Scattering-theoretical method for semiconductor surfaces: Self-consistent formulation and application to Si(001)-(2×1)

P. Krüger and J. Pollmann
Institut für Theoretische Physik II, Universität Münster, D-4400 Münster, West Germany
(Received 28 March 1988; revised manuscript received 27 June 1988)

A self-consistent-field method for the calculation of electronic properties of semi-infinite crystals with reconstructed surfaces is described in detail and applied to the Si(001)-(2×1) surface. The method is based on local-density-functional theory and a Green-function scattering-theoretic formulation is employed. Wave functions and operators are represented in a localized-Gaussian-orbital basis set. The calculations yield the self-consistent surface potential, charge densities, surface band structure, and wave-vector-, atom-, and orbital-resolved layer densities of states with an extreme spectral resolution. Surface bound states and surface resonances are determined unambiguously and accurately even for states whose wave functions are very extended. In order to be able to point out advantages of our method by comparison with the results of other techniques, we have carried out self-consistent slab calculations with varying slab thicknesses as well. The present method is shown to be very efficient and accurate in describing the whole electronic spectrum of the surface. The efficiency of the method stems largely from the fact that it exploits both the full three-dimensional periodicity of the underlying bulk crystal and the short range of the deviation of the surface potential from the bulk or vacuum potentials, respectively. Thus all bulk properties are built in from the start via a band-structure calculation as a well-defined reference and they are preserved. One then focuses on the changes produced by the surface potential. Since the bulk and surface effects are separated analytically in the Dyson equation for the surface Green function, the interpretation of the results is straightforward and unambiguous. The virtues of the scattering-theoretical method are exemplified by a detailed discussion of our results for the technologically most important Si(001)-(2×1) surface in comparison with our own slab calculations and with other results from the literature.

I. INTRODUCTION

The field of electronic-structure theory of solid surfaces has matured within the last decade. A large variety of methods has been developed and successfully applied to study surfaces of metals and semiconductors. Many of these methods are based on model geometries, like slabs or superlattices, in order to artificially reduce the infinitely long unit cell of a surface system to a size which is tractable by standard bulk band-structure techniques. In *bulk physics*, "state of the art" calculations of electronic properties within the local-density approximation[1–3] (LDA), starting from "first principles," have been performed for metals[4] and semiconductors.[5] By employing the translational symmetry of the crystal, self-consistent computations of wave functions and of the energy spectrum is reduced to the solution of the Kohn-Sham equation within the relatively small bulk unit cell. The atomic positions entering these calculations according to the Born-Oppenheimer theorem as parameters, are well known, e.g., from x-ray structure determinations. On the contrary, in *surface physics* we have to deal with broken translational symmetry perpendicular to the surface and, in consequence, with infinitely long unit cells. In addition, the atomic configuration near the surface of a solid is much less precisely known than the bulk configuration. This lack of knowledge of the exact *local* atomic *geometry* in the selvege is a major physical problem. It is currently attacked by many surface scientists using a large variety of theoretical and experimental techniques.[6] The *global geometry* of a surface system with its broken translational invariance perpendicular to the surface constitutes the basic mathematical problem. In their pioneering work on self-consistent electronic-structure theory for semiconductor surfaces, Appelbaum, Baraff, and Hamann[7,8] addressed this problem already more than a decade ago using semi-infinite geometries in their wave-function-matching technique. They evaluated the surface potential, the charge density, the ionization potential, and the dispersion of bound surface states self-consistently. Resonances and layer densities of states have been rarely reported by these authors. Nowadays, in most common slab[9] or supercell[10] techniques a crystal with its surfaces is modeled by either a finite slab or by an infinite periodic repetition of a thin slab and a few vacuum layers, respectively. These techniques allow for a reasonable description of localized surface states, but resonant electronic features whose wave functions can be very extended are less accurately treated. In view of the fact that surface resonances and bulk continuum states are the main part of the electronic spectrum, this situation is unsatisfactory and calls for improvements. In particular, layer densities of states (LDOS), which are a useful tool to analyze and interpret experimental data of angle-resolved photoelectron spectroscopy or of scanning tunneling spectroscopy, are not available with high spectral resolution. To overcome these shortcomings, we have addressed the treat-

ment of truly semi-infinite crystals by a novel approach.

In this paper we present in detail a new and efficient computational scheme for the self-consistent solution of the Kohn-Sham equations describing the electronic structure of semi-infinite, reconstructed semiconductor surfaces. Very brief accounts of the method, as it applies to ideal[11] or reconstructed[12] surfaces, have been reported previously. In our self-consistent approach the surface is treated as a two-dimensionally periodic perturbation which is extremely localized perpendicular to the surface. The surface-electronic structure is determined by scattering of bulk states at the perturbation potential. The advantages of using scattering theory for solving one-particle Schrödinger equations describing localized perturbations of bulk solids were pointed out for the case of *point defects* by Koster and Slater[13,14] 30 years ago. In the late 1970s this technique was developed to a practicable scheme for self-consistent bulk defect calculations.[15–18] Also, in surface physics many electronic-structure studies have been carried out to date on the basis of the original Koster-Slater idea using *empirical tight-binding* descriptions,[19–21] however. The extension of the scattering-theoretic method (STM) for surfaces into a practicable *self-consistent* scheme is described in detail in this paper, whereby both formal and computational aspects are discussed.

We start from a predetermined surface-structure model and present the solution of the Kohn-Sham equations for a semi-infinite crystal. The results of our surface-structure determinations for the (2×1)-reconstructed surfaces of Si(001) and Ge(001), using a semiempirical total-energy ansatz, have been reported previously.[22] The structure optimization in the framework of the method presented in this paper by self-consistent total-energy minimization calculations is briefly addressed in Sec. V.

Our approach makes use of a general concept formulated by Williams, Feibelman, and Lang.[23] The original problem of solving a differential equation is mapped onto the solution of a matrix equation of infinite dimension by expanding the wave functions in terms of a linear combination of atomic orbitals (LCAO's). The determination of the spectrum of this "Hamiltonian matrix" can be achieved by direct comparison with an appropriate reference matrix whose eigenvalues are known. Using the inverse or Green's matrix of the reference matrix leads to matrix equations describing the original system which can be solved numerically.

To exemplify advantages of our technique and to demonstrate its efficiency, we have calculated the surface-electronic structure of Si(001)-(2×1) using both the self-consistent STM and the slab method together with a LCAO basis set. We choose this technologically important prototype surface for our discussions, since it has been studied in detail both experimentally and theoretically over the years. In particular, a detailed account of the surface-electronic structure of Si(001)-(2×1), as it results from plane-wave supercell calculations, has been reported previously by Ihm *et al.*[24] To make our comparisons as meaningful as possible, we have used in all our calculations the same bulk potential as in Ref. 24 and precisely the same surface-structure model, namely the asymmetric dimer model as suggested by Chadi.[25] This choice allows for meaningful comparisons of the results on equal footing and for a critical discussion of methodological differences between STM calculations and slab or supercell calculations. The use of the STM together with norm-conserving pseudopotentials is briefly addressed in Sec. V.

A similar Green-function technique has been used by Feibelman[26] to calculate the electronic properties of a single adatom on a two-dimensional Al(001) monolayer. In addition, Feibelman[26] has applied the technique to study rebonding effects in separation and surface-diffusion barrier energies of Al adatom pairs on Al(100). Benesh and Ingelsfield[27] have worked out a complementary scheme, the embedding approach, to describe point defects and surfaces self-consistently. The interrelationship between the embedding method and the scattering-theoretic method has been discussed by Baraff and Schlüter.[28]

The self-consistent (SC) STM has a number of virtues. Within an effective one-particle picture, it allows one (a) to describe the global geometry of a surface system with correct boundary conditions, (b) to determine bound states and surface resonances with high spectral resolution, and (c) to clearly separate bulk and surface contributions in the electronic spectrum very efficiently.

It should be noted at this point that our calculations suffer from the same shortcomings that are well known from bulk electronic-structure calculations within local-density theory. These problems are currently under discussion.[29–31] New approaches which go beyond the local-density approximation have been worked out and applied to the bulk electronic-structure problem.[32,33] Even a surface system has been studied taking self-energy corrections into account.[34] The latter study was based on the conventional plane-wave supercell approach. Of course, these new techniques also must deal with the broken translational symmetry of a surface system. In particular, they make use of Green functions to describe the self-energy corrections as well. They can thus benefit twice from the virtue of the STM which we describe in this paper.

This paper is organized as follows. In Sec. II we develop the general theory for the self-consistent treatment of reconstructed surfaces by the Green's-matrix approach. The computation approach is outlined in Sec. III. Section IV is devoted to the application of the method to the Si(001)-(2×1) surface and compares the outcome of the calculations for slab, supercell, or semi-infinite geometries in detail. In the concluding discussion of Sec. V we first give a brief summary and then address, in an outlook, the potential of the present method for surface-structure optimizations and its use in the context of norm-conserving pseudopotentials and self-energy-correction calculations.

II. THEORY

A. Basic concepts

The calculation of electronic ground-state properties requires, within local-density theory, the solution of the Kohn-Sham equation,[2]

$$\hat{H}(\mathbf{r})\psi_{s,\mathbf{q}}(\mathbf{r})=[-\nabla_{\mathbf{r}}^2+V_{\text{eff}}(\mathbf{r})]\psi_{s,\mathbf{q}}(\mathbf{r}) =E_s(\mathbf{q})\psi_{s,\mathbf{q}}(\mathbf{r}) \ . \tag{1}$$

The effective potential consists of the lattice potential, the Coulomb potential (Hartree term), and the exchange-correlation potential, i.e.,

$$V_{\text{eff}}(\mathbf{r})=V(\mathbf{r})+V_c(\mathbf{r})+V_{\text{xc}}[\mathbf{n}(\mathbf{r})] \ . \tag{2}$$

and it has to be calculated self-consistently from the charge density.

The semi-infinite system considered is translationally invariant *parallel* to the surface, so that

$$V_{\text{eff}}(\mathbf{r}+\boldsymbol{\rho}_j)=V_{\text{eff}}(\mathbf{r}) \ , \tag{3}$$

where $\boldsymbol{\rho}_j$ is a lattice vector of the two-dimensional surface Bravais lattice. Due to this Bloch symmetry, the solutions of Eq. (1) can be classified by a two-dimensional wave vector $\mathbf{q}$ and a surface band index s. Energies and lengths are measured in rydbergs and Bohr radii, respectively.

The lattice potential $V(\mathbf{r})$ is given by the linear combination of atomic (or in the pseudopotential approach of ionic) potentials $v(\mathbf{r})$as

$$V(\mathbf{r})=\sum_{j,m,\mu} v_{m\mu}(\mathbf{r}-\boldsymbol{\rho}_j-\boldsymbol{\lambda}_{m\mu}) \ , \tag{4}$$

where $\boldsymbol{\lambda}_{m\mu}$ is the position of the μth atom (ion) in a layer unit cell on the mth layer of the semi-infinite crystal.

In Fig. 1 we show a schematic plot of the effective potential of a surface system averaged over surface-parallel planes as a function of the surface-perpendicular direction, i.e., $\bar{V}_{\text{eff}}(z)$. The figure clearly reveals that there is a drastic change from the bulklike crystal potential on the left to the constant vacuum value on the right within the transitions region. For the following discussions it is of essential importance to note that the transition region in Fig. 1 extends only over a few atomic layers due to an efficient screening of the surface effects on the self-consistent potential in a neutral crystal.

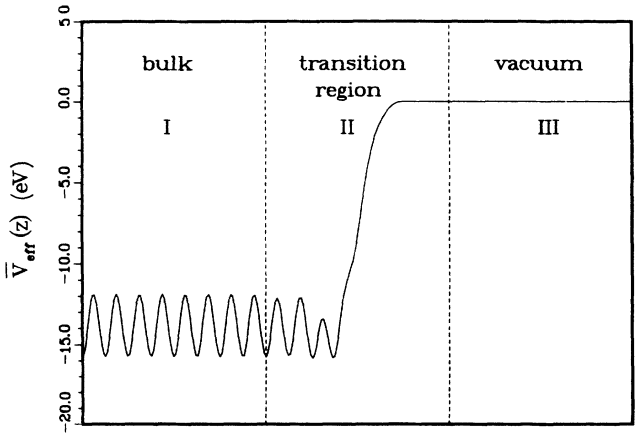


FIG. 1. Schematic plot of the x-y plane-averaged effective potential $\bar{V}_{\text{eff}}(z)$ near the surface of a semi-infinite crystal. The transition region II from the bulklike potential to vacuum is extremely narrow.

Solving Eq. (1) by standard bulk band-structure techniques is impossible due to the infinite extension of the unit cell perpendicular to the surface. If we make use, however, of the small spatial extension of the transition region in $V_{\text{eff}}(\mathbf{r})$ between the bulk and the vacuum potential, the problem can be solved. We formally treat the surface as a localized perturbation of a bulk crystal. The perturbation potential is two-dimensionally periodic parallel to the surface and highly localized perpendicular to it. To treat this perturbation efficiently, we represent the wave functions $\psi_{s,\mathbf{q}}(\mathbf{r})$ by linear combinations of atomic orbitals. These orbitals $\varphi_\alpha(\mathbf{r}-\boldsymbol{\rho}_j-\boldsymbol{\lambda}_{m\mu})$ are localized at all atomic sites $\{\boldsymbol{\rho}_j+\boldsymbol{\lambda}_{m\mu}\}$ of the whole semi-infinite crystal. We use atom-centered Gaussians whose symmetry type is labeled α. Employing the two-dimensional Bloch symmetry, so-called layer orbitals

$$\chi_{\alpha m\mu}(\mathbf{q},\mathbf{r})=\frac{1}{(N_2)^{1/2}}\sum_j e^{i\mathbf{q}\cdot(\boldsymbol{\rho}_j+\boldsymbol{\lambda}_{m\mu})}\varphi_\alpha(\mathbf{r}-\boldsymbol{\rho}_j-\boldsymbol{\lambda}_{m\mu}) \tag{5}$$

are constructed and form the symmetry-adapted LCAO basis set for the representation of the solutions of Eq. (1), i.e.,

$$\psi_{s,\mathbf{q}}(\mathbf{r})=\sum_{\alpha,m,\mu} C^s_{\alpha m\mu}(\mathbf{q})\chi_{\alpha m\mu}(\mathbf{q},\mathbf{r}) \ . \tag{6}$$

In Eq. (5), N_2 is the number of two-dimensional unit cells in the normalization area. For a simplification of the equations, we use the *superindex* $l\hat{=}\alpha,m,\mu$ wherever possible without loss of clarity. Inserting representation (6) into Eq. (1) maps the Kohn-Sham equation onto a matrix equation for the expansion coefficients,

$$\sum_{l'}[H_{l,l'}(\mathbf{q})-E_s(\mathbf{q})S_{l,l'}(\mathbf{q})]C^s_{l'}(\mathbf{q})=0 \ , \tag{7a}$$

or in a more compact matrix notation,

$$[\underline{H}(\mathbf{q})-E_s(\mathbf{q})\underline{S}(\mathbf{q})]\mathbf{C}^s(\mathbf{q})=\mathbf{0} \ . \tag{7b}$$

The matrix elements of the Hamiltonian and overlap matrix are given by

$$H_{l,l'}(\mathbf{q})=\int \chi^*_l(\mathbf{q},\mathbf{r})\hat{H}(\mathbf{r})\chi_{l'}(\mathbf{q},\mathbf{r})d^3r \tag{8a}$$

and

$$S_{l,l'}(\mathbf{q})=\int \chi^*_l(\mathbf{q},\mathbf{r})\chi_{l'}(\mathbf{q},\mathbf{r})d^3r \ . \tag{8b}$$

For a complete set of basis functions, the transformation from Eq. (1) into Eq. (7) is mathematically exact. In a practical calculation, however, we are, of course, limited to a finite number of orbitals at each atom. Yet, for a *given* potential $V_{\text{eff}}(\mathbf{r})$, the eigenvalues of Eq. (7) calculated using a finite basis set are upper limits to the exact eigenvalues of Eq. (1) according to the theorem of Hylleraas and Undheim.[35]

Due to the localization of the Gaussian orbitals, the matrix elements

$$\int \varphi^*_\alpha(\mathbf{r}-\boldsymbol{\lambda}_{m\mu})\hat{H}(\mathbf{r})\varphi_{\alpha'}(\mathbf{r}-\boldsymbol{\rho}_j-\boldsymbol{\lambda}_{m'\mu'})d^3r$$

which occur in Eq. (8a) decrease rapidly with interatomic distance,

$$d = | \lambda_{m\mu} - \rho_j - \lambda_{m'\mu'} | \ . \tag{9}$$

Within a given accuracy of the calculation, the contribution of the integrals in Eqs. (8a) and (8b) can be neglected if the distance d becomes greater than a chosen critical value d_n. This restriction to nth-nearest-neighbor integrals, together with the use of a finite number of basis functions at each atom, is a well-established approximation in bulk band-structure calculations using LCAO's.[36] This approximation leads to quite accurate results, as was confirmed by comparisons with highly converged plane-wave calculations in Ref. 37.

In spite of these approximations, the matrices in Eq. (7) for the surface system are still of infinite dimension since there are infinitely many atoms in the entire unit cell. Therefore the eigenvalues in Eq. (7) cannot be calculated by standard numerical diagonalization techniques. Instead of diagonalizing the matrix $[E\underline{S}(\mathbf{q})-\underline{H}(\mathbf{q})]$, we calculate the inverse, or the resolvent, or the Green function of $[\tilde{E}\underline{S}(\mathbf{q})-\underline{H}(\mathbf{q})]$, where $\tilde{E}=E+i\varepsilon$ is a complex energy with ε positive and real. From this Green function complete information about the one-particle energy spectrum and the charge density $n(\mathbf{r})$ can easily be extracted.[23] The calculation of the Green function $\underline{G}(\tilde{E},\mathbf{q})$, defined by

$$[\tilde{E}\underline{S}(\mathbf{q})-\underline{H}(\mathbf{q})]\underline{G}(\tilde{E},\mathbf{q})=\underline{1} \ , \tag{10}$$

can be achieved by introducing *reference matrices* $\underline{H}^R(\mathbf{q})$ and $\underline{S}^R(\mathbf{q})$ which have the following properties.

(a) $\underline{S}^R(\mathbf{q})$ and $\underline{H}^R(\mathbf{q})$ have the same dimension as $\underline{S}(\mathbf{q})$ and $\underline{H}(\mathbf{q})$.

(b) The *difference matrix*

$$\underline{U}(\tilde{E},\mathbf{q})\equiv\underline{H}(\mathbf{q})-\underline{H}^R(\mathbf{q})-\tilde{E}[\underline{S}(\mathbf{q})-\underline{S}^R(\mathbf{q})] \tag{11}$$

has nonvanishing matrix elements only in a finite subspace B of the infinite vector space.

(c) The inverse matrix

$$\underline{G}^R(\tilde{E},\mathbf{q})=[\tilde{E}\underline{S}^R(\mathbf{q})-\underline{H}^R(\mathbf{q})]^{-1} \tag{12}$$

is known.

With the help of these references matrices, the calculation of $\underline{G}(\tilde{E},\mathbf{q})$ can be reduced to the solution of a matrix equation in subspace B only, as will be described in detail below.

To construct suitable reference matrices, we compare the surface problem with the bulk problem formulated in an analogous way. The bulk problem is described by Eqs. (1)–(4), if we exchange $V_{\text{eff}}(\mathbf{r})$ with $V^0_{\text{eff}}(\mathbf{r})$ and if we introduce ρ^0_j and $\lambda^0_{m\mu}$ as atomic positions in Eq. (4) instead of ρ_j and $\lambda_{m\mu}$. In addition, it should be noted that in the bulk lattice potential $V^0(\mathbf{r})$ [see Eq. (4)] the sum over m runs from $-\infty$ to $+\infty$. Analogously, the layer orbitals for the bulk description,

$$\chi^0_{\alpha m\mu}(\mathbf{q},\mathbf{r})=\frac{1}{(N_2)^{1/2}}\sum_j e^{i\mathbf{q}\cdot(\rho^0_j+\lambda^0_{m\mu})}\varphi^0_\alpha(\mathbf{r}-\rho^0_j-\lambda^0_{m\mu}) \ , \tag{13}$$

are defined for all m from $-\infty \le m \le \infty$ and not only from $-\infty \le m \le -1$, as was the case for the layer orbitals $\{\chi_l(\mathbf{q},\mathbf{r})\}$ of the surface system, as given in Eq. (5). We exclude $m=0$ in our notation. There is one essential difference between the two sets of basis functions given in Eqs. (5) and (13), respectively. The atomic orbitals $\{\varphi_\alpha\}$ in the set $\{\chi_l(\mathbf{q},\mathbf{r})\}$ may be different near the surface from the orbitals $\{\varphi^0_\alpha\}$ in the bulk basis set $\{\chi^0_l(\mathbf{q},\mathbf{r})\}$. In the case of a strong *surface reconstruction,* e.g., the displacement of the atoms near the surface necessitates the use of a basis set with correspondingly displaced orbitals. Also, in *adsorption* studies it may be more efficient to represent the surface wave functions by orbitals of different symmetry type and different decay constants, as compared to the bulk orbitals, because of the possibly different chemical nature of the adsorbed species.[38] Representing the bulk problem in the basis set $\{\chi^0_l(\mathbf{q},\mathbf{r})\}$ leads to

$$[\underline{H}^0(\mathbf{q})-E\underline{S}^0(\mathbf{q})]\mathbf{C}^{0,s}(\mathbf{q})=\mathbf{0} \ . \tag{14}$$

The matrix elements of $\underline{H}^0$ and $\underline{S}^0$ are defined in analogy to Eqs. (8a) and (8b) with χ replaced by χ^0 and V_{eff} replaced by V^0_{eff}. These matrices are also of infinite dimensions since the bulk problem is formulated in the surface-symmetry-adapted reference frame. Of course, for the bulk crystal there is an additional translational symmetry in the surface-perpendicular direction. Thereby, Eq. (14) can simply be reduced to a finite-size matrix equation. We will extensively use this fact below in the construction of the bulk Green function. Yet, it is the special form of Eq. (14) for the bulk description in the layer orbital basis which enables us to solve Eq. (7). For this purpose, we consider the matrices in more detail and compare the structure of $\underline{H}(\mathbf{q})$ and $\underline{S}(\mathbf{q})$ with $\underline{H}^0(\mathbf{q})$ and $\underline{S}^0(\mathbf{q})$. The matrix

$$\underline{M}(\tilde{E},\mathbf{q})=-\underline{H}(\mathbf{q})+\tilde{E}\underline{S}(\mathbf{q}) \tag{15}$$

has the structure of a band matrix as shown schematically in Fig. 2. It is built up by block matrices of size $N_{\alpha\mu}\cdot N_\mu \times N_{\alpha\mu}\cdot N_\mu$, where $N_{\alpha\mu}$ is the number of localized orbitals φ_α at the atomic position μ in the unit cell and N_μ is the number of atoms in each two-dimensional unit cell. The width of the band matrix is determined by d_n. The matrix

$$\underline{M}^0(\tilde{E},\mathbf{q})=-\underline{H}^0(\mathbf{q})+\tilde{E}\underline{S}^0(\mathbf{q}) \tag{16}$$

has precisely the same structure as $\underline{M}$, but its size is twice that of $\underline{M}(\tilde{E},\mathbf{q})$. To be able to take differences of the matrices, we extend $\underline{M}$ to the size of $\underline{M}^0$ by filling the corresponding elements with zeros. Comparing now $\underline{M}$ and $\underline{M}^0$, the following can be stated: within the bulk region inside the crystal on the layers from $m=-m_u$ to $m=-\infty$, the corresponding elements of both matrices are equal.[39] This is a consequence of both the localization of the basis functions and, in particular, of the small spatial extension of the transition area (see Fig. 1). To write the matrices $\underline{M}$ and $\underline{M}^0$ more explicitly, we decompose the full vector space into four regions labeled A–D according to the layer indices spanning the regions

$$A:\ -\infty \le m < -m_u,\quad B:\ -m_u \le m \le -1\ ;$$
$$C:\ 1 \le m \le m_d,\quad D:\ m_d \le m \le \infty\ , \tag{17}$$

with $m_u \ge m_d$. The layer number m_d is determined by the coupling range of the matrix elements in the bulk Hamiltonian and corresponds to the distance d_n as specified below Eq. (9). The number m_u is determined by the range of surface-induced changes in the charge density and in the potential, as compared to the bulk, and, in addition, by the coupling range of the Hamiltonian matrix elements which sense the above-mentioned changes. The matrices now take the following forms:

$$\underline{M} = \begin{pmatrix} \underline{M}_{AA} & \underline{M}_{AB} & \underline{0} & \underline{0} \\ \underline{M}_{BA} & \underline{M}_{BB} & \underline{0} & \underline{0} \\ \underline{0} & \underline{0} & \underline{0} & \underline{0} \\ \underline{0} & \underline{0} & \underline{0} & \underline{0} \end{pmatrix}$$
$$= \begin{pmatrix} \underline{M}^0_{AA} & \underline{M}^0_{AB} & \underline{0} & \underline{0} \\ \underline{M}^0_{BA} & \underline{M}^0_{BB} + \underline{\Delta M}_{BB} & \underline{0} & \underline{0} \\ \underline{0} & \underline{0} & \underline{0} & \underline{0} \\ \underline{0} & \underline{0} & \underline{0} & \underline{0} \end{pmatrix} \tag{18}$$

and

$$\underline{M}^0 = \begin{pmatrix} \underline{M}^0_{AA} & \underline{M}^0_{AB} & \underline{0} & \underline{0} \\ \underline{M}^0_{BA} & \underline{M}^0_{BB} & \underline{M}^0_{BC} & \underline{0} \\ \underline{0} & \underline{M}^0_{CB} & \underline{M}^0_{CC} & \underline{M}^0_{CD} \\ \underline{0} & \underline{0} & \underline{M}^0_{DC} & \underline{M}^0_{DD} \end{pmatrix} . \tag{19}$$

In principle, the extension of region B or the value of m_u, respectively, is defined by the condition $\underline{M}_{BA} \equiv \underline{M}^0_{BA}$. In practice, however, m_u has to be determined by a convergence study where m_u is increased until the effects of the increasingly smaller differences between the calculated matrix elements in $\underline{M}_{BA}$ and $\underline{M}^0_{BA}$ on the outcome of the calculations become negligible. Inspection of Eqs. (18) and (19) shows that $\underline{M}^0$ does not have the property (b) as required for a suitable reference matrix in Eq. (11). However, the upper left-hand part of $\underline{M}^0$, which we introduce as

$$\underline{M}^I = \begin{pmatrix} \underline{M}^0_{AA} & \underline{M}^0_{AB} & \underline{0} & \underline{0} \\ \underline{M}^0_{BA} & \underline{M}^0_{BB} & \underline{0} & \underline{0} \\ \underline{0} & \underline{0} & \underline{0} & \underline{0} \\ \underline{0} & \underline{0} & \underline{0} & \underline{0} \end{pmatrix} , \tag{20}$$

does fulfill the requirements imposed on the reference matrix, *if* it is possible to calculate its Green function $\underline{G}^I(\tilde{E},\mathbf{q})$ defined by

$$\underline{M}^I(\tilde{E},\mathbf{q})\underline{G}^I(\tilde{E},\mathbf{q}) = \begin{pmatrix} \underline{1}_{AA} & \underline{0} & \underline{0} & \underline{0} \\ \underline{0} & \underline{1}_{BB} & \underline{0} & \underline{0} \\ \underline{0} & \underline{0} & \underline{0} & \underline{0} \\ \underline{0} & \underline{0} & \underline{0} & \underline{0} \end{pmatrix} . \tag{21}$$

The inversion of $\underline{M}^I(\tilde{E},\mathbf{q})$ can be achieved in an intermediate step by making use of the properties of $\underline{M}^0(\tilde{E},\mathbf{q})$ and $\underline{G}^0(\tilde{E},\mathbf{q})$. For this purpose we introduce an auxiliary matrix,

$$\tilde{\underline{M}}^I = \begin{pmatrix} \underline{M}^0_{AA} & \underline{M}^0_{AB} & \underline{0} & \underline{0} \\ \underline{M}^0_{BA} & \underline{M}^0_{BB} & \underline{0} & \underline{0} \\ \underline{0} & \underline{0} & \underline{M}^0_{CC} & \underline{M}^0_{CD} \\ \underline{0} & \underline{0} & \underline{M}^0_{DC} & \underline{M}^0_{DD} \end{pmatrix} , \tag{22}$$

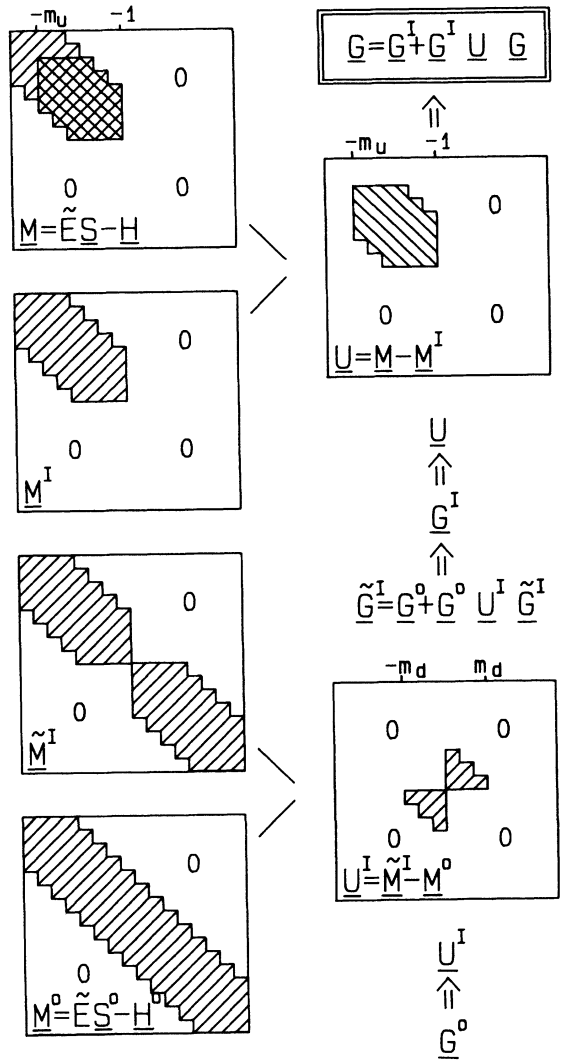


FIG. 2. Schematic plot of the matrices involved in the calculation of the surface Green function $\underline{G}$. The hierarchy develops from bottom to top as $\underline{G}^0$, $\tilde{\underline{G}}^I$, $\underline{G}^I$, and finally $\underline{G}$, which is defined as the inverse matrix of $\underline{M}$. $\underline{G}$ is constructed from $\underline{G}^0$, the inverse of the bulk matrix $\underline{M}^0$, in two consecutive steps involving the appropriate perturbation matrices $\underline{U}^I$ and $\underline{U}$. In the first step, two idealized surfaces described by $\tilde{\underline{M}}^I$ are created from the bulk crystal by the perturbation matrix $\underline{U}^I$, which is the difference matrix of $\tilde{\underline{M}}^I$ and $\underline{M}^0$. Then one of the two decoupled semi-infinite crystals is ignored, leading to $\underline{M}^I$. The corresponding Green function $\underline{G}^I$ is simply given by the upper left-hand block of $\tilde{\underline{G}}^I$. Finally, the true surface is introduced by $\underline{U}=\underline{M}-\underline{M}^I$ and the surface Green function results from the corresponding Dyson equation.

which can be rewritten as a sum of the bulk matrix $\underline{M}^0$ and a *localized perturbation matrix* $\underline{U}^I$, i.e.,

$$\underline{\tilde{M}}^I(\tilde{E},\mathbf{q})=\underline{M}^0(\tilde{E},\mathbf{q})+\underline{U}^I(\tilde{E},\mathbf{q}) , \tag{23}$$

with

$$\underline{U}^I(\tilde{E},\mathbf{q})=\begin{pmatrix} \underline{0} & \underline{0} & \underline{0} & \underline{0} \\ \underline{0} & \underline{0} & -\underline{M}^0_{BC} & \underline{0} \\ \underline{0} & -\underline{M}^0_{CB} & \underline{0} & \underline{0} \\ \underline{0} & \underline{0} & \underline{0} & \underline{0} \end{pmatrix} . \tag{24}$$

There is no coupling between the upper left-hand and lower right-hand blocks in $\underline{\tilde{M}}^I$. This matrix obviously describes two decoupled bulklike semi-infinite solids, i.e., it describes two decoupled *idealized surfaces.* We will define the term "idealized" more precisely further below. The inverse matrix or Green function of $\underline{\tilde{M}}^I$ is simply given by

$$\underline{\tilde{G}}^I(\tilde{E},\mathbf{q})=\begin{pmatrix} \underline{\tilde{G}}^I_{AA} & \underline{\tilde{G}}^I_{AB} & \underline{0} & \underline{0} \\ \underline{\tilde{G}}^I_{BA} & \underline{\tilde{G}}^I_{BB} & \underline{0} & \underline{0} \\ \underline{0} & \underline{0} & \underline{\tilde{G}}^I_{CC} & \underline{\tilde{G}}^I_{CD} \\ \underline{0} & \underline{0} & \underline{\tilde{G}}^I_{DC} & \underline{\tilde{G}}^I_{DD} \end{pmatrix} . \tag{25}$$

The upper left-hand block is identical to the Green function of $\underline{M}^I(\tilde{E},\mathbf{q})$ that we are looking for, i.e.,

$$\underline{G}^I(\tilde{E},\mathbf{q})=\begin{pmatrix} \underline{\tilde{G}}^I_{AA} & \underline{\tilde{G}}^I_{AB} & \underline{0} & \underline{0} \\ \underline{\tilde{G}}^I_{BA} & \underline{\tilde{G}}^I_{BB} & \underline{0} & \underline{0} \\ \underline{0} & \underline{0} & \underline{0} & \underline{0} \\ \underline{0} & \underline{0} & \underline{0} & \underline{0} \end{pmatrix} . \tag{26}$$

It is important to note, that, e.g., $\underline{\tilde{G}}^I_{AA}$ contains the matrix elements of $\underline{G}^I(\tilde{E},\mathbf{q})$ in subspace A and it is *not* the inverse or Green function of $\underline{M}^0_{AA}$. Due to the special form of the perturbation matrix $\underline{U}^I(\tilde{E},\mathbf{q})$, which contains only a few elements of $\underline{M}^0(\tilde{E},\mathbf{q})$ with opposite sign, the Dyson equation for $\underline{G}^I(\tilde{E},\mathbf{q})$,

$$\underline{G}^I(\tilde{E},\mathbf{q})=\underline{G}^0(\tilde{E},\mathbf{q})+\underline{G}^0(\tilde{E},\mathbf{q})\underline{U}^I(\tilde{E},\mathbf{q})\underline{G}^I(\tilde{E},\mathbf{q}) , \tag{27}$$

can easily be solved, as is shown in Appendix A. With the help of Eqs. (18), (20), (26), and (27), we are now able to calculate the Green function. The matrix $\underline{M}(\tilde{E},\mathbf{q})$, the reference matrix $\underline{M}^I(\tilde{E},\mathbf{q})$, and its Green function $\underline{G}^I(\tilde{E},\mathbf{q})$ have nonvanishing elements in subspaces A and B only. Therefore, we can restrict ourselves to these subspaces in the following. The difference matrix $\underline{U}(\tilde{E},\mathbf{q})$, as defined in Eq. (11), now has the form

$$\underline{U}(\tilde{E},\mathbf{q})=\underline{M}^I(\tilde{E},\mathbf{q})-\underline{M}(\tilde{E},\mathbf{q})=\begin{pmatrix} \underline{0} & \underline{0} \\ \underline{0} & \underline{U}_{BB}(\tilde{E},\mathbf{q}) \end{pmatrix} , \tag{28}$$

with

$$\underline{U}_{BB}(\tilde{E},\mathbf{q})=\underline{H}_{BB}(\mathbf{q})-\underline{H}^0_{BB}(\mathbf{q})-\tilde{E}[\underline{S}_{BB}(\mathbf{q})-\underline{S}^0_{BB}(\mathbf{q})] . \tag{29}$$

Inserting (28) into Eq. (10), we obtain, by employing Eq. (21), the Dyson equation for the surface Green function,[23]

$$\underline{G}(\tilde{E},\mathbf{q})=\underline{G}^I(\tilde{E},\mathbf{q})+\underline{G}^I(\tilde{E},\mathbf{q})\underline{U}(\tilde{E},\mathbf{q})\underline{G}(\tilde{E},\mathbf{q}) . \tag{30}$$

Since the difference matrix $\underline{U}(\tilde{E},\mathbf{q})$ is restricted to subspace B, the surface Green function is given by

$$\underline{G}(\tilde{E},\mathbf{q})=\begin{pmatrix} \underline{G}^I_{AA} & \underline{G}^I_{AB} \\ \underline{G}^I_{BA} & \underline{G}^I_{BB} \end{pmatrix} + \begin{pmatrix} \underline{G}^I_{AA} & \underline{G}^I_{AB} \\ \underline{G}^I_{BA} & \underline{G}^I_{BB} \end{pmatrix} \begin{pmatrix} \underline{0} & \underline{0} \\ \underline{0} & \underline{T}_{BB} \end{pmatrix} \begin{pmatrix} \underline{G}^I_{AA} & \underline{G}^I_{AB} \\ \underline{G}^I_{BA} & \underline{G}^I_{BB} \end{pmatrix} , \tag{31}$$

where the scattering matrix $\underline{T}_{BB}$ is given by

$$\underline{T}_{BB}(\tilde{E},\mathbf{q})=\underline{U}_{BB}(\tilde{E},\mathbf{q})[\underline{1}_{BB}-\underline{G}^I_{BB}(\tilde{E},\mathbf{q})\underline{U}_{BB}(\tilde{E},\mathbf{q})]^{-1} . \tag{32}$$

It is important to realize at this point that *only matrix inversions in* the finite *subspace B* have to be carried out. This fact ultimately allows the numerical solution of the Dyson equation.

Let us pause for a moment in the formal description of the method in order to summarize the procedure of calculating the surface Green function described so far from a more physical point of view. Starting from a perfect, three-dimensionally periodic bulk crystal described by the potential $V^0_{\text{eff}}(\mathbf{r})$, we have first constructed the matrices $\underline{M}^0(\tilde{E},\mathbf{q})$ and $\underline{G}^0(\tilde{E},\mathbf{q})$ in a layer orbital basis. In the following step, two *identical* semi-infinite solids have been created by the perturbation $\underline{U}^I(\tilde{E},\mathbf{q})$. Only one of these two decoupled semi-infinite crystals enters the following calculations. The other one can be ignored. In the final and most important step the real surface is created by the perturbation $\underline{U}(\tilde{E},\mathbf{q})$. This perturbation takes into account both the changes of atomic positions in the selvedge due to relaxation or reconstruction and the changes in the surface potential originating from rearrangements of the electronic charge density near the surface. The whole procedure is pictorially summarized in Fig. 2.

To avoid conceivable confusions, we want to specify the term *idealized* surface. In our context it is a purely mathematical, auxiliary construction represented by the matrix $\underline{\tilde{M}}^I(\tilde{E},\mathbf{q})$. Here we calculate the matrix elements of the *bulk* Hamiltonian, containing the bulk potential, in the bulk layer orbital basis χ^0, which is restricted, however, to the two half-spaces. Thus there is no coupling between the left and right half-space in $\underline{\tilde{M}}^I(\tilde{E},\mathbf{q})$ as defined in Eq. (22). In contrast to this idealized semi-infinite crystal with a physically meaningless surface, we can define an *ideal surface* in a more physical sense as follows: In this case the atoms still reside at their bulk positions on the layers from $m=-\infty$ up to $m=-1$, but the rearrangement of the charge density at the geometrically ideal surface, as compared to the bulk situation, is self-consistently taken into account. In a former paper,[11] we have calculated the electronic properties of ideal Si(001)-(1×1) in this way using the Green-function formalism.

Now we continue in the description of the formalism. The perturbation matrix $\underline{U}(\tilde{E},\mathbf{q})$ depends on the charge density $n(\mathbf{r})$ of the semi-infinite crystal with its reconstructed surface. The charge density

$$n(\mathbf{r})=2\int_{-\infty}^{E_F}\sum_{\mathbf{q}}\sum_{s}\psi_{s,\mathbf{q}}(\mathbf{r})\psi^{*}_{s,\mathbf{q}}(\mathbf{r})\delta(E-E_s(\mathbf{q}))dE$$
$$=2\int_{-\infty}^{E_F}\sum_{\mathbf{q}}\sum_{s}\sum_{l,l'}\chi_l(\mathbf{q},\mathbf{r})C^{s}_{l}(\mathbf{q})C^{s}_{l'}{}^{*}(\mathbf{q})\chi^{*}_{l'}(\mathbf{q},\mathbf{r})\delta(E-E_s(\mathbf{q}))dE \tag{33}$$

can be calculated using the surface Green function

$$n(\mathbf{r})=-\frac{2}{\pi}\lim_{\varepsilon\to 0^+}\mathrm{Im}\left[\sum_{\mathbf{q}}\sum_{l,l'}\chi_l(\mathbf{q},\mathbf{r})\int_{-\infty}^{E_F}G_{l,l'}(E+i\varepsilon,\mathbf{q})dE\chi^{*}_{l'}(\mathbf{q},\mathbf{r})\right]. \tag{34}$$

We remind the reader that the superindex l stands for α,m,μ. If we want to calculate $n(\mathbf{r})$ at the surface, the sums over m and m' can be restricted to subspace B due to the strong localization of the layer orbitals. The calculation of the charge density deeper in the selvedge and in the crystal can be achieved using the following representation of the charge density:

$$n(\mathbf{r})=n^{0}(\mathbf{r})+\delta n(\mathbf{r}), \tag{35}$$

with

$$\delta n(\mathbf{r})=-\frac{2}{\pi}\lim_{\varepsilon\to 0^+}\mathrm{Im}\sum_{\mathbf{q}}\sum_{l,l'}\left[\chi_l(\mathbf{q},\mathbf{r})\int_{-\infty}^{E_F}G_{l,l'}(E+i\varepsilon,\mathbf{q})dE\chi^{*}_{l'}(\mathbf{q},\mathbf{r})-\chi^{0}_{l}(\mathbf{q},\mathbf{r})\int_{-\infty}^{E_F}G^{0}_{l,l'}(E+i\varepsilon,\mathbf{q})dE\chi^{0*}_{l'}(\mathbf{q},\mathbf{r})\right]. \tag{36}$$

In principle, we have to extend the sums over m and m' over all layers from $m=-\infty$ to $m=+\infty$. The contributions of $G_{ll'}(\tilde{E},\mathbf{q})$, however, vanish if m or $m'\geq 1$.

If $\mathbf{r}$ is *inside* the crystal, we can restrict the sums over m and m' to subspace B. This results from the fact that due to the localization of the layer orbitals the contribution of the second term in Eq. (36) for m *or* $m'\geq 1$ is negligible. On the other hand, the contributions of the first and second terms in Eq. (36) cancel each other for m *and* $m'<-m_u$. This condition actually defines m_u for a chosen accuracy of the calculations. Thus only matrix elements of $\underline{G}^{I}_{BB}(\tilde{E},\mathbf{q})$ and $\underline{U}_{BB}(\tilde{E},\mathbf{q})$ enter the calculation of the charge density via Eqs. (31), (34), and (36). The dimension of the matrices to be evaluated in the scattering-theoretical method is determined by the spatial extent of the transition region in the surface potential (see Fig. 1), which has to be calculated self-consistently.

The most important single-particle quantity in the outcome of the calculations is the layer density of states (LDOS). It allows one to clearly discern bound surface states from surface resonances in the electronic spectrum of a surface system. Furthermore, it is most important for comparisons of the theoretical results with experimental data of angle-integrated and angle-resolved surface spectroscopy. The LDOS is given by

$$N_l(E,\mathbf{q})=2\sum_{l'}\sum_{s}C^{s}_{l}(\mathbf{q})C^{s}_{l'}{}^{*}(\mathbf{q})S_{l',l}(\mathbf{q})\delta(E-E_s(\mathbf{q})) \tag{37}$$

and can be calculated from the surface Green function. To emphasize the extreme resolution of our approach with respect to the relevant degrees of freedom, we write $l=\alpha,m,\mu$ explicitly in

$$N_{\alpha m\mu}(E,\mathbf{q})=-\frac{2}{\pi}\lim_{\varepsilon\to 0^+}\mathrm{Im}\sum_{\alpha',m',\mu'}G_{\alpha m\mu,\alpha' m'\mu'}(E,\mathbf{q})\times S_{\alpha' m'\mu',\alpha m\mu}(\mathbf{q}). \tag{38}$$

According to Eq. (38), the layer state density can easily be resolved with respect to energy E, wave vector $\mathbf{q}$, orbital-type α, layer index m, and basis atom μ. This is extremely useful for both the physical interpretation of surface features on the basis of theoretical results and for comparisons with experimental data from high-resolution surface spectroscopy.

Finally, we mention that there is an alternative approach to compute the electronic properties of solid surfaces within the concept of self-consistent scattering theory. It is formally somewhat more straightforward than the formalism given in this subsection, but it is much more time consuming in the actual calculations. We defer its description to Appendix B.

B. Bulk Green function G^0

The evaluation of the bulk Green function $\underline{G}^{0}(\tilde{E},\mathbf{q})$ in the layer orbital basis $\{\chi^{0}_{l}(\mathbf{q},\mathbf{r})\}$ is the starting point of the calculations. It is defined by

$$[\tilde{E}\underline{S}^{0}(\mathbf{q})-\underline{H}^{0}(\mathbf{q})]\underline{G}^{0}(\tilde{E},\mathbf{q})=\underline{1}. \tag{39}$$

Evaluating the matrix elements of $\underline{G}^{0}$ in a straightforward way yields precisely the same analytical form as in the tight-binding version[19,20] of the theory. We find

$$G^{0}_{\alpha m\mu,\alpha' m'\mu}(\tilde{E},\mathbf{q})=\sum_{\mathbf{g}_j}{}'\,e^{i\mathbf{g}_j\cdot(\lambda^{0}_{m\mu}-\lambda^{0}_{m'\mu'})}\frac{N_2}{N_3}\sum_{\mathbf{k}_\perp}e^{i\mathbf{k}_\perp\cdot(\lambda^{0}_{m\mu}-\lambda^{0}_{m'\mu'})}\sum_{n}\frac{a^{n}_{\alpha,\nu(m\mu)}(\mathbf{k}_\perp,\mathbf{q}+\mathbf{g}_j)a^{n}_{\alpha',\nu'(m'\nu')}{}^{*}(\mathbf{k}_\perp,\mathbf{q}+\mathbf{g}_j)}{\tilde{E}-E_n(\mathbf{k}_\perp,\mathbf{q}+\mathbf{g}_j)}. \tag{40}$$

Here, $\{E_n\}$ and $\{a^n_{\alpha\nu}\}$ are the eigenvalues and eigenvectors of the bulk electronic-structure problem. The bulk basis atom index ν is uniquely related to layer index m and basis atom index μ in the surface unit cell. The vector $\mathbf{g}_j$ is an element of the two-dimensional reciprocal net, and the prime at the sum over the values $\{\mathbf{g}_j\}$ is meant to indicate that only those $\mathbf{g}_j$ vectors that fulfill the requirement that the vector $\mathbf{k}_\perp+\mathbf{q}+\mathbf{g}_j$ is an element of the first bulk Brillouin zone have to be taken into account. Thus the summation interval in the $\mathbf{k}_\perp$ sum depends, in principle, on $\mathbf{g}_j$. It has been shown that the summation interval for $\mathbf{k}_\perp$ can be transformed using bulk symmetry properties into a sum over $\mathbf{k}_\perp$ with fixed boundaries. Of course, thereby the set of $\mathbf{g}_j$ vectors that must be summed is further reduced. [For more details of the calculation of $\underline{G}^0_{l,l'}(\tilde{E},\mathbf{q})$, see Refs. 19 and 40.] Transforming the $\mathbf{k}_\perp$ sum into an integral and employing an analytic continuation of the integral into the complex $\mathbf{k}_\perp$ plane (conformal mapping) lead to very efficient schemes[41–45] for the evaluation of $\underline{G}^0(\tilde{E},\mathbf{q})$. The numerical effort in our method rapidly increases with increasing maximal range of interactions that need to be retained in the Hamiltonian and overlap matrices. In the empirical tight-binding version of the approach, only interactions up to second-nearest neighbors are typically taken into account. In the self-consistent version we have to deal with longer-ranged interactions. In the calculations presented in this paper, e.g., it turned out to be necessary for convergence to retain interactions up to fifth-nearest neighbors. Therefore it has been expedient to calculate the bulk Green function directly from Eq. (40) on a very dense $\mathbf{k}_\perp$ mesh.

III. COMPUTATIONAL ASPECTS

A. Surface potential

In this subsection we describe the calculation of the effective one-particle potential $V_{\text{eff}}(\mathbf{r})$ from the charge density. The direct computation of the potential in real space is unsuitable due to the nonlocal dependence of the Coulomb potential on the charge density. In our approach we make use of electrostatic neutrality of the system. For this purpose we introduce the total charge density $n_{\text{tot}}(\mathbf{r})$. Since we are using pseudopotentials in the actual calculations for the Si(001)-(2×1) surface, n_{tot} is defined as the sum of the charge density of the valence electrons $n(\mathbf{r})$ and of the ions $n_i(\mathbf{r})$, i.e.,

$$n_{\text{tot}}(\mathbf{r})=n(\mathbf{r})+n_i(\mathbf{r}) . \tag{41}$$

The ionic charge density $n_i(\mathbf{r})$ was constructed from the representation of the full potential given in Eq. (2) as a Hartree plus an exchange-correlation potential,

$$V_{\text{eff}}(\mathbf{r})=2\int \frac{n_{\text{tot}}(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}d^3r'+V_{\text{xc}}[n(\mathbf{r})] . \tag{42}$$

Note that the electronic charge density is counted positive, while the ionic charge density $n_i(\mathbf{r})$ is counted negative, as usual in density-functional theory. We note in passing that, if nonlocal pseudopotentials are used, $n_i(\mathbf{r})$ is determined by the long-range part of these potentials.

The effective potential $V_{\text{eff}}(\mathbf{r})$, averaged over the surface-parallel components of $\mathbf{r}$, is shown in Fig. 3(a). The dots indicate the positions of the layers on which the layer orbitals are localized. Instead of computing the potential $V_{\text{eff}}(\mathbf{r})$ from the charge density $n_{\text{tot}}(\mathbf{r})$ of the semi-infinite crystal, we construct a potential $\tilde{V}_{\text{eff}}(\mathbf{r})$ from the symmetrized charge density,

$$\tilde{n}_{\text{tot}}(\mathbf{r})=\begin{cases} n_{\text{tot}}(\mathbf{r}) & \text{for } z<0 , \\ n_{\text{tot}}(-\mathbf{r}) & \text{for } z>0 . \end{cases} \tag{43}$$

The advantage of introducing this auxiliary potential $\tilde{V}_{\text{eff}}(\mathbf{r})$ [see Fig. 3(b)] lies in the fact that it has inversion symmetry by construction, i.e.,

$$\tilde{V}_{\text{eff}}(\mathbf{r})=\begin{cases} V_{\text{eff}}(\mathbf{r}) & \text{for } z<0 , \\ V_{\text{eff}}(-\mathbf{r}) & \text{for } z>0 , \end{cases} \tag{44}$$

and can thus be represented by the bulk potential $V^0_{\text{eff}}(\mathbf{r})$ and a perturbation $\delta\tilde{V}(\mathbf{r})$, which is localized in the z direction, as shown in Fig. 3(c),

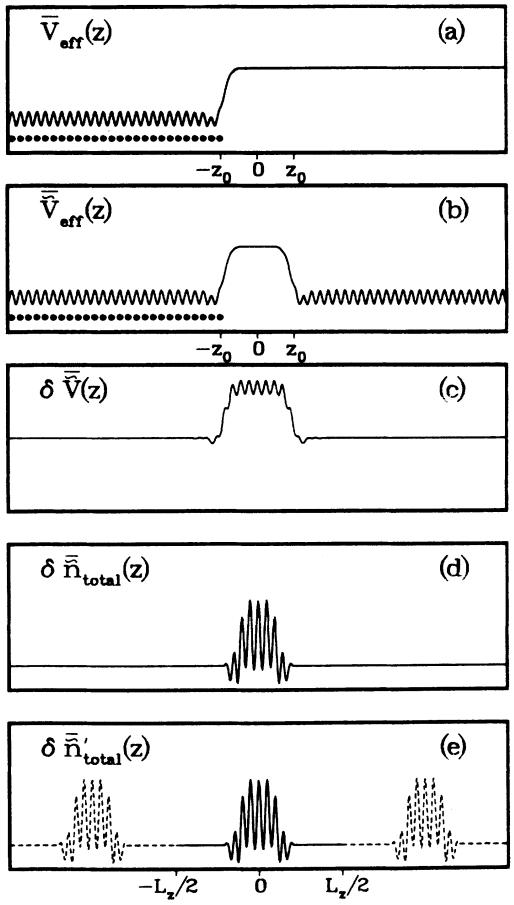


FIG. 3. Schematic plot of (a) the averaged potential $\underline{V}_{\text{eff}}(z)$, (b) the auxiliary potential $\bar{\tilde{V}}_{\text{eff}}(z)$, and (c) the difference potential $\delta\bar{\tilde{V}}(z)$ in the surface region. The charge-density difference $\delta\bar{\tilde{n}}_{\text{tot}}(z)$ is shown in (d). Its periodic continuation beyond the definition interval is shown by dashed lines in (e).

$$\tilde{V}_{\rm eff}(r)=V^0_{\rm eff}(\mathbf{r})+\delta\tilde{V}(\mathbf{r})\ . \tag{45}$$

The corresponding charge-density difference $\delta\tilde{n}_{\rm tot}(\mathbf{r})$ is also localized in the same spatial region, as can be seen in Fig. 3(d). It is important to note that we can use $\tilde{V}_{\rm eff}(\mathbf{r})$ for the calculation of the matrix elements in Eq. (8a) instead of $V_{\rm eff}(\mathbf{r})$ if the distance z_0 between the center of inversion symmetry and the first layer [see Fig. 3(b)] is large enough. This results from the strong localization of the layer orbitals in the left half-space $z\le -z_0$.

Now we address the calculation of $\delta\tilde{V}(\mathbf{r})$ from $\delta\tilde{n}_{\rm tot}(\mathbf{r})$. Parallel to the surface, $\delta\tilde{V}(\mathbf{r})$ and $\delta\tilde{n}_{\rm tot}(\mathbf{r})$ are periodic functions, which rapidly go to zero outside the perturbation region $-L_z/2\le z\le L_z/2$. Standard Fourier-expansion techniques, well known from the *three-dimensional* systems, can be applied for $\tilde{V}_{\rm eff}(\mathbf{r})$, when a little formal trick is used. For this purpose we employ a periodic continuation of $\delta\tilde{n}_{\rm tot}(\mathbf{r})$ along the z axis outside the interval $[-L_z/2, L_z/2]$, labeled $\delta\tilde{n}'_{\rm tot}(\mathbf{r})$, with the following properties [see Fig. 3(e)]:

$$\delta\tilde{n}_{\rm tot}(\mathbf{r})=\begin{cases}\delta\tilde{n}'_{\rm tot}(\mathbf{r}) & \text{for } |z|\le L_z/2\ ,\\ 0 & \text{for } |z|>L_z/2\ ,\end{cases} \tag{46}$$

with

$$\delta\tilde{n}'_{\rm tot}(\mathbf{r})=\sum_{\mathbf{g},g_z}\delta\tilde{\tilde{n}}_{\rm tot}(\mathbf{g},\mathbf{g}_z)e^{i(\mathbf{g}+\mathbf{g}_z)\cdot\mathbf{r}}\ . \tag{47}$$

In this Fourier expansion of the fictitious charge-density difference $\delta\tilde{n}'_{\rm tot}(\mathbf{r})$, $\mathbf{g}$ is a two-dimensional vector of the reciprocal surface net and $\mathbf{g}_z$ is given by

$$\mathbf{g}_z=g_z\hat{\mathbf{z}}=\frac{2\pi}{L_z}n_z\hat{\mathbf{z}}\ \text{ with } n_z=0,\pm1,\pm2,\ldots\ . \tag{48}$$

The electronic charge-density change $\delta n(\mathbf{r})$ is calculated directly in real space from Eq. (36) for $-L_z/2\le z\le 0$ on a mesh of $n_x n_y n_z/2$ points in the surface unit cell. Then it is continued to $0<z\le L_z/2$ due to inversion symmetry. The electronic part of $\delta\tilde{\tilde{n}}_{\rm tot}(\mathbf{g},\mathbf{g}_z)$ is computed by the standard summation technique of the Cooley-Tukey discrete complex fast Fourier transformation[46] (FFT) from $\delta n(\mathbf{r})$. The ionic contributions to the total charge-density difference will be added afterwards.

In the calculation of the Coulomb potential from the charge density, we have to take into account that the Fourier series in Eq. (47) is limited to $|z|\le L_z/2$. Therefore, the representation of $\delta\tilde{V}_{\rm tot}(\mathbf{r})$ in terms of $\delta\tilde{\tilde{n}}_{\rm tot}(\mathbf{g},\mathbf{g}_z)$ contains corrections depending on L_z in addition to the "usual" Fourier representation of the Coulomb potential[47] in the case of three-dimensional periodicity. Our computations show that the main contribution of these corrections is a constant shift C of the potential,

$$C=\sum_{\mathbf{g}_z(\neq\mathbf{0})}8\pi\,\delta\tilde{\tilde{n}}_{\rm tot}(0,\mathbf{g}_z)g_z^{-2}\cos\left[\frac{g_z L_z}{2}\right]\ . \tag{49}$$

The other contributions are extremely small for suitably large L_z values. This is a result of the neutrality of $\delta\tilde{n}_{\rm tot}(\mathbf{r})$ and of the vanishing dipole moment due to inversion symmetry. In the case of the reconstructed Si(001)-(2×1) surface with $L_z=3a_0$, these corrections are vanishingly small and we have therefore neglected them. The difference potential

$$\delta\tilde{V}(\mathbf{r})=\delta\tilde{V}_{\rm tot}(\mathbf{r})+\delta\tilde{V}_{\rm xc}(\mathbf{r})$$

was calculated from

$$\delta\tilde{V}_{\rm tot}(\mathbf{r})=\begin{cases}\displaystyle\sum_{\mathbf{g},g_z}8\pi\delta\tilde{\tilde{n}}_{\rm tot}(\mathbf{g},\mathbf{g}_z)(g^2+g_z^2)^{-1}e^{i(\mathbf{g}+\mathbf{g}_z)\cdot\mathbf{r}}-C & \text{for } |z|\le L_z/2\ ,\\ 0 & \text{for } |z|\ge L_z/2\ .\end{cases} \tag{50}$$

Finally, the exchange-correlation potential $V_{\rm xc}(\mathbf{r})$ is calculated in the local-density approximation from the electronic charge density on the same point mesh as used for the Hartree potential,

$$\delta\tilde{V}_{\rm xc}(\mathbf{r})=\tilde{V}_{\rm xc}(n^0(\mathbf{r})+\delta\tilde{n}(\mathbf{r}))-V_{\rm xc}(n^0(\mathbf{r}))\ , \tag{51}$$

and it is transformed into reciprocal space by FFT. Using Gaussians as localized orbitals, the matrix elements in Eq. (8a) can all be calculated analytically in the Fourier representation of the potential for each $\mathbf{g},\mathbf{g}_z$.

B. Charge-density calculation by complex energy integration

Computation of the charge-density difference $\delta n(\mathbf{r})$ requires the determination of the energy-integrated Green functions according to Eq. (39). These integrations can be carried out most efficiently by an analytical continuation of the integration path from the real energy axis onto a contour in the complex energy plane, as was suggested by Williams *et al.* in Ref. 23. For details, see Ref. 48. The advantage of this method is illustrated in Fig. 4, where we have plotted the negative of the imaginary part of a diagonal element of a Green function as a function of complex energy $E+i\varepsilon$. For $\varepsilon=0$, the picture shows π times a particular layer density of states. Along the real energy axis the Green function mostly shows pronounced structure due to the poles $E_s(\mathbf{q})$. Therefore, an extremely dense mesh of energy points would be necessary for an exact evaluation of the integrals along the real axis. This can be avoided by displacing the contour into the complex plane, yet allowing the exact calculation. We have used the rectangular integration path as shown in Fig. 4(b). Along this path the imaginary part of $G(E,\mathbf{q})$ is a

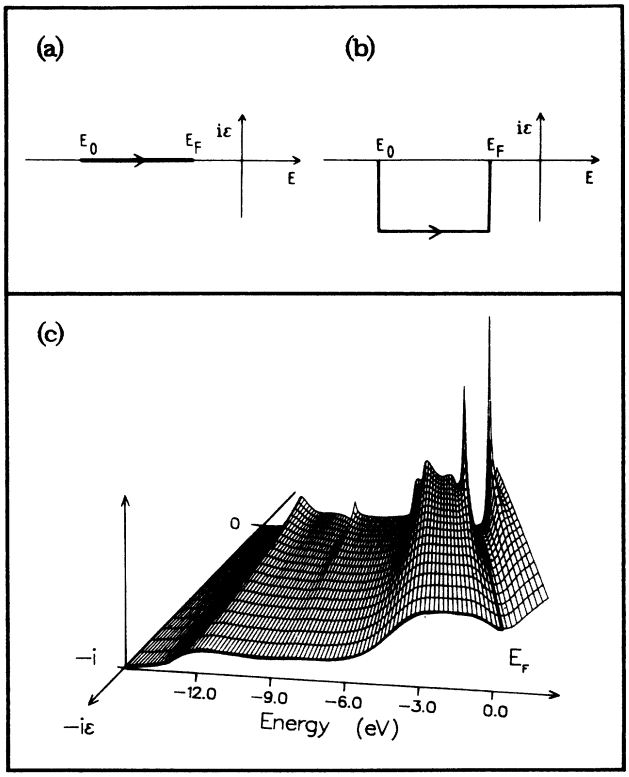


FIG. 4. Displacement of the energy-integration contour in Eq. (36) from the real axis (a) into the complex energy plane (b). The functional form of a typical integrand as a function of ε is shown in the complex plane (c). Note that the integrand has very sharp structures on the real axis (ε=0), while it is well behaved, smooth, and easy to integrate in the complex plane (e.g., for ε=1).

smooth function of energy. The same is true for the real part of $G(E,\mathbf{q})$, which is not shown in Fig. 4. In our case, for ε=1 eV, only 60 sampling points along contour *b* have been sufficient to compute the charge density within 0.5% accuracy. The accuracy can simply be tested by integrating the bulk valence charge density, yielding the number of electrons per unit cell within 0.5%. The requirement on the chosen contour is that it encloses all occupied states. Therefore, this procedure is well suited if the Fermi energy E_F is known *a priori* or if there is a gap between occupied and empty states. To avoid any complications due to either metallic surface-state behavior or the changes of E_F during the self-consistency iterations, we have used the rectangular integration path up to the bulk Fermi level E_F^0. We then continue on the real axis until the Fermi level E_F of the surface system is reached. The very time-consuming calculation of $G^0(E,\mathbf{q})$ has to be done only once on the $E,\mathbf{q}$ mesh because the same rectangular integration contour is used in every self-consistency loop. For advantages or disadvantages of other integration contours, see Ref. 23 or 48, respectively.

C. Iteration scheme

We start our self-consistency loop with an empirical potential $V^{(0)}(\mathbf{r})$ for the semi-infinite crystal. For this purpose we have decomposed the self-consistent bulk potential into a sum of atom-centered spherical pseudopotentials which were fitted to Gaussians. The superposition of these "atomic," short-ranged potentials leads to the starting potential in the surface calculation. The efficiency of this procedure has been demonstrated by Bernholc, Lipari, and Pantelides[16] in their calculations of electronic properties of bulk vacancies in silicon. In the next steps we compute the complete set of matrix elements, solve the Dyson equations for the idealized and reconstructed surfaces, respectively, and calculate the charge density as described in the preceding subsection. The calculation of the Coulomb potential completes the first self-consistency loop. To avoid instabilities[49] in the following iteration steps we have used the iteration scheme due to Anderson.[50] In particular, we have employed a numerical algorithm adapted from Bachelet,[51] which takes up to three preceding iterations into consideration. The iteration process is repeated until input and output potentials agree to within 10^{-5} Ry on our mesh.

IV. APPLICATION TO Si(100)-(2×1)

A. General remarks

In this section we want to present some representative applications of the self-consistent Green-function method for semi-infinite solids and discuss these results from a methodological point of view in comparison with slab or supercell techniques. Further applications of our method for calculations of the electronic properties of (2×1)-reconstructed Si(001) and Ge(001) surfaces and related discussions of the results in comparison with photoemission and scanning-tunneling-spectroscopy data have been given in Refs. 12, 22, 52, and 53. Here, we use Si(001)-(2×1) as an example, since this technologically most important surface has been studied in great detail by many groups. The large variety of experimental and theoretical investigations of this surface has been reviewed, e.g., in Refs. 52 and 54. It is obvious from the literature that a large number of experimental findings can be explained on the basis of the asymmetric dimer model that Chadi[25] proposed on the basis of semiempirical total-energy-minimization calculations. Using this surface configuration, Ihm *et al.*[24] worked out the surface-electronic structure by a self-consistent supercell calculation. Minimizing the total energy in the self-consistent supercell approach, Yin and Cohen[55] arrived at an asymmetric dimer model as well, which showed only quantitative refinements as compared to the former. Later on, more improved models have been studied on the basis of self-consistent total-energy calculations by Northrup[56] and by Pandey.[57] For the following methodological discussion, we choose *one* particular *model*. We have previously discussed the physical aspects of the various different models in comparison with spectroscopy data in Ref. 52.

In this paper we choose the asymmetric dimer model in order to identify the influence of *different basis sets* (plane waves versus Gaussian orbitals) and *different substitution geometries* (slab or supercells versus semi-infinite). All

the self-consistent calculations mentioned above[24,55–57] have been carried out using supercell configurations and plane-wave basis sets. While Refs. 55–57 concentrated on the surface-structure aspect, the self-consistent surface-electronic band structure was only given in Ref. 24. Using these results as a reference for comparisons, we have employed in our calculations the same surface-structure model, the same ionic pseudopotential, and the same exchange-correlation potential as used in Ref. 24.

In a first step we carried out self-consistent Green-function calculations for the semi-infinite crystal with a reconstructed surface using localized Gaussian orbitals. In a second step we carried out supercell calculations—as was done in Ref. 24–but we used the same localized orbitals as in step 1, as opposed to a plane-wave basis set. This allows us to identify basis-set influences on the results. The effect of slab enlargement has also been studied. In our calculations we have used, for comparison sake, the potential of Ref. 10. Exchange and correlation effects were incorporated by a Slater-type LDA potential as in Ref. 24. The reconstruction geometry is described in Ref. 25.

A. Atomic-orbital basis set

In the matrix representations of the Hamiltonian, we use atomic orbitals. The angular part of the localized orbitals is represented by spherical harmonics $Y_{l,m}$ and the radial part is represented by Gaussian functions

$$\phi_\alpha(\mathbf{r})=N_\alpha r^l Y_{l,m}(\vartheta,\varphi)e^{-\gamma r^2} . \tag{52}$$

The superindex α corresponds to l, m, and γ. These indices should not be mixed up with the indices of Sec. II. The factor N_α provides the normalization of each atomic orbital. In our calculations we have used orbitals with $l=0$, 1, and 2 (s,p,d orbitals) and an additional function commonly referred to as the s^* orbital. It is of $l=0$ symmetry, but has a prefactor r^2. These ten orbitals form one shell with the decay constant γ. In bulk calculations, good basis-set convergence is reached by using two shells with $\gamma=0.2$ and 0.6. This corresponds to 20 orbitals per atom. We restrict ourselves to one shell, which is appropriate for the potentials used. Within this shell, however, we allow for *different decay constants* γ_l for each l. These decay constants are determined by minimizing the total energy of the bulk crystal according to the variational principle of density-functional theory. A comparison of calculated bulk band energies at high-symmetry points as calculated using two shells with $\gamma_1=0.2$ and $\gamma_2=0.6$ or one shell with $\gamma_s=0.19$, $\gamma_p=0.18$, $\gamma_d=0.21$, and $\gamma_{s^*}=0.4$ is given in Table I. The results for one shell with one decay constant $\gamma=0.2$ only are also shown for comparison. The agreement between the two-shell calculation and the one-shell calculation with different decay constants for each l is very good, with the exception of the width of the valence band. Nevertheless, a basis set with ten Gaussian orbitals per atom and different decay constants for each l leads to a reasonable description of the bulk. We have employed the same basis set for the calculation of the surface-electronic properties. One might expect that a basis set with modified decay constants for the orbitals at or near the surface could improve the description of the surface band structure. To this end we have included additional orbitals at the outermost layers. This can be done easily within our approach. The resulting changes in the electronic structure of the Si(001)-(2×1) surface have been marginal.

The lattice sums in Eqs. (8a) and (8b) and the corresponding bulk equations have been restricted to fifth-nearest neighbors because of the strong localization of the Gaussian orbitals used. As a consequence, the perturbation matrix $\underline{U}^I$ needs to contain only four layers of silicon atoms (i.e., $m_d=4$) to decouple the two semi-infinite crystals.

C. Symmetry considerations

The numerical effort can be reduced considerably by taking the symmetry of the Si bulk crystal and the symmetry of the (2×1) surface into account. The bulk Green function $\underline{G}^0(\tilde{E},\mathbf{q})$ enters each loop in the self-consistent iteration procedure. Therefore it is expedient to calculate $\underline{G}^0(\tilde{E},\mathbf{q})$ in advance and store it on a disk. Since the matrix elements of $\underline{G}^0$ have to be stored for

TABLE I. Comparison of the band energies of bulk Si as calculated by the self-consistent LCAO method using ten [(a) and (b)], or twenty [(c)] Gaussian orbitals per atom. In calculation (a), the decay constants of the s, p, d, and s^* orbitals have been determined by energy minimization (for details, see text), while in the one-shell basis set (b) the same decay constant $\gamma=0.2$ has been used for all orbital types. Calculation (c) employs a two-shell set of 20 orbitals per atom with $\gamma_1=0.2$ and $\gamma_2=0.6$ for each orbital type.

	(a)	(b)	(c)	Expt.
Γ_1	−12.39	−12.32	−12.83	−12.5[a]
Γ_{25}	0.0	0.0	0.0	0.0
Γ_{15}	2.92	3.02	2.89	3.4[a]
Γ_2	3.55	4.09	3.05	4.15[b]
X_{1v}	−8.63	−8.06	−8.58	
X_{4v}	−3.11	−3.06	−3.14	−2.9[b]
X_{1c}	0.90	1.19	0.92	1.3[c]
L_2	−10.50	−9.90	−10.50	−9.3[a]
L_{1v}	−7.58	−7.31	−7.53	−6.7[a]
L_{3v}	−1.30	−1.32	−1.34	−1.2[a]
L_{1c}	1.60	2.47	1.59	2.04[d]
L_{3c}	3.70	3.80	3.77	3.9[b]
$\Delta_{\min}$	0.74	1.09	0.74	1.17[a]

[a] *Landolt-Börnstein, Zahlenwerte und Funktionen aus Naturwissenschaften und Technik*, edited by K. H. Hellwege, New Series (Springer-Verlag, New York, 1982), Group III, Vol. 17a.
[b] W. E. Spicer and R. C. Eden, in *Proceedings of the Ninth International Conference on the Physics of Semiconductors*, Moscow, 1968, edited by S. M. Ryvkin *et al.* (Nauka, Leningrad, 1968).
[c] M. S. Hybertsen and S. G. Louie, Solid State Commun. **51**, 451 (1984).
[d] R. Hulthén and N. G. Nilsson, Solid State Commun. **18**, 1341 (1976).

many $\tilde{E}$ and $\mathbf{q}$ values, a reduction of both central-processing-unit (CPU) time and of storage space by symmetry considerations is advisable. It is possible to calculate the bulk Green function for the (2×1) surface unit cell from the Green function $\underline{G}^0(\tilde{E},\mathbf{q})$ for the (1×1) unit cell. We note that the $\mathbf{g}_j$ vector sum in Eq. (40) runs only over very few $\mathbf{g}_j$ vectors, fulfilling the required restriction.

For the Si(001)(2×1) surface the sum reduces to two terms,[40] namely $\mathbf{g}_1=\mathbf{0}$ and $\mathbf{g}_2=(\pi/a_0)(1,1,0)$, where a_0 is the bulk lattice constant. The Green function is then given by

$$G^0_{\alpha m\mu,\alpha' m'\mu'}(\tilde{E},\mathbf{q})=\tfrac{1}{2}\sum_{j=1}^{2} e^{i\mathbf{g}_j\cdot(\lambda^0_{m\mu}-\lambda^0_{m'\mu'})}G^{0\,(1\times1)}_{\alpha m,\alpha' m'}(\tilde{E},\mathbf{q}+\mathbf{g}_j)\ , \tag{53}$$

where $G^{0\,(1\times1)}_{\alpha m,\alpha' m'}(\tilde{E},\mathbf{q}+\mathbf{g}_j)$ directly follows from an inspection of Eq. (40).

Here we have introduced the Green function $\underline{G}^{0\,(1\times1)}(\tilde{E},\mathbf{q})$ of the (1×1) unit cell. The factor $\frac{1}{2}$ results from the fact that there is only one basis atom in each two-dimensional (1×1) unit cell. For the same reason we have dropped the indices μ and μ' at $G^{0\,(1\times1)}(\tilde{E},\mathbf{q})$. Thus the bulk Green function of the (2×1)-reconstructed unit cell is a mere superposition of $\underline{G}^{0\,(1\times1)}$ at $\mathbf{q}+\mathbf{g}_1$ and at $\mathbf{q}+\mathbf{g}_2$. This is a simple backfolding effect in $\mathbf{k}$ space, since in real space the unit cell is enlarged. Using this symmetry bisects the necessary storage space since $\underline{G}^0$ is 4 times as large as $G^{0\,(1\times1)}$. In addition, we can make use of cell symmetries of the (1×1) surface. This reduces the $k_\perp$ sums occurring in the evaluation of $\underline{G}^{0\,(1\times1)}$ to positive $\mathbf{k}_\perp$ values. In total, we have used 321 $\mathbf{k}_\perp$ points for the calculations of $\underline{G}^{0\,(1\times1)}$. Further symmetry considerations[58] show that the information contained in the $m\times m$ matrix $\underline{G}^{0\,(1\times1)}$ with $m=1,M$ and $m'=1,M$ can be deduced from a smaller matrix $\underline{G}^0(\tilde{E},\mathbf{q})$ with $m=1,M$ but $m'=1,2$. Therefore, $\underline{G}^{0\,(1\times1)}$ can be stored in a very compressed way.

In the case of the Si(001)-(2×1) surface, the number of independent elements of the Hamiltonian matrix is reduced due to mirror symmetry with respect to the mirror plane spanned by the vectors $(-1,1,0)$ and $(0,0,1)$. Matrix elements of the type

$$H_{\alpha\beta}(\mathbf{a},\mathbf{b})=\int\varphi_\alpha(\mathbf{r}-\mathbf{a})\hat{H}(\mathbf{r})\varphi_\beta(\mathbf{r}-\mathbf{b})d^3r \tag{54}$$

are related to matrix elements $H_{\alpha\beta}(\mathbf{a},\mathbf{b}')$ by

$$H_{\alpha\beta}(\mathbf{a},\mathbf{b})=\sum_{\alpha',\beta'}D_{\alpha\alpha'}H_{\alpha'\beta'}(\mathbf{a},\mathbf{b}')D_{\beta'\beta} \tag{55}$$

if

$$\mathbf{a}=\begin{pmatrix}a_x\\a_y\\a_z\end{pmatrix}\ ,\qquad \mathbf{b}=\begin{pmatrix}a_x+c_x\\a_y+c_y\\b_z\end{pmatrix}\ ,\qquad \mathbf{b}'=\begin{pmatrix}a_x-c_y\\a_y-c_x\\b_z\end{pmatrix}\ . \tag{56}$$

The matrix D describes the transformation of the localized orbitals,

$$\varphi_\alpha(\mathbf{r}')=\sum_{\alpha'}D_{\alpha\alpha'}\varphi_{\alpha'}(\mathbf{r})\ , \tag{57}$$

for $\mathbf{r}=(x,y,z)^T$ transformed to $\mathbf{r}'=(-y,-x,z)^T$.

D. Results and discussion

In this subsection we will discuss the electronic structure of the Si(001)-(2×1) surface in the asymmetric dimer configuration. We compare the results of five different calculations as listed in Table II. Our self-consistent Green-function calculation will be referred to as A. The results of Ihm *et al.*[24] for a ten-layer supercell calculation with a plane-wave basis set is labeled B. Our own self-consistent slab calculations using localized Gaussian orbitals for ten layers in the slabs or 12 layers in the slab are labeled C and D, respectively. Finally, we contrast our calculations with the results of an empirical tight-binding Green-function calculation by Schmeits

TABLE II. Summary of the different self-consistent (SC) calculations for the Si(001)-(2×1) surface, whose results are discussed in comparison in this paper. (PW denotes plane waves.)

	Method	Geometry	Basis set	Reference
A	SC STM	semi-infinite	LCAO	This work
B	SC supercell	10 layers	PW	Ihm *et al.* (Ref. 24)
C	SC slab	10 layers	LCAO	This work
D	SC slab	12 layers	LCAO	This work

et al.[40,59] In the self-consistent calculations A–D the same ionic pseudopotentials and xc potentials have been used. In our SC STM calculation we have taken the rearrangement of the charge density at the topmost four layers ($m_u=4$) into account. The perturbation area has a length of 16.293 Å, which is 3 times the bulk Si lattice constant. The real-space mesh for the charge-density and potential calculation contains $n_{\parallel 1}=11$, $n_{\parallel 2}=21$, and $n_\perp=39$, i.e., 9009, points in total. All **q** integrations have been evaluated using four special points in the irreducible part of the surface Brillouin zone.

The surface band structure following from the Green-function calculation (A) is shown in Fig. 5. The shaded area represents the projected bulk band structure. True bound states have been plotted as solid lines and surface resonances are indicated by dashed lines. We have labeled the various surface states according to their physical origin and character, which can be identified by analyzing atom-, orbital-, and wave-vector-resolved layer densities of states. Together with the corresponding charge-density contours, we obtain a direct correlation between electronic structure and spatial arrangement of the electronic charge density at the surface of a semi-infinite solid.

The surface band structure exhibits a very rich spectrum of surface bands which originate from dangling bonds, backbonds, and dimer bonds. The states S_1-S_5 have backbond character and are mostly *s*-like. The bands B_1-B_5 have backbond character as well, but the related states are strongly *p*-like with small *s* and *d* admixtures. The band D_i stems from the dimer bonds and the D_{up} and D_{down} bands are related to the dangling bonds at the up and down atoms, respectively.

As an example for the information content in the surface Green function and for the analysis of spectral features, we show in Fig. 6 the layer densities of states at the *K* point of the surface Brillouin zone for the first five layers and for a bulk layer in comparison. For clarity's sake, only the energy region from -5 to $+1$ eV is shown. Surface-induced state density peaks can clearly be identified and their spatial localization properties become obvious. The dangling-bond features are most pronounced at the surface layer ($m=1$). The backbond

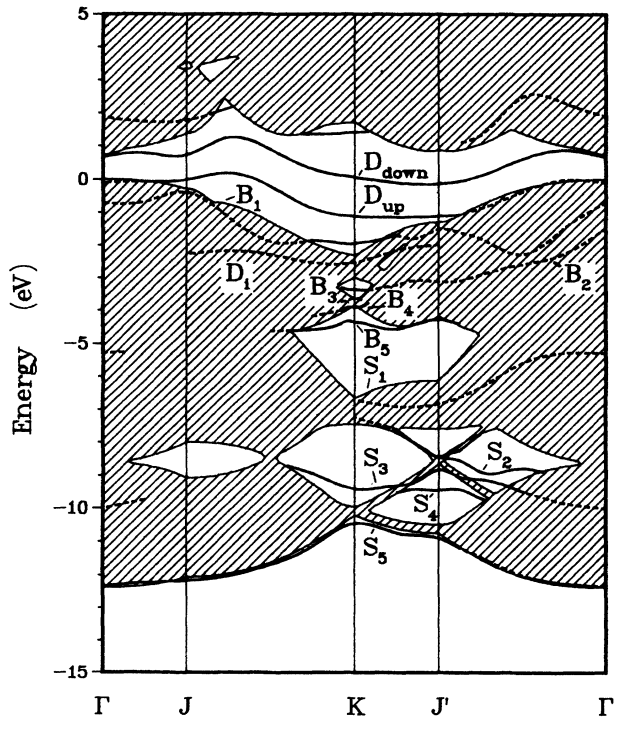


FIG. 5. Self-consistent electronic structure of the Si(001)-(2×1) surface as resulting from our Green-function calculation for the asymmetric dimer model. The bands S_1-S_5 and B_1-B_5 stem from backbond states. D_i is the dimer-bond band, while D_{up} and D_{down} are dangling-bond bands originating from the dangling bonds at the up and down atoms in the surface dimer.

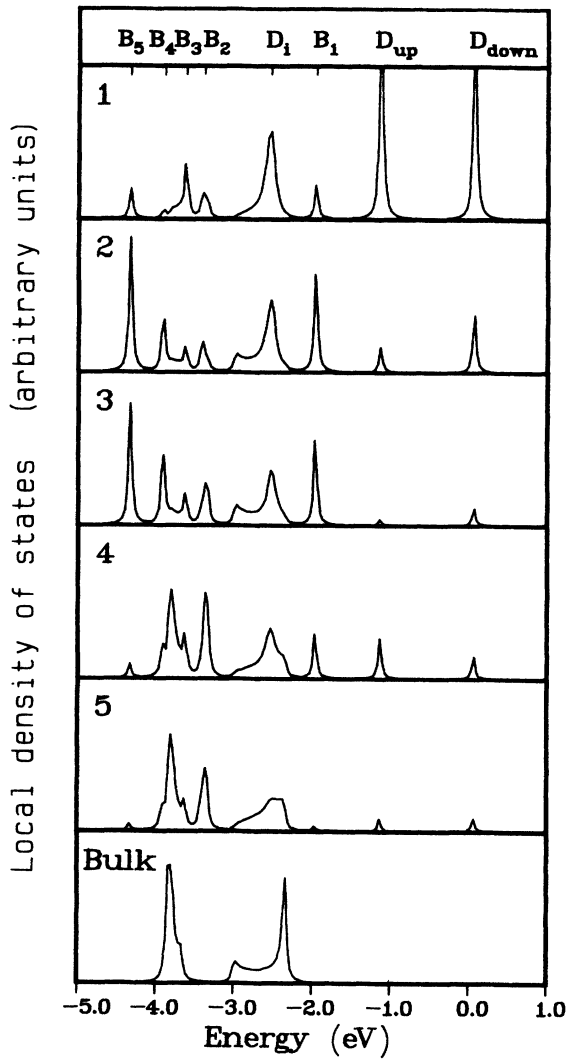


FIG. 6. Wave-vector-resolved layer densities of states at the first five layers of the reconstructed Si(001)-(2×1) surface at the *K* point of the surface Brillouin zone in comparison with a bulk-layer DOS. Note the localization of the dangling bonds at the surface layer, while the backbonds are most pronounced on lower-lying layers.

states are stronger on the second and third layers. The dimer-bond state is a strong resonance which energetically coincides with the bonding p states of the bulk crystal. This is due to the fact that the formation of the surface dimers leads to a strong chemical bond between the up and down atoms in each unit cell.

To further highlight the physical origin of the spectral features at the surface, in Fig. 7 we have decomposed the surface-layer state density (layer 1) of Fig. 6 into its s contribution and the p contributions parallel and perpendicular to the surface. These densities are plotted separately in Fig. 7 for the dimer up atom (solid line) and the down atom (dashed line). The d contribution to the states in this energy region is very small and has not been included in the figure, therefore. The p_z contribution to the LDOS is perpendicular to the surface, while the $p_{1,1,0}$ and $p_{-1,1,0}$ contributions are in the surface plane. The figure clearly reveals that the feature near 0 eV is related to the down atom, while the feature near -1 eV stems from the up atom. The small $p_{-1,1,0}$ components of the LDOS of D_{up} and D_{down} indicate a small tilt angle of the orbital lobes with respect to the [001] direction. The backbonds have only small amplitudes on the surface layer, as mentioned already in the context of Fig. 6.

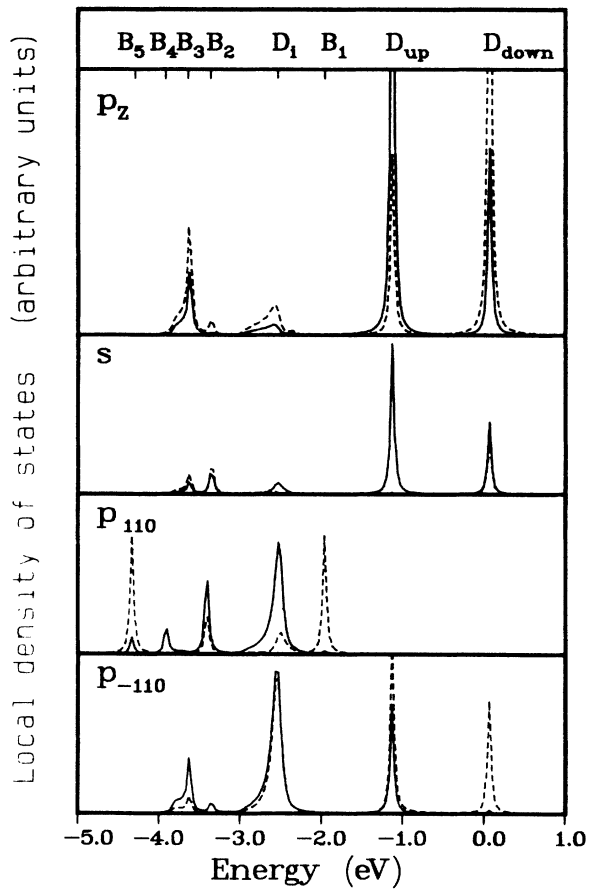


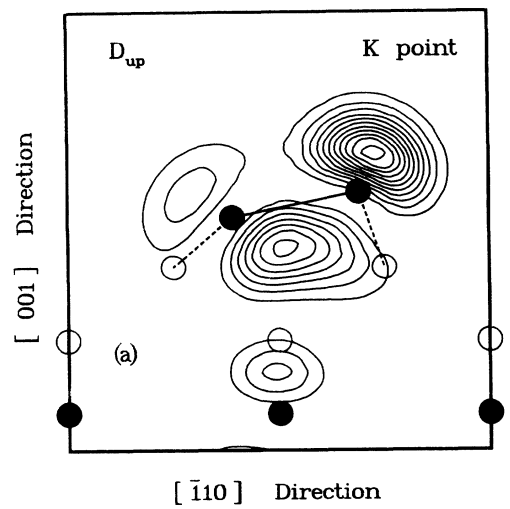
FIG. 7. Wave-vector, orbital-, and atom-resolved surface layer density of states (see layer 1 in Fig. 6 for comparison) at the K point of the Si(001)-(2×1) surface. The solid lines show the LDOS at the up atom and the dashed lines give the contribution of the down atom to the LDOS. The backbond states have only a small amplitude at the surface layer.

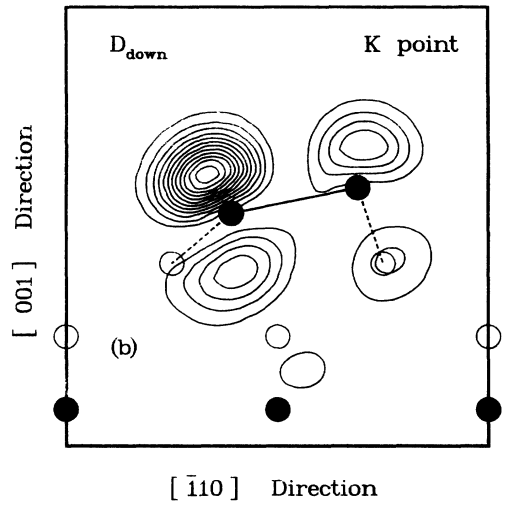
While layer state densities are most useful for analyzing energetic positions of bound states and resonances and for comparisons with surface-spectroscopy data, charge-density contours are still more appropriate to identify the character of a given surface feature. In Figs. 8(a)–8(c) we show the charge density of the predominant surface states at Si(001)-(2×1) plotted for the K point of the surface unit cell. The strong localization of charge density in the dangling bonds at the up atom [Fig. 8(a)] and the down atom [Fig. 8(b)] is evident and confirms the identification as D_{up} and D_{down} states, respectively. Note the strong localization of these features near the surface plane. The dimer bond state D_i [Fig. 8(c)], on the contrary, is strongly resonant with bulk states. Therefore, its charge-density contours are more extended than D_{up} and D_{down}. Nevertheless, the formation of a strong chemical bond between the up and down atoms can clearly be seen.

Due to the change of the chemical environment of the outermost layers, the backbonds of the surface dimers, as well as the bonds between the following layers, are somewhat distorted, giving rise to the backbond resonances B_i. As examples, we show in Figs. 9(a)–9(c) the charge-density contours of the states B_1, B_4, and B_5 at the K point whose layer state densities were given in Figs. 6 and 7, respectively. The drawing plane has been shifted by $a_0\sqrt{2}/4$ in the [110] direction, as compared to Figs. 8(a)–8(c). Therefore, now the atoms of the second and third layers reside in the drawing plane, rather than the dimer atoms and the fourth-layer atoms, as was the case in Figs. 8(a)–8(c). We see that the electrons building up the resonances B_1, B_4, and B_5 are mainly localized at the atoms of the second and third layers. This corresponds to the localization of the predominant layer-state-density peaks in Fig. 6 on the second and third layers for B_1, B_4, and B_5.

The electronic states at the other $\mathbf{q}$ points have been analyzed in the same way. In the energy range from -5 to -13 eV all bands have strong s-like character. The p-like backbonding states B_i are mainly resonant with bulk states. The same holds for the dimer-bond band D_i, as discussed above for the K point, as one example. The situation for the dangling bonds with their energy bands in the gap region or near the edges of the projected bulk bands is quite different. This results from the fact that there is no comparable orbital configuration in the bulk crystal. The dangling bonds of different surface unit cells are second-nearest neighbors in the [110] direction and their distance in the $[\bar{1}10]$ direction is twice as large. Therefore, their interaction along the dangling-bond chains, corresponding to the J–K and J'–Γ directions, leads to a strong dispersion of the bands in these $\mathbf{k}_{\parallel}$ directions. Along the perpendicular direction there is little interaction of the dangling bonds and the bands are therefore rather flat from Γ to J and from K to J'.

We will now turn to a comparison with the results of self-consistent supercell or slab calculations (B)–(D). These calculations yield all the gross features of the surface band structure as well. There are, however, interesting differences which we now address. A comparison





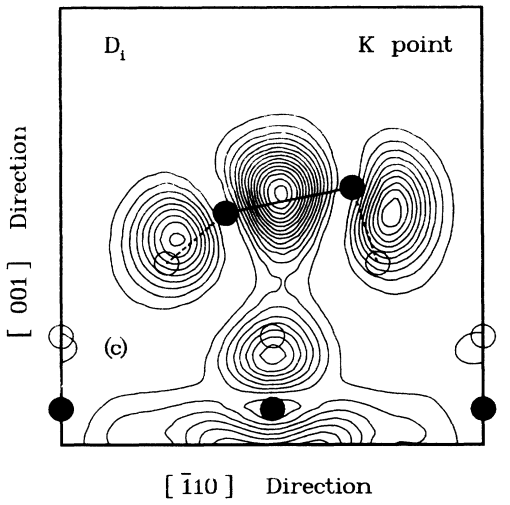
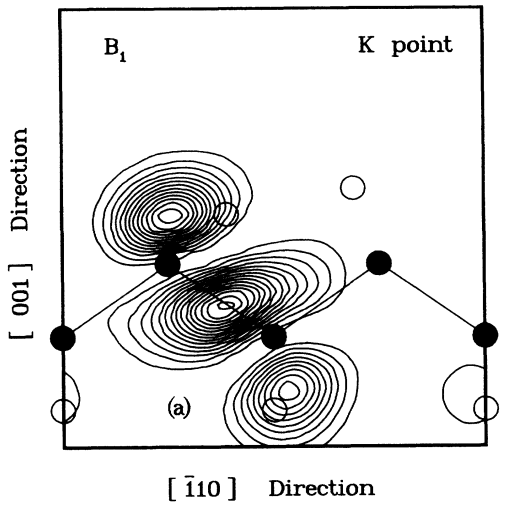
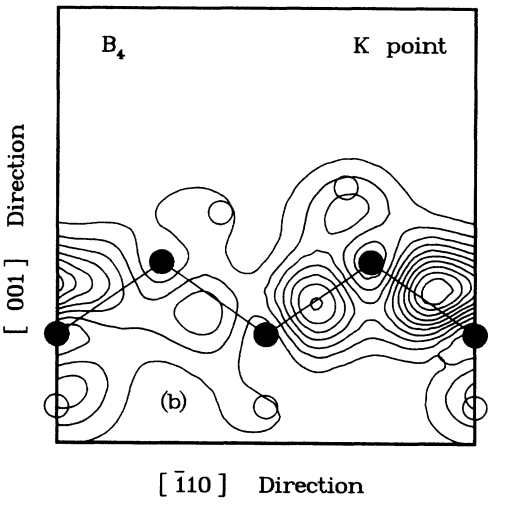


FIG. 8. Charge-density contours of the states (a) D_{up}, (b) D_{down}, and (c) D_i at the K point for Si(001)-(2×1). Atoms represented by solid dots lie within the drawing plane.





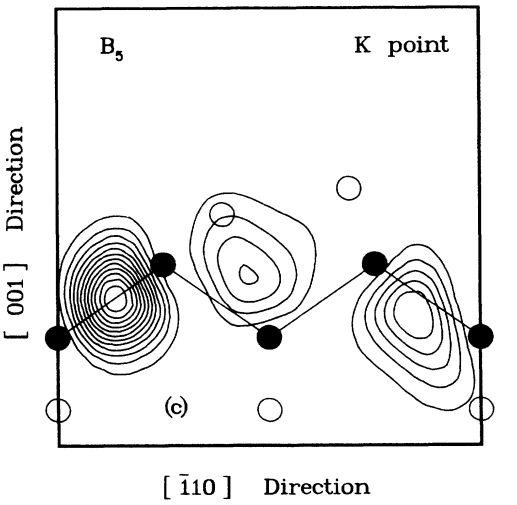


FIG. 9. Charge-density contours of the backbond states (a) B_1, (b) B_4, and (c) B_5 at the K point for Si(001)-(2×1). The drawing plane has been shifted by $a_0\sqrt{2}/4$ in the [110] direction with respect to Fig. 8. Therefore, now the atoms of the second and third layers lie in the drawing plane.

with the surface band structure of the supercell, containing ten Si layers in the unit cell, shows (see Fig. 4 in Ref. 24) that the dangling-bond bands D_{up} and D_{down} result significantly higher in energy from the plane-wave calculation of Ihm *et al.*,[24] as compared to our results in Fig. 5. Opposite this, the backbond state B_1 results at lower energies in calculation B, while it is a localized gap state at the K point in our results. Furthermore, Ihm *et al.* have not identified any bound states in the small "stomach" gap at the K point near -3 eV, opposite our SC STM calculations. Investigation of the origin of these differences in the outcome of the calculations A and B is interesting from a methodological point of view and useful because of the fact that our calculated D_{up} band is in much closer agreement with photoemission data than the band of Ref. 24 (see Fig. 2 in Ref. 22). This deficiency of the supercell result as calculated with a plane-wave basis seems to be a general shortcoming of those calculations since it occurred for C(111)-(2×1),[60] Si(111)-(2×1),[61,62] and Ge(111)-(2×1),[61] as well.

Since the surface-structure model and the underlying ionic and xc potentials are the same in calculations A and B, differences in the results can only stem from the different substitute geometries (supercell versus semi-infinite) or basis sets (plane waves versus localized Gaussians) used. To identify the influence of the substitute geometry as an approximation of the semi-infinite solid, we have carried out slab calculations with ten layers (C) or 12 layers (D) in the unit cell using the same localized basis as in A. The resulting surface band structures are given in Figs. 10 and 11, respectively, and should be compared with our Fig. 5 and with Fig. 4 of Ref. 24. In Figs. 10 and 11 we have plotted the slab bands as resulting from our calculations together with the projected bulk band (hatched area). To be able to superimpose the projected bands in the slab results, we have aligned the energy scales in the different calculations by adjusting the average potential at the slab centers to the average bulk potential. This problem does not occur in Green-function calculations. Comparing Figs. 10 and 11 with Fig. 5, we note that the dangling-bond bands in the slab results are split up by interactions of the surface states across the fairly thin slabs. This splitting, e.g, amounts to 0.4 eV at the Γ point of the ten-layer slab and to 0.1 eV at the Γ point of the 12 layer slab. In general, the splittings of surface bands throughout the surface Brillouin zone are smaller for the 12-layer slab, as is to be expected. This band splitting occurs, of course, also in plane-wave supercell calculations, as we and others have observed. It has become customary, nevertheless, to suppress these artificial splittings and to plot "mean values" of the split bands. Therefore, no splittings are shown in Fig. 4 of Ref. 24.

Comparing Figs. 10 and 11 resulting from calculation C and D, we find a lot of similarities. There is, however, one interesting difference that we want to comment on, i.e., the degeneracy of the D_{up} and D_{down} bands from K to J' in the ten-layer slab result in Fig. 10. Investigating the crystal geometry reveals that in the case of Si(001)-(2×1) slabs containing 10, 14, 18, etc. layers the surface dimer is located in a plane which does not contain the

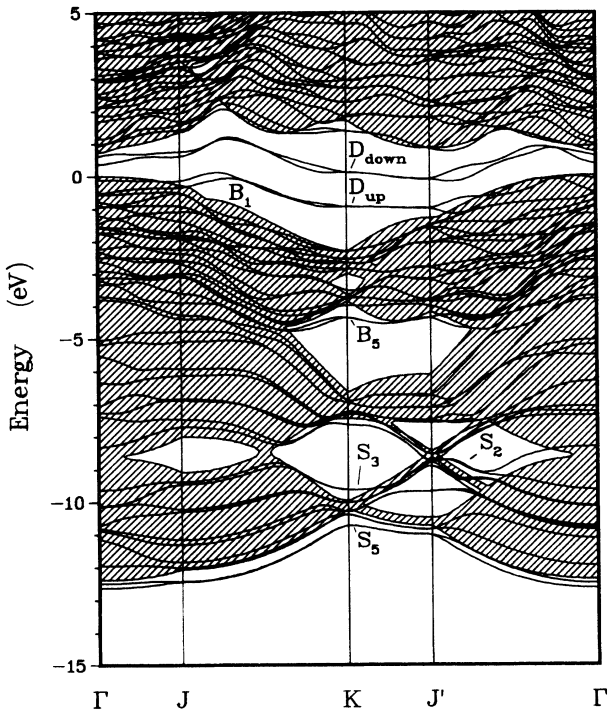


FIG. 10. Self-consistent electronic structure of the reconstructed Si(001)-(2×1) surface as resulting from our ten-layer slab calculation using a localized Gaussian-orbital basis set. The bands have been superimposed by the corresponding projected bulk band structure (hatched areas). Note the twofold degeneracy of all the bands between K and J'.

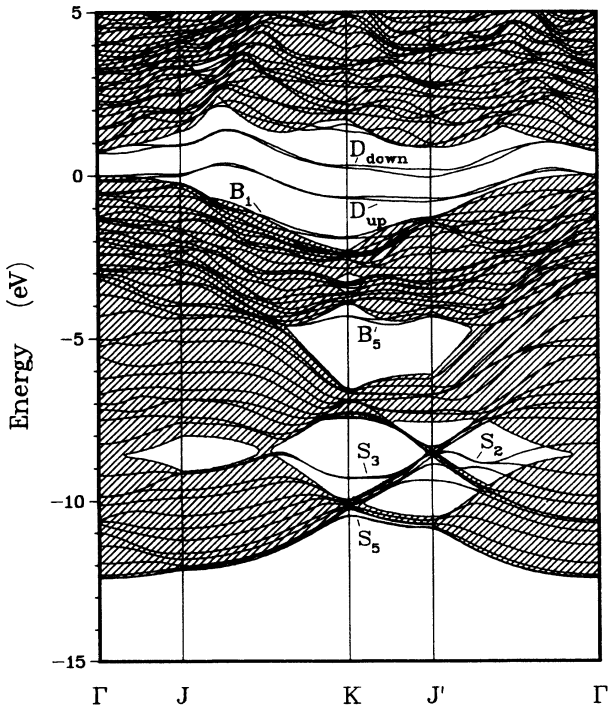


FIG. 11. Same as Fig. 10, but for a 12-layer slab.

center of inversion. According to this glide plane, there are symmetry operations containing pure rotations and nonprimitive translations. The corresponding space group is nonsymmorphic. In this case[63] the representation of the energy bands at J' is two dimensional. Including time-reversal symmetry, this twofold degeneracy of *all* bands pertains for the entire $K-J'$ line. Opposite this, in the case of slabs or supercells with 12, 16, 20, etc. layers of silicon atoms the mirror plane including the surface dimer always contains the center of inversion. All representations of this symmorphic space group are one dimensional and there are no degeneracies of the energy bands by symmetry. In general, surface-state bands without any splittings result from the Green-function calculations for semi-infinite crystals. If we compare the above-discussed "mean values" of the bands in Figs. 10 and 11 with Fig. 5, we find quite good overall agreement. From this result we conclude that it is not the substitute geometry which gives rise to differences in the energy position of pronounced dangling-bond bands.

The second possible cause of discrepancies could be the basis set. The basis-set influence can be identified by comparing our ten-layer result of calculation C in Fig. 10 with the data of Ihm *et al.* in Fig. 4 of Ref. 24. It becomes obvious that the dangling bonds resulting from the plane-wave calculation (B) are lying higher in energy by roughly 0.6 eV as compared to the localized basis calculation (C). In addition, there is a distinct difference in the dispersion of the S_3 band resulting from the two calculations. The basis set used in Ref. 24 includes plane waves up to 2 Ry in kinetic energy only. This basis set leads to an overestimated gap energy $E_g=1.4$ eV in the bulk bands. Increasing the number of plane waves up to an energy cutoff 3.5 Ry shifts the dangling-bond bands down in energy by 0.25 eV with respect to the results of Ihm *et al.*,[24] as was shown by Wolfgarten[64] in our group. From this result and from our localized-orbital calculation (C), we conclude that an improvement of the basis set of calculation C, as compared to B, is responsible for the improvement in the description of the occupied D_{up} band, as compared to experiment. While the small basis set used in Ref. 24 yielded an overestimated gap energy of $E_g=1.4$ eV, we find a value of $E_g=0.7$ eV, which is more typical for self-consistent LDA calculations. We have also checked the basis-set influence on the outcome of our Green-function calculations. When we take only *s*, *p*, and s^* orbitals into account,[58] the projected gap opens up to 1.6 eV and the dangling-bond bands shift up in energy by roughly 0.6 eV as compared to Fig. 5. The dispersion of the dangling-bond bands, on the other hand, is only marginally affected by omission of the *d* orbitals in the basis set.

Summarizing the comparison of the results of the self-consistent calculations A–D, we can state that there is good agreement in the topology of the surface-state bands. There is even quantitative agreement in energy positions of bands related to localized surface states if the computations are carried out with the same level of basis-set convergence using either semi-infinite solids (A) or slabs (C and D). Thus the slab or Green-function techniques are equally well suited to describe localized surface features. When it comes to resonances, however, the Green-function approach is clearly superior. In particular, wave-vector-resolved layer densities of states are described less accurately by slabs or supercells due to the relatively small numbers of layers in the unit cell. This situation is unsatisfactory in view of the fact that high-resolution surface-spectroscopy data are mostly given for the whole energy region of the projected bulk bands, calling for a clear separation of surface and bulk features. Comparing Figs. 10 and 11 with Fig. 5 highlights the better spectral resolution of the Green-function results. But to make our point more transparent, we compare in Fig. 12 the surface-layer density of states at the Γ point for Si(001)-(2×1) as it results from our 12-layer-slab calculation (D) [see Fig. 12(a)] and from our Green-function calculation (A) [see Fig. 12(b)]. For the 12-layer slab,

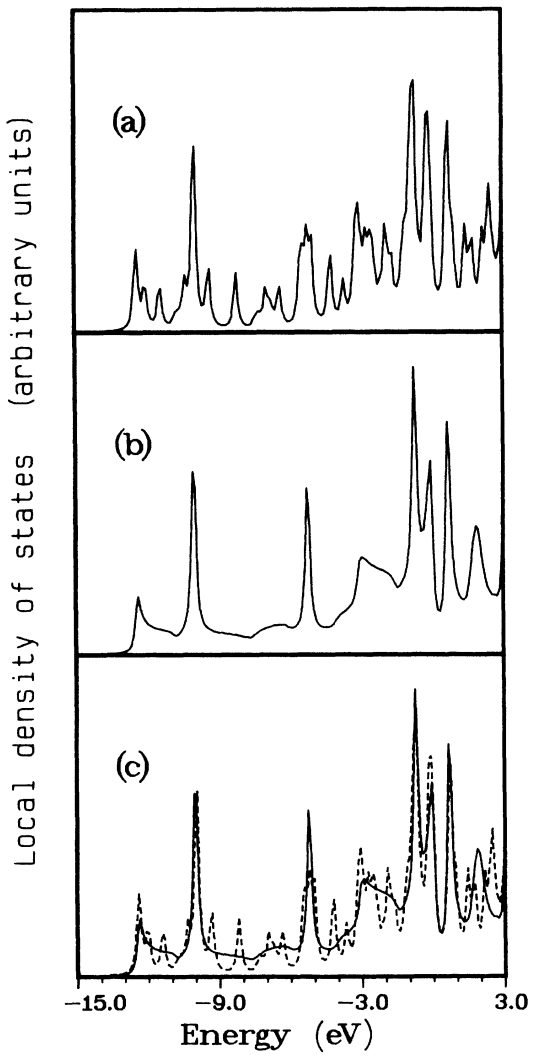


FIG. 12. Comparison of the surface-layer density of states at the zone-center (Γ point) for the Si(001)-(2×1) surface as it results from (a) our 12-layer-slab calculation and (b) our Green-function calculation. The slab spectrum has been created by 0.1-eV Lorentzian broadening of the slab eigenvalues. The direct comparison of (a) and (b) in (c) clearly demonstrates the superiority of the Green-function approach for calculating highly converged layer state densities.

only 48 spin-degenerate states contribute to the spectrum at each $\mathbf{q}$ point in the valence-band energy region. Broadening these states by a Lorentzian of 0.1 eV width yields the spectrum in Fig 12(a). The Green function, on the other hand, yields the true continuous state-density spectrum as plotted in Fig. 12(b). The direct comparison of the two results in Fig. 12(c) clearly demonstrates that the slab spectrum contains a number of spurious peaks which are mere artifacts due to the lack of layer-number convergence in a 12-layer-slab calculation.

In conclusion, our results show that the self-consistent Green-function approach allows for an unambiguous determination of both bound surface states and of surface resonances whose extended wave functions are much more difficult to describe properly by other methods. On the other hand, there is the virtue of slab and supercell calculations that they are more simple both conceptually and computationally.

Calculating the electronic structure of semiconductor surfaces self-consistently from first principles is a rather computer-time-consuming procedure. In many areas of surface physics, e.g., in the investigation of surface phonons, surface phase transitions, or crystal melting, the calculation of the electronic structure for a fixed geometry is only an intermediate step which has to be carried out many times during very elaborate numerical procedures. In such circumstances, realistic empirical tight-binding descriptions of electronic properties with a comparatively low numerical effort can be helpful. The Si(001)-(2×1) surface has been studied previously in our group by empirical tight-binding Green-function calculations as well.[40] Comparing the self-consistent result of Fig. 5 with Fig. 9 of Ref. 40, we find that there is indeed good agreement throughout the valence-band energy region. Considerable deviations occur within the conduction-band region. These can easily be traced back to the limited *s-p* basis set used in the tight-binding calculations.

Finally, we want to comment on the "surface-gap-energy problem." Due to the existence of occupied and empty dangling-bond states in the gap-energy region, the surface gap is typically smaller than the bulk band gap. Since converged LDA calculations underestimate the bulk band gap,[33,34] they are very likely to underestimate the surface band gap as well. There are experimental indications[65,66] that the surface band gap for Si(001)-(2×1) is roughly 0.4 eV. The plane-wave supercell calculations (B) of Ref. 24 yield a small gap of 0.1 eV and our calculations A, C, and D even lead to a metallic surface since the D_{down} band overlaps the D_{up} band to some extent. To understand this difference between calculation B and calculations A, C, and D, the reader should remember that in B the bulk band gap was 1.4 eV (due to the 2-Ry cutoff in the plane-wave basis set), while in our calculations the band-gap energy is 0.7 eV. In any case, the calculations seem to underestimate the surface band gap appreciably. We believe that this deficiency is again a mere artifact of local-density theory and has no physical importance for the surface-structure discussion. A convincing hint regarding this notion is given by recent calculations of Hybertsen and Louie.[32,34] These authors calculated quasiparticle energies of bulk semiconductors[32] and of a surface system[34] in a very encouraging approach based on many-body theory. In the latter application for an As overlayer on Ge(111), they observed a substantial shift of the empty surface states towards higher energies, opening up the surface gap considerably with respect to the LDA result. The same is very likely to occur for an improved treatment of Si(001)-(2×1) when the calculations are extended beyond the local-density approximation. Since the surface-gap energy is related to the energetic position of the D_{down} band which stems mainly from the bulk conduction bands, we expect the surface gap to open up by roughly 0.4 eV in our calculations A, C, and D if the energy spectrum would be calculated more rigorously, going beyond the LDA, so that the bulk gap would open up from our value of 0.7 eV to the correct value of 1.1 eV.

V. CONCLUDING DISCUSSION

A. Summary

In this paper we have presented in detail the Green-function or scattering-theoretical method for self-consistent calculations of the electronic structure of reconstructed semiconductor surfaces within the local-density approximation. Our method describes a surface as a two-dimensionally periodic perturbation of the bulk solid which is highly localized in the surface-perpendicular direction and allows us to treat semi-infinite systems. The principle advantages of this method are the exploitation of (a) the short-range nature of the surface potential, (b) the three-dimensional translational symmetry of the underlying bulk crystal, and (c) the analytic separation between bulk and surface properties. The feasibility and efficiency of the SC STM have been demonstrated by an application to the reconstructed Si(001)-(2×1) surface using ionic pseudopotentials and a Gaussian-orbital basis set. A variety of surface states has been identified and analyzed by wave-vector-, atom-, and orbital-resolved layer densities of states and by the corresponding charge densities. In addition, we have presented the results of self-consistent supercell calculations for Si(001)-(2×1) which were carried out with the same localized Gaussian-orbital basis set as used in the Green-function calculations. A comparison of all the results with the outcome of the plane-wave supercell calculation of Ref. 24 allowed us to trace back differences in the calculated dangling-bond surface bands to differences in the basis sets used. Bound surface-state bands are found to result, in very good agreement from both our supercell or Green-function calculations. The superiority of the Green-function approach for describing surface resonances has been demonstrated.

B. Outlook

In order to be able to compare the new SC STM with current standard methods, we have so far applied the technique for (i) a predetermined structure model with (ii) semiempirical local pseudopotentials on the basis of (iii) the local-density approximation. In this short "outlook," we will indicate how an improvement of the structural

model can be handled within our approach and we will comment on the resulting interrelationship of the above three points.

Structural optimization requires minimization of the ground-state energy of the semi-infinite solid with respect to the coordinates of the atoms. For this purpose it is convenient to calculate the total energy change as compared to the total energy of a semi-infinite solid with an appropriate reference geometry, e.g., the predetermined structure model or the ideal surface. This energy difference[67] consists of the difference of the band-structure energies, the difference of the exchange-correlation energies, the difference of the Coulomb energy of the electrons, and the difference of the Coulomb energy of the ion cores. The atomic coordinates of those two systems differ mainly near the surface. Taking this into account, the band-structure energy difference can be computed directly from the surface Green functions of the two systems. The difference of the exchange-correlation energy caused by the difference of the charge densities near the surface can be calculated using the modified Fourier techniques described in Sec. III A of this work. The same holds for the Coulomb energy of the electrons. It is advantageous to add and subtract "ionic" charges at this stage. The negative term is included in the Coulomb energy of the valence electrons and the positive term is combined with the Coulomb energy of the ion cores. The latter can easily be computed in real space using "ionic" charges of Gaussian distribution.[68] This completes the calculation of the total energy for one atomic configuration. To speed up the minimizing procedure of the ground-state energy, a computation of the forces would be convenient. For the use of advanced iteration techniques, see, e.g., Ref. 69. A method similar to that introduced here has already been used by Feibelman[26] with great success in the calculation of the adsorption energy and the force on adsorbed Al on an Al(001) film.

From the comparison of the SC STM and the slab approach in Sec. IV D, we would expect that *integrated* quantities like the charge density and the band-structure energy are well described also by the slab approach, which is easier to handle. Therefore, it seems very appropriate to use slabs for the structural optimization and to compute the electronic spectrum with high resolution using SC STM afterwards.

For a reliable surface-structure determination, it is mandatory to use the best available, nonlocal *ab initio* pseudopotentials.[70–72] These norm-conserving pseudopotentials have been used widely with great success in calculations of the total energy based either on plane waves[73,74] or localized orbitals[37,68] to predict the lattice constants, compressibilities, and phonon frequencies of periodic crystals. The efficiency of these modern first-principles pseudopotentials for the structure determination of surfaces has been demonstrated by Pandey[75] and by Northrup and Cohen.[61] For reviews, see e.g., Refs. 69 and 76.

Implementation of nonlocal pseudopotentials into our computations is straightforward, leading to matrix equations (7) and (10) of the same structure as in the case of local pseudopotentials. The matrix elements (8) can easily be calculated within the Gaussian basis set. The nonlocal pseudopotentials are no longer "soft-core" potentials, so that at least a double-zeta Gaussian basis is probably necessary for the wave-function expansion, increasing the size of all matrices correspondingly. The first-principles potentials yield excellent ground-state energies in comparison with experimental data, but they give very poor values for direct and indirect gaps, as is known from bulk calculations. Thus for a reasonable description of the electronic spectrum, one would have to go beyond the LDA including self-energy corrections of the type discussed in Refs. 32–34. In general, these self-energy corrections act like complex potentials in the Schrödinger equation (1). The calculation of the electronic states and their lifetime from (1) can be handled by Green functions in a natural way. In our opinion, the present approach is well adapted to these types of investigations. They have to be carried out, though, and they are, of course, more involved than the calculations presented in this paper.

ACKNOWLEDGMENTS

We thank Dr. G. Wolfgarten for numerous helpful discussions and for his contributions in developing the computer code. It is a pleasure to acknowledge many useful discussions with Dr. A. Mazur. This work has partially been supported by the Deutsche Forschungsgemeinschaft (Bonn, Germany) under Contract No. Po-215/2-2 and Po-215/3-1. One of us (P.K.) gratefully acknowledges support from the Graduiertenförderungsprogramm of Northrhine-Westfalia in the early stages of this work.

APPENDIX A: CALCULATION OF $G^I(\tilde{E},\mathbf{q})$

The idealized surface is created from the bulk crystal by the perturbation $\underline{U}^I(\tilde{E},\mathbf{q})$. The computation of the corresponding Green function $\underline{G}^I(\tilde{E},\mathbf{q})$ can be simplified by taking the special form of $\underline{U}^I(\tilde{E},\mathbf{q})$ into account. The size of $\underline{U}^I$ is given by $(m_u+m_d)\times(m_u+m_d)$, as can be seen in Eq. (24). In this appendix we show that the Dyson equation (27) for G^I can be reduced to a similar equation with an $m_d\times m_d$ perturbation matrix only.

The bulk Green function $\underline{G}^0(\tilde{E},\mathbf{q})$ can be expressed formally as

$$\underline{G}^0(\tilde{E},\mathbf{q})=\begin{pmatrix} \underline{G}^0_{AA} & \underline{G}^0_{AB} & \underline{G}^0_{AC} & \underline{G}^0_{AD} \\ \underline{G}^0_{BA} & \underline{G}^0_{BB} & \underline{G}^0_{BC} & \underline{G}^0_{BD} \\ \underline{G}^0_{CA} & \underline{G}^0_{CB} & \underline{G}^0_{CC} & \underline{G}^0_{CD} \\ \underline{G}^0_{DA} & \underline{G}^0_{DB} & \underline{G}^0_{DC} & \underline{G}^0_{DD} \end{pmatrix} \equiv \begin{pmatrix} \underline{G}^0_{LL} & \underline{G}^0_{LR} \\ \underline{G}^0_{RL} & \underline{G}^0_{RR} \end{pmatrix} = \begin{pmatrix} \underline{M}^0_{LL} & \underline{M}^0_{LR} \\ \underline{M}^0_{RL} & \underline{M}^0_{RR} \end{pmatrix}^{-1} . \tag{A1}$$

For clarity's sake, we have compressed subspaces A and B, as defined in Sec. II, to the "left" subspace (L), and

the subspaces C and D correspondingly to the "right" subspace R.

The Green function $\underline{G}^I(\tilde{E},\mathbf{q})$ is defined in Eq. (21) as the inverse of $\underline{M}^I(\tilde{E},\mathbf{q})$, which has the form

$$\underline{M}^I(\tilde{E},\mathbf{q})=\begin{pmatrix}\underline{M}^0_{LL} & \underline{0}\\ \underline{0} & \underline{0}\end{pmatrix} . \tag{A2}$$

as defined in Eq. (20).

At this point we state that $\underline{G}^I$ is given by

$$\underline{G}^I(\tilde{E},\mathbf{q})=\begin{pmatrix}\underline{G}^0_{LL} & \underline{0}\\ \underline{0} & \underline{0}\end{pmatrix} - \begin{pmatrix}\underline{0} & \underline{G}^0_{LR}\\ \underline{0} & \underline{0}\end{pmatrix}\begin{pmatrix}\underline{0} & \underline{0}\\ \underline{0} & \underline{K}_{RR}\end{pmatrix}\begin{pmatrix}\underline{0} & \underline{0}\\ \underline{G}^0_{RL} & \underline{0}\end{pmatrix} , \tag{A3}$$

with

$$\underline{K}_{RR}=\begin{pmatrix}\underline{K}_{CC} & \underline{K}_{CD}\\ \underline{K}_{DC} & \underline{K}_{DD}\end{pmatrix}=\begin{pmatrix}(\underline{G}^0_{CC})^{-1} & \underline{0}\\ \underline{0} & \underline{0}\end{pmatrix} . \tag{A4}$$

Equation (A3) may be written more compactly as

$$\underline{G}^I_{LL}=\underline{G}^0_{LL}-\underline{G}^0_{LR}\underline{K}_{RR}\underline{G}^0_{RL} . \tag{A5}$$

The big advantage of the representation (A3)–(A5) is a reduction of the CPU time by a factor of 8 (or more if m_u is larger than m_d) with respect to the direct solution of Eq. (27) for each $\tilde{E}$ and $\mathbf{q}$. In Eq. (A4) only the $m_d\times m_d$ matrix $\underline{G}^0_{CC}$ needs to be inverted. Calculating the Green function from (A3) to (A5) corresponds to the "atom-removal method" described by Pollmann and Pantelides[19] for the empirical tight-binding version of the STM.

Proof: To prove Eqs. (A3)–(A5), we compute the *bulk* Green function formally from its definition as the inverse of $\underline{M}^0(E,\mathbf{q})$, as defined in Eq. (16),

$$\underline{G}^0(\tilde{E},\mathbf{q})=\begin{pmatrix}\underline{M}^0_{LL} & \underline{M}^0_{LR}\\ \underline{M}^0_{RL} & \underline{M}^0_{RR}\end{pmatrix}^{-1} . \tag{A6}$$

The inversion can easily be done using the following relations, which are valid for regular matrices $\underline{U}$, $\underline{V}$, $\underline{W}$, and $\underline{X}$ with dimensions $(n\times n)$, $(n\times m)$, $(m\times n)$, and $(m\times m)$, respectively,

$$\begin{pmatrix}\underline{U} & \underline{V}\\ \underline{W} & \underline{X}\end{pmatrix}^{-1}=\begin{pmatrix}\underline{Y}^{-1} & -\underline{U}^{-1}\underline{V}\,\underline{Z}^{-1}\\ -\underline{X}^{-1}\underline{W}\,\underline{Y}^{-1} & \underline{Z}^{-1}\end{pmatrix} , \tag{A7a}$$

with

$$\underline{Y}=\underline{U}-\underline{V}\,\underline{X}^{-1}\underline{W} , \tag{A7b}$$

$$\underline{Z}=\underline{X}-\underline{W}\,\underline{U}^{-1}\underline{V} . \tag{A7c}$$

Therefore, the matrices $\underline{G}^0_{LR}$ and $\underline{G}^0_{RL}$ are given by

$$\underline{G}^0_{LR}=-(\underline{M}^0_{LL})^{-1}\underline{M}^0_{LR}\underline{G}^0_{RR} \tag{A8}$$

and

$$\underline{G}^0_{RL}=-(\underline{M}^0_{RR})^{-1}\underline{M}^0_{RL}\underline{G}^0_{LL} . \tag{A9}$$

With these relations, the right-hand side (rhs) of Eq. (A5) can be expressed as

$$\underline{G}^0_{LL}-(\underline{M}^0_{LL})^{-1}\underline{M}^0_{LR}\underline{G}^0_{RR}\underline{K}_{RR}(\underline{M}^0_{RR})^{-1}\underline{M}^0_{RL}\underline{G}^0_{LL} . \tag{A10}$$

To evaluate the underlined term, we use $\underline{M}^0_{LR}$ as given in Eq. (19), yielding

$$\begin{aligned}\underline{M}^0_{LR}\underline{G}^0_{RR}\underline{K}_{RR}&=\begin{pmatrix}\underline{0} & \underline{0}\\ \underline{M}^0_{BC} & \underline{0}\end{pmatrix}\begin{pmatrix}\underline{G}^0_{CC} & \underline{G}^0_{CD}\\ \underline{G}^0_{DC} & \underline{G}^0_{DD}\end{pmatrix}\begin{pmatrix}(\underline{G}^0_{CC})^{-1} & \underline{0}\\ \underline{0} & \underline{0}\end{pmatrix}\\ &=\begin{pmatrix}\underline{0} & \underline{0}\\ \underline{M}^0_{BC} & \underline{0}\end{pmatrix}\begin{pmatrix}\underline{1} & \underline{0}\\ \underline{G}^0_{DC}(\underline{G}^0_{CC})^{-1} & \underline{0}\end{pmatrix}\\ &=\begin{pmatrix}\underline{0} & \underline{0}\\ \underline{M}_{BC} & \underline{0}\end{pmatrix}\\ &=\underline{M}^0_{LR} .\end{aligned} \tag{A11}$$

This leads for the rhs of (A5) to

$$\begin{aligned}&\underline{G}^0_{LL}-(\underline{M}^0_{LL})^{-1}\underline{M}^0_{LR}(\underline{M}^0_{RR})^{-1}\underline{M}^0_{RL}\underline{G}^0_{LL}\\ &\quad=[\underline{1}-(\underline{M}^0_{LL})^{-1}\underline{M}^0_{LR}(\underline{M}^0_{RR})^{-1}\underline{M}^0_{RL}]\underline{G}^0_{LL}\\ &\quad=(\underline{M}^0_{LL})^{-1}[\underline{M}^0_{LL}-\underline{M}^0_{LR}(\underline{M}^0_{RR})^{-1}\underline{M}^0_{RL}]\underline{G}^0_{LL} .\end{aligned} \tag{A12}$$

From Eq. (A7b) it is obvious that the second factor in (A12) is the inverse of $\underline{G}^0_{LL}$. Therefore we have, for the rhs of (A5),

$$(\underline{M}^0_{LL})^{-1}(\underline{G}^0_{LL})^{-1}\underline{G}^0_{LL}=(\underline{M}^0_{LL})^{-1}=\underline{G}^I_{LL} \tag{A13}$$

Q.E.D.

APPENDIX B: ALTERNATIVE COMPUTATION SCHEME FOR THE SURFACE GREEN FUNCTION

In this appendix we will compare the concept for calculating the surface Green function as described in Sec. II with an alternative method we have used previously.[11,12,52,43] In these papers a surface was represented by two semi-infinite solids separated by a small vacuum region. The average potential $\bar{V}_{\text{eff}}(z)$ of this geometry is shown schematically in Fig. 13(a). Contrary to the wave-function expansion used in this paper [see Eq. (6)], the wave function of the "crystal with a vacuum region" was built up by orbitals localized on both the left and right half-crystals and in the vacuum region. The total potential was constructed from the full bulk potential by adding a perturbation potential which is strongly localized perpendicular to the surface. This perturbation potential was determined self-consistently. Within the vacuum region the orbitals were positioned at former bulk-atom sites. The solid dots in Fig. 13(a) indicate the localization planes of these layer orbitals. This approach has the advantage that the bulk crystal directly enters as the reference system. In this case the reference matrices $\underline{H}^R$ and $\underline{S}^R$ in Eqs. (11) and (12) are simply given by $\underline{H}^0$ and $\underline{S}^0$, as defined in Eq. (14) and the separate construction of the idealized surface in an intermediate step is avoided.

However, there is a great disadvantage of this approach related to the dimensions of the matrices involved in solving the Dyson equation for the surface Green func-

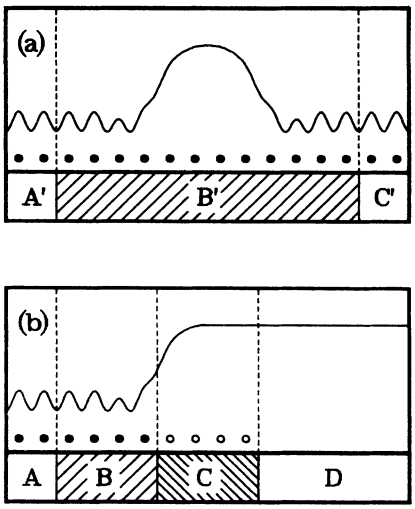


FIG. 13. Schematic plot of the geometry used in our previous STM calculations (a) in Refs. 11, 12, 52, and 53 and in this work, (b). The localization planes of the layer orbitals in the wave function expansion are shown by solid dots. The layer orbitals localized on the planes shown by open circles in (b) are involved in the creation of the idealized surface, i.e., in the decoupling of the two semi-infinite crystals.

tion. Here we have

$$\underline{G}(\tilde{E},\mathbf{q})=\underline{G}^0(\tilde{E},\mathbf{q})+\underline{G}^0(\tilde{E},\mathbf{q})\underline{U}(\tilde{E},\mathbf{q})\underline{G}(\tilde{E},\mathbf{q}) \ . \quad \text{(B1)}$$

If we decompose the vector space in A' and C' for the unperturbed left and right half-crystals and B' for the perturbation region [see Fig. 13(a)], the perturbation matrix has the seemingly simple structure

$$\underline{U}(\tilde{E},\mathbf{q})=\begin{bmatrix}\underline{0} & \underline{0} & \underline{0}\\ \underline{0} & \underline{U}_{B'B'} & \underline{0}\\ \underline{0} & \underline{0} & \underline{0}\end{bmatrix} . \quad \text{(B2)}$$

The subspace B', however, contains the orbitals on m_u layers of the left half-crystal, on m_u layers on the right half-crystal, and on m_d vacuum layers, as illustrated in Fig. 13(a).

Therefore the matrices which need to be inverted to solve the Dyson equation (B1) in analogy to Eqs. (30)–(32) have the dimensions $(2m_u+m_d)\times(2m_u+m_d)$. In the new approach developed in this paper, we have inverted much smaller matrices for all values of $\tilde{E}$ and $\mathbf{q}$. In the new approach, calculation of the Green function of the idealized surface requires inversions of matrices with dimensions $m_d\times m_d$ of subspace C only, as was shown in Appendix A. The calculation of the reconstructed surface Green function is then accomplished in a second step by inversions of matrices with size $m_u\times m_u$ from subspace B. This is illustrated schematically in Fig. 13(b). In our new approach as compared to the earlier method, the computer time needed to calculate the surface Green function $\underline{G}(\tilde{E},\mathbf{q})$ is reduced roughly by a factor of

$$(2m_u+m_d)^3/(m_u^3+m_d^3) \ .$$

For typical values $m_u=4$ and $m_d=4$, this amounts to a factor of 13.5 and thus the new approach saves many CPU hours for a full calculation.[77] The surface-electronic structure resulting from both methods agrees within 0.05 eV for the Si(001)-(2×1) surface discussed in this paper.

[1]P. Hohenberg and W. Kohn, Phys. Rev. **136**, B864 (1964).
[2]W. Kohn and L. J. Sham, Phys. Rev. **140**, A1133 (1965).
[3]W. Kohn and P. Vashishta, in *Theory of the Inhomogeneous Electron Gas,* edited by S. Lundqvist and N. H. March (Plenum, New York, 1983).
[4]See, e.g., V. L. Moruzzi, J. F. Janak, and A. Williams, *Calculated Properties of Metals* (Pergamon, New York, 1978).
[5]See, e.g., M. L. Cohen, in *Computer Based Microscopic Description of the Structure and Properties of Materials,* edited by J. Broughton, W. Krakow, and S. T. Pantelides, MRS Symposia Proceedings Vol. 63 (Materials Research Society, Pittsburgh, 1986), pp. 107; R. M. Martin, in *Festkörperprobleme,* edited by P. Grosse (Vieweg, Braunschweig, 1985), Vol. 25, p. 3.
[6]See, e.g., *The Structure of Surfaces,* edited by M. A. van Hove and S. Y. Tong, Vol. 2 of *Springer Series in Surface Science* (Springer, Berlin, 1985).
[7]J. A. Appelbaum and D. R. Hamann, Phys. Rev. Lett. **32**, 225 (1974).
[8]J. A. Appelbaum, G. A. Baraff, and D. R. Hamann, Phys. Rev. Lett. **35**, 729 (1975); Phys. Rev. B **11**, 3822 (1975); **12**, 5749 (1975); **14**, 588 (1976).
[9]D. J. Chadi, Phys. Rev. B **18**, 1800 (1978).
[10]M. Schlüter, J. R. Chelikowsky, S. G. Louie, and M. L. Cohen, Phys. Rev. B **12**, 4200 (1975).
[11]G. Wolfgarten, P. Krüger, and J. Pollmann, Solid State Commun. **54**, 839 (1985).
[12]P. Krüger, A. Mazur, J. Pollmann, and G. Wolfgarten, Phys. Rev. Lett. **57**, 1468 (1986).
[13]G. F. Koster and J. C. Slater, Phys. Rev. **95**, 1167 (1954).
[14]G. F. Koster, Phys. Rev. **95**, 1436 (1954).
[15]G. A. Baraff and M. Schlüter, Phys. Rev. Lett. **41**, 892 (1978); Phys. Rev. B **19**, 4965 (1979).
[16]J. Bernholc, N. O. Lipari, and S. T. Pantelides, Phys. Rev. Lett. **41**, 895 (1978); Phys. Rev. B **21**, 3545 (1980); **26**, 570 (1982).
[17]R. Zeller and P. H. Dederichs, Phys. Rev. Lett. **42**, 1713 (1979); Phys. Rev. B **28**, 5462 (1983).
[18]M. Scheffler, S. T. Pantelides, N. O. Lipari, and J. Bernholc, Phys. Rev. Lett. **47**, 413 (1981).
[19]J. Pollmann and S. T. Pantelides, Phys. Rev. B **18**, 5524 (1978).
[20]J. Pollmann, in *Festkörperprobleme (Advances in Solid State Physics),* edited by J. Treusch (Vieweg, Braunschweig, 1980), Vol. 20, p. 117.
[21]See, e.g., D. H. Lee and J. D. Joannopoulos, J. Vac. Sci. Technol. **21**, 351 (1982); Y. C. Chang, Phys. Rev. B **26**, 4400 (1982); R. P. Beres, R. E. Allen, and J. D. Dow, *ibid.* **26**, 5702 (1982); C. Mailhiot, C. B. Duke, and Y. C. Chang, *ibid.* **30**, 1109 (1984); see also Refs. 40, 45, 52, and 59.
[22]J. Pollmann, P. Krüger, and A. Mazur, J. Vac. Sci. Technol. B **5**, 945 (1987).
[23]A. R. Williams, P. J. Feibelman, and N. D. Lang, Phys. Rev.

B **26**, 5433 (1982).
[24]J. Ihm, M. L. Cohen, and D. J. Chadi, Phys. Rev. B **21**, 4592 (1980).
[25]D. J. Chadi, Phys. Rev. Lett. **43**, 43 (1979).
[26]P. J. Feibelman, Phys. Rev. Lett. **54**, 2627 (1985); Phys. Rev. B **35**, 2626 (1987); Phys. Rev. Lett. **58**, 2766 (1987).
[27]G. A. Benesh and J. E. Inglesfield, J. Phys. C **17**, 1595 (1984); **19**, L539 (1986); Phys. Rev. B **37**, 6682 (1988); J. E. Inglesfield, Prog. Surf. Sci. **25**, 57 (1987).
[28]G. A. Baraff and M. Schlüter, J. Phys. C **19**, 4383 (1986).
[29]J. Perdew and M. Levy, Phys. Rev. Lett. **51**, 1884 (1983).
[30]L. J. Sham and M. Schlüter, Phys. Rev. Lett. **51**, 1888 (1983).
[31]W. Pickett, Comments Solid State Physics **12**, 1 (1985).
[32]M. S. Hybertsen and S. G. Louie, Phys. Rev. Lett. **55**, 1418 (1985); Phys. Rev. B **34**, 5390 (1986).
[33]R. W. Godby, M. Schlüter, and L. J. Sham, Phys. Rev. Lett. **56**, 2415 (1986).
[34]M. S. Hybertsen and S. G. Louie, Phys. Rev. Lett. **58**, 1551 (1987).
[35]See E. A. Hylleraas, *Mathematical and Theoretical Physics* (Wiley, New York, 1970).
[36]See the references in Ref. 37.
[37]C. T. Chan, D. Vanderbilt, and S. G. Louie, Phys. Rev. B **33**, 2455 (1986).
[38]U. Lindefelt and A. Zunger, Phys. Rev. B **24**, 5913 (1981).
[39]The critical number of layers m_u taken into account has to be chosen such that the neglect of differences between M and M^0 outside this region does not affect the resulting electronic properties within the desired accuracy. This has to be checked by convergence studies with increasing m_u.
[40]M. Schmeits, A. Mazur, and J. Pollmann, Phys. Rev. B **27**, 5012 (1983).
[41]S. Freeman, Phys. Rev. B **2**, 3272 (1970).
[42]R. E. Allen, Phys. Rev. B **19**, 917 (1979); **20**, 1454 (1979).
[43]Y. C. Chang, Phys. Rev. B **25**, 605 (1982); **26**, 4400 (1982).
[44]Y. C. Chang and J. N. Schulman, Phys. Rev. B **25**, 3975 (1982).
[45]P. Krüger and J. Pollmann, Phys. Rev. B **30**, 3406 (1984).
[46]J. W. Cooley and J. W. Tukey, Math. Comput. **19**, 297 (1965).
[47]J. Ihm, A. Zunger, and M. L. Cohen, J. Phys. C **12**, 4409 (1979).
[48]R. Zeller, J. Deutz, and P. H. Dederichs, Solid State Commun. **44**, 993 (1982).
[49]P. H. Dederichs and R. Zeller, Phys. Rev. B **28**, 5462 (1983).
[50]D. G. Anderson, J. Assoc. Comput. Machin. **12**, 547 (1964).
[51]G. Bachelet (private communication).
[52]J. Pollmann, R. Kalla, P. Krüger, A. Mazur, and G. Wolfgarten, Appl. Phys. A **41**, 21 (1986).
[53]J. Pollmann, P. Krüger, A. Mazur, and G. Wolfgarten, in *Proceedings of the 18th International Conference on the Physics of Semiconductors*, edited by Olof Engström (World Scientific, Singapore, 1986), p. 81.
[54]D. E. Eastman, J. Vac. Sci. Technol. **17**, 492 (1980).
[55]M. T. Yin and M. L. Cohen, Phys. Rev. B **24**, 2303 (1981).
[56]J. E. Northrup, Phys. Rev. Lett. **54**, 815 (1986).
[57]K. C. Pandey, in *Proceedings of the 17th International Conference on the Physics of Semiconductors*, edited by D. J. Chadi and W. A. Harrison (Springer, New York, 1985).
[58]P. Krüger, Ph.D. thesis, Universität Dortmund, 1986.
[59]A. Mazur and J. Pollmann, Phys. Rev. B **26**, 7086 (1982).
[60]D. Vanderbilt and S. G. Louie, Phys. Rev. B **29**, 7099 (1984).
[61]J. E. Northrup and M. L. Cohen, Phys. Rev. Lett. **49**, 1349 (1982); Phys. Rev. B **27**, 6553 (1983).
[62]K. C. Pandey, Phys. Rev. Lett. **49**, 223 (1982).
[63]Ch. Kittel, *Quantum Theory of Solids* (Wiley, New York, 1963), p. 213.
[64]G. Wolfgarten (private communication).
[65]Y. J. Chabal, S. B. Christmann, E. E. Chaban, and M. T. Yin, J. Vac. Sci. Technol. A **1**, 1241 (1983).
[66]H. H. Farrell, F. Stucki, J. Anderson, D. J. Frankel, G. J. Lapeyre, and M. Levinson, Phys. Rev. B **30**, 721 (1984).
[67]P. Krüger and J. Pollmann (unpublished).
[68]G. B. Bachelet, H. S. Greenside, G. A. Baraff, and M. Schlüter, Phys. Rev. B **24**, 4745 (1981).
[69]G. P. Srivastava and D. Weaire, Adv. Phys. **36**, 463 (1987).
[70]D. R. Hamann, M. Schlüter, and C. Chiang, Phys. Rev. Lett. **43**, 1494 (1979).
[71]G. Kerker, J. Phys. C **13**, L189 (1980).
[72]G. B. Bachelet, D. R. Hamann, and M. Schlüter, Phys. Rev. B **26**, 4199 (1982).
[73]M. T. Yin and M. L. Cohen, Phys. Rev. Lett. **45**, 1004 (1980); Phys. Rev. B **24**, 6121 (1981); **25**, 7403 (1982); **26**, 3259 (1982); **26**, 5668 (1982); **29**, 6996 (1984).
[74]R. M. Martin, in *Electronic Structure, Dynamics, and Quantum Structural Properties of Condensed Matter*, edited by J. T. Devreese and P. van Camp, NATO Advanced Study Institute Series (Plenum, New York, 1985).
[75]K. C. Pandey, Phys. Rev. Lett. **49**, 223 (1982).
[76]J. Ihm, Rep. Prog. Phys. **51**, 105 (1988).
[77]The calculations in Ref. 52 have been carried out with $m_u = 3$.