Analysis of Mixed Lubrication Effects in Simulated Gear Tooth Contacts

M. J. A. Holmes

H. P. Evans

R. W. Snidle

Cardiff School of Engineering,
Cardiff CF24 0YF, UK

The paper presents results obtained using a transient analysis technique for point contact elastohydrodynamic lubrication (EHL) problems based on a formulation that couples the elastic and hydrodynamic equations. Results are presented for transverse ground surfaces in elliptical point contact that show severe film thinning and asperity contact at the transverse limits of the contact area. This thinning is caused by transverse leakage of the lubricant from the contact in the remaining deep valley features between the surfaces. A comparison is also made between the point contact results on the entrainment center line and the equivalent line contact analysis. The extent of asperity contact is shown to be dependent on the Hertzian contact aspect ratio. It is also shown that transverse waviness (superimposed on the roughness) of even relatively small amplitude can lead to large increases in asperity contact rates over all waviness peaks in the contact.
[DOI: 10.1115/1.1828452]

Introduction

The lubrication mechanism primarily responsible for the protection of gear tooth surfaces from wear and surface distress is elastohydrodynamic lubrication (EHL). In the case of very smooth surfaces (such as those found in rolling element bearings, for example) the EHL mechanism can generate oil films which are thick compared to the height of roughness features present on the surfaces. Under these conditions the thickness of the oil film may be calculated, with reasonable accuracy, using the well-known formula of Dowson and Higginson [1]. An important feature of gear tooth contacts, however, is that the surfaces produced by present day manufacturing methods have roughness features that are significantly greater than the oil film predicted by this formula. Consequently gears tend to operate in a regime described as ''mixed'' or ''micro'' EHL in which there is a significant interaction of roughness asperities on the two surfaces. In theoretical solutions of both the dry and micro EHL situations the presence of roughness leads to significant rippling of the contact pressure distribution with maximum values far in excess of the Hertzian values expected when the surfaces are perfectly smooth. Micro EHL solutions also indicate the presence of very thin films at asperity encounters within the overall rolling/sliding contact. Two practical problems associated with roughness effects and film thinning in gears are micropitting (rolling contact fatigue on the scale of surface asperities) and scuffing (scoring) which is related to the failure of the elastohydrodynamic system. In order to gain a much clearer understanding of these failure mechanisms it is necessary to develop a full theoretical model of lubrication of gear contacts under rough surface/thin film conditions. Such a model must take account of the real operating conditions of gears in terms of loads, speeds, surface roughness and lubricant properties, and be able to predict pressures, local film thickness, temperatures and friction between the teeth. A further important feature that must be considered when roughness is present is the time-dependent effect of roughness: this occurs when roughness features move relative to the overall contact.

The paper presents results from the numerical modeling of transient rough surface point contact problems obtained using a new coupled numerical formulation for solving the elastohydrodynamic (EHL) point contact problem described and validated elsewhere [2,3]. The paper focuses attention on the strong side leakage effects that take place at the edges of contacts that have a surface finish transverse to the direction of rolling/sliding such as that in conventional involute gears. The configuration chosen for analysis is that of a gear simulation disk rig which gives rise to an elliptical contact. In the smooth surface case the EHL contact adopts a self-sealing configuration by developing side constrictions in the form of the familiar horseshoe shape seen in optical interferometry experiments. When transverse roughness features are present, however, this mechanism is unable to seal the pressure-driven transverse flow in the valley features because the closest that the surfaces can be brought together is determined by the physical contact of asperity tip features. Even in this extreme configuration the composite valley features on the surfaces remain open and unsealed, and lubricant can easily escape from the contact area in the transverse direction along these valleys. When oil is lost from the contact due to this sideways leakage mechanism the entrainment of lubricant under the downstream micro contacts is progressively weakened at each successive following contact. This model for EHL failure was proposed earlier by the authors [4] in response to their experimental observation that initial scuffing failure invariably occurs at the transverse edges of such contacts [5]. The detailed results of micro-EHL analysis presented here add further evidence in support of the model. It is of interest to note that the transverse edges, where failure appears to originate in the scuffing experiments, are not subject to extreme temperature or extreme pressure behavior. The identification of contact edges as the location of initial scuffing failure is thus a significant observation indicating that failure of the physical mechanism of EHL is a primary underlying cause of scuffing in gear tooth contacts. In real gears having a finite facewidth effective ''contact edges'' which behave in the way suggested are not limited to the actual face edges of the gears because of the inevitable ''waviness'' present on the surfaces. Results presented in the paper suggest that such waviness, of even small amplitude, can lead to a significantly increased occurrence of film breakdown.

Contact Analysis

The EHL problem is specified by the elastic deflection equation written in differential form [2],

$$\frac{\partial^2 h(x_i,y_j)}{\partial x^2}+\frac{\partial^2 h(x_i,y_j)}{\partial y^2}=\frac{\partial^2 \phi(x_i,y_j)}{\partial x^2}+\frac{\partial^2 \phi(x_i,y_j)}{\partial y^2}+\frac{1}{R_x}+\frac{1}{R_y}$$

$$+\frac{2}{\pi E'}\sum_{all\ k,l}f_{k-i,l-j}p_{k,l} \qquad (1)$$

and the non-Newtonian Reynolds equation

$$\frac{\partial}{\partial x}\left(\sigma_x\frac{\partial p}{\partial x}\right)+\frac{\partial}{\partial y}\left(\sigma_y\frac{\partial p}{\partial y}\right)-\frac{\partial(\rho\bar{U}h)}{\partial x}-\frac{\partial(\rho h)}{\partial t}=0 \qquad (2)$$

The nonlinear dependence of viscosity and density on pressure are taken to be given by the well known isothermal Roelands, and Dowson and Higginson relationships, respectively,

$$\eta=\eta_0\exp\{\ln(\eta_0/\kappa)((1+\chi p)^Z-1)\} \qquad (3)$$

$$\rho=\rho_0\left\{\frac{1+\gamma p}{1+\lambda p}\right\} \qquad (4)$$

For non-Newtonian situations σ_x and $\sigma_y(\neq\sigma_x)$ are determined from the lubricant's pressure, pressure gradients, film thickness, and surface velocities as discussed in [6]. Equation (2) is discretized using linear quadrilateral finite elements with an implicit (Crank–Nicolson) time formulation, and Eq. (1) using finite differences for the Laplacian with the pressure coefficients, $f_{i,j}$, given by the analysis in [7]. The equations are thus written as

$$\sum_{k=0}^{n_c}A_kp_k+\sum_{k=0}^{n_c}B_kh_k=R_{i,j} \qquad (5)$$

$$\sum_{k=0}^{n_c}C_kp_k+\sum_{k=0}^{n_c}D_kh_k=E_{i,j} \qquad (6)$$

where suffix k represents the nodes contributing to the assembled equation at node (i,j) and $k=0$ denotes that node. A_k and B_k are the pressure and film variable coefficients for the Reynolds equation (2), and n_c is the number of neighboring nodes involved in the formulation. Similarly C_k and D_k are the pressure and film variable coefficients for the differential deflection equation (1). Expression $R_{i,j}$ contains information from the previous timestep, and $E_{i,j}$ contains all the contributions to the pressure summation of Eq. (1) that are not explicitly contained on the left hand side of Eq. (6).

Equations (5) and (6) are expressed as a pair of simultaneous equations in the variables p_0 and h_0

$$A_0p_0+B_0h_0=\hat{R}_{i,j}\left\{=R_{i,j}-\sum_{k=1}^{n_c}A_kp_k-\sum_{k=1}^{n_c}B_kh_k\right\}$$

$$C_0p_0+D_0h_0=\hat{E}_{i,j}\left\{=E_{i,j}-\sum_{k=1}^{n_c}C_kp_k-\sum_{k=1}^{n_c}D_kh_k\right\}$$

which are solved to give a coupled iterative scheme to update the values of all the unknown nodal values of p and h in turn, i.e.,

$$p_{i,j}^{new}=\frac{\hat{R}_{i,j}D_0-\hat{E}_{i,j}B_0}{A_0D_0-B_0C_0},\quad h_{i,j}^{new}=\frac{\hat{E}_{i,j}A_0-\hat{R}_{i,j}C_0}{A_0D_0-B_0C_0} \qquad (7)$$

This method is found to be both effective in obtaining rapid convergence in low Λ situations with rough surfaces, and extremely robust. The rapidity with which the influence coefficients $f_{i,j}$ in Eq. (1) fall, as the indices i and j increase from zero [7], is a key advantage of this differential formulation of the deflection equation. This property allows the recalculation of pressure contributions to $E_{i,j}$ to be limited to those that are close to the n_c points used in the iteration sweep [2,8].

During rough surface transient analyses fluid film breakdown can occur resulting in contact between the micro asperities. Where contact occurs between the two surfaces the hydrodynamic film thickness is zero, although in practice there will typically be a
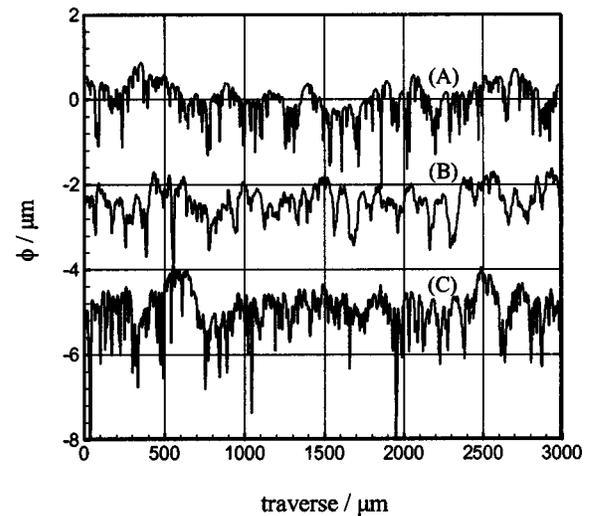
Fig. 1 Profiles adopted for the surfaces used in the numerical investigation. Profiles are offset for clarity and oriented with metal below the curves.

boundary film that controls the local friction coefficient. Equation (2) arises from mass flow continuity of the fluid film, and at locations where the film thickness is zero there is no such mass flow. The physical principle on which Eq. (2) is founded is thus not applicable at micro contact locations. Equation (1), however, is always applicable as it relates the pressure acting on the surfaces with their deflection irrespective of whether the pressure arises from a hydrodynamic film or from direct contact of the surfaces.

Contact situations in the iterative scheme are dealt with as follows. If the iterating equations (7) result in a negative value for $h_{i,j}^{new}$, its value is set to zero and Eqs. (7) are thus replaced by

$$h_{i,j}^{new}=0,\quad p_{i,j}^{new}=\frac{\hat{E}_{i,j}}{C_0} \qquad (8)$$

This effectively replaces the Reynolds equation with the boundary condition $h=0$, and applies the deflection equation subject to that boundary condition.

Thus Eqs. (7) are used at each mesh point during each iterative sweep and are replaced by Eqs. (8) only at mesh points where the current evaluation of Eqs. (7) yield a negative value for $h_{i,j}^{new}$. The ease with which contact conditions can be incorporated using this approach is a further advantage of the coupled differential deflection technique.

This method can be used to solve the dry, elastic contact problem using Eq. (2) as can be seen from the results for contact start-up analysis presented in [9]. This problem involves a simultaneous solution of full film and dry contact areas as liquid is first entrained into a dry contact by motion of the surfaces. The iterative approach described above deals effectively with this situation, maintaining a dry contact pressure that remains essentially Hertzian away from the area where the contact shape is distorted by the entrained fluid. The comparison made in [9] with the elegant experimental work of Glovnea and Spikes [10] for this situation provides confirmation of the validity of the approach adopted.

Results

Behavior at the Transverse Edge of the Contact. The results presented in the current paper are based on an isothermal analysis of elliptical contacts finished in a transverse direction. In each case the contact is between two ellipsoidal bodies whose surface finish is given by one of the three experimental profiles shown in Fig. 1 (solid metal below the profile). Trace (A) is a

Table 1 Operating conditions for the point and line contact comparisons

	Point contact	Line contact
R_x	0.0191 m	0.0191 m
R_y	0.151 m	$\propto$
w, w'	962 N	527 kN/m
a	0.335 mm	0.335 mm
b	1.31 mm	$\propto$
p_{hz}	1.05 GPa	1.0 GPa
E'	227 GPa	
α	11.1 GPa^{-1}	
χ	5.1 GPa^{-1}	
γ	2.27 GPa^{-1}	
η_0	0.005 Pa·s	0.0048 Pa·s
κ	63.2$\times$10^{-6} Pa·s	
λ	1.68 GPa^{-1}	
$\bar{U}$	25 m/s	
ξ	0.1, 0.25, 0.5	
τ_0	10 MPa	

Fig. 2 Pressure/GPa (upper figure) and film thickness/μm contours of the smooth surface result for the conditions analyzed. Central and minimum film thicknesses are 0.48 and 0.42μm, respectively. The heavy curve indicates a Hertzian dry contact area.

profile taken from a well run-in transverse ground disk used in scuffing experiments by Patching [5]. Traces (B) and (C) are taken from micropitting tests on gears and have been run for several load stages and as a result have become run-in to some extent, but a close examination shows that they clearly have larger asperities than profile (A). Traces (A), (B) and (C) have R_a values of 0.32 μm, 0.22 μm and 0.31μm, respectively. Intermediate heights that are required as the surfaces move through the contact are obtained using cubic spline interpolation, which ensures slope continuity at the measured points. A comparison of the line contact behavior of these profiles has been undertaken previously in [11] where it was clearly seen that Profile (C) was the most aggressive of the three profiles in terms of its tendency to produce high pressure ripples and severe film thinning.

The lubricant modeled is Mobil Jet 2, a synthetic gas turbine lubricant used in earlier scuffing experiments [5]. The operating conditions and lubricant parameters adopted are specified in Table 1 and result in a contact whose Hertzian dimension in the transverse direction is four times that in the rolling/sliding direction.
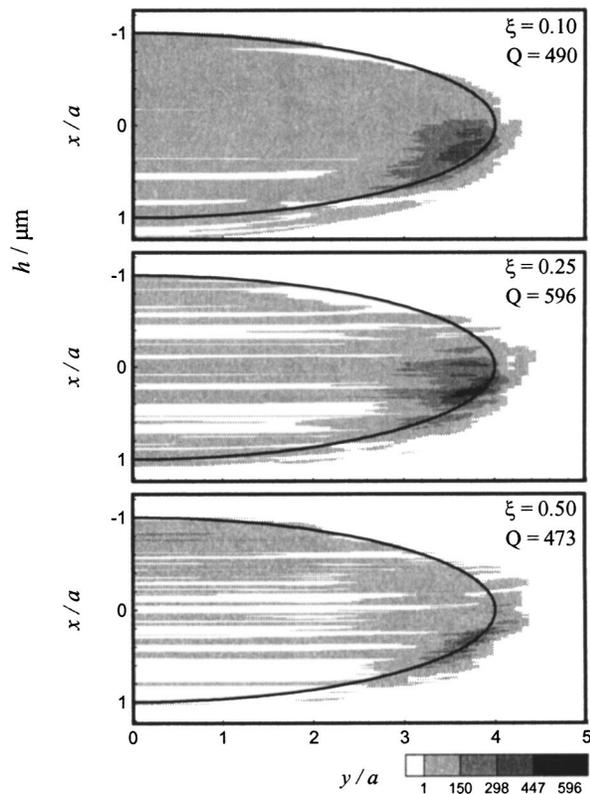
The computing mesh covers the area $-2.5a < x < 1.5a$; $-2b < y < 2b$, with mesh spacing $\Delta x = a/200$; $\Delta y = b/50$. The timestep adopted was $\Delta t = \Delta x/2u_{max}$ so that the faster moving surface moves through one mesh spacing over two timesteps. The transient analysis is started from the smooth steady state result, shown in Fig. 2 which illustrates the pressure and film thickness contours for the operating conditions chosen for analysis. The Moes and Bosma dimensionless groups for the contact are $M = 270$ and $L = 6$. These conditions could be expected to generate an appreciable pressure spike with a Newtonian analysis, but this is diminished into the rudimentary shoulder feature seen in non-Newtonian circumstances; this can be discerned in the very closely spaced (monotonic) pressure contours near the exit of the Hertzian contact area. The maximum pressure developed is 1.03 GPa which is very close to the corresponding Hertzian contact value of 1.05 GPa. The central film thickness "plateau" value is 0.48 μm with a minimum value on the longitudinal center line of 0.42 μm, and transverse edge constrictions where a similar minimum film value of 0.43 μm is developed. When the steady state solution has been established the rough surface features, which make the problem time dependent, are fed in with the moving surfaces from the inlet boundary position. Because of the different speeds of the two surfaces the time taken for both surfaces in the contact to become fully rough is that required for the slowest of the two surfaces to move from the inlet boundary to the exit boundary. Once this has occurred the computation is carried on for a further 2370 timesteps, i.e. further analysis times of 0.076 ms, 0.070 ms and 0.063 ms, respectively, for the three slide/sliding ratios
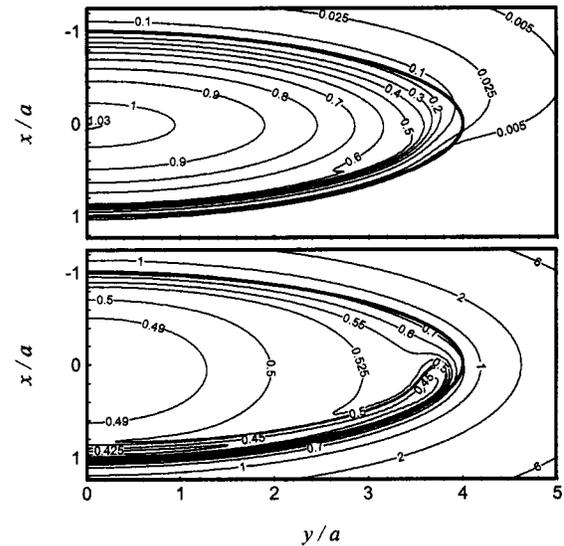
adopted of 0.1, 0.25 and 0.5. The distance moved by the faster moving surface during this further calculation is thus the same in each case, and is equal to $5.9a$.

The presence of roughness on both surfaces causes a significant variation in both pressure and film thickness within the contact area, and Fig. 3 illustrates this effect for one particular timestep in the analysis for two rough surfaces having profile (C) with a slide/roll ratio of $\xi = 0.25$. (This profile was shown in [11] to have the most aggressive EHL response in rough on rough line contact of the three profiles considered.) The figure shows the pressure and film thickness along the entrainment centreline, $y = 0$, and also shows the orientation of the two rough surfaces offset below for clarity. The deviations of pressure from the smooth surface result for this example can be seen to be significant: maximum pressures
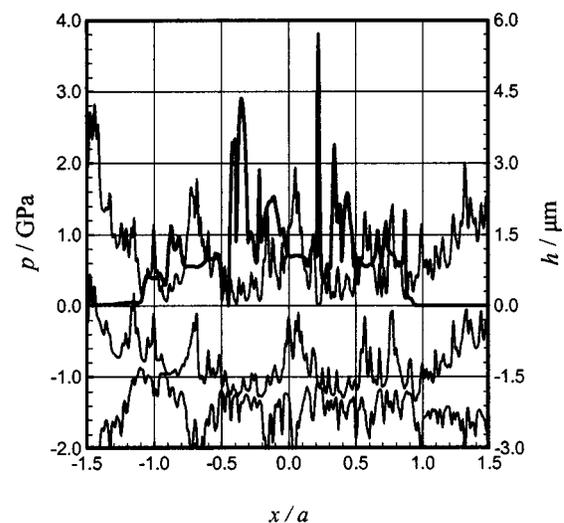
Fig. 3 Pressure (heavy curve) and film thickness on the entrainment axis, $y = 0$, at one timestep in the analysis of contact between two surfaces having Profile (C) with $\xi = 0.25$. Also shown are the two rough surfaces in their contact configuration offset for clarity.

Journal of Tribology

Fig. 4 Contours of contact count rate Q/ms for the transient analysis of two surfaces each having Profile (C). The heavy curve indicates a Hertzian dry contact area.

Table 2 Percentage of transient analysis time for which contact is calculated to occur at one or more mesh points

Surfaces	$\xi = 0.1$	$\xi = 0.25$	$\xi = 0.5$
A and A	38	26	34
B and B	19	20	4.3
C and C	98	95	86
A and C	96	77	55
C and smooth		9.2	
$\sqrt{2} \times$ C and smooth		48	

of the order of 3 GPa can be seen to develop on some asperity/asperity contacts with corresponding extremely small film thickness values. As the analysis progresses contact (as defined in the previous section) occurs occasionally as asperities move past each other. The number of timesteps in which contact occurs is calculated for each mesh point during the fully rough analysis period, and we refer to this as the contact count, denoted Q. Figure 4 shows contours of the contact count for the case where both surfaces have Profile (C) for the three different slide/roll ratios considered. To facilitate comparisons between cases having different sliding speeds, and thus different timesteps, the values of Q obtained are normalized with respect to the total analysis time. High contact count values indicate repeated contact instances at that particular location as the rough surfaces pass through the corresponding smooth surface Hertzian contact area indicated by the semi ellipse. It is clear from the contour values that contact occurs predominantly at the transverse edges of the Hertzian contact area downstream $(x>0)$ of the contact centerline. It is worth noting that the profiles used for the analysis are taken from experimental test disks and that the surface finish has been modified by the action of plastic deformation as the contact has run from its as-manufactured surface finish to the current state. As-manufactured surfaces are considerably rougher, and when finished by grinding have a near Gaussian distribution of surface heights. The height distribution of the current profiles, shown in Fig. 1, have a degree of skewness introduced by the running-in process, and the asperity tips are more rounded as a result. This is the surface configuration that corresponds to contact failure, and as such its EHL behavior is likely to be of more engineering relevance than the more computationally challenging freshly manufactured finish.

In Fig. 4 we see that calculated contact is a relatively frequent occurrence at the transverse margins of the Hertzian area. The mesh point for which contact occurs most frequently experiences contact in 42 of the 2370 timesteps for which the count is carried

out, i.e., about 2% of the calculation time. This is for the intermediate case where $\xi = 0.25$. The number of timesteps for which contact occurs at one or more mesh points during the calculation is, however, a high proportion of the total. Contact occurs for 98%, 95% and 86% of the timesteps for the three analyses as shown in Table 2.

There is little difference between the contact count pattern obtained for the three slide/roll ratios. The highest occurrence is with $\xi = 0.25$, and the area experiencing the higher rates of asperity contact is more pronounced for this and the lower sliding case of $\xi = 0.1$. The higher sliding speed of $\xi = 0.5$ seems to lead to a reduced rate of asperity contact, and with the area over which high asperity contact rates occur also reduced in comparison with the other cases. This feature of the results follows from the fact that the entrainment effect for asperity/asperity collisions within the Hertzian region is effectively given by 0.5 times the sliding velocity as was demonstrated for line contacts in [12]. Thus the asperity/asperity entrainment effect for $\xi = 0.5$ is five times higher than for $\xi = 0.1$. This observation would seem to suggest that higher sliding may be advantageous in preventing asperity/asperity contact. However, we hasten to add that the analysis presented here is isothermal so that the detrimental effects of localized heat generation due to thin film or dry contact sliding between the asperities are not included at this stage.

Figure 5 shows contours of contact count per unit time for the case of contact between two surfaces having Profile (B). The contact behavior of these surfaces is quite different as can be seen by a comparison with Fig. 4. The mesh points experiencing the highest contact counts in Fig. 5 are seen to be located in a very limited area around the boundary of the Hertzian contact region. The contact counts at these locations are about 25% of the peak values seen in Fig. 4. In addition there are almost no contact occurrences in the remainder of the Hertzian area. This is in marked contrast to Figure 4 where contact is seen to occur over the whole width of the Hertzian area, which suggests strongly that contact occurs across the whole Hertzian contact area with some particular asperity collisions. There are bands in Fig. 4 where no contact has occurred but these probably result from a lack of asperity collisions at these locations during the analysis time.

Figure 6 shows contours of contact count per unit time for the case of contact between two surfaces having Profile (A). The peak contact count level is similar for each of the sliding speeds, but this case illustrates the strong effect of the asperity/asperity entrainment due to the sliding velocity. For the highest sliding speed the contact count rate is close to zero over most of the Hertzian region and high values are concentrated at the transverse contact boundary. For the lower sliding speed contact conditions also occur on the centerline and in bands over the exit half of the contact.

Figure 7 shows contact count contours for a contact consisting of Profile (A) running against Profile (C). Comparing this figure with Figs. 4 and 6 shows contact incidences that are intermediate between those of the individual surfaces in contact with themselves. Contact between these two surfaces takes place approximately half as frequently as that between two surfaces having profile (C), and with a similar pattern of contact intensity/location.

Figure 8 examines the contact count obtained for Profile (C) for $\xi = 0.25$ in three different cases. In Fig. 8(a) the surface is in contact with a smooth surface. In Fig. 8(b) it is again in contact
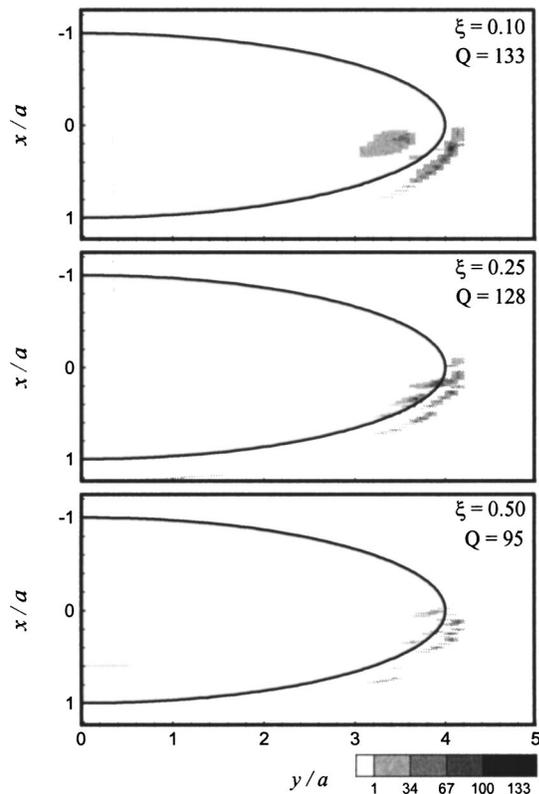
Fig. 5 Contours of contact count rate Q/ms for the transient analysis of two surfaces each having Profile (B). The heavy curve indicates a Hertzian dry contact area.
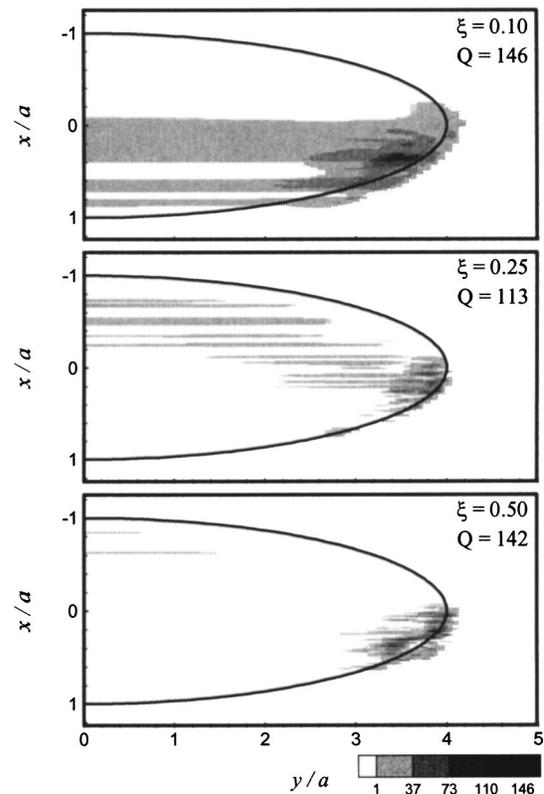
Fig. 6 Contours of contact count rate Q/ms for the transient analysis of two surfaces each having Profile (A). The heavy curve indicates a Hertzian dry contact area.

with a smooth surface, but the scale of the roughness of Profile (C) has been increased by a factor of $\sqrt{2}$. This is so that the composite roughness is of the same order as that for contact between two surfaces having Profile (C), which is the case illustrated in Fig. 8(c). In Fig. 8(a) contact hardly takes place and the little that does occur is limited to the transverse contact boundary. In Fig. 8(b) there is a six-fold increase in the contact count due to the higher roughness and this is again concentrated at the transverse contact boundary. For the case of the two rough surfaces, the maximum contact count at the transverse boundary has increased by a further factor of three, and bands of contact occurrence spread across the entire contact width. Although the composite roughness in the cases shown in Figs. 8(b) and (c) are the same, the relative radius of curvature at the tips of the most aggressive asperities is smaller in the case of Fig. 8(b) than in the case of Fig. 8(c) where they will generally be in contact with surface features having a larger radius of curvature. However contact is less prevalent in Fig. 8(b) than in Fig. 8(c) which supports the view that asperity collision is an important factor in causing contact to occur. This feature of the results, and the percentage contact times given in Table 2, point towards the possibility of experimental verification of the predicted contact effects by measuring fractional contact time using electrical contact resistance. Experiments to investigate this effect are planned for future work.

Figure 9 shows a photograph of part of the surface of a test disk taken from the scuffing program reported in [5]. The disk is from an experiment where the contact load was removed at the first indication of scuffing. The disks are crowned and the Hertzian contact area is illustrated by the ellipse superimposed on the photograph. Grinding marks can be clearly seen extending across the width of the disk but the surface finish has been totally changed in the scuffed part of the running track. The width of the scuffing mark is about 25% that of the running track, and its outer edge corresponds to the transverse limit of the Hertzian contact area.

This pattern has been observed in all the scuffing experiments utilizing transverse-ground crowned disks. (The width of the scuffing mark is dependent on the rapidity with which load is removed when scuffing is detected by its characteristic sudden increase in friction.) We suggest that the correspondence between the position of the scuffing track in Fig. 9 and the location of predicted high contact counts typically shown in Figs. 4 to 7 is striking. This indicates strongly that the primary cause of scuffing in these experiments was the breakdown of the EHL film as a result of direct contact between the surfaces.

Behavior in the Center of the Contact. Although the results given above concentrate on the features of mixed lubrication situations brought about by side flow at the transverse edges of the elliptical contact, it is interesting to compare the center line ($y = 0$) behavior with that of the corresponding line contacts. Figure 10 shows one such example of the film thickness and pressure distribution at a particular timestep. The figure includes both the elliptical contact and equivalent line contact results at the same timestep. It can be seen that the film thickness behavior is identical between the two methods and the minor differences in pressure are no greater than inevitably exist in the comparison of these equivalent smooth surface solutions. This equivalence is found to be generally the case [3] for low Λ conditions with contacts of this aspect ratio. This gives a clear demonstration that line contact transient analyses are a suitable tool for investigating micropitting and scuffing in involute gears, failure occurrences which are not limited to the edges of the gear face width.

The aspect ratio of the contacts considered up to this point are 4:1, i.e. $a/b = 0.25$. For elliptical contacts that have more adverse aspect ratios the proximity of the transverse boundaries exerts a greater influence over the main part of the contact area as might be expected. Figure 11 compares the contact count contours for the case of two rough surfaces each having profile C with
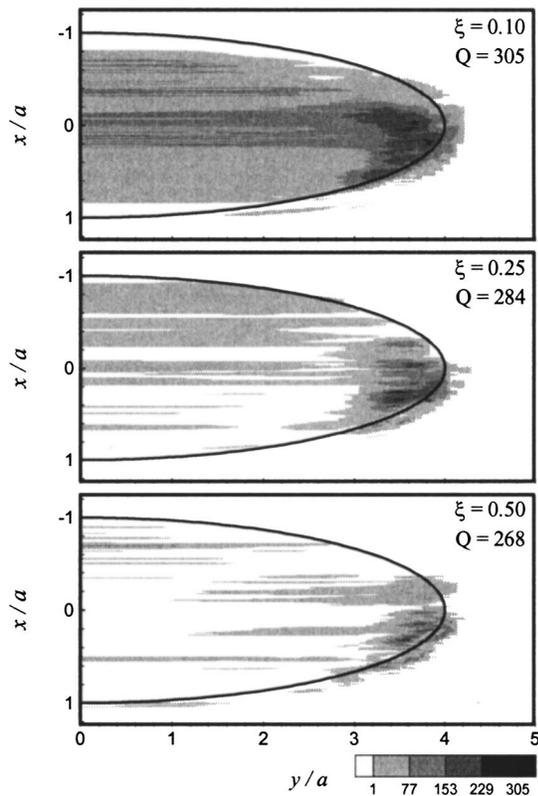
Fig. 7 Contours of contact count rate Q/ms for the transient analysis of two surfaces one having Profile (A) and one with Profile (C). The heavy curve indicates a Hertzian dry contact area.

Fig. 8 Contours of contact count rate Q/ms for the transient analysis of a surface having Profile (C) running against (*a*) a smooth surface, (*b*) a smooth surface with the roughness scaled by a factor of $\sqrt{2}$, (*c*) another surface having Profile (C). The heavy curve indicates a Hertzian dry contact area.

$\xi = 0.25$. The three results shown are for contacts with different aspect ratios of $a/b = 0.25$, 1 and 4. Each contour map has been drawn with the y co-ordinate scaled so that the corresponding Hertzian contact ellipse appears as a circle, and the value of Hertzian dimension a (in the entraining direction) is the same in each case. The proximity of the side boundary is seen to have a significant effect on the degree of calculated contact. The effect of the change in aspect ratio is also significant on the contact center line. Figure 12 compares the film thickness behavior on the entrainment center line at one particular timestep for the three aspect ratios. The rough surfaces are in exactly the same position relative to each other and to the contact point at the timestep illustrated. For $a/b = 0.25$ the surfaces are completely separated by the oil film at this timestep, but as conditions become more adverse due to increased side leakage the surfaces can be seen to become more closely "enmeshed." Some contact instances occur for the circular contact ($a/b = 1$), and this trend increases as a/b increases to 4. This change is associated with a greater proportion of the load being carried at asperity "collisions" as can be seen from Fig. 13 which gives the corresponding comparison of the center line pressure profiles at the same timestep for the three configurations.

The way in which contact is caused by side-leakage of the lubricant is illustrated in detail by the pattern shown in Fig. 14. This figure shows film thickness contours for a region in which there are three asperity encounters near the transverse boundary of a contact. The Hertzian boundary is illustrated by the curve superimposed on the contours. Also included on the figure are lubricant flow vectors calculated at each of the mesh points in the figure. These vectors indicate the resultant direction of flow by their orientation and the magnitude of this flow by their relative size. It is clear that lubricant in the valley features escapes from the load bearing area by transverse motion rather than being forced by
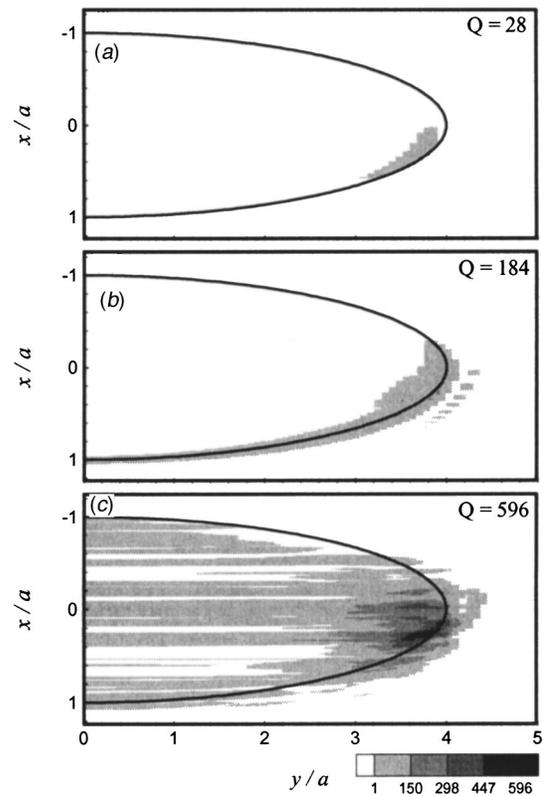
entrainment to maintain the separation of the surfaces. This effective escape of oil leads to direct contact on the asperity micro contact at $x/a = 0.308$ at the timestep shown.

The detrimental effect of transverse leakage is not confined to the extreme edge of the contact, but can also occur due to transverse waviness (i.e., 3D roughness) of the contacting components within the overall contact. This is illustrated in Fig. 15 which shows the effect of transverse waviness on the contact between
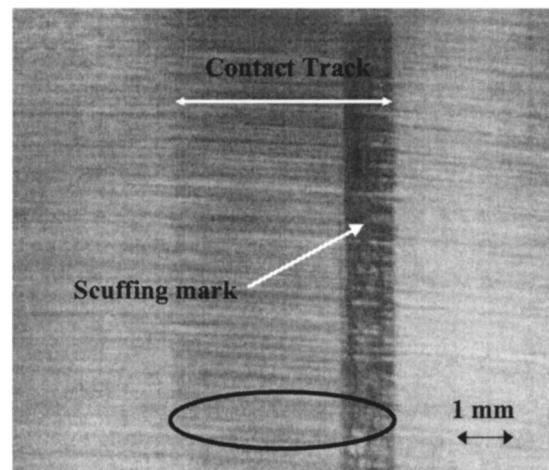
Fig. 9 A photograph of a test disk taken from the experiential scuffing program [5] showing a track subject to scuffing damage. Also shown is the Hertzian contact ellipse for the operating load at which scuffing occurred.
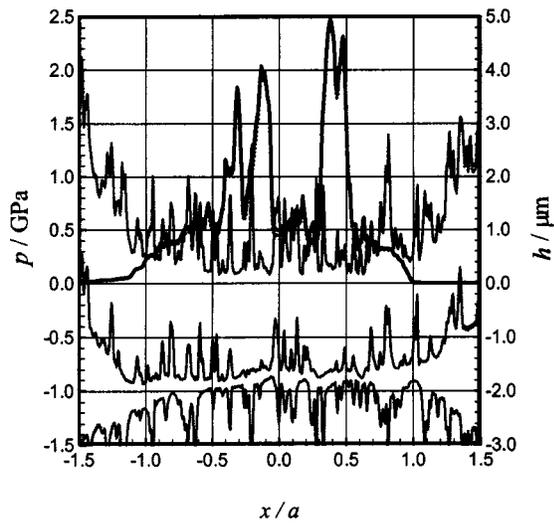
Fig. 10 A comparison of line and point contact center line pressure (heavy curve) and film thickness results which are superimposed in the figure for one timestep in the contact of two surfaces having Profile (A) with $\xi=0.25$. Also shown are the two rough surfaces in their contact configuration, offset for clarity. The line contact results are shown with dotted curves.

two surfaces having profile C at $\xi=0.25$. Different waviness amplitudes are considered and in each case the same waviness profile is applied to both the contacting components with the waviness peaks aligned. The contact aspect ratio is $a/b=0.25$ so that the corresponding result without waviness is shown in Fig. 4 with a maximum value of $Q=596$. The waviness has a wavelength of $0.32b$, i.e., 420 μm, and the effect of waviness amplitudes of 0.05, 0.1, 0.2, 0.3, 0.4 and 0.5 μm are compared in the figure. As the amplitude increases the area of high contact count becomes
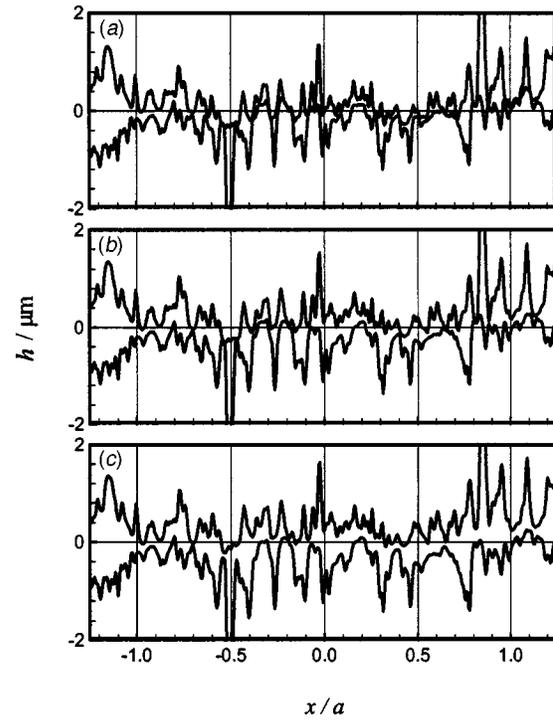

Fig. 12 A comparison of the rough surfaces in their deflected contact position at one timestep of the analysis leading to Fig. 11; (a) $a/b=4$, (b) $a/b=1$, (c) $a/b=0.25$

banded and the extent of contact becomes progressively more severe. For the case with amplitude 0.5 μm the waviness has overcome the cohesion of the contact which has degenerated into seven individual contacts with each subject to a similar degree of
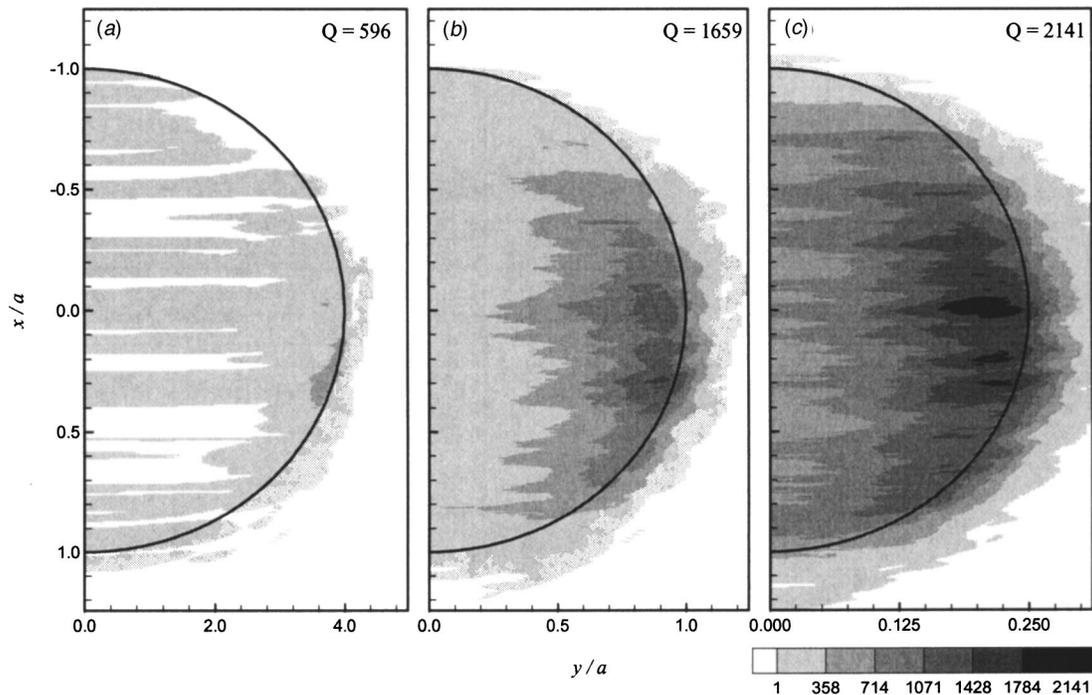

Fig. 11 Contours of contact count rate $Q/$ms for the transient analysis of two surfaces having Profile (C) with $\xi=0.25$ with aspect ratios (a) $a/b=0.25$, (b) $a/b=1$, (c) $a/b=4$. The heavy curve indicates a Hertzian dry contact area.
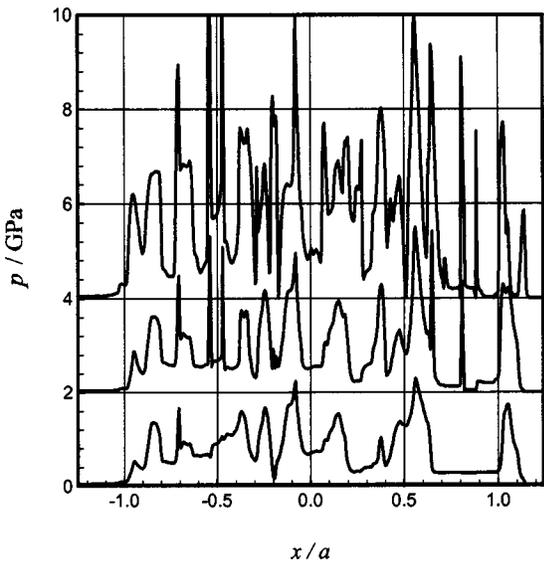
Fig. 13 A comparison of pressure distributions obtained at the same timestep as Fig. 12. The upper profile is for $a/b=4$ and is offset by 4 GPa for clarity; the middle profile is for $a/b=1$ and is offset by 2 GPa; the lower profile is for $a/b=0.25$.

asperity contact with Q values of up to five times higher than those seen at the most vulnerable location of the contact having no transverse waviness.

Discussion and Conclusions

Edge effects have been shown to be significant at the transverse margins of lubricated contacts between components having a ground surface finish transverse to the entrainment direction. The loss of film thickness in these instances is explained by transverse leakage in the valley features of the composite surface. In an earlier, simplified model [4] of film loss in the lubrication of surfaces having transverse roughness we assumed the geometry of rough surfaces in dry contact separated, at the inlet to the contact, by the film thickness calculated on the assumption of smooth surfaces. Steady state conditions were assumed corresponding to a smooth surface running against a stationary rough surface, and
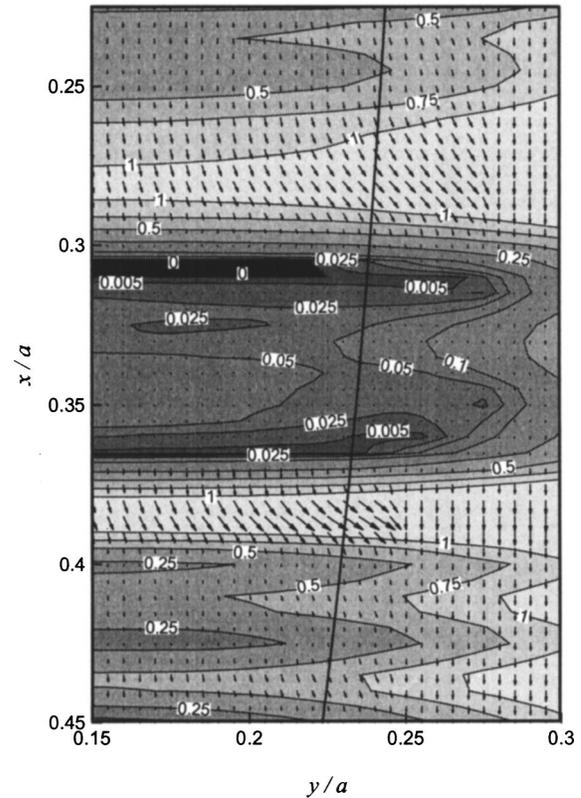

Fig. 14 Film thickness contours and calculated flow vectors for a region comprising three composite asperities near the transverse boundary of a contact. The Hertzian boundary is illustrated by the superimposed curve.

pressure generation in the valley features was determined from a simple line-contact isothermal analysis. The pressure in valleys was then assumed to decay in the transverse direction according to a prescribed semi-elliptical form and the pressure-driven sideways flow calculated for each individual valley. In this way the film thickness that separated the lands of the roughness at the inlet of the overall contact was calculated and seen to gradually reduce at
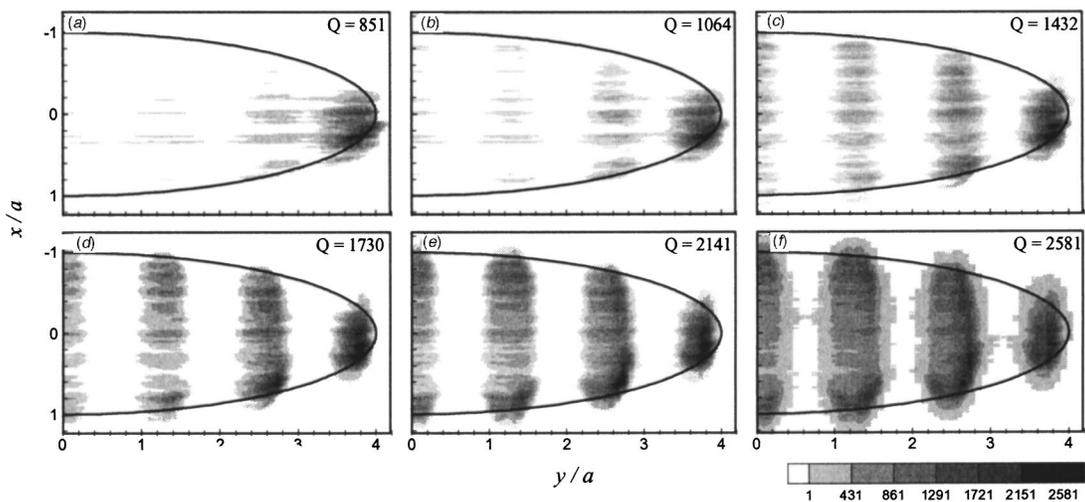

Fig. 15 Contours of contact count rate $Q/$ms for the transient analysis of two surfaces having Profile (C) with $\xi=0.25$, with a transverse waviness of wavelength $0.32b$ applied to each rough surface, and with waviness amplitudes of (a) 0.05 μm, (b) 0.1 μm, (c) 0.2 μm, (d) 0.3 μm, (e) 0.4 μm, (f) 0.5 μm. The heavy curve indicates a Hertzian dry contact area.

each successive valley. Assuming operating conditions under which scuffing had occurred in the corresponding disk machine experiment the analysis typically resulted in complete loss of the film at a longitudinal position somewhere between the center line of the overall contact and its exit. It may be noted that this behavior, in which film loss tends to be concentrated downstream of the contact center line, is also seen in the present work.

The full EHL analysis given here provides a much more detailed insight into film loss. Direct contact between the surfaces at the transverse boundaries of the conjunction is a feature of the solutions, and this is consistent with the location of initial scuffing failure in disk experiments with such surfaces [5]. The extent of contact is dependent on the sliding speed, and high sliding speeds enhance the entrainment mechanism during asperity/asperity collisions. Higher localized heating due to the sliding effect may, however, counterbalance this benefit of sliding that is seen in the current isothermal treatment. Considerations of thermal effects will be investigated in future modeling work.

Results of comparisons between point contact and line contact EHL solutions for low Λ situations show that a line contact analysis is sufficient to determine conditions of pressure and film thickness on the center line of the contact. By implication, line contact analyses are able to determine the film behavior over most of the contacting region in involute gear contacts.

Waviness (even of relatively low amplitude) transverse to the entrainment direction (3D roughness) is seen to give a significant increase in the frequency of EHL film breakdown and localized dry contact.

Acknowledgment

Nomenclature

a,b = Hertz contact dimension m

$A_k,B_k,C_k,D_kE_{i,j},R_{i,j}$ = coefficients in discretized equations (5) and (6)

E' = effective modulus of elasticity Pa

$f_{i,j}$ = pressure coefficient in differential deflection equation m^{-1}

h = film thickness m

L,M = Moes and Bosma nondimensional groups;
$L = \alpha E' (2\eta_0 \bar{U}/E' \bar{R})^{1/4}$,
$M = (w/E' \bar{R}^2, (E' \bar{R}/2\eta_0 \bar{U})^{3/4}$

n_c = number of neighboring mesh points in discretization

u_{max} = surface velocity of the fastest moving surface in the x direction m/s

$\bar{U}$ = mean surface velocity in x-axis direction m s^{-1}

p = pressure Pa

p_{Hz} = maximum pressure in Hertzian contact Pa

Q = number of contacts occurring at a mesh point s^{-1}

R_x,R_y = radii of relative curvature in axis directions m

w = load for point contact N

w' = load per unit length of line contact N/m

x,y = co-ordinates in contact plane m

x_i,y_j = co-ordinates of mesh point with suffices (i,j)

$\Delta x, \Delta y$ = mesh spacing in co-ordinate directions m

Δt = timestep s

α = pressure viscosity coefficient Pa^{-1}

χ = coefficient in viscosity equation (3) Pa^{-1}

γ, λ = coefficients in density equation (4) Pa^{-1}

η_0 = viscosity at ambient pressure Pa·s

κ = coefficient in viscosity equation (3) Pa·s

Λ = film thickness/composite surface roughness

ρ = density kg m^{-3}

σ_x, σ_y = flow coefficients in axis directions m s

τ_0 = non-Newtonian shear stress parameter Pa

ϕ = surface roughness feature m

ξ = slide/roll ratio

References

[1] Dowson, D., and Higginson, G. R., *Elastohydrodynamic Lubrication*, Pergamon, Oxford, 1966.
[2] Holmes, M. J. A., Hughes, T. G., Evans, H. P., and Snidle, R. W., 2003, "Transient Elastohydrodynamic Point Contact Analysis Using a New Coupled Differential Deflection Method: Part 1 Formulation and Validation," Proc. Inst. Mech. Eng., Part J: J. Eng. Tribol., **217**, pp. 289–303.
[3] Holmes, M. J. A., Hughes, T. G., Evans, H. P., and Snidle, R. W., 2003, "Transient Elastohydrodynamic Point Contact Analysis Using a New Coupled Differential Deflection Method: Part 2 Results," Proc. Inst. Mech. Eng., Part J: J. Eng. Tribol., **217**, pp. 305–321.
[4] Evans, H. P., and Snidle, R. W., 1996, "A Model For Elasto-Hydrodynamic Film Failure in Contacts Between Surfaces Having Transverse Finish," Trans. ASME, J. Tribol., **118**, pp. 847–857.
[5] Patching, M. J., Kweh, C. C., Evans, H. P., and Snidle, R. W., 1995, "Conditions For Scuffing Failure of Ground and Super-Finished Steel Disks at High Sliding Speeds Using a Gas Turbine Engine Oil," Trans. ASME, J. Tribol., **117**, pp. 482–489.
[6] Sharif, K. J., Kong, S., Evans, H. P., and Snidle, R. W., 2001, "Contact and Elastohydrodynamic Analysis of Worm Gears: Part 1 Theoretical Formulation," Proc. Inst. Mech. Eng., Part C: J. Mech. Eng. Sci., **215**, pp. 817–830.
[7] Evans, H. P., and Hughes, T. G., 2000, "Evaluation of Deflection in Semi-Infinite Bodies by a Differential Method," Proc. Inst. Mech. Eng., Part C: J. Mech. Eng. Sci., **214**, pp. 563–584.
[8] Holmes, M. J. A., *Transient Analysis of the Point Contact Elastohydrodynamic Lubrication Problem Using Coupled Solution Methods*, Ph.D. thesis, Cardiff University.
[9] Holmes, M. J. A., Evans, H. P., and Snidle, R. W., 2003, "Comparison of Transient EHL Calculations With Start-Up Experiments," *Proc. 29th Leeds–Lyon Tribology Symposium*, Elsevier, pp. 79–89.
[10] Glovnea, R. P., and Spikes, H. A., 2001, "Elastohydro-Dynamic Film Formation at the Start Up of the Motion," Proc. Inst. Mech. Eng., Part J: J. Eng. Tribol., **215**, pp. 125–138.
[11] Tao, J., Hughes, T. G., Evans, H. P., and Snidle, R. W., 2002, "Elastohydrodynamic Response of Transverse Ground Gear Teeth," *Proc. 28th Leeds–Lyon Symp. on Tribology*, Elsevier, Amsterdam.
[12] Elcoate, C. D., Hughes, T. G., Evans, H. P., and Snidle, R. W., 2001, "Transient Elastohydrodynamic Analysis Using a Novel Coupled Differential Deflection Method," Proc. Inst. Mech. Eng., Part J: J. Eng. Tribol., **215**, pp. 319–337.