TEMPERATURE AND SHEAR STRESS IN ROLLING/SLIDING ELASTOHYDRODYNAMIC CONTACTS

R. GRIEVE, H. A. SPIKES
Tribology Section, Department of Mechanical Engineering, Imperial College, London SW7 2BX, UK;
e-mail: h.spikes@ic.ac.uk

SUMMARY
There is growing need to understand the origins and magnitude of friction in full film elastohydrodynamic (EHD) conditions, both to minimise it in some applications, and thus to save energy, but also to maximise and predict it in others, such as traction drives. One barrier to developing reliable models of EHD friction is that measurements of friction actually represent averaged values over the whole, lubricated contact under study. However the fluid film conditions of temperature, pressure and strain rate generally vary over such contacts, which makes it difficult to determine constitutive friction equations. This paper examines the various different techniques used to study the origins of EHD friction and the underlying film rheology. It then focuses on a technique for obtaining detailed maps of shear stress and thus friction across EHD contacts based on the measurement of temperature rise within contacts.

Keywords: EHD, Rheology, Infrared, Shear stress, Traction

1 INTRODUCTION
Because of growing concern to increase the energy-efficiency of machine components it is of considerable practical performance to be able to predict the friction (often called “traction”) of elastohydrodynamic (EHD) lubricant films. In many cases the aim is to reduce this friction to a minimum by appropriate combination of lubricant and operating conditions. In some applications, however, such as traction drives, high EHD friction over a wide range of conditions is sought.

The friction in a full-film contact is determined by the integral over the contact area of the shear stresses generated in the lubricant film and is thus dependent on the rheological response of the lubricant film to the conditions within the contact. In rolling/sliding EHD contacts these conditions are extraordinarily severe, involving the short duration application of both extremely high pressures, of the order of 1 to 3 GPa, and very rapid strain rates, typically 10⁶ to 10⁸ s⁻¹. The high pressure produces a very large increase in the viscosity of the lubricant within the film and the combination of this with the high, applied strain rate, results in the oil experiencing a very high shear stress. At such high shear stresses the lubricant behaves in a highly non-Newtonian fashion. To understand and thus be able to predict friction in EHD contacts we need to have an accurate rheological model to describe this non-Newtonian behaviour.

A number of constitutive equations have been proposed to describe the rheology of liquids in EHD conditions [1], but there is still considerable dispute about their validity. High-pressure viscometry studies have tended to focus on Newtonian behaviour, curtailed, at high shear stress, by a limiting shear stress response. EHD studies have generally placed more emphasis on shear thinning above a certain shear stress and on viscoelastic behaviour.

The currently most widely-used model is the Johnson and Tevaarwerk or Eyring-Maxwell-limiting shear stress model. This, shown in equation 1 below, is based on the simple viscoelastic Maxwell model modified to incorporate a Eyring shear thinning component rather than a Newtonian shear behaviour. Superimposed on this is a limiting shear stress, above which the fluid yields at constant stress [2, 3].

\[
\dot{\gamma} = \frac{\tau \sinh \left( \frac{\tau}{\tau_c} \right)}{\eta_o} + \frac{\dot{\gamma}}{G} \quad \text{for } \tau < \tau_c \quad (1a)
\]

\[
\tau = \tau_c \quad \text{for } \tau \geq \tau_c \quad (1b)
\]

where:
- \(\tau\) is the shear stress
- \(\tau_c\) is the Eyring stress, the stress at which shear thinning begins to take place.
- \(\tau_c\) is the limiting shear stress.
- \(\eta_o\) is the low shear rate viscosity of the fluid.
- \(G\) is the shear modulus of the fluid.
- \(\dot{\gamma}\) is the shear rate.

In principle, given the above six parameters it is possible to integrate equation 1 to extract the average shear stress in the contact for all conditions. There are, however, problems. One is that \(\tau_c, \tau_c, G\) and \(\eta_o\) all vary with temperature and pressure, while the temperature of the film is itself dependent on the shear stress. A second is that the validity of this equation is debatable. The sinh relationship between strain rate and shear stress is based on an activated flow model of fluid shear, which has never been fully validated while the Maxwell viscoelastic model is a small strain model, unlikely to be accurate for high strains [4]. Limiting shear stresses have only been noted for high viscosity or traction fluids, and it is not clear whether they occur for normal hydrocarbon or ester lubricants.

The main problem in validating equations of rheology in EHD conditions is that it is very difficult to measure the shear stress properties of lubricants at conditions that are both representative of those found in EHD contacts and
are accurately defined. This problem will be discussed in the next section. The current paper first examines the main methods used to study lubricant rheology in EHD and thus the origins of EHD friction. It then focuses on a relatively new technique for mapping shear stress in contacts, based on temperature rise measurements.

2 METHODS OF STUDYING EHD RHEOLOGY

There have been five main experimental approaches to study the rheology of lubricant films in EHD conditions.

One is high-pressure viscometry, where conventional methods of measuring fluid rheology, such as a falling ball or rotating, concentric cylinder configuration are applied in a high-pressure device. Most previous work has been at low strain rates [5-10]. This has shown that the viscosity of fluids generally rises exponentially with pressure up to 0.1 GPa, but above this there is often a levelling-out in the viscosity-pressure dependence, followed, above a critical pressure, by a very rapid rise in viscosity [11]. Using high-pressure viscometry it is quite difficult to apply very high shear stresses and thus impose high strain rates on viscous fluids. However Bair and Winer have employed a piston arrangement and a rotating cylinder method to reach shear stresses up to 200 MPa [12-14]. They have shown that, at very high shear stresses, fluids can display a limiting shear stress at which they yield in a fully plastic fashion.

A second approach is based on impact rheometry. In this, a thin film of fluid is sandwiched between the surfaces of a pair of solid bodies and one of these is subjected to a sudden shock wave, typically by impact of a projectile. The resulting displacement of the opposing body is measured to determine the response of the fluid to sudden compression [15] and, by angling the fluid layer, it is possible to measure both the compressive and shear response of the fluid under the transient imposition of a very high pressure and shear wave [16]. This approach has been used to confirm the existence of a limiting shear stress in fluids under shear at very high pressure [17]. In a related approach, a torsion bar has been used to apply very rapid rotational shear to a thin film of super-cooled liquid and shear thinning behaviour investigated [18].

Both of the above have limitations with respect of mimicking EHD conditions. With high-pressure rheometry, it is not yet possible to reach the very high strain rates present in most EHD contacts. This means that most work at high shear stresses has been confined to atypical, high viscosity fluids. There is, indeed, a problem in using very high strain rates, in that these result in large heat generation, making it difficult to disentangle the rheological from the thermal response. Also, high-pressure rheometry fails to match the very sudden, short duration aspect of the imposed shear present in EHD contacts. With impact rheometry, the transient conditions are very well-matched and it is also easy to reach very high pressures, up to 3 GPa and higher. It is, however, more difficult to obtain accurate rheological data at lower pressures, below 1 GPa, because the displacements involved are very small and difficult to measure. Another limitation is that the fluid strain is quite small, much less than in an EHD contact, and it is not certain whether the rheology of liquids at high strain is similar as that at low strain.

The third and most common approach to studying EHD rheology is to set up a steady-state EHD film, generally in a twin disc or ball on flat contact, and to measure friction over a range of rolling/sliding/spin conditions [3, 19-25]. The ratio of the friction to the mean contact pressure then provides a measure of the mean shear stress of the film within the contact. Also, if the film thickness can be reliably estimated, and is reasonably constant over the contact, the corresponding mean strain rate can be estimated. This approach has been used to show the presence of a viscoelastic response of the fluid film in the contact at high pressures and entrainment speeds [20]. It has also clearly demonstrated the occurrence of shear thinning of the fluid in an EHD contact at shear stresses above about 5 MPa [2, 22], and some evidence of a limiting shear stress [19]. Disc machine studies have indeed, led to the most widely-accepted fluid rheological model (equation 1) for EHD conditions [3, 25].

The EHD contact approach has the major advantage in that the fluid is clearly being subjected to realistic EHD conditions. It has, however, several weaknesses. The main one is that the method is based on average measurements of shear stress over a whole contact, whereas the applied conditions of pressure, temperature, and, to a lesser extent shear rate vary across this contact. This has been justified on the grounds that, during shear thinning, shear stress appears to vary linearly with pressure rather than exponentially as expected for a Newtonian fluid [3]. However, there is a danger that the averaging may obscure important details of the rheological response as would be the case, for example, if there were a significant time effect in the response of the fluid to contact entry, so that the shear stress were different in the inlet as compared to the exit half of the contact. Interestingly, the friction method of studying EHD rheology does not, when based on analysis using the Eyring model, predict the very rapid rise in pressure viscosity coefficient noted at very high pressures using high-pressure viscometer [26]. A second disadvantage is, that to interpret results, film temperature must be calculated from theory. This is not straightforward, not least because the oil film temperature rise depends upon the unknown velocity and thus shear gradient across the film, and also because values such as fluid film conductivity at high pressure and shear rate are not known.

A fourth method of studying lubricant rheology is the so-called “bouncing ball” technique [27, 28]. In this, the forces experienced by a ball bouncing on a lubricant-coated flat are determined by monitoring the ball rotation and trajectory. This has been claimed to show the existence of a limiting shear stress in the fluid during impact. The technique has the advantage of easily producing very high contact pressures and minimising
thermal effects due to the short duration of the impact process. However, the method provides only limited control of the contact conditions, since it is not straightforward to vary factors such as slide-roll ratio and film thickness independently and also suffers from the averaging-across-contact limitation of steady state contact. The technique has recently been further developed by monitoring the propagating compression and flexural waves in the support of the flat to derive simultaneous pressure and friction values over time of impact to form a device similar to that used in impact rheometry described above, but using a ball on flat rather than flat on flat contact [29].

The fifth main way to study lubricant EHD rheology is the subject of the current paper. This was developed in the early 1990s and is based, not on measuring friction force, but on measuring temperature rise [30, 31, 32]. Almost all of the energy dissipation and thus temperature rise within a rolling-sliding contact results directly from the shear and thus the friction of the fluid film. While the friction cannot yet be measured locally within a contact, the temperature rise can. From this temperature rise, the friction that produces it can be deduced as a map across the contact. This technique is further described below. The advantages of the method are clear. Local shear stress values can be obtained rather than the averages derived from overall friction measurements, while the corresponding local temperature is also available, reducing one of the main uncertainties in the EHD friction approach. The main disadvantages are difficult experimentation and limited resolution. In the past, this temperature-mapping method has been applied mainly to pure sliding contacts [30, 31]. The current paper shows how it can be applied with higher resolution to mixed rolling-sliding contacts and be used to determine shear stress data at known pressure and shear rate and reasonably well-defined temperature.

3 TEMPERATURE MAPPING METHOD AND TEST RIGS

The first stage of the approach involves mapping temperature rise in a contact. This problem was first solved by Winer, who mapped temperature in EHD contact, using an infrared (IR) microscope to measure thermal emission both from the steel surface and from the lubricant in a sliding contact between steel ball and sapphire disc [33, 34]. These emission measurements were then converted to temperatures using a combined calibration/analysis approach. The current study also used a steel ball on sapphire flat contact but the disc and ball can be independently driven, to obtain any desired sliding/rolling combination, as indicated in figure 1.

Tests were also carried out in a pure sliding system as shown in figure 2.

In both cases, the contact is contained in a temperature-controlled chamber ($\pm 0.5 \, ^{\circ}C$) and the ball is half-immersed in lubricant to ensure fully-flooded conditions. A custom-built, infrared microscope is mounted on an XY table driven by micrometers attached to computer-controlled stepper motors. This microscope is focussed on the steel ball surface within the contact and uses an InSb detector to sample the level of IR emission from the focus area. The microscope is able to focus on a spot approximately $11 \, \mu m$ in diameter. Since the Hertzian diameter is typically $300 \, \mu m$, this means that up to 27 non-overlapping emission and thus temperature readings are possible across the contact diameter.

In an individual experiment, the sliding/rolling motion and bulk temperature are stabilised and the microscope is focussed on the centre of the contact. Then the microscope is moved under computer-control to fully traverse the contact, and IR emission measured at a grid of locations from the inlet to the exit region of the contact, as shown schematically in figure 3.
Figure 3: Grid locations for temperature measurement

A transmission filter is used to remove any IR emission from the hydrocarbon oil film present. The IR emission from the sapphire can be neglected, which means that all measured emission originates from the steel surface. This emission is related to surface temperature by means of a calibration in which IR emission was monitored through a sapphire window from a polished steel surface at a series of temperatures.

4 DETERMINATION OF SHEAR STRESS MAPS

The shear stress calculation technique is based on the principle that, in fluid shear, the local rate of heat generation per unit area in a columnar volume of fluid within the contact, \( \dot{q} \), is given by;

\[
\dot{q} = \tau U_s
\]

(2)

where \( \tau \) is the mean shear stress through the film and \( U_s \) is the sliding speed. It is assumed that this heat passes rapidly into the two bounding surfaces. Then, if \( q_{steel} \) and \( q_{sapp} \) are the rates of heat input into the steel ball and sapphire disc surfaces respectively, as shown in figure 4;

\[
q_{steel} + q_{sapp} = \tau U_s
\]

(3)

If these heat input rates can be obtained, the shear stress can be calculated. The calculation of \( q_{steel} \) and \( q_{sapp} \) is based on the work of Jaeger, who formulated equations for the temperature rise of the surface of a body due to a moving uniform heat input [35].

Jaeger’s equations are for a uniform heat source, while, in practice, heat generation and thus heat input into the solid surfaces varies across the contact. The contact is therefore discretised and it is assumed that the contact is made up of many separate, rectangular heat sources, all moving at the same speed and collectively forming the overall heat input to the surfaces. It is convenient to assume that each measurement point in the temperature map represents the centre of a separate, rectangular heat source.

The heat input at any one location will cause a temperature rise at all points on the surface, which means that the overall temperature rise at any given position results from the sum of the influence of heat input at all locations within the contact. Jaeger provides a means of calculating the influence of heat input at any location on the temperature rise at any other location. These influence coefficients are dependent upon several factors; the vectorial distance between any two points, the thermal properties of the solid surface and the velocity at which the sources are moving with respect to the surface. For a moving surface, they are expressed by:

\[
C_{ij}^{kl} = \frac{1}{2\pi K} \int \int_A \frac{\exp \left[ \frac{U}{2\chi} \left( x_i - x_k \right) + R \right]}{R} \, dx \, dy
\]

(4)

where:
- \( K \) is the thermal conductivity of the surface
- \( \chi \) is the thermal diffusivity of the surface material
- \( R \) is the distance between points
- \( U \) is the speed of the surface relative to the heat source

Once all the coefficients are calculated, a coefficient matrix is assembled, from which the set of individual heat sources can be determined from the set of measured surface temperature rise values by solving the set of equations;

\[
\Delta T_{ij} = \sum_{kl} C_{ij}^{kl} q_{kl}
\]

(5)

Shear stresses can then be determined from heat inputs using equation 2. The main assumptions made in this analysis are; (i) all heat generated passes rapidly by conduction into the bounding surfaces; (ii) the surface temperature of the disc at all points in the contact is the same as that of the ball; (iii) all heat is generated by fluid shear (not by compression). These assumptions and their limitations are discussed in [32].

5 TEST OILS AND CONDITIONS

Two test oils were used, with properties listed in table 1. One was an additive-free mineral oils and the other was a commercial traction fluid. In the rolling-sliding tests, a 19 mm diameter steel ball was used with an applied load of 48 N. In the pure sliding rig, a 25.4 mm diameter ball was used at 60 N.
<table>
<thead>
<tr>
<th></th>
<th>Viscosity $\eta$ cP</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>30 °C</td>
</tr>
<tr>
<td>HVI 115</td>
<td>Mineral oil</td>
</tr>
<tr>
<td>Santotrac 50</td>
<td>Commercial traction fluid</td>
</tr>
</tbody>
</table>

Table 1: Test oils studied

6 VALIDATION

Validation of the shear stress calculation was carried out in the pure sliding rig shown in figure 2, where friction could be monitored simultaneously with the acquisition of temperature maps.

The calculation of shear stress from temperature rise is strongly dependent on the local temperature gradient, so that any random errors in temperature measurement are magnified. This can be alleviated by smoothing the temperature data, but care must be taken not to degrade the integrity of the measurements. Figures 5 and 6 compare temperature rise and shear stress maps using a 20 x 20 grid for unsmoothed temperature data and temperature data smoothed based on the values of neighbouring points (inlet on left). Smoothing was based on a central point weighting of 4 compared to N, S, E and W point weightings of unity. The beneficial effect of the smoothing on the shear stress map can be seen in figure 6.

One deleterious effect of temperature-smoothing was that it tended to reduce the overall temperature across the contact region. Since the shape of the temperature map in this region was convex, smoothing based on neighbouring points produced a net temperature reduction.

This was corrected by applying a multiplier after smoothing (usually of approximately 3 %), which was based on change in the integrated shear stress value across the contact resulting from the temperature-smoothing process. The effect of this is also shown in figure 7. In practice, the net effect on calculated shear stress was quite small, as shown in figure 7.
Tests were carried out using a range of different grid sizes. The choice of grid size is a compromise. Sufficient measurements must be taken to obtain an accurate temperature and thus shear stress map. However both the time of measurement and the computing time for shear stress determination increase rapidly with size of grid. To validate the method and also explore the influence of grid size, shear stress results were compared to measured EHD friction values. The shear stress maps were integrated over the contact area (neglecting the negative shear stress observed at the exit – see section 7 below), and divided by applied load to obtain “calculated” friction coefficients. Table 2 compares measured and calculated friction coefficients for three grid sizes. It can be seen that the difference between the two falls with increasing grid size, to reach approximately 7% for a 28 x 28 grid. Results from three experiments are shown for this grid size, to indicate the level of repeatability.

<table>
<thead>
<tr>
<th>Grid size</th>
<th>Measured T</th>
<th>Calculated T</th>
<th>% diff.</th>
</tr>
</thead>
<tbody>
<tr>
<td>20x20</td>
<td>0.0258</td>
<td>0.0224</td>
<td>13%</td>
</tr>
<tr>
<td>24x24</td>
<td>0.0259</td>
<td>0.0233</td>
<td>10%</td>
</tr>
<tr>
<td>28x28 no.1</td>
<td>0.0259</td>
<td>0.0241</td>
<td>7%</td>
</tr>
<tr>
<td>28x28 no.2</td>
<td>0.0254</td>
<td>0.0239</td>
<td>6%</td>
</tr>
<tr>
<td>28x28 no.3</td>
<td>0.0243</td>
<td>0.0221</td>
<td>9%</td>
</tr>
</tbody>
</table>

Table 2: Difference between measured and calculated EHD friction coefficient

7 RESULTS AND DISCUSSION

Figures 8 shows a typical temperature map for Santotrac 50 at 60 °C, mean rolling speed 0.85 m/s and 100 % slide roll ratio. (In this paper slide roll ratio is defined as the ratio of sliding speed to mean rolling speed, so that pure sliding with the disc surface stationary has slide roll ratio=200 %). These conditions were selected to provide a central EHD film thickness of 100 nm, as measured using optical interferometry.

One problem encountered in mixed sliding/rolling work was that the bulk temperature of the ball rose during a test. This can be seen in figure 9, where the temperature along the sides of the grid rise from inlet to outlet. This is believed to be because the ball was driven by a thin shaft and supported by elliptical contact on rollers, so that, unlike in the pure sliding rig, the ball was thermally isolated. To compensate for this effect, the temperature values along the extreme sides of the map were subtracted from the internal values to flatten the overall profile prior to shear stress analysis.

Figure 9 shows a flattened temperature rise map and figure 10 the calculated shear stress map from the date in figure 9. The shape of the profiles of temperature rise across this map from inlet to outlet were similar to those obtained from microtransducer measurements [36-38].
Figure 10: Calculated shear stress map from temperature rise map shown in figure 9

Figure 11 shows a shear stress map from a test similar to that shown in figure 10 but at 80% slide roll ratio. Here the map was obtained to include more of the contact exit. This shows clearly a feature of all the shear stress maps obtained - the negative shear stress region at the rear of the contact. This signifies that the surfaces are cooling down more rapidly in the exit region than expected simply by conduction of heat from the solid surfaces into their bulk.

Two possible origins for this were considered, (i) evaporation of fluid in the negative pressure exit region and (ii) compression cooling. Both of these would remove heat from the surfaces into the fluid film, (negative local \( \dot{q} \)). The first was discounted, since assuming an appropriate heat of evaporation and that all the liquid passing through the contact evaporates, predicted a negligible \( \dot{q} \). The second was compression cooling. The heat generated by compression in a column of fluid of film thickness \( h \) is given by [39].

\[
\dot{q} = \varepsilon T h \frac{dp}{dx}
\]  

(5)

where \( \varepsilon \) is the coefficient of thermal expansivity of the oil, \( T \) the temperature (degrees Kelvin), \( U \) the mean rolling speed, \( h \) the oil film thickness and \( dp/dx \) is the pressure gradient in the rolling/sliding direction. For typical values of \( \varepsilon = 7.5 \times 10^{-4} \text{ K}^{-1} \), \( T=300 \text{ K} \), \( U=1 \text{ m s}^{-1} \) and \( h=10^{-7} \text{ m} \), this reduces to:

\[
\dot{q} = 2.25 \times 10^{-8} \frac{dp}{dx}
\]  

(6)

Thus the critical value in determining the significance of compression heating/cooling is the pressure gradient \( dp/dx \). In the inlet this is relatively small (e.g. 1 GPa over 100 \( \mu \)m, i.e. \( 10^{13} \text{ Pa/m} \)), suggesting that the compression heat generated per unit area in this zone is \( \dot{q} \approx 2.5 \text{ MJm}^{-2} \). However in the contact exit, both computational studies and microtransducer experiments e.g. [38], suggest that the pressure falls very sharply, over less than 10 \( \mu \)m. This implies that \( dp/dx \approx -10^{14} \text{ Pa/m} \) so that \( \dot{q} \) becomes \(-25 \text{ MJm}^{-2} \), comparable in magnitude to the heat generation per unit area from shear in the contact, albeit over a smaller area. Thus compression cooling in the contact exit may produce the apparently negative shear stresses seen.

8 CONCLUSIONS

This paper discusses the various ways of studying the rheology of lubricants in elastohydrodynamic conditions and identifies the temperature-rise mapping method as being particularly promising.

Some preliminary results for experimentally-determined temperature rise and shear stress maps in mixed rolling/sliding elastohydrodynamic contacts are shown. One practical problem encountered with such measurements is that the steel ball used to form the high pressure contact heated up significantly during acquisition of the required surface temperature rise map. This can be compensated for by subtracting the temperature measured at the contact peripheries where there is negligible flash temperature rise.

9 REFERENCES


