Electrostatic Potential for a Hyperbolic Probe Tip near a Semiconductor

R. M. Feenstra
Dept. Physics
Carnegie Mellon University
Pittsburgh, PA 15213

Abstract

The electrostatic potential resulting from a metallic probe tip near a semiconductor is examined. A solution is formulated assuming circular symmetry and using prolate spheroidal coordinates in the vacuum and Cartesian coordinates in the semiconductor. The result is most directly applied to the case of a hyperbolic probe tip, but other shapes (for example, a small hemispherical protrusion on the tip apex) can also be handled. Numerical results are given for representative cases that might be encountered in scanning probe microscopy.

I. Introduction

From the earliest years of scanning tunneling microscopy it was well recognized that application of the technique to a semiconductor sample would result in some of the applied potential between tip and sample being dropped in the semiconductor itself [1,2]. The resulting electrostatic potential distributions were considered early on in a one-dimensional approximation [2]. Sometime later it was noted that a three-dimensional situation was more applicable, and results for that geometry were reported [3,4]. In a semiclassical approximation, the effect of the varying electrostatic potential in the semiconductor is simply to rigidly shift the energy bands. Thus, this effect of the applied potential extending into the semiconductor, in the limit of low current, has become known generally as "tip induced band bending".

  Electrostatic solutions for the potential around a sharp probe tip close to a semiconductor surface are nontrivial and consequently most of the prior theoretical results for this problem, although quite adequate for their purposes, have nonetheless employed various approximations to the tip shape [4-8]. For example, in our prior work on this problem we assumed that the slope of the equipotentials in the vacuum were small $[(dV/dz)/(dV/dr)<1$ where V is the potential, z is the direction normal to the semiconductor surface, and r is the radial direction] [4,8]. This approximation resulted in a somewhat ill-defined shape for the shank of the probe tip. As the field of scanning probe microscopy progresses it becomes ever more important to rigorously understand the potential distribution between tip and sample, not only in scanning tunneling microscopy but also for other scanning probe microscopies in which the potential distribution can play an important role in determining the measured signal [5,6,9]. Therefore, an improved solution for this electrostatics problem is needed. Modern Poisson solvers have been used for this purpose [6,9], but the solution formulated in the present work, since it is specialized to the particular problem at hand, is likely to be more accurate and efficient.

    In this work we present a solution for the electrostatics problem of a probe tip near a semiconductor, employing a prolate spheroidal coordinate system in the vacuum region

between tip and sample [10]. This type of coordinate system is well-suited to this problem; it has been previously used for the problem of field emission [11] as well as other problems involving a sharp tip in proximity to a metallic surface [12]. Here, we apply such coordinates to the case of the tip-semiconductor problem as encountered in scanning probe microscopy. We present here a complete description of the background theory needed for implementing the solution in a computer program. An important aspect of our solution is that, in additional to the usual specification of tip-sample separation and tip radius of curvature, the slope of the tip shank (or, equivalently, the opening angle of the shank) can be specified as an input parameter. We are thus able to investigate *e.g.* the approach to the one-dimensional limit as this slope approaches zero. Our final solution to this electrostatics problem is an iterative, finite difference one, having the usual form in which the potential at a given point at each iteration is assigned to be a weighted average of the potential of the surrounding grid points [13]. This iterative solution also, of course, includes a contribution arising from the nonzero charge density in the semiconductor; in this case we have found it necessary to devise a novel iterative updating scheme to ensure convergence of the result.

In the following section we present the underlying equations which form the basis for our solution of the electrostatics problem. In section III we present numerical results, considering primarily the regime of semiconductor depletion but also including some comments regarding accumulation and/or inversion. Our work is summarized in section IV.

II. Theory
A. Semiconductor Charge Density
The equations describing the charge density in the semiconductor within effective mass approximation are well known. For the sake of completeness we summarize them here, as taken from Sze [14], written for the case of a single conduction band and a twofold valence band including heavy and light holes. First, for the charge density in the conduction band we have

$$n = N_C \frac{2}{\sqrt{\pi}} F_{1/2}\left(\frac{E_F - E_C}{kT}\right) \tag{1}$$

where E_F is the Fermi-level, E_C is the conduction band minimum, T is the temperature, k is Boltzmann's constant, and $F_{1/2}$ is a Fermi-Dirac integral of order 1/2. Also, $N_C = 2(2\pi m_{de}kT/h^2)^{3/2}$ where m_{de} is the density-of-state effective mass for electrons, and h is Planck's constant. Similarly, for the charge density in the valence band we have

$$p = N_V \frac{2}{\sqrt{\pi}} F_{1/2}\left(\frac{E_V - E_F}{kT}\right) \tag{2}$$

with $N_V = 2(2\pi m_{dh}kT/h^2)^{3/2}$ with m_{dh} being the density-of-state effective mass for holes. The factor of 2 following the equals sign in the expressions for N_V and N_C arises from the spin degeneracy (the degeneracy of heavy hole and light hole bands is included in the definition of m_{dh}). For the charge density due to ionized donors we have for the case of nondegenerate doping concentration,

$$N_D^+ = \frac{N_D}{1 + 2\exp\left(\dfrac{E_F - E_D}{kT}\right)} \tag{3}$$

2

where N_D is the concentration of donors, E_D is the donor energy level and the factor of 2 before the exponential term accounts for the spin degeneracy. In the case of degenerate doping we have simply $N_D^+ = N_D$. Finally, for the charge density due to ionized acceptors we have for nondegenerate doping

$$N_A^- = \frac{N_A}{1 + 4\exp\left(\dfrac{E_A - E_F}{kT}\right)} \tag{4}$$

where N_A is the concentration of acceptors and E_A is the acceptor energy level. The factor of 4 before the exponential in this case arises from the *total* degeneracy, including both spin and heavy and light hole bands. For degenerate doping, $N_A^- = N_A$. From the above equations we can construct the total charge density in the semiconductor, as a function of the Fermi level position E_F,

$$\rho(E_F) = e\,[\,p(E_F) + N_D^+(E_F) - n(E_F) - N_A^-(E_F)\,] \tag{5}$$

where $e = +1.602 \times 10^{-19}$ C and the dependence on E_F of each of the constituent charge densities is clear from Eqns. (1-4). To solve for the value of E_F in the semiconductor, one simply searches for a solution of $\rho(E_F) = 0$ (*i.e.* in the bulk of the semiconductor, for flat-band conditions). The value of E_F thus obtained is a function of temperature, *i.e.*, it is the chemical potential of the system.

In our solution to the tip-semiconductor electrostatics problem discussed below we assume that E_F is constant throughout the semiconductor, even in the presence of a varying electrostatic potential. For situations with nonzero current it is not obvious whether this assumption is valid, *i.e.* one might have a "spreading resistance" type of potential drop in the semiconductor in addition to the electrostatic potential distribution we are computing in the present work. This type of resistive drop was discussed in a very early work within the context of tunneling into the Si(111)7×7 surface [1], although numerous subsequent experiments have demonstrated that the resistive drop (at least up to currents in the nA-range) is negligible for this and other surfaces dominated by surface-states insofar as accurate spectroscopy data can be routinely acquired [15-17]. [A possible exception is the SiC(0001) surface for which the spectroscopy results are found to be quite dependent on the magnitude of the current] [18]. Similar results for the GaAs(110) surface in which the spectroscopy results do not appear to be particularly dependent on current [2] indicate, again, that a resistive type of potential drop in the material is small. Thus, we tentatively conclude that for most scanning tunneling microscopy experiments the transport of carriers within the semiconductor does not appear to produce a significant potential drop, so that E_F can indeed be taken as constant in the material.

To include the effect of a nonzero electrostatic potential on the semiconductor charge density, we use a semiclassical approximation in which the electron energy bands are shifted rigidly by an energy $\phi = -eV$ due to an electrostatic potential V. All of the energies E_C, E_V, E_A, E_D and E_F given above are referenced to some fixed point in the energy bands. If we now shift the bands by ϕ, but maintain a constant E_F, the new Fermi-level position *relative to the energy bands* will be $E_F - \phi$. The resulting charge density is then given by $\rho(E_F - \phi)$. Expressed in terms of the electrostatic potential energy ϕ, Poisson's equation in the semiconductor is

$$\nabla^2 \phi = \frac{e\rho(E_F - \phi)}{\varepsilon \varepsilon_0} \tag{6}$$

where ε_0 is the permittivity of vacuum, and ε is the dielectric constant of the semiconductor. The zero reference for ϕ is taken to be at a point far inside the semiconductor.

The boundary condition on Eq. (6) we use throughout this paper is the difference in electrostatic potential energy between the tip and a point far inside the semiconductor, ϕ_T. This quantity equals the difference in vacuum level positions between the tip and semiconductor, and it is formed from a sum of the applied voltage on the sample relative to the tip, V_S, and the difference between the work functions of sample and tip. The work function of the metallic tip we denote by ϕ_m, and that of the sample is given by $\chi + (E_C - E_F)$ where χ is the electron affinity of the semiconductor. With these definitions, and using the fact that the applied voltage equals the difference in Fermi-level positions of tip and sample, we have

$$\phi_T = eV_S + \phi_m - \chi - (E_C - E_F) \ . \tag{7}$$

With this equation one can, for a specific physical system, relate the experimental parameter V_S to the theoretical quantity ϕ_T used in the present work.

For a one-dimensional geometry the problem of tip induced band bending under depletion conditions can be easily solved analytically, at least for zero temperature. The potential difference ϕ_T between tip and sample will be dropped partly across the vacuum gap of width s and partly within the semiconductor over a depletion distance w. Within the semiconductor the electrostatic potential energy varies quadratically with distance,

$$\phi(z) = \frac{e^2 (N_D - N_A)}{2\varepsilon \varepsilon_0}(w + z)^2 \tag{8}$$

where the semiconductor exists in the region $z < 0$. The potential energy at the surface is then $\phi_{00} = e^2(N_D - N_A)w^2 /(2\varepsilon\varepsilon_0)$ and the electric field at the surface in the semiconductor is

$$E_S = \frac{e(N_D - N_A)}{\varepsilon \varepsilon_0} w \ . \tag{9}$$

In the vacuum the field is uniform, with value $(\phi_T - \phi_{00})/(es)$. Applying the boundary condition of continuity of the electric displacement across the surface, we have

$$e(N_D - N_A) w = \frac{\varepsilon_0}{es}\phi_T - \frac{e(N_D - N_A)}{2\varepsilon s} w^2 \tag{10}$$

thus yielding a quadratic equation for w. This equation is easily solved, yielding one-dimensional values for the depletion width, $w = w_0$, and the surface potential ϕ_{00}.

B. Probe Tip Geometry and Coordinate System

To handle the electrostatic problem of a metallic probe tip in proximity to a semiconducting surface, we use in the vacuum region the well-known prolate spheroidal coordinate system [10-12]. These coordinates form families of confocal hyperbolas and ellipses, as illustrated in Fig. 1(a). For the case of a metallic sample this choice of

coordinates immediately provides an exact solution of the electrostatic problem [11]. For the present problem of a semiconducting sample the coordinate system does not present a trivial solution to the problem, but it nevertheless allows a convenient (and exact) means of specifying the boundary condition in the vacuum and it also enables the use of a moderate size finite-element grid in the vacuum. In this section and the following one we discuss a problem in which the probe tip fits exactly on one of the hyperbolas of the coordinate system; in Section D we consider the cases shown in Figs. 1(b) and (c) for which the probe tip shape does *not* have this constraint.

We label the prolate spheroidal coordinates by (ξ, η), $1 \leq \xi < \infty$ and $0 \leq \eta < 1$, with constant-ξ surfaces consisting of ellipsoids and constant-η surfaces consisting of hyperboloids. In terms of the original cylindrical coordinate (r,z), the constant-ξ ellipsoids satisfy the equation

$$\frac{r^2}{\xi^2 - 1} + \frac{z^2}{\xi^2} = a^2 \tag{12}$$

and the constant-η hyperboloids satisfy

$$-\frac{r^2}{1 - \eta^2} + \frac{z^2}{\eta^2} = a^2 \tag{13}$$

where a is the distance from the semiconductor surface to the common focal point of these iso-surfaces. Transformation equations between the cylindrical and prolate spheroidal coordinates are given by

$$r = a[(\xi^2 - 1)(1 - \eta^2)]^{1/2} \tag{14}$$

and

$$z = a\xi\eta \quad . \tag{15}$$

The specific hyperboloid which corresponds to the probe tip has an η-value of $\eta = \eta_T$. These two parameters a and η_T are determined by the values of the sample-tip separation s and the tip radius of curvature R according to

$$s = a\eta_T \tag{16}$$

and

$$R = a\frac{1 - \eta_T^2}{\eta_T} \tag{17}$$

where R is given by the inverse of d^2z/dr^2 for $\eta = \eta_T$ and at $r = 0$, evaluated from Eqs. (14) and (15). The probe tip thus defined has a specific shank slope, given by

$$b = \sqrt{\frac{s}{R}} \quad . \tag{18}$$

(Tips with other values of b are discussed in Section D). The opening angle of the tip is given by $\theta = 2\tan^{-1}(1/b)$.

C. Poisson's Equation

Assuming circular symmetry, Laplace's equation in the vacuum is given by

$$\frac{\partial}{\partial\xi}\left[(\xi^2 - 1)\frac{\partial\phi}{\partial\xi}\right] + \frac{\partial}{\partial\eta}\left[(1 - \eta^2)\frac{\partial\phi}{\partial\eta}\right] = 0 \quad . \tag{19}$$

On a discrete grid of points (ξ_i, η_j), $i = 1, 2,... m$, $j = 1, 2,... n$, we put this equation into finite-difference form according to

$$\left[\frac{2(\xi_i^2 - 1)}{(\Delta\xi_i)^2} + \frac{2(1 - \eta_j^2)}{(\Delta\eta_j)^2}\right]\phi_{ij} = \frac{(\xi_i^2 - 1)}{(\Delta\xi_i)^2}(\phi_{i+1,j} + \phi_{i-1,j}) + \frac{(1 - \eta_j^2)}{(\Delta\eta_j)^2}(\phi_{i,j+1} + \phi_{i,j-1})$$

$$+ \frac{\xi_i}{\Delta\xi_i}(\phi_{i+1,j} - \phi_{i-1,j}) - \frac{\eta_j}{\Delta\eta_j}(\phi_{i,j+1} - \phi_{i,j-1}) \quad .$$

(20)

At the semiconductor surface ($\eta = 0$) the values of ξ_i are chosen to match values of $r_i = (i - 0.5)\Delta r$, where $\Delta r = r_{max}/m$ with r_{max} being the lateral extent of our electrostatics solution. Thus,

$$\xi_i = \sqrt{1 + \left(\frac{r_i}{a}\right)^2}$$

(21)

so that

$$\Delta\xi_i = \sqrt{1 + \left(\frac{i\Delta r}{a}\right)^2} - \sqrt{1 + \left(\frac{(i-1)\Delta r}{a}\right)^2}$$

$$\approx (i - 0.5)\left(\frac{\Delta r}{a}\right)^2 \quad \text{for} \quad i\Delta r << a \quad .$$

(22)

The values of η_j are chosen to be uniformly spaced, $\eta_j = j\eta_T/n$ so that $\Delta\eta_j \equiv \Delta\eta = \eta_T/n$.

Strictly speaking Eq. (20) is not valid for $i = 1$ or m and for $j = 1$ or n. However, we can formally extend the equation to those cases using the following extended potential values: For $i = 1$ the values of $\phi_{0,j}$ are identical to $\phi_{1,j}$ due to the circular symmetry and for $i = m$ the values of $\phi_{m+1,j}$ are taken to be identical to $\phi_{m,j}$. This latter equality is well satisfied so long as the lateral extent of the solution is large enough so that the potential on the surface at $r = r_m$ is practically zero, since in that case the potential in the vacuum is independent of i (see the discussion at the end of this Section on the initial value used for the potential). For $j = 1$ in Eq. (20) the values of $\phi_{i,0}$ correspond to those on the semiconductor surface, defined below, and for $j = n$ we have $\phi_{i,n} = \phi_T$.

To put Eq. (20) into a form suitable for iterative solution we assign the value of $\phi_{i,j}$ on the left-hand side to be its value on the $(k+1)^{st}$ iteration, with all the ϕ values on the right-hand side of the equation being those on the kth iteration. We have tested this update formula over a wide range of tip geometries and we find it to be stable in all cases. An important aspect of this stability is the choice for the first derivative terms in Eq. (20) being $(\phi_{i+1,j} - \phi_{i-1,j})/(2\Delta\xi_i)$ as opposed to $(\phi_{i+1,j} - \phi_{i,j})/\Delta\xi_i$ or $(\phi_{i,j} - \phi_{i-1,j})/\Delta\xi_i$. This choice gives the exact result for the first derivative value at $i = 1$ in the limit of small Δr for the case of a quadratic dependence of ϕ on r.

In the semiconductor we use a Cartesian grid with cylindrical symmetry to solve Poisson's equation [Eq. (6)]. In discrete form this equation becomes

$$\left[\frac{2}{(\Delta r_i)^2} + \frac{2}{(\Delta z_j)^2} \right] \phi_{ij} = \frac{1}{(\Delta r_i)^2}(\phi_{i+1,j} + \phi_{i-1,j}) + \frac{1}{(\Delta z_j)^2}(\phi_{i,j+1} + \phi_{i,j-1})$$

$$+ \frac{1}{(2r_i \Delta r_i)}(\phi_{i+1,j} - \phi_{i-1,j}) - \frac{e}{\varepsilon \varepsilon_0} \rho_{ij} \quad . \tag{23}$$

with $\rho_{ij} \equiv \rho(E_F - \phi_{ij})$. The values of r_i in this equation are identical to those in Eq. (20), so that $\Delta r_i \equiv \Delta r$ as defined above. The values of z_j, $j = 1, 2, \ldots N$ are chosen to be uniformly spaced with $z_j = -j \, z_{max} / N$ and $\Delta z_j \equiv \Delta z = z_{max} / N$ where z_{max} is the vertical extent of our solution. In practice, we choose both z_{max} and r_{max} to be equal to w_0, or some multiple thereof. The application of Eq. (23) to the values $i = 1$, $i = m$, and $j = 1$ is made in the same way as for Eq. (20). Equation (23) is *not* applied to the case of $j = N$. For the case of $j = N-1$, the values of $\phi_{i,N}$ appearing on the right-hand side of Eq. (23) are assumed to equal zero; this assumption is justified so long as the vertical extent of the solution is sufficiently large.

To formulate Eq. (23) into an update equation for the potential one can envision several possibilities. First, we could assign the left-hand side of Eq. (23) to be the potential at the $(k+1)^{st}$ iteration, $\phi_{ij}^{(k+1)}$, with the right hand side being the values for the potential and charge density at the kth iteration. We denote the resulting updated value of the potential as $\tilde{\phi}_{ij}^{(k+1)}$. We have performed numerical tests of this updating scheme and find that the potential values thus obtained approach their final values in a markedly oscillatory manner. At the very least this behavior makes it difficult to judge the convergence of the result, and in the worst case an instability in the solution may result. This problem arises in part from the nonlinear dependence of $\rho(E_F - \phi)$ on ϕ. We have developed a method by which this behavior is suppressed: In the updating formula, for the charge density, we use its value *at the $(k+1)^{st}$ iteration* rather than at the kth iteration. This leads, of course, to a transcendental equation that must be solved for $\phi_{ij}^{(k+1)}$:

$$\left[\frac{2}{(\Delta r_i)^2} + \frac{2}{(\Delta z_j)^2} \right] \phi_{ij}^{(k+1)} = \frac{1}{(\Delta r_i)^2}(\phi_{i+1,j}^{(k)} + \phi_{i-1,j}^{(k)}) + \frac{1}{(\Delta z_j)^2}(\phi_{i,j+1}^{(k)} + \phi_{i,j-1}^{(k)})$$

$$+ \frac{1}{(2r_i \Delta r_i)}(\phi_{i+1,j}^{(k)} - \phi_{i-1,j}^{(k)}) - \frac{e}{\varepsilon \varepsilon_0} \rho(E_F - \phi_{ij}^{(k+1)}) \quad . \tag{24}$$

At each iteration we perform a one-dimensional search of possible $\phi_{ij}^{(k+1)}$ values that satisfy this equation. The solution is assumed to be bounded by $\phi_{ij}^{(k)}$ and $\tilde{\phi}_{ij}^{(k+1)}$, and the solution is located to an accuracy of $\left| \phi_{ij}^{(k)} - \tilde{\phi}_{ij}^{(k+1)} \right| / 10$. This procedure is found to yield in all cases a stable final solution for the potential, and one in which the final result is approached monotonically.

At the semiconductor surface we apply the boundary condition of continuity of the electric displacement. Using the grids defined above, this condition takes on the form

$$\varepsilon_0 \frac{(\phi_{i,1}^V - \phi_i)}{a\xi_i \Delta\eta} = \varepsilon\varepsilon_0 \frac{(\phi_i - \phi_{i,1}^S)}{\Delta z} \qquad (25)$$

with the $\phi_i \equiv \phi_{i,0}$ values being the potential energy on the surface and where we have added the superscripts V and S to denote the potentials in the vacuum and semiconductor, respectively. The expression $a\xi_i \Delta\eta$ in the denominator of the left-hand side equals the grid spacing in the z-direction between the surface and the first η-value in the vacuum. Equation (25) is solved for ϕ_i, with the resulting expression used as an update formula for ϕ_i in the $(k+1)^{st}$ iteration given values for the potentials in the vacuum and semiconductor in the kth iteration.

To initialize the values of the potential in the 0th iteration, we use $\phi \equiv 0$ everywhere in the semiconductor and on its surface. In the vacuum, we use an expression corresponding to the solution for the potential between a metallic tip and metallic sample [11],

$$\phi_{ij}^{(0)} = \phi_T \frac{\ln\left[\dfrac{1+\eta_j}{1-\eta_j}\right]}{\ln\left[\dfrac{1+\eta_T}{1-\eta_T}\right]} \quad . \qquad (26)$$

With this choice, iterative values of the potential approach their final value, in terms of magnitudes, from below. One could also use some other initialization scheme (e.g. based on a one-dimensional electrostatics solution) in order to approach the correct value from above, thereby providing a useful test on the convergence of the final result.

D. Case of a General Hyperbolic Tip

The theory described in the previous two sections allows us to solve for the electrostatic potential for the case when the probe tip exactly matches one of the constant-η hyperboloids of the coordinate system. Using the tip radius R and the tip-sample separation s as parameters, the slope of the tip shank is then predetermined, with a value of $b = \sqrt{s/R}$ as in Eq. (18). In this section we consider the general case of a hyperbolic tip with arbitrary values of R, s, and b. In this case the tip may *not* correspond to a constant-η surface, and some adjustment of our procedure is required.

We distinguish between two possible cases here: one in which the specified value of shank slope is greater than $\sqrt{s/R}$ and the other in which it is less than $\sqrt{s/R}$. In the first case, illustrated in Fig. 1(b), we choose a coordinate system based on the specified values of R and b but for which the separation between sample and the "coordinate tip" which forms the basis of the coordinate system is *not* given by s. We use the subscript 1 to denote this coordinate tip and we have in this case

$$b_1 = b = \frac{\eta_1}{\sqrt{1-\eta_1^2}} \qquad (27)$$

and

$$R_1 = R = a_1 \frac{1-\eta_1^2}{\eta_1} \quad . \tag{28}$$

These equations determine the values of a_1 and η_1, with identical values applying to the actual tip, $a = a_1$ and $\eta_T = \eta_1$. The value of the separation between sample and coordinate tip is given by $s_1 = R_1 b_1^2$, with $s_1 > s$. Our boundary condition for solving Laplace's equation in the vacuum must now be modified, such that at each grid point i,j in the vacuum we check whether that point falls within the interior of the tip and if so then no action is taken. To achieve an explicit statement for this condition, we transform the values of (ξ_i, η_j) into (r,z) coordinates according to Eqs. (14) and (15),

$$r_{ij} = a_1 [(\xi_i^2 - 1)(1 - \eta_j^2)]^{1/2} \tag{29}$$

and

$$z_{ij} = a \xi_i \eta_j \quad . \tag{30}$$

We then check whether the z_{ij} value thus obtained is greater than that corresponding to the periphery of the actual tip,

$$z_{ij} > \eta_T \left[\frac{r_{ij}^2}{1 - \eta_T^2} + a^2 \right]^{1/2} - a\eta_T + s \quad . \tag{31}$$

If this condition is met then the point falls within the interior of the actual tip and so the value of ϕ_{ij} is left at ϕ_T. If the condition is not met then the usual updating is performed.

The second case we must consider for a general hyperbolic tip is illustrated in Fig. 1(c). Now, the specified value of the shank slope is less than $\sqrt{s/R}$. In this case we choose a coordinate tip with

$$s_1 = s = a_1 \eta_1 \tag{32}$$

and

$$R_1 = R = a_1 \frac{1-\eta_1^2}{\eta_1} \quad . \tag{33}$$

These equations determine the values of a_1 and η_1 for the coordinate tip. For the actual tip, the values of a and η_T are determined by the specified slope b and radius R according to

$$\frac{1}{\eta_T} = \sqrt{1 + \frac{1}{b^2}} \tag{34}$$

and

$$a = \frac{Rb^2}{\eta_T} \quad . \tag{35}$$

The value of the shank slope for the coordinate tip is given by $b_1 = \sqrt{s_1/R_1}$, with $b_1 > b$. Our modification to the boundary condition for solving Laplace's equation then reduces to identically the same form as Eq. (31). If that condition is met then no action is taken at that grid point (*i.e.* the value of ϕ_{ij} is maintained at ϕ_T), and if not then the usual updating is performed.

The case of a general hyperbolic tip also requires some modification of the initial condition, discussed at the end of the previous Section. We modify Eq. (26) by replacing

η_j by η'_{ij} where we define η'_{ij} such that it varies from 0 at the semiconductor surface to η_T at the surface of the actual probe tip. Thus, $\eta'_{ij} = j\,\eta_T / n'_i$ where n'_i is the number of grid points extending between the surface and the actual tip, at the ith value of ξ_i.

E. Case of a Tip with Arbitrary Shape
With the theory developed above for a general hyperbolic tip, it is straightforward to now consider modifications to the probe tip shape. One simply defines the tip periphery as being some new surface which is entirely contained in the vacuum, and the coordinate tip is chosen as the maximal hyperbolic shape which is contained entirely within this actual tip. As an example, let us consider a small protrusion extending beyond the apex of a general hyperbolic tip. The existence of such a protrusion has been previously discussed within the context of scanning tunneling microscopy [4]. Let $p(r)$ describe the z-extent of this protrusion relative to the surface of the hyperbolic tip. Then this protrusion can be included in our computation simply by modifying Eq. (31) to include this additional factor,

$$z_{ij} > \eta_T \left[\frac{r_{ij}^2}{1-\eta_T^2} + a^2 \right]^{1/2} - a\eta_T + s - p(r_{ij}) \quad . \tag{36}$$

As in the prior Section, if this condition is met then no action is taken at that grid point, and if not then the usual updating is performed. For more substantial modifications to the tip shape which significantly modify the shank geometry, *e.g.* using a parabolic rather than a hyperbolic shank, the method described here would not be so applicable and the use of a different coordinate system to match the shank geometry should be explored.

III. Results and Discussion
In this section we discuss results for the distribution of electrostatic potential energy for specific tip-sample systems. Many of our results will be presented in terms of the surface potential energy at a point directly opposite the apex of the probe tip; we denote this value by ϕ_0. As discussed in Section II(C), the radial grid points in our simulation are specified by $r_i = (i-0.5)\Delta r$, $i = 1, 2, ..., m$ so that we do *not* have a grid point precisely at $r = 0$. We therefore use quadratic interpolation to form ϕ_0, $\phi_0 = (9\phi_1 - \phi_2)/8$, where ϕ_1 and ϕ_2 are the surface potentials at r_1 and r_2, respectively.

Figure 2 displays results for the surface potential for n-type material with doping concentration of 1×10^{18} cm^{-3}, in depletion with a tip electrostatic potential energy of 3 eV relative to a point far inside the semiconductor (the semiconductor band gap is chosen to be large enough so that inversion does not occur). The sample-tip separation is taken to be 1 nm. A dielectric constant of $\varepsilon = 12.9$ (corresponding to GaAs) is used for the semiconductor, and for illustrative purposes the donor level is placed 1 meV below the conduction band mimimum, using nondegenerate statistics. The temperature used in the computation is 0 K. The one-dimensional surface potential for these parameters, from Eq. (10), is 2.03 eV with depletion width of $w_0 = 53.8$ nm. In Fig. 2 the surface potential ϕ_0 is plotted as a function of the probe tip radius, using various opening angles for the tip: 30°, 90°, 150°, and a variable angle chosen to exactly match the eccentricity of the coordinate system used in the computation. [The slope b of the tip shank is related to the opening

angle θ by $1/b = \tan(\theta/2)$]. The results shown in Fig. 2, and those below for depletion conditions, vary with the tip potential in a nearly linear manner.

A number of features of the three-dimensional band bending problem are apparent in Fig. 2. First, it is seen that even for rather large tip radii of 1000 nm the one-dimensional limit is not yet reached. Typical tip radii in scanning tunneling microscopy are in the range 10–100 nm [8], so it is clear that the one-dimensional limit is a poor approximation in reality. The effect of shank slope is seen in the results particularly for small values of the tip radius. In the limit of small radii our geometry reduces to that of a wedge of specified opening angle in proximity to the semiconductor. As the opening angle approaches 180° then one must reach the one-dimensional limit, and this behavior is apparent *e.g.* in the difference between the 90° and 150° curves in Fig. 2. Considering the limit of small opening angles for the results plotted in Fig. 2, *i.e.* the lower limit of the potential values, our geometry presumably approaches that of a metallic sphere in proximity to the semiconductor. Comparing the results in Fig. 2 with our prior results of Ref. [8] (which, as discussed in the Introduction, used an approximation resulting in a somewhat ill-defined tip shank geometry), we find good agreement with the results for a tip opening angle of $\approx$90°.

Figure 3 shows specific results for a tip radius of 4 nm and opening angle of 90°, with parameters for the semiconductor identical to those of Fig. 2. Figure 3(a) gives a contour plot of the electrostatic potential energy, for potentials of 0.5, 1.0, and 2.0, and with the probe tip at a potential of 3.0 eV relative to a point far inside the semiconductor. Figure 3(b) shows the potential along the central axis and Fig. 3(c) shows the potential along a line on the surface of the semiconductor and passing through the central axis.

Figure 4 shows results for the bend bending for a lower doped semiconductor, with doping concentration of 1×10^{16} cm^{-3} and all other parameters identical to those of Fig. 2. The qualitative trends in Fig. 4 agree with those of Fig. 2, but the surface potentials in Fig. 4 are all somewhat larger due to the lower doping concentration. The one-dimensional surface potential in this case is 2.88 eV with depletion width of $w_0 = 641$ nm. As for Fig. 2, the results of Fig. 4 emphasize the fact that the one-dimensional limit is inappropriate for realistic tip geometries. Comparing Figs. 2 and 4 we see that the change in surface potential due to this two-order-of-magnitude change in doping concentration is typically 0.5 eV, which is similar to the change in potential when varying the tip radius by one order of magnitude. We thus conclude that the divergence of the potential in the semiconductor is affected by both three-dimensional spreading of the field lines away from the tip as well as the one-dimensional curvature arising from the space charge of the ionized dopant atoms, with the former dominating for small tip radii. Only for very large tip radii and/or shank opening angles very close to 180° is the one-dimensional limit approached.

In Fig. 5 we show another perspective on the band bending problem, namely, its dependence on sample-tip separation. As $s \to 0$ the surface potential ϕ_0 must approach ϕ_T for all values of tip radius and shank angle. The manner in which this limit is approached will, however, vary depending on the tip geometry and the semiconductor doping concentration, as illustrated in Fig. 5. The three-dimensional results $(\theta = 90°)$ are very nonlinear as $s \to 0$, showing a sharp increase towards the value of ϕ_T=3 eV only at

separations significantly less than 1 nm. By the same token, the electric field in the vacuum reaches rather large values at these small separations.

All of the results discussed above have dealt with the semiconductor in depletion. We comment briefly here on the situation when the semiconductor is in accumulation or inversion. As noted previously [8], the tip induced band bending in accumulation is relatively independent of tip geometry, since the semiconductor charge densities near the surface are large and these dominate in determining the solution to Poisson's equation. Using the theory described in Section II we can compute the potential distributions within the semi-classical approximation in which the charge density in the semiconductor is assumed to equal $\rho(E_F - \phi)$ where ϕ is the electrostatic potential energy. In reality the accumulation layers states are, of course, quantized, by confining potentials in both the vertical (z-direction) and lateral (radial) directions. The resulting discrete accumulation layer states have been directly resolved in scanning tunneling spectroscopy at low temperatures [7,20]. The semi-classical approximation is expected to be valid whenever many quantum states are occupied, *i.e.* in conditions of severe accumulation or inversion. The solution thus obtained is also clearly self-consistent.

In accumulation or inversion situations when only a few quantum states are occupied the computation of the electrostatic potential becomes more complicated for two reasons. First, one must perform a Hartree computation in which the charge density in the semiconductor is computed as a sum of state-densities of the occupied states (as in Ref. [7]), and second, the potential and associated charge densities must be computed self-consistently. This type of computation takes us beyond the scope of the present work. Nevertheless, we do present here a few simple semi-classical computations of the energy level separation due to the vertical and lateral confinement, which are useful at least for estimating the number of occupied states. We initially consider only vertical confinement, in which case the quantum states are easily computed using the WKB approximation. Using the potential along the central axis as a barrier for the carriers and assuming no penetration into the vacuum, then as previously described [20] the energies $E_n < 0$ of the confined states satisfy

$$\int_L^0 \sqrt{\frac{2m^*}{\hbar^2}\left[E_n - \phi(z)\right]}\,dz - \frac{3\pi}{4} = \pi n \qquad (37)$$

with $z = L$ being the turning point of the potential where $E_n = \phi(L)$ (recall that $z < 0$ in the semiconductor), and with $n = 0$ for the lowest lying state. Figure 6 displays results for the energy level positions as a function of the tip potential energy, using $m^*=0.1m_0$. We consider a negative tip potential energy of -3 eV, so that the n-type material is in accumulation. For the tip radii displayed in Fig. 6 we find 3 bound states. Since we are considering only vertical confinement each state corresponds to the bottom of a subband of states in which the carriers can freely move parallel to the surface.

As a prelude to computing the complete three-dimensional confinement of the carrier wavefunction, we first consider vertical confinement using a variational approach. We use the well-known form for the wavefunction appropriate to triangular potentials [21], $\psi = A z \exp([z/\beta]^{3/2})$ where the constant A is determined by normalization. We compute the energy in the usual manner and vary the parameter β to minimize the energy. The results are shown by the dashed line in Fig. 6; for example, for $R = 1$ nm we find $\beta = 4.30$

nm and a ground state energy of −54 meV. As seen in Fig. 6 we find good agreement between the variational result and the lowest-lying state from the WKB computation, thus giving us some confidence that the energy splittings and number of states found in the WKB computations are reasonably accurate.

We now consider the full three-dimensional confinement of the wavefunction, again using a variational approach. The ground state wavefunction now assumed to have the form $\psi = A\,z\exp([z/\beta]^{3/2})\exp([r/\gamma]^{3/2})$ where again A is determined by normalization and the parameters β and γ are varied so as to minimize the energy. The result for the ground state energy is shown by the dotted line in Fig. 6; for example, for $R = 1$ nm we find $\beta = 5.5$ nm, $\gamma = 8.3$ nm, and a ground state energy of −24 meV. Comparing in Fig. 6 the variational results with and without lateral confinement we find that this effect produces an upwards shift in energy of 20−30 meV. This amount provides an estimate of the radial confinement energy for the wavefunction. In terms of the total number of bound states, considering both vertical and lateral confinement, we conclude that approximately 2 or 3 of them will exist for the tip radii considered here (*i.e.* about 2 from the lowest subband and perhaps 1 from the next higher subband). The system is thus in a moderately quantum regime; with a greater number of occupied bound states the semi-classical approximation might be valid, whereas with fewer bound states a self-consistent, Hartree approach would certainly be necessary to obtain the correct potential distribution.

IV. Summary

In summary, we have presented a finite-difference technique by which the electrostatic problem of a probe tip in proximity to a semiconductor sample can be solved. Using prolate spheroidal coordinates in the vacuum region enables us to rigorously specify the slope of the tip shank, in addition to the usual parameters of tip radius and tip-sample separation. Numerical results for representative situations are presented, from which the approach to the one-dimensional limit as a function of either tip radius or tip slope is apparent. We conclude, consistent with prior works, that the one-dimensional approximation is not appropriate for realistic tip geometries [3,4,8].

The method presented here can be readily extended to other situations with circular symmetry. For example, the problem of a charged threading dislocation lying in the semiconductor, colinear with the tip axis [9,22,23], could be handled by modifying the boundary condition in the semiconductor at $r = 0$ and as $z \rightarrow -\infty$. Similarly, a nonzero surface charge density can be accommodated by modifying the boundary condition at the surface, Eq. (25), by adding to it the surface charge density (which in general will be a function of the surface potential, ϕ_i). For other situations which do *not* have circular symmetry, such as cross-sectional studies of *pn*-junctions or other devices [5,6,8], a significant extension of the present method to include its full angular dependence is required. Nevertheless the use of prolate spheroidal coordinates is still useful in that case for the same reasons as for circular symmetry, namely, it permits a modest size grid to be used for the vacuum region and it enables convenient and exact specification of the boundary condition in the vacuum.

Acknowledgements

References:
[1] F. Flores, N. Garcia, Phys. Rev. B **30**, 2289 (1984).
[2] R. M. Feenstra and J. A. Stroscio, J. Vac. Sci. Technol. B **5**, 923 (1987).
[3] M. McEllistrem, G. Haase, D. Chen, and R. J. Hamers, Phys. Rev. Lett. **70**, 2471 (1993).
[4] R. M. Feenstra, Phys. Rev. B **50**, 4561 (1994).
[5] V. V. Zavyalov, J. S. McMurray, and C. C. Williams, J. Appl. Phys. **85**, 7774 (1999).
[6] M. L. O'Malley, G. L. Timp, S. V. Moccio, J. P. Garno, and R. N. Kleiman Appl. Phys. Lett. **74**, 272 (1999).
[7] R. Dombrowski, Chr. Steinebach, Chr. Wittneven, M. Morgenstern, and R. Wiesendanger, Phys. Rev. B **59**, 8043 (1999).
[8] Y. Dong, R. M. Feenstra, R. Hey and K. H. Ploog, J. Vac. Sci. Technol. B **20**, 1677 (2002).
[9] B. S. Simpkins, D. M. Schaadt, E. T. Yu, R. J. Molnar, J. Appl. Phys. **91**, 9924 (2002).
[10] *Handbook of Mathematical Functions* (Dover, New York, 1965), ed. M. Abromowitz and I. A. Stegun, p. 752.
[11] A. M. Russell, J. Appl. Phys. **33**, 970 (1962).
[12] N. M. Miskovsky, P. H. Cutler, T. E. Feuchtwang, T.E., and A. A. Lucas, Int. J. Infrared and Millimeter Waves **2**, 739 (1981); B. L. Henson, J. Appl. Phys. **55**, 150 (1984); T. L. Ferrell, Phys. Rev. B **50**, 14738 (1994); V. V. Klimov, M. Ducloy, V. S. Letokhov, European Physical Journal D **20**, 133 (2002).
[13] W. H. Hyatt, Jr., *Engineering Electromagnetics* (McGraw-Hill, New York, 1974), p. 185.
[14] S. M. Sze, *Physics of Semiconductor Devices* (Wiley, New York, 1969), Chap. 2.
[15] R. S. Becker, J. A. Golovchenko, D. R. Hamann, and B. S. Swartzentruber, Phys. Rev. Lett. **55**, 2032 (1985).
[16] R. J. Hamers, R. M. Tromp, and J. E. Demuth, Phys. Rev. Lett. **56**, 1972 (1986).
[17] J. A. Stroscio, R. M. Feenstra, and A. P. Fein, Phys. Rev. Lett. **57**, 2579 (1986).
[18] V. Ramachandran and R. M. Feenstra, Phys. Rev. Lett. **82**,1000 (1999).
[19] Using w_0 as an estimate of the size of the region over which the potential is computed may not be appropriate when $R > |\phi_T|$ or when the semiconductor is in accumulation. In those cases a larger region should be used.
[20] R. M. Feenstra, G. Meyer, F. Moresco and K. H. Rieder, Phys. Rev. B **66**, 165204 (2002); there are two misprints associated with Eq. (1) of this work, the "+" sign on the left-hand side should be a "−" sign, and the first state occurs for n=0, not n=1 (the numerical results of the paper are unaffected by these misprints).
[21] T. Ando, A. B. Fowler, and F. Stern, Rev. Mod. Phys. **54**, 437 (1982), Eq. (3.34).
[22] J. W. P. Hsu, M. J. Manfra, D. V. Lang, K. W. Baldwin, L. N. Pfeiffer, and R. J. Molnar, J. Elec. Mat. **30**, 110 (2001).
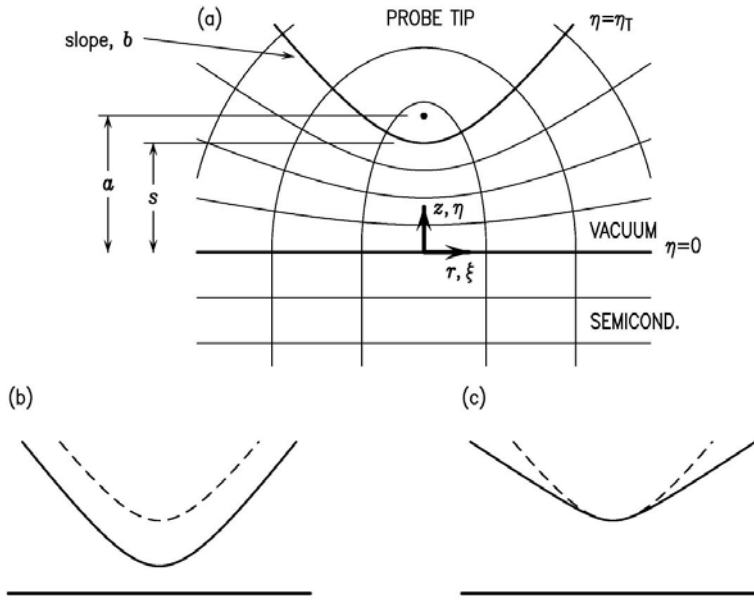[23] C. Tivarus, Y. Ding, and J. P. Pelz, J. Appl. Phys. **92**, 6010 (2002).

Fig. 1. (a) Illustration of prolate spheroidal coordinates (ξ, η) consisting of confocal ellipses and hyperbolas, respectively. The probe tip is given by a hyperbola with $\eta=\eta_t$, having radius of curvature R. The sample-tip distance is s and the sample-focus distance is a. The probe tip shown here exactly corresponds to one of the hyperbolas of the coordinate system, with slope of the tip shank given by $b = \sqrt{s/R}$. Situations when the probe tip does *not* correspond to a hyperbola of the coordinate system are shown in (b) and (c), with the solid line denoting the actual tip and the dashed line showing the "coordinate tip" used as a basis for the coordinate system.
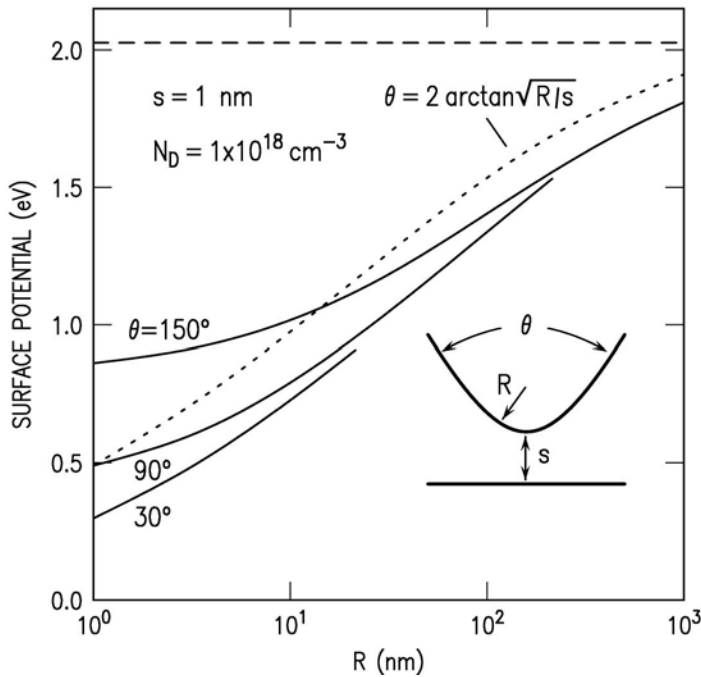
Fig. 2. Electrostatic potential energy at a point on the semiconductor surface directly below the tip apex, for tip-sample separation of 1 nm and n-type doping concentration of 1×10^{18} cm^{-3}. The probe tip has a potential energy of 3 eV relative to a point far inside the semiconductor. Results are shown for shank opening angles of 30°, 90°, 150°, and angles equal to $2\tan^{-1}\sqrt{R/s}$ where R is the tip radius and s the tip-sample separation. The dashed line gives the one-dimensional result (corresponding to an opening angle of 180°).
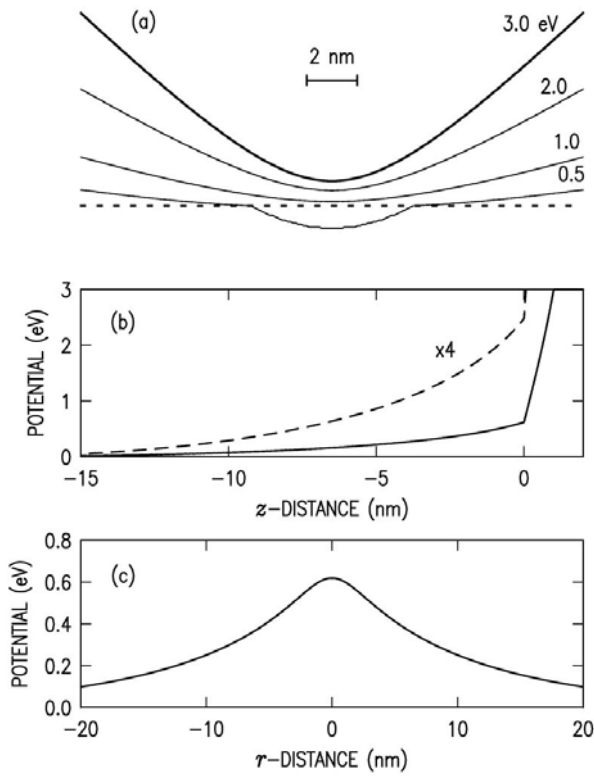
Fig. 3. (a) Potential energy profiles at 0.5, 1.0, and 2.0 eV for a probe tip radius of 4 nm, a tip-sample separation of 1 nm, and an n-type doping concentration of 1×10^{18} cm^{-3}. The probe tip has a potential energy of 3 eV relative to a point far inside the semiconductor. (b) Potential along the central axis, through the semiconductor $(z < 0)$ and the vacuum $(z > 0)$. (c) Potential along a line on the semiconductor surface and passing through the central axis.
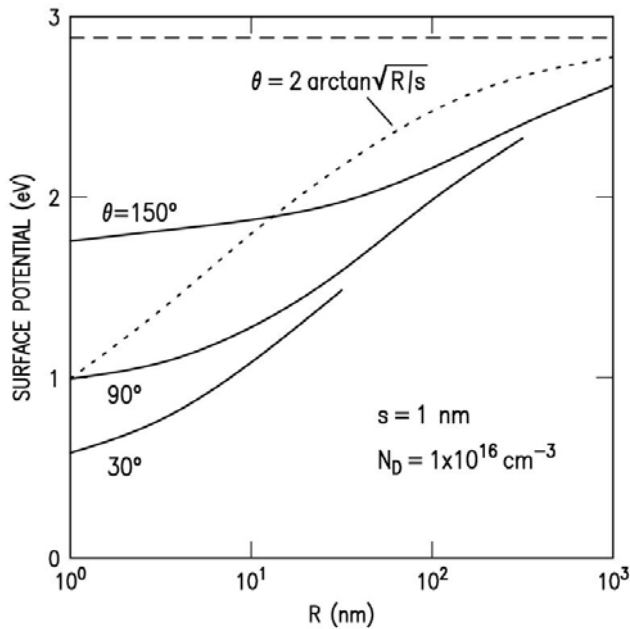

Fig. 4. Same caption as for Fig. 2, but for a n-type doping concentration of 1×10^{16} cm^{-3}.
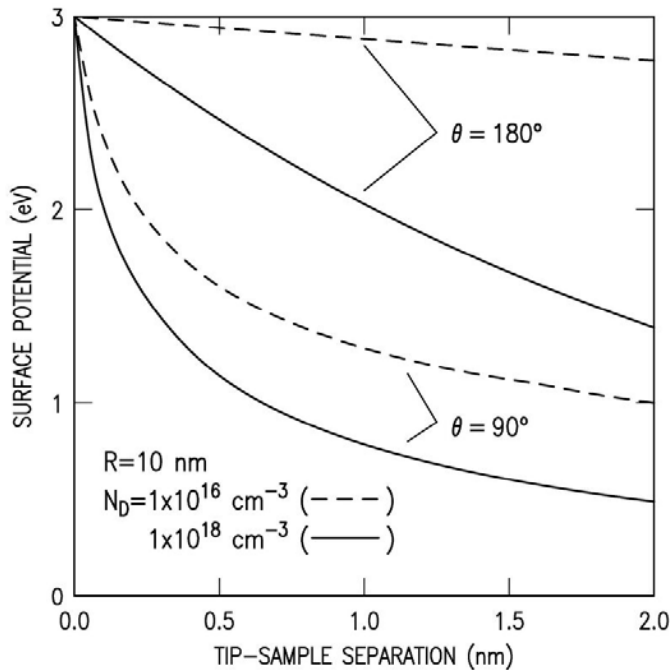
16

Fig. 5. Potential energy at a point on the semiconductor surface directly below the tip apex, as a function of tip-sample separation. The probe tip has a radius of 10 nm and a potential energy of 3 eV relative to a point far inside the semiconductor. Results are shown for shank opening angles of 90° and 180°, and for n-type doping concentrations of 1×10^{16} and 1×10^{18} cm^{-3}.
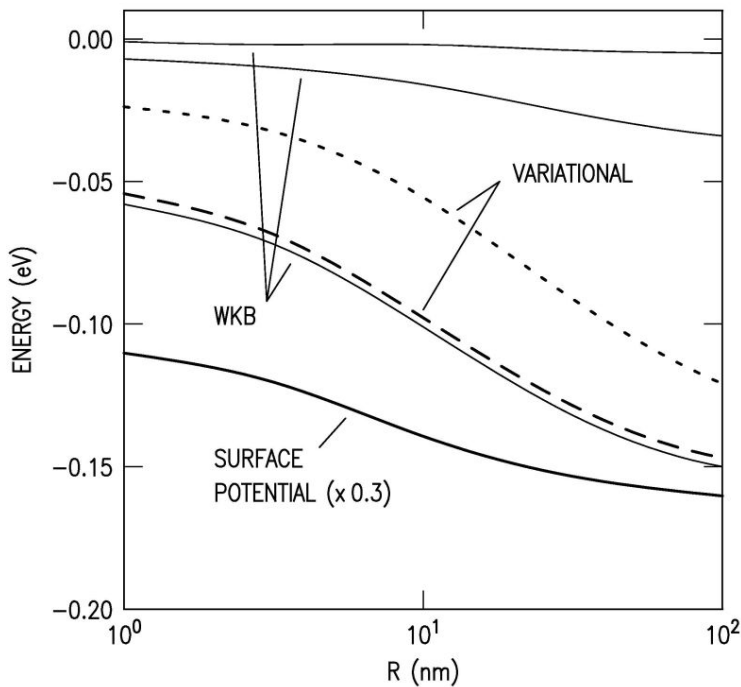

Fig. 6. Potential energy of accumulation layer states for n-type material with a doping concentration of 1×10^{18} cm^{-3} and effective mass of $0.1m_0$. The thick solid line shows the energy of the conduction band minimum at the surface, with thin solid lines showing the energy of localized states as computed in the WKB approximation and considering only vertical confinement. The dashed and dotted lines show variation results for the lowest-lying accumulation layer state, with the former including only vertical confinement and the latter including both vertical and lateral confinement.

17