Process of Boundary Film Formation from Fatty Acid Solution

Ken Nakano1)* and Hugh A. Spikes2)

1)Faculty of Environment and Information Sciences, Yokohama National University
   Yokohama 240-8501, Japan
2)Department of Mechanical Engineering, Imperial College
   London SW7 2BX, UK
*Corresponding author: nakano@ynu.ac.jp

(Manuscript received 30 September 2011; accepted 22 November 2011; published 15 January 2012)

Ultrathin-film interferometry was used to investigate the developing process of boundary films in a steel-on-silica contact lubricated by a solution of \( n \)-hexadecanoic acid dissolved in \( n \)-hexadecane. Immersion of the surfaces in the solution is not sufficient to initiate boundary film formation. However, upon a nominally pure rolling contact, boundary film formation is initiated, and, once initiated, the boundary film develops spontaneously. It is believed that the initial monolayer film adsorbed strongly on the solid surface induces subsequent molecular stacking from the solution. The developed boundary film can be partially destroyed by the rolling contact, although the nominally identical rolling contact acts as the trigger of boundary film formation. These findings reveal that the immobile boundary films to reduce friction of solids under high pressure result from the equilibrium of two competing processes: spontaneous development by molecular stacking and mechanical destruction by relative motion of the contact surfaces.

**Keywords:** boundary lubrication, boundary film, tribofilm formation, optical interferometry

1. Introduction

Oil-soluble surfactants, such as carboxylic acids and their salts, have been used for many years as friction modifiers in lubricating oils. Ever since their discovery and early practical applications around 1920 [1–3], their mechanisms of action have been debated [4]. Many researchers have suggested that they form physically or chemically adsorbed monolayer films on polar metal surfaces and that the films are able either to prevent asperity contacts or to limit adhesive junction growth at the periphery of these contacts. In fact, such a monolayer mode of action has been widely accepted as the basic model of boundary lubrication effects [5]. Meanwhile, a number of reports have advocated a multilayer mode of action. In low-pressure stationary contacts, it has been found that the additives spontaneously form a viscous film with a thickness of up to 1 \( \mu \)m or more on metal surfaces [6–10], possibly due to molecular orientation [11] as in liquid crystals or due to osmotic pressure [12].

Recently, it has become feasible to measure the thickness and some of the properties of thin lubricant films within contacts. One promising technique is ultrathin-film interferometry [13,14], which directly measures film thicknesses down to 1 nm or less in rolling-sliding steel-on-silica contacts under high pressure. This technique was applied to examine the film forming properties of, for example, viscosity index improvers [15] and anti-wear agents [16].

The ultrathin-film interferometry was then applied to examine the film forming properties of some friction modifiers [17]. For a copper carboxylate in mineral oil under a mixed rolling-sliding contact, a boundary film of up to 70 nm in thickness was observed. Just after the entrainment speed was reduced rapidly from a high value, the boundary film vanished but then started to develop once more and returned to its original level after some minutes. Upon halting the motion, most of the developed film remained, separating the stationary contact. Similar behaviors were also observed for other commercial friction modifiers. These results indicate that the boundary film formation is history-dependent and that the boundary lubrication using these friction modifiers is probably based on a multilayer mode of action.

However, some simple friction modifiers (e.g., fatty acids such as octadecanoic acid and oleic acid) just provided a thin boundary film (e.g., 2 nm thick) under rolling conditions [18]. In the presence of water, thick boundary films were formed, suggesting that the friction modifiers react chemically with the surfaces to produce...
insoluble iron carboxylate, which generates thick boundary films (e.g., 10–20 nm thick) [19]. The ultrathin-film interferometry has also been used to monitor the survival of multilayer Langmuir-Blodgett (LB) films of stearic acid in rolling contacts [20]. This showed that deposited layers of up to three monolayers could survive in high pressure rolling contact for considerable test time.

In this study, main attention was focused not on well-developed boundary films but on the developing process of boundary films. Using the ultrathin-film interferometry, we performed two types of tests: (i) the steady elastohydrodynamic (EHD) contact test and (ii) the stationary contact test. From these experiments for a fatty acid solution, we found that the immobile boundary films formed by the solution in the contact involve two competing processes: spontaneous development by molecular stacking and mechanical destruction by relative motion of the contact surfaces.

2. Experimental

2.1. Apparatus

To measure the thickness of thin films formed by lubricating fluids, ultrathin-film interferometry was used [13,14]. The setup is shown schematically in Fig. 1. A lubricated contact is formed between a loaded steel ball and a flat surface of a glass disk. The disk is driven and can drive the ball in nominally pure rolling on a ball carriage having three inclined bearings.

The glass disk surface in contact with the ball is coated with a thin semi-reflecting layer of chromium, on top of which is a spacer layer of silica approximately 500 nm thick. White light is shone into the contact, and some is reflected from the chromium layer, while some passes through the spacer layer and any lubricant film present, to be reflected from the steel ball. The two beams recombine and interfere, where the spacer layer ensures that interference will occur even if no lubricant film is present.

The interfered light from a strip across the contact is then passed into a spectrometer, where it is dispersed and detected by a CCD camera. A frame grabber captures this image at a preselected angular position of the rotating disk and a computer program determines the wavelength of maximum constructive interference in the central region of the contact. The lubricant film thickness is then calculated from the difference between the measured film thickness and the thickness of the silica spacer layer at the identical angular position of the disk.

2.2. Materials

In the experiments, commercially available 19-mm-diameter AISI 52100 steel balls of RMS surface roughness 11 nm were used. As a base fluid in the test lubricants, n-hexadecane was used. It was purified by being stored in a dark glass bottle with activated alumina and silica gel powder for over a week and then filtered prior to use. As an additive in the test lubricants, n-hexadecanoic acid was used as supplied.

2.3. Procedure

(i) Steady EHD contact test

To identify the boundary film formation of test lubricants, film thickness measurements were obtained from the central region of EHD contacts operating continuously under steady speed conditions using ultrathin-film interferometry.

Prior to the test, the ball, disk, and ball carriage were cleaned in toluene in an ultrasonic bath for 15 min and then rinsed first in ethanol and then in acetone. After the ball and disk were dried and installed, a few cm$^3$ of the test lubricant (n-hexadecane or solution of n-hexadecanoic acid dissolved in n-hexadecane) was applied to the contact surfaces. Then, the disk was rotated about its rotationally symmetric axis by a DC motor and a normal load of 20 N was applied to the ball. The rotating disk drove the ball in nominally pure rolling on the ball carriage. The film thickness at the contact center was measured using ultrathin-film interferometry over a series of increasing and decreasing entrainment speeds, until the repeatability of the measurements was confirmed.

(ii) Stationary contact test

To investigate the initial process of boundary film formation, a series of film thickness profile measurements were made across the stationary contact center using ultrathin-film interferometry, as shown in Fig. 2. Seven “stationary contact periods (A, B, C1–C4, and D)” in which film thickness profiles were measured, four “non-contact periods”, and two “rolling contact periods” were defined.

Prior to the test, the ball, disk, and ball carriage were cleaned, dried, and installed in the same manner as for the steady EHD contact test. In period A, the ball was loaded with 20 N and profiles of the dry stationary contact were collected over a time period of 1000 s at 100 s intervals. After separating the contact, a few cm$^2$ of...
test lubricant (10^{-1} mass% solution of \(n\)-hexadecanoic acid dissolved in \(n\)-hexadecane) was injected into the space between the ball and disk surfaces, and the separation was maintained for 2000 s. In period B, the ball was loaded with 20 N and the profile of the wet stationary contact was measured for 1000 s. After separating the contact, disk motion was started, and rolling contact (normal load: 20 N and entrainment speed: 10 mm/s) was maintained for 2000 s. Then, the contact was separated and the disk motion was stopped. Between periods C1 and C4, there were two cycles of a 2000-s non-contact period and a 1000-s stationary contact period (C2 or C3), and these were followed by a long non-contact period of 250000 s (i.e., approximately 70 h). Period C4 was followed by the second rolling contact period of 2000 s. After this, the contact was separated and the disk motion was stopped. Finally, in period D, the ball was loaded to 20 N and the profile of the wet stationary contact was measured for 1000 s. It should be noted that to avoid differences in spacer layer thicknesses, all stationary contacts were made at the same position of the disk surface.

All the above tests (i) and (ii) were carried out at room temperature (i.e., approximately 20–25°C).

3. Results

3.1. Steady EHD contact test

Figure 3 shows logarithmic film thickness versus entrainment speed plots for \(n\)-hexadecane and \(n\)-hexadecanoic acid solution, obtained from a pure rolling contact by using ultrathin-film interferometry. The additive-free \(n\)-hexadecane behavior is represented by “0 mass%”. This is linear, in accord with the prediction of the classical theory of EHD lubrication:

\[ h_c = aU^{0.67} \]  

(1)

where \(h_c\) is the film thickness at the contact center, \(U\) is the entrainment speed, and \(a\) is a constant. The plot remains linear even down to 1 nm film thicknesses, indicating that the viscosity of the fluid being entrained remains constant down to this distance from the surface. This therefore implies that \(n\)-hexadecane gives no measurable boundary films in the contact region.

The behaviors observed on adding small amounts of \(n\)-hexadecanoic acid to \(n\)-hexadecane are represented by plots “10^{-3} mass%” and “10^{-1} mass%” in Fig. 3. The deviation in the \(n\)-hexadecanoic acid solutions from the additive-free \(n\)-hexadecane indicates that thin boundary films of a few nanometers in thickness are formed in the contact region.

The curves in Fig. 3 are fitted to the following equation:

\[ h_c = aU^{0.67} + b \]  

(2)

The values of the fitting constants \(a\) and \(b\) determined using the least-squares method are listed in Table 1, where the bracketed values were treated as invariants in the least-squares method. The high values of the determination coefficient \(R^2\) validate Eq. (2), confirming that the sum of the thicknesses of the “immobile boundary films” formed on the two surfaces results in constant \(b\) (i.e., independent of \(U\)) and also that the behavior of the “fluid film” between the immobile boundary films is in absolute agreement with the additive-free \(n\)-hexadecane behavior. This type of behavior is similar to that when \(n\)-hexadecane is applied to an LB monolayer film deposited on the disk surface.
of the same rolling contact [20].

The boundary film thickness of \( b = 3.0 \) nm for \( 10^{-4} \) mass\% is twice as large as that for \( 10^{-3} \) mass\%. The dependence of boundary film thicknesses on solution concentrations suggests that both positive and negative effects on the boundary film thickness are involved in the rate process occurring in the boundary film formation. It is possible that the difference in the thickness came from the difference in the formation state of the boundary films such as a density and/or an orientation of the adsorbed molecules.

3.2. Stationary contact test

3.2.1 Measured profiles and modified profiles

Figure 4 shows a measured profile \( z(x,t) \) (where \( x \) is the radial position from the contact center and \( t \) is the time) at \( t = 100 \) s in period A (in the dry stationary contact before injecting the test lubricant). Despite the absence of the lubricating fluid in the contact area, the measured profile is non-zero and is convex in a downward direction. This is caused by the inhomogeneous deformation of the spacer layer in the normal direction, due to the pressure profile of the Hertzian contact. According the Hertz theory, the pressure at the contact center \( (x = 0) \) is 0.53 GPa, whereas that at the contact edge \( (x = \pm 135 \) \( \mu m \)) is zero.

The solid line in Fig. 4 shows a function with the following form:

\[
z_0(x) = cx^4 + d
\]  

(3)

Within the Hertzian contact zone \( |x| \leq 135 \) \( \mu m \), the values of the fitting constants \( c \) and \( d \) are determined as \( c = 3.61 \times 10^{-8} \) (nm)\(^4\) (\( \mu m \))^4 and \( d = -0.4 \) nm by the least-squares method, leading to a high determination coefficient \( R^2 = 0.963 \). Consequently, the modified profiles \( h(x,t) \)

\[
h(x,t) = z(x,t) - z_0(x)
\]  

(4)

are the time-varying profiles of boundary film thicknesses. This assumes that the boundary films formed on the surfaces do not affect the normal deformation of the spacer layer.

3.2.2 Change in profiles of boundary film thicknesses

Figure 5 shows the profiles \( h(x,t) \) at \( t = 1000 \) s (represented by solid circles) and \( t = 10000 \) s (represented by open circles) in each period A–D. The data at \( |x| > 100 \) \( \mu m \) were discarded because a slight difference between the origins \( (x = 0) \) of \( z(x,t) \) and \( z_0(x) \) can cause a large error of \( h(x,t) \) in \( |x| > 100 \) \( \mu m \).

In period A under the dry condition, the profile deviates slightly from the dashed line \( h = 0 \); this deviation is believed to be due to the composite surface roughness of the ball and disk in contact with each other. In period B, the situation is quite similar to period A, although it is under the wet condition. In period C1, however, a nominally flat boundary film of approximately 2 nm thickness forms; this is caused by the rolling contact between periods B and C1. From periods C1 to C4, the boundary film develops sequentially, and the profile increases to approximately 5 nm in period C4. Finally, in period D, the thickness of the boundary film decreases to approximately 2 nm; this is caused by the rolling contact between periods C4 and D, although the nominally identical rolling contact increases the boundary film thickness from periods B to C1.

A comparison between the profiles at \( t = 1000 \) s and \( t = 10000 \) s shows that in periods C1–C4, the latter tends to be located slightly below the former.

3.2.3 Change in mean film thickness in each period

The instantaneous mean film thickness was calculated from \( h(x,t) \) in the range \( |x| \leq 100 \) \( \mu m \), and the temporal change in the mean film thickness for each period is shown in Fig. 6.

The deviation in period A from the dashed line \( h = 0 \) is the measurement error, possibly caused by the unexpected vibrations of the apparatus. From this deviation, the RMS error of the mean film thickness measurement is found to be less than 0.1 nm. In period B, although no significant boundary film is formed, the mean film thickness appears to decrease gradually from 0.1 nm. It is possible that the decrease represents a molecular flow to outside the contact region, caused by the pressure gradient; however, as the decrease is comparable to the measurement error, this is difficult to quantify. In periods C1–C4, a gradual but considerable decrease in the mean film thickness is found; the decrease rate is approximately \( 5 \times 10^{-4} \) nm/s. In period D, however, the decrease is unclear; the mean film appears to deviate around the mean value \( h = 2.2 \) nm. This may indicate that some mechanical properties in the boundary film changed after the rolling contact between periods C4 and D.

Fig. 4 Measured profile at \( t = 100 \) s in period A; open circle: measured heights \( z(x,t) \), solid line: base line \( z_0(x) \) determined using the fitting function within the Hertzian contact zone (see Eq. (3))
Fig. 5 Profiles of boundary film thicknesses $h(x,t)$; solid circles: modified heights at $t = 100$ s, open circles: modified heights at $t = 1000$ s in each period (see Eq. (4))

Fig. 6 Temporal changes in mean film thicknesses at stationary contact periods under loading of 20 N
4. Discussion

The boundary film observed in this study reduces friction in boundary lubrication regimes because the test lubricant (10⁻¹ mass% solution of n-hexadecanoic acid in n-hexadecane) provides a favorable boundary lubrication state. For example, using a mini-traction machine [4], the friction coefficient for a steel-on-steel contact lubricated with the test lubricant was found to be 0.1 (under normal load of 20 N, entrainment speed of 0.01 m/s, slide-roll ratio of 50%, and at room temperature), while it was unable to measure the friction coefficient for the same pair lubricated with the additive-free n-hexadecane, as serious material seizure occurred immediately in the contact.

The boundary film thickness measurements obtained in this study are summarized in Fig. 7, where the open histograms are the mean film thicknesses at 1000 s in periods A–D (see Fig. 6) of the stationary contact test and the solid histogram is the immobile film thickness b obtained in the steady EHD contact test (see Table 1). These experimental results reveal important aspects concerning boundary film formation as follows.

First, immersion of the solid surfaces in the test lubricant is insufficient to initiate boundary film formation (see periods A–B); the boundary film is initiated only upon rolling contact and the boundary film formed between periods B–C1 separates the contacts under high applied pressures (maximum pressure is 0.53 GPa). It is believed that mechanical rubbing associated with the initial rolling contact under no boundary film provides situations needed to initiate some chemical reactions of fatty acid molecules with the metal surface, leading to a monolayer film adsorbed strongly on the surface [5,19]. Similar experiments with no rolling contact periods confirmed that the stationary contact alone is insufficient to initiate boundary film formation.

Secondly, once the boundary film formation initiates, the boundary film develops spontaneously under no contact (see periods C1–C4). This indicates that the initial monolayer film adjacent to the solid surface plays the role of a template layer, on top of which additional fatty acid molecules can easily be stacked. This spontaneous development continues after the thickness reaches 3 nm, which is the thickness of the immobile boundary film formed in the steady EHD contact. It is believed that the formed film consists of not only the solute (n-hexadecanoic acid) but also the solvent (n-hexadecane) as it was confirmed under an extremely low pressure in a steel-on-mercury contact [10]; however, ordered molecular orientations perpendicular to the contact surfaces (e.g., like LB films [20]) is not expected because no regularity of boundary film thickness involving molecular lengths was observed.

Thirdly, the well-developed boundary film can be partially destroyed by rolling contact (see periods C4–D), although the nominally identical rolling contact acts as the trigger of boundary film formation. The film thickness after the destruction is less than that of the immobile boundary film formed in the steady EHD contact (i.e., 3 nm). It is believed that the well-developed boundary film in the stationary contact test contains some defects that allow the mechanical rubbing associated with the rolling contact to easily break the boundary film.

It should be noted that, in the steady EHD contact test, the time during which a point on the rolling ball surface experienced high contact pressure was only approximately 0.5% of the whole time because the circumferential length of the ball was approximately 60 mm and the Hertzian contact diameter was approximately 0.3 mm. This indicates that, even in the steady EHD contact test, the boundary film formation progressed under no contact as in the stationary contact test.

Consequently, we can conclude that the immobile boundary films to reduce friction of solids under high pressure involve two competing processes: spontaneous development by molecular stacking and mechanical destruction by relative motion of the surfaces. It is possible that the mechanical action has a positive effect on reducing the defects in the boundary film to make it tougher, rather as a heavy roller improves the condition of a clay tennis court. Further studies are needed to clarify such effects of the mechanical action and explore how it changes the mechanical properties of boundary films; which are inferred to be related to the processes that persist until the friction coefficient is stabilized in the boundary lubrication regime.

5. Conclusions

(1) A stationary contact immersed in a solution of n-hexadecanoic acid in n-hexadecane is insufficient to initiate boundary film formation. A
rolling contact acts as a trigger of boundary film formation to separate the contacts under high pressure.

(2) Once the boundary film formation initiates, the boundary film develops spontaneously. It is believed that the initial monolayer film adsorbed on the solid surface induces subsequent molecular stacking from the solution.

(3) The spontaneously developed boundary film can be partially destroyed by the rolling contact, although the nominally identical rolling contact acts as the trigger of boundary film formation.

(4) The immobile boundary films to reduce friction of solids under high pressure result from the equilibrium of two competing processes: spontaneous development by molecular stacking and mechanical destruction by relative motion of the contact surfaces.

References