Theory of localized defects in solids

J. G. Gay and J. R. Smith

Research Laboratories, General Motors Corporation, Warren, Michigan 48090

(Received 22 June 1973)

Short-range defects are found, for example, at solid surfaces and at deep and isoelectronic traps in the bulk. We have extended a new concept, introduced by Kohn and Onffroy, in order to treat local defects in three-dimensional solids. The new concept involves generalized Wannier functions, which can be obtained variationally. These functions vary from site to site in the (nonperiodic) solid and make a complete orthonormal basis set for the bound and continuum wave functions. Questions as to the validity of the method in three dimensions are answered. The basis functions are calculated via perturbation theory for local defects in a wide-band-gap insulator. An exact set of difference equations similar to the one-band Koster-Slater equations and to those presented by Kohn and Onffroy is derived for the many-band wave functions of the nonperiodic system. It is shown how to obtain charge densities and local densities of states without prior knowledge of the electron wave functions.

I. INTRODUCTION

Many technological applications of solid-state physics depend on the detailed electronic structure in the vicinity of localized defects. The electronic structure of bulk defects strongly influences, for example, the luminescence characteristics of solids. Chemisorption, which is basic to catalytic and corrosion processes, produces a local defect in the surface region of solids.

Defects which produce weak and extended perturbing potentials lead to shallow impurity levels with large electronic orbits. These are relatively well understood on the basis of an effective mass formalism.[1] Those situations in which the impurity wave function is well localized have been a traditional source of frustration, however.

This paper makes use of a new concept introduced by Kohn and Onffroy[2] in the field of localized defects in solids. The concept involves generalized Wannier functions which can be obtained variationally. The generalized-Wannier-function formalism is particularly well suited to deep, sharply localized potentials and thus compliments the effective-mass formalism for shallow potentials.

There are many examples of short-range defects. Isoelectronic impurities are impurity atoms from the same column of the Periodic Table as one of the atoms of the host lattice,[3] e. g., GaP: N. One encounters short-range impurity potentials in these cases. Impurities such as Cu and Au in Si can produce "deep" traps, binding carriers with energies of the order of 0. 5 eV. The orbits associated with these states are well localized.[1] Lattice vacancies[4] are still another example of a localized bulk defect. Surfaces also produce rapidly varying inhomogeneities in solids. Recently experimentalists have been able to measure local densities of states in the surface region.[5,6] Thus the time is ripe for a formalism which can be used to calculate self-consistently the energy spectra of electrons in chemisorption bonds (self-consistent in the sense that the electron wave functions are determined from a potential which depends on the wave functions themselves).

The theoretical means presently at our disposal for solving localized impurity problems is the well known Koster-Slater method.[7-9] Calais and Ribbing[10] review the method, and point out that it requires that one know both the impurity potential and the full set of energy bands and corresponding wave functions for the perfect crystal. This is corroborated by the calculations of Callaway and Hughes,[4] Parada,[4] and Faulkner[3] which showed the necessity of including a number of bands. However, many-band Koster-Slater calculations are difficult and there have been no calculations which treat both bound and virtual levels self-consistently.

The generalized Wannier function concept involves a different approach to the problem. The influence of all of the bands are included directly in the generalized Wannier functions which are then used as a basis for the expansion for the wave functions of the nonperiodic solid.

For a system consisting of an isolated impurity in a one-dimensional lattice, Kohn and Onffroy[2] have shown that the generalized Wannier functions, $a_n(x)$, have exactly the same asymptotic behavior as the Wannier functions of the periodic lattice, $a_n^0(x)$. Further, they showed that for lattice sites n asymptotically far from the perturbation, the $a_n(x)$ approach the $a_n^0(x)$ exponentially in n.

Here we analyze and apply the generalized Wannier function concept to three-dimensional systems with local defects. Des Cloizeaux[11] has enumerated the difficulties that three dimensionality introduces into the analysis of the asymptotic behavior of Wannier functions. We employ an approach based in spirit on Koster-Slater theory to characterize the local behavior of the generalized Wan-

nier functions. The results provide a framework for a systematic variational treatment of solids with local defects.

In Sec. II we consider the question of the localization of the generalized Wannier functions in three dimensions. There we show that the perturbed wave functions of a band can be expressed in terms of a set of functions localized about each lattice site. In addition it is shown that as the distance from the perturbation increases, the generalized Wannier functions reduce to the conventional Wannier functions of the unperturbed periodic system.

In Sec. III generalized Wannier functions are calculated, for the case of local defects in a wide-band-gap insulator. The dependence of the generalized Wannier functions on the bandwidth and band gap is determined.

Sec. IV is devoted to calculational techniques for determining generalized Wannier functions, wave functions, charge densities, and local densities of states. The procedure for calculating the wave functions involves a set of difference equations similar to the one-band Koster-Slater equations. The local density of states and, for filled bands, the charge density may be calculated directly from the generalized Wannier functions without knowledge of the wave functions.

Finally in Sec. V we discuss the advantages of the formalism in application to such local defects as bulk traps and surfaces. Some insight into the chemical-bond theory of solids is also obtained.

II. GENERALIZED WANNIER FUNCTIONS

We begin Sec. II by setting down some familiar results from band theory and the theory of Wannier functions. These results along with a brief discussion of the Koster-Slater formalism for treating crystals with localized perturbations will serve as a background for our demonstration of the localization of the generalized Wannier functions.

A basic assumption of band theory is that the electronic states of crystals are adequately described by a one-electron Hamiltonian of the form

$$H^0 = -\tfrac{1}{2}\nabla^2 + V(\vec{r}), \tag{2.1}$$

where $V(\vec{r})$ is periodic in the direct lattice so that for any lattice vector $\vec{n}$,

$$V(\vec{r}+\vec{n}) = V(\vec{r}) . \tag{2.2}$$

Because of the periodicity of H^0, its eigenfunctions are Bloch waves $\phi_\alpha^0(\vec{k}, \vec{r})$ of band index α and wave vector $\vec{k}$,

$$H^0\phi_\alpha^0(\vec{k}, \vec{r}) = E_\alpha^0(\vec{k})\, \phi_\alpha^0(\vec{k}, \vec{r}). \tag{2.3}$$

Again because of periodicity, these Bloch functions can be expanded in terms of the Wannier functions of band α,

$$\phi_\alpha^0(\vec{k}, \vec{r}) = N^{-1/2}\sum_{\vec{n}} e^{i\vec{k}\cdot\vec{n}} a_\alpha^0(\vec{r} - \vec{n}) . \tag{2.4}$$

In Eq. (2.4) N is the number of unit cells in the crystal and $a_\alpha^0(\vec{r} - \vec{n})$ is the Wannier function of band α located at site $\vec{n}$.

The essential idea of the generalized Wannier function formalism is to develop a relation analogous to (2.4) for nonperiodic systems. Before proceeding, however, we wish to outline the existing Koster-Slater theory[7-9] for nonperiodic systems. To do this let us introduce a localized perturbation $V^\rho(\vec{r})$ into the crystal so that the Hamiltonian of the perturbed nonperiodic system is,

$$H = H^0 + V^\rho(\vec{r}) . \tag{2.5}$$

We denote the eigenfunctions of H by $\phi_\alpha(\vec{q}, \vec{r})$,

$$H\phi_\alpha(\vec{q}, \vec{r}) = E_\alpha(\vec{q})\, \phi_\alpha(\vec{q}, \vec{r}), \tag{2.6}$$

where α is a band index and $\vec{q}$ an intraband quantum number. Since the $\phi_\alpha^0(\vec{k}, \vec{r})$ form a complete set we may expand the $\phi_\alpha(\vec{q}, \vec{r})$ in terms of them, or, equivalently, in terms of the Wannier functions of the various bands,

$$\phi_\alpha(\vec{q}, \vec{r}) = \sum_{\alpha'\vec{n}'} U_{\alpha\alpha'}(\vec{q}, \vec{n}')\, a_{\alpha'}^0(\vec{r} - \vec{n}'). \tag{2.7}$$

This is the fundamental equation of the Koster-Slater method.[7] The $\phi_\alpha(\vec{q}, \vec{r})$ and the corresponding eigenvalues $E_\alpha(\vec{q})$ are determined by converting the Schrodinger equation (2.6) into a set of difference equations for the $U_{\alpha\alpha'}(\vec{q}, \vec{n}')$ as follows.[7,8,12] The Schrodinger equation (2.6) is multiplied by the resolvent operator $[E_\alpha(\vec{q}) - H^0]^{-1}$ to obtain an expression for the wave function,

$$\phi_\alpha(\vec{q}, \vec{r}) = \psi(\vec{r}) + [E_\alpha(\vec{q}) - H^0]^{-1}V^\rho(\vec{r})\phi_\alpha(\vec{q}, \vec{r}), \tag{2.8}$$

where $\psi(\vec{r})$ is any solution to the unperturbed problem with the same energy and boundary conditions as $\phi_\alpha(\vec{q}, \vec{r})$. For $E_\alpha(\vec{q})$ lying in band α, $\phi_\alpha(\vec{q}, \vec{r})$ is a continuum state and $\psi(\vec{r})$ is taken as a Bloch wave $\phi_\alpha^0(\vec{k}, \vec{r})$ with energy $E_\alpha^0(\vec{k}) = E_\alpha(\vec{q})$. For bound states $\psi(\vec{r})$ is zero because $E_\alpha(\vec{q})$ lies in a band gap and there is no unperturbed solution satisfying the boundary conditions.

When (2.8) is expressed in a Wannier-function representation the Koster-Slater difference equations result,

$$U_{\alpha\beta}(\vec{q}, \vec{n}) = \theta N^{-1/2} e^{i\vec{k}\cdot\vec{n}}\delta_{\alpha\beta} + \sum_{\alpha'}\sum_{\vec{m},\vec{n}'} G_\beta^0(\vec{n} - \vec{m}; E_\alpha(q)) V_{\beta\alpha'}^\rho(\vec{m}, \vec{n}') U_{\alpha\alpha'}(\vec{q}, \vec{n}') . \tag{2.9}$$

In (2.9), $G_\beta^0(\vec{n} - \vec{m};\ E)$ is a Green's function for band β obtained by taking matrix elements of the resolvent between Wannier functions,

$$G_\beta^0(\vec{n} - \vec{m};\ E) = \lim_{\epsilon \to 0} \langle a_\beta^0(\vec{r} - \vec{n}) | [E + i\epsilon - H^0]^{-1} | a_\beta^0(\vec{r} - \vec{m}) \rangle = N^{-1} \lim_{\epsilon \to 0} \sum_{\vec{k}'} \frac{e^{i\vec{k}' \cdot (\vec{n} - \vec{m})}}{E + i\epsilon - E_\beta^0(\vec{k}')} , \qquad (2.10)$$

where ϵ is a positive parameter. Similarly,

$$V_{\beta\alpha'}^{\rho}(\vec{m}, \vec{n}') = \langle a_\beta^0(\vec{r} - \vec{m}) | V^\rho(\vec{r}) | a_{\alpha'}^0(\vec{r} - \vec{n}') \rangle , \qquad (2.11)$$

and

$$\theta = 1, \quad \text{if } E_\alpha(\vec{q}) \text{ lies in band } \alpha$$

$$= 0, \quad \text{otherwise} . \qquad (2.12)$$

The fact that $V^\rho(\vec{r})$ is localized means that the matrix elements (2.11) are appreciable only on a certain finite number N_ρ of sites.

Let us first consider the case that α is an isolated band not overlapped by other bands. If $\phi_\alpha(\vec{q}, \vec{r})$ is a continuum state then there is a singularity in $G_\alpha^0(\vec{n} - \vec{m};\ E)$ which is circumvented in (2.10) in such a way that $G_\alpha^0(\vec{n} - \vec{m};\ E)$ is an outgoing Green's function. However there are no singularities in the $G_\beta^0(\vec{n} - \vec{m};\ E)$ of other bands so they are nonpropagating and the contributions to $\phi_\alpha(\vec{q}, \vec{r})$ from these bands are damped. As a consequence of this, sufficiently far from the perturbation, $\phi_\alpha(\vec{q}, \vec{r})$ may be expressed in terms of the Wannier functions of band α alone,

$$\phi_\alpha(\vec{q}, \vec{r}) = \sum_{n'} U_{\alpha\alpha}(\vec{q}, \vec{n}') a_\alpha^0(\vec{r} - \vec{n}'). \qquad (2.13)$$

We shall term the region where (2.13) holds the one-band region and the region of N_0 unit cells where it does not the many-band region. The rate at which the contribution from a band α' in the many-band region decays is determined by the rate of decay of the Green's function of that band. Examples of the behavior of the Green's function of a model band as calculated by Koster and Slater[9] for energies outside its band are shown in Fig. 1. Note that the many-band region extends only a few lattice constants beyond the region for which the perturbation has nonnegligible matrix elements. The decay of the Green's function is of practical importance to the method. We shall show that the generalized Wannier functions need differ from the $a_\alpha^0(\vec{r} - \vec{n})$ only in the many-band region at most. This greatly simplifies the variational methods of Sec. IV. Note we study the *near*-distance localization characteristics of the generalized Wannier functions, that is, the behavior within a few lattice sites of the defect.

The situation is more complex for bound states. In this case since $E_\alpha(\vec{q})$ lies in a band gap, $G_\alpha^0(\vec{n} - \vec{m};\ E_\alpha(\vec{q}))$ is also nonpropagating, and there is no incoming Bloch wave to identify the band to which

the bound state belongs. We identify the bands to which the various bound states belong by seeing with which band they merge as the strength of the perturbation is reduced. We thereby determine the band index α for each bound state.

Equation (2.13) still holds for bound states. For energies $E_\alpha(\vec{q})$ distinctly closer to band α than neighboring bands, the contributions from band α will be less strongly damped than the contributions of other bands (see Fig. 1), and once the contributions from other bands have decayed (2.13) will hold. For deep bound states the contributions from all bands will be sharply damped and the entire bound-state wave function can be confined to the many-band region (by enlarging it if necessary). Equation (2.13) will hold in this case because the bound-state wave function vanishes in the one-band region.

In the case where band α is overlapped by other bands, there will be scattered waves from all
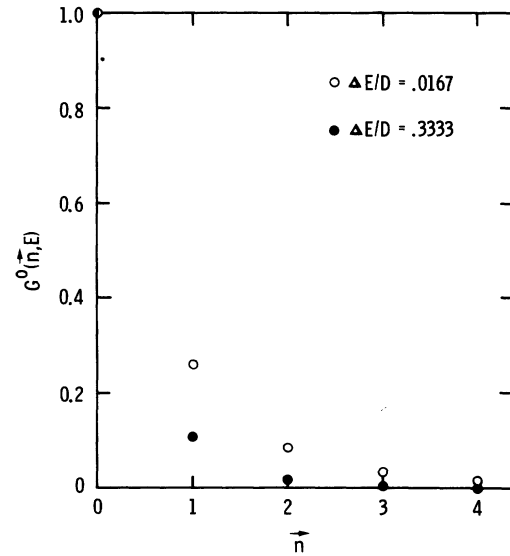

FIG. 1. Plots of $G^0(\vec{n}, E)$ along a [100] direction as computed by Koster-Slater (Ref. 9) for a model band produced by nearest-neighbor interactions in a simple cubic lattice. D is the bandwidth and ΔE is the distance above the top (or below the bottom) of the band of the energy E at which $G^0(\vec{n}, E)$ is evaluated. The decay of $G^0(\vec{n}, E)$ goes asymptotically as $e^{-\lambda|\vec{n}|}/|\vec{n}|$ and λ depends only on the ratio $\Delta E/D$. Note that for E less than 2% of the bandwidth away from the band, $G^0(\vec{n}, E)$ has decayed to negligible values beyond four lattice constants.

these bands contributing to $\phi_\alpha(\vec{q}, \vec{r})$ in the "one-band" region and (2.13) must be replaced by

$$\phi_\alpha(\vec{q}, \vec{r}) = \sum_{\alpha' \vec{n}'}' U_{\alpha\alpha'}(\vec{q}, \vec{n}') a_{\alpha'}^0(\vec{r} - \vec{n}') , \qquad (2.14)$$

where the prime denotes that the sum is only over interconnecting bands.

We shall make considerable use of the formal results of the Koster-Slater theory in the following. However, we note that as a practical calculational procedure the Koster-Slater procedure is difficult. Not only must one know the Wannier functions of all the bands which contribute to $\phi_\alpha(\vec{q}, \vec{r})$ but also the Green's functions of these bands as well. The latter requires a knowledge of $E^0(\vec{k})$ as a function of $\vec{k}$ for all bands which contribute to $\phi_\alpha(\vec{q}, \vec{r})$.

The generalized-Wannier-function formalism circumvents most of these difficulties. The many-band expansion (2.7) is replaced by an equation analogous to the single-band expansion (2.4) for the unperturbed system, namely,

$$\phi_\alpha(\vec{q}, \vec{r}) = \sum_{\vec{n}} C_{\alpha\alpha}(\vec{q}, \vec{n}) a_{\alpha, \vec{n}}(\vec{r} - \vec{n}). \qquad (2.15)$$

In (2.15) $a_{\alpha, \vec{n}}(\vec{r} - \vec{n})$ is a generalized Wannier function of band α. The site subscript $\vec{n}$ denotes that the $a_{\alpha, \vec{n}}(\vec{r} - \vec{n})$ may vary from site to site because the system is not periodic. In addition to this difference, the $C_{\alpha\alpha}(\vec{q}, \vec{n})$ are not plane waves and will differ from band to band. This is denoted by the subscripts $\alpha\alpha$.

As noted in the introduction the practical advantage of the method is that a variational principle can be established for the $a_{\alpha, \vec{n}}(\vec{r} - \vec{n})$ so that they can be calculated by a variational procedure similar to that put forward by Kohn[13] for periodic systems. A knowledge of the $a_{\alpha, \vec{n}}(\vec{r} - \vec{n})$ alone is sufficient to calculate the local density of states for band α, total energies (e. g. , heats of adsorption), and the charge density. There is a further advantage in that we will show that the coefficients $C_{\alpha\alpha}(\vec{q}, \vec{n})$ can be obtained from a set of difference equations similar to the Koster-Slater equations but involving the single band α. Thus the wave functions can be obtained by a simpler procedure than that of Koster-Slater.

The above points will be discussed in subsequent sections. Our purpose in Sec. II is to study the transformation $C_{\alpha\alpha}(\vec{q}, \vec{n})$ and to show that it may be chosen in such a way that the $a_{\alpha, \vec{n}}(\vec{r} - \vec{n})$ are indeed Wannier-function-like. Note that there is no question that one can define a set of $a_{\alpha, \vec{n}}(\vec{r} - \vec{n})$ such that (2.15) is satisfied. In fact, if the $C_{\alpha\alpha}(\vec{q}, \vec{n})$ are the elements of an *arbitrary* unitary matrix, $a_{\alpha, \vec{n}}(\vec{r} - \vec{n})$ defined by

$$a_{\alpha, \vec{n}}(\vec{r} - \vec{n}) = \sum_{\vec{q}} C_{\alpha\alpha}^*(\vec{q}, \vec{n}) \phi_\alpha(\vec{q}, \vec{n}) \qquad (2.16)$$

are orthonormal and automatically satisfy (2.15). What we wish to show is that we may construct the $C_{\alpha\alpha}(\vec{q}, \vec{n})$ so that the $a_{\alpha, \vec{n}}(\vec{r} - \vec{n})$ are localized about their sites. We will actually demonstrate that the $a_{\alpha, \vec{n}}(\vec{r} - \vec{n})$ can be localized and that away from the perturbation they reduce to the $a_\alpha^0(\vec{r} - \vec{n})$ of the unperturbed system.

If we insert the form (2.7) for $\phi_\alpha(\vec{q}, \vec{r})$ into (2.16) we see that the $a_{\alpha, \vec{n}}(\vec{r} - \vec{n})$ are linear combinations of the $a_{\alpha'}^0(\vec{r} - \vec{n})$,

$$a_{\alpha, \vec{n}}(\vec{r} - \vec{n}) = \sum_{\alpha' \vec{n}'} T_{\alpha\alpha'}(\vec{n}, \vec{n}') a_{\alpha'}^0(\vec{r} - \vec{n}'), \qquad (2.17)$$

where

$$T_{\alpha\alpha'}(\vec{n}, \vec{n}') = \sum_{\vec{q}} C_{\alpha\alpha}^*(\vec{q}, \vec{n}) U_{\alpha\alpha'}(\vec{q}, \vec{n}'). \qquad (2.18)$$

Note first that the localization of the $a_{\alpha, \vec{n}}(\vec{r} - \vec{n})$ depends on the localization of the Wannier functions themselves. The question of the localization of the $a_\alpha^0(\vec{r} - \vec{n})$ has been extensively studied.[11,13,14] It has been found that in many cases exponentially localized Wannier functions can be constructed for simple or composite bands isolated from other bands by energy gaps. In many real solids the bands are connected to adjacent bands. In such cases it has been argued[12] that exponentially localized Wannier functions yield arbitrarily accurate results for general points in the interior of the zone, but may give incorrect results at points on or near the zone boundary or at any interior point where there is attachment to another band.

We do not address the question of the localization of the $a_\alpha^0(\vec{r} - \vec{n})$. We will assume that they can be adequately represented by localized functions and will consider the other aspect involved in the localization of the $a_{\alpha, \vec{n}}(\vec{r} - \vec{n})$, namely the form of the transformation matrix T in (2.18). For the $a_{\alpha, \vec{n}}(\vec{r} - \vec{n})$ to be localized it must be that the $T_{\alpha\alpha'}(\vec{n}, \vec{n}')$ diminish sufficiently rapidly as $\vec{n}'$ becomes distant from $\vec{n}$. Further, if $a_{\alpha, \vec{n}}(\vec{r} - \vec{n})$ is to go to $a_\alpha^0(\vec{r} - \vec{n})$ for $\vec{n}$ distant from the perturbation, we must have that $T_{\alpha\alpha'}(\vec{n}, \vec{n}') \to \delta_{\alpha\alpha'} \delta_{\vec{n}, \vec{n}'}$ for $\vec{n}$ distant from the perturbation.

The generalized Wannier function formalism is applicable to perturbations which are localized in at least one dimension. However, in the following development and, for the most part, in the remainder of the paper we will assume that $V^p(\vec{r})$ is localized in all three dimensions, corresponding, for example, to the presence of an impurity.

For the time being we will assume that the bands of H^0 are nonoverlapping. This is an artificial assumption except in one dimension but it is the simpler case and will be dealt with first. Further, for purposes of illustration we will consider just two bands. The derivation can be easily generalized to an arbitrary number of bands. Just how

many bands are needed for a given accuracy will depend upon the perturbing potential. It is the spirit of the Koster-Slater equations (2.9) that only a manageable number of bands will be required. This appears to be born out by the calculations of Callaway and Hughes[4] for vacancies in silicon. They found that the calculation converged as the number of energy bands included was increased, and that low-lying bands were considerably more important than higher bands.

Before proceeding to construct the C matrix we will display schematically the properties of the U matrix set forth above by partitioning it into submatrices $U_{\alpha\alpha'}$,

$$U = \begin{bmatrix} U_{11} & U_{12} \\ U_{21} & U_{22} \end{bmatrix}, \qquad (2.19)$$

where the submatrices $U_{\alpha\alpha'}$ have the elements $U_{\alpha\alpha'}(\vec{q}, \vec{n})$. If we arrange the wave functions and Wannier functions into column vectors similarly partitioned, (2.19) may be written,

$$\begin{bmatrix} \phi_1 \\ \phi_2 \end{bmatrix} = \begin{bmatrix} U_{11} & U_{12} \\ U_{21} & U_{22} \end{bmatrix} \begin{bmatrix} a_1^0 \\ a_2^0 \end{bmatrix} . \qquad (2.20)$$

Let us order the site index $\vec{n}$ so that the perturbed sites are located at the top of each subvector a_α^0, and so that moving down the subvector corresponds to moving to sites more remote from the perturbation. With this ordering we see from our previous discussion that U will have the form

$$(2.21)$$

The hatched regions are the many-band regions and the unhatched regions are the one-band regions. We have used the symbols $A_{\alpha\alpha}$ to indicate that the elements of $U_{\alpha\alpha}$ are nonvanishing in the one-band region and indicated the effective vanishing of the elements of $U_{\alpha\alpha'}$ ($\alpha \neq \alpha'$) by zeros. In some of the matrix manipulations which follow it is conceivable that by assuming the elements of $U_{\alpha\alpha'}$, in the one-band region to be strictly zero could lead to ignoring an infinite sum of terms which are individually negligible but which add up to a non-negligible quantity. To simplify the following presentation we ignore this possibility. However, we have carried through the analysis with exponentially decaying terms retained in the elements of $U_{\alpha\alpha'}$ ($\alpha \neq \alpha'$) in the one-band region and have confirmed that the results below are correct.

We proceed now to construct the C matrix.

Since U is unitary we have $U^\dagger U = I$ and in particular

$$\sum_{\beta, \vec{q}} U_{\beta\alpha}^*(\vec{q}, \vec{n}) U_{\beta\alpha}(\vec{q}, \vec{n}') = \delta_{\vec{n}, \vec{n}'} . \qquad (2.22)$$

If we assume that the form (2.21) for U is exact it follows that, for $\vec{n}$ and $\vec{n}'$ in the one-band region, (2.22) reduces to

$$\sum_{\vec{q}} U_{\alpha\alpha}^*(\vec{q}, \vec{n}) U_{\alpha\alpha}(\vec{q}, \vec{n}') = \delta_{\vec{n}, \vec{n}'} . \qquad (2.23)$$

Equation (2.23) is the key to the construction of the C matrix. It shows that we may consider the $N - N_0$ columns of $U_{\alpha\alpha}$ in the one-band region as comprising a set of $N - N_0$ linearly independent vectors in an abstract vector space of N dimensions, where, as before, N is the number of sites in the crystal, and N_0 the number of sites in the many-band region. It follows that we are at liberty to define N_0 additional vectors in this space which are mutually orthonormal and which are orthogonal to the original $N - N_0$ vectors. Thus we may replace the columns of $U_{\alpha\alpha}$ in the many-band region with these new vectors to define a new $N \times N$ matrix. We do this for each diagonal submatrix $U_{\beta\beta}$ and form a matrix with the resulting matrices as diagonal submatrices and zero off-diagonal submatrices. This is the C matrix. It is block diagonal, unitary and is identical to U in the one-band region. For the two-band example, C has the form

$$(2.24)$$

We now show that this form for C leads to the required form for T. First simply from the form of (2.21) and (2.24) we see that the matrix equation $T = C^\dagger U$ will have the form

$$(2.25)$$

We now show that the elements of T in the one band parts of the hatched regions of (2.25) vanish by making use of the unitarity of T. We consider the product $(TT^\dagger)_{\alpha\alpha}(\vec{n}, \vec{n}')$ for $\vec{n}$ and $\vec{n}'$ in the one-band region and break the site sum in the product operation into a sum over many-band and one-band sites,

$$(TT^\dagger)_{\alpha\alpha}(\vec{n}, \vec{n}') = \sum_\beta \left(\sum_{|\vec{m}| < N_0} T_{\alpha\beta}(\vec{n}, \vec{m}) T_{\alpha\beta}^*(\vec{n}', \vec{m}) + \sum_{|\vec{m}| > N_0} T_{\alpha\beta}(\vec{n}, \vec{m}) T_{\alpha\beta}^*(\vec{n}', \vec{m}) \right) = \delta_{\vec{n}, \vec{n}'} . \qquad (2.26)$$

The sum over one-band sites is itself $\delta_{\vec{n},\vec{n}'}$ because of (2.23). Thus

$$\sum_{\beta} \sum_{|\vec{m}|<N_0} T_{\alpha\beta}(\vec{n}, \vec{m}) T^*_{\alpha\beta}(\vec{n}', \vec{m}) = 0, \qquad (2.27)$$

and in particular for $\vec{n} = \vec{n}'$,

$$\sum_{\beta} \sum_{\vec{m}<N_0} |T_{\alpha\beta}(\vec{n}, \vec{m})|^2 = 0. \qquad (2.28)$$

This establishes the desired form for T,

$$T = \begin{array}{|c|c|c|c|} \hline \boxtimes & 0 & \boxtimes & 0 \\ \hline 0 & I & 0 & 0 \\ \hline \boxtimes & 0 & \boxtimes & 0 \\ \hline 0 & 0 & 0 & I \\ \hline \end{array} . \qquad (2.29)$$

From (2.29) we see that the generalized Wannier functions are indeed localized:

(1) An $a_{\alpha,\vec{n}}(\vec{r} - \vec{n})$ in the many-band region will only contain contributions from $a^0_\beta(\vec{r} - \vec{n}')$ (for any β), which are themselves in the many-band region. Thus these $a_{\alpha,\vec{n}}(\vec{r} - \vec{n})$ are localized to the vicinity of the many-band region.

(2) The $a_{\alpha,\vec{n}}(\vec{r} - \vec{n})$ in the one-band region are just the unperturbed $a^0_\alpha(\vec{r} - \vec{n})$.

We emphasize that the many-band region is not, in general, large. It includes only those lattice sites which fall within the range of $V^b(\vec{r})$, extended by the range of the decaying Green's functions from other bands (see Fig. 1).

The many-band region is an upper limit on the extent of individual $a_{\alpha,\vec{n}}(\vec{r} - \vec{n})$ located in the many-band region. Since, apart from orthonormality requirements, the matrix C is as yet unspecified in the many-band region, it is probable that a judicious choice of those vectors of C in the many-band region can improve the localization of the $a_{\alpha,\vec{n}}(\vec{r} - \vec{n})$ about their sites. Finally, it is possible that the flexibility in the C's may permit the contributions from various upper bands to combine destructively so that the many band region is made effectively smaller.

It is interesting at this point to make contact with the asymptotic behavior derived by Kohn and Onffroy[2] in one dimension. We have shown that as n approaches the one-band region, $a_{\alpha,n}(x - n)$ approaches $a^0_\alpha(x - n)$. It remains to determine the functional dependence on n of this approach for $n \to \infty$ (asymptotic region). It is straightforward to show in one dimension that the $G^0_\beta(n - m; E)$ decay exponentially in $n - m$ as $n - m \to \infty$, provided E lies outside the band β. Further, since $G^0_\beta(n - m; E)$ is the Green's function of a periodic band, the exponent is dependent only on the properties of the unperturbed solid.

From Eqs. (2.9) it is clear that the approach to the one band region is governed by the decay of the

$G_\beta(n - m; E)$, $\beta \neq \alpha$. By replacing the zeroes in Eq. (2.21) with exponentially decaying terms, it is tedious but again straightforward to show that $a_{\alpha,n}(x - n)$ approaches $a^0(x - n)$ exponentially in n, and the exponent is that of the most slowly decaying $G_\beta(n - m; E)$, $\beta \neq \alpha$. The form of the decay is consistent with one of the results listed in Ref. 2. That is, asymptotically $a_n(x - n)$ approaches $a^0(x - n)$ exponentially in n, where the exponent is independent of the perturbing potential.

We now extend these results to the case of composite bands. The wave function in the "one-band" region is now given by (2.14) rather than (2.13), and $\phi_\alpha(\vec{q}, \vec{r})$ will in general contain undamped scattered waves in all bands which overlap band α,

$$\phi_\alpha(\vec{q}, \vec{r}) = \sum_{\alpha',\vec{n}'} C_{\alpha\alpha'}(\vec{q}, \vec{n}') a_{\alpha',\vec{n}'}(\vec{r} - \vec{n}'). \qquad (2.30)$$

We illustrate the form of U with a three-band example with the second band overlapping the first,

$$U = \begin{array}{|c|c|c|} \hline \boxtimes A_{11} & \boxtimes A_{12} & 0 \\ \hline \boxtimes A_{21} & \boxtimes A_{22} & 0 \\ \hline 0 & 0 & \boxtimes A_{33} \\ \hline \end{array} . \qquad (2.31)$$

The matrix C is no longer strictly block diagonal. Only those submatrices connecting nonoverlapping bands can be made to vanish. For the three-band example, C has the form

$$C = \begin{array}{|c|c|c|} \hline \boxtimes A_{11} & \boxtimes A_{12} & 0 \\ \hline \boxtimes A_{21} & \boxtimes A_{22} & 0 \\ \hline 0 & 0 & \boxtimes A_{33} \\ \hline \end{array} . \qquad (2.32)$$

We may now go through the same reasoning which led to (2.29) to obtain a similar form for T,

$$T = \begin{array}{|c|c|c|c|c|c|} \hline \boxtimes & 0 & \boxtimes & 0 & \boxtimes & 0 \\ \hline 0 & I & 0 & 0 & 0 & 0 \\ \hline \boxtimes & 0 & \boxtimes & 0 & \boxtimes & 0 \\ \hline 0 & 0 & 0 & I & 0 & 0 \\ \hline \boxtimes & 0 & \boxtimes & 0 & \boxtimes & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & I \\ \hline \end{array} . \qquad (2.33)$$

It follows that the occurrence of composite bands does not qualitatively alter the localization of the generalized Wannier functions.

In Sec. IV a variational procedure for the $a_{\alpha,\vec{n}}(\vec{r} - \vec{n})$ will be described. Equations (2.29) and

(2.33) show that such a method is feasible. That is, one need vary only those *local* functions in the vicinity of the defect from their periodic counterparts, $a_\alpha^0(\vec{r} - \vec{n})$. This is significant in that it greatly lessens the number of parameters necessary in the variation, as will be discussed in Sec. IV.

In Sec. III we calculate by perturbation theory generalized Wannier functions for the valence band of a wide-band-gap insulator. The calculation shows explicitly how the localization of the generalized Wannier functions depends on the band gap, bandwidths, and the number of neighbors retained in the expansion of the energy bands in the direct lattice. The results also suggest a form for trial functions to be used in a variational calculation of the generalized Wannier functions.

III. A WIDE-BAND-GAP EXAMPLE

In this section we calculate via perturbation theory generalized Wannier functions for a two-band model of an insulator. The two bands are the valence and conduction bands denoted by subscripts v and c. We suppose both bands are simple (not overlapped by other bands) and that they are separated by a gap which is large relative to the width of either band. The assumption of simple bands is a convenience: it is straightforward to generalize to complex bands. The large gap, however, is necessary to the perturbation development.

We decompose the localized perturbing potential $V^\flat(\vec{r})$ as follows:

$$V^\flat(\vec{r}) = V_v^\flat(\vec{r}) + V_c^\flat(\vec{r}) + V_{cv}^\flat(\vec{r}), \quad (3.1)$$

where, with

$$V_{\alpha\alpha'}^\flat(\vec{n}, \vec{n}') = \langle a_\alpha^0(\vec{r} - \vec{n}) | V^\flat(\vec{r}) | a_{\alpha'}^0(\vec{r} - \vec{n}') \rangle, \quad (3.2)$$

$$V_v^\flat(\vec{r}) = \sum_{\vec{n}, \vec{n}'} | a_v^0(\vec{r} - \vec{n}) \rangle V_{vv}^\flat(\vec{n}, \vec{n}') \langle a_v^0(\vec{r} - \vec{n}') |,$$

$$V_c^\flat(\vec{r}) = \sum_{\vec{n}, \vec{n}'} | a_c^0(\vec{r} - \vec{n}) \rangle V_{cc}^\flat(\vec{n}, \vec{n}') \langle a_c^0(\vec{r} - \vec{n}') |, \quad (3.3)$$

$$V_{cv}^\flat(\vec{r}) = \sum_{\vec{n}, \vec{n}'} [| a_c^0(\vec{r} - \vec{n}) \rangle V_{cv}^\flat(\vec{n}, \vec{n}') \langle a_v^0(\vec{r} - \vec{n}') |$$
$$+ | a_v^0(\vec{r} - \vec{n}) \rangle V_{vc}^\flat(\vec{n}, \vec{n}') \langle a_c^0(\vec{r} - \vec{n}') |].$$

$V_v^\flat(\vec{r})$ and $V_c^\flat(\vec{r})$ are thus the intraband parts of $V^\flat(\vec{r})$ while $V_{cv}^\flat(\vec{r})$ is the interband part. Because of the localization of $V^\flat(\vec{r})$ the sums in (3.3) are restricted to be in the neighborhood of some lattice site (for a localized perturbation in the bulk) or in the vicinity of a plane of lattice sites (for a surface perturbation).

The Hamiltonian of the perturbed system is then

$$H = H^0 + V_v^\flat + V_c^\flat + V_{cv}^\flat, \quad (3.4)$$

where H^0 is the Hamiltonian of the unperturbed insulator. Our purpose is to obtain approximate wave functions, and from them approximate generalized Wannier functions, for the valence band. To do this we first solve the Schrödinger equation, $\overline{H}\overline{\phi} = \overline{E}\overline{\phi}$, where

$$\overline{H} = H^0 + V_v^\flat + V_c^\flat. \quad (3.5)$$

This problem is not amenable to perturbation theory because of intraband degeneracy. However, it can be solved exactly by the one-band Koster-Slater resolvent technique. We suppose this has been done and denote the one-band Koster-Slater solutions by

$$\overline{\phi}_v(\vec{q}, \vec{r}) = \sum_{\vec{n}} B_{vv}(\vec{q}, \vec{n}) a_v^0(\vec{r} - \vec{n}),$$
$$\overline{\phi}_c(\vec{q}, \vec{r}) = \sum_{\vec{n}} B_{cc}(\vec{q}, \vec{n}) a_c^0(\vec{r} - \vec{n}), \quad (3.6)$$

and the energies corresponding to these solutions by $\overline{E}_v(\vec{q})$ and $\overline{E}_c(\vec{q})$.

Note that $\overline{H}$ may have bound as well as continuum states. We now write H as

$$H = \overline{H} + V_{cv}^\flat, \quad (3.7)$$

and calculate to first order in $V_{cv}^\flat$ the corrections to the valence-band wave functions in (3.6). The perturbation theory expression for the corrected wave function is

$$\phi_v(\vec{q}, \vec{r}) = \overline{\phi}_v(\vec{q}, \vec{r}) + \sum_{\vec{q}'} \frac{\langle \overline{\phi}_c(\vec{q}', \vec{r}) | V_{cv}^\flat | \overline{\phi}_v(\vec{q}, \vec{r}) \rangle}{\overline{E}_v(\vec{q}) - \overline{E}_c(\vec{q}')} \overline{\phi}_c(\vec{q}', \vec{r}). \quad (3.8)$$

We use (3.6) to write this in terms of Wannier functions,

$$\phi_v(\vec{q}, \vec{r}) = \sum_{\vec{n}'} B_{vv}(\vec{q}, \vec{n}') a_v^0(\vec{r} - \vec{n}') + \sum_{\vec{q}'} \sum_{\vec{m}\vec{m}'\vec{n}'} B_{cc}^*(\vec{q}', \vec{m}') B_{vv}(\vec{q}, \vec{m}) B_{cc}(\vec{q}', \vec{n}') \frac{V_{cv}^\flat(\vec{m}', \vec{m})}{\overline{E}_v(\vec{q}) - \overline{E}_c(\vec{q}')} a_c^0(\vec{r} - \vec{n}'). \quad (3.9)$$

To obtain the generalized Wannier functions we must specify the matrix C (actually $C^\dagger$) which transforms the $\phi_v(\vec{q}, \vec{r})$ into the $a_{\vec{\tau}, \vec{n}}(\vec{r} - \vec{n})$. By inspection of (3.8) or (3.9) we see that, in the one-band region [cf. discussion following Eq. (2.13)], $\phi_v(\vec{q}, \vec{r}) = \overline{\phi}_v(\vec{q}, \vec{r})$. This is because the first-order correction term involves only conduction-band wave functions and therefore cannot contribute a scattered wave. This means that in the one-band region we must have $C = B$. The elements of C in the many-band region may be chosen in infinitely many ways provided only that the columns of C be orthogonal [cf. discussion following Eq. (2.23)]. We make the obvious choice and set $C = B$ everywhere (note that B is block diagonal and unitary). This

choice simplifies the analysis which follows and leads to a neat and easily interpreted expression for the generalized Wannier function. Thus we have

$$a_{v,\vec{n}}(\vec{r}-\vec{n}) = \sum_{\vec{q}} B_{vv}^*(\vec{q},\vec{n})\,\phi_v(\vec{q},r)$$

$$= \sum_{\vec{q}}\sum_{\vec{n}'} B_{vv}^*(\vec{q},\vec{n})\,B_{vv}(\vec{q},\vec{n}')\,a_v^0(\vec{r}-\vec{n}')$$

$$+ \sum_{\vec{q}\vec{q}'}\sum_{\vec{m}\vec{m}'\vec{n}'} B_{vv}^*(\vec{q},\vec{n})\,B_{vv}(\vec{q},\vec{m})\,B_{cc}^*(\vec{q}',\vec{m}')\,B_{cc}(\vec{q}',\vec{n}')\,\frac{V_{cv}^\flat(\vec{m}',\vec{m})}{E_v(\vec{q})-E_c(\vec{q}')}\,a_c^0(\vec{r}-\vec{n}')\ . \tag{3.10}$$

In the Appendix we derive an analog of the familiar expansion of $E_\alpha^0(\vec{k})$ of a periodic system in the direct lattice. This allows us to write the energy denominator as

$$\bar{E}_v(\vec{q}) - \bar{E}_c(\vec{q}') = W(\vec{m}',\vec{m}) + \sum_{\vec{l}\neq\vec{m}} \frac{B_{vv}(\vec{q},\vec{l})}{B_{vv}(\vec{q},\vec{m})}\,\bar{\mathcal{E}}_v(\vec{m},\vec{l})$$

$$- \sum_{\vec{l}\neq\vec{m}'} \frac{B_{cc}^*(\vec{q},\vec{l})}{B_{cc}^*(\vec{q}',\vec{m}')}\,\bar{\mathcal{E}}_c(\vec{m}',\vec{l}),$$

$$\tag{3.11}$$

where

$$W(\vec{m}',\vec{m}) = \bar{\mathcal{E}}_v(\vec{m},\vec{m}) - \bar{\mathcal{E}}_c(\vec{m}',\vec{m}'), \tag{3.12}$$

and where the $\bar{\mathcal{E}}_\alpha(\vec{n},\vec{l})$ are

$$\bar{\mathcal{E}}_\alpha(\vec{n},\vec{l}) = \langle a_\alpha^0(\vec{r}-\vec{n})\,|\,\bar{H}\,|\,a_\alpha^0(\vec{r}-\vec{l})\rangle. \tag{3.13}$$

To clarify the meaning of the quantities appearing in (3.11) we show in Fig. 2 how the corresponding quantities for a one-dimensional nearest-neighbor periodic system are related to its band structure. The $W(\vec{m}',\vec{m})$ are analogous to W of Fig. 2, and will typically be at least as large as the unperturbed band gap. No $W(\vec{m}',\vec{m})$ can be smaller than the energy difference between the lowest perturbed conduction-band state and the highest perturbed valence-band state. Further, the $\bar{\mathcal{E}}_\alpha(\vec{n},\vec{n}')$ appearing in the sums on the right in (3.11) determine the bandwidths as do the $\mathcal{E}_\alpha^0(d)$ of Fig. 2. One can show that these sums are never larger than the widths of the perturbed bands and typically will be less. Then, to first order in the ratio of the sum of the bandwidths to the band gap,

$$\frac{1}{E_v(\vec{q})-E_c(\vec{q}')} \simeq \frac{1}{W(\vec{m}',\vec{m})}\left(1 - \frac{1}{W(\vec{m}',\vec{m})}\sum_{\vec{l}\neq\vec{m}}\frac{B_{vv}(\vec{q},\vec{l})}{B_{vv}(\vec{q},\vec{m})}\,\bar{\mathcal{E}}_v(\vec{m},\vec{l})\right.$$

$$\left. + \frac{1}{W(\vec{m}',\vec{m})}\sum_{\vec{l}\neq\vec{m}'}\frac{B_{cc}^*(\vec{q}',\vec{l})}{B_{cc}^*(\vec{q}',\vec{m}')}\,\bar{\mathcal{E}}_c(\vec{m}',\vec{l})\right). \tag{3.14}$$

When this is used in (3.10) and the unitarity of B is used to simplify the resulting expression we obtain

$$a_{v,\vec{n}}(\vec{r}-\vec{n}) \simeq a_v^0(\vec{r}-\vec{n})$$

$$+ \sum_{\vec{n}'}\left[\frac{V_{cv}^\flat(\vec{n}',\vec{n})}{W(\vec{n}',\vec{n})} - \sum_{\vec{m}\neq\vec{n}}\frac{V_{cv}^\flat(\vec{n}',\vec{m})}{W(\vec{n}',\vec{m})}\frac{\bar{\mathcal{E}}_v(\vec{m},\vec{n})}{W(\vec{n}',\vec{m})} + \sum_{\vec{m}'\neq\vec{n}}\frac{V_{cv}^\flat(\vec{m}',\vec{n})}{W(\vec{m}',\vec{n})}\frac{\bar{\mathcal{E}}_c(\vec{m}',\vec{n}')}{W(\vec{m}',\vec{n})}\right]a_c^0(\vec{r}-\vec{n}')\ . \tag{3.15}$$

This expression for the generalized Wannier function shows that, at every site, $a_{v,\vec{n}}(\vec{r}-\vec{n})$ is predominately the valence band Wannier function of the unperturbed system, but for sites in the vicinity of the perturbation there are correction terms involving the conduction band Wannier functions. The first term in brackets mixes an $a_c^0(\vec{r}-\vec{n}')$ into an $a_{v,\vec{n}}(\vec{r}-\vec{n})$ only when sites $\vec{n}$ and $\vec{n}'$ are connected by the interband perturbation. The second and third terms in brackets involve the $\bar{\mathcal{E}}_\alpha(\vec{m},\vec{l})$ in addition to the $V_{cv}^\flat(\vec{m},\vec{l})$. The second term mixes an $a_c^0(\vec{r}-\vec{n}')$ located on a perturbed site into all $a_{v,\vec{n}}(\vec{r}-\vec{n})$ such that the product $V_{cv}^\flat(\vec{n}',\vec{m})\bar{\mathcal{E}}_v(\vec{m},\vec{n})$ is nonvanishing for some $\vec{m}$. The effect of this

term is to mix a_c^0's from perturbed sites into $a_{v,\vec{n}}$'s located within a certain range of the perturbed region. The range is determined by the width and shape of the valence band, that is, by how rapidly the $\bar{\mathcal{E}}_v(\vec{m},\vec{n})$ fall off with increasing $|\vec{m}-\vec{n}|$. The third term mixes into $a_{v,\vec{n}}(\vec{r}-\vec{n})$ located on perturbed sites all $a_c^0(\vec{r}-\vec{n}')$ such that the product $V_{cv}^\flat(\vec{m}',\vec{n})\bar{\mathcal{E}}_c(\vec{m}',\vec{n}')$ is nonvanishing for some $\vec{m}'$. The effect of this term is to mix a_c^0's located within a certain range of the perturbed region into $a_{v,\vec{n}}$'s on perturbed sites. This range is determined by the width and shape of the conduction band. The situation is illustrated in Fig. 3.
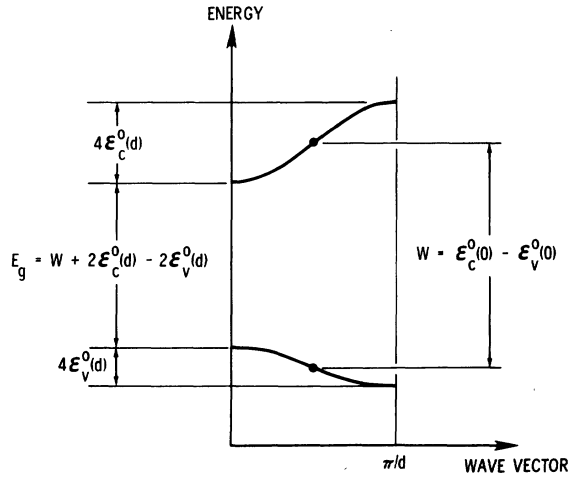
We see, therefore, that for a wide band gap, the

FIG. 2. Illustration of the relation between E_v^0 $(\vec{k})$ and E_c^0 $(\vec{k})$ and the parameters in the direct lattice expansion of these quantities [Eq. (A. 5)]. The illustration is for a one-dimensional periodic system with only nearest-neighbor interactions. E_g is the band gap, and d is the lattice constant.

$a_{v,\vec{n}}(\vec{r} - \vec{n})$ will be strongly localized and will differ from $a_v^0(\vec{r} - \vec{n})$ only very near the perturbation. The narrower the gap, the more influential are the distant sites. Furthermore it need not be that the contributions to an $a_{v,\vec{n}}$ from distant sites serve to delocalize it. They may in fact do the converse in some cases. We expect an attractive potential to accumulate charge and this suggests, at least for an insulator (see Sec. IV), that the $a_{v,n}$ on perturbed sites will be large in the vicinity of their sites. Because the $a_{v,n}$'s must be normalized, this implies that they are small on distant sites.

Equation (3.15) is an example of the transformation $a = Ta^0$ of (2.17). It shows that to first order in the interband perturbation T has the form

$$ T = \begin{array}{|c|c|c|c|} \hline I & 0 & \boxtimes & 0 \\ \hline 0 & I & 0 & 0 \\ \hline \boxtimes & 0 & I & 0 \\ \hline 0 & 0 & 0 & I \\ \hline \end{array} \quad , \qquad (3.16) $$

which is a special case of the general form sketched in (2.29). That is to say, there are no a_v^0's from neighboring sites mixed into $a_{v,\vec{n}}$'s in the vicinity of the perturbation. Note however that were we to carry the perturbation calculation to second order (by calculating also first-order perturbed conduction-band wave functions and then using $\phi_c(\vec{q}, \vec{r})$ and $\phi_v(\vec{q}, \vec{r})$ in a second-order calculation) we would obtain the general form (2.29).

Thus we see that mixing of a_α^0, from other bands into an $a_{\alpha,\vec{n}}$ of a given band is first order in the interband perturbation whereas the mixing of a_α^0 from

the given band but different sites into $a_{\alpha,\vec{n}}$ is second order. This result may prove very helpful in selecting trial functions for the variational calculation of the $a_{\alpha,\vec{n}}$ described in Sec. IV.

IV. CALCULATION OF THE $a_{\alpha,\vec{n}}(\vec{r}-\vec{n})$: WAVE FUNCTIONS, LOCAL DENSITIES OF STATES, AND CHARGE DENSITIES

In Sec. II we demonstrated the localization of the generalized Wannier functions for localized defects in solids. In Sec. III we derived by perturbation theory an explicit expression for the generalized Wannier functions associated with local defects in a wide-band-gap insulator.

The perturbation approach is not appropriate in many cases and it requires knowledge of the periodic Wannier functions for all the bands used in the basis set. Thus a more practical method for obtaining the generalized Wannier functions is required. The lack of an analogous method for periodic systems is the primary reason for the rather academic nature of Wannier functions. It was not until quite recently[13] that a viable variational method for calculating Wannier functions was presented in full generality.

The method we present in the first part of Sec. IV is a straightforward extension to nonperiodic systems of Kohn's[13] variational method.

Once the local functions have been calculated, one must be able to directly obtain wave functions, local densities of states, and charge densities. This is an important point because Koster and Slater[7,8] have shown that a direct algebraic solu-
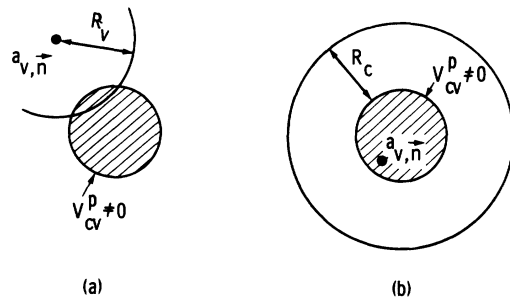

(a)   (b)

FIG. 3. Illustration of the spatial distribution of conduction-band contributions to the $a_{v,\vec{n}}(\vec{r} - \vec{n})$. R_v and R_C are the radii of the most distant sphere of neighbors retained in the expansions of $\bar{E}_v(\vec{q})$ and $\bar{E}_c(\vec{q})$. Fig. 3(a) illustrates the condition for an $a_{v,\vec{n}}$ outside the perturbed region to contain contributions from a_c^0's within the perturbed region, namely, that a portion of the perturbed region falls within a sphere of radius R_v centered at $\vec{n}$. Note that there are no contributions to $a_{v,\vec{n}}$ from a_c^0 outside the perturbed region. Figure 3(b) illustrates the extent of the region over which a_c^0's contribute to $a_{v,\vec{n}}$ located within the perturbed region. The region is a sphere centered in the perturbed region of radius R_c plus the radius of the perturbed region.

tion of finite difference equations is workable only for one-dimensional systems (see especially p. 1173, Ref. 7). This led to their derivation of the Koster-Slater equations (2.9). In the second part of Sec. IV, we derive a set of difference equations for the $C_{\alpha\alpha}(\vec{q}, \vec{n})$ which are the counterparts of the Koster-Slater equations but which are considerably simpler to deal with.

Finally, in the third part of Sec. IV we discuss a powerful means for obtaining the local density of states and charge densities in terms of the generalized Wannier functions. This is a most useful method because it does not require prior determination of the wave functions.

Variational procedure for the local functions

Consider first a simple perturbed band composed of N allowed states, $\phi(\vec{q}, \vec{r})$. The state $\Psi_0(\vec{r})$ of N spinless fermions is given by either of the following Slater determinants:

$$\Psi_0 = (N!)^{-1/2} \det[\phi(\vec{q}, \vec{r}_i)]$$
$$= (N!)^{-1/2} \det[a_{\vec{n}}(\vec{r}_i - \vec{n})], \tag{4.1}$$

since the matrix C relating the ϕ's and a's is unitary. The energy of the band per atom is given by[2]

$$E(\Psi_0) = N^{-1} \langle \Psi_0 | \sum_{i=1}^{N} H_i | \Psi_0 \rangle$$

$$= N^{-1} \sum_{\vec{n}} \langle a_{\vec{n}}(\vec{r} - \vec{n}) | H | a_{\vec{n}}(\vec{r} - \vec{n}) \rangle . \tag{4.2}$$

Note that this energy is specified completely by the generalized Wannier functions. $E(\Psi_0)$ is a stationary value of

$$E(\Psi) = N^{-1} \frac{\langle \Psi | \sum_{i=1}^{N} H_i | \Psi \rangle}{\langle \Psi | \Psi \rangle} . \tag{4.3}$$

Because of the orthogonality requirements between the $a_{\vec{n}}(\vec{r} - \vec{n})$, these functions have oscillatory tails which are difficult to specify in terms of a parametric form. Therefore, following Kohn,[13] we expand the generalized Wannier functions in terms of simple, localized, but nonorthogonal[15,16] trial functions $h_{\vec{n}}(\vec{r} - \vec{n})$ by means of Löwdin's symmetric orthogonalization procedure,[17]

$$a_{\vec{n}}(\vec{r} - \vec{n}) = \sum_{\vec{n}'} F_{\vec{n}',\vec{n}}^{-1/2} h_{\vec{n}}(\vec{r} - \vec{n}'), \tag{4.4}$$

where

$$F_{\vec{n},\vec{n}'} \equiv \langle h_{\vec{n}}(\vec{r} - \vec{n}) | h_{\vec{n}'}(\vec{r} - \vec{n}') \rangle, \tag{4.5}$$

and the matrix $F^{-1/2}$ is the inverse of the square root of F. $F^{-1/2}$ can be constructed via well-known mathematical techniques (e. g., first transform F to diagonal form, replace the diagonal elements with their inverse square roots and then apply the inverse transformation).

Combining Eqs. (4.2) and (4.4), we have

$$E(\Psi) = N^{-1} \sum_{\vec{n},\vec{n}',\vec{n}''} F_{\vec{n},\vec{n}'}^{-1/2} F_{\vec{n}'',\vec{n}}^{-1/2} \langle h_{\vec{n}'}(\vec{r} - \vec{n}') | H | h_{\vec{n}''}(\vec{r} - \vec{n}'') \rangle . \tag{4.6}$$

Equation (4.6) would be difficult to deal with if all the $h_{\vec{n}}(\vec{r} - \vec{n})$ were unknown. However we saw in Secs. II and III that for a localized $V^{p}(\vec{r})$, the $a_{\vec{n}}(\vec{r} - \vec{n})$ deviate significantly from $a^0(\vec{r} - \vec{n})$ only in the vicinity of $V^{p}(\vec{r})$. One can take advantage of this fact by allowing the $h_{\vec{n}}(\vec{r} - \vec{n})$ to deviate from their (known) periodic counterparts only in the vicinity of the perturbation. This will in general mean that only a relatively small number of parameters will need to be varied in the search for the minimum of the energy. For bands other than the lowest there is the problem of restricting the form of the variational functions so that contributions from other bands are excluded. This is accomplished as described for the periodic case in Ref. 13 by making the trial $h_{\alpha,\vec{n}}(\vec{r} - \vec{n})$ orthogonal to the $a_{\alpha',\vec{n}}(\vec{r} - \vec{n})$ of all lower bands. This converts the stationary principle for upper bands to a minimum principle and insures that one cannot get a wrong set of $a_{\alpha,\vec{n}}(\vec{r} - \vec{n})$.

There remains a question of ambiguity. Because of the flexibility in the choice of the $C_{\alpha\alpha}(\vec{q}, \vec{n})$ there are infinitely many correct generalized Wannier functions with, for example, differing degrees of

localization. The problem of finding the most appropriate a's and thus fixing the C-matrix elements is basically a problem of choosing appropriate trial functions $h_\alpha(\vec{r} - \vec{n})$. One would of course like to find the most localized version, since this would lessen the number of neighbors necessary in Eq. (4.6). The kinds of localization to be expected were illustrated in Secs. II and III.

It is straightforward to generalize the preceding to the case of overlapping perturbed bands. The overlapping bands are treated as a composite, and the energy of the composite bands is minimized. If there are p overlapping bands, then the following basis set of generalized Wannier functions is used:

$$a_{\alpha,\vec{n}}(\vec{r} - \vec{n}), \quad \alpha = 1, 2, \ldots, p, \tag{4.7}$$

so that

$$\phi_\alpha(\vec{q}, \vec{r}) = \sum_{\alpha'=1}^{p} \sum_{\vec{n}} C_{\alpha\alpha'}(\vec{q}, \vec{n}) a_{\alpha',\vec{n}}(\vec{r} - \vec{n}). \tag{4.8}$$

The energy per atom is

$$E(\Psi) = N^{-1} \sum_{\alpha=1}^{p} \sum_{\vec{n}} \langle a_{\alpha,\vec{n}}(\vec{r} - \vec{n}) | H | a_{\alpha,\vec{n}}(\vec{r} - \vec{n}) \rangle , \tag{4.9}$$

or

$$E(\Psi) = N^{-1} \sum_{\alpha,\vec{n}} \sum_{\alpha',\vec{n}'} \sum_{\alpha'',\vec{n}''} F^{-1/2}_{\alpha,\vec{n};\alpha',\vec{n}'} F^{-1/2}_{\alpha'',\vec{n}'';\alpha,\vec{n}} \langle h_{\alpha',\vec{n}'}(\vec{r}-\vec{n}') | H | h_{\alpha'',\vec{n}''}(\vec{r}-\vec{n}'') \rangle . \tag{4.10}$$

Many-band wave functions from a one-band resolvent

Once the local functions have been computed, they can be used to obtain the wave functions and the band structure. In the following a Green's-function method is used to derive finite difference equations[2,18] for the $C_{\alpha\alpha}(\vec{q}, \vec{n})$.

The wave functions are given by [see Eq. (2.8) and discussion following it],

$$\phi_\alpha(\vec{q}, \vec{r}) = \psi(\vec{r}) + \int d^3\vec{r}' G^0[\vec{r}-\vec{r}'; E_\alpha(\vec{q})]$$
$$\times V^b(\vec{r}')\phi_\alpha(\vec{q}, \vec{r}'), \tag{4.11}$$

where $\psi(\vec{r})$ is a Bloch wave, $\phi^0_\alpha(\vec{k}, \vec{r})$, for $E_\alpha(\vec{q})$ lying within the energy band α of the periodic

solid. If $E_\alpha(\vec{q})$ lies in a band gap, $\psi(\vec{r}) = 0$.

The Green's function is first written as a spectral resolution,

$$G^0(\vec{r}-\vec{r}'; E) = \lim_{\epsilon \to 0} \sum_{\vec{k}'\alpha'} \frac{\phi^{*0}_{\alpha'}(\vec{k}', \vec{r}')\phi^0_{\alpha'}(\vec{k}', \vec{r})}{E + i\epsilon - E^0_{\alpha'}(\vec{k}')} , \tag{4.12}$$

where ϵ is a small positive number. One can then expand the $\phi^0_\alpha(\vec{k}, \vec{r})$ in terms of Wannier functions and the $\phi_\alpha(\vec{q}, \vec{r})$ in terms of generalized Wannier functions.

Let us first consider the case of simple (non-overlapping) bands. Combining Eqs. (4.11) and (4.12), one obtains

$$\sum_{\vec{n}} C_{\alpha\alpha}(\vec{q}, \vec{n}') a_{\alpha,\vec{n}'}(\vec{r}-\vec{n}') = \theta N^{-1/2} \sum_{\vec{n}'} e^{i\vec{k}\cdot\vec{n}'} a^0_\alpha(r-n')$$

$$+ \lim_{\epsilon \to 0} N^{-1} \int d^3\vec{r}' \sum_{\alpha',\vec{k}'} \sum_{\vec{m},\vec{m}',\vec{n}'} \frac{e^{i\vec{k}'\cdot(\vec{m}-\vec{m}')} a^0_{\alpha'}(\vec{r}-\vec{m}) a^{0*}_{\alpha'}(\vec{r}'-\vec{m}')}{E_\alpha(\vec{q}) + i\epsilon - E^0_{\alpha'}(\vec{k}')}$$

$$\times V^b(\vec{r}') C_{\alpha\alpha}(\vec{q}, \vec{n}') a_{\alpha,\vec{n}'}(\vec{r}'-\vec{n}') , \tag{4.13}$$

where θ is defined in Eq. (2.12). Multiplying by $a^{0*}_\alpha(\vec{r}-\vec{n})$ and integrating, one obtains

$$\sum_{\vec{n}'} \left[T_{\alpha\alpha}(\vec{n}, \vec{n}') - \sum_{\vec{m}'} G^0_\alpha[\vec{n}-\vec{m}'; E_\alpha(\vec{q})] v^b_{\alpha\alpha}(\vec{m}', \vec{n}') \right] C_{\alpha\alpha}(\vec{q}, \vec{n}') = \theta N^{-1/2} e^{i\vec{k}\cdot\vec{n}} , \tag{4.14}$$

where $G^0_\alpha(\vec{n}-\vec{m}'; E)$ is defined in Eq. (2.10), and

$$v^b_{\alpha\alpha}(\vec{m}', \vec{n}) = \langle a^0_\alpha(\vec{r}-\vec{m}') | V^b(\vec{r}) | a_{\alpha,\vec{n}'}(\vec{r}-\vec{n}') \rangle \tag{4.15}$$

and

$$T_{\alpha\alpha}(\vec{n}, \vec{n}') = \langle a^0_\alpha(\vec{r}-\vec{n}) | a_{\alpha,\vec{n}'}(\vec{r}-\vec{n}') \rangle . \tag{4.16}$$

These $T_{\alpha\alpha}(\vec{n}, \vec{n}')$ are just elements in the diagonal sub-blocks of the matrix T of (2.17) or (2.29) which relates the a's to the a^0's.

Equations (4.14) are the analog in the present theory of the Koster-Slater equations (2.9). However they represent a substantial simplification over the Koster-Slater equations because they are one-band equations. No knowledge of the Wannier functions of other bands is required. In the Koster-Slater theory, if one makes the approximation that only m bands contribute to the $\phi_\alpha(\vec{q}, \vec{r})$, one has to deal with mN simultaneous equations. In principle m is infinite and in practice it may be necessary to have m large (Callaway and Hughes[4] found that perhaps eight bands were sufficient to describe vacancies in silicon). One must know the interband and intraband matrix elements, and the Green's functions of all these bands in order to solve the Koster-Slater equations. The difference equations (4.14), on the other hand, comprise a

set of only N simultaneous equations and one need know only the intraband matrix elements and the Green's function for the single band α.

The difference equations (4.14) retain the simplifying feature of the Koster-Slater equations that only a small number of the coefficients have to be solved for simultaneously. To see this let us denote the quantity in brackets in (4.14) by $M_{\alpha\alpha}(\vec{n}, \vec{n}')$, so that

$$\sum_{\vec{n}'} M_{\alpha\alpha}(\vec{n}, \vec{n}') C_{\alpha\alpha}(\vec{q}, \vec{n}') = \theta N^{-1/2} e^{i\vec{k}\cdot\vec{n}} . \tag{4.17}$$

If the sites are ordered as in Sec. II so that those in the vicinity of the perturbation come first, the matrix M of the quantities $M_{\alpha\alpha}(\vec{n}, \vec{n})$ will have the form

$$M = \quad\quad \tag{4.18}$$

Koster and Slater[7] show that the corresponding ma-

trix in their theory has a similar form. Equation (4.18) has the feature that one only need solve the first N' equations simultaneously for the first N' of the $C_{\alpha\alpha}(\vec{q}, \vec{n})$. The remaining values of $C_{\alpha\alpha}(\vec{q}, \vec{n}')$ are given in terms of the first N' by Eqs. (4.14), since $M_{\alpha\alpha}(\vec{n}, \vec{n}') = \delta_{\vec{n}\vec{n}'}$ outside the shaded region. Further, the condition for a bound state is the vanishing of the determinant of the $N' \times N'$ matrix. For the Koster-Slater equations, $N' = m \times N_p$ where N_p is the number of sites on which the perturbing potential has non-negligible matrix elements. For our difference equations, N' is the greater of N_p or N_0, where N_0 is the number of sites in the many-band region (that is, the region in which $T_{\alpha\alpha}(\vec{n}, \vec{n}')$ differs from $\delta_{\vec{n},\vec{n}'}$). N_0 may be less than or greater than N_p, depending on the perturbing potential. The decay of the Green's function (see, e.g., Fig. 1), insures that N_p and

N_0 are nearly equal for a local defect.

Now we generalize to the case of p overlapping bands. In this case[13]

$$\phi_\alpha^0(\vec{k}, \vec{r}) = \sum_{\alpha'=1}^{p} \left[N^{-1/2} \sum_{\vec{n}} a_{\alpha'}^0(\vec{r} - \vec{n}) e^{i\vec{k}\cdot\vec{n}} \right] Q_{\alpha'\alpha}(\vec{k}),$$

(4.19)

where the $Q_{\alpha\alpha'}(\vec{k})$ are elements of a unitary transformation

$$\sum_{\alpha''=1}^{p} Q_{\alpha''\alpha}^*(\vec{k}) Q_{\alpha''\alpha'}(\vec{k}) = \delta_{\alpha\alpha'},$$

(4.20)

whose purpose is to make up linear combinations of the quantities in brackets in (4.19) which are eigenfunctions of the Hamiltonian.

$\phi_\alpha(\vec{q}, \vec{r})$ is given by Eq. (4.8). The difference equations (4.14) become

$$\sum_{\alpha'=1}^{p} \sum_{\vec{n}'} \left[T_{\beta\alpha'}(\vec{n}, \vec{n}') + \sum_{\alpha''=1}^{p} \sum_{m'} G_{\beta\alpha''}^0[\vec{n} - \vec{m}'; E_\alpha(\vec{q})] v_{\alpha''\alpha'}^P(\vec{m}', \vec{n}') \right] C_{\alpha\alpha'}(\vec{q}, \vec{n}')$$

$$= \theta N^{-1/2} e^{i\vec{k}\cdot\vec{n}} Q_{\beta\alpha}(\vec{k}),$$

(4.21)

where

$$G_{\beta\alpha''}^0(\vec{n} - \vec{m}; E) = N^{-1} \lim_{\epsilon \to 0} \sum_{\alpha'=1}^{p} \sum_{\vec{k}'} \frac{e^{i\vec{k}'\cdot(\vec{n}-\vec{m}')}}{E + i\epsilon - E_{\alpha'}^0(\vec{k}')} Q_{\alpha''\alpha'}(\vec{k}') Q_{\beta\alpha'}(\vec{k}'),$$

(4.22)

$$v_{\alpha''\alpha'}^P(\vec{m}', \vec{n}') = \langle a_{\alpha''}^0(\vec{r} - \vec{m}') | V^P(\vec{r}) | a_{\alpha'}(\vec{r} - \vec{n}') \rangle,$$

(4.23)

and $\beta = 1, 2, \ldots, p$.

There are now pN equations to solve.

Charge densities and local densities of states

For a filled band, it follows immediately from Eq. (2.15) and the fact that $C_{\alpha\alpha}(\vec{q}, \vec{n})$ is a unitary transformation that the charge density is given by

$$\rho_\alpha(\vec{r}) = \sum_{\vec{n}} |a_{\alpha,\vec{n}}(\vec{r} - \vec{n})|^2,$$

(4.24)

as shown by Kohn and Onffroy.[22] If there are p overlapping filled bands, the charge density is

$$\rho_p(\vec{r}) = \sum_{\alpha'=1}^{p} \sum_{\vec{n}} |a_{\alpha',\vec{n}}(\vec{r} - \vec{n})|^2.$$

(4.25)

Thus if a perturbation alters the band charge density, then the local functions deviate from the (periodic) Wannier functions of the band. Here we see a powerful result of the method. For filled bands, the charge density as well as the total energy can be obtained directly from a variational calculation, without use of the $E_\alpha(\vec{q})$ and $\phi_\alpha(\vec{q})$. Therefore the analysis of the effect of charge rearrangement on binding energies can be obtained rather simply.

We will see in the following that the local density of states can also be written directly in terms of

the $a_{\alpha,\vec{n}}(\vec{r} - \vec{n})$. The local density of states for band α is

$$n_\alpha(E, \vec{r}) = \sum_{\vec{q}} |\phi_\alpha(\vec{q}, \vec{r})|^2 \delta(E - E_\alpha(\vec{q}))$$

$$= -\pi^{-1} \lim_{\epsilon \to 0} \text{Im} G_\alpha(\vec{r}, \vec{r}, E + i\epsilon),$$

(4.26)

where

$$G_\alpha(\vec{r}, \vec{r}', E) = \sum_{\vec{q}} \frac{\phi_\alpha^*(\vec{q}, \vec{r}') \phi_\alpha(\vec{q}, \vec{r})}{E - E_\alpha(\vec{q})}.$$

(4.27)

Combining Eqs. (2.15), (4.26), and (4.27) and making use of the unitarity of the transformation $C_{\alpha\alpha}(\vec{q}, \vec{n})$, one obtains

$$n_\alpha(E, \vec{r}) = -\pi^{-1} \lim_{\epsilon \to 0} \text{Im} \sum_{\vec{n}} a_{\alpha,\vec{n}}^*(\vec{r} - \vec{n})$$

$$\times \frac{1}{E + i\epsilon - H} a_{\alpha,\vec{n}}(\vec{r} - \vec{n}).$$

(4.28)

The contribution of each local function to the total density of states is

$$n_{\alpha,\vec{n}}(E) = -\pi^{-1} \lim_{\epsilon \to 0} \text{Im} G_{\alpha,\vec{n}}(E + i\epsilon),$$

(4.29)

where

$$G_{\alpha,\vec{n}}(E) = \langle a_{\alpha,\vec{n}}(\vec{r}-\vec{n}) \mid (E-H)^{-1} \mid a_{\alpha,\vec{n}}(\vec{r}-\vec{n}) \rangle$$

$$(4.30)$$

If α is a composite of p overlapping bands, the local density of states is

$$n_\alpha(E,\vec{r}) = -\pi^{-1} \lim_{\epsilon \to 0} \mathrm{Im} \sum_{\alpha'=1}^{p} \sum_{\vec{n}} a_{\alpha',\vec{n}}^{*}(\vec{r}-\vec{n})$$

$$\times \frac{1}{E+i\epsilon-H} \, a_{\alpha',\vec{n}}(\vec{r}-\vec{n}) \qquad (4.31)$$

Equations (4.28) and (4.31) show that the local density of states depends only on the generalized Wannier functions.

Haydock et al.[19,20] have recently provided a means for evaluating terms of the form given in Eq. (4.30). They write Eq. (4.30) as an infinite continued fraction, and use bulk or atomic orbitals for the local functions as an approximation. In the preceding we have seen that an exact local density of states can be obtained if one uses generalized Wannier functions as the local functions. Thus a combination of the variational formalism [Eqs. (4.3) and (4.6) or (4.10)], and their continued-fraction method will allow calculation of self-consistent local densities of states at surfaces and other local defects.

V. DISCUSSION

The generalized-Wannier-function formalism possesses special advantages as a calculation tool for the solid-state defect problem. In Sec. V we wish to reiterate these advantages in the context of certain applications where the formalism is particularly useful.

Consider, for example, the case of a silicon surface. The bulk valence-band Wannier functions can be obtained variationally as described by Kohn[13] using s-type and p-type atomic orbitals for the local functions used in constructing the trial Wannier functions. The introduction of the surface disrupts the tetrahedral bonding of the bulk silicon leading to distortion of the surface orbitals relative to those of the bulk. These distortions can be accounted for by parametrizing the local functions to allow for changes in bond angles and strengths, or by including components of different symmetry in the local functions. These new symmetries correspond to the introduction of Wannier functions of other bands into the generalized Wannier functions of the valence band, as suggested by the analysis of Secs. II and III. By minimizing the band energy as a function of the parameters in these local functions at the surface as described in Sec. IV one can determine quantitatively the distortion in the electronic structure at the surface. Furthermore the atomic positions in the surface can be found by using the ion core positions in the surface region as parameters in the varia-

tional calculation. Thus surface changes such as reconstruction can in principle be predicted theoretically.

A related but more difficult problem is the calculation of potential surfaces for the interaction between molecules and solid surfaces. With these potential surfaces one can in principle determine reaction rates[21] on catalytic surfaces. When computing these potential surfaces as a function of the coordinates of the atoms of the molecules and of the solid surface, one does not even have at his disposal the symmetries of the bare surface because the translational symmetry of the crystal parallel to the surface is destroyed by the chemisorbed molecule. Thus a variational approach like that proposed in this paper is the only viable means of determining these potential surfaces at present. The fact that the generalized Wannier functions need change only in the vicinity of a defect means that only those Wannier functions associated with the molecule and with nearby surface sites need change upon chemisorption. Thus only a small number of local functions need be parametrized and varied in Eqs. (4.6) or (4.10). This makes the calculation tractable.

The fact that the formalism allows for changes in the local functions (and hence the charge density) in the vicinity of a defect means that charge rearrangements can be accounted for. It is well known that when a surface is cleaved there is a charge rearrangement which produces the surface dipole barrier. There is an analogous charge rearrangement around bulk defects. Faulkner[3] has concluded that isoelectronic trap calculations must include polarization of the host crystal. In the context of this discussion this means that the generalized Wannier functions must change in the vicinity of the trap.

That the charge density changes in the vicinity of a defect means that there is a self-consistency problem. Self-consistency has been shown to be of critical importance in the electronic structure of bare surfaces[22] and surfaces with chemisorbed impurities.[23] The generalized-Wannier-function formalism is particularly suited to this self-consistency problem in insulators and semiconductors. It was stated in Sec. IV, but deserves reemphasis, that in these materials with filled bands the charge density and thus the potential is determined entirely by the generalized Wannier functions. Thus calculations may be done self-consistently without ever having to calculate wave functions.

Next, we discuss two consequences of the generalized-Wannier-function formalism which bear on the bond concept in solids. A feature of localized functions made orthogonal by applying the symmetric orthogonalization procedure has been pointed out by Halpern.[24] This is that they decay

smoothly in regions where there are no other functions to which they have to be orthogonal. Thus in a vacancy problem, the generalized Wannier functions will have no unphysical oscillations on the vacancy site whereas the Wannier functions of the host crystal will have such oscillations. Similarly, the generalized Wannier functions in surface layers will have no oscillations in the vacuum. The smoothly decaying tails of these generalized Wannier functions which protrude into the vacuum are the dangling bonds of surface chemistry.

Finally, we mention how the generalized-Wannier-function formalism supports the covalent bond theory of solids. Covalent bond theory presumes that cohesive energies of covalent solids are the sum of the energies of the covalent bonds. That is, it is assumed that the bonds have a rather independent character. Suppose one introduces a perturbing potential $V^{p}(\vec{r})$ which is localized to site $\tilde{n}'$. If the band gap is wide enough relative to the bandwidths, we know from Secs. II and III that the only Wannier function that will be changed significantly is the one centered on site $\tilde{n}'$. From this one can see that the only term in Eq. (4.2) for the total energy which is changed significantly by the perturbation is the one associated with site $\tilde{n}'$. Thus the "bonds" are approximately independent. The wider the band gap, the more valid the concept.

In a recent paper,[25] we have presented calculations on a model insulator surface. Surface- and continuum-state wave functions, charge densities, and surface-state energies are obtained from the generalized Wannier functions. The simplicity and accuracy of the method are demonstrated.

ACKNOWLEDGMENTS

APPENDIX

We derive an analog for nonperiodic systems of the expansion of $E_\alpha^0(\vec{k})$ of a periodic system in the direct lattice. It stems from Eq. (2.6) of the text:

$$E_\alpha(\vec{q})\phi_\alpha(\vec{q},\vec{r}) = H\phi_\alpha(\vec{q},\vec{r}). \tag{A1}$$

Making use of Eq. (2.15), we have

$$E_\alpha(\vec{q})\phi_\alpha(\vec{q},\vec{r}) = \sum_{\tilde{n}} C_{\alpha\alpha}(\vec{q},\tilde{n}) H a_{\alpha,\tilde{n}}(\vec{r}-\tilde{n}). \tag{A2}$$

Taking the inner product of both sides with the generalized Wannier function $a_{\alpha,\tilde{n}'}(\vec{r}-\tilde{n}')$ yields

$$E_\alpha(\vec{q}) = \sum_{\tilde{n}} \frac{C_{\alpha\alpha}(\vec{q},\tilde{n})}{C_{\alpha\alpha}(\vec{q},\tilde{n}')} \mathscr{E}_\alpha(\tilde{n}',\tilde{n}), \tag{A3}$$

where it is assumed that $C_{\alpha\alpha}(\vec{q},\tilde{n}')\neq0$ and where

$$\mathscr{E}_\alpha(\tilde{n}',\tilde{n}) = \langle a_{\alpha,\tilde{n}'}(\vec{r}-\tilde{n}')|H|a_{\alpha,\tilde{n}}(\vec{r}-\tilde{n})\rangle. \tag{A4}$$

This can be compared with the familiar expansion

$$E_\alpha^0(\vec{k}) = \sum_{\tilde{n}} e^{i\vec{k}\cdot(\tilde{n}-\tilde{n}')} \mathscr{E}_\alpha^0(\tilde{n},\tilde{n})$$

$$= \sum_{\vec{R}} e^{i\vec{k}\cdot\vec{R}} \mathscr{E}_\alpha^0(\vec{R}), \tag{A5}$$

where $\vec{R}=\tilde{n}'-\tilde{n}$, and

$$\mathscr{E}_\alpha^0(\tilde{n}',\tilde{n}) = \mathscr{E}_\alpha^0(\tilde{n}'-\tilde{n}) = \langle a_\alpha^0(\vec{r}-\tilde{n}')|H^0|a_\alpha^0(\vec{r}-\tilde{n})\rangle. \tag{A6}$$

Note that the $E_\alpha(\vec{q})$ in Eq. (A3) may be either the energy eigenvalues of bound or continuum states, whereas the $E_\alpha^0(\vec{k})$ include only the periodic continuum states.

[1]W. Kohn, in Solid State Physics, edited by Ehrenreich, Seitz, and Turnbull (Academic, New York, 1957), Vol. 5, p. 257.
[2]W. Kohn and J. Onffroy, Bull. Am. Phys. Soc. 18, 456 (1973); Phys. Rev. B 8, 2485 (1973). See also J. J. Rehr and W. Kohn (unpublished).
[3]R. A. Faulkner, Phys. Rev. 175, 991 (1968).
[4]J. Callaway and A. J. Hughes, Phys. Rev. 156, 860 (1967); Phys. Rev. 164, 1043 (1967); N. Parada, Phys. Rev. B 3, 2042 (1971).
[5]G. E. Becker and H. D. Hagstrum, Surf. Sci. 30, 505 (1972).
[6]D. E. Eastman, Solid State Commun. 10, 933 (1972).
[7]G. F. Koster and J. C. Slater, Phys. Rev. 95, 1167(1954).
[8]G. F. Koster, Phys. Rev. 95, 1436 (1954).
[9]G. F. Koster and J. C. Slater, Phys. Rev. 96, 1208 (1954).
[10]J. Calais and C. Ribbing, Phys. Rev. B 4, 376 (1971).
[11]J. Des Cloizeaux, Phys. Rev. 129, 554 (1962); Phys. Rev. 135, A698 (1964); Phys. Rev. 135, A685 (1964).
[12]J. Callaway, J. Math. Phys. 5, 783 (1964).
[13]W. Kohn, Phys. Rev. B 1, 4388 (1973).
[14]W. Kohn, Phys. Rev. 115, 809 (1959).
[15]Philip W. Anderson, Phys. Rev. Lett. 21, 13 (1968).
[16]Philip W. Anderson, Phys. Rev. 181, 25 (1969); J. D. Weeks, P. W. Anderson, and A. G. H. Davidson, J. Chem. Phys. 58, 1388 (1973).
[17]P. O. Löwdin, J. Chem. Phys. 18, 365 (1950).
[18]An equally useful set of difference equations was derived in Ref. 2 for the system considered here.
[19]R. Haydock, V. Heine, M. J. Kelly, and J. B. Pendry, Phys. Rev. Lett. 29, 868 (1972).
[20]R. Haydock, V. Heine, and M. J. Kelly, J. Phys. C 5, 2845 (1972).
[21]T. L. Hill An Introduction to Statistical Thermodynamics, Chapter 11, (Addison-Wesley, Reading, Mass, 1960).
[22]J. R. Smith, Phys. Rev. 181, 522 (1969); N. D. Lang and W. Kohn, Phys. Rev. B 1, 4555 (1970); J. A. Appelbaum and D. R. Hamann, Phys. Rev. B 6, 2166(1972).
[23]S. C. Ying, J. R. Smith, and W. Kohn, J. Vac. Sci. and Tech. 9, 575 (1972); J. R. Smith, S. C. Ying, and W. Kohn, Phys. Rev. Lett. 30, 610 (1973).
[24]V. Halpern, J. Phys. C 3, 1900 (1970).
[25]J. R. Smith and J. G. Gay, Phys. Rev. Lett. 32, 774 (1974); Bull. Am. Phys. Soc. 18, 334 (1973).