Surface reconstruction and vibrational excitations of Si(001)

O. L. Alerhand and E. J. Mele
Department of Physics, University of Pennsylvania, Philadelphia, Pennsylvania 19104-6396
(Received 17 October 1986)

We present theoretical studies of vibrational excitations on the Si(001) surface. Three different reconstructions based on the surface dimer model are considered. The theoretical model that we use consists of a tight-binding theory for structural energies that we have extended to include explicit electron-electron interactions in the form of an on-site repulsion term. Perturbation theory is applied to calculate the dynamical matrix of the system under study. The phonon spectrum of the Si(001) 2×1 surface is presented and analyzed: we observe a number of modes that are characteristic of the dimer reconstruction. The analysis of the vibrational excitations of the surface provides a microscopic explanation for the driving forces that lead to higher-order reconstructions, namely, $p(2\times2)$ and $c(4\times2)$ structures, and provides an insight into the question of the multiplicity of periodicities that are observed on this surface, even at low temperatures. Using linear-response theory, we calculate the dipole activity of surface phonons on the 2×1, $p(2\times2)$, and $c(4\times2)$ surfaces. The absorption spectra that we obtain can be used to characterize the periodicity of the surface; effects of surface symmetry and surface polarizability on the spectra are also discussed.

I. INTRODUCTION

The nature of the reconstruction of the Si(001) surface has been studied over the last fifteen years and remains a subject of current interest. This surface exhibits a combination of strong short-range reconstruction and weak long-range ordering. Recent experiments on Si(001),[1-6] and on the closely related Ge(001),[7] have explored this question, renewing the interest on these systems. The standard model of the reconstruction of the Si(001) surface[8] involves the formation of tilted surface dimers, as is schematically shown in Fig. 1(b), compared with the bulk-terminated surface of Fig. 1(a). The driving force of this reconstruction is bond formation, reducing the number of dangling bonds at the surface. By forming a tilted dimer the symmetry between the two surface atoms is broken and a gap is opened in the electronic surface states, creating a semiconducting surface, as is observed experimentally.[3,9,10] The value of the tilt, or even whether the dimers are flat or tilted, is a question that is still discussed; it is an important question since the surface charge density and the gap of the electronic surface states strongly depend on this structural parameter.[11] However, once one assumes that the basic building block of the reconstruction of Si(001) is the tilted surface dimer, the possibility exists for different ordering arrangements of these dimers leading to different periodicities. For example, an ordered arrangement where all the dimers are tilted in phase gives the 2×1 structure, while alternating dimers along the $\langle\bar{1}10\rangle$ direction (across the dimers) leads to a surface reconstruction with $p(2\times2)$ periodicity; if the dimers also alternate along the $\langle110\rangle$ direction the reconstruction has $c(4\times2)$ periodicity. The resulting reconstructions are shown in Figs. 1(c) and 1(d), respectively. Of course, one can have many other higher-order periodicities.

There have been experimental and theoretical studies

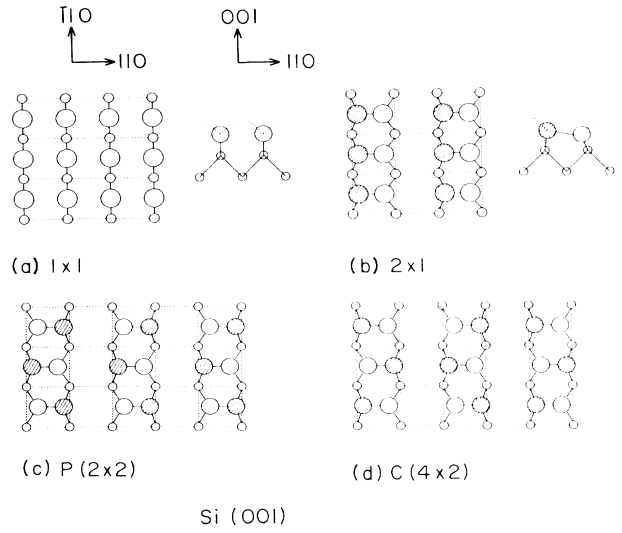

FIG. 1. Various reconstructions of the Si(001) surface based on the surface-dimer model. (a) Ideal surface, top and side view; the surface atoms (large circles) are unpaired, each having two dangling bonds and two backbonds. The smaller circles represent atoms in the first subsurface layer. (b) Pairs of neighboring atoms along the $\langle110\rangle$-direction bond forming tilted surface dimers. Solid circles represent the surface atoms that are displaced out of the surface. The surface dimers form rows along the $\langle\bar{1}10\rangle$ direction. In this 2×1 reconstruction the surface unit cell is doubled along the $\langle110\rangle$ direction. Different ordered arrangements of the tilted surface dimers lead to further reconstruction of the surface: (c) alternating dimers along the $\langle\bar{1}10\rangle$ direction result in a surface with $p(2\times2)$ periodicity, and (d) alternating dimers along both surface directions form a $c(4\times2)$ surface. In both of the last two reconstructions there are two surface dimers (four atoms) per surface unit cell.

that suggest that the equilibrium reconstruction of Si(001) below room temperature has $c(4\times2)$ symmetry, but that the surface dimers are randomly oriented with respect to each other at room temperatures. Furthermore, even at low temperatures the surface exhibits some disorder: weaker signatures of other periodicities are seen superimposed on the $c(4\times2)$ periodicity. Kevan[7] studied this order-disorder phase transition combining low-energy electron-diffraction (LEED) and high-resolution angle-resolved photoemission techniques for the Ge(001) surface; he found that this surface undergoes a two-stage phase transition from the disordered phase to the $c(4\times2)$ phase. This experiment was done with Ge rather than with Si surfaces, but one can expect close similarities. We will return to discuss this experiment in connection with the results we obtain. Recently Sakai, Cardillo, and Hamman[1] reported He-scattering experiments analyzed with a newly proposed modified atomic-charge superposition method; they conclude that the tilted-surface-dimer model with $c(4\times2)$ periodicity is the equilibrium reconstruction of Si(001) at low temperatures. These add to a list of other studies, that include a variety of experimental techniques, which show evidence for higher-order reconstructions on the Si(001) surface. The scanning tunneling microscope (STM) might provide a direct determination of the structure of the surface, though one must be careful in interpreting STM corrugations since these are determined not only by atomic positions but also by the electronic charge density at the surface. Recent experiments by Hamers, Tromp, and Demuth[2] confirm that the reconstruction of Si(001) is indeed based on surface dimers. With respect to the periodicity of the surface, the images they obtained show regions with $p(2\times2)$ and $c(4\times2)$ periodicities, together with regions of apparently symmetric dimers. However, besides the difficulty in interpreting these images mentioned above, one must consider that these experiments were done at room temperature where the equilibrium phase is expected to be disordered.

Theoretical efforts based on total-energy calculations have been directed toward finding the equilibrium structure of the surface. Using a tight-binding theory for structural energies, Chadi[8] found that the dimers tend to tilt and that the $c(4\times2)$ reconstruction is lower in energy than the 2×1 surface. He included in his calculations the Madelung energy, which further stabilizes alternating dimers across or parallel to the dimer direction. Yin and Cohen[21] studied the 2×1 reconstruction using first-principles density-functional theories and concluded that the asymmetric dimers are energetically favorable. However, further density-functional calculations by Pandey[13] suggested that the flat dimer corresponds to the lowest-energy configuration, although the tilted and untilted configurations have very close energies. However, a metallic surface state, implied by symmetric surface dimers, has not been observed experimentally.[3,9,10] Thus one must conclude that the surface dimers cannot be symmetric (unless strong correlation effects are present[14]). Returning to the tight-binding approach, Ihm *et al.*[15] made a theoretical study of this surface where a tight-binding Hamiltonian is mapped onto a spin Hamiltonian that represents the tilted dimers; they found that the low-temperature equilibrium configuration of Si(001) corresponds to a $p(2\times2)$ ordered arrangement of the tilted surface dimers, and that the surface undergoes a second-order phase transition to a disorder phase at about room temperature. The two reconstructions $c(4\times2)$ and $p(2\times2)$ were found to have essentially the same energy, so that the low-temperature phase can be either $p(2\times2)$ or $c(4\times2)$, in agreement with the results of Chadi. In the present study we use the tight-binding approach initially proposed by Chadi,[16] with the addition to the model of an on-site electron-electron repulsion term, similar to the one proposed by Harrison.[17] This addition requires that the electronic structure of the system under study be calculated self-consistently. We find that even in the presence of moderately strong Coulomb repulsion the tilted dimers are favored. With respect to the lattice dynamics and vibrational excitations, there have been very few attempts to study semiconductor surfaces. Tiersten, Ying, and Reinecke[18] used a Keating model to calculate phonon energies. Though this approach is convenient for calculating vibrational correlation functions it lacks any microscopic information regarding the coupling of the surface-electronic and structural degrees of freedom. We have previously reported some results on the phonon spectrum of this same surface[19] and its hydrogenated version,[20] without the presence of electron-electron repulsion in the model. Here we expand the scope of these previous investigations.

In this paper we investigate the question of what is the low-temperature equilibrium configuration of the Si(001) surface within the tilted-surface-dimer model, and how it might be probed by vibrational surface excitations. We find that the two phases $c(4\times2)$ and $p(2\times2)$ are indeed favored over the 2×1 reconstruction. *The main theme that emerges from our studies is the importance of the coupling of surface electrons to surface phonons in providing a microscopic understanding of the equilibrium structure of this surface, and of reconstructed semiconductor surfaces in general.* We find that vibrational fluctuations on the 2×1 surface which strongly couple to the surface electron states comprise the mechanism leading to the order-disorder phase transition, providing an insight into the formation of the reconstructions with higher-order periodicity. A number of modes associated with the presence of surface dimers appear in the phonon spectrum and are analyzed. The theoretical approach that we use also allows us to calculate response functions that can be used to characterize the 2×1 and the higher-order reconstructions: We calculate the phonon-assisted contribution to the surface conductivity and find a substantial difference in the infrared-absorption spectrum of the $c(4\times2)$ and $p(2\times2)$ reconstructions with respect to the 2×1 surface, suggesting that in addition to diffraction one can use spectroscopy to distinguish the different proposed periodicities of the Si(001) surface. We understand this difference as a result of the symmetries of the reconstructions. An interesting result of our studies is that a large part of the charge fluctuations driven by vibrational excitations occurs in subsurface bonds, contrary to what we found on the Si(111) surface,[21] where essentially all these charge oscillations occur at the surface. This is consistent with our

understanding that the $2\times1\rightarrow c(4\times2),p(2\times2)$ reconstruction is a result of local chemical effects, rather than a true Fermi-surface instability, in agreement with the speculations of Kevan.[7]

The structure of this paper is as follows. In the next section we briefly describe the theoretical model, with emphasis on the new addition of the electron-electron repulsion term in the calculations of total energies and vibrational excitations. In Sec. III we present our results for the structure and phonon spectrum of the Si(001) 2×1 surface. In Sec. IV we investigate the higher-order reconstructions $c(4\times2)$ and $p(2\times2)$. The experimental implications of the different periodicities of the reconstructions we are studying are analyzed in Sec. V. Here we present our calculations of response functions, such as the surface conductivity $\sigma(\omega)$, which suggests a way of spectroscopically differentiating between the 2×1 and the higher-order reconstructions. In Sec. VI we study the phase diagram of the surface in the mean-field approximation using an effective spin Hamiltonian that represents the surface structural energy, and compare our results with those of Ihm *et al.*[15] Finally, we summarize our results.

II. FORMALISM

The starting point in our calculations is an expression for structural energies where the valence electrons are treated in the tight-binding (TB) approximation.[16] We have extended this theory to study vibrational excitations of reconstructed Si surfaces and other complex Si structures,[22,19,21] and we have now incorporated explicit electron-electron interactions into the model in the form of a Hubbard-like term. In this section the formalism for calculating the lattice dynamics is presented, leaving for a later section the derivation of surface conductivities as response functions. The total energy of the system is obtained from the Hamiltonian

$$H=\sum_i\sum_{\alpha=s,p}\varepsilon_\alpha a_{i\alpha}^\dagger a_{i\alpha}+\sum_{i,j}\sum_{\alpha,\beta=s,p}h_{i\alpha,j\beta}a_{i\alpha}^\dagger a_{j\beta}+H_U+V_{\text{elastic}}\ , \tag{1}$$

where i,j label ions and α,β denote TB orbitals. We use a set of nearest-neighbor sp_3 parameters obtained by fitting the electronic bands of Si in the diamond structure: $\varepsilon_s=0.00$ eV, $\varepsilon_p=6.54$ eV, $(ss\sigma)=-19.4$ eV, $(sp\sigma)=1.75$ eV, $(pp\sigma)=-1.08$ eV, and $(pp\pi)=3.05$ eV, and we assume that these parameters scale with interatomic separation d as $1/d^2$.

In an effort to incorporate intra-atomic electron-electron interactions, we have added a Hubbard-like term to the model:

$$H_U=\tfrac{1}{2}U\sum_i(\hat{n}_i-n_i^0)(n_i-n_i^0)\ , \tag{2}$$

where the operator $\hat{n}_i$ is defined by

$$\hat{n}_i=\sum_{\alpha=s,p}a_{i\alpha}^\dagger a_{i\alpha}\ ,$$

and n_i is the expectation value $n_i=\langle\hat{n}_i\rangle$, which gives the total valence charge at the atom i; n_i^0 is 4 for sp_3 TB orbitals. Adding this term to H makes geometries with large charge transfers unfavorable. Charge transfers are likely to occur at reconstructed semiconductor surfaces, so the addition of this term represents a significant improvement on the model, as will be shown later when the Si(001) surface is discussed. Also at that point we will discuss how we determine the value of U. Notice that if we make a mean-field approximation to the interaction energy from H_U, the resulting effective electronic Hamiltonian contains only one-particle operators. However, the one-particle equations must now be solved self-consistently.

The total energy of the system can then be expressed as the following sum:

$$E_{\text{tot}}=E_{BS}-H_U+V_{\text{elastic}}\ , \tag{3}$$

where E_{BS} denotes the sum over the eigenvalues of the filled one-particle states obtained from the self-consistent one-body Hamiltonian:

$$H_{\text{el}}=\sum_i\sum_{\alpha=s,p}(\varepsilon_\alpha+U_i)a_{i\alpha}^\dagger a_{i,\alpha}+\sum_{i,j}\sum_{\alpha,\beta=s,p}h_{i\alpha,j\beta}a_{i\alpha}^\dagger a_{j,\beta}\ , \tag{4a}$$

and U_i is the effective on-site repulsion defined by H_U:

$$U_i=U(n_i-n_i^0)\ . \tag{4b}$$

The second term in Eq. (3) corrects for the double counting of the electron-interaction energies defined by H_U in E_{BS}, and V_{elastic} is an empirical elastic potential that partly corrects for the double counting of the TB electron-electron interactions in E_{BS} and also models the ion-ion repulsion; we approximate this term with a nearest-neighbor interaction of the form

$$V_{\text{elastic}}=\sum_b V_1x_b+V_2x_b^2\ , \tag{5}$$

where the sum is over bonds b, and $x_b=d_b/d_0-1$ is the fractional deviation of the bond length d_b from the bulk equilibrium value $d_0=2.35$ Å. $V_1=-16.31$ eV is chosen so that in the diamond structure the total energy is a minimum at $d_b=d_0$, and $V_2=49.26$ eV is a free parameter that we use to fit the bulk optical phonon at Γ.

To study the lattice dynamics of a given system, its equilibrium geometry must first be found. To do this we impose the connectivity of the network that defines the structure and equilibrate it by following the forces defined by H, which include the Hellmann-Feynman forces:

$$F_{i\mu}=-\left\langle\frac{\partial H_{\text{el}}}{\partial x_{i\mu}}\right\rangle-\frac{\partial V_{\text{elastic}}}{\partial x_{i\mu}}\ ,$$

where the angular brackets denote the average over the ground-state density. The index i labels the atomic and μ Cartesian coordinates. This equilibration procedure has to be done self-consistently, since the electronic states enter into the interaction energy through the expectation values n_i. Once the equilibrium configuration has been found, we calculate the dynamical matrix given by the second derivatives with respect to structural degrees of freedom. Given that the structures that we study are periodic and have finite-range interactions, we can im-

mediately express the dynamical matrix in terms of derivatives with respect to spatially modulated displacements with wave vector $\mathbf{q}$, written $x_{i\mu}(\mathbf{q})$, rather than first calculate second-order forces in real space and then Fourier-transform to obtain the dynamical matrix. Following the breakup of H in terms of a "noninteracting" Hamiltonian $H_0=H_{\rm el}+V_{\rm elastic}$, and a repulsion Hamiltonian H_U, we write the dynamical matrix $\mathbf{D}(\mathbf{q})$ as

$$\mathbf{D}(\mathbf{q})=\mathbf{D}^0(\mathbf{q})+\mathbf{D}^U(\mathbf{q}) . \tag{6}$$

The first term of this equation can be itself naturally broken into two parts:

$$D^0_{i\mu,j\nu}(\mathbf{q})=\frac{\partial^2\langle H_{\rm el}\rangle}{\partial x^*_{i\mu}(\mathbf{q})\partial x_{j\nu}(\mathbf{q})}+\frac{\partial^2 V_{\rm elastic}}{\partial x^*_{i\mu}(\mathbf{q})\partial x_{j\nu}(\mathbf{q})} . \tag{7}$$

The contribution from $V_{\rm elastic}$ to $\mathbf{D}^0$ is straightforward to evaluate, and the first term in the above equation is calculated to second order in perturbation theory:

$$\frac{\partial^2\langle H_{\rm el}\rangle}{\partial x^*_{i\mu}(\mathbf{q})\partial x_{j\nu}(\mathbf{q})}=2\sum_n f_n\left\langle n\left|\frac{\partial^2 H_{\rm el}}{\partial x^*_{i\mu}(\mathbf{q})\partial x_{j\nu}(\mathbf{q})}\right|n\right\rangle+2\sum_{n,n'}\frac{f_n-f_{n'}}{E_n-E_{n'}}\left\langle n\left|\frac{\partial H_{\rm el}}{\partial x^*_{i\mu}(\mathbf{q})}\right|n'\right\rangle\left\langle n'\left|\frac{\partial H_{\rm el}}{\partial x_{j\nu}(\mathbf{q})}\right|n\right\rangle , \tag{8}$$

where $|n\rangle$ are the self-consistent electronic states (n labels bands and Brillouin-zone wave vector) and f_n are Fermi factors. The first term in Eq. (8) comes from the harmonic expansion of the electron-phonon interaction; the second term results from linear electron-phonon interaction taken to second order in perturbation theory, and represents the electronic polarization contribution to $\mathbf{D}$. The polarization term is the only part of $\mathbf{D}$ that has long-range character and therefore plays a fundamental role in correctly describing the vibrational excitations in semiconductors. In particular, the electron polarization can describe electronically driven structural instabilities.

The contribution to the dynamical matrix from the repulsion U screens the density fluctuations present in the polarization sum of Eq. (8). It should be noted though that some of the effects of U are already in $\mathbf{D}^0$ since the electronic states that enter Eq. (8) are calculated self-consistently. To evaluate $\mathbf{D}^U(\mathbf{q})$ we first calculate the density fluctuation at the atom l, δn_l, in linear response to the lattice fluctuations $\delta x_{j\nu}(\mathbf{q})$:

$$\delta n^{(0)l}_{j\nu}(\mathbf{q})=2\sum_{n,n'}\frac{f_n-f_{n'}}{E_n-E_{n'}}\langle n\,|\,\hat{n}_l(\mathbf{q})\,|\,n'\rangle \times\left\langle n'\left|\frac{\partial H_{\rm el}}{\partial x_{j\nu}(\mathbf{q})}\right|n\right\rangle . \tag{9}$$

The actual calculation of this term is done in parallel with the polarization sum in Eq. (8). Because of the interaction U this density fluctuation generates an interacting potential $U(\mathbf{q})$,

$$U(\mathbf{q})=U\sum_m \delta n_m(\mathbf{q}) ,$$

where the density fluctuations are themselves screened:

$$\delta n^l_{j\nu}(\mathbf{q})=\delta n^{(0)l}_{j\nu}(\mathbf{q})+2\sum_{n,n'}\frac{f_n-f_{n'}}{E_n-E_{n'}}\langle n\,|\,\hat{n}_l(\mathbf{q})\,|\,\hat{n}'\rangle \times\langle n'\,|\,U(\mathbf{q})\,|\,n\rangle .$$

Solving this set of equations self-consistently [random-phase approximation (RPA)], we obtain

$$\delta n^l_{j\nu}(\mathbf{q})=[\mathbf{I}-U\boldsymbol{\chi}(\mathbf{q})]^{-1}_{l,l'}\delta n^{(0)l'}_{j\nu}(\mathbf{q}) , \tag{10}$$

where χ is the density-density correlation function:

$$\chi_{l,l'}(\mathbf{q})=2\sum_{n,n'}\frac{f_n-f_{n'}}{E_n-E_{n'}}\langle n\,|\,\hat{n}_l(-\mathbf{q})\,|\,n'\rangle \times\langle n'\,|\,\hat{n}_{l'}(\mathbf{q})\,|\,n\rangle . \tag{11}$$

Finally, these fluctuations are combined to give $\mathbf{D}^U(\mathbf{q})$:

$$\mathbf{D}^U_{i\mu,j\nu}(\mathbf{q})=U\delta\mathbf{n}^{(0)\dagger}_{i\mu}(\mathbf{q})[\mathbf{I}-U\boldsymbol{\chi}(\mathbf{q})]^{-1}\delta\mathbf{n}^{(0)}_{j\nu}(\mathbf{q}) . \tag{12}$$

We have tested this model by comparing results with experiments and with results of other calculations. Among the Si systems we have used for comparisons are the diamond structure, the polytope BC8, the Si_2 molecule,[22] and the twin boundary Σ_q.[23] In general, there is good agreement between our results and those of experiments and other more accurate calculations (like density-functional theory), where these are available. The model reproduces the long-range forces that are responsible for the flat TA branch in the bulk phonons of Si; we mentioned that all the long-range behavior in our model is in the electron polarization. These tests were done with no repulsion U present; these results, however, have only a weak dependence on the value of U. This is not surprising since in fourfold-coordinated structures there are no appreciable charge oscillations, contrary to what happens on surfaces where there are dangling bonds. This implies that it is difficult to determine a value of the repulsion U using some bulk property.

III. Si(001) 2×1

A. Structure

We have reported studies of the structure and vibrational excitations of the Si(001) 2×1 surface without including the intra-atomic repulsion U.[19] We now present results of similar studies when on-site electron-electron Coulomb interactions are incorporated into the model in the form described in the preceding section. The consequences of the repulsion U on the structure, surface charge density, and phonon spectrum are analyzed in this section.

In the tilted dimer configuration charge is transferred from the lower to the higher surface atom; without the repulsion U this polar character of the surface dimer and the gap of the electronic surface states are substantially overestimated. The addition of the repulsion U to the electronic Hamiltonian corrects for this deficiency. This repulsion parameter U appearing in Eq. (1) represents an *effective* intra-atomic electron-electron interaction in the sense that with a single intra-atomic term it represents both the intra-atomic Coulomb repulsion U_0, or "bare" interaction, and the Madelung energy of interatomic interactions, K_0, which stabilizes charge transfers:

$$U = U_0 - K_0 \ .$$

Harrison[17] calculated the value of the bare intra-atomic repulsion and obtained $U_0 = 7.64$ eV. If one were to use this value, one has then to calculate the Madelung-energy contribution to the net electrostatic potential for each atom. This represents a significant computational complication for geometries with a large number of atoms per unit cell. Alternatively, one can use an effective U. Tomanek and Schlüter[24] have studied Si clusters using a model very similar to the one we are using here with an effective value of $U=2$ eV; they found that their results did not change noticeably if U is varied by ± 1 eV. We have performed our calculations for two values for the interaction U: 7.6 and 1.9 eV. The first value represents the limit of strong repulsion and the second is chosen to approximately reproduce experimentally measured electronic excitations, as will be explained below. By studying our results for these two values of U, and $U=0$, we can study the general consequences of adding this term to the electronic Hamiltonian, and observe trends as U is increased.

The electronic surface bands obtained in our calculation are shown in Fig. 2. The 2×1 surface has one occupied surface band and one unoccupied surface band. These are, roughly speaking, the bonding and antibonding combinations of the surface dangling bonds. We show in Fig. 2 four bands for each value of U because we model the surface with a slab geometry, which has two identical surfaces. For the present study we use a slab 10 layers thick. The fact that the bands show no splitting throughout most of the surface Brillouin zone (SBZ) indicates that surface-surface coupling is negligible for the 10-layer slab. A large splitting occurs only where surface states are coupled through bulk states. Both the average and the minimum gap between the occupied and the unoccupied surface states decrease as U increases. The arrow in Fig. 2 indicates the minimum gap E_G for $U=7.6$ eV. This gap is from $\bar{\Gamma}$ to $\bar{K}$ for the three values of U that we consider. In the absence of intra-atomic repulsion, we obtain $E_G=0.7$ eV. In the other extreme, for $U=7.6$ eV, $E_G=0.03$ eV. The first value is an overestimate and the second one is too small (metal-like states are not experimentally observed[3,9,10]). With $U=1.9$ eV we obtain $E_G=0.27$ eV; this is a reasonable number according to infrared[5] and high-resolution electron-energy-loss[4] (HREELS) studies of the Si(001) surface.

Even though there is some experimental evidence that the surface dimers are tilted, the question remains whether it is an artifact of the tight-binding model, without the repulsion U, that the symmetry of these dimers is broken. In this work we find that even with moderately strong repulsion the surface dimers remain tilted. The charge transfer δe between the atoms forming a surface dimer,

$$\delta e = (\delta e_1 - \delta e_2)/2 \ , \tag{13}$$

where $\delta e_{1,2}$ is the change in charge at atom 1,2 from the bulk value (4 for sp_3 orbitals), and 1 and 2 denote the up and down atoms of a surface dimer, strongly depends on the value of U: $\delta e/e = 0.39$, 0.23, and 0.08 for $U=0$, 1.9, and 7.6 eV, respectively. However, even for small charge transfers, as occurs in the strong-repulsion limit, the value of the tilt z_0 remains fairly constant, whereas δe and E_G

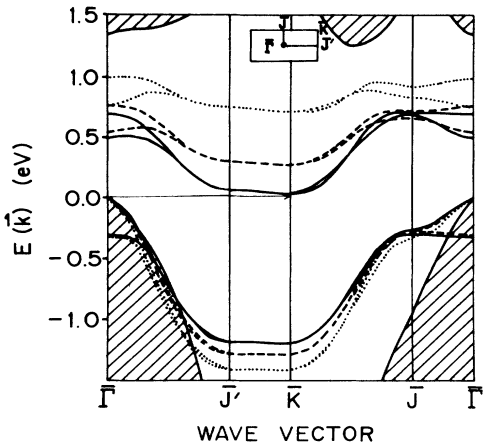

FIG. 2. Surface electron bands of the Si(001) 2×1 surface along a symmetry direction on the SBZ (shown in the inset). Dashed areas correspond to bulk bands projected onto the SBZ. The solid, dashed, and dotted lines correspond to $U=7.6$, 1.9, and 0 eV, respectively. Four bands are shown in each case corresponding to the two identical surfaces present in our slab calculation. Each surface has one empty and one occupied surface band. The arrow indicates the minimum gap E_G, which is indirect from $\bar{\Gamma}$ to $\bar{K}$ for the three values of repulsion U. Electron energies are in eV.

TABLE I. Electronic-surface-state minimum gap E_G, surface-dimer tilt z_0, and charge transfer δe of equilibrium surfaces with different values of repulsion U. E_G is indirect ($\bar{\Gamma}$ to $\bar{K}$) for the 2×1 surface, and direct at $\bar{\Gamma}$ for both $p(2\times2)$ and $c(4\times2)$.

		Si(001)		
Reconstruction	U (eV)	z_0 (Å)	E_G (eV)	$\delta e/e$
2×1	0.0	0.59	0.71	0.39
	1.9	0.58	0.27	0.23
	7.6	0.53	0.03	0.08
$p(2\times2)$	0.0	0.63	0.83	0.36
	1.9	0.64	0.61	0.15
$c(4\times2)$	0.0	0.63	0.84	0.36
	1.9	0.64	0.61	0.16

decrease as U increases. These results are summarized in Table I. Thus the on-site repulsion reduces the charge transfer, but is relatively ineffective in reducing the tilt.

B. Phonons

The phonon spectrum of the Si(001) 2×1 surface, calculated for $U=0$, is shown in Fig. 3. Solid and dashed lines correspond to surface phonons and surface resonances, respectively. We have presented a detailed analysis of some aspects of this spectrum before,[19] so here we will only review the main features, and we will concentrate on the effects of the repulsion U on this phonon spectrum. Besides the usual Rayleigh wave that splits off the bottom of the acoustical continuum, the surface supports modes that are characteristic of the dimer reconstruction. Among these we find in the optical region a mode that has the form of stretching of the surface dimers at $\bar{J}'$ and $\bar{K}$ and appears throughout the SBZ. Along the zone edge $\bar{J}'$ to $\bar{K}$ it is a surface phonon, and a strong resonance elsewhere. There is also a surface mode off the top of the optical continuum that corresponds to a subsurface mode related to the fivefold rings of the reconstruction. A similar fivefold ring mode is observed in our calculated phonon spectrum of the Si(111) 2×1 surface.[21] A particularly important vibrational mode is a rocking oscillation of the surface dimers, also observed at the zone edge, where one of the dimer atoms moves out of the surface and the other atom moves into the surface; we call this the rocking mode. It is found at the lower edge of the first energy gap along the zone edge at 25.0 meV, with almost no dispersion; near $\bar{K}$ it is a true surface mode, but it is a resonance elsewhere. This is an important mode for two reasons: First, this vibrational excitation modulates the gap E_G of the electronic surface states, symmetric dimers produce a metallic surface, and, as the asymmetry is broken by tilting the dimers, a gap is opened in the electron surface states; second, this mode clearly is related to the formation of ordered arrangements of the tilted surface dimers with larger periodicities than 2×1. The energies of these modes, and their symmetry with respect to the mirror plane perpendicular to the $\langle 110\rangle$ direction (the only symmetry of the surface), are listed in Table II.

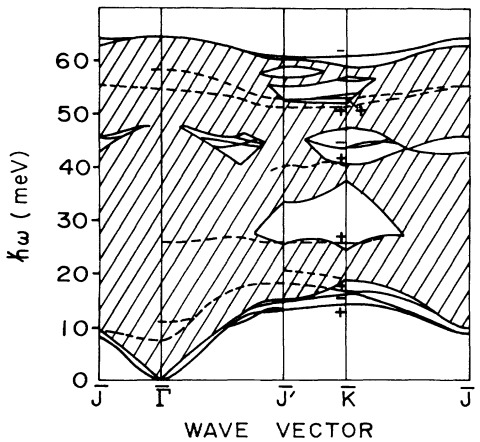

FIG. 3. Phonon spectrum of the Si(001) 2×1 surface calculated with no repulsion, U, present. Dispersions are shown along the symmetry directions of the SBZ indicated in the inset in Fig. 1. The dashed areas correspond to bulk phonons projected onto the SBZ. Solid lines represent surface phonons and dashed lines surface resonances. Phonon energies are in meV.

We now compare these results with those obtained with $U=1.9$ and 7.6 eV. There are two competing effects in the lattice dynamics introduced by U: On one hand, since the ionic character of the surface dimers is reduced by the presence of electron interactions, the value of E_G, and, in general, the average gap, decrease. Therefore the electronic polarization contribution to the restoring forces is enhanced [note the energy denominators in Eq. (8)]. On the other hand, the charge-density fluctuations in this polarization sum are self-consistently screened, which tends to reduce the polarizability. In order to isolate and study the latter effect, we have separately calculated the phonon spectrum of the surface with and without the screening term in the dynamical matrix [Eq. (12)], but with the self-consistent electronic states obtained with U present.

The phonon spectra of the surface for $U=1.9$ and 7.6 eV are shown in Figs. 4 and 5, respectively. The top panels in both figures show the spectra of $\mathbf{D}^0$, and the lower panels show the results with screening included ("complete" calculation). First, let us concentrate on the general consequences of the addition of repulsion U to the model with screening included. Comparing both Figs. 4(b) and 5(b) with the $U=0$, phonons, we see that most of the surface features remain essentially unchanged: the fivefold ring mode appears at the same energy in the three cases; the stretching mode shows similar dispersion across the SBZ for the three values of U, only shifting down in energy by a few meV along the zone edge $\bar{J}'$ to $\bar{K}$, where it drops below the energy gap but remains a strong surface resonance. But contrary to these small changes, the rocking mode seen for the $U=0$ case disappears from the lower edge of the energy gap along the zone edge for both $U=1.9$ and 7.6 eV, and we observe an increase in the number of surface-localized modes near the bottom of the acoustical continuum: the rocking mode undergoes a negative shift in energy and mixes with bulk modes and the Rayleigh wave. As suggested above, it is not surprising that the vibrational excitation that is most affected by the presence of the repulsion U is the rocking mode, since it modulates the structural parame-

TABLE II. Energies (meV) of surface phonons and resonances (R) at $\bar{J}'$ of the Si(001) 2×1 surface for three values of the repulsion U. Symmetry is with respect to the mirror plane perpendicular to the $\langle 110\rangle$ direction: + is even, − is odd.

Vibrational character	U (eV) 0.0	1.9	7.6
Rayleigh (+)	13.6	13.6	13.5
Rocking (+)	25.7	21.8 (R)	19.6 (R)
Stretch (+)	53.4	52.3 (R)	50.9 (R)
Fivefold ring (−)	61.1	61.4	61.4

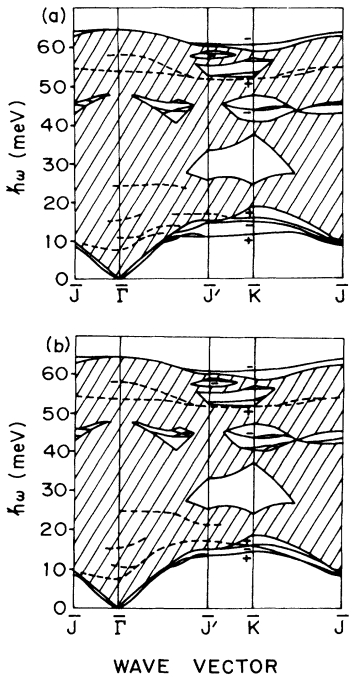

FIG. 4. Phonon spectrum of the Si(001) 2×1 surface calculated with repulsion $U=1.9$ eV. The top panel corresponds to phonons obtained without the effect of screening included in the dynamical matrix [Eq. (10)], and the lower panel corresponds to the screened phonons.

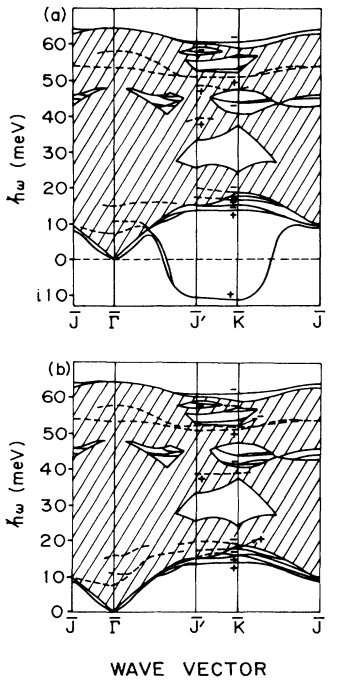

FIG. 5. Si(001) 2×1 phonon spectrum with $U=7.6$ eV. The top and lower panels correspond to unscreened and screened phonons, respectively. Notice the unstable mode (imaginary frequency) along the zone edge $\bar{K}$ to $\bar{J}'$ in the unscreened phonons.

ters that most strongly couple to E_G and the electronic surface-state polarizability. The energies of the relevant modes are listed in Table II; the results obtained with $U=1.9$ eV can be considered our best predictions. The energies listed for the rocking mode for the nonzero values of U correspond to the mixed mode with the largest rocking amplitude.

The unscreened phonons [Figs. 4(a) and 5(a)] show that, similar to the screened phonons, most of the features do not change considerably as U is turned on. The rocking mode is an important exception: For $U=1.9$ eV we see that a fairly flat mode splits off the bottom of the acoustical continuum along the zone edge $\bar{J}'$ to $\bar{K}$ with energy around 10 meV; this mode is a combination of the Rayleigh wave and the rocking mode. In the strong-repulsion limit $U=7.6$ eV we see an unstable mode along the zone edge: this is the rocking mode; the Rayleigh wave has returned to its original position. So the general picture is that as the repulsion U increases, the electron polarizability drives the rocking mode towards lower frequencies, and it passes through the bottom of the acoustical continuum where it mixes with the Rayleigh wave, but as U is increased further it recovers its rocking character and drops in energy until it reaches a point of critical softening, inducing an electronically driven structural instability (recall that we are leaving out screening at the moment). The competing effects of enhanced polarizability versus screening introduced by the repulsion U are quite strong when taken separately, but for most of the phonon modes they cancel and generate only small energy shifts. For the rocking mode, however, this small shift is enough to completely change its character as a strong resonance.

Note that the rocking mode becomes unstable only at the zone edge $\bar{J}'$ to $\bar{K}$, corresponding to the wave vector of the minimum indirect gap E_G. This mode is almost dispersionless along the zone edge, having slightly lower energy at $\bar{K}$ (same as the indirect electronic gap). At $\bar{J}'$ and $\bar{K}$ the rocking mode has the symmetry of the $p(2\times2)$ and $c(4\times2)$ reconstruction, respectively, and along the zone edge it corresponds to a modulated structure with a higher-order periodicity along the $\langle 110\rangle$ direction, perpendicular to the rows of surface dimers. This indicates that the surface, through the surface electron polarizability, has a strong tendency to form alternating surface dimers along the $\langle \bar{1}10\rangle$ direction (along a row of dimers) and a somewhat weaker tendency to forming alternating dimers along the $\langle 110\rangle$ direction, parallel to the dimer rows. We can confront these results with the experimental findings of Kevan on the Ge(001) surface. He found a two-stage phase transition from a disordered phase at high temperature to an ordered phase at low temperatures. The first stage, at about 250 K, corresponds to the ordering of alternating surface dimers along the dimer rows; the second stage, at about 220 K, corresponds to the sharpening of the $c(4\times2)$ spots in the low-energy electron diffraction (LEED) pattern; that is, alternating dimers along the direction perpendicular to the rows of surface dimers. Clearly, the structural instabilities implicit in the electron polarizability that we find are consistent with Kevan's results. Contrary to this, Ihm *et al.* found, using a spin Hamiltonian to represent the surface, that there is a

single-stage phase transition from the disordered phase to either the $p(2\times2)$ or $c(4\times2)$ phase.

IV. HIGHER-ORDER RECONSTRUCTIONS: $c(4\times2)$ and $p(2\times2)$

As we mentioned in the Introduction, there have been a number of experimental and theoretical studies that indicate that at low temperatures the periodicity of the Si(001) surface is $c(4\times2)$ or $p(2\times2)$, rather than 2×1. Both of these two higher-order periodicities of the surface, within the tilted-dimer model, correspond to zone-folding reconstructions of the 2×1 surface: In the $p(2\times2)$ reconstruction the $\bar{J}'$ point at the zone edge folds back to $\bar{\Gamma}$, and in the $c(4\times2)$ reconstruction the zone corner $\bar{K}$ folds back to the zone center. In analogy with the Ising model, as used by Ihm *et al.*[15] to study this surface, the 2×1, $c(4\times2)$, and $p(2\times2)$ reconstructions correspond, respectively, to the ferromagnetic, antiferromagnetic, and layered-antiferromagnetic phases of the Ising spins, where the spin direction represents a tilted surface dimer. In agreement with their calculations and those of Chadi,[8] both based on a similar TB approach as the one we are using here, we find that both the $c(4\times2)$ and $p(2\times2)$ reconstructions are lower in energy than the 2×1 surface. The gains in energy are 0.15 and 0.13 eV per surface dimer, respectively. The value of U that we used in these calculations, and that we will use for the rest of the paper, is $U=1.9$ eV; according to our discussion in the preceding section this is our best estimate of the effective on-site repulsion on the Si(001) surface. Without the repulsion U we find that the energy differences are larger by approximately a factor of 2. It is fair to say that energies obtained with the TB model are not reliable enough to determine stable configurations when energy differences are small, as is the case here: we can only say that the $c(4\times2)$ and $p(2\times2)$ reconstructions have essentially the same energy. This is consistent with the experimental observation of superimposed periodicities of the surface at low temperatures, and with the nearly flat dispersion of the unscreened unstable rocking mode along the zone edge $\bar{J}'$ to $\bar{K}$ discussed in the preceding section. However, the screened phonons in the 2×1 structure do not show any instability. We conclude that the driving force for the formation of the higher-order reconstructions is not a true Fermi-surface instability, but a relaxation of the entire surface band. A similar conclusion was reached by Kevan[7] in his experimental studies of Ge surfaces. As is evident in the dispersion of the electronic surface bands (Fig. 2) (strong dispersion from $\bar{\Gamma}$ to $\bar{J}'$ and weak from $\bar{J}'$ to $\bar{K}$), the interaction between dimers along the rows is strong, whereas across the rows it is weak. Since the electronic surface-state polarizability is then large for wave vectors corresponding to the SBZ edge, this suggests that the tendency to have alternating dimers along the $\langle\bar{1}10\rangle$ direction is strong, but along the $\langle110\rangle$ direction the driving force to form alternating dimers is not as strong. That is, the strong driving force on the reconstruction of the 2×1 surface is to have alternating tilted dimers along the rows of dimers, and there is a weaker tendency to form alternating dimers also between the rows. This is consistent with the fact that the energies of the $c(4\times2)$ and $p(2\times2)$ reconstructions are very similar.

An important consequence of the zone-folding reconstruction in both $c(4\times2)$ and $p(2\times2)$ is that the electronic gap E_G, which is indirect for the 2×1 surface from $\bar{\Gamma}$ to the zone edge $\bar{J}'$ to $\bar{K}$, is now direct at the zone center. The minimum and the average direct gaps are then lower for the two higher-order periodicity surfaces than for the 2×1 surface. This is relevant for optical transitions, and in the next section we will discuss the important effect this has on the infrared absorption of the surface. However, the effect of the zone-folding reconstruction on the direct gaps is somewhat reduced by the fact that the states at the bottom of the empty surface state bands and the states at the top of the valence bands, both occurring now at $\bar{\Gamma}$, can couple through a reciprocal-lattice vector of the superlattice, and therefore the gap widens: instead of having a direct gap of 0.27 eV (the indirect gap of the 2×1 surface), we obtain $E_G=0.61$ eV for both $c(4\times2)$ and $p(2\times2)$. It is still true, however, that the minimum and average gaps are reduced (by about 30%).

We have calculated the zone-center phonons (with $U=1.9$ eV) for the $c(4\times2)$ and $p(2\times2)$ surfaces. In the first case we find that the zone-center phonons essentially correspond to the $\bar{\Gamma}$ and $\bar{K}$ phonons of the 2×1 surface, following the zone-folding reconstruction. Thus, for example, there are two stretching modes, one where the two dimers of the unit cell move in phase (zone center in 2×1) and the other when they move out of phase (zone corner in 2×1); energy shifts with respect to the 2×1 surface are minimal. Similarly, the zone-center modes of the $p(2\times1)$ surface correspond to the $\bar{\Gamma}$ and $\bar{J}'$ modes of the 2×1 surface. In the next section, where we use the vibrational modes that we have calculated to obtain the one-phonon absorption spectrum of the 2×1, $c(4\times2)$, and $p(2\times2)$ surfaces, we will focus on a number of these modes.

V. DIPOLE ACTIVITY OF SURFACE PHONONS

One of the advantages of the theoretical model that we use, with its explicit treatment of electron-phonon coupling, is that one can calculate response functions that probe the interactions between the electronic surface charge density and vibrational excitations. Specifically, in this section we study the phonon-assisted contribution to the surface conductivity, or dipole activity of the surface phonons, of the 2×1, $c(4\times1)$, and $p(2\times2)$ reconstructions of the Si(001) surface. These studies provide an insight into the nature of the electronic states and bonding at the surface, and their relation to structural degrees of freedom. Also, the one-phonon absorption spectra that we obtain can be used to characterize the periodicity of the surface. We have previously presented similar studies on the Si(111) 2×1 surface,[21] where an explanation for the anomalously large dipole-active surface phonon seen in electron-energy-loss-spectroscopy (EELS) experiments was proposed.[25,26] At the end of this section we will briefly comment on some of the similarities and differences between the Si(111) and Si(001) surfaces. Next we briefly outline the procedure we use to calculate surface conductivities, and then present and discuss the results.

A. Formalism

To calculate the dipole activity of the surface, we start by calculating the dynamic response of the current density to a perturbation generated by a lattice fluctuation. Then the current-current correlation function gives the surface conductivity (Kubo's formula). The starting point is then the perturbing potential:

$$H^1(t)=h_i e^{i\omega t}+h_i^+ e^{-i\omega t} , \tag{14a}$$

where $h_i=\partial H/\partial x_i$, and x_i is an ionic degree of freedom (in contrast with the notation used in Sec. II, i here denotes both atom number and Cartesian coordinate). H is the Hamiltonian defined in Eq. (1). Only optical transitions are relevant in the surface conductivity, so that the lattice fluctuation has no spatial modulation ($\mathbf{q}=\mathbf{0}$); this implies that H_μ is real:

$$H^1(t)=h_i(e^{i\omega t}+e^{-i\omega t}) . \tag{14b}$$

Using time-dependent perturbation theory, we obtained the following expression for the linearized response of the current density (after a few straightforward manipulations):

$$\langle \mathbf{J}(t)\rangle_i=\sum_{n,n'}(f_n-f_{n'})\left[e^{i\omega t}\left[\frac{\mathbf{J}(n,n')h_i(n',n)}{E_n-E_{n'}+\omega}+\frac{\mathbf{J}(n',n)h_i(n,n')}{E_n-E_{n'}-\omega}\right]+e^{-i\omega t}\left[\frac{\mathbf{J}(n,n')h_i(n',n)}{E_n-E_{n'}-\omega}+\frac{\mathbf{J}(n',n)h_i(n,n')}{E_n-E_{n'}+\omega}\right]\right] , \tag{15}$$

where the indices n,n' denote the band index and Brillouin-zone vector of the electronic states, f_n are Fermi factors, and $\mathbf{J}(n,n')$, $h_i(n,n')$ are matrix elements of the current operator $\mathbf{J}$ and h_i:

$$\mathbf{J}(n,n')=\langle n\,|\,\mathbf{J}\,|\,n'\rangle,\quad h_i(n,n')=\langle n\,|\,h_i\,|\,n'\rangle .$$

Since the driving frequency ω corresponds to a phonon frequency, we can expand the energy denominators in terms of the small parameter $\omega/(E_n-E_{n'})$. Keeping this expansion to linear order, one obtains

$$\langle \mathbf{J}(t)\rangle_i=\sum_{n,n'}(f_n-f_{n'})\left[\frac{2\cos(\omega t)}{E_n-E_{n'}}[\mathbf{J}(n,n')h_i(n',n)+\mathbf{J}(n',n)h_i(n,n')]+\frac{2i\omega\sin(\omega t)}{(E_n-E_{n'})^2}[\mathbf{J}(n,n')h_i(n',n)-\mathbf{J}(n',n)h_i(n,n')]\right] . \tag{16}$$

Notice that the time dependence of the first term in this expression is in phase with the perturbing potential, whereas the second term is out of phase. As long as the driving frequency is slower than the fastest response of the system, as it is since $\omega<E_G$, the in-phase term ought to vanish. Indeed, the Brillouin-zone integration in Eq. (16), with time reversal, yields this result [if higher-order terms in the $\omega/(E_n-E_{n'})$ expansion are kept, some in-phase terms will not vanish after the $\mathbf{k}$-space integration].

Writing the oscillating current density in terms of a dipole moment, it defines a dynamic charge $\mathbf{e}^*$ through the following equation:

$$\langle \mathbf{J}(t)\rangle_i=\mathbf{e}_i^*\dot{x}_i(t)=\mathbf{e}_i^* x_i^0 i\omega(e^{i\omega t}-e^{-i\omega t}) , \tag{17a}$$

or, using Eq. (16),

$$\mathbf{e}_i^*=\sum_{n,n'}\frac{i(f_n-f_{n'})}{(E_n-E_{n'})^2}[\mathbf{J}(n,n')h_i(n',n)-\mathbf{J}(n',n)h_i(n,n')] . \tag{17b}$$

It is convenient to express the current density response in terms of the dynamic charges $\mathbf{e}_i^*$ because these can then be used to obtain the dynamic charges associated with each vibrational mode, as will be explained below. But before we do this, we have to calculate $\mathbf{e}_i^*$ as given by Eq. (17b). The current operator $\mathbf{J}$ is defined in TB theory through the following commutator:

$$\mathbf{J}=-ie[\mathbf{r},H] , \tag{18}$$

where the position operator $\mathbf{r}$ is itself given by

$$\mathbf{r}=\sum_{j=\text{atom site}}\left[\sum_{\alpha=s,p}\mathbf{R}_j a_{j,\alpha}^\dagger a_{j,\alpha}+\sum_{\mu=x,y,z}\lambda_0(a_{j,s}^\dagger a_{j,p_\mu}+\text{c.c.})\right] .$$

The on-site s-p—polarization constant λ_0 we use is determined by the condition that in tetrahedrally coordinated Si the matrix element of the dipole operator between sp_3 hybrids is approximately half the bond length; that is, we are assuming that the interatomic distance is determined by the spatial extent of the electronic orbitals:

$$\langle h_z\,|\,z\,|\,h_z\rangle=d_0/2 ,$$

where $|\,h_z\rangle$ is an sp_3 hybrid along the $\mathbf{z}$ direction:

$$|\,h_z\rangle=\tfrac{1}{2}(|\,s\rangle+\sqrt{3}\,|\,p_z\rangle) .$$

This gives

$$\lambda_0\equiv\langle s\,|\,z\,|\,p_z\rangle=d_0/\sqrt{3} .$$

By using this condition to fit λ_0, we are incorporating some "solid-state" information into it, rather than obtaining it from a free-atom calculation. In any case, the precise value of λ_0 is not very important: we have varied λ_0 by as much as 15%, obtaining results that vary by not more than a couple of percent.

Note that in an equivalent formulation where static per-

turbation theory is used on the dipole moment $\mathbf{e}^*\mathbf{r}$ induced by a lattice distortion, the relation

$$\langle n \mid \mathbf{r} \mid n' \rangle = \frac{\langle n \mid [\mathbf{r},H] \mid n' \rangle}{E_{n'} - E_n}$$

is used to evaluate the dipole-moment matrix elements. This brings us back to our present formulation.

Having made these remarks, we write $\mathbf{e}_i^*$ as an expression that can be conveniently evaluated:

$$\mathbf{e}_i^* = \sum_{n,n'} \frac{f_n - f_{n'}}{(E_n - E_{n'})^2} [\langle n \mid [\mathbf{r},H] \mid n' \rangle h_i(n',n) - \langle n' \mid [\mathbf{r},H] \mid n \rangle h_i(n,n')] \ . \quad (19)$$

If these dynamic charges are now projected onto the normalized displacement field $Q_i^{(n)}$ of the nth phonon, one obtains the dynamic charge $\mathbf{e}_n^*$ associated with that particular mode:

$$\mathbf{e}_n^* = e \sum_i \mathbf{e}_i^* Q_i^{(n)} \ . \quad (20)$$

The phonon-assisted contribution to the surface conductivity then follows:

$$\mathrm{Re}\sigma_\mu^T(\omega) = \frac{\pi}{2MA} \sum_n |(\mathbf{e}_n^*)_\mu|^2 [\delta(\omega - \omega_n) + \delta(\omega + \omega_n)] \ , \quad (21)$$

where the sum is over $\mathbf{q}=\mathbf{0}$ phonon modes, M is the ionic mass (28 amu), and A is the surface unit-cell area. The index μ denotes Cartesian coordinate, so that we calculate $\sigma_x^T(\omega)$, $\sigma_y^T(\omega)$, and $\sigma_z^T(\omega)$. The directions of the $\mathbf{x},\mathbf{y},\mathbf{z}$ axes with respect to the surface are defined below. The superscript in σ_μ^T indicates that this is the transverse conductivity: transverse charge fluctuation are not screened. But the phonons and self-consistent electron wave functions that enter into the calculation of $\sigma^T(\omega)$ are still obtained with the repulsion U present ($=1.9$ eV). To calculate the longitudinal conductivity $\sigma^L(\omega)$, where the charge fluctuations are screened, we include the following screening correction in Eq. (19):

$$[\mathbf{r},H] \rightarrow (\mathbf{I} - U\chi)^{-1}[\mathbf{r},H] \ .$$

However, we found that the longitudinal and transverse surface conductivities are nearly identical; the direct gaps that enter into the calculation are large enough so that screening has only a small effect. In the next section we will report our results for $\sigma^T(\omega)$, but we will call it simply $\sigma(\omega)$ understanding that the same results correspond to the transverse and longitudinal conductivities.

B. Results

The surface conductivities $\sigma(\omega)$ of the 2×1, $p(2\times2)$, and $c(4\times2)$ reconstructions of Si(001) are shown in Fig. 6. For each reconstruction we show three spectra, corresponding to the three Cartesian coordinates defined as follows: $\mathbf{x}\rightarrow[\bar{1}10]$ (solid line), $\mathbf{y}\rightarrow[110]$ (dashed line), and $\mathbf{z}\rightarrow[001]$ (dashed-dotted line). That is, the polarization of $\sigma_x(\omega)$ is along the rows of surface dimers, $\sigma_y(\omega)$ is perpendicular to these rows, and $\sigma_z(\omega)$ is normal to the surface. We will discuss first the results for the 2×1 surface and then compare them with those of the higher-order reconstructions.

The analysis of the absorption spectra of the surface phonons of the 2×1 surface can be summarized in two points: (1) At low frequencies (20—35 meV) there are a series of peaks that correspond to modes whose vibrational character is a mixture of rocking of the surface dimers and shear of these dimers with respect to the subsurface. We mentioned previously that the $\bar{\Gamma}$ phonons of this sur-

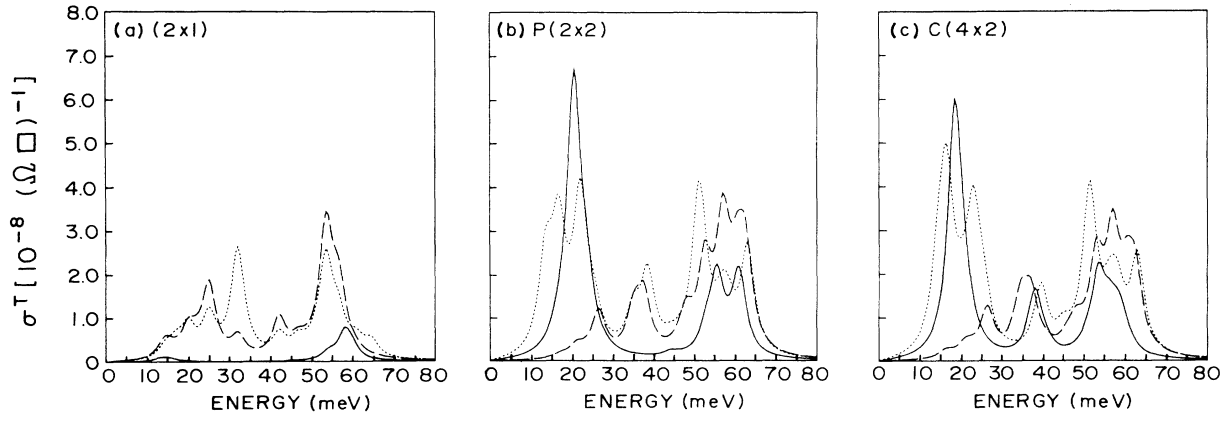

FIG. 6. Surface conductivity $\sigma(\omega)$ of the three reconstructions of the Si(001) surface: (a) 2×1, (b) $p(2\times2)$, and (c) $c(4\times2)$. The three polarizations of $\sigma(\omega)$ are shown in each case: x, solid line; y, dashed line; z, dashed-dotted line. These directions correspond to the $\langle\bar{1}10\rangle$, $\langle110\rangle$, and $\langle001\rangle$ crystallographic axes, respectively. The discrete absorption peaks associated with each vibrational mode were convoluted with a Gaussian with a full width at half maximum (FWHM) equal to 5 meV, comparable to present-day high-resolution spectrometers, to obtain the continuum spectra shown in this figure. The horizontal axis corresponds to phonon energies (in meV).

face include a number of modes with some rocking vibrational character (in the same energy range as above), rather than a unique rocking mode (recall that we are using electronic states and phonons corresponding to $U=1.9$ eV). The polarization of the dipole activity generated by these rock-shear modes is along the **y** and **z** axes. (2) Centered around 55 meV there is a broad peak in both $\sigma_y(\omega)$ and $\sigma_z(\omega)$; this feature is generated not by true surface phonons or resonances, but by bulk optical phonons which become dipole active due to their reflection from the surface potential that breaks the inversion symmetry that forbids dipole excitations in bulk Si. This is a feature that is independent of the particular reconstruction and should appear on any Si surface. Indeed, on all the surfaces we have studied, including different reconstruction models of the Si(111) 2×1 surface, we obtain the same broad absorption peak at 55 meV. In summary, the envelope of the three polarizations of $\sigma(\omega)$ yields a spectrum with two broad features of approximately the same integrated intensity, one centered at about 25 meV and the second, somewhat sharper, centered at 55 meV. Notice that $\sigma_x(\omega)$ is essentially zero compared with $\sigma_y(\omega)$ and $\sigma_z(\omega)$. That is, for the 2×1 surface there is very little dipole absorption with polarization along the rows of surface dimers. This will be an important point of comparison with the $c(4\times2)$ and $p(2\times2)$ surfaces.

In Figs. 6(b) and 6(c) we show $\sigma(\omega)$ for the $p(2\times2)$ and $c(4\times2)$ surfaces, respectively. The first thing one observes is that the spectra of these two surfaces are very similar. Considering the fact that the alternating ordering of the surface dimers along the $\langle\bar{1}10\rangle$ direction is the strong feature of the higher-order reconstructions [common to both $p(2\times2)$ and $c(4\times2)$], compared to the weaker tendency to also reconstruct along the $\langle110\rangle$ direction, it is not surprising that we obtain essentially the same results for both surfaces. There are two main features in $\sigma(\omega)$ that differentiate the higher-order periodicity surfaces from the 2×1 surface: the first is the overall intensity of absorption, and the second is the appearance of structure in $\sigma_x(\omega)$ for the higher-order surfaces. We will analyze each one of these points next.

(1) The overall intensity of the absorption spectra of the $c(4\times2)$ and $p(2\times2)$ surfaces is larger than that for the 2×1 surface by approximately a factor of 2. This increase in the absorption intensity is a result of the zone-folding reconstruction that takes the 2×1 surface to either $c(4\times2)$ or $p(2\times2)$: the indirect gap E_G of the surface electron states in the 2×1 reconstruction becomes direct in the $c(4\times2)$ and $p(2\times2)$ surfaces, as we discussed earlier. Therefore the minimum and average *direct* gaps in the higher-order reconstructions are smaller than in the 2×1 surface, enhancing the polarization sum contribution to the dynamic charges $\mathbf{e}^*$ [Eq. (19)]. In fact, the energy denominators that appear in Eq. (21), used to evaluate $\sigma(\omega)$, appear to the fourth power. Thus the approximately 30% reduction in the direct gap accounts for the twofold increase in the intensity of $\sigma(\omega)$.

Aside from this effect, a number of vibrational modes of the 2×1 surface that occur at $\bar{J}'$ or at $\bar{K}$, and therefore do not contribute to the 2×1 surface conductivity, become zone-center modes of the $p(2\times1)$ and $c(4\times2)$ surfaces, respectively. This leads to a series of new absorption peaks that contribute, to a lesser degree, to the overall increase of absorption intensity. Among these the most prominent is the stretching mode of the surface dimers, with energy 50 meV. As it is to be expected, the dynamic charge carried by this mode ($\approx0.16e$ per surface dimer) is polarized along the $\langle110\rangle$ direction, parallel to the surface dimers.

(2) In the case of the 2×1 surface, the conductivity polarized along the rows of surface dimers was found to be essentially zero. Contrary to this, the *strongest* peak in the absorption spectra for both the $p(2\times2)$ and $c(4\times2)$ surfaces is in the x polarization. This peak is located at 20.3 meV for the $p(2\times2)$ surface and at 19.5 meV for the $c(4\times2)$ surface. In both surfaces this absorption peak is generated by a vibrational mode, which has the same form, in both cases, along the rows of surface dimers. The only difference is that on the $c(4\times2)$ surface the motion of the atoms is out of phase from row to row (i.e., along the $\langle110\rangle$ direction), whereas in the $p(2\times2)$ surface the displacement of the atoms across the rows of surface dimers is in phase. In Fig. 7 we schematically show the displacement field of this phonon along one of the rows of surface dimers. It is interesting to note that it has a large amplitude on the subsurface layer, where the

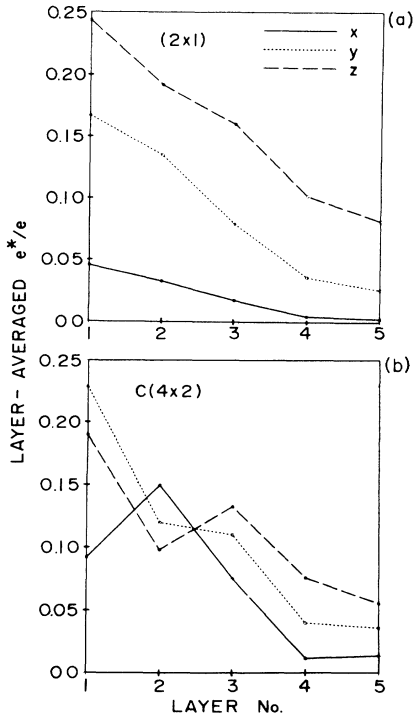

FIG. 7. Layer-averaged dynamic charges $\mathbf{e}_L^*$ as a function of penetration into the bulk. $\mathbf{e}_L^*$ are normalized over the number of atoms per layer. The dynamic charges that are averaged are those defined in Eq. (19). The three polarizations, x, y, and z, are represented by the solid, dashed, and dashed-dotted lines, respectively. The top panel corresponds to the 2×1 surface and the lower to the $c(4\times2)$ surface. The results for the $p(2\times2)$ surface are nearly identical to those for the $c(4\times2)$ surface.

atoms have no dangling bonds; nevertheless, it carries the largest dynamic charge ($\approx 0.26e$ per surface dimer) of all the vibrational modes supported by the surface. We can understand this by studying the layer-averaged dynamic charges $\mathbf{e}_L^*$ defined as

$$\mathbf{e}_L^* = \left[\sum_i (\mathbf{e}_i^*)^2 \right]^{1/2}, \tag{22}$$

where the sum is over the atoms in the Lth layer. We show the results in Fig. 7. The top panel in this figure shows the results for the 2×1 surface: the x component of $\mathbf{e}_L^*$ is much smaller than the dynamic charges along the other two Cartesian components. Notice also that $\mathbf{e}_L$ decays rather slowly as a function of penetration into the bulk. In the lower panel we show the equivalent results for the $c(4\times2)$ surface [for the $p(2\times2)$ surface, the results are almost identical]. The top and lower panels of this figure are rather different; in the case of the $c(4\times2)$ surface the x-polarized dynamic charges are large in the top two layers, and they are even larger in the second layer than at the surface. The vibrational mode shown in Fig. 8 couples very efficiently to these x-polarized dynamic charges, with large vibrational amplitudes in the subsurface layer. Within the dimer model of the reconstruction of the Si(001) surface there are no connected networks of *surface* atoms along the $\langle \bar{1}10 \rangle$ direction (or any other direction); one might suspect then that the charge oscillations generated by the mode of Fig. 8 occur in the subsurface. Indeed, we have calculated the charge oscillations that occur *only* along the surface backbonds, and found that most of the dynamic charge associated with the mode shown in Fig. 8 occurs along these backbonds; these backbonds form the connected network of atoms running along the $\langle \bar{1}10 \rangle$ direction that is closest to the surface. This subsurface charge oscillation is a rather interesting feature, considering that the gap states are expected not to be located at the subsurface. Of course, this, together with the slow decay of $\mathbf{e}_L^*$, suggest that the electronic surface states are not very well localized at the surface, but leak into the bulk. Slow decay of dynamic charges is observed also for the 2×1 surface [Fig. 7(a)], and thus we can conclude that its surface states have the same localization character as in the $c(4\times2)$ and $p(2\times2)$ surfaces.

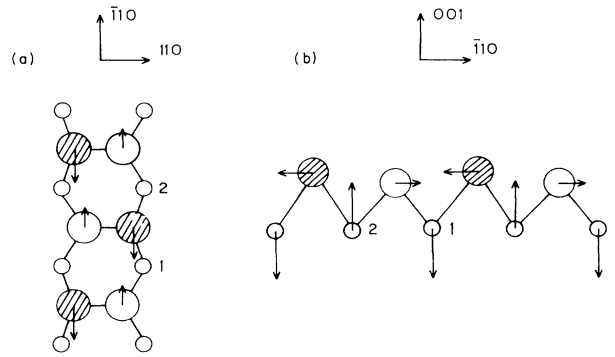

FIG. 8. The vibrational mode shown in this figure (the arrows are proportional to the mode eigenvector) is responsible for the large absorption peak in $\sigma_x(\omega)$ at $\hbar\omega \approx 30$ meV that appears in both the $p(2\times2)$ and $c(4\times2)$ surfaces. This mode appears in both of these surfaces, and along one row of surface dimers it has essentially the same form in both cases. The displacement of the atoms in neighboring rows of dimers is in phase for the $p(2\times2)$ surface and out of phase for the $c(4\times2)$ surface. (a) Top view of the surface; only one row of surface dimers is shown. (b) Side view. The subsurface atoms labeled 1 and 2 are equivalent in the 2×1 surface, but not in the $c(4\times2)$ or $p(2\times2)$ surfaces.

Why is it that this strong x-polarized absorption peak does not occur in the 2×1 surface? The reason is that the surface unit cells in the $c(4\times2)$ and $p(2\times2)$ reconstructions are twice as large as in the 2×1 reconstruction, and this mode is found at the zone center only for the lower-symmetry structures. The atoms labeled 1 and 2 in Fig. 8 are not equivalent in the higher-order reconstructions, as they are in the 2×1 surface. This situation is analogous to a diatomic linear chain where the optical phonon at Γ is dipole active, compared to the monoatomic chain that has no optical mode.

The two comparisons in $\sigma(\omega)$ between the 2×1 surface and the $p(2\times2)$ and $c(4\times2)$ surfaces that we have discussed above can be used to determine whether the periodicity of the surface is 2×1 or of higher order. This possible experimental determination involves spectroscopic studies that would complement recent structural studies (He scattering and STM). Current surface-preparation techniques yield surfaces with a large density of defects, where domains of different periodicities are present. Pandey[13] has even suggested that defects, in the form of missing surface dimers, are an intrinsic property of the Si(001) surface. In principle, however, one might expect that it is possible to prepare high-quality surfaces that could be studied at low temperatures. In particular, in order to uniquely determine the x-polarization spectrum of the surface, one needs single-domain surfaces, or surfaces with terraces where the orientation of the surface dimers is the same across the steps separating these terraces. In some sense, the increase in the absorption intensity of the $c(4\times2)$ and $p(2\times2)$ surfaces with respect to the 2×1 surface suggests an easier experiment to carry through.

The results we have obtained for $\sigma(\omega)$ show very rich structure, particularly for the higher-order periodicity surfaces. A number of the smallest peaks and shoulders in the spectra are associated with modes that are characteristic of the dimer model of the reconstruction of the surface. We mentioned before the stretch mode; there is also, for example, a wagging mode of the surface dimers, which is responsible for the peak in $\sigma_x(\omega)$ close to 55 meV that appears for both $c(4\times2)$ and $p(2\times2)$. However, it is clear that the use of these spectra to differentiate between the $p(2\times2)$ and $c(4\times2)$ reconstructions is not feasible with the experimental resolution currently available.

Finally, we briefly compare these results with those we obtained for the π-bonded chain model of the Si(111) 2×1 surface.[21] Contrary to the extended nature (with respect

to penetration into the bulk) of the electronic surface states on the Si(001) surface, the Si(111) 2×1 surface has states that are very well localized at the surface. Accordingly, we found that the layer-averaged dynamic charges on Si(111) decay extremely fast into the bulk. Also, we found that the phonon absorption spectrum of the surface is dominated by one single mode corresponding to a longitudinal-optical (LO) mode along the surface chains. Unlike the Si(001) surface, on the π-bonded chain model of the Si(111) 2×1 surface there is a connected network of surface atoms, and the large charge oscillation associated with the surface-chain LO mode occurs along this network. As we pointed out, there is no such network of surface atoms on the Si(001) surface.

VI. PHASE COEXISTENCE

The possibility that at low temperature the stable phase of the Si(001) surface involves high-order reconstructions [e.g., $c(4\times2)$ or $p(2\times2)$] has been previously considered in the theoretical work of Ihm *et al.* Since the model underlying our phonon calculations is very similar to that employed in their work, it is useful to compare our results with these previous investigations. An important prediction of the calculations of Ihm *et al.* is that, at finite temperature, transitions between ordered phases [2×1, $p(2\times2)$, or $c(4\times2)$] occur as second-order transitions with a disordered paramagnetic phase separating the ordered phases. According to the structural energies we have calculated, both the $c(4\times2)$ and $p(2\times2)$ reconstructions have lower energy than the 2×1 surface. Nevertheless, the harmonic fluctuations (phonons) about the 2×1 surface do not show any instabilities. That is, both the higher-order and 2×1 phases are simultaneously stable. This would imply first-order transitions with phase-coexistence regions between the ordered phases. The discrepancy between these two results can be attributed to the fact that the renormalization-group (RG) studies of Ihm *et al.* on an effective spin Hamiltonian include fluctuations on all length scales, which our phonon calculations (carried out in the spirit of a mean-field model in which only equilibrium structures of maximum period of four are considered) do not include.

In order to investigate this question in a consistent way with our phonon calculations, we have studied, in the mean-field (MF) approximation, the same spin Hamiltonian used by Ihm *et al.*:

$$\mathscr{H}=-V\sum_{i,j}S_{ij}S_{i+1,j}-H\sum_{i,j}S_{ij}S_{i,j+1}-D\sum_{i,j}S_{ij}S_{i\pm1,j+1}\ , \tag{23}$$

where a spin, with its two possible orientations, represents a tilted surface dimer where one or the other of the surface atoms forming a dimer is up. The spin interactions are defined in the inset of Fig. 9. In this figure we also show the phase diagram of the spin Hamiltonian, in the MF approximation, on the V-H plane; the intersection with the D axis is for D small and positive, a necessary condition for the $c(4\times2)$ surface to be the stable phase. In the calculations of Ihm *et al.* the interaction parameters that correspond to the Si(001) surface are $H=10$, $V=-26$, and $D=4$ meV. This phase diagram is valid at low temperatures. *All* the transition lines are first-order transitions; the point where they meet is a paramagnetic point which grows into a region of finite size with increasing temperature.

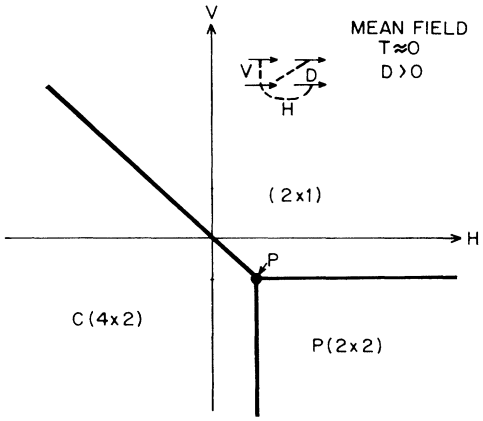

FIG. 9. Low-temperature phase diagram of the spin Hamiltonian defined in Eq. (23), calculated in the MF approximation. The V-H plane shown (interaction parameters defined in the inset) intersects the D axis at a positive and small value. The point P denotes the paramagnetic phase.

An important point to be made is that the transition lines are of first order, so regions of phase coexistence can exist around these transition lines. For example, the 2×1 and $c(4\times2)$ phases can coexist in a narrow band in parameter space around the line $V=-H$. The width of the coexistence region is determined by D. Since (a) the energies of the $c(4\times2)$ and $p(2\times2)$ phases are very close, and (b) our phonon calculations demonstrate the mutual stability of the 2×1 and higher-order phases, we conclude that the appropriate parameters in the effective spin Hamiltonian that represent the Si(001) surface lie very close to the 2×1-$c(4\times2)$ phase boundary, and perhaps very close also to the paramagnetic point P. If this is the case, small perturbations will take the system from one phase to the other, and even a small density of defects may well result in a multidomain surface with domains of different periodicities.

We should stress that the MF calculations neglect long-range fluctuations, which are accounted for in the RG calculations and are responsible for the appearance of the intermediate paramagnetic phase. However, the missing dimer defects, as well as steps, that are observed on the surface limit the size of connected domains and impose a maximum size for the fluctuations. In this sense one might say that the MF approach is more appropriate for studying the short-range correlations on these finite domains. One must be careful, though, in proposing that the coexistence regimes found in the MF phase diagram are responsible for the superimposed periodicities observed on the surface; it might very well be that it is the defects themselves that pin the different reconstructions.

VII. SUMMARY

We have applied a TB theory for structural energies that includes explicit electron-electron interactions in the form of an on-site repulsion term to study vibrational excitations on the Si(001) surface. Three reconstructions of this surface, based on the surface dimer model, were considered: 2×1, $p(2\times2)$, and $c(4\times2)$. We found that even in the presence of strong electron repulsion *tilted* surface dimers are favored, in agreement with most experiments. We present the phonon spectrum of the Si(001) 2×1 surface, and found that the surface supports a number of modes that are characteristic of the dimer-based reconstructions. The study of the phonon spectrum also provides an insight into the nature of the coupling of the surface charge density and the surface structural degrees of freedom, and offers a microscopic picture of the driving forces towards the formation of the higher-order reconstructions, namely $p(2\times2)$ and $c(4\times2)$. Using linear-response theory we have calculated the phonon-assisted contribution to the surface conductivity of the 2×1, $p(2\times2)$, and $c(4\times2)$ surfaces. These results can be used to characterize the periodicity of the surface; in particular, we found a vibrational mode on the $p(2\times2)$ and $c(4\times2)$ surfaces that carries a large (relative to this particular surface) dynamic charge ($\approx0.26e$ per surface dimer) and that does not exist on the 2×1 surface. Our phonon calculations suggest that it is to be expected that multiple periodicities exist on the surface, even at low temperatures. This is consistent with experimental observations. We have complemented these observations with a mean-field study of an effective spin Hamiltonian that represents the surface, finding that the surface maps into a region of the spin-system phase diagram corresponding to a region of coexistence of the three phases 2×1, $p(2\times2)$, and $c(4\times2)$.

ACKNOWLEDGMENT

This work was supported by the U.S. Department of Energy under Grant No. DE-FG02-84ER45118.

[1]A. Sakai, M. J. Cardillo, and D. R. Hamann, Phys. Rev. B **33**, 5744 (1986).
[2]R. M. Hamers, R. J. Tromp, and J. E. Demuth, Phys. Rev. Lett. **55**, 1303 (1985).
[3]F. J. Himpsel and D. E. Eastman, J. Vac. Sci. Technol. **16**, 1297 (1979).
[4]H. H. Farrell, F. Stucki, J. Anderson, D. J. Frankel, G. J. Lapeyre, and M. Levinson, Phys. Rev. B **30**, 721 (1984).
[5]Y. J. Chabal, S. B. Christman, E. E. Chaban, and M. T. Yin, J. Vac. Sci. Technol. A **1**, 1241 (1983).
[6]R. M. Tromp, R. G. Smeenk, F. W. Saris, and D. J. Chadi, Surf. Sci. **133**, 137 (1983).
[7]S. D. Kevan, Phys. Rev. B **32**, 2344 (1985).
[8]D. J. Chadi, Phys. Rev. Lett. **43**, 43 (1979).
[9]J. E. Rowe, Phys. Rev. Lett. **32**, 421 (1974).
[10]R. I. G. Uhrberg, G. V. Hansson, J. M. Nicholls, and S. A. Foldstorm, Phys. Rev. B **24**, 4684 (1981).
[11]D. J. Chadi, J. Vac. Sci. Technol. **16**, 1290 (1979).
[12]M. T. Yin and M. L. Cohen, Phys. Rev. B **24**, 2303 (1981).
[13]K. C. Pandey, in *Proceedings of the International Conference on the Physics of Semiconductors*, edited by W. A. Harrison (Springer, New York, 1985), p. 55.
[14]A. Redondo and W. A. Goddard, J. Vac. Sci. Technol. **21**, 344 (1982).
[15]J. Ihm, D. H. Lee, J. D. Joannopoulos, and J. J. Xiong, Phys. Rev. Lett. **51**, 1872 (1983).
[16]D. J. Chadi, Phys. Rev. Lett. **41**, 1062 (1978).
[17]W. A. Harrison, Phys. Rev. B **31**, 2121 (1985).
[18]S. Tiersten, S. C. Ying, and T. L. Reinecke, Phys. Rev. B **33**, 4062 (1986).
[19]D. G. Allan and E. J. Mele, Phys. Rev. Lett. **53**, 826 (1984).
[20]D. G. Allan and E. J. Mele, Phys. Rev. B **31**, 5565 (1985).
[21]O. L. Alerhand, D. G. Allan, and E. J. Mele, Phys. Rev. Lett. **55**, 2700 (1985).
[22]E. J. Mele, D. G. Allan, O. L. Alerhand, and D. P. DiVincenzo, J. Vac. Sci. Technol. B **3**, 1068 (1985).
[23]D. P. DiVincenzo, O. L. Alerhand, M. A. Schlüter, and J. W. Wilkins, Phys. Rev. Lett. **56**, 1925 (1986).
[24]D. Tomanek and M. A. Schlüter, Phys. Rev. Lett. **56**, 1055 (1986).
[25]H. Ibach, Phys. Rev. Lett. **27**, 253 (1971).
[26]J. N. DiNardo, W. A. Thompson, A. J. Schell-Sorokin, and J. E. Demuth, Phys. Rev. B **32**, 3007 (1986).