NASA TECHNICAL MEMORANDUM

NASA TM X-71598

NASA TM X- 71598

(NASA-TM-X-71598) TWO-DIMENSIONAL GAUSSIAN PROCESSES APPLIED TO THE DETERMINATION OF CONTACT BETWEEN LUBRICATED ROLLING SURFACES (NASA) 42 p HC $3.25 CSCL 13I G3/15

N74-32931

Unclas 48760

TWO-DIMENSIONAL GAUSSIAN PROCESSES APPLIED TO THE DETERMINATION OF CONTACT BETWEEN LUBRICATED ROLLING SURFACES

by Steven M. Sidik
Lewis Research Center
Cleveland, Ohio 44135

TECHNICAL PAPER proposed for presentation at
One hundred and thirty-fourth Annual Meeting
of the American Statistical Association
St. Louis, Missouri, August 26-29, 1974

E-7166

ABSTRACT

We wish to determine how effectively a lubricant film prevents metallic contact between two rolling surfaces (such as in ball bearings) as a function of surface roughness parameters. The parameters considered are the spectral moments of the two-dimensional surface obtained by superposition of the two rolling surfaces. We consider the peak height distribution, estimation of one-dimensional profile spectral moments, and the estimation of two-dimensional surface moments from several profile measurements. Also given is an asymptotic relation between the mean film thickness and contact occurrences.

TWO-DIMENSIONAL GAUSSIAN PROCESSES APPLIED TO THE DETERMINATION OF CONTACT BETWEEN LUBRICATED ROLLING SURFACES

Steven M. Sidik
Lewis Research Center

E-7166

SUMMARY

The basic problem under consideration is that of describing the behavior of two rolling surfaces separated by a thin film of lubricant. We discuss a procedure for determining the mean film thickness from experimentally obtained electrical contact measurements. We are specifically concerned with rolling surfaces such as occur in ball bearings.

The ball surfaces are represented as Gaussian random processes. The development first considers the situation where contact occurs along a one-dimensional line segment. Both exact and approximate relations are developed between the mean film thickness and the amount of electrical contact. The development is then extended to the situation of contact over an area. An approximate relationship between mean film thickness and electrical contact is developed.

The approximation presented involves the determination of the probability density function of peak heights on a random surface and the density of such peaks.

One method of obtaining surface roughness information is to observe relative surface height along one-dimensional profiles. We discuss the estimation of spectral moments of these one-dimensional processes. We then discuss a method of combining this information from several profiles in different directions to compute the spectral moments of the two-dimensional process.

We conclude with a brief discussion of the need for filtering the profile process in order to obtain better agreement with the contact process.

INTRODUCTION

The basic problem under consideration is that of describing the behavior of two rolling surfaces separated by a thin film of lubricant. This situation arises quite naturally in the study of ball bearings. The purpose of the lubricant is to separate the metallic surfaces in order to prevent bearing wear and thereby lengthen the effective life of the bearing. This report discusses a means of experimentally determining the mean film thickness from electrical contact measurements obtained by an experimental technique described in the first section.

The proposed solution of this problem involves representing the ball surfaces as Gaussian random processes as was done in reference 1. Supporting data for this is presented in Williamson (ref. 2). We begin the development by first considering the hypothetical situation where contact takes place along a one-dimensional line segment. That is, we consider two one-dimensional stationary and ergodic Gaussian random processes whose mean levels are separated by a lubricant film of mean thickness h. We then develop a relationship between h and electrical contact occurrences. An approximation is presented for the case when the mean film thickness is large compared to the surface roughness. This part of the development is similar to that of reference 1 except that an erroneous derivation in reference 1 is corrected.

We then extend the study to the real problem where contact actually occurs over an area. That is, we are now concerned with two two-dimensional stationary and ergodic random processes. This presents considerable difficulties in problem definition and analysis. An approximation is developed for the case when the mean film thickness is large compared to the surface roughness. In order to achieve this result a number of results from and extensions of the subject of "statistical geometry" as developed by Longuet-Higgins (refs. 3, 4, and 5) are presented. Also used are some results from Nayak (ref. 6), Cramer and Leadbetter (ref. 7), and Kac and Slepian (ref. 8).

The approximation developed provides an equation for the electrical no-contact time fraction as a function of the mean film thickness, Hertzian contact area, and surface roughness parameters. This equation involves the probability density function of peak heights on a random surface and the number of peaks per unit surface area.

One method of obtaining surface roughness information is to observe relative surface heights along one-dimensional profiles. We discuss three methods of estimating the spectral moments of the resulting one-dimensional processes. The electrical contact prediction is a function of the spectral moments of the two-dimensional surface obtained by superimposing the two rolling surfaces. Thus a method of combining information from several profiles in different directions to compute the spectral moments of the two-dimensional process is discussed.

We conclude with a brief discussion of the need for filtering the profile processes in order to obtain better agreement with the contact process.

It is appropriate to remark at this point that we consider only a simplified version of the real problem. Some of the complications of real life are bearing spin, bearing sliding, viscosity effects, nonuniform lubricant thickness, and several others. A survey of elastohydrodynamic lubrication and related topics is presented by Tallian (ref. 9).

ONE-DIMENSIONAL ANALYSIS

Relation of Film Thickness and Surface Roughness to Electrical Contact Occurrences

The situation is represented in figure 1. There are two balls, conveniently referred to as the upper ball and the lower ball. Under a condition of load, because of local elastic deformation of the ball surfaces, there is a region of contact where the surfaces can be considered as locally parallel. This region is called the region of Hertzian contact and is illustrated in further detail in figure 2.

The authors of reference 1 have shown that, within the Hertzian area, the two surfaces can be closely approximated by two Gaussian random processes separated by an average distance h which represents the average lubricant film thickness. The upper surface is denoted $Z_u(x)$. We use the notation $Z_u(x) \sim N(\mu_u, C_u(\tau))$ to denote that $Z_u(x)$ is distributed as a Gaussian random process with mean μ_u and autocovariance function (acvf)$C_u(\tau)$. The lower surface is denoted by $Z_l(x)$ and we have $Z_l(x) \sim N(\mu_l, C_l(\tau))$. We assume both Z_u and Z_l to be stationary and ergodic. The distance between the surfaces at any point x is described by $Z_u(x) - Z_l(x)$. Thus, we consider the composite process defined by

$$\xi(x) = Z_u(x) - Z_l(x) - h \tag{1}$$

This process is illustrated in figure 3. We have (Bendat, ref. 10)

$$\xi(x) \sim N(0, C(\tau))$$

where $C(\tau) = C_u(\tau) + C_l(\tau)$. Note that $C(0) = \sigma^2$, the variance or mean square of the process. Whenever $\xi(x) \leq -h$, the surfaces are in the state of metallic contact. Obviously, the physical process cannot have $\xi(x) < -h$. We assume that $\xi(x) < -h$ corresponds to a state of elastic deformation of the surface asperities, and that the process of elastically deforming and reforming asperities has no appreciable effect on the characterization of the surfaces in contact by equation (1).

In order to determine, experimentally, when the surfaces are in metallic contact, it is possible to employ the electrical contact method described in reference 1. We describe this technique briefly as follows.

With reference to figure 1, we might apply a voltage across the upper and lower balls. Thus, whenever the balls are in metallic contact, a current flow will occur through the circuit. A continuous recording of the voltage will then provide a record of whenever the balls are in metallic contact. It is important to note that no voltage (i.e., potential) will be recorded whenever there is metallic contact anywhere within the region of Hertzian contact. A voltage will be observed only for those time intervals such that the lubricant film is unbroken throughout the region of Hertzian contact.

We assume that the surfaces are rolling against each other at a constant velocity with no sliding or spinning. Then if a voltage-averaging device is used during a test, the time-averaged fraction of the applied voltage is a function of the distance-averaged fraction of rolled-over-distance which corresponds to the state of metallic contact. We denote the distance-averaged fraction of electrical contact as T_c. The no-contact fraction is thus given by $1 - T_c$. We wish to derive a relation between the average film thickness h and the contact fraction T_c so that experimental measurements of average voltage may be used to estimate h.

Derivation of Contact Fraction

We begin by describing what happens during an electrical no-contact occurrence. Let the length of the Hertzian contact be denoted by d. For each excursion of $\xi(x)$ above the level $\xi(x) = -h$, electrical contact is not broken until the left edge of the Hertzian contact includes the first point at which $\xi(x) > -h$. That is, electrical contact is not broken until the right edge of the Hertzian contact is d units of distance past the first point for which $\xi(x) > -h$. Electrical contact is reestablished as soon as the right edge of the Hertzian contact includes the first point for which $\xi(x) = -h$ again. Thus, if the excursion above $-h$ is of length L, there is a loss of electrical contact only for the length $(L - d)^+$ where $(x)^+ = \mathrm{MAX}\{x, 0\}$.

Define the following random variables (rv) associated with the random process $\xi(x)$:

$\delta_a(h)$	rv denoting lengths of excursions of $\xi(x)$ above level $-h$
$N_a(t, h)$	rv denoting number of crossings of $\xi(x)$ from below $-h$ to above $-h$ in interval $(0, t)$
$N_a(t, h \mid \delta \leq d)$	rv denoting number of crossings of $\xi(x)$ from below $-h$ to above $-h$ in interval $(0, t)$ where excursion is of length less than or equal to d
$N_a(t, h \mid \delta > d)$	rv denoting number of crossings from below $-h$ to above $-h$ where the excursion is of length greater than d
$N(t, h)$	rv denoting total number of crossings of level $-h$ in either direction

Let the random variables $\delta_b(h)$, $N_b(t, h)$, $N_b(t, h \mid \delta \leq d)$, and $N_b(t, h \mid \delta > d)$ be defined similarly for excursions and crossings from above $-h$ to below $-h$. Cramer and Leadbetter (ref. 7, ch. 11) present rigorous proofs that these can, in fact, be treated as random variables. The following developments are presented in a heuristic manner but can be made rigorous through application of the results in reference 7.

We denote the distribution functions of $\delta_a(h)$ and $\delta_b(h)$ by $A_h(\delta_a)$ and $B_h(\delta_b)$, respectively. Although there has been much research directed toward identifying these distributions, they are not known except for special values of h and certain forms of the acvf $C(\tau)$ (e.g., see Slepian, ref. 11). We assume the expectations of $\delta_a(h)$ and $\delta_b(h)$ exist and denote them, respectively, as t_a and t_b.

Of primary interest are the two conditional expectations $\overline{T}_N$ and $\overline{T}_C$ defined by

$$\overline{T}_N = E\left[\delta_a(h) \mid \delta_a(h) > d\right]$$

$$\overline{T}_C = E\left[\delta_a(h) \mid \delta_a(h) \le d\right]$$

where d is the length of the Hertzian contact. Since $\delta_a(h)$ is a positive rv, we have

$$t_a = \overline{T}_N \Pr\left[\delta_a(h) > d\right] + \overline{T}_C \Pr\left[\delta_a(h) \le d\right]$$

For sufficiently large t, the following statements are true ($\approx$ denotes approximately equal):

$$N_a(t,h) \approx N_b(t,h) \approx \frac{1}{2} N(t,h)$$

$$\frac{N_a\left(t,h \mid \delta_a(h) > d\right)}{N_a(t,h)} \approx \Pr\left[\delta_a(h) > d\right]$$

$$\frac{N_a\left(t,h \mid \delta_a(h) \le d\right)}{N_a(t,h)} \approx \Pr\left[\delta_a(b) \le d\right]$$

Because of the stationarity and ergodicity assumptions, the approximate equalities approach equality as $t \to \infty$. Likewise, the following derivations involving approximate equalities become exact in the limit as $t \to \infty$. Thus, for sufficiently large t, the total no-contact distance is approximated by the number of excursions above the level $-h$ which last longer than d, multiplied by the average amount by which such excursions exceed d. Thus,

$$1 - T_c \approx \frac{N_a\left(t, h \middle| \delta_a(h) > d\right)\left(\overline{T}_N - d\right)}{t}$$

But we have

$$t \approx N_a(t, h)t_a + N_b(t, h)t_b$$

$$\approx (t_a + t_b)\, \underline{N(t, h)}$$

Thus

$$1 - T_c \approx \frac{N_a\left(t, h \middle| \delta_a(h) > d\right)\left(\overline{T}_N - d\right)}{(t_a + t_b)\,\frac{N(t, h)}{2}}$$

$$\approx \frac{\left(\overline{T}_N - d\right)}{t_a + t_b}\,\frac{N_a\left(t, h \middle| \delta_a(h) > d\right)}{N_a(t, h)}$$

$$\approx \frac{\overline{T}_N - d}{t_a + t_b}\,\Pr\left\{\delta_a(h) > d\right\}$$

$$= \frac{E\left\{\delta_a(h) \middle| \delta_a(h) > d\right\} - d}{E\left[\delta_a(h)\right] + E\left[\delta_b(h)\right]}\,\Pr\left[\delta_a(h) > d\right] \tag{2}$$

In order to utilize equation (2), we must know the distribution of $\delta_a(h)$ and the expectation of $\delta_b(h)$. In general, these are unknown. There are some asymptotic results for $\delta_b(h)$ as $h \to \infty$. Cramer and Leadbetter (ref. 7) also have a number of results concerning the expectation of the excursion length which are exact. It must be noted that these exact results are, unfortunately, inapplicable since the quantity of interest is $E(L - d)^+$, not $E(L)$. Equation (2) reduces to the expression derived in reference 1 when the excursion lengths are assumed to be exponentially distributed. This assumption is open to question although it is true in the limit as $h \to \infty$.

Computational Method as h Approaches ∞

One of the primary purposes behind studying the EHD process is to design bearing systems where there is little or no metallic contact between surfaces. This implies that primary concern be given to the processes where h is large compared to the surface roughness. That is, where h/σ is large.

For this situation some results from Cramer and Leadbetter (ref. 7) and from Rice (ref. 12) and Kac and Slepian (ref. 8) are valuable. These authors consider (1) the distribution of the lengths between successive upcrossings of a high level and (2) the distribution of the lengths of excursions above a high level. Since the Gaussian process is symmetric about the mean, it is evident that simular results apply to excursions and crossings below a low level.

With regard to item (1), let $F_c(x)$ denote the distribution function of the interval between successive downcrossings of the level $\xi(x) = -h$, and let $1/\theta_c$ denote the mean of this distribution. It is shown in reference 7 that

$$\lim_{h \to \infty} F_c \left(\frac{x}{\theta_c}\right) = 1 - e^{-x}$$

= Pr (the interval between one downcrossing and the next of $\xi(x) = -h$ is $\leq x/\theta_c$)

Since

$$\theta_c = \frac{C''(0)}{2\pi} e^{-h^2/2}$$

we have

$$\theta_c \xrightarrow[h \to \infty]{} 0$$

Hence, the expected length between downcrossings increases rapidly and the stream of downcrossings becomes Poisson in nature.

With regard to item (2), let $F_e(x)$ denote the distribution function of the lengths of excursions below the level $\xi(x) = -h$ and let θ_e denote the mean of this distribution. Then it is shown in reference 7 that

$$\lim_{h \to \infty} F_e(\theta_e x) = 1 - \exp\left(-\frac{\pi x^2}{4}\right)$$

= Pr (excursion length is $\leq \theta_e x$)

Since for large h

$$\theta_e \approx \frac{1}{h}\left(\frac{2\pi}{C''(0)}\right)^{1/2}$$

we have

$$\theta_e \xrightarrow[h \to \infty]{} 0$$

Hence, the expected length of the excursion decreases rapidly.

These two results may now be combined to give a limiting expression for T_c. In particular, for sufficiently large h, we will have excursions below $-h$ occurring quite infrequently; and each excursion causes an electrical contact whose duration will be, on the average, $d + \theta_e$. Thus, for sufficiently large h/σ and t,

$$T_c \approx \frac{N_b(t,h)(d + \theta_e)}{t}$$

We note that Rice (ref. 13) has shown

$$E\left[\frac{N_b(t,h)}{t}\right] \approx \frac{1}{\pi}\left[\frac{-C''(0)}{C(0)}\right]^{1/2} \exp\left[\frac{-h^2}{2C(0)}\right]$$

and hence

$$T_c \approx \left\{\frac{1}{\pi}\left[\frac{-C''(0)}{C(0)}\right]^{1/2} \exp\left[\frac{-h^2}{2C(0)}\right]\right\}\left\{d + \theta_e\right\}$$

Since $C(0) = \sigma^2$, the variance of the Gaussian process, this equation may be rewritten and solved in terms of h/σ. In order to simplify this, we assume θ_e is small relative to d so that

$$\ln T_c = \ln\left\{\frac{d}{\pi}\left[\frac{-C''(0)}{\sigma^2}\right]^{1/2}\right\} - \frac{1}{2}\left(\frac{h}{\sigma}\right)^2$$

or

$$\frac{h}{\sigma} = \sqrt{2}\left(\ln\left\{\frac{d}{T_c \pi}\left[\frac{-C''(0)}{\sigma^2}\right]^{1/2}\right\}\right)^{1/2}$$

TWO-DIMENSIONAL ANALYSIS

Derivation of Contact Fraction

At this point we make the observation that the Hertzian region is (for two balls) a circular area. The authors of reference 1 began their development with the assumption that the Hertzian region is an arbitrarily narrow and approximately rectangular area. They then developed what is essentially a one-dimensional theory. After considering the one-dimensional theory, they then made a rather weak heuristic generalization to the two-dimensional theory.

The results of Nayak (ref. 6) show that a naive analysis assuming that a profile trace in a single direction may be used directly is erroneous. He shows how the two-dimensional moments may be estimated from a one-dimensional trace under the assumption of statistical isotropy. Both Nayak (ref. 6) and Longuet-Higgins (refs. 3, 4, and 5) consider such statistical descriptions as distribution of maxima and minima, distribution of peak heights, mean summit curvature, and so forth. Longuet-Higgins also discusses a method of determining a sequence of estimating functions which converge to the two-dimensional spectrum.

The statistical problem requiring solution before electrical contact analysis can be rigorously applied is now outlined. A complete solution to this problem is yet to be developed. In fact, an appropriate formulation would be useful.

The composite process of interest can be described similarly to equation (1) as

$$\xi(x, y) = Z_u(x, y) - Z_l(x, y) - h \tag{4}$$

where

(x, y) are the coordinates of the reference plane

$$Z_u(x, y) \sim N\left(\mu_u, C_u(\tau_x, \tau_y)\right)$$

$$Z_l(x, y) \sim N\left(\mu_l, C_l(\tau_x, \tau_y)\right)$$

and hence

$$\xi(x, y) \sim N\left(0, C(\tau_x, \tau_y)\right)$$

where

$$C(\tau_x, \tau_y) = C_u(\tau_x, \tau_y) + C_l(\tau_x, \tau_y)$$

and $C(0, 0) = \sigma^2$, the variance or mean square of the process.

Figure 4 presents a visualization of the metallic-contact areas in the plane parallel to the (x, y) plane at level -h. For simplicity, we assume the x-axis corresponds to the direction of rolling. The area over which the Hertzian area passes as the surfaces roll over each other is bounded by the two horizontal dashed lines denoted $y = Y_1$ and $y = Y_2$. The areas for which $\xi(x, y) \leq -h$ are indicated by the enclosed dotted regions. For simplicity, we assume that the Hertzian area is circular.

The left circle corresponds to the position of the Hertzian area when electrical contact is first broken. The right circle corresponds to the position of the Hertzian area when electrical contact is first made after passing over the "hill."

From figure 4 we see that the quantity whose distribution we need to known is X. Since the distribution of X appears too complicated to derive at the current state of knowledge concerning two-dimensional random processes, we will consider a potentially useful approximation for the case when h/σ is large.

Computational Method as h Approaches ∞

As h/σ gets large we might expect a similar asymptotic Poisson character of the number of "valleys" to occur in the two-dimensional process as occurs in the one-dimensional process. This point should be verified. In this case, the metallic-contact occurrences are more appropriately illustrated by figure 5. This figure presents a visualization of the metallic-contact areas in the plane $\xi(x, y) = -h$. The areas for which $\xi(x, y) \leq -h$ are indicated by the enclosed dotted regions. The area over which the Hertzian area passes as the surfaces roll over each other is bounded by the two horizontal dashed lines denoted $y = Y_1$ and $y = Y_2$. The dashed circle on the left of the region denoted I corresponds to the position of the Hertzian area when electrical contact is first made. The point (x_m, y_m) denotes the coordinates of the point where contact is first made. The dashed circle to the right of the region I corresponds to the position of the Hertzian area when electrical contact is finally broken. The point (x_b, y_b) denotes the coordinates of the point where contact is first broken. We let X denote the difference $x_b - x_m$. If X is relatively small compared to the diameter of the Hertzian area, we may introduce the following approximation. It is reasonable to expect y_m and y_b to be quite highly and positively correlated. Thus, assume y_m and y_b are perfectly correlated and that y_m is uniformly distributed over the vertical distance bounded by $y = Y_1$ and $y = Y_2$. In this case we have

$$E(L) = 2E(L_m) + E(X)$$

$$\approx 2E(L_m)$$

If the shape of the Hertzian area and its dimensions are known, it is a simple matter to compute $E(L_m)$.

Let D_{sum} denote the expected number of summits per unit area and $p(\xi^*)$ denote the probability density function of the summit heights, ξ^*, in the composite process defined by $\xi(x, y)$. Since the area over which the Hertzian region passes is of width $w = Y_1 - Y_2$, we obtain the expected number of summits occurring within the strip per unit distance along the x-axis as wD_{sum}. The probability that two peaks occur within the Hertzian region is negligible by the Poisson nature of the peak occurrences. The average fraction of contact distance T_c is thus given by multiplying the expected number of summits that exceed the level h per unit distance along x, times the average distance contact is maintained for each excursion. That is

$$T_c \approx \left\{ wD_{sum} \int_h^\infty p(\xi^*)d\xi^* \right\} 2\,E(L_m) \qquad (5)$$

Peak Height Probability Density Function

The density function $p(\xi^*)$ is really a conditional density of the surface height conditioned on the event that the point is a summit. Since the conditioning event has zero probability, great care must be exercised in the definition of the conditional probability density. This has been discussed by Kac and Slepian (ref. 8).

Nayak (ref. 6) presents a derivation of $p(\xi^*)$. The derivation presented here shows that the results correspond to a "horizontal window" conditional density as defined by Kac and Slepian. This is quite important, as only "horizontal window" arguments lead to the pdf possessing the ergodic property. Let

$$\xi_4 = \xi(x, y)$$

$$\left.\begin{aligned}
\xi_1 &= \frac{\partial^2}{\partial x^2}\xi(x, y) \\
\xi_2 &= \frac{\partial^2}{\partial x \partial y}\xi(x, y) \\
\xi_3 &= \frac{\partial^2}{\partial y^2}\xi(x, y) \\
\xi_5 &= \frac{\partial}{\partial x}\xi(x, y) \\
\xi_6 &= \frac{\partial}{\partial y}\xi(x, y)
\end{aligned}\right\} \tag{6}$$

$$\bar{\xi}^T = [\xi_1, \xi_2, \xi_3, \xi_4, \xi_5, \xi_6]$$

Since $\xi(x, y)$ is assumed to be a stationary and ergodic Gaussian process, there is an autocorrelation function $R(\tau_x, \tau_y)$ and a spectral density function $\mathcal{S}(\varphi_x, \varphi_y)$ associated with the process. The function $R(\tau_x, \tau_y)$ is defined by

$$R(\tau_x, \tau_y) = \frac{C(\tau_x, \tau_y)}{\sigma^2} = \frac{1}{\sigma^2} E\left\{\xi(t_x + \tau_x, t_y + \tau_y)\xi(t_x, t_y)\right\}$$

A duality between $R(\tau_x, \tau_y)$ and $\mathcal{S}(\varphi_x, \varphi_y)$ is defined by

$$R(\tau_x, \tau_y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{i(\tau_x\varphi_x+\tau_y\varphi_y)} \mathscr{S}(\varphi_x, \varphi_y)d\varphi_x d\varphi_y$$

and

$$\mathscr{S}(\varphi_x, \varphi_y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-i(\tau_x\varphi_x+\tau_y\varphi_y)} R(\tau_x, \tau_y)d\tau_x d\tau_y$$

The spectral moments of $\mathscr{S}$ are defined by

$$m_{ij} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \varphi_x^i \varphi_y^j \mathscr{S}(\varphi_x, \varphi_y)d\varphi_x d\varphi_y \tag{7}$$

In terms of these spectral moments, it is well known that $\vec{\xi} \sim N(0, \Sigma)$ where

$$\Sigma = \left[\begin{array}{ccc|c|cc} m_{40} & m_{31} & m_{22} & -m_{20} & 0 & 0 \\ m_{31} & m_{22} & m_{13} & -m_{11} & 0 & 0 \\ m_{22} & m_{13} & m_{04} & -m_{02} & 0 & 0 \\ \hline -m_{20} & -m_{11} & -m_{02} & m_{00} & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & m_{20} & m_{11} \\ 0 & 0 & 0 & 0 & m_{11} & m_{02} \end{array}\right]$$

$$= \begin{bmatrix} F & l & 0 \\ l^T & m_{00} & 0 \\ 0 & 0 & S \end{bmatrix} \tag{8}$$

To derive the distribution of peak heights, consider the area of the x-y plane $dA = (x, x + \delta_x), (y, y + \delta_y)$ where we have at some point (x_0, y_0) in dA

$$\xi_5 = \xi_5(x_0, y_0) = 0 \qquad \xi_1 = \xi_1(x_0, y_0) \leq 0$$

$$\xi_6 = \xi_6(x_0, y_0) = 0 \qquad \xi_3 = \xi_3(x_0, y_0) \leq 0$$

$$\xi_2 = \xi_2(x_0, y_0)$$

$$\xi_1\xi_3 - \xi_2^2 \leq 0$$

The region where the restrictions on ξ_1, ξ_2, and ξ_3 are satisfied is denoted V.

We wish to define the conditional density function

$$p(\xi_4 | \xi_5 = 0, \xi_6 = 0, (\xi_1, \xi_2, \xi_3) \epsilon V) = p(\xi_4 | \text{at a summit})$$

We proceed with a "horizontal window" limiting argument as presented in Kac and Slepian (ref. 8) since we desire ergodic properties of this pdf to hold in the plane (x, y). Thus

$p(\xi_4 | \text{summit})$

$$= \frac{\lim\limits_{dA \to 0} \left[\iiint_V \left\{ \int_{l_5}^{u_5} \int_{l_6}^{u_6} p(\xi_1 \ldots \xi_6) d\xi_5 d\xi_6 \right\} d\xi_1 d\xi_2 d\xi_3 \right]}{\int_{-\infty}^{\infty} \left[\iiint_V \left\{ \int_{l_5}^{u_5} \int_{l_6}^{u_6} p(\xi_1 \ldots \xi_6) d\xi_5 d\xi_6 \right\} d\xi_1 d\xi_2 d\xi_3 \right] d\xi_4}$$

where (l_5, u_5), (l_6, u_6) are lower and upper limits on ξ_5 and ξ_6 imposed by the conditions

$$\xi_5 \approx 0 \text{ in } dA, \ \xi_5(x_0, y_0) = 0$$

and

$$\xi_6 \approx 0 \text{ in } dA, \ \xi_6(x_0, y_0) = 0$$

Now $\xi_5 = \partial\xi/\partial x$ and $\xi_6 = \partial\xi/\partial y$ so that

$$d\xi_5 d\xi_6 = |J| \, dxdy$$

where

$$|J| = |\xi_1\xi_3 - \xi_2^2|$$

Thus $p(\xi_4 | \text{summit})$ is also given by

$$\lim_{\substack{\delta_x \to 0 \\ \delta_y \to 0}} \frac{\left[\iiint\limits_V |J| \left\{\int_0^{\delta_x}\int_0^{\delta_y} p(\xi_1 \ldots \xi_6) dxdy\right\} d\xi_1 d\xi_2 d\xi_3\right]}{\int_{-\infty}^{\infty}\left[\iiint\limits_V |J| \left\{\int_0^{\delta_x}\int_0^{\delta_y} p(\xi_1 \ldots \xi_6) dxdy\right\} d\xi_1 d\xi_2 d\xi_3\right] d\xi_4} \tag{9}$$

The limit is undefined as $\delta_x, \delta_y \to 0$ since the ratio approaches 0/0. Use L'Hospital's rule with respect to δ_x and δ_y, noting that

$$\frac{\partial^2}{\partial\delta_x\partial\delta_y}\int_0^{\delta_x}\int_0^{\delta_y} p(\xi_1 \ldots \xi_6)dxdy\Bigg]_{(0,0)}$$

$$= \lim_{\substack{\delta_x \to 0 \\ \delta_y \to 0}} p(\xi_1, \xi_2, \xi_3, \xi_4, \delta_x, \delta_y) = p(\xi_1, \xi_2, \xi_3, \xi_4, 0, 0)$$

to obtain

$$p(\xi_4 | \text{summit}) = \frac{\left[\iiint_V |J|\, p(\xi_1, \xi_2, \xi_3, \xi_4, 0, 0)d\xi_1 d\xi_2 d\xi_3\right]}{\int^{\infty}\left[\iiint_V |J|\, p(\xi_1, \xi_2, \xi_3, \xi_4, 0, 0)d\xi_1 d\xi_2 d\xi_3\right] d\xi_4}$$

$$\propto \iiint_V \left|\xi_1\xi_3 - \xi_2^2\right| p(\xi_1, \xi_2, \xi_3, \xi_4, 0, 0)d\xi_1 d\xi_2 d\xi_3$$

Note that

$$p(\xi_1, \xi_2, \xi_3, \xi_4, 0, 0) = \frac{1}{\Delta_2\Delta_4} e^{-Q/2}$$

where

$$Q = (\xi_1, \xi_2, \xi_3, \xi_4) \begin{bmatrix} F & l \\ l^T & m_{00} \end{bmatrix}^{-1} \begin{pmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \\ \xi_4 \end{pmatrix} \tag{10}$$

$$\Delta_2 = m_{02}m_{20} - m_{11}^2 = \det S$$

$$\Delta_4 = \det \begin{bmatrix} F & l \\ l^T & m_{00} \end{bmatrix}$$

Nayak (ref. 6) has also presented an argument which concludes that

$$D_{sum} = \int^{\infty} \left[\iiint_V |J| \, p(\xi_1, \xi_2, \xi_3, \xi_4, 0, 0) d\xi_1 d\xi_2 d\xi_3 \right] d\xi_4 \tag{11}$$

And hence, the normalizing factor in equation (9) is simply D_{sum}. The only problems remaining are the evaluation of D_{sum} and $\int_h^{\infty} p(\xi^*) d\xi^*$.

To this end, let

$$R^{-1} = \begin{bmatrix} m_{40} & m_{31} & m_{22} & -m_{20} \\ m_{31} & m_{22} & m_{13} & -m_{11} \\ m_{22} & m_{13} & m_{04} & -m_{02} \\ -m_{20} & -m_{11} & -m_{02} & m_{00} \end{bmatrix} = \begin{bmatrix} F & l \\ l^T & m_{00} \end{bmatrix} \tag{12}$$

Using some results about inverses of partitioned matrices (Graybill, ref. 14)

$$R = \begin{bmatrix} \left[F - \frac{1}{m_{00}} l l^T\right] & -F^{-1} l \left[m_{00} - l^T F^{-1} l\right]^{-1} \\ \frac{-1}{m_{00}} l^T \left[F - \frac{1}{m_{00}} l l^T\right]^{-1} & \left[m_{00} - l^T F^{-1} l\right]^{-1} \end{bmatrix} \tag{13}$$

$$= \begin{bmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{bmatrix}$$

let

$$\vec{\xi}_s = \begin{pmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \end{pmatrix}$$

Then the quadratic form Q of equation (10) is given by

$$\left(\xi_s^T, \xi_4\right) R \begin{pmatrix} \xi_s \\ \xi_4 \end{pmatrix} = \xi_s^T R_{11} \xi_s + \xi_s^T R_{12} \xi_4 + \xi_4 R_{21} \xi_s + \xi_4 R_{22} \xi_4$$

$$= \xi_s^T R_{11} \xi_s + 2\xi_s^T (R_{12} \xi_4) + \xi_4 R_{22} \xi_4$$

$$= \xi_s^T R_{11} \xi_s + 2\xi_s^T R_{11} \left(R_{11}^{-1} R_{12} \xi_4\right) + \left(\xi_4^T R_{21} R_{11}^{-1}\right)$$

$$\times R_{11} \left(R_{11}^{-1} R_{12} \xi_4\right) + \xi_4 R_{22} \xi_4$$

$$- \left(\xi_4^T R_{21} R_{11}^{-1}\right) R_{11} \left(R_{11}^{-1} R_{12} \xi_4\right)$$

$$= \left(\xi_s + R_{11}^{-1} R_{12} \xi_4\right)^T R_{11} \left(\xi_s + R_{11}^{-1} R_{12} \xi_4\right)$$

$$+ \xi_4 \left(R_{22} - R_{21} R_{11}^{-1} R_{12}\right) \xi_4$$

(let $\mathbf{z} = -R_{11}^{-1} R_{12} \xi_4$)

$$= (\xi_s - \mathbf{z})^T \left(F - \frac{1}{m_{00}} l l^T\right)^{-1} (\xi_s - \mathbf{z}) + \frac{1}{m_{00}} \xi_4^2$$

since $\left(R_{22} - R_{21} R_{11}^{-1} R_{12}\right)^{-1} = m_{00}$

Now

$$\left[F - \frac{1}{m_{00}} l l^T\right]^{-1} = F^{-1} - F^{-1} l \left(l^T F^{-1} l - m_{00}\right)^{-1} l^T F^{-1} = F^{-1} - cF^{-1} l l^T F^{-1}$$

where $c = \left(l^T F^{-1} l - m_{00}\right)^{-1}$

And

$$-R_{11}^{-1} R_{12} = - \left[F - \frac{1}{m_{00}} l l^T\right] \left[-F^{-1} l \left(m_{00} - l^T F^{-1} l\right)^{-1}\right]$$

$$= - \left(F - \frac{1}{m_{00}} l l^T\right) (c F^{-1} l)$$

$$= - \left(c l - \frac{c}{m_{00}} l l^T F^{-1} l\right)$$

$$= + \frac{1}{m_{00}} l$$

and hence

$$\xi_s - z = \xi_s - \frac{\xi_4}{m_{00}} l$$

Thus

$(\xi_s - z)^T R_{11} (\xi_s - z)$

$$= \left(\xi_s - \frac{\xi_4}{m_{00}} l\right)^T \left(F^{-1} - c F^{-1} l l^T F^{-1}\right) \left(\xi_s - \frac{\xi_4}{m_{00}} l\right) \tag{14}$$

For fixed ξ_4 we have

$$\xi_s \sim N(0, F)$$

and hence

$$w = \xi_s - \frac{\xi_4}{m_{00}} l \sim N\left(\frac{\xi_4}{m_{00}} l , F\right)$$

The determination of the pdf of the height given we are at a summit thus reduces to the problem of evaluating

$$\iiint_V \left|\xi_1\xi_3 - \xi_2^2\right| e^{-(1/2)(\xi_s-\mu)^T(F^{-1}-cF^{-1}l l^T F^{-1})(\xi_s-\mu)} \, d\xi \qquad (15)$$

where

$$\mu = \frac{\xi_4}{m_{00}} l = \frac{\xi_4}{m_{00}} \begin{bmatrix} -m_{20} \\ -m_{11} \\ -m_{02} \end{bmatrix} \qquad (16)$$

There is, in general, no closed form expression for evaluating this integral. This can be seen from equation (15) which shows the integral to be related to the expectation of a function of normal r.v.'s over a conical region of the space. In fact, the particular conical region is a cone generated by a ray rotated around the line $\xi_1 = \xi_3$ with ξ_1 and ξ_3 both negative. The angle of the ray with $\xi_1 = \xi_3$ is 45^0.

From this representation we can see that the point μ will introduce asymmetry. There is a special case of the process we are considering which retains enough symmetry so that the integral can be evaluated as a closed form. The special case is when the surface is isotropic.

Isotropic Processes

An isotropic random surface is one in which the probabalistic behavior of the surface, along a profile observed at an angle θ, is independent of θ. In particular, this implies that

$$m_{20} = m_{02} = m_2$$

$$m_{11} = m_{13} = m_{31} = 0$$

$$m_{00} = m_0$$

$$3m_{22} = m_{40} = m_{04} = m_4$$

Thus, for an isotropic process, we have

$$R^{-1} = \left[\begin{array}{ccc|cc|c} m_4 & 0 & \frac{m_4}{3} & 0 & 0 & -m_2 \\ 0 & \frac{m_4}{3} & 0 & 0 & 0 & 0 \\ \frac{m_4}{3} & 0 & m_4 & 0 & 0 & -m_2 \\ \hline 0 & 0 & 0 & m_2 & 0 & 0 \\ 0 & 0 & 0 & 0 & m_2 & 0 \\ \hline -m_2 & 0 & -m_2 & 0 & 0 & m_0 \end{array}\right] \qquad (17)$$

Under the isotropy assumption, the integral of equation (15) can be evaluated analytically (e.g., Nayak, ref. 6) to obtain

$$p(\xi^*) = \frac{\sqrt{3}}{2\pi}\left\{\left[e^{-C_1(\xi^*)^2}\right]\left[\frac{3(2\alpha - 3)}{\alpha^2}\right]^{1/2} + \frac{3\pi\sqrt{2}}{2}\left[e^{-1/2(\xi^*)^2}\right]\left[1 + \Phi(\beta)\right]\right.$$

$$\left. \times \left[(\xi^*)^2 - 1\right] + \pi\sqrt{2}\left[\frac{\alpha}{3(\alpha - 1)}\right]^{1/2} \exp\left[-\frac{\alpha(\xi^*)^2}{2(\alpha - 1)}\right]\left[1 + \Phi(\gamma)\right]\right\} \qquad (18)$$

where

$$\xi^* = \left(\frac{\xi_4}{m_0}\right)^{1/2} = \frac{\xi_4}{\sigma}$$

$$\alpha = \frac{m_o m_4}{m_2^2}$$

$$\beta = \left[\frac{3}{2(2\alpha - 3)}\right]^{1/2} \xi^*$$

$$\gamma = \left[\frac{\alpha}{2(\alpha - 1)(2\alpha - 3)}\right]^{1/2} \xi^*$$

$$C_1 = \frac{\alpha}{2\alpha - 3}$$

It may be noted that the properties of an isotropic random surface may be estimated from the statistics of a single profile measured along any direction in the plane. This is not true for non-isotropic surfaces, however.

The surface of a ball may reasonably be assumed to be isotropic because of the way in which they are finished. Hence, when two balls are in contact, the composite process will also be isotropic. This is the case considered in reference 1 and Nayak's (ref. 6) results are applicable.

From the way in which ball bearing raceways are finished, however, one would expect the surface to have grooves and ridges elongated in the direction of grinding. This can be visually verified in some instances. The composite surface process resulting from a ball rolling in a raceway is unlikely to be isotropic. To allow for analyses involving balls rolling in raceways, we now consider methods for estimating the spectral moments m_{ij} so that equation (5) may be used.

DETERMINATION OF SPECTRAL MOMENTS

The use of equation (5) to estimate T_c depends upon the ability to numerically evaluate D_{sum} and $p(\xi^*)$. These quantities, in turn, depend upon the spectral moments m_{ij} of the composite process defined by the sum of the individual surfaces. We outline a method of sampling the individual surfaces of the ball and raceway along selected profiles obtained by use of a profilometer. We begin by discussing methods of estimating the moments along one-dimensional profiles. We then use certain of the results of Longuet-Higgins (ref. 3) which relate profile one-dimensional moments to the two-dimensional moments of the entire surface.

Estimation of Profile Moments

As indicated, the actual surfaces are sampled in a one-dimensional manner with a profilometer. The usual method is to draw a finepointed stylus across the surface and obtain an analog record of the relative surface height along the profile. This analog record may then be digitized and analyzed on a digital computer. There are three possible methods of analysis.

1. Numerically compute the sample one-dimensional spectrum by use of a Fourier transform method (refs. 15 and 16). The moments of the spectrum may then be numerically calculated by quadrature. This method is relatively straightforward and requires no further discussion.

2. Count the number of relative extrema (maxima and minima) and the number of zero-crossings of a profile sample. If we let

D_z = density of zero-crossings
D_e = density of extrema
$\hat{\sigma}^2$ = estimated variance of process

Then it can be shown that estimates for the moments can be defined as

$$\left.\begin{aligned} \hat{m}_0 &= \hat{\sigma}^2 \\ \hat{m}_2 &= \pi^2\hat{\sigma}^2 D_z^2 \\ \text{and} \qquad \hat{m}_4 &= \pi^4\hat{\sigma}^4 D_2^2 D_e^2 \end{aligned}\right\} \tag{19}$$

Naynk (ref. 6) presents a more complete discussion of this method and also indicates a method by which D_z, D_e, and $\hat{\sigma}^2$ can be obtained directly from the analog relative surface height record by electrical devices.

3. Assume some physically realistic but tractably simple family of analytical models for the autocorrelation function. Numerically estimate the autocorrelation function from the digitized profile record via Fourier transform methods. Find the best-fitting member of the

family of autocorrelation functions (i.e., estimate the unknown parameters) and estimate the moments via the well-known relationships (Cramer and Leadbetter, ref. 7)

$$
\begin{aligned}
m_0 &= \frac{1}{\hat{\sigma}^2} R(0) \\
m_2 &= \frac{-1}{\hat{\sigma}^2} R^{(2)}(0) \\
m_4 &= \frac{1}{\hat{\sigma}^2} R^{(4)}(0)
\end{aligned}
\qquad (20)
$$

This method requires further development.

Much of the book by Cramer and Leadbetter is devoted to the conditions under which these moments (and, hence, derivatives of $R(\tau)$ at $\tau = 0$) exist, and the relation of these moments to analytical properties of random processes. Of particular interest is the relation of the moments to such things as the density of zero-crossings and the density of extrema. We refer the reader to Cramer and Leadbetter for full discussion of these topics and limit ourselves to one principal result. This result is that the i^{th} spectral moment exists if and only if $R^{(i)}(\tau)$ exists and is finite at $\tau = 0$. If the second spectral moment is infinite, then the expected number of zero crossings in any finite interval is infinite.

It is thus interesting to note that some authors (refs. 17 and 18) have proposed exponential or exponential-cosine autocorrelation functions as realistic models for surface roughness. Bendat (ref. 10) has also proposed these as being realistic in many other applications. The major theoretically desirable property of a Gaussian random process is that an exponential autocorrelation function implies that the process is Markovian.

Consider the following family of autocorrelation functions

$$R(\tau) = e^{-|\tau|^{\alpha}} \tag{21}$$

where $\alpha = 1$ yields an exponential correlation function.

Then

$$R^{(1)}(\tau) = \begin{cases} \alpha e^{-|\tau|^{\alpha}} |\tau|^{\alpha-1} & \tau < 0 \\ -\alpha e^{|\tau|^{\alpha}} |\tau|^{\alpha-1} & \tau > 0 \end{cases} \tag{22}$$

$$R^{(2)}(\tau) = -e^{-|\tau|^{\alpha}} \left\{\alpha(\alpha - 1)|\tau|^{\alpha-2} + \left(\alpha|\tau|^{\alpha-1}\right)^2\right\} \tag{23}$$

The various values of $R^{(1)}(0)$ and $R^{(2)}(0)$ are given in the following table. From these values, we see that for $\alpha > 2$ the expected number of crossings at any finite level in any finite interval of time is zero. If $\alpha < 2$, then the expected number of crossings of any finite level in any finite interval of time is infinite (or undefined, but undefined in such a manner that we would expect the number of crossings to be infinite). These properties are difficult to accept from physical considerations.

α	$R^{(1)}(0^-)$	$R^{(1)}(0^+)$	$R^{(2)}(0)$
$\alpha > 2$	0	0	0
$\alpha = 2$	0	0	-2
$2 > \alpha > 1$	0	0	$-\infty$
$\alpha = 1$	+1	-1	Undefined
$\alpha < 1$	$+\infty$	$-\infty$	Undefined

In order to resolve this point we propose the following class of models. Let

$$p(\tau) = 1 + a_1\tau^2 + a_2\tau^4 + \ldots + a_n\tau^{2n} \tag{24}$$

$$q(\tau) = 1 + b_1\tau^2 + b_2\tau^4 + \ldots + b_m\tau^{2m} \tag{25}$$

where

$$m - n = k > 0, \; b_i \geq 0, \; b_m > 0$$

and

$$R(\tau) = \frac{p(\tau)}{q(\tau)} \tag{26}$$

Doob (ref. 19) has considered a similar approach to the definition of spectral densities. This representation seems reasonable as an approximation to the exponential-cosine class for the following reasons

1. Since $p(\tau)$ and $q(\tau)$ are both even, we have the required symmetry.

2. Since the degree of q is larger than the degree of p, and q is always positive, $R(\tau)$ will be damped and $R(\tau) \to 0$ as $\tau \to \infty$.

3. If n is sufficiently large, and the roots of p are equally spaced, $R(\tau)$ will have the same qualitative behavior as an exponential cosine function.

4. Since $R^{(j)}(0) = 0$ for $j \geq 2m$, all spectral moments of order $2m$ or greater are equal to zero.

Multiple Profile Analysis

Once the spectral moments of the profiles of the composite surface are obtained along several profiles we wish to use them to estimate the two-dimensional moments. Let

$m_{n,\theta}$ = the n^{th} spectral moment of the one-dimensional profile taken at an angle of θ from the x-axis of the (x, y) plane.

Longuet-Higgins (ref. 3) has shown the following relationship to be true

$$m_{n,\theta} = \sum_{i=n}^{n} m_{n-i,i}\binom{n}{i}\cos^{n-i}(\theta)\sin^{i}(\theta) \tag{27}$$

Suppose that n profiles are taken at the (distinct) angles θ_i, $i = 1, n$. From these profiles we thus obtain

$$m_{0,\theta_i} = m_{00} \qquad (i = 1, n) \tag{28.1}$$

$$m_{2,\theta_i} = m_{20}c_{i1} + m_{11}c_{i2} + m_{02}c_{i3} \qquad (i = 1, n) \tag{28.2}$$

$$m_{4,\theta_i} = m_{40}d_{i1} + m_{31}d_{i2} + m_{22}d_{i3} + m_{13}d_{i4} + m_{04}d_{i5} \qquad (i = 1, n) \tag{28.3}$$

Nayak (ref. 6) claims that three profiles taken at three distinct directions is sufficient information to determine the quantities D_{sum} and $p(\xi^*)$. This seems unlikely to be true for the following reason.

Longuet-Higgins (ref. 5) has shown that many of the gometrical properties of a random surface are determined by the so-called invariants.

Zeroth order invariant:

$$I_0 = m_{00}$$

Second order invariants:

$$S_1 = m_{02} + m_{20}$$

$$S_2 = m_{20}m_{02} - m_{11}^2 = \det|S|$$

Fourth order invariants:

$$F_1 = m_{40} + 2m_{22} + m_{04}$$

$$F_2 = m_{04}m_{40} - 4m_{13}m_{31} + 3m_{22}^2$$

$$F_3 = (m_{40} + m_{22})(m_{04} + m_{22}) - (m_{31} + m_{13})^2$$

$$F_4 = \det F$$

Since there are four invariants of order four, it may be possible that D_{sum} and $p(\xi^*)$ may be expressed in terms of these four invariants. In such a case it would be sufficient to take four profiles and solve equation (28.3) for the two-dimensional moments or the required invariants. It is clear, however, that five distinct profiles will be sufficient since there would then be five equations to be solved for the five unknown fourth order two-dimensional moments. The five equations for the zeroth and second order moments may be solved by least squares. If more than five profiles are obtained, then the fourth order moment equations may also be solved for by least squares.

Prefiltering of Signals

Whitehouse and Archard (ref. 17) have discussed briefly the presence of very low and very high frequency components of the spectrum and the resultant effect upon surface analyses. This deserves some further discussion in relation to the theory of electrical contact occurrences as developed herein.

We first consider high frequencies. A typical profile that might result from a surface which contains primarily low-frequency components superimposed upon a signal of high frequency with small amplitude is illustrated in figure 6. It may easily be recognized that the high-frequency component contributes very heavily toward the number of "summits" contained in the signal. Not all these "summits" have physical significance with relation to electrical contacts, however. We may assume that since the asperities on the surface undergo eleastic deformation, the small asperities will disappear when the larger asperities come into contact. Thus, the data recorded for description of the surface should be filtered so as to remove the high-frequency component.

Figure 7 presents a typical profile that might result from a surface which contains primarily medium-frequency components superimposed on a signal with low frequency and low amplitude. In this case the low-frequency component will be reflected as very long and smooth "hills" and "valleys." But if these are quite large with respect to the Hertzian area, it is more reasonable to assume that the bearings will roll up over the hills and down through the valleys rather than cause very long contact and no-contact occurrences. Thus, the data recorded for description of the surface should also be filtered to remove such low-frequency components.

Since there are little data available in sufficient detail, it is not clear that this problem need arise in practice. The most serious problem is with the high-frequency signals, since they contribute most to the number of summits. Whitehouse and Archard (ref. 17) present some evidence that at least some surfaces may be described as in figure 6. In addition, they also comment that in some instances the smaller peaks undergo plastic rather than elastic deformation and hence disappear rapidly with use. This lends support to the procedure of filtering out the high-frequency components and also presents a potential means for determining where the high-frequency cutoff point should be.

In any event, what may be considered "high" frequencies and "low" frequencies is a subject for further research.

CONCLUDING REMARKS

The basic problem under consideration is that of describing the behavior of two rolling surfaces separated by a thin film of lubricant. We discuss a means of determining the mean film thickness from experimentally obtained electrical contact measurements. We are specifically concerned with rolling surfaces such as occur in ball bearings.

The ball surfaces are represented as Gaussian random processes. The development first considers the situation where contact occurs along a one-dimensional line segment. Both exact and approximate relations are developed between the mean film thickness and the amount of electrical contact. The development is then extended to the situation of contact over an area. An approximate relationship between mean film thickness and electrical contact is developed.

The approximation presented involves the determination of the probability density function of peak heights on a random surface and the density of such peaks.

One method of obtaining surface roughness information is to observe relative surface height along one-dimensional profiles. We discuss the estimation of spectral moments of these one-dimensional processes. We then discuss a method of combining this information from several profiles in different directions to compute the spectral moments of the two-dimensional process.

We conclude with a brief discussion of the need for filtering the profile processes in order to obtain better agreement with the contact process.

REFERENCES

1. Tallian, T. E.; Chiu, Y. P.; Huttenlocher, D. F.; Kamenshire, J. A.; Sibley, L. B.; and Sindlinger, N. E. (1964). Lubricant Films in Rolling Contact of Rough Surfaces. ASLE Trans., 7, pp. 109-126.

2. Williamson J. B. P. (1969). The Shape of Solid Surfaces. Surface Mechanics, ASME, pp. 24-35.

3. Longuet-Higgins, M. S. (1957). The Statistical Analysis of a Random, Moving Surface. Phil. Trans. Roy. Soc., 249, pp. 321-387.

4. Longuet-Higgins, M. S. (1957). Statistical Properties of an Isotropic Random Surface. Phil. Trans. Roy. Soc., 250, pp. 157-174.

5. Longuet-Higgins, M. S. (1962). The Statistical Geometry of Random Surfaces. Proc. Sym. Appl. Math., vol. 13, G. Birkhoff, R. Bellman, and C. C. Lin, eds., pp. 105-143.

6. Nayak, P. Ranganath (1971). Random Process Model of Rough Surfaces. J. Lub. Tech., 93, pp. 398-407.

7. Cramer, H., and Leadbetter, M. R. (1967). Stationary and Related Stochastic Processes, J. Wiley & Sons, New York.

8. Kac, M.; and Slepian, D. (1959). Large Excursions of Gaussian Processes. Ann. Math. Statist., 30, pp. 1215-1228.

9. Tallian, T. E. (1972). The Theory of Partial Elastohydrodynamic Contacts. Wear, 21, pp. 49-101.

10. Bendat, J. S. (1958). Principles and Applications of Random Noise Theory, J. Wiley & Sons, New York.

11. Slepian, David (1962). The One-Sided Barrier Problem for Gaussian Noise. Bell System Tech., J. 41, pp. 463-501.

12. Rice, S. O. (1958). Distribution of the Duration of Fades in Radio Transmission: Gaussian Noise Model. Bell System Tech., J., 37, pp. 581.635.

13. Rice, S. O. (1945). Mathematical Analysis of Randon Noise. Bell System Tech. J., 24, pp. 46-156.

14. Graybill, F. A. (1969). Introduction to Matrices with Applications in Statistics. Wadsworth Publ. Co., New York.

15. Jenkins, G. M., and Watts, D. G. (1968). Spectral Analysis and Its Applications. Holden-Day, San Francisco.

16. Bendat, J. S., and Piersol, A. G. (1971). Random Data: Analysis and Measurement Procedures, John Wiley & Sons, New York.

17. Whitehouse, D. J.; and Archard, J. F. (1969). The Properties of Random Surfaces in Contact. Surface Mechanics, ASME, pp. 36-57.

18. Peklenik, J. (1968). New Developments in Surface Characterization and Measurements by Means of Random Process Analysis. Proc. Inst. Mech. Eng., 128, pp. 108-126.

19. Doob, J. L. (1953). Stochastic Processes, J. Wiley & Sons, New York.

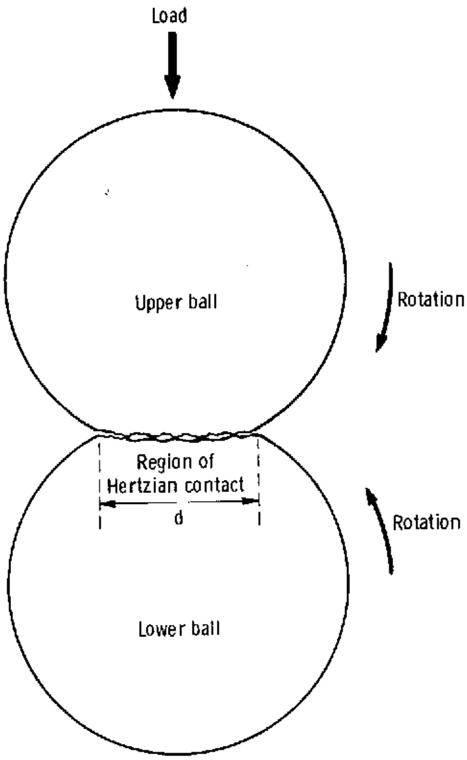

Figure 1. - Idealization of lubricated balls in contact. (Scale of Hertzian contact greatly exaggerated.)

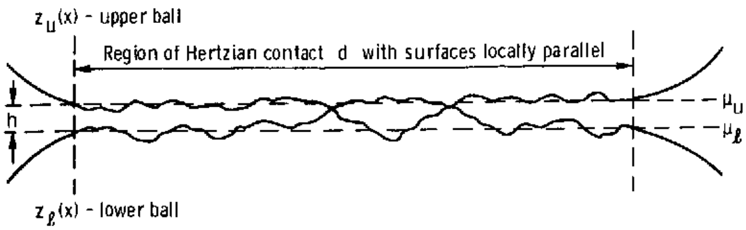

Figure 2. - Idealization of Hertzian region (cross section).

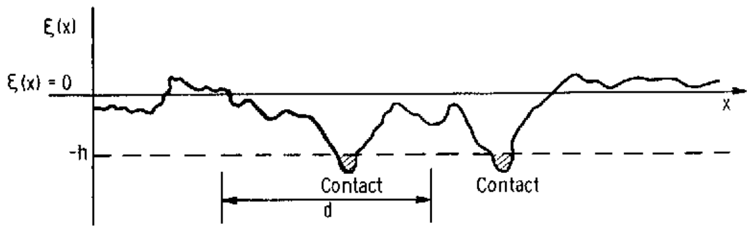

Figure 3. - Composite process defined by $\xi(x) = z_u(x) - z_\ell(x) - h$.

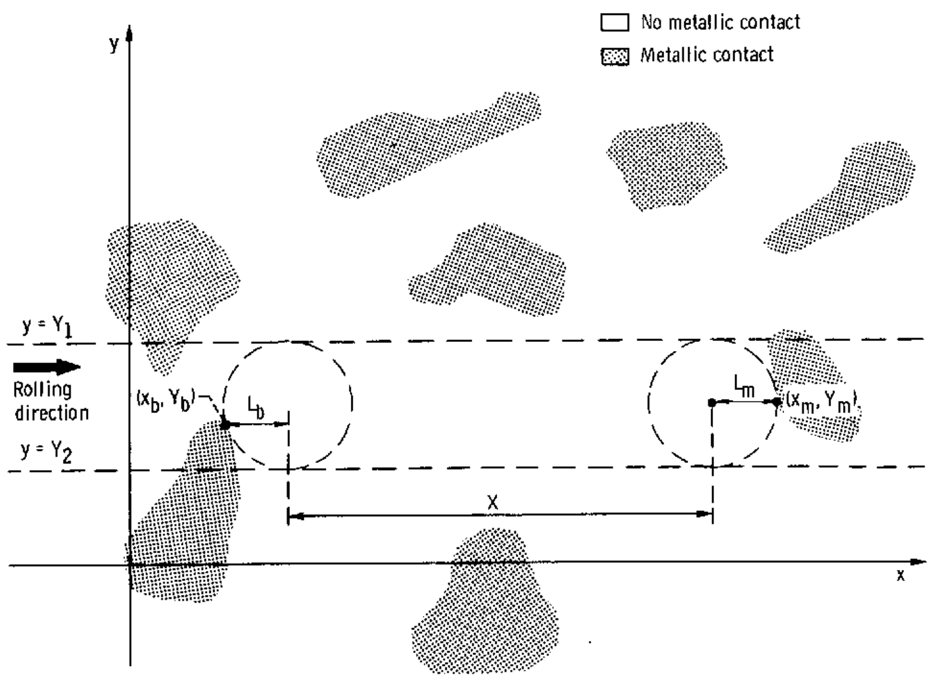

Figure 4. - Typical break-contact occurrence.

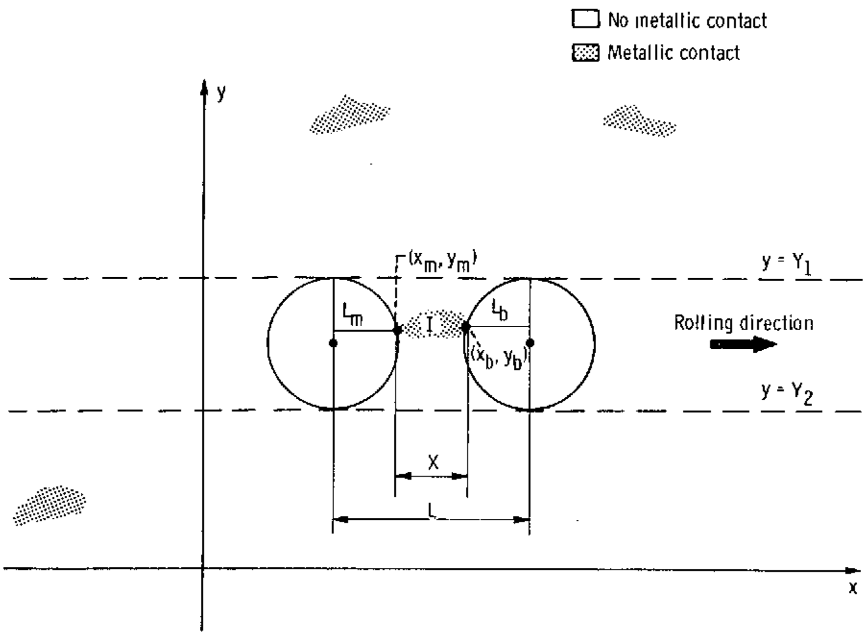

Figure 5. - Typical make-contact occurrence when contact areas are small.

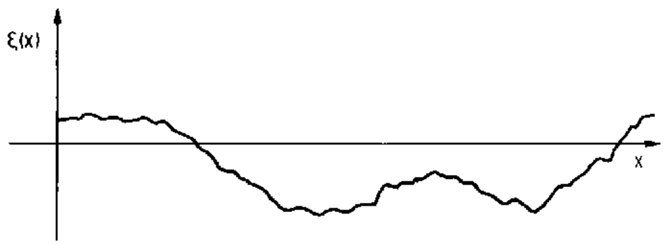

Figure 6. - Typical profile with high frequencies of small amplitude superimposed on a low-frequency signal.

Figure 7. - Typical profile containing primarily medium-frequency components superimposed on a signal with low frequency.