

2007

Influence of tip-induced band bending on tunnelling spectra of semiconductor surfaces

Randall M. Feenstra

Carnegie Mellon University, feenstra@andrew.cmu.edu

Y. Dong

Carnegie Mellon University

M. P. Semstiv

Humboldt-Universität zu Berlin - Humboldt University Berlin

W. T. Masselink

Humboldt-Universität zu Berlin - Humboldt University Berlin

Follow this and additional works at: <http://repository.cmu.edu/physics>

Published In

Nanotechnology , 18, 044015.

This Article is brought to you for free and open access by the Mellon College of Science at Research Showcase @ CMU. It has been accepted for inclusion in Department of Physics by an authorized administrator of Research Showcase @ CMU. For more information, please contact research-showcase@andrew.cmu.edu.

Influence of Tip-induced Band Bending on Tunneling Spectra of Semiconductor Surfaces

R. M. Feenstra and Y. Dong

Department of Physics, Carnegie Mellon University, Pittsburgh, Pennsylvania 15213

M. P. Semtsiv and W. T. Masselink

Department of Physics, Humboldt-Universität zu Berlin, D-10115 Berlin, Germany

Abstract

A theory based on the Bardeen formalism is developed for computing the tunnel current between a metal tip and a semiconductor surface. Tip-induced band bending in the semiconductor is included, with the electrostatic potential computed in a fully three-dimensional model whereas the tunnel current is computed in the limit of large tip radii. Localized states forming at the semiconductor surface as well as wavefunction tailing through the semiconductor depletion region are fully accounted for. Numerical results are provided and compared with data obtained from *p*-type GaAs surfaces, and generalization of the method to semiconductor heterojunctions is discussed.

I. INTRODUCTION

Since the invention of the scanning tunneling microscope (STM), efforts have been made to utilize it for measuring fundamental properties of semiconductor surfaces and heterostructures. Such studies are often performed in a cross-sectional mode, with the probe tip positioned near the heterojunction and tunneling spectra collected. In addition to quantities apparent in the images such as alloy ordering,¹ interface roughness,² or strain in layers,³ the main spectroscopic feature of interest is the relative positions of the valence band (VB) and conduction band (CB) edges in the neighboring layers, i.e. the band offsets. Spatially resolved tunneling spectra (scanning tunneling spectroscopy, STS) are acquired at spatial locations on either side of the interface and are compared in order to determine the band offsets. An early and relatively successful effort of this type was carried out by Salemink and coworkers on AlGaAs/GaAs heterojunctions.⁴ Nevertheless, for that study and all similar ones uncertainty in the results occurs because of the lack of any quantitative description of the tunneling spectra.

A significant impediment to the quantitative description of tunneling spectra from semiconductors is the occurrence of tip-induced band bending in the semiconductor, illustrated in Fig. 1, in which some of the applied potential between the probe tip and the sample is dropped in the sample itself.⁵ In a prior work we have presented a three-dimensional finite-element method for evaluating this band bending.⁶ The main parameters contained in that theory, in addition to the semiconductor parameters such as doping concentration and effective masses, are the tip-sample separation, s , the tip radius, R , and the electrostatic potential energy of the probe tip relative to a point deep within the semiconductor, ϕ_T . This latter quantity is related to the applied voltage on the sample relative to the tip, V , according to

$$\phi_T = eV + \Delta\phi \quad (1a)$$

where the *contact potential* $\Delta\phi$ is given by the work function difference between tip and sample

$$\Delta\phi = \phi_m - \chi - (E_C - E_F) \quad (1b)$$

where ϕ_m is the work function of the metallic tip, χ is the electron affinity of the semiconducting sample, E_C is the CB minimum of the sample, and E_F is the Fermi-level of the sample. Figure 2 shows an energy band diagram for the semiconductor-vacuum-metal tunnel junction illustrating these various quantities. Even in the absence of any applied voltage, a negative contact potential will produce downwards band bending in the semiconductor and a positive contact potential will produce upwards band bending.

Our finite-element method yields the electrostatic potential at all points in the semiconductor and the vacuum. An example of typical results from such a computation is shown in Fig. 3. Of particular interest is the potential energy at the point on the semiconductor surface directly opposite the tip apex, which we denote by ϕ_0 as indicated in Fig. 1(a). Let us consider in what way this band bending will affect a measured tunneling spectrum. For a simple situation in which the spectrum is dominated by surface states (*e.g.* arising from a surface reconstruction), the main influence of the band bending will be to simply shift the spectrum of these states according to the amount of the band bending. Thus, whereas with $\phi_0 \equiv 0$ we would have the usual relationship between

energy of a state E and the sample voltage V , $E - E_F = eV$, with nonzero band bending this equation must be modified so that

$$E - E_F = eV - \phi_0 . \quad (2)$$

An analysis based on this equation could be termed an *energy alignment method*; it assumes that the states being observed shift rigidly in accordance with the band bending, *i.e.* as appropriate for surface states. But, for bulk states, additional quantum effects can occur as illustrated in Figs. 1(b) and (c). Both wavefunction tailing through a barrier region in the semiconductor, Fig. 1(b), as well as localized states formation, Fig. 1(c), can lead to substantial errors in the use of Eq. (2) to relate experiment and theory. In this case one must employ a complete computation of the conductance vs. voltage characteristics, comparing those results with experiment in order to deduce the values of the parameters of interest (e.g. band gap, or heterojunction band offset) from the theory. This method can be viewed as a *lineshape analysis* of the spectra, using curve fitting to compare theory and experiment. It is the purpose of this paper to develop a theory for the computation of tunneling spectra that enables this type of analysis.

One additional point that should be addressed concerns the role of *extrinsic states* at semiconductor surfaces. By extrinsic, we mean any sort of states arising from disorder, defects, or unintentional contamination. Such states, even at low densities of 0.01 monolayers ($\text{ML} \approx 3 \times 10^{14} \text{ cm}^{-2}$) or less, can hold enough charge to significantly affect the magnitude of tip-induced band bending. The extrinsic states produce their own band bending of course, and additional effects due to the tip decrease in significance as the density of extrinsic states increases. Directly spectroscopic observation of extrinsic states is difficult but nevertheless has been accomplished in some cases by STS.⁷ However, it is most common in STS experiments to acquire spectra sufficiently far from the extrinsic defects so that the states are not seen *directly* in the spectra, but they *do* affect the spectroscopic results through a change in the band bending. Thus, although not further discussed in this work, extrinsic states can be included in the computations according to the effect they have on the electrostatic potential.

II. THEORY

A. Background

The main purpose of this work is to develop a theory enabling the computation of tunneling spectra for surfaces such as GaAs(110) for which the role of intrinsic surface states is relatively small [for the (110) surfaces of nearly all III-V semiconductors the intrinsic states associated with the anion and cation dangling bonds form resonances lying within the valence or conduction bands⁸]. In this case the tunnel current can be computed primarily as a summation over the bulk states, as labeled by their wavevectors. In addition to use of the effective mass approximation we make one other significant approximation in all our computations, assuming that the current can be obtained using only the potential along the central axis of our problem, which we denote by $\phi(z)$, and then performing a *planar* tunneling computation with this potential. This approximation essentially amounts to a semiclassical treatment of the lateral (radial) part of the wavefunctions.⁹ In general the effect of curved potential contours as in Fig. 3 will lead to a slight focusing of the wavefunction as one approaches the surface from far inside the semiconductor (*i.e.* for a completely spherical geometry the wavefunction increases as

$1/r$ as the radial coordinate r decreases) with this effect being weakly energy dependent. For localized states at the surface, as in Fig. 1(c), this approximation neglects any quantization of the states arising from lateral variation of the potential; those energy splittings have typical magnitude of 10 meV.^{7,10} For equipotential contours with large radius of curvature as in Fig. 3, both types of effects are expected to be relatively small.

For a planar computation we have, in two prior papers,^{7,11} developed a formalism that is similar to the "exact" computation of tunnel current described by Duke¹² but handles an arbitrary potential variation $\phi(z)$ in the vacuum and semiconductor. The current is written as an integral over incoming states, with some transmission factor for each state and where these transmission factors are computed exactly by integration of the perpendicular part of the Schrödinger equation.^{13,14} Explicitly, for a state with given value of the perpendicular component of the energy E_{\perp} , then on a set of z -values labeled by j with separation Δz , the perpendicular part of the wavefunction ψ is obtained simply from

$$\psi_{j+1} = 2\psi_j - \psi_{j-1} \pm \frac{2m^*(\Delta z)^2}{\hbar^2} \psi_j [U(z_j) - E_{\perp}] \quad (3)$$

where m^* is the effective mass and $U(z)$ is the potential energy. The upper sign in this equation is for a state in the semiconductor CB, the vacuum, or the tip, and the lower sign is for a state in the semiconductor VB. With our zero for the electrostatic potential energy $\phi(z)$ being a point deep inside the semiconductor, $U(z)$ is given by

$$U(z) = \begin{cases} E_C + \phi(z) & \text{semicon. CB} \\ E_V + \phi(z) & \text{semicon. VB} \\ E_C + \chi + \phi(z) & \text{vacuum} \\ E_F + eV - W & \text{tip} \end{cases} \quad (4)$$

where E_V is the VB maximum and W is the energy difference between the Fermi-level in the tip and the bottom of its metallic band (typically about 8 eV). These various potential energies are indicated in Fig. 2. Both m^* and E_{\perp} may have different values in the semiconductor as compared to in the vacuum or the tip as discussed in more detail in the following section. In this type of "exact" computation we are dealing with propagating states in the electrodes and the vacuum, starting in one electrode with an outgoing state and integrating back through the vacuum and into the other electrode. Matching in that electrode to an incoming wave then enables determination of the transmission coefficient for that state. The wavefunctions in the vacuum are thus complex, current carrying functions containing both exponentially decaying and growing parts.

Although the above method does indeed provide an exact solution for the propagating states of the two-electrode system, it has a significant drawback in that it is not clear how to treat states that are localized at the semiconductor surface. For this reason we switch approaches in the present work to the Bardeen formalism for computing tunnel currents.¹⁵ In this case the electrodes are treated separately, integrating the wavefunctions for each (starting from within the vacuum back to each electrode) and dealing only with wavefunctions that are real (and hence *not* current carrying). In this case localized states

can be treated within the same formalism as extended states, as described in the following section. Our results employing the two methods are identical (both formally, and in actual computation) for extended states, but the Bardeen approach also permits a straightforward treatment of localized states.

B. Formalism

We consider tunneling between states in electrodes on the left-hand (L) and right-hand (R) sides of a planar junction. The left-hand electrode is the semiconductor, with Fermi-level at E_F , and the right-hand electrode is the metallic tip, with Fermi-level at $E_F + eV$. Labeling the states by quantum numbers of states in the perpendicular direction, μ and ν , and by their parallel wavevectors, $\mathbf{k}_{||,L}$ and $\mathbf{k}_{||,R}$, respectively, the transition rate is given by^{12,15}

$$P = \frac{2\pi}{\hbar} \sum_{\mathbf{k}_{||,L}, \mathbf{k}_{||,R}, \mu, \nu} \left| \langle \mu \mathbf{k}_{||,L} | \Lambda | \nu \mathbf{k}_{||,R} \rangle \right|^2 [f(E_L) - f(E_R)] \delta(E_L - E_R - eV) \quad (5)$$

where E_L and E_R are the energies of the states and $f(E)$ is a Fermi-Dirac occupation factor. The matrix element for the transition is given by

$$\langle \mu \mathbf{k}_{||,L} | \Lambda | \nu \mathbf{k}_{||,R} \rangle = \frac{\hbar^2}{2m} \delta_{\mathbf{k}_{||,R}, \mathbf{k}_{||,L}} \left(\psi_\mu \frac{d\psi_\nu}{dz} - \psi_\nu \frac{d\psi_\mu}{dz} \right)_{z_1} \quad (6)$$

where ψ_μ and ψ_ν are the perpendicular parts of the wavefunctions of each electrode, and the term in brackets on the far right-hand side of this equation is evaluated at a point z_1 somewhere in the vacuum region. The δ -function in Eq. (6) eliminates one of the sums over $\mathbf{k}_{||}$ in Eq. (5).

We assume a metallic probe-tip, so that its states are all extended ones, and we treat both extended and localized states in the semiconductor sample. Considering the former case first, our derivation for the current follows that given by Duke. Writing the result for the CB of the semiconducting electrode with effective mass m^* we convert the summations over states to integrals as (suppressing the L subscript)

$$\sum_{\mathbf{k}_{||}, \mu} = \frac{A}{(2\pi)^2} 2\pi \int_0^\infty k_{||} dk_{||} \frac{L}{2\pi} \int_{-\infty}^\infty dk_\perp \quad (7a)$$

$$= \frac{AL}{(2\pi)^2} \int_0^\infty k_{||} dk_{||} 2 \int_{-\infty}^\infty \frac{dE_\perp}{\partial E_\perp / \partial k_\perp} \theta(E_\perp - E_C) \quad (7b)$$

$$= \frac{AL}{(2\pi)^2} \int_0^\infty k_{||} dk_{||} \frac{2m^*}{\hbar^2} \int_{-\infty}^\infty \frac{dE_\perp}{k_\perp} \theta(E_\perp - E_C) \quad (7c)$$

$$= \frac{AL}{(2\pi)^2} \frac{2m^*}{\hbar^2} \int_{-\infty}^\infty dE \int_0^\infty k_{||} dk_{||} \frac{1}{k_\perp} \theta(E_\perp - E_C) \quad (7d)$$

where E_C is the CB minimum, L is the thickness of a slab that composes the electrode, and A is the area of the electrode. For a CB the perpendicular part of the energy is given by $E_{\perp} = E - \hbar^2 k_{\parallel}^2 / 2m^*$ and $k_{\perp} = \left[2m^*(E_{\perp} - E_C) / \hbar^2 \right]^{1/2}$ (for simplicity of notation we are assuming equal effective masses for the parallel and perpendicular directions), and the step function $\theta(E_{\perp} - E_C)$ restricts this energy to lie within the band. For a valence band (VB) we have $E_{\perp} = E + \hbar^2 k_{\parallel}^2 / 2m^*$, $k_{\perp} = \left[2m^*(E_V - E_{\perp}) / \hbar^2 \right]^{1/2}$, and the step function is replaced by $\theta(E_V - E_{\perp})$ with E_V being the VB maximum.

For the case of the probe tip the sum over ν , for a given value of k_{\parallel} , is easily evaluated to be (suppressing the R subscript)

$$\sum_{\nu} = \frac{L}{2\pi} \int_{-\infty}^{\infty} dk_{\perp} \quad (8a)$$

$$= \frac{L}{2\pi} 2 \int_{-\infty}^{\infty} \frac{dE_{\perp}}{\partial E_{\perp} / \partial k_{\perp}} \theta(E_{\perp} - E_0) \quad (8b)$$

$$= \frac{L}{2\pi} \frac{2m}{\hbar^2} \int_{-\infty}^{\infty} \frac{dE}{k_{\perp}} \theta(E_{\perp} - E_0) \quad (8c)$$

where $E_{\perp} = E - \hbar^2 k_{\parallel}^2 / 2m$, $k_{\perp} = \left[2m(E_{\perp} - E_0) / \hbar^2 \right]^{1/2}$, and $E_0 = E_F + eV - W$ is the minimum energy of the metallic band in the tip.

Care must be taken in counting the numbers of k_{\perp} states available for tunneling: linear combinations of the plane wave states in each electrode traveling towards and away from the vacuum are formed that satisfy the boundary conditions for an exponentially decaying state in the vacuum. The resulting *standing waves* can have either even or odd parity relative to the center of each slab, so that the total number of states available equals the number of k_{\perp} values including *both* positive and negative k_{\perp} values [hence the factors of 2 preceding the energy integrals in Eqs. (7b) and (8b)]. Additionally, the normalization of the wavefunctions is done on these standing waves just mentioned, *not* on the plane waves, yielding another factor of 2 for each electrode that differs from Duke's derivation. In this way our result is a factor of 16 different from that of Duke, but importantly, our result agrees precisely with that expected from an "exact" treatment of the problem as derived by Duke (in other words, we explicitly resolve the factor of 16 discrepancy, noted by Duke, between his treatments using the Bardeen and the exact formalisms).

Writing the current density as $J = 2eP/A$ (with the factor of 2 being the spin degeneracy), then by combining Eqs. (5-8) we arrive at the final result for extended states of

$$J = \frac{2e}{h} \int_{-\infty}^{\infty} dE [f(E) - f(E - eV)] \int_0^{\infty} \frac{k_{\parallel} dk_{\parallel}}{2\pi} \theta(E_{\perp,L} - E_C) \theta(E_{\perp,R} - E_0)$$

$$\times \frac{L^2 m^*}{k_{\perp,L} k_{\perp,R} m} \left(\psi_{k_{\perp,L}} \frac{d\psi_{k_{\perp,R}}}{dz} - \psi_{k_{\perp,R}} \frac{d\psi_{k_{\perp,L}}}{dz} \right)_{z_1}^2 \quad (9)$$

where the expressions for $k_{\perp,L}$ and $k_{\perp,R}$ are given following Eqs. (7) and (8), respectively. The values for the perpendicular part of the energy are also listed there (with the value in the vacuum being the same as that in the probe tip); these E_{\perp} values are needed for evaluating the wavefunctions in the bracketed expression at the end of the integrand. This expression scales as $1/L^2$, since the wavefunctions are unit normalized on a slab of thickness L , and hence the L^2 term appearing explicitly in Eq. (9) is cancelled.

We compute the wavefunctions numerically by direct integration of the perpendicular part of the Schrödinger equation as in Eq. (3). For each electrode we start the integration in the vacuum with the wavefunction composed of a decaying exponential, and then integrate back to the respective electrode. For the tip, we then simply match the wavefunction to a suitable standing wave. For the semiconductor, we match the wavefunction and its derivative divided by the effective mass across the interface, integrate through any region of the semiconductor with a varying potential, and finally match to a suitable standing wave deep inside the semiconductor. For both the tip and semiconductor, this final matching procedure determines the normalization of the wavefunction, from which its magnitude and derivative values in the vacuum are determined.

All of our computations are made in the effective mass approximation, neglecting the periodic part of the Bloch functions, so that our wavefunctions are just the envelope function part of the actual wavefunctions. In this approximation, BenDaniel and Duke have argued that the boundary condition on the derivative of the envelope function divided by effective mass is suitable, but the matching of the envelope function itself across the semiconductor-vacuum interface *is* a significant approximation.¹⁶ Thus, considering the ratio of our computed currents through the VB compared to the CB, this ratio will be in error by some factor of order unity relating to the detailed period part of the Bloch functions of these states. But importantly, we expect this factor not to deviate too much from unity since for the (110) surface of III-V semiconductors the resonant, dangling bond state associated with the VB (anions) and CB (cations) both have comparable z -height relative to the surface. Several of the parameters in our model affect the ratio of the currents between VB and CB, and thus we believe that small adjustments in these values serve to accommodate our neglect of this factor.

Let us now consider localized states in the semiconducting electrode. The sum on μ now runs over some number of discrete localized states; writing the result for a CB for which the energies of the localized states satisfy $E_{\mu} < E_C$, and including the sum over $\mathbf{k}_{||}$, we have (again, suppressing the L subscript)

$$\sum_{\mathbf{k}_{||}, \mu} = \sum_{\mathbf{k}_{||}, \mu} \int_{-\infty}^{\infty} dE_{\perp} \delta(E_{\perp} - E_{\mu}) \quad (10a)$$

$$= \frac{A}{2\pi} \int_{-\infty}^{\infty} dE \int_0^{\infty} k_{||} dk_{||} \sum_{\mu} \delta(E_{\perp} - E_{\mu}) \quad (10b)$$

with $E_{\perp} = E - \hbar^2 k_{||}^2 / 2m^*$. For the case of a VB the localized states satisfy $E_{\mu} > E_V$ and $E_{\perp} = E + \hbar^2 k_{||}^2 / 2m^*$.

Again using a slab of thickness L to normalize the states in the semiconductor, we note that in principle the sum over μ in the expressions of Eq. (10) includes a two-fold degeneracy for each state (i.e. in addition to the spin degeneracy) since we can form both even and odd parity combinations of the localized states on the two sides of the slab. However, this factor of two is cancelled by the normalization of the wavefunctions if we properly include the portions of the wavefunction on both sides of the slab. It is more intuitive to consider only the portion of the wavefunction on one side of the slab; in this case we can normalize that portion to unity and take the sum over μ in Eq. (10) to include only a *single* state at each value of E_{μ} .

Using Eq. (10b) for a CB in the semiconducting electrode and Eq. (8c) for the metal electrode we arrive at the expression for the current density through localized CB states

$$J = \frac{2e}{h} \int_{-\infty}^{\infty} dE [f(E) - f(E - eV)] \int_0^{\infty} k_{||} dk_{||} \sum_{\mu} \delta(E_{\perp,L} - E_{\mu}) \theta(E_{\perp,R} - E_0) \\ \times \frac{\hbar^2 L}{2m k_{\perp,R}} \left(\psi_{\mu} \frac{d\psi_{k_{\perp,R}}}{dz} - \psi_{k_{\perp,R}} \frac{d\psi_{\mu}}{dz} \right)_{z_1}^2. \quad (11)$$

The means of obtaining the tip wavefunction here is the same as described following Eq. (9), but to obtain the semiconductor wavefunction we now match to a decaying exponential deep inside the semiconductor. We note that Eq. (11) is in a form well suited to computation; the integral over energy is evaluated in the usual way as a discrete sum multiplied by an energy increment ΔE , and the δ -function then makes a contribution of $1/\Delta E$ to those terms for which a localized state first appears at that energy (for a given parallel wavevector) and zero otherwise. It is straightforward to show that the current through localized states reduces to the result for extended states in the limit that the localization distance at the surface becomes infinite.¹⁷

III. NUMERICAL RESULTS

To illustrate our analysis method we consider a spectrum acquired from a *p*-type GaAs surface. This spectrum was actually acquired not too far from a InGaP/GaAs heterojunction, the details of which are discussed elsewhere.¹⁸ The $\text{In}_x\text{Ga}_{1-x}\text{P}$ alloy layer has $x \approx 0.5$ and is lattice matched to GaAs; its bandgap (1.90 eV) is considerably larger than that of GaAs (1.42 eV) so it is useful as an alternative to AlGaAs as a barrier material for heterostructures.¹⁹ The GaAs layer is *p*-type doped at a concentration of $1 \times 10^{18} \text{ cm}^{-3}$ and the 50 nm thick InGaP is undoped. Figure 4 shows the tunneling spectrum acquired at a point located on the GaAs layer located 4 nm from the

InGaP/GaAs heterojunction. At this distance the states of the InGaP do *not* make any significant contribution to the current, although the presence of this layer still affects the results in terms of band bending due to its low doping. The spectrum in Fig. 4 plotted as conductance at constant- s on a logarithmic scale, and it is compared with various computed spectra. The computations are performed using three valence bands (light hole, heavy hole, and split off bands) and a single conduction band, using effective masses (in units of the free electron mass) of 0.0635 for the CB (at 300 K) and 0.081, 0.643, and 0.172 for the light, heavy, and split-off VBs, and a spin-orbit splitting of 0.341 eV.²⁰ Results are shown for a temperature of 300 K and using a modulation voltage of $V_{rms} = 50$ mV (for a sample voltage V the conductance is computed by taking the difference in currents at $V \pm V_{rms}$, divided by $2V_{rms}$).

The fact that the VB onset of the measured spectrum is situated far below 0 V necessitates negative contact potentials with rather large magnitude. The tip radius is also quite important in determining the tip-induced band bending, with larger values producing more band bending. Similarly, smaller values of tip-sample separation produce larger amounts of band bending although this dependence is relatively weak. As shown in Fig. 4(a), a contact potential of -1.4 eV together with a tip radius of 30 nm, at a separation of 0.9 nm, provide a reasonable fit to the experiment. Decreasing the contact potential to -1.3 eV produces a significantly worse fit [Fig. 4(b)], although by decreasing the separation to 0.7 nm [Fig. 4(c)] or increasing the radius to 40 nm [Fig. 4(d)] the good fit is recovered. It is thus seen that considerable correlation occurs among the parameters of the theory, and hence, determination of a band gap alone would have significant uncertainty since various apparent gaps can be achieved in the theory depending on the amount of band bending. However, for a situation of a heterojunction in which we only want to determine a band offset, the uncertainty is expected to be considerably less since we are concerned only with a difference between band edges from neighboring layers. (For determination of a band gap alone, it is useful to have additional features in the spectrum that are fit between experiment and theory, as done *e.g.* in Ref. [11]).

In addition to the three parameters shown in Fig. 4, the fourth parameter used in matching experiment to theory is the overall amplitude of the spectrum. This amplitude is directly related to the *area* of the tunnel junction, a quantity that we expect to be greater than the area of a single atom, ≈ 0.01 nm², and less than or approximately equal to an area of several atomic-spacings on a side, ≈ 1 nm². In the theory of Section II the results are obtained as current density, J , from which the current is given by $I = JA$ for area A . We must however also account for image potential effects, which increase the overall magnitude of the tunnel current by about 3 orders-of-magnitude (with no significant dependence on energy nor tip-sample spacing of this factor).^{21,22} We thus divide the area A by 1000 to obtain an effective area corrected for the image potential. For spectra (a) and (d) of Fig. 4, at tip-sample separation of 0.9 nm, we find a corrected area of 0.7 nm², while for spectrum (c) with a separation of 0.7 nm we find a corrected area of 0.014 nm². Both of these results are in the range consistent with expectations, although separations significantly larger or smaller than these would produce inconsistency between experiment and theory.

It should be noted that for the large, negative contact potentials encountered in this work, wavefunction tailing through a barrier region in the semiconductor, as in Fig. 1(b), dominates the GaAs spectrum for negative voltages. This effect is illustrated in Fig. 5, where we show theoretical results for various parameter sets. The parameters for Fig. 5(a) are identical to those of Fig. 4(a), with the parameters of Figs. 5(b) and (c) also chosen to produce approximate matches [albeit slightly worse than for Fig. 5(a)] with experiment. For the parameters of Fig. 5(a), the inset on the lower, left-hand corner of the figure shows the potential profile in the semiconductor for a sample voltage of -1.1 V. It is seen that the band bending in the semiconductor is rather large, extending nearly all the way down to the tip Fermi-level (at $E_F - 1.1$ eV), so that essentially all the tunnel current (indicated by the arrow in the inset) is formed by carriers tunneling through the barrier in the semiconductor. A similar situation also occurs for Figs. 5(b) and (c). Proper description of the wavefunction extending through the semiconductor barrier region, which in our theory is characterized by the hole effective masses, is therefore essential in obtaining a quantitative description of the spectra.

The other quantum effect described by our theory is the formation of effective-mass states localized at the semiconductor surface, as pictured in Fig. 1(c). Such states make a negligible contribution to the theoretical curve of Fig. 5(a), but they make noticeable contributions to curves (b) and (c). Figure 5(b) is computed with a smaller tip radius and larger-magnitude contact potential than Fig. 5(a), resulting in band bending at voltages near 1 V that is lower (more negative) than for Fig. 5(a). The inset in the lower, right-hand corner of the figure shows the potential profile in the semiconductor for a sample voltage of $+1.1$ V. Two localized states are formed in the CB, at energies of 50.6 and 0.2 meV below the energy of the CB edge deep inside the semiconductor. These states make a significant contribution to the current, as seen by comparing the symbols in Fig. 5(b) with the dotted line.

Figure 5(c) displays a situation with a larger tip radius and smaller-magnitude contact potential compared to Fig. 5(a), resulting in band bending at voltages near 1 V that is higher (more positive) than for Fig. 5(a). Localized states associated with the heavy-hole VB are thus formed, at energies of 7.1 and 0.8 meV above the VB edge as shown in the inset in the upper part of the figure for a sample voltage of $+1.1$ V. These states make a significant contribution to the current as seen by comparing the open symbols in Fig. 5(c) with the dotted line, and this contribution is an *accumulation current* associated with the VB.

For *all* the theoretical curves in Fig. 5 a nonzero conductance is found between 0.1 and 1.0 V; aside from the localized states just mentioned this current arises from electrons tunneling into *extended* VB states that are empty due to the *p*-type doping. This is the so-called dopant-induced component, or D-component, commonly observed on *p*-type GaAs having a relatively low density of extrinsic states and probed with a tip having a contact potential which does *not* have a large negative value. It should be noted that the ordinate in Fig. 5 spans many orders-of-magnitude, with typical experimental data extending only over the top 3 – 4 of these, so that the magnitude of D-component seen in the theory would generally not be visible in a typical experiment [although for Fig. 5(c) the D-component does just reach an observable range]. That expectation is consistent with the data of Fig. 4 in which the D-component is indeed not observed.

IV. SUMMARY

In summary, we have developed a theory based on the Bardeen formalism that allows computation of tunneling current from semiconductor surfaces. The theory employs a fully three-dimensional finite-element solution for the potential, whereas the tunnel current is computed only along the central axis of the problem (i.e. as appropriate for large tip radii). For large tip radii such as that shown in Fig. 3 this approximation seems reasonable, as discussed in Section II(A), but for smaller tip radii a fully three-dimensional treatment of the Schrödinger equation is needed.

Our results can be compared to prior computations of tunnel current that include tip-induced band bending: The main differences between the present results and those of Ref. [5] are, firstly, the potential here is described in a 3D model and, secondly, that a full numerical integration is done in computing the current rather than using the WKB approximation. Both effects influence the conductance vs. voltage characteristics, with the latter one in particular leading to a saturation of the conductance vs. voltage characteristic at high voltages that does not match the experimental data (although it should also be noted that higher lying and/or nonparabolic bands, i.e. beyond the effective mass approximation, will also affect the theory in this voltage range).²³ Comparing our results with those of Dombrowski et al., we note that their results, while providing an excellent treatment of accumulation layer states, are not predictive in terms of the potential (a Gaussian form is assumed, with parameters fit to experiment) and in this sense are rather different than the current computations.

Finally, we comment on the extension of this method to semiconductor heterostructures, a topic that provides a major motivation for the present work. In our prior data obtained from InGaP/GaAs heterojunctions we clearly observe a continuous evolution of the spectra from GaAs-like to InGaP-like as the tip moves across the heterointerface. Using the methods described here we are able to fit spectra located well away (at least several nm) from the interface. Such fits allow a measure of the band offsets between GaAs and InGaP, although since the spectra in the respective materials are acquired at points separated by ≈ 5 nm then some uncertainty occurs in the resulting band offset since a *lateral* field may exist across the interface and shift the respective spectra relative to each other. To reduce this uncertainty, spectra located closer to the interface must be analyzed, but in that case we observe tails of the states from the GaAs extending into the InGaP. Those effects cannot be described by the present theory since we deal only with a homogeneous material. Extension of the technique is required to allow a description of such *evanescent* states near the heterointerface, and such work is currently in progress.²⁴

Acknowledgements

This work was supported by the U.S. National Science Foundation, grant DMR-0503748, and by the Deutsche Forschungsgemeinschaft within the framework of “Forschergruppe 394”.

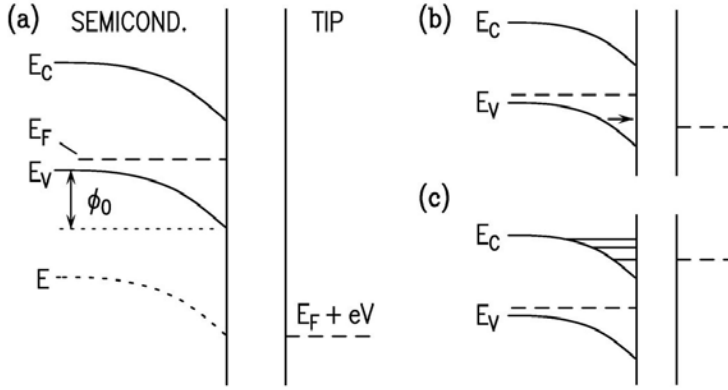


FIG 1. (a) Sketch of energy bands used for computations of the electrostatic potential. The sample Fermi-level is denoted by E_F with the tip Fermi-level at $E_F + eV$ where V is the sample voltage. The band bending at the surface is denoted by ϕ_0 , with V and ϕ_0 both being negative in this diagram. Quantum effects within the semiconductor are illustrated in (b) and (c) for wavefunction tailing through a depletion region and for localized state formation, respectively.

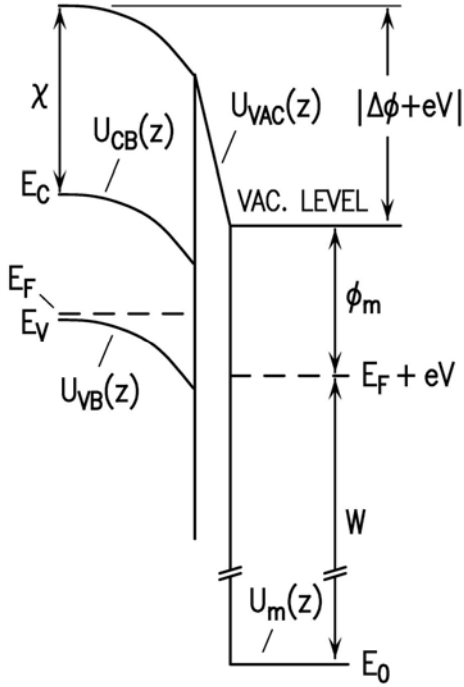


FIG 2. Energy diagram for semiconductor-vacuum-metal junction, with p -type doping in the semiconductor. The sample voltage V is negative in this example, as is the contact potential $\Delta\phi = \phi_m - \chi - (E_C - E_F)$. Other quantities are defined in the text.

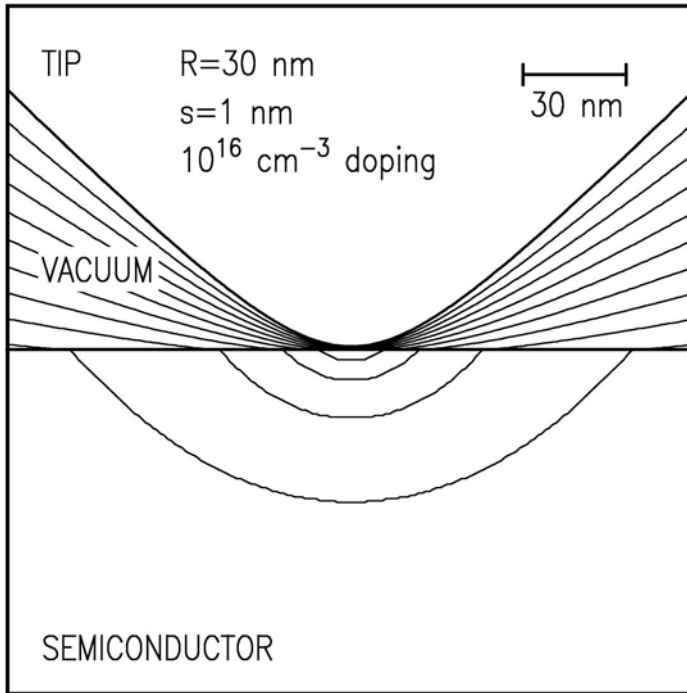


FIG 3. Example of a potential distribution obtained from the theoretical computations, for a tip-sample separation of 1 nm, tip radius of 30 nm, and tip position 1 nm inside the InGaP layer. A situation of depletion in the semiconductor is shown, with potential energy contours separated by one tenth of the potential difference between tip and semiconductor (e.g. with the electrostatic potential energy of the tip being 1 eV relative to a point far inside the semiconductor, the contours are at energies of 0.1, 0.2, ..., 0.9 eV).

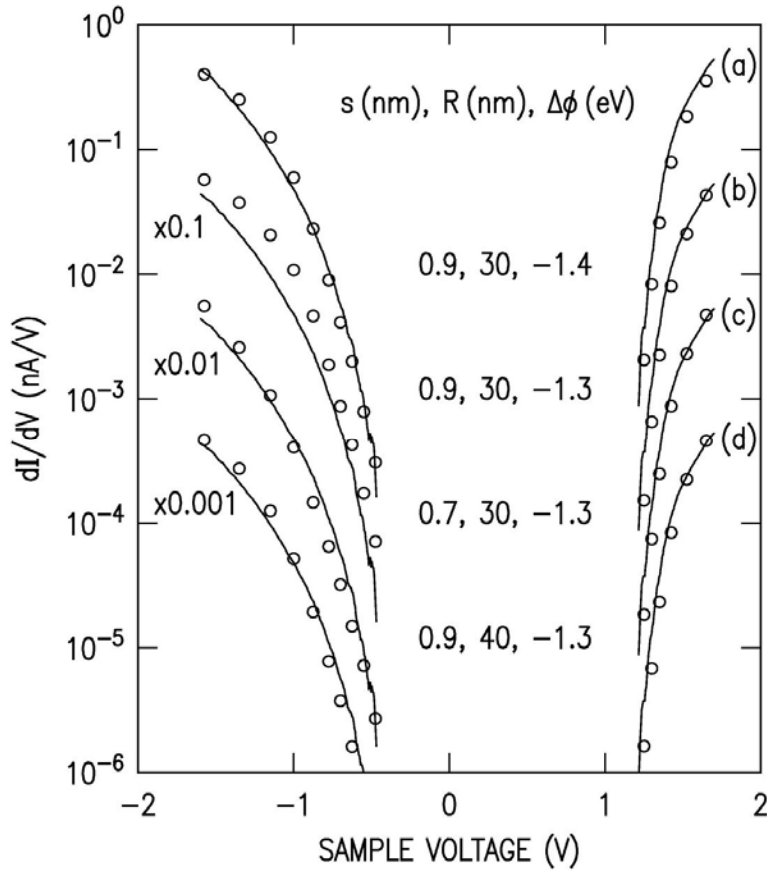


FIG 4. Tunneling spectrum (lines) obtained from a *p*-type GaAs(110) surface, shown repeated four times with each curve compared to a theoretical computation (circles). Consecutive pairs of experimental and theoretical curves are displaced by an order-of-magnitude, for ease of viewing. The theory is computed using various parameter values as indicated; tip-sample separation s , tip radius-of-curvature R , and contact potential $\Delta\phi$.

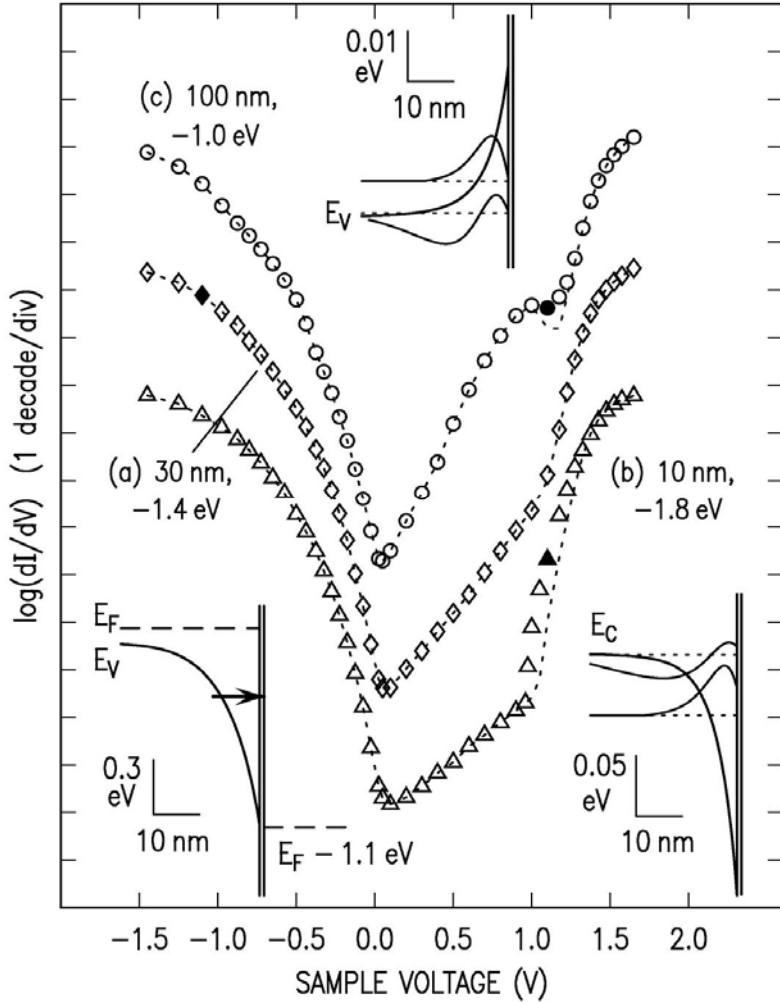


FIG 5. Computed tunnel spectra for the parameter values indicated, using a tip-sample separation of 0.9 nm for all spectra. The open symbols give the result including both extended and localized states, and the dotted lines include extended states only. Insets show the band bending profiles for various sample voltages indicated by the closed symbols: lower left, -1.1 V for curve (a) showing the VB profile; lower right, 1.1 V for curve (b) showing the CB profile and the wavefunctions of localized states; upper center, 1.1 V for curve (c) showing the VB profile and the wavefunctions of localized states.

References:

- ¹ N. Liu, C. K. Shih, J. Geisz, A. Mascarenhas, and J. Olson, Appl. Phys. Lett. **73**, 1979 (1998).
- ² R. M. Feenstra, D. A. Collins, D. Z.-Y. Ting, M. W. Wang, and T. C. McGill, Phys. Rev. Lett. **72**, 2749 (1994).
- ³ H. Chen, R.M. Feenstra, P.G. Piva, R.D. Goldberg, I.V. Mitchell, G.C.Aers, P.J. Poole and S. Charbonneau, Appl. Phys. Lett. **75**, 79 (1999).
- ⁴ H. W. M. Salemink, O. Albrektsen, and P. Koenraad, Phys. Rev. B **45**, 6946 (1992).
- ⁵ R. M. Feenstra and J. A. Stroschio, J. Vac. Sci. Technol. B **5**, 923 (1987).
- ⁶ R. M. Feenstra, J. Vac. Sci. Technol. B **21**, 2080 (2003).
- ⁷ R. M. Feenstra, S. Gaan, G. Meyer, K. H. Rieder, Phys. Rev. B **71**, 125316 (2005).
- ⁸ H. Carstensen, R. Claessen, R. Manzke, and M. Skibowski, Phys. Rev. B **41**, 9880 (1990), and references therein.
- ⁹ For a given energy E_{\perp} of a state (extended or localized) in the z -direction, the wavefunctions in the radial direction in a semiclassical approximation have the form $\exp(i\mathbf{k}_{\parallel} \cdot \mathbf{r})$ in classically allowed regions and zero in forbidden regions, with
$$k_{\parallel} = \left[2m^* |E - E_{\perp} - \phi(x, y, z)| / \hbar^2 \right]^{1/2}$$
 at each point in the semiconductor.
- ¹⁰ R. Dombrowski, Chr. Steinebach, Chr. Wittneven, M. Morgenstern, and R. Wiesendanger, Phys. Rev. B **59**, 8043 (1999).
- ¹¹ R. M. Feenstra, J. Y. Lee, M. H. Kang, G. Meyer, and K. H. Rieder, Phys. Rev. B **73**, 035310 (2006).
- ¹² C. B. Duke, *Tunneling in Solids* (Academic Press, New York, 1969), p. 217-219.
- ¹³ R. S. Becker, J. A. Golovchenko, and B. S. Swartzentruber, Phys. Rev. Lett. **55**, 987 (1985).
- ¹⁴ R. M. Feenstra, J. A. Stroschio, and A. P. Fein, Surf. Sci. **181**, 295 (1987).
- ¹⁵ J. Bardeen, Phys. Rev. Lett. **6**, 57 (1961).
- ¹⁶ D. J. BenDaniel and C. B. Duke, Phys. Rev. **152**, 683 (1966).
- ¹⁷ For example, consider a region near the surface of thickness ℓ in which localized states form. Assume that the potential in this region is constant, and the energy barriers to the neighboring regions in the vacuum and in the semiconductor are large. In that case the states in the region near the surface have the form of a half-integral number of sinusoids, and can be labeled by a perpendicular wavevector given by $k_{\mu} = \pi \mu / \ell$.
For large ℓ the sum over μ in Eq. (11) can be expressed as an integral over E_{μ} times $\ell m^* / (\pi k_{\mu} \hbar^2)$, from which Eq. (11) reduces to Eq. (9).
- ¹⁸ Y. Dong, R. M. Feenstra, M. P. Semtsiv and W. T. Masselink, Appl. Phys. Lett. **84** 227 (2004).
- ¹⁹ H. Kroemer, J. Vac. Sci. Technol. B **1**, 126 (1983).
- ²⁰ I. Vurgaftman, J. R. Meyer, L. R. Ram-Mohan, J. Appl. Phys. **89**, 5815 (2001).
- ²¹ J. G. Simmons, J. Appl. Phys. **34**, 1793 (1963).
- ²² G. Binnig, N. García, H. Rohrer, J. M. Soler, and F. Flores, Phys. Rev. B **30**, 4816 (1984).
- ²³ R. M. Feenstra, Physica B **273-274**, 796 (1999).

²⁴ Y. Dong, R. M. Feenstra, M. P. Semtsiv and W. T. Masselink, to be published.