Total energies of diamond (111) surface reconstructions by a linear combination of atomic orbitals method

David Vanderbilt and Steven G. Louie
Department of Physics, University of California, Berkeley, California 94720
and Lawrence Berkeley Laboratory, Berkeley, California 94720
(Received 15 May 1984)

An *ab initio* linear combination of atomic orbitals approach to local-density theory, capable of handling complex structural geometries, is presented. It incorporates a self-consistent treatment of interatomic charge transfer, which allows an accurate calculation of total energies. The method is applied to study a variety of possible 1×1 and reconstructed 2×1 models of the diamond (111) surface. Among the many models suggested, only the Pandey π-bonded chain model has a lower energy than that of the 1×1 surface. A minimum-energy structure is obtained for this model after extensive consideration of relaxations. No dimerization of the surface chain is found to occur.

I. INTRODUCTION

The $(2\times1)/(2\times2)$ reconstruction of the diamond (111) surface[1] has recently attracted considerable experimental attention.[2–5] This reconstruction appears upon annealing the (hydrogenated[6,7]) 1×1 surface at $\sim1000\,^{\circ}$C, and is presumably the thermodynamically stable structure in the absence of H. A major motivation for these studies is the desire to make sense of the remarkable variety of surface reconstructions which occur for the tetrahedral elements C, Si, and Ge.[8] A possible common denominator may be the occurrence of a similar 2×1 reconstruction on all three elemental surfaces. While clear 2×1 low-energy electron diffraction (LEED) patterns are observed for Si(111) and Ge(111) surfaces, LEED cannot distinguish between a true 2×2 or disordered domains of 2×1 for the diamond (111) surface.[1] However, the similarity of the angle-resolved ultraviolet photoemission (ARUPS) results for C (Ref. 3), Si (Ref. 9), and Ge (Ref. 10) suggests that a common 2×1 structure may be responsible.

Pandey's proposed π-bonded chain model[11] has attracted much attention as a possible candidate for this structure. A perspective view of this structure is shown in Fig. 1(b), along with a view of the ideal 1×1 structure in Fig. 1(a) for comparison. For Si and Ge, energy-minimization calculations[12] identify the Pandey chain structure as the lowest in energy of all tested, and the calculated dispersion of the occupied surface bands is in good agreement with the ARUPS data.[9,10] Also, ion backscattering[13] and optical absorption[14] measurements appear to support this identification. However, contrary indications from LEED (Ref. 15) and, most recently, photoemission[16] experiments have ensured a continued controversy over this assignment.

For the diamond surface, no energy-minimization calculations have previously been done. A comparison of ARUPS results[3] with the calculated energy dispersion of occupied surface states yields some indirect evidence for the Pandey π-bonded chain model,[11,17] possibly with some dimerization along the chain.[11] However, discrepancies in the location and dispersion of the surface state persist, and the model remains controversial. Among some of the alternative models which have been proposed are the Haneman buckled model,[18] the Seiwatz single chain model,[19] and the Chadi π-bonded molecule model.[20] The latter two are shown in perspective views in Figs. 1(c) and 1(d).

In this paper, we report direct energy-minimization calculations for a variety of 1×1 and reconstructed 2×1 models of the C(111) surface. Because of the localized nature of the carbon wave functions, a first-principles linear combination of atomic orbitals (LCAO) approach has been used to calculate total energies in the local-density approximation (LDA). The method is a generalization of

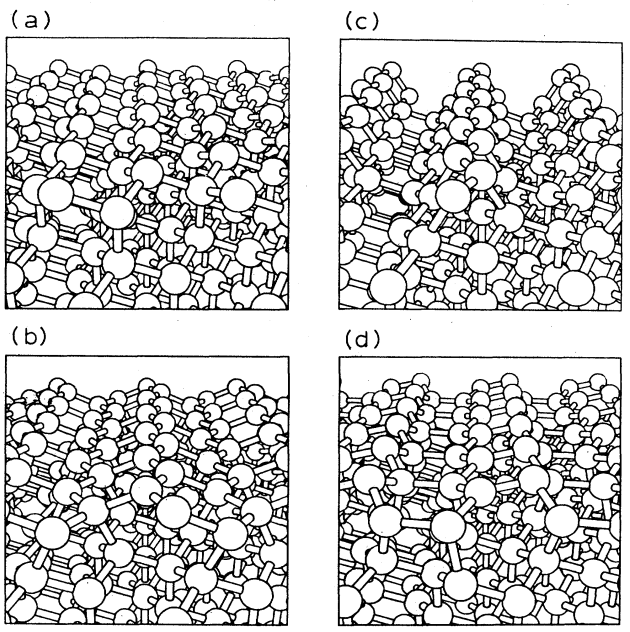

FIG. 1. Perspective views of the structures of (a) the ideal 1×1, (b) the Pandey π-bonded chain, (c) the Chadi π-bonded molecule, and (d) the Seiwatz single chain models.

the first-principles LCAO approach of Chelikowsky and Louie[21] to cases (e.g., surfaces, defects, or heteropolar materials) for which charge transfer between inequivalent atoms must be treated self-consistently.

We have calculated the total energy for a wide variety of surface models, and for those with low total energies we have carried out extensive tests of relaxations which might lower the structural energy. The unrelaxed Pandey π-bonded chain model is found to have a lower energy than any other 2×1 model tested, about equal to that of the ideal 1×1 surface. Relaxation lowers the energy substantially, so that the relaxed Pandey model is ~0.3 eV lower in energy than the relaxed 1×1 surface. Dimerization of the Pandey chain model is always found to *raise* the energy of the system.

The plan of the paper is as follows. In Sec. II we present the details of the general theoretical approach. A discussion of one aspect of the method, namely a modified version of the Broyden convergence accelerator,[22] is deferred to a separate appendix because of its quite general utility. The results for diamond surfaces follow in Sec. III. The relationship of these results to experiment and to other theoretical work is discussed in Sec. IV, and a summary is given in Sec. V.

II. METHOD

A. Charge transfer

1. Self-consistent formulation

The method is a generalization of the first-principles LCAO approach of Chelikowsky and Louie (CL),[21] developed for bulk systems, to the case of systems containing inequivalent atoms. Surfaces, defects, and heteropolar materials are examples of such systems. Because of the inequivalence, interatomic charge transfer must be taken into account and treated self-consistently.

In both the CL and present approaches, the strategy is to make a good approximation, expressed as a sum of Gaussians on the atomic sites, to the LDA potential $V(\vec{r})$. (The basis functions are s-, p-, and d-type Gaussian orbitals, so that all required three-center integrals can then be evaluated analytically.) First-order errors in the potential $V(\vec{r})$ lead to first-order errors in $\rho(\vec{r})$, but the variational property of the LDA total energy ensures that errors in $E_{\rm tot}$ will be second-order small. Thus the method will be capable of giving good structural energies if a good approximation $V(\vec{r})$ can be found.

In the CL approach, the central approximation is

$$V_0(\vec{r})=\sum_{\vec{R},\vec{\tau}} V_{\rm eff}(\vec{r}-\vec{R}-\vec{\tau})\ , \tag{1}$$

where the effective potential $V_{\rm eff}(\vec{r})$ is independent of site $\vec{\tau}$ and is determined for a given element as follows. First, the free (pseudo) LDA atom charge density $\rho_{\rm atom}$ in the appropriate bulk-like configuration (e.g., sp^3 for carbon) is determined. Next, a reference structure, typically the equilibrium bulk structure, is chosen, and $\rho_{\rm atom}$ is superposed to give a first approximation to the crystal charge density:

$$\rho^0_{\rm cryst}(\vec{r})=\sum_{\vec{R},\vec{\tau}} \rho_{\rm atom}(\vec{r}-\vec{R}-\vec{\tau})\ . \tag{2}$$

The crystal potential $V_{\rm LDA}[\rho^0_{\rm cryst}]$ is then evaluated and fit to a sum of spherically symmetric potentials $V_{\rm eff}$ in the manner of Eq. (1); since the ion pseudopotentials and Hartree potentials superpose linearly, we can write

$$V_{\rm eff}=V_{\rm ion}+V_H[\rho_{\rm atom}]+V^{\rm xc}_{\rm eff}\ , \tag{3}$$

where the problem has reduced to fitting

$$V_{\rm xc}[\rho^0_{\rm cryst}](\vec{r})=\sum_{\vec{R},\vec{\tau}} V^{\rm xc}_{\rm eff}(\vec{r}-\vec{R}-\vec{\tau})\ . \tag{4}$$

The fit is done in $\vec{G}$ space, with $V^{\rm xc}_{\rm eff}$ constrained to be a sum of Gaussians. Once $V^{\rm xc}_{\rm eff}$ has been determined, $V_{\rm eff}$ of Eq. (3) is constructed and is itself fitted to a sum of Gaussians. Here $V_{\rm eff}$ is taken to be local, but in general we use nonlocal pseudopotentials $V_{\rm ion}$, with the nonlocal parts fit to Gaussians as well. These and all other Gaussian fits are carried out using a novel Monte Carlo simulated annealing approach to functional fitting.[23]

For a new test structure, $V_0(\vec{r})$ is constructed from Eq. (1) using the new $\vec{R}$'s and $\vec{\tau}$'s, the Schrödinger equation is solved in the LCAO basis, and the output density $\rho^{\rm out}_{\rm cryst}$ is generated. The latter is a much better approximation to the true $\rho_{\rm cryst}$ than the $\rho^0_{\rm cryst}$ formed as in Eq. (2), and the total energy is given by $E_{\rm tot}[\rho^{\rm out}_{\rm cryst};V_{\rm ext}]$, where the external potential is

$$V_{\rm ext}(\vec{r})=\sum_{\vec{R},\vec{\tau}} V_{\rm ion}(\vec{r}-\vec{R}-\vec{\tau})\ . \tag{5}$$

It is important to note that $E_{\rm tot}$ has no explicit dependence upon $\rho^0_{\rm cryst}$ or V_0. This procedure has been shown to give an excellent description of bulk properties, e.g., the lattice constant and zone center phonon frequencies are given to within 1% for diamond.[21]

However, the use of the same $V_{\rm eff}$ on every site $\vec{\tau}$, as in Eq. (1), is no longer valid when the atoms in the unit cell are inequivalent. For example, since the tail of $V_{\rm eff}$ still has some amplitude at the first neighbor sites, a surface atom might have a shallower potential than a bulk atom with more neighbors. This would lead to an unphysical charge transfer off the surface atom in $\rho^{\rm out}_{\rm cryst}$, and a corresponding error in $E_{\rm tot}$. The failure to model the interatomic charge transfer correctly gives rise to errors in the low-q ($q\le G^{\rm min}_{\rm bulk}$) Fourier components of $\rho^{\rm out}_{\rm cryst}(\vec{q})$; these errors are serious because of Coulomb terms $\propto q^{-2}$ in $E_{\rm tot}$. Thus, while the error in $E_{\rm tot}$ is still formally second order in the error in ρ, it is no longer small enough to ignore. The presence of long-range Coulomb interactions requires us to treat the interatomic charge transfer self-consistently.

We begin a given iteration of this self-consistent process with a guess at the electron population Q_τ on each site $\vec{\tau}$, which may differ from the mean value $Q_0=Z$. In analogy with Eq. (2), we approximate $\rho_{\rm cryst}$ by

$$\rho'_{\rm cryst}(\vec{r})=\sum_{\vec{R},\vec{\tau}} \frac{Q_\tau}{Q_0}\rho_{\rm atom}(\vec{r}-\vec{R}-\vec{\tau})\ . \tag{6}$$

The density $\rho'_{\rm cryst}$ now embodies the correct interatomic charge transfer, and has the correct low-q Fourier components, if the Q_τ are correct. The next step is to approximate $V_{\rm LDA}[\rho'_{\rm cryst}]$ by a sum of Gaussians as follows:

$$V(\vec{r})=V_0(\vec{r})+\sum_{\vec{R},\vec{\tau}} c_\tau g(\vec{r}-\vec{R}-\vec{\tau}) . \tag{7}$$

The new correction term on the right-hand side (RHS) is a sum of broad Gaussians, $g(r)=\exp(-\alpha r^2)$, with α chosen as $\sim d_1^{-2}$ (d_1 the first neighbor distance). The coefficients c_τ are chosen in a manner to be described shortly. The correction term contains almost no high Fourier components ($q \gtrsim G^{\rm min}_{\rm bulk}$); its role is to correct the low-q components of V, and hence of $\rho^{\rm out}_{\rm cryst}$. Once $V(\vec{r})$ is obtained, the Schrödinger equation is solved in the LCAO basis, and a Mulliken population analysis[24] is carried out to determine a new set of charges Q_τ. The procedure is then iterated until the input and output Q_τ's are equal, and finally $\rho^{\rm out}_{\rm cryst}$ and then $E_{\rm tot}[\rho^{\rm out}_{\rm cryst};V_{\rm ext}]$ are calculated.

We return to the determination of the coefficients c_τ in Eq. (7). They should be chosen so that V of Eq. (7) has the same low-G Fourier components as $V_{\rm LDA}[\rho'_{\rm cryst}]$. Note that V_0 already contains the correct description of $V_{\rm ion}$; the remainder can be decomposed into a Hartree correction $c_\tau{}^H$ and an exchange-correlation correction $c_\tau{}^{\rm xc}$. We consider these in turn.

The required Hartree correction is

$$\sum_{\vec{R},\vec{\tau}} c_\tau{}^H g(\vec{r}-\vec{R}-\vec{\tau})=V_H[\rho'_{\rm cryst}-\rho^0_{\rm cryst}](\vec{r}-\vec{R}-\vec{\tau}) . \tag{8}$$

Since both sides of Eq. (8) are small ($\rho'_{\rm cryst}\approx\rho^0_{\rm cryst}$) and smooth (high-$q$ contributions to V_H are damped as q^{-2}), it suffices to specify that the functions match at the atom centers $\vec{\tau}$:

$$\sum_{\vec{R},\vec{\tau}'} c^H_{\tau'} g(\vec{\tau}-\vec{R}-\vec{\tau}')=V_H[\rho'_{\rm cryst}-\rho^0_{\rm cryst}](\vec{\tau}-\vec{R}-\vec{\tau}') . \tag{9}$$

The function on the left of Eq. (8) will then interpolate throughout the unit cell in a manner almost identical to the function on the right. Making use of Eqs. (2) and (6) and rewriting as a vector equation, we obtain

$$\overleftrightarrow{g}\cdot\vec{c}^{\,H}=\overleftrightarrow{U}^H\cdot\Delta\vec{Q} , \tag{10}$$

where the notation $\vec{c}$ denotes the N-dimensional vector $(c_{\tau_1},c_{\tau_2},\ldots)$ of dimension N (number of atoms per unit cell). We have defined

$$g_{\tau\tau'}=\sum_{\vec{R}} g(\vec{\tau}-\vec{R}-\vec{\tau}') , \tag{11}$$

$$U^H_{\tau\tau'}=\frac{1}{Q_0}\sum_{\vec{R}} V_H[\rho_{\rm atom}](\vec{\tau}-\vec{R}-\vec{\tau}') , \tag{12}$$

and

$$\Delta Q_\tau=Q_\tau-Q_0 . \tag{13}$$

The solution for $\vec{c}^{\,H}$ is then given by multiplying the RHS of Eq. (10) by $\overleftrightarrow{g}^{-1}$.

The correction to the exchange-correlation potential, given by the vector $\vec{c}^{\,\rm xc}$ of coefficients $c_\tau^{\rm xc}$, is relatively less important because $V_{\rm xc}$ is a weak function of ρ ($V_{\rm xc}\propto\sim\rho^{1/3}$), and because it is local (i.e., no q^{-2} divergence). However, we include it in a manner analogous to the way the Hartree contribution is treated.

We want

$$\sum_{\vec{R},\vec{\tau}} c_\tau{}^{\rm xc} g(\vec{r}-\vec{R}-\vec{\tau})\approx V_{\rm xc}[\rho'_{\rm cryst}] - \sum_{\vec{R},\vec{\tau}} V^{\rm xc}_{\rm eff}(\vec{r}-\vec{R}-\vec{\tau}) . \tag{14}$$

The RHS is rewritten as $\Delta V^{(1)}_{\rm xc}+\Delta V^{(2)}_{\rm xc}$, where

$$\Delta V^{(1)}_{\rm xc}=V_{\rm xc}[\rho^0_{\rm cryst}]-\sum_{\vec{R},\vec{\tau}} V^{\rm xc}_{\rm eff}(\vec{r}-\vec{R}-\vec{\tau}) \tag{15}$$

is independent of $\Delta\vec{Q}$, so that $\Delta V^{(1)}_{\rm xc}(\vec{r})$ is calculated once and for all for a given structure. The second term,

$$\Delta V^{(2)}_{\rm xc}=V_{\rm xc}[\rho'_{\rm cryst}]-V_{\rm xc}[\rho^0_{\rm cryst}] , \tag{16}$$

is linearized as

$$\Delta V^{(2)}_{\rm xc}(\vec{r})\approx\frac{\partial V_{\rm xc}[(Q_\tau/Q_0)\rho_{\rm atom}(\vec{r}-\vec{\tau})]}{\partial Q_\tau}\cdot\Delta Q_\tau \tag{17}$$

for $\vec{r}$ in the neighborhood of $\vec{\tau}$. Unfortunately this function is not smooth in this neighborhood, so we carry out an average using $\rho_{\rm atom}$ as a weighting factor:

$$\Delta V^{(2)}_{\rm xc}(\vec{\tau})=U^{\rm xc}\Delta Q_\tau , \tag{18}$$

$$U^{\rm xc}=\frac{1}{Q_0}\int d^3r\,\rho_{\rm atom}(\vec{r})\frac{\partial V_{\rm xc}[(Q/Q_0)\rho_{\rm atom}(r)]}{\partial Q} . \tag{19}$$

In practice the exchange-correlation corrections [Eqs. (15) and (18)] are only $\sim$2–3 % of the Hartree contribution, Eq. (9). The approximations made here in the treatment of $V_{\rm xc}$ are justifiable in this light.

Collecting all the terms, we have

$$\vec{c}=\overleftrightarrow{g}^{-1}\cdot(\Delta\vec{V}^{(1)}_{\rm xc}+\overleftrightarrow{U}\cdot\Delta\vec{Q}) , \tag{20}$$

where

$$U_{\tau\tau'}=U^H_{\tau\tau'}+\delta_{\tau\tau'}U^{\rm xc} . \tag{21}$$

The U matrix is calculated once and for all for a given structure; the exchange-correlation piece just involves a radial integral, and the Hartree piece is done as a $\vec{G}$-space sum with the $\vec{G}=\vec{0}$ term eliminated to avoid Coulomb divergences. (For surface slab calculations, the slabs are considered to be repeated periodically with a large lattice constant in the z direction, and the change in the Hartree potential is by definition zero in the center of the vacuum region.) $\Delta V^{(1)}_{\rm xc}$ and g are convergent real-space sums. Thus once $\Delta\vec{Q}$ is recalculated from the Mulliken analysis on a given iteration, it is numerically trivial to find $\vec{c}$ from Eq. (20).

We emphasize that the present approach treats the long-range nature of the Coulomb interaction in an essentially exact manner. For example, the self-consistent charge transfer at the surface gives rise to a dipole layer and therefore a potential shift which penetrates infinitely deeply into the bulk; the work function adjusts itself self-consistently, and is a meaningful output of the calculation.

2. Convergence

A side effect of the presence of long-range Coulomb interactions is the fact that the self-consistent iteration process tends to be highly unstable unless heavily damped, because of long wavelength charge oscillations in the unit cell. This problem has been eliminated by using a local version of the dielectric matrix method[25] in which the susceptibility matrix is modeled empirically, and the damping on subsequent iterations is improved using a new version of Broyden's method.[22] We briefly discuss this approach.

Let $\vec{Q}^{(m)}$ be the vector of input Q_τ's for iteration m, $\vec{V}^{(m)}$ the vector of V_τ's giving the potential at the center of site $\vec{\tau}$, and $\vec{Q}_{\rm out}^{(m)}$ the vector of output Q_τ's given by the Mulliken population analysis. We want to choose an effective damping matrix $\overleftrightarrow{\alpha}$ such that

$$\vec{Q}^{(m+1)}=\vec{Q}^{(m)}+\overleftrightarrow{\alpha}\cdot(\vec{Q}_{\rm out}^{(m)}-\vec{Q}^{(m)}) \tag{22}$$

converges quickly to the self-consistent value. Let us define

$$\vec{F}(\vec{Q}^{(m)})=\vec{Q}_{\rm out}^{(m)}-\vec{Q}^{(m)} \tag{23}$$

and expand $\vec{F}$ in linear order as

$$\vec{F}(\vec{Q})=\vec{F}(\vec{Q}^{(m)})-\overleftrightarrow{J}\cdot(\vec{Q}-\vec{Q}^{(m)}) \tag{24}$$

to define the Jacobian matrix $\overleftrightarrow{J}$. The self-consistency criterion is $\vec{F}(\vec{Q}^{(m+1)})=0$, so we choose

$$\vec{Q}^{(m+1)}=\vec{Q}^{(m)}+\overleftrightarrow{J}^{-1}\cdot\vec{F}(\vec{Q}^{(m)})\ , \tag{25}$$

which is just Eq. (22) with $\overleftrightarrow{\alpha}=\overleftrightarrow{J}^{-1}$. The problem thus reduces to finding a good approximation to the Jacobian $\overleftrightarrow{J}$.

Now note that $\overleftrightarrow{J}$ is closely related to the dielectric function. To see this, we follow through one more iteration in linear order. First,

$$\vec{V}^{(m+1)}\approx\vec{V}^{(m)}+\overleftrightarrow{U}\cdot(\vec{Q}^{(m+1)}-\vec{Q}^{(m)})\ , \tag{26}$$

where the matrix $U_{\tau\tau'}$, already determined as discussed in the previous subsection, describes the (unscreened) shift in potential on site $\vec{\tau}$ due to a change in charge on site $\vec{\tau}'$. The output charge is then

$$\vec{Q}_{\rm out}^{(m+1)}\approx\vec{Q}_{\rm out}^{(m)}+\overleftrightarrow{\chi}\cdot(\vec{V}^{(m+1)}-\vec{V}^{(m)})\ , \tag{27}$$

where the unknown electric susceptibility matrix $\chi_{\tau\tau'}$ describes the change in charge on site $\vec{\tau}$ due to the (unscreened) shift in potential on site $\vec{\tau}'$. Combining Eqs. (23), (24), (26), and (27), we obtain

$$\overleftrightarrow{J}=\overleftrightarrow{1}-\overleftrightarrow{\chi}\cdot\overleftrightarrow{U}\ . \tag{28}$$

Thus, $\overleftrightarrow{J}^{-1}$ is the charge dielectric response function, which describes the charge response to an external charge perturbation. Similarly, $\overleftrightarrow{1}-\overleftrightarrow{U}\cdot\overleftrightarrow{\chi}$ is just the dielectric matrix, whose inverse describes the potential response to an external potential perturbation. All this is in close analogy to the plane-wave case,[25] the difference being that a local basis is used here.

To get an initial guess at $\overleftrightarrow{J}$, we use as an empirical model for the susceptibility matrix

$$\chi_{\tau\tau'}=\beta\left[-2\delta_{\tau\tau'}+\frac{1}{m_{\tau'}}\theta_{\tau\tau'}+\frac{1}{N}\right]\ , \tag{29}$$

where β is a constant with dimensions of electrons/Ry, N is the number of atoms per unit cell, $m_{\tau'}$ is the number of first neighbors of $\vec{\tau}'$, and $\theta_{\tau\tau'}$ counts the number of lattice vectors $\vec{R}$ for which $\vec{R}+\vec{\tau}$ is a first neighbor of $\vec{\tau}'$. In words, the model assumes that if the potential on site $\vec{\tau}'$ is raised, half of the charge flowing off is shared equally among first neighbors, the other half being shared equally throughout the unit cell. This model appears to work well for diamond (with $\beta = 3$ electrons/Ry) in most cases, and is probably a reasonable starting point for most insulators and semiconductors.

However, this model for χ seriously underestimates the charge transfer for certain difficult cases, for example, the transfer between dangling bonds in the 2×1 Haneman buckling model.[18] For this reason, the Jacobian is improved on each iteration, starting from $\overleftrightarrow{J}_0$ given by Eq. (28), using a modification of Broyden's method.[22] The modified method is described in the Appendix, and is equally applicable to convergence acceleration in plane-wave calculations or in force-driven structural energy minimization. The advantage of the modified method is that it incorporates information from all previous iterations on an equal footing when updating $\overleftrightarrow{J}$, and converges significantly faster when the initial $\overleftrightarrow{J}_0$ is poor. In practice we find that $|\vec{F}|$ is reduced by a factor of $\sim$ 3—5 per iteration after 2—3 iterations in typical cases, or after $\sim$ 5—6 iterations in the worst cases.

3. Charge density and total energy

During the iterative process, only the Mulliken populations Q_τ are needed to represent the charge configuration. Once self-consistency has been achieved, however, the full $\rho_{\rm cryst}^{\rm out}$ is needed for the calculation of $E_{\rm tot}$. For surface slab calculations, we have found a particularly efficient way to do this. The outer loops are over $\vec{k}=\vec{k}_{||}$ and $z=r_\perp$. For each z, we (i) calculate analytically $\phi_i^{(\vec{k})}(\vec{G}_{||},z)$, the set of basis orbitals expressed in a mixed representation; (ii) fast Fourier transform (FFT) these to

$\phi_i^{(\vec{k})}(\vec{r}_{||},z)$; and (iii) loop over bands n, summing over i to get $\psi_n^{(\vec{k})}(\vec{r}_{||},z)$, and squaring and accumulating these in $\rho(\vec{r}_{||},z)$. At the end, another FFT gives $\rho(\vec{G}_{||},z)$ for a *single* isolated slab. We now repeat the slabs in the z direction and FFT to $\rho(\vec{G}_{||},G_z)=\rho(\vec{G})$ to prepare for the calculation of $E_{\rm tot}$. All arrays with indices $\vec{G}_{||}$ or $\vec{G}$ are carried up to a cutoff $G_{\rm max}$. We have tested convergence with respect to $G_{\rm max}$ and settled on $G_{\rm max}=8$ a.u.$^{-1}$ for the present calculations.

The total energy per atom may be expressed as

$$E_{\rm tot}=\frac{1}{N}\sum_i \epsilon_i + E_{\rm cc} + \Omega_a \sum_{\vec{G}} \rho^{\rm out}_{\rm cryst}(-\vec{G})[\epsilon^{\rm out}_H(\vec{G})+\epsilon^{\rm out}_{\rm xc}(\vec{G})-V^{\rm in}_H(\vec{G})-V^{\rm in}_{\rm xc}(\vec{G})-V^{\rm in}_{\rm sc}(\vec{G})]\ , \tag{30}$$

where

$$\epsilon^{\rm out}_H \equiv \tfrac{1}{2} V_H[\rho^{\rm out}_{\rm cryst}]\ , \tag{31a}$$

$$\epsilon^{\rm out}_{\rm xc} \equiv \epsilon_{\rm xc}[\rho^{\rm out}_{\rm cryst}]\ , \tag{31b}$$

$$\epsilon^{\rm in}_H \equiv \tfrac{1}{2} V_H[\rho^{0}_{\rm cryst}]\ , \tag{31c}$$

$$V^{\rm in}_{\rm xc}(\vec{r}) \equiv \sum_{\vec{R},\vec{\tau}} V^{\rm xc}_{\rm eff}(\vec{r}-\vec{R}-\vec{\tau})\ , \tag{31d}$$

$$V^{\rm in}_{\rm sc}(\vec{r}) \equiv \sum_{\vec{R},\vec{\tau}} c_\tau g(\vec{r}-\vec{R}-\vec{\tau})\ , \tag{31e}$$

and $E_{\rm cc}$ is the core-core Coulomb repulsion. Note that $V^{\rm xc}_{\rm eff}$ is here properly defined as

$$V^{\rm xc}_{\rm eff} \equiv V^{\rm fit}_{\rm eff} - V_{\rm ion} - V_H[\rho_{\rm atom}] \tag{32}$$

so that any fitting errors in going from $V_{\rm eff}$ to $V^{\rm fit}_{\rm eff}$ give canceling contributions in $\Sigma\epsilon_i$ and $V^{\rm xc}_{\rm eff}$. The terms $\epsilon^{\rm out}_H(0)$, $V^{\rm in}_H(0)$, and $E_{\rm cc}$ are the ones containing divergent Coulomb terms, but these can be replaced[26] to give

$$E_{\rm tot}=\frac{1}{N}\sum_i \epsilon_i + \Omega_a \sum_{\vec{G}} \rho^{\rm out}_{\rm cryst}(-\vec{G})[\epsilon^{\rm out}_{\rm xc}(\vec{G})-V^{\rm in}_{\rm xc}-V^{\rm in}_{\rm sc}(\vec{G})] + \Omega_a \sum_{\vec{G}\neq\vec{0}} \rho^{\rm out}_{\rm cryst}(-\vec{G})[\epsilon^{\rm out}_H(\vec{G})-V^{\rm in}_H(\vec{G})] + \alpha_{1H}Z + \gamma_{\rm Ewald}\ . \tag{33}$$

The terms have been grouped so that

$$\Omega_a \alpha_{1H} = \int d^3r \left[\frac{Z}{r} - V_H[\rho_{\rm atom}](r)\right] \tag{34}$$

depends only on $\rho_{\rm atom}$ and is readily calculated analytically for $\rho_{\rm atom}$ expressed as a sum of Gaussians.

4. Tests

We have carried out several tests to check that the method described in this section leads to reasonable total energies.

First, we have checked that the self-consistent formulation leads to the correct charge transfers for the ideal 1×1 surface. We did this by varying the input charges Q_τ on the first few surface layers by hand, while searching for an output total-energy minimum. The total energy obtained by the self-consistent procedure agreed to within $\sim$0.015 eV/surface-atom with that obtained from this explicit variational calculation, and the charge transfers were very similar.

Second, we have tested the sensitivity of $E_{\rm tot}$ to the degree of convergence. In general, we find that when the maximum difference $F_{\rm max}$ between the input and undamped output charges is less than 0.01 electron, the total energy is converged to within $\sim$0.005 eV/surface-atom. The convergence criterion $F_{\rm max}<0.01$ electron was generally adopted, except that for the calculation of the dimerization energy of the Pandey chain model, $F_{\rm max}<0.0025$ electron was used for greater precision.

Third, we have calculated the total energy of a free (pseudo) atom in a supercell configuration, using a three-decay-constant set (3.5, 0.935, and 0.25 a.u.$^{-2}$) and a four-decay-constant set (3.5, 1.070, 0.327, and 0.1 a.u.$^{-2}$). The results were -144.51 and -145.58 eV for the three-decay and four-decay cases, respectively, compared to the corresponding exact LDA result -145.88 eV. We have adopted the three-decay set, for reasons of computational efficiency. We would argue that the error, which is due mostly to the incomplete representation of the carbon wave functions by the Gaussian basis orbitals, will remain roughly constant for modest structural changes such as those involved in surface reconstructions.

Finally, we have checked that the slabs are far enough apart, when they are repeated in the z direction for the construction of $\rho(\vec{G})$, so that $E_{\rm tot}$ is independent of the spacing. We find that this condition is satisfied if we use a lattice constant of 33.9 and 30.4 a.u. for all reconstructed 12-layer and 10-layer slabs, respectively. These values were adopted.

B. Calculation of the matrix elements

A convenient feature of the present approach is that all of the necessary three-center integrals can be calculated

and stored once and for all for a given structure. Then, as the loops over iterations and $\vec{k}$ points are carried out, the construction of the overlap matrix $S^{(\vec{k})}$ and Hamiltonian matrix $H^{(\vec{k})}$ requires a negligible amount of computer time.

The matrix elements which must be stored are

$$s_{ij}^{\vec{R}} = \langle \phi_i(\vec{r}+\vec{R}) \mid \phi_j(\vec{r}) \rangle \ , \tag{35a}$$

$$h_{ij}^{\vec{R}} = \langle \phi_i(\vec{r}+\vec{R}) \mid T \mid \phi_j(\vec{r}) \rangle + \sum_{\vec{R}',\vec{\tau}'} \langle \phi_i(\vec{r}+\vec{R}) \mid V_{\text{eff}}(\vec{r}-\vec{R}'-\vec{\tau}') \mid \phi_j(\vec{r}) \rangle \ , \tag{35b}$$

and

$$u_{ij\tau}^{\vec{R}} = \sum_{\vec{R}'} \langle \phi_i(\vec{r}+\vec{R}) \mid g(\vec{r}-\vec{R}'-\vec{\tau}) \mid \phi_j(\vec{r}) \rangle \ . \tag{35c}$$

Here i and j run over all the Gaussian orbitals centered in one unit cell. Then $S^{(\vec{k})}$ and $H^{(\vec{k})}$ are constructed according to

$$S_{ij}^{(\vec{k})} = \sum_{\vec{R}} e^{i\vec{k}\cdot\vec{R}} s_{ij}^{\vec{R}} \tag{36a}$$

and

$$H_{ij}^{(\vec{k})} = \sum_{\vec{R}} e^{i\vec{k}\cdot\vec{R}} \left[h_{ij}^{\vec{R}} + \sum_{\vec{\tau}} c_\tau u_{ij\tau}^{\vec{R}} \right] . \tag{36b}$$

The real-space sums in Eqs. (35)—(36) are infinite but strongly convergent. In practice the sums are cut off when the distance between any two centers exceeds a cutoff R_{cut}. We have tested convergence with respect to R_{cut}, and have found that a value of 11 a.u. is safe for carbon. For the narrow Gaussians with large decay constants, this cutoff is unnecessarily conservative, and a second cutoff is imposed to speed the calculation. Before a given two- or three-center integral of Gaussians is calculated, its magnitude is estimated by discarding from the integrand the polynomial factors (for p- and d-type orbitals) and angular momentum projection operators (for non-local potential contributions). If this "bare" integral does not exceed a cutoff η_{cut}, it is discarded from the sum. Again, we have tested convergence with respect to η_{cut}, and have settled on a safe value of $\eta_{\text{cut}} = e^{-20}$.

Despite the efficiencies introduced by these two cutoffs, the initial generation of the matrix elements is still an expensive part of the calculation, requiring $\sim 50\%$ of the total computer time. However, it should be noted that the computer time for this portion scales only linearly with cell size, so that for larger cells it becomes relatively less expensive.

C. Solution of the secular equation

Once the overlap and Hamiltonian matrices have been constructed according to Eq. (36), it remains to solve the secular equation

$$(H^{(\vec{k})} - \epsilon_{\vec{k}} S^{(\vec{k})}) \psi_{\vec{k}} = 0 \ . \tag{37}$$

While the standard Choleski decomposition could be used to solve Eq. (37), we have taken a slightly different approach which provides greater numerical stability when the basis orbitals get strongly overcomplete. This problem tends to arise when one of the orbital decay constants becomes too small, and $S^{(\vec{k})}$ becomes nearly singular. The finite cutoffs R_{cut} and η_{cut}, and the finite machine precision, can then give rise to errors in $S^{(\vec{k})}$ so that its near-zero eigenvalues go negative, making Choleski decomposition impossible. Even before they go negative, the ill-behaved near-zero eigenvalues can give rise to spurious results for some of the $\epsilon_{\vec{k}}$. While this problem does not arise for the three-decay set used here for carbon, it does arise for some other cases we have tested, and we have therefore implemented the following corrective scheme.

First, the $n \times n$ matrix S is diagonalized, and the "well-behaved" eigenvectors, those whose eigenvalues are positive and exceed some s_{cut}, are identified. (We use $s_{\text{cut}} = 10^{-3} s_{\max}$.) Let m be the number of these, $m \leq n$, and we are particularly interested in the case $m < n$. Now let U be the $n \times m$ matrix whose columns are just the well-behaved eigenvectors. Then $U^\dagger U = I$ is the $(m \times m)$ identity, and $UU^\dagger$ acts as an $(n \times n)$ projection operator onto the well-behaved space.

The strategy now is to project everything onto the well-behaved space. Thus we define the reduced $m \times m$ matrices

$$S_R \equiv U^\dagger S U \tag{38a}$$

and

$$H_R \equiv U^\dagger H U \ , \tag{38b}$$

and define ψ as a linear combination of the well-behaved eigenvectors of S:

$$\psi \equiv U \psi_R \ . \tag{39}$$

Now inserting the projection operator $UU^\dagger$ into Eq. (37) and multiplying on the left by $U^\dagger$, we obtain

$$(H_R - \epsilon S_R) \psi_R = 0 \ . \tag{40}$$

Note that S_R is a positive definite diagonal matrix whose diagonal elements are just the eigenvalues $s_i > s_{\text{cut}}$; thus it is trivial to construct $S_R^{-1/2}$. Then with the further definitions

$$H' \equiv S_R^{-1/2} H_R S_R^{-1/2} \tag{41}$$

and

$$\psi_R \equiv S_R^{-1/2} \psi' \tag{42}$$

we finally obtain the secular equation

$$(H' - \epsilon) \psi' = 0 \ . \tag{43}$$

The energy eigenvalues are obtained directly from the solution of Eq. (43), and the eigenvectors are obtained with the help of the back-transforms (42) and (39).

The use of this scheme increases the computer time by

~30—40%. However, all these operations are easily vectorizable, so that the solution of the secular equation occupies only ~35% of our total CRAY computer time even though several iterations must typically be performed.

III. RESULTS

We have implemented the method outlined in the previous section on diamond (111) surface models consisting of a single infinite slab with both upper and lower surfaces reconstructed. For the 1×1 models the slab was 12 atomic layers thick (12 atoms/cell), while for the 2×1 models it was 10 atomic layers thick (20 atoms/cell), except for the Seiwatz chain model,[19] which was 8 layers thick (16 atoms/cell). All models were arranged to have inversion symmetry about a point in the center of the slab. As in the bulk calculations of CL,[21] three decay constants were used for each s- or p-type Gaussian orbital in the basis. A Hamann-Schlüter-Chiang pseudopotential[27] is used in conjunction with Hedin-Lundqvist exchange correlation.[28] Sets of 7, 8, and 16 $\vec{k}$ points in the irreducible zone were used for 1×1, mirror-symmetric 2×1, and other 2×1 models, respectively.

The results are summarized in Table I. In addition to the ideal 1×1 model, which is used as a zero of energy,[29] four topologically distinct 2×1 models have been tested. These are the Haneman buckling model,[18] the Pandey π-bonded chain model,[11] the Chadi π-bonded molecule model,[20] and the Seiwatz single chain model.[19] We will discuss these in turn, and then address the question of relaxations in detail.

Buckling of the ideal 1×1 surface is found to raise the energy. This is in agreement with energy-minimization calculations for Si and Ge,[12] and the highly localized nature of the carbon orbitals would be expected to make such a buckling distortion, with its associated charge transfer, even less likely in diamond. Moreover, such a buckled model would be expected to have a metallic surface band entirely within the band gap,[11,17] in contradiction with ARUPS data.[3]

TABLE I. Calculated total energies of diamond (111) 1×1 and 2×1 surface reconstruction models.

Surface model	Energy (eV/surface-atom)
Ideal 1×1	0.00
Relaxed 1×1	−0.37
Buckled[a] ($\Delta z=\pm0.26$ Å)	0.35
Chadi π-bonded molecule[b]	0.28
Seiwatz single chain[c]	1.30
Ideal Pandey π-bonded chain[d]	−0.05
Relaxed Pandey π-bonded chain	−0.47
Same with ±2% dimerization	−0.46
Same with ±4% dimerization	−0.43
Same with ±6% dimerization	−0.38
Fully relaxed Pandey chain	−0.68

[a]Reference 18.
[b]Reference 20.
[c]Reference 19.
[d]Reference 11.

The "ideal" Pandey chain model is defined[11] to have all bond lengths equal to the bulk-diamond bond length, except for the surface-chain bond lengths which are equal to the bulk-graphite bond length. This definition is arbitrary, but will serve as a starting point for consideration of relaxations. We find the ideal Pandey model to have an energy slightly lower than that of the ideal 1×1 surface, and lower than that of any of the other initial models, as we shall see.

The atomic coordinates for the Chadi π-bonded molecule model were provided by Chadi,[30] and include relaxations as determined within his tight-binding energy-minimization approach. Even so, the total energy is found to be higher than for the unrelaxed Pandey chain model. It therefore seems very unlikely that this topology would have a lower energy when relaxations are taken into account. Moreover, the calculated surface-state dispersion[17] is in very poor agreement with the ARUPS data.[3] For these reasons, we feel the model can be ruled out, and we have not tried to relax it further.

Finally, we have tested the Seiwatz chain model, again in an "ideal" configuration, this time with all bonds having the bulk-diamond bond lengths. The resulting total energy is a full 1.30 eV above that of the ideal 1×1 model. This energy is so large that we feel it unlikely that relaxations could lower it dramatically. However, we have also tested a single "relaxed" geometry provided by Chadi,[30] again based upon tight-binding energy minimization. The energy of this geometry was found to be reduced to 0.97 eV, which is still very large. The dispersion of the calculated surface bands[17] was found to bear a weak qualitative resemblance to the ARUPS data, but the position of the band is in error by more than 2 eV. Once again, we feel these objections are sufficient to rule out the model, and we have not considered it further.

We now turn to a consideration of relaxations for the two remaining models which have not been eliminated. We begin with relaxed 1×1 models, which were tested as follows. First, the surface bond length was reduced in steps of 0.24 a.u., and a minimum was extrapolated using a low-order polynomial fit (in this case cubic). Then the surface bond length was held constant while the subsurface bond length was varied in steps of 0.06 a.u., and again a minimum was extrapolated. In principle the surface bond should perhaps have been allowed to relax again, but the additional relaxation would probably be small. The relaxation of these first two surface bonds [Fig. 2(a)] lowers the energy by 0.37 eV.

The energy of the Pandey chain model was minimized by adjusting the four surface-most bond lengths, labeled d_1 to d_4 in Fig. 2(b). Initially we kept $d_2=d_2'$, i.e., we retained maximum symmetry of the surface chain. Once again, different bond lengths were optimized individually in turn while the others were held fixed. We optimized first with respect to d_1, then d_2, then d_3, and then d_4. Surprisingly, the bond length d_4 was found to lengthen by a large amount, ~8%. Therefore, in this case the bond lengths d_1, d_2, and d_3 were readjusted. The resulting "relaxed" structure is shown in Fig. 2(b). The relaxations re-

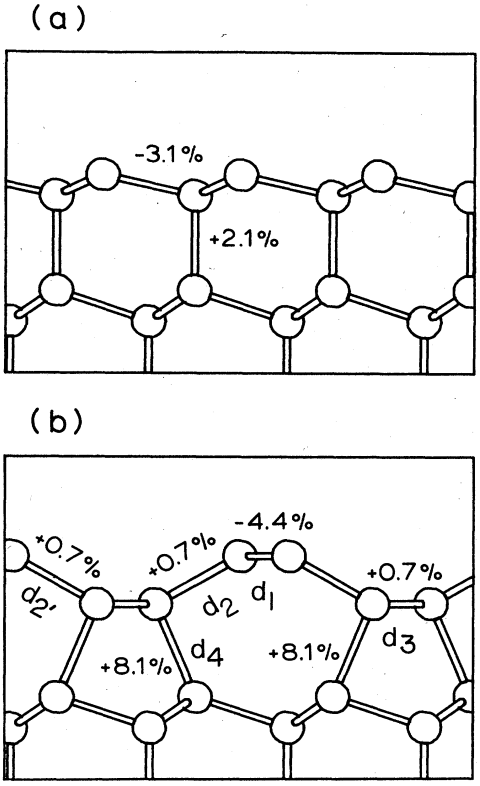

FIG. 2. Illustration of bond-length changes (with respect to bulk) which occur upon relaxation of (a) 1×1 model, and (b) 2×1 Pandey chain model. Certain bond lengths have been labeled for later discussion in the text.

sult in a lowering of the energy to -0.47 eV. The surface chain bond is contracted by $\sim4\%$, a value roughly midway between that of graphite and diamond. Finally, we tested the introduction of asymmetry at the surface chain by calculating the energy for models with $d_2\neq d_2'$, and for models with "tilting" or "buckling" of the surface chain (i.e., bond angle variations only). For both types of distortion, the minimum was found to coincide with the symmetric configuration to the accuracy of the calculation.

Dimerization of this relaxed chain geometry was tested carefully. Table I shows that the energy rises monotonically with the dimerization parameter. In order to test whether the $\vec{k}$-point mesh is fine enough to resolve the band splitting at the $\bar{J}$–$\bar{K}$ zone boundary, where a Peierls gap opens in the surface bands, we have recalculated the energies for small distortions ($\pm1\%$ and $\pm2\%$) and with a larger $\vec{k}$-point set. If the initial 16 $\vec{k}$-point set is described in reciprocal-lattice coordinates as $(i/16,\ \pm j/8)$ $(i=1,3,5,7;\ j=1,3)$, then the 32-point set is derived from it by first replacing the four points $(7/16,\ \pm j/8)$ by eight points $(i/32,\ \pm j/8)$ $(i=13,15)$ with half the weight; and finally by replacing the four points $(15/32,\ \pm j/8)$ by the 16 points $(i/64,\ \pm j/16)$ $(i=29,31;\ j=1,3,5,7)$ with a correspondingly reduced weight. This new $\vec{k}$-point set is very dense in the vicinity of the zone edge. Once again, the energy change was found to be positive. Finally, we tested a model proposed by Chadi[30] in which a radically large dimerization occurs. The energy of this model is found to be 1.1 eV/surface-atom, i.e., it entails a large increase in the energy. Thus we conclude that dimerization always raises the energy.

Finally, the chain geometry was relaxed further using a Keating force-constant model[31] to direct the relaxation of the subsurface atoms in the middle of the slab. This force-constant model contains five on-site terms K_r, K_θ, $K_{rr'}$, $K_{r\theta}$, $K_{\theta\theta'}$, and one additional term $K_{\theta\theta''}$ coupling coplanar bond angles on neighboring sites. It gives a very accurate description of the phonon dispersion throughout the Brillouin zone.[31] A separate computer program was written to relax the chain geometry iteratively so as to minimize the energy within this model, subject to the additional constraint that the four surface-most bond lengths be held fixed at the values already determined. Once this "fully relaxed" geometry was determined, the LDA total energy was again calculated. As expected, the LDA total energy of this final geometry was lowered; it was found to be -0.68 eV. Evidently, this further relaxation relieves some of the bond-angle strains on the third-layer atoms, where the chain topology causes fairly severe departures from the tetrahedral angle. For example, angles of 85° and 134° relax to 93° and 123°, respectively. The coordinates of the atoms in this fully relaxed geometry are given in Table II.

It can be seen from Table II that the structural relaxations due to the two surfaces have not entirely decayed to zero at the slab center. In order to be sure that the interaction between the strain fields of the two surfaces is not significant, we have estimated the interaction within the Keating model by carrying out parallel calculations for the 20-atom slab and for a thicker 28-atom slab containing four additional bulk layers, but with identical surfaces. Both models were fully relaxed, and the resulting Keating energies differed by only 0.015 eV. (The atomic displacements at slab center were an order of magnitude smaller for the 28-atom slab, so any remaining elastic interactions are presumably negligible.) A correction of order ~0.02 eV is unimportant, and this effect need not be considered further.

Figure 3 shows the calculated surface band structure in

TABLE II. Coordinates of atoms for fully relaxed Pandey π-bonded chain geometry. The magnitudes of the lattice vectors in the x and y directions are 4.758 60 a.u. and 8.242 14 a.u., respectively. Only the top half of the slab is given because of the presence of inversion symmetry.

Atom	x (a.u.)	y (a.u.)	z (a.u.)
1	2.379 30	1.407 11	8.716 02
2	0.0	2.846 67	8.789 99
3	2.379 30	−1.138 60	7.270 79
4	0.0	5.389 89	7.329 44
5	2.379 30	−0.100 27	4.285 62
6	0.0	4.198 15	4.420 88
7	0.0	−1.397 79	3.174 27
8	2.379 30	2.764 67	3.630 38
9	0.0	−1.391 60	0.372 06
10	2.379 30	2.764 09	0.596 10

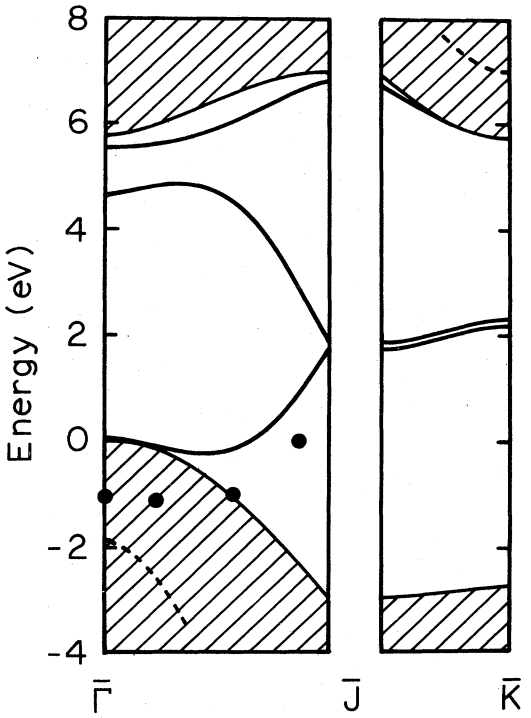

FIG. 3. Calculated surface bands (solid lines) and resonances (dashed lines) for fully relaxed Pandey chain model. The bulk projected band structure (shaded) and the experimental ARUPS data of Ref. 2 (black dots) are shown for comparison.

the gap region for this fully relaxed model. The major features in the gap are the occupied and unoccupied surface bands associated with the surface chains. The dispersion of the occupied surface band along $\overline{\Gamma J}$ has been greatly reduced from that of Pandey,[11] who found a difference of $\gtrsim 3$ eV between $E(\bar{\Gamma})$ and $E(\bar{J})$. This was already reduced to ~ 2.3 eV in our previous calculation on the ideal structure,[17] and has now been reduced further to ~ 1.7 eV due to relaxations (primarily the lengthening of the surface chain bonds by $\sim 4\%$ from their "ideal" graphitic values). The dispersion is thus in good agreement with experiment without the need for dimerization, thereby removing one of the major motivations for the dimerized model. The calculated band is too high by a rigid shift of ~ 1 eV, but this is also true (by ~ 0.3 and ~ 0.8 eV, respectively) for Si and Ge.[12] This discrepancy will be discussed further in the following section. Note that we find a second unoccupied surface band in the neighborhood of the conduction band edge; it is similar to the bulk conduction band in character, and weakly localized in the vicinity of the lengthened subsurface interlayer bonds.

Finally, the charge density has been calculated for this fully relaxed chain geometry, and is shown in Fig. 4. The total charge density is shown in Fig. 4(a). It is evident that the bond charge is somewhat reduced along the elongated subsurface interlayer bond d_4, as should be expected. Figure 4(b) shows the charge density of the occupied surface band, and clearly indicates the dangling bond nature of this state. The degree of localization at the surface chain is evidently very high. The orientation of the

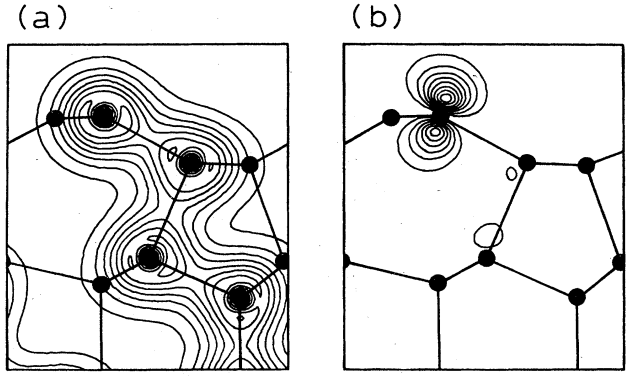

FIG. 4. Charge-density contour plots in the plane $x=0$ (perpendicular to the chain). Atom positions are indicated by filled circles; only half are in the plane of the plot. (a) Total charge density. (b) Charge density of the occupied surface band in the gap.

dangling bonds is close to vertical, the deviation being only $\sim 18°$. It is therefore quite plausible that π interactions between neighboring dangling bonds would be strong enough to stabilize this structure.

IV. DISCUSSION

The driving force for the reconstruction of the (111) surface is presumably the presence, on every surface atom, of an energetically unfavorable dangling bond containing an unpaired electron. The Haneman buckling model attempts to pair electrons via a kind of charge-density-wave distortion, while the Pandey chain, Chadi molecule, and Seiwatz single chain models make use of topological reconstructions which arrange for the dangling bonds to become nearest neighbors, thereby allowing π bonding between the dangling bonds. However, all these pairing mechanisms involve a cost. In the case of the buckled model, it is the Coulomb correlation energy of putting a second electron in one dangling bond; for the topological models, it is the elastic strains associated with the rebonding. Based on our total-energy calculations, it appears that the costs outweigh the gains in all cases except that of the π-bonded chain model. One can make a plausibility argument in each case: the highly localized nature of the carbon wave functions should make the Coulomb correlation relatively large for the buckled model; the elastic strains should be largest for the Chadi π-bonded molecule, which requires the most severe subsurface rebonding; and the π-bonding energy gain should not be very large for the Seiwatz single chain model, whose dangling bonds are so far from vertical. The π-bonded chain model evidently strikes a profitable balance between favorable π bonding and unfavorable strain energies.

It is instructive to compare to the case of Si or Ge for the π-bonded chain model. In carbon, both the π bonding and elastic energies are relatively larger. The fact that π bonding is much more common in carbon chemistry is one indication of the relatively stronger π bonding in C. The elastic energies are also expected to be larger, because

the principal strains are due to bond bending in the surface region. Because the bonding is more directional in carbon, the cost should be larger. A measure of this directionality is the ratio of bond bending to bond stretching force constants, which is 0.23 and 0.11 $\text{Å}^2\,\text{rad}^{-2}$ for C and Si, respectively, in the Keating model used above.[31] One consequence is that consideration of relaxations for the Pandey chain model is considerably more important for C than for Si. Energy-minimization calculations for Si indicate[12] that the ideal Pandey chain model is already at -0.22 eV with respect to the ideal 1×1, while relaxations only lower this to -0.36 eV; for C the corresponding numbers are -0.05 eV and -0.68 eV, respectively.

The possible dimerization of the Pandey chain model is a question of considerable interest, because it would open up the possibility of observing soliton formation and propagation on a crystal surface. The surface chain in this model is analogous to *trans*-polyacetylene, $(CH)_x$, with the bonds to hydrogen atoms replaced by bonds to the second-layer atoms. In suggesting the dimerization of the chain model, Pandey was motivated by the need to reduce the dispersion of the occupied surface band, as well as by the analogy to $(CH)_x$. As we showed in the previous section, the dispersion has been corrected without the need for dimerization. Moreover, the analogy with $(CH)_x$ is only approximate, and there are several reasons why dimerization might be expected to be less likely here. First, the elastic restoring forces against dimerization must be stronger because of the bonding to the subsurface bulk; the corresponding bonds to H atoms in $(CH)_x$ are entirely unconstrained. Second, the π interactions are weaker here because the dangling bonds are not entirely parallel. Third, the interactions between the chains introduce some dispersion along the $\bar{J}$–$\bar{K}$ zone boundary, and the inequivalence between the two surface atoms making up the chain introduces a splitting (see Fig. 3); while these effects are smaller than for the chain model in Si and Ge,[12] they still allow us to circumvent the Peierls argument that sufficiently small dimerization is always favorable in truly one-dimensional systems.

Antiferromagnetic (AF) ordering is an alternative mechanism which could open a gap along the $\bar{J}$–$\bar{K}$ zone boundary. This possibility has been considered within a mean-field picture,[32] and gives rise to a gap and to a splitting between spin-up and spin-down occupied surface bands which is small as long the inequivalence of the two chain atoms is small. Thus the fact that only a single band is observed in the ARUPS data could well be consistent with this picture. Moreover, the dispersion of the theoretical surface bands would be slightly reduced, and the position would be slightly lower. If the chains are weakly interacting, a one-dimensional Hubbard model might be a more appropriate starting point; here there is no true long-range AF order, but a gap occurs nevertheless.[33] We think it likely that some such AF ordering occurs; it has not been included in our calculations, but would have the effect of lowering the energy of the π-bonded chain model further.

We end this section by reviewing the relationship of our results to the experimental work for the diamond (111) $(2\times2)/(2\times1)$ surface. The most serious discrepancy is in the location of the occupied surface band, which is ~1 eV higher than indicated by the ARUPS data.[3] There are three sources of error which may all be responsible in part for the discrepancy. First, the position of the experimental surface band is affected by the location of the Fermi level, which can be determined only approximately.[3] Second, inclusion of the AF ordering mentioned above might improve the agreement. Finally, the local-density theory itself may be at fault. In the density-functional theory, upon which it is based, the ground-state *total energies* are given exactly, while the *eigenvalues* have *no* physical meaning as electron removal energies. The photoemission experiment measures just such a removal energy, which contains a component, namely the interaction between the electron being removed and the hole left behind, that is not correctly taken into account in LDA. Qualitatively, a correction for this effect would be such as to lower the theoretical occupied surface band, which is in the right direction. In any case, the total-energy-minimization procedure should be more reliable, in light of these arguments, than a detailed indirect comparison of the experimental and theoretical eigenvalues.

Other meaningful experimental work is largely unavailable. We are not aware of any attempt at a LEED refinement for this structure. An analysis of missing spots in the LEED pattern has led Yang and Jona[5] to the conclusion that the surface structure should be constructed out of chain-like units, but this ought to be consistent with either the Pandey or Seiwatz chain models. Electron energy loss experiments by Pepper[4] indicate the existence of unoccupied surface states near midgap, which is consistent with the π-bonded chain model. They also appear to indicate the existence of a sizable gap, although the width of the elastic scattering peak makes this somewhat uncertain; in any case, a gap is consistent with a π-bonded chain model with AF ordering.

We hope the present work will serve to stimulate further experimental investigation, which might confirm our identification of the π-bonded chain model as the correct structure for this surface.

V. SUMMARY

We have developed a new *ab initio* LCAO approach to the calculation of total energies within local-density theory. The method relies on the variational principle for the treatment of intra-atomic charge transfer, but incorporates a fully self-consistent treatment of interatomic charge transfer. All wave functions and potentials are represented in terms of Gaussians on the site centers, so that all three-center integrals can be taken analytically in an efficient fashion. An improved version of Broyden's method is used to control the convergence upon the self-consistent solution.

The method is applied to study a variety of proposed reconstructions on the diamond (111) surface. A single infinite slab geometry is used to realize these models, which include the Haneman buckling model, the Pandey π-bonded chain model, the Chadi π-bonded molecule model, and the Seiwatz single chain model. The *undimerized* π-bonded chain model is found to have the lowest en-

ergy, ~0.3 eV/surface-atom lower than for the relaxed 1×1 surface. The other models are found to be implausible on the basis of total energies and surface-state dispersion. The dispersion of the calculated surface band is found to be in good agreement with experiment for the fully relaxed π-bonded chain model, without the need for dimerization. The position of the band is shifted by ~1 eV; several factors are discussed which may be responsible for this. Finally, we have argued that the π-bonded chain model is plausible in terms of energetic considerations, and is consistent with all experiments reported to date.

ACKNOWLEDGMENTS

Support for this work was provided by National Science Foundation Grant No. DMR-83-19024 and by a program development fund from the Director of the Lawrence Berkeley Laboratory. Cray-1 computer time was provided by the U.S. Office of Energy Research of the Department of Energy. We (D.V. and S. G. L.) would also like to acknowledge support by the Miller Institute and Alfred P. Sloan Foundation, respectively.

APPENDIX: MODIFIED BROYDEN METHOD FOR CONVERGENCE ACCELERATION

We propose here a modified version of Broyden's method[22] for updating the Jacobian in iterative improvement. The method is equally applicable to convergence acceleration for plane-wave and LCAO calculations, and to structural relaxation based on force calculations.

In general, one wants to minimize a vector function $\vec{F}(\vec{x})$ of a vector variable $\vec{x}$. For the LCAO case, $\vec{x}$ is the vector $\vec{Q}$ of site populations, and

$$\vec{F}(\vec{Q})=\vec{Q}_{\rm out}(\vec{Q})-\vec{Q}\ , \tag{A1}$$

where $\vec{Q}_{\rm out}(\vec{Q})$ is the output charge after one iteration starting from $\vec{Q}$. Similarly, for the plane-wave case, $\vec{x}$ is the vector $\vec{V}$ of Fourier coefficients of the potential, $V_i=V(\vec{G}_i)$, and

$$\vec{F}(\vec{V})=\vec{V}_{\rm out}(\vec{V})-\vec{V}\ . \tag{A2}$$

For structural relaxation, $\vec{x}$ is the vector of site coordinates, and $\vec{F}$ is the vector of forces. In all cases the problem is to find $\vec{x}$ such that $\vec{F}(\vec{x})=\vec{0}$.

First an initial $\vec{x}^{\,0}$ is chosen and $\vec{F}^{(0)}=\vec{F}(\vec{x}^{\,(0)})$ is determined. Then we may predict that, in linear order, a new $\vec{x}^{\,(1)}$ will result in an $\vec{F}^{(1)}$ satisfying

$$\vec{F}^{(1)}-\vec{F}^{(0)}\approx-\overleftrightarrow{J}\cdot(\vec{x}^{\,(1)}-\vec{x}^{\,(0)})\ , \tag{A3}$$

where the Jacobian is defined as

$$J_{ij}\equiv-\frac{\partial F_i}{\partial x_j}\ . \tag{A4}$$

An initial guess $\overleftrightarrow{J}^{(0)}$ is made to $\overleftrightarrow{J}$, and since we want $\vec{F}^{(1)}=0$, $\vec{x}^{\,(1)}$ is chosen as

$$\vec{x}^{\,(1)}=\vec{x}^{\,(0)}+[\overleftrightarrow{J}^{(0)}]^{-1}\cdot\vec{F}^{(0)} \tag{A5}$$

to satisfy Eq. (A3). In general, on the mth iteration we choose

$$\vec{x}^{\,(m+1)}=\vec{x}^{\,(m)}+[\overleftrightarrow{J}^{(m)}]^{-1}\cdot\vec{F}^{(m)}\ . \tag{A6}$$

Then $\vec{F}^{(m+1)}$ is computed from $\vec{x}^{\,(m)}$ and the normalized differences

$$\Delta\vec{x}^{\,(m)}=(\vec{x}^{\,(m+1)}-\vec{x}^{\,(m)})/\,|\,\vec{x}^{\,(m+1)}-\vec{x}^{\,(m)}\,|\ , \tag{A7}$$

$$\Delta\vec{F}^{(m)}=(\vec{F}^{(m+1)}-\vec{F}^{(m)})/\,|\,\vec{x}^{\,(m+1)}-\vec{x}^{\,(m)}\,| \tag{A8}$$

are obtained.

Now in the usual Broyden's method we observe that, from Eq. (A3), $\overleftrightarrow{J}$ ought to satisfy

$$\overleftrightarrow{J}\cdot\Delta\vec{x}^{\,(m)}=-\Delta\vec{F}^{(m)} \tag{A9}$$

although $\overleftrightarrow{J}^{(m)}$ fails to do so. Therefore $\overleftrightarrow{J}^{(m+1)}$ is chosen so that (i) $\overleftrightarrow{J}^{(m+1)}$ satisfies Eq. (A9), and (ii) $||\overleftrightarrow{J}^{(m+1)}-\overleftrightarrow{J}^{(m)}||$ is minimized subject to constraint (i). ($||\overleftrightarrow{A}||\equiv\sum_{ij}|A_{ij}|^2$.) The result is

$$\overleftrightarrow{J}^{(m+1)}=\overleftrightarrow{J}^{(m)}-[\Delta\vec{F}^{(m)}+\overleftrightarrow{J}^{(m)}\cdot\Delta\vec{x}^{\,(m)}]\otimes\Delta\vec{x}^{\,(m)T}\ , \tag{A10}$$

which corresponds to changing one column of $\overleftrightarrow{J}^{(m)}$ in a representation in which $\Delta\vec{x}^{\,(m)}$ is one of the basis vectors. However, this new $\overleftrightarrow{J}^{(m+1)}$ does not generally satisfy

$$\overleftrightarrow{J}^{(m+1)}\cdot\Delta\vec{x}^{\,(l)}=-\Delta\vec{F}^{(l)} \tag{A11}$$

for $l<m$, as it should. Thus, information from the most recent iteration is used to update $\overleftrightarrow{J}$, but is allowed to override arbitrarily information from earlier iterations.

We propose here a modification in which information is incorporated from all previous iterations on each update of $\overleftrightarrow{J}$. A weight $w^{(l)}$ is associated with each previous iteration, and a weight w' is assigned to the initial guess $\overleftrightarrow{J}^{(0)}$. The $\overleftrightarrow{J}^{(m+1)}$ is determined according to a least-squares minimization:

$$E=\sum_{l=0}^{m}w_l^2\,|\,\overleftrightarrow{J}^{(m+1)}\cdot\Delta\vec{x}^{\,(l)}+\Delta\vec{F}^{(l)}\,|^2 +w'^{\,2}||\overleftrightarrow{J}^{(m+1)}-\overleftrightarrow{J}^{(0)}||^2\ . \tag{A12}$$

Setting $0=\partial E/\partial J_{ij}^{(m+1)}$ gives

$$\overleftrightarrow{J}^{(m+1)}=\overleftrightarrow{\gamma}^{(m+1)}\cdot(\overleftrightarrow{\beta}^{(m+1)})^{-1}\ , \tag{A13}$$

where

$$\overleftrightarrow{\beta}^{(m+1)}=w'^{\,2}\overleftrightarrow{1}+\sum_{l=0}^{m}w_l^2\Delta\vec{x}^{\,(l)}\otimes\Delta\vec{x}^{\,(l)T} \tag{A14}$$

and

$$\overleftrightarrow{\gamma}^{(m+1)}=w'^{\,2}\overleftrightarrow{J}^{(0)}-\sum_{l=0}^{m}w_l^2\Delta\vec{F}^{(l)}\otimes\Delta\vec{x}^{\,(l)T}\ . \tag{A15}$$

The appearance of the weights w' and w_l gives the method considerable flexibility. For example, early iterations for which $\vec{x}$ was far from optimal, or for which smaller $\vec{k}$-point sets or basis sets were used, can be

TABLE III. Convergence test comparing Broyden convergence accelerator (Ref. 22) with present method.

Iteration	$\|\overleftrightarrow{J}-\overleftrightarrow{D}\|\|$	$\|\vec{x}\|$	x_1	x_2	x_3	x_4	x_5
Broyden							
0	2.345	2.236	1.000	1.000	1.000	1.000	1.000
1	2.345	2.358	0.490	−0.010	−0.510	−1.010	−2.010
2	1.745	0.700	0.388	−0.006	−0.192	−0.168	0.524
3	1.645	0.432	0.217	−0.001	0.090	0.244	−0.268
4	1.304	0.154	0.141	0.000	0.021	−0.046	0.034
5	1.251	0.075	0.072	0.000	−0.013	0.015	0.007
6	1.129	0.032	0.029	0.000	0.008	−0.010	0.005
7	0.968	0.003	−0.002	0.000	−0.001	0.002	0.000
8	0.961	0.000	0.000	0.000	0.000	0.000	0.000
Present							
0	2.345	2.236	1.000	1.000	1.000	1.000	1.000
1	2.345	2.358	0.490	−0.010	−0.510	−1.010	−2.010
2	1.745	0.703	0.389	−0.006	−0.192	−0.167	0.529
3	1.157	0.305	0.259	−0.002	0.021	0.143	−0.072
4	0.776	0.135	0.122	0.001	0.034	−0.047	0.009
5	0.419	0.035	0.030	0.000	−0.015	0.009	−0.002
6	0.056	0.000	0.000	0.000	0.000	0.000	0.000

weighted more lightly than recent iterations. Note that in the limit that $w_m \gg w'$ and $w_m \gg w_{l<m}$, the present method reduces to the standard Broyden method.

It is convenient, but not necessary, for the weights to be fixed (i.e., each w_l is independent of m). If this is done, $\overleftrightarrow{\beta}$ and $\overleftrightarrow{\gamma}$ can be updated easily from one iteration to the next, and there is no need to save the $\Delta\vec{x}^{(l)}$'s and $\Delta\vec{F}^{(l)}$'s from earlier iterations. Moreover, $\overleftrightarrow{\beta}^{-1}$ can be updated using the identity

$$(A+\vec{u}\otimes\vec{v}^{\,T})^{-1}=\overleftrightarrow{A}^{-1}-\sigma^{-1}\overleftrightarrow{A}^{-1}\cdot\vec{u}\otimes\vec{v}^{\,T}\cdot\overleftrightarrow{A}^{-1}\,, \tag{A16}$$

$$\sigma=1+\vec{v}^{\,T}\cdot\overleftrightarrow{A}^{-1}\cdot\vec{u}\,.$$

Thus no matrix inversion need ever be done explicitly; the operations are all computationally efficient, even for large-dimensional problems.

We have tested the method on a sample five-dimensional problem in which $\vec{F}(\vec{x})$ is given by

$$F_i=-d_i x_i-c x_i^3\,. \tag{A17}$$

Clearly the solution is $\vec{x}=\vec{0}$ with Jacobian $\overleftrightarrow{D}$, where $\overleftrightarrow{D}$ is the diagonal matrix formed from the d_i. The cubic term was included to simulate some nonlinearity. We used $\vec{d}=(3,2,1.5,1,0.5)$ and $c=0.01$. The initial guess $\overleftrightarrow{J}^{(0)}$ was set equal to the identity matrix, and $\vec{x}^{(0)}$ was chosen as $(1,1,1,1,1)$. Convergence to the solution was tested for the Broyden and the present methods; the results are presented in Table III. Here we used $w'=0.01$ and $w^{(l)}=1$ for all l. As can be seen, the convergence to the correct $\vec{x}$ occurs more rapidly using the present scheme. Moreover, while $\overleftrightarrow{J}$ does not converge to the true value at all in the standard Broyden scheme, it does so now.

In summary, the present approach improves the performance of the Broyden method, provides added flexibility, and is no more difficult to implement. We see no reason why its use should not become widespread in iterative calculations.

[1] J. J. Lander and J. Morrison, Surf. Sci. **4**, 241 (1966).
[2] B. B. Pate *et al.*, J. Vac. Sci. Technol. **19**, 349 (1981).
[3] F. J. Himpsel, D. E. Eastman, P. Heimann, and J. F. van der Veen, Phys. Rev. B **24**, 7270 (1981).
[4] S. V. Pepper, Surf. Sci. **12**, 47 (1982).
[5] W. S. Yang and F. Jona, Phys. Rev. B **29**, 899 (1984).
[6] B. B. Pate *et al.*, J. Vac. Sci. Technol. **21**, 364 (1982).
[7] B. J. Waclawski, D. T. Pierce, N. Swanson, and R. J. Celotta, J. Vac. Sci. Technol. **21**, 368 (1982).
[8] For a recent review, see D. Haneman, Adv. Phys. **31**, 165 (1982).
[9] F. Houzay *et al.*, Surf. Sci. **132**, 40 (1983) and references therein.
[10] J. M. Nicholls, G. V. Hansson, R. I. G. Uhrberg, and S. A. Flodstrom, Phys. Rev. B **27**, 2594 (1983).
[11] K. C. Pandey, Phys. Rev. Lett. **47**, 1913 (1981); Phys. Rev. B **25**, 4338 (1982).
[12] J. E. Northrup and M. L. Cohen, Phys. Rev. Lett. **49**, 1349 (1982); Phys. Rev. B **27**, 6553 (1983).
[13] R. M. Tromp, L. Smit, and J. F. van der Veen, Phys. Rev. Lett. **51**, 1672 (1983).
[14] P. Chiaradia *et al.*, Phys. Rev. Lett. **52**, 1145 (1984); M. A. Olmstead and N. Amer, *ibid.* **52**, 1148 (1984).
[15] R. Feder, W. Mönch, and P. P. Auer, J. Phys. C **12**, L179 (1979).
[16] F. Solal *et al.*, Phys. Rev. Lett. **52**, 360 (1984).
[17] D. Vanderbilt and S. G. Louie, J. Vac. Sci. Technol. B **1**, 725 (1983).
[18] D. Haneman, Phys. Rev. **121**, 1093 (1961).
[19] R. Seiwatz, Surf. Sci. **2**, 473 (1964).

[20]D. J. Chadi, Phys. Rev. B **26**, 4762 (1982).
[21]J. R. Chelikowsky and S. G. Louie, Phys. Rev. B **29**, 3470 (1984).
[22]C. G. Broyden, Math. Comp. **19**, 577 (1965); see also P. Bendt and A. Zunger, Phys. Rev. B **26**, 3114 (1982).
[23]D. Vanderbilt and S. G. Louie, J. Comput. Phys. (to be published).
[24]R. S. Mulliken, J. Chem. Phys. **23**, 1833 (1955).
[25]K. M. Ho, J. Ihm, and J. D. Joannopoulos, Phys. Rev. B **25**, 4260 (1982).
[26]J. Ihm, A. Zunger, and M. L. Cohen, J. Phys. C **12**, 1979; see also D. Vanderbilt, Ph. D. thesis, MIT, 1981.
[27]D. R. Hamann, M. Schlüter, and C. Chiang, Phys. Rev. Lett. **43**, 1494 (1979).
[28]L. Hedin and B. I. Lundqvist, J. Phys. C **4**, 2064 (1971).
[29]The energy of the ideal 1×1 surface was found to be 3.2 eV/surface-atom compared to bulk.
[30]D. J. Chadi (private communication); see also in *Proceedings of the IXth International Vacuum Congress and Vth International Conference on Solid Surfaces,* edited by J. L. de Segovia (Imprenta Moderna, Madrid, Spain, 1983), pp. 80—88; J. Vac. Sci. Technol. B **2**, 948 (1984).
[31]R. Tubino, L. Piseri, and G. Zerbi, J. Chem. Phys. **56**, 1022 (1972).
[32]M. Lannoo and G. Allan, Solid State Commun. **47**, 153 (1983).
[33]A. Martín-Rodero *et al.,* Phys. Rev. B **29**, 476 (1984).