N73-15620

NASA TECHNICAL NOTE

NASA TN D-7134

NASA TN D-7134

CASE FILE COPY

ANALYSIS OF ELECTRICAL CONTACT OCCURRENCES BETWEEN ROLLING SURFACES WITH APPLICATION TO ELASTOHYDRODYNAMIC LUBRICATION

by Steven M. Sidik

Lewis Research Center

Cleveland, Ohio 44135

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION • WASHINGTON, D. C. • JANUARY 1973

1. Report No. NASA TN D-7134	2. Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle ANALYSIS OF ELECTRICAL CONTACT OCCURRENCES BETWEEN ROLLING SURFACES WITH APPLICATION TO ELASTOHYDRODYNAMIC LUBRICATION		5. Report Date January 1973
		6. Performing Organization Code
7. Author(s) Steven M. Sidik		8. Performing Organization Report No. E-7166
		10. Work Unit No. 501-24
9. Performing Organization Name and Address Lewis Research Center National Aeronautics and Space Administration Cleveland, Ohio 44135		11. Contract or Grant No.
		13. Type of Report and Period Covered Technical Note
12. Sponsoring Agency Name and Address National Aeronautics and Space Administration Washington, D.C. 20546		14. Sponsoring Agency Code
15. Supplementary Notes		

16. Abstract

A mathematical model enabling the relation of electrical contact occurrences between rolling surfaces to the mean thickness of an elastohydrodynamic lubricant film is discussed. A previous development by other authors is shown to contain an error invalidating their results. Some potentially useful approximations are developed for the case when the lubricant film is large with respect to the surface roughness. Complete solutions that are valid for all film thicknesses are not yet available.

17. Key Words (Suggested by Author(s))		18. Distribution Statement	
Elastohydrodynamic lubrication Randon noise processes Rolling surfaces		Unclassified - unlimited	
19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price*
Unclassified	Unclassified	25	$3.00

*For sale by the National Technical Information Service, Springfield, Virginia 22151

ANALYSIS OF ELECTRICAL CONTACT OCCURRENCES BETWEEN ROLLING SURFACES WITH APPLICATION TO ELASTOHYDRODYNAMIC LUBRICATION

by Steven M. Sidik

Lewis Research Center

SUMMARY

A mathematical model enabling the relation of electrical contact occurrences between rolling surfaces to the mean thickness of an elastohydrodynamic lubricant film is discussed. A previous development by other authors is shown to contain an error invalidating their results. Some potentially useful approximations are developed for the case when the lubricant film is large with respect to the surface roughness. Complete solutions that are valid for all film thicknesses are not yet available.

INTRODUCTION

It has been shown in reference 1 that a mathematical model for the elastohydrodynamic (EHD) lubrication of rolling surfaces can be facilitated by representing the surfaces as two-dimensional, stationary and ergodic, Gaussian random processes. This model has been used to estimate the average lubricant film thickness from experimental measurements by electrical contact methods. This report points out a fundamental error in reference 1, derives the correct method of analysis, and briefly discusses the areas where further theoretical developments must be made to make the mathematical model more realistic. Some approximate methods are proposed for both one- and two-dimensional analyses when the film thickness is large compared to the bearing roughness.

Besides the supporting data presented in reference 1, Williamson (ref. 2) has discussed the applicability of Gaussian processes to the analysis of rough surfaces.

A fundamental fault in several previous analyses of two-dimensional random surfaces (refs. 1 and 3, for example) is the assumption that the properties of surfaces are directly described by one-dimensional profile traces. Nayak (ref. 4) has discussed the appropriate techniques for characterizing two-dimensional surfaces from one-dimensional profile traces under the assumption of statistically isotropic surfaces. He shows that a naive analysis assuming that profile statistics may be directly used is erroneous. The bulk of Nayak's results are applications or extensions of results obtained by Longuet-Higgins in two lengthy and excellent papers (refs. 5 and 6).

In this report, we begin with the one-dimensional case to develop the correct method of analysis. Then a generalization of the electrical contact occurrence method is made to two dimensions. Although a complete solution is not developed herein, approximate solutions are presented for both the one- and two-dimensional cases when the film thickness is large compared to the surface roughness. A brief discussion of the assumption of isotropic surfaces is also presented.

It is appropriate to remark at this point that we consider only a somewhat simplified version of the real problem. Some of the complications arising in real life are bearing spin, bearing sliding, viscosity effects, nonuniform lubricant thickness, temperature-related effects, and so forth. However, until an adequate solution to the simplified problem which ignores these effects is found, little progress will be made on the more complicated problems.

ONE-DIMENSIONAL ANALYSIS

Relation of Film Thickness and Surface Roughness to Electrical Contact Occurrences

The basic problem under consideration here is that of describing the behavior of two rolling surfaces separated by a thin film of lubricant. 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. Under differing conditions of load, speed, surface roughness, and so forth, the effectiveness of the lubricant film in achieving this purpose differs. This report provides a means of experimentally determining the mean film thickness. A first step in determining the average film thickness from experiment data and analyzing its effectiveness was presented in reference 1.

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(a).

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 $y_u(x)$. We use the notation $y_u(x) \sim N(\mu_u, C_u(\tau))$ to denote that $y_u(x)$ is distributed as a Gaussian random process with mean μ_u and autocovariance function (acvf)$C_u(\tau)$. The lower surface is denoted by $y_l(x)$ and we have $y_l(x) \sim N(\mu_l, C_l(\tau))$. We assume both y_u and y_l to be stationary and ergodic. The distance between the surfaces at any point x is described by $y_u(x) - y_l(x)$. Thus, we consider the composite process defined by

$$s(x) = y_u(x) - y_l(x) - h \tag{1}$$

This process is illustrated in figure 2(b). We have (ref. 5)

$$s(x) \sim N(0, C_s(\tau))$$

where $C_s(\tau) = C_u(\tau) + C_l(\tau)$. Note that $C_s(0) = \sigma_s^2$. Whenever $s(x) > -h$, the surfaces are not in the state of metallic contact. Whenever $s(x) \leq -h$, the surfaces are in the state of metallic contact. Obviously, the physical process cannot have $s(x) < -h$. We assume that $s(x) < -h$ 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 will be recorded whenever there is metallic contact anywhere within the region of Hertzian contact. A voltage will be observed only whenever 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 $s(x)$ above the level $s(x) = -h$, electrical contact is not broken until the left edge of the Hertzian contact includes the first point at which $s(x) > -h$. 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 $s(x) > -h$. Electrical contact is reestablished as soon as the right edge of the Hertzian contact includes the first point for which $s(x) = -h$ 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 $(x)^+ = \text{MAX}\{x, 0\}$.

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

$\delta_a(h)$	rv denoting lengths of excursions of $s(x)$ above level $-h$
$N_a(t,h)$	rv denoting number of crossings of $s(x)$ from below $-h$ to above $-h$ in interval $(0,t)$
$N_a(t,h \mid \delta \leq d)$	rv denoting number of crossings of $s(x)$ 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 $\delta_b(h)$, $N_b(t,h)$, and $N_b(t,h \mid \delta \leq d)$ 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 ref. 8). 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):

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

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

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

Because of the stationarity and ergodicity assumptions, the approximate equalities become equality when $t \to \infty$. 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,

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

But we have

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

$$\approx (t_a + t_b)\,\frac{N(t, h)}{2}$$

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

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

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

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

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 quantities are unknown. There are some asymptotic results for $\delta_b(h)$ as $h \to \infty$, however. 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 $E(L)$. Before we proceed to consider the application of the asymptotic results, we present a comparison of the results obtained here and those given in reference 1.

Discussion of Results of Reference 1

Reference 1 and reference 9 were the first to discuss the relation of electrical contact measurements to the average film thickness. The mathematical and probabilistic arguments forming the basis for the discussions in reference 1 are given in reference 9 and a series of subsequent progress reports (refs. 10 and 11). Their results for the no-contact fraction for the one-dimensional case are not in agreement with ours and some discussion of this disagreement is in order.

In reference 9 there are actually four slightly different statements for the expression for $1 - T_c$. Each of these statements is given here and each is then discussed. The notation of reference 9 is changed to agree with the notation of this report.

(1) Reference 9, page 13: "Thus, the (time) average fraction of the no-contact time is equal to the probability of finding intervals longer than d where $s(x) > -h$."

(2) Reference 9, page 13: $1 - T_c = \Pr\left\{s(x) > -h \mid x_0 < x < x_0 + d\right\}$

(3) Reference 9, pages 13 and 14: "The probability (given in item (2) above) is equal to the probability of finding intervals of length $L > d$ within which the random process does not cross over the line $s(x) = -h$, whereas it crosses this line upward ahead of the interval and downward behind the interval. This probability can be computed if the distribution of "spacings" between subsequent upward and downward crossings of the process over the line $s(x) = -h$ is known."

(4) Reference 9, page 16: "From $\Pr\left\{\delta_a(h) > d\right\}$ we get the probability of finding a spacing $\delta_a(h) > d$ at an arbitrary point x by observing that this event requires

(a) That $s(x) > -h$ at x

(b) That, given (a), the x point be part of an interval of length $L > d$ in which $s(x) > -h$ at all points

Thus, the average no-contact time fraction $1 - T_c$ is $1 - T_c = \left[1 - \Phi(-h/\sigma)\right]\Pr\left\{\delta_a(h) > d\right\}$."

Statement (1) can be easily proved false by a simple counter example. Consider a Gaussian process defined by

$$s(x) = \sin\left(\frac{x}{3d}\right) + \epsilon(x) - h$$

where $\epsilon(x)$ is a band-limited white noise of arbitrarily small variance. That is the $s(x)$ process may be well represented by a sine wave of period $3d$. This is illustrated in figure 3. Note that for each excursion above $s(x) = -h$, there is loss of electrical contact only for the distance $(L - d)^+ = d/2$. For each period of length $3d$, therefore, there is a no-contact fraction 1/6, and hence $1 - T_c = 1/6$. It is evident, however, that the probability that any x chosen at random lies within an excursion whose length exceeds d is 1/2. But $1 - T_c$ does not equal 1/2, and hence the statement is false.

Statement (2) appears to be correct if interpreted properly. If we let all x be equally likely (not clearly possible since it would involve an improper distribution), then consider the probability of choosing an x_0 such that $s(x) > -h$ for all $x \epsilon (x_0, x_0 + d)$. The requirement that $s(x) > -h$ for all $x \epsilon (x_0, x_0 + d)$ effectively subtracts the length d from the excursions above $s(x) = -h$ and hence accomplishes the same thing as computing the conditional expectation $E\left\{\delta_a(h) \mid \delta_a(h) > d\right\}$.

Statement (3) is almost identical to statement (2) and the same comments should apply.

Statement (4) is the one from which the actual evaluation of $1 - T_c$ is developed in reference 9. This statement suffers from the same inaccuracy as statement (1). Namely, there is no provision for subtracting d from the length of the excursion. This may be seen by comparing their result to ours. From arguments presented in reference 9 it can be seen that

$$1 - \Phi\left(\frac{-h}{\sigma}\right) = \frac{t_a}{t_a + t_b}$$

Thus, from reference 9 we get

$$1 - T_c = \frac{t_a}{t_a + t_b} \Pr\left\{\delta_a(h) > d\right\}$$

whereas equation (2) of this report yields

$$1 - T_c = \frac{\overline{T}_N - d}{t_a + t_b} \Pr\left\{\delta_a(h) > d\right\}$$

The difference is clearly in the factors t_a and $\overline{T}_N - d$ in the numerator.

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_s is large.

For this situation some results from Cramer and Leadbetter (ref. 7) and from Rice (ref. 12) and Kac and Slepian (ref. 13) 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 symmetrical about the mean, it is evident that similar results apply to excursions and crossings below a very low level.

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

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

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

Since

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

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 $s(x) = -h$ and let θ_e denote the mean of this distribution. Then it is shown in reference 7 that

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

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

Since for large h

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

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_s and t,

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

We note that Rice (ref. 14) has shown

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

and hence

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

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

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

or

$$\frac{h}{\sigma_s} = \sqrt{2}\left(\ln\left\{\frac{d}{T_c\pi}\left[\frac{-C_s''(0)}{\sigma_s^2}\right]^{1/2}\right\}\right)^{1/2} \qquad (3)$$

In order to apply this equation, σ_s^2 and $C_s''(0)$ must be estimated. One method of estimating $C_s''(0)$ is to note that this value is also the negative of the second spectral moment. Thus, the spectrum could be estimated and its moments found.

This completes the discussion of the one-dimensional analysis. The two-dimensional case is of much more benefit practically; so we now proceed to consider this, using the insights developed from this one-dimensional analysis.

TWO-DIMENSIONAL ANALYSIS

Derivation of Contact Fraction

At this point we reiterate 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. 4) 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. 4) and Longuet-Higgins (refs. 5 and 6) 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. Although these are important quantities in surface analyses, they do not directly contribute to the problem of electrical contact analyses.

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.

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

$$s(x, z) = y_u(x, z) - y_l(x, z) - h$$

where

$$y_u(x, z) \sim N\left(\mu_u, C_u(\tau_x, \tau_z)\right)$$

$$y_l(x, z) \sim N\left(\mu_l, C(\tau_x, \tau_z)\right)$$

and hence

$$s(x, z) \sim N\left(0, C_s(\tau_x, \tau_z)\right)$$

where

$$C_s(\tau_x, \tau_z) = C_u(\tau_x, \tau_z) + C_l(\tau_x, \tau_z)$$

Figure 4 presents a visualization of the metallic-contact areas in the plane parallel to the (x, z) 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 $z = Z_1$ and $z = Z_2$. The areas for which $s(x, z) \leq -h$ are indicated by the enclosed dotted regions. For simplicity, we also assume that the process is statistically isotropic and that the Hertzian area is circular. These assumptions are discussed further in the section Isotropy Assumption.

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 know 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_s is large.

Computational Method as h Approaches ∞

As h/σ_s gets large we might expect the same asymptotic Poisson character of the number of "valleys" to occur in the two-dimensional process as occurs in the one-dimensional process. 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 $s(x, z) = -h$. The areas for which $s(x, z) \leq -h$ 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 $z = Z_1$ and $z = Z_2$. 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, z_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, z_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 z_M and z_B to be quite highly and positively correlated. Thus, assume z_M and z_B are perfectly correlated and that z_M is uniformly distributed over the vertical distance bounded by $z = Z_1$ and $z = Z_2$. 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 $E(L_M)$. For example, if the Hertzian area is a circle of radius $d/2$, then

$$E(L_M) = \frac{1}{Z_2 - Z_1} \int_{z=Z_2}^{z=Z_1} \left[\frac{d^2}{4} - \left(z - z_A\right)^2\right]^{1/2} dz$$

where

$$z_A = \frac{Z_1 + Z_2}{2}$$

Some of the results of Longuet-Higgins (refs. 5 and 6) and Nayak (ref. 4) concerning the distribution of summit heights may now be employed if we assume the surface is isotropic. We first introduce some additional notation and results about two-dimensional random processes which will be needed. We again note that because the Gaussian process is symmetric, similar results apply to depths of valleys. For the sake of notational correspondence we will continue in terms of peaks and summits rather than nadirs. This is equivalent to discussing the process $-s(x)$.

We let $\mathcal{S}_\rho(k)$ denote the power spectral density function of a profile taken at an angle ρ along the (x, z) plane. Nayak shows that this power spectral density and its moments are all the information needed to describe the distribution of summit heights if the surface is isotropic. In particular, let m_i denote the i^{th} moment of $\mathcal{S}_\rho(k)$. That is,

$$m_i = \int_{-\infty}^{\infty} \mathcal{S}_\rho(k) k^i \, dk = C_\rho^{(i)}(0)\left(\sqrt{-1}\right)^i$$

where $C_\rho^{(i)}(0)$ denotes the i^{th} derivative of the autocovariance function for the profile along ρ evaluated at zero.

Let $s^* = s/m_o^{1/2} = s/\sigma_s$

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

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

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

and

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

Let D_{sum} be defined as the expected number of summits per unit area and $p(s^*)$ denote the probability density function of the summit heights in a normalized random process defined by $s(x,z)/\sigma_s$. Then

$$D_{sum} = \frac{1}{6\pi\sqrt{3}}\left(\frac{m_4}{m_2}\right)$$

and

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

Since the area over which the Hertzian region passes is of width $z_A = Z_1 - Z_2$, we simply multiply D_{sum} by z_A to get the expected number of summits occurring within the strip per unit distance along the x-axis, $\bar{s}_h$. We recall the summits are few and far between. Thus, the probability that two peaks occur within the Hertzian area is negligible.

The average fraction of contact distance T_c is thus given by $\bar{s}_h$, the expected number of summits that exceed the level $h(h/\sigma_s$ in the normalized process) per unit distance along x, multiplied by the average distance contact is maintained for each excursion. That is,

$$T_c \approx \bar{s}_h 2E(L_M)$$

$$\approx \frac{1}{3\pi\sqrt{3}}\left(\frac{m_4}{m_2}\right)E(L_M)\int_{h/\sigma_s}^{\infty} p(s^*)ds^*$$

This quantity may be numerically evaluated.

Isotropy Assumption

It is of interest to note that the assumption of isotropy may not be necessary. This would be of some importance in practice. Isotropy is quite reasonable when considering balls rolling on each other. However, balls rolling in a raceway will not, in general, exhibit isotropy because of the particular processes used to finish the surfaces. In particular, raceway surfaces would typically be expected to have long grooves and hills running in the direction of grinding and much shorter grooves and hills in the normal direction. Assume that the grinding direction is parallel to the rolling direction of the ball in the raceway. If we assume the acvf is of the form

$$C_s(\tau_x, \tau_z) = \sigma_s^2 \exp\left[-\alpha\ \tau_x^2 + \beta\tau_z^2\right]^{1/2}$$

then it is easily seen that the process defined by $s(x, z/\ \beta)$ is isotropic. The only effect such a transformation of variables has on the analysis is that the shape of the Hertzian area will be changed. In particular, a circle would be transformed to an ellipse. This produces no fundamental changes in the analysis of electrical contact occurrences.

An alternative method of considering nonisotropic surfaces which does not involve knowledge of β is to extend the results of Nayak to nonisotropic surfaces as outlined in his paper (ref. 4). Such an extension does not involve any new principles. Some additional work would need to be done regarding the expressions for D_{sum} and $p(s^*)$.

Prefiltering of Signals

Whitehouse and Archard (ref. 3) 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 elastic 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 there to be 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 even arise in practice. The most serious problem is with the high-frequency signals, since they contribute the most to the number of summits. Whitehouse and Archard (ref. 3) 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 open for research. Clearly, the cutoff frequencies for filtering must be related to such things as the Hertzian area dimensions, the rolling speed of the surfaces, the load, and the elasticity of the materials.

CONCLUDING REMARKS

The purpose of this report is to show that reference 1 presents an erroneous development of the relation between average film thickness h and contact fraction and to present a correct development of the relation. This has been done. The general analysis that is valid for all values of h is seen to lead to an unsolved (so far) problem in random noise theory. The problem is that of describing the distribution of excursion lengths above a specified level. This has not yet been solved even in the one-dimensional case except for certain forms of the autocovariance function and certain levels.

We have, however, presented approximate solutions for the case when the average film thickness is large. These approximations should prove of considerable value in practice since the area of primary interest in bearing analyses is the case when h/σ_s is large.

Lewis Research Center,
National Aeronautics and Space Administration,
Cleveland, Ohio, October 20, 1972,
501-24.

APPENDIX - SYMBOLS

$A_h(\delta_a(h))$	distribution function of $\delta_a(h)$
$B_h(\delta_b(h))$	distribution function of $\delta_b(h)$
$C(\tau_x)$, $C(\tau_x, \tau_z)$	autocovariance functions
$C'(\tau)$, $C''(\tau)$, $C^{(i)}(\tau)$	denote derivatives of $C(\tau)$
D_{sum}	average number of summits per unit area
d	length of Hertzian area
E(A)	expectation of A
$E(A \mid B)$	conditional expectation of A given event B
F_c, F_e	distribution functions of crossings and excursions, respectively
h	average lubricant film thickness
L	length of an excursion of random noise process above a given level
m_i	i^{th} spectral moment
$N_a(t,h)$, $N_b(t,h)$	random variable denoting number of crossings of s(x) from below (above) -h to above (below) -h in interval (0,t)
$N_a(t,h \mid \delta \leq d)$, $N_b(t,h \mid \delta \leq d)$	random variable denoting number of crossings from below (above) -h to above (below) -h in interval (0,t) where excursion is of length less than or equal to d
$N_a(t,h \mid \delta > d)$, $N_b(t,h \mid \delta > d)$	random variable denoting number of crossings from below (above) -h to above (below) -h in interval (0,t) where excursion is of length greater than d
N(t,h)	random variable denoting number of either-way crossings of level -h in interval (0,t)
$N(\mu, C(\tau))$	normal random noise process with mean level μ and autocovariance function $C(\tau)$
Pr(A)	probability of event A
$p(s^*)$	probability density function of s^*
$\mathcal{S}_\rho(k)$	power spectral density of process corresponding to a profile taken at angle ρ

$s(x)$, $s(x, z)$	random noise process in one and two dimensions, respectively
s^*	normalized random noise process
$\bar{s}_h$	expected number of summits per unit distance in normalized random process whose heights exceed h
$\bar{T}_N$, $\bar{T}_C$	conditional expectation of excursion lengths given $\delta_a(h)$ as greater than d and less than or equal to d, respectively
T_c	contact fraction
t_a, t_b	expected length of excursion above (below) level $-h$
X	distance between center points of Hertzian area at make-contact and break-contact locations
$y(x)$, $y(x, z)$	random noise processes
$\delta_a(h)$, $\delta_b(h)$	random variable denoting length of excursion of $s(x)$ above (below) $-h$
$\epsilon(x)$	white noise process
θ_c	mean of distribution of length between successive downcrossings of $s(x) = -h$
θ_e	mean of distribution of excursion lengths below $s(x) = -h$
μ	mean level of random noise process
σ	standard deviation of random process
$\Phi(x)$	error function or cumulative normal distribution function

Subscripts:

a	crossings and excursions above $-h$
b	crossings and excursions below $-h$
B	break-contact occurrences
l	lower ball
M	make-contact occurrences
p	profile statistics
s	composite process $s(x)$ or $s(x, z)$
u	upper ball

$[x]^+$	positive part of x, that is, MAX (0,x)
ρ	profile statistics in direction ρ
$\sim$	"distributed as"
$\approx$	"approximately equal"

REFERENCES

1. Tallian, T. E.; Chiu, Y. P.; Huttenlocher, D. F.; Kamenshine, J. A.; Sibley, L. B.; and Sindlinger, N. E.: Lubricant Films in Rolling Contact of Rough Surfaces. ASLE Trans., vol. 7, no. 2, Apr. 1964, pp. 109-126.
2. Williamson, J. B. P.: The Shape of Solid Surfaces. Surface Mechanics. ASME 1969, pp. 24-35.
3. Whitehouse, D. J.; and Archard, J. F.: The Properties of Random Surfaces in Contact. Surface Mechanics. ASME, 1969, pp. 36-57.
4. Nayak, P. Ranganath: Random Process Model of Rough Surfaces. J. Lub. Tech., vol. 93, no. 3, July 1971, pp. 398-407.
5. Longuet-Higgins, M. S.: The Statistical Analysis of a Random, Moving Surface. Phil. Trans. Roy. Soc., Ser. A, vol. 249, no. 966, Feb. 21, 1957, pp. 321-387.
6. Longuet-Higgins, M. S.: Statistical Properties of an Isotropic Random Surface. Phil. Trans. Roy. Soc., Ser. A, vol. 250, no. 975, Oct. 17, 1957, pp. 157-174.
7. Cramér, Harald; and Leadbetter, M. R.: Stationary and Related Stochastic Processes. John Wiley & Sons, Inc., 1967.
8. Slepian, David: The One-Sided Barrier Problem for Gaussian Noise. Bell System Tech. J., vol. 41, no. 2, Mar. 1962, pp. 463-501.
9. Sindlinger, N. E.; Martin, J. A.; and Huttenlocher, D. F.: Influence of Lubrication on Endurance of Rolling Contacts. Rep. AL62T013, SKF Industries, Inc., 1962.
10. Schmidt, W.; et al.: Influence of Lubrication on Endurance of Rolling Contacts. Rep. AL63T016, SKF Industries, Inc., 1963.
11. Brady, E. F.; McCool, J.; Sibley, L. B.; and Valori, R.: Influence of Lubrication on Endurance of Rolling Contacts. Rep. AL64T003, SKF Industries, Inc. (AD-430287), 1963.
12. Rice, S. O.: Distribution of the Duration of Fades in Radio Transmission: Gaussian Noise Model. Bell System Tech. J., vol. 37, no. 3, May 1958, pp. 581-635.
13. Kac, M.; and Slepian, D.: Large Excursions of Gaussian Processes. Ann. Math. Statist., vol. 30, 1959, pp. 1215-1228.
14. Rice, S. O.: Mathematical Analysis of Random Noise. Bell System Tech. J., vol. 24, no. 1, Jan. 1945, pp. 46-156.

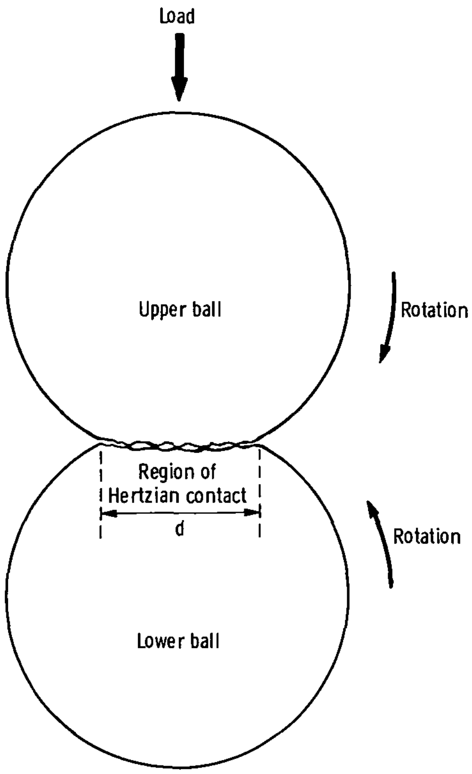
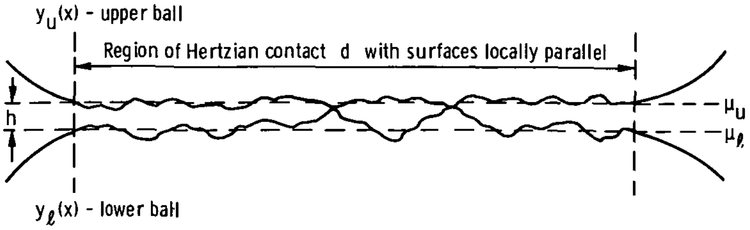
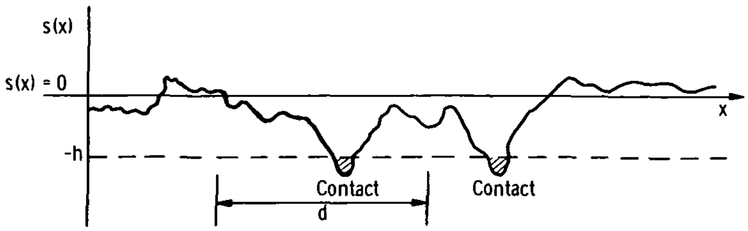


Figure 1. - Idealization of lubricated balls in contact. (Scale of Hertzian contact greatly exaggerated.)



(a) Idealization of Hertzian region (cross section).



(b) Composite process defined by $s(x) = y_u(x) - y(x) - h$.

Figure 2. - Detail of a representative cross section of the Hertzian region and the corresponding composite process.

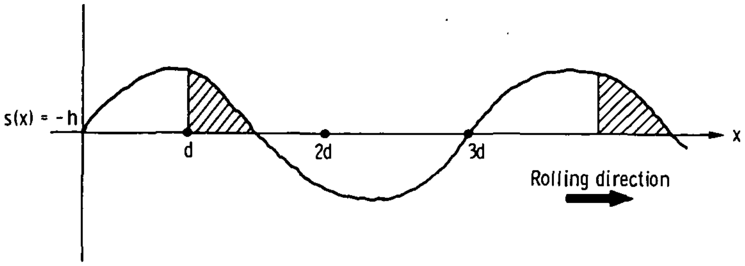


Figure 3. - Process $s(x) = \sin\left(\frac{x}{3d}\right) + \varepsilon(x)$.

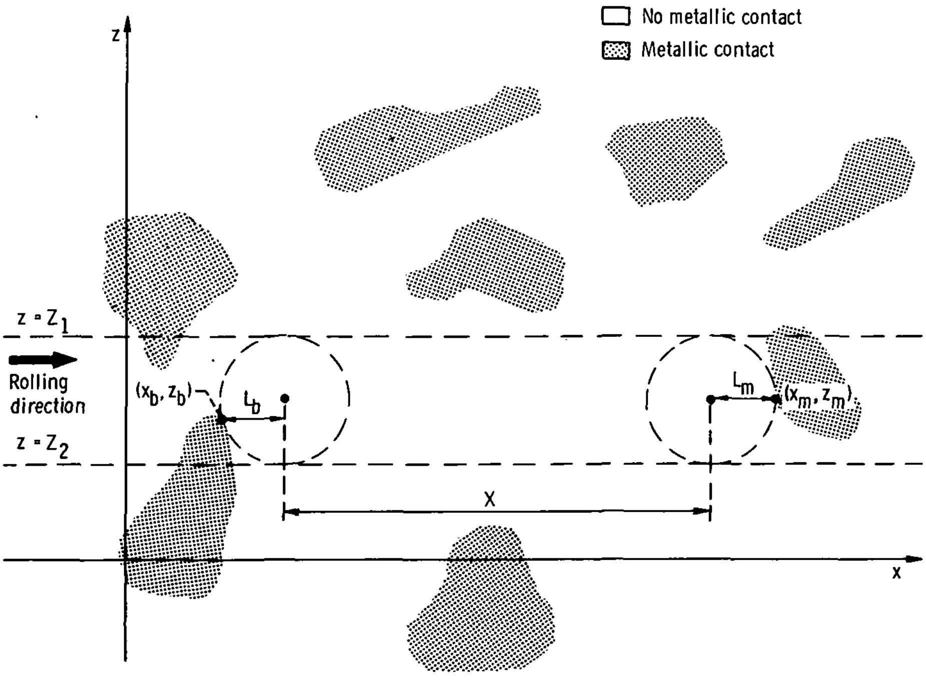


Figure 4. - Typical break-contact occurrence.

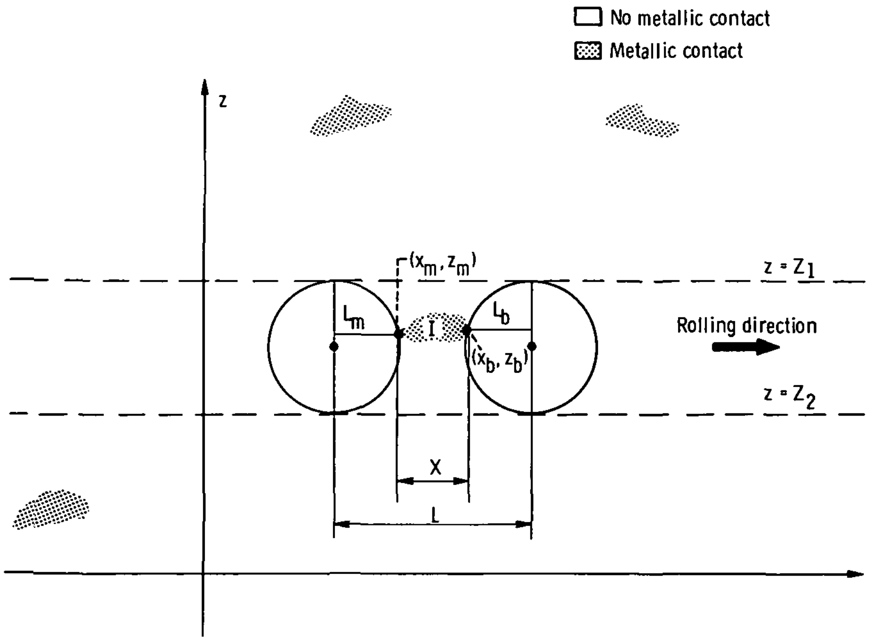


Figure 5. - Typical make-contact occurrence when contact areas are small.

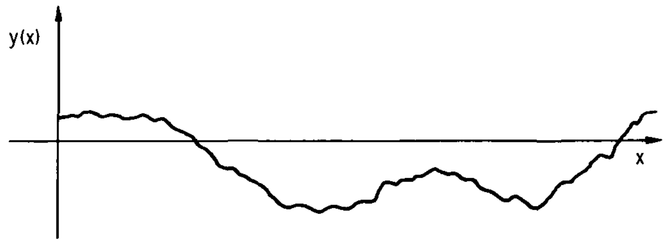


Figure 6. - Typical profile with high frequencies of small amplitude superimposed on a low-frequency signal.

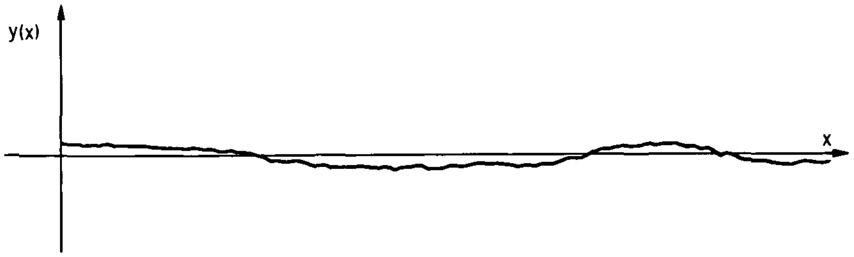


Figure 7. - Typical profile containing primarily medium-frequency components superimposed on a signal with low frequency.

NASA-Langley, 1973 —— 19 E-7166

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION
WASHINGTON, D.C. 20546

OFFICIAL BUSINESS
PENALTY FOR PRIVATE USE $300

SPECIAL FOURTH-CLASS RATE
BOOK

POSTAGE AND FEES PAID
NATIONAL AERONAUTICS AND
SPACE ADMINISTRATION
451

U.S.MAIL

POSTMASTER: If Undeliverable (Section 158 Postal Manual) Do Not Return

---

"The aeronautical and space activities of the United States shall be conducted so as to contribute . . . to the expansion of human knowledge of phenomena in the atmosphere and space. The Administration shall provide for the widest practicable and appropriate dissemination of information concerning its activities and the results thereof."

—NATIONAL AERONAUTICS AND SPACE ACT OF 1958

NASA SCIENTIFIC AND TECHNICAL PUBLICATIONS

TECHNICAL REPORTS: Scientific and technical information considered important, complete, and a lasting contribution to existing knowledge.

TECHNICAL NOTES: Information less broad in scope but nevertheless of importance as a contribution to existing knowledge.

TECHNICAL MEMORANDUMS: Information receiving limited distribution because of preliminary data, security classification, or other reasons. Also includes conference proceedings with either limited or unlimited distribution.

CONTRACTOR REPORTS: Scientific and technical information generated under a NASA contract or grant and considered an important contribution to existing knowledge.

TECHNICAL TRANSLATIONS: Information published in a foreign language considered to merit NASA distribution in English.

SPECIAL PUBLICATIONS: Information derived from or of value to NASA activities. Publications include final reports of major projects, monographs, data compilations, handbooks, sourcebooks, and special bibliographies.

TECHNOLOGY UTILIZATION PUBLICATIONS: Information on technology used by NASA that may be of particular interest in commercial and other non-aerospace applications. Publications include Tech Briefs, Technology Utilization Reports and Technology Surveys.

Details on the availability of these publications may be obtained from:

SCIENTIFIC AND TECHNICAL INFORMATION OFFICE

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION

Washington, D.C. 20546