Theory of the overlayer plasmon on the Si(001)2×1-K surface

Hiroshi Ishida, Nobuyuki Shima, and Masaru Tsukada
Department of Physics, University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113, Japan
(Received 3 June 1985)

The dispersion of the overlayer plasmon on Si(001)2×1-K surface is investigated theoretically based on structure models in which alkali-metal adatoms form one-dimensional chains on the ridge of Si dimers. First, a simplified rod model is introduced to clarify qualitative features of the overlayer plasmon in such systems. Then a first-principles band-structure calculation is performed to elucidate the detailed electronic structure of Si(001)2×1-K surface. Based on the surface bands obtained, the dispersion of the overlayer plasmon on this surface is calculated nonempirically by the use of the random-phase approximation. The interband plasmon mode between K $4s$—like and K $4p_x$—like bands is found to reproduce the observed plasmon dispersion very well.

I. INTRODUCTION

Overlayer plasmon is a collective excitation of electrons which has been often observed in alkali-metal atom monolayers adsorbed on the surface of other materials.[1–5] Its dispersion, namely plasmon energy versus two-dimensional wave vector $\mathbf{Q}$, is quite sensitive to the microscopic structure of surface and thus can be a powerful tool for analyzing alkali-metal overlayer systems. Adsorption of alkali-metal atoms on solid surface has been intensively studied as an example of simple chemisorption systems. In contrast to the cases of transition-metal adatoms, covalent interaction between alkali-metal atoms and the substrate is not very strong and alkali-metal atoms do not have as much influence on the structure of the substrate. At the full coverage, the alkali-metal overlayer becomes metallic due to the formation of surface bands.

Excitations of valence electrons within a single or between two surface bands are the origin of the overlayer plasmon. The collective excitation within a partially occupied surface band is called as intra-surface-band plasmon whereas that between two different surface bands is called as inter-surface-band plasmon. The former is characterized by the energy dispersion rising from zero in proportion to $Q^{1/2}$ ($Q=|\mathbf{Q}|$) in small Q regions while the dispersion of the latter starts linearly from a finite value. The energy of the inter-surface-band plasmon at $Q=0$ is pushed up at somewhat higher value than the individual excitation energy owing to the depolarization effect.

When the Si(001) surface is clean, it takes the 2×1 or the $c(4\times2)$ reconstructed structure. Supported by numbers of experiments and theoretical studies, the dimer model has been most widely accepted for the reconstructed structure of this surface. Based on low-energy electron diffraction (LEED) observation, Levine proposed a structure model for Si(001)2×1—alkali-metal-atom surfaces shown in Fig. 1.[6] In his model, alkali-metal atoms sit on the fourfold hollow sites between two dimers and form raised rows of one-dimensional chains which extend along [110] direction. When oxygenation is carried out on this surface, O atoms must be located in large caves between alkali-metal chains. Therefore the large dipole between alkali-metal and O atoms required to achieve NEA (negative electron affinity) is expected for this model. Actually among the crystal surfaces of Si, the (001) surface is the only one that can be activated to NEA by the adsorption of alkali-metal atoms followed by the O atom chemisorption.[7]

Recently Aruga, Tochihara, and Murata[8] have exhaustively investigated the dispersion of overlayer plasmons on Si(001)2×1-K surface by AREELS (angle-resolved electron-energy-loss spectroscopy). The measured plasmon dispersion is shown in Fig. 2. Here it cannot be determined experimentally which of the two azimuths ([110] or [1$\bar{1}$0]) the plasmon dispersion in the left half of the figure corresponds to, since Si(001)2×1-K surface is composed of two kinds of domain structures. The plasmon energy corresponding to the main loss peak has a finite energy (~1.7 eV) at $Q=0$ and its dispersion depends positive linearly on Q when Q is small. The slope of the observed plasmon dispersion shows an azimuth-dependent anisotropy, which is expected to reflect the peculiar "chain" structure of the alkali-metal overlayer. In the energy-loss spectrum of the [100] azimuth they also found a shoulder on the lower tail of the main loss peak that might be assigned to another excitation mode.

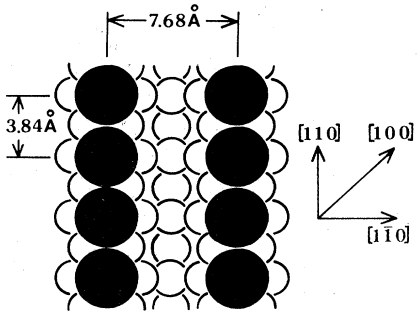

FIG. 1. Structure model of alkali-metal overlayers on the Si(001)2×1 surface. Solid and open circles indicate alkali-metal and Si atoms, respectively.

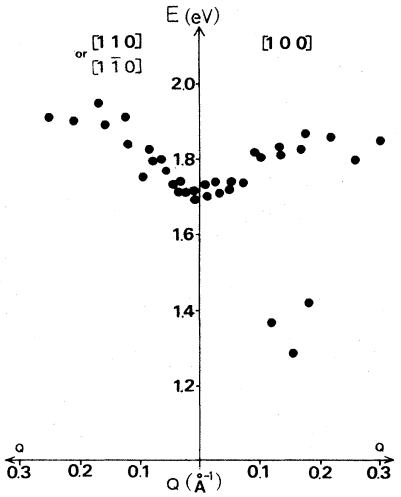

FIG. 2. Dispersion of the overlayer plasmon on the Si(001)2×1-K surface measured by Aruga *et al.* (Ref. 8).

Microscopic theory of the overlayer plasmon was initiated by Newns.[9] He used the "box" model where alkali-metal adatom layers are modeled by an uniform thin film which confines free electrons by infinite potential barriers at both sides. In such an uniform model, microscopic structure of the electron wave function parallel to the surface is completely neglected and the observed anisotropy of the plasmon dispersion cannot be explained. Moreover, the box model always predicts the negative linear dispersion at $Q=0$, which is in disagreement with the observation by Aruga *et al.*[8] Recently Nakayama *et al.* discussed some general properties of the inter-surface-band plasmon at $Q=0$ with the random phase approximation (RPA).[10] According to them, the microscopic shape of the wave functions of electrons in surface bands plays a crucially important role to determine the plasmon dispersion, since the polarization field of plasmon is ruled by it.

In this paper we calculate the dispersion of the overlayer plasmon on the Si(001)2×1-K surface nonempirically based on the self-consistent band calculation and discuss the validity of the structure model through the comparison between the theory and the experiment. The plan of this paper is as follows. In Sec. II we introduce a simplified model which simulates alkali-metal overlayers on the Si(001) surface and clarify the features of the overlayer plasmon in such systems. In Sec. III we reveal the surface-band structure of the Si(001)2×1-K surface with the self-consistent LCAO-$X\alpha$ (LCAO denotes linear combination of atomic orbitals) method. Based on this band calculation the dispersion of the overlayer plasmon is calculated nonempirically in Sec. IV.

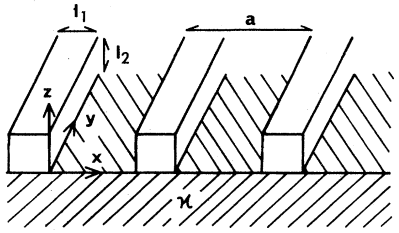

FIG. 3. Parallel-rod model which consists of rectangular rods confining free electrons on the semi-infinite dielectric medium with a dielectric constant κ.

II. PARALLEL-ROD MODEL

To simulate alkali-metal adatom chains on the Si(001)2×1 surface, a simplified parallel-rod model is introduced in this section.[11] As depicted in Fig. 3, the model consists of rectangular rods confining free electrons. The rods have side lengths l_1, l_2 ($\sim$diameter of an alkali-metal atom) and are arrayed in parallel at intervals of a (7.2 Å) on the surface of the uniform substrate which is modeled by a semi-infinite dielectric medium with a dielectric constant κ (~ 15). The eigenfunction and the energy of electrons confined in the nth rod are written as

$$\phi_{kn}^{pq}(\mathbf{r})=\frac{2}{(l_1 l_2 L)^{1/2}}e^{iky}\sin\left[\frac{p\pi}{l_1}(x-na)\right]\sin\left[\frac{q\pi}{l_2}z\right] \tag{2.1}$$

and

$$\epsilon_k^{pq}=\frac{k^2}{2m}+\frac{1}{2\mu_1}\left[\frac{p\pi}{l_1}\right]^2+\frac{1}{2\mu_2}\left[\frac{q\pi}{l_2}\right]^2, \tag{2.2}$$

where L is the quantization length in the y direction, m is the free-electron mass, and μ_1 and μ_2 are the effective masses for x and z directions, respectively. The integers $p-1$, $q-1$ mean the number of nodes of the wave function in the x,z direction and thus the modes with $(p,q)=(1,1),(1,2),(2,1)$ approximately correspond to alkali-metal s,p_z,p_x valence orbitals. In the following, we consider these lowest three modes. We assume the higher bands ϵ_k^{12}, ϵ_k^{21} to be empty while assuming the lowest one ϵ_k^{11} to be occupied if the absolute value of k is less than the Fermi wave number k_f. Then three kinds of excitation modes appear; (i) s-s intraband, (ii) s-p_z interband, and (iii) s-p_x interband excitation modes. Each consists of individual excitation and collective excitation, namely, "plasmon."

The induced charge $\rho_{\text{ind}}(\mathbf{r},\omega)$ due to the external one $\rho_{\text{ext}}(\mathbf{r},\omega)$ is obtained by the standard RPA equation

$$\rho_{\text{ind}}(\mathbf{r},\omega)=2\sum_{i,j}\frac{f_i-f_j}{\epsilon_i-\epsilon_j+\omega+i\delta}\phi_i^*(\mathbf{r})\phi_j(\mathbf{r})\int d\mathbf{r}_1\int d\mathbf{r}_2\,\phi_i(\mathbf{r}_1)\phi_j^*(\mathbf{r}_1)G(\mathbf{r}_1,\mathbf{r}_2)[\rho_{\text{ind}}(\mathbf{r}_2,\omega)+\rho_{\text{ext}}(\mathbf{r}_2,\omega)], \tag{2.3}$$

where f_i, f_j are the Fermi distribution functions and i,j represent the set of indices (k,n,p,q). The interaction potential between two unit charges $G(\mathbf{r}_1,\mathbf{r}_2)$ is given by

$$G(\mathbf{r}_1,\mathbf{r}_2)=\frac{1}{|\mathbf{r}_1-\mathbf{r}_2|}-\frac{\kappa-1}{\kappa+1}\frac{1}{[(\mathbf{x}_1-\mathbf{x}_2)^2+(|z_1|+|z_2|)^2]^{1/2}}. \tag{2.4}$$

$\rho_{\rm ext}(\mathbf{r},\omega)$ is set equal to $\exp(i\mathbf{Q}\cdot\mathbf{X})\rho_{\rm ext}(z)$ to discuss the charge fluctuation with a crystal wave vector $\mathbf{Q}$. Then we find $\rho_{\rm ind}(\mathbf{r},\omega)$ can be expanded as

$$\rho_{\rm ind}(\mathbf{r},\omega)=\sum_{\nu}D(\nu)\Phi(\mathbf{Q},\nu,\mathbf{r}) , \tag{2.5}$$

$$\Phi(\mathbf{Q},\nu,\mathbf{r})=\frac{1}{L^{1/2}}e^{iQ_y y}\frac{1}{N^{1/2}}\sum_{n}e^{iQ_x na}\sin\left[\frac{p_1\pi}{l_1}(x-na)\right]\sin\left[\frac{p_2\pi}{l_1}(x-na)\right]\sin\left[\frac{q_1\pi}{l_2}z\right]\sin\left[\frac{q_2\pi}{l_2}z\right] , \tag{2.6}$$

where N is the number of rods and ν represents a set of indices (p_1,p_2,q_1,q_2). The correspondence of the pair band indices (p_1,p_2,q_1,q_2) and the mode index ν is listed in Table I. Substituting Eq. (2.5) into Eq. (2.3) yields the matrix equation

$$\sum_{\nu'}[\delta_{\nu\nu'}-\chi(Q_y,\omega,\nu)I(\mathbf{Q},\nu,\nu')]D(\nu')=\chi(Q_y,\omega,\nu)J(\mathbf{Q},\nu) . \tag{2.7}$$

In the above, $\chi(Q,\omega,\nu)$ is the polarization function defined by

$$\chi(Q,\omega,\nu)=\frac{2\pi}{L}\sum_{k}\frac{f_{k-Q}^{p_2q_2}-f_k^{p_1q_1}}{\epsilon_{k-Q}^{p_2q_2}-\epsilon_k^{p_1q_1}+\omega+i\delta}+(1-\delta_{p_1p_2}\delta_{q_1q_2})\frac{2\pi}{L}\sum_{k}\frac{f_{k-Q}^{p_1q_1}-f_k^{p_2q_2}}{\epsilon_{k-Q}^{p_1q_1}-\epsilon_k^{p_2q_2}+\omega+i\delta} . \tag{2.8}$$

$I(\mathbf{Q},\nu,\nu')$ and $J(\mathbf{Q},\nu)$ are the matrix elements defined by

$$I(\mathbf{Q},\nu,\nu')=\frac{16}{\pi l_1^2 l_2^2}\int d\mathbf{r}_1\int d\mathbf{r}_2\,\Phi^*(\mathbf{Q},\nu,\mathbf{r}_1)G(\mathbf{r}_1,\mathbf{r}_2)\Phi(\mathbf{Q},\nu',\mathbf{r}_2) . \tag{2.9}$$

$$J(\mathbf{Q},\nu)=\frac{16}{\pi l_1^2 l_2^2}\int d\mathbf{r}_1\int d\mathbf{r}_2\,\Phi^*(\mathbf{Q},\nu,\mathbf{r}_1)G(\mathbf{r}_1,\mathbf{r}_2)\times e^{i\mathbf{Q}\cdot\mathbf{X}_2}\rho_{\rm ext}(z_2) . \tag{2.10}$$

The plasmon energy is determined by the condition

$$\det[\delta_{\nu\nu'}-\chi(Q_y,\omega,\nu)I(\mathbf{Q},\nu,\nu')]=0 . \tag{2.11}$$

As we are considering three coupled excitation modes, the matrix in Eq. (2.7) becomes 3×3. But nondiagonal matrix elements are smaller than the diagonal ones in small Q regions, so at first we use the approximation to treat the three modes independently. Such a procedure is quite helpful to emboss the physical picture of each plasmon model.

A. s-s ($\nu=1$) intraband mode

When Q is small, the plasmon dispersion is given by

$$\omega^2(\mathbf{Q})=\frac{2k_f}{ma}\left[\frac{2}{\kappa+1}+\left[\frac{5}{4\pi^2}-\frac{1}{3}+\frac{\kappa-1}{\kappa+1}\right]l_2Q\right]Q\cos^2\theta , \tag{2.12}$$

where θ is the angle between $\mathbf{Q}$ and rod axis. The dispersion is almost the same as that of A mode of Newns's model[9] except for the factor $\cos\theta$. This mode is essentially caused by the charge fluctuation parallel to the rods. Thus it can oscillate most rapidly along the chains ($\theta=0°$) and cannot oscillate perpendicularly to them ($\theta=90°$).

TABLE I. Correspondence of the mode index ν and pair-band indices (p_1,p_2,q_1,q_2).

ν	Mode	p_1	p_2	q_1	q_2
1	s-s	1	1	1	1
2	s-p_z	1	1	1	2
3	s-p_x	1	2	1	1

B. s-p_z ($\nu=2$) interband mode

When Q is small, plasmon dispersion is given by

$$\omega^2(\mathbf{Q})=\omega_{12}^2+4\omega_{12}k_f\left[I(\mathbf{Q}=\mathbf{0},2,2)-\frac{2\kappa}{\kappa+1}\left[\frac{16l_2}{9\pi^2}\right]^2\frac{2Q}{a}\right] , \tag{2.13}$$

where $\omega_{12}=3\pi^2/(2\mu_2 l_2^2)$ is the energy difference between s and p_z bands. The plasmon energy exceeds the individual excitation energy ω_{12} by the depolarization shift. The polarization field is perpendicular to the surface and this mode is essentially the same as the B mode of Newns's model.[9] The dispersion is linear with negative

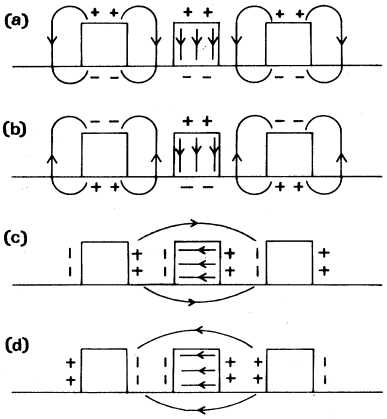

FIG. 4. Schematic illustration of the polarization fields by s-p_z mode [(a) and (b)] and s-p_x mode [(c) and (d)]. $\mathbf{Q}$ is equal to (0,0) for (a) and (c) and $(\pi/a,0)$ for (b) and (d).

gradient and moreover it is isotropical in small Q regions though the electronic structure is extremely anisotropic. Here we clarify the physical meaning of the negative linear dispersion. It means the decrease of the depolarization shift as Q deviates from zero. Actually, as shown in Fig. 4(a), all the rods oscillate in phase at $Q=0$ and the polarization field of a rod is reinforced by those of neighbors. This results in the largest depolarization shift at $Q=0$. On the other hand, the adjacent rods begin to oscillate in antiphase as Q increases and the polarization field of a rod is partially canceled by those of neighbors [Fig. 4(b)]. Therefore the depolarization shift becomes smaller and the negative linear dispersion occurs.

C. s-p_x ($\nu=3$) interband mode

When Q is small, the plasmon dispersion is given by

$$\omega^2(\mathbf{Q})=\omega_{12}^2+4\omega_{12}k_f\left[I(\mathbf{Q}=\mathbf{0},3,3)+\frac{2}{\kappa+1}\left[\frac{16l_1}{9\pi^2}\right]^2\frac{2Q}{a}\sin^2\theta\right], \tag{2.14}$$

where ω_{12} is equal to $3\pi^2/(2\mu_1 l_1^2)$. This mode is characteristic of the chain structure of the present model and does not exist in uniform thin-film models advanced so far. The dispersion depends on Q positive linearly. The slope of the dispersion at $Q=0$ is proportional to $\sin^2\theta$ and thus becomes the largest at $\theta=90°$. The origin of the positive linear dispersion is easily understood. In contrast to the case of s-p_z mode, the polarization field of a rod is weakened most by those of neighbors when they oscillate in phase [Fig. 4(c)] and the depolarization shift becomes the smallest at $Q=0$. On the other hand, the cancellation of the polarization field becomes smaller with the increasing Q and correspondingly the depolarization shift becomes larger by degrees [Fig. 4(d)]. This mode reproduces the features of the observed plasmon dispersion well and is expected to be responsible for the main loss peak of the AREELS observation. Thus the plasmon dispersion in the left half of Fig. 2 may well be assigned to that of [1$\bar{1}$0] azimuth.

We have solved Eq. (2.11) numerically in general Q regions and obtained dispersions of three coupled plasmon modes. Fig. 5 shows the plasmon dispersion for $\theta=45°$. The interband transition energies ω_{12} are adjusted to 0.59 eV in order that interband plasmon modes roughly reproduce the observed loss peak and l_1, l_2 are chosen to be 3.7 Å. Near $Q=0$ the plasmon modes $\omega_1,\omega_2,\omega_3$ correspond to s-s ($\nu=1$), s-p_z ($\nu=2$), s-p_x ($\nu=3$) modes, respectively. The shaded regions indicate the individual excitation energies. The s-p_x mode reproduces the observed plasmon dispersion corresponding to the main loss peak in small Q regions well. Subsidiary structure appearing in the tail of the main loss peak may be assigned to the s-s intraband mode.

III. ELECTRONIC STRUCTURE OF THE Si(001)2×1-K SURFACE

In Sec. II we showed that the observed plasmon dispersion can be explained fairly well provided that there are

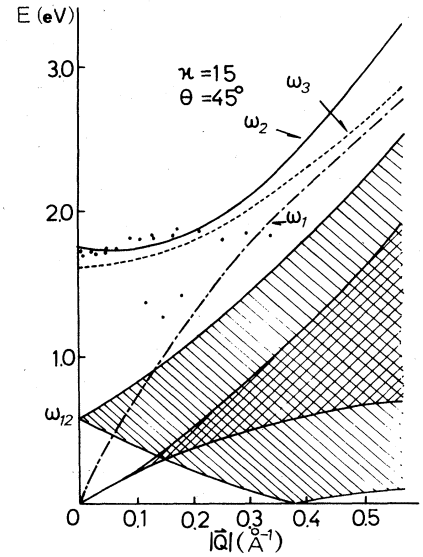

FIG. 5. Calculated plasmon dispersion of the parallel-rod model for the case of $\kappa=15$, $\theta=45°$. Solid circles show the experimental data.

two surface bands near the Fermi level, which have characters of K 4s and 4p_x states. In the real system the band structure of the alkali-metal overlayer might change drastically from such a simplified model due to the charge transfer and the orbital hybridization between the K adatoms and the substrate Si atoms. To elucidate the electronic structure of the Si(001)2×1-K surface qualitatively as well as quantitatively, the self-consistent band calculation is performed in this section.

A. Model and the method of the band calculation

We adopt two structure models as candidates for the Si(001)2×1-K surface. Model 1 is the original Levine model,[6] in which Si atoms in the top layer are assumed to form symmetric dimers and the K atom sits on the fourfold hollow site between two dimers. Model 2 is identical to model 1 except that the K atom sits on the bridge site of a dimer instead of the hollow site. The reason why we take two structure models is as follows. Though it is almost established that alkali-metal atoms are arrayed on the ridge of the Si dimers, the precise adsorption site of the K atom still remains unknown. So it is interesting to examine the change of the electronic structure when the construction of the surface is slightly modified in order to extract common features of the band structure in such chain systems.

The surface band structures are calculated with the thin-film models KSi_8H_4. Figure 6(a) shows the side view of the thin film model 1, which is the same as that of model 2 except for the height of the K atom. Figures 6(b) and 6(c) show the top views of models 1 and 2, respectively. The thin film consists of the K overlayer, four Si layers, and the H overlayer on the backside surface. The hydrogen atoms are adsorbed to eliminate the dangling bonds of the backside surface. The structure of the sub-

strate Si is assumed identical as that of the bulk Si except for the top dimer layer. The bond length between neighboring Si atoms is 2.3 Å and we use the same value for the Si—Si bond length of a dimer. The distance between the K atom and the closest Si atom of a dimer is chosen to be equal to the sum of K and Si atomic radii, 3.52 Å. The band structure is calculated with the numerical-LCAO-$X\alpha$ method. The details of the numerical procedures are given in Ref. 12.

Since the asymmetric dimer is presumed superior to the symmetric one with respect to the bare Si(001)2×1 surface,[13,14] it is intriguing to calculate the electronic structure of the Si(001)2×1-K surface with the asymmetric dimers. However, it should be remarked that the mechanism which stabilizes the asymmetric dimer as compared to the symmetric one might not work for the alkali-metal covered surface because of the charge transfer between the alkali-metal atom and the dimer.

B. Results of the band calculation

To test the validity of our band calculation, we first calculated the band structure of the bare Si(001)2×1 surface with the thin film model Si_8H_4 which is obtained from KSi_8H_4 by removing the K atom. The resultant band structure near the Fermi level E_F is shown in Fig. 7 together with the Brillouin zone (BZ) and the notation of symmetry points. The bands with bulklike characters are

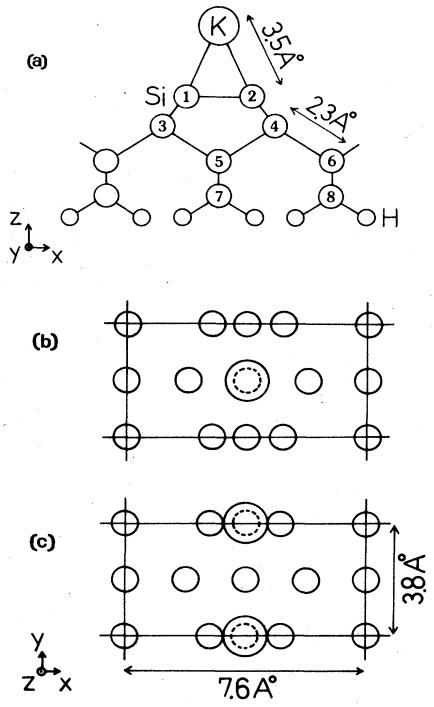

FIG. 6. (a) Side view of the thin-film model KSi_8H_4 seen from the [110] direction, which is common to models 1 and 2, except for the height of the K atom. (b) Top view of model 1. (c) Top view of model 2.

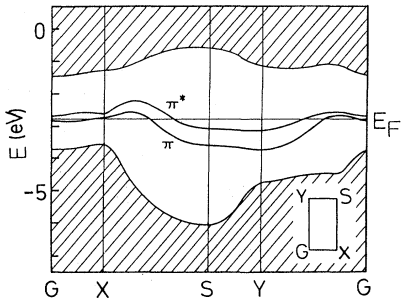

FIG. 7. Surface bands of the thin-film Si_8H_4. The BZ and the notation of symmetry points are also shown.

located in shaded regions. There appear two surface bands in the gap region. The lower band has the character of the bonding π orbital while the upper one has the character of the antibonding π^* orbital composed of two dangling bonds of a dimer. The obtained band structure is almost the same as that proposed by Ihm, Cohen, and Chadi[15] using the pseudopotential method. Hence, the present slab model KSi_8H_4 is also supposed to work well to reveal the band structure of the Si(001)2×1-K surface.

The calculated band structures of KSi_8H_4 for models 1 and 2 are shown in Figs. 8(a) and 8(b), respectively, in the energy region near the Fermi level E_F. The surface becomes metallic and E_F crosses the band b. Three electrons per unit cell are accommodated in the bands a and b. The bands a, b, and c are the surface states composed mainly of K valence states and those of a dimer as discussed below. Other bands with bulklike characters are situated in the shaded regions. It should be noted that our band calculations do not reproduce the band gap of the substrate Si because they are obtained with thin-film models which include only four Si layers. But this is not crucial as we are mainly interested in the surface band structures, which can be reproduced with thin-film models so long as the charge transfer between the overlayer and the substrate is adequately described by them. By comparing Fig. 8(a) with Fig. 8(b), we find overall structures of surface bands are common to models 1 and 2 except for some detailed points such as width of each band and relative energy differences between them. Owing to the chain structure of the Si(001)2×1-K surface, the energy dispersion of each surface band is generally smaller along the G—X and Y—S axes than along the G—Y and X—S axes. Yet, the band b has a rather large energy change along the G—X axis as compared with others due to the interaction between chains mediated by the substrate Si.

Next, we will discuss the orbital characters of the three surface bands as they are closely related with the dispersion of the overlayer plamon. Reflecting the peculiar chain structure, the orbital components of each band do not change along the direction perpendicular to the chain. The origin of the bands a, b, and c near the G—X axis in the BZ is common to models 1 and 2 and can be understood schematically as illustrated in Fig. 9. The wave function of the surface band a is mainly made of the antibonding combination of the K 4s and the bonding σ orbi-

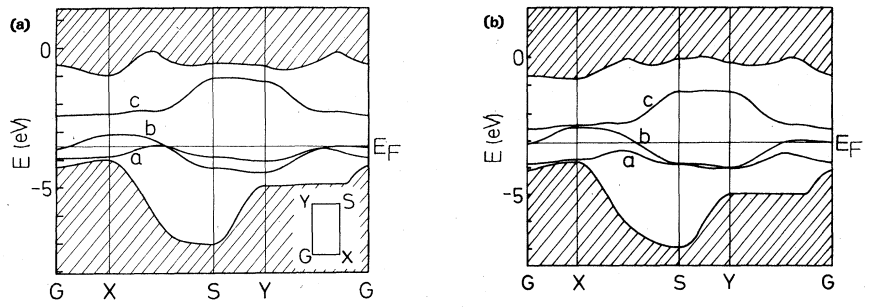

FIG. 8. Surface bands of the thin-film KSi_8H_4 for (a) model 1 and (b) model 2.

tal of a dimer. It also contains the π orbital of a dimer with a little smaller weight. The wave function of the band b is mostly contributed from the bonding combination of K $4p_x$ and the π^* orbital of a dimer. On the other hand, the wave function of the band c is mainly made of the bonding combination of K $4p_z$ orbital and the π orbital of a dimer with a considerable admixture of K $4s$ orbital. The interaction between K $4s,4p_x$ states and the π^*,σ states of a dimer is effectively larger in model 1 than in model 2 since the K atom has two adjacent dimers in model 1 while it has only one in model 2. Thus model 1 exceeds model 2 both in the energy lowering of the band b and in the energy rising of the band a. Therefore the energy difference between the bands a and b of model 1 at the G point ($\sim$0.7 eV) becomes smaller than the corresponding value of model 2 ($\sim$1.1 eV). In contrast to this, the interaction between the K $4p_z$ state and the π state of a dimer becomes larger in model 2 than in model 1 owing to the large overlap of the two orbitals. This results in the larger energy lowering of the band c near the G—X axis for model 2.

Near the Y—S axis, the orbital components of the valence states of the substrate Si are for the most part the same as those near the G—X axis for both models 1 and 2. On the other hand, weights of K valence states in the bands a, b, and c are considerably reduced and each surface band comes to have the character of the dangling-bond band of a dimer. The large energy increase of the band c along the X—S axis is due to the K $4s$ component in it while the large energy decrease of band b reflects 3-p_y of the substrate Si atoms.

Finally, we mention the charge distribution of the surface. According to Mulliken's population analysis, the charge transfer from K atom to the substrate is about 0.13 electron for model 1 and about 0.09 electron for model 2. The Si atom of a dimer becomes negative about -0.08 electron in model 1 and about -0.09 electron in model 2.

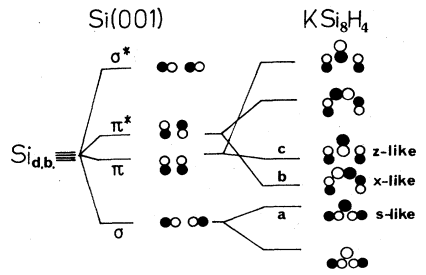

FIG. 9. Orbital character of the surface bands a, b, and c near the G—X axis. $Si_{d.b.}$ denotes the dangling bond states of Si atoms before the formation of dimers.

IV. DISPERSION OF THE OVERLAYER PLASMON BASED ON THE SURFACE BAND CALCULATION

A. Theoretical formulation

In this section we calculate the dispersion of the overlayer plasmon by a nonempirical way based on the surface bands calculated in Sec. III. For this purpose we solve the random-phase-approximation (RPA) Eq. (2.3). The wave function $\phi_{n\mathbf{k}}$ ($=\phi_i$) for the state with the wave vector $\mathbf{k}$ in the nth band is written in terms of the Wannier function of the nth band as

$$\phi_{n\mathbf{k}}(\mathbf{r})=\frac{1}{N^{1/2}}\sum_l e^{i\mathbf{k}\cdot l}W_n(\mathbf{r}-l)\ , \tag{4.1}$$

where l is a lattice vector. $\rho_{\text{ext}}(\mathbf{r})$ is hereafter set equal to $\exp(i\mathbf{Q}\cdot\mathbf{X})\rho_{\text{ext}}(z)$. Then the induced charge can be expanded as

$$\rho_{\text{ind}}(\mathbf{r},\omega)=\sum_{m,n,d} D(m,n,\mathbf{d})\Phi(\mathbf{Q},m,n,\mathbf{d},\mathbf{r})\ , \tag{4.2}$$

$$\Phi(\mathbf{Q},m,n,\mathbf{d},\mathbf{r})=\frac{1}{N^{1/2}}\sum_l e^{i\mathbf{Q}\cdot l}W_m(\mathbf{r}-l)W_n(\mathbf{r}-l-\mathbf{d})\ , \tag{4.3}$$

In the above, $\mathbf{d}$ is a lattice vector and only the terms with $n\geqq m$ are included in Eq. (4.2) as $\Phi(\mathbf{Q},n,m,\mathbf{d},\mathbf{r})$ is equal to $\Phi(\mathbf{Q},m,n,-\mathbf{d},\mathbf{r})$. The following matrix equation is derived for the expansion coefficients $D(m,n,\mathbf{d})$:

$$[\delta_{m_1n_1\mathbf{d}_1,m_2n_2\mathbf{d}_2}-\chi(\mathbf{Q},\omega,m_1,n_1,\mathbf{d}_1,\mathbf{d}_3)I(\mathbf{Q},m_1,n_1,\mathbf{d}_3;m_2,n_2,\mathbf{d}_2)]D(m_2,n_2,\mathbf{d}_2)=\chi(\mathbf{Q},\omega,m_1,n_1,\mathbf{d}_1,\mathbf{d}_3)J(\mathbf{Q},m_1,n_1,\mathbf{d}_3)\ , \tag{4.4}$$

where repeated indices are understood to imply summation. In the above, $\chi(\mathbf{Q},\omega,m,n,\mathbf{d},\mathbf{d}')$ is the polarization function defined by

$$\chi(\mathbf{Q},\omega,m,n,\mathbf{d},\mathbf{d}')=\frac{2}{N}\sum_{\mathbf{k}}\frac{f(m,\mathbf{k}-\mathbf{Q})-f(n,\mathbf{k})}{\epsilon(m,\mathbf{k}-\mathbf{Q})-\epsilon(n,\mathbf{k})+\omega+i\delta}e^{i\mathbf{k}\cdot(\mathbf{d}-\mathbf{d}')}+\frac{2}{N}\sum_{\mathbf{k}}\frac{f(n,\mathbf{k})-f(m,\mathbf{k}+\mathbf{Q})}{\epsilon(n,\mathbf{k})-\epsilon(m,\mathbf{k}+\mathbf{Q})+\omega+i\delta}e^{-i\mathbf{k}\cdot(\mathbf{d}-\mathbf{d}')}\quad(n>m)$$

$$=\frac{2}{N}\sum_{\mathbf{k}}\frac{f(m,\mathbf{k}-\mathbf{Q})-f(n,\mathbf{k})}{\epsilon(m,\mathbf{k}-\mathbf{Q})-\epsilon(n,\mathbf{k})+\omega+i\delta}(e^{i\mathbf{k}\cdot\mathbf{d}}+e^{i(\mathbf{Q}-\mathbf{k})\cdot\mathbf{d}}-\delta_{\mathbf{d},0})(e^{-i\mathbf{k}\cdot\mathbf{d}'}+e^{i(\mathbf{Q}-\mathbf{k})\cdot\mathbf{d}'}-\delta_{\mathbf{d}',0})\quad(n=m)\ . \tag{4.5}$$

$I(\mathbf{Q},m,n,\mathbf{d};m',n',\mathbf{d}')$ and $J(\mathbf{Q},m,n,\mathbf{d})$ are the matrix elements defined by

$$I(\mathbf{Q},m,n,\mathbf{d};m',n',\mathbf{d}')=\int d\mathbf{r}_1\int d\mathbf{r}_2\,\Phi^*(\mathbf{Q},m,n,\mathbf{d},\mathbf{r}_1)G(\mathbf{r}_1,\mathbf{r}_2)\Phi(\mathbf{Q},m',n',\mathbf{d}',\mathbf{r}_2)\ , \tag{4.6}$$

$$J(\mathbf{Q},m,n,\mathbf{d})=\int d\mathbf{r}_1\int d\mathbf{r}_2\,\Phi^*(\mathbf{Q},m,n,\mathbf{d},\mathbf{r}_1)G(\mathbf{r}_1,\mathbf{r}_2)e^{iQ\cdot X_2}\rho_{\text{ext}}(z_2)\ . \tag{4.7}$$

The plasmon energy is determined by the condition that Eq. (4.4) for no external field ($J=0$) has the nontrivial solution, i.e.,

$$\det[\delta_{m_1n_1\mathbf{d}_1,m_2n_2\mathbf{d}_2}-\chi(\mathbf{Q},\omega,m_1,n_1,\mathbf{d}_1,\mathbf{d}_3)I(\mathbf{Q},m_1,n_1,\mathbf{d}_3;m_2,n_2,\mathbf{d}_2)]=0\ . \tag{4.8}$$

However, if the plasmon energy is merged in the continuous energy spectrum of other individual excitation modes, the plasmon peak becomes broad owing to the coupling with them and the zero point of the determinant becomes undecided. In such a case, we calculate the energy loss spectrum of the incident electron beam directly by means of the "classical trajectory approximation."[16,17]

The energy loss W of the specularly reflected electron beam outgoing with velocity $(\mathbf{v}_{||},v_z)$ is then given by

$$W=-\int_0^\infty d\omega\frac{\omega}{2\pi}\int_{\text{BZ}}\frac{d^2Q}{(2\pi)^2}\sum_{\mathbf{G}_1,\mathbf{G}_2}\Gamma(\mathbf{Q}+\mathbf{G}_1,\omega)\Gamma(\mathbf{Q}+\mathbf{G}_2,\omega)\text{Im}\Lambda(\mathbf{Q},\omega,\mathbf{G}_1,\mathbf{G}_2)\ , \tag{4.9}$$

where

$$\Gamma(\mathbf{Q},\omega)=\frac{4\pi v_z}{(\omega-\mathbf{v}_{||}\cdot\mathbf{Q})^2+v_z^2Q^2}\ , \tag{4.10}$$

and $\mathbf{G}_1,\mathbf{G}_2$ are reciprocal lattice vectors. In the above, $\Lambda(\mathbf{Q},\omega,\mathbf{G}_1,\mathbf{G}_2)$ is defined by

$$\Lambda(\mathbf{Q},\omega,\mathbf{G}_1,\mathbf{G}_2)=\frac{1}{S}(1-\chi I)^{-1}_{m_1n_1\mathbf{d}_1,m_2n_2\mathbf{d}_2}\chi(\mathbf{Q},\omega,m_2n_2,\mathbf{d}_2,\mathbf{d}_3)J^*(\mathbf{Q}+\mathbf{G}_1,m_1,n_1,\mathbf{d}_1)J(\mathbf{Q}+\mathbf{G}_2,m_2,n_2,\mathbf{d}_3)\ , \tag{4.11}$$

where S is the area of the unit cell and the exponentially decaying field $\exp(Qz)$ should be used as an external potential in calculating $J(\mathbf{Q},m,n,\mathbf{d})$. The quantity

$$-\sum_{\mathbf{G}_1,\mathbf{G}_2}\Gamma(\mathbf{Q}+\mathbf{G}_1,\omega)\Gamma(\mathbf{Q}+\mathbf{G}_2,\omega)\text{Im}\Lambda(\mathbf{Q},\omega,\mathbf{G}_1,\mathbf{G}_2) \tag{4.12}$$

is interpreted as a loss function which gives the probability that an excitation mode with energy ω and with crystal momentum $\mathbf{Q}$ is generated in the system by an incident electron beam. All information of the elementary excitation peculiar to the system is contained in $\Lambda(\mathbf{Q},\omega,\mathbf{G}_1,\mathbf{G}_2)$.

B. Numerical procedure

The Bloch function $\phi_{n\mathbf{k}}(\mathbf{r})$ is expanded as

$$\phi_{n\mathbf{k}}(\mathbf{r})=\sum_{\alpha,\lambda}C_{\alpha\lambda}(n,\mathbf{k})\frac{1}{N^{1/2}}\sum_l e^{i\mathbf{k}\cdot(l+\mathbf{r}_\alpha)}\psi_{\alpha\lambda}(\mathbf{r}-l-\mathbf{r}_\alpha)\ , \tag{4.13}$$

where $\mathbf{r}_\alpha$ is the coordinate vector of the αth atom in the unit cell, $\psi_{\alpha\lambda}$ is the λth atomic orbital of the αth atom, and $C_{\alpha\lambda}(n,\mathbf{k})$ is the expansion coefficient. The Wannier function is composed of linear combination of atomic orbitals,

$$W_n(\mathbf{r})=\sum_{\alpha,\lambda,l}A(\alpha,\lambda,l)\psi_{\alpha\lambda}(\mathbf{r}-l-\mathbf{r}_\alpha)\ , \tag{4.14}$$

where

$$A(\alpha,\lambda,l)=\frac{1}{N}\sum_{\mathbf{k}}C_{\alpha\lambda}(n,\mathbf{k})e^{i\mathbf{k}\cdot(l+\mathbf{r}_\alpha)}\ . \tag{4.15}$$

We select the phase of $C_{\alpha\lambda}(n,\mathbf{k})$ so that $W_n(\mathbf{r})$ may be localized at $r=0$ as well as possible. This is achieved fairly well if $C_{\alpha\lambda}(n,\mathbf{k})$ corresponding to the dominant orbital of the nth band is chosen to be a real positive number. In this case, $W_n(\mathbf{r})$ becomes a real function since $C_{\alpha\lambda}(n,\mathbf{k})$ coincides with $C^*_{\alpha\lambda}(n,-\mathbf{k})$ due to the time reversal symmetry.

In evaluating matrix elements $I(\mathbf{Q},m,n,d;m',n',d')$ and $J(\mathbf{Q},m,n,d)$, $G(\mathbf{r}_1,\mathbf{r}_2)$ is replaced by the bare Coulomb potential as it is not certain how the image potential approximation holds on an atomic scale. To compute them analytically, each numerical atomic base which constructs the Wannier functions is fitted by the linear combination of Gaussian orbitals. Six Gaussian orbitals are found sufficient to reproduce the features of the valence orbitals of Si and K perfectly. The **k**-space integral of the polarization function is calculated with the tetrahedron method of Lehmann and Taut.[18]

C. Results of the calculation

There are three surface bands a, b, and c near the Fermi level E_F and thus following four excitation modes are possible, in principle: (1) $b\rightarrow b$ intraband, (2) $a\rightarrow b$ interband, (3) $a\rightarrow c$ interband, and (4) $b\rightarrow c$ interband excitation modes. At first the dispersion of the $a\rightarrow b$ plasmon is calculated following the procedure described above. This mode corresponds to the s-p_x mode of the parallel rod model and is expected to reproduce the features of the observed plasmon dispersion. As we are interested in the $a\rightarrow b$ excitation mode, we adopt an approximation to retain only the matrix components with $(m,n)=(a,b)$ neglecting the couplings with other excitation modes.

The Wannier functions of surface bands a and b are calculated from Eqs. (4.14) and (4.15). In order to localize them at $\mathbf{r}=\mathbf{0}$, $C_{\alpha\lambda}(n,\mathbf{k})$ corresponding to the bonding σ and π^* orbital of a dimer are chosen to be a real positive number in the BZ for surface bands a and b, respectively. Since the obtained Wannier functions are well localized along the direction perpendicular to K chains, the overlap of two Wannier functions, $W_m(\mathbf{r})W_n(\mathbf{r}-\mathbf{d})$, becomes large only for the three vectors in units of lattice parameters: $\mathbf{d}=(0,0)$, $(0,\pm1)$. Thus these three vectors are incorporated into the sets of indices (m,n,d) and as a result the matrix in Eq. (4.4) becomes 3×3. The calculated dispersion of the $a\rightarrow b$ interband plasmon of models 1 and 2 is shown in Figs. 10(a) and 10(b), respectively, for the $[1\bar{1}0]$ and the $[100]$ azimuth. The shaded region indicates the maximum of the individual excitation energy. The depolarization shift at $Q=0$ amounts to about 0.5 eV. The obtained plasmon dispersion exquisitely reproduces both the positive linear dispersion and the azimuth-dependent anisotropy of the experimental dispersion. The slope of the dispersion at $Q=0$ becomes the largest for the $[1\bar{1}0]$ azimuth as in the case of the parallel-rod model. Moreover, saturation of the plasmon energy in large Q regions, which was not reproduced by the parallel-rod model, is also realized very well. The difference of the plasmon energy between models 1 and 2 is attributed to the difference of the maximum single-particle excitation energy that amounts to 0.4 eV.

Next, we consider the couplings of the $a\rightarrow b$ excitation and other surface excitation modes. Especially for model 2, the energy region of the $a\rightarrow b$ plasmon overlaps with the individual excitation spectrum of the $a\rightarrow b$ and $b\rightarrow c$ excitation modes. Thus these modes might conceal the $a\rightarrow b$ interband plasmon from the energy loss spectrum of the incident electron beam. So we solve the matrix equation (4.4) taking three excitation modes $(m,n)=(a,b),(a,c),(b,c)$ into consideration. But in this case we must be reconciled to taking only $\mathbf{d}=(0,0)$ in sets of indices (m,n,d) so as to make the calculation tractable.

The obtained energy loss spectrum by Eq. (4.12) is shown in Fig. 11 for a crystal momentum $\mathbf{Q}=(0.2,0)$ Å^{-1}. The sharp peak at $\omega\sim1.5$ eV corresponds to the $a\rightarrow b$ plasmon. Though the plasmon energy is somewhat smaller than that of Fig. 10(b), this is an effect caused by neglecting the matrix elements with $\mathbf{d}=(0,\pm1)$. The continuous loss spectrum below the $a\rightarrow b$ plasmon corresponds to the $a\rightarrow b$ individual excitation spectrum, while that around and above it is due to the $a\rightarrow c$ and $b\rightarrow c$ individual excitations. The additional small peak near the maximum of the $b\rightarrow c$ individual excitation spectrum is roughly regarded as the $b\rightarrow c$ plasmon. However, the matrix element $I(\mathbf{Q},b,c,\mathbf{d}=0;b,c,\mathbf{d}'=0)$, in other words, the polarization field by the $b\rightarrow c$ excitation, is rather small and this peak appears only after taking account of the coupling of the $b\rightarrow c$ and $a\rightarrow b$ modes. We find another plasmon pole at $\omega\sim4.0$ eV which is identified as the $a\rightarrow c$ plasmon. The Wannier functions of the surface bands a,c have the same symmetry and thus the matrix element $I(\mathbf{Q},a,c,\mathbf{d}=0;a,c,\mathbf{d}'=0)$ becomes large. This leads to the large depolarization shift of the $a\rightarrow c$

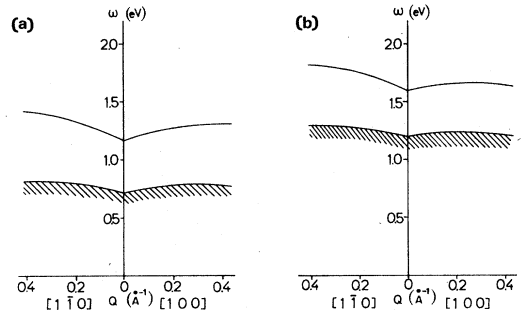

FIG. 10. Calculated dispersions of the $a\rightarrow b$ interband plasmon for (a) model 1 and (b) model 2. The shaded region indicates the maximum of the $a\rightarrow b$ individual excitation energy.

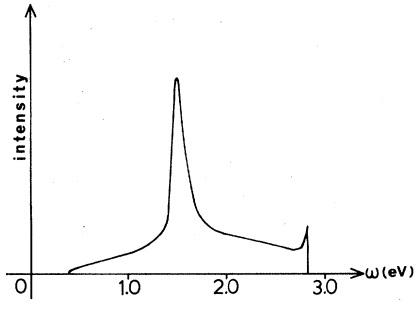

FIG. 11. Calculated energy-loss spectrum of model 2 for $\mathbf{Q}=(0.2,0.0)$ (Å^{-1}). Three excitation modes with $(m,n)=(a,b),(a,c),(b,c)$ are considered. The sharp peak corresponds to the $a\rightarrow b$ interband plasmon.

plasmon which comes to about 1.2 eV. From the figure we can expect that the $a \rightarrow b$ plasmon is clearly observed in the energy loss spectrum of the electron beam even if it is merged in energy regions of other excitation modes and it seems reasonable to assign the observed plasmon peak to the $a \rightarrow b$ interband plasmon in the present structure models.

Finally we will comment briefly on the absolute value of the $a \rightarrow b$ interband plasmon energy at $Q=0$. It is about 1.2 eV for model 1 and about 1.6 for model 2. So regarding the agreement with the experimental value (~ 1.7 eV), model 2 seems superior to model 1. But it is not conclusive, since we have not calculated the total energies of both systems yet.

V. SUMMARY

Collective modes in the array of alkali-metal adatom chains are elucidated using the simplified rod model. The s-p_x interband plasmon peculiar to the chain structure reproduces the features of the observed plasmon dispersion well. The subsidiary structure of the AREELS at the lower tail of the main loss peak is considered to correspond to the s-s intraband mode. Detailed band structure of Si(001)2×1-K surface is obtained by the first-principles band calculation. Three surface bands, a, b, and c, appear near the Fermi level. The bands a, b, and c have the K $4s$—, K $4p_x$—, and K $4p_z$—like characters near the G—X line, respectively. These surface bands are found not to be influenced against slight changes of the surface structure. The dispersions of the overlayer plasmon are calculated nonempirically based on this band calculation. The $a \rightarrow b$ inter-surface-band plasmon mode, which reproduces the observed plasmon dispersion very well, is assigned to the main loss peak of the AREELS.

In conclusion, the observed plasmon dispersion of the Si(001)2×1-K surface can be explained fairly well based on the structure models with alkali-metal adatom chains riding on the ridge of Si dimers. This fact provides a supporting evidence of such models from the theoretical point of view.

ACKNOWLEDGMENTS

The authors would like to express their sincere thanks to Professor Y. Murata, Dr. H. Tochihara, Dr. T. Aruga, and Professor M. Nakayama for valuable discussions. Numerical calculations were performed at the Computer Center of the Institute for Molecular Science and that of the University of Tokyo.

[1]A. U. MacRae, K. Muller, J. J. Lander, J. Morrison, and J. C. Phillips, Phys. Rev. Lett. **22**, 1048 (1969).
[2]U. Jostell, Surf. Sci. **82**, 333 (1979).
[3]S. A. Lindgren and L. Wallden, Phys. Rev. B **22**, 5969 (1980).
[4]H. Tochihara and Y. Murata, J. Phys. Soc. Jpn. **51**, 2920 (1982).
[5]H. Tochihara, Surf. Sci. **126**, 523 (1983).
[6]J. D. Levine, Surf. Sci. **34**, 90 (1973).
[7]R. U. Rartinelli, Appl. Phys. Lett. **17**, 313 (1970).
[8]T. Aruga, H. Tochihara, and Y. Murata, Phys. Rev. Lett. **53**, 372 (1984).
[9]D. M. Newns, Phys. Lett. **39A**, 341 (1972).
[10]M. Nakayama, T. Kato, and K. Ohtomi, Solid State Commun. **50**, 409 (1984).
[11]M. Tsukada, H. Ishida, and N. Shima, Phys. Rev. Lett. **53**, 376 (1984).
[12]N. Shima, J. Phys. Soc. Jpn. **52**, 578 (1983).
[13]M. Aono, Y. Hou, C. Oshima, and Y. Ishijima, Phys. Rev. Lett. **49**, 567 (1982).
[14]D. J. Chadi, Phys. Rev. Lett. **43**, 43 (1979).
[15]J. Ihm, M. L. Cohen, and D. J. Chadi, Phys. Rev. B **21**, 469 (1972).
[16]A. A. Lucas and M. Sunjic, Phys. Rev. Lett. **26**, 229 (1971).
[17]J. Harris, Solid State Commun. **16**, 683 (1971).
[18]G. Lehmann and M. Taut, Phys. Status Solidi **54**, 469 (1972).