[27] A. Many, J. Shappir, and U. Shaked, Surface Sci. 14, 156 (1969).

[28] A. Many, Y. Goldstein, and N. B. Grover, *Semiconductor Surfaces* (North-Holland, Amsterdam, 1965).

[29] D. R. Frankl, *Electrical Properties of Semiconductor Surfaces* (Pergamon, New York, 1967).

[30] J. D. Levine, J. Vac. Sci. Technol. 6, 549 (1969).

[31] S. G. Davison and J. D. Levine, in *Solid State Physics*, edited by H. Ehrenreich, F. Seitz, and D. Turnball (Academic, New York, 1970), Vol. 25, p. 7.

[32] R. K. Swank (private communication).

[33] R. Nosker (private communication).

[34] S. Kurtin, T. C. McGill, and C. A. Mead, Phys. Rev. Letters 22, 1433 (1969).

[35] H. C. Gatos and M. C. Lavine, J. Electrochem. Soc. 107, 433 (1963).

[36] I. Bartoš, Surface Sci. 15, 94 (1969).

[37] C. M. Chaves, N. Majlis, and M. Cardona, Solid State Commun. 4, 271 (1966).

PHYSICAL REVIEW B VOLUME 2, NUMBER 8 15 OCTOBER 1970

Surface States for a Molecular Orbital Model Hamiltonian: Resolvent Methods and Application to Zinc Blende

Smith Freeman
RCA Laboratories, Princeton, New Jersey 08540

(Received 16 March 1970)

A careful discussion is given of the application of the resolvent or Koster-Slater method to a surface-state problem described by a model Hamiltonian of the tight-binding type. Because of a special symmetry, peculiar to the surface-state problem, it is possible to reduce the degree of the characteristic determinant by a factor of 2. The usual, "truncation" method of doing this is shown to sometimes lead to spurious solutions. The origin and properties of these are determined and a prescription for recognizing them presented. An alternative, "pre-severed" approach to the reduction of the determinant is described. This latter is less useful for numerical calculation, but has advantages in obtaining analytical results. The general form of the surface-state wave function is derived from the integral representation of the resolvent. It proves possible to characterize this form using only simple properties of the bulk band structure as represented by the model Hamiltonian. This then permits the popular "ansatz" method of treating the surface states to be routinely used in problems having an arbitrary degree of complexity. The results of the general discussion are applied to determine the surface states associated with the (110) cleavage face of a semiconductor having the zinc-blende structure. The "pre-severed" resolvent method is used to derive some preliminary analytical results, and a program is set up for the numerical calculation of the surface-state properties for more realistic models of the surface perturbation.

I. INTRODUCTION

The (110) face of a II-VI or III-V semiconductor having the zinc-blende structure is an especially interesting surface for theoretical study since it is found experimentally *not* to be reconstructed, i.e., the physical surface retains the full translation symmetry of a parallel bulk plane. This paper, together with its companion work,[1] describes the results of a theoretical investigation of the properties of the electronic surface states associated with such a surface. We have used the molecular orbital (MO) scheme[2] to describe the one-electron states of the crystal and have treated the effects of the surface perturbation via the resolvent technique. This present paper sets up the required formal machinery and applies it to the system of interest. The appropriate matrix elements of the resolvent are here obtained; some results which can be obtained analytically are derived, and the method used in the remaining numerical calculation is outlined. The subsequent paper[1] will display the surface bands derived by numerical calculation for some physically interesting cases. Threshold conditions, which can be obtained by several routes, will be studied there. A discussion will be given there of the physical justification for the particular MO representations chosen to describe the bulk band structure and the surface perturbation, and the results will be compared to experiment.

Section II of the present paper contains a complete and elementary description of the methods used. Various somewhat intricate points, as well as some useful extensions of the basic method which are not required in the present work, are pursued in the appendices. Section III then applies the method to the system of interest. Portions of Sec. II describe methods which are not new and have been used by other authors.[2,7,13,17,18,20,22] However, the central technique used here to obtain the surface electronic states has not been developed or employed before. In addition, the various preliminaries and subordinate techniques characteristic of the surface-state problem, which are usually glossed over, are here described in a systematic fashion. It seems desirable that such a pedagogically oriented treatment should be available. A number of alternative methods of solving the equations of motion for a surface one-electron state are ultimately obtained. They are all exact within the context of the MO model Hamiltonian and thus give the same answer.[3] However, the interrelations among these different approaches have not previously been displayed, as can be done here. Finally, a number of rather central results are derived here for the first time. Our systematic presentation permits these new features to be fitted into the development of the subject in their proper logical context.

In detail, the organization of Sec. II is as follows: II A the tight-binding-model representation of the bulk states is set up. II B the surface is introduced. II C the translation symmetry parallel to the surface is exploited to reduce the dimensionality of the problem. II D a general discussion of the resolvent method is given. II E the resolvent method is applied to the task of obtaining the surface electronic states, and the special features of this particular application are studied in detail. II F the integral representation of the resolvent is derived and used to determine the general form of the surface-state wave function. The subsections correspond to this division of the subject.

II. FORMAL TREATMENT

A. MO Description of Bulk Bands

In this paper we shall use the MO approach to describing the surface electronic properties of a solid.[2] This technique achieves a marked simplification of the equations of motion by truncating the band structure and the surface perturbation, retaining only the part which can be described in terms of a limited basis of states. The resolvent or Koster-Slater method can then be employed to solve the equations of motion without further approximation. Although this approach is well known and widely used, it will be described here in some detail. We hope thus to make this work, together with its companion paper,[1] reasonably self-contained, and also to bring out certain features of the method which are not widely appreciated. Some new formal results will be given; for the most part, their detailed derivation is relegated to Appendix A. Appendix B extends the method to the treatment of reconstructed surfaces and surface defect states.

The essential feature of the MO description of the electronic properties of a solid is that the one-electron states are assumed to arise as linear combinations of a limited basis of atomic orbitals present on the various ions in a unit cell. This is analogous to the fashion in which MO's are built up out of atomic orbitals in the treatment of the one-electron states of a molecule. The Hamiltonian is given by specifying its various diagonal and off-diagonal elements in terms of these localized basis states. The diagonal elements are commonly called "Coulomb integrals" and the off-diagonal ones "transfer" or "resonance integrals." We shall use a MO *model* Hamiltonian approach following the example of Levine and Davison.[2] That is, the Hamiltonian will be specified *solely* through these matrix elements between localized "atomic orbital" states. The values of these matrix elements will be obtained by fitting various known properties of the bulk band structure rather than as explicitly calculated integrals involving an actual one-electron Hamiltonian and some sort of atomic orbital wave functions.

Let us consider in more detail the description of the band structure of the bulk crystal – a perfect crystal infinite in all directions and thus free of surfaces. The basis states will be described in the Dirac notation $|\vec{m}\beta\rangle$; they are taken to be orthonormal. Here $\vec{m}$ specifies a particular unit cell, and β enumerates the various basis states associated with that cell; β runs from 1 to ν. This abstract notation is appropriate since the dynamical problem is specified by the matrix elements of the Hamiltonian in this basis, and the spatial wave functions of the orbitals are never employed.[4] The vector cell index $\vec{m}$ is defined: $\vec{m} = m_1\vec{a} + m_2\vec{b} + m_3\vec{c}$, where $\vec{a}$, $\vec{b}$, and $\vec{c}$ are the primitive translation vectors of the lattice. In the usual fashion, $\vec{m}$ represents the translation which carries the unit cell containing the origin into that characterized by the triple of integers (m_1, m_2, m_3). We now specify the bulk Hamiltonian $\mathcal{H}_0$:

$$\langle\vec{m}\beta|\mathcal{H}_0|\vec{l}\alpha\rangle = W_{\beta\alpha}(\vec{m}-\vec{l}) \quad . \tag{1}$$

As indicated on the right, these matrix elements depend only on the difference, $\vec{m}-\vec{l}=\vec{n}$. This is an expression of the translation symmetry of the bulk crystal, i.e., of $\mathcal{H}_0$. It is essential to our treatment of the surface problem that $W_{\beta\alpha}(\vec{n})$ vanishes unless $\vec{n}$ is one of a finite set of near-neighbor displacements. The nonzero values of $W_{\beta\alpha}(\vec{n})$ are picked to fit the known band structure.

The eigenvectors of $\mathcal{H}_0$ satisfy

$$\mathcal{H}_0\psi = E\psi \quad , \tag{2}$$

where ψ stands for a column vector with components $\psi(\vec{m}\beta)$:

$$\psi = \sum_{\vec{m}\beta} |\vec{m}\beta\rangle\psi(\vec{m}\beta) \quad . \tag{3}$$

The translation symmetry of $\mathcal{H}_0$ permits us to characterize these solutions by a wave vector $\vec{k} = k_1\vec{a}^* + k_2\vec{b}^* + k_3\vec{c}^*$ and a band index ρ, running from 1 to ν. Here $\vec{a}^*$, $\vec{b}^*$, and $\vec{c}^*$ are basis vectors for the reciprocal-lattice dual to $\vec{a}$, $\vec{b}$, and $\vec{c}$. Each component k_i will be assumed to lie in the range $-\pi < k_i \le \pi$. This defines a primitive cell of the reciprocal lattice which is not the first Brillouin zone (FBZ). We shall seldom have occasion to employ the FBZ. In terms of the local basis $|\vec{m}\beta\rangle$, the components of such a solution $\psi_{\vec{k}\rho}$, have the form

$$\psi_{\vec{k}\rho}(\vec{m}\beta) = e^{i\vec{k}\cdot\vec{m}} e_\beta(\vec{k},\rho) \quad . \tag{4}$$

Here $e_\beta(\vec{k},\rho)$ is determined as a normalized solution of the ν-dimensional eigenvalue equation,

$$\sum_{\alpha=1}^{\nu} X_{\beta\alpha}(\vec{k}) e_\alpha(\vec{k},\rho) = e_\beta(\vec{k},\rho)\lambda_\rho(\vec{k}) \quad . \tag{5}$$

The ν-dimensional matrix $\overleftrightarrow{X}(\vec{k})$ is just the Fourier transform of $\overleftrightarrow{W}(\vec{n})$:

$$X_{\beta\alpha}(\vec{k}) = \sum_{\vec{n}} e^{-i\vec{k}\cdot\vec{n}} W_{\beta\alpha}(\vec{n}) \quad . \tag{6}$$

The eigenvalues $\lambda_\rho(\vec{k})$ represent the energy bands of the bulk, and the vectors $\psi_{\vec{k}\rho}$ are the corresponding band states. It is worth noting explicitly that Schrödinger's equation, which provides the exact description of the one-electron states, is a differential equation, and its solutions form a complete set. Any arbitrary form of spatial variation of a wave function can thus be described. Corresponding to this feature is the necessary existence of an infinite number of bands of solutions. The MO approximation replaces the differential equation by a difference equation based on a constrained spatial variation for the wave functions.[5] This in turn leads to the presence of only a finite number of bands of solutions. The number ν of basis states $|\vec{m}\beta\rangle$ used to describe the spatial variation of the wave function within a unit cell is just equal to the number of bands. The constrained variation implicit in the finite value of ν is central to the resolvent method.

B. Introduction of Surface

We must now proceed from the description of the bulk to the description of a crystal with a surface. The basis states $|\vec{m}\beta\rangle$ can still be used to define the Hamiltonian and to expand its eigenfunctions. Now, however, some of these correspond to sites "inside" the crystal and some to sites "outside." This will be reflected in the matrix elements of the Hamiltonian $\langle\vec{m}\beta|\mathcal{H}|\vec{l}\alpha\rangle$, which are specified in such a way as to represent the surface. The usual way in which this is done, starting from the infinite crystal characterized by $\mathcal{H}_0$, is to sever the two halves of the crystal on opposite sides of some crystallographic plane. One thus converts an infinite crystal into two semi-infinite crystals, each of which has a surface, as desired. In particular, one includes a potential $\mathcal{V}_S$ in $\mathcal{H}$ which just cancels out those resonance integrals in $\mathcal{H}_0$ connecting the two halves of the crystal. Then $\mathcal{H}_0 + \mathcal{V}_S$, unlike $\mathcal{H}_0$ alone, is block diagonal in subspaces referring to the "left" and "right" half-crystals, as we shall call them. The two half-crystals will be essentially identical if the surface plane is a reflection plane or a glide plane, otherwise the two halves represent two different systems which can be treated simultaneously.

Let us define a set S_0 of left-side surface orbitals consisting of those basis states $|\vec{m}\beta\rangle$, which are associated with the left half-crystal, but to which $\mathcal{H}_0$ assigns nonzero transfer integrals to basis states on the right. The corresponding set of right-side surface orbitals will be called S_1. According to the scheme described, $\mathcal{V}_S$ has matrix elements only between S_0 and S_1 states, and its matrix elements between such states are just the negatives of the corresponding matrix elements of $\mathcal{H}_0$. If we introduce a projection operator P_0, for the states S_0 and another P_1 for the states S_1, then the definition of $\mathcal{V}_S$ can be conveniently summarized:

$$\mathcal{V}_S = -\left(P_0\mathcal{H}_0P_1 + P_1\mathcal{H}_0P_0\right) \quad . \tag{7}$$

Now, in general, we will take

$$\mathcal{H} = \mathcal{H}_0 + \mathcal{V}_S + \mathcal{V}_T \quad , \tag{8}$$

where $\mathcal{V}_T$ provides for changes in the Coulomb and transfer integrals in the vicinity of the surface. Like $\mathcal{H}_0 + \mathcal{V}_S$, $\mathcal{V}_T$ is block diagonal in left and right half-spaces. It might represent, for example, the changes in the Madelung energies of the surface ions or changes in the transfer integrals resulting from a distortion of the surface layers compared to those in the bulk. Surface states which result simply from severing the crystal without the presence of a surface perturbation $\mathcal{V}_T$ will be called Shockley states. When it is desired to emphasize the presence of $\mathcal{V}_T$, a surface state will be called a Tamm state. This terminology is consistent with the customary use of these terms.[6]

The method of decoupling described above is employed in this work and in Ref. 1, but it is by no means the only one possible. Instead of severing the bonds connecting the two halves via $\mathcal{V}_S$, we could have introduced a potential barrier. Thus, consider $\mathcal{H}_0 + \mathcal{V}_B$, where

$$\mathcal{V}_B = (1/\epsilon)P_0 \; ; \tag{9}$$

the effect of $\mathcal{V}_B$ is to increase the Coulomb integrals of the basis states in S_0 by ϵ^{-1}. As $\epsilon \to 0+$, this amounts to an infinite potential barrier, which will prevent an electron in the right half-crystal from entering the left. The electronic states of the right half-crystal resulting from $\mathcal{H}_0 + \mathcal{V}_B$ must be identical to those resulting from $\mathcal{H}_0 + \mathcal{V}_S$. This fact will be exploited in Appendix A. It has been pointed out[7] that a *finite* and position-dependent potential barrier would be the physically most plausible representation of the surface. Unfortunately, such a scheme becomes rapidly more difficult to treat as the thickness of the effective surface region increases. Finally, it should be mentioned that, in any case, there is no reason why the severing of the two half-crystals must take place across a lattice plane. Instead, one might leave certain unit cells or individual ions sticking out farther than others, thus, representing a reconstructed or faceted surface. This possibility will be explored further in Appendix B and in a subsequent paper.[8]

We must now find the eigenvalues of the Hamiltonian $\mathcal{H}$ defined in Eq. (8). Solely as a matter of convenience, it will be supposed that the matrix elements of $\mathcal{V}_T$ are nonzero only within S_0 and S_1.[9] As a first step, let us exploit the symmetry properties of $\mathcal{H}$. Of course, the translational symmetry of $\mathcal{H}_0$ perpendicular to the surface plane has been lost, but that parallel to it remains. The symmetry of $\mathcal{V}_T$ parallel to the surface is assumed to be as high as that of $\mathcal{H}_0$, otherwise we would be dealing with a reconstructed surface. For such a surface the translational periods would be increased, and the low-energy electron-diffraction (LEED) pattern would show fractional order spots.

C. Exploitation of Surface Symmetry

We can choose two of the three primitive translation vectors $\vec{a}$ and $\vec{b}$ to be parallel to the surface plane. The third, $\vec{c}$, is taken to point into the right half-space. Labeling the various primitive cells in the fashion $m_1\vec{a}+m_2\vec{b}+m_3\vec{c}$, the right half-crystal is defined to include $m_3 \geq 1$ and the left $m_3 \leq 0$. For an unreconstructed surface, this will always be possible when the primitive cell is appropriately chosen (We shall often drop the subscript 3 from m_3 and its conjugate variable k_3.) The reciprocal-lattice vector $\vec{c}^*$ is now perpendicular to the surface plane. By assumption, $\mathcal{H}$ is unchanged by a translation through $m_1\vec{a}+m_2\vec{b}$. It follows that the components of the wave vector in the $\vec{a}^*$ and $\vec{b}^*$ directions are conserved, and the eigenstates of the system can be classified according to their values of $\vec{k}_\parallel = k_1\vec{a}^* + k_2\vec{b}^*$. Those states belonging to a particular value of $\vec{k}_\parallel$ can be expanded in terms of the "layer orbitals" $|\vec{k}_\parallel; m\beta\rangle$, defined

$$|\vec{k}_\parallel; m\beta\rangle = \sum_{m_1, m_2} e^{i\vec{k}_\parallel\cdot\vec{m}}\,|\vec{m}\beta\rangle \quad . \tag{10}$$

An eigenstate $\Psi(\vec{k}_\parallel)$ has components $\Psi(\vec{k}_\parallel; m\beta)$ along these basis elements:

$$\Psi(\vec{k}_\parallel) = \sum_{m\beta} |\vec{k}_\parallel; m\beta\rangle\Psi(\vec{k}_\parallel; m\beta) \quad . \tag{11}$$

The eigenvalue equation now takes the form

$$\mathcal{H}(\vec{k}_\parallel)\Psi(\vec{k}_\parallel) = E\Psi(\vec{k}_\parallel) \quad . \tag{12}$$

For states of fixed $\vec{k}_\parallel$, the Hamiltonian has reduced to $\mathcal{H}(\vec{k}_\parallel)$, which has the matrix elements

$$\langle\vec{k}_\parallel; m\beta|\mathcal{H}(\vec{k}_\parallel)|\vec{k}_\parallel; l\alpha\rangle = \sum_{m_1, m_2} e^{-i\vec{k}_\parallel\cdot(\vec{m}-\vec{l})}\langle\vec{m}\beta|\mathcal{H}|\vec{l}\alpha\rangle \quad . \tag{13}$$

The relation between $\mathcal{H}$ and $\mathcal{H}(\vec{k}_\parallel)$ can be written explicitly:

$$\langle\vec{k}'_\parallel; m\beta|\mathcal{H}|\vec{k}_\parallel; l\alpha\rangle = (2\pi)^2\delta^{(2)}(\vec{k}'_\parallel - \vec{k}_\parallel) \times\langle\vec{k}_\parallel; m\beta|\mathcal{H}(\vec{k}_\parallel)|\vec{k}_\parallel; l\alpha\rangle \quad . \tag{14}$$

Replacing $\mathcal{H}$ by $\mathcal{H}(\vec{k}_\parallel)$ and regarding $|\vec{k}_\parallel; m\beta\rangle$ as normalized to unity amounts to factoring out the δ-function normalization of the surface states corresponding to the parallel degrees of freedom m_1 and m_2. The operators $\mathcal{H}_0$, $\mathcal{V}_S$, and $\mathcal{V}_T$ have analogous reduced forms within the restricted set of states having fixed $\vec{k}_\parallel$. Similarly, in this context, the projection operators, P_0 and P_1, will be considered to project on the appropriate, finite collection of layer orbitals.

The surface states are describable in terms of the two-dimensional Brillouin zone of the wave vector $\vec{k}_\parallel$. The equation of motion, (12) plus (13), depends on the coordinate $m\beta$ only and is simply parameterized by $\vec{k}_\parallel$. It describes an effectively one-dimensional problem, and we shall often drop explicit mention of the label $\vec{k}_\parallel$.[10] Note that $\mathcal{H}_0(\vec{k}_\parallel)$, which is defined in terms of the matrix elements of $\mathcal{H}_0$ by an equation of the form of (13), describes the bulk states of fixed $\vec{k}_\parallel$. It gives rise to the energy bands and corresponding bulk eigenstates associated with a one-dimensional transection at fixed $\vec{k}_\parallel$ through the three-dimensional bands of the bulk. The sets $S_0(\vec{k}_\parallel)$ and $S_1(\vec{k}_\parallel)$ are now *finite* collections of layer orbitals, $|\vec{k}_\parallel; m\beta\rangle$, and the matrices $\mathcal{V}_S(\vec{k}_\parallel)$ and $\mathcal{V}_T(\vec{k}_\parallel)$ have nonzero entries only within the finite dimensional subblock whose rows and columns refer to these finite sets.[11]

D. Resolvent Method - General

The states of principal interest here are those for which the electron remains localized in the vicinity of the surface. In terms of the effective one-dimensional problem, such a state is a normalizable bound state – in contrast to the plane-wave eigenstates of $\mathcal{H}_0(\vec{k}_\parallel)$. Now the solutions of Eq. (12) are of two types: (1) plane-wave-like states which scatter off the surface but whose energy is unaffected thereby and thus lies within one of the continuous bands of energy eigenvalues of $\mathcal{H}_0(\vec{k}_\parallel)$ and (2) states whose energy lies outside these bands. These latter states are necessarily normalizable; they are surface bound states. Moreover, barring additional unexploited symmetries of the Hamiltonian $\mathcal{H}(\vec{k}_{\parallel})$ these will be the only such localizable states. Our task is thus to find those solutions of Eq. (12) for which E lies outside the transect of the bulk bands at fixed $\vec{k}_{\parallel}$. The technique we shall use is the so-called resolvent or Koster-Slater method. In principal, it permits the problem to be reduced to the evaluation of some integrals and solution of an algebraic equation. A general description of the method can be found in the original paper of Koster and Slater.[12] Its application to surface-state problems is described by Koutecký,[7,13] among others.

The basic problem is simply stated. The Hamiltonian $\mathcal{H}$ is the sum of $\mathcal{H}_0$ and $\mathcal{V} = \mathcal{V}_S + \mathcal{V}_T$. The bulk term $\mathcal{H}_0$ takes a simple form in a plane-wave representation, in which it reduces to block diagonal form. The remaining diagonalization problem (5) is ν dimensional and thus tractable. On the other hand, $\mathcal{V}$, which represents the effect of the surface, has a simple form only in a position basis, in which it vanishes except for a single finite-dimensional subblock. The over-all Hamiltonian $\mathcal{H}$ has neither of these simplifying properties; yet we must exploit both to solve the problem. The technique is very simple. Since we are seeking a solution whose energy E lies outside the effective one-dimensional energy bands, the operator $E - \mathcal{H}_0$ has no zero eigenvalues and therefore its inverse exists. This inverse is the so-called resolvent and will be written here $\mathcal{R}(E)$. As will be seen below, the block diagonal form assumed by $\mathcal{H}_0$ in a plane-wave representation permits us to explicitly evaluate the necessary matrix elements of $\mathcal{R}(E)$. We now write Eq. (12) in the form

$$(E - \mathcal{H}_0)\Psi = \mathcal{V}\Psi \tag{15}$$

and multiply through by the inverse of $(E - \mathcal{H}_0)$ to obtain

$$\Psi = \mathcal{R}(E)\mathcal{V}\Psi \quad . \tag{16}$$

Now $\mathcal{V}$ has matrix elements only within the rows and columns referring to the finite set of surface layers, $S \equiv S_0 + S_1$. Defining the projection operator P corresponding to S, $P = P_0 + P_1$, we can write those components of Ψ within the set S as $P\Psi$. Clearly, only the finite dimensional vector $P\Psi$ is required to evaluate the right-hand side of Eq. (16). Using Eq. (16) to evaluate $P\Psi$ itself, we find that it must satisfy the equation

$$P\Psi = P\mathcal{R}(E)\mathcal{V}(P\psi) \quad . \tag{17}$$

Writing N to represent the dimensionality of S_0 and of S_1, assumed to be the same,[14] this is seen to be a family of $2N$ linear equations for the $2N$ components of $P\Psi$. We know that a necessary and sufficient condition for a solution of such a system to exist is the vanishing of the determinant

$$\Delta(E) \equiv \det P[1 - \mathcal{R}(E)\mathcal{V}] = 0 \quad . \tag{18}$$

Here, 1 stands for the unit matrix. We have introduced a notation here which will be frequently employed below: The determinant is to be taken of that finite-dimensional matrix obtained by restricting the rows and columns of the operator $[1 - \mathcal{R}(E)\mathcal{V}]$ to the subspace projected on by P. Equation (18) is the fundamental equation of the Koster-Slater method. It states that the surface-state energies are just the values of E for which $\Delta(E)$ vanishes. It is worth noting that $\Delta(E)$ is real.[15] Having found such an energy, $P\Psi$ can be determined from Eq. (17). The other components of Ψ are then found from (16) after $P\Psi$ has been substituted on the right. Thus the zeros of the determinant $\Delta(E)$ are just the surface-state energies. The motivation for the various approximations implicit in use of the resolvent method should now be clear. In order to keep the dimension $2N$ of the system (18) sufficiently small to be dealt with, we try to minimize the size of the set S. This involves neglecting interband transitions involving bands much higher or lower in energy than the states sought; it involves neglecting transfer integrals coupling distant neighbors, neglecting surface perturbations of extended range, and truncating the distance to which the electrons can tunnel outside the crystal. Any of these constraints could be quantitatively relaxed at the price of increasing the dimension of S.

E. Adaptation of Resolvent Method to Surface-State Problem

So far our application of the resolvent method to surface states has not differed in any essential respect from its application to other sorts of bound defect states. At this point, however, it is helpful to recognize that ultimately the system described by $\mathcal{H}$ consists of two disjoint noncommunicating half-crystals. A surface state pertains to one particular half, and its amplitude vanishes identically on the other. Equation (18), on the other hand, involves both sides of the crystal. One might expect that it would be possible to reexpress the problem so as to depend explicitely on the properties of only a single half-crystal. This is indeed the case and has the effect of reducing the dimensionality of the determinant to be calculated by half. In fact, we can argue, as above: If $P\Psi$ satisfies Eq. (17), then Ψ constructed via Eq. (16) is an eigenstate of $\mathcal{H}$. Since $\mathcal{H}$ has no matrix elements connecting the two halves of the crystal, Ψ can always be assumed to vanish on one side or the other.[16] Suppose, for definiteness, that Ψ vanishes on the left; then $P_0\Psi = 0$. Equation (17) now implies that

$$P_1\Psi = P_1\mathcal{R}(E)\mathcal{V}(P_1\Psi) \quad , \tag{19}$$

and hence, equating the determinant of the transformation to zero,

$$\Delta_1(E) \equiv \det P_1[1 - \mathcal{R}(E)\mathcal{V}] = 0 \quad . \tag{20R}$$

This determinant $\Delta_1(E)$ is N dimensional, and Eq. (20R) represents the desired reduction of the problem. This truncation of the surface-state problem has been used by Koutecký and co-workers.[13,17,18] Evaluating $P_0\Psi$, which is supposed to be identically zero, from Eq. (17), we obtain the additional relation

$$0 = P_0\mathcal{R}(E)\mathcal{V}P_1\Psi \tag{21}$$

from which it follows that

$$\det[\mathcal{R}(E)\mathcal{V}]_{01} = 0 \quad . \tag{22}$$

(The subscript 01 indicates that the square matrix with rows from S_0 and columns from S_1 is denoted.) Such relations as this have sometimes been found useful.[18]

We have shown that Eq. (20R) or its left-side analog,

$$\Delta_0(E) \equiv \det P_0[1 - \mathcal{R}(E)\mathcal{V}] = 0 \quad , \tag{20L}$$

must hold for any surface-state energy; that is, one or the other of these equations holds whenever (18) is satisfied. However, we have yet to establish that (20L) or (20R) is actually *sufficient* in order that E correspond to the energy of a surface state. We must show that (18) actually follows from either of these conditions. This question of the sufficiency of the truncated determinantal condition has been generally overlooked. An affirmative answer has been implicitly assumed.[17,18] It turns out, however, to be quite difficult to prove that (20L) or (20R) implies (18) for the very good reason that it is *not quite* true. Instead one finds that the following identity holds:

$$\det P_0(1 - \mathcal{R}\mathcal{V}) \;\; \det P_1(1 - \mathcal{R}\mathcal{V}) = \det P(1 - \mathcal{R}\mathcal{V}) \;\; \det P(1 - \mathcal{R}\mathcal{V}_S) \quad . \tag{23}$$

Note the presence of the subscript S in the rightmost determinant. This equation is proved in Appendix A. It shows that among the zeros of $\Delta_0(E)$ and $\Delta_1(E)$, the factors on the left, are included not only the zeros of $\Delta(E)$, which we know correspond to the true surface states, but also the zeros of $\Delta_S(E) \equiv \det P[1 - \mathcal{R}(E)\mathcal{V}_S]$. These latter give the energies of the Shockley states, the surface states present when $\mathcal{V}_T$ is set to zero.

Using Eq. (23), it is possible to deduce the properties of $\Delta_0(E)$ and $\Delta_1(E)$ in detail. Specifically, $\Delta_1(E)$, the right-side truncated determinant, has zeros at the energies of the surface (Tamm) states of the right half-crystal plus zeros at the energies of the Shockley states of the *left* half-crystal. These latter are spurious roots, which do not imply the presence of surface states at all. These two categories of zeros exhaust the roots of $\Delta_1(E)$; $\Delta_0(E)$ has analogous properties. It is instructive to go through the argument leading to these conclusions. We shall need to make use of the additional identity,

$$\det P_0[1 - \mathcal{R}(E)\mathcal{V}_S] = \{\det P_1[1 - \mathcal{R}(E)\mathcal{V}_S]\}^* \quad , \tag{24}$$

which is also proved in Appendix A.[19] If we set $\mathcal{V}_T$ to zero in Eq. (23) and use Eq. (24), then we see that

$$\begin{aligned} |\det P_0[1 - \mathcal{R}(E)\mathcal{V}_S]| &= |\det P_1[1 - \mathcal{R}(E)\mathcal{V}_S]| \\ &= |\det P[1 - \mathcal{R}(E)\mathcal{V}_S]| \quad . \end{aligned} \tag{25}$$

This last determinant is real, like $\Delta(E)$ in general; its zeros give the Shockley state energies for both half-crystals. Clearly, each of the subdeterminants, $\Delta_0(E)$ and $\Delta_1(E)$, must now vanish at the energies of all Shockley states, both left and right. Now reintroduce $\mathcal{V}_T' = \lambda\mathcal{V}_T$ and consider what happens to $\Delta_1(E)$ when λ is gradually increased from 0 to 1. For $\lambda = 0$, the zeros of $\Delta_1(E)$ give the Shockley states for both half-crystals. As λ increases, the part of $\mathcal{V}_T'$ which perturbs the right half causes the energies of those erstwhile Shockley states located on the right to shift. Additional zeros of $\Delta_1(E)$ may appear or some may be lost. In any event, none of these displaced energies coincides with a Shockley state energy. In Eq. (23), they can only match up to zeros of $\Delta(E)$ and thus represent the true bound states. On the other hand, those zeros of $\Delta_1(E)$ associated initially with Shockley states of the left half-crystal are left to match up to zeros of $\Delta_S(E)$ in Eq. (23). These zeros of $\Delta_1(E)$ do not shift and remain at the energies of the left Shockley states when $\lambda = 1$. They can be easily identified as spurious by this very property. To reiterate, each of the truncated determinantal conditions, (20L) and (20R), yields as its solutions the true surface-state energies for the side to which it nominally refers plus spurious solutions at the energies of the Shockley states of the wrong side.

There is another, quite different method of reducing the $2N$-dimensional problem (18), to N dimensions. Suppose we had started from $\mathcal{H}_0 + \mathcal{V}_S$ as unperturbed Hamiltonian and then applied $\mathcal{V}_T$ as the surface perturbation. The surface-state energies for the right half-crystal would then be given by the solutions of the determinantal equation

$$\det P_1[1 - \mathcal{R}'(E)\mathcal{V}_T] = 0 \quad . \tag{26}$$

Here $\mathcal{R}'(E) = (E - \mathcal{H}_0 - \mathcal{V}_S)^{-1}$ is the resolvent for the Hamiltonian $\mathcal{H}_0 + \mathcal{V}_S$. Since the latter is block diagonal, so is $\mathcal{R}'(E)$, and the two half-crystals are decoupled right from the start. This approach is closely related to that employed by Baldock[20]; he dealt with a model crystal so simple that the solutions of $\mathcal{H}_0+\mathcal{V}_S$ could be determined by inspection. This permitted $\mathcal{R}'(E)$ to be constructed by essentially the same method we shall use for $\mathcal{R}(E)$ below. However, in general, this is not feasible. Fortunately, there is a general prescription which permits $\mathcal{R}'(E)$ to be routinely derived from $\mathcal{R}(E)$ itself.

We define $\mathcal{D}(E)$,

$$\mathcal{D}(E)=1-\mathcal{R}(E)\mathcal{V}_S\ . \tag{27}$$

If $\mathcal{D}(E)^{-1}$ could be found, then $\mathcal{R}'(E)$ could be obtained as $\mathcal{D}^{-1}\mathcal{R}$:

$$\begin{aligned}\mathcal{D}(E)^{-1}\mathcal{R}(E)&=\{\mathcal{R}(E)^{-1}[1-\mathcal{R}(E)\mathcal{V}_S]\}^{-1}\\&=(E-\mathcal{H}_0-\mathcal{V}_S)^{-1}\ .\end{aligned} \tag{28}$$

Here $\mathcal{R}(E)^{-1}=(E-\mathcal{H}_0)$, which is essentially the definition of $\mathcal{R}(E)$, has been used to obtain the second line. Now $\mathcal{D}(E)$ is just the unit matrix except for those *columns* referring to the surface basis states, i.e., to the set S. This property is summarized by the relation

$$\mathcal{D}(E)(1-P)=(1-P)\ . \tag{29}$$

The projection operator $1-P$ just picks out those components which do *not* involve S. Multiplying Eq. (29) through by $\mathcal{D}(E)^{-1}$ on the left shows that $\mathcal{D}(E)^{-1}$ also has this property

$$\mathcal{D}(E)^{-1}(1-P)=(1-P)\ . \tag{30}$$

Thus only the columns of $\mathcal{D}(E)^{-1}$ referring to S have to be determined. From (26) and (28), we see that it is only the rows of $\mathcal{R}'(E)$, and thus of $\mathcal{D}(E)^{-1}$, referring to S which are of immediate interest; that is, we want to find $P\mathcal{D}(E)^{-1}$. We multiply Eq. (30) by P on the left and use the property of projection operators $P(1-P)=0$ to obtain

$$P\mathcal{D}(E)^{-1}=P\mathcal{D}(E)^{-1}P\ . \tag{31}$$

In short, only the finite dimensional submatrix of $\mathcal{D}(E)^{-1}$ with both rows and columns referring to S is required, and the relevant portion of $\mathcal{R}'(E)$ to use in Eq. (26) is found by multiplying *finite*-dimensional matrices as follows:

$$\begin{aligned}P\mathcal{R}'(E)P&=P\mathcal{D}(E)^{-1}\mathcal{R}(E)P\\&=P\mathcal{D}(E)^{-1}P\mathcal{R}(E)P\ .\end{aligned} \tag{32}$$

It is perhaps worth noting explicitly that although $\mathcal{R}'(E)$ has no matrix elements connecting S_0 and S_l, both $\mathcal{D}(E)^{-1}$ and $\mathcal{R}(E)$ do. The multiplication in Eq. (32) involves the components of both these sets. Finally, we use Eq. (31) and the analogous equation for $\mathcal{D}(E)$ itself to determine $P\mathcal{D}(E)^{-1}P$. Multiply the definition, $\mathcal{D}\mathcal{D}^{-1}=1$, on left and right by P, then

$$P\mathcal{D}(E)\mathcal{D}(E)^{-1}P=P\mathcal{D}(E)P\mathcal{D}(E)^{-1}P=P\ . \tag{33}$$

This shows that, regarded as $2N$-dimensional finite matrices,

$$P\mathcal{D}(E)^{-1}P=[P\mathcal{D}(E)P]^{-1}\ . \tag{34}$$

The inverse of the finite-dimensional matrix $P\mathcal{D}(E)P$ can be routinely constructed.

In Appendix A, it is shown that we can improve on Eq. (34) and write out a more explicit expression for $P\mathcal{D}(E)^{-1}P$. Using subscripts to denote the submatrices involving different combinations of rows and columns from S_0 and S_1, we find that

$$\begin{Bmatrix}[\mathcal{D}^{-1}]_{00}\ ; & [\mathcal{D}^{-1}]_{01}\\ [\mathcal{D}^{-1}]_{10}\ ; & [\mathcal{D}^{-1}]_{11}\end{Bmatrix}=\begin{Bmatrix}1_{00} & ;\ -\mathcal{R}_{01}(\mathcal{R}_{11})^{-1}\\ -\mathcal{R}_{10}(\mathcal{R}_{00})^{-1} & ;\ 1_{11}\end{Bmatrix}\ . \tag{35}$$

The E dependence is to be understood. Note that $(\mathcal{R}_{00})^{-1}$, say, stands for the matrix inverse of the finite-dimensional matrix $\mathcal{R}_{00}$. From Eq. (32), we now obtain

$$\mathcal{R}'_{11}=\mathcal{R}_{11}-\mathcal{R}_{10}(\mathcal{R}_{00})^{-1}\mathcal{R}_{01}\ . \tag{36}$$

Though this rather resembles second-order perturbation theory, it is not an approximation.

From the identities derived above, together with the definition of $\mathcal{D}(E)$, and the usual rules for manipulating the determinants of finite-dimensional matrices, we derive that

$$\begin{aligned}\frac{\det P[1-\mathcal{R}(E)\mathcal{V}]}{\det P[1-\mathcal{R}(E)\mathcal{V}_S]}&=\det P\{\mathcal{D}(E)^{-1}[1-\mathcal{R}(E)\mathcal{V}]\}\\&=\det P[1-\mathcal{R}'(E)\mathcal{V}_T]\\&=\det P_0[1-\mathcal{R}'\mathcal{V}_T]\det P_1[1-\mathcal{R}'\mathcal{V}_T]\ .\end{aligned} \tag{37}$$

The first line required the use of (34) and (31) and the second (28) or (32). The third then follows from the fact that both $\mathcal{R}'(E)$ and $\mathcal{V}_T$ are block diagonal, having no matrix elements connection S_0 and S_1. Suppose we are interested in the right half-crystal and choose $\mathcal{V}_T$ to be nonvanishing only on the right. Then it is clear from Eq. (37) that Eq. (26) is identical to the condition

$$[\Delta(E)/\Delta_S(E)]=0\ . \tag{38}$$

Here $\Delta_S(E)$ stands for $\det P[1-\mathcal{R}(E)\mathcal{V}_S]$ as usual. As far as the Tamm states of the right half-crystal are concerned, Eq. (26) in the form (38) is equivalent to our original fundamental condition, Eq. (18).

When the definition of $\mathcal{D}(E)$ given in Eq. (27) and the form of $P\mathcal{D}(E)^{-1}P$ given in Eq. (35) are substituted into the relation (33), various identities involving the matrix elements of $\mathcal{R}(E)$ are generated. These identities are sometimes useful in simplifying the various determinantal conditions: (18), (20), and (26). In general, it requires less effort to determine the surface states from (20) than from (26).

The latter equation does have the advantage, however, that the necessary effects of breaking the crystal are explicitly segregated from the optional effects of the surface perturbation $\mathcal{V}_T$. In addition, it avoids spurious solutions at the Shockley states of the wrong half-crystal, though at the price of poles at the Shockley states of the right half.

F. Matrix Element of Resolvent

It remains only to describe how to obtain explicitly the required matrix elements of $\mathcal{R}(E)$. These are expressed below in the integral form, Eq. (47), which is used together with Eq. (44) or Eq. (45).

As the inverse of $(E-\mathcal{H}_0)$, $\mathcal{R}(E)$ has essentially all the symmetry properties of $\mathcal{H}_0$ itself: It is Hermitian, invariant under "time reversal" and under the space group of the crystal. In particular, we shall exploit its invariance under the translation group of the crystal, which implies that this operator is essentially diagonal in a plane-wave representation. Specifically, let us define a basis of plane-wave states $|\vec{k}\beta\rangle$:

$$|\vec{k}\beta\rangle=\sum_{\vec{m}} e^{i\vec{k}\cdot\vec{m}}\,|\vec{m}\beta\rangle\ . \tag{39}$$

We shall require the relation between these states and our basis of layer orbitals $|\vec{k}_\parallel;\, m\beta\rangle$:

$$|\vec{k}\beta\rangle=\sum_m e^{ikm}\,|\vec{k}_\parallel;\, m\beta\rangle \tag{40}$$

and

$$|\vec{k}_\parallel;\, m\beta\rangle=\int_{-\pi}^{\pi}\frac{dk}{2\pi}\,e^{-ikm}\,|\vec{k}\beta\rangle\ . \tag{41}$$

The wave vector $\vec{k}$ is supposed to have the components $\vec{k}=k_1\vec{a}^*+k_2\vec{b}^*+k\vec{c}^*$. Because of the translation symmetry of $\mathcal{H}_0$, the matrix elements of $(E-\mathcal{H}_0)$ in the $|\vec{k}\beta\rangle$ basis have the form

$$\langle\vec{k}'\beta|(E-\mathcal{H}_0)|\vec{k}\,\alpha\rangle=(2\pi)^3\,\delta^3(\vec{k}'-\vec{k})\,[E\overleftrightarrow{1}-\overleftrightarrow{X}(\vec{k})]_{\beta\alpha}, \tag{42}$$

where $\overleftrightarrow{X}(\vec{k})$ is the ν-dimensional matrix defined in Eq. (6), and $\overleftrightarrow{1}$ is the ν-dimensional unit matrix: $(\overleftrightarrow{1})_{\beta\alpha}=\delta_{\beta\alpha}$. Equation (42) was implicitly used in transforming the bulk eigenvalue equation (2) to the form (5). Now the inverse of $(E-\mathcal{H}_0)$ has matrix elements of analogous form

$$\langle\vec{k}'\beta|(E-\mathcal{H}_0)^{-1}|\vec{k}\alpha\rangle=(2\pi)^3\delta^3(\vec{k}'-\vec{k})$$
$$\times\{[E\,\overleftrightarrow{1}-\overleftrightarrow{X}(\vec{k})]^{-1}\}_{\beta\alpha}\ . \tag{43}$$

The last term on the right is just the $\beta\alpha$ matrix element of the ν-dimensional matrix inverse of $[E\,\overleftrightarrow{1}-\overleftrightarrow{X}(\vec{k})]$. There are standard methods by which this inverse matrix can be found. For instance, we can write

$$\{[E\overleftrightarrow{1}-\overleftrightarrow{X}(\vec{k})]^{-1}\}_{\beta\alpha}=\frac{\Delta_{\alpha\beta}(\vec{k};E)}{\Delta(\vec{k};\,E)}\quad . \tag{44}$$

Here $\Delta(\vec{k};E)$ is the determinant of the matrix $[E\,\overleftrightarrow{1}-\overleftrightarrow{X}(\vec{k})]$ and $\Delta_{\alpha\beta}(\vec{k};E)$ is the minor of its $\alpha\beta$ entry. Another representation of the inverse has the form of a sum over the eigenvectors of $\overleftrightarrow{X}(\vec{k})$:

$$\{[E\,\overleftrightarrow{1}-\overleftrightarrow{X}(\vec{k})]^{-1}\}_{\beta\alpha}=\sum_{\rho=1}^{\nu}\frac{e_\beta(\vec{k},\rho)\,e^*_\alpha(\vec{k},\rho)}{E-\lambda_\rho(\vec{k})}\quad . \tag{45}$$

The eigenvectors $\vec{e}(\vec{k},\rho)$ were introduced in Eqs. (4) and (5); they are orthonormal and complete in the ν-dimensional space of states having fixed $\vec{k}$. These forms are, of course, well known, and both have been used elsewhere to calculate this inverse matrix.[7,18] The form (44) is ordinarily the more convenient for calculation.

We need to find the matrix elements in the layer orbital basis of that factored form of $\mathcal{R}(E)$ which is appropriate to the effective one-dimensional problem. That is, we want $\mathcal{R}(E;\vec{k}_\parallel)$, which bears the same relation to $\mathcal{R}(E)$ that $\mathcal{H}_0(\vec{k}_\parallel)$ does to $\mathcal{H}_0$:

$$\langle\vec{k}'_\parallel;m\beta|\mathcal{R}(E)|\vec{k}_\parallel;\,l\alpha\rangle=(2\pi)^2\delta(\vec{k}'_\parallel-\vec{k}_\parallel)$$
$$\times\langle\vec{k}_\parallel;m\beta|\mathcal{R}(E;\vec{k}_\parallel)|\vec{k}_\parallel;l\alpha\rangle\ . \tag{46}$$

Using the relation and the expression for $\mathcal{R}(E)$, Eq. (43), we obtain

$$\langle\vec{k}_\parallel;m\beta|\mathcal{R}(E;\,\vec{k}_\parallel)|\vec{k}_\parallel;l\alpha\rangle$$
$$=\int_{-\pi}^{\pi}\frac{dk}{2\pi}\,e^{ik(m-l)}\{[E\,\overleftrightarrow{1}-\overleftrightarrow{X}(\vec{k})]^{-1}\}_{\beta\alpha}\ . \tag{47}$$

Either of the two forms, Eqs. (44) and (45), can be used in Eq. (47) to complete the evaluation of the integrand.[21]

Thanks to the translation invariance of $\mathcal{R}(E)$, the matrix element evaluated in Eq. (47) depends only on the difference, $m-l$; this feature is explicitly displayed by the integral expression on the right. We shall denote this matrix element by $\mathcal{R}_{m-l;\beta\alpha}(E;\,\vec{k}_\parallel)$. It is useful to consider the implications for this quantity of the symmetry properties of $\mathcal{R}(E)$ enumerated previously. Hermiticity now takes the form

$$R^*_{n;\,\alpha\beta}(E;\vec{k}_\parallel)=R_{-n;\beta\alpha}(E;\vec{k}_\parallel)\quad , \tag{48}$$

and time reversal invariance

$$R^*_{n;\,\alpha\beta}(E;\vec{k}_\parallel)=R_{n;\,\alpha\beta}(E;\,-\vec{k}_\parallel)\,. \tag{49}$$

Those operations of the factor group of the crystal's space group which carry the surface plane into itself will give further relations; examples of such will be given in Sec. III.

It is usually possible to perform the integral in Eq. (47) by regarding k, the variable of integration, as a complex variable and deforming the contour of integration. From Eq. (6) we see that the matrix elements $X_{\beta\alpha}(\vec{k})$ are all analytic functions of k

throughout the finite k plane. Hence, $\Delta(\vec{k}, E)$ and its minors also have this property. Now it suffices to consider positive values of n, since R_{-n} can be obtained from R_n with the aid of Eq. (48). For large enough positive values of n, the factor e^{ikn} causes the integrand of $R_{n;\beta\alpha}(E;\vec{k}_{\parallel})$ to become exponentially small as the imaginary part of k approaches $+\infty$. We deform the contour to run from $(-\pi, 0)$ to $(-\pi, \kappa)$ along the line $\mathrm{Re}k=-\pi$ then to (π, κ) along the line $\mathrm{Im}k=\kappa>0$, and finally to $(\pi, 0)$ along $\mathrm{Re}k=\pi$. No contribution to the integral is obtained from the sum of the line segments along $\mathrm{Re}k=-\pi$ and $\mathrm{Re}k=+\pi$ because these segments are transversed in opposite directions, and the integrand is periodic in $\mathrm{Re}k$ with period 2π. Letting κ approach $+\infty$, we see that the value of the integral is just $2\pi i$ times the sum of the residues of the integrand at its poles in the upper half k plane. These poles are just the zeros of $\Delta(\vec{k}, E)$ regarded now as a function of the complex variable, k. For small values of $n \geq 0$, there may be an additional contribution to R_n from the integral along the line segment, $\mathrm{Im}k=\kappa$, in the limit $\kappa \to +\infty$. This contribution, as well as those stemming from the zeros of the denominator, can be more easily studied if the variable of integration is changed to $\mu \equiv e^{ik}$.[22] From Eq. (6), we see that $X_{\beta\alpha}(\vec{k})$ is a polynomial of finite order in positive and negative powers of this new variable. This is therefore true of $\Delta(\vec{k}, E)$ and its minors as well. Equation (47) now becomes

$$R_{n;\beta\alpha}(E;\vec{k}_{\parallel})=\frac{1}{2\pi i}\oint \frac{d\mu}{\mu}\, \mu^{n}\, \frac{d_{\alpha\beta}(\mu)}{d(\mu)} \quad . \tag{50}$$

The function $d(\mu)$ is just $\Delta(\vec{k}, E)$ expressed in terms of μ and with the parameters E and $\vec{k}_{\parallel}$ no longer displayed explicitly. Similarly, $d_{\alpha\beta}(\mu)$ is $\Delta_{\alpha\beta}(\vec{k}, E)$. The integral in Eq. (50) is over the unit circle in the counter-clockwise sense. The upper half k plane has become the interior of the unit circle, and for $n \geq 0$, we shall shrink the contour to zero, picking up the various residues at the poles of the integrand within the unit circle. Note that the point $\mu=0$ now corresponds to the limit $\mathrm{Im}k \to +\infty$, so that what was previously the limiting contribution on the contour, $\mathrm{Im}k=\kappa$, has now become simply the contribution from a pole of finite order at $\mu=0$.

Now for real k, $\Delta(\vec{k}, E)$ is real since $\overleftrightarrow{X}(\vec{k})$ is Hermitian. This means that $d(\mu)$ is real for μ on the unit circle, or, more generally,

$$d(\mu)^{*}=d(1/\mu^{*}) \quad . \tag{51}$$

Let A be an integer representing the degree of the highest power of μ in $d(\mu)$. From the reality of $d(e^{ik})$ for real k, this must also be the degree of the highest power of μ^{-1}, so that $d(\mu)$ can be written

$$d(\mu)=\mu^{-A}p(\mu) \quad , \tag{52}$$

where $p(\mu)$ is a polynomial in μ of degree $2A$. The roots of $p(\mu)$ locate the poles of the integrand in Eq. (50). From its definition as the determinant of $[E\overleftrightarrow{1}-\overleftrightarrow{X}(\vec{k})]$, it is obvious that the zeros of $\Delta(\vec{k}, E)$ for real k give the bulk band energies $E=\lambda_{\rho}(\vec{k})$, $\rho=1, \ldots, \nu$. Since we have stipulated that E lies outside the range of all the bands $\lambda_{\rho}(\vec{k})$ for given $\vec{k}_{\parallel}$, $d(\mu)$ cannot vanish for μ on the unit circle. Moreover, given any root of $d(\mu)$ at μ_1, say, Eq. (51) implies that there must be another at $\mu_1'=1/\mu_1^{*}$. If one of these is inside the unit circle, the other must be outside. Thus $p(\mu)$ must have exactly A roots within the unit circle: μ_j, $j=1, \ldots, A$. The jth root contributes the residue

$$\mu_j^{n-1}\,\frac{d_{\alpha\beta}(\mu_j)}{d'(\mu_j)} \tag{53}$$

to the integral (50), where $d'(\mu)$ stands for the derivative of $d(\mu)$ with respect to μ. Note the nature of the dependence on n: writing $\mu_j=e^{iq_j}$ with $\mathrm{Im}q_j>0$, the contribution (53), to $R_{n;\beta\alpha}$ is proportional to e^{inq_j}; that is, it goes exponentially to zero with increasing n, the decay constant being just the imaginary part of q_j. We now return to considering the $\mu=0$ contribution. Let B be the degree of the highest power of μ^{-1} appearing in any of the minors, $d_{\alpha\beta}(\mu)$. Then as μ tends to zero, the most singular term in any of the integrands for given n behaves as $\mu^{n-1+A-B}$. If $A>B$, then this is a positive power of μ for all $n \geq 0$, and there is no contribution to any $R_{n;\beta\alpha}$ from the origin. If $B \geq A$, the most singular term is a pole of order $B-A+1-n$. The integrand will ordinarily include a first-order pole at $\mu=0$ if this exponent is greater than or equal to one. Thus for all n, $0 \leq n \leq n_0 \equiv (B-A)$, there will be an additional term in $R_{n;\beta\alpha}$ for at least some subscripts $\beta\alpha$, representing the residue at this pole at the origin. Commonly, we find $n_0 \leq 0$, so that only $\overleftrightarrow{R}_0$ has an extra contribution.[23] Below we shall set $n_0=-1$ in case $A>B$.

Having determined the form of $\overleftrightarrow{R}_n$, it is now possible to describe the nature of a surface-state wave function. From the fundamental equation (16), we see that $\mathcal{V}\Psi$, which is nonzero only in the surface region, acts as a "source" for the wave function throughout the crystal. Let l_s define the region in which $\mathcal{V}\Psi$ is nonzero: $(\mathcal{V}\Psi)_l=0$ except for $-l_s+1 \leq l \leq l_s$. (The number of surface orbitals, N, considered above is just νl_s, recall Ref. 14.) Then for $m>l_s+n_0$, a surface state Ψ has the form

$$\Psi(m\beta)=\sum_{j=1}^{A} e^{imq_j}\phi_{\beta}^{(j)} \quad , \tag{54}$$

where the roots of $p(\mu)$ are written $\mu_j=e^{iq_j}$, $j=1, \ldots, A$, as discussed above, and $\phi_{\beta}^{(j)}$ is the βth component of a ν-dimensional polarization vector $\vec{\phi}^{(j)}$. Of course, $\vec{\phi}^{(j)}$ depends on the components of Ψ in the surface region:

$$\phi_\beta^{(j)} = \sum_{l=-l_s+1}^{l_s} \sum_{\alpha=1}^{\nu} \mu_j^{-1-l} \frac{d_{\alpha\beta}(\mu_j)}{d'(\mu_j)} (\mathcal{V}\Psi)_{l\alpha} \; . \tag{55}$$

The point worth noting is that Eq. (54) expresses Ψ interior to the surface region as a sum of A linearly independent solutions of the bulk equations of motion. That is, each term in the sum can be obtained by analytically continuing an ordinary band state to the energy of the surface state. Of course, the number of such linearly independent states A is not simply related to the number of bands, ν. Having obtained the asymptotic form, Eq. (54), which is a sum of A decaying exponentials, and knowing its region of validity $m > l_s + n_0$ we now have the assurance that it we make the *ansatz* that the wave function has this form and treat E, μ_j, $\vec{\phi}^{(j)}$, and $\vec{\Psi}_m$ for $1 \le m \le l_s + n_0$ as parameters to be determined, in the usual fashion, then a solution of this form can indeed be found and the surface-state energy determined thereby. (If $n_0 = -1$ and $l_s = 1$, there is no surface region in which $\vec{\Psi}_m$ takes an exceptional form and the asymptotic form is valid everywhere.) No simpler ansatz will suffice. To carry out this program, which is based on knowing in advance the correct form of the surface-state wave function, only the exponents A and B are required. These can be easily determined from the matrix $\overleftrightarrow{X}(\vec{k})$ describing the bulk band structure. For nontrivial examples, this ansatz method is more difficult to carry through than straightforward application of the resolvent technique. Nonetheless, for high symmetry points in the two-dimensional zone of the surface states, the ansatz approach can be useful, especially in deriving threshold conditions. In obtaining such conditions otherwise, the energy is fixed at a band edge value and the determinantal bound-state condition, (18), (20), or (26), is regarded as a function of the parameters characterizing the surface perturbation.

As a final note, it is worth mentioning that including overlaps in treating the surface states causes little or no increase in computational difficulty. If the local atomic basis states are not taken to be orthogonal, then the transfer integrals appearing parametrically in $\overleftrightarrow{X}(\vec{k})$ are effectively replaced by linear functions of $E: \gamma \to \gamma - E\sigma$, say, where σ represents the overlap integral between the pair of orbitals whose transfer integral is γ. Since in either the ansatz or the resolvent method, we use the bulk dispersion condition, $\det[E\overleftrightarrow{1} - \overleftrightarrow{X}(\vec{k})] = 0$, to solve for $\mu \equiv e^{ik}$ as a function of E rather than vice versa, it suffices to consider the solutions $\mu_j(E)$ already found. These depend parametrically on the transfer integrals γ, for which we now simply substitute the expressions $\gamma - E\sigma$. In the ansatz scheme, this suffices, and we proceed as before; otherwise, it remains only to make a similar substitution wherever γ appears in the expressions for $\mathcal{R}$ and $\mathcal{V}_S$ or for $\mathcal{R}'$. Ordinarily we will wish to modify the surface perturbation $\mathcal{V}_T$ as well, in so far as it represents changes in the bonds in the surface region. As an example, if γ is replaced by $\gamma - E\sigma$ in Eqs. (62) and (86) below, then $\mathcal{R}'$ given in Eq. (86) correctly represents the resolvent for the presevered crystal including overlaps. Corresponding modifications in the off-diagonal terms proportional to δ_1, δ_2, and δ_3 would be appropriate in Eq. (72).

III. APPLICATION

In this section, the formalism derived above will be applied to a physically interesting system. The surface electronic states associated with a (110) face of a crystal having the zinc-blende structure will be considered. The Bravais lattice for this crystal is face-centered cubic; the basis may be taken to consist of a cation at the origin and an anion displaced by one quarter of a cube diagonal along one of the [111] directions parallel to the surface (110) plane. The appearance of a (110) plane is shown in Fig. 1. We shall make use of the right-handed Cartesian coordinate system depicted there having x axis along a $[1\bar{1}0]$ direction in the surface plane, y axis along the [001] direction is this plane, and z axis pointing along [110] *into* the crystal. Unit vectors in these directions are written $\hat{\imath}$, $\hat{\jmath}$, and $\hat{k}$, respectively. Basis vectors for the Bravais lattice

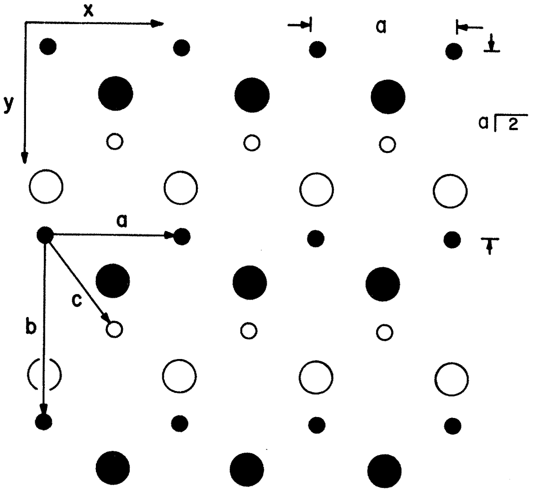

FIG. 1. The shaded circles represent atoms in the surface plane, the cation being depicted as the smaller of the two. The open circles are the projected images of the next layer of ions, displaced into the paper, which is the $+z$ direction, by $a/2$. The edge of the cubic unit cell is $a\sqrt{2}$.

may be chosen: $\vec{a}=a\hat{i}$, $\vec{b}=a\sqrt{2}\,\hat{j}$, and $\vec{c}=\frac{1}{2}a(\hat{i}+\sqrt{2}\,\hat{j}+\hat{k})$. Here a is the nearest *like* neighbor distance so that the length of the cube edge is $a\sqrt{2}$. With the origin located at a cation site in the surface layer of the left half-crystal, the half-crystal of interest will consist of all sites having $m \geqslant 1$. Of course, both halves have identical properties since there is a glide plane parallel to (110) which interchanges them. As discussed in Ref. 11, it will be convenient to write $\vec{k}=(1/a)(k_1\hat{i}+k_2\hat{j}/\sqrt{2}+2k\hat{k})$ rather than using the actual reciprocal-lattice vectors: $\vec{a}^*=(\hat{i}-\hat{k})/a$ and $\vec{b}^*=(\hat{j}/\sqrt{2}-\hat{k})/a$. This amounts to making a canonical transformation, changing the relative phases of the layer orbitals $|\vec{k}_{\parallel};m\beta\rangle$. The wave-vector components, k_1 and k_2, are nontheless conserved. The basis states included in each unit cell represent an s-like state on the cation, $\beta=0$, and three p-like states on the anion, $\beta=1$, 2, and 3, labeled in the order p_x, p_y, and p_z. These subscripts refer to our Cartesian coordinate system. Writing rows and columns in order of increasing β from $\beta=0$ to $\beta=3$, $\overleftrightarrow{X}(\vec{k})$, the submatrix of the Hamiltonian for states of wave vector $\vec{k}$, takes the form

$$\overleftrightarrow{X}(\vec{k})=\begin{bmatrix} H & C(-\vec{k}) & D(-\vec{k}) & F(-\vec{k}) \\ C(\vec{k}) & -H & 0 & 0 \\ D(\vec{k}) & 0 & -H & 0 \\ F(\vec{k}) & 0 & 0 & -H \end{bmatrix} , \tag{56}$$

where

$$\begin{aligned} C(\vec{k}) &= i\sqrt{8}\,\gamma\,\sin(k_1/2)e^{-ik_2/4} , \\ D(\vec{k}) &= 2\gamma[-\cos(k_1/2)e^{-ik_2/4}+\cos(k)e^{ik_2/4}] , \\ F(\vec{k}) &= i\sqrt{8}\,\gamma\,\sin(k)e^{ik_2/4} . \end{aligned} \tag{57}$$

The transformation from position to wave-vector representation has been based on the convention discussed in Ref. 11, taking $\vec{u}_0=0$ and $\vec{u}_\beta=\frac{1}{2}\vec{a}+\frac{1}{4}\vec{b}$ for $\beta=1$, 2, and 3. The zero of energy has been taken at the midpoint of the zone-center bandgap. The magnitude of this gap is $2H$. The parameter γ describes the magnitude of the nearest-neighbor s-p transfer integral; transfer integrals to more distant neighbors are neglected. This form of the Hamiltonian is discussed more fully in the companion paper.[1] Note that $C(-\vec{k})=C(\vec{k})^*$, etc., for real k.

The bulk bands resulting from Eq. (56) include two flat bands with energy $E=-H$. The conduction band and the remaining, light hole valence band are symmetrical about zero energy: $E=\pm\lambda_c(\vec{k})$, where

$$\lambda_c(\vec{k})=\{H^2+4\gamma^2[4-\cos^2(k_1/2)\,\sin^2(k_2/2)] - 4\gamma^2[\cos k+\cos(k_1/2)\,\cos(k_2/2)]^2\}^{1/2} . \tag{58}$$

At zone center, the three valence bands are tangent at $E=-H$, and the conduction band has its minimum $E=+H$. This bulk band structure is illustrated in Ref. 1. The determinant of $[E\overleftrightarrow{1}-\overleftrightarrow{X}(\vec{k})]$ has the value $(E+H)^2[E^2-\lambda_c(\vec{k})^2]$. From Eq. (58), we see that this quantity is quadratic in $\cos k$, so $A=2$ in the notation of the previous section. From Eq. (56), it is clear that these quadratic terms arise from contributions to the determinant proportional to $D(\vec{k})D(-\vec{k})$ and $F(\vec{k})F(-\vec{k})$. These products also appear in some of the minors; hence $B=2$ as well. Thus we can say at once that $\overleftrightarrow{R}_n$ is in general the sum of two contributions, each of which decays exponentially with increasing $|n|$. In addition, there will be an extra contribution for $n=0$. The integral indicated in Eq. (47) was performed using in turn each of the methods of evaluating $[E\overleftrightarrow{1}-\overleftrightarrow{X}(\vec{k})]^{-1}$, Eqs. (44) and (45). When the sum over eigenvectors, Eq. (45), was employed, the integration was performed *before* summing except that the two flat bands were lumped together. The expressions for $\overleftrightarrow{R}_n$ obtained by these two methods agreed, of course. The result for $n\geqslant 0$ may be expressed succinctly

$$\overleftrightarrow{R}_n(E;\vec{k}_{\parallel})=\overleftrightarrow{U}\cdot\left\{\delta_{n0}\overleftrightarrow{\mathcal{E}}^{(0)}-g\sum_{i=1}^{2}\frac{(\sigma_i)^n e^{-n\kappa_i}}{\sinh(\kappa_i)}[\vec{V}_i\otimes\vec{V}_i^\dagger]\right\} . \tag{59}$$

Here the left-hand side symbolizes the 4×4 matrix with matrix elements $R_{n;\beta\alpha}$. The matrix $\overleftrightarrow{U}$ is just

$$\overleftrightarrow{U}=\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} ; \tag{60}$$

multiplying by $\overleftrightarrow{U}$ on the left, as in Eq. (59), simply reverses the sign of the last row ($\beta=3$). The first term inside the curly brackets on the right-hand side in Eq. (59) represents the extra contribution in the $n=0$ case. The matrix $\overleftrightarrow{\mathcal{E}}^{(0)}$ is defined

$$\overleftrightarrow{\mathcal{E}}^{(0)}=(E+H)^{-1}\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & \sqrt{2} \\ 0 & 0 & \sqrt{2} & 1 \end{pmatrix} . \tag{61}$$

The remaining two terms arise from the zeros of $\det[E\overleftrightarrow{1}-\overleftrightarrow{X}(\vec{k})]$ in the upper half k plane. Specifically, the two zeros are written $q_1=i\kappa_1$ and $q_2=\pi+i\kappa_2$ and are determined from the relation

$$\cos q_i+\cos(k_1/2)\cos(k_2/2)=(\delta_i\{H^2+4\gamma^2 \times[4-\cos^2(k_1/2)\sin^2(k_2/2)]\}-E^2)^{1/2} . \tag{62}$$

Here and in Eq. (59), $\sigma_i=+1$ for $i=1$, and -1 for $i=2$. From Eq. (62), it is clear that

$$\cosh\kappa_2 - \cosh\kappa_1 = 2\cos(k_1/2)\cos(k_2/2) \quad . \tag{63}$$

The factor g before the sum in Eq. (59) stands for

$$1/\{4\gamma^2[\cosh\kappa_2 + \cosh\kappa_1](E+H)\} \quad ; \tag{64}$$

the sum $\cosh\kappa_1 + \cosh\kappa_2$ is just twice the magnitude of the right-hand side of Eq. (62). Finally, the matrix in square brackets represents the outer product of the four-dimensional vector $\vec{\mathrm{V}}_i$ with itself. These matrix elements are, specifically,

$$[\vec{\mathrm{V}}_i \otimes \vec{\mathrm{V}}_i^\dagger]_{\beta\alpha} = V_{i,\beta} V_{i,\alpha}^* \quad . \tag{65}$$

The vectors $\vec{\mathrm{V}}_i$, $i=1$ and 2, are defined

$$\vec{\mathrm{V}}_i = \left\{ \begin{array}{l} E+H \\ i\sqrt{8}\,\gamma\,\sin(k_1/2)e^{-ik_2/4} \\ 2\gamma[-\cos(k_1/2)e^{-ik_2/4} + \sigma_i \cosh(\kappa_i)e^{i\kappa_2/4}] \\ \sqrt{8}\,\gamma\sigma_i \sinh(\kappa_i)e^{ik_2/4} \end{array} \right\} \quad ; \tag{66}$$

the $\beta = 1$, 2, and 3 elements here are just $C(\vec{\mathrm{k}})$, $D(\vec{\mathrm{k}})$, and $-F(\vec{\mathrm{k}})$ evaluated for $k = q_i$. Aside from normalization, both vectors, $\overrightarrow{\mathrm{UV}}_1$ and $\overrightarrow{\mathrm{UV}}_2$, can be obtained by analytic continuation of the eigenvector associated with the conduction band. The former arises in continuing to complex k from the minimum of the one-dimensional band at $k=0$ (k_1 and k_2 fixed), the latter in continuing from the relative minimum present for nonzero k_1 or k_2 at $k=\pi$.

The bulk crystal is symmetric under reflection in a (110) plane passing through a layer of ion sites. [As usual, the Miller indices here refer to a cube edge basis. In our surface adapted coordinate basis, this would be a (001) plane.] Under this operation $\vec{\mathrm{k}}_{\parallel}$ is unchanged and n goes into $-n$. The orbitals labeled by $\beta = 0$, 1, and 2 are unchanged, while that labeled by $\beta = 3$ is multiplied by -1 under this operation. Thus a four-dimensional vector whose components refer to these basis elements, $\beta = 0$, 1, 2, and 3, is multiplied by $\overleftrightarrow{\mathrm{U}}$, defined in Eq. (60). Under this reflection, the 4×4 matrix $\overleftrightarrow{\mathrm{R}}_n(E;\vec{\mathrm{k}}_{\parallel})$ is transformed into $\overleftrightarrow{\mathrm{U}} \cdot \overleftrightarrow{\mathrm{R}}_{-n}(E;\vec{\mathrm{k}}_{\parallel}) \cdot \overleftrightarrow{\mathrm{U}}$. Since the crystal is invariant under this symmetry operation,

$$\overleftrightarrow{\mathrm{R}}_n(E;\vec{\mathrm{k}}_{\parallel}) = \overleftrightarrow{\mathrm{U}} \cdot \overleftrightarrow{\mathrm{R}}_{-n}(E;\vec{\mathrm{k}}_{\parallel}) \cdot \overleftrightarrow{\mathrm{U}} \quad . \tag{67}$$

This provides the simplest means of obtaining $\overleftrightarrow{\mathrm{R}}_n$ for negative n. For $\overleftrightarrow{\mathrm{R}}_0$, this relation implies that the off-diagonal elements in the last row and column vanish identically. This can be verified using Eq. (59). We also find that the general relations, Eqs. (48) and (49), hold. All these identities provide further tests of the correctness of the expression (59) for $\overleftrightarrow{\mathrm{R}}_n$.

The decoupling potential $\mathcal{V}_S$ has the following form in our present example:

$$\begin{aligned} \langle \vec{\mathrm{k}}_{\parallel}; m\beta | \mathcal{V}_S(\vec{\mathrm{k}}_{\parallel}) | \vec{\mathrm{k}}_{\parallel}; l\alpha \rangle &= V_{10;\beta\alpha} \quad \text{for} \quad m=1 \text{ and } l=0 \\ &= V_{01;\beta\alpha} \quad \text{for} \quad m=0 \text{ and } l=1 \\ &= 0 \ , \quad \text{otherwise} \end{aligned} \tag{68}$$

where

$$\overleftrightarrow{\mathrm{V}}_{10} = \overleftrightarrow{\mathrm{V}}_{01}^\dagger = -\left\{ \begin{array}{cccc} 0 & 0 & \gamma e^{-ik_2/4} & \sqrt{2}\,\gamma e^{-ik_2/4} \\ 0 & 0 & 0 & 0 \\ \gamma e^{ik_2/4} & 0 & 0 & 0 \\ -\sqrt{2}\,\gamma e^{ik_2/4} & 0 & 0 & 0 \end{array} \right\} . \tag{69}$$

The reflection symmetry employed above implies that

$$\overleftrightarrow{\mathrm{V}}_{01} = \overleftrightarrow{\mathrm{U}} \cdot \overleftrightarrow{\mathrm{V}}_{10} \cdot \overleftrightarrow{\mathrm{U}} \quad . \tag{70}$$

We assume that the surface layer, in addition to *not* being reconstructed, is not distorted in such a way as to sacrifice the reflection plane which passes through each ion site normal to the vector $\vec{\mathrm{a}}$. In that case, the most general form of surface perturbation – given that it affects only the surface layer itself – has the form

$$\begin{aligned} \langle \vec{\mathrm{k}}_{\parallel}; m\beta | \mathcal{V}_T(\vec{\mathrm{k}}_{\parallel}) | \vec{\mathrm{k}}_{\parallel}; l\alpha \rangle &= V_{11;\beta\alpha} \quad \text{for} \quad m=1 \text{ and } l=1 \\ &= V_{00;\beta\alpha} \quad \text{for} \quad m=0 \text{ and } l=0 \\ &= 0 \ , \quad \text{otherwise} \end{aligned} \tag{71}$$

where

$$\overleftrightarrow{\mathrm{V}}_{11} = \left\{ \begin{array}{cccc} \epsilon_0 & 2i\delta_1 \sin(k_1/2)e^{ik_2/4} & 2\delta_2\cos(k_1/2)e^{ik_2/4} & 2\delta_3\cos(k_1/2)e^{ik_2/4} \\ -2i\delta_1\sin(k_1/2)e^{-ik_2/4} & \epsilon_1 & 0 & 0 \\ 2\delta_2\cos(k_1/2)e^{-ik_2/4} & 0 & \epsilon_2 & \epsilon_{23} \\ 2\delta_3\cos(k_1/2)e^{-ik_2/4} & 0 & \epsilon_{23} & \epsilon_3 \end{array} \right\} . \tag{72}$$

This expression will be employed in Ref. (1). The nature of $\overleftrightarrow{\mathrm{V}}_{00}$ has no effect on the surface states of the right half-crystal, which we are studying here. It is convenient, however, to define it according to

$$\overleftrightarrow{V}_{00} = \overleftrightarrow{U} \cdot \overleftrightarrow{V}_{11} \cdot \overleftrightarrow{U} \quad . \tag{73}$$

With this definition, the left and right half-crystals are interchanged by a glide plane midway between the $m = 0$ and $m = 1$ planes of ions. The complete system, composed of the two half-crystals, is invariant under this operation.

The surface-state energies are determined by the fundamental equation (18). The surface set S onto which P projects here comprises all four basis states, $\beta = 0$, 1, 2, and 3, of each of the two surface layers, $m = 0$ and $m = 1$. The rows and columns of the determinant $\Delta(E)$ will be written with $m = 1$ following $m = 0$ so that Eq. (18) now takes the form

$$\Delta(E) = \det \left\{ \begin{matrix} \overleftrightarrow{1} - \overleftrightarrow{R}_{-1}\overleftrightarrow{V}_{10} - \overleftrightarrow{R}_0\overleftrightarrow{V}_{00} & -\overleftrightarrow{R}_0\overleftrightarrow{V}_{01} - \overleftrightarrow{R}_{-1}\overleftrightarrow{V}_{11} \\ -\overleftrightarrow{R}_0\overleftrightarrow{V}_{10} - \overleftrightarrow{R}_1\overleftrightarrow{V}_{00} & \overleftrightarrow{1} - \overleftrightarrow{R}_1\overleftrightarrow{V}_{01} - \overleftrightarrow{R}_0\overleftrightarrow{V}_{11} \end{matrix} \right\} = 0 \quad . \tag{74}$$

Here an 8×8 matrix has been specified in terms of 4×4 submatrices whose positions in the 2×2 array indicate the layers they connect. In Sec. II, two methods were developed for factoring this determinantal equation into portions referring individually to the two, disjoint half-crystals: The determinant can simply be truncated to one of the 4×4 submatrices located on the diagonal of the above array. This may be called the truncation method and corresponds to Eqs. (20L) and (20R) above. Otherwise, the surface-state condition can be reexpressed in terms of $\mathcal{R}'(E)$, the resolvent for the already served crystal. This is the approach of Eq. (26) above and will be referred to as the presevered technique. It is this latter method which will be principally employed in the present work. We have used the former only to check certain key results. The prescription derived in Sec. II for the rearrangement of Eq. (74) requires that the 8×8 matrix appearing there be multiplied on the left by the inverse of the 8×8 matrix $\mathfrak{D}$, defined

$$\mathfrak{D} = \left\{ \begin{matrix} \overleftrightarrow{1} - \overleftrightarrow{R}_{-1}\overleftrightarrow{V}_{10} & -\overleftrightarrow{R}_0\overleftrightarrow{V}_{01} \\ -\overleftrightarrow{R}_0\overleftrightarrow{V}_{10} & \overleftrightarrow{1} - \overleftrightarrow{R}_1\overleftrightarrow{V}_{01} \end{matrix} \right\} \quad . \tag{75}$$

[$\mathcal{D}$ is just $P\mathfrak{D}P$, where $\mathfrak{D}$ was defined in Eq. (27).] This procedure amounts to dividing $\Delta(E)$ by $\Delta_S(E) = \det\mathcal{D}$, and converts Eq. (74) to the form

$$\det \left\{ \begin{matrix} \overleftrightarrow{1} - \overleftrightarrow{R}'_{00}\overleftrightarrow{V}_{00} & \overleftrightarrow{0} \\ \overleftrightarrow{0} & \overleftrightarrow{1} - \overleftrightarrow{R}'_{11}\overleftrightarrow{V}_{11} \end{matrix} \right\} = 0 \quad , \tag{76}$$

which clearly factors into the product of the determinants of the 4×4 matrices on the diagonal. These refer to the individual half-crystals, and when $\overleftrightarrow{V}_{00}$ is defined by Eq. (73), are actually identical. The 8×8 matrix of $\mathcal{R}'(E)$ restricted to the surface set is obtained in the form

$$\mathfrak{R}' \equiv \left\{ \begin{matrix} \overleftrightarrow{R}'_{00} & \overleftrightarrow{0} \\ \overleftrightarrow{0} & \overleftrightarrow{R}'_{11} \end{matrix} \right\} = \mathfrak{D}^{-1} \cdot \mathfrak{R} \quad , \tag{77}$$

where $\mathfrak{R}$ stands for the matrix of $\mathcal{R}(E)$ restricted to this set:

$$\mathfrak{R} = \left\{ \begin{matrix} \overleftrightarrow{R}_0 & \overleftrightarrow{R}_{-1} \\ \overleftrightarrow{R}_1 & \overleftrightarrow{R}_0 \end{matrix} \right\} \quad . \tag{78}$$

Ordinarily obtaining the inverse of the 8×8 matrix $\mathfrak{D}$ would be a very tiresome task. Here it is greatly simplified by the existance of the reflection symmetry discussed above. The properties of the two half-crystals must be identical. Every energy eigenvalue of the joint system is doubly degenerate with one eigenvector belonging to each side, the two being interchanged under the reflection. Now we can describe these eigenstates equivalently in terms of linear combinations even and odd under this reflection. All operators, $\mathcal{H}_0$, $\mathcal{V}_S$, and $\mathcal{V}_T$ are *separately* invariant under this reflection. Therefore, when a unitary transformation is applied to $\mathfrak{D}$ and $\mathfrak{R}$ to convert them to this new basis, they immediately become block diagonal, the even and odd parts completely decoupled. Specifically, $\mathfrak{D}$ transforms:

$$\mathfrak{D} \to \mathcal{O}^{\dagger}\mathfrak{D}\mathcal{O} = \left\{ \begin{matrix} \overleftrightarrow{1} - (\overleftrightarrow{R}_0 + \overleftrightarrow{U}\,\overleftrightarrow{R}_1)\overleftrightarrow{U}\,\overleftrightarrow{V}_{10} & \overleftrightarrow{0} \\ \overleftrightarrow{0} & \overleftrightarrow{1} + (\overleftrightarrow{R}_0 - \overleftrightarrow{U}\,\overleftrightarrow{R}_1)\overleftrightarrow{U}\,\overleftrightarrow{V}_{10} \end{matrix} \right\} \quad , \tag{79}$$

where the unitary transformation $\mathcal{O}$ is defined

$$\mathcal{O} = \frac{1}{\sqrt{2}} \left\{ \begin{matrix} \overleftrightarrow{1} & \overleftrightarrow{1} \\ \overleftrightarrow{U} & -\overleftrightarrow{U} \end{matrix} \right\} \quad . \tag{80}$$

The inverse matrices of the 4×4 submatrices appearing in (79) can be constructed relatively easily. These are now multiplied into the corresponding submatrices, $\overleftrightarrow{R}_0 \pm \overleftrightarrow{U}\,\overleftrightarrow{R}_1$, of the transformed version of $\mathfrak{R}$ and the result converted back to the original basis by the unitary transformation inverse to that in (79). Equation (77) now takes the form

$$\begin{Bmatrix} \overleftrightarrow{R}'_{00} & \overleftrightarrow{0} \\ \overleftrightarrow{0} & \overleftrightarrow{R}'_{11} \end{Bmatrix} = \frac{1}{2}\begin{Bmatrix} \overleftrightarrow{R}'_{(+)} + \overleftrightarrow{R}'_{(-)} & (\overleftrightarrow{R}'_{(+)} - \overleftrightarrow{R}'_{(-)})\overleftrightarrow{U} \\ \overleftrightarrow{U}\,(\overleftrightarrow{R}'_{(+)} - \overleftrightarrow{R}'_{(-)}) & \overleftrightarrow{U}\,(\overleftrightarrow{R}'_{(+)} + \overleftrightarrow{R}'_{(-)})\overleftrightarrow{U} \end{Bmatrix} , \tag{81}$$

where

$$\overleftrightarrow{R}'_{(\pm)} = [\overleftrightarrow{1} \mp (\overleftrightarrow{R}_0 \pm \overleftrightarrow{U}\,\overleftrightarrow{R}_1)\overleftrightarrow{U}\,\overleftrightarrow{V}_{10}]^{-1}(\overleftrightarrow{R}_0 \pm \overleftrightarrow{U}\,\overleftrightarrow{R}_1) . \tag{82}$$

The vanishing of the off-diagonal matrix entries in Eq. (81) requires that $\overleftrightarrow{R}'_{(+)}$ and $\overleftrightarrow{R}'_{(-)}$ be identical. This just reflects the fact that the even and odd states must be pairwise degenerate with essentially identical wave functions if, by taking linear combinations of them in turn, we are to produce states localized on each of the two sides.

We have performed the calculation indicated in Eq. (82). The ± sign was carried along throughout the calculation, and it could thus be explicitly verified that all terms depending on this sign drop out of the final result. This gives us a powerful check of the correctness of the algebraic manipulations as well as of the form, Eq. (59), of $\overleftrightarrow{R}_n$ itself. An additional check of the former aspect of the derivation was provided by a computer program which calculated numerically the various matrices and performed the indicated operations to obtain $\overleftrightarrow{R}'_{(\pm)}$ and thence $\overleftrightarrow{R}'_{00}$ and $\overleftrightarrow{R}'_{11}$. The numerical values of the matrix elements so obtained could then be compared to the corresponding values assumed by the expressions presented in Eq. (86) below.

Logically, as well as chronologically, the first interesting result obtained is $\Delta_S(E) = \det\mathfrak{D}$. The zeros of this quantity determine the Shockley states, which give rise to poles in the matrix elements of $\mathcal{R}'$ and, of course, have an intrinsic interest as well. The determinant of $\mathfrak{D}$ is secured as the product of the determinants of the even and odd submatrices in Eq. (79). These latter are found to be

$$\det[\overleftrightarrow{1} \mp (\overleftrightarrow{R}_0 \pm \overleftrightarrow{U}\,\overleftrightarrow{R}_1)\overleftrightarrow{U}\,\overleftrightarrow{V}_{10}] = (T_0 \pm T_1)\,\frac{2\mathcal{S} - \mathcal{C}}{2\mathcal{C}(\mathcal{C} - \mathcal{S})} , \tag{83}$$

from which we find that

$$\det\mathfrak{D} = \frac{(C + \mathcal{S})(2\mathcal{S} - \mathcal{C})^2}{4s_1s_2\mathcal{C}^2(\mathcal{C} - \mathcal{S})} . \tag{84}$$

Here and subsequently we use the following abbreviations:

$$c_i = \cosh\kappa_i \quad \text{and} \quad s_i = \sinh\kappa_i , \quad i = 1 \text{ or } 2 ;$$

$$\mathcal{C} = c_1 + c_2 \quad \text{and} \quad \mathcal{S} = s_1 + s_2 ;$$

also

$$T_0 = \frac{\mathcal{S}}{s_1s_2} \quad \text{and} \quad T_1 = \frac{c_1}{s_1} - \frac{c_2}{s_2} . \tag{85}$$

The even and odd linear combinations of Shockley states arise from the zeros of the quantities in Eq. (83) when one takes, respectively, the upper or lower sign. Though these expressions are not identical, their zeros in the region between the bands are so since they arise solely from the factor $(2\mathcal{S} - \mathcal{C})$. This is clear from their product, Eq. (84). The Shockley states will be discussed further below. Before leaving the subject of $\det\mathfrak{D}$, however, it is worth noting that one may here verify the correctness of an identity stated in Sec. II. Because of the reflection symmetry, the truncated forms of the determinant of $\mathfrak{D}$, $\det P_0[1 - \mathcal{R}(E)\mathcal{V}_S]$ and $\det P_1[1 - \mathcal{R}(E)\mathcal{V}_S]$, are real and, hence, equal to each other and to $\det\mathfrak{D}$ itself. The 4 × 4 determinant of one such projected form $\det(\overleftrightarrow{1} - \overleftrightarrow{R}_1\overleftrightarrow{V}_{01})$ which is trivially equal to the other projected quantity, was directly calculated. It was found to be indeed identical to Eq. (84). This represents one more check on the correctness of our results.

We now list the matrix elements of $\overleftrightarrow{R}'_{11}$:

$$R'_{11;00} = -\frac{E+H}{2\gamma^2}\left(\frac{\mathcal{C} - \mathcal{S}}{2\mathcal{S} - \mathcal{C}}\right) ,$$

$$R'_{11;10} = -\frac{i\sqrt{2}\,\sin(k_1/2)e^{-ik_2/4}}{\gamma}\left(\frac{\mathcal{C} - \mathcal{S}}{2\mathcal{S} - \mathcal{C}}\right) ,$$

$$R'_{11;20} = \frac{e^{+ik_2/4}}{\gamma}\left(\frac{\mathcal{C} - \mathcal{S}}{2\mathcal{S} - \mathcal{C}}\right)\{\tfrac{1}{2}(s_1 - s_2) + \cos(k_1/2) \times [2\cos(k_2/2) - i\sin(k_2/2)]\} ,$$

$$R'_{11;30} = \frac{\sqrt{2}\,e^{ik_2/4}}{\gamma}\left(\frac{\mathcal{C} - \mathcal{S}}{2\mathcal{S} - \mathcal{C}}\right)[\tfrac{1}{2}(s_1 - s_2) + \cos(k_1/2)\cos(k_2/2)] ,$$

$$R'_{11;11} = (E+H)^{-1}\left[1 - 4\sin^2\left(\frac{k_1}{2}\right)\left(\frac{\mathcal{C} - \mathcal{S}}{2\mathcal{S} - \mathcal{C}}\right)\right] , \tag{86}$$

$$R'_{11;21} = -\frac{i\sqrt{8}\,\sin(k_1/2)e^{ik_2/2}}{(E+H)}\left(\frac{\mathcal{C} - \mathcal{S}}{2\mathcal{S} - \mathcal{C}}\right)\{\tfrac{1}{2}(s_1 - s_2) + \cos(k_1/2)[2\cos(k_2/2) - i\sin(k_2/2)]\} ,$$

$$R'_{11;31} = -\frac{4i\sin(k_1/2)e^{ik_2/2}}{(E+H)}\left(\frac{\mathcal{C} - \mathcal{S}}{2\mathcal{S} - \mathcal{C}}\right)[\tfrac{1}{2}(s_1 - s_2) + \cos(k_1/2)\cos(k_2/2)] ,$$

$$\begin{aligned} R'_{11;22} = \frac{-2\mathcal{S}(2\mathcal{S} - \mathcal{C})}{\mathcal{C}^2(E+H)} + \frac{2(\mathcal{C} - \mathcal{S})}{\mathcal{C}^2(E+H)(2\mathcal{S} - \mathcal{C})}\{&(\mathcal{C}^2 - \mathcal{S}^2) \\ &\times [\tfrac{1}{4}(s_1 - s_2)^2 + \tfrac{1}{4}(c_1 - c_2)^2 + 2 - c_1c_2 - 2s_1s_2] \\ &- \mathcal{C}^2[+\cos^2(k_1/2) - c_1c_2 - 2s_1s_2] \\ &- (\mathcal{C} - \mathcal{S})^2\tfrac{1}{2}(s_1 - s_2)^2\} , \end{aligned}$$

$$R'_{11;32}=\frac{\sqrt{2}\,(\mathcal{C}-\mathcal{S})(2\mathcal{S}-\mathcal{C})}{\mathcal{C}^2(E+H)}+\frac{\sqrt{2}\,(\mathcal{C}-\mathcal{S})^2}{\mathcal{C}^2(E+H)(2\mathcal{S}-\mathcal{C})}$$

$$\times\{-i\mathcal{C}({}_1-s_2)\cos(k_1/2)\sin(k_2/2)$$

$$+\mathcal{S}[\tfrac{1}{4}(s_1-s_2)^2+\tfrac{1}{4}(c_1-c_2)^2$$

$$+2-c_1c_2-2s_1s_2]+(\mathcal{S}-\mathcal{C})\tfrac{1}{2}(s_1-s_2)^2\}\ ,$$

$$R'_{11;33}=\frac{(2\mathcal{C}-\mathcal{S})(2\mathcal{S}-\mathcal{C})}{\mathcal{C}^2(E+H)}-\frac{(\mathcal{C}-\mathcal{S})^3}{\mathcal{C}^2(E+H)(2\mathcal{S}-\mathcal{C})}$$

$$\times[\tfrac{3}{4}(s_1-s_2)^2+\tfrac{1}{4}(c_1-c_2)^2+2-c_1c_2-2s_1s_2]\ .$$

The remaining off-diagonal elements may be obtained by using the fact that $\overleftrightarrow{\mathrm{R}}'_{11}$ is Hermitian: $R'_{11;\alpha\beta}=R'^{*}_{11;\beta\alpha}$. Now in the form (82), the hermiticity of $\overleftrightarrow{\mathrm{R}}_{11}$ is not trivially apparent. Therefore, both $R'_{11;\alpha\beta}$ and $R'_{11;\beta\alpha}$ were calculated. The hermiticity was thus verified and a further check on the algebra achieved.

The surface perturbation (72) can now be combined with $\overleftrightarrow{\mathrm{R}}'_{11}(E)$ to yield the surface-state condition

$$\det[\overleftrightarrow{\mathrm{I}}-\overleftrightarrow{\mathrm{R}}'_{11}(E)\overleftrightarrow{\mathrm{V}}_{11}]=0\ . \tag{87}$$

In most cases of physical interest, the expression on the left is so complicated that it can only be dealt with numerically. There are, however, a few analytically obtainable results which are worth deriving here. Of course, the surface states belong to two-dimensional bands characterized by the wave vector (k_1, k_2). The energy function of such a band $\mathcal{E}(k_1, k_2)$ must be periodic in both k_1 and k_2 with periods of 2π. As discussed in Ref. 1, the first Brillouin zone of the two-dimensional reciprocal lattice is just the region $-\pi<k_1, k_2\leq\pi$. It is worth noting that $\mathcal{E}(k_1, k_2)$ is an even function both of k_1 and of k_2:

$$\mathcal{E}(k_1,k_2)=\mathcal{E}(-k_1,k_2)=\mathcal{E}(-k_1,-k_2)\ . \tag{88}$$

The former equality follows from the reflection symmetry of the lattice in a plane perpendicular to the lattice vector $\vec{a}$ and passing through a lattice point. The second equality is a consequence of time reversal symmetry.

As mentioned above, the Shockley states are determined by the condition

$$2\mathcal{S}=\mathcal{C}\ , \tag{89}$$

where $\mathcal{S}$ and $\mathcal{C}$ were defined in Eq. (85). It should be noted that the zeros of Eq. (84) are double. This is because they represent the Shockley states of both half-crystals. Since the two sides have identical properties, all the Shockley states have twofold degeneracy. The quantities $\mathcal{S}$ and $\mathcal{C}$ in Eq. (89) must be expressed in terms of the energy, as well as of the components k_1 and k_2 of the two-dimensional wave vector. Writing ξ for $\cos(k_1/2)\times\cos(k_2/2)$, the required relations are most naturally expressed in the form

$$\mathcal{S}=[(\tfrac{1}{2}\mathcal{C}+\xi)^2-1]^{1/2}+[(\tfrac{1}{2}\mathcal{C}-\xi)^2-1]^{1/2} \tag{90}$$

and

$$E^2+\gamma^2\mathcal{C}^2=H^2+4\gamma^2[4-\cos^2(k_1/2)\sin^2(k_2/2)]\ . \tag{91}$$

The values of $\mathcal{C}$ and $\mathcal{S}$ increase as the energy E recedes from the edge of the conduction or valence band into the forbidden gap. They are unchanged when the sign of E is reversed. Their minima at the conduction-band edge are, respectively, $2(1+\xi)$ and $2[\xi(1+\xi)]^{1/2}$. We can eliminate $\mathcal{S}$ between (89) and (90) to obtain a relation between $\mathcal{C}$ and ξ. By then requiring that the value of $\mathcal{C}$ for the surface state exceed the band edge value, we find that Shockley states are present only near the edges of the two-dimensional zone, in particular, only for

$$\xi\equiv\cos(k_1/2)\cos(k_2/2)\leq\tfrac{1}{3}\ . \tag{92}$$

When present, these states appear symmetrically below the conduction band and above the light hole valence band. Because of the form of Eq. (91), the simplest way to express the binding energy of the Shockley state is in terms of the difference in the values of E^2 for band edge and surface state:

$$(E^2)_{\mathrm{BE}}-(E^2)_{\mathrm{ss}}=\tfrac{4}{3}\gamma^2(3\xi-1)^2\ . \tag{93}$$

Some simple properties of the Tamm states can be investigated by permitting only ϵ_0 in Eq. (72) to be nonzero. This quantity represents the change in the value of the Coulomb integral for the surface cation. As discussed in Ref. 1, it is expected to be negative and rather smaller in magnitude than H. When only ϵ_0 is nonzero, only the first column of the 4×4 matrix product in Eq. (87) has nonzero entries. Thus this determinantal equation reduces to

$$1-\epsilon_0R'_{11;00}(E)=0\ . \tag{94}$$

When the expression for $R'_{11;00}$ is substituted from (86), this condition can be rewritten in a form closely analogous to the Shockley state condition, Eq. (89):

$$(2+\alpha)\mathcal{S}=(1+\alpha)\mathcal{C}\ , \tag{95}$$

where α is defined

$$\alpha=-\epsilon_0(E+H)/2\gamma^2\ . \tag{96}$$

For negative ϵ_0, α is positive for energies above the flat bands at $E=-H$ and negative below them. Once again $\mathcal{S}$ can be eliminated between Eqs. (95) and (90). Now, however, the relation between $\mathcal{C}$ and ξ involves α as well. The threshold condition may nonetheless be easily evaluated in closed form. The surface states are present only for

$$\xi\leq(1+\alpha)^2/(3+2\alpha)\ , \tag{97}$$

where α is evaluated from Eq. (96) using the expression for the band edge energy. Where Eq. (97) holds as an equality, the surface-state band meets the edge of the bulk continuum. Equation (97) can be used to study the qualitative behavior of the Tamm state bands. As the magnitude of ϵ_0 increases from zero, the binding energy of the erstwhile Shockley state below the conduction band increases, and its region of existence expands to larger values of ξ in accordance with Eq. (97). The Γ point is the last point in the two-dimensional zone to be included in the domain of this band of surface states. The threshold value of α for a surface state at this last point is $\sqrt{2}$, which corresponds to

$$(\epsilon_0)_{\text{Threshold}} = -\sqrt{2}\,\gamma^2/H \quad . \tag{98}$$

In the meantime, since α is negative below the flat band where the lower Shockley state was found, the binding energy of this state now diminishes and its region of existence shrinks toward the zone edge. Since $\xi = 0$ defines the zone edge, this state finally disappears at a given point on the zone boundary when $\alpha = -1$ there or $\epsilon_0 = -\gamma^2/(|E_{\text{BE}}| - H)$. The state persists longest at that point at the zone edge where the magnitude of the band edge energy is smallest. This is the X' point,[24] $(k_1, k_2) = (0, \pi)$, for which $E_{\text{BE}} = -(H^2 + 12\gamma^2)^{1/2}$. As a matter of interest, if ϵ_0 is positive, the domain of the upper surface state shrinks, vanishing for $\alpha = -1$ while that of the lower state grows. However, the surface state associated with the lower valence band is constrained to lie below the flat bands; hence, though its domain of existence expands towards zone center as ϵ_0 increases, it never actually arrives there.

The relation between $\mathcal{C}$ and ξ for the Tamm state found from Eqs. (95) and (90) is

$$\mathcal{C}^2 = 4(2+\alpha)^2[\xi^2/(1+\alpha)^2 + 1/(3+2\alpha)] \quad , \tag{99}$$

valid in the region defined by Eq. (97). This relation can be used to find the binding energy of the surface state by substituting in Eq. (91) and solving the resulting equation simultaneously with Eq. (96), defining α. Two useful results can be extracted without numerical computation: (i) The dependence of $\mathcal{C}^2$ on ξ^2 alone insures that the resulting surface state band has the proper repeat period, 2π, in k_1 and k_2; (ii) just at threshold, when the surface-state band is tangent to the edge of the conduction band at zone center, the masses of the surface-state band and of the bulk conduction band at zone center are equal.[25]

For a final analytic result, we observe that all the off-diagonal matrix elements $R'_{11;\beta\alpha}$ in Eq. (86) which have $\alpha = 1$ or $\beta = 1$ are proportional to $\sin(k_1/2)$ and thus vanish for $k_1 = 0$. The same is true for the corresponding elements of $\overleftrightarrow{V}_{11}$ given in Eq. (72). Thus along the line $k_1 = 0$ the determinant in Eq. (87) factors into the product of the (1, 1) element and its minor. The former factor has the form

$$[1 - \epsilon_1 (E+H)^{-1}] \quad . \tag{100}$$

Since ϵ_1 is expected to be positive,[1] there will surely be a band of surface states which is flat along this line in the two-dimensional zone with energy $E = -H + \epsilon_1$. By specializing yet further to the center of the two-dimensional zone $k_2 = 0$ and setting E equal to H the energy at the edge of the conduction band, one may readily derive a threshold condition from Eq. (87). That is, one finds the zeros of the determinant considered now as a function of the parameters appearing in $\overleftrightarrow{V}_{11}$. At such a zero, there is a surface state just at the edge of the continuum. Since the energy of a surface state will, for the most part, be a continuous function of the parameters of the perturbation, and since such a state can appear or disappear only through the band edge, this threshold condition and the corresponding one for the lower band edge can be used to characterize the maximum possible number of surface states present at zone center for any given set of parameter values. We shall not do this here since such threshold conditions will be thoroughly discussed in Ref. 1.

The analytic results in these few special cases exemplify many of the expected features of the surface-state bands. Obviously, however, we cannot extract by these means the physically interesting, quantitative consequences of a realistic model of the surface perturbation. For this reason, a computer program was written to evaluate the determinant in Eq. (87) numerically and find its roots. A number of checks on the correctness of the algebraic procedures culminating in Eq. (86) have already been mentioned. In debugging the computer program, further tests were carried out with the primary goal of verifying that the program correctly evaluated the determinant in Eq. (87) using Eqs. (86) and (72) and accurately found its roots, but which also served to test yet further the correctness of Eq. (86) itself. Most notably the ansatz method of solving the surface-state problem was adapted to verify the correctness of the threshold condition obtained via the resolvent scheme, and its predictions were then tested against the computer program. Furthermore this ansatz technique was used to calculate zone center binding energies for selected parameter values.[26] The binding energies directly and stringently tested both Eq. (86) and the computer program. The ansatz method employed and the nature of the tests based on it are described in detail in Ref. 1.

IV. CONCLUSION

Within the context of a MO description of the electronic states of a semiconductor, we have discussed three general techniques of solving the equations of motion for the surface states without introducing additional approximations. Two are based on the resolvent or Koster-Slater method and simply represent different techniques for taking advantage of a symmetry peculiar to the surface problem to reduce the dimensionality of the Koster-Slater determinant by half. The truncation scheme has been the more commonly employed of the two. It will ordinarily be the more convenient, especially for numerical calculations carried out via computer. We found that this approach can lead to spurious solutions. Fortunately, these can be readily recognized. The energies of the spurious solutions were shown to coincide with those of the Shockley states for the wrong half-crystal. The alternative, pre-severed method lacks spurious solutions but requires more algebraic manipulations. The characteristic determinant in this case also has poles at the Shockley states of the right half-crystal. These can be inconvenient in a computer calculation. However, the pre-severed technique can be very helpful in obtaining analytical results, especially if the surface perturbation involves fewer surface orbitals than are affected by the severing *per se*. The third method studied was the familiar ansatz approach, in which one guesses the general form of the surface-state wave function and substitutes it in the Schrödinger equation. Using the integral representation of the resolvent, the correct general form of the trial wave function was determined. It turned out that this form could be predicted from simple properties of the MO representation of the bulk band structure, and for this purpose the explicit computation of the resolvent was unnecessary. The ansatz technique, which is now of general applicbility, will ordinarily be less useful than the resolvent methods, but it always provides a powerful check on the correctness of results otherwise obtained.

These general methods were applied to the study of the surface states associated with a (110) face of a partially ionic semiconductor having the zinc-blende structure. The resolvent was constructed for a particular representation of the valence and conduction bands using a basis adapted to a (110) face. The pre-severed resolvent technique was mainly employed, but both of the other methods were used to check the algebraic correctness of the results. The necessary steps were carried through to set up a program for the subsequent numerical calculation of the properties of the surface states using a computer.[1] In addition, a number of analytical results were obtained, based on highly over-simplified representations of the surface perturbation. These illustrate features of the surface-state bands to be expected also for the more sophisticated models of the surface perturbation. Salient properties include (i) an acceptorlike surface-state band relatively close below the conduction band, which is present in the forbidden gap only if a threshold condition is satisfied. This threshold condition requires a sufficiently large decrease of the Coulomb integral for the surface cations. The zone center mass of this surface-state band approximates that of the conduction band. (ii) A light donorlike surface-state band would lie compressed beneath the heavy hole band and would probably be unobservable by the experimental techniques usually employed. (iii) Heavy donorlike surface-states bands are to be expected in the forbidden gap when the Coulomb integrals for the surface anions are increased. The mass of such a band should be comparable to that of a heavy hole band. Its binding energy relative to the valence-band edge should exceed that of the light, acceptorlike band minimum relative to the conduction-band edge.

ACKNOWLEDGMENTS

This work was largely motivated by Jules D. Levine's interest in treating the specific problem discussed in Sec. III. I am also indebted to him for extensive discussions concerning the more physical aspects of the material in that section.

A number of formal and mathematical points relating to the material in Appendix A were brought out or clarified during discussions with Edward Williams. Some of the mathematical results were also discussed with Christopher B. Davis.

APPENDIX A

The object of this Appendix is to prove Eq. (23) together with some interesting related identities. The proof proceeds in several steps, of which the first is to establish the relation

$$\det P_0(1-\mathcal{R}\mathcal{V})\det P_1(1-\mathcal{R}\mathcal{V})=\det P(1-\mathcal{R}\mathcal{V}) \times \frac{\det P_0\mathcal{R}\det P_1\mathcal{R}}{\det P\mathcal{R}} \quad . \tag{A1}$$

The reasoning involved is essentially the same as that used in Sec. II to prove Eqs. (32) and (34). We now define

$$\mathcal{R}''(E)=(E-\mathcal{H}_0-\mathcal{V}_S-\mathcal{V}_T)^{-1} \tag{A2}$$

and

$$\mathcal{D}'(E)=1-\mathcal{R}(E)(\mathcal{V}_S+\mathcal{V}_T) \quad . \tag{A3}$$

It is then possible to show, just as before, that

$$P\mathcal{R}''(E)P=P[P\mathcal{D}'(E)P]^{-1}P\mathcal{R}(E)P \quad . \tag{A4}$$

After rearranging Eq. (A4) to read

$$P\mathcal{D}'(E)P = P\mathcal{R}(E)P[P\mathcal{R}''(E)P]^{-1}P \quad , \qquad \text{(A5)}$$

we can make use of the fact that, in turn, $\mathcal{H}_0+\mathcal{V}$, $\mathcal{R}''(E)$, $P\mathcal{R}''(E)P$, and the last factor on the right in Eq. (A5) have no matrix elements connecting the two half-crystals. Thus, we see that

$$P_i\mathcal{D}'(E)P_j = P_i\mathcal{R}(E)P_j[P_j\mathcal{R}''(E)P_j]^{-1}P_j \quad , \qquad \text{(A6)}$$

where i and j run over the indices 0 and 1 denoting the surface sets on the two sides. On taking determinants, we find from Eq. (A5) that

$$\det P(1-\mathcal{R}\mathcal{V}) = \frac{\det P\mathcal{R}}{\det P_0\mathcal{R}''\det P_1\mathcal{R}''} \qquad \text{(A7)}$$

and from Eq. (A6) that

$$\det P_i(1-\mathcal{R}\mathcal{V}) = \frac{\det P_i\mathcal{R}}{\det P_i\mathcal{R}''} \quad . \qquad \text{(A8)}$$

Equation (A1) follows immediately from Eqs. (A7) and (A8).

It is worth noting that there is a good deal of freedom in the choice of P_0 and P_1 in these equations. Suppose we write $P(\mathcal{V})$ for the projection operator onto the *smallest* possible surface set S consistent with the defining requirements: $\mathcal{V}P=\mathcal{V}$. Similarly, $P_0(\mathcal{V})$ and $P_1(\mathcal{V})$ resolve $P(\mathcal{V})$ into portions referring to the left and right half-crystals, as usual. Now our proof of Eq. (A1) requires only that P_0 be at least as large as $P_0(\mathcal{V})$, that is, $P_0\cdot P_0(V)=P_0(V)$ and the same for P_1 and $P=P_0+P_1$. Note, however, that each of the two factors on the left in Eq. (A1) as well as the first factor on the right are individually independent of the choice of P_0 and P_1. For example, in the case of $\det P(1-\mathcal{R}\mathcal{V})$, all columns except those projected on by $P(\mathcal{V})$ contain only a single unit entry on the diagonal. Expanding in minors of each extra column in turn, one reduces this determinant to $\det P(\mathcal{V})(1-\mathcal{R}\mathcal{V})$. Consequently, the factor in square brackets on the right side of Eq. (A1) must also be independent of the choice of P_0 and P_1 even though its numerator and denominator individually may be expected to depend on this choice.

With the aid of Eqs. (A7) and (A8), it is possible to proceed some distance in characterizing the zeros and poles of the various determinants appearing in Eq. (A1). The strongest results are obtained when the original, unsevered crystal characterized by $\mathcal{H}_0$ is already finite though large, so that *all* its eigenvalues are discrete. Note that no periodic boundary conditions are employed. We shall assume the crystal to be finite in this Appendix except when it is explicitly stated otherwise. From the definition of $\mathcal{R}''(E)$, Eq. (A2), it is clear that $\det P_0\mathcal{R}''(E)$ has poles precisely at the eigenvalues of the left half-crystal. Similarly, $\det P_1\mathcal{R}''(E)$ has poles at the eigenvalues of the right half-crystal and nowhere else. Now in Eq. (A7), we know that the zeros of $\det P(1-\mathcal{R}\mathcal{V})$ correspond in location and order to the eigenvalues of the system described by $\mathcal{H}_0+\mathcal{V}$. In view of the possibility of shifting any eigenvalue of this final system by modifying $\mathcal{V}_T$, it is clear that the orders of the poles of $\det P_i\mathcal{R}''(E)$ must correspond to the multiplicities of the corresponding eigenvalues. Thus the zeros of the left-hand side of Eq. (A7) match up with the poles of the denominator of the right-hand side: Any zeros of $\det P\mathcal{R}(E)$ must be cancelled by corresponding zeros in $\det P_0\mathcal{R}''(E)$ or $\det P_1\mathcal{R}''(E)$. Now the poles of the left-hand side of Eq. (A7) are located at the eigenvalues of $\mathcal{H}_0$. In view of possibility of including in $\mathcal{H}_0$ a perturbation which affects only one side of the crystal, it is clear that these poles must be matched as to order as well as location by the poles of $\det P\mathcal{R}(E)$, so that there are no zeros of $\det P_i\mathcal{R}''(E)$ *except* those which correspond to the zeros of $\det P\mathcal{R}(E)$. By exactly the same argument, the poles of $\det P_i(1-\mathcal{R}\mathcal{V})$ on the left side of Eq. (A8) are matched by the poles of $\det P_i\mathcal{R}$. Thus the zeros of $\det P_i\mathcal{R}''$ must be cancelled by zeros in $\det P_i\mathcal{R}$. The zeros of $\det P_i(1-\mathcal{R}\mathcal{V})$ include zeros at the poles of $\det P_i\mathcal{R}''$. These latter correspond in location and order to the true eigenstates of the corresponding half-crystal. This is not new information though it has been obtained in a new fashion. However, $\det P_i(1-\mathcal{R}\mathcal{V})$ may have additional zeros if $\det P_i\mathcal{R}$ has more zeros than $\det P_i\mathcal{R}''$. These will *not* correspond to eigenstates of the corresponding half-crystal. That such extra zeros do indeed exist can be shown with the aid of Eq. (24) of the text. This relation follows from the identity

$$\{\det P[1-(P_0\mathcal{R}P_1)(P_1\mathcal{V}_SP_0)]\}^* = \det P[1-(P_1\mathcal{R}P_0)(P_0\mathcal{V}_SP_1)] \quad , \qquad \text{(A9)}$$

which is established by taking the Hermitian conjugate of the matrix in square brackets on the left and then using the lemma of Ref. 15 to reverse the order of the operators. Equation (24) is now obtained when one eliminates the extra columns on each side in the fashion discussed in the previous paragraph. From Eq. (24) alone it follows that $\det P_i(1-\mathcal{R}\mathcal{V}_S)$ has zeros corresponding in energy and multiplicity to the Shockley states of *both* sides. (This can be made completely unequivocal by including in $\mathcal{H}_0$ an infinitesimal perturbation to lift any possible degeneracy between Shockley states of left and right half-crystals.) When $\mathcal{V}_T$ is set to zero in Eq. (A8), the poles of $\det P_i\mathcal{R}'$ account only for the zeros corresponding to the Shockley states of the i side; the additional zeros of $\det P_i(1-\mathcal{R}\mathcal{V}_S)$, which correspond in location and order to the Shockley states of the other side,[27] must arise as zeros of $\det P_i\mathcal{R}$. These spurious zeros at the Shockley states of the wrong side obviously persist in $\det P_i\mathcal{R}$ and thus in $\det P_i(1-\mathcal{R}\mathcal{V})$ whatever the potential $\mathcal{V}_T$ may be.

It remains to prove that

$$\frac{\det P_0\mathcal{R}\,\det P_1\mathcal{R}}{\det P\mathcal{R}}=\det P(1-\mathcal{R}\mathcal{V}_S)\ . \tag{A10}$$

The results of the previous paragraph suffice to show that the quantity on the left has a zero wherever $\det P(1-\mathcal{R}\mathcal{V}_S)$ does. Since we have stipulated that the original crystal is finite though very large, these zeros give the eigenvalues of *all* the eigenstates of the severed system. It is important to note that under these circumstances $\det P(1-\mathcal{R}\mathcal{V}_S)$ is analytic except for a finite number of poles at the eigenvalues of $\mathcal{H}_0$. Obviously the same is true of the left-hand side of Eq. (A10). By making use of (i) the characterization of the poles of $\det P_i\mathcal{R}$ in relation to those of $\det P_i(1-\mathcal{R}\mathcal{V}_S)$ and of those of $\det P\mathcal{R}$ in relation to those $\det P(1-\mathcal{R}\mathcal{V}_S)$ which was carried out in the last paragraph, (ii) the freedom to redefine P_i as larger than $P_i(\mathcal{V}_S)$ discussed in the paragraph before, and (iii) the possibility of considering in turn the consequences of severing the crystal to the *right* of S_1 or to the *left* of S_0, it is possible to prove that the orders of the poles of the quantities on the two sides of Eq. (A10) are identical. It follows that

$$\left[\frac{\det P_0\mathcal{R}\,\det P_1\mathcal{R}}{\det P\mathcal{R}}\right]\Big/\det P(1-\mathcal{R}\mathcal{V}_S) \tag{A11}$$

is an entire analytic function of complex E. Since its limit as $|E|\to\infty$ is unity, Eq. (A10) follows from Liouville's theorem.[28] Having proved Eq. (A10), we could now complete the characterization of the zeros of $\det P\mathcal{R}$ and $\det P_i\mathcal{R}$ using arguments similar to those employed above. It turns out that $\det P_i\mathcal{R}$ has a zero for each eigenvalue of *either* of the half-crystals which one obtains on *removing* from the original crystal all the ions corresponding to the surface set S_i. Note that in the case of $\det P_0\mathcal{R}$, when one removes S_0, the half-crystal to the right is just our usual right half-crystal, but that to the left does not coincide with the left half-crystal which we have studied previously. A more general version of this theorem will be proved below by another method.

There is another method of proving Eq. (A10) which is rather indirect, but is more physical and requires less deviousness than that sketched in the preceding paragraph. This alternate method makes use of the possibility, discussed in Sec. II, of using a potential barrier $\mathcal{V}_B$ to decouple the crystal, rather than breaking the bonds connecting the two halves via $\mathcal{V}_S$. If $\mathcal{V}_B$ is defined as in Eq. (9) of the text, then the eigenvalues and eigenfunctions of the states belonging to the right half-crystal in the limit $\epsilon\to 0+$ will be the same as those obtained when the crystal is broken by $\mathcal{V}_S$.[29] To avoid questions of convergence, it will be assumed here that the crystal is finite in the dimension normal to the surface. After Eq. (A10) has been proved for this case, we can let this dimension become infinite again. The invariance of the eigenvalues E_ρ and corresponding eigenvector components $\Psi_\rho(m\beta)$ under this change of Hamiltonian suffice to establish this same property for the matrix elements of the resolvent for the severed crystal $\mathcal{R}'(E)$, since these matrix elements can be represented in the form

$$\langle m\beta|\mathcal{R}'(E)|l\alpha\rangle=\sum_\rho\frac{\Psi_\rho(m\beta)\Psi_\rho(l\alpha)^*}{E-E_\rho}\ . \tag{A12}$$

The basis elements $m\beta$ and $l\alpha$ here refer to the right half-crystal. (As usual, the transverse wave vector $\vec{k}_\parallel$ is suppressed and a one-dimensional terminology employed.) In short, we can equate those matrix elements of $\mathcal{R}^{(1)\prime}(E)=(E-\mathcal{H}_0-\mathcal{V}_S)^{-1}$ and of $\mathcal{R}^{(2)\prime}(E)=(E-\mathcal{H}_0-\mathcal{V}_B)^{-1}$ which refer exclusively to the right half-crystal. As we shall see below this identification implies the relation (A10) directly.

To relate the matrix elements of these operators, we shall need to consider further the inverse of operators of the form $\mathcal{D}^{(1)}=1-\mathcal{R}\mathcal{V}_S$, where the essential property of $\mathcal{V}_S$ is that $\mathcal{V}_S(1-P)=0$. As shown in Sec. II, the inverse of such an operator characteristically satisfies identities of the form given in Eqs. (30) and (31) of the text. It remains to determine $(1-P)\mathcal{D}^{(1)-1}P$: Multiply the defining relation $\mathcal{D}^{(1)-1}\mathcal{D}^{(1)}=1$ through by $(1-P)$ on the left and by P on the right. Then it follows that

$$[(1-P)\mathcal{D}^{(1)-1}P](P\mathcal{D}^{(1)}P)+(1-P)\mathcal{D}^{(1)-1}P=0\ , \tag{A13}$$

where Eq. (30) has been employed. Substituting $\mathcal{D}^{(1)}=1-\mathcal{R}\mathcal{V}_S$, this becomes

$$(1-P)\mathcal{D}^{(1)}(E)^{-1}P=(1-P)\mathcal{R}(E)\mathcal{V}_S P[P\mathcal{D}^{(1)}(E)P]^{-1}P\ . \tag{A14}$$

This is the required relation. The analogously defined quantity $\mathcal{D}^{(2)}(E)=1-\mathcal{R}(E)\mathcal{V}_B$ satisfies relations having the forms of Eqs. (30), (31), and (A14) with P replaced by P_0 and $\mathcal{V}_S$ by $\mathcal{V}_B$.

We can now construct the relevant matrix elements of $\mathcal{R}^{(2)\prime}(E)$. We find that

$$\begin{aligned}P_0\mathcal{D}^{(2)}(E)^{-1}P_0&=P_0\{\epsilon[P_0\mathcal{R}(E)P_0]^{-1}\}P_0\ ,\\ (1-P_0)\mathcal{D}^{(2)}(E)^{-1}P_0&=-(1-P_0)\mathcal{R}(E)P_0[P_0\mathcal{R}(E)P_0]^{-1}P_0\ ,\end{aligned} \tag{A15}$$

and

$$\mathcal{D}^{(2)}(E)^{-1}(1-P_0)=(1-P_0)\ ,$$

of which the first two lines are correct only to lowest order in ϵ. Forming $\mathcal{R}'^{(2)}(E)=\mathcal{D}^{(2)}(E)^{-1}\mathcal{R}(E)$, we obtain

$$(1-P_0)\mathcal{R}''^{(2)} = (1-P_0)\mathcal{R}[1-P_0(P_0\mathcal{R}P_0)^{-1}P_0\mathcal{R}] \quad , \tag{A16}$$

which is exact in the limit $\epsilon \to 0+$. In particular, we shall use

$$\mathcal{R}_{11}''^{(2)} = \mathcal{R}_{11} - \mathcal{R}_{10}(\mathcal{R}_{00})^{-1}\mathcal{R}_{01} \quad , \tag{A17}$$

where matrix subscripts are written in place of projection operators P_0 and P_1.

In the case of $\mathcal{R}''^{(1)}(E)$, we can use the relation

$$\mathcal{R}_{10}''^{(1)} = 0 = (\mathcal{D}^{(1)-1})_{11}\mathcal{R}_{10} + (\mathcal{D}^{(1)-1})_{10}\mathcal{R}_{00} \quad , \tag{A18}$$

to eliminate $(\mathcal{D}^{(1)-1})_{10}$ from the expression for $\mathcal{R}_{11}''^{(1)}$. We then find

$$\mathcal{R}_{11}''^{(1)} = (\mathcal{D}^{(1)-1})_{11}[\mathcal{R}_{11} - \mathcal{R}_{10}(\mathcal{R}_{00})^{-1}\mathcal{R}_{01}] \quad . \tag{A19}$$

Comparing Eq. (A17) with Eq. (A19), it is clear that

$$(\mathcal{D}^{(1)-1})_{11} = P_1 \quad , \tag{A20}$$

the unit matrix within the S_1 subspace. From Eq. (A18) it now follows

$$(\mathcal{D}^{(1)-1})_{10} = -\mathcal{R}_{10}(\mathcal{R}_{00})^{-1} \quad . \tag{A21}$$

Since $\mathcal{D}^{(1)}$ treats the S_0 and S_1 subspaces on a symmetrical footing, it is clear that

$$P\mathcal{D}^{(1)-1}P = \begin{Bmatrix} P_0 & -\mathcal{R}_{01}(\mathcal{R}_{11})^{-1} \\ -\mathcal{R}_{10}(\mathcal{R}_{00})^{-1} & P_1 \end{Bmatrix} \quad . \tag{A22}$$

Taking the determinants of both sides, we obtain

$$[\det P(1-\mathcal{R}\mathcal{V}_S)]^{-1} = \frac{\det P\mathcal{R}}{\det P_0\mathcal{R}\,\det P_1\mathcal{R}} \quad , \tag{A23}$$

which is just Eq. (A10).

There are a number of corollaries which are worth mentioning. Equation (A17) and (A22) are of interest in their own right as explicit expressions for $\mathcal{R}_{11}'$ and $P\mathcal{D}^{-1}P$ of the text. They were given there in Eqs. (36) and (35), respectively. Another expression for $\mathcal{R}_{11}'$ can be obtained from Eq. (A6) by setting $\mathcal{V}_T$ to zero and rearranging:

$$\mathcal{R}_{11}' = (\mathcal{D}_{11})^{-1}\mathcal{R}_{11} \quad . \tag{A24}$$

Note that $(\mathcal{D}_{11})^{-1}$ and $(\mathcal{D}^{-1})_{11}$ are quite different quantities. An explicit expression for the former can be obtained by combining (A24) with (A17). Obviously an expression for $\mathcal{R}_{11}''$, in case $\mathcal{V}_T$ is not zero, can be obtained from (A6) in the same manner. It differs from (A24) only in having $\mathcal{D}$ replaced by $\mathcal{D}'$ defined in Eq. (A3).

When the perturbation $\mathcal{V}_B$ is introduced into the original (finite) crystal described by $\mathcal{H}_0$, the eigenvalues of the resulting Hamiltonian $\mathcal{H}_0 + \mathcal{V}_B$ are determined by the resolvent technique to be the roots of the determinant, $\det P_0[1-\mathcal{R}(E)P_0 1/\epsilon]$. Those roots which approach finite limits as $\epsilon \to 0+$ become in this limit the zeros of $\det P_0\mathcal{R}(E)$. As was shown in Ref. 29, these eigenvalues correspond to eigenvectors having vanishing amplitude on the set S_0. Thus the roots of $\det P_0\mathcal{R}(E)$ are just the eigenvalues of the two disconnected half-crystals obtained by removing the set S_0 entirely. This theorem was alluded to following Eq. (A11) above. It can be easily generalized and the proof sketched here made quite rigorous. The general statement is that if Σ is any set of basis elements whatsoever and Π projects on Σ, then the zeros of $\det\Pi\mathcal{R}(E)$ are just the eigenvalues of the crystal described by $\mathcal{H}_0$ but with the set Σ removed, i.e., all rows and columns of $\mathcal{H}_0$ referring to this set are to be deleted.

The construction of the matrix elements of $\mathcal{D}^{-1}$ summarized in Eq. (A14) can be of considerable utility when the surface set on which $\mathcal{V}_T$ acts, $P(\mathcal{V})$, is actually larger than that implicated in the decoupling of the two half-crystals $P(\mathcal{V}_S)$. In this case, the finite-dimensional matrix $[P(\mathcal{V}_S)\mathcal{D}P(\mathcal{V}_S)]^{-1}$ can first be constructed, and then the remaining matrix elements required in the construction of $P_1(\mathcal{V})\mathcal{R}'P_1(\mathcal{V})$, say for use in Eq. (26), can be obtained from (A14).

Finally, Eqs. (A16) and (A22) give rise to numerous identities involving the matrix elements of the resolvent. In the case of (A22), these are precisely the identities required to prove directly that $\mathcal{R}'$ is block diagonal. These identities can be used to check the correctness of the algebraic expressions for the components of $\mathcal{R}$ derived according to the prescription presented at the end of Sec. II. Implicitly, this was one of the checks performed in the case of the example described in Sec. III.

APPENDIX B

In this Appendix, two useful extensions of the methods developed in Sec. II will be presented: The modifications required in order to treat reconstructed surfaces will be described, and the method used by Baldock[20] to treat defects located on or near a crystal's surface will be generalized. In addition, the use of the resolvent to study virtual surface states will be briefly discussed.

By definition, a reconstructed surface has a periodicity lower than that of a parallel crystallographic plane in the bulk. Instead of primitive translations $\vec{a}$ and $\vec{b}$, the potential $\mathcal{V} = \mathcal{V}_S + \mathcal{V}_T$ which characterizes the surface has periods $\mu_1\vec{a}$ and $\mu_2\vec{b}$. Here μ_1 and μ_2 are positive integers, of which at least one is greater than unity. We adopt a new description of the crystal lattice adapted to this lowered symmetry. The basis state $|\vec{l}\alpha\rangle$, where $\vec{l} = l_1\vec{a} + l_2\vec{b} + l\vec{c}$, is now described as $|\vec{L}\alpha\alpha_1\alpha_2\rangle$. Here $\vec{L} = L_1(\mu_1\vec{a}) + L_2(\mu_2\vec{b}) + l\vec{c}$ with $l_i = \mu_i L_i + \alpha_i$, for $i = 1$ and 2. The integer α_i runs from 0 to $\mu_i - 1$. Clearly, the primitive cell has simply been expanded in the $\vec{a}$ and $\vec{b}$ directions by factors μ_1 and μ_2, respectively. The reciprocal lattice is now based on the dual vectors $\mu_1^{-1}\vec{a}^*$, $\mu_2^{-1}\vec{b}^*$ and $\vec{c}^*$. We write a wave vector $\vec{p}$ in the primitive cell of the reciprocal lattice in the form $\vec{p}=p_1(\mu_1^{-1}\vec{a}^*)+p_2(\mu_2^{-1}\vec{b}^*)+k\vec{c}^*$. Here, p_1, p_2, and k are chosen to run from 0 to 2π. Each band $\lambda_\rho(\vec{k})$ according to our old scheme now goes over into $\mu_1\mu_2$ bands, $\lambda_\rho(\gamma_1,\gamma_2,\vec{p})$. Here γ_i is an integer between 0 and μ_i-1, and we identify $\lambda_\rho(\gamma_1,\gamma_2,\vec{p})$ with $\lambda_\rho(\vec{k})$ for $k_i=\mu_i^{-1}[p_i+2\pi\gamma_i]$, $i=1$ and 2. Layer orbitals are now defined in analogy with Eq. (10);

$$|\vec{p}_\parallel;l\alpha\alpha_1\alpha_2\rangle=\sum_{L_1,L_2} e^{i\vec{p}_\parallel\cdot\vec{L}}|\vec{L}\alpha\alpha_1\alpha_2\rangle \tag{B1}$$

and the reduced matrix elements of $\mathcal{V}$, in analogy to Eq. (13), have the form

$$\langle\vec{p}_\parallel;m\beta\beta_1\beta_2|\mathcal{V}(\vec{p}_\parallel)|\vec{p}_\parallel;l\alpha\alpha_1\alpha_2\rangle=\sum_{M_1,M_2} e^{-i\vec{p}_\parallel\cdot(\vec{M}-\vec{L})}\times\langle\vec{M}\beta\beta_1\beta_2|\mathrm{V}|\vec{L}\alpha\alpha_1\alpha_2\rangle\cdot \tag{B2}$$

Here the matrix elements of $\mathcal{V}(\vec{p}_\parallel)$ are independent of L_1 and L_2 since those of $\mathcal{V}$ itself depend only on the differences M_1-L_1 and M_2-L_2; it was to secure this property that our notation was modified to presuppose only the lower symmetry. The resolvent is also reduced to the $\vec{p}_\parallel$ subspace in the usual fashion, and the expression supplementing Eq. (47) is

$$R_{n;\beta\beta_1\beta_2,\alpha\alpha_1\alpha_2}(E;\vec{p}_\parallel)=\exp\left[i\left(\frac{p_1}{\mu_1}(\beta_1-\alpha_1)+\frac{p_2}{\mu_2}(\beta_2-\alpha_2)\right)\right]\mu_1^{-1}\sum_{\nu_1=0}^{\mu_1-1}\exp\left(i\frac{2\pi\nu_1}{\mu_1}(\beta_1-\alpha_1)\right)\times\mu_2^{-1}\sum_{\nu_2=0}^{\mu_2-1}\exp\left(i\frac{2\pi\nu_2}{\mu_2}(\beta_2-\alpha_2)\right)R_{n;\beta\alpha}(E;\vec{k}_\parallel)\quad, \tag{B3}$$

where k_i stands for $\mu_i^{-1}(p_i+2\pi\nu_i)$ in the last factor. Note that no integration other than Eq. (47) need be performed to treat reconstructed surfaces. The reason is that $\mathcal{R}$ is determined by the properties of the bulk alone, which is unchanged. We have merely reexpressed $\mathcal{R}$ in a slightly different basis. This formalism will be applied in a future study.[8]

Baldock[20] has applied the Koster-Slater method to defects on a surface. The unperturbed Hamiltonian then represents the crystal with surface, and the perturbation is just the localized perturbation of the defect. Baldock's method requires that the eigenvalues and eigenfunctions of the unperturbed, bounded crystal be first obtained. He was able to do this for simple examples by inspection, but this is not possible in most cases of real physical interest. Fortunately, the techniques developed in Sec. II can be applied to the general treatment of this problem.

As usual, $\mathcal{H}_0$ represents the Hamiltonian of the infinite, unsevered crystal and $\mathcal{V}=\mathcal{V}_S+\mathcal{V}_T$ introduces the surface with its accompanying perturbation of the near surface orbitals. Now, however, the total Hamiltonian is $\mathcal{H}=\mathcal{H}_0+\mathcal{V}+\mathcal{V}_D$, where $\mathcal{V}_D$ is the localized defect potential. A bound electronic state localized near the defect has an energy lying completely outside both bulk and surface bands. The Koster-Slater determinantal equation determining the energy of such a localized state has the form

$$\det P_D[1-\mathcal{R}''(E)\mathcal{V}_D]=0\quad, \tag{B4}$$

where P_D is a projection operator such that $\mathcal{V}_DP_D=\mathcal{V}_D$, and $\mathcal{R}''(E)$ is the resolvent for $\mathcal{H}_0+\mathcal{V}$: $(E-\mathcal{H}_0-\mathcal{V})^{-1}$, as in Appendix A. The general matrix element of $\mathcal{R}''(E)$ can be expressed in terms of the reduced form having diagonal $\vec{k}_\parallel$, with which we have been primarily concerned till now,

$$\langle\vec{m}\beta|\mathcal{R}''(E)|\vec{l}\alpha\rangle=\int\frac{d^2\vec{k}_\parallel}{(2\pi)^2}e^{i\vec{k}_\parallel\cdot(\vec{m}-\vec{l})}\mathcal{R}''_{m\beta,l\alpha}(E;\vec{k}_\parallel)\quad. \tag{B5}$$

In analogy with Eq. (47), we here use the symmetry of $\mathcal{H}_0+\mathcal{V}$ under translations parallel to the surface to express its resolvent in terms of an integral of a more easily evaluated quantity. Now $\mathcal{R}''(E;\vec{k}_\parallel)$ can be found by the method used in Appendix A,

$$\mathcal{R}''(E;\vec{k}_\parallel)=[\mathcal{D}'(E;\vec{k}_\parallel)]^{-1}\mathcal{R}(E;\vec{k}_\parallel)\quad. \tag{B6}$$

The problem of inverting $\mathcal{D}'(E;\vec{k}_\parallel)=1-\mathcal{R}(E;\vec{k}_\parallel)\mathcal{V}(\vec{k}_\parallel)$ is no different in principle or in degree of difficulty from that of inverting $\mathcal{D}$. This latter task, of course, has been extensively discussed in Sec. II and Appendix A. If only the matrix elements of $\mathcal{R}''(E;\vec{k}_\parallel)$ within the surface layer projected on by P_1 are required, we can use the simpler formula [see Eq. (A24)]

$$P_1\mathcal{R}''(E;\vec{k}_\parallel)P_1=P_1[P_1\mathcal{D}'(E;\vec{k}_\parallel)P_1]^{-1}P_1\mathcal{R}(E;\vec{k}_\parallel)P_1\quad. \tag{B7}$$

This requires only the inversion of a finite-dimensional matrix. In practice, it may not be possible to do the integral in (B5) in closed form; this was true even for Baldock's relatively simple examples. Thus numerical integration will ordinarily be required.

The resolvent $\mathcal{R}''$ for the severed crystal with physically perturbed surface is also relevant to the question of the possible existence of virtual surface states or resonances lying within the energy region occupied by bulk states. In this regime, the eigenstates of $\mathcal{H}=\mathcal{H}_0+\mathcal{V}_S+\mathcal{V}_T$ are scattering states which can be labeled by the wave vector and band index of the incident plane wave. A scattering state of energy E will be said to represent a virtual surface state if an electron occupying such a state has an especially large relative probability of being found in the surface region, i. e., in one or more surface layer orbitals. More specifically we shall consider the quantity,

$$Q_{m\beta}(E;\vec{k}_{\parallel})=\sum_{\rho=1}^{\nu}\int_{-\pi}^{0}\frac{dk}{2\pi}\,|\Psi_{k\rho}(\vec{k}_{\parallel};m\beta)|^2\,\delta[E-\lambda_\rho(\vec{k})] \;. \tag{B7}$$

This represents precisely such a relative probability defined for the particular surface layer orbital $|\vec{k}_{\parallel};m\beta\rangle$. For convenience we have chosen to sum over all scattering states of energy E and have weighted the result by the density of states at E. In this equation, $\vec{k}$ has its usual decomposition: $\vec{k}=\vec{k}_{\parallel}+k\vec{c}^{\,*}$. Within the subspace of states of definite $\vec{k}_{\parallel}$, the eigenstate $\Psi_{k\rho}$ represents that scattering state whose incoming plane-wave part has the bulk quantum numbers $k\rho$ – k being negative for an incoming wave in the right half-crystal. The amplitude of this state on the layer orbital $|\vec{k}_{\parallel};m\beta\rangle$ is written $\Psi_{k\rho}(\vec{k}_{\parallel};m\beta)$. When $m\beta$ is appropriately chosen, a virtual surface state will be represented by a peak in $Q_{m\beta}(E;\vec{k}_{\parallel})$ centered on the virtual state's energy and having a width related to the reciprocal of the lifetime of the virtual state in the usual fashion.

The quantity $Q_{m\beta}$ can be expressed as a matrix element of $\mathcal{R}''$:

$$Q_{m\beta}(E;\vec{k}_{\parallel})=-(1/\pi)\,\mathrm{Im}(\mathcal{R}''_{m\beta,m\beta}(E+i\epsilon;\vec{k}_{\parallel}) \;, \tag{B8}$$

where ϵ is a positive infinitesimal. Using the method described above, the matrix elements of $\mathcal{R}''$ may be expressed in terms of those of $\mathcal{R}$, the resolvent of the bulk crystal. These latter may be presumed known outside the bulk energy bands and may be found within them, as presently required, by analytic continuation. In particular, the integral performed in Sec. II F gave us $\mathcal{R}$ outside the bands as a function of E and of certain functions of E, the decaying exponentials $\mu_j(E)=e^{iq_j(E)}$. Outside the bands, the q_j's are complex with positive imaginary parts. At the band edge, at least one of these imaginary parts goes to zero and that $q_j(E)$ has a branch point there. The others may have branch points there as well or elsewhere within the band. In any case, the positive infinitesimal ϵ in Eq. (B8) specifies how these branch points are to be circumvented. For example, suppose we want to continue from a band gap to a higher-lying conduction band, and suppose $\mathrm{Im}q_j(E)\to 0$ as $E\to E_0$ from below. Then typically, for $E<E_0$, $q_j(E)\approx q_{j0}+$ positive imaginary part, and for $E>E_0$, $q_j(E)\approx q_{j0}+$ positive real part. Here q_{j0} is the real wave vector $q_j(E_0)$ and the added terms go to zero at E_0 as the square root of $|E-E_0|$. In any case, if the expressions $q_j(E)$ are known explicitly, the analytic continuation can be easily carried out, and it will not be necessary to perform anew integrals like those in Sec. IIF.

---

[1] J. D. Levine and S. Freeman, preceding paper, Phys. Rev. B 2, 3255 (1970).

[2] J. D. Levine and S. G. Davison, Phys. Rev. 174, 911 (1968); this approach was used earlier in another context by J. C. Slater and G. F. Koster, Phys. Rev. 94, 1498 (1959).

[3] A comparative survey of a number of exact methods has been given by J. D. Levine and P. Mark, Phys. Rev. 182, 926 (1969).

[4] The matrix elements of the MO model Hamiltonian are determined so as to fit a portion of the observed band structure, including its symmetry properties, as closely as possible, rather than as the actual matrix elements of a real one-electron Hamiltonian in a tight-binding basis. One would thus expect that the MO representation of the bulk band structure could be used in the study of surface and defect electronic states even when an *ab initio* tight-binding derivation of the band structure was not feasible. Nonetheless, decisions as to how many local basis states $|\vec{m}\beta\rangle$ are to be included, which particular transfer integrals are to be nonzero, and how the effect of the surface is to be represented are undoubtedly made on the basis of a tight-binding analogy. How to go beyond this analogy, and what may then be the nature of the one-electron states to which the local basis elements correspond are questions which seem not to have been answered. For such reasons as these, it may be preferable for accurate calculations to use a model Hamiltonian based on a nonorthogonal set of local basis states. Such an approach can represent the bulk band structure economically and extremely accurately. [See A. H. Kellner, Acta Phys. Austrica 18, 48 (1964).] The basis states in this case should be well localized, so that the effects of the surface could be relatively easily introduced. The generalization of the formalism presented here to nonorthogonal basis states is straightforward. It is outlined at the very end of Sec. II.

[5] This point has been emphasized by Levine and Mark in the paper cited in Ref. 3.

[6] The origin and history of these terms is discussed in S. G. Davison and J. D. Levine, *Surface States in Solid State Physics*, edited by F. Seitz, D. Turnbull, and H. Ehrenreich (Academic, New York, 1970), Vol. 25, p. 1.

[7] J. Koutecký, Phys. Rev. 108, 13 (1957).

[8] E. Williams and S. Freeman (unpublished).

[9] Otherwise one would simply redefine S_0 and S_1 as somewhat larger sets for which this property held.

[10] The reduced Hamiltonian $\mathcal{H}(\vec{k}_\parallel)$ differs from a true one-dimensional Hamiltonian in its behavior under time reversal: $\mathcal{H}(\vec{k}_\parallel)^* = \mathcal{H}(-\vec{k}_\parallel)$. Here the matrix elements of the operator on the left are complex conjugated without transposition.

[11] It is necessary at this point to mention two minor modifications of the formalism which are almost invariably made in practice. Indeed, we shall employ them in Sec. III. First, we have described the location of a unit cell by a vector, $\vec{m} = m_1\vec{a} + m_2\vec{b} + m_3\vec{c}$, which represents the displacement required to bring the cell containing the origin into coincidence with the cell labeled by $\vec{m}$. The origin is fixed at some convenient reference point within the unit cell. We now imagine that the βth orbital is associated with an ion whose location relative to the reference point is $\vec{u}_\beta$. It produces somewhat more symmetrical expressions if the transformation from real to $\vec{k}$ space is based on $\exp[i\vec{k}\cdot(\vec{m}+\vec{u}_\beta)]$ rather than on $\exp(i\vec{k}\cdot\vec{m})$ itself. This amounts to changing the phase of the plane-wave basis vector $|\vec{k}\beta\rangle$ by a factor $\exp(i\vec{k}\cdot\vec{u}_\beta)$. [This basis vector is explicitly defined in Eq. (39).] The various formulas presented in the text can be adapted to this new convention by uniformly replacing $\vec{l}$, $\vec{m}$ and $\vec{n}$ when they appear in exponential expressions by $\vec{l}+\vec{u}_\alpha$ $\vec{m}+\vec{u}_\beta$ and $\vec{n}+\vec{u}_\beta-\vec{u}_\alpha$, respectively. Second, we have chosen to describe a wave vector $\vec{k}$ in terms of its components along the reciprocal-lattice vectors $\vec{a}^*$, $\vec{b}^*$, and $\vec{c}^*$. Then, when the surface is introduced the components along $\vec{a}^*$ and $\vec{b}^*$ are conserved, and each surface state is characterized by a value of $\vec{k}_\parallel = k_1\vec{a}^* + k_2\vec{b}^*$. If we had *not* started from the bulk but simply noted that the crystal with surface has a translation group based on primitive translations $\vec{a}$ and $\vec{b}$, we would naturally have constructed a two-dimensional reciprocal lattice based on $\vec{a}^{*\prime}$ and $\vec{b}^{*\prime}$, which are dual to $\vec{a}$ and $\vec{b}$ within the $\vec{a}$, $\vec{b}$ plane. Recall that $\vec{c}^*$ is perpendicular to $\vec{a}$ and $\vec{b}$; clearly, the vectors $\vec{a}^{*\prime}$ and $\vec{b}^{*\prime}$ differ from $\vec{a}^*$ and $\vec{b}^*$ simply by multiples of $\vec{c}^*$. We may write $\vec{k} = k_1\vec{a}^* + k_2\vec{b}^* + k_3\vec{c}^* = k_1'\vec{a}^{*\prime} + k_2'\vec{b}^{*\prime} + k_3'\vec{c}^*$, where $k_1 = k_1'$, $k_2 = k_2'$, and $k_3 \neq k_3'$. The linear relation between k_3' and k_1, k_2, and k_3 is easily found by, say, taking the scalar product of both expressions for $\vec{k}$ with $\vec{c}$. The primitive cell of the bulk reciprocal lattice can be taken to be $-\pi < k_i' \leq \pi$ just as well as $-\pi < k_i \leq \pi$. This amounts to replacing a parallelepiped with edges along $\vec{a}^*$, $\vec{b}^*$, and $\vec{c}^*$ by a *right* parallelepiped of equal volume with edges along $\vec{a}^{*\prime}$, $\vec{b}^{*\prime}$, and $\vec{c}^*$. Note that when expressed in terms of these new coordinates, the bulk bands $\lambda_\rho(\vec{k})$ are not periodic in k_1' and k_2' with period 2π except when k_3' is simultaneously displaced in an appropriate fashion. However, the surface-state bands *are* so periodic, as is the bulk density of states. When $\vec{k}_\parallel$ is redefined as $k_1'\vec{a}^{*\prime} + k_2'\vec{b}^{*\prime}$ only the phases of the layer orbitals, $|\vec{k}_\parallel, m\beta\rangle$, are affected. In all the formulas of Sec. II, either convention can be used so long as one is consistent.

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

[13] J. Koutecký, in *Advances in Chemical Physics*, edited by I. Prigogine (Interscience, New York, 1965), Vol. 9.

[14] For formal and pedagogic purposes, it is convenient to think of S_0 and S_1 as including one or more complete unit cells. This can always be arranged by increasing the dimension of one or both sets, though to do so may not be the most economical approach. If unnecessary basis elements have been included in S through such a procedure, the corresponding columns of the determinant (18) will have only zeros except for 1 on the diagonal.

[15] This follows from the Hermiticity of $\overleftrightarrow{\mathcal{R}}(E) = (E - \mathcal{H}_0)^{-1}$ and $\overleftrightarrow{\mathcal{V}}$, together with the lemma: $\det(\overleftrightarrow{1} - \overleftrightarrow{A}\cdot\overleftrightarrow{B}) = \det(\overleftrightarrow{1} - \overleftrightarrow{B}\cdot\overleftrightarrow{A})$. This holds for finite-dimensional determinants even if neither $\overleftrightarrow{A}$ nor $\overleftrightarrow{B}$ has an inverse. It is easily proved by replacing $\overleftrightarrow{A}$ by $\overleftrightarrow{A}(\lambda) = \overleftrightarrow{A} + \lambda\overleftrightarrow{1}$. The lemma must hold for all λ for which $\overleftrightarrow{A}(\lambda)$ has an inverse. Since this excepts only isolated values of λ, the equation $\det[1 - \overleftrightarrow{A}(\lambda)\cdot\overleftrightarrow{B}] = \det[\overleftrightarrow{1} - \overleftrightarrow{B}\cdot\overleftrightarrow{A}(\lambda)]$ must be an identity in λ. We can equate the coefficients of the various powers of λ on the two sides. The zeroth-order terms give the required relation.

[16] Let L be a projection operator for the left half-crystal, then $[L, \mathcal{H}] = 0$ expresses the fact that $\mathcal{H}$ does not connect the two half-crystals and implies that L and $\mathcal{H}$ can be simultaneously diagonalized. When this is done, an eigenstate Ψ of $\mathcal{H}$ satisfies either $L\Psi = 0$, in which case Ψ refers to the right half-crystal, or $L\Psi = \Psi$, in which case it refers to the left.

[17] J. Koutecký and S. G. Davison, Inter. J. Quantum Chem. 2, 73 (1968).

[18] J. Koutecký and M. Tomášek, Phys. Rev. 120, 1212 (1965).

[19] It is worth noting that, in a number of interesting cases, $\det P_0[1 - \mathcal{R}(E)\mathcal{V}_S]$ and $\det P_1[1 - \mathcal{R}(E)\mathcal{V}_S]$ are real and consequently identical. These cases include all truly one-dimensional problems, as well as those three-dimensional cases for which the factor group contains a twofold rotation, proper or improper, normal to the surface crystallographic plane.

[20] G. R. Baldock, Proc. Cambridge Phil. Soc. 48, 457 (1952). This author used $\mathcal{R}'(E)$ to study bound states associated with *defects* in a surface rather than surface states *per se*. In this sense, his method is generalized in Appendix B.

[21] Recalling the redefinition of the phases of the plane-wave basis states discussed in Ref. 11, it is worth pointing out that as far as the $\vec{c}$ components of the vectors $\vec{u}_\beta$ are concerned, either convention gives the identical integrand in Eq. (47). Complex-conjugate phase factors are assigned to the two factors in the integrand and thus cancel out. The same observation applies to the remaining formulas in this subsection.

[22] W. Kohn, Phys. Rev. 115, 809 (1959).

[23] We use $\overleftrightarrow{R}_n$ to represent the $\nu\times\nu$ matrix with matrix elements $R_{n;\,\beta\alpha}$.

[24] For the labeling of the symmetry points of the two-dimensional zone, see R. O. Jones, Phys. Rev. Letters 20, 992 (1968) or Ref. 1.

[25] This same model having only ϵ_0 nonzero is used in Ref. 1 to study the properties of resonant scattering states within the continuum. An outline of how this is done, in general, is given at the end of Appendix B.

[26] This calculation using the ansatz method was carried out by J. D. Levine, see Ref. 1.

[27] Since we have not yet proved Eq. (23) of the text at this point in the argument we are not yet able to conclude that these are the *only* additional zeros. Note that Shockley state here stands for *any* eigenstate of $\mathcal{H}_0 + \mathcal{V}_S$ since they are all discrete.

[28] E. T. Whittaker and G. N. Watson, *A Course of Modern Analysis*, 4th ed. (Cambridge U. P., Cambridge, England, 1958), p. 105.

[29] The identity of the eigenvalues and eigenfunctions obtained for the right half-crystal in the two cases can easily be proved in perturbation theory. Consider as unperturbed Hamiltonian $\mathcal{H}_0 + \mathcal{V}_S + \mathcal{V}_B$. The presence of $\mathcal{V}_B$ has no effect on the solutions for the right half-crystal, which thus represent the eigenstates of $\mathcal{H}_0 + \mathcal{V}_S$. Now impose the perturbation $-\mathcal{V}_S$, which converts the Hamiltonian to $\mathcal{H}_0 + \mathcal{V}_B$. Represent by γ the typical magnitude of one of the transfer integrals in $\mathcal{V}_S$. Then according to perturbation theory, the lowest-order modifications of energy eigenvalues and wave-function amplitudes are of order $\epsilon\gamma^2$ and $\epsilon\gamma$, respectively. Obviously these go to zero with ϵ. This argument can be made rigorous without much difficulty.

PHYSICAL REVIEW B VOLUME 2, NUMBER 8 15 OCTOBER 1970

Extension of a Linear Diatomic-Chain Model for the Calculation of Local-Mode Frequencies in Real Crystals

G. Lucovsky
Xerox Research Laboratories, Rochester, New York 14603
and
M. H. Brodsky
IBM Thomas J. Watson Research Center, Yorktown Heights, New York 10598
and
E. Burstein*
University of Pennsylvania, Philadelphia, Pennsylvania 19104

(Received 20 April 1970)

Calculations by Mazur, Montroll, and Potts (MMP) have shown that local modes above the optical branch of the host crystal are predicted by a linear diatomic-chain model for all positive values of the mass-defect parameter ϵ. Three-dimensional calculations show that local modes exist only for values of ϵ greater than some critical value. However, these three-dimensional calculations require a knowledge of the eigenvalues and eigenvectors of all the phonon states of the host lattice. We show that the simpler MMP model can be applied to three-dimensional crystals by inclusion of the LO-phonon frequency. In a given system of host crystal and impurity, the determinant parameters are the mass defect of the impurity and the width of the host-crystal reststrahlen band. Calculations on approximately 20 solid solution systems of the form $AB_{1-x}C_x$ have successfully predicted the existence or nonexistence of a local mode when x is large and the mass of B is less than the mass of C. The modified one-dimensional model gives quantitative results for local-mode frequencies which agree with full three-dimensional calculations for local modes in NaI, CdS, and Si.

I. INTRODUCTION

It is well known that small concentrations of substitutional impurities in crystalline lattices can result in local, resonance, or gap vibrational modes.[1] Experimentally these modes are observable by impurity-induced infrared absorption[1] or Raman scattering.[2] Recently, it has been pointed out that certain aspects of the behavior of the long-wavelength optical phonons in pseudobinary mixed crystals could be predicted from the nature of the impurity modes at the low composition limits of the alloy system.[3-6] This paper reports a one-dimensional model calculation for the frequencies of local modes. Our model gives results in good quantitative agreement with the more complicated full three-dimensional calculations, and also reasonably accounts for the experimental observations in many cases where extensive three-dimensional calculations are not available. As developed below, the model is applied to the simpler diatomic crystals. Extension to other systems is straightforward but not considered in detail here.

Calculations of local-mode frequencies[7] (as well as gap- and resonance-mode frequencies) require a knowledge of the eigenfrequencies and eigenvectors of the host-crystal vibrational modes as well as the mass and force-constant changes introduced by the impurity atom. Such calculations were first performed[7] for substitutional impurities in Si, using a mass-defect formalism in which force-constant changes were neglected. The local-mode frequencies were calculated as a function of the mass-defect parameter ϵ_j, where $\epsilon_j = 1 - M/M_j$. M is the mass of the impurity atom and M_j the mass of the host-crystal atom. These calculations were later extended to include changes in the effective force constant at the impurity site.[8] For the polar diatomic crystals, three-dimensional mass-defect calculations have been performed for a limited number