

New self-consistent approach to the electronic structure of localized defects in solids

G. A. Baraff and M. Schlüter

Bell Laboratories, Murray Hill, New Jersey 07974

(Received 3 November 1978)

We have developed a new method to calculate electronic states associated with localized defects. The method yields energy levels, wave functions, and electron charge densities with a precision which is equal to state-of-the-art results for bulk band structure or surface calculations and thus should eventually allow meaningful detailed comparison with experiments. The calculations are based on a scattering-type Green's-function formulation which describes the system of a single isolated defect in an otherwise perfect crystal. Self-consistency is achieved by an iterative recalculation of the valence charge and then of the defect potential. The Green's function, which expresses all of the necessary information about the perfect crystal, is here evaluated using an eigenfunction expansion employing high-precision wave functions and band structures obtained from a self-consistent, pseudopotential, local-density-functional calculation. The defect potential is calculated in the same approximation, using occupied electron states of the perturbed system. The use of this method is not restricted, however, to pseudopotentials. The possibility of extending it, e.g., in the direction of crystal ionic potentials with core electrons, seems quite real in the event that such extension should be needed, say, for studying the energetics of relaxation or reconstruction around the vacancy. The method is applied to the example of an isolated Si vacancy. This system has been chosen to facilitate comparison to earlier non-self-consistent, or self-consistent but artificially periodic, calculations. The results generally agree with these earlier works but are improved in various aspects.

I. INTRODUCTION

The precise determination of the electronic structure of localized or deep defects represents a problem which is still unsolved. While much progress has been made in quantitatively explaining delocalized or shallow defects¹⁻³ by employing effective-mass-type techniques, the extension of these approaches to localized defects fails. Several formal approaches have been proposed to describe localized defects¹⁻⁵ but only a few quantitative calculations have been attempted. Among those, the defect molecule or cluster calculations are most frequently seen,⁶ but these usually provide only semiquantitative information due to the absence of any connection to the host-crystal band structure. Qualitative perturbation calculations, based on the idea of Slater and Koster (Ref. 4) have been carried out in several cases.⁷ Quantitative calculations based on the Slater-Koster scheme, have been carried out by Jaros and Brand to study vacancies and isoelectronic impurities in semiconductors.⁸ Though the calculations focused mainly on energies, were not self-consistent, and yielded only rather imprecise wave functions, Jaros has indicated⁹ that he has made substantial progress in the direction of developing a self-consistent version of his scheme.

Another approach, actually a bulk-crystal calculation based on a periodically repeated large unit cell containing one vacancy per cell, and using self-consistent pseudopotentials, has been published recently.¹⁰ Due to its simplicity this

approach is very appealing. Its main disadvantage, however, is that the defect wave functions, though highly localized, still overlap sufficiently from one cell to the next to give rise to a "defect band structure" with a width of the order of 1 eV. To remedy this situation, one would have to significantly increase the size of the large unit cell, which at this time is prohibitively expensive, even on the most advanced computers. While the defect wave functions and thus also the eigenenergies of individual states are affected by the proximity of neighboring defects, the total charge density is not. This is a well-known and important result which has also been found in various surface calculations^{11,12}; namely, that the self-consistent potential and charge density heals rather fast (within one or two bond lengths) as one leaves the surface (or neutral defect) while individual states can still have rather long tails of several bond lengths extension. This point must be borne in mind when we compare our test example with the ideal, neutral Si vacancy as considered in Ref. 10.

The study of the various theoretical approaches cited above was motivation to develop a new quantitative method for describing localized defects which should be free of the inherent disadvantages of existing methods. In particular, the method should describe a truly isolated defect. It should be capable of yielding wave functions, it should be quantitative, it should be related to the host band structure (as the various defect cluster, or large-molecule calculations are not) and it should be self-consistent.

Two groups—ourselves¹³ and Bernholc, Lipari, and Pantelides at IBM¹⁴—have responded to this motivation in similar ways and have decided that the Slater-Koster, or Green's-function technique, can be used as the key idea in a practical computation scheme. The schemes developed by the two groups differ, however, in the basic equations which are proposed and solved, and perhaps in other subtler details as well. Hence, detailed explanation of the methods and comparison of the results obtained using them is of interest.

This paper represents our contribution to such an explanation and comparison. Its object is to present the scheme we have developed for performing quantum-mechanical calculations of the self-consistent potential, charge density, electronic wave functions, and energies associated with an isolated, localized point defect in an otherwise perfect infinite crystal. Having described our technique, we present the first results we have obtained using it to study the neutral vacancy in silicon.

Basically, our technique can be described as a quantum-chemical linear-combination-of-atomic-orbitals (LCAO) calculation in which the coefficients in the LCAO expansion are determined by making use of the Green's-function equations, rather than by using the Schrödinger equation directly. The advantage in using the Green's function, as was emphasized by Koster and Slater from the outset, is that one needs to consider the problem only in that region of space for which the impurity potential exists. Such a region is typically much smaller than the region over which the wave functions extend, at least for those types of point defects for which a Green's-function formulation is appropriate. The resulting problem, solving for the LCAO expansion coefficients in the limited region of space, is simple enough computationally that one can obtain the bound states and the Green's function of the perturbed crystal in sufficient detail to allow synthesis of the charge density. This charge density then gives rise to a new impurity potential and the whole process is repeated until input and output potentials agree.

Choice of the ideal vacancy in silicon as a first test of the method is an appropriate one for several reasons. First of all, the electronic structure of atomic silicon, of bulk crystalline silicon, and of various silicon surfaces has been sufficiently studied, both experimentally and theoretically, that there now exist reliable pseudopotentials which describe the interaction between the silicon (pseudo-) core (charge +4) and the valence electrons surrounding it. Secondly, the neutral vacancy is perhaps among the strongest and most localized defects which can be expected in a crystal

and therefore ideally suited for a Green's-function study. Finally, there exists a large number of earlier theoretical calculations with which we compare the results of the technique presented here. This comparison is initially more important than is any comparison between a new theoretical calculation of vacancy energy levels and experiment, because of uncertainty as to the exact position of atoms in the neighborhood of the vacancy. Indeed, one of the long-range goals of a calculation of the electronic structure of a defect is to provide additional information about atomic positions nearby by studying the relation between the energy spectrum and the position of atoms near the defect. In the present work, however, we shall be concerned only with the calculation of the ideal neutral silicon vacancy—that is, a crystal from which a single atom has been removed without allowing the adjacent atoms to readjust position, and with the comparison of these results to the results of earlier calculations of the same system. The comparison indicates that the technique we use is capable of a precision on the order of that obtained in state-of-the-art calculations of bulk semiconductor crystals and surfaces. For this reason, we anticipate that the method will be able to provide useful information on relaxation and reconstruction around the defects as well.

We are aware that localized states, such as those associated with a vacancy, might possibly give rise to specific correlation effects and multiplet structures which lie beyond the single-particle approximation we are using here.¹⁵ The first step in estimating the size of such effects is to calculate, in a basis of reliable one-particle states, the matrix elements of the appropriate energy operator. States such as those we calculate here will be useful for that purpose but we see no point in estimating these effects until after a thorough study of the relaxed vacancy has been carried out.

In Sec. II, we describe the formalism used for obtaining the charge density, electronic states, change in energy density of states, and bound-state energy. Section III describes briefly the construction of the defect potential. In Sec. IV, we give specific details about the vacancy calculation itself. Results of the calculation are presented and analyzed in Sec. V.

II. METHOD OF CALCULATION

A. Individual wave functions

A crystalline (point) defect is a break in the translational invariance of the lattice and is characterized by a defect potential which tends to zero as one leaves the disturbance. This situation can generally be described in the one-electron ap-

proximation by a Schrödinger equation of the form

$$(H_0 + U)\psi(r) = E\psi(r), \quad (1)$$

where $H_0 = -\frac{1}{2}\nabla^2 + V_c$ denotes the Hamiltonian of the perfect crystal and U the additional potential due to the disturbance. The case of strongly localized defects is characterized by U being a strong but localized perturbation.

The form of the potential V_c describing the periodic crystal will depend on what level of sophistication one uses to reduce the actual many-body problem to a one-electron problem. In any case, it will have the translational periodicity of the perfect infinite crystal, and thus the eigenfunctions of H_0 can be Bloch states of the perfect crystal,

$$H_0\psi_n(k, r) = E_n(k)\psi_n(k, r). \quad (2)$$

Let these be normalized in such a way that the orthogonality and completeness relations are

$$\int \psi_n^*(k, r)\psi_{n'}(k', r)d^3r = \delta_{nn'}\delta^3(k - k'), \quad (3a)$$

$$\sum_n \int_{\text{BZ}} d^3k \psi_n(k, r)\psi_n^*(k, r') = \delta^3(r - r'). \quad (3b)$$

Then a useful representation of the Green's function for the perfect crystal is

$$G_E(r, r') \equiv \sum_n \int_{\text{BZ}} d^3k \frac{\psi_n(k, r)\psi_n^*(k, r')}{E - E_n(k)}. \quad (4)$$

Using this Green's function, the Schrödinger equation (1) can be transformed into well-known integral equations for bound states and scattering states.

(i) For bound states,⁴ where E is not in the spectrum of H_0 ,

$$\psi(r) = \int G_E(r, r')U(r')\psi(r')d^3r'. \quad (5a)$$

(ii) For scattering states,^{4,5} where E is in the spectrum of H_0 ,

$$\psi(r) = \psi_{n_0}(k_0, r) + \int G_E(r, r')U(r')\psi(r')dr', \quad (5b)$$

$$E \equiv E_{n_0}(k_0) + i0^+.$$

These equations are exactly equivalent to Eq. (1); the task is to devise a practical and accurate method to solve them. We proceed in two steps as follows.

In the first step, we regard $U(r)\psi(r)$, not $\psi(r)$ itself, as the basic unknown. Equations expressing this point of view are obtained by multiplying Eqs. (5) by $U(r)$:

$$U(r)\psi(r) = U(r) \int G_E(r, r')U(r')\psi(r')d^3r', \quad (6a)$$

$$U(r)\psi(r) = U(r)\psi_{n_0}(k_0, r) + U(r) \int G_E(r, r')U(r')\psi(r')d^3r', \quad (6b)$$

$$E \equiv E_{n_0}(k_0) + i0^+.$$

When $\psi(r)$ is a bound state, the normalization condition on $U(r)\psi(r)$ can be obtained from Eq. (5a) and the normalization condition on $\psi(r)$ itself. By writing Eq. (5a) in the symbolic form $\psi = G_E U \psi$, the normalization of the bound state becomes

$(\psi U G_E G_E U \psi) = (\psi \psi) = 1$. However, the properties of the Green's function contained in (3a) and (4) ensure that, for any value of E not in the spectrum of H_0 , i.e., for any value of E at which a bound state might possibly occur, $G_E G_E = -dG_E/dE$. Thus, the normalization condition to be used with (6a) is

$$-\iint \psi(r)U(r) \left[\frac{dG_E(r, r')}{dE} \right] U(r')\psi(r')drdr' = 1. \quad (7)$$

Let us now think of $\psi(r)$ as a linear combination of atomic or other conveniently localized orbitals. Each atomic orbital is centered on an atomic position and each atomic position will be assigned a number of atomic orbitals. The exact number and exact types of orbitals at each location is to be determined either by convergence studies or by intuition based on the chemistry of the system under investigation but in any case, if the atomic positions are shifted to represent relaxation of the crystal around a defect, the orbital centers shift also. A vacancy location has no orbitals associated with it, because there is no atom there to which they belong. Thus, we consider an LCAO expansion

$$\psi(r) \approx \sum_i C_i \phi_i(r), \quad (8)$$

where $\phi_i(r)$ is an atomic orbital whose form and location are fixed at the outset, and the C_i are the set of coefficients which are to be determined. Because $U(r)\psi(r)$ is our unknown, however, we need not satisfy (8) in all space but instead, we need only include enough atomic centers and enough orbitals on each center that the following, weaker, approximation is valid:

$$U(r')\psi(r') \approx U(r') \sum_i C_i \phi_i(r'). \quad (9)$$

It is of course possible to satisfy requirement (9) using a reasonable number of orbitals $\phi_i(r)$ when $U(r)$ is a localized potential. No orbital which lies beyond the range of the potential U is needed. Few orbitals are needed near the outer limits of the potential because in satisfying Eq. (9), it is wasted effort to approximate $\psi(r')$ accurately

where $U(r')$ is small. The greatest orbital flexibility, i.e., largest number of orbitals per location, should be permitted where the overlap between the defect potential and the wave function is expected to be greatest.

Having described the choice of orbitals to be used, we now consider the equations which the coefficients C_i must satisfy. For this purpose, we start with the bound-state equation (6a) and a variational principle associated with it, namely,

$$\delta\Lambda[\phi, \epsilon] / \delta\phi(r) = 0 \quad (10a)$$

and

$$\Lambda[\phi, E] \equiv N[\phi] / D[\phi, E] = 1, \quad (10b)$$

with

$$N[\psi] \equiv \int \psi^*(r) U(r) \psi(r) dr \quad (11a)$$

and

$$D[\psi, E] \equiv \int \psi^*(r) U(r) G_E(r, r') U(r') \psi(r') dr dr' \quad (11b)$$

The conditions (10) are satisfied when $\phi(r)$ is the solution of Eq. (6a). By choosing (3) as the trial function ϕ and varying with respect to the C_i , we obtain

$$\sum_j [N_{ij} - D_{ij}(E)] C_j = 0, \quad (12)$$

$$N_{ij} = \int \phi_i^*(r) U(r) \phi_j(r) dr, \quad (13a)$$

$$D_{ij}(E) \equiv \int \phi_i^*(r) U(r) G_E(r, r') \times U(r') \phi_j(r') dr dr'. \quad (13b)$$

Equation (12) has no solution unless

$$\text{Det}[N_{ij} - D_{ij}(E)] = 0, \quad (14)$$

which determines the energy E to second-order accuracy in the wave-function error. No first-order error occurs because our matrices N and D each contain an additional factor of $U(r)$ which renders the energy variational. The variational aspect of the energy is discussed in the Kohn-Rostoker¹⁶ paper on a method for computing Bloch waves using a Green's-function technique but the underlying mathematics is so similar to what we have here that the Kohn-Rostoker discussion applies without change.

Normalization of the coefficients C_j is fixed by inserting (9) into (7) with the result

$$-\sum_{ij} C_i^* \left[\frac{d}{dE} D_{ij}(E) \right] C_j = 1. \quad (15)$$

For the scattering states described by Eq. (6b), the functional to be varied is

$$J[\phi] = 2\text{Re}F[\phi] - N[\phi] + D[\phi, E], \quad (16)$$

where

$$F[\phi] \equiv \int \phi^*(r) U(r) \psi_0(r) dr, \quad (17)$$

and the algebraic equations which govern the C_i coefficients are

$$\sum_j [N_{ij} - D_{ij}(E)] C_j = F_i(n_0, k_0), \quad (18)$$

where

$$F_i(n_0, k_0) \equiv \int \phi_i^*(r) U(r) \psi_{n_0}(k_0, r) dr. \quad (19)$$

Note that no use has been made of orbital orthogonality, and that choice of either orthogonalized or nonorthogonal orbitals is possible. Our preference, to use the simple localized LCAO's directly, is based on the observation that the longer-ranged oscillatory tails which orbitals acquire in the orthogonalization process will cause orbitals from more-distant locations to extend into the region of finite $U(r)$. They will therefore have to be included in the set of orbitals being used in (9), increasing the size of the computation, with no compensating gain in accuracy, computational simplicity, understanding, or ease of interpretation. The choice of orbitals whose centers move with the position of the atoms themselves is motivated by the same consideration; keeping the size of the LCAO set as small as possible consistent with describing the particular defect under study.

We now consider the second step of the calculation, evaluation of $\psi(r)$ itself once $U(r)\psi(r)$ has been determined. For this, we must evaluate the integrals which appear in Eq. (5). The procedure we have found most convenient here is closely related to the method by which the matrix elements $D_{ij}(E)$ are evaluated, as we now explain.

Let us introduce a second set of localized orbitals, infinite, periodic (again, not necessarily orthogonal) but complete enough in every unit cell to provide an approximation to the Bloch waves in the form

$$\psi_n(k, r) = \sum_m B_m(n, k) \Phi_m(r). \quad (20)$$

The localized orbitals $\Phi_m(r)$ have to be complete enough so that (20) is valid for each of the $\psi_n(k, r)$ which are needed for construction of the Green's function in the form (4). Periodicity imposes a far more severe requirement than is imposed on the LCAO set used in (8) or (9), because the LCAO's have the additional flexibility of being al-

lowed to move with the atom or to change shape if the atom is a chemically different species. The two sets of orbitals, $\phi_i(r)$, the LCAO or "inner" set, and $\Phi_m(r)$, the periodic or "outer" set, are not related to each other in any way.

Coefficients $B_m(n, k)$ for the expansion (20) are fixed by carrying out a least-squares fit subject to the constraint of strict normalization. If this least-squares fit is carried out in Fourier space, then the number of linear simultaneous algebraic equations to be solved at each value of n and k is equal to the number of localized orbitals $\Phi_m(r)$ per unit cell, and the effort involved in this step is about equal to that of a band-structure calculation. The functions $\psi_n(k, r)$ to which the fitting is done and the associated energies $E_n(k)$ were calculated previously using a self-consistent pseudopotential Hamiltonian and a plane wave basis for representing the wave function, a feature which simplifies the subsequent determination of the $B_m(n, k)$ coefficients.

Using this local representation of the $\psi_n(k, r)$, the Green's function (4) takes the form

$$G_E(r, r') = \sum_{mm'} \Phi_m(r) G_{mm'}(E) \Phi_m^*(r'), \quad (21)$$

where the matrix

$$G_{mm'}(E) \equiv \sum_n \int d^3k \frac{B_m(n, k) B_{m'}^*(n, k)}{E - E_n(k)}, \quad (22)$$

is evaluated via a Gilat-Raubenheimer-Kam¹⁷ technique for the imaginary part and a Kramers-Kronig transform for the real part. Although costly, calculation of this matrix is a one-time expense, and it can be stored and used for a variety of defect calculations.

Note that the $G_{mm'}(E)$ are *not matrix elements* of the Green's function in any particular set of states. They are *expansion coefficients* by which the Green's function can be conveniently expressed in a position representation. If the set $\Phi_m(r)$ had been an orthonormal set of functions, then the matrix elements and the expansion coefficients would be identical numerically, but again there is no apparent gain to be obtained by orthogonalizing the set of outer orbitals while there is a problem if long-range tails introduce orbitals whose centers lie beyond the region of interest.

Having developed the Green's function and wave function in terms of the outer set of orbitals and having found $U(r)\psi(r)$ in terms of the inner set of orbitals, we have the scattering state in (5b) expressed in the form

$$\psi(r) = \sum_m \Phi_m(r) B_m(n_0, k_0) + \sum_{mnj} \Phi_m(r) G_{mn}(E) U_{nj} C_j, \quad (23)$$

$$E = E_{n_0}(k_0) + i0^+,$$

where

$$U_{nj} \equiv \int \Phi_n^*(r) U(r) \phi_j(r) dr \quad (24)$$

is the matrix element of the impurity potential taken between an orbital in the outer set and one in the inner set. The evaluation of the bound state in (5a) differs only in that the energy E in (23) is to be set equal to that of the bound state as determined by (14), and the first term—the incoming state in (23)—is absent.

These matrix elements of impurity potential between outer- and inner-set orbitals appear again in the evaluation of $D_{ij}(E)$ —Eq. (13b). After using (21) and (24), this matrix takes the form

$$D_{ij}(E) = \sum_{mn} U_{ni}^* G_{mn}(E) U_{nj}. \quad (25)$$

Note that matrix elements between two inner-set orbitals alone have already appeared, in Eq. (13a), as the matrix N_{ij} . Thus the equation (14) determining the bound-state energy could have been written

$$\det \left[U_{ij} - \sum_{mn} U_{ni}^* G_{mn}(E) U_{nj} \right] = 0, \quad (26a)$$

where the indices ij refer to the inner set orbitals, few in number, and the indices mn refer to the outer set orbitals, many in number.

There is, of course, a relationship between Eq. (26a) and the equation for the bound state which emerges when the Green's-function equation (5a) is simply expressed in an orthonormal basis,¹⁴ namely,

$$\det \left\| \delta_{\alpha\beta} - \sum_{\gamma} G_{\alpha\gamma}^0(E) U_{\gamma\beta} \right\| = 0. \quad (26b)$$

However, these equations are certainly not identical in terms of the size of the determinants which must be considered to achieve a given level of precision in determining E . We suspect strongly that the variational nature of E determined using (26a) favors this form over (26b) for a given size determinant but this matter has not been studied in detail.

B. Evaluation of the charge density

The state $\psi(r)$ given in (23) is one which corresponds to and is labeled by a particular incoming state, n_0, k_0 . We also label the coefficients C_j by the incoming state n_0, k_0 . Let us also introduce a matrix which will appear frequently below, namely,

$$\tilde{D}_{mj}(E) \equiv \sum_n G_{mn}(E) U_{nj}. \quad (27)$$

Then the state in (23) is

$$\psi(n_0 k_0, r) = \sum_m \Phi_m(r) \left(B_m(n_0, k_0) + \sum_j \tilde{\mathfrak{D}}_{mj}(E) C_j(n_0, k_0) \right). \quad (28)$$

The charge density is obtained by summing over all such scattering states (n_0, k_0) whose energies

$$\rho_{mm'} \equiv \sum_{n_0} \int_{\text{BZ}} d^3 k_0 \Theta(E_F - E_{n_0}(k_0)) \left(B_m(n_0, k_0) + \sum_j \tilde{\mathfrak{D}}_{mj}(\epsilon) C_j(n_0, k_0) \right) \times \left(B_{m'}(n_0, k_0) + \sum_{j'} \tilde{\mathfrak{D}}_{m'j'}(\epsilon) C_{j'}(n_0, k_0) \right)^*, \quad \epsilon \equiv E_{n_0}(k_0) + i0^+. \quad (30)$$

The coefficients C_j may be evaluated from (18) and (19), making use of the localized representation (20) and the definition of the matrix elements (24) between the inner and outer sets of orbitals:

$$C_j(n_0, k_0) = \sum_{in} M_{ji}^{-1}(E_{n_0}(k_0)) U_{ni}^* B_n(n_0, k_0), \quad (31)$$

where

$$M_{ij}(E) \equiv N_{ij} - D_{ij}(E + i0^+). \quad (32)$$

We insert (31) into (30) and before integrating, insert a

$$\int dE \delta(E - E_{n_0}(k_0)) = 1$$

into the integrand. The only terms in the integrand which depend on (n_0, k_0) , as opposed to being functions of $E_{n_0}(k_0)$, are the $B_m(n_0, k_0)$ terms. The n_0 sum and the k_0 integration can be performed first before the E integration, using

$$\sum_{n_0} \int d^3 k_0 \delta(E - E_{n_0}(k_0)) B_m(n_0, k_0) B_n^*(n_0, k_0) = -\pi^{-1} \text{Im} G_{mn}(E + i0^+),$$

a result which follows from (22). This procedure lets (30) be expressed as

$$\rho_{mm'} = -\frac{1}{\pi} \int^{E_F} dE \left[\rho_{mm'}^{(0)}(E) + \rho_{mm'}^{(1)}(E) + \rho_{mm'}^{(2)}(E) + \rho_{mm'}^{(3)}(E) \right], \quad (33a)$$

with

$$\rho_{mm'}^{(0)}(E) \equiv \text{Im} G_{mm'}(E + i0^+), \quad (33b)$$

$$\rho_{mm'}^{(1)}(E) \equiv \sum_{ijn} \tilde{\mathfrak{D}}_{mj}(E + i0^+) \times M_{ji}^{-1}(E) U_{ni}^* \text{Im} G_{nm'}(E + i0^+), \quad (33c)$$

$$\rho_{mm'}^{(2)}(E) \equiv \sum_{i'j'n'} \text{Im} G_{mn'}(E + i0^+) U_{n'i'} \times M_{j'i'}^{-1}(E) \tilde{\mathfrak{D}}_{m'j'}^*(E + i0^+), \quad (33d)$$

lie below the Fermi energy E_F :

$$\rho_{\text{scat}}(r) \equiv \sum_{n_0} \int_{\text{BZ}} d^3 k |\psi(n_0 k_0 r)|^2 \Theta(E_F - E_{n_0}(k_0)) = \sum_{mm'} \Phi_m(r) \rho_{mm'} \Phi_{m'}^*(r), \quad (29)$$

where

$$\rho_{mm'}^{(3)}(E) \equiv \sum_{\substack{ijn \\ i'j'n'}} \tilde{\mathfrak{D}}_{mj}(E + i0^+) M_{ji}^{-1}(E) \times U_{ni}^* \text{Im} G_{nn'}(E + i0^+) U_{n'i'} \times M_{j'i'}^{-1}(E) \tilde{\mathfrak{D}}_{m'j'}^*(E + i0^+). \quad (33e)$$

The important points here are that the sum over all incoming states has been reduced to a one-dimensional integral over energy, and that any resonant behavior in the integrand, i.e., sharp peaks, is contributed by the term $M^{-1}(E)$. The integral is evaluated numerically on a moderately fine mesh. If there is a resonance or other near-singular behavior of $M^{-1}(E)$ within a particular mesh interval, then energy-dependent numerators and denominators are expanded as power series in E and the integral is evaluated analytically over that interval.

We can indicate a heuristic derivation of (33) which may be useful in suggesting the relationship between this form for the density and that used in Ref. 14: The Dyson equation relating the unperturbed and perturbed Green's functions

$$G = G_0 + G_0 U G,$$

can be rewritten

$$G - G_0 = G_0 U (U - U G_0 U)^{-1} U G_0. \quad (34)$$

If we identify $G_0 U$ with $\tilde{\mathfrak{D}}$, $U G_0 U$ with D , and U with N , as is suggested by Eqs. (27), (13b), and (13a), then this relationship takes the form

$$G - G_0 = \tilde{\mathfrak{D}}(E) [N - D(E)]^{-1} \tilde{\mathfrak{D}}(E)^+. \quad (35a)$$

The change in density is given by

$$\Delta \rho(r) = -\frac{1}{\pi} \int^{E_F} dE \text{Im} [G_E(r, r) - G_E^0(r, r)]. \quad (35b)$$

Expressing G and G_0 in terms of outer-set orbitals, $N - D(E)$ in terms of inner-set (LCAO) orbitals, and $\tilde{\mathfrak{D}}(E)$ as a matrix between the two sets allows the terms in (35b) to be put into correspondence with those in (33).

The change in density used in Ref. 14 was obtained by rewriting the Dyson relationship as

$$G - G_0 = -[1 - (1 - G_0 U)^{-1}]G_0,$$

and using a complete set of localized states to evaluate all the operators involved. Quantitative comparison between the two approaches reduces to asking whether the smaller size of the matrix $[N - D(E)]^{-1}$ evaluated in the LCAO basis has cost any computational loss of accuracy relative to use of the larger size matrix $[1 - G_0(E)U]^{-1}$ in the localized basis. We believe that the greater flexibility of the LCAO basis permits equal accuracy with fewer orbitals but again, a detailed comparison has not been carried out and the question is still an open one.

Expression (33) is only the scattering state contribution to the charge density. Any charge associated with bound states whose energy lies below the Fermi energy must also be included. The total density matrix $\rho_{mm'}$ can be shown to satisfy the Kohn-Majumdar¹⁸ theorem: it is continuous as a function of impurity potential even if a bound state is split off from, or returns to, a band edge as the potential varies. This is essential for a self-consistent treatment.

C. Evaluation of the change in density of states

Callaway⁵ has shown that the change in density of states caused by a localized defect is given by

$$\Delta n(E) = -\frac{1}{\pi} \frac{d\phi(E)}{dE}, \quad (36)$$

where $\phi(E)$ is the scattering phase shift which is obtained from the determinant of $1 - G_0(E + i0^+)U$ by using the definition

$$\det\|1 - G_0(E + i0^+)U\| = R(E)e^{i\phi(E)},$$

with R and ϕ real.

We noted in Secs. IIA and IIB that the operator $1 - G_0(E)U$ in the standard formal theory⁵ and our matrix $N - D(E)$, the matrix of $U - UG_0(E)U$ with respect to the LCAO states, play similar roles in the two formulations. Bound-state energies are determined by the vanishing of their determinants; resonant behavior is determined by the sharp peak in their inverses. By analogy, we expect that $N - D(E)$ contains all information required to extract $\Delta n(E)$. A suggestive, but far from rigorous argument can be given by considering the formal operator relationship

$$\begin{aligned} \det\|U - UG_0(E + i0^+)U\| &= \det\|U\| \det\|1 - G_0U\| \\ &= \det\|U\| R(E)e^{i\phi(E)}. \end{aligned}$$

Evaluating the determinants in our LCAO basis

gives

$$\det\|N_{ij} - D_{ij}(E + i0^+)\| = \hat{R}(E)e^{i\phi(E)}, \quad (37)$$

where $\hat{R}(E) = \det\|N_{ij}\| R(E)$ is real because the impurity potential $U(r)$, being real, causes $\det\|N_{ij}\|$ to be real.

We calculate the phase shift $\phi(E)$ by using Eq. (37) and we use (36) to infer the change in the density of states.

D. Use of symmetry

Even though the introduction of a point defect destroys translational symmetry, certain point-group operations will in general still leave the crystal invariant. The largest point group which will do so depends on the type of atomic reconstruction or Jahn-Teller-type deformation around the defect. It is computationally very convenient to take advantage of these remaining symmetry properties and to use symmetrized basis functions in the various expansion sets. To create symmetrized orbitals the projection operator technique¹⁹ may be used. A symmetrized orbital is then symbolically given by

$$\phi_\mu^\Gamma(r) = \sum_R D_{\mu a}^{\Gamma*}(R) O_R \phi_i(r), \quad (38)$$

where the sum runs over all operations R of the defect point group. The symmetrized orbital is labeled by two symmetry indices, the irreducible representation Γ and the partner index μ (in case of multidimensional representations) and by a repetition index a . D represents the unitary representation matrices. Due to the systematic nature of Eq. (38) the construction of symmetrized local orbitals can most conveniently be done on a computer in analogy to the construction of symmetrized plane waves.²⁰

The merits of using symmetrized orbitals which bring the matrices in Eqs. (13), (22), and (30) into block-diagonal form are twofold: first, the computation time and storage to solve the linear systems of Eqs. (12) and (18) is drastically reduced and, second, the solutions can be identified directly in terms of their symmetry. In particular, each representation gives rise to its own contribution to the scattering phase shift and thus, its own change in the density of states. Callaway⁵ gives a full discussion of this matter.

III. DEFECT POTENTIAL

The defect potential $U(r)$ is the difference between the total potentials for the imperfect crystal and the perfect crystal. The physical approximation we use is a one-electron potential, consisting of a local pseudopotential to model the in-

teraction between the ionic cores and the valence electrons we are interested in, a Coulomb potential to describe how the valence electrons influence each other, and a local function of the valence-electron density to account for the exchange and correlation effects which are not contained in the Coulomb potential as calculated from the charge density.

Evaluation of all the terms for both perturbed and unperturbed systems and subtracting one from the other offers no special difficulty or novelty, except for construction of the Coulomb potential. We use the same ionic pseudopotential as reported in Ref. 10 and the same form of local exchange and correlation potential, X_α with $\alpha = 0.79$, so as to allow us to study how the results obtained by our new technique compare to those obtained by state-of-the-art self-consistent pseudopotential band-structure calculations when both are applied to exactly the same physical and mathematical problem.

For constructing the Coulomb potential, we have found it convenient to use fast Fourier transforms (FFT), an exceedingly efficient computer algorithm for computing discrete Fourier transforms of a variable given on an evenly spaced mesh of points. We transform $\delta\rho(r) \equiv \rho(r) - \rho_0(r)$ into Fourier space, solve Poisson's equation there, and then transform the resulting potential back into position space. Use of discrete Fourier transforms, however, implies a periodicity in real space which the true charge and potential do not have. To avoid difficulty from this source, we find it useful to apply the FFT technique only to those parts of the charge distribution which give rise to relatively short-range potentials. In this way, the potential which arises from the charge distribution (spuriously) replicated into adjacent periodic cells does not extend into the cell in which the potential is desired. We accomplish this by first approximating the actual charge distribution by a distribution which is (a) simple enough analytically that the resulting potential can be evaluated analytically and exactly and (b) has exactly the same low-order moments (charge, dipole, and quadrupole) as does the actual distribution. The FFT technique is then applied to the difference between the actual charge distribution and the approximating charge distribution. The potential which results from use of the FFT is both weak and short ranged, and, when added to the analytically evaluated part of the potential, corrects for the small difference between the true Coulomb potential and the analytically calculated one.

All three components of the defect-potential pseudopotential, Coulomb, and exchange correlation are obtained in numerical form and the needed matrix elements of U between symmet-

rized inner- and outer-set orbitals are accumulated numerically.

The loop of self-consistency is entered by taking an ad hoc defect potential, calculating a complete set of matrix elements, and evaluating a charge density as described in the previous subsections. A new defect potential is calculated using the charge density. A portion of the new potential is blended with the old and the process is repeated until input potentials and output potentials agree to within some specified limit at all points of space.

IV. DETAILS OF THE VACANCY CALCULATION

Having described the formalism and method by which the calculations were performed, we now give specific details concerning the vacancy calculation being reported here.

A. Choice of orbitals

Gaussian orbitals, of the form

$$\phi_{lm\alpha}(\vec{r}) = Y_l^m(\vec{\Omega}) r^l e^{-\alpha r^2}, \quad (39)$$

were used in both the LCAO set ϕ_i and in the outer set ϕ_n . Kane²¹ has investigated the band structure of Si using orbitals of this form and has provided data on the number of orbitals and their decay constants versus the accuracy with which the band structure from a known Hamiltonian can be reproduced using them.

1. LCAO orbitals

A portion of the Si crystal structure in the neighborhood of the vacancy is given in Fig. 1, and in

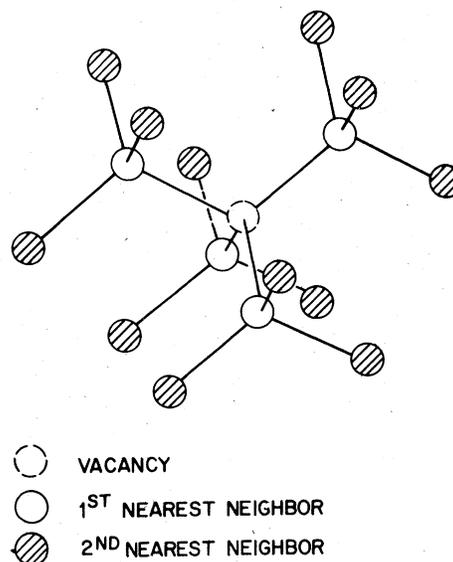


FIG. 1. Perspective view of the diamond structure with a single vacancy in "ideal" environment.

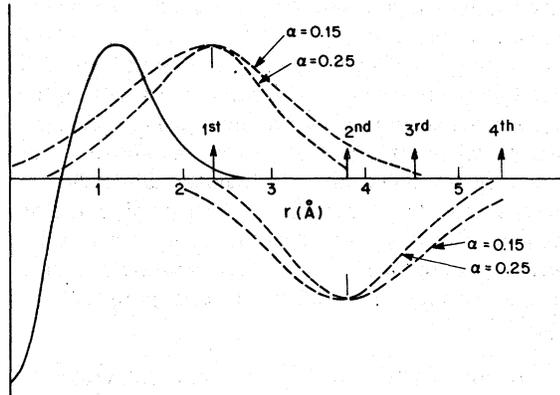


FIG. 2. Radial variation of the defect potential (full line) used in our calculations. Also indicated is the radial extent of Gaussian orbitals with various decay constants as used in the calculations.

Fig. 2 we show the spherical average of the vacancy potential. Also shown in Fig. 2 are the distances of the first- and second-nearest neighbors, and around each of these, the spread of the Gaussian wave functions of various decay constants. We have taken 11 orbitals for each atom. *s*, *p*, and *d* with $\alpha = 0.2$ a.u., and two additional *s* functions, with $\alpha = 0.15$ and $\alpha = 0.25$ a.u.. No LCAO's are needed at the vacancy site because there is no atom there, none are needed at second-nearest-neighbor sites because there is insignificant overlap with the vacancy potential there and so we are left with orbitals only on the four nearest neighbors. There are 44 orbitals in all and the size of the largest matrix $[N - D(E)]$ to be inverted, that for representation T_2 (see Sec. IV A 2) is only 8×8 .

2. Outer-set orbitals

For the present case we used the same 11 orbitals per atom for the outer-set orbitals. Because the outer-set orbitals are used to construct states which extend beyond the impurity potential, we have studied the effect of using orbitals out as far as fourth-nearest neighbors from the vacancy. The results presented here have been calculated using orbitals out as far as second-nearest neighbors, 17 sites in all. This leads to a total of 187 outer orbitals. The size of the largest density matrix ρ_{mn} , that for representation T_2 , is 29×29 .

B. Evaluation of unperturbed wave functions

The host-crystal eigenfunctions $\psi_n(k, r)$ are obtained from self-consistent pseudopotential calculations. The calculations are based on an ionic-silicon pseudopotential as used in Ref. 10, which has the following analytical form in momentum

space:

$$V_{\text{ion}}(q) = -(a_1/q^2)(4\pi e^2 Z/\Omega) \times [\cos(a_2 q) + a_3] \exp(a_4 q^4), \quad (40)$$

with the parameters $a_1 = 1.5432$, $a_2 = 0.79065$, $a_3 = -0.35201$, $a_4 = -0.01807$ (a.u.). The wave functions are expanded in plane waves, i.e.,

$$\psi_n(k, r) = \sum_G C_n(k, G) e^{i(\vec{k} + \vec{G}) \cdot \vec{r}}, \quad (41)$$

with the sum running over reciprocal-lattice vectors G with $|G|^2 \leq 6$ Ry. We choose this particular representation, since it yields rather accurate potentials and wave functions for semiconductors. With the choice of Gaussians and plane waves, all integrals needed for the evaluation of the $B_m(n, k)$ coefficients can be evaluated analytically. Wave functions were evaluated on a mesh of 70 points per $\frac{1}{48}$ th of BZ, energies on a grid of 203 points per $\frac{1}{48}$ th of BZ.

C. Evaluation of the Green's function

The Green's function $G_E(r, r')$ is evaluated in a spectral representation as given by Eq. (4). For this purpose the complex function is separated into its real and imaginary parts. The imaginary part may be written

$$\text{Im}G_E(r, r') = -\pi \sum_n \int d\vec{k} \psi_n(\vec{k}, r') \psi_n(\vec{k}, r) \times \delta(E - E_n(\vec{k})). \quad (42)$$

The real part is then obtained by Kramers-Kronig transformation:

$$\text{Re}G_E(r, r') = \frac{1}{\pi} \text{P} \int dE' \frac{\text{Im}G_{E'}(r, r')}{E' - E}. \quad (43)$$

Since $G_E(r, r')$ is a function of r and r' the actual evaluation is not done for $G_E(r, r')$ but for the individual matrix elements $G_{mm'}(E)$ of Eq. (22).

First the imaginary part of the integrand of $G_{mm'}(E)$ is evaluated on the grid of 70 k points in the irreducible part of the Brillouin zone of the host crystal for which the wave functions and the $B_m(nk)$ coefficients had been evaluated. The size of this grid was dictated by the phase fluctuations in the wave-function product in Eq. (4) at the maximum distance $r - r'$ of interest. At the same time, a reasonable number of valence and conduction bands have to be included to warrant sufficient convergence of $G_E(r, r')$ for the energy region of interest. The matrix elements $M_{mn}(\vec{k}, n)$ defined by

$$\text{Im}G_{mm'}(E) = \sum_n \int_{\text{BZ}} d^3k M_{mn}(\vec{k}, n) \delta(E - E_n(\vec{k})) \quad (44)$$

are then fitted, band by band and element by element, to periodic Fourier expansions of the form

$$M_{mm'}(\vec{k}, n) = \sum_{\lambda} M_{\lambda} e^{i\vec{k} \cdot \vec{R}_{\lambda}}, \quad (45)$$

where the sum runs over real lattice vectors \vec{R}_{λ} . There are actually far fewer coefficients M_{λ} than there are lattice vectors R_{λ} because $M_{mm'}$ is symmetric as a function of \vec{k} under all the point-group operations and we utilize this symmetry to the fullest extent. With the aid of this expansion the matrix elements $M_{mm'}(\vec{k}, n)$ are now easily evaluated at a larger number of grid points (i.e., 203 points per $\frac{1}{48}$ th of BZ). The integration in k space in Eq. (44) is then performed using the Gilat-Kam¹⁷ method which requires linear expansions of $E_n(\vec{k})$ and $M_{mm'}(\vec{k}, n)$. This procedure yields highly converged matrix element $\text{Im}G_{mm'}(E)$ as a function of energy. It consists essentially of calculating density of states functions weighted by some k -dependent functions.

In a next step a Kramers-Kronig transformation of each matrix element yields the real part of the Green's-function matrix elements. We have experimented with up to 20 bands in the Green's function. The present results were computed using 15 bands, 4 valence and 11 conduction.

D. Self-consistency

After the one-time evaluation of the Green's-function matrix [Eq. (23)], charge-density and potential iterations were carried out. Each complete iteration involves evaluating matrix elements of the defect potential $U(r)$, multiplications to produce $D(E)$ and $\mathfrak{D}(E)$ [Eqs. (25) and (27)], inversion of $N - D(E)$ [Eq. (31)], multiplication to obtain $\rho^{(1)}$, $\rho^{(2)}$, and $\rho^{(3)}$ [Eqs. (33c)–(33e)], and construction of the defect potential. This can typically be done at a small fraction, one-fifth of the cost of a supercell calculation.¹⁰ A self-consistent solution was assumed to be reached after input and output potentials agreed on average to 0.016 eV. The largest potential fluctuation occurring at the origin was 0.36 eV.

V. RESULTS FOR THE IDEAL Si VACANCY

As a first structural model we have studied the neutral Si lattice vacancy in its "ideal" structure. In this structure, the atoms surrounding the vacancy site remain in their bulk crystalline positions after the vacancy is created and the structure retains T_d point-group symmetry. A portion of the Si crystal structure with the vacancy is shown in Fig. 1. Every Si atom is tetrahedrally coordinated and the valence electrons form covalent bonds linking the neighboring atoms. As a result of cre-

ating a vacancy, four bonds are broken. Four valence electrons which previously participated in the broken bonds are removed together with the neutralizing Si^{4+} core. The atoms around the neutral vacancy now experience a different potential and localized electron states may occur. In fact one may view the vacancy as a small internal surface and the four surrounding atoms as surface atoms. In the simplest conceptual model, one may assume that only electrons of these four "surface" atoms are perturbed. This amounts to a total of $4 \times 4 = 16$ s and p localized electronic states which might combine to form bound states or resonances in the valence electron spectrum. Whether these 16 states are sufficient, or whether more states involving second- and third-nearest-neighbor atoms are actually needed, is determined self-consistently by the extent of the defect potential.

The defect point group T_d has five different representations and thus allows for five different symmetry types of defect states. The character table of T_d is shown in Table I together with atomic transformation properties for s , p and d functions centered at the origin, i.e., at the vacancy site. In Table II we list the symmetry decomposition of s -, p -, and d -like basis functions centered at the vacancy site, and at the first-, second-, and third-nearest-neighbor shells. One can, e.g., see that s - and p -like functions centered at the first-neighbor sites induce $(A_1) \Gamma_1$, Γ_3 , Γ_4 , and $(T_2) \Gamma_5$ representations. Some of the states formed by linear combination of these functions are bonding combinations and should be found among the valence bands as resonances or true bound states. Some of them are antibonding and will be located higher up in the conduction bands.

Another simple conceptual model, but one which is somewhat easier to reason about, is obtained by considering bonding orbitals rather than s -, p -, and d -like atomic functions. In a simple four atom approximation, the four dangling bonds pointing towards the vacancy induce the $(A_1) \Gamma_1$ and the $(T_2) \Gamma_5$ representations. If one includes the back bonds (or dehybridization), additional Γ_3 and Γ_4 contributions appear and the situation is equivalent to the 16 orbital model mentioned above. Since the dangling bonds, not the back bonds, should be most af-

TABLE I. Character table of the point group T_d .

T_d	E	$8C_3$	$3C_2$	$6S_4$	$6\sigma_d$	Basis function
$(A_1) \Gamma_1$	1	1	1	1	1	s
$(A_2) \Gamma_2$	1	1	1	-1	-1	
$(E) \Gamma_3$	2	-1	2	0	0	$d(y^2 - z^2, 3x^2 - r^2)$
$(T_1) \Gamma_4$	3	0	-1	1	-1	
$(T_2) \Gamma_5$	3	0	-1	-1	1	$p, d(xy, yz, zx)$

TABLE II. Symmetry decomposition according to the irreducible representation of T_d of local atomic orbitals centered at the vacancy site, the first-, second-, and third-nearest-neighbor shells of the diamond structure.

Vacancy	<i>s</i>	Γ_1
	<i>p</i>	Γ_5
	<i>d</i>	$\Gamma_3 + \Gamma_5$
First neighbor	<i>s</i>	$\Gamma_1 + \Gamma_5$
	<i>p</i>	$\Gamma_1 + \Gamma_3 + \Gamma_4 + 2\Gamma_5$
	<i>d</i>	$\Gamma_1 + 2\Gamma_3 + 2\Gamma_4 + 3\Gamma_5$
Second neighbor	<i>s</i>	$\Gamma_1 + \Gamma_3 + \Gamma_4 + 2\Gamma_5$
	<i>p</i>	$2\Gamma_1 + \Gamma_2 + 3\Gamma_3 + 4\Gamma_4 + 5\Gamma_5$
	<i>d</i>	$3\Gamma_1 + 2\Gamma_2 + 5\Gamma_3 + 7\Gamma_4 + 8\Gamma_5$
Third neighbor	<i>s</i>	$\Gamma_1 + \Gamma_3 + \Gamma_4 + 2\Gamma_5$
	<i>p</i>	$2\Gamma_1 + \Gamma_2 + 3\Gamma_3 + 4\Gamma_4 + 5\Gamma_5$
	<i>d</i>	$3\Gamma_1 + 2\Gamma_2 + 5\Gamma_3 + 7\Gamma_4 + 8\Gamma_5$

ected by creation of the vacancy, one expects the most important changes in the density of states to occur in the A_1 and T_2 representations.

In Fig. 3 we display the calculated bulk electron density of states for silicon and on the same energy scale, the A_1 and T_2 phase shifts induced by the self-consistent vacancy potential. The phase shifts have been calculated according to Eq. (37) and describe the energy distribution of changes induced in the density of states via Eq. (36). Note that the integral of Eq. (36)

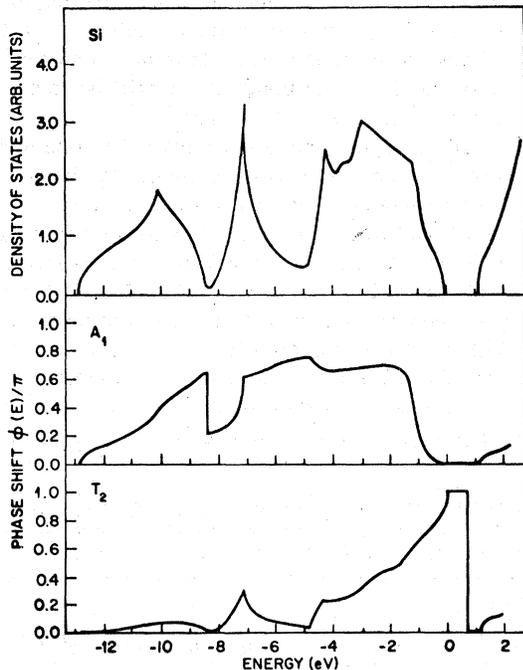


FIG. 3. Bulk density of states of Si and the A_1 and T_2 phase shifts induced by a neutral isolated vacancy.

$$\int_{E_1}^{E_2} \Delta n(E) dE = -\frac{1}{\pi} [\phi(E_2) - \phi(E_1)]$$

shows that the number of states gained (or lost) in the energy interval $E_1 < E < E_2$ is equal to π^{-1} times the difference in phase shift at the energy limits of the interval. Since the imaginary part of the Green's function vanishes in the band gaps, the phase $\phi(E)$ must be a multiple of π for any energy in the gap and thus the total number of states gained by an entire band (the integral above with E_2 in the gap above the band and E_1 in the gap below it) must be an integer, positive, negative, or zero. The same statement applies to states of a particular symmetry type,⁵ provided that $\phi(E)$ is calculated using that block of $N - D(E)$ appropriate to the particular symmetry, which is how the A_1 and T_2 phase shifts are calculated.

From Fig. 3 it follows that for the A_1 symmetry, the neutral ideal vacancy has neither gained nor lost a state from its valence band. This is in contrast to the T_2 symmetry, for which a net phase change of π is found for the valence bands, indicating the loss of one threefold-degenerate state. This state is recovered as a true bound state about 0.7 eV above the valence-band edge. The value is somewhat higher than the 0.5 eV extracted by Louie et al.¹⁰ as the center of gravity of the "super-cell" defect band and it is very close to the 0.8 eV calculated by Bernholc et al.¹⁴ in their self-consistent scattering-type calculation.

To obtain a neutral vacancy, we have assumed that each of the threefold-degenerate T_2 states in the gap is filled to one-third occupancy, resulting in a system which is both electrically neutral and invariant under T_d . Such occupancy is unstable with respect to symmetry-lowering Jahn-Teller distortions⁶ and so direct comparison with experiment is not yet appropriate. However, the position of the boundstate at 0.7 eV is consistent with experimental observations and their interpretations according to which there is a symmetry-lowering Jahn-Teller distortion which lowers the band state by about 1 eV,²² and according to which the bound-state energy level of the neutral vacancy is found near the top of the valence band.

The overall shape of the phase shifts displayed in Fig. 3 is qualitatively similar to the non-self-consistent tight-binding results of Kauffer et al.⁷ and in nearly quantitative agreement with the results of Bernholc et al.¹⁴

As seen by comparison with the density of states, structure in the phase shifts is usually correlated with structure in the density of states. This correlation arises because the imaginary part of a Green's-function matrix element is the product of the density of states of the appropriate symmetry

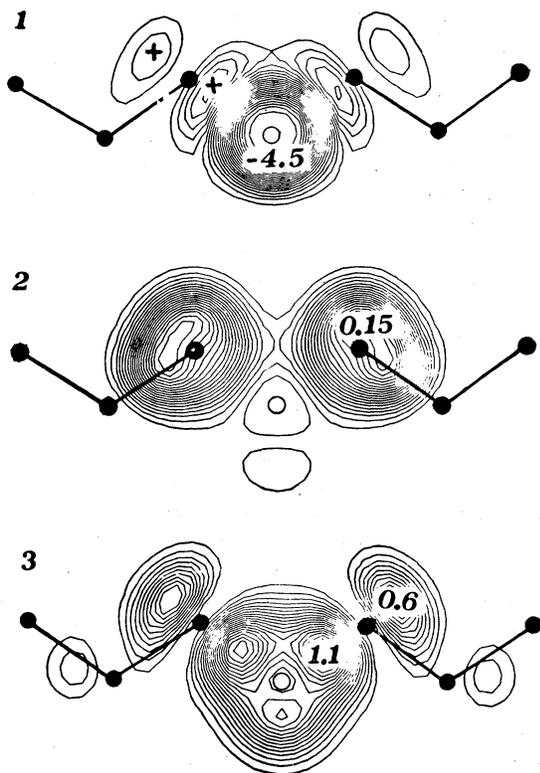


FIG. 4. Charge-density contour plots displayed in a (110) plane of atoms containing one vacancy. The total integrated change in charge density of A_1 symmetry (1) is compared to charge-density fluctuations associated with the A_1 resonance at -8.4 eV (2) and the A_1 resonance at -1.1 eV (3). The values are given in units of electrons per Si bulk unit cell.

(a function with sharp structural features) and a rather smooth function of energy. This structure influences the phase of the determinant of $N - D(E)$. The most striking example of this correlation is a low-energy A_1 resonance, which occurs at the minimum in the density of states of virtually every calculation of the ideal vacancy in diamond, Si, or Ge^{5, 7, 14, 23, 24} regardless of how much these various calculations differ among themselves in the placement of the bound state. The bound-state energy is, like the strength of the resonance (the phase change across the resonance) more sensitive to the strength of the vacancy potential.

Our A_1 phase shift here is dominated by two strong resonances. A sharp A_1 resonance at -8.4 eV has a strength of 0.4 states and a width of 0.1 eV. A broader A_1 resonance having a strength of about 0.6 states and a width of about 0.6 eV is located at about -1.1 eV. These two resonances correspond to A_1 -symmetry combinations of mostly- s -like (-8.4 eV) and mostly- p -like (-1.1 eV) orbitals centered at the four nearest-neighbor atoms. This behavior can clearly be seen in real-

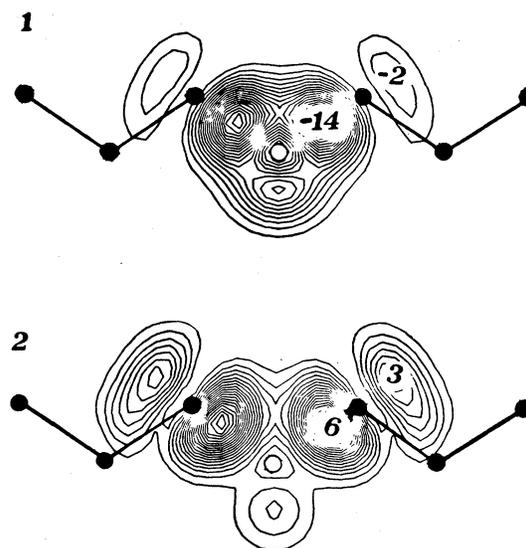


FIG. 5. Contour plots of charge fluctuations of T_2 symmetry induced over the total valence-band region (1) and for the bound state at 0.7 eV (2). Units as in Fig. 4.

space charge plots as shown in Fig. 4. Here the total A_1 -type valence-charge perturbation (1) is shown together with the -8.4 eV resonance distribution (2) and the -1.1 eV resonance distribution (3). The densities are given as contours in a (110) plane containing parts of a silicon chain and one vacancy. While the total A_1 perturbation yields zero net charge (i.e., no loss or gain of any states) and thus contains both negative and positive charge fluctuations, both resonances (2) and (3) contain only positive charge fluctuations whose integral over all space must equal 0.4 and 0.6 states, respectively. These numbers are obtained from the phase-shift analysis. Direct real-space integration of the charge fluctuations in Fig. 4 is limited by the spatial extent of our outer orbitals. For the two resonances, only 50% of the correct charge is contained in our "outer-set" basis set. The remaining charge extends beyond the range of our basis, i.e., beyond second-nearest neighbors. However, since the self-consistent procedure is solely based on the potential and therefore on the total charge perturbation, the incomplete description of individual states is without consequence.

Similar arguments apply to the T_2 symmetry. In Fig. 5 the integrated valence-band charge perturbation of T_2 symmetry (1) is compared to the distribution of the (totally) occupied bound state (2). While phase-shift arguments predict integrated net charges of -3 and $+3$ states, real-space integration yields only 90% and 60% of these values, respectively. The trend towards higher localization for the total charge perturbation can be seen

by the opposite sign of long-range charge fluctuations in the scattering-band charge (1) and bound-state charge (2), respectively. The charge distributions displayed in Figs. 4 and 5 combined with the phase-shift curves of Fig. 3 give us an intuitive picture of the main changes in the electronic structure of Si induced by a neutral vacancy. While the simple dangling-hybrid argument, given before would predict the occurrence of one A_1 and one T_2 feature, dehybridization effects (or equivalently the involvement of back bonds) modifies that picture slightly. There is a T_2 and an A_1 combination of mostly- p -like states in the vicinity of the gap (i.e., at 0.7 and -1.1 eV), the T_2 state lying above the A_1 state because of its larger kinetic energy, and there is an other A_1 combination of s -like states at -8.4 eV. The corresponding T_2 combination of s -like states does not give rise to a single sharp resonance.

No significant phase-shift structures appear for the representations Γ_3 and Γ_4 which involve mainly states at larger distance from the vacancy and beyond the effective range of the potential. The induced small charge fluctuations, however, are included in the self-consistent procedure; their real-space distributions are shown in Fig. 6. No Γ_2 -type charge perturbation exists for the present calculations which, because of the short range of the defect potential, involve s -, p -, and d -type "inner-set" LCAO's at only the four nearest-neighbor atoms (see Table II).

If the calculated charge perturbances of the various representations are added to produce a total charge disturbance at the vacancy, a rather short-

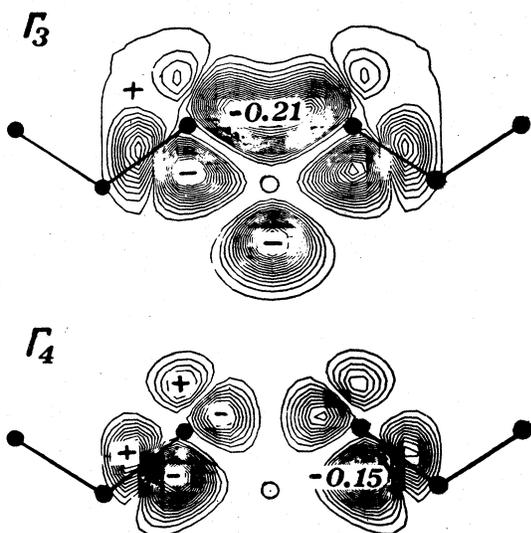


FIG. 6. Contour plots of charge fluctuation of Γ_3 (top) and Γ_4 (bottom) symmetry integrated over all valence bands. Units as in Fig. 4.

range charge density is obtained (see Fig. 7, panel 3), which essentially is confined to a cavity terminated by the nearest-neighbor atoms. This is in contrast, as we have seen above, to the behavior of individual states which can be quite extended. This behavior of fast "healing" of a perturbation has also been found for semiconductor surface systems¹² and is one of the fundamental concepts underlying the present formalism. The same range of charge disturbance, of the order of a bond length, is also found in calculations of the spatially dependent static dielectric function using various simple models of the valence charge of a semiconductor.²⁵

Also shown in Fig. 7 are contour plots of the total unperturbed silicon charge (1) and the total charge in the presence of a vacancy (2). As for the charge perturbances, shown before, the displays have been prepared by expanding the Bloch waves in an infinite set of orbitals $\Phi_m(r)$ and retaining only those orbitals ("outer set") which influence the density in the neighborhood of the de-

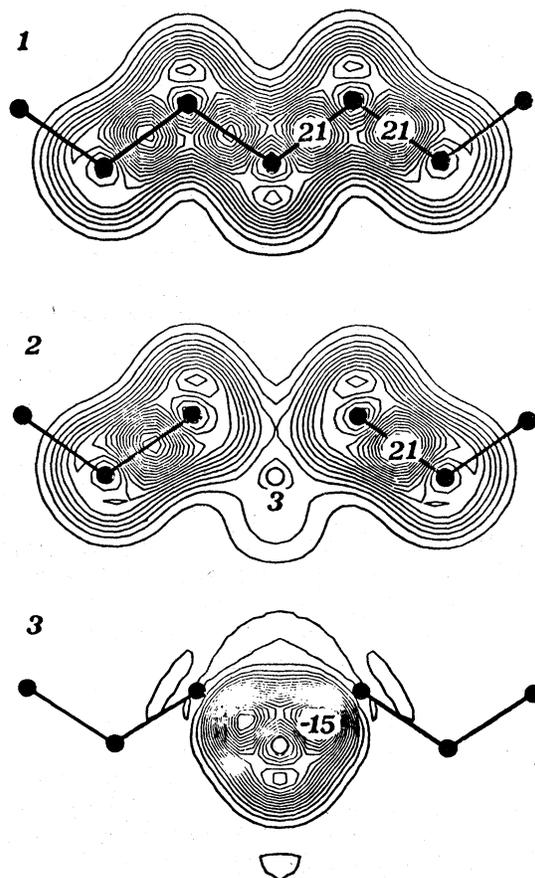


FIG. 7. Contour plots of the unperturbed (1), perturbed (2) and total change (3) of charge densities induced by a silicon vacancy. Units as in Fig. 4.

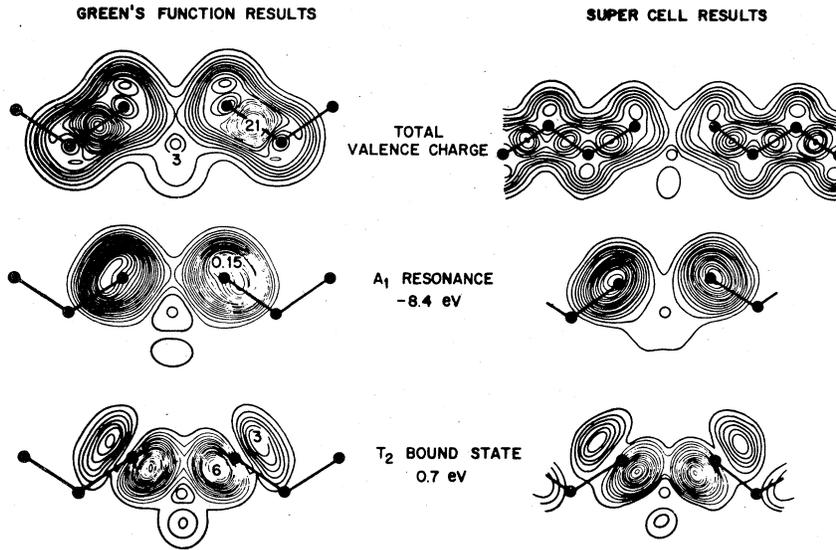


FIG. 8. Comparison of charge densities calculated in the present work with some "supercell" results of Ref. 10. The total perturbed charge density (top), the lower (-8.4 eV) A_1 resonance (middle) and the T_2 bound state (bottom) are shown. Units as in Fig. 4.

fect. No information is lost in this way but the display lacks the translational periodicity one would expect to see.

The total valence charge in the presence of the vacancy (Fig. 7, panel 2) clearly shows the disappearance of bonding charge with the removal of an atom. This situation is very reminiscent of that of the (111) surface¹² which also showed a "smeared out" and rounded charge distribution which decays rapidly into "vacuum." The charge densities calculated here agree closely with those calculated in Ref. 10 using a supercell-band-structure technique (see Fig. 8). It is clear to us that the supercells chosen in that work were large enough to isolate the periodic defect potentials but not large enough to remove the overlap of individual bound states and scattering resonances, which can be more extended than the potential.

In Fig. 9 we show the spherical average of various components of the calculated defect potential. While the present method does not use any spherical approximation to the potential, angular anisotropy is small ($\sim 10\%$) and the spherical average does contain all the physically significant features. The removal of one (pseudo-) Si^{4+} ion is described by the spherically symmetric potential $-V_{\text{ion}}$ according to Eq. (40). The self-consistent arrangement of four missing electrons yields a Hartree potential V_H which compensates the long-range Coulomb tail of $-V_{\text{ion}}$. A repulsive effective exchange-correlation potential V_{xc} results as the difference between exchange-correlation potentials with and without vacancy. The total defect perturbation potential V_{tot} is of very short range as can be seen from Fig. 9. Due to the dielectric screening of the silicon host crystal V_{tot} is of

shorter range than the corresponding (negative) silicon-atom pseudopotential.

Also indicated in Fig. 9 is the total defect potential of the "supercell" calculation of Ref. 10, where the potential was evaluated using a relatively low cutoff (2.8 a.u.) of Fourier components. This introduced spurious oscillations into the "supercell" ionic potential. Aside from the oscillatory differences due to this low cutoff, the two calculated defect potentials agree closely.

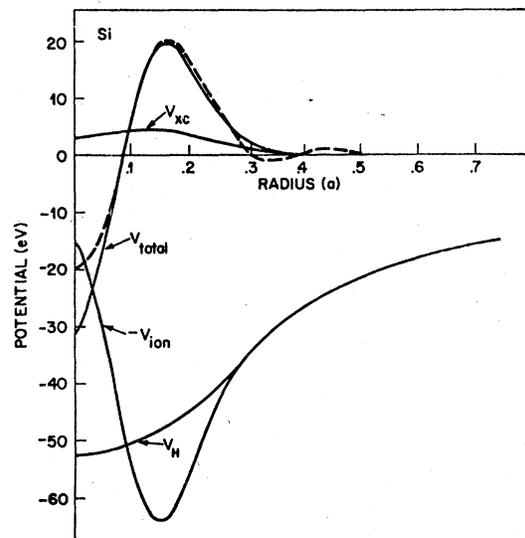


FIG. 9. Spherical average of components of the self-consistent vacancy potential and the total potential of the "supercell" calculation (Ref. 10, dashed line). Angular anisotropy of the potentials is generally smaller than $\pm 10\%$.

VI. CONCLUSION

In this paper, we have tried to do three things. First, we have described a new method for performing quantum-mechanical calculations of the self-consistent potential, charge density, electronic wave functions, and energies associated with an isolated point defect. Second, we have tried to indicate how it is related to and differs from the work of Bernholc, Lipari, and Pantelides¹⁴ which was being pursued independently at the same time as our method was being developed. Third, we have illustrated our method by applying it to a well-studied problem, the ideal neutral vacancy in silicon. This application shows that this Green's-function technique yields self-consistent densities and potentials which agree with those calculated self-consistently via a bulk-band-structure calculation of a periodic vacancy array. Its

real strength, however, is that it produces bound-state energies, resonance energies, and widths that are more reliable than, or inaccessible to, the periodic-array technique and at less cost. The type of information and the degree of accuracy made available by this new technique will be needed to study the effects of lattice distortions and their coupling to the electronic system.

ACKNOWLEDGMENTS

We should like to thank E. O. Kane for a long series of useful discussions during the course of this work. In particular, his work on the use of Gaussian wave functions for quantitative band-structure calculations strongly influenced our thinking in the early stages of our development of the method reported here.

-
- ¹F. Bassani and G. Pastori-Parravicini, *Electron States and Optical Transitions in Solids*, edited by R. A. Ballinger (Pergamon, Oxford, 1975), and references therein.
- ²F. Bassani, G. Iadonisi, and B. Preziosi, *Rep. Prog. Phys.* **37**, 1099 (1974), and references therein.
- ³See, for example, the recent review article by S. T. Pantelides, *Rev. Mod. Phys.* (1978).
- ⁴G. F. Koster and J. C. Slater, *Phys. Rev.* **96**, 1208 (1954).
- ⁵J. Callaway, *J. Math. Phys.* **5**, 783 (1964); J. Callaway and A. J. Hughes, *Phys. Rev.* **156**, 860 (1967); J. Callaway, *Phys. Rev. B* **3**, 2556 (1971).
- ⁶See, e.g., R. P. Messmer and G. D. Watkins, in *Radiation Damage and Defects in Semiconductors* (Institute of Physics, London, 1972), No. 16, p. 255, and references therein.
- ⁷F. P. Larkins and A. M. Stoneham, *J. Phys. C* **4**, 143 (1971); or E. Kauffer, P. Pecheur, and M. Gerl, *ibid.* **9**, 2319 (1976); or William Y. Hsu *et al.*, *Phys. Rev. B* **16**, 1597 (1977).
- ⁸M. Jaros and S. Brand, *Phys. Rev. B* **14**, 4494 (1976), references therein.
- ⁹M. Jaros (private communication).
- ¹⁰S. G. Louie, M. Schlüter, J. R. Chelikowsky, and Marvin L. Cohen, *Phys. Rev.* **13**, 1654 (1976).
- ¹¹J. A. Appelbaum and D. R. Hamann, *Rev. Mod. Phys.* **48**, 479 (1976).
- ¹²M. Schlüter, J. R. Chelikowsky, S. G. Louie, and Marvin L. Cohen, *Phys. Rev. B* **12**, 4200 (1975).
- ¹³G. A. Baraff and M. Schlüter, *Phys. Rev. Lett.* **41**, 892 (1978).
- ¹⁴J. Bernholc, N. Lipari, and S. T. Pantelides, *Phys. Rev. Lett.* **41**, 895 (1978).
- ¹⁵G. D. Watkins and R. P. Messmer, *Phys. Rev. Lett.* **32**, 1244 (1974), and references therein.
- ¹⁶W. Kohn and N. Rostoker, *Phys. Rev.* **94**, 1111 (1954).
- ¹⁷G. Gilat and Z. Kam, *Phys. Rev. Lett.* **22**, 715 (1969).
- ¹⁸W. Kohn and C. Majumdar, *Phys. Rev.* **138**, A1617 (1965).
- ¹⁹J. C. Slater, *Symmetry and Energy Bands in Crystals* (Dover, New York, 1972).
- ²⁰B. Renaud and M. Schlüter, *Helv. Phys. Acta* **45**, 66 (1972).
- ²¹E. O. Kane, *Phys. Rev. B* **13**, 3478 (1976).
- ²²G. D. Watkins, *Inst. Phys. Conf. Ser. No. 23*, 1 (1975).
- ²³J. D. Joannopoulos and Eugene J. Mele, *Solid State Commun.* **20**, 729 (1976).
- ²⁴M. Lannoo and P. Lenglar, *J. Phys. Chem. Solids* **30**, 2409 (1969).
- ²⁵R. Resta, *Phys. Rev. B* **16**, 2717 (1977); and G. Srinivasan, *Phys. Rev.* **178**, 1244 (1969).