Localized Defects in Semiconductors*

JOSEPH CALLAWAY
Department of Physics, University of California, Riverside, California
AND
A. JAMES HUGHES
Applied Research Laboratory, Aeronutronic Division of Philco-Ford Corporation, Newport Beach, California
(Received 10 November 1966)

A method of calculating the energies of bound states associated with imperfections described by short-range potentials in semiconductors is presented and applied. The method is based on solid-state scattering theory and involves expansion of physical quantities in terms of the Wannier functions for a many-band system. The properties of Bloch functions and Wannier functions are investigated to enable the numerical computation of matrix elements. Application is made to the neutral vacancy in silicon. A pseudopotential is employed to represent the change in the crystal potential produced by the vacancy. It is found that the vacancy produces a localized state in which low-lying energy bands and interband couplings are important, in contrast to the shallow donor and acceptor states which have been studied in the past.

I. INTRODUCTION

AN imperfection or an impurity in a semiconductor may produce a state with an energy within the band gap of the crystal. The states associated with many common dopants—the shallow donors and acceptors—have been well understood for many years and are amenable to theoretical treatment.[1] The potential experienced by an electron at large distances from the center is just the electrostatic potential of the excess (or deficit) ionic charge screened by the static dielectric constant of the perfect crystal. If the defect state lies close to a nondegenerate band, it is possible to write a simple Schrödinger equation for the envelope of the wave function in which the effective mass for the nearby band appears.

Some impurity systems, for instance transition metals in germanium, have long been known to produce levels lying farther within the band gap; these levels are not given correctly by the simple theory mentioned above.[2] Other interesting examples are the so-called isoelectronic traps.[3] In the case of certain isoelectronic substitutional atoms of which nitrogen and bismuth in gallium phosphide is an example, discrete levels are produced within the gap. This situation is altogether inconsistent with the usual simple picture of the defect state due to the lack of a long-range Coulomb potential. There is no adequate theoretical treatment of these systems.

Defect states which are produced in experiments concerning radiation damage do not conform to the predictions of the simple theory either.[4,5] Of particular interest to us is the isolated vacancy which can apparently exist in a number of charge states. It is our intention to present in this paper a general method for dealing with localized states. It treats those potentials which, although strong, are of short range as compared to the effective-mass theory which treats weak long-range potentials. We will apply this method to the isolated, neutral vacancy in silicon. Our interest in the vacancy is due to the circumstance that, if lattice relaxation is neglected, it is possible to represent the change in the crystal potential produced by the vacancy in a simple way by a pseudopotential.

There have been several previous studies of energy levels associated with vacancies in materials with the diamond structure.[6–12] One of these calculations, that of Brennemann[11] is similar in some respects to our work. Brennemann empolys a t-matrix method, which is close to ours in spirit, but makes the drastic initial assumption of using free-electron wave functions in zero order. This approach might be more appropriate for a polyvalent metal. The semiconducting character of the host crystal evolves in higher order through multiple-scattering theory. Our method involves an extension of the Koster-Slater approach[13] to a multiband system. The general theory has been discussed elsewhere in connection with the theory of scattering in solids.[14] We expand functions relating to the defect in terms of Wannier functions for the perfect crystal. Similar methods have been applied to other defect systems. A few references are

* The research reported in this paper was sponsored in part by the U. S. Air Force Cambridge Research Laboratories, Office of Aerospace Research, under Contract No. AF-ESD/AF19(628)-5660.

[1] W. Kohn, in *Solid State Physics*, edited by F. Seitz and D. Turnbull (Academic Press Inc., New York, 1957), Vol. 5, p. 258.
[2] J. A. Burton, Physica **20**, 845 (1954).
[3] D. G. Thomas, J. Phys. Soc. Japan Suppl. **21**, (1966).
[4] G. D. Watkins, in *Proceedings of the Symposium on Radiation Damage in Semiconductors* (Dunod Cie., Paris, 1965), p. 97.
[5] J. W. Corbett, *Electron Radiation Damage in Semiconductors and Metals* (Academic Press Inc., New York, 1966).
[6] H. M. James and K. Lark-Horovitz, Z. Physik Chem. (Leipzig) **198**, 107 (1951).
[7] E. I. Blount, Phys. Rev. **113**, 995 (1959).
[8] M. C. M. O'Brien and M. H. L. Pryce, in *Proceedings of the Conference on Defects in Crystalline Solids, Bristol, 1954* (The Physical Society, London, 1955), p. 88.
[9] R. A. Coulson and M. J. Kearsley, Proc. Roy. Soc. (London) **A241**, 433 (1957).
[10] T. Yamaguchi, J. Phys. Soc. Japan **17**, 1359 (1962); **18**, 923 (1963).
[11] K. H. Brennemann, Phys. Rev. **137**, A1497 (1965).
[12] A. M. Stoneham, Proc. Phys. Soc. (London) **88**, 135 (1966).
[13] G. F. Koster and J. C. Slater, Phys. Rev. **96**, 1208 (1954).
[14] J. Callaway, J. Math. Phys. **5**, 783 (1964).

given below.[15–17] In contrast to previous calculations, however, we have obtained matrix elements of physical quantities on the basis of Wannier functions for a many-band system by numerical calculation. In order to make such calculations possible, we have investigated the symmetry properties of Bloch functions and Wannier functions. Since the calculations require a very substantial computing effort, we have been forced to adopt a pseudopotential approach. The wave functions and energy levels for the unperturbed system are obtained from a pseudopotential band calculation and the vacancy is represented by the (negative of) an atomic pseudopotential.

Rather reasonable results are obtained. We find that Wannier functions can be used in practical computations as well as in formal arguments. It is now possible to study systems, like the vacancy, for which interband couplings are important. The plan of this paper is as follows: The remainder of the introduction contains a qualitative discussion of experimental information concerning the levels associated with vacancies in silicon. The formal machinery with which defect levels are located in a multi-band system is presented in Sec. II. Section III contains a discussion of the properties of Bloch functions and Wannier functions that are required in the calculation. In Sec. IV, we describe the manner in which the pseudopotential is employed. The band structure of the perfect silicon crystal is discussed in Sec. V. with regard to the identification of bands and the construction of Wannier functions. Application to the neutral silicon vacancy is made in Sec. VI. In Sec. VII, our results for the neutral vacancy in silicon are presented, and our conclusions are summarized in Sec. VIII.

Much of the experimental information concerning vacancies in silicon is based on electron-paramagnetic-resonance (EPR) measurements. A detailed discussion can be found in Refs. 4 and 5. Unfortunately, defects which involve an even number of spin-paired electrons—and this includes the neutral vacancy—cannot be studied directly by this technique.

EPR studies of p-type silicon irradiated at low temperatures with 1.5-MeV electrons have revealed spectra associated with two different charge states of the isolated lattice vacancy. In Watkin's notation[4] these two defects and their associated EPR spectra are labeled Si-$G1$ and Si-$G2$. They have been identified as $+1$ and -1 charge states, respectively, of the single vacancy. For the Si-$G1$ (V^+) spectra three defect electrons are to be localized at the vacancy. In the molecular orbital picture, two of the electrons go into the same orbital with spins paired. The resonant third electron is considered to be in a localized orbital associated with all four of the nearest-neighbor silicon atoms consistent with a four-silicon hyperfine interaction. The location of the V^+ state is not certain. Since neither it nor the V^- state represents the stable charge state when the Fermi level is pinned to shallow acceptors near the valence band, it is believed that the V^+ state must lie lower than $E_v+0.05$ eV, where E_v is the valence-band energy. It seems likely at the present time that this metastable charge state may be a resonance or scattering state lying inside the valence band.[18]

The energy level for the neutral vacancy V^0 is somewhat more certain. This charge state is stable when the Fermi level is locked on the shallow acceptors at approximately $E_v+0.05$ eV. The energy level for V^0 is believed to lie somewhere below $E_v+0.05$ eV and may be quite close to the valence bands. Attempts to move the Fermi level closer to the valence band by heavier doping in order to further delimit the energy level unfortunately would involve deterioration of the EPR spectra due to the increased doping.

Further electron irradiation raises the Fermi level in the forbidden gap. When the Fermi level rises to approximately 0.25 eV above the valence band the Si-$G2$ (V^-) spectra appear as the stable charge state of the single vacancy. We therefore tentatively consider the V^- defect to be a deep acceptor located at roughly 0.25 eV above the valence band. One additional charge charge of the single vacancy may be inferred. Neither the V^+ nor V^- spectra are observed in low resistivity in n-type silicon. It is presumed that in this case the V^{--} charge state is the stable charge. Watkins indicates this schematically as lying somewhere below the conduction band.

Estimates for the migration energy of both V^0 and V^- are available. The vacancy migration energy in p-type silicon is about 0.33 eV and disappears in a 15-min anneal at about 170°K. From the energy-level structure and a consideration of the Fermi level as related to the stable charge state, it was deduced that this migration energy was that of the neutral vacancy. The V^- spectra disappeared after a 15-min anneal at about 60°K and suggested an activation energy for V^- of less than 0.16 eV.

II. GENERAL THEORY

The computation of the properties of localized defect states in semiconductors can be based on the general methods of solid-state scattering theory.[14] In this section we will review the formal theory which is applied in our calculation.

Let the Hamiltonian of the perfect crystal including the periodic potential be denoted by H_0 and the change crystal potential produced by the introduction of the defect be denoted by V. We define a Green's function G, by

$$G=(E-H_0)^{-1}. \tag{2.1}$$

[15] A. M. Clogston, Phys. Rev. **136**, A1417 (1964).
[16] R. E. Turner and D. A. Goodings, Proc. Phys. Soc. (London) **86**, 87 (1965).
[17] A. Seeger, in *Metallic Solid Solution*, edited by J. Friedl and A. Guinier (W. A. Benjamin, Inc., New York, 1963), p. VII-1.

[18] We are grateful to Dr. James W. Corbett for a very helpful discussion.

The quantity E which appears here may be taken to be a real energy as long as we are concerned with bound states: that is, with states whose energies do not coincide with any energy in the spectrum of H_0. In a scattering problem in which continuum states are of interest, E must be allowed to have an infinitesimal imaginary part. Consider the determinantal function $D(E)$ defined by

$$D(E)=\det[I-GV], \tag{2.2}$$

where I is a unit operator. For every real E_0 for which the determinantal function vanishes, i.e.,

$$D(E_0)=0, \tag{2.3}$$

the operator $I-GV$ is singular. Hence, there must exist some vector $|\phi\rangle$ which is annihilated by the operator $I-G(E_0)V$;

$$[1/(E_0-H_0)]V|\phi\rangle=|\phi\rangle, \tag{2.4}$$

which implies that $|\phi\rangle$ is a solution of the Schrödinger equation for energy E_0:

$$(H_0+V)|\phi\rangle=E_0|\phi\rangle. \tag{2.5}$$

Alternatively, the problem of solving the Schrödinger equation for the energy of a defect state may be approached by looking for the roots of $D(E)$.

It is necessary to have a suitable basis for the representations of operators. For localized defect problems the Wannier functions are most appropriate. In this section, we will neglect complications caused by band degeneracies, and treat the Wannier functions as if they were derived from Bloch functions belonging to simple noncrossing bands. Let the Bloch function for band n and wave vector $\mathbf{k}$ be denoted by $\Psi_n(\mathbf{k},\mathbf{r})$. It is an eigenfunction of H_0.

$$H_0\Psi_n(\mathbf{k},\mathbf{r})=E_n(\mathbf{k})\Psi_n(\mathbf{k},\mathbf{r}), \tag{2.6}$$

in which $E_n(\mathbf{k})$ is the energy-band function. Then the Wannier function is defined by

$$a_n(\mathbf{r}-\mathbf{R}_\mu)=\frac{\Omega^{1/2}}{(2\pi)^{3/2}}\int d^3k\, e^{-i\mathbf{k}\cdot\mathbf{R}_\mu}\Psi_n(\mathbf{k},\mathbf{r}), \tag{2.7}$$

in which $\mathbf{R}_\mu$ is a lattice vector, Ω is the volume of the unit cell, and the integration volume is the Brillouin zone. As is well known, the Wannier functions are orthonormal:

$$\int a_n^*(\mathbf{r}-\mathbf{R}_\mu)a_l(\mathbf{r}-\mathbf{R}_\nu)d^3r=\delta_{nl}\delta_{\mu\nu}$$

(the integral includes all space). The Wannier function $a_n(\mathbf{r}-\mathbf{R}_\mu)$ is, approximately, localized around site $\mathbf{R}_\mu$.[19–21]

[19] W. Kohn, Phys. Rev. **115**, 809 (1959).
[20] E. I. Blount, in *Solid State Physics*, edited by F. Seitz and D. Turnbull (Academic Press Inc., New York, 1962), Vol. 13.
[21] J. des Cloizeaux, Phys. Rev. **135**, A685 (1964); **135**, A698 (1964).

There is a double infinity of Wannier functions, since both the band index and the site index must be considered. Therefore, it is possible to evaluate the determinant only approximately in contrast to the case of localized spin-wave modes in a Heisenberg ferromagnet, for example. We must replace V by a finite matrix. It is therefore necessary to investigate the convergence of the results, and this will be described subsequently.

The elements of the Green's function on the Wannier basis may be expressed as

$$(n\mu|G|l\nu)=\int a_n^*(\mathbf{r}-\mathbf{R}_\mu)\frac{1}{E-H_0}a_l(\mathbf{r}-\mathbf{R}_\nu)d^3r$$
$$=\frac{\delta_{nl}\Omega}{(2\pi)^3}\int d^3k\frac{e^{i\mathbf{k}\cdot(\mathbf{R}_\mu-\mathbf{R}_\nu)}}{E-E_n(\mathbf{k})}. \tag{2.8}$$

These integrals may be calculated in a straightforward way if the band structure is known. Since we are concerned here with states in the band gap the denominator in Eq. (2.8) will not vanish.

In order for the present method to be a useful one it is necessary to approximate the defect potential V by a matrix which, on the Wannier basis, has only a finite number of nonzero elements. Let us suppose that we decide to consider elements of the potential involving n_b bands and n_s sites. Then the nonzero portion of the V matrix is of dimension $N\times N$, where $N=n_bn_s$, before the factorization of D due to symmetry is considered. For instance, in the present problem in which 8 bands and 10 sites are included, the matrix V is 80×80. The Green's-function matrix has, however, an infinite number of nonzero elements. Fortunately it is necessary to consider only an $N\times N$ portion of G in evaluating $D(E)$. To see this, it is necessary only to write G and V in block form. Let us denote the nonzero portion of V by V_{NN}, and the corresponding diagonal portion of G by G_{NN}. The remaining portions of G will be labeled G_{NT}, G_{TN}, and G_{TT}. Thus,

$$G=\begin{pmatrix}G_{NN} & G_{NT}\\ G_{TN} & G_{TT}\end{pmatrix}. \tag{2.9}$$

Then we find

$$I-GV=\begin{pmatrix}I_{NN}-G_{NN}V_{NN} & 0\\ -G_{TN}V_{NN} & I_{TT}\end{pmatrix}. \tag{2.10}$$

It is now easy to see, by expanding the determinant $D(E)$ according to the minors of the columns in the right-hand portion, that

$$D(E)=\det[I\text{-}GV]=\det[I_{NN}-G_{NN}V_{NN}]. \tag{2.11}$$

For computational purposes, the matrix $I_{NN}-G_{NN}V_{NN}$ has the disadvantage that it is not Hermitian, even though both G_{NN} and V_{NN} are Hermitian. It is convenient to rewrite Eq. (2.11) as

$$D(E)=\det(G_{NN})\det(G_{NN}^{-1}-V_{NN}), \tag{2.12}$$

in which G_{NN}^{-1} is the inverse of the submatrix G_{NN} (and this not the same as the NN portion of the inverse matrix G^{-1}). In the present calculation, we are interested only in bound states, so we can therefore restrict our considerations to the determinant

$$\bar{D}(E)=\det[G_{NN}^{-1}-V_{NN}]. \tag{2.13}$$

We see from the preceding argument that if we find an energy E_0 such that

$$\bar{D}(E_0)=0, \tag{2.14}$$

then there is a solution of the full Schrödinger equation with energy E_0. Also, obviously, the matrix $G_{NN}^{-1}-V_{NN}$ has a zero eigenvalue for $E=E_0$. Let us consider the corresponding eigenvector, which we will now call ϕ_N. It satisfied

$$G_{NN}^{-1}\phi_N=V_{NN}\phi_N. \tag{2.15}$$

ϕ_N is not the complete eigenvector of the full Hamiltonian, since it has only N rows, but the required eigenvector of H can be found directly from ϕ_N. Let us write $\phi=\{\phi_N,\phi_T\}$. Then $(I-GV)\phi=0$ yields

$$\phi_T=G_{TN}V_{NN}\phi_N. \tag{2.16}$$

Therefore, to determine the defect wave function, we must compute not only the elements G_{NN}, but the G_{TN} as well.

Computations with this method may be substantially simpler when the defect potential has some symmetry if appropriately symmetrized functions are introduced. Then one finds that the determinant $D(E)$ factors into a product of terms coming from the irreducible representations of the symmetry group. The subdeterminants involved may be substantially smaller than the original $D(E)$. To investigate how this symmetry analysis may be applied in the present problem, we need some results concerning the symmetries of Bloch functions and Wannier functions, which are contained in Sec. III.

III. SYMMETRY PROPERTIES OF BLOCH FUNCTIONS AND WANNIER FUNCTIONS

In this section we will investigate some of the symmetry properties of Bloch functions and Wannier functions. We will consider only *isolated* energy bands; that is energy bands which do not touch or cross other bands. This situation can be analyzed quite simply and rigorously.

We will denote an operation in the space group of the crystal by $\{\alpha|\mathbf{t}_\alpha+\mathbf{R}_\mu\}$. In this notation α is a rotation or reflection, $\mathbf{t}_\alpha$ is a nonprimative translation associated with α, and $\mathbf{R}_\mu$ is a lattice translation. That is, acting on an arbitrary vector $\mathbf{r}$, we have

$$\{\alpha|\mathbf{t}_\alpha+\mathbf{R}_\mu\}\mathbf{r}=\mathbf{r}'=\alpha\mathbf{r}+\mathbf{t}_\alpha+\mathbf{R}_\mu. \tag{3.1}$$

We recall that the law of multiplications is

$$\{\beta|\mathbf{t}'\}\{\alpha|\mathbf{t}\}=\{\beta\alpha|\beta\mathbf{t}+\mathbf{t}'\}, \tag{3.2}$$

where now we use $\mathbf{t}$ for $\mathbf{t}_\alpha+\mathbf{R}_\mu$, and that the inverse operator is given by

$$\{\alpha|\mathbf{t}\}^{-1}=\{\alpha^{-1}|-\alpha^{-1}\mathbf{t}\}. \tag{3.3}$$

In order to determine the symmetry properties of Wannier functions, we must begin by studying the effect of a general space-group operation on a Bloch function. This has, of course, been discussed by other authors.[22] However, we require the results in a more specific form. In particular, since Wannier functions may be altered by a change in phase of the Bloch functions from which they are constructed, it is necessary to specify phase factors brought in by transformations completely. Let us consider the Bloch function $\Psi_n(\mathbf{k},\mathbf{r})$ for a state of wave vector $\mathbf{k}$ in band n. It satisfies Bloch's theorem in the form

$$\{\epsilon|\mathbf{R}_\mu\}\Psi_n(\mathbf{k},\mathbf{r})=\Psi_n(\mathbf{k},\mathbf{r}-\mathbf{R}_\mu)=e^{-i\mathbf{k}\cdot\mathbf{R}_\mu}\Psi_n(\mathbf{k},\mathbf{r}), \tag{3.4}$$

where $\{\epsilon|\mathbf{R}_\mu\}$ is a pure lattice translation. Now consider the more general operator $\{\alpha|\mathbf{t}_\alpha+\mathbf{R}_\mu\}$:

$$\{\alpha|\mathbf{t}_\alpha+\mathbf{R}_\mu\}\Psi_n(\mathbf{k},\mathbf{r})=\Psi_n(\mathbf{k},\alpha^{-1}(\mathbf{r}-\mathbf{t}_\alpha-\mathbf{R}_\mu)).$$

But $\alpha^{-1}\mathbf{R}_\mu$ is a lattice translation if $\mathbf{R}_\mu$ is, so we have

$$\begin{aligned}\{\alpha|\mathbf{t}_\alpha+\mathbf{R}_\mu\}\Psi_n(\mathbf{k},\mathbf{r})&=\{\epsilon|\mathbf{R}_\mu\}\{\alpha|\mathbf{t}_\alpha\}\Psi_n(\mathbf{k},\mathbf{r})\\&=e^{-i\alpha\mathbf{k}\cdot\mathbf{R}_\mu}\Psi_n(\mathbf{k},\alpha^{-1}(\mathbf{r}-\mathbf{t}_\alpha)).\end{aligned} \tag{3.5}$$

The exponential factor in Eq. (3.5) tells us, after comparison with Eq. (3.4), that the function on the left satisfies Bloch's theorem for a state of wave vector $\alpha\mathbf{k}$. We may now drop the $\mathbf{R}_\mu$ and consider just the operation $\{\alpha|\mathbf{t}_\alpha\}$. This operation, being a space-group operation leaves the crystal lattice unchanged, and this must also leave the charge density of the electron system unchanged. We see from Eq. (3.5) that the operation must interchange members of the star of $\mathbf{k}$, but the result must be a wave function belonging to the star. Since we have assumed we are dealing with a simple band, there is only one wave function for each $\mathbf{k}$ in the star. Thus, we must have

$$\{\alpha|\mathbf{t}_\alpha\}\Psi_n(\mathbf{k},\mathbf{r})=e^{i\Theta_\alpha(\mathbf{k})}\Psi(\alpha\mathbf{k},\mathbf{r}), \tag{3.6}$$

where Θ_α is a (real) phase. It is possible to set $\Theta_\alpha=0$ but this turns out to be an undersirable choice as we will see below. In fact, we will choose

$$\{\alpha|\mathbf{t}_\alpha\}\Psi_n(\mathbf{k},\mathbf{r})=\chi_n^{(j)}(\alpha)\,e^{-i\alpha\mathbf{k}\cdot\mathbf{t}_\alpha}\Psi_n(\alpha\mathbf{k},\mathbf{r}), \tag{3.7}$$

in which $\chi_n^{(j)}(\alpha)$ is the character for the operation α in the jth one-dimensional representation of the point group and has numerical value ±1. The particular representation j is chosen in order to insure, as far as possible, smooth behavior of the wave function as its argument, $\mathbf{k}$, goes around the Brillouin zone. We shall show that the factor $e^{-i\alpha\mathbf{k}\cdot\mathbf{t}_\alpha}$, which would appear if we replaced the Bloch function $\Psi_n(\mathbf{k},\mathbf{r})$ by a plane wave, has

[22] J. S. Lomont, *Applications of Finite Groups* (Academic Press Inc., New York, 1960).

a desirable effect on the expansion coefficients of the wave function when expanded in plane waves.

Since we are using a pseudopotential method, it is desirable to expand the Bloch function in the form

$$\Psi_n(\mathbf{k},\mathbf{r})=\frac{1}{(2\pi)^{3/2}}\sum_s b_n(\mathbf{k},\mathbf{K}_s)e^{i(\mathbf{k}+\mathbf{K}_s)\cdot\mathbf{r}}, \quad (3.8)$$

in which $\mathbf{K}_s$ is a reciprocal lattice vector. The quantity b_n is the momentum wave function for band n. It is possible to show that[23]

$$b_n(\mathbf{k},\mathbf{K}_s)=b_n(\mathbf{k}+\mathbf{K}_s), \quad (3.9)$$

but since we prefer to use the reduced zone scheme in which the range of variation of $\mathbf{k}$ is confined to the first Brillouin zone, the notation of Eq. (3.8) is more appropriate for our purposes. The factor of $(2\pi)^{-3/2}$ in Eq. (3.8) insures that the Bloch functions have the conventional delta-function normalization

$$\int\Psi_n^*(\mathbf{k},\mathbf{r})\Psi_l(\mathbf{k}',\mathbf{r})d^3r=\delta_{nl}\delta(\mathbf{k}-\mathbf{k}')$$

if the b_n are orthonormal:

$$\sum_s b_n^*(\mathbf{k},\mathbf{K}_s)b_l(\mathbf{k},\mathbf{K}_s)=\delta_{nl}.$$

It is frequently possible, as in the case for the diamond lattice, to choose the original of coordinates or the calculation of energy bands in the prefect crystal in such a way that the perfect-crystal Hamiltonian is represented by a real symmetric matrix on a plane-wave [or orthogonalized-plane-wave (OPW)] basis. Then it is possible to find eigenvectors $b_n(\mathbf{k},\mathbf{K}_s)$ which are real for all $\mathbf{k}$. We will assume below that this has been done, so that the b_n are always real. It will be immediately observed that this choice reduces the indeterminacy of phase of the Bloch functions to a question of algebraic sign, and this question can be settled by the requirement that the $b_n(\mathbf{k},\mathbf{K}_s)$ be smooth functions of $\mathbf{k}$.

TABLE I. The components $b_n(\mathbf{k},\mathbf{K}_s)$ are given for the highest valence band and lowest conduction band at the points $\mathbf{k}=(2\pi/a)(\frac{1}{2},\frac{1}{8},\frac{1}{8})$ and $\mathbf{k}=(2\pi/a)(-\frac{1}{8},-\frac{1}{8},\frac{1}{2})$ as obtained from the band calculation described in Sec. V. These results illustrate the behavior of eigenvector components of planes on symmetry in the zone, and also show the relation between eigenvectors for two members of the star of $\mathbf{k}$. The reader may verify that Eqs. (3.12) and (3.13) are satisfied.

$\mathbf{K}_s$	$k=(2\pi/a)(\frac{1}{2},\frac{1}{8},\frac{1}{8})$ Highest valence band	Lowest conduction band	$k=(2\pi/a)(-\frac{1}{8},-\frac{1}{8},\frac{1}{2})$ Highest valence band	Lowest conduction band
0 0 0	0.00000	0.22181	0.00000	0.22181
1 1 1	0.00000	−0.10624	0.00000	−0.02104
−1 1 1	0.00000	−0.50770	0.08412	0.12900
1−1 1	−0.08412	0.12900	−0.08412	0.12900
1 1−1	0.08412	0.12900	0.00000	−0.18339
−1−1 1	0.65376	−0.50758	0.00000	0.10624
1−1−1	0.00000	0.02104	0.65376	−0.50758
−1 1−1	−0.65376	−0.50758	−0.65376	−0.50758
−1−1−1	0.00000	0.18339	0.00000	0.50771
2 0 0	0.00000	−0.00710	−0.21677	0.15603
−2 0 0	0.00000	0.22237	−0.13615	0.02155
0 2 0	0.13615	−0.02155	0.21677	0.15603
0−2 0	0.21677	−0.15603	0.13615	0.02155
0 0 2	−0.13615	−0.02155	0.00000	−0.00740
0 0−2	−0.21677	−0.15603	0.00000	0.22237

[23] J. C. Slater, Rev. Mod. Phys. 6, 209 (1934).

Integrals involving Wannier functions and/or Bloch functions may be expressed in terms of integrals involving the functions $b_n(\mathbf{k},\mathbf{K}_s)$. We must therefore determine how the b_n vary throughout the zone. To examine this in detail, let us consider the effect of $\{\alpha|\mathbf{t}_\alpha\}$ on $\Psi_n(\mathbf{k},\mathbf{r})$ when the wave function is given by Eq. (3.8):

$$\{\alpha|\mathbf{t}_\alpha\}\Psi_n(\mathbf{k},\mathbf{r})=(2\pi)^{-3/2}\sum_s b_n(\mathbf{k},\mathbf{K}_s)\{\alpha|\mathbf{t}_\alpha\}e^{i(\mathbf{k}+\mathbf{K}_s)\cdot\mathbf{r}}$$

$$=(2\pi)^{-3/2}e^{-i\alpha\mathbf{k}\cdot\mathbf{t}_\alpha}\sum_s b_n(\mathbf{k},\mathbf{K}_s) \quad (3.10)$$

$$\times e^{-i\alpha\mathbf{K}_s\cdot\mathbf{t}_\alpha}e^{i\alpha(\mathbf{k}+\mathbf{K}_s)\cdot\mathbf{r}}.$$

On the other hand, if we substitute the plane-wave expansion directly into the right side of Eq. (3.7) we get

$$\{\alpha|\mathbf{t}_\alpha\}\Psi_n(\mathbf{k},\mathbf{r})=\frac{\chi_n^{(j)}(\alpha)}{(2\pi)^{3/2}}e^{-i\alpha\mathbf{k}\cdot\mathbf{t}_\alpha}$$

$$\sum_t b_n(\alpha\mathbf{k},\mathbf{K}_t)e^{i(\alpha\mathbf{k}+\mathbf{K}_t)\cdot\mathbf{r}}. \quad (3.11)$$

Comparison of Eqs. (3.10) and (3.11), plus the fact that $[\chi_n^{(j)}(\alpha)]^{-1}=\chi_n^{(j)}(\alpha)$, yields

$$b_n(\alpha\mathbf{k},\mathbf{K}_s)=\chi_n^{(j)}(\alpha)b_n(\mathbf{k},\alpha^{-1}\mathbf{K}_s)e^{-i\mathbf{K}_s\cdot\mathbf{t}_\alpha}. \quad (3.12)$$

In the diamond lattice, the factor $\exp[-i\mathbf{K}_s\cdot\mathbf{t}_\alpha]$ is always real.

We illustrate this relation in Table I in which a few eigenvectors in silicon are compared.

We have not yet discussed how the characters $\chi_n^{(j)}(\alpha)$ are determined. This may be done by inspection of the behavior of the components $b_n(\mathbf{k},\mathbf{K}_s)$ on planes of symmetry in the Brillouin zone. Such planes include ones in which $k_x=k_y$, with k_z arbitrary; $k_x=0$, k_y and k_z arbitrary; etc. Whenever $\mathbf{k}$ is in such a plane, there is an operation, say α_1, such that $\alpha_1\mathbf{k}=\mathbf{k}$. Then we have in this case from Eq. (3.12)

$$b_n(\mathbf{k},\mathbf{K}_s)=\chi_n^{(j)}(\alpha_1)b_n(\mathbf{k},\alpha_1^{-1}\mathbf{K}_s)e^{-i\mathbf{K}_s\cdot\mathbf{t}_{\alpha_1}}. \quad (3.13)$$

We may use this relation between components of the eigenvector b_n at the same point of the Brillouin zone to determine the $\chi_n^{(j)}(\alpha)$. Let us consider the specific reciprocal lattice vector $K_s=0$. In this case, Eq. (3.13) reduces still further to

$$b_n(\mathbf{k},0)=\chi_n^{(j)}(\alpha_1)b_n(\mathbf{k},0). \quad (3.14)$$

Thus, if $b_n(\mathbf{k},0)\neq 0$, $\chi=1$. In order to have $\chi=-1$, we must have $b_n(\mathbf{k},0)=0$. Thus, we may look at the $\mathbf{K}_s=0$ component of b_n on a symmetry plane to determine the character to be associated with a given band for some symmetry operation.

On the basis of these results, we can understand why the character $\chi_n^{(j)}(\alpha)$ should appear in Eq. (3.7). The basic subzone (containing 1/48 of the volume of the Brillouin zone) in which wave functions and energies are calculated is bounded by symmetry planes of the type mentioned above. Suppose $\mathbf{k}$ is in such a plane. Then $\Psi_n(\mathbf{k},\mathbf{r})$ must be a basis function for an irreducible representation of the space group for that $\mathbf{k}$, and in view of the existence of a symmetry operator α_1, as above, the plane-wave expansion coefficients $b_n(\mathbf{k},\mathbf{K}_s)$ will obey Eq. (3.13) in which $\chi(\alpha_1)$ is the character of α_1 in the small group of $\mathbf{k}$. We now demand that the wave function on either side of the symmetry plane must join smoothly with that on the plane. This requirement forces $\Psi_n(\mathbf{k},\mathbf{r})$ to transform in accord with Eq. (3.7) even when $\mathbf{k}$ is not on a symmetry plane.

In the cubic point group, there are four one-dimensional representations, usually denoted by Γ_1, Γ_2, Γ_1', and Γ_2'. To determine which of these should be used in in Eq. (3.7) it is necessary to consider only two symmetry planes in the Brillouin zone, and apply Eq. (3.14). The planes (1) $k_z=0$ and (2) $k_x=k_y$ suffice. A Γ_1 function has nonvanishing $b_n(\mathbf{k},0)$ on both planes; whereas $b_n(\mathbf{k},0)$ vanishes on both planes in the case of Γ_1'. A function belonging to Γ_2 has $b_n(\mathbf{k},0)=0$ when $k_x=k_y$ but not when $k_z=0$; while Γ_2' has $b_n(\mathbf{k},0)=0$ if $k_z=0$ but not if $k_x=k_y$.

These considerations permit us to choose the phases of the Bloch functions so as to give satisfactory Wannier functions in the case of an isolated band. Choice of the proper characters in Eq. (3.7) insures that the Bloch function varies smoothly going around the zone. If this choice is made in a manner inconsistent with Eq. (3.14), a Wannier function can still be defined, but the components $b_n(\mathbf{k},\mathbf{K}_s)$ will not go smoothly through the planes of symmetry in the zone, and as a result, the Wannier function will not be properly localized. It is therefore necessary only to choose the sign of the eigenvector b_n within 1/48 of the zone so that the components vary continuously from point to point within this region, and the functions will then vary smoothly throughout the zone. Since the choice of phase was reduced to one of algebraic sign by the requirement that the b_n be real, an unambiguous prescription for the phase has, in effect, been given.

We can now proceed to a determination of the transformation properties of the Wannier functions. We consider only space-group operations for which $\mathbf{t}_\alpha$ may be taken to be zero, for only in this case is a simple result obtained. We use the definition of the Wannier function given in Eq. (2.7), and apply Eq. (3.7):

$$\{\alpha|0\}a_n(\mathbf{r}-\mathbf{R}_\mu)=\frac{\Omega^{1/2}}{(2\pi)^{3/2}}\times\int d^3k\, e^{-i\mathbf{k}\cdot\mathbf{R}_\mu}\{\alpha|0\}\Psi_n(\mathbf{k},\mathbf{r}), \quad (3.15)$$

$$\begin{aligned}\{\alpha|0\}a_n(\mathbf{r}-\mathbf{R}_\mu)&=\frac{\Omega^{1/2}\chi_n^{(j)}(\alpha)}{(2\pi)^{3/2}}\times\int e^{-i\mathbf{k}\cdot\mathbf{R}_\mu}\Psi_n(\alpha\mathbf{k},\mathbf{r})d^3k\,,\\&=\frac{\Omega^{1/2}\chi_n^{(j)}(\alpha)}{(2\pi)^{3/2}}\times\int e^{-i\mathbf{k}'\cdot\alpha\mathbf{R}_\mu}\Psi_n(\mathbf{k}',\mathbf{r})d^3k' \quad (3.16)\\&=\chi_n^{(j)}(\alpha)a_n(\mathbf{r}-\alpha\mathbf{R}_\mu)\,.\end{aligned}$$

On the other hand, we have directly from Eq. (3.15) that

$$\{\alpha|0\}a_n(\mathbf{r}-\mathbf{R}_\mu)=a_n(\alpha^{-1}\mathbf{r}-\mathbf{R}_\mu)\,. \quad (3.17)$$

Equations (3.16) and (3.17) hold for all α; therefore, we may proceed as follows: Equate (3.16) and (3.17); replace α by α^{-1}, and let $\mathbf{R}_\mu=\alpha\mathbf{R}_\nu$. Then we have

$$a_n(\mathbf{r}-\mathbf{R}_\nu)=\chi_n^{(j)}(\alpha)a_n[\alpha(\mathbf{r}-\mathbf{R}_\nu)]. \quad (3.18)$$

These equations describe the transformation properties of the Wannier functions for a simple isolated band.

We can use Eq. (3.18) to establish a relation between certain matrix elements of the defect potential on the basis of Wannier functions. Suppose that the potential is unchanged by some operation $\{\beta|0\}$. We consider the matrix element

$$(n\mu|V|l\nu)=\int a_n^*(\mathbf{r}-\mathbf{R}_\mu)V(\mathbf{r})a_l(\mathbf{r}-\mathbf{R}_\nu)d^3r. \quad (3.19)$$

Let $\mathbf{R}_\mu=\beta\mathbf{R}_\rho$ and $\mathbf{R}_\nu=\beta\mathbf{R}_\tau$. Then, since $V(\beta\mathbf{r})=V(\mathbf{r})$, we find that

$$(n\mu|V|l\nu)=\chi_n^{(j)}(\beta)\chi_l^{(k)}(\beta)(n\rho|V|l\tau). \quad (3.20)$$

As an example of the use of Eq. (3.20), suppose that the defect potential is invariant under the group C_{3v}. Then the central cell matrix element of this potential ($\mathbf{R}_\mu=\mathbf{R}_\nu=0$) vanishes if band n is either Γ_1 or Γ_2' and band 1 is either Γ_2 or Γ_1'. Other useful relations are obtained immediately.

We now want to consider the factorization of the determinant $\bar{D}(E)$, which is defined in Eq. (2.13), through the use of symmetry considerations. Since the Wannier functions for a given band form an orthonormal set, Eq. (3.16) tells us that the Wannier functions for a given lattice vector type (by type, we mean that set of lattice vectors which can be formed from any one of them by applying all the operations of the group considered) are basis functions for a (reducible) representation of the group of the operator $\{\alpha|0\}$ (or of any subgroup of this group). We may now apply standard techniques to construct symmetrized linear combinations of Wannier functions which transform according to ir-

reducible representations of the symmetry group of the defect potential. This was implicitly assumed in Ref. 14.

In particular, if we want a linear combination of Wannier functions which belong to the σth row of the sth irreducible representation of the group of the potential (we will denote the operators in this group by β), we form the combination

$$\sum_\beta D_{\sigma\sigma}{}^{(s)}(\beta)\{\beta|0\}a_n(\mathbf{r}-\mathbf{R}_\mu) = \sum_\beta D_{\sigma\sigma}{}^{(s)}(\beta)a_n(\mathbf{r}-\beta\mathbf{R}_\mu)\chi_n{}^{(j)}(\beta). \quad (3.21)$$

If there is more than one combination of Wannier functions from band n and lattice vectors of type $\mathbf{R}_\mu$, which belong to row σ of representation s, the functions obtained from (3.21) for (σ,s) will usually have to be orthogonalized to each other, possibly by the Schmidt process. Apart from this possible complication, the elements of the unitary transformation $U(\mathbf{s}_\sigma,\mathbf{R}_\mu)$ introduced in Ref. 14, may be determined immediately.

According to the general principle of group theory, there will be no matrix elements of the Green's function or potential between combinations of Wannier functions which belong to different rows of the same representation. Consequently, the determinant $\bar{D}(E)$ defined in Eq. (2.13) will factor into a product of subdeterminants coming from the various irreducible representations.

$$\bar{D}(E)=\prod_s [\bar{D}_s(E)]^{g_s}, \quad (3.22)$$

in which g_s is the degeneracy of representation s. Consequently, in order to solve Eq. (2.14) it is necessary only to examine the simpler equation

$$\bar{D}_s(E_0)=0. \quad (3.23)$$

A solution of Eq. (3.21) belongs to representation s.

IV. THE PSEUDOPOTENTIAL

In order to determine the energy levels of the bound states (if any) associated with a vacancy, it is necessary to have some expression for the change in potential produced by the removal of a silicon atom. To begin, we will ignore the effects produced by the relaxation of atoms near the vacancy. Then since the total crystal potential can be represented as the sum of potentials due to individual atoms, the vacancy perturbation will be the negative of the potential of a single atom.

The true potential of a silicon atom is quite strong and gives rise to core eigenstates as well as to the valence states of principal interest. Wave functions for valence states, including those associated with a defect, must be orthogonal to the wave functions of core electrons. This leads to inconvenient complications, which are circumvented through the pseudopotential method (see Harrison[24] for a review of this procedure). Pseudopotential calculations have been quite successful in reproducing the essential features of the observed band structures for the common semiconductors. Harrison,[25] using the general analysis of the pseudopotential method given by Austin, Heine, and Sham[26] has shown that a crystal pseudopotential can be expressed as a sum of contributions from individual atoms. Therefore, we may simplify our calculations not only by using a pseudopotential method to determine energy bands in the perfect crystal, but we may represent the vacancy as the negative of an atomic pseudopotential. We will also use pseudopotential wave functions to construct the Wannier function of our calculation.

Brust[27] has found empirical pseudopotential parameters for silicon which yield energy bands in good agreement with experiment. This is a local potential in the sense that the matrix elements of this pseudopotential V_{ps} between two plane-wave states $|\mathbf{k}\rangle$ and $|\mathbf{q}\rangle$ are functions only of the vector difference $\mathbf{q}-\mathbf{k}$. We write

$$\langle\mathbf{k}|V_{ps}|\mathbf{q}\rangle=U(\mathbf{q}-\mathbf{k}), \quad (4.1)$$

with

$$V_{ps}|\mathbf{q}\rangle=\sum_{\mu j} v_{ps}(\mathbf{r}-\mathbf{R}_{\mu j})|\mathbf{q}\rangle, \quad (4.2)$$

in which $v_{ps}(\mathbf{r}-\mathbf{R}_{\mu j})$ is the atomic pseudopotential for an atom at the jth site in the unit cell located at $\mathbf{R}_\mu$. To fix the normalization, we will suppose that the plane-wave states $|\mathbf{k}\rangle$ and $|\mathbf{q}\rangle$ satisfy

$$\langle\mathbf{k}|\mathbf{q}\rangle=\delta(\mathbf{k}-\mathbf{q}). \quad (4.3)$$

Strictly speaking, the assumption that the pseudopotential is local is not correct, but it is not known if the neglect of nonlocal character is of experimental importance, and we will assume we have a local potential.

By combining Eqs. (2.7) and (3.8), we obtain the following expression for the matrix elements of the vacancy potential between Wannier functions:

$$(n\mu|V|l\nu)=-\int a_n{}^*(\mathbf{r}-\mathbf{R}_\mu)v_{ps}(\mathbf{r}-\mathbf{d})a_l(\mathbf{r}-\mathbf{R}_\nu)d^3r$$

$$=\frac{-\Omega}{(2\pi)^6}\sum_{st}\int\int e^{i(\mathbf{k}\cdot\mathbf{R}_\mu-\mathbf{p}\cdot\mathbf{R}_\nu)}b_n{}^*(\mathbf{k},\mathbf{K}_s)b_l(\mathbf{p},\mathbf{K}_t)$$

$$\times\left[\int d^3r\, e^{-i(\mathbf{k}+\mathbf{K}_s)\cdot\mathbf{r}}v_{ps}(\mathbf{r}-\mathbf{d})e^{i(\mathbf{p}+\mathbf{K}_t)\cdot\mathbf{r}}\right]d^3k\,d^3p.$$

We have placed the perturbation in the cell at the origin. Each unit cell in the diamond lattice contains 2 atoms, located at $\pm\mathbf{d}$ with respect to the cell center; hence we obtain the argument $\mathbf{r}-\mathbf{d}$ if the atom at $\mathbf{d}$ is removed. The pseudopotential coefficients given by Brust are

[24] W. A. Harrison, *Pseudopotentials in the Theory of Metals* (W. A. Benjamin, Inc., New York, 1964).

[25] Reference 24, p. 19.

[26] B. J. Austin, V. Heine, and L. J. Sham, Phys. Rev. **127**, 276 (1962).

[27] D. Brust, Phys. Rev. **134**, A1337 (1964).

effectively given by

$$U(\mathbf{q}-\mathbf{k})=\cos[(\mathbf{q}-\mathbf{k})\cdot\mathbf{d}] \times u_{at}(\mathbf{q}-\mathbf{k})\sum_s \delta(\mathbf{q}-\mathbf{k}-\mathbf{K}_s), \quad (4.4)$$

where

$$u_{at}(\mathbf{q}-\mathbf{k})=\frac{2}{\Omega}\int e^{-i\mathbf{k}\cdot\mathbf{r}}v_{ps}(\mathbf{r})e^{i\mathbf{q}\cdot\mathbf{r}}d^3r, \quad (4.5)$$

Ω being, as before, the cell volume, and $\mathbf{K}_s$ a reciprocal lattice vector. Then we find

$$(n\mu|V|l\nu)=\frac{-\Omega^2}{2(2\pi)^6}\sum_{st}\int\int \exp\{i[(\mathbf{k}+\mathbf{K}_s)\cdot(\mathbf{R}_\mu-d) -(\mathbf{p}+\mathbf{K}_t)\cdot(\mathbf{R}_\nu-\mathbf{d})]\}b_n(\mathbf{k},\mathbf{K}_s)u_{at}(\mathbf{p}+\mathbf{K}_t-\mathbf{k}-\mathbf{K}_s) \times b_l(\mathbf{p},\mathbf{K}_t)d^3kd^3p. \quad (4.6)$$

A problem arises immediately in that band calculations for perfect crystals use (and hence "determine") the Fourier coefficients of the atomic pseudopotential only when the argument is a reciprocal lattice vector. This is apparent from Eq. (4.4). In the present calculation, we must know $u_{at}(\mathbf{k})$ for all $\mathbf{k}$. In order to approximate this, we simply make a four term polynomial fit to the pseudopotential parameters given by Brust. It is necessary to obtain $u_{at}(0)$ by other means since this quantity is not determined in a calculation which is concerned with the relative position of the bands. We have used the prescription of Harrison to find a value of -0.61 Ry for this quantity, and it has been included in the construction of the polynomial below:

$$u_{at}(\mathbf{k})=\sum_n \alpha_n\left(\frac{|\mathbf{k}|a}{2\pi}\right)^{2n}, \quad |\mathbf{k}|^2<48\pi^2/a^2$$
$$u_{at}(\mathbf{k})=0, \quad |\mathbf{k}|^2\geq 48\pi^2/a^2. \quad (4.7)$$

The coefficients α_n are given in Table II.

The matrix elements of the pseudopotential were computed numerically as will be described in the next section.

TABLE II. The coefficients α_n in a polynomial fit to Brust's silicon pseudopotential [see Eq. (4.7)].

n	α_n(Ry)
0	−0.610000
1	0.177311
2	−0.016250
3	0.000530

V. APPLICATION TO SILICON: ENERGY BANDS

We shall now discuss the extension of Sec. III concerning the symmetries of Bloch functions and Wannier functions for isolated simple bands to silicon. The actual band structure of silicon contains both degeneracies required by symmetry (which occur at isolated symmetry points and along certain symmetry axes in the Brillouin zone), accidental degeneracies and quasidegeneracies which occur when energy bands approach each other at general points of the Brillouin zone. Quasidegeneracies are found to be rather numerous. Near a quasidegeneracy the wave functions behave as if their $E(\mathbf{k})$ surfaces would cross but in fact they do not (von Neumann and Wigner).[28] The existence of degeneracies and quasidegeneracies in the energy bands of real crystal requires extension of the analysis of the previous section. Attempts in this direction have been made by Takeuti,[29] Blount[20] and des Cloizeaux.[21] However, at this time, a usable and complete analysis does not seem to exist. In this section we describe the manner in which we have proceeded to form the energy bands. The discussion will be primarily descriptive. It is hoped that a more detailed theoretical analysis can be given in the future.

In order to utilize the results of Secs. III and IV we have chosen to define energy bands pertaining to a single symmetry. Quasidegeneracies are rather uncommon in the four valence bands, so that it is possible simply to arrange these states in order of increasing energy. The conduction bands, on the other hand, are more complicated and exhibit many near crossings which shall henceforth be called quasidegeneracies. In the presence of quasidegeneracies we have departed, where necessary, from the ordering of bands according to increasing energy. We have instead followed the symmetry of the wave functions as indicated by the behavior of the coefficients $b_n(\mathbf{k},\mathbf{K}_s)$. Our Green's functions are calculated consistently with this procedure, and thus the conduction-band Green's functions may not, strictly speaking, decay exponentially at large distances. Within the limits already imposed by the numerical accuracy we can achieve, this problem is probably not particularly serious. It turns out, as will be seen, the conduction bands are not particularly important in the determination of the energy of the vacancy bound state we are considering. The vital question from our point of view concerns the convergence of the bound-state energy value, and this will be seen to be satisfactory.

To obtain energy bands and wave functions, we have employed the pseudopotential plane-wave method as discussed in Sec. IV. In the pseudopotential diagonalization procedure we seek the energy eigenvalues $E_n(\mathbf{k})$ and the pseudo-wave function $\Psi_n(\mathbf{k},\mathbf{r})$ of the equation

$$H_p\Psi_n(\mathbf{k},\mathbf{r})=E_n(\mathbf{k})\Psi_n(\mathbf{k},\mathbf{r}), \quad (5.1)$$

where $H_p=(-\hbar^2/2m)\nabla^2+V_{ps}(\mathbf{r})$. On a plane-wave

[28] J. von Neumann and E. P. Wigner, Z. Physik **30**, 467 (1929).

[29] Y. Takeuti, Progr. Theoret. Phys. (Kyoto) Suppl. **12**, 75 (1959).

basis the secular equation is

$$|H_p{}^{ij}-E\delta_{ij}|=0, \tag{5.2}$$

where

$$H_p{}^{ij}=\langle\Psi_{\mathbf{k}+\mathbf{K}_i}|H_p|\Psi_{\mathbf{k}+\mathbf{K}_j}\rangle,$$

with $\Psi_{\mathbf{k}+\mathbf{K}_i}=\exp[i(\mathbf{k}+\mathbf{K}_i)\cdot\mathbf{r}]$. In order to keep the numerical labor within reasonable bounds we have truncated the secular equation after fifteen plane waves. The fifteen plane waves employed were the (000) plane wave, the eight plane waves of the type (111), and the six plane waves of the type (200) (units $2\pi/a$).

The diagonalization of the secular equation was carried out numerically by the Jacobi method at a number of points in the Brillouin zone. As solutions to the 15×15 secular determinant at a point $\mathbf{k}$ we obtain numerical eigenvalues $E_m(\mathbf{k})$ and the associated eigenvectors $(b_m(\mathbf{k},\mathbf{K}_1),b_m(\mathbf{k},\mathbf{K}_2),\cdots,b_m(\mathbf{k},\mathbf{K}_{15}))$ with $m=1,\cdots,15$. The subscript m here labels columns of a unitary (and real) matrix S which diagonalizes the Hamiltonian matrix H_p, and in general does not rank the eigenvalues in any particular order. In particular, m is not a band index. It is found that usually two or three subscripts m will be involved in obtaining a band over the entire Brillouin zone. The significance of the m subscript is clear in the limiting case of free electrons. In that case m would refer to an energy level $\sim(\mathbf{k}+\mathbf{K}_m)^2$ and a wave function with coefficients given by $b_m(\mathbf{k},\mathbf{K}_j)=\delta_{mj}$.

It is, of course, not necessary to diagonalize the secular determinant over the entire Brillouin zone. In particular we need to consider only the fundamental subzone of the Brillouin zone containing 1/48 of the volume of the Brillouin zone and characterized by k_x, k_y, k_z. The bounding interior planes of this subzone D are $k_z=0$, $k_x=k_y$, and $k_y=k_z$.

It follows as a general consequence of the theory of the space group of the Hamiltonian that if the energy bands and wave functions within the 1/48 subzone D are known, then energy bands and wave functions throughout the entire Brillouin zone may be obtained from those in D by a suitable transformation.

Before proceeding to numerical examples of the method used to define the energy bands in silicon, let us consider briefly what extensions and modifications to Sec. III are required by the various types of degeneracies present in silicon. As in Sec. III, we shall restrict ourselves to situations in which only one-dimensional irreducible representations need to be considered. We therefore exclude from detailed consideration those $E(\mathbf{k})$ points in D which are degenerate, either by symmetry or by accident. Examples would include the various threefold degenerate energy levels at Γ, the twofold degeneracy at the point X and any accidental degeneracies along symmetry axes and symmetry planes. Apparently points of quasidegeneracy must be similarly excluded. This point of view is similar to that of Bouckaert, Smoluchowski, and Wigner[30] who, in a somewhat different context, excluded from the proper definition of an energy band those points where two energy surfaces "stick together." We emphasize that points $\mathbf{k}$ in the vicinity of points of quasidegeneracy $\mathbf{k}_0$ are not excluded.

Consider a point $\mathbf{k}$ in D, an energy $E_n(\mathbf{k})$, and the wave function $\Psi_n(\mathbf{k},\mathbf{r})=\sum_{\mathbf{K}_i} b_n(\mathbf{k},\mathbf{K}_i)\exp[i(\mathbf{k}+\mathbf{K}_i)\cdot\mathbf{r}]$. The subscript n will be referred to loosely as the "band" index but its precise meaning in the present discussion is merely that of a label between $E_n(\mathbf{k})$ and $\Psi_n(\mathbf{k},\mathbf{r})$ at $\mathbf{k}$ in D. Since $\Psi_n(\mathbf{k},\mathbf{r})$ at $\mathbf{k}$ is nondegenerate, Eq. (3.12) must be applicable at the point $\mathbf{k}$. To emphasize the applicability at the point $\mathbf{k}$ we write Eq. (3.12) in a slightly different form:

$$b_n(\alpha\mathbf{k},\mathbf{K}_i)=\chi_n{}^{(j)}(\mathbf{k};\alpha)b_n(\mathbf{k},\alpha^{-1}\mathbf{K}_i)\exp(-i\mathbf{K}_i\cdot t_\alpha). \tag{5.3}$$

$\chi_n{}^{(j)}(\mathbf{k};\alpha)$ is the character of the operation α for the particular one-dimensional representation j which pertains to $\Psi_n(\mathbf{k},\mathbf{r})$ and $E_n(\mathbf{k})$ at the point $\mathbf{k}$. Next consider a collection of points $E_n(\mathbf{k})$, one for each point $\mathbf{k}$ in D, selected so that taken together the $E_n(\mathbf{k})$ points form a fairly smooth $E(\mathbf{k})$ surface in the Brillouin zone. We emphasize again that the subscript n is merely a label and neither ranks the energy level $E_n(\mathbf{k})$ at $\mathbf{k}$ nor is required to have the same indicial value at the different points $\mathbf{k}$ in D.

Define a given energy band, say the mth, to consist of all the energy values $E_n(\mathbf{k})$ selected above and their wave functions $\Psi_n(\mathbf{k},\mathbf{r})$, and characterized by the symmetry characters $\chi_n{}^{(j)}(\mathbf{k};\alpha)$. Two types of situations will arise. Either the characters depend on and vary with $\mathbf{k}$, or else they are independent of $\mathbf{k}$. In the first case the energy band, as defined, consists of Bloch functions of different symmetries and is a "mixed" band. Equations (3.14) through (3.20) therefore are not applicable. In the second case the energy band consists only of Bloch functions of the same symmetry and Eqs. (3.14)–(3.20) are applicable. We have consistently defined our energy bands such that they are always of the single symmetry type so that the "simple" band analysis is applicable.

This prescription is of considerable convenience in that group-theoretic analysis and simplifications may be applied to the Wannier functions and to the determination of the bound state of the defect. Since both approaches yield orthonormal Wannier functions, neither the single symmetry prescription nor the mixed symmetry prescription can be considered to be the more "correct" procedure. Any preference would apparently have to come from a consideration of the relative amounts of labor required to obtain convergent solutions for the defect problems to which the general method is applicable. It is possible that other, more complicated, constructions of Wannier functions and energy bands might also prove useful in some problems.

Let us begin the discussion of the results for the energy-band calculation by considering the four valence bands (levels Γ_1 and $\Gamma_{25'}$ at $\mathbf{k}=0$) lying below the for-

[30] L. P. Bouckaert, R. Smoluchowski, and E. Wigner, Phys. Rev. **50**, 58 (1936).

bidden energy gap. These four bands are, at general points in the Brillouin zone, separated from one another and seldom exhibit either degeneracies or quasidegeneracies. These are therefore "simple" bands and the analysis of Sec. III is applicable in toto. These four bands are labeled in order of increasing energy. The behavior of the coefficients $b_n(\mathbf{k},\mathbf{K}_i)$ on the bounding planes was examined and the symmetry assignments made in accordance with Sec. III. These symmetry assignments are given in Table III.

We next describe the conduction bands. For definiteness we must describe how the bands are to be labeled. It is observed that through much of the 1/48 subzone D the bands may be arranged and labeled in order of increasing energy. In particular most of the volume of $\mathbf{k}$ space lying in the vicinity of the Γ-Q line in D (end points $\Gamma[0,0,0]$ and Q $[(2\pi/a)(\frac{3}{4},\frac{1}{2},\frac{1}{4})]$ are points of degeneracy and are excluded) does not appear to exhibit degeneracies or quasidegeneracies. We have therefore determined an energy ranking and labeling of the energy bands in this region. The four conduction bands we denote by bands 5, 6, 7, and 8 are 5th, 6th, 7th, and 8th in order of increasing energy at approximately the midpoint of the Γ-Q line. We shall see however that the bands as defined do not maintain this ranking over the entire subzone D.

For purposes of discussion we lay out a regularly spaced mesh of general points in the interior of zone D. We choose a grid such that we obtain 60 general points in the interior of D. In order to label the 60 points with integers a mltiplying scale factor of 960 has been used. We now start at our reference point and follow the previously defined bands out in all directions along the mesh, maintaining the same energy ordering until either a degeneracy or quasidegeneracy is encountered. Figure 1 represents the four conduction bands along a line between the points (240,144,48) and (528,144,48) with a scale factor of 960. This graph is fairly typical and exhibits two quasidegeneracies. The energy bands and the symmetry assignments are labeled consistently in our notation. The quasidegeneracies are identified from a detailed consideration of the coefficients $b_n(\mathbf{k},\mathbf{K}_i)$ and the energy graphs. An example will illustrate the procedure. Let the column vector $(b_n(\mathbf{k},\mathbf{K}_1),b_n(\mathbf{k},\mathbf{K}_2),\cdots,$

TABLE III. Symmetry assignments for the calculated energy bands as required for the energy-band analysis (O_h) and for the vacancy potential (C_{3v}). All Wannier functions formed from a given band are thus either pure real or pure imaginary as indicated in the fourth column.

Band	(O_h)	(C_{3v})	
1	Γ_1	Λ_1	R
2	Γ_2'	Λ_1	I
3	Γ_1	Λ_1	R
4	Γ_1'	Λ_2	I
5	Γ_1	Λ_1	R
6	Γ_2'	Λ_1	I
7	Γ_2'	Λ_1	I
8	Γ_2	Λ_2	R

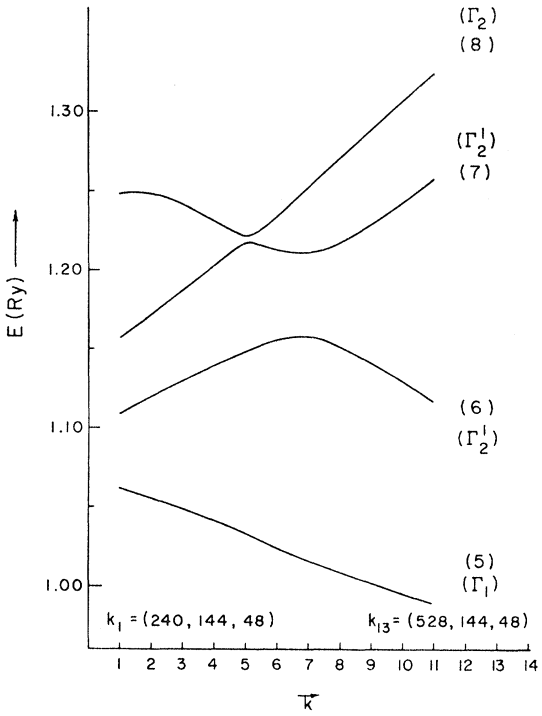

FIG. 1. A graph of the four lowest conduction bands along a line in the Brillouin zone between the points $k_1=(240,144,48)$ and $k_{13}=(528,144,48)$ for a scale factor of 960. The figure illustrates a very close quasidegeneracy between two bands (Γ_2 and Γ_2') of different symmetry and a more widely separated quasidegeneracy between two bands (Γ_2') of the same symmetry. At the point k_1 the labeling of the bands is 6, 8, 7, 5.

$b_n(\mathbf{k},\mathbf{K}_{15})$) for band n at the point $\mathbf{k}$ in D be denoted by $B_n(\mathbf{k};\mathbf{K}_i)$. Let us now compare the vector $B_7(\mathbf{k}_5;\mathbf{K}_i)$ with both $B_6(\mathbf{k}_8;\mathbf{K}_i)$ and $B_7(\mathbf{k}_8;\mathbf{K}_i)$, where $\mathbf{k}_5$ and $\mathbf{k}_8$ are as labeled at the bottom of the figure. It is found that, coefficient by coefficient, $B_7(\mathbf{k}_5;\mathbf{K}_i)$ is recognizably more similar to $B_7(\mathbf{k}_8;\mathbf{K}_i)$ than it is to $B_6(\mathbf{k}_8;\mathbf{K}_i)$. The evaluation of scalar products of these vectors shows that $B_7(\mathbf{k}_5;\mathbf{K}_i)\cdot B_7(\mathbf{k}_8;\mathbf{K}_i)$ is greater than $B_7(\mathbf{k}_5;\mathbf{K}_i)\cdot B_6(\mathbf{k}_8;\mathbf{K}_i)$. In many instances this difference may be as large as a factor of 10. Of course at the point of closest approach $\mathbf{k}_0$ between bands 6 and 7, the dot product is zero since band 6 and band 7 represent two different eigenvalues of a Hermitian matrix and are hence orthogonal. In addition, the reference points (here $\mathbf{k}_5$ and $\mathbf{k}_8$) cannot be located too far apart and still retain their significance. Apparently band 6 should be regarded as the "first" solution with symmetry Γ_2' and band 7 as the "second" solution with symmetry Γ_2'.

The band structure is then analyzed point by point until the bands have been identified at each of the 60 general grid points. In most instances two neighboring points have been connected together by two or more different paths. This is done to achieve internal consistency since any particular path may not always yield a coefficient analysis which is conclusive. After the bands at the 60 points have been labeled, then the

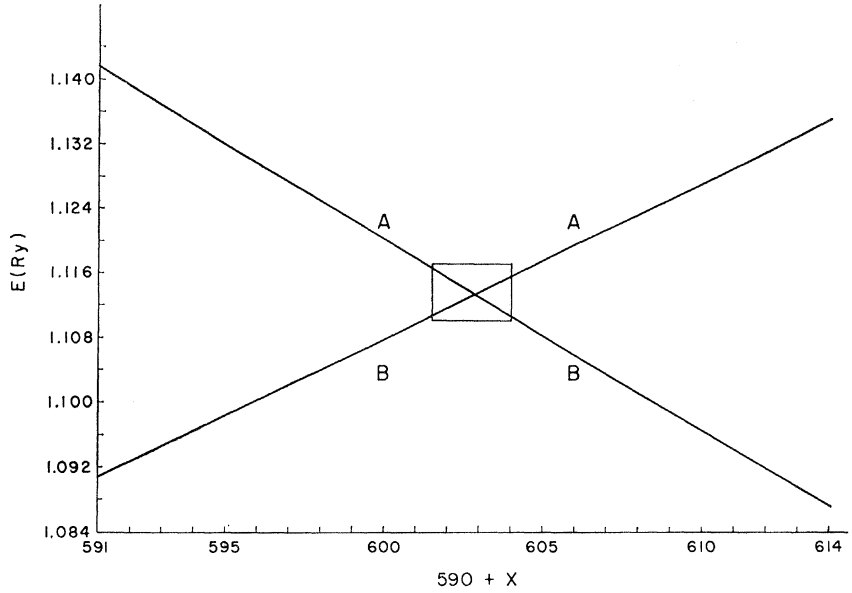

FIG. 2. A graph of band 5(Γ_1) and band 7(Γ_2') along a line given by $(590+x, 410-2x, 80+x)$ with a close quasidegeneracy in the vicinity of the point (603,384,93) for a scale factor of 960. The energy scale is that of Fig. 1. The energy of the two bands appears, on the scale of this graph, to be linear in k and a crossing would seem to be indicated.

energy bands and wave functions at those general points lying adjacent to the bounding planes of D are continued on into the planes and the symmetries identified. The symmetry assignments for the four conduction bands we have treated are given in Table III.

As an example of these considerations, we show in Fig. 2 a close quasidegeneracy between two bands which we will temporarily call A and B. The energy bands appear to vary linearly near the point (606,384,93) (scale factor of 960) where an accidental degeneracy appears to occur. The region within the box in Fig. 2 is shown on an expanded scale in Fig. 3. It is apparent from the expanded diagram that the bands do not actually cross. The minimum energy separation is, however, extremely small—approximately 0.001 Ry. The behavior of the wave function in the vicinity of this quasidegeneracy is quite interesting. From this point of view, the bands appear to switch. The eigenvector $B_n(\mathbf{k},\mathbf{K}_s)$ belonging to band A at the point (600,390,90) resembles that of band B at the point (606,378,96) much more closely than it does that of the lower band at the latter point. Similarly, the band B eigenvector at (600,390,90) resembles that of A at (606,378,96). This resemblance can be described quantitatively by the scalar products of the eigenvectors as was mentioned above. We find

Point (600,390,90)	Point (606,378,96)	Scalar Product
Band A	Band B	0.99604
Band B	Band A	0.99630
Band A	Band A	−0.08314
Band B	Band B	0.08323

With present computational techniques and equipment it is impossible to consider rapid changes in the wave function and energy in regions of this size in the determination of matrix elements of the Green's function and the potential. We are therefore forced to treat the bands in the vicinity of a quasidegeneracy as if they went smoothly through the degeneracy, making the A-B and B-A connection illustrated above.

We have found examples of quasidegeneracies with minimum energy separations in the range between 0.05 and 0.001 Ry. The large separations are generally associated with attempted crossings of bands of the same symmetry [as described by the representation character in Eq. (3.7)], and the extremely small separations with bands of different symmetry. In addition, the quasidegeneracies tend to occur in $\mathbf{k}$ space close to points at which actual degeneracies would occur for free electrons. The pseudopotential couples free electron bands in first order unless the Fourier coefficient of the potential for a wave vector equal to the difference of the vectors of the degenerate pair vanishes, and thus replaces the degeneracy by a quasidegeneracy. The interaction between free-electron wave functions for a general $\mathbf{k}$ can never be zero on the basis of symmetry alone in distinction to the case of free-electron wave functions for symmetry planes, axes, or points.

Let us finally indicate how the Green's functions were calculated. Once the energy bands are known, the Green's functions can be calculated in a straightforward manner from Eq. (2.8). By properly symmetrizing the exponential, it is of course possible to work only within 1/48 of the zone. Our conduction energy bands contain discontinuities at the various quasidegeneracy points. We might anticipate that the conduction-band Green's functions would have somewhat different asymptotic properties at large distances than would the valence-band Green's functions. We have, however, not investigated this. The asymptotic properties of the Green's functions would be difficult to obtain numerically.

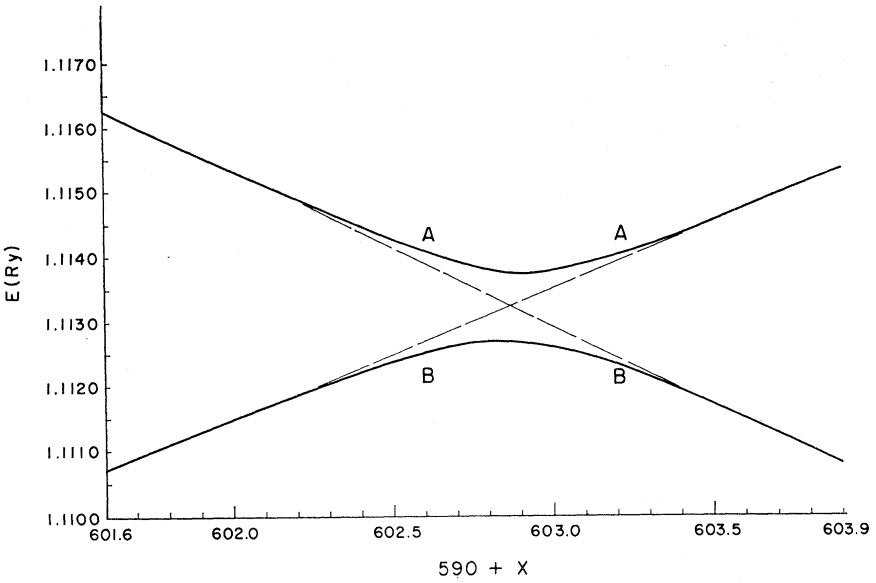

FIG. 3. An expanded scale graph of the region of Fig. 2 enclosed by the rectangle. The two energy bands are shown along the same line as in Fig. 2. On this expanded scale it is clear that the bands do not cross. The minimum energy separation is, however, quite small and is approximately 0.001 Ry.

We note from the denominator in Eq. (2.8) that the Green's functions for the nth band will tend to receive important contributions from those portions of the $E(\mathbf{k})$ surface which lie nearest the forbidden energy gap. For the conduction-band Green's functions we have employed a mesh of 60 general points in 1/48 of the zone. This mesh contains a reasonable number of sample points in the vicinity of the local minimum near the point X and the (100) axis.

For the valence bands a finer mesh containing 356 sample points in 1/48th of the Brillouin zone was employed in order to more adequately represent the important region near $\mathbf{k}=0$. In summary, the Green's functions were calculated for eight bands (four valence and four conduction) at eighteen different uniformly spaced energies within the forbidden gap. Calculations were done for lattice vectors, in units of $(2\pi/a)$, of the five types (000), (110), (200), (211), and (220).

VI. APPLICATION TO THE NEUTRAL SILICON VACANCY

We now turn to the numerical problem for the neutral silicon vacancy. As described in Secs. III and IV, the origin for the Bloch functions and hence for the defect problem is located midway between the two atoms in a silicon unit cell. The two atoms are then located at $\pm\mathbf{d}$. We then remove the atom at $+\mathbf{d}$ to create a vacancy. The vacancy potential is then of the form $V_d(\mathbf{r}-\mathbf{d})$. In Cartesian coordinates $\mathbf{d}$ is $(a/2)(\frac{1}{4},\frac{1}{4},\frac{1}{4})$ so that there are six group operations, each of the form $(\alpha|0)$, within the silicon space group which leave the vacancy invariant. The group of the defect potential is therefore C_{3v} (isomorphic to D_3). Table IV lists the operations and the character table.

Our choice of origin, which leads to C_{3v} symmetry, requires some comment. It is not the most obvious choice. If we were to take the origin at the site of the vacancy, the group of the defect potential would be T_d. We would then be able to reduce somewhat the number of independent matrix elements. However, a price would have to be paid. When the origin is at an atomic site, some of the matrix elements of the crystal potential are complex, and this means that the plane-wave expansion coefficients $b_n(\mathbf{k},\mathbf{K}_s)$ will be complex. Complex b's would greatly complicate our calculation of matrix elements, compensating for the reduction in number. In addition, the use of real b's greatly facilitates investigation of the transformation properties of wave functions.

No obvious relation between the expansion coefficients of the defect wave function using our origin is to be expected on account of the "hidden" T_d symmetry. This results because a shift of origin leads to substantially different Wannier functions: Each shifted Wannier function is a linear combination of all the original Wannier functions of the various bands. Thus, to make apparent the hidden T_d symmetry of the final wave function, we would have to obtain the complete defect wave function and perform a transformation. However, the partitioning technique described in Sec. II enables us to avoid computing the entire wave function, and we have used this option.

The matrix of the defect potential expressed on the basis of Wannier functions has been described in Secs.

TABLE IV. Character table for the group C_{3v} with operations expressed as substitutions on the Cartesian coordinates.

	xyz	zxy yzx	yxz zyx xzy
	E	$2C_3$	$3\sigma_v$
A_1	1	1	1
A_2	1	1	−1
A_3	2	−1	0

TABLE V. Column 1 contains the site group index. Column 2 lists the vectors in the three-site groups. Column 3 is the distance squared from a site $\mathbf{R}_n$ to the defect at $+\mathbf{d}$. $\mathbf{R}_n$ is a Wannier function lattice site vector. Column 4 is the decomposition ($\Lambda = P_1\Lambda_1 + P_2\Lambda_2 + P_3\Lambda_3$) of a reducible representation Λ into irreducible components $\Lambda_1, \Lambda_2, \Lambda_3$, for the site groups, and for Γ_1 and Γ_2' bands.

Site group	$\{\mathbf{R}_n\}$	$\lvert\mathbf{R}_n-\mathbf{d}\rvert^2$	Symmetries
1	(0 0 0)	$\frac{3}{16}$	$P_1=1, P_2=0, P_3=0$
2	(1 1 0)	$1\frac{3}{16}$	$P_1=1, P_2=0, P_3=1$
	(1 0 1)		
	(0 1 1)		
3	(1 $\bar{1}$ 0)	$2\frac{3}{16}$	$P_1=1, P_2=1, P_3=2$
	($\bar{1}$ 1 0)		
	(0 1 $\bar{1}$)		
	(0 $\bar{1}$ 1)		
	(1 0 $\bar{1}$)		
	($\bar{1}$ 0 1)		

III and IV. We note that as a consequence of the manner in which we have chosen to define our energy bands, all Wannier functions for any given band are either purely real or purely imaginary (Table III). Thus, after performing a simple transformation on the matrix, only real arithmetic need be considered.

A total of eight bands and ten lattice site vectors have been considered. The ten lattice vectors and the site groups to which they belong are given in Table V. As the problem is now phrased, the defect potential matrix would be of order 80×80; Hermiticity reduces this to a consideration of 3240 elements. The labor that would be involved in evaluating this many matrix elements would be entirely prohibitive. Fortunately, the number of matrix elements which must be calculated may be reduced by about a factor of 9 by an application of the group theory developed in Sec. III. The actual number of matrix elements which must actually be computed is smaller than 3240 due to three related circumstances. First, matrix elements may not be independent; secondly, certain matrix elements can be shown to be zero; and thirdly, although independent and nonzero, certain matrix elements will not contribute to the symmetrized combinations of matrix elements which we shall subsequently form. The first two cases are consequences of Eq. (3.20) and the last case is a consequence of Eq. (3.21).

TABLE VI. The fourteen types of potential matrix elements which must be considered for the three-site group problem involving ten lattice site vectors. For particular bands s and t certain of the above matrix elements are zero as explained in the text.

Group index	Matrix element	Band indices
1	$(s, 000\lvert V\rvert t, 000)$	$t\geq s$
2	$(s, 110\lvert V\rvert t, 110)$	$t\geq s$
3	$(s, 110\lvert V\rvert t, 101)$	$t\geq s$
4	$(s, \bar{1}10\lvert V\rvert t, \bar{1}10)$	$t\geq s$
5	$(s, \bar{1}10\lvert V\rvert t, \bar{1}01)$	$t\geq s$
6	$(s, \bar{1}10\lvert V\rvert t, 0\bar{1}1)$	$t\geq s$
7	$(s, \bar{1}10\lvert V\rvert t, 1\bar{1}0)$	$t\geq s$
8	$(s, \bar{1}10\lvert V\rvert t, 10\bar{1})$	$t\geq s$
9	$(s, \bar{1}10\lvert V\rvert t, 01\bar{1})$	$t\geq s$
10	$(s, 000\lvert V\rvert t, 110)$	t, s
11	$(s, 000\lvert V\rvert t, \bar{1}10)$	t, s
12	$(s, 110\lvert V\rvert t, \bar{1}10)$	t, s
13	$(s, 110\lvert V\rvert t, \bar{1}01)$	t, s
14	$(s, 110\lvert V\rvert t, 01\bar{1})$	t, s

When we apply Eq. (3.20) to determine the number of independent matrix elements for the ten site problem, we find that there are fourteen different site vector pairs which must be considered. These are listed in Table VI. For eight bands Table VI represents 644 matrix elements.

Using Eq. (3.20) several of the matrix elements can be shown to vanish by symmetry. In particular the matrix elements in groups 1, 2, and 10 vanish if s (or t) is Γ_1 or Γ_2' symmetry and t (or s) is Γ_2 or Γ_1' symmetry. Therefore matrix elements belonging to groups 1, 2, and 10 involving bands 4 and 8 vanish except for the band pairs (4,4), (4,8), and (8,8). Apparently no other matrix elements vanish identically.

No further simplifications can be achieved using single Wannier functions as basis functions. We therefore take as basis functions symmetrized combinations of Wannier functions constructed according to Eq. (3.21) and the discussion in Sec. III. The new basis functions are characterized by a band index, a site group index and are labeled according to the various irreducible representations of the defect potential. This labeling must of course include rows if the representation is degenerate and must include all independent combinations that can be formed. In particular we note from Table V that for site group 3 it is possible to form two linearly independent functions transforming according to Λ_3 and both must of course be included.

A preliminary examination of our numerical matrix indicated that the localized state, if one exists, would probably belong to the totally symmetric Λ_1 representation. We therefore limited our attention to this representation although we do not completely rule out the other representations from possible interest. All subsequent discussion and the numerical results we present will pertain to the Λ_1 representation.

Valence band 4 (Γ_1') and conduction band 8 (Γ_2) have small matrix elements of the potential and do not couple strongly to other bands. In particular, bands 4 or 8 are not at all coupled to each other or to the other six bands by site groups 1 or 2. This is a consequence of the fact that it is not possible to form any (nonvanishing) Λ_1 combinations of Wannier functions with site groups 1 or 2 for band 4 or 8. Site group 3 does provide some interband coupling but this is expected to be smaller than the couplings between the other six bands. It appears to be sufficient therefore to consider six bands only. For six bands and the Λ_1 representation Table VI contains 369 matrix elements which must be

evaluated for the three-site group problems. A restriction to 6 bands and site groups 1 and 2 only requires 99 matrix elements while the single-site group problem requires 21 matrix elements. It is seen that the number of matrix elements which must be calculated increases rapidly with increasing number of site groups.

Let us next describe briefly how the matrix elements themselves were evaluated. This portion of the program involved most of the numerical effort. Referring to Eq. (4.6) it is seen that a double integral over the entire Brillouin zone is required and hence represents a major undertaking. This has been reduced considerably by transforming the problem so that we could work only over the subzone. We shall not go through the transformation because it is quite lengthy, but merely indicate what was done.

We note that each of the single integrals over the Brillouin zone can be written as an integral over the subzone and a sum over group operations. Thus two separate and independent sums over group operations are involved. Actually one of these two sums can be written in closed form. The considerably simplified computer program resulting from this fortunate circumstance saves about a factor of 20 in computer time over the program in which the two sums were both programmed.

In calculating the matrix elements we were able to include a maximum of ten points in 1/48 of the zone. Even with this limited number of points, each potential matrix element required about 6 min of high-speed computing time. Comparisons were made between values of some potential matrix elements evaluated with only three points in the subzone and the values obtained using the full ten points in the subzone. The values were in much closer agreement than would have been expected. It is tempting to speculate that the group-theoretic transformation around the zone plays a more important role in the matrix elements of the potential than does the variation within any individual subzone. At any rate, the problem has been handled in as much detail as appears practical at this time.

We then proceeded to combine the matrix elements which had been calculated with the previously calculated Green's functions in order to form the determinant in Eq. (2.13). From this the zeros of the determinant were obtained and the bound states located. The results and analysis thereof is contained in the next section.

VII. RESULTS

In this section we present our results for the bound state of the neutral vacancy in silicon. The root of the determinantal equation, Eq. (2.11), which lies in the band gap is the energy of this state. We will consider how the energy of this state depends on the size of the determinant, that is, on the number of bands and the number of site groups which are included. This amounts to an investigation of the convergence of our method. It is important to recall, however, that the method we use yields an exact eigenvalue of the Hamiltonian for a given potential matrix V.

In addition, it turns out to be quite instructive to introduce another degree of freedom into our problem. We multiply the potential matrix V by a parameter λ without changing the elements in any other way. This allows us to study, for a potential matrix of given dimension, the dependence of the bound-state energy on the strength of the potential. Thus, we consider instead of Eq. (2.11) a modified equation

$$\det(I-\lambda GV)=0. \tag{7.1}$$

The root of this equation in the band gap will be a function of λ, and we will describe its behavior below. The symmetry analysis is, of course, unchanged.

To determine the convergence of the method with respect to the number of site groups employed we have determined the bound-state energy for the single-site group problem (1-site vector), and two-site group problem (4-site vectors), and the three-site group problem (10-site vectors). The convergence of the method with respect to the energy bands included was also investigated by deleting various bands from the numerical problem. It is possible in this manner to establish which energy bands are of primary importance in making up the defect state.

Table VII contains our results for the single-site problem with six bands included. They are shown as the curve labeled (1: 1,2,3,5,6,7) in Fig. 4. It is seen that a potential multiplier of about 1.6 is required to localize a bound state within the energy gap. Apparently the effective vacancy potential for this restricted problem is much too weak to produce a bound state. Table VIII contains the approximate partial eigenvector which is obtained at $E_b=0.06$ eV with a multiplier $\lambda=1.620$. The size of the components for bands 1 and 2 should be noted. Apparently bands 1 and 2 play an important role in the formation of the defect state. The eigenvector given in Table VIII is partial in the sense of Sec. II and represents effectively ϕ_N of Eq. (2.16). If the full eigenvector $\phi=(\phi_N,\phi_T)$ were to be obtained,

TABLE VII. Bound-state energy E_b(eV) as a function of the potential parameter λ for the single-site problem. The energy scale has been shifted so that the valence-band maximum is at 0.0 eV and the valence-band minimum at 0.92 eV in accord with Brust (Ref. 27). Our energy gap was approximately 0.05 eV larger than that of Brust. We have ignored this small difference.

	Six bands (1,2,3,5,6,7)		
λ	E_b	λ	E_b
1.60	0.03	1.90	0.55
1.620	0.06	1.95	0.63
1.65	0.11	2.00	0.71
1.70	0.19	2.05	0.78
1.75	0.285	2.10	0.85
1.80	0.38	2.15	0.91
1.85	0.46	...	...

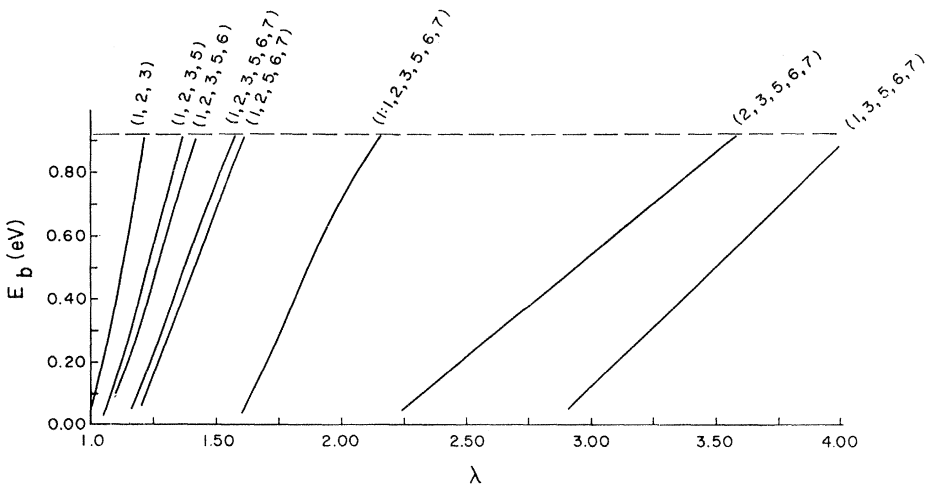

FIG. 4. The bound-state energy E_b as a function of the potential parameter λ for the single-site group problem and for the two-site group problem. For the two-site problem the results for a variety of band combinations are shown and the curves are labeled according to the bands included. For the single-site group problem only the results for the six-band combination are shown. This curve is labeled (1:1,2,3,5,6,7). The scale of energy employed places the valence-band maximum at 0.0 eV and the conduction-band minimum at 0.92 eV.

ϕ_N would be renormalized but the relative size of the individual components would be unaffected.

Table IX shows the results for the two-site group problem involving four-site vectors. A variety of different band combinations were investigated. Several features should be noted. The six bands result in Fig. 4 show that a bound state is localized in the gap for a potential multiplier of about 1.15. The multiplying factor of 1.15 might be crudely interpreted as a measure of the error in our calculation. A change of the order of 15% in the assumed vacancy potential might not be an unreasonable modification in view of the neglect of lattice relaxation in the neighborhood of the defect.

Next we note that as the various conduction bands are successively dropped out of the numerical problem, the curves of E_b versus λ shift toward the origin and if only valence bands (1,2,3) are considered, a bound state at about 0.06 eV above the valence band occurs. It is clear that the conduction bands act to inhibit the formation of a localized state, a situation in agreement with what one expects on the basis of perturbation theory when higher levels are included.

The change in E_b versus λ is considerably larger (as much as a factor of 30) when the various valence bands are dropped from the numerical problem. Deleting band 3 results in only a slight shift to larger λ. Dropping band 2 from the six-band problem causes a very large shift to higher λ with a value of about 3 required to produce a bound state. Dropping band 1 shifts the bound-state energy noticeably less than for band 2, but the shift is still large compared to the shift when band 3 was dropped. While we do not attach great significance to the bound states which result for large values of λ, it is clear that the valence bands play the predominant role in the formation of the defect state.

TABLE VIII. Approximate partial eigenvector for the single-site problem. λ=1.620, E_b=0.06 eV, and ΔE=0.0043 eV. The zeros of the numerical determinant are located by observing the point at which one of the eigenvalues of the determinant goes through zero. ΔE is the actual value of the eigenvalue used instead of zero.

Band	Value
1	0.4551
2	0.7100
3	−0.2570
5	0.3683
6	−0.1395
7	−0.2598

TABLE IX. Bound-state energy E_b as a function of the potential parameter λ for the two-site group problem involving four-site vectors. The zero of energy is such that $E_{\text{valence}}=0.0$ and $E_{\text{conduction}}=0.92$ eV.

λ	E_b(eV)	λ	E_b(eV)
3 bands (1,2,3)		5 bands (1,3,5,6,7)	
1.00	0.06	3.50	0.50
1.05	0.21	4.00	0.88
1.10	0.41		
1.15	0.63	5 bands (1,2,3,5,6)	
1.20	0.86		
4 bands (1,2,3,5)		1.10	0.10
		1.20	0.32
		1.30	0.61
1.05	0.03	1.40	0.86
1.10	0.14		
1.15	0.28	6 bands (1,2,3,5,6,7)	
1.20	0.43		
1.25	0.58		
1.30	0.73	1.16	0.05
1.35	0.87	1.17	0.07
		1.18	0.08
5 bands (2,3,5,6,7)		1.19	0.10
		1.20	0.12
		1.25	0.22
2.40	0.15	1.30	0.33
2.60	0.28	1.35	0.45
2.80	0.41	1.40	0.56
3.00	0.54	1.45	0.66
		1.50	0.77
5 bands (1,2,5,6,7)		1.55	0.86
		1.58	0.91
1.20	0.06		
1.40	0.47		
1.60	0.88		

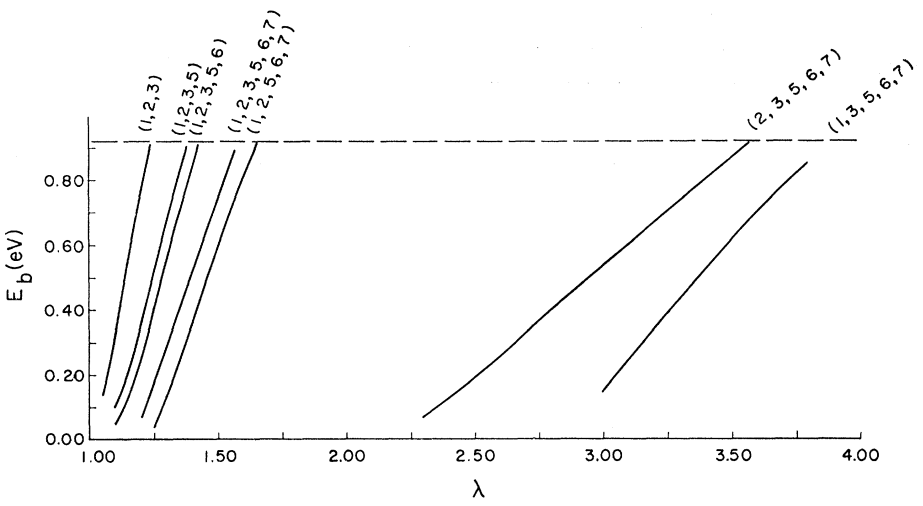

FIG. 5. The bound-state energy E_b as a function of the potential parameter λ for the three-site group problem employing ten lattice site vectors.

This is reasonable on intuitive grounds. What is somewhat surprising is the extreme importance of bands 1 and 2. Having obtained this result, an explanation can of course be given. In general, the matrix elements of the potential are larger for the lower-lying bands while the Green's functions are largest for those bands lying nearest the gap. The coupling between these two numerical situations selects band 2 as being of prime importance in localizing an energy level within the energy gap. Table X shows the partial eigenvector for the two-site group problem. The size of the components pertaining to bands 1 and 2 is again noted and a general decrease in going from site group 1 to site group 2 is observed.

Table XI contains the results for the three-site group problem, again for a variety of band combinations. These are shown in Fig. 5. The changes brought about by increasing the number of site vectors from four to ten are small. In particular E_b versus λ for bands (1,2,3,5,6,7) is very nearly identical for the two- and three-site group problems. For three-site groups, the (1,2,3) band combination apparently produces a bound state for $\lambda=1$ that lies somewhere between 0 and 0.05 eV above the valence band. A major conclusion is that the two-site group bound-state energy is changed very little in more than doubling the number of lattice sites included in the calculation. Apparently the convergence of the method is satisfactory.

Table XII contains partial eigenvectors for three different values of E_b. These were obtained in order to see what changes take place as one moves from an

TABLE X. Approximate partial eigenvector for the two-site group problem. ($\lambda=1.165$, $E_b=0.06$ eV, $\Delta E=+0.0002$ eV.)

Site group	Band	Component	Value
1	1	1	0.3742
1	2	2	0.6215
1	3	3	−0.2106
1	5	4	0.2927
1	6	5	−0.1202
1	7	6	−0.2141
2	1	7	0.1788
2	2	8	0.4734
2	3	9	−0.1284
2	5	10	0.0389
2	6	11	−0.0426
2	7	12	−0.0868

TABLE XI. Bound-state energy E_b as a function of the potential parameter λ for the three-site group problem with ten-site vectors. The zero of energy is such that $E_{\text{valence}}=0.0$ eV and $E_{\text{conduction}}=0.92$ eV.

λ	E_b(eV)	λ	E_b(eV)
3 bands (1,2,3)		5 bands (1,2,3,5,6)	
1.05	0.14	1.10	0.05
1.10	0.332	1.15	0.13
1.15	0.553	1.20	0.26
		1.25	0.41
4 bands (1,2,3,5)		1.30	0.55
		1.35	0.70
		1.40	0.83
1.10	0.10		
1.15	0.21	5 bands (2,3,5,6,7)	
1.20	0.36		
1.25	0.52		
1.30	0.68	2.30	0.07
1.35	0.83	2.40	0.13
		2.50	0.19
5 bands (1,3,5,6,7)		2.60	0.26
		2.70	0.33
		2.80	0.40
3.00	0.15	2.90	0.47
3.20	0.34	3.00	0.53
3.60	0.70	3.10	0.60
3.80	0.85		
		6 bands	
5 bands (1,2,5,6,7)		(1,2,3,5,6,7)	
1.25	0.04	1.20	0.0725
1.30	0.13	1.25	0.170
1.35	0.24	1.30	0.285
1.40	0.36	1.40	0.517
1.45	0.48	1.45	0.633
1.50	0.59	1.50	0.74
1.55	0.71		
1.60	0.81		
1.65	0.91		

TABLE XII. Approximate partial eigenvectors for the three-site group problem at three different energies E_b within the forbidden-energy gap.

Site group	Band	Component number	$E_b=0.06$ $\lambda=1.193$ $\Delta E=-0.0013$	$E_b=0.46$ $\lambda=1.376$ $\Delta E=-0.0023$	$E_b=0.86$ $\lambda=1.558$ $\Delta E=+0.0008$
1	1	1	0.3567	0.3879	0.3835
1	2	2	0.5784	0.5776	0.5430
1	3	3	−0.2543	−0.2155	−0.1848
1	5	4	0.2694	0.3596	0.4602
1	6	5	−0.1016	−0.1232	−0.1370
1	7	6	−0.2010	−0.2500	−0.2862
2	1	7	0.1765	0.1933	0.1922
2	2	8	0.4099	0.3728	0.3333
2	3	9	−0.1472	−0.0953	−0.0730
2	5	10	0.0269	0.0389	0.0507
2	6	11	−0.0478	−0.0556	−0.0577
2	7	12	−0.0818	−0.1010	−0.1139
3	1	13	−0.0449	−0.0475	−0.0456
3	2	14	0.1953	0.1291	0.0937
3	3	15	−0.2770	−0.1983	−0.1561
3	5	16	0.0502	0.0587	0.0555
3	6	17	−0.0096	−0.0109	−0.0107
3	7	18	0.0149	0.0202	0.0265

energy near the valence band ($E_b=0.06$ eV) to an energy near the center of the gap ($E_b=0.46$ eV) and finally to an energy near ($E_b=0.86$ eV) the conduction band. Only small changes are evident. The state vector appears to be dominated by the valence-band contributions for all energies in the gap.

VIII. CONCLUSIONS

Although one expects that the most appropriate way of treating localized defects is to employ localized basis functions, i.e., Wannier functions, these functions in the past have been employed principally in formal arguments. Their utilization in realistic calculations has not been attempted, partially because of the presumed ambiguities concerning the phase of the Bloch functions, and partially because of the computational labor involved. A more difficult question concerns the proper and useful definition of an energy band.

One important conclusion to emerge from this investigation is that these problems can be solved. It is possible to use Wannier functions in practical computations concerning localized defects in crystals, and thereby obtain numerical results in moderate agreement with experimental results in the one important case investigated in detail.

The theoretical methods we have developed appear applicable to the calculation of formation and migration energies of localized defects. The formation energy may be considered as the change in the one-electron energy of the system and is obtained from an expression involving a determinant identical in form to the determinant in Eq. (2.11). For problems of this class, as opposed to the bound-state energy problem, the Green's functions must be evaluated for energies lying within the valence bands. The evaluation and the resulting Green's functions are consequently more complicated. However, in the case of the formation energy of the silicon vacancy, the numerical values of the matrix elements of the potential previously calculated would be directly applicable. Other logical extensions of the present method include the silicon divacancy, the vacancy-oxygen pair and the vacancy-phosphorous pair. We hope that the results of our present work will open the door for many further investigations of the properties of localized defects.

The application of the general method we have developed to the vacancy, the particular defect chosen for detailed calculation, has shown that a localized state is produced in which low-lying energy bands—in particular, the lowest two valence bands—and interband couplings are important. This contrasts with the results for the shallow donor and acceptor states which have been studied extensively by different methods in the past, for which only the band or bands closest in energy to the impurity state are important, and for which, except in cases of band degeneracies, interband couplings may be neglected.

The method we have developed appears to be the most powerful approach presently available for the investigation of localized defects. It should yield the most realistic results since, in contrast to methods previously employed by others, it incorporates detailed information concerning the energy band structure of the perfect crystal.