Theory and Application for the Scanning Tunneling Microscope

J. Tersoff and D. R. Hamann

Bell Laboratories, Murray Hill, New Jersey 07974

A theory is presented for vacuum tunneling between a real solid surface and a model probe with a locally spherical tip, applicable to the recently developed "scanning tunneling microscope." Calculations for 2×1 and 3×1 reconstructions of Au(110) are in excellent agreement with recent experimental results, if an effective radius of curvature of 9 Å is assumed for the tip.

One of the most fundamental problems in surface physics is the determination of surface structure. Recently a new and uniquely promising technique, the "scanning tunneling microscope" (STM), was introduced.[1-4] This method offers, for the first time, the possibility of *direct*, *real-space* determination of surface structure, including nonperiodic structures. A small metal tip is brought near enough to the surface that the vacuum tunneling resistance between surface and tip is finite and measurable. The tip scans the surface in two dimensions, while its height is adjusted to maintain a constant tunneling resistance. The result is essentially a contour map of the surface.

The one-dimensional tunneling problem (i.e., through two-dimensionally-uniform barriers) has been treated extensively,[5] and field emission from a tip is well understood. The usefulness of STM stems from the fact that it is neither one-dimensional nor field emission, but is instead sensitive to the full three-dimensional structure of the surface. Little is known about tunneling in this case.[4] Here we present the first quantitative theory for the scanning tunneling microscope, and apply the theory to Au(110). Results are in excellent agreement with experiment. We infer that under actual experimental conditions of Ref. 4, the tip was roughly 6 Å from the surface and had an effective radius of curvature of about 9 Å.

Unfortunately, little is known about the structure of the tunneling probe tip, which is at present prepared in a relatively uncontrolled and non-reproducible manner.[1,4] Similar results have been obtained with W, Mo, and stainless steel tips,[6] so that details of the tip electronic structure do not appear to be important. We model the tip as a locally spherical potential well where it approaches nearest to the surface, as illustrated in Fig. 1. R is the local radius of curvature about the center located at $\vec{r}_0$, and d is the distance of nearest approach to the surface.

The tunneling current by first-order perturbation theory is

$$I = (2\pi e/\hbar)\sum_{\mu\nu} f(E_\mu)[1 - f(E_\nu + eV)] \times |M_{\mu\nu}|^2 \delta(E_\mu - E_\nu), \quad (1)$$

where $f(E)$ is the Fermi function, V is the applied voltage, $M_{\mu\nu}$ is the tunneling matrix element between states ψ_μ of the probe and ψ_ν of the surface, and E_μ is the energy of state ψ_μ in the absence of tunneling. Since the experiments are performed at room temperature or below and at small voltage (~ 10 meV for metal-metal tunneling), we take the limits of small voltage and temperature,

$$I = (2\pi/\hbar)e^2 V \sum_{\mu\nu} |M_{\mu\nu}|^2 \delta(E_\nu - E_F)\delta(E_\mu - E_F), \quad (2)$$

where E_F is the Fermi level. The essential problem is to calculate $M_{\mu\nu}$. Bardeen[7] has shown that

$$M_{\mu\nu} = -(\hbar^2/2m)\int d\vec{S} \cdot (\psi_\mu^* \nabla \psi_\nu - \psi_\nu \nabla \psi_\mu^*), \quad (3)$$

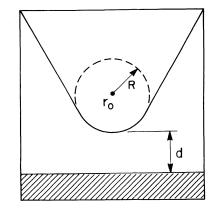where the integral is over any surface lying entirely within the vacuum (barrier) region separat-

FIG. 1. Schematic picture of tunneling geometry. Probe tip has arbitrary shape but is assumed locally spherical with radius of curvature R, where it approaches nearest the surface (shaded). Distance of nearest approach is d. Center of curvature of tip is labeled $\vec{r}_0$.

ing the two sides. The quantity in parentheses is simply the current operator.

To evaluate $M_{\mu\nu}$, we expand the surface wave function in the form

$$\psi_\nu = \Omega_s^{-1/2} \sum_G a_G \exp[-(k^2 + |\vec{k}_{\parallel} + \vec{G}|^2)^{1/2} z]$$

$$\times \exp[i(\vec{k}_{\parallel} + \vec{G}) \cdot \vec{x}], \quad (4)$$

which is a completely general expression for ψ in the region of negligible potential. Here Ω_s is sample volume, $k = \hbar^{-1}(2m\varphi)^{1/2}$ is the inverse decay length for the wave functions in vacuum, φ is the work function, $\vec{k}_{\parallel}$ is the surface Bloch wave vector of the state, and $\vec{G}$ is a surface reciprocal-lattice vector. For a nonperiodic surface the sum over G becomes an integral. The first few a_G are of order unity.

The wave functions of the tip are, in general, not known. In the region of interest, however, they have the asymptotic spherical form

$$\psi_\mu = \Omega_t^{-1/2} c_t kR e^{kR} (k|\vec{r} - \vec{r}_0|)^{-1} e^{-k|\vec{r} - \vec{r}_0|}, \quad (5)$$

where Ω_t is the probe volume, k is defined as above, and R is the radius of curvature. (We assume for simplicity that the work function φ for the tip is equal to that of the surface.) The form is chosen to be correctly normalized when the parameter c_t (which is determined by the tip geometry, detailed electronic structure, and tip-vacuum boundary condition) is of order 1, with the assumption $R \gg k^{-1}$. We have neglected the possible angular dependence of ψ_μ, which introduces some quantitative modifications discussed below.

We expand the tip wave function (5) in the same form as the surface (4) using the fact that

$$(kr)^{-1} e^{-kr} = \int d^2q \, b(\vec{q}) \exp[-(k^2 + q^2)^{1/2}|z|]$$

$$\times \exp(i\vec{q} \cdot \vec{x}), \quad (6)$$

$$b(\vec{q}) = (2\pi)^{-1} k^{-2}(1 + q^2/k^2)^{-1/2}. \quad (7)$$

The matrix element is then almost trivial to evaluate. Substituting the surface and the tip wave functions in (3) and evaluating the expansion term by term in G, one finds

$$M_{\mu\nu} = (\hbar^2/2m)4\pi k^{-1} \Omega_t^{-1/2} kR e^{kR} \psi_\nu(\vec{r}_0), \quad (8)$$

where $\vec{r}_0$ is the position of the *center of curvature* of the tip. On substitution into (2) the desired result is

$$I = 32\pi^3 \hbar^{-1} e^2 V \varphi^2 D_t (E_F) R^2 k^{-4} e^{2kR}$$

$$\times \sum_\nu |\psi_\nu(\vec{r}_0)|^2 \delta(E_\nu - E_F), \quad (9)$$

where D_t is the density of states per unit volume of the probe tip. Note that (8) does *not* imply that the value of the surface wave function ψ_ν at $\vec{r}_0$ is *physically* relevant. The matrix element is determined by an integral entirely within the gap region. However, because of the analytic properties of (4) and (5), the formal evaluation of ψ_ν at distance $R + d$ correctly describes the lateral averaging due to finite tip size.

The spherical-tip approximation entered only in the normalization of (5). The crucial approximation was evaluating the matrix element only for an s-wave tip wave function. The $\vec{q}$ dependence of $b(\vec{q})$ in (7) then cancelled that of the z derivative in the matrix element (3), so that (9) involved only undistorted wave functions of the surface. For tip wave functions with angular dependence $(l \neq 0)$, it is sufficient to include the $m = 0$ term (other m give a node towards the surface). In that case the terms in the Fourier expansion of ψ_ν are weighted by a factor $\sim(1 + q^2/k^2)^{1/2}$ in the matrix element, which for relevant values of q can be neglected for small l. (In the example below the relevant $q^2/k^2 \approx 0.1$.) The tip model therefore becomes less accurate for large R, where higher l values become more important. A more exact treatment would probably be far less useful, since it would require more specific information about the tip wave functions, and would not reduce to an explicit equation such as (9) or (10) and (11) below.

Substituting typical metallic values into (9), one obtains for the tunneling conductance

$$\sigma \approx 0.1 R^2 e^{2kR} \rho(\vec{r}_0; E_F), \quad (10)$$

$$\rho(\vec{r}_0; E) \equiv \sum_\nu |\psi_\nu(\vec{r}_0)|^2 \delta(E_\nu - E), \quad (11)$$

where σ is in Ω^{-1}, distances are in atomic units, and energy in electron volts. Since $|\psi_\nu(\vec{r}_0)|^2 \propto \exp[2k(R + d)]$, we see from (10) that $\sigma \propto e^{-2kd}$ as expected. Because of the exponential dependence on distance, it is not essential that the coefficient in (10) be every accurate.

By consideration of general aspects of the behavior of $\rho(\vec{r}; E_F)$, it is possible to draw some conclusions concerning the sensitivity of STM. In particular, the suppression of higher Fourier components in (4) and (9) is equivalent to an instrumental resolution of roughly $[2k^{-1}(R + d)]^{1/2}$, at least for components $q \ll 2k$. Since $2k^{-1} \approx 1.6$ Å, if $R + d \approx 15$ Å (as in the example below), the resolution is about 5 Å, so that the 8 Å periodicity of Au(110) (2×1) is resolved. It is important to note that for periodic structures the measured corru-
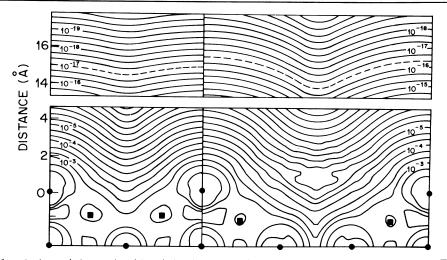
FIG. 2. Calculated $\rho(r; E_F)$ for Au(110)(2×1) (left) and (3×1) (right) surfaces. Figure shows (1$\bar{1}$0) plane through outermost atoms. Positions of nuclei are indicated by solid circles (in plane) and squares (out of plane). Contours of constant $\rho(r; E_F)$ are labeled in units of a.u.$^{-3}$ eV^{-1}. Note break in distance scale. Peculiar structure around contour 10^{-5} of (3×1) is due to limitations of the plane-wave part of the basis in describing the exponential decay inside the deep troughs. Center of curvature of probe tip follows dashed line.

gation amplitude decreases *exponentially* with resolution. This Debye-Waller-like behavior explains the puzzle[4] of why in Ref. 2, with somewhat poorer resolution, nonperiodic structures such as steps were clearly resolved on Au(110), but the periodic reconstruction was not.

The Au(110) surface normally exhibits 2×1 reconstruction with a missing-row geometry.[8] A 3 × 1 reconstruction has also been observed.[9] Recently Binnig *et al.*[4] reported high-resolution STM measurements for an Au(110) surface with regions of both 2×1 and 3×1 structure, and concluded that the 3×1 structure consisted of (111) microfacets analogous to the 2×1. Measured STM corrugations were 0.45 and 1.4 Å for 2×1 and 3×1, respectively. (The two phases occurred together and were measured in the same scan with the same tip, permitting direct comparison.)

We have calculated $\rho(\vec{r}; E_F)$ for both 2×1 and 3×1 surfaces using a recently developed linearized augmented plane-wave method described elsewhere.[10] For the 2×1 surface we used a slab geometry of three complete layers with a half-layer [alternate (1$\bar{1}$0) rows missing] on either side. The 3×1 geometry suggested by Binnig *et al.*[4] was employed; an asymmetric slab was constructed of two complete layers, a third layer with one missing row, and a fourth layer with two missing rows (see Fig. 2). The calculation is similar to that in Ref. 10, with $\rho(\vec{r}; E_F)$ approximated by the charge in states within 0.5 eV of E_F, divided by the finite interval width of 1 eV. (The

error from this finite-interval approximation can be estimated quantitatively and should not affect any conclusions here.) Despite some difficulty in obtaining convergence with respect to the plane-wave part of our basis set for these open structures and large unit cells, we believe the numerical accuracy to be quite adequate for the 3×1 and still better for the 2×1 geometry. We are currently investigating approximate methods for treating more complicated structures, since the greatest strength of STM is for systems with large unit cells and for nonperiodic systems.

Figure 2 shows the calculated $\rho(\vec{r}; E_F)$ for Au(110). Since the actual tip geometry is not known, we consider a tip radius $R = 9$ Å, so that (10) predicts a (2×1) corrugation of 0.45 Å at tunneling resistance[6] 10^7 Ω in agreement with experiment. Then d is found to be 6 Å, measured from the surface Au nuclei to the edge of the tip potential well (i.e., the shell at which the tip wave function becomes decaying in character). This value is consistent[6] with experimental estimates of d based on resonant tunneling oscillations.[4] Given R, (10) yields a corrugation of 1.4 Å for the (3×1) surface in excellent agreement with experiment.

In the analysis above d was determined rather directly by (10) since the dependence of current upon R largely cancels as noted above. However, R was inferred by fitting the experimental corrugation, which depends on $R + d$, and subtracting d. The corrugation is more susceptible to errors,

both experimental and theoretical, than is the current. Moderate errors ($\sim 20\%$) in either the calculated or measured corrugation amplitude have little effect on our conclusions. Nevertheless, since this is the first such calculation for STM, we believe it would be premature to rule out a tip consisting in effect of one or two atoms. For a sufficiently small cluster of atoms, the effective value of R depends on the precise geometry.

We conclude that a relatively simple model for the tip, in conjunction with detailed calculations for the surface, gives excellent agreement with experimental results of STM and provides insight into the method's resolution and sensitivity. With improved characterization of the probe tip structure in the future, more precise comparisons will become possible, furthering the detailed understanding of STM.

[1]G. Binnig, H. Rohrer, Ch. Gerber, and E. Weibel, Appl. Phys. Lett. **40**, 178 (1982).

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

[3]G. Binnig, H. Rohrer, Ch. Gerber, and E. Weibel, Phys. Rev. Lett. **50**, 120 (1983).

[4]G. Binnig, H. Rohrer, Ch. Gerber, and E. Weibel, Bull. Am. Phys. Soc. **28**, 461 (1983), and to be published.

[5]C. B. Duke, *Tunneling in Solids* (Academic, New York, 1969).

[6]H. Rohrer, private communication.

[7]J. Bardeen, Phys. Rev. Lett. **6**, 57 (1961).

[8]I. K. Robinson, Phys. Rev. Lett. **50**, 1145 (1983), and references therein.

[9]W. Moritz and D. Wolf, Surf. Sci. **88**, L29-34 (1979); Young Kuk, Bull. Am. Phys. Soc. **28**, 260 (1983), and private communication.

[10]D. R. Hamann, Phys. Rev. Lett. **46**, 1227 (1981).