Total energies in Se. II. Vacancy in the crystal

David Vanderbilt
Department of Physics, University of California at Berkeley, Berkeley, California 94720

J. D. Joannopoulos
Department of Physics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139
(Received 7 September 1982)

Ab initio total-energy calculations are used to determine theoretically the structural configuration of the vacancy in trigonal Se. The method consists of calculating the forces, as well as the total energies, within the local-density and frozen-core approximations, for a superlattice structure containing a vacancy. In this way, relaxations at the vacancy can be fully taken into account, including a possible self-healing of the vacancy. A slightly relaxed symmetric version of the ideal vacancy is the lowest-energy structure found; neither asymmetric relaxation nor valence alternation appears to occur. A simple Hubbard Hamiltonian is used to analyze the spin configuration of the lowest-energy structure.

I. INTRODUCTION

In the preceding paper (hereafter referred to as paper I), it was shown that accurate structural information on trigonal selenium may be obtained by calculating the total energy as a function of structure, using local-density theory and *ab initio* pseudopotentials. Here, this approach is extended to the study of the neutral vacancy in trigonal Se.

The vacancy is an interesting defect because its structure is not easily guessed *a priori*. The unreconstructed vacancy has a single atom removed from one of the chains of the trigonal structure, and therefore gives rise to two dangling bonds, one on each of the chain-termination atoms above and below the vacancy. In this symmetric vacancy structure V_s, shown in Fig. 1(a), each of the two defect atoms is neutral and is therefore expected to have a singly occupied "dangling-bond" gap state, very much like the C_1^0 defect proposed for glassy Se.[1–3] The structure has an exact C_2 symmetry about an axis passing through the vacancy site, so that (in the absence of spin splitting) one has a state of even symmetry and one of odd symmetry; the lower of these is fully occupied, while the higher is unoccupied. Thus no Jahn-Teller distortion is required.

Nevertheless, several interesting reconstructions are possible candidates for the lowest-energy (stable) vacancy. One possibility is a charge transfer from one dangling bond to the other. While this is unfavorable in terms of Coulomb interactions, the ensuing lattice distortions might stabilize the structure, in analogy to the proposed stability of the D^+D^- pair in chalcogenide glasses.[1,2] If this distortion takes the form of simply introducing an asymmetry in the lengths of the bonds adjoining the two oppositely charged defect atoms, the asymmetric vacancy structure V_a of Fig. 1(b) results. On chemical grounds, however, the positively charged atom might be expected to bond to a neighboring chain to form a threefold-coordinated site, analogous to the proposed C_3^+ defect in the glass.[2] The reconstructed vacancy V_r of Fig. 1(c) then results.

An intriguing alternative candidate for the stable vacancy structure is shown in Fig. 1(d). In this

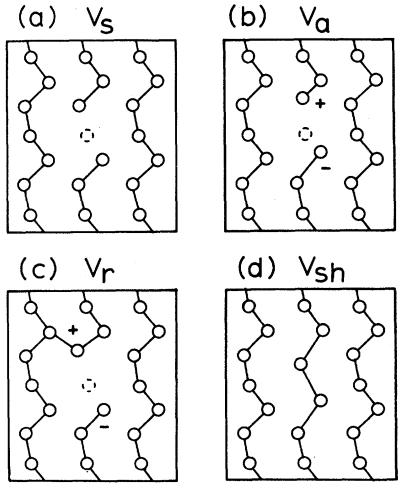

FIG. 1. Schematic view of proposed vacancy structures. (a) Symmetrically relaxed version of ideal vacancy. (b) Asymmetrically relaxed version. (c) Reconstructed structure embodying a threefold-coordinated Se site. (d) Self-healed structure, in which bridging bonds span the vacancy.

self-healed vacancy structure V_{sh}, the two dangling bonds have bonded together to reconstitute the original chain. This requires substantial displacements of several atoms on each side of the original vacancy site; the resulting structure has a region over which one chain is "stretched" to make up for the missing atom. In the limit that the interchain interactions are negligible compared to the intrachain interactions, this self-healed structure will almost certainly occur, because the energy cost in intrachain elastic energy can be made arbitrarily small by making the stretched region arbitrarily long. In trigonal Se, the interchain forces are weak but by no means negligible (see paper I). If the solitonlike V_{sh} structure were stable in trigonal Se, it would presumably be extremely mobile and would soon migrate to a surface, defect, or impurity. Thus vacancies created e.g., by bombardment would appear to self-annihilate rapidly on experimental time scales.

Finally, it is possible that the vacancy has the symmetric structure V_s, but reconstructs electronically by acquiring a spin-polarization in which a spin-up electron resides on one dangling bond while a spin-down electron resides on the other. This "antiferromagnetic" model bears a resemblance to models for spin ordering among dangling bonds on the Si(111) surface.[4,5] However, for a finite spin system such as the vacancy, no sharp transition occurs; the question here is not the existence, but the degree, of such ordering.

In this paper, we calculate the total energy of the various candidates for the stable vacancy, taking relaxations fully into account by calculating the Hellman-Feynman forces. We find that among the non-spin-polarized candidates, the symmetric V_s is more stable than the charge-transfer configurations V_a and V_r. Our results with respect to V_{sh} are not definitive; we did not find a configuration more stable than V_s, but a more thorough search would have to be done before V_{sh} could be ruled out. Finally, a model spin Hamiltonian with parameters taken from the realistic calculation indicates that spin polarization will be present but weak at V_s.

In Sec. II we describe the structural supercell models used in the calculations. Section III contains a discussion of some of the details of the total energy and force calculations. The results are given and discussed in Sec. IV. Finally, Sec. V contains a summary and conclusions.

II. STRUCTURAL MODELS

In order to model the vacancy structure, we have constructed a superlattice whose supercell contains eight atoms and one vacancy. This is done by removing every ninth atom from the trigonal crystal structure in the pattern shown in Fig. 2. The cells have been chosen in this staggered manner so as to maximize the distance between neighboring vacancies.

Let the lattice vectors for trigonal Se be

$$\vec{a}=\left[\frac{a}{2},\frac{a\sqrt{3}}{2},0\right],$$
$$\vec{b}=\left[\frac{a}{2},-\frac{a\sqrt{3}}{2},0\right], \tag{1}$$
$$\vec{c}=(0,0,c).$$

The lattice vectors for the superlattice structure of Fig. 2 can be written

$$\vec{a}_s=\left[\frac{a}{2},-\frac{a\sqrt{3}}{2},-c\right],$$
$$\vec{b}_s=\left[\frac{a}{2},\frac{a\sqrt{3}}{2},-c\right], \tag{2}$$
$$\vec{c}_s=(-a,0,-c).$$

This forms a convenient set because the superlattice, while strictly trigonal, has $\beta=66.0°$ and is therefore

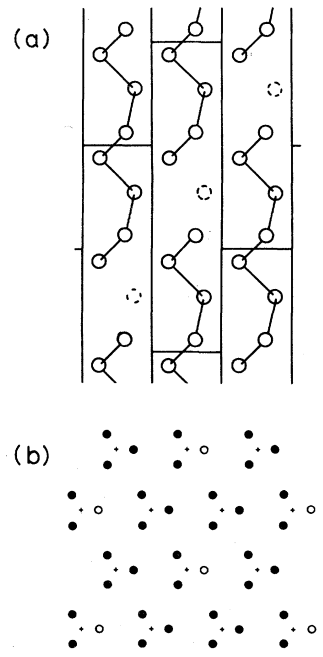

FIG. 2. (a) Side view of superlattice structure containing a vacancy (dotted circles) and eight atoms (full circles) per cell. (b) Top view of a slab three atoms thick taken from the same structure; vacancies (open circles) are maximally spaced among normal Se atoms (black dots). Plus signs (+) indicate chain axes.

almost fcc. Thus special k points[6] can be chosen as though for an fcc structure.

For the lattice constants a and c, we use the *calculated* values from paper I, so that in the absence of a vacancy we have a fully relaxed structure. The calculated values $a=3.97$ Å and $c=4.91$ Å imply that each vacancy site has six nearest-neighbor vacancies 6.32 Å away in the directions $\pm\vec{a}_s$, etc, and six second-neighbor vacancies 6.88 Å away in the $\vec{a}$-$\vec{b}$ plane of Fig. 2(b). It should be pointed out that for the case of V_s, the C_2 symmetry of the vacancy is honored by the *basis,* but not by the *lattice,* of our model. Thus only the local environment of a given vacancy (up to a radius less than 6.32 Å) has C_2 symmetry with respect to atom locations. Thus the C_2 symmetry may be said to be weakly broken for the superlattice geometry of Fig. 2(a).

The structures used to model the various hypothetical vacancy types V_s, V_a, V_r, and $V_{\rm sh}$ were all obtained by keeping the lattice constants of the superlattice fixed and moving the eight atoms within the cell. The choice of atom locations was usually guided by a previous force calculation, in a manner to be described shortly.

III. CALCULATIONAL METHOD

The calculation of the total energy for the vacancy structures was done using the method of paper I. This uses the momentum-space representation to calculate the total energy[7] with Wigner exchange correlation[8] and an *ab initio* norm-conserving pseudopotential.[9] The energy cutoffs for the plane-wave basis were reduced slightly from those of paper I to make the calculation tractable for the larger cells used here. The search for the minimum energy structure was carried out using $E_A=2.5$ Ry and $E_B=8.33$ Ry for the lower and upper Löwdin cutoffs;[10] the total energy for the relaxed structure was then evaluated more accurately using $E_A=4.30$ Ry and $E_B=14.33$ Ry (corresponding to $N_A\approx 205$ and $N_B\approx 1040$ plane waves, respectively). Tests were carried out at higher cutoffs to insure that adequate convergence had been obtained.

In order to carry out the present calculation, two extensions of the method were necessary. Firstly, the large size of the unit cell and the presence of degenerate half-occupied gap states give rise to a charge instability in the iterative procedure which is eliminated using a dielectric matrix method. Secondly, the lack of symmetry and consequent large number of independent structural degrees of freedom make it almost impossible to find the variational minimum of the total energy without calculating the forces. The calculations of the forces using the Löwdin perturbation scheme[10] require some care, but can be used in conjunction with a simple force-constant model to predict displacements which lead rapidly to the variational minimum. This section will be concerned primarily with these two extensions. In addition, strain energies associated with the supercell periodic boundary conditions are discussed, and are shown to be unimportant.

Let us start with the dielectric matrix method. The charge instability problem referred to above can be understood as follows. Let $V^{(n)}$ be the mean-field potential obtained on the nth iteration, $\rho^{(n)}$ the charge density obtained by solving the Schrödinger equation for $V^{(n)}$, and $U^{(n)}$ be the new potential constructed by screening with $\rho^{(n)}$. If $V^{(n)}$ were the converged potential $V^{(0)}$, we would have $U^{(n)}=V^{(n)}=V^{(0)}$. Now suppose we add a perturbation:

$$V^{(n)}(r)=V^{(0)}(r)+\delta V\cos(Gr)\ . \tag{3}$$

Then we expect, by the definition of the dielectric function $\epsilon(q,q')$, that

$$U^{(n)}(r)=V^{(0)}(r)+\sum_{G'}[\delta_{GG'}-\epsilon(G,G')]\delta V\cos(G'r)\ . \tag{4}$$

Here, the perturbing potential $\delta V\cos(Gr)$ is the sum of the external perturbation $V^{(n)}-U^{(n)}$ and the induced perturbation $U^{(n)}-V^{(0)}$. If we approximate

$$\epsilon(G,G')\approx\epsilon(G)\delta_{GG'} \tag{5}$$

then

$$U^{(n)}(r)=V^{(0)}(r)+[1-\epsilon(G)]\delta V\cos(Gr)\ . \tag{6}$$

Clearly for $\epsilon>2$, the perturbation has an oscillatory divergent behavior. Since $\epsilon(q\to 0)\approx 10$–$20$ in many semiconductors, this is a severe problem for the small reciprocal lattice vectors that occur for large unit cells. This problem can be solved by employing a wave-vector–dependent damping parameter $\alpha(G,G')$ in the construction of the new potential:

$$V^{(n+1)}(G)=V^{(n)}(G)+\sum_{G'}\alpha(G,G')\times[U^{(n)}(G')-V^{(n)}(G')] \tag{7}$$

or, using the approximate Eq. (5),

$$V^{(n+1)}(G)=\alpha(G)U^{(n)}(G)+[1-\alpha(G)]V^{(n)}(G)\ . \tag{8}$$

From Eq. (6), the best α is $\alpha(G)=\epsilon^{-1}(G)$. We use a Fermi-Thomas (FT) dielectric function with a small-q cutoff:

$$\epsilon(G)=1+k_{\rm FT}^2/(G^2+q_0^2)\ . \tag{9}$$

By treating k_{FT} and q_0 as adjustable parameters, it is possible to find values which give fast convergence. A scheme of this type was first proposed by Kerker.[11]

The use of the diagonal $\alpha(G)$ above is adequate for bulk calculations in large supercells, but there is an even more severe instability which arises because of the presence of partially occupied defect states on the two dangling bonds. In this case, charge can oscillate from one defect state to the other in the self-consistent iteration process. Consider, for example, a simple two-state model in which the self-consistent bonding state

$$\psi_b=\frac{1}{\sqrt{2}}(\phi_1+\phi_2) \tag{10a}$$

is doubly occupied, while the antibonding state

$$\psi_a=\frac{1}{\sqrt{2}}(\phi_1-\phi_2) \tag{10b}$$

is empty. Let the splitting between ϵ_b and ϵ_a be $2v$, where v is the hopping matrix element between ϕ_1 and ϕ_2. Now consider a small perturbation

$$H'=-\beta a_1^\dagger a_1+\beta a_2^\dagger a_2 , \tag{11}$$

where $a_1^\dagger$ and $a_2^\dagger$ are creation operators for ϕ_1 and ϕ_2. Then perturbation theory gives that the charge on state ϕ_1 will increase by

$$\Delta q_1=(\beta/v)q_e \tag{12}$$

on the next iteration. This then gives rise to a new perturbation H' of the form (11) with

$$\begin{aligned}\beta^{\text{new}}&=U\,\Delta q_1\\&=(U/v)\beta^{\text{old}} ,\end{aligned} \tag{13}$$

where U is some unscreened Coulomb correlation energy for adding a second electron to the dangling bond. For a free atom, $U\approx 8$ eV, and the vacancy dangling bonds interact by $v\approx 0.2$ eV, so that the charge instability of Eq. (13) can grow by a factor of ~ 40 for certain reciprocal lattice vectors of the vacancy superlattice structure. The result is that the diagonal $\alpha(G)$ of Eq. (9) is inadequate to damp the oscillation.

The problem is solved by going back to Eq. (7) and using $\alpha(G,G')=\epsilon^{-1}(G,G')$.[12] Since the oscillatory behavior comes primarily from the small G vectors, we calculate $\epsilon(G,G')$ only for $|\vec{G}|, |\vec{G}'|<G_{\max}$. Moreover, since most of the divergent behavior comes from small energy denominators, we impose a cutoff $E_{\max}$ upon the size of the energy denominators considered in calculating $\epsilon(G,G')$. We add to the calculated $\epsilon(G,G')$

$$\epsilon(G,G')=\delta_{GG'}-\frac{4\pi e^2}{G^2\Omega}\sum_{k,n,n'}\Theta(E_{n'k}-\epsilon_F)\Theta(\epsilon_F-E_{nk})\times\frac{\langle n'k|e^{iG'r}|nk\rangle\langle nk|e^{-iGr}|n'k\rangle+\langle nk|e^{iG'r}|n'k\rangle\langle n'k|e^{-iGr}|nk\rangle}{E_{nk}-E_{n'k}} \tag{14}$$

a contribution $\gamma\epsilon_{FT}(G)$ to make up for contributions from denominators greater than $E_{\max}$. For $|\vec{G}|$ or $|\vec{G}'|>G_{\max}$, $\epsilon(G,G')$ is taken to be diagonal, and equal to $\epsilon_{FT}(G)$. The parameters $G_{\max}=1.8$ a.u.$^{-1}$ (~ 60 $\vec{G}$ vectors), $E_{\max}=5$ eV, and $\gamma=0.8$ are found to allow rapid convergence (i.e., reduction of maximum errors in $V(G)$ by a factor of 3 on each iteration), without increasing the overall cost of the computer calculations by more than a few percent.

We consider now the problem of calculating the Hellman-Feynman forces. The expression for the forces in the momentum-space representation has been given previously.[7] However, one comment about the calculation of forces using Löwdin perturbation theory[10] should be made. In the Löwdin theory, one can choose whether or not to work with perturbed *wave functions* as well as eigenvalues (folding out the wave function). We have found empirically that folding out the wave functions is not very important in obtaining good total energies (paper I). However, it is *essential* if one wants to obtain reasonable Hellman-Feynman forces. Presumably this is because there is no variational principle on the force. The use of the folded-out wave functions in conjunction with nonlocal pseudopotentials leads to an expensive double sum in the evaluation of the forces. Therefore, we calculate the *local* contribution to the forces using the *folded-out* wave functions, but the *nonlocal* contribution is calculated using the *unperturbed* wave functions. When the full double sum was done in test cases, the correction was found to be negligible.

Once the forces are in hand, we need to use them to predict atomic displacements which will relieve the forces. That is, we wish to guess a structure which will give rise to vanishing forces (the equilibrium or lowest-energy structure). Let F_i^μ be the force in the ith Cartesian direction on atom μ; we can expand

$$F_i^\mu(\vec{R}+\vec{\Delta})\approx F_i^\mu(\vec{R})-\sum_{j\nu}K_{ij}^{\mu\nu}\Delta_j^\nu , \tag{15}$$

where $K_{ij}^{\mu\nu}$ is the dynamical matrix. Clearly we want to choose

$$\Delta_i^\mu = \sum_{j,\nu} (K^{-1})_{ij}^{\mu\nu} F_j^\nu(\vec{R}) . \tag{16}$$

The key obtaining rapid convergence to the equilibrium structure is the choice of a model for $K_{ij}^{\mu\nu}$. The simplest choice is

$$K_{ij}^{\mu\nu} = \delta_{\mu\nu}\delta_{ij}\kappa . \tag{17}$$

This corresponds to moving each atom in the direction the force is pointing, by an amount proportional to the force. This model is empirically found to be highly inadequate in our case. Because of the strong intrachain bonds and weak interchain bonds in Se, the force constants are highly anisotropic. We have found that using a simple Born—von Kármán force-constant model improves drastically upon Eq. (7) in terms of convergence speed. It contains three bond-stretching parameters fit to some test calculations of the total energy of trigonal Se. For nearest neighbors, second-intrachain neighbors, and nearest-interchain neighbors, they are $k_r = 1.19$ Ry Å^{-2}, $k_\theta = 0.1$ Ry Å^{-2}, and $k_R = 0.1$ Ry Å^{-2}, respectively. An additional contribution of the form (17), with $\kappa = 0.25$ Ry Å^{-2}, was added in to eliminate vanishing eigenvalues of K.

In some of the vacancy models studied (e.g., V_r or $V_{\rm sh}$, Fig. 1) it is not possible to uniquely define the four nearest interchain neighbors of a given atom. We have therefore made it a practice to identify a set of possible interchain neighbors for each atom, corresponding to those neighbors (other than first- and second-intrachain neighbors) at a distance less than some cutoff radius $R_{\rm max}$. For these neighbors, we include a force constant

$$k_R' = k_R \exp[-\alpha(R - R^0)] , \tag{18}$$

where R^0 is the crystalline nearest interchain neighbor distance (d_2 of paper I) and α is a constant. In addition, we have found it useful to define a fictitious force F_t of the form

$$F_f = k_F \Theta(R^0 - R)(R - R^0)\exp[-\alpha(R - R^0)] \tag{19}$$

from each such possible interchain neighbor, and to add these contributions to the Hellman-Feynman forces before inserting in Eq. (16). These repulsive forces are usually quite small or zero (only neighbors closer than R^0 contribute), but they do prevent the chains from cross-linking in certain topologies which are unavailable to the true vacancy. (That they are otherwise allowed, in our calculation, is an artifact of the superlattice geometry.) The parameters $R_{\rm max} = 3.9$ Å, $\alpha = 3$ Å^{-1}, and $k_F = 5.0$ Ry Å^{-2} have been adopted.

The use of this simple model in conjunction with Eq. (16) allows us to predict displacements which converge quickly upon the equilibrium geometry. Initial geometries were guessed for V_s, V_r, and $V_{\rm sh}$ of Fig. 1 by requiring that all nearest-neighbor bond lengths and bond angles (including those at the threefold-coordinated site) be equal to those of the crystal. (For V_a asymmetric displacements away from V_s were made by hand.) Typically, ~ 3—4 cycles using Eq. (16) were needed to remove $\sim 90\%$ of the strain energy of the initial configuration. In no case did the total energy increase as a result of a predicted set of displacements.

The dielectric matrix and the force-constant matrix discussed in this section are philosophically similar. Both are aids for rapidly approaching the (electronic or structural) variational minimum. The various approximations made in modeling $\epsilon(G,G')$ and $K_{ij}^{\mu\nu}$ do not effect the final result; they only effect the speed with which the calculation converges on this result (and whether it converges at all). For the vacancy in Se, the simple models for $\epsilon(G,G')$ and $K_{ij}^{\mu\nu}$ presented here make possible an otherwise intractable calculation.

Finally, we have also tested how the structural relaxation energy changes when the constraints arising from periodic boundary conditions (PBC's) are removed. The tests were carried out for V_r, whose cross-linked geometry might give rise to substantial PBC-related stresses. The total energy was monitored as these stresses were relieved by going to larger unit cells, up to 72 sites per cell (71 atoms and a vacancy). Because it is prohibitive to carry out the pseudopotential calculations for such large cells, we have instead used an eight-parameter Keating-type force constant Hamiltonian[13] for Se to monitor the relaxation energy. The Keating model is implemented by deleting any couplings involving the removed vacancy atom, and by treating the extraordinary bond at the threefold site on the same footing as all other intrachain bonds. For each supercell size, the structure is relaxed iteratively until the minimum energy configuration is reached. The remaining strain energy for V_r is found to be 0.068 eV for the supercell of Fig. 2, and reduces to 0.060 eV for the 71-atom supercell. Thus the PBC-related strain energy is expected to be $\lesssim 0.01$ eV for the Keating model. Of course the Keating model may not be accurate at the defect sites, but we may take this as an order-of-magnitude estimate of the true PBC-related strain energy, which we claim must therefore be $\lesssim 0.05$ eV. For V_s, it must be much less. As we shall see, corrections of this magnitude cannot change our conclusions in any way, and they will hereafter be dropped.

IV. RESULTS AND DISCUSSION

Consider the unrelaxed version of the symmetric vacancy structure V_s of Fig. 1(a). Since this structure is an important reference point for what follows, we begin by describing its electronic states in some detail.

In an earlier paper,[14] we showed that many essential features of the defect states in Se could be understood in terms of a very simple tight-binding model, having three p orbitals per site interacting with first neighbors only, and having bond angles and dihedral angles set equal to 90°. For this model, the p_x, p_y, and p_z subsystems of orbitals on a chain decouple from each other and from those of neighboring chains. In the bulk, each such chain consists of alternating σ-bond orbitals and nonbonding p orbitals, coupled by π interactions. When the chain is interrupted by a vacancy, as in Fig. 3, one obtains a dangling-bond state on the end of each remaining half-chain. Each such dangling-bond state is made out of that subsystem (p_x, p_y, or p_z) for which there are two *adjacent* nonbonding p orbitals at the end of the chain without an intervening σ-bond orbital. (These are shown in Fig. 3.) The dangling-bond state consists primarily of a π^* combination of these two nonbonding p orbitals, with most of the weight on the site adjoining the vacancy.[14] For the real vacancy, the two dangling-bond states on either side should interact weakly. The overall vacancy defect states are thus expected to be symmetric and antisymmetric combinations (ϕ_g and ϕ_u) of the two dangling-bond states, the lower being doubly occupied and the upper being empty.

Figure 4 shows that this is indeed the case. The charge density[15] of the filled symmetric state is shown at the end of a chain [Fig. 4(a)] and across the vacancy [Fig. 4(b)]. (These views correspond to the x-z plane passing through sites 1-2, and the y-z plane passing through sites 1-V-1′, respectively, in Fig. 3.) The unfilled antisymmetric state, not shown, lies $\sim$0.5 eV above the symmetric state, indicating a hopping matrix element of $\sim$0.25 eV between the dangling-bond states on either side of the vacancy. (The absolute energy location of these states in the gap is not uniquely defined for a superlattice calculation of this kind, because of an ambiguity in the zero of the Coulomb potential.)

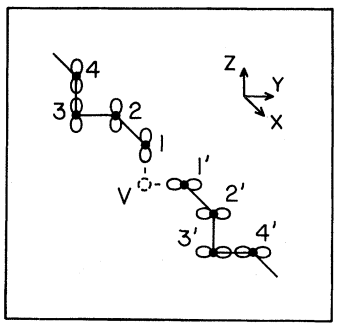

FIG. 3. Schematic view of p orbitals which participate in dangling-bond gap states in simple tight-binding model. Two dangling-bond gap states consist primarily of π^* combinations of p_z orbitals on sites 1 and 2, and of p_y orbitals on sites 1′ and 2′.

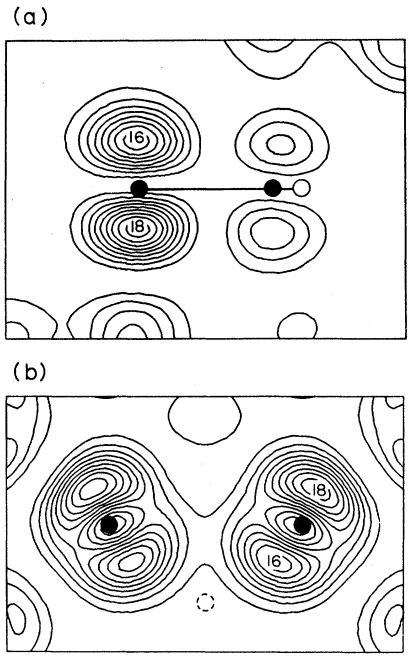

FIG. 4. Charge densities for occupied gap state of ideal vacancy: (a) for plane passing through the last two atoms on chain (black circles), and normal to bond angle with the third atom (open circle, out of plane); (b) for plane passing through vacancy (dotted circle) and two neighboring sites (black circles).

Note that the covalent bond between each of the onefold-coordinated atoms and its nearest neighbor is a unique one. For the neutral case, there is not only a σ-bond component, but also a half π bond. (For a positively charged chain end, this would be a full π bond, because the π^* state would be empty.) This partial π-bonding character does not occur in the bulk, and therefore gives rise to a *new kind of bond* at the dangling chain ends by the vacancy.

We have calculated the total energy of the unrelaxed vacancy to be 1.44 eV, measured relative to the energy of the same number of bulk atoms. (That is, 1.44 eV is the activation energy for creating vacancies in the bulk, or, the energy released per vacancy as vacancies migrate to the surface.)

So far, we have considered only the ideal (unrelaxed) vacancy structure. If we use the methods of the preceding section to calculate the forces on this structure, we find that they are small but not zero. After three cycles of calculating forces, we find that the total energy stabilizes at 1.34 eV. The bonds adjoining the one-fold-coordinated atoms have shortened slightly to 2.290 and 2.300 Å. The lengths of the other bonds along the chain remain near the 2.367-Å bulk value,[16] with an rms deviation of 0.028 Å from this value. The last bond angles at the chain ends have the values 105.80° and 103.91°, compared to 102.60° in the bulk. The small differences between the relaxations on either side of the vacancy are due to the broken C_2 symmetry of the superlattice; for the true vacancy, this would be a symmetric relaxation, giving rise to a relaxed V_s [Fig. 1(a)].

To summarize, the relaxations giving rise to V_s are small, are dominated by a shortening of the final bonds by ~ 0.07 Å, and are reflected in a lowering of the total energy by ~ 0.1 eV. The electronic structure is essentially unchanged from that of the unrelaxed vacancy described above.

The V_s structure found here corresponds to a local minimum of the total energy in structural space, but it need not be a global minimum. Even if it is, it would be interesting if some of the other structures (eg., V_a or V_r) were local minima; they would then be metastable structural configurations. We have therefore "manually" constructed structures of the kind V_a, V_r, and $V_{\rm sh}$ of Figs. 1(b)–1(d), and then allowed the forces to relax the structure with these guesses as starting points.

Consider first the V_a structure of Fig. 1(b). The motivation for considering such a structure is related to the partial π bond at the onefold-coordinated atom, discussed above. The neutral chain end has a half π bond. If an electron were to be transferred from one chain end to the other, the positive chain end would have a full π bond, while the negative chain end would have no π bonding. Thus the final bond on the positive chain end might be expected to shorten, while that of the negative would lengthen. It is possible that this structural relaxation could support the charge transfer, in the manner of a "negative U." We find that this is not the case; the asymmetric structure V_a simply relaxes back to the symmetric structure V_s. The shorter bond of the V_a does become somewhat positively charged, but the effect is evidently not strong enough to overcome the strain energy of creating the asymmetry.

The motivation for considering the V_r structure of Fig. 1(c) is similar. As mentioned in Sec. I, the situation is very reminiscent of the negative U defect model for glassy chalcogenides[1,2] in which the formation of a "valence alternation pair" (consisting of a positive threefold and a negative onefold-coordinated site) is proposed to be energetically favorable. We constructed an initial guess for the V_r structure by demanding that all of the nearest-neighbor bond lengths and bond angles (including those at the threefold-coordinated site) be identical to those of the trigonal chain in the bulk. This actually turns out to be a rather bad guess at a relaxed geometry; after 2–3 cycles of calculating forces and relaxing the structure, we find that the total energy has been lowered by ~ 0.8 eV from the initial guess. However, it is still ~ 0.3 eV higher in energy than the relaxed V_s structure. At this point, the structure is still threefold coordinated on one side of the vacancy, although the extraordinary bond is becoming longer than the others. Further cycles of relaxation simply drive the structure back to the V_s configuration.

Finally, an initial guess at the self-healed vacancy structure $V_{\rm sh}$ was constructed by demanding that all of the bond lengths (including the one bridging the vacancy) be equal to that of the bulk, and that the bond angles be as close as possible to that of the bulk. In our superlattice model, a chain made out of eight atoms per cell has to span a distance which is covered by nine atoms in a normal chain, so that the bond angles must be widened. The resulting "initial guess" structure of $V_{\rm sh}$ is actually a perfectly helical chain which rotates 135° per atom (as opposed to 120° in the bulk) and has bond angles of 109.37°. The calculated total energy for this structure is 2.04 eV, or ~ 0.7 eV *higher* than the relaxed V_s.

This structure is perhaps not the most appropriate starting point. For an isolated $V_{\rm sh}$, one would expect one or a few bonds to be stretched most, with displacements decaying to zero outside this region (a kind of domain wall). We have also constructed structures in which we identify the displacements $\Delta\vec{R}$ which would take the initial $V_{\rm sh}$ structure into the V_s structure, and then displace the initial $V_{\rm sh}$ by $\alpha\Delta\vec{R}$, where $\alpha=\frac{1}{12}$ and $\frac{1}{4}$. For these structures, the bond bridging at the vacancy is longer than the others (2.464 and 2.771 Å for $\alpha=\frac{1}{12}$ and $\frac{1}{4}$, respectively), thus modeling the case in which one bond is stretched the most. The total energies are calculated as 1.99 and 1.88 eV, respectively, for these two cases. For both these intermediate cases, the forces indicate that further stretching of the weak bond is desirable. Thus the energy appears to be falling monotonically as we approach the V_s structure.

Because the energies are all so much higher than that of the V_s, a global minimum at a structure resembling $V_{\rm sh}$ seems unlikely. Moreover, no evidence of a local minimum has been found. However, we have not performed a sufficiently exhaustive search of structural configurations to rule out the

existence of a stable or (especially) metastable V_{sh} center.

Finally, we return to the V_s, which now appears to be the most likely structure for the vacancy, and consider its spin configuration. This is modeled using a simple two-state Hubbard Hamiltonian H_H:

$$H_H=\epsilon_0\sum_{i\sigma}n_{i\sigma}+v\sum_{\sigma}(a^{\dagger}_{1\sigma}a_{2\sigma}+\text{c.c.})+U'\sum_i n_{i\uparrow}n_{i\downarrow}\ , \tag{20}$$

where v is the hopping matrix element and U' is the screened Coulomb correlation energy. The subscripts $i=1,2$ label the dangling-bond states ϕ_1 and ϕ_2 with self-energies ϵ_0 on either side of the vacancy, σ labels the spin, $a^{\dagger}_{i\sigma}$ are creation operators, and $n_{i\sigma}$ are number operators. The eigenstates of the two-electron problem consist of a spin triplet with energy $E_T=2\epsilon_0$, and three spin singlet states. The spatial basis states for the singlets can be taken to be either

$$\begin{aligned}\Phi_1&=\phi_1(1)\phi_1(2)\ ,\\ \Phi_2&=\phi_2(1)\phi_2(2)\ ,\\ \Phi_3&=\frac{1}{\sqrt{2}}[\phi_1(1)\phi_2(2)+\phi_2(1)\phi_1(2)]\ ,\end{aligned} \tag{21}$$

or

$$\begin{aligned}\Phi'_1&=\phi_g(1)\phi_g(2)\ ,\\ \Phi'_2&=\phi_u(1)\phi_u(2)\ ,\\ \Phi'_3&=\frac{1}{\sqrt{2}}[\phi_g(1)\phi_u(2)+\phi_u(1)\phi_g(2)]\ ,\end{aligned} \tag{22}$$

where

$$\phi_{u,g}=\frac{1}{\sqrt{2}}[\phi_1\pm\text{sgn}(v)\phi_2]\ . \tag{23}$$

If the mixing angle θ is defined by

$$\tan\theta=4v/U' \tag{24}$$

then the lowest singlet state has energy

$$E_S=2\epsilon_0-2v\tan(\theta/2)\ . \tag{25}$$

Thus this singlet state is the ground state. For this state, the probability P_I of finding the electrons in one of the "ionic" configurations Φ_1 or Φ_2 is

$$P_I=\sin^2(\theta/2)\ . \tag{26}$$

This probability approaches $\frac{1}{2}$ for $U'\ll 4v$, and 0 for $U'\gg 4v$, and is therefore a measure of the importance of correlation between spins. The case $P_I\to 0$ corresponds to the highly correlated antiferromagnetic limit.

These spin-correlation effects should give rise to a correction to the non-spin-polarized local-density total energy E_{NSP}. To estimate this correction, we associate E_{NSP} with the expectation value of H_H in state Φ'_1 (two electrons in the "bonding" state):

$$\langle\Phi'_1\,|\,H_H\,|\,\Phi'_1\rangle=2\epsilon_0-2\,|\,v\,|+U'/2\ . \tag{27}$$

This association is valid because Φ'_1 is composed of only a single Slater determinant.[17] The spin-polarization correction ΔE_{SP} is thus given by the difference between Eqs. (27) and (25).

For the vacancy in Se, we estimate $v\sim 0.25$ eV, as discussed earlier in this section. To obtain an approximate U', we take the Coulomb correlation energy U for putting two electrons on an isolated dangling-bond state, and then reduce this by the attractive Coulomb energy arising from the oppositely charged sites on either side of the vacancy. Using the methods of the following paper (paper III of this series) (see the discussion of the density matrix method and the Madelung terms), we estimate 0.92 and -0.59 eV for the repulsive and attractive terms, respectively, giving an overall U' of ~ 0.33 eV. This is to be understood as a rough estimate only. Using these values, we obtain $\theta=72°$, $P_I=0.34$, and $\Delta E_{\text{SP}}=-0.03$ eV. The overall estimate for the total energy of the relaxed symmetric vacancy V_s is now 1.31 eV.

Thus the spin-polarization energy correction is quite small, and the ionic contributions to the ground state are only weakly suppressed. In short, the V_s exhibits a weak tendency towards antiferromagnetic ordering.

V. SUMMARY

A superlattice structure containing vacancies is constructed to model the vacancy in trigonal Se. Total energies and Hellman-Feynman forces are then calculated in the local-density approximation, using an *ab initio* nonlocal pseudopotential to represent the Se cores.

Several new features of the calculational method are discussed in some detail. First, it is shown that an approximate calculation of the dielectric matrix allows fast convergence of the self-consistent iterative procedure, whereas oscillatory divergences occur otherwise. Then, the implementation of the force calculation in the Löwdin perturbation scheme is discussed. Finally, a simple Born—von Kármán force-constant model is used to predict displacements which will relieve the forces, thereby relaxing the structure quickly and efficiently to its equilibrium configuration.

A variety of structures are considered for the Se vacancy. The simplest, a slightly relaxed version of the ideal symmetric vacancy structure, turns out to have the lowest energy of those studied. The activation energy to create this vacancy is ~ 1.3 eV. Structures which embody asymmetric relaxation, valence alternation (i.e., a threefold-coordinated Se

site), and self-healing, all have higher energy and appear to relax directly to the symmetric structure without passing over an energy barrier. These conclusions are tentative for the case of the self-healed vacancy, for which a more extensive search of structural configurations would be desirable.

Finally, the spin configuration of the lowest-energy symmetric vacancy structure is analyzed using a simple Hubbard Hamiltonian. The ground state is a spin singlet whose spatial wave function is a linear combination of ionic and neutral configurarations. It is estimated that the ionic configurations are only weakly suppressed, and that the spin polarization energy is only a few hundredths of an eV, indicating that the system is only weakly correlated.

ACKNOWLEDGMENTS

This work was supported through the MIT Center for Materials Science and Engineering via National Science Foundation (NSF) Grant No. DMR-76-80895. Additional University of California (Berkeley) support in the latter stages of the work was provided by NSF Grant No. DMR-78-22465 and by the Director, Office of Energy Research, Office of Basic Energy Sciences, Materials Sciences Division of the U.S. Department of Energy under Contract No. DE-AC03-76SF00098. The authors (D.V. and J.D.J.) would also like to thank the Miller Institute and the John S. Guggenheim Foundation, respectively, for receipt of fellowships.

[1]R. A. Street and N. F. Mott, Phys. Rev. Lett. 35, 1293 (1975); N. F. Mott, E. A. Davis, and R. A. Street, Philos. Mag. 32, 961 (1975).

[2]M. Kastner, D. Adler, and H. Fritzsche, Phys. Rev. Lett. 37, 1504 (1976); M. Kastner and H. Fritzsche, Philos. Mag. B 37, 199 (1978).

[3]D. Vanderbilt and J. D. Joannopoulos, Solid State Commun. 35, 535 (1980).

[4]C. B. Duke and W. K. Ford, Surf. Sci. 111, L685 (1981).

[5]J. E. Northrup, J. Ihm, and M. L. Cohen, Phys. Rev. Lett. 47, 1910 (1981).

[6]J. D. Chadi and M. L. Cohen, Phys. Rev. B 8, 5747 (1973).

[7]J. Ihm, A. Zunger, and M. L. Cohen, J. Phys. C 12, 4409 (1979).

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

[9]D. R. Hamann, M. Schluter, and C. Chiang, Phys. Rev. Lett. 43, 1494 (1979).

[10]P. Löwdin, J. Chem. Phys. 19, 1396 (1951).

[11]G. Kerker, Phys. Rev. B 23, 3082 (1981).

[12]K. M. Ho, J. Ihm, and J. D. Joannopoulos, Phys. Rev. B 25, 4260 (1982).

[13]R. M. Martin, G. Lucovsky, and K. Helliwell, Phys. Rev. B 13, 1383 (1976).

[14]D. Vanderbilt and J. D. Joannopoulos, Phys. Rev. B 22, 2927 (1980).

[15]The charge densities have been averaged over the supercell Brillouin zone, and then symmetrized with respect to the exact C_2 symmetry of the vacancy. The latter step is required because of the weak symmetry breaking introduced by the superlattice.

[16]All comparisons are made with respect to the *calculated* bulk structure so that differences will be meaningful.

[17]U. von Barth, Phys. Rev. A 20, 1693 (1979).