Influence of surface topography on friction, film breakdown and running-in in the mixed lubrication regime

P M Lugt[1]*, **R W M Severt**[2], **J Fogelström**[3] and **J H Tripp**[2]
[1]SKF Engineering and Research Centre BV, Nieuwegein, The Netherlands
[2]University of Twente, Enschede, The Netherlands
[3]SKF Sverige AB, Development Centre, Göteborg, Sweden

Abstract: The influence of surface topography on the lubricant film build-up ability and the friction characteristics of potential rolling bearing surfaces has been investigated by experiments on two-disc rigs. Traction–friction torque measurements were made for a variety of surface combinations, together with measurement of the electrical resistance between the discs as an indication of surface separation. For all disc combinations, running-in of the surfaces under load at any slide–roll ratio led eventually to full film separation. Contrary to results reported in the literature, film breakdown did not always increase with slip but depended on certain aspects of the surface structure. Friction torque measurements in the mixed lubrication regime also confirmed that friction is not determined simply by an R_a value. By suitable modification of the surface topography, keeping R_a constant, friction can be varied by as much as 10 per cent.

Keywords: surface topography, roughness, engineering surfaces, running-in, lift-off, mixed lubrication

NOTATION

BG	ground
CrH	cross-honed
H	lubrication number (m)
p	average contact pressure (Pa)
P	isotropically machined
R_a	arithmetic average roughness (m)
SP	polished
TSP	polished
U	sum speed (m/s)
V	vibro finished
η	dynamic viscosity (Pa s)

1 INTRODUCTION

The most important function of bearings is to reduce friction and, to achieve this, roller bearing contacts ideally operate in the full lubricant film regime. In practice, however, for a variety of reasons they frequently operate in the mixed regime, because the surfaces are not sufficiently smooth, the load or temperature is too high, or the viscosity or speed is too low. Even where the *steady* operating conditions lie in the full film regime, contacts still pass *transiently* through the mixed regime at start-up or stop, or during impulse loading. In this regime, the contacting regions give rise to high local pressures, which may lead to noise, fatigue damage and high wear rates. Hence the study of surface topography and mixed lubrication is of great practical interest. The breakdown of separating films in the mixed lubrication regime is, of course, determined not only by the R_a values of the surfaces; it will be shown here that the full three-dimensional topography, the shape of the features, plays an important role. At low speeds, this separability of the surface asperities influences the friction in bearing contacts. Thus, friction can be controlled through choice of surface structure.

1.1 Background

As early as 1978, Patir and Cheng [**1**] calculated the film thickness in lubricated contacts and showed that a transversely oriented roughness was superior to a longitudinally oriented roughness in the forming of a lubricant film. Using advanced numerical techniques it later became possible to include elastic deformation of the roughness and to make elastohydrodynamic lubrication (EHL) calculations where roughness was modelled as sinusoidal waves,

The MS was received on 29 September 2000 and was accepted after revision for publication on 19 March 2001.
**Corresponding author: SKF Engineering and Research Centre BV, PO Box 2350, 3430 DT Nieuwegein, The Netherlands.*

either transverse or longitudinal (see, for example, Venner and Lubrecht [**2**]). In reference [**2**] it was shown that film breakdown was most probably to be found near the side lobes caused by pressure-induced flow. Earlier Venner [**3**] showed that in the case of slip a phase shift occurs between pressure and film thickness, as a result of which the minimum film thickness no longer coincides with the (local) high-pressure region. Thus, whereas in the case of pure rolling the solid-like behaviour of the lubricant associated with high pressure occurs where the film is thinnest, this may not be true in the case of slip.

From electrical resistance measurements, Johnson and Higginson [**4**] and Ishibashi and Sonoda [**5**] demonstrated a reduction in film thickness of the micro-EHL film under sliding conditions. Jacobson [**6**] took a non-Newtonian lubricant which, if sliding occurs in the contact, quickly approaches its limiting shear. The shear strength of the lubricant will then be reached mainly by its shearing in the sliding direction, leaving little shear resistance in the perpendicular direction. Thus a transverse pressure gradient, according to the conventional description, can easily displace the lubricant sideways, leading to film collapse. Hamer *et al.* [**7**] modelled film collapse by simultaneously solving the plastic extrusion equations and the elastic pressure equations for the film trapped between approaching asperities. They showed that the speed of collapse is sensitive to surface asperity properties, such as mean spacing, a result already derived by Jacobson [**6**].

Of course, the choice of surface finish in rolling element bearings will be based not only on lubrication criteria. Generally, bearing surfaces have to fulfil several other functions as well, e.g. resistance to contact fatigue. From dry contact calculations, Tripp and Ioannides [**8**] showed how fatigue life is influenced by surface slope and skewness. Ultimately, the selection of surface topography becomes a problem of optimization between the various functional requirements.

Frequently it is the operation of the bearing itself which effectively provides the final finishing process. This so-called *running-in* phenomenon occurs when the bearing runs under severe conditions, or during starting and stopping, so that the elements are not fully separated by a lubricant film. Plastic deformation or mild wear can then occur by locally high normal or shear stresses. In favourable cases, this process leads to a surface topography that both reduces friction and extends bearing life.

1.2 Lubrication regimes

The well-known *Stribeck curve* describes how the traction (friction) between sliding solids varies with slip speed in the presence of lubricant. The sliding speed is varied, keeping it a constant percentage of the roll or average speed of the two surfaces. At low speeds, very little hydrodynamic pressure is generated and surface separation is small. The shear stress at the interface is thus largely due to *solid contact* between the surface summits mediated by any thin *boundary layers* present, such as oxides, physically or chemically absorbed molecules, and so forth. The traction coefficient then often shows approximately Coulombic behaviour, independent of normal load and speed. This low-speed regime is known as the *boundary lubrication regime*. At high speeds, hydrodynamic pressure can support the whole applied normal load, the surfaces are well separated by the lubricant film and *solid contact* is virtually absent. In this *full film regime*, traction depends mainly on the rheology of the bulk lubricant and has been extensively studied by conventional, i.e. macroscopic (elasto)hydrodynamic techniques. At intermediate speeds, the load is shared between contact pressure at the elastically deformed summits and hydrodynamic pressure in the lower-lying valley regions. This portion of the Stribeck curve is the regime described as *mixed lubrication*. The transition from full film to mixed lubrication sometimes shows an intermediate regime where surface roughness still clearly influences traction but no physical contact between summits can be measured. In this *micro-EHL regime*, a very thin lubricant film apparently still separates the summits. Generally in this regime, as the hydrodynamic film increases with speed, the traction drops rapidly from its *dry* level to the fully lubricated level.

2 EXPERIMENTAL PROCEDURES

In rolling bearing operation there are always two loaded contacts to consider: outer ring–roller and inner ring–roller. While these contacts operate as closely as possible to the ideal pure rolling condition, all real contacts involve some degree of sliding. In the present work, therefore, the individual contacts were simulated on disc machines, where the slide–roll ratio could be precisely controlled.

2.1 Disc machines

Figures 1 and 2 show the disc machines that have been used here, in which the disc material is steel. Film formation is detected from the electrical signal from a resistance measurement. The electrical circuit can be found in, for example, Schipper [**9**]. At zero speed the surfaces touch and the resistance is very small. At high speeds a lubricant film builds up, eventually causing full separation and thus a very high resistance between the running surfaces. At intermediate speeds, partial contact between surface asperities results in a fluctuating signal corresponding to a finite average electrical resistance. The output voltage signal is normalized and indicates the *level of separation*, a quantity defined here as linearly proportional to the voltage. Thus, the minimum potential difference at zero speed corresponds to 0 per cent separation (maximum degree of contact and minimum contact resistance), while the open circuit voltage indicates 100 per cent separation (minimum degree of contact and maximum resistance). As the speed

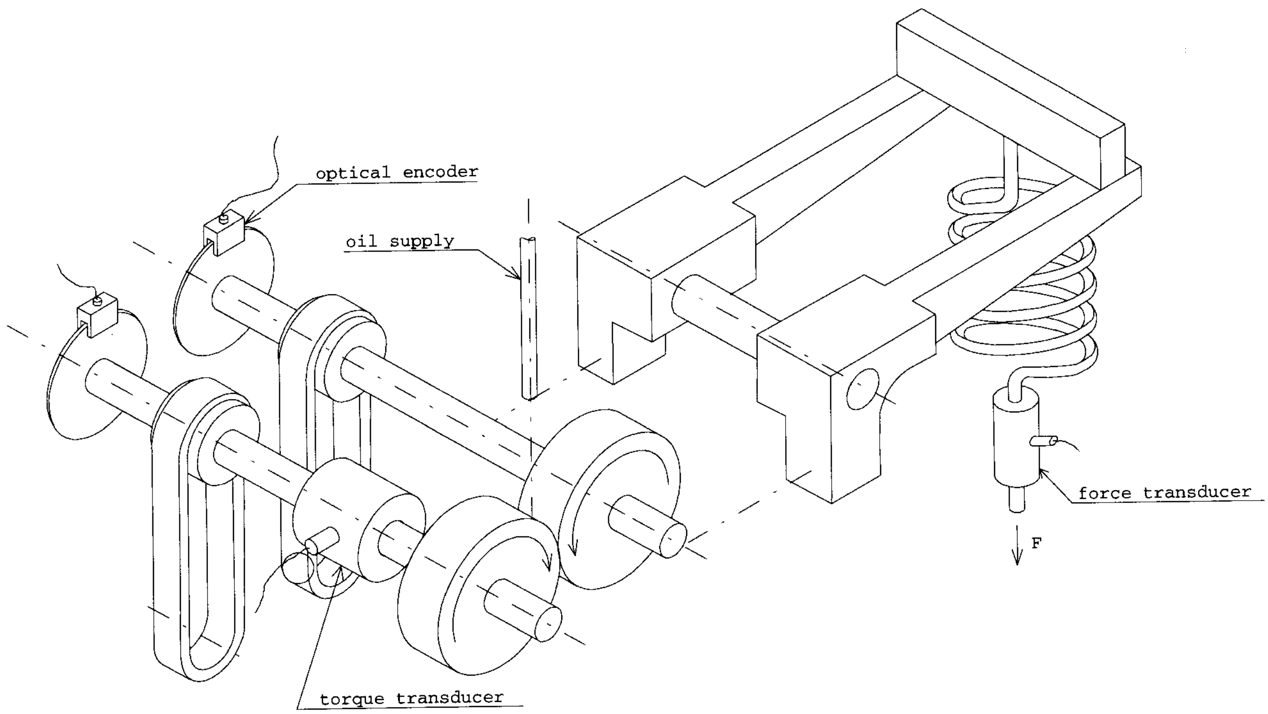

Fig. 1 The Twente two-disc machine

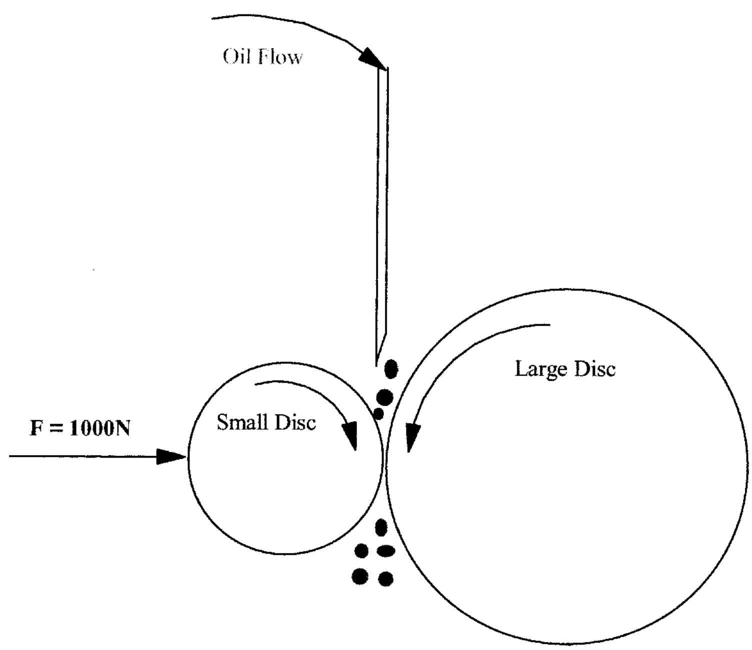

Fig. 2 The ERC two-disc machine

is ramped up (in stepwise fashion), the *lift-off speed* is defined as the speed where the output signal crosses 90 per cent. As the speed is ramped down, the 90 per cent level defines the *breakthrough speed*. While the level of separation thus defined is determined by the area of real asperity contact, the relationship is not assumed to be linear.

The Twente machine operates with two independently driven shafts, so that the sliding and rolling speeds can be separately controlled. This control is necessary to be able to record traction as a function of slide–roll ratio. On this machine, then, both separation and friction torque were measured. A radial load was applied to the discs (rings) by means of tension in a spring and the friction torque of one of the shafts was measured by a torque transducer. This torque is the sum of the friction in the contact between the rings and in the support bearing of the shaft. The maximum contact pressure is always 2.1 GPa; the contact ellipse radii are 0.6 mm in the running direction and 1.9 mm in the transverse direction.

The ERC two-disc rig has only one drive and uses a system of toothed pulleys and belts to vary the slip ratio. A maximum pressure of 2 GPa is used for all tests. The contact ellipse radii are 0.27 mm in the running direction and 0.88 mm in the transverse direction. The ERC machine was thus used only to measure the lift-off and breakthrough speeds. For these observations, special attention was paid to temperature effects. Lubricant film build-up ability is assumed to be related to the product of the speed and the lubricant viscosity, a strongly temperature-dependent quantity. Changes in temperature are therefore expected to have a large impact on the lift-off results. Accordingly, a *40 °C equivalent* or *temperature corrected lift-off speed* is defined by multiplying the lift-off speed and the viscosity at the measured temperature, obtained from thermocouples sliding on the large disc, and dividing by the viscosity at 40 °C. Viscosity values are read from the supplier's chart.

2.2 Materials and surface finish

For both machines, one of the two discs was crowned (the larger in the case of the ERC rig), while the counter-disc was uncrowned. All discs were manufactured from martensitic hardened steel. On the Twente rig, the surface of the crowned disc was always circumferentially honed ($R_a = 0.05$ μm) while the finish of the uncrowned disc was varied. Grinding and honing were always carried out in the circumferential direction, except for the cross-honed surface where the honing grooves make an angle of approximately 25° with the running direction. The abbreviations used in this paper to describe the different finishes used are explained in Table 1.

The lubricants that were used were Shell Turbo T9 and T32 for the Twente and ERC rigs respectively. These oils contained no additives other than antifoam and anticorrosion.

2.3 Roughness measurements

All three-dimensional roughness measurements were performed on a Prolap interference microscope. The objective used has a magnification factor of 10 and a lateral resolution of 1.1 μm. The measurement area was 440 μm × 339 μm and the number of measurement points was 304 × 228. In addition to the high-pass filter, a 3 × 3 smoothing function was applied, replacing each point by a weighted average of itself and its eight nearest neighbours.

Table 1 Abbreviations

Abbreviation	Finish	R_a (μm)
BG	Ground	0.04
P	Isotropically machined	0.05
V	Vibro finished	0.03
TSP	Polished	0.01
SP	Polished	0.001
CrH	Cross-honed	0.02

3 RUNNING-IN EXPERIMENTS

Since the principal objective of this work was to investigate the effect of surface topography on traction and surface separation, it is necessary to make measurements under operating conditions for which the topography is stable. If a tribo system is operating under conditions where the film thickness is insufficient for full surface separation, then they touch, and locally high stresses (both normal but particularly shear) cause the surfaces to deform plastically and/or to wear. If these processes are not severe, the plastic deformation shakes down and the wear rate converges to zero with successive over-rollings. The process is then referred to as *running-in*. In practice, running-in occurs precisely in the regime of interest, i.e. under mixed lubrication conditions where the degree of contact increases as the surface speed drops. The strategy adopted was thus to ensure that all surfaces were run in at speeds lower than the lowest subsequently used for traction measurements. To observe running-in, the conditions, either pure rolling or 2.8 per cent slip, were chosen to be mild in terms of the degree of expected asperity contact, as in most rolling bearing applications. While this procedure obviously shifts the lubrication regimes, the new boundaries are fixed, allowing consistent measurement across the new regimes. Some interesting observations of running-in follow.

3.1 Running-in: pure rolling

Under pure rolling conditions both lift-off and breakthrough speeds, measured on the ERC rig, decreased between successive runs. For instance, the lift-off speed of the BG–BG combination was measured 20 times, decreasing from 175 r/min in the first test to 110 r/min in the last. A number of tests on running-in under pure rolling conditions were made, with results depicted in Fig. 3. A V–P disc combination was run under pure rolling conditions for 80 000 over-rollings at 70 per cent of its initial lift-off speed, with the result that, at the end of the test, full film separation had been achieved. The new lift-off speed had become equal to the running-in speed. Repeating the test at 70 per cent of this new lift-off speed, then after some 220 000 over-rollings there was again lift-off. In another test, the TSP–SP combination was run for 173 000 over-rollings at 50 r/min, about 60 per cent of the 85 r/min initial lift-off speed. Here too, after 56 000 revolutions, lift-off occurred.

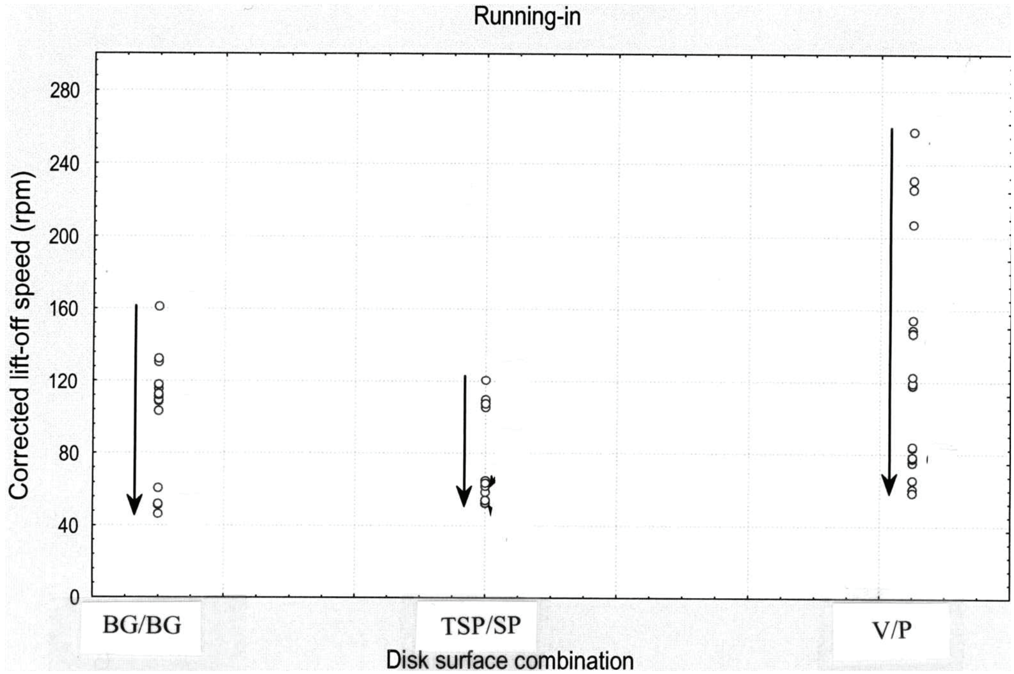

Fig. 3 Lift-off speed during running-in

3.2 Running-in with slip

A test was made on two ‘fresh’ BG surfaces, with results illustrated in Fig. 4. The vertical arrows denote running-in, i.e. reduction in lift-off speed with number of over-rollings at fixed slip ratio, while the crossed arrows indicate switching between slip and pure rolling modes without dismounting the discs. The numbers in the graph show the sequence followed in the test.

Measurement on these discs with 2.8 per cent slip initially showed no lift-off within the speed range, up to 300 r/min, available in these tests. They were then run at 40 r/min (still with 2.8 per cent slip) until eventually, after 172 000 large disc revolutions, lift-off was achieved at a (corrected) lift-off speed of 71.2 r/min. Switching now to pure rolling, the lift-off speed was observed to be 177 r/min. Then after running at pure rolling for 48 000 revolutions at 40 r/min, the lift-off speed decreased to 47 r/min; so the surfaces were behaving as though they were well run in for pure rolling. Returning to 2.8 per cent

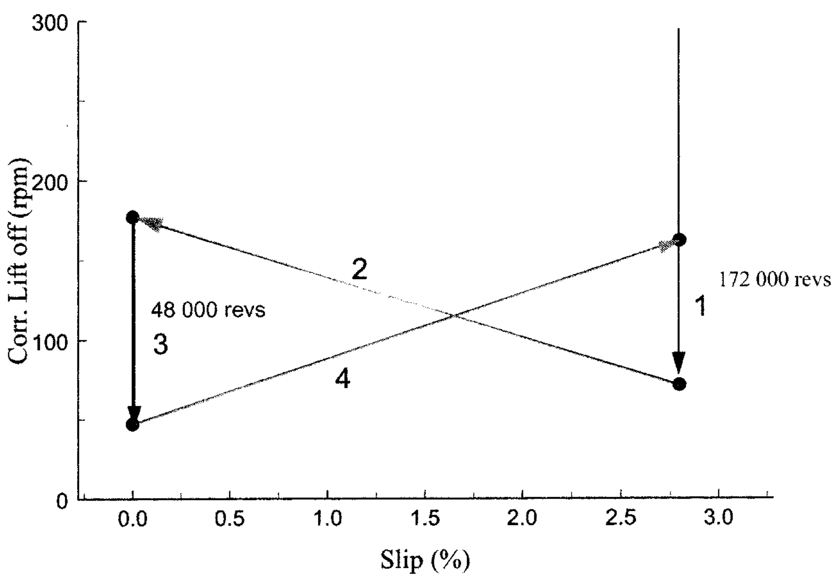

Fig. 4 Running-in (BG–BG surfaces)

slip conditions, however, these surfaces could not reach lift-off until 162 r/min.

To verify this unexpected result a second similar test was made, depicted in Fig. 5. Again the discs were run in with slip (40 r/min) until lift-off occurred at 40 r/min after 180 000 revolutions. On changing to pure rolling, the lift-off, as before, became very high (no lift-off within the range of the machine). Again, by letting the discs run at 40 r/min for 48 000 revolutions in pure rolling, the discs became run in. This time, switching to slip once more increased the lift off speed, but by not so much as before. After a further 122 400 revolutions (step 5), lift-off returned to 40 r/min, where it remained after switching again to pure rolling (step 6).

A CrH–CrH combination was also run in using a similar test sequence. The result is depicted in Fig. 6. The surfaces were first run in at 2.8 per cent slip, until the lift-off and run-in speeds were equal. However, in contrast with the previous tests, changing the slip mode to pure rolling for these surfaces (step 2), did not result in a higher lift-off speed. In fact, the lift-off speed was even somewhat lower than at 2.8 per cent slip. Changing back to slip the lift-off speed returned to its higher initial value but, after a further 122 400 revolutions, once again became equal to the running-in speed.

It may be hypothesized that, when running in, either at pure rolling or with slip, the lift-off speed becomes equal to the running-in (sum) speed, which means that the 90 per

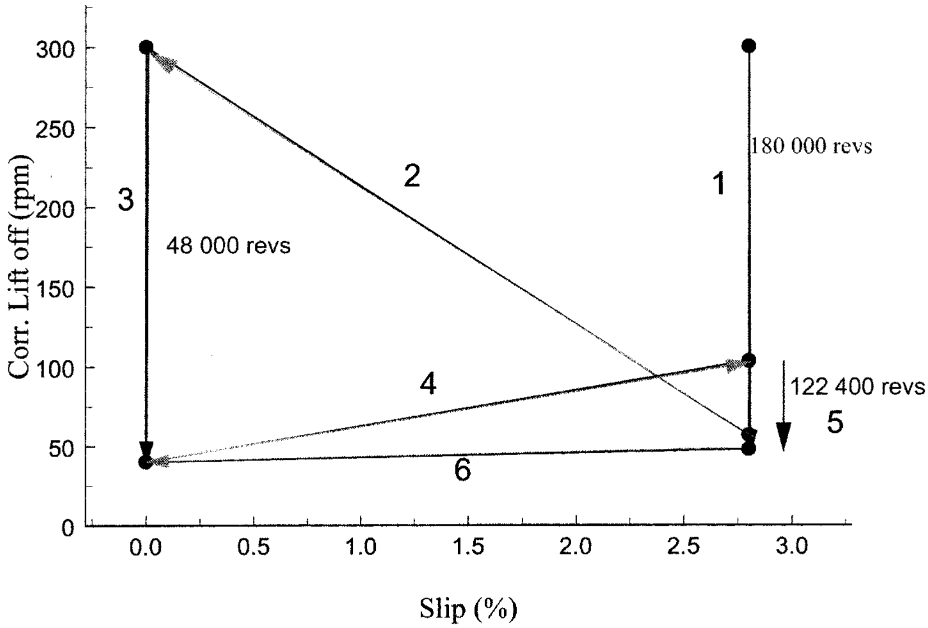

Fig. 5 Running-in (BG–BG surfaces)

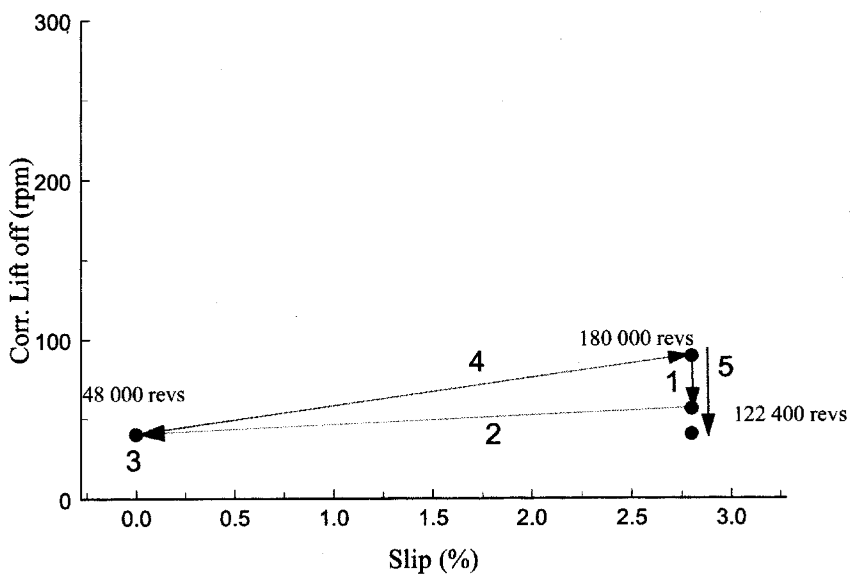

Fig. 6 Running-in (CrH–CrH surfaces)

cent level of separation is reached at a sum speed independent of the slip ratio. By extension, it may be further hypothesized that, for a given surface combination, i.e. given surface topography, the relationship between level of separation and sum speed for *fully run-in* surfaces becomes independent of slip speed.

4 TRACTION EXPERIMENTS

The friction and traction measurements were made on the Twente two-disc rig, using discs with ground, honed and two different isotropic finishes [vibro finished (V) and surface isotropically machined (P)]. The counter-surface was always circumferentially honed.

The surfaces were subjected to running-in in two stages, both at 3.5 per cent slip. During the first stage the surfaces were run in at the same $\Lambda = h/R_q = 1$ (h is the nominal film thickness and R_q the r.m.s. composite surface height) for 60 000 revolutions, to remove the sharpest peaks in a mild way. Note that this implies that the various surfaces were run in at different sum speeds. In order to be able to compare the performance of the various surfaces the second running-in stage of 220 000 rev was performed at the same speed for all combinations, chosen such that the nominal central film thickness was 0.04 μm (requiring a sum speed of 1.3 m/s). For all surface combinations the friction signal became constant during this second stage, indicating a steady state, at least so far as traction is concerned.

4.1 Varying slip ratio; constant sum speed

In the previous section, the slip was switched between 2.8 per cent and pure rolling. In this section it will be continuously varied between −3.5 and 3.5 per cent, the surfaces having been run in at 3.5 per cent. The separation voltage signal and the traction coefficient were measured simultaneously. Since running-in was performed at only one slip ratio, the surfaces are not *fully run in* in the sense discussed in Section 3.2. This is clearly seen from Figs 7 to 11 showing the friction and various levels of separation for the different disc combinations as a function of slip ratio. All measurements presented in these figures were obtained with equal nominal film thickness h, i.e. the same constant sum speed U taken to be larger than the speed at which the surfaces were run in, as discussed in Section 3.

The quantity H, defined as

$$H = \frac{\eta U}{p}$$

where η is the dynamic viscosity and p is the average contact pressure, has dimensions of length and, like Λ, is an inverse measure of the severity of the operating

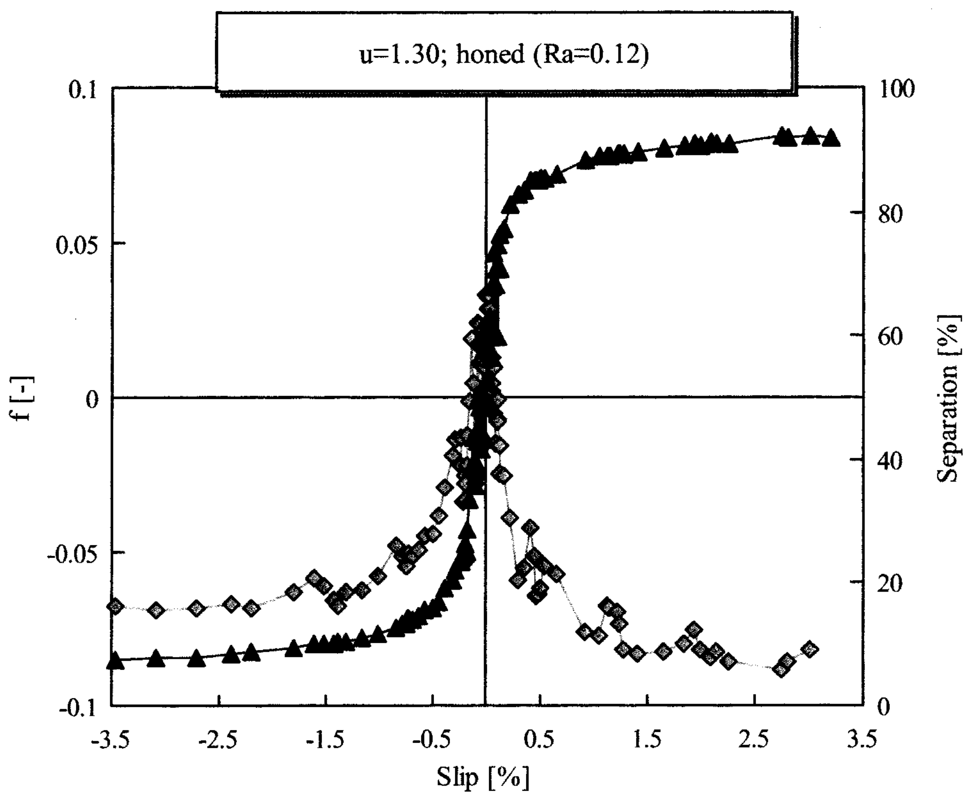

Fig. 7 Traction coefficients of the honed (rough)–honed combination

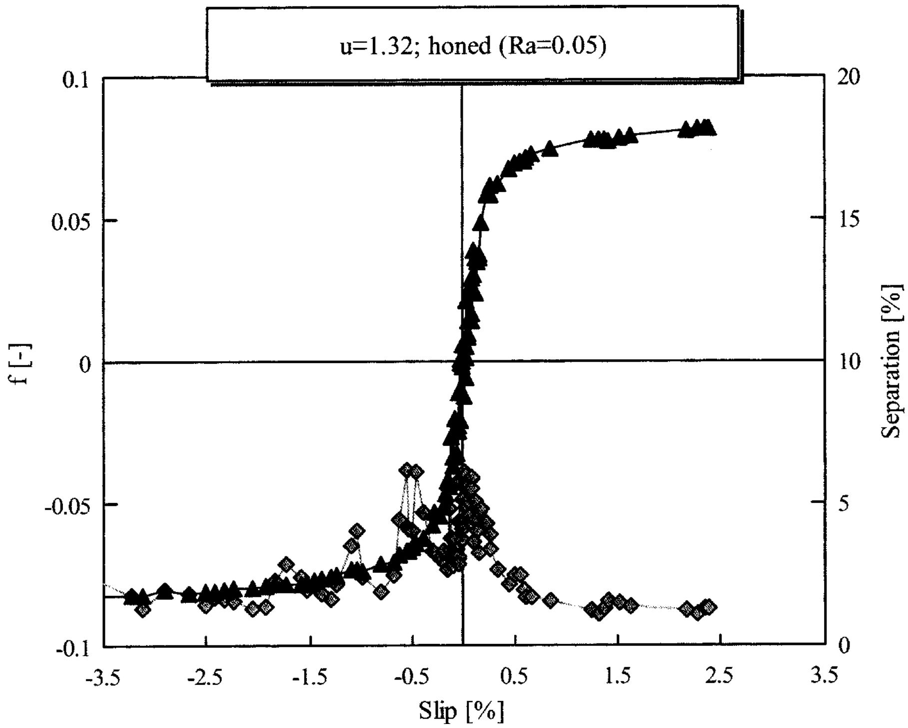

Fig. 8 Traction coefficients of the honed (smooth)–honed combination

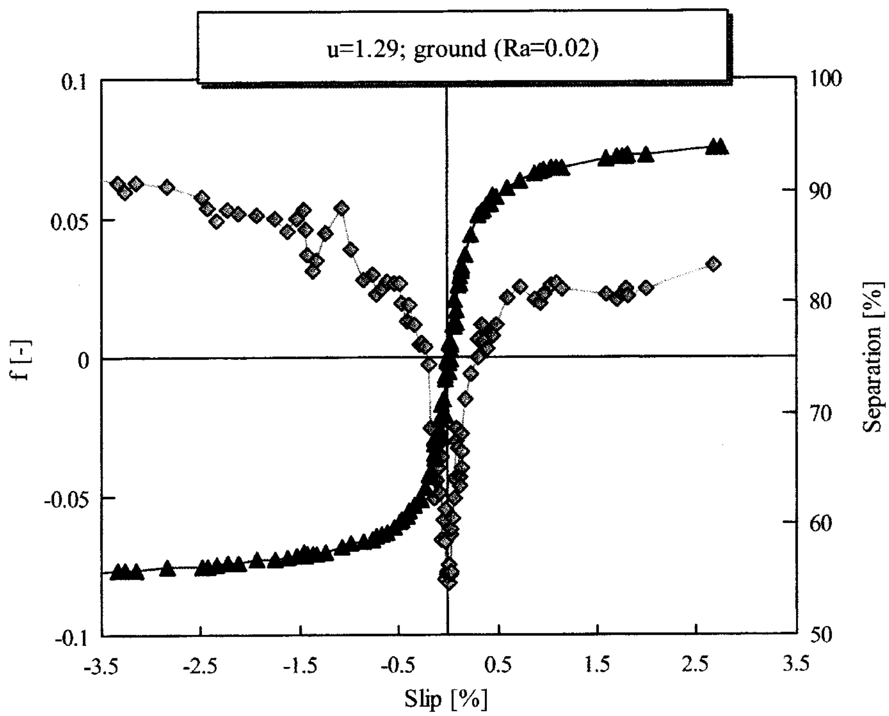

Fig. 9 Traction coefficients of the ground (smooth)–honed combination

conditions in a contact. It is often used as the independent variable in plotting a Stribeck curve. For the traction measurements displayed in the figures, $H = 4.5\times 10^{-12}$ m while, during the second stage of running-in, $H = 3.3 \times 10^{-12}$ m.

4.1.1 *Friction*

All measured friction versus slip curves have the same shape, characterized by a linear region around the origin (Newtonian fluid behaviour: shear stress $\propto$ strain rate),

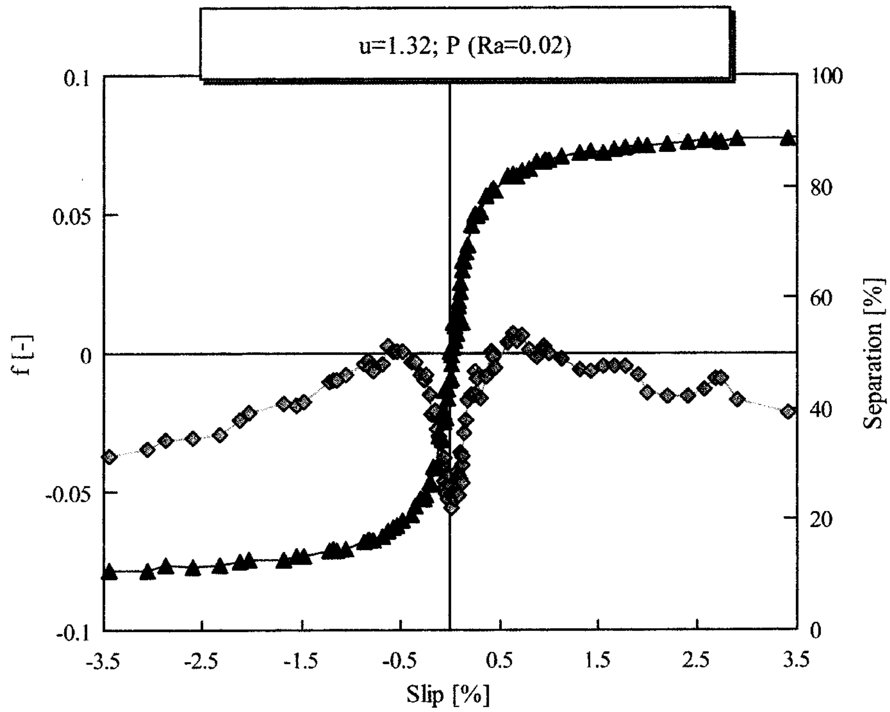

Fig. 10 Traction coefficients of the P ($R_a = 0.02$ μm)–honed combination

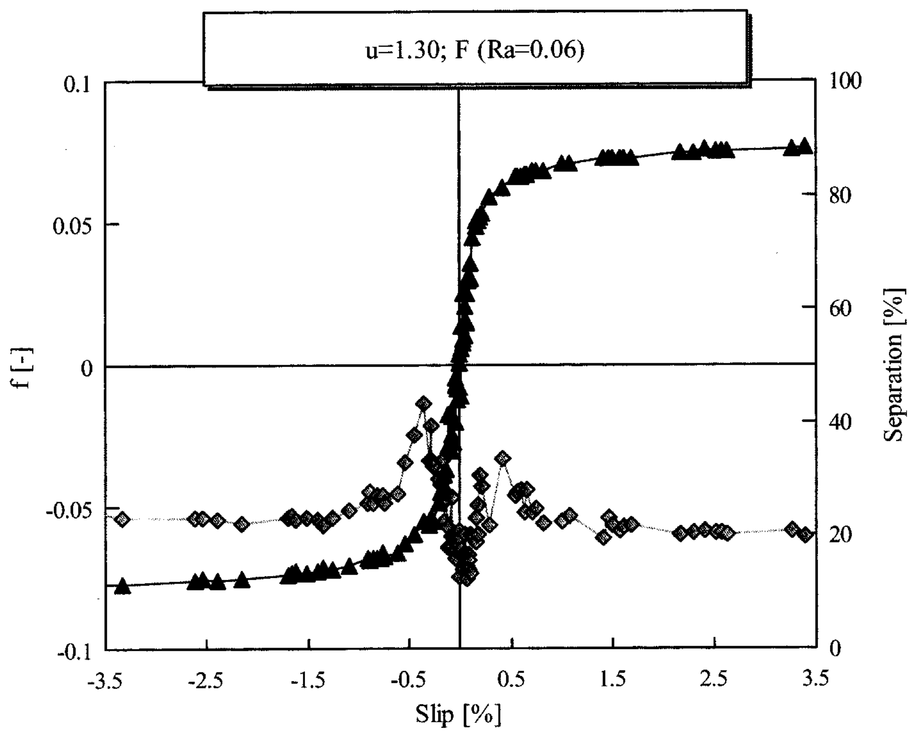

Fig. 11 Traction coefficients of the vibro finished ($R_a = 0.06$ μm)–honed combination

followed by an approach to a stress limit. If the strain rate were increased beyond about 5 per cent slip, traction would decrease again due to thermal effects. Comparison of Figs 7 to 11 shows that surface roughness does not have a dominant influence after the degree of running-in experienced by these surfaces. The friction measured at the largest slip, however, varies by about 10 per cent between the different surfaces. A variation of this magnitude is of considerable significance for bearing performance.

As already remarked, torque measurements on a two-disc machine include both the friction in the contact between the discs and the friction in the support bearings. To remove the contribution from the bearings, the friction–slip curves presented here have been shifted vertically to bring the friction level to zero at pure rolling. In practice, this was achieved by making the levels of friction at 3.0 and −3.0 per cent slip equal.

4.1.2 *Separation: the three types*

The electrical resistance measurements from Figs 7 to 11 show that, despite the fixed sum speed U, the separation is not constant but a strong function of the slip between the contacts, demonstrating that the surfaces are not yet fully run in. Comparison of the results for different surfaces shows that this phenomenon depends on the type of surface involved. The most obvious distinction is between the ground and honed surfaces. In the case of a honed surface the separation decreases more or less continuously with increasing slip, designated type I behaviour, whereas in the case of a ground surface it continuously increases, designated type II behaviour.

This confirms in a general way the measurements on the ERC two-disc machine, as presented in Figs 4 to 6. For example, Fig. 6 shows that, for honed surfaces, the lift-off speed after the running-in with slip becomes (slightly) lower when switching over to pure rolling. Hence the separation at slip is lower than for pure rolling, as shown in Figs 7 and 8 for the honed combinations, also run in with slip. Again (Figs 4 and 5), both ground combinations show the opposite: better separation at slip than at pure rolling if running-in has been done with slip. This is similar to what can be seen in Fig. 9 for the ground–honed combination.

The isotropic surfaces (Figs 10 and 11) manifest yet a third type of behaviour as slip increases, called type III. While, like the ground combination, the separation initially increases, for the isotropic surfaces it then passes through a maximum at around 0.5 per cent slip and thereafter decreases.

4.2 Varying sum speed; constant slip ratio

Figures 12 and 13 show typical Stribeck curves for isotropic and honed surface combinations, the result of measurements where the sum speed is (stepwise) increased while keeping the slip ratio constant, here at 3.5 per cent slip. As mentioned in Section 4.1, running-in was done at an H value to the left of any point belonging to the curves. Hence, throughout the measurements the speed, which varies from 1.3 to 6.6 m/s, was never lower than the running-in speed.

From the separation signal plotted in Figs 12 and 13, the film is gradually building up as the speed (or H) increases. Clearly then, the running-in procedure had not yet produced the fully run-in state. In Fig. 13 for the rough honed surface, the friction at first decreases steeply with increasing speed in the region where the surface separation is still growing. This decrease is generally attributed to the diminishing number of asperity contacts. Since for $H > 2 \times 10^{-11}$ (lift-off) there are effectively no remaining contacts, the continued (now approximately linear) decrease must here be attributed to thermal effects. The

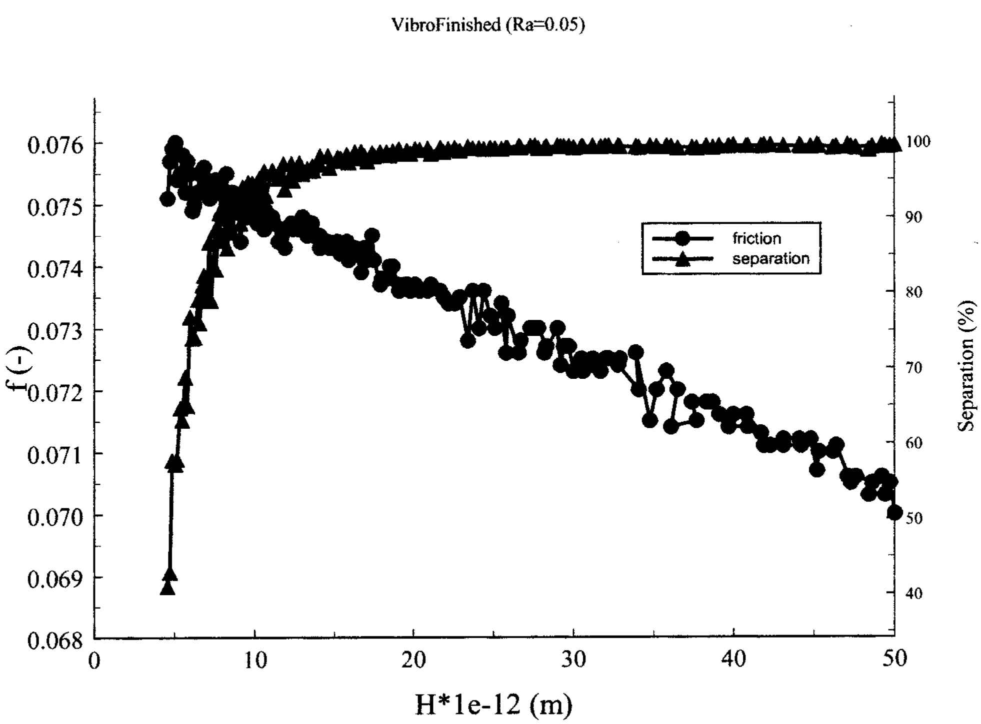

Fig. 12 Stribeck–separation curves for an isotropic vibro finished surface (slip, 3.5 per cent)

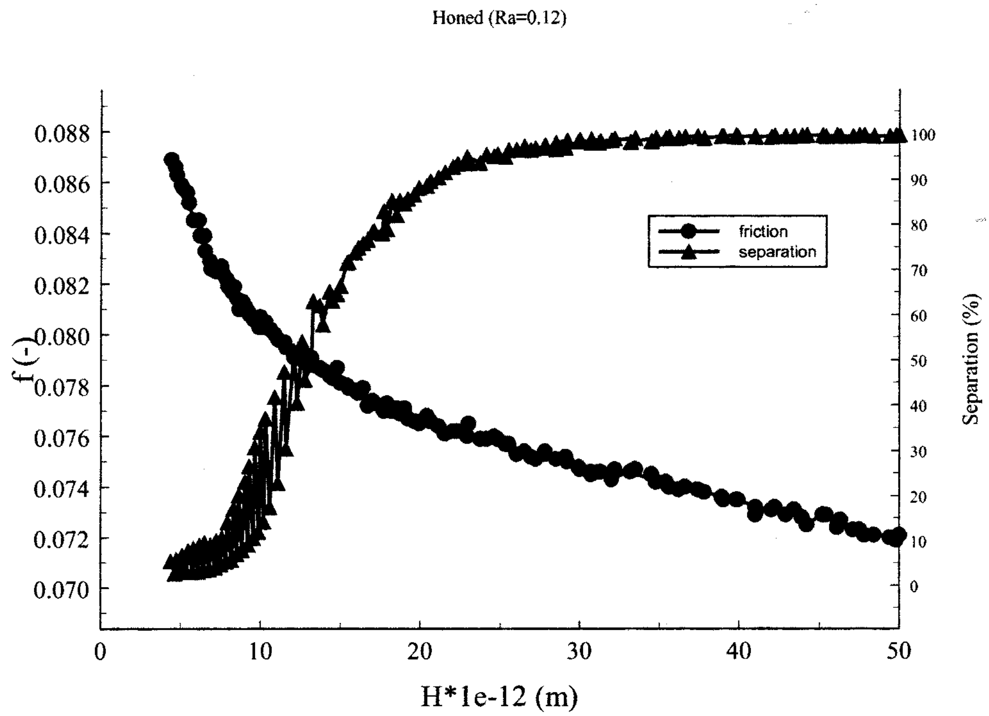

Fig. 13 Stribeck–separation curves for a rough honed surface (slip, 3.5 per cent)

increase in speed results in increases in both film thickness and heat generation. The net effect is a rise in temperature inside the contact, resulting in lower viscosity and hence also lower traction. In this full film regime, prediction of the traction coefficient at any speed simply from the viscous shearing of the film, whose temperature is known from thermocouple measurements and whose thickness is found from a film thickness formula, readily yields the observed linear decreasing behaviour.

By contrast, Fig. 12 shows the Stribeck curve of an isotropic vibro finished surface, where the high-speed linear character of the Stribeck curve is maintained even in the low-speed area, $H < 10^{-11}$, where there is still electrical contact between the surface asperities. Apparently in this case the surface roughness interaction contributes negligibly to the friction, whose behaviour continues to be dominated by the thermal mechanism. By extrapolation of the linear part of the Stribeck curve in Fig. 13, it may be estimated that the contribution of asperity interaction to friction is approximately 10 per cent at the lowest speed measured.

The friction–slip measurements for the various disc surface combinations shown in Figs 7 to 11 were repeated for five additional values of H (by varying U), all chosen from their Stribeck curves to fall near the transition from full film to mixed lubrication. By reading the traction coefficient at 3 per cent slip from each traction plot at each sum speed, the Stribeck curves in Fig. 14 were constructed. For all surface combinations, full separation of the surfaces occurs, according to the resistance signal, at some H within the range shown in this figure. Each plotted point represents the average of the measurements from two or three pairs of discs, where the largest difference between the two measurements was 0.002.

5 DISCUSSION

5.1 Film collapse and running-in

The conventional view that lubricant film build-up under slip conditions is worse than with pure rolling (see, for example, references [**5**], [**7**] and [**10**]) obviously is not always correct. Such type I behaviour is not universal. The results presented here show that it depends on how the surfaces have been run in and, more critically, on the surface structure. The running-in experiments show that all surfaces initially adapt to the prevailing running conditions, if these are sufficiently mild, whereby the running-in and lift-off speeds become approximately equal. Changing to another slip mode then alters the lift-off speed but, after a number of such changes, all surfaces achieve a more or less slip-independent run-in condition, showing good lift-off either with or without slip.

Differences observed between the various surface finishes concern only the manner in which each approaches its final state. For the ground surfaces (Figs 4 and 5), every switch of slip mode results in an immediate increase of lift-off speed, followed by a decrease as running-in at the new slip speed continues. These changes with every switch, regardless of direction, become progressively smaller, thereby leading eventually to the final run-in state.

A different sequence applies to the honed surface combination (Fig. 6). Here the run-in surface produced

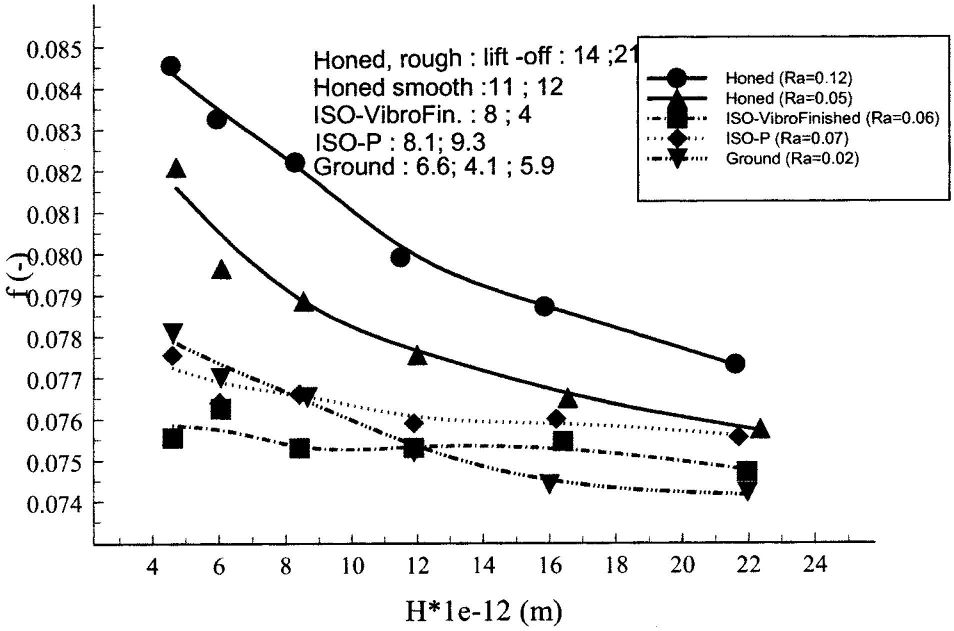

Fig. 14 Stribeck curves extracted from friction–slip curves (slip, 3.0 per cent)

under slip conditions also behaves as though run in after switching to pure rolling, giving in fact a small drop in lift-off speed rather than an increase. Despite this, however, it is clear that modification of the surface structure continues under these pure rolling conditions. Even though this is not reflected by any modification in lift-off speed at pure rolling, switching back to the slip condition results in an immediate increase in lift-off speed. Subsequently, as with the ground surface combinations, this increased lift-off speed falls back to the running speed. For these surfaces then, the final run-in state is approached as this one-way-only increase at switching becomes progressively smaller.

From these tests it follows that running under slip or pure rolling conditions may be thought of as two different but interdependent surface-modifying processes. The surface topography produced by the one depends on the topography that it is presented with by the other. In such a situation, the optimum surface will always be arrived at by iteration of the two processes.

Three broadly different types of surface separation versus slip speed behaviour during approach to full run-in have been identified in the experiments reported here: type I, for honed surfaces; type II, for ground surfaces; type III, for isotropic surfaces. The most striking difference in behaviour is observed between types I and II, the two finishes which have the most striking similarity: their linear (i.e. circumferential) lay. Clearly then, such a patterning on this scale is not responsible for the difference in lift-off behaviour. Rather, attention should be given to the full three-dimensional microgeometry of the surfaces.

The type III behaviour of the isotropic surfaces, at least at very small slip, is similar to the ground surfaces. One property distinguishing types II and III from the type I finish is the relatively high slope of the surface asperity summits. A plausible explanation, then, is that under slip these summits cause a squeeze effect, building pressure to contribute to surface separation. This squeeze effect is much smaller in the case of honing, where the more plateau-like summits have lower slopes. For these surfaces, the more gentle summits produce less violent hydrodynamic pressure fluctuations and thus a thicker film under pure rolling but, in the presence of slip, the conventional non-Newtonian lubricant effect causes the film to collapse. These explanations are in part confirmed by numerical calculations by Lubrecht *et al.* [**11**]. They showed that short-wavelength asperities (asperities with a relatively high slope) are less flattened if the rough surface is moving faster than the smooth surface and more flattened if the rough surface is the slower, giving in other words an asymmetrical behaviour around pure rolling. Later Jacod *et al.* [**12**] showed that this is a consequence of the assumed Newtonian fluid model, using instead an Eyring model to calculate the deformation of single Fourier components inside an EHL contact. This showed that in the presence of slip the effective wavelength increases from the inlet of the contact towards the centre, as a result of which the amplitude of these waves is more reduced than in pure rolling conditions. This increase in wavelength was found to be a function of the *absolute value* of the slip! This offers an explanation of the symmetrical type II behaviour observed here in the experiments of Fig. 9.

Interesting comparisons may be drawn with the experiments of Ishibashi and Sonoda [**5**]. Taking similar precautions to run-in at a lower sum speed, they made simultaneous measurements of traction and D_{EHL} versus slip, where D_{EHL} is the percentage of time that full EHL conditions prevail in the contact. It is essentially the inverse of the separation level and is also deduced from the electrical contact resistance. The surface roughness produced either by grinding or by grinding and polishing, was characterized by R_t, the peak-to-valley height, yielding a wide range from 0.03 to 3.0 μm. Before running-in, all surfaces showed the largest separation at pure rolling, decreasing monotonically as slip increased. Not surprisingly, the roughest surfaces never fully ran in but continued to show this type I behaviour. By contrast, the separation for the smoother surfaces after run-in became close to 100 per cent over the whole slip range from -20 to $+20$ per cent, i.e. fully run-in behaviour. In one case, starting from the pure rolling maximum the separation showed a small minimum at about 1 per cent slip and a small second maximum at about 2 per cent before the monotonic decrease. Thus, apart from the narrow central maximum, this case resembles type III. The distinction between these earlier results and the present is now only too obvious; all the surfaces measured in the previous work were based on grinding and showed only type I or perhaps type III behaviour, while all the ground surfaces in the present work showed only type II.

It should be noted that, in the earlier work, a different running-in procedure was used: 8000 rev at pure rolling followed by 27 000 rev at 15 per cent slip. Subsequently, traction measurements were made up to 20 per cent slip. This relatively small number of over-rollings with high slip and subsequently making measurements at even higher slip values may be one reason for the contradictory results.

Another reason could be ascribed to the lubricant (Shell T9 and T32) used in the present tests. When this was changed to Shell Tellus R5 oil, a lubricant containing anti-wear additives, both honed (type I) and isotropic (type III) surfaces changed their behaviour, now showing an increase in separation with slip (type II). The additives are activated by (local) heat produced by sliding in the high-shear-rate regions associated with summits, forming electrically insulating protective boundary layers which simulate an increase in separation. When the sliding speed is reduced, these are rubbed off while no new protective layers are formed. As observed, then, in the presence of such additives, (apparent) type II behaviour would be universally expected. None of the finishes investigated behaved any longer like the type I ground surfaces reported by Ishibashi and Sonoda [**5**].

Finally, it should of course be recalled that, in the attempt to model any tribological behaviour on the basis of surface topography, there are always two surfaces involved. While scalar composite amplitude parameters such as R_a can be defined to describe the two surface topographies, it is by no means clear how this might be generalized to vector and tensor parameters, such as slope and curvature or, on another scale, to the combination of two different surface patternings or lays. Even though one of the two surfaces in the present experiments was, at least statistically, always the same, the mechanism proposed here to account for the observed differences in behaviour of various types of surface should, nevertheless, be regarded as tentative.

5.2 Friction

While the surface topography clearly has a large impact on film collapse its effect on the friction level is also important. Figs 12 to 14 reveal a number of interesting phenomena. Most striking, perhaps, is the good correlation between the lift-off speed and the level of friction. The rough-honed surface has a high lift-off speed ($14 \times 10^{-12} < H < 21 \times 10^{-12}$) followed by the smooth-honed surface ($11 \times 10^{-12} < H < 12 \times 10^{-12}$). The other surfaces have about the same lift-off speed and also comparable levels of friction. These observations support the calculations of Patir and Cheng [**1**], who demonstrated the tendency for a longitudinal lay to give a reduced film thickness.

This influence of lay and isotropy on film thickness is also the most obvious explanation of the quite large differences in friction seen at the upper end of the H range, where direct asperity contact no longer occurs [**13**]. Thus, the surface topography gives differences in mean film thickness, and hence also in mean shear rate and traction, in the full film regime. In the low-H regime, it gives differences in the degree of local contact and hence also of micro-EHL and boundary friction. The observed differences in friction level at the two ends of the H range are now seen to depend on quite different aspects of the surface topography: at low H, shape of summits; at high H, anisotropy and lay.

The correlation between lift-off speed and friction is in marked contrast with the influence of R_a, the roughness parameter commonly assumed to determine the level of friction. Figure 14 indicates otherwise. The rough-honed surfaces are indeed much rougher than the other surfaces and show a higher friction level but there is no further correlation with R_a. The smooth-honed surface is smoother than the P surface but gives a higher level of friction while, although the vibro finished surface is rougher than the ground surface, the measured friction coefficients for these surfaces are about equal.

The three surface finish types identified on the basis of run-in behaviour are also distinguished by their Stribeck curves. The slope of the curves is clearly much larger for type I surfaces than for type III throughout the whole H domain where, in fact, type III surfaces show only 1 or 2 per cent change in traction. Type II surfaces have an intermediate behaviour, with a slope which becomes small only in the region where no electrical contact exists, i.e. $19 \times 10^{-12} < H < 22 \times 10^{-12}$.

6 SUMMARY AND CONCLUSIONS

This work represents a semiquantitative investigation into the effects of surface roughness on the running-in and friction of EHL contacts operating in or close to the mixed lubrication regime. While with modern surface topographs it is straightforward to evaluate dozens of different surface roughness parameters, the difficulties of finding the appropriate parameters to characterize and discriminate between these functions are well known. Giving actual numerical values for surface parameters has therefore been deliberately quite restricted. Instead, the various surfaces are generally described and distinguished by the finishing processes themselves, while qualifiers such as *rough* or *smooth* acquire relative significance. This procedure conveniently parallels the manufacture of the surfaces, which in the present preliminary investigation were chosen to represent surfaces that could be readily produced and which consequently are of potential interest to the bearing industry. For this practical reason, and also because in any case there exist already a number of models for predicting the tribological behaviour of rough EHL contacts, it was decided not to use a full or even partial factorial design for the selection of the various surface parameters. Qualitatively, the primary effects described are reproducible but, except in those few stated cases where averaging was possible, the results presented relate to a single operating contact, ruling out any attempt at error analysis. In particular, for the running-in process outlined in this paper, its iterative nature shows that it is not uniquely related to the instantaneous values of the roughness parameters of either of the two surfaces, even though the fully run-in surfaces may each show systematic changes from their initial state. This is perhaps fortunate, since changes in the parameters cannot in any case be monitored during the experiment without disturbing the measurements.

The observed phenomena are of course related to the combination of the two surfaces, so that the task of understanding their behaviour is simplified if only one of the two is varied. Therefore, as described, in most cases one of the pair has been kept of the same type. However, even if the initial roughness parameters of one surface are held as constant as possible, their influence on the composite roughness parameters of the combination which, it must be assumed, determine the mixed EHL regime, depends not only on the corresponding roughness parameters of the second surface but also crucially on the cross-correlation between the two relatively moving surfaces. Changes in this correlation could well explain the sudden changes in lift-off seen when the slip is toggled between two different values. Similarly, it may partially explain the three types of separation behaviour observed when the slip is changed continuously.

6.1 Running-in and film breakdown

Running-in under mild conditions will eventually always lead to full separation, such that the lift-off and running-in speeds become equal. Initially, lubrication performance (surface separation) is strongly dependent on the running-in conditions. To achieve good surface separation under all slip conditions, the running-in process needs to be repeatedly switched between pure rolling and slip.

With sufficient iterations, a fully run-in topographical state is produced, showing full film separation for any slip, provided that the rolling speed at least equals that at which the surfaces have been run in.

6.2 Continuously varying slip and film breakdown

Prior to achieving the fully run-in state, the separation of surfaces as a function of the ratio of slip between them behaves in one of three distinct ways, determined by the surface topography. The particular behaviour manifested appears to depend on the magnitude of the curvature or slope of the surface summits. For honed surfaces with more flattened summits, film build-up is always found to be better at pure rolling than with slip: type I behaviour. For the sharp summits on ground or isotropic surfaces, the film build-up observed is always worse at pure rolling: type II or III behaviour. The effects are confirmed by a numerical analysis by Jacod *et al.* [**12**]. This distinction is, however, lost if a lubricant with anti-wear additives is used, whereupon type I behaviour disappears altogether and all surfaces show most breakdown at pure rolling.

6.3 Friction

Designing a proper three-dimensional surface structure is an effective way to reduce friction and to improve lubrication performance. Thus, relatively rough isotropic surfaces perform well compared with either rough or smooth surfaces with longitudinal linear lay. Isotropic surface finish leads to a less steep Stribeck curve.

By changing the three-dimensional surface topography while keeping the roughness amplitude R_a fixed, the level of friction can be influenced by as much as 10 per cent. To achieve such a variation simply by proportionally scaling the surface height, the R_a value would have to be changed by a factor of at least 3!

ACKNOWLEDGEMENTS

The authors would like to thank Dr H.H. Wittmeyer, Managing Director of SKF Engineering and Research Centre, BV, for his kind permission to publish this paper. We would also like to thank Professor Jacobson (at present at Lund University) for many fruitful discussions on the work presented here.

REFERENCES

1 **Patir, N.** and **Cheng, H. S.** An average flow model for deterministic effects of three-dimensional roughness on partial hydrodynamic lubrication. *Trans. ASME, J. Lubric. Technol.*, 1978, **100**, 12–17.

2 **Venner, C. H.** and **Lubrecht, A. A.** Numerical analysis of the influence of waviness on the film thickness of a circular EHL contact. *Trans. ASME, J. Tribology*, 1996, **118**, 153–161.

3 **Venner, C. H.** Multilevel solution of the EHL line and point contact problems. PhD thesis, University of Twente, The Netherlands, 1991.

4 **Johnson, K. L.** and **Higginson, J. G.** A non-Newtonian effect of sliding in micro-EHL. *Wear*, 1988, **128**, 249–264.

5 **Ishibashi, A.** and **Sonoda, K.** Effects of surface roughness and type of oil on traction characteristics between steel rollers. *Wear*, 1994, **175**, 39–49.

6 **Jacobson B.** Redistribution of solidified films in rough Hertzian contacts. Part II: experimental. In Proceedings of the 14th Leeds–Lyon Symposium on *Tribology*, Lyon, France, 1987, 1988, pp. 59–63 (Elsevier, Amsterdam).

7 **Hamer, J. C., Sayles, R. S.** and **Ioannides, E.** The collapse of sliding micro-EHL films by plastic extrusion. *Trans. ASME, J. Tribology*, 1991, **113**, 805–810.

8 **Tripp, J. H.** and **Ioannides, E.** Effects of surface roughness on rolling bearing life. In Proceedings of the Japan International Tribology Conference, Nagoya, Japan, 1990, pp. 797–802.

9 **Schipper, D. J.** Transitions in the lubrication of concentrated contacts. PhD thesis, University of Twente, The Netherlands, 1988.

10 **Jacobson, B.** Thin film lubrication of non-smooth surfaces. In Proceedings of the 82nd Meeting of the AGARD SMP on *Tribology for Aerospace Systems*, Sesimbra, Portugal, 1996, pp. 6.1–6.7.

11 **Lubrecht, A. A., Graille, D., Venner, C. H.** and **Greenwood, J. A.** Waviness amplitude reduction in EHL line contacts under rolling–sliding. *Trans. ASME, J. Tribology*, 1998, **120**, 705–709.

12 **Jacod, B., Lugt, P. M., Dumont, M.-L, Tripp, J. H.** and **Venner, C. H.** Amplitude reduction of waviness in elastohydrodynamic lubrication using an Eyring fluid model. *Proc. Instn Mech. Engrs, Part J, Journal of Engineering Tribology*, 2000, **214**(J4), 343–349.

13 **Tripp, J. H.** and **Hamrock, B. J.** Surface roughness effects in elastohydrodynamic contacts. In Proceedings of the 11th Leeds–Lyon Symposium on *Tribology*, Leeds, 1984, 1985, pp. 30–39 (Butterworth, London).

