PROCEEDINGS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

The Properties of Random Surfaces of Significance in their Contact

D. J. Whitehouse and J. F. Archard

Proc. R. Soc. Lond. A 1970 **316**, doi: 10.1098/rspa.1970.0068, published 31 March 1970

Proc. Roy. Soc. Lond A. **316**, 97–121 (1970)
Printed in Great Britain

The properties of random surfaces of significance in their contact

By D. J. Whitehouse
Rank Precision Industries Ltd

and J. F. Archard
Department of Engineering, The University of Leicester

(*Communicated by D. Tabor, F.R.S.—Received* 11 *August* 1969—
Revised 27 *October* 1969)

In recent work it has been shown that many types of surfaces used in engineering practice have a random structure. The paper takes, as a representation of the profile of such a surface, the waveform of a random signal; this is completely defined by two parameters, a height distribution and an auto-correlation function. It is shown how such a representation can be transformed into a model, appropriate for the study of surface contact, consisting of an array of asperities having a statistical distribution of both heights and curvatures. This theory is compared with the results of an analysis of surface profiles presented in digital form. The significance of these findings for the theory of surface contact and for the measurement and characterization of surface finish is discussed.

1. Introduction

All surfaces are rough. This is the starting-point from which current ideas about friction, wear, and other aspects of surfaces in contact have evolved. Because surfaces are rough the true area of contact, which is much smaller than the apparent area in contact, must support pressures so large that they are comparable with the strengths of the materials of the contacting bodies. In their earlier work Bowden & Tabor (1954) suggested that these contact pressures are equal to the flow pressure of the softer of the two contacting materials and the normal load is then supported by plastic flow of its asperities. The true area of contact, A, is then proportional to the load, W; thus it was possible to provide a simple and elegant explanation of Amontons's laws of friction. However, if the asperities are plastically deformed the details of the surface finish seem relatively unimportant since the total area of contact and the contact pressure do not depend upon surface topography.

More recently it has been recognized that surface contact must often involve an appreciable proportion of asperity contacts at which the deformation is entirely elastic. It has been shown that, under conditions of multiple contacts, even if the deformation were entirely elastic, the true area of contact, A, can increase almost proportionally with the load, W (Archard 1957; Greenwood & Williamson 1966); thus a satisfactory explanation of Amontons's laws of friction is not dependent upon the assumption of plastic deformation. It therefore becomes more important

to understand, in some detail, the role which surface topography plays in the contact of surfaces. For example, it is clear that surface finish will play a large part in determining the proportions of elastic and plastic deformation which will occur under any given set of conditions.

Knowledge of the topography of surfaces has been derived from the use of many techniques of surface examination. However, in considering the contact of nominally flat surfaces, the most relevant information has come from the use of profile meters in which a lightly loaded stylus is moved across the surface. In the past it has sometimes been assumed that the only significant information thus obtained could be expressed as the r.m.s. or centre line average (c.l.a.) value of the profile. However, in more recent years the outputs of profile meters have been analysed in greater detail by both analogue and digital techniques. In the field of production engineering this information has been presented in many different ways; height distributions, slope distributions, power spectral density curves, and auto-correlation functions are but a few of the characteristics which have been displayed (Peklenik 1967–8).

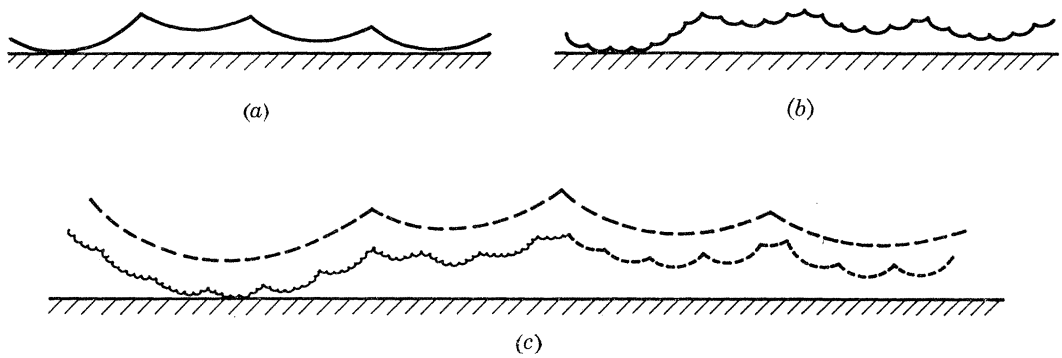

FIGURE 1. Models of surfaces containing asperities of differing scales of size. When the deformation is elastic the relationships between the area of contact (A) and the load (W) are as follows: (*a*) $A \propto W^{\frac{4}{5}}$; (*b*) $A \propto W^{\frac{14}{15}}$; (*c*) $A \propto W^{\frac{44}{45}}$.

On the other hand, those concerned with the problems of surface contact have used models of surfaces based upon many different assumptions about the nature of surface topography. Thus Archard (1957) postulated a series of models (figure 1) which were used to provide the first explanation of Amontons's laws of friction for elastic deformation of asperities. Although it was admitted that these models were artificial, they contain an important feature to which we shall revert later; this is the assumption that there exists upon the surfaces superposed asperities of widely differing scales of size.

It is, of course, important that the models used in theories of surface contact be more closely based upon knowledge gained from the examination of surface topography. Greenwood & Williamson (1966), and others, using information obtained by digital analysis of profile meter outputs, have shown that for many

surfaces the distribution of heights is very close to Gaussian. Greenwood & Williamson also made an investigation of the height distribution of peaks; the most common technique used in this investigation was three point analysis, a peak being defined when the central of three successive sampled heights lies above those on either side. Thus the distribution of peak heights was also shown to be close to Gaussian but both the mean value and the standard deviation of this distribution differed from that of the heights of the ordinates. In addition, by the same techniques, a distribution of peak curvatures was obtained; this was skewed towards lower values of curvature. As a result of these observations Greenwood & Williamson postulated a model, representing a rough nominally flat surface, consisting of a series of spherical peaks, each having the same radius of curvature, and having a Gaussian distribution of heights. On the basis of this model, and the assumption that the deformation was elastic, it was shown that the relation between A and W was close to direct proportionality; thus a second theoretical derivation of Amontons's laws for elastic deformation conditions was provided.

The theory of Greenwood & Williamson (1966), although representing a notable advance, is still far from a complete or accurate representation of random surfaces such as those analysed in their work; in particular the assumption of a single radius of curvature for surface asperities is clearly a major simplification of the model. Moreover, it will be shown below that the use, in their examination of surface profiles, of three point analysis together with a single sampling interval severely limits the information obtained from the surface profile.

The present paper considers the representation of a surface profile as a random signal. Although such a representation can be a complete description of the profile the problem lies in its transformation into a form appropriate to the study of surface contact. From earlier work this requirement is seen as a model consisting of a distribution of asperities; therefore the paper is concerned with the distribution of the heights of the peaks and their radii of curvature.

2. The model

A surface profile (figure 2), if it is of a random type, can be defined completely (in a statistical rather than a deterministic sense) by two characteristics: the height distribution and the autocorrelation function (see, for example, Bendat 1958). In the main body of this paper we shall confine our attention to the particular example of a surface profile having a Gaussian distribution of heights and an exponential autocorrelation function. There are a number of reasons for using this particular model. First, and foremost, analysis by one of us (D. J. Whitehouse, to be published) of a large number of surfaces used in engineering practice has shown that a not insignificant proportion of such surfaces fits this model or is a reasonable approximation to it. Moreover, this model has been widely used in the theory of random processes; it has also been used to represent surfaces in studies of the scattering of electromagnetic radiation (Beckmann & Spizzichino 1963).

The model also simplifies some aspects of the mathematics and allows a clearer statement of the important physical principles to emerge. Throughout this analysis it will be assumed that the surfaces are isotropic although it is possible, at least in principle, to extend the theory to surfaces having an anisotropic structure.

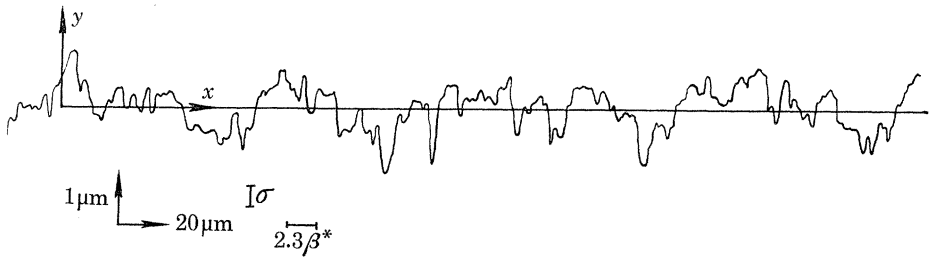

FIGURE 2. Surface profiles of Aachen 64–13 showing coordinate system. The profile is of a surface chosen for a detailed analysis which is described later in the paper. The magnitude of the r.m.s. value of the height distribution (σ) and the correlation distance β^* are shown for comparison with the profile.

The system of coordinates used is shown in figure 2; the mean line through the profile will be taken as $y = 0$. In practice the d.c. level, the general slope and the curvature of the surface are removed by a filter eliminating the longest wavelengths. This does not substantially affect the autocorrelation function. The probability of finding an ordinate at a height between h and $(h+\mathrm{d}h)$ is $f(h)\,\mathrm{d}h$. When the height distribution is Gaussian the height probability density function is

$$f(y) = (2\pi)^{-\frac{1}{2}} \exp\left(-\tfrac{1}{2}y^2\right); \tag{1}$$

the heights h have now been expressed in the normalized form $y = h/\sigma$, where σ is the standard deviation of the height distribution.

The autocorrelation function of the profile is defined as

$$C(\beta) = \lim_{L\to\infty} \frac{1}{L} \int_{-\frac{1}{2}L}^{+\frac{1}{2}L} y(x)\, y(x+\beta)\, \mathrm{d}x, \tag{2}$$

where $y(x)$ is the height of the profile at a given coordinate x and $y(x+\beta)$ is the height at an adjacent coordinate $(x+\beta)$. In the theory which follows it will be assumed that

$$C(\beta) = \exp\left(-\beta/\beta^*\right), \tag{3}$$

where β^* will be called the correlation distance.† When $\beta = 2.3\beta^*$, $C(\beta)$ has declined to 10 %; in what follows we shall, arbitrarily, take this spacing as that at which the two points on the profile have just reached the conditions where they can be regarded as independent events.

† We use the term 'correlation distance' to mark a distinction between ourselves and Peklenik (1967–8) who uses the term 'correlation length' for $2.3\beta^*$.

There exists a Fourier transform relation between the autocorrelation function $C(\beta)$ and the power spectral density function $P(\omega)$ of a random waveform given by

$$C(\beta) = \left(\frac{1}{2\pi}\right)\int_{-\infty}^{+\infty} P(\omega)\exp(i\omega\beta)\,d\omega. \tag{4}$$

For an exponential autocorrelation function the power spectral density function is represented by white noise limited only in the upper frequencies by a cut-off of 6 dB per octave. This is illustrated in figure 3. Thus the physical meaning of our

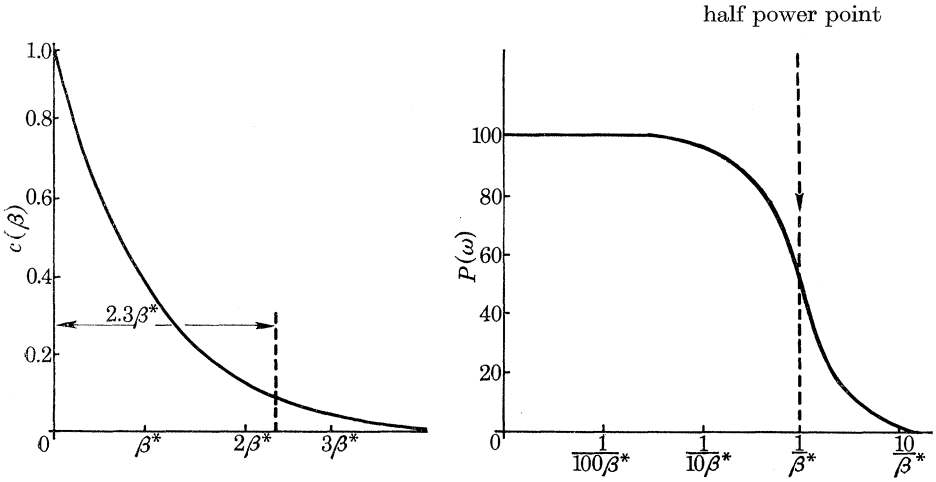

FIGURE 3. The model; autocorrelation function and power spectral density.

model is that the main components of the surface profile consists of a band covering the lower frequencies (longer wavelengths). Shorter wavelength components exist but their magnitude declines with increasing frequency so that, in this range, the amplitude is proportional to wavelength. Therefore, in broad terms, the random signal representation of a profile which has been introduced here has features akin to superposed asperities of differing scales of size; it introduces the multiple scale of size features of figure 1*c* (cf. figure 1*a*) into a random model representation.

We are concerned here with the properties of surfaces of significance in their contact. These are the heights of the peaks and their curvatures; they will be defined, as in the earlier work, by three point analysis. This technique will first be applied to the theoretical model of our random profile and the results of this theory compared with experimental results obtained from the digital presentation of profilometer records.

3. THEORY

As a starting point in the development of the theory we assume that the profile has been sampled as a sequence of effectively independent events; therefore the ordinates considered are separated by lengths, $l \geqslant 2.3\beta^*$. Thus a peak at a height

between y and $y+\delta y$ may be defined by three such consecutive events shown in figure 4*a* with the following restrictions: (*a*) the central event lies between y and $y+\delta y$; (*b*) event (1) has a value of less than y; (*c*) event (3) also has a value of less than y. Thus the probability that the central ordinate represents a peak between y and $y+\delta y$ is the multiplication of P_1, P_2 and P_3 where the P's refer to the shaded areas of the height distributions.

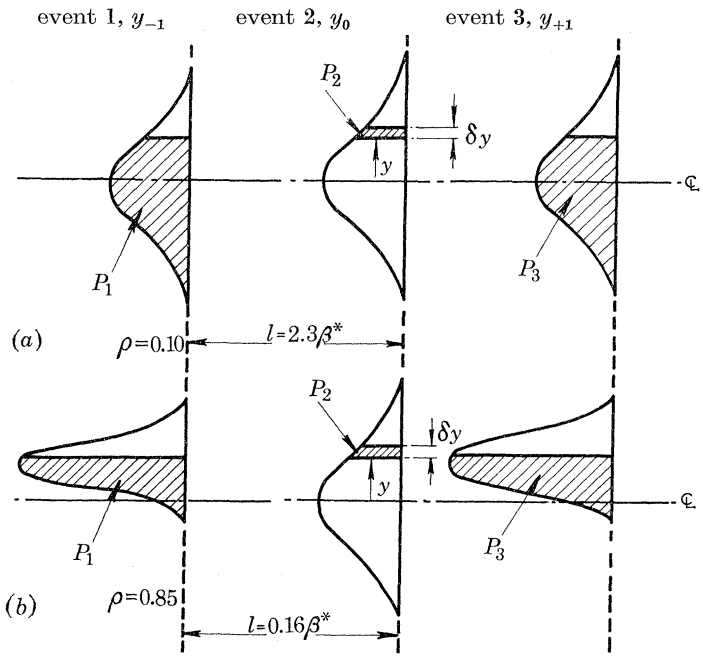

FIGURE 4. Model used in deducing distribution of peaks. (*a*) Sampling interval, $l = 2.3\beta^*$; correlation, $\rho = 0.10$; (*b*) sampling interval, $l = 0.16\beta^*$; correlation, $\rho = 0.86$.

Using this simple definition of a peak and equation (1) to define the probability density of a height distribution we can show (see the appendix) that the probability density of an ordinate being a peak at height y is given by

$$f^*(y) = [1/4\sqrt{(2\pi)}]\,[1+\mathrm{erf}\,(y/\sqrt{2})]^2 \exp(-\tfrac{1}{2}y^2),$$
$$= [1/\sqrt{(2\pi)}]\,\Phi^2(y) \exp(-\tfrac{1}{2}y^2), \qquad (5)$$

where here, and subsequently, f^* is used to indicate that properties of peaks are considered, f being retained for properties of the whole profile.

This argument can be extended to include the situation where the sampled ordinates are taken too close to be considered independent of each other. Such a situation is shown in figure 4*b*. In this case, also, the probability of an ordinate being a peak at height between y and $y+\delta y$ is given by $P_1P_2P_3$. However, the ordinates adjacent to the central y_0 are now not allowed to take all the values of the original height distribution; they take values that fit into a modified distribution

whose shape depends upon the sampling interval. The modified height distributions shown in figure 4*b* are the result of having ordinates y_{+1} and y_{-1} so close to the central ordinate y_0, taken as a reference, that they are dominated by it. Thus, for short sampling intervals, y_{+1} cannot differ greatly from y_0 because of the inability of the signal level to change rapidly; this, in turn, is a consequence of the limitation of the power spectrum at high frequencies (figure 3*b*). In general, the distribution of y_{+1} is influenced not only by y_0, but also, to a lesser degree, by y_{-1}. However, for the particular example of the model used here (an exponential correlation function) y_{+1} can be considered as influenced by y_0 only and not by y_{-1}; this is a property indicative of a first order Markov process (Bendat 1958, p. 215) and the distributions of y_{+1} and y_{-1} drawn in figure 4*b* are dependent only upon the sampling interval and the value of y_0. (See the appendix, where the derivation of the equations of this section is outlined and the basis of a theory for any form of correlation function is indicated.)

On the more general theory associated with figure 4*b*, the probability density of an ordinate being a peak at height y is given by

$$f^*(y,\rho) = \frac{1}{4\sqrt{(2\pi)}}\left[1+\operatorname{erf}\left(\frac{y}{\sqrt{2}}\sqrt{\frac{1-\rho}{1+\rho}}\right)\right]^2 \exp\left(-\tfrac{1}{2}y^2\right)$$

$$= \frac{1}{\sqrt{(2\pi)}}\Phi^2\left(y\sqrt{\frac{1-\rho}{1+\rho}}\right)\exp\left(-\tfrac{1}{2}y^2\right). \tag{6}$$

Figure 5 shows plots of this peak height distribution for a high and a low value of the correlation, ρ; the height distribution of the ordinates is also shown for comparison. The trends with varying values of ρ will be observed. As $\rho \to 0$ (large sampling interval), the shape of the peak height distribution becomes slightly skewed, its mean value approaches $+0.85$ and its standard deviation approaches a value of 0.70. Thus, when one uses larger sampling intervals the main, longer wavelength, structure of the profile is revealed (neglecting, here, the problem of aliasing (Bendat 1958)) and the peaks tend to lie above the centre line. As $\rho \to 1$ the shape of the peak height distribution, and its mean value and standard deviation, approaches those of the height distribution of the ordinates. Thus, when using short sampling intervals one is concerned with the shorter wavelength structure of the profile and therefore the peaks revealed by three point analysis follow closely the broad scale surface structure (cf. figure 1*b*, *c*).

The mean value of the peak height density curve $y^*(\rho)$ is found by taking the first moment of $f^*(y,\rho)$ in the normalized version of equation (6) and yields

$$\bar{y}^*(\rho) = \frac{1}{2N}\left(\frac{1-\rho}{\pi}\right)^{\frac{1}{2}}. \tag{7}$$

Similarly, the variance of the peak heights is the second central moment. Thus

$$[\sigma^*(\rho)]^2 = \left[1-\frac{(1+\rho)^{\frac{1}{2}}}{2N\pi\tan^2(N\pi)}-\frac{(1-\rho)}{4\pi N^2}\right], \tag{8}$$

where the normalizing factor N, giving the ratio of number of peaks to number of ordinates is

$$N = (1/\pi)\tan^{-1}\sqrt{\{(3-\rho)/(1+\rho)\}}. \tag{9}$$

It will be noted that equations (5) and (6), when divided by equation (9), are the probability densities of peak heights. Equation (9) shows that as the correlation, ρ, increases from zero to unity, N falls from $\frac{1}{3}$ to $\frac{1}{4}$. These limiting values have a simple explanation. As the sampling interval is increased, $\rho \to 0$ and $N \to \frac{1}{3}$; the three events are then effectively independent (figure 4*a*) and the chance that any

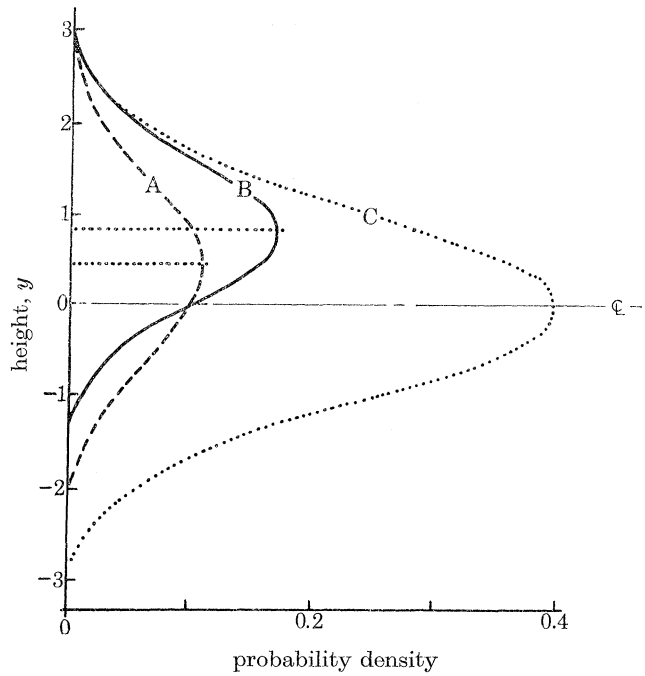

FIGURE 5. Probability densities of an ordinate being a peak at height y. The height, y, is normalized by the r.m.s. value (σ) of the ordinates. Results are shown for two different values of the sampling interval (l). (A) $l = 2.3\beta^*$; correlation, $\rho = 0.10$; average peak height = 0.82; (B) $l = 0.16\beta^*$; correlation, $\rho = 0{\cdot}86$; average peak height = 0·41; (C) Gaussian distribution of ordinates.

one of them (e.g. the centre one) is the highest becomes one-third. On the other hand, as the sampling interval is decreased $\rho \to 1$ and $N \to \frac{1}{4}$. The modified distributions of the two outer events are now centred upon the central ordinate (figure 4*b*); the areas P_1 and P_3 have values of $\frac{1}{2}$ and the probability that the central event is a peak is $\frac{1}{4}$.

To provide an adequate description of a surface in terms of a distribution of asperities it is also necessary to specify their radii of curvature. It is more convenient to discuss this in terms of a distribution of curvatures and we shall follow Greenwood & Williamson (1966) in deriving curvatures from the digital presentation of the profile. The assumption here is that one is justified in fitting a parabola to the profile by three point analysis; the problems involved in this assumption

and the limits within which this analysis is justified will be discussed later. Consider, first, the example of three independent events (figure 4*a*). Figure 6*a* shows one possible arrangement of the three events which will give a peak at height y with a curvature C given by

$$C = 2y_0 - y_{+1} - y_{-1}. \tag{10}$$

In this equation C is non-dimensional but the true value of curvature depends upon the sampling interval l; to obtain a true value, C must be multiplied by σ/l^2. We designate the sampling interval as l, rather than β of equation (2), to emphasize the important point that the properties of the profile are being deduced by finite difference methods from data obtained by a sampling technique.

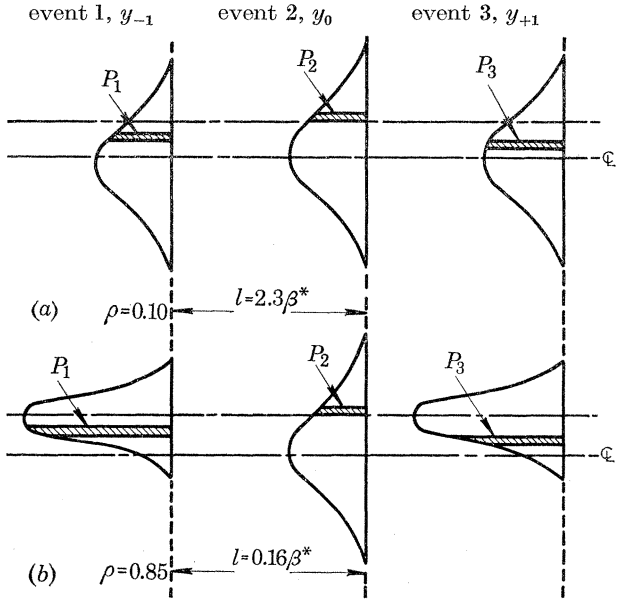

FIGURE 6. Model used in deducing the distribution of curvatures. (*a*) Sampling interval, $l = 2.3\beta^*$; correlation, $\rho = 0.10$; (*b*) sampling interval, $l = 0.16\beta^*$; correlation, $\rho = 0.86$.

This treatment assumes that the second derivative of the profile is an acceptable approximation to the curvature. Then the probability of the configuration shown in figure 6*a* is $P_1P_2P_3$, where P_1, P_2 and P_3, are given by the shaded areas shown. In order to obtain the *total* probability of a peak with curvature C at a height between y and $y+\mathrm{d}y$ many configurations, similar to that shown in figure 6*a*, must be taken into account. It is shown in the appendix that this total probability is given by a convolution integral. Thus the probability density function that any ordinate is a peak of curvature C at height y is

$$f^*(y, C) = \frac{\exp(-\frac{1}{2}y^2)}{2\pi\sqrt{2}} \exp[-(y-\tfrac{1}{2}C)^2] \operatorname{erf}(\tfrac{1}{2}C). \tag{11}$$

As before, the argument can be repeated for a shorter sampling interval (figure 6*b*). Thus

$$f^*(y, C, \rho) = \frac{\exp(-\frac{1}{2}y^2)}{2\pi[2(1-\rho^2)]^{\frac{1}{2}}} \exp\left[-\frac{[(1-\rho)\,y - \frac{1}{2}C]^2}{(1-\rho^2)}\right] \operatorname{erf}\left[\frac{C}{2\sqrt{(1-\rho^2)}}\right]. \tag{12}$$

The probability density function that any ordinate is a peak of curvature C (at any height) is obtained by integrating equation (12) to give

$$f^*(C, \rho) = \left[\frac{1}{4\pi(3-\rho)\,(1-\rho)}\right]^{\frac{1}{2}} \exp\left[\frac{-C^2}{4(3-\rho)\,(1-\rho)}\right] \operatorname{erf}\left[\frac{C}{2\sqrt{(1-\rho^2)}}\right]. \quad (13)$$

This distribution is skewed towards zero curvature; this is in general accord with the distribution of curvatures obtained by Greenwood & Williamson (1966) from digital analysis of a bead blasted surface. These authors suggest a Γ function as a suitable description of the distribution but equation (13) is nearer to a Rayleigh distribution than a Γ function and for large curvatures is very nearly Gaussian. A further comparison of these equations with results derived from surface profiles is given below.

The mean curvature $\bar{C}^*$ for all peaks is obtained by finding the first moment of $f^*(C, \rho)$ in equation (13). This yields

$$\bar{C}^* = (3-\rho)\,(1-\rho)^{\frac{1}{2}}/2N\sqrt{\pi}, \quad (14)$$

where the distribution has been normalized by N, the ratio of peaks to ordinates (equation (9)). The curvature (or, more strictly, the second differential) of the *profile* as a whole is given by

$$f(C, \rho) = \frac{1}{[4\pi(3-\rho)\,(1-\rho)]^{\frac{1}{2}}} \exp\left[\frac{-C^2}{4(3-\rho)\,(1-\rho)}\right], \quad (15)$$

which is equation (13) with the error function removed. This is a Gaussian distribution having a mean of zero and a standard deviation of $[2(3-\rho)\,(1-\rho)]^{\frac{1}{2}}$. A simple check on this value of the standard deviation (or, more strictly, the variance) is obtained by finding the square of the expected value of the curvature from equation (10). Thus $E[2y_0-(y_{-1}+y_{+1})]^2 = 6-8\rho+2\rho^2$.

The distribution of slopes is of importance because one widely used criterion for the onset of plastic flow (Blok 1952; Halliday 1955) uses the mean slope of the flanks of the asperities. The distribution of slopes (m) on the profile is easily obtained from the fact that the formula involves a simple linear relation of the Gaussian variates y_{-1} and y_{+1} and so is itself Gaussian with a mean of zero and a variance $|2\sigma^2(1-\rho^2)|/4l^2$. Hence

$$f(m, \rho) = \exp\left[\frac{-m^2l^2}{\sigma^2(1-\rho^2)}\right] \Big/ [4\pi(1-\rho^2)]^{\frac{1}{2}}, \quad (16)$$

from which one derives the average upward or downward slope (mean value of the modulus)

$$\overline{m} = \frac{\sigma}{l}\left[\frac{1-\rho^2}{\pi}\right]^{\frac{1}{2}}. \quad (17)$$

These formulae can also be obtained by the same type of procedure as that used in the derivation of equation (15).

4. Results of analysis of surface profiles

The validity of the theory given above has been checked by digital analysis of profile meter outputs. In order to present a coherent picture we give below a fairly complete analysis of the results obtained from one surface. The surface chosen (Aachen 64–13) is a typical ground surface used in an O.E.C.D. cooperative research programme (O.E.C.D, to be published). The surface profiles were taken at right angles to the direction of grinding.

The main experimental results were derived from surface profiles measured on a Talysurf 4 stylus surface roughness instrument in which a lightly loaded diamond stylus is traversed across the surface under examination. A normal stylus with a nominal tip dimension of 2.5μm was used. In these experiments the stylus was a diamond in the form of a truncated pyramid and the term 'tip dimension' refers to the linear dimension of the flat region at the tip in the direction of motion. A horizontal magnification of 500 was obtained by using the $500\times$ drive unit. Coupled to the Talysurf was a data logging system, specially devised for it, consisting of a Solartron A/D convertor and serializer together with a Data Dynamics 110 paper tape punch. From this equipment the amplified analogue signal of the surface profile was converted into a sequence of ordinates on the tape. Using this system ordinates were sampled at intervals of 1μm. The information on the paper tape was subsequently processed in an ICL 1905 computer.

Additional results were also obtained, with a special sharp stylus having a tip dimension of 0.25μm, the Talystep instrument being coupled to the same data logging equipment. The Talystep apparatus is capable of vertical magnifications of up to 10^6 and horizontal magnifications of up to 2×10^3. This latter facility made it possible to sample the profile at intervals of 0.25μm. Another feature of this instrument is the normal load on the stylus which was reduced to only 10^{-3} g (one-hundredth of the load on the Talysurf stylus) and this made the use of a very sharp stylus possible. Figure 7 shows an electron micrograph of this sharp stylus.†

A typical set of Talysurf results, from a single profile, consisted of some 10000 ordinates with a sampling interval of 1μm; of these, some 7000 were available for use after filtering to remove the d.c. level, the general slope and wavelengths comparable with the length of the profile. The results presented below are based upon five profiles and statistical analysis shows that the normalized standard errors are *ca.* 2 % for mean values (e.g. equations (7) to (9)) and *ca.* 5 % for points on the probability distributions (e.g. equations (4), (5) and (13)).

By suitable selection of data it was possible to present results for sampling intervals between 0.25 and 15μm. It was thus found (figure 8) that the model used in this paper was a good representation of the data obtained from the surface profiles; the distribution of ordinates was very close to Gaussian, with an r.m.s. value (σ) of 0.5μm, and the autocorrelation function was close to exponential

† We are indebted to Mr J. Jungles (Research Department Rank Precision Industries Ltd) for his skill in making and measuring this stylus.

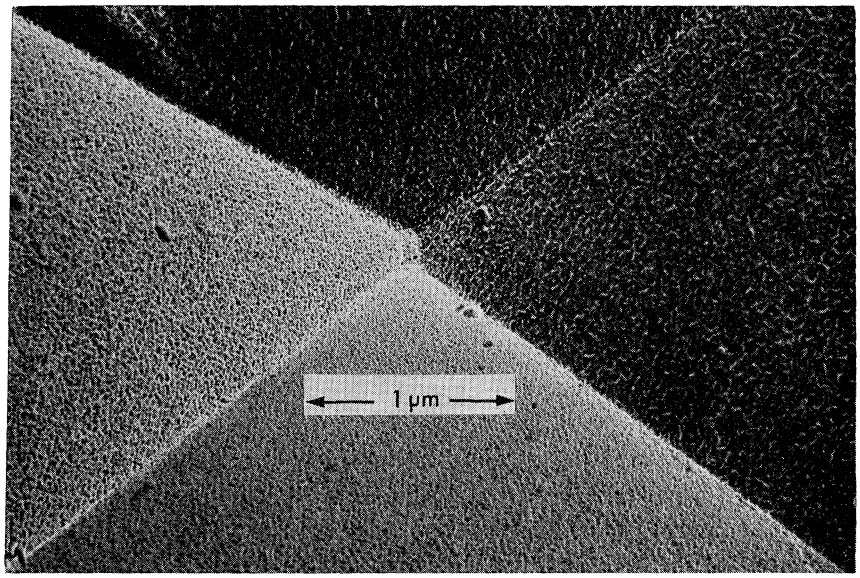

FIGURE 7. Electron micrograph of sharp stylus used in the experiments.

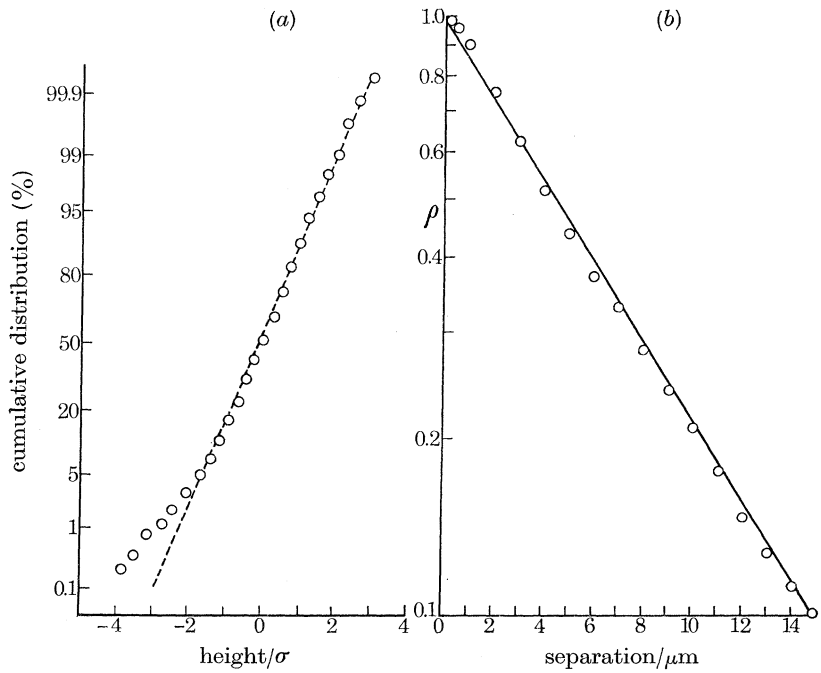

FIGURE 8. Characteristics of profiles of Aachen 64–13 represented as a random signal. (*a*) Cumulative distribution of heights (normal probability paper); (*b*) Correlation as a function of sampling interval (logarithmic-linear plot).

with a correlation distance (β^*) of 6.5μm. In subsequent discussion we shall use theoretical values of the correlation (ρ) for the selected values of the sampling interval as shown in table 1. It will be observed that any divergences between these values and those obtained from the profiles are, for the most part, within the limits of experimental error. In the results, presented below in graphical form, sampling intervals (l) of 15, 3 and 1μm (corresponding to correlations (ρ) of 0.10, 0.63 and 0.86 respectively) have been selected to display certain important features.

TABLE 1. RELATION BETWEEN SAMPLING INTERVAL (l) AND CORRELATION (ρ) BETWEEN SUCCESSIVE SAMPLES FOR AACHEN 64–13

sampling interval, l/μm	15	6.0	3.0	2.0	1.0	0.5	0.25
correlation, ρ	0.10	0.40	0.63	0.74	0.86	0.92	0.96

Figure 9 shows a comparison of theory and experiments for the probability that an ordinate is a peak at a height y (equation (6)). It will be observed that for $l = 15\mu$m and $l = 3\mu$m the agreement between theory and experiment is remarkably good. However, for $l = 1\mu$m (figure 9*c*) there is a marked divergence, the number of peaks detected falling significantly below the theoretical values. The results for all values of the sampling interval are shown in figure 10 in which the mean value and the standard deviation of the distribution of peaks (equations (7) and (8)) are plotted against the value of the correlation between successive samples. The most significant divergence between theory and experiment is the fact that, for the shorter sampling intervals, the mean values lie above the theoretical predictions (see also figure 9*c*).

Figure 11 presents theory and experiment for the probability than an ordinate is a peak of given curvature. As before, for $l = 15\mu$m and $l = 3\mu$m the agreement is excellent but there are significant differences for the shorter sampling interval of $l = 1\mu$m. It will be observed from the magnitudes of the curvatures shown in figures 11*a* to *c*, that, as the sampling interval is decreased, one is concerned with asperities of smaller and smaller radius. This is made quite clear in figure 12 which compares theoretical values of the mean curvature of the peaks with the values found from the profiles for differing sampling intervals. Once more, the only significant divergence between theory and experiment occurs at the shortest sampling interval ($l = 1\mu$m).

The results obtained at shorter sampling intervals (see, in particular, figures 9*c*, 11*c*) suggest that the measurements of the surface profiles are affected by the finite size of the stylus. In figure 11*c* a value of the nominal stylus curvature has been indicated; this is taken as the reciprocal of the nominal tip dimension of the stylus. The character of the divergence between theory and experiment shown in figure 11*c* is certainly consistent with the assumption that it arises from the finite size of the stylus. The total number of peaks detected is less than that forecast by the theory and the distribution has apparently been distorted towards smaller values of the curvature.

It is, of course, equally possible that the surface used in this work does not conform, at these shorter wavelengths, to the model assumed in this paper. Figures 9*c* and 11*c* would then imply that the structure of shorter wavelengths, although present in the model, does not exist upon the surface. In an attempt to

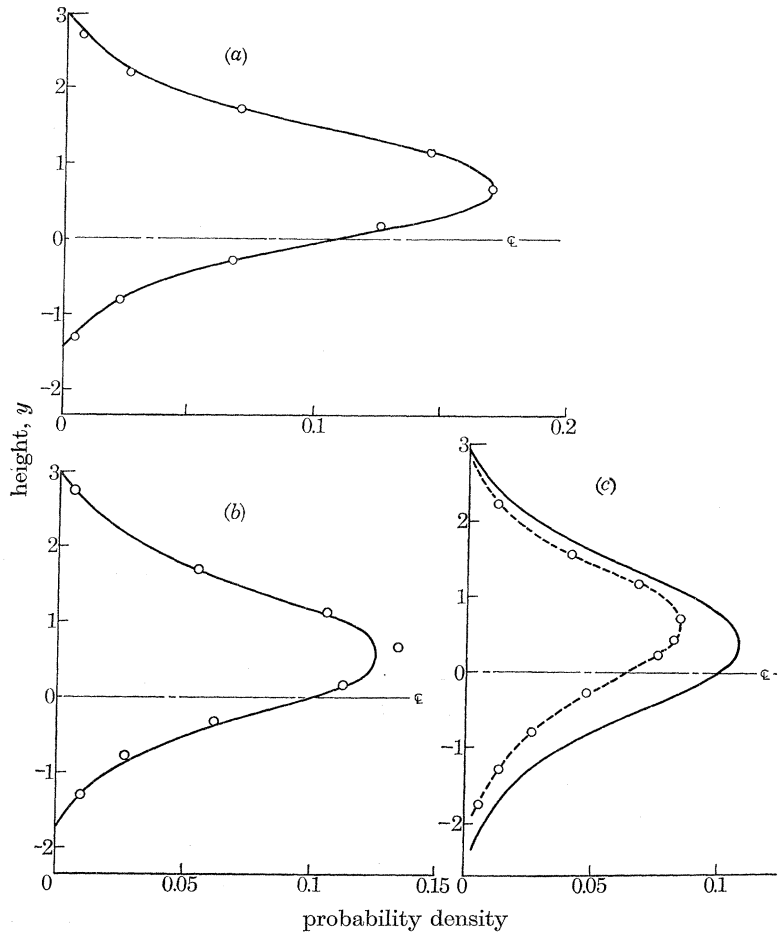

FIGURE 9. Probability densities of an ordinate being a peak at a height y. The full lines give the theory (equation (6)). The experimental points are derived from digital analysis of profiles of Aachen 64–13. Results are shown for the values of the sampling interval (l) corresponding to differing values of the correlation (ρ) between successive samples. (*a*) $l = 15\ \mu\text{m}$, $\rho = 0.10$; (*b*) $l = 3.0\,\mu\text{m}$, $\rho = 0.63$; (*c*) $l = 1.0\,\mu\text{m}$, $\rho = 0.86$.

resolve this question experiments were performed with a stylus with a smaller tip dimension. The results are shown in figure 13 where the ratio (N) of peaks to ordinates is plotted against the correlation (ρ) between successive samples. It will be recalled that the theory (equation 9) forecasts that this ratio varies between 0.33 ($\rho = 0$) and 0.25 ($\rho = 1$). Figure 13 shows, once more a divergence between theory and experiment for sampling intervals of less than $2\,\mu\text{m}$; in this region the

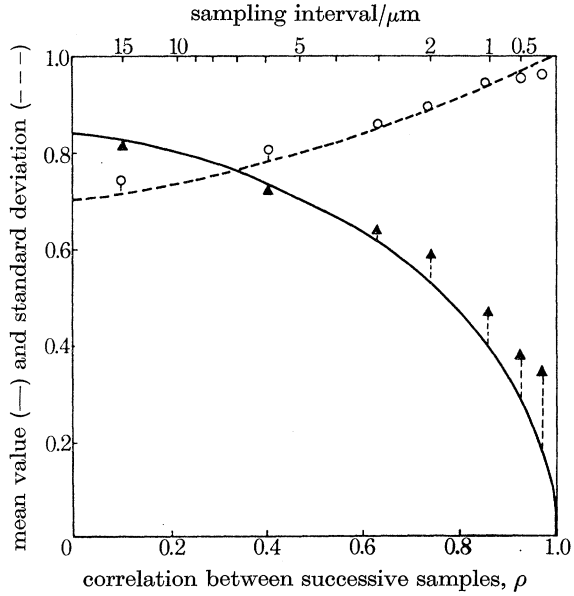

FIGURE 10. Characteristics of the distribution of peaks. The full line gives the mean value (equation (7)) and the broken line the standard deviation (equation (8)); they are normalized by the standard deviation of the ordinates (σ). The experimental points are derived from digital analysis of profiles from Aachen 64–13.

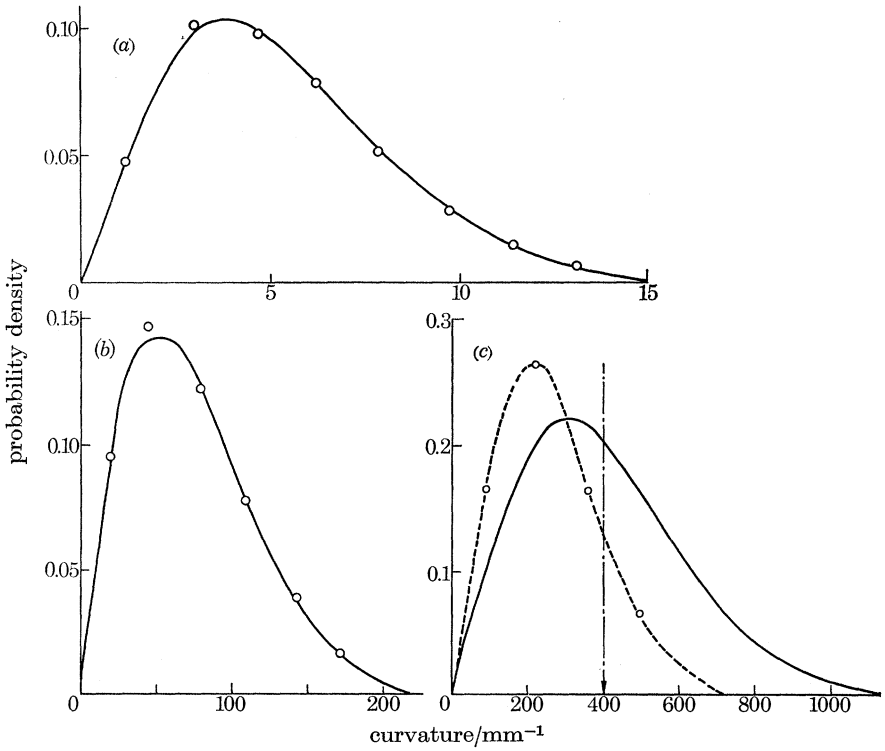

FIGURE 11. Probability densities of an ordinate being a peak of a given curvature. The full lines give the theory (equation (13)). The experimental points are from digital analysis of profiles from Aachen 64–13 ($\sigma = 0.5$ μm, $\beta^* = 6.5$ μm). Results are shown for three values of the sampling interval (l) corresponding to differing values of the correlation (ρ) between successive samples. (*a*) $l = 15$ μm, $\rho = 0.10$; (*b*) $l = 3.0$ μm, $\rho = 0.63$; (*c*) $l = 1.0$ μm, $\rho = 0.86$; the arrow indicates the nominal stylus curvature, 400 μm^{-1}.

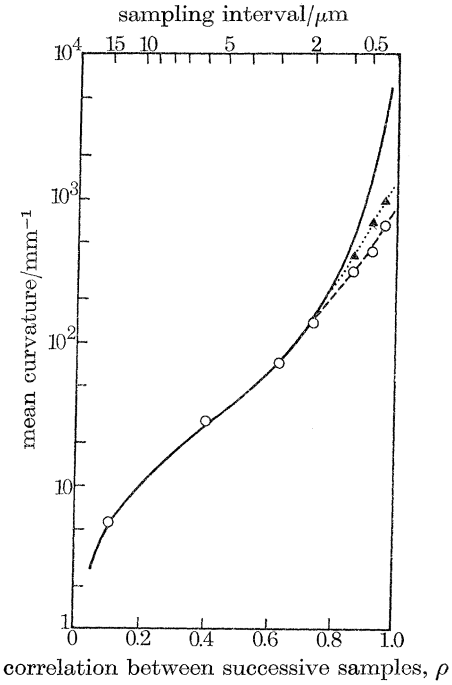

FIGURE 12. Mean curvature of the peaks as a function of the correlation (ρ) between successive samples. The full line gives the theory (equation (14)). The experimental points are derived from digital analysis of profiles from Aachen 64–13 ($\sigma = 0.5$ μm, $\beta^* = 6.5$ μm). ○, Normal stylus, nominal tip dimension 2.5 μm; ▲, special stylus, nominal tip dimension 0.25 μm.

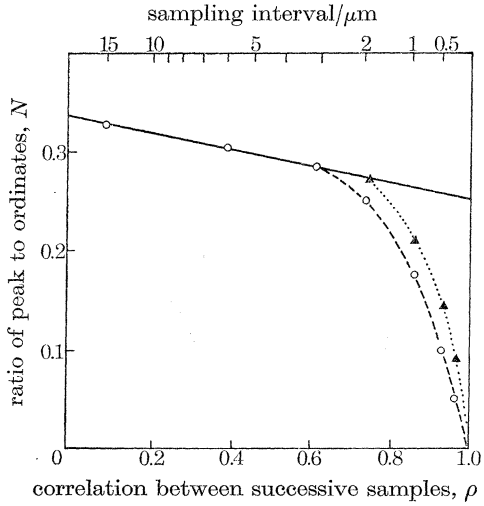

FIGURE 13. Ratio (N) of peaks to ordinates as a function of the correlation (ρ) between successive samples. The full line gives the theory (equation (9)). The experimental points are derived from digital analysis of profiles from Aachen 64–13. ○, Normal stylus, nominal tip dimension 2.5 μm; ▲, special stylus, nominal tip dimension 0.25 μm.

numbers of peaks detected falls well below the theoretical values. Similar plots showing a decline in the number of peaks detected at short sampling intervals have been presented by Sharman (1967–8), but no explanation of the cause has been advanced. Figure 13 also shows that when using a stylus with a smaller tip dimension the decline is delayed to smaller values of the sampling intervals. Clearly, therefore, stylus resolution is a significant factor affecting the behaviour in this region.

5. Discussion

The starting-point of the present work is the concept, well accepted in the theory of random processes but unexplored in the field of surfaces and their contact, that a random profile can be completely defined by two parameters: the height distribution and the autocorrelation function. For the particular example of the model used in this paper, these two parameters become simply two lengths, the r.m.s. value of the height distribution (σ) and the correlation distance (β^*). The statistical distributions of all significant geometric characteristics of the surface profile, for example, slopes, peaks and curvatures, can then be predicted from these two independent parameters. It is significant that the present work, which has arisen from markedly theoretical considerations, is set against a particular background of comment from those concerned with practical aspects of the measurement of surface finish and its use in engineering. This includes many comments that a measurement of surface finish, widely adopted hitherto, based solely upon the height distribution (r.m.s. or c.l.a. roughness) is not an adequate description of the functional significance of surface roughness (Reason 1967–8). Therefore, the ideas outlined here seem very relevant in any attempt to provide a more complete specification of surface finish. In practice, for reasons connected with ease of measurement, it may be desirable to measure a derived parameter such as the average upward or downward slope. However, although many surfaces do not conform exactly to the simple model used in this paper, the principle of specifying surface finish in terms of two independent parameters seems capable of wider application.

The importance of β^* in the specification of surface finish has been stressed, but it has an additional significance in its own right. Consider the measurement of the c.l.a. or r.m.s. roughness value of a random surface; it is desired to know the confidence limits of such a measurement and this is easily expressed if one knows the standard deviation of a large number of such measurements made upon the same surface. Alternatively this can be deduced if β^* is known. It can be shown that the standard deviation of such measurement of the c.l.a. roughness of a random surface, when normalized as a ratio of the mean value, is $\approx 1/\sqrt{(2M)}$. In this expression M is the ratio of the assessment length (L), used in the measurement of c.l.a., to the distance between points on the surface ($2.3\beta^*$) which just provides effectively independent events. (This ratio M is analogous to the band width $\times$ duration product used in communications theory to estimate the reliability

of data.) For example, on the 0.03 in cut-off range, the Talysurf 4 instrument has an assessment length, of 0.15 in (3.81 mm). The value of $2.3\beta^*$ for Aachen 64–13 is 15 μm. Thus the normalized standard deviation of c.l.a. readings on this surface should be 4.5 %. A measured value of the standard deviation for Aachen 64–13, based upon a large number of readings, was 4.3 %.

The comparison of theory and experiment outlined in this paper shows that the model which has been adopted can provide a satisfactory description of the geometric features of profiles from a surface typical of engineering practice; the statistical distribution of surface characteristics are accurately forecast over a wide range. This range covers an order of magnitude in the linear dimensions of the asperities and more than two orders of magnitude in their curvatures. Divergences appear only at shorter wavelengths and these arise, at least in part, from the resolution of the stylus. This is clear from a more detailed consideration of the results obtained with the sharp stylus.

First, consider the results of figure 13. The model adopted in the present paper requires that at short sampling intervals ($N \rightarrow \frac{1}{3}$) the number of peaks detected on a given length of profile should be inversely proportional to the sampling interval. However, with the normal stylus, reducing the sampling interval from 1 to 0.5 μm increases the number of peaks by only 16 % and a further reduction to 0.25 μm causes no detectable increase in the number of peaks. These results suggest that either the fine scale structure does not exist upon the surface, or it is present and is not detected by the stylus. The results obtained with the sharp stylus show, clearly, that stylus resolution is a significant factor. At a sampling interval of 1 μm the replacement of the normal stylus by the sharp stylus causes an increase of 20 % in the number of peaks detected. In addition, when using the sharp stylus, a reduction in the sampling interval to 0.5 μm and to 0.25 μm causes increases in the number of peaks by 37 and 75 % respectively. Figure 12 shows that use of the sharp stylus also results in an increase in the mean curvature of the peaks; in other words the sharp stylus reveals more detail and finer detail. The suggestion that the finite size of the stylus has an influence upon the resolution of the finest detail is supported by other experiments (R. E. Reason, unpublished) in which a sharp stylus has revealed considerable detail upon smooth surfaces which the normal stylus does not reveal.

A detailed discussion of the effect of the size and shape of the stylus upon resolution is outside the scope of the present paper. However, some general considerations are worthy of comment. First, the effect of the finite size of the flat tip of the stylus is not likely to produce a sharp cut-off in resolution but should exercise an influence over a range of wavelengths somewhat as has been indicated in the discussion of figure 13. Nevertheless, assuming that the profile corresponds to the model of this paper, the change in the stylus tip dimension from 2.5 to 0.25 μm produces a smaller change in the resolution than might be expected; perhaps the fine scale structure is present but its magnitude is less than is required by the theoretical model. However, it should be noted that the tip dimension may

not be the only factor which determines the resolution of the stylus. Equation (17) shows that as the sampling interval is reduced, and structures of shorter wavelength are involved, the local slope of the surfaces becomes steeper (cf. figures 1 *a*, *c*). Thus, although existing instruments reveal the details known to be of functional significance, to resolve the finest detail on random surfaces it may be necessary to consider the complete shape of the stylus, the sides as well as the tip.

The present work has also focused attention upon a problem associated with the representation of random surfaces by models. It has been explained that the random signal representation used in this paper is a complete description of the profile, in a statistical rather than a deterministic sense. However, models appropriate to the problems of surface contact are expressed as a distribution of peaks and such models have been derived from digital analysis of surface profiles; consideration of the results obtained in this paper shows, immediately, the difficulties associated with any definition of a peak. The simpler forms of digital analysis (for example, the three-point system which has been widely used) cause, inevitably a loss of information compared with that provided in the original random signal representation, or that which is present in the surface profile. Thus, a close sampling technique collects the maximum amount of information about the profile; but three-point analysis of these data restricts the information to structures of this same small order of size. Only by the use of more sophisticated techniques, such as digital filtering, can the total information in the sample be utilized. An alternative is to use differing selections of the same information, by rejecting some data, but still to use three-point analysis; e.g. to present, as we have done, information from three-point analysis for a wide range of sampling intervals.

However, these techniques present problems of rigour. When using the longer sampling intervals one is presenting statistical information about the longer wavelength structure of the profile obtained by drawing a smooth curve through widely separated points upon it. Therefore the results presented above, in particular, the values of curvature obtained from long sampling intervals, should be interpreted with these reservations in mind.

Because the description of the profile as a random signal is statistical in its form an overemphasis upon the limitations of three-point sampling is not, perhaps, the most significant problem. The more important, and difficult, question is deciding the range of sampling intervals from which it is acceptable to use information in devising models for theories of surface contact. The significance of the shorter wavelength structures (revealed by the use of shorter sampling intervals) in the contact and rubbing of surfaces may be questioned; in any event, a lower limit to the acceptable range of sampling intervals is set by stylus resolution. At the other extreme it is clear that the use of very long sampling intervals will give results which have little physical significance.

A more relevant question is the spacing of events that will define the dominant or main structure of the profile; such a spacing might be considered as the upper

limit of sampling intervals to be used in devising models for the study of surface contact. To outline this structure implies that the profile be considered as a series of events which have just reached the spacing where they can be regarded as effectively independent. As discussed earlier, this suggests a spacing of events of $2.3\beta^*$ and also implies that the main structure of the profile has a wavelength of *ca.* $10\beta^*$. Similarly the profile can be considered in terms of the power spectral density function (figure 3*b*). This suggests that the most significant information is contained in a low-pass band of frequencies having a cut-off frequency of $(2\pi\beta^*)^{-1}$. This approach suggests a wavelength of *ca.* $6\beta^*$ for the main structure of the profile. Finally, we can arrive at an estimate of the broad scale structure from a consideration of the number of times the profile crosses the centre line; the mean distance between such crossings might be regarded as half the wavelength of this main structure. To obtain a meaningful result with this approach one

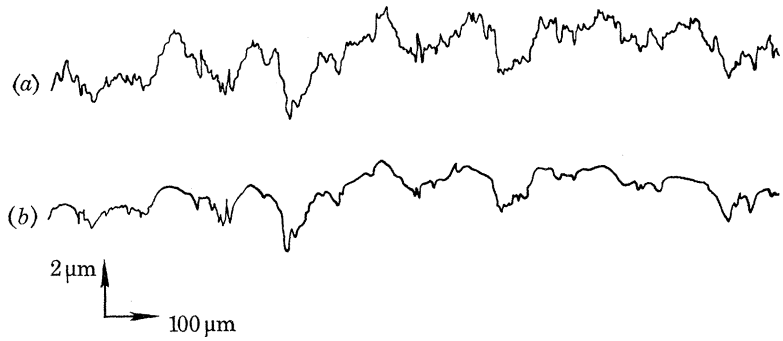

FIGURE 14. Talysurf profiles of cylindrical specimens used in lubricated friction experiments. (*a*) Original surface profile; (*b*) profile of the same line after one traversal of the load. Specimens 0.5 % C steel; 300 d.p.n.; 0.635 mm diameter; load 25 N.

must remove the high frequency components. If we replace the power spectral density distribution of figure 3*b* by one with a sharp cut-off at an angular frequency of $1/\beta^*$ (low-pass filtered white noise) it is possible to use a result derived by Rice (see Bendat 1958, p. 128). For this situation the mean distance between centre line crossings is $5.5\beta^*$ and a representative value of the wavelength of the main structure is $11\beta^*$. It can be argued that, when considering the results of three-point analysis, a somewhat shorter interval is appropriate because in surface contact one is concerned with the tips of the higher asperities. As a general guide to the main structure of relevance in contact problems we shall therefore use the results derived from a sampling interval of $2.3\beta^*$.

The significance of this distinction between the differing scales of size involved in surface structure and the importance of the main, broad scale, structure is illustrated in figure 14. The upper record shows the original profile of a cylindrical specimen used in a low speed boundary lubricated friction experiment using a crossed-cylinders friction machine (Archard 1958). The lower profile is of the *same*

track upon the surface after just one traversal of the load. It will be seen that the smaller scale structure has been lost by plastic flow during the first traversal but the longer wavelength structure remains upon the surface and is the dominant factor in determining the contact conditions in subsequent traversals of the surface. A more complete account of these experiments will be published elsewhere (D. J. Whitehouse & J. F. Archard, to be published).

An important aim of recent studies of the topography of surfaces has been to provide an estimate of the chances that a given surface will be subjected to plastic flow during contact. Blok (1952) and Halliday (1955) considered the shape of asperities which could be pressed flat without recourse to plastic deformation. It was shown that this criterion could be expressed in the form

$$\overline{m} \leqslant KH/E', \tag{18}$$

Where H is the hardness and $E' = E/(1-\nu^2)$, E being the Young modulus and ν the Poisson ratio; $\overline{m}$, as above, is the average upward or downward slope and K is a numerical factor, in the range 0.8 to 1.7, which depends only upon the assumed shape of the asperity. Greenwood & Williamson (1966) assessed the probability of plastic deformation using their model in which asperities, each of radius R, are disposed in a Gaussian distribution of heights of standard deviation σ^*. In this model there is always a finite chance of plastic flow; however, it was shown that it depended very little upon the load but was critically dependent upon a plasticity index, ψ, given by

$$\psi = \left(\frac{E'}{H}\right)\left(\frac{\sigma^*}{R}\right)^{\frac{1}{2}}. \tag{19}$$

The plasticity criteria of equations (18) and (19) are similar in form. The Blok–Halliday criterion (equation (18)) is, however, unduly severe because it assumes complete depression of the asperities. The plasticity index of Greenwood & Williamson (1966) takes account of the fact that only the tips of asperities are normally involved in contact. The present work emphasizes the simplifications which have been made in these plasticity criteria because they take no account of the existence, upon surfaces, of superposed asperities of differing scales of size. The plasticity calculations of equation (18) and (19) assume that the deformation of each of the asperities is independent. Therefore the plasticity index of equation (19) has a significance only if it is applied to the main, long wavelength structure; it should then indicate the probability of plastic flow over regions associated with this scale of size. If values of R corresponding to smaller scale structure are used the arguments involved in the derivation of equation (19) become invalid because the deformation of adjacent asperities interact, as in the model of figures 1*b*, *c*.

To summarize the results derived from the model of the present paper, table 2 shows the way in which the significant characteristics of a surface profile depend upon the two independent parameters σ and β^*. To emphasize the importance of the scale of size used in the analysis each characteristic (except ψ, for reasons outlined above) is shown for two scales. The main structure of the profile is derived

assuming a sampling interval, l, of $2.3\beta^*$ and the fine scale structure assumes asperity dimensions one order of magnitude smaller ($l = 0.23\beta^*$). For Aachen 64–13, used in our digital analysis described above, this is just within the limits of resolution of the Talysurf instrument using the normal stylus. In deriving a value of the plasticity index, ψ, a value of the mean curvature of the peaks, derived from equation (14), has been used. The value of ψ derived in this way somewhat underestimates the probability of plastic flow because, as equation (11) shows, the curvature of the peaks increases with increasing height. Thus the highest peaks, which are those involved in contact, have a smaller radius than the total population.

TABLE 2. CHARACTERISTICS OF A RANDOM PROFILE IN TERMS OF σ AND β^*

characteristic of the profile	main structure ($l = 2.3\beta^*$)	fine structure ($l = 0.23\beta^*$)
mean of peak distribution	$+0.82\sigma$	$+0.47\sigma$
standard deviation of peak distribution, σ^*	0.71σ	0.9σ
ratio of peaks to ordinates, N	0.33	0.26
average upward or downward slope, $\overline{m}$	$0.24\sigma/\beta^*$	$1.66\sigma/\beta^*$
mean peak curvature, $\bar{c}^*$	$0.45\sigma/\beta^{*2}$	$20\sigma/\beta^{*2}$
plasticity index, ψ	$0.3\left(\frac{E'}{H}\right)\left(\frac{\sigma}{\beta^*}\right)$	—

A complete theory of the contact of surfaces on the basis the model of this paper cannot be presented here. However, one interesting feature of the model and its comparison with that of Greenwood & Williamson (1966) is worthy of comment. The Greenwood & Williamson model is specified by three parameters; σ^*, the standard deviation of the peak height distribution; R, the radius of curvature of the asperities; and η, the density of asperities per unit area. In the model of this paper the required parameters are completely defined by σ, the standard deviation of the height distribution and β^*, the correlation distance. The theory of contact based on our model involves a statistical distribution of both peak heights and peak curvatures. Comparing the two models: σ^* is proportional to σ, R is proportional to β^{*2}/σ, and η is proportional to $1/\beta^{*2}$. Therefore for all random surfaces, when the Greenwood & Williamson model is used the parameters should be related by the equation

$$\sigma^* R\eta = \text{constant}.$$

There is some evidence (J. A. Greenwood, private communication) from the analysis of bead-blasted surfaces that this relation is, indeed, obeyed.

Finally, a brief comment upon the generation, by mechanical processes, of surfaces having a random structure. Such surfaces are generated by multiple

contacts between particles and the surfaces. Thus in grinding and sand blasting the unit event is the interaction of a grit with the surface resulting in the displacement or removal of material. If one postulates a random element in these events it seems likely that the surfaces thus generated would have a random structure in which the standard deviation of the height distribution bears a simple relationship to the depth of the unit event and the correlation distance similarly bears a simple relation to the width of the unit event.

We are indebted to Dr J. A. Greenwood and Dr J. B. P. Williamson for discussions of their published and unpublished work. Our thanks are due to our colleague Dr P. H. Phillipson for helpful discussions of the theory of random processes. We also thank the Directors of Rank Precision Industries Ltd for financial support and for permission to publish.

Appendix. General form of the theory

In the derivation of the relevant expressions use will be made of the multi-dimensional normal distribution (m.n.d.) which is concerned with the joint probability density function of a number of Gaussian variates; in our problem these variates will be simply profile ordinates. The m.n.d. will be Gaussian because any linear combination of Gaussian variates is itself Gaussian.

Definition (see Bendat 1958). If the ordinates $y_1, y_2, \ldots, y_N$ have Gaussian height distributions, having a mean of zero and a variance unity then their combined joint density function is given by

$$f(y_1, y_2, \ldots, y_N) = \frac{1}{(2\pi)^{\frac{1}{2}N}|M|^{\frac{1}{2}}} \exp\left[\frac{-\sum_{i,j=1}^{N} M_{ij} y_i y_j}{2|M|}\right]. \tag{A 1}$$

Where $|M|$ is the determinant of M; M is given by the square matrix

$$M = \begin{pmatrix} d_{11} & d_{12} & \ldots & d_{1N} \\ \ldots & d_{ij} & \ldots & \ldots \\ d_{N1} & d_{N2} & \ldots & d_{NN} \end{pmatrix},$$

d_{ij} being the second moment of the variables $y_i y_j$. M_{ij} is the co-factor of d_{ij} in M.

Take, for example, the joint probability density of two ordinates, say y_{-1} and y_0 correlated by ρ. Then

$$f(y_0, y_{-1}) = \frac{1}{\sqrt{(2\pi)}} \exp\left(-\tfrac{1}{2}y_0^2\right) \frac{1}{\sqrt{\{2\pi(1-\rho^2)\}}} \exp\left[-\frac{(y_{-1}-\rho y_0)^2}{2(1-\rho^2)}\right].$$

Similarly, for three ordinates, y_{-1}, y_0, y_{+1}, having a correlation of ρ_1 between adjacent ordinates and ρ_2 between extreme ordinates the joint probability density is given by

$$f(y_{-1}, y_0, y_{+1}) = f(y_0)\, f(y_{-1}/y_0)\, f[y_{+1}/(y_0, y_{-1})], \tag{A 2}$$

where

$$f(y_0) = \frac{1}{\sqrt{(2\pi)}} \exp\left(-\tfrac{1}{2}y_0^2\right)$$

$$f(y_{-1}/y_0) = \frac{1}{\sqrt{\{2\pi(1-\rho_1^2)\}}} \exp\left[-\frac{(y_{-1}-\rho_1 y_0)^2}{2(1-\rho_1^2)}\right]$$

$$f[y_{+1}/(y_0, y_{-1})] = \frac{(1-\rho_1^2)}{\sqrt{\{2\pi(1-\rho_2)\,(1+\rho_2-2\rho_1^2)\}}} \times \exp\left[-\frac{\{y_1(1-\rho_1^2)-y_0\rho_1(1-\rho_2)-y_{-1}(\rho_1^2-\rho_2)\}^2}{2(1-\rho_1^2)\,(1-\rho_2)\,(1+\rho_2-2\rho_1^2)}\right].$$

For the model under consideration the autocorrelation is exponential. Hence $\rho_1^2 = \rho_2$. Then $f[y_{+1}/y_0, y_{-1}]$ reduces to $f(y_{+1}/y_0)$ which is a criterion for a first-order Markov process. Hence for an exponential correlation function equation (A 2) becomes

$$f(y_{-1}, y_0, y_{+1}) = \frac{1}{\sqrt{(2\pi)}} \exp\left(-\tfrac{1}{2}y_0^2\right) \frac{1}{\sqrt{\{2\pi(1-\rho^2)\}}} \times \exp\left[-\frac{(y_{-1}-\rho y_0)}{2(1-\rho^2)}\right] \frac{1}{\sqrt{\{2\pi(1-\rho^2)\}}} \exp\left[-\frac{(y_{+1}-\rho y_0)^2}{2(1-\rho^2)}\right]. \quad \text{(A 3)}$$

It will be noted that the simplified form of the theory used in this paper arises because $\rho_2 = \rho_1^2$ and for a *particular* sampling interval this result does not depend upon the fact that the autocorrelation function is exponential. However, the exponential correlation function is an essential requirement for the general applicability of the simplified theory at *all* sampling intervals.

Peak height distribution

The probability of an ordinate being a peak at a height between y and $y+\mathrm{d}y$ is written in terms of a restriction of the joint probability density range in the following form (see theory):

$$\text{prob}\,[y_{-1} < y, y < y_0 < y+\mathrm{d}y, y_{+1} < y] = \int_{-\infty}^{y}\int_{y}^{y+\mathrm{d}y}\int_{-\infty}^{y} f(y_{-1}, y_0, y_{+1})\,\mathrm{d}y_{-1}\,\mathrm{d}y_0\,\mathrm{d}y_{+1}.$$

Which is the general equation of a peak using the three ordinate model; when the correlation is exponential the probability density reduces to the following form

$$f^*(y, \rho) = \frac{\exp\left(-\tfrac{1}{2}y^2\right)}{\sqrt{(2\pi)}} \frac{1}{2\pi(1-\rho^2)} \times \int_{-\infty}^{y} \exp\left[\frac{-(y_{-1}-\rho y_0)^2}{2(1-\rho^2)}\right]\mathrm{d}y_{-1} \int_{-\infty}^{y} \exp\left[\frac{-(y_{+1}-\rho y_0)^2}{2(1-\rho^2)}\right]\mathrm{d}y_{+1}.$$

This reduces to equation (6) of the theory and, for $\rho = 0$, gives equation (5).

The ratio (N) of the number of peaks to ordinates, for any value of the correlation ρ, is obtained simply by integrating equation (6).

Peak curvature distribution

For a given curvature, C, as defined in the text, the ordinates are related by an expression

$$C = 2y_0 - y_{-1} - y_{+1}.$$

Hence the total probability of an ordinate being a peak between y and $y+\mathrm{d}y$ and describing a curvature, C covering all possible configurations of y_{-1} and y_{+1}, is given by

$$f^*(y, C, \rho) = \int_y^{y+\mathrm{d}y} \int_{y-C}^{y} f(y_0)\, f(y_{-1}/y_0)\, f\left[\frac{2y_0 - y_{-1} - C}{y_0, y_{-1}}\right] \mathrm{d}y_{-1}\mathrm{d}y_0. \qquad \text{(A 4)}$$

Hence the probability density of an ordinate being a peak at height y with curvature C is

$$f^*(y, C, \rho) = f(y) \int_{y-C}^{y} f[y_{-1}/y]\, f\left[\frac{2y - y_{-1} - C}{y, y_{-1}}\right] \mathrm{d}y_{-1}. \qquad \text{(A 5)}$$

This is a convolution integral which enables simple graphical equivalents to be constructed; equation (A 5) corresponds to the general curvature formula with a general autocorrelation function. However, for the simple exponential correlation it becomes

$$f^*(y, C, \rho) = \frac{1}{\sqrt{(2\pi)}} \exp\left(-\frac{y^2}{2}\right) \int_{y-C}^{y} \frac{1}{2\pi(1-\rho^2)} \times \exp\left[-\frac{(y_{-1} - \rho y)^2}{(12-\rho^2)}\right] \exp\left[-\frac{(2y - y_{-1} - C - \rho y)^2}{2(1-\rho^2)}\right] \mathrm{d}y_{-1}.$$

This reduces to equation (12) of the theory and when $\rho = 0$ one obtains equation (11).

References

Archard, J. F. 1957 *Proc. Roy. Soc. Lond.* A **243**, 190.

Archard, J. F. 1958 *Wear* **2**, 21.

Beckmann, P. & Spizzichino, A. 1963 *The Scattering of electromagnetic radiation from rough surfaces*. London: Pergamon.

Bendat, J. S. 1958 *Principles and applications of random noise theory*. New York: Wiley.

Blok, H. 1952 *Proc. Roy. Soc. Lond.* A **212**, 480.

Bowden, F. P. & Tabor, D. 1954 *Friction and lubrication of solids*. Oxford University Press.

Greenwood, J. A. & Williamson, J. B. P. 1966 *Proc. Roy. Soc. Lond.* A **295**, 300.

Halliday, J. S. 1955 *Proc. Instn mech. Engrs* **109**, 777.

O. E. C. D. Sub-group Typology and Topology (to be published).

Peklenik, J. 1967–8. *Proc. Instn mech. Engrs* **182**, part 3K (conference on the properties and metrology of surfaces), p. 108.

Reason, R. E. 1967–8 *Proc. Instn mech. Engrs* **182**, part 3K (conference on the properties and metrology of surfaces), p. 300.

Sharman, H. B. 1967–8 *Proc. Instn mech. Engrs* **182**, part 3K (conference on the properties and metrology of surfaces), p. 416.

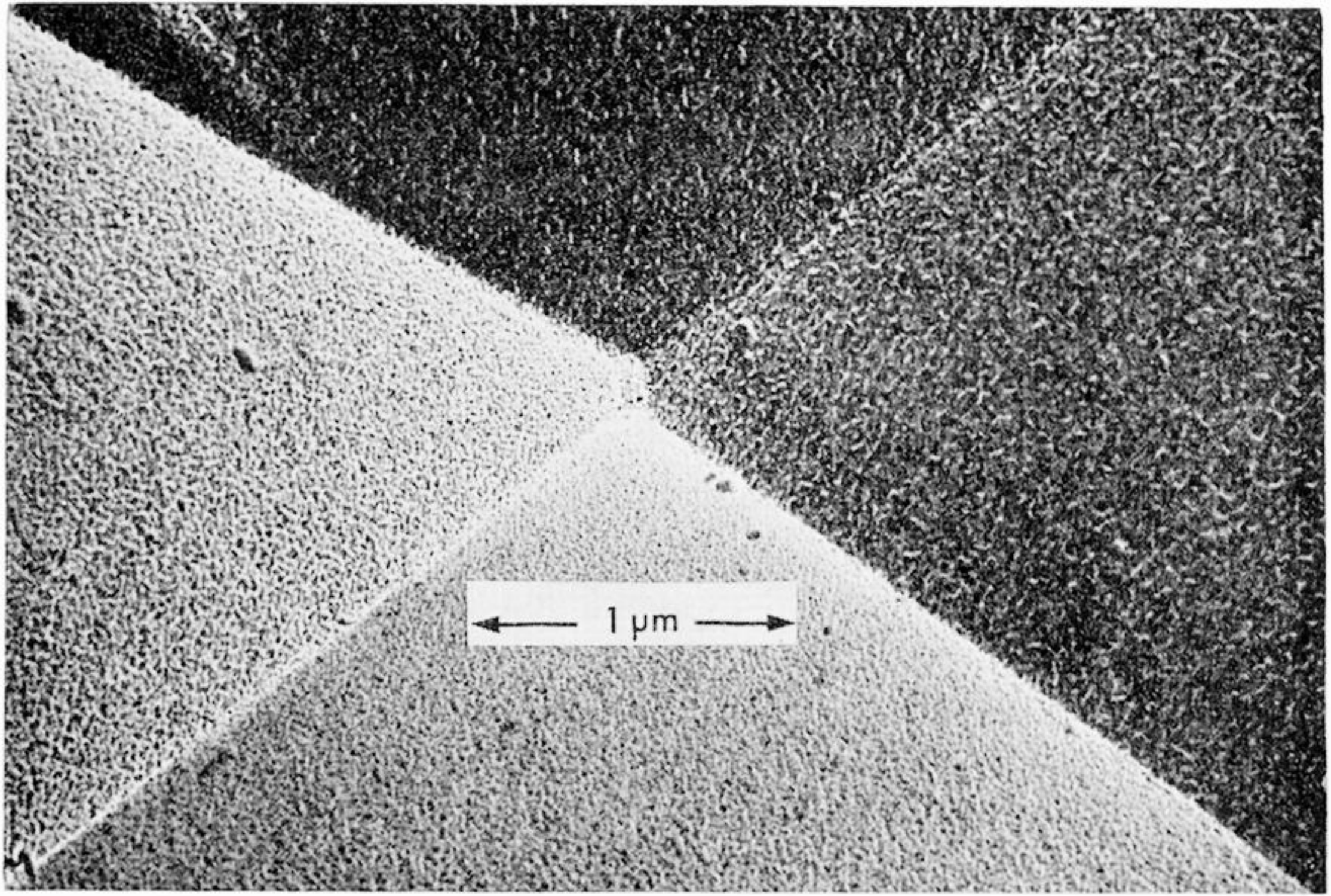

FIGURE 7. Electron micrograph of sharp stylus used in the experiments.