PHYSICAL REVIEW B VOLUME 14, NUMBER 10 15 NOVEMBER 1976

Localized defects in III-V semiconductors

M. Jaros* and S. Brand
Department of Theoretical Physics, School of Physics, The University, Newcastle Upon Tyne, United Kingdom

(Received 9 March 1976)

A Green's-function approach, closely related to that of Koster and Slater, is developed, within the one-electron pseudopotential formalism. This is applied to single vacancies, the divacancy and vacancy-oxygen pairs in GaAs, and single vacancies in GaP and InSb. A number of localized states are found. It is shown that the properties of the host-crystal band structure, the electron-electron interaction, and lattice relaxation play an important part in the formation of these states. In contrast with the Koster-Slater calculations, our results are *not* very sensitive to the strength of the vacancy potential. Our calculations indicate that As and Ga vacancies in GaAs should behave as single donors and acceptors, respectively. The V_{Ga}-O complex behaves as a deep neutral center. The wave functions associated with the levels introduced into the band gap are highly localized. Since they contain admixtures of both valence- and conduction-band functions, and exhibit large lattice distortion, they are likely to take part in nonradiative recombination processes. Finally, we note that nonradiative recombination via some deep neutral centers is likely to be very efficient since our calculation also indicates a possibility of localized excited states introduced by these centers.

I. INTRODUCTION

The III-V compound semiconductors have been the object of a great deal of applied and pure research over the last decade. This is because of their importance in the technology of electron-transfer devices, light-emitting diodes, etc., and also because of their interesting fundamental properties.

In most applications one must take account of the native defects and chemical impurities. The bound states in the gap associated with some common dopants—the shallow donors and acceptors—have been well understood for some years and are amenable to theoretical analysis.[1] This is because of the simple long-range electrostatic potentials they are associated with. However, there is much evidence indicating a number of localized states lying further within the band gap.[2] The position of these states in the gap often depends on the method of crystal preparation and their chemical nature is in many cases difficult to establish. Nor are these defects and their effects on the local electronic properties of the solid well understood from a theoretical standpoint. In particular, most of the existing studies on the properties of vacancies (and their complexes) in semiconductors have been confined to the group-IV materials, namely, diamond and silicon.[3] Probably the only exceptions are the works of Parada,[4] Pratt,[5] and more recently Hemstreet,[6] on single vacancies in IV-VI compound materials. The main reason why so little theoretical work on these problems has been done is that there is no generally suitable formalism available for dealing with such systems.

The existing calculations on vacancies may be divided into two groups: (i) Koster-Slater-type and (ii) cluster calculations. The former[7] is rooted in a formalism exploiting our knowledge of the electronic bands of the perfect material in $\vec{k}$ space. Its advantage is that the position of the localized defect states with respect to the band edges of the host crystal is well defined. However, a great deal of both numerical and algebraic work is required in using this method. In addition, the calculations of Callaway and Hughes[8] exhibit an alarming sensitivity to the strength of the vacancy potential. Since a good vacancy potential cannot be easily obtained, the final error in the result must be dangerously large. It is, therefore, not surprising that a cluster calculation seemed a natural alternative. A number of schemes have emerged[9] with the same basic idea in common. They focus on a small part of the solid surrounding the defect. The rest of the solid enters the problem only via the boundary conditions on the cluster wave functions. The emphasis on local properties of the defect seems to indicate that the cluster methods should offer a faster and less complicated calculational technique. In particular, it promises a self-consistent result within its own limitations. However, the most recent studies of Messmer and Watkins, Hemstreet, and others indicate a number of problems. For example, the size of the cluster, the choice of the surface and boundary conditions, may significantly affect the very existence of any localized states in the gap. Also, it is not always obvious that the cluster calculation for the "perfect" solid does provide a reliable starting point for the calculation with the "imperfect" cluster. Such a situation may become particularly serious in the case of many-valley direct-gap materials like GaAs or ZnSe. The conduction bands of semiconductors derived from a tight-binding-like scheme are known to be in error.

Recently, we have developed a method of calculation which has something in common with both the Koster-Slater and cluster schemes. It has the advantage of the first in that it allows one to define precisely the position of the impurity (defect) level with respect to the true band edges of the host (unperturbed) crystal. However, it avoids the complexities and ambiguities which arise in working with the Wannier functions. It is really a Green's-function[10] approach (closely related to that of Koster and Slater), in which the role of the Wannier functions is taken over by a more suitable complete set (s) of functions localized at a center (s) of interest. The band structure of the perfect crystal and atomic self-consistent pseudopotential are required as input. The method has been applied[11] to calculate localized (one- and two-electron) states associated with deep donor oxygen in GaP and GaAs. The results of these calculations lead us to believe that some useful predictions concerning vacancies and their complexes in III-V semiconductors can be made.[12] In the absence of any theoretical prediction concerning these defects, such an attempt seems highly desirable. Furthermore, it may allow us to compare the nature of chemical impurities with that of the native defects since the input data representing the effect of the host-crystal band structure is exactly the same for both. The fact that we have been able to apply our method to a variety of defects *and* chemical impurities should not be overlooked. Defect wave functions can be generated either in terms of a table of coefficients of the unperturbed Bloch pseudofunctions or in terms of a convenient trial set of (localized) functions. Our calculations may well be used to help in distinguishing spectra associated with various deep levels in semiconductors like GaAs. Our calculations can be made self-consistent and allow for lattice-distortion effects. However, a truly quantitative description of these effects has not been achieved.

In Sec. II we give a detailed description of our method of calculation and its convergence properties. In Sec. III A, we present results of our calculation of localized states associated with vacancies in an unrelaxed lattice. The vacancy potential is represented by an empirical pseudopotential. The results are reported for single vacancies in GaP, GaAs, and InSb, and the divacancy and V_{Ga} -O complex in GaAs. In Sec. III B we investigate the electron-electron interaction and the effect of Jahn-Teller distortion on the above states. On the basis of these results we suggest in Sec. IV a prediction which is amenable to experimental verification. The prediction for defects in GaAs is summarized in Fig. 12.

In Sec. IV we also comment on some general characteristic properties of deep states. In particular, we show that some neutral centers may introduce more than one deep state in the forbidden gap. This is likely to enhance the probability of nonradiative recombination via such centers.

II. DESCRIPTION OF THE METHOD OF CALCULATION

In order to find the energy levels produced by the introduction of an impurity potential h into a pure-crystal environment we begin with the Schrödinger equation:

$$(H_0+h)\psi=\epsilon\psi. \tag{1}$$

The pure-crystal problem in which $H_0\varphi_{\vec{k},n}=E_{\vec{k},n}\varphi_{\vec{k},n}$ where $\varphi_{\vec{k},n}$ and $E_{\vec{k},n}$ are, respectively, the Bloch functions and associated energies, is assumed to be already solved. The impurity pseudopotential h will be initially expressed as a general impurity complex such that

$$h(\vec{r})=\sum_i h_i(|\vec{r}-\vec{R}_i|), \tag{2}$$

where $\vec{R}_i$ indicates the position of a potential h_i at a site i, with respect to the band-structure origin. An expansion in terms of the host-crystal Bloch functions is used to represent the impurity wave function, i.e.,

$$\psi(\vec{r})=\sum_n \int_{\text{BZ}} d\vec{k}\, A_{\vec{k},n}\varphi_{\vec{k},n}(\vec{r}), \tag{3}$$

where we have used the reduced-zone notation. [In pseudopotential representation, $\varphi_{\vec{k},n}(\vec{r})=\sum_{\vec{g}} a_{\vec{g}} e^{j(\vec{k}+\vec{g})\cdot\vec{r}}$, where $j=\sqrt{-1}$ and where $\vec{g}$ stands for a reciprocal-lattice vector.] Inserting the above expansion for ψ into the Schrödinger equation leads to the consistency relations

$$\begin{aligned} A_{\vec{k}',n'}&(E_{\vec{k}',n'}-\epsilon) \\ &+\sum_i\sum_n\int_{\text{BZ}} d\vec{k}\,A_{\vec{k},n} \\ &\qquad\times\int d\vec{r}\varphi^*_{\vec{k}',n'}(\vec{r})h_i(\vec{r})\varphi_{\vec{k},n}(\vec{r})=0\,. \end{aligned} \tag{4}$$

If a finite-sampling-point technique is used then the equations can be solved as they stand, but we prefer to follow a method outlined by Bassani *et al.*[10] Therefore we proceed as follows. Consider the above real space integral, $\int d\vec{r}\,\varphi_{\vec{k}',n'}(\vec{r})h_i(\vec{r})\varphi_{\vec{k},n}(\vec{r})$. This can be expressed as

$$\int d(\vec{r}-\vec{R}_i)\varphi^{i*}_{\vec{k}',n'}(\vec{r}-\vec{R}_i)h_i(|\vec{r}-\vec{R}_i|)\varphi^{i}_{\vec{k},n}(\vec{r}-\vec{R}_i)$$

$$=\int d\vec{r}'\,\varphi^{i*}_{\vec{k}',n'}(\vec{r}')h^a_i(\vec{r}')h^b_i(\vec{r}')\varphi^{i}_{\vec{k},n}(\vec{r}'). \quad (5)$$

The coordinate change modifies the appearance of the Bloch functions and gives

$$\varphi^{i}_{\vec{k},n}(\vec{r}')=\sum_{\vec{g}}a_{\vec{g}}\,e^{j(\vec{k}+\vec{g})\cdot\vec{r}'}e^{j(\vec{k}+\vec{g})\cdot\vec{R}_i} \quad (6)$$

with respect to the new origin at $\vec{R}_i$. The spherically symmetric pseudopotential h_i has been expressed as a product of the radial functions h^a_i, h^b_i as this form proves to be useful. Making use of the complete orthonormal set of functions $g_{m,i}$ the identity

$$1=\int d\vec{r}''\,\delta(\vec{r}'-\vec{r}'')=\sum_m\int d\vec{r}''\,g^*_{m,i}(\vec{r}'')g_{m,i}(\vec{r}') \quad (7)$$

is inserted. This results in a separation of the integral into a product of integrals containing either $\vec{k}$ or $\vec{k}'$ alone. The integral (5) now becomes

$$\sum_m\int d\vec{r}'\varphi^{i*}_{\vec{k}',n'}(\vec{r}')h^a_i(\vec{r}')g_{m,i}(\vec{r}')$$

$$\times\int d\vec{r}''\,g^*_{m,i}(\vec{r}'')h^b_i(\vec{r}'')\varphi^{i}_{\vec{k},n}(\vec{r}'')$$

$$=\sum_m f^a_{m,i}(\vec{k}',n')f^b_{m,i}(\vec{k},n). \quad (8)$$

Introducing new coefficients,

$$a_{m,i}=\sum_n\int_{\text{BZ}}d\vec{k}A_{\vec{k},n}f^b_{m,i}(\vec{k},n), \quad (9)$$

the consistency relations take the form

$$A_{\vec{k}',n'}+\sum_i\sum_m\frac{a_{m,i}f^a_{m,i}(\vec{k}',n')}{E_{\vec{k}',n'}-\epsilon}=0. \quad (10)$$

The transformation to a formulation in terms of the $a_{m,i}$ is achieved by multiplying by $f^b_{m',i'}(\vec{k}',n')$ and integrating over all $\vec{k}$ space to give

$$a_{m',i'}+\sum_i\sum_m\sum_{n'}a_{m,i}\int_{\text{BZ}}d\vec{k}'\,\frac{f^a_{m,i}(\vec{k}',n')f^b_{m',i'}(\vec{k}',n')}{E_{\vec{k}',n'}-\epsilon}=0. \quad (11)$$

We find solutions by looking for ϵ such that

$$\left|\delta_{m',m}\delta_{i',i}+\sum_{n'}\int_{\text{BZ}}d\vec{k}'\,\frac{f^a_{m,i}(\vec{k}',n')f^b_{m',i'}(\vec{k}',n')}{E_{\vec{k}',n'}-\epsilon}\right|=0. \quad (12)$$

Our main interest lies with values of ϵ in the band gap, i.e., when $E_{\vec{k}',n'}-\epsilon\neq 0$. However, we also use this method without modification to locate values of ϵ just inside the band edges.

The purpose of introducing the $a_{m,i}$ rather than persevering with the $A_{\vec{k},n}$ is to reduce the number of coefficients and thus make the calculation more tractable. This can be done by making a suitable choice for the g_m.[11] We now reveal the actual form of the functions g_m. The familiar spherical harmonics $Y_{l,m}(\theta,\phi)$ are an appropriate choice for the angular part. The radial part of the functions, which we denote by $G^l_n(r)$, must satisfy

$$\int dr\,r^2G^l_n(r)G^l_{n'}(r)=\delta_{n,n'},\quad n,n'=0,1,2.\ldots \quad (13)$$

This can be achieved by choosing

$$G^l_n(r)=\left[\Gamma(l+1)\binom{n}{n-l}\right]^{-1/2}\frac{e^{-r/2}}{r}r^{l/2}L^l_{n-l}(r), \quad (14)$$

where the L's are the associated Laguerre polynomials. We denote the number of functions g_m per site by K. If, for example, we use ten s and nine p functions, then $K=10$. In our calculations we only deal with g_m containing the spherical harmonics Y_0, Y_1.

In calculating the real space integrals we expand the components, $e^{j(\vec{k}+\vec{g})\cdot\vec{r}}$ of the Bloch functions as products of spherical Bessel functions and spherical harmonics, i.e.,

$$e^{j(\vec{k}+\vec{g})\cdot\vec{r}}=4\pi\sum_{l,m}j^l B_l(|\vec{k}+\vec{g}|r)$$

$$\times Y^*_{l,m}(\theta_{\vec{k}+\vec{g}},\phi_{\vec{k}+\vec{g}})Y_{l,m}(\theta,\phi). \quad (15)$$

Consider now the impurity pseudopotentials $h_i(r)$. These are found by fitting a real space radial function to give tabulated pseudopotential[13] coefficients. The factorizable form $h_i(r)=h^a_ih^b_i=(c^2r^2-d^2)r^{2n}e^{-\alpha r}$ is fairly flexible for fitting data and has been used for the vacancy calculations we have carried out. Figure 1 shows the potentials we used for V_{Ga} and oxygen in GaAs.

We now direct our attention towards the calculation of the elements of the determinant

$$\sum_{n'}\int_{\text{BZ}}d\vec{k}'\,\frac{f^a_{m,i}(\vec{k}',n')f^b_{m',i'}(\vec{k}',n')}{E_{\vec{k}',n'}-\epsilon}. \quad (16)$$

This has been performed for single vacancies in GaAs and other materials and for the divacancy and vacancy-oxygen complexes also in GaAs. An outline will be given here for the calculation in the case of the general double impurity in such a material. The $\vec{R}_i$ are situated at adjacent pair positions, that is at $\pm(\frac{1}{8}a)(1,1,1)$, a being the lattice constant. The band structure is centered about the midpoint of the pair and for the purposes of the calculation we use the (1, 1, 1) direction for the z axis. Beginning with Bloch functions in the irreducible $\frac{1}{48}$ segment of the Brillouin zone we cover the entire $\vec{k}$ space by applying the 48 symmetry operations, denoted by α to give

$$\sum_{n'} \int_{1/48} \frac{d\vec{k}'}{E_{\vec{k}',n'} - \epsilon} \left(\sum_{\alpha} f^{a}_{m,i}(\alpha\vec{k}', n') f^{b}_{m',i'}(\alpha\vec{k}', n') \right). \tag{17}$$

Expanded in full, the bracketed term is

$$\sum_{\alpha} \left(4\pi(-j^{l}) \sum_{\vec{g}} a_{\vec{g}} e^{-j\alpha\vec{q}\cdot\vec{R}_i} Y_{l,m}(\theta_{\alpha\vec{q}}, \phi_{\alpha\vec{q}}) \int dr\, G^{l}_{n}(r) B_{l}(qr) r^2 h^{a}_{i}(r) \right)$$
$$\times \left(4\pi j^{l'} \sum_{\vec{g}'} a_{\vec{g}'} e^{j\alpha\vec{q}'\cdot\vec{R}_{i'}} Y^{*}_{l',m'}(\theta_{\alpha\vec{q}'}, \phi_{\alpha\vec{q}'}) \int dr'\, G^{l'}_{n'}(r') B_{l'}(q'r') r'^2 h^{b}_{i'}(r') \right), \tag{18}$$

where $\vec{q} = \vec{k} + \vec{g}$ and $\theta_{\alpha\vec{q}}$, $\phi_{\alpha\vec{q}}$ are the angular coordinates of $\alpha\vec{q}$.

In considering the operation-dependent part of this product we begin by looking at the separate terms of form $e^{-j\alpha\vec{q}\cdot\vec{R}_i} Y_{l,m}(\theta_{\alpha\vec{q}}, \phi_{\alpha\vec{q}})$. As $\vec{R}_i$ is oriented along the (1, 1, 1) direction, the z axis, the C_{3v} operations have no effect upon the complex exponential. Nor do they affect θ. This means that all products of Y_1, $Y_{1,0}$ spherical harmonics remain invariant under these six operations. Thus we need only perform the remaining eight operations, which will be denoted by β in the calculation of these products. Further scrutiny reveals that products of either Y_0 or $Y_{1,0}$ terms with $Y_{1,\pm 1}$ terms vanish when a summation is made over the C_{3v} operations γ. This also holds true between $Y_{1,1}$ and $Y_{1,-1}$ products which can be seen to lead to doubly degenerate energy states.

When the $Y_{1,\pm 1}$ terms are considered the f^a, f^b are certainly not invariant upon application of the γ. However, if the sum of the products of the terms is calculated then we obtain

$$\sum_{\gamma} e^{-j\gamma\vec{q}\cdot\vec{R}_i} Y_{l,m}(\theta_{\gamma\vec{q}}, \phi_{\gamma\vec{q}}) e^{j\gamma\vec{q}'\cdot\vec{R}_{i'}} Y^{*}_{l',m'}(\theta_{\alpha\vec{q}'}, \phi_{\alpha\vec{q}'}) = 6\,\mathrm{Re}[e^{-j\vec{q}\cdot\vec{R}_i} Y_{l,m}(\theta_{\vec{q}}, \phi_{\vec{q}}) e^{j\vec{q}'\cdot\vec{R}_{i'}} Y^{*}_{l',m'}(\theta_{\vec{q}}, \phi_{\vec{q}'})]. \tag{19}$$

Thus all the calculations for both singlet and doubly degenerate states can be performed by operating with only the eight remaining operations β. The determinant finally reduces to the form shown below, where columns are terms in f^a and rows are terms in f^b, and N means terms that are nonzero:

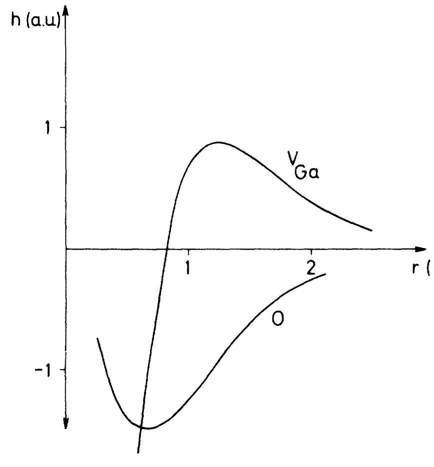

FIG. 1. V_{Ga} and substitutional oxygen (at arsenic site) potentials used in GaAs calculations. The real space potentials were found by fitting to pseudopotential coefficients.

$$\begin{array}{c|cccc|} & Y_0 & Y_{1,0} & Y_{1,1} & Y_{1,-1} \\ \hline Y_0 & N & N & 0 & 0 \\ Y_{1,0} & N & N & 0 & 0 \\ Y_{1,1} & 0 & 0 & N & 0 \\ Y_{1,-1} & 0 & 0 & 0 & N \end{array}.$$

We look for zeros for each of the blocks of the determinant with elements

$$\delta_{m',m}\delta_{i',i} + 6\,\mathrm{Re}\left(\sum_{n'} \int_{1/48} \frac{d\vec{k}'}{E_{\vec{k}',n'} - \epsilon} \times \sum_{\beta} f^{a}_{m,i}(\beta\vec{k}', n') f^{b}_{m',i'}(\beta\vec{k}', n') \right). \tag{20}$$

Each of the blocks in the determinant is built up from elements coming from two impurity centers and hence the possibility is allowed that we may find two zeros for each. There is then the possibility of finding two singlet states and two doubly degenerate states within the band gap. In our results we never find all of these states. For the divacancy and vacancy-oxygen pair only one singlet state is found, the other being pushed up into the conduction-band region in both cases. For the vacancy-oxygen complex one of the degenerate states is also lost.

For the single impurity the band structure is centered at the impurity site and the higher symmetry leads to a simpler calculation. In such cases the determinant partitions to give one singlet state together with a triply degenerate state.

We now consider the di-impurity results with regard to convergence with number of functions and the effect of scaling the potential. Radial functions G_n^l were used for values of K up to 10. This entails the calculation of a 38×38 matrix (19 functions per site, ten for $l=0$, nine for $l=1$) in the case of the singlet state for each of the trial energies ϵ used. In Fig. 2 the convergence of the result as a function of K value is illustrated. It can be seen that by about $K=5$ (an 18×18 submatrix) the position of the level is beginning to emerge. The position of the level is obtained to a degree sufficiently adequate for our purposes by $K=10$. At this stage the error in the potential is the factor which essentially sets the confidence we may put on the final result. It appears that the choice of functions is a good one. Figure 3 shows that the results are sensitive to scaling (i.e., change in the potential achieved by multiplying the potential by a constant P) but not unduly so. It is essential that this is so, otherwise the error in the potential would be disastrous.

The convergence properties of the degenerate states, shown in Fig. 4, are better but of a similar nature to those for the singlet state. A particularly interesting feature of these levels is their response to scaling (Fig. 5). The remarkable insensitivity of these is consistent with the properties of the degenerate states for the single vacancy and seems to be a general result. This indicates that the error in the potential may not be so crucial in dealing with states of this type. For the single vacancy the convergence of the results with respect to the number of functions used is similar to that found for GaP:O.[11] Previous calculations[11] have indicated what constitutes an adequate sampling procedure and this has been adopted throughout. We use a grid of 41 points in the $\frac{1}{48}$ segment and sum over four valence and six conduction bands. Each Bloch function is formed from an expansion containing 65 reciprocal-lattice vectors.

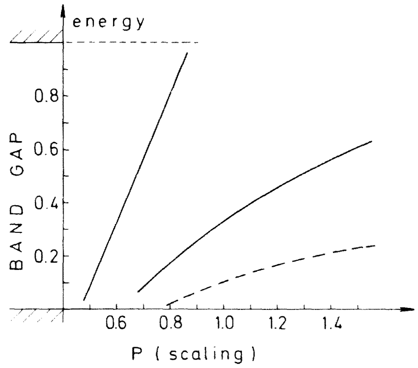

FIG. 3. Effect of scaling upon the singlet-state results for the di-impurity in GaAs. Equal scaling is applied to the potentials on both sites. The solid curves are for the divacancy and the dashed curve for the V_{Ga}-O complex.

Typically, the calculation for 10 trial energies in the band gap for the 38×38 singlet matrix requires about 1000 sec of central-processing-unit time on an IBM 370 computer.

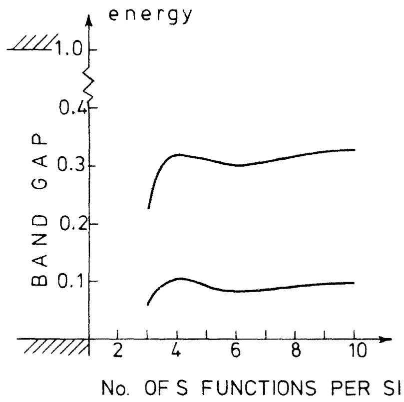

FIG. 2. Convergence of the singlet state results for di-impurities in GaAs. The K value is equal to the number of s functions on each site ($K=10$ implies a total of 19 functions per site). The upper curve is for the divacancy, the lower is V_{Ga}-O.

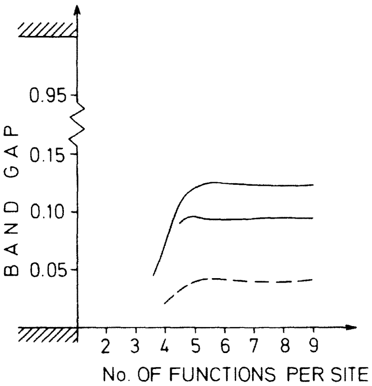

FIG. 4. Convergence properties of the degenerate states. In the dashed curve the effects upon the V_{Ga}-O complex in GaAs can be seen. The other two curves are for the divacancy.

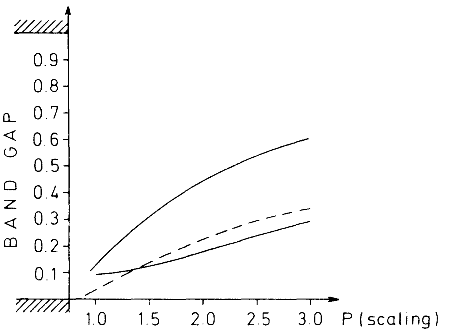

FIG. 5. Effect of scaling upon the di-impurity degenerate energy levels in GaAs. Labeling is the same as Fig. 4.

III. RESULTS

A. Vacancy as an "isoelectronic" impurity

One of the first calculations of defect levels in a semiconductor was that of Coulson and Kearsley.[14] They treated the vacancy in diamond as a "defect molecule" made up of electrons in the sp^3 "dangling" orbitals from each of the four neighboring atoms. They then constructed many-electron states from these orbitals and found the electron energies with respect to the energies of the "original" dangling electrons. Since they created a vacancy by simply removing a carbon atom from an otherwise perfect lattice of diamond, they had really four one-electron "defect" molecular orbitals, one of a_1 symmetry, and the other three forming a triply degenerate t_2 state.

In an attempt to improve upon the result of Coulson and Kearsley, larger molecules have been constructed.[9] As a result more-complicated systems of molecular orbitals are obtained. The "defect" orbitals tend to spill out of the molecule and a simple one to one connection with the "defect" molecule is impossible. The number of the molecular orbitals and their localization depend very much on the details of the model employed in the calculation.

However, even in the limit of an infinite cluster of atoms, the basic problem remains the same. Since we create a vacancy by removing a neutral carbon atom, we must find at least four orbitals, more or less localized in the area of the vacancy. For a neutral state of the vacancy, these orbitals will be populated with four ("dangling") electrons. In our method which we described in Sec. II, a solid is made of overlapping spherical pseudoatoms. We create a vacancy by removing a neutral pseudoatom from the otherwise perfect lattice. With the formalism of Sec. II we generate bound (or resonance) states which transform according to the irreducible representations of the relevant symmetry group. The wave functions associated with these states are the particular solutions of the integral equation of our scattering formalism. The wave functions are localized in the area of the vacancy and are dominated by the contributions from the valence-band Bloch functions. It is clearly these states that must be occupied with the "dangling" electrons. The first step in our calculation must be to find these states and their properties in the otherwise perfect crystal lattice.

Of course, the above-mentioned states may have very little in common with the states created by vacancies in real crystals and observed by experimentalists. This may be so even if we disregard for the moment the problem of lattice distortion. For if we now occupy the "defect" states with electrons, and if the wave functions concerned are highly localized, the self-consistent vacancy potential must be quite different from that which we used to generate these states. Furthermore, the very pseudopotential which helps us to make the first step in our calculation, in not error free. It is noted that in their Koster-Slater calculation, Callaway and Hughes discovered that their results were extremely sensitive to the strength of the pseudopotential representing a vacancy in silicon. In Figs. 6–9 we can see the levels associated with

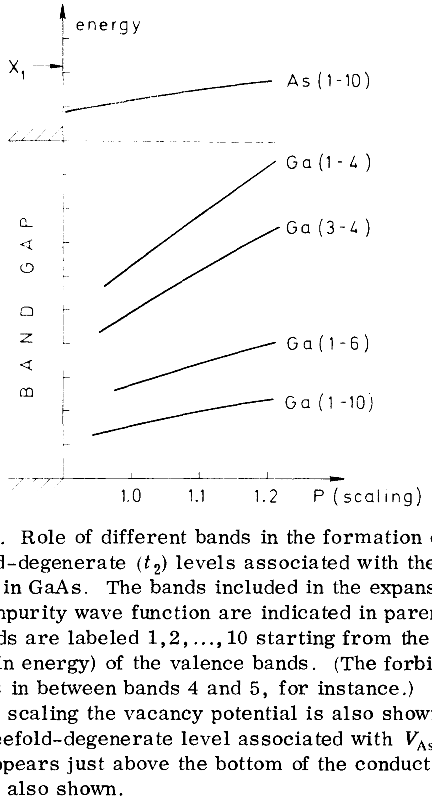

FIG. 6. Role of different bands in the formation of the threefold-degenerate (t_2) levels associated with the Ga vacancy in GaAs. The bands included in the expansion of the impurity wave function are indicated in parenthesis. The bands are labeled 1, 2, ..., 10 starting from the lowest (in energy) of the valence bands. (The forbidden gap falls in between bands 4 and 5, for instance.) The effect of scaling the vacancy potential is also shown. The threefold-degenerate level associated with V_{As}, which appears just above the bottom of the conduction band, is also shown.

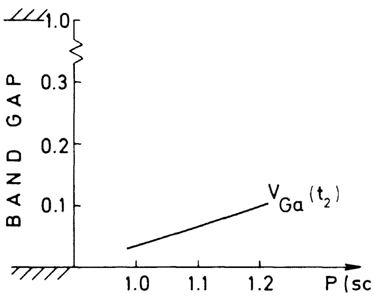

FIG. 7. Position of the threefold-degenerate levels associated with the Ga vacancy in GaP, as a function of the scaling parameter P.

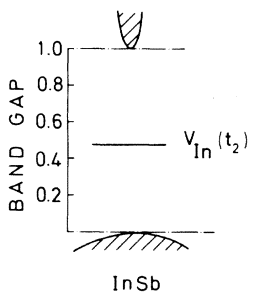

FIG. 9. Position of the threefold-degenerate level introduced by an indium vacancy into the band gap of InSb. Although no results are available for the singlet, it is expected to occur below the degenerate level shown in this figure.

single vacancies in GaP, GaAs, and InSb, obtained from our pseudopotential calculation. This is the result of the first step in our calculation in which the vacancies are treated as "substitutional impurities" in an otherwise perfect lattice. We will now describe these results (and their sensitivity to the strength of the potential) in some detail.

As expected, the threefold-degenerate p-like (t_2) states appear above the s-like (a_1) states, which in GaP and GaAs were found to be resonances in the upper part of the valence band. The Ga vacancy potential is weak and the p states are close to the valence-band edge. However, as we go from GaAs to InSb the valence band becomes more narrow and the t_2 level rises further into the gap. It is not clear whether the separation between a_1 and t_2 states follows the same trend. The a_1 states are particularly sensitive to the properties of the lowest valence band. Unfortunately, this band is not well represented by the empirical-pseudopotential band-structure calculation which provides an input for our defect level calculation. Under these circumstances the extra effort which is required to locate a resonance level in the valence band can hardly be justified. However, it would seem that the localization of the states in the valence band is not sufficient to bring about a meaningful change in our vacancy potential.

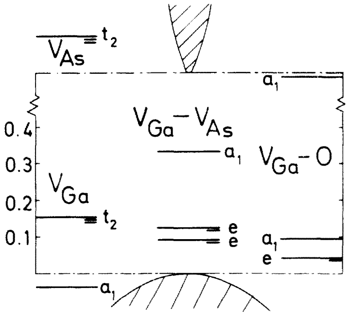

FIG. 8. A summary of results for GaAs described in Sec. III A. The column on the left shows the threefold-degenerate (resonance) level for V_{As} (above the conduction-band edge) and for V_{Ga} (above the valence-band edge). The level shown *under* the top of the valence band indicates approximately positions of the singlet (resonance) levels associated with V_{Ga} and V_{As}. The column on the right shows levels associated with V_{Ga}-O complex. The degree of degeneracy is marked by short lines beneath the levels concerned. The levels associated with the divacancy are shown in the middle column.

The As vacancy potential is strong and the t_2 state disappears into the conduction band. The a_1 state remains in the valence band. Although we only applied our calculation to V_{As}, we expect similar results for other III-V compound semiconductors.

We also applied our scheme to calculate states associated with the divacancy and V_{Ga}-O complexes in GaAs. In both cases the symmetry of the center is lower and the t_2 state is split into a singlet and a twofold-degenerate state. The states in the gap are dominated by the gallium vacancy. The states dominated by the potential on the As site are pushed into the conduction band.

In Fig. 6 we demonstrate the role of individual bands in the formation of the t_2 state associated with V_{Ga} in GaAs. The bands included in the expansion of the impurity wave function are indicated. It is clear that both the valence and conduction bands are required in order to achieve convergence. The upper valence bands dominate the state. This figure and Figs. 5 and 7 also show the sensitivity of our results to the strength of the potential. It should be noted that the energy levels which correspond to the fully convergent calculation (i.e., ten bands included in the expansion of the wave function ψ) are less sensitive to scaling than those corresponding to calculations with a truncated expansion of ψ.

It is, however, most important for us to note that all energy levels are surprisingly insensitive to the strength of the potential. It means that the over-all picture presented in Fig. 8 can hardly be altered by an error in the choice of the vacancy potential.

B. Effect of Jahn-Teller distortion and electron-electron interaction

As we indicated in the first part of this section the results in Figs. 7–9 represent vacancies in an unrelaxed lattice. The s-like (a_1) states in the valence band are assumed to be occupied with two electrons so that the levels above the valence-band edge should be related to the possible states of the third "dangling" electron. For P, As, and Sb vacancies this is their "neutral" state since be removing, say, an As atom we create three dangling electrons. In the case of Ga and In, we have five electrons to accommodate and the levels in Figs. 7–9 might represent the doubly ionized states (V^{2+}) of the vacancies in an unrelaxed lattice. If we wish to find the position of levels related to V^+, we must take into account the electron-electron interaction involving the two electrons in the localized states in the gap and repeat the calculation until the potential is self-consistent with the wave function. This was done in the Hartree approximation for the Ga vacancy in GaAs, and the resulting (one-electron) energy is shown in Fig. 11 (solid lines). It shows that the correlation energy is comparable to the electron energy in the one-electron configuration. A similar calculation was performed in a study of the two-electron levels associated with O impurity in GaP.[11] In both cases the wave functions are well localized although their details are difficult to picture. The characteristic features of these functions are very similar. The maximum of the radial charge density tends to occur near the nearest-neighbor distance.[11] The tail of the function is very complicated because the function is obtained from an expansion in Bloch states. In the present study, we chose to generate the new potential with a smooth trial function in the form $R(r) \sim r^{\alpha} e^{-\beta r}$. The parameters α and β were found by demanding that the secular equation with $R(r)$ yields the required eigenvalue, and satisfies the virial theorem.

The ground state of t_2 symmetry is threefold degenerate and is a candidate for Jahn-Teller distortion. As is customary in the treatment of Jahn-Teller problems, a considerable simplification is achieved by exploiting the symmetry of the problem. Since we are interested here only in the minimum-energy configuration we can consider the individual distortion modes separately. To reduce the amount of computation we simulate a trigonal mode of distortion by moving the vacancy along the trigonal axis indicated in Fig. 10. The calculation was performed for the Ga vacancy in GaAs, in the V^{2+} charge state, and the change in the one-electron energy as a function of the trigonal displacement is shown in Fig. 11 (dashed line). We expect this effect to be similar for the other vacancies since the localization of the states in the gap is very much the same.

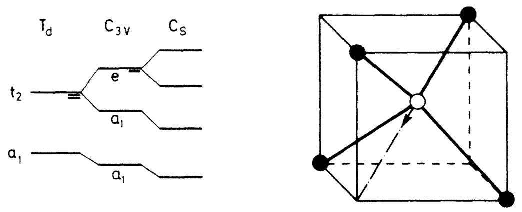

FIG. 10. A trigonal displacement (indicated by the arrow) and the corresponding effect upon the levels transforming according to the representation of the tetrahedral group (T_d) symmetry group. The trigonal group is C_{3v}. The splitting due to a distortion from C_{3v} to C_s symmetry is also shown.

At this point a comment on the problem of self-consistency may be in order. We begin our calculation with a self-consistent pseudopotential obtained from empirical-band-structure calculations. This potential has a simple meaning in the pseudopotential theory of perfect solids if we invoke the concept of screening.[13] However, in a distorted lattice the crystal potential in the vacancy area may not be smooth enough to justify the assumptions on which the screening formalism is based. In practice this means that the screening charge density should be recalculated and the vacancy potential altered in the spirit of the self-

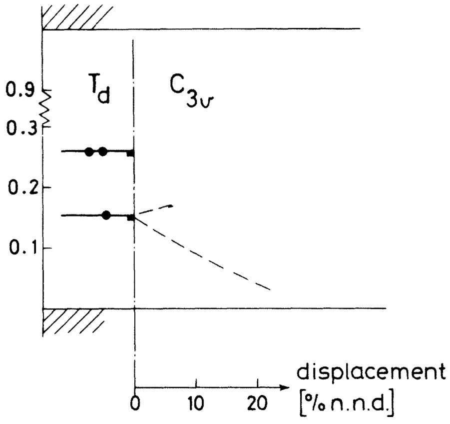

FIG. 11. Level with one full circle is the degenerate level due to V_{Ga} in GaAs (occupied with one dangling electron). If another electron is added the one-electron energy changes. This energy (per electron) is marked by two full circles. The effect upon the level (when occupied by one electron) of the trigonal displacement, indicated by the arrow in Fig. 10, is also shown (the dashed line). The displacement is measured as a percentage of the nearest-neighbor distance (nnd) of the perfect GaAs lattice.

consistent-field method.[13] In principle, such a calculation can be done within the formalism of Sec. II. However, we have so far failed to obtain a numerically meaningful appreciation of this process. The reason for it is not difficult to grasp. Let us recall the form of the wave-vector-dependent dielectric function $\epsilon(\vec{q})$. The most important part of the screened impurity (vacancy) potential comes from an area near $|\vec{q}| \approx k_{\text{Fermi}}$. At this point the polarizability contribution is only $\epsilon(k_F) - 1 \approx 0.5$.[15] Now, our results reported in the Sec. II indicate that the calculation is fairly insensitive to the strength of the vacancy potential. It would require a drastic change in the polarizability (50% or more) in order to produce a change of the order of, say 0.1 eV in the vacancy energy levels. It is difficult to see how such a change could arise in a frozen lattice. Therefore, if the changes in the valence charge density do matter then the contribution must come from a large number of smaller terms at higher $\vec{q}$ or as a result of lattice relaxation. The convergence problems in both instances are great. A similar convergence problem arises when we wish to evaluate the total energy since we have to investigate all occupied states. We have not attempted a complete calculation of the vacancy levels along these lines. However, our present calculations do provide an insight into the nature of the lattice distortion and electron-electron interaction and yield a rough quantitative evaluation of these effects. In view of the actual position of the vacancy levels in the gap (Fig. 8)—i.e., the Ga vacancy levels appear in the proximity of the valence band, whereas V_{As} levels are just above the bottom of the conduction band—we can estimate the overall change in the position of these levels which might result from the above corrections.

IV. DISCUSSION AND CONCLUSIONS

A. States associated with vacancies

To the best of our knowledge, no detailed calculations concerning electronic levels associated with vacancies in III-V semiconductors have been reported. However, the electrical properties of defects in GaAs have been studied experimentally by a number of researchers. Some authors[16] have concluded that Ga vacancies are acceptors while others[17] that As vacancies are acceptors or even[18] that both V_{Ga} and V_{As} are acceptors. More recently, Hasegawa and Majerfeld[19] reported levels which appear to be associated with Ga and As vacancies, respectively. They found an electron trap at 0.83 eV from the conduction band of high-purity vapor-phase epitaxy n-type GaAs, and a hole trap level at 0.64 eV above the valence band of liquid-phase epitaxy p-type GaAs.

Also, there is a recent work by Chiang and Pearson[20] who indicate that Ga and As vacancies behave as single acceptors and donors, respectively. It is difficult to see whether the levels observed in these experiments really belong to single vacancies or whether they are related to some more complicated centers. A similar lack of consensus seems to prevail in the case of other III-V compound semiconductors.

In Secs. III A and III B we described the results of our calculation concerning levels introduced by vacancies in some III-V semiconductors. The properties of these levels were studied under certain restrictive assumptions, with a view to establishing the role of the vacancy potential, lattice distortion, and electron-electron interaction in the formation of these states. We must now attempt to put the individual pieces of information together and build a simple picture of the energy levels which may be observed in an experiment.

Our calculation showed that both the axial distortion and the electron-electron interaction represent a substantial contribution to the energy of a state in the forbidden gap of GaAs. We also know from a previous study[11] on GaP:O that the symmetric lattice relaxation in response to charge localization in deep states lowers their energy and more than compensates for the repulsive electron-electron interaction. This effect can only be more pronounced in the case of vacancy states which are strongly hybridized with the valence band. Therefore, our results from Sec. III A, summarized in Figs. 7–9 should be reconsidered in the light of the above mentioned corrections.

1. P, As, Sb *vacancies*

There is no doubt that the vacancy potential is strong enough to push the t_2 state as far as the conduction band. This is a decisive factor. These vacancies should (in their neutral state) behave as single donors. If they capture an extra electron, a two-electron state may appear in the upper part of the gap. In materials with wider valence bands (GaP) the level should be deeper, whereas in InSb these levels are likely to remain in (or in the proximity of) the conduction band.

2. Ga, In *vacancies*

The vacancy potential is weak and even in the absence of lattice relaxation of any kind, the t_2 states appear close to the valence band. When these states are occupied with electrons, the combined effect of symmetric and Jahn-Teller lattice relaxation most probably brings the lowest two levels into the valence band. This creates a hole in the valence band and leaves an empty state in the lower part of the gap. The situation which corresponds to a neutral gallium vacancy in GaAs is pictured in a summary in Fig. 12. It follows that the Ga vacancy in GaP and GaAs should behave as a single acceptor. In materials with narrow valence bands (or near the surface) the vacancy potential becomes more effective and one of the fully occupied states may remain in the gap. The wave functions of the states in the band gap are highly localized and composed of contributions not only from the valence but also from the conduction bands (see Fig. 6). These levels are therefore likely to play a prominent part in recombination processes.

Our results for V_{Ga}-O can be interpreted in a similar way. The extra electron supplied by oxygen fills the hole in the valence band and the complex in its neutral state does not behave as an acceptor.

We can only speculate about the position of levels introduced by the divacancy V_{Ga}-V_{As}. Presumably the complex behaves as a deep acceptor producing a number of levels in the upper part, or in the vicinity of, the valence band.

B. Properties of deep neutral centers

There is probably no need to emphasize that the states associated with predominantly short-range potentials have wave functions extending over a very large area of the $\vec{k}$ space. The extent to which this is so has been studied in a number of papers and the reader may wish to compare our Fig. 6 with similar results for oxygen levels.[11]

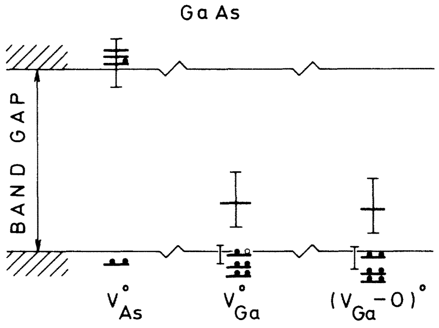

FIG. 12. Final assessment of the available impurity levels for neutral impurity systems (Ga vacancy, As vacancy, Ga-vacancy–oxygen, and divacancy) in GaAs when both the electron-electron interaction and Jahn-Teller distortion have been taken into account. All energy levels shown are singlets. The bars indicate approximately the degree of uncertainty in the position of the levels concerned.

In all cases the final result depends on a delicate process of cancellation involving a large number of numerically significant contributions. Since every part of the band structure has its own symmetry properties it is clear that the convergence in bands is absolutely necessary. It is, of course, possible to choose only a limited area of $\vec{k}$ space and still arrive at seemingly correct impurity energies. There is now plenty of evidence to show to what degree such an agreement with experiment is merely accidental.[11,21] Although one or two bands may play a dominant role in the formation of the state in question, the contribution due to the rest of the band structure may lead to an error ranging from 0 to 100%. The position of an impurity level in the gap is not in itself a good indication of the degree of localization of the wave function or its composition. Nor does it necessarily indicate which bands are important. We believe that the band structure must be well represented if one is to achieve unambiguous understanding of the problem or if one simply requires an accurate knowledge of the impurity wave function.

One of the important parameters which enables us to assess the role of a deep center in affecting carrier densities in a given material is the ability of such a center to capture electrons and holes. A complete description of the process of capture requires the knowledge of *all* localized states in or in the vicinity of the gap. Indeed it is well known that, for instance, the cascade model[22] is not applicable to centers with short-range potentials since they do not possess a large number of excited states separated by less than one phonon energy. In fact, it has been argued that a deep neutral center cannot possess any excited states at all.[23] This result has been widely accepted.[2] It means that a carrier must dispose of a large amount of energy if it is to be captured. This imposes a severe limitation on the efficiency of phonon-assisted processes. As we can see in Fig. 13 the probability might be increased if an excited state with a well-localized wave function was present. In view of the recent revival of

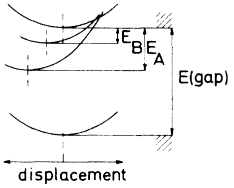

FIG. 13. Configuration coordinate diagram involving the valence and conduction bands and two impurity levels. E_A and E_B indicate the equilibrium positions. Efficiency of nonradiative capture at high temperatures is sensitive to the position of the crossing of the potential-well curves.

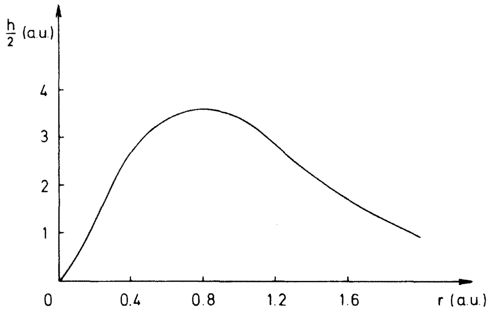

FIG. 14. Impurity potential which corresponds to the results in Fig. 15.

interest in multiphonon processes[24] it is worth asking whether the result of Bonch-Bruevich and Glasko is generally valid. They considered a neutral impurity in a semiconductor, such as silicon. They chose a potential following an earlier paper by Lax,[22] and determined its strength by demanding that there be a deep level (split off the valence band) in the middle of the gap. The band structure of the host crystal was represented by a single parabolic band, in the spirit of the well-known effective-mass theory.[1] The Schrödinger equation was then solved numerically. The calculation showed that no excited state was possible for realistic values of polarizability.

We set out to investigate the problem at an equally general level. It seems to us that the final verdict concerning the bound states associated with any particular neutral center can only be provided by a detailed calculation for that center. The purpose of our investigation is to find out whether under some physically meaningful circumstances the excited states may appear. If so, we should be in a position to show why our result disagrees with those of Bonch-Bruevich and Glasko. We chose a potential which might well be a suitable realistic potential (attractive to holes) for a deep neutral center (Fig. 14). Its maximum is just outside the core of the gallium atom and its strength characteristic for deep centers (compare with Fig. 1). The impurity is positioned at the gallium site and the calculation is performed according to Sec. II. The results are shown in Fig. 15. The states of $a_1(A)$ and $t_2(T)$ symmetry are shown. The number of bands in the expansion of the impurity wave functions is indicated. The bands are bands are labeled 1, 2, . . . , 10 starting from the lowest valence band. It is evident that the fully convergent calculation, with a realistic potential which produces a deep ground state in the gap, can also yield an excited state. The impurity energies vary slowly with the strength of the potential. On the other hand, the levels which correspond to a truncated wave function are much more sensitive to the strength of the potential and the separation between a_1 and t_2 levels is large. Therefore, a calculation based on a truncated band model is less likely to produce excited states for a physically meaningful short-range impurity potential. It is worth noting that the above result is not peculiar to the potential in Fig. 14. For instance, a slightly scaled Ga-vacancy potential shown in Fig. 1 would also yield two states in the band gap.[12] Consequently, we are led to believe that some neutral centers may exhibit several localized states in the forbidden band gap.

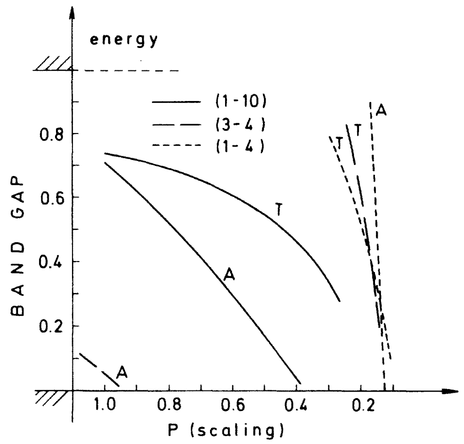

FIG. 15. Realistic impurity potential is used to study the formation of a_1 and t_2 levels as a function of bands which are labeled as in Fig. 6 ($a_1 \cdots A; t_2 \cdots T$).

ACKNOWLEDGMENTS

One of us (S.B.) wishes to acknowledge the financial assistance of the Science Research Council (U.K.). We are grateful to Dr. A. Majerfeld for providing his results prior to publication.

---

*Presently on leave of absence at Dept. of Phys. and Astronomy, Univ. of Mass. Amherst, Mass. 01002.

[1] W. Kohn, in *Solid State Physics*, edited by F. Seitz and D. Turnbull (Academic, New York, 1957), Vol. 5, p. 257.

[2] A. G. Milnes, *Deep Impurities in Semiconductors* (Wiley, New York, 1973).

[3] A. M. Stoneham, *Theory of Defects in Solids* (Oxford U. P., London, 1975).

[4] N. J. Parada, Phys. Rev. B 3, 2042 (1971).

[5] G. W. Pratt, J. Nonmetals 1, 103 (1973).

[6] L. A. Hemstreet, Phys. Rev. B 11, 2260 (1975); 12,

1212 (1975).

[7] G. F. Koster and J. C. Slater, Phys. Rev. 95, 1167 (1954); 96, 1208 (1954).

[8] J. Callaway and A. J. Hughes, Phys. Rev. 156, 860 (1967).

[9] See, for instance, R. P. Messmer and G. D. Watkins, Phys. Rev. B 7, 2568 (1973).

[10] F. Bassani, G. Iadonisi, and B. Preziosi, Phys. Rev. 186, 735 (1969).

[11] M. Jaros, J. Phys. C 8, 2455 (1975).

[12] M. Jaros, J. Phys. C 8, L550 (1975).

[13] M. L. Cohen and V. Heine, in *Solid State Physics*, edited by H. Ehrenreich, F. Seitz, and D. Turnbull (Academic, New York, 1970), Vol. 24, p. 37.

[14] C. A. Coulson and M. J. Kearsley, Proc. R. Soc. A 241, 433 (1957).

[15] See, for instance, P. K. W. Vinsome and M. Jaros, J. Phys. C 3, 2140 (1970).

[16] R. M. Logan and D. T. J. Hurle, J. Phys. Chem. Solids 32, 1739 (1971).

[17] H. Kressel, J. U. Dunse, H. Nelson, and F. Z. Hawrylo, J. Appl. Phys. 39, 2006 (1968).

[18] E. Munoz, W. L. Snyder, and J. L. Moll, Appl. Phys. Lett. 16, 262 (1970).

[19] F. Hasegawa and A. Majerfeld, Electron. Lett. 11, 286 (1975); and *ibid.* (to be published).

[20] S. Y. Chiang and G. L. Pearson, J. Appl. Phys. 46, 2986 (1975).

[21] M. Jaros and S. F. Ross, in *Proceedings of the Twelfth International Conference on the Physics of Semiconductors, Stuttgart, 1974,* edited by M. H. Pilkuhn (Teubner, Stuttgart, 1974), p. 401.

[22] M. Lax, Phys. Rev. 119, 1502 (1960).

[23] V. L. Bonch-Bruevich and V. B. Glasko, Fiz. Tverd. Tela 4, 510 (1962) [Sov. Phys.–Solid State 4, 371 (1962)].

[24] C. H. Henry and D. V. Lang, in Ref. 21, p. 411.