

Theory of the scanning tunneling microscope

J. Tersoff* and D. R. Hamann

AT&T Bell Laboratories, Murray Hill, New Jersey 07974

(Received 25 June 1984)

We present a theory for tunneling between a real surface and a model probe tip, applicable to the recently developed "scanning tunneling microscope." The tunneling current is found to be proportional to the local density of states of the surface, at the position of the tip. The effective lateral resolution is related to the tip radius R and the vacuum gap distance d approximately as $[(2 \text{ \AA})(R+d)]^{1/2}$. The theory is applied to the 2×1 and 3×1 reconstructions of Au(110); results for the respective corrugation amplitudes and for the gap distance are all in excellent agreement with experimental results of Binnig *et al.* if a 9-Å tip radius is assumed. In addition, a convenient approximate calculational method based on atom superposition is tested; it gives reasonable agreement with the self-consistent calculation and with experiment for Au(110). This method is used to test the structure sensitivity of the microscope. We conclude that for the Au(110) measurements the experimental "image" is relatively insensitive to the positions of atoms beyond the first atomic layer. Finally, tunneling to semiconductor surfaces is considered. Calculations for GaAs(110) illustrate interesting qualitative differences from tunneling to metal surfaces.

I. INTRODUCTION

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 in three dimensions, 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.

For electronic states at the Fermi level, the surface represents a potential barrier whose height is equal to the work function ϕ . As expected by analogy with planar tunneling, the current varies exponentially with the vacuum gap distance, with decay length $\pi(8m\phi)^{-1/2}$. For typical metallic work functions, this length is about 0.4 Å. Thus, aside from issues of *lateral* resolution, in the constant-current scanning mode the tip may be expected to follow the surface height to 0.1 Å or better. It can be seen from the data that the new microscope designs have sufficient mechanical stability to achieve this in practice.¹⁻⁴

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 operating as a field emitter, but is instead sensitive to the full three-dimensional structure of the surface. Little was known quantitatively about tunneling in this case, until the recent development of STM motivated the present work (parts of which were reported briefly elsewhere⁶), and other approaches,^{7,8} which

are discussed briefly below.

Here we develop a theory of STM which is at once sufficiently realistic to permit quantitative comparison with experimental "images," and sufficiently simple that the implementation is straightforward. The surface is treated "exactly," while the tip is modeled as a locally spherical potential well where it approaches nearest the surface. This treatment is intuitively reasonable and is consistent with the current poor understanding of the actual microscopic geometry of the tip, which is prepared in an uncontrolled and nonreproducible manner.⁹

In Sec. II we present the formal development of the theory. The tunneling current is found to be proportional to the (bare) surface local density of states (LDOS) at the Fermi level (E_F) at the position of the tip. The effective lateral resolution is roughly $[(2 \text{ \AA})(R+d)]^{1/2}$, where R is the tip radius of curvature and d is the vacuum gap. General features of the surface LDOS are discussed, as are the various approximations. Some other recent approaches^{7,8} to the problem are also considered.

Section III describes a calculation for the 2×1 and 3×1 reconstructions of the Au(110) surface. The results are in quantitative agreement with recent measurements of Binnig *et al.*⁴ if a 9-Å tip radius is assumed. General features and limitations of the numerical implementation are discussed. In particular, self-consistent electronic structure calculations of vacuum charge far from the surface are at present only feasible for systems with small unit cells.

We therefore introduce in Sec. IV a crude approximation for the surface LDOS, which permits convenient calculation of the STM image even for large unit cells or nonperiodic systems. Comparison with results of Sec. III shows that the approximation works rather well, at least for Au(110). Using this approximation, we compare the images expected for different possible structures of Au(110). We conclude that STM is rather insensitive to

the position of the surface layer relative to the underlying layers. For the Au(110) 3×1 surface, even the presence or absence of a missing row in the second layer cannot be reliably distinguished.

Finally, in Sec. V, we consider the case of a semiconducting surface. The theory is expected still to apply, though with some modifications. In particular, the image may be qualitatively different for different tunneling polarity or sample doping. This effect is illustrated with calculations for GaAs(110).

II. THEORY OF STM

While it is easy to write down a formal expression for the tunneling current, many approximations are needed to derive an expression which permits practical computation. Some of the approximations made below are sufficiently drastic that they can be justified only because of the relative insensitivity of any conclusions to the resulting errors. It is therefore not convenient to justify the various approximations as they are introduced. Instead, we first present the theory in Sec. II A. Then in Sec. II B we consider general features of the surface local density of states and, hence, of the tunneling current as a function of tip position. These results determine the intrinsic resolution and sensitivity of STM. Finally, in Sec. II C we consider the various approximations and their possible effect.

A. Tunneling current

The tunneling current is given to first order in Bardeen's¹⁰ formalism by

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

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. Note that while (1) resembles ordinary first-order perturbation theory, it is formally different in that ψ_μ and ψ_ν are nonorthogonal eigenstates of different Hamiltonians. For high temperatures there is a corresponding term for reverse 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,

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

Before attempting a realistic treatment, it is worthwhile to consider the limit where the tip is replaced with a point probe. This case represents the ideal of a nonintrusive measurement of the surface, with the maximum possible resolution. If the tip wave functions are arbitrarily localized, then the matrix element is simply proportional to the amplitude of ψ_ν at the position \vec{r}_0 of the probe, and (2) reduces to

$$I \propto \sum_\nu |\psi_\nu(\vec{r}_0)|^2 \delta(E_\nu - E_F).$$

The quantity on the right is simply the surface local density of states (LDOS) at E_F , i.e., the charge density from states at E_F . Thus the tunneling current is proportional to the surface LDOS at the position of the point probe, and the microscope image represents a contour map of constant surface LDOS. This almost trivial result anticipates major features of the more complete treatment below.

In handling (2) in general, the essential problem is to calculate $M_{\mu\nu}$. Bardeen¹⁰ has shown that

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

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

$$\psi_\nu = \Omega_s^{-1/2} \sum_G a_G \exp[(\kappa^2 + |\vec{k}_G|^2)^{1/2} z] \exp(i\vec{k}_G \cdot \vec{x}), \quad (4)$$

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

Since the microscopic structure of the tip is not yet known, we model it 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. In the region of interest, the wave functions of the tip are taken to have the asymptotic spherical form

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

where Ω_t is the probe volume and κ is defined as above.

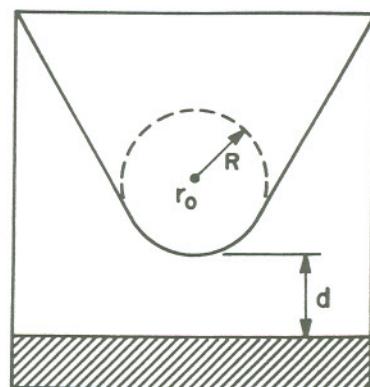


FIG. 1. Schematic picture of tunneling geometry. Probe tip has arbitrary shape but is assumed locally spherical with radius of curvature R , where it approaches nearest to the surface (shaded). Distance of nearest approach is d . Center of curvature of tip is labeled \vec{r}_0 .

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. 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

$$(\kappa \vec{r})^{-1} e^{-\kappa \vec{r}} = \int d^2 q b(\vec{q}) \exp[-(\kappa^2 + q^2)^{1/2} |z|] \times \exp(i \vec{q} \cdot \vec{x}), \quad (6)$$

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

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

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

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

$$I = 32\pi^3 \hbar^{-1} e^2 V \phi^2 D_t(E_F) R^2 \kappa^{-4} e^{2\kappa R} \times \sum_v |\psi_\nu(\vec{r}_0)|^2 \delta(E_v - E_F), \quad (9)$$

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 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 canceled that derivative in the matrix element (3), so that (9) involves only undistorted wave functions of the surface or 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 $(1 + q^2/\kappa^2)^{1/2}$ in the matrix element, which for relevant

values of q can be approximated by unity for small l . (In the example below the relevant $q^2/\kappa^2 \approx 0.1$.) 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) below.

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

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

$$\rho(\vec{r}_0, E) \equiv \sum_v |\psi_\nu(\vec{r}_0)|^2 \delta(E_v - E),$$

where σ is in ohms $^{-1}$, distances are in a.u., and energy in eV. Since $|\psi_\nu(\vec{r}_0)|^2 \propto e^{-2\kappa(R+d)}$, we see from (10) that $\sigma \propto e^{-2\kappa d}$ as expected. Because of the exponential dependence on distance, it is not essential that the coefficient in (10) be very accurate.

We considered above the limit of a point probe. Realistically, the sharpest tip imaginable is a single atom, supported on a cluster or small plateau. The form (5) is not really appropriate for determining the normalization of ψ_μ in that case. However, because of the insensitivity of results to the coefficients, an adequate estimate for the single-atom case may be obtained simply by taking $R = 2\kappa^{-1}$ (roughly the metallic radius for most metals) in (10).

Note that $\rho(\vec{r}, E_F)$ is simply the surface local density of states (at E_F) at the point \vec{r} or, equivalently, the charge per unit energy from states of the surface at E_F . According to (10), at constant current the tip follows a contour of constant $\rho(\vec{r}, E_F)$. We therefore consider the behavior of $\rho(\vec{r}, E_F)$ in some detail.

B. General features of $\rho(\vec{r}, E_F)$

Within the approximations above, the microscope image is simply a contour of constant $\rho(\vec{r}, E_F)$ of the surface. The behavior of $\rho(\vec{r}, E_F)$, along with the tip radius, therefore determines the resolution and sensitivity of STM. Moreover, a detailed picture of $\rho(\vec{r}, E_F)$ is essential in assessing the approximate method described in Sec. IV.

The starting point here is Eq. (4) for the surface wave function. A given wave function ψ_ν contributes a charge density

$$|\psi_\nu|^2 = \Omega_s^{-1} \sum_{G,G'} a_G a_G^* \exp[-(\kappa^2 + \kappa_G^2)^{1/2} z - (\kappa^2 + \kappa_{G'}^2)^{1/2} z + i(\vec{k}_G - \vec{k}_{G'}) \cdot \vec{r}]. \quad (11)$$

Since $\vec{k}_G - \vec{k}_{G'} = \vec{G} - \vec{G}'$, $|\psi_\nu|^2$ has the periodicity of the lattice and can be Fourier expanded,

$$|\psi_\nu|^2 = \sum_G u_{\nu G}(z) e^{i \vec{G} \cdot \vec{r}}. \quad (12)$$

The total $\rho(\vec{r}, E)$ may similarly be written

$$\rho(\vec{r}, E) = \sum_v |\psi_\nu|^2 \delta(E_v - E) = \sum_G \rho_G(z, E) e^{i \vec{G} \cdot \vec{r}}, \quad (13)$$

$$\rho_G(z, E) = \sum_v u_{\nu G}(z) \delta(E_v - E). \quad (14)$$

At sufficiently large distances $\rho(\vec{r}, E_F)$ becomes rather smooth, and only the lowest nonzero Fourier component need be retained for this discussion. This is so in the example of Au(110) in Sec. III, where the STM image is practically sinusoidal. (The case where the image is highly structured is considered below.) Then

$$\rho(\vec{r}, E_F) = \rho_0(z, E_F) + 2\rho_{G_1}(z, E_F) \cos(\vec{G}_1 \cdot \vec{x}), \quad (15)$$

where we have assumed a reflection symmetry and where G_1 is the smallest G . Far from the surface $\rho_0(z, E_F)$ is dominated by states near the center ($\vec{0}$) of the surface Brillouin zone, since $\kappa_G=0$ gives the longest decay length (decay constant $= 2\kappa$), in (4) and (11). It can be seen from minimizing the exponents in (11) that the longest decay length for ρ_{G_1} occurs at $\vec{k}_{||} = \frac{1}{2}\vec{G}_1$. Then $|\kappa_G| = \frac{1}{2}G_1$ for $G=0$ or $-G_1$. The corresponding asymptotic decay constant for ρ_{G_1} is

$$\begin{aligned} \frac{d}{dz} \ln[\rho_{G_1}(z, E_F)] &= 2(\kappa^2 + \frac{1}{4}G_1^2)^{1/2} \\ &\approx 2\kappa + \frac{1}{4}\kappa^{-1}G_1^2 \end{aligned} \quad (16)$$

using $\frac{1}{2}G_1 \ll \kappa$. [For Au(110), $G_1/2\kappa \approx 0.1$.]

The extremal values of z for constant current [constant $\rho(\vec{r}, E_F)$] occur at $\cos(\vec{G}_1 \cdot \vec{x}) = \pm 1$, and these are denoted z_{\pm} here. Then from (15),

$$\rho(\vec{r}) = \rho_0(z_{\pm}) \pm 2\rho_{G_1}(z_{\pm}), \quad (17)$$

where the argument E_F is omitted for simplicity. Defining the corrugation amplitude $\Delta = z_+ - z_-$ and using $\rho_0(z_+) \approx e^{-2\kappa\Delta} \rho_0(z_-)$, (17) gives

$$e^{-2\kappa\Delta} \approx \frac{\rho_0(z) - 2\rho_{G_1}(z)}{\rho_0(z) + 2\rho_{G_1}(z)}, \quad (18)$$

where z is some average value between z_+ and z_- . At distances where the image is sufficiently smooth ($\kappa\Delta \ll 1$), using (16), (18) becomes

$$\begin{aligned} \Delta &\approx 2\kappa^{-1} \rho_{G_1}(z)/\rho_0(z) \propto e^{-\beta z}, \\ \beta &\equiv 2(\kappa^2 + \frac{1}{4}G_1^2)^{1/2} - 2\kappa \approx \frac{1}{2}\kappa^{-1}G_1^2. \end{aligned} \quad (19)$$

Thus, the corrugation decreases exponentially with distance from the surface, the decay length β being very sensitive to the surface lattice constant. This corrugation decay length is in agreement with numerical calculations described below. Of course, the result applies only far from the surface ($\beta z > 1$), and is not strictly correct (though it still works well in practice) if there is a gap in the projected one-dimensional density of states at E_F for $\vec{k}_{||}=0$ or $\frac{1}{2}\vec{G}_1$.

If the surface unit cell is large, then the features of interest in the image may be well localized within the unit cell. This is the case, for example, in images of the Si(111) 7×7 surface.³ Then the Bloch wave vector may be neglected, and it is easy to show that for any G such that $G_1 \ll G < \kappa$, the most slowly decaying term in (11) contributing to ρ_G has the same falloff as given in (16), with G_1 replaced by G , to lowest order in (G^2/κ^2) . Thus the

decay constant (16) seems to have rather general validity.

These results may be used to define an effective real-space resolution for STM. The suppression of the Fourier term for $G \neq 0$ by a factor $\exp(-\frac{1}{4}\kappa^{-1}G^2 z)$ is precisely the effect of averaging over a Gaussian resolution function of rms width $(0.5\kappa^{-1}z)^{1/2}$, i.e., full width at half maximum $1.66(\kappa^{-1}z)^{1/2}$. Recall that for the relevant contours, $z=R+d$; if $R \gg d$, the resolution is determined by the tip radius but is nonetheless much smaller than R , since $\kappa^{-1} < 1 \text{ \AA}$. For $R \ll d$, as in the case of a single-atom effective tip, the resolution is limited by d and, therefore, by how small a tunneling resistance is experimentally feasible. Note, however, that reducing d from 6 to 4 Å requires a decrease in tunneling voltage, or an increase in current, by roughly a factor of 200, and yet gives only a 20% increase in resolution.

In the plane of the surface atoms, $z=0$, $\rho(r)$ is rather localized within the unit cell. It is reasonable, therefore, to assume $\rho_0(z=0) \approx \rho_G(z=0)$, as long as there is only one atom per surface cell. Then at large z the corrugation (19) becomes

$$\Delta \approx 2\kappa^{-1} e^{-\beta z}. \quad (20)$$

This crude approximation, in fact, gives a good semiquantitative description of the results of the self-consistent calculations for Au(110) described below. A more systematic (though similarly crude) prescription for approximating $\rho(r, E_F)$ is suggested below.

C. Assessment of approximations

We now return to the theory developed in Sec. II A and consider the accuracy and generality of the many approximations made there.

The most crucial point is that rather large errors can be tolerated in the coefficient in (9) and (10). A factor of $e^2 \approx 7$ error in the coefficient shifts the inferred value of \vec{r}_0 for a given current by only $\kappa^{-1} < 1 \text{ \AA}$. The corresponding change in the corrugation Δ is, using (19), roughly a factor of $\exp(-\frac{1}{4}G^2/\kappa^2) \approx 0.92$ for Au(110) 2×1 ($G \approx 0.8 \text{ \AA}^{-1}$), an error under 10%. The substitution of typical metallic values in (10) is thus quite adequate.

As mentioned above, the use of an s -wave tip wave function is adequate if the real wave functions are restricted to small angular momentum l . For a sufficiently large effective tip the approximation is expected to lose its detailed validity. In any case, the s -wave treatment here is not intended as an accurate description of a real tip, but rather as a useful way of parametrizing the effect of finite tip size, which is otherwise relatively intractable.

In Sec. II A we implicitly assumed that the potential goes to zero in a region between the surface and tip and that the integral (3) is taken in that region. Actually, the electron is never more than about 3–6 Å from the surface, so the magnitude of the potential is never less than $\sim 1 \text{ eV}$. Locally the effective value of κ is

$$\kappa(r) = \hbar^{-1} (2m)^{1/2} [\phi + V(r)]^{1/2}.$$

The resulting modest ($\sim 10\%$) change in κ is unimportant

except as it affects the wave function [as opposed to the coefficient in (9)]. The surface wave functions are calculated using the full potential, so we anticipate no problems except that we neglect the contribution of the tip to the potential. Without a more precise description of the real tip, we see no way to incorporate the tip potential in a consistent fashion.

(The local-density potential and thin slab geometries used in Sec. III below may be a source of inaccuracy in the implementation. However, such problems of *implementation* are a separate issue from the *intrinsic* limitations of the model presented here.)

When the tip and surface have different work functions, the resulting potential gradient causes a smooth variation in the effective $\kappa(\vec{r})$ between surface and tip. Again, as long as the difference is not great (< 1 eV), we expect no crucial changes.

While the nominal current density may be quite large, the nanoamp current used corresponds to one electron per 1.6×10^{-10} sec. This is long compared to relevant transit times¹¹ as well as phonon vibration and relaxation times.

Electrons may therefore be viewed as tunneling one at a time, and effects such as space charge and sample heating should be negligible.

We conclude that the approximations of Sec. II A are adequate for a quantitative understanding of STM, within the constraint of nearly complete ignorance of the microscopic structure of the tip. We hope that in the future improved characterization of the tip will permit a better evaluation of the model presented here.

Some other theoretical treatments of STM have been reported recently.^{7,8} Garcia *et al.*⁷ have applied methods developed for atom diffraction to calculate the current between two periodic metal surfaces. One surface is taken as strongly corrugated, and represents the tip. The potential is taken as flat throughout, with abrupt discontinuities at the two surfaces. A major drawback of this approach is that it is strictly numerical, and gives no direct insight or intuition into what is measured in STM. Moreover the quantitative results must be viewed with some caution. The model form for the potential is somewhat arbitrary;

more important, the tip is treated in a peculiar manner. The tip is apparently taken as infinitely extended in one dimension; moreover it is periodically repeated, which could give rise to irrelevant and unphysical interference effects. The model of Garcia *et al.* is appropriate for studying qualitative aspects of vacuum tunneling, but it is not clear how it could be usefully applied to aid in making structural inferences from experimental images. We prefer the approach taken here because it is at once more quantitative and more conceptually transparent.

Feuchtwang *et al.*⁸ have pointed out that, instead of assuming a specific form for the tip wave function, one may represent the current as a convolution of spectral functions of the surface and tip. This suggests a possible avenue of investigation, but may not be applicable to imaging in the constant current mode (the only mode of STM imaging now considered practical). In any case the spectral function of the tip is not known, so implementing this approach would probably require approximations similar in spirit to those here. We prefer to retain the explicit asymmetry between surface and tip, reflecting the asymmetry both in our interest and in our understanding of the two.

III. AN EXAMPLE: Au(110)

In this section we describe a self-consistent calculation for the Au(110) surface, and compare our results with recent measurements of Binnig *et al.*⁴ Agreement is excellent if a tip radius of 9 Å is assumed. We also discuss factors limiting the accuracy of the calculation and conclude that such calculations are feasible for relatively few systems of interest for STM. In the next section we consider an alternative approach which is less reliable but is feasible even for extremely complex systems.

The Au(110) surface normally exhibits 2×1 reconstruction with a missing-row geometry.¹² A 3×1 reconstruction has also been observed.¹³ 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.)

Since Au(110) is the only surface with a tractable unit-cell size for which high-resolution STM images are available, we have chosen it for detailed study in this section and in the next. We have calculated $\rho(\vec{r}, E_F)$ for both 2×1 and 3×1 surfaces using a recently developed linearized augmented-plane-wave (LAPW) method described elsewhere.¹⁴ For the 2×1 surface we used a slab geometry of three complete layers with a half layer [alternate (110) 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. 13, 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.

Figure 2 shows the calculated $\rho(\vec{r}, E_F)$ for Au(110). Since the actual tip geometry is not known, we choose a tip radius $R = 9$ Å, so that (10) gives a (2×1) corrugation of 0.45 Å at tunneling resistance⁹ 10^7 Ω to fit 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 consistent⁹ 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.

The agreement here is gratifying; with one parameter, R , we obtain good agreement with two experimental corrugations, as well as with the gap distance. Nevertheless, it is worth briefly considering the numerical aspects of the calculation, which limit its accuracy.

In the surface LAPW method the wave functions are expanded in a Laue basis beyond the last plane of atoms, so the exponential decay poses no problems for simple surfaces. However, for the very "open" Au(110) surface,

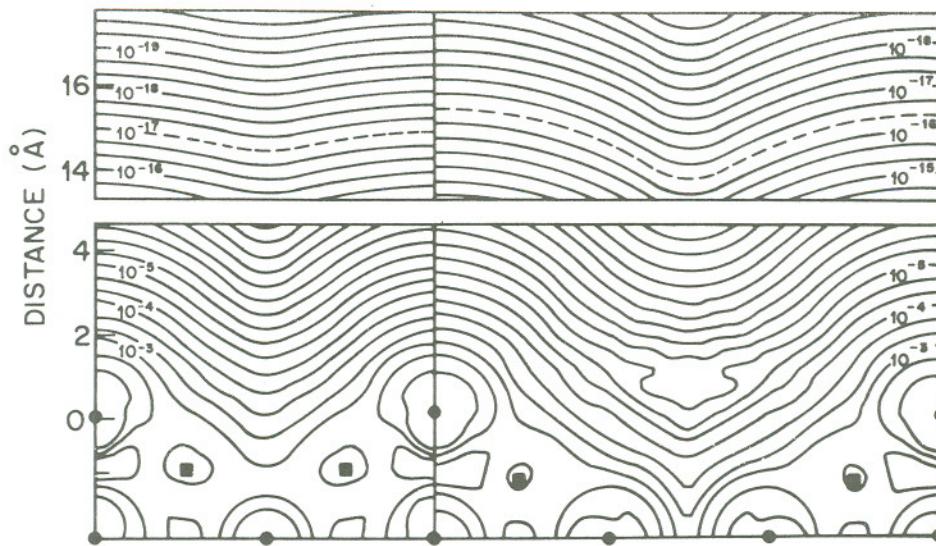


FIG. 2. Calculated $\rho(\vec{r}, E_F)$ for Au(110) (2×1) (left) and (3×1) (right) surfaces. Figure shows $(\bar{1}\bar{1}0)$ plane through outermost atoms. Positions of nuclei are indicated by solid circles (in plane) and squares (out of plane). Contours of constant $\rho(\vec{r}, E_F)$ are labeled 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.

we are obliged to expand the wave functions in a plane-wave basis in the "trough" region where surface atoms are missing. Since the wave functions decay exponentially there, the expansion converges slowly. The persistence of Gibbs's oscillations in the charge (Fig. 2) suggests that the convergence is still imperfect, but the 400-plane-wave expansion is the maximum possible with a CRAY-1 computer and our current code. (Some improvement could be obtained by taking advantage of inversion and reflection symmetry for a suitable slab geometry.)

The other major source of inaccuracy is the very thin "slab" geometry employed. This might result in an inaccurate work function, which would certainly affect the results to some extent. The calculated work functions are 5.7 and 5.2 eV for 2×1 and 3×1 surfaces, respectively. Also, the thin slab gives only a few discrete states for a given wave vector. This sparse sampling of the bulk continuum leads to a numerical noise in energy-projected quantities. For this reason we included states from a rather large (1-eV) interval to approximate the charge $\rho(\vec{r}, E_F)$ from states at E_F .

The local-density approximation used here does not reproduce the correct image form of the correlation potential at large distances from the surface; presumably it also gives incorrect lateral structure in the correlation potential in this region. Neither of these shortcomings has a significant effect on the results, however. The "crossover" from the high-density regime to the image-potential regime occurs well outside the classical turning point for electrons at the Fermi level, where the potential is small compared to the (negative) kinetic energy. The structure in the wave functions is determined by the strong potential near the atom cores, and the evolution of the wave functions at large distances from the surface is determined primarily by kinetic energy, as discussed above.

None of these sources of inaccuracy can be expected to

greatly alter the results obtained; the calculation certainly gives a good overall representation of the true $\rho(\vec{r}, E_F)$. Nevertheless, the accumulation of numerical uncertainties dictates some care in drawing conclusions. 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.

IV. APPROXIMATE METHODS FOR STM

The unique strength of STM is that it is a truly local real-space probe of surface geometry. As such it can resolve isolated steps, defects, and impurities. The direct computation of electronic structure for such nonperiodic structures is not, in general, feasible. Conversely, STM provides little information for relatively smooth low-Miller-index surfaces, the only kind which are presently susceptible to accurate calculation of the vacuum charge. It is therefore imperative that methods for treating more complex structures be developed, if the theoretical analysis of STM results is to progress. Such methods need not be highly accurate to be useful.

A. Atom superposition for $\rho(\vec{r}, E_F)$

The calculated $\rho(\vec{r}, E_F)$ in Fig. 2 bears a strong resemblance to total charge densities, which have been em-

ployed in understanding helium scattering.^{14,15} It is known^{15,16} that the charge is sometimes well approximated by the superposition of atom charge densities,

$$\rho(\vec{r}) = \sum_R \phi(\vec{r} - \vec{R}), \quad (21)$$

where $\phi(r)$ is the charge density of the free atom, and R are atom positions, which need not form a periodic lattice. While this approximation has never been tested at the large distances relevant for STM, we show below that it is well worth trying.

A natural next assumption is

$$\rho(\vec{r}, E_F) \approx \rho(\vec{r})/E_0, \quad (22)$$

where $\rho(\vec{r})$ is the total charge. To estimate E_0 , we write

$$\rho(\vec{r}, E) \sim A \exp[-\hbar^{-1}(2mE)^{1/2}z], \quad (23)$$

where the variation of A with E is assumed small over the range contributing to $\rho(\vec{r})$. Then using

$$\rho(\vec{r}) = \int_{-\infty}^{E_F} \rho(\vec{r}, E) dE$$

and evaluating the integral, we find $E_0 \approx E_F/\kappa z$. (The derivation assumes $\kappa z \gg 1$, so $E^{1/2}$ can be expanded about E_F .) For example, if $\kappa z \approx 10$, then $E_0 \approx \frac{1}{2}$ eV. The precise value is unimportant, as discussed in Sec. II C. Note that the most drastic assumption is not (22), but rather the use of (21) at distances so great that only states near E_F contribute.

The use of (21) and (22), however crude, is not totally without justification. If the atom wave function is $\phi(r) \sim e^{-\kappa r}$, it is not hard to show¹⁶ that the asymptotic decay length of the corrugation is identical to (16) and (19), to first order in $(G/2\kappa)^2$. For Au the atom eigenvalue is close to the work function in the local density approximation, and so κ for the atom and for the surface are almost the same. Were this not the case, one could replace the true atom charge with a model charge having the decay length appropriate for the surface. The fractional error in the atom superposition estimate of the corrugation therefore approaches an asymptotic value rather than growing without bound for large z . The decay length for the charge is correct by construction, so the method gives an excellent estimate of the gap distance d .

Intuitively one expects the greatest success for noble metals such as Au, where directional bonding is minimal. In other cases, $\rho(\vec{r}, E_F)$ may show marked electronic structure effects, even when the total charge does not. For interesting examples, see Ref. 17 and Sec. V below.

B. Comparison with self-consistent results

We have repeated the calculation of Sec. III using the atom superposition approximation described above. Assuming as before a 9-Å tip radius, the calculated 2×1 and 3×1 corrugations are 0.30 and 0.93 Å, both about 30% less than in the self-consistent calculation. (While the accuracy of that calculation could not be calibrated quantitatively, better agreement would probably be fortuitous in any case.) This level of accuracy is enough to permit semiquantitative estimates of the STM

image to be expected for a given atomic geometry.

We can also compare these atomic results directly with experiment, as we did in Sec. III. Using R as a fitting parameter as before, we obtain excellent agreement with both of the two measured corrugation amplitudes, and a gap of $d = 6$ Å as before, by assuming $R = 4$ Å. Thus the atom superposition calculation is entirely consistent with the experimental data but leads to a different (and presumably less reliable) conclusion regarding the magnitude of R .

C. Structure sensitivity of STM

We are now in a position to calculate conveniently (albeit crudely) the STM image for an arbitrary geometry. By comparing the images expected from different geometries, we can judge what structural conclusion can (or cannot) be drawn from experimental data.

As the first example, we consider the Au(110) 3×1 surface. Binnig *et al.*⁴ inferred a geometry with two rows missing in the first layer and one in the second layer (see Fig. 2), to account for the deep observed corrugation of the 3×1 surface (1.4 Å versus 0.45 Å for the 2×1). While this inference is quite reasonable, we consider now a more quantitative test. We have calculated $\rho(\vec{r}, E_F)$ with approximations (21) and (22) for two 3×1 geometries, one with and one without a row missing in the second layer. The results are shown in Fig. 3. While the

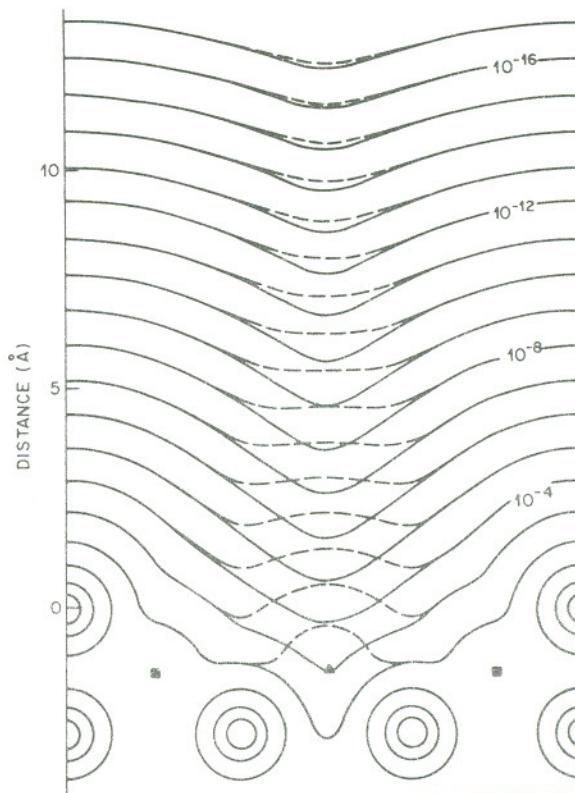


FIG. 3. Atom superposition charge density (a.u. $^{-3}$) for two possible geometries of Au(110) 3×1. Solid lines are for same geometry as in Fig. 2. Dashed lines are for geometry with no atom missing in second layer. Triangle shows site of atom (out of plane of figure) present only in latter case; squares show position of other out-of-plane atoms.

charge densities look radically different close to the surface, at a distance of 10 Å (where the corrugation is 1.4 Å) the corrugation amplitudes differ by only 15%; at 15 Å (appropriate for a 9-Å tip) the corrugation amplitudes differ by less than 5%. Realistically, even the 15% difference is far too little to reliably distinguish the two geometries. The greater corrugation in the 3×1 case (as compared to the 2×1) is attributable entirely to the greater surface lattice constant, which permits clearer resolution of the peaks and troughs. According to (19), a smaller surface lattice constant (as for the 2×1) results in an exponentially smaller corrugation at large distances.

As another sensitivity test, we compared the charge for the 2×1 surface to that for the (half-filled) first layer alone. At distances greater than 8 or 9 Å, the removal of the second and all subsequent layers has no noticeable effect. While atom superposition neglects the electronic changes for such a monolayer, the result at least tells us that the STM data carry no useful information whatever on the position of the first layer relative to the underlying substrate. A more methodical study of the relationship between geometry and charge density (and hence the STM image), also within the atom superposition approximation, is presented by Tersoff *et al.*¹⁶

V. SEMICONDUCTOR SURFACES

As noted above, for small voltages the tunneling occurs between states at E_F . At semiconducting surfaces, E_F lies near the conduction- or valence-band edge, depending on whether the doping is *n* or *p* type. The character of the states at E_F , and hence the form of $\rho(\vec{r}, E_F)$, may be drastically different for these two cases, giving correspondingly different STM images.

For low doping or high voltages, the voltage polarity rather than the doping may determine whether tunneling involves valence or conduction states. In the one reported example, Binnig *et al.*³ found that a measurable tunneling current for the Si(111) surface required a large voltage, over 2.5 V. These measurements were repeated with heavily doped Si samples, however, and comparable STM images were obtained with voltages in the 10-meV range used for Au.⁹ The high voltage in the first instance was probably developed across a non-Ohmic contact, a surface barrier due to band-bending, or both. We now have no reason to believe that the tunneling conductance in the STM imaging regime is significantly different for semiconductors and metals.

One of the simplest semiconductor surfaces is the cleaved GaAs(110) surface. The geometry of the 1×1 reconstruction is reasonably well established, and there are known to be no surface states in the band gap. We therefore use this surface to illustrate the difference between expected STM images for tunneling involving valence and conduction states. Figure 4 shows $\rho(\vec{r}, E_F)$ calculated for the GaAs(110) surface, based on the charge in states within 1 eV of the respective band edges. The total charge density is also shown. Far from the surface, the valence-edge charge looks quite similar to the total charge density. The charge is concentrated on the As atoms, which are raised above the Ga by the reconstruction. As a

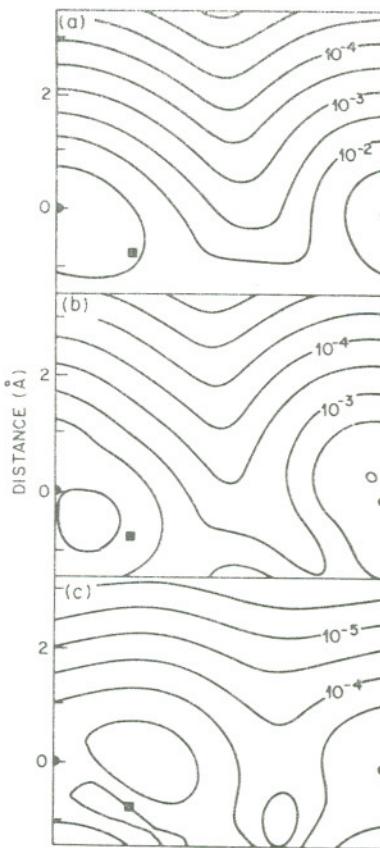


FIG. 4. Projected charge densities at the GaAs(110) surface, in a (110) plane midway between the Ga and As atoms, in units of bohr⁻³. The vacuum charge density is much smoother in the direction perpendicular to the figure. The three panels show (a) total charge; (b) charge in states within 1 eV of the valence-band edge; (c) charge in states within 1 eV of the conduction-band edge. Positions of the surface atoms, projected into the plane of the figure, are given by circles (As) and squares (Ga). Horizontal direction is ⟨001⟩, vertical is ⟨110⟩.

result, the image is well approximated by a superposition of As atom charge densities.

The conduction-band charge, however, looks quite different. Charge is concentrated on the Ga atoms; but these are lower than the As, with the net effect that both contribute comparably to the vacuum charge. The total corrugation is thus much smaller than for the valence charge, with the charge density peaking weakly above the Ga sites.

The surface lattice constant of GaAs(110) is less than 6 Å, which may be beyond the power of STM to resolve. However, the qualitative difference predicted for valence and conduction-band tunneling here should be observable in a wide variety of semiconducting surfaces.

VI. CONCLUSION

We have presented a simple theory for STM, which includes fully the detailed electronic structure of the surface and yet is computationally tractable. The tunneling current is found to be proportional to the surface LDOS at the position of the tip. The approximations made appear to introduce relatively little inaccuracy, except that

the tip is treated in a model way; even this approximation probably cannot be significantly improved until the microscopic structure of the tip is better understood.

The theory provides explicit expressions for the intrinsic spatial resolution and for the dependence of the tunneling current on tip size and position. When applied to the Au(110) surface, the theory agreed well with experiment. Moreover, the accuracy appeared to be limited by the computational implementation rather than by intrinsic factors.

Motivated both by the difficulty in carrying out accurate electronic structure calculations for STM, and by the need for a local technique for treating nonperiodic structures now observable with STM, we have proposed a simple approximate technique based on atom superposition, analogous to methods often used for the helium-diffraction problem. Despite its crudeness, this method has some analytic justification in terms of its asymptotic behavior, and gives good results for Au(110). It also pro-

vides a convenient way to test the sensitivity of STM to details of the surface structure, such as the presence or absence of a missing row in the second layer of Au(110) 3×1 . We conclude that STM is [at least for the Au(110) surface] quite insensitive to positions of atoms beyond the

first layer.

Finally, we have considered in a qualitative way the novel effects which may be observable in tunneling to semiconductor surfaces, where valence and conduction states have very different charge distributions. For the Au(110) surface, the LDOS (and hence the image) reflected surface topography in a relatively straightforward way. For GaAs(110), on the other hand, this was true only if tunneling involved the valence band (as for tunneling out of *p*-type GaAs). For conduction-band tunneling (into *n*-type GaAs), the image bore no simple relation to the topography, since the (lower) Ga atoms were emphasized by electronic structure effects.

We have concentrated on the use of STM to determine surface atomic structure, since this has been the main application to date. However, the fact that STM really measures the surface LDOS may be exploited in the future to give novel information about not only the topography of surfaces, but their electronic structure as well.

ACKNOWLEDGMENT

We are grateful to H. Rohrer and A. Baratoff for stimulating discussions.

¹Present address: IBM Thomas J. Watson Research Center, Yorktown Heights, New York 10598.

²G. Binnig, H. Rohrer, Ch. Gerber, and E. Weibel, *Appl. Phys. Lett.* **40**, 178 (1982).

³G. Binnig, H. Rohrer, Ch. Gerber, and E. Weibel, *Phys. Rev. Lett.* **49**, 57 (1982); G. Binnig and H. Rohrer, *Surf. Sci.* **126**, 236 (1983).

⁴G. Binnig, H. Rohrer, Ch. Gerber, and E. Weibel, *Phys. Rev. Lett.* **50**, 120 (1983).

⁵G. Binnig, H. Rohrer, Ch. Gerber, and E. Weibel, *Surf. Sci.* **131**, L379 (1983).

⁶C. B. Duke, *Tunneling in Solids*, Suppl. 10 of *Solid State Physics*, edited by F. Seitz and D. Turnbull (Academic, New York, 1969), p. 1.

⁷J. Tersoff and D. R. Hamann, *Phys. Rev. Lett.* **50**, 1998 (1983).

⁸N. Garcia, C. Ocal, and F. Flores, *Phys. Rev. Lett.* **50**, 2002 (1983).

⁹T. E. Feuchtwang, P. H. Cutler, and N. M. Miskovsky, *Phys. Lett.* **99A**, 167 (1983).

¹⁰H. Rohrer (private communication).

¹¹J. Bardeen, *Phys. Rev. Lett.* **6**, 57 (1961).

¹²I. K. Robinson, *Phys. Rev. Lett.* **50**, 1145 (1983), and references therein; L. D. Marks and D. J. Smith, *Nature (London)* **303**, 316 (1983).

¹³W. Moritz and D. Wolf, *Surf. Sci.* **88**, L29 (1979); Y. Kuk, *Bull. Am. Phys. Soc.* **28**, 260 (1983), and unpublished.

¹⁴D. R. Hamann, *Phys. Rev. Lett.* **46**, 1227 (1981).

¹⁵R. B. Laughlin, *Phys. Rev. B* **25**, 2222 (1982); M. Manninen, J. F. Nørskov, and C. Umrigar, *Surf. Sci.* **119**, L393 (1982).

¹⁶J. Tersoff, M. J. Cardillo, and D. R. Hamann (unpublished).

¹⁷P. J. Feibelman and D. R. Hamann, *Phys. Rev. Lett.* **52**, 61 (1984).

Tunneling microscopy of graphite in air

Sang-II Park and C. F. Quate

Edward L. Ginzton Laboratory, Stanford University, Stanford, California 94305

(Received 20 September 1985; accepted for publication 4 November 1985)

Images of graphite have been recorded with a scanning tunneling microscope operating in air at ambient pressure. The results, which are in agreement with theory and previous experiments in vacuum, confirm that with a surface such as graphite the tunneling microscopy in air can be used to examine the geometry of the surface with a resolution that is less than 2 Å.

The scanning tunneling microscope (STM) as introduced by Binnig *et al.*¹ provides us with a superb tool for surface science. The scanning mechanism, the control of vibration, and the formation of the tip are all problems that have been solved. It is now known that the STM can be used to define the spatial positions of atoms on surfaces of metals and semiconductors. Much of the definitive work has been carried out with the STM mounted in UHV stations in order to conform to the requirements of surface science. A notable exception is the work of Coleman *et al.*² in liquid nitrogen.

In this letter we report that the STM operating in air can generate images with atomic resolution. Hansma *et al.*³ have reported previously on their preliminary results with an STM operating in air. We have selected graphite as the sample crystal because it is inert and free from contamination. For bulk graphite the atomic structure is well known and electronic states are simple to calculate.^{4,5} However, the structure of the surface layer is less well known and it is important to gather information on these surfaces.

Graphite has been studied by others. Baro *et al.*⁶ have pointed out that this is an ideal substrate for the STM since large atomically flat net planes are available. The most definitive study was carried out by Binnig *et al.*⁷ in a UHV station. They have reported on the hexagonal structure of the surface layer. Their results are in agreement with the theory of Selloni *et al.*⁸ who used a model calculation to show that the height corrugation is voltage dependent.

Our experimental results are similar to those of Binnig *et al.*⁷ and agree with the predictions of Selloni *et al.*⁸ They demonstrate that our instrument can generate images with a resolution that is better than 2 Å. This suggests that the "effective tip" must consist of a single atom.

The experiments were a test run of our new STM that was designed to fit into a UHV system together with surface physics instruments. Our instrument was operated without special precautions in ambient pressure where the sample turn-around time is measured in minutes rather than hours. The condensation of water molecules on the tip does not appear to be a problem with the field strength used here.⁹ The images are easily repeated with different samples and different tips. We believe that the results are significant since we can record definitive information on surface structure in this environment. We wonder what it is about the graphite surface that allows us to do this.

The instrument that we are using has several innovative features that make it useful for our purposes, but it is derived from the design that was shown to us by the IBM group in

Zurich. The details of this STM will be described elsewhere; here we will only outline the essential features. The entire microscope is mounted on a 10-in. o.d. UHV flange which can be inserted into the vacuum chamber through an 8-in. i.d. port from the side. For vibration isolation, we use conventional double stage spring suspension with eddy-current damping. The stray magnetic field is reduced by encasing the entire microscope within a mu-metal sheet and mounting the magnets in the form of octopoles. The mu-metal cover can be opened for sample transfer. The sample holder is mounted on a sturdy translation stage and the two vibration isolation stages have a lock to provide rigid support during the sample transfer operation. The translation stage is driven by a UHV compatible stepper motor via a worm gear. It moves in steps of 1500 Å over a range of 5 mm. The scanner is formed with three orthogonal PZT bars each with a maximum range of 1 μm. It was calibrated and tested for orthogonality with laser interferometer. We used commercial tungsten tips with an approximate radius as determined by the SEM of 1000 Å.

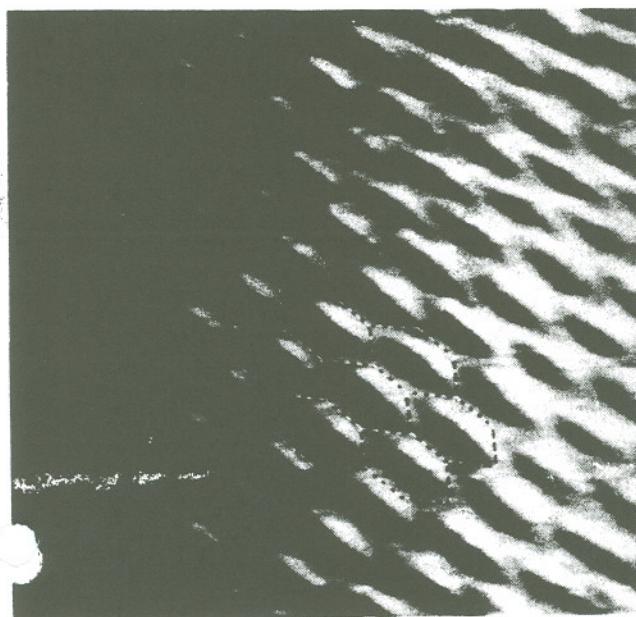
The thermal drift, one of the most important factors in successful tunneling microscopy, becomes more critical when the instrument is operated in air. In order to provide thermal insulation, we used multilayers of lint-free paper and vinyl wrap on the mu-metal cover of the microscope. The temperature of the unit was maintained at 27 °C with a temperature controller driving a built-in heater. The temperature variations were less than a milli degree over the period of several minutes.

The sample was highly oriented pyrolytic graphite as described by Moore.⁵ This layered material is easily cleaved. Alternatively, Scotch tape can be used to remove the top layers. We used these techniques to expose fresh layers and found that the quality of the image was unchanged after several days of exposure to air.

In Figs. 1 and 2 we display two images obtained from the raw data (a) and from the digital filtered data (b). All images span 20 Å in both *x* and *y*. The scanning speed along a line, the *x* axis, was 4 Hz in Fig. 1 and 10 Hz in Fig. 2. The tip with 10 mV was positive with respect to the graphite and the average value of the tunneling current was 5 nA. The variations in tunneling current during the scan provided the signal that we used for the image. In order to remove the noise from the raw data, we have used the Weiner optimum filter¹⁰ with the assumption that we have $1/f$ noise along the *x* axis and white noise along the *y* axis. This power spectrum of the noise was deduced from the momentum space spectrum of the data and the noise could be removed more efficiently



(a)



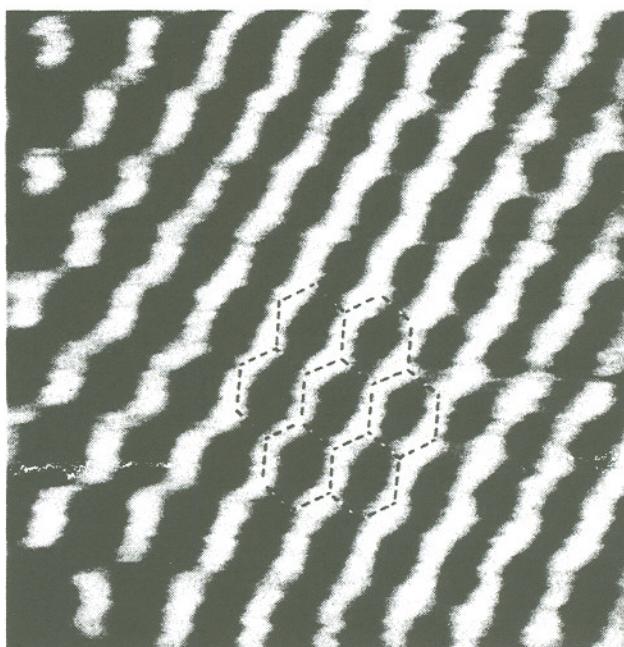
(b)

FIG. 1. STM image of the basal plane of graphite obtained with (a) the raw data and (b) the digital filtered data. The scanning rate is 4 Hz and the scale is $20 \text{ \AA} \times 20 \text{ \AA}$. The hexagonal grid (shown in Fig. 3) is superimposed so as to visualize the unit cell.

with this assumption.

We found that the quality of these images was not degraded with time and it could be recorded repeatedly with various scanning range and speed. The hexagonal pattern of the atomic structure is evident in all figures even though the image in Fig. 1 is elongated diagonally as a result of the residual thermal drift.

The dark regions represent the depressed hollows in the middle of the unit hexagon. The depth of this hollow is about 0.8 \AA . In Fig. 1 the bright regions just next to this hollow



(b)

FIG. 2. Another image of graphite via STM at 10 Hz scanning rate. The image in (b) is digital filtered and the original image appears in (a). Each image is $20 \text{ \AA} \times 20 \text{ \AA}$.

may look strange, but this is a result of the response time of the feedback circuit. It is somewhat slow compared to the scanning speed and the probing tip cannot accurately follow the contour when the sample surface changes abruptly. This enhances the tunneling current and produces some distortion in the image. We can decrease the response time of the feedback circuit by increasing the gain or changing the RC filter, but there is a limit imposed by the instabilities of the loop. The image in Fig. 2 with the higher scan rates was taken with the time constant in the feedback circuit increased to the point where the tip could not follow sharp

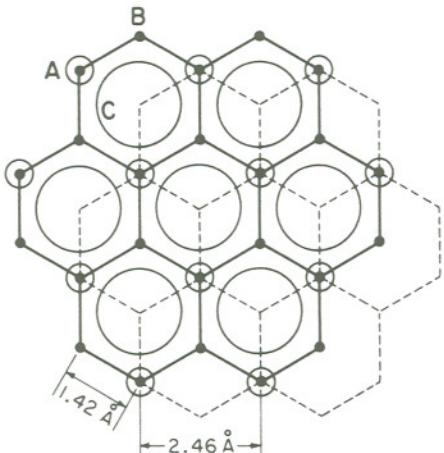


FIG. 3. Surface structure of graphite. There is a hollow in the middle of the unit hexagon (site C). The atom at site A has a neighboring atom in the lower layer (dashed hexagon). It is slightly higher than atom at site B which does not have a neighbor in the lower layer.

changes in the surface contour. This system is only suitable for flat surfaces such as graphite where the tip height can remain constant over the area of the image.

These contours are consistent with the calculations of Selloni *et al.*⁸ They discuss the differences in height for three different sites marked A, B, and C in Fig. 3. At site A the atoms in the top layer coincide with those in the underneath layer, whereas the atom at site B does not have a corresponding atom in the underneath layer. At site C there is an atom in the underneath layer, but nothing in the top layer. Their calculations indicate that for electrons near the Fermi energy (zero bias on the tip) the maximum corrugation, with a height difference near 1 Å, should occur between the peak at site A and the hollow at site C. This corresponds with what we measure. They also find that the atom at site A is about 0.15 Å above the atoms at site B. This difference is difficult to distinguish in our results where the estimated vertical resolution is 0.1 Å. At the higher voltages they predict a change in this pattern with a reversal in the positions of the peak and the depressions. We could not verify this since the surface corrugations in our images decreased when we increased the tip voltage. Neither could we find the surface states that are predicted¹¹ for energies above 2.5 or 3 V. This is not unexpected since the resolution in the STM decreases at the higher voltages because of the increase in gap spacing.

The high resolution came as a surprise since the tip with a macroscopic radius of 1000 Å was used "as is"¹² without resorting to any of the procedures of "forming" that have been used with instruments operating in vacuum. The calculations relating to the resolution of the STM¹³⁻¹⁶ suggest that our effective tip is a single atom. If this were not the case it is more probable, considering that the nearest neighbor distance in tungsten is 2.74 Å, that we would be imaging the tip rather than the graphite. This tip, where the single atom seems to be sitting at the edge of a terrace on the tungsten probe, may not be stable since we do see some variations in the texture of the image.

Finally, we want to point out that the remarkable effectiveness of the Weiner filter suggests that the noise must come from a physical process occurring in the tunneling region itself rather than random events in the electromechanical system. We have in mind the suggestion put forward by Gomer¹⁷ that a study of the fluctuations in the tunneling current would yield important information on the surface diffusion of adsorbed molecules. A related problem that has been studied is the fluctuation in field emission current. Kleint,¹⁸ who has measured the spectral density of fluctuation in field emission current, reports that the noise spectra has a $1/f$ character in the presence of surface diffusion and "reveals mainly shot noise" when emitters are cleaned. It may turn out that with the Weiner filter we are removing the effects of surface diffusion and simulating the high vacuum environment.

We want to acknowledge the support from R. Nemanich and A. W. Moore in providing us with the Graphite, from A. Bryant and D. Smith for their discussions in tunneling, and from Ch. Gerber who looked at our data and said "Yes, those are the atoms!" G. Binnig offered us words of wisdom and encouragement. S. Doniach was helpful in our discussions on digital filtering. This program is supported by the Defense Advanced Research Projects Agency, and in part by a grant from the Office of Naval Research and from the IBM Corporation.

¹G. Binnig, H. Rohrer, Ch. Gerber, and E. Weibel, Phys. Rev. Lett. **49**, 57 (1982).

²R. V. Coleman, B. Drake, P. K. Hansma, and G. Slough, Phys. Rev. Lett. **55**, 394 (1985).

³P. K. Hansma, R. Sonnenfeld, J. Schneir, B. Drake, and J. Hadzicki, Bull. Am. Phys. Soc. **30**, DI3, 309 (1985).

⁴G. S. Painter and D. E. Ellis, Phys. Rev. B **1**, 4747 (1970).

⁵A. W. Moore, in *Chemistry and Physics of Carbon*, edited by P. L. Walker, Jr. and P. A. Thrower (Dekker, New York, 1973 and 1981), Vol. 11, p. 69; Vol. 17, p. 233.

⁶A. Baro, R. Miranda, J. Alaman, N. Garcia, G. Binnig, H. Rohrer, Ch. Gerber, and J. L. Carrascosa, Nature **315**, 253 (1985).

⁷G. Binnig, H. Fuchs, Ch. Gerber, H. Rohrer, E. Stoll, and E. Tosatti, Europhys. Lett. **1**, XX (1986).

⁸A. Selloni, P. Carnevali, E. Tosatti, and C. D. Chen, in *Proceedings of the 17th International Conference on the Physics of Semiconductors*, edited by J. D. Chadi and W. A. Harrison (Springer, New York, 1985), p. 11.

⁹L. Swanson (private communication).

¹⁰K. R. Castleman, in *Digital Image Processing* (Prentice-Hall, Englewood Cliffs, NJ, 1979), p. 201.

¹¹M. Posternak, A. Baldereschi, A. J. Freeman, and E. Wimmer, Phys. Rev. Lett. **52**, 863 (1984).

¹²Model 7× Probe from Micromanipulator Sales and Service, Inc., Escondido, CA.

¹³J. Tersoff and D. Hamann, Phys. Rev. Lett. **50**, 1998 (1983).

¹⁴N. Garcia, C. Ocal, and F. Flores, Phys. Rev. Lett. **50**, 2002 (1983).

¹⁵E. Stoll, A. Baratoff, A. Selloni, and P. Carnevali, J. Phys. C **17**, 3073 (1984).

¹⁶E. Stoll, Surf. Sci. **143**, L411 (1984).

¹⁷R. Gomer, communicated at the IBM Europe Institute, Oberlech, Austria, July 1985.

¹⁸C. L. Kleint, Surf. Sci. **25**, I, 394; II, 411 (1971).

Observation of Atomic Corrugation on Au(111) by Scanning Tunneling Microscopy

V. M. Hallmark, S. Chiang, J. F. Rabolt, J. D. Swalen, and R. J. Wilson

IBM Almaden Research Center, San Jose, California 95120

(Received 23 July 1987)

This work presents the first real-space images which resolve individual close-packed metal atoms on the surface of a Au(111) thin film. Measurements made in both air and ultrahigh-vacuum environments reveal the same highly ordered, laterally extensive close-packed structure with a corrugation amplitude of $\approx 0.3 \text{ \AA}$ at a tip bias of $+30 \text{ mV}$ and a tunneling current of 3 nA . The possible role of surface electronic states on Au(111) in producing the corrugations imaged by the scanning tunneling microscope is discussed.

PACS numbers: 61.16.Di, 68.35.Bs

Resolution of individual close-packed metal atoms with the scanning tunneling microscope (STM) has not previously been achieved, although images of monolayer-height steps¹⁻³ and corrugation effects arising from surface reconstructions⁴⁻⁸ have been obtained. Close-packed surfaces such as Au(111) typically exhibit extensive terraces separated by broad steps of monolayer height, and the absence of corrugation along these terraces has been interpreted as atomic flatness of these metal surfaces.^{1,2} STM images of reconstructions of other low-index planes, e.g., (110) and (100) faces of Au^{4,5,8} and Pt,⁷ reveal long parallel hills and channels, again with no observable corrugation along the channel direction which could be assigned to individual atoms. This inability to resolve individual close-packed metal atoms has been attributed to an intrinsic smoothing of electron-density distributions at the surface due to the free-electron nature of metals.

In this Letter, we report the first observation of atomic corrugation for a close-packed metal surface. In contrast to the measurements referenced above, individual metal atoms are clearly resolved in our images of Au(111) thin films. These novel results are obtained in both air and ultrahigh-vacuum (UHV) environments and reveal a stable, inert surface with a highly ordered, close-packed gold lattice. The observation of a substantial corrugation, approximately 0.3 \AA , has important implications for future studies of clean and adsorbate-covered metal surfaces and raises significant questions about the theoretical understanding of electronic structure and corrugation in these systems.

Au(111) samples were prepared by epitaxial evaporative deposition⁹ of a 2500-\AA thickness of gold onto cleaved mica substrates maintained at 300°C . Deposition rates near 5 \AA/s were monitored by a quartz microbalance. The samples were allowed to cool radiatively in the evaporator system (base pressure $\approx 10^{-6} \text{ Torr}$) before transfer to one of two different STM systems.

Images in air were obtained with a "pocket-size" microscope¹⁰ including a stack of stainless-steel plates separated by Viton spacers, a piezoelectric tube tripod

for scanning the tunneling tip, and a micrometer for coarse approach of the sample to the tip. Samples were cemented to a quartz insulator mounted on the micrometer stage, and electrical contact was provided by the attachment of a wire to the gold surface with conductive epoxy. Tunneling images were obtained in the fast-scan mode¹¹ by scanning of the X and Y piezoelectric elements at rates of ≈ 40 and 0.04 Hz , respectively, while maintaining the tip at a constant average distance above the sample, so that the tunneling current was modulated by the corrugation of the surface electron density. Tip voltages were typically 50 to 500 mV , positive or negative, and the dc level of the tunneling current was maintained at 2 nA . Tunneling-current images were digitized and recorded directly onto video tape with an overall acquisition time of $\approx 10 \text{ s/frame}$.

Measurements were also performed in a multichamber UHV STM system with extensive surface preparation and analysis facilities,¹² previously employed to investigate structures of metal overlayers on silicon.¹³ The sample was fastened by tantalum clamps lined with gold foil to a tantalum carrier, then loaded via an airlock into the UHV system (base pressure $\leq 10^{-10} \text{ Torr}$). Auger-electron spectroscopy disclosed a substantial carbon contamination on the untreated Au(111) sample. Although diffuse low-energy electron diffraction (LEED) 1×1 patterns were acquired for the uncleaned sample, Ar^+ sputtering followed by annealing to 360°C via electron bombardment of the Ta carrier resulted in sharper LEED spots with lower background. This treatment also reduced contamination, as measured with Auger-electron spectroscopy, to levels less than 1% of a monolayer. Scanning electron microscopy disclosed a mosaic structure, with grain size varying from 0.5 to 0.02 mm depending on sample thermal history. Within such domains, previous field-emission scanning-electron-microscopy images of uncleaned samples had shown no resolvable structure $\geq 50 \text{ \AA}$. After preparation and analysis, the sample and carrier were loaded into the STM. Images from this instrument were obtained in the slow-scan mode¹⁴ (100 \AA/s) by our monitoring the

movement of the Z piezoelectric element while maintaining a constant tunneling current of 3 nA. Acquisition times were approximately 5 min per picture. Tip voltages again varied between 50 and 500 mV, both positive and negative.

The first atomic corrugation was observed for Au(111) with the air STM. The representative tunneling-current image in Fig. 1 reveals hexagonal close packing over dimensions of $\approx 25 \times 15 \text{ \AA}^2$.² The observed atomic spacing of $\approx 2.8 \pm 0.3 \text{ \AA}$, as calibrated against graphite measurements from this instrument, compares well with the known gold interatomic distance of 2.88 \AA . This close-packed structure persisted for tip translations of hundreds of angstroms in the lateral directions. The magnitude of the observed current modulation in this fast-scan mode was $\approx 10\%$ of the dc tunneling current. Images with this instrument were readily obtainable from four separate Au(111) samples, at various positions on each sample and with a series of different tunneling tips. The observed corrugation may represent either the metal surface itself or an ordered contamination layer since there is no independent assessment of sample cleanliness for this air measurement. Samples coated with controlled contamination layers, in the form of organic monolayer films deposited immediately after removal from the evaporator system, were also imaged in the air microscope. Films of cadmium arachidate transferred from a Langmuir-Blodgett trough¹⁵ and of octadecyltrichlorosilane self-assembled from solution¹⁶ onto Au(111) presented the same highly ordered close-packed STM

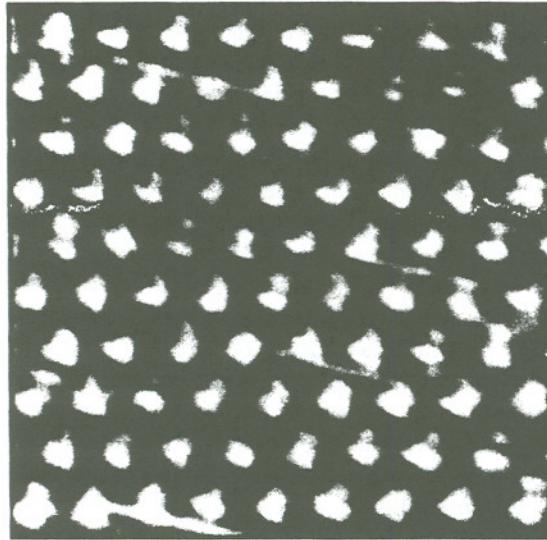


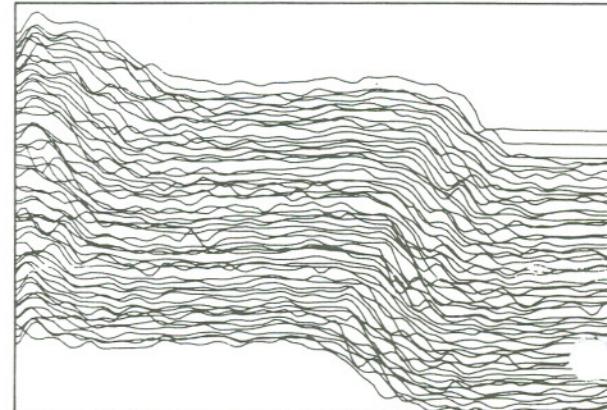
FIG. 1. STM tunneling-current image of a $\approx 25 \times 25 \text{ \AA}^2$ region of a Au(111) thin film measured in air, with an atomic spacing of $2.8 \pm 0.3 \text{ \AA}$. The current modulation is $\approx 10\%$ with a dc level of 2 nA and a tip bias of +50 mV. This figure has been corrected for thermal drift and piezoelectric creep. The diagonal streaks correspond to sudden jumps in the dc current level along scan lines in the fast direction.

images over most of their surfaces as observed for the uncoated gold samples. Since atmospheric contamination seems unlikely to result in well-ordered films comparable to those formed by the Langmuir-Blodgett or self-assembly methods, we believe that the observed images originate from the gold substrate. These results are consistent with reports of imaging of gold surfaces with no interference from submersion in paraffin oil.¹⁷

To address the contamination question unequivocally, a Au(111) thin-film sample was transferred into the UHV STM system. Our few attempts to image the surface corrugation in the slow-scan, constant-current mode before cleaning the surface were not successful because of tip instabilities, although fast-scan tunneling-current images indicated the presence of atomic corrugation.



(a)



(b)

FIG. 2. STM topographical line scans from a clean Au(111) surface measured in UHV, with a tip bias of +20 mV and 3-nA tunneling current. (a) In this $\approx 150 \times 90 \text{ \AA}^2$ region, a step of height $\approx 2.5 \text{ \AA}$ and width about 40 \AA is located at the right of the image. The evident low lateral resolution precludes interpretation of the lack of corrugation on the terraces as atomic flatness, even though the vertical resolution is $\approx 0.1 \text{ \AA}$. (b) An area of $\approx 90 \times 50 \text{ \AA}^2$, with the same vertical scale as in (a), showing two monolayer steps. Some corrugation is apparent along many scan lines, and the width of the step in the center of the image is much narrower than in (a), ranging from 5 to 10 \AA for various scan lines.

After sputtering and annealing of the sample to remove the carbon contamination observed by Auger-electron-spectroscopy, initial slow scans resembled those reported by other researchers: Large terraces, often extending for thousands of angstroms, were separated by broad steps of monolayer height, as shown in Fig. 2(a). Although step-terrace structures were evident, the poor lateral resolution at the step edge makes interpretations of the terraces as defect-free, atomically flat surfaces highly questionable. Improvements in tip resolution may sometimes be induced by the enlargement of the tip bias from a typical operating value of ≈ 50 mV to > 1 V for several seconds, or by an increase of the tunneling current by up to a factor of 3 for ≈ 30 s, followed by lateral movement of the tip to a new position. Indeed, this sharpening procedure resulted in observable corrugation along the terraces and in narrowing of the step width, as demonstrated by Fig. 2(b), although tip instabilities during this transition period appear to cause many line-to-line fluctuations. Nevertheless, after we attained still higher lateral and vertical resolution, clear atomic corrugation could be stabilized over reasonable scan areas, as in Fig. 3. Images of these close-packed features were obtained while we varied the lateral tip position for many hundreds of angstroms in any direction. At a tip bias of ≈ 50 mV, identical images were obtained with either polarity of the tip bias. The vertical corrugation was typically ≈ 0.3 Å, but ranged from 0.05 Å when first becoming resolvable to over 0.5 Å during short periods of high tip instability. In comparison to the constant-current corrugation amplitude of ≈ 0.3 Å at a tip bias of 30 mV and a tunneling current of 3 nA, successive constant-height measurements under identical conditions showed

the same 10% current modulation as obtained on the contaminated samples in air. Calibration of the Au(111) images against Si(111) 7×7 images obtained with this UHV microscope give a near-neighbor distance of 3.0 ± 0.3 Å. The absence of any measurable contamination, combined with the appropriate hexagonal packing and spacing for the Au(111) surface, conclusively supports assignment of the observed corrugation to the metal lattice itself.

These images represent the first observation of close-packed atomic corrugation for a metal surface. The expectation of little or no discernible corrugation for this surface because of the free-electron nature of gold has contributed to the oversight of such structure. Other researchers¹ have indeed reported flatness down to a few tenths of 1 Å other than at steps, but lateral resolution in these experiments, as monitored by step widths, may be insufficient to observe atomic corrugation. Also, our experimental conditions of ≈ 30 -mV tip bias and 3-nA tunneling current yield a tunneling resistance of 10^7 Ω, which is much smaller than the typical value of 10^9 Ω used in other STM experiments. Indeed, we were unable to image the atomic corrugation at a tunneling resistance of 2×10^8 Ω: With parameters of > 0.5 V and 2 nA the tip became unstable, and at 50 mV and 0.2 nA the corrugation was overwhelmed by noise. This order-of-magnitude increase in the tunneling resistance can be expected to correspond to about a 1-Å increase in the tip-to-surface distance, suggesting that the images presented here sample an electronic corrugation with spatial extent too small for previous STM studies of metals to have detected. We believe that the combination of high lateral and vertical resolution, low tunneling resistance, and excellent instrumental stability accounts for our novel observation of metal atomic corrugation.

Since our observed corrugation is much larger than the atomic corrugation for close-packed metals,¹⁸ a significant electronic enhancement must exist. Such electronic enhancements have been described previously for semiconductor¹⁹ and semimetal surfaces.²⁰ For gold, the presence of the bulk band gap at the *L* point²¹ suggests that a surface state near the Fermi level (E_F) may contribute strongly to the tunneling current. Experimental evidence for such a surface state on Au(111) with an energy peak at -0.5 V exists from both tunneling spectroscopy¹ and photoemission.²² Inverse photoemission studies²³ on Au(111) indicate the continuation of this surface state above E_F , so that tunneling into or out of this state under our experimental conditions is possible. Since the peak in the density of states due to the surface state is ≈ 0.3 eV from the bulk band edge, it seems reasonable to suggest that the state is localized to approximately atomic dimensions.

In summary, we have imaged individual close-packed metal atoms on Au(111) thin films by scanning tunneling microscopy performed with high lateral and vertical

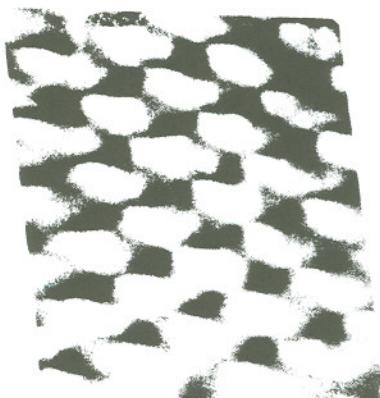


FIG. 3. STM topographical image of $\approx 12 \times 14$ -Å² region of a clean Au(111) thin film measured in UHV with high lateral and vertical resolution. The atomic spacing is 3.0 ± 0.3 Å. The corrugation is 0.3 Å at a tip bias of +30 mV and a tunneling current of 3 nA. This figure has been corrected for thermal drift and piezoelectric creep. We attribute the elongation of the observed atoms in the slow-scan direction to an asymmetric tip.

resolution and low tunneling resistance. Thin epitaxial metal films grown on mica represent a simple way to obtain (111) metal surfaces with very large, flat terraces separated by monolayer steps, and are amenable to cleaning and analysis in UHV with standard surface techniques. The similarity between images obtained in air, both with and without adsorbed organic monolayer films, and from clean samples in UHV indicates that the Au(111) surface is valuable as an alternative system to highly oriented pyrolytic graphite to obtain atomic-resolution images in air or fluid ambients for purposes of instrumental calibration. In response to this unexpected success in the utilization of the STM to image individual atoms on metal surfaces, STM studies of metals and chemical reactions occurring at their surfaces should receive renewed attention.

We would like to thank I. P. Batra, D. M. Eigler, R. M. Feenstra, F. Herman, N. V. Smith, and J. Tersoff for helpful theoretical discussions, and G. Borges, J. Duran, and A. Leone for assistance in sample preparation. We are grateful to M. Salmeron for providing preprints of Refs. 2 and 3. Two of the authors (V.M.H. and J.F.R.) are pleased to acknowledge partial support for this work from the Chemistry Division of the Office of Naval Research.

¹W. J. Kaiser and R. C. Jaklevic, Surf. Sci. **181**, 55 (1987), and **182**, L227 (1987).

²M. Salmeron, D. S. Kaufman, B. Marchon, and S. Ferrer, to be published.

³M. Salmeron, B. Marchon, S. Ferrer, and D. S. Kaufman, to be published.

⁴G. Binnig, H. Rohrer, Ch. Gerber, and E. Weibel, Surf. Sci.

131, L379 (1983).

⁵G. K. Binnig, H. Rohrer, Ch. Gerber, and E. Stoll, Surf. Sci. **144**, 321 (1984).

⁶A. M. Baro, G. Binnig, H. Rohrer, Ch. Gerber, E. Stoll, A. Baratoff, and F. Salvad, Phys. Rev. Lett. **52**, 1304 (1984).

⁷R. J. Behm, W. Hosler, E. Ritter, and G. Binnig, Phys. Rev. Lett. **56**, 228 (1986).

⁸Y. Kuk and P. J. Silverman, Appl. Phys. Lett. **48**, 1597 (1986).

⁹K. Reichelt and H. O. Lutz, J. Cryst. Growth **10**, 103 (1971).

¹⁰Ch. Gerber, G. Binnig, H. Fuchs, O. Marti, and H. Rohrer, Rev. Sci. Instrum. **57**, 221 (1986).

¹¹A. Bryant, D. P. E. Smith, and C. F. Quate, Appl. Phys. Lett. **48**, 832 (1986).

¹²S. Chiang and R. J. Wilson, J. Vac. Sci. Technol. A (to be published).

¹³R. J. Wilson and S. Chiang, Phys. Rev. Lett. **58**, 369 (1987), and **58**, 2575 (1987).

¹⁴G. Binning, H. Rohrer, Ch. Gerber, and E. Weibel, Phys. Rev. Lett. **49**, 57 (1982).

¹⁵G. L. Gaines, *Insoluble Monolayers at Liquid-Gas Interfaces* (Interscience, New York, 1966).

¹⁶H. O. Finklea, L. R. Robinson, A. Blackburn, B. Richter, D. Allara, and T. Bright, Langmuir **2**, 239 (1986).

¹⁷B. Drake, R. Sonnenfeld, J. Schneir, P. K. Hansma, G. Slough, and R. V. Coleman, Rev. Sci. Instrum. **57**, 441 (1986).

¹⁸I. P. Batra, J. A. Barker, and D. J. Auerbach, J. Vac. Sci. Technol. A **2**, 943 (1984).

¹⁹R. M. Feenstra, W. A. Thompson, and A. P. Fein, Phys. Rev. Lett. **56**, 608 (1986).

²⁰J. Tersoff, Phys. Rev. Lett. **57**, 440 (1986).

²¹N. E. Christensen and B. O. Seraphin, Phys. Rev. B **4**, 3321 (1971).

²²Z. Hussain and N. V. Smith, Phys. Lett. **66A**, 492 (1978).

²³D. P. Woodruff, W. A. Royer, and N. V. Smith, Phys. Rev. B **34**, 764 (1986).

Tunneling microscopy of 2H-MoS₂: A compound semiconductor surface

M. Weimer, J. Kramar, C. Bai, and J. D. Baldeschwieler

California Institute of Technology, Pasadena, California 91125

(Received 8 September 1987; revised manuscript received 23 December 1987)

Molybdenum disulfide, a layered semiconductor, is an interesting material to study with the tunneling microscope because two structurally and electronically different atomic species may be probed at its surface. We report on a vacuum scanning tunneling microscopy study of 2H-MoS₂. Atomic resolution topographs and current images show the symmetry of the surface unit cell and clearly reveal two distinct atomic sites in agreement with the well-known x-ray crystal structure.

Over the last several years, scanning tunneling microscopy (STM) has emerged as a powerful tool for the real-space visualization of surface structures on an atomic scale.¹ While the technique has enjoyed enormous success in elucidating several classical problems of geometric and electronic structure on homogeneous metallic and semiconducting surfaces, there remains a central problem with respect to the study of compound surfaces: to what extent can the chemical identity of individual atomic species be ascertained? Recently, progress in this area was made by Feenstra, Stroscio, Tersoff, and Fein² who succeeded in selectively imaging either Ga or As atoms in voltage-dependent STM scans of GaAs(110), but so far this is the only case where such selectivity has been achieved. In an effort to add to our understanding of compound surface imaging by tunneling microscopy, we have undertaken a vacuum STM study of 2H-MoS₂, a layered semiconducting transition-metal dichalcogenide.

Layered compounds in general, and the transition-metal dichalcogenides in particular, have played a prominent role in tunneling microscopy. Much of the emphasis has been on metallic members of this family, such as TaS₂ and TaSe₂, where the surface charge rearrangement accompanying charge-density-wave formation has been observed in the low-temperature work of Coleman *et al.*³ The atomic corrugation of NbSe₂ has also been observed at room temperature, by Bando *et al.*, who distinguished three different sites which they interpreted in terms of the two nonequivalent halves of the surface unit cell.⁴ Those results, obtained at tunnel gap resistances of $3 \times 10^5 \Omega$, or less, showed unusually large corrugation amplitudes which the authors attributed to strong forces between the probe tip and sample surface.⁵

Molybdenum disulfide is a new surface for STM studies. Since it is semiconducting, one expects its surface electronic wave functions to be preferentially localized over specific atoms or bonds. In this paper we show that it is possible to distinguish two distinct atomic sites at the surface of 2H-MoS₂ by tunneling microscopy, both in the conventional constant-current, variable-height mode of operation (without obtaining anomalously large corrugation amplitudes) as well as in the variable-current, constant-height mode.⁶

The crystal structure of molybdenum disulfide as determined by x-ray diffraction⁷ is shown in Fig. 1. From the point of view of the (001) surface projection, the top lay-

er is a hexagonal lattice of sulfur atoms with lattice constant 3.16 Å. Immediately below this plane is an identical hexagonal lattice of molybdenum atoms laterally displaced relative to the top layer. The position of the Mo atoms reduces the sixfold S planar rotational symmetry to threefold symmetry, producing a diamond-shaped surface unit cell with a Mo atom centered in one triangular half and a hollow located in the other. Viewed from the side, one finds a repeated structure of two, alternating, S-Mo-S sandwiches separated by a van der Waals gap along which cleavage occurs.

MoS₂ in the 2H polytype is naturally occurring and mineralogical samples come in both *n*- and *p*-type with varying levels of doping. The highly anisotropic layer structure produces a substantial reduction in electrical conductivity perpendicular to the planes relative to that exhibited in plane.⁸ Though we have not always found it necessary to do so, the data reported here were obtained

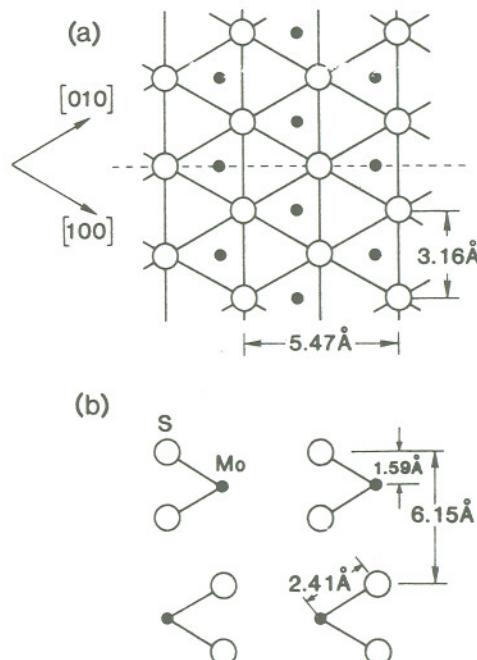


FIG. 1. (a) The (001) surface structure and (b) the (110) cross section [along the dashed line in (a)] of 2H-MoS₂, as determined by x-ray diffraction.

with crystals whose edges were coated with silver paint in order to take advantage of this higher in-plane conductivity. The doping level was not characterized, so the contribution of the sample spreading resistance to the total gap resistances we measure cannot be estimated.

The tunnel microscope used in these studies is patterned after the IBM Zurich "pocket-size" STM,⁹ and is supported by a dual-stage spring suspension for additional vibration isolation. Scanning, data acquisition, and image display are all done under microcomputer control. After mounting and cleaving in air, our samples were pumped down to approximately 2×10^{-8} Torr without baking. All of the scans illustrated here were performed with an electrochemically etched W tip, but similar images have been obtained with other W tips and with electrochemically etched Au tips, on several different MoS₂ samples.

Figure 2 shows an atomic resolution constant-current topograph of the surface of 2H-MoS₂ obtained at positive sample bias so that electrons tunneled from the tip into occupied states of the sample. The image was acquired in 24 sec at a scan rate of 125 Å/sec and subsequently low-pass filtered. A centered hexagon of bright spots is evident in the figure, as are three distinct sites corresponding to the two constituent atom types and a surface hollow. An overlay of the surface unit cell is illustrated for comparison with Fig. 1. We see four bright spots at the corners, a secondary site in one half of the unit cell and a hollow in the other half, in agreement with the x-ray crystal structure.

Figure 3 displays the change in surface height along cross-sectional cuts indicated by the dashed lines in Fig. 2. Figure 3(a) shows the height versus distance along a [1̄10] cell diagonal with a major peak, a secondary peak, and a hollow clearly visible. The peak-to-hollow corrugation amplitude is roughly 0.5 Å while the difference in height between the major and secondary peaks is approximately half as large. We note that the corrugation inferred from tunneling is much smaller than the 1.59 Å separating atomic centers in the top layer from those in a second. Figure 3(b) depicts the corrugation along a

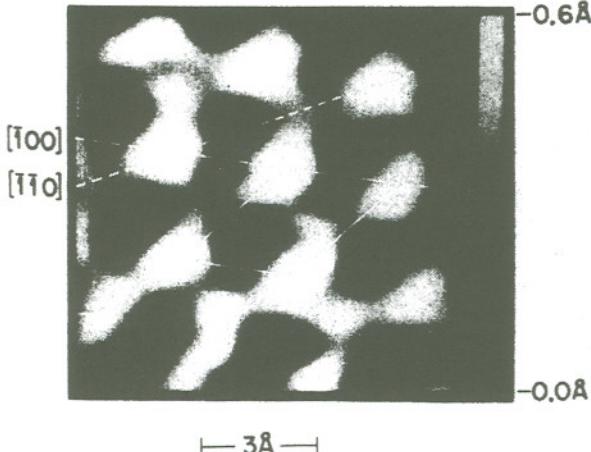


FIG. 2. Topographic image of the surface of 2H-MoS₂ at +140 mV sample bias and 2-nA tunnel current. Grey scale indicates vertical range.

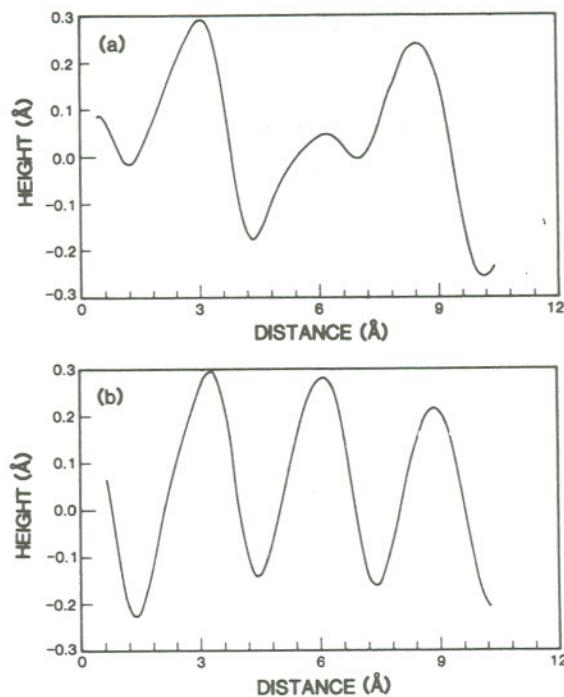


FIG. 3. Interpolated cross sections through the data of Fig. 2: (a) along the [1̄10] cell diagonal and (b) the [1̄00] cell edge.

[1̄00] cell edge, where only a simple sinusoidal modulation is evident. The peak-to-hollow amplitude measured there is almost as large as that along the [1̄10] cell diagonal.

Figure 4 shows an atomic resolution current-contrast image of the same surface. Again, tunneling is into the sample. The image was acquired in 1.6 sec at a scan rate of 1350 Å/sec and then low-pass filtered. The vertical scale reflects the ac variation in tunnel current relative to

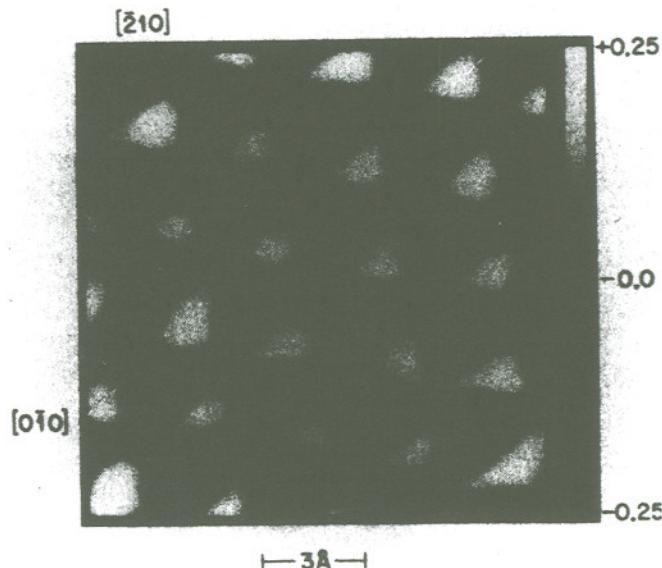


FIG. 4. Variable current image of the surface of 2H-MoS₂ at +250 mV sample bias and 2 nA mean tunnel current. Grey scale indicates ac current modulation relative to mean current.

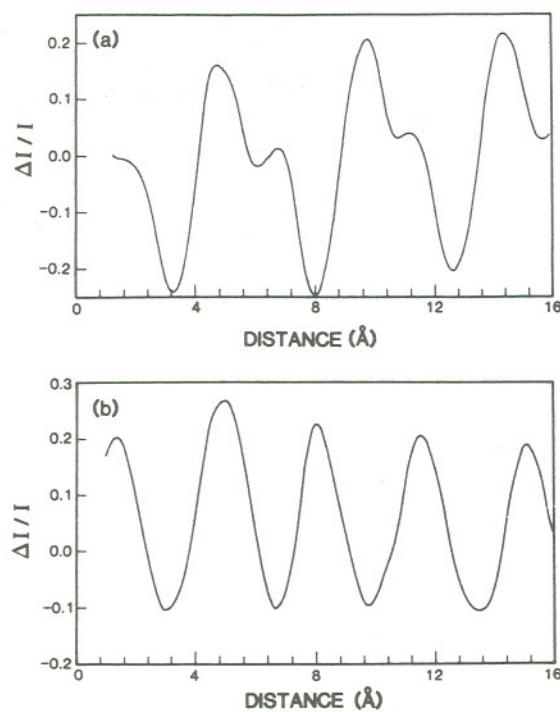


FIG. 5. Interpolated cross sections through the data of Fig. 4: (a) along the $[2\bar{2}0]$ cell diagonal and (b) the $[0\bar{1}\bar{0}]$ cell edge.

its mean value.

The centered hexagonal pattern of bright spots is especially vivid, and the definition in the three distinct sites superior to that of Fig. 2. This difference presumably arises from the greater external noise immunity associated with the fast scan mode⁶ and the complete absence of thermal drift. The bright spots and hollows appear roughly circular whereas threefold symmetry is apparent in the secondary sites. An overlay of the familiar surface unit cell shows the expected structure. When we plot the variation in tunnel current relative to its mean value along the directions indicated in Fig. 4 we find a repeated pattern of three sites along the cell diagonal, and sinusoidal modulation along the cell edge, as illustrated in Fig. 5. Along the $[2\bar{1}0]$ cell diagonal the total peak-to-hollow corrugation represents a variation of 50% in $\Delta I / I$ while along the $[0\bar{1}\bar{0}]$ cell edge it is less, as required by symmetry.

The lateral distance scales in Figs. 3 and 5 are those inferred from the nominal expansion coefficients of our ceramics and none of the figures has been corrected for piezo nonorthogonality. As a result, the unit-cell overlays of Figs. 2 and 4 appear slightly distorted while the peak spacings and hollow locations in Fig. 5 all vary by about $\pm 10\%$ from the correct values indicated in Fig. 1.

It was our experience that we did not usually obtain such high-quality images of $2H\text{-MoS}_2$ on initial approach of the tip to the sample, but had to spend some time scanning over the surface first. Images appeared more easily and consistently in the current-contrast mode than in the topographic mode, but atomic resolution was by no means automatic in either case. This may have been due in part

to our particular samples, which displayed a high degree of surface roughness relative to graphite, making the task of atomic resolution imaging more difficult. Qualitatively similar results were observed at negative sample bias.

We cannot rule out the possibility that a multiple-atom tip¹⁰ is responsible for the structure we report on here, as a number of different images were observed as a function of time and presumably tip condition. Nevertheless, the Fourier transform of Fig. 4 produces the sixfold symmetric power spectral density expected for this surface. Information on the two sites per unit cell is contained in the relative phases associated with each of these six directions, reducing the apparent sixfold symmetry to the appropriate threefold one. The symmetry of the Fourier transform together with the accurate position of the second site within the unit cell requires a precise spatial phase relationship between hypothesized multiple tips. In our view, such an artifice is rendered less likely by the frequency with which this configuration is observed independent of probe-tip composition.

A more straightforward interpretation of our results directly identifies the two sites in these images with the two surface atomic species. Simple geometric considerations would then imply the brightest points are due to sulfur atoms, which are closest to the tip, while the secondary peaks arise from molybdenum atoms in the second layer. The predominant factor governing tunneling microscope images, however, is the contribution made by each atom's valence orbitals to the position- and energy-dependent density of states (near E_F) at distances above the surface normally associated with tunneling.¹¹ Numerous theoretical investigations of the band structure of $2H\text{-MoS}_2$ (Ref. 12) point to strongly covalent bonding with a substantial Mo $4d$ contribution at the top of the valence band and bottom of the conduction band. Thus, one cannot *a priori* ignore the possibility that Mo $4d$ levels, rather than S $3p$ levels, are primarily responsible for the tunnel current.

In conclusion, it appears possible to clearly resolve two chemically and structurally distinct atomic sites in a layered compound by tunneling microscopy. Since there are no intrinsic surface states on $2H\text{-MoS}_2$ (Ref. 13) one should be able to directly probe the differences between tunneling into the conduction band (as we have done here) or out of the valence band, as first suggested by Tersoff and Hamann¹¹ and realized in the experiments of Feenstra *et al.*² It will be interesting to see whether such experiments, combined with appropriate calculations, can uniquely establish the position of the transition-metal atom in this and other layered semiconductors.

The authors are indebted to Ch. Gerber for guidance and to W. J. Kaiser for constant advice and encouragement. We have benefited from stimulating conversations with D. Denley, T. Coley, and W. A. Goddard III. This work was supported by the Office of Naval Research, under Contract No. N00014-86-K-0214, by the National Institutes of Health, under Contract No. RO1 GM37226-01, and by a gift from the Shell Companies Foundation.

- ¹G. Binnig and H. Rohrer, IBM J. Res. Dev. **30**, 355 (1986).
- ²R. M. Feenstra, J. A. Stroscio, J. Tersoff, and A. P. Fein, Phys. Rev. Lett. **58**, 1192 (1987).
- ³R. V. Coleman, W. W. McNairy, C. G. Slough, P. K. Hansma, and B. Drake, Surf. Sci. **181**, 112 (1987); C. G. Slough, W. W. McNairy, R. V. Coleman, B. Drake, and P. K. Hansma, Phys. Rev. B **34**, 994 (1986); R. V. Coleman, B. Drake, P. K. Hansma, and C. G. Slough, Phys. Rev. Lett. **55**, 394 (1985).
- ⁴H. Bando, H. Tokumoto, W. Mizutani, K. Wantanabe, M. Okano, M. Ono, H. Murakami, S. Okayama, Y. Ono, S. Wakiyama, F. Sakai, K. Endo, and K. Kajimura, Jpn. J. Appl. Phys. **26**, L41 (1987).
- ⁵J. M. Soler, A. M. Baro, N. Garcia, and H. Rohrer, Phys. Rev. Lett. **57**, 444 (1986).
- ⁶A. Bryant, D. P. E. Smith, and C. F. Quate, Appl. Phys. Lett. **48**, 832 (1986).
- ⁷R. G. Dickinson and L. Pauling, J. Am. Chem. Soc. **45**, 1466 (1923).
- ⁸A. J. Grant, T. M. Griffiths, G. D. Pitt, and A. D. Yoffe, J. Phys. C **8**, L17 (1975).
- ⁹Ch. Gerber, G. Binnig, H. Fuchs, O. Marti, and H. Rohrer, Rev. Sci. Instrum. **57**, 221 (1986).
- ¹⁰H. A. Mizes, Sang-il Park, and W. A. Harrison, Phys. Rev. B **36**, 4491 (1987).
- ¹¹J. Tersoff and D. R. Hamann, Phys. Rev. B **31**, 805 (1985); Phys. Rev. Lett. **50**, 1998 (1983).
- ¹²E. Doni and R. Girlanda, in *Electronic Structure and Electronic Transitions in Layered Materials*, edited by V. Grasso (Reidel, Dordrecht, 1986), p. 72; R. Coehoorn, C. Haas, J. Dijkstra, C. J. F. Flipse, R. A. deGroot, and A. Wold, Phys. Rev. B **35**, 6195 (1987).
- ¹³A. Koma and K. Enari, in *Physics of Semiconductors—1978*, edited by B. L. H. Wilson, IOP Conference Series No. 43 (Institute of Physics, Bristol, 1979), p. 895.

Theory and Application for the Scanning Tunneling Microscope

J. Tersoff and D. R. Hamann

Bell Laboratories, Murray Hill, New Jersey 07974

(Received 17 March 1983)

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.

PACS numbers: 68.20.+t, 07.80.+x, 73.20.-r, 73.40.Gk

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 $M_{\mu\nu}$. Bardeen⁷ 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-

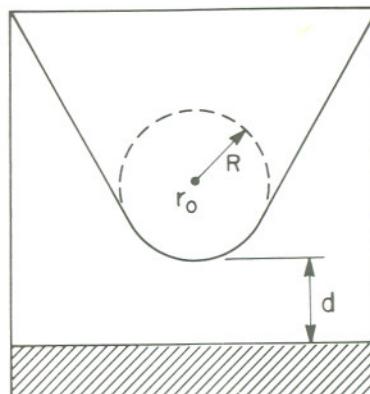


FIG. 1. Schematic picture of tunneling geometry. Probe tip has arbitrary shape but is assumed locally spherical with radius of curvature R , where it approaches nearest the surface (shaded). Distance of nearest approach is d . Center of curvature of tip is labeled \vec{r}_0 .

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

$$\begin{aligned}\psi_\nu = \Omega_s^{-1/2} \sum_G a_G \exp[-(k^2 + |\vec{k}_\parallel + \vec{G}|^2)^{1/2} z] \\ \times \exp[i(\vec{k}_\parallel + \vec{G}) \cdot \vec{x}],\end{aligned}\quad (4)$$

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

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

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

$$\begin{aligned}(kr)^{-1} e^{-kr} = \int d^2q b(\vec{q}) \exp[-(k^2 + q^2)^{1/2}|z|] \\ \times \exp(i\vec{q} \cdot \vec{x}),\end{aligned}\quad (6)$$

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

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

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

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

$$\begin{aligned}I = 32\pi^3 \hbar^{-1} e^2 V \varphi^2 D_t(E_F) R^2 k^{-4} e^{2kR} \\ \times \sum_\nu |\psi_\nu(\vec{r}_0)|^2 \delta(E_\nu - E_F),\end{aligned}\quad (9)$$

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 very 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 $2k^{-1} \approx 1.6 \text{ \AA}^{-1}$, if $R + d \approx 15 \text{ \AA}$ (as in the example below), the resolution is about 5 \AA , so that the 8 \AA periodicity of Au(110) (2×1) is resolved. It is important to note that for periodic structures the measured corru-

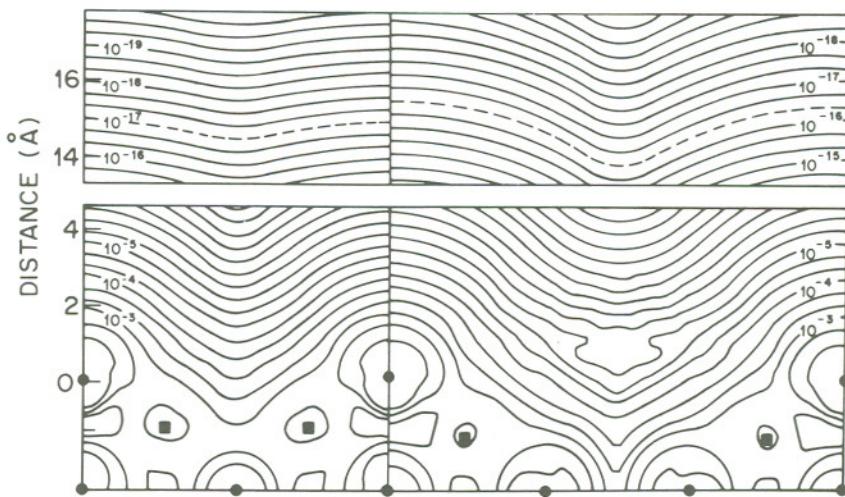


FIG. 2. Calculated $\rho(\mathbf{r}; E_F)$ for Au(110)(2×1) (left) and (3×1) (right) surfaces. Figure shows (1̄10) plane through outermost atoms. Positions of nuclei are indicated by solid circles (in plane) and squares (out of plane). Contours of constant $\rho(\mathbf{r}; E_F)$ are labeled in units 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 puzzle⁴ 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 $\rho(\mathbf{r}; E_F)$ 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 (1̄10) 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 $\rho(\mathbf{r}; E_F)$ 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 $\rho(\mathbf{r}; E_F)$ for Au(110). Since the actual tip geometry is not known, we consider a tip radius $R = 9 \text{ \AA}$, so that (10) predicts a (2×1) corrugation of 0.45 Å at tunneling resistance⁶ $10^7 \Omega$ 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 consistent⁶ 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.

We are grateful to H. Rohrer for making results available prior to publication, and to him and A. Baratoff for stimulating discussion and

valuable suggestions.

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

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

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

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

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

⁶H. Rohrer, private communication.

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

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

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

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

triangles. Reasonable agreement is thus obtained between the experimental form factor values and those computed assuming the unpaired electrons to be in a state of pure e_g symmetry.

It is important to observe that the scale factor, k , here determined experimentally, is such as to make the unpaired $3d$ charge distribution for Ni^{++} much more compact in the solid than it is for the free atom. This is in contrast to the case of Mn^{++} where experiments¹³ show that the charge distribution is expanded in the solid.

These experimental results may be compared with the recent calculations of Watson and Freeman⁵ for the Ni^{++} ion. These Hartree-Fock self-consistent field calculations allow the wave functions of electrons with opposite spins to have different radial dependencies (spin polarization) and lead to a contraction of the charge distribution (both for the free-atom case and the case where the Ni^{++} ion is placed in an octahedral array of point charges).¹⁴ Unfortunately, the magnitude of the contraction is much too small to explain the observations reported here. Nevertheless, the important fact that the relation of the observed f_S to the free-atom f_S is just opposite for the cases of Ni^{++} and Mn^{++} would lead one to look for the origin of this effect in the outstanding difference between the two ions: namely their differing spin configurations. These experimental results also serve to suggest that whereas effects such as spin polarization and crystalline environment have a large influence on f_S , their effect on f_A is small.

The author wishes to thank Professor E. Uchida and Dr. Y. Nakazumi for the single crystal of

NiO .

[†]Performed under the auspices of the U. S. Atomic Energy Commission.

¹R. Nathans, C. G. Shull, G. Shirane, and A. Andresen, *J. Phys. Chem. Solids* **10**, 138 (1959).

²R. Nathans and A. Paoletti, *Phys. Rev. Letters* **2**, 254 (1959).

³S. J. Pickart and R. Nathans, *Suppl. J. Appl. Phys.* **31**, 372S (1960).

⁴R. Nathans, S. J. Pickart, and H. A. Alperin, *Bull. Am. Phys. Soc.* **5**, 455 (1960).

⁵R. E. Watson and A. J. Freeman, *Phys. Rev.* **120**, 1125 (1960); R. E. Watson and A. J. Freeman, *Phys. Rev.* **120**, 1134 (1960).

⁶R. Nathans, *Suppl. J. Appl. Phys.* **31**, 350S (1960).

⁷H. A. Alperin, U. S. Naval Ordnance Laboratory Report NAVORD-6907, June, 1960 (unpublished).

⁸J. D. Dunitz and L. E. Orgel, *J. Phys. Chem. Solids* **3**, 20 (1957).

⁹We are assuming here that the radial part of the wave functions is the same for all eight electrons so that the three electrons of t_{2g} symmetry with their spins in one direction "cancel" the three electrons of like symmetry with oppositely directed spins.

¹⁰R. J. Weiss and A. J. Freeman, *J. Phys. Chem. Solids* **10**, 147 (1959).

¹¹R. E. Watson and A. J. Freeman, *Acta Cryst.* (to be published).

¹²The form factor of Ni^{++} in KNiF_3 has been measured up to the value $\sin\theta/\lambda = 0.32$ by V. Scatturin, L. Corliss, N. Elliott, and J. Hastings, *Acta Cryst.* (to be published). Their result, which is for a powder measurement, is in approximate agreement with the curve presented here.

¹³J. M. Hastings, N. Elliott, and L. M. Corliss, *Phys. Rev.* **115**, 13 (1959).

¹⁴It should be noted that Mn^{++} which has all its spins in the same direction has no spin polarization effect.

TUNNELLING FROM A MANY-PARTICLE POINT OF VIEW*

J. Bardeen

University of Illinois, Urbana, Illinois

(Received December 16, 1960)

Giaever¹ and more recently Nicol, Shapiro, and Smith² have observed the tunnelling current flowing between two metals separated by a thin oxide layer. The most interesting results are obtained when one or both of the metals are superconducting, in which case they find direct evidence for a gap in the quasi-particle spectrum of the superconductor. They were able to account for the data quantitatively on the assumption that the only relevant factor is the density of states in energy.

This is to be expected if the transition probability for transfer of an electron from one side to the other is given by the familiar expression $(2\pi/\hbar)|M|^2\rho_f$, where M is the matrix element and ρ_f the energy density of final states, and if it is further assumed that M can be treated as a constant. It is implied that M is not only independent of energy for the small energy differences involved, but is also unchanged when the metal goes from normal to superconducting.

However, it is not immediately obvious that these assumptions are justified. We give here a discussion of tunnelling from a many-particle point of view and show that it is plausible to treat M as a constant in the interpretation of the experiments.

We suppose that the barrier extends from x_a to x_b , with metal a to the left of x_a and metal b to the right of x_b . Consider two many-particle states of the entire system, Ψ_0 and Ψ_{mn} , which differ in the transfer of an electron from a to b . We suppose that Ψ_0 and Ψ_{mn} may be defined in terms of quasi-particle occupation numbers of metals a and b , so that Ψ_{mn} differs from Ψ_0 in the transfer of an electron from state m in a to state n in b , all other occupation numbers remaining the same. Of course m may correspond to a normally occupied state, in which case there will be a hole in m in Ψ_{mn} .

The quasi-particles do not correspond to plane waves, but to waves which are reflected at the barrier and which drop exponentially with distance into the barrier region. For example, in the free-electron approximation for the normal state the wave function would be of the form (WKB approximation):

$$\psi_m = C p_x^{-1/2} e^{i(p_y y + p_z z)} \sin(p_x x + \gamma), \quad x < x_a \quad (1a)$$

$$\begin{aligned} \psi_m = \frac{1}{2} C |p_x|^{-1/2} e^{i(p_y y + p_z z)} \\ \times \exp\left(-\int_{x_a}^x |p_x| dx\right), \quad x_a < x < x_b \end{aligned} \quad (1b)$$

where $C = (2p_x/L)^{1/2}$ is a normalization constant, and in the barrier region, $|p_x| = (2\mu U - p_y^2 - p_z^2)^{1/2}$, where $U(x)$ is the potential energy. We have taken units such that $\hbar = 1$. Beyond x_b , we assume that the wave function representing the electron in state m drops smoothly to zero, instead of oscillating, so that it is not a good solution for $x > x_b$.

Thus we assume that Ψ_0 is a solution of the Schrödinger equation with energy W_0 for $x < x_b$, but there is a region to the right of x_b where it is not a good solution. Similarly, we assume that Ψ_{mn} with energy W_{mn} is a solution for $x > x_a$, but not for the region to the left of x_a where the wave function for quasi-particle n drops to zero. Both Ψ_0 and Ψ_{mn} are good solutions in the barrier region $x_a < x < x_b$.

We form a time-dependent solution as a linear combination of Ψ_0 and various final states, Ψ_{mn} , by the usual method:

$$\Psi = a(t)\Psi_0 e^{-iW_0 t} + \sum_{mn} b_{mn}(t)\Psi_{mn} e^{-iW_{mn} t}, \quad (2)$$

and substitute into the Schrödinger equation. This gives for the matrix element for the transition

$$M_{mn} = \int \Psi_0^* (H - W_{mn}) \Psi_{mn} d\tau. \quad (3)$$

Since the integrand vanishes except over a region to the left of x_a , we need to integrate only over this region.

We may express the integral in a more symmetric form by subtracting $\Psi_{mn}(H - W_0)\Psi_0^*$, which vanishes to the left of x_b . Since we are interested only in final states such that $W_{mn} \approx W_0$, the result may be written

$$M_{mn} = \int_a [\Psi_0^* H \Psi_{mn} - \Psi_{mn}^* H \Psi_0] d\tau, \quad (4)$$

where the subscript a indicates that the integration is to be taken over the region to the left of x_a .

The matrix element may be expressed in terms of that of the current density operator, J , in the barrier region as follows. We introduce into the integrand a step function $S(x)$ which is equal to unity between a point x_0 to the left of the important region of integration and a point x_1 in the barrier and which vanishes elsewhere. Then integrations with respect to y_i and z_i vanish and integration with respect to x_i gives

$$\begin{aligned} M_{mn} = -\frac{1}{2\mu} \sum_i \int \dots \int (\Psi_0^* \nabla_i^2 \Psi_{mn} - \Psi_{mn}^* \nabla_i^2 \Psi_0) S(x_i) d\tau_1 \dots d\tau_N \\ = -i [J_{mn}(x_1) - J_{mn}(x_0)], \end{aligned} \quad (5)$$

where $J_{mn}(x)$ is the matrix element of the x component of the current density operator defined by

$$J_{mn}(x) = -\frac{i}{2\mu} \sum_i \int \dots \int [\Psi_0^* \partial \Psi_{mn} / \partial x_i - \Psi_{mn}^* \partial \Psi_0 / \partial x_i] \delta(x - x_i) d\tau_1 \dots d\tau_N. \quad (6)$$

According to our assumptions, $J_{mn}(x_0) = 0$ and further, $J_{mn}(x_1)$ is independent of position as long as x_1 is in the barrier region. Thus

$$M_{mn} = -iJ_{mn}(x_1). \quad (7)$$

It is easily verified that this method leads to the usual results for barrier penetration problems.

The quasi-particle energy in a superconductor is $E = (\epsilon^2 + \Delta^2)^{1/2}$, where $\epsilon(k)$ is the normal state energy measured from the Fermi surface. In calculating the density of final states, it is simplest to take the conventions that $f = 1$ and E is negative for the normally occupied states in the Fermi sea, so that for $k < k_F$, the probability of a hole excitation is $1-f$ and the energy of a hole is $-E$. This is the procedure which has been used in the interpretation of the experimental data.^{1,2} In each region of k space the density of states in energy in the superconductor differs from that in the normal metal by the factor

$$\rho_s = \rho_n [|E| / (E^2 - \Delta^2)^{1/2}], \quad (8)$$

with $\rho_s = 0$ in the gap, $|E| < \Delta$. Agreement with experiment is obtained if it is assumed that M_{mn} (or J_{mn}) is the same for the corresponding transitions in normal and superconducting states.

Usually coherence factors, which have a marked effect on transition probabilities, are introduced in calculations of matrix elements between quasi-particle states of a superconductor.³ To see why such factors are not expected to occur in tunnelling, one must consider the pairing in the vicinity of the barrier. Coherence factors would be introduced if one simply paired complex conjugate wave functions of the type (1a), (1b), so that pairing extends into the barrier region. However, if one looks at the problem more closely, from the viewpoint of the more general Gor'kov equations⁴ which allow for a variation of the energy gap parameter with position, one sees that Δ will drop to zero very rapidly in the barrier. In effect electrons in this region are not paired and the wave function is essentially the same as in the normal state.

For an effective interaction $v(\vec{r}_1, \vec{r}_2)$, the position-dependent energy gap function is defined by

$$\Delta(\vec{r}_1, \vec{r}_2) = \langle N - 2 | \psi(\vec{r}_1) \psi(\vec{r}_2) | N \rangle_{av} v(\vec{r}_1, \vec{r}_2). \quad (9)$$

Simply from the fact that ψ is very small in the barrier region, one expects that $\Delta(\vec{r}_1, \vec{r}_2)$ is small when either \vec{r}_1 or \vec{r}_2 is in the barrier. Note that the tail in the barrier of the wave function of a typical electron at the Fermi surface is very much smaller than that of one of the few electrons moving normal to the interface which has an appreciable probability of penetrating. However, $\Delta(\vec{r}_1, \vec{r}_2)$ is expected to rise very rapidly to normal values in the superconductor, $\Delta(\vec{r}_1, \vec{r}_2) \approx \delta(\vec{r}_1 - \vec{r}_2)\Delta$, with $\Delta = \text{const.}$

Since, according to (6), the matrix element depends only on the wave function in the barrier, it would be expected to be the same as for the corresponding transition in the normal state. Actually, the wave function in the barrier would be changed slightly because of the difference in quasi-particle energies in normal and superconducting states, but this would have a negligible effect on the matrix elements. Thus the only significant difference in the tunnelling current comes from the density of states factor.

The method described here can also be used for calculating the tunnelling current between the valence and conduction bands of a semiconductor, as observed in the Esaki diode. It can be generalized to take indirect transitions into account.

The author is indebted to Dr. W. A. Harrison and Dr. J. R. Schrieffer for discussions of the theory of tunnelling in a superconductor.

*This work was supported in part by the Office of Ordnance Research, U. S. Army.

¹Ivar Giaever, Phys. Rev. Letters **5**, 147, 464 (1960).

²J. Nicol, S. Shapiro, and P. H. Smith, Phys. Rev. Letters **5**, 461 (1960).

³J. Bardeen, L. N. Cooper, and J. R. Schrieffer, Phys. Rev. **108**, 1175 (1957).

⁴L. P. Gor'kov, J. Exptl. Theoret. Phys. U.S.S.R. **34**, 735 (1958) [translation: Soviet Phys.-JETP **34**(7), 505 (1958)].