PHYSICAL REVIEW B VOLUME 30, NUMBER 4 15 AUGUST 1984

Calculation of the total energy of charged point defects using the Green's-function technique

G. A. Baraff and M. Schlüter
AT&T Bell Laboratories, Murray Hill, New Jersey 07974

(Received 24 January 1984; revised manuscript received 11 April 1984)

We present a method for calculating the energy of an isolated, charged, deep-level point defect in an otherwise-perfect infinite crystal. To simplify the evaluation, one can make the usual replacement of one-particle energy terms by a sum over occupied eigenvalues plus corrections for overcounting. However, this simplification leads to conceptual difficulties when the defect is charged. These are overcome by truncating the long-range tail of the defect potential. The correct screening charge does not appear automatically when a truncated potential is used, and so a constant potential shift is added to guarantee proper screening. Alternatively, one can evaluate the original kinetic-energy form of the functional. We do so and compare it with the eigenvalue formulation. A careful study of the effects of truncation indicates that, although the two formulations are equivalent for neutral systems, they are not so for charged systems. The calculated energy of charged defect must differ slightly when evaluated by the two methods. The truncation error is greater for the eigenvalue formulation than for the kinetic-energy formulation. However, this difference is expected to be in the (0.1—0.2)-eV range for reasonable truncation radii, and to be quite insensitive to atomic displacements. This may be sufficiently small and insensitive, depending on the situation being studied, that the greater numerical simplicity of the eigenvalue formulation would make it the method of choice. If not, the kinetic-energy formulation presents no major difficulties. Consequences of shifting the conduction-band eigenvalues, as a way of overcoming the small-band-gap problem inherent in the use of local-density-functional theory are explored.

I. INTRODUCTION

This paper deals with how to formulate and evaluate the energy of an isolated point defect, that is, the difference in total energy between an infinite crystal containing a single isolated point defect and the total energy of the same infinite, but perfect crystal. This difference, the defect energy, plays the central role in controlling the behavior of the defect, e.g., its equilibrium configuration, its vibrational properties, and the various reactions in which it may participate. In the case of semiconductor crystals, most defects can exist in one of several charge states. Their behavior will then be charge-state dependent, giving rise to an enormous range of interesting and occasionally technologically important effects.[1-9]

Calculating the defect energy for an isolated *charged* point defect is complicated by the presence of a long-range Coulomb tail in the defect potential. We have recently called attention to this problem and have described a method of calculating the defect energy within the local-density-functional formalism.[10] That method made use of the self-consistent Green's-function technique[11,12] for the electronic structure of the defect crystal and the *eigenvalue* formulation of the total-energy expression.[13] The problem addressed in that paper was the inherent contradiction of using the Green's-function technique when the defect potential had a Coulomb tail. Two purposes of the present paper are (a) to present the formulation of Ref. 10 in much fuller detail, and (b) to describe improvements, either in accuracy or efficiency, which could not be included within the format constraints of Ref. 10.

In this paper we also examine the calculational alternative of using the Green's-function technique for the electronic structure of the defect crystal in conjunction with the *kinetic-energy* formulation of the total-energy expression. We shall show that the practical necessity of truncating the Green's function when the defect is charged causes the two formulations to give different values of the defect energy. In the course of this study we find that the kinetic-energy formulation is consistent with the requirements of dielectric response theory, while the eigenvalue formulation is, in principle, not. Our preliminary studies suggest that, for reasonable truncation radii, the difference between the two formulations may lie in the (0.1—0.2)-eV range, and should be quite insensitive to atomic displacements. This may be sufficiently small and insensitive, depending on the situation being studied, to recommend use of the numerically simpler eigenvalue formulation. If not, use of the kinetic-energy formulation presents no major difficulties.

The central problem for formulating the defect energy is how to separate and recombine short-range and long-range effects. There are several aspects to this, but before treating them, let us outline the general viewpoint and then outline the ideas we shall explore. To describe a point defect, we first consider an infinite, perfect semiconductor crystal. The point defect is created by, e.g., removing one crystal atom, or replacing it by an impurity atom, or inserting some atom into an interstitial location, or by similar operations. We define a sphere of radius R_1 containing region 1 in which strong chemical effects occur as a result of creation of the defect (see Fig. 1). These effects are, e.g., the formulation of dangling bonds near a vacan-

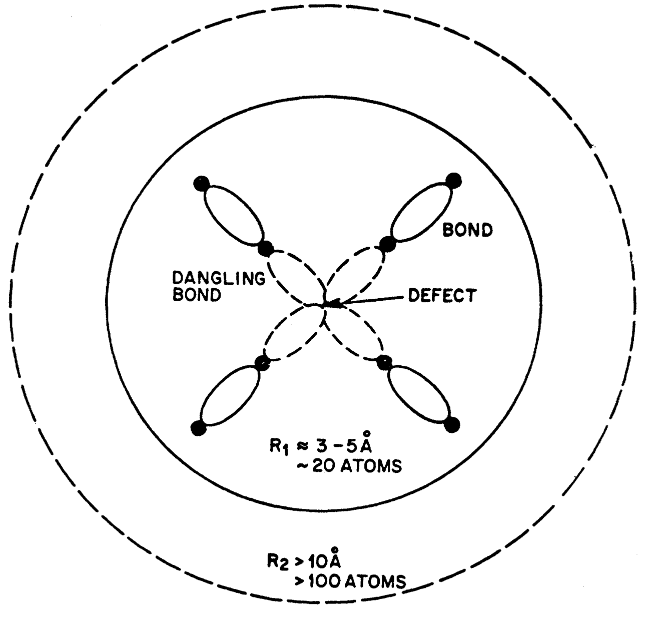

FIG. 1. Operational subdivision of space around a localized lattice defect. Region 1 contains all the primary changes in local chemistry; region 2 can be regarded as reacting macroscopically to changes in region 1.

cy, or conversely, the attempt to form new bonds between an extra interstitial atom and its neighbors, or simply the modification in the strength of bonds around a substitutional impurity. In general, atoms in region 1 will distort significantly from their original perfect-crystal position. R_1 is thus operationally defined as being the range of direct chemical action leading to strong changes in the electronic structure. In semiconductors, R_1 is typically a few screening lengths ($\approx 3-5$ Å) and region 1 contains of the order of 30 atoms.

Surrounding region 1 is a much larger region, region 2, in which the crystal's response to the defect can best be described by elastic and dielectric means. The point is that small distortions and polarizations can persist, in principle, up to infinite range without significant perturbations on a local scale. A useful description has to include both regions 1 and 2, although treated with different means and emphasis.

The tools we use for treating region 1 include the local-density-functional formalism,[14,15] nonlocal norm-conserving[16] pseudopotentials, and a Green's-function technique for implementing the calculations in an isolated-point-defect geometry. Two conceptual problems arise at this stage: (a) the energy-gap problem, and (b) the truncation problem. The energy-gap problem arises because local-density-functional theory used with first-principles pseudopotentials or in all-electron calculations of semiconductor band structure gives an eigenvalue difference $\epsilon_c - \epsilon_v$ between the bottom of the conduction band and the top of the valence band which is consistently smaller than E_g, the experimental band gap.[17,18] Although there is fundamental research in progress aimed at resolving this problem,[19–22] there is at present no scheme for doing so which is sufficiently simple to be used with self-consistent defect calculations. Among the various palliative techniques that have been applied to defect calculations for coping with the band-gap problem are *ad hoc* adjustment of the local exchange-correlation potential,[23] *ad hoc* adjustment of the pseudopotential,[24] and *ad hoc* adjustment of the conduction-band eigenvalues themselves.[25] We have been doing the latter for some time now.[26] In this work we examine the effect of the adjustment of the conduction-band eigenvalues on the calculated defect-energy and bound-state eigenvalues.

The truncation problem arises because the Green's-function method really presupposes the use of a defect potential with finite range.[27] Consequently, the form chosen to represent the Green's function, no matter how accurate it is in the neighborhood of the defect, is allowed to fail completely, e.g., identically, far from the defect.[11,12,23] Clearly, one should carefully consider the consequences of using such a truncated Green's function with an infinite-range defect potential. This we will do here.

The tools we use for investigating the outer region are macroscopic dielectric theory for the long-range polarization and a spring model for the long-range stress. We mention the latter for completeness only. Actually, there is no difficulty in joining the atomic displacements at the inner part of region 2 with those at the outer part of region 1 because the energy associated with relaxation of the more distant atoms can be subtracted afterwards from a total-energy calculation in which the energy is calculated with the distant atoms held fixed in their unrelaxed positions. A wide variety of sping models would probably serve equally well here. No specific model will be discussed in this paper.

Unlike the situation for the long-range mechanical stresses, the discussion of the role of the long-range electric polarization is somewhat subtle. Although there are accepted models for calculating the polarization around point defects in insulators and for its contribution to the defect energy,[28,29] we find here that if the inner region is large enough to fully screen the defect, reducing its nominal charge n to n/ϵ, where ϵ is the dielectric constant of the crystal, then the polarization in the outer region, though finite, makes no contribution to the energy of the system.

Having finished the survey of what the paper will contain, we describe the organization of the paper itself: In Sec. II we obtain a general expression for the energy functional which is compatible with our use of shifted eigenvalues. In Sec. III we discuss the truncation problem. We explore how truncating the Green's function leads to *loss* of screening at distances of the order of $2R_c$ from the defect. This happens even though the screening is fully developed at a distance of order R_c. The loss of screening between $R_c < r < 2R_c$ has no adverse effect on the calculation because charge in this region is not going to be included in the construction of the defect potential. In Sec. IV we apply the ideas developed in Sec. III to an evaluation of the difference of energies described in Sec. II. In Sec. V we describe the correction term which arises when "smeared" ions, which are convenient numerically, replace the point ions, which are a better physical description. In Sec. VI we recast the energy expression into the familiar and convenient eigenvalue form. Section VII is devoted to the comparison of the two basic energy formalisms, eigenvalue and kinetic energy, when the truncated Green's function is used for both, and that comparison serves as the conclusion of the paper.

II. TOTAL-ENERGY EXPRESSION, INCLUDING EIGENVALUE SHIFTS

Use of local-density-functional theory for describing exchange and correlation interactions between the electrons, use of nonlocal pseudopotentials to describe the interactions between the electrons and pseudoions, and use of point-charge repulsion between the pseudoions themselves gives a total-energy expression of the form

$$\begin{aligned} E[\rho_e,\{R_A\}]=&\sum_i n_i\int\psi_i^*(r)(-\tfrac{1}{2}\nabla^2)\psi_i(r)dr+\tfrac{1}{2}\int\int\frac{dr\,dr'}{|r-r'|}[\rho_e(r')+\rho_i(r')][\rho_e(r)+\rho_i(r)]\\ &+\sum_i n_i\int\int\psi_i^*(r)U_A^{\mathrm{ps}}(r,r')\psi_i(r')dr\,dr'+E_{\mathrm{xc}}+\tfrac{1}{2}\left[\sum_A\sum_{B(\neq A)}\frac{Q_AQ_B}{|R_A-R_B|}-\int\int\frac{dr\,dr'}{|r-r'|}\rho_i(r')\rho_i(r)\right] \end{aligned} \tag{2.1a}$$

where

$$E_{\mathrm{xc}}\equiv\int\rho_e(r)\epsilon_{\mathrm{xc}}(\rho_e(r))dr\ . \tag{2.1b}$$

The electronic density $\rho_e(r)$ is defined in terms of independent-particle occupation numbers n_i and wave functions $\psi_i(r)$:

$$\rho_e(r)=\sum_i n_i\psi_i(r)\psi_i^*(r)\ . \tag{2.2}$$

The ionic charge $\rho_i(r)$ arises as a result of the way we break up the electron-ion interaction into a nonlocal short-range part $U_A^{\mathrm{ps}}(r,r')$ referring to ion A at position R_A, and a long-range but local potential whose source is a smeared charge distribution with total charge Q_A. This latter has spherical symmetry and a limited Gaussian spread about location R_A.

The last term in (2.1a), enclosed in large square brackets, corrects for the difference between the smeared ion-ion repulsion included as a convenience in the Coulomb-energy term, and the point-ion interaction we actually use. That correction takes the form

$$\tfrac{1}{2}[\cdots]=\sum_A\sum_{B\neq A}C_{AB}-\sum_A J_A\ . \tag{2.3a}$$

The term C_{AB} expresses the fact that smeared ions A and B do not repel each other as strongly when they start to overlap as do point ions A and B. It vanishes rapidly with increasing R_{AB}. The term J_A, which is independent of atomic position, is the interaction of the smeared ion A with itself. Expressing the smeared ionic charge as a superposition of Gaussian functions is useful in that it makes possible an analytic evaluation of C_{AB}, of J_A, of $\rho_i(r)$, and of $\rho_i(q)$, the Fourier transform of $\rho_i(r)$. That evaluation appears in Sec. V. If C_{AB} is appreciable, then the cores are starting to overlap and the frozen-core approximation, on which the pseudopotentials are based, is starting to break down.

The density $\rho_e(r)$ to be used in (2.1) is the one which minimizes the functional. Carrying out this minimization subject to the constraint of wave-function normalization leads to a Schrödinger-type equation,[15]

$$[-\tfrac{1}{2}\nabla^2+\Phi(r)+U^{\mathrm{ps}}(r)+V_{\mathrm{xc}}(r)]\psi_i(r)=\epsilon_i\psi_i(r)\ , \tag{2.3b}$$

where

$$\Phi(r)\equiv\int\frac{dr'}{|r-r'|}[\rho_i(r')+\rho_e(r')]\ , \tag{2.3c}$$

$$U^{\mathrm{ps}}(r)\psi_i(r)\equiv\sum_A\int U_A^{\mathrm{ps}}(r,r')\psi_i(r')dr'\ , \tag{2.3d}$$

and

$$V_{\mathrm{xc}}(r)\equiv\frac{d}{d\rho}[\rho\epsilon_{\mathrm{xc}}(\rho)]_{\rho=\rho_e(r)}\ . \tag{2.3e}$$

Equations (2.3) and (2.2) must be solved self-consistently using a given $\epsilon_{\mathrm{xc}}(\rho)$, $U_A^{\mathrm{ps}}(r,r')$, and $\rho_i(r)$.

The defect energy we wish to calculate, ΔE, contains the difference between two energy expressions of the form (2.1), one for the defect crystal with its ions located at $\{R_A\}$ giving rise to an electronic charge density $\rho_e(r)$, and the other for the perfect crystal $[\{R_A^0\},\rho_e^0(r)]$. In order that the reference level of the electrostatic potential (say its average value $\langle\Phi\rangle$ far from the defect) should not influence the defect energy, the initial and final situations must have the same total charge. Since the perfect initial crystal is most naturally taken to be neutral, the defect crystal should also be taken as neutral. A defect with nominal charge n (n more electrons than protonic charges in the system; n an integer, positive, negative, or zero) must then be thought of as being compensated by some distant reservoir states which are now lacking n electrons. If we denote the change in reservoir energy as μ per added electron, then the change in reservoir energy when the defect has nominal charge n will be $-\mu n$. This change is also a component of the defect energy. Thus, the full definition of the defect energy is

$$\Delta E=E_{\mathrm{tot}}-E_{\mathrm{tot}}^0-\mu n\ , \tag{2.4}$$

where

$$E_{\mathrm{tot}}^0\equiv E[\rho_e^0,\{R_A^0\}]\ , \tag{2.5a}$$

$$E_{\mathrm{tot}}\equiv E[\rho_e,\{R_A\}]\ . \tag{2.5b}$$

In the initial situation, each unit cell of the perfect crystal is identical. Any numerical scheme for performing self-consistent band-structure calculations may be used to solve the Schrödinger equations. This gives wave functions $\psi_i^0(r)$ in the form of Bloch waves $\psi_n(k,r)$, eigenvalues $\epsilon_i^0=\epsilon_n(k)$, and an unperturbed electronic charge density,

$$\rho_e^0(r)=\sum_i n_i^0\psi_i^0(r)[\psi_i^0(r)]^* \ . \qquad (2.6)$$

When the defect is present, the Schrödinger equation differs from that for the perfect crystal. The difference in Schrödinger operators is the defect potential U. The perturbed electronic charge density $\rho_e(r)$ can be expressed as an integral over the defect-crystal Green's function $G_E(r,r')$,

$$\rho_e(r)=-\frac{1}{\pi}\int_{-\infty}^{E_F} dE\,\mathrm{Im}\,G_{E+i\eta}(r,r) \quad (\eta\rightarrow 0+) \ , \qquad (2.7)$$

or as a related contour integral in the plane of complex E.[30] The Green's function G_E is obtained from the defect potential U and the perfect-crystal Green's function G_E^0,

$$G_E^0(r,r')=\sum_n\int d^3k\frac{\psi_n(k,r)\psi_n^*(k,r')}{E-\epsilon_n(k)} \qquad (2.8)$$

by solving Dyson's equation,[27]

$$G_E(r,r')=G_E^0(r,r')+\int\int G_E^0(r,r_1)U(r_1,r_2) \times G_E(r_2,r')dr_1dr_2 \ . \qquad (2.9)$$

A substantial amount of literature already exists relating to various aspects of the calculational procedure implied by these equations.

The energy-gap problem has been mentioned. We have chosen to deal with this by making an *ad hoc* shift of all the conduction-band eigenvalues.[25,26] After the shift, the difference $\epsilon_c-\epsilon_v$ is equal to the experimental band gap E_g. [In actual use, we have shifted every conduction-band state upward by the same amount Δ. It would have been equally possible to shift $\epsilon_n(k)$ by $\Delta_n(k)$ if this had been desired.] We then evaluate the perfect-crystal Green's function G_E^0 Eq. (2.8), using the shifted eigenvalues, and with no alteration to the defect potential U we carry out the iterative solution to self-consistency. This procedure cannot be justified at any fundamental level, but, given the seriousness of the energy-gap problem, it is probably better than doing nothing. Therefore, in the remainder of this section, we motivate it as best we can and examine what effect it has on the evaluation of the defect energy (2.4) and the bound-state eigenvalue.

We start by recognizing that the problem arises from the use of a local-density-approximation (LDA) potential in the density-functional treatment. Rather than trying to find improved potentials (i.e., local operators) within local-density theory, we conjecture that an improved theory might have a nonlocal, but still *energy-independent,* potential $\Sigma(r,r')$ similar to the self-energy operator of the quasiparticle equation when its energy dependence is small over the energy range of interest. In that case, the E_{xc} term in the total-energy functional (2.1) would have the form

$$E_{\mathrm{xc}}=\sum_i n_i\int\int\psi_i^*(r)\Sigma(r,r')\psi_i(r')dr\,dr' \ . \qquad (2.10)$$

The nonlocal kernel Σ is in principle a functional of the electron density ρ_e.

The impressive successes of *local*-density-functional theory[31,32] suggest that Σ has a local component equal to the usual exchange-correlation energy. It would then be useful to write the kernel as

$$\Sigma(\rho_e;r,r')=\delta(r-r')E_{\mathrm{xc}}(\rho_e(r))+S(\rho_e;r,r') \ . \qquad (2.11)$$

This conjecture, used in (2.10), yields

$$E_{\mathrm{xc}}=\int\rho_e(r)\epsilon_{\mathrm{xc}}(\rho_e(r))dr + \sum_i n_i\int\int\psi_i^*(r)S(\rho_e;r,r')\psi_i(r')dr\,dr' \ . \qquad (2.12)$$

There is no theory to dictate a form for the "scissors operator" S. However, if we want to build an *ad hoc* correction term which does not affect the valence-band density or eigenvalues for the perfect infinite crystal, then the Bloch-wave expansion of S,

$$S(\rho_e,r,r')=\sum_n\sum_{n'}\int\int dk\,dk'\psi_n(k,r) \times S[\rho_e]_{nn}^{kk'}\psi_{n'}^*(k',r') \qquad (2.13)$$

must have coefficients which vanish if n or n' is in the valence band and ρ_e is the perfect-crystal density ρ_e^0.

The Schrödinger equation which results from minimizing the energy acquires an extra term which results because of the nonlocal kernel S. It is

$$\frac{1}{n_i}\frac{\delta}{\delta\psi_i(r)^*}\sum_j n_j\int\int\psi_j^*(r)S(\rho_e;r,r')\psi_j(r')dr\,dr'=\int S(\rho_e;r,r')\psi_i(r')dr' + \left[\sum n_j\int\psi_j^*(r_1)\left[\frac{\delta}{\delta\rho_e(r)}S(\rho_e;r_1,r_2)\right]\psi_j(r_2)dr_1dr_2\right]\psi_i(r) \ . \qquad (2.14)$$

S depends on ρ_e, but, lacking a theory of the dependence, we are forced to ignore that dependence and to proceed as though $S(\rho_e)=S(\rho_e^0)$. That means that there is no $S(\rho_e)-S(\rho_e^0)$ contribution to the defect potential U. It also means that the second term in (2.14) vanishes. Only the first goes into the Schrödinger equation, where, because of the restriction that S have no valence-band components, it can, at most, change the conduction-band Bloch states and eigenvalues. In the absence of any theory for how the Bloch states change, we arrange to leave them undisturbed by choosing the expansion coefficients for S to be diagonal, i.e.,

$$S[\rho_e]=S[\rho_e^0]=\Delta_n(k)\delta_{nn'}\delta(k-k') \ , \qquad (2.15)$$

where $\Delta_n(k)$ is zero for n in the valence band. The effect of including (2.12), (2.14), and (2.15) in the Schrödinger equation *for the perfect* crystal is to shift the eigenvalues from $\epsilon_n(k)$ to $\epsilon_n(k)+\Delta_n(k)$. The effect of including (2.13) and (2.15) in the energy functional (2.1), in general, is to bring in an extra term which is added to the energy,

$$E_{\rm xc}^{\rm NL}=\sum_i n_i \sum_n \int dk\, \Delta_n(k) \left| \int \psi_n^*(k,r)\psi_i(r)dr \right|^2 . \tag{2.16}$$

This raises the energy of the system according to how much of its density is contributed by perfect-crystal conduction-band wave functions. The term (2.16) does not appear overtly in the energy when the eigenvalue form of total energy is used. Nonetheless, its value is included automatically once the perfect-crystal eigenvalues have been shifted. The kernel S will appear overtly in the Schrödinger equation for both the perfect and the defect crystal. Its effect on ϵ_B, the eigenvalue for a bound state whose wave function is $\psi_B(r)$, is, in lowest order, given by the expectation value

$$\begin{aligned}\epsilon'_B &\approx \epsilon_B + \int \int \psi_B^*(r)S(r,r')\psi_B(r')dr\,dr' \\ &= \epsilon_B + \sum_n \int d^3k\, \Delta_n(k) \left| \int \psi_n^*(k,r)\psi_B(r)dr \right|^2 .\end{aligned} \tag{2.17}$$

The shift in ϵ_B depends on the fraction of ψ_B derived from conduction-band states.

III. TRUNCATION PROBLEM

In this section we discuss the consequences of using a truncated Green's function with a defect potential which, in principle, has infinite range.

A. Description of Green's-function truncation

The perfect crystal has a Green's function (2.8) which we denote as $G_E^T(r,r')$, the superscript T denoting "true." Another notation is

$$G_E^T(r,r') \equiv \langle r \,|\, (E-H_0)^{-1} \,|\, r' \rangle . \tag{3.1}$$

In almost all of the recent work this Green's function has been represented by an expansion in terms of a limited set of localized orbitals,[11,12,23,30,33]

$$G_E^0(r,r')=\sum_{a,b} G_{ab}^0(E)\phi_a(r)\phi_b^*(r') . \tag{3.2}$$

The localized orbital $\phi_a(r)$ is centered at location R_A. The set of locations $\{R_A\}$ need not be the same as the set of atomic positions of the perfect or defect crystal. However, the types of orbitals and their locations are chosen so that $G_E^0(r,r')$ is an adequate approximation to $G_E^T(r,r')$ for r and r' both near the defect. The set of locations $\{R_A\}$ does not extend far beyond the defect region. Therefore G_E^0, unlike G_E^T, will vanish if either r or r' becomes too large, even if r and r' are close to each other. Because of the construction, the two Green's functions are related to each other by

$$G_E^0(r,r')=A_E(r)G_E^T(r,r')A_E^*(r') , \tag{3.3}$$

where $A_E(r)$ is some sort of cutoff function, equal to 1 for r close to the defect and equal to zero for r far from it. The details of how $A(r)$ goes from 1 to 0 depend on the energy, but the energy dependence is not important, and, for simplicity, we can ignore it. The r dependence, however, is important. It may be useful to regard $A(r)$ as being approximately equal to $\theta(R_c-r)$, where θ is the unit step function and R_c lies at the edge of region 1.

B. Effect of Green's function truncation on the defect potential

The Green's function for the defect crystal is supposed to satisfy the Dyson equation (2.9), which, in symbolic form, is

$$G_E=G_E^T+G_E^TUG_E . \tag{3.4a}$$

Having solved (3.4a), the *change* in electronic charge density is given by an energy integration of

$$\Delta^T G_E \equiv G_E - G_E^T = G_E^TUG_E . \tag{3.4b}$$

However, we can neither solve (3.4a) nor evaluate (3.4b) because we do not have G_E^T available. We do have G_E^0 available and we use it instead. Our experience has been that, without further adjustment, use of G_E^0 instead of G_E^T in (3.4) produces a change in electronic charge density $\Delta\rho_e(r)$ whose spatial integral over the neighborhood of the defect (say, $r<R_c$) leads to a charge neutral defect. This is true even if the occupation numbers n_i are such that the defect should have been nominally charged. When the nominal charge on the defect is n, the net charge in the neighborhood of the defect should be n/ϵ, where ϵ is the dielectric constant of the crystal. The charge we find empirically is always a small fraction of n/ϵ.

The simplest adjustment which restores the proper charge n/ϵ to the defect region while doing the least to alter the distribution of that charge is to add a constant potential $\hat{\Phi}$ to U, choosing $\hat{\Phi}$ by trial and error if necessary. Thus, instead of (3.4), the equations we actually solve and use are

$$G_E=G_E^0+G_E^0(U+\hat{\Phi})G_E , \tag{3.5a}$$

$$\Delta^0 G_E \equiv G_E - G_E^0 = G_E^0(U+\hat{\Phi})G_E . \tag{3.5b}$$

Given the form (3.3) and Eq. (3.5a) for G_E, it is evident that $G_E(r,r')$ can be written as

$$G_E(r,r')=A(r)\hat{G}_E(r,r')A^*(r') \tag{3.6a}$$

because substitution of (3.3) and (3.6) into (3.5a) leads to

$$A\hat{G}_EA^*=AG_E^TA^*+AG_E^TA^*(U+\hat{\Phi})A\hat{G}_EA^* .$$

At all values of r and r' for which $A(r)\neq 0$ and $A^*(r')\neq 0$, dividing the above equation by $A(r)A^*(r')$ is legitimate. We do so, and the resulting equation for $\hat{G}_E$ is

$$\hat{G}_E=G_E^T+G_E^TU_m\hat{G}_E , \tag{3.6b}$$

where

$$U_m \equiv A^*(U+\hat{\Phi})A . \tag{3.6c}$$

This shows that the defect Green's function G_E is a truncated approximation to that defect Green's function $\hat{G}_E$ which, by (3.6b) and (3.4a), would have been exact if the defect potential had been U_m.

The defect potential U_m (henceforth, the truncated or model potential) differs from the real defect potential U in two ways. First, its long-range tail, if present, has been cut off by $A^*(r_1)A(r_2)$. Second, within the region where it is not cut off, the model potential appears to be shifted by $\hat{\Phi}$ from the true one. This apparent shift, however, is vastly reduced by the self-consistency requirement. That is, the average value of U_m calculated without $\hat{\Phi}$ present and iterated to self-consistency is rather close to the average of U_m calculated self-consistently *with* the correct $\hat{\Phi}$ present.

C. Effect of truncation on the charge disturbance

In analogy with (3.4b) and (3.5b), it is conceptually useful to consider the change in the *un*truncated Green's function G_E associated with the model potential U_m,

$$\Delta^m G_E \equiv \hat{G}_E - G_E^T = G_E^T U_m \hat{G}_E \ . \tag{3.7}$$

Correspondingly, there are three charge disturbances which are of interest, namely,

$$\Delta^T \rho_e(r) \equiv -\frac{1}{\pi}\int_{-\infty}^{E_F} dE\ \mathrm{Im}\Delta^T G_E(r,r)\ , \tag{3.8a}$$

$$\Delta^m \rho_e(r) \equiv -\frac{1}{\pi}\int_{-\infty}^{E_F} dE\ \mathrm{Im}\Delta^m G_E(r,r)\ , \tag{3.8b}$$

$$\Delta^0 \rho_e(r) \equiv -\frac{1}{\pi}\int_{-\infty}^{E_F} dE\ \mathrm{Im}\,\Delta^0 G_E(r,r)\ . \tag{3.8c}$$

The charge disturbance $\Delta^T\rho_e(r)$ is that which the true crystal would support: There is no truncation involved in its definition. Even though we cannot calculate it, we can make some general statements about it. (These statements are simpler to make if we temporarily ignore the presence of the reservoir states and regard the defect crystal as charged.) Suppose that we have a charged defect with nominal charge n. There is an integer change in the ionic charge caused by the creation of the defect,

$$\Delta N_i \equiv \int_{\text{crystal}} \Delta\rho_i(r)dr\ . \tag{3.9a}$$

There is also an integer change in the electronic charge for the same reason,

$$\Delta^T N_e \equiv \int_{\text{crystal}} \Delta^T\rho_e(r)dr\ . \tag{3.9b}$$

The nominal charge n associated with the defect is

$$n \equiv \Delta^T N_e + \Delta N_i\ . \tag{3.9c}$$

The polarizable nature of the semiconductor is such that the total charge change near the defect, say within $r<R_c$ will be n/ϵ,

$$\int_{r<R_c}[\Delta\rho_i(r)+\Delta^T\rho_e(r)]dr$$
$$=\Delta N_i + \int_{r<R_c}\Delta^T\rho_e(r)dr = n/\epsilon\ . \tag{3.10}$$

(In writing 3.10, we have assumed that R_c fully encloses the dielectric screening region 1 and that all of the change in the ionic charge is confined to region 1.) Combining (3.9) and (3.10) gives

$$\int_{r>R_c \text{ in crystal}} \Delta^T\rho_e(r)dr = n\left[1-\frac{1}{\epsilon}\right] . \tag{3.11}$$

Since there is no net accumulation in the bulk of a uniform dielectric crystal, i.e., since the $n/\epsilon r$ defect potential implied by (3.10) persists out the surface of the crystal, the net charge accumulation implied by (3.11) resides at the outer surface of the crystal. There is certainly no significant net accumulation of charge at, or slightly beyond, $r=R_c$.

Now let us consider the second type of electronic charge disturbance, $\Delta^m\rho_e(r)$, given by (3.8b). This is the charge disturbance in a crystal governed by the exact Green's function G_E^T, but acted upon by the model, truncated defect potential U_m [Eq. (3.6b)]. The value of $\hat{\Phi}$ for a defect of nominal charge n has been adjusted so that the analog to (3.10) is still true. The occupation numbers n_i have been chosen so that (3.9) is still true. Thus, analogous to (3.11), it is still true that

$$\int_{r>R_c \text{ in crystal}} \Delta^m\rho_e(r)dr = n\left[1-\frac{1}{\epsilon}\right] . \tag{3.12}$$

Evan Kane[34] has pointed out to us that the charge being integrated in (3.12) is really located rather close to R_c, say in the region $R_c<|r|<2R_c$. His argument can be most simply expressed as follows: All experience with calculations of the charge disturbance around neutral defects with a strong defect potential near $r=0$ shows that the charge disturbance vanishes by $r=R_c$. Therefore a potential which is zero beyond $r=R_c$, as is the truncated potential, will produce a charge disturbance which cannot persist beyond $r=2R_c$.

This argument may need to be strengthened because one might object that the limited range of charge disturbance surrounding a strong but neutral defect potential comes about because of self-consistency and the Coulomb potential. These two effects are absent when the potential is truncated. Nonetheless, it is very likely true that the perfect-crystal Green's function $G_E^T(r,r')$ effectively does have a limited range, vanishing when r and r' are too far apart. Certainly, Callaway's[35] work shows that if E is in the gap, then G_E^T decays exponentially with $|r-r'|$, the rate of decay being governed by the separation between E and the nearest allowed bands. His work also shows that if E lies in a band, then G_E^T behaves like

$$G_E^T(r,r') \approx |r-r'|^{-1}\exp(iq\,|r-r')\ ,$$

where q depends on E. An energy integration, such as that in (3.8) for evaluating the charge, will cause the imaginary exponential to phase-oscillate to zero at large $|r-r'|$, thereby introducing an effective cutoff. An effective range to $G_E^T(r,r')$ is all that is needed to establish that the charge disturbance vanishes at some finite distance beyond the range of a truncated potential. Therefore, Kane's conclusion is basically correct. Defining the distance where the charge disturbance vanishes as $2\tilde{R}_c$, i.e., defining $2\tilde{R}_c \equiv$ potential range + Green's-function range, we have $\Delta^m\rho_e(r)\equiv 0$ for $r>2\tilde{R}_c$. Equations (3.9) and (3.12) can then be written with finite outer limits as

$$\int_{R_c<r<2\tilde{R}_c} \Delta^m \rho_e(r)dr = n\left[1-\frac{1}{\epsilon}\right], \tag{3.13}$$

and

$$\int_{r<2\tilde{R}_c} [\Delta\rho_i(r)+\Delta^m\rho_e(r)]dr = n \ . \tag{3.14}$$

This is a startling result. It means that the screening disappears at $2\tilde{R}_c$, not at the surface of crystal if a truncated potential is used. It implies that $\Delta^m\rho_e(r)$ and $\Delta^T\rho_e(r)$ differ significantly in the region $R_c<r<2\tilde{R}_c$, and therefore that it is wrong, in principle and in practice, to use the charge density arising from a truncated potential as an approximation to the actual charge in the region beyond the cutoff.

The third type of charge disturbance is $\Delta^0\rho_e(r)$, Eq. (3.8c). It is clear from Eqs. (3.5)—(3.8) that

$$\Delta^0\rho_e(r) = A(r)\Delta^m\rho_e(r)A^*(r) \ . \tag{3.15}$$

$\Delta^0\rho_e(r)$ will then be very much like the true distribution $\Delta^T\rho_e(r)$. They should both agree for $r<R_c$. For $r>R_c$ they both vanish on the average—$\Delta^0\rho_e(r)$ identically, $\Delta^T\rho_e(r)$ only on the average. They should both give rise to the same defect potential $U(r)$ except for the microscopic-scale, finite-amplitude fluctuations in the region $r>R_c$. The effect of these fluctuations will vanish on the average because of the following argument.

D. Dielectric polarization at large distances

The true charge disturbance $\Delta^T\rho_e(r)$ in the response region $r>R_c$ is a polarization charge, varying on a length scale of the bonds, having an overall amplitude which drops like $1/r^2$, existing because the electric field of the defect pushes charge along bonds or across them,[36] but piling up no macroscopic charge except perhaps at the outer surface of the crystal. It might be tempting to regard it as being the divergence of a macroscopic polarization vector $\vec{P}$, averaging over some macroscopic (but still small region) as is conventionally done,[37]

$$\bar{\rho}(r) \equiv \int f(|r-r'|)\Delta^T\rho_e(r')dr' = -\vec{\nabla}\cdot\vec{P}(r)$$

$$\text{with } f\geq 0, \quad \int f(r)dr = 1 \ . \tag{3.16}$$

Macroscopic dielectric theory provides an expression for the energy of a polarizable dielectric in an electric field,

$$\delta U = -\tfrac{1}{2}\int \vec{P}(r)\cdot\vec{E}(r)dr \ , \tag{3.17}$$

and, considering the ease of evaluating (3.17), it is tempting to use this expression, carried over the outer region, for its contribution to the energy.

It is wrong to do so, because this formalism implies that $\vec{P}(\vec{r})\equiv\vec{0}$ outside the region of integration. The implied discontinuity in $\vec{P}$ then gives rise to a surface charge density equal to the outgoing normal component,[38]

$$\vec{\sigma} = \vec{P}_n \ . \tag{3.18}$$

In the Appendix we present a simple evaluation of (3.17) and (3.18) to show that the correct amount of charge [cf. (3.11)] resides on the outer surface of the crystal, but that the same amount, but of opposite sign, spuriously resides on the inner surface at $r=R_c$, rendering the system to which (3.17) refers neutral. This is in conflict with (3.11). For that reason, it is better to use the alternative formulation of (3.17), namely

$$\delta U = \tfrac{1}{2}\int \rho_{\rm def}(r)\phi_p(r)dr \ , \tag{3.19}$$

where $\phi_p(r)$ is the potential which arises from the polarization charge,

$$\phi_p(r) = \int_{r'>R_c} \frac{dr'}{|r-r'|}[-\vec{\nabla}\cdot\vec{P}(r')] \ , \tag{3.20}$$

and $\rho_{\rm def}(r)$ is the charge associated with the defect for $r<R_c$. Doing so allows one to ignore the spurious surface charge at $r=R_c$. The remaining contribution to the energy δU goes to zero as the crystal becomes infinite (see Appendix).

Correctly applied then, the macroscopic dielectric response theory yields no contribution to the energy, except within that region near the defect where charge has actually been moved to do the screening. If the microscopic quantum-mechanical calculation is formulated so that the energy is expressed as the integral of an energy density composed of a sum of partial energy densities (e.g., kinetic, exchange, etc.) then in the defect situation macroscopic response theory justifies confining that integral to the inner active region. For the charged defect, the long-range polarization means that there is finite $\Delta^T\rho_e(r)$, and demonstrably, nonzero partial energy densities, in the outer region. The partial densities must, to conform to the macroscopic theory, sum to zero on the average.

IV. EVALUATION OF THE DEFECT ENERGY

All the major ideas needed for the evaluation of the defect energy are now in place. We evaluate energies with no relaxation of the atomic positions in region 2. As a result, the ionic charge disturbance

$$\Delta\rho_i(r) \equiv \rho_i(r) - \rho_i^0(r) \tag{4.1a}$$

is identically zero for $r>r_c$. The electronic charge disturbance

$$\Delta\rho_e(r) \equiv \rho_e(r) - \rho_e^0(r) \tag{4.1b}$$

is also identically zero in region 2, because we evaluate $\Delta\rho_e(r)$ using the truncated Green's function G_E^0. Thus $\Delta\rho_e(r)$ is really $\Delta^0\rho_e(r)$ [cf. (3.5) and (3.15)].

In this section we shall merely assemble the terms needed for evaluation of the defect energy (2.4). For dealing with the nonlocal operators in (2.1) and (2.12), it is useful to make use of the one-particle density matrices for the perturbed and unperturbed systems,

$$n(r,r') \equiv \sum_i n_i \psi_i(r)\psi_i^*(r') \ , \tag{4.2a}$$

$$n^0(r,r') \equiv \sum_i n_i^0 \psi_i^0(r)[\psi_i^0(r')]^* \ , \tag{4.2b}$$

$$\Delta n(r,r') \equiv n(r,r') - n^0(r,r') \ . \tag{4.2c}$$

Once the Green's function $G_E^0(r,r')$ is expressed in terms of localized orbitals, as in (3.2), the change in density matrix quite naturally takes the form of an expansion based on the same orbitals,

$$\Delta n(r,r')=\sum_{a,b}\rho_{ab}\phi_a(r)\phi_b^*(r') , \qquad (4.3)$$

and so, this is also a truncated expansion. The change in electronic density is then evaluated as

$$\Delta\rho_e(r)=\Delta n(r,r) . \qquad (4.4)$$

We now turn to an enumeration of the terms in the defect energy, ΔE [Eq. (2.4)], using the basic energy functional (2.1), augmented by the scissors operator S [Eqs. (2.15) and (2.12)], which represents exactly the effect of the eigenvalue shifts.

The change in the kinetic-energy term is, using (4.2),

$$\Delta T=\int dr[-\tfrac{1}{2}\nabla_r^2\Delta n(r,r')]_{r'=r} . \qquad (4.5)$$

We should not include the integral over the entire crystal, only over region 1. Therefore, use of the truncated density matrix of the form (4.3) in (4.5) is helpful for it does introduce a cutoff just where it is needed. Inserting (4.3) into (4.5) gives

$$\Delta T=\sum_{a,b}\Delta\rho_{ab}T_{ba} , \qquad (4.6a)$$

where

$$T_{ba}\equiv\int\phi_b^*(r)(-\tfrac{1}{2}\nabla^2)\phi_a(r)dr . \qquad (4.6b)$$

Next, consider the contribution of the scissors operator S. It is

$$\Delta S=\int\int S(\rho_e^0;r,r')\Delta n(r',r)dr\,dr' . \qquad (4.7)$$

Again, we limit the integration over r and r' to region 1, and so again use of (4.3) is desired.

We have the change in the pseudopotential, expressible as

$$\begin{aligned}\Delta U_{\mathrm{ps}}&=\int\int[U_{\mathrm{ps}}(r,r')n(r',r)-U_{\mathrm{ps}}^0(r,r')n^0(r',r)]dr\,dr'\\&=\int\int[\Delta U_{\mathrm{ps}}(r,r')n^0(r',r)+U_{\mathrm{ps}}(r,r')\Delta n(r'r)]dr\,dr' ,\end{aligned} \qquad (4.8a)$$

where

$$U_{\mathrm{ps}}(r,r')=\sum_A U_A^{\mathrm{ps}}(r,r') , \qquad (4.8b)$$

$$U_{\mathrm{ps}}^0(r,r')=\sum_{A^0} U_{A^0}^{\mathrm{ps}}(r,r') , \qquad (4.8c)$$

$$\Delta U_{\mathrm{ps}}(r,r')\equiv U_{\mathrm{ps}}(r,r')-U_{\mathrm{ps}}^0(r,r') \qquad (4.8d)$$

refer to the short-range parts of the pseudopotential in the defect crystal (pseudoions located at $\{R_A\}$) and the perfect crystal (pseudoions located at $\{R_A^0\}$). Because of the short range and the fact that only atoms in region 1 are involved in creating the defect, $\Delta U_{\mathrm{ps}}(r,r')$ will automatically restrict the integrations to region 1. If we then have an $n^0(r'r)$ expressed in the localized-orbital form (and hence truncated) as

$$n^0(r,r')=\sum_{a,b}\rho_{ab}^0\phi_a(r)\phi_b^*(r') , \qquad (4.9)$$

this can be used in evaluating the first term in (4.8a) with no approximation caused by truncation. For the second term in (4.8a) we count on the use of the form (4.3) to provide the desired cutoff. This term is the most troublesome of all numerically because it involves the evaluation of the matrix elements of many nonlocal potentials in the local-orbital basis.

The evaluation of the exchange and correlation part of (2.1) is straightforward, as is the change in the sums involving C_{AB} and J_A [Eq. (2.5)]. In particular, the change in the J_A term is

$$\Delta\left[-\sum_A J_A\right]=\sum_{A\text{ removed}}J_A-\sum_{B\text{ added}}J_B , \qquad (4.10)$$

where "A removed" denotes a count over all atoms which are removed from the crystal in forming the defect and "B added" denotes a count over all atoms which are added to the crystal in forming the defect. The individual terms in C_{AB} vanish unless atoms A and B are close enough for their pseudocores to overlap. In practical terms, this restriction rules out all AB pairs except first-nearest neighbors. The pairs surviving in $\Delta\sum_{A,B}C_{AB}$ are thus only those whose bond $A-B$ is created, broken, stretched, or compressed in the process of creating the defect.

The change in the electrostatic Coulomb term is

$$\begin{aligned}\Delta E_{\mathrm{es}}&=\tfrac{1}{2}\int\int\frac{dr\,dr'}{|r-r'|}[\rho(r')\rho(r)-\rho^0(r')\rho^0(r)]\\&=\int\int\frac{dr\,dr'}{|r-r'|}[\rho^0(r')\Delta\rho(r)+\tfrac{1}{2}\Delta\rho(r')\Delta\rho(r)] ,\end{aligned} \qquad (4.11a)$$

where

$$\rho(r)\equiv\rho_e(r)+\rho_i(r) , \qquad (4.11b)$$

$$\rho^0(r)\equiv\rho_e^0(r)+\rho_i^0(r) , \qquad (4.11c)$$

and

$$\Delta\rho(r)\equiv\rho(r)-\rho^0(r) . \qquad (4.11d)$$

In this case, restricting the contribution of ΔE_{es} to region 1 applies to the terms involving $\Delta\rho(r)$ only. The $\rho^0(r')$ term has to be carried over the entire crystal. In analogy to (2.3b), we write

$$\int\frac{dr'}{|r-r'|}\rho^0(r')=\Phi^0(r) , \qquad (4.12)$$

which is the periodic electrostatic potential of the perfect crystal. It has an expansion in reciprocal-lattice vectors,

$$\Phi^0(r)=\sum_{\vec{G}}\Phi^0(G)e^{iG\cdot r} , \qquad (4.13a)$$

where

$$\Phi^0(G)=\frac{4\pi}{G^2}\rho^0(G),\quad G\neq0 \qquad (4.13b)$$

$$\rho^0(G)=\frac{1}{\Omega_c}\int_{\Omega_c}\rho^0(r)e^{-iG\cdot r}dr , \qquad (4.13c)$$

and

$$\Phi^0(G)=\langle\Phi^0\rangle,\quad G=0\ . \tag{4.13d}$$

Ω_c is the volume of the unit cell in the crystal. $\langle\Phi^0\rangle$, the average value of the potential in the infinite crystal, is arbitrary. The value of the first integral in (4.11a) is thus

$$\int\int\frac{dr\,dr'}{|r-r'|}\rho^0(r')\Delta\rho(r)=\langle\Phi^0\rangle\int\Delta\rho(r)dr+\sum_{G\neq 0}\frac{4\pi}{G^2}\rho^0(G)\Delta\rho(G)\ , \tag{4.14a}$$

where

$$\Delta\rho(G)=\int_{r<R_c}\Delta\rho(r)e^{iG\cdot r}dr,\quad G\neq 0\ . \tag{4.14b}$$

The last term to be included in the energy of the defect is that of the reservoir states, namely $-\mu n$.

We consider the question of what value to assign to the first term on the right-hand side (rhs) of (4.14a). This term depends on the reference level of the potential. The only other term which depends on this reference level is $-\mu n$, because μ, the energy needed to add an electron to the reservoir, depends additively on $\langle\Phi^0\rangle$, the average electrostatic potential in the reservoir region. We group the two terms together as

$$\langle\Phi^0\rangle\int\Delta\rho(r)dr-\mu n=\langle\Phi^0\rangle\left[\int\Delta\rho(r)dr-n\right]-\mu_0 n\ , \tag{4.15}$$

where

$$\mu_0\equiv\mu-\langle\Phi^0\rangle\ . \tag{4.16}$$

Since μ depends additively on $\langle\Phi^0\rangle$, μ_0 is independent of $\langle\Phi^0\rangle$. Its value can be assigned: Recall that it is proper to measure μ relative to the eigenvalue $\epsilon_n(k)$, because μ is the systems energy change upon adding one electron to the reservoir, while $\epsilon_n(k)$ is the energy change upon adding one electron to a particular band state. Subtracting $\langle\Phi^0\rangle$ from each of these, we conclude that $\epsilon_n(k)-\langle\Phi^0\rangle$ is the appropriate scale to measure μ_0. However, obtaining a numerical value for $\epsilon_n(k)$ in an actual calculation can be carried out only by assigning a numerical value to $\langle\Phi^0\rangle$, so $\epsilon_n(k)-\langle\Phi^0\rangle$ is always known and independent of $\langle\Phi^0\rangle$. Therefore, μ_0 can be assigned a numerical value, according to whether the reservoir is at the bottom of the conduction band, or top of the valence band, etc.

Returning to Eq. (4.15), this result, and therefore ΔE, is independent of $\langle\Phi^0\rangle$ only if

$$\int\Delta\rho(r)dr=n\ . \tag{4.17}$$

We therefore make the assignment (4.17). This means that, in spite of the $r<R_c$ limitation, we have put on $\Delta\rho(r)$, for this one usage, an integral over the entire crystal is needed. Combining (4.14)–(4.17) gives

$$\int\int\frac{dr\,dr'}{|r-r'|}\rho^0(r')\Delta\rho(r)-\mu n=\sum_{G\neq 0}\frac{4\pi}{G^2}\rho^0(G)\Delta\rho(G)-\mu_0 n\ , \tag{4.18}$$

and there is no $\langle\Phi^0\rangle$ dependence anywhere in ΔE.

V. POINT-ION VERSUS SMEARED-ION CORRECTION

We start with the general expression for the electron-ion interaction in which a nonlocal pseudopotential $U_A(r,r')$ describes the interaction of an electron with a single pseudoatom of type A located at R_A. It is convenient to express this interaction as being the sum of a truly short-range nonlocal pseudopotential $U_A^{\text{ps}}(r,r')$ which vanishes if $|\vec{r}-\vec{R}_A|$ or $|r'-R_A|$ exceeds (typically) a core radius, and a long-range local potential, the Coulomb potential of a smeared charge distribution with total charge Q_A centered at R_A. That charge distribution is conveniently described as a sum of Gaussians. We have

$$U_A(r,r')=\delta(r-r')Q_A\times\sum_j C_A^j\int\frac{dr''}{|r-r''|}g_A^j(|r''-R_A|)+U_A^{\text{ps}}(r,r')\ , \tag{5.1a}$$

where

$$\sum_j C_A^j=1 \tag{5.1b}$$

and

$$g_A^j(r)\equiv(\alpha_A^j/\pi)^3\exp(-\alpha_A^j r^2)\ . \tag{5.1c}$$

Then the electron-ion—interaction energy is

$$\begin{aligned}V_{\text{e-i}}&\equiv\sum_i n_i\int\int\psi_i^*(r)\sum_A U_A(r,r')\psi_i(r')dr\,dr'\\&=\int\int\frac{dr\,dr'}{|r-r'|}\rho_e(r')\rho_i(r)\\&\quad+\int\int U^{\text{ps}}(r,r')n(r',r)dr\,dr'\ ,\end{aligned} \tag{5.2a}$$

where

$$\rho_i(r)\equiv\sum_A Q_A\sum_j C_A^j g_A^j(|r-R_A|)\ , \tag{5.2b}$$

$$n(r',r)\equiv\sum_i n_i\psi_i(r')\psi_i^*(r)\ , \tag{5.2c}$$

$$\rho_e(r)=n(r,r)\ , \tag{5.2d}$$

and

$$U^{\text{ps}}(r,r')\equiv\sum_A U_A^{\text{ps}}(r,r')\ . \tag{5.2e}$$

The point-ion interaction between all of the distinct ion pairs can be described as the self-interaction of the ionic charge density $\rho_i(r)$ and two correction terms, C_{AB} and J_A. We have

$$\frac{1}{2}\sum_A\sum_{B\neq A}\frac{Q_AQ_B}{|R_A-R_B|}=\frac{1}{2}\int\int\frac{dr\,dr'}{|r-r'|}\rho_i(r')\rho_i(r)+\sum_A\sum_{B\neq A}C_{AB}-\sum_A J_A\ , \tag{5.3a}$$

where, using (5.1),

$$C_{AB} \equiv \tfrac{1}{2}\frac{Q_A Q_B}{R_{AB}} \sum_{j,k} C_A^j C_B^k \operatorname{erfc}[R_{AB}(\gamma_{AB}^{ik})^{1/2}] , \tag{5.3b}$$

$$J_A \equiv Q_A^2 \sum_{j,k} C_A^j C_A^k (\gamma_{AA}^{ik}/\pi)^{1/2} , \tag{5.3c}$$

$$R_{AB} \equiv | R_A - R_B | , \tag{5.3d}$$

$$\gamma_{AB}^{ik} \equiv \alpha_A^j \alpha_B^k / (\alpha_A^j + \alpha_B^k) . \tag{5.3e}$$

VI. FORMULA FOR THE DEFECT ENERGY IN THE EIGENVALUE FORMULATION

We now evaluate the defect energy (2.4) using the eigenvalue form of the energy functional. First, however, it is useful to assume that complete self-consistency has not yet been achieved. We start with a defect potential U_m^{in} which is input to the Schrödinger equation. The solution to that equation gives a $\Delta\rho_e$. This $\Delta\rho_e(r)$ is truncated at $r = R_c$ and the truncated form is used to give a model output potential U_m^{out}. Lack of self-consistency means that U_m^{out} and U_m^{in} still differ by an unacceptable amount and that quantities being studied still fluctuate from iteration to iteration. The first step is to use the eigenvalue sum to replace the one-particle terms, following Ihm *et al.*[13] We multiply the defect Schrödinger equation by $n_i\psi_i(r)$, integrate over r and sum over i to obtain

$$\int dr[-\tfrac{1}{2}\nabla_r^2 n(r,r')]_{r'=r} + \int [\Phi^0(r) + V_{\text{xc}}^0(r)]\rho_e(r)dr + \int\int [U_{\text{ps}}^0(r,r') + U_m^{\text{in}}(r,r')]n(r',r)dr\,dr' = \sum_i n_i\epsilon_i . \tag{6.1}$$

The model potential U_m here is that given in (3.6c) where

$$U \equiv [\Delta\Phi(r) + \Delta V_{\text{xc}}(r)]\delta(r-r') + \Delta U_{\text{ps}}(r,r') ,$$

and

$$\Delta\Phi(r) \equiv \Phi(r) - \Phi^0(r) = \int \frac{dr'}{|r-r'|}\Delta\rho(r'), \quad \Delta\rho(r) = \Delta\rho_e(r) + \Delta\rho_i(r) ,$$

$$\Delta V_{\text{xc}}(r) \equiv V_{\text{xc}}(r) - V_{\text{xc}}^0(r), \quad \Delta\rho_e(r) \equiv \rho_e(r) - \rho_e^0(r) .$$

Using this expression to replace the kinetic-energy term in the total-energy expression (2.1) gives

$$\begin{aligned} E[\rho_e,\{R_A\}] = {} & \sum_i n_i\epsilon_i + \int dr[\tfrac{1}{2}\Phi(r)\rho(r) - \Phi^0(r)\rho_e(r)] + \int [U_{\text{ps}}(r,r') - U_{\text{ps}}^0(r,r') - U_m^{\text{in}}(r,r')]n(r',r)dr\,dr' \\ & + \int [\epsilon_{\text{xc}}(\rho_e(r)) - V_{\text{xc}}^0(r)]\rho_e(r)dr + \sum_A \sum_{B \neq A} C_{AB} - \sum_A J_A . \end{aligned} \tag{6.2}$$

Rearranging slightly,

$$\int dr[\tfrac{1}{2}\Phi(r)\rho(r) - \Phi^0(r)\rho_e(r)] = \int \Delta\Phi(r)\rho_e(r)dr + \tfrac{1}{2}\int\int \frac{dr\,dr'}{|r-r'|}[\rho_i(r)\rho_i(r') - \rho_e(r')\rho_e(r)] , \tag{6.3a}$$

$$U_{\text{ps}}(r,r') - U_{\text{ps}}^0(r,r') = \Delta U_{\text{ps}}(r,r') . \tag{6.3b}$$

Using the definition of the defect potential U,

$$\int \Delta\Phi(r)\rho_e(r)dr + \int \Delta U_{\text{ps}}(r,r')n(r',r)dr\,dr' = \int U^{\text{out}}(r,r')n(r',r)dr\,dr' - \int \Delta V_{\text{xc}}(r)\rho_e(r)dr . \tag{6.4a}$$

Recalling (3.6c), the definition of the model potential U_{m}, right-hand side (rhs) of (6.4a) can also be written as

$$\begin{aligned} & \int \Delta\Phi(r)\rho_e(r)dr + \int \Delta U_{\text{ps}}(r,r')n(r',r)drdr' \\ & = \int U_m^{\text{out}}(r,r')n(r',r)dr\,dr' - \hat{\Phi}\int_{r<R_c}\rho_e(r)dr + \int_{r>R_c}\int_{r'>R_c} U^{\text{out}}(r,r')n(r',r)dr\,dr' - \int \Delta V_{\text{xc}}(r)\rho_e(r)dr . \end{aligned} \tag{6.4b}$$

In the next-to-last term in (6.4b), the $\Delta U_{\text{ps}}(r,r')$ and $\Delta V_{\text{xc}}(r)$ parts of $U^{\text{out}}(r,r')$ do not contribute because ΔU_{ps} is confined to the region $r < R_c$, and because $\Delta V_{\text{xc}}(r)$ vanishes when $\Delta\rho_e(r)$ vanishes beyond R_c by reason of truncation. Thus,

$$\int_{r>R_c}\int_{r'>R_c} U^{\text{out}}(r,r')n(r',r)dr\,dr' = \int_{r>R_c} dr\,\Delta\Phi(r)\rho_e(r)dr . \tag{6.4c}$$

Inserting (6.4) and (6.3) into (6.2) gives

$$\begin{aligned} E[\rho_e\{R_A\}] = {} & \sum_i n_i\epsilon_i + \tfrac{1}{2}\int\int \frac{dr\,dr'}{|r-r'|}[\rho_i(r)\rho_i(r') - \rho_e(r)\rho_e(r')] \\ & + \int\int [U_m^{\text{out}}(r,r') - U_m^{\text{in}}(r,r')]n(r',r)dr\,dr' - \hat{\Phi}\int_{r<R_c}\rho_e(r)dr \\ & + \int_{r>R_c} dr\,\Delta\Phi(r)\rho_e(r) + \int [\epsilon_{\text{xc}}(\rho_e) - V_{\text{xc}}(r)]\rho_e(r)dr + \sum_A \sum_{B\,(\neq A)} C_{AB} - \sum_A J_A . \end{aligned} \tag{6.5}$$

Since the $\Delta U_{\mathrm{ps}}(r,r')$ part of U_m^{in} and of U_m^{out} is the same, $U_{\mathrm{out}}^m - U_{\mathrm{in}}^m$ is a local potential, and we can define

$$\int\int dr\,dr'[U_m^{\mathrm{out}} - U_m^{\mathrm{in}}]n(r',r) = \int_{r<R_c}[U_{\mathrm{scr}}^{\mathrm{out}}(r) - U_{\mathrm{scr}}^{\mathrm{in}}(r)]\rho_e(r)dr \equiv E_{\mathrm{trim}} , \tag{6.6}$$

where

$$U_{\mathrm{scr}}(r) = \Delta\Phi(r) + \Delta V_{\mathrm{xc}}(r) + \hat{\Phi} \tag{6.7}$$

is the local screening part of the model potential. It is $U_{\mathrm{scr}}(r)$ that has to be iterated to self-consistency. At self-consistency, E_{trim} vanishes. Finally, we define

$$E_{\mathrm{xc}} \equiv \int[\epsilon_{\mathrm{xc}}(r) - V_{\mathrm{xc}}(r)]\rho_e(r)dr , \tag{6.8a}$$

$$C\{R_A\} \equiv \sum_A \sum_{B\,(\neq A)} C_{AB} , \tag{6.8b}$$

$$J \equiv \sum_A J_A , \tag{6.8c}$$

and the energy of the defect crystal is then

$$\begin{aligned} E[\rho_e,\{R_A\}] = {} & \sum_i n_i\epsilon_i + E_{\mathrm{trim}} + E_{\mathrm{xc}} + C\{R_A\} - J + \int_{r>R_c}\Delta\Phi(r)\rho_e(r)dr \\ & + \tfrac{1}{2}\int\int\frac{dr\,dr'}{|r-r'|}[\rho_i(r)\rho_i(r') - \rho_e(r)\rho_e(r')] - \hat{\Phi}\int_{r<R_c}\rho_e(r)dr . \end{aligned} \tag{6.9}$$

The corresponding energy for the initial situation, having been obtained via a band-structure calculation, is fully-self-consistent. It is

$$E[\rho_e^0,\{R_A^0\}] = \sum_i n_i^0\epsilon_i^0 + E_{\mathrm{xc}}^0 + C\{R_A^0\} - J^0 + \tfrac{1}{2}\int\int\frac{dr\,dr'}{|r-r'|}[\rho_i^0(r)\rho_i^0(r') - \rho_e^0(r)\rho_e^0(r')] . \tag{6.10}$$

In forming the defect energy (2.4), we subtract these last two expressions. The difference in eigenvalues is evaluated using $\Delta n(E)$, the change in density of states, as

$$\sum_i n_i\epsilon_i - \sum_i n_i^0\epsilon_i^0 = -\frac{1}{\pi}\int_{-\infty}^{E_F} E\frac{d}{dE}\phi(E)dE \tag{6.11}$$

with the phase shifts $\phi(E)$ given in Ref. 39. Evaluation of the terms E_{trim}, and $\Delta E_{\mathrm{xc}} \equiv E_{\mathrm{xc}} - E_{\mathrm{xc}}^0$ cause no particular difficulty because the integrands are confined to the region $r<R_c$.

Next, we consider the change in the electrostatic terms in (6.9) and (6.10). In rearranging these, it is helpful to add and subtract the following finite term:

$$\int_{r<R_c}\int_{\mathrm{all}\,r'}\frac{dr\,dr'}{|r-r'|}\Delta\rho_i(r')\rho_e^0(r) .$$

Having done so, it becomes possible to group terms to show that

$$\begin{aligned} & \tfrac{1}{2}\int\int\frac{dr\,dr'}{|r-r'|}(\rho_i\rho_i - \rho_i^0\rho_i^0 - \rho_e\rho_e + \rho_e^0\rho_e^0) + \int_{r>R_c}\Delta\Phi(r)\rho_e(r)dr \\ & \quad = \int\Phi^0(r)\Delta\rho_i(r)dr - \int_{r<R_c}\Delta\Phi(r)\rho_e^0(r)dr + \tfrac{1}{2}\int\int\frac{dr\,dr'}{|r-r'|}[\Delta\rho_i(r)\Delta\rho_i(r') - \Delta\rho_e(r)\Delta\rho_e(r')] . \end{aligned} \tag{6.12}$$

In arriving at (6.12), we have used the fact that $\Delta\rho_i(r)=0$ for $r>R_c$ because the region R_c has been chosen to be large enough to fully enclose $\Delta\rho_i(r)$, and that $\Delta\rho_e(r)=0$ for $r>R_c$, because that is part of the prescription for evaluating $\Delta\rho_e(r)$.

The first term on the rhs of (6.12) is evaluated using the periodic charge density of the perfect crystal as

$$\int\Phi^0(r)\Delta\rho_i(r)dr = \sum_{G\neq 0}\frac{4\pi}{G^2}[\rho_i(G) + \rho_e(G)]\Delta\rho_i(G) + \langle\Phi^0\rangle\Delta N_i . \tag{6.13}$$

In (6.13),

$$\rho_{i,e}(G) = \frac{1}{\Omega_c}\int_{\Omega_c}\rho_{i,e}^0(r)e^{-iG\cdot r}dr \tag{6.14a}$$

is a Fourier component of the charge density of the perfect crystal, evaluated over the periodic cell of volume Ω_c, and available either analytically (ρ_i^0) or in numeric form from a previous band-structure calculation (ρ_e^0): The term

$$\Delta\rho_i(G) = \int\Delta\rho_i(r)e^{iG\cdot r}dr \tag{6.14b}$$

is a Fourier integral of the change in ionic density, a quantity which, with the choice (5.1) and (5.2), is available in analytic form. $\langle\Phi^0\rangle$ is the spatial average of the electrostatic part of the perfect-crystal potential. It is arbitrary. Its value affects the eigenvalues in (6.11) and the reservoir energy μ additively. The zero of the eigenvalue scale *must* be taken compatible with the choice of $\langle\Phi^0\rangle$. The quantity ΔN_i is the change in ionic charge upon creation of the defect.

The second term on the rhs of (6.10) is usefully combined with the last remaining term in (6.9) as

$$-\int_{r<R_c}[\Delta\Phi(r)\rho_e^0(r)+\hat{\Phi}\rho_e(r)]dr=-\int_{r<R_c}[U_{\rm scr}^{\rm out}(r)-\Delta V_{\rm xc}(r)]\rho_e^0(r)dr-\hat{\Phi}\int_{r<R_c}\Delta\rho_e(r)dr\ . \tag{6.15}$$

Finally, adding the reservoir energy $-\mu n$ and collecting all of the term described above, we obtain the defect energy (2.4) in the following form

$$\begin{aligned}\Delta E=&-\frac{1}{\pi}\int_{-\infty}^{E_F}E\frac{d\phi(E)}{dE}dE+\Delta C\{R_A\}+\sum_{\rm removed}J_R-\sum_{\rm added}J_A+\tfrac{1}{2}\int\int\frac{dr\,dr'}{|r-r'|}[\Delta\rho_i(r)\Delta\rho_i(r')-\Delta\rho_e(r)\Delta\rho_e(r')]\\&+\sum_{G\neq0}\frac{4\pi}{G^2}[\rho_i(G)+\rho_e(G)]\Delta\rho_i(G)+\langle\Phi^0\rangle\Delta N_i-\mu n-\int_{r<R_c}[U_{\rm scr}^{\rm out}(r)-\Delta V_{\rm xc}(r)]\rho_e^0(r)dr\\&-\hat{\Phi}\int_{r<R_c}\Delta\rho_e(r)dr+E_{\rm trim}+\Delta E_{\rm xc}\ .\end{aligned} \tag{6.16}$$

Equation (6.16) is the expression we have used to evaluate the defect energies reported in Ref. 40. We conclude this section with three comments about the above expression.

(1) Although the term $E_{\rm trim}$, Eq. (6.6), vanishes at self-consistency, including it in (6.16) stabilizes the calculated value of ΔE remarkably as self-consistency is approached. Our experience has been that the largest value of $U_{\rm scr}^{\rm out}(r)-U_{\rm scr}^{\rm in}(r)$ that can be tolerated is typically a factor of 10^2 greater when $E_{\rm trim}$ is included than when it is not. This can be traced back to the variational nature of the energy expression (2.4).[41] Fortunately, much of the underlying variational insensitivity is retained here, even though there has been truncation of the Green's function, of the potential, and of the charge.

(2) The phase shift $\Phi(E)$ often has sharp structure, exhibiting resonances, etc., that make difficult an accurate evaluation of (6.11), the integral giving the eigenvalue change. A similar difficulty occurs in evaluating $\Delta\rho_e(r)$, Eq. (3.8), the integrand having the same sort of structure. Just as Williams *et al.*[30] and earlier workers have noted that (3.8) can be replaced by a contour integral in the complex plane surrounding the occupied spectrum, exactly the same replacement is possible for (6.11), requiring only that the expression (5.3) for phase shift be generalized to the complex plane. The appropriate generalization of (6.11) is

$$\sum_i n_i\epsilon_i-\sum n_i^0\epsilon_i^0=\frac{1}{2\pi i}\int z\left[\frac{d\phi(z)}{dz}\right]dz\ , \tag{6.17a}$$

where

$$\phi(z)=\ln[\det(U-UG_z^0)/\det(U)]\ . \tag{6.17b}$$

A contour parallel to the imaginary axis, intersecting the real axis at E_F, and completed by the semicircle at infinity, provides the smoothest integrand, one for which it is particularly easy to devise integration schemes which make use of our knowledge of the asymptotic (large-z) behavior of the integrand. We have been using this contour consistently instead of the real contour in (6.16). We find that it has very much to recommend it, although at the price of losing all the physical interpretive information contained in $\phi(E)$ for real E. A compromise, reported to us by Scheffler,[42] is to use a contour parallel to, but somewhat above, the real axis, which will smooth, but not obliterate, the structure in $\phi(E)$.

(3) To demonstrate that (6.16) is independent of $\langle\Phi^0\rangle$, imagine that $\langle\Phi^0\rangle$ is replaced by $\langle\Phi^0\rangle+\chi$, where χ is some constant. This replaces the eigenvalues ϵ_i by $\epsilon_i+\chi$, and ϵ_i^0 by $\epsilon_i^0+\chi$, which changes the integral in (6.16) by

$$-\frac{1}{\pi}\int_{-\infty}^{E_F}\chi\frac{d\phi}{dE}dE=\chi\int_{-\infty}^{E_F}\Delta n(E)dE=\chi\Delta N_e\ .$$

The reservoir-energy change μ is measured on the same energy scale as the eigenvalues. Therefore, μ is replaced by $\mu+\chi$, and $-\mu n$ is changed by $-\chi n$. Finally, $\langle\Phi^0\rangle\Delta N_i$ is changed by $\chi\Delta N_i$. There are no other $\langle\Phi^0\rangle$-dependent quantities and the total change in ΔE is

$$\delta\Delta E=\chi(\Delta N_e+\Delta N_i-n)\ .$$

This is zero because $\Delta N_e+\Delta N_i=n$ is the nominal charge on the defect, with charge supplied by the n particles lost from the reservoir.

VII. COMPARISON BETWEEN THE EIGENVALUE AND KINETIC-ENERGY FORMULATIONS

The kinetic-energy formalism is consistent in ignoring all contributions (except the purely elastic) to the energy density in region 2. All kinetic-energy, exchange-energy, and Coulomb terms were systematically discarded in this region. In that way, macroscopic dielectric theory and microscopic linear-response theory, which must agree with each other, are made to agree in this situation.

The eigenvalue formulation of the defect energy is not consistent in this regard: In all derivations of the eigenvalue form of energy functional there is a step where the Schrödinger equation for $\psi_i(r)$ is multiplied by $n_i\psi_i^*(r)$, summed over i, *and integrated over all space.* The normalization of the wave functions $\int\psi_i\psi_i^*dr=1$ then brings in the familiar eigenvalue sum $\sum_1 n_i\epsilon_i$, which, after correction for overcounting, replaces the one-particle terms in the energy functional. Therefore the eigenvalue sum necessarily contains contributions from some, but not all of the individual energy terms in region 2. Since the underlying Schrödinger equation which gave rise to the wave functions ψ_i in the defect crystal contained the potential U_m rather than U, the eigenvalue sum will pick up spurious energy contributions associated with $\Delta^m \rho_e(r)$ in the region $R_c < r < 2\tilde{R}_c$.

Formal expressions for the difference in energy between the eigenvalue and kinetic-energy formulations can easily be derived. Evaluating them demands knowledge of $\Delta^m \rho_e(r)$ (and its density-matrix equivalent) in the region $R_c < r < 2\tilde{R}_c$, something which we do not have. This situation calls for a direct numerical evaluation of the difference. This is work in progress whose results will be reported elsewhere.[43]

In summary, the choice is between use of a kinetic-energy formulation whose treatment of the outer response region, ignoring it, is compatible with the consequences of using macroscopic response theory in this region and use of the numerically simpler eigenvalue formulation which does include some of the energy contributions from this region in a way which cannot be justified formally but whose effects, at this preliminary stage, seem to be small.

ACKNOWLEDGMENT

We are happy to acknowledge many useful discussions with E. O. Kane during the course of this work.

APPENDIX: ILLUSTRATIVE EVALUATION OF DIELECTRIC ENERGY

For a defect with nominal charge n, the macroscopic averaged defect potential is $\phi(r) = n/\epsilon r$, the E, D, and P fields will be directed radially outward with $E = n/\epsilon r^2$, $D = \epsilon E$, and

$$P = (D - E)/4\pi = n \left[1 - \frac{1}{\epsilon} \right] \left[\frac{1}{4\pi r^2} \right] .$$

Let the crystal be spherical, with an outer radius R_0, and let the inner radius of the response region be R_c. Then the integral (3.17) is

$$\delta U = -\frac{1}{2} \frac{n^2}{\epsilon} \left[1 - \frac{1}{\epsilon} \right] \left[\frac{1}{R_c} - \frac{1}{R_0} \right] . \tag{A1}$$

The surface charge density (3.18) on the outer surface is

$$\sigma^{\text{out}} = n \left[1 - \frac{1}{\epsilon} \right] \left[\frac{1}{4\pi R_0^2} \right] , \tag{A2}$$

giving a total charge on the outer surface of $n(1 - 1/\epsilon)$, in agreement with (3.11). The energy expression (A1) will be recognized as being of the form (3.19) provided that there are two sources of polarization potential ϕ_p [Eq. (3.20)], namely σ^{out}, at R_0, and

$$\sigma^{\text{in}} = -n \left[1 - \frac{1}{\epsilon} \right] \left[\frac{1}{4\pi R_c^2} \right] ,$$

the charge density (3.18) located at R_c. The polarization potential $\phi_p(r)$ acts on the defect charge for which

$$\int_{r < R_c} \rho_{\text{def}}(r) dr = n/\epsilon . \tag{A3}$$

The σ^{in} is spurious, arising from improper use of the polarization formalism: Dropping its contribution to (A1) gives

$$\delta U = \frac{1}{2} n^2 \epsilon R_0 \left[1 - \frac{1}{\epsilon} \right] ,$$

which goes to zero as $R_0 \to \infty$.

---

[1] L. C. Kimerling, in *Defects and Radiation Effects in Semiconductors, Nice, 1978,* edited by J. H. Albany (IOP, London, 1979), p. 56.

[2] P. J. Dean and W. J. Choyke, Adv. Phys. **26**, 1 (1977).

[3] J. R. Troxell, A. P. Chatterjee, G. D. Watkins, and L. C. Kimerling, Phys. Rev. B **19**, 5336 (1979).

[4] C. H. Henry, H. Kukimoto, G. L. Miller, and F. R. Merritt, Phys. Rev. B **7**, 2499 (1973).

[5] J. L. Newton, A. P. Chatterjee, R. D. Harris, and G. D. Watkins, in *Proceedings of the 12th International Conference on Defects in Semiconductors* edited by C. A. J. Ammerlaan (North-Holland, Amsterdam, 1983), p. 219.

[6] D. V. Lang and R. A. Logan, Phys. Rev. Lett. **39**, 635 (1977).

[7] Y. Toyozawa, Solid-State Electron. **21**, 1313 (1978).

[8] M. Levinson, M. Stavola, J. L. Benton, and L. C. Kimerling, Phys. Rev. B **28**, 5848 (1983).

[9] G. Vincent and D. Bois, Solid State Commun. **27**, 431 (1978).

[10] G. A. Baraff and M. Schlüter, Phys. Rev. B **28**, 2296 (1983).

[11] J. Bernholc, N. O. Lipari, and S. T. Pantelides, Phys. Rev. B **21**, 3545 (1980).

[12] G. A. Baraff and M. Schlüter, Phys. Rev. B **19**, 4965 (1979).

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

[14] P. C. Hohenberg and W. Kohn, Phys. Rev. **136**, B864 (1964).

[15] W. Kohn and L. Sham, Phys. Rev. **140**, A1133 (1965).

[16] D. R. Hamann, M. Schlüter, and C. Chiang, Phys. Rev. Lett. **43**, 1494 (1979).

[17] D. R. Hamann, Phys. Rev. Lett. **42**, 662 (1979).

[18] M. T. Yin and M. L. Cohen, Phys. Rev. B **26**, 5668 (1982).

[19] J. P. Perdew and M. Levy, Phys. Rev. Lett. **51**, 1884 (1983).

[20] L. J. Sham and M. Schlüter, Phys. Rev. Lett. **51**, 1888 (1983).

[21] C. S. Wang and W. E. Pickett, Phys. Rev. Lett. **51**, 597 (1983).

[22] G. Strinati, H. J. Mattausch, and W. Hanke, Phys. Rev. B **25**, 2867 (1982).

[23] U. Lindefelt and A. Zunger, Phys. Rev. B **26**, 846 (1982).

[24] M. Scheffler, in *Festkörperprobleme (Advances in Solid State Physics),* edited by J. Treusch (Vieweg, Braunschweig, 1982), Vol. XXII, p. 115.

[25] G. A. Baraff and M. Schlüter, in *Proceedings of the XIth International Conference on Defects and Radiation Effects in Semiconductors, Oiso, Japan, 1980,* edited by R. R. Hasiguti (IOP, Bristol, 1981), p. 287.

[26] G. B. Bachelet, M. Schlüter, and G. A. Baraff, Phys. Rev. B **27**, 2545 (1983).

[27] G. F. Koster and J. C. Slater, Phys. Rev. **95**, 1167 (1954).

[28] N. F. Mott and R. W. Guerney, *Electronic Processes in Ionic Crystals* (Oxford University Press, New York, 1948), Chap. II.

[29] K. Weiser, Phys. Rev. **126**, 1427 (1962).

[30] A. R. Williams, P. J. Feibelman, and N. D. Lang, Phys. Rev. B **26**, 5433 (1982).

[31] M. Schlüter and L. J. Sham, Physics Today **35**(2), 36 (1982), and references therein.

[32] A. R. Williams and U. von Barth, in *Theory of the Inhomogeneous Electron Gas,* edited by S. Lundquist and N. H. March (Plenum, New York, 1983).

[33] G. A. Baraff, M. Schlüter, and G. Allan, Phys. Rev. B **27**, 1010 (1983); U. Lindefelt and A. Zunger, *ibid.* **24**, 5913 (1981).

[34] E. O. Kane (private communication), and unpublished.

[35] J. Callaway, J. Math Phys. **5**, 783 (1964).

[36] For a contour plot of $\Delta\rho_e(r)$ which illustrates this bond displacement in a uniform electric field, see Fig. 2(a) of A. Baldereschi, R. Car, and E. Tosatti, Solid State Commun. **32**, 757 (1979).

[37] N. W. Ashcroft and N. D. Mermin, *Solid State Physics* (Holt, Rinehart and Winston, New York, 1976), Chap. 27.

[38] L. D. Landau and E. M. Lifshitz, *Electrodynamics of Continuous Media* (Addison-Wesley, Reading, Mass., 1960), Chap. II.

[39] G. A. Baraff, E. O. Kane, and M. Schlüter, Phys. Rev. B **21**, 5662 (1980).

[40] G. A. Baraff and M. Schlüter (unpublished).

[41] J. R. Chelikowsky and S. G. Louie, Phys. Rev. B **29**, 3470 (1984).

[42] M. Scheffler (private communication).

[43] G. A. Baraff and M. Schlüter (unpublished).