PHYSICAL REVIEW B VOLUME 21, NUMBER 12 15 JUNE 1980

Theory of the silicon vacancy: An Anderson negative-U system

G. A. Baraff, E. O. Kane, and M. Schlüter
Bell Laboratories, Murray Hill, New Jersey 07974
(Received 26 November 1979)

We present a quantitative theory of electron states associated with strongly lattice-coupled localized defects. Using the total energy functional, we derive optical transition energies, occupancy levels, and activation energies. The parameters of this theory are calculated for silicon vacancies. Electron levels and deformation potentials are derived by a self-consistent Green's-function technique. The elastic lattice response is calculated using a modified valence force model. We find that the charge states V^0, V^+, and V^{++} form an "Anderson negative-U system," and we predict two-electron transitions between V^0 and V^{++}. To test these results, the three parameters of our theory are then treated as fully adjustable and fitted to electron paramagnetic resonance and deep-level transient spectroscopy experiments. The fitted values also lead to a prediction of the two-electron transition. Thus, experiment alone strongly suggests that the silicon vacancy is an Anderson negative-U system. The results of the fitting confirm the correctness of our general theory and also demonstrate the accuracy and usefulness of our *a priori* calculations.

I. INTRODUCTION

The problem of calculating the electronic and structural properties of deep-level defects in semiconductors is an inherently interesting and difficult one, involving as it does the need to specify energies to within a fraction of the energy gap (the extent of the energy spectrum available to present-day experimental probes), the need to define levels in terms of total energy changes in a strongly coupled defect-lattice system, and the absence of a single dominant energy (rather than several energies of approximately equal size), which would aid in the construction of an appropriate zero-order theory.

For several reasons, the isolated vacancy in silicon is an ideal system on which to test such calculations. From the standpoint of *ab initio* input, the system is composed solely of atoms of a single type, silicon, for which there exist a number of well-investigated pseudopotentials.[1-3] From the standpoint of experimental richness, the vacancy has been studied for over fifteen years by electron paramagnetic resonance[4] (EPR), for which temperature, uniaxial stress, light, and a variable Fermi energy have been used as external probes and more recently, by deep-level transient spectroscopy[5] (DLTS) for which most of the same external probes have been employed. These studies have shown that the vacancy can exist in several charge states, that these charge states are associated with specific symmetry-lowering lattice distortions, and that the stress coupling coefficients for the various observed charge states stand in suggestively simple integer ratios to each other.[6]

Based on these and related studies, Watkins[7] has devised a qualitative description which nicely connects the form of the lattice distortion to the charge state of the vacancy. What has been missing, however, is a quantitative theory capable of calculating *occupancy levels* (carefully defined here as a value of the Fermi energy above and below which the equilibrium charge state of the vacancy is different) and *activation energies* (carefully defined here as a description of the exponential dependence of transition rates on inverse temperature), both of which can either be measured or bounded experimentally.

In the present work, we give more details of our recent progress towards such a theory. The first section of this paper describes the previously accepted picture of the vacancy and some of our new findings. It contains a warning about use of eigenvalues alone, without considering the elastic energy, when discussing experiments in the strongly coupled system. In Sec. II, we summarize some very general ideas which relate the levels to the total energy functional of the system. In constructing that functional, we build on insights of Watkins on which the previously accepted picture is based. In Sec. III, we use our Green's-function technique[8,9] to calculate directly some of the parameters of the functional and we use a modified Keating model[10] to evaluate semiempirically some of the others. From these emerge a picture of the vacancy which differs from the previously accepted interpretation in several interesting ways. First, for p-type silicon, we predict the stability of the doubly positively charged state of the vacancy V^{++}, a state which, being EPR invisible, has never been directly observed nor previously considered but whose existence is crucial in explaining some recently reported experiments. Second, we find that the states V^0, V^+, and V^{++} form an "Anderson negative-effective-U system,"[11] by

which is meant that the level (as defined above) is associated with a *two*-electron transition between V^0 and V^{++}. This is unusual, because normally one would expect that electron-electron repulsion requires an extra energy, call it U, to add a second electron to a singly occupied state. Thus, as the Fermi energy is raised, the sequence of *expected* equilibrium states would be V^{++}, V^+, and V^0. The essence of Anderson's "effective-negative-U" idea, however, is that an energy-lowering structural distortion may be sufficiently enhanced by the presence of a second electron that the energy gain more than compensates the e-e repulsive energy cost. In this case, a second electron will rapidly follow the first one on or off the center. The parameters we have calculated suggest that this indeed occurs for the states V^{++}, V^+, and V^0 and that consequently, the transition induced by a change of Fermi energy is between equilibrium states V^0 and V^{++}, while V^+ is *never* the stable state of lowest energy.

In the fourth section of the paper, we discuss and compare with experiment. First, we treat the parameters of our formal theory as fully adjustable and determine them to the extent we can by fitting to two of Watkins's experiments. The parameters determined in this way come close enough to what we had calculated *a priori* to confirm that (a) the general scheme relating the energy functional and the experimental observables is substantially correct and (b) our method of calculating the parameters of the theory is accurate enough to be useful. We emphasize here that the parametrized theory depends on only three parameters which can be bounded, using purely experimentally determined numbers, to yield the conclusion that the silicon vacancy is indeed an Anderson-negative-U system. Therefore this important prediction of our theory is substantially strengthened. We then turn to a consideration of Watkins's other experimental findings and show that they are in accord with the theoretical predictions. Finally, in the last section of the paper, we discuss shortcomings of the present approach, suggest lines for further theoretical and experimental investigation, and summarize our conclusions.

A. The Watkins vacancy model

Let us now review the previously accepted picture of the vacancy, which is due to the work of Watkins. The underlying model is the Coulson-Kearsley defect molecule,[12] where attention is focused on those orbitals in the crystal which are most perturbed by removal of the single atom to form the vacancy. As is illustrated in Fig. 1,

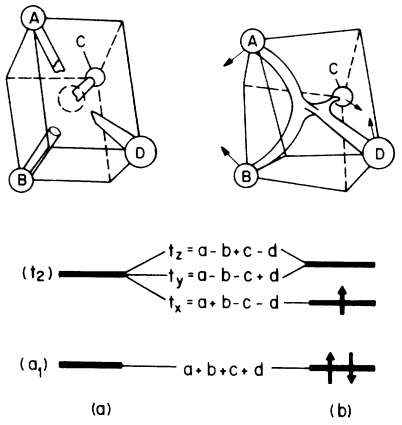

FIG. 1. One-electron orbital associated with the lattice vacancy: (a) undistorted, in tetrahedral symmetry; (b) after Jahn-Teller lattice distortion into a tetragonal symmetry.

there are four such orbitals, one on each of the four atoms nearest the vacancy, each associated with one of the four valence bonds broken by removal of the central atom. One-electron molecular orbitals are constructed as linear combinations of the four broken bond orbitals a, b, c, and d, associated with the atoms A, B, C, and D, respectively. Four such independent molecular orbitals can be constructed. The completely symmetric combination (a_1) is expected to be lowest in energy because it is nodeless. The other three are partner functions for a three-dimensional representation (t_2) of the T_d symmetry group of the undistorted crystal.

Most of the recent one-electron calculations of the electronic states of the undistorted silicon vacancy[9,13-16] confirm that the threefold-degenerate t_2 level lies in the gap and that the a_1 state, or the resonance to which it gives rise, is located about 1 eV below the top of the valence band. For the vacancy to be in the singly positively charged state V^+, three electrons must be distributed among these four states. Two of them go into the a_1 state (or that resonance) with their spins paired. The third goes into the t_2 level, a situation which, because of the orbital degeneracy, is unstable with respect to Jahn-Teller distortions.[17] The type of distortion, which drops the energy of the t_x state and raises the t_y and t_z energy *without* splitting them, is illustrated in Fig. 1(b). The atoms are shown pulling together in pairs, a distortion tending to rebond the four broken bonds into two unbroken ones. This distortion is tetragonal, with a $\langle 100 \rangle$ axis. Its symmetry group is D_{2d}.

For the vacancy to be in its neutral state, V_0, an additional electron is needed. This too goes into

the t_x state, further increasing the amount of tetragonal distortion and pairing the spins.

An additional electron is needed to produce the singly negatively charged state V^-. It goes into one of the unsplit t_2 levels, a situation which, because of the orbital degeneracy still remaining, is again unstable with respect to another Jahn-Teller distortion. The distortion needed this time to remove the degeneracy is trigonal, as is illustrated in Fig. 2. Its symmetry group is C_{2v}.

The doubly negatively charged state V^{--} has an additional electron in the b_1 state, increasing still further the amount of trigonal distortion. The two electrons in this state have their spins paired so that it, like V^0, is EPR invisible.

Experimental support for this picture is provided by Watkins's EPR experiments, out of which has also emerged the level scheme illustrated in Fig. 3. In that figure the range of Fermi energy between valence band and conduction band is shown divided by horizontal bars representing the occupancy levels. The region between levels denotes the range of Fermi energy over which the indicated charge state is stable. Figures such as this frequently appear in which the levels are labeled by an activation energy for some transition rate. Although the activation energy and occupancy level (as we have defined it above) may be numerically close, the concepts are so different that we wish to use such a diagram only to denote the level itself, stressing the role played by the Fermi energy in determining the equilibrium charge state.

Watkins placed the level $E(0/+)$ (above and below which the stable states are V^0 and V^+, respectively) closer than 0.05 eV to the valence band because, in boron-doped silicon, for which $E_F = 0.045$ eV, he observed that V^+ was unstable. He thus concluded that the region of stability of V^+ did not extend as high as 0.045 eV. The assignment and nature of this level is contradicted by our calculation. Our claim, in this paper and in the accounts of the work leading to it,[18,19] is that the lowest electrical level is not $E(0/+)$, but is $E(0/++)$.

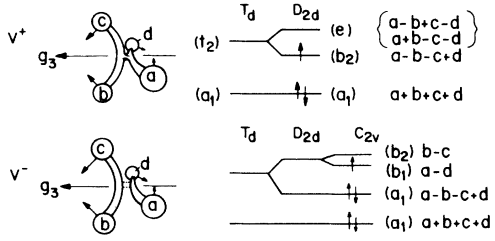

FIG. 2. One-electron molecular orbitals for the EPR active states of the silicon vacancy.

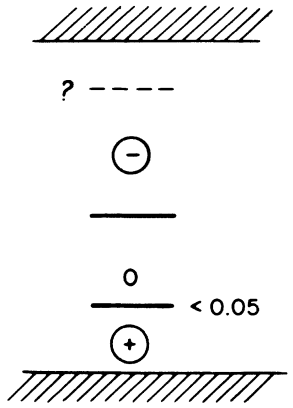

FIG. 3. Currently accepted occupancy level structure of the silicon vacancy. Horizontal bars are occupancy levels, above and below which the stable charge states are different. The circled charge states are those observed by EPR.

We assert that $E(0/++)$ lies below 0.16 eV (the Fermi level in In-doped silicon), probably below 0.10 eV, but certainly above 0.065 eV (the Fermi energy in Ga-doped silicon).

There is a close parallel between the tetragonal distortion associated with the sequence of states V^{++}, V^+, V^0, and the trigonal distortion associated with the sequence of states V^0, V^-, V^{--}. Although we are not going to calculate explicitly any energies connected to the trigonal distortion here, it is tempting to speculate on the possibility that there is but a single level, $E(--/0)$ in the upper part of Fig. 3, giving a level structure as in Fig. 4. Perhaps V^-, like V^+, is metastable at all values of Fermi energy, because the trio of states V^0, V^-, and V^{--} also form an Anderson negative-effective-U system, with the trigonal distortion playing the same role as does the tetragonal distortion for the trio of states V^{++}, and V^+, and V^0. We shall comment more about this possibility elsewhere.[20]

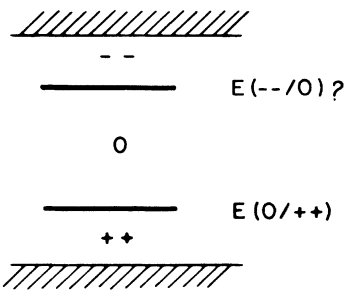

FIG. 4. Occupancy level structure which corresponds to having an Anderson negative-U system as predicted by our calculations.

B. The concept of level: A warning

The use of the term "level" in the context of these last two paragraphs and the symbol we choose to represent it is sufficiently nonstandard that we feel it deserves some elaboration. The term "level" as normally used describes in part the energy difference between two states of a system composed of a defect, for example, and a reservoir, in which the defect occupancy changes between $n+1$ and n. We make this explicit by using the notation $E(n+1/n)$ for the "level." The energy change in the transition is then $E(n+1/n)-\mu$, where μ is the energy per particle of the reservoir. We also consider two-electron transitions for which the "level" is denoted $E(n+2/n)$ and the energy change is $2[E(n+2/n)-\mu]$. With this slight generalization we call the level so defined an *occupancy level*, since it gives the value of the Fermi level for which the ground-state occupancy changes from n to $n+1$ (or n to $n+2$ in the case of an Anderson negative-U system). *This definition includes both electronic and lattice readjustment energies and thus is not given by a one-electron eigenvalue calculation.* It is *only* these levels (characteristic of two charge states), and *not* the one-electron eigenvalues, which can be measured in a quasiequilibrium experiment.

DLTS and related experiments measure yet another energy, an activation energy, which is a conceptually different quantity. Where there is strong lattice relaxation, failure to distinguish between activation energy and occupancy level, or worse, to compare either of them with an eigenvalue, leads to a complete misunderstanding and obscures whether theory and experiment are in agreement. A recent example of this confusion may be found in Ref. 38.

II. MODEL FOR THE TOTAL ENERGY OF THE SYSTEM

A detailed understanding of the electronic structure of the vacancy is complicated by the possibility that the associated electronic states, being rather localized, may be subject to sizable e-e correlations. The general question of the importance of correlations in dangling-bond states in semiconductors has been studied recently by Lannoo,[21] and there is still no definitive statement as to the importance of specific multiplet effects. However, Surratt and Goddard[22] have performed self-consistent configuration-interaction calculations on silicon vacancies using four-atom clusters and found ΔE, an averaged singlet-triplet splitting, ≈ 0.2 eV. This value certainly represents a generous upper limit since (a) silicon vacancy states have been found to extend further than four atoms[9,13,15] and (b) no crystal polarizability is included in their four-atom cluster calculation. The inclusion of both effects should lower ΔE considerably, according to estimates of ΔE vs cluster size as calculated by Watkins and Messmer.[23] Experimentally,[24] EPR findings strongly suggest that for vacancies in silicon, Jahn-Teller energies override or at least compensate e-e correlation effects.

Here we develop an account of the isolated vacancy in silicon based on this point of view, i.e., the neglect of correlation-induced spin-multiplet structure. The theoretical construct linking calculation and experiment is the total energy functional,[25] which we use in the local-density-functional (LDF) formalism.[26] Its dependence on position of atoms near the vacancy provides, for each set of electronic occupation numbers $\{n_i\}$, the proper potential energy for the lattice motion. It is not, however, the full Hamiltonian for the system because it does not contain the kinetic energy operator for the nuclear motion. Therefore, by using it as the fundamental object of interest, we are adopting a Born-Oppenheimer[27] approach. Specifically, its electronic states are adiabatic states, and transitions between them, which are induced by the nuclear kinetic energy operator treated as a perturbation,[28] are simply not a part of this limited description.

The energy functional $E_{\rm tot}$ expresses the total energy of the system of electrons and stationary nuclei in terms of the single-particle density matrix ρ,

$$\rho(\vec{r}',\vec{r}) \equiv \sum_i n_i \psi_i(\vec{r}')\psi_i^*(\vec{r}) \tag{2.1}$$

and $\vec{R}_A$, the nuclear or ion core positions. The functional itself is given by

$$E_{\rm tot}(\rho,\{R_A\}) = T + V + \Phi + V_I + E_{\rm xc} - N_T\mu, \tag{2.2}$$

where T, V, Φ, V_I, and $E_{\rm xc}$ are the electronic kinetic, electron-ion, electron-electron, ion-ion, and exchange-correlation energies, respectively:

$$T \equiv -\int \left[\tfrac{1}{2}\nabla_r^2 \rho(r,r')\right]_{r'=r} d^3r\,, \tag{2.3a}$$

$$V \equiv \int \rho(r)v(r)d^3r, \tag{2.3b}$$

$$\rho(r) \equiv \rho(r,r)\,, \tag{2.3c}$$

$$v(r) \equiv \sum_A U_{\rm ps}(|r-R_A|), \tag{2.3d}$$

$$\Phi \equiv \tfrac{1}{2}\int \frac{\rho(r)\rho(r')}{|r-r'|}\,d^3r\,d^3r' \equiv \tfrac{1}{2}\int \rho(r)\phi(r)d^3r, \tag{2.3e}$$

$$V_I \equiv \frac{Z^2}{2}\sum_A \sum_{B\neq A} \frac{1}{|R_A - R_B|}, \tag{2.3f}$$

$$E_{xc} \equiv \int \rho(r)\epsilon_{xc}(\rho)_{\rho=\rho(r)}\, d^3r\,. \tag{2.3g}$$

The last term in (2.2) has been introduced so that we can use E_{tot} to calculate occupancy levels: We consider our system as being connected to a reservoir containing noninteracting electrons, each of which has an energy μ, the Fermi energy. For a given value of μ, electrons will distribute themselves between the reservoir and the system in such a way as to achieve the lowest energy of the total entity—system plus reservoir. In this manner, we determine how the occupancy of the localized states depends on μ. We choose our energy reference corresponding to all valence-band states occupied and all levels above the valence band empty. Transferring $N_T = 0$, 1, or 2 electrons to the bound state in the gap changes the energy of the reservoir by $-N_T\mu$. This reservoir energy is included in E_{tot}.

A Schrödinger-type equation results from the requirement that E_{tot} be stationary with respect to variations of the density matrix which conserve the number of electrons in the system:

$$[-\tfrac{1}{2}\nabla^2 + v(r) + \phi(r) + U_{xc}(r) - \epsilon_i]\psi_i(r) = 0\,, \tag{2.4a}$$

$$U_{xc}(r) \equiv [(d/d\rho)\rho\epsilon_{xc}(\rho)]_{\rho=\rho(r)}\,. \tag{2.4b}$$

Solving this equation and evaluating (2.1) gives a density matrix ρ and eigenvalues ϵ_i which depend on the ionic positions $\{R_A\}$ and on the electronic occupation numbers $\{n_i\}$. Having determined these, it is possible *in principle* to obtain the numerical value of E_{tot} itself; i.e.,

$$E_{tot} = E_{tot}(\{n_i\}, \{R_A\}, \mu)\,, \tag{2.5a}$$

$$\epsilon_i = \epsilon_i(\{n_j\}, \{R_A\})\,. \tag{2.5b}$$

Just as the electronic occupation numbers at a particular value of μ are those which minimize E_{tot}, so too the ionic positions in equilibrium are those which minimize E_{tot}. Forces $F_\alpha(\{n_i\})$ and spring constants $K_{\alpha\beta}(\{n_i\})$ can then (again in principle) be obtained by expanding E_{tot} to second order in displacements Q_α away from some reference configuration $\{R_A^0\}$ and evaluating first and second derivatives at $Q_\alpha = 0$:

$$E_{tot}(\{n_i\}, \{R_A\}, \mu) = \mathcal{E}(\{n_i\}, \{R_A^0\}) - \sum_\alpha Q_\alpha F_\alpha(\{n_i\}) + \tfrac{1}{2}\sum_{\alpha\beta} Q_\alpha Q_\beta K_{\alpha\beta}(\{n_i\}) - N_T\mu. \tag{2.6}$$

With a reference configuration of sufficient symmetry, the spring-constant matrix is diagonal, and the minimum energy configuration is described by

$$Q_\alpha^0 = F_\alpha/K_{\alpha\alpha}\,, \tag{2.7a}$$

$$\Delta E_\alpha(\{n_i\}) = -\tfrac{1}{2}F_\alpha^2/K_{\alpha\alpha}\,, \tag{2.7b}$$

$$E_{min}(\{n_i\}, \mu) = \mathcal{E}(\{n_i\}, \{R_\alpha^0\}) + \sum_\alpha \Delta E_\alpha(\{n_i\}) - N_T\mu. \tag{2.8}$$

The meaning and use of Eqs. (2.6) and (2.8) is illustrated in Fig. 5 for the simplifying example of a single coordinate Q and a single occupation number N_T. Here we have plotted $E_{tot}(N_T, Q, \mu)$ vs Q, for three curves labeled by their value of N_T, using one value of μ in the top panel and another value of μ in the middle panel. Each of the two curves separately has the appearance of a configuration coordinate diagram, and the equilibrium value of Q for each N_T is at the minimum of that particular curve. The lowest of the three minima gives the ground state of the system. Because Eq. (2.6) involves μ, the state for which the system is lowest depends on μ, as is illustrated in

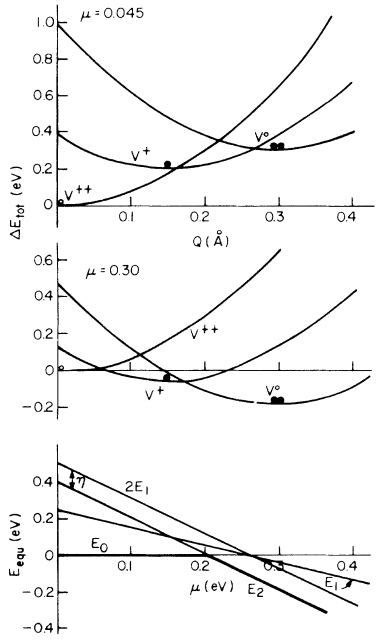

FIG. 5. Total energy functional $E_{tot}(N_T, Q, \mu)$ as a function of Q for three different charge states, $N_T = 0, 1, 2$. Two situations for different Fermi-level positions μ are shown. The bottom panel shows the equilibrium energies $(E_{equ}) = E_{min}(N_T, \mu)$ as a function of μ. The system ground state is indicated by a heavy line, coincident with $E_{min}(0, \mu)$ for $\mu < 0.2$ and coincident with $E_{min}(2, \mu)$ for $\mu > 0.2$.

Figs. 5(a) and 5(b). By plotting that minimum energy $E_{\min}(N_T, \mu)$ vs μ as done in Fig. 5(c), we can determine which N_T is the ground state for any value of μ and thus determine the μ dependence of the charge state of the vacancy. In spite of their similarity to configuration coordinate diagrams, *these μ-dependent curves are not to be used for calculating activation energies.*

Calculation of the activation energies involves use of the distribution function for the total entity, $\exp(-E_{\text{tot}}/kT)$, and study of its dependence on $\{n_i\}$ and $\{R_A\}$. We have investigated these elsewhere[19] for the specific example of the states V^{++}, V^{+}, V^{0}, although not in the language of the total energy functional. There is nothing to be gained by generalizing that work, which gave the same prescription as here for specifying occupancy levels. We shall instead simply quote, from Ref. 19, formulas for activation energies when they are needed, and shall give, at the end of this section, a configuration coordinate diagram from which both occupancy levels and activation energies can be read.

Let us reconsider the value of E_{tot}, Eq. (2.5a), which led to (2.6) and (2.8) and start to specialize it into a model for the states of the vacancy. There are three types of distortion which come immediately to mind. The first is a breathing distortion: When an atom is plucked from the crystal to create a vacancy, each of the four nearest neighbors is left with a single dangling bond pointing into the vacancy position, a situation not unlike that on the unreconstructed (111) surface of silicon, where each surface atom is left with a single dangling bond pointing perpendicularly out into the vacuum. The surface atoms undergo a relaxation one of whose components is motion of the surface plane towards the bulk of the crystal.[29] The back bonds of the surface atoms then shorten and strengthen.

The same sort of relaxation is to be expected at the vacancy. The four atoms nearest the vacancy should have an outward breathing component of motion which moves them into the bulk, strengthening their bonds with the rest of the crystal. Their outward motion will induce outward motion in *their* neighbors, and so on. We refer to this entire pattern of outward displacements as the breathing mode. It has the full symmetry of the T_d point group, and transforms as the A_1 (identity) representation.

Independent of this breathing mode, certain charge states of the vacancy have tetragonal symmetry. We can consider the four atoms nearest the vacancy moving to a tetragonally symmetric configuration with no breathing component of distortion, as depicted schematically in Fig. 1. Their motion will induce tetragonal displacements in *their* neighbors and so on. We refer to this entire pattern of displacements as the tetragonal mode. It transforms like the E representation of T_d.

Independent of the breathing mode and the tetragonal mode are purely trigonal modes.[17] Experiment[24] shows they play no role in the states in which we are interested here so we shall not consider them further. We have then *two* modes of interest, a breathing mode whose amplitude we denote Q_b, and a tetragonal (E) mode, whose amplitude we denote Q_E. We define Q_b and Q_E as the displacement of any one of the four nearest-neighbor atoms in the particular mode, although a different convention is often used in discussions of the Jahn-Teller distortion in XY_4 tetrahedral molecules.[17] The coordinates of our model are Q_E, Q_b, and N_T, the number of electrons on the bound state, i.e., B_2 state in the gap.

Let us now define $\epsilon_{1(2)}(Q_E, Q_b) - \mu$ as the additional energy needed to move the first (second) electron to the B_2 state from the reservoir when the lattice distortion is Q_E, Q_b. That is, we define

$$\epsilon_1(Q_E, Q_b) - \mu \equiv E_{\text{tot}}(N_T = 1, Q_E, Q_b, \mu) - E_{\text{tot}}(N_T = 0, Q_E, Q_b, \mu), \quad (2.9a)$$

$$\epsilon_2(Q_E, Q_b) - \mu \equiv E_{\text{tot}}(N_T = 2, Q_E, Q_b, \mu) - E_{\text{tot}}(N_T = 1, Q_E, Q_b, \mu), \quad (2.9b)$$

and for convenience in notation, let us denote the value of the *total* energy functional for the state V^{++} by $E_0(Q_E, Q_b)$,

$$E_0(Q_E, Q_b) \equiv E_{\text{tot}}(N_T = 0, Q_E, Q_b, \mu). \quad (2.9c)$$

Then the value of E_{tot} for the states of interest here are, for V^{++}, V^{+}, and V^{0}, respectively,

$$E_{\text{tot}} = \begin{cases} E_0(Q_E, Q_b), & N_T = 0 \quad (2.10a) \\ E_0(Q_E, Q_b) + \epsilon_1(Q_E, Q_b) - \mu, & N_T = 1 \quad (2.10b) \\ E_0(Q_E, Q_b) + \epsilon_1(Q_E, Q_b) + \epsilon_2(Q_E, Q_b) - 2\mu, & N_T = 2. \quad (2.10c) \end{cases}$$

Forces $F_\alpha(N_T)$ and spring constants $K_{\alpha\beta}(N_T)$ are obtained by expanding $E_{\text{tot}}(N_T)$ to second order about $Q_E, Q_b = 0$, as in (2.6). For example, let us write such an expansion for the case $N_T = 0$. This is the state V^{++}, and there is sufficient symmetry so that there is no mixed second-order term in the energy. Then we have

$$E_0(Q_E, Q_b) = \mathcal{E}_0 - \alpha Q_E - \beta Q_b + \tfrac{1}{2}(k_E Q_E^2 + k_b Q_b^2) . \tag{2.11}$$

Equilibrium values of the displacements, Q_E^0 and Q_b^0, are determined by minimizing $E_0(Q_E, Q_b)$, giving

$$Q_E^0(V^{++}) = \alpha/k_E , \tag{2.12a}$$

$$Q_b^0(V^{++}) = \beta/k_b . \tag{2.12b}$$

Since there is no Jahn-Teller mechanism to drive any tetragonal distortion for V^{++}, we know that $Q_E^0 = 0$, and infer then that $\alpha \equiv 0$. We shall return later to consider the remaining constants $\mathcal{E}_0$, β, k_E, and k_b in Eq. (2.11).

Now consider the expansion for the case $N_T = 1$, the state V^+. Instead of evaluating ϵ_1 in (2.9a) as the small difference between two large total energies, neither of which we can evaluate to the required accuracy, we make use of Slater's transition-state idea.[30] The occupation number N_T is treated as a continuous variable so that $E_{\rm tot}(N_T)$ can be regarded as an analytic function of N_T. This gives

$$\epsilon_1(Q_E, Q_b) - \mu = [\partial E_{\rm tot}(N_T, Q_E, Q_b, \mu)/\partial N_T]_{N_T = 1/2} + O(\Delta N)^3,$$

$$\epsilon_2(Q_E, Q_b) - \mu = [\partial E_{\rm tot}(N_T, Q_E, Q_b, \mu)/\partial N_T]_{N_T = 3/2} + O(\Delta N)^3,$$

where $O(\Delta N)^3$ indicates that terms of order $(\Delta N)^3$ are being ignored. A general property of any local density functional is that its derivative with respect to a particular occupation number equals the eigenvalue for that particular state.[31] Thus, to within the accuracy we have indicated, $\epsilon_1(Q_E, Q_b)$ and $\epsilon_2(Q_E, Q_b)$ are equal to the eigenvalue for the B_2 state in the gap, calculated self-consistently with that state occupied by $\frac{1}{2}$ and $\frac{3}{2}$ electrons, at a lattice distortion Q_E, Q_b.

At this point, we reason that ϵ_2, the energy to add a second electron, is greater than ϵ_1, the energy to add the first, because of e-e repulsion in the B_2 state. Since the spatial extent of that state is large compared to the amplitude of the distortion, we do not expect that energy difference to be strongly dependent on lattice distortion. Therefore a useful approximation is to assume that

$$\epsilon_2(Q_E, Q_b) = \epsilon_1(Q_E, Q_b) + U , \tag{2.13}$$

where U, the e-e repulsion term, is a constant, independent of Q_E, Q_b. We evaluate it by calculating the $N_T = \frac{3}{2}$ and $\frac{1}{2}$ eigenvalues at $Q_E = 0$, $Q_b = 0$.

We make use of this same approximation again (namely, that the *difference* in eigenvalues is independent of distortion) by studying the distortion dependence of the eigenvalue $\epsilon(Q_E, Q_b)$, not at an occupation of $\frac{1}{2}$ or $\frac{3}{2}$ electrons in the gap state, but at $N_T = 2$. The self-consistent potential is then that of the neutral state V^0, a situation which we can treat somewhat more reliably than we can a charged state, as explained in the next section.

The eigenvalue for the B_2 state, calculated in the V^0 potential at various distortions, shows only a small quadratic distortion dependence over the range we have studied. Therefore we shall use a linear expression of the form

$$\epsilon_1(Q_E, Q_b) = \epsilon_L - V_E Q_E - V_b Q_b . \tag{2.14}$$

Using the forms (2.11), (2.13), and (2.14), and choosing the overall reference energy zero at $\mathcal{E}_0$, the equations (2.10) become

$$E_{\rm tot} = \begin{cases} \tfrac{1}{2}(k_E Q_E^2 + k_b Q_b^2) - \beta Q_b , & N_T = 0 \qquad (2.15a) \\ \tfrac{1}{2}(k_E Q_E^2 + k_b Q_b^2) - \beta Q_b + \epsilon_L - V_E Q_E - V_b Q_b - \mu , & N_T = 1 \qquad (2.15b) \\ \tfrac{1}{2}(k_E^2 Q_E^2 + k_b Q_b^2) - \beta Q_b + 2(\epsilon_L - V_E Q_E - V_b Q_b - \mu) + U , & N_T = 2. \qquad (2.15c) \end{cases}$$

The forces $F_\alpha(N_T)$ for the three charge states are given by

$$F_b(0) = \beta, \quad F_b(1) = \beta + V_b, \quad F_b(2) = \beta + 2V_b \tag{2.16a}$$

for the breathing mode and by

$$F_E(0) = 0, \quad F_E(1) = V_E, \quad F_E(2) = 2V_E \tag{2.16b}$$

for the tetragonal mode. The spring constants $K_{\alpha\alpha}(N_T)$ are $K_b = k_b$ for the breathing mode and $K_E = k_E$ for the tetragonal, independent of the value of N_T. The N_T dependence of the spring constant, which we had anticipated in writing $K_{\alpha\beta}(\{n_i\})$ in (2.6), is important only to the extent that the eigenvalue exhibits a *quadratic* dependence on distortion.

Note the implications of this approach: *Changes* in the forces and spring constants caused by a change in charge state can be obtained by studying an eigenvalue for half-integral occupancy, *without* evaluating a complicated total energy. This follows directly from Eq. (2.10) and will be applicable to many charge-driven distortions.

To this point, the constants whose evaluation we have discussed, ϵ_L, U, V_E and V_b, are those associated with *adding additional* electrons. The constants whose evaluation we have not yet discussed, β, k_E and k_b, are associated with a *total* energy itself, that of the V^{++} state, and they have to be treated in a different manner.

Our numerical finding that we can ignore the N_T dependence of the spring constants means that we can evaluate k_E and k_b for the V^0 neutral state of the system. We do this, as explained in the next section, by considering a cluster of about 100 atoms surrounding the vacancy. The atoms in it interact according to a modified Keating model (quadratic terms, no linear terms) and are free to move, minimizing their energy, in response to a displacement imposed on the four atoms nearest the vacancy. The energy of the cluster depends quadratically on the amplitude of the displacements Q_E and Q_b imposed, and from that, we obtain k_E and k_b. The approach is identical in spirit, though not in detail, with that used by Larkins and Stoneham.[32] We do not try to calculate β, the force driving the outward breathing, but we estimate, in analogy with the (111) surface, the displacement to which it gives rise.

We may now calculate the occupancy levels resulting from the total energy functional in the form given by Eq. (2.15). For each value of N_T, the equilibrium value of distortion is determined by minimizing $E_{\rm tot}$. The resulting distortions are as follows:

$$Q_b^0 = \begin{cases} \beta/k_b, & N_T = 0 \\ (\beta + V_b)/k_b, & N_T = 1 \\ (\beta + 2V_b)/k_b, & N_T = 2 \end{cases} \qquad Q_E^0 = \begin{cases} 0, & N_T = 0 \quad (2.17a) \\ V_E/k_E, & N_T = 1 \quad (2.17b) \\ 2V_E/k_E, & N_T = 2 . \quad (2.17c) \end{cases}$$

The associated minimum values of $E_{\rm tot}$ are

$$E_{\rm tot} = \begin{cases} -\beta^2/2k_b, & N_T = 0 \\ -(\beta + V_b)^2/2k_b - V_E^2/2k_E + \epsilon_L - \mu, & N_T = 1 \\ -(\beta + 2V_b)^2/2k_b - (2V_E)^2/2k_E + 2\epsilon_L + U - 2\mu, & N_T = 2 . \end{cases}$$

Let us define the energies

$$E_{JT} \equiv V_E^2/2k_E , \quad (2.18a)$$

$$E_b \equiv \beta^2/2k_b , \quad (2.18b)$$

$$\hat{E}_{JT} = E_{JT} + V_b^2/2k_b , \quad (2.18c)$$

$$\hat{\epsilon}_L \equiv \epsilon_L - V_b Q_b^0(V^{++}) . \quad (2.18d)$$

Note that $\hat{\epsilon}_L$ is the eigenvalue for the B_2 level at occupancy $N_T = \frac{1}{2}$, no Jahn-Teller distortion, and the outward breathing distortion $Q_b^0(V^{++}) = \beta/k_b$. In terms of these energies, we reexpress the total energies as follows:

$$E_{\rm tot} = \begin{cases} -E_b, & N_T = 0 \quad (2.19a) \\ -E_b + \hat{\epsilon}_L - \hat{E}_{JT} - \mu, & N_T = 1 \quad (2.19b) \\ -E_b + 2\hat{\epsilon}_L + U - 4\hat{E}_{JT} - 2\mu, & N_T = 2 . \quad (2.19c) \end{cases}$$

The additive constant E_b plays no role in determining occupancy levels or activation energies. For our purposes, it is unobservable and can be ignored. There are thus three parameters of interest here, $\hat{\epsilon}_L$, U, and $\hat{E}_{JT}$.

We now consider which $E_{\rm tot}$ is lowest at a given μ. We can either plot $E_{\rm tot}$ vs μ, as in Fig. 5, or we can give a description in terms of levels. The use of "levels" is so standard in semiconductor physics that we think it desirable to try to describe our results in these terms even though certain extensions of the concept are required when the effective correlation energy of two electrons is negative.

The term level actually refers to part of the energy difference between *two* states of the system, as already suggested by our notation $E(n/n')$. In the normal situation, the "level" is $E(n+1/n)$ and its energy difference relative to the Fermi energy, $E(n+1/n) - \mu$, is the energy to take an electron from the Fermi reservoir and add it to the defect, raising the defect occupation from n to $n+1$.

Consider the possible energy differences between the three charge states of the vacancy, as

given in (2.19). Letting $E_{N_T}(\mu)$ denote the value of E_{tot}, we have levels defined by

$$E_1(\mu) - E_0(\mu) \equiv E(+/{+}{+}) - \mu ,$$

$$E_2(\mu) - E_1(\mu) \equiv E(0/+) - \mu ,$$

$$E_2(\mu) - E_0(\mu) \equiv 2[E(0/{+}{+}) - \mu] ,$$

so that

$$E(+/{+}{+}) = \hat{\epsilon}_L - \hat{E}_{JT} , \tag{2.20a}$$

$$E(0/+) = \hat{\epsilon}_L + U - 3\hat{E}_{JT} \tag{2.20b}$$

$$= E(+/{+}{+}) - \eta . \tag{2.20c}$$

We have defined

$$\eta \equiv 2\hat{E}_{JT} - U . \tag{2.21}$$

It is also clear that

$$E(0/{+}{+}) = \hat{\epsilon}_L + U/2 - 2\hat{E}_{JT} \tag{2.22a}$$

$$= E(+/{+}{+}) - \eta/2 \tag{2.22b}$$

$$= \tfrac{1}{2}[E(0/+) + E(+/{+}{+})] . \tag{2.22c}$$

Although our calculations are going to show that η is a positive constant for the silicon vacancy, it is useful to think of η as a variable, and plot the levels $E(n/n')$ vs η as in Fig. 6(a).

On the left-hand side of the diagram, η is negative, and [according to (2.20c)] we have the conventional situation of $E(0/+) > E(+/{+}{+})$. These levels then divide the Fermi energy (the vertical direction) into three regions where V^{++}, V^{+}, and V^{0} are the ground-state configuration of the vacancy.

On the right-hand side with η positive, there are only two stable configurations, ++ and 0, separated by the level $E(0/{+}{+})$ as shown. We call these levels where the occupation number of the ground state changes "occupancy levels." Since there are two such levels on the left-hand side and only one on the right, we have clearly lost some information from the diagram, namely, the description of how much the energy of V^{+} exceeds the ground-state energy, i.e., the V^{+} excitation energy. We can therefore introduce an excitation level, defined in such a way that its position relative to the Fermi level gives the energy by which the metastable state V^{+} lies above the ground state. The excitation level is $E(+/{+}{+})$ for $\mu \leq E(0/{+}{+})$ or $E(0/+)$ for $\mu \geq E(0/{+}{+})$. The two cases are shown by the dashed lines in Fig. 6.

With these definitions the "occupancy" and "excitation" levels provide the same information for $\eta > 0$ as the two normal (occupancy) levels for $\eta < 0$. They are to be used in the same way. Because $E(0/{+}{+})$ refers to a two-electron occupancy change, however, the thermal occupancy of V^{0} is determined by $2[E(0/{+}{+}) - \mu]/kT$ rather than by $(E - \mu)/kT$ as for a one-electron change.

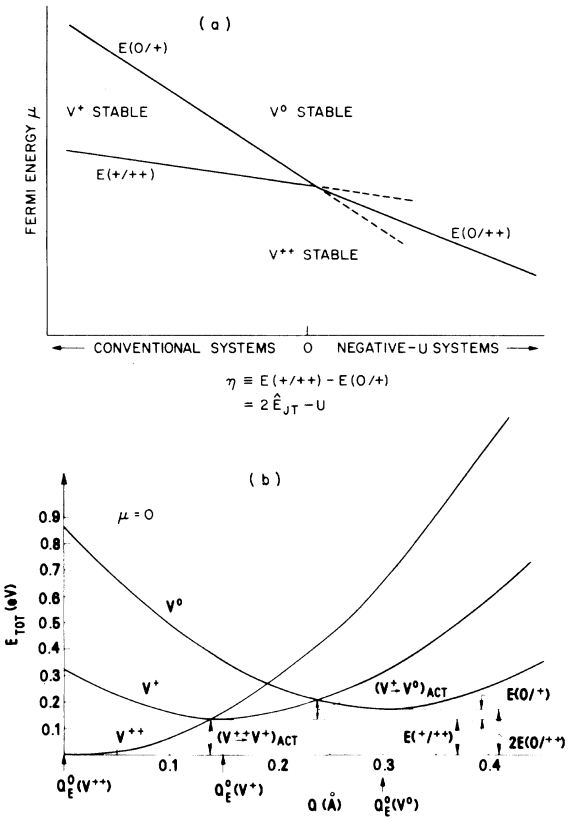

FIG. 6. (a) Occupancy level structure for conventional systems, $\eta < 0$, and for negative-effective-U systems, $\eta > 0$, shown schematically as a function of η. As η approaches zero from negative values, the region of Fermi energy for which V^{+} is the ground state shrinks and disappears. For $\eta > 0$, V^{+} is never the state of lowest energy and the only occupancy level is that between V^{++} and V^{0}. The excitation energy by which V^{+} fails to be the ground state, is given, for $\eta > 0$, by $E(+/{+}{+}) - \mu$ if $\mu \leq E(0/{+}{+})$ and by $\mu - E(0/+)$ if $\mu \geq E(0/{+}{+})$. (b) Configuration coordinate diagram for transfer of electrons between the vacancy and valence band. The levels $E(0/+)$ and $E(+/{+}{+})$ are indicated as the energy required to add an electron to the vacancy from the top of the valence band. If $E(0/+) > E(+/{+}{+})$, then both of these are occupancy levels and electrons can be added to the vacancy one at a time. If $E(0/+) < E(+/{+}{+})$, the electrons must be added *two* at a time, at an energy cost per electron of $E(0/{+}{+}) = \frac{1}{2}[E(0/+) + E(+/{+}{+})]$. In this case, $E(0/{+}{+})$ is the only occupancy level. Activation energies for transitions $N_T \rightarrow N_T + 1$ are indicated. For transitions $N_T \rightarrow N_T - 1$, the electron can return to the valence band only if there is a preexisting hole there, so activation energies for such transitions contain an additive constant μ (not shown in the figure), the activation energy of hole creation.

For transfer of electrons between the vacancy and the valence band, activation energies, optical energies and occupancy levels can be read from a

conventional configuration coordinate diagram, constructed in the usual manner with $\mu = 0$. Figure 6(b) is such a diagram, constructed using values close to those we have calculated as described in the next section, and denoted by Calc. 2, a set of theoretically predicted values displayed in Table VIII, in Sec. V. The levels $E(+/++)$ and $E(0/+)$ are indicated. Each level is the energy required to add an electron to the vacancy (from a reservoir coincident with the top of the valence band). If $E(0/+) > E(+/++)$, then both of these are occupancy levels, and we can add electrons to the vacancy one at a time. If $E(0/+) < E(+/++)$, then we must add electrons *two* at a time, and the energy per electron to add them is $E(0/++) = \frac{1}{2}[E(0/+) + E(+/++)]$. In this case, $E(0/++)$ is the only occupancy level.

Activation energies for transitions $N_T \rightarrow N_T \pm 1$ can be obtained from this diagram by noting the energy difference between the minimum of the N_T curve and its crossing with the $N_T \pm 1$ curve. For the transition $N_T \rightarrow N_T + 1$, the activation energy equals this energy difference. For the transition $N_T \rightarrow N_T - 1$, the electron cannot return to the valence band unless there is a preexisting hole there, so the activation energy for the transition $N_T \rightarrow N_T - 1$ is the sum of that energy difference and μ, the activation energy for hole creation.[19]

III. COMPUTATIONAL METHODS AND RESULTS

A. Self-consistent Green's-function formalism

Recall that the electronic structure of an isolated defect is described by

$$[H_0 + U(r)]\psi(r) = E\psi(r) , \tag{3.1}$$

where H_0 is the Hamiltonian of the perfect crystal and where $U(r)$ is the self-consistent defect potential which tends to zero at large distance. A detailed description of the techniques used to solve Eq. (3.1) is given in a previous paper.[9] Here we briefly recapitulate some key steps in our method of solution.

In the first step, the wave function $\psi(r)$ is approximated as a linear combination of atomic orbitals (LCAO)

$$\psi(r) \approx \sum_i c_i \phi_i(r) , \tag{3.2}$$

with orbitals centered on actual atomic sites. The number and type of orbitals at each site is chosen with the intent that

$$U(r)\psi(r) \approx U(r) \sum_i c_i \phi_i(r) ; \tag{3.3}$$

i.e., we choose a "chemical" LCAO-like set of local orbitals $\phi_i(r)$ ("inner set") that approximates the solutions of Eq. (3.1) within the range of $U(r)$.

We consider Eq. (3.1) in the Green's-function form:

$$\psi(r) = \psi_0(r) + \int G_E^0(r, r')U(r')\psi(r')\,dr' . \tag{3.4}$$

We define a second set of localized functions $\Phi_m(r)$ ("outer set"), distinct from and independent of the LCAO basis $\phi_i(r)$, and use it to express the eigenfunctions of H_0:

$$\psi_n(k, r) = \sum_m B_m(n, k)\Phi_m(r) . \tag{3.5}$$

In contrast to our earlier procedure,[9] we now use the set $\Phi_m(r)$ *directly* in a local-orbital band-structure calculation to obtain eigenvalues $E_n(k)$ and eigenfunctions $\psi_n(k, r)$. These in turn are used to construct the perfect-crystal Green's function (BZ denotes Brillouin zone)

$$G_E^0(r, r') = \sum_n \int_{BZ} dk \frac{\psi_n(k, r)\psi_n^*(k, r')}{E - E_n(k)} , \tag{3.6}$$

which then takes the form

$$G_E^0(r, r') = \sum_{mm'} \Phi_m(r) G^0_{mm'}(E) \Phi_{m'}(r') . \tag{3.7}$$

Formally, an infinite matrix $G^0_{mm'}$ is needed. Practically, the size of this matrix is determined by two considerations:

(a) the set $\Phi_m(r)$ should accurately reproduce the host band structure and

(b) the Green's function [Eq. (3.7)] should be complete enough to respond accurately to the shape of $U(r')\psi(r')$ in Eq. (3.4).

We shall return to these completeness criteria in more detail below.

Once $G^0_{mm'}(E)$ is known, Eq. (3.4) can be written as a matrix equation which is either homogeneous for energies E outside the spectrum of H_0, i.e., for bound states, or inhomogeneous for scattering states [see also Eqs. (22)–(28) in Ref. 9]. Furthermore, using the matrix representation of Eq. (3.4), we can evaluate the total charge density:

$$\rho(r) = \sum_{mm'} \Phi_m(r)\rho_{mm'}\Phi_{m'}(r). \tag{3.8}$$

The original form of $\rho_{mm'}$ given in Ref. 9 can be simplified using some matrix identities as outlined in Ref. 33. With knowledge of the density $\rho(r)$ for both the perfect and the defect crystal, the defect (difference) potential $U(r)$ can be evaluated. We use a local approximation for exchange and correlation (see Ref. 9), identical to the one used in the perfect-crystal band-structure calculation. The calculation of $\rho(r)$ is then iterated until self-consistency in $U(r)$ is reached.

B. Convergence and completeness of Green's function

The solution of Eq. (3.4) requires that each state ψ_0 of the unperturbed crystal and each state ψ of the perturbed crystal be expressible in terms of outer-set orbitals $\Phi_n(r-R)$:

$$\psi_j(r)=\sum_n\sum_R c_j(n,R)\Phi_n(r-R)\,, \tag{3.9}$$

where the sum n runs over orbitals in the unit cell and R runs over all cells. While in principle both sums are infinite, in practice a finite number of terms is retained. In this context we distinguish an energy or angular momentum cutoff for the orbital index n and a real-space cutoff for the lattice site index R, both of which characterize the completeness of $G_E^0(r,r')$ in Eqs. (3.6) and (3.7).

To test for completeness we define

$$\Delta(r,r')=\frac{1}{\pi}\int dE\,\mathrm{Im}G_E^0(r,r')\,. \tag{3.10}$$

For any exact Green's function, $\Delta(r,r')$ defined this way is the delta function $\delta(r-r')$. The completeness test is a study of how well $\Delta(r,r')$ acts like $\delta(r-r')$ in the function space of interest. $\Delta(r,r')$ can also be written as

$$\Delta(r,r')=\sum_m\sum_{m'}\Phi_m(r)S^{-1}_{mm'}\Phi_{m'}(r')\,, \tag{3.11}$$

where m stands for n and R, and where $S^{-1}_{mm'}$ is the inverse of the *infinite* overlap matrix

$$S_{mm'}=\int\Phi^*_m(r)\Phi_{m'}(r)\,dr\,. \tag{3.12}$$

To actually evaluate $S^{-1}_{mm'}$ we make use of periodicity by first calculating $S_{mm'}$ in Fourier space, which allows us to invert a finite matrix. After transforming back into real space, only a finite number of elements of $S^{-1}_{mm'}$ are retained, namely, those we intend to use in $G^0_{mm'}(E)$.

To test the completeness of $G_E^0(r,r')$ or $\Delta(r,r')$ we calculate

$$\xi=\frac{\int|f(r)|^2dr-\int\int f^*(r)\Delta(r,r')f(r')dr\,dr'}{\int|f(r)|^2dr}\,, \tag{3.13}$$

a quantity which vanishes if Δ is indeed a δ function. From Eq. (3.4) it follows that completeness is in our case needed only for functions of the type

$$f(r)=U(r)\psi(r)\,. \tag{3.14}$$

This implies that the spatial extent of $U(r)$ determines the lattice space cutoff R in Eqs. (3.9) and (3.11) and that the specific form of $f(r)$, i.e., short-range oscillations, etc., determine the energy cutoff. In practice, this test is carried out for each inner-set orbital separately replacing ψ in (3.14), to insure that G^0 is complete for every function of the type $f=U\psi$.

Figure 7 illustrates the spatial behavior of $U(r)\psi(r)$ for the cases of an unrelaxed (top) and a relaxed (bottom) vacancy. The unrelaxed case whose results were presented in Ref. 9 is characterized by two simplifying features: The overlap $U(r)\psi(r)$ is small, and the defect "dangling" bonds expressed by $\phi_j(r)$ are centered at the same sites as are the outer-set perfect-crystal orbitals $\Phi_m(r)$. Thus satisfactory completeness can be reached with relatively few orbitals of rather long range (see Table I). The situation changes drastically when an atom is displaced. Figure 7 (bottom) schematically shows the resulting potential perturbation which corresponds to the removal of an atom at the "old" lattice site and the insertion of an atom at the "new" lattice site. The overlap with the outer-set orbitals Φ_m is large. Completeness requires the reproduction of $U\phi_j$ by the set Φ_m.

Differently formulated, the electronic wave function $\psi=\sum_j c_j\phi_j$, which moves with the displaced atom, has to be expanded in terms of outer-set functions Φ_m which are still centered at the original lattice sites. Simple linear-expansion arguments show that a displaced s-like Gaussian function in first order is described by an s-like and three p-like Gaussian functions, while displaced p-like Gaussian functions require the complete set of p- and d-like functions augmented by the symmetric function r^2. The linear description of dis-

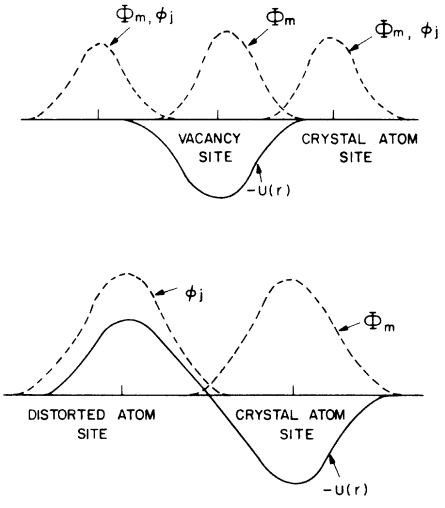

FIG. 7. Schematic view of defect potential $U(r)$, outer-set orbitals $\Phi_m(r)$, and inner-set orbitals $\phi_j(r)$. Top figure illustrates the situation of the unrelaxed vacancy (simple removal of an atom), bottom figure the situation created by the displacement of one atom only.

TABLE I. Values of the Green's-function completeness test [Eq. (3.13)] for different orbital sets for undistorted ($\Delta Q = 0$) and distorted ($\Delta Q_b = 0.2$ Å) vacancy geometries. Seventeen lattice sites were included in each test.

	$U = U_{\text{vacancy}}$ $\Delta Q_b = 0$	$U = U_{\text{vacancy}} + U_{\text{distortion}}$ $\Delta Q_b = 0.2$ Å
$\alpha = 0.20$ 10 orbitals per atom	$\xi \approx 0.1$–0.2	$\xi \approx 0.6$–1.0
$\alpha = 0.20, 0.60$ 20 orbitals per atom	$\xi \approx 0.1$	$\xi \approx 0.1$–0.2

placed sp^3 hybrids thus requires at least ten basic functions Φ_m. Furthermore, the specific short-range fluctuations of U representing atom displacement require inclusion of additional short-range orbitals supplementing the usual set of orbitals needed for a band-structure calculation. Table I shows several ξ values [Eq. (3.13)] characteristic for chosen orbital basis sets. It is evident from these tests that the use of only long-range ($\alpha = 0.2$) orbitals is sufficient to describe the undistorted vacancy (Ref. 9), while the distorted vacancy requires additional s-, p-, and d-like short-range orbitals.

To test the lattice space cutoff for the R sum in Eq. (3.9), we have experimented with 17, 35, and 41 lattice sites. Satisfactory convergence was found with 17 sites for distortions restricted to the four nearest-neighbor atoms of the vacancy.

C. Tight-binding band structure of silicon

The representation of silicon pseudopotential wave functions by Bloch sums of s, p, and d Gaussians centered on atom sites has been discussed in full length by Kane.[34] It has been shown that reasonable valence band structures can be obtained with single s and p Gaussians of one decay constant α. Conduction bands, however, and also to some extent the position of the top Γ_{25}' valence band improve if d-like Gaussians are added.

Here we have carried out test calculations using ten Gaussians of s-, p-, d-type and an additional r^2-type. We used one decay constant $\alpha = 0.20$ a.u. and two constants $\alpha = 0.20, 0.60$ a.u., respectively. The potential was the self-consistently screened "soft-core" pseudopotential used in Ref. 2. This potential has also been used in our previous calculation of the unrelaxed vacancy.[9] In Table II we compare energy eigenvalues of our local-orbital calculations with independent linearized-augmented-plane-wave (LAPW) calculations by Hamann,[35] based on the same ionic pseudopotential. The agreement is generally within a few tenths of an eV and similarly good for one and two decay constants. Note, however, that the inclusion of short-range Gaussians proved to be necessary for Green's function completeness. As seen from Table II, the indirect gap of ~0.8 eV is about 0.3 eV too small as compared to experiment. We corrected this by slightly adjusting the crystal potential while minimizing other changes occurring throughout the band structure. In Fig. 8 we present the calculated charge profiles plotted along an axis connecting two silicon atoms. The comparison is made between Hamann's LAPW results[35] and our ($\alpha = 0.2$) and ($\alpha = 0.2, 0.6$) local-orbital results. The charge suppression at the atom sites is improved by adding short-range ($\alpha = 0.6$) orbitals and is brought into approximate coincidence with the LAPW results. Figure 9

TABLE II. Comparison of energy eigenvalues obtained by LAPW scheme (using the Wigner interpolation formula for exchange) with local-orbital calculations (using $X\alpha$ for exchange). All calculations are based on a soft-core ionic pseudopotential (Ref. 2). For later use the potential was adjusted to yield the correct experimental gap value of 1.1 eV.

		10 Gaussians	
	LAPW Ref. 35	One decay 0.2 a.u.	Two decays 0.2, 0.6 a.u.
Γ_1	−12.79	−12.89	−12.88
Γ_{25}'	0	0	0
Γ_{15}'	2.80	2.84	2.98
Γ_2'	3.37	3.32	3.23
X_1	−8.49	−8.63	−8.62
X_4	−3.16	−3.18	−3.13
$X_1(E_g)$	0.64	0.78	0.79
X_4	11.94	11.87	12.13
Γ_2'	−10.39	−10.55	−10.54
L_1	−7.57	−7.60	−7.58
L_3'	−1.35	−1.34	−1.30
L_1	1.49	1.52	1.50
L_3	3.63	3.70	3.79

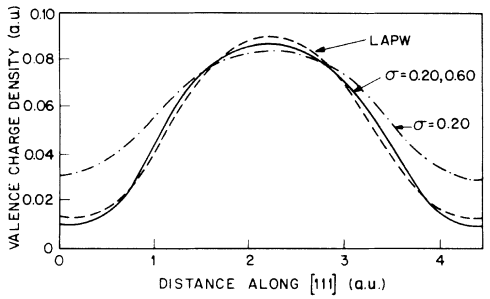

FIG. 8. Valence charge density profile plotted along the bond connecting two Si atoms. Dashed line, LAPW calculation by Hamann (Ref. 35); dotted line, $\alpha = 0.20$ local-orbital calculation; full line, $\alpha = 0.20$, $\alpha = 0.60$ local-orbital calculation. All calculations are based on the same core pseudopotential (Ref. 2).

shows the density of states for silicon derived from the ($\alpha = 0.2, 0.6$) 40 orbitals per cell basis set by integrating over 203 k points in the irreducible part of the Brillouin zone. The same integration scheme, i.e., the Gilat-Raubenheimer[36] technique, is used to calculate the matrix elements of the imaginary part of the Green's function. While the band-structure integration is carried out over 203 k points, the wave functions are interpolated from 70 k points.

D. General results of Green's-function calculation

The results for the neutral Si lattice vacancy in its hypothetical "ideal" unrelaxed structure of T_d symmetry are described in detail in Ref. 9. Briefly, the removal of an atom creates four broken bonds. Four valence electrons which previously participated in the bonds are removed together with the neutralizing Si^{4+} core. The atoms around the defect now experience a different (repulsive) potential and localized resonances and bound states occur. Describing these states in the T_d point group of the undistorted vacancy, one finds a strong A_1-type resonance located in the "pseudogap" at −8.5 eV followed by another A_1 resonance around −1.1 eV. These resonances correspond to symmetric (A_1) combinations of mostly s-like orbitals (−8.5 eV) and mostly p-like orbitals (−1.1 eV) centered at the four nearest-neighbor atoms surrounding the vacancy. In addition, one finds a T_2 bound state at 0.7 eV in the forbidden gap. This degenerate state can accommodate up to six electrons; its numerical occupation defines the various charge states of the vacancy.

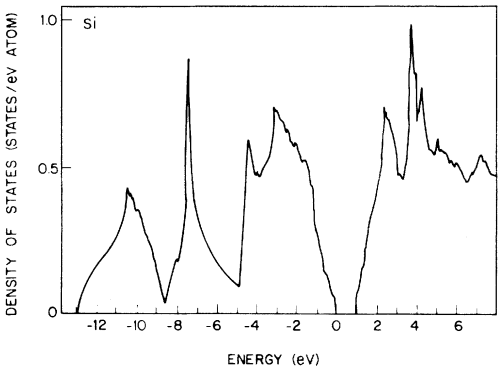

FIG. 9. Valence density of states for Si calculated with a local-orbital basis of $\alpha = 0.20$, $\alpha = 0.60$ (40 orbitals per cell) as used in the evaluation of the Green's function.

Incomplete occupancy creates a situation which is unstable with respect to static symmetry-lowering Jahn-Teller distortions as we described in Secs. I and II. The charge-state-dependent distortions are summarized in Table III. We have added the undistorted V^{++} state to the four charge states proposed by Watkins. Higher negative charge states than V^{--} have not been observed and can probably be ruled out because of excessive Coulomb repulsion and insufficient Jahn-Teller lowering, as we approach the fully occupied multiplet which has no tendency to distort in a symmetry-lowering fashion.

As we mentioned in Sec. II, symmetry-lowering distortions are in general accompanied by symmetric breathing-type distortions whose magnitude may in general depend on the charge state. We have thus investigated by our Green's-function technique symmetric breathing-type and asymmetric tetragonal "pairing-antipairing"-type distortions. For the calculation of energy levels and deformation potentials, we have restricted the distortions to the four nearest neighbors of the vacancy, as in Fig. 10, an approximation which is motivated by the actual localization of the bound-state wave function. While the bound-state deformation potential is directly linked to the localized

TABLE III. List of various charge states of vacancy. The level splittings, degeneracies, and occupancies are shown. The point group symmetries are experimentally determined (Ref. 6). V^{++}, V^{+}, and V^{0} are found in p-type silicon, while V^{-} and V^{--} are found in n-type silicon.

Charge state	Number of electrons in bound state	Crystal symmetry
V^{++}	0	T_d
V^{+}	1	D_{2d}
V^{0}	2	D_{2d}
V^{-}	3	C_{2v}
V^{--}	4	C_{2v}

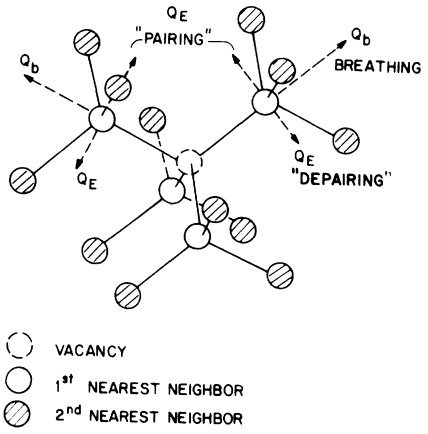

FIG. 10. Perspective view of the diamond structure with a single vacancy. Symmetric breathing-type (Q_b) and symmetry-lowering–pairing-depairing–type (Q_E) distortion directions of nearest-neighbor atoms are indicated.

nature of the bound-state wave function, any quantities involving electron states other than the bound state as, e.g., the *driving force* for breathing-type distortions or the lattice elastic restoring forces, are strongly dependent on the *long-range* nature of distortions (see Sec. III E below).

The calculation of charged defects with our Green's-function scattering approach deserves more comment. The technique of solving Eq. (3.4) with a finite basis of localized functions requires the defect potential $U(r)$ to be of short range. The long-range Coulomb potential of the form $Z_{eff}/\epsilon r$ arising from charging the defect is conveniently treated perturbatively. For this purpose we define an outer potential radius $R_c \approx 4.0$ Å within the short-range defect potential decays to its "dielectric" long-range value $Z_{eff}/\epsilon r$. The value $U(R_c)$ is set equal to $Z_{eff}/\epsilon R_c$ and $U(r>R_c)$ is set equal to zero during the self-consistency iteration procedure. We then calculate the energy contribution arising from the tails of the bound-state wave function in the screened Coulomb potential perturbatively. Since these tails are not explicitly known, we model them by an exponentially decaying (p-like) function, $r\exp(-\alpha r)$, and obtain for the correction (in a.u.),

$$\Delta\epsilon(Z_{eff}) = \frac{Z_{eff}\Theta}{\epsilon R_c}\left(\frac{1+\chi}{1+2\chi+2\chi^2}\right), \tag{3.15}$$

with $\chi = 1/(\alpha R_c)$, and $\Theta = 0.4$, the fraction of the vacancy wave function lying beyond R_c. A simple upper bound to this correction is obtained for $\chi = 0$, as though all tail charge is concentrated at R_c. Evaluating Eq. (3.15) we obtain $\Delta\epsilon(Z_{eff}=1) \approx -0.15$ eV for the silicon vacancy. Together with the results of the Green's-function treatment we obtain $\epsilon(n=\frac{1}{2}) = 0.32$ eV, $\epsilon(n=\frac{3}{2}) = 0.57$ eV, yielding $\epsilon_L = 0.32$ eV and $U = 0.25$ eV. The Green's-function calculations were done for full T_d symmetry, occupying each of the three T_2 orbitals equally. Even at zero distortion this is only an approximation to the situation in which one of the three T_2 states (B_2) is preferentially occupied and self-consistently screened. We estimate that it will not affect U, and its effect on ϵ_L is likely to be less than 0.1 eV.

The same value of U, 0.25 eV, can be extracted from an interesting recent calculation of the silicon vacancy by Kauffer, Pecheur, and Gerl.[37] They report that the ϵ_{T_2} eigenvalue is 0.14, 0.37, and 0.64 eV for the unrelaxed vacancy states V^+, V^0, and V^-. The differences in these eigenvalues are 0.23 eV and 0.27 eV. We regard this agreement as a useful confirmation of the reasoning which they used to arrive at their parametrized self-consistency.

We now discuss the results for various mode distortions. As reported in Ref. 9, the change in the density of electronic states caused by a localized defect is given by

$$\Delta n(E) = -\frac{1}{\pi}\frac{d\Phi(E)}{dE}, \tag{3.16}$$

where $\Phi(E)$ is the scattering phase shift which is obtained from the determinant of the Green's-function matrix equation [Eq. (37) in Ref. 9]. In Fig. 11 we display the A_1 scattering phase shift (top) for V^0 as a function of energy for several breathing distortions. While the A_1 s-like resonance around −8.5 eV becomes only slightly weakened, the p-like resonance below the top of the valence band gets

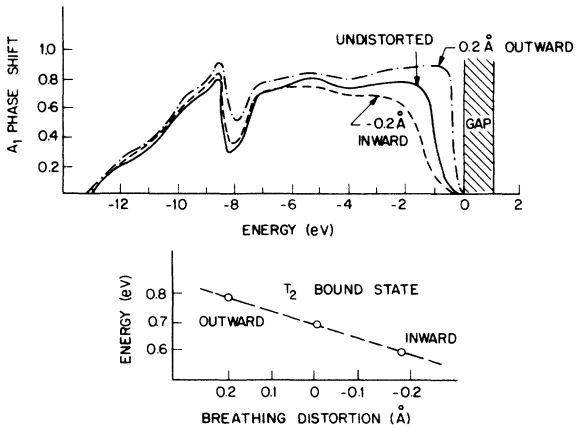

FIG. 11. Calculated A_1 phase shifts and T_2 bound-state energies induced by neutral vacancy for different breathing-type distortions.

shifted to higher energies and sharpens strongly upon outward breathing (away from the vacancy). It becomes a genuine A_1 bound state in the gap for distortions $Q_b \gtrsim 0.3$ Å.

Breathing away from the vacancy renders the back-bond geometry more graphitelike, i.e., sp^2 like, which increases the p-orbital content in the dangling bond directed towards the vacancy. Increased p-orbital content increases the energy of the upper A_1 resonance and produces a fading of the lower (s-like) resonance.

Conversely, inward breathing (towards the vacancy) sharpens the s-like resonance and weakens the p-like resonance along with a downward shift in energy. This is caused by the increase in p^3-like bonding of the back bonds which renders the dangling bond more s-like. The T_2 bound state which corresponds to a "p-like" combination of dangling-bond states responds in a similar way to breathing distortions as does the upper A_1 resonance. Its energy position as a function of distortion is shown in Fig. 11 (bottom). For it we derive a rather small linear deformation potential of $V_b \equiv \Delta\epsilon/\Delta Q_b = 0.5$ eV/Å. This value agrees reasonably well with results by Lipari, Bernholc, and Pantelides[38] and also results by Jaros, Rodrigues, and Brand.[39] The above description is very reminiscent of the arguments proposed by Haneman[40] to rationalize the up and down motion of surface atoms on silicon (111). Moreover, the calculated deformation potential value V_b qualitatively agrees with theoretical findings on this surface.[41] We shall, therefore, use this analogy to the (111) surface to estimate approximate breathing distortion values.

We now turn to the calculation of pure tetragonal, symmetry-breaking distortions leading to a lowering of symmetry from T_d to D_{2d}.

Figure 12 (top) illustrates the behavior of the $A_1(T_d)$ and $A_1(D_{2d})$ phase shifts of the neutral vacancy V^0 upon pairing and depairing E-type distortions. As can be seen, the effects on position and strength of resonances are minor. The splitting of the T_2 bound state into the (B_2,E) pair is shown in the bottom panel. Pairing lowers the onefold B_2 level while depairing lowers the twofold E level. Also shown (as dotted line) is the deformation dependence of the center of gravity of the (B_2,E) pair. Its slight dependence on the amount or sense of the distortion checks the assumption of linearity for small atomic motions of pure symmetry-lowering character. Again approximating the (B_2,E) pair for zero distortion by the degenerate T_2 level, we derive a linear deformation potential for tetragonal distortions of V_0 to be $V_E \equiv d\epsilon/dQ_E = -2.25 \pm 0.2$ eV/Å for the B_2 level. We find that this calculated value depends

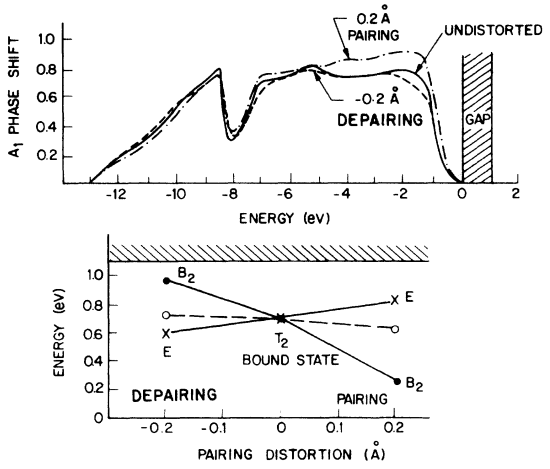

FIG. 12. Calculated A_1 phase shifts and T_2-(B_2,E) bound-state energies induced by the neutral vacancy for different symmetry-lowering pairing, depairing-type distortions.

quantitatively on the "completness" of the Green's function used to evaluate it. In particular, neglect of the short-range orbitals in G^0 results in a decrease of $d\epsilon/dQ_E$ by about 30%. In this context we note some discrepancies between our results and other calculations. Lipari *et al.*[38] have reported $d\epsilon/dQ_E$ values 40% smaller than ours while Jaros *et al.*[39] have reported $d\epsilon/dQ_E$ values considerably smaller than these.

In the following we shall present results of defect potentials and charge densities which were calculated for a combined breathing–symmetry-lowering distortion. The chosen distortion of +0.2-Å outward breathing is representative. The +0.3-Å atom pairing distortion was derived from total energy minimization. The phase shifts induced by this combined distortion are shown in Fig. 13 where they are compared to the results of the ideal undistorted vacancy. The distortion splits the T_2 level by 0.68 eV, i.e., lowers B_2 by 0.45 eV from its undistorted value.

In Fig. 14 we present contour plots of the defect potential $U(r)$, displayed in a (110) plane containing the vacancy and nearest-neighbor atoms. The potential for the unrelaxed vacancy (bottom) is nearly spherical and roughly corresponds to a "negative" screened Si-atom pseudopotential (see also Fig. 9 in Ref. 9). It is of short range, decaying within nearest-neighbor distance. The dipolelike short-range potential fluctuations induced by atom relaxation are clearly visible in the top figure. They can be thought of as originating from removing atoms at the perfect-crystal lattice sites and inserting atoms at the relaxed lattice sites.

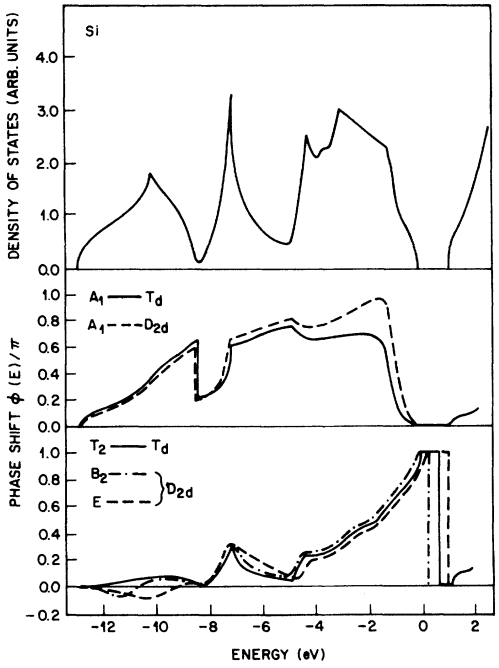

FIG. 13. Bulk density of states of Si and the A_1 and $T_2(B_2, E)$ phase shifts induced by neutral isolated vacancy. Full line, undistorted vacancy; broken lines, distorted vacancy. The distortion is a superposition of 0.2 Å outward breathing and 0.3 Å atom pairing.

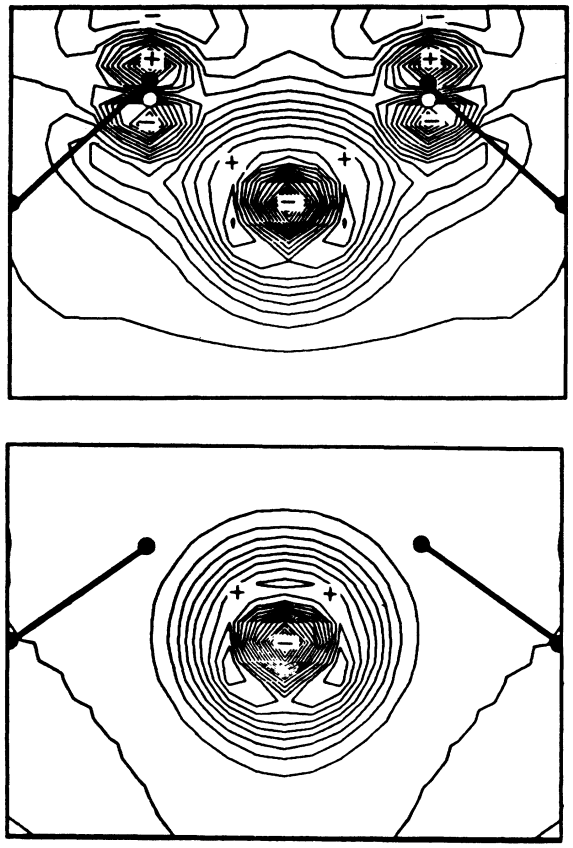

FIG. 14. Contour plot in a (110) plane of total self-consistent defect potential $U(r)$ for the undistorted vacancy (bottom) and the distorted vacancy (top).

For distortions ≈0.3 Å the dipole potential amplitudes are roughly comparable to the vacancy potential itself.

Figure 15 shows contour plots of pseudopotential total valence charge densities around the undistorted vacancy (bottom) and distorted vacancy (top). The displays have been prepared by expanding the Bloch waves in an infinite set of orbitals Φ_m [Eq. (3.5)] and retaining only those orbitals (outer set) which influence the density in the neighborhood of the defect. No information is lost, but the display lacks the translational periodicity one would expect to see (see also Ref. 9). The plots clearly show the disappearance of bonding charge with the removal of an atom. Moreover, little change occurs upon atom relaxation, essentially following our intuitive expectations derived from the Born-Oppenheimer picture. Some quantitative changes, however, are noticeable. The displayed back bonds lose charge and strength upon this particular relaxation since they are stretched. Conversely, back bonds (not shown here) which are compressed show increased bonding charges. The amount of distortion appearing here represents a slight overestimate. Since our Green's-function calculations were intended to yield bound-state eigenvalues and their deformation dependence, it was a reasonable approximation for us to ignore atom relaxation beyond nearest neighbors.

We notice that the charge density at atom centers and bond maxima follows the motion of the atoms. This is of particular importance, since the charge density is derived from outer set orbitals Φ_m which are centered at perfect-crystal lattice sites. Figure 15 thus is an illustration of the adequacy of the Green's function [Eq. (3.7)] in regard to completeness. The effect is even more emphasized in Fig. 16 where we present difference charge densities $\Delta\rho = \rho_{\text{perfect crystal}} - \rho_{\text{defect crystal}}$ for both, unrelaxed (bottom) and relaxed (top) vacancy configurations. The top figure clearly shows the decrease in back bonding charge and the shift of charge in the immediate neighborhood of the relaxed atoms. The charge perturbation is consequently more spread out than for the unrelaxed case.

We next consider the charge distribution of the defect bound state. For the unrelaxed case each of the three partner functions of the T_2 bound state

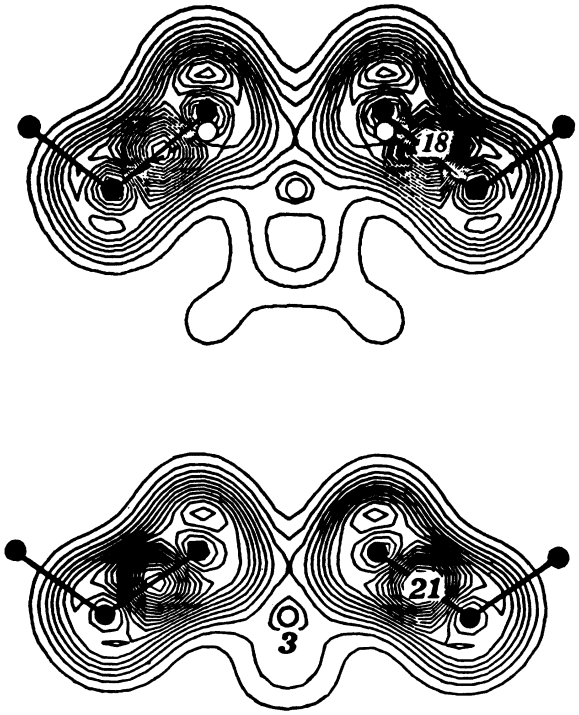

FIG. 15. Total valence charge distribution around neutral vacancy for undistorted (bottom) and distorted (top) geometries. The values are given in units of electrons per Si bulk unit cell.

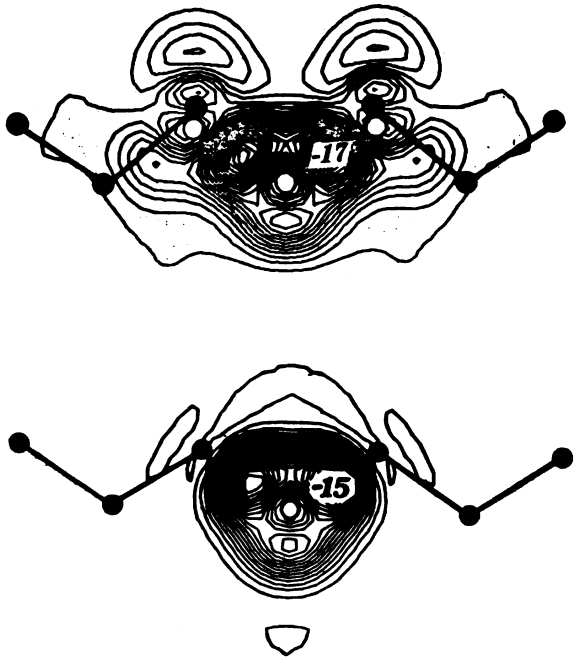

FIG. 16. Change in total valence charge density induced by neutral vacancy for undistorted (bottom) and distorted (top) geometries. Units as in Fig. 14.

was occupied to equal amounts to produce a charge distribution invariant under T_d. Its real-space distribution is shown in Fig. 17 (bottom). It exhibits dangling-bond character with the individual dangling bonds located at the four nearest-neighbor atoms. Upon symmetry-breaking relaxation ($T_d \rightarrow D_{2d}$) it splits into an occupied ($2e^-$ for V^0) B_2 level (top panel) and an empty E level (middle panel). The wave-function symmetry is such as to exhibit no node for B_2 between the paired atoms (shown) but nodes between the unpaired atoms (not shown). Conversely for the E states, nodes exist between the paired atoms (shown) and no nodes between the unpaired atoms (not shown). The difference in self-consistent potential along lines connecting paired atoms (more attractive) and unpaired atoms (more repulsive) is the origin of the B_2-E splitting. This difference also modifies, though only slightly, the resulting wave-function shapes. The more attractive potential

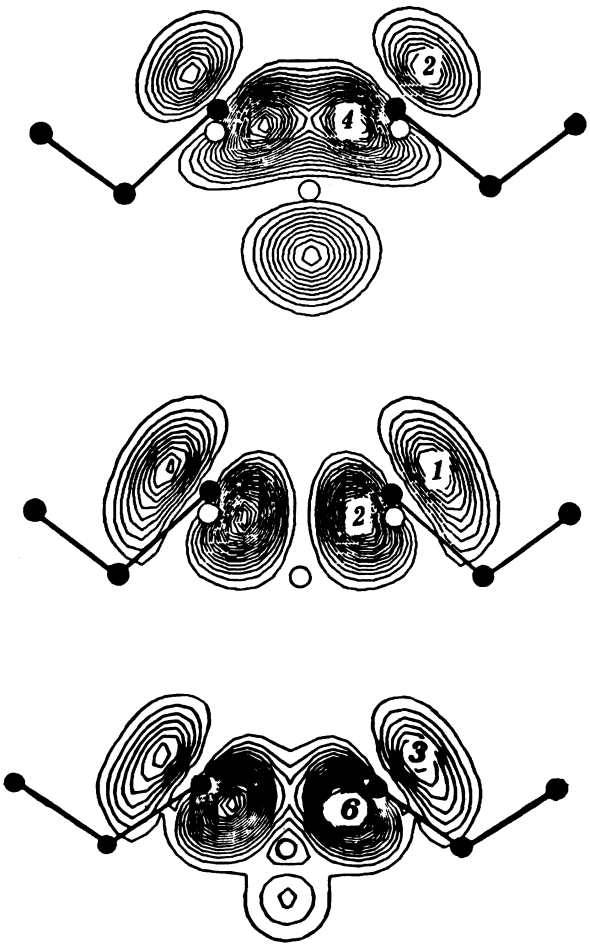

FIG. 17. Charge density of T_2 bound state (bottom) for undistorted vacancy and E (middle) and B_2 (top) bound states for distorted vacancy.

between paired atoms "pulls in" some charge resulting in a small tilt (~10°) of the dangling bonds *towards* each other for the lower B_2 state. The opposite effect (i.e., a tilt away from each other) is found for the higher E state. On the basis of these small wave-function changes the (B_2-E) pair may also be viewed as bonding and antibonding partners of a new weak "bond" between paired second-nearest-neighbor atoms. The pairing may be regarded to be the tendency to rebond threefold coordinated Si atoms around the vacancy in order to "heal" the defect.

The findings described above and illustrated in Fig. 17 seem to contradict findings from extended Hückel calculations on diamond vacancies by Messmer and Watkins.[42] They reported tilting of the occupied dangling-bond axes towards the axis of distortion, i.e., tilting of dangling bonds away from each other upon pairing. They furthermore reported that this tilt direction was necessary to account for experimental hyperfine tensor symmetries observed on silicon vacancies.[43] Further investigations are planned to resolve this discrepancy.

E. Model calculations of elastic restoring forces

As mentioned in Sec. II, the model we use here is based on the valence-force-field method of Keating[10] in which *all* interatomic forces are resolved into nearest-neighbor bond-stretching and bond-bending forces. Following Martin's[44] notation the model total energy is given as

$$E_{\text{tot}} = \tfrac{1}{2}\alpha\left(\frac{3}{4d^2}\right)\sum_{i=1}^{4}[\Delta(\vec{r}_i^{\,1}\cdot\vec{r}_i^{\,1})]^2 + \tfrac{1}{2}\beta\left(\frac{3}{4d^2}\right)\sum_{s=1}^{2}\sum_{i,j>i}[\Delta(\vec{r}_i^{\,s}\cdot\vec{r}_j^{\,s})]^2 , \qquad (3.17)$$

where the bonds about each atom in the unit cell are denoted by $i,j=1,\ldots,4$, d is the equilibrium bond length and $\Delta(\vec{r}_i^{\,s}\cdot\vec{r}_j^{\,s})$ are the scalar variations of bond vectors about atoms. Equation (3.17) yields terms up to quartic in the displacements. We retain only the quadratic terms.

The Keating parameters α and β describe pure bond stretching (α) and bond bending with some stretching component (β). The values of α and β published by Martin[44] (Table IV) are fitted to reproduce the experimental long-wavelength elastic constants c_{11}, c_{44}, and c_{12} with an internal consistency of better than 1%. However, the simplicity of the model causes inaccuracies for other distortion modes. Deviations of about 5% are reported by Martin[44] for the $k=0$ transverse optic mode frequency TO(Γ). For finite wave vectors, large deviations occur. In particular the "flatness" of the TA branch towards the Brillouin-zone boundary is very inadequately reproduced[45] (see Table IV).

In this work we have to describe localized distortions around defects. Fourier analysis would involve distortions with wave vectors distributed throughout the Brillouin zone, with considerable weighting for larger wave vectors. To obtain an optimum description of these distortions within the simple Keating model, we modify Martin's values of α and β to reproduce finite wave vector modes throughout large portions of the Brillouin zone. The modifications are restricted to retain the correct compressibility $K=3/(c_{11}+2c_{12})$. The resulting values α and β are listed in Table IV together with values for elastic constants and phonon frequencies. We notice that the modifications essentially correspond to a reduction of the bond-bending force by a factor of 3. This simulates the TA-mode softening so characteristic of Si and Ge, at the expense of unrealistic softening of the long-wavelength shear mode which, because of the Brillouin-zone weighting argument given above, is of minor importance here.

With these values of α and β we have evaluated the total energy of a cluster containing about 100 atoms (up to 11 shells) and one vacancy. To obtain an effective lattice restoring force $k_{\text{eff}}=2E_{\text{tot}}/Q_\Gamma^2$ we have *imposed* distortions Q_Γ of the four-nearest-neighbor atoms to the vacancy, allowed all remaining atoms in the cluster to relax freely, and evaluated the minimum in the total energy according to Eq. (3.17) exactly up to quadratic terms. The results for breathing-type, tetragonal and trigonal distortions are given in Table V. To investigate the dependence of k_{eff} on local variations of α and β in the neighborhood of the vacancy, E_{tot} has been evaluated for two extreme limits: (a) with *no* bond-bending forces between the dangling

TABLE IV. Comparison of Keating parameter values given by Martin (Ref. 43) and those used in the present work. The resulting elastic constants (in 10^{11} dyn/cm^2), compressibility K (in 10^{-11} cm^2/dyn) and mode-frequency ratios are given.

	α	β	c_{11}	c_{12}	c_{44}	K	$\frac{\omega_{\text{expt}}}{\omega_{\text{calc}}}$/TO($\Gamma$)	$\frac{\omega_{\text{expt}}}{\omega_{\text{calc}}}$/TA($X$)
Reference 43	48.50	13.81	16.57	6.39	7.96	0.102	0.96	0.58
This work	51.51	4.70	12.07	8.60	3.17	0.102	1.01	1.00

TABLE V. Calculated effective force constants of Si for breathing-type, tetragonal, and two different trigonal distortions of nearest-neighbor atoms around the vacancy. The units of k are eV/Å^2 for actual atom displacements, *not* normal-mode amplitudes (to normalize to normal-mode amplitudes divide the values in Table V by a factor of 4).

	k_b	k_E	k_{T_2}	
This work	7.47	14.81	7.27	14.47
Reference 32	~4.0	5.20	3.88	

bonds and the back bonds and (b) with full bulklike bond-bending forces. The resulting k_{eff} values differed by roughly 20% indicating relative insensitivity to local variations. This insensitivity results from the relatively weak bond-bending forces and from quite large second-, third-, etc., nearest-neighbor distortions occurring when the crystal is allowed to relax freely. Figure 18 shows the relative displacement amplitudes of atoms in subsequent shells plotted versus shell radius. The decay of the amplitudes is compared to a $1/r^2$ behavior. Second-nearest-neighbor amplitudes of up to ~60% are observed. Physically, because of the relatively soft bond-bending forces, all atoms avoid bond-length changes when possible. This results in large displacements even for distant neighbors. Test calculations without free-atom relaxation, i.e., fixed neighbor positions, show a dramatic increase (~ factor of 2) in k_{eff}.

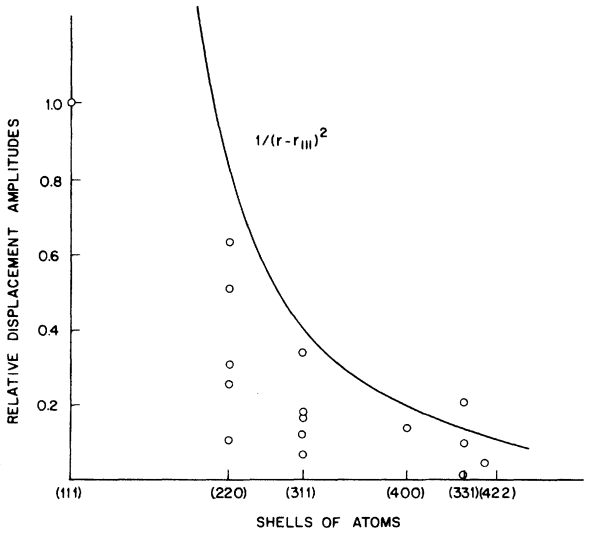

FIG. 18. Relative displacement amplitudes of atoms in subsequent shells surrounding vacancy. A pairing-type distortion of amplitude 1 is imposed onto the four nearest neighbors [(111) shell]; all other atoms are allowed to relax freely within the Keating force model. The particular choice of parameters for this model is discussed in the text.

Also in Table V are indicated effective force constants evaluated by Larkins and Stoneham.[32] These authors used a four-parameter valence force model to describe the interaction potential. They then solved for the total energy minimum as a function of imposed distortion by solving iteratively the dynamical equations of motion of a system of about 2000 atoms. The differences between their results and ours are considerable. In particular, k_E differs by a factor of about 3. We find from our calculations that both k_b and k_E scale roughly with the bond-bending force constant β. A decrease of β by a factor of 2–3, which would bring our results in rough agreement with those of Larkins and Stoneham, yields finite wave-vector phonon modes [such as, e.g., TA(X)] too low in frequency by about a factor of 1.5–2.0. Unfortunately no TA(X) mode-frequency calculation with the four-parameter model has been published by Larkins and Stoneham.

Another independent, *a posteriori*, argument in favor of our $k_E \approx 15$ eV/Å value comes from the analysis of experimental data on silicon vacancies[23] to be discussed in detail in Sec. IV. Consistency between several independent experiments requires formally that $E_{JT} = V_E^2/2k_E < 0.21$ eV, independent of any theoretical value for V_E or k_E. Combining our independently calculated value of $V_E \equiv dE/dQ_E = 2.25$ eV/Å with this upper bound we arrive at a lower bound for $k_E \approx 12$ eV/Å, about twice the value proposed by Larkins and Stoneham but in agreement with our Keating model calculations.

IV. DISCUSSION AND COMPARISON WITH EXPERIMENT

A. Summary of calculated values

In Sec. II, we formally defined the parameters needed to describe the total energy of the vacancy in various charge states and gave equations relating these quantities to the occupancy levels. In Sec. III, we obtained the following numerical values for them:

$$\begin{aligned} k_E &= 14.8 \text{ eV/Å}^2, \\ k_b &= 7.5 \text{ eV/Å}^2, \\ V_E &= 2.25 \text{ eV/Å}, \\ V_b &= 0.5 \text{ eV/Å}, \\ \epsilon_L &= 0.32 \text{ eV}, \\ U &= 0.25 \text{ eV}, \end{aligned} \tag{4.1}$$

and, by analogy with the (111) surface,[41,46] we have assumed an outward breathing equilibrium distortion for the neutral system of 0.1 Å. This gives

$$(\beta + 2V_b)/k_b = 0.1 \text{ Å}. \tag{4.2}$$

Remembering that V_b is an inward directed force (i.e., it has a sign opposite to that of β, which is an outward one) we use the values in (4.1) and obtain $\beta = 1.75$ eV/Å. From these values and Eq. (2.17) we obtain the following distortions for V^{++}, V^{+}, and V^{0}, respectively:

$$Q_b^0 = \begin{cases} 0.23\ \text{Å}, \\ 0.17\ \text{Å}, \\ 0.10\ \text{Å}, \end{cases} \quad Q_E^0 = \begin{cases} 0, \\ 0.15\ \text{Å}, \\ 0.3\ \text{Å}. \end{cases} \tag{4.3}$$

The Jahn-Teller energy, the total relaxation energy, and the shifted level, as given by (2.18), are $E_{JT} = 0.17$ eV, $\hat{E}_{JT} = 0.19$ eV, and $\hat{\epsilon}_L = 0.44$ eV. The value of $\hat{\epsilon}_L$ contains an upward shift of 0.12 eV arising from the term $V_b Q_b^0$. This may be too generous, as we shall see by comparison with experiment at the end of this section. The energies of the states as a function of μ is given by (2.19) with $\hat{E}_{JT} = 0.19$ eV and $\eta \equiv 2\hat{E}_{JT} - U = 0.13$ eV. The energy levels are given by (2.20) and by (2.22). The positive value of η means that the only occupancy level is $E(0/{+}{+})$ and that $E(+/{+}{+})$ or $E(0/+)$ is the excitation level which is a measure of the metastable V^{+}.

The condition for $\eta > 0$ is that $\hat{\epsilon}_{JT} > U/2$, which will be recognized as Anderson's condition[11] for a negative-effective U, generalized so as to include the two modes which are simultaneously driven by the presence of one and two extra electrons. With η calculated to be > 0, the predicted level $E(0/{+}{+})$ is at 0.18 eV. $E(0/+)$ is 0.12 eV and $E(+/{+}{+})$ is 0.25 eV. The values quoted here differ somewhat from those in our Letter[18] because that earlier work did not include the charge dependence of breathing; i.e., it used E_{JT} where we now find it proper to use $\hat{E}_{JT}$. The two differ, however, by only 0.02 eV, which is less than the precision we put on these calculations.

B. Fitting to experiment

We now approach the problem from a complementary direction. We ignore theoretical estimates of model parameters and use the kinetics and statistical mechanics of the model itself to fit the parameters to various experiments. At the end, this procedure will allow us to check the calculated values of model parameters. Instead of using k_E, ϵ_L, V_E, and U, it is convenient to use $\hat{\epsilon}_L$, U, and $\hat{E}_{JT}$ as independent model parameters. The kinetics and statistics of a system of vacancies in contact with a reservoir of electrons has been discussed by us in Ref. 19. We here summarize the results for the possible activation energies for transitions of the type $N_T \to N_T + 1$ or $N_T + 1 \to N_T$ by hole capture or emission to the valence band:

$$(0 \to 1)_{\text{act}} = \left(0 - \frac{\hat{\epsilon}_L}{2\hat{E}_{JT}}\right)^2 \hat{E}_{JT}, \tag{4.4a}$$

$$(1 \to 2)_{\text{act}} = \left(1 - \frac{\hat{\epsilon}_L + U}{2\hat{E}_{JT}}\right)^2 \hat{E}_{JT}, \tag{4.4b}$$

$$(1 \to 0)_{\text{act}} = \left(1 - \frac{\hat{\epsilon}_L}{2\hat{E}_{JT}}\right)^2 \hat{E}_{JT} + \mu, \tag{4.4c}$$

$$(2 \to 1)_{\text{act}} = \left(2 - \frac{\hat{\epsilon}_L + U}{2\hat{E}_{JT}}\right)^2 \hat{E}_{JT} + \mu. \tag{4.4d}$$

The extra energy μ in (4.4c) and (4.4d) is the energy needed to produce the hole which is captured in the $N_T + 1 \to N_T$ transition.[19] The terms in parentheses are the differences in energy from the initial-state minimum (N_T) or ($N_T + 1$) to the crossing of the $N_T \leftrightarrow N_T + 1$ configuration coordinate curves, evaluated using (2.15) at $\mu = 0$.[19] Again, notice that knowledge of lattice elastic energy (via $\hat{E}_{JT}$) is needed to evaluate activation energies, just as it was needed to evaluate occupancy levels.

The first experiment to be fitted is the activation energy for the decay of the V^{+} EPR signal in In-doped silicon ($\mu = 0.16$ eV) which Watkins has measured[6] as 0.057 eV, and which he ascribes to hole emission (a transition $N_T = 1 \to 2$ via capture of an electron from the top of the valence band). Once we postulate existence of V^{++}, another channel of V^{+} decay, namely, $V^{+} \to V^{++}$ must be considered. However, the activation energy for this channel (4.4c) is greater than μ because a hole must be present in the valence band before it can be captured by the vacancy. The numbers just quoted rule out this channel for V^{+} decay. The possibility of V^{+} decay via direct transfer between the vacancy and an In acceptor should also be small, and so we are led to accept Watkins's assignment of the 0.057 activation energy to V^{+} decay via hole emission. From (4.4b), we have $[(\hat{\epsilon}_L + U)/2\hat{E}_{JT} - 1]^2 \hat{E}_{JT} = 0.057$ eV. This condition is conveniently regarded as fixing $\hat{\epsilon}_L + U$ as a function of $\hat{E}_{JT}$, and we exhibit this in columns 1 and 2 of Table VI.

The second experiment to be fitted is the activation energy of 0.13 eV, observed by DLTS measurements in p-type silicon.[47] Watkins has suggested to us privately that the transition being observed here might be the $V^{++} \to V^{0}$ transition which we had proposed,[18] and has suggested that at these DLTS temperatures, after the initial $V^{++} \to V^{+}$ transition, the subsequent $V^{+} \to V^{0}$ transition is so rapid that only the $V^{++} \to V^{+}$ step is observed as an activated process in DLTS. The reason that the first transition $V^{++} \to V^{+}$ is the rate-limiting process can be seen in Fig. 19, taken from Ref. 47, which shows measured emission rate versus inverse temperature, both for the 0.13-eV

TABLE VI. Characteristic energies of Jahn-Teller distorted vacancies obtained by fitting two independent experiments to the three parameters of the model (see text). The magnitude of E_{JT} (column 1) is to be regarded as the remaining independent variable. Values below the dotted line can be ruled out by the existence of a DLTS signal in gallium-doped silicon.

1 $\hat{E}_{JT}$	2 $\hat{\epsilon}_L + U$	3 $\hat{\epsilon}_L$	4 U	5 η	6 $E(0/++)$
0.0500	0.2068	0.1612	0.0455	0.0545	0.0840
0.1000	0.3510	0.2280	0.1230	0.0770	0.0895
0.1500	0.4849	0.2793	0.2056	0.0944	0.0821
0.2000	0.6135	0.3225	0.2911	0.1089	0.0680
0.2500	0.7387	0.3606	0.3782	0.1218	0.0497
0.3000	0.8615	0.3950	0.4666	0.1334	0.0283
0.3500	0.9825	0.4266	0.5559	0.1441	0.0046
0.4000	1.1020	0.4561	0.6459	0.1541	−0.0210

DLTS signal and for the V^+ decay. Extrapolating the V^+ rate upwards, we see that at the temperatures for which the DLTS signal was observed, the V^+ decay rate is of the order of 10^4–10^6 times greater than the $V^{++} \rightarrow V^+$ decay which we identify with the 0.13-eV signal. With this suggestion we obtain from (4.4a) $\hat{\epsilon}_L^2/4\hat{E}_{JT} = 0.13$. This condition is conveniently regarded as fixing $\hat{\epsilon}_L$ as a function of $\hat{E}_{JT}$, and we exhibit this in column 3 of Table VI.

Columns 4, 5, and 6 of Table VI give certain quantities which can be calculated from entries in columns 1, 2, and 3. Column 4 gives the Coulomb repulsion term $U = (\hat{\epsilon}_L + U) - \hat{\epsilon}_L$. Column 5 gives $\eta \equiv 2\hat{E}_{JT} - U$. If $\eta > 0$, then there is no value of μ for which V^+ is the ground state of the system, that is, the equilibrium configuration changes directly from V^{++} to V^0 as the Fermi energy μ is raised. Notice that $\eta > 0$ for all entries in the table. This implies that within our model we do, in fact, have an Anderson negative-U system in this entire range of possible $\hat{E}_{JT}$ values. Column 6 gives $E(0/++) = \hat{\epsilon}_L - 2\hat{E}_{JT} + U/2$, which is the value of μ at which the ground state changes from V^{++} to V^0 as μ is raised.

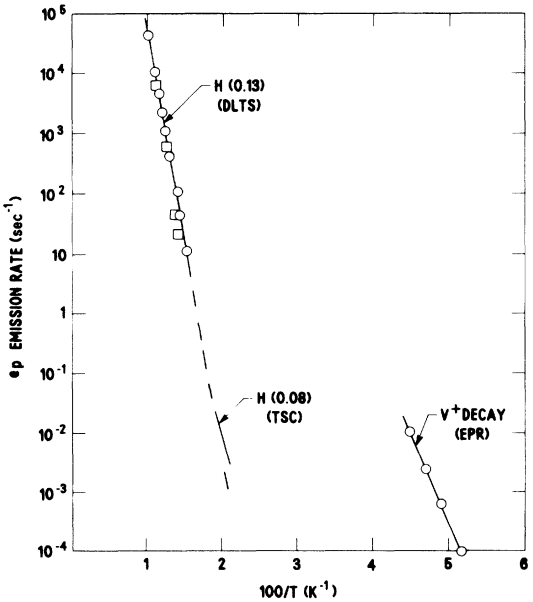

FIG. 19. Hole emission rate versus temperature for the V^+ state measured by EPR and for the $H(0.13)$ DLTS level which we associate with the activated decay $V^{++} \rightarrow V^0$. Thermally stimulated capacitance (TSC) data are also shown. Figure taken from Ref. 47.

As a third independent experimental input we consider now the implications of being able to observe a $V^{++} \rightarrow V^+$ transition in DLTS. The essence of a DLTS experiment is that the charge state of the deep level should alter (producing an observable change in the capacitance of a *pn* junction) when all the mobile carriers are swept away from the deep level. This means that its charge state under equilibrium conditions (as determined by the Fermi energy μ) and its charge state under depletion (when all mobile carriers are swept away) *cannot* be the same. The experiments make it clear that the state is a hole trap close to the valence band, so that in depletion, it is occupied with electrons, just as all the nearby valence-band state are fully occupied by electrons when there are no holes present. For this level to be observed, it must therefore be in the state V^+ or V^{++} during that brief time at the beginning of the experiment when its occupation is determined by μ. This requires that the value of μ be *less* than $E(0/++)$ in the sample in which the DLTS signal is observed. Watkins *et al.* report[47] seeing the 0.13-eV activation energy in p-type silicon samples doped with boron ($\mu = 0.045$), aluminum ($\mu = 0.057$), and gallium ($\mu = 0.065$). We therefore conclude that $E(0/++) > 0.065$ eV. According to Table VI, this implies a value of $\hat{E}_{JT}$ which is less than 0.21 eV, and thereby establishes allowable ranges of $\hat{\epsilon}_L$ and U (above dotted line).

Watkins also describes an attempt to measure the excitation energy for the state V^+ in boron-

TABLE VII. Comparison of calculated model parameters with bounds obtained by fitting to experimental data (from Refs. 6 and 47).

Model calculations	Parameter fit to experiments (Refs. 6 and 47)
$\hat{E}_{JT} = 0.19$ eV	$\hat{E}_{JT} \leq 0.21$ eV
$\hat{\epsilon}_L = 0.32 + 0.12$ eV	$\hat{\epsilon}_L \leq 0.33$ eV
$U = 0.25$ eV	$U < 0.33$
V^+ metastable	V^+ metastable

doped silicon by observing the strength of the EPR signal from V^+ at a finite temperature and comparing it with the strength of the EPR signal when the number of vacancies in the V^+ state has been enhanced by flooding the sample with infrared photons.[6] The excitation energy for V^+ (i.e., the amount by which the energy of the system in the state V^+ exceeds the energy of the ground state [which we believe is V^{++} because the Fermi energy μ is less than $E(0/++)$]) is

$$E_{ex} = E_1(\mu) - E_0(\mu) = \hat{\epsilon}_L - \hat{E}_{JT} - \mu . \tag{4.5}$$

For a value of $\mu = 0.045$ eV and values of $\hat{\epsilon}_L$ and $\hat{E}_{JT}$ from the allowed part of Table VII ($\hat{E}_{JT} < 0.21$ eV) we find $0.066 < E_{ex} < 0.077$. Watkins finds 0.006 eV. The discrepancy of 0.06 to 0.07 eV between our prediction and his measurement serves as a measure of the validity of the assumptions, such as those spelled out in Sec. II, on which our model is based, provided that the experimental estimate is valid.

There is, interestingly, a qualitatively new feature which should be considered in redoing the excitation energy experiment. If the equilibrium charge state in the boron-, or aluminum-, or gallium-doped samples is V^{++} as we believe, then the mechanism for generating V^+ by flooding the sample with radiation cannot be capture of the holes produced by ionizing the acceptors. It must be instead direct optical excitation of an electron from the top of the valence band to the vacancy. This process requires a photon energy $\hbar\omega > \hat{\epsilon}_L$, and so measuring the optical threshold for V^+ production will give a direct measurement of $\hat{\epsilon}_L$, one of the parameters of our model. In In-doped silicon, on the other hand, where $\mu = 0.16$, the value $E(0/++) < 0.09$ indicated in Table VII predicts that the ground state of the vacancy is V^0. This can be converted to V^+ by hole capture, so that *any* illumination capable of creating holes in the valence band should be effective in generating the V^+ in an In-doped sample. In one of the early experiments on the silicon vacancy,[43] Watkins mentions that the V^+ signal can be regenerated reproducibly in boron-doped silicon by illuminating with light passed through an InAs window. This would indicate that $\hat{\epsilon}_L \leq 0.35$ eV, which is compatible with the range of parameters we have deduced in Table VI by fitting the other two experiments.

In Table VII, we compare the parameters we have calculated with those we obtain by an *a posteriori* fit to experiment. The largest discrepancy is in the value of $\hat{\epsilon}_L$, which we have determined in part from the Green's-function self-consistent calculations for an unrelaxed vacancy and in part from a guess that the atoms near the vacancy might relax outward (breathing mode) by an amount comparable to that on the silicon (111) surface. The comparison with experiment seems to indicate that our guess of an outward breathing relaxation and the rise of 0.1 eV in the level position from 0.32 to 0.44 eV is too generous an estimate.

There are two important conclusions to be drawn here. First, we have obtained a rather convincing demonstration (by fitting to experiment and by direct calculation) that the Anderson effective U for silicon vacancies is indeed negative. Second, we have obtained a rather convincing limit on $\hat{E}_{JT}$ of 0.21 eV without relying on estimates of lattice response.

V. SUMMARY AND CONCLUSIONS

This study of the silicon vacancy can be divided conceptually into two complementary parts. First, in Sec. II, we describe a simple model for the occupancy levels and activation energies of this strongly lattice-coupled defect using three underlying parameters, $\hat{\epsilon}_L$, $\hat{E}_{JT}$, and U. $\hat{\epsilon}_L$ is a transition-state eigenvalue, shifted upwards by the outward breathing distortion. $\hat{E}_{JT}$ is a relaxation energy. It is composed of two parts: $V_E^2/2k_E$, the usual symmetry-lowering Jahn-Teller energy, and $V_b^2/2k_b$, an energy due to that part of the breathing relaxation which is charge dependent. U is the electron-electron repulsion energy for two electrons in the same state. Second, in Sec. III we described the extensive numerical calculations required to evaluate the parameters of the model.

What emerges from this study is a level structure at variance with the previously accepted one but *not* at variance with the experiments on which that previously accepted view is based. To substantiate this last point, we display, in Table VIII, predictions of the theory based on using the *calculated* parameters. Column 2 of that table gives the equation number (this paper) in which the quantity in column 1 is defined. Columns 3 and 4 are two sets of theoretically predicted values, both using the parameters calculated in Sec. III. The first, Calc. 1, uses $U = 0.25$ eV, $\hat{E}_{JT} = 0.19$ eV, and $\hat{\epsilon}_L = 0.44$ eV, an estimate based on the maximum value of breathing-induced upward level shift. The second, Calc. 2, uses $\hat{\epsilon}_L = 0.32$ eV, making *no* allowance for this shift. The next column gives the experimental value and the last is the reference which contains the data or statement on which the experimental value is based. In some cases, the logical connection between statement and experimental value is contained in the reference itself. In others, the logical connection is to be found in Sec. IV of the present work.

TABLE VIII. Predictions of the theory based on calculated parameters. Calc. 1 allows for maximum breathing distortion while Calc. 2 is based on zero breathing distortion.

Quantity	Formula	Calc. 1 eV	Calc. 2 eV	Expt. eV	Reference
$(V^+ \rightarrow V^0)_{act}$	Eq. (4.4b)	0.13	0.05	0.057	6
$(V^{++} \rightarrow V^+)_{act}$	Eq. (4.4a)	0.25	0.13	0.13	47
$E(0/++)$	Eq. (2.22)	0.18	0.06	>0.065	47
$V^+_{excitation}$	Eq. (4.5)	0.21	0.09	0.006	6
$\hat{\epsilon}_L$	Eq. (2.18d)	0.44	0.32	<0.35	43
η	Eq. (2.21)	0.13 eV $\eta > 0$ implies V^+ is always metastable		V^+ not stable	6, 43, 47

Agreement between Calc. 1 and experiment is of the order of 0.1 eV for all of the quantities compared, except for the excitation energy for V^+ in boron-doped silicon, which is off by 0.2 eV. The agreement becomes virtually exact (except for the V^+ excitation energy, which is then in error by 0.1 eV) if we lower $\hat{\epsilon}_L$ to the value it would have without the upward breathing shift as in Calc. 2. There is no point in trying to use this argument to improve the fit or *determine* the breathing shift because the eigenvalue itself can probably not be realistically calculated to better than 0.1 eV. There *is* reason to trust the calculation of the other parameters, however: They express how the eigenvalue *changes* in response to small variations in the system coordinates Q_E, Q_b and N_T.

The 0.1-eV discrepancy remaining between the predicted and experimental value of $V^+_{excitation}$ *cannot* be substantially reduced by a more careful calculation. We have demonstrated this in Sec. IV where we treated $\hat{\epsilon}_L$, $\hat{E}_{JT}$, and U as fully adjustable parameters and showed, by fitting the first two experimental values in Table VIII, that there was at least an error of 0.07 eV in the predicted V^+ excitation energy. This discrepancy points to small energy dependences of parameters that we have taken as constants in the three-parameter theory.

The parameter-fitting exercise of Sec. IV pointed to an internal inconsistency in the three-parameter theory of about 0.07 eV. It also provided some useful bounds on the other parameters, establishing limits for them as described in Table VII. These limits confirm (a) that the unit of relaxation energy $\hat{E}_{JT}$ cannot appreciably exceed 0.21 eV, which is considerably smaller than had been inferred[6] and (b) that the system *is* an Anderson U system for the states V^{++}, V^+, and V^0, so that the occupancy level is a two-electron level $E(0/++)$.

At this stage of development of the theory, it is reasonable to regard unification of the disparate experimental attributes of the silicon vacancy into a simple theoretical framework as a useful advance, even to within an accuracy of about ~0.1 eV. The basic ideas here can readily be generalized to discuss other systems, e.g., the negative charge states of the vacancy and separately, other strongly lattice-coupled point defects with near-degenerate states in the gap. Future work will undoubtedly seek more accuracy in their description, and we would like to comment on this.

Our starting point was neglect of spin-induced correlations. Even had we included them, it would still have been possible to formulate a theory along exactly the same lines as we followed in setting up Eq. (2.10), namely, that there is a *total* energy associated with V^{++} and that there are energies associated with *adding* one more and two more electrons. In such case, we could still *define* an e-e repulsion term by

$$U(Q_E, Q_b) = \epsilon_2(Q_E, Q_b) - \epsilon_1(Q_E, Q_b). \tag{5.1}$$

From this point of view, the only approximation intentionally restricting the flexibility of our parametrized theory was taking U to be independent of distortion. Clearly, a careful study of ϵ_1 and ϵ_2 would allow us to determine the size of the next most important missing terms, the linear dependence of U and the quadratic dependence of ϵ. Until these are evaluated, it would be premature to introduce a many-electron starting point *unless it turned out that the basic symmetry of the state in the gap were different in a many-electron theory from what it is in the local density theory used here.*

It might be useful to contemplate a serious attempt to evaluate the total energy functional, say for the state V^{++}, and thus obtain the constant β and another independent estimate of k_E and k_b. Most attempts to evaluate the total energy functional find it convenient to express its numerical value by replacing the kinetic energy by a term which depends on the eigenvalues, a step which can be carried out by multiplying (2.4a) by $n_i\psi_i^*(r)$, integrating over r and summing on i. The result

is

$$E_{\text{tot}}(\{n_i\},\{R_A\},\mu) = E_{\text{band}} + V_I - \Phi + \Delta E_{\text{xc}} - N_T\mu\,, \quad (5.2)$$

where

$$E_{\text{band}} \equiv \sum_i n_i \epsilon_i(\{n_j\},\{R_A\})\,, \quad (5.3)$$

and where ΔE_{xc} arises from the nonlinear dependence of exchange correlation on the electronic density:

$$\Delta E_{\text{xc}} = \int \rho(r)(\epsilon_{\text{xc}} - U_{\text{xc}})dr\,. \quad (5.4)$$

Several authors, among them Varma and Weber[48] and Chadi,[49] have studied forces and spring constants by calculating the displacement dependence of E_{band}, and parametrizing the displacement dependence of the remaining term $(V_I - \Phi + \Delta E_{\text{xc}})$. Similar parametrization might be useful in obtaining β, the linear term in the energy of the V^{++} breathing distortion. Recent attempts have also been reported to determine structural distortions, e.g., bulk lattice constants and surface relaxations, by direct evaluation of the displacement dependence of the $V_i - \Phi + \Delta E_{\text{xc}}$ term.[3] When these are fully explored, it should prove possible to use a similar approach for the point defects.

In Sec. IV, we have suggested that study of the energy $\hbar\omega$ necessary to photopopulate the V^+ state in boron-doped silicon would measure a useful parameter of the theory. Clearly, a direct measurement of the number of electrons captured by each vacancy in a DLTS transition would be a definitive test of the *two*-electron nature of the transition $V^0 \rightarrow V^{++}$ we predict. Watkins has suggested to us privately that stress coupling experiments, similar to those by which he infers the presence of V^0 by studying the anisotropic distribution of V^- generated from V^0 under stress, might be applied to V^+. In such case, V^+ generated from V^{++} under stress would be isotropically distributed while V^{++} produced from V^0 would be *an*isotropically distributed, because unlike V^0, V^{++} has *no* anisotropic stress coupling coefficient. On the basis of the experimental determination $0.065\text{ eV} < E(0/{+}{+}) < 0.16\text{ eV}$, we would expect V^+ to reveal no pre-existing stress alignment in boron-, aluminum-, or gallium-doped samples but to exhibit it in In-doped samples. Measurements in samples where the Fermi energy was swept through this range would be of particular interest.

Note added in proof. Watkins and Troxell[50] have very recently reported that two electrons per vacancy do seem to be involved in the 0.13-eV DLTS transition. This observation provides the *direct* confirmation of our prediction of the Anderson-negative-U nature of the Si vacancy.

ACKNOWLEDGMENTS

It is a pleasure to acknowledge stimulating discussions with many people during the course of this work, in particular, George D. Watkins, V. Narayanamurti, D. V. Lang, L. C. Kimerling, C. H. Henry, J. Bernholc, J. van Vechten, and M. Jaros. We also express special thanks to Sue Marta McDonald for critical discussions of this manuscript.

[1] J. A. Appelbaum and D. R. Hamann, Phys. Rev. B 8, 777 (1973).
[2] M. Schlüter, J. R. Chelikowsky, S. G. Louie, and M. L. Cohen, Phys. Rev. B 12, 4200 (1975).
[3] J. Ihm and M. L. Cohen, Solid State Commun. 29, 711 (1979).
[4] G. D. Watkins, in *Point Defects in Solids*, edited by J. H. Crawford, Jr. and L. M. Slifkin (Plenum, New York, 1975), Vol. 2, p. 333.
[5] G. L. Miller, D. V. Lang, and L. C. Kimerling, in Ann. Rev. Mater. Sci. 7, 1977, pp. 377–448.
[6] G. D. Watkins, in *Lattice Defects in Semiconductors—1974*, edited by F. A. Huntley (Institute of Physics, London, 1975), p. 1.
[7] G. D. Watkins, in *Radiation Damage in Semiconductors* (Dunod, Paris, 1964), p. 97.
[8] G. A. Baraff and M. Schlüter, Phys. Rev. Lett. 41, 892 (1978).
[9] G. A. Baraff and M. Schlüter, Phys. Rev. B 19, 4965 (1979).
[10] P. N. Keating, Phys. Rev. 145, 637 (1966).
[11] P. W. Anderson, Phys. Rev. Lett. 34, 953 (1975).
[12] C. A. Coulson and M. J. Kearsley, Proc. R. Soc. London A 241, 433 (1957).
[13] J. Bernholc, N. Lipari, and S. T. Pantelides, Phys. Rev. Lett. 41, 895 (1978).
[14] E. Kauffer, P. Pecheur, and M. Gerl, J. Phys. C 9, 2319 (1976).
[15] S. G. Louie, M. Schlüter, J. R. Chelikowsky, and M. L. Cohen, Phys. Rev. B 13, 1654 (1976).
[16] J. Bernholc and S. T. Pantelides, Phys. Rev. B 18, 1780 (1978).
[17] M. D. Sturge, in *Solid State Physics*, edited by F. Seitz, D. Turnbull, and H. Ehrenreich (Academic, New York, 1967), Vol. 20, p. 91.
[18] G. A. Baraff, E. O. Kane, and M. Schlüter, Phys. Rev. Lett. 43, 956 (1979).
[19] G. A. Baraff, E. O. Kane and M. Schlüter, Phys. Rev. B 21, 3563 (1980).
[20] G. A. Baraff, E. O. Kane and M. Schlüter, (unpublished).
[21] M. Lannoo, in *Defects and Radiation Damage in Semiconductors—1978*, edited by J. H. Albany (Institute of Physics, London, 1979), p. 1.

[22]G. T. Surratt and W. A. Goddard, Phys. Rev. B 18, 2831 (1978).
[23]G. D. Watkins and R. P. Messmer, Phys. Rev. Lett. 32, 1244 (1974).
[24]G. D. Watkins, in *Radiation Defects and Damage in Semiconductors—1972*, edited by J. E. Whitehouse (Institute of Physics, London, 1973), p. 228.
[25]P. Hohenberg and W. Kohn, Phys. Rev. 136, B864 (1964).
[26]W. Kohn and L. Sham, Phys. Rev. 140, A1133 (1965).
[27]M. Born and R. Oppenheimer, Ann. Phys. (Paris) 84, 457 (1927).
[28]R. Kubo and Y. Toyazawa, Prog. Theor. Phys. 13, 160 (1955).
[29]H. D. Shih, F. Jona, D. W. Jepsen, and P. M. Marcus, Phys. Rev. Lett. 37, 1622 (1976).
[30]J. C. Slater, *The Self Consistent Field for Molecules and Solids* (McGraw-Hill, New York, 1974).
[31]J. F. Janak, Phys. Rev. B 18, 7165 (1978).
[32]F. P. Larkins and A. M. Stoneham, J. Phys. C 4, 143 (1971).
[33]G. A. Baraff and M. Schlüter, Phys. Rev. B 20, 4363 (1979).
[34]E. O. Kane, Phys. Rev. B 13, 3478 (1976).
[35]D. R. Hamann, Phys. Rev. Lett. 42, 662 (1979).
[36]G. Gilat and Z. Kam, Phys. Rev. Lett. 22, 715 (1969).
[37]P. Pecheur, E. Kauffer, and M. Gerl, in *Defects and Radiation Effects in Semiconductors—1978*, edited by J. H. Albany (Institute of Physics, London, 1979), p. 174.
[38]N. O. Lipari, J. Bernholc, and S. T. Pantelides, Phys. Rev. Lett. 43, 1354 (1979).
[39]M. Jaros, C. O. Rodrigues, and S. Brand, Phys. Rev. B 19, 3137 (1979).
[40]D. Haneman, Phys. Rev. 121, 1093 (1961).
[41]J. Ihm and M. L. Cohen, Solid State Commun. 29, 711 (1979); A. Redondo, W. A. Goddard, T. C. McGill, and G. T. Surratt, *ibid.* 20, 733 (1976).
[42]R. P. Messmer and G. D. Watkins, Phys. Rev. B 7, 2568 (1973).
[43]G. D. Watkins, J. Phys. Soc. Jpn. 18, Suppl. II 22 (1963).
[44]R. M. Martin, Phys. Rev. B 1, 4005 (1970).
[45]R. M. Martin and D. J. Chadi, in *Proceedings of the XIII International Conference on the Physics of Semiconductors, Rome, 1976*, edited by F. G. Fumi (North-Holland, Amsterdam, 1977), p. 187.
[46]J. V. Florio and W. D. Robertson, Surf. Sci. 24, 173 (1971) and H. D. Shih, F. Jona, D. W. Jepsen, and P. M. Marcus, Phys. Rev. Lett. 37, 1622 (1976).
[47]G. D. Watkins, J. R. Troxell, and A. P. Chatterjee, in *Defects and Radiation Effects in Semiconductors—1978*, edited by J. H. Albany, (Institute of Physics, London, 1979), p. 16.
[48]C. Varma and W. Weber, Phys. Rev. B 19, 6142 (1979).
[49]D. J. Chadi, Phys. Rev. Lett. 41, 1062 (1978).
[50]G. D. Watkins and J. R. Troxell, Phys. Rev. Lett. 44, 593 (1980).