

Surface topography in scanning tunneling microscopy: A free-electron model

W. Sacks, S. Gauthier, S. Rousset, and J. Klein

Groupe de Physique des Solides de l'Ecole Normale Supérieure, Université Paris VII Tour 23, 2 place Jussieu, 75251 Paris Cédex 05, France

M. A. Esrick

Physics Department, Loyola College in Maryland, Baltimore, Maryland 21210

(Received 16 December 1986)

The topographic image as given by scanning tunneling microscopy (STM) is deduced in analytic form in a free-electron, or Sommerfeld model. The method is non-numerical and employs perturbed wave functions for an arbitrarily modified plane metal surface to approximate the local density of states (LDOS), at the Fermi level. The curves of constant LDOS, hence also the contours followed by the probe in an s -wave tip model, are calculated in terms of $h(\mathbf{x}_{\parallel})$, a surface profile function. The image of an arbitrary periodic or nonperiodic surface structure is determined by contours of the form $z(\mathbf{x}_{\parallel}) = \bar{z} + \Delta(\mathbf{x}_{\parallel}, \bar{z})$ where \bar{z} is the average probe-surface separation, and $\Delta(\mathbf{x}_{\parallel}, \bar{z})$ is a convolution over $h(\mathbf{x}_{\parallel})$. We also discuss the parallel and perpendicular resolution of surface structures such as a one- or two-dimensional Gaussian, a perfect step, and a cosine surface, as a function of distance and tip radius. We find there is considerable smoothing of the image in STM for finite surface defects for typical tip-surface separations and tip radii.

I. INTRODUCTION

The work of Tersoff and Hamann^{1,2} provides both a fundamental understanding of scanning tunneling microscopy (STM)³ at low voltages and, in particular, a detailed discussion of the measured surface topography in the constant-current mode. At low voltages, as is well known, the current is Ohmic and depends on tunneling between states at or near the Fermi level of each electrode. In particular, Tersoff and Hamann (TH) find the tunneling conductance for STM to be proportional to the local density of states (LDOS) of the surface, at the Fermi level.

This result is obtained by applying Bardeen's formalism⁴ to the tunneling problem in the low-bias limit and relies on two basic assumptions. Firstly, the tip wave function is modeled by an $l=0$ state external to a spherical potential well of radius R . Secondly, it is assumed that the effect of the image potential in the gap can be neglected. In addition, the Bardeen method requires that the overlap between tip and surface wave functions be small. While the tip is given a model form, the surface wave function remains quite general and thus can be calculated independently. For the tunneling conductance ($\sigma = I/V$), the authors find,

$$\sigma \simeq 0.1 R^2 e^{2kR} \rho(\mathbf{r}, E_F) \text{ ohms}^{-1}, \quad (1)$$

with

$$\rho(\mathbf{r}, E_F) = \sum_n |\psi_n(\mathbf{r})|^2 \delta(E_n - E_F), \quad (2)$$

where \mathbf{r} is the location of the center of curvature of the tip, a distance $R + d$ from the surface.

It follows from (1) that the curves of constant conductance, in the s -wave tip model, are identical to the curves of constant LDOS, defined by Eq. (2). For this judicious choice of tip wave function, i.e., s wave, the density of states of the tip is a multiplicative factor and is not convoluted with the LDOS of the surface, as one might expect. The further work by Feuchtwang *et al.*⁵ and the calculations of Lang⁶ on tunneling between two adsorbates across a vacuum gap tend to confirm that the surface topography, as measured by STM, is close to a contour of the surface LDOS.

An alternative approach to STM is provided by Garcia *et al.*^{7,8} and by Stoll *et al.*⁹ in which an analogy is drawn between tunneling electrons at nonplanar metal surfaces and scattering theory. The electrodes are treated as free-electron metals with a periodically corrugated abrupt potential step at each interface. In representing the tip electrode by a periodic structure, whose repeat distance is large compared to the surface lattice constant, it is implicitly assumed that the current between a single probe and surface is reproduced. One drawback in this method is that it remains essentially numerical, while employing a simple model for both tip and surface. However, in Ref. 10 Stoll derives an approximate expression for the tunneling current at a free-electron metal with a weakly sinusoidal boundary. The resolution for this particular case is also discussed.

In spite of the improved knowledge of the tip structure^{11,12} that has been obtained since TH's work, we feel that model calculations of the STM system are instructive (see, for example, the discussions in Refs. 13–15). For this reason, in this work we retain the free-electron model, as in Refs. 7–10, but calculate $\rho(\mathbf{r}, E_F)$ in a perturbative expansion in terms of $h(\mathbf{x}) = h(x, y)$, the surface profile

function. This is achieved, in Sec. II, by first calculating $\psi_n^{(1)}$, the first-order correction to the perfectly-plane-surface wave function $\psi_n^{(0)}$, which is known. We then apply Eqs. (1) and (2) to obtain the tunneling conductance (Sec. III). The expression obtained, valid for large distances compared to the inverse decay length of the surface electron wave function, has the advantage of being easily solvable for $z(\mathbf{x})$, the position of the probe in the constant-conductance mode. In Sec. IV, we derive the general form

$$z(\mathbf{x}) = \bar{z} + \Delta(\mathbf{x}, \bar{z}), \quad (3)$$

where

$$\Delta(\mathbf{x}, \bar{z}) = \int d^2x' h(\mathbf{x}') f(\mathbf{x} - \mathbf{x}', \bar{z}),$$

i.e., a convolution over the surface, and \bar{z} is the average distance between tip center and the surface. As Eq. (3) can be determined for any $h(\mathbf{x})$, it provides a general solution for the surface topography at a distance \bar{z} from a periodic, or nonperiodic, surface. Finally, in Sec. V, we discuss and compare Eq. (3) for a variety of possible surface structures such as a cosine, a one- or two-dimensional Gaussian, and a single perfect step. The resolution of such structures, both in the parallel and perpendicular directions, as a function of distance from the surface and radius R of the probe, is also discussed.

II. CALCULATION OF THE WAVE FUNCTION

Let ψ_n be the wave function of an electron in the presence of a nonplanar step barrier potential of the form $V(\mathbf{x}, z) = -V_0 \Theta(h(\mathbf{x}) - z)$, where $h(\mathbf{x})$ is the surface profile, and $\Theta(x)$ is the usual Heaviside step function. The equation to be solved is:

$$\left[-\frac{\hbar^2 \nabla^2}{2m} + V(\mathbf{x}, z) \right] \psi_n = E_n \psi_n, \quad (4)$$

where ψ_n is subject to the usual boundary conditions at $z = h(\mathbf{x})$. For $z \geq h(\mathbf{x})$ the solution to (4) is that of a free particle with an abrupt barrier of height V_0 at $z = h(\mathbf{x})$. Assuming that $h(\mathbf{x})$ is periodic, ψ_n can be expanded in the general form:

$$\psi_n^> = \frac{1}{\sqrt{A}} e^{i\mathbf{q} \cdot \mathbf{x}} c_{\mathbf{q}}^{(0)} e^{-\alpha_{\mathbf{q}} z} + \frac{1}{\sqrt{A}} \sum_{\mathbf{G}} a_{\mathbf{G}}^{(1)} e^{i(\mathbf{q} + \mathbf{G}) \cdot \mathbf{x}} e^{-\alpha_{\mathbf{q} + \mathbf{G}} z}, \quad (8a)$$

$$\psi_n^< = \frac{1}{\sqrt{A}} e^{i\mathbf{q} \cdot \mathbf{x}} c_{\mathbf{q}}^{(0)} \frac{\cos[\beta_{\mathbf{q}} z - \delta(\mathbf{q})]}{\cos \delta(\mathbf{q})} + \frac{1}{\sqrt{A}} \sum_{\mathbf{G}} e^{i(\mathbf{q} + \mathbf{G}) \cdot \mathbf{x}} (b_{\mathbf{G}}^{(1)} e^{i\beta_{\mathbf{q} + \mathbf{G}} z} + c_{\mathbf{G}}^{(1)} e^{-i\beta_{\mathbf{q} + \mathbf{G}} z}), \quad (8b)$$

where $c_{\mathbf{q}}^{(0)} = B e^{i\delta(\mathbf{q})} \cos \delta(\mathbf{q})$. $\delta(\mathbf{q})$ is an energy-dependent phase due to the step potential, defined by $\tan \delta(\mathbf{q}) = -\alpha_{\mathbf{q}} / \beta_{\mathbf{q}}$, and B is a normalization constant.

Two of the three coefficients, $a_{\mathbf{G}}^{(1)}$, $b_{\mathbf{G}}^{(1)}$, and $c_{\mathbf{G}}^{(1)}$ are determined using the boundary conditions (6b). Solving

$$\psi_n^> = \frac{1}{\sqrt{A}} \sum_{\mathbf{G}} a_{\mathbf{G}} e^{i(\mathbf{q} + \mathbf{G}) \cdot \mathbf{x}} e^{-\alpha_{\mathbf{q} + \mathbf{G}} z}, \quad (5a)$$

$$\psi_n^< = \frac{1}{\sqrt{A}} \sum_{\mathbf{G}} e^{i(\mathbf{q} + \mathbf{G}) \cdot \mathbf{x}} (b_{\mathbf{G}} e^{i\beta_{\mathbf{q} + \mathbf{G}} z} + c_{\mathbf{G}} e^{-i\beta_{\mathbf{q} + \mathbf{G}} z}), \quad (5b)$$

where

$$\alpha_{\mathbf{q}} = \left[q^2 - \frac{2mE}{\hbar^2} \right]^{1/2}, \quad \beta_{\mathbf{q}} = \left[\frac{2m}{\hbar^2} (E + V_0) - q^2 \right]^{1/2}$$

and n stands for (\mathbf{q}, E) .

More generally, $h(\mathbf{x})$ is nonperiodic and the sum over \mathbf{G} is replaced by a Fourier integral. Although the present problem could be solved numerically, an analytic expression for ψ_n can be determined perturbatively in orders to $h(\mathbf{x})$. Thus, both expanding $\psi_n = \psi_n^{(0)} + \psi_n^{(1)} + \dots$ in the usual fashion, and further doing a Taylor-series expansion about $z=0$, i.e., $\psi_n^{(n)}|_{z=h} = \psi_n^{(n)}|_{z=0} + h \psi_n^{(n)}|_{z=0} + \dots$, the boundary conditions are (order by order)

$$(\psi_n^{(0)<} - \psi_n^{(0)>})_{z=0} = 0, \quad (6a)$$

$$\left[\frac{\partial}{\partial z} \psi_n^{(0)<} - \frac{\partial}{\partial z} \psi_n^{(0)>} \right]_{z=0} = 0,$$

$$(\psi_n^{(1)<} - \psi_n^{(1)>})_{z=0} = 0, \quad (6b)$$

$$\left[\frac{\partial}{\partial z} \psi_n^{(1)<} - \frac{\partial}{\partial z} \psi_n^{(1)>} \right]_{z=0} = -h(\mathbf{x}) \left[\frac{\partial^2}{\partial z^2} \psi_n^{(0)<} - \frac{\partial^2}{\partial z^2} \psi_n^{(0)>} \right]_{z=0}.$$

In addition we expand the coefficients in (5a) and (5b)

$$a_{\mathbf{G}} = a_{\mathbf{G}}^{(0)} + a_{\mathbf{G}}^{(1)} + \dots \quad (7)$$

and likewise for $b_{\mathbf{G}}$ and $c_{\mathbf{G}}$. The energy E is not expanded as, by inspection of Eq. (4), together with the explicit form of the potential, E is independent of $h(\mathbf{x})$. The coefficients $a_{\mathbf{G}}^{(0)}$, $b_{\mathbf{G}}^{(0)}$, and $c_{\mathbf{G}}^{(0)}$, which are determined up to a normalization constant using the boundary conditions (6a), are simply those corresponding to the plane-surface wave function $\psi_n^{(0)}$. Hence through first order ($\psi_n = \psi_n^{(0)} + \psi_n^{(1)}$):

for $b_{\mathbf{G}}^{(1)}$ and $c_{\mathbf{G}}^{(1)}$ in terms of $a_{\mathbf{G}}^{(1)}$, the solution is

$$c_{\mathbf{G}}^{(1)} = \frac{(i\beta_{\mathbf{q} + \mathbf{G}} + \alpha_{\mathbf{q} + \mathbf{G}})}{2i\beta_{\mathbf{q} + \mathbf{G}}} a_{\mathbf{G}}^{(1)} - \frac{c_{\mathbf{q}}^{(0)}}{2i\beta_{\mathbf{q} + \mathbf{G}}} h_{\mathbf{G}} (2mV_0/\hbar^2), \quad (9a)$$

$$b_{\mathbf{G}}^{(1)} = \frac{(i\beta_{\mathbf{q}+\mathbf{G}} - \alpha_{\mathbf{q}+\mathbf{G}})}{2i\beta_{\mathbf{q}+\mathbf{G}}} a_{\mathbf{G}}^{(1)} + \frac{c_{\mathbf{q}}^{(0)}}{2i\beta_{\mathbf{q}+\mathbf{G}}} h_{\mathbf{G}} (2mV_0/\hbar^2), \quad (9b)$$

where $h_{\mathbf{G}} = (1/a) \int d^2x e^{-i\mathbf{G}\cdot\mathbf{x}} h(\mathbf{x})$. This leaves $a_{\mathbf{G}}^{(1)}$ to be

$$\psi_n^{(1)>} = \frac{c_{\mathbf{q}}^{(0)}}{\sqrt{A}} \sum_{\mathbf{G}} h_{\mathbf{G}} e^{i(\mathbf{q}+\mathbf{G})\cdot\mathbf{x}} \alpha_{\mathbf{q}+\mathbf{G}} e^{-\alpha_{\mathbf{q}+\mathbf{G}} z}, \quad (10a)$$

$$\psi_n^{(1)<} = \frac{c_{\mathbf{q}}^{(0)}}{\sqrt{A}} \sum_{\mathbf{G}} h_{\mathbf{G}} e^{i(\mathbf{q}+\mathbf{G})\cdot\mathbf{x}} \left[\frac{\alpha_{\mathbf{q}+\mathbf{G}}}{\cos\delta(\mathbf{q}+\mathbf{G})} \cos[\beta_{\mathbf{q}+\mathbf{G}} z - \delta(\mathbf{q}+\mathbf{G})] + \frac{2mV_0}{\hbar^2 \beta_{\mathbf{q}+\mathbf{G}}} \sin(\beta_{\mathbf{q}+\mathbf{G}} z) \right]. \quad (10b)$$

As a simple check, if $h(\mathbf{x})=h$, i.e., a constant shift of the plane surface in the z direction, then $h_{\mathbf{G}}=h\delta_{\mathbf{G},0}$. Equation (10) then reduces to $\psi_n^{(1)} = -h\partial\psi_n^{(0)}/\partial z$ which is the first-order term of the exact solution $\psi_n^{(0)}(z-h)$. The general result for $\psi_n^{(1)}$ in Eq. (10) can equivalently be derived using the alternative technique of perturbation theory on the boundary condition,¹⁶ as applied in Ref. 17 to surface states. In principle, the perturbation expansion can be carried out to any desired order.

III. CALCULATION OF THE LDOS

As we have calculated ψ_n through first order, we proceed to derive an expression for the LDOS at the Fermi level also to first order. Expanding ψ_n in Eq. (3) in terms of $h(\mathbf{x})$ we have

$$\rho^{(0)}(\mathbf{r}, E_F) = \sum_n |\psi_n^{(0)}|^2 \delta(E_n - E_F), \quad (11a)$$

$$\rho^{(1)}(\mathbf{r}, E_F) = \sum_n \psi_n^{(0)*} \psi_n^{(1)} \delta(E_n - E_F) + \text{c.c.} \quad (11b)$$

$\rho^{(0)}$ and $\rho^{(1)}$ are particularly easy to calculate asymptotically for $\kappa z \gg 1$. This limit is consistent, however, with the simplified choice for the potential, and the original use of the Bardeen formalism for the tunneling current.

From Sec. II, the relevant solution for ψ_n is

$$\psi_n^{(0)>} = \frac{1}{\sqrt{A}} e^{i\mathbf{q}\cdot\mathbf{x}} c_{\mathbf{q}}^{(0)} e^{-\alpha_{\mathbf{q}} z}, \quad (12a)$$

$$\psi_n^{(1)>} = \frac{1}{\sqrt{A}} c_{\mathbf{q}}^{(0)} \sum_{\mathbf{G}} h_{\mathbf{G}} e^{i(\mathbf{q}+\mathbf{G})\cdot\mathbf{x}} \alpha_{\mathbf{q}+\mathbf{G}} e^{-\alpha_{\mathbf{q}+\mathbf{G}} z}, \quad (12b)$$

and $\rho(\mathbf{r}, E_F)$ for the flat-surface LDOS is

$$\rho^{(0)}(z, E_F) = C \int_{q < q_F} d^2q \frac{1}{(q_F^2 - q^2)^{1/2}} \cos^2\delta(\mathbf{q}) e^{-2\alpha_{\mathbf{q}} z}, \quad (13)$$

where $\alpha_{\mathbf{q}} = (q^2 + \kappa^2)^{1/2}$, $q_F^2 = 2m(E_F + V_0)/\hbar^2$, and C is $m/\pi^3\hbar^2$. The factor $(q_F^2 - q^2)^{-1/2} = \beta_{\mathbf{q}}^{-1}$ arises from the one-dimensional density of states evaluated at $E_F - \hbar^2 q^2/2m$. For large κz , the integral in (13) can be approximated by expanding the argument of the exponential about $\mathbf{q}=0$ and extending the upper limit to infinity. As the exponential decay dominates, the prefactor can be evaluated at $\mathbf{q}=0$,

$$\rho^{(0)}(z, E_F) \simeq C \cos^2\delta(0) \frac{\pi\kappa}{q_F} \frac{e^{-2\kappa z}}{z}. \quad (14)$$

The tunneling conductance for a sphere and perfect plane, using Eq. (1), is then

determined by the orthonormalization of $\{\psi_n\}$, which imposes a further condition on $\psi_n^{(1)}$. One finds $a_{\mathbf{G}}^{(1)} = c_{\mathbf{q}}^{(0)} h_{\mathbf{G}} \alpha_{\mathbf{q}+\mathbf{G}}$, and the final result for $\psi_n^{(1)}$ is

$$\sigma = \text{const} \times \frac{R^2}{R+d} e^{-2\kappa d}, \quad (15)$$

where we have used $z=R+d$. To lowest order in d/R , the conductance is linear in the tip radius, $\sigma \propto R e^{-2\kappa d}$, and not quadratic. On the other hand, if $d \gg R$, $\sigma \propto R^2 e^{-2\kappa d}/d$, and a factor of d appears in the denominator.¹⁸

In a similar manner, we calculate $\rho^{(1)}$ given by (11b) using the expressions for both $\psi^{(0)}$ and $\psi^{(1)}$ in Eq. (12). We obtain

$$\rho^{(1)}(\mathbf{r}, E_F) = 2C \sum_{\mathbf{G}} h_{\mathbf{G}} e^{i\mathbf{G}\cdot\mathbf{x}} \int_{q < q_F} d^2q \frac{\cos^2\delta(\mathbf{q})}{(q_F^2 - q^2)^{1/2}} \alpha_{\mathbf{q}+\mathbf{G}} \times e^{-(\alpha_{\mathbf{q}} + \alpha_{\mathbf{q}+\mathbf{G}}) z} \Theta(q_F - |\mathbf{q} + \mathbf{G}|), \quad (16)$$

where the step function ensures the proper cutoff in the sum. $\rho^{(1)}$, in contrast to $\rho^{(n)}$, contains a single Fourier sum, and therefore a first-order calculation is considerably simplified. An approximate expression for $\rho^{(1)}$ is likewise obtained for $\kappa z \gg 1$ by expanding the exponent about $\mathbf{q} = -\mathbf{G}/2$ and extending the integration to ∞

$$\rho^{(1)}(\mathbf{r}, E_F) \simeq 2C \sum_{\mathbf{G}} h_{\mathbf{G}} e^{i\mathbf{G}\cdot\mathbf{x}} \frac{\cos^2\delta(\mathbf{G}/2)}{\beta_{\mathbf{G}/2}} \times \frac{\pi\alpha_{\mathbf{G}/2}^2}{z} e^{-2\alpha_{\mathbf{G}/2} z} \left[1 + \frac{1}{2\alpha_{\mathbf{G}/2} z} \right], \quad (17)$$

where the sum is cut off at $|\mathbf{G}| = 2q_F$. Putting together (14) and (17) we obtain for the LDOS,

$$\rho(\mathbf{r}, E_F) \simeq A \left[\frac{e^{-2\kappa z}}{z} + 2 \sum_{\mathbf{G}} h_{\mathbf{G}} e^{i\mathbf{G}\cdot\mathbf{x}} \frac{\beta_{\mathbf{G}/2} \alpha_{\mathbf{G}/2}^2}{q_F \kappa} \times e^{-2\alpha_{\mathbf{G}/2} z} \left[1 + \frac{1}{2\alpha_{\mathbf{G}/2} z} \right] \right], \quad (18)$$

where A is a constant. Again, (18) is exact to first order for a ‘‘raised’’ surface. Indeed, if we substitute $h_{\mathbf{G}} = h\delta_{\mathbf{G},0}$:

$$\rho(z, E_F) = A \left[\frac{e^{-2\kappa z}}{z} + 2\kappa h \frac{e^{-2\kappa z}}{z} \left[1 + \frac{1}{2\kappa z} \right] \right], \quad (19)$$

which is precisely $\rho^{(0)}(z) - h\partial\rho^{(0)}(z)/\partial z$ the desired result.

Using (18) the tunneling conductance, consistent through first order, is

$$\sigma(\mathbf{r}) = \text{const} \times R^2 e^{2\kappa R} \frac{e^{-2\kappa[z - \Delta(\mathbf{x}, z)]}}{(z - \Delta(\mathbf{x}, z))}, \quad (20)$$

where $\Delta(\mathbf{x}, z)$ is defined by

$$\begin{aligned} \Delta(\mathbf{x}, z) &= \sum_{\mathbf{G}} h_{\mathbf{G}} e^{i\mathbf{G} \cdot \mathbf{x}} \frac{\beta_{\mathbf{G}/2} \alpha_{\mathbf{G}/2}^2}{q_F \kappa^2} \\ &\quad \times e^{-2(\alpha_{\mathbf{G}/2} - \kappa)z} \left[\frac{1 + 1/2\alpha_{\mathbf{G}/2} z}{1 + 1/2\kappa z} \right] \\ &\simeq \sum_{\mathbf{G}} h_{\mathbf{G}} e^{i\mathbf{G} \cdot \mathbf{x}} \frac{\beta_{\mathbf{G}/2} \alpha_{\mathbf{G}/2}^2}{q_F \kappa^2} e^{-2(\alpha_{\mathbf{G}/2} - \kappa)z}, \end{aligned} \quad (21)$$

neglecting the term in parentheses for $\kappa z \gg 1$. The form given in (20) for the conductance is *better* than a first-order calculation since, for example, it is exact for the raised surface, for which $\Delta(\mathbf{x}, z) = h$, and (20) reduces to

$$\sigma(z) = \text{const} \times R^2 e^{2\kappa R} \frac{e^{-2\kappa(z-h)}}{(z-h)}.$$

However, Δ is still a first-order correction to z in the argument of the exponential.

$\Delta(\mathbf{x}, z)$ can further be expressed in the form of a convolution:

$$\Delta(\mathbf{x}, z) = \int d^2 \mathbf{x}' h(\mathbf{x}') f(\mathbf{x} - \mathbf{x}', z), \quad (22)$$

where $f(\mathbf{x}, z)$ is a function that smooths the surface geometry. By inspection of Eqs. (20) and (21) we see that the LDOS, or the conductance, is very nearly flat with $\Delta(\mathbf{x}, z)$ defining the relevant modulation. For example, as $z \rightarrow \infty$, we only pick up the $\mathbf{G} = \mathbf{0}$ Fourier component in (21), i.e., $\Delta \approx h_0 = \langle h \rangle_{\text{av}}$. The conductance has no \mathbf{x} dependence,

$$\sigma \propto \frac{e^{-2\kappa(z - \langle h \rangle_{\text{av}})}}{z - \langle h \rangle_{\text{av}}}, \quad (23)$$

and a mere readjustment of the probe-surface distance results. The next term in (21) is the \mathbf{G}_1 coefficient corresponding to a corrugation of the largest wavelength. Each subsequent Fourier component is further damped by the exponential, as also noted by Tersoff and Hamann in Ref. 2. A similar approximation for $\Delta(\mathbf{x}, z)$ is obtained for nonperiodic surfaces:

$$\Delta(\mathbf{x}, z) \simeq \int_{q < 2q_F} d^2 q h(\mathbf{q}) e^{i\mathbf{q} \cdot \mathbf{x}} \frac{\beta_{\mathbf{q}/2} \alpha_{\mathbf{q}/2}^2}{q_F \kappa^2} e^{-2(\alpha_{\mathbf{q}/2} - \kappa)z}, \quad (24)$$

with

$$h(\mathbf{q}) = \int d^2 \mathbf{x} e^{-i\mathbf{q} \cdot \mathbf{x}} h(\mathbf{x}).$$

IV. STM IMAGE

The topographic image as seen using STM is given by the contours of the surface $z(\mathbf{x})$ that maintains constant conductance. Setting $\sigma(\mathbf{r}) = \text{const}$ in Eq. (20) leads to the expression:

$$z - \Delta(\mathbf{x}, z) = \text{const}, \quad (25)$$

which we can solve for $z(\mathbf{x})$ through first order. The above constant is obtained by averaging over the surface [over a unit cell if $h(\mathbf{x})$ is periodic or over the entire surface if $h(\mathbf{x})$ is nonperiodic]. Using the fact that $\langle \Delta \rangle_{\text{av}} = h_0 = \langle h \rangle_{\text{av}}$, gives $\text{const} = \langle z \rangle_{\text{av}} - \langle h \rangle_{\text{av}} \equiv \bar{z}$. The latter constant is thus the average distance from the center of the probe to the *actual* surface. The equation to be solved is then

$$z = \bar{z} + \Delta(\mathbf{x}, z). \quad (26)$$

It is clear that for a first-order expression for $z(\mathbf{x})$ a first iteration of z in (26) is sufficient, as Δ is first order in h . A single iteration gives $\bar{z} + \Delta(\mathbf{x}, z) \approx \bar{z} + \Delta(\mathbf{x}, \bar{z})$ and the probe position is then found to be

$$z(\mathbf{x}) = \bar{z} + \Delta(\mathbf{x}, \bar{z}), \quad (27)$$

where Δ , defined in Eqs. (21) and (24), is evaluated at \bar{z} , a quantity independent of the original coordinate system chosen to define $\psi_n^{(0)}$.

The probe thus follows a perturbed path approximately a distance \bar{z} from the surface. The deviation about this path, given by $\Delta(\mathbf{x}, \bar{z})$, can be calculated for an arbitrary $h(\mathbf{x})$ which we illustrate in the following section. The image a distance $\bar{z} = R + \langle d \rangle$ from the surface is a damped and smoothed version of $h(\mathbf{x})$ and, as \bar{z} increases, the image of an arbitrary surface structure becomes increasingly flat ($z \sim \bar{z}$). By combining Eqs. (27) and (22), an alternative expression can be found:

$$z(\mathbf{x}) = \bar{z} + \int d^2 \mathbf{x}' h(\mathbf{x}') f(\mathbf{x} - \mathbf{x}', \bar{z}). \quad (28)$$

One can think of the ideal image as being $h(\mathbf{x})$, while $f(\mathbf{x})$ is an independent resolution function. More specifically, one can show that, for $\kappa z \gg 1$, $f(\mathbf{x}) \approx (\kappa/\pi\bar{z}) \exp(-x^2\kappa/\bar{z})$, i.e., a Gaussian with full width at half maximum $2(\ln 2\bar{z}/\kappa)^{1/2}$ which characterizes the lateral resolution. The nominal resolution thus diminishes rapidly for large tip radii (R), or large tip-surface separations ($\langle d \rangle$), as discussed in Ref. 19.

In the following section we make use of two characteristic parameters Δ_{\perp} and Δ_{\parallel} of the actual image $z(\mathbf{x})$. We define Δ_{\perp} and Δ_{\parallel} as the response of the tip at a distance \bar{z} such that Δ_{\perp} is the maximum change in $\Delta(\mathbf{x}, \bar{z})$ and Δ_{\parallel} is the parallel distance over which Δ_{\perp} is defined. For example, if $\Delta(\mathbf{x}, \bar{z})$ is a Gaussian, then Δ_{\perp} and Δ_{\parallel} are the amplitude and width, respectively. These two characteristic lengths are related to an alternative notion of "resolution": the ability to image a given surface (not to be confused with the nominal resolution discussed above). For example, Δ_{\perp} is related to the resolution by requiring that the signal-to-noise ratio, Δ_{\perp}/n , be greater than one (as in Ref. 19). If the STM system noise is larger than $\Delta(\bar{z})$, then no real resolution of the surface occurs. In addition, the parameter Δ_{\parallel} indicates the characteristic width or broadening of the image as a function of \bar{z} . As the following examples illustrate, Δ_{\perp} and Δ_{\parallel} are dependent on the surface geometry and also dependent on the lateral resolution: $2(\ln 2\bar{z}/\kappa)^{1/2}$.

V. EXAMPLES

In this section we consider a few examples of the surface geometry to illustrate the use of Eq. (27).

a. Cosine surface. Choosing $h(\mathbf{x}) = h \cos(\mathbf{G}_1 \cdot \mathbf{x})$, Eq. (27) then gives (provided $G_1 < 2q_F$)

$$z(\mathbf{x}) = \bar{z} + h \cos(\mathbf{G}_1 \cdot \mathbf{x}) \frac{\beta_{G_1/2} \alpha_{G_1/2}^2}{a_F \kappa^2} e^{-2(\alpha_{G_1/2} - \kappa)\bar{z}}. \quad (29)$$

These contours are illustrated in Fig. 1 for various distances \bar{z} from the surface. The probe thus follows the surface with a \bar{z} -dependent amplitude:

$$A = h \frac{\beta_{G_1/2} \alpha_{G_1/2}^2}{q_F \kappa^2} e^{-2(\alpha_{G_1/2} - \kappa)\bar{z}}. \quad (30)$$

Tersoff and Hamann² obtained the identical exponential dependence, i.e., $A \propto e^{-\beta z}$. They further suggest that in the limit of a small tip radius $A \approx 2\kappa^{-1} e^{-\beta z}$.

In the small- $G_1/2\kappa$ limit, we obtain the very simple result

$$A \simeq h e^{-G_1^2 \bar{z} / 4\kappa}, \quad (31)$$

having used $\beta_{G_1/2} \alpha_{G_1/2}^2 \approx q_F \kappa^2$ and $\alpha_{G_1/2} - \kappa \approx G_1^2 / 8\kappa$. For the cosine surface, the perpendicular resolution is twice the amplitude

$$\Delta_{\perp} \simeq 2h e^{-G_1^2 \bar{z} / 4\kappa}, \quad (32)$$

which coincides with the expression of Stoll *et al.*^{9,10} $\Delta_{\parallel} = \pi/G_1$ is independent of \bar{z} , reflecting the fact that at any given distance (including R) such that Δ_{\perp} is larger than the system noise, or $(\Delta_{\perp}/n) > 1$, the probe feels the lateral periodicity of the surface.

b. Stepped surface. $h(\mathbf{x}) = h\Theta(x)$ defines a step of height h whose edge is parallel to the y axis. Here we assume that $\kappa\bar{z} \approx q_F \bar{z} \gg 1$ such that we can evaluate an approximate form for $\Delta(\mathbf{x}, \bar{z})$ in Eq. (24)

$$\Delta(\mathbf{x}, \bar{z}) \simeq \int \frac{d^2 q}{(2\pi)^2} h(\mathbf{q}) e^{iq \cdot \mathbf{x}} e^{-\gamma q^2}, \quad (33)$$

where $\gamma = \bar{z}/4\kappa$. Due to the exponential, the q dependence of the prefactor is small. Expanding for small $q/2\kappa \approx q/2q_F$ there are corrections to the integrand of (33) of order $(q/2\kappa)^2$, which can be neglected for the present purposes of illustration.

For the step

$$h(\mathbf{q}) = 2\pi \delta(q_y) h [\pi \delta(q_x) + P(1/iq_x)],$$

where P is the Cauchy principal value. Using this to evaluate (33) the probe position becomes:

$$z(x) = \bar{z} + \frac{h}{2} [1 + \Phi(x/2\sqrt{\gamma})], \quad (34)$$

where $\Phi(y) = \text{erf}(y)$ is the error function.²⁰

At the surface ($\bar{z} = 0$) we have a sharp step, while as a function of \bar{z} the step becomes progressively smoother. Indeed, for $|x|/2\sqrt{\gamma} \ll 1$, $\Phi(x/2\sqrt{\gamma}) \approx x/\sqrt{\pi\gamma}$ and in this limit:

$$z(x) \simeq \bar{z} + \frac{h}{2} (1 + x/\sqrt{\pi\gamma}), \quad |x|/2\sqrt{\gamma} \ll 1. \quad (35)$$

The slope of the tip path near $x=0$ is thus proportional to $(\bar{z})^{-1/2}$. In the opposite limit, i.e., $|x|/2\sqrt{\gamma} \gg 1$, $\Phi(x/2\sqrt{\gamma}) \approx \text{sgn}(x)$, and using $1 + \text{sgn}(x) = 2\Theta(x)$,

$$z(x) \simeq \bar{z} + h\Theta(x), \quad |x|/2\sqrt{\gamma} \gg 1. \quad (36)$$

The probe thus maps out a step of height h , as it is swept from large negative to positive values in x . However, as Fig. 2 illustrates, the step is considerably broadened. For the resolution of a step: $\Delta_{\perp} = h$, and is independent of \bar{z} , while Δ_{\parallel} can be estimated by noting that the derivative of $z(\mathbf{x})$ with respect to x is the Gaussian,

$$\frac{h}{2\sqrt{\pi\gamma}} e^{-x^2/4\gamma}, \quad (37)$$

and by taking $\Delta_{\parallel} = 2x_c$, where x_c reduces the derivative by a factor of e^{-2} ,

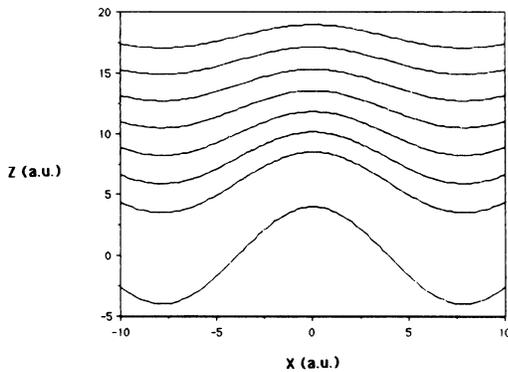


FIG. 1. Isoconductance traces for a cosine surface $h(\mathbf{x}) = h \cos(\mathbf{G}_1 \cdot \mathbf{x})$, along G_1 direction, for distances $\bar{z} = 6$ to 18 a.u. in steps of 2 a.u. ($G_1 = \pi/8$, $\kappa = \frac{1}{2}$ a.u.⁻¹), $\bar{z} = R + \langle d \rangle$, where R is the tip radius and $\langle d \rangle$ is the average tip-surface separation.

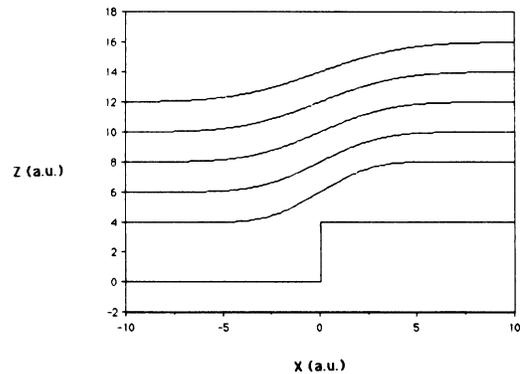


FIG. 2. Isoconductance traces along the x direction for a step of height 4 a.u. (see Table I). \bar{z} is chosen at 4, 6, 8, 10, and 12 a.u. ($\kappa = \frac{1}{2}$ a.u.⁻¹)

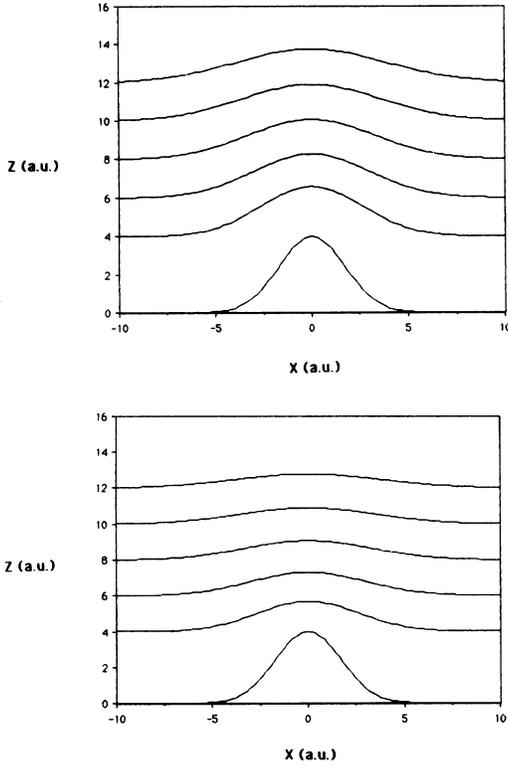


FIG. 3. Isoconductance traces along the x direction for (a) a 1D Gaussian and (b) a 2D Gaussian, both of amplitude 4 a.u. and width 4 a.u. \bar{z} is chosen identically to Fig. 2, and $\kappa = \frac{1}{2}$ a.u.⁻¹.

$$\Delta_{\parallel} = \left[\frac{8\bar{z}}{\kappa} \right]^{1/2}. \quad (38)$$

For example if $\kappa = \frac{1}{2}$ a.u.⁻¹, $\bar{z} = 8$ a.u., then $\Delta_{\parallel} \approx 11.3$ a.u. At this distance the step edge is already broadened to 6 Å laterally.

c. One- and two-dimensional Gaussians. Equation (33) can likewise be solved for a one-dimensional Gaussian $h(\mathbf{x}) = he^{-ax^2}$, or a two-dimensional Gaussian $h(\mathbf{x}) = he^{-ar^2}$, where $r^2 = x^2 + y^2$. The width in both

cases is $\Gamma = 2\sqrt{\ln 2}/\sqrt{a}$. For the two cases, respectively, the probe position is

$$z(x) = \bar{z} + \frac{h}{(1+4a\gamma)^{1/2}} e^{-ax^2/(1+4a\gamma)}, \quad (39a)$$

$$z(r) = \bar{z} + \frac{h}{(1+4a\gamma)} e^{-ar^2/(1+4a\gamma)}. \quad (39b)$$

As illustrated in Figs. 3(a) and 3(b), the probe also follows a Gaussian curve but of much smaller amplitude and larger width, as expressed above. For the perpendicular resolution, Δ_{\perp} , we have

$$\Delta_{\perp} = h/\sqrt{1+4a\gamma}$$

and

$$\Delta_{\perp} = h/(1+4a\gamma),$$

respectively, while Δ_{\parallel} is the same for both Gaussians,

$$\Delta_{\parallel} = \left[\Gamma^2 + \frac{4 \ln 2 \bar{z}}{\kappa} \right]^{1/2}, \quad (40)$$

and gives the broadening as a function of \bar{z} (see Table I). Taking, for example, $\bar{z} = 8$ a.u., $\kappa = \frac{1}{2}$ a.u.⁻¹, and $\Gamma = 4$ a.u., Δ_{\perp} is 0.51 h and 0.26 h , respectively. The amplitude is reduced by 49% for the 1D Gaussian as compared to 84% for the 2D Gaussian. This could be compared to 0% for the step, and indicates that it is easier to resolve a step edge (in the perpendicular direction) than a finite deformation of the surface. The lateral resolution of the Gaussian is comparable to that of the step, however,

VI. CONCLUSION

Our model indicates that, for simple metals, the STM probe moves along a contour $z(\mathbf{x})$ which is to first order a single convolution over the surface profile:

$$z(\mathbf{x}) = \bar{z} + \int d^2x' f(\mathbf{x} - \mathbf{x}', \bar{z}) h(\mathbf{x}'), \quad (41)$$

where \bar{z} is the average distance between tip center and surface. This simple result is not appropriate where band-structure effects^{14,21} or image-potential effects²² may become important. However the model could be extended to include relevant corrections.

TABLE I. The resolution of four simple surface geometries, defined by $h(\mathbf{x})$, is expressed in terms of $\Delta_{\perp}(\bar{z})$ and $\Delta_{\parallel}(\bar{z})$, where $\bar{z} = R + \langle d \rangle$ (tip radius plus average tip-surface separation).

Surface	$h(\mathbf{x})$	Δ_{\perp}	Δ_{\parallel}
Cosine	$h \cos(\mathbf{G}_1 \cdot \mathbf{x})$	$2he^{-G_1^2 \bar{z}/4\kappa}$	π/G_1
Step	$h\Theta(x)$	h	$\left[\frac{8\bar{z}}{\kappa} \right]^{1/2}$
1D Gaussian	he^{-ax^2}	$h / \left[1 + \frac{a\bar{z}}{\kappa} \right]^{1/2}$	$\left[\Gamma^2 + \frac{4 \ln 2 \bar{z}}{\kappa} \right]^{1/2}$
2D Gaussian	he^{-ar^2}	$h / \left[1 + \frac{a\bar{z}}{\kappa} \right]$	$\left[\Gamma^2 + \frac{4 \ln 2 \bar{z}}{\kappa} \right]^{1/2}$

In the limit that the corrugation is very flat, the above expression simplifies to $z \approx \bar{z} + h(\mathbf{x}) + (\bar{z}/4\kappa)\nabla^2 h(\mathbf{x}) + \dots$, which is obtained by expanding $f(\mathbf{x}-\mathbf{x}',\bar{z})$ about \mathbf{x} in (41). The first two terms are in agreement with Ref. 17, and in agreement with the conclusion that the tip follows the surface in this limit. We also note that Eq. (41) can be inverted giving $h(\mathbf{x})$ as a function of the probe displacement. Given an experimental $\delta z(\mathbf{x})$, one can in principle calculate an effective $h(\mathbf{x})$ provided an estimate of \bar{z} can be obtained.

An interesting feature of STM is the sample or geometry dependence of the resolution, as summarized in Table I. Provided Δ_{\perp} is larger than the experimental resolution, system noise, etc., the scan will detect a periodic surface of largest wavelength $2\pi/G_{\perp}$, or a step of height h , as in these examples Δ_{\parallel} and Δ_{\perp} are independent of \bar{z} , re-

spectively. This perhaps reflects the relative ease at imaging stepped surfaces with large terraces.²³ However the Δ_{\parallel} for the step indicates that the edge is considerably broadened ($\Delta_{\parallel} \gg \Delta_{\perp}$). The 1D Gaussian, having an infinite dimension along one axis, is easier to resolve than the 2D Gaussian, which represents a finite deformation of the surface. Indeed, for large \bar{z} we have $\Delta_{\perp} \sim h\sqrt{\kappa/\bar{z}}$ and $\Delta_{\parallel} \sim h(\kappa/\bar{z})$, respectively.

The wave functions present in Sec. II may have further applications. For example, we have calculated the tunneling conductance, and corresponding isoconductance contours, in the case of both an arbitrary probe and arbitrary surface geometry. This work, which includes a comparison between spherical and Gaussian tip electrodes, and the resulting probe position $z(\mathbf{x})$, will be presented in a future report.²⁴

¹J. Tersoff and D. R. Hamann, Phys. Rev. Lett. **50**, 1998 (1983).

²J. Tersoff and D. R. Hamann, Phys. Rev. B **31**, 805 (1985).

³G. Binnig, H. Rohrer, Ch. Gerber, and E. Weibel, Phys. Rev. Lett. **49**, 57 (1982).

⁴J. Bardeen, Phys. Rev. Lett. **6**, 57 (1961).

⁵T. E. Feuchtwang, P. H. Cutler, and N.M. Miskowsky, Phys. Lett. **99A**, 167 (1983); S. Chung, T. E. Feuchtwang, and P. H. Cutler, Surf. Sci. **181**, 412 (1987).

⁶N. D. Lang, Phys. Rev. Lett. **56**, 1164 (1986).

⁷N. Garcia, C. Ocal, and F. Flores, Phys. Rev. Lett. **50**, 2002 (1983).

⁸N. Garcia and F. Flores, Physica **127B**, 137 (1984).

⁹E. Stoll, A. Baratoff, A. Selloni, and P. Carnevali, J. Phys. C **17**, 3073 (1984).

¹⁰E. Stoll, Surf. Sci. **143**, L411 (1984).

¹¹Y. Kuk and P. J. Silverman, Appl. Phys. Lett. **48**, 1597 (1986); G. Van de Walle, H. Van Kempen, and P. Wyder, Surf. Sci. **167**, L219 (1986).

¹²H. W. Fink, IBM J. Res. Dev. **30**, 5 460 (1986).

¹³J. Bono and R. H. Good, Surf. Sci. **151**, 543 (1985).

¹⁴A. Baratoff, Physica **127B**, 143 (1984).

¹⁵T. E. Feuchtwang, P. H. Cutler, and E. Kazes, J. Phys. (Paris) **48**, C9-111 (1984).

¹⁶H. Feshbach and A. M. Clogston, Phys. Rev. **59**, 189 (1940).

¹⁷W. L. Clinton, M. A. Esrick, and W. S. Sacks, Phys. Rev. B **35**, 4074 (1987).

¹⁸There is some disagreement in the literature on the R dependence of the tunneling current (see, for example, Refs. 2 and 8).

¹⁹P. K. Hansma and J. Tersoff, J. Appl. Phys. **61**, R1 (1987).

²⁰*Handbook of Mathematical Functions with Formulas, Graphs and Mathematical Tables*, edited by M. Abramowitz and I. A. Stegun (U.S. Dept. Commerce, Washington, D.C., 1964).

²¹J. Tersoff, Phys. Rev. Lett. **57**, 440 (1986).

²²M. C. Payne, J. Phys. C **19**, 781 (1986).

²³G. Binnig, H. Rohrer, Ch. Gerber, and E. Stoll, Surf. Sci. **144**, 321 (1984); R. J. Behm, W. Hösler, E. Ritter, and G. Binnig, Phys. Rev. Lett. **56**, 228 (1986).

²⁴W. Sacks, S. Gauthier, S. Rousset, and J. Klein (unpublished).