Electronic structure of point defects in A^4B^6 semiconductors

B. A. Volkov and O. A. Pankratov

P. N. Lebedev Physics Institute, Academy of Sciences of the USSR

(Submitted 5 July 1984)
Zh. Eksp. Teor. Fiz. **88**, 280–293 (January 1985)

Taking into account the origin of the band states of the A^4B^6 semiconductors from the atomic orbitals of p symmetry, we solve analytically the problem of the localized states of a defect with deep levels. It is shown that the ground state of the chalcogen (B) vacancy is a spin-orbit split triplet. A symmetry classification of the levels is given. A calculation of the localized electron density redistribution is carried out, making it possible to give an intuitive interpretation of the various charge states of the A and B vacancies.

1. INTRODUCTION

In A^4B^6 semiconductors the concentration and type of charge carriers are almost always governed by point defects, principally vacancies.[1] It is assumed that shallow impurity levels in these materials are generally absent, since the small effective masses $m^* \sim 0.1\, m_0$ and the large dielectric constant $\varkappa_0 \sim 10^3$ result in a negligible binding energy $m^*e^4/\hbar^2 \varkappa_0^2 \sim 10^{-6}$ eV.[1] This estimate, however, does not allow for spatial dispersion of $\varkappa_0$, the result of which is that screening is unimportant at small distances. As a result the Coloumb potential of a charged defect is short-range, as is the perturbation $V(\mathbf{r})$ that is introduced into the crystal by a neutral defect. In this case the effective mass method is in principal inapplicable, and calculation of the bound state requires knowledge of the band structure over the entire Brillouin zone.

Even for a rather small $V(\mathbf{r})$, where the radius of the localized state is large it is not possible to restrict the considerations to the vicinity of the band edges. The latter determines only the asymptotic form of wave function but not the binding energy.[1)]

This difficulty, which is typical of the problem of deep levels in semiconductors, can, as a rule, be overcome only through the use of numerical calculations. For the A^4B^6 compounds such calculations were undertaken by Parada and Pratt,[4] who calculated the vacancy levels in PbTe. The lead vacancies were found to be doubly ionized acceptors, while the tellurium vacancies proved to be doubly charged donors. It is evident that this result agrees with experimental data.[1] The energy levels that were obtained lie in the band gaps, so that, in agreement with experiment,[1] the charge carriers are not frozen out even at the lowest temperatures. According to calculation,[4] almost all the levels are located far from the band edges. The exceptions are three states that belong to the tellurium vacancy. As reported in Ref. 4, they lie 27, 55, and 218 meV above the bottom of the conduction band. In a number of investigations,[1,5] these states were invoked to interpret the experimental data.

However, it is quite clear that the accuracy attributed to the determination of the energy levels in Ref. 4 was exaggerated. Indeed, in order to calculate the positions of the deep levels it is necessary to know the energy spectrum of the ideal crystal (and Ref. 4 contains a numerical calculation of the bands). At the present time the accuracy that can be obtained in "first principles" band structure calculations is about 0.5 eV. Since the band gap ε_g in the A^4B^6 semiconductors is about $\varepsilon_g \sim 0.1$ eV, it is usually taken as an adjustable parameter from experiment.[1] This, however, does not guarantee the best accuracy for the subsequent bands, or for the values of $\varepsilon(\mathbf{k})$ at the other points of the Brillouin zone. The calculation scheme used by Parada and Pratt also did not take into account entirely the requirements of symmetry. As shown below, two of the three levels of the tellurium vacancy are degenerate purely from considerations of symmetry.

Moreover, the A^4B^6 semiconductors allow the rare possibility of an analytic treatment of the deep levels. This is due to the simple origin of their bands, which are constructed from the atomic p orbitals (the p model).[6] The p model is based on the idea of a metallic parent phase—a crystal with a simple cubic lattice and an electron spectrum composed of three quasi-one-dimensional bands ξ_{ik} (the index i refers to the p orbitals). These bands are formed by the overlap of the orbitals $|p_i\rangle$ oriented along the cubic axes. Since in the A^4B^6 compounds there are on the average three valence p electrons per atom, the $\xi_{i\mathbf{k}}$ bands are half filled. After taking into account the ionicity potential $\Delta(\mathbf{r})$, which characterizes the difference in the A and B atoms, the spectrum of the parent phase becomes dielectricized. In this process each of the bands $\xi_{i\mathbf{k}}$ is split into two bands $\pm\varepsilon_i(\mathbf{k})$, separated by an energy gap. In order to obtain the true spectrum it is still necessary to take into account mixing of the p bands by the crystal field (hybridization $\hat{W}$) and by the spin-orbit interaction $\hat{\lambda}$. In this approach the band structures of the various A^4B^6 compounds are characterized by various values of Δ, $\hat{W}$ and $\hat{\lambda}$, while the parameters of the parent phase spectrum $\xi_{i\mathbf{k}}$ are practically universal.[6]

The characteristic values of λ and W are considerably smaller than Δ.[6] Therefore, in the problem of the deep impurity, where the perturbation V is much larger than λ or W, $V \gg \lambda$, W, it is reasonable to treat the defect within the model of a dielectrized quasi-one-dimensional parent phase and then take into account the effect of $\hat{\lambda}$ and $\hat{W}$ directly on the localized states. As a result, the problem in the zeroth approximation reduces to a quasi-one-dimensional problem.

Within the framework of the p model the location of the defect in the lattice is characterized via the ionicity potential.

0038-5646/85/010164-08$04.00

Actually, for a substitutional defect, the potential $\Delta(\mathbf{r})$ appears to be even as viewed from the center point of the defect. Here, according to the definition of $\Delta(\mathbf{r})$ (Ref. 6), it is of opposite sign on the different sites (on the chalcogen (B) $\Delta < 0$ and on the metal (A) $\Delta > 0$). For an interstitial impurity the function $\Delta(\mathbf{r})$ is odd.

Modeling of the vacancy with the use of a localized perturbation $V(\mathbf{r})$ in general lacks self-consistency. This means that in the potential $V(\mathbf{r})$ one must take into account the perturbation of the electron density near the defect. At the least it is necessary to make sure that the disturbance is strongly localized and corresponds to the removal of two (for A-type vacancies) or four (for B-type vacancies) p electrons.

In this paper we calculate, within the framework of the p model, the energy spectrum of point defects (vacancies) in the A^4B^6 semiconductors. We determine the distribution of the valence electron density around the defect and find the radius of the bound state. In contrast to numerical calculations, the analytic theory permits a qualitative understanding of the nature of the difference in the charge states of the A-and B-type vacancies, it allows one to give a rigorous symmetry classification of the levels and to predict how they move as functions of composition in solid solutions of the type $Pb_{1-x}Sn_xTe$. Furthermore, experience in using the p model in band structure calculations[6] allows us to anticipate better accuracy in numerical estimates of the energy spacings between localized levels than in calculations from first principles.[4]

2. LOCALIZED STATES IN THE QUASI-ONE-DIMENSIONAL MODEL

In the one-electron approximation all the information about a system is contained in the single-particle Green's function (the resolvent operator) $\tilde{G}$, which satisfies the equation

$$\tilde{G}=\hat{G}+\hat{G}V\tilde{G}, \tag{1}$$

where $\hat{G}$ is the Green's function of the defect-free crystal. The solution of (1) has the form

$$\tilde{G}=\hat{G}+\hat{G}V(1-\hat{G}V)^{-1}\hat{G}. \tag{2}$$

The energy spectrum of the bound states is determined by the equation

$$\det\|1-\hat{G}V\|=0, \tag{3}$$

and the wave function by the relation

$$|\psi\rangle=\hat{G}V|\psi\rangle. \tag{4}$$

Within the framework of the p model the representation $|\mathbf{n}, i, \sigma\rangle$ of the localized p orbitals is used for the Green's functions (the vector $\mathbf{n}$ runs over the sites of the simple cubic lattice and σ is the spin index). In this representation the operators $\tilde{G}_{\mathbf{nn}'}$, $\hat{G}_{\mathbf{nn}'}$, and $\hat{V}_{\mathbf{nn}}$ are matrices in the indices i and σ.

To begin with we shall confine the discussion to the zeroth approximation of the p model, where the three p bands are considered noninteracting (the effects of mixing of the p bands are examined in section 4). The zeroth approximation corresponds to the Green's function

$$\hat{G}_{\mathbf{nn}'}=\begin{pmatrix} G^x_{\mathbf{nn}'} & 0 & 0 \\ 0 & G^y_{\mathbf{nn}'} & 0 \\ 0 & 0 & G^z_{\mathbf{nn}'} \end{pmatrix} \otimes \hat{1}, \tag{5}$$

where $\hat{\mathbf{1}}$ is the unit matrix in spin space. In this approximation the matrix $\hat{V}_{\mathbf{nn}}$ also has no off-diagonal elements in i. Therefore the relations (3) and (4) decompose into three independent equations for each of the p bands.

For the simplest single-site defect model, proposed by Koster and Slater[7], the matrix $\hat{V}_{\mathbf{nn}}$ has the form

$$\hat{V}_{\mathbf{nn}'}=\delta_{i,i'}\delta_{\sigma,\sigma'}\delta_{\mathbf{n},0}\delta_{\mathbf{n}',0}V_0. \tag{6}$$

Taking into account (5) and (6), we obtain from formulas (3) and (4)

$$1-G_{00}{}^iV_0=0, \tag{7}$$

$$c_{\mathbf{n}}{}^i=c_0{}^iG_{\mathbf{n}0}{}^iV_0, \tag{8}$$

where $c^i_{\mathbf{n}} = \langle \mathbf{n}, i, \sigma|\psi\rangle$ is the wave function in the $\mathbf{n}$ representation.

Koster and Slater considered the case where the impurity state is formed from the Bloch functions of only one band. In subsequent numerical calculations of course, the multiband character of the problem has been taken into account,[8] but it is not clear what qualitatively new consequences this has led to.

In the p model the spectrum of a semiconductor originates in the doubling of the period of the present phase as a result of the ionicity. Therefore the Green's function $G^i_{\mathbf{nn}'}$ characterizes both branches $\pm\varepsilon_i(\mathbf{k})$ of the dielectric spectrum and it is automatically a two-band Green's function. This circumstance is much more important than is allowance for a large number of sites in the perturbation matrix $\hat{V}_{\mathbf{nn}'}$. The one-site approximation (6) is quite sufficient to understand the reason for the radical difference in the energy spectra and charge states of the A and B vacancies.

The Hamiltonian of the ith quasi-one-dimensional band has the form

$$\hat{H}_i=\begin{pmatrix} \xi_{i,\mathbf{k}} & \Delta \\ \Delta & \xi_{i,\mathbf{k}+\mathbf{q}} \end{pmatrix}. \tag{9}$$

The matrix element of the ionicity Δ is real if we consider lattice site defects for which the potential $\Delta(\mathbf{r})$ is even. The parent phase spectrum $\xi_{i,\mathbf{k}}$ of, for example, the z band is given, according to Ref. 6, by the formula

$$\xi_{z,\mathbf{k}}=\xi_0[\cos k_z+\xi_\perp(\cos k_x+\cos k_y)]. \tag{10}$$

The spectrum (10) satisfies the condition $\xi_{i,\mathbf{k}+\mathbf{q}} = -\xi_{i,\mathbf{k}}$, where $\mathbf{q} = \pi(111)$; (the lattice constant a of the cubic parent phase is taken to be unity). From (9) it follows that

$$G^i_{\mathbf{nn}'}(\omega)=\frac{1}{N}\sum_{\mathbf{k}} e^{i\mathbf{k}(\mathbf{n}-\mathbf{n}')}\frac{\omega+(-1)^{n'}\Delta+\xi_{i,\mathbf{k}}}{\omega^2-|\Delta|^2-\xi^2_{i,\mathbf{k}}}. \tag{11}$$

The zeros of the denominator on the right hand side of (11) determines the allowed bands of the quasi-one-dimensional model after dielectrization of the parent phase:

$$\varepsilon_i(\mathbf{k})=\pm(\xi^2_{i,\mathbf{k}}+|\Delta|^2)^{1/2}. \tag{12}$$

The actual values of $\xi_\perp$ and Δ/ξ_0 are of the order 0.2 to 0.3

(Ref. 6). Therefore in the calculation of (11) we can assume that $\xi_\perp^2 \ll 1$ and the quantity

$$\varkappa^2 \equiv (|\Delta|^2 - \omega^2)/\xi_0^2$$

is also small, since we are interested in the energy region within the ionicity gap: $|\omega| < |\Delta|$. As a result of the calculations we find

$$G_{\mathbf{n}0}{}^z(\omega) = -\frac{\exp(-\varkappa|n_z|)}{\xi_0\varkappa} J_{|n_x|}(\xi_\perp|n_z|) J_{|n_y|}(\xi_\perp|n_z|) \begin{cases} \left[(\omega+\Delta)\cos\frac{|n_x|+|n_y|}{2}\pi + \varkappa\sin\frac{|n_x|+|n_y|}{2}\pi\right] \\ \times(-1)^{|n_z|/2} \\ \left[(\omega+\Delta)\sin\frac{|n_x|+|n_y|}{2}\pi - \varkappa\cos\frac{|n_x|+|n_y|}{2}\pi\right] \\ \times(-1)^{(|n_z|+1)/2} \end{cases} \quad (13)$$

where $J_n(z)$ are Bessel functions of the first kind. Of the expressions following the curly bracket, the upper one should be used when n_z is even and the lower one when n_z is odd. Formula (13) is valid under the condition

$$|n_z(\xi_\perp^3/3 - \xi_\perp\varkappa^2)| \ll 2\pi. \quad (14)$$

For large n_z the preferred status of the z axis becomes less significant. From (14) it can be seen that the anisotropic structure of the Green's function (13) is preserved at large distances from the zeroth site, out to $n_z \approx 2\pi/\xi_\perp^3 \approx 300$–$500$. The quantity G_{00}^i, which according to (7) determines the spectrum of the localized states is equal to

$$G_{00}{}^i \approx -\frac{\omega+\Delta}{\xi_0\varkappa}\left(1 + \frac{\xi_\perp{}^2 - \varkappa^2}{2}\right). \quad (15)$$

This formula follows from (11) if terms of order $\xi_\perp^2$ and $\varkappa^2$ are retained. As can be seen from (15), G_{00}^i is identical for all three bands. Consequently, if Eq. (7) has a solution, then the corresponding level is three-fold degenerate. It is easy to show that a solution of (7) exists when the signs of Δ and the effective coupling constant $g \equiv V_0/\xi_0$ are opposite. The energy of the level is

$$\varepsilon_0(g) = \Delta(1-g^2)/(1+g^2). \quad (16)$$

If $g\Delta > 0$ then a level does not arise. In the derivation of (16), the corrections $\xi_\perp^2$ and $\varkappa^2$ in (15) were dropped. As is shown below (section 4) they are important only in the calculation of the spin-orbit splitting of the triplet ε_0.

For vacancies it is natural to assume that the constant $g > 0$, i.e., it corresponds to repulsion. In the following section we shall see that actually in this case a distribution of electron density similar to the atomic electron density distribution is expelled from the vicinity of the site occupied by the vacancy. The entire difference between A and B vacancies reduces to the sign of the ionicity Δ. For the chalcogen vacancy (B) $\Delta < 0$ and therefore the level ε_0 exists (Fig. 1), while for the metal vacancy (A) $\Delta > 0$, and there is no level.

According to formula (8) the coefficients c_n^i are proportional to $G_{\mathbf{n}0}^i$ (13) and are even functions of $\mathbf{n}$. Because the p orbitals $f_i(\mathbf{r})$ are themselves odd, the wave functions in the coordinate representation

$$\psi_i(\mathbf{r}) = \sum_{\mathbf{n}} c_{\mathbf{n}}{}^i f_i(\mathbf{r}-\mathbf{n}) \quad (17)$$

are odd and transform according to the irreducible representation Γ_{15} of the cubic group. Consequently the threefold degeneracy of the level ξ_0 is due to symmetry and can be lifted only to the extent of the spin-orbit interaction.

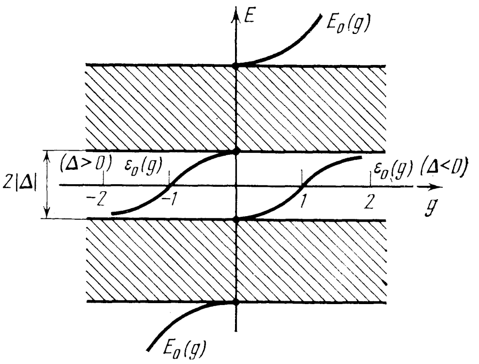

FIG. 1 Dependence of the levels of the localized states on the coupling constant in the quasi-one-dimensional model. The level $\varepsilon_0(g)$ lies in the ionicity gap $2|\Delta|$; for $\Delta > 0$ (defect on an A site) it exists only for an attractive potential $g < 0$ (left branch) and for $\Delta < 0$ (B site) it exists only for a repulsive potential $g > 0$ (right branch). The state E_0 is strongly localized and influences the electron density in the unit cell containing the defect (see section 3).

The decay of the wave function in, for instance, the z band along the z axis is characterized by a distance $a_0 \approx 1/\varkappa$. Because of the oscillations of the term after the curly bracket in (13) the amplitude of the wave function at the nearest-neighbor site is greater than at the zeroth site. In the transverse plane (n_x, n_y) the rate of falloff of the envelope of $c_{\mathbf{n}}^z$ is determined by the Bessel functions and depends on $|n_z|$. For $|n_x| \gg \xi_\perp |n_z|$ the function has the asymptotic form

$$J_{|n_x|}(\xi_\perp|n_z|) \approx (2\pi|n_x|)^{-1/2}(e\xi_\perp|n_z|/2|n_x|)^{|n_x|}.$$

If we consider the matrix elements $V_{\mathbf{nn}'} = \langle\mathbf{n}|V|\mathbf{n}'\rangle$ for $|\mathbf{n}|, |\mathbf{n}'| = 0, \pm 1$ then from Eqs. (7) and (8) we obtain the state ε_- with an odd envelope of $c_n^{i(-)}$. The condition for existence of this state, $g_-\Delta > 0$, is the opposite of the condition for the existence of the even state, and the effective coupling constant

$$g_- = (V_{1,-1} - V_{1,1})/\xi_0 \approx -V_{1,1}/\xi_0 \quad (18)$$

determines $\varepsilon_-(g_-)$ according to formula (16), in which the replacement $g \to g_-$ should be made. Thus, an A vacancy can create an even localized state that is slightly split off from the valence band (according to the smallness of g_-).

In some semiconductors of type A^4B^6 (GeTe and SnTe) a ferroelectric phase transition takes place. It is of interest to determine the effect of this transition on the depth of the level.

The relative shift of the A and B sublattices creates an odd doubling potential $\Delta_{ph}(\mathbf{r})$. This potential increases the energy gap due to the ionicity. In order to allow for the structural distortion, the replacement $\Delta \to \Delta + i\Delta_{ph}\sin k_i$ should be made in (11) (Ref. 9), where Δ_{ph} is a constant proportional to the displacement of the sublattices. It is found that in the presence of the distortion a bound state for the one-side perturbation (6) exists for any relation between the signs of g and Δ. The energy of this state is given by

$$\varepsilon_0(g) = -\frac{\Delta}{1+g^2}\left[\operatorname{sign}(g\Delta)\left(1+(1+g^2)\frac{\Delta_{ph}^2}{\Delta^2}\right)^{1/2}+g^2\right].$$

In this way the structural rearrangement makes the level of the localized state deeper and can cause the appearance of a level at an A vacancy.

According to (16) a chalcogen (B) vacancy creates a localized state within the ionicity gap $2|\Delta| \gtrsim 3$–1 eV.[6] As will be shown below, the interactions $\hat{\lambda}$ and $\hat{W}$, which cause a mixing of the quasi-one-dimensional bands, bring about only a small correction of (16) if the coupling constant $g \approx 1$. However, these interactions have an important effect on the band energies near the points of degeneracy of the bare bands $\xi_{i,\mathbf{k}}$ (Ref. 6). At the L point where the extrema of the bands are located, these interactions cause a strong reduction in the gap $2|\Delta|$ and to a small observable band gap ε_g. The L terms of the bands for various A^4B^6 compounds have been calculated within the framework of the p model in Ref. 6. The positions of these terms, as well as of the level ε_0 are referenced to the Fermi surface of the metallic parent phase. This makes it possible to determine the position of the level relative to the band edges. The coupling constant g cannot be found from the theory without a self consistent calculation of $V(\mathbf{r})$. However, it can be estimated from experiment for some specific A^4B^6 compounds. According to experimental data[5] the level of the Te vacancy in PbTe lies near the conduction band edge. For this it is necessary that $g_{Te} \approx 1$. Another possibility of determining g is associated with the fact that this constant enters into the triplet ε_0 splitting (see section 4), which can be determined experimentally. It is probable that the constant g is not very different from one chalcogen to another. For chalcogen vacancies in a series of lead salts, PbTe, PbSe, and PbS, the constants g_ν (ν = S, Se, Te) can be evaluated from the relation $g_\nu = g_{Te}\, I_\nu / I_{Te}$, where I_ν is the atomic ionization potential of the νth chalcogen. Attempts to detemine more accurately the values of g_ν are meaningful only if lattice relaxation is taken into account. This is because, owing to the degeneracy of the level ε_0, the crystalline environment of the defect undergoes a rearrangement, and as a result there is a shift and a splitting of the level (the Jahn-Teller effect).[10]

Figure 2 shows the position of the level ε_0 calculated from formula (16) for $g_{Te} = 1.15$ (and, correspondingly, $g_{Se} = 1.24$, and $g_S = 1.32$) for the different A^4B^6 compounds. From this figure it can be seen how the level can be expected to move in solid solutions of the A^4B^6 compounds. The shift of the level ε_0 relative to the band edges in the alloy $Pb_{1-x}Sn_xTe$ agrees qualitatively with experimental data.[5,11]

The difference in the energy spectra of the A and B vacancies is caused by the difference in their charge states. As can be seen from Figs. 1 and 2, a chalcogen vacancy in the lead salts brings about a transfer of the triplet ε_0 from the valence band to the conduction band. Of the six electrons that have been "freed" four p electrons are removed along with the B atom and the other two are transferred to the conduction band. In the case of a metal vacancy, the state ε_0 does not exist, and the level ε_-, on account of the smallness of the constant g_- necessarily lies below the edge of the valence band. Therefore transfer of the states across the gap $2|\Delta|$ does not occur, and the removal of two p atoms with the metal atom leads to the formation of two holes in the valence band. In SnTe the level ε_0 lies in the valence band (Fig. 2). Therefore a Te vacancy in this compound will not be a donor, but an acceptor, just like an Sn vacancy. It is probable that this is the reason that only p-type SnTe is known.

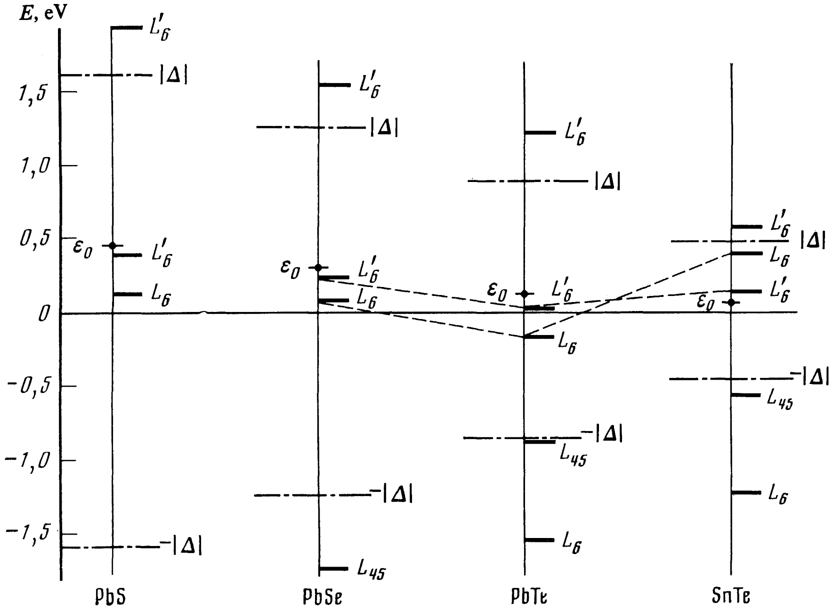

FIG. 2. Position of the level ε_0 of the chalcogen vacancy relative to the band L terms for the various A^4B^6 compounds. The coupling constant $g_{Te} = 1.15$. The band gap ε_g is formed by the levels L_6 and L_6'. The dot-dash lines show the boundaries of the ionicity gap. The motion of the L terms in the solid solutions of the corresponding binary compounds is shown by the dashed lines (from Ref. 6).

A similar explanation of the charge activity of vacancies in PbTe was given by Parada and Pratt[4] on the basis of a numerical calculation. In the numerical approach, the presence of three close-lying levels at a tellurium vacancy appears accidental. Furthermore, the complete lifting, obtained in Ref. 4, of the threefold degeneracy of the bare level ε_0 is inconsistent with symmetry considerations.

3. ELECTRON DENSITY NEAR A DEFECT

The potential $V(\mathbf{r})$ of a defect not only produces localized states, but also causes a perturbation of the valence electron density. A calculation of the electron density is required in order to check the adequacy of the description of a vacancy by a repulsive potential $V(\mathbf{r})$. Moreover, such a calculation makes it possible to give an intuitive interpretation (in coordinate space) of the charge activity of vacancies.

In the lattice site representation the variation of the electron density is determined by the standard formula

$$\delta\rho(\mathbf{n})=\oint\frac{d\omega}{2\pi i}\left[\tilde{G}_{\mathbf{nn}}(\omega)-G_{\mathbf{nn}}(\omega)\right], \tag{19}$$

where the integration is carried out over the closed contour that corresponds to the occupied states. Since, in the calculation of (19) terms of order $\varkappa^2$ cannot be neglected, we shall investigate the purely one-dimensional case ($\xi_\perp=0$), in which the problem can be solved exactly. Hence, we shall omit the band index i in the treatment below. Using formulas (2) and (6), we find from (19)

$$\delta\rho(\mathbf{n})=V_0\oint\frac{d\omega}{2\pi i}G_{\mathbf{n}0}{}^2(1-G_{00}V_0)^{-1}. \tag{20}$$

Calculating G_{n0} for the one-dimensional band $\xi_k=\xi_0\cos k$ from formula (11) and substituting the result into (20), we obtain

$$\delta\rho(n)=V_0\oint\frac{d\omega}{2\pi i}\frac{\alpha^{2|n|}(\omega)}{(\xi_0{}^2+\Delta^2-\omega^2)Z(\omega)}\begin{cases}(\Delta+\omega)/(\Delta-\omega)\\ 1\end{cases}, \tag{21}$$

where

$$Z(\omega)=1+V_0(\Delta+\omega)\left[(\xi_0{}^2+\Delta^2-\omega^2)(\Delta^2-\omega^2)\right]^{-1/2},$$
$$\alpha(\omega)=\left[(\Delta^2-\omega^2)^{1/2}-(\xi_0{}^2+\Delta^2-\omega^2)^{1/2}\right]/\xi_0. \tag{22}$$

The meaning of the curly bracket in (21) is the same as in (13). The integrand in (21) is analytic in the ω plane with cuts along the segments, $|\Delta|<|\omega|<(\xi_0^2+\Delta^2)^{1/2}$ of the real axis, corresponding to the band gaps (12). The zeros of the function $Z(\omega)$ determine the levels of the localized states. One of them, $\varepsilon_0(g)$, Eq. (16), lies in the ionicity gap. The energy of the other is

$$E_0(g)=\xi_0\,\mathrm{sign}\,g(1+\Delta^2/\xi_0{}^2+g^2)^{1/2}. \tag{23}$$

The state $E_0(g)$ is situated outside of both band gaps (Fig. 1) and it exists both in the metallic parent phase and in the dielectric phase regardless of the relation of the signs of g and Δ. This is, in fact, the level that was obtained by Koster and Slater[7] in the one-dimensional one-band model. The distribution of the electron density of the localized states is associated with the residues in (21) from the poles of (16) and (23). For the state ε_0 the distribution has the form

$$\delta\rho_{\varepsilon_0}(n)=\frac{4\theta(-g\Delta)|g^3\Delta|}{\xi_0(1+g^2)^2}\left(1-\frac{2|g\Delta|}{\xi_0(1+g^2)}\right)^{2|n|}\begin{cases}g^{-2}\\ 1\end{cases}, \tag{24}$$

where $\theta(x)$ is the step function. The distribution of electrons associated with the level (23) and lying outside of the band continuum is described by the formula

$$\delta\rho_{E_0}(n)=[g^2/(1+g^2)]^{1/2}[(1+g^2)^{1/2}-|g|]^{2|n|}.$$

For $g\gtrsim1$ the function $\delta\rho_{E_0}(n)$ is localized on a scale of order the lattice constant a of the cubic parent phase (we recall that we take $a=1$). The distribution $\delta\rho_{\varepsilon_0}$ is characterized by a radius

$$a_0=|\ln^{-1}[1-2|g\Delta|/\xi_0(1+g^2)]|, \tag{26}$$

which is a minimum ($a_0\approx\xi_0/|\Delta|$) for $g=1$, when the level of the bound state (16) is exactly at the middle of the ionicity gap. It is interesting to note that the radius a_0 is large both in the case of a small and a large interaction constant g. Within these limits, however, the distribution of the electron density is quite variable. In the case of a weak potential ($|g|\ll1$) there is a substantial change in the density, according to (24) at the even sites, i.e., at sites of the same species as that corresponding to the site occupied by the defect. In the case $|g|\gg1$ the electron density is greatly changed at the odd sites.

In order to determine the change in the density of band electrons, it is necessary to calculate the integral (21) over the contour that encloses the cut corresponding to the valence band. For a strong perturbation, $g\gg1$ the result has the form

$$\delta\rho(n)=-\frac{1}{2}\delta\rho_{E_0}(n)+\frac{\Delta}{\xi_0}\sum_{l=|n|}^{\infty}B_{|n|}^{l}\left[\left(1+\frac{\Delta^2}{\xi_0{}^2}\right)^{1/2}-\frac{|\Delta|}{\xi_0}\right]^{2l}, \tag{27}$$

where the numbers $B_{|n|}^{l}$ are defined by the expansion of the Legendre polynomials[12]:

$$P_l(\cos x)=\sum_n B_n{}^l\cos nx.$$

The first term in (27) corresponds to a density decrease localized within the unit cell (for $g>0$ the level $E_0(g)$ is empty and the contribution from (25) is absent). In accordance with the condition that the bands of the parent phase are half filled, in the absence of the perturbation V_0 the electron density at a lattice site is $\approx1/2$ (with accuracy up to terms of order Δ/ξ_0 and disregarding spin). The first term in (27) exactly cancels this quantity, so that in the case of strong repulsion there are no electrons at the site where the defect is located. This result is completely obvious. The second term is of great interest. In the first place, its sign depends on the sign of the ionicity, i.e., on the kind of site occupied by the defect. In the second place, it decays in a distance of order

$$a_{ion}\approx\xi_0/|\Delta|\approx2-4. \tag{28}$$

This dimension can be interpreted as the radius of localization of the charge that is transferred from an A site to a B site after the ionicity is turned on. It is easy to determine the total number of electrons that are concentrated in a cloud of dimension a_{ion}. In fact, the change δN in the number of particles in the valence band is equal to the number of levels that are split off. Therefore, in the case of an A vacancy, when $\delta N = 0$ (there is no ε_0 level), a cloud of radius $a_{\rm ion}$ contains an increase of 1/2 in the number of particles, and when this is added to the first term in (27) the result is zero. For a chalcogen vacancy $\delta N = -1$, and the second term in (27) contributes $-1/2$ to the total number of particles.

Formula (27) is valid for $|g| \gg 1$. In the intermediate case $g \sim 1$ there appears in the density distribution an oscillatory part that decays exponentially in a distance $a_{\rm ion}$. This part is a remnant of Friedel oscillations that are suppressed by the dielectric gap $2|\Delta|$. In the absence of this gap (for $\Delta = 0$) we obtain from (21)

$$\delta\rho(n) = -\frac{1}{2}\delta\rho_{E_0}(n) + \operatorname{sign} g \frac{4}{\pi}$$

$$\times(-1)^n |n| \sum_{l=0}^{\infty} \frac{(-1)^l[(1+g^2)^{1/2}-|g|]^{2l+1}}{(2n)^2-(2l+1)^2}. \qquad (29)$$

Here the second term, which is proportional to $(-1)^n/|n|$ for $|n| \gg 1$, represents Fridel oscillations in a one-dimensional metal.

In the weak binding limit, as one would expect, the change in the electron density in the two-band model that we are considering coincides with the results of the one-band model. It is equal to the electron density at the localized level, but with opposite sign.

It is now possible to give a clear picture of how the electron density in a crystal changes when a vacancy is created. For vacancies of both types, three electrons are expelled from the unit cell (the first term in (27) multiplied by two for the spin and by the number of p bands). This corresponds to the removal of an atom of the parent phase along with its electron cloud. The difference between a metal and a chalcogen vacancy is contained in the second term in (27) and is due to the ionicity. In the case of a metal vacancy, this term gives an increase in the electron density within the radius $a_{\rm ion}$, the integrated density increase being equal to three particles (Fig. 3a). This integrated increase in the electron density compensates for the density decrease in the central cell. The removal of two atomic p electrons is accompanied by the appearance of two holes at infinity (a doubly charged acceptor).

For a chalcogen vacancy the second term in (27) has the opposite sign and corresponds to the expulsion of three electrons from a region of size $a_{\rm ion}$ (Fig. 3b). When the central cell is taken into account, altogether six electrons are removed from the vicinity of the vacancy. Four of them are removed with the B atom and two free electrons remain at infinity (a doubly charged donor). Generally speaking, these electrons could be localized at the level ε_0. However, as shown above, in the lead chalcogenides, this level lies above the bottom of the conduction band. Therefore the electrons turn out to be free and are not frozen out at low temperatures.

Thus, the point potential $V(\mathbf{r})$ actually leads to the same electron density distribution that one would expect from qualitative considerations. At the same time the presence of a region of increased (for an A vacancy) or decreased (for a B vacancy) electron density of dimensions on the order of a few lattice constants indicates that there is some "spreading-out" of the bare perturbation. This must be taken into account in a self consistent determination of the potential $V(\mathbf{r})$. For our purposes this is not important, since we take the constant g as a phenomenological parameter, to be determined experimentally.

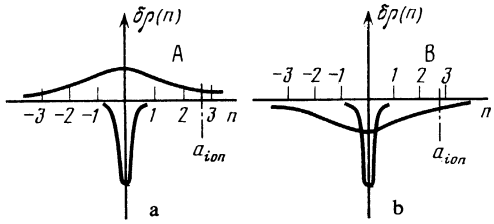

FIG. 3. Redistribution of electron density in the vicinity of the A and B vacancies.

4. THE EFFECTS OF THREE-DIMENSIONALITY; SYMMETRY CLASSIFICATION OF THE LEVELS

The number of levels of localized states increases with the number of sites that are touched upon by the potential of the defect, since the dimensionality of the matrix $\hat{V}_{\mathbf{nn}'}$ determines the degree of the secular equation (3).

The classification of the localized states and their energies depends on the location of the central point of the potential $V(\mathbf{r})$. For a defect located at a B site the one-site repulsive potential (6) creates in each of the p bands a localized state ε_0 (16). In the spectrum of the states arising when the dimensionality of the matrix $\hat{V}_{\mathbf{nn}'}$ is increased, this level is the ground state level. In accordance with the oscillation theorem its envelopes of $c^i_{\mathbf{n}}$ are even and have no nodes. Altogether the level ε_0 is threefold degenerate and the corresponding wave functions ψ_i (17) transform according to the vector representation Γ_{15} of the cubic group. This is due to the vector character of the localized p orbitals on which the envelope is defined. Mixing of the p bands by the crystal field (hybridization $\hat{W}$) cannot lift the degeneracy, since the representation Γ_{15} is irreducible. Only as a result of the spin-orbit interaction is the triplet ε_0 split into a singlet (Γ_6) and a doublet (Γ_8) level:

$$\Gamma_{15}\otimes D_{1/2}=\Gamma_6+\Gamma_8. \qquad (30)$$

This result is due only to the symmetry of the p orbitals and is therefore exact.

For the repulsive potential (6) centered on an A site the state ε_0 does not occur. The first localized level arises only when the neighboring sites $n = \pm 1$ are taken into account in each of the p-chains. This state can be interpreted as the first excited state. The state ε_- is threefold degenerate like the level ε_0. However its envelope is odd and therefore the functions ψ_i (17) transform according to the representation Λ (as x^2, y^2, and z^2). Calculating the characters, we obtain

$$\Lambda=\Gamma_1+\Gamma_{12},$$

which implies a splitting of the triplet ε_- into a singlet (Γ_1) and a doublet (Γ_{12}) already as a result of hybridization. The spin orbit interaction does not change the degree of the degeneracy:

$$\Gamma_1 \otimes D_{1/2} = \Gamma_6, \qquad \Gamma_{12} \otimes D_{1/2} = \Gamma_8.$$

Thus, the structure of the energy spectrum of the localized states is in the form of a series of doublet and singlet terms of different parity. Obviously, this is valid only for a centrally symmetric defect. In a structurally distorted lattice the degeneracy of the levels is completely lifted and the states ε_0 and ε_- can be realized at a single center.

The treatment above refers to the spectrum that arises at once in the quasi-one-dimensional model. In principle, in a three-dimensional model, other levels having a more complicated envelope symmetry are possible. For these levels the classification according to parity is preserved. However, since these levels are determined by overlap integrals with distant coordination spheres, their binding energies are extremely small, and they cannot be responsible for the charge activity of the vacancies.

Let us calculate the spin-orbit splitting of the ground state ε_0 for the one-site model of a B vacancy. In the **n** representation the spin-orbit interaction has the form

$$\hat{\lambda} = \sum_{\mathbf{n}} \hat{\lambda}_{\mathbf{n}}, \tag{31}$$

where the operators $\hat{\lambda}_{\mathbf{n}} = \lambda_{\pm} J$, and J is the matrix of the spin-orbit interaction in the basis of the p functions.[6] The constants $\lambda_{\pm}$ and λ_- refer to the A and B sites. The contribution λ_0 of the zeroth site to the sum (31) should vanish in the case of a vacancy.

The Green's function in the absence of the potential $\hat{V}_{\mathbf{nn}'}$ satisfies the equation

$$\hat{G}_{\mathbf{nn}'}^{(\lambda)} = \hat{G}_{\mathbf{nn}'} + \sum_{\mathbf{m}} \hat{G}_{\mathbf{nm}} \hat{\lambda}_{\mathbf{m}} \hat{G}_{\mathbf{mn}'}^{(\lambda)}, \tag{32}$$

where $\hat{G}_{\mathbf{nn}'}$ is defined by formulas (5) and (13). Equation (32) is easily solved in the momentum representation.[13] As a result the quantity $G_{00}^{(\lambda)}$ is expressed by a complicated triple integral which can be calculated analytically only for the case $g \equiv V_0/\xi_0 \ll 1$. For the range of values of $g \approx 1$ that we are interested in, a more effective way is the direct calculation of $\hat{G}_{00}^{(\lambda)}$ by the successive iterations of Eq. (32):

$$\hat{G}_{00}^{(\lambda)} = \hat{G}_{00} + \sum_{\mathbf{m}} \hat{G}_{0\mathbf{m}} \hat{\lambda}_{\mathbf{m}} \hat{G}_{\mathbf{m}0} + \sum_{\mathbf{m},\mathbf{m}_1} \hat{G}_{0\mathbf{m}} \hat{\lambda}_{\mathbf{m}} \hat{G}_{\mathbf{mm}_1} \lambda_{\mathbf{m}_1} \hat{G}_{\mathbf{m}_1 0} + \ldots \tag{33}$$

For $g \approx 1$ it follows from (16) that $\varepsilon_0 \ll |\Delta|$. In this energy region the series (33) converges according to how small $\lambda_{\pm}/|\Delta| < 1$ is. For a substitutional impurity in the zeroth site, $\lambda_0 \neq \lambda_{\pm} \neq 0$, and the main contribution would be given by the first order $\lambda_0 G_{00}^2$. Taking into account that the eigenvalues of $\hat{\lambda}_{\mathbf{n}}$ are $-2\lambda_{\mathbf{n}}$ (for the singlet) and $\lambda_{\mathbf{n}}$ (for the doublet) we find from Eq. (3) the magnitude of the splitting:

$$\delta\varepsilon_0 \approx 12|\Delta|\lambda_0/(1+g^2)^2\xi_0. \tag{34}$$

However, for a vacancy $\lambda_0 = 0$. In this case the first order in (33) is nonvanishing only in proportion to the size of $\xi_\perp$ which breaks down the one-dimensionality of the p band (Eq. 10). Actually, for $\xi_\perp = 0$, the functions $G_{\mathbf{nn}'}^i$ correspond to one-dimensional motion along the axes $i = x$, y, and z. Rotations are possible only at sites containing the operator $\hat{\lambda}_{\mathbf{n}}$. Therefore in first order it is impossible to construct closed trajectories which emerge from and terminate at the zeroth site. In second order the only closed paths that are possible are those that begin and end in the same band i. These trajectories do not give a nondiagonal contribution to $\hat{G}_{00}^{(\lambda)}$. Only in third order are there nontrivial trajectories which contribute to $G_{00}^{(\lambda)i,j}$ (for $i \neq j$). Thus, in (33) we must calculate the first order taking into account $\xi_\perp$ and the third order in the one-dimensional approximation. As a result we obtain

$$\hat{G}_{00}^{(\lambda)} \approx G_{00} + \hat{\lambda}_+ G_{00}^2 \left\{ \frac{\xi_\perp^2}{2} + \frac{\lambda_+^2 G_{00}^2 - \lambda_-^2/\xi_0^2 e^{4\varkappa}}{\operatorname{sh}^2 2\varkappa e^{4\varkappa}} \right\}. \tag{35}$$

From Eq. (3) it follows that in the second and third terms in (35) we can set $G_{00} \approx 1/V_0$. The magnitude of the splitting is given by formula (34), in which, in place of λ_0, one should substitute the product of λ_+ and the factor in curly brackets in (35).

In actual fact it is found that the contribution of the first order in $\hat{\lambda}$ is dominant. For a Te vacancy in PbTe, setting $\lambda_+ = 0.28$, $\lambda_- = 0.42$, $\xi_0 = 3.41$, $|\Delta| = 0.876$, and $\xi_0\xi_\perp = -0.9$ (all quantities are in eV)[6] we find that the first term in the curly brackets in (35) is more than an order of magnitude greater than the second term. An estimate of the splitting gives $\delta\varepsilon_0 \approx 10$ meV, with the singlet level lying above the doublet (Fig. 4). The magnitude of $\delta\varepsilon_0$ is clearly not accessible to numerical calculations, the accuracy of which in the best case is ~ 100 meV. This is the reason for the considerably greater splitting obtained by Parada and Pratt[4] (Fig. 4). The overlap of the levels of the localized states with the continuous spectrum (Figs. 3 and 4) results in broadening some of them. A rigorous calculation of the broadening Γ is possible within the framework of the p model, since the wave functions of both the localized and the band states are known. In the present discussion we shall limit ourselves to just an estimate of the broadening.

The magnitude of Γ should be compared with the splitting of the level ε_0 from the valence band, and for $g \approx 1$ this is of order $|\Delta|$. The standard procedure of the theory of virtual states[14] gives

$$\Gamma/|\Delta| \approx V_0 N_L,$$

where $N_L \approx 10^{-3}$ eV^{-1} is the density of band states per atom near the L point. The ratio $\Gamma/|\Delta| \sim 10^{-2}/10^{-3}$ is small to

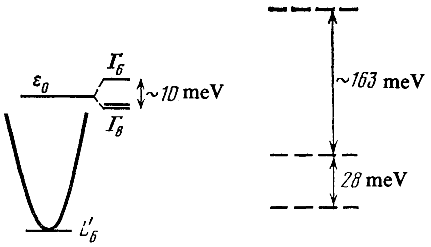

FIG. 4. Spin-orbit splitting of the triplet level ε_0 of a Te vacancy in PbTe. The dashed lines show the levels obtained by Parada and Pratt[4] from a numerical calculation.

the extent that N_L is small. Allowing for the actual form of the wave functions (that is, the selection rule for transitions between the localized level and the band states) can only make this number smaller. Moreover, the broadening Γ may be comparable with the spacing $\delta\varepsilon_0$ between the singlet and the doublet.

5. CONCLUSIONS

The A^4B^6 cubic semiconductors comprise a class of materials that can be considered model substances. Practically all their electronic properties can be understood within the framework of the p model.[6,9]

A point defect such as a vacancy is the ultimate localized probe, being sensitive to the electronic structure at distances of the order of the lattice constant. It is at just these distances (large momenta) that the zeroth approximation of the p model, the approximation in which the mixing of the p bands is weak, is effective. This makes it possible first to construct the localized state in the quasi-one-dimensional p bands and then take into account their interaction. When the p-orbital origin of the localized states is taken into account, then an exact symmetry classification of the levels can be given. Together with the simplest considerations of valency, this is sufficient to account for the different charge states of the A and B vacancies.

The parameters of the p model which characterize the band spectrum are determined from experimental data on the effective masses and energy gaps.[6] Similarly, the constants $V^{ij}_{\mathbf{nn}'}$, which characterize the defect, also can be determined from experiment. In the one-site approximation this is a single parameter $g \equiv V_0/\xi_0$. For the chalcogen vacancy the constant $g \approx 1$, and it determines not only the position of the level ε_0 but its spin-orbit splitting $\delta\varepsilon_0 \approx 10$ meV. Because of the smallness of $\delta\varepsilon_0$ the level is a quasidegenerate triplet. Therefore the capture of electrons in this level will be accompanied by a Jahn-Teller distortion of the lattice.[10] This circumstance must be taken into account in the analysis of experimental data.

Because of the high degree of degeneracy of the level ε_0, it is evident that multicharged defect states are possible. Each of them is characterized by a particular Jahn-Teller distortion symmetry. In this case an important problem is the self-consistent determination of the perturbation potential $V(\mathbf{r})$, which is different for the different charge states.

---

(1) In a number of articles (e.g., Refs. 2 and 3) this circumstance has been ignored. The authors investigated a short-range perturbing potential in a Kane-type band model, assuming the matrix element $\langle \mathbf{nk}|V|n'\mathbf{k}'\rangle$ to be independent of the quasimomentum of the Bloch states. In this approach the position of the defect in the crystal lattice drops out of the calculation, and this, for a strongly localized potential is clearly erroneous.

[1] G. Nimtz and B. Schlicht, in: Narrow Gap Semiconductors, Springer Tracts in Modern Physics **98**, (1983).
[2] A. Mauger and J. Friedel, Phys. Rev. **B12**, 2412 (1975).
[3] L. Liu and C. Vérié, Phys. Rev. Lett. **37**, 453 (1976).
[4] N. J. Parada and G. W. Pratt, Phys. Rev. Lett. **22**, 180 (1969).
[5] K. H. Gresslehner and L. J. Palmetshofer, J. Appl. Phys. **51**, 4735 (1980).
[6] B. A. Volkov, O. A. Pankratov, and A. V. Sazonov, Zh. Éksp. Teor. Fiz. **85**, 1395 (1983) [Sov. Phys. JETP **58**, 809 (1983)].
[7] G. F. Koster and J. C. Slater, Phys. Rev. **94**, 1392 (1954); Phys. Rev. **95**, 1167 (1954).
[8] S. T. Pantelides, J. Bernhole, and N. O. Lipari, J. Phys. Soc. Jpn. **49**, 235 (1980).
[9] B. A. Volkov and O. A. Pankratov, Zh. Éksp. Teor. Fiz. **75**, 1362 (1978) [Sov. Phys. JETP **48**, 687 (1978)].
[10] B. A. Volkov and O. A. Pankratov, Dokl. Akad. Nauk SSSR **255**, 93 (1980) [Sov. Phys. Dokl. **25**, 922 (1980)].
[11] B. A. Akimov, L. I. Ryabova, O. B. Yatsenko, and S. M. Chudinov, Fiz. Tekh. Polyprovodn. **13**, 752 (1979) [Sov. Phys. Semicond. **13**, 441 (1979)].
[12] I. S. Gradshtein and I. M. Ryzhik, Tablitsy integralov, summ, ryadov, i proizvedenii, Fizmattiz, Moscow (1963) [Tables of Integrals, Series, and Products, Academic Press, New York (1965)].
[13] B. A. Volkov and O. A. Pankratov, Preprints FIAN No. 127, No. 130, Moscow (1980) [Preprints No. 127 and 130 of the P. N. Lebedev Institute of Physics of the Academy of Sciences, USSR].
[14] C. Kittel, Quantum Theory of Solids, Wiley, New York (1963).

Translated by J. R. Anderson