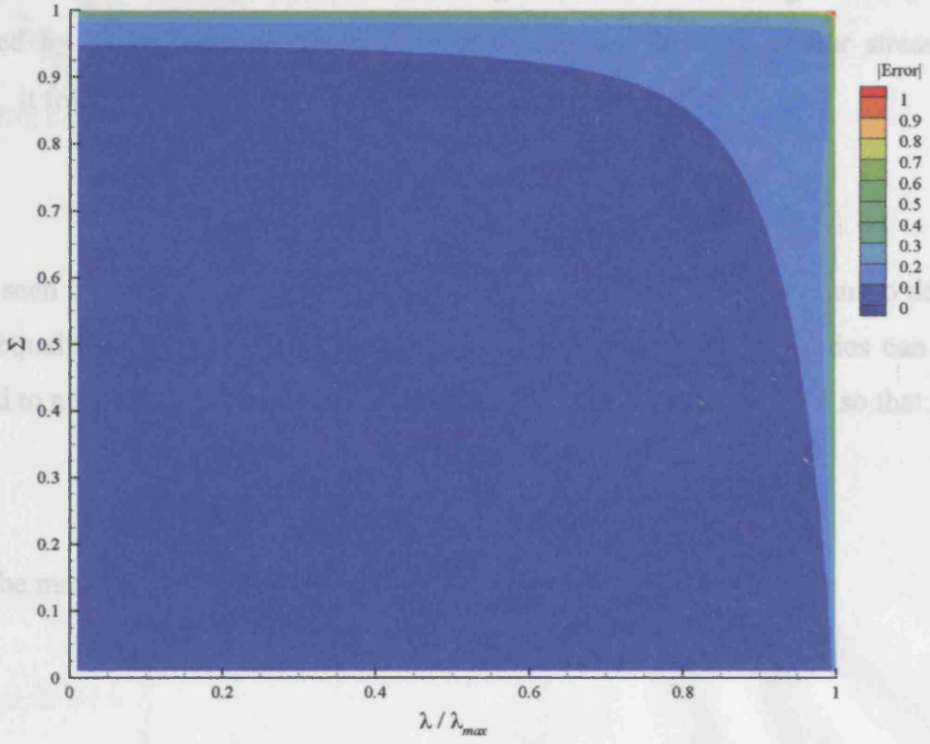

Figure 3.1 – Accuracy of numerical integration for $m = 0.5$ Cardiff model

3.6 Maximum Sliding and Non-Newtonian Viscosity Correction

The previous sections have outlined an expression that relates the amount of sliding imposed on the oil to the two non-dimensional parameters λ and Σ. For rheological models having a limiting shear stress, the values that λ and Σ can take are constrained.

The maximum shear stress will occur at one of the two surfaces, as $|z|$ is greatest at these locations. This allows the maximum value of the shear stress to be established from equation (3.8) as:

$$\frac{\tau}{\tau_l} = \lambda \pm \Sigma$$

The signs of λ and Σ are determined from the sign of the pressure gradient and the direction of the sliding velocity. The pressure gradient defines the sign of Σ and it can be seen from appendix A that the direction of the sliding velocity defines the sign of λ. It should also be noted from appendix A that the sign of Σ does not affect the direction of the sliding velocity and that the signs of λ and Σ do not affect S. For simplicity, the correction to the viscosity, S, and the magnitude of the sliding velocity can be determined by considering the magnitude of λ alone. Since the shear stress cannot exceed τ_l it follows that two conditions must be satisfied:

$$|\lambda| + \Sigma \leq 1 \qquad \text{and} \qquad |\lambda| - \Sigma \leq 1$$

It can be seen from these two inequalities that the sign of Σ is only relevant to determine which inequality constrains the values of λ and Σ. The two inequalities can thus be simplified to a single relationship by considering only the magnitude of Σ so that:

$$|\lambda| + |\Sigma| \leq 1$$

This can be manipulated to define the maximum magnitude of λ:

$$|\lambda| = 1 - |\Sigma| \tag{3.23}$$

The inequalities have been manipulated in terms of λ since this term is unknown in the solution scheme. Σ is already defined by the film thickness, pressure gradient and limiting shear stress. Equation (3.23) can be applied to determine the maximum value of the correction to the viscosity as well as the magnitude of the maximum sliding as both terms increase with increasing magnitude of λ.

3.6.1 Bair and Winer Model

Equation (3.10) can be simplified using a non-dimensional sliding velocity U_s to:

$$U_s = \frac{1}{2\Sigma}\left[(1-|\lambda+\Sigma|)\ln(1-|\lambda+\Sigma|)-(1-|\lambda-\Sigma|)\ln(1-|\lambda-\Sigma|)+|\lambda+\Sigma|-|\lambda-\Sigma|\right]$$

Where: $U_s = \frac{\eta u_s}{\tau_l h}$

Substituting equation (3.23) into the above equation gives the magnitude of the maximum sliding as:

$$|U_s|_{max} = \frac{1}{2\Sigma}\left[\begin{array}{c}\left(1-|(1-|\Sigma|)+\Sigma|\right)\ln\left(1-|(1-|\Sigma|)+\Sigma|\right)-\left(1-|(1-|\Sigma|)-\Sigma|\right)\ln\left(1-|(1-|\Sigma|)-\Sigma|\right)\\ +|(1-|\Sigma|)+\Sigma|-|(1-|\Sigma|)-\Sigma|\end{array}\right]$$

Considering the magnitude of Σ as its sign has no effect on the sliding velocity:

$$|U_s|_{max} = \frac{1}{2|\Sigma|}\left[1-|1-2|\Sigma||-(1-|1-2|\Sigma||)\ln(1-|1-2|\Sigma||)\right]$$

Applying the same approach to S allows the maximum correction to the viscosity, S_{max}, to be determined. As there are two different equations for this, equations (3.13) and (3.14), there will be two equations. Considering equation (3.13) and applying equation (3.23):

$$S_{max} = \frac{3}{4|\Sigma|^3}\left[\left(1-2(1-|\Sigma|)+(1-|\Sigma|)^2-\Sigma^2\right)\ln\left(\frac{1-(1-|\Sigma|)-|\Sigma|}{1-(1-|\Sigma|)+|\Sigma|}\right)+2|\Sigma|(1-(1-|\Sigma|))\right]$$

Which simplifies to:

$$S_{max} = \frac{3}{2|\Sigma|}$$

This is only valid provided $|\lambda| \geq |\Sigma|$, hence is only applicable for cases where $|\Sigma| \leq 0.5$.

For all other cases equation (3.14) must be used, applying equation (3.23):

$$S = \frac{3}{4|\Sigma|^3}\left[\begin{array}{c}\left(1-2(1-|\Sigma|)+(1-|\Sigma|)^2-\Sigma^2\right)\ln(1-(1-|\Sigma|)-|\Sigma|)\\ +\left(1+2(1-|\Sigma|)+(1-|\Sigma|)^2-\Sigma^2\right)\ln(1+(1-|\Sigma|)-|\Sigma|)+\Sigma^2+2|\Sigma|-\lambda^2\end{array}\right]$$

Which simplifies to:

$$S_{max} = \frac{3}{4|\Sigma|^3}\left[4(1-|\Sigma|)\ln(2-2|\Sigma|)+8|\Sigma|-2\Sigma^2-3\right]$$

3.6.2 Lee and Hamrock's Circular Model

Combining equations (3.15) and (3.23) and simplifying:

$$|U_s|_{max} = \frac{1}{2|\Sigma|}\sqrt{4|\Sigma|(1-|\Sigma|)}$$

The maximum S term can be determined from (3.16) and (3.23) as:

$$S_{max} = \frac{3}{4|\Sigma|^3}\left[\pi - \sin^{-1}(1-2|\Sigma|)\right]$$

3.6.3 Gecim and Winer Model

Substituting equation (3.23) into equation (3.18):

$$|U_s|_{max} = \frac{1}{2|\Sigma|}\ln\left(\frac{1}{1-|\Sigma|}\right) - \frac{1}{2}\ln\left(\frac{|\Sigma|}{1-|\Sigma|}\right)$$

Using (3.19) and (3.23) S_{max} can be obtained as:

$$S_{max} = \frac{3}{2|\Sigma|^3}\left[(1-|\Sigma|)\ln(1-|\Sigma|)+|\Sigma|\right] \qquad (3.24)$$

3.6.4 Cardiff Model

Using equations (3.21) and (3.23):

$$|U_s|_{max} = \frac{1}{4|\Sigma|(1-m)}\left[4|\Sigma|(1-|\Sigma|)\right]^{1-m}$$

Since there is no analytical expression for S there can be no analytical expression for S_{max}, instead the numerical integration must be performed for the case outlined in equation (3.23).

3.7 Slip

As the previous section shows, for a given Σ there is a certain amount of sliding that the oil can tolerate. The equations presented in section 3.6 lead to an equation that links the magnitude of the maximum non-dimensional sliding to the pressure gradient. This non-dimensional sliding can be related back to a physical sliding value, which is the difference in the surface speeds for the various points along the contact. However, only the magnitude of the maximum physical sliding can be determined since only the magnitude of the maximum non-dimensional sliding is known. The magnitude of the maximum physical sliding, $u_{s\max}$, that the oil can maintain is given by:

$$|u_{s\max}| = \frac{\tau_l h}{\eta}|U_{s\max}|$$

If the difference in the surface speeds results in a greater sliding demand than the oil can withstand, the oil must compensate for this is some way. The oil cannot change the velocities of the surfaces themselves, but can move at a different speed to the surfaces, hence leading to slip. In slipping the oil now has different boundary conditions to those outlined in section 2.8. The Couette flow will alter due to the slip and needs to be re-defined for this case. Slip can be expected to occur at the surfaces as this is where the shear stress reaches the limiting value first.

Considering a positive pressure gradient and sliding velocity, u_s (hence λ), without slip, then the velocity profile will be of the form shown in Figure 3.2.

Considering the shear stress on the two surfaces:

$$\tau_1 = \tau_l(\lambda + \Sigma) \qquad \text{and} \qquad \tau_2 = \tau_l(\lambda - \Sigma)$$

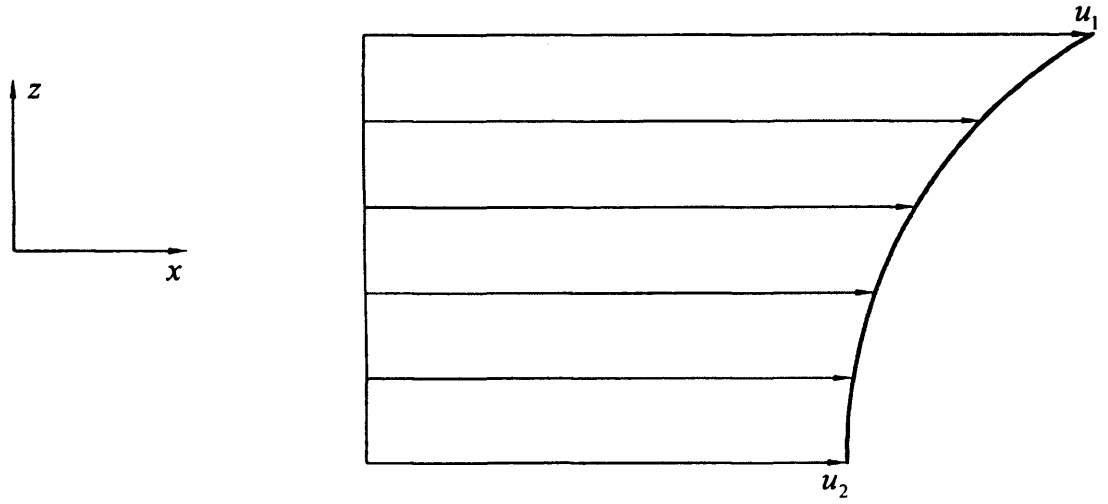

Figure 3.2 – Velocity profile for positive pressure gradient without slip

λ increases with the sliding velocity which increases the shear stress. When λ reaches its maximum value, the shear stress reaches the limiting value, τ_l, at surface 1 (upper surface). Increases in the sliding velocity cannot be accommodated by the oil as the value of λ needed would violate the rheological equation. If λ was increased it would lead to the shear stress exceeding τ_l at surface 1 which is the only location within the oil film that can exceed the limiting value.

If the maximum sliding velocity is exceeded then slip will occur at the upper surface, subscript 1, so that the velocity profile becomes as shown in Figure 3.3.

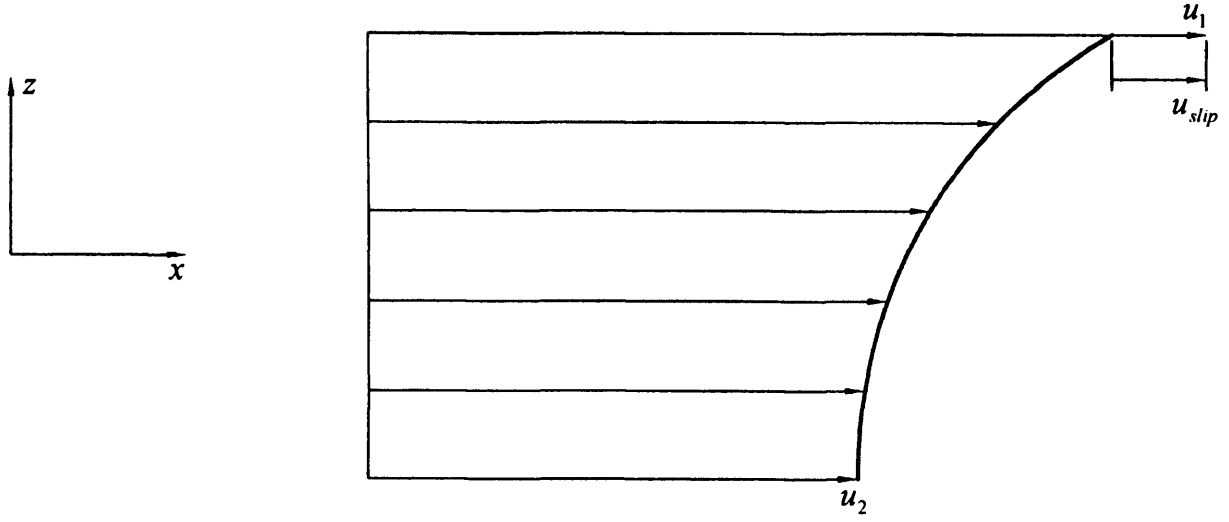

Figure 3.3 – Velocity profile for positive pressure gradient including slip velocity

The slip velocity, u_{slip}, always acts in the positive x direction and can be determined from:

$$u_{slip} = |u_1 - u_2| - |u_{s\max}|$$

The boundary conditions that are used to determine the Couette flow can be re-defined as:

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

From section 2.8 the Couette flow was defined by:

$$Q_c = \rho \int_{-\frac{h}{2}}^{\frac{h}{2}} u dz$$

Applying the boundary conditions this becomes:

$$Q_c = \rho(u_1 - u_{slip})\frac{h}{2} + \rho u_2 \frac{h}{2} = \rho \bar{u} h - \rho u_{slip} \frac{h}{2}$$

The Couette flow can be simplified by using a modified mean entrainment velocity, $\hat{u}$, thus:

$$Q_c = \rho \hat{u} h$$

Where: $\hat{u} = \bar{u} - \frac{u_{slip}}{2}$

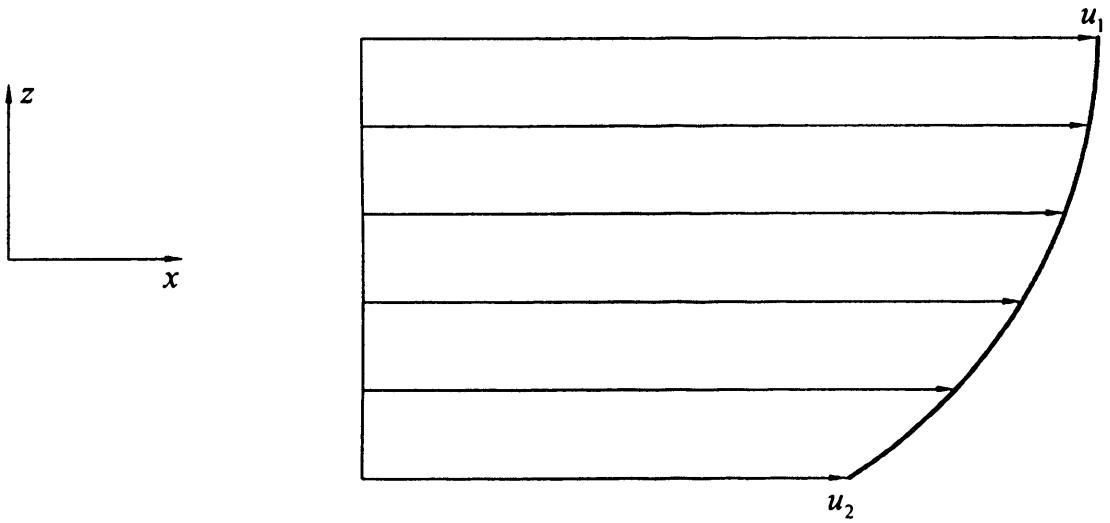

Figure 3.4 – Velocity profile for negative pressure gradient without slip

For cases having a negative pressure gradient, the velocity profile without slip will be of the form shown in Figure 3.4.

The change in pressure gradient results in the shear stress reaching the limiting value at surface 2 (lower) first. If the maximum no slip sliding velocity, $u_{s\max}$, is exceeded slip will occur at the lower surface, surface 2, so that the velocity profile becomes as shown in Figure 3.5.

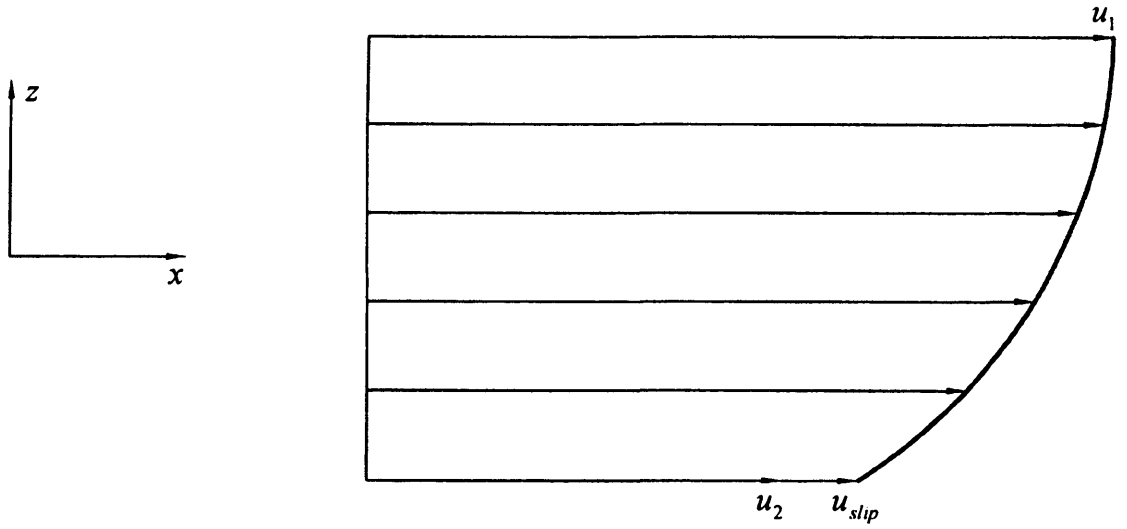

Figure 3. 5 – Velocity profile for negative pressure gradient including slip velocity

The boundary conditions for the Couette flow change to:

$$u = u_1 \text{ at } z = \frac{h}{2} \qquad \text{and} \qquad u = u_2 + u_{slip} \text{ at } z = -\frac{h}{2}$$

The change in boundary conditions modify the Couette flow to:

$$Q_c = \rho u_1 \frac{h}{2} + \rho\left(u_2 + u_{slip}\right)\frac{h}{2} = \rho \bar{u} h + \rho u_{slip} \frac{h}{2}$$

The modified mean entrainment velocity now becomes:

$$\hat{u} = \bar{u} + \frac{u_{slip}}{2}$$

The modified mean entrainment velocity changes the Couette flow, which means the hydrodynamic Reynolds equation changes to:

$$\frac{d}{dx}\left(\sigma\frac{dp}{dx}\right)-\frac{d}{dx}(\rho\hat{u}h)-\frac{d}{dt}(\rho h)=0 \tag{3.25}$$

Where: $\sigma=\frac{\rho h^3}{12\eta}S$ and $\hat{u}=\bar{u}\pm\frac{u_{slip}}{2}$

The influence that the slip velocity has on the modified mean entrainment velocity can be linked to the pressure gradient. For positive pressure gradients the slip acts to reduce the modified mean entrainment velocity and for negative pressure gradients the slip acts to increase the modified mean entrainment velocity. This can be summarised by the following:

$$\hat{u}=\bar{u}-\frac{u_{slip}}{2}\ \text{for}\ \frac{dp}{dx}>0 \qquad \text{and} \qquad \hat{u}=\bar{u}+\frac{u_{slip}}{2}\ \text{for}\ \frac{dp}{dx}<0$$

It should be noted that making the lower surface faster than the upper, $u_2 > u_1$, will make λ negative, but it will not affect the equations developed for the modified mean entrainment velocity. Changing the surface velocities in this manner will only alter the surface at which slip occurs. This can be accounted for by re-expressing the manner in which slip if defined.

For positive pressure gradients, slip may occur on the faster moving surface and will tend to reduce the modified mean entrainment velocity. For negative pressure gradients, if slip occurs it will take place on the slower moving surface and will act to increase the modified mean entrainment velocity.

The approach presented here and the conclusions made are the same as those found by other authors; Jacobson and Hamrock (1984), Zhang, Wen and Wei (2001) and Zhang and Wen (2002).

The effect that slip has on the Couette flow has been introduced, however, its influence on the temperature has not. In section 2.7 it was found that contact produced a change in the heat flux that flows between the two surfaces. Slip produces a similar effect, which

means that there is no longer a conservation of flux. It was shown that the heat generated at the interface was given by:

$$heat\ gen = \tau \times u_s$$

Considering the case where slip is occurring, the velocity that is generating the heat will be the slip velocity. The shear stress will be the shear stress at the point where slip is occurring, this will be $\pm \tau_l$. However, care must be taken with the shear stress, it was noted earlier that u_{slip} is always positive and it can be seen that the shear stress can be either positive or negative. A situation could occur where the shear stress is negative which would lead conceptually to negative heat generation, where the slip would remove heat from the interface. This is not the case, slip will always generate heat, therefore it is easier to consider the heat generated at points where slip occur as the product of the positive limiting shear stress and the slip velocity:

$$heat\ gen = \tau_l \times u_{slip}$$

This can be written in the same form as equation (2.20) to give the change in flux due to the slip as:

$$\dot{q}_1 - \dot{q}_2 = \tau_l \times u_{slip} \tag{3.26}$$

3.8 Limiting Shear Stress Behaviour at the Extreme Values of Σ

Section 3.6 defined the limits of λ and Σ, the limits were used to develop two general conditions that must be obeyed:

$$|\lambda| + \Sigma \le 1 \qquad \text{and} \qquad |\lambda| - \Sigma \le 1$$

These conditions can be used to develop two inequalities that define the behaviour of λ and Σ:

$$|\lambda| \le 1 - |\Sigma| \qquad \text{where} \qquad 0 \le |\Sigma| \le 1$$

The equations that have been derived work within this particular range and the effects that the first inequality have were covered in section 3.6. The behaviour of the equations at the extreme values of Σ need further consideration

3.8.1 Σ Approaching Zero

As Σ approaches zero the intrinsic numerical procedures that are used to determine S begin to 'break down'. The term must be investigated further so that its response within this region can be established. This can be done using a Maclaurin expansion, which is defined by:

$$f(x)_{x\to 0} = f(0) + f'(0).x + f''(0)\frac{x^2}{2!} + f'''(0)\frac{x^3}{3!} + \ldots.$$

Recalling the definition of Σ:

$$\Sigma = \frac{h}{2\tau_l}\frac{dp}{dx}$$

This term approaches zero because the product of the pressure gradient and the film thickness will tend to 'small' values. The other non-dimensional shear stress component, λ, is related to the mid-film shear stress. λ is not affected by small pressure gradients or small film thicknesses, it is only influenced by the amount of sliding produced. It can be assumed for most cases that a modest amount of sliding will be present ensuring a value of λ larger than the small values of Σ as it approaches zero. Therefore it is reasonable to assume that in the regions where Σ approaches zero λ will remain unaffected and will not approach zero. Thus for the Bair and Winer model only equation (3.13) needs evaluating as $|\lambda| > |\Sigma|$. The expansion of this equation to three terms is:

$$S_{\Sigma\to 0} = \frac{1}{1-\lambda} + \frac{\Sigma^2}{5(1-\lambda)^3} + \frac{3\Sigma^4}{35(1-\lambda)^5} \tag{3.27}$$

The Circular model also gives rise to numerical problems at low Σ and the Maclaurin expansion for S is:

$$S_{\Sigma\to 0} = \frac{1}{(1-\lambda^2)^{\frac{3}{2}}} + \frac{3(1+4\lambda^2)}{10(1-\lambda^2)^{\frac{7}{2}}}\Sigma^2 + \frac{9(1+12\lambda^2+8\lambda^4)}{56(1-\lambda^2)^{\frac{11}{2}}}\Sigma^4$$

Similarly the Gecim and Winer model expansion to three terms is:

$$S_{\Sigma\to 0} = \frac{1}{1-\lambda^2} + \frac{1+3\lambda^2}{5(1-\lambda^2)^3}\Sigma^2 + \frac{1+10\lambda^2+5\lambda^4}{35(1-\lambda^2)^5}\Sigma^4$$

In addition, equation (3.24) also breaks down. This is due to the numerical instability of the following term:

$$\ln(1-|\Sigma|)$$

Expanding this term to the fourth order differential:

$$\ln(1-|\Sigma|)_{\Sigma\to 0} = -\Sigma - \frac{\Sigma^2}{2} - \frac{\Sigma^3}{3} - \frac{\Sigma^4}{4} \tag{3.28}$$

Interestingly a Maclaurin expansion can be obtained for the Cardiff model in spite of the fact that S is not available as a closed form function. The terms that contribute to S can be differentiated with respect to Σ and then the resulting equation can be integrated with respect to z. The resulting expansion for the Cardiff model to three terms is:

$$S_{\Sigma\to 0} = \frac{1+(2m-1)\lambda^2}{(1-\lambda^2)^{m+1}} + \frac{m\left[3+(12m+6)\lambda^2+(4m^2-1)\lambda^4\right]}{5(1-\lambda^2)^{m+3}}\Sigma^2 + \frac{m\left[\begin{array}{c}15(m+1)+(90m^2+225m+135)\lambda^2+(60m^3+180m^2+165m+45)\lambda^4 \\ +(8m^4+20m^3+10m^2-5m-3)\lambda^6\end{array}\right]}{70(1-\lambda^2)^{m+5}}\Sigma^4$$

3.8.2 Σ Equals Zero

Section 3.8.1 introduced equations for S that can be used as Σ tends towards zero. Consideration was not given to the effect that Σ approaching zero has on the terms for the non-dimensional sliding velocity, namely equations (3.11), (3.15), (3.18) and (3.21). For the case of $\Sigma = 0$ these equations cannot be used as there is a Σ term on the denominator, instead another approach has to be taken to solve for these values. In addition, the Maclaurin expansions have some power of the term $1-\lambda$, for high sliding cases where Σ approaches zero this term can also approach zero, instead the equations for S are re-expressed where they are determined from the non-dimensional sliding velocity which removes the problem.

Σ has been defined by:

$$\Sigma = \frac{h}{2\tau_l}\frac{dp}{dx}$$

When this term is zero one of two things has happened. The first is that the film thickness has become zero, in which case there is no fluid between the surfaces and hence no hydrodynamic equation to be solved. The second is that there is no pressure gradient, this case means that there can be no Poisueille flow unless S equals infinity. From the previous section it can be seen that S approaches a finite value that is dependent on λ. In this case λ can be evaluated and S determined.

An expansion could be carried out for the velocity terms so that they could be evaluated for $\Sigma = 0$, however it is easier to re-define the velocity component from first principles.

Consider the Bair and Winer model where $\Sigma = 0$:

$$\frac{du}{dz} = -\frac{\tau_l}{\eta}\ln\left(1-\frac{\tau}{\tau_l}\right) = -\frac{\tau_l}{\eta}\ln(1-\lambda)$$

The corresponding sliding velocity can be obtained as:

$$u_s = -\frac{\tau_l h}{\eta}\ln(1-\lambda)$$

Re-arranging for λ:

$$\lambda = 1 - e^{-U_s} \tag{3.29}$$

The S term can also be re-defined using this term. From the previous section the Maclaurin expansion for the Bair and Winer model can be used giving:

$$S_{\Sigma=0} = \frac{1}{1-\lambda}$$

This can be manipulated using equation (3.29) to:

$$S_{\Sigma=0} = \frac{1}{e^{-U_s}} = e^{U_s}$$

The Circular model produces the expressions:

$$\lambda = \frac{U_s}{\sqrt{1+U_s^{\,2}}}$$

$$S_{\Sigma=0} = \left(1+U_s^{\,2}\right)^{\frac{3}{2}}$$

The Gecim and Winer model gives the following:

$$\lambda = 1 - \tanh(U_s)$$

$$S_{\Sigma=0} = \frac{\left(1+e^{2U_s}\right)^2}{4e^{2U_s}}$$

The Cardiff model cannot be manipulated like the other models. Considering the equation that defines the model:

$$\frac{du}{dz} = \frac{\tau_l}{\eta} \frac{\frac{\tau}{\tau_l}}{\left(1-\frac{\tau^2}{\tau_l^2}\right)^m}$$

Substituting $\Sigma = 0$:

$$\frac{du}{dz} = \frac{\tau_l}{\eta} \frac{\lambda}{\left(1-\lambda^2\right)^m}$$

This can be written in terms of the non-dimensional sliding velocity as:

$$U_s = \frac{\lambda}{\left(1-\lambda^2\right)^m}$$

This cannot be manipulated to give a functional form for λ, instead the equation is solved by iterative methods for λ.

The value of S for $\Sigma = 0$ can be manipulated from the Maclaurin expansion in terms of λ as:

$$S_{\Sigma=0} = \frac{1+(2m-1)\lambda^2}{\left(1-\lambda^2\right)^{m+1}}$$

3.8.3 Σ Equals Unity

As $|\Sigma| = 1$, it follows from equation (3.23) that $\lambda = 0$ and substituting these values into the velocity terms in section 3.6 gives zero sliding velocity. Since there is no sliding velocity then the oil at the two surfaces must move at the same speed, which will inevitably lead to slip. Section 3.7 examined slipping and it was determined that slip

occurs on the faster surface for positive pressure gradients and the slower surface for negative pressure gradients. The same approach is adopted for $|\Sigma| = 1$, where the slip will occur at one surface alone.

Other authors [Jacobson and Hamrock, 1984, Zhang and Wen, 2002] have allowed slip to occur on both surfaces for $|\Sigma| \geq 1$. However, in the approach adopted here, slip is a phenomenon that alters the velocity of the oil at the surfaces to ensure the mean shear stress, $\tau_m = \tau_l \lambda$, is an acceptable value so that limiting shear stress is not exceeded within the film. Slip only occurs at a surface where the shear stress would be exceeded, since this was the same criteria set out in section 3.7, the same conclusions apply. It should be noted that although both surfaces have a shear stress equal to the limiting shear stress it does not mean that both surfaces would slip. In addition, $|\Sigma| > 1$ is not considered a valid solution, as this will result in a shear stress greater than the limiting shear stress to be developed at both surfaces. Therefore Σ is limited to the range -1 to 1 in the solution process. A 'converged' solution is considered invalid if the solution contains a value of Σ that is outside these limits.

The effect of $\Sigma = 1$ on the S term can be evaluated by substituting into the equations in section 3.6. This is straight-forward and does not require limit consideration.

Considering the Bair and Winer model:

$$S_{max} = \frac{3}{4|\Sigma|^3}\left[4(1-|\Sigma|)\ln(2-2|\Sigma|)+8|\Sigma|-2|\Sigma|^2-3\right]$$

Substituting $\Sigma = 1$:

$$S_{\Sigma=1} = \frac{3}{4}\left[4(1-1)\ln(2-2)+8-2-3\right] = \frac{9}{4}$$

The Gecim and Winer model reduces to:

$$S_{\Sigma=1} = \frac{3}{2}$$

Similarly the Circular model yields:

$$S_{\Sigma=1} = \frac{3\pi}{4}$$

The Cardiff model does not have a finite value, instead requiring evaluation of:

$$S_{\Sigma=1} = \frac{6}{h^2} \int_{-\frac{h}{2}}^{\frac{h}{2}} \frac{z^2}{\left(1 - z^2\right)^m} dz$$

Chapter 4

Numerical Methods

4.0 Introduction

The equations developed in Chapter 3 allow many different non-Newtonian limiting shear stress rheological models to be utilised. These equations are investigated further, their stability is assessed and a robust and quick solution process determined. The response of the different rheological models are compared and contrasted to give some indication of their different features. The numerical approaches used to solve the thermal equations as well as the general solution method for the problem as a whole are presented.

4.1 Maclaurin Expansion

Since there are expansions for low Σ, it is necessary to investigate where the expansions need to be used. Figure 4.1 illustrates a typical break down of the intrinsic procedures used in equation (3.13) for the Bair and Winer model compared to the Maclaurin expansion, equation (3.27). The figure shows that the equation oscillates at small values of Σ, it can also be seen that the Maclaurin expansion gives good agreement at small Σ but begins to deviate slightly as Σ increases. The only way to improve this would be to include more terms in the expansion, however, this will increase the time taken to perform the calculation which is undesirable. Considering only three terms, each of the expansions can then be investigated and the range for which it is necessary to use the expansion determined.

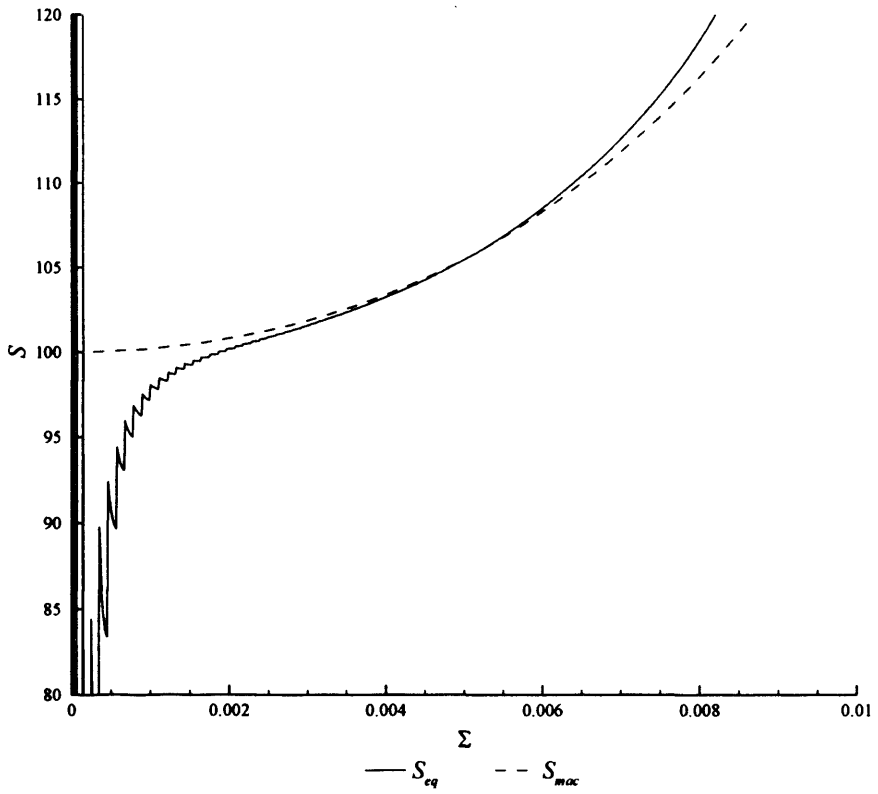

Figure 4.1 – Response of Equation on Maclaurin expansion for $\lambda = 0.9$

Defining the value taken for S from sections 3.2 to 3.4 based on intrinsic procedures as S_{eq} and the value based on the Maclaurin expansions as S_{mac}, the error can be expressed as:

$$error = 100 * \frac{S_{eq} - S_{mac}}{S_{eq}}$$

Figure 4.2 shows contours of the magnitude of the error in S for the Bair and Winer model over the whole parameter range of λ and Σ (note that $|\lambda| + |\Sigma| \leq 1$). The region where the Maclaurin expansion is required because of a breakdown in S_{eq} evaluation using intrinsic procedures is the red area at the bottom of the figure. This region can be identified with the elliptical boundary shown in the figure. The error that occurs as λ approaches 1 is the inability of the Maclaurin expansion to accurately follow the actual equation. In contrast, the error that occurs as λ approaches 0 is related to a change in the equation used to calculate S_{eq}. The Macluarin expansion was conducted for equation (3.13) where $|\lambda| \geq |\Sigma|$, it can be seen from the area that is red that this condition is not met

and there is some inaccuracy since the equation used to determine S_{eq} is equation (3.14). The region that is bounded by the ellipse does not suffer this error and the approach of using this as the method selection boundary is deemed acceptable.

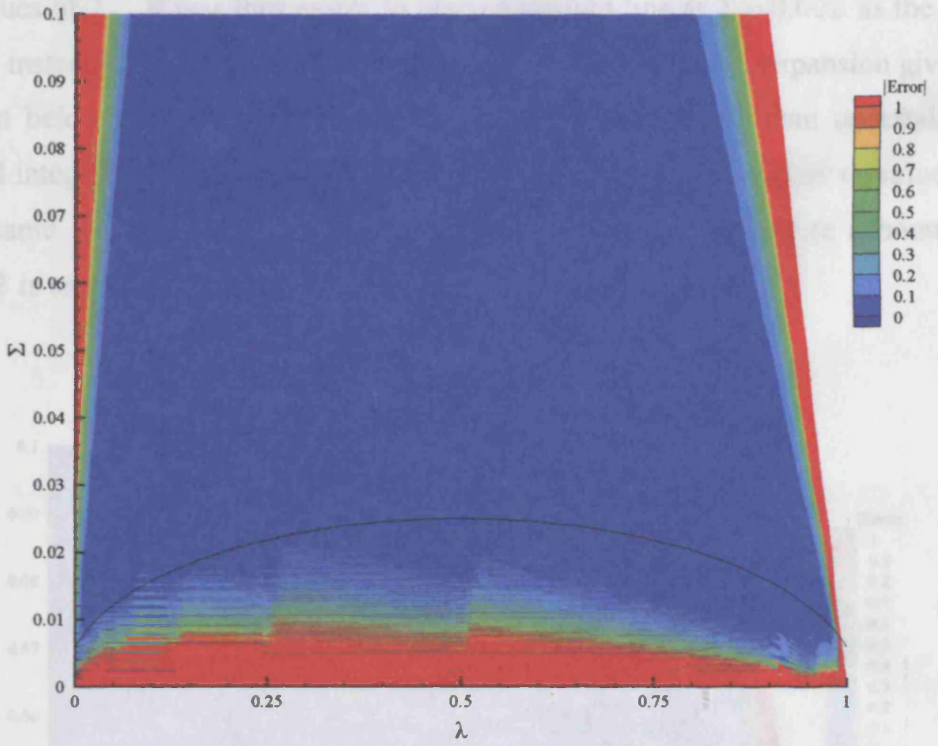

Figure 4.2 – Bair and Winer model accuracy including ellipse used as boundary for small Σ expansion

The equation of the ellipse shown in figure 4.2 is given by:

$$\frac{(\lambda-0.5)^2}{0.515^2}+\frac{\Sigma^2}{0.025^2}=1$$

The Circular and Gecim and Winer models are shown in Figures 4.3 and 4.4 respectively, where suitable boundary ellipses are also shown. The ellipses used are:

Circular: $$\frac{(\lambda-0.5)^2}{0.51^2}+\frac{\Sigma^2}{0.02^2}=1$$

Gecim and Winer: $$\frac{(\lambda-0.95)^2}{1^2}+\frac{\Sigma^2}{0.007^2}=1$$

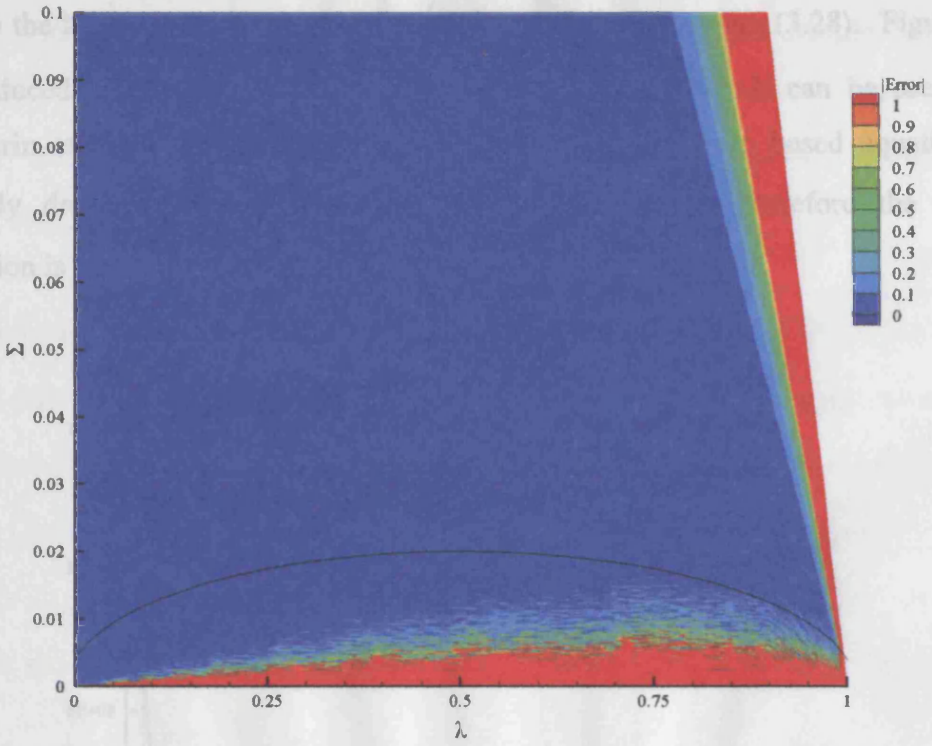

Figure 4.3 – Circular model accuracy including ellipse used as boundary for small Σ expansion

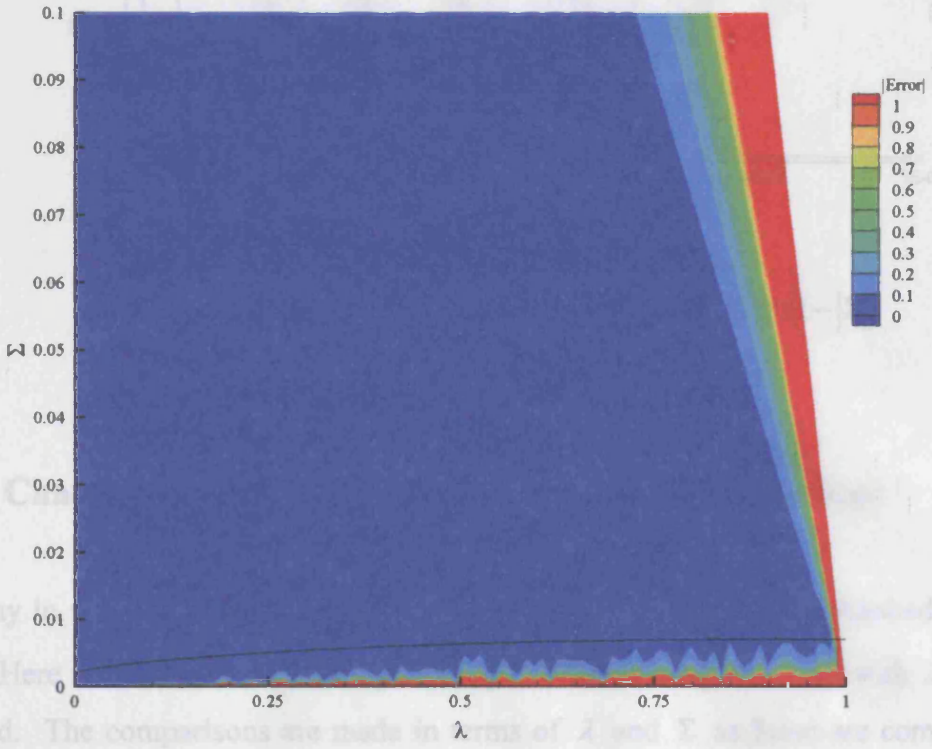

Figure 4.4 – Gecim and Winer accuracy including ellipse used as boundary for small Σ expansion

The break down that occurs for the Cardiff model is much less pronounced than the other models considered. This is illustrated in Figure 4.5 where the Cardiff model is used for $m = 0.5$ and was chosen to allow comparison with the Circular model. The reduced volatility of the S_{eq} term is due to it being calculated by numerical integration rather than using an equation based on intrinsic procedures. It does however break down at much lower values of Σ. It was thus easier to place a straight line at $\Sigma = 0.002$ as the method boundary instead of seeking an elliptical boundary. The Maclaurin expansion gives good agreement below this value of Σ and the calculation is quicker than undertaking the numerical integration. The same comparison was also made for a variety of values of m and the same response as shown in figure 4.5 was observed, therefore a boundary of $\Sigma = 0.002$ is used for all values of m.

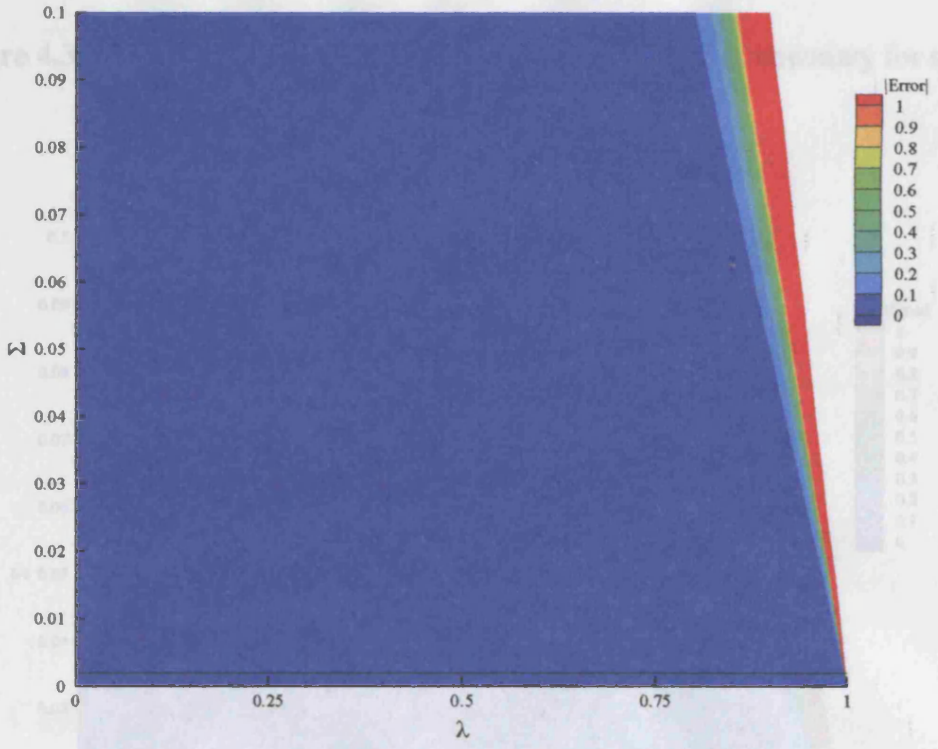

Figure 4.5 – Cardiff model accuracy including boundary for small Σ expansion with $m = 0.5$

The next consideration is the break down of $\ln(1-|\Sigma|)$ in the $S_{\max}$ term for the Gecim and Winer model, equation (3.24). Setting S_{eq} to the term $\ln(1-|\Sigma|)$ from equation (3.24) and

S_{mac} to the Maclaurin expansion for $\ln(1-|\Sigma|)$, which is equation (3.28). Figure 4.6 can be produced comparing the response of the two equations. It can be seen that the Maclaurin expansion gives very good agreement, the intrinsic based equation results suddenly drop and become unstable at $\Sigma \approx 1.50\times10^{-9}$. Therefore the Maclaurin expansion is used for $\Sigma \le 2.00\times10^{-8}$.

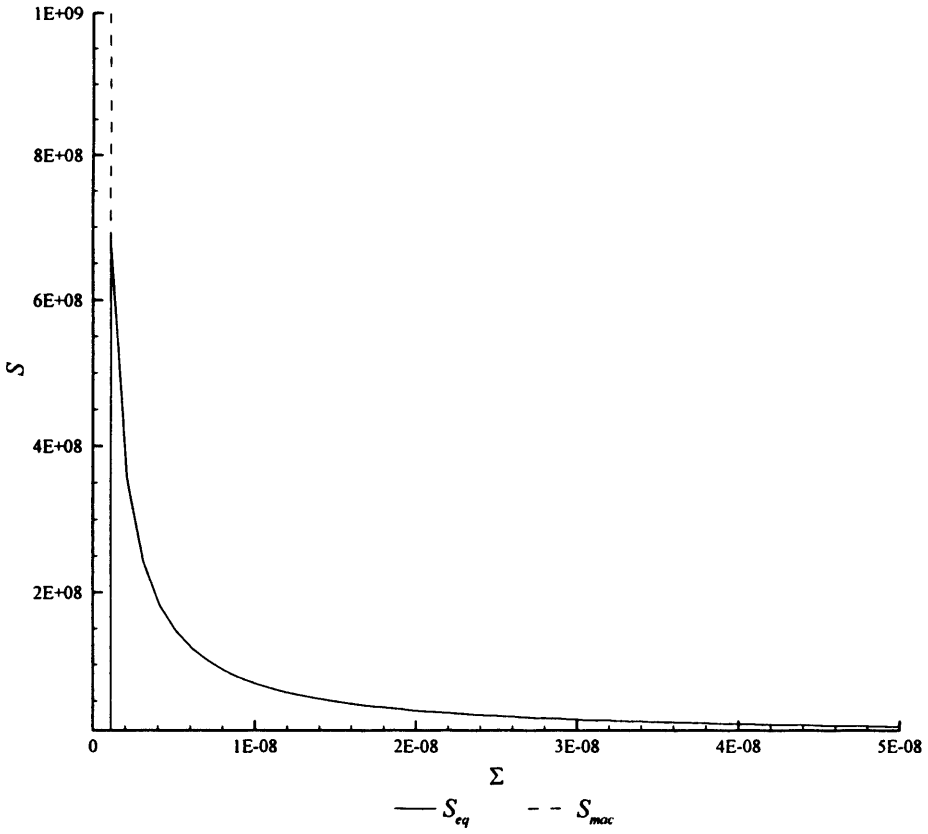

Figure 4.6 – Maclaurin expansion comparison for $\ln(1-|\Sigma|)$

4.2 Characteristic Behaviour of Non-Newtonian Functions

The way in which each model behaves in terms of $\dot{\gamma}$ and τ was established in section 1.7.1. Here the way in which the equations presented in chapter 3 vary with λ and Σ is outlined. The comparisons are made in terms of λ and Σ as these are common to all equations and are related to one another by $|\lambda|+|\Sigma| \le 1$. The response of different models can be compared and contrasted using these non-dimensional variables to determine how

they will influence EHL lubricated contacts as well as to give insight into the best approaches to solve the non-Newtonian problem. The terms that are compared are U_s, $U_{s\max}$, S and $S_{\max}$ for the Bair and Winer, Circular and Gecim and Winer models as well as the Cardiff $m = 0.3$ and $m = 0.7$ models.

Since only U_s and Σ are known, λ will have to be determined. Before a method to achieve this can be developed, knowledge of how the function U_s changes with λ and Σ is required. Figure 4.7 shows the variation of U_s for the Bair and Winer model, it should be noted that one axis is $\lambda/\lambda_{\max}$ where $\lambda_{\max} = 1 - \Sigma$ (taken from equation 3.23) and thus when $\lambda/\lambda_{\max} = 1$ the term that is being plotted is $U_{s\max}$ taken from section 3.6. It can be seen that the general trend is an increase in U_s for decreasing Σ and increasing λ.

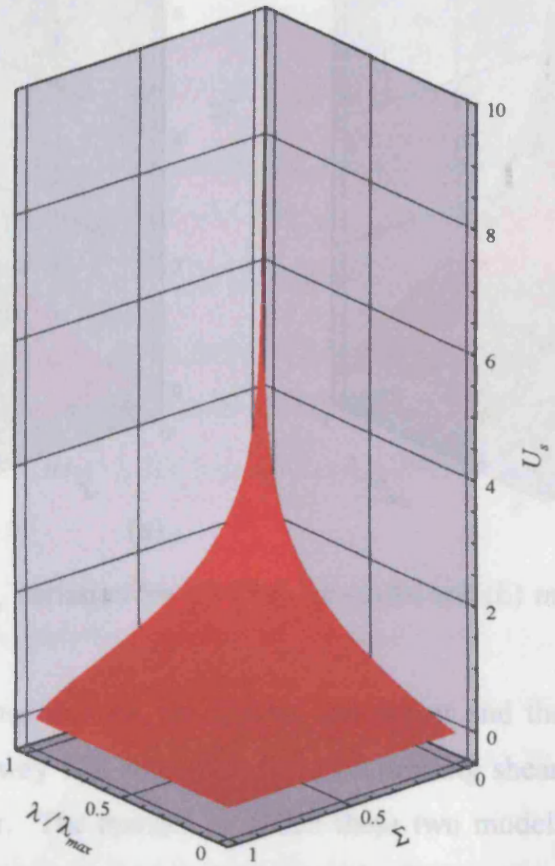

Figure 4.7 – U_s variation for Bair and Winer model

The same trend is evident in the Circular model, shown in Figure 4.8 (a), however the term increases more rapidly for decreasing Σ and increasing λ. The values plotted have been cut off at low Σ to prevent the values going above the axis limits, the value used was $\Sigma < 0.01$. Figure 4.8 (b) shows the response for the $m = 0.7$ model. This illustrates an even more dramatic increase in U_s than the Circular or Bair and Winer models show, there is also a cut off placed on this graph of $\Sigma < 0.05$.

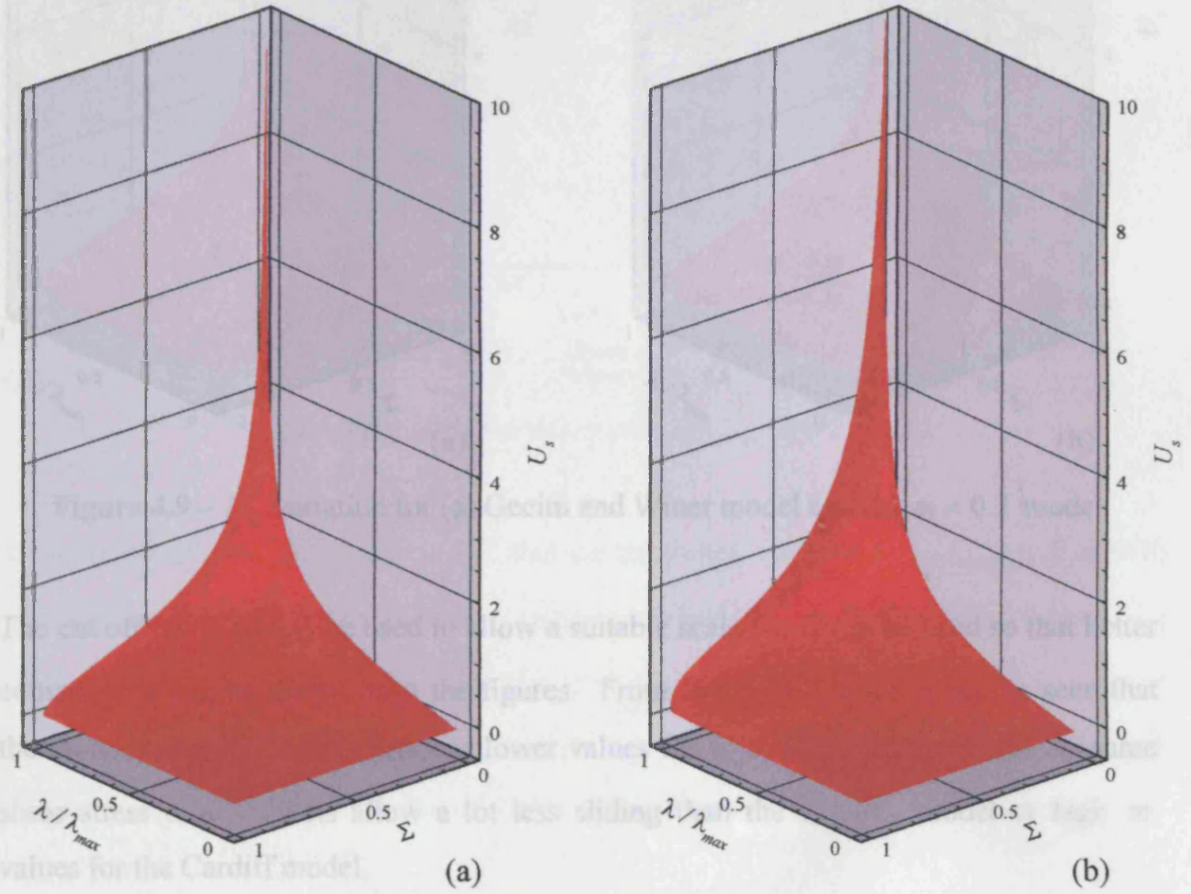

Figure 4.8 – U_s variation for (a) Circular model and (b) $m = 0.7$ model

In contrast to the previous models, the Gecim and Winer and the $m = 0.3$ models are more aggressive in the way that they approach the limiting shear stress than the three models consider thus far. The manner in which these two models approach the liming shear stress is very similar, as is evident in figure 1.2. The same trend is illustrated here, where their responses for U_s are very alike as shown in Figure 4.9.

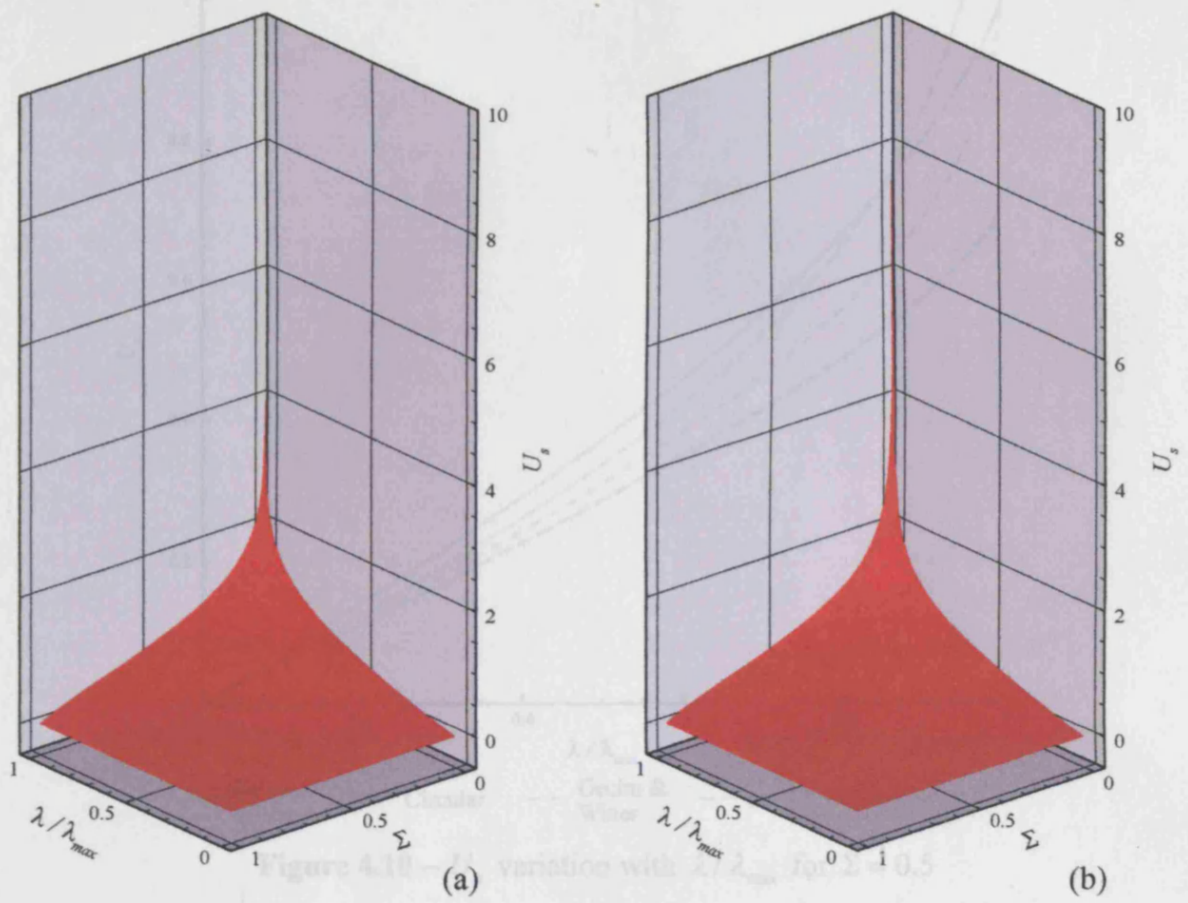

Figure 4.9 – U_s variation for (a) Gecim and Winer model and (b) $m = 0.3$ model

The cut off for Σ had to be used to allow a suitable scale for U_s to be used so that better comparisons can be drawn from the figures. From figures 4.7 to 4.9 it can be seen that the more aggressive models produce lower values for U_s. This means that for the same shear stress these models allow a lot less sliding than the Circular model or high m values for the Cardiff model.

This can be further illustrated by plotting U_s against λ/λ_{max} for a fixed value of Σ, which essentially is the same as taking sections through the previous figures. The value of Σ used is 0.5 and the results are shown in Figure 4.10. It can be seen that the different shape of the Bair and Winer model, shown in figure 1.1, leads to a slightly different response to the other models. The difference between the models becomes greater as Σ is reduced below this fixed value of 0.5.

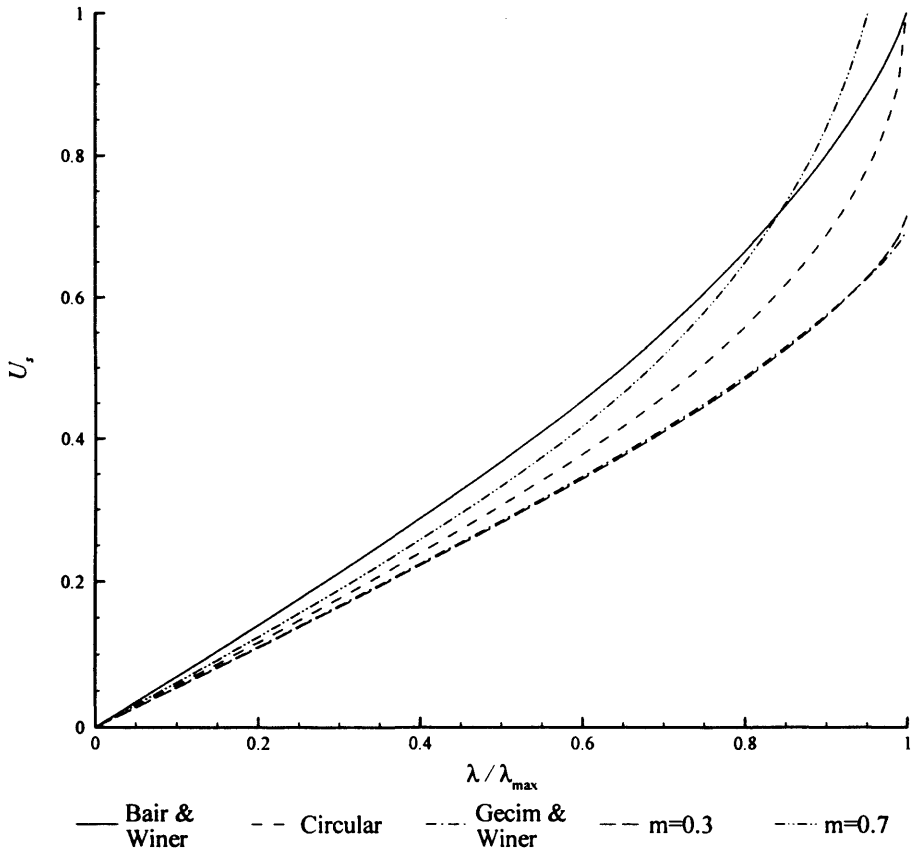

Figure 4.10 – U_s variation with λ / λ_{max} for $\Sigma = 0.5$

It can be seen from figures 4.7 to 4.10 that the maximum value of U_s, $U_{s\max}$, varies with Σ. The trend seen in these figures is an increase in $U_{s\max}$ as Σ decreases, this can be illustrated better by plotting $U_{s\max}$ against Σ which is shown in Figure 4.11. As expected the more aggressive the rheology the less sliding the model can allow, also the Bair and Winer model gives a different shaped curve to the other models.

The sliding that the surfaces impose on the oil will dictate what the shear stress distribution will be within the film. It can be seen that high sliding values can be achieved with lower shear stress distribution with less aggressive rheology, however, the influence that the shear stress has on the correction to the viscosity, S, still has to be established.

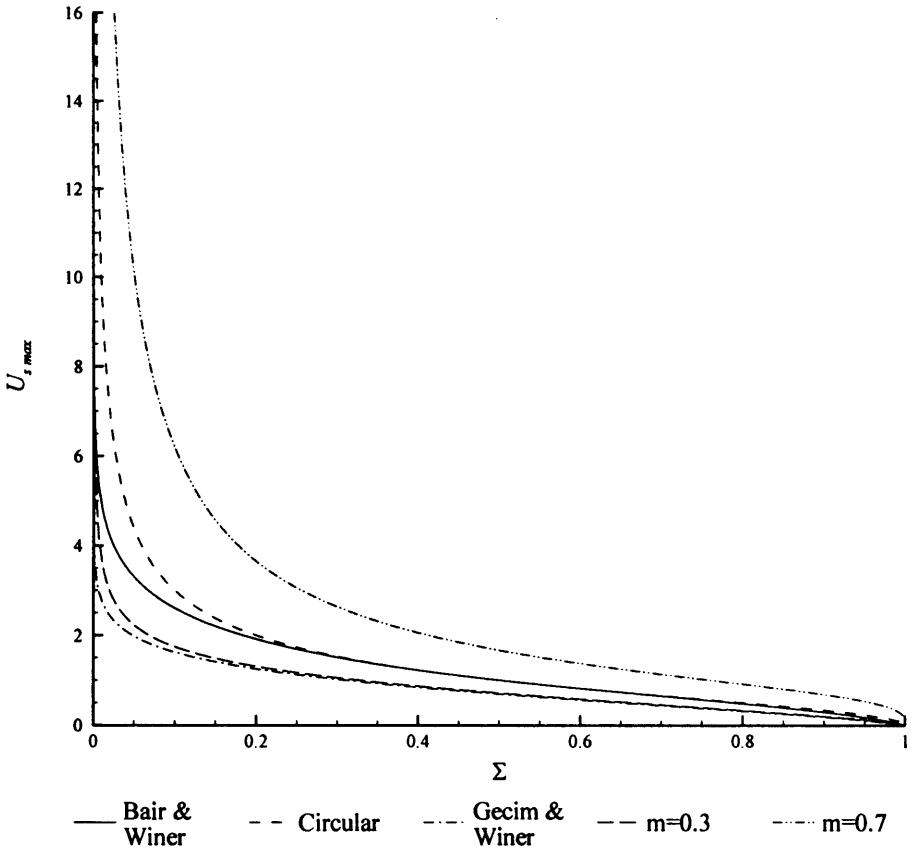

Figure 4.11 – $U_{s\,max}$ variation with Σ

Figure 4.12 shows the variation of S for varying combinations of shear stress, ie plotting against Σ and λ / λ_{max}, for the Bair and Winer model. The trend is similar to U_s where there is an increase for increasing λ and decreasing Σ, however S is more sensitive to this. S does not reach 0, instead its lowest value is 1, which is the same as a Newtonian response as there is no correction to the viscosity. It can be seen for low values of λ that there is a slight increase in S with increasing Σ. As there was for figure 4.8, there is a cut off used for Σ, for figure 4.12 the cut off is $\Sigma < 0.02$.

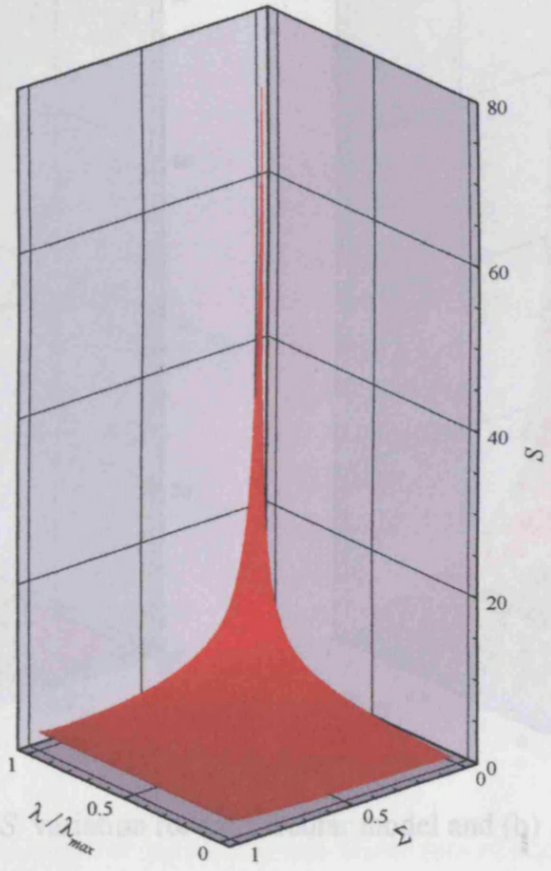

Figure 4.12 – S variation for Bair and Winer model

The same trends can be seen for the Circular and $m = 0.7$ models, shown in Figure 4.13. These models are less aggressive than the Bair and Winer model and it can be seen that S increases more for the less aggressive models, a similar trend to that seen for U_s. There is a cut off for the Circular model of $\Sigma < 0.05$ and a cut off of $\Sigma < 0.107$ for the $m = 0.7$ model.

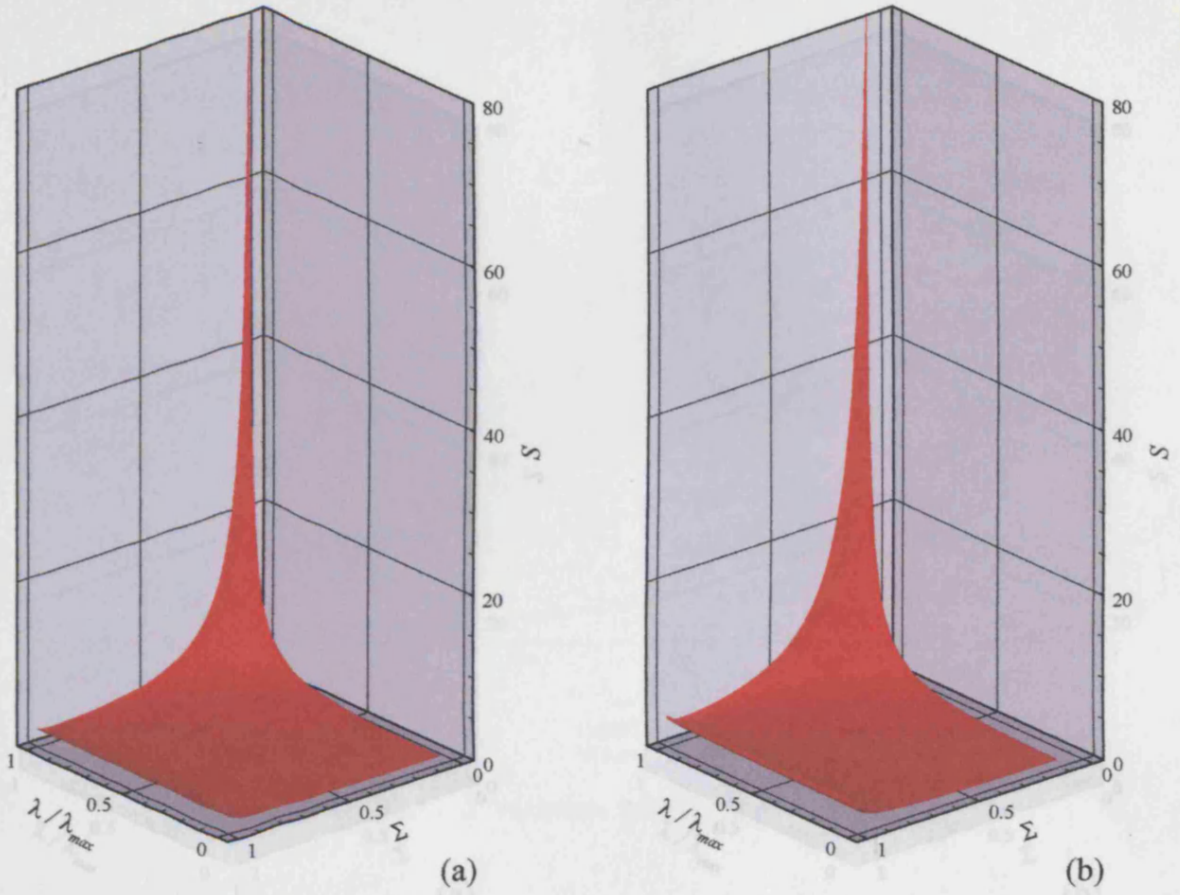

Figure 4.13 – S variation for (a) Circular model and (b) $m = 0.7$ model

Plotting the two more aggressive rheological models in Figure 4.14 it can be seen that the two models still give a similar response to one another. It can also be seen that the increase in S is a lot less pronounced than that seen for the less aggressive rheological models. There is no cut off imposed on either of these models.

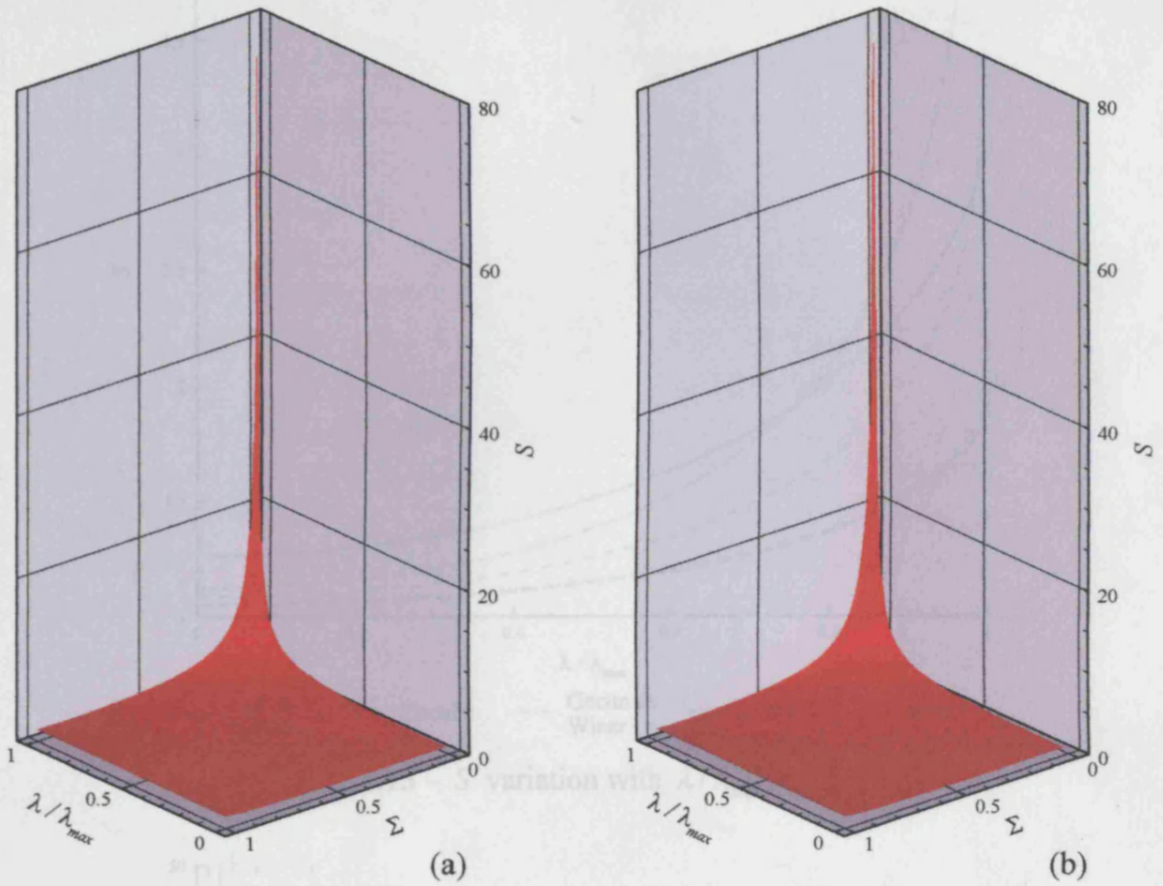

Figure 4.14 – S variation for (a) Gecim and Winer model and (b) $m = 0.3$ model

Figure 4.15 shows S against λ / λ_{max} for $\Sigma = 0.5$, a similar response to that seen in figure 4.10 is evident here. The Bair and Winer model has a different response to the other models and the less aggressive rheological models produce larger values of S than the more aggressive rheological models.

Figure 4.16 shows the variation of S_{max} against Σ, the trend is similar to figure 4.11, where there is an increase in S_{max} as Σ reduces. The more aggressive models have more dramatic change of slope at low Σ and produce lower values of S_{max} for all values of Σ. A feature that is evident in the $m = 0.7$ model, shown in figure 4.13 (b), is that there is a turning point as $\Sigma \Rightarrow 1$, this is still evident in figure 4.16. It can also be seen the the Bair and Winer model and Circular model exhibit the same trend, where there is a slight increase in S_{max} as $\Sigma \Rightarrow 1$.

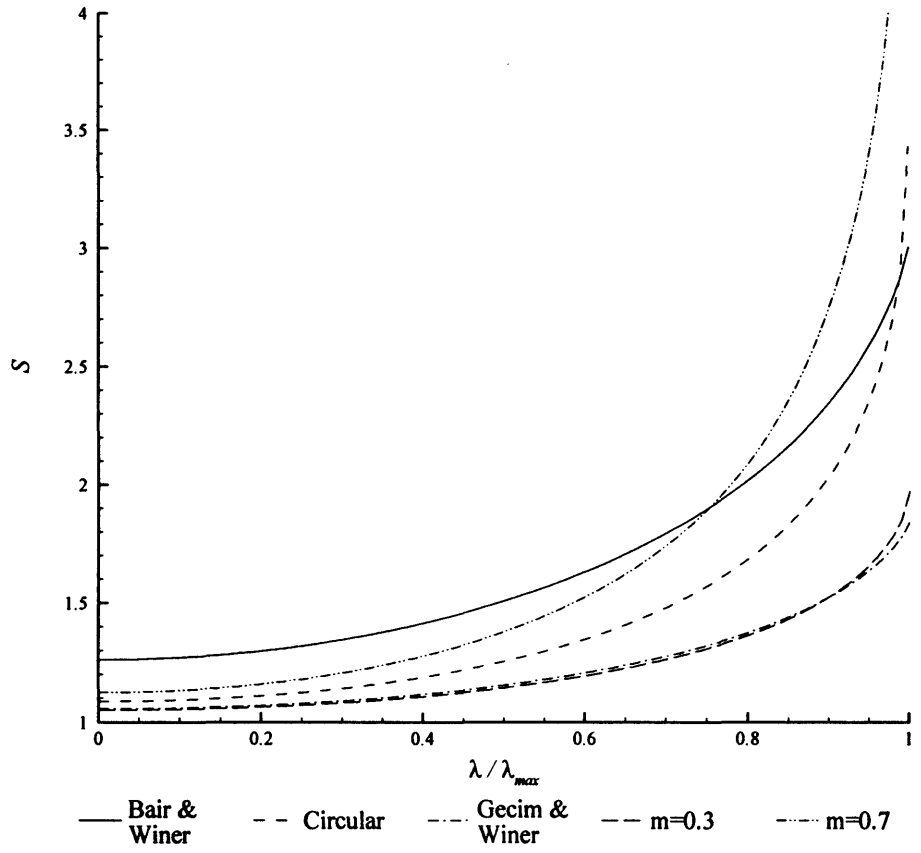

Figure 4.15 – S variation with λ / λ_{max} for $\Sigma = 0.5$

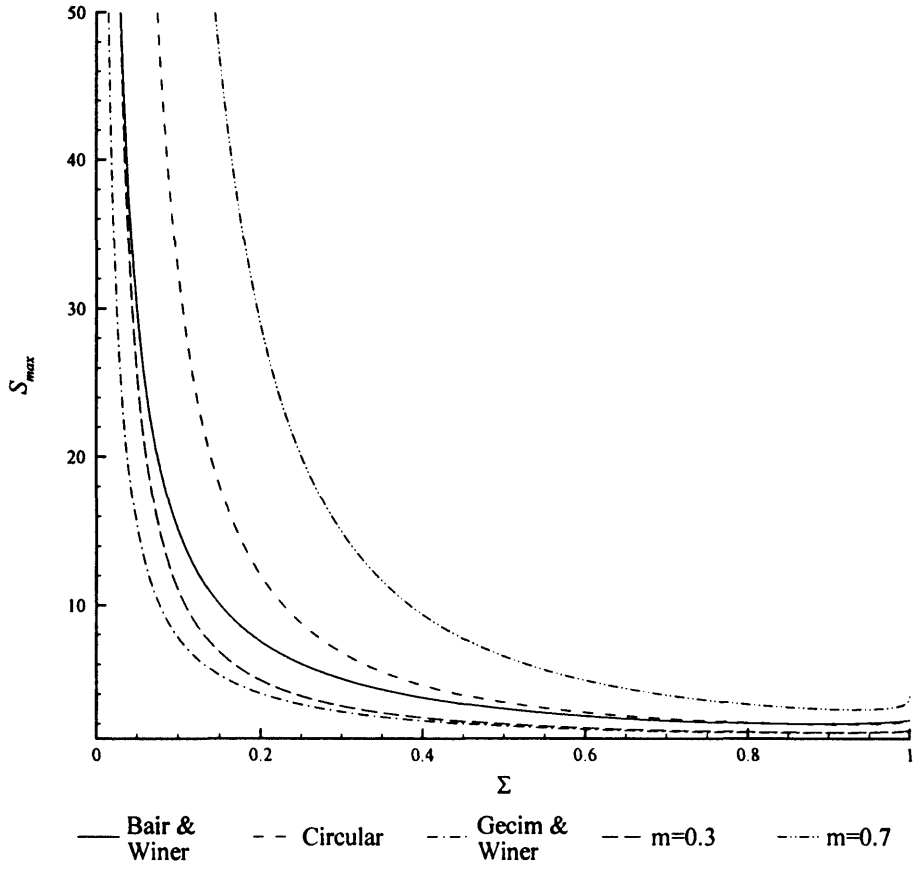

Figure 4.16 – S_{max} variation with Σ

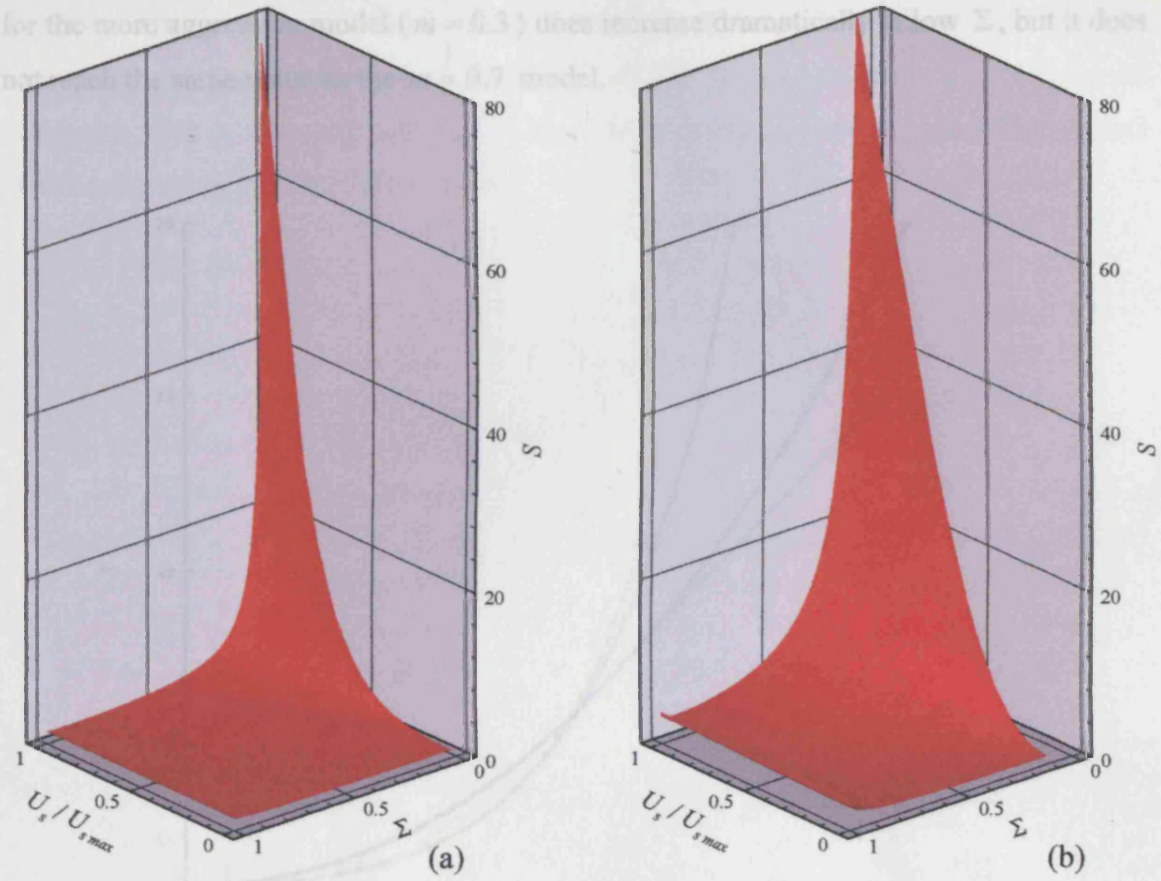

Figure 4.17 – S variation with $U_s / U_{s\max}$ for (a) $m = 0.3$ and (b) $m = 0.7$ models

It can be seen from figures 4.7 – 4.9 that different rheological models produce different shear stress variations through the oil film to deal with the conditions that are imposed on it by the surfaces. The more aggressive rheological models produce greater shear stresses for the same operating conditions. The influence that the shear stress distribution has on the effective viscosity via S has been determined in figures 4.12 – 4.14. Indirectly, the effect that the operating conditions (U_s and Σ) have been linked to S through the figures 4.7 – 4.9 and 4.12 – 4.14, but it is difficult to draw comparisons. To give a better indication of the effect that the operating conditions have on the viscosity correction, S can be plotted against U_s and Σ, which is shown in Figure 4.17 for the $m = 0.3$ and $m = 0.7$ models. The $m = 0.3$ and $m = 0.7$ models are chosen as the $m = 0.3$ model is one of the most aggressive models and the $m = 0.7$ is the least aggressive, thus the results will give a good indication of how the remainder of the models will respond. It is seen that the less aggressive model ($m = 0.7$) has a greater influence on the viscosity. S for

the $m = 0.7$ model begins to increase at a larger value of Σ than the $m = 0.3$ model. S for the more aggressive model ($m = 0.3$) does increase dramatically at low Σ, but it does not reach the same value as the $m = 0.7$ model.

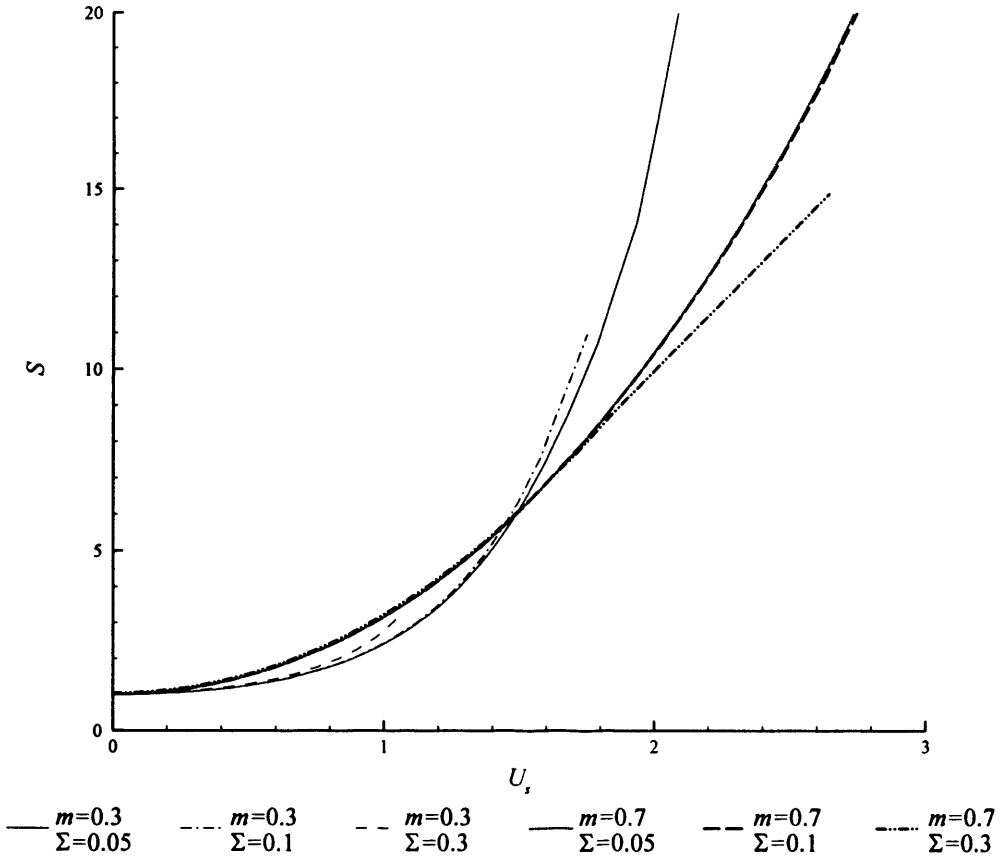

Figure 4.18 – S variation against U_s for fixed values of Σ

It was seen in figure 4.11 that less aggressive models withstand much more non-dimensional sliding without causing slip. The previous comparisons had the sliding scaled to the maximum, this does not allow a direct comparison between the actual conditions imposed and the effect that the different models have on viscosity. Figure 4.18 shows responses for several fixed values of Σ as U_s varies. The less aggressive model ($m = 0.7$) produces larger values of S than the more aggressive model ($m = 0.3$) at higher values of Σ. At low values of Σ and U_s, the more aggressive model develops higher values of S than the less aggressive model. There is a changeover in the responses, where the more aggressive model produces higher values of S than the less aggressive model when $U_s \approx 1.5$. This is due to the more aggressive model approaching

the maximum sliding and developing a large increase in S, the less aggressive model is far away from its maximum sliding so the increase in S is much more gradual. It can be seen that increasing the sliding imposed on the oil will produce an increase in S up to the maximum sliding, then slip will occur, the effects of this cannot be quantified without considering the complete EHL solution.

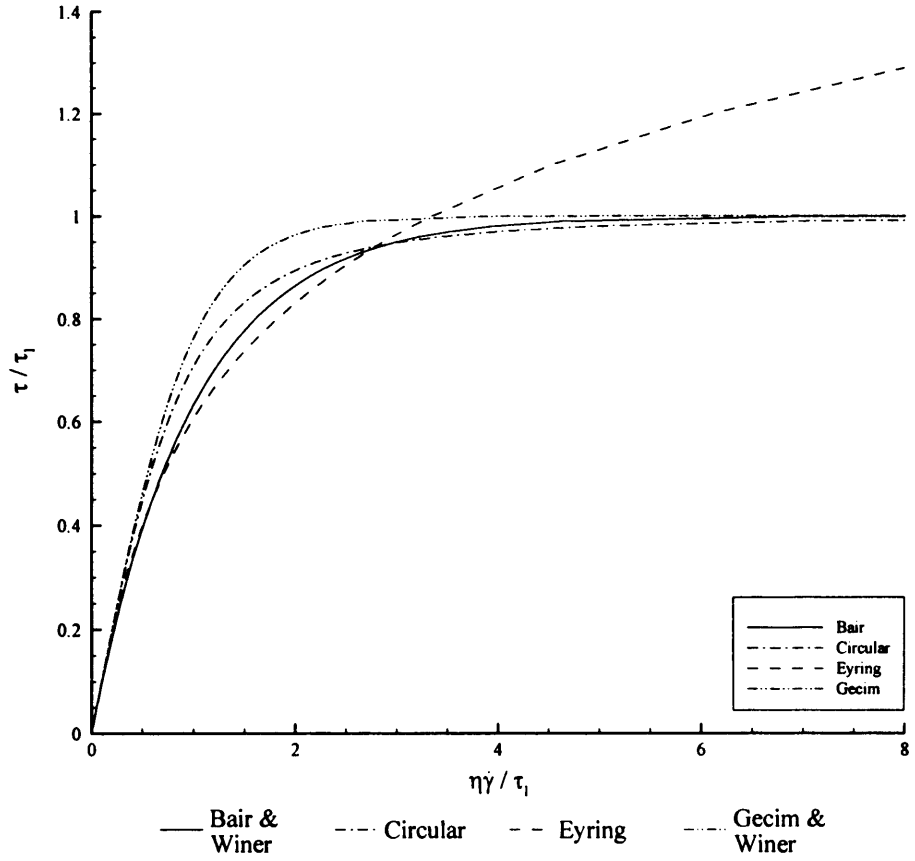

Figure 4.19 – Comparison of Eyring and limiting shear stress models where $\tau_l = 3\tau_0$

Although the various limiting models have been compared and contrasted the Eyring shear thinning rheological model has not been considered. A comparison requires the shear stress and shear strain response of the Eyring model to be somewhat similar to that of the limiting shear stress models. τ_0 and τ_l can be scaled to one another so that the limiting shear stress and Eyring models give similar responses, this is shown in Figure 4.19 where $\tau_l = 3\tau_0$.

To allow comparisons of the Eyring model to the limiting shear stress models the Eyring model has to be adapted so that it is formulated in terms of τ_l, λ and Σ. Considering the definition of Σ_E:

$$\Sigma_E = \frac{h}{2\tau_0}\frac{dp}{dx}$$

Substituting for τ_0:

$$\Sigma_E = \frac{3h}{2\tau_l}\frac{dp}{dx}$$

It follows that Σ_E and Σ can be related by:

$$\Sigma_E = 3\Sigma$$

Similarly λ_E is determined from:

$$\lambda_E = \frac{\tau_m}{\tau_0} = \frac{3\tau_m}{\tau_l}$$

Which allows λ_E and λ to be related by:

$$\lambda_E = 3\lambda$$

The sliding for the Eyring model was defined as:

$$u_s = \frac{\tau_0 h}{\Sigma_E \eta}\sinh(\lambda_E)\sinh(\Sigma_E)$$

Which can be modified substituting for τ_0, λ_E and Σ_E:

$$u_s = \frac{\tau_l h}{9\Sigma\eta}\sinh(3\lambda)\sinh(3\Sigma)$$

Allowing the non-dimensional sliding to be determined as:

$$U_s = \frac{1}{9\Sigma}\sinh(3\lambda)\sinh(3\Sigma)$$

Similarly, S can be expressed as:

$$S = \frac{[3\Sigma\cosh(3\Sigma) - \sinh(3\Sigma)]}{9\Sigma^3}\cosh(3\lambda)$$

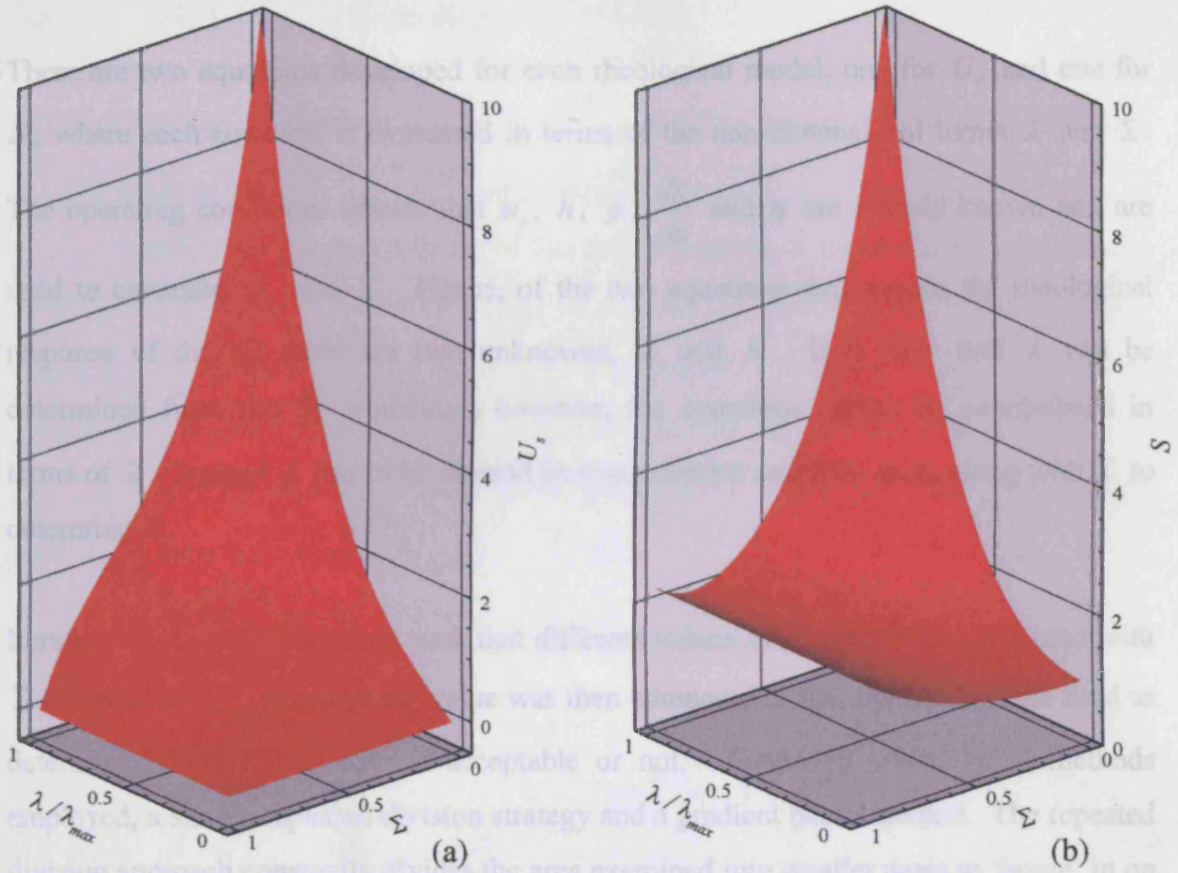

Figure 4.20 – Variation of (a) U_s and (b) S for Eyring model

These relationships can be used to compare the limiting shear stress models with the Eyring model. Figure 4.20 (a) shows the sliding velocity response for the Eyring model, it can be seen that the model allows much more sliding at larger values of λ and Σ than any of the limiting shear stress models. The results for the viscosity correction are shown in Figure 4.20 (b), it can be seen that the shape that is produced is somewhat similar to the Cardiff $m = 0.7$ model. However, it can also be seen that the results for the Eyring model do not reach the same extreme values that the $m = 0.7$ model gives.

Considering the response in figure 4.20, it can be seen that for a fixed sliding demand a smaller shear stress variation will result for the Eyring model compared to the limiting models. This smaller shear stress will also result in a much smaller viscosity correction compared to any of the limiting models. It can thus be seen that the limiting shear stress models will affect the EHL solution much more than the Eyring shear thinning model.

4.3 Non-Newtonian Solution Process

There are two equations developed for each rheological model, one for U_s and one for S, where each equation is expressed in terms of the non-dimensional terms λ and Σ. The operating conditions ensure that u_s, h, p, $\frac{dp}{dx}$ and η are already known and are used to calculate U_s and Σ. Hence, of the two equations that govern the rheological response of the oil there are two unknowns, λ and S. It is clear that λ can be determined from the U_s equations, however, the equations cannot be manipulated in terms of λ. Instead λ has to be iterated in some manner and then used, along with Σ to determine S.

Iteration of λ, was conducted such that different values of λ were taken and used with Σ to calculate U_s, the resulting value was then compared to that imposed on the fluid to determine whether the value is acceptable or not. There are two iterative methods employed, a simple repeated division strategy and a gradient based method. The repeated division approach constantly divides the area examined into smaller areas to 'home' in on the target value. The gradient based method calculates the gradient between two known values in the range and uses these to determine the new point, this is then used with one of the previous points to generate a new gradient and another point determined, this process is repeated until an answer is reached.

It is clear that this solution process is time consuming as for each point in the computational mesh λ has to be iterated to attain the correct sliding velocity. To overcome this an interpolation method was developed to attain the solution quicker. The known parameters, U_s and Σ, are used to determine the value of λ and S that satisfy the equations. Setting up a mesh in the U_s Σ plane, the values of λ and S are obtained at each 'mesh point' using the iteration techniques mentioned previously. The values, essentially, form a database or a two dimensional look up table that is used for interpolation. The time taken to carry out the calculations and their accuracy are compared for three possible approaches: iterating λ then using this value to calculate S;

interpolating λ then calculating S; and interpolating both λ and S. These comparisons allow a suitable strategy to be developed.

The variation of S with U_s and Σ was shown in figure 4.17 for the $m = 0.3$ and $m = 0.7$ models, where it was seen that a dramatic change in S occurs at small values of Σ. To try and make the interpolation of this function easier, it was decided to investigate interpolating in terms of $\ln(S)$ or $\log(S)$ as well as S. Interpolating using $\ln(S)$ and $\log(S)$ give the same result, but the two were compared to determine if one method was quicker than the other due to the intrinsic procedure used. It was found that using $\ln(S)$ was slightly quicker than $\log(S)$. Interpolating in terms of $\ln(S)$ gave results which were much better then interpolating in terms of S alone.

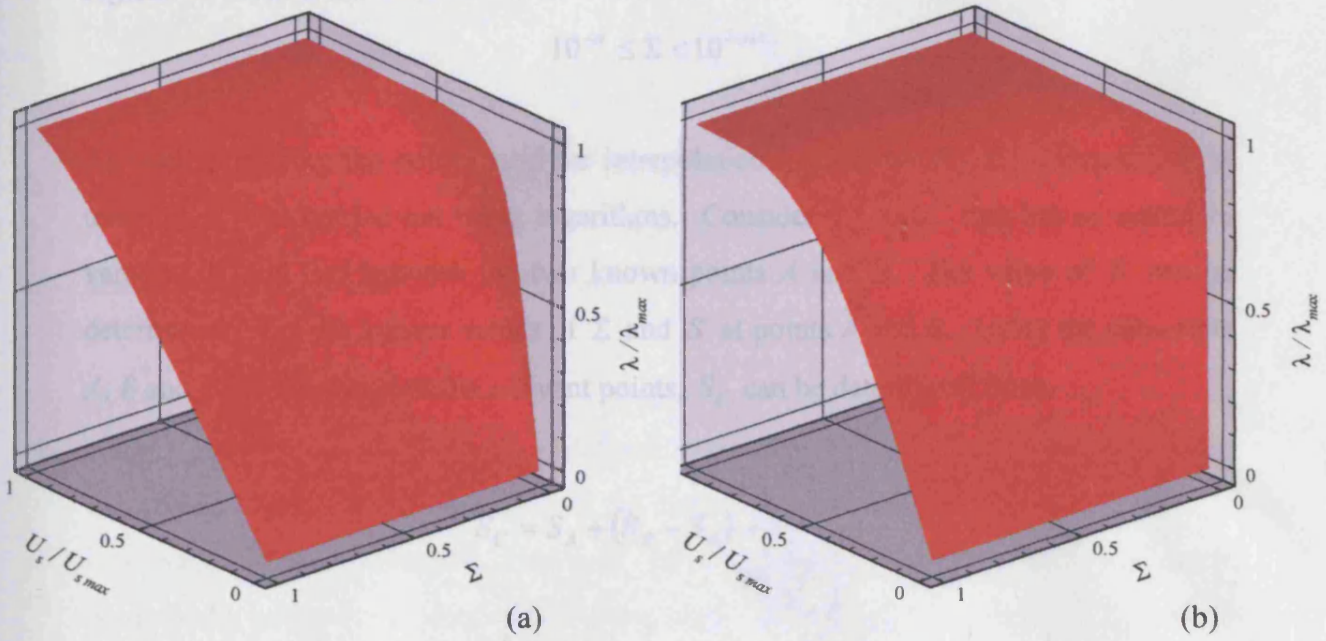

Figure 4.21 – λ variation with $U_s / U_{s\,max}$ for (a) $m = 0.3$ and (b) $m = 0.7$ models

The manner in which λ varies with U_s and Σ is illustrated for the $m = 0.3$ and $m = 0.7$ models in Figure 4.21, where it is seen that λ changes dramatically at low Σ. This is seen by the way that the function changes at the far top right of each figure, the $m = 0.7$ model shows a curve whereas the $m = 0.3$ shows a sharp transition. It is evident that interpolating these functions becomes more difficult as Σ reduces.

It was found that placing the points used for the interpolating in a regular arrangement, ie with all the points having the same spacing between them, was only suitable to describe the functions for reasonably large values of Σ. For smaller values of Σ the functions change rapidly and many more points at a much smaller spacing were necessary. To overcome this a different method of interpolation was proposed. The method employed was to space the points used for Σ with a varying distance between them. This was achieved by spacing them on a logarithmic scale. For example, the points can be spaced such that they correspond to the values 0.001, 0.002, 0.003....0.009, 0.010, 0.020....0.090, 0.100 and so on. This spacing can be achieved using two constants, a and b, in the following:

$$\Sigma = \left(9 \times \frac{b-1}{b_{\max}} + 1\right) \times 10^{-a}$$

For a fixed value of a, this equation spaces a number of points equal to $b_{\max}$ within the region:

$$10^{-a} \leq \Sigma < 10^{-a+1}$$

As well as spacing the points used for interpolation logarithmically the interpolation in terms of Σ was carried out using logarithms. Consider a point C that has an unknown value of S and lies between the two known points A and B. The value of S can be determined using the known values of Σ and S at points A and B. Using the subscripts A, B and C for the values at the relevant points, S_C can be determined from:

$$S_C = S_A + (S_B - S_A)\frac{\ln\left(\frac{\Sigma_B}{\Sigma_A}\right)}{\ln\left(\frac{\Sigma_C}{\Sigma_A}\right)}$$

It should be noted that the logarithmic interpolation methods were not used for $\Sigma \geq 0.1$ as many points were required, in addition there is the added problem of dealing with $\Sigma \Rightarrow 1$ where $\ln(\Sigma)$ approaches 0. Therefore, logarithmic interpolation was used for $\Sigma < 0.1$ and linear interpolation used for $\Sigma \geq 0.1$. The logarithmic approach was only used in Σ as the functions changed more dramatically in that direction. Linear interpolation was used for the interpolation in the U_s direction as the functions generally change less in this

direction. Methods based on U_s^2 and $\sqrt{U_s}$ were also investigated but their results were not as accurate as the linear interpolation method.

To assess the accuracy of the different solutions reference values needed to be established. As U_s is a function of λ and Σ, and S is a function of λ and Σ, the best approach was to use known values of λ and Σ to calculate initial values of U_s and S. The value of U_s calculated and Σ were treated as the known values in the interpolation and iteration methods. The interpolation and iteration methods return values of λ and S, which were then compared to the values calculated initially.

The value of S returned by the various methods, termed S_{calc}, could be compared to the value calculated on the mesh values of λ and Σ, termed S_{orig}. This allows a percentage error to be determined from:

$$error = 100 \times \frac{|S_{orig} - S_{calc}|}{S_{orig}}$$

The error was based on S as this function is highly non-linear and hence the error produced will be greater than if it was based on λ.

The mesh used was set up in terms of Σ from 0 to 1. For each Σ value the mesh was set up for λ ranging from 0 to $1-\Sigma$, as defined by equation (3.23). The maximum error for each method could be determined over the mesh and used to get a general indication of how accurate each method is. However, this can be slightly misleading as one errant point could cause the method to be discounted. It was therefore decided to include another measure of accuracy where the measure of the variation of the error was assessed over the whole region. This was achieved by recording the number of times the error produced was above threshold levels. The count could then be related to a percentage of the total number of points, which is the percentage of the total area of the whole computational region above the threshold. The threshold levels taken for the analysis corresponded to errors of 0.1% and 0.5%.

To establish the accuracy of each method, an analytical mesh of 1500 points in Σ and 1000 points in U_s was used to test several rheological models. The models analysed are the Bair and Winer model, Gecim and Winer model, Hamrock's circular model as well as the Cardiff model for $m = 0.2, 0.3, 0.4, 0.5, 0.6, 0.7$ and 0.8.

Considering the accuracy for the Bair and Winer model, shown in Table 4.1, it can be seen that interpolation gives results that are reasonably accurate. Interpolating λ and using this to calculate S gives slightly better overall response than interpolating both values, seen by a smaller area above the threshold levels. However, this method does produce a much larger maximum error than interpolating both values. It should be noted that the area above the 0.5% cut off point for the logarithmic interpolation for the Interpolate λ, Calc S method is 0.0002%, essentially rounding to 0. Linear and log are used to distinguish between the two regions where the different interpolation methods are employed.

Table 4.1 – Accuracy for Bair and Winer model

Method	Max Error / %		Points above threshold value / %			
			0.1%		0.5%	
	Linear	Log	Linear	Log	Linear	Log
Iterate λ, Calc S	0.014	0.033	0.000	0.000	0.000	0.000
Interpolate λ, Calc S	0.096	14.286	0.000	0.044	0.000	0.000
Interpolate λ and S	0.125	0.379	0.014	0.115	0.000	0.000

The Circular model gives a similar trend, shown in Table 4.2. In this case the maximum errors for all but one of the interpolation methods are higher than in table 4.1. For this model it is generally more accurate to interpolate both values as interpolating λ and using this to calculate S gives slightly higher errors. Interpolating both values does produce a slightly higher maximum error than the other interpolation method employed.

Table 4.2 – Accuracy for the Circular model

Method	Max Error / %		Points above threshold value / %			
			0.1%		0.5%	
	Linear	Log	Linear	Log	Linear	Log
Iterate λ, Calc S	0.013	0.015	0.000	0.000	0.000	0.000
Interpolate λ, Calc S	0.287	10.427	0.428	0.203	0.000	0.057
Interpolate λ and S	0.331	11.696	0.334	0.160	0.000	0.054

Table 4.3 gives the results for the Gecim and Winer model. Interpolating both λ and S gives lower errors over the range considered than interpolating λ and calculating S for both regions. The maximum error for the logarithmic range is higher than the linear range. Interpolating both values, although generally more accurate, does produce a higher maximum error than interpolating λ and calculating S. The area above the cut off values for the logarithmic interpolation of both λ and S is 0.00033% and 0.00013% for the cut of values 0.1% and 0.5% respectively.

Table 4.3 – Accuracy for Gecim and Winer model

Method	Max Error / %		Points above threshold value / %			
			0.1%		0.5%	
	Linear	Log	Linear	Log	Linear	Log
Iterate λ, Calc S	0.018	0.034	0.000	0.000	0.000	0.000
Interpolate λ, Calc S	0.546	0.647	0.252	0.213	0.003	0.003
Interpolate λ and S	0.203	2.243	0.064	0.000	0.000	0.000

The Cardiff model produced different results dependent on the value of m used, this is illustrated in Tables 4.4 to 4.10. For low values of m the functions are easier to interpolate, however, as m increases the functions become increasingly difficult to interpolate. This is illustrated by a decrease in the accuracy obtained by iterative methods, which is due to the rise in the amount of sliding tolerated. To try and attain the best accuracy possible it was necessary to change the mesh on which interpolation was conducted once $m \geq 0.6$. For the higher power m values it was necessary to increase the number of points used for the interpolation along the velocity axis.

Changing the mesh of values on which the interpolation is conducted provided significant improvement in accuracy, but it should be noted that it is very difficult to get the interpolation to perform to the same accuracy as the iterative method. This is due to the fact that the points used for the interpolation are determined using the iterative method. Taking this into account it can be seen that the higher powers of m do give good accuracy when both values are interpolated in the linear range. This is particularly evident for the case where $m = 0.8$, shown in table 4.10.

The logarithmic region changes more rapidly than the linear region which causes the maximum error produced by interpolation to become high. Despite this, the results for $m \leq 0.7$ illustrate that the area that is above the 0.5% cut off values is still less than 0.25% and just over 1.5% of the total area considered is over the 0.1% cut off point. Due to the way in which S varies interpolating λ then using this to calculate S amplifies any slight error. This is seen, particularly in the linear region, as a very large proportion of the area is over the 0.5% cut off level for the interpolate λ, calc S method.

Table 4.4 – Accuracy for $m = 0.2$

Method	Max Error / %		Points above threshold value / %			
			0.1%		0.5%	
	Linear	Log	Linear	Log	Linear	Log
Iterate λ, Calc S	0.023	0.030	0.000	0.000	0.000	0.000
Interpolate λ, Calc S	9.082	0.107	42.338	0.000	13.103	0.000
Interpolate λ and S	0.229	0.217	0.125	0.001	0.000	0.000

Table 4.5 – Accuracy for $m = 0.3$

Method	Max Error		Points above threshold value / %			
			0.1%		0.5%	
	Linear	Log	Linear	Log	Linear	Log
Iterate λ, Calc S	0.043	0.021	0.000	0.000	0.000	0.000
Interpolate λ, Calc S	7.612	0.166	51.624	0.012	20.062	0.000
Interpolate λ and S	0.087	0.202	0.000	0.000	0.000	0.000

Table 4.6 – Accuracy for $m = 0.4$

Method	Max Error / %		Points above threshold value / %			
			0.1%		0.5%	
	Linear	Log	Linear	Log	Linear	Log
Iterate λ, Calc S	0.105	0.017	0.006	0.000	0.000	0.000
Interpolate λ, Calc S	7.490	1.177	58.679	0.057	26.877	0.021
Interpolate λ and S	0.196	1.093	0.076	0.051	0.000	0.026

Table 4.7 – Accuracy for $m = 0.5$

Method	Max Error / %		Points above threshold value / %			
			0.1%		0.5%	
	Linear	Log	Linear	Log	Linear	Log
Iterate λ, Calc S	0.235	0.015	0.131	0.000	0.000	0.000
Interpolate λ, Calc S	9.241	10.427	64.582	0.106	33.742	0.057
Interpolate λ and S	0.407	11.696	0.477	0.087	0.000	0.054

Table 4.8 – Accuracy for $m = 0.6$

Method	Max Error / %		Points above threshold value / %			
			0.1%		0.5%	
	Linear	Log	Linear	Log	Linear	Log
Iterate λ, Calc S	0.494	0.013	0.275	0.000	0.000	0.000
Interpolate λ, Calc S	6.193	7.607	59.147	0.213	25.375	0.055
Interpolate λ and S	0.788	31.638	0.363	0.177	0.026	0.046

Table 4.9 – Accuracy for $m = 0.7$

Method	Max Error / %		Points above threshold value / %			
			0.1%		0.5%	
	Linear	Log	Linear	Log	Linear	Log
Iterate λ, Calc S	0.978	0.125	1.036	0.003	0.093	0.000
Interpolate λ, Calc S	9.369	56.837	65.659	0.299	33.277	0.062
Interpolate λ and S	1.441	41.763	1.430	0.293	0.155	0.066

Table 4.10 – Accuracy for $m = 0.8$

Method	Max Error / %		Points above threshold value / %			
			0.1%		0.5%	
	Linear	Log	Linear	Log	Linear	Log
Iterate λ, Calc S	2.926	2.641	11.140	0.155	3.885	0.090
Interpolate λ, Calc S	15.853	96.131	72.517	0.681	43.847	0.219
Interpolate λ and S	2.846	70.429	12.390	0.661	3.780	0.223

All the models show that reasonable accuracy can be achieved using the interpolation methods. The Gecim and Winer and Bair and Winer are the easiest models to interpolate. The Circular and Cardiff models, which are based on the same expression, are more difficult to interpolate. Due to the increased difficulty a greater number of points have to be used in the interpolation meshes. The actual number of points that had to be used was up to 4 times greater for the Σ axis in the logarithmic range and the U_s axis in both ranges.

A further comparison that could be made is to compare the methods of calculation for $m = 0.5$ to the Circular model and is shown in Table 4.11. It can be seen that interpolating λ and S gives similar results to iterating λ and calculating S.

Table 4.11 – Accuracy for $m = 0.5$ compared to Circular model

Method	Max Error		Area above Cut off value			
			0.1%		0.5%	
	Linear	Log	Linear	Log	Linear	Log
Iterate λ, Calc S	1.823	43.929	13.231	0.044	0.109	0.010
Interpolate λ, Calc S	9.240	43.929	69.023	0.151	34.088	0.067
Interpolate λ and S	1.821	43.930	13.229	0.130	0.105	0.064

The interpolation method's greatest strength is in reducing the calculation time. The Eyring rheological model is one of the simplest models to implement and the current version of the software uses an optimised method based on this rheological model. It was used as the benchmark for the timing of all of the limiting shear stress rheological models.

Table 4.12 shows the time taken in seconds for 50000 calls to the subroutine, where the various values were taken over the range of U_s and Σ. The different times taken to interpolate both λ and S occur due to the time required to calculate the maximum sliding. It can also be seen that the calculation of S, equation (3.22), for the Cardiff model takes a substantial amount of time, due to the numerical integration required.

Table 4.12 – Time in seconds for 50000 calculations of λ and S for the different methods

Model	Iterating λ and S	Interpolating λ, calculating S	Interpolating λ and S	Eyring Model (Benchmark)
Bair	0.339	0.083	0.062	0.047
Circular	0.484	0.105	0.061	0.047
Gecim	0.383	0.107	0.068	0.047
$m = 0.3$	1.250	0.689	0.069	0.047
$m = 0.5$	1.493	0.691	0.069	0.047
$m = 0.7$	2.012	0.691	0.069	0.046

Table 4.13 shows the percentage time difference between the various methods and the benchmark. It is seen that interpolating both λ and S is evidently a lot quicker than the other methods employed. The Cardiff model clearly benefits the most from interpolating.

Table 4.13 – Percentage difference between different methods and benchmark

Model	Iterating λ and S	Interpolating λ, calculating S	Interpolating λ and S
Bair	628.1	78.5	32.9
Circular	941.1	124.8	30.9
Gecim	724.4	131.1	46.4
$m = 0.3$	2588.4	1381.8	47.3
$m = 0.5$	3109.4	1385.8	47.6
$m = 0.7$	4225.9	1385.5	47.5

It is apparent that interpolating both λ and S has a significant saving on computation time. It also produces an acceptable error, where less than 0.5% of the total area has an

error of over 0.1% for the Bair and Winer, Circular, Gecim and Winer and $m < 0.5$ models. The other values of m still produce very good accuracy, particularly taking into account the error that is present in the iteration method. Another feature of interpolation is that it gives a smooth transformation from point to point. Iteration methods could converge low for one point and high for the next, creating a step, this could promote instability within the hydrodynamic equation.

Now that the methods have been introduced and validated, the actual solution method needs to be defined. The film thickness, pressure and sliding speeds are known. From this the viscosity, limiting shear stress and the pressure gradient can be calculated. These are used to determine Σ and U_s. The way these are used to determine the other variables is outlined in the flow chart in Figure 4.22.

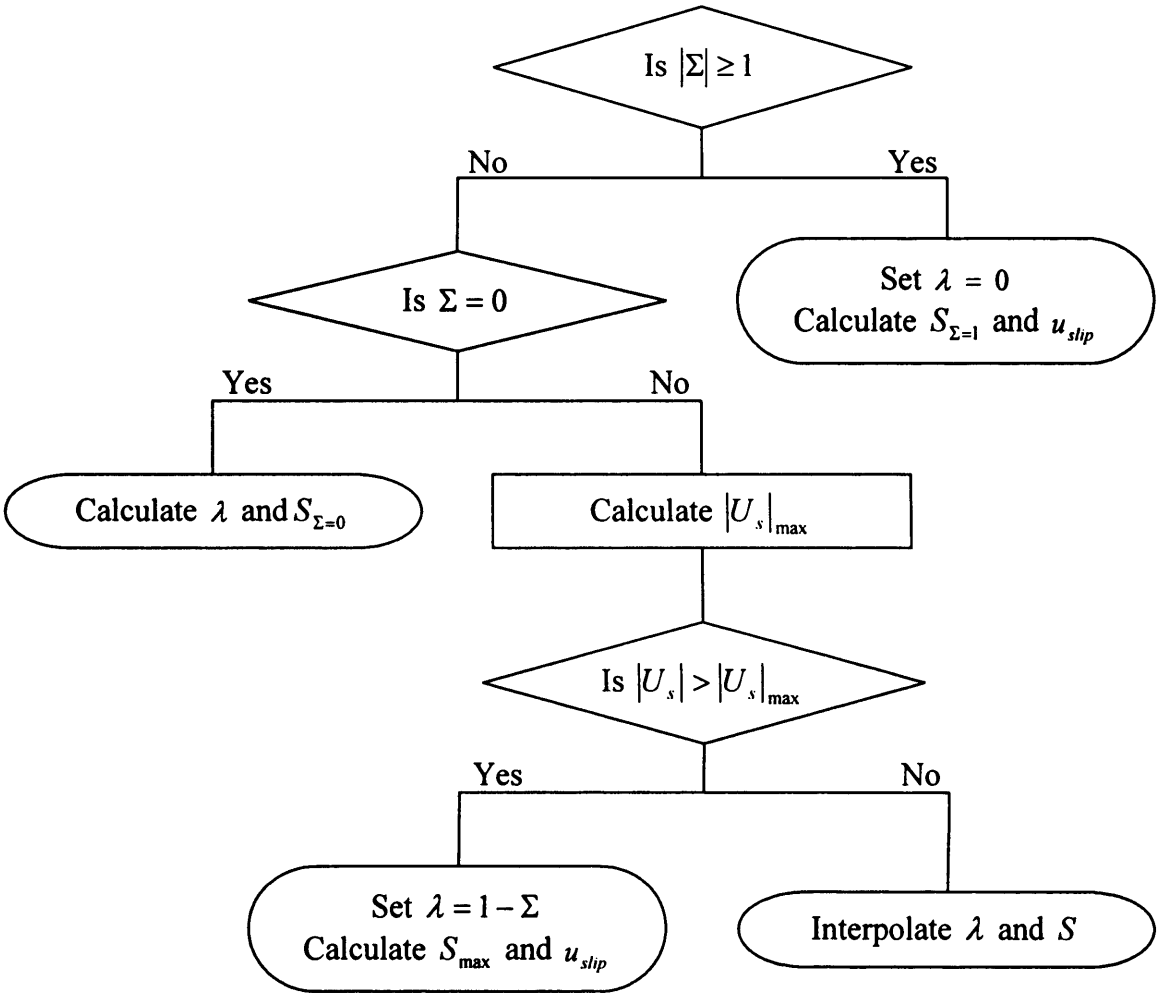

Figure 4.22 – Flow Chart showing solution process for limiting shear stress models

4.4 Thermal Solution Technique

The thermal equations are introduced in section 2.7. These equations are formulated in columns through the contact area where they form a tri-diagonal matrix. This matrix is solved as a whole and the next column is then considered. The major complication in this method is the linking of the fluid and solid equations at the surfaces.

The approach used has benefited from the thermal analysis technique developed by Sharif [Sharif et al., 2001a, 2001b, 2001c, 2004]. In the current thesis the need to deal with transient situations incorporating direct contact between the solid bodies at asperities led to adaptation and development of these methods.

The approach for all regions, except those near the surface, is to adopt a control volume method at each point in the mesh. In each volume the various terms are specified by finite difference schemes and the time dependence of the temperature is determined by a Crank-Nicholson scheme.

Equation (2.18) can be written as:

$$\rho c\frac{d\theta}{dt}=k\left(\frac{d^2\theta}{dx^2}+\frac{d^2\theta}{dz^2}\right)-\rho cu\frac{d\theta}{dx}+\varepsilon\left(u\frac{dp}{dx}+\frac{dp}{dt}\right)\theta+\tau\dot{\gamma}=eqn$$

Evaluating the time dependent terms using a Crank-Nicholson scheme and adopting a time averaged density $\hat{\rho}$ gives:

$$\frac{\hat{\rho}c}{\delta t}\left(\theta^{t}-\theta^{t-1}\right)=\frac{eqn^{t}}{2}+\frac{eqn^{t-1}}{2}$$

The specific heat, c, is not time averaged as it is fixed and is not considered to change with pressure or temperature.

Re-arranging this and using finite difference approximations for the remaining terms the fluid thermal equation can be written as:

$$0=\frac{k}{\delta z^2}\left(\theta_{i,j-1}+\theta_{i,j+1}-2\theta_{i,j}\right)+\frac{k}{\delta x^2}\left(\theta_{i-1,j}+\theta_{i+1,j}-2\theta_{i,j}\right)-\frac{\rho cu}{\delta x}\left(\theta_{i,j}-\theta_{i-1,j}\right)$$
$$+\varepsilon\left(u\frac{dp}{dx}+\frac{dp}{dt}\right)\theta_{i,j}-2\frac{\hat{\rho}c}{\delta t}\theta_{i,j}+\tau\dot{\gamma}+rhs$$

Where *rhs* is the influence of the terms at the previous time step and is given by:

$$rhs = \frac{k}{\delta z^2}\left(\theta_{i,j-1}^{t-1} + \theta_{i,j+1}^{t-1} - 2\theta_{i,j}^{t-1}\right) + \frac{k}{\delta x^2}\left(\theta_{i-1,j}^{t-1} + \theta_{i+1,j}^{t-1} - 2\theta_{i,j}^{t-1}\right) - \frac{\rho c u}{\delta x}\left(\theta_{i,j}^{t-1} - \theta_{i-1,j}^{t-1}\right)$$
$$+ \varepsilon\left(u\frac{dp}{dx} + \frac{dp}{dt}\right)^{t-1}\theta_{i,j}^{t-1} + 2\frac{\hat{\rho}c}{\delta t}\theta_{i,j}^{t-1} + \tau\dot{\gamma}^{t-1}$$

Similar equations can be developed for the solids:

$$0 = \frac{k}{\delta z^2}\left(\theta_{i,j-1} + \theta_{i,j+1} - 2\theta_{i,j}\right) + \frac{k}{\delta x^2}\left(\theta_{i-1,j} + \theta_{i+1,j} - 2\theta_{i,j}\right)$$
$$- \frac{\rho c u}{\delta x}\left(\theta_{i,j} - \theta_{i-1,j}\right) - 2\frac{\rho c}{\delta t}\theta_{i,j} + rhs$$

$$rhs = \frac{k}{\delta z^2}\left(\theta_{i,j-1}^{t-1} + \theta_{i,j+1}^{t-1} - 2\theta_{i,j}^{t-1}\right) + \frac{k}{\delta x^2}\left(\theta_{i-1,j}^{t-1} + \theta_{i+1,j}^{t-1} - 2\theta_{i,j}^{t-1}\right)$$
$$- \frac{\rho c u}{\delta x}\left(\theta_{i,j}^{t-1} - \theta_{i-1,j}^{t-1}\right) + 2\frac{\rho c}{\delta t}\theta_{i,j}^{t-1}$$

There is no time average on the density of the solids as it is considered constant. It should also be noted that in all of the above a backward difference scheme was used for the convection terms, but a forward difference scheme was used when the velocity was negative.

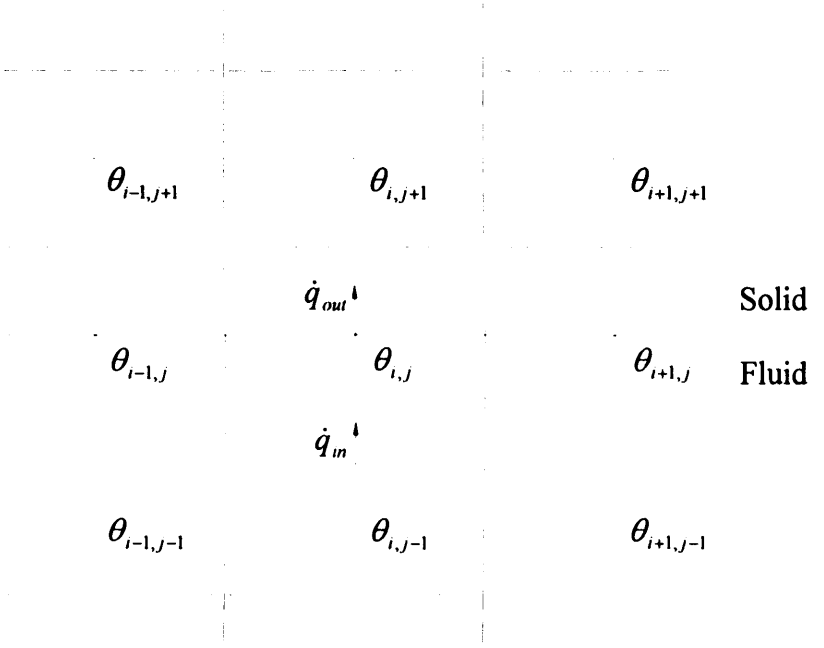

Figure 4.23 – Control volumes and temperature points used in finite difference schemes

In linking the two equations the problem must be addressed in a slightly different manner. There is no longer a control volume centred at the temperature mesh point as this would contain both the solid and the fluid, instead two 'half volumes' are considered. Figure 4.23 illustrates this for the upper surface interface, also shown are the fluxes that flow into and out of the fluid half volume.

In formulation of the terms for the tri-diagonal matrix the two half volumes must be linked, this is achieved using the conservation of flux between the solid and fluid. To enable this, the conduction term in the z direction for each half volume must be expressed in a different manner:

$$k\frac{d^2\theta}{dz^2} = \frac{d}{dz}(-\dot{q}) = 2\frac{\dot{q}_{in} - \dot{q}_{out}}{\delta z}$$

For the case of the last fluid volume at the upper surface, where the fluxes $\dot{q}_{in}$ and $\dot{q}_{out}$ are shown in figure 4.23, the energy equation can be written as:

$$0 = \frac{2}{\delta z}(\dot{q}_{in} - \dot{q}_{out}) + \frac{k}{\delta x^2}(\theta_{i-1,j} + \theta_{i+1,j} - 2\theta_{i,j}) - \frac{\rho c u}{\delta x}(\theta_{i,j} - \theta_{i-1,j}) + \varepsilon\left(u\frac{dp}{dx} + \frac{dp}{dt}\right)\theta_{i,j} - 2\frac{\hat{\rho}c}{\delta t}\theta_{i,j} + \tau\dot{\gamma} + rhs$$

If the flux flowing out of the fluid, $\dot{q}_{out}$, is replaced with $\dot{q}_2$ and $\dot{q}_{in}$ is replaced by a finite difference formulation, then the above equation can be re-written as:

$$\dot{q}_2 = -k\frac{\theta_{i,j} - \theta_{i,j-1}}{\delta z} + \frac{\delta z}{2}\left[\begin{array}{l}\frac{k}{\delta x^2}(\theta_{i-1,j} + \theta_{i+1,j} - 2\theta_{i,j}) - \frac{\rho c u}{\delta x}(\theta_{i,j} - \theta_{i-1,j}) \\ + \varepsilon\left(u\frac{dp}{dx} + \frac{dp}{dt}\right)\theta_{i,j} - 2\frac{\hat{\rho}c}{\delta t}\theta_{i,j} + \tau\dot{\gamma} + rhs\end{array}\right] \qquad (4.1)$$

In a similar manner the energy equation for the upper solid surface half volume can be written as:

$$0 = \frac{2}{\delta z}(\dot{q}_{in} - \dot{q}_{out}) + \frac{k}{\delta x^2}(\theta_{i-1,j} + \theta_{i+1,j} - 2\theta_{i,j}) - \frac{\rho c u}{\delta x}(\theta_{i,j} - \theta_{i-1,j}) - 2\frac{\rho c}{\delta t}\theta_{i,j} + rhs$$

Defining the flux that flows into the solid half volume, $\dot{q}_{in}$, as $\dot{q}_1$ and replacing $\dot{q}_{out}$ with a finite difference scheme, the above equation can be written in terms of $\dot{q}_1$ as:

$$-\dot{q}_1 = k\frac{\theta_{i,j+1}-\theta_{i,j}}{\delta z} + \frac{\delta z}{2}\left[\begin{array}{c}\frac{k}{\delta x^2}\left(\theta_{i-1,j}+\theta_{i+1,j}-2\theta_{i,j}\right)-\frac{\rho c u}{\delta x}\left(\theta_{i,j}-\theta_{i-1,j}\right)\\ -2\frac{\rho c}{\delta t}\theta_{i,j}+rhs\end{array}\right] \tag{4.2}$$

Due to the conservation of flux, $\dot{q}_1 = \dot{q}_2$, equations (4.1) and (4.2) can be added together to essentially form a complete volume that spans both the fluid and the solid:

$$\begin{aligned}0 = &-k_f\frac{\theta_{i,j}-\theta_{i,j-1}}{\delta z_f} + \frac{\delta z_f}{2}\left[\begin{array}{c}\frac{k_f}{\delta x^2}\left(\theta_{i-1,j}+\theta_{i+1,j}-2\theta_{i,j}\right)-\frac{\rho_f c_f u_f}{\delta x}\left(\theta_{i,j}-\theta_{i-1,j}\right)\\ +\varepsilon\left(u_f\frac{dp}{dx}+\frac{dp}{dt}\right)\theta_{i,j}-2\frac{\hat{\rho}_f c_f}{\delta t}\theta_{i,j}+\tau\dot{\gamma}+rhs_f\end{array}\right]\\ &+k_s\frac{\theta_{i,j+1}-\theta_{i,j}}{\delta z_s} + \frac{\delta z_s}{2}\left[\begin{array}{c}\frac{k_s}{\delta x^2}\left(\theta_{i-1,j}+\theta_{i+1,j}-2\theta_{i,j}\right)-\frac{\rho_s c_s u_1}{\delta x}\left(\theta_{i,j}-\theta_{i-1,j}\right)\\ -2\frac{\rho_s c_s}{\delta t}\theta_{i,j}+rhs_s\end{array}\right]\end{aligned}$$

If slip occurs at the surface, with a velocity u_{slip1}, then there is no conservation of flux, instead equation (3.26) can be applied:

$$\begin{aligned}-\tau_1 u_{slip1} = &-k_f\frac{\theta_{i,j}-\theta_{i,j-1}}{\delta z_f} + \frac{\delta z_f}{2}\left[\begin{array}{c}\frac{k_f}{\delta x^2}\left(\theta_{i-1,j}+\theta_{i+1,j}-2\theta_{i,j}\right)-\frac{\rho_f c_f u_f}{\delta x}\left(\theta_{i,j}-\theta_{i-1,j}\right)\\ +\varepsilon\left(u_f\frac{dp}{dx}+\frac{dp}{dt}\right)\theta_{i,j}-2\frac{\rho_f c_f}{\delta t}\theta_{i,j}+\tau\dot{\gamma}+rhs_f\end{array}\right]\\ &+k_s\frac{\theta_{i,j+1}-\theta_{i,j}}{\delta z_s} + \frac{\delta z_s}{2}\left[\begin{array}{c}\frac{k_s}{\delta x^2}\left(\theta_{i-1,j}+\theta_{i+1,j}-2\theta_{i,j}\right)-\frac{\rho_s c_s u_1}{\delta x}\left(\theta_{i,j}-\theta_{i-1,j}\right)\\ -2\frac{\rho_s c_s}{\delta t}\theta_{i,j}+rhs_s\end{array}\right]\end{aligned}$$

In a similar manner the equations can be linked at the lower surface to give:

$$0 = k_f \frac{\theta_{i,j+1} - \theta_{i,j}}{\delta z_f} + \frac{\delta z_f}{2}\left[\begin{array}{c} \frac{k_f}{\delta x^2}\left(\theta_{i-1,j} + \theta_{i+1,j} - 2\theta_{i,j}\right) - \frac{\rho_f c_f u_f}{\delta x}\left(\theta_{i,j} - \theta_{i-1,j}\right) \\ + \varepsilon\left(u_f \frac{dp}{dx} + \frac{dp}{dt}\right)\theta_{i,j} - 2\frac{\hat{\rho}_f c_f}{\delta t}\theta_{i,j} + \tau\dot{\gamma} + rhs_f \end{array}\right]$$

$$- k_s \frac{\theta_{i,j} - \theta_{i,j-1}}{\delta z_s} + \frac{\delta z_s}{2}\left[\begin{array}{c} \frac{k_s}{\delta x^2}\left(\theta_{i-1,j} + \theta_{i+1,j} - 2\theta_{i,j}\right) - \frac{\rho_s c_s u_2}{\delta x}\left(\theta_{i,j} - \theta_{i-1,j}\right) \\ - 2\frac{\rho_s c_s}{\delta t}\theta_{i,j} + rhs_s \end{array}\right]$$

If slip occurs at the lower surface, with slip velocity u_{slip2}, then the two energy equations can be linked using:

$$-\tau_l u_{slip2} = k_f \frac{\theta_{i,j+1} - \theta_{i,j}}{\delta z_f} + \frac{\delta z_f}{2}\left[\begin{array}{c} \frac{k_f}{\delta x^2}\left(\theta_{i-1,j} + \theta_{i+1,j} - 2\theta_{i,j}\right) - \frac{\rho_f c_f u_f}{\delta x}\left(\theta_{i,j} - \theta_{i-1,j}\right) \\ + \varepsilon\left(u_f \frac{dp}{dx} + \frac{dp}{dt}\right)\theta_{i,j} - 2\frac{\hat{\rho}_f c_f}{\delta t}\theta_{i,j} + \tau\dot{\gamma} + rhs_f \end{array}\right]$$

$$- k_s \frac{\theta_{i,j} - \theta_{i,j-1}}{\delta z_s} + \frac{\delta z_s}{2}\left[\begin{array}{c} \frac{k_s}{\delta x^2}\left(\theta_{i-1,j} + \theta_{i+1,j} - 2\theta_{i,j}\right) - \frac{\rho_s c_s u_2}{\delta x}\left(\theta_{i,j} - \theta_{i-1,j}\right) \\ - 2\frac{\rho_s c_s}{\delta t}\theta_{i,j} + rhs_s \end{array}\right]$$

As mentioned in section 2.7, in the event that contact occurs the equations for the solids are linked by equation (2.20). Using this equation and the above approach the two solids can be linked by:

$$-p\mu|u_1 - u_2| = k_{s1} \frac{\theta_{i,j+1} - \theta_{i,j}}{\delta z_{s1}} + \frac{\delta z_{s1}}{2}\left[\begin{array}{c} \frac{k_{s1}}{\delta x^2}\left(\theta_{i-1,j} + \theta_{i+1,j} - 2\theta_{i,j}\right) - \frac{\rho_{s1} c_{s1} u_1}{\delta x}\left(\theta_{i,j} - \theta_{i-1,j}\right) \\ - 2\frac{\rho_{s1} c_{s1}}{\delta t}\theta_{i,j} + rhs_{s1} \end{array}\right]$$

$$- k_{s2} \frac{\theta_{i,j} - \theta_{i,j-1}}{\delta z_{s2}} + \frac{\delta z_{s2}}{2}\left[\begin{array}{c} \frac{k_{s2}}{\delta x^2}\left(\theta_{i-1,j} + \theta_{i+1,j} - 2\theta_{i,j}\right) - \frac{\rho_{s2} c_{s2} u_2}{\delta x}\left(\theta_{i,j} - \theta_{i-1,j}\right) \\ - 2\frac{\rho_{s2} c_{s2}}{\delta t}\theta_{i,j} + rhs_{s2} \end{array}\right]$$

The equations that describe the temperature distribution are applied in a series of columns, the terms in the x direction are not active, hence, only the terms in the z direction are active and form the tri-diagonal matrix. This is solved and the temperatures

updated before moving to the next column in the contact region. Figure 4.24 illustrates the solution method for the temperatures within the region.

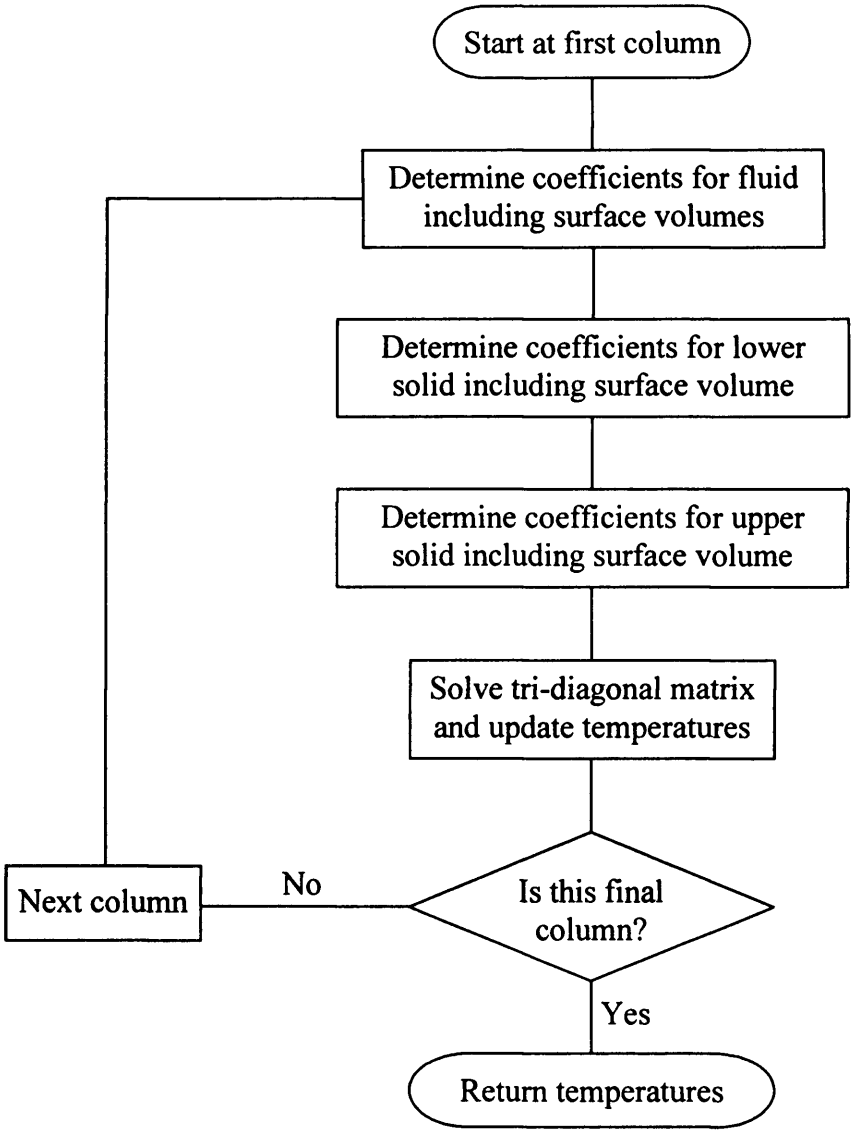

Figure 4.24 – Flowchart illustrating solution method for temperatures

4.5 Complete Coupled EHL Solution Method

The two governing equations, the hydrodynamic and elastic film thickness equations, are coupled creating a set of simultaneous equations. These simultaneous equations are solved to determine the pressure and film thickness through the contact.

From section 2.4 it was seen that the elastic film thickness equation, equation (2.17), can be presented in the form:

$$\frac{d^2h(x_i)}{dx^2} = \sum_{all\ k} f_{k-i}p_k + \frac{1}{R} + \frac{d^2\phi(x)}{dx^2}$$

Applying a central difference scheme for the second order differentials:

$$\frac{h_{i+1} + h_{i-1} - 2h_i}{\delta x^2} = \sum_{all\ k} f_{k-i}p_k + \frac{1}{R} + \frac{\phi_{i+1} + \phi_{i-1} - 2\phi_i}{\delta x^2}$$

Assuming that 'local' pressures have a reasonable effect on the deflection generated at the point of interest and are considered active in the solution:

$$\frac{h_{i+1} + h_{i-1} - 2h_i}{\delta x^2} - \sum_{|k-i|\le n_c} f_{k-i}p_k = \sum_{|k-i|>n_c} f_{k-i}p_k + \frac{1}{R} + \frac{\phi_{i+1} + \phi_{i-1} - 2\phi_i}{\delta x^2} \qquad (4.3)$$

Where n_c defines the number of mesh points that are considered 'local'

If the geometry chosen was not parabolic, as in the case of an involute, then equation (4.3) could be written as:

$$\frac{h_{i+1} + h_{i-1} - 2h_i}{\delta x^2} - \sum_{|k-i|\le n_c} f_{k-i}p_k = \sum_{|k-i|>n_c} f_{k-i}p_k + \frac{h_{ui+1} + h_{ui-1} - 2h_{ui}}{\delta x^2} + \frac{\phi_{i+1} + \phi_{i-1} - 2\phi_i}{\delta x^2}$$

Applying the Galerkin weighted residual approach to the hydrodynamic equation, equation (3.25), gives:

$$\int N_i \frac{d(\rho h)}{dt}dx + \int N_i \frac{d(\rho \hat{u} h)}{dx}dx - \int N_i \frac{d}{dx}\left(\sigma \frac{dp}{dx}\right)dx = 0 \qquad (4.4)$$

Note that the limits of integration have been omitted for simplicity and N_i is the shape function.

Applying the 'weak' formulation to equation (4.4) and manipulating the Couette flow component:

$$\int N_i \frac{d(\rho h)}{dt}dx + \int N_i\left(\rho h \frac{d\hat{u}}{dx} + \hat{u}h\frac{d\rho}{dx} + \rho\hat{u}\frac{dh}{dx}\right)dx + \int \frac{dN_i}{dx}\sigma\frac{dp}{dx}dx = 0$$

Applying the Crank-Nicholson scheme for the time dependent term:

$$\int N_i \rho h dx\Big|^{t+\delta t} - \int N_i \rho h dx\Big|^{t} + \frac{\delta t}{2}\left[\int N_i\left(\rho h\frac{d\hat{u}}{dx} + \hat{u}h\frac{d\rho}{dx} + \rho\hat{u}\frac{dh}{dx}\right)dx + \int\frac{dN_i}{dx}\sigma\frac{dp}{dx}dx\right]^{t+\delta t}$$
$$+\frac{\delta t}{2}\left[\int N_i\left(\rho h\frac{d\hat{u}}{dx} + \hat{u}h\frac{d\rho}{dx} + \rho\hat{u}\frac{dh}{dx}\right)dx + \int\frac{dN_i}{dx}\sigma\frac{dp}{dx}dx\right]^{t} = 0$$

Using standard interpolation techniques and re-arranging this can be expressed as:

$$\left\{\int N_i\tilde{\rho}\left(\sum_{j=1}^{n_e}N_j h_j\right)dx + \frac{\delta t}{2}\left(\int N_i\left[\left(\tilde{\rho}\frac{d\tilde{\hat{u}}}{dx} + \tilde{\hat{u}}\frac{d\tilde{\rho}}{dx}\right)\sum_{j=1}^{n_e}N_j h_j + \tilde{\rho}\tilde{\hat{u}}\sum_{j=1}^{n_e}\frac{dN_j}{dx}h_j\right]dx + \int\frac{dN_i}{dx}\tilde{\sigma}\sum_{j=1}^{n_g}\frac{dN_j}{dx}p_j dx\right)\right\}^{t+\delta t} \quad (4.5)$$
$$= \left\{\int N_i\tilde{\rho}\tilde{h}dx - \frac{\delta t}{2}\left[\int N_i\left(\tilde{\rho}\tilde{h}\frac{d\tilde{\hat{u}}}{dx} + \tilde{\hat{u}}\tilde{h}\frac{d\tilde{\rho}}{dx} + \tilde{\rho}\tilde{\hat{u}}\frac{d\tilde{h}}{dx}\right)dx + \int\frac{dN_i}{dx}\tilde{\sigma}\frac{d\tilde{p}}{dx}dx\right]\right\}^{t}$$

Where n_e is the number of points in the element.

The average terms are evaluated at the gauss points based on the local nodal values from:

$$\tilde{h} = \sum_{k=1}^{n_g}N_k h_k \qquad \tilde{\hat{u}} = \sum_{k=1}^{n_g}N_k\hat{u}_k \qquad \tilde{\rho} = \sum_{k=1}^{n_g}N_k\rho_k \qquad \tilde{\sigma} = \sum_{k=1}^{n_g}N_k\sigma_k$$

$$\frac{d\tilde{h}}{dx} = \sum_{k=1}^{n_g}N_k\frac{dh}{dx}_k \qquad \frac{d\tilde{p}}{dx} = \sum_{k=1}^{n_g}N_k\frac{dp}{dx}_k \qquad \frac{d\tilde{\hat{u}}}{dx} = \sum_{k=1}^{n_g}N_k\frac{d\hat{u}}{dx}_k \qquad \frac{d\tilde{\rho}}{dx} = \sum_{k=1}^{n_g}N_k\frac{d\rho}{dx}_k$$

Where n_g is the number of gauss points used.

The approaches used to determine equations (4.3) and (4.5) are covered in more detail in Elcoate et al (2001), however, a slight variation was implemented here due to the effect of slip leading to a modified mean entrainment velocity.

It can be shown that equations (4.3) and (4.5) can be expressed as:

$$\sum_{k=0}^{n_c}A_k p_k + \sum_{k=0}^{n_c}B_k h_k = R_i \quad (4.6)$$

$$\sum_{k=0}^{n_c}C_k p_k + \sum_{k=0}^{n_c}D_k h_k = E_i \quad (4.7)$$

A_k and B_k are the pressure and film thickness coefficients for the hydrodynamic Reynolds equation and R_i encompasses the values from the previous time step. C_k and D_k are the coefficients for the pressure and film thickness determined from the differential deflection approach, E_i includes the influence of the other pressures in the mesh. n_c is the number of nodes involved in the formulation which is determined from the 'local' points in the elastic film thickness equation.

These equations are solved at each point in the mesh by either elimination or iterative techniques. Iterative techniques are adopted as they can be modified to deal with contact easily. For the iterative approach equations (4.6) and (4.7) can be manipulated and expressed as a pair of simultaneous equations applied at a mesh point i with pressure p_i and film thickness h_i:

$$A_i p_i + B_i h_i = R_i - \sum_{k=1}^{n_c} A_k p_k - \sum_{k=1}^{n_c} B_k h_k = \hat{R}_i \tag{4.8}$$

$$C_i p_i + D_i h_i = E_i - \sum_{k=1}^{n_c} C_k p_k - \sum_{k=1}^{n_c} D_k h_k = \hat{E}_i \tag{4.9}$$

The new nodal pressure, p_i^{new}, and film thickness, h_i^{new}, can be determined by the block iteration:

$$p_i^{new} = \frac{D_i \hat{R}_i - B_i \hat{E}_i}{A_i D_i - B_i C_i} \tag{4.10}$$

$$h_i^{new} = \frac{A_i \hat{E}_i - C_i \hat{R}_i}{A_i D_i - B_i C_i} \tag{4.11}$$

If the solution to equation (4.11) produces a zero or negative value then a different approach is adopted. Since there is no fluid film there is no hydrodynamic equation, instead only the elastic deflection equation is applied to determine the pressure. Equations (4.10) and (4.11) can be modified to:

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

The solution process is outlined in the flowchart shown in Figure 4.25. The problem is started from the Hertzian response, with a semi-elliptical pressure distribution and flat parallel film thickness in the Hertz contact area, the values of which are determined from the load, geometry, material and oil properties. These conditions are combined with the temperature to determine the density, viscosity, S and $\hat{u}$, which are then used to calculate the coefficients in equations (4.6) and (4.7).

These equations can be solved by elimination and back substitution or by iterative methods which were outlined earlier. If iterative methods are used, then equations (4.10) and (4.11) are solved, the values obtained are compared to the original values. If the resulting values are considered converged then the values are returned, if not another iteration is performed with the new values forming the starting point for the next set of calculations. This iterative loop is termed the inner loop (not shown in figure 4.25) and the values are not relaxed before being used in the re-calculation. More information on the iterative method is presented in Holmes (2002)

The initial outer loop values are relaxed towards the values calculated from the coupled equations. These new pressure and film thickness values are used to update the temperature. If the pressure and film thickness are converged in the outer loop then the next time step is started with the converged solution forming the starting point.

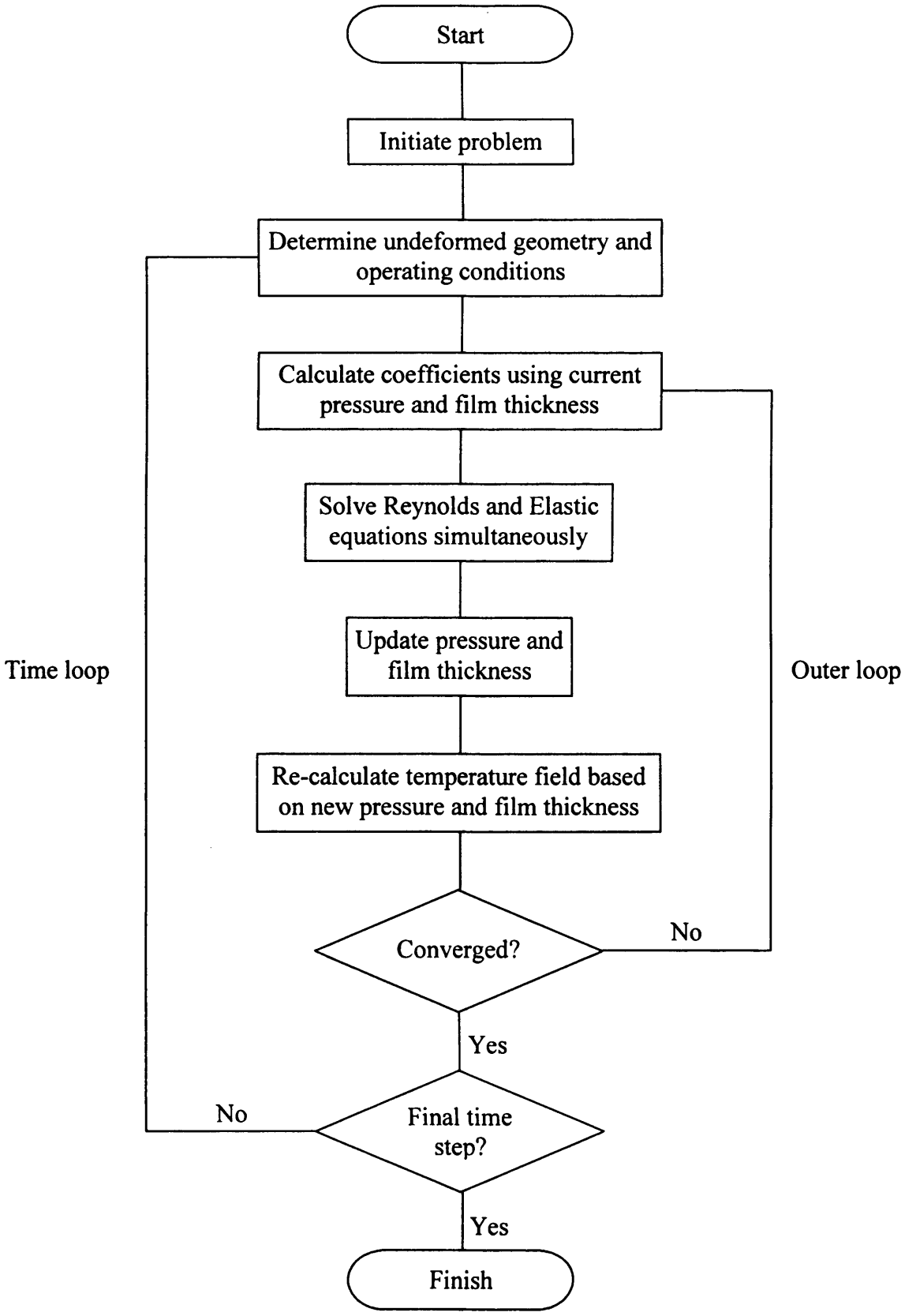

Figure 4.25 – Flowchart illustrating coupled solution method

Chapter 5

Steady State Isothermal Results

5.0 Introduction

This section takes the methods and equations introduced in the previous chapter for the limiting shear stress non-Newtonian models and uses them to examine smooth surface steady state problems. The results obtained are compared with those obtained using the Eyring non-Newtonian model which is simple to implement in a line contact model.

The responses of the different models are compared with varying operating conditions to determine which parameters have a major effect on the results. An approach is presented that allows the amount of influence that the limiting shear stress models have on the pressure and film thickness to be determined from the operating conditions.

For all the cases presented the operating conditions have been taken from Patching (1994) not withstanding the parameter changes introduced to give different responses. The operating conditions simulate test conditions that are run experimentally at Cardiff University. The experimental arrangement is almost identical to that covered later in Chapter 9. The tests involve two test disks, of the same dimensions, which are driven at different speeds and loaded against one another. For the test case considered, both disks are made of identical material. The material properties of the disks as well as the operating conditions and oil properties for the investigations are given in Table 5.1.

For all the Eyring results, the Eyring shear stress has been scaled to the limiting shear stress in order to achieve reasonably good comparisons. The scaling used is the same as that used in section 4.2, where it was determined that $\tau_l = 3\tau_0$ gives good agreement.

Table 5.1 – Operating conditions for test case

Operating Condition	Value
E of disks	206.9GPa
ν of disks	0.30
R of disks	38.1mm
$\bar{u}$	$25ms^{-1}$
ζ	1.25
w'	$526.6kNm^{-1}$
Oil	MobilJet II
Bulk oil temperature	100°C
η_0 at bulk oil temperature	0.0048Pas
α at bulk oil temperature	$11.1GPa^{-1}$

The changes in the operating parameters will change the film thickness, this change can be quantified using the Dowson and Higginson [Dowson, 1968] minimum film thickness equation:

$$h_{min} = 2.65\alpha^{0.54}(\eta_0\bar{u})^{0.7}R^{0.43}E'^{-0.03}w'^{-0.13} \tag{5.1}$$

The changes in the parameters will also change the maximum pressure according to [Hertz, 1881]:

$$p_0 = \sqrt{\frac{E'w'}{2\pi R}} \tag{5.2}$$

Using Newtonian behaviour, the shear stress variation through the film is given by:

$$\tau = \frac{\eta u_s}{h} + z\frac{h}{2}\frac{dp}{dx} \tag{5.3}$$

The changes in the parameters will affect the pressure distribution and the film thickness according to the equation (5.1) and equation (5.2). The pressure will influence the viscosity, increasing the pressure acts to increase the viscosity. Hence, the response of the Newtonian shear stress, equation (5.3), to the changes in the parameters can be approximated.

5.1 Influence of Slide Roll Ratio

From the equations presented in Chapter 3 it can be seen that the non-dimensional sliding imposed on the oil gives the amount of influence that the limiting shear stress model has. The equation that gives the non-dimensional sliding is re-written below for convenience:

$$U_s = \frac{\eta u_s}{\tau_l h} = \frac{\eta \bar{u} \zeta}{\tau_l h} \tag{5.4}$$

The slide roll ratio, ζ, does not directly influence the film thickness or the pressure generated. However, the amount of non-dimensional sliding imposed on the oil is proportional to ζ and the response that the rheological models have to changes in ζ can thus be assessed. Starting with pure rolling, $\zeta = 0.0$, the relative speeds of the two surfaces can be increased in subsequent simulations so that the slide roll ratio increases. This is conducted until a simple sliding condition is achieved, where $\zeta = 2.0$ and one of the surfaces is stationary.

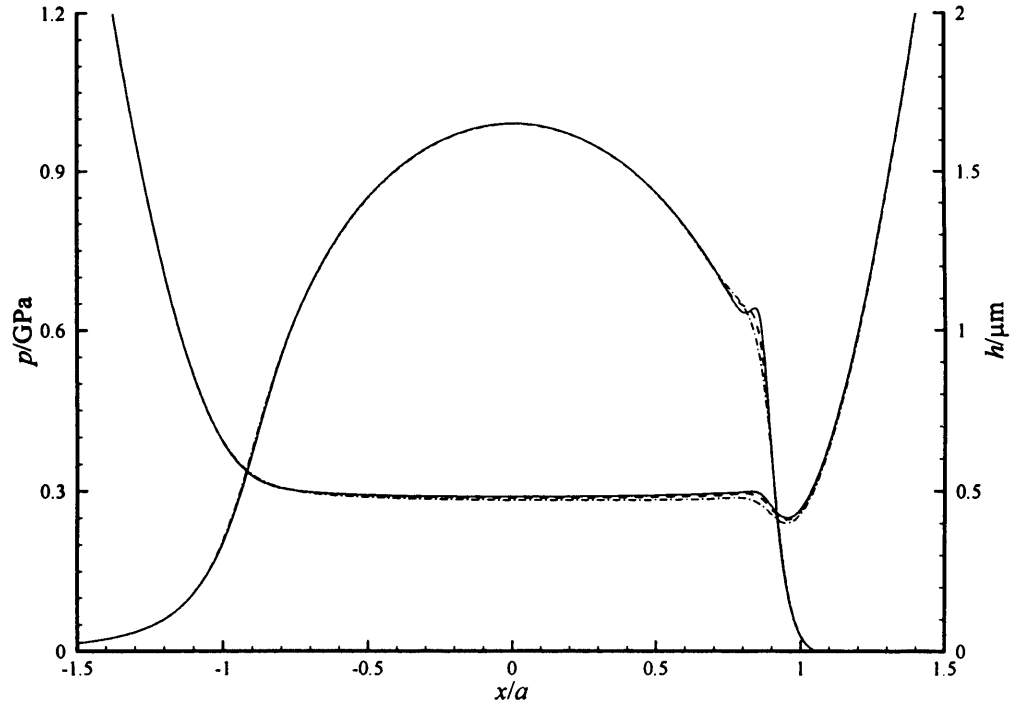

Figure 5.1 – Pressure and film thickness variation using the Eyring model for slide roll ratios of $\zeta = 0.00$ (solid), $\zeta = 1.00$ (dashed) and $\zeta = 2.00$ (dash-dotted)

Figure 5.1 shows the effect that increasing the slide roll ratio has on the pressure and film thickness produced assuming the Eyring model. It can be seen that small changes are produced in the film shape and pressure distribution as the amount of sliding is increased. The changes in the film thickness are noticeable in the central region. There is a downward slope formed from inlet to exit as the sliding velocity rises and the angle of the slope increases with slide roll ratio. The pressure does not change perceptibly through most of the contact, but there is a small difference seen in the spike feature which becomes less apparent as the slide roll ratio increases. The difference between the cases presented at the intermediate slide roll ratio results are small, thus the results are omitted from figure 5.1 for clarity.

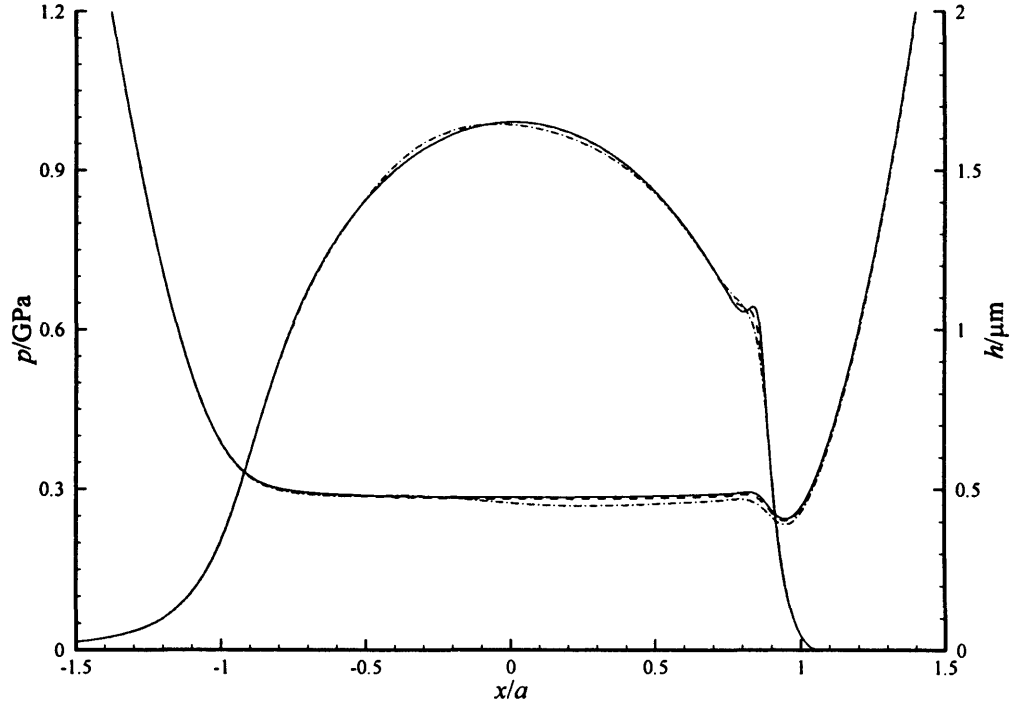

Figure 5.2 – Pressure and film thickness variation using the Bair and Winer model for slide roll ratios of $\zeta = 0.00$ (solid), $\zeta = 1.00$ (dashed) and $\zeta = 2.00$ (dash-dotted)

Figure 5.2 shows the corresponding results obtained using the Bair and Winer model. This shows a somewhat larger effect with different film shapes and pressure distributions developed at large slide roll ratios. The film thickness does not slope downward from inlet to exit, instead it develops an 's' shaped kink that is more prominent as the slide roll

ratio increases. The kink results from a slight thickening of the film in the inlet, at $x/a = -0.5$, and the localised decrease in the film thickness through most of the exit, $0 \le x/a \le 1$. The decrease in the film thickness in the exit region is far more noticeable than the increase created in the inlet region. The changes in the film thickness are accompanied by changes in the pressure, as the film thickness increases in the inlet the pressure rises and where the film thickness reduces in the exit the pressure falls. As with the Eyring results, the pressure spike at the exit constriction becomes less apparent as the sliding velocity increases. These features of the rheology produce pressures that are slightly different from the Eyring and the Hertzian pressure profiles, also the film thickness deviates from the parallel film thickness that is widely accepted.

The reasons for these responses can be examined in detail by considering the parameters that form part of the limiting shear stress rheology. Figure 5.3 shows the non-dimensional sliding, U_s, maximum non-dimensional sliding, $|U_s|_{\max}$, maximum non-dimensional shear stress, viscosity correction, S, $|\tau|_{\max}/\tau_l$, λ and Σ for the various slide roll ratios. The maximum non-dimensional sliding is closely linked to the pressure gradient, where an increase in the maximum non-dimensional sliding occurs for a decrease in the pressure gradient and vice versa. It should be noted that the maximum non-dimensional sliding is taken to be very large past the cavitation point, as the pressure gradient is zero and the maximum non-dimensional sliding would be infinite. The change in the pressure spike is also seen in the maximum non-dimensional sliding. The non-dimensional sliding is zero for the pure rolling case, due to the slide roll ratio being zero in equation (5.4). For the other cases it can be seen that the non-dimensional sliding increases towards the centre of the contact, this is due to the pressure generating high viscosities which cause the non-dimensional sliding to increase through equation (5.4). The difference between the non-dimensional sliding velocity for the $\zeta = 1.00$ and the $\zeta = 2.00$ cases is due to the increase in the slide roll ratio in equation (5.4).

The closer the non-dimensional sliding gets to its maximum value the more the pressure and film thickness move away from the pure rolling result. This indicates that the limiting shear stress rheology is influencing the pressure and film thickness developed and the influence increases as the non-dimensional sliding approaches the maximum value.

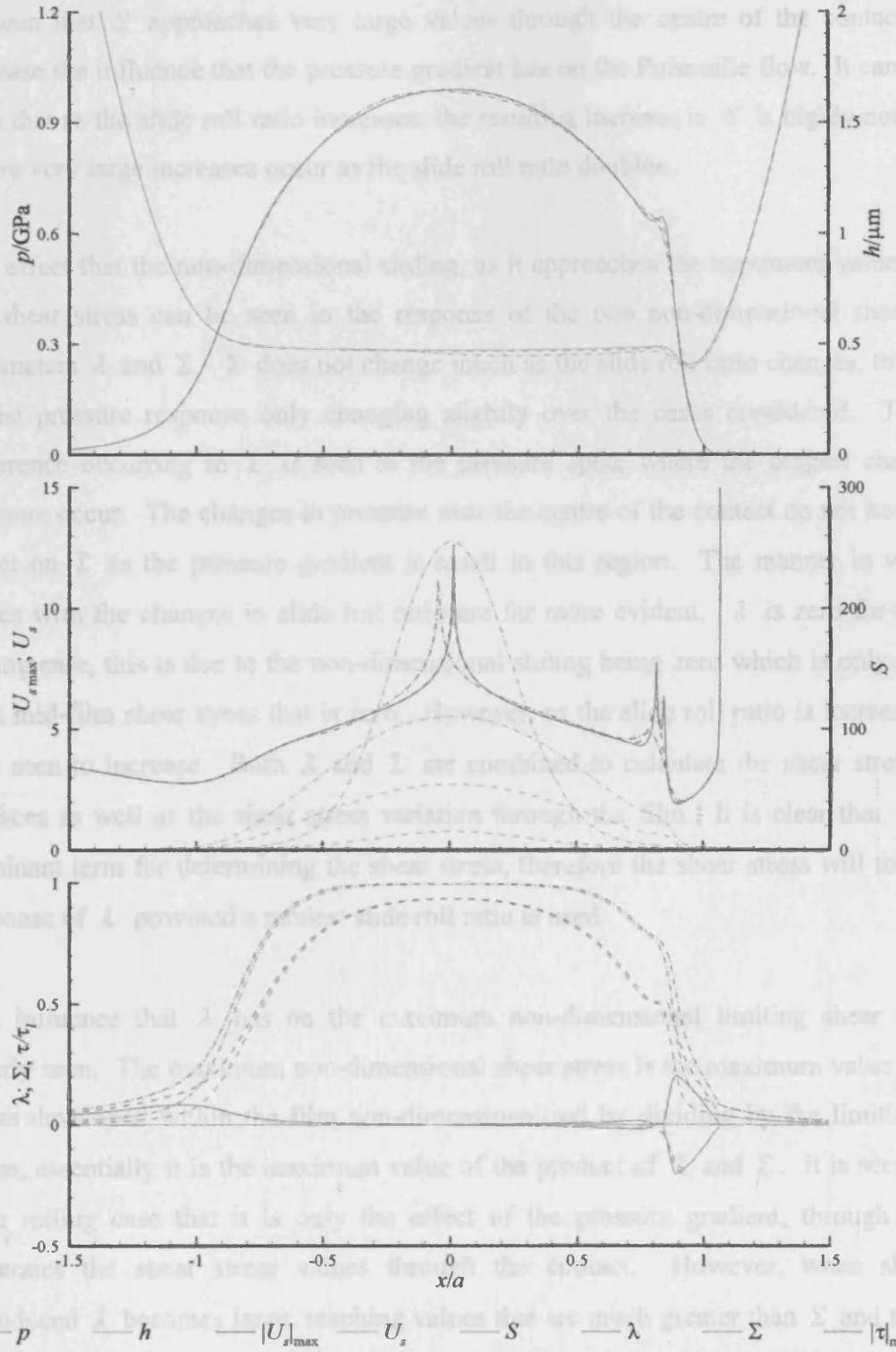

Figure 5.3 – Pressure, film thickness, maximum non-dimensional sliding, non-dimensional sliding, S, λ, Σ and maximum non-dimensional shear stress variation using the Bair and Winer model for slide roll ratios of $\zeta = 0.00$ (solid), $\zeta = 1.00$ (dashed) and $\zeta = 2.00$ (dash-dotted)

The change in the pressure and film thickness developed result from changes in S. It can be seen that S approaches very large values through the centre of the contact which increase the influence that the pressure gradient has on the Poiseuille flow. It can also be seen that as the slide roll ratio increases, the resulting increase in S is highly non-linear, where very large increases occur as the slide roll ratio doubles.

The effect that the non-dimensional sliding, as it approaches the maximum value, has on the shear stress can be seen in the response of the two non-dimensional shear stress parameters λ and Σ. Σ does not change much as the slide roll ratio changes, this is due to the pressure response only changing slightly over the cases considered. The only difference occurring in Σ is seen at the pressure spike where the biggest changes in pressure occur. The changes in pressure near the centre of the contact do not have much effect on Σ as the pressure gradient is small in this region. The manner in which λ varies with the changes in slide roll ratio are far more evident. λ is zero for the pure rolling case, this is due to the non-dimensional sliding being zero which is only satisfied by a mid-film shear stress that is zero. However, as the slide roll ratio is increased λ is also seen to increase. Both λ and Σ are combined to calculate the shear stress at the surfaces as well as the shear stress variation through the film. It is clear that λ is the dominant term for determining the shear stress, therefore the shear stress will follow the response of λ provided a modest slide roll ratio is used.

The influence that λ has on the maximum non-dimensional limiting shear stress is clearly seen. The maximum non-dimensional shear stress is the maximum value of shear stress developed within the film non-dimensionalised by dividing by the limiting shear stress, essentially it is the maximum value of the product of λ and Σ. It is seen for the pure rolling case that it is only the effect of the pressure gradient, through Σ, that generates the shear stress values through the contact. However, when sliding is introduced λ becomes large, reaching values that are much greater than Σ and thus hide its influence. For the $\zeta = 2.00$ case the shear stress is very close to the limiting value for the majority of the contact region, which gives rise to the changes in the film thickness and pressure.

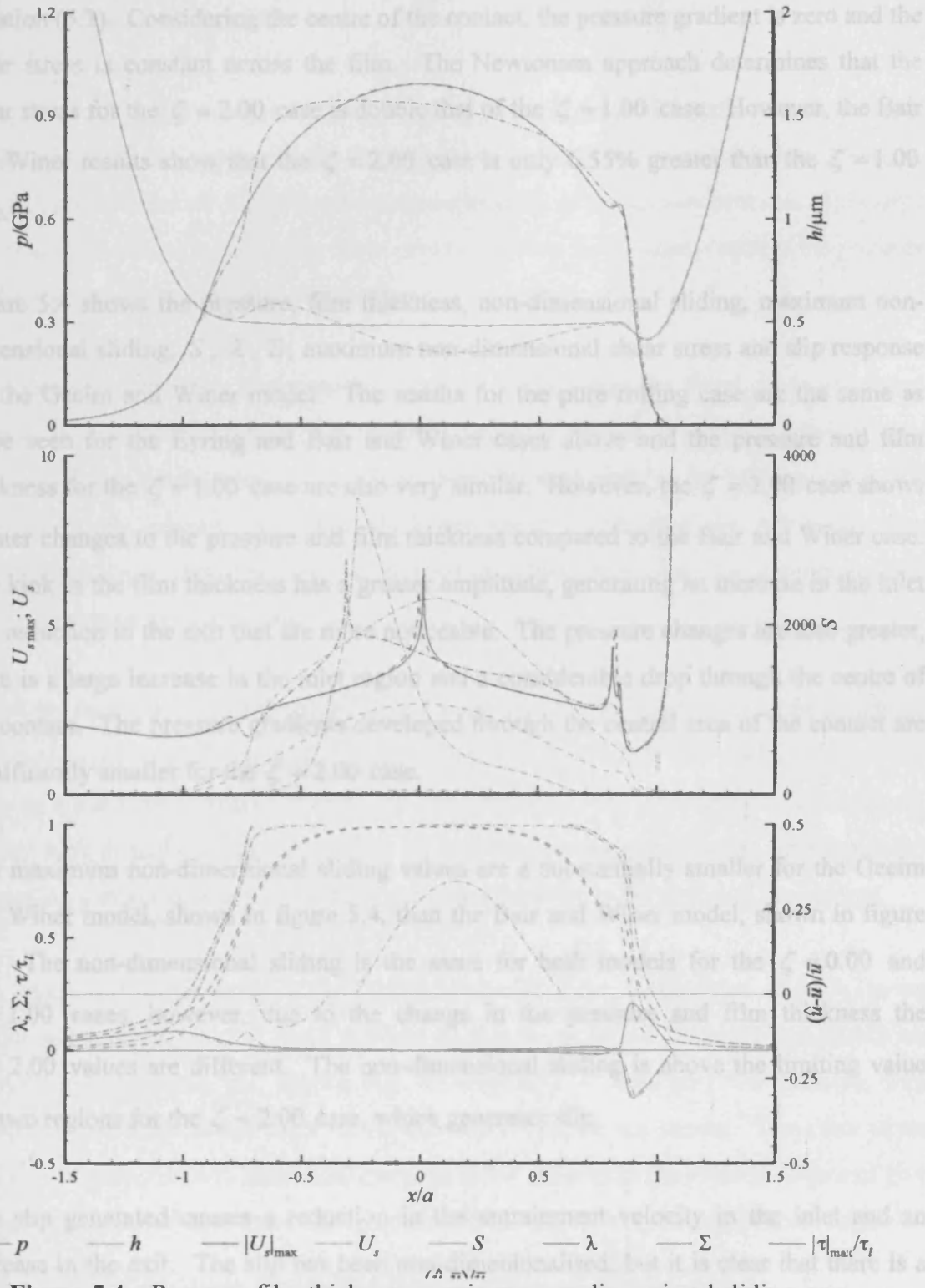

Figure 5.4 – Pressure, film thickness, maximum non-dimensional sliding, non-dimensional sliding, S, λ, Σ, maximum non-dimensional shear stress and slip variation using the Gecim and Winer model for slide roll ratios of $\zeta = 0.00$ (solid), $\zeta = 1.00$ (dashed) and $\zeta = 2.00$ (dash-dotted)

The shear stress response is unlike that predicted using the Newtonian shear stress from equation (5.3). Considering the centre of the contact, the pressure gradient is zero and the shear stress is constant across the film. The Newtonian approach determines that the shear stress for the $\zeta = 2.00$ case is double that of the $\zeta = 1.00$ case. However, the Bair and Winer results show that the $\zeta = 2.00$ case is only 6.55% greater than the $\zeta = 1.00$ case.

Figure 5.4 shows the pressure, film thickness, non-dimensional sliding, maximum non-dimensional sliding, S, λ, Σ, maximum non-dimensional shear stress and slip response for the Gecim and Winer model. The results for the pure rolling case are the same as those seen for the Eyring and Bair and Winer cases above and the pressure and film thickness for the $\zeta = 1.00$ case are also very similar. However, the $\zeta = 2.00$ case shows greater changes to the pressure and film thickness compared to the Bair and Winer case. The kink in the film thickness has a greater amplitude, generating an increase in the inlet and reduction in the exit that are more noticeable. The pressure changes are also greater, there is a large increase in the inlet region and a considerable drop through the centre of the contact. The pressure gradients developed through the central area of the contact are significantly smaller for the $\zeta = 2.00$ case.

The maximum non-dimensional sliding values are a substantially smaller for the Gecim and Winer model, shown in figure 5.4, than the Bair and Winer model, shown in figure 5.3. The non-dimensional sliding is the same for both models for the $\zeta = 0.00$ and $\zeta = 1.00$ cases, however, due to the change in the pressure and film thickness the $\zeta = 2.00$ values are different. The non-dimensional sliding is above the limiting value for two regions for the $\zeta = 2.00$ case, which generates slip.

The slip generated causes a reduction in the entrainment velocity in the inlet and an increase in the exit. The slip has been non-dimenionalised, but it is clear that there is a large amount of slip occurring, reaching a maximum velocity of $8.75\,\mathrm{ms}^{-1}$. There is a region of slip in the inlet occurring on the faster upper surface and to maintain the fluid flow the film thickness has to increase. There is another larger region of slip occurring on

the slower lower surface which increases the mean speed and to maintain the fluid flow the film thickness must reduce.

The values of S are substantially higher than those developed for the Bair and Winer model, which is the reason for the large changes in the pressure and the film thickness. The maximum value of S still occurs near the location of the highest pressure. This large increase in S allows substantially more pressure driven flow, which requires the pressure gradient to reduce to maintain continuity.

The changes in the film thickness and pressure occur when the shear stress developed within the contact approach the limiting value. As the non-dimensional sliding approaches the limiting value, the shear stress is very near to its limiting value and changes in the pressure and film thickness occur. Therefore the maximum non-dimensional shear stress response can be used to indicate how close the non-dimensional sliding is to its maximum value.

It can be seen that the more aggressive Gecim and Winer model has a much greater influence on the pressure and film thickness than the Bair and Winer model. Both of the limiting shear stress models generate the same features, with a tendency to create an 's' shaped kink in the film thickness. The models also change the pressure distribution by generating a larger pressure in the inlet and lower pressure in the centre of the contact.

The Gecim and Winer model is used in the remainder of the comparisons with the Eyring model as it emphasises the small differences that occur as the different parameters are changed. In addition, the figures can be simplified so that only the pressure, film thickness, non-dimensional shear stress and slip responses are shown. The shear stress and the slip response will show how close the shear stress is to the limiting value and give a good indication of how close the non-dimensional sliding is to the maximum value.

5.2 Operating Conditions

This section takes the work in section 5.1 further, the reponse that the different rheological models have to various operating conditions is determined. The operating conditions that are investigated are the load, w', and the mean entrainment velocity, $\overline{u}$.

5.2.1 Load Variation

The operating load given in table 5.1 is increased and decreased by 10% of the original value. Considering equation (5.1) it can be determined that increasing the load will generate small reductions in the film thickness. Larger loads will also generate larger pressures, indicated by equation (5.2). The increase in pressure will generate an increase in the viscosity and will increase the shear stress developed, seen in equation (5.3). Larger loads also lead to an increase in the Hertzian contact dimension. The response to the change in load of the Eyring model is shown in Figure 5.5.

The effect of the load on the film thickness is small, but it is seen to change the pressure distribution, where higher pressures are generated for the higher loads. The expected response of shear stress to the change in load can be seen, where higher loads increase the maximum value of the non-dimensional shear stress and lower loads reduce the maximum value. This is related to the way that the pressure influences the viscosity, higher pressure generates higher viscosity which increases the non-dimensional sliding. The shape of the pressure and the film thickness generated do not change as the load is varied. The increase in the shear stress for the +10% case is 6.12%.

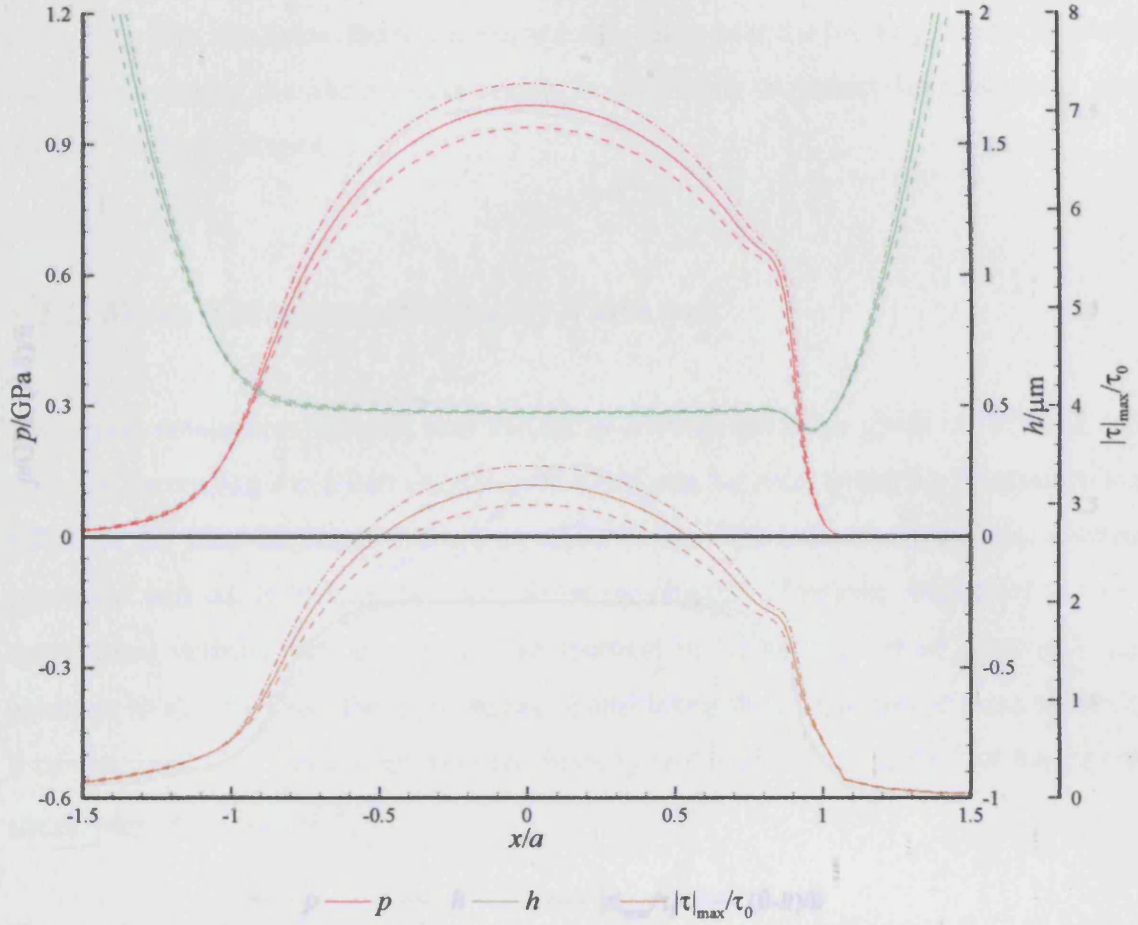

Figure 5.5 – Response of the Eyring model to different loads, original load (solid), load - 10% (dashed) and load +10% (dash-dotted)

The response of the Gecim and Winer model to the load variation is shown in Figure 5.6. The pressure is seen to change with load in a different manner to the Eyring model and it is immediately clear that the load affects the film thickness variation unlike the Eyring model. The film thickness kink is seen to increase in amplitude and moves further towards the inlet as the load increases. The change in film thickness is accompanied by corresponding changes in pressure. The pressure increases as necessitated by higher loads, but it is also seen to increase more in the inlet and develop a lower pressure gradient through the centre of the contact due to the limiting shear stress effect. The changes seen are a result of the rheology having more influence in the hydrodynamic equation.

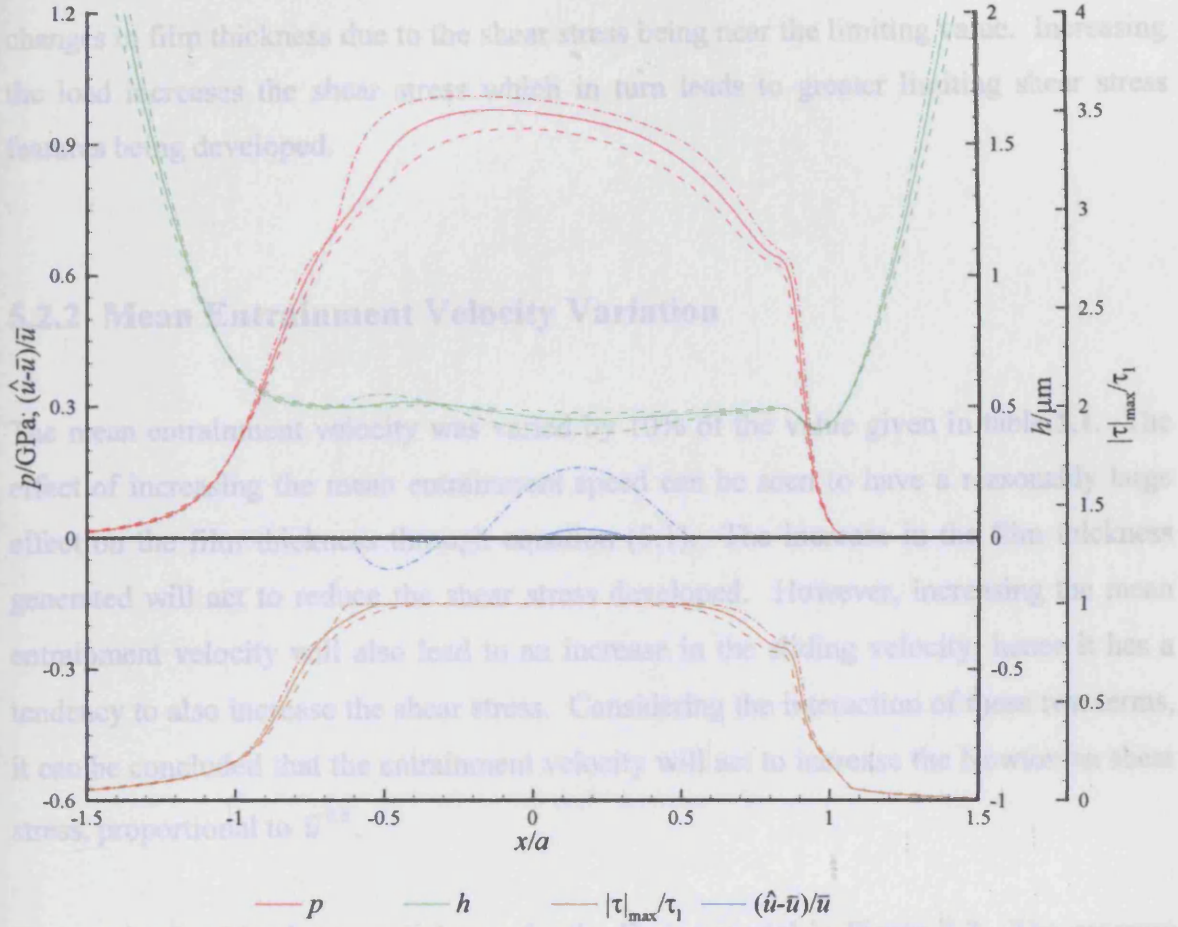

Figure 5.6 – Response of the Gecim and Winer model to different loads, original load (solid), load -10% (dashed) and load +10% (dash-dotted)

The shear stress increases with load and it is also seen that the region of the contact area that is at the limiting shear stress increases with load. The increased region that is at the limiting shear stress leads to a greater region that slips. In addition to the slip covering a greater area, it is also seen to increase in magnitude. Slip is only generated at the original load and the load +10% cases, where a large region of slip is seen for the load +10% case on the lower surface in the exit. The load -10% case does not develop slip and the shear stress only approaches the limiting value for a small region from $x/a = -0.3$ to $x/a = 0.3$.

Considering the Eyring and Gecim and Winer responses in figure 5.5 and 5.6 respectively, it can be seen that the load influences the maximum shear stress developed. For the Eyring model the shear stress increases with increasing load, but the influence that

the load has on the film thickness is very small, which is to be expected. In contrast, the Gecim and Winer model shear stress increases with load and is seen to produce large changes in film thickness due to the shear stress being near the limiting value. Increasing the load increases the shear stress which in turn leads to greater limiting shear stress features being developed.

5.2.2 Mean Entrainment Velocity Variation

The mean entrainment velocity was varied by 10% of the value given in table 5.1. The effect of increasing the mean entrainment speed can be seen to have a reasonably large effect on the film thickness through equation (5.1). The increase in the film thickness generated will act to reduce the shear stress developed. However, increasing the mean entrainment velocity will also lead to an increase in the sliding velocity, hence it has a tendency to also increase the shear stress. Considering the interaction of these two terms, it can be concluded that the entrainment velocity will act to increase the Newtonian shear stress, proportional to $\bar{u}^{0.3}$.

The results for this change are shown for the Eyring model in Figure 5.7. The pressure developed does not change, but the film thickness increases with larger entrainment velocities. The increase in the film thickness is to be expected as the entrainment mechanism is improved by larger velocities. The film shape developed for the different speeds are consistent, they all develop a slight downward slope from inlet to exit.

As predicted, the maximum shear stress developed is seen to increase slightly as the speed increases. The shape of the shear stress response also remains the same. The changes developed are much smaller than the changes seen for the load variation. The increase in the shear stress is 0.98%, which is roughly of the order predicted by $\bar{u}^{0.3}$.

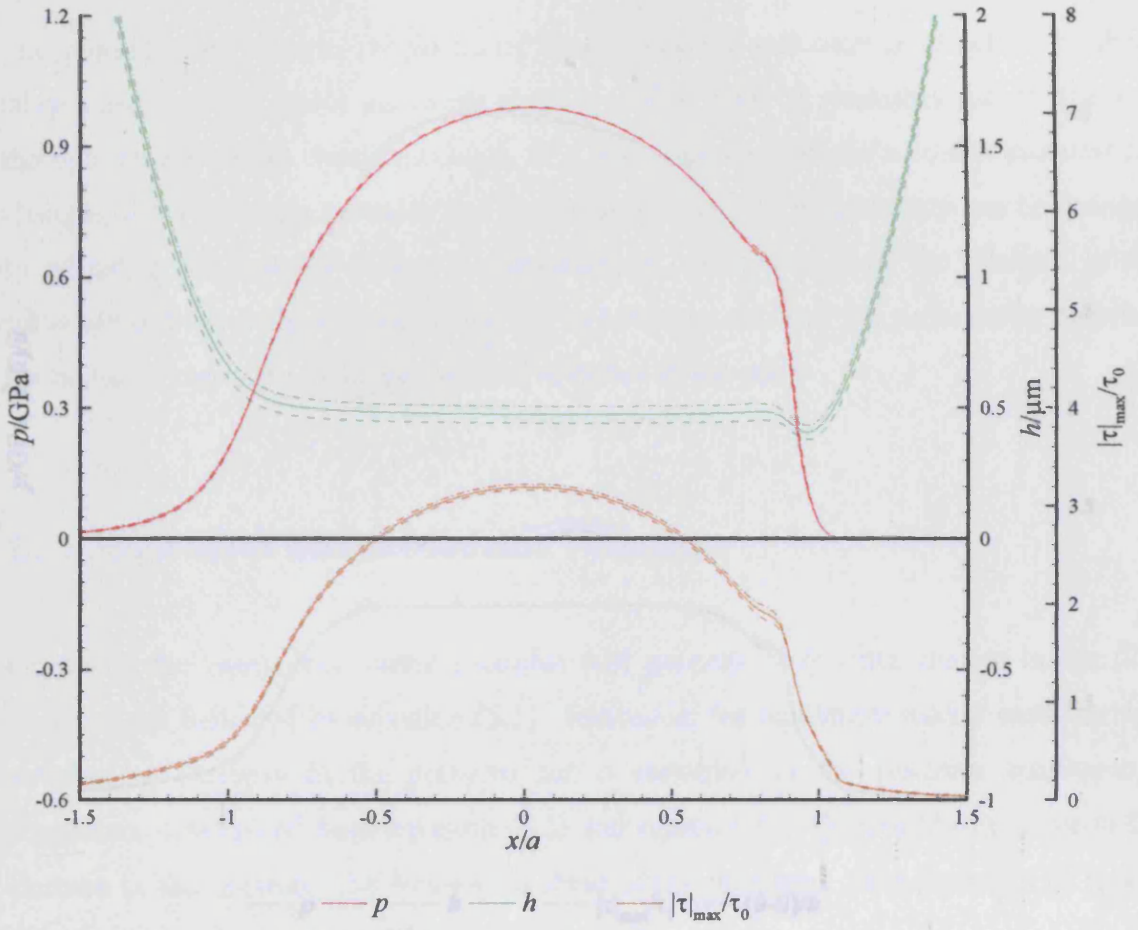

Figure 5.7 – Response of the Eyring model to different entrainment velocities, original (solid), -10% (dashed) and +10% (dash-dotted)

The Gecim and Winer response is shown in Figure 5.8, where the changes seen for the variation in entrainment speed are also small. The pressure profile shows minor changes in the centre of the contact and at $x/a = -0.5$. The film thickness increases with speed and retains the same features. There is a small increase in the amplitude of the kink in the film thickness, most noticeably for the +10% case.

The shear stress developed is the same for all of the cases through the centre of the contact, this is due to the shear stress reaching the limiting value through this region for each case. There are changes in the shear stress response at the edges of the contact, but like the Eyring results these changes are small. There is no region of slip produced for the -10% case, but it is seen that the slip generated increases for the +10% case. The increase in the slip is much smaller than that seen for the load variation in figure 5.6.

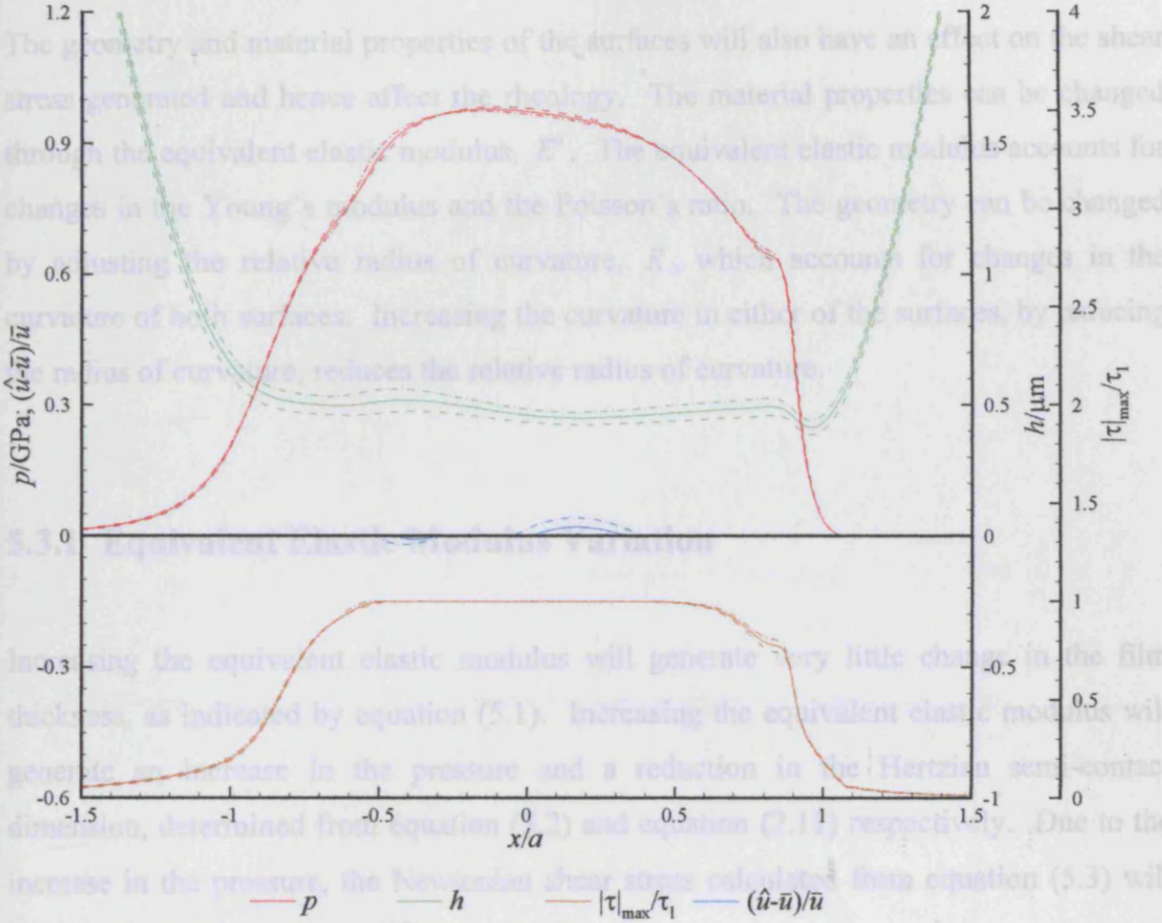

Figure 5.8 – Response of the Gecim and Winer model to different entrainment velocities, original (solid), -10% (dashed) and +10% (dash-dotted)

Increasing the entrainment velocity does increase the shear stress, but the influence is much lower than the effect that the load has. The effect that the mean entrainment velocity has on the shear stress is lower due to the influence it has on the film thickness, as covered earlier. Higher speeds develop thicker films, when considering equation (5.4) it can be seen that the influence of higher speeds on the numerator will be diminished by the thicker films on the denominator. This may be contrasted with the influence of changes in pressure on the viscosity; small changes in large loads increase the viscosity a large amount due to the exponential relationship. Large changes in viscosity change the non-dimensional sliding, through equation (5.4), by a large amount and thus produce much higher shear stress changes.

5.3 Geometry and Material Properties

The geometry and material properties of the surfaces will also have an effect on the shear stress generated and hence affect the rheology. The material properties can be changed through the equivalent elastic modulus, E'. The equivalent elastic modulus accounts for changes in the Young's modulus and the Poisson's ratio. The geometry can be changed by adjusting the relative radius of curvature, R, which accounts for changes in the curvature of both surfaces. Increasing the curvature in either of the surfaces, by reducing the radius of curvature, reduces the relative radius of curvature.

5.3.1 Equivalent Elastic Modulus Variation

Increasing the equivalent elastic modulus will generate very little change in the film thickness, as indicated by equation (5.1). Increasing the equivalent elastic modulus will generate an increase in the pressure and a reduction in the Hertzian semi-contact dimension, determined from equation (5.2) and equation (2.11) respectively. Due to the increase in the pressure, the Newtonian shear stress calculated from equation (5.3) will increase due to the pressure effect on the viscosity.

The response of the Eyring model to the variation in the equivalent elastic modulus is shown in Figure 5.9. Increasing the equivalent elastic modulus generates higher pressures whilst producing very little change in the film thickness, as expected. The pressures change by 5% of the original value when the elastic modulus is changed by 10%, which is of the order predicted by Hertzian contact theory through equation (5.2). The film thickness experiences very small changes as the equivalent elastic modulus is varied, where the larger equivalent elastic modulus develops a smaller film thickness.

The shear stress developed also increases with increasing equivalent elastic modulus. This response is related to the increased viscosity developed through the contact due to higher pressures. This was also noted for the effect that increasing load has on the shear stress. The changes in the shear stress are not as large as the changes for the load due to

the reduced sensitivity of the film thickness to the increase in the equivalent elastic modulus.

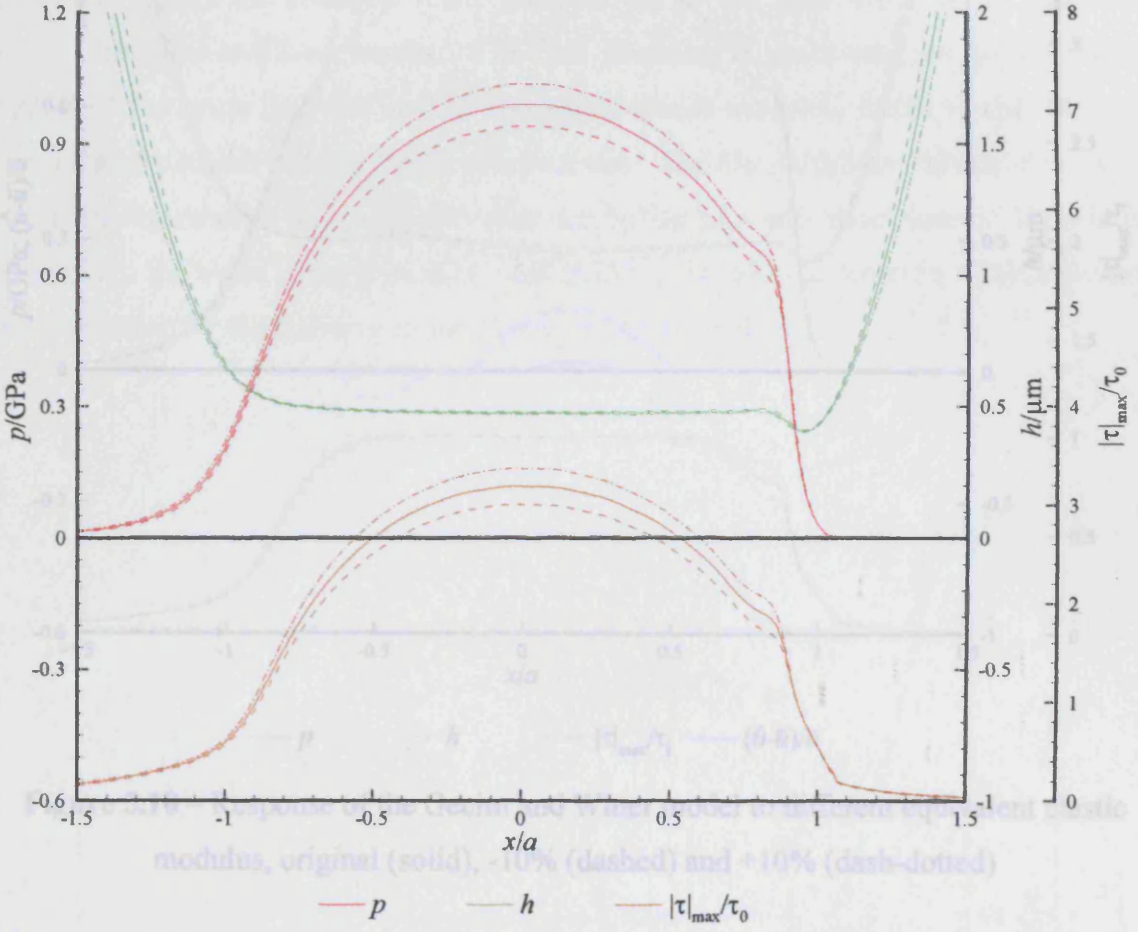

Figure 5.9 – Response of the Eyring model to different equivalent elastic modulus, original (solid), -10% (dashed) and +10% (dash-dotted)

The response of the Gecim and Winer model to the equivalent elastic modulus changes are shown in Figure 5.10. Again, the changes in equivalent elastic modulus develop similar changes to those seen when the load was varied. Increasing the equivalent elastic modulus develops higher pressures, but due to the limiting shear stress effect the pressure gradient reduces through the centre of the contact and the pressure increases in the inlet. Larger amplitude film thickness kinks are generated when the equivalent elastic modulus is increased.

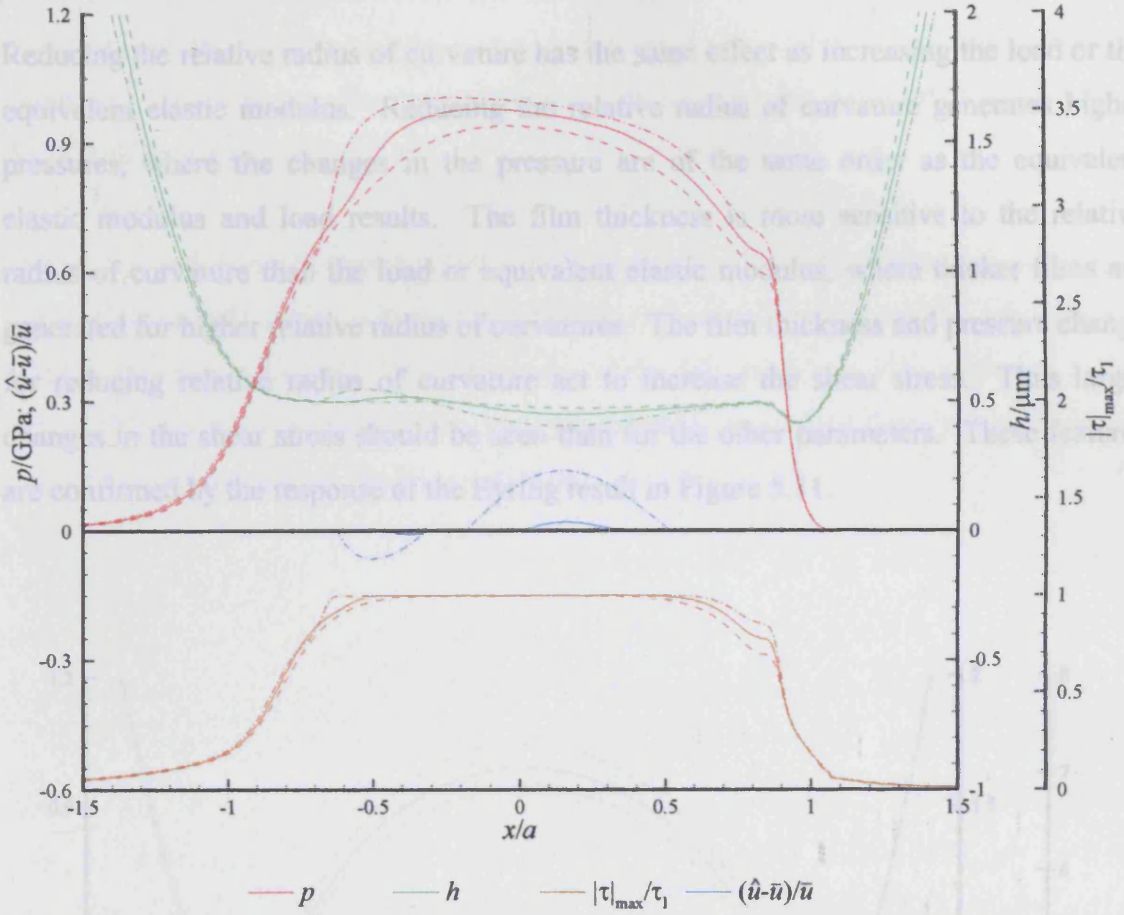

Figure 5.10 – Response of the Gecim and Winer model to different equivalent elastic modulus, original (solid), -10% (dashed) and +10% (dash-dotted)

The part of the contact region that is at the limiting shear stress enlarges as the equivalent elastic modulus increases. The increase in the shear stress also produces greater amounts of slip, where the regions of slip generated also cover more of the Hertzian contact area.

Generally, increasing the equivalent elastic modulus causes the pressure to rise, which is predicted using equation (5.2). Similarly to the load effect the increase in the pressure generates higher viscosities and the changes in equivalent elastic modulus do not have any significant influence on the film thickness. The change in shear stress for the Eyring model is 5.67% as the equivalent elastic modulus is increased by 10%. This is lower than the increase in the shear stress for the load of 6.12%.

5.3.2 Relative Radius of Curvature Variation

Reducing the relative radius of curvature has the same effect as increasing the load or the equivalent elastic modulus. Reducing the relative radius of curvature generates higher pressures, where the changes in the pressure are of the same order as the equivalent elastic modulus and load results. The film thickness is more sensitive to the relative radius of curvature than the load or equivalent elastic modulus, where thicker films are generated for higher relative radius of curvatures. The film thickness and pressure change for reducing relative radius of curvature act to increase the shear stress. Thus larger changes in the shear stress should be seen than for the other parameters. These features are confirmed by the response of the Eyring result in Figure 5.11.

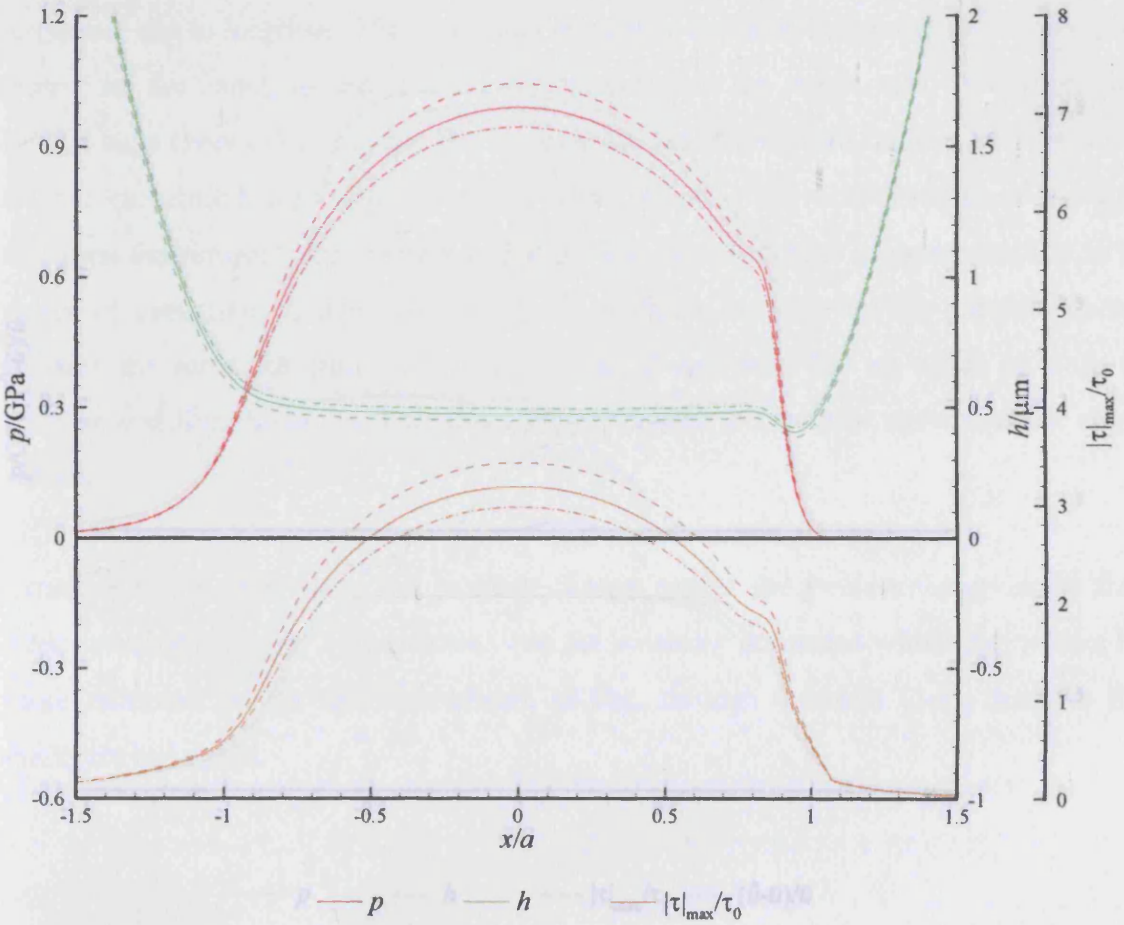

Figure 5.11 – Response of the Eyring model to different relative radius of curvature, original (solid), -10% (dashed) and +10% (dash-dotted)

The shear stress rises as the pressure increases due to the pressure influence on viscosity. Despite developing the same changes in pressure as the load and equivalent elastic modulus, the shear stress shows greater variation for the relative radius of curvature variation. This is due to the interaction of the thinner films in equation (5.4) increasing the non-dimensional sliding.

The Gecim and Winer response is shown in Figure 5.12. The same response is seen as for the Eyring model, where reducing the radius of curvature increases the pressure and reduces the film thickness. The amplitude of the film thickness kink increases as the radius of curvature is reduced. The kink produced is larger than those produced by the load and equivalent elastic modulus variations.

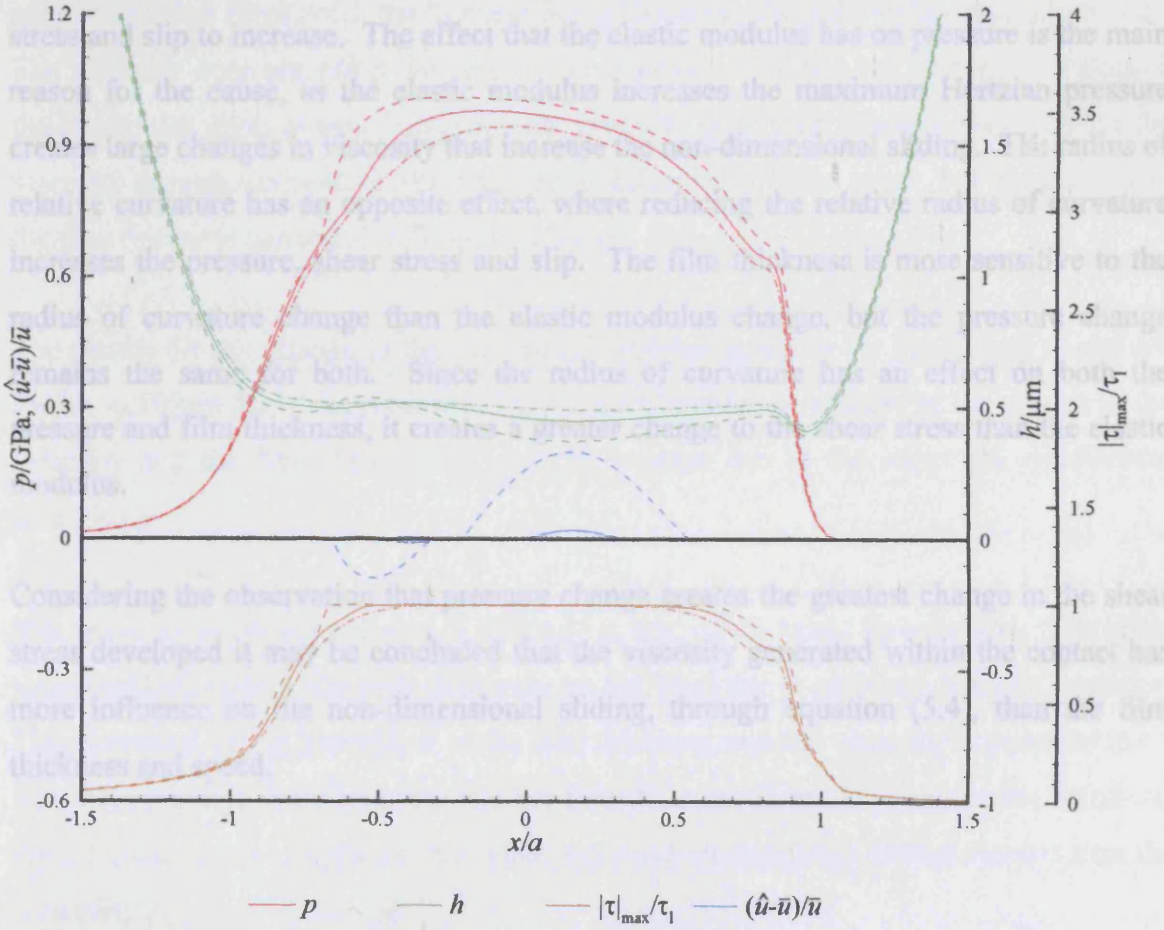

Figure 5.12 – Response of the Gecim and Winer model to different relative radius of curvature, original (solid), -10% (dashed) and +10% (dash-dotted)

The shear stress increases as the radius of curvature is reduced, creating a greater region that is at the limiting shear stress. The slip responds to the shear stress change, increasing as the radius of curvature reduces.

The relative radius of curvature creates the same change to the pressure as load and equivalent elastic modulus, seen by comparing figure 5.11 to figures 5.9 and 5.5. The response of the pressure follows the changes predicted by Hertzian equations. The radius of curvature does create a noticeable change in the film thickness, where greater radii of curvature generate thicker films. Combining the two results and considering equation (5.4) it can be seen that both trends will act to increase the non-dimensional sliding. This is confirmed when examining the change seen in the shear stress, for the Eyring model the maximum value of non-dimensional shear stress increases by 7.71%.

Generally, it can be seen that increasing the equivalent elastic modulus causes the shear stress and slip to increase. The effect that the elastic modulus has on pressure is the main reason for the cause, as the elastic modulus increases the maximum Hertzian pressure creates large changes in viscosity that increase the non-dimensional sliding. The radius of relative curvature has an opposite effect, where reducing the relative radius of curvature increases the pressure, shear stress and slip. The film thickness is more sensitive to the radius of curvature change than the elastic modulus change, but the pressure change remains the same for both. Since the radius of curvature has an effect on both the pressure and film thickness, it creates a greater change to the shear stress than the elastic modulus.

Considering the observation that pressure change creates the greatest change in the shear stress developed it may be concluded that the viscosity generated within the contact has more influence on the non-dimensional sliding, through equation (5.4), than the film thickness and speed.

5.4 Oil Properties

It has been seen that the viscosity of the lubricant is one of the important properties when calculating the non-dimensional sliding and hence in determining the nature of limiting shear stress behaviour developed. The oil properties can be varied which allow the viscosity and the limiting shear stress values to be considered. The viscosity response is determined from two properties, the viscosity at ambient pressure, η_0, and the viscosity pressure coefficient, α.

5.4.1 Viscosity at Ambient Pressure Variation

It can be seen in equation (5.1) that increasing the viscosity at ambient pressure will generate thicker films, with the increase being similar to the entrainment velocity results. The viscosity does not effect the pressure and contact area, but can be seen to influence the Newtonian shear stress. Increasing the viscosity at ambient pressure will increase the viscosity through the contact which will act to increase the shear stress, but the increase in the film thickness generated will act to reduce it.

The results for the change in the viscosity at ambient pressure for the Eyring model are shown in Figure 5.13. As predicted, the pressure remains insensitive to the change in the viscosity and the film thickness is seen to increase due to the improved entrainment conditions.

The response of the shear stress is different to that which might be expected. The shear stress reduces, unlike the film thickness, as the viscosity at ambient pressure increases. This is caused by the interaction of the film thickness and the viscosity in equation (5.4). The viscosity is in the numerator and the film thickness is in the denominator, therefore, for the shear stress to drop the film thickness must increase by a greater amount than the viscosity.

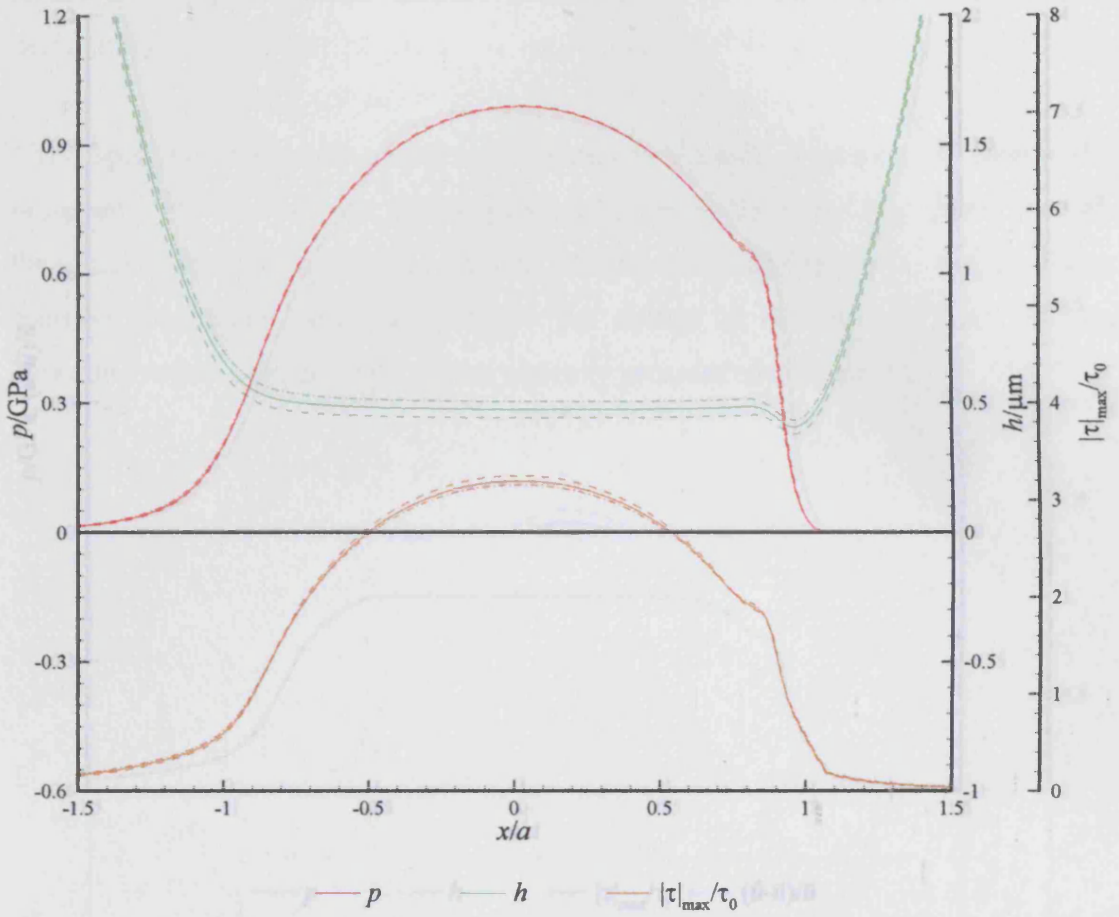

Figure 5.13 – Response of the Eyring model to different viscosities at ambient pressure, original (solid), -10% (dashed) and +10% (dash-dotted)

The Gecim and Winer response is shown in Figure 5.14, where similar trends to that seen for the Eyring model are also present. The pressure remains roughly the same, the only changes that occur in the pressure are a result of the rheology influence. The film thickness is seen to increase as the viscosity at ambient pressure increases. As the film thickness increases the kink that is formed reduces in amplitude, as a result of the shear stress dropping.

Despite the trend seen for the film thickness, the shear stress does not experience large changes. The shear stress that is at the limiting value does not appear to change as the viscosity varies, but the slip generated does reflect the trend seen in the shear stress for the Eyring model. The slip reduces as the viscosity at ambient pressure is increased, the changes are small, but they account for the changes seen in the film thickness.

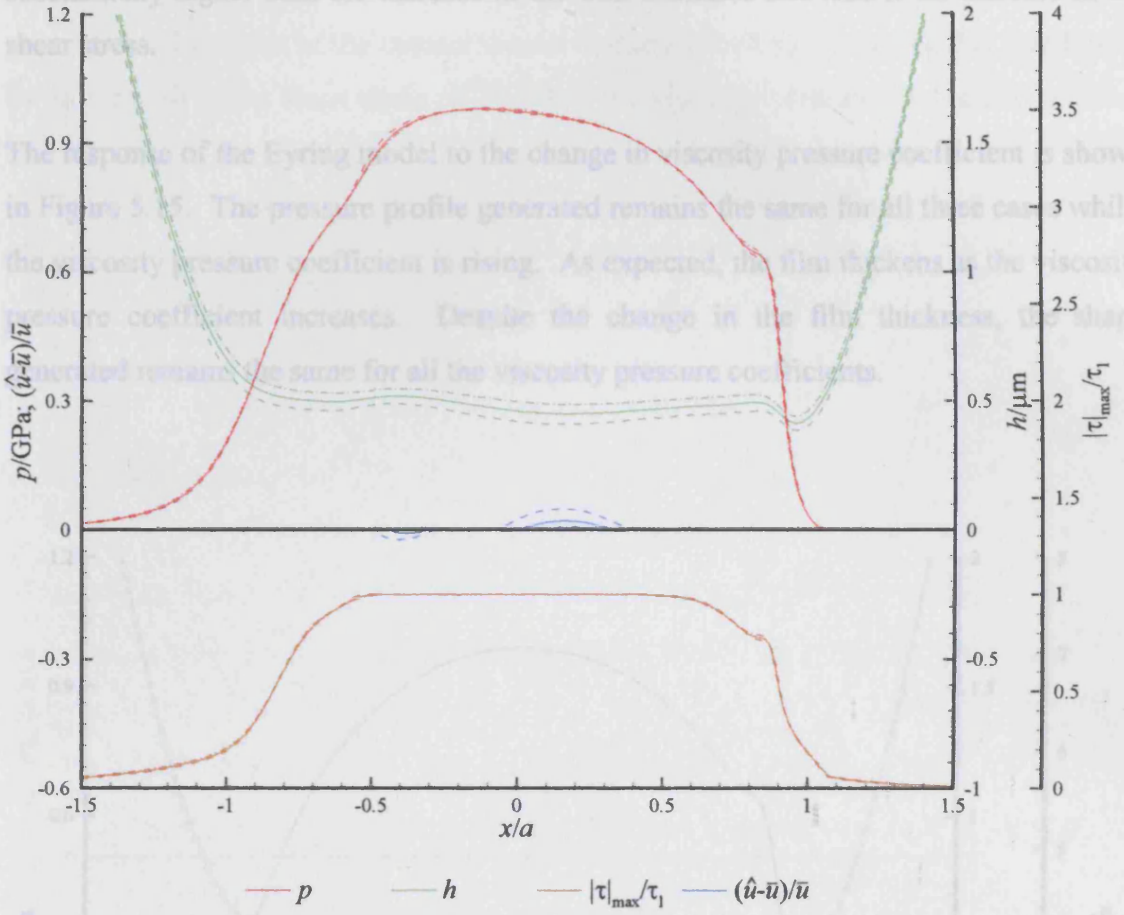

Figure 5.14 – Response of the Gecim and Winer model to different viscosities at ambient pressure, original (solid), -10% (dashed) and +10% (dash-dotted)

In contrast to what might be expected, increasing the ambient pressure viscosity does not increase the amount of limiting shear stress features seen. The increase actually causes the shear stress to drop and reduces the amplitude of the film thickness kink. The changes are small, where the drop in the Eyring shear stress is only 1.17%. The shear stresses for the Gecim and Winer model were very similar, differences were only seen in the slip response.

5.4.2 Viscosity Pressure Coefficient Variation

Increasing the viscosity pressure coefficient will generate substantially higher viscosities in the centre of the contact. The viscosity pressure coefficient will also generate thicker

films, similar to the viscosity at ambient pressure. The increase in the viscosity should be substantially higher than the increase in the film thickness and lead to an increase in the shear stress.

The response of the Eyring model to the change in viscosity pressure coefficient is shown in Figure 5.15. The pressure profile generated remains the same for all three cases whilst the viscosity pressure coefficient is rising. As expected, the film thickens as the viscosity pressure coefficient increases. Despite the change in the film thickness, the shape generated remains the same for all the viscosity pressure coefficients.

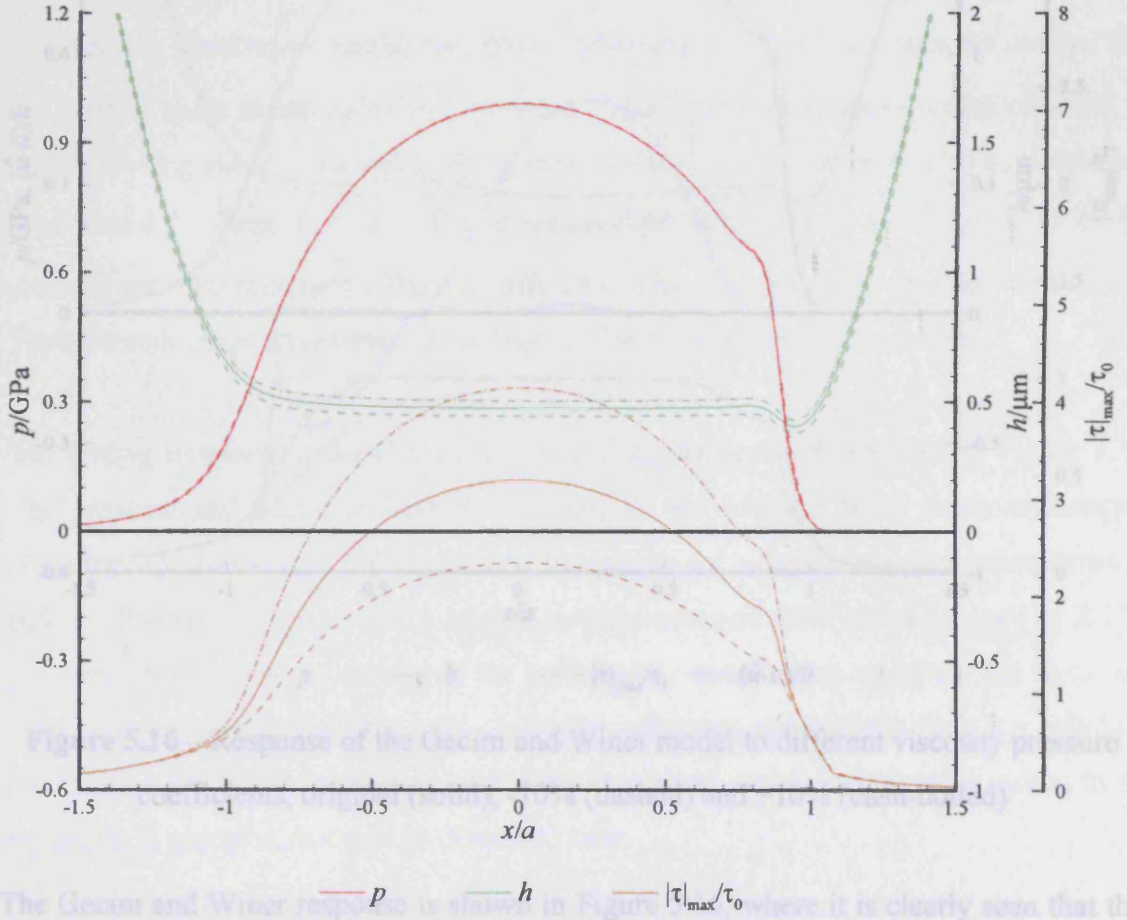

Figure 5.15 – Response of the Eyring model to different viscosity pressure coefficients, original (solid), -10% (dashed) and +10% (dash-dotted)

The shear stress changes by large amounts as the viscosity pressure coefficient is varied, despite the modest changes in film thickness and almost identical pressures developed.

The larger the viscosity pressure coefficient the higher the viscosity, which in turn develops higher non-dimensional sliding generating higher shear stresses. The viscosity generated at the centre of the contact is seen to increase by large amounts, this is indicated by an increase in the shear stress of 29.9% as the viscosity pressure coefficient increases by 10%.

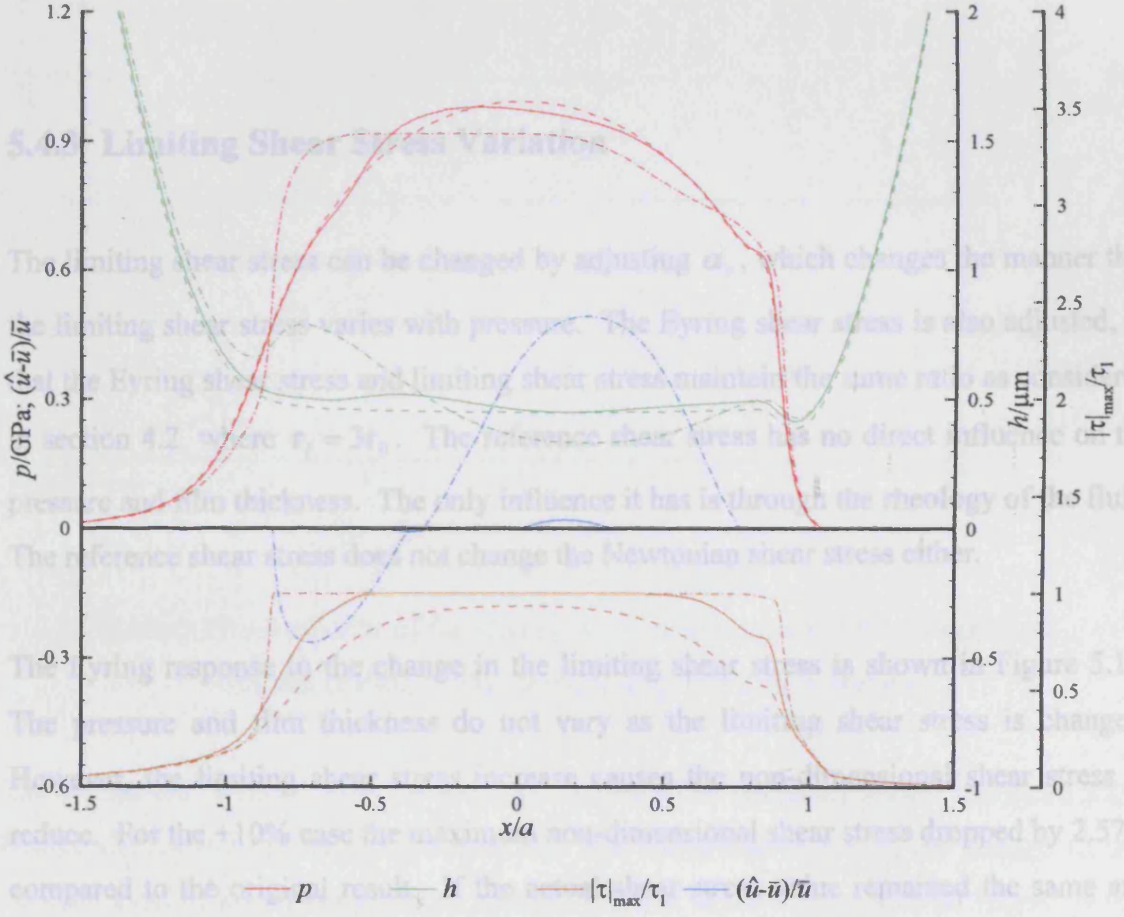

Figure 5.16 – Response of the Gecim and Winer model to different viscosity pressure coefficients, original (solid), -10% (dashed) and +10% (dash-dotted)

The Gecim and Winer response is shown in Figure 5.16, where it is clearly seen that the pressure and film thickness change drastically as the viscosity pressure coefficient is varied. The maximum pressure generated in the contact reduces as the viscosity pressure coefficient is increased, this is accompanied by a reduction in the pressure gradient through the centre of the contact. The film thickness kink increases in amplitude by a

substantial amount as the viscosity pressure coefficient increases. The pressure and film thickness developed for the -10% case are very similar to the Eyring results in figure 5.15.

The region over which the shear stress is at the limiting value increases by almost 80% as the viscosity increases. As the viscosity reduces, it is seen that the shear stress does not actually reach the limiting value at any point within the contact. The slip occurring through the contact region increases by large amounts as the viscosity increases, for the +10% case almost 80% of the Hertzian region experiences slip.

5.4.3 Limiting Shear Stress Variation

The limiting shear stress can be changed by adjusting α_τ, which changes the manner that the limiting shear stress varies with pressure. The Eyring shear stress is also adjusted, so that the Eyring shear stress and limiting shear stress maintain the same ratio as considered in section 4.2, where $\tau_l = 3\tau_0$. The reference shear stress has no direct influence on the pressure and film thickness. The only influence it has is through the rheology of the fluid. The reference shear stress does not change the Newtonian shear stress either.

The Eyring response to the change in the limiting shear stress is shown in Figure 5.17. The pressure and film thickness do not vary as the limiting shear stress is changed. However, the limiting shear stress increase causes the non-dimensional shear stress to reduce. For the +10% case the maximum non-dimensional shear stress dropped by 2.57% compared to the original result. If the actual shear stress value remained the same and then was non-dimensionalised using a higher reference shear stress then the reduction would be 9.0%. The interaction of the limiting shear stress, or reference stress, in the equations is complex, but will be discussed later.

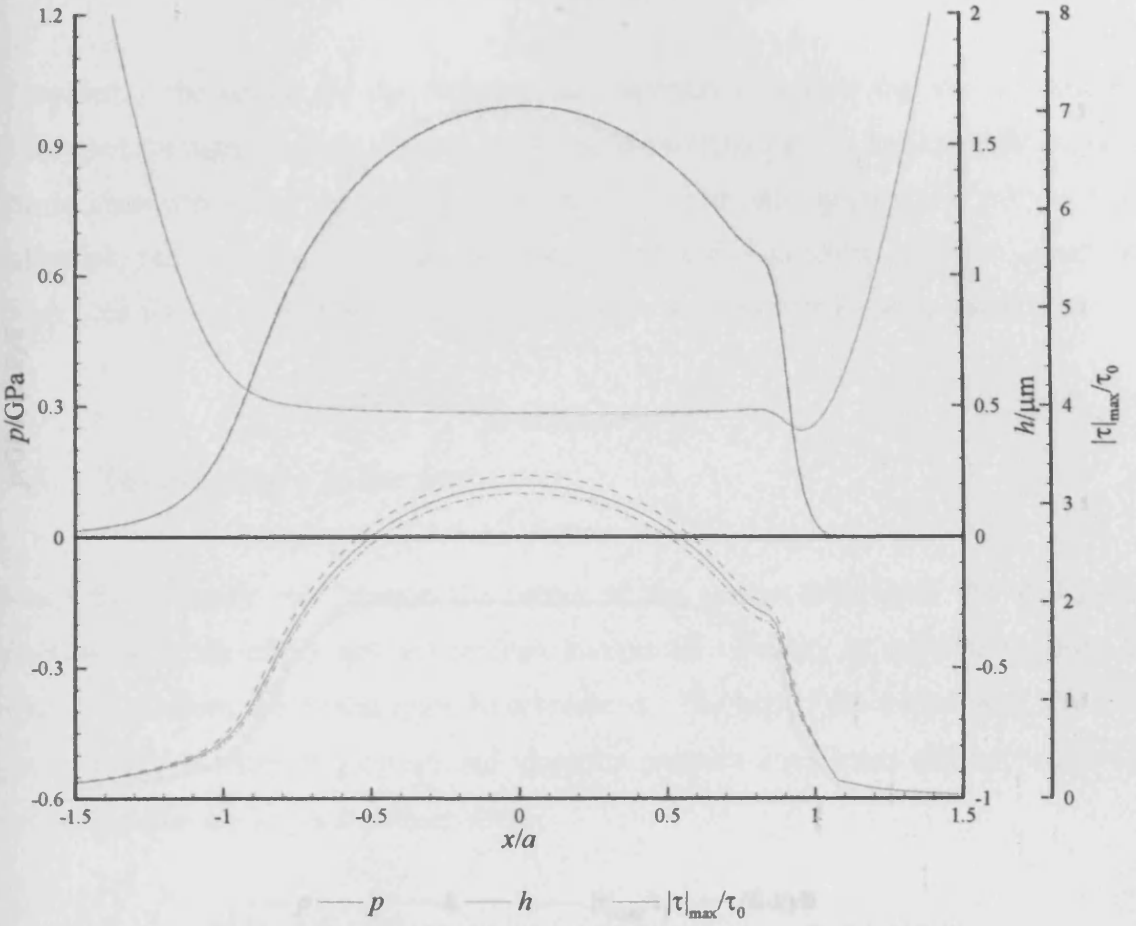

Figure 5.17 – Response of the Eyring model to different limiting shear stress, original (solid), -10% (dashed) and +10% (dash-dotted)

The Gecim and Winer results for varying limiting shear stress are shown in Figure 5.18. Unlike the Eyring results the pressure and film thickness vary with the change in the limiting shear stress. However, the changes are due to the shear stress developed being near the limiting value, thus the differences seen are purely related to the rheology. The limiting shear stress features are seen to diminish as the limiting shear stress is increased, which is expected.

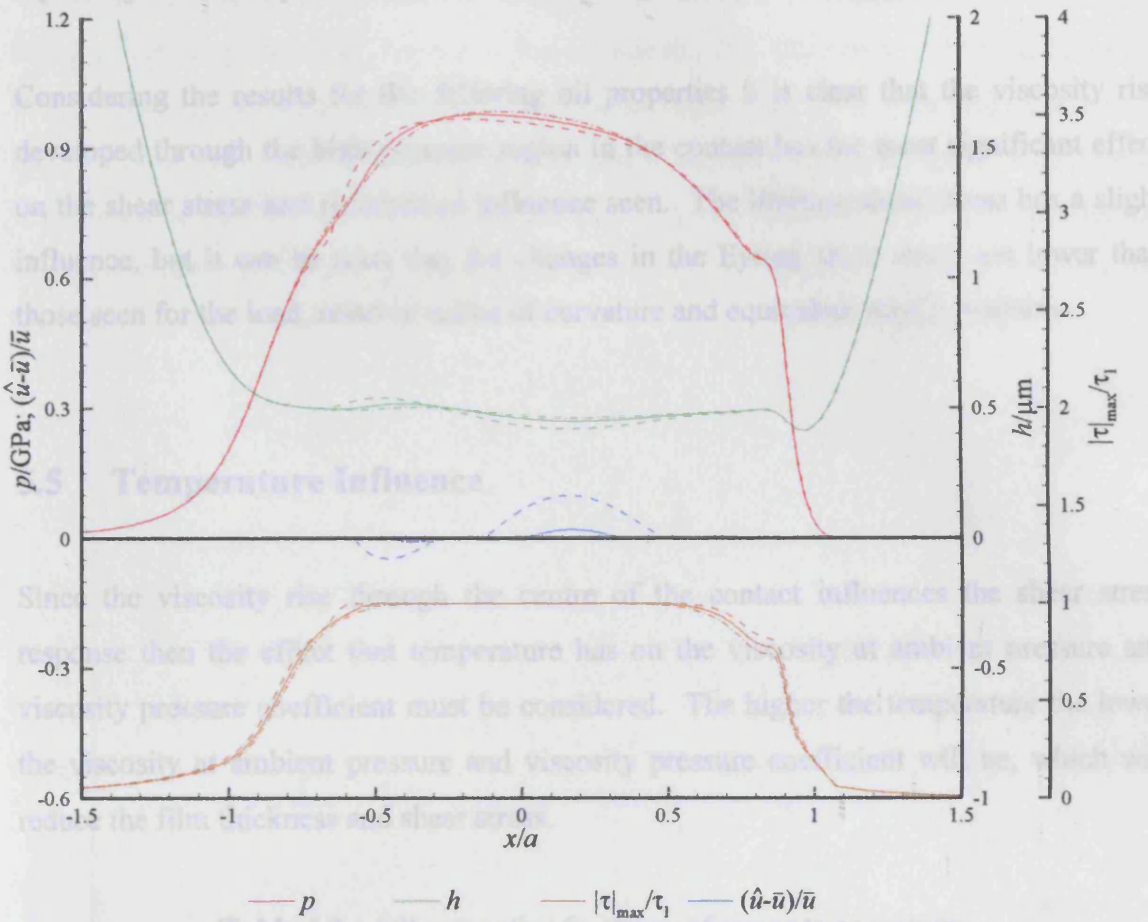

Figure 5.18 – Response of the Gecim and Winer model to different limiting shear stress, original (solid), -10% (dashed) and +10% (dash-dotted)

The shear stress and slip are both seen to fall as the limiting shear stress is increased. The reason for this is quite complex as the limiting shear stress forms part of three terms which interact. The first use is in determining Σ, the limiting shear stress is in the denominator so increasing it will reduce Σ. The limiting shear stress is also used in equation (5.3) to calculate the non-dimensional sliding, again it is present on the denominator and will reduce the value. Due to the reduction in Σ the maximum sliding allowed by the oil will increase. Since the non-dimensional sliding has reduced the shear stress will become smaller, but this interaction is highly non-linear due to the manner in which the terms interact. Finally, the limiting shear stress is used with the maximum shear stress developed through the contact to determine the maximum non-dimensional shear stress. In all of the interactions, increasing the limiting shear stress will allow the

non-dimensional shear stress to fall. It can be seen that the actual reduction is small, especially when compared to the results for changing the viscosity pressure coefficient.

Considering the results for the differing oil properties it is clear that the viscosity rise developed through the high pressure region in the contact has the most significant effect on the shear stress and rheological influence seen. The limiting shear stress has a slight influence, but it can be seen that the changes in the Eyring shear stress are lower than those seen for the load, relative radius of curvature and equivalent elastic modulus.

5.5 Temperature Influence

Since the viscosity rise through the centre of the contact influences the shear stress response then the effect that temperature has on the viscosity at ambient pressure and viscosity pressure coefficient must be considered. The higher the temperature the lower the viscosity at ambient pressure and viscosity pressure coefficient will be, which will reduce the film thickness and shear stress.

Table 5.2 – Oil properties for three reference temperatures

Oil Property	90°C	100°C	110°C
η_0	0.0059Pas	0.0048Pas	0.0039Pas
α	12.54GPa^{-1}	11.78GPa^{-1}	11.47GPa^{-1}

Three different temperatures of 90°C, 100°C and 110°C are considered. The oil properties given in table 5.1 are for 100°C and using known viscosities at other temperatures the oil properties at 90°C and 110°C can be determined using the approach of Wu, Klaus and Duda (1989). The properties for the three temperatures are presented in Table 5.2. It should be noted that the viscosity pressure coefficient for 100°C is different in table 5.2 to that in table 5.1. This is because the viscosity pressure coefficients above are those given using the approach of Wu, Klaus and Duda (1989). Despite the difference, the values are taken as they can be used to give a good indication of how the temperatures will effect the shear stress response.

Figure 5.19 shows the response of the Eyring model to the three different temperatures. The pressure remains unaffected by the change in the bulk oil temperature, however, the effect of the temperature is clearly seen to change the film thickness. As the temperature reduces the viscosity increases developing thicker films. Comparing the film thickness of the 90°C and 110°C results with the 100°C case, it can be seen that the greater difference is achieved for the 90 °C result than for the 110 °C result. This is a result of the difference in the viscosity at ambient pressure in table 5.2, where the difference is greater for the 90°C result.

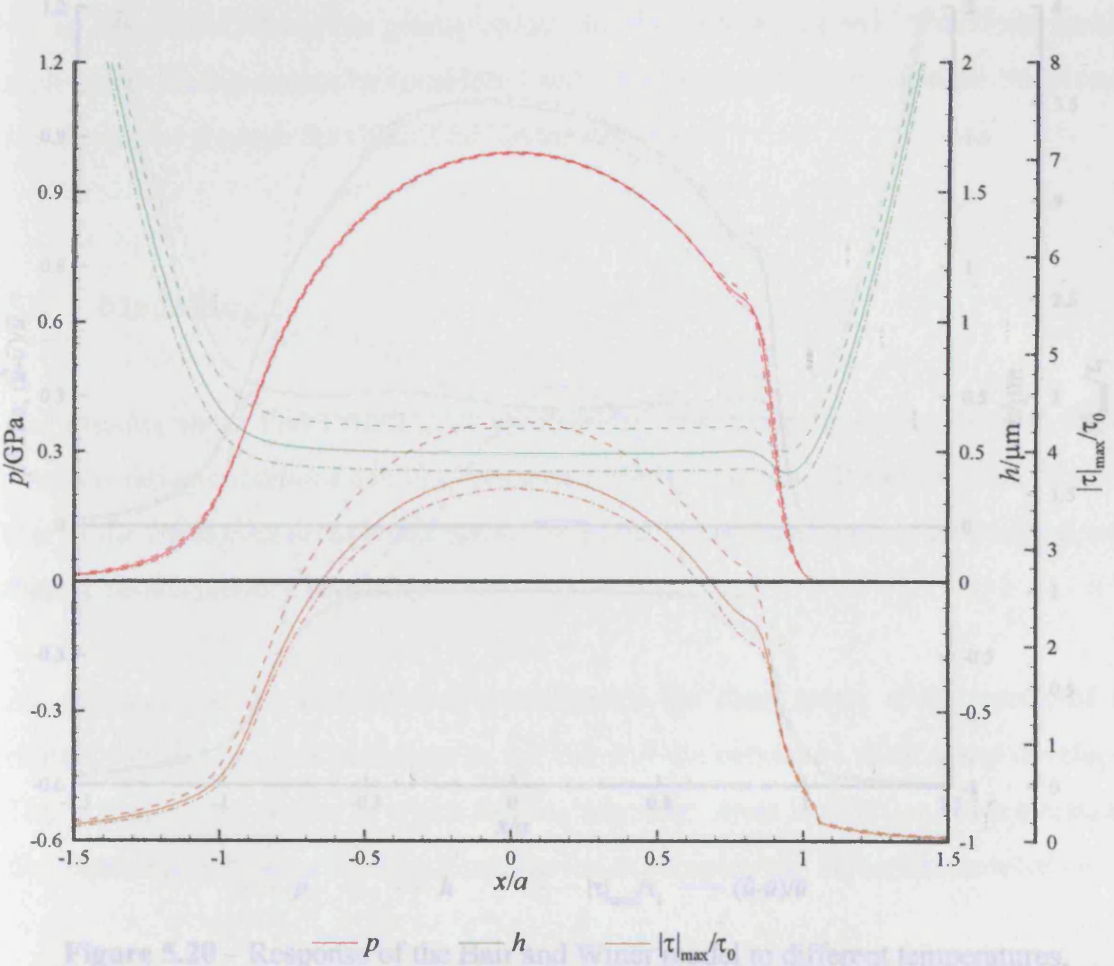

Figure 5.19 – Response of the Eyring model to different temperatures, 100°C (solid), 90°C (dashed) and 110°C (dash-dotted)

The shear stress increases as the temperature drops, which is what was predicted previously. It can be seen that the increase in the shear stress as the temperature drops from 110°C to 100°C is lower than the increase developed when the temperature drops

from 100°C to 90°C. This is similar to that seen in the film thickness and must be a combination of the film thickness and the viscosity pressure coefficient.

A converged result for the Gecim and Winer model could not be achieved at 90°C, the pressure distribution becomes very flat and the solution becomes very sensitive to small changes in the pressure and film thickness. Instead, the Bair and Winer result is shown in Figure 5.20.

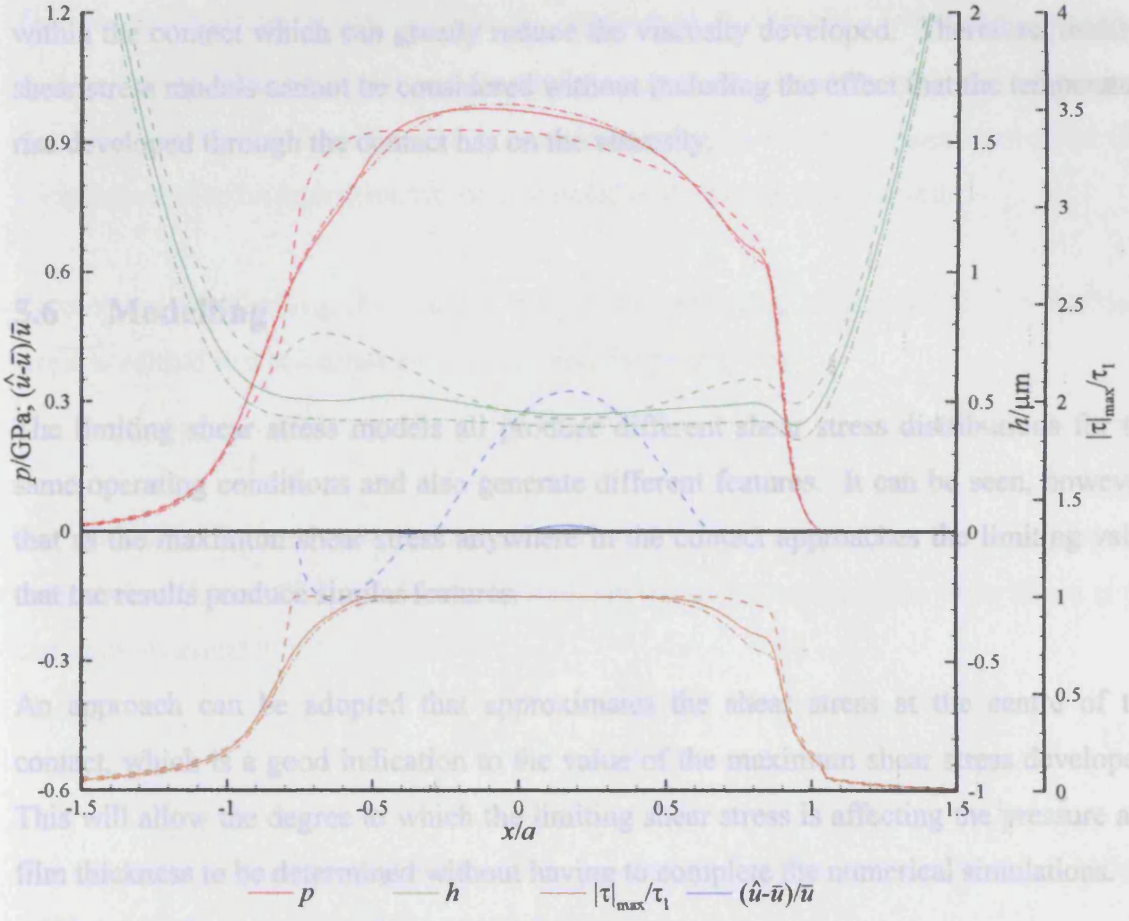

Figure 5.20 – Response of the Bair and Winer model to different temperatures, 100°C (solid), 90°C (dashed) and 110°C (dash-dotted)

Unlike the Eyring result, large changes in the pressure occur due to the influence that the temperature has over the amount of limiting shear stress influence developed. The amplitude of the film thickness kink is seen to rise as the temperature falls, the kink is also seen to move further into the inlet region as the temperature reduces.

The region over which the shear stress is at the limiting value increases as the temperature falls, this is a result of the viscosity increasing the non-dimensional sliding throughout the contact. The 110°C result does produce a shear stress near the limiting value, but the limiting value is not reached and thus no slip is generated within the region. The kink that is seen in the film thickness is a result of the limiting shear stress approaching the limiting value, the same result that was seen in figure 5.3.

Considering the response of the limiting shear stress to the temperature, in figure 5.20, and the viscosity, in figures 5.14 and 5.16, it is clear that the viscosity determines the extent of the limiting shear stress developed. Significant temperatures can be generated within the contact which can greatly reduce the viscosity developed. Therefore, limiting shear stress models cannot be considered without including the effect that the temperature rise developed through the contact has on the viscosity.

5.6 Modelling

The limiting shear stress models all produce different shear stress distributions for the same operating conditions and also generate different features. It can be seen, however, that as the maximum shear stress anywhere in the contact approaches the limiting value that the results produce similar features.

An approach can be adopted that approximates the shear stress at the centre of the contact, which is a good indication to the value of the maximum shear stress developed. This will allow the degree to which the limiting shear stress is affecting the pressure and film thickness to be determined without having to complete the numerical simulations.

At the centre of the contact it can be assumed that the pressure gradient is zero, or at least very close to it. This ensures that the pressure gradient term, Σ, is zero which allows the equations developed in section 3.8.2 to be used. Considering the equations developed for the Bair and Winer model, equation (3.29) can be used and is re-written below for clarity:

$$\lambda = 1 - e^{-U_s}$$

Since the pressure gradient term, Σ, is zero, the shear stress is equal throughout the film and is determined by λ alone from:

$$\tau = \lambda \tau_l$$

The non-dimensional shear stress, $|\tau|_{max} / \tau_l$, at the centre of the contact is simply equal to λ. Therefore, the equations developed in section 3.8.2 can be used to calculate the shear stress at the centre of the contact. However, the non-dimensional sliding demand, U_s, written in equation (5.4) still needs to be determined.

The non-dimensional sliding demand, U_s, presented in equation (5.4) comprises of five terms. The entrainment speed and the slide roll ratio are simple to calculate, as they are given directly by the operating conditions. However, the viscosity, shear stress and film thickness need to be approximated or calculated at the centre of the contact.

Considering the limiting shear stress first, it was shown in section 1.7.2 that the shear stress is related to a reference shear stress and the pressure by:

$$\tau_l = \tau_{l0} + \alpha_\tau p$$

The pressure at the centre of the contact will be close to the maximum Hertzian pressure, as given in equation (5.2). Substituting this into the above equation the shear stress at the centre of the contact is determined by:

$$\tau_l = \tau_{l0} + \alpha_\tau \sqrt{\frac{E'w'}{2\pi R}} \tag{5.5}$$

The viscosity relationship that is used was presented in section 2.5, the Roelands (1966) relationship was adopted and is re-written below:

$$\eta = \eta_0 e^{\ln(\eta_0 + 9.67)\left[\left(1+5.1\times10^{-9}p\right)^Z - 1\right]}$$

The parameter Z is determined using the viscosity pressure coefficient, α, and the viscosity at ambient pressure, η_0. Therefore this value is fixed for the whole of the contact region, it is only the pressure that influences the viscosity through the contact.

The viscosity at the centre of the contact can be determined using the maximum Hertzian pressure, in equation (5.2), allowing the viscosity to be written as:

$$\eta = \eta_0 e^{\ln(\eta_0+9.67)\left[\left(1+5.1\times10^{-9}\,p_0\right)^Z-1\right]} \tag{5.6}$$

The film thickness at the centre of the contact can be approximated using equation (5.1). Using equations (5.1), (5.5) and (5.6) and the operating conditions the non-dimensional sliding demand at the centre of the contact, equation (5.4), can be determined. This sliding demand can then be used in the equations in section 3.8.2 to calculate the shear stress at the centre of the contact. The results of this approach can be compared to the actual maximum value of the shear stress developed for the Bair and Winer, Circular and Gecim and Winer models over the range of slide roll ratios is shown in Figure 5.22.

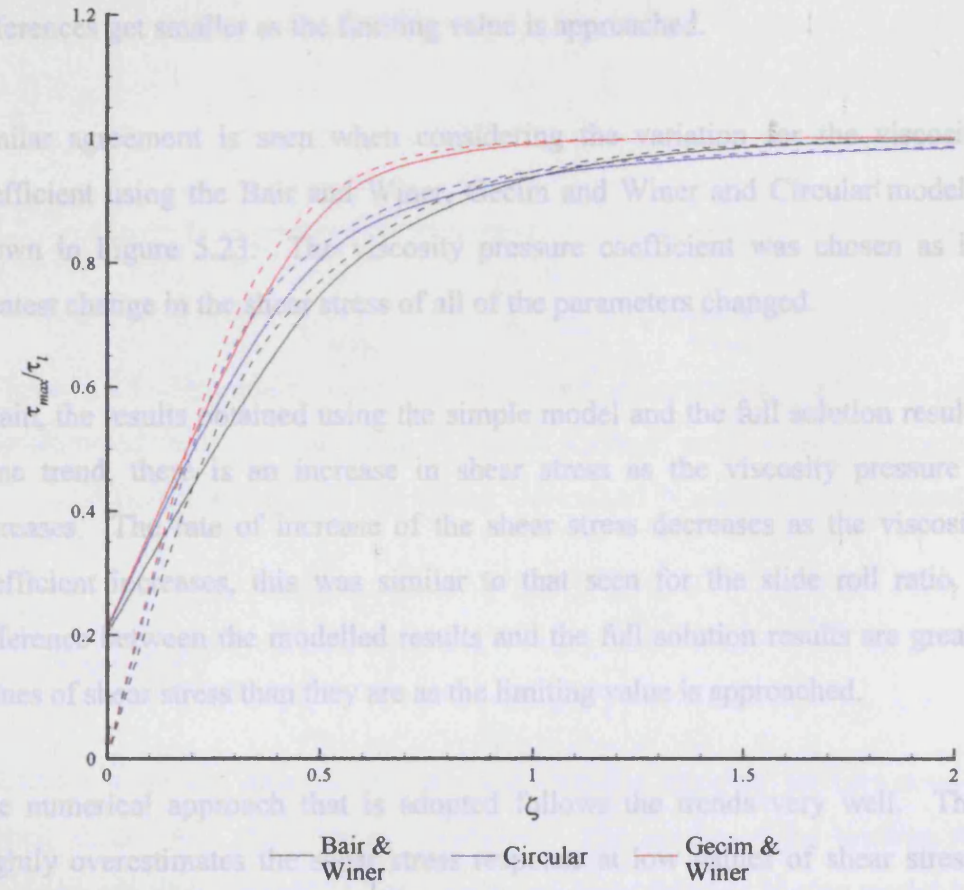

Figure 5.22 – Maximum non-dimensional shear stress throughout contact (solid) and approximated shear stress at centre of contact (dashed) variation with slide roll ratio for the Bair and Winer, Gecim and Winer and Circular models

The maximum value of the non-dimensional shear stress is dominated by the pressure gradient term, Σ, at low slide roll ratios. For low slide roll ratios the non-dimensional sliding is not large enough to produce a value of λ that is greater than Σ. This response is seen in the modelled result, where the values begin at zero and increase as the slide roll ratio is increased.

The modelled approach gives the same response as the actual solution, where the rate at which the shear stress increases reduces as the slide roll ratio is increased. It is seen that the more aggressive Gecim and Winer model increases more rapidly at lower slide roll ratio and flattens off before the Bair and Winer and Circular models. The Circular model approaches the limiting value quickly initially, but after $\zeta = 1.0$ the rate at which it approaches the limiting value decreases. All of the results obtained by using the simple model presented have a tendency to overestimate the shear stress. The differences between the two results are most noticeable at low values of shear stress and the differences get smaller as the limiting value is approached.

Similar agreement is seen when considering the variation for the viscosity pressure coefficient using the Bair and Winer, Gecim and Winer and Circular models, which is shown in Figure 5.23. The viscosity pressure coefficient was chosen as it gives the greatest change in the shear stress of all of the parameters changed.

Again, the results obtained using the simple model and the full solution results show the same trend, there is an increase in shear stress as the viscosity pressure coefficient increases. The rate of increase of the shear stress decreases as the viscosity pressure coefficient increases, this was similar to that seen for the slide roll ratio. Also, the difference between the modelled results and the full solution results are greater at small values of shear stress than they are as the limiting value is approached.

The numerical approach that is adopted follows the trends very well. The approach slightly overestimates the shear stress response at low values of shear stress, but good agreement is seen as the limiting value is approached. Nevertheless, the approach gives a conservative estimate of the maximum shear stress developed and can be used to

determine the influence that the limiting shear stress model has on the pressure and film thickness in the full solution.

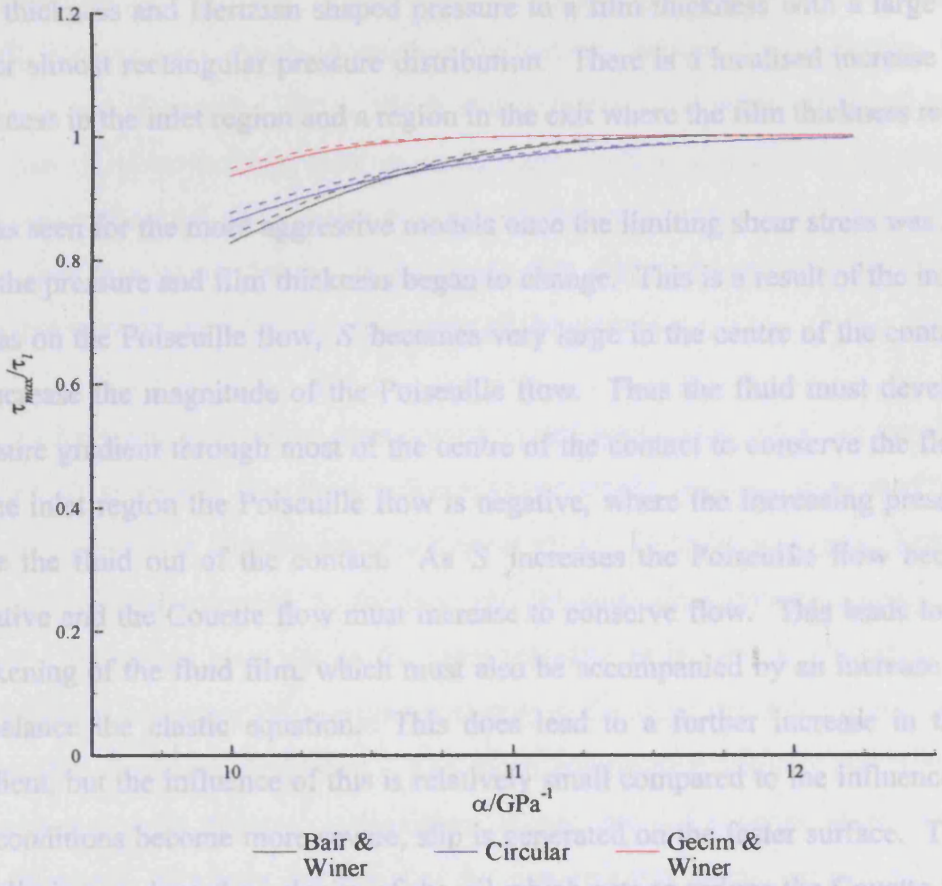

Figure 5.23 – Maximum non-dimensional shear stress throughout contact (solid) and approximated shear stress at centre of contact (dashed) variation with viscosity pressure coefficient for the Bair and Winer, Gecim and Winer and Circular models

5.7 Summary

Results are presented under isothermal conditions that show the features developed when limiting shear stress rheology is considered. The results of the limiting shear stress rheologies are compared to the non-Newtonian shear thinning Eyring model to allow the differences to be compared to another non-Newtonian rheological model.

The limiting shear stress models give similar results to an Eyring model with $\tau_l = 3\tau_0$ as long as the shear stress in the fluid does not approach the limiting value. Once the shear stress begins to approach the limiting value the rheology changes the response of the pressure and film thickness. The pressure and film thickness change from a flat parallel film thickness and Hertzian shaped pressure to a film thickness with a large kink and a flatter almost rectangular pressure distribution. There is a localised increase in the film thickness in the inlet region and a region in the exit where the film thickness reduces.

It was seen for the more aggressive models once the limiting shear stress was approached that the pressure and film thickness began to change. This is a result of the influence that S has on the Poiseuille flow, S becomes very large in the centre of the contact and acts to increase the magnitude of the Poiseuille flow. Thus the fluid must develop a lower pressure gradient through most of the centre of the contact to conserve the flow of fluid. In the inlet region the Poiseuille flow is negative, where the increasing pressure tries to force the fluid out of the contact. As S increases the Poiseuille flow becomes more negative and the Couette flow must increase to conserve flow. This leads to a localised thickening of the fluid film, which must also be accompanied by an increase in pressure to balance the elastic equation. This does lead to a further increase in the pressure gradient, but the influence of this is relatively small compared to the influence of S. As the conditions become more severe, slip is generated on the faster surface. The effect of the slip is to reduce the velocity of the oil which acts to reduce the Couette flow. Thus the film thickness must increase further so that flow continuity is obtained, which continues to make the feature more prominent.

In the exit region, the Poiseuille flow is positive and the increase in the flow generated by S must be balanced by a reduction in the Couette flow. This is achieved by reducing the film thickness, which must also develop a smaller pressure to balance the elastic deflection equation. When slip is generated it occurs on the slower surface acting to increases the Couette flow. In a similar manner to the inlet, this requires greater change in the film thickness to maintain flow.

The interaction of both flows is complex, the film thickness must change shape and the pressure gradient must reduce to allow conservation of flow. Both factors have a knock

on effect on other terms, some of which help balance the flows and others that do not. The solution attained is the overall response that maintains the flow continuity.

The kink in the film thickness seen for the limiting shear stress models is not the first illustration of such a feature. The circular model was first presented by Lee and Hamrock (1990), but was later shown to produce the features associated with limiting shear stress by Shieh and Hamrock (1991). Shieh and Hamrock used higher slide roll ratios which give rise to more limiting shear stress features. They showed that the film thickness increases in the inlet and reduces at a location past the centre of the contact towards the exit. They also found that the change in the film shape was accompanied by a reduction in the pressure gradient through the centre of the contact.

Various operating parameters were varied to determine what influence they have on the limiting shear stress. The load and equivalent elastic modulus generated very similar changes in pressure, film thickness and shear stress. The relative radius of curvature also generated similar responses to the load and the equivalent elastic modulus, except that reducing the radius gave rise to the same response as increasing the other two factors. The radius of curvature led to a greater change in the shear stress due to the larger change in the film thickness developed. The mean entrainment velocity did not change the pressure and shear stress an appreciable amount, but did generate large changes in the film thickness. Similarly, the effect of the viscosity at ambient pressure did not generate large changes in the shear stress and pressure, but did generate thicker films. The viscosity pressure coefficient gave rise to very large increases in the shear stress without changing the pressure and only had a slight influence on the film thickness. Changing the reference shear stress did not change the pressure or film thickness, but did have a noticeable effect on the shear stress developed. From the results, it is clear that the shear stress response is sensitive to changes in the viscosity, this is illustrated by the changes produced when the pressure changes for the load, radius of curvature, elastic modulus and pressure viscosity coefficient. It is the viscosity pressure coefficient, however, that generates the largest change in the shear stress. As a result of this the bulk oil temperature was increased and reduced by 10°C, where the results showed that increasing temperature reduces the limiting shear stress effects.

Comparison of the change in entrainment velocity (figure 5.7 and 5.8) and the change in ambient viscosity (figures 5.13 and 5.14) show that although the film thickness and pressure depend on the product of $\eta\bar{u}$, the shear stress and slip behaviour is different for cases with the same value of $\eta\bar{u}$ but different viscosity and velocity combinations. This indicates that the non-dimensionalisation of Reynolds equation does not carry through into non-Newtonian considerations.

An approach was presented that allowed the non-dimensional shear stress at the centre of the contact to be determined. The shear stress at the centre of the contact is a very good indication of the maximum value of the non-dimensional shear stress that is developed within the contact. Therefore, the approach allows the maximum non-dimensional shear stress to be approximated for any operating conditions. The approach gives very good agreement with the different models over a range of different conditions. This allows the operating conditions and oil properties to be used and the expected maximum non-dimensional shear stress to be calculated. The closer this value is to unity the more prominent the limiting shear stress features will be.

Since the temperature generated within the contact is dependent on the shear stress developed, the limiting shear stress models may not produce the film thickness kink and pressure changes seen. Under conditions that cause the shear stress to increase a greater amount of heat would be generated. The amount of heat generated may be large enough to reduce the severity of the conditions allowing the shear stress to reduce. Therefore the interaction of the two terms is very important.

The friction force produced by all the models over the range of conditions considered was over 20N. The difference between the different models for all the conditions was less than 3N. The only exception to these rules were the viscosity pressure coefficient results. The -10% case developed a friction force of under 20N and the +10% case had larger differences between the friction forces. The Eyring model for the +10% viscosity pressure coefficient developed a friction force of 33N, whereas the limiting shear stress models did not exceed 28N. Considering the response of the models as a whole over the large range of different operating conditions, the differences are not very significant. It

should be noted that the friction forces presented correspond to unit length in the y direction.

Chapter 6

Thermal Non-Newtonian Results

6.0 Introduction

The results presented in Chapter 5 are taken a stage further by applying a full energy equation to the fluid and solids contained within and in the vicinity of the contact area. This allows the effects that temperature changes have on the viscosity to be introduced to the non-Newtonian solutions presented in the previous chapters.

The different rheological models produce different shear stress distributions within the oil film and the thermal approach allows the effects that the shear stress has on the temperature developed to be assessed. The temperature distributions obtained using limiting shear stress models can be compared to the Eyring shear thinning results. In addition, the influence that slip has on the temperature developed both within the oil film and at each surface is also presented.

In the first sections, results are presented at high speed corresponding to the conditions introduced in Chapter 5, allowing comparison of thermal and isothermal behaviour. These operating conditions are used to conduct a study into the influence that the viscosity has on the pressure, film thickness, temperature and shear stress generated. In the last section, the operating conditions are changed from the two disk arrangement used throughout Chapter 5, to a combination of a steel ball and a glass disk. The change in the operating conditions allows simulations to be undertaken that model the environment under which optical interferometry experiments are carried out. The speeds of the surfaces are varied to promote limiting shear stress features, which may be investigated experimentally.

6.1 Thermal – Isothermal Comparisons

Using the standard test case in Chapter 5 the thermal approach can be applied allowing the effect that the temperature increase in the contact has on the shear stress, film thickness and pressure developed to be assessed. As mentioned previously, the increase in temperature reduces the viscosity so it is expected that in general the temperature changes will reduce the amount of non-Newtonian behaviour developed.

The cases presented compare the rheological models under the same operating conditions and using the same oil properties as those presented in table 5.1, except for the slide roll ratio, which is chosen to give $\zeta = 1.50$. The results for the Eyring rheological model can be seen in Figure 6.1 showing the pressure, film thickness, temperature, viscosity and shear stress response. The temperatures illustrated are the mid-film, θ_m, upper surface, θ_1, and lower surface, θ_2, temperatures. The Eyring results are developed with the Eyring shear stress and limiting shear stress maintained at a fixed ratio of $\tau_l = 3\tau_0$.

It can be seen that the mid-film temperature follows the pressure response quite closely, it increases towards the centre of the contact and then begins to fall as the exit is approached. The mid-film temperature falls more dramatically when the restriction is reached near the exit. The surface temperatures increase up to the constriction at the exit, then they both fall. The lower (slower) surface experiences higher temperatures than the upper (faster) surface. This is due to the increased amount of time that the surface spends in the contact area and thus the heat is applied for a longer time causing a higher temperature rise and penetrating deeper into the surface.

The mid-film temperature is still calculated after the oil has cavitated, however, the heating terms that form part of the full energy equation are no longer considered. This essentially reduces the energy equations so that only the conduction and convection terms are balanced. The conductivity and density used in this case are those of the oil, even though it is accepted that they will generally be lower than this due to the presence of air. Although there may be slight inaccuracies, the effect that the temperatures calculated have on the remainder of the temperatures is very small. It should be noted that the calculations for both of the solids remain unaffected.

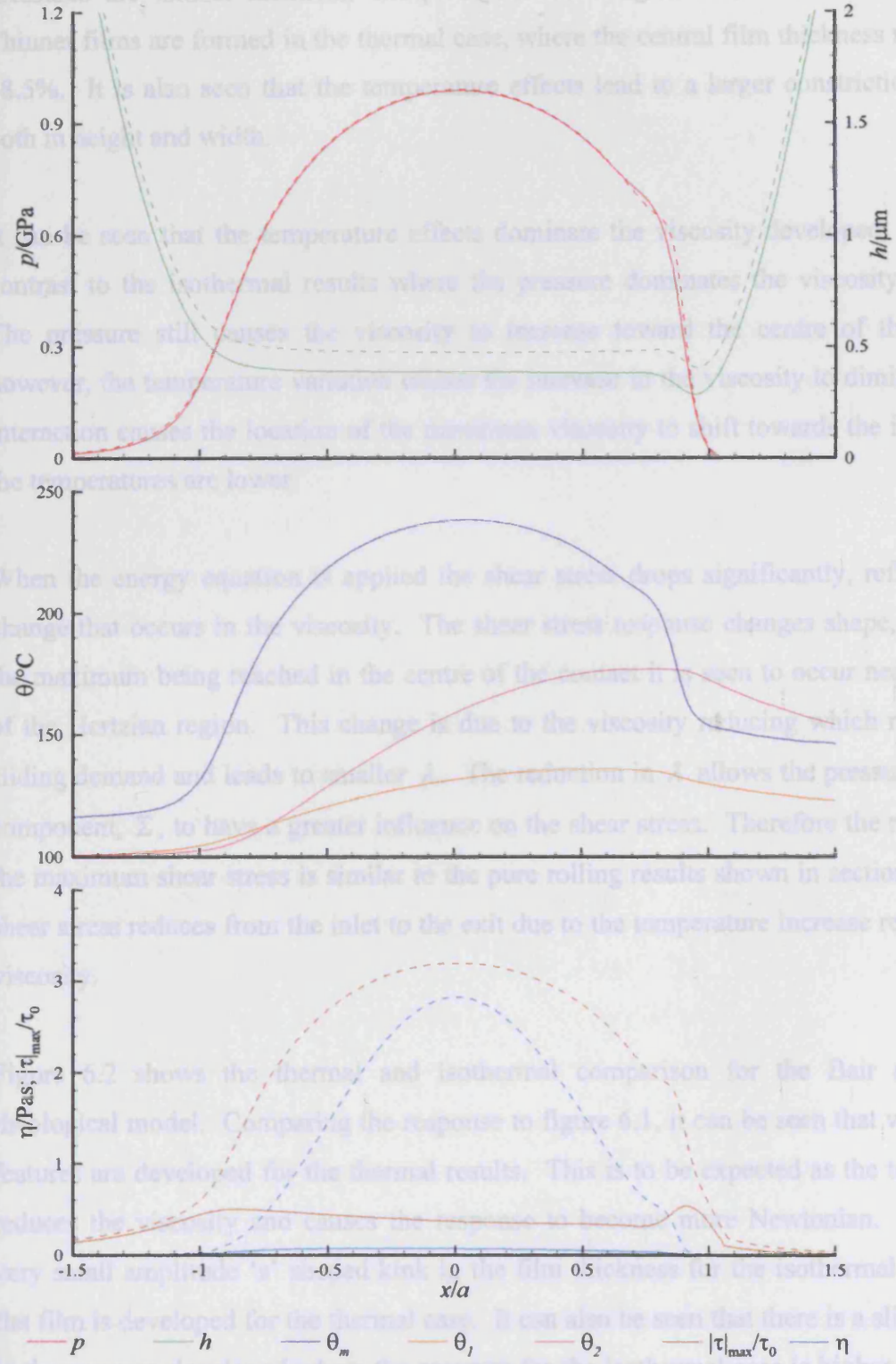

Figure 6.1 – Comparison of thermal (solid) and isothermal (dashed) conditions for the Eyring model showing the pressure, film thickness, mid-film temperature, surface temperatures, shear stress and viscosity

The temperature effects have very little influence on the pressure developed. These pressures are almost identical, except for a small region near the exit constriction. Thinner films are formed in the thermal case, where the central film thickness reduces by 18.5%. It is also seen that the temperature effects lead to a larger constriction feature, both in height and width.

It can be seen that the temperature effects dominate the viscosity developed, in marked contrast to the isothermal results where the pressure dominates the viscosity response. The pressure still causes the viscosity to increase toward the centre of the contact, however, the temperature variation causes the increase in the viscosity to diminish. This interaction causes the location of the maximum viscosity to shift towards the inlet where the temperatures are lower.

When the energy equation is applied the shear stress drops significantly, reflecting the change that occurs in the viscosity. The shear stress response changes shape, instead of the maximum being reached in the centre of the contact it is seen to occur near the edge of the Hertzian region. This change is due to the viscosity reducing which reduces the sliding demand and leads to smaller λ. The reduction in λ allows the pressure gradient component, Σ, to have a greater influence on the shear stress. Therefore the response of the maximum shear stress is similar to the pure rolling results shown in section 5.1. The shear stress reduces from the inlet to the exit due to the temperature increase reducing the viscosity.

Figure 6.2 shows the thermal and isothermal comparison for the Bair and Winer rheological model. Comparing the response to figure 6.1, it can be seen that very similar features are developed for the thermal results. This is to be expected as the temperature reduces the viscosity and causes the response to become more Newtonian. There is a very small amplitude 's' shaped kink in the film thickness for the isothermal case but a flat film is developed for the thermal case. It can also be seen that there is a slight change in the pressure developed where the pressure for the isothermal case is higher in the inlet than the thermal result.

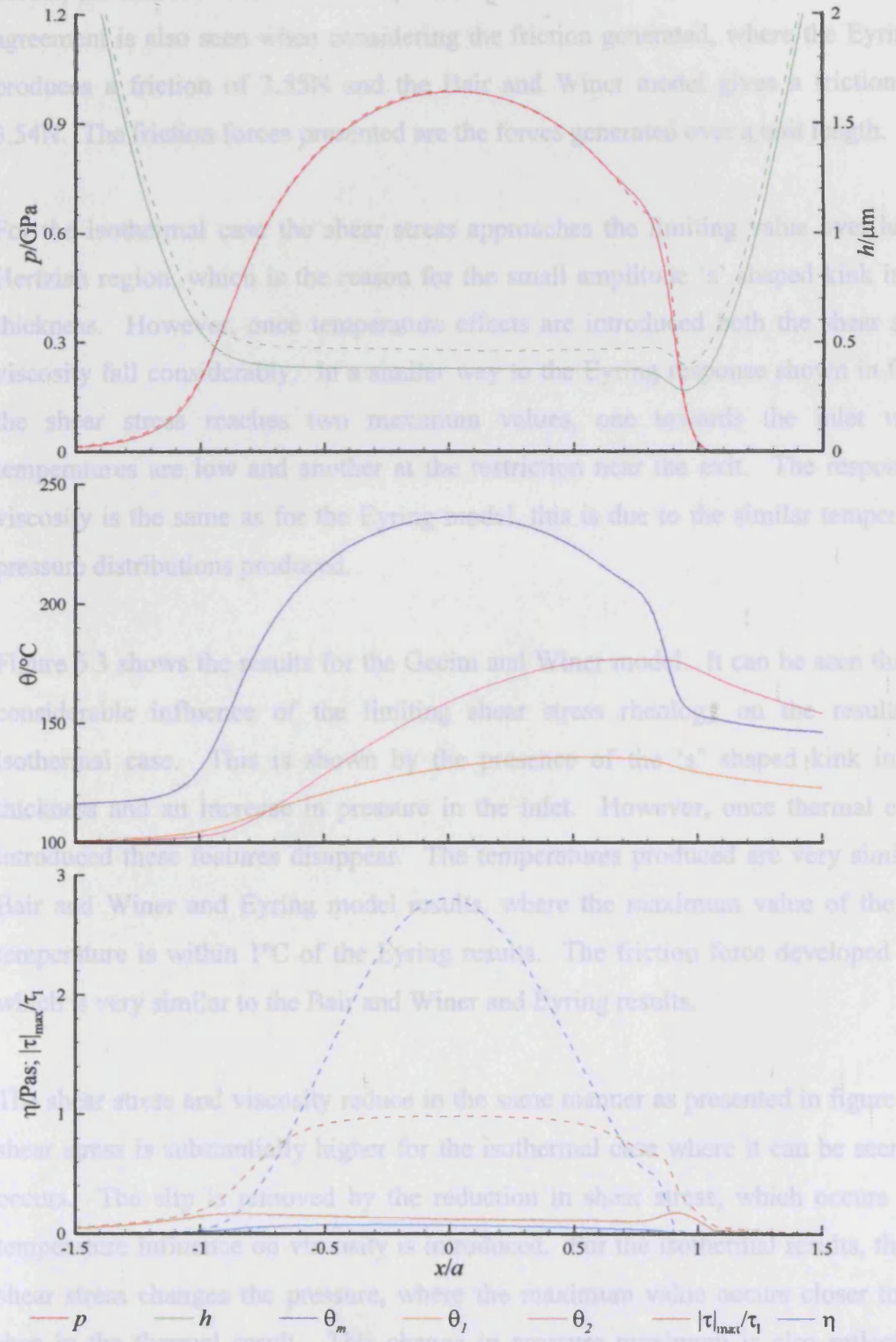

Figure 6.2 – Comparison of thermal (solid) and isothermal (dashed) conditions for the Bair and Winer model showing the pressure, film thickness, mid-film temperature, surface temperatures, shear stress and viscosity

The temperatures developed for the Bair and Winer model are similar to the Eyring model, the maximum mid-film temperature reached are within 2°C of one another. This agreement is also seen when considering the friction generated, where the Eyring model produces a friction of 3.55N and the Bair and Winer model gives a friction force of 3.54N. The friction forces presented are the forces generated over a unit length.

For the isothermal case the shear stress approaches the limiting value over half of the Hertzian region, which is the reason for the small amplitude 's' shaped kink in the film thickness. However, once temperature effects are introduced both the shear stress and viscosity fall considerably. In a similar way to the Eyring response shown in figure 6.1, the shear stress reaches two maximum values, one towards the inlet where the temperatures are low and another at the restriction near the exit. The response of the viscosity is the same as for the Eyring model, this is due to the similar temperature and pressure distributions produced.

Figure 6.3 shows the results for the Gecim and Winer model. It can be seen that there is considerable influence of the limiting shear stress rheology on the results for the isothermal case. This is shown by the presence of the 's' shaped kink in the film thickness and an increase in pressure in the inlet. However, once thermal effects are introduced these features disappear. The temperatures produced are very similar to the Bair and Winer and Eyring model results, where the maximum value of the mid-film temperature is within 1°C of the Eyring results. The friction force developed is 3.55N, which is very similar to the Bair and Winer and Eyring results.

The shear stress and viscosity reduce in the same manner as presented in figure 6.2. The shear stress is substantially higher for the isothermal case where it can be seen that slip occurs. The slip is removed by the reduction in shear stress, which occurs when the temperature influence on viscosity is introduced. For the isothermal results, the limiting shear stress changes the pressure, where the maximum value occurs closer to the inlet than in the thermal result. This change in pressure maximum is also reflected in the viscosity response, where the maximum value moves towards the inlet.

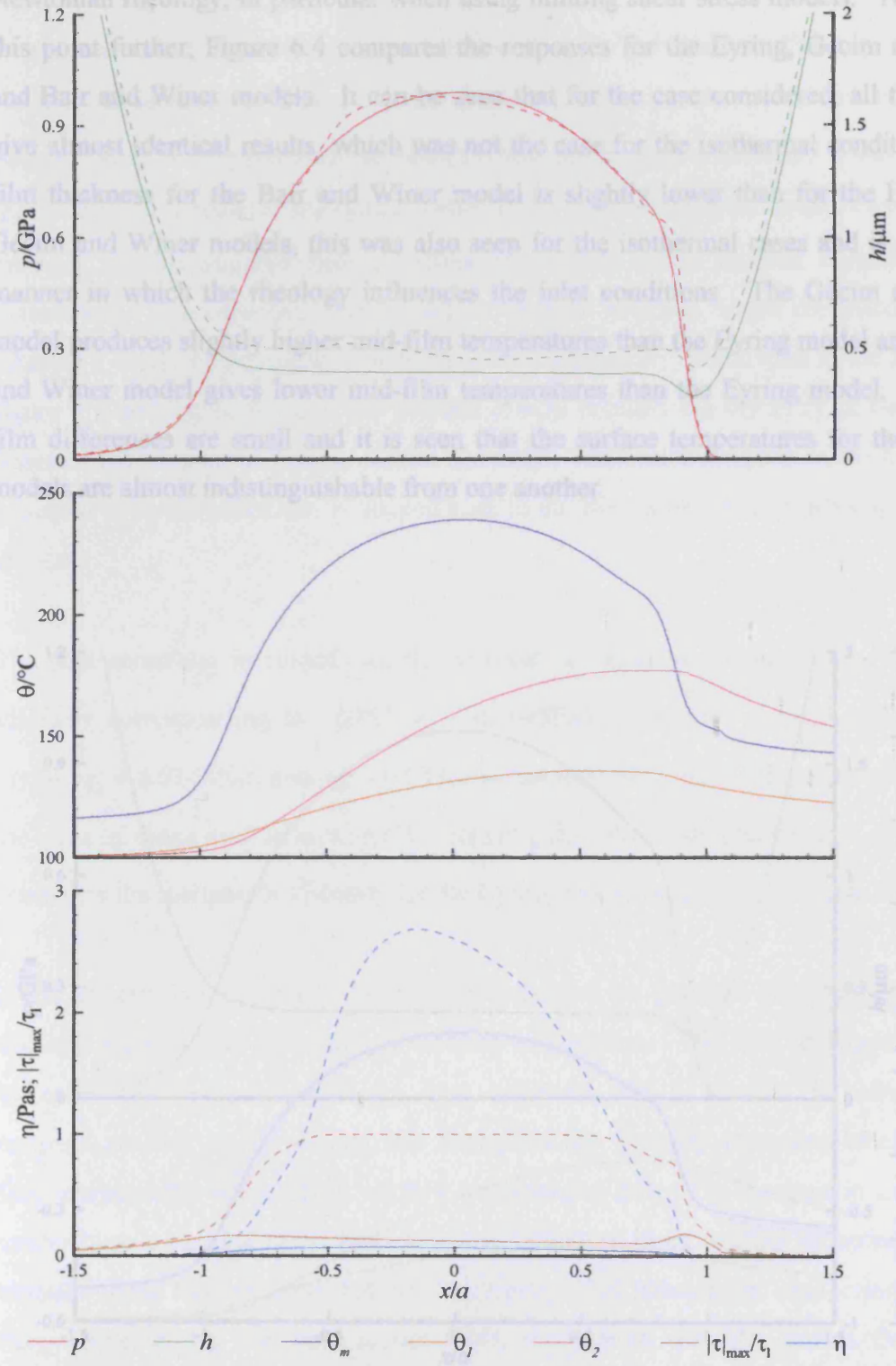

Figure 6.3 – Comparison of thermal (solid) and isothermal (dashed) conditions for the Gecim and Winer model showing the pressure, film thickness, mid-film temperature, surface temperatures, shear stress and viscosity

Figures 6.1 to 6.3 confirm that temperature effects must be considered when using non-Newtonian rheology, in particular when using limiting shear stress models. To illustrate this point further, Figure 6.4 compares the responses for the Eyring, Gecim and Winer and Bair and Winer models. It can be seen that for the case considered, all the models give almost identical results, which was not the case for the isothermal conditions. The film thickness for the Bair and Winer model is slightly lower than for the Eyring and Gecim and Winer models, this was also seen for the isothermal cases and is due to the manner in which the rheology influences the inlet conditions. The Gecim and Winer model produces slightly higher mid-film temperatures than the Eyring model and the Bair and Winer model gives lower mid-film temperatures than the Eyring model. The mid-film differences are small and it is seen that the surface temperatures for the different models are almost indistinguishable from one another.

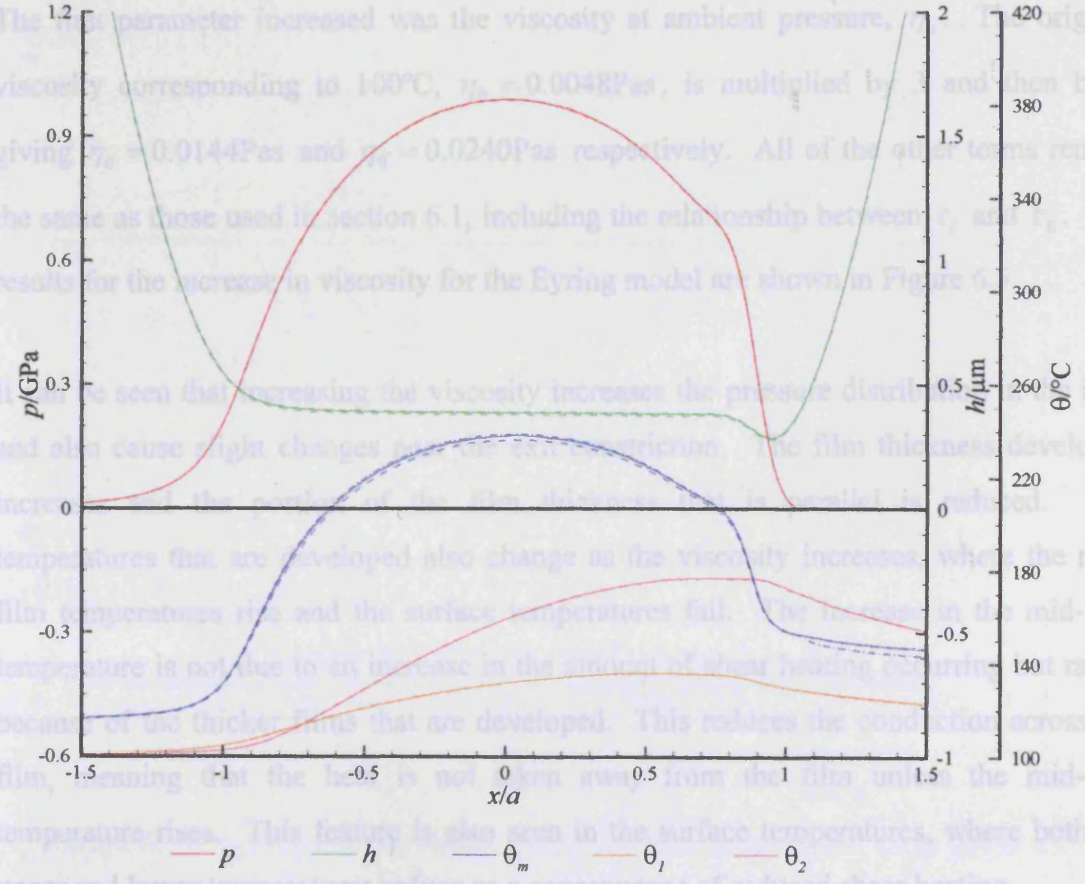

Figure 6.4 – Comparison of thermal results for the Eyring (solid), Bair and Winer (dashed) and Gecim and Winer (dash-dot) rheological models

6.2 Viscosity Influence

It was concluded from Chapter 5 that the viscosity pressure coefficient, α, and the viscosity at ambient pressure, η_0, both have an effect on the maximum value of the limiting shear stress produced. Although it was determined in section 5.4 that increasing η_0 resulted in a reduced maximum value of the non-dimensional shear stress, viscosity changes are investigated here to allow the effects under thermal conditions to be determined. The objective was to seek to create a situation where the features seen in Chapter 5 might be reproduced under thermal conditions. It was seen in the figures 6.1, figure 6.2 and figure 6.3 that the thermal effects reduced the viscosity of the fluid and lessened the pressure influence on the viscosity. Thus increasing the viscosity might be expected to re-introduce the 's' shaped kink in the film thickness and associated pressure response.

The first parameter increased was the viscosity at ambient pressure, η_0. The original viscosity corresponding to 100°C, $\eta_0 = 0.0048$Pas, is multiplied by 3 and then by 5 giving $\eta_0 = 0.0144$Pas and $\eta_0 = 0.0240$Pas respectively. All of the other terms remain the same as those used in section 6.1, including the relationship between τ_l and τ_0. The results for the increase in viscosity for the Eyring model are shown in Figure 6.5.

It can be seen that increasing the viscosity increases the pressure distribution in the inlet and also cause slight changes near the exit constriction. The film thickness developed increases and the portion of the film thickness that is parallel is reduced. The temperatures that are developed also change as the viscosity increases, where the mid-film temperatures rise and the surface temperatures fall. The increase in the mid-film temperature is not due to an increase in the amount of shear heating occurring but rather because of the thicker films that are developed. This reduces the conduction across the film, meaning that the heat is not taken away from the film unless the mid-film temperature rises. This feature is also seen in the surface temperatures, where both the upper and lower temperatures reduce as a consequence of reduced shear heating.

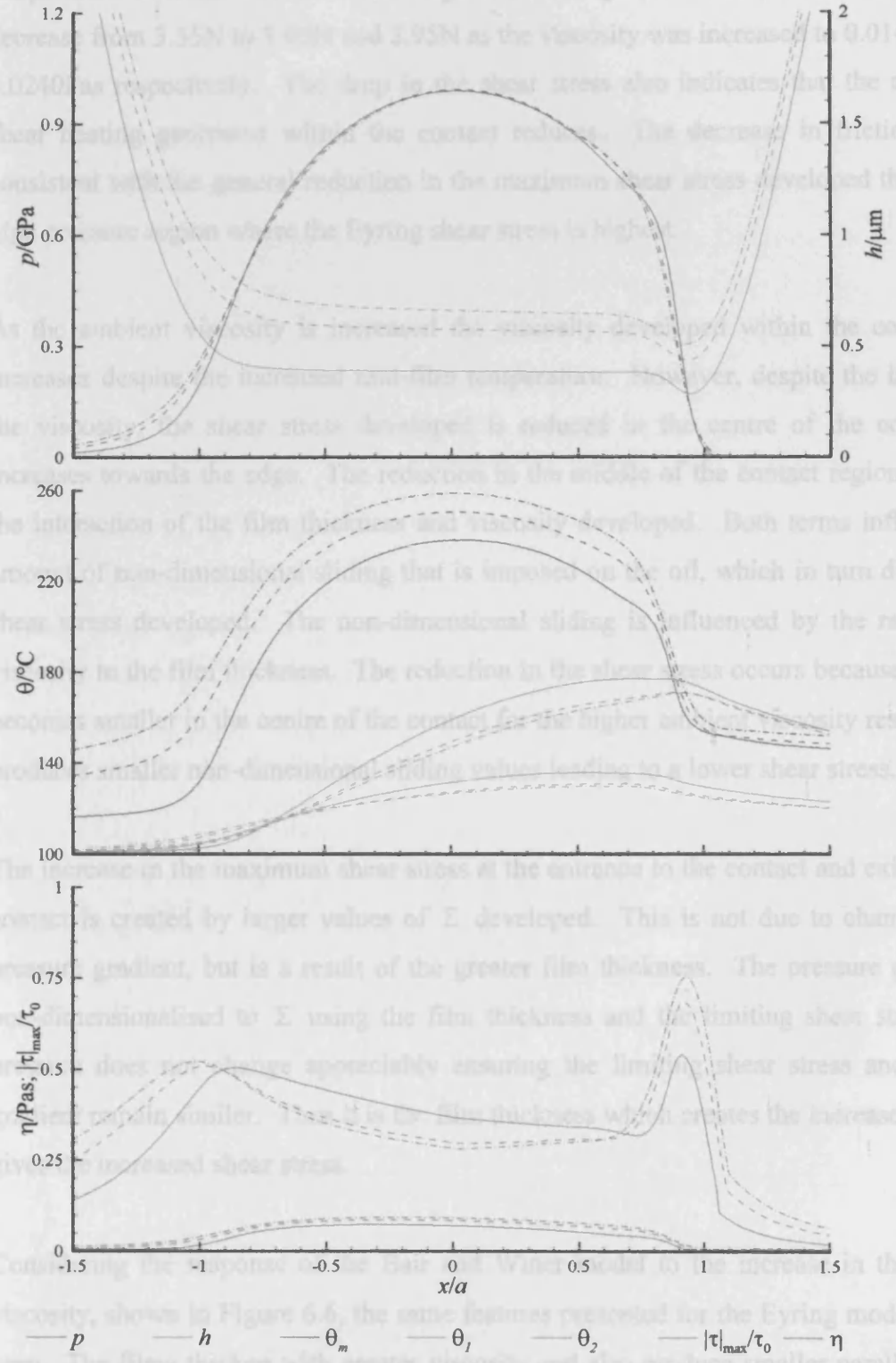

Figure 6.5 – Eyring model pressure, film thickness, temperature, shear stress and viscosity response for different viscosities, $\eta_0 = 0.0048$Pas (solid), $\eta_0 = 0.0144$Pas (dashed) and $\eta_0 = 0.0240$Pas (dash-dot)

Considering the friction force developed for the Eyring results, it is clear that the mid-film temperature does not increase due to greater heating. The friction force was seen to decrease from 3.55N to 3.05N and 2.95N as the viscosity was increased to 0.0144Pas and 0.0240Pas respectively. The drop in the shear stress also indicates that the amount of shear heating generated within the contact reduces. The decrease in friction is also consistent with the general reduction in the maximum shear stress developed through the high pressure region where the Eyring shear stress is highest.

As the ambient viscosity is increased the viscosity developed within the contact also increases despite the increased mid-film temperature. However, despite the increase in the viscosity, the shear stress developed is reduced in the centre of the contact and increases towards the edge. The reduction in the middle of the contact region is due to the interaction of the film thickness and viscosity developed. Both terms influence the amount of non-dimensional sliding that is imposed on the oil, which in turn dictates the shear stress developed. The non-dimensional sliding is influenced by the ratio of the viscosity to the film thickness. The reduction in the shear stress occurs because this ratio becomes smaller in the centre of the contact for the higher ambient viscosity results. This produces smaller non-dimensional sliding values leading to a lower shear stress.

The increase in the maximum shear stress at the entrance to the contact and exit from the contact is created by larger values of Σ developed. This is not due to changes in the pressure gradient, but is a result of the greater film thickness. The pressure gradient is non-dimensionalised to Σ using the film thickness and the limiting shear stress. The pressure does not change appreciably ensuring the limiting shear stress and pressure gradient remain similar. Thus it is the film thickness which creates the increase in Σ and gives the increased shear stress.

Considering the response of the Bair and Winer model to the increase in the ambient viscosity, shown in Figure 6.6, the same features presented for the Eyring model are also seen. The films thicken with greater viscosity and also produce smaller parallel regions for higher viscosities. The pressure changes very slightly, the mid-film temperature increases and the surface temperatures reduce.

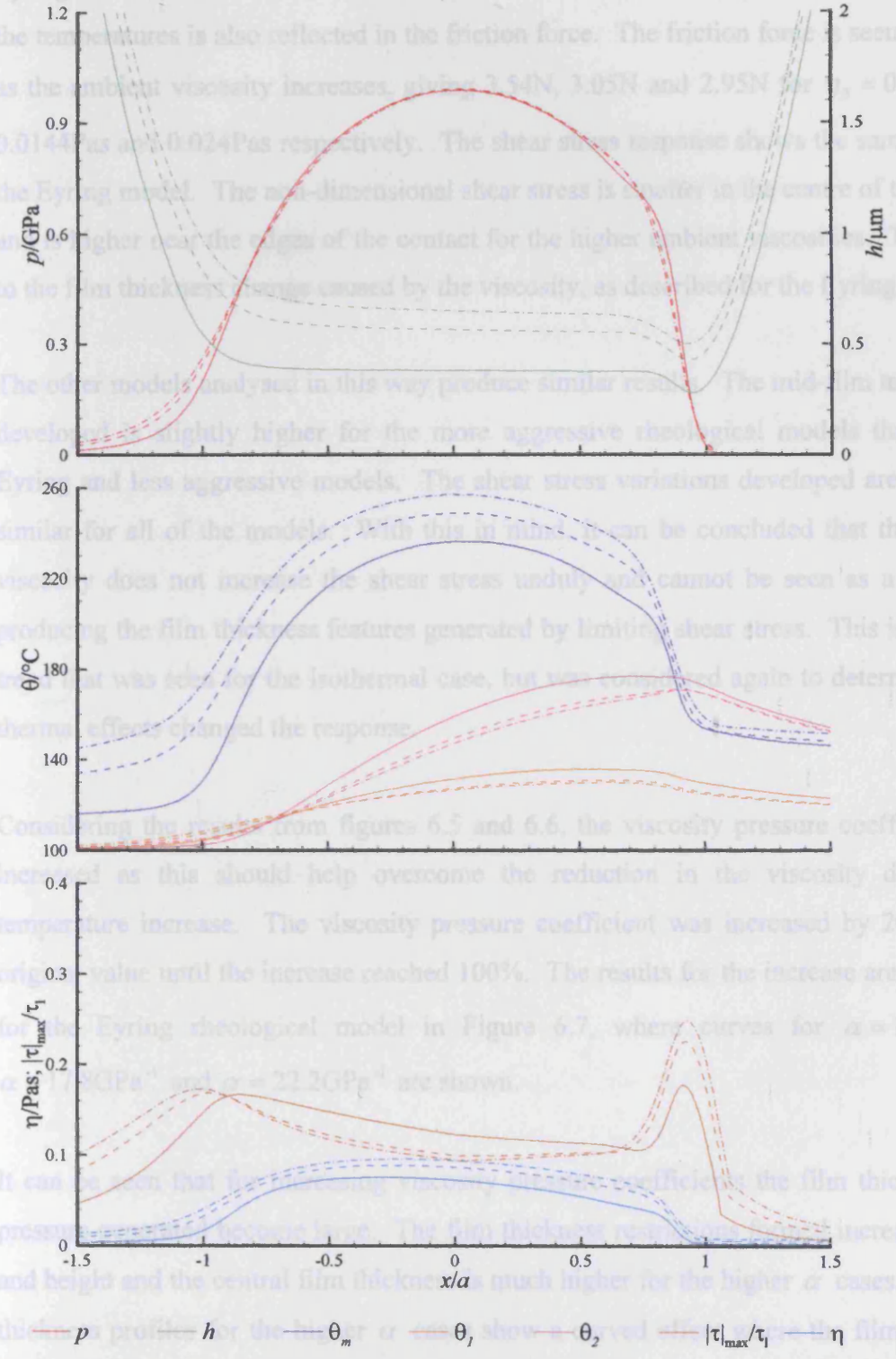

Figure 6.6 – Bair and Winer model pressure, film thickness, temperature, shear stress and viscosity response for different viscosities, $\eta_0 = 0.0048$Pas (solid), $\eta_0 = 0.0144$Pas (dashed) and $\eta_0 = 0.0240$Pas (dash-dot)

The temperatures produced are slightly lower for the Bair and Winer model than for the Eyring model, of the order 2°C at the centre of the contact. The good agreement seen in the temperatures is also reflected in the friction force. The friction force is seen to reduce as the ambient viscosity increases, giving 3.54N, 3.05N and 2.95N for $\eta_0 = 0.0048\text{Pas}$, 0.0144Pas and 0.024Pas respectively. The shear stress response shows the same trend as the Eyring model. The non-dimensional shear stress is smaller in the centre of the contact and is higher near the edges of the contact for the higher ambient viscosities. This is due to the film thickness change caused by the viscosity, as described for the Eyring model.

The other models analysed in this way produce similar results. The mid-film temperature developed is slightly higher for the more aggressive rheological models than for the Eyring and less aggressive models. The shear stress variations developed are also very similar for all of the models. With this in mind, it can be concluded that the ambient viscosity does not increase the shear stress unduly and cannot be seen as a means of producing the film thickness features generated by limiting shear stress. This is the same trend that was seen for the isothermal case, but was considered again to determine if the thermal effects changed the response.

Considering the results from figures 6.5 and 6.6, the viscosity pressure coefficient was increased as this should help overcome the reduction in the viscosity due to the temperature increase. The viscosity pressure coefficient was increased by 20% of the original value until the increase reached 100%. The results for the increase are presented for the Eyring rheological model in Figure 6.7, where curves for $\alpha = 11.1\text{GPa}^{-1}$, $\alpha = 17.8\text{GPa}^{-1}$ and $\alpha = 22.2\text{GPa}^{-1}$ are shown.

It can be seen that for increasing viscosity pressure coefficients the film thickness and pressure generated become large. The film thickness restrictions formed increase in area and height and the central film thickness is much higher for the higher α cases. The film thickness profiles for the higher α cases show a curved effect where the film thickness increases toward the centre of the contact. The pressures generated are slightly higher in the centre of the contact and the pressure spike developed is more noticeable as α is increased.

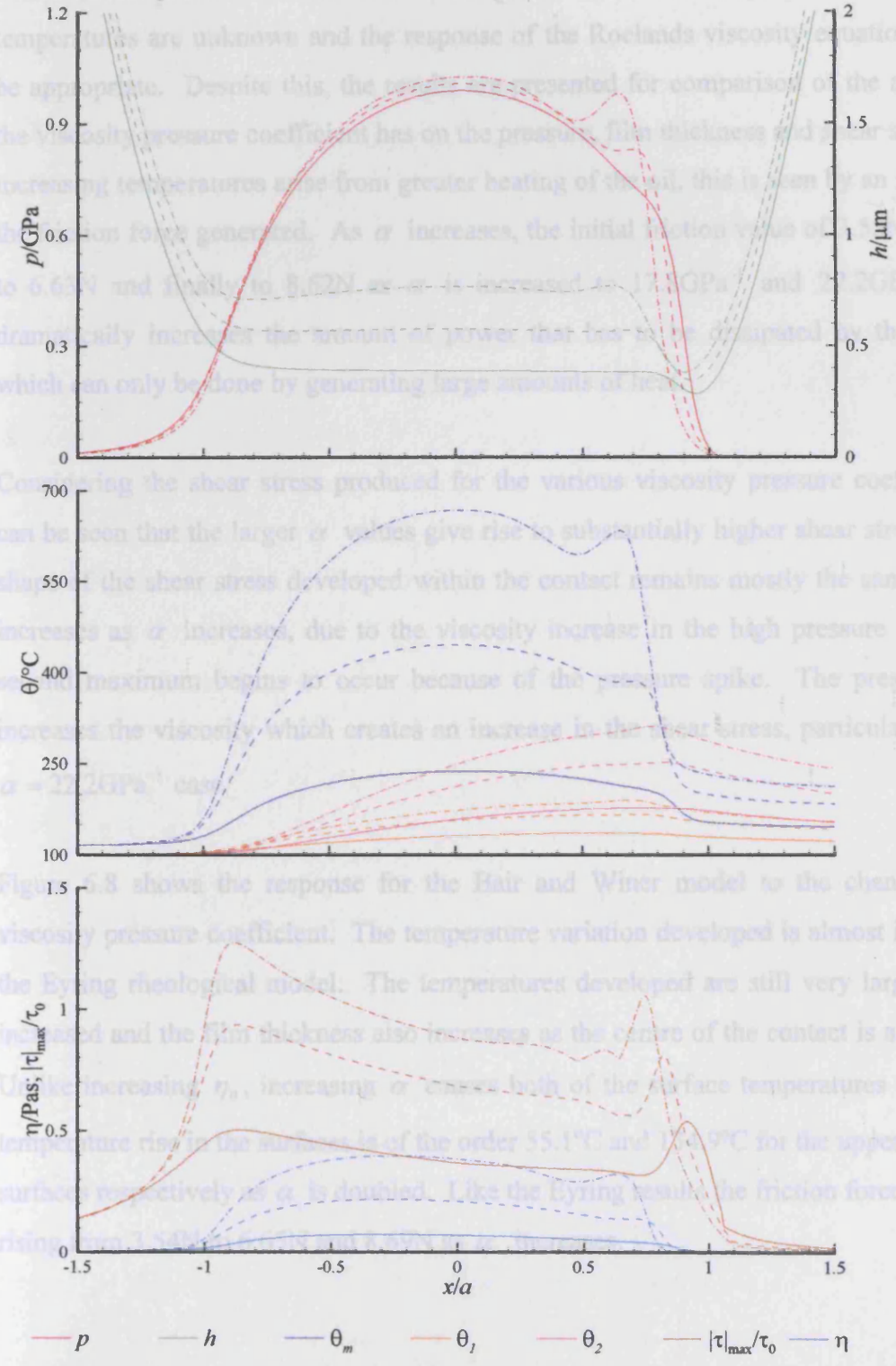

Figure 6.7 – Eyring model pressure, film thickness, temperature, shear stress and viscosity response for different viscosity pressure coefficients, $\alpha = 11.1\text{GPa}^{-1}$ (solid), $\alpha = 17.8\text{GPa}^{-1}$ (dashed) and $\alpha = 22.2\text{GPa}^{-1}$ (dash-dot)

The temperatures developed increase rapidly, in particular the mid-film temperature, where the maximum value rises 450°C as α is doubled. The oil properties at these high temperatures are unknown and the response of the Roelands viscosity equation may not be appropriate. Despite this, the results are presented for comparison of the affects that the viscosity pressure coefficient has on the pressure, film thickness and shear stress. The increasing temperatures arise from greater heating of the oil, this is seen by an increase in the friction force generated. As α increases, the initial friction value of 3.55N increases to 6.63N and finally to 8.62N as α is increased to 17.8GPa^{-1} and 22.2GPa^{-1}. This dramatically increases the amount of power that has to be dissipated by the oil film, which can only be done by generating large amounts of heat.

Considering the shear stress produced for the various viscosity pressure coefficients, it can be seen that the larger α values give rise to substantially higher shear stresses. The shape of the shear stress developed within the contact remains mostly the same and just increases as α increases, due to the viscosity increase in the high pressure region. A second maximum begins to occur because of the pressure spike. The pressure spike increases the viscosity which creates an increase in the shear stress, particularly for the $\alpha = 22.2\text{GPa}^{-1}$ case.

Figure 6.8 shows the response for the Bair and Winer model to the changes in the viscosity pressure coefficient. The temperature variation developed is almost identical to the Eyring rheological model. The temperatures developed are still very large as α is increased and the film thickness also increases as the centre of the contact is approached. Unlike increasing η_0, increasing α causes both of the surface temperatures to rise, the temperature rise in the surfaces is of the order 55.1°C and 134.9°C for the upper and lower surfaces respectively as α is doubled. Like the Eyring results the friction force increases, rising from 3.54N to 6.65N and 8.69N as α increases.

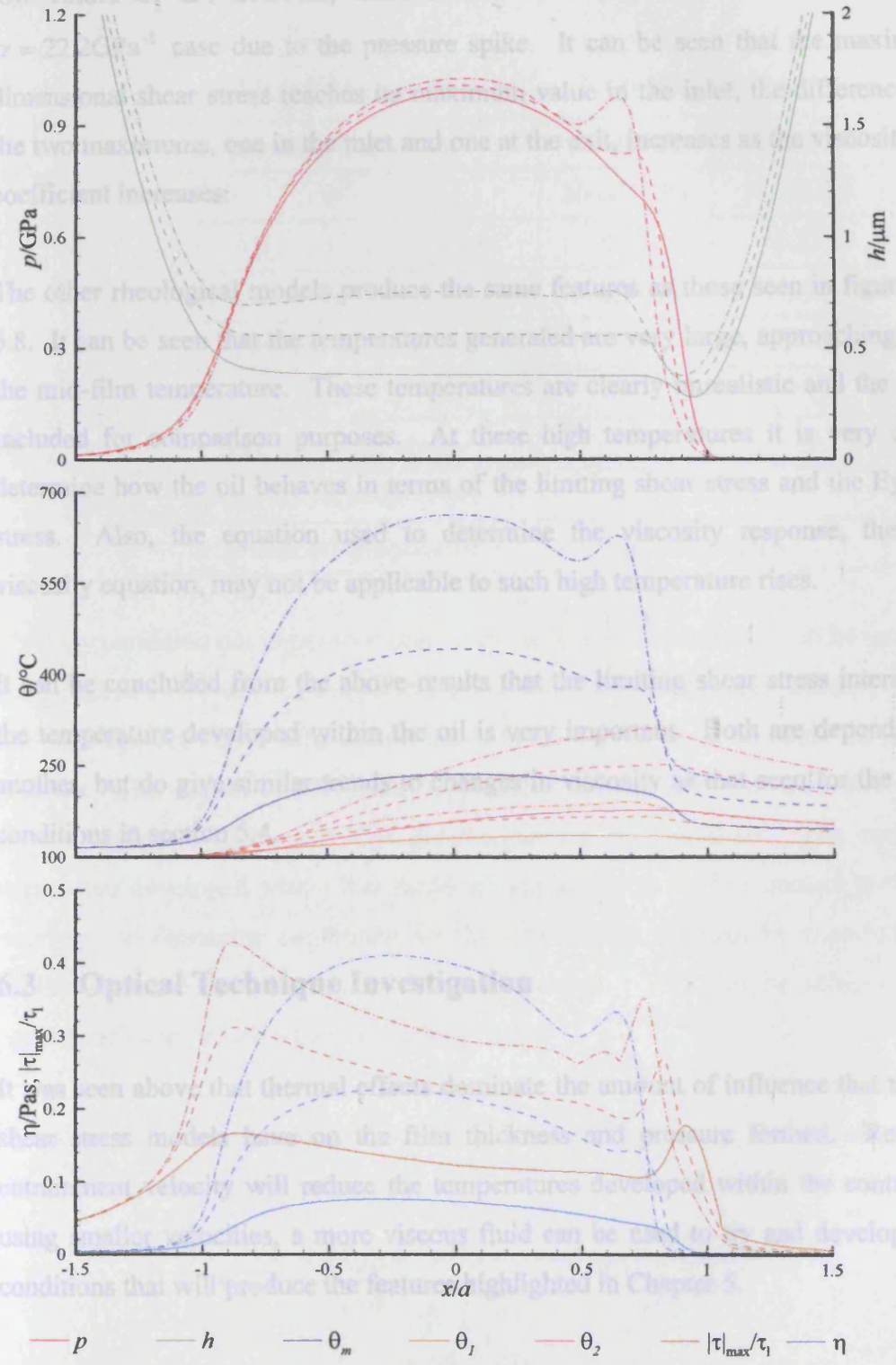

Figure 6.8 – Bair and Winer model pressure, film thickness, temperature, shear stress and viscosity response for different viscosity pressure coefficients, $\alpha = 11.1\text{GPa}^{-1}$ (solid), $\alpha = 17.8\text{GPa}^{-1}$ (dashed) and $\alpha = 22.2\text{GPa}^{-1}$ (dash-dot)

The limiting shear stress variation at high α values retains the same features as seen for low values of α, however, there is still an increase in the shear stress for the $\alpha = 22.2\text{GPa}^{-1}$ case due to the pressure spike. It can be seen that the maximum non-dimensional shear stress reaches its maximum value in the inlet, the difference between the two maximums, one in the inlet and one at the exit, increases as the viscosity pressure coefficient increases.

The other rheological models produce the same features as those seen in figures 6.7 and 6.8. It can be seen that the temperatures generated are very large, approaching 650°C for the mid-film temperature. These temperatures are clearly unrealistic and the results are included for comparison purposes. At these high temperatures it is very difficult to determine how the oil behaves in terms of the limiting shear stress and the Eyring shear stress. Also, the equation used to determine the viscosity response, the Roelands viscosity equation, may not be applicable to such high temperature rises.

It can be concluded from the above results that the limiting shear stress interaction with the temperature developed within the oil is very important. Both are dependent on one another, but do give similar trends to changes in viscosity as that seen for the isothermal conditions in section 5.4.

6.3 Optical Technique Investigation

It was seen above that thermal effects dominate the amount of influence that the limiting shear stress models have on the film thickness and pressure formed. Reducing the entrainment velocity will reduce the temperatures developed within the contact. When using smaller velocities, a more viscous fluid can be used to try and develop operating conditions that will produce the features highlighted in Chapter 5.

In order to allow the results to be validated, the operating conditions used correspond to an optical experimental rig and a specific oil chosen so that its properties are known. Spikes and co-workers [Glovnea and Spikes, 2001] used an optical interferometry

arrangement which they used to study the film thickness formed between a steel ball and a glass disk, the operating conditions and oil properties are listed in Table 6.1.

Table 6.1 – Operating condition for optical test case

Operating Condition	Steel Ball	Glass Disk
E	206.85GPa	72.00GPa
ν	0.30	0.23
R	9.525mm	∞
Applied load, W	20N	
Oil	5P4E	
Bulk oil temperature	50°C	
η_0 at bulk oil temperature	0.15Pas	
α at bulk oil temperature	$\alpha = 30.30\text{GPa}^{-1}$	

The only condition not imposed is that of the surface velocities, this is to be varied to give some indication of how the temperature, film thickness and pressure change. Before this can be done, the above information needs to be expressed in terms of the line contact problem. The optical interferometry approaches involve the study of a steel ball being pressed against a glass disk, this gives a circular point contact. The equations and approaches developed within this thesis are applicable to the line contact problem. It is necessary to determine conditions for the line contact that can be expected to give a similar response to that seen in the point contact. This can be achieved using an equivalencing approach which is outlined below.

Using suffix '*point*' to denote the point contact parameters the Hertz contact equations [Hertz,1881] are:

$$a_{point} = \left(\frac{3R_{point}W}{2E'}\right)^{\frac{1}{3}} \qquad a = \sqrt{\frac{8}{\pi}\frac{Rw'}{E'}}$$

$$p_{0\,point} = \left(\frac{3WE'^2}{2\pi^3 R_{point}^2}\right)^{\frac{1}{3}} \qquad p_0 = \sqrt{\frac{E'w'}{2\pi R}}$$

Equating the expressions for p_0 and a, R can be eliminated giving the equivalent load for the line contact, w', as:

$$w' = \left(\frac{9}{32} \frac{E'W^2}{R_{point}} \right)^{\frac{1}{3}} \tag{6.1}$$

Substituting equation (6.1) back into the equations gives the equivalent line contact relative radius of curvature, R, as:

$$R = \frac{\pi}{4} R_{point} \tag{6.2}$$

Using the data in table 6.1 the appropriate line contact load and relative radius of curvature can be determined from equations (6.1) and (6.2) and the values obtained are given in Table 6.2.

Table 6.2 – Equivalenced line contact conditions

w'	1.10kNm^{-1}
R	7.48mm
p_0	0.517GPa
a	0.136mm

Using these values in the line contact analysis equal pressure and contact dimensions can be achieved. Figure 6.9, shows the equivalent isothermal line contact compared to the isothermal point contact result. The velocity used for the both cases shown is low at 0.01 ms^{-1} and the slide roll ratio is low at 0.1. It should be noted that there is very good agreement between the two results within the Hertzian contact region, but due to the difference in the relative radius of curvature, developed through the equivalencing, there is slight variation to be seen outside of the contact dimension. It can be seen that the differences are relatively small and thus it can be concluded that the equivalencing will give reasonably good agreement with experimental observations.

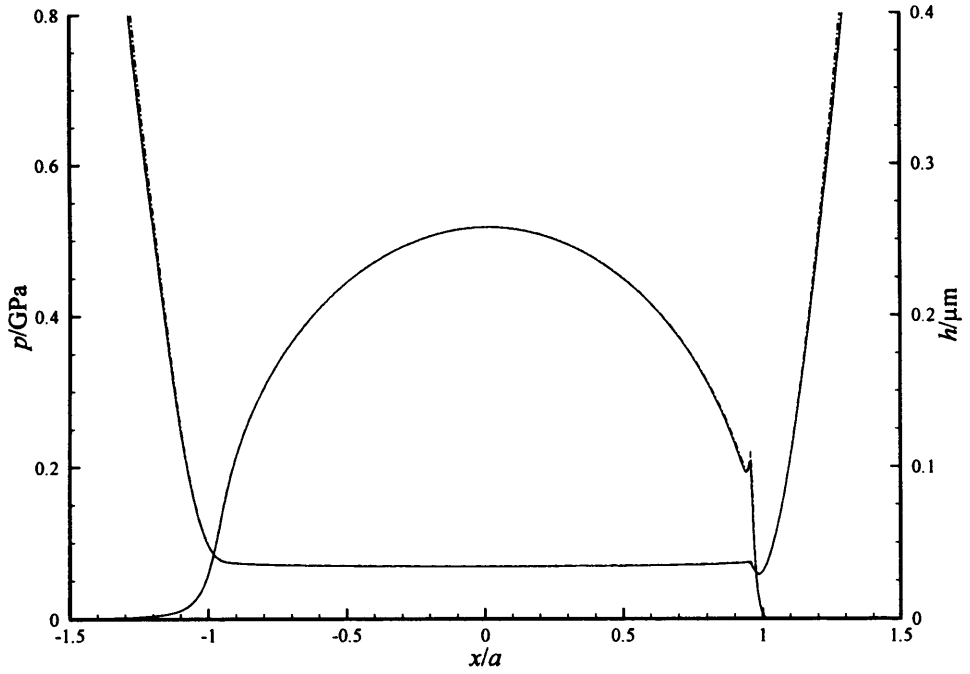

Figure 6.9 – Comparison of centre line point contact result (solid) with equivalent line contact result (dashed) at low speed

Using the conditions outlined in tables 6.1 and 6.2, the different rheological models can be investigated over a range of surface speeds to show the effects that the different models have over a range of conditions. The limiting shear stress and Eyring shear stress are oil properties taken from Evans and Johnson (1986). Both these non-Newtonian parameters were taken to vary with temperature and pressure.

6.3.1 Preliminary Results

For the initial cases considered, the slide roll ratio, ζ, and mean entrainment velocity, $\bar{u}$, were kept constant. The operating conditions were fixed to allow comparisons between the different models under the operating conditions. Later sections will investigate how the models respond to different entrainment speeds and slide roll ratios.

Results for the Eyring rheological model are shown in Figure 6.10. The response does not show typical Newtonian features, such as a sharp spike at the constriction, instead there is a levelling off of the pressure at the constriction. It can be seen that there is no noticeable difference between the surface temperatures and the mid film temperature, this indicates that all of the heat generated within the film is conducted to the surfaces. Since the film thickness is small, it is very difficult to generate a temperature difference between the surfaces. For thin films conduction across the film is relatively easily achieved. Thus the difference between the surface and mid-film temperatures remains small as the heat generated is easily conducted away.

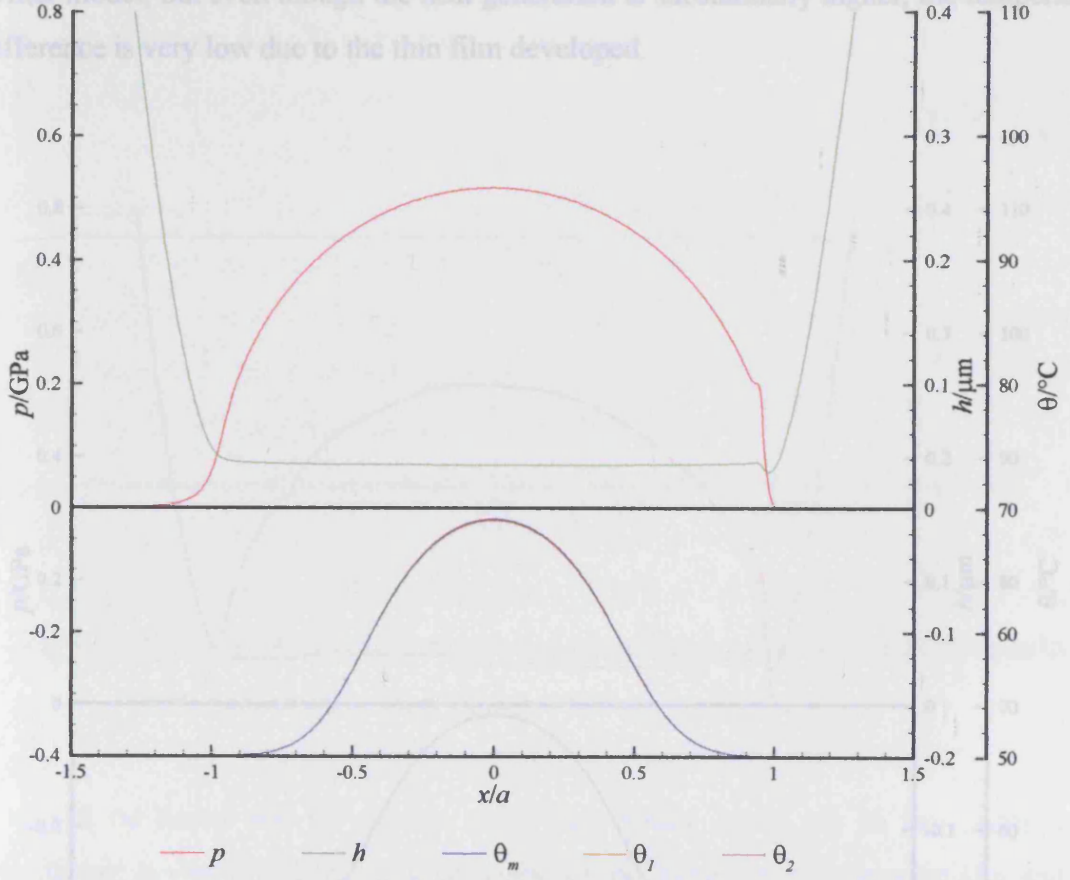

Figure 6.10 – Film thickness, pressure and temperature response using the Eyring rheological model for $\bar{u} = 0.01\text{ms}^{-1}$ and $\zeta = 0.50$

Considering the response for the Bair and Winer model for $\bar{u} = 0.01\text{ms}^{-1}$ shown in Figure 6.11, there are different features compared to figure 6.10. The film thickness for the Bair

and Winer model is not flat through the centre of the contact. It can be seen that there is a 's' shaped feature, this was seen in the isothermal results in Chapter 5. It can be seen that the pressure also deviates from the Eyring response at several locations. There is a kink formed in the pressure at the location where the film thickness begins to increase, there are also other changes of slope in the pressure when the film thickness kinks begin and end. There is a very slight difference between the temperatures developed, where the Bair and Winer model generates higher temperatures. The difference between the two results in less than a degree and results from an increase in the amount of heat generated within the film for the Bair and Winer result. The friction force for the Eyring model is 2.46N, however, the Bair and Winer model produces a much greater friction force of 9.66N. This clearly suggests that more heat is produced within the film for the Bair and Winer model, but even though the heat generation is substantially higher, the temperature difference is very low due to the thin film developed.

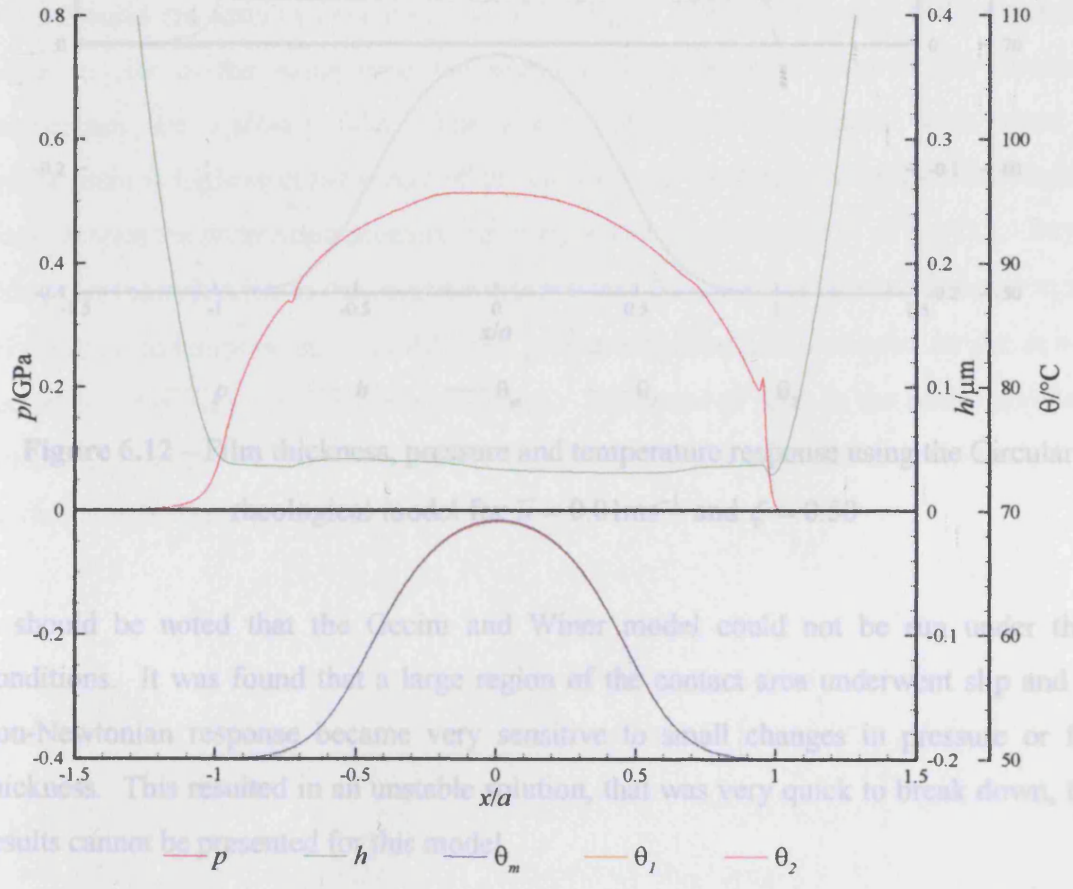

Figure 6.11 – Film thickness, pressure and temperature response using the Bair and Winer rheological model for $\overline{u} = 0.01 \text{ms}^{-1}$ and $\zeta = 0.50$

Comparing the above results with the Circular model shown in Figure 6.12, it can be seen that the less aggressive Circular model gives the same trends as the Eyring model. There is a slight reduction in the film thickness around $x/a = 0.5$, but there is no noticeable effects on pressure or any other features present in the film thickness.

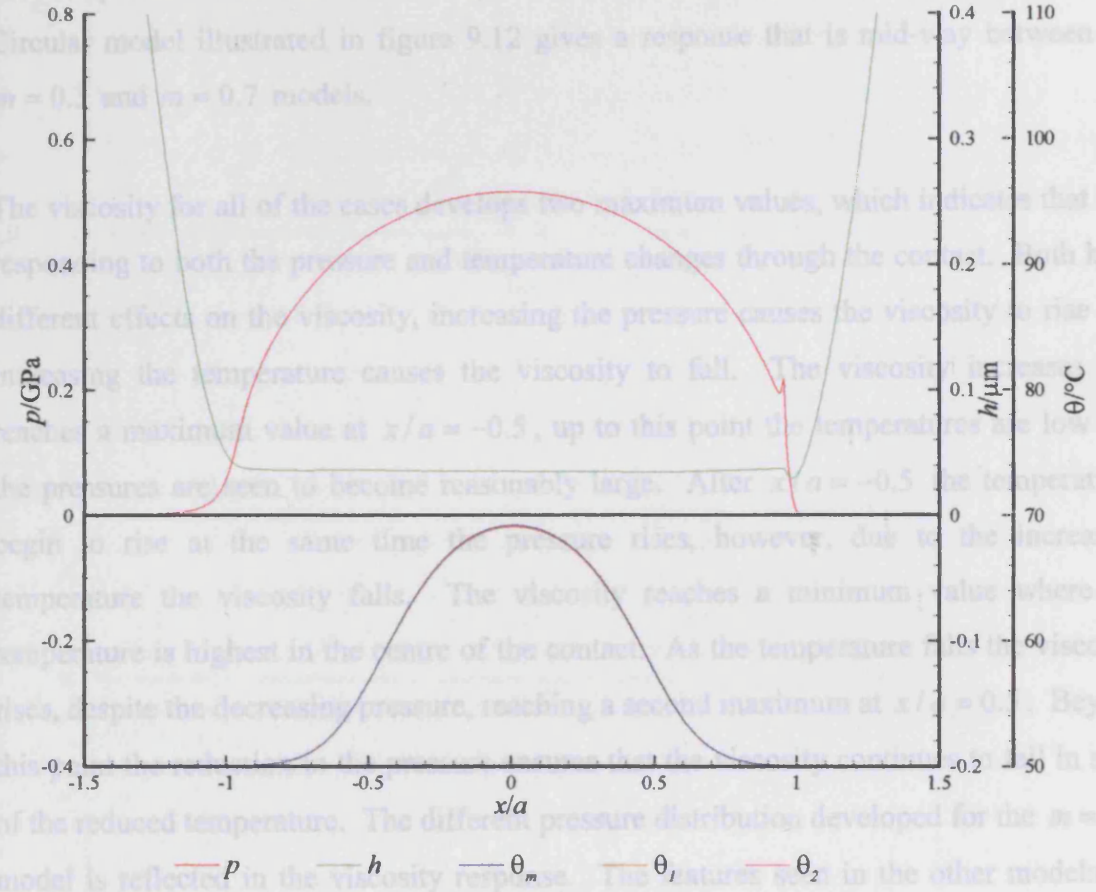

Figure 6.12 – Film thickness, pressure and temperature response using the Circular rheological model for $\bar{u} = 0.01\text{ms}^{-1}$ and $\zeta = 0.50$

It should be noted that the Gecim and Winer model could not be run under these conditions. It was found that a large region of the contact area underwent slip and the non-Newtonian response became very sensitive to small changes in pressure or film thickness. This resulted in an unstable solution, that was very quick to break down, thus results cannot be presented for this model.

Figure 6.13 compares the results for the Eyring, $m = 0.3$ and $m = 0.7$ models. It can be seen that the $m = 0.3$ model gives similar pressure and film thickness results to the Bair and Winer model. The $m = 0.7$ model does not show any features of the Bair and Winer model, the response is very similar to the Eyring model where there is no 's' shaped feature in the film thickness. Despite the increase in the scale, all the models produce the same temperature distribution, where there are only small differences between the mid-film temperature and the surface temperatures for any of the results. It is clear that the Circular model illustrated in figure 9.12 gives a response that is mid-way between the $m = 0.3$ and $m = 0.7$ models.

The viscosity for all of the cases develops two maximum values, which indicates that it is responding to both the pressure and temperature changes through the contact. Both have different effects on the viscosity, increasing the pressure causes the viscosity to rise and increasing the temperature causes the viscosity to fall. The viscosity increases and reaches a maximum value at $x/a = -0.5$, up to this point the temperatures are low and the pressures are seen to become reasonably large. After $x/a = -0.5$ the temperatures begin to rise at the same time the pressure rises, however, due to the increasing temperature the viscosity falls. The viscosity reaches a minimum value where the temperature is highest in the centre of the contact. As the temperature falls the viscosity rises, despite the decreasing pressure, reaching a second maximum at $x/a = 0.5$. Beyond this point the reduction in the pressure ensures that the viscosity continues to fall in spite of the reduced temperature. The different pressure distribution developed for the $m = 0.3$ model is reflected in the viscosity response. The features seen in the other models are retained, but larger values of viscosity are developed in the inlet and lower in the exit due to the changed pressure distribution.

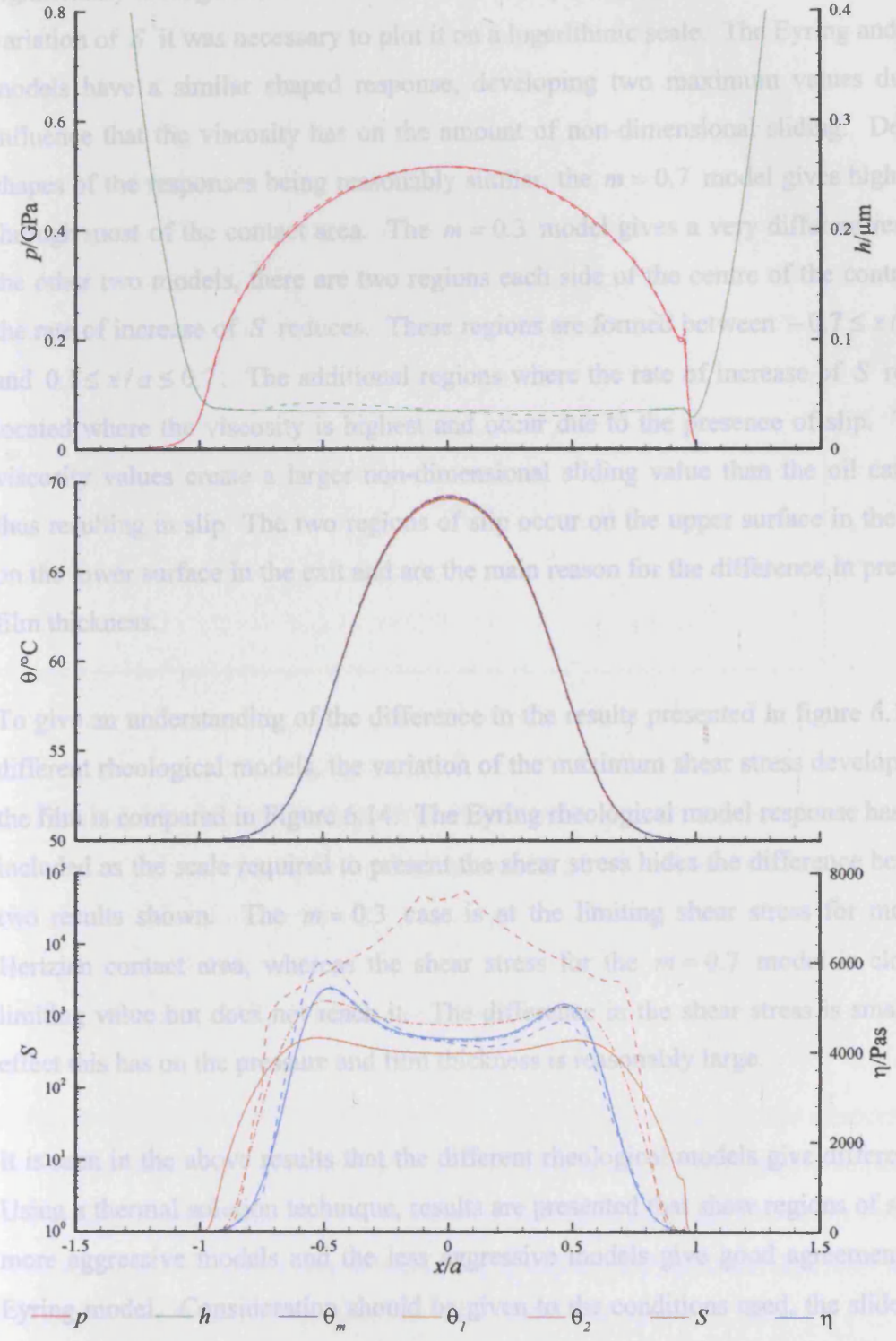

Figure 6.13 – Pressure, film thickness, temperature, S and η response for the Eyring (solid), $m = 0.3$ (dashed) and $m = 0.7$ (dash-dotted) rheological models for $\overline{u} = 0.01\text{ms}^{-1}$ and $\zeta = 0.50$

The correction to the viscosity due to the non-Newtonian behaviour, S, varies significantly through the contact. To allow comparison of the models and to show the variation of S it was necessary to plot it on a logarithmic scale. The Eyring and $m = 0.7$ models have a similar shaped response, developing two maximum values due to the influence that the viscosity has on the amount of non-dimensional sliding. Despite the shapes of the responses being reasonably similar, the $m = 0.7$ model gives higher values through most of the contact area. The $m = 0.3$ model gives a very different response to the other two models, there are two regions each side of the centre of the contact where the rate of increase of S reduces. These regions are formed between $-0.7 \le x/a \le -0.2$ and $0.1 \le x/a \le 0.7$. The additional regions where the rate of increase of S reduce are located where the viscosity is highest and occur due to the presence of slip. The large viscosity values create a larger non-dimensional sliding value than the oil can sustain, thus resulting in slip The two regions of slip occur on the upper surface in the inlet and on the lower surface in the exit and are the main reason for the difference in pressure and film thickness.

To give an understanding of the difference in the results presented in figure 6.13 for the different rheological models, the variation of the maximum shear stress developed within the film is compared in Figure 6.14. The Eyring rheological model response has not been included as the scale required to present the shear stress hides the difference between the two results shown. The $m = 0.3$ case is at the limiting shear stress for most of the Hertzian contact area, whereas the shear stress for the $m = 0.7$ model is close to the limiting value but does not reach it. The difference in the shear stress is small, but the effect this has on the pressure and film thickness is reasonably large.

It is seen in the above results that the different rheological models give different results. Using a thermal solution technique, results are presented that show regions of slip for the more aggressive models and the less aggressive models give good agreement with the Eyring model. Consideration should be given to the conditions used, the slide roll ratio was modest, but the entrainment velocity was very small.

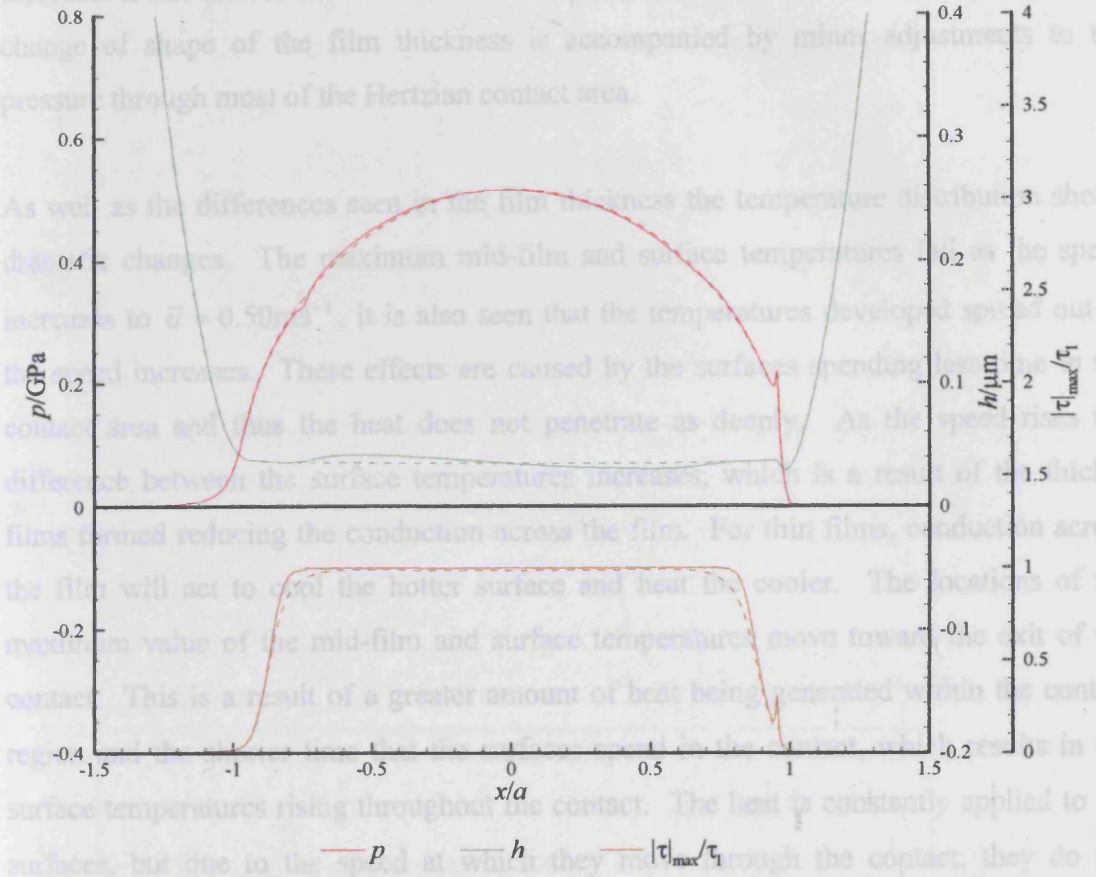

Figure 6.14 – Maximum shear stress response for the $m = 0.3$ (solid) and $m = 0.7$ (dashed) rheological models for $\bar{u} = 0.01 \text{ms}^{-1}$

6.3.2 Influence of Entrainment Speed

It was seen in Chapter 5 that increasing the velocity increases the amount of non-Newtonian behaviour. Therefore the effect that the increased entrainment velocity has on the results is considered, where it is expected that more limiting shear stress features will be seen. Results for increasing velocities of $\bar{u} = 0.01$, 0.10, 0.50 and 2.00 ms^{-1} are shown in Figure 6.15 for the Eyring model for a fixed slide roll ratio of 0.50.

It can be seen that the film thickness increases as the velocity increases, which is to be expected. The film thickness constriction formed at the exit of the contact moves further towards the centre of the contact as the speed rises. The film thickness formed through

the centre of the contact begins to change shape, for small velocities it is flat, as the speed increases it can be seen that the film thickness slopes down from the inlet to the exit. The change of shape of the film thickness is accompanied by minor adjustments to the pressure through most of the Hertzian contact area.

As well as the differences seen in the film thickness the temperature distribution shows dramatic changes. The maximum mid-film and surface temperatures fall as the speed increases to $\bar{u} = 0.50\text{ms}^{-1}$, it is also seen that the temperatures developed spread out as the speed increases. These effects are caused by the surfaces spending less time in the contact area and thus the heat does not penetrate as deeply. As the speed rises the difference between the surface temperatures increases, which is a result of the thicker films formed reducing the conduction across the film. For thin films, conduction across the film will act to cool the hotter surface and heat the cooler. The locations of the maximum value of the mid-film and surface temperatures move toward the exit of the contact. This is a result of a greater amount of heat being generated within the contact region and the shorter time that the surfaces spend in the contact, which results in the surface temperatures rising throughout the contact. The heat is constantly applied to the surfaces, but due to the speed at which they move through the contact, they do not respond quickly to the changes in the heat flux developed through the contact region. As the velocity is increased from $\bar{u} = 0.10\text{ms}^{-1}$ to $\bar{u} = 0.50\text{ms}^{-1}$ the faster surface temperature falls in general, whereas the slower surface and mid-film temperatures fall in the inlet and increase in the exit.

The temperatures developed within the contact generally continue to rise as the velocity increases further to $\bar{u} = 2.00\text{ms}^{-1}$, only the faster surface temperature is seen to fall in the inlet region. It can be seen that the mid-film temperature increases and that it also bears a resemblance to the pressure distribution to which it corresponds. A de-compressive cooling effect is seen for the higher entrainment velocities by a sudden drop in the mid-film temperature near the exit.

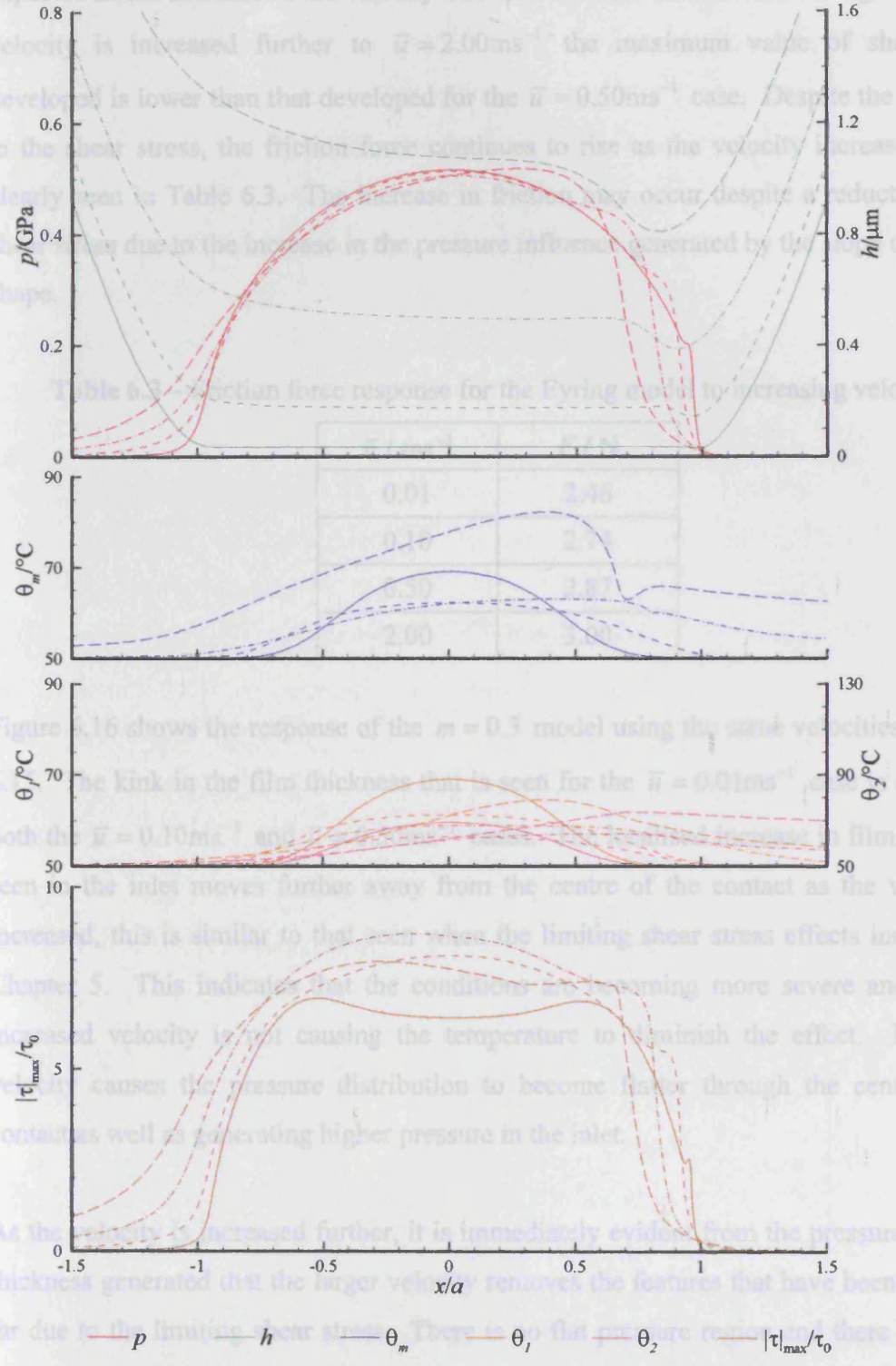

Figure 6.15 – Pressure, film thickness, temperature and shear stress response using the Eyring rheological model for $\bar{u} = 0.01\text{ms}^{-1}$ (solid), $\bar{u} = 0.10\text{ms}^{-1}$ (dashed), $\bar{u} = 0.50\text{ms}^{-1}$ (dash-dotted) and $\bar{u} = 2.00\text{ms}^{-1}$ (long dashed)

The shear stress is seen to increase as the velocity is increased to $\bar{u} = 0.50\text{ms}^{-1}$, this is expected as the increase in the velocity increases the non-dimensional sliding. When the velocity is increased further to $\bar{u} = 2.00\text{ms}^{-1}$ the maximum value of shear stress developed is lower than that developed for the $\bar{u} = 0.50\text{ms}^{-1}$ case. Despite the reduction in the shear stress, the friction force continues to rise as the velocity increases, this is clearly seen in Table 6.3. The increase in friction may occur despite a reduction in the shear stress due to the increase in the pressure influence generated by the slope of the film shape.

Table 6.3 – Friction force response for the Eyring model to increasing velocity

$\bar{u}$ / ms^{-1}	F / N
0.01	2.46
0.10	2.74
0.50	2.87
2.00	3.00

Figure 6.16 shows the response of the $m = 0.3$ model using the same velocities as figure 6.15. The kink in the film thickness that is seen for the $\bar{u} = 0.01\text{ms}^{-1}$ case is evident in both the $\bar{u} = 0.10\text{ms}^{-1}$ and $\bar{u} = 0.50\text{ms}^{-1}$ cases. The localised increase in film thickness seen in the inlet moves further away from the centre of the contact as the velocity is increased, this is similar to that seen when the limiting shear stress effects increased in Chapter 5. This indicates that the conditions are becoming more severe and that the increased velocity is not causing the temperature to diminish the effect. Increasing velocity causes the pressure distribution to become flatter through the centre of the contact as well as generating higher pressure in the inlet.

As the velocity is increased further, it is immediately evident from the pressure and film thickness generated that the larger velocity removes the features that have been seen thus far due to the limiting shear stress. There is no flat pressure region and there is also no indication of a kink in the film thickness. In fact, the film thickness is almost flat for most of the Hertzian contact region.

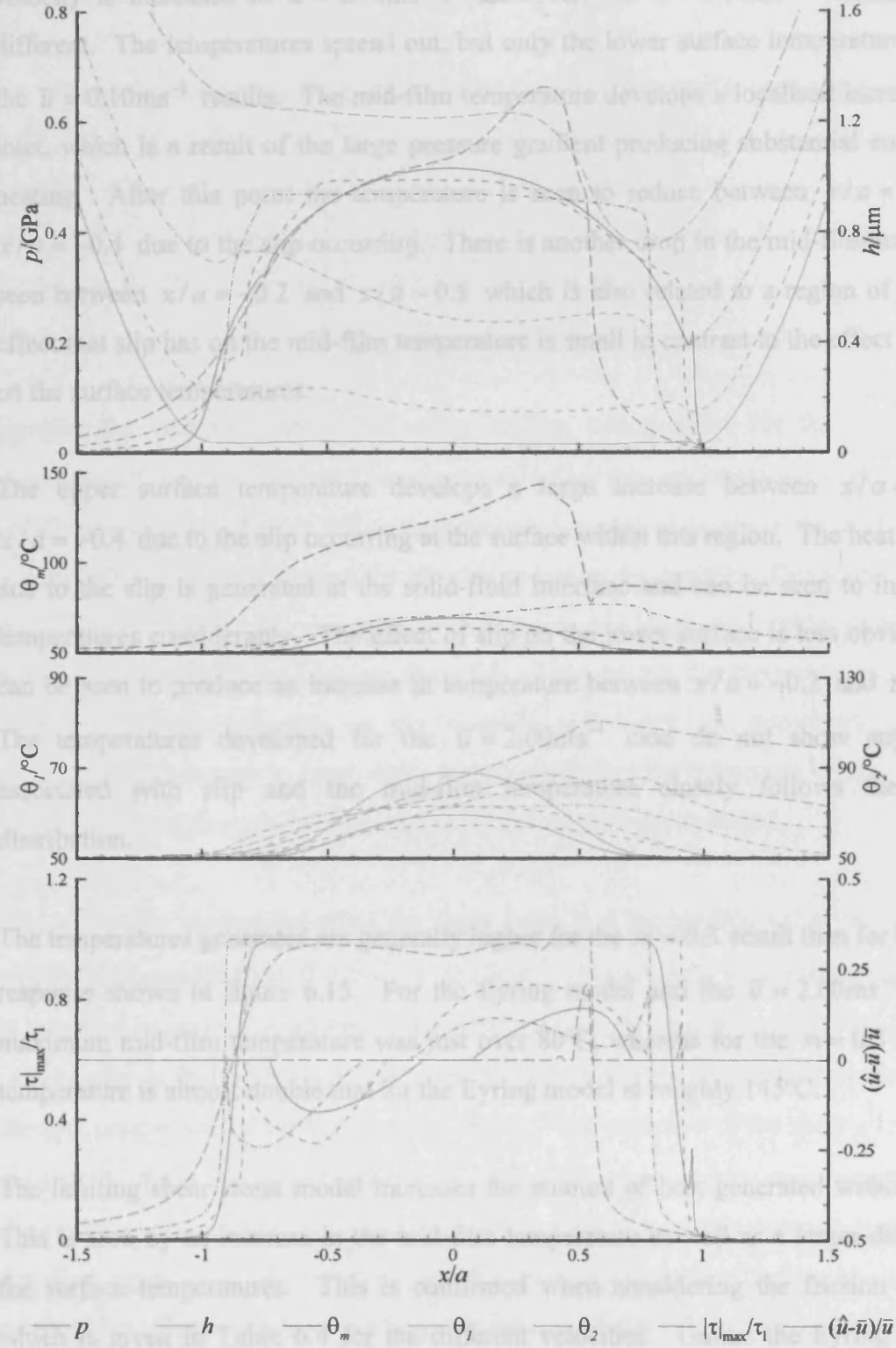

Figure 6.16 – Pressure, film thickness, temperature, shear stress and slip response using the $m = 0.3$ rheological model for $\bar{u} = 0.01\text{ms}^{-1}$ (solid), $\bar{u} = 0.10\text{ms}^{-1}$ (dashed), $\bar{u} = 0.50\text{ms}^{-1}$ (dash-dotted) and $\bar{u} = 2.00\text{ms}^{-1}$ (long dashed)

In a similar manner to the Eyring response, the maximum temperatures fall when the velocity is increased to $\bar{u} = 0.10\text{ms}^{-1}$. However, the $\bar{u} = 0.50\text{ms}^{-1}$ results are very different. The temperatures spread out, but only the lower surface temperature is below the $\bar{u} = 0.10\text{ms}^{-1}$ results. The mid-film temperature develops a localised increase in the inlet, which is a result of the large pressure gradient producing substantial compressive heating. After this point the temperature is seen to reduce between $x/a = -0.9$ and $x/a = -0.4$ due to the slip occurring. There is another drop in the mid-film temperature seen between $x/a = -0.2$ and $x/a = 0.5$ which is also related to a region of slip. The effect that slip has on the mid-film temperature is small in contrast to the effect that it has on the surface temperatures.

The upper surface temperature develops a large increase between $x/a = -0.9$ to $x/a = -0.4$ due to the slip occurring at the surface within this region. The heat generated due to the slip is generated at the solid-fluid interface and can be seen to increase the temperatures considerably. The effect of slip on the lower surface is less obvious, but it can be seen to produce an increase in temperature between $x/a = -0.2$ and $x/a = 0.5$. The temperatures developed for the $\bar{u} = 2.00\text{ms}^{-1}$ case do not show any features associated with slip and the mid-film temperature closely follows the pressure distribution.

The temperatures generated are generally higher for the $m = 0.3$ result than for the Eyring response shown in figure 6.15. For the Eyring model and the $\bar{u} = 2.00\text{ms}^{-1}$ case, the maximum mid-film temperature was just over 80°C, whereas for the $m = 0.3$ model the temperature is almost double that for the Eyring model at roughly 145°C.

The limiting shear stress model increases the amount of heat generated within the film. This is seen by an increase in the mid-film temperature as well as a larger difference in the surface temperatures. This is confirmed when considering the friction generated, which is given in Table 6.4 for the different velocities. Unlike the Eyring result, the friction force decreases as the viscosity increases. Even though the friction decreases, the friction force for the $m = 0.3$ model is almost three times the value of the Eyring result at $\bar{u} = 2.00\text{ms}^{-1}$. The difference in the friction force indicates that much more shear heating is developed for the $m = 0.3$ model than the Eyring model. This is probably the main

reason for the differences seen in the mid-film temperature values developed at higher speeds, especially the $\bar{u} = 2.00\text{ms}^{-1}$ results.

Table 6.4 – Friction force response for the $m = 0.3$ model to increasing speed

$\bar{u}$ / ms^{-1}	F / N
0.01	9.77
0.10	9.74
0.50	9.56
2.00	8.14

Despite the increased amount of shear heating taking place for the $\bar{u} = 0.01\text{ms}^{-1}$, $\bar{u} = 0.10\text{ms}^{-1}$ and $\bar{u} = 0.50\text{ms}^{-1}$ cases, the mid-film temperature is not higher than both solid's temperature. This is due to the heat being generated nearer to the surfaces, this is a result of the shear heating and the slip that occurs at the surfaces. Due to the thin films, the heat generated is still conducted into the solids and large temperature differences are not necessary to drive the heat into the solids. This ensures that the temperature does not have a noticeable influence on the viscosity and hence the non-Newtonian effects dominate. This is seen in more detail when considering the maximum value of the shear stress developed through the contact and the amount of slip produced.

The area of the contact for which the shear stress is at the limiting value increases as the velocity is increased from 0.01ms^{-1} to 0.10ms^{-1}. In contrast, as the entrainment velocity is increased again to $\bar{u} = 0.50\text{ms}^{-1}$, it can be seen that the area over which the shear stress is at or near to its limiting value reduces. Although the area reduces, it can be seen that the area progresses further into the inlet region. The reduction of the shear stress towards the exit for the $\bar{u} = 0.50\text{ms}^{-1}$ case is caused by the pressure drop moving closer toward the centre of the contact. This reduces the viscosity within this area which ensures that the shear stress within the oil is not near the limiting value. The limiting shear stress and temperature effects lead to a very large flat pressure distribution. Due to the magnitude of the pressure generated in this region, the area over which this pressure is generated must reduce in order to carry the correct load.

The maximum temperatures reached reduce as the velocity rises, allowing the viscosity to increase. Combining this effect with the increase in the entrainment velocity, the non-dimensional sliding imposed on the oil is seen to increase. Thus increasing the entrainment velocity produces a greater region that is at the limiting shear stress as well as producing greater regions of slip. As the entrainment velocity is increased the maximum non-dimensional slip is also seen to increase. The region of slip generated in the inlet, which reduces the entrainment mechanism, moves further into the inlet as the entrainment velocity is increased. Higher speeds produce greater amounts of slip which weaken the oil entrainment further. However, as the velocity is increased the region over which the slip in the inlet occurs does not appear to increase. The slip that occurs in the exit part of the contact, allowing the oil to flow out of the contact quicker, is also seen to move further towards the inlet as the velocity increases. For the $\bar{u} = 0.50\text{ms}^{-1}$ case, the slip on the lower surface is seen to change a significant amount, the magnitude of the non-dimensional slip almost halves for most of the region and is combined with a reduction in the area over which the slip occurs. For the $\bar{u} = 2.00\text{ms}^{-1}$ case there is only a small region of slip that occurs near the exit constriction.

Figure 6.17 shows the results for the $m = 0.7$ rheological model. It can be seen that features from the Eyring model (figure 6.15) and the $m = 0.3$ model (figure 6.16) are present. There is an 's' shaped kink in the film thickness, for the $\bar{u} = 0.10\text{ms}^{-1}$ and $\bar{u} = 0.50\text{ms}^{-1}$ cases. The 's' shaped kink is accompanied by an increase in pressure in the same region as the localised increase in the film thickness. A higher amplitude kink is seen for the $\bar{u} = 0.50\text{ms}^{-1}$ case, which also develops a pressure distribution that is much flatter through the centre of the contact. These features were seen for the $m = 0.3$ model and are associated with the shear stress reaching the limiting value. Despite the lack of these features for the $\bar{u} = 2.00\text{ms}^{-1}$ case, the results for the $m = 0.7$ case are more like the $m = 0.3$ result than the Eyring result. The Eyring result developed a slope in the film thickness, however the $m = 0.7$ response is more like the flat film thickness produced for the $m = 0.3$ model.

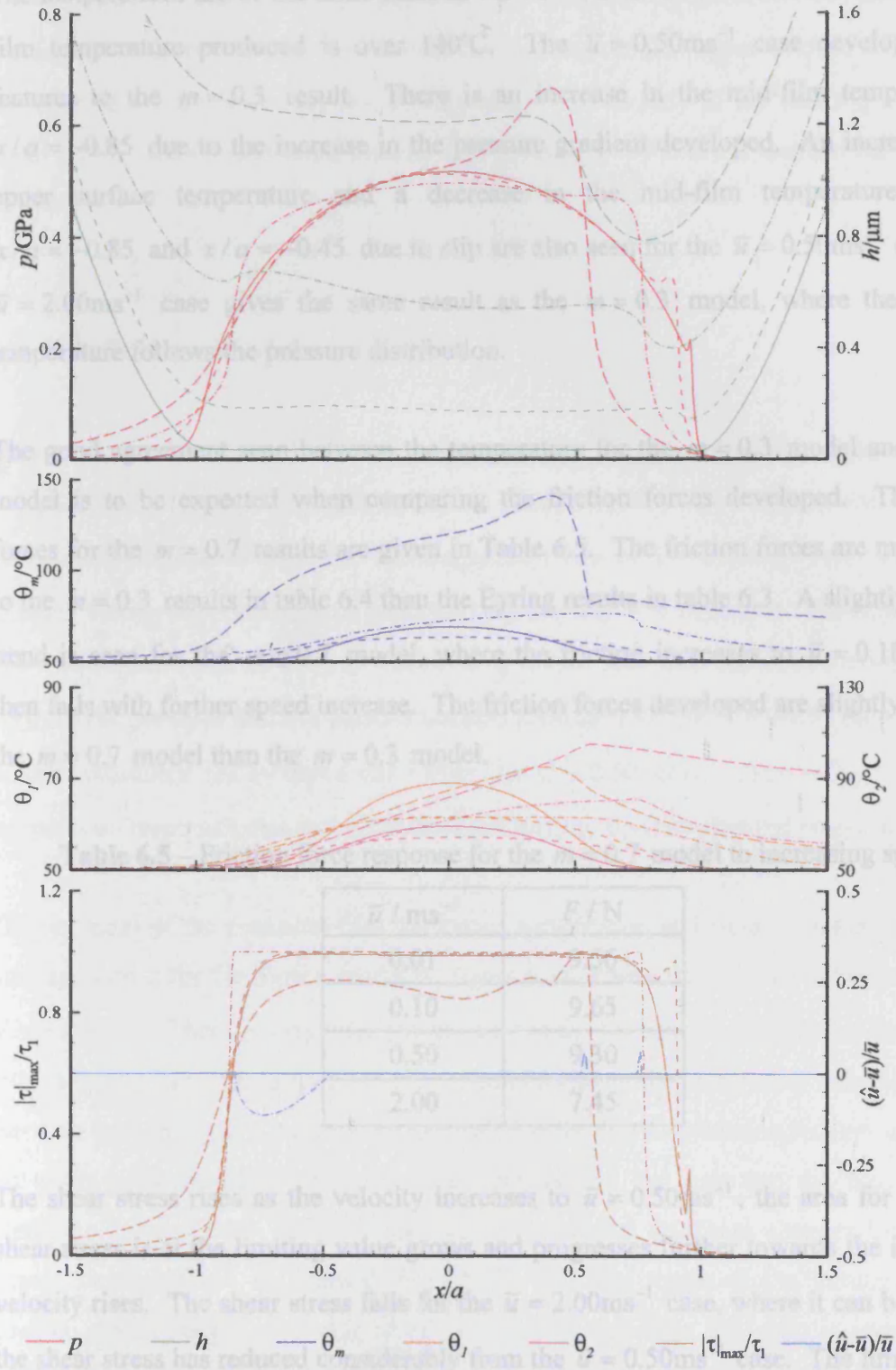

Figure 6.17 – Pressure, film thickness, temperature shear stress and slip response using the $m = 0.7$ rheological model for $\bar{u} = 0.01\text{ms}^{-1}$ (solid), $\bar{u} = 0.10\text{ms}^{-1}$ (dashed), $\bar{u} = 0.50\text{ms}^{-1}$ (dash-dotted) and $\bar{u} = 2.00\text{ms}^{-1}$ (long dashed)

The temperatures generated are also more like the $m = 0.3$ model than the Eyring model. The temperatures are of the same scale as the $m = 0.3$ model, where the maximum mid-film temperature produced is over 140°C. The $\bar{u} = 0.50\text{ms}^{-1}$ case develops similar features to the $m = 0.3$ result. There is an increase in the mid-film temperature at $x/a = -0.85$ due to the increase in the pressure gradient developed. An increase in the upper surface temperature and a decrease in the mid-film temperature between $x/a = -0.85$ and $x/a = -0.45$ due to slip are also seen for the $\bar{u} = 0.50\text{ms}^{-1}$ case. The $\bar{u} = 2.00\text{ms}^{-1}$ case gives the same result as the $m = 0.3$ model, where the mid-film temperature follows the pressure distribution.

The good agreement seen between the temperature for the $m = 0.3$ model and $m = 0.7$ model is to be expected when comparing the friction forces developed. The friction forces for the $m = 0.7$ results are given in Table 6.5. The friction forces are much closer to the $m = 0.3$ results in table 6.4 than the Eyring results in table 6.3. A slightly different trend is seen for the $m = 0.7$ model, where the friction increases to $\bar{u} = 0.10\text{ms}^{-1}$ and then falls with further speed increase. The friction forces developed are slightly lower for the $m = 0.7$ model than the $m = 0.3$ model.

Table 6.5 – Friction force response for the $m = 0.7$ model to increasing speed

$\bar{u}$ / ms^{-1}	F / N
0.01	9.56
0.10	9.65
0.50	9.30
2.00	7.45

The shear stress rises as the velocity increases to $\bar{u} = 0.50\text{ms}^{-1}$, the area for which the shear stress is at the limiting value grows and progresses further towards the inlet as the velocity rises. The shear stress falls for the $\bar{u} = 2.00\text{ms}^{-1}$ case, where it can be seen that the shear stress has reduced considerably from the $\bar{u} = 0.50\text{ms}^{-1}$ case. The limiting shear stress is only attained for a very small region near the exit constriction at $x/a = 0.5$. This gives good agreement with the results seen for the $m = 0.3$ model.

Unlike the $m = 0.3$ model, the $m = 0.7$ model only produces a modest amount of slip for the $\bar{u} = 0.50\text{ms}^{-1}$ case. There is only one large region of slip generated for all the speeds considered, this occurs for the $\bar{u} = 0.50\text{ms}^{-1}$ case on the upper surface in the inlet. Although the shear stress for the $\bar{u} = 2.00\text{ms}^{-1}$ case is only at the limiting value for a small part of the contact area, there is slip generated at the lower surface for this small region.

6.3.3 Influence of Slide Roll Ratio

Extending the work that has been conducted thus far, film thickness patterns can be compared for different rheologies at various slide roll ratios. Two entrainment speeds are considered, one low to minimise the temperature effects, $\bar{u} = 0.10\text{ms}^{-1}$, and another that is larger where the temperatures generated are modest, $\bar{u} = 1.00\text{ms}^{-1}$. The range of slide roll ratios considered is from $\zeta = 0.25$ to $\zeta = 0.75$ and thus includes the value of 0.5 used for the previous results. Results are also presented for the $\zeta = -0.50$ case where the surface velocities are swapped over from the $\zeta = 0.50$ case. This will give different temperature responses due to the two surfaces having different thermal properties.

The variation of the pressure, film thickness, temperature and shear stress with slide roll ratio are shown for the Eyring model in Figure 6.18, where the velocity is held constant at $\bar{u} = 0.10\text{ms}^{-1}$. There is very little change seen in the pressure distribution as the slide roll ratio is varied, however, slight differences can be seen at the entrance to and exit from the Hertzian contact area. The pressure generated at the exit constriction reduces as the slide roll ratio is increased, giving rise to a difference in the way that the film thickness minimum is generated. At low slide roll ratios the film thickness changes from the flat almost parallel shape maintained through the Hertzian region to the minimum feature quickly, developing a relatively sharp corner. As the slide roll ratio is increased this transition becomes much smoother. Also seen is the change of slope of the film thickness through the whole contact region. When $\zeta = 0.25$ the film thickness is almost flat, but as the slide roll ratio increases there is a downward slope generated form inlet to exit that increases with slide roll ratio.

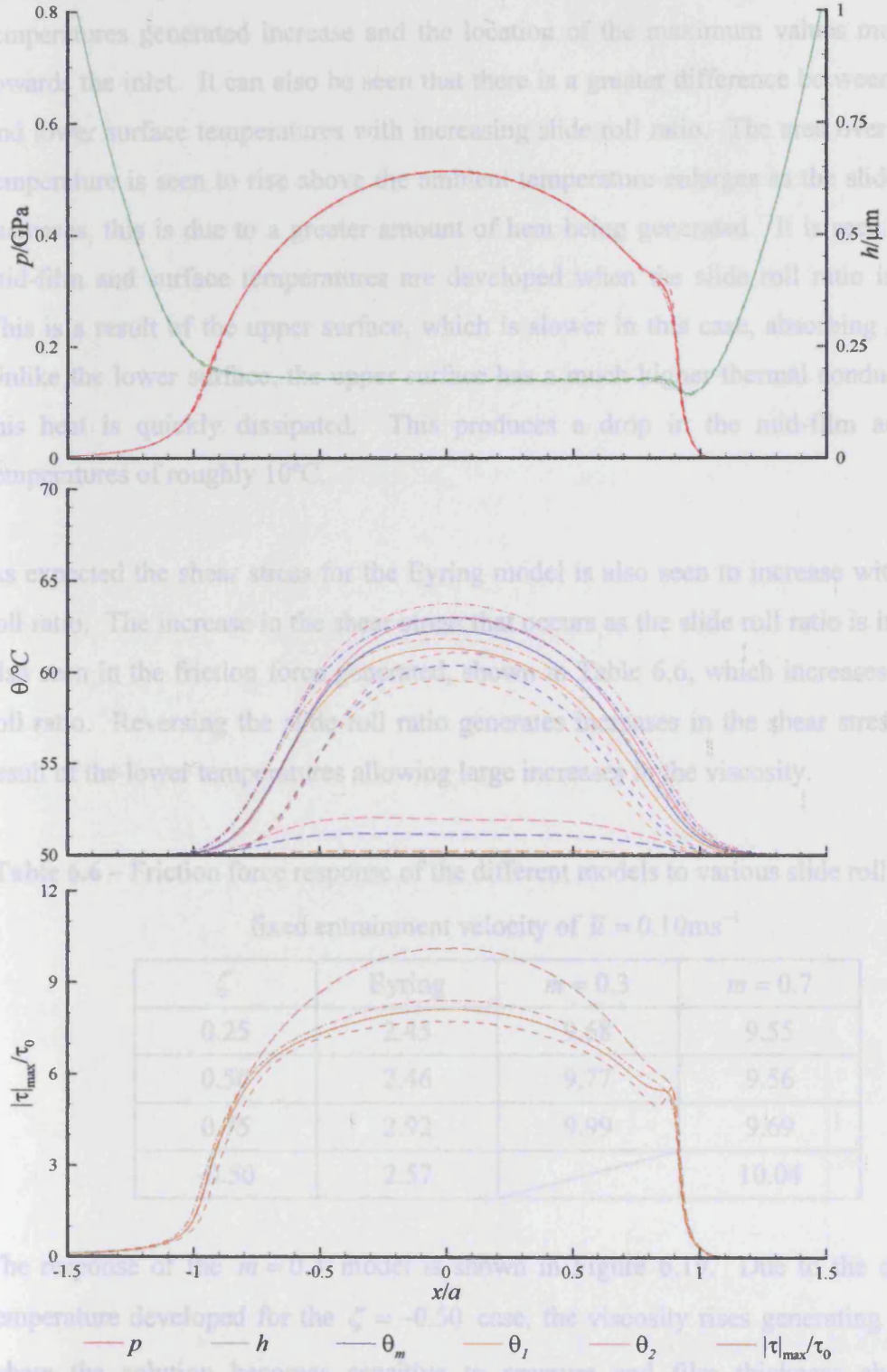

Figure 6.18 – Pressure, film thickness, temperature and shear stress response using the Eyring rheological model for fixed entrainment speed of $\bar{u} = 0.10\text{ms}^{-1}$ and slide roll ratios of $\zeta = 0.50$ (solid), $\zeta = 0.25$ (dashed), $\zeta = 0.75$ (dash-dotted) and $\zeta = -0.50$ (long dashed)

The changes in temperatures observed are more apparent than those seen in the film thickness and pressure. As the slide roll ratio increases from 0.25 to 0.75 all the temperatures generated increase and the location of the maximum values move further towards the inlet. It can also be seen that there is a greater difference between the upper and lower surface temperatures with increasing slide roll ratio. The area over which the temperature is seen to rise above the ambient temperature enlarges as the slide roll ratio increases, this is due to a greater amount of heat being generated. It is seen that lower mid-film and surface temperatures are developed when the slide roll ratio is reversed. This is a result of the upper surface, which is slower in this case, absorbing more heat. Unlike the lower surface, the upper surface has a much higher thermal conductivity and this heat is quickly dissipated. This produces a drop in the mid-film and surface temperatures of roughly 10°C.

As expected the shear stress for the Eyring model is also seen to increase with the slide roll ratio. The increase in the shear stress that occurs as the slide roll ratio is increased is also seen in the friction force generated, shown in Table 6.6, which increases with slide roll ratio. Reversing the slide roll ratio generates increases in the shear stress, this is a result of the lower temperatures allowing large increases in the viscosity.

Table 6.6 – Friction force response of the different models to various slide roll ratios at a fixed entrainment velocity of $\bar{u} = 0.10\text{ms}^{-1}$

ζ	Eyring	$m = 0.3$	$m = 0.7$
0.25	2.45	9.68	9.55
0.50	2.46	9.77	9.56
0.75	2.92	9.99	9.69
-0.50	2.57		10.04

The response of the $m = 0.3$ model is shown in Figure 6.19. Due to the drop in the temperature developed for the $\zeta = -0.50$ case, the viscosity rises generating a situation where the solution becomes sensitive to pressure and film thickness changes. A converged result could not be achieved for this case and cannot be included in the results in figure 6.19.

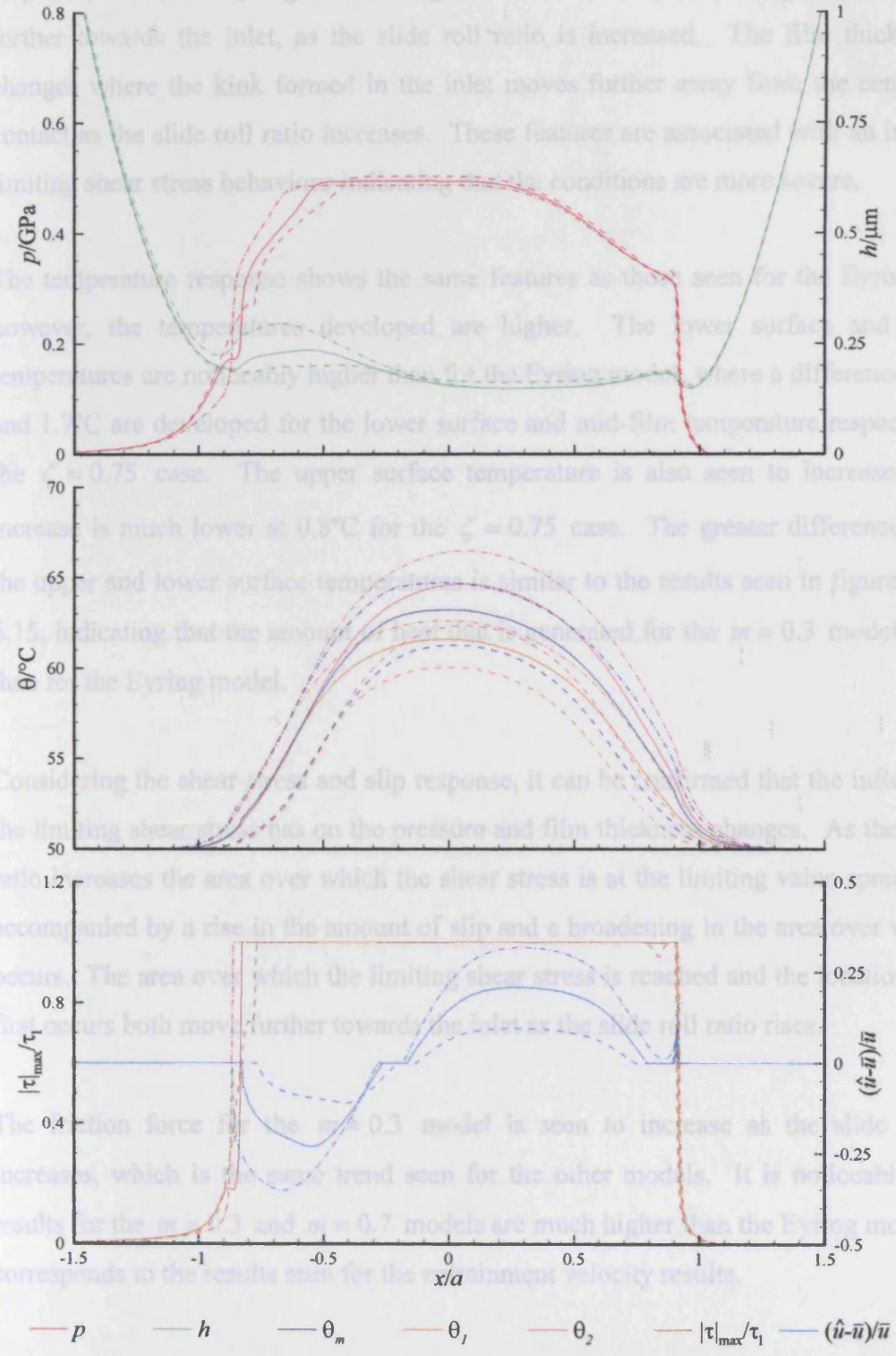

Figure 6.19 – Pressure, film thickness, temperature, shear stress and slip response using the $m = 0.3$ rheological model for fixed entrainment speed of $\bar{u} = 0.10\text{ms}^{-1}$ and slide roll ratios of $\zeta = 0.50$ (solid), $\zeta = 0.25$ (dashed) and $\zeta = 0.75$ (dash-dotted)

It is clear that the slide roll ratio has large effects on the film thickness and pressure response, unlike the Eyring model in figure 6.18. The flat pressure region grows, moving further towards the inlet, as the slide roll ratio is increased. The film thickness also changes where the kink formed in the inlet moves further away from the centre of the contact as the slide roll ratio increases. These features are associated with an increase in limiting shear stress behaviour indicating that the conditions are more severe.

The temperature response shows the same features as those seen for the Eyring model, however, the temperatures developed are higher. The lower surface and mid-film temperatures are noticeably higher than for the Eyring model, where a difference of 2.7°C and 1.7°C are developed for the lower surface and mid-film temperature respectively for the $\zeta = 0.75$ case. The upper surface temperature is also seen to increase, but this increase is much lower at 0.8°C for the $\zeta = 0.75$ case. The greater difference between the upper and lower surface temperatures is similar to the results seen in figure 6.16 and 6.15, indicating that the amount of heat that is generated for the $m = 0.3$ model is higher than for the Eyring model.

Considering the shear stress and slip response, it can be confirmed that the influence that the limiting shear stress has on the pressure and film thickness changes. As the slide roll ratio increases the area over which the shear stress is at the limiting value spreads and is accompanied by a rise in the amount of slip and a broadening in the area over which slip occurs. The area over which the limiting shear stress is reached and the location that slip first occurs both move further towards the inlet as the slide roll ratio rises.

The friction force for the $m = 0.3$ model is seen to increase as the slide roll ratio increases, which is the same trend seen for the other models. It is noticeable that the results for the $m = 0.3$ and $m = 0.7$ models are much higher than the Eyring model. This corresponds to the results seen for the entrainment velocity results.

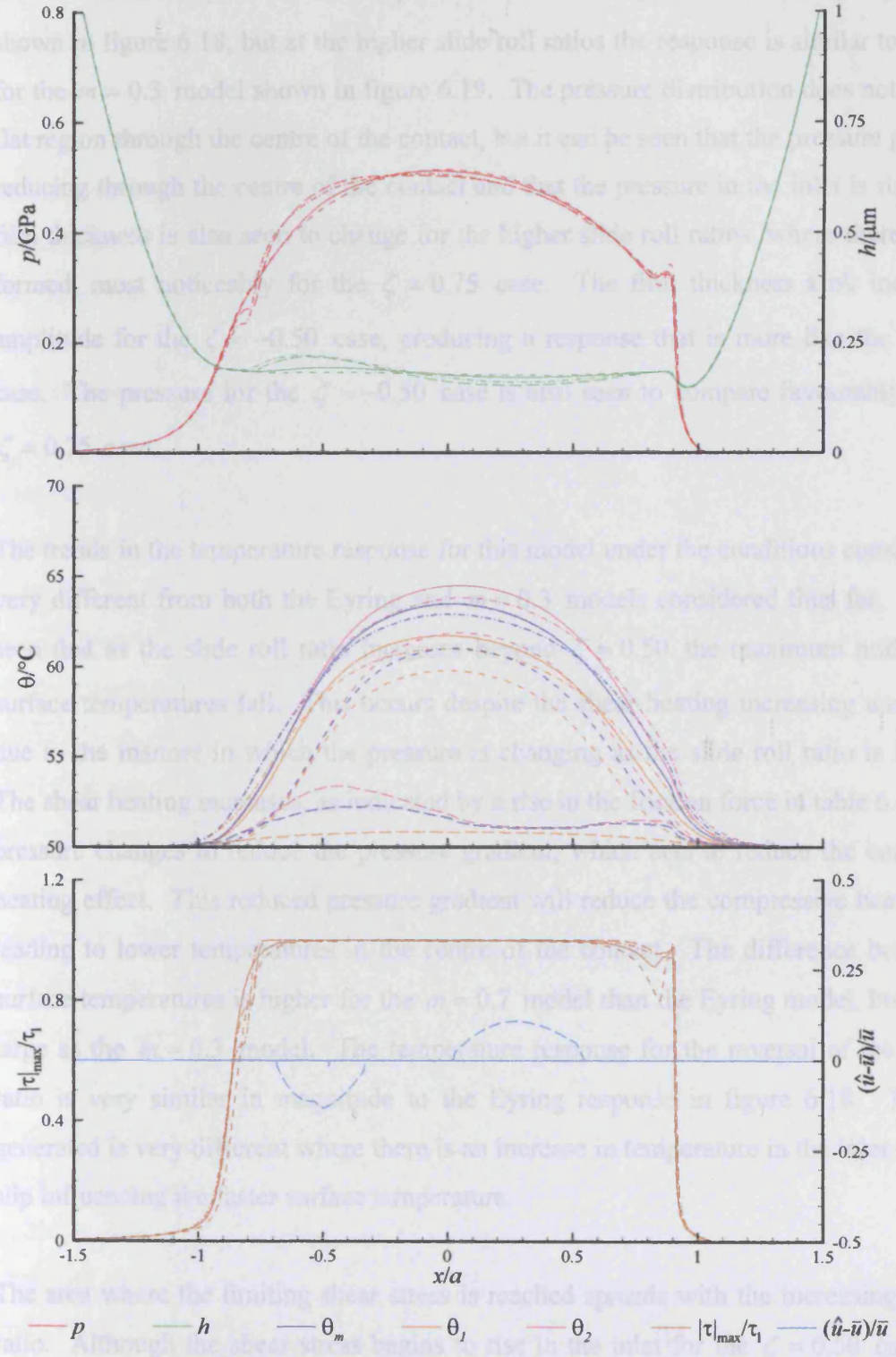

Figure 6.20 – Pressure, film thickness, temperature, shear stress and slip response using the $m = 0.7$ rheological model for fixed entrainment speed of $\bar{u} = 0.10\text{ms}^{-1}$ and slide roll ratios of $\zeta = 0.50$ (solid), $\zeta = 0.25$ (dashed), $\zeta = 0.75$ (dash-dotted) and $\zeta = -0.50$ (long dashed)

Figure 6.20 shows the response for the $m = 0.7$ model to different slide roll ratios. The film thickness and pressure for the $\zeta = 0.25$ case are very similar to the Eyring response, shown in figure 6.18, but at the higher slide roll ratios the response is similar to that seen for the $m = 0.3$ model shown in figure 6.19. The pressure distribution does not contain a flat region through the centre of the contact, but it can be seen that the pressure gradient is reducing through the centre of the contact and that the pressure in the inlet is rising. The film thickness is also seen to change for the higher slide roll ratios, where there is a kink formed, most noticeably for the $\zeta = 0.75$ case. The film thickness kink increases in amplitude for the $\zeta = -0.50$ case, producing a response that is more like the $\zeta = 0.75$ case. The pressure for the $\zeta = -0.50$ case is also seen to compare favourably with the $\zeta = 0.75$ case.

The trends in the temperature response for this model under the conditions considered are very different from both the Eyring and $m = 0.3$ models considered thus far. It can be seen that as the slide roll ratio increases beyond $\zeta = 0.50$ the maximum mid-film and surface temperatures fall. This occurs despite the shear heating increasing and must be due to the manner in which the pressure is changing as the slide roll ratio is increased. The shear heating increases, as indicated by a rise in the friction force in table 6.6, and the pressure changes to reduce the pressure gradient, which acts to reduce the compressive heating effect. This reduced pressure gradient will reduce the compressive heating term, leading to lower temperatures in the centre of the contact. The difference between the surface temperatures is higher for the $m = 0.7$ model than the Eyring model, but is not as large as the $m = 0.3$ model. The temperature response for the reversal of the slide roll ratio is very similar in magnitude to the Eyring response in figure 6.18. The shape generated is very different where there is an increase in temperature in the inlet due to the slip influencing the faster surface temperature.

The area where the limiting shear stress is reached spreads with the increasing slide roll ratio. Although the shear stress begins to rise in the inlet for the $\zeta = 0.50$ case sooner than for the $\zeta = 0.75$ case, it can be seen that the $\zeta = 0.75$ case reaches the limiting value before the $\zeta = 0.50$ case. Also, the shear stress response at the exit changes with slide roll ratio. At low slide roll ratios the shear stress drops away from the limiting value

before the exit constriction, this dip decreases as the slide roll ratio is increased. The shear stress is substantially higher for the $\zeta = -0.50$ case than the $\zeta = 0.50$ case, which was the same trend seen for the Eyring model.

The $\zeta = -0.50$ case produces two large regions of slip, one on the lower surface (faster in this case) in the inlet and another region on the upper surface in the exit. There is a small region of slip generated for the $\zeta = 0.75$ case, which occurs on the upper surface in the inlet region. Although slip occurs for some of the results, it can be seen in the results without slip that the limiting shear stress rheology still generates changes in the film thickness and pressure response.

Increasing the entrainment velocity to $\bar{u} = 1.00\text{ms}^{-1}$ and using the same slide roll ratios gives rise to different results. Consider the response of the Eyring model illustrated in Figure 6.21. The pressure and film thickness change as the slide roll ratio is increased, the differences are larger at this entrainment speed than they were for the lower entrainment speed, shown in figure 6.18. The film thickness slopes down more from inlet to the exit as the slide roll ratio is increased and it can also be seen that the minimum film thickness reduces as the slide roll ratio is increased. The pressure is seen to decrease near the exit constriction as the slide roll ratio increases, which is the same effect as that seen in figure 6.18. In addition to the decrease in the pressure near the exit, the pressure at the entrance to the contact is seen to rise as the slide roll ratio increases. Reversing the slide roll ratio does not produce a large change in the pressure and film thickness. There are only slight changes seen in the pressure and film thickness near the exit constriction.

The temperatures generated at the higher entrainment velocity are substantially higher and have a different shape compared to the temperatures developed for the lower entrainment velocity. The surface with the lower velocity experiences a greater temperature rise as the slide roll ratio increases and the faster (upper, steel) surface is seen to produce the lowest rise. The shape of the temperature profiles remain the same and it is noticeable in the mid-film temperature that a considerable amount of compressive cooling occurs at the beginning of the exit constriction. The difference between the surface temperatures increases as the slide roll ratio increases, but unlike the $\bar{u} = 0.10\text{ms}^{-1}$ case, the mid-film temperature does not remain between the surface temperatures.

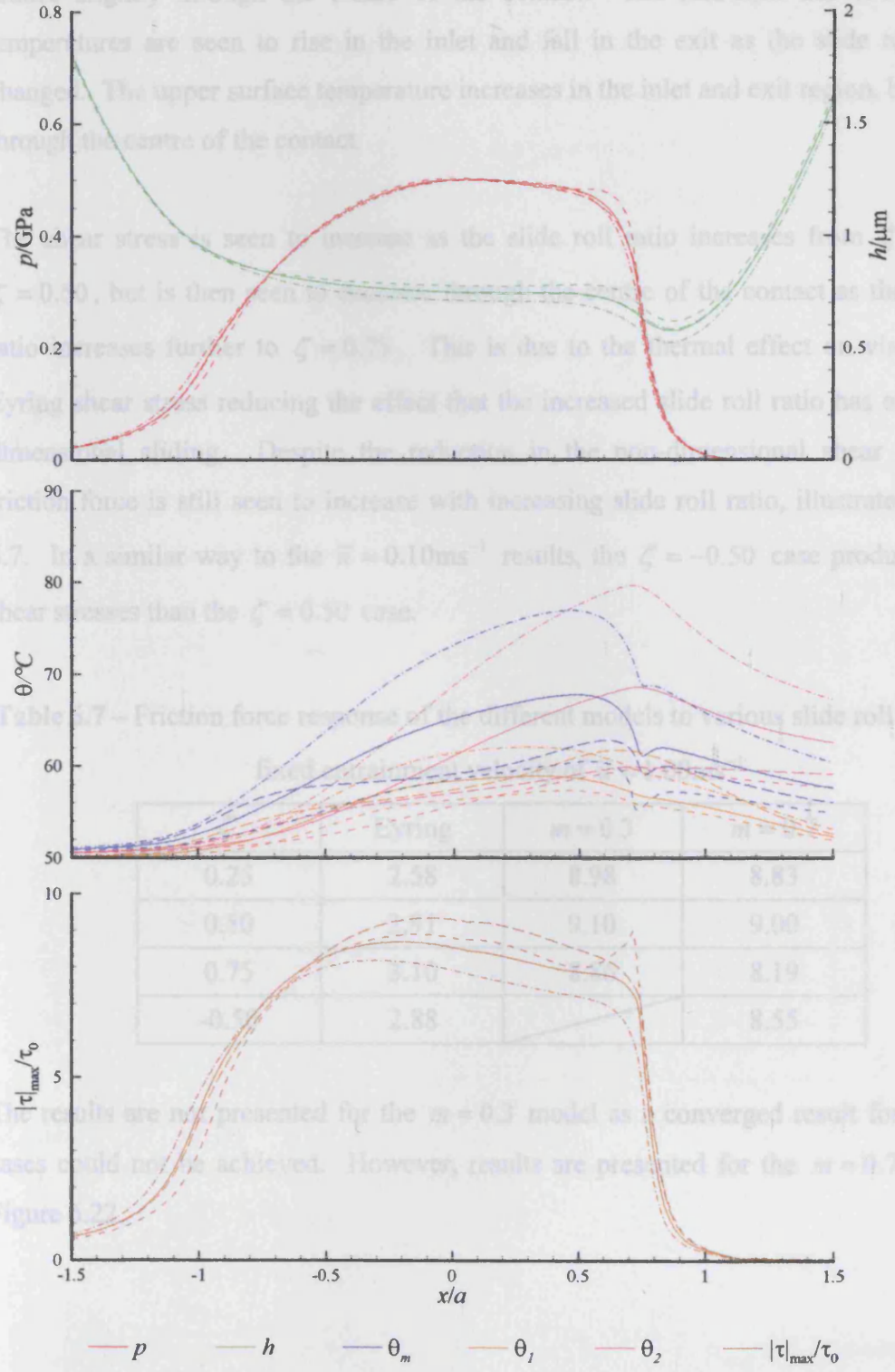

Figure 6.21 – Pressure, film thickness, temperature and shear stress response using the Eyring rheological model for fixed entrainment speed of $\bar{u} = 1.00\text{ms}^{-1}$ and slide roll ratios of $\zeta = 0.50$ (solid), $\zeta = 0.25$ (dashed), $\zeta = 0.75$ (dash-dotted) and $\zeta = -0.50$ (long dashed)

The temperature profiles change shape for the $\zeta = -0.50$ case, where all the temperatures reduce slightly through the centre of the contact. The mid-film and lower surface temperatures are seen to rise in the inlet and fall in the exit as the slide roll ratio is changed. The upper surface temperature increases in the inlet and exit region, but reduces through the centre of the contact.

The shear stress is seen to increase as the slide roll ratio increases from $\zeta = 0.25$ to $\zeta = 0.50$, but is then seen to decrease through the centre of the contact as the slide roll ratio increases further to $\zeta = 0.75$. This is due to the thermal effect on viscosity and Eyring shear stress reducing the effect that the increased slide roll ratio has on the non-dimensional sliding. Despite the reduction in the non-dimensional shear stress, the friction force is still seen to increase with increasing slide roll ratio, illustrated in Table 6.7. In a similar way to the $\bar{u} = 0.10\text{ms}^{-1}$ results, the $\zeta = -0.50$ case produces higher shear stresses than the $\zeta = 0.50$ case.

Table 6.7 – Friction force response of the different models to various slide roll ratios at a fixed entrainment velocity of $\bar{u} = 1.00\text{ms}^{-1}$

ζ	Eyring	$m = 0.3$	$m = 0.7$
0.25	2.58	8.98	8.83
0.50	2.91	9.10	9.00
0.75	3.10	8.86	8.19
-0.50	2.88		8.55

The results are not presented for the $m = 0.3$ model as a converged result for all of the cases could not be achieved. However, results are presented for the $m = 0.7$ model in Figure 6.22.

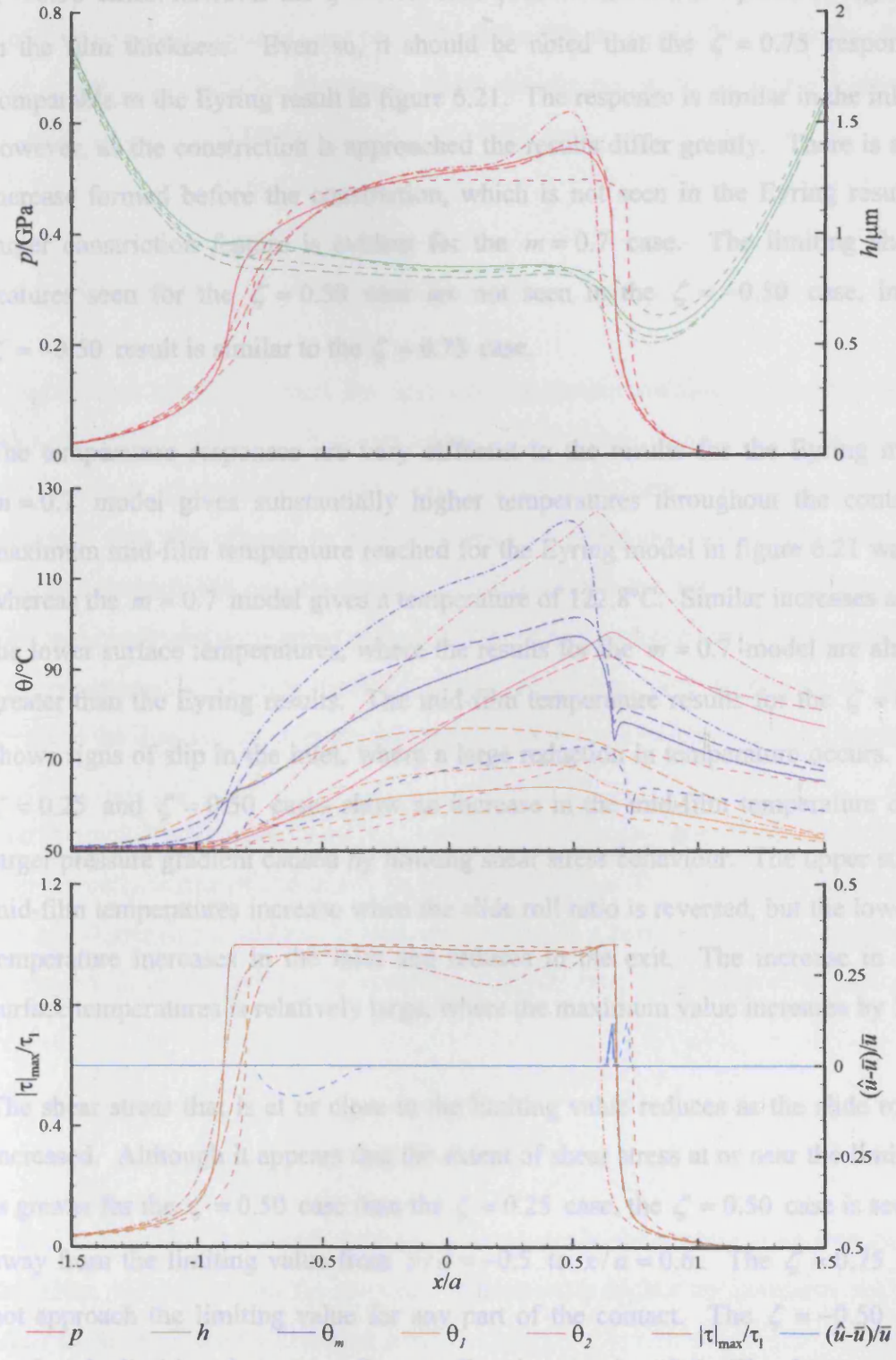

Figure 6.22 – Pressure, film thickness, temperature, shear stress and slip response using the $m = 0.7$ rheological model for fixed entrainment speed of $\bar{u} = 1.00\text{ms}^{-1}$ and slide roll ratios of $\zeta = 0.50$ (solid), $\zeta = 0.25$ (dashed), $\zeta = 0.75$ (dash-dotted) and $\zeta = -0.50$ (long dashed)

The pressure and film thickness show limiting shear stress features for the $\zeta = 0.25$ and $\zeta = 0.50$ cases, however the $\zeta = 0.75$ case does not show a flat pressure region or kink in the film thickness. Even so, it should be noted that the $\zeta = 0.75$ response is not comparable to the Eyring result in figure 6.21. The response is similar in the inlet region, however, as the constriction is approached the results differ greatly. There is a pressure increase formed before the constriction, which is not seen in the Eyring results, and a larger constriction feature is evident for the $m = 0.7$ case. The limiting shear stress features seen for the $\zeta = 0.50$ case are not seen in the $\zeta = -0.50$ case, instead the $\zeta = -0.50$ result is similar to the $\zeta = 0.75$ case.

The temperature responses are very different to the results for the Eyring model, the $m = 0.7$ model gives substantially higher temperatures throughout the contact. The maximum mid-film temperature reached for the Eyring model in figure 6.21 was 76.9°C, whereas the $m = 0.7$ model gives a temperature of 122.8°C. Similar increases are seen in the lower surface temperatures, where the results for the $m = 0.7$ model are almost 50% greater than the Eyring results. The mid-film temperature results for the $\zeta = 0.25$ case shows signs of slip in the inlet, where a large reduction in temperature occurs. Both the $\zeta = 0.25$ and $\zeta = 0.50$ cases show an increase in the mid-film temperature due to the larger pressure gradient caused by limiting shear stress behaviour. The upper surface and mid-film temperatures increase when the slide roll ratio is reversed, but the lower surface temperature increases in the inlet and reduces in the exit. The increase in the upper surface temperatures is relatively large, where the maximum value increases by 13°C

The shear stress that is at or close to the limiting value reduces as the slide roll ratio is increased. Although it appears that the extent of shear stress at or near the limiting value is greater for the $\zeta = 0.50$ case than the $\zeta = 0.25$ case, the $\zeta = 0.50$ case is seen to drop away from the limiting value from $x/a = -0.5$ to $x/a = 0.6$. The $\zeta = 0.75$ case does not approach the limiting value for any part of the contact. The $\zeta = -0.50$ case only reaches the limiting shear stress for a small region at $x/a = 0.6$. The shear stress is lower for the $\zeta = -0.50$ than the $\zeta = 0.50$ case.

Slip is generated for the $\zeta = 0.25$ and $\zeta = 0.50$ cases, but only the $\zeta = 0.25$ case produces slip in the inlet to the contact. Both cases generate slip in the exit half of the contact, the magnitude and area over which the slip occurs in this region is identical for both cases. However, the location at which the slip occurs in the exit is further towards the inlet for the $\zeta = 0.50$ case than the $\zeta = 0.25$ case.

6.4 Summary

Results have been presented for five non-Newtonian rheological models of the sort proposed within the EHL literature where a full energy equation has been applied and solved as part of the solution technique. Initial investigations conducted in section 5.5 indicated that the change in viscosity with temperature significantly influenced the manner in which the non-Newtonian rheology effects the pressure and film thickness. This was confirmed by running one of the test cases considered in Chapter 5 using the full thermal approach. It was found that the temperature rise developed in the contact removed the limiting shear stress features that were presented throughout Chapter 5.

The thermal effects were found to dominate the viscosity generated within the contact. The viscosity is one of the major factors that influence the limiting shear stress through the non-dimensional sliding. It was determined in Chapter 5 that the maximum shear stress that is developed by the rheology is determined by the non-dimensional sliding, U_s:

$$U_s = \frac{\eta u_s}{\tau_l h} \tag{6.3}$$

It can be seen that the actual viscosity at the point in the contact is an important part of this term. Therefore, as the viscosity reduces with higher temperatures the maximum shear stress produced will become smaller.

This effect is clearly seen in section 6.1 where the limiting shear stress features that have been widely seen in Chapter 5 are no longer present. It was found that the temperature, as expected, dominated the viscosity and greatly reduced the maximum shear stress

generated. In an effort to promote the limiting shear stress effects seen with isothermal models, it was decided to adjust the oil properties to make the oil more viscous. In this it was hoped to increase the viscosity of the oil within the centre of the contact, to generate a greater maximum shear stress.

The viscosity at ambient pressure, η_0, was adjusted first, which produced the same trend as that presented for isothermal behaviour. Instead of producing greater shear stresses, the increases in η_0 actually produced lower shear stresses, as predicted by the isothermal results in section 5.4.1. This effect results from the thicker films formed due to the improved entrainment mechanism. Despite the temperature generated within the film and the solids changing, thicker films are formed that reduce the shear stress. The increase in the viscosity at ambient pressure does not increase the viscosity at the centre of the film to the same degree that the film thickness increases. The result of this is a lower ratio of viscosity over film thickness, it can be seen from equation (6.3) that this will reduce the non-dimensional sliding within the contact, hence creating the lower shear stresses seen.

Investigations into the viscosity pressure co-efficient, α, were then undertaken, and results were presented for values of $\alpha = 11.1\text{GPa}^{-1}$, $\alpha = 17.8\text{GPa}^{-1}$ and $\alpha = 22.2\text{GPa}^{-1}$. It was seen that increasing α generated much larger viscosities within the contact. The increase in the viscosity created thicker films, but the films were not thick enough to reduce the ratio of the viscosity over the film thickness. The progressive rise of α resulted in an increase in the maximum shear stress developed within the film. Despite the increase, the values reached were not large enough to promote any features associated with the shear stress approaching the limiting value.

It was seen that the shear stresses that were closest to the limiting values were generated where the pressure changed most rapidly at the edges of the Hertzian contact area. This is probably due to the manner in which the limiting shear stress was calculated, as the limiting shear stress was applied as a function of pressure only. Therefore, any temperature variations did not influence the value of the limiting shear stress used. As the temperatures increase quite substantially within the contact, it may well be the case that this would reduce the limiting shear stress value and lead to the shear stress approaching the limiting value.

A final point that should be made about the results for increasing α is that the temperatures that were developed within the contact approached 667.4°C. Even for $\alpha = 17.75\text{GPa}^{-1}$ it was seen that the temperatures reached 446.0°C. These temperatures are very large and the Roelands viscosity approach used is not valid for such high temperature increases. The oil properties would also change and the oil approaches its boiling point at the high temperatures reached, therefore, the operating condition used were not appropriate and other conditions had to be found.

To remove the temperature effects it is desirable to use as low an entrainment velocity as possible. With this in mind, operating conditions were chosen that represent optical interferometry approaches to enable the results to be investigated. This should allow the different models used to be compared to experimental results and some form of agreement established.

The optical measurements are made in point contacts and the model adopted is for line contact arrangements. To overcome this, a method of equivalencing the line contact to the point contact centre line was presented. The equivalencing ensures that the same Hertzian contact dimension is achieved as well as attaining the same maximum Hertzian pressure. This is accomplished by adjusting the load, w', and the relative radius of curvature, R, for the line contact case. A comparison was made under isothermal conditions, where it was seen that the equivalenced line contact approach gave very good agreement with the point contact results.

Initially, results were presented at a low entrainment speed, $\bar{u} = 0.10\text{ms}^{-1}$, and slide roll ratio of 0.5 for the Eyring, Bair and Winer, Circular, $m = 0.3$ and $m = 0.7$ models. The temperatures produced for all the models were low and there was no variation between the mid-film temperature and the upper and lower surface temperatures. It was seen that the Bair and Winer and $m = 0.3$ models produced 's' shaped kinks that were in line with the features seen in Chapter 5. The Circular and $m = 0.7$ models did not give rise to such features, instead giving responses that were very similar to the Eyring model.

Further results were presented, using the same slide roll ratio, for increases in velocity using the $m = 0.3$, $m = 0.7$ and Eyring models. As the velocity increases it was noticed

that differences occur between the two surface temperatures. This temperature difference was greater for the more aggressive $m = 0.3$ model than for the less aggressive $m = 0.7$ model and the Eyring model. For cases where $\bar{u} = 0.10\text{ms}^{-1}$ or less, the mid-film temperature for all of the models remains between the two surface temperatures due to the small films present. The small films allow heat to be conducted across the film and also do not necessitate large temperature gradients to drive the heat into the solids. For velocities below $\bar{u} = 0.10\text{ms}^{-1}$ the location of the maximum temperature for the mid-film and surface temperatures occurred at the centre of the contact. This is due to the increase in the heat generated in the film and the shorter period of time that the surfaces spend in the contact region.

As the velocity was increased up to $\bar{u} = 0.50\text{ms}^{-1}$, the area over which the shear stress was at the limiting value increased. The area that was at the limiting shear stress progressed further into the inlet region as the speed increased, resulting in the localised film thickness maximum moving further towards the inlet. At the highest velocity, however, the area that was at the limiting shear stress value was seen to reduce drastically.

The slip was non-dimensionalised by dividing by the entrainment velocity, which allowed the effect that the velocity increase has on the slip generated to be assessed. Up to $\bar{u} = 0.50\text{ms}^{-1}$, it was seen that the magnitude of the non-dimensional slip increased as the entrainment velocity increased. It was also seen that the region over which slip occurred moved further toward the inlet with the increasing velocity, in a similar manner to the shear stress. The movement of the slip generates greater film thickness maximums in the inlet. For the $\bar{u} = 0.50\text{ms}^{-1}$ case it was seen that the slip generated on the lower surface in the exit of the contact reduces in amplitude compared to the other cases. This is a result of the temperature increase reducing the viscosity, creating a lower non-dimensional sliding demand resulting in less slip.

As well as affecting the film thickness, shear stress and slip, the velocity increase also influenced the pressure generated. It was seen in Chapter 5 that the pressure distribution becomes flat when the limiting shear stress effect is large, the same feature was seen here as the entrainment velocity was increased. This feature continued, particularly for the

$m = 0.3$ model, until the pressure distribution was more of a rectangular shape than the normal parabolic shape seen for most cases using the Eyring model.

It is seen for $\bar{u} = 0.50\text{ms}^{-1}$ that the temperature distribution contains some interesting features. There are noticeable dips in the mid-film temperature that occur due to the presence of slip. It can also be seen within the regions experiencing slip that the surface temperature increases at the surface where slip occurs. This is due to the heating effect that slip generates, extra heat is generated at the fluid-surface interface leading to a rise in temperature. The feature is not seen at speeds below $\bar{u} = 0.50\text{ms}^{-1}$ as the slip velocity is small, which produces less heat. It is seen for the $\bar{u} = 2.00\text{ms}^{-1}$ case that the temperature distribution was different from the other cases in that the mid-film temperature was above both of the surface temperatures for more of the region.

To improve the understanding of the interaction between the limiting shear stress and the temperature the slide roll ratio was varied. Cases were presented for two velocities, one low ($\bar{u} = 0.10\text{ms}^{-1}$) to minimise the amount of heat generated and another that is larger ($\bar{u} = 1.00\text{ms}^{-1}$) creating more heat. The initial slide roll ratio that was used for the velocity variation cases ($\zeta = 0.50$) was increased and decreased by 50% and reversed. Reversing the slide roll ratio swaps the surface velocities over.

At the low entrainment velocity the pressure and film thickness for the Eyring model were very similar, but there was a downward slope in the film thickness that developed as ζ increased. For the more aggressive model the pressure and film thickness show limiting shear stress features, which become more prominent as the slide roll ratio is increased. For the less aggressive model, the response moved from something similar to the Eyring model to a response closer to the more aggressive limiting shear stress model. The pressure gradient was seen to decrease through the centre of the contact and the film thickness developed a kink as the slide roll ratio increased. For the Eyring and $m = 0.3$ models the temperatures were seen to rise as the slide roll ratio increased. However, for the $m = 0.7$ model the temperatures grew up to $\zeta = 0.50$ but were seen to fall for $\zeta = 0.75$ due to the reduction in the compressive heating caused by the fall in the pressure gradient. Reversing the slide roll ratio caused the temperature to drop considerably, allowing the viscosity to rise. This has a very large influence on the

limiting shear stress results, where more limiting shear stress effects were seen due to the increase in the viscosity.

At the higher entrainment speed the Eyring model gave similar features to that seen at the low entrainment speed, however, the features are more prominent. Converged results for the $m = 0.3$ model could not be achieved, but it was clearly seen from the $m = 0.7$ model the manner in which the limiting shear stress models behave. It was seen that as the slide roll ratio increased, the amount of limiting shear stress fell due to the rise in the temperature. When the slide roll ratio was increased to 0.75, the shear stress did not reach the limiting value for any portion of the contact region. It should be noted, that even though there were no limiting shear stress features for the $m = 0.7$ model at the high slide roll ratio, the response was not similar to the Eyring results. Reversing the slide roll ratio did not affect the Eyring results, but it did create large changes to the $m = 0.7$ response. Unlike low speeds, the reversal causes the temperature to rise substantially, which reduces the viscosity. The $m = 0.7$ model develops substantial temperature rises which create lower limiting shear stress effects than that seen for the $\zeta = 0.50$ case.

Considering the response of the limiting shear stress models to the increase in the entrainment velocity and slide roll ratio it can be seen that there is a region where the limiting shear stress effects dominate. It was seen that as the entrainment velocity increased above $\bar{u} = 0.50\text{ms}^{-1}$ that the limiting shear stress effects began to fall with further increases in the velocity. This goes against the manner expected based on kinematics and must be due to the temperature effect dominating the response. As the velocity was increased to $\bar{u} = 0.50\text{ms}^{-1}$, it was seen that the limiting shear stress behaviour increased agreeing with the expected response from the isothermal conditions. Even though modest temperature rises of 25°C were reached, the response still followed the isothermal results. At low entrainment speeds the shear stress was seen to increase with slide roll ratio, but at the higher entrainment speed the shear stress was seen to fall with an increase in slide roll ratio.

This response indicates that there is a regime within which the limiting shear stress effects dominate, then another regime within which the temperature effects dominate. This allows the features associated with limiting shear stress to be investigated experimentally

and the possibility of this change over to be seen. The results presented in section 6.1 and 6.2 also indicate that there are two regimes, where even large changes in the oil properties do not produce large changes in the limiting shear stress response. Instead the changes lead to very large increases in the temperature.

Since the limiting shear stress features can still be established under thermal conditions, then the effect that they have on rough surface results should be considered. It was noted by Tao et al (2003) that higher Eyring shear stresses developed higher, narrower pressure at the asperity tips. This is a result of a more Newtonian response which does not allow the oil to flow near the asperity tips, whereas greater non-Newtonian effects allow the oil to move around the asperity tips generating lower broader pressures. The effect of the reduction in the pressure is to reduce the sub-surface shear stress and reduce the wear and fatigue effects. Limiting shear stress will allow this feature to occur to an even greater degree which will also have a large influence on the wear, fatigue and sub-surface shear stress.

The trend seen in the mid-film temperature for the limiting shear stress results that produced a large pressure gradient in the inlet can be seen experimentally in the work of Ausherman et al. (1976). Figure 12 [Ausherman et al., 1976] for the $V_s = 2.54$ and $V_s = 1.387$ cases show a localised increase in the mid-film temperature in the inlet region. This feature was not discussed, but shows very good agreement to the features seen in figures 6.16, 6.17 and 6.20. The same feature can be seen to a lesser degree in the results of Yagi, Kyogoku and Nakahara (2004) in Fig. 10. The feature presented in their work is much lower than that seen in the results of Ausherman at al. and could be a spurious measurement. Despite this possibility, the strong correlation with the results of Ausherman at al. could be investigated further.

It should be noted that care must be taken when interpreting the results when the mid-film temperatures become high. When a large difference in temperature occurs through the film the equations applied may not be applicable. During the derivation of the equations used for the limiting shear stress models, it was assumed that the viscosity did not vary across the film. However, when the temperatures generated within the film become very large this assumption is no longer true. As the mid-film temperatures rise, it is clear that

the oil properties in terms of the viscosity and the limiting shear stress value will change. This could lead to the oil reaching the limiting value at a location within the film and slip would occur at this point rather than at the surface. Currently, the average temperature generated is used to determine the viscosity and limiting shear stress value that is used in the various models. This approach is deemed acceptable, but care must be taken when interpreting the results when the temperatures vary greatly across the film.

The effect that the temperature variation will have on the limiting shear stress is discussed. In the middle of the film the higher temperature will lower the viscosity and the limiting shear stress, the viscosity will reduce most which will reduce the severity at this location. However, near the surfaces the oil will be cooler which will increase the viscosity, which may lead to slip in these locations. Hence, limiting shear stress features may occur when the difference between the surface and mid-film temperatures became large, such as the conditions considered here where the thermal effects began to dominate the response. It is the author's opinion that this may allow more of the limiting shear stress features to be seen for longer, but the features will still reduce as higher temperatures are reached.

Another possible limitation of the approach used is that the solutions are achieved using line contact techniques and applied to a point contact problem. This simplification has two effects; firstly there will be heating terms that will increase the temperature due to the change of pressure in the y direction, perpendicular to the entrainment direction, as well as conduction in the y direction. Secondly, the change of pressure in the y direction will also have a component that will increase the shear stress within the film, which will cause the model to reach limiting behaviour before that seen for the line contact. It should also be noted that velocity effects in the y direction have not been considered here. The velocity in the y direction would only occur due to side leakage, which is felt to be small under these conditions. However, isothermal results for the Eyring model have been run for the line contact and point contact problems and good agreement seen. Thus the point contact problem needs further development to investigate limiting shear stress models under the conditions presented here for line contacts.

Chapter 7

Surface Contact Thermal Solution Results

7.0 Introduction

This chapter is concerned with developing and refining the approach that has been adopted to calculate the temperature variation within the solids for areas where direct surface contact occurs. The equations and solution technique were presented in sections 2.7 and 4.4 where the Crank-Nicholson scheme was used to evaluate the time dependent terms.

Different numerical methods that can be used to approximate the thermal equations are investigated and the most stable method determined. Also presented is a brief investigation into the effect that the coefficient of friction imposed at direct contact areas has on the surface temperatures produced in the contact for a sinusoidal roughness.

7.1 Thermal Solution Refinement

In an effort to try and promote contact the standard test case that was presented in Chapter 5 was adopted, this is the same case as that used by Patching (1994). There were two slight modifications made, the entraining velocity was reduced and two sinusoidal roughnesses introduced. The conditions imposed are summarised in the Table 7.1, where the oil properties correspond to MobilJet II with a bulk oil temperature of 100°C.

The EHL mesh and thermal mesh information is detailed in Appendix C.

Table 7.1 – Operating conditions for thermal contact test case

a / mm	0.3352
E' / GPa	227.3
R / mm	19.05
$\bar{u} / ms^{-1}$	7.00
w' / kNm^{-1}	526.6
α / GPa^{-1}	11.10
ζ	0.50
η_0 / Pas	0.0048

The two sinusoids that were used to generate contact are very aggressive, in that they have short wavelengths and large amplitudes to make it more difficult for them to be deflected. The two sinusoids are damped over the first two wavelengths, this arrangement is used so that the preceding points can be used as some comparison for the contacting points. Due to the speed difference of the surfaces, the points on the surfaces that form the start of the sinusoidal roughnesses will move away from one another. This acts to enhance the damping effect, consider the fast surface (upper surface) origin being further towards the centre of the contact than the slower surface (lower surface). The upper surface sinusoid will initially interact with a smooth surface until the point of origin of the lower surface sinusoid. These progressively larger sinusoids are used as they make the conditions in the inlet increasingly more severe. This approach was adopted as contact primarily occurs in the inlet region [Holmes, Evans and Snidle, 2004] thus the increasingly aggressive sinusoids will first contact in the inlet region.

The sinusoids imposed on the surfaces are not identical, the sinusoid on the upper surface increases at a larger rate and has a larger final amplitude than the lower surface sinusoid. The same function is used to generate both sinusoids and the larger rate of increase is achieved on the upper surface due to the larger amplitude used. The function used to generate the sinusoidal features on both surfaces is:

$$\phi = amp \times damp \times \cos\left(2\pi \frac{xtime}{period}\right)$$

amp is the amplitude of the sinusoid, *xtime* is the distance from the current point to the point of origin of the function (it is positive in the entrainment direction) and *period* is the wavelength of the sinusoid.

The damping function, *damp*, used to reduce the initial amplitude of the sinusoids, is only applied to regions where *xtime* is positive and is given by:

$$damp = 10^{-\left(\frac{xtime}{period}\right)^2}$$

The wavelength adopted for both features is $0.25a$ and the amplitudes of the sinusoids for the upper and lower surfaces are $1.0\mu m$ and $0.7\mu m$ respectively. The amplitudes are very large compared to the minimum film thickness, given by Dowson and Higginson [Dowson, 1968] formula, of $0.187\mu m$. It can be seen that the combined amplitude of the sinusoids is nearly ten times greater than the minimum film thickness generated within the contact under smooth surface conditions. The two sinusoids used are shown in Figure 7.1, also shown is the minimum film thickness.

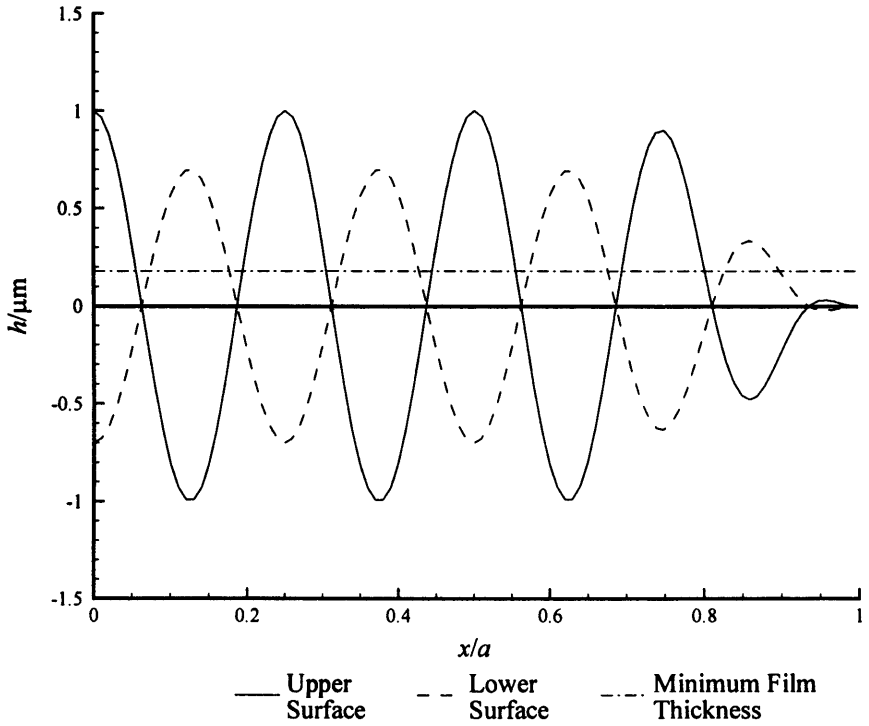

Figure 7.1 – Comparison of the two sinusoids to the minimum film thickness

The computational mesh used to solve the problem has 400 points in the semi Hertz dimension, a, this gives 100 points to define each sinusoidal feature which is more than sufficient. The mesh ranges from $-2.5a$ to $1.5a$, hence, the total number of points in the mesh is 1601.

Figures 7.2 (a) to 7.2 (c) show the results corresponding to the above conditions for a number of time steps, the solutions are isothermal and are presented to indicate how the pressure and film thickness vary with time. It can be seen that the pressures (shown in red) are very high and the film thickness (shown in green) is very small at the asperity tips but, for the isothermal case considered, there is no contact developed. In the regions where the two valley features are aligned the pressures are very low and can be seen to cavitate due to the restriction of flow into the area due to the constriction formed at the asperity tips. The regions increase in size as the sinusoids move out of phase and the volume in the regions increases. However, due to the small film thickness upstream there is not sufficient fluid flow to fill the void, leading to cavitation over a large area. As the asperities continue to move out of phase, the pressures at the asperity tips drop and the film thickness increases.

Since it is seen that regions of cavitation occur in the isothermal solution to the problem, the manner in which this is to be dealt with within the thermal solution approach needs to be determined. Since cavitated regions are a mixture of air and oil, it can be very difficult to determine the exact behaviour within these regions. As a simplification, within cavitated regions the oil is considered to have the same density that it has at ambient pressure. In addition, the shear heating effects are ignored as the air will shear a lot easier than the oil, thus if the oil does not shear then there can be no shear heating.

During initial investigations using this case and the thermal solution technique there were numerous problems with the thermal solution technique breaking down and causing the whole program to crash. Significant work was undertaken to try and determine the cause of the instability within the calculations. It was seen that there were oscillations occurring within the temperature profiles developed across the film which propagated in time. The oscillations grew until they became dominant and influenced the Reynolds equation, leading to a break down in the overall solution scheme.

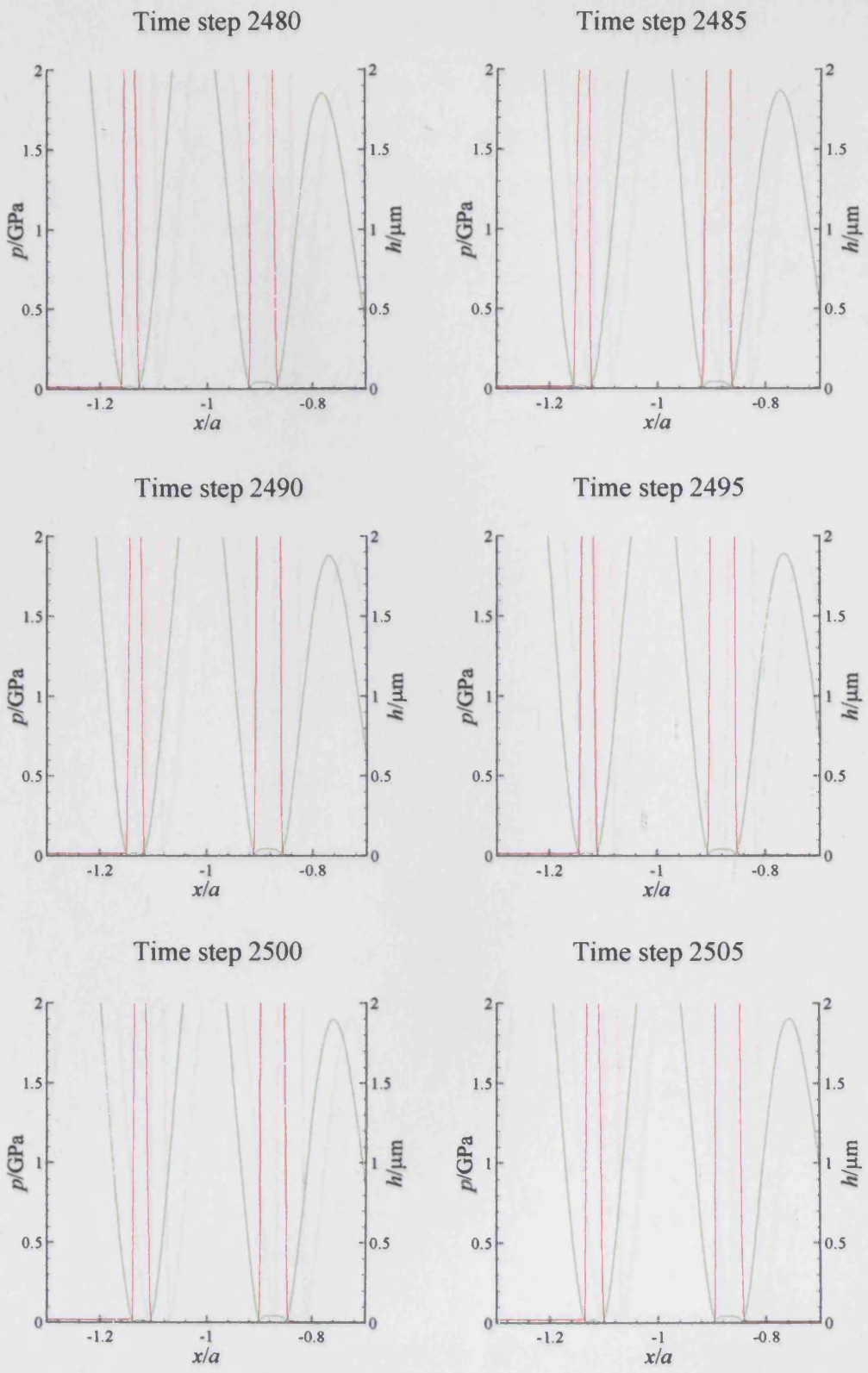

Figure 7.2 (a) – Isothermal results of pressure (red) and film thickness (green) for the thermal test case every 5 time steps from 2480 to 2505

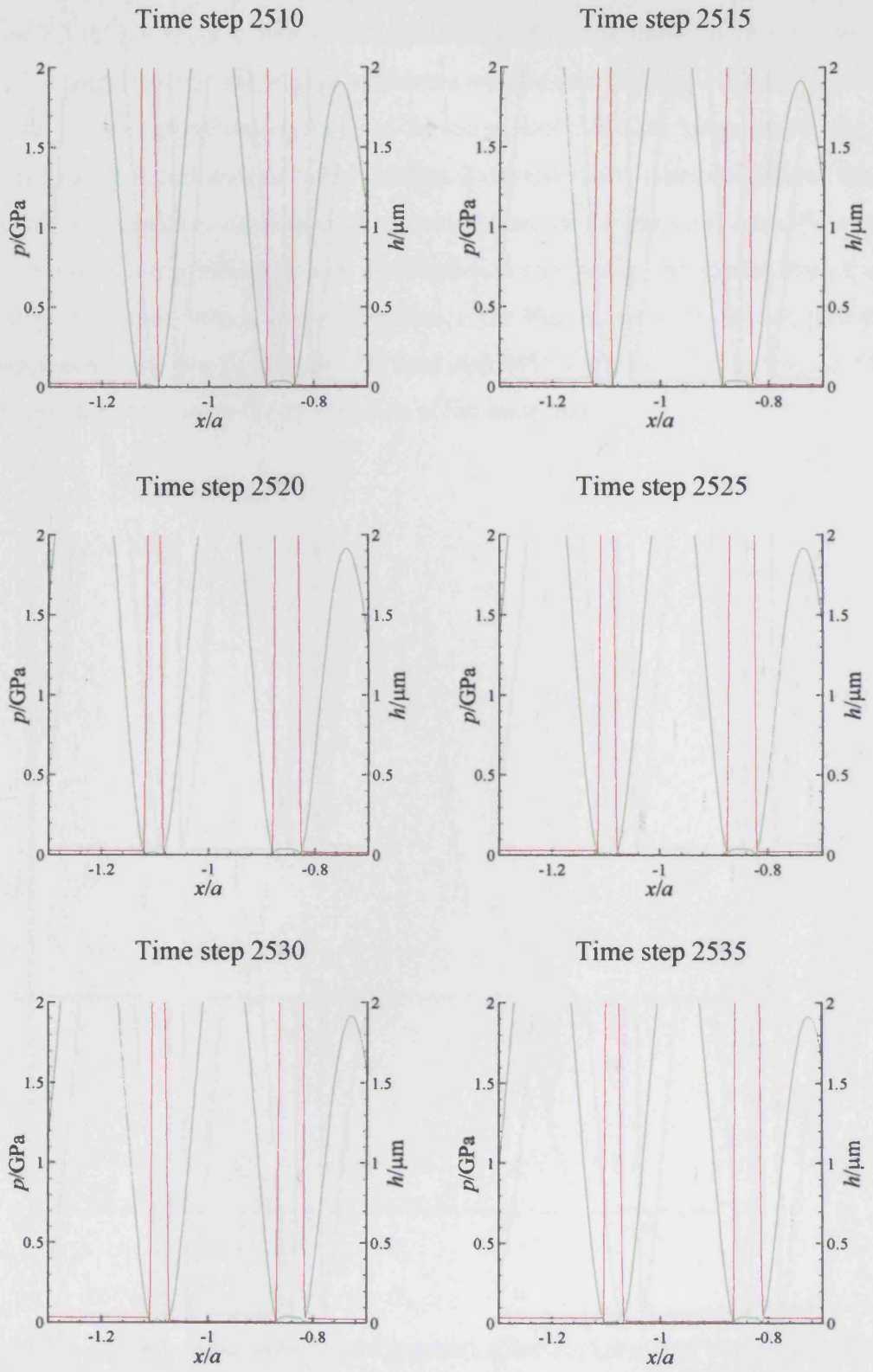

Figure 7.2 (b) – Isothermal results of pressure (red) and film thickness (green) for the thermal test case every 5 time steps from 2510 to 2535

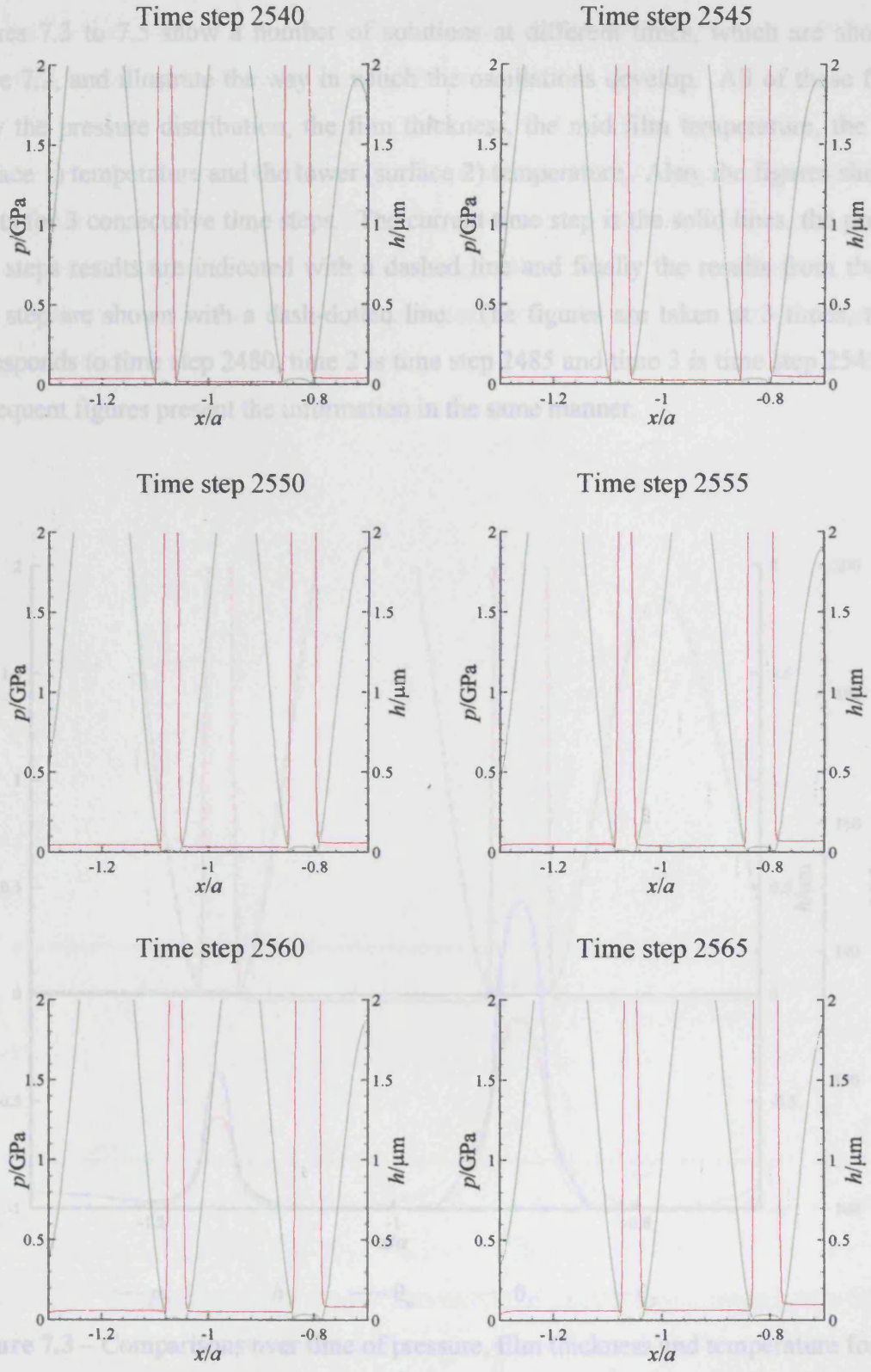

Figure 7.2 (c) – Isothermal results of pressure (red) and film thickness (green) for the thermal test case every 5 time steps from 2540 to 2565

Figures 7.3 to 7.5 show a number of solutions at different times, which are shown in figure 7.2, and illustrate the way in which the oscillations develop. All of these figures show the pressure distribution, the film thickness, the mid film temperature, the upper (surface 1) temperature and the lower (surface 2) temperature. Also, the figures show the results for 3 consecutive time steps. The current time step is the solid lines, the previous time steps results are indicated with a dashed line and finally the results from the next time step are shown with a dash-dotted line. The figures are taken at 3 times, time 1 corresponds to time step 2480, time 2 is time step 2485 and time 3 is time step 2545. All subsequent figures present the information in the same manner.

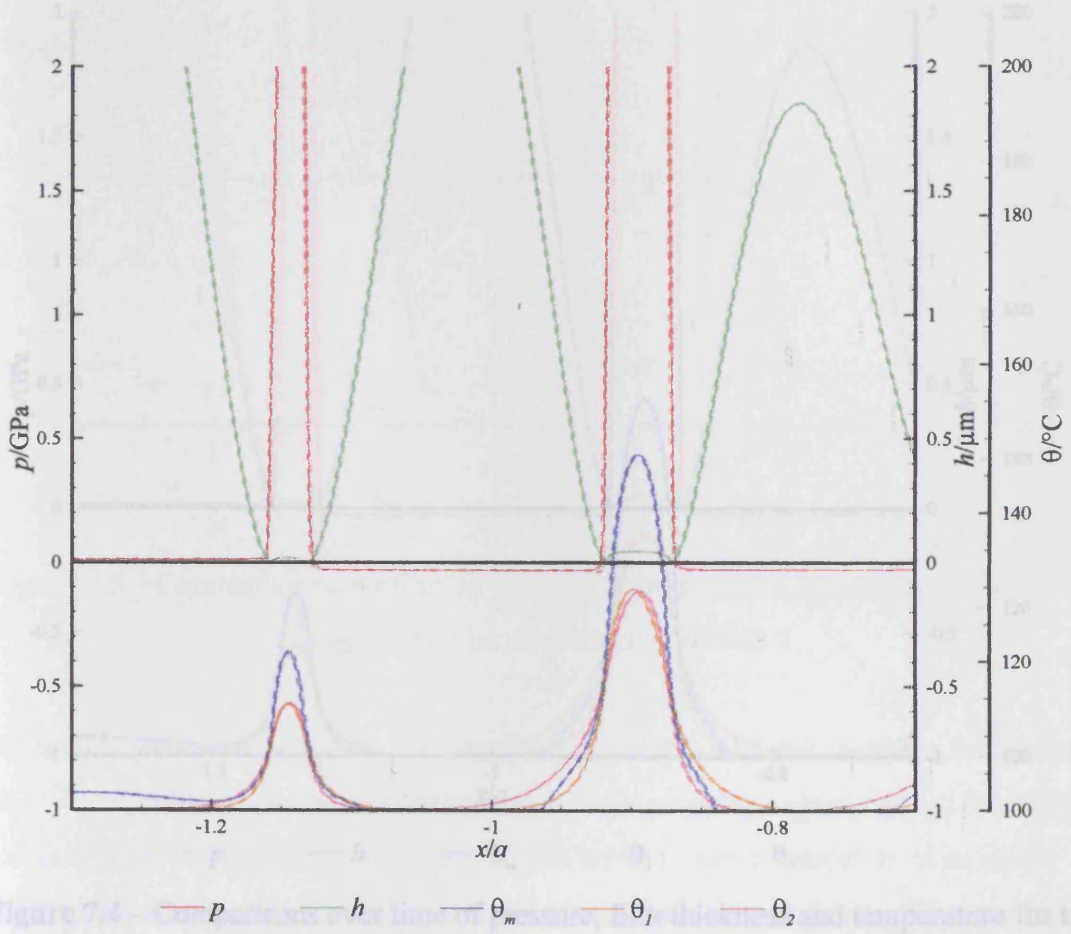

Figure 7.3 – Comparisons over time of pressure, film thickness and temperature for two large sinusoidal roughnesses for time 1

Figure 7.3 shows no evidence of an oscillation and can be compared to figure 7.2. It can be seen that the pressure and film thickness show the same variations, however, the film

thickness is seen to be smaller in the thermal case due to the shear heating effect in the inlet region reducing the film thickness globally.

It can be seen that for the next time step, shown in Figure 7.4, there is a series of different slopes occurring at $x/a = -1.15$. It can be seen that a kink is formed in the mid-film temperature, this develops over a short number of time steps. This variation in the mid-film temperature indicates that the temperature profile developed across the film keeps changing in time.

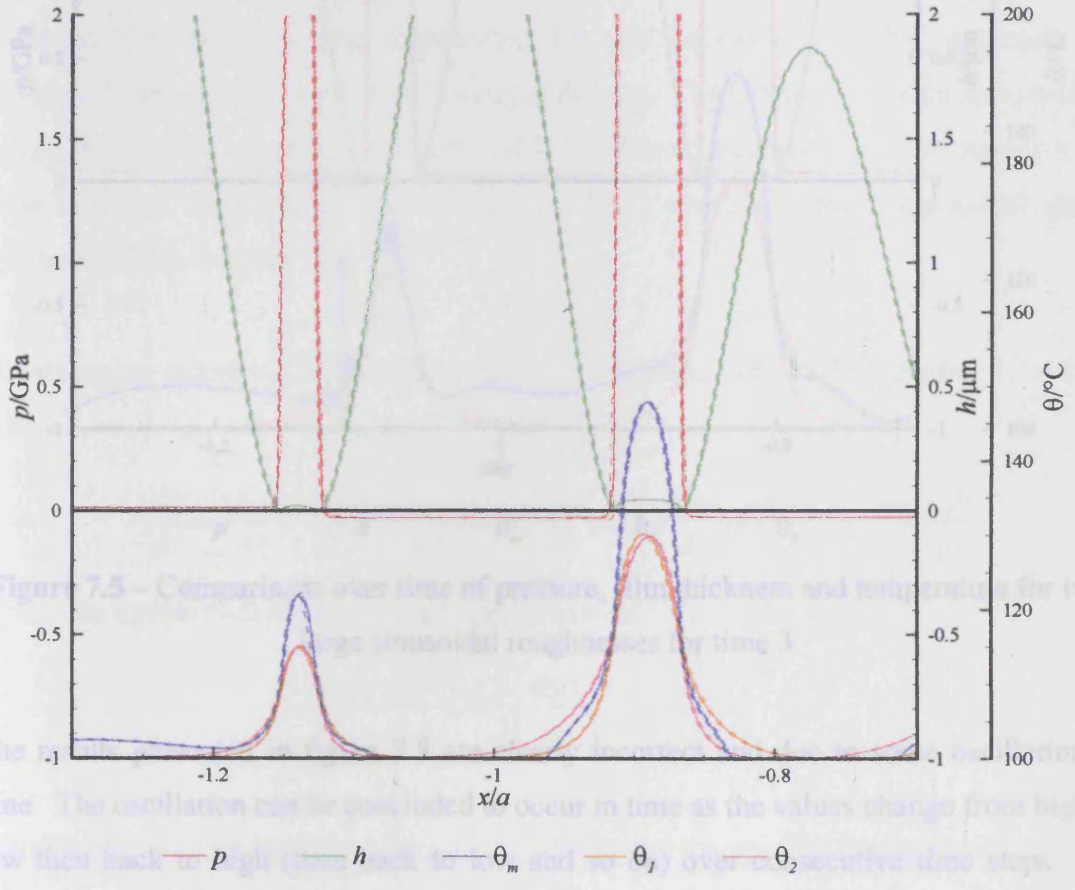

Figure 7.4 – Comparisons over time of pressure, film thickness and temperature for two large sinusoidal roughnesses for time 2

The slight oscillation illustrated in figure 7.4 propagates in time, further oscillations in other locations also occur leading to the situation shown in Figure 7.5, which shows the results occurring at time 3. The slight oscillation has progressed and several other spikes have formed in the mid-film temperature, however, these features are not seen in the

surface temperatures and do not appear to have a marked effect on the pressure and film thickness calculated.

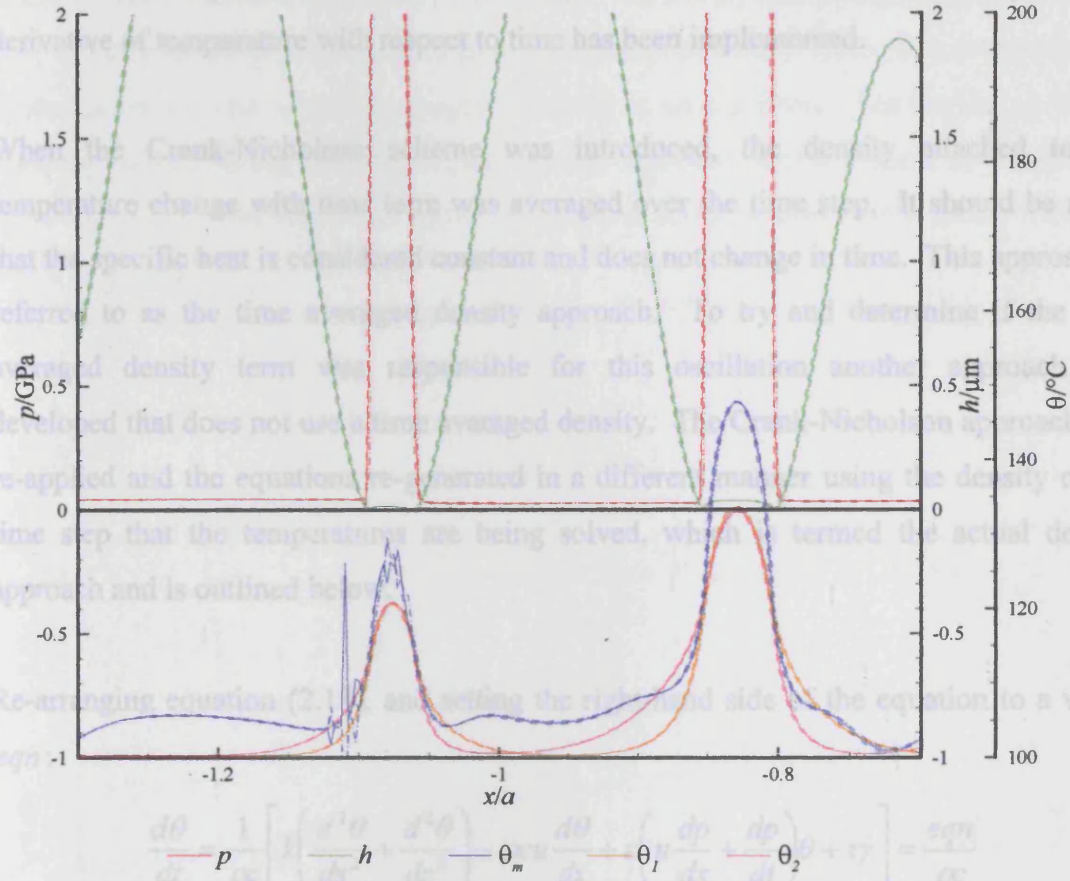

Figure 7.5 – Comparisons over time of pressure, film thickness and temperature for two large sinusoidal roughnesses for time 3

The results presented in figure 7.5 are clearly incorrect and due to some oscillation in time. The oscillation can be concluded to occur in time as the values change from high to low then back to high (then back to low and so on) over consecutive time steps. An investigation was undertaken to try and determine the origin of this oscillation. All the terms that are involved in the thermal equation were examined over time for the regions where the oscillations occur. It was seen that there were slight changes occurring in the compressive heating and shear heating terms, the other spatial terms were also investigated and the changes were found to be negligible.

To try and determine which of the two terms was more volatile they were independently removed from the calculation. It was found that removing either or both of the terms did not influence the results, there was still evidence of oscillations, although they were slightly smaller than those presented in figure 7.5. Since the heating terms and spatial terms are not responsible for the oscillation then it must be related to the way that the derivative of temperature with respect to time has been implemented.

When the Crank-Nicholson scheme was introduced, the density attached to the temperature change with time term was averaged over the time step. It should be noted that the specific heat is considered constant and does not change in time. This approach is referred to as the time averaged density approach. To try and determine if the time averaged density term was responsible for this oscillation another approach was developed that does not use a time averaged density. The Crank-Nicholson approach was re-applied and the equations re-generated in a different manner using the density of the time step that the temperatures are being solved, which is termed the actual density approach and is outlined below.

Re-arranging equation (2.18), and setting the right-hand side of the equation to a value, eqn:

$$\frac{d\theta}{dt} = \frac{1}{\rho c}\left[k\left(\frac{d^2\theta}{dx^2} + \frac{d^2\theta}{dz^2}\right) - \rho c u \frac{d\theta}{dx} + \varepsilon\left(u\frac{dp}{dx} + \frac{dp}{dt}\right)\theta + \tau\dot{\gamma}\right] = \frac{eqn}{\rho c}$$

Using the Crank-Nicholson approach this becomes:

$$\frac{\theta^t - \theta^{t-1}}{\delta t} = \frac{1}{2}\left[\frac{eqn}{\rho c}\right]^t + \frac{1}{2}\left[\frac{eqn}{\rho c}\right]^{t-1}$$

Multiplying through by the product of the density and the specific heat at the current time step:

$$2[\rho c]^t \frac{\theta^t - \theta^{t-1}}{\delta t} = eqn^t + \frac{[\rho c]^t}{[\rho c]^{t-1}} eqn^{t-1}$$

The results obtained using this actual density approach for times 1, 2 and 3 are shown in Figures 7.6 to 7.8. It is evident from all the figures that the change made to the thermal equation has had no effect on the temperature developed, in addition, it can also be seen

that the pressure and film thickness generated have remained the same. This was encouraging as it meant that the formulation change did not affect the values calculated, but did influence the stability of the approach.

Figure 7.6 does not show any signs of oscillation. However, closer inspection reveals that there is a change in slope of the mid film temperature at $x/a = -1.13$, this is the start of an oscillation and can be seen to progress quickly as time evolves. The beginning of the oscillation occurs earlier than the average density approach, this is evident by comparing figure 7.6 to 7.3. The oscillation grows quickly and develops into a large spike, which is shown in Figure 7.7.

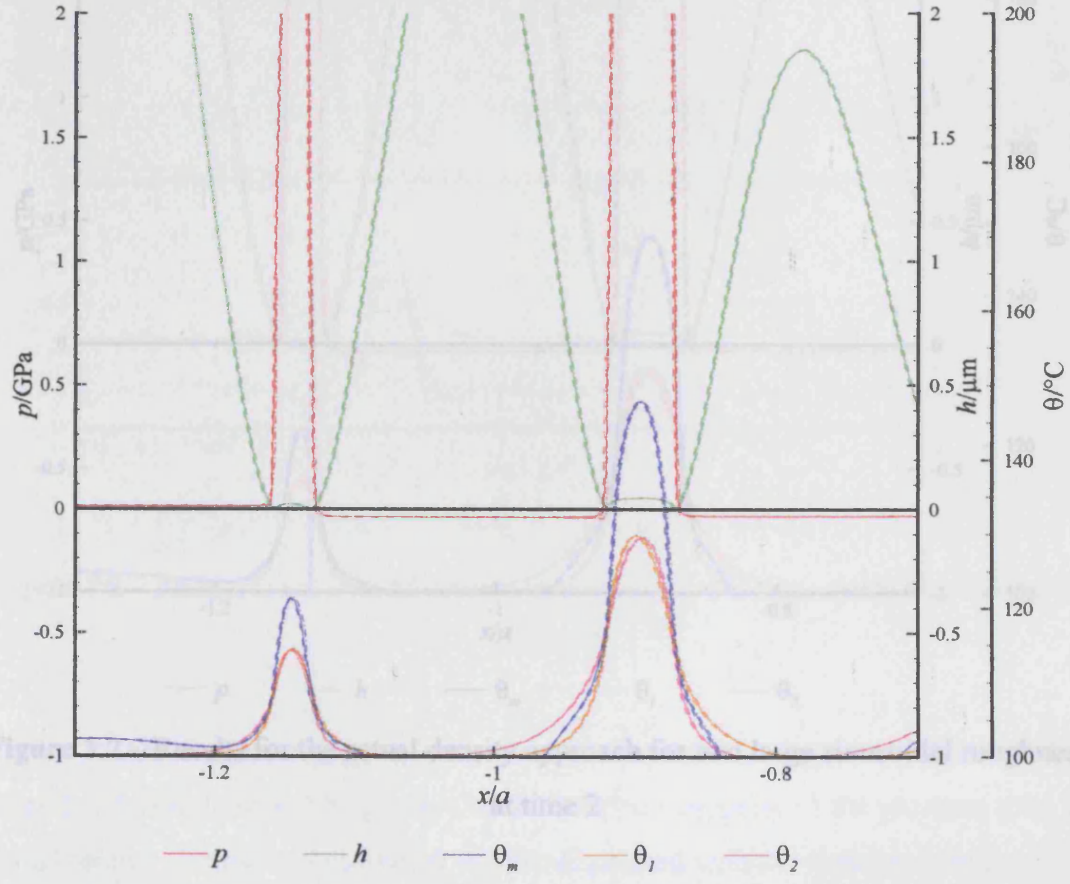

Figure 7.6 – Results for the actual density approach for two large sinusoidal roughnesses at time 1

It can be seen from figure 7.7 that the spike, which results from the beginning of the oscillation shown in figure 7.6, has a large amplitude but is localised, only influencing the

two adjacent mesh points. Comparing this to figure 7.4, it can be seen that the average density approach gives a much smaller oscillation than the actual density approach. The average density approach oscillation produces a temperature difference of 1°C at time 2, whereas the oscillation developed for the actual density approach produces a temperature difference of 52°C at time 2. It is clearly evident that the average density approach is more stable than the actual density approach. However, considering the result at time 3 in Figure 7.8 the oscillation has almost disappeared.

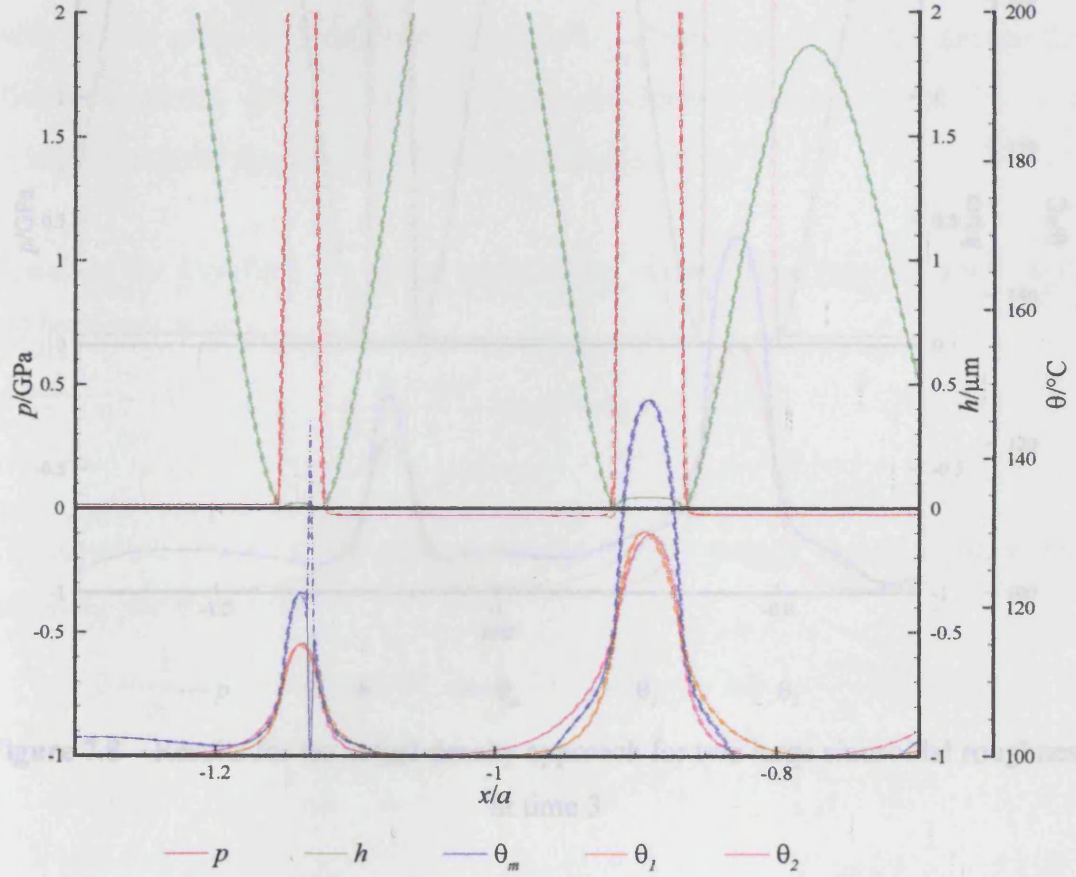

Figure 7.7 – Results for the actual density approach for two large sinusoidal roughnesses at time 2

Figure 7.8 does show a slight kink in the mid-film temperature at $x/a = 1.06$, this is much better than the result from the average density approach shown in figure 7.5. There is only the influence of the one oscillation considered in figures 7.6 and 7.7, which has moved downstream at the mean entrainment speed of the oil. This clearly illustrates that the actual density approach has the ability to return back to a more reasonable result and

does not allow the oscillation to have an effect on the surrounding points, which can be seen for the average density approach. It may be a better approach, but the oscillation developed means that it is not a valid result and a better solution method needs to be developed.

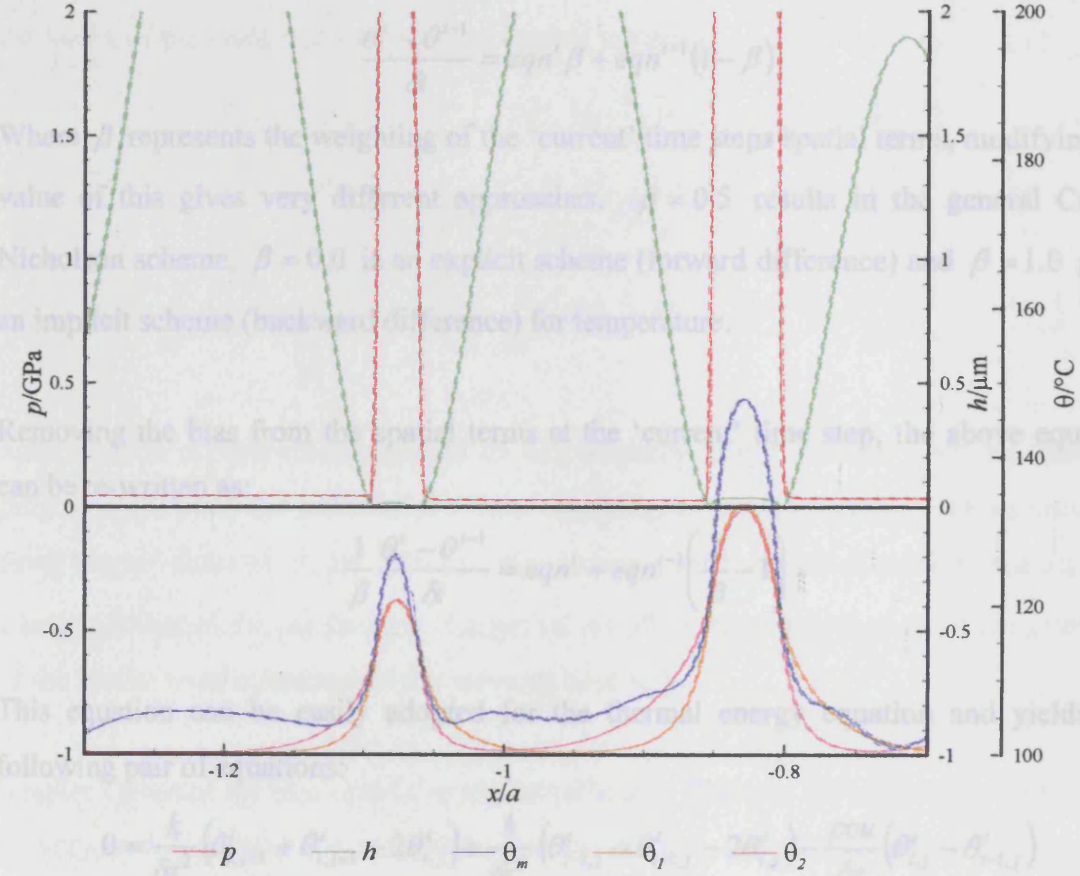

Figure 7.8 – Results for the actual density approach for two large sinusoidal roughnesses at time 3

It was thought that the change in pressure over time was having some influence on the temperature calculations through the Crank-Nicholson coupling of the previous time steps spatial terms. The thermal equations were re-expressed with a weighting function used to allow the influence that the previous time step's spatial terms have on the calculations to be increased or decreased to asses the effect on the observed oscillations.

Consider a simple time dependent equation of the form:

$$\frac{d\theta}{dt} = eqn$$

The Crank-Nicholson scheme averages the spatial terms, eqn, over the two time steps. As a variation on this, the ratio of the averaging was amended so that a bias could be introduced. Considering this approach then the previous equation can be re-written as:

$$\frac{\theta^t - \theta^{t-1}}{\delta t} = eqn^t \beta + eqn^{t-1}(1-\beta)$$

Where β represents the weighting of the 'current' time steps spatial terms, modifying the value of this gives very different approaches. $\beta = 0.5$ results in the general Crank-Nicholson scheme, $\beta = 0.0$ is an explicit scheme (forward difference) and $\beta = 1.0$ gives an implicit scheme (backward difference) for temperature.

Removing the bias from the spatial terms at the 'current' time step, the above equation can be re-written as:

$$\frac{1}{\beta}\frac{\theta^t - \theta^{t-1}}{\delta t} = eqn^t + eqn^{t-1}\left(\frac{1}{\beta} - 1\right)$$

This equation can be easily adopted for the thermal energy equation and yields the following pair of equations:

$$0 = \frac{k}{\delta z^2}\left(\theta^t_{i,j-1} + \theta^t_{i,j+1} - 2\theta^t_{i,j}\right) + \frac{k}{\delta x^2}\left(\theta^t_{i-1,j} + \theta^t_{i+1,j} - 2\theta^t_{i,j}\right) - \frac{\rho c u}{\delta x}\left(\theta^t_{i,j} - \theta^t_{i-1,j}\right)$$
$$+ \varepsilon\left(u\frac{dp}{dx} + \frac{dp}{dt}\right)\theta^t_{i,j} - \frac{1}{\beta}\frac{\hat{\rho}c}{\delta t}\left(\theta^t_{i,j} - \theta^{t-1}_{i,j}\right) + \tau\dot{\gamma} + \left(\frac{1}{\beta} - 1\right)rhs$$

$$rhs = \frac{k}{\delta z^2}\left(\theta^{t-1}_{i,j-1} + \theta^{t-1}_{i,j+1} - 2\theta^{t-1}_{i,j}\right) + \frac{k}{\delta x^2}\left(\theta^{t-1}_{i-1,j} + \theta^{t-1}_{i+1,j} - 2\theta^{t-1}_{i,j}\right) - \frac{\rho c u}{\delta x}\left(\theta^{t-1}_{i,j} - \theta^{t-1}_{i-1,j}\right)$$
$$+ \varepsilon\left(u\frac{dp}{dx} + \frac{dp}{dt}\right)^{t-1}\theta^{t-1}_{i,j} + \tau\dot{\gamma}^{t-1}$$

In moving the bias from the 'current' time step there is a weighting factor applied to the rhs, it can be seen that this scaling factor goes to zero when the implicit scheme, $\beta = 1.0$, is used. The temperature at the previous time step that comes from the term representing the time derivative of temperature has previously been linked in with the rhs

and has been moved out from the *rhs* in the above equations. This was to ensure that the influence it has is not removed when the implicit approach is used. This approach and manipulation ensures that the equations that have already been developed do not need comprehensive change, the effect of the weighting function only modifies the coefficient of the time derivative of temperature terms and a general scaling of the *rhs* term

A similar pair of equations can be determined for the solid, the only difference being the exclusion of the shear and compressive heating terms:

$$0 = \frac{k}{\delta z^2}\left(\theta^t_{i,j-1} + \theta^t_{i,j+1} - 2\theta^t_{i,j}\right) + \frac{k}{\delta x^2}\left(\theta^t_{i-1,j} + \theta^t_{i+1,j} - 2\theta^t_{i,j}\right)$$
$$- \frac{\rho c u}{\delta x}\left(\theta^t_{i,j} - \theta^t_{i-1,j}\right) - \frac{1}{\beta}\frac{\rho c}{\delta t}\left(\theta^t_{i,j} - \theta^{t-1}_{i,j}\right) + \left(\frac{1}{\beta} - 1\right) rhs$$

$$rhs = \frac{k}{\delta z^2}\left(\theta^{t-1}_{i,j-1} + \theta^{t-1}_{i,j+1} - 2\theta^{t-1}_{i,j}\right) + \frac{k}{\delta x^2}\left(\theta^{t-1}_{i-1,j} + \theta^{t-1}_{i+1,j} - 2\theta^{t-1}_{i,j}\right) - \frac{\rho c u}{\delta x}\left(\theta^{t-1}_{i,j} - \theta^{t-1}_{i-1,j}\right)$$

This allowed different formulations to be compared easily without need for different programs allowing the effect of the bias weighting to be investigated. It was seen that using larger values of β, i.e. $\beta = 1.0$, did allow a greater degree of stability and stopped a large number of the oscillations. Larger values of the bias remove most of the influence of the spatial term influences at the previous time step.

Smaller values of the bias lead to an explicit scheme. This was investigated and proved to be very unstable under the conditions considered here. Therefore, when looking at the effect that the bias had, the investigation only considered values where $0.5 \leq \beta \leq 1.0$ and was concerned with the actual density approach. It was seen that increasing the value of β did not seem to provide a greater degree of stability than the Crank-Nicholson scheme, it was not until $\beta > 0.95$ that the magnitude of the spike seen in figure 7.7 began to reduce.

For the implicit case where $\beta = 1.0$, no oscillations occurred in the solutions, as can be seen in Figures 7.9 to 7.11 for time 1, time 2 and time 3.

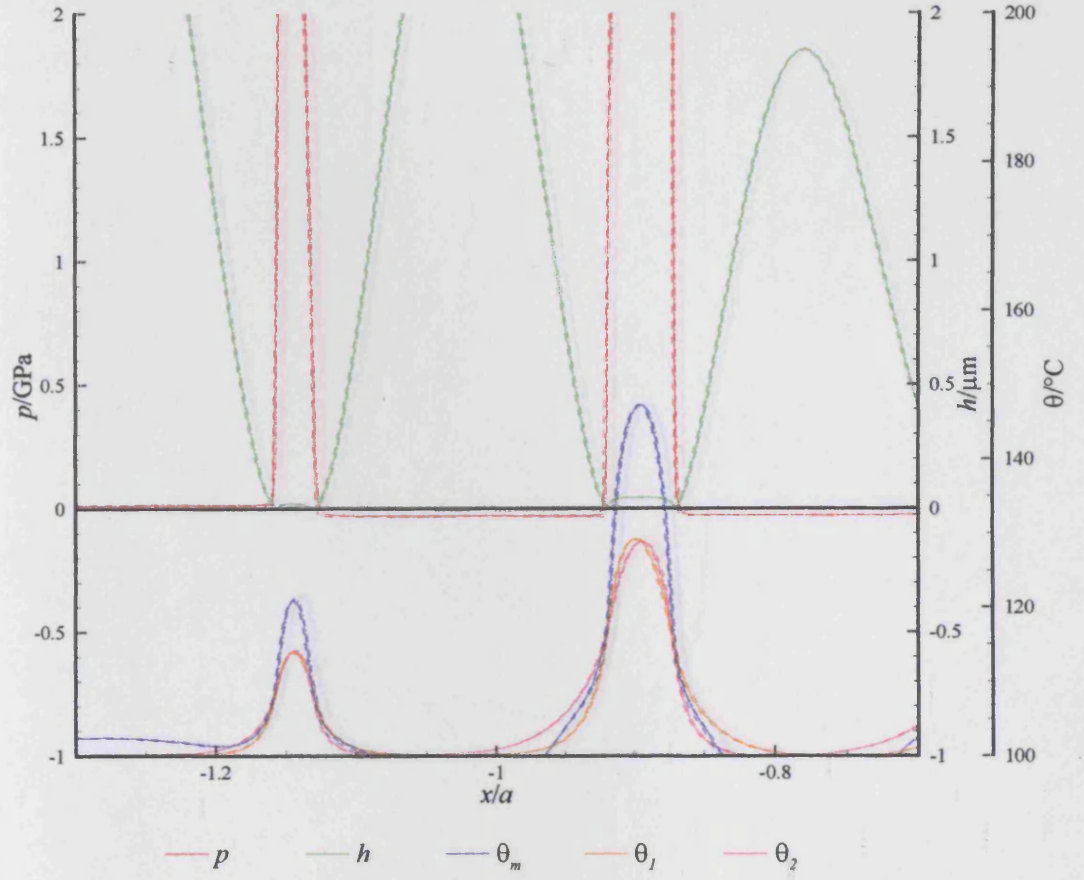

Figure 7.9 – Results for the implicit scheme for two large sinusoidal roughnesses at time 1

It can be seen in Figure 7.9 that there is no evidence to suggest the beginning of any oscillations. It can also be seen that the results for the mid-film temperature and both surface temperatures give very good agreement with the Crank-Nicholson scheme, this can be seen by comparing figure 7.9 to figure 7.3.

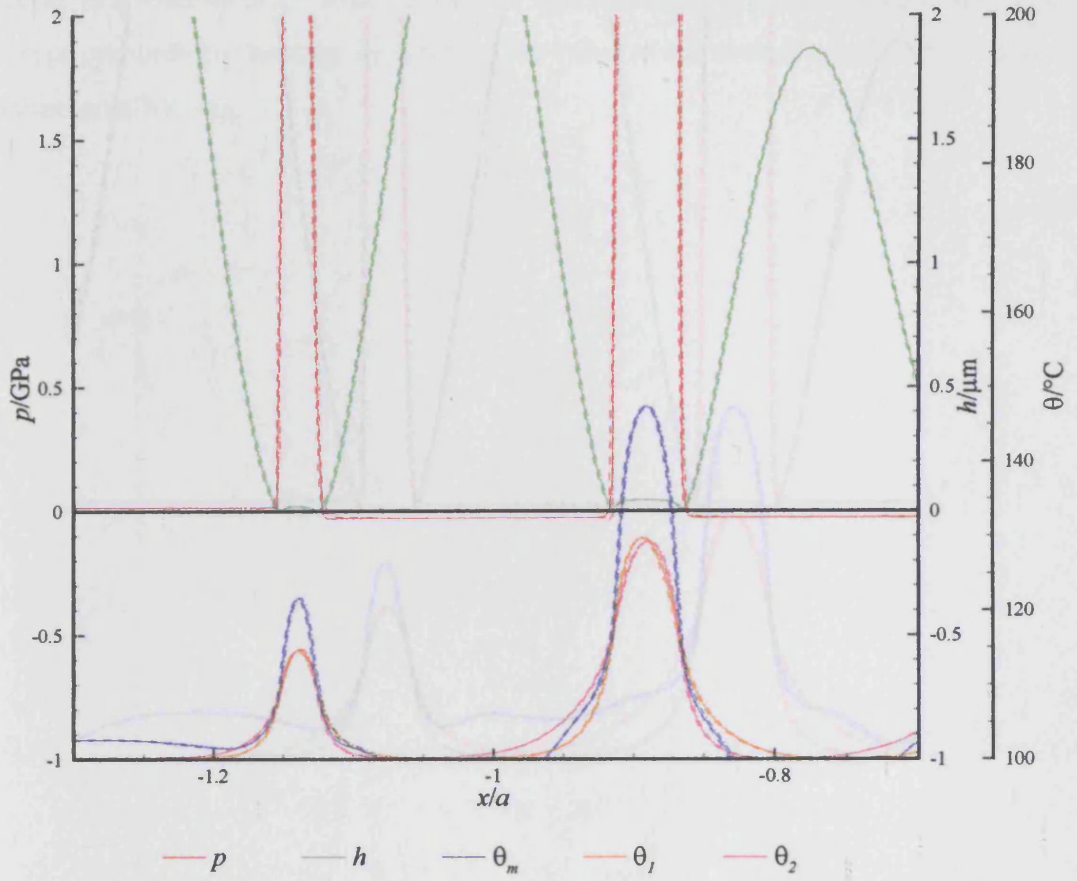

Figure 7.10 – Results for the implicit scheme for two large sinusoidal roughnesses at time 2

Even a few time steps later, at time 2 shown in Figure 7.10, there is nothing to suggest that any oscillation is occurring. Comparing this result to figures 7.4 and 7.7, it is clear that the implicit method is much more stable and robust. This is further evident by looking at the results at time step 3 shown in Figure 7.11. There are still no kinks in the mid-film of surface temperatures that would have remained had any oscillation or other slight anomaly occurred. This method has eliminated the problems seen thus far with the Crank-Nicholson scheme. The Crank-Nicholson scheme is widely used due to its improved accuracy, so further investigation was undertaken to try and determine why this is not the case here.

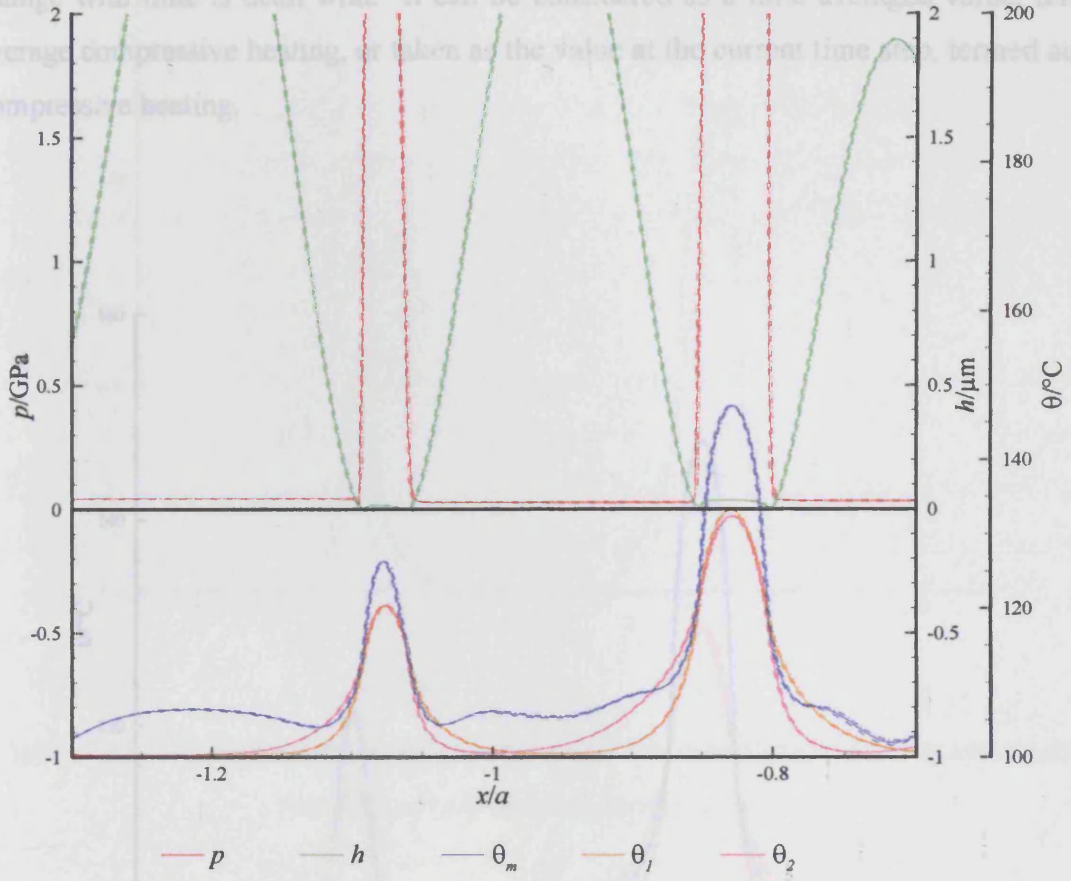

Figure 7.11 – Results for the implicit scheme for two large sinusoidal roughnesses at time 3

By definition the Crank-Nicholson scheme should be applied to all of the time dependent terms. The manner in which it has been applied thus far has been to apply it to the derivative of temperature with time term, the derivative of pressure with time term in the compressive heating term has been ignored.

Re-arranging equation (2.18):

$$\frac{d\theta}{dt} - \frac{\varepsilon}{\rho c}\frac{dp}{dt}\theta = \frac{k}{\rho c}\left(\frac{d^2\theta}{dx^2} + \frac{d^2\theta}{dz^2}\right) - u\frac{d\theta}{dx} + \frac{\varepsilon u}{\rho c}\frac{dp}{dx}\theta + \tau\dot{\gamma} = \frac{eqn}{\rho c}$$

Applying the Crank-Nicholson approach to both time derivatives this becomes:

$$\frac{\theta^t - \theta^{t-1}}{\delta t} - \frac{\varepsilon}{\rho c}\frac{p^t - p^{t-1}}{\delta t} = \frac{1}{2}\left[\frac{eqn}{\rho c}\right]^t + \frac{1}{2}\left[\frac{eqn}{\rho c}\right]^{t-1}$$

Consideration has to be given to the manner in which the coefficient of the pressure change with time is dealt with. It can be considered as a time averaged value, termed average compressive heating, or taken as the value at the current time step, termed actual compressive heating.

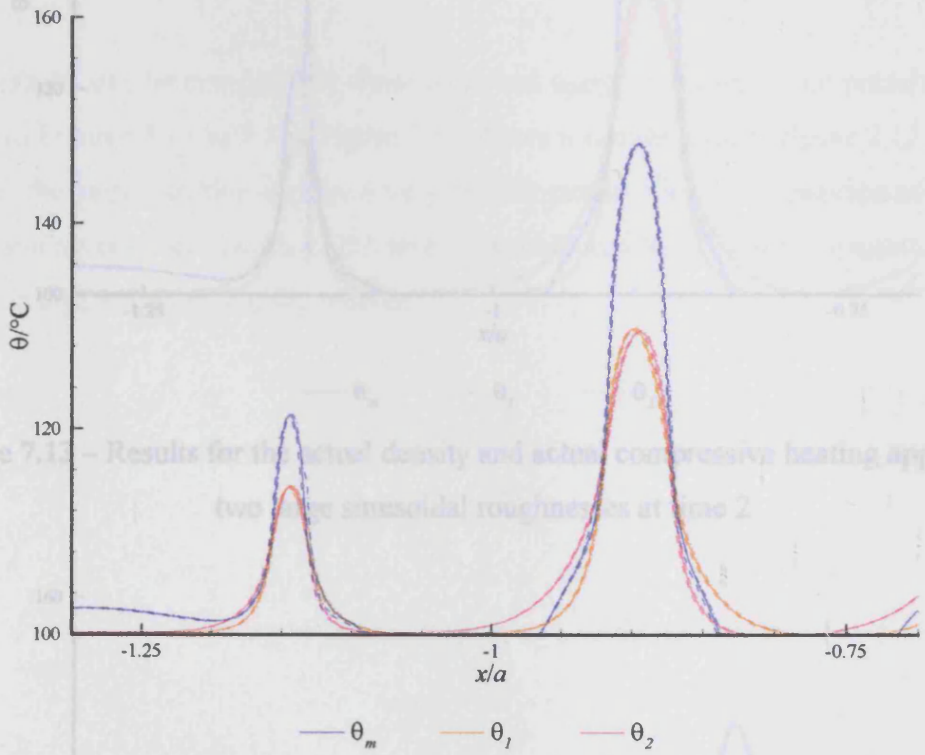

Figure 7.12 – Results for the actual density and actual compressive heating approach for two large sinusoidal roughnesses at time 1

Using the actual compressive heating and the actual density approach gives the results shown in Figures 7.12 to 7.14. Figures 7.12 to 7.14 are slightly different to those presented in figures 7.3 to 7.11 as they only show the temperatures and do not show pressure and film thickness, this was done as the pressure and film thickness do not change for the different methods and it is merely the changes that are occurring in the temperature that are important. Figure 7.12 shows that there is a slight kink occurring in the mid-film temperature at the same position, but with a larger amplitude, as that shown in figure 7.6.

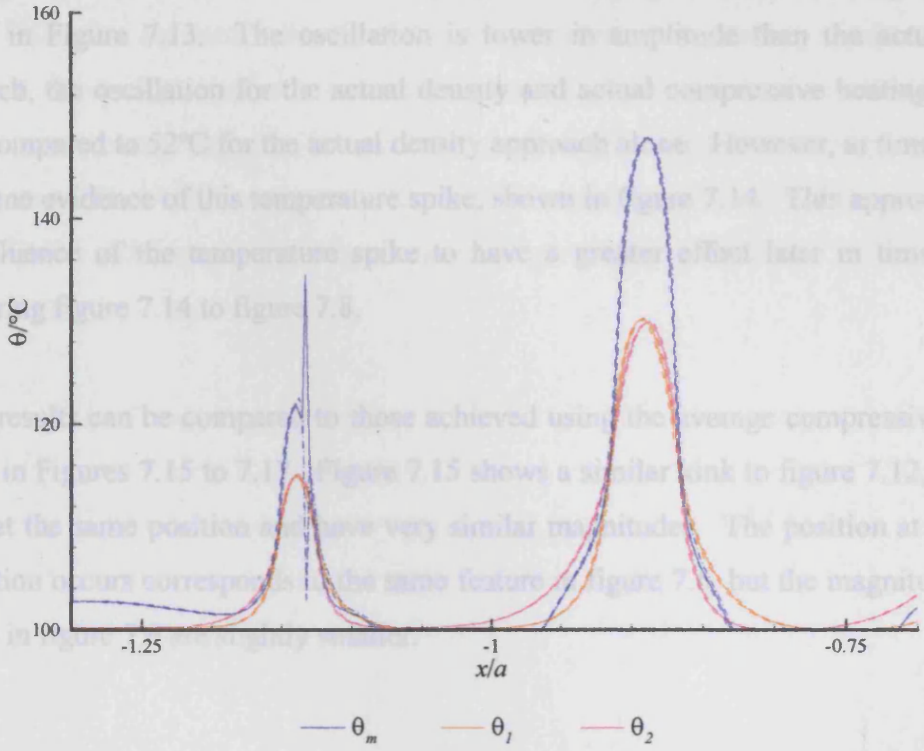

Figure 7.13 – Results for the actual density and actual compressive heating approach for two large sinusoidal roughnesses at time 2

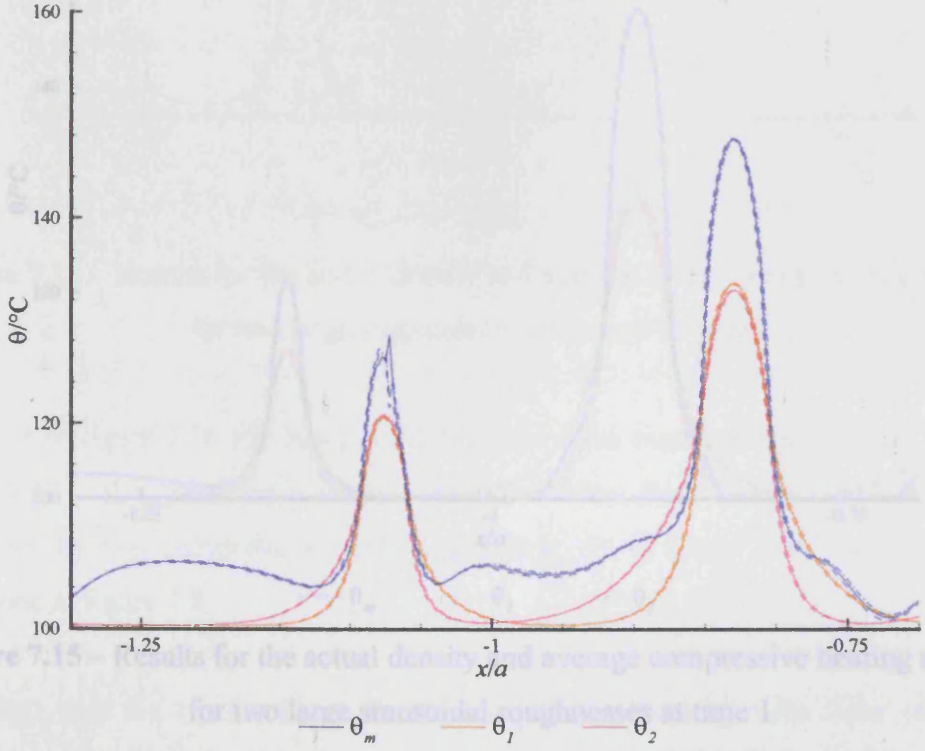

Figure 7.14 – Results for the actual density and actual compressive heating approach for two large sinusoidal roughnesses at time 3

The slight oscillation seen in figure 7.12 progresses very quickly into the large oscillation shown in Figure 7.13. The oscillation is lower in amplitude than the actual density approach, the oscillation for the actual density and actual compressive heating measures 31°C compared to 52°C for the actual density approach alone. However, at time 3 there is still some evidence of this temperature spike, shown in figure 7.14. This approach allows the influence of the temperature spike to have a greater effect later in time, seen by comparing figure 7.14 to figure 7.8.

These results can be compared to those achieved using the average compressive heating, shown in Figures 7.15 to 7.17. Figure 7.15 shows a similar kink to figure 7.12, they both occur at the same position and have very similar magnitudes. The position at which the oscillation occurs corresponds to the same feature in figure 7.6, but the magnitudes of the feature in figure 7.6 are slightly smaller.

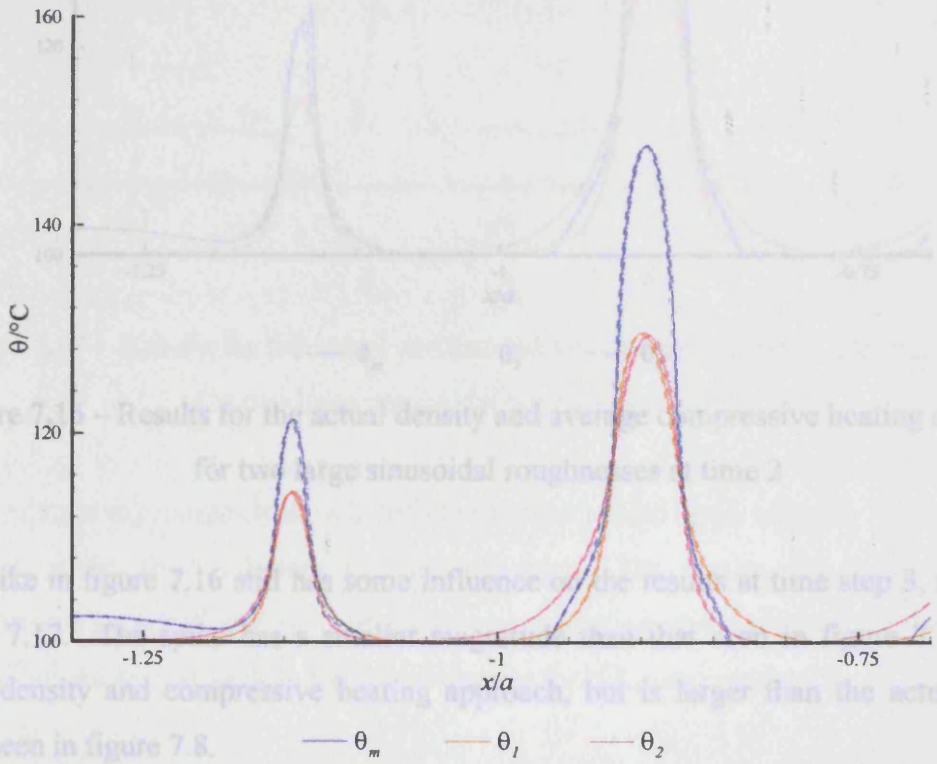

Figure 7.15 – Results for the actual density and average compressive heating approach for two large sinusoidal roughnesses at time 1

The oscillation seen in figure 7.15 again develops quickly into a larger spike shown in Figure 7.16. For the actual density and average compressive heating approach the spike produced is smaller than for the actual density alone and the actual density and compressive heating cases. The spike for the actual density and average compressive heating is 11°C, which compares to the actual density and actual compressive heating value of 31°C and the actual density alone which gave a spike of 52°C.

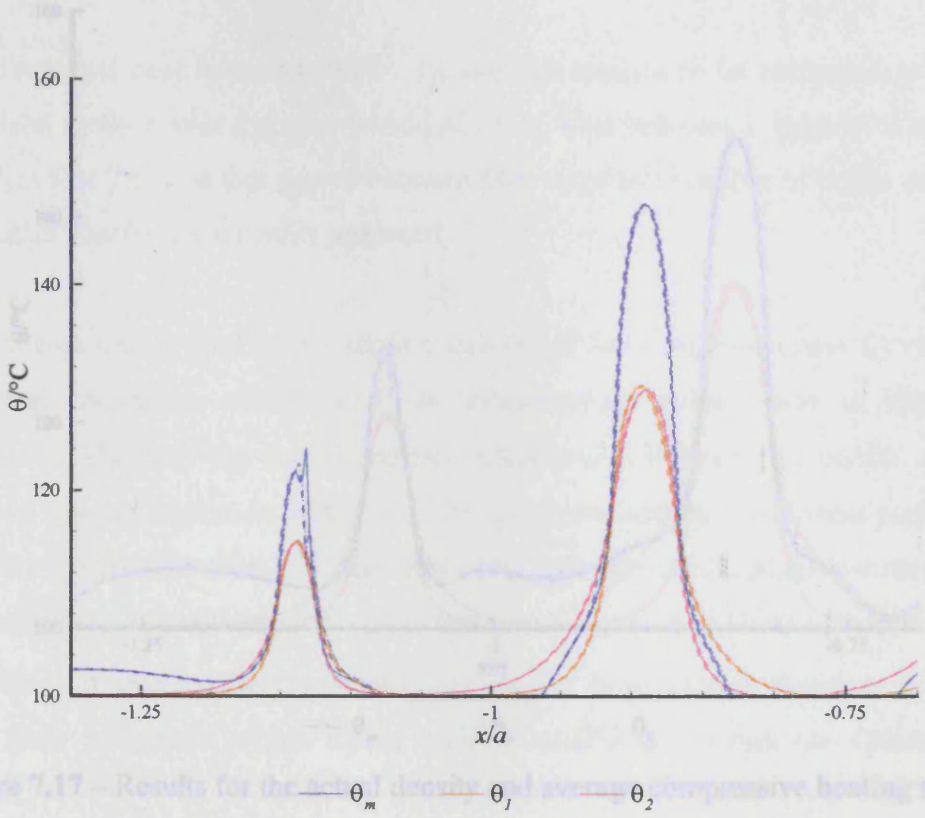

Figure 7.16 – Results for the actual density and average compressive heating approach for two large sinusoidal roughnesses at time 2

The spike in figure 7.16 still has some influence on the results at time step 3, as seen by Figure 7.17. The spike has a smaller magnitude than that seen in figure 7.14 for the actual density and compressive heating approach, but is larger than the actual density alone seen in figure 7.8.

It is clear that the modifications to the compressive heating term have reduced the magnitude of the temperature spikes seen at time 2. The modifications have, however, allowed the spikes developed to remain with a greater magnitude within the solutions

developed at time 3. The oscillations that are seen to develop at time 1 are larger when the modifications to the compressive heating terms are introduced. In addition, it can be seen that the effect of using the average compressive heating term has reduced the magnitude of all of the oscillations seen at the three time steps considered.

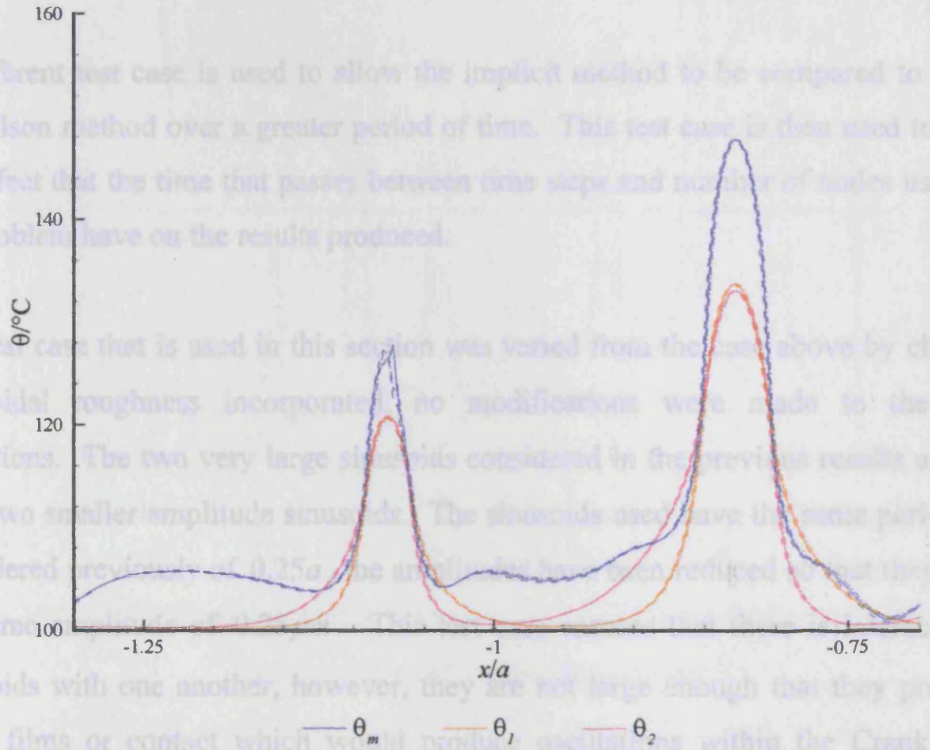

Figure 7.17 – Results for the actual density and average compressive heating approach for two large sinusoidal roughnesses at time 3

The compressive heating approach outlined here was also used with the average density approach, although the figures are not presented. The results, for the two different combinations of actual and average compressive heating, did not resist the formation of the oscillations seen as well as the actual density formulation presented.

The problems that arise with the Crank-Nicholson scheme only occur within regions where the film thickness is extremely small. It can be seen where the film thickness does not approach such values that the results are stable and no oscillations occur.

7.2 Implicit Method Validation

The results have been presented for a variety of different formulations, they all gave good agreement to one another until small film thicknesses were reached, where it was seen that this led to problems with some of the formulations. The implicit method was not sensitive to this and was found to be very robust and able to deal with these extreme conditions. This section thus covers the validation of the implicit method.

A different test case is used to allow the implicit method to be compared to the Crank-Nicholson method over a greater period of time. This test case is then used to determine the effect that the time that passes between time steps and number of nodes used to solve the problem have on the results produced.

The test case that is used in this section was varied from the case above by changing the sinusoidal roughness incorporated, no modifications were made to the operating conditions. The two very large sinusoids considered in the previous results are replaced with two smaller amplitude sinusoids. The sinusoids used have the same period as those considered previously of $0.25a$, the amplitudes have been reduced so that they both have the same amplitude of $0.25\mu m$. This test case ensures that there is interaction of the sinusoids with one another, however, they are not large enough that they produce very small films or contact which would produce oscillations within the Crank-Nicholson scheme.

The results for the Crank-Nicholson scheme using the actual density and average compressive heating approach are shown in Figure 7.18. The figures are generated in the same manner as figures 7.3 to 7.11 where there are three results shown for consecutive time steps.

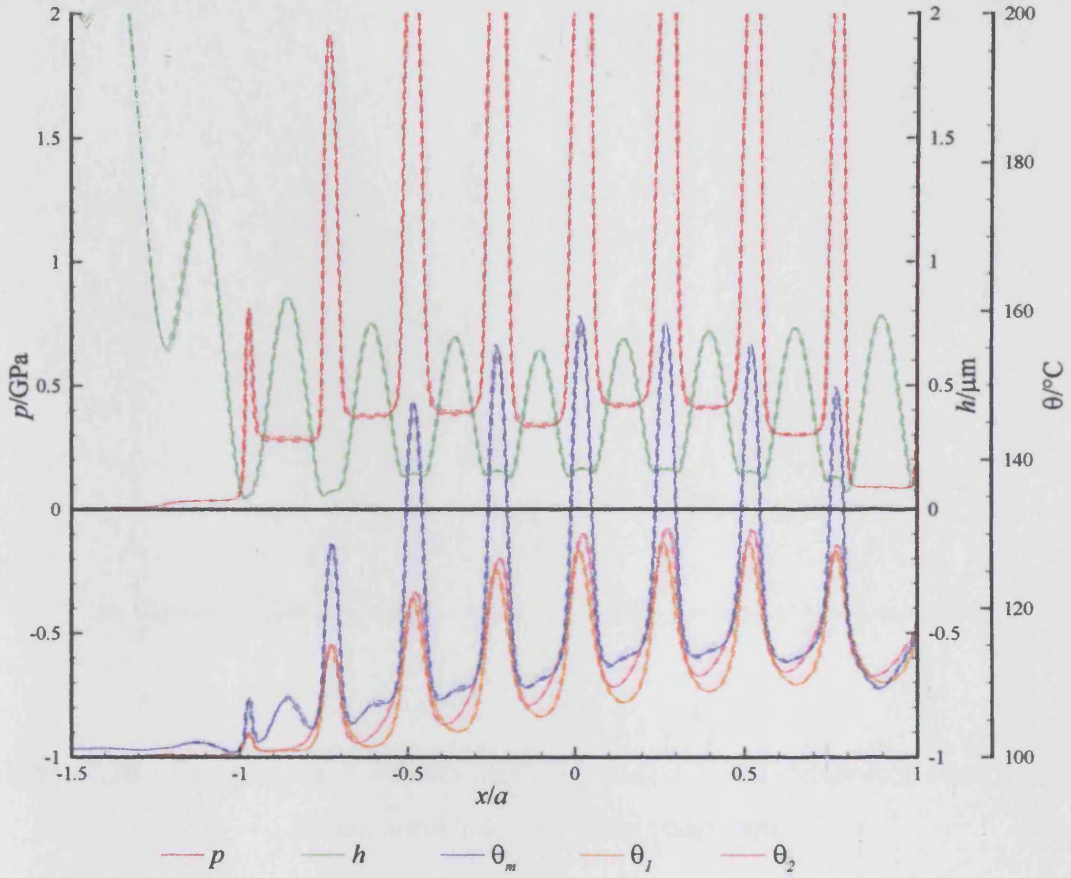

Figure 7.18 – Low amplitude sinusoid results for the Crank-Nicholson actual density and average compressive heating approach

The lower amplitude sinusoids produce larger film thicknesses, but they do generate similar pressures. The case considered does not produce small films and it can be seen that there is no sign of any oscillations within the results for the Crank-Nicholson scheme. The temperatures produced for the test case are similar to those produced for the thermal contact test case used for figures 7.3 to 7.17.

Figure 7.19 shows the result for the implicit method. It can be seen that there are no signs of oscillations or any other features that are different to that seen in figure 7.18. The temperature profiles are similar and there is no difference in the pressure and film thickness developed.

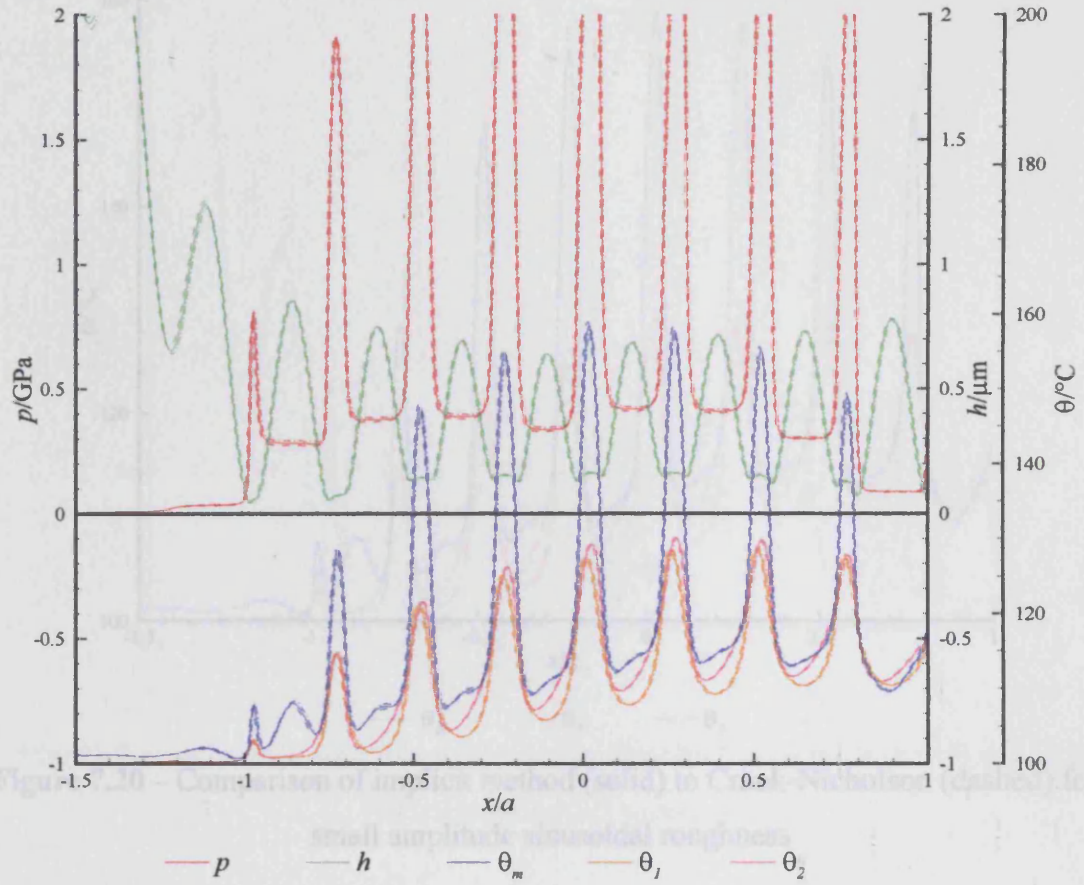

Figure 7.19 – Low amplitude sinusoid results for the implicit method

To give a better comparison, the temperature results for both approaches are plotted together in Figure 7.20. The difference between the two approaches is very slight, of the order of a couple of degrees. It is interesting to note that there is no noticeable temperature difference in the mid-film temperature responses for the high pressure regions, but there is a slight difference in the low pressure regions. It can also be seen that there is a slight difference in the temperatures developed at the two surfaces when comparing the two methods, but the difference is small. Assuming that the Crank-Nicholson results are more accurate, due to the second order accuracy of the temperature-time term, the implicit method tends to overestimate all the temperatures at low pressures and underestimate the surface temperatures in the high pressure region.

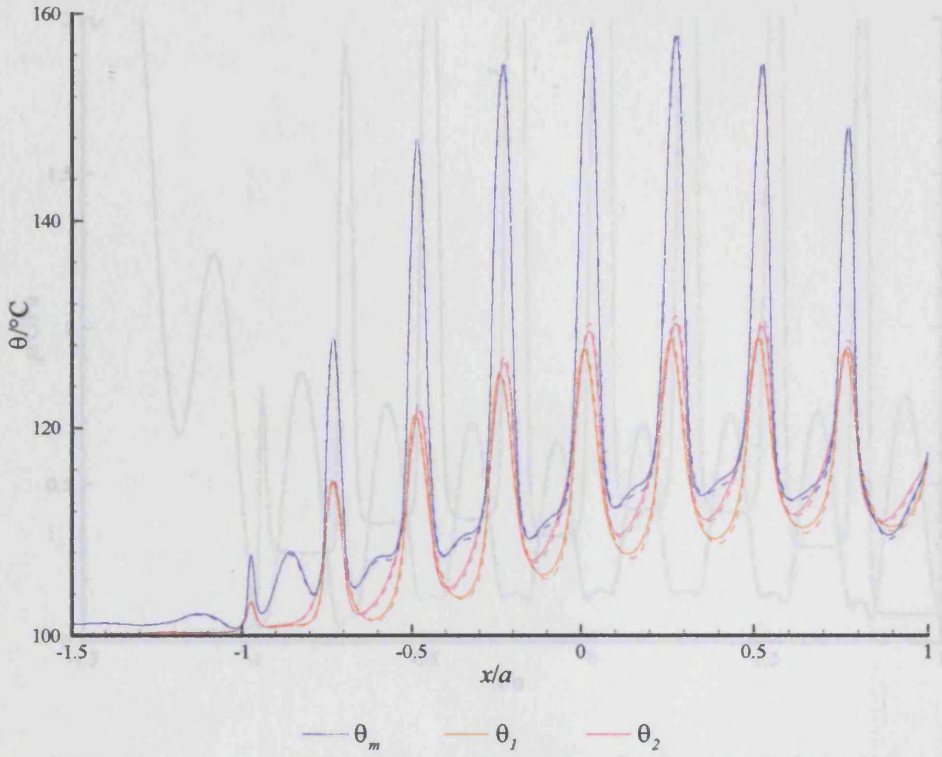

Figure 7.20 – Comparison of implicit method (solid) to Crank-Nicholson (dashed) for a small amplitude sinusoidal roughness

A large amount of time has passed allowing the differences between the different formulations for the time derivative of temperature to arise. Considering this, it can be seen that the differences that are presented are small, thus it could be argued that the results are essentially the same as the error is small. This is emphasised by comparing the pressure and film thickness response for the two approaches which is shown in Figure 7.21, where the results for both approaches are indistinguishable.

It should also be noted that the Crank-Nicholson scheme solves the above problem in less time that the implicit method. The Crank-Nicholson scheme was nearly 4% quicker than the implicit scheme, this is due to less iterative cycles being used in solving the problem. The Crank-Nicholson method performed 7.97% less iterative cycles than the implicit method. It should be noted that the implicit method does not require calculation of the previous time steps spatial terms, $rhs = 0$, this is the reason why it is quicker to perform an iteration than the Crank-Nicholson approach.

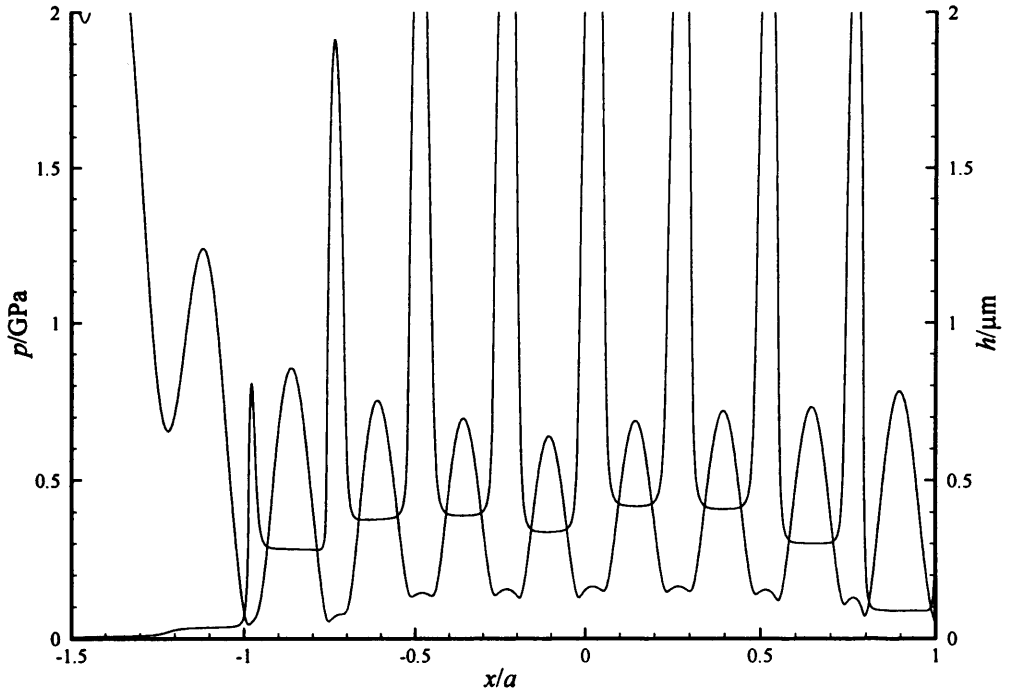

Figure 7.21 – Pressure and film thickness response for the implicit (solid) and Crank-Nicholson (dashed) approaches

Since the difference between the Crank-Nicholson scheme and the implicit method are small, it can be assumed that the implicit method is a valid, accurate solution to the thermal energy equation. Further investigation, however, is necessary to determine how sensitive the results presented are to the mesh spacing used as well as the time step used for the solutions.

The mesh spacing is considered first, where the results are presented for different mesh sizes. In each solution the time step used in the solution was kept the same which allows the effect that the mesh spacing has alone to be determined. Since the features used are sinusoidal and are generated using an analytic function, the mesh spacing should not have any influence on the way in which they are represented within the elastic equation. The features are smooth and the way in which they are represented in the elastic film thickness equation through equation (2.17) should not be sensitive to the number of points used to define them. The mesh change should only give rise to the sensitivity of the equations to the spatial approximations used.

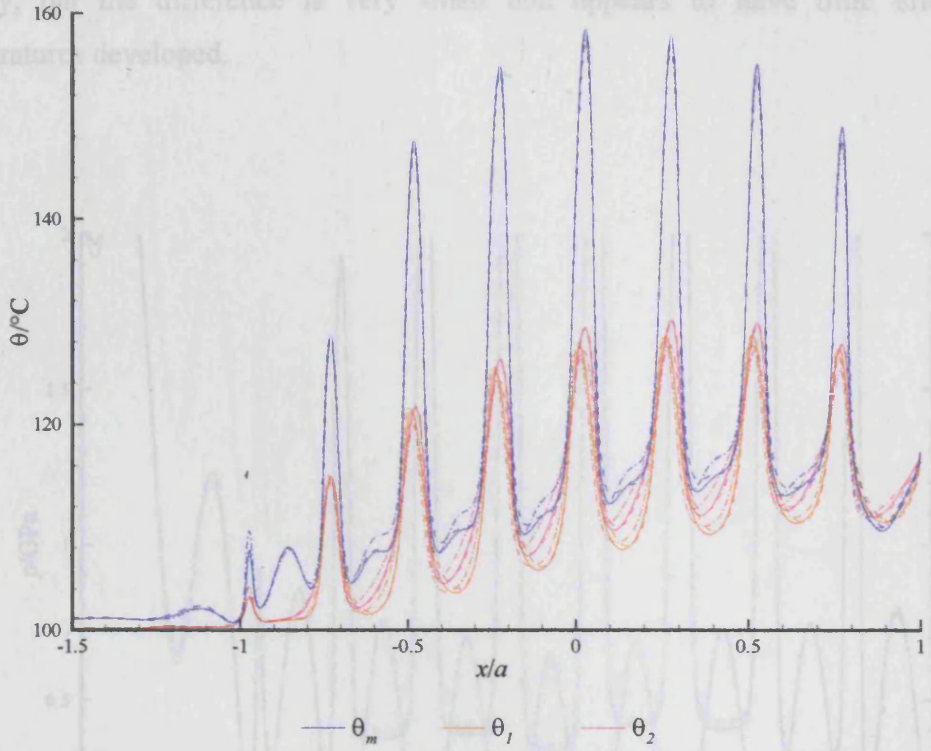

Figure 7.22 – Temperature results for the implicit method for mesh spacings of 400 nodes (solid), 200 nodes (dashed) and 800 nodes (dash-dotted)

Figure 7.22 shows the effect that the mesh spacing has on the temperatures developed, the original mesh spacing used in figure 7.20 was halved to 200 points in the half Hertz dimension, a, and also doubled to 800 points in the half Hertz dimension. The variations between the different methods arise in similar regions to the difference between the implicit method and the Crank-Nicholson scheme. The differences arise at low pressures and the differences in the high pressure regions are small. The variation seen is of the order of 5°C which is larger than the couple of degrees seen in figure 7.20.

To allow a better insight into why there is such a large difference in the results, the pressure and film thickness are compared in Figure 7.23. It can be seen that there are slight differences in the valley pressures, this could give rise to the difference in the temperatures seen in figure 7.22. The different pressures generated in the valleys will create different shear and compressive heating effects, which will cause the temperature to increase or decrease slightly. In addition to the difference in the pressure, there is also a very slight difference in the film thickness developed at the asperity tips. This is due to

the number of nodes that resolve the features, the difference occurs for the largest mesh density, but the difference is very small and appears to have little effect on the temperatures developed.

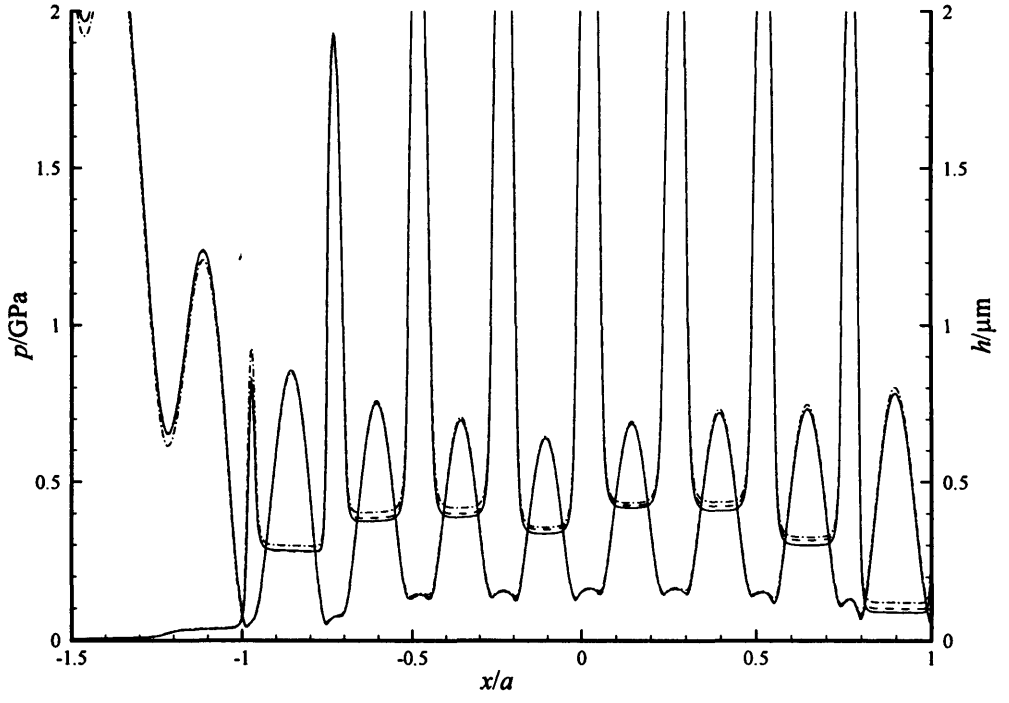

Figure 7.23 – Pressure and film thickness results for the implicit method for mesh spacings of 400 nodes (solid), 200 nodes (dashed) and 800 nodes (dash-dotted)

Comparing the results in figure 7.23 to the isothermal results for different mesh spacing shown in Figure 7.24, it can be seen that similar differences are seen. There are differences occurring in the low pressure regions for the isothermal case, but they are not quite as large as the differences seen in the thermal case. This is due to the different temperature distribution generated by the pressure feeding back into itself making the differences a bit larger. Despite this interaction, there is no evidence that the thermal solution is particularly sensitive to the number of mesh points used.

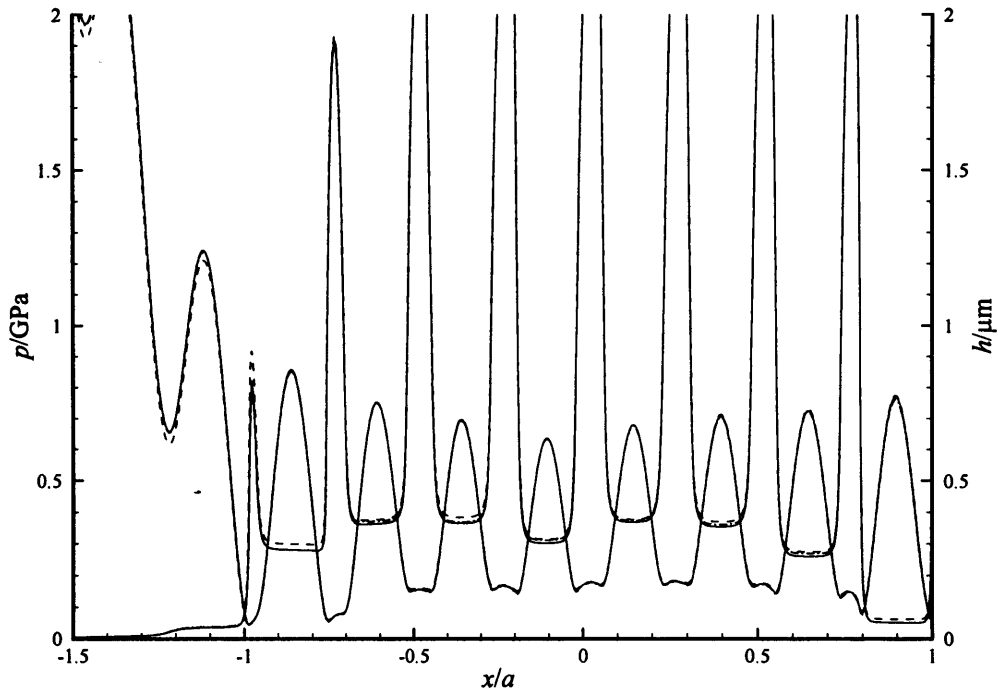

Figure 7.24 – Pressure and film thickness results for the isothermal case for mesh spacings of 400 nodes (solid), 200 nodes (dashed) and 800 nodes (dash-dotted)

All of the results for the different mesh spacing have been obtained using the same time step, the effect of the time that passes between time steps, δt, still needs to be determined. It was seen in figure 7.20 that the Crank-Nicholson approach and the implicit method gave very good agreement. Since the Crank-Nicholson scheme gives more accurate results for larger time step values, it can be concluded that the time step for the implicit method is sufficiently small that no further refinement would be achieved with a smaller δt. In addition, since δt is related to the speed of the surfaces, the time step involved for the case considered here is reasonably large as the surface speeds are small.

It is known that implicit methods can break down when the time that passes between solutions becomes large. Therefore, it follows that δt should be increased to determine how sensitive the method is to δt. To determine the response, δt was progressively increased and compared to previous values for a specific time step. The mesh spacing is kept constant at 400 nodes in the half Hertz dimension to ensure that there is no influence of different mesh spacings. Figure 7.25 shows the results for three different time step

values of δt, $2\times\delta t$ and $3\times\delta t$, where it can be seen that there is very little difference between any of the results.

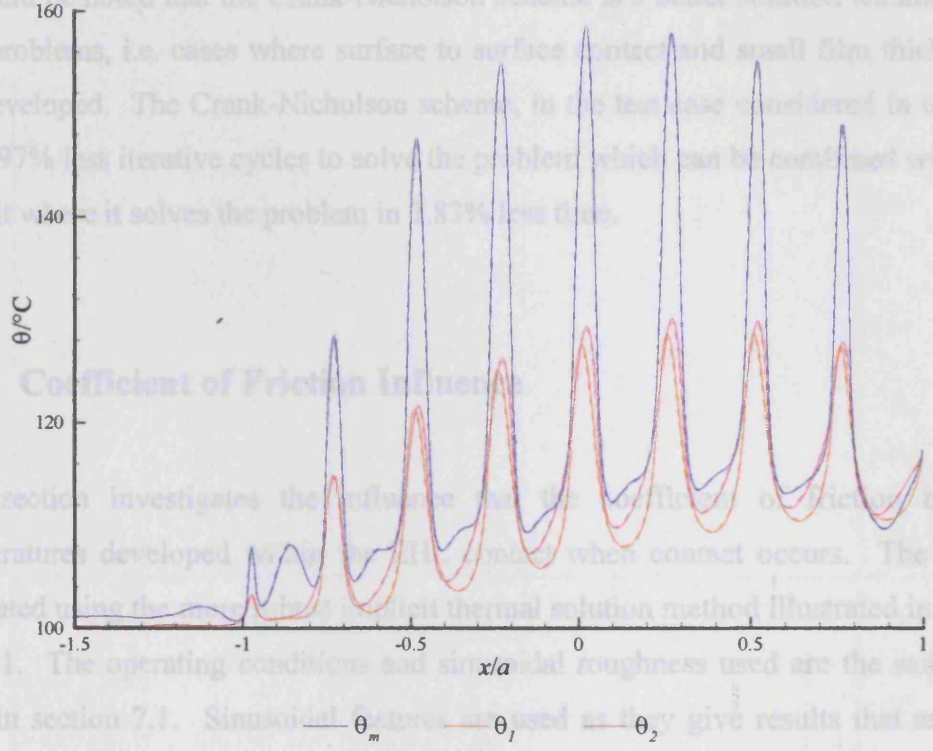

Figure 7.25 – Temperature variation for varying δt; δt (solid), $2\times\delta t$ (dashed) and $3\times\delta t$ (dash-dotted)

The results presented in figures 7.20 to 7.25 show good agreement with one another, thus, it can be concluded that the implicit method adopted is suitable for the solving the thermal contact problem. No matter how many mesh points are used, there is very little variation seen in the results. This indicates that the spatial terms are not unduly sensitive to the mesh spacing and also that the implicit method does not show sensitivity to the coarsening of the mesh. Figure 7.25 also allows the time step to be investigated, where it was seen that the solution method is reasonably insensitive to δt at this scale.

In addition, since the method is developed to examine rough surface EHL, it will be applied when the surface features will not change rapidly from time step to time step. It can also be assumed that the mesh spacing will be reasonably dense to allow a good definition of the rough surface features and the value of δt used will also be small so that

the surfaces do not change rapidly from one solution to the other. Combining this with the information presented here, it is clear that the implicit solution method is a realistic and viable method to solve the thermal contact problem for real rough surfaces.

It should be noted that the Crank-Nicholson scheme is a better solution technique for full film problems, i.e. cases where surface to surface contact and small film thicknesses are not developed. The Crank-Nicholson scheme, in the test case considered in this section, did 7.97% less iterative cycles to solve the problem which can be combined with the time benefit where it solves the problem in 3.83% less time.

7.3 Coefficient of Friction Influence

This section investigates the influence that the coefficient of friction has on the temperatures developed within the EHL contact when contact occurs. The results are generated using the more robust implicit thermal solution method illustrated in figures 7.9 to 7.11. The operating conditions and sinusoidal roughness used are the same as those used in section 7.1. Sinusoidal features are used as they give results that are easier to analyse. Real rough surfaces are unpredictable, making conclusions based on results for these features more problematic.

Comparisons are made between three different coefficients of friction, $\mu = 0.2$, $\mu = 0.4$ and $\mu = 0.6$, these are chosen as the general kinetic coefficient of friction for steel on steel varies from 0.2 to 0.6 [Booser, 1993]. The friction results are evaluated at the beginning of contact, midway through the time that contact occurs, at the end of contact and many time steps after contact has ceased.

Figure 7.26 shows the results for a time step shortly after contact has begun, contact occurs at the exit edge of the first micro contact at $x/a = -1.11$. It can be seen that there is a slight difference between the results in the region that contact occurs. Differences in the mid film temperature and both of the surface temperatures can be seen, where all three show an increase as the coefficient of friction is increased. This is expected, as higher coefficients of friction will increase the amount of heat generated at the interface between

the two solids. It is evident that the heat generated is conducted into the solid in both the x and z directions due to the trend that the surrounding points also exhibit and increase in temperature.

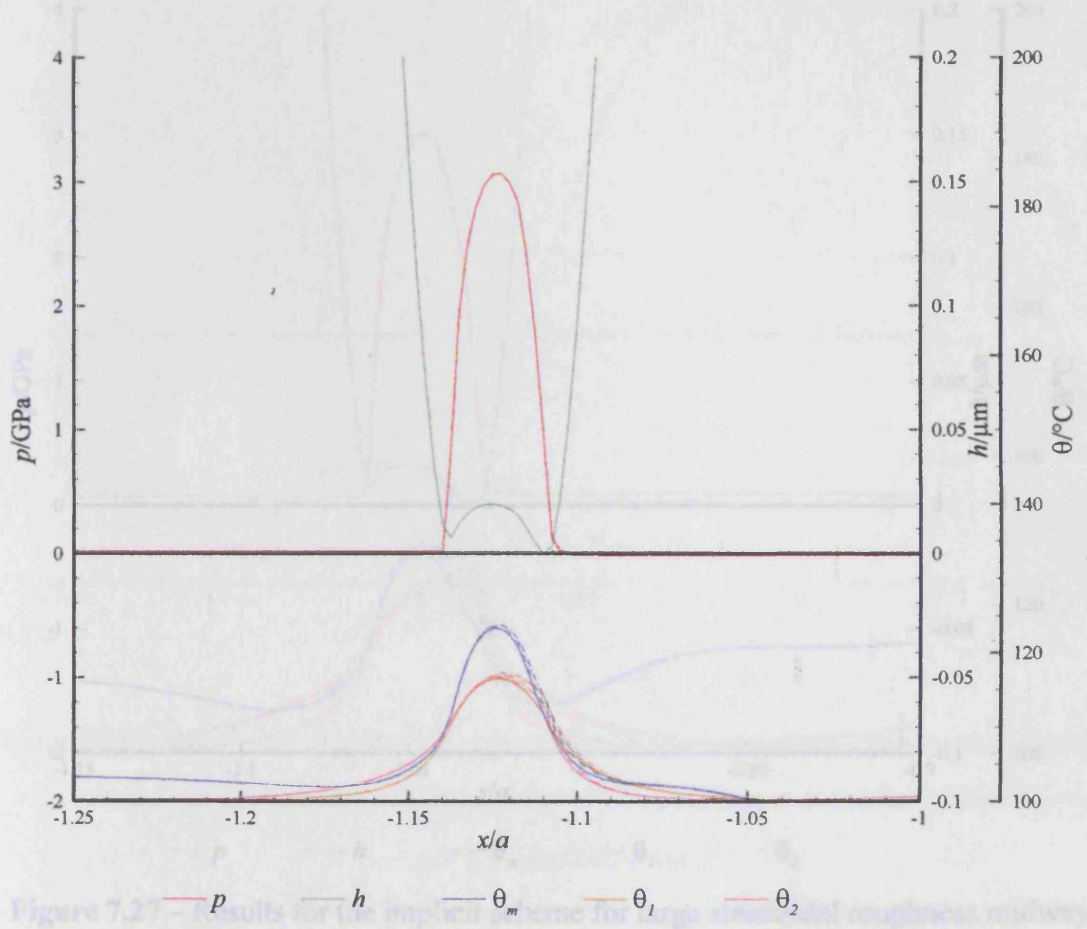

Figure 7.26 – Results for the implicit scheme for large sinusoidal roughness shortly after the beginning of contact for coefficients of friction of $\mu = 0.2$ (solid), $\mu = 0.4$ (dashed) and $\mu = 0.6$ (dash-dotted)

Figure 7.27 shows the results for a time step midway through contact, where contact is occurring at $x/a = -1.03$. The results show a greater variation in temperature for the different coefficients of friction than that shown in figure 7.26. It can also be seen that the differences occur over a greater area than that shown in figure 7.26. For the higher value of the coefficient of friction, $\mu = 0.6$, there is a slight spike in the mid-film and surface temperatures. This is due to the heat generation occurring at one point and the conduction is not sufficient to transport the heat away from the point of generation. There

is slight evidence of the same feature occurring for the $\mu = 0.4$ case, where there is a slight local increase at the contact point.

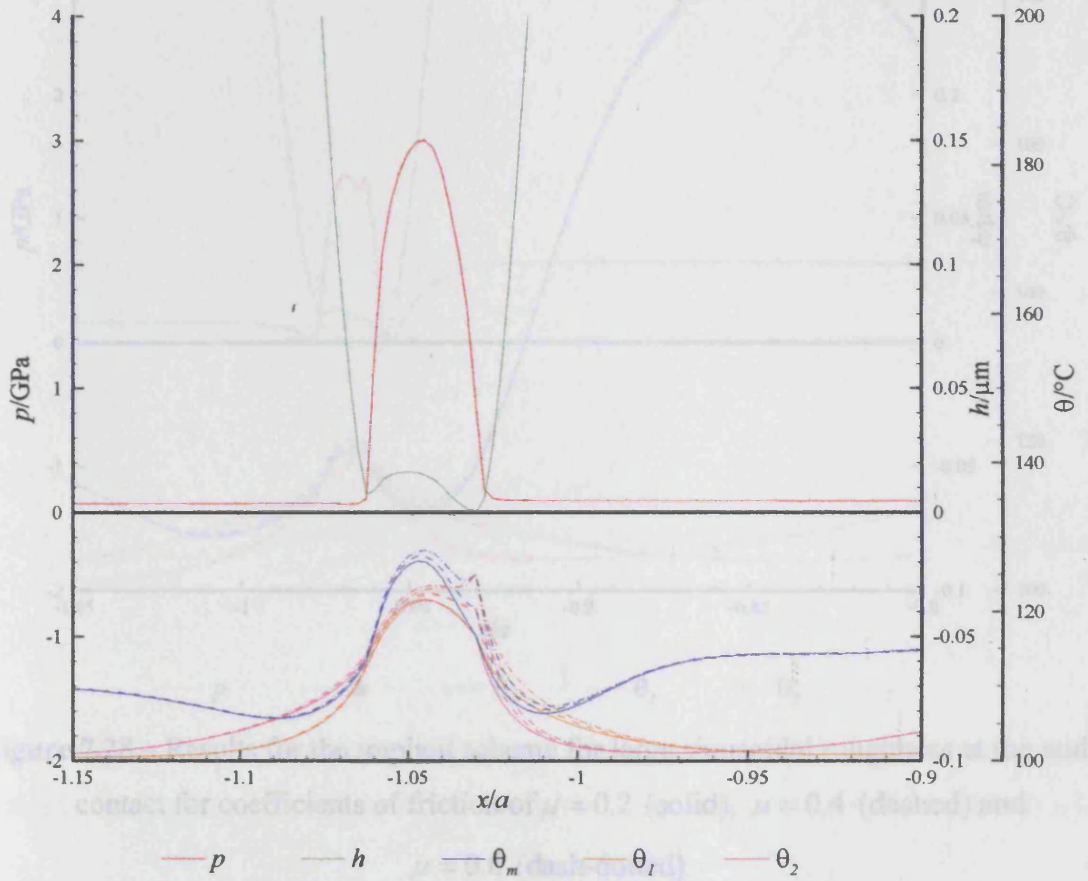

Figure 7.27 – Results for the implicit scheme for large sinusoidal roughness midway through the contacting time for coefficients of friction of $\mu = 0.2$ (solid), $\mu = 0.4$ (dashed) and $\mu = 0.6$ (dash-dotted)

Figure 7.28 shows the results for the end of the contact period, where contact occurs at $x/a = -0.96$. The pressure at the contact has fallen considerably from that occurring in the earlier time steps shown in figures 7.26 and 7.27. Consequently the heat generation at the contact point has also fallen. Most of the heat generated has been dissipated into the solids, but there is still a slight localised temperature peak for the $\mu = 0.6$ case. The higher coefficients of friction still generate higher temperatures and it can be seen that the area over which there is a difference in temperature is extended from that shown in figure 7.27.

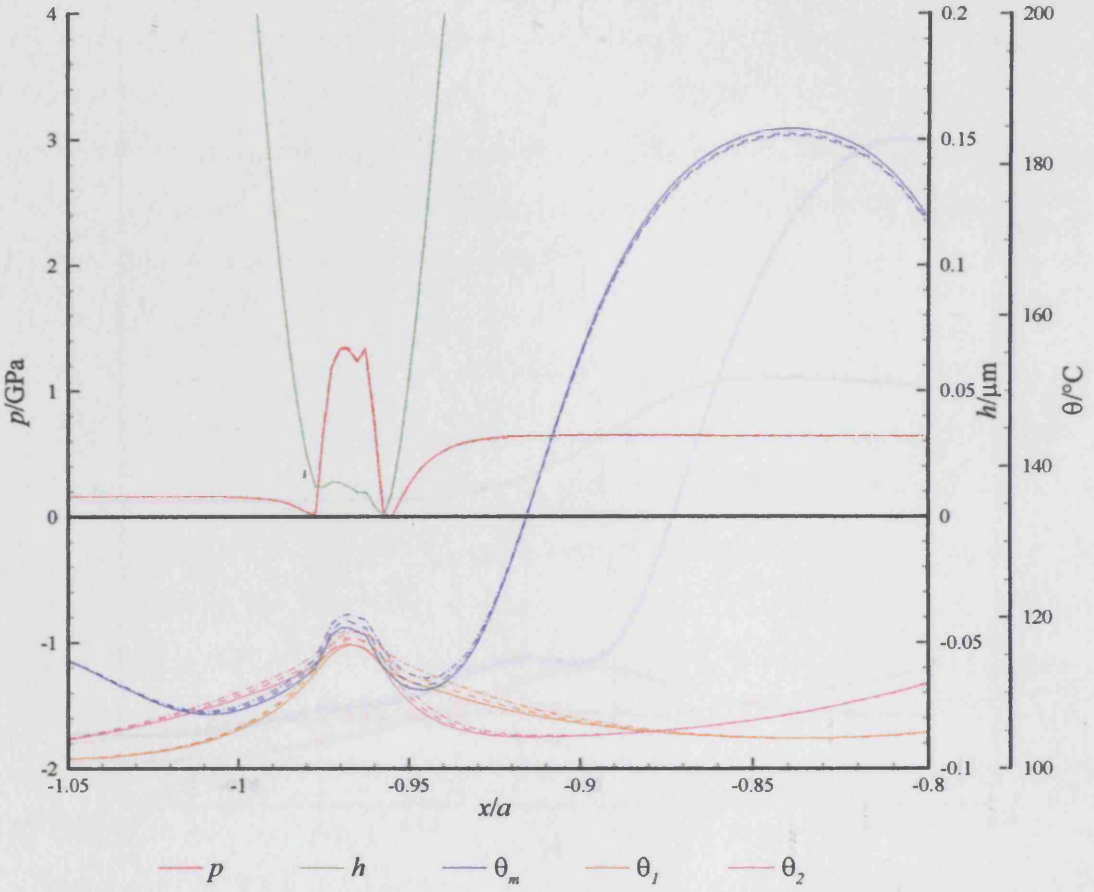

Figure 7.28 – Results for the implicit scheme for large sinusoidal roughness at the end of contact for coefficients of friction of $\mu = 0.2$ (solid), $\mu = 0.4$ (dashed) and $\mu = 0.6$ (dash-dotted)

Figure 7.29 shows the results for 100 time steps after contact had finished. The contact considered in figures 7.26 to 7.28 occurred over some 150 time steps. It can be seen that the effects of the different velocities of the surfaces, through convection, has moved the temperature difference further towards the centre of the contact for the upper faster surface. Due to the heat being conducted into the surfaces the difference in the surface temperatures for the different coefficients of friction decreases. It can also be seen in figure 7.29 that the pressure and temperature have dropped considerably from figures 7.26 to 7.28, this is also accompanied by an increase in the film thickness. This occurs due to the two sinusoidal roughness features moving out of phase with one another.

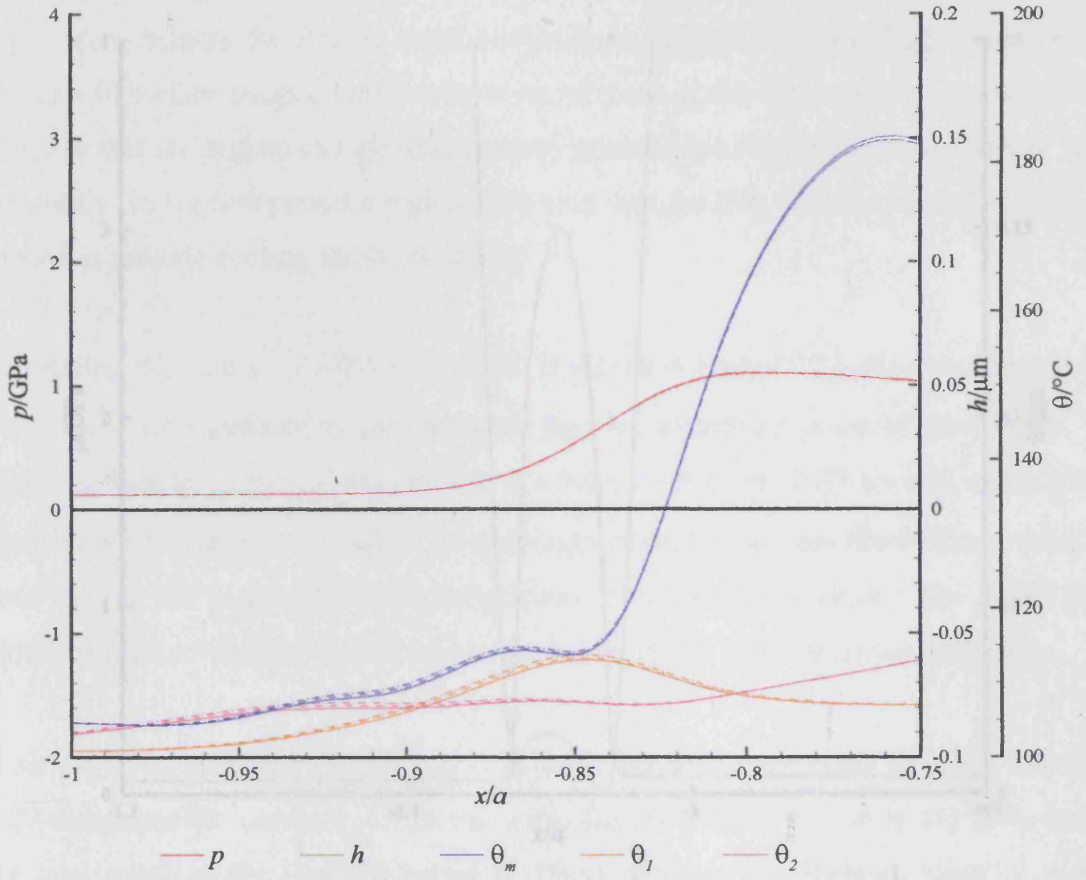

Figure 7.29 – Results for the implicit scheme for large sinusoidal roughness many time steps after contact for coefficients of friction of $\mu = 0.2$ (solid), $\mu = 0.4$ (dashed) and $\mu = 0.6$ (dash-dotted)

To give a better idea of the influence that the coefficient of friction has on the temperature variation within the solids and the fluid, the two dimensional variation of temperature within the film and the solids can be generated. Figure 7.30 shows the results for the $\mu = 0.2$ case midway through contact.

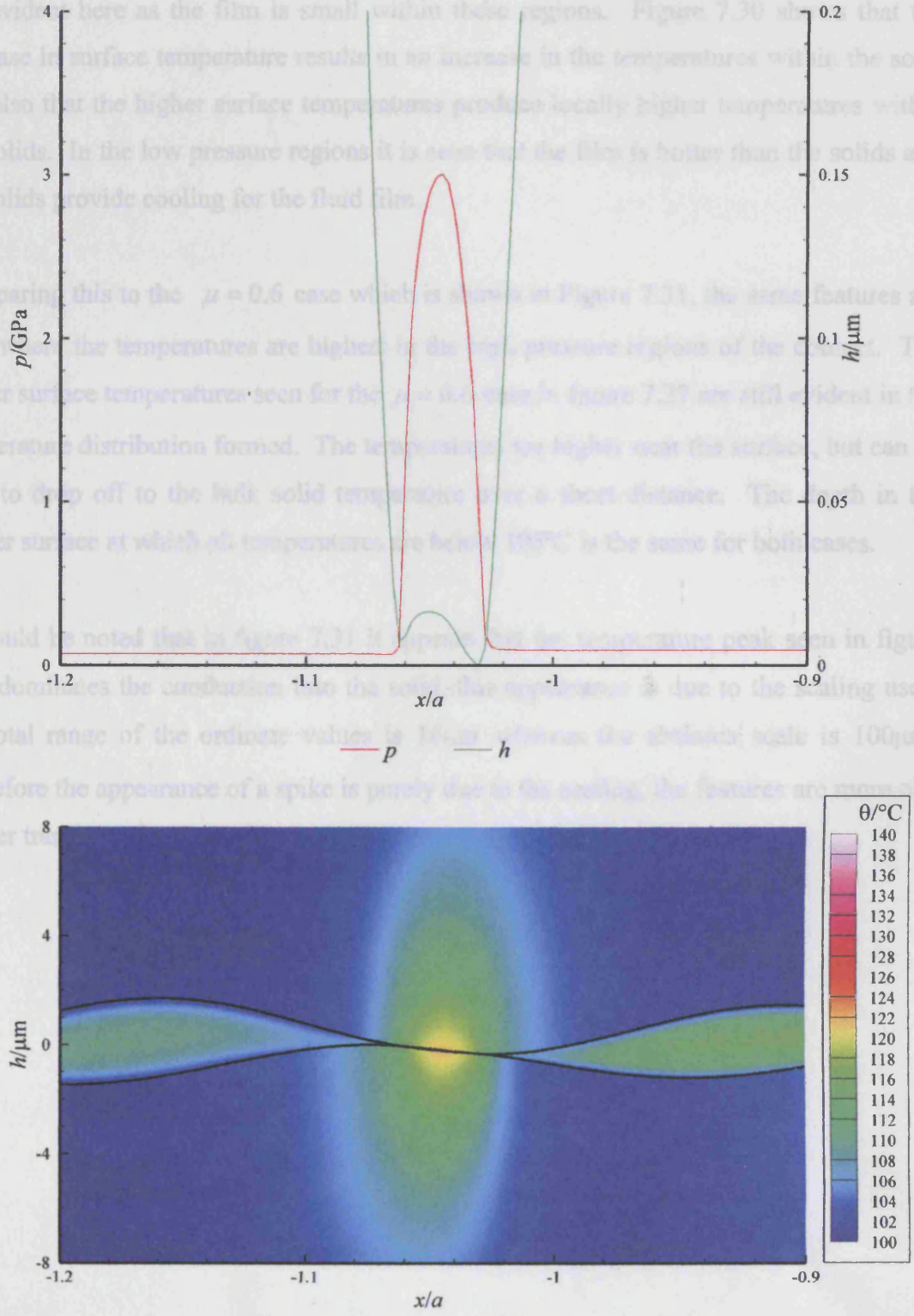

Figure 7.30 – Temperature variation through solid midway through contact for the $\mu = 0.2$ case

It can be seen that the heat generated in the high pressure region of the film is absorbed into the bounding surfaces. The high film temperatures that are seen in figure 7.27 are not evident here as the film is small within these regions. Figure 7.30 shows that the increase in surface temperature results in an increase in the temperatures within the solid and also that the higher surface temperatures produce locally higher temperatures within the solids. In the low pressure regions it is seen that the film is hotter than the solids and the solids provide cooling for the fluid film.

Comparing this to the $\mu = 0.6$ case which is shown in Figure 7.31, the same features are seen where the temperatures are highest in the high pressure regions of the contact. The higher surface temperatures seen for the $\mu = 0.6$ case in figure 7.27 are still evident in the temperature distribution formed. The temperatures are higher near the surface, but can be seen to drop off to the bulk solid temperature over a short distance. The depth in the slower surface at which all temperatures are below 105°C is the same for both cases.

It should be noted that in figure 7.31 it appears that the temperature peak seen in figure 7.27 dominates the conduction into the solid, this appearance is due to the scaling used, the total range of the ordinate values is 16μm whereas the absissca scale is 100μm. Therefore the appearance of a spike is purely due to the scaling, the features are more of a gentler transition than perceived.

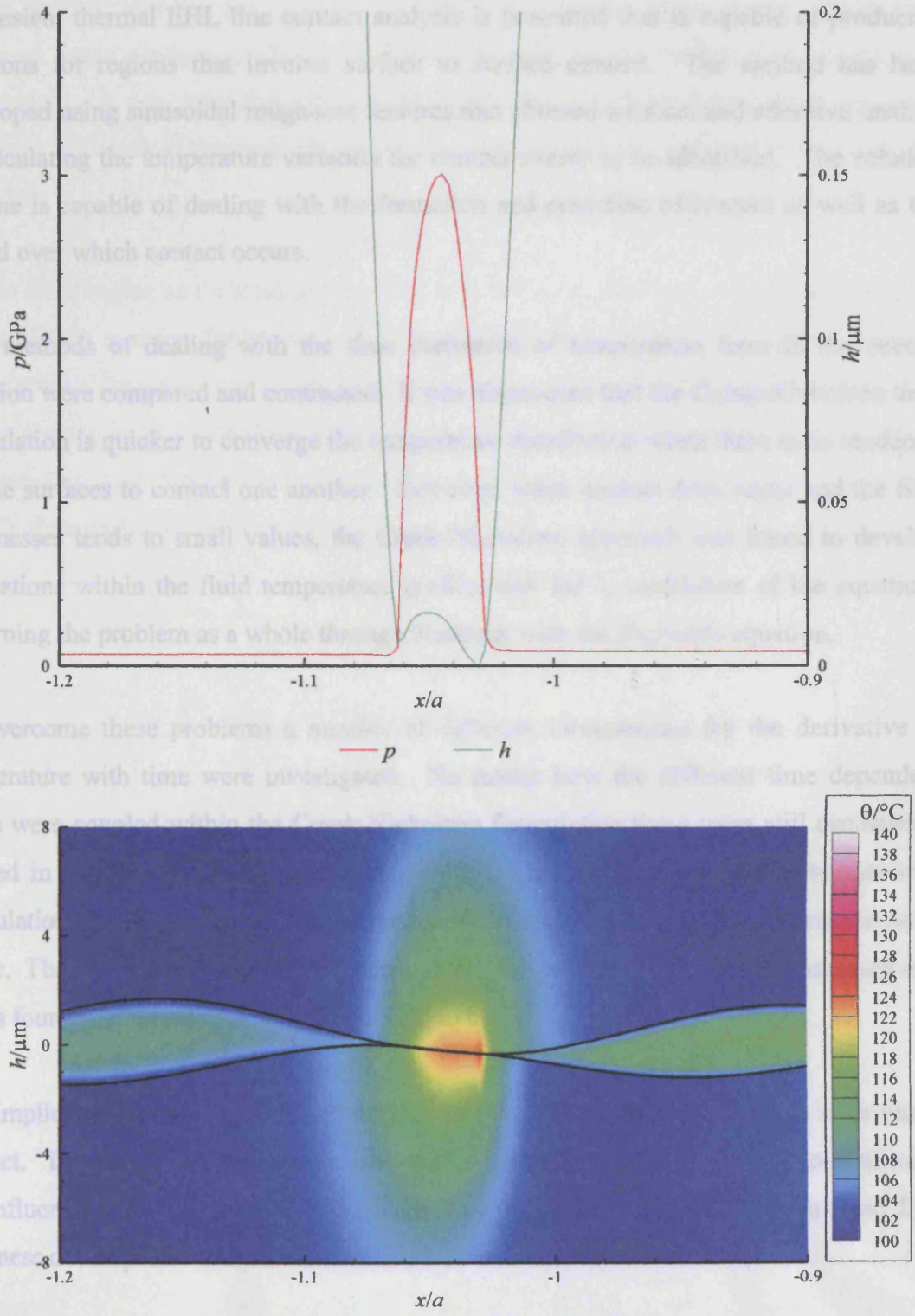

Figure 7.31 – Temperature variation through solid midway through contact for the $\mu = 0.6$ case

7.4 Summary

A transient thermal EHL line contact analysis is presented that is capable of producing solutions for regions that involve surface to surface contact. The method has been developed using sinusoidal roughness features that allowed a robust and effective method of calculating the temperature variation for contact events to be identified. The solution scheme is capable of dealing with the formation and cessation of contact as well as the period over which contact occurs.

Two methods of dealing with the time derivative of temperature term in the energy equation were compared and contrasted. It was discovered that the Crank-Nicholson time formulation is quicker to converge the temperature distribution when there is no tendency for the surfaces to contact one another. However, when contact does occur and the film thicknesses tends to small values, the Crank-Nicholson approach was found to develop oscillations within the fluid temperature profiles that led to oscillation of the equations governing the problem as a whole through feedback with the Reynolds equation.

To overcome these problems a number of different formulations for the derivative of temperature with time were investigated. No matter how the different time dependent terms were coupled within the Crank-Nicholson formulation there were still oscillations formed in time in the film temperature profiles. Other time derivative of temperature formulations were investigated and it was found that an implicit formulation was the most stable. The sensitivity of the implicit method to mesh spacing and time were assessed and it was found that it was relatively insensitive to both.

The implicit method was taken and applied to two large sinusoidal features to promote contact. The coefficient of friction used within the formulation was varied to determine the influence that the coefficient of friction has on the temperatures, pressure and film thickness developed.

The results presented indicate that the coefficient of friction affects the temperatures that are developed and that the higher the coefficient of friction the greater the temperatures generated. The difference that results from different coefficients of friction is localised

both at the time of direct contact and subsequently. The temperature difference between different coefficients of friction is reasonably small, of the order of a few degrees. These results indicate that the coefficient of friction has a modest effect on the temperature produced, large changes in the coefficient of friction do not lead to large changes in temperature. It was seen that whilst contact occurs the differences in temperature are moderate, but these differences diminish with time as the heat is conducted away into the surfaces.

This is encouraging as it means that contact does not generate huge temperature increases that are likely to cause further contact due to the thinning of the oil due to the increased temperature. The effect that contact has is to increase the local temperature, of the order of ten degrees or less, and does not give rise to a dramatic temperature spike. There is evidence of a slight temperature spike for the highest coefficient of friction considered but is only of the order of 3 or 4 degrees.

Chapter 8

EHL Simulation of the Gear Meshing Cycle

8.0 Introduction

Section 2 illustrated the manner in which spur gears contact. In this chapter this theory is taken further and an approach developed that allows computational modelling of such contacts. The main difficulty in modelling gear meshes is that the surface speeds, radius of relative curvature and load change throughout the mesh and need to be accounted for.

This chapter outlines how gear meshes can be modelled, as well as introducing an approach that allows the actual measured roughness of the gear teeth to be incorporated. Results are then presented for the smooth problem as well as the rough surface problem for different meshes.

8.1 Operating Conditions

The contacting point on each face moves at the same angular speed as the gear itself, the rotation is centred at the point of tangency of the line of action with the base circle. This was illustrated in figure 2.5 and allows the velocity tangential to the surfaces at the contacting point for the two surfaces to be calculated from:

$$u_1 = (r_{b1}\tan\psi + s)\omega_1 \qquad \text{and} \qquad u_2 = (r_{b2}\tan\psi - s)\omega_2$$

The starting point for the solution process is to determine the first and last contact points, s_f and s_l respectively. These allow the length of contact, Δs, to be determined as:

$$\Delta s = s_l - s_f$$

Using the velocity at which the contact point moves the time taken, t, for the contact to progress from first to last point is given from:

$$t = \frac{\Delta s}{v_{contact}}$$

This gives the time increment, δt, between each point that makes up the meshing cycle as:

$$\delta t = \frac{t}{no\ steps}$$

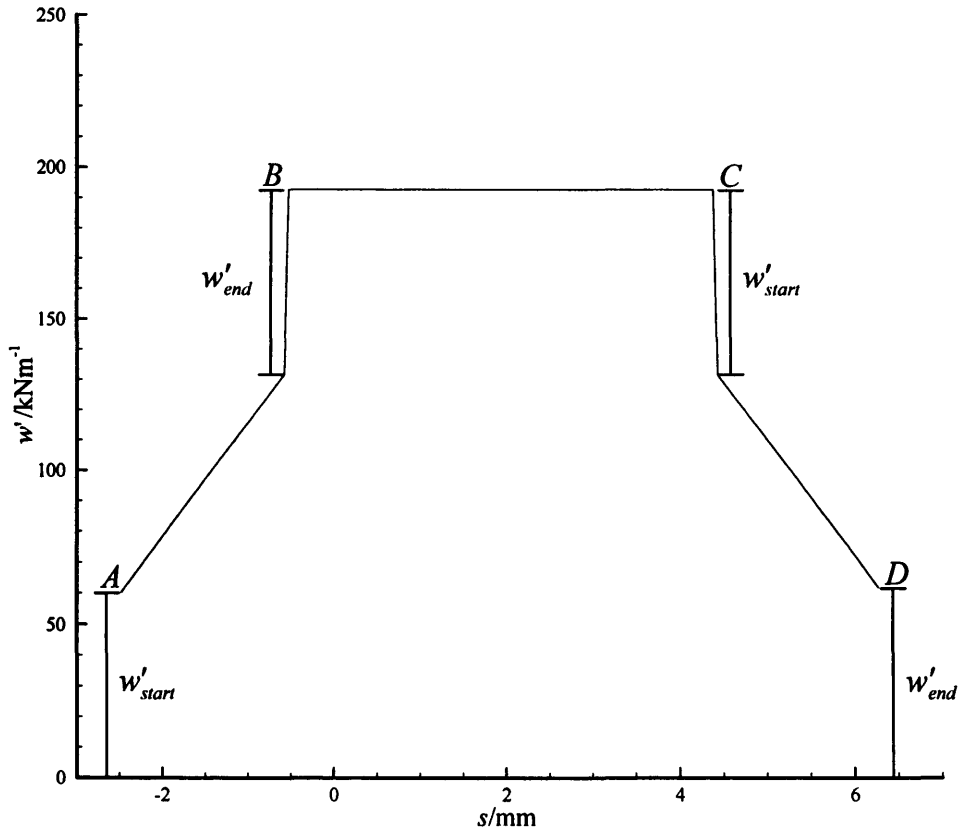

Figure 8.1 – Load variation for a gear meshing cycle with change points at positions *A*, *B*, *C* and *D*

The main difficulty of the gear meshing problem is determining the load that is applied to the gear teeth through the meshing cycle. The load variation for a typical gear mesh is shown in Figure 8.1. As gear teeth come into and out of contact there are very large increases and decreases in the load carried by each individual gear tooth. The points during the meshing cycle that gear teeth come into and out of contact are termed the

change points. The first point to be analysed is such a change point, where the teeth under consideration first come into contact denoted *A* in figure 8.1. After the first point in the meshing cycle, the load on the gear teeth rises until the first change point within the meshing cycle is reached at *B*. At this location, the meshing between the preceding pair of teeth comes to an end creating a sudden increase in load. The load increases to the maximum load at this point and continues at the maximum load until the second change point within the meshing cycle is reached at *C*. At the second change point, the following pair of teeth begin to mesh, which reduce the load carried by the modelled pair. As the meshing cycle continues to progress, the load continues to decrease until the end of the meshing cycle at *D*, where a final value of load is carried.

It can be seen that the load that the gear teeth carry initially, w'_{start}, is the same value as the load decrease at the second change point. Similarly, the end load, w'_{end}, is the same value as the load increase at the first change point. This must be the case, as otherwise the load is not constant between the gears. The maximum load, w'_{max}, may not be achieved at the pitch point, it is all dependent on the location of the change points.

The distance in s to the first change point, s_{c1}, occurring at *B* can be determined from:

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

Similarly the second distance to the change point, s_{c2}, occurring at *C* is given by:

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

The terms n_1 and n_2 are the number of teeth on the driving and driven teeth respectively.

Once the time taken to complete the meshing cycle, contact path length and load variation are determined the problem can be initiated. The EHL computing mesh is non-dimensionally scaled to the Hertz dimension, a, that corresponds to the pitch conditions. The value of R used to calculate a is determined using simplified cylindrical geometry.

8.2 Solution Techniques

The meshing cycle solution techniques can be split into two, the rough surface technique and the smooth surface approach. The solution to the smooth surface problem is necessary before the rough surface problem can be considered. This is to allow the variation in the first node film thickness, h_1, for the smooth case to be used as an input for the rough solution. Both solution techniques use the same equations and solution methods established in Chapters 2 and 4 and combine this with a technique for changing the operating conditions as the meshing cycle progresses.

One component that is common to both techniques is the way that the geometry of the gear teeth is taken into account. There are two types of geometry that can be used, one is the equivalent roller approximation and the other is the true involute form. The equations that are used for the equivalent rollers give the radii of the two rollers that can be used in equation (2.1) to determine the undeformed geometry. The involute equations are somewhat more complex.

There are four equations that define the involute shape, comprising of two equations for the driven gear, equations (2.7) and (2.8), and two equations for the driving gear, equations (2.9) and (2.10). There are only two unknowns in each of these pairs of equations, γ_1 and z_1 for the driving gear and γ_2 and z_2 for the driven gear. These equations give x and z in terms of the reference angles γ_1 and γ_2. What is required for the model are coordinate pairs (x, z) for each surface at each computing mesh point x_i. Unfortunately, γ_1 and γ_2 cannot be obtained analytically from equations (2.7) and (2.9) which are transcendental. They can be determined numerically and a Newton-Rhapson technique is used to find γ_1 and γ_2 for each value of x required. These angles can then be used in equations (2.8) and (2.10) to determine the corresponding values of z_1 and z_2.

8.2.1 Smooth Problem Solution Method

The problem is initiated using the geometry, speed and loading on the gears to calculate the contact path, time for the meshing cycle and load variation. The mesh is also scaled to the Hertz dimension at the pitch point, this is calculated using the actual load and relative radius of curvature at the pitch point.

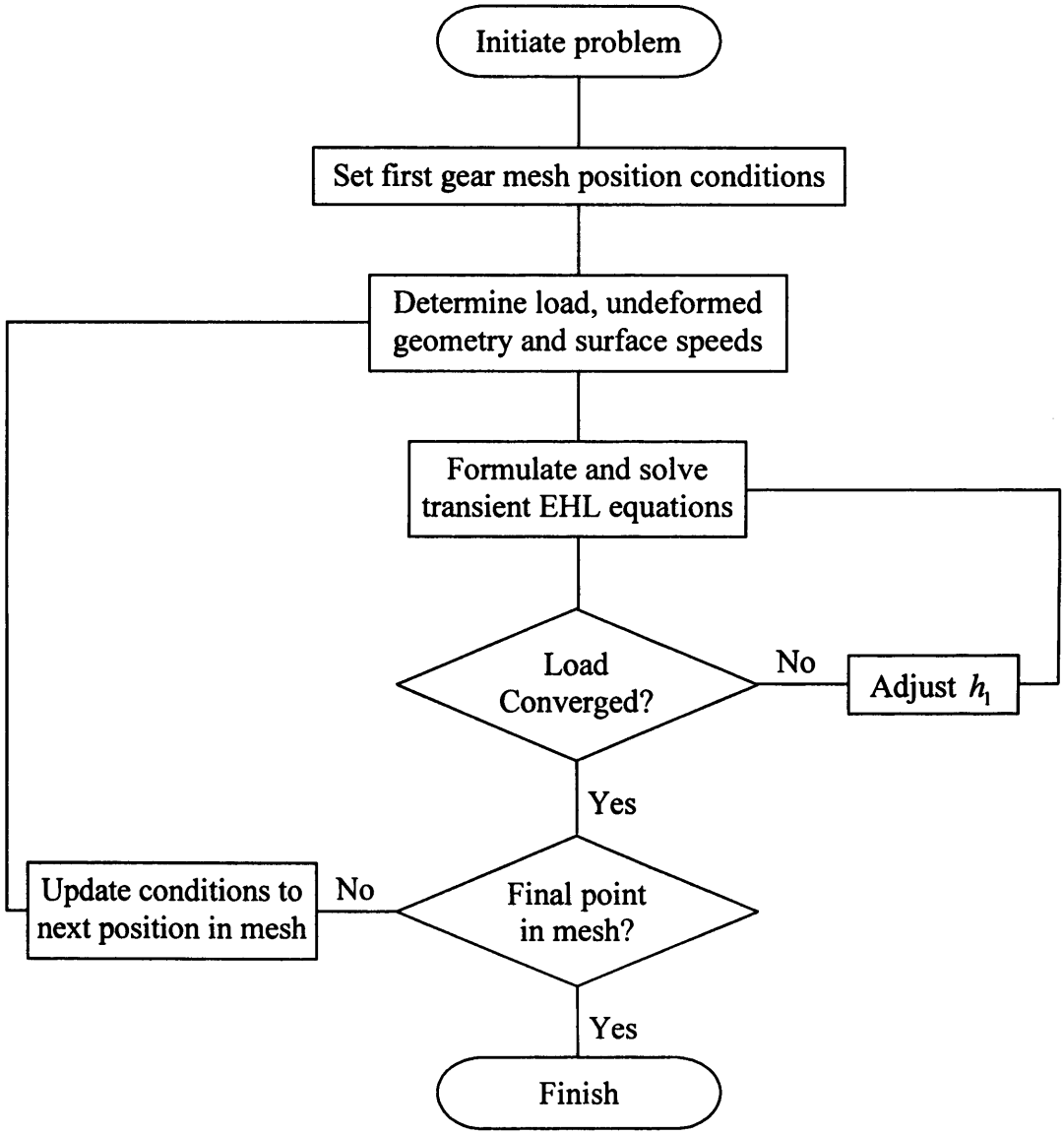

Figure 8.2 – Flow chart of solution process

The operating conditions are set to the conditions at the first gear mesh position. The load, speed and geometry are determined and the equations formulated. The equations do not include any time dependant terms, the first mesh position solution is a steady state solution and is used as the initial condition for the subsequent transient analysis. Once

the solution to the first mesh position is achieved, the next point in the meshing cycle is solved using the full transient equations. When moving to the next position in the meshing cycle the undeformed geometry, speed and load are all re-established.

The load that is carried at the positions in the meshing cycle analysed are already known from the load variation as illustrated in figure 8.1. The first node film thickness, h_1, essentially controls the approach distance of the components and thus the load developed. This is changed over the meshing cycle to ensure that the load returned from the solved equations is converged to the desired value. This approach is adopted and repeated for all other solution positions in the meshing cycle. This process is shown in Figure 8.2 and it is repeated until a solution is obtained for all positions in the meshing cycle.

8.2.2 Rough Surface Problem Solution Method

Once the smooth surface problem has been solved it gives some idea of the change necessary to h_1 to attain the specified load through the meshing cycle. This is used as an input to the rough surface problem, this is necessary as load seeking with rough surfaces is unrealistic and is very difficult to achieve. Due to the interaction of the asperities, the surfaces would have to be brought closer together and pushed further apart at each time step during the meshing cycle. This very rapid bulk movement of the teeth will generate squeeze film effects which will change the response in the contact area and possibly lead to instabilities in the equations developed due to the constant change in h_1. It is difficult to quantify the amount of change in h_1 necessary to attain the correct load due to the changes in the surface features, some points may require greater changes in h_1 than others. In addition, the gear teeth are very large elastic bodies which deflect at the root when they are brought into contact. Thus any increase or decrease in the load generated will be accommodated by small changes in the deflection at the root of the gear teeth.

The strategy developed is thus to vary h_1 over the meshing cycle in line with the variation determined in order to achieve the specified load variation in the smooth surface case. The values of h_1 used in the rough surface simulations are not identical to the values

obtained from the smooth surface solution, instead they are altered slightly due to the effect that the asperity interaction has on the load. The variation of h_1 for the load variation shown in figure 8.1 is shown in Figure 8.3.

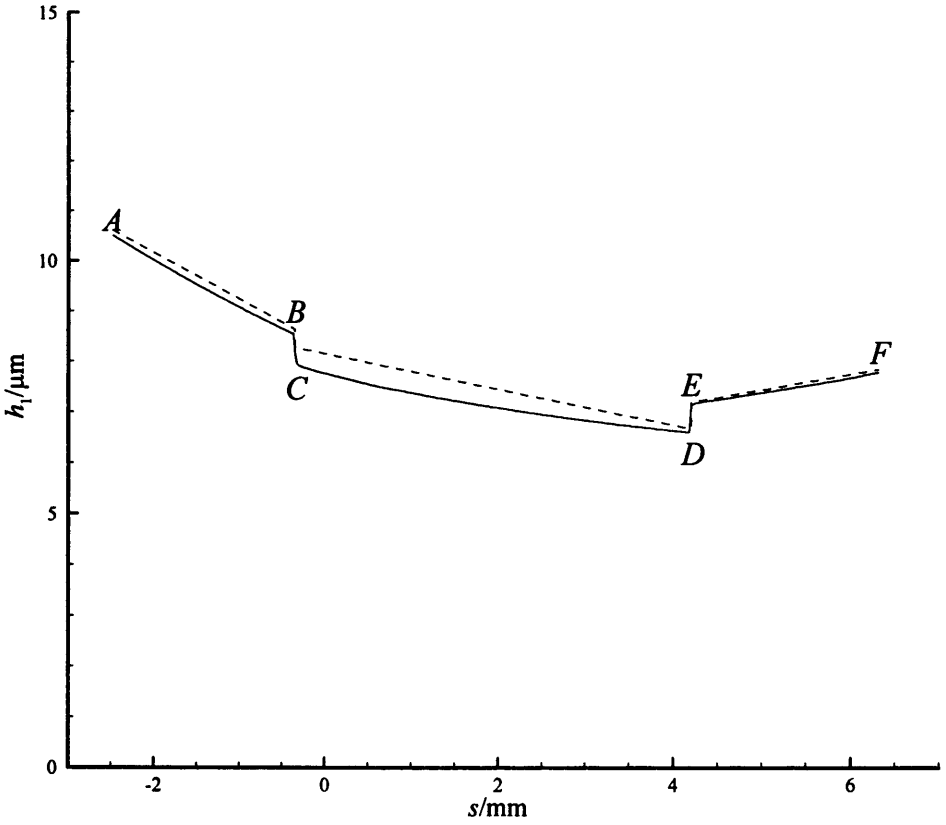

Figure 8.3 – Variation of first node film thickness over meshing cycle for smooth surfaces (solid) and rough surfaces (dashed)

The variation in h_1 is represented for the rough surface solution by defining h_1 at six points $A - F$. Points B and C are taken for the same value of s, relating to the first change point in the meshing cycle. Similarly, points D and E are taken to represent the change in the load that occurs at the second change point. The values of h_1 at the other points during the meshing cycle are determined by linear interpolation of the defined points.

The variation of h_1 for a rough surface solution is also shown in figure 8.3. There is a slight positive offset in the variation in h_1 to account for the interaction of the asperities. A similar shape is retained for the rough surface result, there is essentially an offset

applied to the smooth surface result. The value of the offset that is required for the rough surface results to ensure that the load follows the desired load is obtained by trial and error over several complete analyses.

The most difficult part of the rough surface solution is determining which roughness points to use in the EHL mesh during the meshing cycle. The roughness profiles that are used as inputs for the analysis are deviations from a mean line and contain no information as to which part of the gear tooth they correspond. An approach was developed that allowed the points in the roughness data to be linked to the point on the tooth that they were taken from. The way in which the roughness is linked to a part of the tooth is discussed first, then the manner in which the roughness is introduced back onto the surface in the simulation is covered.

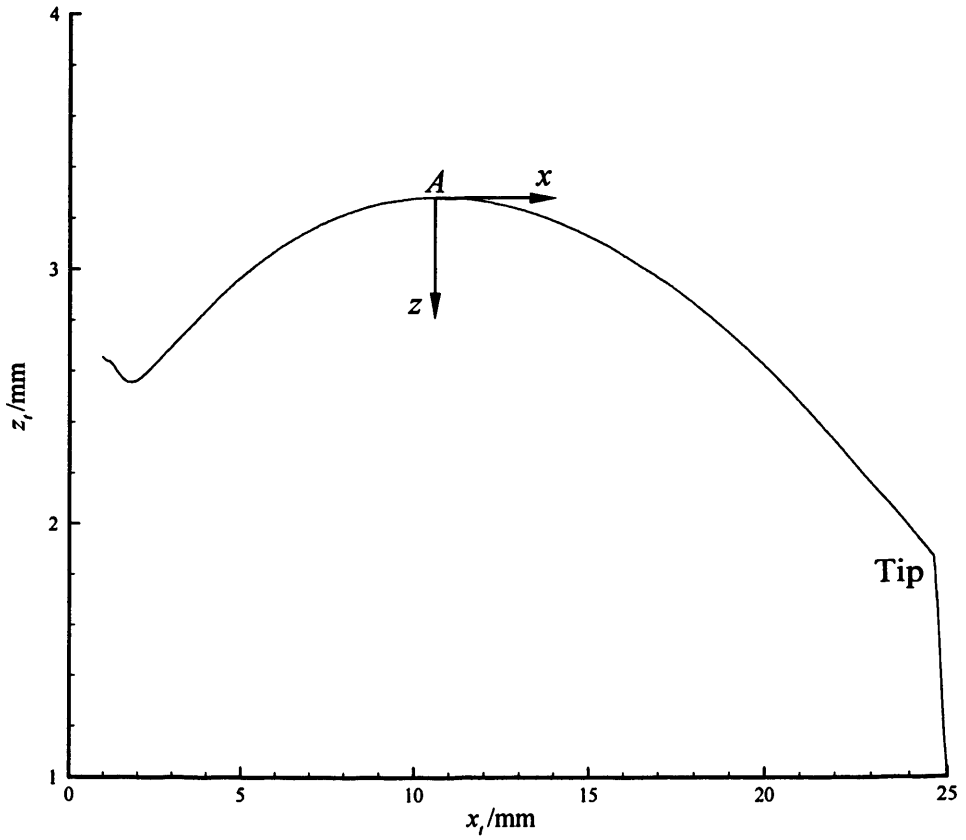

Figure 8.4 – Talysurf profile of gear tooth illustrating tangential point and involute axis

First consider Figure 8.4 showing a trace of a gear tooth form taken from a Talysurf. This instrument has sufficient dynamic range to measure the shape of the tooth profile as well

as the surface roughness. The root and tip features are clearly visible as is the active profile (Note that the shape of the tooth is distorted in the figure because of the different magnitudes in the two directions). The active profile is the part of the gear tooth that is generated from the involute and is where the contact between the gear teeth occurs. The Talysurf axes are denoted by x_t and z_t. The alignment of the axes to the gear tooth depends on the manner in which the component is mounted in the instrument.

The Talysurf measurements are taken perpendicular to a datum that is tangential to one point on the involute, the tangential point for the profile presented in figure 8.4 is denoted *A*. Equations were presented in section 2.2.2 that allowed the distance to any point on the involute along a tangential datum to be determined. Fitting a mathematical involute to the Talysurf profile allows the involute angle that defines the tangential point, *A*, to be determined. This is denoted γ_{ref} and can be used together with the involute equations so as to determine the distance along the roughness profile to any point of interest on the involute curve.

Equations were obtained in section 2.2.2 for the involute curve in terms of Cartesian coordinates *x* and *z*, where *x* was measured along the tangent to the curve at the nominal contact point. Since the measured tooth profile, illustrated in figure 8.4, starts at the root and progresses toward the tip in the positive *x* direction then a similar involute can be generated from equations (2.7) and (2.8), which are re-written below:

$$\begin{aligned} x = {} & r_{b2}\cos\gamma_{c2}(\cos\gamma_2 + \gamma_2\sin\gamma_2 - \cos\gamma_{c2} - \gamma_{c2}\sin\gamma_{c2}) \\ & + r_{b2}\sin\gamma_{c2}(\sin\gamma_2 - \gamma_2\cos\gamma_2 - \sin\gamma_{c2} + \gamma_{c2}\cos\gamma_{c2}) \end{aligned}$$

$$\begin{aligned} z = {} & r_{b2}\sin\gamma_{c2}(\cos\gamma_{c2} + \gamma_{c2}\sin\gamma_{c2} - \cos\gamma_2 - \gamma_2\sin\gamma_2) \\ & + r_{b2}\cos\gamma_{c2}(\sin\gamma_2 - \gamma_2\cos\gamma_2 - \sin\gamma_{c2} + \gamma_{c2}\cos\gamma_{c2}) \end{aligned}$$

To determine the angle at the tangential point, the two Cartesian coordinates have to be aligned parallel with the Talysurf coordinates. The Cartesian coordinates for the mathematical involute have their origin at the tangential point, aligning them with the coordinates on which the Talysurf readings were taken produces the arrangement shown in figure 8.4. If the two coordinate systems are aligned, then the resulting shape of the involute from the mathematical approach will give good agreement with that measured using the Talysurf.

The above equations can be manipulated in terms of the angle, γ_{ref}, that corresponds to the reference point, A. Replacing γ_{c2} with γ_{ref} allows the involute shape to be generated relative to the datum tangential to this point on the involute. The other terms used in the equations are the angle on the involute at which its Cartesian coordinates are needed, γ_2, and the base radius, r_{b2}. The numerical subscript on both terms are used to distinguish between the driven and driving gear, they can be omitted for convenience allowing the equations to be written in the form:

$$\begin{aligned} x &= r_b \cos\gamma_{ref}\left(\cos\gamma + \gamma\sin\gamma - \cos\gamma_{ref} - \gamma_{ref}\sin\gamma_{ref}\right) \\ &\quad + r_b \sin\gamma_{ref}\left(\sin\gamma - \gamma\cos\gamma - \sin\gamma_{ref} + \gamma_{ref}\cos\gamma_{ref}\right) \\ z &= r_b \sin\gamma_{ref}\left(\cos\gamma_{ref} + \gamma_{ref}\sin\gamma_{ref} - \cos\gamma - \gamma\sin\gamma\right) \\ &\quad + r_b \cos\gamma_{ref}\left(\sin\gamma - \gamma\cos\gamma - \sin\gamma_{ref} + \gamma_{ref}\cos\gamma_{ref}\right) \end{aligned}$$

It can be seen from figure 8.4 that the equations are formulated so that the distances are relative to the tangential point, i.e. $x = 0$ when $\gamma = \gamma_{ref}$. However, the Talysurf axis, x_t, is zero at the first point in the Talysurf profile. An offset needs to be used that gives the distance from the first point in the roughness to the tangential point. Terming the distance between the first point and the tangential point as x_{shift}, then the mathematical involute can be expressed in terms of the Talysurf coordinate system x_t by:

$$\begin{aligned} x_t &= x_{shift} + r_b \cos\gamma_{ref}\left(\cos\gamma + \gamma\sin\gamma - \cos\gamma_{ref} - \gamma_{ref}\sin\gamma_{ref}\right) \\ &\quad + r_b \sin\gamma_{ref}\left(\sin\gamma - \gamma\cos\gamma - \sin\gamma_{ref} + \gamma_{ref}\cos\gamma_{ref}\right) \end{aligned} \tag{8.1}$$

The z axis is orientated opposite to the z_t axis, this can be easily changed by changing the sign of z. Similarly to the x axis, an offset is used that corresponds to the height that the tangential point, A, is above the datum. Terming this offset z_{shift}, the mathematical involute can be expressed on the z_t axis by:

$$\begin{aligned} z_t &= z_{shift} - r_b \sin\gamma_{ref}\left(\cos\gamma_{ref} + \gamma_{ref}\sin\gamma_{ref} - \cos\gamma - \gamma\sin\gamma\right) \\ &\quad - r_b \cos\gamma_{ref}\left(\sin\gamma - \gamma\cos\gamma - \sin\gamma_{ref} + \gamma_{ref}\cos\gamma_{ref}\right) \end{aligned} \tag{8.2}$$

Equations (8.1) and (8.2) can be used to fit a theoretical involute onto the roughness profile taken from the Talysurf. The dimensions x_{shift} and z_{shift} are obtained from the

coordinates of the 'highest point' on the measured profile in the x_t, z_t coordinate system. This is determined with the high frequency roughness suppressed, effectively by drawing the profile on the scale of figure 8.4 so that the form is revealed. To obtain the value of γ_{ref}, equation (8.1) is used together with the estimated positions of the root and tip features from figure 8.4. They cannot be accurately located as the scale of the data precludes the exact points from being determined.

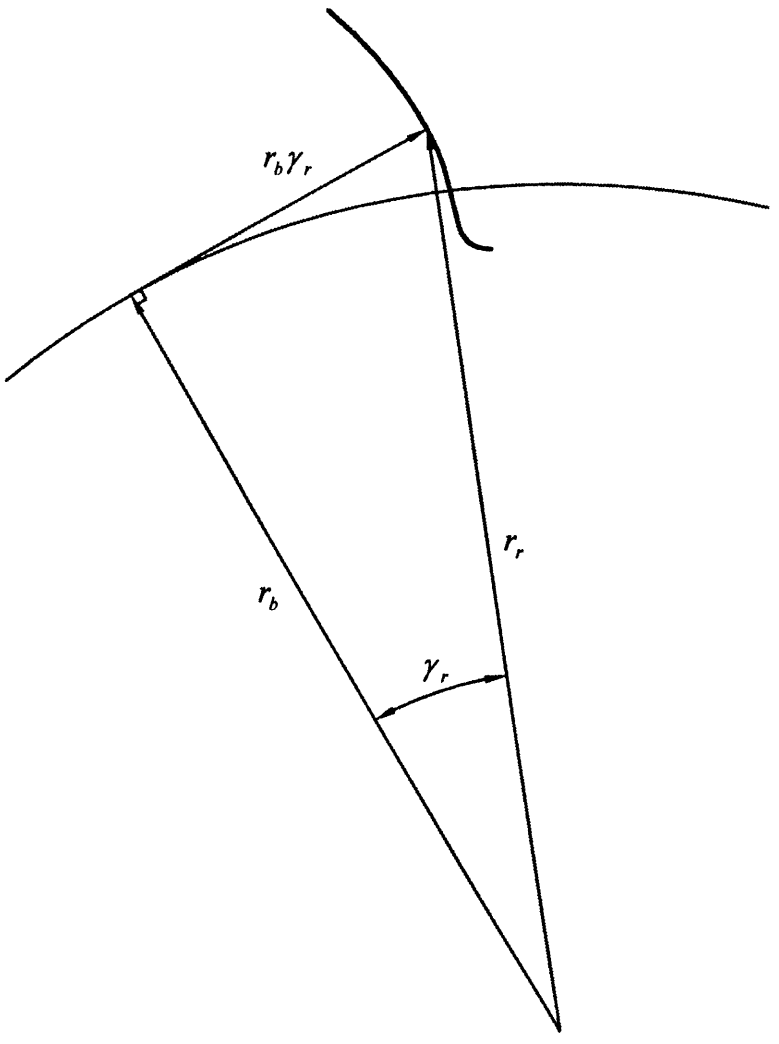

Figure 8.5 – Component dimension that are used to determine involute angle at the root

The involute angle that corresponds to the root feature can be obtained by trigonometry. Consider figure 8.5, which shows the known root radius, r_r, the angle on the involute

corresponding to the root feature, γ_r, and the base radius, r_b. The three terms are related by:

$$r_r^{\,2} = r_b^{\,2}\gamma_r^{\,2} + r_b^{\,2}$$

This can be re-arranged to give:

$$\gamma_r = \sqrt{\left(\frac{r_r}{r_b}\right)^2 - 1}$$

Using this value in equation (8.1) the distance to the root, x_r, can be determined as:

$$\begin{aligned} x_r = x_{shift} &+ r_b \cos\gamma_{ref}\left(\cos\gamma_r + \gamma_r \sin\gamma_r - \cos\gamma_{ref} - \gamma_{ref}\sin\gamma_{ref}\right) \\ &+ r_b \sin\gamma_{ref}\left(\sin\gamma_r - \gamma_r\cos\gamma_r - \sin\gamma_{ref} + \gamma_{ref}\cos\gamma_{ref}\right) \end{aligned} \tag{8.3}$$

Similarly the tip radius, r_t, can be used to determine the tip angle, γ_t, as:

$$\gamma_t = \sqrt{\left(\frac{r_t}{r_b}\right)^2 - 1}$$

Which gives the distance to the tip, x_t, as:

$$\begin{aligned} x_t = x_{shift} &+ r_b \cos\gamma_{ref}\left(\cos\gamma_t + \gamma_t \sin\gamma_t - \cos\gamma_{ref} - \gamma_{ref}\sin\gamma_{ref}\right) \\ &+ r_b \sin\gamma_{ref}\left(\sin\gamma_t - \gamma_t\cos\gamma_t - \sin\gamma_{ref} + \gamma_{ref}\cos\gamma_{ref}\right) \end{aligned} \tag{8.4}$$

The relative distance between the root and tip features, *dist*, can be determined as:

$$dist = x_t - x_r$$

Substituting equations (8.3) and (8.4) into the above gives:

$$\begin{aligned} dist = \; & r_b \cos\gamma_{ref}\left(\cos\gamma_t + \gamma_t \sin\gamma_t - \cos\gamma_r - \gamma_r\sin\gamma_r\right) \\ & + r_b \sin\gamma_{ref}\left(\sin\gamma_t - \gamma_t\cos\gamma_t - \sin\gamma_r + \gamma_r\cos\gamma_r\right) \end{aligned}$$

This can be simplified by substituting for the known bracketed terms:

$$dist = r_b\left(A\cos\gamma_{ref} + B\sin\gamma_{ref}\right) \tag{8.5}$$

Where:

$$A = \cos\gamma_t + \gamma_t\sin\gamma_t - \cos\gamma_r - \gamma_r\sin\gamma_r \quad \text{and} \quad B = \sin\gamma_t - \gamma_t\cos\gamma_t - \sin\gamma_r + \gamma_r\cos\gamma_r$$

Using trigonometric substitutions, equation (8.5) becomes:

$$\gamma_{ref} = \sin^{-1}\sqrt{\frac{dist^2 - r_b^2 A^2}{r_b^2\left(B^2 - A^2\right)}}$$

The dimension *dist* is taken from the tooth profile and used in the above equation to give γ_{ref}. The dimension *dist* can only be determined approximately from figure 8.4, thus can only give an approximation to γ_{ref}. The magnitude of γ_{ref}, x_{shift} and z_{shift} can be manually adjusted to obtain the best fit to the measured profile.

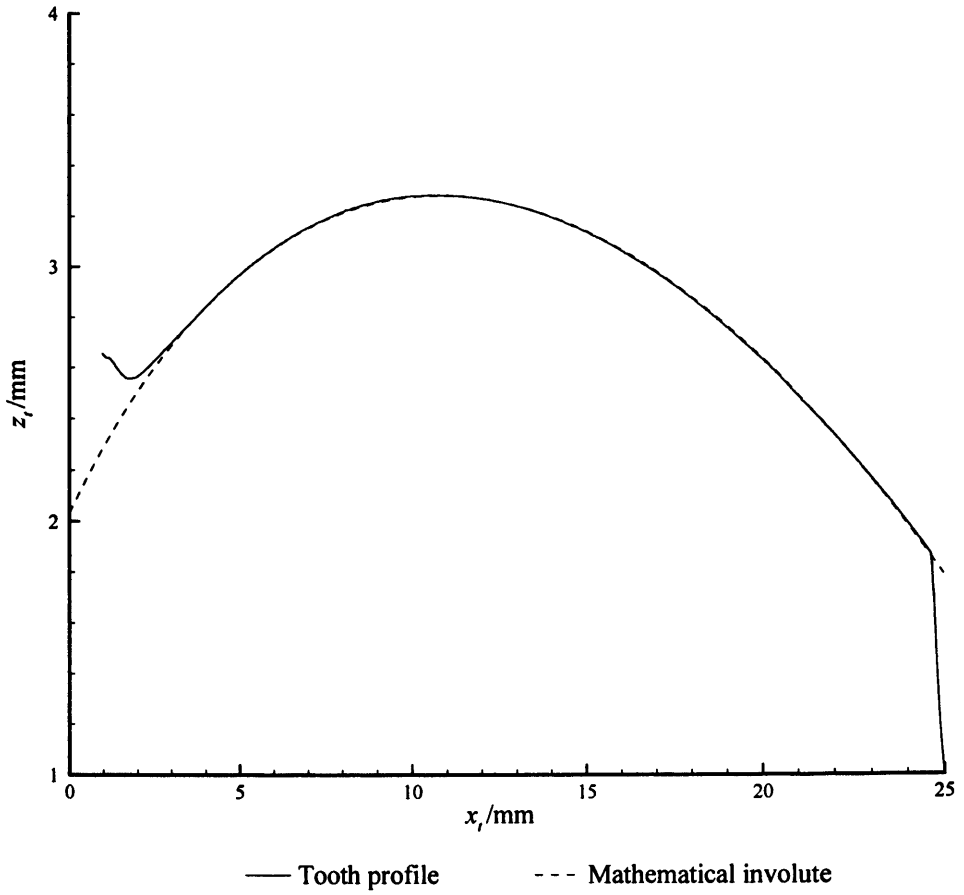

Figure 8.6 – True involute and mathematical approximation comparison

Figure 8.6 shows the curve fit obtained by this method compared to the gear tooth profile previously presented in figure 8.4. It can be seen that there is very good agreement over the majority of the involute. There are slight deviations occurring near the root and tip regions, this is probably due to the tooth relief features that are used. This result shows

that the mathematical involute equations developed give good agreement with actual involute profile. Also, the result confirms that the approach presented above can be used to generate a good agreement with the actual involute.

Once this fitting process has been completed, the reference angle, γ_{ref}, and the point at which this occurs in the Talysurf profile, x_{shift}, are known. The Talysurf profile is filtered to remove the form and to give the roughness profile. During this process a known quantity of data are removed from the start and end of the profile. If this removed quantity is expressed as x_{rem} then the distance in the roughness file that corresponds to the tangential point can be determined from:

$$x_{ref} = x_{shift} - x_{rem}$$

This value can be used along with the reference angle, γ_{ref}, to link the roughness data, which has no form and is just deviations from a mean line, to the part of the involute that they were originally taken from.

It was determined in section 2.2.2 that the contact between the two gear teeth is centred upon the involute angles γ_{c1} and γ_{c2}. Considering a general angle, γ_c, that corresponds to the involute angle at which the contact is centred, then the point in the roughness profile that corresponds to this point, x_{ruff}, is determined from:

$$\begin{aligned} x_{ruff} = x_{ref} &+ r_b \cos\gamma_{ref}\left(\cos\gamma_c + \gamma_c \sin\gamma_c - \cos\gamma_{ref} - \gamma_{ref}\sin\gamma_{ref}\right) \\ &+ r_b \sin\gamma_{ref}\left(\sin\gamma_c - \gamma_c\cos\gamma_c - \sin\gamma_{ref} + \gamma_{ref}\cos\gamma_{ref}\right) \end{aligned} \tag{8.6}$$

This allows the point in the roughness profile that contact is centred upon to be determined. The surrounding points in the roughness profile need to be used to generate the other features that are present within the contact area. This is done by considering the orientation of the EHL axis relative to the datum on which the roughness was measured. The EHL x-axis is tangential to the point at which contact is centred, so the angle between it and the roughness datum, θ_{axis}, is given by:

$$\theta_{axis} = \gamma_{ref} - \gamma_c$$

This angle allows the distance between the nodes in the computational mesh to be transformed onto the roughness axis. The distance between the nodes in the computational mesh is δx, this can be used along with the angle between the two axes, θ_{axis}, to calculate the spacing of the nodes on the roughness axis, δ_{ruff}, from:

$$\delta_{ruff} = \delta x \cos\theta_{axis} \tag{8.7}$$

It should be noted that there has been no adjustment to the height of the asperities due to the angle between the axes. This may be necessary as there is a measurement error due to the angle at which the stylus traverses the surface. A preliminary investigation was conducted to account for this effect by increasing the height of the asperities by the secant of θ_{axis}. This generated small increases in the asperity heights, but was found to have no significant effect on the results obtained.

In addition to the complication of putting the roughness onto the correct part of the involute during the meshing cycle, there is also difficulty in getting the roughness features into the contact area. Simply imposing the surface roughness onto the undeformed geometry and solving this as a steady state problem would flatten the roughness completely. Instead it was deemed better to fix the operating conditions to those at the first point and gradually bring the roughness into the contact in a transient manner. The method brings the roughness data in progressively on each surface until it is in the correct position for the first meshing position. Due to the different speeds of the two surfaces, the roughness data must be started at different points so that the correct parts of the roughness meet at the centre of the contact.

It is apparent from this that there are two separate methods required to incorporate the roughness of the surfaces into the gear meshing problem. Initially the points in the two roughness profiles that correspond to the centre of the contact have to be determined along with the spacing of the nodes in the roughness plane. This is then used to calculate the length of profile that has to run through the contact, this is combined with the speed at which each surface moves to determine points on the surface where they first become rough. This information is used in a transient solution method until the roughness is at the correct point, then the method of incorporating the roughness is changed for subsequent time steps. Equation (8.6) is used to determine the point in the roughness

related to the centre of contact and the other points are determined using the spacing given by equation (8.7). It should be noted that extra roughness data are needed due to the necessity to run some of the data through the contact region. The extra data required are created by mirroring the roughness data about the last point, this ensures that the roughness data used are representative of the region. Except for the changes necessary to deal with the incorporation of the roughness, the remainder of the solution is achieved in the same manner as the smooth solution technique.

8.3 Results

The approaches presented can be used for any involute gear configuration. Three different gear meshes are presented in this section, one corresponds to two small gears at high speed and modest load, the second involves larger gears at higher loads and modest speed and the final mesh involves very large gears operating under heavy loads and low speeds. The details of the meshes are given in Table 8.1.

Table 8.1 – Operating conditions for the three gear meshes

Mesh	A		B		C	
Gear	Pinion	Wheel	Pinion	Wheel	Pinion	Wheel
r_b / mm	33.9	97.5	67.3	481	141	1050
r_t / mm	39.5	106	80.1	520	165	1110
n	31	89	27	193	32	235
ω / RPM	22800	7960	5100	714	714	97
ψ / °	22.24		21.28		19.07	
w'_{max} / kNm^{-1}	193		508		1210	
w'_{start} / kNm^{-1}	60.2		169		402	
w'_{end} / kNm^{-1}	60.2		169		402	
η_0 / mPas	4.80		41.1		41.1	
α / GPa^{-1}	11.1		18.4		18.4	
Temperature / °C	100		50		50	

All the configurations are such that they operate as speed reduction gear boxes, so the pinion drives the wheel and the resultant speed is lower but has a greater amount of torque. This arrangement is common in real engineering systems, particularly in gear box

designs for ships, submarines and motor cars and auxiliary gear box configurations in a variety of aircraft.

The results for the meshes are split into two sections, one for the smooth surface results and another for the rough surface results. Rough surface results are only presented for mesh A and mesh C. Mesh B is not considered as there was no actual roughness data available for this particular gear mesh.

8.3.1 Smooth Surface Results

Two different approaches for the geometry of the gear teeth were presented in Chapter 2, one using equivalent roller geometry and another that uses the involute form. The results for both of the geometries are presented for mesh A in Figure 8.7 under isothermal conditions. This figure, and others shown subsequently, illustrate the variation in the load, w', maximum pressure, p_{max}, and minimum film thickness, h_0, obtained from the transient EHL analysis over the meshing cycle. The entrainment speed and the relative radius of curvature both increase over the meshing cycle, generating the general increase in minimum film thickness and reduction in maximum pressure.

The maximum pressure rises quickly when the load increases at the first change point. This change in the load causes the contact to spread rapidly, creating a smaller minimum film thickness. The increase in the contact area traps a large quantity of oil which progresses through the contact at the mean surface speed. The effect of this is seen in the increase in the minimum film thickness at the pitch point, as the trapped oil passes through the exit of the contact region. The effect of the trapped oil is also seen in the maximum pressure response, where the greater amount of oil generates larger pressures.

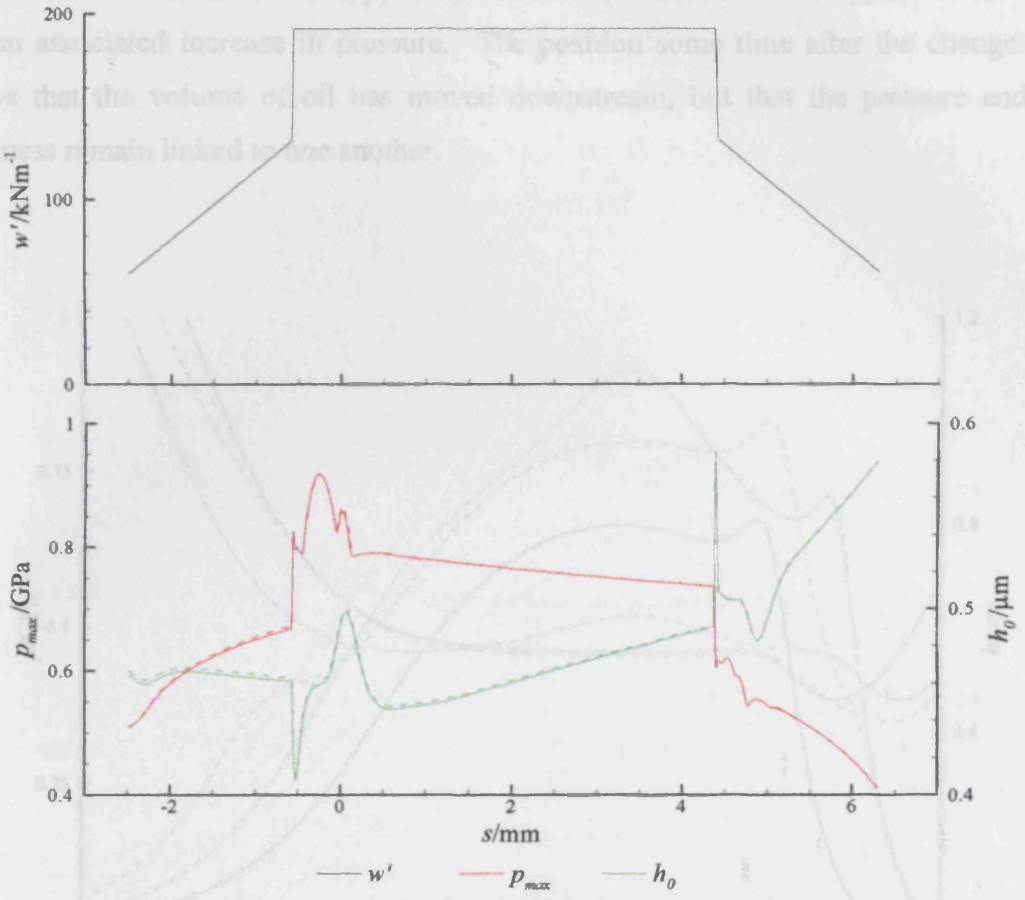

Figure 8.7 – Load, maximum pressure and minimum film thickness variation throughout the meshing cycle for mesh A using equivalent roller (solid) and involute (dashed) geometries

These effects are shown in more detail in Figure 8.8, where a series of EHL results for individual time steps are shown around the first change point for mesh A. The results show the calculated pressure and film thickness through the contact region for a position in the meshing cycle before the change point, immediately after the change point and some time after the change point. The film thickness and pressure before the change point are as expected for standard smooth steady state results. The results following the change point show that the pressure increases rapidly and spreads out over the contact. Comparing the film thickness for the position immediately after the change point with the result some time after the change point it can be seen that there is a trapped volume of oil in the inlet. The load carried by the two positions is identical and only slight changes in the relative radius of curvature have developed. The film thickness for the position immediately after the change point in the region $-1.2 \le x/a \le -0.7$ is higher as a result

of an increased volume of oil trapped by the increase in load. The trapped volume of oil has an associated increase in pressure. The position some time after the change point shows that the volume of oil has moved downstream, but that the pressure and film thickness remain linked to one another.

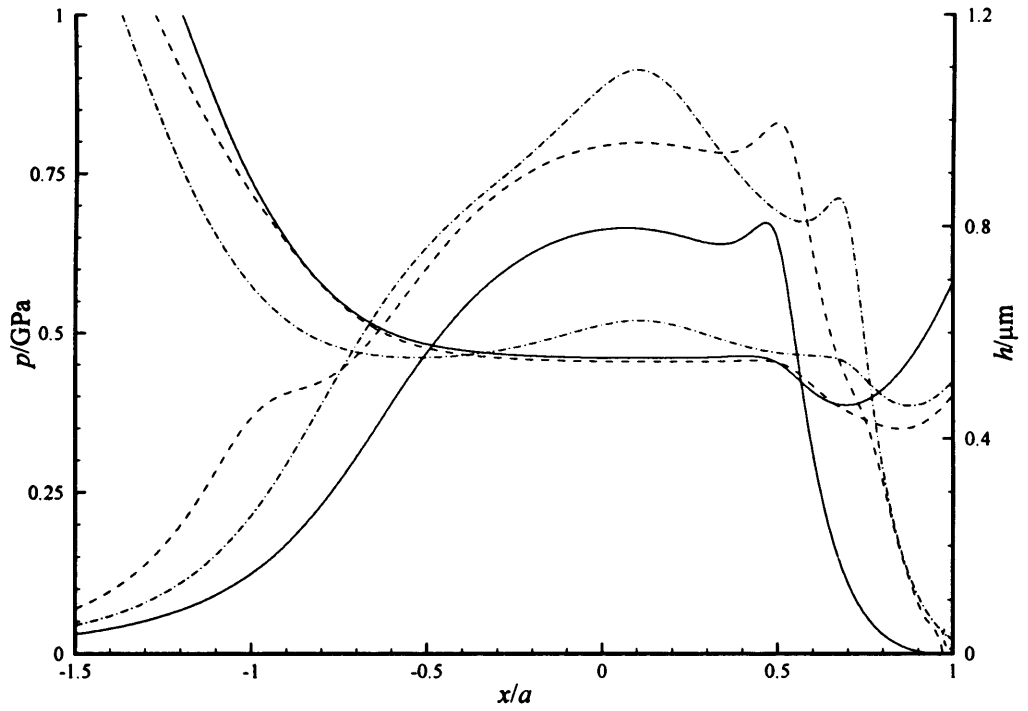

Figure 8.8 – Pressure and film thickness variation through the contact during the meshing cycle near the first change point; just before the change point (solid), immediately after the change point (dashed) and some time after the change point (dash-dotted)

The results taken around the second change point, shown in Figure 8.9, can be compared with those for the first change point in figure 8.8. It can be seen that the reduction in the load causes the minimum film thickness to rise rapidly as the kink formed at the exit is removed shortly after the change point. The contact area reduces rapidly resulting in a reduction in the film thickness in the inlet, which is opposite to that seen for the first change point. The drop in the film thickness continues to move through the contact in much the same way that the increased film thickness does in figure 8.8. The film thickness minimum develops a localised reduction in pressure, but it can also be seen that it develops a very slight pressure increase formed slightly ahead of the film thickness drop. The pressure features follow the film thickness feature through the contact and are

responsible for the small kink in the maximum pressure developed at $s \approx 4.5\,\text{mm}$ in figure 8.7. When the film thickness feature reaches the end of the contact it creates the drop in the minimum film thickness, which is seen at $s \approx 5.0\,\text{mm}$.

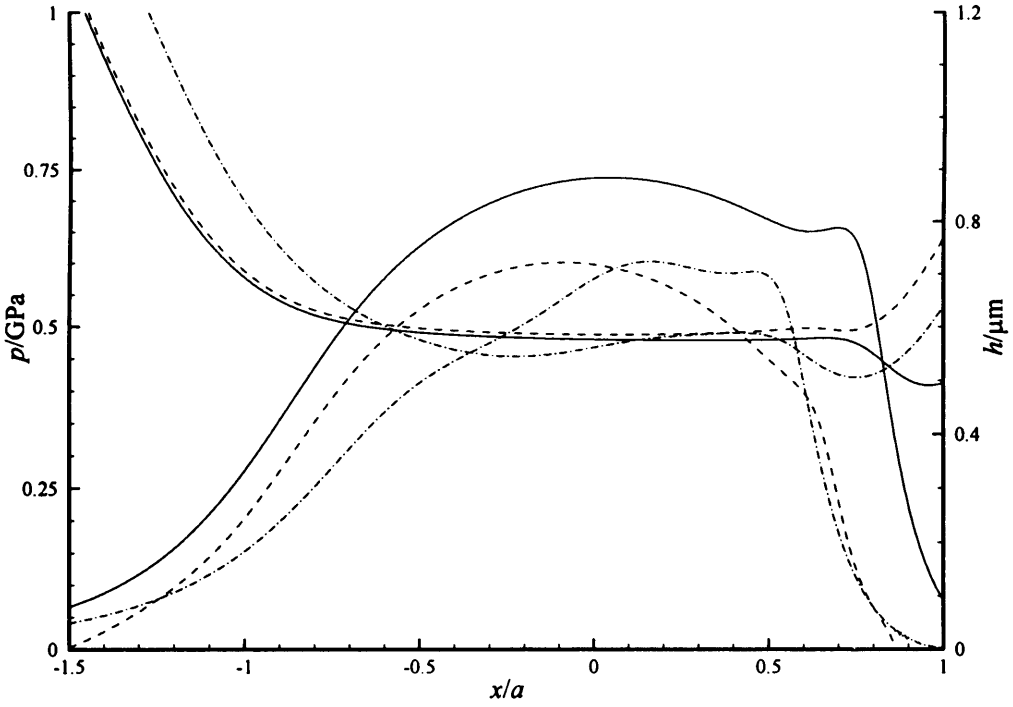

Figure 8.9 – Pressure and film thickness variation through the contact during the meshing cycle near the second change point; just before the change point (solid), immediately after the change point (dashed) and some time after the change point (dash-dotted)

There are slight differences seen between the cylindrical approximation and the true involute geometries in figure 8.7. The involute geometry develops slightly larger minimum film thicknesses throughout the meshing cycle and also gives rise to slightly higher maximum pressures. The changes are most noticeable at the beginning of the meshing cycle due to the smaller radii of curvature, but the differences are small.

Figure 8.10 shows a comparison of the EHL results for a position early on in the meshing cycle. Both geometries develop the same features, but slight differences between the two results can be seen. The change in the curvature for the involute is clearly seen, where the involute and equivalent roller geometries are different outside of the contact area. The equivalent roller approximation has too small a radius of curvature in the inlet, and too

large a radius of curvature in the exit. This is to be expected, but it does lead to slightly different pressure and film thickness results. The pressure generated is broader for the equivalent roller approximation, which generates thinner films. Although the difference is small for this mesh, it should be noted that the differences become more noticeable when higher loads are used due to the larger contact areas developed.

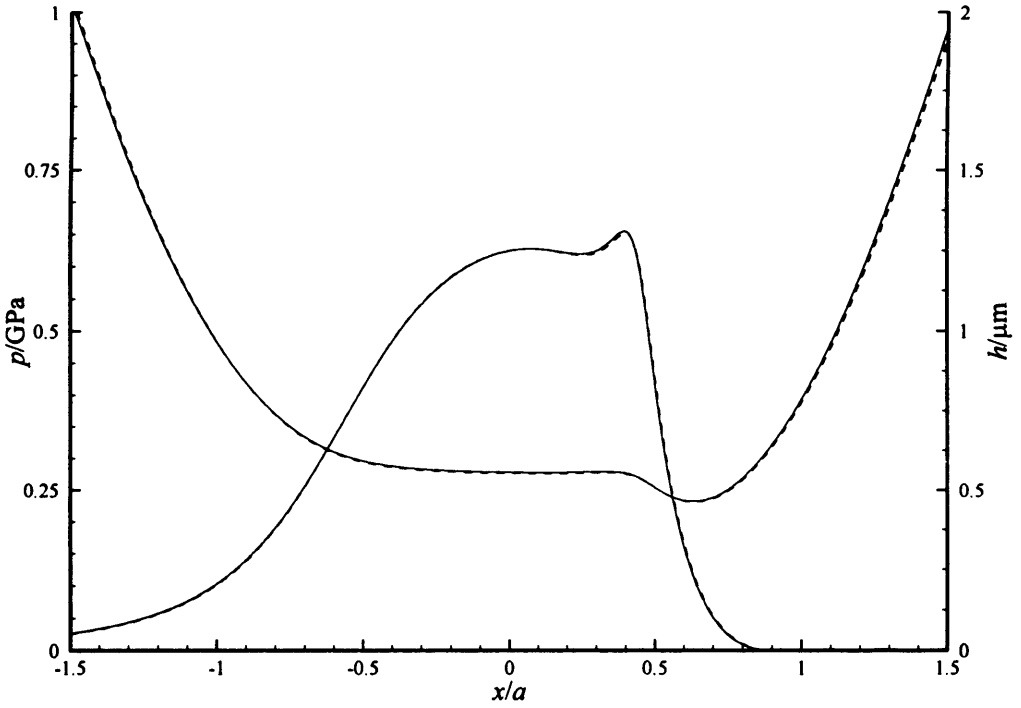

Figure 8.10 – Pressure and film thickness variation through the contact for true involute (solid) and equivalent roller (dashed) geometries

A comparison of the results for the thermal and isothermal conditions for mesh A is shown in Figure 8.11. In addition to the load, maximum pressure and minimum film thickness, the figure shows the variation of the maximum surface temperatures, $\theta_{1\,max}$ and $\theta_{2\,max}$, and the maximum mid-film temperature, $\theta_{m\,max}$. The inclusion of the temperature effects reduces the film thickness and the maximum pressure, but despite the reduction the features generated remain substantially the same. The thermal results do not show the same sensitivity to the change points as the isothermal results, this is seen by a slightly lower change in the minimum film thickness at and following the changes in the load. The increase in the minimum film thickness during the mesh, generated by the increase in the entrainment velocity, is lower for the thermal results than the isothermal results. The

minimum film thickness does not increase as rapidly due to the increasing temperature generated as the mesh progresses past the pitch point. The temperatures increase because the entrainment velocity and slide roll ratio both increase once the pitch point is passed generating greater shear heating. The higher temperatures developed act to reduce the viscosity and therefore lessen the effect of the increased entrainment velocity.

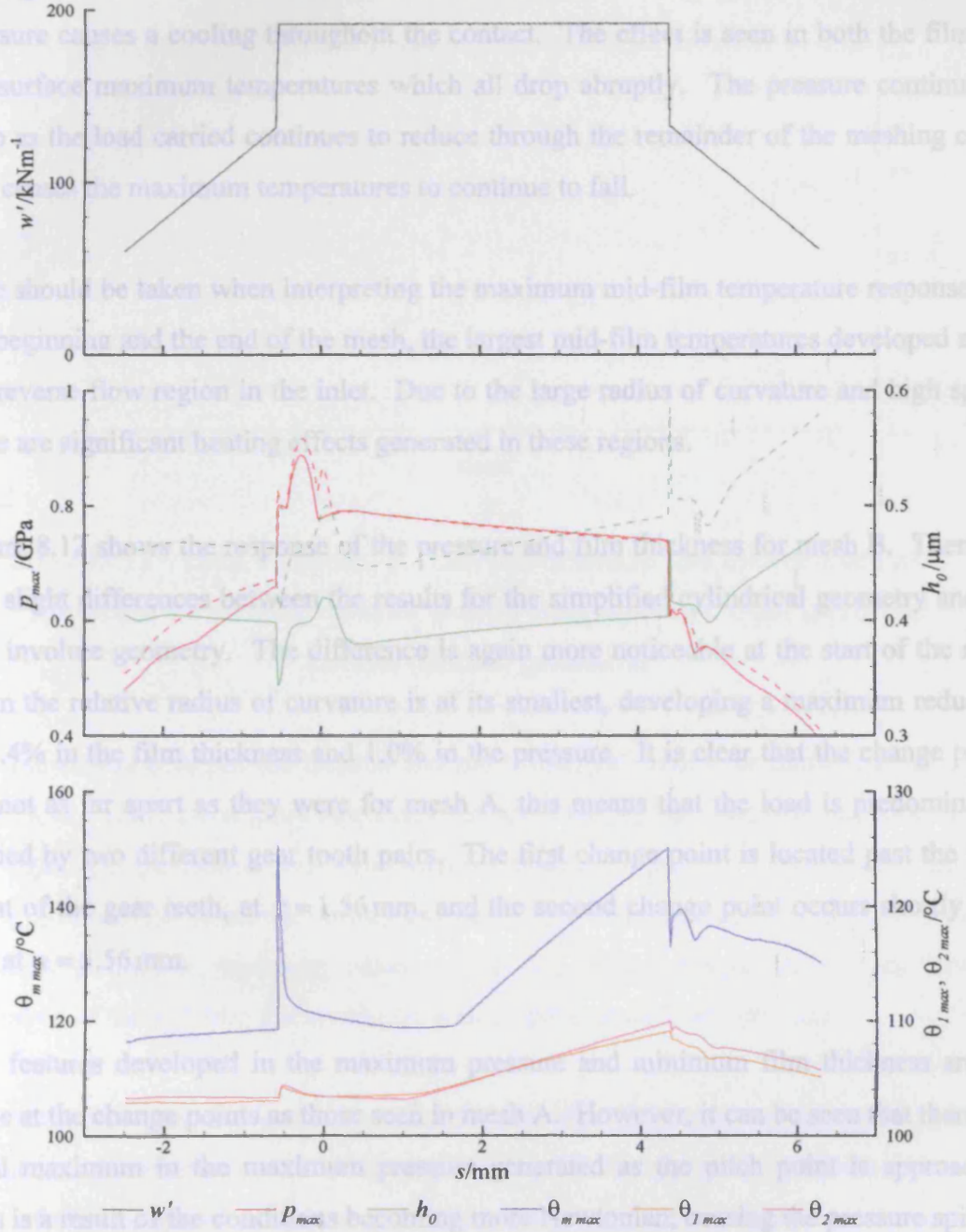

Figure 8.11 – Load, maximum pressure, minimum film thickness and maximum temperature response for the thermal (solid) and isothermal (dashed) condition for mesh A

The maximum mid-film and surface temperatures rise rapidly at the first change point, this is generated by the increase in the pressure which increases the compressive heating throughout the contact. After this initial rise, the maximum temperatures reduce as the pitch point is approached. The fall in the temperatures results from a lower shear heating effect due to the reduction in the sliding velocity. The maximum temperatures remain constant after this point until the sliding velocity increases generating a greater shear heating effect. When the load reduces at the second change point, the change in the pressure causes a cooling throughout the contact. The effect is seen in both the film and the surface maximum temperatures which all drop abruptly. The pressure continues to drop as the load carried continues to reduce through the remainder of the meshing cycle, this causes the maximum temperatures to continue to fall.

Care should be taken when interpreting the maximum mid-film temperature response. At the beginning and the end of the mesh, the largest mid-film temperatures developed are in the reverse flow region in the inlet. Due to the large radius of curvature and high speed, there are significant heating effects generated in these regions.

Figure 8.12 shows the response of the pressure and film thickness for mesh B. There are still slight differences between the results for the simplified cylindrical geometry and the true involute geometry. The difference is again more noticeable at the start of the mesh when the relative radius of curvature is at its smallest, developing a maximum reduction of 1.4% in the film thickness and 1.0% in the pressure. It is clear that the change points are not as far apart as they were for mesh A, this means that the load is predominantly carried by two different gear tooth pairs. The first change point is located past the pitch point of the gear teeth, at $s = 1.56\,\text{mm}$, and the second change point occurs shortly after this at $s = 5.56\,\text{mm}$.

The features developed in the maximum pressure and minimum film thickness are the same at the change points as those seen in mesh A. However, it can be seen that there is a local maximum in the maximum pressure generated as the pitch point is approached. This is a result of the conditions becoming more Newtonian, causing the pressure spike to rise. Similarly to mesh A, the maximum pressure is lower at the end of the mesh than the beginning and the minimum film thickness generally increases.

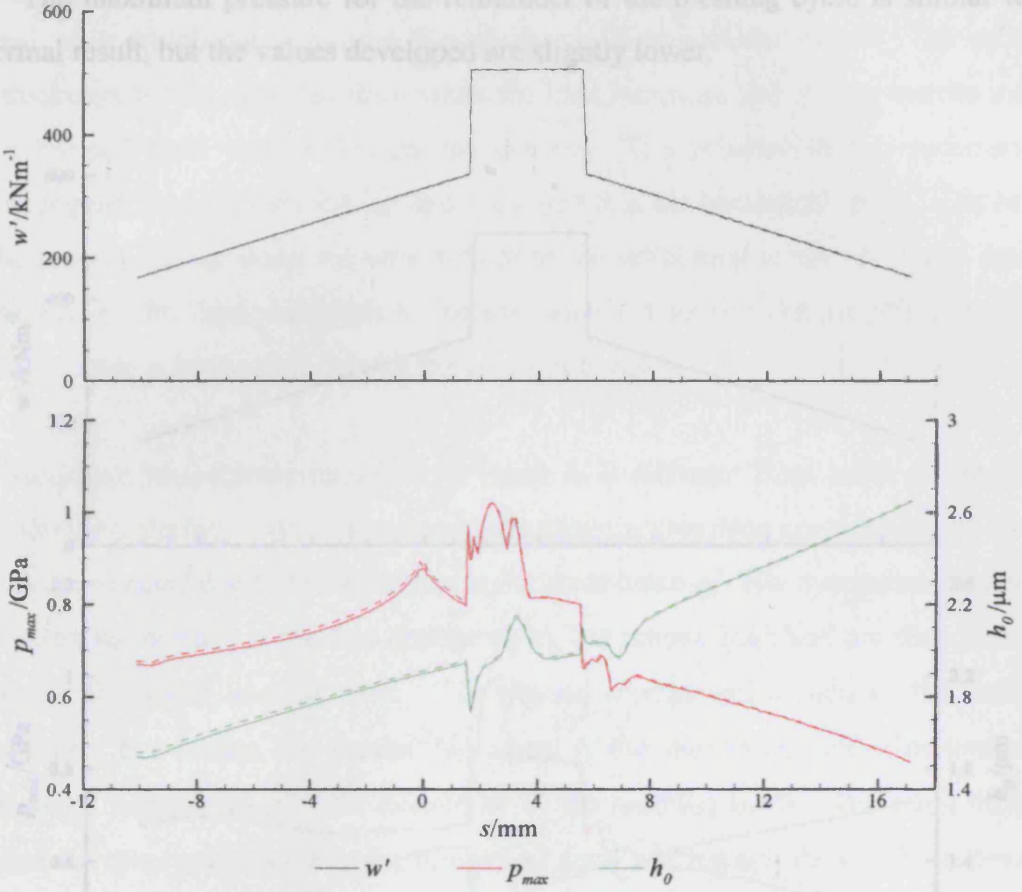

Figure 8.12 – Load, maximum pressure and minimum film thickness variation throughout the meshing cycle for mesh B using equivalent roller (solid) and involute (dashed) geometries

The comparison between the thermal and the isothermal results for mesh B is shown in Figure 8.13. There are slightly larger differences in the pressure and film thickness due to the temperature interaction than those seen for mesh A. Figure 8.11 showed that the film thickness was offset to smaller values and only slight differences occurred in the features developed in the film thickness. However, figure 8.13 shows larger differences between the shapes of the two film thicknesses as well as differences in the pressures.

The maximum pressure generally falls for the thermal result as the pitch point is approached, in contrast to the isothermal result where the pressure increased. This is a result of the temperature response generating lower viscosities which allow the pressure spike to reduce. The maximum pressure for the thermal case increases quickly when the first change point is reached and continues to increase shortly after due to the trapped

fluid. The maximum pressure for the remainder of the meshing cycle is similar to the isothermal result, but the values developed are slightly lower.

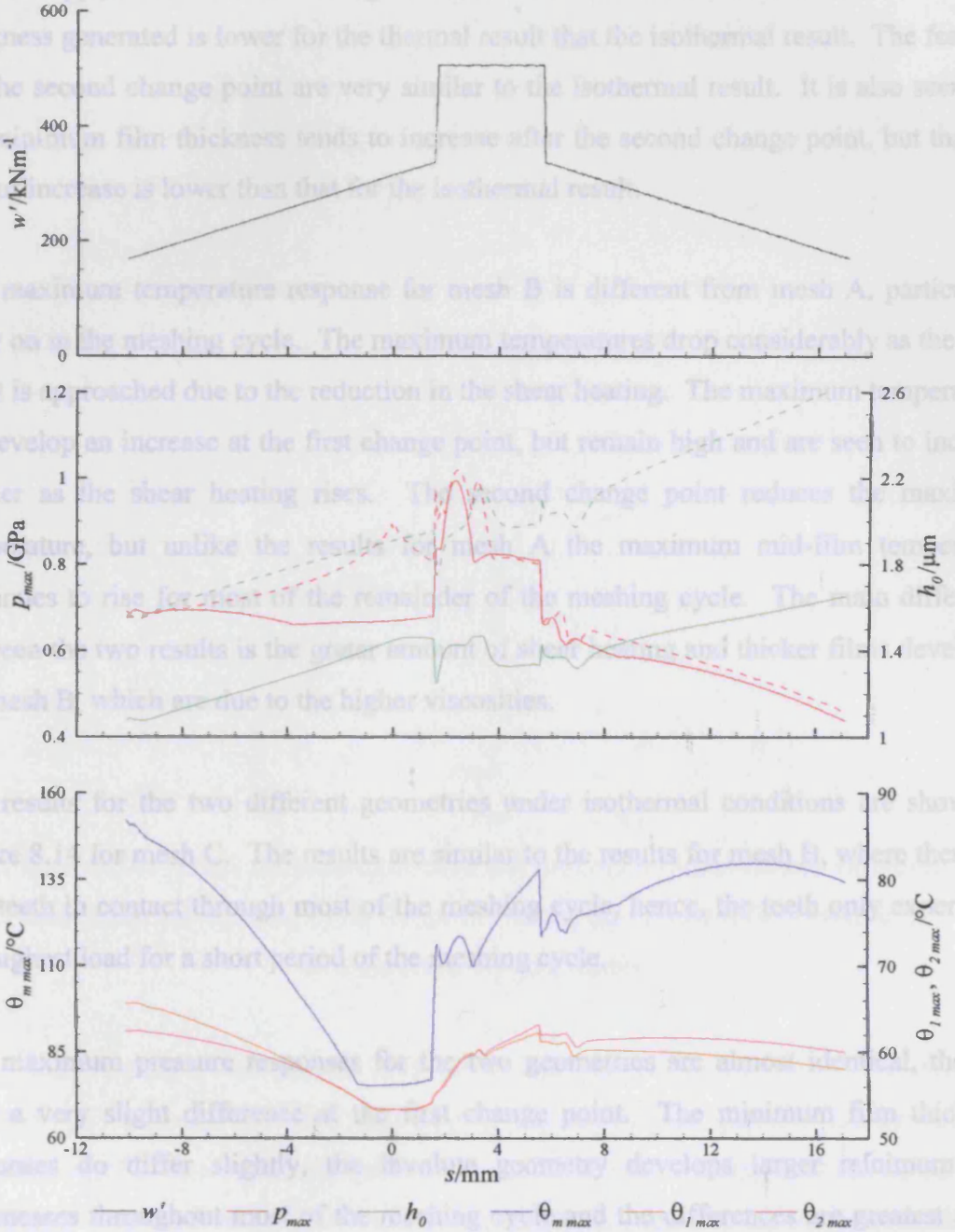

Figure 8.13 – Load, maximum pressure, minimum film thickness and maximum temperature response for the thermal (solid) and isothermal (dashed) condition for mesh B

The minimum film thickness increases for the thermal result until the pitch point is reached, after which the minimum film thickness is seen to flatten off. The minimum film thickness shows a sudden drop when the load increases and is then seen to increase as the trapped fluid moves through the contact. The increase in the minimum film thickness generated is lower for the thermal result that the isothermal result. The features for the second change point are very similar to the isothermal result. It is also seen that the minimum film thickness tends to increase after the second change point, but that the rate of increase is lower than that for the isothermal result.

The maximum temperature response for mesh B is different from mesh A, particularly early on in the meshing cycle. The maximum temperatures drop considerably as the pitch point is approached due to the reduction in the shear heating. The maximum temperatures do develop an increase at the first change point, but remain high and are seen to increase further as the shear heating rises. The second change point reduces the maximum temperature, but unlike the results for mesh A the maximum mid-film temperature continues to rise for most of the remainder of the meshing cycle. The main difference between the two results is the grater amount of shear heating and thicker films developed for mesh B, which are due to the higher viscosities.

The results for the two different geometries under isothermal conditions are shown in Figure 8.14 for mesh C. The results are similar to the results for mesh B, where there are two teeth in contact through most of the meshing cycle, hence, the teeth only experience the highest load for a short period of the meshing cycle.

The maximum pressure responses for the two geometries are almost identical, there is only a very slight difference at the first change point. The minimum film thickness responses do differ slightly, the involute geometry develops larger minimum film thicknesses throughout most of the meshing cycle and the differences are greatest at the start of the mesh. The reduction in the minimum film thickness that occurs when the load is removed is far greater for the involute geometry than for the equivalent roller geometry.

Due to the speed at which the oil moves through the contact and the time between the load changes, there is an interaction between the effects of the two load changes. The

result is that there is no increase in the minimum film thickness for the second load change, a reduction in film thickness occurs instead.

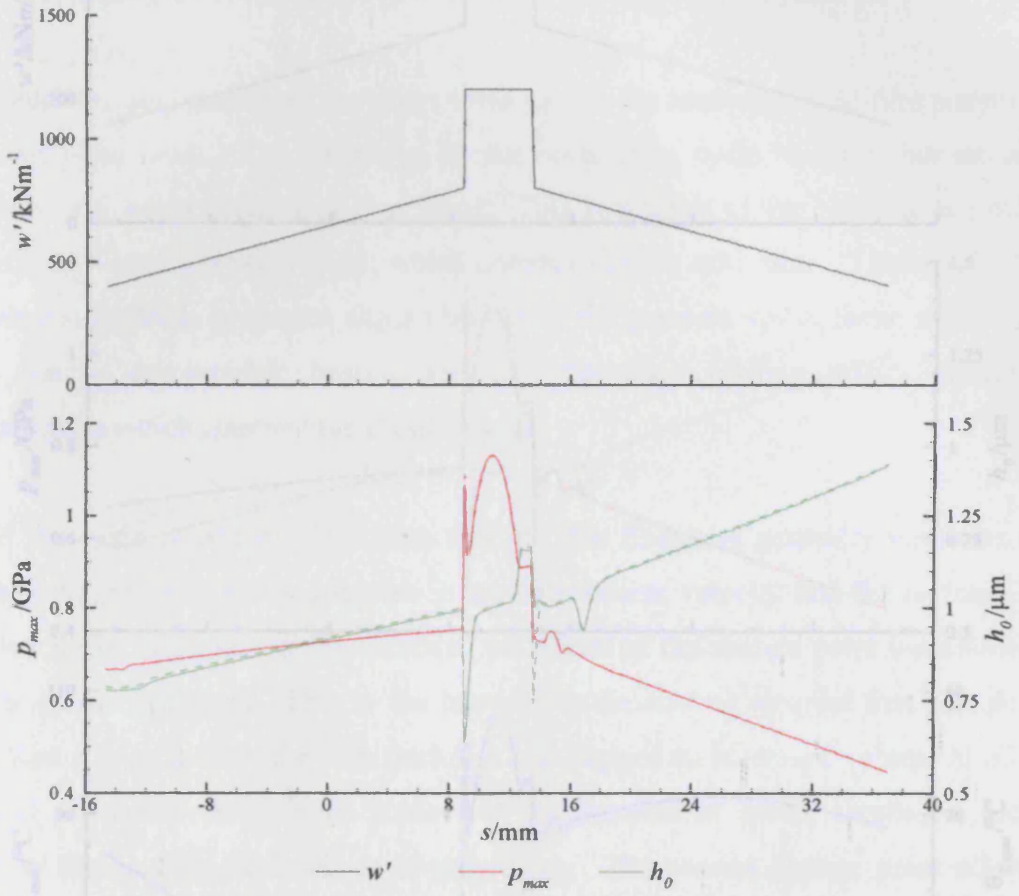

Figure 8.14 – Load, maximum pressure and minimum film thickness variation throughout the meshing cycle for mesh C using equivalent roller (solid) and involute (dashed) geometries

Figure 8.15 compares the thermal and isothermal results for mesh C. The thermal approach generates thinner film thicknesses, but retains the features present in the isothermal result. The maximum pressure is almost the same throughout the mesh with only a small region when the load reaches the maximum value that the thermal results are lower than the isothermal. The rate at which the minimum film thickness increases towards the end of the mesh is again lower for the thermal results than the isothermal.

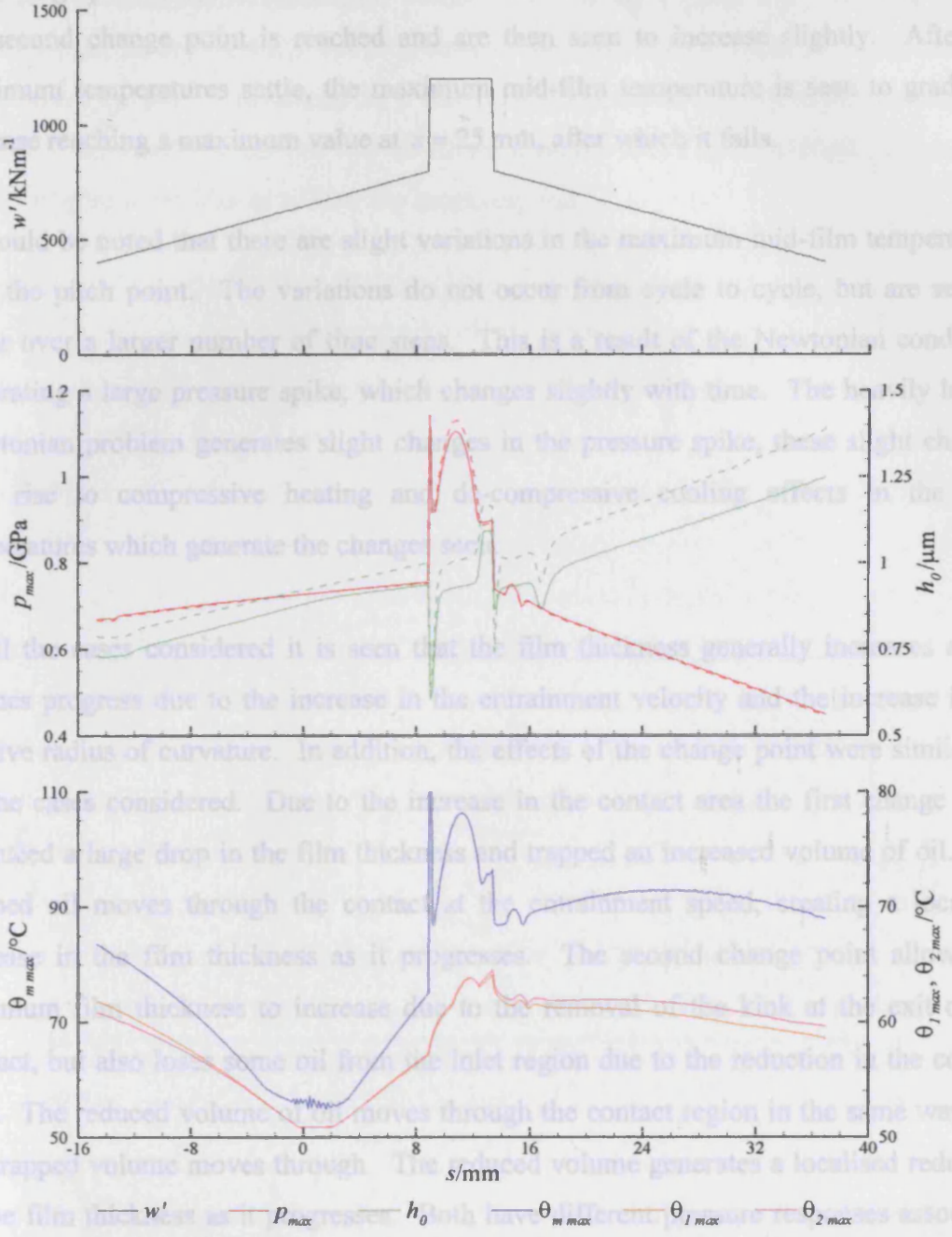

Figure 8.15 – Load, maximum pressure, minimum film thickness and maximum temperature response for the thermal (solid) and isothermal (dashed) condition for mesh C

The maximum temperature response for mesh C is reasonably similar to mesh B. The maximum temperatures fall as the pitch point is approached due to the reduction in the shear heating. After the pitch point the maximum temperatures rise as the sliding velocity increases generating more shear heating. When the first change point is reached there is a large increase in the maximum temperatures, there is also another region of heating after

that change point that was also seen for mesh B. The maximum temperatures fall when the second change point is reached and are then seen to increase slightly. After the maximum temperatures settle, the maximum mid-film temperature is seen to gradually increase reaching a maximum value at $s = 25$ mm, after which it falls.

It should be noted that there are slight variations in the maximum mid-film temperatures near the pitch point. The variations do not occur from cycle to cycle, but are seen to occur over a larger number of time steps. This is a result of the Newtonian conditions generating a large pressure spike, which changes slightly with time. The heavily loaded Newtonian problem generates slight changes in the pressure spike, these slight changes give rise to compressive heating and de-compressive cooling effects in the fluid temperatures which generate the changes seen.

In all the cases considered it is seen that the film thickness generally increases as the meshes progress due to the increase in the entrainment velocity and the increase in the relative radius of curvature. In addition, the effects of the change point were similar for all the cases considered. Due to the increase in the contact area the first change point produced a large drop in the film thickness and trapped an increased volume of oil. The trapped oil moves through the contact at the entrainment speed, creating a localised increase in the film thickness as it progresses. The second change point allows the minimum film thickness to increase due to the removal of the kink at the exit of the contact, but also loses some oil from the inlet region due to the reduction in the contact area. The reduced volume of oil moves through the contact region in the same way that the trapped volume moves through. The reduced volume generates a localised reduction in the film thickness as it progresses. Both have different pressure responses associated with them, the increased volume of oil generates a localised increase in the pressure, whereas the reduced volume of oil generates a reduction in pressure.

Since the minimum film thicknesses increase over the meshing cycle the overall minimum film thickness occurs early in the meshing of the gear teeth. For meshes A and C the smallest minimum film thickness occurs at the first change point, however, mesh B develops the lowest minimum film thickness at the beginning of the meshing cycle.

It is worth considering the manner in which these meshes respond when they are run as speed increase (as opposed to reduction) drives. The result of this modification is that the wheel drives the pinion, reversing the direction that the contact moves to that considered thus far. The results from the two different arrangements can be plotted against one another to give some idea as to how the gears respond.

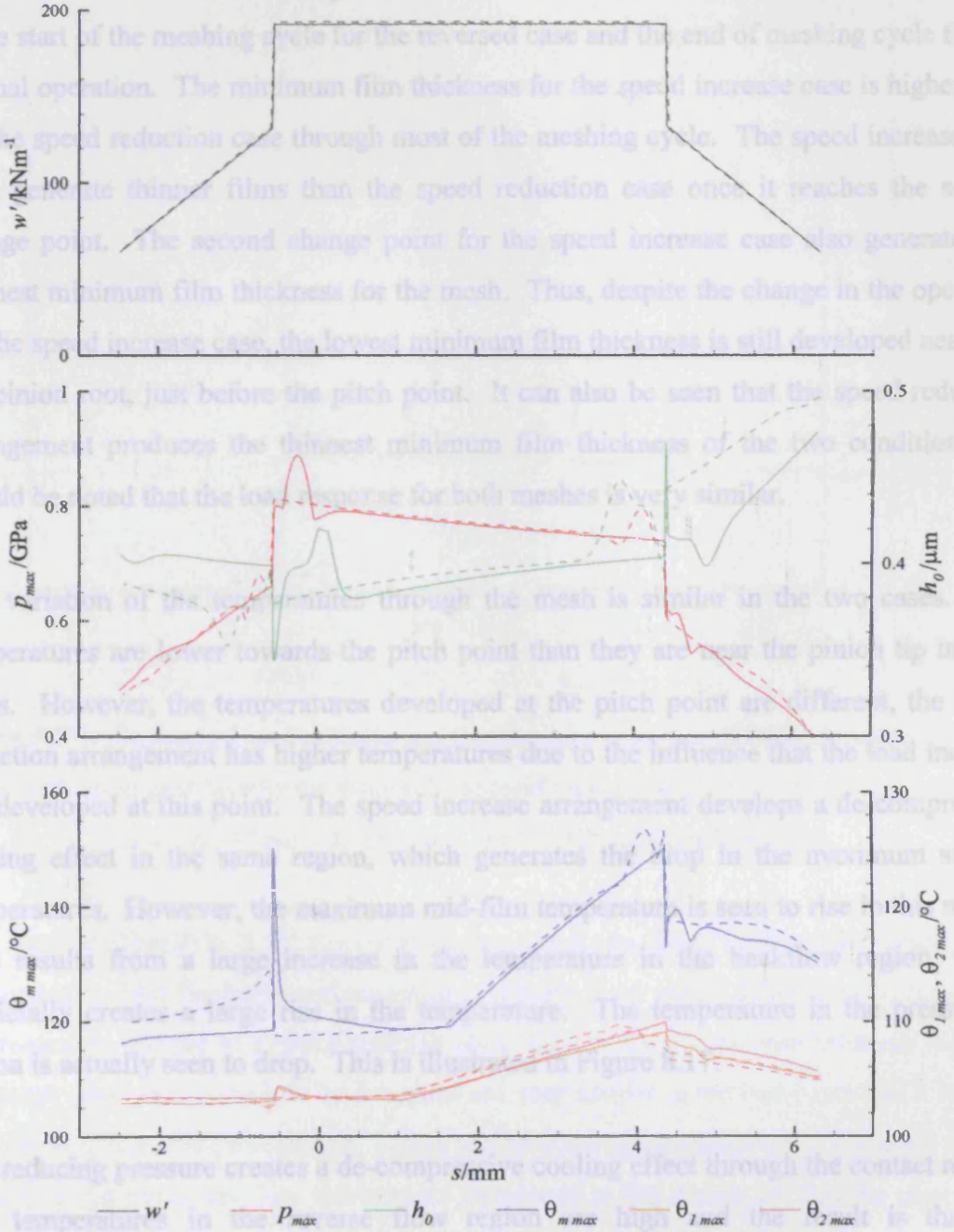

Figure 8.16 – Load, maximum pressure, minimum film thickness and maximum temperature response of mesh A operating as a speed reduction (solid) and speed increase (dashed) gear sets

The result of reversing the driving gear is shown in Figure 8.16 for mesh A. The same features are generated at the change points in both cases, but it is seen that they are opposite to one another due to the reversal of the contact path. The maximum pressure response for both cases is also very similar, but the minimum film thicknesses developed are different.

The minimum film thickness is greater for the reversed case towards the pinion tip, which is the start of the meshing cycle for the reversed case and the end of meshing cycle for the normal operation. The minimum film thickness for the speed increase case is higher than for the speed reduction case through most of the meshing cycle. The speed increase case does generate thinner films than the speed reduction case once it reaches the second change point. The second change point for the speed increase case also generates the thinnest minimum film thickness for the mesh. Thus, despite the change in the operation for the speed increase case, the lowest minimum film thickness is still developed nearer to the pinion root, just before the pitch point. It can also be seen that the speed reduction arrangement produces the thinnest minimum film thickness of the two conditions. It should be noted that the load response for both meshes is very similar.

The variation of the temperatures through the mesh is similar in the two cases. The temperatures are lower towards the pitch point than they are near the pinion tip in both cases. However, the temperatures developed at the pitch point are different, the speed reduction arrangement has higher temperatures due to the influence that the load increase has developed at this point. The speed increase arrangement develops a de-compressive cooling effect in the same region, which generates the drop in the maximum surface temperatures. However, the maximum mid-film temperature is seen to rise in this region. This results from a large increase in the temperature in the backflow region, which artificially creates a large rise in the temperature. The temperature in the pressurised region is actually seen to drop. This is illustrated in Figure 8.17.

The reducing pressure creates a de-compressive cooling effect through the contact region. The temperatures in the reverse flow region are high and the result is that the temperatures in this region influence the maximum mid-film temperature shown in figure 8.16. It can be seen that as the load is removed at $s = -0.587\,\text{mm}$ the mid-film temperature in the high pressure region reduces generating the reduction in the maximum

surface temperatures. This is accompanied by an increase in the temperature, shown in figure 8.16, in the reverse flow region and is the reason for the increase in the maximum mid-film temperature when the load is removed.

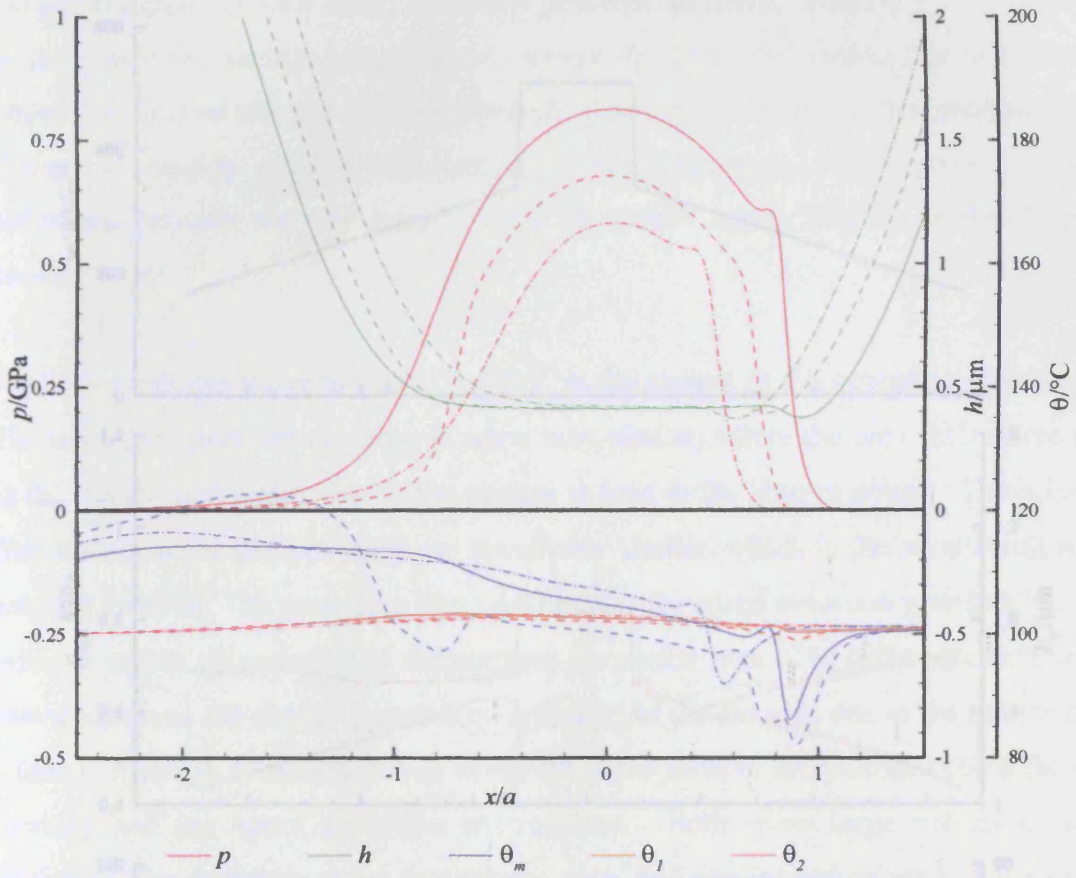

Figure 8.17 – Pressure, film thickness and temperature response of mesh A operating as a speed reduction gear arrangement near the second change point; $s = -0.543$ mm (solid), $s = -0.587$ mm (dashed) and $s = -1.423$ mm (dash-dotted)

The response of mesh B to the change in the operating conditions is shown in Figure 8.18. The minimum film thickness and maximum pressure show slight differences as the gears are changed to operate as a speed increase drive, which was not the case for mesh A. The maximum pressures in the low load regions are very similar in the two cases and it is also seen the minimum film thicknesses in this region show good agreement. There are slight changes in the minimum film thicknesses developed at the pinion tip, but the minimum film thicknesses developed near the pinion root are almost identical. There are differences in the minimum film thickness near the change points, but the features seen at the first and second change points are consistent. The location of the smallest minimum

film thickness for the speed increase arrangement is still seen at the pinion root near the end of the meshing cycle, but the case does not reach as low a value as the speed reduction case.

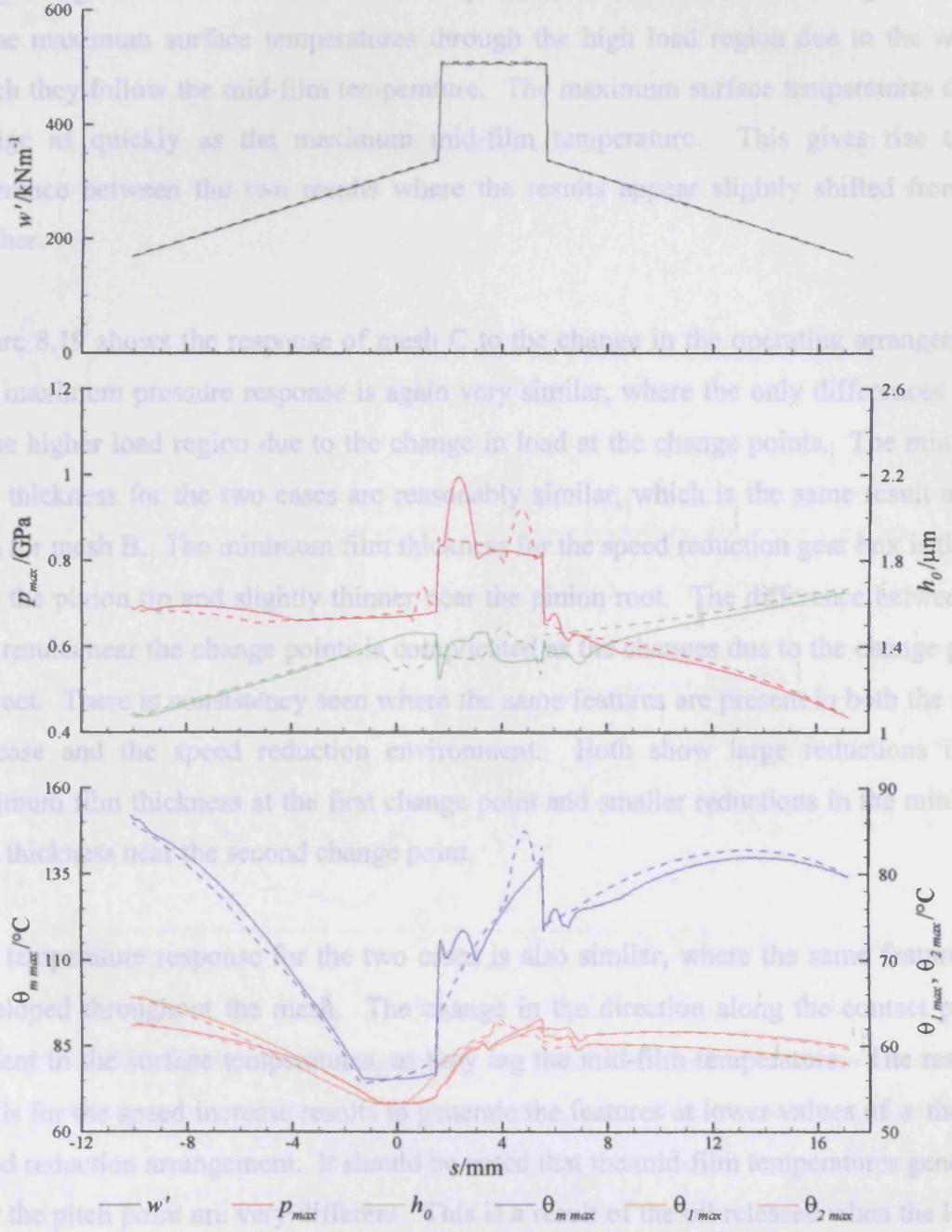

Figure 8.18 – Load, maximum pressure, minimum film thickness and maximum temperature response of mesh B operating as a speed reduction (solid) and speed increase (dashed) gear sets

The maximum mid-film and surface temperatures developed also show good agreement outside of the high load region. The maximum surface temperatures are almost identical near the pinion tip and only slight changes are seen in the temperatures at the pinion root. The maximum mid-film temperature developed for both cases are very similar, reflecting the good agreement seen in the maximum pressure variation. There is a large difference in the maximum surface temperatures through the high load region due to the way in which they follow the mid-film temperature. The maximum surface temperatures do not change as quickly as the maximum mid-film temperature. This gives rise to the difference between the two results where the results appear slightly shifted from one another.

Figure 8.19 shows the response of mesh C to the change in the operating arrangements. The maximum pressure response is again very similar, where the only differences occur in the higher load region due to the change in load at the change points. The minimum film thickness for the two cases are reasonably similar, which is the same result as that seen for mesh B. The minimum film thickness for the speed reduction gear box is thinner near the pinion tip and slightly thinner near the pinion root. The difference between the two results near the change points is complicated as the changes due to the change points interact. There is consistency seen where the same features are present in both the speed increase and the speed reduction environment. Both show large reductions in the minimum film thickness at the first change point and smaller reductions in the minimum film thickness near the second change point.

The temperature response for the two cases is also similar, where the same features are developed throughout the mesh. The change in the direction along the contact path is evident in the surface temperatures, as they lag the mid-film temperature. The result of this is for the speed increase results to generate the features at lower values of s than the speed reduction arrangement. It should be noted that the mid-film temperatures generated near the pitch point are very different. This is a result of the oil released when the load is removed slowly moving through the reverse flow region, which artificially increases the maximum mid-film temperature.

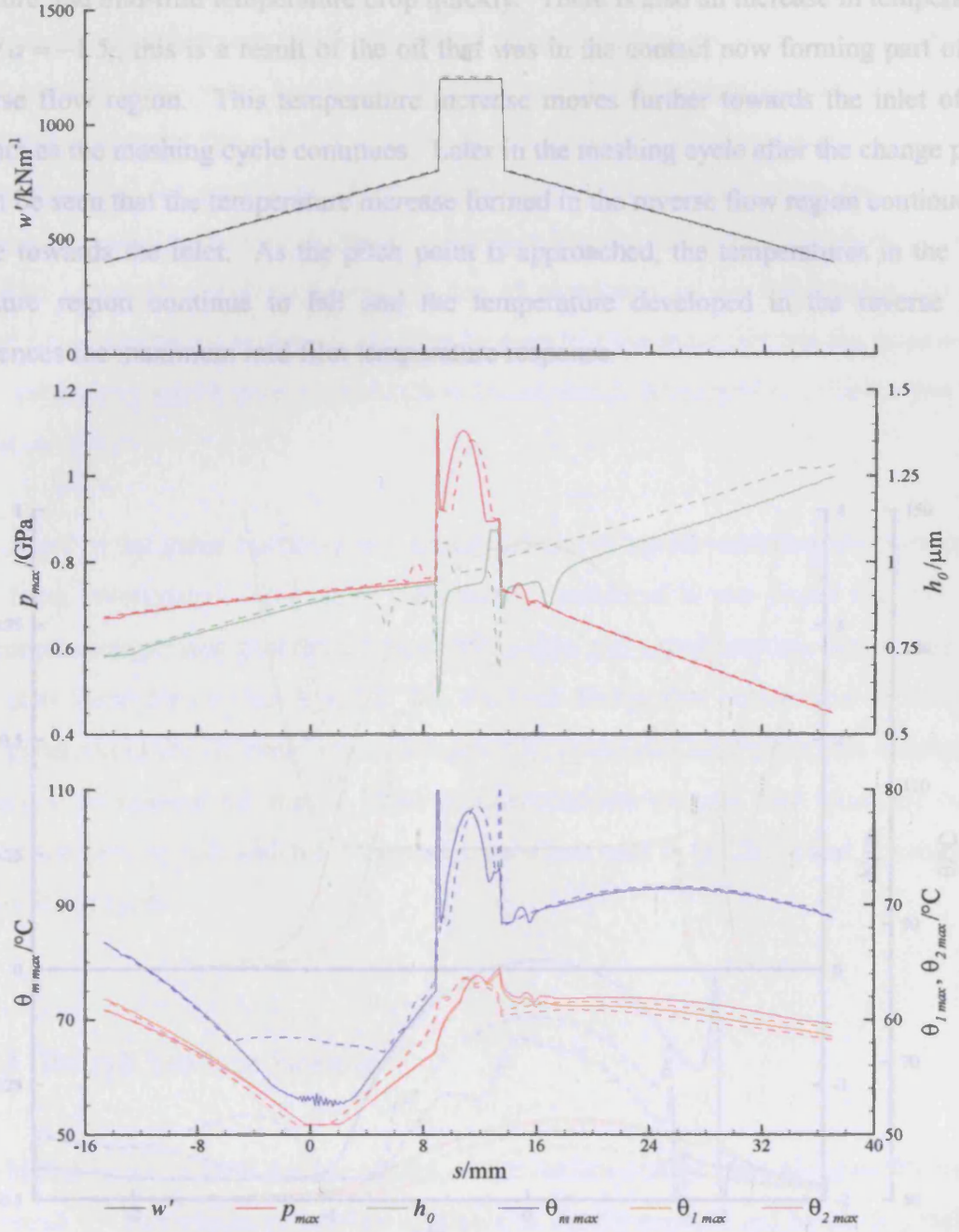

Figure 8.19 – Load, maximum pressure, minimum film thickness and maximum temperature response of mesh C operating as a speed reduction (solid) and speed increase (dashed) gear sets

This increase in the maximum mid-film temperature in figure 8.19 near the pitch point is the same feature seen for mesh A in figure 8.17, it is illustrated again for mesh C in Figure 8.20. Before the load change, the highest temperatures are reached in the centre of the contact. At $s = 8.736$mm the load has been removed and it can be seen that the

pressure and mid-film temperature drop quickly. There is also an increase in temperature at $x/a = -1.5$, this is a result of the oil that was in the contact now forming part of the reverse flow region. This temperature increase moves further towards the inlet of the contact as the meshing cycle continues. Later in the meshing cycle after the change point it can be seen that the temperature increase formed in the reverse flow region continues to move towards the inlet. As the pitch point is approached, the temperatures in the high pressure region continue to fall and the temperature developed in the reverse flow influences the maximum mid-film temperature response.

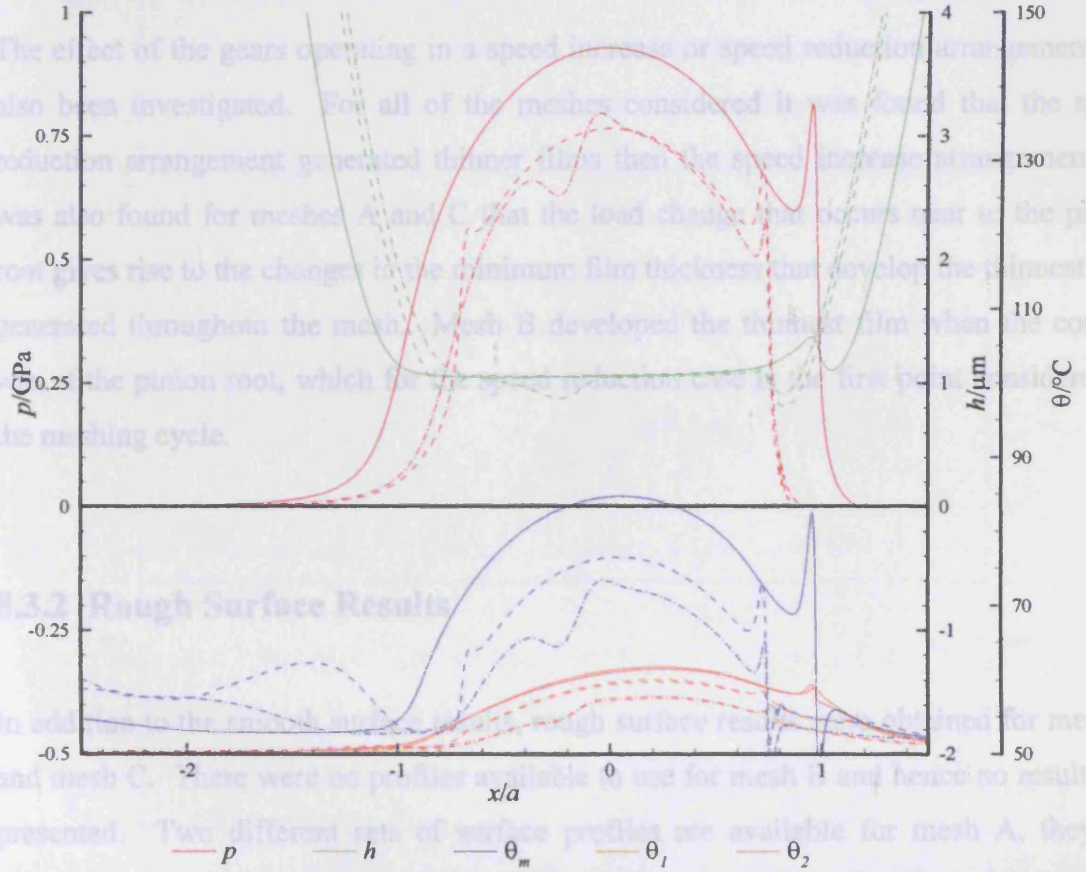

Figure 8.20 – Pressure, film thickness and temperature response of mesh C operating as a speed reduction gear arrangement near the second change point; $s = 8.995$ mm (solid), $s = 8.736$ mm (dashed) and $s = 1.119$ mm (dash-dotted)

The results indicate that the variation in the load developed at the change points gives rise to large changes in minimum film thickness and that they can be responsible for developing the thinnest films throughout the gear mesh. Comparisons between simplified

cylindrical geometry and true involute geometry have been conducted and it was seen that the overall differences are small. However, when the actual pressure and film thickness are compared between the two results it can be seen that the simplified geometry does not accurately represent the actual curvature of the surfaces. The load is supported over a larger area for the equivalent roller geometry generating lower pressures. The film thickness is smaller for the equivalent roller than the true involute. This is related to the involute geometry developing larger radius of curvatures in the inlet region that have better film forming properties. It can also be seen that the same features are generated by both geometries which gives confidence in the approach developed to calculate the shape of the involute.

The effect of the gears operating in a speed increase or speed reduction arrangement has also been investigated. For all of the meshes considered it was found that the speed reduction arrangement generated thinner films then the speed increase arrangement. It was also found for meshes A and C that the load change that occurs near to the pinion root gives rise to the changes in the minimum film thickness that develop the thinnest film generated throughout the mesh. Mesh B developed the thinnest film when the contact was at the pinion root, which for the speed reduction case is the first point considered in the meshing cycle.

8.3.2 Rough Surface Results

In addition to the smooth surface results, rough surface results were obtained for mesh A and mesh C. There were no profiles available to use for mesh B and hence no results are presented. Two different sets of surface profiles are available for mesh A, they are termed set A1 and set A2. The profiles for mesh C are termed set C and roughness values for all three sets of surfaces are given in Table 8.2 and illustrated in Appendix D.

Table 8.2 – Ra values in μm for gear set A1, set A2 and set C

	Set A1	**Set A2**	**Set C**
Pinion	0.18	0.20	0.19
Wheel	0.096	0.090	0.23

The results for the mesh of set A1 is shown in Figure 8.21. The rough surface results are compared to the smooth surface results and it can be seen that reasonably good agreement is seen in the load, maximum pressure and maximum temperature responses between the rough and smooth results. The minimum film thickness is very different, where the roughness of the surfaces changes the response considerably as might be expected.

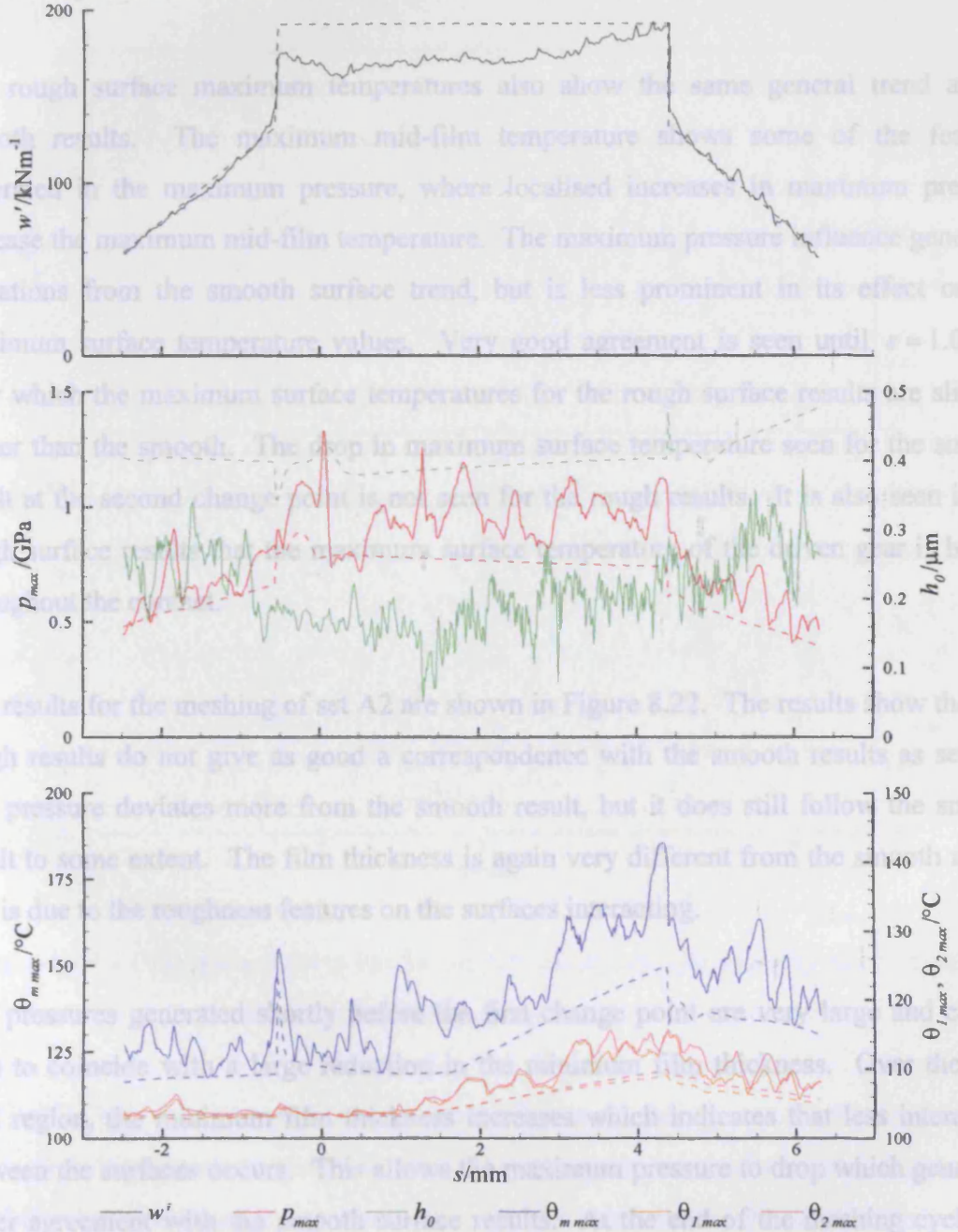

Figure 8.21 – Comparison of results for set A1 (solid) with the smooth results (dashed) over the meshing cycle

The maximum pressure generated is influenced by the surface features, but it is seen to follow the same trend as the smooth result. The maximum pressure gradually rises at the start of the meshing cycle until the first change point is reached. The load increases slightly, but the bump features seen in the smooth result is much more prominent in the rough surface result. The maximum pressure remains almost constant over the high load region and then falls when the second change point is reached. After the change point the maximum pressure continues to fall.

The rough surface maximum temperatures also show the same general trend as the smooth results. The maximum mid-film temperature shows some of the features generated in the maximum pressure, where localised increases in maximum pressure increase the maximum mid-film temperature. The maximum pressure influence generates deviations from the smooth surface trend, but is less prominent in its effect on the maximum surface temperature values. Very good agreement is seen until $s = 1.0\,\text{mm}$, after which the maximum surface temperatures for the rough surface results are slightly higher than the smooth. The drop in maximum surface temperature seen for the smooth result at the second change point is not seen for the rough results. It is also seen in the rough surface results that the maximum surface temperature of the driven gear is higher throughout the contact.

The results for the meshing of set A2 are shown in Figure 8.22. The results show that the rough results do not give as good a correspondence with the smooth results as set A1. The pressure deviates more from the smooth result, but it does still follow the smooth result to some extent. The film thickness is again very different from the smooth result, this is due to the roughness features on the surfaces interacting.

The pressures generated shortly before the first change point are very large and can be seen to coincide with a large reduction in the minimum film thickness. Over the high load region, the minimum film thickness increases which indicates that less interaction between the surfaces occurs. This allows the maximum pressure to drop which generates better agreement with the smooth surface results. At the end of the meshing cycle the minimum film thickness drops generating more surface interaction, leading to contact. The increase in the surface interaction generates very large pressures, which dramatically increase the maximum mid-film and maximum surface temperatures.

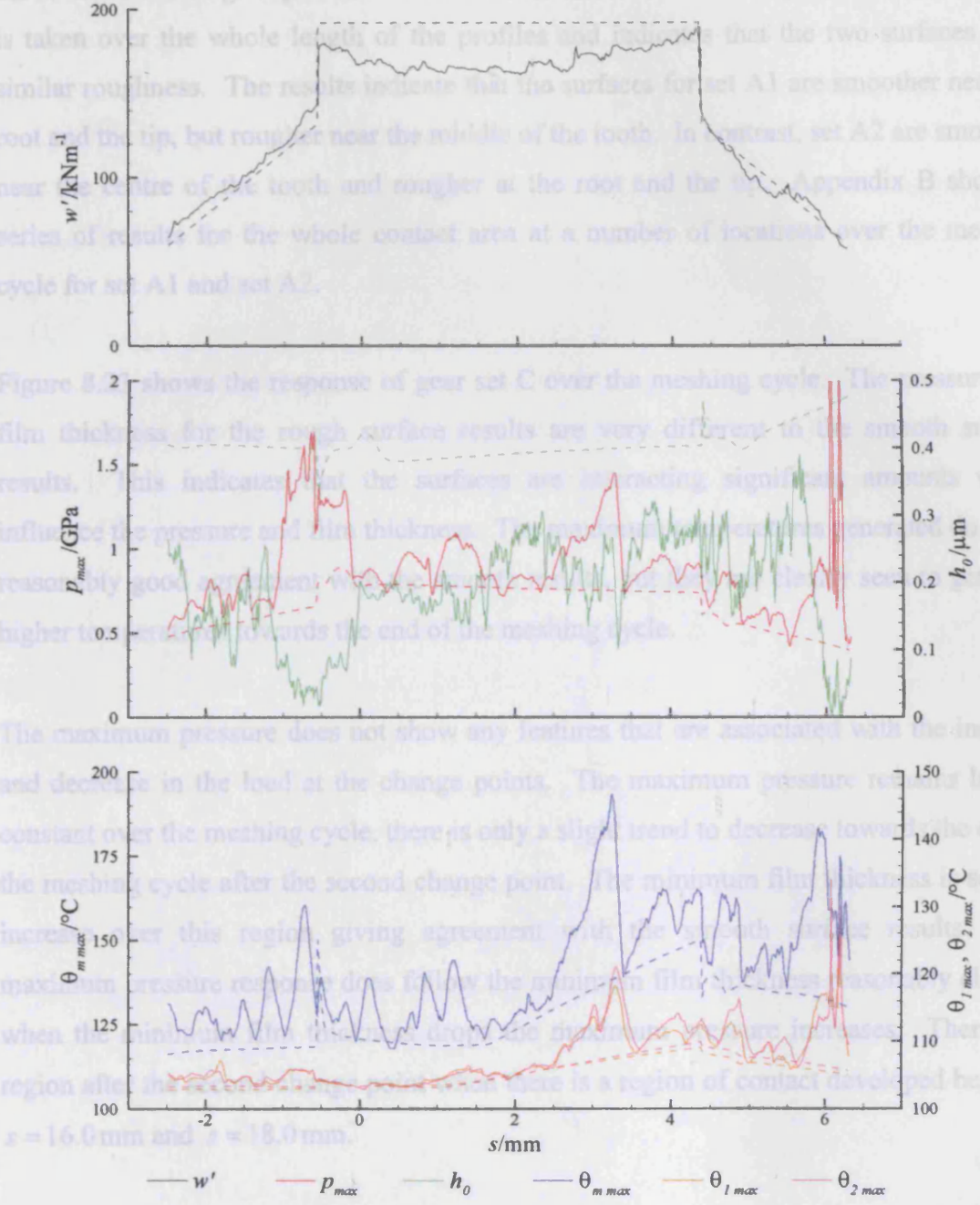

Figure 8.22 – Comparison of results for set A2 (solid) with the smooth results (dashed) over the meshing cycle

It is clear that far more interaction is generated for set A2 than for set A1, but the amount by which the interaction increases is not expected considering the Ra values in table 8.2. The difference between the two results is quite large considering the small difference in the Ra value. The results for set A1 show greater variation in maximum pressure and thinner minimum film thicknesses in the high load region of the meshing cycle. In contrast set A2 develops thicker films and lower pressures in the high load region, but has

thinner films and higher pressures at the start and end of the meshing cycle. The Ra value is taken over the whole length of the profiles and indicates that the two surfaces have similar roughness. The results indicate that the surfaces for set A1 are smoother near the root and the tip, but rougher near the middle of the tooth. In contrast, set A2 are smoother near the centre of the tooth and rougher at the root and the tip. Appendix B shows a series of results for the whole contact area at a number of locations over the meshing cycle for set A1 and set A2.

Figure 8.23 shows the response of gear set C over the meshing cycle. The pressure and film thickness for the rough surface results are very different to the smooth surface results. This indicates that the surfaces are interacting significant amounts which influence the pressure and film thickness. The maximum temperatures generated do show reasonably good agreement with the smooth results, but they are clearly seen to generate higher temperatures towards the end of the meshing cycle.

The maximum pressure does not show any features that are associated with the increase and decrease in the load at the change points. The maximum pressure remains largely constant over the meshing cycle, there is only a slight trend to decrease towards the end of the meshing cycle after the second change point. The minimum film thickness is seen to increase over this region giving agreement with the smooth surface results. The maximum pressure response does follow the minimum film thickness reasonably closely, when the minimum film thickness drops the maximum pressure increases. There is a region after the second change point when there is a region of contact developed between $s = 16.0\,\text{mm}$ and $s = 18.0\,\text{mm}$.

The temperatures that are generated for the rough case before the first change point are very similar to the smooth results. The same general trends are seen in both the surface and mid-film temperatures, where all are seen to decrease as the pitch point is approached and then all begin to increase after the pitch point is passed. The temperatures generated for the rough surfaces do not show the features that are seen in the smooth result at the change points. After the second change point the temperatures for the rough surface results are mostly much higher than the smooth results, this is due to the pressure being higher than the smooth result.

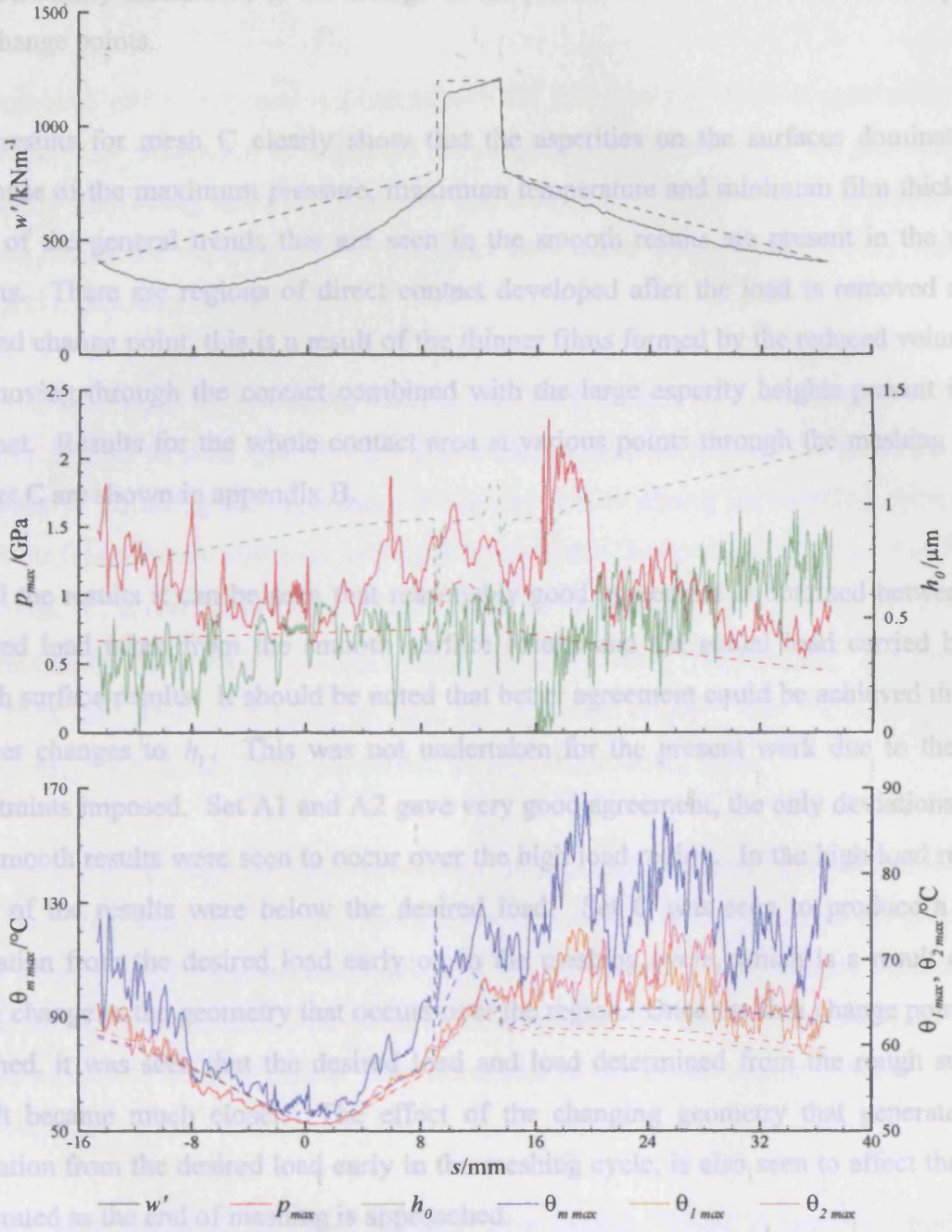

Figure 8.23 – Comparison of results for set C (solid) with the smooth results (dashed) over the meshing cycle

The maximum mid-film and maximum surface temperatures do not show the same features that are seen at the change points when rough surfaces are introduced. For the smooth surface conditions, the changes in load produce film thickness and pressure features that influence the temperatures. However, when rough surfaces are present they generate changes in the pressure and film thickness that are larger than the changes introduced at the change points. The result is that the maximum temperatures developed

are reasonably insensitive to the change in the pressure and film thickness developed at the change points.

The results for mesh C clearly show that the asperities on the surfaces dominate the response of the maximum pressure, maximum temperature and minimum film thickness. Few of the general trends that are seen in the smooth results are present in the rough results. There are regions of direct contact developed after the load is removed at the second change point, this is a result of the thinner films formed by the reduced volume of oil moving through the contact combined with the large asperity heights present in the contact. Results for the whole contact area at various points through the meshing cycle for set C are shown in appendix B.

In all the results it can be seen that reasonably good agreement is obtained between the desired load taken from the smooth surface results and the actual load carried by the rough surface results. It should be noted that better agreement could be achieved through further changes to h_1. This was not undertaken for the present work due to the time constraints imposed. Set A1 and A2 gave very good agreement, the only deviations from the smooth results were seen to occur over the high load region. In the high load region, both of the results were below the desired load. Set C was seen to produce a large deviation from the desired load early on in the meshing cycle, which is a result of the large change in the geometry that occurs over the region. Once the first change point was reached, it was seen that the desired load and load determined from the rough surface result became much closer. The effect of the changing geometry that generates the deviation from the desired load early in the meshing cycle, is also seen to affect the load generated as the end of meshing is approached.

The effect of asperity interactions are seen in the load for sets A1 and A2, where there is a noise component generated in the load. This effect is more noticeable in the results for set A2. The effect of asperity interactions are not as clearly seen in the results for set C as the scale required to present the load hides the changes that occur in the load. Even so, there are slight instances where the load is seen to fluctuate, but the amplitude of the changes are much smaller than that seen for set A1 and A2.

8.6 Summary

An approach has been presented that allows the full meshing cycle of gear teeth to be modelled. The simulation incorporates the changes in the geometry of the gear teeth, the change in the rolling and sliding speeds and also allows different load variations to be applied during the meshing cycle. The load variation is defined using a start load, end load and a maximum load.

The simulation can be run using smooth gear teeth or can be run with real roughness profiles applied to each gear tooth. The smooth approach follows the prescribed load variation by adjusting the separation of the two bodies during the meshing cycle. The rough surface solution uses a different approach where the variation of the first node film thickness, h_1, is read in at four locations during the meshing cycle. The variation of h_1 is determined for the remainder of the meshing cycle by linear interpolation. The first node film thickness gives information on how the separation changes over the meshing cycle, thus giving the different loads prescribed through the meshing cycle.

In order to run the rough surface simulations, an approach was developed that allowed the roughness data taken from a gear tooth to be related back to the points on the gear tooth from which it came. When the simulation is run, the contact is centred upon a known part of the involute, which allows the point in the roughness file that corresponds to the centre of the contact to be determined. The remainder of the points can be determined using trigonometry.

Results were presented for three different gear meshes. Mesh A involves small high speed gears operating at relatively low load, mesh B involves larger gears operating under a higher load and also operating at modest speed, and mesh C incorporates large gears that move slowly and carry a very high load. Due to the different designs of each gear mesh, they developed very different load variations and responded in different ways.

The equivalent roller and true involute geometries were compared for each mesh under isothermal conditions. It was seen that the involute develops slight differences in the maximum pressure and minimum film thickness developed throughout the meshing cycle.

Despite the slight differences in the results the two geometries show the same general features. Investigating the actual pressure and film thickness developed at a point in the meshing cycle, it was seen that the constant change in the curvature of the involute leads to a variation in the film shape. The equivalent roller approach overestimates the radius of curvature in the direction of the pinion root and underestimates the radius of curvature in the direction of the pinion tip.

The difference between the equivalent roller and involute geometries is greater near the pinion root as the radius of curvature is smallest in this region. The results at the pinion tip are very similar, in some cases the two are indistinguishable. The effect of the involute geometry is to develop slightly thicker films throughout the contact region and also to develop higher pressures. The pressures generated for the involute are narrower than the pressures developed for the equivalent roller approach, which is the reason for the increase in the maximum pressure. The difference between the two results becomes greater when higher loads are used.

Both geometries showed that there is a general trend for the film thickness to increase as the mesh progresses, which is a result of the higher surface velocities and larger radius of relative curvature. The maximum pressure was also seen to be lower at the end of the mesh than at the beginning, even though the applied loads were the same. This is a result of the larger radius of curvature developing a larger contact area, thus the pressure spreads out and reduces in magnitude.

The changes in load that occur at the change points give similar features in all of the meshes. At the first change point, where the load increases, the contact area spreads quickly and the pressure increases. This generates an immediate drop in the minimum film thickness and also traps a large volume of liquid in the inlet. The volume of liquid trapped moves through the contact at the mean speed of the oil, locally increasing the film thickness as it moves. The trapped oil has a localised pressure increase associated with it. This entrapment and associated pressure give rise to the increase in the minimum film thickness and maximum pressure seen after the change point. The second change point gives rise to a reduction in the contact area and a drop in pressure. This causes the minimum film thickness to increase, but also loses a volume of oil from the inlet as the contact area reduces. This creates a reduction in entrainment generating a localised

reduction in film thickness. The dip formed in the film thickness also generates a drop in the pressure, located at the centre of the feature. The film thickness dip and the localised pressure reduction both move through the contact, giving rise to the reduction in film thickness and pressure seen after the second change point.

The effect that the temperature has on the pressure and film thickness developed through the contact was also investigated. The temperatures generated within the contact area reduce the viscosity and lead to smaller film thicknesses. The thermal effects do not remove the features that are generated by the change points, but do cause the magnitude of the features to reduce. The temperatures also act to reduce the rate at which the film thickness increases over the meshing cycle.

The maximum temperatures developed throughout the meshing cycle were different for the different meshes. Generally, it could be seen that the maximum temperatures fell as the pitch point was approached due to a reduction in the shear heating. The first change point generated a rapid increase in the maximum temperatures, most noticeably for the mid-film temperature, as a result of the large compressive heating effect. Through the high pressure region all the models showed that the maximum mid-film and surface temperatures increased linearly as the slide roll ratio and velocity increased. At the second change point the maximum temperatures fall due to the de-compressive cooling effect generated by the fall in the pressure. After the second change point the response of mesh A is different to meshes B and C. The maximum temperatures in mesh A continue to reduce after the second change point. However, the maximum mid-film temperatures for meshes B and C both increase slightly before reducing over the remainder of the meshing cycle.

All the gear meshes considered are designed to operate in a speed reduction arrangement. This means that the pinion drives the wheel, generating a reduction in speed and increase in torque. This arrangement was changed so that the wheel was driving the pinion which results in the gears operating as a speed increase arrangement. This change causes the first contact to occur between the pinion tip and the wheel root and the contact moves towards the wheel tip. For the speed reduction case the first contact point occurs between the wheel tip and the pinion root and the contact moves towards the pinion tip.

The two different cases were examined for all the meshes and it was found that the same features existed whether the gears were run in a speed increase or speed reduction arrangement. The first and second change point occurred at different points on the teeth, but the features generated by each change point remained the same. It was seen for meshes B and C that the change to a speed increase arrangement generated slightly thicker films throughout the meshing cycle. However, mesh A developed thicker films at the pinion tip and thinner films at the pinion root for the speed increase arrangement than for the speed reduction. The large difference between the two results can only occur due to the squeeze film effect associated with the large increase or decrease in load. When considering the change in the manner in which the gear pair operate, it was found that the speed reduction case developed the lowest minimum film thickness.

Results were also presented for meshes A and C using roughness data taken from actual gear teeth used in these meshes. There were two different sets of profiles available for mesh A and one set of profiles available for mesh C. The two sets of profiles for mesh A were termed A1 and A2 and the profiles for mesh C were termed set C.

Considering the results for mesh A, it was seen that the features on the surfaces reduce the minimum film thickness, but good agreement with the smooth surface maximum pressure response is developed. The maximum pressures generally increased at the beginning of the mesh and remained constant through the high load region before reducing after the second change point. The interaction of the roughness features on the two surfaces give rise to small increases in the maximum pressure response, which are also reflected in the maximum mid-film temperatures developed. The maximum temperatures developed do remain similar to the smooth surface results, only slight changes occur due to the variation in the pressure response.

The two sets of surfaces give different results. Set A1 gave good agreement with the smooth surface pressure through the low load regions. However, set A2 gave good agreement with the smooth surface pressure through the high load region. This was also reflected in the minimum film thickness values developed through the meshing cycle, where the minimum film thickness was thicker in the low load regions for set A1 than the high load region. The film thickness was greater through the high load region for set A2 than in the low load region. Set A2 developed a region of contact near the end of the

meshing cycle at the pinion tip, which generated high pressures and temperatures at these locations. It was also seen that set A2 came close to developing contact near the first change point, which also generated high pressures.

The results for mesh C did not show any of the features or trends seen in the maximum pressure and minimum film thickness for the smooth result. The maximum pressure remained large throughout the meshing cycle and the minimum film thickness fell at many locations. It was also seen that contact was developed at two different regions of the meshing cycle. The maximum pressure and minimum film thickness response indicate that significant surface interaction occurs. Despite the large differences in the maximum pressure and minimum film thickness, the maximum temperature response remained reasonably similar to the smooth surface results.

Results for the rough surface problems, showing the pressure and film thickness developed through the contact region at various points during the meshing cycle, are shown in Appendix B. It can be seen that the roughness features are seen to move through the contact and that the pressure and temperature show no signs of any oscillations. Since the pressure and temperature are not seen to oscillate, it gives confidence that the approach adopted is representing the changes in the surface roughness.

The smooth surface thermal results indicate that the maximum temperature and maximum pressure increase rapidly at the first change point in the mesh. This is the region where the load increases rapidly and also gives rise to a reduction in minimum film thickness. The changes in these parameters give rise to an increase in the possibility of surface fatigue. The increase in the pressure will generate rapid increases in the sub-surface shear stresses which may lead to crack formation or propagation. Also, the increased thermal effects and the reduced film thickness allow a greater degree of surface interaction.

The smooth surface approach works very well for all of the test cases considered here. The load seeking approach does not promote instabilities and follows the desired load closely. The only possible limitation to the approach is the manner in which the problem is started. The first point is solved as a steady state solution, which as far as the temperature is concerned can only be seen as an approximation. The manner in which the

load is applied to this point also needs to be carefully considered. Despite the possible inaccuracies of the first point solution, the manner in which the remainder of the gear meshing cycle develops from this point is reasonably modelled.

The first node film thickness is used for the rough surface simulations. Six values at specific points in the mesh were input into the simulation and used to determine the first node film thickness at any point during the meshing cycle by linear interpolation. This may not be the optimum approach as the load is seen to vary from the desired load quite substantially at locations between the points entered, most noticeably for mesh C. This deficiency may be addressed by using more points to define the variation in the first node film thickness and making further adjustments. This detail was not pursued during the study due to time pressures.

The results obtained for the smooth surface results can be compared to the results presented by Wang et al. (2004) and Larsson (1997). The isothermal and thermal results presented in this Chapter show the same features that are presented by Larsson (1997) and also show good agreement with the thermal results of Wang et al. (2004). However, the results that are presented for by Wang et al. for the isothermal conditions are different to the responses presented in this work and those reported by Larsson. There are variations in the minimum film thickness developed over the meshing cycle outside of the high load region. When the load is increasing at the start of the meshing cycle the minimum film thickness presented in Fig. 5.(b) [Wang et al., 2004] develops very large changes. Similar features are also seen when the load is reducing after the change point. This conflicts with the results presented in this Chapter and those by Larsson, where the minimum film thickness shows a near linear increase as the meshing cycle progresses.

Another point to note is the effect that the thermal conditions have on the central and minimum film thickness presented in Fig 5 [Wang et al., 2004]. During the second half of the meshing cycle the minimum and central film thicknesses for the thermal results are above those of the isothermal results. This response is unlike the results presented in this chapter in figures 8.11, 8.13 and 8.15. This is confusing as the increase in the temperature reported would act to reduce the viscosity and generate thinner films for the thermal case, however, in stark contrast thicker films are formed. The results presented by Wang et al. for the isothermal conditions do not appear correct, hence the comparison

made between the isothermal and thermal results may be incorrect. The results presented here for a variety of different meshes and those presented by Larsson all show consistent features that tend to contradict those seen in the isothermal results of Wang et al. Despite the differences seen outside the high load region, the features generated at the change points by both Wang et al. and Larsson give very good agreement with the results seen in this Chapter.

Chapter 9

Experimental Investigations into 'Running In'

9.0 Introduction

This chapter introduces the experimental work to investigate the process of 'running in'. 'Running in' is generally seen as a beneficial process as it reduces the roughness of the surfaces and therefore improves the possibility of generating an effective oil film. On the other hand this process can create residual stresses that later lead to micropitting. Current knowledge of running in is not based on science and understanding and there is a need for detailed study of the phenomenon.

The experimental work investigates the contact between various rough and smooth disks. The rough disks are axially ground, so that the finish is transverse to the entrainment direction, with a typical roughness value, R_a, of $0.40\mu m$. The smooth disks are initially prepared in the same manner as the rough disks but are then superfinished giving $R_a = 0.05\mu m$ typically.

The experimental rig that is to be used was designed by Mr Marc Alanou of the School of Engineering and in the current work has undergone modification to allow insight into the process of 'running in'. These modifications allow on-line electrical resistance measurements to be made as well as 'in situ' measurements of surface profiles of the disks using a Talysurf profilometer.

9.1 Experimental Rig

The rig is a two disk machine, using the same ideas as presented by Merritt (1935) and was originally designed to investigate micro-pitting. Figure 9.1 shows the general layout of the test head, illustrating the disks and the manner in which the load is applied to them.

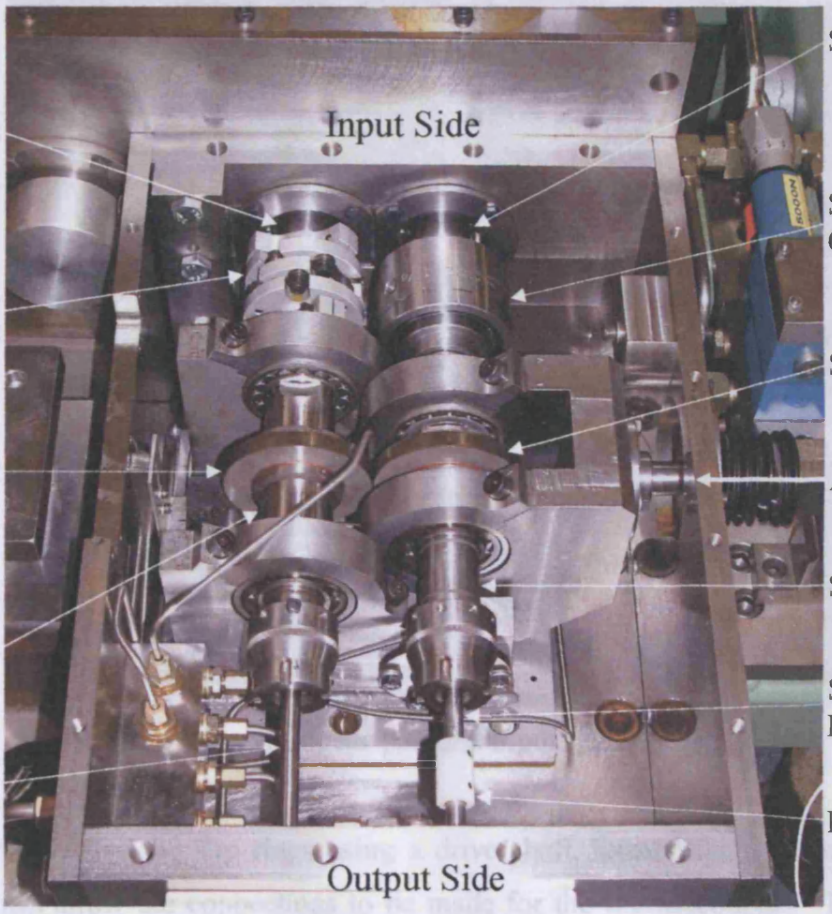

Figure 9.1 – General view of the test head

The loading mechanism used is a hydraulic ram, this allows the applied load to be varied whilst minimising shock loading, which can result from the placement of dead weights as used in some other disk machines. The hydraulic ram can provide a force of up to 8KN to the disks.

The disks are firmly mounted onto hardened steel shafts, termed test shafts. Each test shaft is driven through a coupling, termed the input coupling, which is driven by an input shaft. The two input shafts are interconnected using a pair of spur gears. The spur gear pair can be changed allowing the relative speeds of the two shafts to be changed. This

assembly is shown in Figure 9.2. The ratio of these gears for the cases considered here is fixed at 1.483.

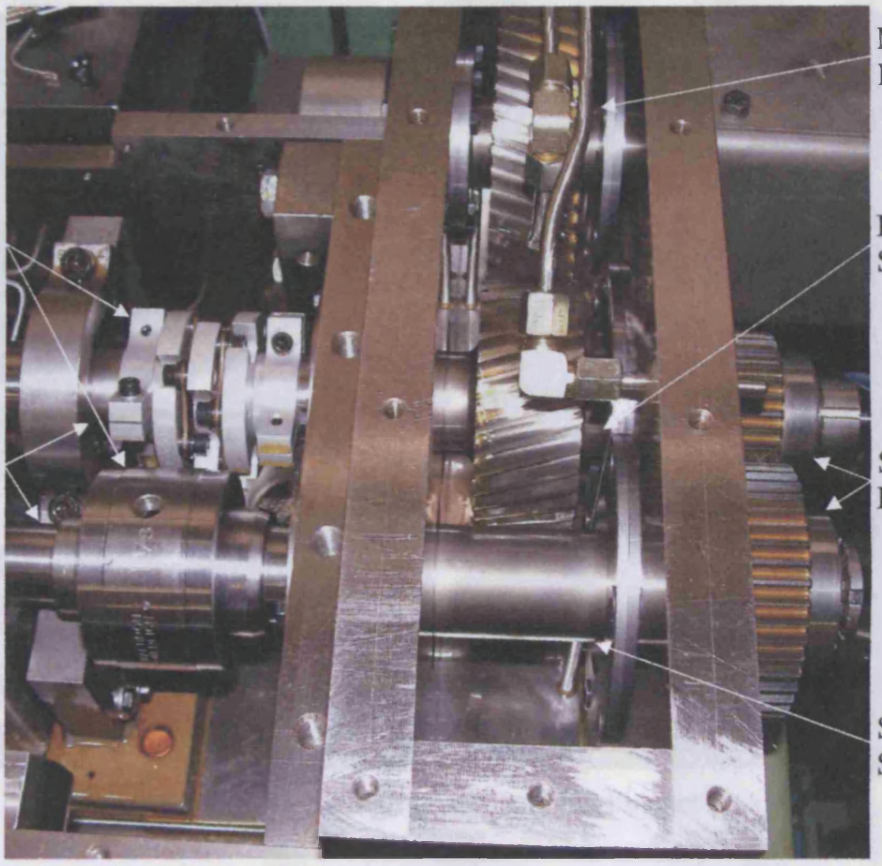

Figure 9.2 – Spur gear pair and input shaft arrangement

The test shafts drive the slip rings using a drive shaft, termed the slip ring drive shaft. The slip rings allow the connections to be made for the thermocouples. The test shafts and the slip ring drive shafts have a central hole bored. This allows the thermocouple wire to run from the disk through the test shaft, through the slip ring drive shaft and into the slip ring itself.

Each of the test shafts is firmly located using a self aligning spherical roller bearing and a plain roller bearing which are held in supports. The fast shaft supports are bolted to the base of the test head, but the slow shaft supports are not located directly onto the base of the test head, instead they are held in a swinging yoke. The yoke constrains the shaft so that it can follow an arc, centred at the pivot of the yoke, allowing the load to be applied

so that the slow disk approaches the fast disk. This arrangement is shown in Figure 9.3 and ensures that the load applied by the hydraulic ram is transferred solely onto the disks.

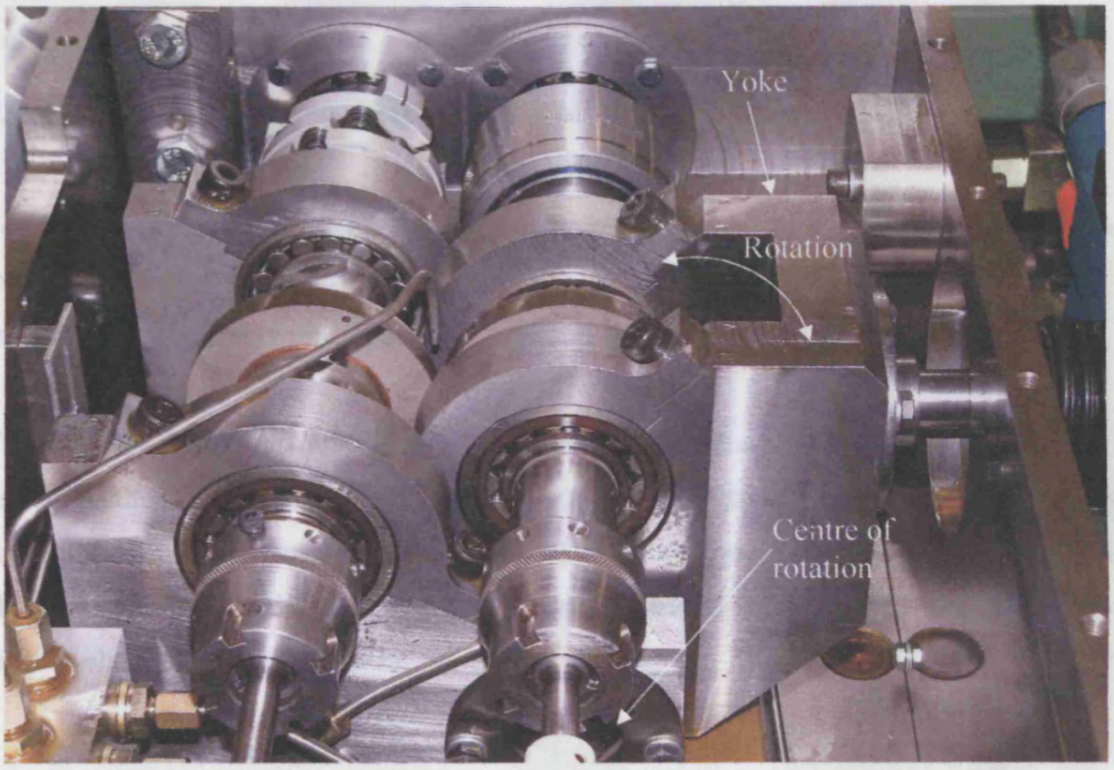

Figure 9.3 – Swinging yoke and loading set up

The lubrication oil used for the tests is MobilJet 2 and the disks used are hardened alloy steel (HV 650) of diameter of 76.2mm (3") with a crown radius in the axial direction of 304.8mm (12"). The motor speed is controlled by an inverter and can be varied to drive the fast shaft at speeds between 1500 and 6000rpm, which gives slow shaft speeds of 1012 to 4047rpm.

9.2 Rig Modifications

The modifications made to the rig in the current study can be divided into two categories, those allowing electrical contact resistance measurements to be made and others allowing 'in situ' measurements to be made of the surface profiles of the disks.

The electrical resistance modifications can also be broken down further, the mechanical modifications to the rig enabling measurements to be taken and the computational and electrical modifications that enabled acquisition of the signals.

9.2.1 Electrical Contact Resistance Mechanical Modifications

To allow measurements of the electrical contact resistance between the test disks, one of the test shafts was electrically isolated from its surroundings. This was achieved by a variety of different methods which are outlined below:

The yoke in which the slow speed shaft is mounted was electrically isolated at its pivot by replacing the conventional pivot ball bearings. The ball bearings were replaced with ceramic rolling element bearings, where Silicon Nitride was used for the ceramic elements. A Silicon Nitride ball was also used to isolate the loading push rod from the yoke as shown in Figure 9.4.

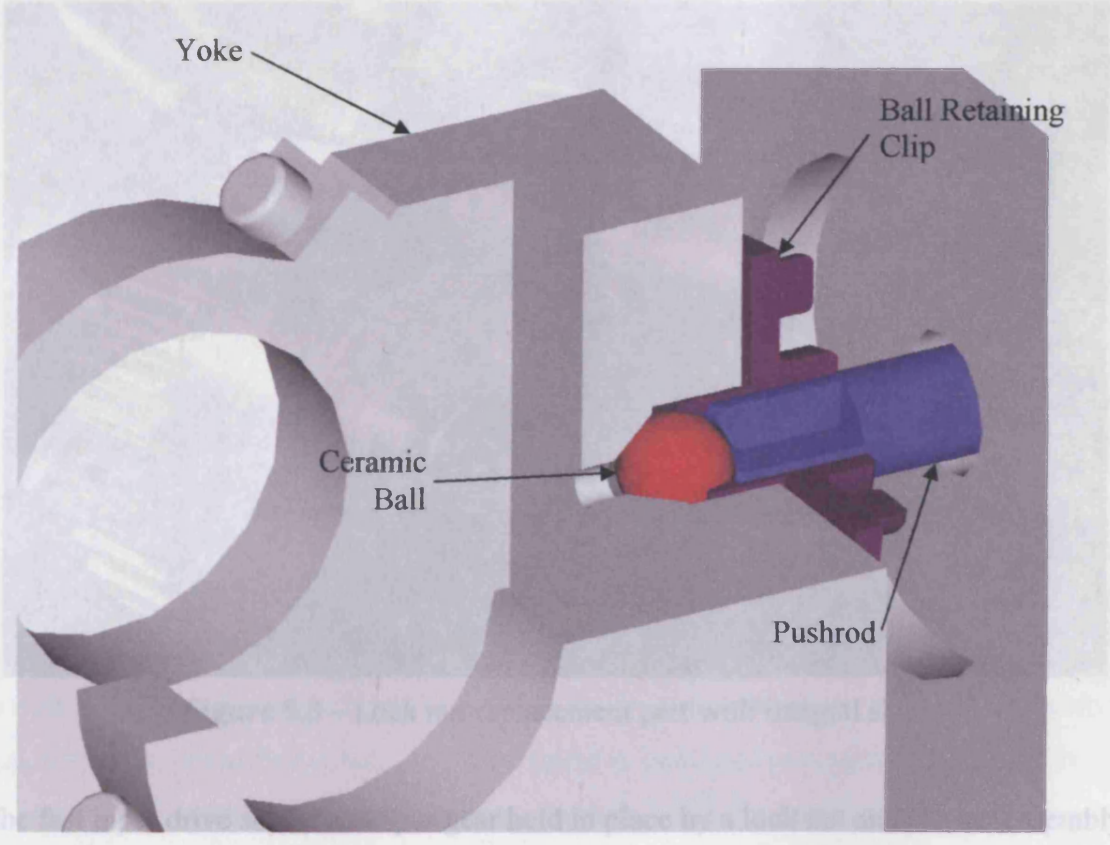

Figure 9.4 – 3D CAD drawing of ball and loading rod

The lower speed test shaft was isolated at its driven end by a PEEK sleeve and washer fitted to the hub of the gear coupling. A PTFE muff coupling was used between the test shaft and the slip ring drive, this coupling is shown at the bottom of figure 9.1.

The electrical connections were made to each of the disks using a brass washer with a solder tab. The washer is clamped between the disk and its locating shoulder on the test shaft. The wires from the connections can be attached to the solder tabs and then run down the central bore in the shaft and through the slip ring drive shaft.

During the course of the experiments the electrical resistance between the disks changes rapidly. In order to detect these changes a shaft encoder was fitted to allow the resistance to be sampled at specific points during the revolution. A rigid (zero backlash) connection between the encoder and the test disk was necessary, requiring the fast input gear coupling to be replaced with a rigid diaphragm coupling. The fast shaft was selected because its axis is fixed and it has good alignment with its drive shaft, which enables the use of a torsionally rigid coupling.

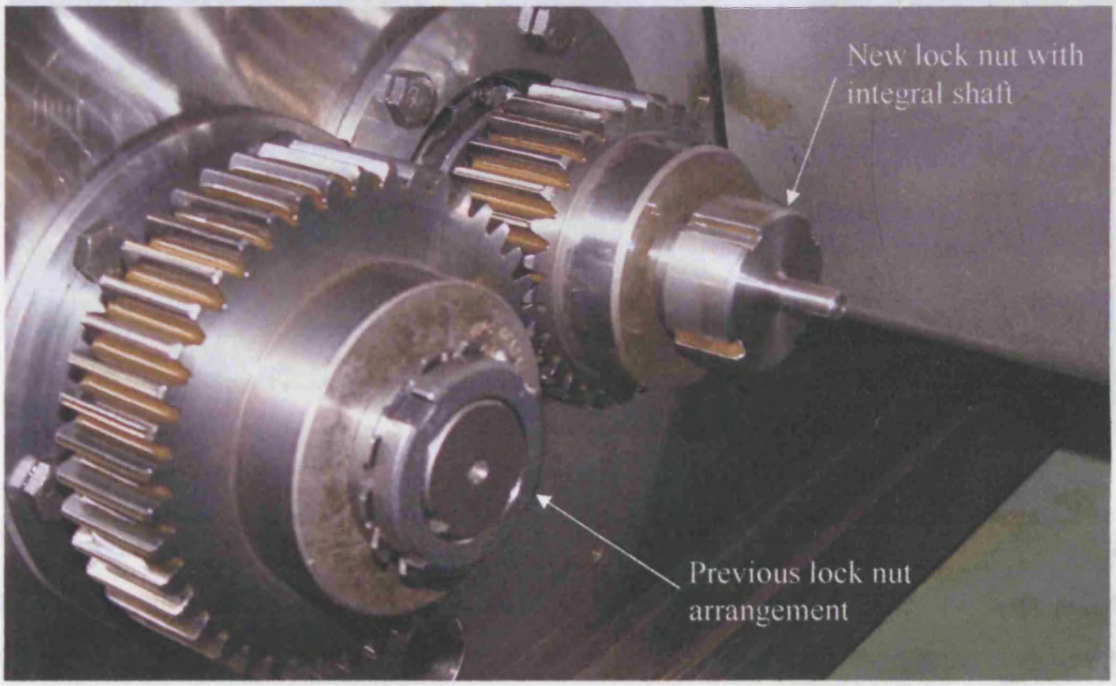

Figure 9.5 – Lock nut replacement part with integral shaft

The fast input drive shaft has a spur gear held in place by a lock nut and washer assembly. The shaft is hardened making it difficult to find some way of attaching another shaft to

drive the shaft encoder. The solution was to manufacture a new nut with an integral spigot shaft as shown in Figure 9.5.

The shaft encoder is designed to clamp onto some form of drive shaft. The connection between the shaft encoder drive shaft and the spigot shaft is made using a flexible coupling, which is shown in Figure 9.6. The flexible coupling is a zero backlash coupling, which ensures that there is a backlash free connection between the test shaft and the shaft encoder.

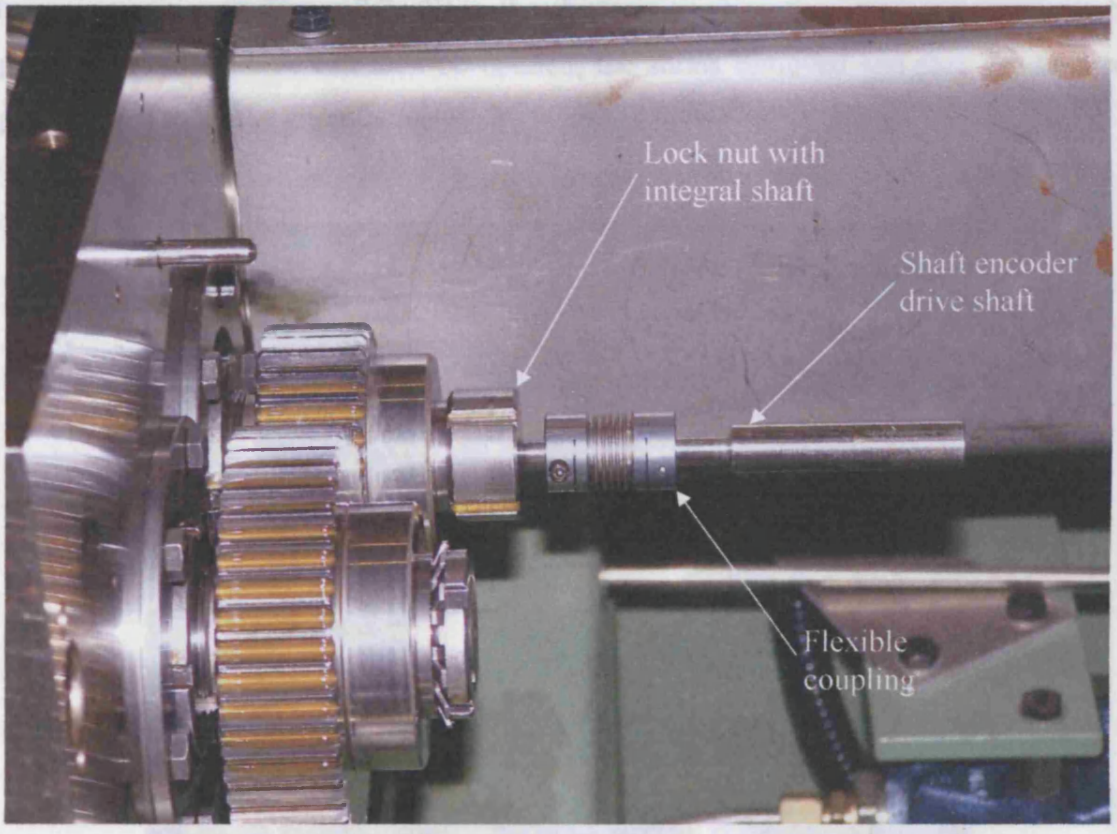

Figure 9.6 – Shafts and coupling used to drive the shaft encoder

These modifications cover those that were necessary to insulate the slow shaft as well as those necessary to allow data to be sampled throughout the revolution. There was one further consideration needed. Since running in is being modelled, the first few revolutions are critical and great care has to be taken to ensure that the operating conditions are controlled accurately. The speed is computer controlled and the disks can be run up to speed without contacting, however, the manner in which the load is introduced must be considered in more detail.

The magnitude of the load applied to the disks is manually controlled by a variable pressure relief valve. It was decided to set this control valve initially and introduce a ball valve after the control valve that could be opened to remove the load and then closed subsequently to re-apply the load in a repeatable manner.

The control valve pre-set cannot be determined by loading the two disks together as it will deform the asperities and influence the results obtained. Instead, two spacer shims were manufactured that can be located between the fast test shaft and the slow test shaft bearing support so that the disks are separated when the load is applied. The shims are shown in position in Figure 9.7. With the shims in position the appropriate load can be applied. The ball valve can then be opened, the shims removed and then the load re-applied when the disks are rotating.

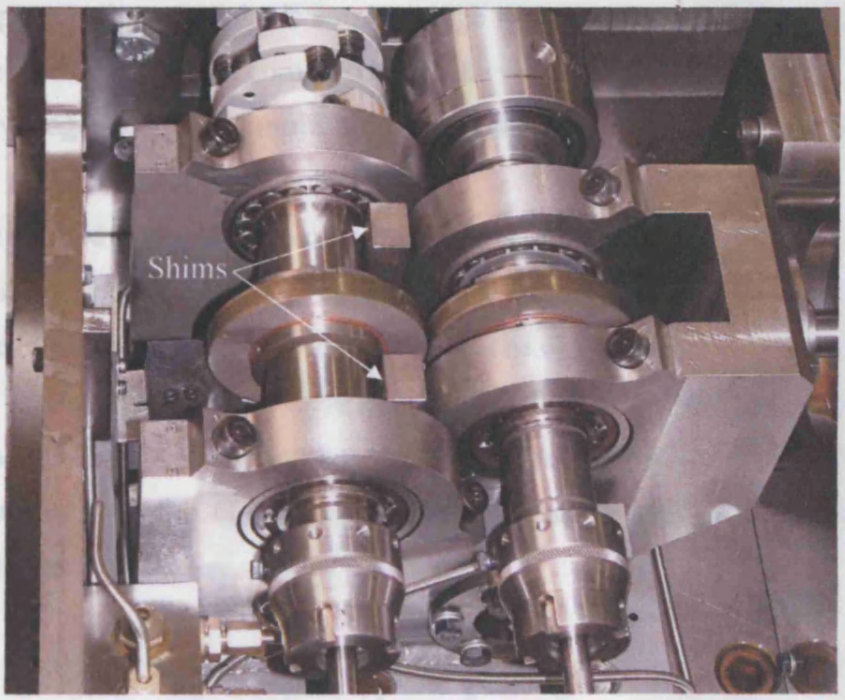

Figure 9.7 – Shims used to separate the disk during the setting of the load

9.2.2 Electrical Contact Resistance Electrical Modifications

The technique employed for electrical contact resistance measurements in this work is based on that conducted by Furey (1961). The basic layout of the circuit is shown in Figure 9.8.

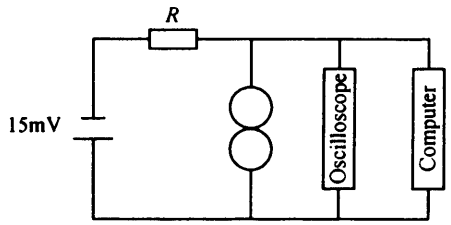

Figure 9.8 – General layout of electrical contact resistance circuit

The voltage applied to the disks is 15mV. With this voltage conduction across the fluid film that separates the two disks is small enough to ensure the amount of metallic contact is not influenced. This can be compared with the work undertaken by Furey, where it was shown that voltages of 1.5V or more effect the amount of metallic contact observed.

The voltage applied to the disks is obtained by dividing a 15V supply across 2 resistors in series to form a potential divider. This voltage is then used with a series of resistors allowing the internal resistance, *R*, to be varied. The circuit layout is shown in Figure 9.9. This allows the resistance circuit to be modified so that the contact resistance of the disks that gives a mid-scale voltage output can be changed. This change can be done very quickly using one of the connections A – D, ensuring that no modification to the components used is necessary.

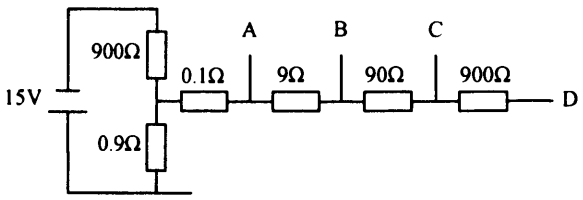

Figure 9.9 – Resistance Bridge layout

The voltage measured across the test disks is transferred through slip rings. This can introduce another resistance into the circuit and it is important that it is not on the scale

that is examined. Furey (1961) concluded that the resistance of the instances that corresponded to contact 'was found to be approximately equal to the static resistance'. This means that if the disks are cleaned (have the oil removed) and loaded against one another, the contact resistance measured will be roughly equal to the resistance produced by contact when the disks are running under load with a lubricant. This conclusion forms the basis of the slip ring resistance tests.

The preliminary resistance tests involved cleaning the disks and then loading them against one another with a force of 4KN. The resistance between any two points in the circuitry can be measured using a digital multimeter and the results analysed to show the influence that different components have on the resistance. The resistance was measured at four points: across the disks, across the brass washers, across one slip ring and finally across both slip rings. Measuring one slip ring's influence then that of the other allowed validation of whether either slip ring was developing high resistances that could affect the results. This test was repeated several times, each time the shafts were rotated to examine a different part of the slip ring contacts.

The resistances measured across the disks was generally less than 6Ω this may be compared to the increase in resistance when measuring across both slip rings of 0.1Ω to 0.3Ω. Both slip rings gave similar results, this indicates that both are in good working order and that one slip ring is not responsible for the difference in the resistances measured. The results from these preliminary tests clearly show that the slip ring resistances had very little influence on the resistance measured across the disks when contact occurs. In addition to examining the slip rings, the brass washers that connect the resistance measurement circuit to the disks were tested. For all the tests considered there was no difference between the resistance measured across the washers and that measured across the disks themselves.

Preliminary work was undertaken into the influence that the surrounding electronic interference has on the signals that are to be recorded. Although the circuitry was enclosed in a sealed aluminium box and all the cables were screened, there was still evidence of electrical interference. The main source of this interference was found to be the 3 phase inverter that drives the motor. A mains filter was fitter to the inverter which eliminated this problem.

Despite the removal of the noise from the inverter using the filter, there were still other noise signals present. A spectrum analyser was used and the signals could not be linked to a specific frequency and appeared to be random. The source of the signals could not be identified which meant the only way to remove the signals was to use a low pass filter in the circuitry between the disks and the recording devices.

In addition to the modifications to the rig and the inclusion of new circuitry there was also a need to implement a data acquisition system. The signals monitored are those corresponding to the electrical contact resistance, friction, load, optical reflection (for future projects) and bulk temperature of the two disks. The system employed samples these signals when triggered by the shaft encoder. The shaft encoder provides various trigger pulses for the sampled signals; there are 5000 trigger pulses per revolution generated on channel 1 and another trigger pulse once per revolution on channel 2. The major concern was the speed at which these signals were to be sampled as the system needed to operate at a maximum fast shaft speed of 6000 RPM.

The major problem with this application was the number of channels and the speed at which they were sampled. The only signals that changed rapidly over a revolution were the electrical resistance and the optical reflection, other signals such as the bulk temperature, friction and load vary relatively slowly and are sampled at a lower rate. Applying these considerations, the best arrangement was determined where two different data acquisition cards were used. One card sampled the high speed signals (electrical resistance and optical reflection) and another sampled at lower frequencies (friction, temperature, load). The cards chosen were National Instruments M-series data acquisition cards, the high speed acquisition is handled by the NI-PCI 6250 and the low speed by the NI-PCI 6220.

The data acquisition software was specially designed. Since the electrical resistance was sampled at fractions of a revolution, outputting this data for a whole test would stretch any modern computer system to the limits, especially at the high speeds considered. To overcome this situation it was decided to sample the contact resistance at fractions of a revolution, but to average the values taken over the revolution, giving an indication of how the amount of contact over each revolution changes as the test progresses. At certain

revolutions, specified by the user, the data acquired through the revolution needs to be recorded.

Furey (1961) defined the percentage metallic contact as the amount of time that there is contact over the total time considered, which in his work was 10 ms. Here, however, the number of contact points through a revolution could be determined instead. This was done by monitoring the number of points sampled during the revolution that have a resistance below a value specified by the user.

Similarly the optical reflection signal was sampled at fractions of a revolution and could also be averaged over the revolution, this gives some indication of how the surface changes as the disks run together.

The software also inferred the speed of the fast shaft from the once per revolution pulse, this was used as a check to ensure that the shaft encoder was running at the same speed as the fast shaft and there was no slipping in the couplings connecting them. The necessity for this was emphasised by a failure of the flexible coupling linking the shaft encoder to the input shaft.

9.2.3 Electrical Contact Resistance Software Development

The acquisition of the signals during the course of the experiments was controlled by the LabVIEW software package. The package allows the user to create a 'virtual instrument' using a graphical user interface approach. Care was necessary in the programming of the environment, especially since the signals were to be acquired at different rates and on different cards. The necessity to output the data at specific revolutions also makes the programming slightly more complicated.

For the electrical contact resistance measurements, a resistance value must be determined that signifies contact is occurring. During the acquisition, the data were compared to this value and the number of points per revolution below this value determined, allowing the percentage contact to be calculated for each revolution.

The environment was set up so that the user specifies the filenames and paths for the output files to be saved as well as other information relating to the test itself, such as rotational speed, test number, fast and slow disk number, their surface finish and material. The revolutions for which the data acquired through the revolution is to be recorded to file, the resistance value below which contact is considered to occur and the voltage applied to the contact also have to be specified.

During the acquisition the friction and load were averaged over a number of revolutions, this is due to the noise on the signals causing some slight variation in the values. The noise could not be sourced, instead the signals had to be averaged. However, since the load and friction are slow transient signals they are not affected by the averaging.

The temperature of the disks, friction, load, average optical reflection, average contact resistance voltage and the percentage contact were all plotted. All of these values are written to file once the user specifies a start point to record the data. The actual load is also written to file so that the point at which the load is applied and removed can be determined.

During the setup of the system some problems were encountered. The first computer used did not have a motherboard that provided a 3.3V power line to the PCI slots, this meant that neither card would work, necessitating the use of a more modern computer. The terminal block that was used with the slow data acquisition card had a built in cold junction temperature sensor. However, due to the cards purchased being modern designs, there was an incompatibility between the two components. This meant that the cold junction compensation could not be used and thermocouple amplification units with built in cold junction compensation were necessary. The units chosen were the Analogue Devices AD594AQ which provide an output scaled to 10mV/°C compensated to a cold junction.

It was also found during the development that the switching from channel to channel was causing some loss in accuracy of the measurements made. The fast data acquisition card is programmed to acquire data from two different channels with the shaft encoder acting as an external clock, which determines when the samples are taken. Since the card only has one analogue to digital converter, it constantly changes the input values to the

converter. The converter has a settling time, which is affected by the impedance of the circuit and is expressed as a fraction of the step size between samples. The contact resistance has an impedance of approximately 100Ω and the optical reflection circuit has a substantially larger impedance of 4kΩ. More importantly the difference between the scales of the two measurements is large. The contact resistance is less than 15mV and the optical reflection is of the order 2V, it was seen at modest speeds that there was an interaction between the two signals. When switching from the contact resistance to the optical reflection, the measurements taken were lower than the actual value and then when switching back to the contact resistance the measurements taken were artificially high.

At first this was thought to be due to a mistake in the programming or the occurrence of some noise within the computer system. A test program was written that sampled only one of the signals, the results produced gave very good agreement with the values taken from an oscilloscope. It could therefore be concluded that the error seen was due to the switching between the two channels sampled.

To prevent this situation two solutions were considered; The first was to reduce the impedance of both circuits which reduces the settling time and hence the error. The second solution was to ensure that both signals have similar ranges. Reducing the impedance of the circuits would be very difficult, it was far easier to ensure that both signals were in the same range. There were two ways to achieve this, one was to use a potential divider circuit to reduce the optical reflection signal down to the same level as the contact resistance, the other was to use an instrumentation amplifier to increase the signal from the contact resistance so that it was on the same scale as the optical reflection output. The potential divider circuit would increase the impedance of the optical reflection circuit, although the values would be on the same scale, this increase in impedance would further increase the settling time. Using an amplifier reduces the impedance of the contact resistance signal as well as reducing the difference between the two signals. This was the better alternative and was easily incorporated. It should be noted that another solution was to purchase a further data acquisition card with simultaneous sampling, but due to the high speed involved this would be expensive.

The amplifier chosen was the AD627AN which gives gains in the range of 5 – 1000, which is set by an external resistor. The resistor used was 2100kΩ which gave a gain of 100, which increases the contact resistance signal to 1.5V. This gain was taken into account in the data acquisition program so that the resulting amplified signal could be related back to the voltage applied to the disks. The amplifier also allowed a reference voltage to be applied to remove the offset created by the amplification. This arrangement removed the error due to the switching from one channel to the other.

When the software side of the acquisition system was being developed it was found that there were problems with the speed at which the data was acquired. Although data were being acquired successfully by the cards and stored to memory, the data were not processed sufficiently quickly by the computer and thus the memory buffer filled up too quickly. Simply increasing the buffer does not solve the problem, as the faster the disks rotate the faster the buffer fills. Also, the longer the test run the more data are stored in the buffer. It was necessary to adapt the programming to make it faster as well as looking into using a more modern computer with a faster processor and higher bus frequency. This allowed the data to be transferred from memory quicker as well as reducing the time to perform the calculations.

The data acquisition program was set up to perform a number of loops, in each loop the data were read from memory and then manipulated to give the desired output. One of the calculations made was to determine the frequency from the once per revolution pulse of the shaft encoder. This was used to infer the speed of the shaft and hence determine whether any slip occurred between the fast input shaft and the shaft encoder. Changing this arrangement so that the frequency was calculated from the 5000^{th} of a revolution pulse, the loop time was halved. This was significant, as it allowed twice the number of revolutions to be acquired without modifying or changing the computer itself.

Other investigations focused on the other aspects that made up the data acquisition environment. The main aim was to determine the best method to acquire and process the data, this was necessary as it could be done in one of two ways. Firstly, the data for an entire revolution were read from memory, stored as an array and the manipulation carried out in various loops. The other alternative was to read the values from memory one sample at a time, manipulate the value and then append it to an array so that it can be

written to file. It was found that the former option was quicker as the computer is not constantly accessing the memory.

Every revolution the data acquired were written to screen in the form of numerical indicators as well as graphs. It was found that this process takes a large amount of time and significant saving could be achieved if the front panel values were not updated each revolution. The program was re-written to update the graphs and indicators at a specific number of revolutions, the value used is entered by the user.

The computer used for the initial tests was an Intel PIII 800MHz machine with 512MB of RAM. When replaced with a newer Intel P4 1.5GHz machine with 512MB of RAM it was found that the execution time halved. This enabled the data to be acquired and processed at the same speed as it was generated up to an operating speed of 3000RPM, which corresponds to 20ms per revolution. At higher rotational speeds the execution time did not reduce, indicating that this is the minimum execution time.

The only way to reduce the execution time further was to either invest in a new computer system, or to acquire one signal at the 5000^{th} of a revolution resolution. Since the optical reflection is incorporated for future experiments to be conducted, it was felt that the acquisition of this signal was not necessary within the scope of this work. This allows the contact resistance alone to be sampled, which reduced the execution time of the program sufficiently to allow data generated at 6000RPM to be processed at the speed at which it became available.

If the optical reflection was sampled in addition to the contact resistance at speeds above 3000RPM, it should be undertaken using a higher specification machine otherwise some of the data may not be available due to the buffer filling up. However, to allow initial work to be undertaken without the necessity to purchase a more modern computer system the software was amended to allow 1000 samples to be taken per revolution. This enables the current configuration to run at speeds of 6000RPM without filling the buffer. This also required no further electrical modification, as a signal was present that gave 1000 pulses per revolution. This signal was generated by dividing the 5000 pulses per revolution signal and was necessary during initial tests using a data logger that couldn't sample at high frequency.

9.2.4 'In Situ' Talysurf Measurements

Surface profiles are taken from the test disks using a Taylor Hobson Form Talysurf Intra instrument shown in Figure 9.10. In order that profiles can be taken at various stages during a running in experiment an in situ arrangement was devised.

Figure 9.10 – Taylor Hobson Form Talysurf Intra

A stable support for the talysurf was required as well as the ability to relocate the Talysurf easily and accurately so that it could measure and re-measure specific parts of either disk. The design had to allow movement in the axial direction so that traces could be taken at different axial positions across the disk, perpendicular to the entrainment direction.

Various arrangements were considered for supporting the Talysurf that would allow measurement of both disks. Initial ideas involved elevating the talysurf, this was necessary as the sides of the test head were higher than the top of the disks which meant, if located horizontally, the Talysurf stylus could not get access to measure either disk. This introduced problems in that some form of complex slide would have to be used to get from one disk to another, plus the Talysurf should ideally be set within a few degrees of horizontal.

For the Talysurf to be set horizontal some material on the side of the test head had to be removed. The 'cut out' made was lower than the top of the disks as the guides that allow the vertical movement of the pick up on the Talysurf are lower than the tip of the stylus and some clearance is necessary.

This modification allowed the Talysurf to be used to measure the fast disk, however, problems were still met when the Talysurf was used to measure the slow disk. The fast disk obstructed the pick up of the Talysurf preventing the stylus being lowered into a position to measure the slower disk. The solution to this was rather simple, requiring the use of a deep recess stylus. The deep recess stylus can measure features located 10.5mm below the bottom of the pick up rather than the standard conical stylus which can only measure features located 0.5mm below the bottom of the pick up.

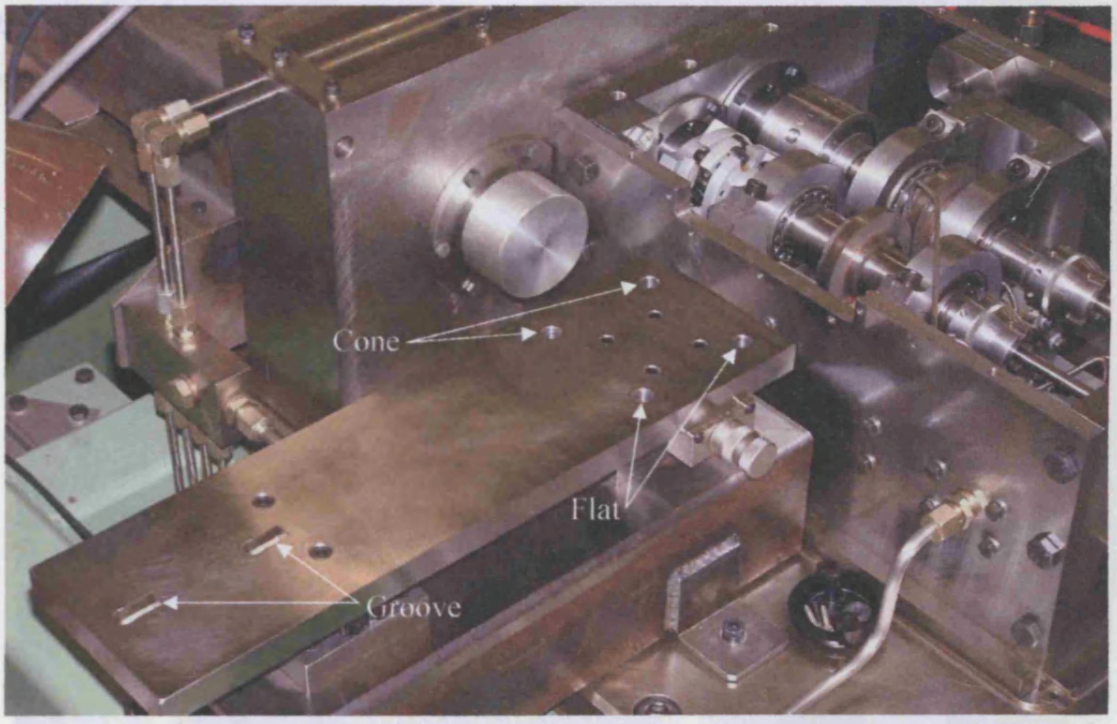

Figure 9.11 – Stage to support Talysurf

When not in use to measure profiles of the two disks the Talysurf was removed from its position close to the rig. A means was therefore provided to allow re-location of the unit so that a series of profiles could be taken, the unit removed and then replaced following the operation of the rig. The stage on which the Talysurf was placed included a three point kinematic location arrangement. This consisted of a vee grove, a conical seating

and a flat that engages with the spherical ends of the feet of the Talysurf unit. Two separate locations were provided corresponding to the 76.2mm spacing of the test disks. The Talysurf support stage is shown in Figure 9.11.

The stage was placed upon slides which allow the unit to be traversed in the direction parallel to the axis of the test disks. The motion of the stage was actuated by a manually operated lead screw. By traversing the support stage it was possible to take circumferential profiles of the two test disks at different axial positions.

The edges of the test disks provide the datum for relocation of the Talysurf stylus in the axial direction. Profiles were taken at different circumferential positions of the disks, where relocation was provided by etched lines in the sides of the disks. The marks on the disks were aligned using a locating bar as shown in Figure 9.12. The bar had two marks 76.2mm apart and it was placed in a reference position by resting it in contact with the disks with the end of the bar touching the slow speed yoke as shown in figure 9.12.

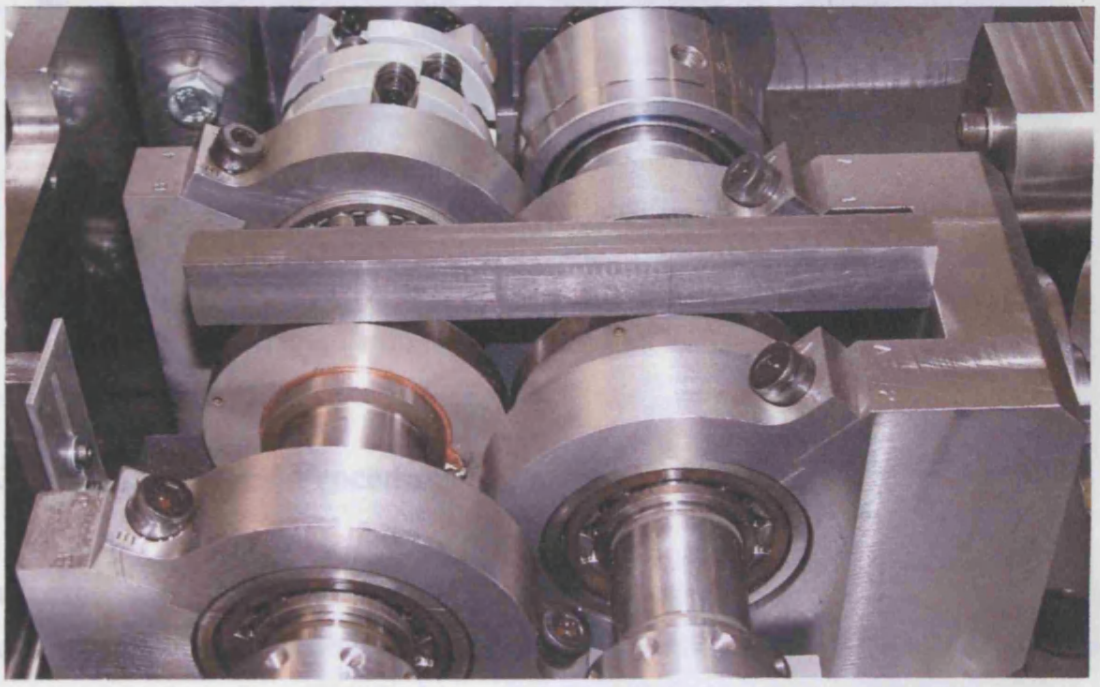

Figure 9.12 – Circumferential relocation assembly

This arrangement only provided a nominal relocation in the circumferential direction. Accurate relocation of profiles taken at different stages of the experiments was subsequently achieved by manual alignment of the unique valley features of the profiles.

To ensure that the measurements taken are accurate, the Talysurf was re-calibrated each time it was used. This involves traversing the stylus across a calibration ball of known radius of curvature. The calibration was carried out with the Talysurf positioned on the table. To enable the 'in-situ' calibration, a calibration jig was designed that spans the gap between the bearing supports of the slow shaft and the cut out made to allow the Talysurf access to the disks. This can be seen in Figure 9.13, where the calibration jig, ball and Talysurf stylus are all shown.

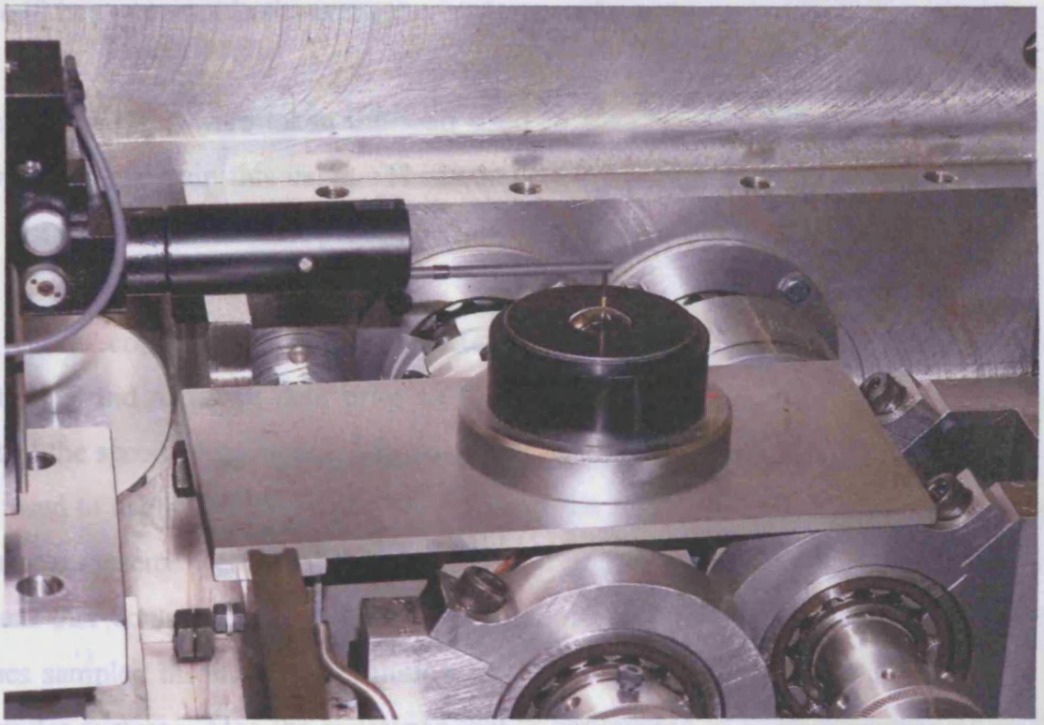

Figure 9.13 – Talysurf being calibrated using calibration jig and calibration ball

9.3 Experimental Procedure

Each experiment consisted of a number of tests on the same pair of disks. Three different loads were used and a number of tests were run at each load. The first test at each load stage was run for a short period of time and then further tests were run at the same load for increasing running times. In this way the effect of running time (at constant load) on surface modification was investigated.

Before the test head was assembled for each experiment, the circumferential reference points on the two disks were marked. These were made using a diamond tipped scribe, for the tests considered here four reference marks were made 90° apart. When new disks are fitted into the test head initial surface profiles are taken at the reference points. The axial distance between the profiles taken depends on the load that is applied during the running in test. Applying light loads to the disks results in a small contact dimension across the disks, thus the spacing between the traces was smaller than those used when a high load is applied. This ensured that the profiles taken were within the contact area, since if they were not in the contact area they would not be affected during running.

Once the profiles were taken the load used in the test was determined and the control valve set to maintain this value. This was done by placing the spacer shims between the two test shafts and using the procedure described earlier.

The rig is controlled by two computer systems, the original computer system that controlled and acquired data prior to the modifications detailed here was still used to control the speed of the disks. The computer system that was developed during this work was used to acquire the data generated during the test. Before a test was started the two computer systems were configured. The desired speed of the disks was entered into the system that controls the speed. The revolutions during which the electrical resistance values sampled through the revolution were to be acquired were entered into the 'new' computer system. Also the resistance that is considered to indicate contact, the internal resistance of the contact resistance circuit, shown in figure 9.9, and the voltage applied to the disks were entered into the 'new' computer system.

The oil systems were started well in advance of the start of the test to allow the oil to reach the desired operating temperature, through the use of a heater, as well as to provide lubrication to all components of the system prior to any movement. To begin the test the motor inverter was switched on, the computer software that controls the speed was started and the disks run to the pre-set speed. The next step was to close the ball valve and start the 'new' computer software running to acquire the data generated.

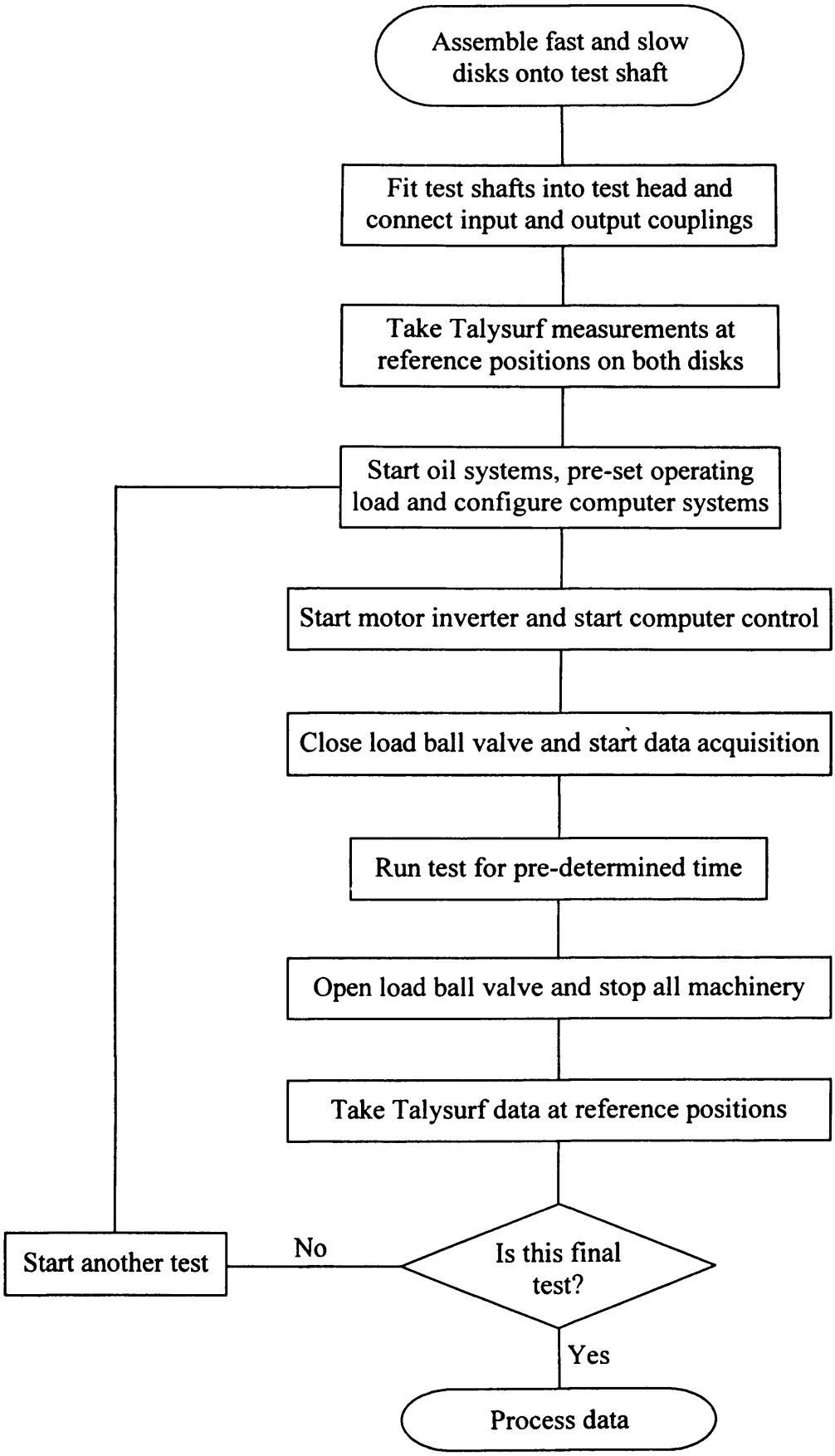

Figure 9.14 – Flow chart to show experimental 'running in' investigation procedure

The time for which a test was run changed from test to test, for the first test of each load stage the running time was very short, in the region of 15 seconds. To stop the system, the ball valve was opened to remove the load and the motor inverter switched off to stop the rotation. After this, Talysurf profiles were taken at the reference positions to give an indication of how the surfaces have changed during the test. This comparison was made on both disks.

The other tests undertaken were approached in the same manner. So that the Talysurf profiles could be taken the pumps and other machinery were stopped to eliminate the vibration, which would give false surface features. Since the tests had to be re-started, the oil system had to run for a period of time and the load re-set before the test continued. This process is illustrated using the flow chart shown in Figure 9.14.

9.4 Experimental Results

The first series of tests were conducted using a superfinished disk on the slow shaft and an axially ground disk on the fast shaft. There were initial problems with the electrical circuitry and the computer data acquisition, but the first test allowed these approaches to be refined.

After running a rough disk against a smooth disk, the arrangement was modified to consider a rough disk running against another rough disk. The two configurations were compared in terms of the manner in which the disks modify over the tests as well as the temperature, friction and contact resistance response.

The loads stages used in the experiments are given in Table 9.1.

Table 9.1 – Load stages

Stage	p_{max} / GPa^{-1}	w' / N	a / mm	b / mm
1	0.7	288	0.876	0.224
2	1.5	2834	1.877	0.481
3	2.1	7776	2.503	0.673

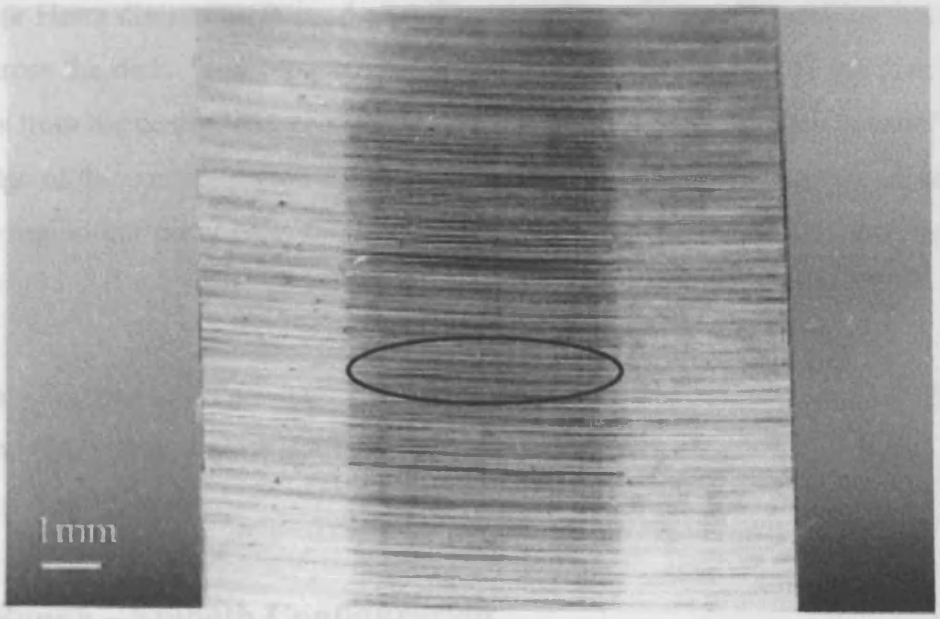

Figure 9.15 – Contact area superimposed on top of run in test disk

The dimensions of the disks given in section 9.1 can be used in the Hertzian equations to determine the contact ellipse. The relative radius of curvature in the entrainment direction is 19.05mm and the radius of curvature perpendicular to the entrainment direction is 152.4mm. This generates a contact ellipse where the major Hertz semi-dimension, a, gives the width of the contact across the disk and the minor Hertz semi-dimension, b, gives the length of the contact. The ratio of the major dimension to the minor dimension is 4:1, giving rise to the ellipse shown in Figure 9.15.

Table 9.2 – Axial locations of the Talysurf profiles

Profile	Distance / mm
1	-2.4
2	-1.8
3	-0.8
4	-0.2
5	0
6	0.2
7	0.8
8	1.8
9	2.4

The major Hertz dimension is used to determine the spacing of the profiles that are to be taken across the disk. The profiles are spaced from the centre line of the disk and their distances from the centre line, profile 5, are given in Table 9.2. Profiles 1-3 and 7-9 occur at the edge of the contact region for the three load stages. Profiles 4 and 6 are in the high pressure region but away from the centre of the contact and profile 5 is the centre of the contact.

In all tests the speed was constant. The motor speed was set to give a fast shaft speed of 3000rpm, which gives a slow shaft speed of 2023rpm.

9.4.1 Rough – Smooth Configuration

The details of the rough – smooth tests are outlined in Table 9.3.

Table 9.3 – Test details for the rough – smooth configuration

Test	Fast Shaft Finish	Slow Shaft Finish	Load Stage	Position in Sequence	Duration /minutes: seconds
1-S	Rough	Smooth	1	1	0:15
2-S	Rough	Smooth	1	2	0:30
3-S	Rough	Smooth	1	3	5:00
4-S	Rough	Smooth	1	4	60:00
5-S	Rough	Smooth	2	1	0:15
6-S	Rough	Smooth	2	2	1:00
7-S	Rough	Smooth	2	3	4:00
8-S	Rough	Smooth	2	4	15:00
9-S	Rough	Smooth	2	5	60:00
10-S	Rough	Smooth	3	1	0:15
11-S	Rough	Smooth	3	2	1:00

There were initial problems with the computer system for the data acquisition, the result of this was that most of the data for the first test was lost. Other problems also developed during the tests, the thermocouple for the slow disk stopped working after the first test and could not be replaced without stripping the entire rig and embedding another thermocouple. It was decided to continue the series of tests with one thermocouple as the time taken to replace the thermocouple would create considerable delays.

The computer data acquisition program takes time to configure the channels and set up arrays that are used to output data. During this time period, the data that are generated cannot be measured which resulted in some of the data being lost from the first test. The program was re-written so that the configuration could be completed and the system armed ready to acquire signals. The acquisition of the signals was started using a manual trigger, which removes the chances of losing data.

Whilst the computer program was being re-written, the contact resistance voltage recorded during the revolution was analysed. Unlike the results of Furey (1961), it was found that the voltage measured did not go from very high to very low values. There was a gentler transition and the values were at an intermediate level. This was studied in more detail and was found to be an averaging effect. The source of the averaging was initially thought to be related to the data acquisition hardware, but was found to be a result of the low pass filter. Upon further investigation it was found that the time constant of the low pass filter was 0.1ms, which is too large and influencing the data being acquired. The effect of the filter is to average the contact resistance over a number of points.

Despite this averaging, there were still repeatable trends seen in the data acquired, but it could not be used to determine the percentage of contact that occurs through the revolution. However, the value that is averaged over the revolution will still give an indication of what is happening to the amount of contact over the duration of the test. This value can be used as a qualitative indicator of the amount of contact that is occurring, the lower the value measured over the revolution the more the contact that is occurring. Instead of re-developing all of the circuitry related to the contact resistance, the average of the values taken over the revolution was taken as qualitative indicator of the amount of contact occurring.

The results for the second test, which was run for 30 seconds, can be seen in Figure 9.16. Figure 9.16 and all subsequent figures show the variation of the fast disk bulk temperature, *TF*, friction, *FR*, contact resistance voltage, *CR*, and the load, *W*. It is seen that the contact resistance voltage measured across the disks falls when the load is applied, indicating that there is contact occurring between the two disks. The voltage does not drop a significant amount, which means that the amount of contact developed is reasonably small. The friction and temperature are almost constant over the test.

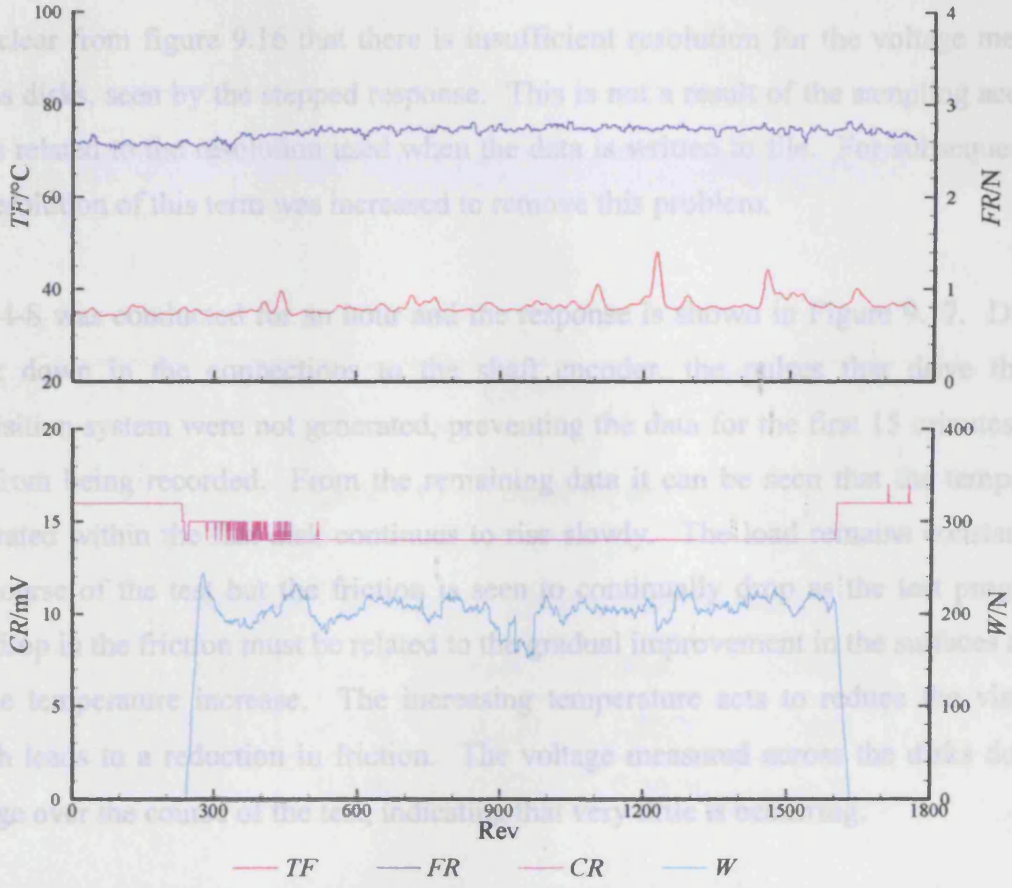

Figure 9.16 – Results for test 2-S

The periods of running for each test were not defined initially, the test periods were increased over the first load stage by an arbitrary amount based on the results from the previous test. The test lengths used were 15 seconds, 30 seconds, 5 minutes and finally an hour for test1-S, test 2-S, test 3-S and test 4-S respectively. In addition, the time necessary to take the Talysurf profiles was determined, it was found that taking the readings at four circumferential positions was too time consuming. The time taken to

take all of the axial readings at the four circumferential positions was 4 hours, if only 5 readings were taken then the time reduced to 3 hours. To reduce the time further, the circumferential positions were reduced to 2 and only 3 axial profiles were taken at each circumferential point after each test. At the end of each load stage the profiles that are next to the outer profiles are also taken. For instance, considering the first load stage the outer most profiles are 3 and 7. At the end of the load stage, profiles 4 and 6 were taken to give an indication of the change in the profiles near the edge and profiles 2 and 8 were taken to ensure that they had not been influenced during the load stage.

It is clear from figure 9.16 that there is insufficient resolution for the voltage measured across disks, seen by the stepped response. This is not a result of the sampling accuracy, but is related to the resolution used when the data is written to file. For subsequent tests the resolution of this term was increased to remove this problem.

Test 4-S was conducted for an hour and the response is shown in Figure 9.17. Due to a break down in the connections to the shaft encoder, the pulses that drive the data acquisition system were not generated, preventing the data for the first 15 minutes of the test from being recorded. From the remaining data it can be seen that the temperature generated within the fast disk continues to rise slowly. The load remains constant over the course of the test but the friction is seen to continually drop as the test progresses. The drop in the friction must be related to the gradual improvement in the surfaces as well as the temperature increase. The increasing temperature acts to reduce the viscosity, which leads to a reduction in friction. The voltage measured across the disks does not change over the course of the test, indicating that very little is occurring.

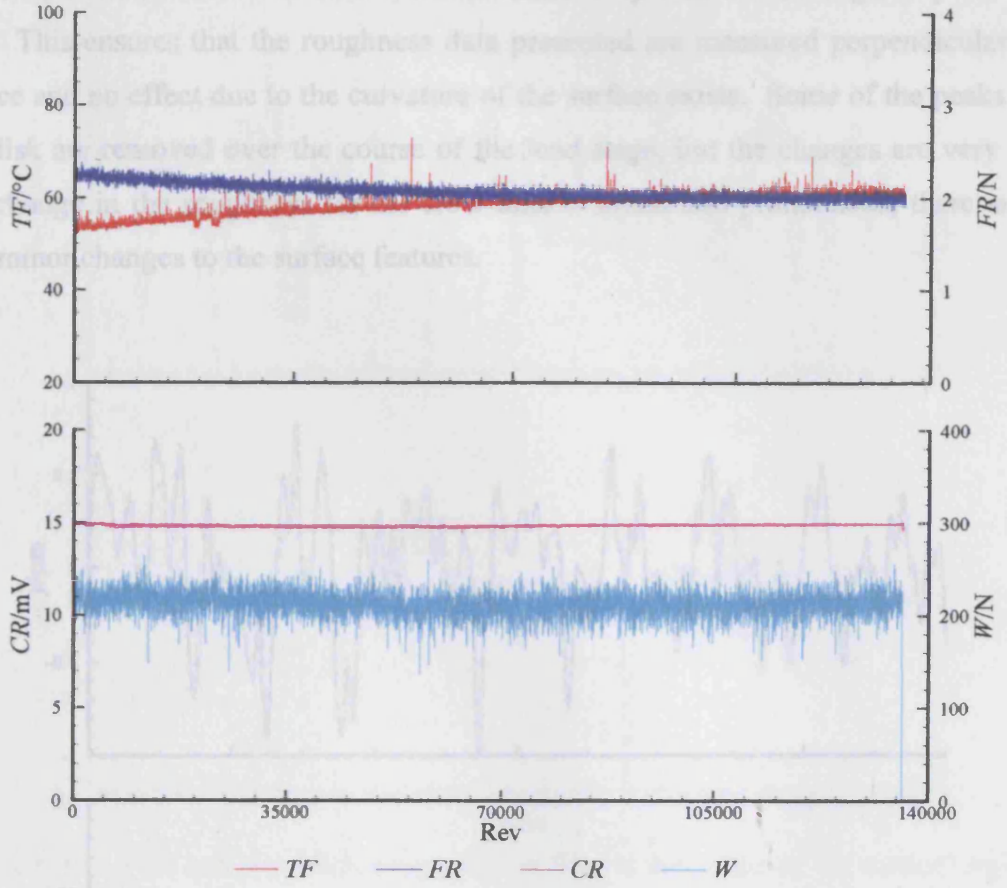

Figure 9.17 – Results for test 4-S

Table 9.4 – Ra values in μm from test 0-S and 4-S

Test	Fast 3	Fast 5	Slow 3	Slow 5
0-S	0.330	0.330	0.0596	0.0568
4-S	0.316	0.325	0.0619	0.0573

The surfaces did not modify significantly during the first load stage, this can be seen when considering the change in the Ra value of the profiles shown in Table 9.4. Test 0-S refers to the profile taken before the disks were run against one another. It can be seen that there are only slight changes in the Ra value for the centre line profiles. There is a greater change in the Ra value at the edge of the contact on both surfaces. The change in Ra values are small, this is reflected in the surface profiles.

The roughness profile at the edge of the contact for the fast and slow disks is shown in Figure 9.18. The profiles for both disks are presented for a region close to the centre of

the trace, therefore the data illustrated are taken very close to the highest point on the disk. This ensures that the roughness data presented are measured perpendicular to the surface and no effect due to the curvature of the surface exists. Some of the peaks on the fast disk are removed over the course of the load stage, but the changes are very minor. The change in the roughness for the slow disk is much less pronounced, there are only very minor changes to the surface features.

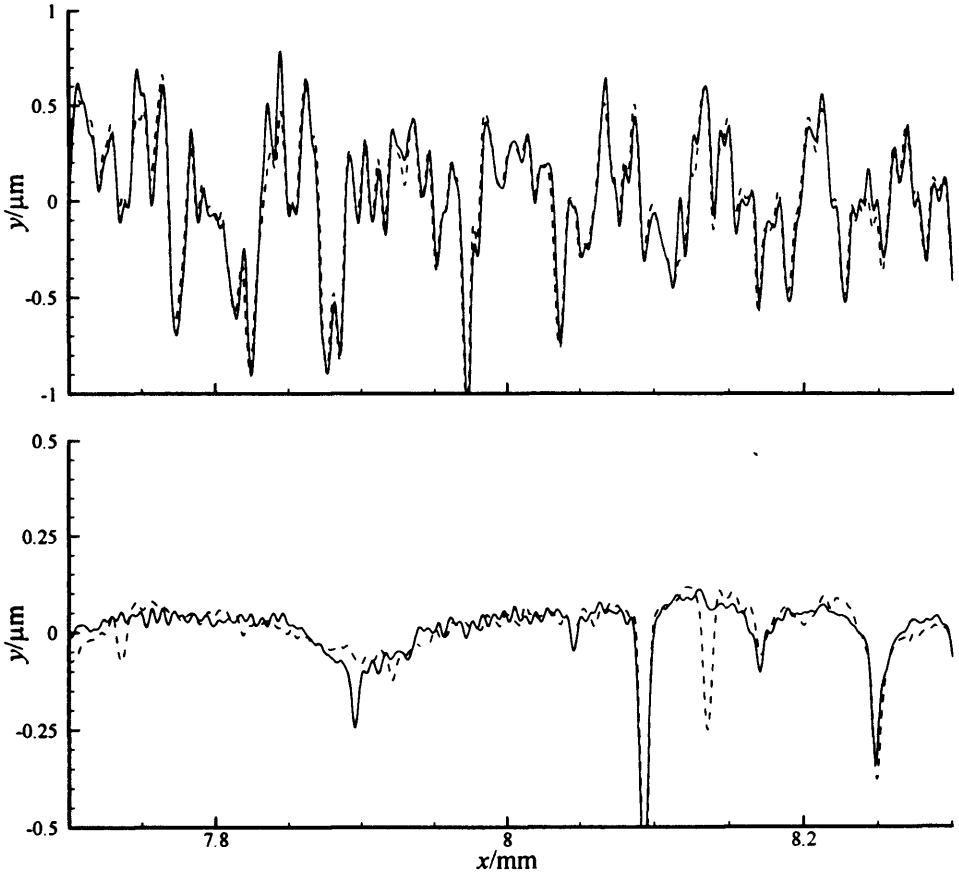

Figure 9.18 – Fast and slow disk roughness profiles at the edge of the contact region for test 0-S (solid) and test 4-S (dashed)

The centre line profiles for the two disks are shown in Figure 9.19. The change in the roughness for the centre line response is similar to the roughness taken from the edge of the contact. A number of the peak features of the fast disk remain unaffected during the test, only the very large peaks have been modified during the test.

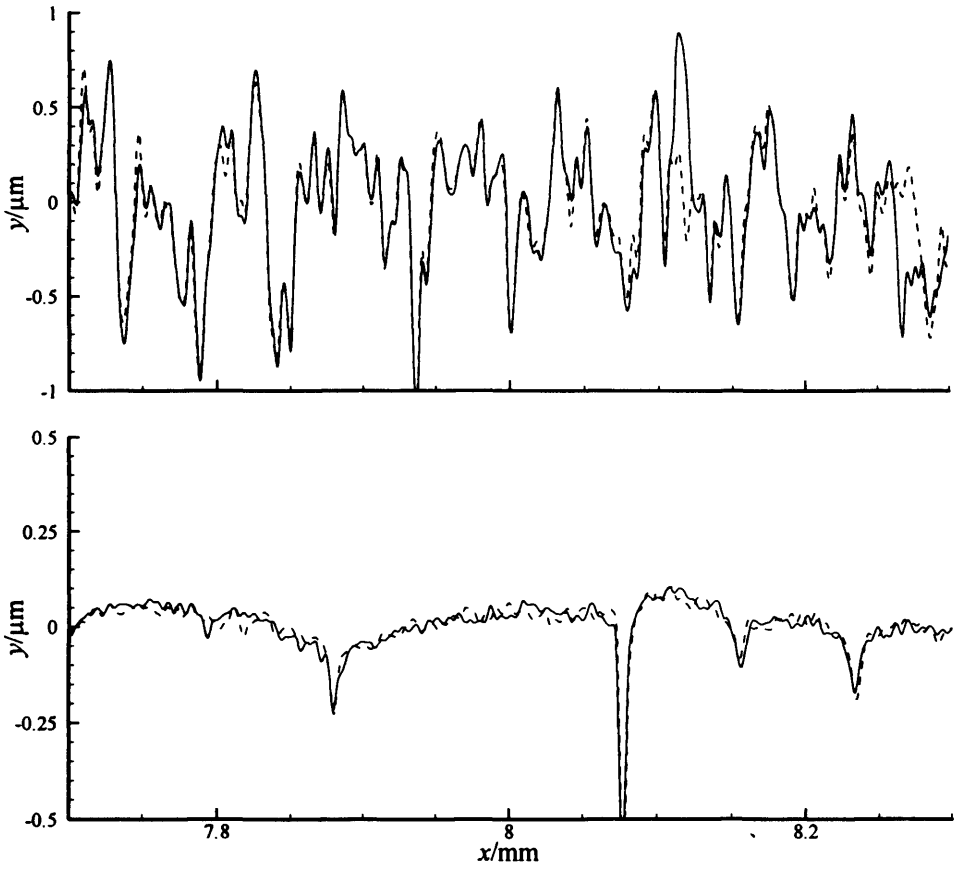

Figure 9.19 – Fast and slow disk roughness profiles at the centre of the contact region for test 0-S (solid) and test 4-S (dashed)

After the first load stage had been completed, it was decided to use the results obtained to choose a running time for all of the subsequent load stages. The approach adopted was to run for periods of 15 seconds, 1 minute, 4 minutes, 15 minutes and 60 minutes.

The response for test 5-S is shown in Figure 9.20. The increased load causes a corresponding increase in friction. It can also be seen that a higher temperature was developed during the test and the voltage measured across the disks dropped significantly. The thinner films that are formed allow more interaction between the surfaces. Due to the increase in the temperature, the film thickness reduces during the test, which generates more interaction of the surfaces. This is reflected in the voltage measured across the disks, where it continues to reduce as the test progresses.

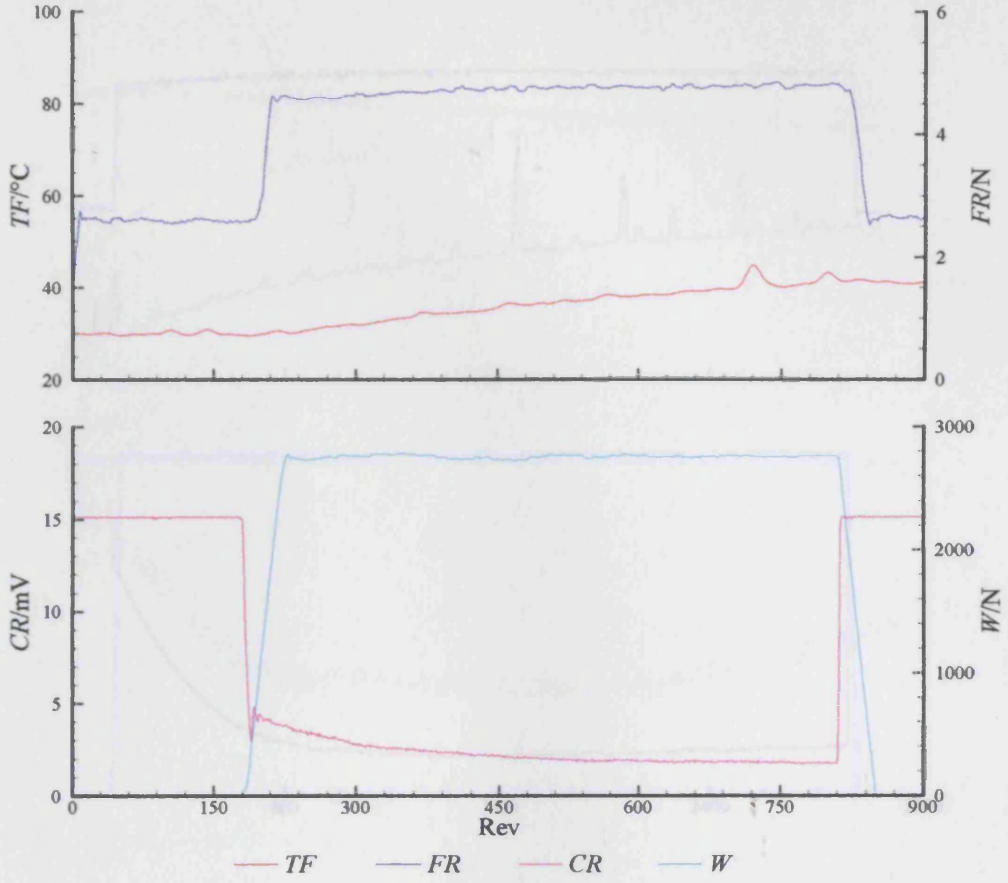

Figure 9.20 – Results for Test 5-S

The same features are seen when considering the results from test 6-S, shown in Figure 9.21. The friction force shows the same features as that seen in figure 9.20. The temperature continues to rise over the duration of the test and reaches a higher value than for test 5-S, where a temperature of over 50°C is reached towards the end of the test. The contact resistance drops as the temperature rises, reaching its lowest value at 1400 revolutions.

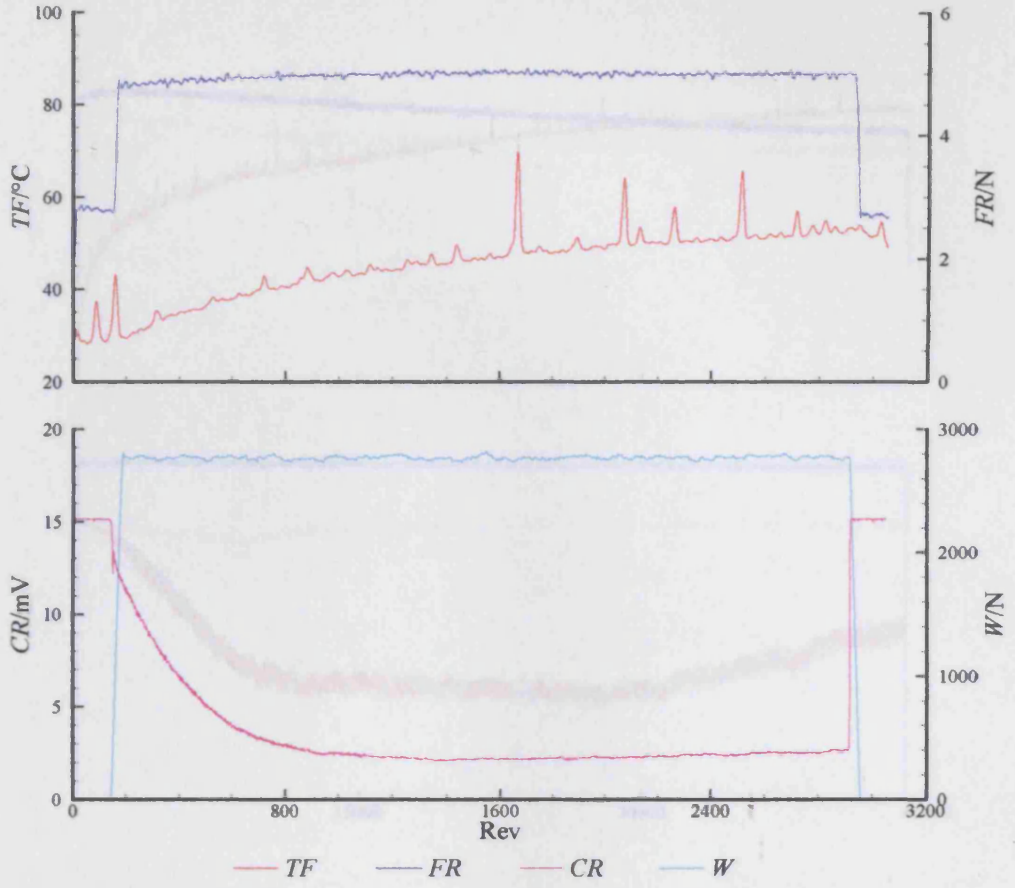

Figure 9.21 – Results for Test 6-S

Test 7-S generates similar features, the only difference is seen in the friction, where the friction reduces slightly as the test progresses. The reduction in the friction is similar to the response of the friction seen for test 4-S. The contact resistance voltage develops the same general trend as test 6-S, but does not reach the same minimum value, the values are roughly 1mV higher in test 7-S than in test 6-S. The temperature reached during test 7-S was nearly 65°C, which is about 15°C higher than test 6-S.

Similar results to tests 6-S and 7-S are also seen for test 8-S, which is shown in Figure 9.22. The friction develops the same maximum value as test 6-S at 5N, but is then seen to drop over the duration of the test as the temperature rises. The friction finally reaches a value of 4N when the temperature approaches 80°C. The contact resistance drops off as the test begins, it reaches a constant minimum value of 6mV between revs 15,000 and 30,000. After this point the voltage is seen to increase, indicating the amount of metallic contact reduces.

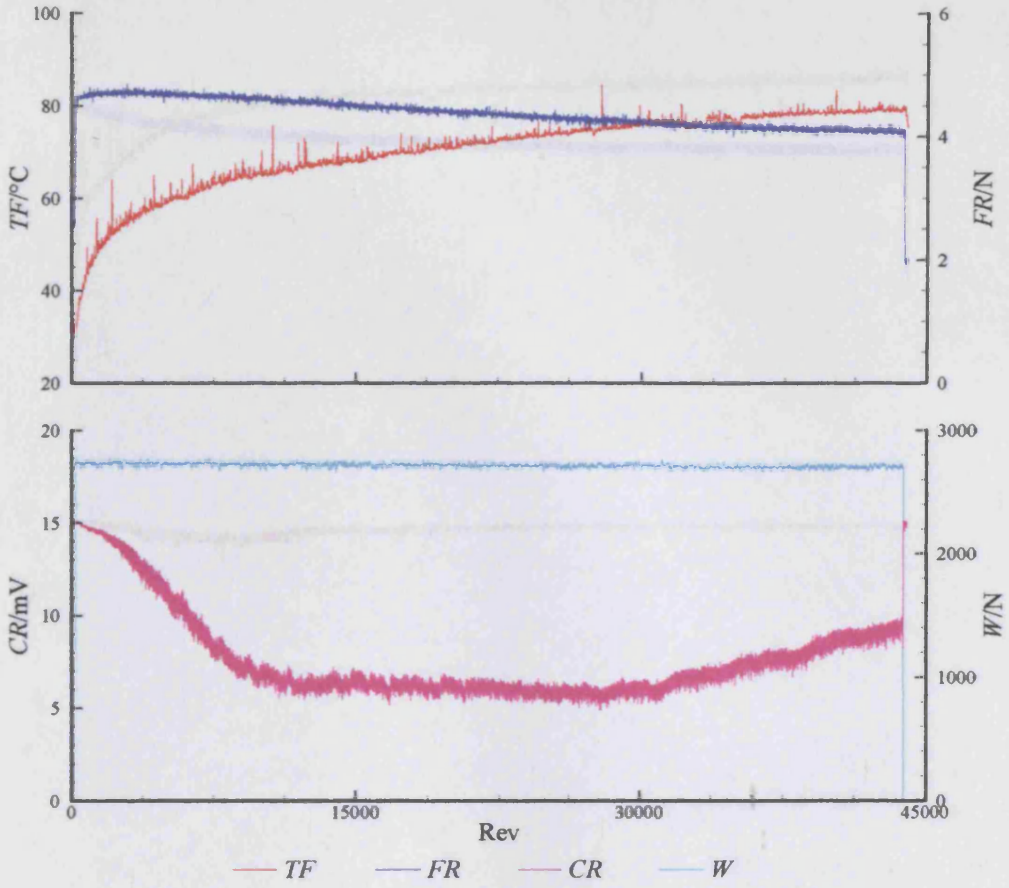

Figure 9.22 – Results for Test 8-S

The results for test 9-S show the same features as test 4-S, the results for test 9-S are shown in Figure 9.23. The friction force and the temperature appear to be related, as the temperature levels off the decline in the friction is also seen to reduce. The temperature levels off after a value of 80°C is reached, this corresponds with a minimum friction force of just under 4N. The contact resistance drops slightly at the beginning of the test, reaching a minimum value of 13mV at revolution 30,000. The contact resistance continues to rise from this point until revolution 60,000, after this point the contact resistance remains reasonably constant. There is a slight upwards slope in the voltage measured across the contact from revolution 60,000 to the end of the test.

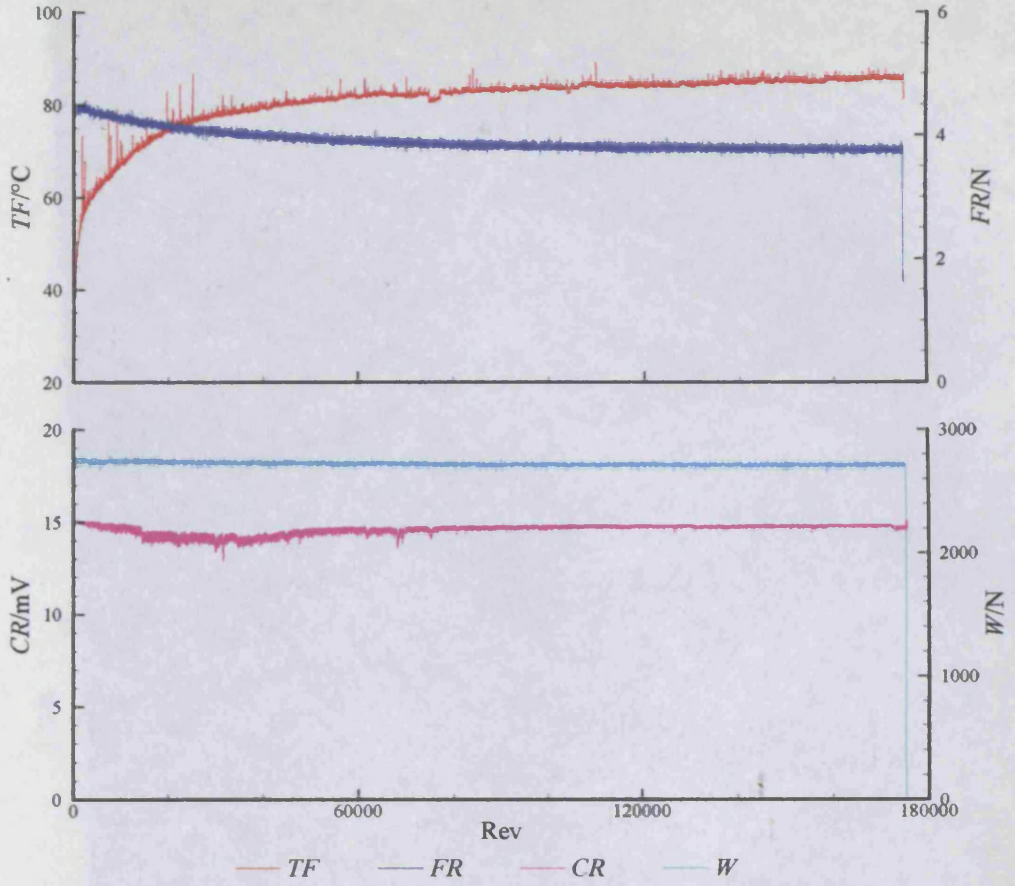

Figure 9.23 – results for test 9-S

The roughness change over the load stage was more noticeable than at the lower load stage, this is clearly seen in Table 6.5. The smooth surface remains reasonably unaffected during the test, but there was a slight roughening of both traces. The rough disk shows a reduction in Ra over test at both profile locations, where the edge of the contact area developed a smoother profile than the centre line. This indicates that most of the contact occurring during the tests occurs at this location. Tests 5-S to 8-S develop large amounts of contact for most of the test duration, which is reflected in the change in the Ra of the fast disk. The traces from the edge of the contact area produce greater changes in Ra for the longer duration test, indicating that more material is removed or displaced. For test 9-S, where there were only slight changes in the voltage measured across the disk, the Ra values do not change significantly in any of the traces considered.

Table 9.5 – Ra values in μm from test 4-S to 9-S

Test	Fast 2	Fast 5	Slow 2	Slow 5
4-S	0.326	0.325	0.0621	0.0573
5-S	0.317	0.321	0.0625	0.0575
6-S	0.316	0.322	0.0625	0.0602
7-S	0.305	0.312	0.0656	0.0588
8-S	0.277	0.312	0.0669	0.0591
9-S	0.272	0.309	0.0679	0.0609

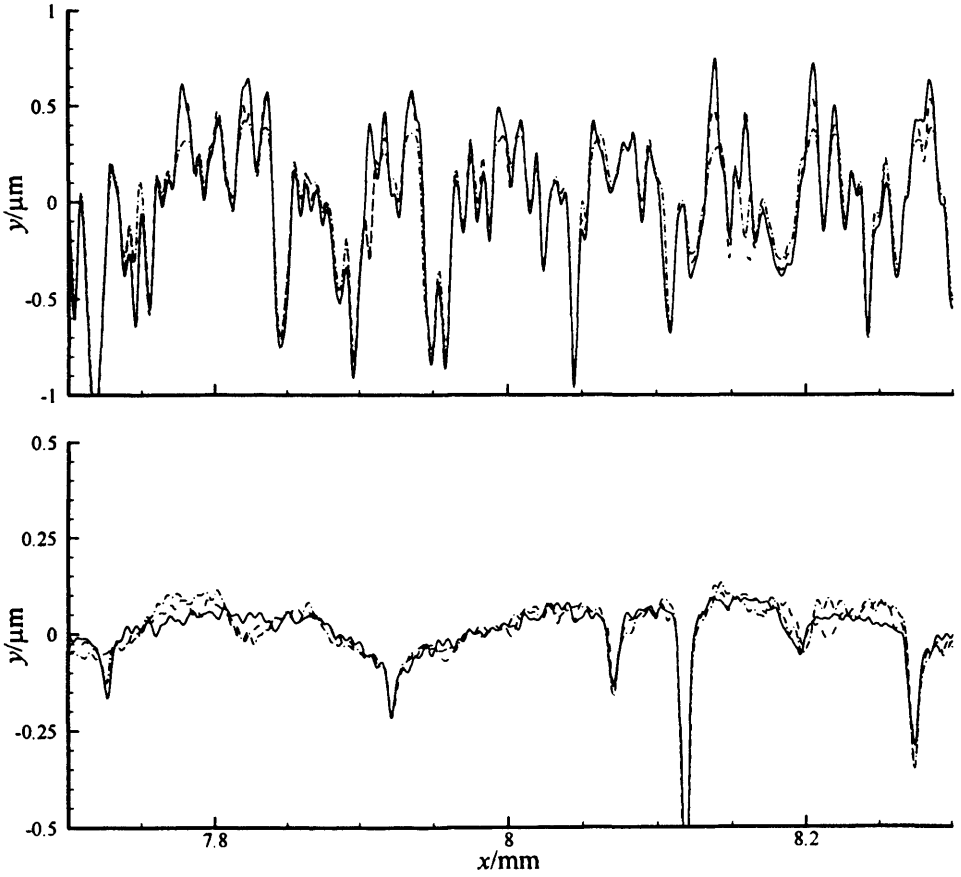

Figure 9.24 – Fast and slow disk roughness profiles at the edge of the contact region for test 4-S (solid), test 7-S (dashed) and test 9-S (dash-dotted)

Figure 9.24 shows the change in the surface features at the edge of the contact on the fast and slows disks over the load stage. The slow disk does experience proportionally large changes, but the peaks on the fast disk are continually rounded as the tests progress. There are large peaks present before the surface has run, but after running the surfaces for

a small period of time, just over 5 minutes, the large peaks have been reduced in amplitude and are more rounded.

The results for test 10-S, which is the first test at the highest load stage, are shown in Figure 9.25.

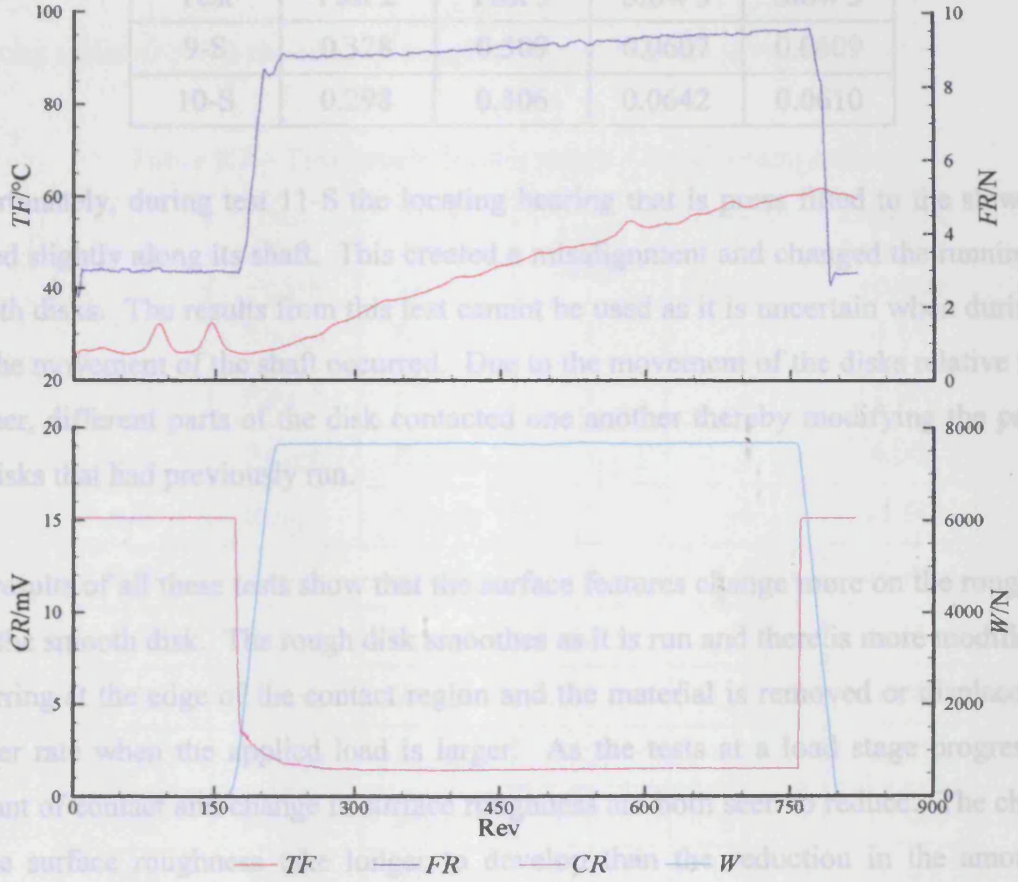

Figure 9.25 – Results for Test 10-S

The features seen in the friction, temperature and contact resistance are very similar to those seen for test 5-S. The temperature increases rapidly, rising to 60°C before the end of the 15 second test. The friction force is much higher at this load, increasing throughout the test and exceeding 9.5N. The voltage measured across the disks drops rapidly to a value of 1.5mV and remains constant through most of the test duration.

The change in the roughness over test 10-S is shown in Table 9.6. The change in roughness at the edge of the contact over test 10-S was much greater at this load than at

any of the other load stages. An addition to the large amount of smoothing occurring on the rough surface there was a noticeable roughening of the smooth disk, most notably at the edge of the contact. The change in roughness along the centre line remains low, but there is still the tendency to roughen the smooth disk and smooth the rough disk.

Table 9.6 – Ra values in μm for test 9-S and 10-S

Test	Fast 2	Fast 5	Slow 3	Slow 5
9-S	0.328	0.309	0.0607	0.0609
10-S	0.298	0.306	0.0642	0.0610

Unfortunately, during test 11-S the locating bearing that is press fitted to the slow shaft moved slightly along its shaft. This created a misalignment and changed the running part of both disks. The results from this test cannot be used as it is uncertain when during the test the movement of the shaft occurred. Due to the movement of the disks relative to one another, different parts of the disk contacted one another thereby modifying the parts of the disks that had previously run.

The results of all these tests show that the surface features change more on the rough disk than the smooth disk. The rough disk smoothes as it is run and there is more modification occurring at the edge of the contact region and the material is removed or displaced at a greater rate when the applied load is larger. As the tests at a load stage progress, the amount of contact and change in surface roughness are both seen to reduce. The changes in the surface roughness take longer to develop than the reduction in the amount of contact.

The voltage measured across the disks at the end of one test is not the same as the minimum value measured in the next. Since the temperatures generated at the end the following test are higher, then the feature cannot be related to temperature. There may be a higher surface temperature generated when there is more contact that could locally thin the film, but these effects are thought to be small. This effect could be the result of further modification of the surfaces occurring at lower temperatures. Due to the changes at lower temperatures in the next test the surfaces are less rough, when they reach the higher temperatures the surfaces are smoother and produce less contact.

Higher loads are seen to develop greater friction and higher temperatures. The higher loads and temperatures generate thinner films that allow more interaction of the surfaces. It is also seen that the amount of surface modification occurring increases with load.

9.4.2 Rough – Rough Configuration

The tests undertaken for the rough – rough configuration are given in Table 9.7.

Table 9.7 – Test details for the rough – rough configuration

Test	Fast Shaft Finish	Slow Shaft Finish	Load Stage	Position in Sequence	Duration /minutes: seconds
1-R	Rough	Rough	1	1	0:15
2-R	Rough	Rough	1	2	1:00
3-R	Rough	Rough	1	3	4:00
4-R	Rough	Rough	1	4	15:00
5-R	Rough	Rough	1	5	60:00
6-R	Rough	Rough	2	1	0:15
7-R	Rough	Rough	2	2	1:00
8-R	Rough	Rough	2	3	4:00
9-R	Rough	Rough	2	4	15:00
10-R	Rough	Rough	2	5	60:00
11-R	Rough	Rough	3	1	0:15
12-R	Rough	Rough	3	2	1:00
13-R	Rough	Rough	3	3	4:00

The amount of surface modification seen in the rough – smooth configuration was small. It is only possible to comment of the changes in the fast disk as it is rough and the slow disk does not undergo much change in roughness. In an effort to generate more modification of the surfaces and produce more interaction two rough disks were used.

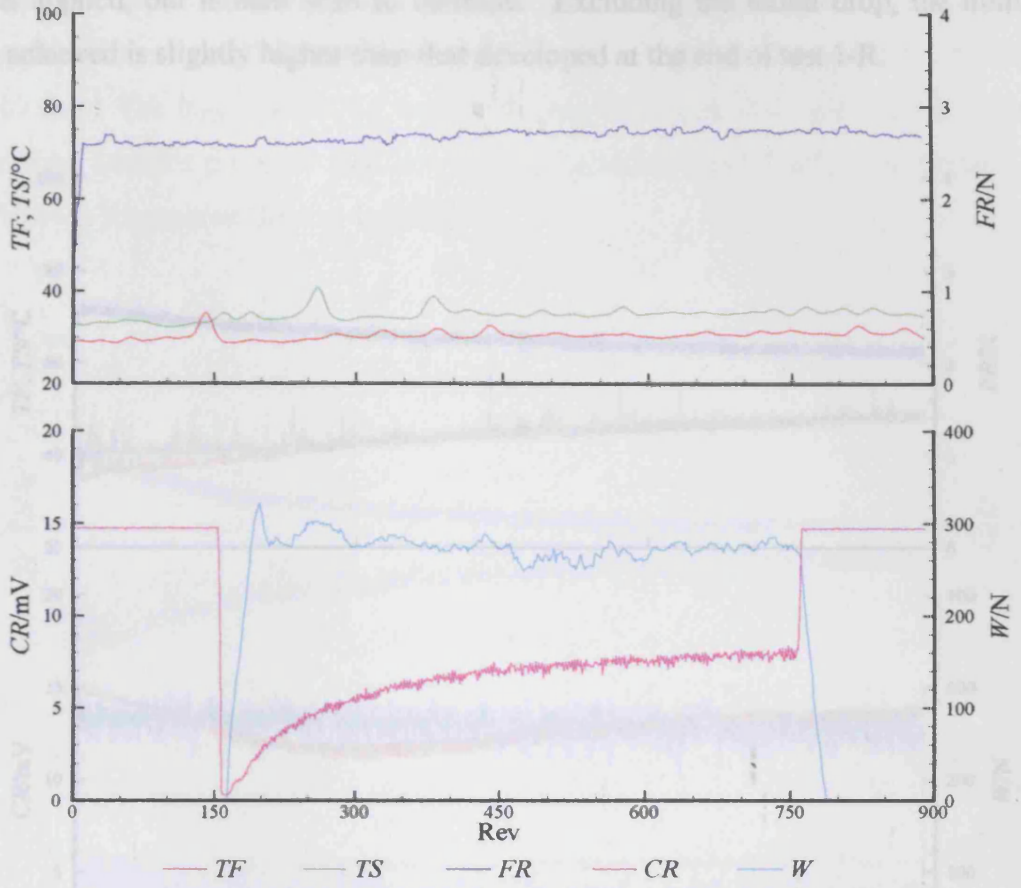

Figure 9.26 – Results for rough on rough test 1-R

The results for the first test are shown in Figure 9.26. In addition to the data shown in previous figures, figure 9.26 and subsequent figures show the variation in the slow disk bulk temperature, *TS*. At the low load it can be seen that the slow disk temperature is nearly 5°C higher than the fast disk throughout the test. During the test the temperature in both disks increases slightly and the friction is also seen to increase a small amount. The contact resistance drops quickly when the load is applied, but is then seen to increase quite quickly. This indicates that the surfaces modify quickly reducing the amount of contact developed.

The results for test 2-R are very similar to test 1-R. The temperature in the disks increases a little more over the test, but the slow disk temperature is still a few degrees hotter than the fast disk. The friction force rises to revolution 600, but is then seen to drop over the remaining revolutions. The contact resistance drops very quickly when the

load is applied, but is then seen to increase. Excluding the initial drop, the minimum value achieved is slightly higher than that developed at the end of test 1-R.

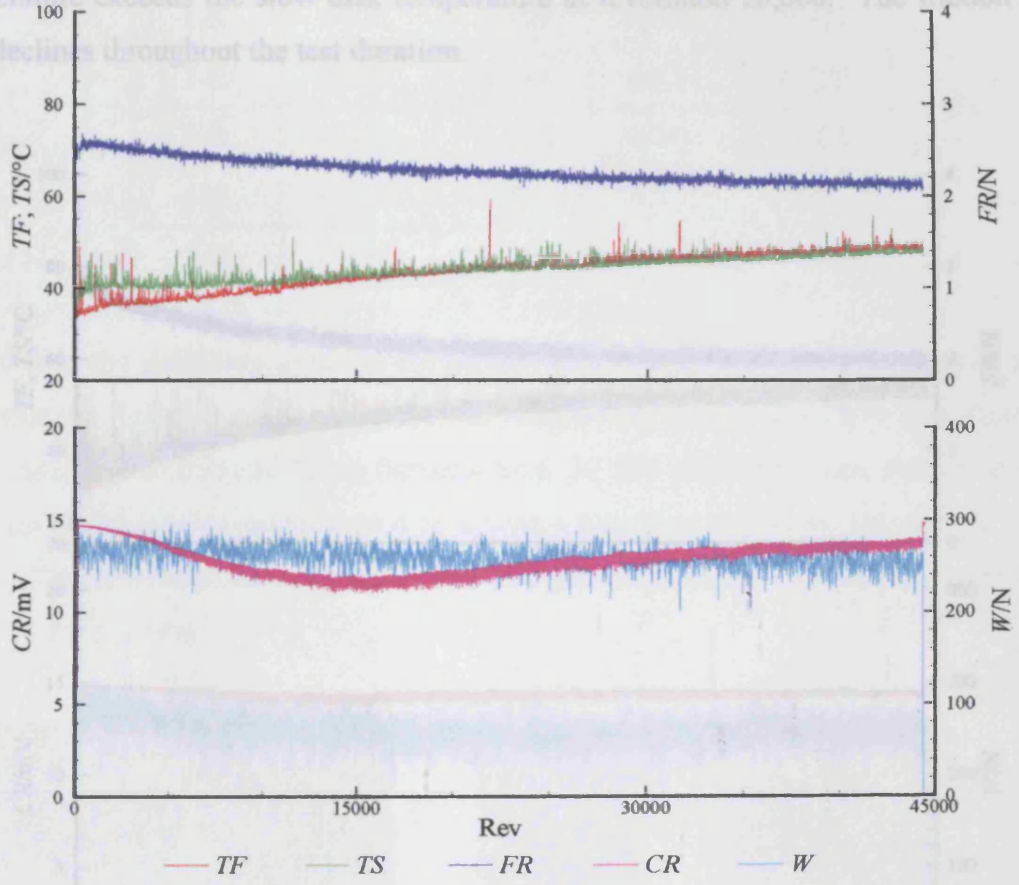

Figure 9.27 – Results for rough on rough test 4-R

Slightly different features are seen during test 3-R, where the contact resistance no longer rises and begins to fall over the test. The temperatures in the disks converge, thus at the end of the test the temperatures are almost the same. The temperatures do reach higher values than that experienced in test 1-R and test 2-R. The friction continues to drop during the test, as expected due to the increasing temperature. The contact resistance develops a different response to that seen in test 1-R and test 2-R, instead of increasing the contact resistance is seen to fall. This may be related to the increase in temperature in reducing the film thickness, allowing the asperities on the surfaces to come into closer contact.

Test 4-R, shown in Figure 9.27, also shows a reduction in the contact resistance over the first period of the test. The contact resistance reduces until revolution 18,000 but is then seen to rise. The temperatures of both disks increase, it is also seen that the fast disk temperature exceeds the slow disk temperature at revolution 20,000. The friction force also declines throughout the test duration.

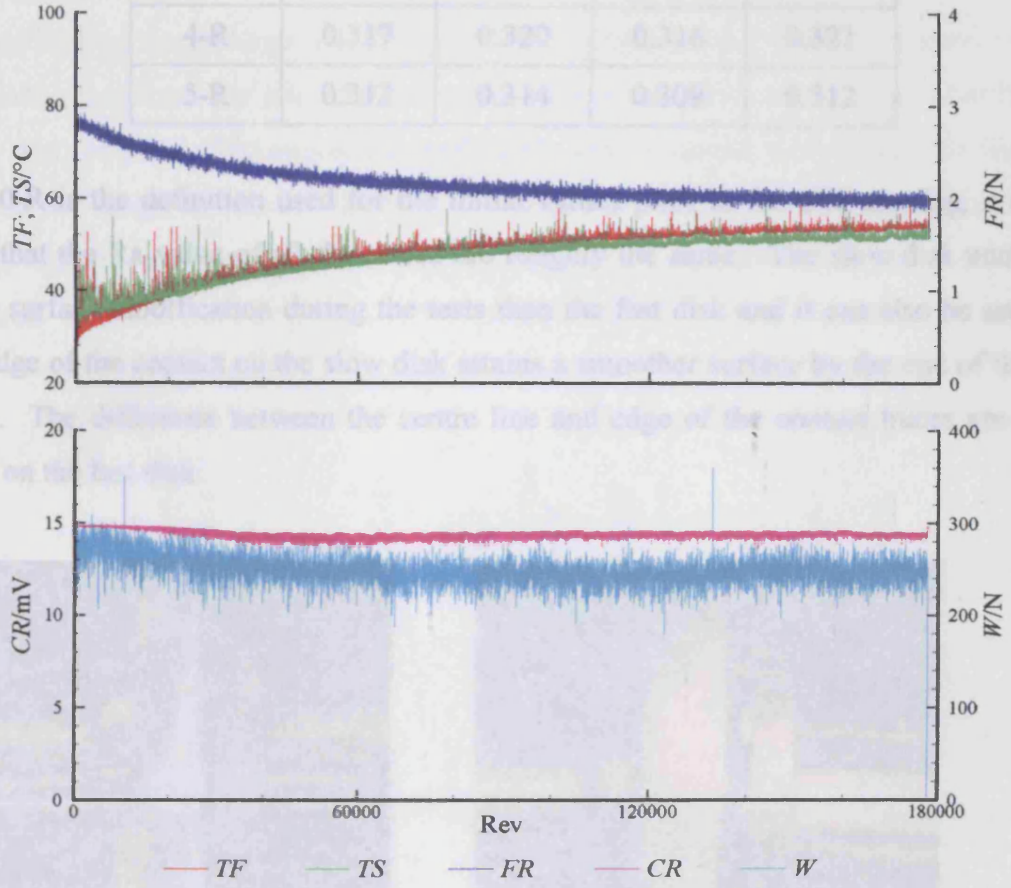

Figure 9.28 – Results for rough on rough test 5-R

The results for test 5-R are shown in Figure 9.28. The temperatures cross over and continue to rise, reaching 55°C by the end of the test. The friction drops off, showing the same features to that seen for test 9-S that is shown in figure 9.23. The contact resistance generates a similar response to test 4-R, except that the drop in the resistance is much less noticeable.

The change in the roughness over the first load stage can be seen in Table 9.8.

Table 9.8 – Ra values in μm for first load stage

Test	Fast 3	Fast 5	Slow 3	Slow 5
0-R	0.335	0.338	0.347	0.349
1-R	0.325	0.329	0.333	0.331
2-R	0.329	0.330	0.331	0.329
3-R	0.322	0.325	0.326	0.330
4-R	0.317	0.320	0.316	0.321
5-R	0.312	0.314	0.309	0.312

Test 0-R is the definition used for the initial values prior to the disk running, it can be seen that the Ra value of all the traces are roughly the same. The slow disk undergoes more surface modification during the tests than the fast disk and it can also be seen that the edge of the contact on the slow disk attains a smoother surface by the end of the load stage. The difference between the centre line and edge of the contact traces are not as large on the fast disk.

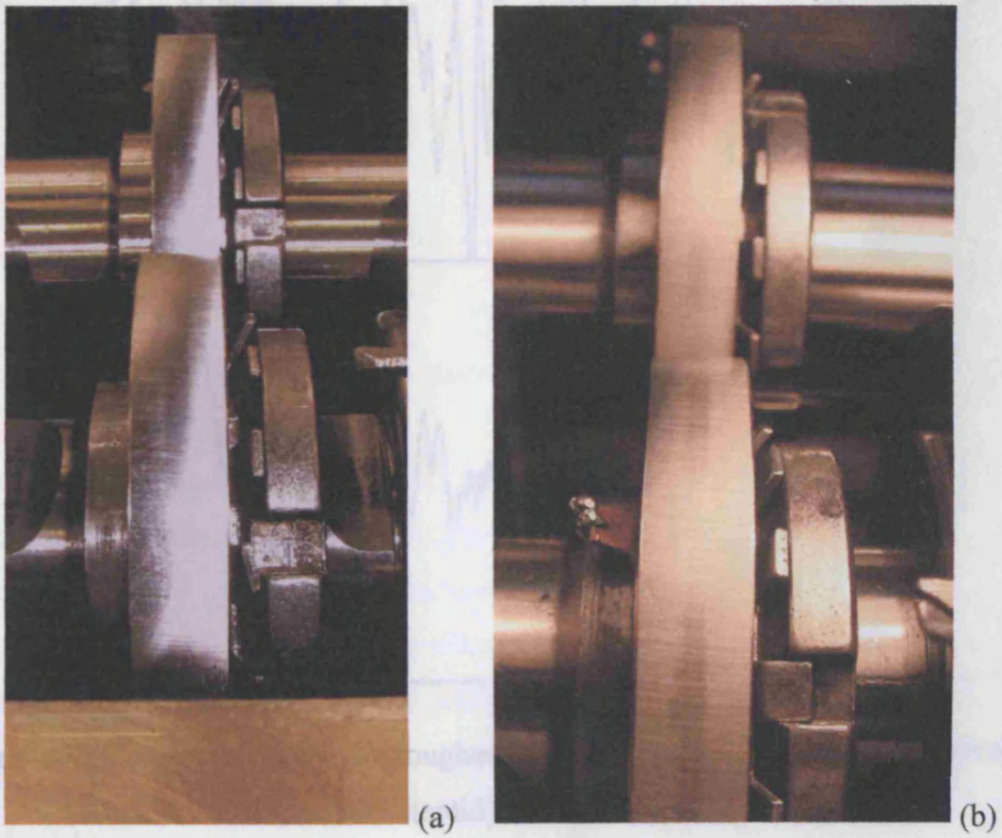

Figure 9.29 – Photographs of disks; (a) after test 1-R, (b) after test 5-R

The change in the surfaces over the first test can be clearly seen in Figure 9.29 (a). The running track is formed very quickly, as seen by a bright area developed near the centre of the disks. Figure 9.29 (b) shows the surfaces after the end of the test 5-R. As the load stage develops, it can be seen that the surfaces smooth more, but the smoothing at the edge of the slow disk becomes far more apparent.

Despite the changes to the appearance of the surface, the actual features that make up the surface do experience large modification over the load stage. Consider Figure 9.30 that shows the roughness variation for the centre line traces before running and after the load stage was complete. The aggressive peaks have been rounded, as they were on the rough disk in the rough on smooth tests. The changes are slightly more noticeable than the results shown in figure 9.19, but still do not show very large changes in the surface features.

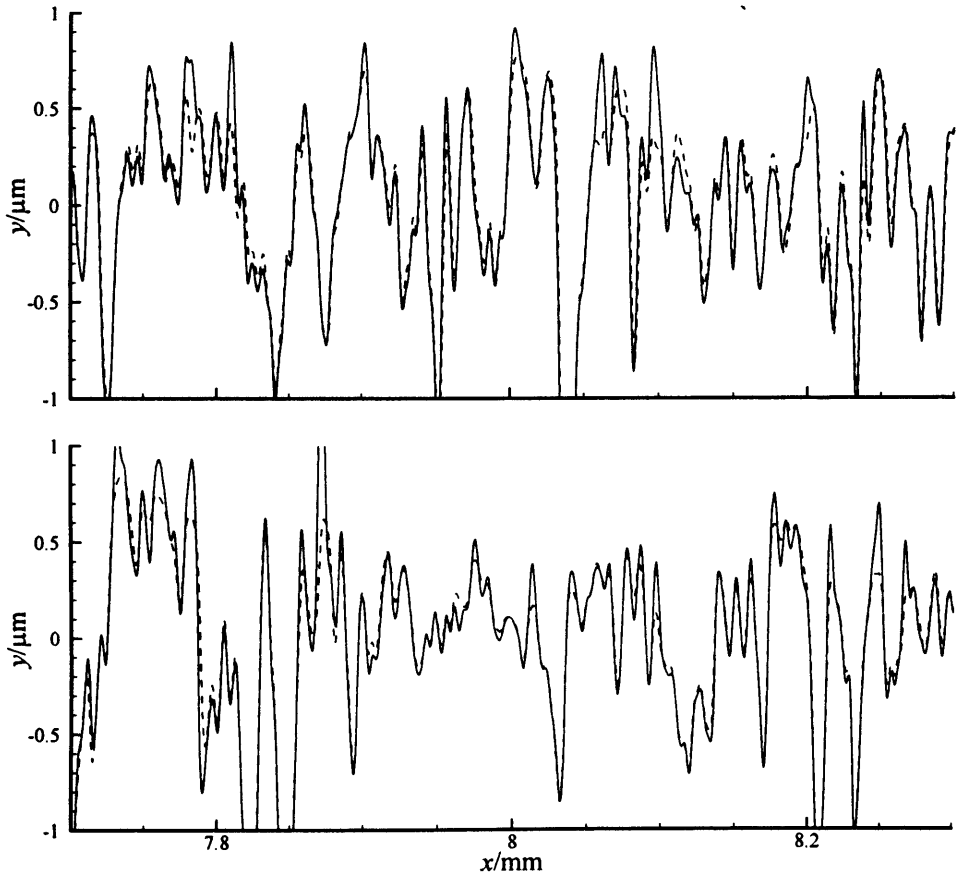

Figure 9.30 – Fast and slow disk roughness profiles at the centre of the contact region for test 0-R (solid) and test 5-R (dashed)

The change in the Ra value over the first load stage is slightly higher for the rough on rough configuration than the rough on smooth configuration. The friction response for both configurations is very similar and the features seen in the contact resistance are also similar. However, in contrast to what would be expected, the temperatures for the rough on smooth configuration are higher than for the rough on rough configuration.

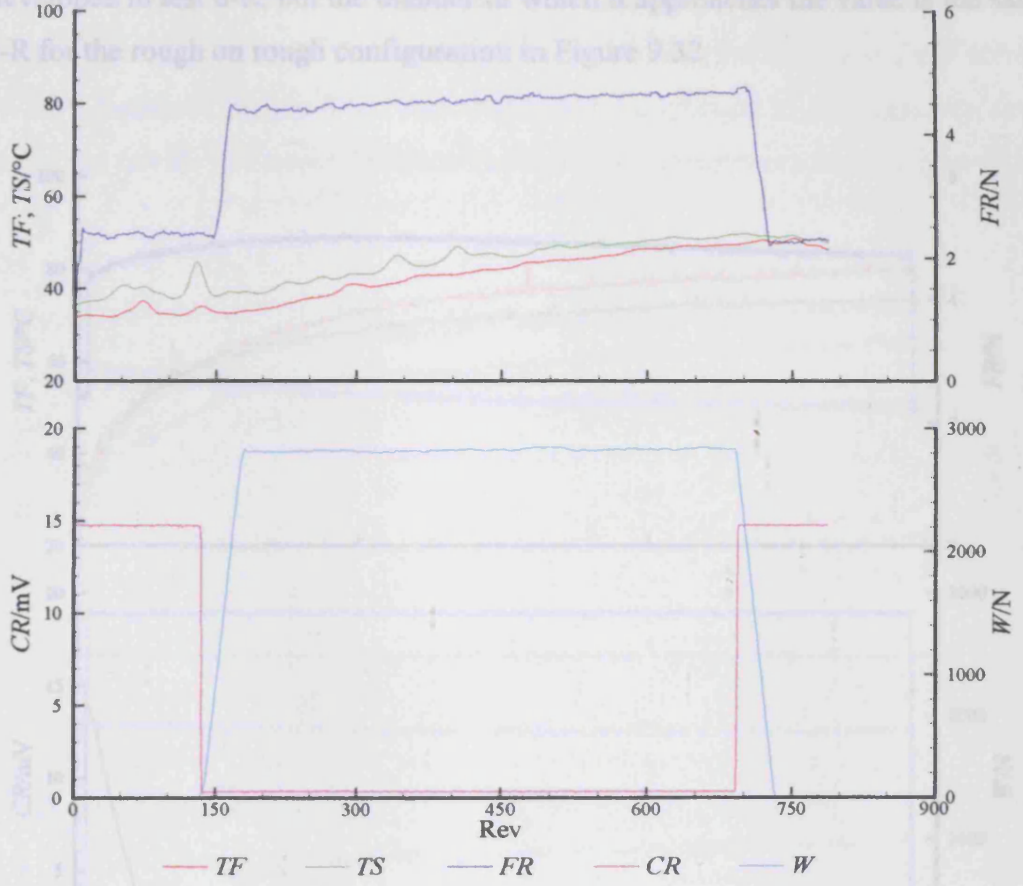

Figure 9.31 – Results for rough on rough test 6-R

The result for test 6-R, at the intermediate load stage, is shown in Figure 9.31. The temperatures developed are much higher than those seen for test 1-R. The slow disk temperature is still above the fast disk temperature for the test, but they converge towards the end of the test. The contact resistance response is different from all the previous results. When the load is applied the voltage measured across the disks drops significantly, reaching a value of less than 0.3mV. The voltage remains at this level until the load is removed, which is very different from the response seen in test 1-R. This

indicates that contact occurs for the majority of the time that the disks are in contact with one another.

Test 7-R develops the same response, where the friction continues to rise gently in spite of the rising temperatures. The temperatures of the disks cross over during the test and reach a final value of over 55°C. The contact resistance reaches the same low value as that developed in test 6-R, but the manner in which it approaches the value is the same as test 8-R for the rough on rough configuration in Figure 9.32.

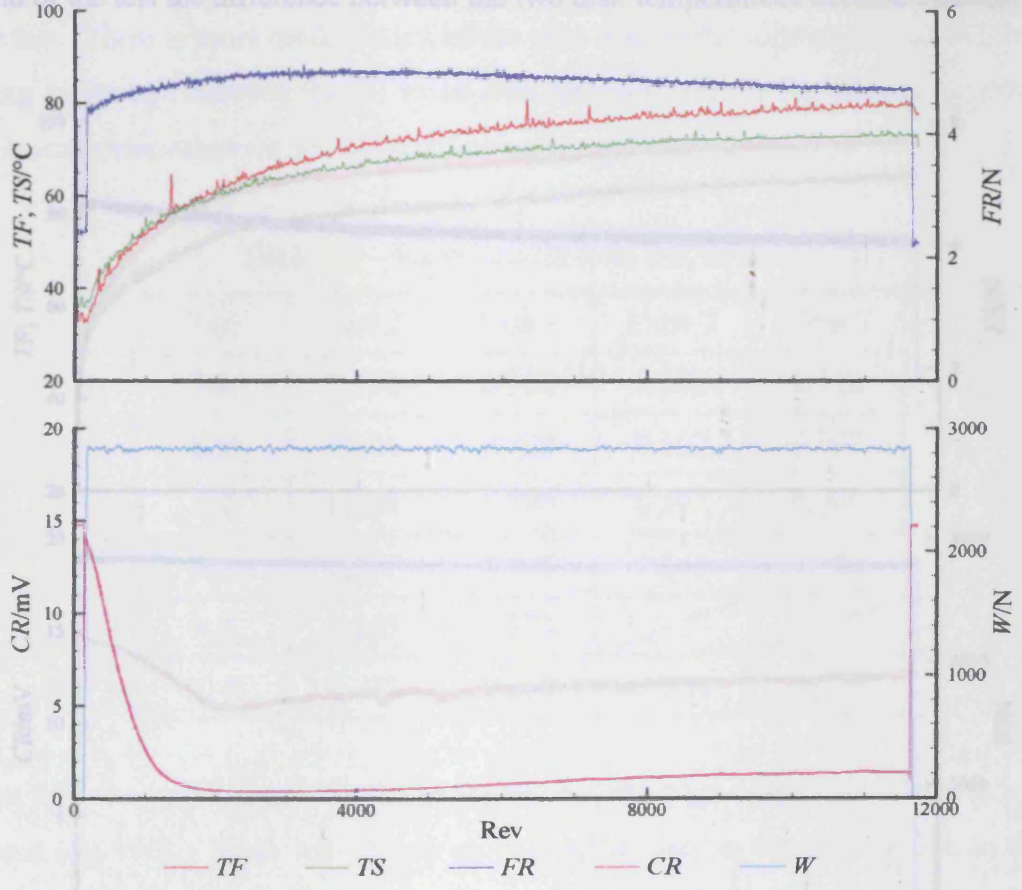

Figure 9.31 – Results for rough on rough test 8-R

Test 8-R, which is shown in figure 9.32, develops different responses in the contact resistance and friction to test 6-R and test 7-R. The friction continues to rise until revolution 3000, where it reaches a maximum value and then begins to fall for the remainder of the test. At the same point that the friction reaches it maximum value the contact resistance reaches its minimum value. After revolution 3000 the contact

resistance slowly begins to rise for the remainder of the test. The temperature of the fast disk continues to rise, it exceeds the temperature of the slow disk, and it then diverges from the slow disk temperature after this point finally reaching 75°C.

The response in test 9-R was very similar to test 8-R. The friction developed a maximum value and the contact resistance developed a minimum value, however, the two did not coincide. The friction was reducing when the contact resistance reached its minimum value. The minimum value of the voltage measured across the disks was 2mV compared with 0.3mV measured for test 8-R. The temperatures continued to rise, however, towards the end of the test the difference between the two disk temperatures became constant.

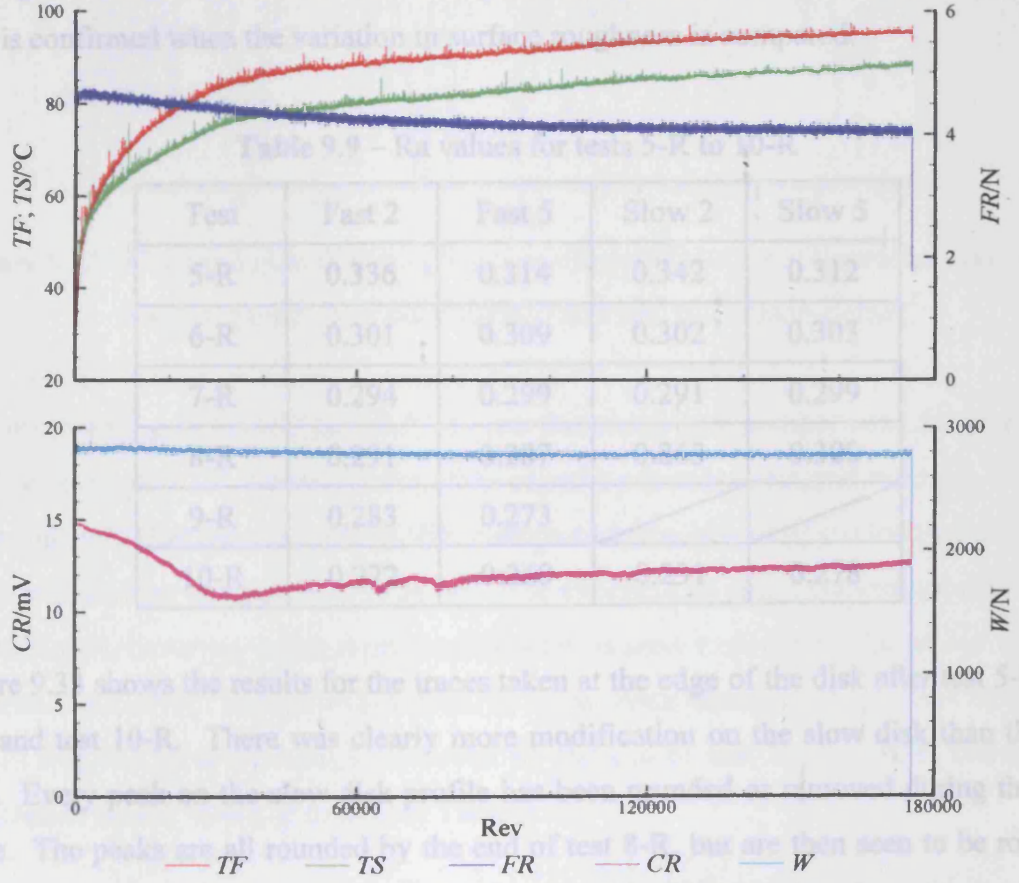

Figure 9.32 – Results for rough on rough test 10-R

The results taken from test 10-R are shown in Figure 9.32. The friction can be seen to reduce over the test, which is similar behaviour to that seen for test 5-R. The temperatures develop a constant difference of roughly 10°C after revolution 60,000. The

contact resistance develops a minimum value of 10mV, which is much higher than the minimum value reached in test 9-R.

The change in roughness for the centre line and edge of the contact are given in Table 9.9. Due to a software problem, the profiles taken after test 9-R on the slow disk were corrupted and cannot be presented. The relationship between the profiles at the centre of the contact and the profiles at the edge of the contact are different for each disk. The profiles at the edge of the slow disk are smoother than the profiles at the centre of the disk, however, the profiles at the edge of the fast disk are rougher than the profiles at the centre line. This trend was developed over the load stage, the results after test 7-R clearly show this. There is more modification of the slow disk at the edge of the contact, as there is a big difference between the Ra value after test 5-R and the Ra value after test 10-R. This is confirmed when the variation in surface roughness is compared.

Table 9.9 – Ra values for tests 5-R to 10-R

Test	Fast 2	Fast 5	Slow 2	Slow 5
5-R	0.336	0.314	0.342	0.312
6-R	0.301	0.309	0.302	0.303
7-R	0.294	0.299	0.291	0.299
8-R	0.291	0.287	0.263	0.300
9-R	0.283	0.273		
10-R	0.272	0.260	0.231	0.278

Figure 9.33 shows the results for the traces taken at the edge of the disk after test 5-R, test 8-R and test 10-R. There was clearly more modification on the slow disk than the fast disk. Every peak on the slow disk profile has been rounded or removed during the load stage. The peaks are all rounded by the end of test 8-R, but are then seen to be rounded and flattened more by the end of test 10-R. The fast disk also has the peaks removed, however, the amount of material removed or displaced from the peaks is much less than that seen for the slow disk.

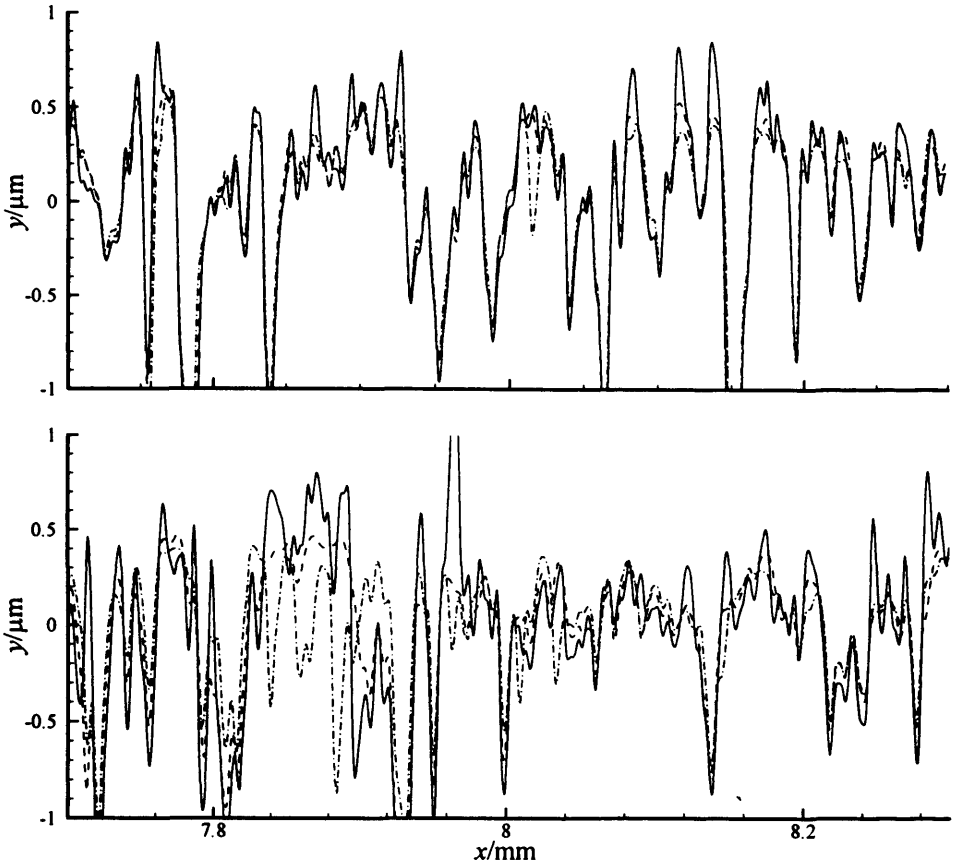

Figure 9.33 – Fast and slow disk roughness profiles at the edge of the contact region for test 5-R (solid), test 8-R (dashed) and test 10-R (dash-dotted)

In contrast to figure 9.33, Figure 9.34 shows the centre line comparisons from test 5-R, test 8-R and test 10-R where the fast disk experiences more material removal. The Ra value indicates that there is more profile change over the course of the load stage and the roughness profiles confirm this. The fast disk experiences slight flattening of the peaks after test 8-R, however, much more modification is seen to occur by the end of the load stage after test 10-R. The slow disk does not show much change in the surface by the end of test 8-R, but more material is removed by the end of test 10-R. These features are reflected in the Ra values presented in Table 9.7.

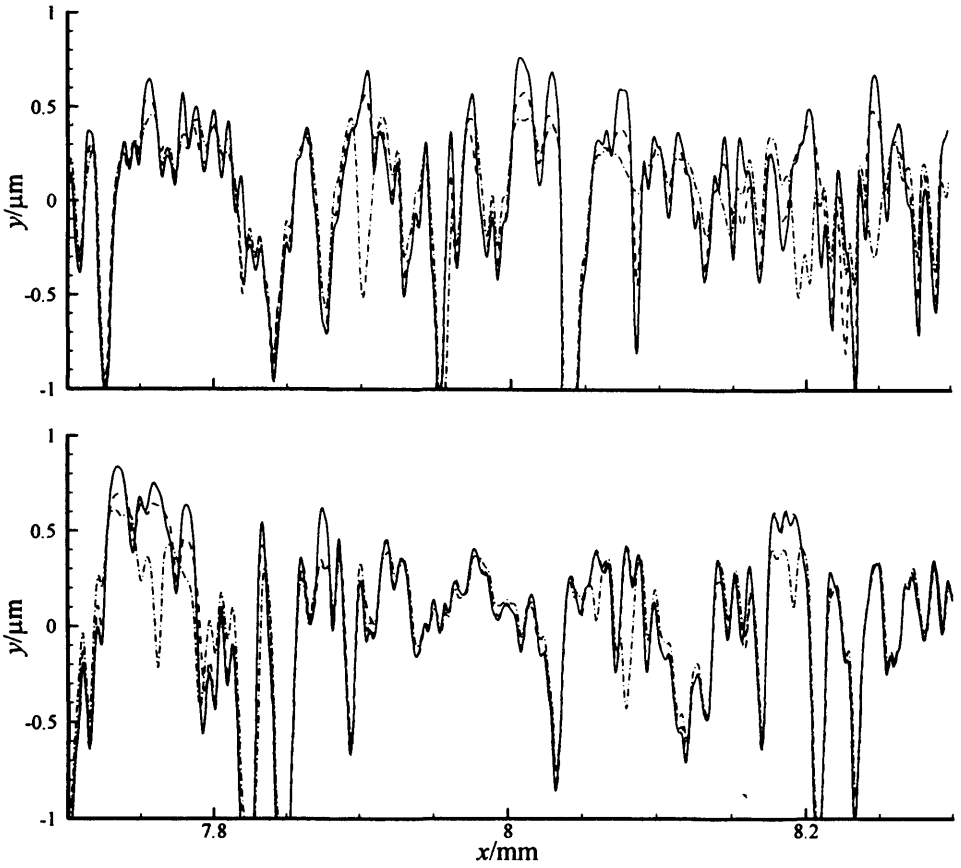

Figure 9.34 – Fast and slow disk roughness profiles at the centre of the contact region for test 5-R (solid), test 8-R (dashed) and test 10-R (dash-dotted)

The result for test 11-R, which is the first test at the highest load stage, is shown in Figure 9.35. The temperature and friction increase rapidly over the very short running period of 15 seconds. The fast disk and slow disk temperatures cross and reach 75°C before the temperatures drop when the load is removed. The friction rises quickly when the load is applied and is then seen to continue to rise as the test continues. The contact resistance voltage drops very quickly to a fraction of a milivolt, which is the same response as that seen in test 6-R.

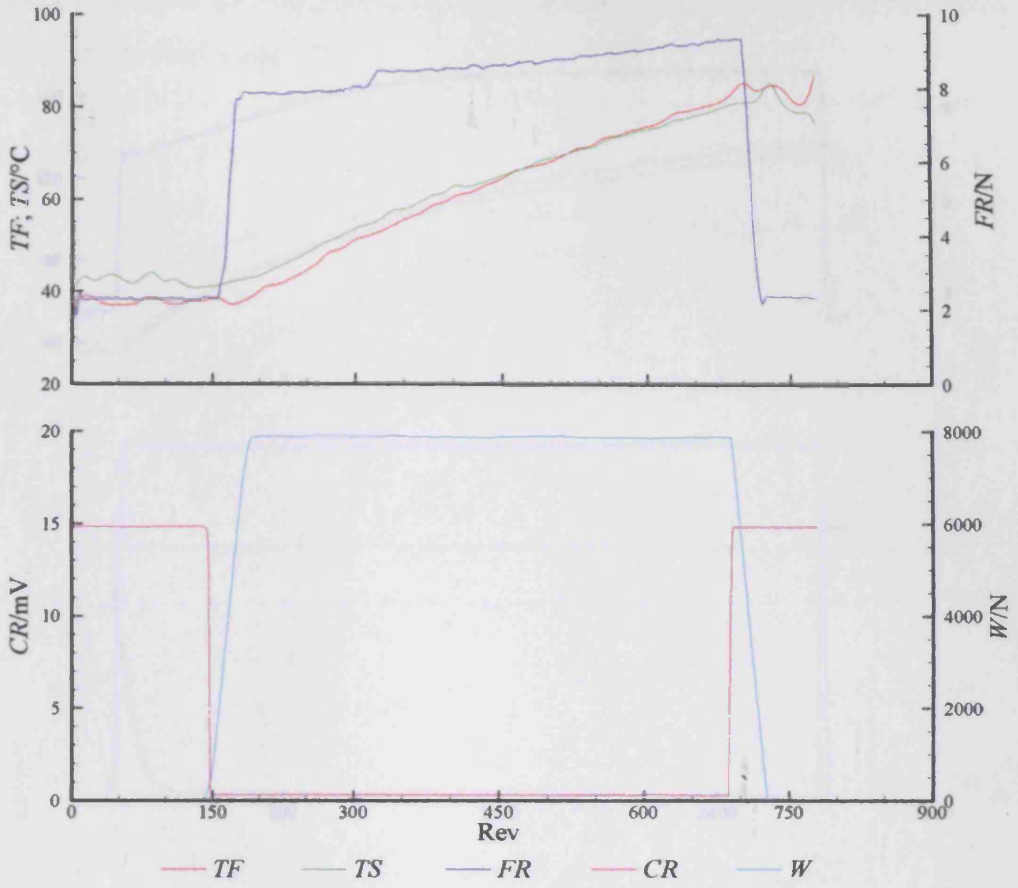

Figure 9.35 – Results for rough on rough test 11-R

The results of test 12-R are shown in Figure 9.36 and have the same features to those seen in test 11-R. The fast disk and slow disk temperatures cross over and both temperatures rise very rapidly, in this case reaching over 120°C. The friction rises until revolution 1600 after which it remains reasonably constant, however, the contact resistance is unaffected by this. The contact resistance reaches the same low value as in test 11-R and remains at this level for the remainder of the test. The maximum friction force is 10.3N, which is much higher than that seen for the second load stage of 5.1N.

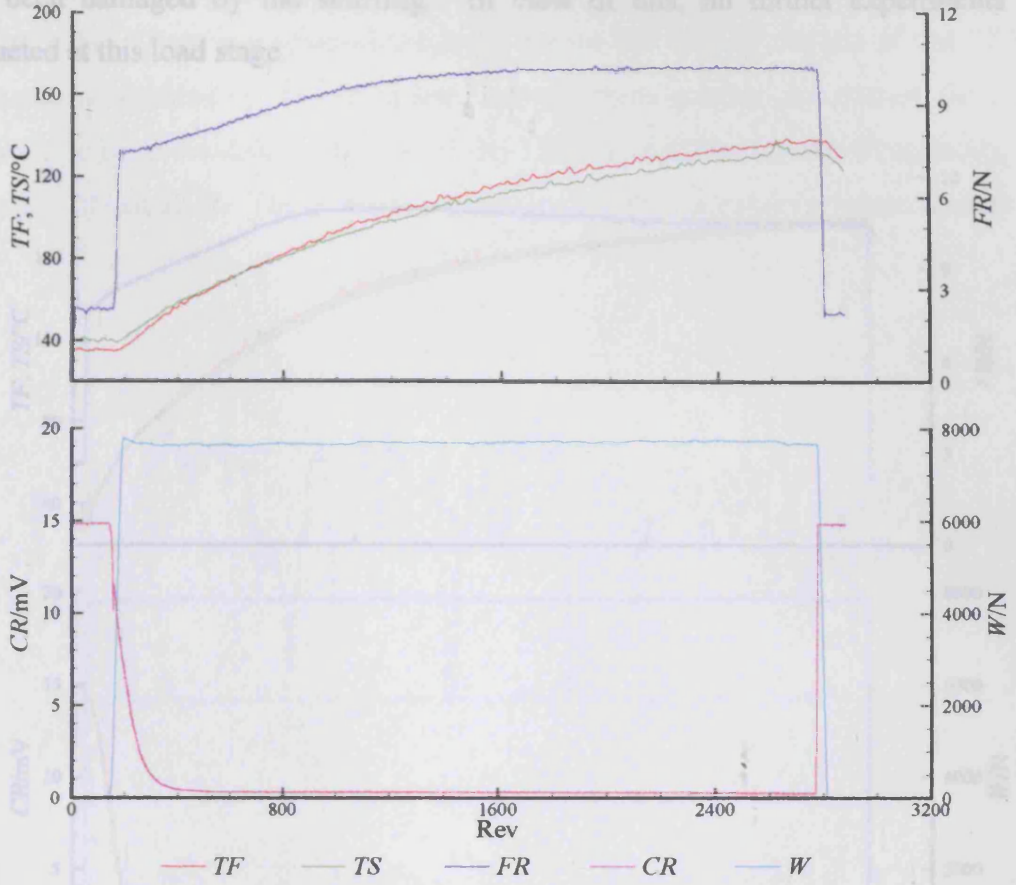

Figure 9.36 – Results for rough on rough test 12-R

The result for test 13-R is shown in Figure 9.37. The temperatures continue to rise very quickly, in under 4 minutes they exceed 170°C. Despite the dramatic temperature rise, the two temperatures do not differ much unlike the results seen at the other load stages. The friction response is very similar to test 12-R, rising and reaching a maximum, but in this case the friction does not remain constant after the maximum is reached instead it is seen to fall. The contact resistance response shows the same features as test 12-R, where the resistance falls to a minimum value that is maintained for the remainder of the test.

At the temperatures reached in this high load case there is a danger of scuffing as seen by Patching (1994). The experimental rig is not designed for scuffing experiments and the friction measurement was not sensitive enough to detect the onset of scuffing. If scuffing were to occur then it would take some time for the effects to be noticed in the friction and temperature response. It is feared that during this time most of the disks' surfaces would

have been damaged by the scuffing. In view of this, no further experiments were conducted at this load stage.

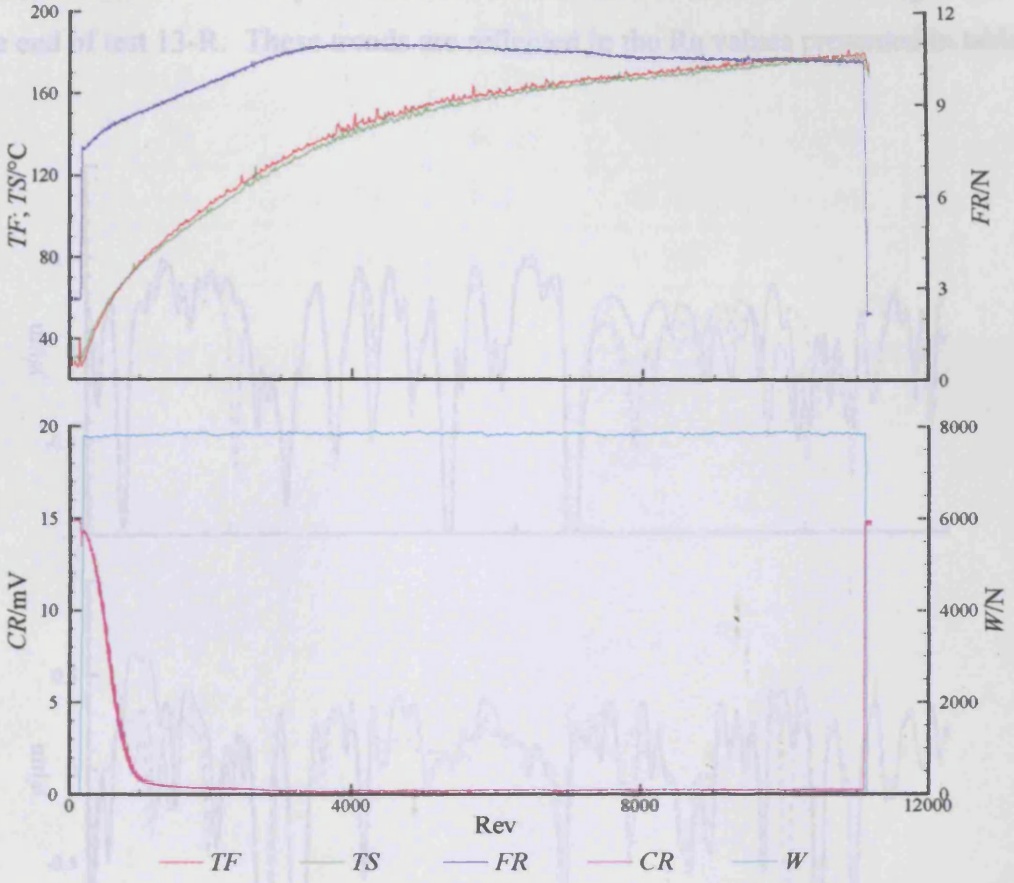

Figure 9.37 – Results for rough on rough test 13-R

The change in the roughness during these three tests is given in Table 9.10.

Table 9.10 – Ra values in μm for the highest load stage

Test	Fast 1	Fast 5	Slow 1	Slow 5
10-R	0.336	0.260	0.338	0.278
11-R	0.283	0.254	0.276	0.276
12-R	0.264	0.234	0.258	0.263
13-R	0.253	0.230	0.247	0.243

The variation in the roughness at the centre line for test 10-R, test 12-R and test 13-R is given in Figure 9.38. It can be seen that the peaks are already reasonably rounded by the

end of test 10-R, however, further modification occurred to the peaks during test 12-R and test 13-R. There is a noticeable change for the fast disk by the end of test 12-R, but the changes that occur by the end of test 13-R are much smaller. In contrast, the changes that occur to the slow disk by the end of test 12-R are smaller than the changes that occur by the end of test 13-R. These trends are reflected in the Ra values presented in table 9.8.

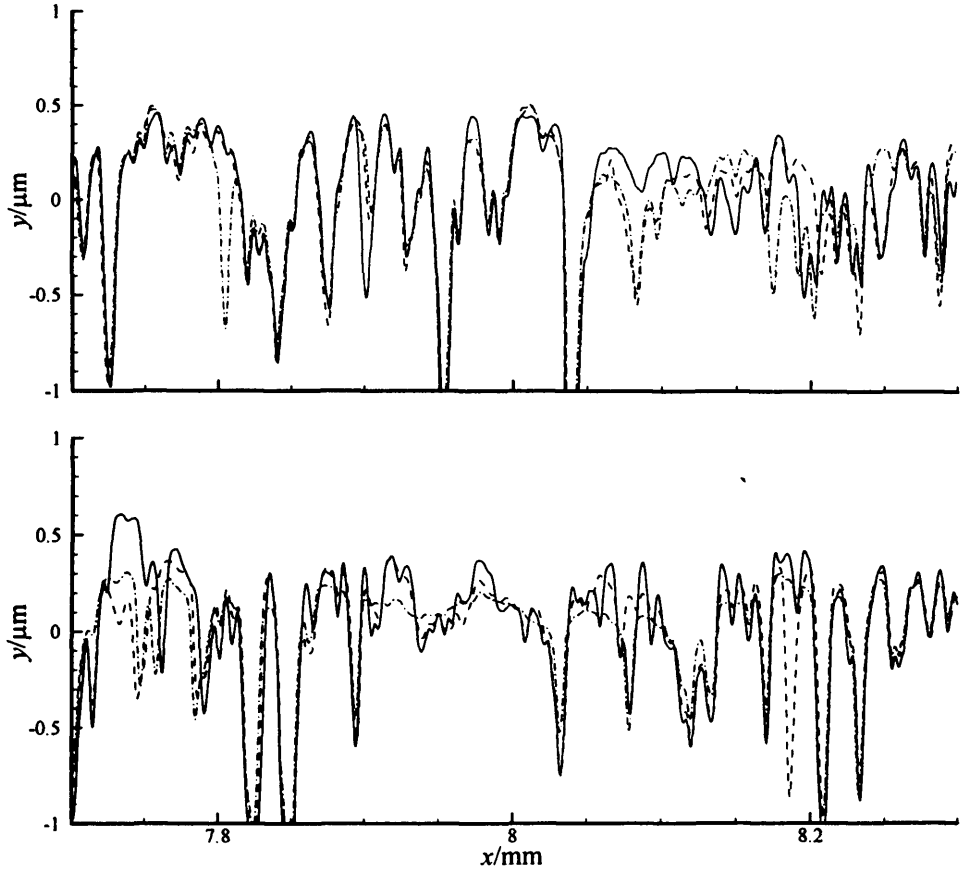

Figure 9.38 – Fast and slow disk roughness profiles at the centre of the contact region for test 10-R (solid), test 12-R (dashed) and test 13-R (dash-dotted)

It is worth considering the changes in the roughness that occur over the disk as a whole. Table 9.11 shows the change in the Ra value for some of the profiles taken over the tests considered.

Table 9.11 – Ra values in μm for different profiles for tests 0-R and 13-R

Profile	Fast Disk		Slow Disk	
	Test 0-R	Test 13-R	Test 0-R	Test 13-R
1	0.338	0.253	0.342	0.247
2	0.336	0.239	0.344	0.203
3	0.335	0.229	0.347	0.250
4	0.336	0.230	0.350	0.244
5	0.338	0.230	0.349	0.243

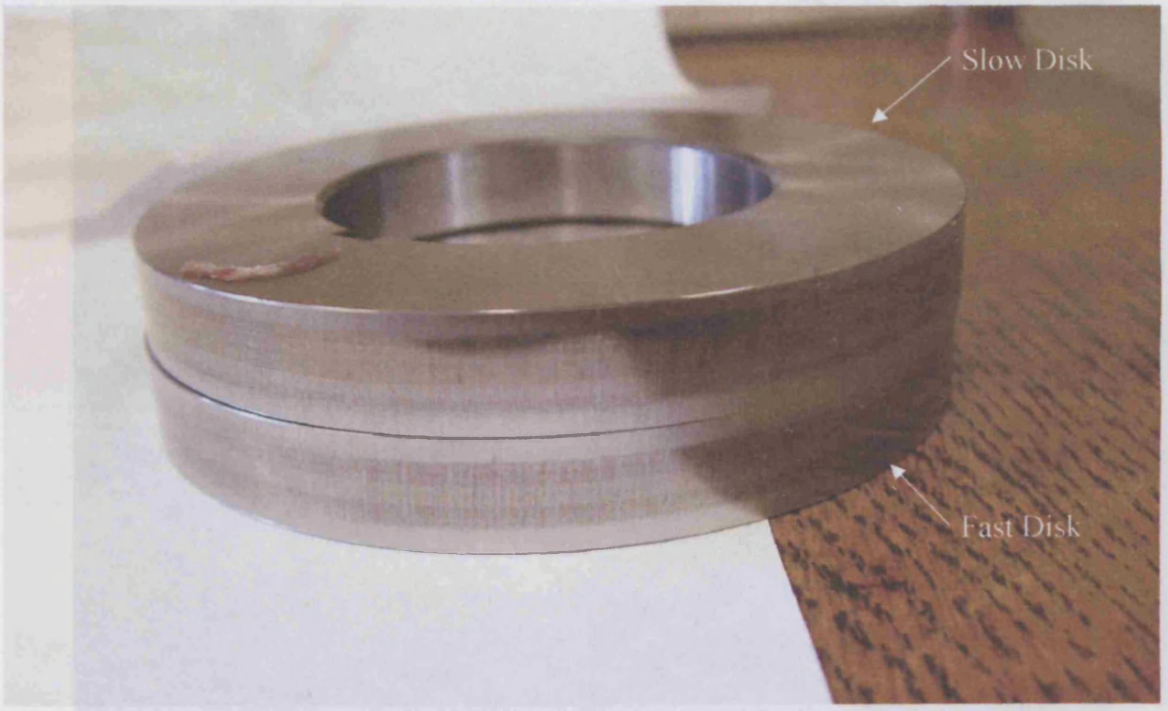

Figure 9.39 – Photograph of the two disks after test 13-R

The Ra values can be linked with Figure 9.39, which shows the surfaces at the end of running. It can be seen that there is a dark region in the centre of the slow disk, with a lighter region outside of it and then another darker region. The changes in the shade indicate the roughness of the surfaces, the lighter the surface the smoother the surface is. Therefore, the two lighter bands formed on the slow disk are related to profile 2, which has a lower Ra value than any of the other profiles. The fast disk has dark fringes and a light centre to the running track, which is to be expected considering profiles 3, 4 and 5 have a lower Ra value than profiles 1 and 2. There are marks in the middle of the dark regions on the fast disk that relate to the edge of the contact region for the intermediate

load. Outside of this line the surface has a darker appearance, which is also reflected in the Ra values.

Ignoring the profiles that make up the edge of the contact for the higher load stage, it can be seen that the profiles have an opposite effect to one another. Where the slow disk develops a smoother region towards the edge of the contact area for the intermediate load stage, the fast disk develops a rougher region. This effect is noticeable in the intermediate load stage results in table 9.9 and can be seen pictorially in the photograph of the surfaces.

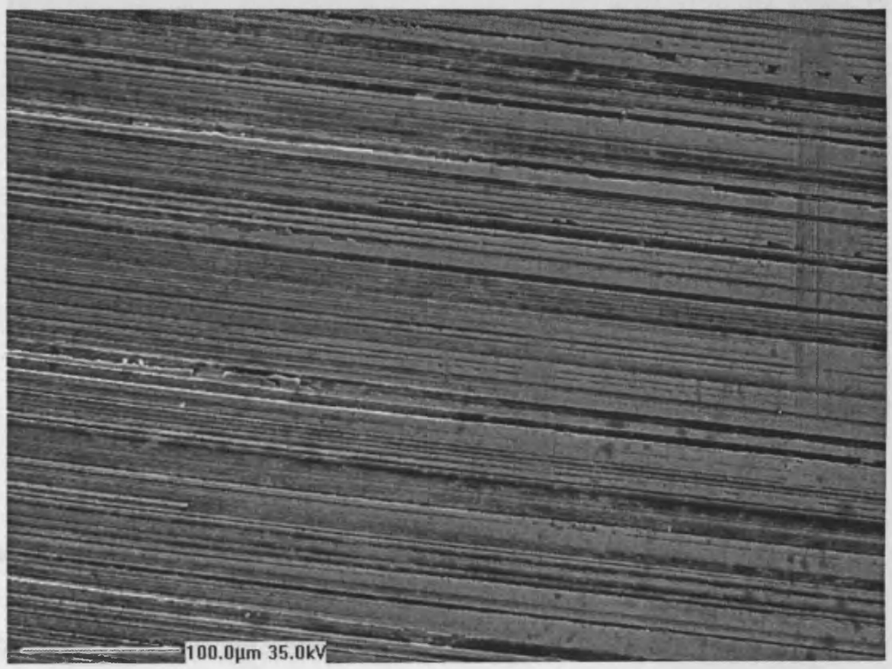

Figure 9.40 – SEM image of fast test disk after test 13-R

Figure 9.40 shows an SEM image taken from the edge of the contact region from the fast disk after the end of test 13-R. The right hand side of the image shows the smoothing of the surface, where the asperities are clearly more rounded. The modification that had taken place has caused plastic flow of the material, this causes the asperities to fold over slightly along their edges. This feature is particularly seen at the top right hand section of the figure.

The edge of the contact region is not defined by a straight line, but it can be seen that the surface becomes much rougher to the left of the image. The two regions where there are lines running vertically suggest that significant interaction of the surfaces has occurred, where there is significant smoothing as well as a scoring effect.

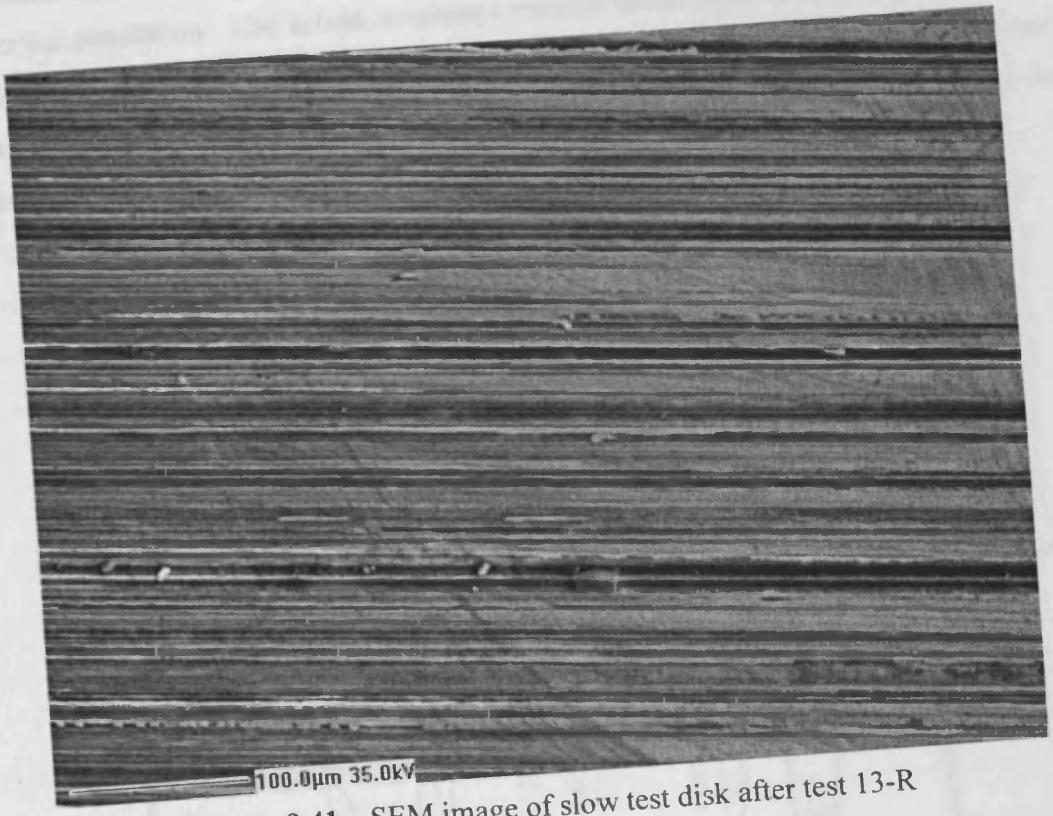

Figure 9.41 – SEM image of slow test disk after test 13-R

Figure 9.41 shows an SEM image taken from the slow disk after the end of test 13-R. The same features that are seen for the fast disk are also seen for the slow disk. The edge of the contact region is not better defined, but it could be argued that the surface to the right of the image is not as smooth as that for the fast disk. This feature also coincides with the Ra values presented in table 9.11, where it is seen that the fast disk is smoother at the edge of the contact than the slow disk.

9.4.3 Numerical Simulations

The numerical approach introduced in this work may be used to give an indication of how the asperities on the surface interact during the tests. The line contact parameters can be

equivalenced to give the same response as the centre line of the elliptical contact generated between the two disks. This is achieved in the same way that the line contact was equivalenced for the point contact in section 6.3. The bulk oil temperature and the bulk surface temperatures developed during the test can be used as an input to the numerical simulation. The actual roughness profiles taken from the surfaces can also be used which gives an indication of how the surfaces interacted during the test. The two tests of longest duration are analysed as they allow the widest temperature range to be investigated.

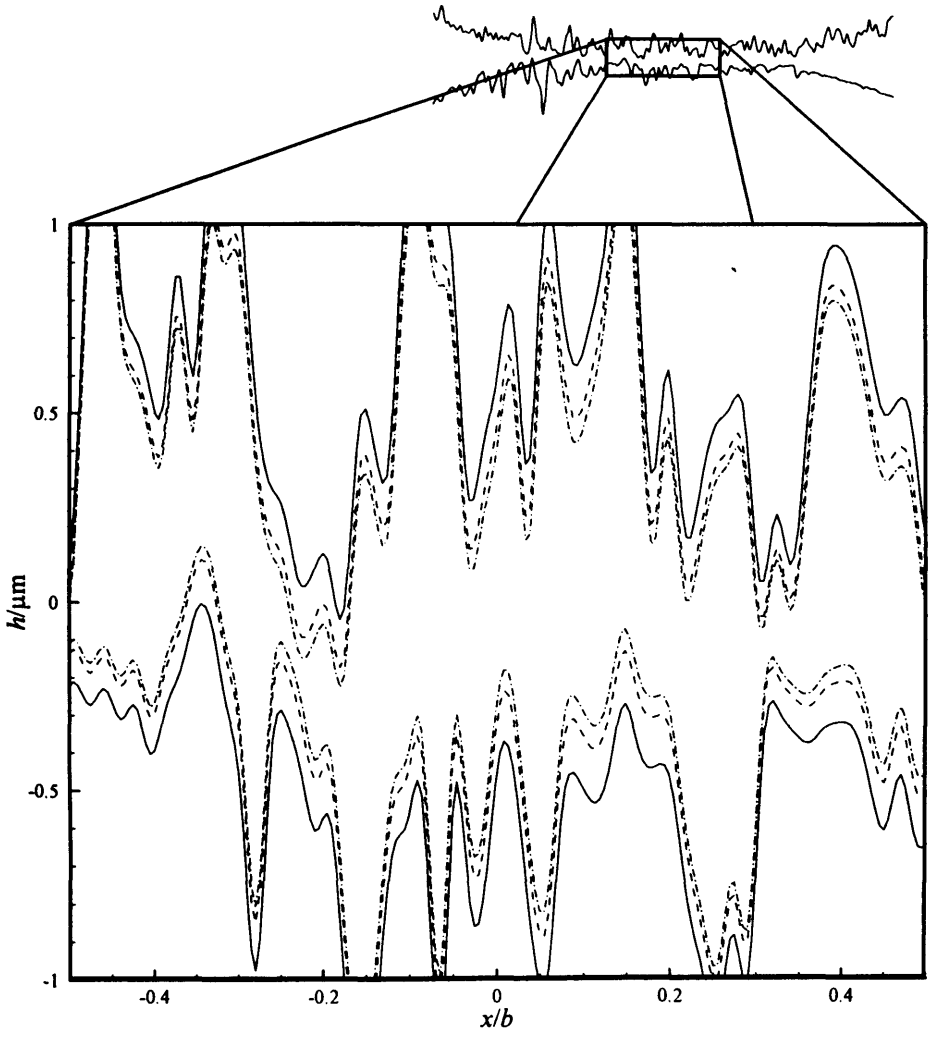

Figure 9.42 – Film thickness variation for test 5-R at revolution 0 (solid), 50,000 (dashed) and 180,000 (dash-dotted)

Figure 9.42 shows the oil film developed using the operating conditions measured at three points during the test. The profiles of the two surfaces are shown in the central region of the contact, from $x/b = -0.5$ to $x/b = 0.5$ at a particular time step in a transient analysis. The points used are revolution 0, revolution 50,000 and revolution 180,000. The profiles used in the simulation are the centre line profiles taken at the end of test 5-R, where the maximum Hertzian contact pressure was 0.7GPa.

The results of the simulations show that the surfaces are kept apart by the oil film throughout the region and it can be seen that the results for revolution 0 have the thickest films. This is due to the lower temperatures and corresponding higher viscosity. It can be seen as the test progresses that the surfaces move closer to one another as the temperatures increase.

The results also suggest that very little contact is developed even as the temperatures reach high values at the end of the test. This shows good agreement with the contact resistance voltage results obtained for test 5-R, where the voltage measured across the disks throughout the entire test was relatively high indicating that little contact occurs.

Figure 9.43 shows a series of results using the numerical approach for different points during test 10-R. The roughness profiles used are those taken at the centre line at the end of test 10-R. The bulk oil and surface temperatures are taken from the measured temperatures for test 10-R for the corresponding revolution.

It is clear that the effect of increasing temperature is to produce thinning of the film that allows the asperities from the opposing surfaces to contact one another. The temperatures are higher for test 10-R than for test 5-R and this is seen in the amount that the film thickness reduces over the duration of the test. It should be noted that the features on the surfaces seem rougher than those presented in figure 9.42, this is due to the larger contact dimension developed in test 10-R so a greater actual length of the profiles are shown in the results.

There is little difference between the results at revolution 50,000 and 180,000, despite the difference in the temperatures. The difference in temperature between the two results does give rise to a difference in film thickness of 0.05μm, but due to the scale used this

difference cannot be seen. It can be seen that the results for both of these revolutions show regions of contact, this is consistent with the results presented in figure 9.32 where contact occurred throughout the test.

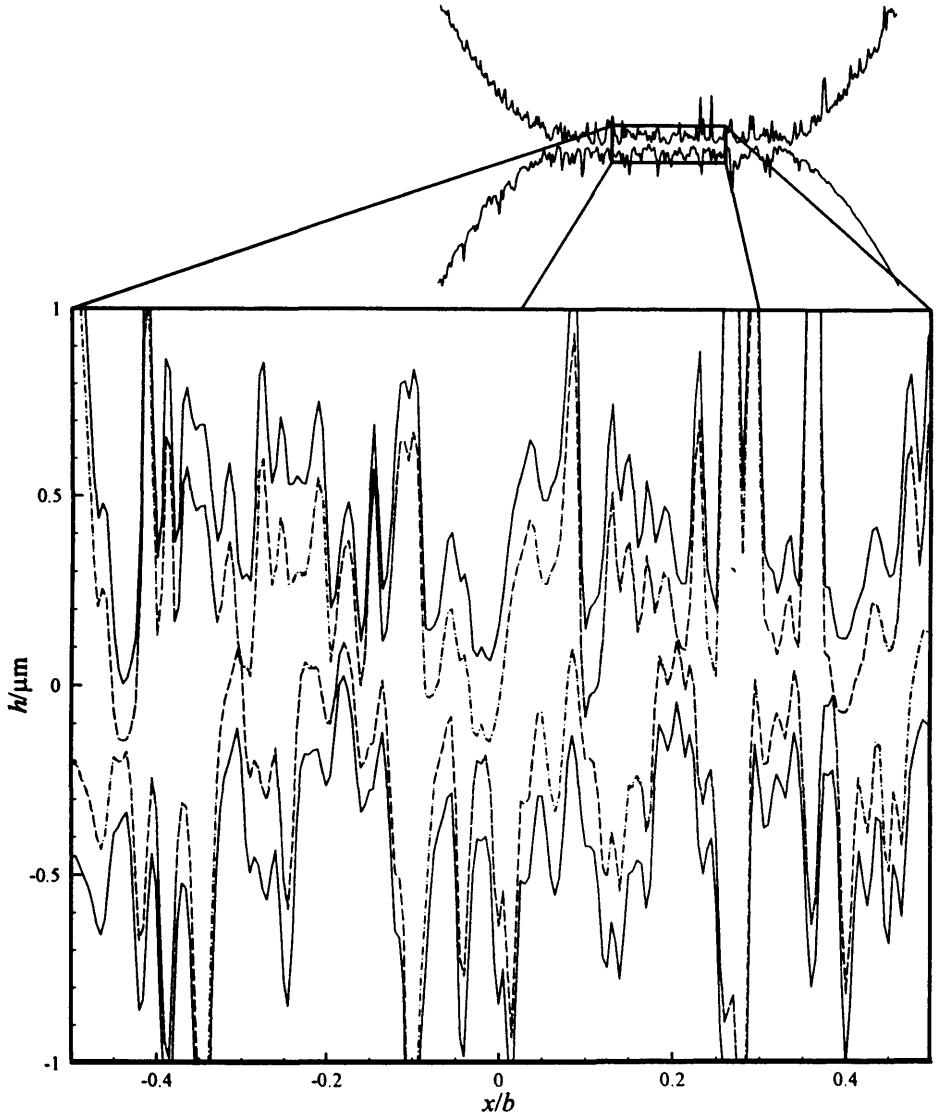

Figure 9.43 – Film thickness variation for test 10-R at revolution 0 (solid), 50,000 (dashed) and 180,000 (dash-dotted)

Reasonable agreement is observed between the numerical simulations and the actual results obtained during the test. It can be seen that the rise in the bulk disk temperatures over the duration of the test result in large changes in the film thickness, which allows the surfaces to generate greater amounts of contact. This response is seen in figure 9.32, where the amount of contact is seen to increase at the start of the test before reaching a

minimum value and then increasing again. This indicates that the temperature rise brings the surfaces closer together generating increased amounts of contact. As the test progresses the rate at which the temperatures rise begins to fall, slowing down the rate at which the film thickness reduces. The rate of film thickness reduction becomes lower than the rate at which the surfaces modify, allowing the amount of contact to reduce.

9.5 Summary

Modifications to an existing experimental rig have been conducted that allow both 'in situ' Talysurf measurements and electrical contact resistance measurements to be made. The modifications allow investigations into 'running in' to be undertaken, where the amount of metallic contact can be detected using contact resistance measurements and correlated to the changes in the surface roughness that occur as the disks run against one another.

The modifications necessary to undertake the contact resistance measurements required electrical insulation of one of the test disks and the shaft on which it is mounted. This was achieved using a variety of methods to remove the current paths that existed within the rig. Further mechanical modifications were also necessary to allow the electrical connections to be made to the disks.

Electrical circuitry was designed to allow the temperature and contact resistance to be measured during a test. Other electrical circuitry was incorporated to allow the friction and load to be acquired using a computer data acquisition system, as well as more complex circuits that handle shaft encoder pulses. The electrical circuitry was combined with a PC based computer acquisitions system using LabVIEW.

A virtual instrument was developed that acquires the load, friction, bulk disk temperatures, contact resistance voltage as well as an optical reflection signal during a test. The acquisition of the signals is controlled by a high resolution shaft encoder allowing data to be acquired 5000 times a revolution. The virtual instrument allows the contact resistance and optical reflection signals to be acquired at this resolution, as well as

acquiring the friction, load and temperatures once per revolution. The software also determines the speed of rotation of the disks using trigger pulses supplied from the shaft encoder.

The modification to allow the 'in situ' measurements of the disks involved providing a means of accessing the disks using a Taylor Hobson Talysurf Intra. In addition to the design problem of accessing the disks, accurate repeatability and the necessity to measure at various axial positions along the disk were also necessary. A solution was found that used a stage with pre-drilled holes to locate the Talysurf and slides to move the table so that the Talysurf could measure at any location along the width of the disks. Relocation in the axial direction was provided by determining the edge of the disk. Relocation of the disks in the circumferential sense was conducted by eye, marks were made on the disks and then aligned with a set datum for each profile. Comparisons of profiles taken from the same nominal position could then be aligned by reference to clearly discernible valley features.

The modifications allow a series of test to be conducted where, at a set speed, the load is increased and the effect that this has on the friction, temperature and amount of contact can be observed. Each load that the disks are to be run at can be broken down into running times that steadily increase. At a fixed load a number of tests with increasing duration are run and the data acquired. In addition, at the end of each test Talysurf profiles can be taken from the surfaces that allow the effect that the test has on the surface features of the disks to be considered.

Due to the manner in which the tests are to be conducted, changes needed to be made to the way in which the load was applied. The load was applied manually using a control valve, setting this at the beginning of each test introduced a large source of error. Instead a ball valve was fitted so that the load could be applied and removed quickly, removing any human error. Also, an approach was developed in which the load could be applied without the disks touching. This ensured that the load could be pre-set without the disks interacting.

Results are presented for three maximum Hertzian pressures of 0.7GPa, 1.5GPa and 2.1GPa. At each of these loads the width of the contact could be determined and the

Talysurf measurements spaced to allow measurement of the centre line of the contact as well as regions near the edges of the contact areas. The running times generally used were 15 seconds, 1 minute, 4 minutes, 15 minutes and 60 minutes, which gave a total running time at each load stage of 80 minutes 15 seconds.

Two experiments were run, one with a rough disk on the fast shaft against a smooth disk and another where two rough disks were used. Case-carburised alloy steel disks are used which have a surface hardness of 650Hv. The rough disks were axially ground and the smooth disk was axially ground and then superfinshed. The Ra values of the rough and smooth disks were 0.35μm and 0.05μm respectively.

The rough on smooth test was the first experiment conducted using the rig and some initial problems were encountered. The slow disk thermocouple failed during the first test and there were also software and electrical problems during the first load stage tests. In spite of the problems, meaningful results were obtained that showed the changes in the roughness and amount of contact developed for Hertzian pressures of 0.7GPa and 1.5GPa. The results for the highest load stage could not be achieved due to a movement in the locating bearing in the slow shaft leading to a misalignment of the two disks.

At low loads and short running times the temperatures developed were low and the films relatively thick. As a result, the surfaces modify quickly and the amount of contact developed rapidly reduced. As the tests were run for longer, or higher loads applied, the temperatures begin to increase and the amount of contact developed did not reduce as quickly. The friction force was found to increase initially and then reduce as the temperatures increased, the reduction in the friction occurred as a result of lower viscosity developed at higher temperatures.

The rough surface developed small reductions in Ra at the lightest load stage, but there was little difference seen in the values taken from the smooth disk. At the intermediate load stage it was found that the rough surface experienced greater modification, particularly at the edge of the contact region. The trend on the rough surface was to reduce the Ra value by removing the aggressive asperity peaks. On the smooth disk, the interaction acted to roughen the surface slightly.

For both load stages it was seen that as the load was re-applied and the tests run for longer, the amount of contact developed was lower than the amount developed at the end of the previous test. The temperature may have some slight influence, as contact will increase the surface temperatures, however, this effect was thought to be small. The drop is more likely a result of the surface modification that occurs at the beginning of the test when the temperatures are rising. Over this period the surfaces modify further, which will then develop less interaction when the higher temperature and thinner films are reached.

When two rough surfaces were run against one another, the amount of contact developed was much higher than that developed between a rough disk and a smooth disk. This is to be expected, as the combined amplitude of the roughness for the rough on rough case is almost double that for the rough on smooth case. The slow disk gave temperatures that were higher than the fast disk, but the fast disk was seen to generate higher temperatures as the test continued and exceeded those of the slow disk as the test progressed. The crossover point of the two temperatures reduced, in terms of time, with increasing load. The friction response was found to show the same features as the rough on smooth cases and also produced similar values.

The amount of contact developed for the rough on rough conditions was much larger than the rough on smooth. For most of the intermediate load stage there were large amounts of contact developed. All of the tests prior to test 9-R developed periods where the voltage measured across the disks was below 1mV.

At the highest load stage even more contact occurred. Unfortunately, the load stage could not be completed; the last test run was for 4 minutes and the bulk disk temperatures exceeded 170°C. At this temperature the disks were thought to be prone to scuffing, subsequent tests would have exceeded this temperature and possibly scuffed the disks. Since it was important to retain the features that were present on the disks, scuffing should be avoided. Thus the tests were stopped at this point.

It was seen that the features developed on the disks are markedly different. There are regions within the running track where there are rougher and smoother areas. The rougher and smoother areas do not occur at the same location on both disks, in fact they

are opposite. Where the fast disk develops a smoother area, the slow disk develops a rougher area. This was clearly seen in figure 9.39, where there were light and dark bands developed around the circumference of the disks. The light band corresponded with smooth areas and the dark bands with rough. The shading can also be linked to the Ra values taken from the surfaces. The light and dark bands were not generated in the same locations on both disks. It was found that in the regions where the slow disk developed light bands the fast disk developed dark band and vice versa.

It had been hoped to use the high resolution shaft encoder and high speed data acquisition system to detect exact instances of contact that could be quantified using numerical approaches. However, the presence of a low pass filter averaged the values needed to make this judgement. Instead a qualitative appreciation of contact could only be developed. Despite this problem, the hardware and software has been designed to allow this to be conducted and only the low pass filter circuitry would need to be re-designed.

Despite the inability to quantify the amount of contact, numerical results were presented that confirmed the contact resistance values obtained. Numerical simulations of test 5-R and test 10-R were conducted at different points during the test duration. The results for test 5-R did not produce large drops in film thickness as the temperatures increased and did not develop contact. Test 10-R experienced a greater thinning of the film and developed instances of contact. Both of these simulations showed good agreement with the results from the actual experiments.

The 'in situ' Talysurf measurements with the ability to relocate on surfaces have proven to be very useful in the study of profile changes during running. The same features are present in the profiles and the profiles clearly show how the surfaces are modified as the disks run against one another. Combined with the remainder of the changes developed it is clear that a full study could be conducted into the amount of surface modification and contact using the existing arrangement.

Chapter 10

Conclusions and Recommendations

10.0 Introduction

This chapter outlines the major conclusions that can be drawn from the work described in the preceding chapters. Also presented are some suggestions for future work that could build upon the results and developments presented in this thesis.

10.1 Conclusions

The line contact EHL problem was formulated to include the effects of different rheological models on the pressure and film thickness within the contact. It was seen that limiting shear stress behaviour causes a reduction in the pressure gradient near the centre of the contact, leading to an increase in the pressure generated in the inlet region. Limiting shear stress effects also change the film thickness profile, where 's' shaped kinks are developed. Instead of developing the usual flat parallel region through the centre of the contact, limiting shear stress models change the response so that a localised increase in the film thickness is developed in the inlet and a localised decrease is formed in the exit region.

The introduction of limiting shear stress rheology was also found to develop slip between the surfaces and the fluid at their interface. Slip was found to occur on the faster of the two surfaces in the inlet, which acts to reduce the entrainment mechanism. Slip is also seen to occur in the exit part of the contact on the slower surface, which increases the fluid flow in this region. These effects increase the amplitude of the kink developed in the film thickness and also lead to a flatter pressure distribution.

Different operating parameters were varied and the response that this had on different rheological models assessed. It was found that rheological models that change from viscous to plastic behaviour aggressively develop larger amplitude film thickness kinks than those which have a more gentle transition. It was also found that the results obtained using the different models were very sensitive to the viscosity generated within the contact. This was illustrated by changing the bulk oil temperature and comparing the isothermal results. It was seen that small changes in temperature could generate large changes in the pressure and film thickness profiles.

As a result of the change in the response of the limiting shear stress models to temperature, a full energy equation was introduced to determine the temperature variation through the contact. The energy equation was developed to account for the heating influence that slip has as well as allowing the shear heating effects of the different rheological models to be included. The temperature rise generated through the contact reduced the limiting shear stress features that were generated in an isothermal environment. The only way to introduce the limiting shear stress effect was to reduce the velocity so that the temperatures generated through the contact became low.

Using low entrainment velocities and operating conditions that correspond to optical interferometry experiments, it was found that the limiting shear stress features could be reproduced. The velocity was increased and it was found that there was a region in which the limiting shear stress effects dominated the response of the pressure and film thickness. At low speeds the temperatures are low, which allow the viscosity to remain high in the centre of the contact thereby generating large limiting shear stress effects. As the speed increases the limiting shear stress increases up to a point, after which it reduces due to the increase in the temperature. When higher temperatures are reached the viscosity drops substantially in the centre of the contact, removing the influence that the limiting shear stress has. Clearly two domains exist, one where the limiting shear stress effects dominate the response and another where the temperature effects dominate the response.

In addition to including the effects that the limiting shear stress rheology and slip have on the temperature, the full energy approach was adapted to include the heating effect generated when the two solid surfaces contact one another. This allows the temperature to be calculated throughout the contact region even when the fluid film breaks down.

Two sinusoidal features (many times larger than the smooth surface minimum film thickness) that developed instances of contact were used to test the approach adopted. The surface temperature developed at the contact points begin to rise over time, as the contact develops a greater heat input to the surfaces than when there is a film present. The heat penetrates deeper into the surface, resulting in an increase in temperature over time. The effect that the assumed coefficient of friction at 'dry' contacts has on the temperature was also briefly investigated. As expected, higher coefficients of friction generate greater temperature increases near the contacting points, but the results are not markedly different. Despite the large increase in the heat generated at the interface the temperature rise was small, indicating that the heat generated is easily dissipated within the solids.

A transient EHL solution technique was developed to simulate the meshing of gear teeth, which allows both smooth surface and rough surface conditions to be simulated. The simulations could use equivalent roller or true involute geometries and showed slight differences between the two geometries. The roughness on the gear teeth was related to the parts on the gear teeth on which it was present. This allowed the simulations to reproduce the exact conditions at gear tooth contacts as they varied during the meshing cycle, in terms of load, relative geometry, surface speeds and local roughness features.

It was found under thermal results that the maximum pressure and maximum temperature generated increase suddenly when the first change point within the meshing cycle is reached. This also gives rise to a reduction in minimum film thickness. All of these features act to increase the possibility of surface fatigue effects being generated within this region. This may well describe the reason for the micro-pitting that occurs near the pitch point of gears, as the change points are generally located within this region.

In addition to the above theoretical work, experimental work was conducted to investigate running in of ground steel surfaces under EHL conditions. Electrical contact resistance and 'in situ' Talysurf profiles were taken during several tests to show the degree of contact and the surface modification developed. The results indicated that the modification of the surfaces occurred quickly, but that the temperatures of the surfaces increased reducing the film thickness and prolonging the time that contact occurred. At

the higher loads more contact was generated which persisted for longer periods of time, which is the result of the temperature increase developed over the test duration.

Upon examining the disks after the experiment, it was found that there were different degrees of surface modification developed across the running track. The difference in the roughness of the disks was also clearly seen by eye, where light regions occurred for smooth parts of the disks and dark regions for rougher parts of the disk. The light regions on the slow disk corresponded with the dark regions on the fast disk and vice versa.

It should also be noted that numerical simulations showed good agreement with the experimental results presented. Numerical simulations were achieved using roughness profiles taken from the disk and also included the measured bulk oil and bulk disk temperatures developed through the test. Different simulations were undertaken to represent different periods of the test duration. It was seen that there was good agreement between the numerical simulations and the experimental results, where the amount of contact developed in the numerical simulations correlates well with the contact resistance voltage acquired during the tests.

10.2 Recommendations

The equations developed for the fluid film did not account for the detailed variation in the viscosity and density across the film. In deriving the hydrodynamic Reynolds equation it was assumed that the viscosity and density were constant across the film. This assumption is acceptable for the isothermal case, but when a thermal approach is used this may not be appropriate. The work described used an average temperature when determining the viscosity and the density in the Reynolds equation. Similarly, the equations developed for the non-Newtonian rheology did not take into account the variation in the viscosity across the film; the viscosity used was determined using the average temperature. In future work both the non-Newtonian and Reynolds equations should be re-expressed to account for the change in viscosity and density with temperature.

Experimental work could be conducted under the operating conditions outlined in section 6.3 to determine if there are any kinks present in the film thickness. This may allow the features associated with limiting shear stress to be seen in an optical interferometry experimental arrangement. The results could then be used to determine which limiting shear stress model is most appropriate. It may transpire that different oils require different limiting shear stress equations to describe their response. For a true representation and validation, the limiting shear stress models should be adapted as described above and employed within a point contact problem.

The different limiting shear stress models could be compared using rough surfaces or sinusoidal features. This would allow further understanding of the effects that they will have in real EHL situations. In addition, the effects that different models have on surface damage, sub-surface shear stress and fatigue could all be determined.

The thermal approach adopted has to use an implicit method for the derivative of temperature with time term when the film thickness becomes small. This implicit method takes longer to converge than the Crank-Nicholson scheme, but is more robust when small film thicknesses are encountered. A hybrid approach incorporating the two methods may be the best option. The method could use the implicit method for points where the film thickness is small and the Crank-Nicholson approach for all other points. This would provide the possibility of exploiting the time saving effects of the Crank-Nicholson scheme and the robust features of the implicit method. The thermal approach may also benefit from a finite element formulation rather than the difference formulation used in this work.

A more detailed study of rough surface effects under thermal conditions should be conducted. The work here considered sinusoidal roughness features, which do not generate large discontinuities in the film thickness, whereas real rough surfaces generate very rapid changes in film thickness. Large pressures are needed to deflect the asperities on the surfaces, which could give rise to very large surface temperatures if the asperities contacted one another.

The simulation of gear meshing would benefit from a refinement in the way that the load is controlled in the rough surface solutions. Using the separation of the two surfaces or

some other variable may remove the deviations from the desired load over the meshing cycle. In addition, the variation of the quantity used to describe the load over the meshing cycle will need to be determined. Values taken from the smooth results would not give the correct load as the asperities on the surfaces interact and generate higher pressures. Currently the variation in the first node film thickness is set and the simulation run, the results of the load carried are compared to the desired load and the first node film thickness adjusted where necessary. This approach is very time consuming, and a better approach needs to be determined. It may be possible to run steady state comparisons of individual points within the meshing cycle and the separation adjusted to attain the correct load. This would allow the variation of the desired parameter to be known at certain points of the meshing cycle and the simulation run using these values.

The gear meshing simulation could be refined further where changes could be made to the conditions used at the first point in the meshing cycle. The temperatures would not, in practice, correspond to the steady state values as they would be quickly heated from the bulk temperature. In addition, there will be some squeeze effect due to the manner in which the load is applied to the teeth quickly. More careful examination of the results from finite element analysis may yield a better idea of what conditions to use.

The experimental approach adopted does require some modification to the contact resistance circuitry if quantitative studies of contact are to be undertaken. The low pass filter is currently preventing such studies being conducted. However, removing the filter will subject the results to background noise. Either filtering out this noise using a tuned filter or trying to locate and eliminate the source of the noise would allow the current filter to be removed.

Further experiments could be conducted to investigate the influence that the speed has on the amount of contact and surface modification developed. Lower operating speeds will generate thinner films but lower temperatures, which may allow the surfaces to modify quickly. Lower speeds may develop contact over a shorter period of time than that seen at higher speeds due to the lower influence of the temperature. Also, further load stages could be used, this would allow the different finishes seen on the two disks to be studied further.

Finally, the results presented could be plotted in terms of non-dimensional pressure and film thickness. This will allow comparisons to be made between the results of Lee and Hamrock (1990), Shieh and Hamrock (1991) and Hsiao and Hamrock (1992) using the circular limiting shear stress model. This would allow validation of the method and approaches presented within this work with another approach.

Appendix A

This is a continuation of the work conducted in section 3.2. In section 3.2 the derivations only considered one case, however, there are eight cases that need to be considered. Here each case will be covered in turn and general equations developed. From Section 2.8 equation (2.22):

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

Using the same substitution outlined in sections 3.2, 3.3, 3.4 and 3.5 gives:

$$\tau = \tau_l\left(\lambda + \frac{2\Sigma}{h}z\right) \tag{A.1}$$

This yields the shear stresses at the upper surface τ_1 and lower surface τ_2 as:

$$\tau_1 = \tau_l(\lambda + \Sigma) \qquad \tau_2 = \tau_l(\lambda - \Sigma)$$

Consider the following figure where λ, Σ and the shear stresses are illustrated:

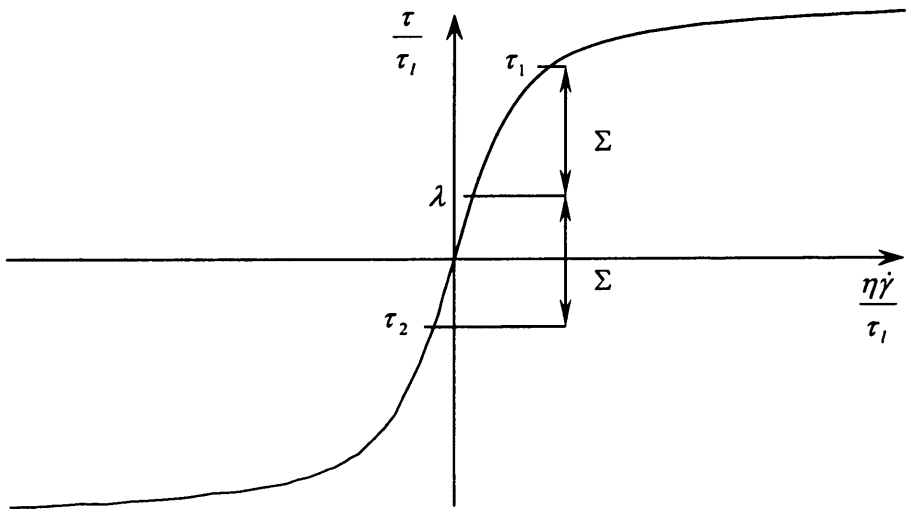

Figure A.1 – Illustration of the effect that λ and Σ have on the surface shear stresses

Figure A.1 shows positive λ and Σ, where Σ is greater than λ. The curve shown is constructed from two functions, one for positive τ and another for negative τ. When the x axis is crossed $\tau = 0$ within the film, requiring the two functions to be combined and the limits of integration changed. One limit of integration will be at a surface and the

other will be at the position in the fluid film that $\tau = 0$. This point can be determined from equation (A.1) as:

$$z = -\frac{\lambda h}{2\Sigma} \tag{A.2}$$

Case 1 – λ and Σ positive and $\lambda > \Sigma$

The surface shear stresses are positive and one function describes shear stress within the film. This covered in section 3.2 where the velocity components were determined as:

$$u = u_2 + \frac{\tau_1 h}{2\Sigma\eta}\left[\left(1-\lambda-\frac{2\Sigma}{h}z\right)\ln\left(1-\lambda-\frac{2\Sigma}{h}z\right)-(1-\lambda+\Sigma)\ln(1-\lambda+\Sigma)+\Sigma\left(\frac{2}{h}z+1\right)\right] \tag{A.3}$$

and

$$u_s = \frac{\tau_1 h}{2\Sigma\eta}\left[(1-\lambda-\Sigma)\ln(1-\lambda-\Sigma)-(1-\lambda+\Sigma)\ln(1-\lambda+\Sigma)+2\Sigma\right] \tag{A.4}$$

The corresponding S equation is:

$$S = \frac{3}{4\Sigma^3}\left[(1-2\lambda+\lambda^2-\Sigma^2)\ln\left(\frac{1-\lambda-\Sigma}{1-\lambda+\Sigma}\right)+2\Sigma(1-\lambda)\right] \tag{A.5}$$

Case 2 - λ positive, Σ negative and $\lambda > |\Sigma|$

From Figure A.1 it can be seen that making Σ negative will swap the shear stresses over. This change does not affect the function that needs to be integrated nor the limits used in the integration. It is just a matter of changing the sign of Σ in equations (A.3), (A.4) and (A.5):

$$u = u_2 - \frac{\tau_1 h}{2|\Sigma|\eta}\left[\left(1-\lambda-\frac{2|\Sigma|}{h}z\right)\ln\left(1-\lambda-\frac{2|\Sigma|}{h}z\right)-(1-\lambda+|\Sigma|)\ln(1-\lambda+|\Sigma|)+|\Sigma|\left(\frac{2}{h}z+1\right)\right]$$

$$u_s = -\frac{\tau_1 h}{2|\Sigma|\eta}\left[(1-\lambda+|\Sigma|)\ln(1-\lambda+|\Sigma|)-(1-\lambda-|\Sigma|)\ln(1-\lambda-|\Sigma|)-2|\Sigma|\right]$$

$$S = -\frac{3}{4\Sigma^3}\left[(1-2\lambda+\lambda^2-|\Sigma|^2)\ln\left(\frac{1-\lambda+|\Sigma|}{1-\lambda-|\Sigma|}\right)-2|\Sigma|(1-\lambda)\right]$$

This change does not affect the result of the calculations; within the square bracket the terms alter sign which is then cancelled out by the change in sign of Σ outside these brackets.

Case 3 - λ negative, Σ positive and $|\lambda| > \Sigma$

This means that both shear stresses are below the x axis in figure A.1. The previous function is no longer valid and the second function for shear stress must be used:

$$\dot{\gamma} = \frac{\partial u}{\partial z} = \frac{\tau_l}{\eta} \ln\left(1 + \lambda + \frac{2\Sigma}{h} z\right)$$

As the x axis isn't crossed the function presented above is used for the whole of the film. Integrating the function gives:

$$u = \frac{\tau_l h}{2\Sigma\eta}\left[\left(1 + \lambda + \frac{2\Sigma}{h} z\right)\ln\left(1 + \lambda + \frac{2\Sigma}{h} z\right) - \left(1 + \lambda + \frac{2\Sigma}{h} z\right)\right] + const$$

Applying the lower surface velocity as a boundary condition gives the velocity variation as:

$$u = u_2 + \frac{\tau_l h}{2\Sigma\eta}\left[\left(1 + \lambda + \frac{2\Sigma}{h} z\right)\ln\left(1 + \lambda + \frac{2\Sigma}{h} z\right) - (1 + \lambda - \Sigma)\ln(1 + \lambda - \Sigma) - \Sigma\left(\frac{2}{h} z + 1\right)\right]$$

Since λ is negative this can be expressed as:

$$u = u_2 + \frac{\tau_l h}{2\Sigma\eta}\left[\begin{array}{l}\left(1 - |\lambda| + \frac{2\Sigma}{h} z\right)\ln\left(1 - |\lambda| + \frac{2\Sigma}{h} z\right) \\ -(1 - |\lambda| - \Sigma)\ln(1 - |\lambda| - \Sigma) - \Sigma\left(\frac{2}{h} z + 1\right)\end{array}\right] \qquad \text{(A.6)}$$

Applying the upper surface boundary condition gives the sliding velocity as:

$$u_s = \frac{\tau_l h}{2\Sigma\eta}\left[(1 - |\lambda| + \Sigma)\ln(1 - |\lambda| + \Sigma) - (1 - |\lambda| - \Sigma)\ln(1 - |\lambda| - \Sigma) - 2\Sigma\right] \qquad \text{(A.7)}$$

This gives the same magnitude as cases 1 and 2, but has the opposite sign. This is due to the lower surface moving faster than the upper, hence reversing the mean shear stress direction.

In a similar manner S has to be re-evaluated using the different shear strain function:

$$Q_p = -\frac{\rho\tau_l}{\eta}\int_{-\frac{h}{2}}^{\frac{h}{2}} \ln\left(1+\lambda+\frac{2\Sigma}{h}z\right).dz$$

$$= -\frac{\rho\tau_l h^2}{8\Sigma^2\eta}\left\{\begin{array}{l}\left(-1-2\lambda-\lambda^2+\frac{4\Sigma^2}{h^2}z^2\right)\ln\left(1+\lambda+\frac{2\Sigma}{h}z\right) \\ -\frac{1}{2}\left[-3-6\lambda-3\lambda^2-\frac{4\Sigma(1+\lambda)}{h}z+\frac{4\Sigma^2}{h^2}z^2\right]\end{array}\right\}_{-\frac{h}{2}}^{\frac{h}{2}} \qquad \text{(A.8)}$$

$$Q_p = -\frac{\rho\tau_l h^2}{8\Sigma^2\eta}\left[\left(1+2\lambda+\lambda^2-\Sigma^2\right)\ln\left(\frac{1+\lambda-\Sigma}{1+\lambda+\Sigma}\right)+2\Sigma(1+\lambda)\right]$$

This gives S as:

$$S = \frac{3}{4\Sigma^3}\left[\left(1+2\lambda+\lambda^2-\Sigma^2\right)\ln\left(\frac{1+\lambda-\Sigma}{1+\lambda+\Sigma}\right)+2\Sigma(1+\lambda)\right]$$

Since λ is negative this becomes:

$$S = \frac{3}{4\Sigma^3}\left[\left(1-2|\lambda|+\lambda^2-\Sigma^2\right)\ln\left(\frac{1-|\lambda|-\Sigma}{1-|\lambda|+\Sigma}\right)+2\Sigma(1-|\lambda|)\right] \qquad \text{(A.9)}$$

This is the same as the terms utilised in the first two cases.

Case 4 - λ and Σ negative and $|\lambda| > |\Sigma|$

This is similar to that shown in case 2, where the shear stresses swap over. Equations (A.6), (A.7) and (A.8) are still valid and it is just a matter of swapping the sign of Σ:

$$u = u_2 - \frac{\tau_l h}{2|\Sigma|\eta}\left[\begin{array}{l}\left(1-|\lambda|-\frac{2|\Sigma|}{h}z\right)\ln\left(1-|\lambda|-\frac{2|\Sigma|}{h}z\right) \\ -\left(1-|\lambda|+|\Sigma|\right)\ln\left(1-|\lambda|+|\Sigma|\right)-|\Sigma|\left(\frac{2}{h}z+1\right)\end{array}\right]$$

$$u_s = -\frac{\tau_l h}{2|\Sigma|\eta}\left[\left(1-|\lambda|-|\Sigma|\right)\ln\left(1-|\lambda|-|\Sigma|\right)-\left(1-|\lambda|+|\Sigma|\right)\ln\left(1-|\lambda|+|\Sigma|\right)+2|\Sigma|\right]$$

$$S = -\frac{3}{4|\Sigma|^3}\left[\left(1-2|\lambda|+\lambda^2-|\Sigma|^2\right)\ln\left(\frac{1-|\lambda|+|\Sigma|}{1-|\lambda|-|\Sigma|}\right)-2|\Sigma|(1-|\lambda|)\right]$$

The signs inside and outside the square brackets for u_s and S change and cancel one another out; hence the equations are the same as case 3.

Case 5 - λ and Σ positive and $\lambda < \Sigma$

This is the case shown in figure A.1, where one shear stress is above the x axis and one below. Two different functions have to be used to construct the variation of the velocity across the film. The positive shear stress occurs at the upper surface and the negative at the lower, applying the lower boundary condition and considering negative shear stresses an equation the same as equation (A.6) results:

$$u = u_2 + \frac{\tau_1 h}{2\Sigma\eta}\left[\left(1+\lambda+\frac{2\Sigma}{h}z\right)\ln\left(1+\lambda+\frac{2\Sigma}{h}z\right)-(1+\lambda-\Sigma)\ln(1+\lambda-\Sigma)-\Sigma\left(\frac{2}{h}z+1\right)\right]$$

This is only applicable to the region where the shear stress is negative, given by $z \le -\frac{\lambda h}{2\Sigma}$. This is used as a subscript on the velocity for clarity giving:

$$u_{z\le-\frac{\lambda h}{2\Sigma}} = u_2 + \frac{\tau_1 h}{2\Sigma\eta}\left[\left(1+\lambda+\frac{2\Sigma}{h}z\right)\ln\left(1+\lambda+\frac{2\Sigma}{h}z\right)-(1+\lambda-\Sigma)\ln(1+\lambda-\Sigma)-\Sigma\left(\frac{2}{h}z+1\right)\right]$$

If the shear stress is positive then the strain rate can be integrated, as in section 3.2, to give:

$$u = \frac{\tau_1 h}{2\Sigma\eta}\left[\left(1-\lambda-\frac{2\Sigma}{h}z\right)\ln\left(1-\lambda-\frac{2\Sigma}{h}z\right)-\left(1-\lambda-\frac{2\Sigma}{h}z\right)\right]+const$$

Applying $z = -\frac{\lambda h}{2\Sigma}$ to both of the above equations, *const* can be determined and substituted back in to give:

$$u_{z>-\frac{\lambda h}{2\Sigma}} = u_2 + \frac{\tau_1 h}{2\Sigma\eta}\begin{bmatrix}\left(1-\lambda-\frac{2\Sigma}{h}z\right)\ln\left(1-\lambda-\frac{2\Sigma}{h}z\right)\\ -(1+\lambda-\Sigma)\ln(1+\lambda-\Sigma)+\Sigma\left(\frac{2}{h}z-1\right)+2\lambda\end{bmatrix}$$

Applying the upper surface boundary condition this gives the sliding velocity as:

$$u_s = \frac{\tau_1 h}{2\Sigma\eta}[(1-\lambda-\Sigma)\ln(1-\lambda-\Sigma)-(1+\lambda-\Sigma)\ln(1+\lambda-\Sigma)+2\lambda] \qquad (A.10)$$

The two functions for shear strain must be used and appropriate limits applied. Using this the Poiseuille flow can be determined from equations (3.12) and (A.8) as:

$$Q_p = -\frac{\rho\tau_l h^2}{8\Sigma^2\eta}\left(\begin{array}{l}\left\{\begin{array}{l}\left(1-2\lambda+\lambda^2-\frac{4\Sigma^2}{h^2}z^2\right)\ln\left(1-\lambda-\frac{2\Sigma}{h}z\right)\\ -\frac{1}{2}\left[3-6\lambda+3\lambda^2-\frac{4\Sigma(1-\lambda)}{h}z-\frac{4\Sigma^2}{h^2}z^2\right]\end{array}\right\}_{-\frac{\lambda h}{2\Sigma}}^{\frac{h}{2}}\\ +\left\{\begin{array}{l}\left(-1-2\lambda-\lambda^2+\frac{4\Sigma^2}{h^2}z^2\right)\ln\left(1+\lambda+\frac{2\Sigma}{h}z\right)\\ -\frac{1}{2}\left[-3-6\lambda-3\lambda^2-\frac{4\Sigma(1+\lambda)}{h}z+\frac{4\Sigma^2}{h^2}z^2\right]\end{array}\right\}_{-\frac{h}{2}}^{-\frac{\lambda h}{2\Sigma}}\end{array}\right)$$

$$Q_p = -\frac{\rho\tau_l h^2}{8\Sigma^2\eta}\left[\begin{array}{c}\left(1-2\lambda+\lambda^2-\Sigma^2\right)\ln(1-\lambda-\Sigma)\\ +\left(1+2\lambda+\lambda^2-\Sigma^2\right)\ln(1+\lambda-\Sigma)+\Sigma^2+2\Sigma-3\lambda^2\end{array}\right]$$

Re-arranging to get the standard form of the flow factor, S now becomes:

$$S = \frac{3}{4\Sigma^3}\left[\begin{array}{c}\left(1-2\lambda+\lambda^2-\Sigma^2\right)\ln(1-\lambda-\Sigma)\\ +\left(1+2\lambda+\lambda^2-\Sigma^2\right)\ln(1+\lambda-\Sigma)+\Sigma^2+2\Sigma-3\lambda^2\end{array}\right] \qquad \text{(A.11)}$$

Case 6 - λ negative, Σ positive and $|\lambda| < \Sigma$

The same equations can be used; in fact the limits of integration remain the same since the shear stresses on the upper and lower surfaces are still positive and negative respectively even though the mean shear stress is negative. Thus only the sign of λ needs to be changed, the equations developed in Case 5 become:

$$u_{z\le\frac{|\lambda|h}{2\Sigma}} = u_2 + \frac{\tau_l h}{2\Sigma\eta}\left[\left(1-|\lambda|+\frac{2\Sigma}{h}z\right)\ln\left(1-|\lambda|+\frac{2\Sigma}{h}z\right)-(1-|\lambda|-\Sigma)\ln(1-|\lambda|-\Sigma)-\Sigma\left(\frac{2}{h}z+1\right)\right]$$

$$u_{z>\frac{|\lambda|h}{2\Sigma}} = u_2 + \frac{\tau_l h}{2\Sigma\eta}\left[\begin{array}{c}\left(1+|\lambda|-\frac{2\Sigma}{h}z\right)\ln\left(1+|\lambda|-\frac{2\Sigma}{h}z\right)\\ -(1-|\lambda|-\Sigma)\ln(1-|\lambda|-\Sigma)+\Sigma\left(\frac{2}{h}z-1\right)-2|\lambda|\end{array}\right]$$

$$u_s = \frac{\tau_l h}{2\Sigma\eta}\left[(1+|\lambda|-\Sigma)\ln(1+|\lambda|-\Sigma)-(1-|\lambda|-\Sigma)\ln(1-|\lambda|-\Sigma)-2|\lambda|\right]$$

$$S = \frac{3}{4\Sigma^3}\left[\begin{array}{c}(1+2|\lambda|+\lambda^2-\Sigma^2)\ln(1+|\lambda|-\Sigma) \\ +(1-2|\lambda|+\lambda^2-\Sigma^2)\ln(1-|\lambda|-\Sigma)+\Sigma^2+2\Sigma-3\lambda^2\end{array}\right]$$

It can be seen that the sliding velocity term has the same magnitude but has the opposite sign to equation (A.10). S still gives the same value as equation (A.11) regardless of the magnitude of λ.

Case 7 - λ positive, Σ negative and $\lambda < |\Sigma|$

The shear stresses on the upper and lower surface swap over from the previous two cases. The shear stress at the upper surface is negative and the shear stress at the lower surface is positive. Applying the lower surface boundary condition to the velocity equation developed for positive shear stress, gives the same as equation (A.3):

$$u_{z\le-\frac{\lambda h}{2\Sigma}} = u_2 + \frac{\tau_1 h}{2\Sigma\eta}\left[\left(1-\lambda-\frac{2\Sigma}{h}z\right)\ln\left(1-\lambda-\frac{2\Sigma}{h}z\right)-(1-\lambda+\Sigma)\ln(1-\lambda+\Sigma)+\Sigma\left(\frac{2}{h}z+1\right)\right]$$

Re-arranging this so that the effect of Σ being negative gives:

$$u_{z\le\frac{\lambda h}{2|\Sigma|}} = u_2 - \frac{\tau_1 h}{2|\Sigma|\eta}\left[\left(1-\lambda+\frac{2|\Sigma|}{h}z\right)\ln\left(1-\lambda+\frac{2|\Sigma|}{h}z\right)-(1-\lambda-|\Sigma|)\ln(1-\lambda-|\Sigma|)-|\Sigma|\left(\frac{2}{h}z+1\right)\right]$$

In a similar manner to that presented in Case 5, the velocity profile for the remainder of the film is given by:

$$u_{z>\frac{\lambda h}{2|\Sigma|}} = u_2 - \frac{\tau_1 h}{2|\Sigma|\eta}\left[\begin{array}{c}\left(1+\lambda-\frac{2|\Sigma|}{h}z\right)\ln\left(1+\lambda-\frac{2|\Sigma|}{h}z\right) \\ -(1-\lambda-|\Sigma|)\ln(1-\lambda-|\Sigma|)+|\Sigma|\left(\frac{2}{h}z-1\right)-2\lambda\end{array}\right]$$

Applying the boundary condition at the upper surface the sliding velocity is obtained as:

$$u_s = -\frac{\tau_1 h}{2|\Sigma|\eta}\left[(1+\lambda-|\Sigma|)\ln(1+\lambda-|\Sigma|)-(1-\lambda-|\Sigma|)\ln(1-\lambda-|\Sigma|)-2\lambda\right]$$

Re-arranging gives:

$$u_s = \frac{\tau_1 h}{2|\Sigma|\eta}\left[(1-\lambda-|\Sigma|)\ln(1-\lambda-|\Sigma|)-(1+\lambda-|\Sigma|)\ln(1+\lambda-|\Sigma|)+2\lambda\right] \qquad \text{(A.12)}$$

A similar approach to that used in Case 5 can be used to determine S except that the limit of integration swap over:

$$Q_p = -\frac{\rho\tau_l h^2}{8\Sigma^2\eta}\left(\begin{array}{l}\left\{\begin{array}{l}\left(-1-2\lambda-\lambda^2+\frac{4\Sigma^2}{h^2}z^2\right)\ln\left(1+\lambda+\frac{2\Sigma}{h}z\right)\\ -\frac{1}{2}\left[-3-6\lambda-3\lambda^2-\frac{4\Sigma(1+\lambda)}{h}z+\frac{4\Sigma^2}{h^2}z^2\right]\end{array}\right\}_{-\frac{\lambda h}{2\Sigma}}^{\frac{h}{2}}\\ +\left\{\begin{array}{l}\left(1-2\lambda+\lambda^2-\frac{4\Sigma^2}{h^2}z^2\right)\ln\left(1-\lambda-\frac{2\Sigma}{h}z\right)\\ -\frac{1}{2}\left[3-6\lambda+3\lambda^2-\frac{4\Sigma(1-\lambda)}{h}z-\frac{4\Sigma^2}{h^2}z^2\right]\end{array}\right\}_{-\frac{h}{2}}^{-\frac{\lambda h}{2\Sigma}}\end{array}\right)$$

$$Q_p = -\frac{\rho\tau_l h^2}{8\Sigma^2\eta}\left[\begin{array}{c}\left(-1-2\lambda-\lambda^2+\Sigma^2\right)\ln(1+\lambda+\Sigma)\\ -\left(1-2\lambda+\lambda^2-\Sigma^2\right)\ln(1-\lambda+\Sigma)-\Sigma^2+2\Sigma+3\lambda^2\end{array}\right]$$

Re-arranging, S is obtained as:

$$S = \frac{3}{4\Sigma^3}\left[\begin{array}{c}\left(-1-2\lambda-\lambda^2+\Sigma^2\right)\ln(1+\lambda+\Sigma)\\ -\left(1-2\lambda+\lambda^2-\Sigma^2\right)\ln(1-\lambda+\Sigma)-\Sigma^2+2\Sigma+3\lambda^2\end{array}\right]$$

Since Σ is negative this can be re-written as:

$$S = \frac{3}{4|\Sigma|^3}\left[\begin{array}{c}\left(1-2\lambda+\lambda^2-\Sigma^2\right)\ln(1-\lambda-|\Sigma|)\\ +\left(1+2\lambda+\lambda^2-\Sigma^2\right)\ln(1+\lambda-|\Sigma|)+\Sigma^2+2|\Sigma|-3\lambda^2\end{array}\right] \qquad \text{(A.13)}$$

This is the same as equation (A.11).

Case 8 - λ and Σ negative and $|\lambda| < |\Sigma|$

In a similar manner to case 6 the axis is crossed and still have to use both functions for $\dot{\gamma}$. The mean shear stress changes sign but the surface shear stresses remain the same as case 7. The equations derived above are still applicable since the limits of integration are the same. Accounting for the change in λ the equations become:

$$u_{z\le-\frac{|\lambda|h}{2|\Sigma|}} = u_2 - \frac{\tau_l h}{2|\Sigma|\eta}\left[\left(1+|\lambda|+\frac{2|\Sigma|}{h}z\right)\ln\left(1+|\lambda|+\frac{2|\Sigma|}{h}z\right)-\left(1+|\lambda|-|\Sigma|\right)\ln\left(1+|\lambda|-|\Sigma|\right)-|\Sigma|\left(\frac{2}{h}z+1\right)\right]$$

$$u_{z>-\frac{|\lambda|h}{2|\Sigma|}} = u_2 - \frac{\tau_l h}{2|\Sigma|\eta}\left[\begin{array}{c}\left(1-|\lambda|-\frac{2|\Sigma|}{h}z\right)\ln\left(1-|\lambda|-\frac{2|\Sigma|}{h}z\right)\\ -(1+|\lambda|-|\Sigma|)\ln(1+|\lambda|-|\Sigma|)+|\Sigma|\left(\frac{2}{h}z-1\right)+2|\lambda|\end{array}\right]$$

$$u_s = \frac{\tau_l h}{2|\Sigma|\eta}\left[(1+|\lambda|-|\Sigma|)\ln(1+|\lambda|-|\Sigma|)-(1-|\lambda|-|\Sigma|)\ln(1-|\lambda|-|\Sigma|)-2|\lambda|\right]$$

$$S = \frac{3}{4|\Sigma|^3}\left[\begin{array}{c}(1+2|\lambda|+\lambda^2-\Sigma^2)\ln(1+|\lambda|-|\Sigma|)\\ +(1-2|\lambda|+\lambda^2-\Sigma^2)\ln(1-|\lambda|-|\Sigma|)+\Sigma^2+2|\Sigma|-3\lambda^2\end{array}\right]$$

The sliding term has the same magnitude as equations (A.10) and (A.12) but opposite sign. S is the same as equation (A.13), hence equals equation (A.11).

Conclusion

Taking into account all the different equations produced, four general equations can be determined that account for all of the cases considered. The velocity variation across the film thickness can be accounted for using:

$$u = u_2 + \frac{\tau_l h}{2\Sigma\eta}\left[\begin{array}{c}\left(1-\left|\lambda+\frac{2\Sigma}{h}z\right|\right)\ln\left(1-\left|\lambda+\frac{2\Sigma}{h}z\right|\right)+\left|\lambda+\frac{2\Sigma}{h}z\right|-|\lambda-\Sigma|\\ -(1-|\lambda-\Sigma|)\ln(1-|\lambda-\Sigma|)\end{array}\right]$$

The sliding velocity can be determined from:

$$u_s = \frac{\tau_l h}{2|\Sigma|\eta}\left[(1-|\lambda+\Sigma|)\ln(1-|\lambda+\Sigma|)-(1-|\lambda-\Sigma|)\ln(1-|\lambda-\Sigma|)+|\lambda+\Sigma|-|\lambda-\Sigma|\right]$$

This term allows the sliding velocity to become negative when λ is negative. This equation holds true for all combinations of λ and Σ.

The first general equation for S is derived from cases 1 – 4 and only applicable provided $|\lambda| > |\Sigma|$, i.e. τ doesn't equal zero in the film. Thus yields the following expression:

$$S = \frac{3}{4|\Sigma|^3}\left[\left(1-2|\lambda|+\lambda^2-\Sigma^2\right)\ln\left(\frac{1-|\lambda|-|\Sigma|}{1-|\lambda|+|\Sigma|}\right)+2|\Sigma|\left(1-|\lambda|\right)\right]$$

The second equation for S is taken from the other four cases and is valid provided $|\lambda| \le |\Sigma|$ and can be determined as:

$$S = \frac{3}{4|\Sigma|^3}\left[\begin{array}{c}\left(1-2|\lambda|+\lambda^2-\Sigma^2\right)\ln\left(1-|\lambda|-|\Sigma|\right) \\ +\left(1+2|\lambda|+\lambda^2-\Sigma^2\right)\ln\left(1+|\lambda|-|\Sigma|\right)+\Sigma^2+2|\Sigma|-3\lambda^2\end{array}\right]$$

Appendix B

This appendix presents four solutions from the full gear meshing cycle for the three rough cases considered in Chapter 8. The three cases comprised of two sets taken for mesh A, termed set A1 and A2, and a set for mesh C, termed set C.

In addition to the 'snap shots' taken from the full meshing cycle, a CD is enclosed that contains animations of the full meshing cycle for all of the thermal cases considered. Thus, it comprises of the normal operation and reversed (speed increase) operating arrangements for the three gear meshes of mesh A, B and C. It should be noted that s is presented within the animations, which is measured in the direction of rotation of the gear. This means that the value is taken as positive towards the pinion tip for the normal (speed reduction) operation, but is negative towards the pinion tip for the reversed (speed increase) arrangement.

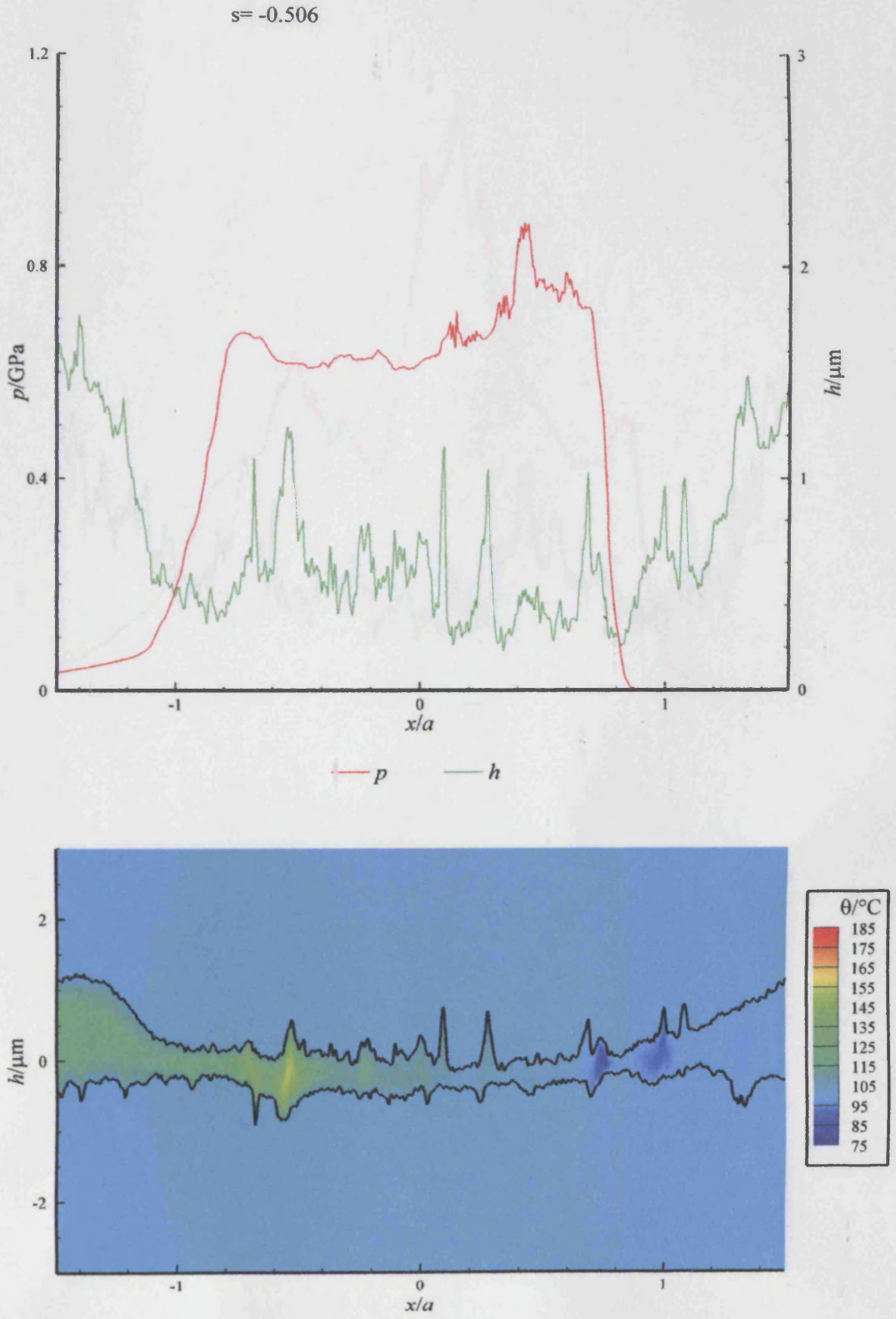

Figure B.1 – Pressure, film thickness and temperature variation for set A1

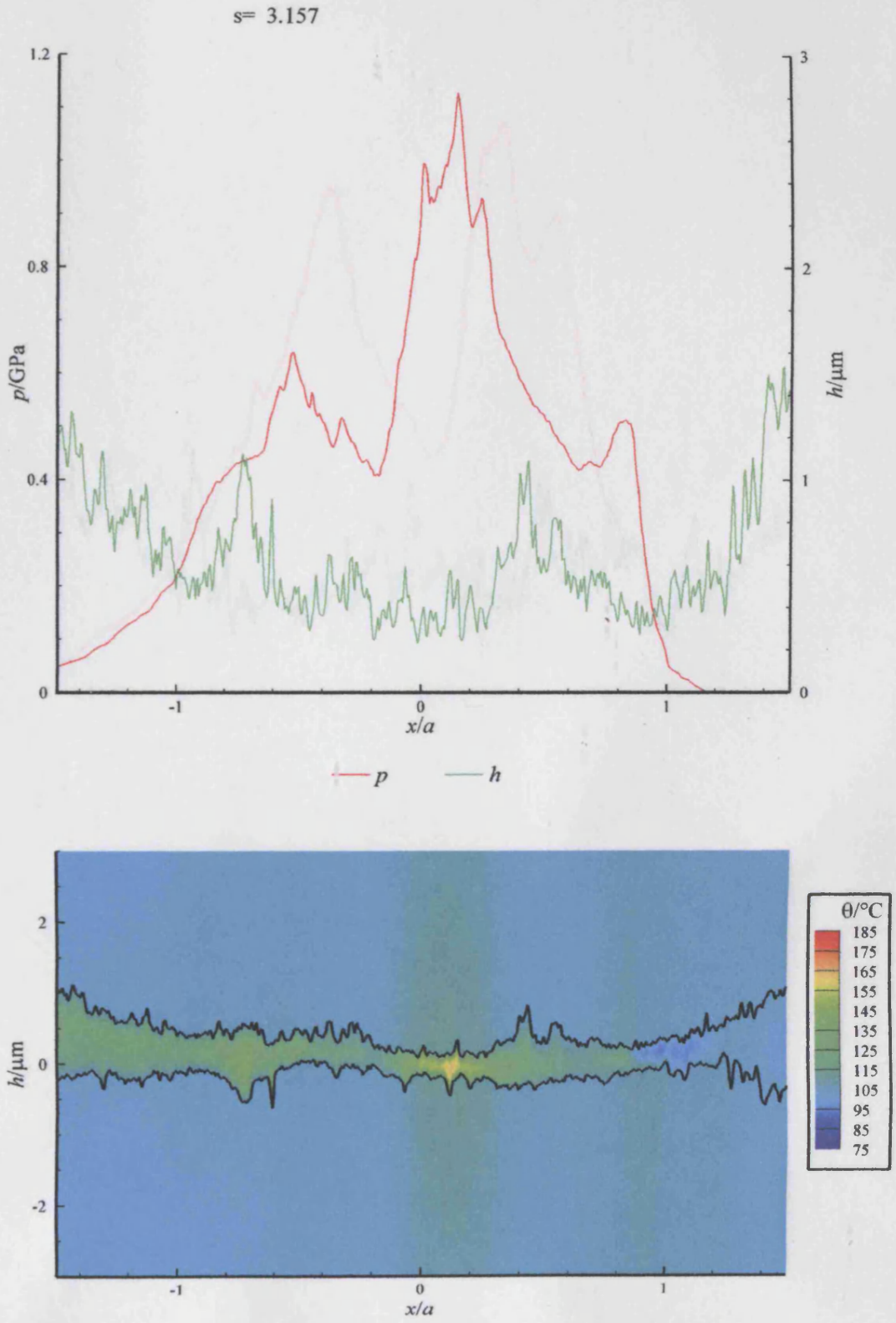

Figure B.2 – Pressure, film thickness and temperature variation for set A1

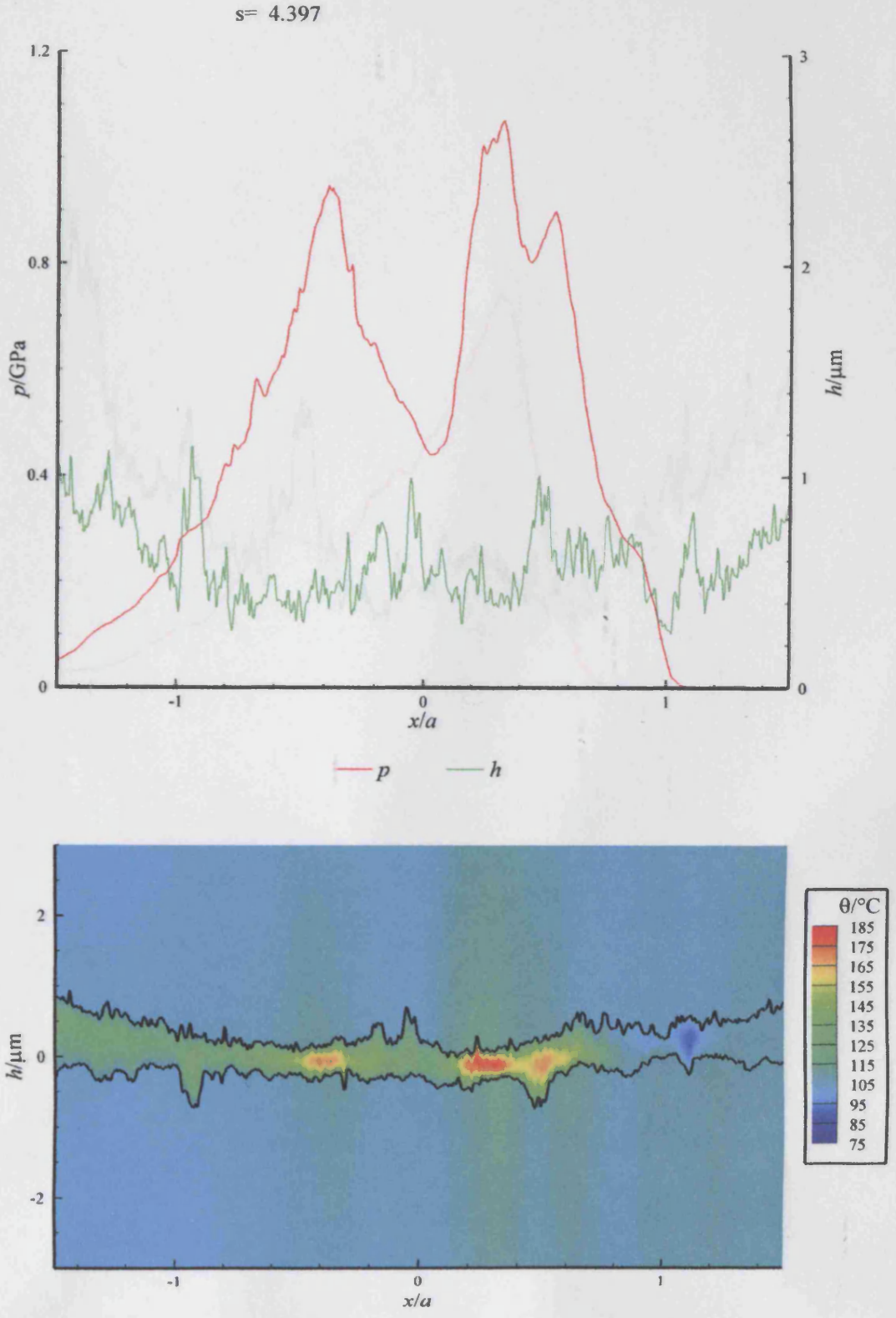

Figure B.3 – Pressure, film thickness and temperature variation for set A1

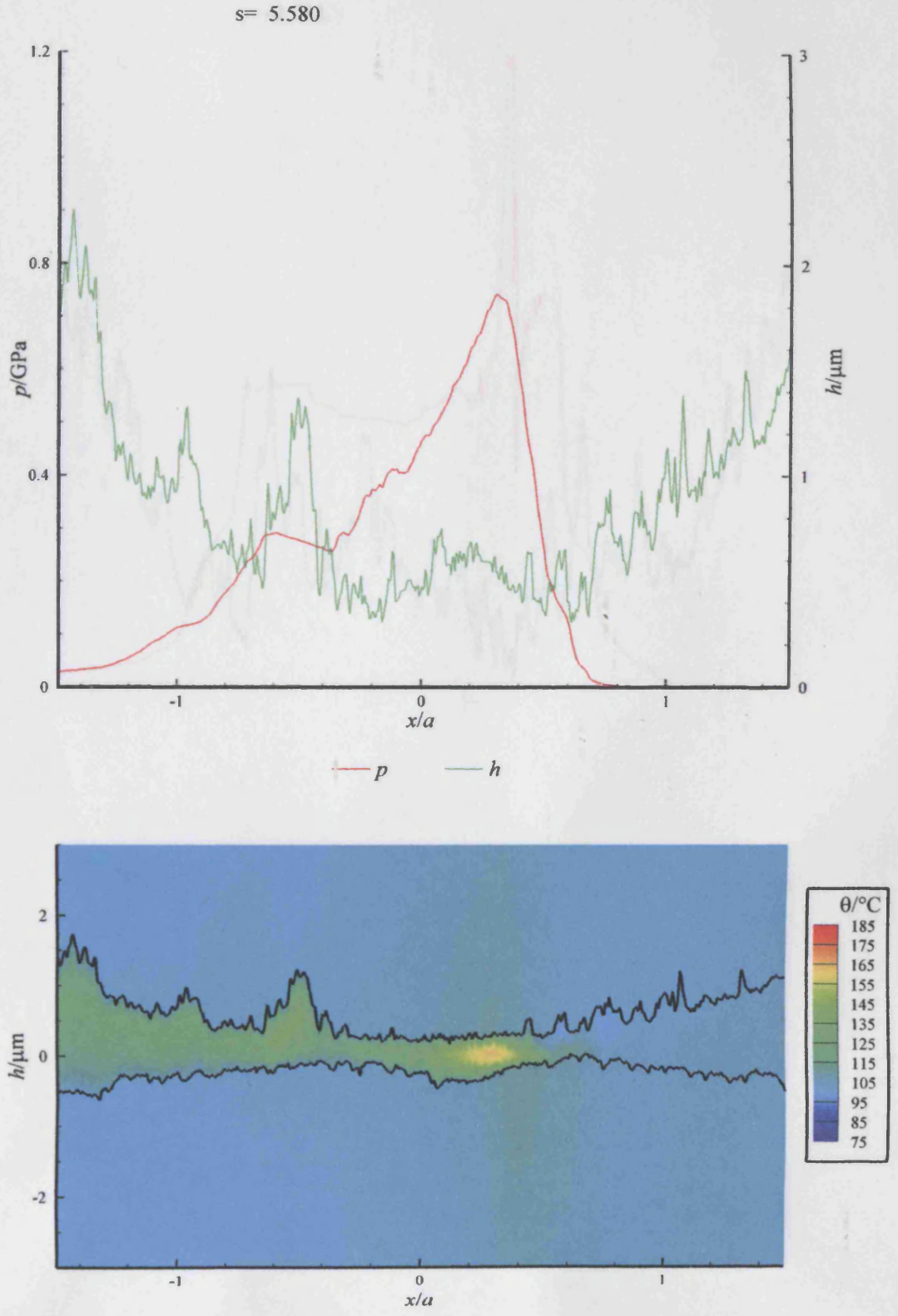

Figure B.4 – Pressure, film thickness and temperature variation for set A1

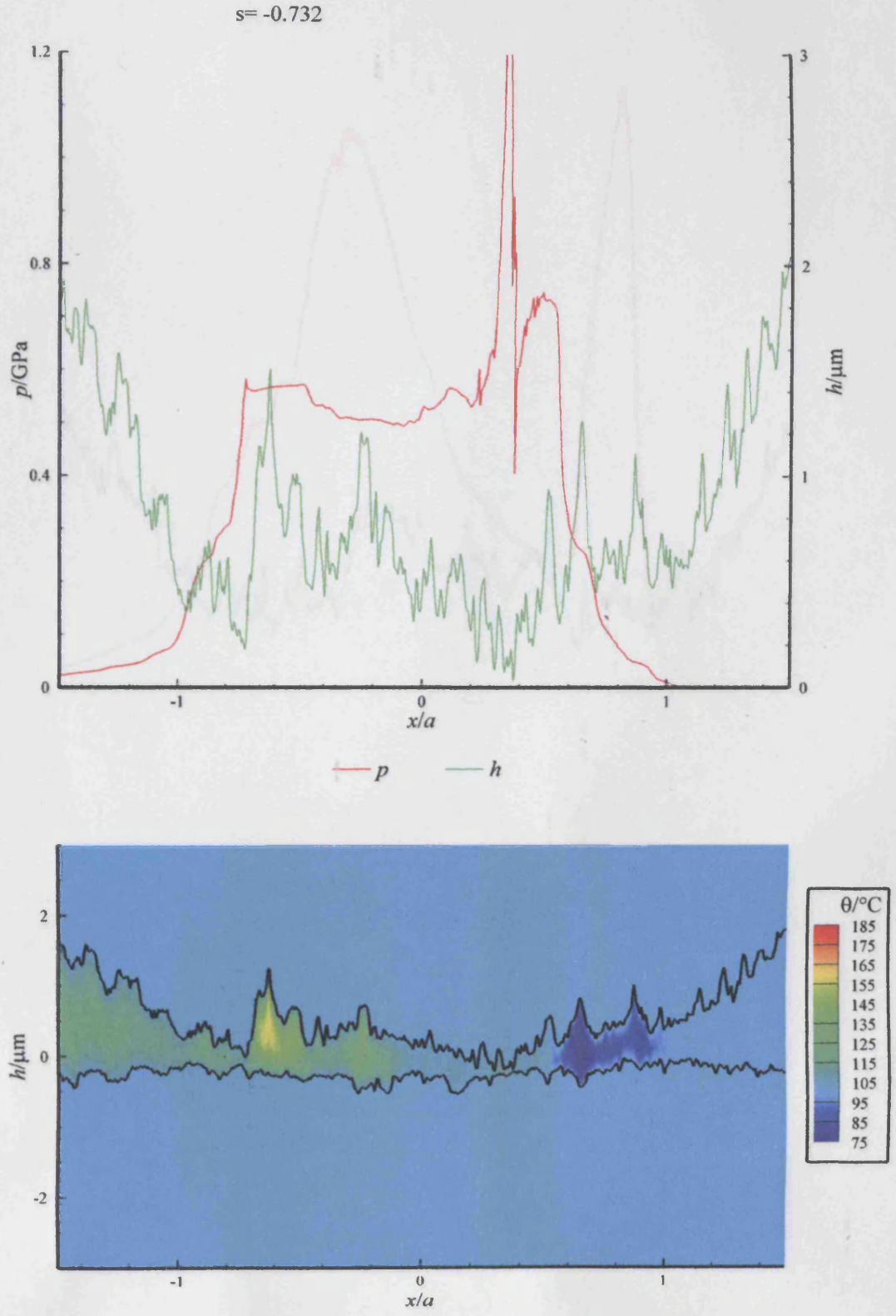

Figure B.5 – Pressure, film thickness and temperature variation for set A2

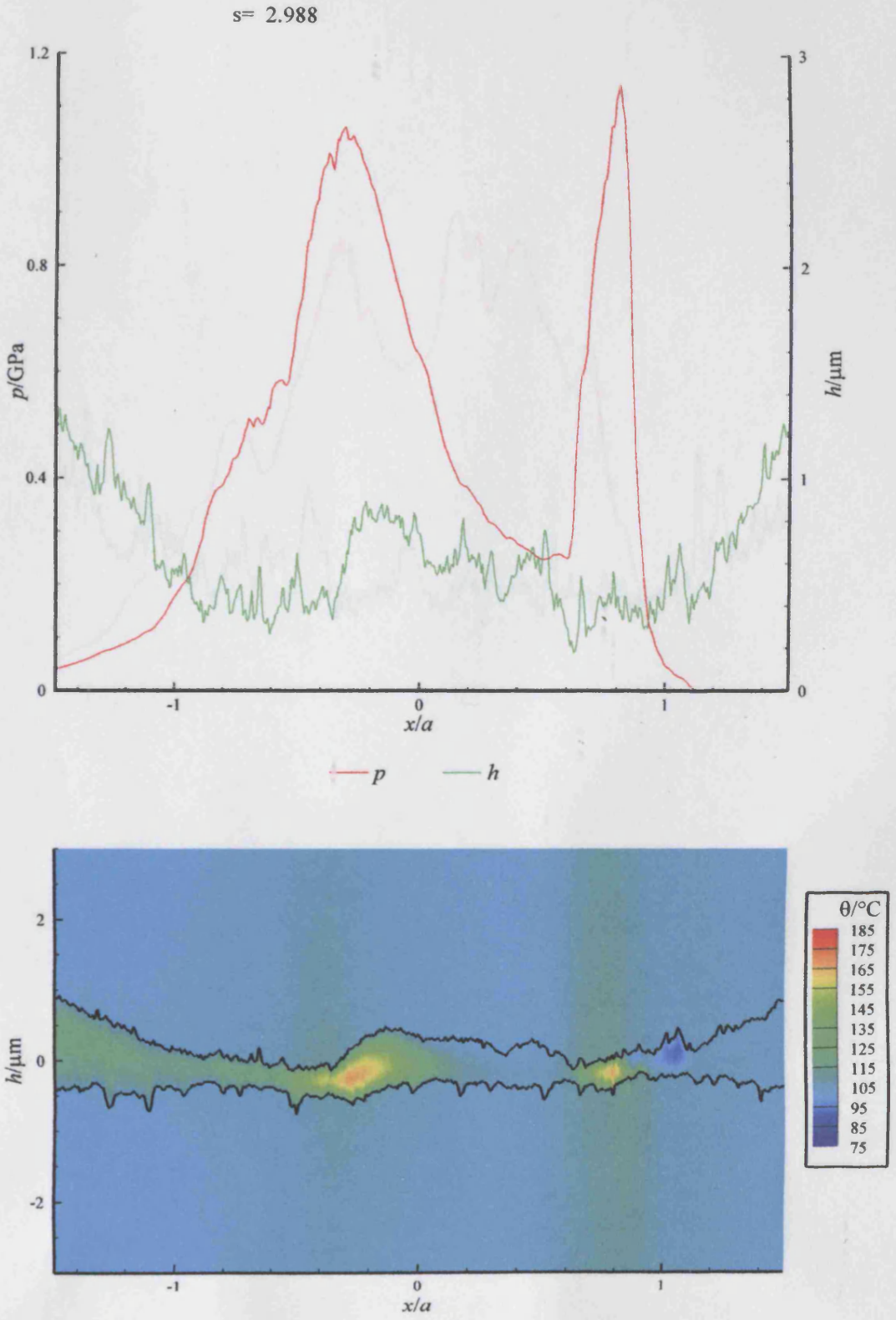

Figure B.6 – Pressure, film thickness and temperature variation for set A2

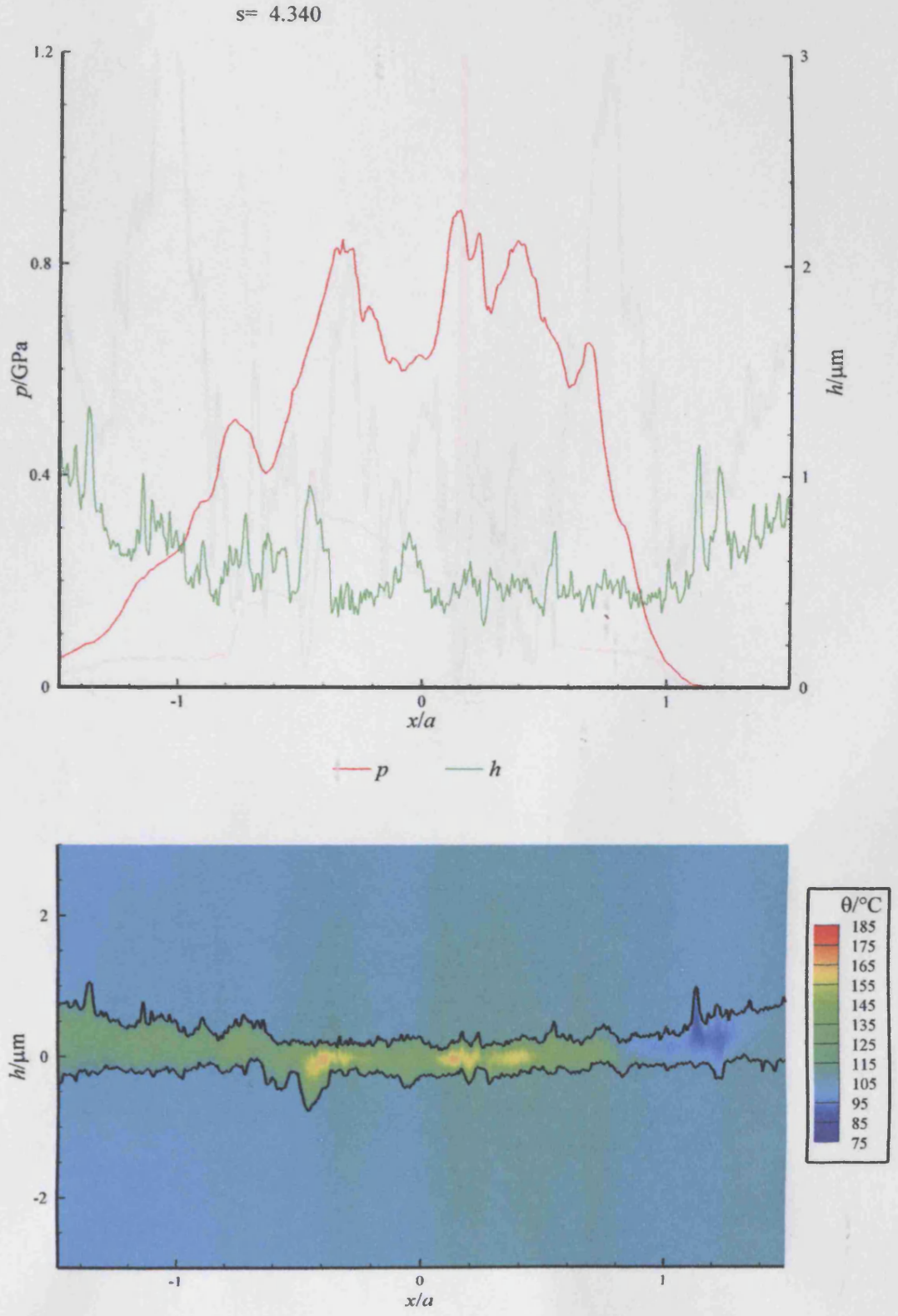

Figure B.7 – Pressure, film thickness and temperature variation for set A2

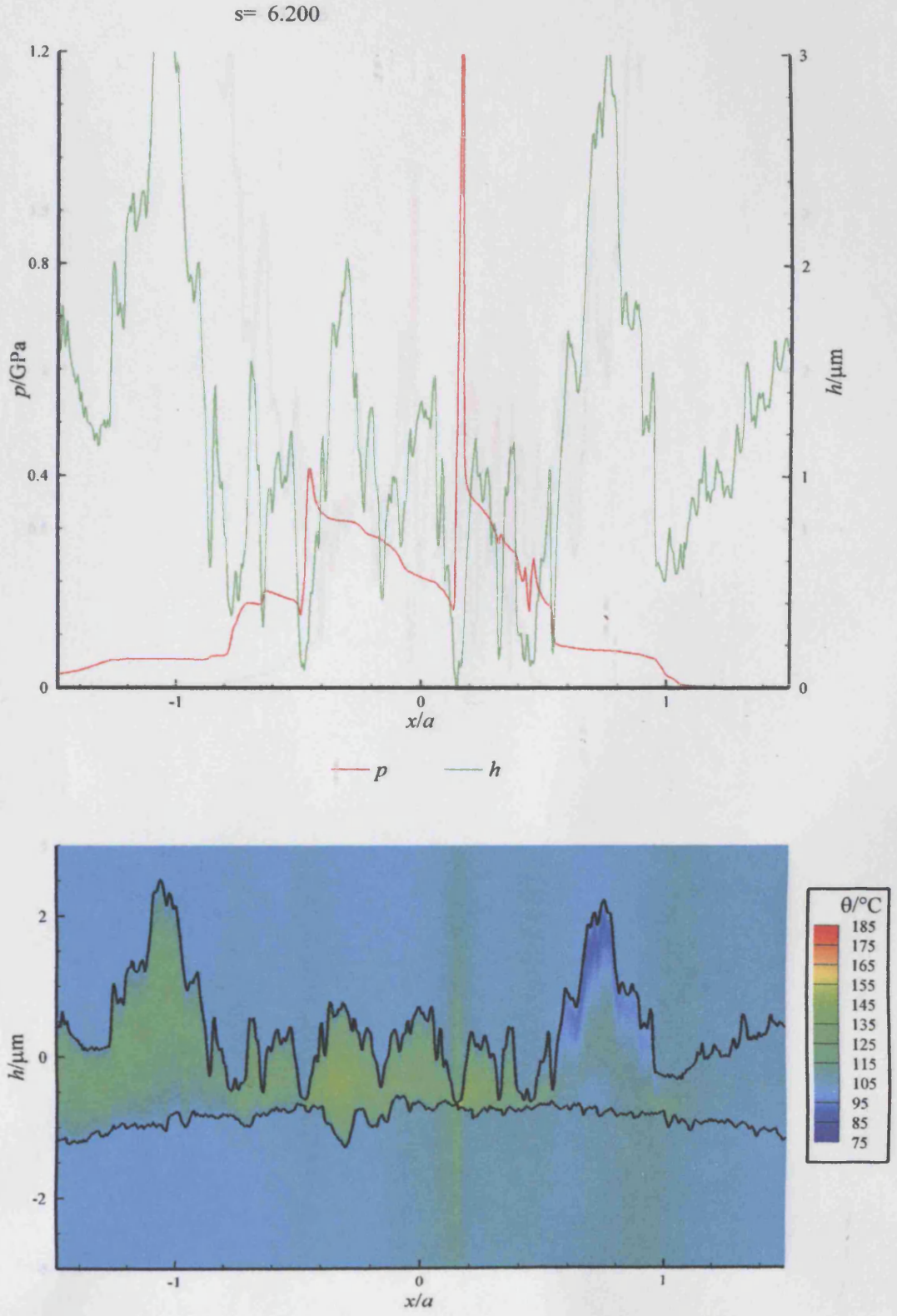

Figure B.8 – Pressure, film thickness and temperature variation for set A2

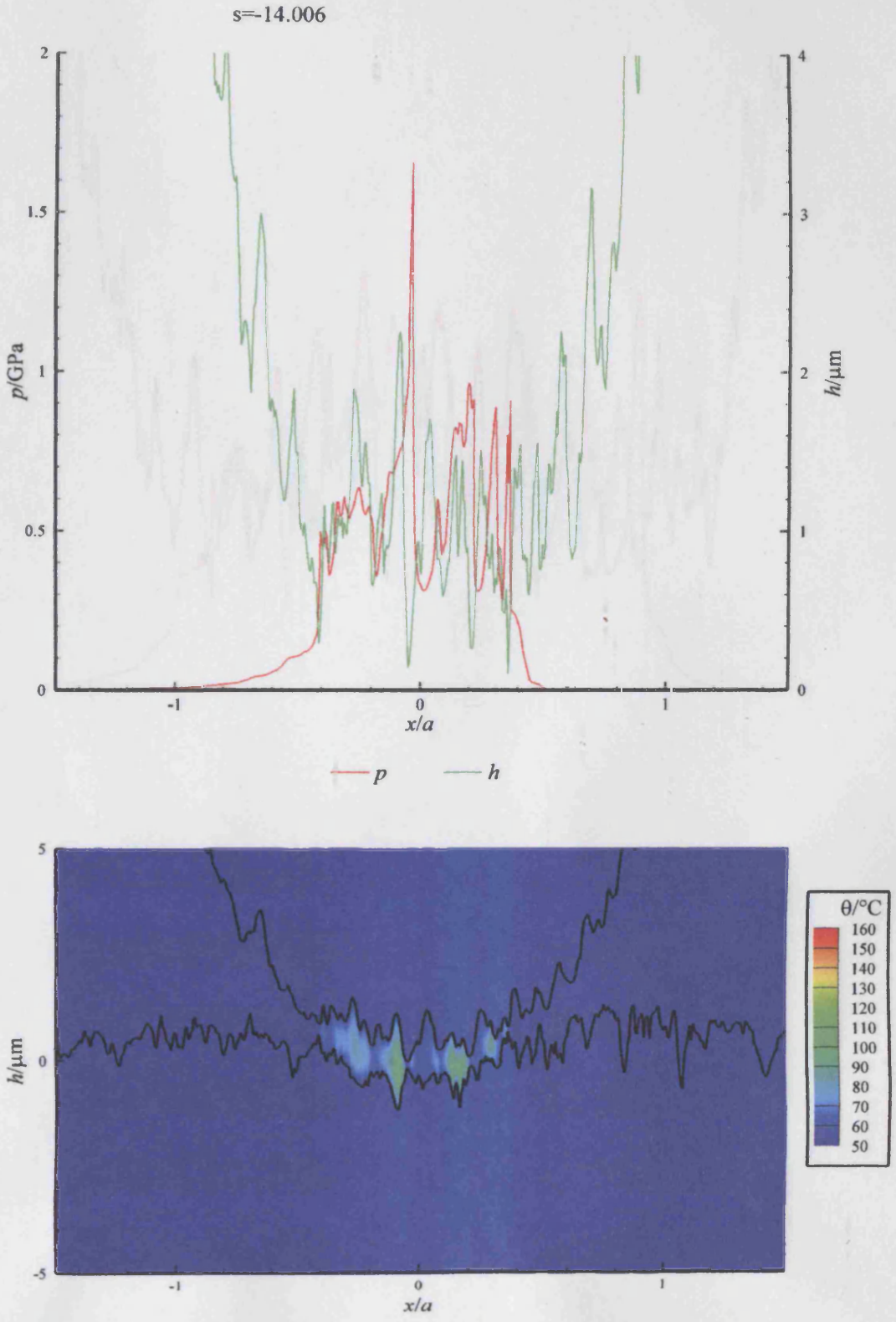

Figure B.9 – Pressure, film thickness and temperature variation for set C

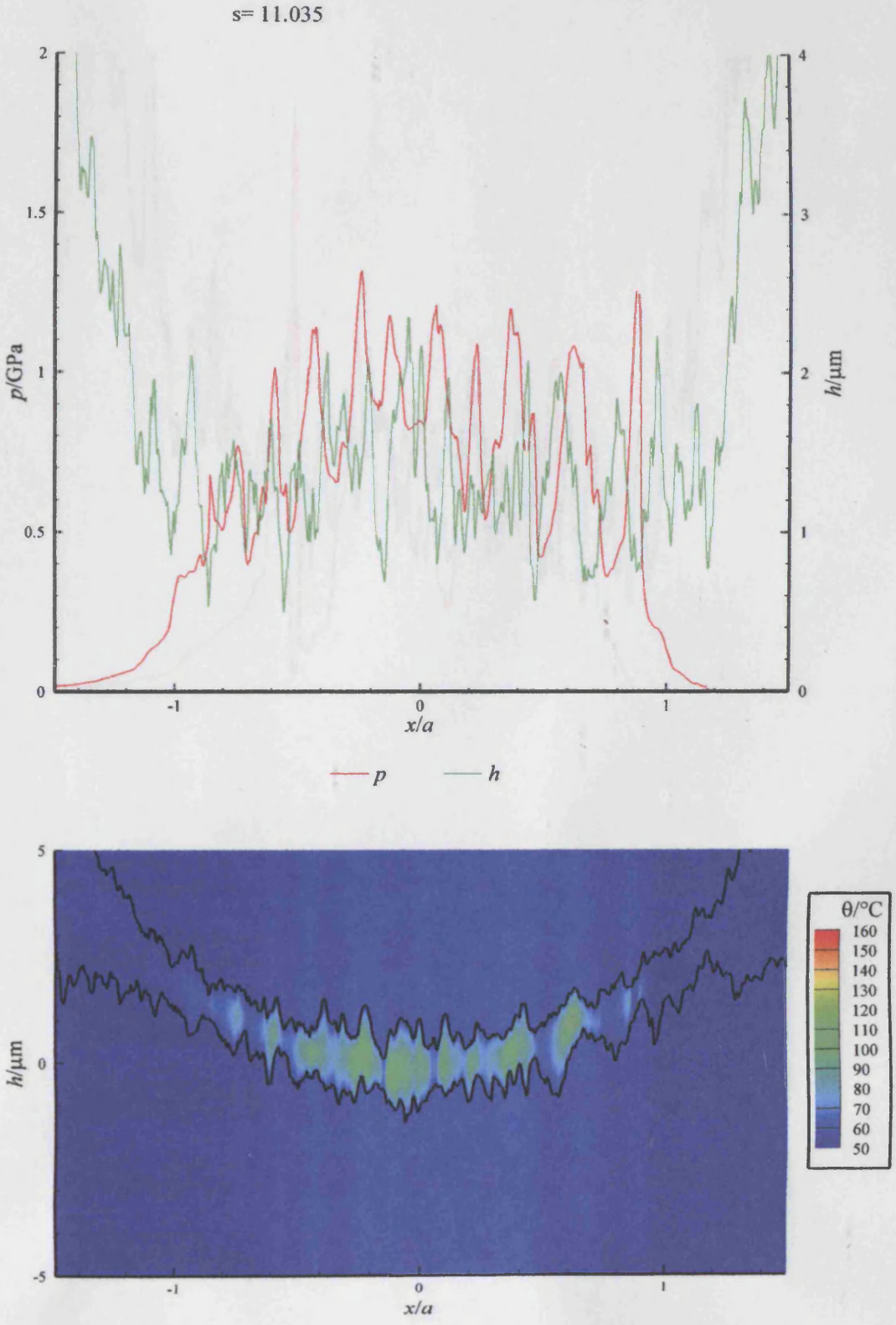

Figure B.10 – Pressure, film thickness and temperature variation for set C

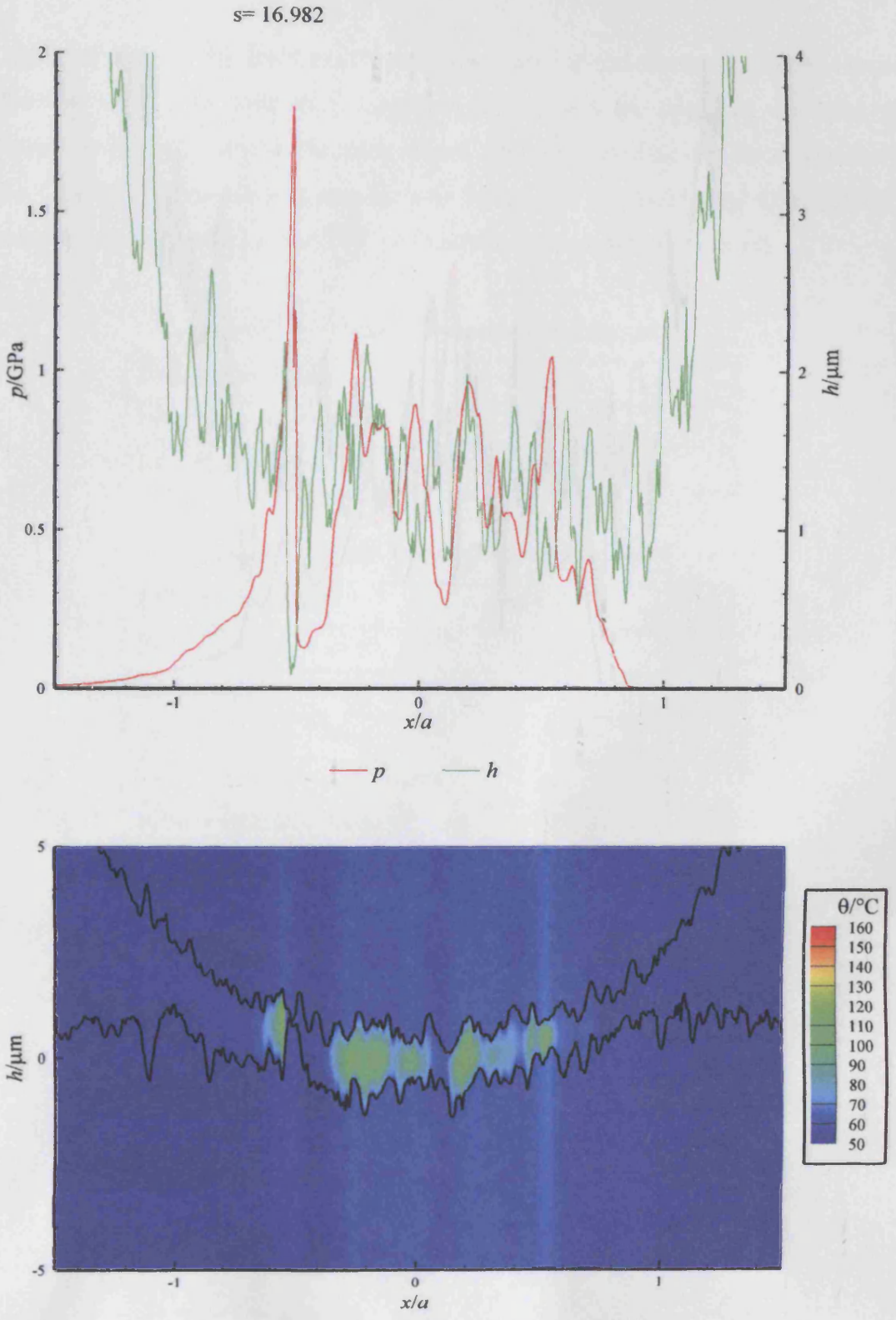

Figure B.11 – Pressure, film thickness and temperature variation for set C

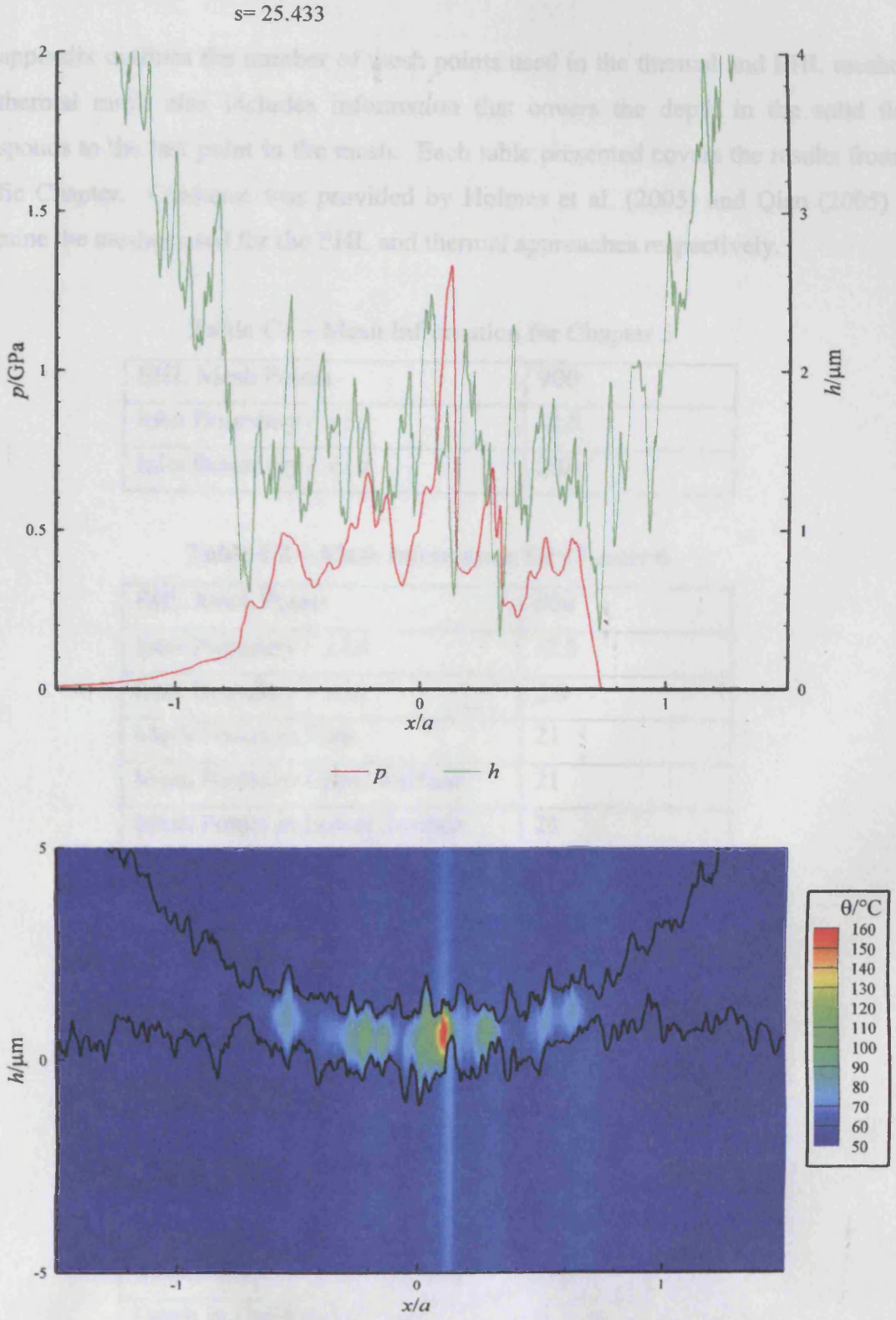

Figure B.12 – Pressure, film thickness and temperature variation for set C

Appendix C

This appendix outlines the number of mesh points used in the thermal and EHL meshes. The thermal mesh also includes information that covers the depth in the solid that corresponds to the last point in the mesh. Each table presented covers the results from a specific Chapter. Guidance was provided by Holmes et al. (2005) and Qiao (2005) to determine the meshes used for the EHL and thermal approaches respectively.

Table C1 – Mesh Information for Chapter 5

EHL Mesh Points	900
Inlet Boundary / x/a	-2.5
Inlet Boundary / x/a	2.0

Table C2 – Mesh Information for Chapter 6

EHL Mesh Points	900
Inlet Boundary / x/a	-2.5
Inlet Boundary / x/a	2.0
Mesh Points in Film	21
Mesh Points in Upper Surface	21
Mesh Points in Lower Surface	21
Depth in Upper Solid / μm	100
Depth in Lower Solid / μm	1000

Table C3 – Mesh Information for Chapter 7

EHL Mesh Points	1801
Inlet Boundary / x/a	-2.5
Inlet Boundary / x/a	2.0
Mesh Points in Film	21
Mesh Points in Upper Surface	21
Mesh Points in Lower Surface	21
Depth in Upper Solid / μm	50
Depth in Lower Solid / μm	50

Table C4 – Mesh Information for Chapter 8

EHL Mesh Points	900
Inlet Boundary / x/a	-3.0
Inlet Boundary / x/a	2.0
Mesh Points in Film	21
Mesh Points in Upper Surface	21
Mesh Points in Lower Surface	21
Depth in Upper Solid / μm	50
Depth in Lower Solid / μm	50

Appendix D

The roughness profiles are illustrated for the gear meshing simulations considered in Chapter 8. The roughness profiles are shown as a pair of traces within each figure. The upper profile is taken from the pinion and the lower profile corresponds to the wheel. The profiles for set A1 and A2 are measured from the root to the tip of the tooth, therefore the first point in the roughness profile is adjacent to the root feature. The profiles for set C are measured form the tip to the root and the fist point in the profile is adjacent to the tip feature.

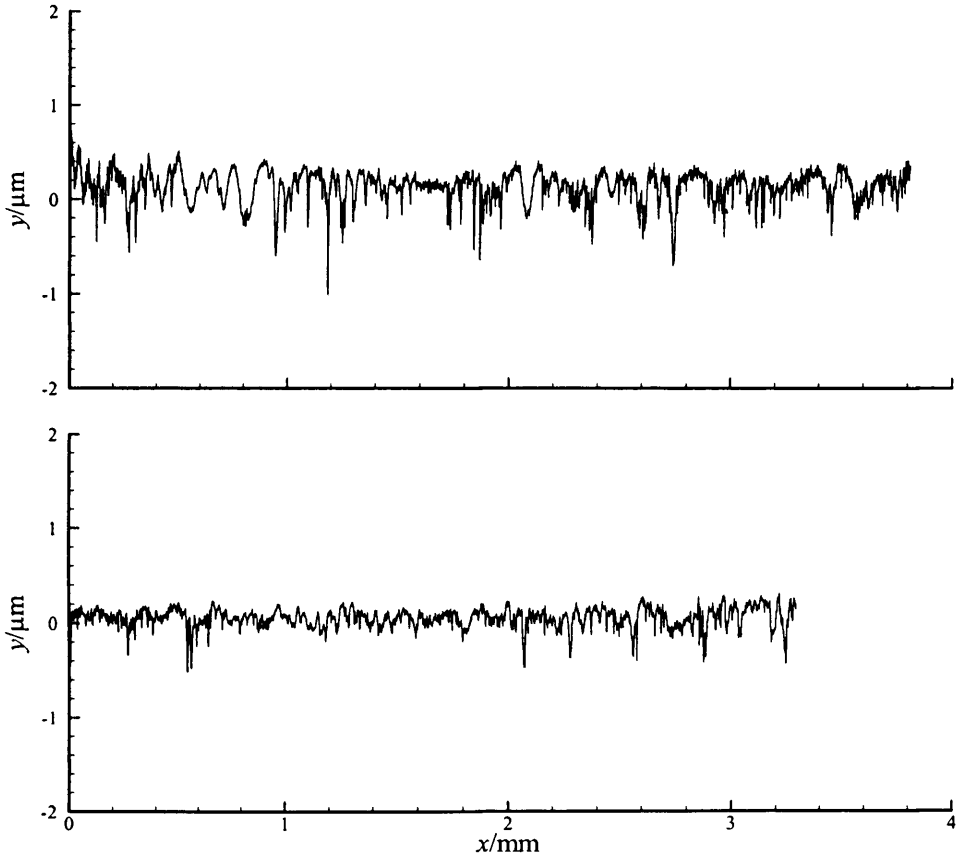

Figure D1 – Roughness profiles for Set A1, Pinion is upper profile and Wheel is lower profile

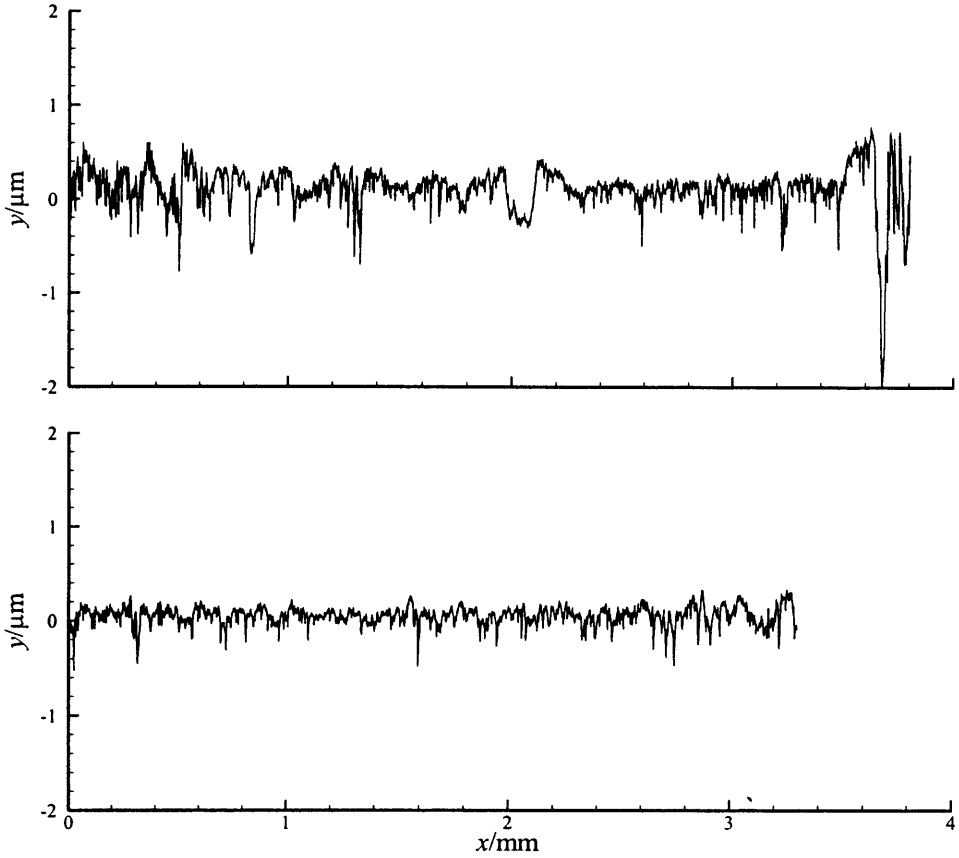

Figure D2 – Roughness profiles for Set A2, Pinion is upper profile and Wheel is lower profile

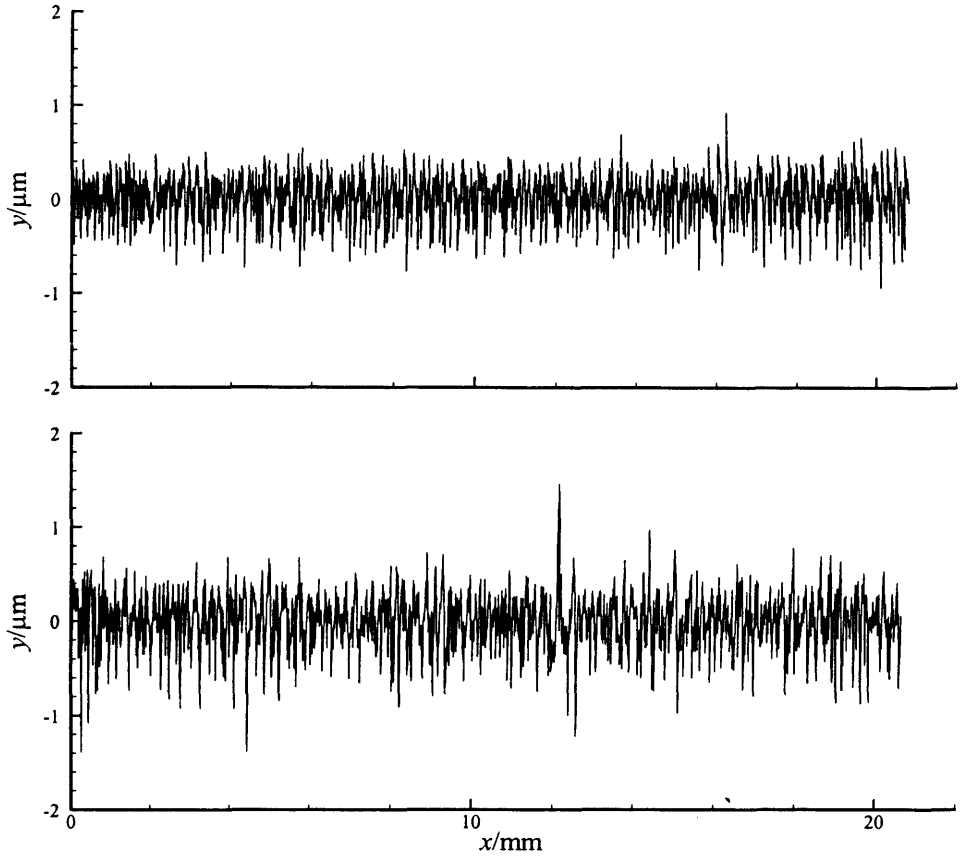

Figure D3 – Roughness profiles for Set C, Pinion is upper profile and Wheel is lower profile

References

Ai, X. and **Cheng, H.S.**, 1994, "A Transient EHL Analysis for Line Contacts with Measures Surface Roughness Using Multigrid Technique", *Trans. ASME, Journal of Tribology*, Vol. 116, No. 3, pp. 549 – 556.

Archard, J.F. and **Kirk, M.T.**, 1961, "Lubrication at Point Contacts", *Proc. R. Soc. Lond. A.*, Vol. 261, pp. 532 – 550.

Archard, J.F. and **Kirk, M.T.**, 1963, "Influence of Elastic Modulus on the Lubrication of Point Contatcs", *Proc. Lubrication and Wear Convention, Instn Mech. Engrs*, Paper 15, pp. 181 – 189.

Ausherman, V.K., Negaraj, H.S., Sanborn, D.M. and **Winer, W.O.**, 1976, "Infrared Temperature Mapping in Elastohydrodynamic Lubrication", *Trans. ASME, Journal of Lubrication Technology*, Vol. 98, No. 2, pp. 236 – 243.

Bair, S., 2004, "Roelands' Missing Data", *Proc. Instn Mech. Engrs, Part J, Journal of Engineering Tribology*, Vol. 218, pp. 57 – 60.

Bair, S. and **Winer, W.O.**, 1979, "A Rheology Model for Elastohydrodynamic Contacts Based on Primary Laboratory Data", *Trans. ASME, Journal of Lubrication Technology*, Vol. 101, pp. 258 – 265.

Barus, C., 1893, "Isothermal, Isopiestics and Isometrics in relation to Viscosity", *American Journal of Science, 3rd Series*, Vol. 45, pp. 87 – 96.

Booser, R.E., 1993, *CRC Handbook of Lubrication and Tribology: Monitoring, Materials, Synthetic Lubricants, and Applications, Volume III*, CRC Press.

Brandt, A., 1977, "Multi-Level Adaptive Solutions to Boundary-Value Problems", *Mathematics of Computation*, Vol. 31, No. 138, pp. 333 – 389.

Cameron, A. and **Gohar, R.**, 1966, "Theoretical and Experimental Studies of the Oil Film in Lubricated Point Contacts", *Proc. R. Soc. Lond. A.*, Vol. 291, pp. 520 – 536.

Carslaw, H.S. and **Jaeger, J.C.**, 1959, *Conduction of Heat in Solids*, Clarendon Press, Oxford.

Cann, P.M., Spikes, H.A. and **Hutchinson, J.**, 1996, "The Development of a Spacer layer Imaging Method (SLIM) for Mapping Elastohydrodynamic Contacts" *Tribology Transactions*, Vol. 39, No. 4, pp. 915 – 921.

Chang, L. and **Farnum, C.**, 1992, "A Thermal Model for Elastohydrodynamic Lubrication of Rough Surfaces", *Tribology Transactions*, Vol. 35, No. 2, pp.281 – 286.

Conry, T.F., Wang, S. and **Cusano, C.A.**, 1987, "A Reynolds-Eyring Equation for Elastohydrodynamic Lubrication in Line Contacts", *Trans. ASME, Journal of Tribology*, Vol. 109, pp. 648 – 658.

Crook, A.W., 1961a, "The Lubrication of Rollers, II. Film Thickness with Relation to Viscosity and Speed", *Philos. Trans. R. Soc. Lond. A.*, Vol 254, pp. 223 – 236.

Crook, A.W., 1961b, "The Lubrication of Rollers, III. A Theoretical Discussion of Friction and Temperatures in the Oil Film", *Philos. Trans. R. Soc. Lond. A.*, Vol 254, pp. 237 – 258.

Department of Education and Science, 1966, *Lubrication (Tribology) Education and Research. A Report on the Present Position and Industries Needs*, Department of Education and Science, HMSO, London.

Dowson, D., 1968, "Elastohdrodynamics", *Proc. Instn Mech. Engrs*, Vol. 182, Part 3A, pp 151 – 167.

Dowson, D., 1979, *History of Tribology*, Longman Group Limited, London.

Dowson, D. and **Higginson, G.R.**, 1959, "A Numerical Solution to the Elastohydrodynamic Problem", *Journal of Mechanical Engineering Science*, Vol. 1, No. 1, pp. 6 – 15.

Dowson, D. and **Higginson, G.R.**, 1966, *Elastohydrodynamic Lubrication*, Second Edition, Pergammon Press, Oxford.

Ehret, P., Dowson, D. and **Taylor, C.M.**, 1998, "On Lubricant Transport Conditions in Elastohydrodynamic Conjunctions", *Proc. R. Soc. Lond. A.*, Vol. 454, pp. 763 – 787.

Ehret, P., Dowson, D., Taylor, C.M. and **Wang, D.**, 1997, "Analysis of Isothermal Elastohydrodynamic Point Contacts Lubricated by Newtonian Fluids using Multigrid Methods", *Proc. Instn Mech. Engrs, Part C, Journal of Mechanical Engineering Science*, Vol. 211, No. 7, pp 493 – 508.

Elcoate, C.D., Evans, H.P., Hughes, T.G. and Snidle, R.W., 2001, "Transient Elastohydrodynamic Analysis of Rough Surfaces Using a Novel Coupled Differential Deflection Method", *Proc. Instn Mech. Engrs, Part J, Journal of Engineering Tribology*, Vol. 215, pp 319 – 337.

Evans, H.P. and **Hughes, T.G.**, 2000, "Evaluation of Deflection in a Semi-infinite Bodies by a Differential Deflection Method", *Proc. Instn Mech. Engrs, Part C, Journal of Mechanical Engineering Science*, Vol. 214, pp. 563 – 584.

Evans, C.R. and **Johnson, K.L.**, 1986, "The Rheological Properties of Elastohydrodynamic Lubricants", *Proc. Instn Mech. Engrs, Part C, Journal of Mechanical Engineering Science*, Vol. 200, pp. 303 – 312.

Eyring, H., 1936, "Viscosity, Plasticity and Diffusion as Examples of Absolute Reaction Rates", *Journal of Chemical Physics*, Vol. 4, pp. 283 – 291.

Furey, M.J., 1961, "Metallic Contact and Friction between Sliding Surfaces", *ASLE Transactions*, Vol. 4, pp. 1 – 11.

Gecim, B. and **Winer, W.O.**, 1980, "Lubricant Limiting Shear Stress Effect on EHD Film Thickness, *Trans. ASME, Journal of Lubrication Technology*, Vol. 102, pp. 213 – 221.

Glovnea, R.P. and **Spikes, H.A.**, 2001, "Elastohydrodynamic Film Formation at the Start-Up of the Motion", *Proc. Instn Mech. Engrs, Part J, Journal of Engineering Tribology*, Vol. 215, pp 125 – 138.

Gohar, R. and **Cameron, A.**, 1967, "The Mapping of Elastohydrodynamic Contacts", *ASLE Transactions*, Vol. 10, pp. 215 – 225.

Greenwood, J.A., 2000, "Two-dimensional Flow of a Non-Newtonian Lubricant", *Proc. Instn Mech. Engrs, Part J, Journal of Engineering Tribology*, Vol. 214, pp. 29 – 41.

Greenwood, J.A. and **Kauzlarich, J.J.**, 1973, "Inlet Shear Heating in Elastohydrodynamic Lubrication", *Trans. ASME, Journal of Lubrication Technology*, Vol. 95, pp. 417 – 426.

Grubin, A.N. and **Vinogradova, I.E.**, 1949, "Investigation of the Contact of machine Components", *Central Scientific Research Institute for Technology and Mechanical Engineering*, Book No. 30, Moscow [DSIR Translation No. 337]

Hertz, H., 1881, "Über die Berührung fester Elastischer Körper", [On the Contact of Elastic Solids], *J. reine und angewandte Mathematik*, 92, 156 – 171, [English Translation: **Jones** and **Schott**, Eds., 1896, *Misc. Papers by H. Hertz*, Macmillan, London.].

Holmes, M.J.A., Evans, H.P., Hughes, T.G. and **Snidle, R.W.**, 2003a, "Transient Elastohydrodynamic Point Contact Analysis Using a New Coupled Differential Deflection Method Part 1: Theory and Validation", *Proc. Instn Mech. Engrs, Part J, Journal of Engineering Tribology*, Vol. 217, No. 4, pp. 289 – 304.

Holmes, M.J.A., Evans, H.P., Hughes, T.G. and **Snidle, R.W.**, 2003b, "Transient Elastohydrodynamic Point Contact Analysis Using a New Coupled Differential Deflection Method Part 2: Results", *Proc. Instn Mech. Engrs, Part J, Journal of Engineering Tribology*, Vol. 217, No. 4, pp. 305 – 322.

Holmes, M.J.A., Evans, H.P. and **Snidle, R.W.**, 2004, "Surface Contact in Micro-EHL", *IUTAM Symposium on Elastohydrodynamics and Micro-Elastohydrodynamics, Cardiff, Sept. 2004* (To be published).

Holmes, M.J.A, Qiao, H., Evans, H.P. and **Snidle, R.W.**, 2005, "Surface Contact and Damage in Micro-EHL", *Proc. of the 31st Leeds-Lyon Symp. on Tribology*, pp. 605 – 616.

Holmes, M.J.A, 2002, *Transient analysis of the Point Contact Elastohydrodynamic Lubrication Problem Using Coupled Solution Methods*, PhD Thesis, Cardiff University.

Houpert, L. and **Hamrock, B.J.**, 1986, "Fast Approach for Calculating Film Thickness and Pressure in Elastohydrodynamically Lubricated Contacts", *Trans. ASME, Journal of Tribology*, Vol. 107, pp. 128 – 131.

Hsiao, H.S.S. and **Hamrock, B.J.**, 1992, "A Complete Solution for Thermal-Elastohydrodynamic Lubrication of Line Contacts Using Circular Non-Newtonian Fluid Model", *Trans. ASME, Journal of Tribology*, Vol. 114, pp. 540 – 552.

Hu, Y.Z. and **Zhu, D.**, 2000, "A Full Numerical Solution to the Mixed Lubrication in Point Contacts", *Trans. ASME, Journal of Tribology*, Vol. 122, pp. 1 – 9.

Iivonen, H. and **Hamrock, B.J.**, 1991, "A new Non-Newtonian Fluid Model for Elastohydrodynamic Lubrication of Rectangular Contacts", *Wear*, Vol. 143, pp. 297 – 305.

Jackson, A. and **Cameron, A.**, 1976, "An Interferometric Study of the EHL of Rough Surfaces", *ASLE Transactions*, Vol. 19, pp. 50 – 60.

Jacobson, B.O. and **Hamrock, B.J.**, 1984, "Non-Newtonian Fluid Model Incorporated into Elastohydrodynamic Lubrication of Rectangular Contacts", *Trans. ASME, Journal of Tribology*, Vol. 106, pp. 275 – 284.

Johnson, K.L., 1985, *Contact Mechanics*, Cambridge University Press, Cambridge.

Johnson, K.L. and **Tavaarwerk, J.L.**, 1977, "Shear Behaviour of Elastohydrodynamic Films", *Proc. R. Soc. Lond. A.*, Vol. 356, pp. 215 – 236.

Johnston, G.J, Wayte, R., and **Spikes, H.A.**, 1991, "The Measurement and Study of Very Thin Lubricant Films in Concentrated Contacts", *Tribology Transactions of STLE*, Vol. 34, pp. 187 – 194.

Kaneta, M. and **Cameron, A.**, 1980, "Effects of Asperities in Elastohydrodynamic Lubrication", *Trans. ASME, Journal of Lubrication Technology*, Vol. 102, pp. 374 – 379.

Kaneta, M. Sakai, T. and **Nishikawa, H.**, 1992, "Optical Interferometric Observations of the Effects of a Bump on Point Contact EHL", *Trans. ASME, Journal of Tribology*, Vol. 114, pp. 779 – 784.

Kaneta, M. Sakai, T. and **Nishikawa, H.**, 1993, "Effects of Surface Roughness on Point Contact EHL", *Tribology Transactions*, Vol. 36, pp. 605 – 612.

Kannel, J.W., Bell, J.C. and **Allen, C.M.**, 1964, "Methods for Determining Pressure Distribution in Lubricated Rolling Contact", *ASLE Paper No. 64 LC-24, ASME – ASLE International Lubrication Conference, Washington DC.*

Kirk, M.T., 1962, "Hydrodynamic Lubrication of Perspex", *Nature*, Vol. 194, No. 4832, pp. 965 – 967.

Kweh, C.C., Patching, M.J., Evans, H.P. and **Snidle, R.W.**, 1992, "Simulation of Elastohydrodynamic Contacts Between Rough Surfaces", *Trans. ASME, Journal of Tribology*, Vol. 114, No. 3, pp. 412 – 419.

Lane, T.B. and **Hughes, J.R.**, 1952, "A Study of the Oil Film Formation in Gears by Electrical Resistance Measurements", *Journal of Applied Physics*, Vol. 3 pp. 315 – 318.

Larsson, R., 1997, "Transient Non-Newtonian Elastohydrodynamic Lubrication Analysis of an Involute Spur Gear", *Wear*, Vol. 207, pp. 67 – 73.

Larsson, R., Larsson, P.O, Eriksson, E., Sjöberg, M. and **Höglund, E.**, 2000, "Lubricant Properties for Input to Hydrodynamic and Elastohydrodynamic Lubrication Analyses", *Proc. Instn Mech. Engrs, Part J, Journal of Engineering Tribology*, Vol. 214, pp 17 – 27.

Lee, R.T. and **Hamrock, B.J.**, 1990, "A Circular Non-Newtonian Fluid Model: Part I – Used in Elastohydrodynamic Lubrication", *Trans. ASME, Journal of Tribology*, Vol. 112, pp. 486 – 496.

Lee, K.M. and Cheng, H.S., 1973, "Effects of Surface Asperity on Elastohydrodynamic Lubrication", *NASA CR-2195, Washington, D.C.*

Lubrecht, A.A., ten Napel, W.E. and **Bosma, R.**, 1986, "Multigrid, an Alternative Method of Solution for Two-Dimensional Elastohydrodynamically Lubricated Point Contact Calculations", *Trans. ASME, Journal of Tribology*, Vol. 108, No. 3, pp. 551 – 556.

Lubrecht, A.A., ten Napel, W.E. and **Bosma, R.**, 1988, "The Influence of Longitudinal and Transverse Roughness on the Elastohydrodynamic Lubrication of Circular Contacts", *Trans. ASME, Journal of Tribology*, Vol. 110, No. 3, pp. 421 – 426.

Martin, H.M., 1916, "The Lubrication of Gear Teeth", *Engineering*, Vol. 102, pp. 199 – 204.

Merritt, H.E., 1935, "Worm Gear Performance", *Proc. Instn Mech. Engrs*, Vol. 129, pp. 127 – 194.

Okamura, H., 1983, "A Contribution to the Numerical Analysis of Isothermal Elastohydrodynamic Lubrication", *Proc. of the 9th Leeds-Lyon Symp. on Tribology*, pp. 313 – 320.

Ostvik, R. and **Christensen, H.**, 1968, "Changes in Surface Topography with Running-In", *Proc. Instn Mech. Engrs*, Vol. 183, Pt. 3, pp 57 – 65.

Paliwal, M.C., 1987, *Running-in and Scuffing Failure of Marine Gears*, PhD Thesis, University of Wales, Cardiff.

Patching, M.J., 1994, *The Effects of Surface Roughness on the Micro-Elastohydrodynamic Lubrication and Scuffing Performance of Aerospace Gear Tooth Contact*, PhD Thesis, University of Wales, Cardiff.

Patir, N. and **Cheng, H.S.**, 1978, "An Average Model for Determining Effect of Three Dimensional Roughness on Partial Hydrodynamic Lubrication", *Trans. ASME, Journal of Lubrication Technology*, Vol. 100, pp. 12 – 17.

Petrov, N., 1883, "Friction in Machines and the Effect of the Lubricant", *Inzh. Zh., St-Peterb.*, Vol. 1, pp. 71 – 140; Vol. 2, pp. 227 – 279; Vol. 3, pp. 377 – 436; Vol. 4, pp.535 – 64.

Petrusevich, A.I., 1951, "Fundamental Conclusions From the Contact-Hydrodynamic Theory of Lubrication", *Izv. Akad. Nauk. SSSR (OTN)*. Vol. 2, pp. 1 -11. [Ministry of Defence Translation No. 293]

Poon, C.Y. and **Sayles, R.S.**, 1994, "Numerical Contact Model of a Smooth Ball on an Isotropic Rough Surface", *Trans. ASME, Journal of Tribology*, Vol. 116, No. 2, pp. 194 – 201.

Qiao, H., 2005, *Prediction of Contact Fatigue for the Rough Surface Elastohydrodynamic Lubrication Line Contact Problem under Rolling and Sliding Conditions*, PhD Thesis, Cardiff University. (To be published)

Reddi, M.M., 1969, "Finite Element Solution of the Incompressible Lubrication Problem", *Trans. ASME, Journal of Lubrication Technology*, Vol. 91, pp. 524 – 532.

Reddi, M.M., 1970, "Finite Element Solution of the Steady-State Compressible Lubrication Problem", *Trans. ASME, Journal of Lubrication Technology*, Vol. 92, No. 3, pp. 495 – 503.

Rennie, G., 1829, "Experiments on the Friction and Abrasion of the Surfaces of Solids", *Philos. Trans. R. Soc. Lond.*, Vol. 34, part 1, pp. 143 – 170.

Reynolds, O., 1886, "On the Theory of Lubrication and its Application to Mr. Beauchamp Tower's Experiments, Including an Experimental Determination of the Viscosity of Olive Oil", *Philos. Trans. R. Soc. Lond.*, Vol. 177, pp.157 – 234.

Rohde, S.M. and **Oh, K.P.**, 1975, "A Unified Treatment of Thick and Thin Elastohydrodynamic Problems by Using Higher Order Element Methods", *Proc. R. Soc. Lond. A.*, Vol. 343, pp. 315 – 331.

Roelands,, C.J.A., 1966, *Correlation Aspects of the Viscosity – Temperature – Pressure Relationship of Lubricating Oils*, PhD Thesis, Delft University, Delft [Druk, V.R.B., Groningen.].

Salehizadeh, H. and **Saka, N.**, 1991, "Thermal Non-Newtonian Elastohydrodynamic Lubrication of Rolling Line Contatcs", *Trans. ASME, Journal of Tribology*, Vol. 113, pp. 481 – 491.

Sharif, K., Kong, S., Evans, H.P. and **Snidle, R.W.**, 2001a, "Contact and Elastohydrodynamic Analysis of Worm Gears: Part 1 Theoretical Formulation", *Proc. Instn Mech. Engrs, Parts C, Journal of Mechanical Engineering Science*, Vol. 215, pp. 817 – 830.

Sharif, K., Morris, S.J., Evans, H.P. and **Snidle, R.W.**, 2001b, "Comparison of Non-Newtonian EHL Models in High Sliding Applications", *Proc. of the 27th Leeds-Lyon Symp. on Tribology*, pp. 787 – 796.

Sharif, K., Kong, S., Evans, H.P. and **Snidle, R.W.**, 2001c, "Elastohydrodynamic Analysis of Worm Gear Contacts", *Proc. International Tribology Conference, Japanese Society of Tribologists, Nagasaki*, pp. 1867 – 1872.

Sharif, K., Evans, H.P., Snidle, R.W. and **Newall, J.P.**, "Modelling of Film Thickness and Traction in a Variable Ratio Traction Drive Rig", *Trans. ASME, Journal of Tribology*, Vol. 126, No. 1, pp. 92 – 104.

Shieh, J.A. and **Hamrock, B.J.**, 1991, "Film Collapse in EHL and Micro-EHL", *Trans. ASME, Journal of Tribology*, Vol. 113, pp. 372 – 377.

Sibley, L.B., Bell, J.C., Orcutt, F.K. and **Allen, C.M.**, 1960, "A Study of the Influence of Lubricant Properties on the Performance of Aircraft Gas Turbine Engine Rolling-Contact Bearings", *WADD Tech. Rep.*, pp. 60 - 189

Smith, C.F., 1959, "Lubrication Behaviour in Concentrated Contact Systems – the Castor Oil-Steel System", *Wear*, Vol. 2, No. 4, pp 250 – 263.

Sui, P.C. and **Sadeghi, F.**, 1991, "Non-Newtonian Thermal Elastohydrodynamic Lubrication", *Trans. ASME, Journal of Tribology*, Vol. 113, pp. 390 – 397.

Tao, J., Hughes, T.G., Evans, H.P., Snidle, R.W., Hopkinson, N.A., Talks, M. and **Starbuck, J.M.**, 2003, "Elastohydrodynamic Lubrication Analysis of Gear Tooth Surfaces From Micropitting Tests", *Trans. ASME, Journal of Tribology*, Vol. 125, No. 2, pp. 267 – 274.

Taylor, C. and **O'Callaghan, J.F.**, 1972, "A Numerical Solution to the Elastohydrodynamic Lubrication Problem Using Finite Elements", *Journal of Mechanical Engineering Science*, Vol. 14, No. 4, pp. 229 – 237.

Tower, B., 1883, "First Report on friction Experiments (Friction of Lubricated Bearings)", *Proc. Instn Mech. Engrs*, Nov., pp. 632 – 659.

Tower, B., 1885, "Second Report on friction Experiments (Experiments on the Oil Pressure in a Bearing)", *Proc. Instn Mech. Engrs*, Jan., pp. 50 – 70.

Turcina, V., Sanborn, D.M. and **Winer, W.O.**, 1974, "Temperature Measurements in Sliding Elastohydrodynamic Point Contacts", *Trans. ASME, Journal of Lubrication Technology*, Vol. 94, No. 3, pp. 464 – 471.

Venner, C.H. and **Lubrecht, A.A.**, *Multilevel Methods in Lubrication*, Tribology Series 37, Elsevier, 2000.

Venner, C.H., 2005, "EHL Film Thickness Computations at Low Speeds: Risk of Artificial Trends as a Result of Poor Accuracy and Implications for Mixed Lubrication Modelling", *Proc. Instn Mech. Engrs, Part J, Journal of Engineering Tribology*, Vol. 219, pp 285 – 289.

Wang, Y., Li, H., Tong, J and **Yang, P.**, 2004, "Transient Thermoelastohydrodynamic Lubrication Analysis of an Involute Spur Gear", *Tribology International*, No. 37, pp.773 – 782.

Wedeven, L.D., 1970, *Optical Measurements in Elastohydrodynamic Rolling-Contact Bearings*, PhD Thesis, University of London.

Wedeven, L.D. and **Cusano, C.**, 1979, "Elastohydrodynamic Film Thickness Measurements of artificially Produced Surface Dents and Grooves", *ASLE Transactions*, Vol. 22, pp. 396 – 381.

Westlake, F.J. and **Cameron, A.**, 1967, "Fluid Film Interferometry in Lubrication Studies", *Nature*, Vol. 214, pp. 633 – 634.

Wu, C.S., Klaus, E.E. and **Duda, J.L.**, 1989, "Development of a Method for the Prediction of Pressure-Viscosity Coefficient of Lubricating Oils Based on Free-Volume Theory", *Trans. ASME, Journal of Tribology*, Vol. 111, pp. 121 – 128.

Yagi, K., Kyogoku, K. and **Nakahara, T.**, 2004, "Relationship Between Temperature Distribution in EHL Film and Dimple Formation", Proc. 2004 ASME/STLE International Joint Tribology Conference, California.

Zhang, Y.B. and **Wen, S.Z.**, 2002, "An Analysis of Elastohydrodynamic Lubrication with Limiting Shear Stress: Part 1 – Theory and Solutions", *Tribology Transactions*, Vol. 45, No. 2, pp. 135 – 144.

Zhang, Y.B., Wen, S.Z. and **Wei, D.P.**, 2001, "About the Load – Carrying Capacity of Elastohydrodynamic Lubrication Film", *Tribology Transactions*, Vol. 44, No. 1, pp. 1 – 10.