The Si (100) surface: A theoretical study of the unreconstructed surface

Joel A. Appelbaum, G. A. Baraff, and D. R. Hamann
Bell Laboratories, Murray Hill, New Jersey 07974
(Received 11 December 1974)

The potential, charge density, and energy spectrum of an ideal unreconstructed Si (100) surface is obtained self-consistently. Two bands of surface states that arise from broken surface bonds are found in the energy gap between the valence and conduction bands. One band ($\lesssim 1$ eV wide) has its electron density localized about a line through and normal to the surface atoms, with a node on the atoms; the other band (~ 2.5 eV wide) has its charge located around a line through a row of surface atoms, again having nodes on the surface atoms. The role these surface states play in surface reconstruction as well as in understanding the surface-state spectrum of stepped Si (111)-surface bands is discussed. Additional surface states within the valence band are found and their significance discussed.

I. INTRODUCTION

This is the first in a series of three papers that will attempt to investigate theoretically the electronic properties of the Si (100) surface. The electronic and structural properties of this surface have been studied experimentally over the last 15 years in some detail.[1,2] The surface is known to always occur in a reconstructed form; the so-called 2×1 room-temperature form being the most common.[3]

For our first theoretical investigation of the (100) surface, we considered it prudent to begin with an ideal (1×1) unreconstructed and unrelaxed arrangement of surface atoms. While it is our belief that the electronic spectroscopy of the (2×1) Si (100) surface will differ significantly from that of the hypothetical (1×1) structure, valuable insights into the origins of the structural deformations that occur on this surface can be obtained from a study of the "ideal surface geometry." A study of surface relaxation will be deferred to the second paper in this series and surface reconstruction to a third paper.

The (100) surface, aside from its intrinsic interest, represents a way of studying defect structures on other semiconductor surfaces. For example, the edges of the stepped terraces that occur in nonideal cleaves of the (111) surface[4] are identical to the rows of double broken bonds that occur on the (100) surface, except for their degree of isolation.

The theoretical methods used in this study closely parallel those developed by Appelbaum and Hamann.[5] They employ a numerical integration of Schrödinger's equation together with a model-potential approach to the electronic potential in the surface region to obtain a realistic and self-consistent description of the electronic spectrum, charge density, and potential in this region. In Sec. II we shall briefly review these methods. The implementation of any self-consistent electronic-structure calculation requires an efficient algorithm for determining the charge density from a select sample of occupied eigenstates. This topic as well as related matters are discussed in Sec. III. The electronic spectrum of the Si (100) surface is described in Sec. IV with particular emphasis on two bands of surface states that have energies in the gap between the bulk conduction and valence bands. In that section we make contact between the spectrum calculated here and the results of previous theoretical investigations, none of which have been self-consistent or have been concerned with the calculation of a surface potential or charge density. This latter topic is the subject of Sec. V.

In our final section, VI, we explore some of the implications of the present work for the origin and nature of the (2×1) reconstruction that occurs on the physical (100) surface as well as implications for the nature of surface states on stepped (111) surfaces.

II. THEORETICAL METHODS

The theoretical approach used in this paper follows that of Appelbaum and Hamann[5] (AH) and we briefly review it here for the convenience of the reader. The surface region is shown in Fig. 1 in profile. It contains the vacuum, two layers of atoms, and a bulk-surface matching plane midway between the second and third atom planes.

The potential seen by an electron in the surface region is given by

$$V(\vec{x}) = V_{es}(\vec{x}) + V_{xc}(\vec{x}) + V_{core}(\vec{x}) . \tag{2.1}$$

The electrostatic potential V_{es} is produced by the valence electrons and by the $(+4e)$ core charge. It is obtained by solving Poisson's equation

$$\nabla^2 V_{es}(\vec{x}) = -4\pi[\eta(x) + \rho_I(\vec{x})] , \tag{2.2}$$

where $\eta(x)$ is the valence charge density and where $\rho_I(\vec{x})$, the Si-ion core charge density, is parame-

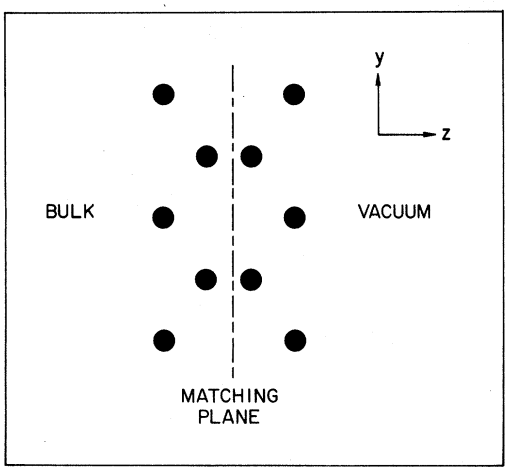

FIG. 1. Position of the bulk-surface matching plane relative to the four planes of atoms nearest the surface is shown together with the coordinate system. The four planes of atoms normal to the x axis are not all in the same x = const plane.

trized for each ion as

$$\rho_I(\vec{x}) = 4(\alpha/\pi)^{3/2} e^{-\alpha x^2} . \tag{2.3}$$

The exchange and correlation potential is treated in the local approximation using the Wigner interpolation formula[6]:

$$V_{xc}(\vec{x}) = -\left(0.984 + \frac{8.77 f(\vec{x}) + 0.9437}{[1 + 12.57 f(\vec{x})]^2}\right) f(\vec{x}) , \tag{2.4}$$

$$f(\vec{x}) = \eta^{1/3}(\vec{x}) , \tag{2.5}$$

where we have adopted atomic units ($e = \hbar = m_e = 1$).

The ion core potential V_{core} is a model potential having the form

$$V_{core}(\vec{x}) = (v_1 + v_2 |\vec{x}|^2) e^{-\alpha |\vec{x}|^2} . \tag{2.6}$$

We adopt here the same parameters used by Appelbaum and Hamann,[7] viz.,

$$\alpha = 0.61018, \quad v_1 = 3.0417, \quad v_2 = -1.372 .$$

These produce a state-of-the-art fit to the experimentally known bulk-Si energy band gaps. Equations (2.1)–(2.6) allow one to calculate $V(\vec{x})$ given a knowledge of $\eta(x)$ and the surface-atom positions. The latter have been fixed by fiat in this paper; the former is calculated via Schrödinger's equation.

The Schrödinger equation is solved in the Laue representation.[8] Expanding the potential $V(x)$ and wave function Ψ as

$$V(\vec{x}) = \sum_{\vec{G}_\parallel} V_{\vec{G}_\parallel}(z) e^{i\vec{G}_\parallel \cdot \vec{x}_\parallel} , \tag{2.7}$$

$$\Psi_{\vec{k}_\parallel}(\vec{x}) = e^{+i\vec{k}_\parallel \cdot \vec{x}_\parallel} \sum_{\vec{G}_\parallel} U_{\vec{G}_\parallel}(z) e^{i\vec{G}_\parallel \cdot \vec{x}_\parallel} , \tag{2.8}$$

where $\vec{k}_\parallel$ is the two-dimensional Bloch wave vector and $\{\vec{G}_\parallel\}$ the set of two-dimensional reciprocal-lattice vectors, Schrödinger's equation takes the form

$$\left(-\frac{1}{2}\frac{d^2}{dz^2} + V_0(z) + \frac{|\vec{k}_\parallel + \vec{G}_\parallel|^2}{2} - \epsilon_{\vec{k}_\parallel}\right) U_{\vec{G}_\parallel}(z) + \sum_{\vec{G}'_\parallel} V_{\vec{G}_\parallel - \vec{G}'_\parallel}(z)\, U_{\vec{G}'_\parallel}(z) = 0 . \tag{2.9}$$

Equation (2.9) represents a set of coupled differential equations which are integrated numerically for 30 different $U_{\vec{G}_\parallel}(z)$'s, most of which are handled by an iterative Green's-function technique. A spacing of 0.32 Å is used in the numerical integration. The Schrödinger equation is solved for a number of $\vec{k}_\parallel$'s, from which, following the discussion in Sec. III, a charge density is determined. A potential is calculated from this charge density and the process is repeated until the input and output potentials are equal to within 0.002 hartree, rms. Eight iterations were required to achieve this accuracy, with the iterative loop begun by choosing a "best guess" starting potential.

III. CHARGE-DENSITY SUMMATION SCHEME

A key step in the self-consistent calculation is the determination of the surface charge density from a sample of surface and continuum states calculated at symmetry points in the surface Brillouin zone (SBZ). As efficient a computational scheme as possible is desirable because of the complexity of the calculations being performed. The charge density $\eta(\vec{x})$ is given by

$$\eta(\vec{x}) = \sum_n \sum_{\vec{k}_\parallel} \sum_{k_z} |\Psi^{B,n}_{\vec{k}_\parallel, k_z}(\vec{x})|^2 \Theta(\epsilon_F - \epsilon^{B,n}(\vec{k}_\parallel, k_z)) + \sum_n \sum_{\vec{k}_\parallel} |\Psi^{S,n}_{\vec{k}_\parallel}(\vec{x})|^2 \Theta(\epsilon_F - \epsilon^{S,n}(\vec{k}_\parallel)) , \tag{3.1}$$

where $\Psi^{B,n}_{\vec{k}_\parallel, k_z}$ is a band state with energy $\epsilon^{B,n}(\vec{k}_\parallel, k_z)$ and wave vector $\vec{k}_\parallel$ in the SBZ, and k_z, the bulk wave vector normal to the surface, labels the incident-wave part of the various band states for a given band n and fixed $\vec{k}_\parallel$. The surface states $\Psi^{S,n}_{\vec{k}_\parallel}$ have energy $\epsilon^{S,n}_{\vec{k}_\parallel}$, the Fermi energy is ϵ_F, and n is a band index.

The k_z summations in (3.1) are performed in a completely straightforward way by a uniform sampling of k_z, typically using 10 to 16 points for each $(n, \vec{k}_\parallel)$. Convergence studies have indicated this to be quite adequate.

The $\vec{k}_\parallel$ integrations are handled in a different way. Following Baldereschi[9] we expand $\rho^n(\vec{k}_\parallel)$, the partial charge density for band n, with momentum $\vec{k}_\parallel$ in a Fourier series in real space:

$$\rho^n(\vec{k}_\parallel) = \sum_{i,j} \rho_{ij} e^{i\vec{k}_\parallel \cdot \vec{R}_{ij}} , \tag{3.2}$$

where

$$\vec{R}_{ij} = (a/\sqrt{2})(i\hat{x} + j\hat{y}), \tag{3.3}$$

and i and j are integers and a is the length of the crystallographic unit cell for bulk Si ($a = 10.265$). In Fig. 2 we have drawn both the real-space surface lattice and the reciprocal lattice. The (100) surface has C_{mm} symmetry (two perpendicular mirror planes) so that some of the ρ_{ij} are equal. Exploiting this and retaining only the first four coefficients in the Fourier expansion (3.2) of $\rho^n(\vec{k}_{\parallel})$ one obtains

$$\rho^n(\vec{k}_{\parallel}) = \rho_{00} + 2\rho_{10}\cos k_1\pi + 2\rho_{01}\cos k_2\pi + 4\rho_{11}\cos k_1\pi \cos k_2\pi , \tag{3.4}$$

where

$$\vec{k}_{\parallel} = (2\pi/a\sqrt{2})(k_1\hat{x} + k_2\hat{y}). \tag{3.5}$$

The validity of truncating (3.2) by including only the first few terms has been discussed by Phillips,[10] among others,[11,12] in the context of bulk solids. We believe (3.4) to be an excellent approximation to the charge density. We shall also use it for the surface band energy $\epsilon^{n,S}(\vec{k}_{\parallel})$, for which an expansion analogous to (3.2) is valid.

As recently observed by Baldereschi, if we are interested in a filled band of states, the sum $\sum_{\vec{k}_{\parallel}} \rho^n(\vec{k}_{\parallel})$ is rigorously proportional to ρ_{00}. If we assume the validity of (3.4)

$$\rho^n(\tfrac{1}{2}, \tfrac{1}{2}) = \rho_{00} , \tag{3.6}$$

so that the calculation of states for $\vec{k}_{\parallel} = (\frac{1}{2}, \frac{1}{2})$ would allow us to determine the surface charge density for filled bands. This summation scheme corresponds to the use of a single special point, and represents the best possible one-point integration scheme, as discussed in a systematic fashion by Cunningham.[11] There are disadvantages with this choice, however, which prompt us to opt for calculating states at the points (0, 0) and (1, 1). With

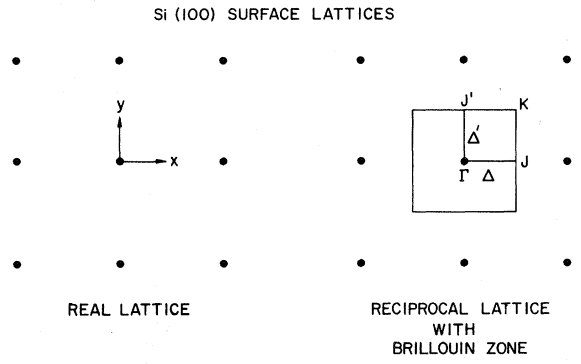

FIG. 2. Real- and reciprocal-space lattice for the Si (100) surface are shown together with the surface Brillouin zone.

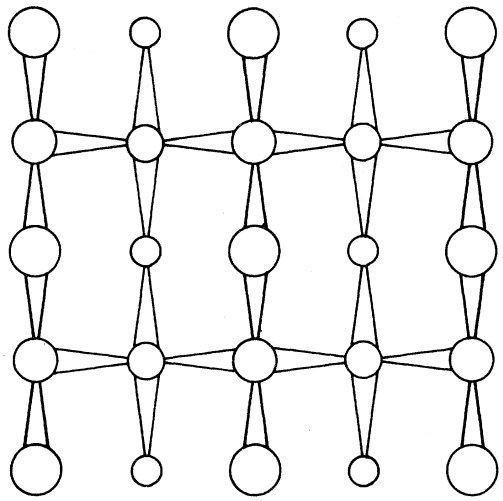

FIG. 3. Orientation of the first four planes of atoms nearest vacuum is shown when viewed normal to the surface and toward the bulk. Atoms of equal size are in the same plane. The atom size is used to indicate how far atoms are from vacuum (the largest atoms are closest to vacuum).

these states

$$\rho_{00} = \tfrac{1}{2}[\rho^n(0, 0) + \rho^n(1, 1)] + O(\rho_{11}^n) , \tag{3.7}$$

so that we once again can calculate ρ_{00}, but only by neglecting the ρ_{11} term in (3.4). We believe this to be much smaller than either ρ_{10} or ρ_{01} and we adopt this summation scheme for calculating ρ_{00} for filled bands. The smallness of ρ_{11} stems from the chain nature of atoms arranged in planes parallel to the (100) surface (see Fig. 3) and the tight-binding interpretation of the expansion (3.4). The argument proceeds as follows. To pass from one surface atom to the next nearest surface atom along either the [10] or [01] directions requires one to pass through two bonds. To pass to the next nearest surface atom, i.e., that along the [11] direction, requires one to pass through four bonds. In this respect the ρ_{11} is the same order as ρ_{20} and ρ_{02}, which also require one to pass through four bonds. The above argument ignores direct overlap between sites, but since this should fall off very rapidly with distance, we expect it will not invalidate the above conclusions.

The question remains why we have chosen a two-point scheme such as (3.7) instead of (3.6). The reason is twofold. First, the (0, 0) and (1, 1) states, denoted by Γ and K in Fig. 2, are symmetry states, and their calculation proceeds much more efficiently than the state at $(\frac{1}{2}, \frac{1}{2})$. Second, the calculation of the energy spectrum at Γ and K gives us a better picture of the total energy spectrum than does the single point $(\frac{1}{2}, \frac{1}{2})$.

For bands that intersect the Fermi surface, the

integration scheme (3.7) is clearly inadequate. What must be done there is to calculate $\rho(\vec{k}_{\parallel})$ and $\epsilon(\vec{k}_{\parallel})$ at Γ, K, and two additional points J and J' shown in Fig. 2. This gives us the full expansion, (3.4), and allows us to calculate ρ. We defer until Sec. IV, where the electronic spectroscopy of the (100) surface is presented, a discussion of how we proceed in this more complex case.

IV. SPECTROSCOPY

The formation of an ideal Si (100) surface results in the breaking of two bonds per surface atom. Such severe distortions of the bonding arrangements at the surface are expected to manifest themselves in the formation of one or more bands of surface states. We find this to indeed be the case, with two bands of surface states located in the bulk energy gap between valence and conduction states.

The energy dispersion of these two bands are plotted in Figs. 4 and 5 along principal symmetry directions in the SBZ.

These curves were generated by fitting the interpolation formula

$$\epsilon^S(\vec{k}_{\parallel}) = \epsilon_{00} + 2\epsilon_{10}\cos k_1\pi + 2\epsilon_{01}\cos k_2\pi + 4\epsilon_{11}\cos k_1\pi\cos k_2\pi \qquad (4.1)$$

at the four symmetry points Γ, J, J', and K. The values of the ϵ_{ij} so obtained are listed in Table I. It is important to point out that the lower-gap surface-state band (LGSS) ceases to exist as a *bona-fide* surface state at Γ, having merged into the valence band there. Indeed, when we examine the valence band of the appropriate symmetry at Γ, we find a strong resonance close to the top of the band; the position of this sharp resonance has been used in determining $\epsilon^S(\vec{k}_{\parallel})$. Notice, with particular reference to Fig. 4, where we have plotted the bulk valence-band energy continuum of appropriate symmetry versus $\vec{k}_{\parallel}$, that the region of the SBZ where the LGSS is a resonance is extremely narrow. Although there is no way of showing it simply in Fig. 5, the upper band is nearly mid-gap (with respect to bands of the same symmetry) throughout the SBZ.

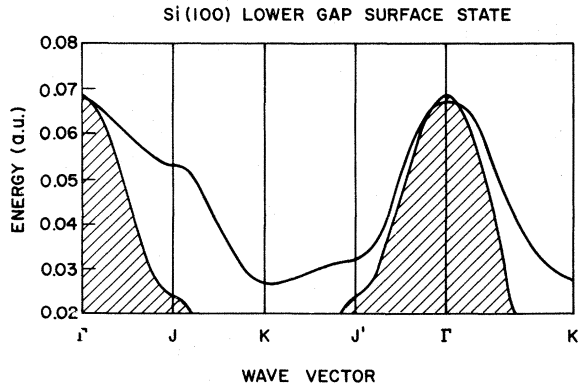

FIG. 4. Plot of the LGSS energy versus $\vec{k}_{\parallel}$ along the principal symmetry axes in the SBZ. The continuum of bulk energy levels corresponding to the top portion of the valence band are shown as striped regions.

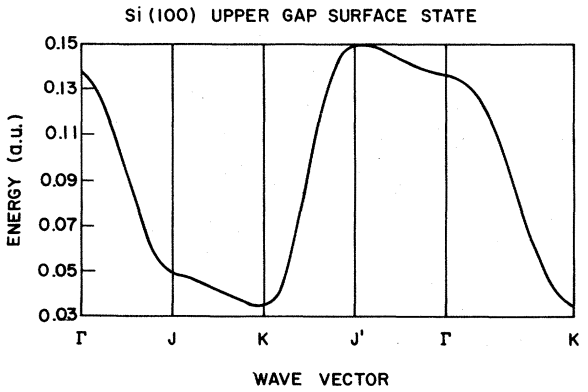

FIG. 5. Plot of the UGSS energy versus $\vec{k}_{\parallel}$ along the principal symmetry axes in the SBZ.

The dispersion of both bands of gap surface states is quasi-one-dimensional, with the upper-gap surface state (UGSS) more nearly one-dimensional. The directions of maximum dispersion for these bands are at right angles to each other.

What are the natures of these two states? To best answer that, we have plotted contours of constant charge density $\rho^S_{\vec{k}_{\parallel}}(\vec{x})$, on select planes normal to the (100) plane and passing through surface atoms. Such contour plots, with the contour plane passing through a row of broken surface bonds, are shown in Figs. 6, 7, and 8 for the LGSS at $\vec{k}_{\parallel} = J$, J', and K, respectively. This band is clearly a dangling-bond band, pointing into vacuum at right angles to the surface and having a node just below the surface atoms. It is similar to the dangling-bond surface band on Si (111) in this respect, [5] but whereas the latter surface state was directed along the broken-bond direction, the present state is not. The degree of anisotropy about the surface normal of this surface state is revealed by plotting the K surface-state charge density in Fig. 9 on a plane at right angles to that which was used in Fig. 8,

TABLE I. Values of ϵ_{ij} in the expansion (4.1) of the lower-gap (LGSS) and upper-gap (UGSS) surface-state band energy as a function of $\vec{k}_{\parallel}$.

	ϵ_{00}	$2\epsilon_{01}$	$2\epsilon_{10}$	$4\epsilon_{11}$
LGSS	0.04531	0.01562	0.00482	0.00242
UGSS	0.09299	0.000248	0.05093	−0.00698

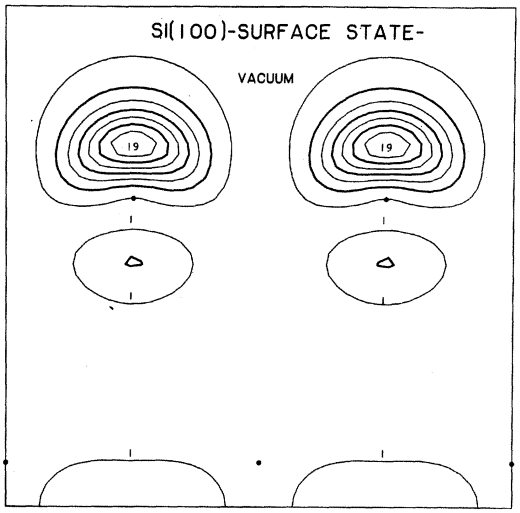

FIG. 6. Contour plot of the LGSS charge density for $\vec{k}_{\parallel}=J$ on a plane normal to the surface and passing through a row of broken bonds. In this and all subsequent charge contour plots, charge densities are in 10^{-3} a.u., and the plotting region extends from the midpoint between the 4th and 5th atom planes to the vacuum. Atom sites are indicated by dots.

i.e., a plane which passes through a row of completed bonds between the first and second planes of atoms. Notice that while the forward lobe is highly symmetric, the back lobe has considerable anisotropy.

Much of the nodal structure exhibited in these contour plots has its explanation in the form that a single orbital Bloch wave function assumes, viz, .

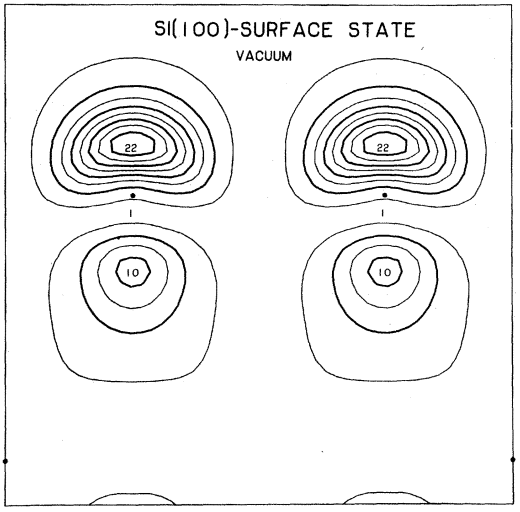

FIG. 8. Contour plot of the LGSS charge density for $\vec{k}_{\parallel}=K$ on a plane normal to the surface and passing through a row of broken bonds.

$$\Psi^{S}_{\vec{k}_{\parallel}}=\sum_{i=1}^{N} e^{i\vec{k}_{\parallel}\cdot\vec{R}_i}\,\varphi(\vec{x}-\vec{R}_i)\,. \tag{4.2}$$

Some information about $\varphi(\vec{x})$ can be obtained by fitting (3.4) to the charge density above a surface atom for the various momentum states. Assuming that only nearest-neighbor overlap contributes to $\rho_k(\vec{x})$, one finds that φ is not spherically symmetric at near-neighbor distance, that φ is 20 times larger at the atom site than at the near-neighbor distance, and that φ changes sign there as well. The latter

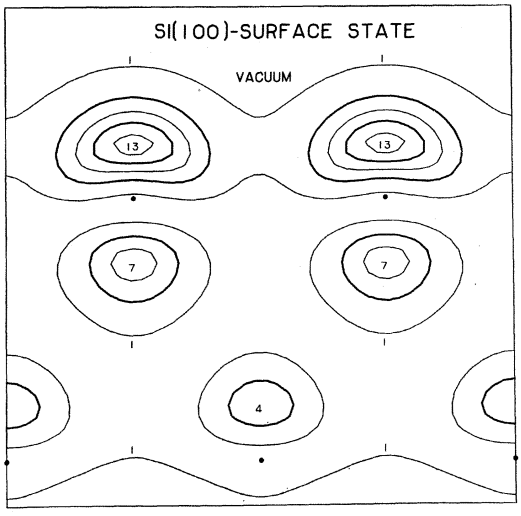

FIG. 7. Contour plot of the LGSS charge density for $\vec{k}_{\parallel}=J'$ on a plane normal to the surface and passing through a row of broken bonds.

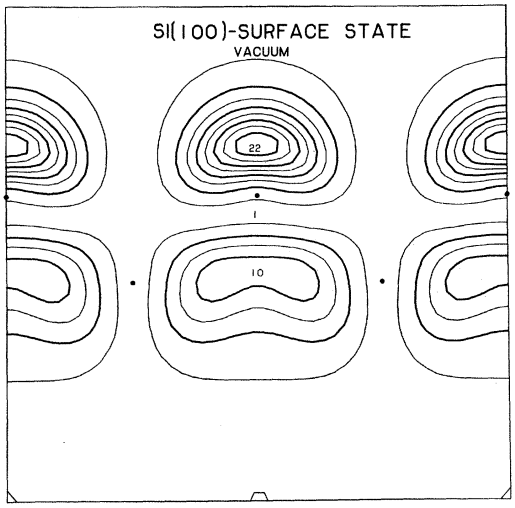

FIG. 9. Plot of the LGSS charge density for $\vec{k}_{\parallel}=K$ on a plane normal to the surface and passing through a row of complete (back) bonds between atom in the first and second planes.

property is to be expected of Wannier functions.

Consider now the UGSS. States at K, Γ, and J are plotted in Figs. 10, 11, and 12 on a contour plane which passes through a row of broken bonds. This band has a maximum charge-density accumulation that is confined symmetrically about a line passing through a row of surface atoms and in the direction of the broken bonds. Once again the nodal structure revealed in Figs. 10–12 can be understood with reference to a single localized orbital expansion of $\Psi^S_{\vec{k}_{\parallel}}$, i.e., (4.2).

Both bands of surface states have their origin in the two broken bonds that are present on the (100) surface. The broken bonds can be thought of as a linear combination of S, P_z, and P_x orbitals; viz.,

$$\alpha_s S + \alpha_z P_z - \alpha_x P_x \tag{4.3}$$

and

$$\alpha_s S + \alpha_z P_z + \alpha_x P_x \,. \tag{4.4}$$

The symmetric combination gives an $(\alpha_s S + \alpha_s P_z)$-type dangling-bond state, i.e., the lower band, while the antisymmetric combination yields a P_x-type band, the upper band.

Because of overlap among the P_x-type orbitals, the upper band is considerably wider than the lower band; so much so that the upper and lower bands overlap in energy. This, as we shall soon see, complicates the populating of these bands and the determination of a surface Fermi level in our self-consistent calculation.

Before elaborating on this subject we shall complete our discussion of the surface states found on Si(100). In addition to the gap surface states isolated pockets of surface states were found at K at

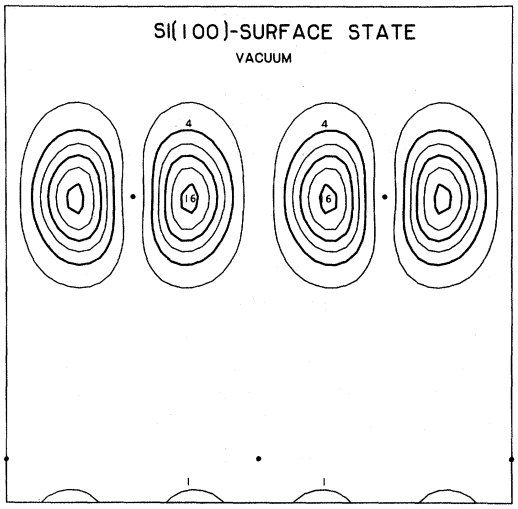

FIG. 11. Contour plot of the UGSS charge density for $\vec{k}_{\parallel}=\Gamma$ on a plane normal to the surface and passing through a row of broken bonds.

energies -0.2424 and -0.2197. These states were weakly split from bulk bands that are themselves extremely narrow. The charge density in these surface states is essentially identical to that contributed by the bulk band. The phenomenom we observe here results from the fact that the periodic parts of the wave functions for the bulk energy bands at $\vec{k}_{\parallel}=K$ are highly localized. A small shift in the effective potential energy in the vicinity of one of these localized orbitals takes it out of the band from which it is derived, creating in a formal sense a surface state, but causing almost no change

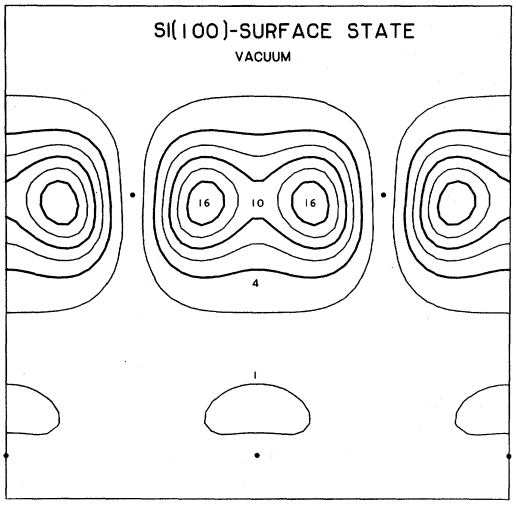

FIG. 10. Contour plot of the UGSS charge density for $\vec{k}_{\parallel}=K$ on a plane normal to the surface and passing through a row of broken bonds.

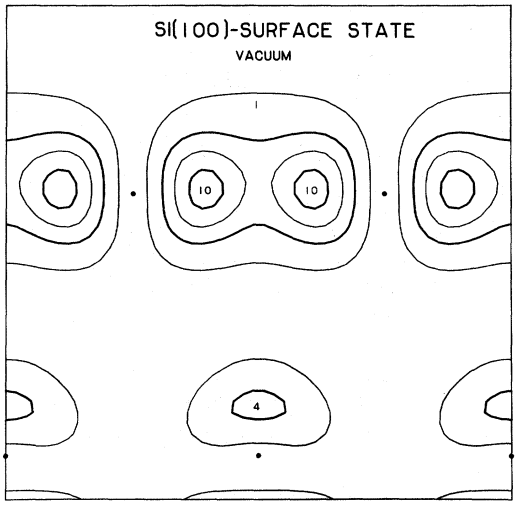

FIG. 12. Contour plot of the UGSS charge density for $\vec{k}_{\parallel}=J$ on a plane normal to the surface and passing through a row of broken bonds.

in the local spatial nature of the charge density.

Such states are also present on the Si(111) surface[13] and one can expect that their presence on almost all semiconductor surfaces is assured by the rather remarkable abundance of lines in $\vec{k}$ space along which $\epsilon(\vec{k})$ has almost no variation. It is highly unlikely that these states will have any significant physical influence on surface electronic properties.

We turn now to the difficult question of populating the gap or broken-bond surface states. If the UGSS were not so wide, we could ignore it entirely in filling our states with electrons. The lower-gap surface state would then contain two electrons per surface unit cell, and the surface could be described as semiconducting. This is to be contrasted with the situation on the ideal Si (111) surface, where a single dangling-bond band exists, containing only 1 electron per surface unit cell. In the calculations reported by Appelbaum and Hamann on Si (111), occupancies for the dangling-bond state were quoted as 0.6 to 0.7 electrons, in violation of the at that time unappreciated and unproved charge-quantization sum rule.[14] Had the states been occupied to 1 full electron, charge neutrality violation ~ 0.3 electron would have occurred in the surface region. Subsequent refinements in the procedures for normalizing surface states, replacing an approximate scheme by an exact analytic one, and improvements in the numerical solutions of Schrödinger's equation, both incorporated in the present calculation, account for most of this discrepancy. The remainder is due to the placement of the matching plane; had that plane been placed further into the bulk, charge neutrality would be more nearly satisfied. (See Sec. V.) Subsequent work[15] on Si (111) with the above refinements indicate little change in the quantitative features of the spectrum, and no qualitative changes.

The presence of the UGSS changes the character of the surface from semiconducting to semimetallic. Charge flows out of the LGSS into the UGSS. At the same time, some charge flows out of the continuum states near Γ into the UGSS. The actual amount of such charge from continuum states in the surface region is very small, since the density of states in the surface-state band is much higher than in the continuum states. In addition, the charge in the continuum states near a band extremum is greatly reduced by the surface barrier. (This statement is weakened somewhat if there exists a resonance in the continuum near that extremum but since we would treat that resonance as if it were part of the surface-state band, the spirit of the argument here is unchanged.)

Depopulating continuum states, unlike surface states, can produce a charge disturbance deep into the bulk and long-range effects such as band bending. Our calculational scheme is not equipped to handle such effects nor are we interested in them per se. To avoid having to deal with them, we can postulate that the crystal has been doped p type, so that the small amount of charge missing from the bulk continuum bands is exactly neutralized by the dopant.

If we ignore the continuum charge, the problem of determining a surface Fermi level involves ensuring that (i), the phase space for the hole states created in the LGSS is the same as that for the electron states in the UGSS, and (ii), the electron and a hole Fermi levels are equal. The quasi-one-dimensional nature of the surface states makes this extremely simple. The flow of electrons from the lower to the upper surface bands depopulates a strip of LGSS states parallel to the Δ direction (see Fig. 2), whose width $2k_1 = 0.56$, and populates a strip of UGSS states parallel to the Δ' direction of equal width.

With the extent of populated and empty states determined, the problem of sampling them in order to calculate a charge density remains. We proceed in two steps: First, we ignore the charge flow between the two broken-bond surface states. In that case the summation scheme using Γ and K to determine ρ_{00} presented in Sec. III is adequate. To correct $\rho(\vec{x})$ for the charge transfer we add to $\rho(\vec{x})$ the approximately $\frac{1}{2}$ an electron in the sheet about Δ', using an average of Γ and J' states as representative of the charge in that sheet, and subtract an equal number of electrons from the LGSS.

One would like to use an average of the Γ and J states for this subtraction, but because of the resonant nature of the Γ state, it was felt that using only the J states was both simpler and avoided the problem that Γ may be unrepresentative of a major portion of the hole sheet.

Before leaving the topic of surface spectroscopy we shall compare the results of this section with previous theoretical calculations on the Si (100) surface. No self-consistent calculations exist for this surface but a number of tight-binding calculations and a pseudopotential calculation using an abrupt-potential-step model have been done. The latter calculation, due to Jones,[16] finds two bands of gap surface states. Both bands are wider than we find, with the LGSS band over 2 eV wide. The one-dimensional dispersion found in our calculation is not present in Jones's. We believe this is primarily caused by his inadequate solution of Schrödinger's equation.

The earliest tight-binding calculation of Si (100) is due to Hirabayashi.[17] The spectrum found in Hirabayashi's calculation, which employed a nearest-neighbor tight-binding formalism, consisted of two bands, the upper band less than ~ 0.1 eV wide and split from the lower band by ~ 1 eV. This

TABLE II. Comparison of the LGSS and UGSS energies at $\vec{k}_{\parallel}=\Gamma$, K, and J' obtained in this calculation (ABH) and by Pandey and Phillips (PP). The dagger indicates a resonance energy. The energies are quoted in hartrees, on a scale where Γ_{25}', the top of the valence band, is 0.06884, and in eV, where the energies are quoted relative to Γ_{25}'.

	Γ		K		J'		
	(hartree)	(eV)[a]	(hartree)	(eV)[a]	(hartree)	(eV)[a]	
LGSS	0.0861	0.47	0.0383	−0.83	0.0453	−0.64	PP
	0.0682	−0.02†	0.0273	−1.13	0.0321	−1.00	ABH
UGSS	0.1390	1.91	0.0446	−0.66	0.1259	1.54	PP
	0.1372	1.86	0.0348	−0.93	0.150	2.23	ABH

[a]The energies quoted in eV are referenced to the valence-band maximum.

lower band was ~0.3 eV in width. There is essentially total disagreement between this calculation and ours.

This extreme narrowness in the surface-state bands found by Hirabayashi is due to his use of only nearest-neighbor interactions in his tight-binding scheme. A similar situation occurred on the Si (111) surface, where Hirabayashi found a very narrow surface state that was significantly broadened when he introduced next nearest neighbors.

Very recently, empirical tight-binding methods have been used by Pandey and Phillips[18] to fit the Si (111) calculations of AH.[5] They have also calculated the surface state spectrum of the Si (100) surface with a 0.45-Å relaxation inward of the outer surface plane simulated by a Hückel-type adjustment of the surface-atom parameters. While this is not the structure calculated here, this calculation and the one reported here are in reasonable agreement. They find two bands of gap surface states, one 0.048 hartree wide, and one 0.095 hartree. A comparison is made in Table II between the calculated values at Γ, J', and K. For some reason they did not calculate J. (Note: they referred to our J' point as J.) We consider this to be quite reasonable agreement although it must be emphasized that the two calculations are for different geometries.

V. SELF-CONSISTENT CHARGE DENSITY AND WORK FUNCTION

The self-consistent potential, whose spectrum was discussed in Sec. IV, is shown in Figs. 13 and 14. As with the partial charge densities, we have plotted the contours of constant potential on two planes normal to the (100) surface and passing in one case (Fig. 13) through a row of broken bonds and in the other (Fig. 14) through a row of completed bonds between the first and second atom planes.

The potential is shown only on the vacuum side of the matching plane, where it is self-consistent to 0.002 hartree, rms. Noticethe spherical nature of the potential above the surface atoms and the strong tetrahedrally directed backbond potential between the first and second layers. The ionization potential is 5.9 eV, higher than is the case for the actual (100) (2×1) surface.[19] It is our present belief that this discrepancy has its origin in the surface reconstruction on the (100) surface, which, as we will discuss in Sec. VI, should lower the surface dipole moment considerably.

The total surface charge density, calculated from the surface spectrum as discussed in Secs. III and IV, is plotted on planes identical to those used for the potential in Figs. 15 and 16. Notice the total breakdown of tetrahedral bonding in the plane of the broken bonds. The charge in that plane is cen-

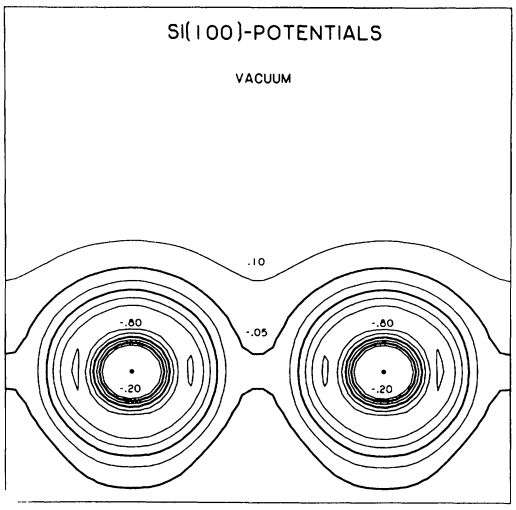

FIG. 13. Contour plot of the surface potential on a plane normal to the surface passing through a row of broken bonds. Only the region outside the matching plane is shown here. The potential is in atomic units, and the placement of the scale is such that the valence-band maximum falls at +0.069. Atom sites are indicated by dots.

tered about the atoms, with charge lobes above and below the surface atoms which reflect the almost complete occupation that we find of the dangling-bond surface-state band.

The charge in the bonds between the first and second planes, shown in Fig. 16, has returned to its bulk value, indicating how rapid the screening is of the disturbance set up by the surface potential barrier. In spite of the rapid screening, complete charge neutrality is not achieved to the right (vacuum side) of the matching plane. There is approximately 0.2 of an excess electron in this region. If we study the charge disturbance behind the matching plane, there is a deficit of charge. Between the matching plane and a plane a distance $a/4$ behind the matching plane, that deficit is ~0.07 electrons; going back a distance $a/2$, that deficit increases to ~0.11. Assuming an exponential falloff in the charge disturbance,[20] one finds much of the excess to the right of the matching planes is canceled by a deficit to the left, with ~0.045 electrons in excess. We consider this an excellent indication of the numerical accuracy of our Schrödinger-equation solution.

The presence of an excess charge in the surface region would lead to a finite field in vacuum if nothing were done to compensate for it. There are a number of procedures one could use. First a field could be introduced at the matching plane equal and opposite to the excess. Another approach, one we adopt here, is to scale the surface charge density to achieve neutrality.

We believe that achieving a zero-field vacuum condition by introducing charge into the surface region rather than through a field at the matching plane corresponds more nearly to what the crystal in fact would do. The deficit charge we now find behind the matching plane would be screened by the crystal if we allowed the potential in that region to adjust. This in turn would push charge into the surface region. Studies by Appelbaum and Hamann[21] have shown that for the case of H adsorbed on Si(111), the above procedure yields results highly insensitive to the position of the matching plane.

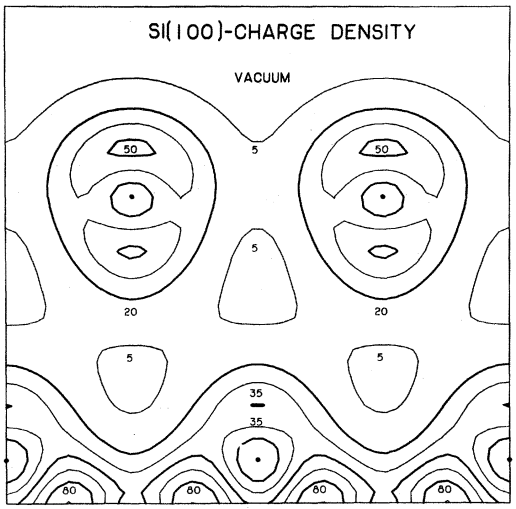

FIG. 15. Contours of total charge density are plotted on a plane normal to the surface and passing through a row of broken bonds.

VI. RECONSTRUCTION AND STEPPED SURFACE

As already emphasized, the actual Si (100) surface is reconstructed into a 2×1 structure. The

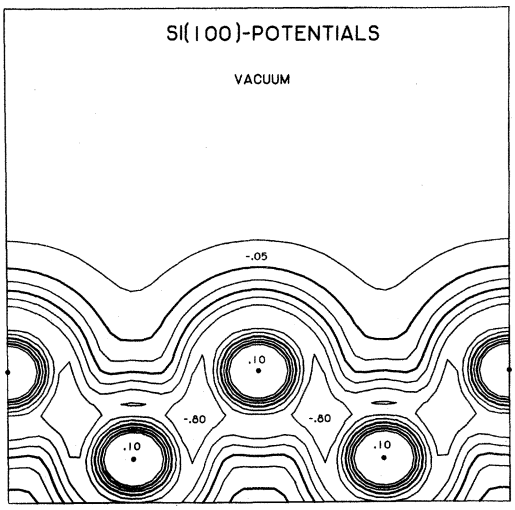

FIG. 14. Contour plot of the surface potential on a plane normal to the surface passing through a row of bonds between atoms in the first and second planes nearest vacuum. Only the region outside the matching plane is shown here.

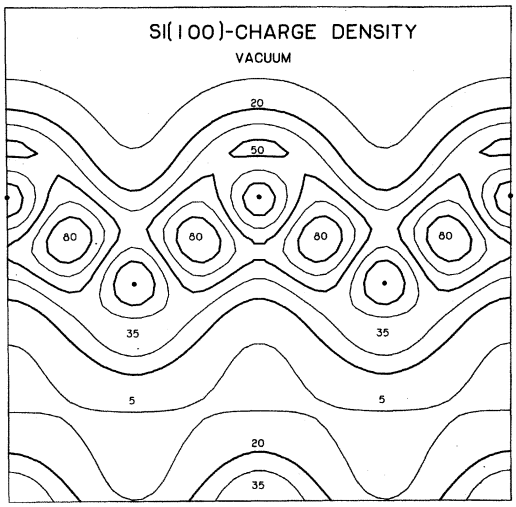

FIG. 16. Contours of total charge density are plotted on a plane normal to the surface and passing through a row of back bonds between the first and second atom planes.

nature of this reconstruction is at present a matter of considerable speculation. Two very different hypotheses have been advanced, each having a number of elaborations. The first hypothesis, suggested by Schlier and Farmsworth,[3] is a pairing of rows of surface atoms in a chemically familiar dimerization process. The second hypothesis originally due to Lander suggests that the structure has 50% vacancies. We presently favor the former hypothesis, where a simple pairing reconstruction occurs. As emphasized by Levine,[22] the pairing of atoms on the (100) surface can be achieved with no bond-length distortions and modest bond-angle modifications. For example, the most severely distorted angle from ideal tetrahedral is that between the pair bond parallel to the surface and the backbonds, and that is ~ 93°.

Such a dimerization is easily rationalized in terms of the spectrum we find for the ideal (100) surface. The broad UGSS is in a unique position to supply the energy that drives the transition. With pairing, the surface band splits into two bands, one having states whose density is on the pair bond. This band, either completely or partially occupied, will have a considerably lower energy than the band it derived from. The dangling-bond band also splits, but not nearly so much, and one expects considerable charge transfer from this band of states to the now lowered UGSS. This has the effect of moving electronic charge from a region considerably above the surface atoms to a region in the plane of the surface atoms, and should lower the calculated ionization potential. It has been suggested by Ibach and Rowe[23] that because the adsorption of H on the Si (100) surface in monolayer coverages does not reorder the surface, the vacancy model for this surface is more probable.

We see, however, from our discussion above, that the band of states that participates in the pairing transformation is the UGSS and not the dangling-bond states. Consequently, one can have H absorbing by means of dangling-bond orbitals without disturbing the pairing that occurs through the upper-gap surface-state band. A more complete discussion of the 2×1 reconstruction will be postponed to the third paper in this series.

A final topic we wish to briefly consider has to do with stepped (111) surfaces. Stepped surfaces are produced as a consequence of imperfect cleaves—intentional or otherwise. On the (111) surface a two-atom step exposes a row of atoms identical to the rows of atoms on the (100) surface. Some of the spectroscopic information obtained for the (100) surface should be applicable to a discussion of step-induced surface states.

We expect there to be two bands of one-dimensional surface states. One band should be similar to the UGSS, which itself is quasi-one-dimensional, and has its direction of dispersion along the edge direction. This band should be quite wide. The second band is analogous to the LGSS. Since that band disperses only about 0.3 eV (0.012 hartree) along the row of broken bonds, the second step band should be extremely narrow; so narrow, in fact, that the band picture may be inappropriate. One would then have isolated levels which would be occupied by an electron. It is this electron, we believe, that could be responsible for the EPR signal observed on the Si (111) surface.[24] It has been observed that a strong correlation exists between the EPR signal and the step density on the Si (111).[25]

[1] An excellent review of the experimental literature is contained in W. Mönch, Adv. Solid State Phys. 13, 241 (1973).

[2] J. E. Rowe and H. Ibach, Phys. Rev. Lett. 32, 421 (1974).

[3] R. E. Schlier and H. E. Farnsworth, J. Chem. Phys. 30, 917 (1959). J. Lander, in *Progress in Solid State Chemistry*, edited by H. Reiss (Pergamon, New York, 1965).

[4] M. Henzler and J. Clabes, in *Surface Physics of Crystalline Materials*, edited by J. M. Blakeley (Academic, New York, to be published).

[5] J. A. Appelbaum and D. R. Hamann, Phys. Rev. B 6, 2166 (1972); Phys. Rev. Lett. 31, 106 (1973); 32, 225 (1974).

[6] E. P. Wigner, Phys. Rev. 46, 1002 (1934).

[7] J. A. Appelbaum and D. R. Hamann, Phys. Rev. B 8, 1777 (1973).

[8] M. Von Laue, Phys. Rev. 37, 53 (1931).

[9] A. Baldereschi, Phys. Rev. B 7, 5212 (1973).

[10] J. C. Phillips, Comments Solid State Phys. 5, 115 (1973).

[11] S. L. Cunningham, Phys. Rev. B 10, 4988 (1974).

[12] D. J. Chadi and M. L. Cohen, Phys. Rev. B 7, 692 (1973); 8, 5747 (1973).

[13] J. A. Appelbaum and D. R. Hamann, in *Proceedings of the Twelfth International Conference on the Physics of Semiconductors*, edited by M. H. Pilkuhn (Teubner, Stuttgart, 1974), p. 675.

[14] H. Heine, Phys. Rev. 145, 593 (1966); L. Kleinman, Phys. Rev. B 11, 858 (1955); J. A. Appelbaum and D. R. Hamann, Phys. Rev. B 10, 4973 (1974).

[15] J. A. Appelbaum and D. R. Hamann (unpublished).

[16] R. O. Jones, J. Phys. C 5, 1615 (1972).

[17] K. Hirabayashi, J. Phys. Soc. Jap. 27, 1475 (1969).

[18] K. C. Pandey and J. C. Phillips, Phys. Rev. Lett. 32, 1433 (1974).

[19] J. C. Riviere, in *Solid State Surface Science*, edited by M. Green (Dekker, New York, 1970).

[20] J. A. Appelbaum and D. R. Hamann, Phys. Rev. B 10, 4973 (1974).

[21] J. A. Appelbaum and D. R. Hamann, Phys. Rev. Lett. 34, 806 (1975).

[22] J. Levine, Surf. Sci. 34, 90 (1973).

[23] H. Ibach and J. E. Rowe, Surf. Sci. 43, 481 (1974).

[24] M. F. Chung and D. Haneman, J. Appl. Phys. 37, 1879 (1966).

[25] M. Henzler, as reported in Ref. 1.