

Quasiband crystal-field method for calculating the electronic structure of localized defects in solids

Ulf Lindelfelt and Alex Zunger

Solar Energy Research Institute, Golden, Colorado 80401

(Received 30 September 1981)

We introduce an *ab initio* self-consistent approach—the quasiband crystal-field (QBCF) method—to calculate the electronic structure of localized defect states in solids within a density-functional Green's-function approach. The method is simple, yet it produces very accurate self-consistent solutions both for *s-p* as well as for the hyperlocalized transition-atom *d*-electron impurities. This is made possible by four ideas: (1) Following the pioneering work (1929) of Bethe and Van Vleck and results of modern computations, it is recognized that whereas the defect and host wave functions may be extended and highly anisotropic in coordinate space, for deep defects the density and potential perturbations $\Delta\rho(\vec{r})$ and $\Delta V(\vec{r})$ are considerably more localized and have a reduced directional anisotropy. We therefore describe the latter in a crystal-field one-center expansion $\sum_l F_l(|\vec{r}|)K_l(\hat{r})$, anchored at the defect site, with a separation of radial $F_l(|\vec{r}|)$ and angular $K_l(\hat{r})$ variables. The defect problem, treated by contemporary Green's-function techniques as a multicenter scattering problem, is then transformed into a far simpler atomic-like problem, characterized by analytic angular integrals (Gaunt coefficients) and simple one-dimensional radial integrals. This permits a simple and highly precise treatment of self-consistency, incorporation of accurate (first-principles) nonlocal pseudopotentials, and the use of variationally flexible and computationally simple single-site basis functions introduced in chemistry in 1933 by Mulliken. (2) The standard Koster-Slater Green's-function approach to defects uses an expansion of the impurity wave functions in terms of (often chemically and physically unrelated) host-crystal Bloch eigenfunctions. It is shown that when the perturbation approaches a characteristic atomic length scale (or when the impurity is chemically sufficiently different from the host atom), such expansions converge exceedingly slowly. We have reformulated the Koster-Slater resolvent problem in terms of quasiband wave functions that incorporate from the outset not only aspects of the host, but also the characteristics of the defect. A large number of conduction-band wave functions, which would have been needed for an adequate representation of localized defect states, are renormalized into a much smaller number of quasiband wave functions. Expansion in terms of quasibands results in a rapidly convergent and efficient description even of very localized defect wave functions. (3) A new Newton-Raphson Jacobian update technique is used to establish self-consistency in the screening potential. It does not require any new information; it "remembers" information from all past iterations, but automatically discounts information from the distant past and is hence not confused by nonlinearities. The method is far more efficient than all standard self-consistency methods and permits a precise assessment of charge-redistribution effects in the system. (4) The lengthy summations over the Brillouin zone encountered in spectral Green's-function methods are transformed into a simpler rapidly convergent series in a supercell representation. This allows one to treat impurities in large supercells (e.g., 2662 atoms per cell) by treating only small matrices (36×36) whose sizes do not depend on the dimensions of the supercell. In this representation the poles of the Green's function are real and can be efficiently located using a new and fast algorithm introduced here. This paper describes these four ideas in physical terms. Full mathematical details are given in a series of Appendices that can be used by the reader to independently reproduce the method. The method is applied to study the electronic structure of the unrelaxed silicon vacancy as well as to the far more difficult problem of a substitutional transition-metal impurity in silicon. The electronic structure of Cu in an extended crystal is described for the first time.

I. INTRODUCTION

Substantial progress has been made over the past few years in advancing the one-electron theory of localized defects in solids to a similar level of sophistication with which contemporary electronic band-structure models describe the periodic states of the host crystals. This has been accomplished by applying the classical Koster-Slater resolvent method^{1,2} in a Bloch representation, using local pseudopotentials.^{3,4}

These methods were very successful in treating the electronic structure of defects whose potential perturbation $\Delta V(\vec{r})$ has a characteristic size of either many or few lattice constants (shallow effective-mass defects or *s-p* electron impurities, respectively). It would be of interest to inquire whether the same approaches could be used for defects whose characteristic length scale approaches that of an isolated atomic core, (e.g., the technologically important transition-atom impurities in semiconductors). We find that at this limit, not treated before by such methods, the conventional Green's-function approaches pose at least four major difficulties. The problem becomes highly intractable and would involve exceedingly complex and time consuming computational procedures, accessible only through large high-speed computers. We analyze the sources of these difficulties and find that they mostly result from the fact that the conventional Green's-function approaches have extended to the study of *localized* perturbations many of the concepts used in the past to treat the *extended* states of ideal crystals [e.g., expansion of perturbed wave functions in terms of eigenstates of the ideal host crystal, a multicentered linear combination of atomic orbitals (LCAO) description of wave functions and potentials, etc.]. In this paper we present an approach which overcomes these difficulties in a simple and effective manner. This approach can be used both for intermediate-radius *s-p* defects (e.g., vacancies, or main-row impurities), as well as for hyperlocalized defects (e.g., transition-atom impurities and core excitons). We will describe a new *ab initio* self-consistent approach to localized defects, based on the density-functional formalism. Our objective is to provide a method that will consistently reflect the predictions of the density-functional formalism (rather than those of the computational approximations to it), to a precision of about 0.1 eV in energies and about 2% in charge densities. The method includes four ideas that simplify the problem considerably, yet it produces extremely precise self-consistent solutions.

These are: (1) A Green's-function generalization of the classical crystal-field and ligand-field methods. This reduces the defect problem treated by contemporary Green's-function methods^{3,4} as a complex *multicenter scattering problem* (requiring $\sim 10^6$ multicenter integrals) into a computationally simpler and physically far more transparent one-center *atomiclike* problem, described by simple one-dimensional integrals. This is made possible by rigorously separating the multicenter problem into (i) an analytical (nonspherical) angular part describing the directional anisotropy and (ii) a one-dimensional radial part treated numerically. This permits a very precise, and at the same time, computationally simple handling of self-consistency, pseudopotential nonlocality, and basis set variational flexibility. (2) Rather than expanding the localized defect wave functions in terms of the extended host-crystal eigenfunctions (requiring $10^2 - 10^4$ host bands for hyperlocalized defects), we describe the former in terms of "quasiband wave functions" that incorporate from one outset not only the characteristics of the host but also the localized chemical aspects of the impurity. This results in a rapidly convergent (requiring 1–10 quasibands) and physically transparent representation of the defect states. (3) The complicated \vec{k} -space summations (200–300 \vec{k} points) involved in Green's-function models in the spectral representation are transformed into rapidly convergent (10–30 \vec{k} points) summations using a supercell representation. Large supercells (e.g., 2662 atoms) are then easily treated by handling only very small matrices. (4) The slowly convergent self-consistency procedure used for calculating the system's screening potential (20–40 iterations using mixing of input and output potentials and an atomic initial guess) is accelerated substantially ($\sim 3 - 7$ iterations) by using a new and simple Newton-Raphson Jacobian update technique, introduced by Bendt and Zunger.⁵ No human intervention is required to achieve self-consistency.

The main body of this paper emphasizes the new physical ideas, rather than detailed mathematical formulation, leading to our simple solution to the problem. It analyzes the fundamental difficulties encountered in a frontal attack on the problem using Green's-function methods and points at the same time to the physical simplicity of the classical Bethe–Van Vleck crystal-field approach to the problem, largely overlooked by modern computational methods. It then offers a synthesis of the two approaches that leads to a simple and accurate method. Mathematical details are presented in

their entirety in a series of Appendices that completely characterize the method; given this information, we believe that the results of the quasiband crystal-field (QBCF) method can be fully reproduced by others. The method is illustrated by applications to the unrelaxed silicon vacancy (characteristic of intermediate-radius s - p defects), as well as to the Cu impurity in silicon (characteristic to small-radius hyperlocalized defects). Application to all $3d$ impurities in silicon will be presented in a future publication.

II. STATEMENT OF THE PROBLEM

Our description of the problem addressed in this paper consists of (1) a formulation of the host- (H) crystal problem, (2) the definition of the potential perturbation associated with the defected (D) crystal, and (3) the statement of the defect problem. In this section we define the basic quantities appearing in the forthcoming discussions and in the presentation of our method.

The electronic structure of the ideal host is expressed in terms of a single-particle equation that provides the host wave functions $\phi_j^0(\vec{k}, \vec{r})$ and band energies $\epsilon_j^0(\vec{k})$ for a translationally invariant host-crystal potential $V_H(\rho_H(\vec{r}))$:

$$\left[-\frac{1}{2}\nabla^2 + V_H(\rho_H(\vec{r}))\right]\phi_j^0(\vec{k}, \vec{r}) = \epsilon_j^0(\vec{k})\phi_j^0(\vec{k}, \vec{r}). \quad (1)$$

Here the host bands form both an orthogonal

$$\int \phi_j^{0*}(\vec{k}, \vec{r})\phi_j^0(\vec{k}', \vec{r})d\vec{r} = \delta_{jj'}\delta_{\vec{k}\vec{k}'}, \quad (2)$$

and a complete set

$$\sum_j \int_{\text{BZ}} \phi_j^0(\vec{k}, \vec{r})\phi_j^{0*}(\vec{k}, \vec{r}')d\vec{k} = \delta^3(\vec{r} - \vec{r}'), \quad (3)$$

where its occupied (occ) portion in the first Brillouin zone (BZ) forms the basis for constructing the ground-state host charge density $\rho_H(\vec{r})$ using the band occupation numbers $N_j(\vec{k})$:

$$\rho_H(\vec{r}) = \sum_j \sum_{\vec{k}}^{\text{occ}} N_j(\vec{k}) |\phi_j^0(\vec{k}, \vec{r})|^2. \quad (4)$$

The crystal potential has two contributions: a static external field $V_H^{\text{ext}}(\vec{r})$ that, with the number of electrons N_e^H , defines the physical system, and a screening potential $V_H^{\text{scr}}(\vec{r})$ that describes the response of the N_e^H electrons to this external field. The external potential $V_H^{\text{ext}}(\vec{r})$ is specified in terms of the lattice vectors $\{\vec{R}_p\}$ of the P unit cells at \vec{R}_p , the site vectors $\{\vec{\tau}_\beta\}$ of the Q atoms β per primitive cell, and their chemical identity in terms

of the atomic numbers $\{Z_\beta\}$. The screening potential $V_H^{\text{scr}}(\vec{r})$ is defined in terms of the electron density $\rho_H(\vec{r})$ containing N_e^H electrons in the neutral ground state. The translationally invariant total host potential is hence given by

$$V_H(\vec{r}) = V_H^{\text{ext}}(\vec{r}; \{\vec{R}_p, \vec{\tau}_\beta\}; \{Z_\beta\}) + V_H^{\text{scr}}(\rho_H(\vec{r})). \quad (5)$$

A pseudopotential representation is used here to describe the external potential. In this form the external potential is the sum of a core attraction term $-Z_\beta^v/r$ due only to the valence (v) charge Z_β^v and to an angular-momentum (L) dependent pseudizing (ps) term $W_{\text{ps}}^{(L,\beta)}(\vec{r})$ ("repulsive pseudopotentials"). Here the charge density contains only valence electrons. Thus,

$$V_H^{\text{ext}}(\vec{r}) = \sum_{\vec{R}_p} \sum_{\vec{\tau}_\beta} \left[\frac{-Z_\beta^v}{|\vec{r} - \vec{R}_p - \vec{\tau}_\beta|} + \sum_L W_{\text{ps}}^{(L,\beta)}(\vec{r} - \vec{R}_p - \vec{\tau}_\beta) \hat{P}_L \right]$$

and

$$\int \rho_H(\vec{r})d\vec{r} = P \sum_\beta Z_\beta^v. \quad (6)$$

The terms in the large parentheses define the L component of the "nonlocal" core pseudopotential of atom β as

$$v_{\text{ps}}^{(L,\beta)}(r) \equiv -Z_\beta^v/r + W_{\text{ps}}^{(L,\beta)}(r)\hat{P}_L, \quad (7)$$

and \hat{P}_L is an angular-momentum projection operator.⁶ [In an all-electron representation, $W_{\text{ps}}^{(L,\beta)} = 0$ and the number of valence electrons Z_β^v is replaced by the atomic number Z_β .] The screening potential contains, in general, an interelectronic (ee) Coulomb term $V_H^{\text{ee}}(\rho_H(\vec{r}))$, and an exchange-correlation (xc) term $V_H^{\text{xc}}(\rho_H(\vec{r}))$. It may be formulated by a wide range of approaches (e.g., simple dielectric screening, random-phase approximation, Hartree-Fock, and Thomas-Fermi). Here we choose to represent it by the density-functional formalism⁷ in which

$$V_H^{\text{scr}}(\vec{r}) = V_H^{\text{ee}}(\rho_H(\vec{r})) + V_H^{\text{xc}}(\rho_H(\vec{r})), \quad (8a)$$

where the individual components are [for $\rho_H(\vec{r}) = \rho(\vec{r})$]

$$V_H^{\text{ee}}(\rho(\vec{r})) = \int \frac{\rho(\vec{r}')}{|\vec{r} - \vec{r}'|} d\vec{r}'$$

and

$$V_H^{\text{xc}}(\rho(\vec{r})) = \frac{\partial}{\partial \rho}(\rho \epsilon_{\text{xc}}[\rho]). \quad (8b)$$

The term $\epsilon_{xc}[\rho]$ is the exchange (with an exchange coefficient α_x) and correlation energy per particle of the homogeneous electron gas.

Equations (1)–(8) define the electronic structure of the host crystal; given $V_H^{\text{ext}}(\vec{r})$ and N_e^H alone, one can solve for the self-consistent response $V_H^{\text{scr}}(\vec{r})$, the charge density $\rho_H(\vec{r})$, and the single-particle spectrum $\{\epsilon_j^0(\vec{k})\}$ in Eq. (1). The external potential $V_H^{\text{ext}}(\vec{r})$ contains the space group and the core pseudopotentials $\{v_{\text{ps}}^{(L,\beta)}(r)\}$. The latter quantities can be derived in a first-principles approach in the density-functional formalism for any atom.⁸

In the defect-containing system, the external potential $V_H^{\text{ext}}(\vec{r})$ is replaced by $V_D^{\text{ext}}(\vec{r})$; the electrons respond to the changed potential, producing a screening $V_D^{\text{scr}}(\vec{r})$ replacing $V_H^{\text{scr}}(\vec{r})$. The potential perturbation due to a defect is defined here to include both contributions:

$$\begin{aligned} \Delta V(\vec{r}) &\equiv \Delta V^{\text{ext}}(\vec{r}) + \Delta V^{\text{scr}}(\vec{r}) \\ &= [V_D^{\text{ext}}(\vec{r}) - V_H^{\text{ext}}(\vec{r})] + [V_D^{\text{scr}}(\vec{r}) - V_H^{\text{scr}}(\vec{r})], \end{aligned} \quad (9a)$$

and the associated density fluctuation is

$$\Delta\rho(\vec{r}) \equiv \rho_D(\vec{r}) - \rho_H(\vec{r}). \quad (9b)$$

The change $\Delta V^{\text{ext}}(\vec{r})$ in the external potential may include changes in the chemical identity of a host atom from Z_β to Z_I (i.e., substitutional impurity), replacement of a host Z_β by zero (vacancy), addition of a Z_I to the set $\{Z_\beta\}$ (interstitial impurity), a change in the vectors $\{\vec{R}_\beta, \vec{\tau}_\beta\}$ around a defect site to new site vectors $\{\vec{R}_I\}$ (defect-induced lattice relaxation), a change in $\{\vec{R}_\beta\}$ in a line or a plane (line and plane dislocations), or even replacement of Z_β electronically by $Z_\beta - 1$ in an atomic subshell (e.g., a core hole). The defect potential is hence represented by

$$V_D(\vec{r}) = V_D^{\text{ext}}(\vec{r}; \{\vec{R}_I\}; \{Z_I\}) + V_D^{\text{scr}}(\rho_D(\vec{r})), \quad (10)$$

where $V_D^{\text{ext}}(\vec{r})$ and $V_D^{\text{scr}}(\rho_D(\vec{r}))$ have the same functional forms as in Eqs. (6) and (7) and Eq. (8), respectively. The charge density of the defected system is given in terms of its wave functions $\{\psi_i(\vec{r})\}$ and occupation numbers N_i (containing N_e^D electrons) as

$$\rho_D(\vec{r}) = \sum_i^{\text{occ}} N_i |\psi_i(\vec{r})|^2, \quad (11)$$

and the defect screening potential $V_D^{\text{scr}}(\rho_D(\vec{r}))$ is self-consistently described via the forms in Eq. (8), by using $\rho(\vec{r}) = \rho_D(\vec{r})$.

In what follows we discuss defect systems in which $[V_D^{\text{ext}}(\vec{r}) - V_H^{\text{ext}}(\vec{r})]$ decays asymptotically away from its center, leading to what we define as localized defect states. These are characterized by an asymptotically decaying response perturbation $[V_D^{\text{scr}}(\vec{r}) - V_H^{\text{scr}}(\vec{r})]$ whose characteristic size scale depends on $\Delta V^{\text{ext}}(\vec{r})$.

Using the definitions of the host electronic structure [Eqs. (1)–(8)] and the defect perturbation [Eqs. (9)–(11)] we address in this paper the problem of finding solutions to the defect single-particle equation,

$$\begin{aligned} &[-\frac{1}{2}\nabla^2 + V_H(\rho_H(\vec{r})) \\ &+ \Delta V(\rho_H(\vec{r}), \rho_D(\vec{r}))]\psi_i(\vec{r}) = \epsilon_i \psi_i(\vec{r}). \end{aligned} \quad (12)$$

The inputs to the problem are: (i) the pseudopotentials $v_{\text{ps}}^{(L,H)}(r)$ and $v_{\text{ps}}^{(L,D)}(r)$, (ii) the number of electrons N_e^H and N_e^D , and (iii) the atomic structure coordinates $\{\vec{R}_\beta, \vec{\tau}_\beta, \vec{R}_I\}$. The output we are most interested in includes the single-particle energies $\{\epsilon_i\}$ and orbitals $\{\psi_i(\vec{r})\}$, as well as the density $\Delta\rho(\vec{r})$ and screening $\Delta V^{\text{scr}}(\vec{r})$ fluctuations. Our objective is to design a method of solution that will reflect the predictions of the physical input and the underlying density-functional description of interelectronic screening, rather than the artifacts of a computational scheme.

III. DIFFERENT VIEWS ON DEFECT WAVE FUNCTIONS

In this section we will describe the two approaches taken in the past to describe the defect wave functions of Eq. (12): expansion in terms of the extended host-crystal Bloch eigenfunctions $\{\phi_j^0(\vec{k}, \vec{r})\}$ and expansion in terms of local orbitals. The first approach has been used in various perturbationlike models^{9–13} whereas the second approach has been used in crystal field,^{14,15} ligand field,^{16–19} and in various cluster methods.^{20–25} It has been recently shown²⁶ that the Green's-function approach to defects is derivable by requiring that both descriptions of the defect wave functions be *simultaneously and independently satisfied* (the “dual representation”). We will summarize these results below for completeness and for establishing both the notation to be used later and the basis of our new approach to the problem. In Sec. IV we will use these results to show that the dual representation in Green's-function approaches is responsible for a fundamental difficulty in describing defects whose characteristic length scale approaches

that of atomic dimensions. This analysis will lead to a new formulation of the problem in terms of quasibands (rather than pure host bands), which resolves these difficulties in a simple manner. It will be described and illustrated in Sec. V.

The first approach to defect wave functions involves its spectral decomposition into the eigen-

states $\{\phi_j^0(\vec{k}, \vec{r})\}$ of the host crystal classified by bands j and wave vectors \vec{k} :

$$\psi_i(\vec{r}) = \sum_j^M \sum_{\vec{k}}^{\text{BZ}} A_{ij}(\vec{k}) \phi_j^0(\vec{k}, \vec{r}). \quad (13)$$

The operator eigenvalue problem in Eq. (12) is hence transformed into a matrix secular problem

$$[\epsilon_j^0(\vec{k}') - \epsilon_i] A_{ij'}(\vec{k}') + \sum_j^M \sum_{\vec{k}}^{\text{BZ}} A_{ij}(\vec{k}) \langle \phi_j^0(\vec{k}', \vec{r}) | \Delta V(\vec{r}) | \phi_j^0(\vec{k}, \vec{r}) \rangle = 0, \quad (14)$$

from which the coefficients $\{A_{ij}(\vec{k})\}$ are found. The appeal of this direct diagonalization method²⁷ stems from the fact that it permits the analysis of the evolution of defect wave functions from host wave functions as the perturbation $\Delta V(\vec{r})$ is turned on. It implies, however, that the physical characteristics of a defect wave function $\psi_i(\vec{r})$ could be described by a (hopefully) manageable number M of host wave functions $\{\phi_j^0(\vec{k}, \vec{r})\}$, which may, in general, be chemically and physically unrelated to the impurity. To take some extreme examples, in using Eq. (13) one needs to hope that the wave functions of a heavy impurity atom (e.g., uranium) in a hydrogen lattice, could be effectively spanned by a few hydrogen bands, or that d -like impurities could be described by a few s - p -like host wave functions, etc. The conventional Green's-function approach¹⁻⁴ tacitly assumes this condition.²⁶ The degree of convergence of the results on the number M becomes then a crucial issue. This difficulty has been recognized quite early,²⁸ but has not been discussed in contemporary calculations.

An alternative approach to describing $\psi_i(\vec{r})$ has been offered by the crystal field, ligand field, and cluster approaches. Inspired by a localized molecular view, the defect wave functions are expanded in a set of *local functions* $\{g_a(\vec{r})\}$, usually related to the impurity and ligand atoms:

$$\psi_i(\vec{r}) = \sum_a^N C_{ia} g_a(\vec{r}). \quad (15a)$$

This transforms Eq. (12) to a secular equation,

$$\sum_a^N C_{ib} [\langle g_a(\vec{r}) | -\frac{1}{2} \nabla^2 + V_H(\vec{r}) + \Delta V(\vec{r}) | g_b(\vec{r}) \rangle - \epsilon_i \langle g_a(\vec{r}) | g_b(\vec{r}) \rangle] = 0. \quad (15b)$$

In contrast to the previous approach [cf. Eq. (14)] the full defect Hamiltonian, rather than the perturbation $\Delta V(\vec{r})$ alone, appears in the secular problem. The appeal of this approach (e.g., the remarkable results by Hemstreet²³) stems from the ability to implement one's intuition on the system by selecting chemically reasonable orbitals $g_a(\vec{r})$ in this expansion (e.g., symmetry-adapted atomic wave functions of the impurity and first coordination shell ligand atoms). Its weakness is rooted in the fact that N in Eq. (15) is often limited in practice to small values due to the need to treat the full defect Hamiltonian, and that expansion in terms of local orbitals often describes poorly the extended host states with which the defect interacts.

In the standard Green's-function approach to localized defect problems, one is tacitly requiring that the two views expressed in Eqs. (13) and (15a) be *simultaneously* fulfilled: Not only must enough bands M be included to make the expansion in Eq. (13) complete, but also the expansion in Eq. (15a)

in terms of N localized basis functions $\{g_a(\vec{r})\}$ must be valid in the subspace of the perturbation. That is, if $\Theta(\vec{r} - \vec{R}_c)$ denotes a step function that equals unity in that part of space ($r \leq R_c$) where the defect potential $\Delta V(r)$ is nonzero and is zero (or decays to zero) otherwise, both Eq. (13) and

$$\Theta(\vec{r} - \vec{R}_c) \psi_i(\vec{r}) = \sum_a^N C_{ia} g_a(\vec{r}) \quad (16)$$

must be valid.

This has been proved in Ref. 26 which also demonstrated that the basic Green's-function matrix equations^{3,4} for calculating one-particle energies and wave functions are derived from Eq. (14) under the assumption that Eqs. (13) and (16) are simultaneously satisfied. It was further demonstrated that if Eq. (13) was assumed^{3,4} both conditions are, in fact, necessary.

According to Ref. 26, the one-particle energies ϵ_i and wave-function expansion coefficients C_{ia} for state i are determined by the set of simultaneous

equations:

$$\sum_a [\delta_{a'a} - \sum_{a''} \tilde{G}^0(\epsilon_i)_{a'a''} \langle g_{a''} | \Delta V | g_a \rangle] C_{ia} = 0, \quad (17a)$$

where the nonorthogonal representation of the Green's-function matrix is

$$\tilde{G}^0(\epsilon)_{aa'} = [\underline{S}^{-1} \underline{G}^0(\epsilon) \underline{S}^{-1}]_{aa'}, \quad (17b)$$

and the elements of the basis-set overlap matrix are denoted by $S_{ab} = \langle g_a | g_b \rangle$. Here the host-crystal Green's-function matrix is given in its standard form as

$$G^0(\epsilon)_{ab} = \sum_{j=1}^M \sum_{\vec{k}}^{\text{BZ}} \frac{\langle g_a | \phi_j^0(\vec{k}, \vec{r}) \rangle \langle \phi_j^0(\vec{k}, \vec{r}) | g_b \rangle}{\epsilon - \epsilon_j^0(\vec{k})}. \quad (17c)$$

The overlap matrix \underline{S} has been introduced for a general nonorthogonal basis. Furthermore, the requirements in Eqs. (13) and (16) lead to the well-known Lippman-Schwinger equation relating the defect wave function at \vec{r} to its value at \vec{r}' :

$$\psi_i(\vec{r}) = \int G^0(\vec{r}, \vec{r}'; \epsilon_i) \Delta V(\vec{r}') \psi_i(\vec{r}') d\vec{r}', \quad (18)$$

with

$$G^0(\vec{r}, \vec{r}'; \epsilon) = \sum_j^M \sum_{\vec{k}}^{\text{BZ}} \frac{\phi_j^0(\vec{k}, \vec{r}) \phi_j^{0*}(\vec{k}, \vec{r}')}{\epsilon - \epsilon_j^0(\vec{k})}, \quad (19)$$

so that the wave function $\psi_i(\vec{r})$ is determined in all space once it has been calculated in the subspace of the perturbation.

The standard Green's-function representation in Eqs. (17)–(19) forms the basis for a number of recent calculations on defect problem.^{2–4,27,29} In these calculations, various numerical computational schemes have been devised to solve the underlying Eqs. (17)–(19) for simple defects (i.e., vacancies and non-transition-atom impurities containing s and p , but not the truly localized d orbitals). With such defects, where the perturbation extends over a few bond distances, the consequences of the underlying dual representation of the defect wave functions cannot be fully appreciated. We discuss the implications of this representation in the next section and show that it leads to fundamental difficulties when truly deep defects are considered (i.e., when the perturbation approaches atomic dimensions). The deep impurity center problem was treated in the past both by methods that rely on the expansion in Eq. (15a) and by Green's-function methods. We discuss in the next two sections,

IV A and IV B, respectively, the underlying difficulties associated with these two approaches. We then use this analysis in Sec. V to suggest a new method that complements the previous methods in that it works effectively even in situations where the previous methods are inapplicable.

IV. WHY IS THE PROBLEM DIFFICULT?

A. Obvious reasons

Many of the obvious reasons for the difficulties associated with solving the localized defect problem via Eqs. (15) can be appreciated by considering the one-electron energy scales involved. While the relevant matrix elements of $V_H(\vec{r})$ and $V_D(\vec{r})$ span a range of a few Rydbergs, many physically interesting point defects produce energy levels located in a far narrower energy range: the fundamental band gap (e.g., 1–2 eV in semiconductors). Yet, it is the small variations (often ~ 0.1 eV) in the energy levels of chemically different defects^{30(a),(b)} that have immense technological consequences for solar cells,³¹ diodes, light detectors,³² etc. Clearly, many of the simplifying approximations often used successfully to describe semiconductor band structures may be entirely inadequate for treating localized defects. In particular, even if the potential of the host crystal is a smooth function in coordinate space (i.e., in the pseudopotential representation) and hence can be conveniently described in a basis of plane waves, the perturbation potential $\Delta V(\vec{r})$ associated with a defect is not. Furthermore, the quantum interference with the extended host wave functions leads to some wave functions having components that far exceed the range of $\Delta V(\vec{r})$ and consequently complicate considerably the integration of the one-particle equations. The substantial amounts of electronic charge and spin polarization experimentally observed to reside outside the central cell [e.g., Ref. 30(a)] even for deep levels make questionable the validity of the approaches that rely on expanding $\psi_i(\vec{r})$ exclusively in terms of a few local basis functions surrounding the defect site [e.g., Eq. (16) but not Eq. (13)]. At the same time this effect causes the models using small ($\lesssim 50$ atoms) repeated cells^{33–35} to introduce a large and spurious dispersion in the defect energy ϵ_i . Although the interference of the impurity atom orbitals with the host-crystal wave functions is often large enough to invalidate a purely atomiclike approach to the problem (e.g., crystal field^{14,15}), the short-

range part of the perturbing potential is frequently too strong to permit low-order perturbation approaches based on the host wave functions as zero-order states.^{12,13}

Most of these difficulties associated with solving the deep defect problem are overcome by the Green's-function approach.¹⁻⁴ The knowledge of $\psi_i(\vec{r})$ only in the range where the perturbing potential $\Delta V(\vec{r})$ is nonzero is sufficient [c.f. Eq. (18)] to generate $\psi_i(\vec{r})$ everywhere. Consequently, the energies and wave functions of both localized and extended states can be found by using the localized representation in Eq. (16) if the set $\{g_a(\vec{r})\}$ is sufficiently complete only in the subspace of the perturbation. This is in sharp contrast to methods that rely only on the expansion in Eq. (15a) where, as can be seen from Eq. (15b), one needs to span the entire space of $V_H(\vec{r}) + \Delta V(\vec{r})$.

B. Less obvious reasons

There are at least four reasons that considerably complicate the use of the standard Green's-function approach to localized defect states. These are reviewed below in Secs. IV B 1 to IV B 4. Our resolution of these issues will be presented in Secs. V A – V D, respectively.

1. The dual description by localized functions and host band wave functions

The advantages offered by the Green's-function approach to defects in combining the two different views [expressed in Eqs. (13) and (16)] on the defect wave functions (e.g., the existence of a Lippman-Schwinger relationship) are also the sources of its weakness: The two rather opposing views on the defect wave functions need to be equally and simultaneously fulfilled. It can be demonstrated that even if one had a formally exact description of $\psi_i(\vec{r})$ in terms of some local orbitals $\{g_a(\vec{r})\}$ (e.g., if one knew the exact wave function or the Wannier orbitals of the defect system) but an additional expansion of the form of Eq. (13) is used,^{3,4} one needs to make the latter expansion *independently* complete in order to recover the correct energies. This can be illustrated by using a simple, analytically solvable model.²⁶ Consider a free-electron model for a silicon crystal with a potential $V_H(\vec{r})=0$ and Bloch eigenfunctions

$$\phi_j^0(\vec{k}, \vec{r}) = \Omega^{-1/2} e^{i(\vec{k} + \vec{G}) \cdot \vec{r}}$$

in Eq. (1), where Ω is the unit-cell volume and \vec{G} are reciprocal-lattice vectors. Assume we are seeking the energy eigenvalues associated with a parabolic defect potential $V_D(\vec{r}) = K^2(\vec{r} - r_0)^2$ centered on the lattice site $\vec{R} = \vec{r}_0$ in this crystal.

The exact solutions (denoting i as nlm) that we will attempt to recover using the Green's-function method are

$$\epsilon_{nl} = K(n + \frac{3}{2}) \quad (20a)$$

and

$$\psi_{nlm}(\vec{r}) = N_{nl} (Kr^2)^{l/2} e^{-Kr^2/2} L_{(n-l)/2}^{l+1/2} \times (Kr^2) Y_{lm}(\theta, \phi). \quad (20b)$$

Here, N_{nl} is the normalization factor, $L_p^q(Kr^2)$ is the Laguerre polynomial, and $Y_{lm}(\theta, \phi)$ is a spherical harmonic. Assume further that the defect wave functions are already known exactly in terms of a local orbital expansion of Eq. (15a), i.e., use for $g_a(\vec{r})$ the solutions $\psi_{nlm}(r)$ of Eq. (20b). We now use the formalism of Eqs. (13), (16), and (17) to solve for the defect energies, inquiring how many pure host bands M are required to recover the exact energies in Eq. (20a).

The upper part of Table I compares the exact energy eigenvalues [Eq. (20a)] for the lowest defect levels of symmetry a_1 , t_2 , and e_2 (angular momentum of $l=0, 1$, and 2 , respectively), with those obtained by the conventional defect Green's-function (DGF) approach, using the first $M=10, 20, 32$, and 41 host bands for each \vec{k} value. These host bands span a very large energy range of 1.7, 3.0, 3.7, and 4.5 Ry, respectively. Yet, the errors involved in the defect energy levels ϵ_i are seen to be enormously large on any relevant scale. Clearly, the conventional Green's-function approach with its underlying dual description of the impurity wave functions fails in reproducing effectively the defect structure even if $\psi_i(\vec{r})$ is inputted *exactly* in terms of the local orbitals [Eq. (16)]. We may, in general, expect the defect wave functions to converge very slowly with respect to the number of host bands used whenever the defect is chemically and physically sufficiently different from the host. This includes cases such as impurities that have an angular momentum or principal valence quantum numbers that are absent from the host (e.g., d -electron impurities in an s,p electron host; a $5s$ impurity in a $4s$ host), or systems where the defect-induced lattice relaxation is substantial. For example, it can be shown²⁶ that a description of the d -like defect states associated with a transition atom such as Cr in a silicon free-electron host requires

TABLE I. The calculated defect energy levels of a parabolic perturbation potential $\Delta V(\vec{r})=K(r-r_0)^2$ in a silicon free-electron host, comparing the results of the conventional defect Green's-function (DGF) method and the present quasiband Green's-function method. The average number of host band wave functions included for each \vec{k} point is M ; six quasibands are used corresponding to the total dimensionality of the three lowest defect levels $a_1+t_2+e_2$. The potential perturbation is characterized by $K=1$ Ry/a.u.² and $|\vec{r}-\vec{r}_0|\leq 4$ a.u.; energies are given with respect to the zero of the potential well (1 Ry = 13.605 eV). Notice the extremely slow convergence of the values obtained with the conventional DGF method, relative to the immediate convergence obtained with the present quasiband method.

Number of bands	a_1 (eV)	t_2 (eV)	e_2 (eV)
Conventional DGF method			
$M=10$	69.413	130.553	186.946
$M=20$	52.175	96.759	146.240
$M=32$	46.257	83.507	127.139
$M=41$	43.835	77.440	115.180
Exact results	40.815	68.025	95.235
Quasiband DGF method			
$M=1$	40.815	68.025	95.235
$M=10$	40.815	68.025	95.235
$M=20$	40.815	68.025	95.235
$M=32$	40.815	68.025	95.235

the inclusion of no less than 70 000 host bands in the expansion of Eq. (13) to obtain a precision of 0.2 eV in the defect energy level (i.e., $\sim 20\%$ of the silicon band gap).

We notice here that in the tight-binding Green's-function (TBGF) approach to defects^{36,37} one is using a fixed and small number $M \cong 8-10$ of host bands in the expansion of Eq. (13) for all the perturbation potentials $\Delta V(\vec{r})$ considered. Our discussion here indicates that such an approach cannot legitimately recover the known chemical trends in the impurity energies $\epsilon_i(\Delta V)$ with the chemical identity of the defect (given by ΔV). This is so because $M \cong 8-10$ host wave functions are insufficient to describe equally defect states associated with both extended perturbations $\Delta V(\vec{r})$ (i.e., impurities with small $Z_I^v - Z_H^v$ values) and with localized perturbations $\Delta V(\vec{r})$ (i.e., impurities with large $Z_I^v - Z_H^v$ values). This failure of the TBGF concept has been demonstrated quantitatively by Singh *et al.*³⁸

This basic difficulty with the standard Green's-function approach to hyperlocalized impurities may be resolved by brute force by increasing the number of host bands M to either the convergence

limit or the computer's capability limit. The physical transparency of the model is, however, lost. In Sec. V A we describe a simple and elegant method to overcome the difficulty; only a few quasiband wave functions are required to achieve a result that is exact, in principle, and very precise, in practice (cf. lower part of Table I).

2. The multicenter nature of the problem— insights from the crystal-field approach

The conventional Green's-function approach to localized defects [Eqs. (13), (16), and (17)] tacitly assumes a dual discretization of space into the host crystal-lattice sites $\{\vec{R}_p\}$ and the sites $\{\vec{R}_I\}$ in the perturbation subspace. The first set is used to define $\{\phi_j^0(\vec{k}, \vec{r})\}$ in Eq. (13) while the second is used to define $\{g_a(\vec{r})\}$ in Eq. (16). This discretization is rooted in LCAO-type approaches in molecular chemistry where the central quantity, the molecular or crystal potential $V(\rho(\vec{r}))$, is indeed extended over many atoms. This approach transforms the defect problem into a multicenter representation: a large number of integrals such as

$$\langle g_a(\vec{r}-\vec{R}_a) | \Delta V(\vec{r}-\vec{R}_I) | g_b(\vec{r}-\vec{R}_b) \rangle$$

[cf. Eq. (17a)] involving three-dimensional, nonspherical multicentered functions need to be evaluated. The task of obtaining self-consistency in $\Delta V^{\text{scr}}(\Delta\rho(\vec{r}))$ also involves the calculation of a multicentered potential and charge density. The number of such integrals required in Eq. (17) (which equals the number of host bands times the number of \vec{k} points times the number of local orbitals $g_a(\vec{r})$ times the number of lattice sites) extends into the millions. Although large numbers of multicentered integrals can be handled numerically on big high-speed computers, it leaves something to be desired in terms of simplicity and physical transparency. In addition, it often requires limiting $\{g_a(\vec{r})\}$ to simple forms (e.g., Gaussians) that individually do not have the physical properties of the quantum system at hand. (For example, Huzinaga^{39(a)} and Ten Hoor^{39(b)} needed no less than ten Gaussian functions and a simultaneous optimization of its ten *nonlinear* coefficients together with ten linear coefficients to reproduce the energy of the hydrogen 1s orbital. This optimization involves a search in a space with multiple minima and saddle points. Also, 3d orbitals of transition elements require more Gaussians.)

As early as 1929, Bethe¹⁴ and subsequently Van Vleck¹⁵ realized, however, that while $[V_D(\vec{r}), V_H(\vec{r})]$, and $[\rho_D(\vec{r}), \rho_H(\vec{r})]$ are individually multicentered in character and highly anisotropic, the corresponding *differences* $\Delta V(\vec{r})$ and $\Delta\rho(\vec{r})$ [Eqs. (9a) and (9b)] resemble a one-center object (much like a nonspherical open-shell atom or an atom in an electrical field) with a far reduced spatial anisotropy. This led to the formulation of the crystal-field theory^{16,17} in which both the defect potential and the impurity-anchored orbital basis functions $\{g_a(\vec{r})\}$ are expanded in Kubic harmonics *around the impurity site*. In the simple crystal-field approach, the potential of the defected system is expressed as

$$\begin{aligned} V_D(\vec{r}) &= \{V_H(\vec{r}) - v_a(\vec{r}-\vec{R}_H)\} + v_I(\vec{r}-\vec{R}_I) \\ &\equiv \Delta v(\vec{r}) + v_I(\vec{r}-\vec{R}_I), \end{aligned} \quad (21)$$

(denoting atomic potentials by v and crystalline potentials by V). Here $v_a(\vec{r}-\vec{R}_H)$ is the potential of a single host atom removed from site \vec{R}_H , and $v_I(\vec{r}-\vec{R}_I)$ is the potential of the impurity atom introduced into site \vec{R}_I (in the absence of lattice distortions, $\vec{R}_H = \vec{R}_I$). The first term $\Delta v(\vec{r})$ is hence the crystal field experienced at the point \vec{r} due to all atoms except the one at \vec{R}_H . Bethe and Van

Vleck have realized that the deep defect problem can, nevertheless, be treated as a single-site, atom-like problem, provided that the crystal field is expanded around the defect site alone in a few partial waves:

$$\Delta v(\vec{r}) = \sum_l D_l(|\vec{r}|) K_l^{a_1}(\hat{r}), \quad (22a)$$

with a similar expansion for the impurity orbitals in terms of radial atomic orbitals R_{nl} with atomic principal quantum number n :

$$g_{nl}^{\alpha,\lambda}(\vec{r}) = R_{nl}(|\vec{r}|) K_l^{\alpha,\lambda}(\hat{r}), \quad (22b)$$

where α and λ are the representation and partner indices, respectively, and a_1 denotes the totally symmetric representation. [If the crystal field $\Delta v(\vec{r})$ is modeled approximately by the electrostatic field due to point charges q_i on lattice sites, D_l is independent of r .] The orbital energies of an impurity atom interacting with the crystal field are then obtained from the secular problem derived from Eqs. (21) and (22); the separation of variables underlying Eqs. (22) transforms this secular equation into a simple atomlike problem. The experimentally observed splitting between defect energy levels may be characterized empirically by qD_l . This approach has not only made possible systematization of a vast number of optically observed defect levels in many molecular complexes and solids,¹⁶ but it has also established a quantitative basis for understanding the stability of complex ions, the vibronic coupling of impurities to hosts, the multiplet structure of impurities, the hyperfine splitting of magnetic impurities, the paramagnetism of the complexes of the first transition series, and the anisotropy and temperature variation of the magnetic susceptibility.¹⁶

The success of this simple model of Eqs. (21) and (22) is at first surprising—if the orbitals $g_{nl}(\vec{r})$ of the impurity atom extend outside the range of its atomic potential $v_I(\vec{r})$ into the space where the host-crystal field $\Delta v(\vec{r})$ is strong, they will attempt to variationally mimic hostlike states (“variational collapse”). The orbital energies derived from Eqs. (22) would then contain large errors. A scrutiny of the details of the formalism reveals, however, the source of its success: Whenever orbital energies ϵ_{nl} are calculated for matrix elements such as

$$\langle R_{nl}(r) | -\frac{1}{2}\nabla^2 + V_D(\vec{r}) | R_{nl}(r) \rangle,$$

an implicit cut-off distance R_c is invoked; outside R_c , where variational collapse may occur, $V_D(\vec{r})$ is set to zero (hence, by construction, hybridization with host states is excluded).

The first to realize that this is, in effect, a generalized pseudopotential problem was Phillips⁴⁰ who showed that if one orthogonalizes $\psi_i(\vec{r})$ to the ligand orbitals, the crystal-field problem can still be solved in terms of the one-center impurity-atom orbitals $g_{nl}(\vec{r})$ provided that one adds to $V_D(\vec{r})$ in Eq. (21) a new term $V_R[\vec{r}, \epsilon_{nl}]$ —the crystal-field pseudopotential. A more general formulation of this basic idea is given in Appendix A. The implicit cutoff radius R_c used in the Bethe—Van Vleck crystal-field model is then a simplified realization of the radius *outside* which $V_R[\vec{r}, \epsilon_{nl}]$ has a strong repulsive character.

This then shows two ways by which one could circumvent the multicenter nature of the defect problem and reduce it to a one-center atomiclike problem. It is sufficient to use a single-site, defect-centered basis set as in crystal-field approaches, provided one adds to the potential $V_D(\vec{r})$ the crystal-field effective pseudopotential derived in Appendix A. Alternatively, in a Green's-function approach where the central quantity is the potential perturbation $\Delta V(\vec{r})$ rather than the full potential $V_H(\vec{r}) + \Delta V(\vec{r})$, it is sufficient to use a variationally complete but one-center description of the wave function $\Theta(\vec{r} - \vec{R}_c)\psi_i(\vec{r})$ inside the range R_c of the perturbation $\Delta V(\vec{r})$ and utilize the Lippman-Schwinger Eq. (18) to translate $\psi_i(\vec{r})$ outside \vec{R}_c . Although the wave function $\psi_i(\vec{r})$ may change outside \vec{R}_c as more basis functions are added, this change does not alter the defect energies if the description inside \vec{R}_c is sufficiently complete. We will choose this approach. This approach offers not only a high numerical precision, but also shares with the (far less complete) crystal-field methods its physical simplicity. This will be described in Sec. V B.

3. The \vec{k} summation problem

In the conventional application of the Green's-function method to solid-state problems involving a continuous energy spectrum,⁴¹ the matrix representation of $\tilde{G}^0(\epsilon)$ requires a \vec{k} -space integration of a nontrivial complex function. If this spectral representation is used, one usually (e.g., Ref. 3) first separates this complex function into its real and imaginary parts; the imaginary part is then obtained through BZ sampling techniques (using about 200—300 \vec{k} points) with a microcube (e.g., Gilat-Raubenheimer⁴²) or the tetrahedron (e.g., Lehman and Taut⁴³) partitioning of the BZ, fol-

lowed by a Kramers-Kronig transformation to obtain the real part. The requirement for inclusion of a large number of bands (cf. Sec. IV B 1) and local basis functions together with a fine resolution in the energy parameter ϵ needed to describe a large set of defect wave functions, makes these two operations extensive and costly. In addition, these often require fitting the host bands $\phi_j^0(\vec{k}, \vec{r})$ in terms of an auxiliary basis set³ and the imaginary part of the Green's functions in terms of yet another basis set³ to perform the spectral sampling and the Kramers-Kronig transformation interpolatively.⁴⁴ While the need to maintain these two additional basis sets add complexity, the fitting procedure may introduce errors. In fact, both are not needed if the continuous spectral representation is replaced by a discrete representation.

This problem of extensive BZ summations has been previously avoided in supercell schemes^{32–34,45,46} in which one places an impurity at a center of a large unit cell containing N_Ω host atoms and solves for the associated electronic structure in the corresponding, small Brillouin zone (SBZ) by using band-structure techniques. Only the \vec{k} points folded from the primitive BZ into the SBZ are used; the defect energies then converge rather rapidly with the number of \vec{k} points.^{34,45} The interdefect interaction introduced by the artificial periodic boundary conditions is then reduced by increasing N_Ω . The problem becomes, however, rapidly intractable as the size of the Hamiltonian matrices increases as a high power of N_Ω . For example, although a defect in each of the *two-dimensional*, 50-atom unit cells of planar graphite^{34(c)} produces only a weak (0.1 eV) unphysical defect-defect interaction, in the more usual case of *three-dimensional* solids (e.g., vacancy in silicon³⁵ or zinc in silicon⁴⁵), a 54-atom unit cell produces an intolerable interdefect interaction of 1 eV.

We avoid the slow convergence of the \vec{k} sums in the spectral Green's-function approaches as well as the cumbersome dependence of matrix size on N_Ω in the supercell approach by using a discrete representation for the Green's-function \vec{k} summation. This technique is formally equivalent to a supercell model but the size of the Hamiltonian matrix *does not depend on the dimension of the supercell*; large supercells (e.g., 2662 atoms^{29(a),(b)}) may be easily treated. This special \vec{k} summation has been previously used by Zunger^{34(a)} and Zunger and Englman^{34(b)} in the context of the small periodic cluster method and by Evarestov and Lovchikov,⁴⁶ and Lindelfelt⁴⁵ in the context of the large unit-cell

method. It was first used in the context of Green's-function calculations by Lindefelt.^{29(a),(b)}

This \vec{k} -space summation method simplifies the calculations enormously: Only a small number of \vec{k} points are needed, and the poles of the Green's function are real. At the same time, it provides a direct link to the repeated cell models for point defects: It shows how many atoms per supercell are needed to reach the limit of noninteracting defects. A simple analytical model introduced below can then be used to obtain the energies of the gap levels at the limit of an infinite supercell, by using the results of small supercells. This method provides a direct way of performing calculations on the electronic structure of defects as a function of the concentration of the defects. The method is described in Sec. V C.

4. The self-consistency problem

The solution of the single-particle defect problem, Eq. (12), clearly involves the construction of an operator $\Delta V(\rho_H, \rho_D)$ that requires the knowledge of the solution $\rho_D = \sum_i N_i |\psi_i(\vec{r})|^2$ itself [Eqs. (11) and (12)]. Whereas *ad hoc* prescriptions that fix $\Delta V(\rho_H, \rho_D)$ by using the properties of the isolated atoms [e.g., $\Delta V = v_I(r) - v_a(r)$ in Eq. (21)] are possible,^{35,37} they often require an ingenious guess to correctly reflect the shape and chemical trends in the screened perturbations. The nontriviality of this guess is illustrated in Fig. 1 for an unrelaxed silicon vacancy. Here the inverse coordinate-space screening function

$$\frac{1}{\epsilon(r)} = \frac{\Delta V^{\text{ext}}(r) + \Delta V^{\text{scr}}(r)}{\Delta V^{\text{ext}}(r)} \quad (23)$$

is plotted for: (a) a free Si atom, in which $V^{\text{scr}}(r)$ is obtained self-consistently from $V^{\text{ext}}(r)$ within the density-functional formalism,⁷ (b) a dielectrically screened Si atom, where

$$V^{\text{scr}}(r) = \int \frac{V^{\text{ext}}(q)}{\epsilon(q)} e^{i\vec{q} \cdot \vec{r}} d\vec{q},$$

and $\epsilon(q)$ is the silicon host dielectric function,⁴⁷ and (c) a self-consistent Si vacancy, calculated within the formalism of the present paper (Sec. VI). All three functions $\epsilon^{-1}(r)$ are obtained by using for $V^{\text{ext}}(r)$ the same silicon semiempirical local pseudopotential.³⁵ As seen in Fig. 1, the self-consistent vacancy screening function cannot be approximated in any simple way by the self-consistent atom or by linear dielectric response forms: The small- r behavior, the position of the

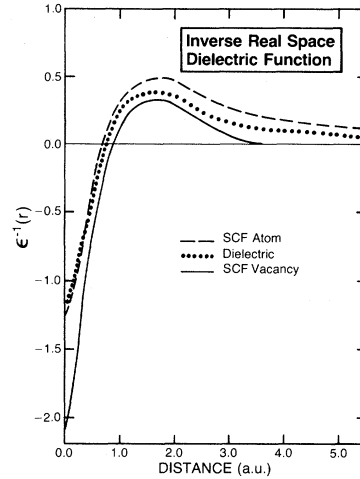


FIG. 1. Inverse screening function $\epsilon^{-1}(r)$ in coordinate space obtained from: (a) a self-consistent density-functional calculation for the free-space Si atom (—), (b) the dielectrically screened Si atom using Nara's (Ref. 47) dielectric function $\epsilon(q)$ (···), and (c) a self-consistent calculation for a Si vacancy, c.f. Sec. VI and Fig. 6 (---). All calculations use the same pseudopotential $V^{\text{ext}}(\vec{r})$ of Ref. 35.

node, and the attenuation of $\epsilon^{-1}(r)$ near the Wigner-Seitz sphere radius are not related in any simple way to the properties of the atomic or dielectric screening function. The problem is far more acute for highly localized impurities (e.g., transition atoms) where small errors in ρ_D translate into substantial errors²⁶ in ΔV as well as ϵ_i . In general, a self-consistent solution to Eq. (12) is therefore needed.

Two basic difficulties are associated with attaining self-consistency: (i) obtaining $\Delta V(\rho_H, \rho_D)$ from a given set of solutions $\{\psi_i(r)\}$, (ii) iterating over $\Delta V(\rho_H, \rho_D)$ for different ρ_D . We have already indicated that problem (i) becomes nearly trivial if the solutions $\{\psi_i(\vec{r})\}$ and the potential $\Delta V(\vec{r})$ are represented in a single-center generalized crystal-field form. The quantities expressed in Eq. (8b) are then obtained through a one-dimensional radial integration and a simple linear fit (Sec. V B). The remaining question—how many single-site functions are required for an accurate representation—is satisfactorily resolved through actual tests (Sec. VI). Problem (ii), however, as treated by existing approaches, turns out to be the single most important source of computational expense: at least 20–40 iterations are needed to obtain a stable self-consistent solution for substitutional Cu in silicon if the initial guess for $\Delta V(\vec{r})$ is constructed

from atomic potentials and if linear mixing^{3,4,35} between input and output $\Delta V(\vec{r})$ is used.

We have overcome this difficulty by using a new and efficient method developed by Bendt and Zunger⁵ for general self-consistency problems. This method uses a Newton-Raphson technique with Jacobian updates. It is far more effective than potential mixing,^{3,4,35} the Pratt scheme,^{48,49} Kerker's scheme,⁵⁰ or Ferreira's accelerator.⁵¹ It enables the efficient self-consistent solution of the defect problem even with modest computational facilities. The method is described in Sec. V D.

V. THE PRESENT RESOLUTION OF THE DIFFICULTIES

We describe in this section our QBCF approach to resolving the difficulties described in the preceding section. The discussion parallels the structure of Secs. IV B 1—IV B 4.

A. The quasiband structure representation of the Green's-function problem

In the standard derivation of the Green's-function approach to the localized defect problem (e.g., Refs. 1—4,13,52), one tacitly assumes that the orthogonal [Eq. (2)] crystal wave functions $\phi_j^0(\vec{k}, \vec{r})$ are *eigenstates to the host crystal Hamiltonian operator* $\hat{H}_0 = -\frac{1}{2}\nabla^2 + V_H$ [Eq. (1)]. Expanding then the defect wave functions in this set $\{\phi_j^0(\vec{k}, \vec{r})\}$ [Eq. (13)], one gets the basic Eq. (14). As discussed in Sec. IV B 1, the requirement in Eqs. (1) and (13) will often result in a computationally intractable and physically undesirable large number M of host eigenfunctions required to represent a localized defect state $\psi_i(\vec{r})$. The central point of our argument is that although the condition that $\phi_j^0(\vec{k}, \vec{r})$ be an eigenstate to \hat{H}_0 is sufficient, it is not a necessary condition for deriving the underlying formalism [i.e., Eq. (14)]. *The far weaker condition that an arbitrarily chosen orthogonal set of Bloch functions $\{\phi_j^{\text{QB}}(\vec{k}, \vec{r})\}$ diagonalize the host crystal Hamiltonian matrix is a sufficient condition.* One is then free to choose the set $\{\phi_j^{\text{QB}}(\vec{k}, \vec{r})\}$ [quasiband (QB) functions] with its eigenvalues $\epsilon_j^{\text{QB}}(\vec{k})$ (quasiband structure) so that the impurity wave function $\psi_i(\vec{r})$ can be expanded exactly, in principle, and very accurately, in practice, by a *small number* \bar{M} of such quasibands:

$$\psi_i(\vec{r}) = \sum_j^{\bar{M}} \sum_{\vec{k}}^{\text{BZ}} A_{ij}^{\text{QB}}(\vec{k}) \phi_j^{\text{QB}}(\vec{k}, \vec{r}). \quad (24)$$

The conventional choice $\phi_j^{\text{QB}}(\vec{k}, \vec{r}) = \phi_j^0(\vec{k}, \vec{r})$ [i.e., Eq. (13)] used in previous models^{3,4} is only a particular case. However, far better choices that incorporate in $\phi_j^{\text{QB}}(\vec{k}, \vec{r})$ the physical characteristics of the defect's wave functions from the outset are possible. Our approach will hence be to renormalize a large number of conduction-band wave functions that would have been needed to spectrally describe $\psi_i(\vec{r})$ through Eq. (13), into a far smaller number of local quasiband wave functions, which contain from the outset the relevant chemical information needed to describe $\psi_i(\vec{r})$.

We have previously proved²⁶ that in fact the set $\{\phi_j^{\text{QB}}(\vec{k}, \vec{r})\}$ used in the spectral representation of the Green's function need only diagonalize the matrix of the host Hamiltonian operator, but need not form eigenstates to it. This means that even if the high-energy conduction-band wave functions are not accurate eigenstates to \hat{H}_0 (as is often the case in practical band-structure calculations using a limited basis set), this fact is of no consequence as long as these wave functions were obtained by diagonalizing the (finite) Hamiltonian matrix.

This proof has an important implication for defect calculations, in that it offers a new way to circumvent the difficulty associated with expanding $\psi_i(\vec{r})$ in host Bloch states. Instead of increasing continuously the number of pure host bands needed in an accurate expansion of Eq. (13), one can augment a set of few band-structure (BS) wave functions $\{\phi_j^{\text{BS}}(\vec{k}, \vec{r})\}$ by a few local wave functions $\{\chi_j(\vec{k}, \vec{r})\}$ containing the relevant chemical information on the impurity, provided that a unitary transformation $\{\phi_j^{\text{QB}}(\vec{k}, \vec{r})\}$ of the combined set $\{\phi_j^{\text{BS}}(\vec{k}, \vec{r}), \chi_j(\vec{k}, \vec{r})\}$ diagonalizes the host Hamiltonian matrix.

The quasiband wave functions are then given as

$$\phi_j^{\text{QB}}(\vec{k}, \vec{r}) = \sum_{j'}^{M_b} b_{jj'}(\vec{k}) \phi_{j'}^{\text{BS}}(\vec{k}, \vec{r}) + \sum_{j''}^{M_a} a_{jj''}(\vec{k}) \chi_{j''}(\vec{k}, \vec{r}), \quad (25)$$

where $\{b_{jj'}(\vec{k}), a_{jj''}(\vec{k})\}$ are given from the diagonalization of the matrix of \hat{H}_0 , with eigenvalues $\epsilon_j^{\text{QB}}(\vec{k})$. The local quasiband orbitals $\chi_j(\vec{k}, \vec{r})$ may be generated from a LCAO representation,

$$\chi_j(\vec{k}, \vec{r}) = N^{-1/2} \sum_{\vec{R}_p} e^{i\vec{k} \cdot \vec{R}_p} f_j(\vec{r} - \vec{R}_p), \quad (26)$$

where $\{f_j(\vec{r})\}$ are some local orbitals. In band-structure calculations where a mixed-basis representation such as Eq. (25) is used,⁵³ one is choosing

$\{f_j(\vec{r})\}$ to have characteristics of the *host* atoms, so that $\{\phi_j(\vec{k}, \vec{r})\}$ becomes a better approximation to the true eigenfunctions of \hat{H}_0 than $\{\phi_j^{\text{BS}}(\vec{k}, \vec{r})\}$ alone. We have shown, however, that if the defect problem is solved via a Green's-function formalism, full convergence of the band wave-functions set $\{\phi_j^{\text{BS}}(\vec{k}, \vec{r})\}$ to the true eigenstates $\{\phi_j^0(\vec{k}, \vec{r})\}$ is not required formally. Instead of seeking, therefore, approximate completeness of $\phi_j^{\text{QB}}(\vec{k}, \vec{r}) \rightarrow \phi_j^0(\vec{k}, \vec{r})$ as in band-structure calculations, we will take the opposite view: We will use a manifestly incomplete (i.e., underconverged) finite set $\{\phi_j^{\text{QB}}(\vec{k}, \vec{r})\}$ of the form in Eq. (25) but we will choose $\{f_j(\vec{r})\}$ of Eq. (26) to share the characteristics of the perturbation $\Delta V(\vec{r})$ rather than those of the host potential $V_H(\vec{r})$. By the variational principle, augmenting the band-structure basis set $\{\phi_j^{\text{BS}}(\vec{k}, \vec{r})\}$ by another set $\{\chi_j(\vec{k}, \vec{r})\}$ (even if it is physically unrelated) can do no harm. At the same time, a judicious choice of $\{f_j(\vec{r})\}$ can make the expansion of the defect wave function $\psi_i(\vec{r})$ in terms of the quasibands $\{\phi_j^{\text{QB}}(\vec{k}, \vec{r})\}$ [Eq. (24)] converge much faster than the corresponding expansion in terms of host wave functions alone [Eq. (13)].

Note that if $f_j(\vec{r})$ were to be the actual defect wave function $\psi_i(\vec{r})$, (e.g., the crystal-field limit), the set $\{\phi_j^{\text{QB}}(\vec{k}, \vec{r})\}$ of quasiband wave functions will give us the exact solution and therefore the exact energy eigenvalue. This is demonstrated below for a simple case that can be solved analytically. If, however, the trial orbital $f_j(\vec{r})$ is only an approximation for the true solution $\psi_i(\vec{r})$, sufficient components of host wave functions $\{\phi_j^{\text{BS}}(\vec{k}, \vec{r})\}$ will be introduced in Eq. (25) to make the expansion in Eq. (24) complete. However, these have to reproduce merely the *difference* between $\psi_i(\vec{r})$ and the trial orbitals $f_j(\vec{r})$. This is far easier to accomplish than expanding $\psi_i(\vec{r})$ directly in extended states.^{1-4,9-13}

The power of the quasiband approach can be illustrated for a case that has closed-form analytic solutions: an harmonic oscillator impurity in a diamond-lattice, free-electron host discussed in Sec. IV B 1. Using $M_a=6$ local quasibands, where $f_j(\vec{r})$ are the harmonic oscillator orbitals, we solve the Green's-function problem in Eq. (17), replacing $\{\phi_j^0(\vec{k}, \vec{r})\}$ by $\{\phi_j^{\text{QB}}(\vec{k}, \vec{r})\}$ and $\{\epsilon_j^0(\vec{k})\}$ by $\{\epsilon_j^{\text{QB}}(\vec{k})\}$. We now ask how many host bands M_b are required to obtain the exact defect energy level given by Eq. (20a).

The upper part of Table I, discussed in Sec. IV B 1, shows the poor convergence of the defect energy levels obtained in the standard Green's-

function approach to the problem. The lower half of Table I shows the results obtained in the quasiband approach. Even a single band-structure wave function ($M_b=1$) required merely to establish the energy origin is sufficient to recover the exact results for these defect levels. Further discussion of this example and illustrative applications to a Cr impurity in a silicon empty lattice are given elsewhere²⁶ (for the latter case, only $M_b=14$ host wave functions and $M_a=2$ local functions were used to obtain very precise defect energies; in the conventional Green's-function approach, more than 70 000 host bands are needed for a precision of 0.2 eV). These examples demonstrate the extreme efficiency of the quasiband approach.

We close the discussion on the quasiband approach by a technical remark. The diagonalization of the host Hamiltonian matrix within a set of \bar{M} quasibands

$$\{\phi_j^{\text{QB}}(\vec{k}, \vec{r})\} = \{\phi_j^{\text{BS}}(\vec{k}, \vec{r}); \chi_j(\vec{k}, \vec{r})\}$$

of Eq. (25) requires the computation of matrix elements of the form

$$\langle \phi_j^{\text{BS}}(\vec{k}, \vec{r}) | \hat{H}_0 | \chi_j(\vec{k}, \vec{r}) \rangle ,$$

$$\langle \chi_j(\vec{k}, \vec{r}) | \hat{H}_0 | \chi_j(\vec{k}, \vec{r}) \rangle ,$$

and the corresponding overlap elements. The expressions for these elements where $f_j(r)$ are arbitrary radial functions are given in Appendix B. The basic Green's function, Eq. (17), requires the evaluation of the overlap integrals $\langle g_a(\vec{r}) | \phi_j^{\text{QB}}(\vec{k}, \vec{r}) \rangle$. An efficient method for evaluating this is described in Appendix C.⁵⁴⁻⁵⁶

B. The one-center crystal-field representation of potentials and wave functions

1. The representation

From our foregoing discussion (cf. Sec. IV B 2) we see that the localized set $\{g_a(\vec{r})\}$ needs to be completely only in the subspace of the perturbation $\Delta V(\vec{r})$ [i.e., to span $\Theta(\vec{r} - \vec{R}_c)\psi_i(\vec{r})$, where R_c is the range of $\Delta V(\vec{r})$ and $\Theta(r)$ is some cut-off functions]. Following Bethe's idea,¹⁴ we choose $\{g_a(\vec{r})\}$ to have a single-center form (with origin at the defects center) with a separation of angular and radial variables. The single index a used so far to characterize $\{g_a(\vec{r})\}$ is consequently decomposed into four indices ($\mu l \alpha \lambda$), where μ is the radial "principal" quantum number, l is the angular momentum in the system's point-group representa-

tion, α is the irreducible representation of the defect's point-symmetry group, and λ is the partner index in the α th irreducible representation:

$$g_{\mu l}^{\alpha\lambda}(\vec{r}) = F_{\mu l}(|\vec{r}|) K_l^{\alpha,\lambda}(\hat{r}) \quad (27)$$

and

$$\Theta(\vec{r} - \vec{R}_c) \psi_i^{\alpha,\lambda}(\vec{r}) = \sum_l^{L_{\max}^{(1)}} \sum_{\mu}^N C_{\mu li}^{\alpha} g_{\mu l}^{\alpha\lambda}(\vec{r}), \quad (28)$$

where $K_l^{\alpha,\lambda}(\hat{r})$ are lattice harmonics. We use the notation $|\vec{r}|$ to emphasize that $F_{\mu l}(|\vec{r}|)$ does not depend on the direction, but merely on a scalar radial quantity.

Single-center, central-field-type basis functions are not new in molecular chemistry. They have been used in 1929 by Bethe¹⁴ for crystal-field studies of molecular complexes and suggested in 1933 by Mulliken⁵⁷ and in 1941 by Buckingham *et al.*⁵⁸ for molecular calculations. Since then they have been extensively applied in molecular studies, most notably by Cohen and Coulson,⁵⁹ Joy and Parr,⁶⁰ and recently by Desclaux and Pyykkö⁶¹ for relativistic calculation. An extensive review on this subject was given by Bishop^{62(a)} and recently by Kranz and Steinborn^{62(b)}. In solid-state applications, single-center-basis functions have been recently used by Song and Malwal⁶³ and Slater⁶⁴ for cellular-type band-structure calculations, and by Jaros and Brand,²⁷ and Lindelfelt,^{29(a),(b)} and Lindelfelt and Pantelides^{29(c)} to calculate deep defect gap state levels.

The choice of $g(\vec{r})$ in the form (27) parallels the standard central-field representation in atomic physics: The single atomic radial orbital $R_{nl}(|\vec{r}|)$ [cf. Eq. (22b)], which is an eigenstate to an isolated impurity atom Hamiltonian $-\frac{1}{2}\nabla^2 + v_I(r)$, is no longer an eigenstate in the presence of the crystal field $\Delta v(\vec{r})$ [cf. Eq. (22a)] which can mix different orbitals. This is hence replaced in Eq. (28) by sums of $g_{\mu l}^{\alpha\lambda}(r)$ where $F_{\mu l}(|\vec{r}|)$ is a general radial function (i.e., not necessarily R_{nl}). Similarly the spherical harmonics $Y_{lm}(\hat{r})$ used in atomic systems with spherical symmetry is replaced in Eq. (27) by the symmetry-adapted lattice harmonics $K_l^{\alpha,\lambda}(\hat{r})$ transforming like the λ th partner of the α th irreducible representation of the point-symmetry group underlying the crystal field. In analogy to atomic physics, the separation of variables in Eq. (27) will reduce the defect problem into analytic angular integrals over $K_l^{\alpha,\lambda}(\hat{r})$ and to simple one-dimensional, one-center radial integrals over $F_{\mu l}(|\vec{r}|)$. The choice of specific radial functions $F_{\mu l}(|\vec{r}|)$ is immaterial for the development of the

formalism; its discussion will be postponed to Secs. VI and VII. It will suffice to say here that we use a large and flexible enough set $\{F_{\mu l}(|\vec{r}|)\}$ that couples to the spherical as well as nonspherical components of the potential and that convergence of various physical quantities with respect to the size and type of the set is examined and achieved.

The central advantage of our choice in Eqs. (27) and (28) is that the defect wave functions are given in a simple form, separable in radial and angular coordinates:

$$\Theta(\vec{r} - \vec{R}_c) \psi_i^{\alpha,\lambda}(\vec{r}) = \sum_{l=0}^{L_{\max}^{(1)}} G_{il}^{\alpha}(|\vec{r}|) K_l^{\alpha,\lambda}(\hat{r}), \quad (29)$$

where

$$G_{il}^{\alpha}(|\vec{r}|) = \sum_{\mu} C_{\mu li}^{\alpha} F_{\mu l}(|\vec{r}|)$$

is a radial function of $|\vec{r}|$. The coefficients $\{C_{\mu li}^{\alpha}\}$ will be determined from the Green's-function problem defined by Eqs. (17a)–(17c). The complementary quasiband representation of the defect wave function in Eq. (24) is

$$\psi_i^{\alpha,\lambda}(\vec{r}) = \sum_j^{\bar{M}} \sum_{\vec{k}}^{\text{BZ}} A_{ij}^{\alpha\lambda}(\vec{k}) \phi_j^{\text{QB}}(\vec{k}, \vec{r}), \quad (30)$$

where we have dropped for clarity the obvious notation QB from the expansion coefficients $A_{ij}^{\alpha\lambda}$. One needs to determine only the set $\{C_{\mu li}^{\alpha}\}$ in Eq. (28) since the set $\{A_{ij}^{\alpha\lambda}(\vec{k})\}$ is related to it by the Lippman-Schwinger Eq. (18). This is shown in Appendix D, which also discusses the relevant normalization conditions. Equations (29) and (30) form the basis of our QBCF representation: They provide a dual description of defect states in terms of a simple central-field series and an expansion in terms of noninteracting impurity and host quasibands.

Notice that the basis $\{g_{\mu l}^{\alpha\lambda}(\vec{r})\}$ is formally not restricted to be orthogonal. In fact, our general formulation of the Green's-function problem in Eq. (17) includes the basis overlap matrix \underline{S} which does not have to be the unit matrix. In practice, however, a nonorthogonal basis may include approximately linearly dependent components that do not add variational degrees of freedom and merely create the illusion of a large basis. The fact that certain calculated quantities (e.g., orbital energies) do not vary by adding more such basis functions is then not an indication for convergence to the correct limit but merely a statement of the variational irrelevance of such additional basis func-

tions. The common critique of basis set orthogonalization³ states that the long-range oscillations introduced by this procedure can penetrate the domain of $\Delta V(\vec{r})$ without adding extra variational flexibility. In the generalized crystal-field representation, however, the existence of a physical distance scale in the problem—the perturbation radius R_c —allows one to overcome this difficulty easily. The basis $\{g_{\mu l}^{\alpha \lambda}(\vec{r})\}$ is Schmidt orthogonalized *within a sphere of radius R_c* such that the orthogonalized function vanishes outside R_c . All members of the set that are linearly dependent to within a prescribed tolerance (i.e., the wave-function amplitude inside R_c after orthogonalization but before normalization is smaller than a prescribed value) are then discarded. The degree of completeness of the basis is monitored by assuring that $\sum_{jk} |\langle j\vec{k} | g_a(\vec{r}) \rangle|^2$ is close to unity for all a values. This guarantees that no false convergence with basis-set size will occur.

2. Potentials and charge densities in the crystal-field representation

The introduction of the single-center, generalized crystal-field basis in Eqs. (27)–(29) simplifies enormously the calculation of the ingredients of the Green's-function problem in Eq. (17). The charge density [Eq. (11)] of the defected system has the simple central-field form,

$$\rho_D^\dagger(\vec{r}) = \sum_{l=0}^{L_{\max}^{(2)}} \rho_{lD}^\dagger(|\vec{r}|) K_l^{a_1}(\hat{r}), \quad (31)$$

with a similar expression for spin-down (\downarrow). Here $(\alpha, \lambda) = a_1$ is the totally symmetric representation of the defect's point group. The expression for the radial coefficients $\rho_{lD}^\dagger(|\vec{r}|)$ in terms of the radial basis $\{F_{\mu l}\}$, the variational coefficient $\{C_{\mu li}^\alpha\}$, and the Gaunt coefficient $\{D_l^\alpha[l', l'']\}$ is given in Appendix E. Similarly, the host charge density $\rho_H(\vec{r})$ [Eq. (4), assumed here to be nonmagnetic] is given in the same basis by

$$\rho_H(\vec{r}) = \sum_{l=0}^{L_{\max}^{(2)}} \rho_{lH}(|\vec{r}|) K_l^{a_1}(\hat{r}), \quad (32)$$

where $\rho_{lH}(|\vec{r}|)$ is expressed in closed form in terms of the radial basis $\{F_{\mu l}\}$, the host quasiband wave functions $\{\phi_j^{\text{QB}}(\vec{k}, \vec{r})\}$, and the Gaunt coefficients $\{D_l^\alpha[l', l'']\}$ in Appendix E. The charge-density perturbation therefore has the simple form,

$$\Delta\rho(\vec{r}) = \sum_{l=0}^{L_{\max}^{(2)}} \Delta\rho_l(|\vec{r}|) K_l^{a_1}(\hat{r}), \quad (33)$$

where the large cancellation between the defect and host densities is affected *algebraically* in

$$\Delta\rho_{l(r)} = \rho_{lD}(r) - \rho_{lH}(r).$$

The various components of the perturbation potential in Eq. (9a) take a similar crystal-field form. The external potential perturbation [Eq. (9a)] is given as

$$\begin{aligned} \Delta V^{\text{ext}}(\vec{r}) &= \sum_{L=0}^{L_{\max}^{(1)}} \Delta v_L^{\text{ext}}(|\vec{r}|) \hat{P}_L \\ &= \sum_{L=0}^{L_{\max}^{(1)}} [v_{\text{ps}}^{(L,D)}(|\vec{r}|) \\ &\quad - v_{\text{ps}}^{(L,H)}(|\vec{r}|)] \hat{P}_L, \end{aligned} \quad (34)$$

where D and H denote impurity and host atoms, respectively, $v_{\text{ps}}^{(L,\beta)}(|\vec{r}|)$ is the nonlocal core pseudopotential of Eq. (7), and lattice relaxation has been omitted for simplicity of notation. The perturbation in the screening potentials [Eqs. (8) and (9)] is partitioned into the interelectronic Coulomb part $\Delta V^{\text{ec}}(\vec{r})$ and the exchange-correlation part $\Delta V^{\text{xc}}(\vec{r})$. The crystal-field form of the density perturbation in Eq. (33) yields immediately a similar form for $\Delta V^{\text{ec}}(\vec{r})$ obtained by solving Poisson's equation for $\Delta\rho(\vec{r})$:

$$\Delta V^{\text{ec}}(\vec{r}) = \sum_{l=0}^{L_{\max}^{(2)}} \Delta v_l^{\text{ec}}(|\vec{r}|) K_l^{a_1}(\hat{r}), \quad (35)$$

where $\Delta v_l^{\text{ec}}(|\vec{r}|)$ is given as a sum of two atomic-like one-dimensional radial integrals discussed in Appendix E. The exchange-correlation perturbation cannot be obtained directly from the difference $\Delta\rho(\vec{r})$ owing to its nonlinear nature. It is nevertheless expressible in the same form

$$\Delta V^{\text{xc}}(\vec{r}) = \sum_{l=0}^{L_{\max}^{(2)}} \Delta v_l^{\text{xc}}(|\vec{r}|) K_l^{a_1}(\hat{r}), \quad (36)$$

where the radial functions $\Delta v_l^{\text{xc}}(|\vec{r}|)$ are obtained by a simple *one-dimensional* fitting procedure described in Appendix E.

3. Variational flexibility in the generalized crystal-field representation

Conceptually, our choice of a generalized crystal-field expansion of the wave functions and potential perturbations in partial waves is not merely a specialization of the familiar multicenter expansion technique to a single center. In prin-

principle, we are constructing a transformation on the multicentered wave-function expansion

$\sum_{\vec{R}_n} f_{\mu L}(\vec{r} - \vec{R}_n)$, characterized by an atomic angular momentum L (not a good quantum number) into a new basis $\sum_l F_{\mu l}(|\vec{r}|) K_l^{\alpha\lambda}(\hat{r})$, characterized by the systems angular momentum l (a good basis-set quantum number). Our central observation here is that only a few *partial wave components are needed to characterize the potential and density perturbations* $\Delta V(\vec{r})$ and $\Delta\rho(\vec{r})$ as well as the *perturbed wave functions in the central-cell region, but that each of these components requires for its description a large number of different radial basis functions*. This is illustrated in Sec. VI for the Si vacancy and in Sec. VII for the Cu impurity in silicon.

In the conventional, multicentered expansion technique, one places a few atomic orbitals $f_{\mu L}(\vec{r})$ on the sites surrounding the defect; the number of different radial functions associated with these is usually very small (e.g., $N=1$ for $L=1,2$ and $N=3$ for $L=0$ in Ref. 3). The basic variational coordinates manipulated in this approach are $\{f_{\mu L}(\vec{r})\}$. The shell orbitals produce a large number of high l components in the system's point-group representation, even though L is restricted to small values (e.g., the $L=0$ s orbitals on the nearest neighbors can contribute $L=0,3,4,6,7,8, \dots$, to a_1 representation in T_d). However, an overwhelming proportion of these l waves is variationally irrelevant since $\Delta V(\vec{r})$ and $\Delta\rho(\vec{r})$ physically contain contributions only from the lowest few l values and since the amplitude of the high- l components in the wave functions are significantly attenuated relative to the spherical component in the central-cell region. The remaining, variationally relevant, low-index partial waves then have only a very limited *radial* flexibility to describe the shape of $\Delta V(|\vec{r}|)$ and $\{\psi_i(\vec{r})\}$.

In the present generalized crystal-field expansion, we identify the radial l components $\Delta V_l(|\vec{r}|)$ [Eqs. (35) and (36)], $\Delta\rho_l(|\vec{r}|)$ [Eq. (33)], and $G_{ij}^\alpha(|\vec{r}|)$ [Eq. (29)] in $\Delta V(\vec{r})$, $\Delta\rho(\vec{r})$, and $\psi_i(\vec{r})$, respectively, as the fundamental independent variational coordinates of the problem. We therefore treat each l component *independently, in the system's angular-momentum representation* l , irrespective of the (physically irrelevant) parent-hood of each such partial wave in an atomic L state. Each such variational coordinate is then described independently by many radial-basis functions $F_{\mu l}(|\vec{r}|)$ (e.g., $N \leq 14$ for each l). We suggest that the present crystal-field representation for

localized defects is physically far more plausible and computationally much more efficient than the conventional multicentered representation since the former approach places the variational flexibility of the basis directly and explicitly where it is needed (i.e., in effect, projects the relevant pieces of the LCAO wave functions onto a single-site expansion) and, at the same time, avoids all multicenter integrals.

Previous experience with one-center expansion techniques in molecular chemistry⁵⁷⁻⁶² and solid-state applications^{63,64} has often indicated a slow convergence with respect to the number and type of basis orbitals. Examination of the results indicates, however, that such difficulties are related predominantly to the failure of a small one-center basis to reproduce *core orbitals* on neighboring sites and the placement of the expansion center on a bond, not an atomic site.^{62(b)} In fact, the popularity⁶² of single-center expansion techniques for calculation of the electronic structure of hydride molecules AH_n can be traced to the efficiency of the method to span nearest-neighbor orbitals of *coreless* atoms (hydrogen). In the QBCF method, however, we replace all core orbitals by pseudopotentials. We find that one-center expansions converge very rapidly as the off-center components of the pseudowave functions are, by construction, spatially very smooth. We suspect that a similar combination of a one-center expansion technique with pseudopotentials may prove very effective in molecular calculations.

4. Matrix elements in the crystal-field representation

The structure of our basis set $F_{\mu l}(|\vec{r}|) K_l^{\alpha\lambda}(\hat{r})$ and all potential and charge-density perturbations

$$\Delta(\vec{r}) = \sum_l \Delta_l(|\vec{r}|) K_l^{\alpha\lambda}(\hat{r})$$

in Eqs. (33)–(36) makes the calculations of all matrix elements nearly trivial—much like in atomic physics. For example, the matrix element of the external nonlocal pseudopotential perturbation is simply:

$$\begin{aligned} \langle g_{\mu l}^{\alpha\lambda}(\vec{r}) | \Delta V^{\text{ext}}(\vec{r}) | g_{\mu' l'}^{\alpha'\lambda'}(\vec{r}) \rangle \\ = \delta_{\alpha\alpha'} \delta_{\lambda\lambda'} \delta_{ll'} \langle F_{\mu l} | \Delta v_l^{\text{ext}} | F_{\mu' l'} \rangle, \end{aligned} \quad (37)$$

i.e., a one-dimensional radial integral, diagonal in all but the principal quantum number indices. Hence, the pseudopotential nonlocality that is

essential for describing transition elements⁸ and, at the same time, results in an involved multiple-scattering form when an overlapping multicenter representation is used, becomes diagonal in all but

the simple radial index μ . The matrix elements of the screening perturbation $\Delta V^{AB}(\vec{r})$ [$AB=ee$, or xc in Eq. 8(b)] is also given by a sum of one-dimensional radial integrals (nondiagonal in l),

$$\langle g_{\mu l}^{\alpha\lambda}(\vec{r}) | \Delta V^{AB}(\vec{r}) | g_{\mu'\lambda'}^{\alpha\lambda'}(\vec{r}) \rangle = \delta_{\alpha\alpha'} \delta_{\lambda\lambda'} \frac{1}{d_\alpha} \sum_{l''} D_{l''}^\alpha[l, l'] \langle F_{\mu l} | \Delta V_l^{AB} | F_{\mu' l'} \rangle, \quad (38)$$

where d_α is the dimensionality of the α th irreducible representation, and $D_{l''}^\alpha[l, l']$ are the Gaunt coefficients. Notice that the sum over l'' is determined by the triangular conditions $l + l' + l''$ which equals an even integer, and $|l - l'| \leq l'' \leq l + l'$. Further, the number of l'' is limited by the site symmetry of the defect. For a T_d site symmetry, for example, one has in the a_1 representation for the potential and density expansions of Eqs. (31)–(36) only the terms $l=0, 3, 4, 6, 7$, and 8 if $L_{\max}^{(1)}=4$. Similarly, t_2 orbitals contain only $l=1, 2, 3, 4, \dots$, e orbitals contain only $l=2, 4, \dots$, and t_1 orbitals include $l=3, 4, \dots$. The coefficients $D_{l''}^\alpha[l, l']$ are given in Table II for this group and $\alpha=a_1, e, t_1$, and t_2 , $0 \leq l', l \leq 4$, and $l''=0, 3, 4, 6, 7, 8$.

Our one-center, generalized crystal-field representation permits (i) easy inclusion of pseudopotential nonlocality, (ii) reduction of the self-consistency problem to one-dimensional integration [Eq. (35)] and fitting [Eq. (36)], (iii) the possibility to include simply and efficiently any set of radial functions $\{F_{\mu l}(|\vec{r}|)\}$. In particular, the physically appealing “exact” (numerical) impurity atomic orbitals or the quantum defect orbitals⁶⁵ are as easy to handle as the variationally less effective analytic orbitals (e.g., Gaussians), and (iv) direct effective cancellation of the difference between perturbed and unperturbed densities and potentials. The formalism lends itself to simple comparison with the empirical crystal-field model in which the expansion coefficients of the perturbation

are replaced by constants qD_l , which are adjusted

$$[\Delta V_L^{\text{ext}}(|\vec{r}|) + \Delta v_l^{\text{ee}}(|\vec{r}|) + \Delta v_l^{\text{xc}}(|\vec{r}|)]$$

are replaced by constants qD_l , which are adjusted

TABLE II. The Gaunt coefficients $D_{l''}^\alpha[l', l]$ for representations a_1, e, t_1 , and t_2 of the point-group T_d , defined in terms of the Kubic harmonics

$$\sum_{\lambda=1}^{d_\alpha} K_{l''}^{\alpha\lambda}(\hat{r}) K_l^{\alpha\lambda}(\hat{r}) = \sum_{l''} D_{l''}^\alpha[l', l] K_{l''}^{\alpha 1}(\hat{r}) \quad [\text{Eq. (E5)}]$$

l''	l', l	$D_{l''}^{a_1}[l', l]$	l', l	$D_{l''}^e[l', l]$	l', l	$D_{l''}^{t_1}[l', l]$	l', l	$D_{l''}^{t_2}[l', l]$
0	0,0	0.282 095	2,2	0.564 190	3,3	0.846 284	1,1	0.846 284
	3,3	0.282 095	4,4	0.564 190	4,4	0.846 284	2,2	0.846 284
	4,4	0.282 095					3,3	0.846 284
3	0,3	0.282 095			3,4	0.455 153	4,4	0.846 284
	3,4	-0.235 040					1,2	0.554 023
							1,4	-0.639 731
4	0,4	0.282 095	2,2	0.369 349	3,3	-0.117 520	2,3	-0.564 190
	3,3	-0.235 040	2,4	0.387 716	4,4	0.244 080	3,4	0.133 255
	4,4	0.162 720	4,4	0.046 491			1,3	0.564 190
6	3,3	0.201 176	2,4	0.263 35	3,3	-0.452 646	2,2	-0.369 349
	4,4	0.201 176	4,4	-0.321 881			2,4	0.290 787
							3,3	0.352 560
7	3,4	0.162 971			3,4	0.378 710	4,4	-0.453 292
	4,4	0.192 393	4,4	0.269 350	4,4	-0.461 744	2,4	-0.522 670
							3,3	0.251 470
8	3,4	0.162 971			3,4	0.378 710	4,4	0.150 882
	4,4	0.192 393	4,4	0.269 350	4,4	-0.461 744	3,4	-0.431 181
							4,4	0.0

from the spectroscopically observed splitting of the various defect levels.

C. The simplification of the \vec{k} -space summation problem

The calculation of the Green's function in Eq. (17) involving extensive BZ sampling (cf. Sec. IV B 3) is dramatically simplified by replacing the continuous representation of \vec{k} with a discrete representation \vec{k}_ν , $\nu=1 \cdots N_\Omega$ which corresponds to the center of the BZ of a supercell.^{29(a),(b)} This is illustrated below.

$$G_{ab}^0(\epsilon) = \sum_j \sum_{\vec{k}_\nu} w_\nu \frac{\langle g_a(\vec{r}) | \phi_j^{\text{QB}}(\vec{k}_\nu, \vec{r}) \rangle \langle \phi_j^{\text{QB}}(\vec{k}_\nu, \vec{r}) | g_b(\vec{r}) \rangle}{\epsilon - \epsilon_j^{\text{QB}}(\vec{k}_\nu)}. \quad (40)$$

Here the weight factors w_ν are given by $w_\nu = (N_\omega/N_\Omega)m_\nu$, and m_ν is the number of members in the star of \vec{k}_ν . (The factor N_ω/N_Ω comes from our assumption that the Bloch functions $|j\vec{k}_\nu\rangle$ are normalized to unity over the primitive unit cell ω .) Furthermore, since we are dealing with a truly discrete set of energies $\{\epsilon_j^{\text{QB}}(\vec{k}_\nu)\}$, we can assume that the energy spectrum of \hat{H} does not coincide with any of the elements in the set $\{\epsilon_j^{\text{QB}}(\vec{k}_\nu)\}$. We have therefore omitted the "small" imaginary number in the denominator of the usual Green's-function spectral representation.

In Table III we show the relation between the number of \vec{k}_ν points and the number N_Ω of atoms in the supercell for the fcc structure. It is clear from this table that a relatively small number of \vec{k}_ν points correspond to large supercells. In this method, the size of the Hamiltonian matrices involved does not depend on the size of the supercell; large supercells can be treated by using a reasonably small number of \vec{k}_ν points. Notice also that in the present representation all the poles of the Green's function are real. This is an important advantage over the Koster-Slater technique.¹⁻⁴

The problem of obtaining the energies $\epsilon_i(d = \infty)$ of gap levels for defects or surfaces at the limit of an isolated perturbing species has often been handled in supercell calculations by increasing the size of the supercell to the computer's capability limit.^{34,35} A simpler method that involves the calculation of $\epsilon_i(d)$ only for some *small* supercells is, however, possible. The dependence of such energy levels on the interdefect separation d results from the overlap of the exponentially decaying, gap-state

Let \vec{R}_p denote the primitive translations vectors of the host, and let f be an integer. The new translation vectors $\vec{T}_p = f\vec{R}_p$ generate a superlattice where each supercell Ω contains $N_\Omega = f^3 N_\omega$ atoms. Here N_ω denotes the number of atoms in the primitive unit cell. We then construct the reciprocal-lattice vectors \vec{k}_ν of the superlattice, i.e., we choose those \vec{k} -points in the irreducible section of the Brillouin zone (IBZ) that satisfy

$$\vec{k}_\nu \cdot \vec{r} = 2\pi I, \quad (39)$$

where \vec{k}_ν belongs to the IBZ and I is an integer. We use these points to perform the \vec{k} -space summation in the Green's-function matrix elements:

wave function with the potential from the nearest defects.^{29(a),45} Hence, one can estimate $\epsilon_i(\infty)$ from $\epsilon_i(d) = \epsilon_i(\infty) + ae^{-ad}$ by using data on $\epsilon_i(d)$ from three calculations with finite d values and solving for a value $\epsilon_i(\infty)$ that makes the plot $\ln[\epsilon_i(d) - \epsilon_i(\infty)]$ vs d linear. This method gives an excellent approximation to $\epsilon_i(\infty)$, as demonstrated in Secs. VID and VII C. It is a convenient tool for estimating the residual error in the supercell model.

D. The self-consistency problem

The angular-momentum components of the perturbation

TABLE III. Relations between the number of \vec{k} points used in the IBZ to the size of the supercell and interdefect separation in a silicon lattice (lattice constant $a = 10.2646$ a.u.).

Number of \vec{k} points	d Interdefect separation (a.u.)	N_Ω Number of atoms in supercell
3	14.5	16
4	21.8	54
10	36.3	250
20	50.8	686
35	65.3	1458
56	79.8	2662
84	94.4	4394

$$\Delta V_l^{\text{scr}}(|\vec{r}|) = \Delta V_l^{\text{ec}}(|\vec{r}|) + \Delta V_l^{\text{xc}}(|\vec{r}|)$$

in the screening potential [Eqs. (35) and (36)] can be constructed from any given densities $\rho_D(\vec{r})$ and $\rho_H(\vec{r})$, as described in Sec. VB 2. A self-consistent solution to the defect Green's function Eqs. (17) and (18) then requires that the density perturbation $\Delta\rho^{(m)}[\Delta V(\Delta\rho^{(m-1)})]$, obtained at the m th iteration from the previous $(m-1)$ th step, equals $\Delta\rho^{(m+1)}[\Delta V(\Delta\rho^{(m)})]$. If the values of $\Delta V_l^{\text{scr}}(|\vec{r}|)$ at K different radial points are used to construct a vector

$$\vec{\mu}_l = \{\Delta V_l^{\text{scr}}(r_1), \dots, \Delta V_l^{\text{scr}}(r_K)\},$$

the self-consistency problem reduces to finding the zeros of a vector function measuring the departure from potential self-consistency for all of the system's l components:

$$\vec{F}(\vec{\mu}_0, \dots, \vec{\mu}_L) = 0; \quad L = L_{\text{max}}^{(2)}. \quad (41)$$

This can be written in a compact form also as $\vec{F}(\vec{\mu}) = 0$, where $\vec{\mu}$ denotes a column vector made of all r_i points and l components. Previous approaches^{3,4,35} have attempted to solve this equation by constructing a guess $\vec{\mu}^{(m+1)}$ from a linear mixing of two previous steps, i.e.,

$$\vec{\mu}^{(m+1)} = \beta \vec{\mu}^{(m-1)} + (1-\beta) \vec{\mu}^{(m)},$$

where $\beta \leq 1$ is a mixing parameter. This procedure is not only not guaranteed to converge, but it also requires a large number of iterations when it does converge.

To solve Eq. (41), we use a Newton-Raphson approach introduced by Bendt and Zunger.⁵ In this approach, Eq. (41) is written (for the m th iteration) as

$$\vec{F}(\vec{\mu}) \simeq \vec{F}(\vec{\mu}^{(m)}) + \underline{J}^{(m)}(\vec{\mu} - \vec{\mu}^{(m)}) = 0, \quad (42)$$

where $\underline{J}^{(m)} = \partial \vec{F}(\vec{\mu}^{(m)}) / \partial \vec{\mu}^{(m)}$ is the Jacobian matrix. The right-hand side of Eq. (42) is then made zero by setting

$$\vec{\mu} = \vec{\mu}^{(m)} - [\underline{J}^{(m)}]^{-1} \vec{F}(\vec{\mu}^{(m)}). \quad (43)$$

At each iteration, $\vec{\mu}^{(m)}$ and its corresponding $\vec{F}(\vec{\mu}^{(m)})$ are generated; after a couple of iterations, an approximation \underline{J}_0 can be constructed for \underline{J} . This is done by Jacobian update method,⁶⁶ well known in optimization theory. Thus, the true Jacobian \underline{J} is not needed. Notice that approximating \underline{J}_0 by a constant is equivalent to simple potential mixing.^{3,4} More details on this general self-consistency method can be found in Ref. 5.

This method is particularly suited for the prob-

lem at hand for several reasons. First, this method does not require any new information; it uses only the quantities needed for the Green's-function calculation. Second, the method "remembers" information from all past iterations and is able to use this information effectively in constructing the best guess for $\{\mu_l(m)\}$. This property results in better guesses (i.e., more rapid convergence), than any method using only the most recent iterations, making it superior to potential mixing^{3,4,35} or the accelerations of Ferreira.⁵¹ A third advantage is that this method automatically discounts information from the distant past and hence it is not confused by nonlinearity as is the subspace inversion method advocated by Puley.⁶⁷

E. The mechanics of the solution

The input to the problem consists of $v_{\text{ps}}^{(L,H)}(r)$, $v_{\text{ps}}^{(L,D)}(r)$, the crystal structures $\{\vec{R}_p, \vec{\tau}_p, \vec{R}_I\}$, and the number of valence electrons in the host and impurity atoms. The functions that need to be chosen are (i) the host local quasiorbitals $f_j(\vec{r})$ in Eq. (26), and (ii) the radial local orbitals $F_{\mu l}(|\vec{r}|)$ in Eq. (27).

With this input the QBCF method provides: (i) the quasiband structure $\{\phi_j^{\text{QB}}(\vec{k}, \vec{r}); \epsilon_j^{\text{QB}}(\vec{k})\}$ in Eq. (25) (details in Appendix B), (ii) the elements

$$\langle g_{\mu l}^{\alpha \lambda}(\vec{r}) | G^0(\epsilon_i) | g_{\mu' l'}^{\alpha' \lambda'}(\vec{r}) \rangle$$

of the Green's-function matrix in Eq. (17a) (details in Appendix C), and (iii) the potential perturbations $\Delta V^{\text{ec}}, \Delta V^{\text{xc}}$ [Eqs. (35) and (36), details in Appendix E], and ΔV^{ext} [Eq. (34)] as well as the potential matrix elements

$$\langle g_{\mu l}^{\alpha \lambda}(\vec{r}) | \Delta V^{\text{ext}}(\vec{r}) | g_{\mu' l'}^{\alpha' \lambda'}(\vec{r}) \rangle$$

and

$$\langle g_{\mu l}^{\alpha \lambda}(\vec{r}) | \Delta V^{\text{ec}} + \Delta V^{\text{xc}} | g_{\mu' l'}^{\alpha' \lambda'}(\vec{r}) \rangle$$

in Eqs. (37) and (38), respectively. This step requires the knowledge of an initial guess to $\Delta V^{\text{ec}} + \Delta V^{\text{xc}}$. This does not determine the final result, but only the computational speed with which it is achieved. The wave functions and energies of the defected system are obtained from Eq. (17) through an efficient root-finding algorithm (Appendix F).^{68,69} The wave functions are then normalized (details in Appendix D) and the charge density $\rho_D(\vec{r})$ obtained through Eq. (31). Given the host charge density of the form in Eq. (32) the new screening potential is obtained through Eqs.

(33), (35), and (36). The process is iterated until the screening perturbations $\Delta V^{\text{scr}}(\vec{r})$ obtained in successive iterations do not differ by more than a prescribed tolerance. For convenience, we give in Appendix G the explicit expressions for the $l=4$ Kubic harmonics used in this work. The Kubic harmonics for $l=0, 1, 2$, and 3 were taken from Ref. 70.

The convergence parameters of this model are: (i) the number M_b of hostlike bands $\{\phi_j^{\text{BS}}(\vec{k}, \vec{r})\}$ in Eq. (25), (ii) the number M_a of local Bloch functions $\{\chi_j(\vec{k}, \vec{r})\}$ in Eq. (25), (iii) the total number N of local orbitals $\{g_{\mu l}^{\alpha}(r)\}$ in Eq. (28) with a maximum angular momentum of $L_{\text{max}}^{(1)}$ (iv) the highest angular momentum $L_{\text{max}}^{(2)} = 2L_{\text{max}}^{(1)}$ kept in the density and potential expansion [Eqs. (31)–(33) and (35) and (36)], (v) the number f of \vec{k}_v points used to sample the Green's function [Eq. (40), related to the size of the supercell, cf. Table III], and (vi) the admissible self-consistency error

$$\sigma = N^{-1} \sum_{i=1}^N |V_{ii}^{\text{scr}}(\text{output}) - V_{ii}^{\text{scr}}(\text{input})| .$$

We will show in the next section that it is possible to find in the parameter space $\{M_a, M_b, N, L_{\text{max}}^{(1)}, L_{\text{max}}^{(2)}, f, \sigma\}$ values that produce a high precision in the results for deep defect levels and yet require only a modest computational effort. With this precision we can assume that the results reflect, to within a good approximation, the physical input to the problem and the underlying density-functional formalism used to describe the interelectronic interactions.

VI. THE UNRELAXED SILICON VACANCY

We first illustrate our method for calculating the electronic properties of deep defects on a simple case previously treated in the literature^{3,4,35}—the unrelaxed silicon vacancy. We will not focus on the discussion of the physical properties of the silicon vacancy, since this has been previously done extensively.^{3,4,27} Instead, we intend to establish the convergence properties of the new QBCF method on a known system, before applying it to new problems (Sec. VII). We therefore use for the host-crystal pseudopotential $V_H^{\text{ext}}(\vec{r})$ [Eq. (6)] the same local pseudopotential used in Refs. 3 and 35. The screening function used is given by the density-functional approach [Eq. (8b)]. In Refs. 3, 4, and 35, however, the exchange coefficient used for $\epsilon_x[\rho]$ was set to $\alpha_x = 0.8$. This produces a rather poor agreement between the calculated band

structure and experiment, e.g., the indirect band gap is 0.6 eV instead of the experimental value of 1.15 eV, and the $\Gamma_{25,v} \rightarrow \Gamma_{15,c}$ transition is at 2.8 eV, instead of the experimental value of 4.18 eV. This discrepancy is partially fixed in the calculation of Ref. 4 by using an imperfect fit of the band structure to a localized basis: The fitting deviations then raise the band gap to a value of 0.97 eV. In fact, the semiempirical pseudopotential of Refs. 3 and 35 was designed to reproduce the experimental results for Si only if an exchange parameter of $\alpha_x = 1$ is used.⁷¹ Since the defect energy levels of primary interest lie in the narrow band gap, it is felt that a physically accurate description of the host band gap is essential. Using $\alpha_x = 1$ in a self-consistent band-structure calculation with an energy cut-off value of 12.1 Ry (181 plane waves at Γ), we get an indirect band gap of 1.2 eV and an overall band structure that is in very good agreement with optical and photoemission data. The perturbing external potential $\Delta V^{\text{ext}}(\vec{r})$ [Eq. (9a)] in the case of an unrelaxed vacancy equals $-v_{\text{ps}}^{(H)}(r)$. Using this input alone, we calculate the wave functions and energy levels of the vacancy self-consistently. We examine below the convergence properties of the QBCF method with respect to the parameters described in Sec. V E. In the simple case of a silicon vacancy no quasibands are needed to attain convergence with respect to bands, hence $M_a = 0$ in Eq. (25).

A. Convergence with respect to host band wave functions

Table IV shows the convergence of the energy of the t_2 gap level with respect to the number M_b of host-crystal band wave functions $\{\phi_j^{\text{BS}}(\vec{k}, \vec{r})\}$ used in the expansion of Eq. (25). The results are illustrated for different sizes of the basic supercell, ranging from $N_{\Omega} = 54$ to $N_{\Omega} = 2662$ atoms per supercell. Also shown is the orbital-density localization,

$$q_i^{\alpha} = \sum_l q_{il}^{\alpha} = \sum_l N_l \int_0^{R_c} |G_{il}^{\alpha}(r)|^2 r^2 dr , \quad (44)$$

measuring the amount of charge enclosed within a sphere of radius (nearest neighbor) $R_c = 4.44$ a.u. in the level $|i\alpha\rangle$. This quantity includes a sum over all the l components that contribute to this level. We refer to q_i^{α} as the orbital-density localization parameter. Notice that q_i^{α} can not exceed 2.0,

TABLE IV. Convergence of the energy ϵ_{t_2} (in eV) and orbital-density localization $q_i^{t_2}$ of Eq. (44) (in electrons) of the silicon vacancy gap level with respect to the number M_b of host-crystal bands and the number N_Ω of atoms in a supercell. Also shown is the smallest distance d between adjacent vacancies in neighboring supercells. In Eq. (44), $R_c=4.44$ a.u. (the nearest-neighbor distance). The zero of energy is the valence-band maximum. The basis set consists of 14 radial Coulombic functions with $Z^*=10$ for each of the l components used ($l=1,2,3,4$). The maximum value q can attain equals the occupation of the t_2 level, i.e., 2.0.

Interdefect distance		N_Ω Atoms/cell	$M_b=10$ bands		$M_b=21$ bands		$M_b=30$ bands	
d (a.u.)	ϵ_{t_2}		$q_i^{t_2}$	ϵ_{t_2}	$q_i^{t_2}$	ϵ_{t_2}	$q_i^{t_2}$	
21.8	54	1.227	0.48	1.061	0.66	1.025	0.70	
36.3	250	0.906	0.66	0.715	0.70	0.682	0.70	
50.8	686	0.817	0.66	0.632	0.70	0.602	0.72	
65.3	1458	0.802	0.66	0.610	0.70	0.581	0.72	
79.8	2662	0.798	0.66	0.604	0.70	0.575	0.72	

whereas q_i^e and $q_i^{t_2}$ can not exceed 4.0 and 6.0, respectively, for fully occupied levels and 2.0 for the t_2 gap level.

For the largest supercell considered, the gap state is converged to within only 0.22 eV (or 20% of the band gap) if 10 host bands are used (four valance bands and six conduction bands). The convergence error reduces to 2.5% when 21 host bands are used. For 25 bands the error is 1% with respect to a further increase in the number of bands. The orbital-density localization parameter q_i^a converges much faster than the energy. The upper a_1 resonance at -1.09 eV (Table V) converges to within 0.1 eV of its correct value when

TABLE V. Convergence of the self-consistent energy ϵ_{a_1} (in eV) and orbital-density localization parameter $q_i^{a_1}$ of Eq. (44) (in electrons) of the silicon vacancy upper a_1 resonance with respect to the number M_b of host-crystal bands. The zero of energy is the valence-band maximum. The basis set consists of 14 radial Coulombic functions with $Z^*=10$ for each of the l components ($l=0,3,4$). A supercell of 250 atoms is used. $R_c=4.44$ a.u. (the nearest-neighbor distance) is used in Eq. (44). The maximum value that q can attain equals the occupation number, i.e., 2.0.

$M_b=10$ bands		$M_b=21$ bands		$M_b=30$ bands	
ϵ_{a_1}	$q_i^{a_1}$	ϵ_{a_1}	$q_i^{a_1}$	ϵ_{a_1}	$q_i^{a_1}$
-0.985	0.81	-1.086	0.82	-1.090	0.82

10 host bands are included; about 20 bands are needed to reduce the convergence error to below 0.005 eV.

The rather slow convergence with respect to the number of host bands is clearly reflected in the spectral decomposition of the t_2 gap-state wave function in terms of host-crystal wave functions. Shown in Fig. 2 are the expansion coefficients $A_{ij}^{t_2}(\vec{k})$ of Eq. (13) of $\psi_i^{t_2}(\vec{r})$ in terms of the Bloch states $\phi_j^{\text{BS}}(\vec{k}, \vec{r})$. The normalization condition is such that

$$\sum_j \sum_{\vec{k}_v \in \text{IBZ}} d_v |A_{ij}^{t_2}(\vec{k}_v)|^2 = 1,$$

where d_v is the number of vectors in the star of \vec{k}_v . The spectral decomposition is given at the Γ point, as well as for three other points in the (0,0,1), (0,1,1), and (1,1,1) directions. Clearly, an acceptor state such as the vacancy t_2 level is composed not only from valence-band states, but contains sizable contributions also from the conduction bands. In fact, even high-energy conduction-band states at ~ 25 eV above the valence-band maximum have sufficient t_2 character to couple into the t_2 gap state. More specifically, summing the contributions over the Brillouin zone

$$P_j^{t_2} = \sum_{\vec{k}_v \in \text{IBZ}} d_v |A_{ij}^{t_2}(\vec{k}_v)|^2,$$

we find that this generalized Mulliken charge is 0.004, 0.0425, 0.266, and 0.490 for the four valence bands, respectively (i.e., 80% of the t_2 wave func-

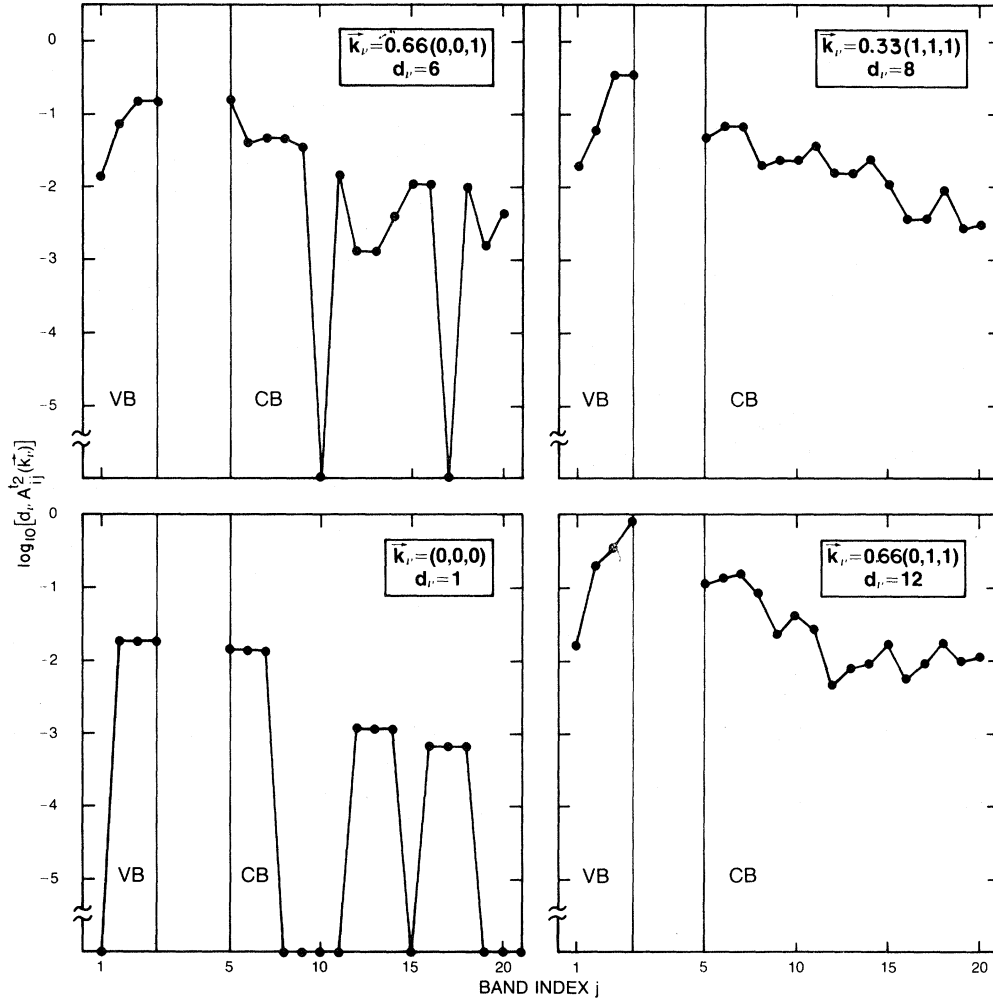


FIG. 2. Spectral decomposition of the t_2 gap-state wave function $\psi_i^{t_2}(\vec{r})$ of the silicon vacancy in terms of host-crystal Bloch wave functions $\phi_j^{\text{BS}}(\vec{k}, \vec{r})$. The coefficients are given by $\psi_i^{t_2}(\vec{r}) = \sum_j \sum_{\vec{k}_v} A_{ij}^{t_2}(\vec{k}_v) \phi_j^{\text{BS}}(\vec{k}_v, \vec{r})$ and are normalized such that $\sum_j \sum_{\vec{k}_v} d_v |A_{ij}^{t_2}(\vec{k}_v)|^2 = 1$. Here d_v is the degeneracy factor of the star of \vec{k}_v . \vec{k}_v is given in units of $2\pi/a$.

tions are derived from the valence band), whereas the conduction bands contribute a Mulliken charge of 20%. This contribution, however, couples effectively about 20 conduction bands.

B. Convergence with respect to local orbitals

We next examine the convergence of the defect energies and wave functions with respect to the number (N) and type of the local orbitals $\{g_{\mu l}^{\alpha\lambda}(\vec{r})\}$ used in the expansion of Eq. (28). We demonstrate the results for two types of local orbitals. The harmonic oscillator (HO) basis has a radial part given by

$$F_{\mu l}(|\vec{r}|) = \left[\frac{2K\sqrt{K}k!}{\Gamma(k+l+\frac{3}{2})} \right]^{1/2} (Kr^2)^{l/2} \times e^{-Kr^2/2} L_k^{l+(1/2)}(Kr^2), \quad (45)$$

where $k = \frac{1}{2}(\mu - l)$ are integers, and $L_q^p(x)$ are the associated Laguerre polynomials. The nonlinear parameter is given by K , and $\Gamma(p)$ denotes the gamma function. The Coulomb orbitals have a radial part given by the orthogonal basis [cf. A function, Ref. 62(b)]:

$$F_{\mu l}(|\vec{r}|) = \left[\alpha^3 \frac{(\mu-l-1)!}{2\mu(\mu+l)!} \right]^{1/2} (\alpha r)^l \times e^{-\alpha r/2} L_{\mu-l-1}^{2l+1}(\alpha r), \quad (46)$$

where $\alpha = 2Z^*/\mu$ and Z^* is the effective nuclear charge.

The two types of basis sets have very different spatial properties (e.g., cusp behavior and localization), yet Fig. 3 shows that the energy levels of the t_2 and a_1 vacancy levels obtained with these basis functions agree to within 0.01 eV. Notice that we use the same number N of radial orbitals for *each* of the angular-momentum components of the wave function ($l=1,2,3,4$ for t_2 and $l=0,3,4$ for a_1). This basis has therefore an enormous radial variational flexibility: In multicenter expansions used previously,^{3,4} more typically only one to three different radial functions are allowed for each atomic angular orbital. Figure 3 shows that for $N=6-8$ a convergence of better than 0.1 eV is obtained. Optimization of K or Z^* in Eqs. (45) and (46), respectively, allows us to reduce this number to 4-5. Similarly, even a small number of chemically and physically suitable orbitals [e.g., $G_{ii}^{\alpha}(r)$ in Eq. (29) or quantum-defect orbitals⁶⁵] can be chosen.

However, in the present work we will not describe the details of such optimizations since the computational efficiency of the present method allows one to use larger N values with only moderate extra cost.

Since the defect wave functions are obtained in the present method analytically, we are able to provide the final self-consistent wave function compactly. Table VI gives the expansion coefficients $\{C_{\mu li}^{\alpha}\}$ in Eq. (28) for the $\alpha=t_2$ and $\alpha=a_1$ vacancy states, in terms of the HO basis orbitals. These analytic orbitals may be used to calculate photoionization cross sections, position annihilation profiles, etc.

C. Convergence with respect to angular momenta

Figures 4 and 5 depict the decomposition of the wave functions of the t_2 gap state and the upper a_1 resonance in terms of the radial l components $G_{ii}^{\alpha}(|\vec{r}|)$ of Eq. (29). While the lowest two allowed l components ($l=1,2$ for t_2 and $l=0,3$ for a_1) are significant, an overall attenuation of the amplitude of higher l components is apparent. This is illustrated in Figs. 6 and 7 that show the wave functions of these two states, plotted in the $\pm\langle 111 \rangle$ crystal directions and calculated with

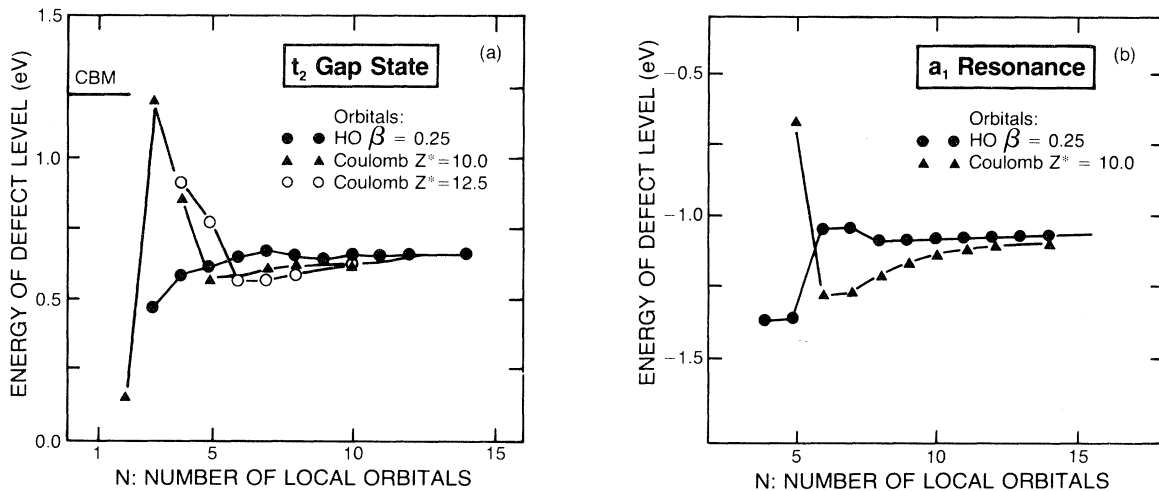


FIG. 3. Convergence of the energies of the Si vacancy t_2 gap state (a) and the upper a_1 resonance (b) with respect to the number N and type of local orbitals used in the wave-function expansion [Eqs. (27) and (28)]. The zero of energy is at the valence-band maximum; 21 host bands are used in the expansion of Eq. (13). N local orbitals are used for *each* of the l components ($l=1,2,3,4$ for the t_2 state and $l=0,3,4$ for the a_1 state). The total number of orbitals is therefore $\sum_l N_l(2l+1)$ (i.e., for $N=10$ this includes 170 orbitals for a_1 and 240 orbitals for t_2). A supercell of 686 and 250 atoms, respectively, is used for (a) and (b). The harmonic oscillator (HO) orbitals [Eq. (45)] have an exponent $K=0.25$, while the two sets of Coulomb orbitals [Eq. (46)] have exponents of $Z^*=10$ and $Z^*=12.5$.

TABLE VI. Expansion coefficients [Eq. (28)] for the self-consistent wave functions of the upper a_1 resonance and the t_2 gap state for the silicon vacancy in terms of harmonic oscillator basis functions [Eq. (45)] with $K=0.25$. This expansion is valid in the domain of $\Delta V(\vec{r})$ only, i.e., $r \lesssim 5$ a.u. The numbers in parentheses are the exponents for the corresponding expansion coefficient [e.g., $-0.4544 (-2)$ means -0.4544×10^{-2}]. The expansion coefficients are normalized to unity, i.e., $\sum_{\mu} \sum_l |C_{i\mu l}^{\alpha}|^2 = 1$, and the "normal" factors at the bottom of the table are the factors with which the expansion coefficients have to be multiplied to give the normalization over all space.

k	Expansion coefficients for the upper a_1 resonance			Expansion coefficients for the t_2 gap state			
	$l=0$	$l=3$	$l=4$	$l=1$	$l=2$	$l=3$	$l=4$
0	0.7278	-0.4544 (-2)	0.1262	0.6435	0.3699	-0.1220 (-1)	0.9778 (-1)
1	-0.1335	0.3587	-0.2241	0.7900 (-1)	0.2124	-0.2124	-0.2246
2	-0.4181	-0.7631 (-1)	0.2676 (-1)	-0.3980	-0.2720	0.9028 (-1)	0.3823 (-2)
3	0.2336	-0.6184 (-1)	0.3723 (-1)	0.1563	0.7553 (-1)	0.2270 (-1)	0.3629 (-1)
4	-0.6776 (-2)	0.3479 (-1)	-0.1813 (-1)	0.6815 (-1)	0.5951 (-1)	-0.3434 (-1)	-0.7637 (-2)
5	-0.1109	0.2234 (-1)	-0.1778 (-1)	-0.7623 (-1)	-0.3934 (-1)	-0.7878 (-2)	-0.1587 (-1)
6	-0.2865 (-2)	-0.1056 (-1)	0.6760 (-3)	-0.3011 (-1)	-0.2573 (-1)	0.1307 (-1)	-0.2729 (-2)
7	0.3080 (-1)	-0.1370 (-1)	0.5491 (-2)	0.2221 (-1)	0.1029 (-1)	0.7947 (-2)	0.5136 (-2)
8	-0.4130 (-2)	-0.4074 (-2)	0.2031 (-2)	0.1155 (-1)	0.1045 (-1)	-0.5746 (-3)	
9	-0.1890 (-1)	0.4191 (-3)	0.4076 (-3)	-0.1013 (-1)	-0.3278 (-2)	-0.1511 (-2)	
10	-0.5135 (-2)	-0.3820 (-3)	0.1413 (-2)	-0.1012 (-1)	-0.6293 (-2)	0.8733 (-3)	
11	0.6898 (-2)	-0.1275 (-2)	0.1935 (-2)	0.8833 (-3)	-0.1647 (-3)	0.1527 (-2)	

Normal factor 0.7883 Normal factor 0.7326

various truncation of the angular-momentum sum in Eq. (29). Clearly, retaining only the lowest l component introduces significant errors. In particular, the wave function of the a_1 resonance is symmetric with respect to reflection at the origin (Fig. 6) if only the $l=0$ component is retained, whereas the accurate wave function (including $l=0,3,4,\dots$) has a higher amplitude in the bonding $\langle 111 \rangle$ region than in the antibonding $-\langle 111 \rangle$

region. However, in the central-cell region the contributions of the higher l components decreases rapidly as l increases.

Notice that both these dangling-bond wave functions are nodeless in the central-cell region (i.e., within a radius of the nearest-neighbor distance) and have their maximum amplitude at 65% and 57% of the distance to the nearest neighbor for the t_2 and a_1 states, respectively. Yet, only 36% and

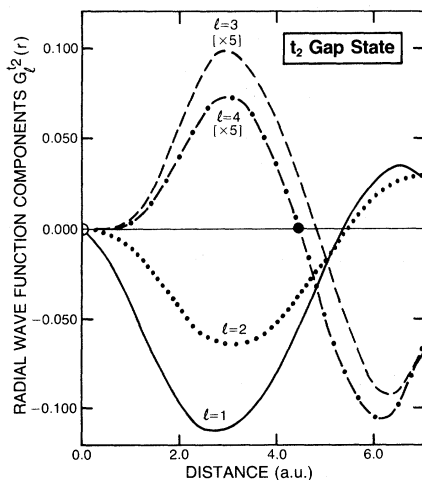


FIG. 4. Radial components $G_{l\mu}^{\alpha}(|\vec{r}|)$ [Eq. (29)] of the Si vacancy t_2 gap level wave function.

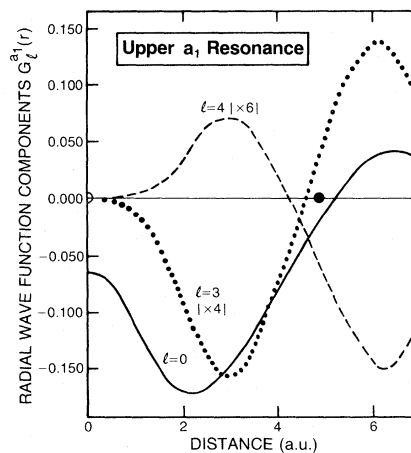


FIG. 5. Radial components $G_{l\mu}^{\alpha}(|\vec{r}|)$ [Eq. (29)] of the Si vacancy upper a_1 resonance wave function.

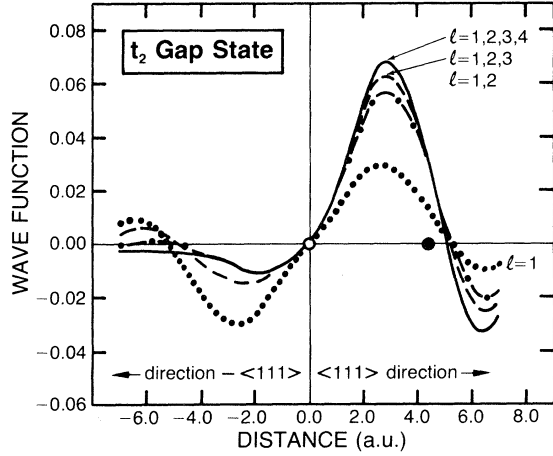


FIG. 6. Wave function of the t_2 gap state of a silicon vacancy, shown in the $\langle 111 \rangle$ and $-\langle 111 \rangle$ directions. The origin is at the vacancy site (open circle). The nearest-neighbor silicon atom is indicated as a full circle. Shown are the calculated wave functions using different sizes of truncated l series in Eq. (28). —: $l=1,2,3,4$. ---: $l=1,2,3$. -·-·: $l=1,2$. ···: $l=1$. 21 host bands, 14 harmonic oscillator orbitals with $K=0.25$, and a supercell of 686 atoms are used.

41% of the orbital densities of these wave functions are enclosed in the central cell (c.f. Tables IV and V), whereas the remaining 64% of the t_2 wave function and 59% of the a_1 charge is extended beyond the first-nearest-neighbor shell. As indicated previously,^{3,4} this behavior highlights the difficulty to obtain reliable solutions to the deep defect problem using small clusters, small supercells,³⁵ a simple, ligand-field model,^{16–18} or a defect molecule model.²⁰

The relative contributions of the nonspherical components are further reduced in taking the sum of squares of these wave functions (i.e., in obtaining ρ_D). For example, the contribution of the $G_3^{a_1}(r)$ term in the wave function to the $l=0$ component of $\rho_D(\vec{r})$ is $(1/4\pi)[G_3^{a_1}(r)]^2$; not only is the square of a small $G_3^{a_1}(r)$ even smaller, but also the projection of it onto the $l=0$ density (resulting in the factor $1/\sqrt{4\pi}$) reduces it further (e.g., in the example of Fig. 5, the total reduction factor is 60). Hence the contribution of high- l components in the wave functions to the low- l components of the density is small. Furthermore, in calculating the change in the charge density $\Delta\rho(\vec{r})=\rho_D(\vec{r})-\rho_H(\vec{r})$, the contributions of the high- l components of the wave functions to the high- l components of $\Delta\rho(\vec{r})$ is canceled effectively since the

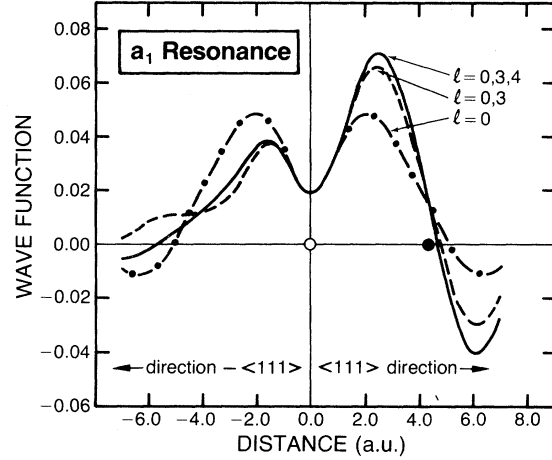


FIG. 7. Wave function of the a_1 resonance at -1.1 eV of a silicon vacancy, plotted in the $\langle 111 \rangle$ and $-\langle 111 \rangle$ directions. Origin at the vacancy site (open circle). The nearest-neighbor silicon atom is indicated by the full circle. Shown are the calculated wave functions using different sizes of truncated l series in Eq. (28). —: $l=0,3,4$. ---: $l=0,3$. -·-·: $l=0$. 21 host bands, 14 harmonic oscillator orbitals with $K=0.25$, and a supercell of 250 atoms are used. Notice that in the $l=0$ approximation the wave function is symmetric around the origin whereas the accurate wave function ($l=0,3,4$...) is not.

host crystal and the defected crystal have similar high- l components in the central-cell region. Figure 8 shows the l components $\Delta\rho_l(r)$ of the self-consistent density perturbation. As expected from our preceding discussion, its nonspherical components are small. We also present in this figure two different calculations of $\Delta\rho_{l=0}(r)$: one using an harmonic oscillator basis set [Eq. (45)] for the wave-function expansion (full line in Fig. 8) and one using Coulombic basis orbitals (dashed line in Fig. 8) [Eq. (46)]. The very close agreement in $\Delta\rho_0(r)$ obtained from these very different basis sets confirms the excellent convergence properties of the model. Figure 9 shows the two lowest l components $\Delta V_l(r)$ of the total potential perturbation. Although the $l=3$ component cannot be neglected, higher components are negligible.

The fact that the perturbations $\Delta V(\vec{r})$ and $\Delta\rho(\vec{r})$ are nearly spherical does not imply that a spherical atomlike model for the defect problem is adequate, since the defect wave functions correspond to the Hamiltonian $-\frac{1}{2}\nabla^2 + V_H(\vec{r}) + \Delta V(\vec{r})$, not to $-\frac{1}{2}\nabla^2 + \Delta V(\vec{r})$. This is illustrated in Fig. 10 that shows the l components of the host screening $V_H^{\text{scr}}(\vec{r})$, the host pseudopotential $V_H^{\text{ext}}(\vec{r})$ (both are

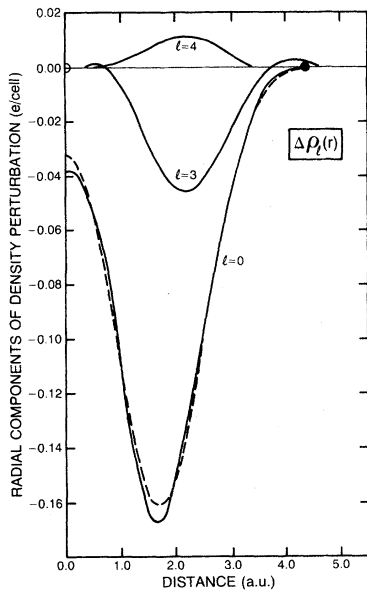


FIG. 8. Three lowest angular-momentum components of the self-consistent density perturbation $\Delta\rho(\vec{r})$ [Eq. (33)] for the silicon vacancy. See text for discussion.

given with respect to the average potential, i.e., the $\vec{G}=0$ Fourier component), and the screened host pseudopotential $V_H(\vec{r})$. As seen, the high- l components of the host potential have large amplitudes near the central-cell boundaries. We emphasize here that the existence of high- l components in

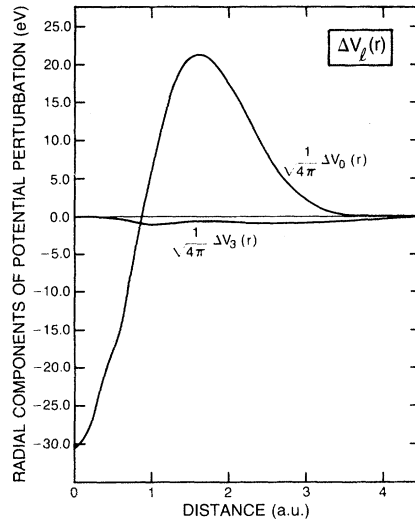


FIG. 9. Two lowest angular-momentum components of the self-consistent potential $\Delta V^{\text{ext}}(\vec{r}) + \Delta V^{\text{cc}}(\vec{r}) + \Delta V^{\text{xc}}(\vec{r})$ for the silicon vacancy. The exchange parameter used is $\alpha_x=1$.

$V_H^{\text{scr}}(\vec{r})$ and $V_H^{\text{ext}}(\vec{r})$ results from the multicenter nature of the crystal, not from pseudopotential nonlocality—a local (i.e., $L=0$) pseudopotential³⁵ is used here. The spatial anisotropy of $V_H(\vec{r})$ is then reflected in the nonspherical components of the defect wave functions. It is hence necessary to construct a model for the wave functions that in-

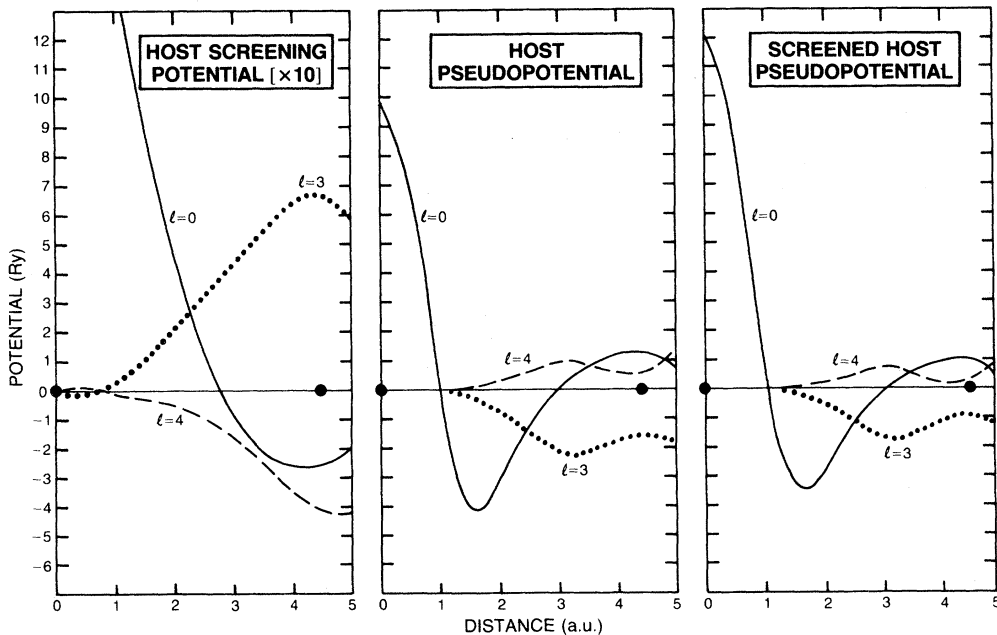


FIG. 10. Angular-momentum decomposition of the self-consistent screening potential, local pseudopotential, and screened local pseudopotential of the Si crystal.

cludes the lowest nonspherical components even though $\Delta V(\vec{r})$ and $\Delta\rho(\vec{r})$ are nearly spherically symmetric. The fact that only the first few l components of the wave function need to be retained in a Green's-function approach is illustrated in Table VII. It illustrates the convergence of the energies and orbital-density localization parameters of the t_2 and a_1 vacancy levels with respect to the number of angular-momentum components $L_{\max}^{(1)}$ included in the wave-function expansion of Eq. (28). If one restricts the wave-function representation to the single lowest l component (i.e., s orbitals for the a_1 states and p orbitals for the t_2 state), the resulting errors in the corresponding energy levels are as high as 10% and 20% of the band gap, respectively. Including, however, the next highest l component (f orbitals for a_1 and d orbitals for t_2) reduces these errors to 2%.

The results presented in this section highlight the physical basis of our generalized crystal-field model: While the potentials of the host crystal and the defect system may include high nonspherical components due to the multicentered nature of these systems, the potential perturbation $\Delta V(\vec{r})$ is nearly spherically symmetric.^{3,4} Furthermore, the one-electron energy spectrum is determined only by the first few l components of the wave functions in the central-cell region. These conclusions, established through modern and precise calculations, highlight the ingenuity of Bethe's crystal-field model, which dates to 1929.

D. Convergence with respect to supercell size

The convergence of our results with respect to the number of \vec{k} points used to construct the Green's function in Eq. (17c) is illustrated in Table IV. Since in the present method the number of \vec{k} points is uniquely related to the number of atoms N_Ω in a supercell, we show only the latter in Table IV (4, 10, 20, 35, and 56 \vec{k} points yield N_Ω values of 54, 250, 686, 1458, and 2662 atoms per supercell, respectively). Clearly, even a modestly small number of \vec{k} points already corresponds to very large supercells. Using a supercell of 54 atoms only,³⁵ one obtains a t_2 defect energy level that is 0.45 eV too high (i.e., 40% of the band gap). The error drops to 10% for a supercell containing 250 atoms and to 2% and 0.5%, respectively, for a cell containing 686 and 1458 atoms.

As suggested in Sec. V D, it is possible to estimate how well the gap level is converged with respect to N_Ω by using only data from few calculations on small supercells. For example, taking from Table IV the values $\epsilon_1=1.025$ eV, $\epsilon_2=0.682$ eV, and $\epsilon_3=0.602$ eV, corresponding to interdefect separations of $d_1=21.8$ a.u., $d_2=36.3$ a.u., and $d_3=50.8$ a.u., respectively, one can solve for the constant $\epsilon(\infty)$ that makes $\ln[\epsilon(d)-\epsilon(\infty)]$ vs d linear. This immediately yields (in eV) for the gap level

$$\epsilon(d)=0.575+3.7e^{-0.097d}.$$

TABLE VII. Errors in the a_1 and t_2 orbital energies (in eV) and orbital-density localization parameters (electrons) resulting from using a truncated angular-momentum series in the wave-function expansion [Eq. (28)]. 250 and 686 atoms per supercell are used for the a_1 and t_2 states, respectively. 14 Coulomb local orbitals with $Z^*=10$ are used for each l component. 30 host-crystal bands are used in the expansion of Eq. (13).

Angular momenta retained	Error in orbital energy (eV)		Error in localization parameter (e)
		a_1 resonance	
Using $l=0$	0.103		0.04
Using $l=0,3$	0.027		0.00
Using $l=0,3,4$	<0.01		0.00
		t_2 gap state	
Using $l=1$	0.200		0.304
Using $l=1,2$	-0.020		0.020
Using $l=1,2,3$	0.005		0.006
Using $l=1,2,3,4$	<0.000		0.000

This describes the three energy levels within a remarkable accuracy of 0.3% and correctly predicts the large supercell limit of $\epsilon=0.575$ eV (c.f. Table IV). An accuracy of 0.1 eV ($\sim 10\%$ of the band gap) in the energy level hence requires a calculation with a supercell with $d=37$ a.u. (~ 250 atoms), whereas an accuracy of 0.01 eV requires $d=61$ a.u. (~ 1200 atoms). This highlights the advantage of the present supercell method relative to direct diagonalization supercell approaches^{33–35} which are presently limited to ~ 50 atoms.

In practice, we first achieve self-consistency in the potential perturbation $\Delta V(\vec{r})$ for a small supercell and then use the results for a few additional self-consistency iterations with a larger cell size. This saves considerable computation time. Table IV also illustrates the variation in the defect energy level with the concentration of the (periodic) vacancies in the solid. At a very high concentration of 1.8% vacancies the t_2 level at the zone center appears just 0.2 eV below the conduction-band minimum (CBM), while as the concentration decreases to 0.4%, 0.14%, and 0.04%, the level drops to 0.53, 0.61 and finally to 0.635 eV, respectively, from the CBM.

E. Convergence in the self-consistency iterations

The measure used for the quality of self-consistency is

$$\sigma_l^{(m)} = \frac{1}{K} \sum_{i=1}^K |\Delta V_l^{(m-1)}(r_i) - \Delta V_l^{(m)}(r_i)|, \quad (47)$$

where m and $m-1$ refer to iteration numbers. If one uses the conventional mixing of input and output^{3,4,35} and a starting screened atomic potential, about 15 iterations are required to bring $\sigma_0^{(m)}$ to 10 mRy. However, using the present Newton-Raphson method with Jacobian updates⁵ it takes two to three iterations for the Jacobian to “learn” the sensitivity of the output potential to given changes in the input; each of the subsequent two iterations lowers $\sigma_l^{(m)}$ by an order of magnitude. Our final $\sigma_0^{(m)}$ after five iterations is 0.9 mRy. The largest deviation $\Delta V_{l=0}^{(m)}(\vec{r}) - \Delta V_{l=0}^{(m-1)}(\vec{r})$ is 1.4 mRy at $r=0.36$ a.u. This is a better convergence than that attained in previous studies.

F. Comparison of results with previous studies

In this section we briefly compare our results for the unrelaxed silicon vacancy with previous

ab initio self-consistent results. The calculation of the unrelaxed silicon vacancy reported here was done with the same pseudopotential used in Ref. 3 (which differs somewhat from that used in Ref. 4). As discussed in Sec. VI, we use, however, an exchange coefficient of $\alpha_x=1$ to obtain a physically correct host-crystal band structure, whereas in Refs. 3 and 4 an exchange coefficient of $\alpha_x=0.8$ has been used. This effect can be estimated. We show in Fig. 11 the $l=0$ component of our final self-consistent perturbation potential, together with that obtained in Ref. 3, using a completely independent computational technique, but employing the same physical inputs (pseudopotential and lattice constant). We have scaled the exchange contribution calculated in Ref. 3 by $1/0.8$ to take into account in an approximate manner the different exchange parameters used (this, however, makes little difference). The agreement between the two calculations is excellent, not only in the bond region ($\pm 0.3\%$ deviation) but also near the origin (1.5% difference). The agreement is far better than that obtained by Jaros, Rodriguez, and Brand⁷² in which the self-consistency procedure was approximated by a first-order perturbation correction using spherical potentials and a two-atom supercell.

The calculations of Refs. 3 and 4 do not report convergence tests with respect to the number of host bands and type and number of local orbitals. It is difficult to assess this aspect of their precision. Taken at face value, the defect energy levels of Ref. 3 (0.70 and -1.1 eV for the upper t_2

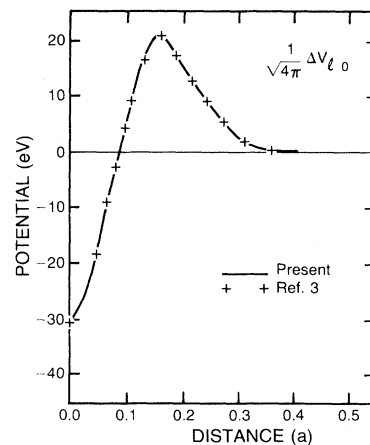


FIG. 11. Comparison of the $l=0$ component of the present self-consistent perturbation potential (—) with the calculation of Ref. 3 (+ + +).

and a_1 levels) agree closely with the present results (0.575 and -1.09 eV, respectively). The deep a_1 level obtained in Ref. 3 at -8.4 eV appears in the present calculation as a fairly broad a_1 resonance at -8.0 ± 0.3 eV. No line plots of the defect wave functions are given in Refs. 3 and 4; it is difficult to use their contour plots for a quantitative comparison with better than 10% precision. We find that the maximum amplitude in their t_2 and upper a_1 wave function occur at 53% and 68% of the distance to the nearest-neighbor atom in the $\langle 111 \rangle$ direction, while our results yield 57% and 65%, respectively. We conclude that the two sets of results are in very good agreement.

VII. SUBSTITUTIONAL COPPER IN SILICON

In this section we present the first calculation of a transition-atom impurity (i.e., retaining its s - p as well as d orbitals) in an extended (periodic) semiconductor. In Secs. VII A – VII D we illustrate how the basic four ideas underlying the QBCF method work. This discussion parallels the description of the four basic problems encountered in the description of deep defects in solids (Secs. IV B 1 – IV B 4), as well as the discussion in Secs. V A – V D of our present resolution of these difficulties. Finally, in Sec. VII E we discuss the physical properties of the substitutional Cu impurity.

In the present application to Si:Cu we use the same silicon host band structure and pseudopotential^{3,35} $V_H^{\text{ext}}(\vec{r})$ as we used for the calculation of the vacancy. For the Cu atom, we use the first-principles density-functional nonlocal pseudopotential $v_{\text{ps}}^{(L)}(r)$ derived in Ref. 8. Although used in a

tabulated numerical form, convenient analytical fits to these atomic pseudopotentials are given in Ref. 73. Note that due to the absence of $L=2$ electrons of the core of Cu, our $L=2$ pseudopotential and wave function equal the all-electron quantities, respectively. We use the exchange parameter $\alpha_x=1$ as we used in Sec. VI. Our present calculation is performed consistently on the level of the Kohn-Sham density-functional⁷ formalism. Extension to the level of the self-interaction-corrected density-functional formalism⁷⁴ will be discussed in a future publication. The energies and orbital localization parameters [c.f. Eq. (44)] of the major defect levels introduced by a substitutional Cu impurity are summarized in Table VIII.

A. Quasibands for Si:Cu

Since Cu is likely to produce defect levels associated with its very localized $3d$ orbital, a quasiband representation must be used to produce a converged description of the defect wave functions with respect to bands [Eq. (24)]. We include $M_a=5$ quasiband wave functions for each wave vector \vec{k} (corresponding to the dimensionalities of the $e+t_2$ representations). The local orbital $f_j(\vec{r})$ in Eq. (26) is chosen as the (numerical) $3d$ orbital of the Cu atom, obtained in a self-consistent local-density calculation for the ground state of the Cu atom with $\alpha_x=1$. The calculated quasiband structure $\epsilon_j^{\text{QB}}(\vec{k})$ of Si:Cu is shown in Fig. 12. The bands below 2 Ry are those of pure Si (to within 10^{-3} Ry), as obtained in self-consistent local pseudopotential calculations. Their wave functions $\{\phi_j^{\text{BS}}(\vec{k}, \vec{r})\}$ are grossly insufficient to produce a

TABLE VIII. Energy levels (with respect to the valence-band maximum) and orbital-density localization parameters q_i^α of Eq. (44) for the major defect states associated with a substitutional Cu impurity in silicon. Only states with localization parameters greater than $\sim 0.10e$ are shown. Resonances do not have, in general, a sharply defined energy and localization parameter. For the broad t_2 resonance we quote, therefore, the total localization parameter, summed over the resonance. Results are given for a 250 atom supercell, using an exchange parameter $\alpha_x=1$ and $R_c=4.44$ a.u. (nearest-neighbor sphere radius) in Eq. (44). The maximum values that q_i^α can attain equals the occupation number, i.e., 3.0, 2.0, 4.0, and 6.0 for the t_2 gap state and a_1 , e , and t_2 resonances, respectively.

Symmetry and type	Energy (eV)	Orbital localization parameters (e)					$q_i^\alpha(e)$
		$l=0$	$l=1$	$l=2$	$l=3$	$l=4$	
t_2 gap state	0.580	0.0	0.88	0.240	0.016	0.0075	1.143
$a_1(1)$ resonance	-1.7	0.66	0.0	0.0	0.002	0.0002	0.662
$a_1(2)$ resonance	-3.8	0.214	0.0	0.0	0.0092	0.0012	0.224
$a_1(3)$ resonance	-8.2	0.110	0.0	0.0	0.02	0.0024	0.132
e resonance	-5.15	0.0	0.0	3.752	0.0	0.0008	3.755
t_2 resonance	4.6 to 5.7						5.26

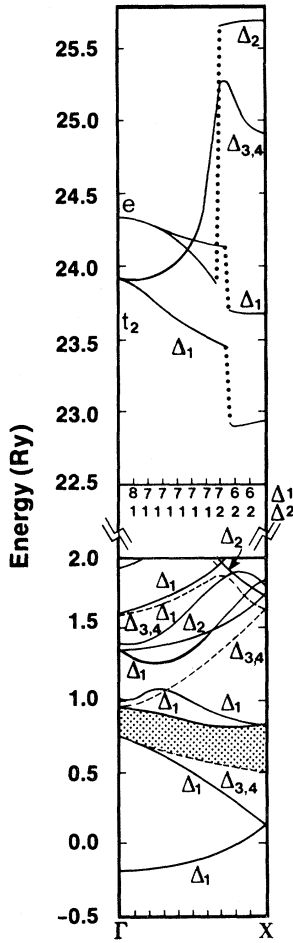


FIG. 12. Quasiband structure for substitutional Cu in silicon. All the band-structure bands below 2 Ry are included. Five local quasibands, constructed from the numerical ground-state Cu 3*d* atomic orbitals are used.

converged description of the localized defect wave functions via Eq. (13). The five upper bands in Fig. 12 (“localized quasibands”) correspond to hybridized solutions between $\{\phi_j^{\text{BS}}(\vec{k}, \vec{r})\}$ and $\{\chi_j(\vec{k}, \vec{r})\}$ in Eq. (25). They are the major contributors to the description of the defect wave functions in terms of bands. This is demonstrated in Fig. 13 which shows the expansion coefficients $A_{ij}^{\alpha\lambda}(\vec{k})$ of Eq. (30) for $|\alpha i\rangle = e$ resonance at four \vec{k} points. Shown are the contributions of the first $M_b = 8$ hostlike bands (open circles), as well as the contributions of the $M_a = 5$ local quasibands (solid circles). It is clear that the inclusion of the first four valence and four conduction bands of silicon is far from sufficient to describe the *e* resonance of Cu in silicon. A simple quantitative measure to that is provided by the proportion of the coeffi-

cients given in Eq. (30): One can partition the normalization of the defect wave function into a contribution from the band-structure (BS) wave functions, and a contribution from the local quasibands (LQB):

$$\begin{aligned} \int |\psi_i^{\alpha,\lambda}(\vec{r})|^2 d\vec{r} &= 1 \\ &= \sum_{j=1}^{M_b} \sum_{\vec{k}_v}^{\text{IBZ}} |A_{ij}^{\alpha\lambda}(\vec{k}_v)|^2 d_v \\ &\quad + \sum_{j=1}^{M_a} \sum_{\vec{k}_v}^{\text{IBZ}} |A_{ij}^{\alpha\lambda}(\vec{k}_v)|^2 d_v \\ &= a_{\text{BS}}^{i\alpha\lambda} + a_{\text{LQB}}^{i\alpha\lambda}. \end{aligned} \quad (48)$$

For the *e* resonance we find $a_{\text{LQB}} = 0.72$ and $a_{\text{BS}} = 0.28$, even when as many as 32 BS bands are used. Hence, on average, each band-structure band contributes less than 1% to the normalization, whereas each local quasiband contributes about 15%. It is estimated that without the local quasibands, several thousand silicon band wave functions will be needed to adequately represent this localized level. This is so because the energy denominator in the Green’s function causes *each* of the high-energy conduction bands to have a small effect. The *t*₂ gap state, on the other hand, can be adequately represented by 25–30 band-structure bands, similar to the situation for the silicon vacancy. This illustrates the extent of the added difficulty in treating a *d*-like versus a simpler *sp*-like defect.

B. The one-center crystal-field representation for Si:Cu

Figure 14(a) shows the spherical component of the self-consistent perturbation potential $\Delta V(\vec{r})$ for the pseudopotential components $L = 0, 1$, and 2 [c.f. (9a)]. Compared with the same quantity for the silicon vacancy (Fig. 9, where an $L = 0$ local pseudopotential is used), it can be seen that Cu constitutes a smaller repulsive perturbation relative to the vacancy. Furthermore, at smaller distances, the $L = 2$ component of $\Delta V_0^L(r)$ becomes strongly attractive (approaching the origin as $-2Z/r$) since no pseudopotential cancellation takes place for 3*d* orbitals. Figure 14(b) shows the *r*-multiplied angular-momentum components of the self-consistent screening perturbation $\Delta V_l^{\text{scr}}(r)$. The large-*r* limit of the $l = 0$ component equals

$$2(Z_{\text{Cu}}^v - Z_{\text{Si}}^v) = 2(11 - 4) = 14.$$

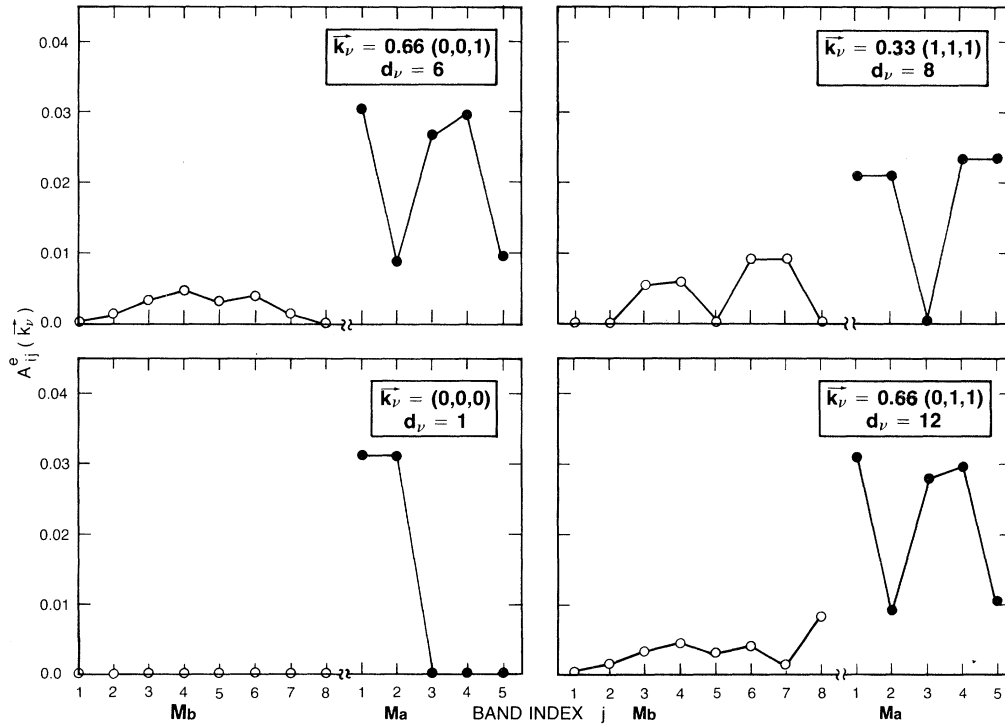


FIG. 13. Spectral decomposition of the wave function of the e resonance in Si:Cu in terms of contributions of the eight lowest hostlike bands (open circles) and the five local quasibands (full symbols). d_ν denotes the degeneracy factor of the star of \vec{k}_ν . \vec{k}_ν is given in units of $2\pi/a$.

The striking result of this figure is that the non-spherical ($l=3,4$) components are attenuated by more than a factor 100 relative to the spherical ($l=0$) component. This substantiates dramatically the basic physical picture underlying the one-center crystal-field approach (cf. Sec. V B 3)—that the expansion of the perturbation in terms of the systems angular momenta l (not the atomic angular states L) converges very rapidly. A similar result is apparent for the charge-density perturbation (Fig. 15): The spherical component [Fig. 15(a)] of $\Delta\rho_l(r)$ is 300–400 larger than the nonspherical components [Fig. 15(b); note change in scale] over the entire central-cell region. As discussed in Sec. VIC, however, this fact does not imply that a spherical model can be used successfully to describe the electronic structure of such systems. The *individual* wave functions may include non-negligible, nonspherical contributions from the lowest few l values. This is illustrated in Figs. 16 and 17 that show the radial l components $G_{il}^a(|\vec{r}|)$ [Eq. (29)] for the a_1 and e resonances as well as the t_2 gap state. Clearly, the a_1 resonances can not be characterized as “ s states,” but rather as s - f hybrids, and the t_2 gap state is likewise not a

“ p state” but a p - d hybrid. The e resonance, on the other hand, is essentially a pure d state; its $l=2$ component is very similar to the Cu $3d$ orbital. It is this fact that makes its description in terms of extended hostlike wave functions (Sec. VII A) virtually impossible. Notice that while both the defect and the host wave functions have non-negligible, nonspherical components, they interfere destructively in a way that produces an essentially spherical screening perturbation [Fig. 14(b)]. Further, as was the case in the vacancy calculation, the rapid attenuation of the contributions of the high- l components to the wave-function amplitude in the central-cell region causes the one-electron energy spectra to be determined almost exclusively by $l=0, 1$, and 3. It is this recognition that made the Bethe–Van Vleck crystal-field idea so fruitful.

The structure of the present crystal-field representation of the defect wave functions [Eqs. (27) and (29)] provides a simple means for constructing variationally efficient basis functions. One can use the radial components $G_{il}^a(|\vec{r}|)$ (Figs. 16 and 17) obtained in an early self-consistency iteration as the basis orbital $f_j(\vec{r})$ [Eq. (26)] for future iterations. Such an iterative update of the basis is not

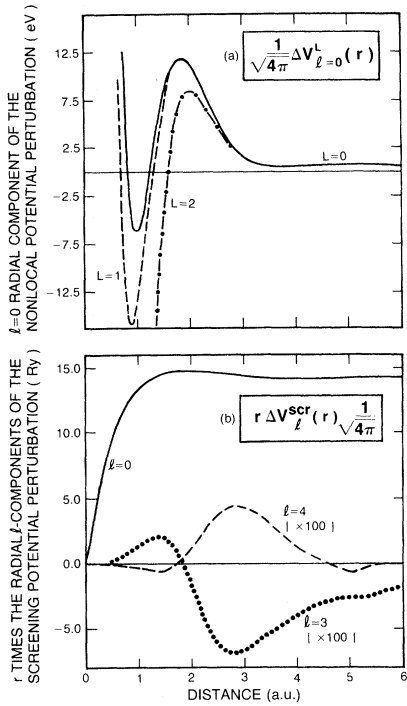


FIG. 14. (a) $l=0$ spherical component of the total self-consistent perturbation potential for Si:Cu. L denotes the nonlocality index of the atomic pseudopotential [c.f. Eq. (7)]. (b) The radial l components of the r multiplied self-consistent screening perturbation for Si:Cu. Notice the scale factor of 100 for the nonspherical components.

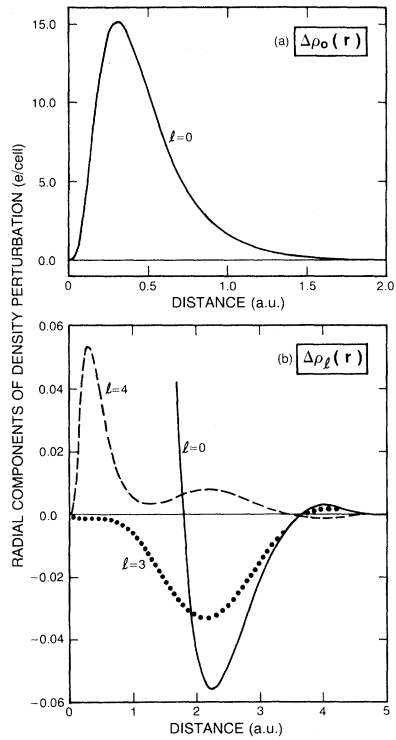


FIG. 15. (a) $l=0$ component of the self-consistent charge-density perturbation for Si:Cu. The spherical average may be obtained by multiplying this function by $1/\sqrt{4\pi}$. (b) An enlarged view of the radial l components of the charge-density perturbation in Si:Cu. Notice that the nonspherical components are much smaller than the spherical part.

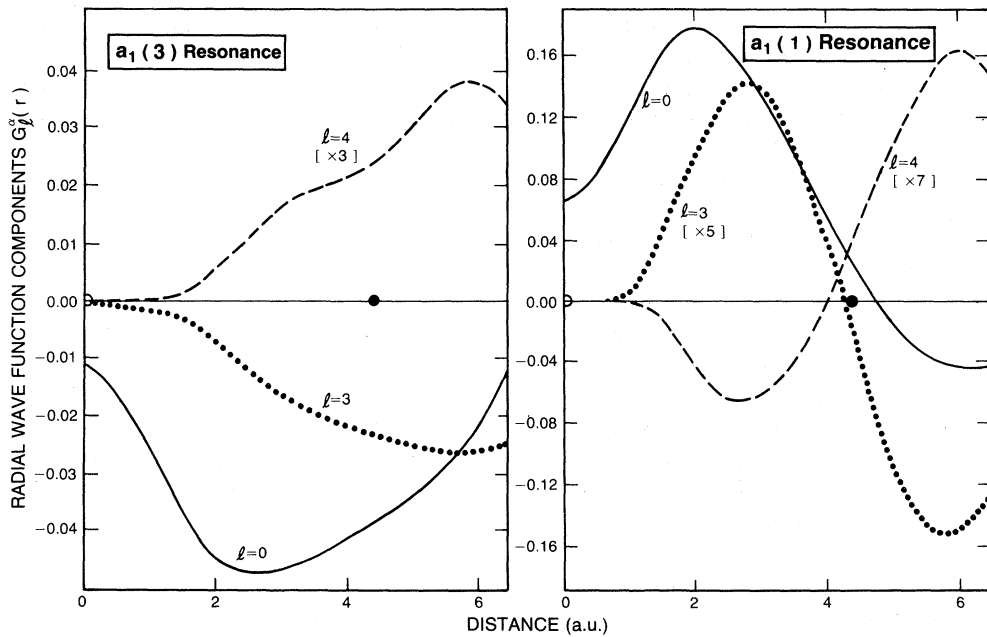


FIG. 16. Radial l components $G_l^a(|\vec{r}|)$ [Eq. (29)] of the a_1 resonances for Si:Cu.

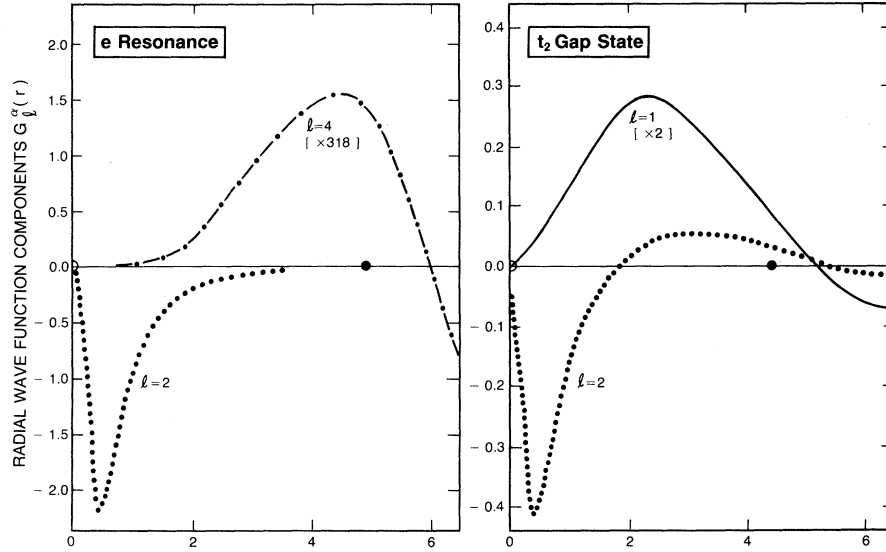


FIG. 17. Radial l components $G_{il}^{\alpha}(|\vec{r}|)$ [Eq. (29)] of the e resonance and t_2 gap state for Si:Cu.

possible in multicentered expansions that do not employ an l representation.

C. Supercells for Si:Cu

The size N_{Ω} of the supercell has been varied to examine its effects on the calculated electronic structure. For the t_2 gap state, we find that supercells of size $N_{\Omega}=54, 250, 686,$ and 1458 atoms produce energy levels of $0.990, 0.577, 0.490,$ and 0.466 eV, respectively, with wave-function localization parameters of $1.23, 1.14, 1.14,$ and 1.14 electrons, respectively. Fitting the first three energy values to $\epsilon(d)=\epsilon(\infty)+ae^{-\alpha d}$ yields $\epsilon(\infty)=0.465$ eV, $a=5.14$ eV, and $\alpha=0.105$ a.u. $^{-1}$, with a precision of a fraction of a percentage. The calculation with $N_{\Omega}=1458$ atoms confirms the prediction for the energy value $\epsilon(\infty)=0.465\pm 0.001$ eV at the limit of an isolated defect. Hence, a precision of 0.1 eV in the determination of the energy of this gap state requires a supercell with an interdefect separation of $d=37.5$ a.u., while a precision of 0.01 eV requires $d\approx 60$ a.u. At present, a calculation corresponding to a precision of 0.1 eV (i.e., few hundreds of atoms per supercell) cannot be performed with the conventional *direct-diagonalization* supercell techniques^{33–35} using state-of-the-art methods of electronic structure. The present method can handle 200 – 2000 atoms per unit cell quite easily.

Since resonance states do not correspond, in gen-

eral, to a sharp energy value, it is more difficult to assess accurately their variations with the supercell size. We find for the sharp e resonance that a calculation with 250 and 686 atoms per supercell produces energy values of -5.150 and -5.148 eV, respectively, indicating excellent convergence. For the broader t_2 resonance, we find three main peaks in the local density of states. The center and width of this resonance are $-5.2\pm 0.5,$ $-5.15\pm 0.55,$ and -5.29 ± 0.3 eV for supercells with $54, 250,$ and 686 atoms, respectively. Our final values for the t_2 gap level, e and t_2 resonances are $0.465,$ $-5.15,$ and -5.3 ± 0.3 eV, respectively. We find that the wave functions of these three levels vary by less than 1% over the range $r\leq 7.5$ a.u. if the supercell size is increased from 250 to 686 atoms.

D. Self-consistency for Si:Cu

In the absence of any previous calculation for a Cu impurity in an extended host crystal, one must guess the initial perturbation potential $\Delta V(\vec{r})$ to be used in the self-consistency iterations. The simplest approximation for $\Delta V(\vec{r})$ in Eq. (9a) can be taken from atomic physics, i.e.,

$$\Delta V^A(r) = (v_{ps}^{(L,Cu)} - v_{ps}^{(Si)}) + (v_{Cu}^{scr} - v_{Si}^{scr}), \quad (49)$$

where $v_{ps}^{(L,B)}$ are atomic pseudopotentials and v_{β}^{scr} is the self-consistent screening of the free atom β . In Fig. 18 we show the difference between the output potential perturbation $\Delta V^{out}(r)$ and the input atom-

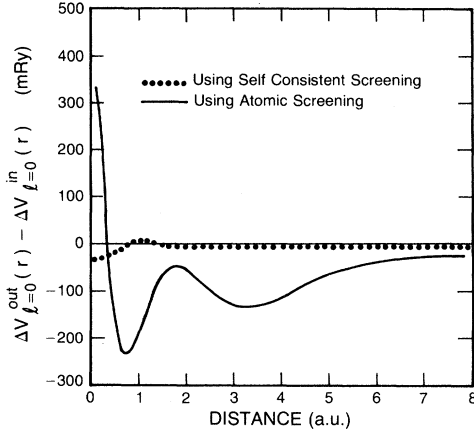


FIG. 18. Difference between the spherical component of the output and input perturbation potentials in Si:Cu using an atomic guess for the input [Eq. (49), full line] and a self-consistent perturbation potential (dotted line).

ic guess $\Delta V^A(r)$ producing it. Also shown in this figure is the corresponding difference between $\Delta V^{\text{out}}(r)$ and $\Delta V^{\text{in}}(r)$ for the final self-consistency iteration. It is seen that the simple atomistic guess produces not only very large errors ($\sim \pm 200$ mRy), but also a nontrivial structure in $\Delta V^{\text{out}}(r) - \Delta V^A(r)$ that could not have been guessed *a priori*. Even much larger errors are obtained with lighter $3d$ impurities. This structure indicates an attenuation of the screening in both the $3d$ and the bond center region and a screening enhancement near the origin. We note that by using in Eq. (49) a bulk screening⁴ instead of $v_{\text{Si}}^{\text{scr}}$ produces even larger errors. The self-consistent result (dotted line in Fig. 18) shows, however, small (~ 5 mRy), uniform errors.

This is achieved in the present calculation in seven iterations, using the Newton-Raphson Jacobi-an update method. We find that it takes two to three initial iterations for the Jacobian to “learn” the iteration path. Using the standard mixing^{3,4,35} of ΔV^{in} with ΔV^{out} , we need at least 30–40 iterations. The difficulty in achieving self-consistency in the screening for a transition-atom impurity can be appreciated from Fig. 15: A small error in the input potential can produce a small adjustment in the charge-density redistribution; even if this change occurs in the region where $\Delta\rho(\vec{r})$ is as small as 1% from its average value ($r \geq 1.5$ a.u.), it results in very large changes in the output potential.

E. A simple chemical picture for Si:Cu

In this section we describe the salient features of the electronic structure of a substitutional Cu im-

purity in silicon, obtained in the present calculation. A comparison with experiment and the discussion of the chemical trends are postponed to a future publication that will discuss the results for all $3d$ transition-atom impurities in silicon.⁷⁵

In discussing the bonding mechanisms in the system, it is useful to consider an electron population analysis of the wave functions. The one-center expansion technique suggests a simple and physically appealing population analysis. The charge contained within the central-cell region in the l component of all occupied representations of type α is

$$Q_l^\alpha = \sum_i^{\text{occ}} N_i \int_0^{R_c} |G_{ii}^\alpha(r)|^2 r^2 dr, \quad (50)$$

where the occupation number N_i includes the representation degeneracy factor. The total charge below the Fermi energy for each l component is

$$Q_l = \sum_\alpha Q_l^\alpha \text{ orbital occupation}, \quad (51)$$

whereas the total charge for a given representation is

$$Q^\alpha = \sum_l Q_l^\alpha \text{ representation occupation}. \quad (52)$$

Finally, the total electronic charge in the central cell is

$$Q^{\text{tot}} = \sum_l Q_l = \sum_\alpha Q^\alpha \text{ total charge}. \quad (53)$$

Each of these charges contains, for the defected systems, both the contributions of the four ligands and the central impurity atom. The corresponding charges for the vacancy contain only contributions from the ligands. It is hence useful to consider the appropriate differences between the various charges for the impurity and the vacancy. These differences correspond to the effective occupation numbers for an impurity atom embedded in a (healed) vacant site in the semiconductor matrix. Table IX gives the various occupation numbers for Si:Cu and the Si vacancy. It shows that the vacancy has a total central-cell charge of six electrons, corresponding to 1.5 electrons for each of the four ligands, while the Si:Cu system has a total charge of 17 electrons in the central cell, corresponding to the eleven Cu valence electrons and the six dangling-bond electrons.

Taking the appropriate differences, one finds that the Cu atom in silicon has an effective orbital occupation of $s^{0.14}p^{0.84}d^{9.92}$ (with additional 0.04 electrons in the fg shells and the rest outside the

TABLE IX. Population analysis (in electrons) of the occupied states for the Si:Cu and the silicon vacancy. The definitions are given in Eqs. (50)–(53). Q_l^α denotes the l th component of the density of all occupied α representations. Q^α is the representation occupation, Q_l is the orbital occupation, and Q^{tot} is the total electronic charge. The sphere radius used in Eqs. (50)–(53) is $R_c = 4.44$ a.u. which corresponds to the nearest-neighbor distance.

l	$Q_l^{a_1}$	Q_l^e	$Q_l^{t_1}$	$Q_l^{t_2}(\text{VB})$	$Q_l^{t_2}(\text{gap})$	Q_l
Si:Cu						
0	1.7086	0.0	0.0	0.0	0.0	1.7086
1	0.0	0.0	0.0	2.098	0.880	2.978
2	0.0	4.293	0.0	6.670	0.239	11.202
3	0.1406	0.0	0.4536	0.1397	0.0162	0.750
4	0.0163	0.086	0.0698	0.1811	0.0075	0.361
Q^α	1.8655	4.379	0.5234	9.089	1.143	$Q^{\text{tot}} = 17.00$
Silicon vacancy						
0	1.516	0.0	0.0	0.0	0.0	1.516
1	0.0	0.0	0.0	1.663	0.473	2.136
2	0.0	0.3074	0.0	0.7734	0.205	1.286
3	0.167	0.0	0.309	0.1310	0.0156	0.623
4	0.024	0.1155	0.0825	0.2206	0.0069	0.449
Q^α	1.707	0.423	0.391	2.788	0.7005	$Q^{\text{tot}} = 6.01$

central-cell region), compared with the ground-state atomic configuration of $s^1 p^0 d^{10}$ and the effective configuration $(sp)^{1.3} d^{9.7}$ in the metallic state.⁷⁶ One observes that the effect of the host is to keep Cu almost neutral within the central cell and merely shift some of the s electrons into the p shell without breaking the d shell. This is in marked contradiction with the Ludwig-Woodbury^{30(a)} model which suggested that a substitutional transition atom will have two of its d electrons promoted into the sp shells to form sp^3 hybrids (i.e., $s^1 p^3 d^7$ for Cu). The basic picture that emerges from our population analysis is that, although the discrete atomic levels of Cu, when brought into the crystal, broaden and hybridize into various resonances with a nontrivial distribution of energy levels and orbital charge densities, its “effective” electronic configuration remains remarkably very much like an excited free atom.

One can further deduce from the population analysis the distribution of the charge of the effective Cu atom into various representations. This gives $a_1^{0.16} e^{3.96} t_1^{0.13} t_2^{6.75}$ showing that the atomic configuration $a_1^1(e, t_2)^{10}$ has lost 0.84 of its a_1 electrons into the other representations and how the total non- s charge is redistributed into these crystal-field states.

Although a population analysis can indicate the overall redistribution of charge into quantum

states, it provides no information on the actual bonding (i.e., spatial extent of the populated quantum states). This can be assessed from inspecting the wave functions. Figure 19 shows the wave function of the t_2 gap state, and Fig. 20 shows the orbital density of the e and t_2 resonances. Comparing the t_2 gap-state wave function of Si:Cu with the corresponding wave function for the vacancy (Fig. 6), one notices that the former has a node in the central cell (resulting from its orthogonality to the lower t_2 resonance). In fact, this node is the “fingerprint” of the transition-atom impurity and is introduced by the $l=2$ component of the wave function (cf. Fig. 17). This component has the op-

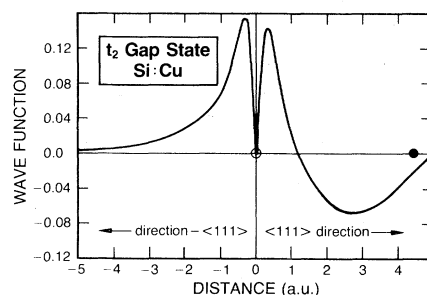


FIG. 19. Wave function of the t_2 gap state for Si:Cu, plotted in the $\pm\langle 111 \rangle$ directions.

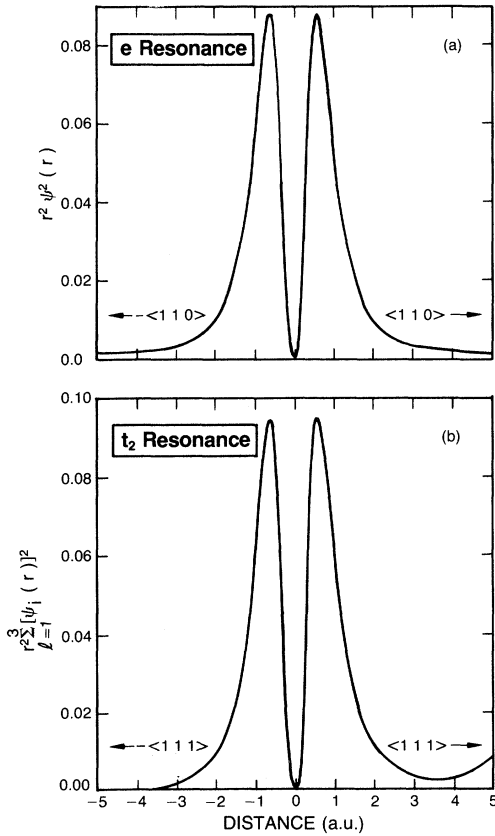


FIG. 20. Radial densities of the Si:Cu e resonance in the $\pm\langle 110 \rangle$ directions (a), and the sum of radial densities for the three components of the t_2 resonance, shown in the $\pm\langle 111 \rangle$ directions (b).

posite sign of the $l=1$ component in Si:Cu, but the same sign in the case of silicon vacancy (cf. Fig. 4). The gap state wave function in Si:Cu is hence distinctly different from the dangling-bond t_2 wave function of the silicon vacancy.

The t_2 and e resonances (Fig. 20) retain very much the overall characteristics of the atomic $3d$ orbital. Since the lobes of the e orbital point away from the nearest-neighbor ligands, this state is almost a pure unhybridized d state (cf. Table VIII), while the t_2 resonance hybridizes somewhat with the ligands and produces a broad resonance. Nevertheless, the *sum* of the orbital densities over this broad resonance [Fig. 20(b)] produces a localized atomiclike state.

The basic picture that emerges from this description is that the bonding in the system is produced by the sp states, while the d states are merely renormalized into a set of weakly hybridized resonances that combine to produce a simple atomic $3d$ density. Although the d orbitals are

essentially passive, a simple, atomic view of this impurity is likely to be wrong since extensive charge redistribution and hybridization occurs in the s - p manifold.

VIII. COMPUTER TIME

We quote here the computer time needed for a fully self-consistent calculation of a substitutional transition-atom impurity in a semiconductor, using the method described in this paper. A CDC 7600 computer system is used with a NOS 1.4 level 531.00 operating system. We note that, at present, no effort was done to optimize the computer program with respect to execution efficiency; such optimization is expected to reduce the computer time substantially. The specific example is a Cu substitutional impurity characterized by a nonlocal pseudopotential with angular-momentum classes of $L=0,1$ and $L=2$, in a 54-atom silicon host supercell. The number of host bands is $M_b=30$ per \mathbf{k} point plus $M_a=5$ local quasibands per \mathbf{k} point. The local basis set is composed of eleven radial s orbitals, ten radial p orbitals, ten radial d orbitals, eight radial f orbitals, and eight radial g orbitals, all in the form of Coulomb functions, Schmidt-orthogonalized in a sphere of radius $R_c=7.5$ a.u. In addition, a single numerical radial Cu orbital is used in the local basis $\{g_a(\vec{r})\}$. All local functions are mutually orthogonal. If we count both radial and angular components, the basis has 11, 30, 50, 56, and 72 orbitals of angular momentum $l=0, 1, 2, 3,$ and 4 , respectively, or a total of $N=219$, with $L_{\max}^{(1)}=4$. The highest angular momentum kept in the density and potential expansions is $L_{\max}^{(2)}=8$. The one-electron energy range searched in the Green's-function calculation is 15 eV; the energy resolution is $\Delta\epsilon=0.00015$ eV (i.e., levels separated by $\Delta\epsilon$ can be solved). In searching for solutions we attempt seven successive iterations for each interval. The wave-function representations that are calculated are $a_1, e, t_1,$ and t_2 . The partial wave components of the potential and density $\Delta V_l(|r_i|)$ and $\Delta\rho_l(|r_i|)$ are tabulated on 145 logarithmic grid points. The maximum l component for fitting the exchange potential is $L_{\max}^{(3)}=12$. Nineteen angular directions are used in this fit.

The computer time needed is distributed among the various operations as follows: (1) The iteration-independent operations (i.e., a one-time-expense) are (i) calculation of the Cu pseudopotential from scratch, 15 sec, (ii) calculation of the host band structure, starting from a self-consistent host

potential, 200 sec, and (iii) calculation of all overlap integrals and quasibands entering the Green's function, 300 sec. (2) Iteration-dependent calculations, given here per iteration, are (i) calculation of the potential perturbation from defect wave function, 45 sec and (ii) the solution of the Green's-function problem, 110 sec. If we start from an atomic potential as an initial guess, the Jacobian update method requires seven iterations to converge to a few mRy accuracy. The calculations cost is hence 8.6 min for the one-time expense, plus 18 min for the full self-consistency. Using somewhat less stringent parameters lowers the cost substantially (e.g., if only 25 bands, rather than 35 are used, the iteration time drops from 18 to 13 min). A similar calculation for a silicon vacancy (Sec. VI) does not require quasibands and is converged to full self-consistency (starting from atoms) in only five iterations. The computer time required for this calculation is a small fraction of that required for a Cu impurity.

IX. LIMITATIONS OF THE METHOD

In this section we provide a critique of our proposed QBCF method, based on our experience with it to date. It has three main weaknesses. First, the method is well suited for potential and density perturbations $\Delta V(\vec{r})$ and $\Delta\rho(\vec{r})$ that are localized within a radius R_c of an order of 4–6 a.u. Any strong component beyond this radius requires that the angular-momentum summations in Eqs. (33), (35), and (36) be extended beyond $l=8$. This, however, is not a severe limitation for treating deep defects: Even the vacancy gap state that is considerably extended (since it is a dangling-bond orbital rather than an impurity-anchored orbital) has an $R_c \cong 4.5$ a.u. Substantial lattice relaxations may, however, extend the range of $\Delta V(\vec{r})$ making the present method somewhat more costly to apply. Second, the number of partial wave components in the wave functions and potential increases rapidly as the site symmetry of the defect is lowered. For example, for a tetrahedral site (i.e., substitutional or tetrahedral interstitial in zinc-blende hosts) only the $l=0,3,4, \dots$, components appear in the a_1 representation; however, a trigonally distorted lattice also needs the $l=2$ component. Third, due to the use of a discrete set of energy levels, it is difficult to calculate the exact width of defect resonances with a precision better than ~ 0.2 eV.

Note added in proof. Recently, we have been able to use the QBCF method for calculation of all

substitutional and interstitial $3d$ impurities in silicon, as well as to predict the magnitude of the a_1 lattice distortion in the silicon vacancy, with high precision. The results indicate that our concerns about the limitations (increased computer costs) of the method were pessimistic.

ACKNOWLEDGMENTS

We are grateful to P. Bendt for collaboration on the Newton-Raphson Jacobian update method (Ref. 5) and for helpful discussions throughout the course of this work. We thank V. A. Singh for his assistance in the vacancy calculations at the early stage of this application and for very stimulating discussions.

APPENDIX A: THE CRYSTAL-FIELD PSEUDOPOTENTIAL

In this appendix we show how the standard ligand field or cluster approach to the defect problem can be transformed into a one-center problem by adding to the defect potential a crystal-field pseudopotential derived here. Our derivation follows the idea suggested by Phillips.⁴⁰

Consider a defect-containing system described by an unperturbed host potential $V_H(\vec{r})$ and a defect-associated perturbation $\Delta V(\vec{r})$, which has a localized (e.g., gap state) defect wave function $\psi_i^\alpha(\vec{r})$ with energy E_i and a manifold of occupied wave functions $\{\phi_j^\alpha(\vec{r})\}$ with energies ϵ_j below E_i . The potentials $V_H(\vec{r})$ and $V_H(\vec{r}) + \Delta V(\vec{r})$ may be described either as all electron or as screened pseudopotentials. In the former case $\phi_j^\alpha(\vec{r})$ will include corelike states, whereas in the latter case $\phi_j^\alpha(\vec{r})$ describes the valence pseudowave functions. In either cases, the wave function $\psi_i^\alpha(\vec{r})$ is orthogonal to all $\phi_j^\alpha(\vec{r})$ and, therefore, might include nodal structure outside the range of $\Delta V(\vec{r})$. In the classical crystal-field approach,^{14,15} one attempts to solve

$$[\hat{T} + V_H(\vec{r}) + \Delta V(\vec{r})]\psi_i^\alpha(\vec{r}) = E_i\psi_i^\alpha(\vec{r}) \quad (\text{A1})$$

by expanding $\psi_i^\alpha(\vec{r})$ in an impurity-anchored one-center expansion

$$\psi_i^\alpha(\vec{r}) = \sum_l R_{nl}^\alpha(|\vec{r}|)K_l^{\alpha,\lambda}(\vec{r}) \quad (\text{A2})$$

in the crystal-field approximation, where $R_{nl}^\alpha(|\vec{r}|)$ is the radial impurity atomic orbital in the α th point-group representation. Often the same radial

orbital R_{nl}^α is used for all l components. Clearly, however, the simple choice of Eq. (A2) will fail for distances outside the range of $\Delta V(\vec{r})$ where $V_H(\vec{r})$ is strong (e.g., on the ligands), since the limited basis set in Eq. (A2) would attempt to reproduce the ligandlike nodal structure of $\phi_j^\alpha(\vec{r})$ rather than $\psi_i^\alpha(\vec{r})$. This difficulty can be circumvented in a way that still permits a convenient one-center expansion of the localized wave function. This will lead to a generalized crystal-field pseudopotential (CFP) equation,

$$[\hat{T} + V_H(\vec{r}) + \Delta V(\vec{r}) + V_{\text{CFP}}^\alpha(\vec{r}, E_i)] \chi_i^\alpha(\vec{r}) = E_i \chi_i^\alpha(\vec{r}), \quad (\text{A3})$$

where a new, representation-dependent (“nonlocal”) term $V_{\text{CFP}}^\alpha(\vec{r}, E_i)$ is added to the original Hamiltonian. We now derive a form for the CFP based on a transformation of the defect wave functions of Eq. (A1) into a crystal-field pseudowave function $\chi_i^\alpha(\vec{r})$. The latter can be written as a linear combination of the original occupied wave functions,

$$\chi_i^\alpha(\vec{r}) = B_{ii}^\alpha \psi_i^\alpha(\vec{r}) + \sum_{j \neq i} B_{ij}^\alpha \phi_j^\alpha(\vec{r}), \quad (\text{A4})$$

where the coefficients $\{B_{ij}^\alpha\}$ are at this stage arbitrary. From Eqs. (A1), (A3), and (A4), one obtains the CFP that satisfies Eq. (A3) by construction, as:

$$V_{\text{CFP}}^\alpha(r, E_i) = \frac{\sum_{j \neq i} B_{ij}^\alpha (E_i - \epsilon_j) \phi_j^\alpha(\vec{r})}{B_{ii}^\alpha \psi_i^\alpha(\vec{r}) + \sum_{j \neq i} B_{ij}^\alpha \phi_j^\alpha(\vec{r})}. \quad (\text{A5})$$

Notice that *any* choice of $\{B_{ij}^\alpha\}$ in Eq. (A4) will provide a solution to Eq. (A3) with the correct energy E_i . However, the resulting wave function $\chi_i^\alpha(\vec{r})$ may not share the properties of the correct solution $\psi_i^\alpha(\vec{r})$. Note further that with the CFP of Eq. (A5) one can solve Eq. (A3) even with a limited basis set without fearing that the solution E_i will converge to a low-lying ligandlike state ϵ_j , which is the case in the standard crystal-field and ligand-field approach. This permits then an expansion of the crystal-field pseudowave functions $\chi_i^\alpha(\vec{r})$ in a defect-anchored, one-center expansion,

$$\chi_i^{\alpha\lambda}(\vec{r}) = \sum_l G_{il}^{\alpha\lambda}(|\vec{r}|) K_l^\alpha(\hat{r}), \quad (\text{A6})$$

with coefficients $G_{il}^{\alpha\lambda}(|\vec{r}|)$ of Eq. (34). This, however, provides solutions to our generalized crystal-field problem in Eq. (A3), but not to the original equation (A1) and (A2), as assumed simplistically in the classical crystal-field method.

The coefficients $\{B_{ij}^\alpha\}$ in Eqs. (A4) and (A5)

may be chosen as

$$B_{ij}^\alpha = \langle \chi_i^\alpha(r) | \phi_j^\alpha(\vec{r}) \rangle$$

to satisfy orthogonality. There are, however, other useful choices, e.g., select $\{B_{ij}^\alpha\}$ so that the second term of Eq. (A4) is minimized outside the range of $\Delta V(\vec{r})$ [that is, minimize the projections of $\phi_j^\alpha(\vec{r})$ on $\chi_i^\alpha(\vec{r})$ at this range] or determine $\{B_{ij}^\alpha\}$ to minimize the kinetic energy of $\chi_i^\alpha(\vec{r})$. These two choices may allow one to use in the expansion of Eq. (A6) a localized set or a smooth set, respectively. Any such choice still permits one to recover the correct wave function $\psi_i^\alpha(\vec{r})$ from the occupied states through

$$\psi_i^\alpha(\vec{r}) = \chi_i^\alpha(\vec{r}) - \sum_{j \neq i} \langle \chi_i^\alpha | \phi_j^\alpha \rangle \phi_j^\alpha(\vec{r}). \quad (\text{A7})$$

It is instructive to compare the CFP of Eq. (A5) to standard atomic pseudopotentials $v_{\text{ps}}^{(L)}(\vec{r})$. In atomic pseudopotential theory, one orthogonalizes the valence states to the core states that are localized at small r , resulting in a $v_{\text{ps}}^{(L)}(\vec{r})$ that is often repulsive at small r and weak at large r . In the CFP we orthogonalize to the ligand states $\phi_j^\alpha(\vec{r})$ that are extended to large r . For a simple two-level model, it can be seen from Eq. (A5) that at the distance where the localized defect wave function $\psi_i^\alpha(\vec{r})$ is vanishingly small (i.e., large r) the CFP is repulsive (as $E_i > \epsilon_j$). Notice that at such distances one can model $V_{\text{CFP}}^\alpha(\vec{r}, E_i)$ without recourse to any information on $\psi_i^\alpha(\vec{r})$.

Clearly, much like the situation in atomic pseudopotential theory the calculation of $V_{\text{CFP}}^\alpha(\vec{r}, E)$ requires the knowledge of the solutions E_i . Hence Eq. (A5) is of little *direct* help. However, as has been the case in applications of pseudopotentials to the determination of crystal band structure, the form of Eq. (A5) lends itself to modeling in terms of effective potentials that share with Eq. (A5) its overall properties. Such models [the simplest is a replacement of $V_{\text{CFP}}^\alpha(\vec{r}, E_i)$ by a model square well that is repulsive outside a radius R_c , in the spirit of the Heine-Abarenkov atomic model potential] may prove to be very useful.

APPENDIX B: CONSTRUCTION OF THE MIXED BASIS MATRIX ELEMENTS FOR THE CALCULATION OF THE QUASIBAND STRUCTURES

The calculation of the quasiband wave functions,

$$\phi_j^{\text{QB}}(\vec{k}, \vec{r}) = \sum_{j'}^{M_b} b_{jj'}(\vec{k}) \phi_j^{\text{BS}}(\vec{k}, \vec{r}) + \sum_{j''}^{M_a} a_{jj''}(\vec{k}) \chi_{j''}(\vec{k}, \vec{r}), \quad (\text{B1})$$

and energies $\{\epsilon_j^{\text{OB}}(\vec{k})\}$ requires the diagonalization of the host-crystal Hamiltonian \hat{H}_0 in the nonorthogonal basis of band-structure wave functions $\{\phi_j^{\text{BS}}(\vec{k}, \vec{r})\}$ and local Bloch functions $\{\chi_j(\vec{k}, \vec{r})\}$ [Eq. (25)]. This then requires the evaluation of the matrix elements

$$\begin{aligned} &\langle \phi_j^{\text{BS}}(\vec{k}, \vec{r}) | \hat{O} | \phi_j^{\text{BS}}(\vec{k}, \vec{r}) \rangle, \\ &\langle \chi_j(\vec{k}, \vec{r}) | \hat{O} | \chi_j(\vec{k}, \vec{r}) \rangle, \end{aligned}$$

and

$$\langle \phi_j^{\text{BS}}(\vec{k}, \vec{r}) | \hat{O} | \chi_j(\vec{k}, \vec{r}) \rangle,$$

for $\hat{O} = \hat{H}_0$ or unity. This appendix describes the calculation of these elements. We will assume that the band-structure wave functions $\{\phi_j^{\text{BS}}(\vec{k}, \vec{r})\}$ are available in a Fourier series:

$$\phi_j^{\text{BS}}(\vec{k}, \vec{r}) = N^{-1/2} \sum_{\vec{G}} B_j(\vec{k} + \vec{G}) e^{i(\vec{k} + \vec{G}) \cdot (\vec{r} - \vec{\tau})}, \quad (\text{B2})$$

which can be obtained by many standard band-structure programs. Here $\vec{\tau}$ is the displacement vector between the band-structure origin (e.g., bond center) and the impurity site. The final working expressions derived in this appendix are Eqs. (B8) and (B12)–(B14).

The calculation of the matrix elements between $\phi_j^{\text{BS}}(\vec{k}, \vec{r})$ is trivial since these wave functions already diagonalize \hat{H}_0 :

$$\langle \phi_j^{\text{BS}}(\vec{k}, \vec{r}) | \hat{H}_0 | \phi_j^{\text{BS}}(\vec{k}, \vec{r}) \rangle = \epsilon_j(\vec{k}) \delta_{jj'}$$

and

$$\langle \phi_j^{\text{BS}}(\vec{k}, \vec{r}) | \phi_j^{\text{BS}}(\vec{k}, \vec{r}) \rangle = \delta_{jj'}.$$

We then proceed to calculate the other matrix elements. As shown in Eq. (26), the local Bloch function $\chi_j(\vec{k}, \vec{r})$ is given as an LCAO Bloch sum of some local orbitals $f_j(\vec{r})$:

$$\begin{aligned} \chi_j(\vec{k}, \vec{r}) &= N^{-1/2} \sum_{\vec{R}_p} e^{i\vec{k} \cdot \vec{R}_p} f_j(\vec{r} - \vec{R}_p) \\ &\equiv e^{i\vec{k} \cdot \vec{r}} u_j(\vec{k}, \vec{r}), \end{aligned} \quad (\text{B3})$$

where $u_j(\vec{k}, \vec{r})$ is periodic with respect to translations by \vec{R}_p . It can therefore be Fourier-analyzed as

$$\begin{aligned} u_j(\vec{k}, \vec{r}) &\equiv N^{-1/2} \sum_{\vec{R}_p} e^{-i\vec{k} \cdot (\vec{r} - \vec{R}_p)} f_j(\vec{r} - \vec{R}_p) \\ &= \sum_{\vec{G}} F_j(\vec{k}, \vec{G}) e^{i\vec{G} \cdot \vec{r}}. \end{aligned} \quad (\text{B4})$$

Defining the Fourier transform of the local orbital $f_j(\vec{r})$ as

$$f_j(\vec{q}) = (2\pi)^{-3/2} \int e^{-i\vec{q} \cdot \vec{r}} f_j(\vec{r}) d\vec{r}, \quad (\text{B5})$$

one obtains the relation between $f_j(\vec{q})$ and $F_j(\vec{q}, \vec{G})$ as

$$F_j(\vec{k}, \vec{G}) = \frac{(2\pi)^{3/2}}{\Omega_a N^{1/2}} f_j(\vec{k} + \vec{G}), \quad (\text{B6})$$

where Ω_a is the unit-cell volume. The LCAO Bloch function $\chi_j(\vec{k}, \vec{r})$ can then be written as a Fourier series:

$$\begin{aligned} \chi_j(\vec{k}, \vec{r}) &= \left[\frac{(2\pi)^3}{N\Omega_a^2} \right]^{1/2} \\ &\times \sum_{\vec{G}} f_j(\vec{k} + \vec{G}) e^{i(\vec{k} + \vec{G}) \cdot (\vec{r} + \vec{\rho})}, \end{aligned} \quad (\text{B7})$$

where we have shifted the origin from the impurity site ($\vec{r} = 0$) by a constant displacement vector $\vec{\tau}$ to the bond center position ($\vec{\rho} = 0$), hence $\vec{r} = \vec{\rho} + \vec{\tau}$. Using Eq. (B2), we see that the Hamiltonian matrix elements in the mixed basis become

$$\begin{aligned} &\langle \chi_j(\vec{k}, \vec{r}) | \hat{H}_0 | \phi_j^{\text{BS}}(\vec{k}, \vec{r}) \rangle \\ &= \frac{(2\pi)^{3/2}}{\Omega_a} \sum_{\vec{G}, \vec{G}'} f_j^*(\vec{k} + \vec{G}') B_j(\vec{k} + \vec{G}) e^{-i(\vec{k} + \vec{G}') \cdot \vec{\tau}} \\ &\quad \times H_0(\vec{k} + \vec{G}'; \vec{k} + \vec{G}), \end{aligned} \quad (\text{B8a})$$

where $H_0(\vec{k} + \vec{G}'; \vec{k} + \vec{G})$ is the Hamiltonian matrix in a pure plane wave basis,

$$\begin{aligned} \Omega_a^{-1} H_0(\vec{k} + \vec{G}; \vec{k} + \vec{G}') &= (\vec{k} + \vec{G})^2 \delta_{\vec{G}, \vec{G}'} \\ &\quad + V_H(\vec{G} - \vec{G}'). \end{aligned} \quad (\text{B8b})$$

Here $V_H(\vec{G} - \vec{G}')$ is the Fourier transform of the host potential $V_H(\vec{\rho})$ of Eq. (5). Both $\{V_H(\vec{q})\}$ and $\{B_j(\vec{k} + \vec{G})\}$ are readily available as output from many band-structure programs. The calculation of the matrix element of Eq. (B8) hence requires performing a one-dimensional radial integral to obtain $f_j(\vec{q})$ of Eq. (B5) and the double reciprocal space sum of Eq. (B8a). Note, however, that the number of terms in the double reciprocal space sum is related to the size of the host Hamiltonian matrix and not to the number of terms required to converge the Fourier transform of $f_j(\vec{r})$. The latter may be very extended in reciprocal space.

For constructing the matrix elements $\langle \chi_j(\vec{k}, \vec{r}) | \hat{H}_0 | \chi_j(\vec{k}, \vec{r}) \rangle$, it is useful to assume that the local orbitals $f_j(\vec{r})$ are confined within a

unit cell. This poses no serious restriction since the object of the quasiband approach is to have the set $\{\chi_j(\vec{k}, \vec{r})\}$ reproduce the localized parts of the defect wave functions; the extended pieces of it are then spanned by $\{\phi_j^{\text{BS}}(\vec{k}, \vec{r})\}$. Using this fact, one obtains

$$\langle \chi_j(\vec{k}, \vec{r}) | \hat{H}_0 | \chi_{j'}(\vec{k}, \vec{r}) \rangle = \langle f_j(\vec{r}) | \hat{H}_0 | f_{j'}(\vec{r}) \rangle. \quad (\text{B9})$$

We now represent the local orbitals $f_j(\vec{r})$ in a crystal-field manner as in Eq. (27), breaking the single index j into $(\mu, \alpha, \lambda, l)$, where μ denotes the radial "principal" quantum number, α is the irreducible representation of the defect's point-symmetry group, λ is the partner index in this representation, and l is the angular momentum:

$$f_{\mu l}^{\alpha \lambda}(\vec{r}) = F_{\mu l}(|\vec{r}|) K_l^{\alpha, \lambda}(\hat{r}). \quad (\text{B10})$$

Similarly, the host potential $V_H(\vec{r})$ is decomposed in partial waves as

$$V_H(\vec{r}) = \sum_l v_l(|\vec{r}|) K_l^{\alpha}(\hat{r}). \quad (\text{B11})$$

In analogy to atomic physics, the matrix element of Eq. (B9) breaks into an angular part that results in sums over Gaunt coefficients $D_l^{\alpha}[l', l'']$ (defined in Appendix E and given in Table II) and simple one-dimensional radial integrals:

$$\begin{aligned} \langle \chi_j(\vec{k}, \vec{r}) | \hat{H}_0 | \chi_{j'}(\vec{k}, \vec{r}) \rangle &= -\delta_{l'l} \delta_{\alpha\alpha} \delta_{\lambda\lambda} \langle F_{\mu l} | \hat{T} | F_{\mu l} \rangle \\ &+ \frac{1}{d_{\alpha}} \delta_{\lambda\lambda} \delta_{\alpha\alpha} \sum_{l''} D_l^{\alpha}[l', l''] \langle F_{\mu l'} | v_{l''} | F_{\mu l} \rangle. \end{aligned} \quad (\text{B12})$$

Here

$$\hat{T} = \frac{1}{r^2} \frac{\partial}{\partial r} \left[r^2 \frac{\partial}{\partial r} \right] - \frac{l(l+1)}{r^2}$$

is the radial part of the Laplacian and d_{α} is the dimensionality of the α th representation.

The various overlap matrix elements can be obtained by replacing in Eqs. (B8) and (B12) the Hamiltonian operator by unity. Hence we obtain:

$$\begin{aligned} \langle \chi_j(\vec{k}, \vec{r}) | \phi_{j'}^{\text{BS}}(\vec{k}, \vec{r}) \rangle &= (2\pi)^{3/2} \sum_{\vec{G}} f_j^*(\vec{k} + \vec{G}) B_{j'}(\vec{k} + \vec{G}) e^{-i(\vec{k} + \vec{G}) \cdot \vec{r}} \end{aligned} \quad (\text{B13})$$

and

$$\langle \chi_j(\vec{k}, \vec{r}) | \chi_{j'}(\vec{k}, \vec{r}) \rangle = \delta_{l'l} \delta_{\alpha\alpha} \delta_{\lambda\lambda} \langle F_{\mu l} | F_{\mu l} \rangle. \quad (\text{B14})$$

Again, in Eq. (B13), the number of terms in the \vec{G} summation is limited by the size of the Hamiltonian matrix, whose dimensionality equals the number of \vec{G} components in $B_j(\vec{k} + \vec{G})$. Given the matrix elements in Eqs. (B8) and (B12)–(B14), one generalizes the eigenvalue problem

$$\underline{H} \begin{bmatrix} b \\ a \end{bmatrix} = \epsilon \underline{S} \begin{bmatrix} b \\ a \end{bmatrix}$$

to obtain as solutions the quasiband wave functions [Eq. (B1)] and energies.

APPENDIX C: CALCULATION OF GREEN'S-FUNCTION MATRIX ELEMENTS

The calculation of the Green's-function matrix elements

$$G^0(\epsilon)_{ab} = \sum_j \sum_{\vec{k}} \frac{\langle g_a | \Theta | j\vec{k} \rangle \langle j\vec{k} | \Theta | g_b \rangle}{\epsilon - \epsilon_j^{\text{QB}}(\vec{k})}, \quad (\text{C1})$$

in the local basis representation $\{g_a(\vec{r})\}$ requires the calculation of the overlap integrals $\langle g_a | \Theta | j\vec{k} \rangle$ where $\phi_j^{\text{QB}}(\vec{k}, \vec{r}) = \langle \vec{r} | j\vec{k} \rangle$ are the quasiband wave functions. According to Eq. (25), we have sums over two types of elementary overlap integrals:

$$\langle g_a | \Theta | j\vec{k} \rangle = \sum_{j'=1}^{M_b} b_{jj'}(\vec{k}) \langle g_a(\vec{r}) | \Theta(\vec{r}) | \phi_{j'}^{\text{BS}}(\vec{k}, \vec{r}) \rangle + \sum_{j''=1}^{M_a} a_{jj''}(\vec{k}) \langle g_a(\vec{r}) | \Theta(\vec{r}) | \chi_{j''}(\vec{k}, \vec{r}) \rangle. \quad (\text{C2})$$

We denote these two types of elementary overlap integrals, as

$$S_1 = \langle g_a(\vec{r}) | \Theta(\vec{r}) | \phi_{j'}^{\text{BS}}(\vec{k}, \vec{r}) \rangle$$

and

$$S_{II} = \langle g_a(\vec{r}) | \Theta(\vec{r}) | \chi_j(\vec{k}, \vec{r}) \rangle .$$

From Eqs. (27) and (B2) we obtain for the first matrix element in Eq. (C2),

$$S_I = \frac{4\pi}{\sqrt{N}} \sum_{\vec{G}} B_j(\vec{k} + \vec{G}) e^{-i(\vec{k} + \vec{G}) \cdot \vec{\tau}} K_l^{\alpha\lambda}(\vec{k} + \vec{G}) i^l \int_0^\infty F_{\mu l}(r) j_l(|\vec{k} + \vec{G}|r) \Theta(r) r^2 dr , \quad (C3)$$

where $\vec{\tau}$ is the displacement vector between the origin of the band-structure wave functions and the impurity site. Here we have expanded the plane waves in terms of Bessel functions $j_l(qr)$ and Kubic harmonics.

For the matrix element S_{II} in Eq. (C2), one could use the same approach, i.e., express $\chi_j(\vec{k}, \vec{r})$ in a Fourier series and sum Fourier elements as in Eq. (C3). Since, however, $\chi_j(\vec{k}, \vec{r})$ is constructed from highly localized functions (e.g., $3d$ atomic orbitals), the Fourier transform of $\chi_j(\vec{k}, \vec{r})$ is extremely slowly converging (often requiring more than 70 000 \vec{G} components for $3d$ transition-atom orbitals²⁶). A computationally much faster procedure is to calculate the two-center integrals directly from the definition of $\chi_j(\vec{k}, \vec{r})$ in Eq. (26):

$$S_{II} = \frac{1}{\sqrt{N}} \sum_{\vec{R}_p} e^{i\vec{k} \cdot \vec{R}_p} \times \int g_a^*(\vec{r}) \Theta(\vec{r}) f_j(\vec{r} - \vec{R}_p) d\vec{r} . \quad (C4)$$

Note that from this equation the advantage of introducing the cut-off function $\Theta(\vec{r})$ becomes apparent: It reduces the number of two-center integrals that need to be calculated.

The calculation of the integrals in Eq. (C4) involves the standard α -Löwdin expansion.⁵⁶ However, this technique has not been previously generalized to include the treatment of *arbitrary* orbi-

tals $f_j(\vec{r} - \vec{R}_p)$ and $g_a(\vec{r})$, where the polar axes of the two-coordinate systems do not point towards each other. In this appendix, we give a generalization of this technique. The final results of this technique are given in Eqs. (C7)–(C11).

The local orbital $g_a(\vec{r}) = g_{\mu l}^{\beta\lambda}(\vec{r}) = R_{\mu l}(r) K_l^{\beta\lambda}(\hat{r})$, centered at the impurity site at $\vec{r}=0$ and the quasi-band orbital $f_j(\vec{\rho})$ where $\vec{\rho} = \vec{r} - \vec{R}_p$, centered at $\vec{r} = \vec{R}_p$, are both given as products of radial functions and Kubic harmonics [c.f. Eqs. (27) and (B10), respectively]. By expanding the function

$$f_{\mu' l'}^{\beta' \lambda'}(\vec{\rho}) = F_{\mu' l'}(\rho) K_l^{\beta' \lambda'}(\hat{\rho}) \quad (C5)$$

in products of radial functions (to be given below) and Kubic harmonics around $\vec{r}=0$, we can easily calculate the two-center integrals in Eq. (C4) in terms of one-dimensional radial integrals. (In this appendix we let β , not α , denote the irreducible representation to avoid confusion with the α -expansion functions.) Thus, if we write the expansion of $f_j(\vec{r} - \vec{R}_p)$ around $\vec{r}=0$ as

$$f_{\mu' l'}^{\beta' \lambda'}(\vec{r}) = \sum_{l'_1=0}^{\infty} \sum_{\beta'_1 \lambda'_1} \alpha_{l'_1 \beta'_1 \lambda'_1}(\mu' l' \beta' \lambda', R_p | r) \times K_{l'_1}^{\beta'_1 \lambda'_1}(\hat{r}) , \quad (C6)$$

we get for the matrix element S_{II} in Eq. (C4),

$$S_{II} = \frac{1}{\sqrt{N}} \sum_{\vec{R}_p} e^{i\vec{k} \cdot \vec{R}_p} \int R_{\mu l}(r) \Theta(r) \alpha_{l\beta\lambda}(\mu' l' \beta' \lambda', R_p | r) r^2 dr . \quad (C7)$$

After some lengthy algebra, one obtains for the α -expansion functions the expression

$$\alpha_{l\beta\lambda}(\mu' l' \beta' \lambda', R_p | r) = \sum_{m'=-l'}^{l'} [\mathcal{D}^{(l)\dagger}(a, b, c)]_{\beta\lambda, m'} I_m^l(\mu' l', R_p | r) \cdot \mathcal{D}^{(l')}(a, b, c)_{m' \beta' \lambda'} . \quad (C8)$$

The quantities in this expression are defined as

$$I_m^l(\mu' l', R_p | r) = 2\pi k_{lm} k_{l'm} \frac{1}{r R_p} \int_{|R_p-r|}^{R_p+r} R F_{\mu' l'}(R) P_l^m \left[\frac{r^2 - R^2 - R_p^2}{2RR_p} \right] P_{l'}^m \left[\frac{r^2 - R^2 + R_p^2}{2rR_p} \right] dR , \quad (C9)$$

i.e., one-dimensional integrals, and

$$\mathcal{D}^{(l)}(a, b, c) = D^{(l)}(a, b, c) U^{(l)} , \quad (C10)$$

where the unitary matrix $D^{(l)}$ is the transformation matrix of spherical harmonics in the Condon-Shortley phase convention⁵⁴ under rotation of the Euler angles (a, b, c) :

$$Y_l^m(\theta, \phi) = \sum_{m'=-l}^l D_{m'm}^{(l)}(a, b, c) Y_l^{m'}(\theta', \phi'). \quad (\text{C11})$$

Furthermore, $U^{(l)}$ is the unitary matrix relating spherical and Cubic harmonics of angular-quantum number l :

$$K_l^{\beta\lambda}(\theta, \phi) = \sum_{m=-l}^l U_{m,\beta\lambda}^{(l)} Y_l^m(\theta, \phi). \quad (\text{C12})$$

The Euler angles (a, b, c) in Eqs. (C10) and (C11) correspond to a rotation of the coordinate system at $\vec{r}=0$ [the origin of the $g_a(\vec{r})$ functions] such that the polar axis (z axis) points along the vector \vec{R}_p . The explicit expression for the D matrices is⁵⁵

$$D_{m'm}^{(l)}(a, b, c) = e^{im'c} e^{ima} \sum_{\kappa} \frac{(-1)^{\kappa} [(l+m)!(l-m)!(l+m')!(l-m')!]^{1/2}}{\kappa!(l+m-\kappa)!(l-m'-\kappa)!(\kappa+m'-m)!} \\ \times \left[\cos \frac{b}{2} \right]^{2l-2\kappa-m'+m} \left[\sin \frac{b}{2} \right]^{2\kappa+m'-m}, \quad (\text{C13})$$

where the summation over κ includes all values that make the arguments of the factorials non-negative. Furthermore, $P_l^m(x)$ are the associated Legendre polynomials and k_{lm} are the spherical harmonic normalization constants (in the Condon-Shortley phase convention):

$$Y_l^m(\theta, \phi) = k_{lm} P_l^m(\cos\theta) e^{im\phi}, \quad m \geq 0 \\ Y_l^{-m}(\theta, \phi) = (-1)^m Y_l^{m*}(\theta, \phi), \quad m \geq 0 \quad (\text{C14})$$

$$k_{lm} = (-1)^m \left[\frac{2l+1}{4\pi} \right]^{1/2} \left[\frac{(l-m)!}{(l+m)!} \right]^{1/2}.$$

In deriving these expressions, we have assumed that the coordinate system at $\vec{r}=\vec{R}_p$ is related to the fixed coordinate system at $\vec{r}=0$ through a translation by \vec{R}_p (i.e., the two-coordinate systems have the same orientation). Hence, their polar axes are parallel, but do not generally point along the direction of \vec{R}_p . The interpretation of the various factors in Eq. (C8) is then the following: The \mathcal{D} -matrix elements to the right of the “transfer integral” I first express the function $f_j(\vec{\rho})$ in Eq. (C5) at $\vec{r}=\vec{R}_p$ in terms of spherical harmonics and then express this function in a rotated coordinate system whose polar axis is in the direction of \vec{R}_p . Then the transfer integral I gives the resulting function in a rotated coordinate system at $\vec{r}=0$, whose polar axis is also in the direction of \vec{R}_p . This integral alone (without the \mathcal{D} matrices) is the only ingredient in the α -expansion technique derived by Löwdin.⁵⁶ Finally, the \mathcal{D} matrix to the left of the I gives the function in terms of spheri-

cal harmonics in a coordinate system rotated back to its original orientation and expresses this function in terms of Cubic harmonics in that coordinate system. Finally, we note that a general symmetrized form of the Green’s-function matrix elements with the \vec{k} -space summation running only over the irreducible part of the Brillouin zone was given in Ref. 29(a).

APPENDIX D: WAVE-FUNCTION EXPANSION COEFFICIENTS AND NORMALIZATION CONDITIONS

The dual representation of the defect wave functions $\psi_i^{\alpha,\lambda}(\vec{r})$ is given in our QBCF model by Eqs. (29) and (30):

$$\Theta(\vec{r}-\vec{R}_c) \psi_i^{\alpha,\lambda}(\vec{r}) = \sum_l \left[\sum_{\mu} C_{\mu li}^{\alpha} F_{\mu l}(|\vec{r}|) \right] K_l^{\alpha,\lambda}(\hat{r}) \quad (\text{D1})$$

and

$$\psi_i^{\alpha,\lambda}(\vec{r}) = \sum_j^{\bar{M}} \sum_{\vec{k}}^{\text{BZ}} A_{ij}^{\alpha,\lambda}(\vec{k}) \phi_j^{\text{QB}}(\vec{k}, \vec{r}), \quad (\text{D2})$$

i.e., expansion in a crystal-field series (D1) and in quasibands (D2). In this appendix we show (i) only the coefficients $\{c_{\mu li}^{\alpha}\}$ in Eq. (D1) need to be solved since $\{A_{ij}^{\alpha,\lambda}\}$ in Eq. (D2) can be obtained from them through the Lippman-Schwinger equation, (ii) the resulting wave-function normalization conditions, and (iii) how to reduce sums over the

full BZ [Eq. (D2)] to sums over the irreducible BZ (IBZ). The final results of this appendix are given in Eqs. (D3), (D5), and (D8).

Since the expansion in (D1) is valid only in the subspace $\vec{r} \leq \vec{R}_c$ of the perturbation, we cannot directly use this expression to normalize the wave functions over all space. Instead, we substitute the

expression (19) for the Green's function $G^0(\vec{r}, \vec{r}', \epsilon)$ into the Lippman-Schwinger equation (18), using for $\psi_i(\vec{r})$ in the left-hand side of Eq. (18) the expansion (D2) and for $\psi_i(\vec{r})$ in the right-hand side of Eq. (18) the expansion (D1). This provides a relationship between the two sets of expansion coefficients:

$$A_{ij}^{\alpha\lambda}(\vec{k}) = \frac{1}{\epsilon_i - \epsilon_j^{\text{QB}}(\vec{k})} \sum_{\mu l} C_{\mu l i}^{\alpha} \sum_{\mu''} \sum_{\mu' l'} \sum_{\alpha' \lambda'} \langle \phi_j^{\text{QB}}(\vec{k}, \vec{r}) | \Theta(\vec{r}) | g_{\mu'' l'}^{\alpha' \lambda'} \rangle (\underline{S}^{-1})_{\mu'' \mu'} \langle g_{\mu' l'}^{\alpha' \lambda'}(\vec{r}) | \Delta V(\vec{r}) | g_{\mu l}^{\alpha\lambda}(\vec{r}) \rangle. \quad (\text{D3})$$

It is important to emphasize that these results do not require formally that the local orbital set $\{g_{\mu l}^{\alpha\lambda}\}$ be complete. Instead, it is required that the finite set [see Eq. (D2)] of quasiband wave functions can be expanded in terms of the set $\{g_{\mu l}^{\alpha\lambda}\}$:

$$\Theta(\vec{r}) \phi_j^{\text{QB}}(\vec{k}, \vec{r}) = \sum_{\mu'} \sum_{\mu l} \sum_{\alpha \lambda} \langle \phi_j^{\text{QB}}(\vec{k}, \vec{r}) | g_{\mu l}^{\alpha\lambda}(\vec{r}) \rangle^* (\underline{S}^{-1})_{\mu' \mu} g_{\mu l}^{\alpha\lambda}(\vec{r}). \quad (\text{D4})$$

The accuracy of this expansion is examined through convergence tests. Reference 26 discussed an example (a Cr impurity in a silicon free-electron host) and illustrates numerically the high degree of precision obtained in such an expansion with a moderate size basis set.

Using Eq. (D3), one can now normalize the wave function in the following fashion:

$$\frac{N_{\Omega}}{N_{\omega}} \sum_j \sum_{\vec{k}_v}^{\text{BZ}} |A_{ij}^{\alpha\lambda}(\vec{k}_v)|^2 = 1, \quad (\text{D5})$$

where N_{Ω} and N_{ω} are, respectively, the number of atoms per supercell and primitive cell, and $\{\vec{k}_v\}$ denotes the discrete wave vectors included in the expansion (D2). We have assumed that the quasiband wave functions $\{\phi_j^{\text{QB}}(\vec{k}, \vec{r})\}$ are normalized to unity over the primitive unit cell.

We now seek to modify Eq. (D5) for a summation over the IBZ rather than the full BZ. For concreteness we illustrate this for a defect with T_d symmetry in a zinc-blende host crystal. Let m_v denote the number of vectors in the star of \vec{k}_v with respect to the group O_h of order 48 ($O_h = I \times T_d$ where I is the inversion group), and let R_m denote operations in the group O_h . We get from Eq. (D5),

$$1 = \frac{N_{\Omega}}{N_{\omega}} \sum_j \sum_{\vec{k}_v}^{\text{IBZ}} \sum_m^{m_v} |A_{ij}^{\alpha\lambda}(R_m \vec{k}_v)|^2 \\ = \frac{N_{\Omega}}{N_{\omega}} \sum_j \sum_{\vec{k}_v}^{\text{IBZ}} \frac{m_v}{48} \sum_{m=1}^{24} \{ |A_{ij}^{\alpha\lambda}(R_m \vec{k}_v)|^2 \\ + |A_{ij}^{\alpha\lambda}(-R_m \vec{k}_v)|^2 \}. \quad (\text{D6})$$

One can prove that

$$\sum_j \sum_{m=1}^{24} |A_{ij}^{\alpha\lambda}(\pm R_m \vec{k}_v)|^2 \\ = \frac{24}{d_{\alpha}} \sum_j \sum_{\lambda=1}^{d_{\alpha}} |A_{ij}^{\alpha\lambda}(\vec{k}_v)|^2, \quad (\text{D7})$$

where d_{α} is the dimensionality of the α th irreducible representation. Thus, the normalization condition Eq. (D5) becomes

$$\frac{N_{\Omega}}{N_{\omega}} \sum_j \sum_{\vec{k}_v}^{\text{IBZ}} \left(\frac{m_v}{d_{\alpha}} \right) \sum_{\lambda=1}^{d_{\alpha}} |A_{ij}^{\alpha\lambda}(\vec{k}_v)|^2 = 1, \quad (\text{D8})$$

for each state i belonging to irreducible representation α in the point-symmetry group T_d .

APPENDIX E: CRYSTAL-FIELD KUBIC HARMONICS REPRESENTATIONS FOR DEFECT POTENTIALS AND CHARGE DENSITIES

We have stated in the text [Eqs. (33)–(36)] that the self-consistent perturbations $\Delta(\vec{r})$ [$\Delta(\vec{r})$ is either the charge-density perturbation $\Delta\rho(\vec{r})$, Eq. (33), the interelectronic Coulomb potential perturbation $\Delta V^{\text{ec}}(\vec{r})$, Eq. (35), or the perturbation in the exchange-correlation potential $\Delta V^{\text{xc}}(\vec{r})$, Eq. (36)] can be written in a one-center crystal-field form as

$$\Delta(\vec{r}) = \sum_l \delta_l(|\vec{r}|) K_l^{\alpha}(\hat{r}), \quad (\text{E1})$$

where $\delta_l(|\vec{r}|)$ is a one-dimensional radial func-

tion. In this appendix we derive explicit closed-form expressions for $\delta_l(|\vec{r}|)$ for T_d defects in a zinc-blende host crystal. The final results of this appendix are given in Eqs. (E16)–(E20), and (E26).

1. Charge density of the defected system

Using our expression (29) for the defect wave functions $\psi_i^{\alpha,\lambda}(\vec{r})$ in terms of local orbitals, we obtain

$$\Theta(\vec{r}-\vec{R}_c)\psi_i^{\alpha,\lambda}(\vec{r}) = \sum_l G_{il}^\alpha(|\vec{r}|)K_l^{\alpha,\lambda}(\hat{r}). \quad (\text{E2})$$

The charge density within the region defined by \vec{R}_c for spin σ ($\sigma = \uparrow$ or \downarrow) is

$$\begin{aligned} \rho_D^\sigma(\vec{r}) &= \sum_{i,(\sigma)}^{\text{occ}} \sum_{\alpha}^{d_\alpha} \sum_{\lambda=1}^{d_\alpha} |\Theta(\vec{r}-\vec{R}_c)\psi_i^{\alpha,\lambda}(\vec{r})|^2 \\ &= \sum_l \rho_{lD}^\sigma(|\vec{r}|)K_l^{\alpha,1}(\hat{r}), \end{aligned} \quad (\text{E3})$$

where

$$\begin{aligned} \rho_{lD}^\sigma(|\vec{r}|) &= \sum_{i(\sigma)}^{\text{occ}} \sum_{\alpha} \sum_{l''} D_l^\sigma[l',l''] \\ &\quad \times G_{il}^\alpha(|\vec{r}|)G_{il''}^\sigma(|\vec{r}|) \end{aligned} \quad (\text{E4})$$

and the Gaunt coefficients (Table II) are defined by

$$\sum_{\lambda=1}^{d_\alpha} K_l^{\alpha,\lambda}(\hat{r})K_{l'}^{\alpha,\lambda}(\hat{r}) = \sum_{l''} D_l^\sigma[l',l'']K_{l''}^{\alpha,1}(\hat{r}). \quad (\text{E5})$$

The calculation of $\rho_D^\sigma(\vec{r})$ in Eq. (E3) requires, therefore, the knowledge of the variational coefficients $\{C_{\mu i}^\alpha\}$ and radial basis functions $\{F_{\mu l}(|\vec{r}|)\}$ that together specify $\{G_{il}^\alpha(|\vec{r}|)\}$ [Eq. (29)], as well as the (tabulated) Gaunt coefficients $\{D_l^\sigma[l',l'']\}$.

2. Charge density of the host crystal

We need to calculate the host-crystal charge density $\rho_H(\vec{r})$ in a crystal-field representation [Eq. (32)] so that: (i) The expression will follow closely the expansion (E2) for the defected system (i.e., in calculating the charge-density perturbation $\rho_D-\rho_H$, systematic errors will tend to cancel algebraically rather than numerically), and (ii) the expansion

needs to be valid only in the domain $|\vec{r}| \leq |\vec{R}_c|$, much like in Eq. (E3). Since the host wave functions are often available as a Fourier series [c.f. Eq. (B2)], one might construct from it a similar series for the host charge density and proceed to l decompose each of the plane waves in that series to achieve the form (E1). This procedure does not, however, follow the previous conditions (i) and (ii) above. Instead, we expand the quasiband wave functions in our local crystal-field basis $\{g_a(\vec{r})\}$ in the domain $|\vec{r}| \leq |\vec{R}_c|$,

$$\Theta(\vec{r})\phi_j^{\text{QB}}(\vec{k},\vec{r}) = \sum_a C_{ja}(\vec{k})g_a(\vec{r}), \quad (\text{E6})$$

[much like the expansion of the defect wave functions in Eq. (16)] and calculate the charge density from the sum of squares of Eq. (E6):

$$\begin{aligned} \rho_H(\vec{r}) &= 2 \sum_j^{\text{occ}} \sum_{\vec{k}}^{\text{BZ}} |\Theta(\vec{r})\phi_j^{\text{QB}}(\vec{k},\vec{r})|^2 \\ &= 2 \sum_{a,b} g_a^*(\vec{r}) \left[\sum_j^{\text{occ}} \sum_{\vec{k}}^{\text{BZ}} C_{ja}^*(\vec{k})C_{jb}(\vec{k}) \right] g_b(\vec{r}), \end{aligned} \quad (\text{E7})$$

followed by partitioning $g_a(\vec{r})$ to a product of a radial and angular parts [Eq. (32)] that breaks $\rho_H(\vec{r})$ in Eq. (E7) into the desired crystal-field form of Eq. (E1). We now derive explicit equations for the terms in Eq. (E7) in terms of known quantities.

The final results are given by Eqs. (E15) and (E16).

The expansion coefficients $\{C_{ja}(\vec{k})\}$ of Eq. (E6) can be found by taking the inner product in Eq. (E6), i.e.,

$$\begin{aligned} \langle g_a(\vec{r}) | \Theta(\vec{r}) | j\vec{k} \rangle &= \sum_b C_{jb}(\vec{k})S_{ab}, \\ S_{ab} &\equiv \langle g_a(\vec{r}) | g_b(\vec{r}) \rangle. \end{aligned} \quad (\text{E8})$$

The expansion coefficients in Eq. (E6) are then simply obtained as

$$C_{ja}(\vec{k}) = \sum_b (\underline{S}^{-1})_{ab} \langle g_b(\vec{r}) | \Theta(\vec{r}) | j\vec{k} \rangle. \quad (\text{E9})$$

With these expansion coefficients, the host charge density of Eq. (E7) becomes

$$\rho_H(\vec{r}) = 2 \sum_{a,b} g_a(\vec{r})(\underline{S}^{-1}\underline{B}\underline{S}^{-1})_{ab}g_b^*(\vec{r}), \quad (\text{E10})$$

where the elements of the matrix \underline{B} are given by [c.f. Ref. 29(a)]:

$$\begin{aligned} B_{ab} &= \sum_j^{\text{occ}} \sum_{\vec{k}}^{\text{BZ}} \langle g_a(\vec{r}) | \Theta(\vec{r}) | j\vec{k} \rangle \langle j\vec{k} | \Theta(\vec{r}) | g_b(\vec{r}) \rangle \\ &= \delta_{\alpha\alpha'}\delta_{\lambda\lambda'} \frac{N_\omega}{N_\Omega} \sum_{\vec{k}_v}^{\text{IBZ}} \frac{m_v}{d_\alpha} \sum_j^{\text{occ}} \sum_{\lambda=1}^{d_\lambda} \text{Re} \{ \langle g_{\mu'}^{\alpha\lambda'} | \Theta | j\vec{k}_v \rangle \langle j\vec{k}_v | \Theta | g_{\mu}^{\alpha\lambda} \rangle \} = B_{\mu'\mu}^{\alpha\lambda}(\alpha). \end{aligned} \quad (\text{E11})$$

We can make contact with the standard continuum Green's-function formalism by taking $\Theta(\vec{r})=1$, $\underline{S}=\underline{1}$. One then recognizes that the matrix elements B_{ab} of Eq. (E11) correspond to

$$B_{ab} = \frac{1}{\pi} \text{Im} \int_0^{E_F} G_{ab}^0(\epsilon) d\epsilon, \quad (\text{E12})$$

where E_F is the Fermi energy and that $\rho_H(\vec{r})$ of Eq. (E10) corresponds to

$$\rho_H(\vec{r}) = 2 \sum_{a,b} g_a(\vec{r}) B_{ab} g_b^*(\vec{r}), \quad (\text{E13})$$

as in standard derivations (e.g., Refs. 3 and 4).

To bring $\rho_H(\vec{r})$ in Eq. (E10) to a crystal-field form, we use our standard separation of variables for $\{g_a(\vec{r})\}$ [Eq. (27)], where, as in Eq. (11), the single index a is broken into $(\mu l \alpha \lambda)$ (c.f. Sec. V B 1):

$$g_a(\vec{r}) \equiv F_{\mu l}(|\vec{r}|) K_l^{\alpha, \lambda}(\hat{r}). \quad (\text{E14})$$

Using this form in Eq. (E10), one obtains the final result

$$\rho_H(\vec{r}) = \sum_{l''} \rho_{l'' H}(|\vec{r}|) K_{l''}^{\alpha_1}(\hat{r}), \quad (\text{E15})$$

where

$$\rho_{l'' H}(|\vec{r}|) = 2 \sum_{\alpha} \sum_{\mu l} \sum_{\mu' l'} D_{l''}^{\alpha} [l, l'] F_{\mu l}(|\vec{r}|) \tilde{B}_{l''}^{\mu \mu'}(\alpha) F_{\mu' l'}(|\vec{r}|). \quad (\text{E16})$$

Here $D_{l''}^{\alpha} [l, l']$ are the Gaunt coefficients [Eq. (E5)], and the matrix \tilde{B} is $(\underline{S}^{-1} \underline{B} \underline{S}^{-1})$ of Eq. (E10). Notice that the matrix \underline{B} is diagonal in the indices α and λ and that the matrix \underline{S} is diagonal in α , λ , and l . The calculation of $\rho_H(\vec{r})$ in Eq. (E15) requires the computation of the overlap integrals $\langle g_a(\vec{r}) | \Theta(\vec{r}) | j \vec{k} \rangle$. These were described in Appendix C.

3. The change in the charge density

Using expression (E3) for $\rho_D(\vec{r})$ and expression (E15) for $\rho_H(\vec{r})$ one obtains the final crystal-field form for the (spin-unpolarized) density perturbation [Eq. (33)]:

$$\Delta \rho(\vec{r}) = \sum_l \Delta \rho_l(|\vec{r}|) K_l^{\alpha_1}(\hat{r}), \quad (\text{E17})$$

with the coefficients:

$$\Delta \rho_l(|\vec{r}|) = \sum_i^{\text{occ}} \sum_{\alpha} \sum_{l'' \mu''} \sum_{\mu' l'} D_l^{\alpha} [l', l''] F_{\mu' l'}(|\vec{r}|) F_{\mu'' l''}(|\vec{r}|) (C_{\mu' l' i}^{\alpha} C_{\mu'' l'' i}^{\alpha} - 2 \tilde{B}_{l''}^{\mu' \mu''}). \quad (\text{E18})$$

Only the product $C_{\mu' l' i}^{\alpha} C_{\mu'' l'' i}^{\alpha}$ changes in self-consistency iterations; the other terms are specified through the fixed radial basis functions, the Gaunt coefficients, and the quasiband wave functions. In the classical crystal-field approach,^{14,17} $\Delta \rho_l$ is approximated by a set of point charges. Expression (E18) gives the generalization of this for realistic perturbation potentials.

4. The change in the Coulomb interelectronic repulsion

Given the density perturbation $\Delta \rho(\vec{r})$ in the crystal-field form of Eq. (E17), the change in the interelectronic Coulomb repulsion can be calculated directly from the definition of $V^{\text{ee}}[\rho]$ as

$$\Delta v^{\text{ee}}(\vec{r}) = \sum_l \Delta v_l^{\text{ee}}(|\vec{r}|) K_l^{\alpha_1}(\hat{r}), \quad (\text{E19})$$

where

$$\Delta v_l^{\text{ee}}(|\vec{r}|) = \frac{4\pi}{2l+1} \left[\frac{1}{r^{l+1}} \int_0^r r_1^{l+2} \Delta \rho_l(r_1) dr_1 + r^l \int_r^{\infty} r_1^{l-1} \Delta \rho_l(r_1) dr_1 \right]. \quad (\text{E20})$$

This involves performing two simple one-dimensional radial integrals, exactly like in atomic calculations. This procedure was tested successfully on a number of analytically solvable problems.

5. The change in the exchange-correlation potential

The exchange-correlation potential $\Delta V^{\text{xc}}[\rho_D, \rho_H]$ is nonlinear with respect to $\rho_D - \rho_H$ [cf. Eq. (8b)] and therefore cannot be evaluated from $\Delta\rho$. Instead, we use a simple procedure to bring it into the central-field form of Eq. (36). For concreteness we illustrate the method for a T_d site-symmetry. The expansion in Eq. (36) is then

$$\Delta V^{\text{xc}}(\vec{r}) = \Delta v_0^{\text{xc}}(r)K_0^{a_1}(\theta, \phi) + \Delta v_3^{\text{xc}}(r)K_3^{a_1}(\theta, \phi) + \Delta v_4^{\text{xc}}(r)K_4^{a_1}(\theta, \phi) + \Delta v_6^{\text{xc}}(r)K_6^{a_1}(\theta, \phi) + \dots, \quad (\text{E21})$$

where

$$\Delta V^{\text{xc}}(\vec{r}) = V^{\text{xc}}(\rho_D(\vec{r})) - V^{\text{xc}}(\rho_H(\vec{r})). \quad (\text{E22})$$

The familiar spin-polarized (for spin σ) form of $V^{\text{xc}}(\rho, \xi)$ is given as

$$V_{\sigma}^{\text{xc}}(\rho, \xi) = V_{\text{xc}}^{\text{u}}(\rho) + f(\xi)[V_{\text{xc}}^{\text{p}}(\rho) - V_{\text{xc}}^{\text{u}}(\rho)] + \{\epsilon_{\text{xc}}^{\text{p}}[\rho] - \epsilon_{\text{xc}}^{\text{u}}[\rho]\}[\text{sgn}(\sigma) - \xi] \frac{df}{d\xi}, \quad (\text{E23})$$

where the spin-polarization $\xi(\vec{r})$ is

$$\xi(\vec{r}) = \frac{\rho_+(\vec{r}) - \rho_-(\vec{r})}{\rho_+(\vec{r}) + \rho_-(\vec{r})},$$

and

$$f(\xi) = \frac{(1+\xi)^{4/3} + (1-\xi)^{4/3} - 2}{2^{4/3} - 2}, \quad (\text{E24})$$

and V_{xc}^{u} , V_{xc}^{p} , $\epsilon_{\text{xc}}^{\text{u}}$, and $\epsilon_{\text{xc}}^{\text{p}}$ are, respectively, the unpolarized and polarized potentials and the unpolarized and polarized exchange-correlation energies per particle. Given $\rho_D^{\sigma}(\vec{r})$ and $\rho_H(\vec{r})$, one can calculate the left-hand side of Eq. (E21) through the expression (E22)–(E24). Having n terms in the right-hand side of Eq. (E21), we choose n different directions (θ_i, ϕ_i) , $i = 1, \dots, n$ (which are inequivalent under the point-group operations) for each radial point $|\vec{r}_j|$. Defining

$$C_l^{(i)} = K_l^{a_1}(\theta_i, \phi_i) \quad (\text{E25})$$

(independent on the radial coordinate), we obtain for each radial point a set of n inhomogeneous with n unknowns $\Delta v_0^{\text{xc}}(r_j)$, $\Delta v_3^{\text{xc}}(r_j)$, $\Delta v_4^{\text{xc}}(r_j)$, $\Delta v_6^{\text{xc}}(r_j)$, etc., of the form:

$$\Delta V^{\text{xc}}(\vec{r}_j^{(i)}) = \Delta v_0^{\text{xc}}(r_j)C_0^{(i)} + \Delta v_3^{\text{xc}}(r_j)C_3^{(i)} + \dots \quad (\text{E26})$$

Arranging the $C_l^{(i)}$'s [Eq. (E25)] in a matrix \underline{C} [with indices (i, l)], the unknown coefficients $\Delta v_l^{\text{xc}}(r_j)$ in a vector \vec{P} (with index l), and the known values $\Delta V^{\text{xc}}(r_j^{(i)})$ in a vector \vec{Q} (with indices i), one obtains for each radial value r_j the linear

matrix equation,

$$\underline{C}\vec{P} = \vec{Q}, \quad (\text{E27})$$

which can be inverted directly since $\det |\underline{C}| \neq 0$ for n inequivalent directions. This yields the unknown coefficients $\Delta v_l^{\text{xc}}(|\vec{r}|)$ in Eq. (E21). In practice, we choose in the expansion (E21) more angular directions than l components so that Eq. (E27) becomes an overdetermined system of equations constructed from $l_{\text{max}} = 12$ in the expansion in Eq. (E21) (corresponding to 11 terms) and 19 uniformly spaced directions in the irreducible $\frac{1}{24}$ th of the unit cell. The method was tested successfully on a number of analytically solvable problems.

APPENDIX F: SOLVING THE GREEN'S-FUNCTION DETERMINENTAL EQUATIONS

To find the solutions to Eq. (17a) one needs to solve the determinental problem

$$D(\epsilon) \equiv \det \underline{Q}_1(\epsilon) = 0$$

and

$$\underline{Q}_1(\epsilon) = \underline{1} - \underline{G}^0(\epsilon)\underline{V}, \quad (\text{F1})$$

where \underline{V} denotes the matrix of the perturbation $\Delta V(\vec{r})$ in a local basis. This appendix discusses a very efficient method for solving these equations based on an extension of the idea suggested by Van der Avoird *et al.*⁶⁸

A straightforward solution of Eq. (F1) is very inefficient since it requires calculation of $D(\epsilon)$ on

the full energy scale with an energy mesh that can resolve any two roots. There exists however a very elegant algorithm, first used by Van der Avoird *et al.*,⁶⁸ for finding the roots of $D(\epsilon)$ efficiently. These authors considered the matrix

$$\underline{Q}_2(\epsilon) = \underline{V} \underline{G}^0(\epsilon) \underline{V} - \underline{V} = -\underline{V} \underline{Q}_1(\epsilon), \quad (\text{F2})$$

where \underline{V} is a nonsingular matrix, and showed that the eigenvalues of $\underline{Q}_2(\epsilon)$ always pass through zero with a negative slope, unless ϵ equals one of the host-crystal energies $\epsilon_j(\vec{k})$. In this case, a number of eigenvalues [equal to the degeneracy of $\epsilon_j(\vec{k})$] pass from $-\infty$ to $+\infty$.

From these properties of the eigenvalues of $\underline{Q}_2(\epsilon)$, it is clear that the number of roots $n(\epsilon_1, \epsilon_2)$ of Eq. (F1) in an energy interval $[\epsilon_1, \epsilon_2]$ not containing a pole in the Green's function is given by

$$n(\epsilon_1, \epsilon_2) = n(\epsilon_1) - n(\epsilon_2), \quad (\text{F3})$$

where $n(\epsilon)$ denotes the number of positive eigenvalues of $\underline{Q}_2(\epsilon)$. Thus, the energy roots of Eq. (F1) can be determined by repeated bisection of the intervals until a desired accuracy is obtained.

To find the number of positive eigenvalues $n(\epsilon)$ of \underline{Q}_2 at some energy ϵ , one can, of course, diagonalize \underline{Q}_2 directly, using standard techniques. A much faster method, however, is to bring \underline{Q}_2 into upper triangular form \underline{Q}_2^T by the Gauss elimination process and use the fact⁶⁹ that the number of positive eigenvalues of any real symmetric matrix \underline{A} equals the number of positive diagonal elements in \underline{A}^T .

In our applications, we have chosen to work with the matrix

$$\underline{Q}_3(\epsilon) = \underline{G}^0(\epsilon) - \underline{V}^{-1} \equiv -\underline{Q}_1(\epsilon) \underline{V}^{-1}, \quad (\text{F4})$$

rather than with $\underline{Q}_2(\epsilon)$. The reason for this is that the construction of the matrix $\underline{Q}_2(\epsilon)$ from the basic quantities $\underline{G}^0(\epsilon)$ and \underline{V} requires a double matrix multiplication $\underline{V} \underline{G}^0(\epsilon) \underline{V}$ for each trial energy $\epsilon = \epsilon_i$. The computer time for performing these operations is much longer than inverting \underline{V} once. Since $\underline{Q}_3(\epsilon)$ is a symmetric matrix with only real eigenvalues, the equation $\det \underline{Q}_3(\epsilon) = 0$ has the same roots as $D(\epsilon) = 0$ since $\det \underline{V} \neq 0$.

We now prove that the eigenvalues $\mu_i(\epsilon)$ of $\underline{Q}_3(\epsilon)$ obey the condition

$$\frac{d\mu_i(\epsilon)}{d\epsilon} \leq 0 \quad (\text{F5})$$

for all energies different from the unperturbed one-electron energies, so that the method described above for finding the roots of $D(\epsilon) = 0$ can still be

applied.

Proof. Let $\vec{a}_j(\epsilon)$ denote the eigenvectors of $\underline{Q}_3(\epsilon)$:

$$\underline{Q}_3(\epsilon) \vec{a}_j(\epsilon) = \mu_j(\epsilon) \vec{a}_j(\epsilon), \quad (\text{F6})$$

with the normalization requirement

$$\vec{a}_j^\dagger(\epsilon) \vec{a}_j(\epsilon) = 1. \quad (\text{F7})$$

Using the Hellmann-Feynman theorem, one gets that for all energies

$$\frac{d\mu_j(\epsilon)}{d\epsilon} = \vec{a}_j^\dagger(\epsilon) \frac{d\underline{G}^0(\epsilon)}{d\epsilon} \vec{a}_j(\epsilon). \quad (\text{F8})$$

Since the matrix $d\underline{G}^0(\epsilon)/d\epsilon$ is symmetric, there exists some unitary matrix $\underline{U}(\epsilon)$ such that $\underline{U}(\epsilon) \times [d\underline{G}^0(\epsilon)/d\epsilon] \underline{U}^\dagger(\epsilon)$ is diagonal. We get from Eq. (F8) using the notation:

$$\vec{b}_j(\epsilon) = \underline{U}(\epsilon) \vec{a}_j(\epsilon), \quad (\text{F9})$$

that

$$\frac{d\mu_j(\epsilon)}{d\epsilon} = \vec{b}_j^\dagger(\epsilon) \left[\underline{U}(\epsilon) \frac{d\underline{G}^0(\epsilon)}{d\epsilon} \underline{U}^\dagger(\epsilon) \right] \vec{b}_j(\epsilon). \quad (\text{F10})$$

We shall now show that the eigenvalues of $d\underline{G}^0(\epsilon)/d\epsilon$ (i.e., the diagonal elements in the matrix $\underline{U}(\epsilon) [d\underline{G}^0(\epsilon)/d\epsilon] \underline{U}^\dagger(\epsilon)$) are all nonpositive (the off-diagonal elements are, of course, zero). Consider a typical matrix element in the matrix $d\underline{G}^0(\epsilon)/d\epsilon$:

$$\frac{dG^0(\epsilon)_{\alpha\beta}}{d\epsilon} = - \sum_j \sum_{\vec{k}}^{\text{BZ}} \frac{\langle \alpha | j\vec{k} \rangle \langle j\vec{k} | \beta \rangle}{[\epsilon - \epsilon_j(\vec{k})]^2}. \quad (\text{F11})$$

After diagonalization we have

$$\begin{aligned} \frac{dG^0(\epsilon)_{\tilde{\alpha}\tilde{\beta}}}{d\epsilon} &\equiv \left[\underline{U}(\epsilon) \frac{d\underline{G}^0(\epsilon)}{d\epsilon} \underline{U}^\dagger(\epsilon) \right]_{\tilde{\alpha}\tilde{\beta}} \\ &= -\delta_{\tilde{\alpha},\tilde{\beta}} \sum_j \sum_{\vec{k}}^{\text{BZ}} \frac{\langle \tilde{\alpha} | j\vec{k} \rangle \langle j\vec{k} | \tilde{\beta} \rangle}{[\epsilon - \epsilon_j^0(\vec{k})]^2}. \end{aligned} \quad (\text{F12})$$

where

$$\langle \tilde{\alpha} | = \underline{U}(\epsilon) \langle \alpha |$$

and

$$| \tilde{\beta} \rangle = | \beta \rangle \underline{U}^\dagger(\epsilon). \quad (\text{F13})$$

Thus, the eigenvalues of $d\underline{G}^0(\epsilon)/d\epsilon$ are of the form

$$\frac{dG^0(\epsilon)_{\tilde{\alpha}\tilde{\alpha}}}{d\epsilon} = - \sum_j \sum_{\vec{k}}^{\text{BZ}} \frac{|\langle \tilde{\alpha} | j\vec{k} \rangle|^2}{[\epsilon - \epsilon_j^0(\vec{k})]^2} = -h_{\alpha\alpha}(\epsilon), \quad (\text{F14})$$

which are all nonpositive [$h_{\alpha\alpha}(\epsilon) \geq 0$], and when substituted into Eq. (F10), they give

$$\frac{d\mu_j(\epsilon)}{d\epsilon} = - \sum_{\alpha} h_{\alpha\alpha}(\epsilon) |b_j^{\alpha}|^2 \leq 0, \quad (\text{F15})$$

where b_j^{α} are the components in the vector \vec{b}_j .

Next we address the problem of solving the homogeneous system of linear equations

$$\underline{Q}_1(\epsilon_i) \vec{C}_i = 0, \quad (\text{F16})$$

where the components of the column vector \vec{C}_i are the expansion coefficients of the wave function $\psi_i(\vec{r})$ in terms of the local basis set $\{g_a(\vec{r})\}$:

$$\psi_i(\vec{r}) = \sum C_{ia} g_a(\vec{r}). \quad (\text{F17})$$

If we multiply Eq. (F16) from the left with the transpose of the real matrix $\underline{Q}_1(\epsilon_i)$, we get

$$\underline{Q}_1^t(\epsilon_i) \underline{Q}_1(\epsilon_i) \vec{C}_i = 0. \quad (\text{F18})$$

The matrix $\underline{Q}_1^t(\epsilon) \underline{Q}_1(\epsilon)$ is a real symmetric non-negative definite matrix; hence, all its real eigenvalues are larger than or equal to zero.

In practice, one rarely finds an energy $\epsilon = \epsilon_i$ such that the determinant of $\underline{Q}_1(\epsilon)$ is exactly zero. Therefore, Eqs. (F16) and (F18) have only the trivial solution $C_i = 0$. If, however, ϵ is sufficiently close to ϵ_i , the eigenvector that belongs to the smallest eigenvalue will be an excellent approximation to C_i . That is, for a certain root $\epsilon = \epsilon'$ we solve for the eigenvector $\vec{C}_i(\epsilon')$ that satisfies

$$\underline{Q}_1^t(\epsilon') \underline{Q}_1(\epsilon') \vec{C}_i(\epsilon') = \lambda^{\min}(\epsilon') \vec{C}_i(\epsilon'), \quad (\text{F19})$$

where $\lambda_{(\epsilon')}^{\min}$ is the smallest eigenvalue to the matrix $\underline{Q}_1^t(\epsilon') \underline{Q}_1(\epsilon')$. It is obvious that this vector $\vec{C}_i(\epsilon')$ minimizes the norm of the inhomogeneous term $\lambda^{\min} \vec{C}_i(\epsilon')$. This procedure requires some caution, however. If the calculated root ϵ' is too far from the true root ϵ_i , it need not necessarily be true that the smallest eigenvalue is the one that becomes zero at $\epsilon = \epsilon_i$. If this happens, the solution vector $\vec{C}_i(\epsilon')$ may be almost orthogonal to $\vec{C}_i(\epsilon_i)$. In practice we generally iterate the energy to a precision better than 0.00015 eV, so that the smallest eigenvalue is typically less than 10^{-10} , and the ratio between the smallest and next smallest eigenvalue is typically less than 10^{-8} .

APPENDIX G: KUBIC HARMONICS

Kubic harmonics of order $l=4$ belonging to the irreducible representations a_1 , e , t_1 and t_2 in the point-symmetry group T_d , are given in terms of Cartesian coordinates (x, y, z) with $r^2 = x^2 + y^2 + z^2$. Also given are the a_1 Kubic harmonics for $l=6, 7$, and 8 . The corresponding Kubic harmonics for $l=0, 1, 2$, and 3 have been previously published in the literature and can be found in Ref. 70. Notice that the phase convention we are using (leading, for example to the different signs of the radial wave-function components in Figs. 4 and 5, and the Gaunt coefficients in Table II) is that of Ref. 70 and those given next. The lowest Kubic harmonic belonging to the irreducible representation a_2 in the group T_d is of order $l=6$,

$$\begin{aligned} K_4^{a_1}(x, y, z) &= \left[\frac{1}{4\pi} \right]^{1/2} \left[\frac{21}{4} \right]^{1/2} \left[1 - \frac{5(x^2y^2 + x^2z^2 + y^2z^2)}{r^4} \right], \\ K_4^{e,1}(x, y, z) &= \left[\frac{1}{4\pi} \right]^{1/2} \left[\frac{5}{3} \right]^{1/2} \frac{3}{4} \left[\frac{3z^4 + 14x^2y^2 - 4x^2z^2 - 4y^2z^2}{r^4} - 1 \right], \\ K_4^{e,2}(x, y, z) &= \left[\frac{1}{4\pi} \right]^{1/2} \sqrt{5} \frac{3}{4} \frac{1}{r^4} [x^4 - y^4 - 6(x^2 - y^2)z^2], \\ K_4^{t_1,1}(x, y, z) &= \left[\frac{1}{4\pi} \right]^{1/2} \sqrt{35} \frac{3}{2} \frac{1}{r^4} (y^2 - z^2)yz, \\ K_4^{t_1,2}(x, y, z) &= \left[\frac{1}{4\pi} \right]^{1/2} \sqrt{35} \frac{3}{2} \frac{1}{r^4} (x^2 - z^2)xz, \\ K_4^{t_1,3}(x, y, z) &= \left[\frac{1}{4\pi} \right]^{1/2} \sqrt{35} \frac{3}{2} \frac{1}{r^4} (x^2 - y^2)xy, \end{aligned}$$

$$\begin{aligned}
K_4^{t_2,1}(x,y,z) &= \left[\frac{1}{4\pi} \right]^{1/2} \sqrt{5} \frac{3}{2} \frac{1}{r^4} (r^2 - 7x^2)yz, \\
K_4^{t_2,2}(x,y,z) &= \left[\frac{1}{4\pi} \right]^{1/2} \sqrt{5} \frac{3}{2} \frac{1}{r^4} (r^2 - 7y^2)xz, \\
K_4^{t_2,3}(x,y,z) &= \left[\frac{1}{4\pi} \right]^{1/2} \sqrt{5} \frac{3}{2} \frac{1}{r^4} (r^2 - 7z^2)xy, \\
K_6^a(x,y,z) &= \left[\frac{1}{4\pi} \right]^{1/2} \left[\frac{13}{2} \right]^{1/2} \frac{1}{4} \left[\frac{7(x^6 + y^6 + z^6 + 30x^2y^2z^2)}{r^6} - 5 \right], \\
K_7^a(x,y,z) &= \left[\frac{1}{4\pi} \right]^{1/2} \left[\frac{455}{3} \right]^{1/2} \frac{3}{4} \frac{1}{r^7} xyz [11(x^4 + y^4 + z^4) - 5r^4], \\
K_8^a(x,y,z) &= \left[\frac{1}{4\pi} \right]^{1/2} \sqrt{561} \frac{1}{8} \frac{1}{r^8} [x^8 + y^8 + z^8 - 14(x^6y^2 + x^6z^2 + y^6z^2 \\
&\quad + x^2y^6 + x^2z^6 + y^2z^6) + 35(x^4y^4 + x^4z^4 + y^4z^4)].
\end{aligned}$$

- ¹G. F. Koster and J. C. Slater, *Phys. Rev.* **96**, 1208 (1954); L. M. Roth, Ph.D. thesis, Harvard University, 1957 (unpublished); J. Callaway, *J. Math. Phys.* **5**, 783 (1964); P.O. Löwdin, *J. Mol. Spectrosc.* **14**, 112 (1964).
- ²J. Callaway and A. J. Hughes, *Phys. Rev.* **156**, 860 (1967).
- ³G. A. Baraff and M. Schlüter, *Phys. Rev. Lett.* **41**, 892 (1978); *Phys. Rev. B* **19**, 4965 (1979); G. A. Baraff, E. O. Kane, and M. Schlüter, *ibid.* **21**, 5662 (1980).
- ⁴J. Bernholc, N. Lipari, and S. T. Pantelides, *Phys. Rev. Lett.* **41**, 895 (1978); *Phys. Rev. B* **21**, 3545 (1980); J. Bernholc, S. T. Pantelides, N. O. Lipari, and A. Baldereschi, *Solid State Commun.* **37**, 705 (1981).
- ⁵P. Bendt and A. Zunger, *Phys. Rev. B* (in press).
- ⁶This formulation differs from the pseudopotential formulation of the defect problem introduced by S. T. Pantelides and C. T. Sah [*Phys. Rev. B* **10**, 638 (1974)] in which *core* projection operators are used.
- ⁷W. Kohn and L. J. Sham, *Phys. Rev.* **140**, A1133 (1965).
- ⁸A. Zunger and M. L. Cohen, *Phys. Rev. B* **18**, 5449 (1978); **20**, 4082 (1979); A. Zunger, *ibid.* **22**, 649 (1980).
- ⁹W. Kohn, in *Solid State Physics*, edited by F. Seitz and D. Turnbull (Academic, New York, 1957), Vol. 5.
- ¹⁰A. Baldereschi and J. Lipari, in *Proceedings of the Thirteenth International Conference on the Physics of Semiconductors, Rome, 1976*, edited by F. G. Fumi (Typographia Marves, Rome, 1977), p. 595; M. Jaros, *Adv. Phys.* **29**, 409 (1980).
- ¹¹S. T. Pantelides, *Rev. Mod. Phys.* **50**, 797 (1978).
- ¹²W. A. Harrison, *Pseudopotentials in the Theory of Metals* (Benjamin, New York, 1966).
- ¹³M. Altarelli and F. Bassani, in *Handbook of Semiconductors*, edited by W. Paul (North-Holland, Amsterdam, 1981).
- ¹⁴H. Bethe, *Ann. Phys. (Leipzig)* **3**, 135 (1929).
- ¹⁵J. H. Van Vleck, *Phys. Rev.* **41**, 208 (1932).
- ¹⁶C. J. Ballhausen, *Ligand Field Theory* (McGraw-Hill, New York, 1962); S. Sugano, Y. Tanabe, and H. Kamimura, *Multiplets of Transition-Metal Ions in Crystals*, (Academic, New York, 1970); J. S. Griffith, *The Theory of Transition-Metal Ions* (Cambridge University Press, London, 1971).
- ¹⁷H. Watanabe, *Operator Methods in Ligand Field Theory* (Prentice-Hall, Englewood Cliffs, 1966), p. 84.
- ¹⁸W. H. Kleiner, *J. Chem. Phys.* **20**, 1784 (1952).
- ¹⁹R. G. Shulman and S. Sugano, *Phys. Rev. Lett.* **7**, 157 (1961).
- ²⁰C. A. Coulson and M. J. Kearsley, *Proc. R. Soc. London Ser. A* **241**, 433 (1957).
- ²¹G. D. Watkins and R. P. Messmer, *Phys. Rev. Lett.* **25**, 656 (1970); *Phys. Rev. B* **7**, 2568 (1978).
- ²²C. Weigel, D. Peak, J. C. Corbett, G. D. Watkins, and R. P. Messmer, *Phys. Rev. B* **8**, 2906 (1973).
- ²³L. A. Hemstreet, *Phys. Rev. B* **15**, 834 (1977).
- ²⁴B. Cartling, B. Roos, and U. Wahlgren, *Chem. Phys. Lett.* **21**, 380 (1973); B. Cartling, *J. Phys. C* **8**, 3171 (1975); **8**, 3183 (1975).
- ²⁵A. Zunger and A. Katzir, *Phys. Rev. B* **11**, 2378 (1975); A. Zunger, *J. Phys. Chem. Solids* **36**, 229 (1975).
- ²⁶U. Lindefelt and A. Zunger, *Phys. Rev. B* **24**, 5913 (1981).
- ²⁷M. Jaros and S. Brand, *Phys. Rev. B* **14**, 4494 (1976); S. Brand, M. Jaros, and C. O. Rodrigues, *J. Phys. C* **14**, 1243 (1981).
- ²⁸J. L. Beeby, *Phys. Rev.* **137**, A933 (1965); G. E. Kil-

- by, Proc. Phys. Soc. London 90, 181 (1967); W. B. Fowles, Phys. Status Solidi 26, 639 (1968). See, also, discussion by A. M. Stoneham, in *Theory of Defects in Solids* (Clarendon, Oxford, 1975), pp. 87–114. We are grateful to M. Stoneham for pointing out these references to us.
- ²⁹(a) U. Lindefelt, J. Phys. C 11, 3651 (1978); Corrigendum J. Phys. C. 12, 959 (1979); (b) U. Lindefelt, J. Phys. C. 12, L419 (1979); (c) U. Lindefelt and S. T. Pantelides, Solid State Commun. 31, 631 (1979).
- ³⁰(a) G. H. Ludwig and H. H. Woodbury, *Solid State Physics*, edited by F. Seitz and D. Turnbull (Academic, New York, 1961), p. 223; (b) J. W. Chen and A. G. Milnes, Ann. Rev. Mater. Sci. 10, 157 (1980).
- ³¹H. J. Hovel, *Semiconductors and Semimetals Vol. II Solar Cells* (Academic, New York, 1975).
- ³²A. A. Bergh and J. P. Dean, *Light Emitting Diodes* (Oxford University Press, Clarendon, 1976).
- ³³G. D. Watkins and R. P. Messmer, in *Computational Methods for Large Molecules and Localized States in Solids*, edited by F. Herman, A. D. McLean, and R. K. Nesbet (Plenum, New York, 1972), p. 133.
- ³⁴(a) A. Zunger, J. Chem. Phys. 62, 1861 (1975); (b) 7, 96 (1974); (c) A. Zunger and R. Englman, Phys. Rev. B 17, 642 (1978).
- ³⁵S. G. Louie, M. Schlüter, J. R. Chelikowsky, and M. L. Cohen, Phys. Rev. B 13, 1654 (1976).
- ³⁶J. Bernholc and S. T. Pantelides, Phys. Rev. B 18, 1780 (1978).
- ³⁷H. P. Hjalmarson, P. Vogel, D. J. Wolford, and J. D. Dow, Phys. Rev. Lett. 44, 810 (1980).
- ³⁸V. Singh, U. Lindefelt, and A. Zunger, Phys. Rev. B 25, 2781 (1982).
- ³⁹(a) S. Huzinaga, J. Chem. Phys. 42, 1293 (1965); (b) M. J. Ten Hoor, Int. J. Quantum Chem. 17, 1099 (1980).
- ⁴⁰J. C. Phillips, J. Phys. Chem. Solids 11, 226 (1959).
- ⁴¹F. Garcia-Moliner, in *Theory of Imperfect Crystalline Solids; Trieste Lectures, 1970* (IAEA, Vienna, 1971).
- ⁴²G. Gilat and L. J. Raubenheimer, Phys. Rev. 144, 390 (1966).
- ⁴³G. Lehmann and M. Taut, Phys. Status Solidi. 54B, 469 (1972).
- ⁴⁴In a recent version of their model, M. Scheffler, S. T. Pantelides, N. O. Lipari, and J. Bernholc [Phys. Rev. Lett. 47, 413 (1981)] are no longer fitting their band structure but rather using it directly to construct the Green's function, as done here.
- ⁴⁵U. Lindefelt, J. Phys. C 11, 85 (1978).
- ⁴⁶R. A. Evarestov and V. A. Lovchikov, Phys. Status Solidi 79B, 743 (1977); R. A. Evarestov, *ibid.* 72B, 569 (1975).
- ⁴⁷H. Nara and A. Morita, J. Phys. Soc. Jpn. 21, 1852 (1966). We use $\epsilon(q=0)=11.4$.
- ⁴⁸F. Herman and S. Skillman, *Atomic Structure Calculations* (Prentice-Hall, Englewood Cliffs, 1963).
- ⁴⁹G. W. Pratt, Phys. Rev. 88, 1217 (1952).
- ⁵⁰G. P. Kerker, Phys. Rev. B 23, 3082 (1981).
- ⁵¹L. G. Ferriera, J. Comp. Phys. 36, 198 (1980).
- ⁵²F. Bassani, G. Iadonisi, and B. Preziosi, Phys. Rev. 186, 735 (1969).
- ⁵³R. N. Euwema, Phys. Rev. B 4, 4332 (1971); R. A. Deegan and W. D. Twose, Phys. Rev. 164, 993 (1967); A. B. Kunz, J. Phys. C 3, 1542 (1970).
- ⁵⁴E. U. Condon and G. H. Shortley, *The Theory of Atomic Spectra* (Cambridge University Press, London, 1967).
- ⁵⁵M. Tinkham, *Group Theory and Quantum Mechanics* (McGraw-Hill, New York, 1964).
- ⁵⁶P. O. Löwdin, Adv. Phys. 5, 1 (1956).
- ⁵⁷R. S. Mulliken, J. Chem. Phys. 1, 492 (1933). See also, 36, 3428 (1962); P. Politzer and R. S. Mulliken, *ibid.* 55, 5135 (1971).
- ⁵⁸R. A. Buckingham, H. S. Massey, and S. R. Tibbs, Proc. R. Soc. London Ser. A 178, 119 (1941).
- ⁵⁹M. Cohen and C. A. Coulson, Proc. Cambridge Philos. Soc. 57, 96 (1961).
- ⁶⁰H. W. Joy and R. G. Parr, J. Chem. Phys. 28, 448 (1958).
- ⁶¹J. P. Desclaux and P. Pyykkö, Chem. Phys. Lett. 39, 300 (1976); 50, 503 (1977); Chem. Phys. 34, 261 (1978).
- ⁶²(a) D. M. Bishop, in *Advances in Quantum Chemistry*, edited by P. O. Löwdin (Academic, New York, 1967), Vol. 3, p. 25; (b) H. H. Kranz and E. O. Steinborn, Phys. Rev. A 25, 66 (1982).
- ⁶³K. S. Song and K. L. Malwal, Phys. Rev. B 13, 4314 (1976).
- ⁶⁴J. C. Slater, in *Quantum Science*, edited by J. L. Calais, O. Goscinski, J. Linderberg, and Y. Öhrn (Plenum, New York, 1976), p. 215.
- ⁶⁵E. Fues, Ann. Phys. (Leipzig) 80, 367 (1926); J. D. Settler and R. A. Shats, Int. J. Quantum Chem. S5, 277 (1971); V. Singh, U. Lindefelt, and A. Zunger (unpublished).
- ⁶⁶R. Fletcher and M. J. D. Powell, Comput. J. 6, 163 (1963); R. Fletcher, *ibid.* 13, 317 (1970); W. C. Davidson, *ibid.* 10, 406 (1968).
- ⁶⁷P. P. Puley, Chem. Phys. Lett. 73, 393 (1980).
- ⁶⁸A. Van der Avoird, S. P. Liebmann, and D. J. M. Fassaert, Phys. Rev. B 10, 1230 (1974).
- ⁶⁹J. H. Wilkinson, *The Algebraic Eigenvalue Problem* (Clarendon, Oxford, 1965), p. 203.
- ⁷⁰F. Bassani and G. Pastori Parravicini, *Electronic States and Optical Transitions in Solids* (Pergamon, New York, 1973), pp. 24.
- ⁷¹J. R. Chelikowsky (private communication).
- ⁷²M. Jaros, C. O. Rodriguez, and S. Brand, Phys. Rev. B 19, 3137 (1979).
- ⁷³P. K. Lam, M. L. Cohen, and A. Zunger, Phys. Rev. B 22, 1698 (1980).
- ⁷⁴A. Zunger, J. P. Perdew, and G. L. Oliver, Solid State Commun. 34, 933 (1980); J. P. Perdew and A. Zunger, Phys. Rev. B 23, 5048 (1981).
- ⁷⁵A. Zunger and U. Lindefelt (unpublished).
- ⁷⁶R. Nieminen and J. C. Hodges, J. Phys. F 6, 573 (1976).