A quantitative approach to the effects of surface topography on tunnelling current between two large rough metal bodies

1991 J. Phys.: Condens. Matter 3 4655

(http://iopscience.iop.org/0953-8984/3/25/013)

J. Phys.: Condens. Matter 3 (1991) 4655–4675. Printed in the UK

A quantitative approach to the effects of surface topography on tunnelling current between two large rough metal bodies

Frédéric Houzé and Lionel Boyer

Laboratoire de Génie Électrique de Paris, Unité Associée au CNRS 0127, Universités Paris VI & Paris XI, École Supérieure d'Électricité, Plateau de Moulon, 91192 Gif-sur-Yvette Cedex, France

Received 14 November 1989, in final form 5 February 1991

Abstract. A convenient method is proposed to deal with the effects of some large-scale surface roughness when calculating the tunnelling current between two macroscopic rough metal bodies—for instance the two members of an electrical 'quasi-contact'. Surface topography is introduced through the distribution of local distances between electrodes, which describes both their nominal shape and their local roughness. Even for surfaces as simple as a plane, cylinder or sphere, this distribution is quite intricate and exhibits a highly tortuous shape. A few periodic roughness models, and a more realistic approach by means of fractals, are successively considered. In any case, we prove that only the beginning of the distance distribution is essential from the viewpoint of the tunnelling current. A computed example allows us to assess quantitatively how the local topography may change drastically any prediction concerning the value of the current. Comparison between some reliable sphere/plane experimental results and the corresponding theoretical predictions proves to be very good and attests to the relevance of the proposed method.

1. Introduction

The work we present below forms part of a study of some types of electric contacts. More precisely, our overall purpose is to develop some models of electronic transport phenomena through electrical 'quasi-contacts'. By quasi-contact we mean a situation in which both contact members are very close together, but remain separated at any point by a layer of an intermediate insulating body that is sufficiently thin to allow tunnelling of electrons. This intermediate body may be either a lubricant introduced into the contact to reduce wear [1–3], an oxide which has formed on the metal surfaces [4], or an organic polymer deliberately deposited on the electrodes [5,6]. We think that this 'quasi-contact' formulation gives a quite good description of real contacts, since in most cases actual metal/metal junctions are known to be rare or even non-existent [7].

As a whole, the problem thus formulated is rather complex. We present, in this paper, the original method we have developed to calculate the tunnelling current contribution in a 'quasi-contact' involving an easily describable intermediate medium (i.e. vacuum or any insulator that can be approximated by the widely used 'conduction-band model'). This method mainly allows us, considering some not too severely rough electrodes, to take into account the effects of roughness, which prove to be crucial

0953-8984/91/254655+21$03.50

from the viewpoint of the tunnelling current. Our purpose is, of course, not to bring an universal answer to the general and intricate problem of tunnelling between two rough surfaces, but rather to make possible quantitative predictions about current values in some topographically well defined and experimentally feasible situations. The commonly used sphere/plane geometry has been chosen for contact members chiefly because it is widely accepted to be representative and because we were able to obtain some reliable physical measurements in such a configuration. As the local radius of curvature at any point of both surfaces is supposed to be very large compared with atomic radii, the theoretical approach we propose in the following is *very different* from the scanning-tunnelling-microscope formalism [8–12].

To begin with, we show that the tunnelling current through the sphere/plane 'quasi-contact' can be worked out by considering separately one function describing the physical phenomena involved and another dealing with the topographic surface data. This latter function, hereafter termed $\partial S/\partial h$, expresses the distribution of distances between the interacting contact members. We then study its shape and behaviour for simple systems like plane/plane, cylinder/plane or sphere/plane, first with some periodic roughness models, and then with a more realistic approach to irregular surface topography by means of fractals. Following this substantial development, some concrete calculations and comments concerning the tunnelling current through a golden sphere/plane 'vacuum quasi-contact' are given by way of illustration. The last section is devoted to a convincing experimental validation involving an original sphere/plane apparatus acting in air. To conclude, the key contributions of our approach are stressed, and we consider other possible applications of our method—especially how to extend it to situations in which the thin intermediate insulating layer requires a more complex energy-band description.

2. Introduction of the $\partial S/\partial h$ distribution

To begin with, let us suppose that both metal surfaces are ideally smooth. Let us assume the plane to be the (x, y) plane and the sphere to be situated in the half-space $(z > 0)$. Let R be its radius, d the shortest distance between 'quasi-contact' members, and $h(x, y)$ the distance between two points on the plane and on the lower half-sphere with the same (x, y) coordinates. In the proximity of the base of the sphere, this distance is not very different from the shortest geometrical one. R is assumed to be large (typically a few millimetres) compared with atomic radii, whereas d must be small enough to allow a finite flow of tunnelling electrons. (For a given voltage drop between two equipotential electrodes, the tunnelling current density J_t is known to decrease exponentially as the separation between the surface elements considered increases [13].)

With the above notation, it is obvious that

$$h(x, y) = d + R - \sqrt{R^2 - x^2 - y^2} \tag{1}$$

and there is a one-to-one mapping between $r = \sqrt{x^2 + y^2}$ and h. Therefore the tunnelling current density J_t can be written either as $J_t(h)$ or as $J_t(r)$. So long as the local slope on the sphere remains very small, we can split the sphere/plane geometry up into a succession of plane/plane tunnelling junctions resulting in concentric rings [14],

and the tunnelling current I_t can then be expressed as

$$I_t = \int_0^{r_{max}} 2\pi r J_t(r)\,dr \tag{2}$$

where r_{max} is the limit beyond which any contribution to the integral may be considered as negligible. Since it is found, in practice, that $r_{max} \ll R$, we justify the plane/plane approximation *a posteriori*.

To deal with real electrodes, we must now take into account the inescapable roughness of surfaces, $\rho_{pl}(x,y)$ and $\rho_{sp}(x,y)$. Let us suppose, as is commonly assumed in such problems, that all the corrugations of both surfaces can be gathered onto one of them, for instance onto the plane [15]. Then the equation of the resulting 'rough plane' becomes $z = \rho_{pl}(x,y) - \rho_{sp}(x,y) = \rho(x,y)$ (instead of $z = 0$). The roughness parameters are assumed to be such that the local slope at any point of the rough plane is small enough to allow the assumption of juxtaposed elementary plane/plane junctions to be retained. The pseudo-distance $h(x,y)$ is now expressed as

$$h(x,y) = d + R - \sqrt{R^2 - x^2 - y^2} - \rho(x,y). \tag{3}$$

Thus the mapping between r and h is no longer one-to-one, and I_t cannot be calculated from (2).

The most judicious way to overcome this difficulty is then to introduce the distribution of local distances between the surfaces, $\partial S/\partial h$, and to express the tunnelling current by

$$I_t = \int_{h_{min}}^{h_{max}} J_t(h)\frac{\partial S}{\partial h}(h)\,dh \tag{4}$$

where h_{min} and h_{max} depend on the roughness $\rho(x,y)$ features; furthermore h_{max} is linked to the convergence of the integral. The current density $J_t(h)$ depends only on the theoretical approach to this specific tunnelling problem. Its exact expression is linked to which energy band model, potential barrier shape, etc, you choose. On the other hand, the function $(\partial S/\partial h)$ conveys only the topographic data. It combines, in fact, the classicaly encountered 'height distributions' [16] relative to each of the two interacting surfaces. In that way, it describes both their nominal shape and their local roughness.

Such a separation between the physical phenomena and the surface topography may also be convenient when studying other kinds of short-distance interactions between two 'large' solids. For instance, we show in appendix A the way it could be used to calculate the mechanical interaction energy and forces between two large, not-too-rough bodies.

3. Typical $\partial S/\partial h$ distributions

In this section we first present some typical $\partial S/\partial h$ distributions for simple periodic roughness models. Although at first glance these models look rather far from reality [16], they are very helpful in understanding the genesis of the $\partial S/\partial h$ distributions, their strange shape and behaviour, and their utmost sensitivity to roughness features. Then we propose a more realistic approach to irregular topographies by means of fractal surfaces. The corresponding $\partial S/\partial h$ functions are given and compared with the previous ones.

3.1. Periodic roughness models

Let us choose initially for the rough plane the two simplest periodic models which come to mind: the 'sinusoidal corrugation' model

$$\rho(x,y) = a\cos\left(2\pi\frac{x}{\lambda_x} + \varphi_x\right) \tag{R1}$$

and the 'crossed sinusoidal corrugation' model

$$\rho(x,y) = a\cos\left(2\pi\frac{x}{\lambda_x} + \varphi_x\right)\cos\left(2\pi\frac{y}{\lambda_y} + \varphi_y\right). \tag{R2}$$

The 'wavelengths' λ_x and λ_y considered at present (typically 0.5 to 2.5 μm) are large compared with atomic radii and small compared to the sphere radius. Optimistic values of the amplitude a for a naturally smooth or well polished surface may be taken to be in the range 10 to 25 Å [17].

To lay the foundations of the $\partial S/\partial h$ distributions, let us consider, first, the parallel smooth plane/rough plane geometry. In this configuration, our $\partial S/\partial h$ for both smooth and rough planes are identical to the classical 'height distribution' for the rough plane alone, with a shift of d.

For a rectangular field of analysis on the smooth plane, λ_x wide in the x-direction and L_y in y-direction, one can obtain (see appendix B) for the roughness model (R1)

$$\frac{\partial S}{\partial h} = \frac{\lambda_x L_y}{\pi a}\left[1 - \left(\frac{h-d}{a}\right)^2\right]^{-1/2} \tag{5}$$

or, using a series expansion

$$\frac{\partial S}{\partial h} = \frac{\lambda_x L_y}{\pi a}\left[1 + \sum_{n=1}^{+\infty}\frac{(2n-1)!!}{2^n n!}\left(\frac{h-d}{a}\right)^{2n}\right]. \tag{6}$$

For a rectangular field of analysis with sides λ_x and λ_y, the expression relative to roughness model (R2) is (see appendix B)

$$\frac{\partial S}{\partial h} = \frac{\lambda_x \lambda_y}{\pi a}\left\{1 + \sum_{n=1}^{+\infty}\left(\frac{(2n-1)!!}{2^n n!}\right)^2\left[1 - \left(\frac{h-d}{a}\right)^2\right]^n\right\}. \tag{7}$$

One may note the similarity of the equations (6) and (7). The two $\partial S/\partial h$ distributions are plotted in figure 1 as histograms ((*a*) and (*d*)). In fact, they reveal a significant difference: for the roughness model (R1), the surface density is maximum around the tops of 'hills' and the bottoms of 'valleys' (i.e. for h around $(d \pm a)$), whereas for the roughness model (R2), the maximum density occurs near the median plane (i.e. for h around d). In addition, it is clear that both histograms (especially (*a*)) are quite far from a Gaussian distribution, which is reputed to be a good approximation for many real surface structures [16].

There is no problem in finding the expression for $\partial S/\partial h$ within configurations involving a smooth plane and any other smooth simple surface like a cylinder or a

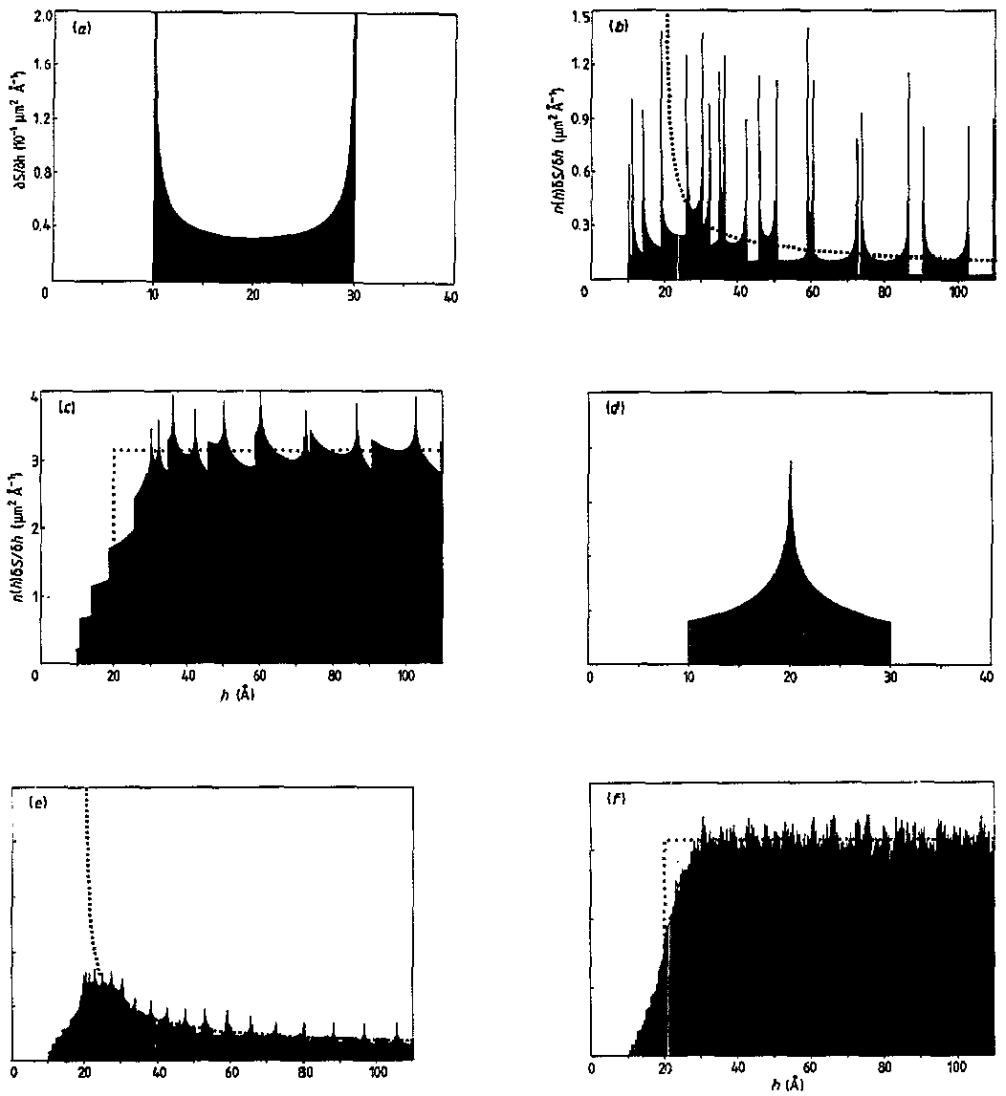


Figure 1. (*a*), (*b*), (*c*) and (*d*), (*e*), (*f*) refer to the roughness models (R1) and (R2), respectively. The distance between nominal surfaces is $d = 20$ Å; roughness parameters are $a = 10$ Å, $\lambda_x = \lambda_y = 1$ μm and $\varphi_x = \varphi_y = 0$. (*a*), (*d*): smooth plane/rough plane geometry. Plot of the surface density $\partial S/\partial h$ from theoretical distributions (6) and (7) as histograms. It can be seen that for (R1) $\partial S/\partial h$ is a maximum around the tops of 'hills' and the bottoms of 'valleys' ($h \simeq d \pm a$), whereas for (R2), the maximum density occurs near the median plane ($h \simeq d$). (*b*), (*e*): smooth cylinder/rough plane geometry. Plot of $n(h)\delta S/\delta h$ obtained numerically with $\delta S = 1$ Å × 1 Å and $\delta h = 0.05$ Å. We consider a 1 μm-long portion of a cylinder of circular cross section with curvature $R = 5$ mm and its axis of revolution parallel to the y-axis. Note the multiplicity of juxtaposed and sometimes overlapping basic patterns. When excluding peaks, the general appearance of these histograms, particularly clear in (e), is due to the $\partial S/\partial h$ distribution (8) for smooth solids (dotted curves). (*c*), (*f*): smooth sphere/rough plane geometry. Plot of $n(h)\delta S/\delta h$ obtained numerically with $\delta S = 4$ Å × 4 Å and $\delta h = 0.05$ Å for a sphere with radius $R = 5$ mm. Comments are as before, the $\partial S/\partial h$ distribution for smooth solids being given in this case by (9) instead of (8).

sphere. For a smooth plane and a portion of length L of a smooth cylinder of circular section with radius R and its axis parallel to the plane, we obtain

$$\frac{\partial S}{\partial h} = 2L\frac{R+d-h}{\sqrt{(h-d)(2R+d-h)}} \approx L\sqrt{\frac{2R}{h-d}} \qquad \text{when } h-d \ll R \tag{8}$$

and, for a smooth sphere with radius R facing a smooth plane

$$\frac{\partial S}{\partial h} = 2\pi(R+d-h) \approx 2\pi R \qquad \text{when } h-d \ll R. \tag{9}$$

The complexity of the expressions (6) and (7) for such a simple configuration as the plane/plane one suggests that their combination with (8) and (9) when a cylinder or a sphere is facing a rough plane will make an analytic solution impossible. This is why it is essential to deal with these more intricate situations by numerical calculation.

The way the computation is performed is quite simple. We split up the field of analysis on the (x, y) plane into a vast number of tiny squares with area $\delta S = \delta x \delta y$. Then, the pseudo-distance $h(x, y)$ is calculated at the centre of each elementary square and classified with a step δh. In this way, we finally get the successive numbers $n(h_k)$ of elementary squares whose h belongs to the interval $[h_{\min} + (k-1)\delta h,\ h_{\min} + k\delta h[$. Calculation parameters δx, δy and δh must, however, be chosen in such a way that any successive 'contour lines', $(h = h_k)$ and $(h = h_{k+1})$, are always separated by a few elementary squares, which definitely occurs when

$$\left(\frac{\partial h}{\partial x}\right)_{\max} \delta x \ll \delta h \qquad \left(\frac{\partial h}{\partial y}\right)_{\max} \delta y \ll \delta h. \tag{10}$$

Under such conditions, we may take it as certain that, for a given h

$$\lim_{\delta h \to 0} n(h)\frac{\delta S}{\delta h} = \frac{\partial S}{\partial h}(h) \tag{11}$$

and thus this numerical method does allow us to plot a faithful shape of intricate $\partial S/\partial h$ distributions. Unfortunately, δh cannot, in practice, be made as small as one would wish, since the conditions (10) result in corresponding small values for δx and δy, and therefore result in a prohibitively long calculation time.

We present in figure 1 some results obtained within cylinder/plane and sphere/plane configurations (1(b), 1(e) and 1(c), 1(f), respectively). In both situations, the general appearance (i.e. excluding peaks) is due to functions (8) or (9) relative to smooth solids (dotted curves).We notice that the basic pattern obtained with the plane/plane geometry, for roughness models (R1) or (R2), (figures 1(a) and 1(d), respectively), is repeated with occasional overlapping, as if some kind of spatial convolution had occurred. In fact, the origin of such tortuous shapes may be quite easily understood, at least in the case of the cylinder/plane configuration, as we explain in figure 2: a discontinuity, for a given h, conveys that a 'hill' or a 'valley' begins or ceases to participate in the distribution of local distances.

Once suitable programs have been developed, there is nothing to prevent one from studying more complex roughness models. We present for instance in figure 3 the $\partial S/\partial h$ distributions concerning plane/plane and sphere/plane configurations for two

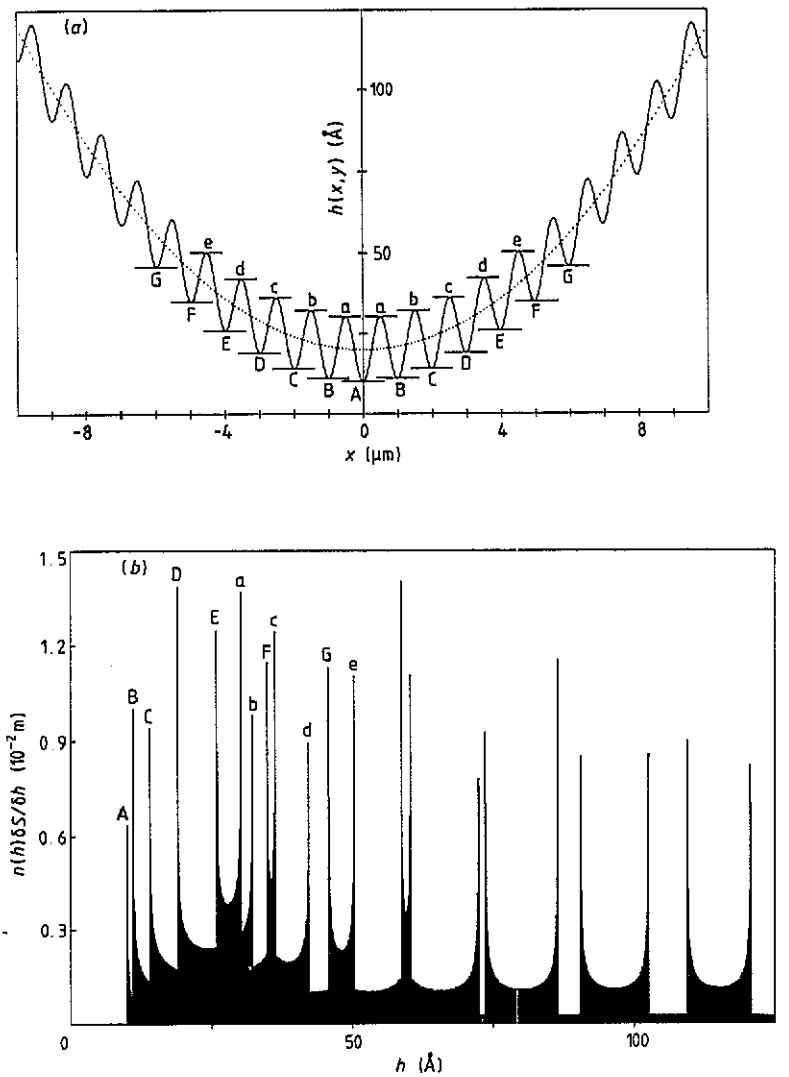


Figure 2. Origin of the tortuous shape of a $n(h)\delta S/\delta h$ histogram in the case of the smooth cylinder/rough plane geometry with roughness model (R1). All parameters concerning calculation and topography are as given in figure 1(*b*). (*a*) Plot of $h(x) = d + R - \sqrt{R^2 - x^2} - a\cos(2\pi x/\lambda_x)$ versus x. NB. We emphasize that the very dissimilar scales on the x- and h-axes make the local slope look, in places, very steep, whereas in practice it never exceeds 5%. (*b*) Plot of the corresponding histogram. We see that a discontinuity, for a given h, signifies that a 'hill' (upper case letters), or a 'valley' (lower case letters), begins or ceases to participate in the distance distribution. For h increasing from 0, the 'hills' A–E successively begin to contribute to the histogram, each of them adding the beginning of a slightly distorted basic pattern on the histogram. Then the 'valleys' a and b occur: elementary tiny square areas relating, first to the central corrugation, then to the next two, cease to be taken into account, inducing two successive ends of the basic pattern on the histogram. This continues, with alternating 'hills' and 'valleys' generating a succession of truncated, slightly distorted and overlapping basic histograms. Such a structure is obviously very sensitive to parameters such as a, λ_x or φ_x: a small change in one of these parameters may result in a completely different arrangement of the truncated patterns.

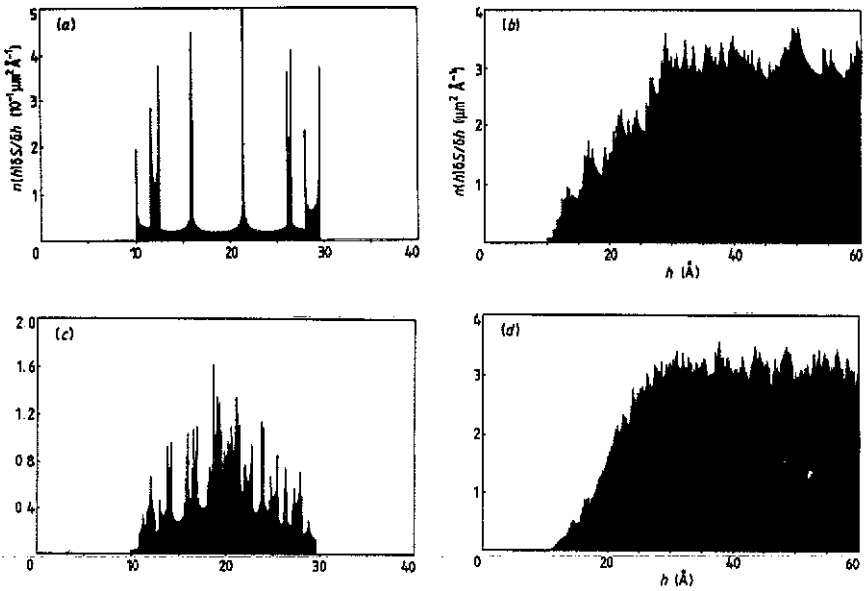


Figure 3. (*a*), (*b*) and (*c*), (*d*) refer to the double periodic roughness models (R3) or (R4), respectively. Roughness parameters are $a_1 = 9$ Å, $a_2 = 1$ Å, $\lambda_{1x} = \lambda_{1y} = 1$ μm, $\lambda_{2x} = \lambda_{2y} = 0.1$ μm, and $\varphi_{1x} = \varphi_{1y} = \varphi_{2x} = \varphi_{2y} = 0$. (*a*), (*c*): smooth plane/rough plane geometry. Plot of $n(h)\delta S/\delta h$ numerically obtained with $\delta S = 1$ Å × 1 Å and $\delta h = 0.05$ Å. As could be logically expected, superposing a smaller periodic component onto the previous one on each histogram in figures 1(*a*) and 1(*d*) results in small replicas of the basic patterns themselves. Note that the 'double crossed sinusoidal corrugation' leads to a histogram (*c*) whose shape is a little nearer to a Gaussian distribution. (*b*), (*d*): smooth sphere/rough plane geometry. Plot of $n(h)\delta S/\delta h$ obtained numerically with $\delta S = 4$ Å × 4 Å and $\delta h = 0.05$ Å. Comments are as before, but the parts refer to the histograms in figures 1(*c*) and 1(*f*).

double-periodic roughness models derived from (R1) and (R2):

$$\rho(x,y) = a_1 \cos\left(2\pi\frac{x}{\lambda_{1x}} + \varphi_{1x}\right) + a_2 \cos\left(2\pi\frac{x}{\lambda_{2x}} + \varphi_{2x}\right) \tag{R3}$$

and

$$\rho(x,y) = a_1 \cos\left(2\pi\frac{x}{\lambda_{1x}} + \varphi_{1x}\right) \cos\left(2\pi\frac{y}{\lambda_{1y}} + \varphi_{1y}\right) + a_2 \cos\left(2\pi\frac{x}{\lambda_{2x}} + \varphi_{2x}\right) \cos\left(2\pi\frac{y}{\lambda_{2y}} + \varphi_{2y}\right). \tag{R4}$$

One may notice that the 'double crossed sinusoidal corrugation' (R4) leads to a histogram (figure 3(*c*)) whose shape is nearer to the Gaussian distribution. Therefore it is clear that any multiple-periodic roughness model could allow to render quite correctly some features of actual surfaces. However, a fractal viewpoint seems to be a more judicious and neat approach to the real complexity of surface topography.

3.2. Example of fractal modelling

Mandelbrot's fractal geometry has blossomed widely in the past few years, since it provides both a description and a mathematical model for many of the complex shapes

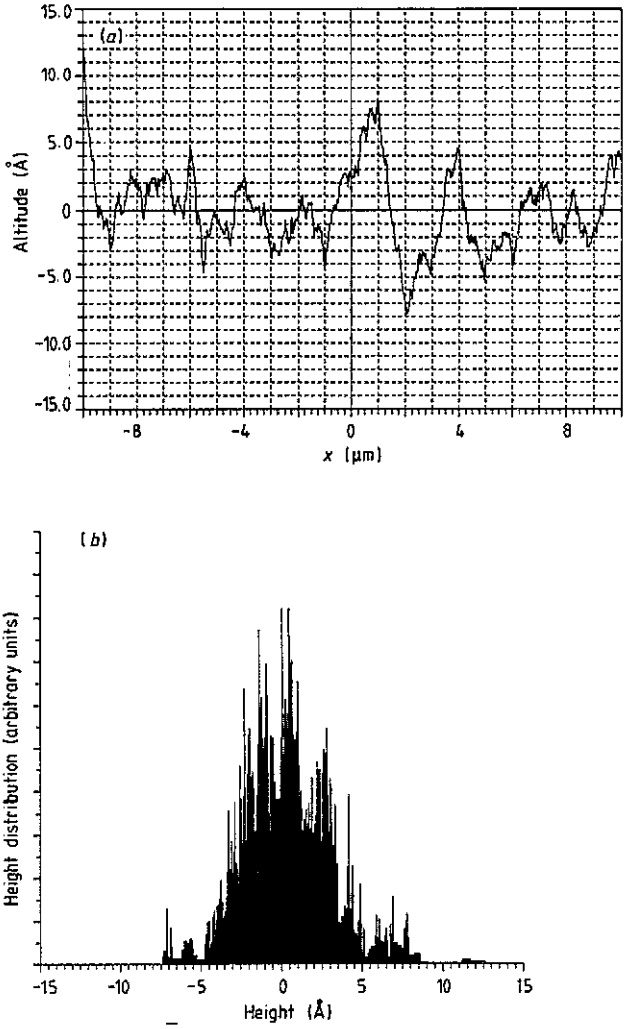


Figure 4. (*a*) Fractal rough profile generated as explained in subsection 3.2 from 21 equidistant positions on the x-axis between -10 μm and $+10$ μm. The reduction ratio is $r = \frac{1}{2}$ and the Hurst exponent $H = 0.65$. This profile looks quite realistic and could have been recorded experimentally. NB. The very dissimilar scales on the x- and z-axes exaggerate the local slope, which never exceeds 5% in practice. (*b*) The classical height distribution corresponding to the profile in (*a*). The smoothed figure of the histogram presents as expected a pleasing Gaussian shape.

and seemingly self-similar patterns in nature [18]. Of course, it appears very suitable for the realistic simulation of the topography of surfaces [19–21].

In this subsection we present a bi-dimensional fractal model for the rough plane (i.e. such that its calculated profile is the same in any section perpendicular to a given direction—for instance the y-axis). This simplification results in notably shorter computation times for $\partial S/\partial h$ distributions. Although it may seem quite restrictive, we think it can give a correct representation of the tiny area—typically a few μm^2 (see below)—that is active from the tunnelling current viewpoint, since an actual physical surface does often present a local bi-dimensional structure [16].

The way we obtain our fractal rough profile is inspired by the 'successive random addition' algorithm introduced by Voss [22] to generate fractional Brownian motion with an arbitrary resolution. The starting point is a sequence of positions $x_1, x_2, \ldots, x_N$ with corresponding altitudes $z_1, z_2, \ldots, z_N$. To be concrete, let us choose $N = 3$; let us assume that the x_i are equidistant (for instance 0, $\frac{1}{2}$ and 1) and the initial altitudes are set equal to zero. Next, these three altitudes are given a random addition chosen from a normal distribution with zero mean and unit variance $\sigma_1^2 = 1$. Then the midpoints of the x_i intervals become additional positions at which the altitudes are estimated by interpolation. The positions are now $x_1, x_2, \ldots, x_5 = 0, \frac{1}{4}, \frac{1}{2}, \frac{3}{4}, 1$. All five altitudes are then given a random addition with zero mean and a reduced variance $\sigma_2^2 = r^{2H}\sigma_1^2$. H is called the 'Hurst exponent' and it lies in the range 0 to 1; r is a reduction ratio. The five positions obtained are again interpolated to the midpoints of the intervals to give nine new positions, and so on. After n such random addition–interpolation cycles, $(1+2^n)$ points of the profile are obtained, the variance of the addition in the nth generation of the process being $\sigma_n^2 = r^{2(n-1)H}\sigma_1^2$. Voss has shown that such a process leads to self-affine curves with a fractal dimension $D_c = 2 - H$. Therefore the surface resulting from sliding such a curve along the y-axis has a fractal dimension $D_s = 3 - H$.

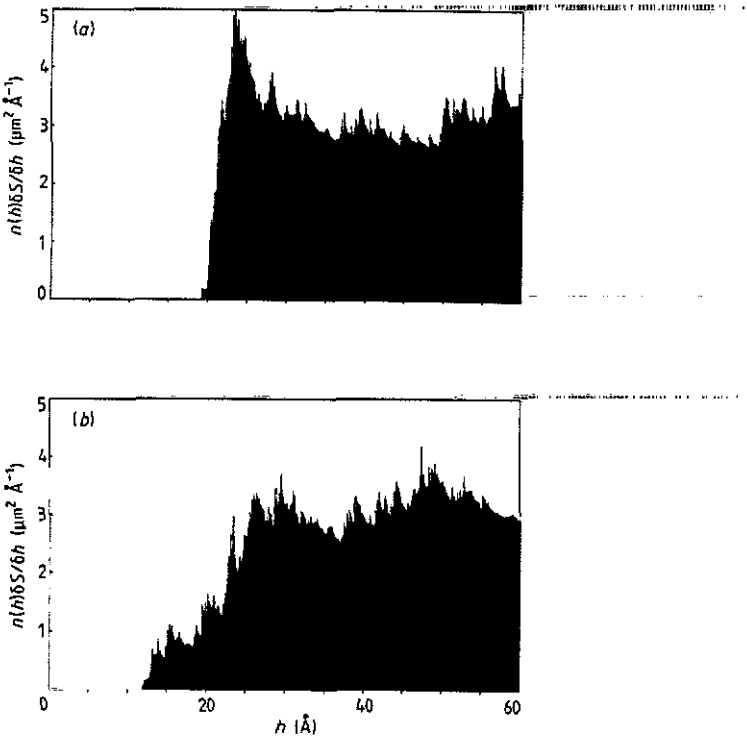


Figure 5. Two examples of $\partial S/\partial h$ distributions relating to the smooth sphere/fractal rough plane geometry. The smooth sphere ($R = 5$ mm) is 20 Å from the median plane ($z = 0$); the abscissa of its centre is (a) -2 μm and (b) $+2$ μm, respectively. One can see the considerable effect of the lateral position of the sphere upon $\partial S/\partial h$ shape. Huge discrepancies occur, particularly at the beginning of the distribution, which can be very steep (a) or rather slowly increasing (b).

We have generated such fractal curves from $N = 21$ equidistant positions on the x-axis in the range -10 μm to $+10$ μm and for various values of r and H, bearing in mind the conditions (10) on the local slope. We show in figure 4(a) a pleasing profile solution obtained using $n = 5$ generations of the above algorithm with $r = \frac{1}{2}$ and $H = 0.65$ (the profile obtained from the algorithm has been rescaled in order to have zero mean and to be such that $z_{max} - z_{min} = 2a$, i.e. the peak-to-valley amplitude of periodic models (R1) and (R2)). This computed profile looks quite similar to real measured ones. It must be borne in mind that the *very dissimilar scales* on the x- and z-axes *exaggerate* the local slope, which in practice never exceeds 5%. The corresponding height distribution is shown in figure 4(b): it presents, as it could be expected, a pleasant Gaussian-like shape.

Then numerous $\partial S/\partial h$ distributions can easily be obtained for the smooth sphere with radius R facing our fractal rough plane: for a given distance d between nominal surfaces, each small lateral displacement of the sphere leads to a new relative topographic situation. Despite unchanged individual height distributions for each body, the effect on the behaviour of $\partial S/\partial h$ is quite notable, as one may realize by looking at figure 5. All this proves that *it is impossible to use reasoning based on the usual statistical parameters when the local distance between smoothed surfaces is variable.* Huge discepancies occur, especially at the start of the distance distributions, which is precisely the most crucial part of them from the viewpoint of tunnelling current, as we show immediately below.

4. Application: quantitative calculations of tunnelling current through a sphere/plane vacuum 'quasi-contact'

The $\partial S/\partial h$ distributions enable us to calculate in concrete terms the effects that the different roughness models described above have on the tunnelling current flowing between two very close metal electrodes, one being a plane and the other a sphere, separated by vacuum. The tunnelling current, as expressed in equation (4), may be approximated numerically, considering (11), as

$$I_t \approx \sum_k J_t(h_k) n(h_k)\, \delta S. \tag{12}$$

In practice, the terms are summed until the quantity $|1 - I_{t(p+1)}/I_{t(p)}|$ becomes less than a given η, which occurs very quickly because of the exponential decrease of the current density $J_t(h)$. We present in figure 6 the typical behaviour of $J_t(h)$, $\partial S/\partial h(h)$ and the partial integral $I_t(h) = \int_{h_{min}}^{h} J_t(u)\partial S/\partial u(u)\,du$. The tunnelling current density shown relates to gold electrodes and a voltage drop set at 100 mV. We used for its calculation the free-electron model and the true potential barrier shape, including the image force expression due to Simmons [23]. For h greater than 4 Å, $J_t(h)$ has been found to decrease exponentially as $J_t(h) = J_t(0)\exp(-h/h_0)$, with $J_t(0) = 9.745 \times 10^{15}$ A m^{-2} and $h_0 = 0.47$ Å. The $\partial S/\partial h$ distribution plotted relates to a given position of the smooth sphere ($R = 5$ mm) above our fractal rough plane described in subsection 3.2 (distance between nominal surfaces: $d = 15$ Å). One can verify that the exponential decrease of $J_t(h)$ literally overshadows the tortuous shape of the $\partial S/\partial h$ distribution and that the integral in (4) (or its discrete equivalent expression in (12)) does converge very quickly, i.e. for h varying within a few angströms.

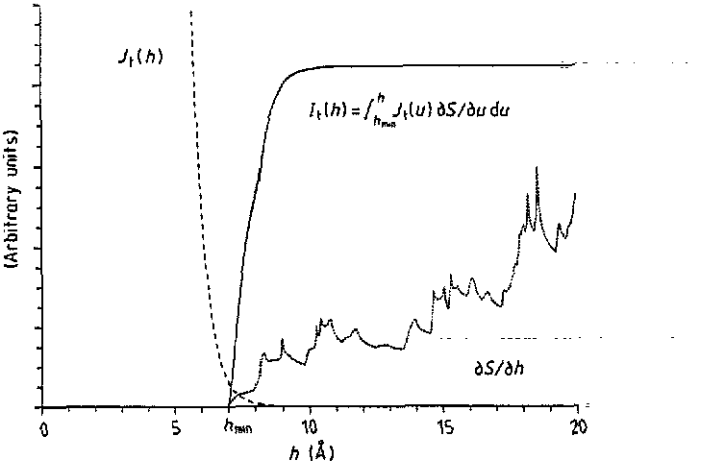


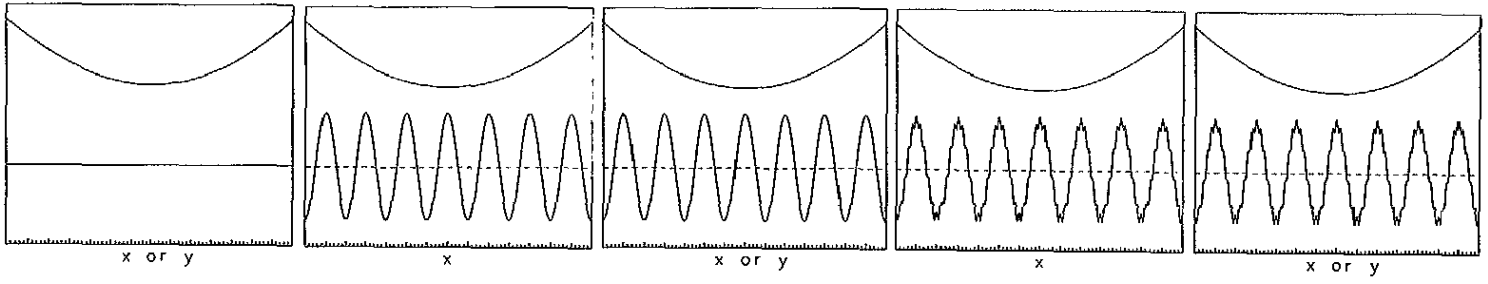
Figure 6. Typical behaviour of the tunnelling current density $J_t(h)$, the combined height distribution $\partial S/\partial h(h)$ and the partial integral $I_t(h) = \int_{h_{min}}^{h} J_t(u)(\partial S/\partial u)(u)\,du$, for the smooth sphere/rough plane geometry. The $J_t(h)$ shown relates to gold electrodes and a voltage drop set to 100 mV. For h greater than 4 Å, we get the exponentially decreasing law $J_t(h) = J_t(0)\exp(-h/h_0)$, with $J_t(0) = 9.745 \times 10^{15}$ A m^{-2} and $h_0 = 0.47$ Å. The $\partial S/\partial h$ distribution shown refers to a given position of the smooth sphere above the fractal rough plane described in subsection 3.2 (abscissa of the sphere centre, +2 μm; radius, 5 mm; distance between nominal surfaces, 15 Å). One can verify that the exponential decrease of $J_t(h)$ literally overshadows the tortured shape of $\partial S/\partial h$, and that the integral giving I_t converges within a few angströms. Therefore, only the start of the $\partial S/\partial h$ distribution is essential from the viewpoint of tunnelling current.

Therefore, *only the start of the $\partial S/\partial h$ distribution is essential from the viewpoint of tunnelling current.* This important result, that is in accordance with intuition, is brought out through our method in a *quantitative* way.

The value h^* for which one gets 99% of the total current may be easily worked out. Then, the projection onto the (x, y) plane of the electrodes areas such as $h_{min} \leqslant h \leqslant h^*$ can be regarded as 'tunnelling current spots'. Results that relate to the periodic and non-periodic roughness models described in section 3 are summarized in tables 1 and 2. The physical conditions are those described in figure 6 (gold/vacuum/gold configuration, $V_{app} = 100$ mV, $d = 15$ Å). In connection with this we show the calculated current values (second row), the profiles of both electrodes within and

Table 1. (Following folio) Summarized results of the tunnelling current for the periodic roughness models presented in subsection 3.1 (gold sphere/plane separated by vacuum, $V_{app} = 100$ mV, $d = 15$ Å). The case of two smooth electrodes is given as a reference. We present the calculated current values (second row), the profiles of both considered electrodes within and around the tunnelling active zone (third row), and the resulting 'current spots' (fourth row) through which 99% of the tunnelling current is flowing. The profiles are plotted in a rectangular frame corresponding to a 7 μm distance along the x- (or y-) axis and a variation from −15 Å to +30 Å along the z-axis. The current spots are presented within square frames which represent 7 μm×7 μm areas of the (x, y) plane. From these results, it is clear that even a very small periodic roughness is able to cause a huge increase in the tunnelling current.

Smooth surfaces (reference)	Sinusoidal roughness	Crossed sinusoidal roughness	Double sinusoidal roughness	Double crossed sinusoidal roughness
	$\rho(x,y)=a\cos(2\pi x/\lambda_x)$	$\rho(x,y)=a\cos(2\pi x/\lambda_x)\times\cos(2\pi y/\lambda_y)$	$\rho(x,y)=a_1\cos(2\pi x/\lambda_{1x})+a_2\cos(2\pi x/\lambda_{2x})$	$\rho(x,y)=a_1\cos(2\pi x/\lambda_{1x})\times\cos(2\pi x/\lambda_{1y})+a_2\cos(2\pi x/\lambda_{2x})\times\cos(2\pi y/\lambda_{2y})$
	a=10 Å, λ_x=1 μm	a=10 Å, $\lambda_x=\lambda_y$=1 μm	a_1=9 Å, a_2=1 Å λ_{1x}=1 μm, λ_{2x}=0.1 μm	a_1=9 Å, a_2=1 Å, $\lambda_{1x}=\lambda_{1y}$=1 μm $\lambda_{2x}=\lambda_{2y}$=0.1 μm
$(I_t)_{smooth}$ = 0.23 nA	$I_t/(I_t)_{smooth} \approx 1.2 \times 10^8$	$I_t/(I_t)_{smooth} \approx 2.1 \times 10^7$	$I_t/(I_t)_{smooth} \approx 4.1 \times 10^7$	$I_t/(I_t)_{smooth} \approx 4.9 \times 10^6$



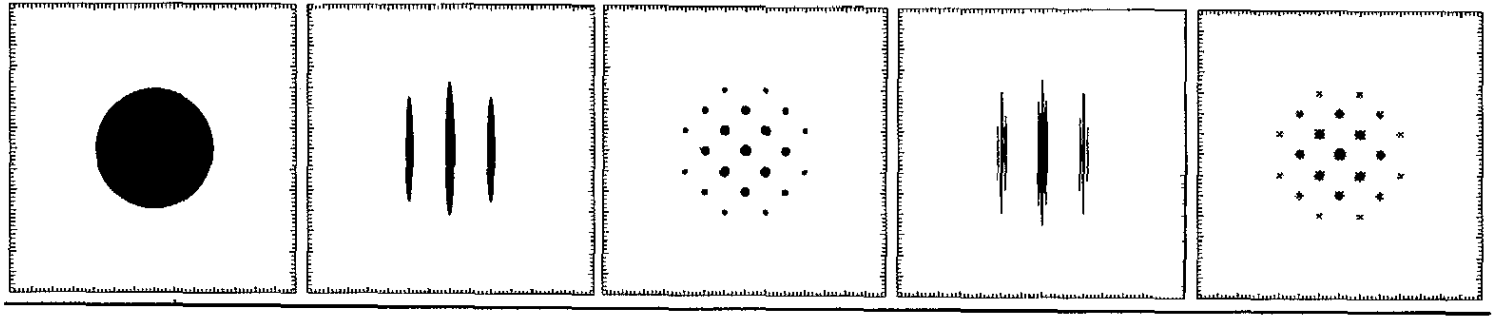
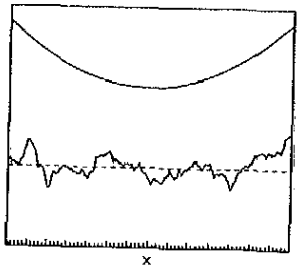
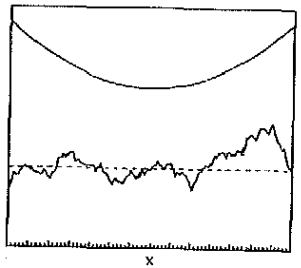
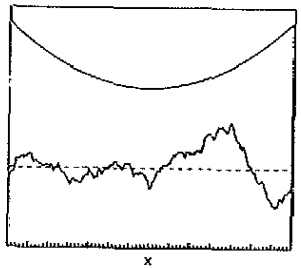
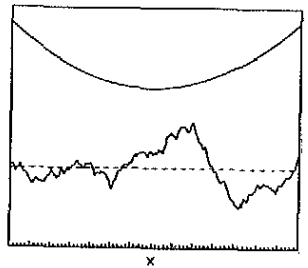
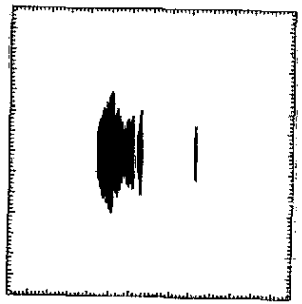
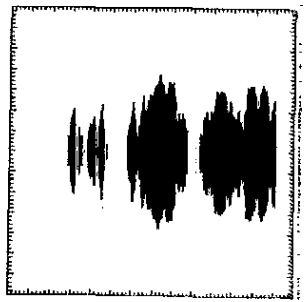
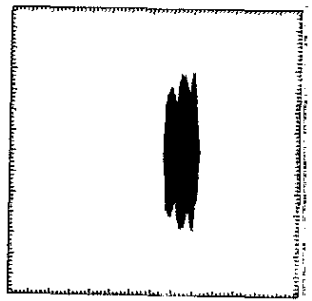


Table 2. Summarized results of the tunneling current for the fractal roughness model presented in subsection 3.2. Conditions and presentation are as in table 1. The abscissa of the sphere centre is successively -3 μm, -2 μm, -1 μm and 0. These different lateral positions allow us to assess how, for a given geometry of nominal surfaces, the local topography modifies tremendously both the current value and the shape of the spots.

Fractal roughness (cf subsection 3.2 and figure 4).
The smooth sphere takes up different lateral positions above the rough plane.

$I_t/(I_t)_{smooth} \approx 7.3$	$I_t/(I_t)_{smooth} \approx 0.5$	$I_t/(I_t)_{smooth} \approx 26.6$	$I_t/(I_t)_{smooth} \approx 3.6 \times 10^4$
			
			

around the tunnelling active zone (third row) and the resulting 'current spots' (fourth row). One can assess quantitatively how the local topography significantly modifies both the current values and the shapes of the spots. As these spots show a slight lateral expansion around the peaks of dominant bumps and therefore mainly concern electrodes areas of smaller local slopes, *the technique described actually works for surfaces hillier than those hitherto considered* (it will be enough if the local slope in the active zones remains below a tolerable threshold). The case of some real, more severely rough surfaces can thus be tackled through our method (cf section 5).

To complete this section, let us mention that the exponential decrease of $J_t(h)$ (for h greater than 4 Å) causes the exponential decrease of the tunnelling current I_t as a function of the distance d between nominal surfaces. This appears to hold true whatever $\partial S/\partial h$ distribution one considers, since changing d is, in fact, equivalent to shifting the $\partial S/\partial h$ distribution along the h-axis (see appendix C for the development of this argument).

5. Experimental validation

Some actual physical measurements have recently allowed us to test the validity of our theoretical predictions in a topographically well-controled gold sphere/air/gold plane situation. This experimental side is the main spin-off from our collaboration with researchers at the Ecole Centrale de Lyon (Professor J M Georges, A Tonck and J L Loubet), who designed an original and very reliable sphere/plane apparatus allowing several mechanical and electrical characteristics to be measured simultaneously for very slight separations.

The principles and features of the quite exceptional instrument developed by Tonck *et al* have already been detailed in recent papers [24, 25]. Briefly, a feedback loop allows to move a macroscopic spherical body towards and away from a plane object, with a resolution of about 0.15 Å in terms of the relative displacement of the solids. Electrical measurements can be performed by simply using metal or plated bodies. The determination of the relative separation between surfaces is carried out through capacitive transducers, the origin being extrapolated from the sphere/plane capacitor variations as a function of the relative displacement. It can be shown that the separation thus measured is nearly identical to the distance d between the smoothed surfaces [26]. The electrodes used to get the data reported below were a silicon plane and a pyrex sphere of radius 3.5 mm, both covered with a 400 Å cobalt sticking layer and a 600 Å terminal gold coating. A STM examination of such surfaces showed a gently undulating aspect—like an irregular version of our 'crossed-sinusoidal model'—similar to the mesoscopic roughness already observed on some Au thin layers [27], with peak-to-valley heights of 50–60 Å and a mean distance between adjacent bumps of about 500 Å. An exhaustive presentation of the main mechanical and electrical results obtained with these electrodes in ambiant air will be given in a forthcoming letter [28].

Figure 7 shows some current versus distance experimental characteristics (broken curves) measured for applied direct voltages V_{app} of 10 mV and 50 mV. The arrows indicate the inward and outward approach, carried out with a constant speed of relative displacement of 0.5 Å s^{-1} (quasi-static operation). In view of the considerable variation of the current, $\log(I_{mes})$ is plotted instead of I_{mes}. The linearity of the curves

thus obtained conveys a law of the form

$$I(d, V_{app}) = I_0(V_{app}) \exp(-d/d_0) \tag{13}$$

which is similar to the general prediction of appendix C for tunnelling current between two gently rough electrodes. The experimental value found for d_0 is 0.87 Å.

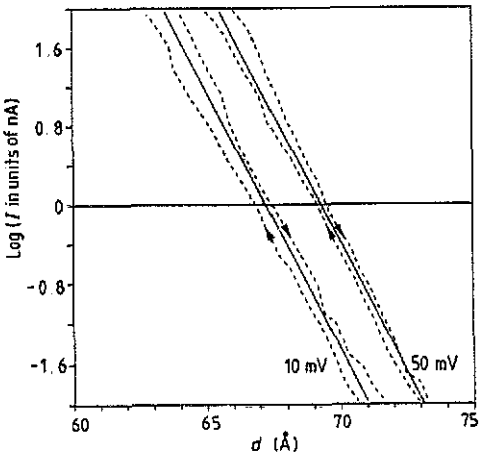


Figure 7. Plot of the current versus distance characteristics obtained for low direct voltages in a gold sphere/air/gold plane experiment (broken curves), and the corresponding theoretical curves calculated with our method (full lines). The fitting is all the more convincing since the roughness parameters considered in the model are very similar to the real ones observed by STM.

To fit the measured characteristics with theoretical curves stemming from our model requires one to choose suitably: (i) the main physical parameters such as the 'effective work function' Φ_{eff} of the golden electrodes and the permittivity ϵ_{eff} of the inter-electrode medium that are necessary for the determination of the current density function, and (ii) a roughness model of the surfaces, essential for any current value calculation. Concerning physical parameters, the best agreement has been found by taking $\Phi_{eff} = 1.27$ eV and $\epsilon_{eff} = 8\epsilon_0$ [26,28]; a plausible interpretation is proposed in these references in terms of the existence of a condensed water micromeniscus in the interface region. As regards surface roughness, a 'crossed-sinusoidal model' with $a = 29$ Å and $\lambda_x = \lambda_y = 500$ Å—i.e. a topography very similar to the real one observed with STM—has proved to be very good, as one may be persuaded by considering the theoretical curves (full lines) in figure 7. Such a nearly perfect agreement between calculation and data has also be obtained for current versus voltage characteristics [26,28].

6. Conclusion

We propose in this paper an original method allowing to quantify the effects of a large-scale surface roughness when calculating the tunnelling current between two large rough metal bodies. Based on the assumption of small local slopes, this method

proposes that it is possible to separate the physical and topographic contributions—in this case the current density function $J_t(h)$ and the distribution of local distances between electrodes $\partial S/\partial h$. As the slope conditions actually only concern tunnelling active areas, some quite severely rough surfaces can be tackled perfectly well in this way, provided that the immediate vicinity of the peaks of the dominant bumps do not curve too steeply. Some rather academic periodic models have been considered to bring out the physical content and intrinsic interest of $\partial S/\partial h$ distributions. A more realistic approach of irregular surfaces has also been carried out by means of fractals. In any case, the illustrative calculations clearly bring to the fore quantitatively: (i) the necessity of a sound knowledge of the combined local height distribution $\partial S/\partial h$ in order to obtain a credible prediction of current values, (ii) the crucial role of the low part of this distribution and (iii) the small size of the tunnelling active zones (which we call the 'tunnelling current spots'). Even more realistic models of irregular surfaces could quite easily be envisaged, for instance generalizing Voss' algorithm to three-dimensional irregularities [20, 22]. However, some of the 'naive' periodically bumped surfaces presented in the text can give a quite relevant description of certain real topographies, as shown convincingly by the nearly perfect agreement between calculation and experimental data in the previous section.

Parameters that require a significantly wider computation field, like capacitance (which involves the whole surface of electrodes), can also be dealt with perfectly satisfactorily via our method, provided that this 'local' technique and the ' macroscopic' calculation are carefully coupled together [26]. The specific case of the sphere/plane capacitance should in the coming months give rise to some experimental investigations using the original apparatus described briefly in the text.

The extension of the method to situations in which the thin intermediate insulating layer requires a more elaborate electronic model can be envisaged within the same global principle. However, the calculation of the current density then becomes more complex, since, if we insist on regarding a band structure as still being valid, two transport processes must now be considered [29]. The first one involves the bottom of the insulator conduction band, whereas the second one involves the top of its valence band ('two-band' models [30–33]). The crucial parameters are then the electron affinity and the energy gap of the insulator, as well as, of course, the metal work function. Under suitable circumstances, it can be expected that one of the barriers involved may be lowered sufficiently to increase the tunnelling current density significantly. Taking these ideas into consideration with the prospect of adding to the understanding of electric contact interfaces forms part of our objectives in the near future.

Acknowledgments

We thank Professor J M Georges, A Tonck and J L Loubet (Ecole Centrale de Lyon), Professor B Sapoval and M Rosso (Ecole Polytechnique), for stimulating discussions, and Professor G Fournet (Université Paris VI) for his careful reading of the manuscript and for constructive comments.

Appendix A

To illustrate the fact that the separation between the physical phenomena and the surface topography may also be useful in calculating the mechanical interaction energy

and forces between two large, gently rough solids, we shall simply develop the broad lines of the calculation relative to mechanical interaction between a sphere and a rough plane, proceeding as Israelachvili does with ideally smooth surfaces [34].

Let us keep the same notation for the geometrical parameters. The potential energy of two atoms or small molecules separated by a distance u is assumed to be purely attractive and of the form $w(u) = -C/u^n$. Both solids are supposed to be made of the same material with a concentration ρ of atoms or molecules. The net interaction energy, for an atom or a molecule at a distance h away from the surface of a plane, may be written

$$w(h) = -\frac{2\pi C\rho}{(n-2)(n-3)h^{n-3}}. \tag{A1}$$

One may easily obtain for the net interaction energy between the whole sphere and the rough plane the expression

$$W = -\frac{2\pi C\rho^2}{(n-2)(n-3)}\int_{h_{\min}}^{h_{\max}} \frac{S(h)}{h^{n-3}}\,\mathrm{d}h. \tag{A2}$$

Integrating by parts and considering the fact that $R \gg h_{\min}$ and, therefore, that only small values of h contribute to the integral, we then obtain

$$W = -\frac{2\pi C\rho^2}{(n-2)(n-3)(n-4)}\int_{h_{\min}}^{h_{\max}} \frac{1}{h^{n-4}}\left(\frac{\partial S}{\partial h}\right)\mathrm{d}h \tag{A3}$$

which is a function of d through $h_{\min}$, $h_{\max}$ and $\partial S/\partial h$. The attractive force F_z in the z-direction may be obtained by the well known expression $F_z = -\partial W/\partial d$. The case $n = 6$ corresponds to van der Waals' law.

Appendix B

We present here the way we get the analytical expressions (6) and (7) for $\partial S/\partial h$ in the smooth plane/rough plane situation with the periodic roughness models (R1) and (R2).

Let us consider, first, the roughness model (R1). It is clear that $h(x, y)$ only depends on the x coordinate and is given by

$$h(x, y) = d - a\cos\left(2\pi\frac{x}{\lambda_x} + \varphi_x\right). \tag{B1}$$

For a rectangular field of analysis on the smooth plane, λ_x wide in the x-direction and L_y long in the y-direction, the area $S(h)$ which is less than h away from the rough plane is expressed as

$$S(h) = \frac{\lambda_x L_y}{\pi}\cos^{-1}\left(\frac{d-h}{a}\right) \tag{B2}$$

from which we get

$$\frac{\partial S}{\partial h} = \frac{\lambda_x L_y}{\pi a}\left[1 - \left(\frac{h-d}{a}\right)^2\right]^{-1/2}. \tag{B3}$$

This last equation may also be written in the form of a series expansion as

$$\frac{\partial S}{\partial h} = \frac{\lambda_x L_y}{\pi a}\left\{1 + \sum_{n=1}^{+\infty} \frac{(2n-1)!!}{2^n n!}\left(\frac{h-d}{a}\right)^{2n}\right\} \tag{B4}$$

writing $(2n-1)!! = (2n-1) \times (2n-3) \times \cdots \times 3 \times 1$.

The case of the second roughness model (R2), in the same geometric configuration, is somewhat more complicated. Now h depends on both the x and y coordinates, and is given by

$$h(x,y) = d - a\cos\left(2\pi\frac{x}{\lambda_x} + \varphi_x\right)\cos\left(2\pi\frac{y}{\lambda_y} + \varphi_y\right). \tag{B5}$$

For a rectangular field of analysis with sides λ_x and λ_y on the smooth plane, one may write $S(h)$ as

$$S(h) = 8\int_{-(\varphi_x/2\pi)\lambda_x}^{x_M(h)} y(x,h)\,\mathrm{d}x \tag{B6}$$

where $x_M(h)$ is such that $h = d - a\cos[2\pi(x_M(h)/\lambda_x) + \varphi_x]$. Differentiation of (B6) and a few transformations lead us to

$$\frac{\partial S}{\partial h} = \frac{4\lambda_y}{\pi a}\int_{-(\varphi_x/2\pi)\lambda_x}^{x_M(h)} \left[\cos^2\left(2\pi\frac{x}{\lambda_x} + \varphi_x\right) - \left(\frac{h-d}{a}\right)^2\right]^{-1/2} \mathrm{d}x \tag{B7}$$

which, by the variable change $\theta = 2\pi(x/\lambda_x) + \varphi_x$, may be turned into

$$\frac{\partial S}{\partial h} = \frac{2\lambda_x\lambda_y}{\pi^2 a|\sin\theta_M(h)|}\int_0^{\theta_M(h)} (1 - \sin^2\theta/\sin^2\theta_M(h))^{-1/2}\,\mathrm{d}\theta\,. \tag{B8}$$

This latter equation involves a complete elliptic integral of the first kind, and therefore we finally get

$$\frac{\partial S}{\partial h} = \frac{2\lambda_x\lambda_y}{\pi^2 a} K\left[\sin(\theta_M(h))\right]$$

or, making the elliptic integral K more explicit [35]

$$\frac{\partial S}{\partial h} = \frac{\lambda_x\lambda_y}{\pi a}\left\{1 + \sum_{n=1}^{+\infty}\left(\frac{(2n-1)!!}{2^n n!}\right)^2\left[1 - \left(\frac{h-d}{a}\right)^2\right]^n\right\}. \tag{B9}$$

Appendix C

Let us prove that the exponential decrease of the current density $J_t(h)$ results in the exponential decrease of the tunnelling current I_t as a function of the distance d between nominal surfaces.

Let us assume that

$$J_t(h) = J_t(0)\exp(-h/h_0)\,. \tag{C1}$$

For given sphere-to-plane distances d_1 and d_2, the tunnelling current is given by

$$I_{t(d_i)} = \int_{h_{\min(d_i)}}^{\infty} J_t(u)\left(\frac{\partial S}{\partial u}\right)_{(d_i)}(u)\,\mathrm{d}u \qquad (i = 1, 2)\,. \tag{C2}$$

It is clear that, from the $\partial S/\partial h$ distribution viewpoint, changing the distance d from d_1 to d_2 is equivalent to shifting the whole $\partial S/\partial h$ distribution $(d_2 - d_1)$ along the h-axis. Therefore

$$h_{\min(d_2)} = h_{\min(d_1)} + (d_2 - d_1) \tag{C3}$$

and

$$\left(\frac{\partial S}{\partial h}\right)_{(d_2)}(h + (d_2 - d_1)) = \left(\frac{\partial S}{\partial h}\right)_{(d_1)}(h)\,. \tag{C4}$$

Using (C3) and the variable change $v = u - \Delta d$, in which $\Delta d = d_2 - d_1$, $I_{t(d_2)}$ as expressed in (C2) may be written as

$$I_{t(d_2)} = \int_{h_{\min(d_1)}}^{\infty} J_t(v + \Delta d)\left(\frac{\partial S}{\partial v}\right)_{(d_2)}(v + \Delta d)\,\mathrm{d}v\,. \tag{C5}$$

From (C1) and (C4), this expression may be turned into

$$I_{t(d_2)} = \exp(-\Delta d/h_0)\int_{h_{\min(d_1)}}^{\infty} J_t(v)\left(\frac{\partial S}{\partial v}\right)_{(d_1)}(v)\,\mathrm{d}v \tag{C6}$$

i.e.

$$I_{t(d_2)} = \exp\bigl(-(d_2 - d_1)/h_0\bigr)\,I_{t(d_1)}\,. \tag{C7}$$

References

[1] Antler M 1963 *Wear* **6** 44
[2] Antler M 1987 *IEEE Trans. Comp. Hybrids Manuf. Technol.* **CHMT-10** 24
[3] Hsue E Y and Bayer R G 1989 *IEEE Trans. Comp. Hybrids Manuf. Technol.* **CHMT-12** 206
[4] Holm R 1967 *Electric Contacts* (Berlin: Springer)
[5] Reynaud C, Boiziau C, Juret C, Leroy S, Perreau J and Lécayon G 1985 *Synth. Met.* **11** 159
[6] Boiziau C and Lécayon G 1988 *La Recherche* **19** 888
[7] Périé J, Périé M, Boyer L and Noël S 1988 *Wear* **128** 153
[8] Tersoff J and Hamann D R 1983 *Phys. Rev. Lett.* **50** 1998
[9] Garcia N, Ocal C and Flores F 1983 *Phys. Rev. Lett.* **50** 2002
[10] Baratoff A 1984 *Physica* **127B** 143
[11] Tersoff J and Hamann D R 1985 *Phys. Rev.* B **31** 805
[12] Sacks W, Gautier S, Rousset S, Klein J and Esrick M A 1987 *Phys. Rev.* B **36** 961
[13] Duke C B 1969 *Tunnelling in Solids* (New York: Academic) and references therein

[14] Teague E C 1986 *J. Res. NBS* **91** 171
[15] Greenwood J A, Johnson K L and Matsubara E 1984 *Wear* **100** 47
[16] Thomas T R 1982 *Rough Surfaces* (London: Longman)
[17] Tonck A 1989 private communication
[18] Mandelbrot B B 1982 *The Fractal Geometry of Nature* (New York: Freeman)
[19] Gagnepain J J and Roques-Carmes C 1986 *Wear* **109** 119
[20] Peitgen H O and Saupe D 1988 *The Science of Fractal Images* (Berlin: Springer)
[21] Majumdar A and Bhushan B 1990 *ASME J. Tribol.* **112** 205
[22] Feder J 1988 *Fractals* (New York: Plenum)
[23] Simmons J G 1969 in *Tunnelling Phenomena in Solids* ed E Burstein and S Lundqvist (London: Academic) pp 135–48
[24] Tonck A, Georges J M and Loubet J L 1988 *J. Coll. Interf. Sci.* **126** 150
[25] Montfort J P, Tonck A, Loubet J L and Georges J M 1991 *J. Polym. Sci.* B **29**
[26] Houzé F 1990 *Thèse de Doctorat*, Université Paris VI
[27] Reiss G, Schneider F, Vancea J and Hoffmann H 1990 *Appl. Phys. Lett.* **57** 867
[28] Tonck A, Houzé F, Boyer L, Loubet J L and Georges J M 1991 *J. Phys.: Condens. Matter* submitted
[29] Habib S E D and Simmons J G 1980 *Solid-State Electron.* **23** 87
[30] Franz W 1956 *Handbuch der Physik* vol 17 ed S Flügge (Berlin: Springer) pp 155–264
[31] Stratton R, Lewicki G and Mead C A 1966 *J. Phys. Chem. Solids* **27** 1599
[32] Christov S G 1972 *Contemp. Phys.* **13** 199
[33] Gundlach K H 1973 *J. Appl. Phys.* **44** 5005
[34] Israelachvili J N 1985 *Intermolecular and Surface Forces* (London: Academic)
[35] Gradshteyn I S and Ryzhik I M 1965 *Table of Integrals, Series and Products* (London: Academic)