General Disclaimer

One or more of the Following Statements may affect this Document

- This document has been reproduced from the best copy furnished by the organizational source. It is being released in the interest of making available as much information as possible.

- This document may contain data, which exceeds the sheet parameters. It was furnished in this condition by the organizational source and is the best copy available.

- This document may contain tone-on-tone or color graphs, charts and/or pictures, which have been reproduced in black and white.

- This document is paginated as submitted by the original source.

- Portions of this document are not fully legible due to the historical nature of some of the material. However, it is the best reproduction available from the original submission.

Produced by the NASA Center for Aerospace Information (CASI)

TWO-DIMENSIONAL RANDOM SURFACE
MODEL FOR ASPERITY-CONTACT IN
ELASTOHYDRODYNAMIC LUBRICATION

John J. Coy
Propulsion Laboratory
AVRADCOM Research and Technology Laboratories

and

Steven M. Sidik
Lewis Research Center
Cleveland, Ohio

TWO-DIMENSIONAL RANDOM SURFACE MODEL FOR

ASPERITY-CONTACT IN ELASTOHYDRODYNAMIC

LUBRICATION

John J. Coy
Propulsion Laboratory
AVRADCOM Research and Technology Laboratories

and

Steven M. Sidik
Lewis Research Center
National Aeronautics and Space Administration
Cleveland, Ohio

ABSTRACT

Relations for the asperity-contact time fraction during elastohydrodynamic lubrication of a ball bearing are presented. The analysis is based on a two-dimensional random surface model, and actual profile traces of the bearing surfaces are used as statistical sample records. The results of the analysis show that transition from 90 percent contact to 1 percent contact occurs within a dimensionless film thickness range of approximately four to five. This thickness ratio is several times larger than reported in the literature where one-dimensional random surface models were used. It is shown that low pass filtering of the statistical records will bring agreement between the present results and those in the literature.

INTRODUCTION

Elastohydrodynamic lubrication (EHD) is the term used to describe part of the technology concerning lubrication of concentrated mechanical contacts. In essence, EHD lubricant film formation depends on the coupled effects of physical changes in the lubricant, which are caused by high pressures in the Hertzian contact area, and elastic changes in the shape of the Hertzian contact area, which affect the pressure distribution. The high pressures in the EHD contact area act to squeeze out the lubricant. However, the lubricant becomes thicker (more viscous) with increasing pressure and resists being squeezed out. The net result is

the formation of a thin lubricant film that is beneficial in preventing seizure and rapid wear of the contacting parts (refs. 1 to 4). For many applications the EHD film thickness is the same order of magnitude as the surface rms roughness. Experimental measurement of the film thickness is very difficult because films are so thin. Various methods that have been used are optical (interferometry), X-ray, and electrical capacitance and conductance techniques (ref. 5). Of the aforementioned measurement methods, the capacitance and conductance methods are most suited to measurement of film thickness in full-scale bearings (refs. 6 and 7). The conductance method of measurement depends on having a known relation between film thickness and contact time fraction. The contact time fraction is directly related to the normalized average voltage observed when a low voltage is applied across the lubricant film.

In 1964, Tallian and his coworkers (ref. 7) formulated a statistical model of bearing surface roughness and used the model to infer EHD film thicknesses, based on electrical conductance measurements. Their results were applicable to the regime of "partial EHD contact," where the load is shared by the EHD film and the high points or asperities of the metal surfaces that momentarily interrupt the lubricant film (ref. 8).

By the early 1970's it was generally accepted that asperity contact must be viewed as a random process (refs. 8 to 14). Most researchers used stylus traces of the surface to obtain profile statistics for the random process models. In 1971, Nayak (ref. 14) explained how Longuet-Higgens' theory of ocean surfaces (refs. 15 to 17) could be used to model rough surfaces as two-dimensional, isotropic, Gaussian random processes. He showed that significant differences exist between surface statistics and profile statistics and that a naive analysis assuming that profile statistics may be directly used is erroneous (refs. 14 and 18). Sidik has extended the theory of Nayak to obtain a model for asperity-contact time fraction as a function of film thickness in partial EHD contact lubrication (ref. 19). In reference 20, the theory is generalized to nonisotropic Gaussian surfaces. In reference 21 the computational methods are developed.

The objective of this paper is to apply the relevant results of two-dimensional random surface analysis to obtain a relation between asperity-contact time fraction and average EHD film thickness for a typical ball bearing and to compare the results with those expressed in references 6 and 11. Also, recent work by Sayles and Thomas (ref. 22) is used to correlate the results.

Asperity-Contact Model

The ball bearing for which this analysis was performed has a 20-millimeter bore and three 7.15-millimeter- (9/32-in.-) diameter balls. Three different thrust loads were considered in the analysis of contact time fraction. Table I gives the calculated Hertzian stresses and contact ellipse dimensions corresponding to the different loads.

Under loaded conditions, assume that the ball and race surfaces are two-dimensional ergodic Gaussian processes and that within the Hertzian contact zone the mean planes are parallel and separated by a lubricant film of thickness h. A cross section of a single ball-race contact is presented in figure 1. Coordinate x is in the direction of rolling. The ball surface is denoted by $z_b(x,y)$ and the race surface by $z_r(x,y)$. The two processes z_b and z_r are independent, with mean levels $\mu_b = 0$ and $\mu_r = 0$, correlation functions R_b and R_r, and variances σ_b^2 and σ_r^2, respectively.

The composite process $z = z_b + z_r$ is also an ergodic Gaussian process with mean zero and correlation function R and variance σ^2, where

$$\sigma^2 R(\tau_x, \tau_y) = \sigma_b^2 R_b(\tau_x, \tau_y) + \sigma_r^2 R_r(\tau_x, \tau_y) \tag{1}$$

With this notation, then (as shown in fig. 2) any metallic contact occurrence is represented by the composite surface rising above the level h.

An approximation to the time fraction during which there is metallic contact anywhere within the Hertzian zone will now be derived. Consider the process z above the x-y plane. At a level h above the reference plane, pass a cutting plane that will occasionally intersect z. The sets of points in the reference plane where $z(x,y) \geq h$ are called excursion sets. Such excursion sets are represented as the crossed areas in figure 3. Superimposed upon this plane is an elliptical region that represents the Hertzian contact area. For constant rolling velocity, this elliptical region moves to the right at a constant velocity v through the region bounded by the parallel dashed lines y_1 and y_2. At the termination of a test period of time T the Hertzian area is at the elliptical region at the right in figure 3.

If it is assumed that $\Lambda = h/\sigma$ is large there will be few excursions of z above h and, hence, few metallic contacts. The contacts will be small in area, and the probability of two or more contacts in a small area is negligible. The contact occurrence is as follows: The dashed ellipse on the left in figure 3 represents the location of the Hertzian area

when the contact is first made. The dashed ellipse on the right represents the location of the Hertzian area when the contact is broken. The two points $P(x_M, y_M)$ and $P(x_B, y_B)$ denote coordinates of the make-contact and break-contact occurrences. The distance of $P(x_B, y_B)$ to the centerline of the ellipse on the right is termed L_B. Thus, contact exists for a total distance L that is composed of three parts. Two of these parts are L_M and L_B; the third is termed X and is the distance $x_M - x_B$. It is assumed that X is negligible with respect to L_M and L_B and the excursions are uniformly distributed with respect to the y-axis. Also L_M and L_B are approximately equal. As a result, the statistical expectation of the contact distance L is

$$E\{L\} \cong 2E\{L_M\} = \frac{\pi \ell}{2} \qquad (2)$$

where $\pi \ell / 2$ is the average length of the Hertzian ellipse. The expected total contact time $E\{T^*\}$ can be approximated by the product of the average number of excursions and the average time of contact for each. Equivalently, $E\{T^*\}$ can be approximated by the expected number of excursions per unit area $E\{\chi\}$ times the area rolled over, times the average time of contact for each. If T is large, the rolled-over area is approximately vTw, so that

$$E\{T^*\} = E\{\chi\} \, (vTw) \left(\frac{\pi \ell}{2v}\right) \qquad (3)$$

The expected contact time fraction is obtained as

$$E\{T_c\} = \frac{E\{T^*\}}{T} = \frac{\pi \ell w}{2} E\{\chi\} \qquad (4)$$

Thus, $E\{T_c\}$ as a function of Λ can be calculated directly from a computation of $E\{\chi\}$ as a function of Λ. The determination of $E\{\chi\}$ is presented next.

Following the method of Nayak (ref. 14), let the following variables $\xi_i (i = 1, 6)$ be defined:

$$\xi_1 \quad \xi_1(x,y) \equiv \frac{\partial^2}{\partial x^2} z(x,y)$$

$$\xi_2 = \xi_2(x,y) \equiv \frac{\partial^2}{\partial x \, \partial y} z(x,y)$$

$$\xi_3 = \xi_3(x,y) \equiv \frac{\partial^2}{\partial y^2} z(x,y)$$

$$\xi_4 = \xi_4(x,y) \equiv z(x,y) \qquad\qquad (5)$$

$$\xi_5 = \xi_5(x,y) \equiv \frac{\partial}{\partial x} z(x,y)$$

$$\xi_6 = \xi_6(x,y) \equiv \frac{\partial}{\partial y} z(x,y)$$

The expected number of summits of height ξ_4 within a unit area is given by the triple integral

$$f(\xi_4) = \frac{1}{|\mathcal{F}|^{1/2}|S|^{1/2}(2\pi)^3} \int\!\!\int\limits_V\!\!\int \left|\xi_1\xi_3 - \xi_2^2\right| e^{-(1/2)Q} \, d\xi_1 \, d\xi_2 \, d\xi_3$$

$$(6)$$

where the region of integration V is defined by

$$\left.\begin{aligned} \xi_1 &\leq 0 \\[6pt] \xi_3 &\leq 0 \\[6pt] \xi_1\xi_3 - \xi_2^2 &\geq 0 \end{aligned}\right\} \qquad\qquad (7)$$

and $\mathcal{F}$ and S by the moments of the power spectral density (see appendix)

$$\mathcal{F} = \begin{bmatrix} m_{40} & m_{31} & m_{22} & -m_{20} \\ m_{31} & m_{22} & m_{13} & -m_{11} \\ m_{22} & m_{13} & m_{04} & -m_{02} \\ -m_{20} & -m_{11} & -m_{02} & m_{00} \end{bmatrix} \tag{8}$$

$$S = \begin{bmatrix} m_{20} & m_{11} \\ m_{11} & m_{02} \end{bmatrix} \tag{9}$$

$$Q = (\xi_1, \xi_2, \xi_3, \xi_4) \mathcal{F}^{-1} (\xi_1, \xi_2, \xi_3, \xi_4)^T \tag{10}$$

Equation (6) is transformed to cylindrical coordinates followed by a rotation. The transformation equations are

$$\xi_1 = \frac{r - \rho \cos \varphi}{\sqrt{2}}$$

$$\xi_2 = \frac{\rho \sin \varphi}{\sqrt{2}}$$

$$\left. \begin{array}{l} \\ \\ \\ \end{array} \right\} \tag{11}$$

$$\xi_3 = \frac{r + \rho \cos \varphi}{\sqrt{2}}$$

$$\xi_4 = \Lambda m_{00}^{1/2}$$

It now becomes evident that the region V describes a semi-infinite cone, as shown by the limits of integration on the transformed equation.

$$f(\Lambda) = \frac{1}{|\mathcal{F}|^{1/2} |S|^{1/2} (2\pi)^3} \int_{r=0}^{\infty} \int_{\rho=0}^{r} \int_{\varphi=0}^{2\pi} \frac{(r^2 - \rho^2)\rho}{2^{3/2}} e^{-(1/2)Q} \, d\varphi \, d\rho \, dr \tag{12}$$

$$Q = \begin{bmatrix} \dfrac{r - \rho\cos\varphi}{\sqrt{2}} \\[2em] \dfrac{\rho\sin\varphi}{\sqrt{2}} \\[2em] \dfrac{r + \rho\cos\varphi}{\sqrt{2}} \\[2em] \Lambda \end{bmatrix}^T \begin{bmatrix} m_{40} & m_{31} & m_{22} & \dfrac{-m_{20}}{m_{00}^{1/2}} \\[1.5em] m_{31} & m_{22} & m_{13} & \dfrac{-m_{11}}{m_{00}^{1/2}} \\[1.5em] m_{22} & m_{13} & m_{04} & \dfrac{-m_{02}}{m_{00}^{1/2}} \\[1.5em] \dfrac{-m_{20}}{m_{00}^{1/2}} & \dfrac{-m_{11}}{m_{00}^{1/2}} & \dfrac{-m_{02}}{m_{00}^{1/2}} & 1 \end{bmatrix} \begin{bmatrix} \dfrac{r - \rho\cos\varphi}{\sqrt{2}} \\[2em] \dfrac{\rho\sin\varphi}{\sqrt{2}} \\[2em] \dfrac{r + \rho\cos\varphi}{\sqrt{2}} \\[2em] \Lambda \end{bmatrix} \qquad (13)$$

Equation (12) was evaluated numerically on the digital computer. The expected number of summits per unit area D_{sum} is given by the integral

$$D_{sum} = \int_{-\infty}^{\infty} f(\Lambda)d(\Lambda) \qquad (14)$$

The probability density for summit heights is given by the ratio

$$p^*(\Lambda) = \frac{f(\Lambda)}{D_{sum}} \qquad (15)$$

The expected number of excursions above level Λ per unit area is approximated by the product of peaks per unit area and the proportion of such peaks that exceed the level Λ.

$$E\{\chi(\Lambda)\} \approx D_{sum} \int_{\xi=\Lambda}^{\infty} p^*(\xi)d\xi = \int_{\xi=\Lambda}^{\infty} f(\xi)d\xi \qquad (16)$$

This approximation is valid only in the limiting sense as $\Lambda \to \infty$.

Adler and Hasofer (ref. 23) also provide an approximation for the upcrossings of a process $z(x,y)$ over the level Λ. In terms of the present notation, their results are given by the relation

$$E\{\chi(\Lambda)\} \approx \frac{\Lambda|S|^{1/2} \exp\left(-\frac{\Lambda^2}{2}\right)}{(2\pi)^{3/2}\sigma^2} \tag{17}$$

It is interesting to note that this expression does not involve any fourth-order moments.

The derivation to this point provides contact time fraction as a function of h for a single Hertzian contact. Next, the results are applied to a ball bearing with three balls for three different loads. It is assumed that at each of the six ball-race contacts the mean film thickness is the same and that each contact is statistically independent of the others. Because of the geometry of the balls and the race, however, the nominal Hertzian areas at the inner and outer race contacts are different. Table I presents the calculated conditions at these contacts for three different loads. From equation (4) it is evident that $E\{T_c\}$ is simply the area of the Hertzian contact times $E\{\chi\}$.

Let $T_{c,in}$ and $T_{c,out}$ denote the expected contact time fractions at the inner and outer races for a single ball. Thus the probability of no contact on a single ball is $1 - T_{c,in}T_{c,out}$. From the independence assumption, the probability of no contact on any of the three balls is the quantity

$$T_o = 1 - T_{c,ov} = (1 - T_{c,in}T_{c,out})^3 \tag{18}$$

Surface Measurements

The first step in determining the expected number of excursions per unit area is to obtain and analyze surface profile traces from the bearing surfaces. By using these profile traces, the important surface parameters are computed as outlined in the previous section. As explained in the appendix, in order to characterize the surface statistics, one must first obtain profile traces in at least three different directions. The surfaces of the ball and race specimens were traced to obtain records of the surface profile. Several traces in different directions on the ball surface showed that it was an isotropic surface; the race surface was nonisotropic.

It was not possible to obtain a sample record on the race surface in any direction other than the rolling direction and the cross groove direction. In order to obtain the necessary additional traces, a flat specimen was prepared by material and finishing methods identical to those

used in making the bearing race. The flat specimen was approximately 2.5 by 5.0 centimeters (1 by 2 in.) with a 0.13- to 0.25-micrometer (5- to 10-μin.) CLA surface finish. Similarly, a surrogate ball, 14.3 millimeters (9/16 in.) in diameter with a 0.03- to 0.05-micrometer (1- to 2-μin.) CLA surface finish, was used to obtain sample ball records of sufficient length.

Figures 4 and 5 show the microtopography of the ball and the flat specimen. Traces to be analyzed were recorded for three different directions on the ball and six different directions on the flat specimen. Nominally, the traces were taken 45° apart for the ball and 18° apart for the flat specimen. The first trace on the flat specimen was taken in the direction of the lay of the surface finish. This is the rolling direction for the ball in the race. Five more traces were taken, with the last trace being at 90° to the lay of the surface finish. The traces were recorded in analog form as an FM signal on magnetic tape. This tape was sampled at equal intervals and written in digital format on another tape. The sampling intervals and the number of sampled points, along with the total sampled length, are presented in table II.

Each of the digitized profile traces was processed by a computer program that performed the following steps:

Plotted input. - All the plots were examined for faithful reproduction of the original record and were visually found to be identical.

Moving-average trend removal. - The raw data required detrending for two reasons. First, the stylus head does not follow a path parallel to the mean line of the profile when tracing. This causes a linear trend. The other reason for detrending was to anticipate the detrending that occurs naturally in the lubrication process. Therefore, it was decided to remove trends with wavelengths longer than the Hertzian contact. The moving-average trend remover, which is essentially a high-pass filter, was effective in removing these trends.

The number of points in the moving average is a function of the sample interval and the load since the contact ellipse dimensions change with load. The number of points for each load and each profile are given in table III.

Normal probability plot. - A simple graphical test for normality is to plot the sample cumulative distribution function on Gaussian probability paper. The plots indicate general agreement with Gaussian distribution (fig. 6).

Estimated spectral moments. - Spectral moments are estimated by the variance of the derivatives where the derivatives are approximated by the finite difference method. Table IV presents the estimated moments for the three load conditions.

Adding ball moments to flat profile moments. - According to the definitional equatioi. for spectral moments (eq. (A8)), the spectral moments of a profile of the composite surface are simply the sums of the appropriate individual surface moments.

Estimating spectral moments by least squares. - The following relations between the two-dimensional moments m_{ij} and the profile moments $m_{n,\theta}$ are obtained from equation (A9).

$$m_{0,\theta} = m_{00} = \sigma_\theta^2 \tag{19}$$

$$m_{2,\theta} = m_{20} \cos^2 \theta + 2m_{11} \cos \theta \sin \theta + m_{02} \sin^2 \theta \tag{20}$$

$$m_{4,\theta} = m_{40} \cos^4 \theta + 4m_{31} \cos^3 \theta \sin \theta + 6m_{22} \cos^2 \theta \sin^2 \theta$$

$$+ 4m_{13} \cos \theta \sin^3 \theta + m_{04} \sin^4 \theta \tag{21}$$

Equation (19) implies that the best estimator for m_{00} is simply the average of the $m_{0,\theta}$. The results for m_{00} are 5.97×10^{-2}, 6.41×10^{-2}, and 5.90×10^{-2} square micrometer, respectively, for loads of 90, 445, and 3100 newtons (20, 100, and 700 lbf).

Equation (20) provides one equation for each θ, or a total of seven equations in three unknowns. For each load the estimates for m_{20}, m_{11}, and m_{02} are given here in a matrix format, emphasizing that this represents the covariance matrix of φ_x and φ_y in the spectral density function. The matrices are nearly diagonal, which means that the distributions of φ_x and φ_y are effectively uncorrelated. For the 90-newton (20-lbf) load,

$$\begin{bmatrix} m_{20} & m_{11} \\ m_{11} & m_{02} \end{bmatrix} = \begin{bmatrix} 2.39\times10^{-3} & 3.4\times10^{-5} \\ 3.4\times10^{-5} & 5.73\times10^{-3} \end{bmatrix} \tag{22}$$

For the 445-newton (100-lbf) load,

$$\begin{bmatrix} m_{20} & m_{11} \\ m_{11} & m_{02} \end{bmatrix} = \begin{bmatrix} 2.22\times10^{-3} & 1.31\times10^{-4} \\ 1.31\times10^{-4} & 5.58\times10^{-3} \end{bmatrix} \tag{23}$$

For the 3100-newton (700-lbf) load,

$$
\begin{bmatrix} m_{20} & m_{11} \\ m_{11} & m_{02} \end{bmatrix} = \begin{bmatrix} 2.17\times10^{-3} & -4.4\times10^{-5} \\ -4.4\times10^{-5} & 5.66\times10^{-3} \end{bmatrix}
\tag{24}
$$

Equation (21) provides one equation for each θ, hence, seven equations in the five unknowns m_{40}, m_{31}, m_{22}, m_{13}, and m_{04}. For each load the estimates of these moments are as follows: For the 90-newton (20-lbf) load,

$$
\begin{bmatrix} m_{40} & m_{31} & m_{22} \\ m_{31} & m_{22} & m_{13} \\ m_{22} & m_{13} & m_{04} \end{bmatrix} = \begin{bmatrix} 6.65\times10^{-2} & 3.88\times10^{-4} & 1.89\times10^{-2} \\ 3.88\times10^{-4} & 1.89\times10^{-2} & 2.14\times10^{-3} \\ 1.89\times10^{-2} & 2.14\times10^{-3} & 7.27\times10^{-2} \end{bmatrix} \mu m^{-2}
\tag{25}
$$

For the 445-newton (100-lbf) load,

$$
\begin{bmatrix} m_{40} & m_{31} & m_{22} \\ m_{31} & m_{22} & m_{13} \\ m_{22} & m_{13} & m_{04} \end{bmatrix} = \begin{bmatrix} 6.61\times10^{-2} & 3.13\times10^{3} & 1.55\times10^{-2} \\ 3.13\times10^{3} & 1.55\times10^{-2} & 4.39\times10^{-3} \\ 1.55\times10^{-2} & 4.39\times10^{-3} & 7.20\times10^{-2} \end{bmatrix} \mu m^{-2}
\tag{26}
$$

For the 3100-newton (700-lbf) load,

$$
\begin{bmatrix} m_{40} & m_{31} & m_{22} \\ m_{31} & m_{22} & m_{13} \\ m_{22} & m_{13} & m_{04} \end{bmatrix} = \begin{bmatrix} 6.49\times10^{-2} & 4.50\times10^{-3} & 1.51\times10^{-2} \\ 4.50\times10^{-3} & 1\,51\times10^{-2} & 5.33\times10^{-3} \\ 1.51\times10^{-2} & 5.33\times10^{-3} & 7.05\times10^{-2} \end{bmatrix} \mu m^{-2}
\tag{27}
$$

RESULTS AND DISCUSSION

The final result of this analysis was obtained by using the moment estimates for the two-dimensional process to calculate the expected number of excursions per unit area (eqs. (12) to (16)) for various film thickness ratios. These expectations were then used to obtain the no-contact time fraction T_0 (eq. (18)). The results are presented in table V. Curves of no-contact time fraction against film thickness ratio are plotted in figure 7. The calculations for these results required some 30 hours of computer time.

For comparison, values of $E\{\chi\}$ were calculated from the spectral moments in table IV by using the method of Adler and Hasofer (eq. (17)). The resulting no-contact time fraction for the 90N (20 lb) load is also plotted in figure 7 for comparison. The spacing between curves for the 445N and 3100N loads was similar, but are not shown on figure 7. The Adler and Hasofer method gives curves that are shifted toward lower values of Λ by an amount approximately equal to 0.35. The reason for this difference is unclear. Both methods should be asymptotically equivalent, and values of Λ greater than 4 should be sufficiently large for the asymptotic results to hold. Adler and Hasofer's "upcrossings" seem closer to the required "excursion sets" in the development of contact time fraction than the approximation based on peak height distribution. Nevertheless both are approximations.

Measurements of the no-contact time fraction by Tallian, et al. (ref. 7) on a 4-ball apparatus and by Poon and Haines (ref. 12) on a point contact disc machine are shown in figure 7 for comparison. The measurements indicate that the no-contact time fraction begins to increase from zero at $\Lambda \approx 1$ and monotonically increases until there is 100-percent film at $\Lambda \approx 3.5$. By comparison the results of this investigation show this change in contact time fraction occurring for film thicknesses several times larger. Also, the incremental change in film thickness corresponding to the incremental change from no-contact to 100-percent contact is smaller.

The theoretical differences in probability distributions for peak heights of summits on a two-dimensional surface and peak heights on profile traces along a fixed direction on the surface may partially account for these differences. However, the recent work of Sayles and Thomas (ref. 22) provides much more insight. They have shown that the asperity density and spectral moments are strongly dependent on the surface sample interval. Their work shows that an order of magnitude decrease in sample interval causes approximately an order of magnitude increase in the second-order moments and two orders of magni-

tude increase in the fourth-order moments. Also, as the sample interval becomes smaller than 10 μm, the summit density seems to increase without approaching an asymptote. Therefore, it was decided to compare the present asperity-contact model to several models in the literature within the framework of effects of asperity summit density and order of magnitude changes in the spectral moments. Basically these are the effects of low pass filtering of the surface traces.

The surface traces used for the work in this paper were sampled at 0.31 μm intervals which was sufficient to accurately reproduce the analog traces. It is not practical at this time to filter and reprocess the original surface traces. However, the effect of filtering can be produced by using Adler and Hasofer's result with the second-order moments reduced by one or two orders of magnitude. The result of this procedure is shown in figure 7. The effect is to bring the T_0 curve closer to the result generally reported in the literature.

Curves were calculated using the method of Johnson, Greenwood, and Poon (ref. 11), and Kannel and Snediker (ref. 6) which is based on an earlier paper by Greenwood (ref. 9). Reference 11 assumes that there are 20,000 summits/cm^2 and reference 6 assume 0,000 summits/ cm^2 where the assumption was "used to obtain a reasonable fit with the experimental results" (ref. 11) which were reported by Tallian (ref. 7) and Poon and Haines (ref. 12).

Results were also obtained by substituting the values of $E\{\chi\}$ from table V into Johnson, Greenwood, and Poon's Poisson process model. In addition, results were obtained using the Kannel and Snediker model but it is now assumed that there are $\sim 10^8$ summits/cm^2. The results were identical for the two methods and there was exact agreement with the method of Coy and Sidik near the full film conditions as shown in figure 7.

From the comparisons presented it is clear that the main reason for the disparity in results is due to the effects of sample interval which is equivalent to low pass filtering.

Some comments are needed regarding the usefulness of these curves as a means of measuring film thickness by the electrical conductance method. The results presented show the incremental change in film thickness corresponding to the incremental change from no contact to 100-percent contact to be small. Therefore, the usefulness of the model is limited to a narrow range of film thickness for any given constant load. In addition the location of the curves is a strong function of surface sampling interval. At this time, there is no known rational way to choose the sampling interval consistent with the asperity-contact phenomenon in elastohydrodynamic lubrication processes.

CONCLUSIONS

Relations for asperity-contact time fraction as a function of nominal elastohydrodynamic film thickness have been presented. The calculations were based on a two-dimensional random surface model. Results were obtained for a 20-millimeter-bore ball bearing with three 7.15-millimeter (9/32-in.) balls. Surface traces were obtained by using a profilometer, and a statistical analysis was performed in which the profile traces were used as statistical sample records. The results were compared with previously reported results based on simple one-dimensional models. The investigation showed that

1. Using the two-dimensional asperity-contact model, with surface traces at small sample interval, the no-contact time fraction varied from almost full contact (90 percent) to almost no contact (1 percent) in the dimensionless film thickness range 4 to 5. Full contact occurred at a film thickness ratio several times larger than commonly reported in the literature.

2. Choosing larger sample intervals in digitizing surface traces is the same as low-pass filtering. The effect is to lower the asperity count. Low-pass filtering of surface traces will shift the no-contact time fraction curves to lower values of film thickness.

3. All the asperity no-contact time fraction results may be correlated on the basis of asperity density counts, which is connected with a sample interval effect.

4. A rational basis for selection of the surface sampling interval consistent with the elastohydrodynamic lubrication process does not exist at the present time.

5. The usefulness of the curves of contact time fraction as a means of determining film thickness by electrical conductance measurements is limited to a narrow range of film thickness ratio.

REFERENCES

1. Dowson, D.; and Higginson, G. R.: Elastohydrodynamic Lubrication. Pergamon Press, 1966.

2. Zaretsky, E. V.; and Anderson, W. J.: EHD Lubrication. Mach. Des., vol. 40, Nov. 7, 1968, pp. 167-173.

3. Tallian, T. E.: Elastohydrodynamic Hertzian Contacts - Part 1. Mech. Eng., vol. 93, no. 11, Nov. 1971, pp. 14-18.

4. Tallian, T. E.: Elastohydrodynamic Hertzian Contacts - Part 2. Mech. Eng., vol. 93, no. 12, Dec. 1971, pp. 17-22.

5. Seth, B. B.; and Willis, T.: Techniques for Film Thickness Measurements in Elastohydrodynamic Lubrication. ASME Paper 76-DET-79, Sep. 1976.

6. Kannel, J. W.; and Snediker, D. K.: Elastohydrodynamic Lubrication in an Instrument Ball-Bearing. J. Lubr. Technol., vol. 98, no. 2, Apr. 1976, pp. 244-250.

7. Tallian, T. E.; et al.: Lubricant Films in Rolling Contact of Rough Surfaces. Trans. ASLE, vol. 7, no. 2, Apr. 1964, pp. 109-126.

8. Tallian, T. E.: The Theory of Partial Elastohydrodynamic Contacts. Wear, 21, 1972, pp. 49-101.

9. Greenwood, J. A.: On the Area of Contact Between Rough Surfaces and Flats. J. Lubr. Technol., vol. 89, no. 1, Jan. 1967, pp. 81-91.

10. Whitehouse, D. J.; and Archard, J. F.: The Properties of Random Surfaces of Significance in Their Contact. Proc. R. Soc. (London), Ser. A, vol. 316, no. 1524, Mar. 31, 1970, pp. 97-121.

11. Johnson, K. L.; Greenwood, J. A.; and Poon, S. Y.: A Simple Theory of Asperity Contact in Elastohydrodynamic Lubrication. Wear, vol. 19, 1972, pp. 91-108.

12. Poon, S. Y.; and Haines, D. J.: Frictional Behaviour of Lubricated Rolling-Contact Elements. Proc. Inst. Mech. Eng. (London), vol. 181, 1966-67, pp. 363-376.

13. Archard, J. F.: Surface Topography and Tribology. Tribol. Int., vol. 7, no. 5, Oct. 1974, pp. 213-220.

14. Nayak, P. Ranganath: Random Process Model of Rough Surfaces. J. Lubr. Technol., vol. 93, no. 3, July 1971, pp. 398-407.

15. Longuet-Higgins, M. S.: The Statistical Analysis of a Random, Moving Surface. Phil. Trans. Roy. Soc. (London), Ser. A, vol. 249, no. 966, 1956-1957, pp. 321-387.

16. Longuet-Higgins, M. S.: Statistical Properties of an Isotropic Random Surface. Phil. Trans. Roy. Soc. (London), Ser. A, vol. 250, no. 975, 1957-1958, pp. 157-174.

17. Longuet-Higgins, M. S.: The Statistical Geometry of Random Surfaces. Hydrodynamic Instability: 13th Symposia in Applied Mathematics, R. Bellman, ed., American Mathematical Society, 1962, pp. 105-143.

18. Nayak, P. Ranganath: Some Aspects of Surface Roughness Measurement. Wear, vol. 26, 1973, pp. 165-174.

19. Sidik, Steven M.: Analysis of Electrical Contact Occurrences Between Rolling Surfaces with Application to Elastohydrodynamic Lubrication. NASA TN D-7134, 1973.

20. Sidik, Steven M.: Two-Dimensional Gaussian Processes Applied to the Determination of Contact Between Lubricated Rolling Surfaces. NASA TM X-71598, 1974.

21. Sidik, S. M.; and Coy, J. J.: Statistical Model for Asperity-Contact Time Fraction in Elastohydrodynamic Lubrication. NASA TP-1130, 1978.

22. Sayles, R. S.; and Thomas, T. R.: Measurements of the Statistical Microgeometry of Engineering Surfaces. ASME Paper 78-LUB-15, Oct. 1978.

23. Adler, Robert J.; and Hasofer, A. M.: Level Crossings for Random Fields. Ann. Probability, vol. 4, no. 1. Feb. 1976, pp. 1-12.

APPENDIX

The probabilistic behavior of an ergodic Gaussian random surface is entirely defined by either the correlation function R or the power spectral density (psd) function $\mathcal{S}$. They are Fourier transform pairs. A Gaussian random process $z(x,y)$ is one (1) that follows a Gaussian distribution with mean μ and variance σ^2 and (2) for which, for all finite n and values of $\tau_{x,i}$ and $\tau_{y,i}$ ($i = 1,n$), the variables $z(x + \tau_{x,i}, y + \tau_{y,i})$ follow joint multivariate normal distributions. In this report the correlation function is defined as

$$R(\tau_x, \tau_y) = \frac{1}{\sigma^2} E\left\{ \left[z(x + \tau_x, y + \tau_y) - \mu \right] \left[z(x,y) - \mu \right] \right\} \quad (A1)$$

This function measures the degree of relation between the heights of the random surface above two points of the reference plane that are a fixed distance and direction apart.

The ergodic assumption states that the statistics of the process are not a function of x and y. Therefore, the expectation may be taken in the ensemble sense as in equation (A1) or as an average over the x-y plane as follows:

$$R(\tau_x, \tau_y) = \lim_{\substack{T_x \to \infty \\ T_y \to \infty}} \frac{1}{\sigma^2} \frac{1}{2T_x} \frac{1}{2T_y} \int_{-T_x}^{T_x} \int_{-T_y}^{T_y} \left[z(x + \tau_x, y + \tau_y) - \mu \right]$$

$$\times \left[z(x,y) - \mu \right] dy\, dx \quad (A2)$$

The psd is the Fourier transform of R given by

$$\mathcal{S}(\varphi_x, \varphi_y) = \frac{1}{(2\pi)^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \exp\left[-i(\tau_x \varphi_x + \tau_y \varphi_y) \right] R(\tau_x, \tau_y) d\tau_x\, d\tau_y$$

$$(A3)$$

Hence, by means of the inverse Fourier transform the correlation function is obtained.

$$R(\tau_x, \tau_y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \exp\left[i(\tau_x \varphi_x + \tau_y \varphi_y)\right] \mathcal{S}(\varphi_x, \varphi_y) d\varphi_x \, d\varphi_y \tag{A4}$$

The correlation function is a characterization of the surface in the x-y plane; the psd is a characterization of the surface in the frequency domain, where φ_x and φ_y are frequencies. That is, a random surface z(x,y) may be thought of as a superposition of many surface waves of different wavelength. Frequencies at which the psd is largest contribute more terms; frequencies at which it is smallest contribute the least.

Spectral moments. - The interpretation of the psd as a probability density function leads naturally to considering the moments of that distribution as descriptors of its shape. These moments, m_{ij}, are defined here as

$$m_{ij} = \sigma^2 \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \varphi_x^i \varphi_y^j \mathcal{S}(\varphi_x, \varphi_y) d\varphi_x \, d\varphi_y \tag{A5}$$

The psd $\mathcal{S}(x,y)$, as defined here, is truly a probability density function. Thus, strictly speaking, the m_{ij} as defined by equation (A5) should not include the factor σ^2. The reason for including it in this way is so the spectral moments defined herein will correspond to the definition of spectral moments as presented by Longuet-Higgins and Nayak (refs. 14 to 18). Nayak (ref. 14) has discussed a method of estimating the two-dimensional moments from a series of one-dimensional profiles obtained in at least three different directions. This method is explained next.

Relation between surface and profile power spectral densities. - Consider a straight line through the origin in the x-y plane and at an angle θ to the x-axis. The height of the surface above this line is a one-dimensional random function of r, the distance from the origin. The correlation function and psd are defined by

$$R_\theta(\tau) = \frac{1}{\sigma^2} E\left\{\left[z(r) - \mu\right]\left[z(r+\tau) - \mu\right]\right\} \tag{A6}$$

and

$$\mathcal{S}_\theta(\varphi) = \frac{1}{2\pi} \int_{-\infty}^{\infty} R_\theta(\tau) e^{-i\varphi\tau} d\tau \tag{A7}$$

The moments of the profile spectrum are calculated by

$$m_{n,\theta} = \int_{-\infty}^{\infty} \varphi^n \mathscr{G}_\theta(\varphi) d\varphi \tag{A8}$$

The moments of profile psd's and the surface psd are related by the following equation (ref. 14):

$$m_{n,\theta} = \sum_{k=0}^{n} \binom{n}{k} m_{n-k,k} \sin^k\theta \cos^{n-k}\theta \tag{A9}$$

where $\binom{n}{k}$ denotes the number of combinations of n things taken k at a time.

TABLE I. - HERTZIAN CONTACT CONDITIONS AT INNER AND OUTER

RACES FOR THREE THRUST LOADS

[Width of rolling track is determined by major axis width.]

Race	Contact condition	Thrust load, N (lbf)		
		90 (20)	445 (100)	3100 (700)
Inner	Maximum Hertzian stress, GPa (ksi)	1.28 (185)	2.09 (303)	3.63 (527)
	Semimajor axis, cm (in.)	0.0510 (0.0200)	0.0840 (0.0330)	0.1500 (0.0570)
	Semiminor axis, cm (in.)	0.0066 (0.0026)	0.0110 (0.0043)	0.0190 (0.0076)
Outer	Maximum Hertzian stress, GPa (ksi)	1.13 (164)	1.85 (269)	3.27 (474)
	Semimajor axis, cm (in.)	0.0460 (0.0180)	0.0740 (0.0290)	0.1300 (0.0500)
	Semiminor axis, cm (in.)	0.0086 (0.0034)	0.0140 (0.0055)	0.0250 (0.009.)

TABLE II. - DIGITIZATION SAMPLE INTERVALS, NUMBER OF SAMPLE

POINTS, AND TOTAL LENGTH OF SURFACE PROFILE SAMPLED

Trace	Profile	Sample interval, Δ, μm	Sample points, N	Sampled length, cm
1	Ball, 0°	0.94	14 220	1.30
2	Ball, 45°	.75	32 232	2.40
3	Ball, 90°	.71	9 954	.70
4	Flat, 0°	0.31	29 388	0.90
5	Flat, 18°		25 122	.77
6	Flat, 36°		27 492	.84
7	Flat, 54°		29 388	.90
8	Flat, 72°		29 388	.90
9	Flat, 90°		25 122	.77
10	Flat, 90° (repeat)		28 440	.87

TABLE III. - TOTAL NUMBER OF SAMPLE POINTS

USED IN COMPUTING MOVING AVERAGE

Trace	Profile	Thrust load, N (lb)		
		90 (20)	445 (100)	3100 (700)
1	Ball, 0°	2163	3 567	6 163
2	Ball, 45°	2711	4 475	7 729
3	Ball, 90°	2857	4 715	8 143
4	Flat, 0°	6623	10 927	18 875
5	Flat, 18°			
6	Flat, 36°			
7	Flat, 54°			
8	Flat, 72°			
9	Flat, 90°			
10	Flat, 90° (repeat)	↓	↓	↓

TABLE IV. - ESTIMATES OF SPECTRAL MOMENTS

Trace	Profile	Thrust load, N (lb)								
		Profile spectral moments								
		m_0, μm^2			m_2, dimensionless			m_4, μm^{-2}		
		90 (20)	445 (100)	3100 (700)	90 (20)	445 (100)	3100 (700)	90 (20)	445 (100)	3100 (700)
1	Ball, 0°	1.90×10^{-3}	1.92×10^{-3}	1.91×10^{-3}	8.46×10^{-5}	8.49×10^{-5}	7.56×10^{-5}	4.07×10^{-5}	4.21×10^{-5}	3.04×10^{-5}
2	Ball, 45°	1.86×10^{-3}	2.01×10^{-3}	2.27×10^{-3}	1.66×10^{-4}	1.67×10^{-4}	1.68×10^{-4}	1.22×10^{-4}	1.25×10^{-4}	1.24×10^{-4}
3	Ball, 90°	1.81×10^{-3}	1.69×10^{-3}	1.29×10^{-3}	1.66×10^{-4}	1.63×10^{-4}	1.38×10^{-4}	1.13×10^{-4}	1.12×10^{-4}	1.0×10^{-4}
4	Flat, 0°	7.94×10^{-2}	9.84×10^{-2}	6.12×10^{-2}	2.12×10^{-3}	2.11×10^{-3}	2.06×10^{-3}	6.64×10^{-2}	6.61×10^{-2}	6.49×10^{-2}
5	Flat, 18°	6.59×10^{-2}	6.99×10^{-2}	4.22×10^{-2}	2.65×10^{-3}	2.69×10^{-3}	2.57×10^{-3}	6.55×10^{-2}	6.67×10^{-2}	6.71×10^{-2}
6	Flat, 36°	5.55×10^{-2}	6.19×10^{-2}	6.91×10^{-2}	3.62×10^{-3}	3.64×10^{-3}	3.40×10^{-3}	6.44×10^{-2}	6.47×10^{-2}	6.56×10^{-2}
7	Flat, 54°	6.06×10^{-2}	6.37×10^{-2}	7.58×10^{-2}	4.38×10^{-3}	4.46×10^{-3}	4.42×10^{-3}	6.77×10^{-2}	6.74×10^{-2}	6.83×10^{-2}
8	Flat, 72°	4.45×10^{-2}	4.42×10^{-2}	4.97×10^{-2}	5.07×10^{-3}	5.19×10^{-3}	5.06×10^{-3}	7.20×10^{-2}	7.25×10^{-2}	7.2×10^{-2}
9	Flat, 90°	4.99×10^{-2}	4.73×10^{-2}	5.18×10^{-2}	5.65×10^{-3}	5.62×10^{-3}	5.82×10^{-3}	7.14×10^{-2}	7.12×10^{-2}	6.89×10^{-2}
10	Flat, 90° (repeat)	4.96×10^{-2}	4.98×10^{-2}	5.05×10^{-2}	5.72×10^{-3}	5.68×10^{-3}	5.66×10^{-3}	7.39×10^{-2}	7.29×10^{-2}	7.2×10^{-2}

TABLE V. - DIMENSIONLESS CONTACT FRACTIONS

Load		Dimensionless film thickness,	Summits of height Λ per square centimeter,	Excursions above level Λ per square centimeter,	Dimensionless contact fractions		
N	lbf	Λ	$f(\Lambda)$	$E(\chi)$	$T_{c,in}$	$T_{c,out}$	T_o
90	20	4.35	3018	679.8	0.717	0.843	0.062
90	20	4.40	2440	543.9	.573	.675	.230
90	20	4.50	1584	345.7	.364	.429	.601
90	20	4.60	1017	217.5	.229	.270	.826
90	20	4.70	647	135.6	.143	.168	.9296
90	20	4.80	407	83.7	.088	.104	.9727
90	20	4.90	254	51.2	.054	.064	.9897
90	20	5.00	157	31.1	.033	.039	.99622
90	20	5.10	96	18.7	.020	.023	.99863
445	100	4.55	1242	267.8	0.771	0.866	0.037
445	100	4.60	994	212.2	.610	.686	.196
445	100	4.70	632	132.2	.380	.428	.587
445	100	4.80	398	81.6	.235	.264	.825
445	100	4.90	248	49.9	.144	.161	.9231
445	100	5.00	153	30.3	.087	.098	.9747
445	100	5.10	93	18.2	.052	.059	.99082
3100	700	4.80	403	82.8	0.727	0.814	0.068
3100	700	4.90	251	50.6	.444	.498	.472
3100	700	5.00	155	30.7	.269	.302	.776
3100	700	5.10	95	18.4	.162	.181	.9149
3100	700	5.20	57	10.9	.096	.108	.9693
3100	700	5.30	34	6.4	.057	.063	.9893
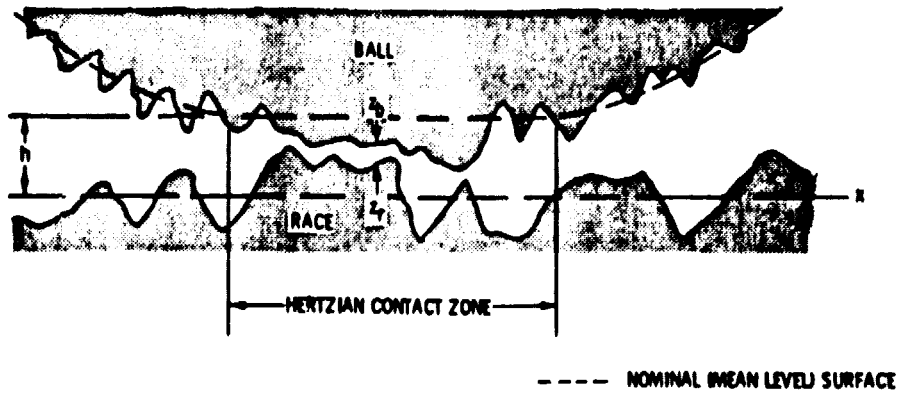
Figure 1. - Cross section of ball-race contact. (Surface roughness is greatly exaggerated.)

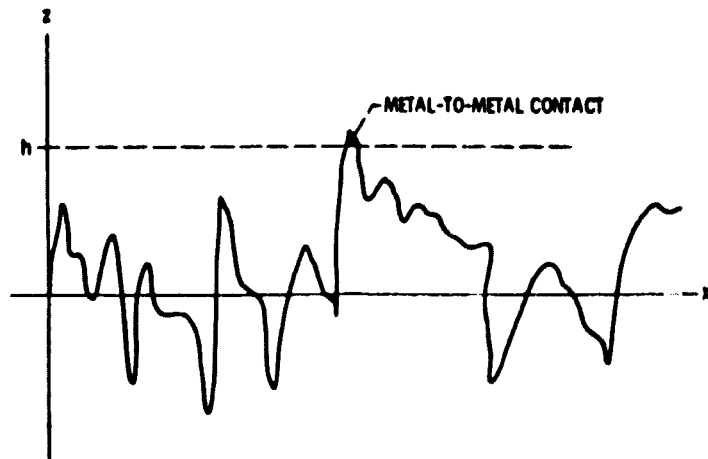---- NOMINAL (MEAN LEVEL) SURFACE

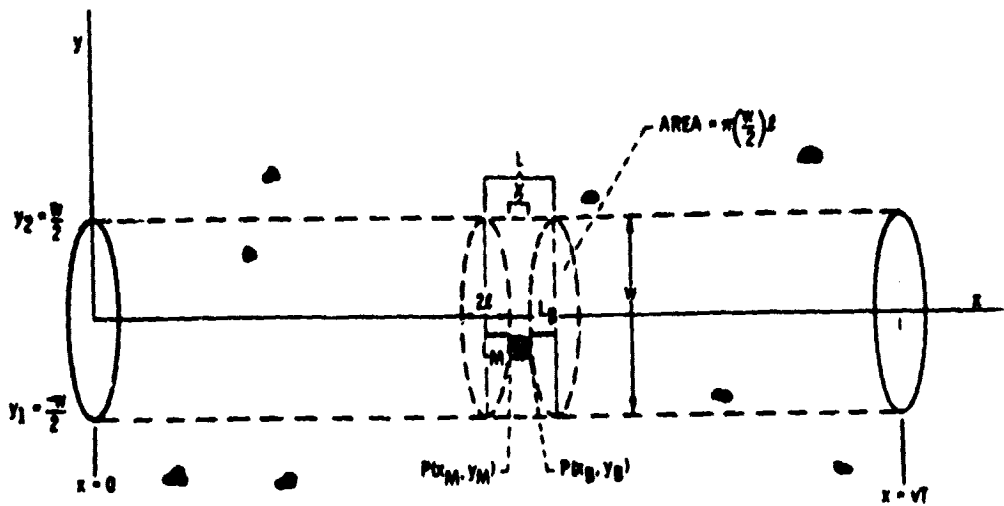Figure 2. - Cross section of a composite surface roughness process.

Figure 3. - Typical contact occurrence when contact areas are small.

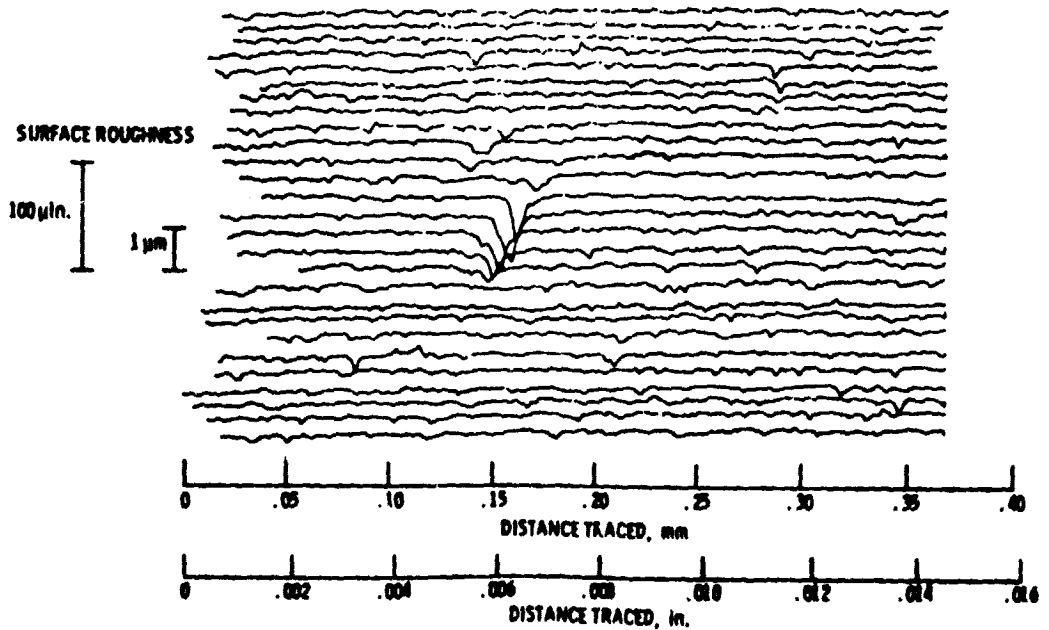Figure 4. - Microtopography of ball specimen. Nominal centerline average (CLA) roughness, 0.03 to 0.05 micrometer (1 to 2 μin.). Nominal trace spacing, 6 micrometers (0.00025 in.); vertical magnification, approximately 20 times horizontal scale. (Small surface scratches in random directions are evident.)
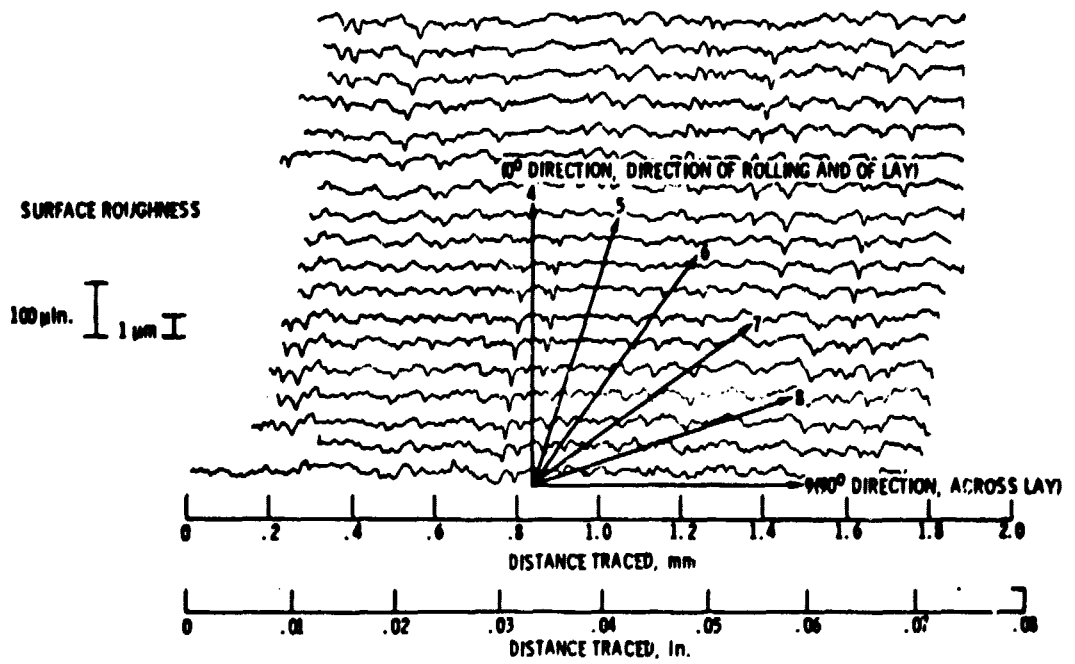
Figure 5. - Microtopography of flat specimen. Nominal centerline average (CLA) roughness, 0.13 to 0.25 micrometer (5 to 10 µin.) Nominal trace spacing, 64 micrometers (0.0025 in.); vertical magnification, 50 times horizontal scale; tracing directions, denoted by arrows and numbers. (Unidirectional scratches showing lay of surface are clearly evident.)
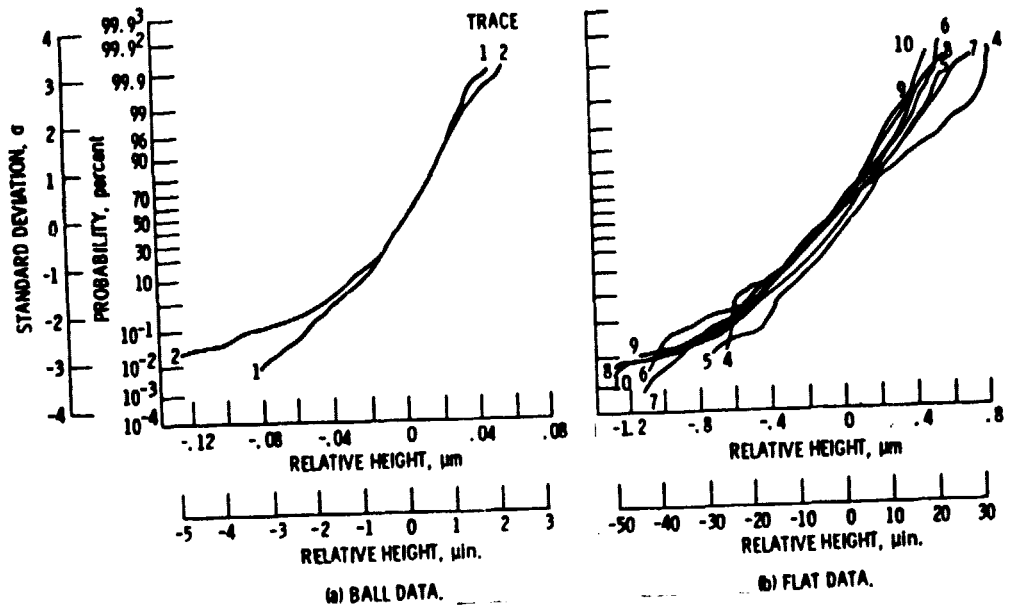
Figure 6. - Cumulative distribution of relative surface height as determined from profile traces in various directions.

CURVES OF T_0 VS Λ FOR 90N (20 lb) LOAD USING MODELS IN LITERATURE CITED

A ADLER AND HASOFER (REF. 23) WITH SPECTRAL MOMENTS FROM TABLE IV

B MOMENTS FROM TABLE IV REDUCED BY 10

C MOMENTS FROM TABLE IV REDUCED BY 100

D JOHNSON, GREENWOOD, AND POON (REF. 11) (20 000 summits/cm^2)

E KANNEL AND SNEDIKER (REF. 6) (50 000 summits/cm^2)

F $\begin{cases} \text{JOHNSON, ET AL. (REF. 11) (E}\beta\text{x) FROM TABLE V)} \\ \text{KANNEL AND SNEDIKER (REF. 6) } (\text{-}10^8 \text{ summits/cm}^2) \end{cases}$

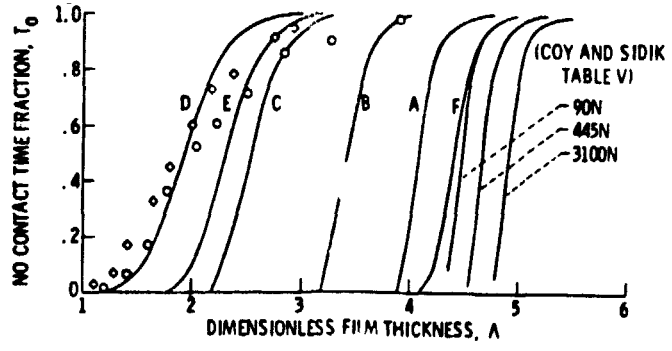o TALLIAN (REF. 7)

□ POON AND HAINE: 12)

Figure 7. - Comparison of results for various models of contact time fraction.