

Deformation theory of the tellurium energy spectrum

B. A. Volkov, O. A. Pankratov, and S. V. Pakhomov

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

(Submitted 28 December 1983)

Zh. Eksp. Teor. Fiz. **86**, 2293–2303 (June 1984)

The energy spectrum of a group of eight bands in the vicinity of the extremal point H of the Brillouin zone is found by symmetry theory. Use is made of the fact that the tellurium space lattice is close to simple cubic and of the concept that the band states originate on atomic p -orbitals.

1. The semiconducting tellurium and selenium chalcogenides have an anisotropic chainlike crystal structure. It is usually represented as an aggregate of helical chains packed in a triangular lattice.¹ The unit cell contains three atoms that lie on one turn of the helix. Such a space lattice, despite its rather complicated form, can be obtained by slight distortion of a simple cubic lattice. To this end it is necessary to displace the (1,1,1) atomic planes of the cubic lattice along the twofold axes $U_{1,2,3}$ (Fig. 1) in such a way that the displacement vector of each is rotated through an angle $2\pi/3$ relative to the displacement vector of the preceding plane. As a result, the atoms that are closest to one another form helical chains (Fig. 1). To obtain the tellurium structure it suffices to stretch somewhat (to compress in the case of selenium) the initial lattice along the C_3 axis.¹⁾

The displacement of the planes, which we shall call optical distortion, leads to tripling of the period and to a lowering of the symmetry to trigonal. The subsequent acoustic deformation does not change the symmetry.

The degree of distortion of a simple cubic lattice can be assessed from the data listed in Table I. The relative value of the shift $u/A = (r_2^2 - r_1^2)/3\sqrt{2} A^2$ is expressed in terms of the distances between the nearest (r_1) and next-to-nearest (r_2) neighbors in the real chalcogen structures; A is the translation period in a plane perpendicular to the trigonal axis. The relative change of the distance between plane on account of acoustic deformation is $\varepsilon = C/C_0 - 1$, where C is the period along the trigonal axis, $C_0 = a\sqrt{3}$, a is the period of the initial cubic lattice. The value of a can be determined from the unit-cell volume, $a = (CA^2/2\sqrt{3})^{1/3}$, assuming that there is no hydrostatic compression on going over to the tellurium lattice (this assumption is immaterial for the sequel, since the change of the specific volume does not influence the functional form of the dispersion law). It can be seen from the table that with increasing atomic number the crystal lattice of the chalcogens of the selenium–tellurium–polonium series becomes close to cubic, so that polonium in its α modification already has a simple cubic structure. The dielectric gap ε_g of the electron spectrum is simultaneously decreased (see Table I), and polonium is a metal. This indicates that the dielectric character of the tellurium or selenium spectrum is due to the tripling of the simple-cube structure.

A similar situation is realized also in semimetals of the bismuth group.³ Their space lattices are obtained by doubling of the period of the cubic “parent phase” and acoustic deformation along C_3 . However, whereas in the case of bismuth the metallic character of the spectrum of the “parent

phase” is obvious (odd number of valence electrons per atom (p^3)), the “parent phase” of $\text{Te}(p^4)$ need generally speaking not be a metal. Nonetheless, even in this case the “parent phase” is metallic. This is evident from the existence of metallic α -Po.

For substances with unfilled p shell it is difficult to understand the metallic character of the “parent-phase” spectrum if it is constructed with the aid of the tight-binding method out of localized p -symmetry states (p model).^{4,5} Owing to the threefold degeneracy of the p level, the energy spectrum consists of three overlapping bands p_x , p_y , and p_z (X , Y , and Z are the cubic coordinate axes, Fig. 1). These bands are degenerate at $\mathbf{k} = 0$ (if no account is taken of spin-orbit interaction; \mathbf{k} is the quasimomentum), so that all are partly filled if the number of p electrons is less than six. If the bands p_x , p_y , and p_z do not interact (zeroth approximation of the p model), each of them is half-filled in the “parent phase” of bismuth and $2/3$ filled in that of tellurium. Dielectrization of the spectrum in semimetals and in chalcogens calls therefore for respectively doubling and tripling of the period. The dielectrization itself is possible because the Fermi surface of the “parent phase” has, with sufficient accuracy, the congruence property (is “quasicongruent”).^{4,5} The initial structure is then realigned as a result of the Peierls instability (Ref. 4).²⁾

These simple considerations explain the genesis of the electron spectrum and of the crystal structures of not only elements of the tellurium and bismuth group, but also of a

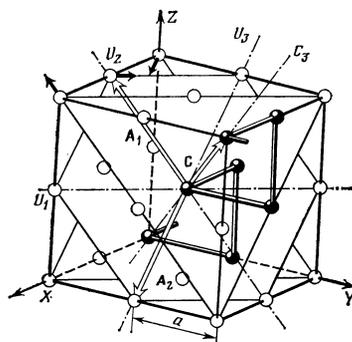


FIG. 1. Atom displacements in the formation of the tellurium structure from a simple cubic lattice. Shown is only the optical distortion [thick arrows—shifts of the (1,1,1) planes] that lead to tripling of the period (it must be supplemented by a small stretching of the lattice along the C_3 axis). The atoms connected by thick lines form helical chains of the right-circular structure of tellurium. C and $A_{1,2}$ are the new translation vectors.

Table I.

| | Se | Te | Po |
|-------------------------|--------|-------|----|
| u/A | 0,074 | 0,049 | 0 |
| ϵ | -0,049 | 0,056 | 0 |
| ϵ_g, eV | 1,2 | 0,33 | - |

number of compounds. These include IV–VI semiconductors (doubling of the period on account of the difference between the constituent atoms),^{4,5} compounds with cinnabar (HgS) structure (sextupling of the period), Bi₂Te₃ (quintupling of the period), and a number of more complicated compounds.

We confine ourselves to a determination of the form of the energy spectrum of tellurium in the vicinity of the extremal points H , regarding the distortion of the cubic “parent phase” as a perturbation. One can expect the deformation method to apply better to tellurium than to selenium, in view of the smaller deformation. Since the spectrum near the H points is determined only by the symmetry of the basis functions, no approximations connected with the tight-binding method within the p -model framework will become necessary. At the same time, only the p model will yield the correct spectrum. The point is that at the H points of the “parent phase” all the other levels are far from the Fermi surface and the correct spectrum cannot be obtained.

In the case of bismuth the effectiveness of the p model is not as evident. Abrikosov and Fal’kovskii³ obtained the energy spectrum by starting only from one irreducible representation. Certain difficulties, however, were found (nonsatisfaction of the Luttinger theorem, impossibility of explaining the observed values of the g factor); these could be avoided only after taking into account the threefold quasidegeneracy, predicted by the p model,^{5,6} at the L point of the “parent phase.” This quasidegeneracy is also extremely important for the determination of the spectrum of IV–VI semiconductors.

2. The Brillouin zone that results from the tripling of the period is of the form of a hexagonal prism (Fig. 2). The extrema of the upper occupied and the lower free energy bands in tellurium are located near its vertices—the H points.¹ In the initial cubic Brillouin zone these points lie on the diagonals of the cube faces such as to divide each diagonal into three equal parts. In the Brillouin zone of the “parent phase” all the points $H_{1,2,3}$ and $H'_{1,2,3}$ are nonequivalent. In the tellurium structure they are divided into two triplets $H_{1,2,3}$ and $H'_{1,2,3}$ of equivalent points (the states H_i and H'_i are connected by time-reversal transformation).

The wave functions φ_i in a cubic crystal are transformed in accordance with the irreducible representations of the small group C_{2v} of the point H_i . For the point H_1 (Fig. 2) this group contains the transformations of rotation of C_2 around the Z_1 axis and of reflection in planes passing through the X_1 and Y_1 axes. All the irreducible representations of the C_{2v} group are one-dimensional: $A_1\{z_1\}$, $B_1\{x_1\}$, $A_2\{x_1 y_2\}$, $B_2\{y_1\}$ (the braces contain type of basis function;

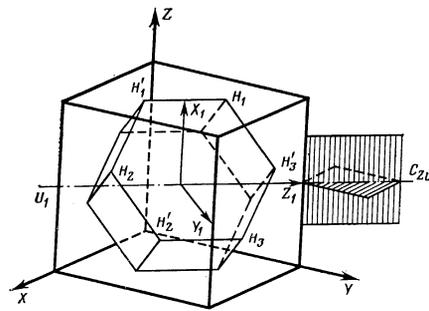


FIG. 2. Brillouin zone of the tellurium structure. Outer cube—Brillouin zone of initial cubic lattice. Owing to the acoustic deformation, the entire picture should be slightly compressed along the $[1, 1, 1]$ axis. The symmetry elements of the small group C_{2v} of the point H_1 in the initial simple cubic structure are shown.

the notation for the representations is in accord with Ref. 7). To identify the particular representations that give rise to the H terms of tellurium, we turn to the p model. In the absence of interaction between the bands p_x , p_y , and p_z , two of them are degenerate at the points H (e.g., p_x and p_y at the point H_1). The third band is separated downward by a distance $3\xi_0/2 \sim 5 \text{ eV}$, where ξ_0 is the width of the allowed band in the metallic “parent phase.” The random degeneracy of the bands p_x and p_y is lifted by the cubic-symmetry crystal potential. The doublet is split into levels with energy $+W$ and wave functions

$$|y_1\rangle = (p_x + p_y)/\sqrt{2}, \quad |z_1\rangle = (p_y - p_x)/\sqrt{2},$$

which are transformed respectively in accord with the representations B_2 and A_1 . The wave function of the separated band $|x_1\rangle = p_x$ is transformed in accord with the representation B_1 . Since $2/3$ of each of the unrenormalized quasi-one-dimensional bands is filled, the Fermi surface of the “parent phase” at the point H passes near a doublet. The forbidden band is therefore formed between sets of levels that stem from the representations A_1 and B_2 .

After the tripling of the period, the points H_i ($i = 1, 2, 3$) become equivalent and the representations $A_1^{(i)}$ and $B_2^{(i)}$ that correspond to them must be combined. This results in the unrenormalized triplets \mathcal{H}_{A_1} and \mathcal{H}_{B_2} , (Fig. 3).

The wave functions $|n_i\rangle$ ($n = z, y, x$) serve as the bases of the irreducible representations \mathcal{H}_{A_1} , \mathcal{H}_{B_2} , and \mathcal{H}_{B_1} . The latter break up into irreducible representations of the small group G_H of the point H in the tellurium lattice

$$\mathcal{H}_{A_1} = H_1 + H_3, \quad \mathcal{H}_{B_2(B_1)} = H_2 + H_3, \quad (1)$$

where the representations $H_{1,2}$ are one-dimensional and H_3 is two-dimensional.⁸ To verify the validity of (1), it is necessary to construct a table of the characters of the representations \mathcal{H}_{A_1} and \mathcal{H}_{B_1, B_2} , after subjecting the functions $|ni\rangle$ to transformations from the G_H group. This group contains a threefold screw axis and two twofold axes combined with partial translations along C_3 (Ref. 8). Recognizing, however, that the $|ni\rangle$ are defined as Bloch functions in a cubic lattice, we find it convenient to choose a C_3 axis passing through the “parent-phase” lattice sites (Fig. 1). All the partial translations included in the elements of the G_H group are then sim-

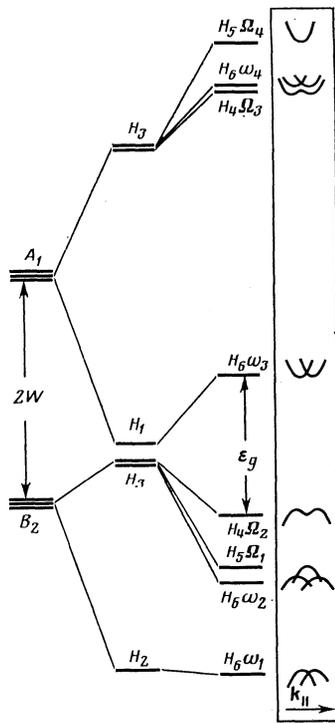


FIG. 3. Evolution of the tellurium spectrum from the energy levels A_1 and B_2 of a cubic crystal after tripling of the period and turning-on the spin-orbit interaction. The terms $\Omega_{1,2}$ and $\omega_{1,2}$ pertain to the valence bands, the others to the conduction bands.

ple combinations of the basis vectors $\mathbf{a}_{x,y,z}$ of the cubic lattice, and their action on the functions $|n_i\rangle$ reduce to multiplication by the phase factors $\exp(\pm i\pi/3)$. The two-fold axes $U_{1,2,3}$ can then be chosen respectively along the directions $[-1, 1, 0]$, $[0, -1, 1]$ and $[1, 0, -1]$ (Fig. 1) such as to make up together with the C_3 axis the point group D_3 . In the upshot, the elements of the group G_H take the following form: $\{C_3|\mathbf{a}_y\rangle$, $\{C_3^2|-\mathbf{a}_x\rangle$, $\{U_3|\mathbf{a}_y\rangle$, $\{U_2|-\mathbf{a}_x\rangle$, $\{U_1|0\rangle$, the unit operator E , the translation $T_3 = \{\varepsilon|\mathbf{a}_x + \mathbf{a}_y + \mathbf{a}_z\rangle$, and the products of T_3 by all the preceding elements.³⁾

All the projective representations of the point group of the directions D_3 of the group G_H are projectively equivalent to the vector representations, i.e., to the usual D_3 representations (Ref. 9). It suffices therefore to determine the transformation of the functions $|ni\rangle$ under the action of the operations D_3 and then compare the result with the characters of the irreducible representations $A_{1,2}$ and E of this group,⁷ which assume the role of the representations $H_{1,2}$ and H_3 of the G_H group.

According to (1), the unrenormalized triplets are split by the optical-deformation potential $\Delta(\mathbf{r})$ into a singlet level and a doublet level. The spin-orbit interaction splits the doublet H_3 , and since there is no inversion center, it is partially lifted and the Kramers degeneracy [the representations are one-dimensional for the double group $H_{4,5}$ and two-dimensional for H_6 (Fig. 3)].⁸ It follows from relation (1) that the symmetry function H_3 that pertains to the edge of the valence band can be made up of functions that transform in a cubic crystal in accord with any of the possible represen-

tations of the group of the wave vector of the point H . This means that, confining ourselves to some definite initial representations (A_1 and B_2) we construct a function H_3 of particular form. This leads to additional (compared with the usual $\mathbf{k}\cdot\mathbf{p}$ method) selection rules for the matrix elements of the $\mathbf{k}\cdot\mathbf{p}$ Hamiltonian and makes it possible to characterize the dispersion of the eight bands at the point H by a relatively small number of parameters.

3. Neglecting the split-off band B_1 , we obtain the matrix of the $\mathbf{k}\cdot\mathbf{p}$ Hamiltonian in the representation of the functions $\hat{y}_i = \langle y_i | \uparrow, \downarrow \rangle$, $\hat{z}_i = \langle z_i | \uparrow, \downarrow \rangle$ (here \uparrow and \downarrow are spin functions, while \hat{y}_i and \hat{z}_i are spinors). To this end it is necessary to make up invariants of the basis functions and of the vectors \mathbf{k}_i , reckoned from the points H_i of the spin-orbit interaction operator $\hat{\lambda}$, of the tripling potential $\Delta(\mathbf{r})$, and of the strain tensor ε_{ij} .

The quantity $\Delta(\mathbf{r})$ is invariant to transformations of the group G_H , and the tensor ε_{ij} takes in terms of the axes X_1 , Y_1 , and Z_1 the form

$$\varepsilon_{ij} = \frac{\varepsilon}{2} \begin{bmatrix} 0 & \sqrt{2} & 0 \\ \sqrt{2} & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix}. \quad (2)$$

All the diagonal components ε_{ii} are invariants of the group C_{2v} . Therefore the diagonal matrix elements of the acoustic deformation are expressed in terms of linear combinations of ε_{y_1, y_1} and ε_{z_1, z_1} :

$$\langle z_1 | \varepsilon | z_1 \rangle = a_1 \varepsilon_{y_1, y_1} + b_1 \varepsilon_{z_1, z_1},$$

$$\langle y_1 | \varepsilon | y_1 \rangle = a_2 \varepsilon_{y_1, y_1} + b_2 \varepsilon_{z_1, z_1}.$$

These invariants, together with the crystal potential of the cubic "parent phase," contribute to the splitting $2W$ of the states $|y_i\rangle$ and $|z_i\rangle$ that are degenerate in the zeroth approximation of the p model. They can therefore be taken into account by simply redefining the quantity W . The off-diagonal elements ε_{ij} make it possible to set up only a second-order invariant $\langle y_1 | \varepsilon_{x_1, y_1} k_{x_1} | z_1 \rangle \equiv p_\varepsilon k_{x_1}$, which leads to a weak dependence of the longitudinal (along C_3) momentum matrix element on ε . The smallness of this invariant is due also to the fact that it stems from the split-off band $|x_1\rangle$ and contains a large energy denominator.

We confine ourselves to invariants that are linear in Δ , ε , and λ , since all these quantities are of the same order of smallness. Therefore when setting up the invariants that contain the spin-orbit interaction operators it can be assumed that $\hat{\lambda}$ has cubic symmetry. The nonzero matrix elements of $\hat{\lambda}$ are equal to $\langle y_1 | \hat{\lambda} | z_1 \rangle = -i\lambda \sigma_{x_1}$, where for $i = 1, 2, 3$ the values of the index of the Pauli matrices σ_{x_i} coincide respectively with the indices of the cubic axes Z , X , and Y (Fig. 1).⁵⁾

By virtue of the translational symmetry, only the potential $\Delta(\mathbf{r})$ has matrix elements between the functions $|n_i\rangle$ and $|m_i\rangle$ that belong to the different points H . The symmetry properties of the functions $|n_i\rangle$ are illustrated in Fig. 4. The states $|n_{2,3}\rangle$ are obtained from $|n_1\rangle$ by successive rotations about the C_3 axis. Therefore by applying the operators $\{C_3^2|-\mathbf{a}_x\rangle$ and $\{C_3|\mathbf{a}_y\rangle$ from the G_H group to the wave

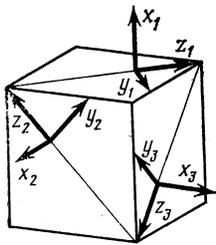


FIG. 4. Symmetry of basis functions (shown by arrows) in the point subgroup of the directions of the small group G_H .

functions, we obtain

$$\langle n_2 | \Delta | m_3 \rangle = e^{i\varphi} \langle n_3 | \Delta | m_1 \rangle = e^{-i\varphi} \langle n_1 | \Delta | m_2 \rangle,$$

where $\varphi = 2\pi/3$. With the aid of the operation $\{U_1|0\}$ it is easy to show that the matrix elements that are diagonal in n are real, $\langle n_2 | \Delta | n_3 \rangle = \langle n_3 | \Delta | n_2 \rangle$ and obtain the relation

$$\langle n_2 | \Delta | m_3 \rangle = -\langle n_3 | \Delta | m_2 \rangle \quad (n \neq m).$$

The potential $\Delta(\mathbf{r})$ can be represented in the form

$$\Delta(\mathbf{r}) = - \sum_{\mathbf{R}} \mathbf{u}_{\mathbf{R}} \nabla v(\mathbf{r}-\mathbf{R}),$$

where the summation is over the sites of the simple cubic lattice, $v(\mathbf{r})$ has cubic symmetry, and the shifts of the atoms upon tripling of the period are

$$\mathbf{u}_{\mathbf{R}} = \frac{1}{2} u (\mathbf{e}_1 + i\mathbf{e}_2) e^{-i\mathbf{q}\cdot\mathbf{R}} + \text{c.c.}, \quad \mathbf{q} = (2\pi/3)(1, 1, 1),$$

\mathbf{e}_1 and \mathbf{e}_2 are unit vector along the axes $[-1, 1, 0]$ and $[-1, -1, 2]$. For the matrix elements given above we have then

$$\Delta_1 = \langle y_2 | \Delta | y_3 \rangle \sim u_2, \quad \Delta_2 = \langle z_2 | \Delta | z_3 \rangle \sim u_2,$$

$$\delta = \langle y_2 | \Delta | z_3 \rangle \sim u_2 + i c u_1,$$

where $u_{1,2}$ is the projection of the displacement vector $\mathbf{u}_{\mathbf{R}=0}$ on the axes $\mathbf{e}_{1,2}$.

Arranging the basis functions in the sequence $\hat{y}_1, \hat{z}_1, \hat{y}_2, \hat{z}_2, \hat{y}_3, \hat{z}_3$, we write down the matrix of the Hamiltonian:

$$S^* H S = \begin{bmatrix} 2\Delta_+ + (\mathbf{n}\tau) + \alpha k_{\parallel} \tau_x & e^{-i\varphi} \Lambda & e^{-i\varphi} \Lambda^+ \\ e^{i\varphi} \Lambda^+ & -\Delta_+ + (\mathbf{n}_+ \tau) + \alpha k_{\parallel} \tau_x & \Lambda \\ e^{i\varphi} \Lambda & \Lambda^+ & -\Delta_+ + (\mathbf{n}_- \tau) + \alpha k_{\parallel} \tau_x \end{bmatrix}, \quad (6)$$

where

$$\Lambda = \left(\frac{2}{3}\right)^{1/2} \sigma_+ \tau_y + \frac{k_{\perp}}{2} \begin{bmatrix} \alpha_1 & i(p - p_{\varepsilon} \sqrt{2})/\sqrt{3} \\ i(p - p_{\varepsilon} \sqrt{2})/\sqrt{3} & \alpha_2 \end{bmatrix}, \quad (7)$$

and the vector $\tau = (\tau_x, \tau_y, \tau_z)$ is made up of Pauli matrices that act in the two-dimensional space of the functions $|y\rangle$ and $|z\rangle$. In Eqs. (6) and (7) are defined in this space the operators τ and Λ , while σ_+ and σ_{\parallel} are the spin operators

$$\sigma_+ = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad \sigma_{\parallel} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}. \quad (8)$$

The vector \mathbf{n} and \mathbf{n}_{\pm} have the following components:

$$\begin{aligned} \mathbf{n} &= (0, \lambda \sigma_{\parallel} / \sqrt{3}, W + 2\Delta_-), \\ \mathbf{n}_{\pm} &= (\pm \delta_r \sqrt{3}, \lambda \sqrt{3} \mp \delta_R \sqrt{3}, W - \Delta_-), \end{aligned} \quad (9)$$

$$H = \begin{bmatrix} K_1 & e^{i\varphi} K & e^{i\varphi} K^+ \\ e^{-i\varphi} K^+ & K_2 & K \\ e^{-i\varphi} K & K^+ & K_3 \end{bmatrix}, \quad (3)$$

where K and $K_{1,2,3}$ are the following matrices:

$$K = \begin{bmatrix} \Delta_1 & \delta \\ -\delta^* & \Delta_2 \end{bmatrix},$$

$$K_i = \begin{bmatrix} W + \alpha_1 k_{z_i} & -i\lambda \sigma_{x_i} + p k_{y_i} + p_{\varepsilon} k_{x_i} \\ i\lambda \sigma_{x_i} + p k_{y_i} + p_{\varepsilon} k_{x_i} & -W + \alpha_2 k_{z_i} \end{bmatrix}. \quad (4)$$

The parameters $\Delta_{1,2}$ and $\alpha_{1,2}$ are real (recognizing that the basis states stem from the functions $|F_x\rangle$ and $|p_y\rangle$ of the cubic parent phase, it is easy to show also that λ, p , and p_{ε} are real). All the elements of matrices (4), which contain no Pauli operators, are multiples of the unit operator in spin space; k_n are the projections of the vector \mathbf{k}_i on the axes X_i, Y_i , and Z_i .

The diagonal elements of the matrices K_i (4) do not depend on the projection of \mathbf{k} on the C_3 axis, since the vectors Z_i are coplanar (Fig. 4). Therefore, were we to confine ourselves only to some one initial representation, the Hamiltonian matrix (3) would contain no dependence whatever on the longitudinal component k (in the approximation linear in \mathbf{k}), and the corresponding effective mass would become infinite. The finite value of this mass is due to the off-diagonal elements of (4) and is connected with the quasidegeneracy of the initial levels A_1 and B_1 .

We subject the matrix (3) to the unitary transformation

$$S = \frac{1}{\sqrt{3}} \begin{bmatrix} e^{i\varphi} & 1 & 1 \\ 1 & e^{i\varphi} & 1 \\ 1 & 1 & e^{i\varphi} \end{bmatrix}, \quad (5)$$

which effects a transition to basis functions that transform in accord with the irreducible representations of the group G_H [according to (1)]. As a result we get

where δ_I and δ_R are the imaginary and real parts of the matrix element δ [see (4)], and $\Delta_{\pm} = (\Delta_1 \pm \Delta_2)/2$. The coefficient of k_{\parallel} in (6) is $\alpha = (p\sqrt{2} + p_{\varepsilon})/\sqrt{3}$. Finally, the longitudinal and transverse components k_{\parallel} and k_{\perp} of the quasi-momentum are connected with k_i as follows:

$$\begin{aligned} k_{\parallel} &= \frac{1}{\sqrt{6}} \sum_{i=1}^3 k_{y_i} = \frac{1}{\sqrt{3}} \sum_{i=1}^3 k_{x_i}, \\ k_{\perp} &= \frac{2}{3} (k_{z_1} + e^{i\varphi} k_{z_2} + e^{-i\varphi} k_{z_3}) = k_1 + i k_2, \end{aligned} \quad (10)$$

where k_1 and k_2 are the projections of the quasimomentum on the mutually perpendicular axes \mathbf{e}_1 and \mathbf{e}_2 that lie in a plane perpendicular to the C_3 axis.

We introduce the spin variable explicitly into the Hamiltonian (6) (the spin quantization axis is along C_3). With the aid of Eqs. (8), following an hermiticity-conserving rearrangement of the rows and columns, we have

$$\hat{H} = \begin{bmatrix} A_+ & C_1 & C_2^+ \\ C_1^+ & A_- & C_3 \\ C_2 & C_3^+ & D \end{bmatrix}, \quad (11)$$

where

$$A_+ = \begin{bmatrix} 2\Delta_+ + (\mathbf{n}_+\boldsymbol{\tau}) + \alpha k_{\parallel}\boldsymbol{\tau}_x & e^{-i\varphi}\boldsymbol{\nu}\boldsymbol{\tau}_y \\ e^{i\varphi}\boldsymbol{\nu}\boldsymbol{\tau}_y & -\Delta_+ + (\mathbf{n}_+\boldsymbol{\tau}) + \alpha k_{\parallel}\boldsymbol{\tau}_x \end{bmatrix},$$

$$D = \begin{bmatrix} -\Delta_+ + (\mathbf{n}_+\boldsymbol{\tau}) + \alpha k_{\parallel}\boldsymbol{\tau}_x & \boldsymbol{\nu}\boldsymbol{\tau}_y \\ \boldsymbol{\nu}\boldsymbol{\tau}_y & -\Delta_+ + (\mathbf{n}_+\boldsymbol{\tau}) + \alpha k_{\parallel}\boldsymbol{\tau}_x \end{bmatrix},$$

$$C_1 = K_{\perp}^+ \otimes \begin{bmatrix} 0 & e^{-i\varphi} \\ e^{i\varphi} & 0 \end{bmatrix}, \quad C_2 = K_{\perp}^+ \otimes \begin{bmatrix} e^{i\varphi} & 0 \\ 0 & 1 \end{bmatrix},$$

$$C_3 = K_{\perp}^+ \otimes \begin{bmatrix} 0 & -e^{i\varphi} \\ 1 & 0 \end{bmatrix}.$$

The matrix A_- is obtained from A_+ by replacing the indices \uparrow and $+$ of the vectors \mathbf{n} with their inverses \downarrow and $-$; $\boldsymbol{\nu} = (2/3)^{1/2}\lambda$; K_{\perp} is the matrix that enters as the second term in (7). The spin index of the vectors \mathbf{n} and \mathbf{n}_{\pm} means that it is necessary to replace σ_{\parallel} in (9) by $+1$ for its \uparrow value and by -1 for \downarrow .

It can be seen from (11) that at $k_{\perp} = 0$ the matrix of the Hamiltonian breaks up into three independent blocks. The first in the lower right corner describes the levels Ω_j (see Fig. 3) and the two remaining blocks the levels ω_j . The determination of the roots of the secular equation for the matrix (11) is substantially facilitated by the fact that the operators $(\mathbf{n}\cdot\boldsymbol{\tau})$ are easily diagonalized by rotating the vector $\boldsymbol{\tau}$ along \mathbf{n} . At $k = 0$ the roots Ω_j are determined exactly:

$$\Omega_j = -\Delta_{\pm} \pm \Omega_{\pm}, \quad (12)$$

where the four possible combinations of the signs correspond to the four roots $j = 1$ to 4, and the Ω_{\pm} are equal to

$$\Omega_{\pm} = [(W - \Delta_{\pm})^2 + \lambda^2 + 3|\delta|^2 - 2\lambda\delta_R \mp 2\sqrt{2}\lambda\delta_I]^{1/2}. \quad (13)$$

All the roots Ω_j are nondegenerate and pertain to two one-dimensional representations H_4 and H_5 of the binary group. On the contrary, the terms ω_j pertain to the two-dimensional representations H_6 and are doubly degenerate. This can be seen from (11), since the two corresponding diagonal blocks have identical roots. They cannot be diagonalized exactly because of the nondiagonal spin-orbit interaction $\boldsymbol{\nu}\boldsymbol{\tau}_y$. In the absence of this interaction the level energies are

$$\omega_{2,4}^{(0)} = -\Delta_{\pm} \mp |\mathbf{n}_{\pm}|, \quad \omega_{1,3}^{(0)} = 2\Delta_{\pm} \mp |\mathbf{n}_{\pm}|. \quad (14)$$

Numerical analysis shows that the level pairs that interact strongly through $\boldsymbol{\nu}\boldsymbol{\tau}_y$ are mainly $\omega_2^{(0)}$, $\omega_3^{(0)}$ and $\omega_1^{(0)}$, $\omega_4^{(0)}$. In this approximation we get

$$\omega_{2,3} \approx {}^{1/2}(\omega_2^{(0)} + \omega_3^{(0)}) \pm {}^{1/2}[(\omega_2^{(0)} - \omega_3^{(0)})^2 + 8\lambda^2/3]^{1/2}, \quad (15)$$

$$\omega_{1,4} \approx {}^{1/2}(\omega_1^{(0)} + \omega_4^{(0)}) \pm {}^{1/2}[(\omega_1^{(0)} - \omega_4^{(0)})^2 + 8\lambda^2/3]^{1/2}.$$

More accurate values of ω_i are obtained by a numerical iteration procedure.

We obtain now the dispersion of the bands along the C_3 axis for the Ω_j terms. Account of the interaction $\alpha k_{\parallel}\boldsymbol{\tau}_x$ between the nearest valence bands $\Omega_{1,2}$ must be exact, and the coupling via αk_{\parallel} with the levels $\Omega_{3,4}$ should be treated by perturbation theory. As a result we have $\Omega_{1,2}(k_{\parallel}) + \Delta_+ = -[\Omega_{3,4}(k_{\parallel}) + \Delta_+]$ and

$$\Omega_{1,2}(k_{\parallel}) = -\Delta_+ - {}^{1/2}(\Omega_+ + \Omega_-) \mp [{}^{1/4}(\Omega_+ - \Omega_-)^2 + F^2 k_{\parallel}^2]^{1/2} - A k_{\parallel}^2, \quad (16)$$

where

$$F^2 = \frac{\alpha^2}{2} \left(1 - \frac{\Omega_+^2 + \Omega_-^2 - 12\delta_I^2}{2\Omega_+\Omega_-} \right), \quad (17)$$

$$A = -\frac{\alpha^2}{2(\Omega_+ + \Omega_-)} \left(1 + \frac{\Omega_+^2 + \Omega_-^2 - 12\delta_R^2}{2\Omega_+\Omega_-} \right). \quad (18)$$

The dispersion of the bands $\omega_j^{(0)}$ is given by Eqs. (14) and (15), in which \mathbf{n}_{\pm} must be replaced by the k_{\parallel} -dependent vectors $\mathbf{n}(k_{\parallel})$ and $\mathbf{n}_{\pm}(k_{\parallel})$, which are obtained by adding the term αk_{\parallel} to the x components in Eqs. (9). From (14) we get

$$\omega_{2,4}^{(0)}(k_{\parallel}) = -\Delta_{\pm} \mp [(\pm\alpha k_{\parallel} + \delta_I\sqrt{3})^2 + (3^{-1/2}\lambda + \delta_R\sqrt{3})^2 + (W - \Delta_{\pm})^2]^{1/2}, \quad (19)$$

$$\omega_{1,3}^{(0)}(k_{\parallel}) = 2\Delta_{\pm} \mp [\alpha^2 k_{\parallel}^2 + \lambda^2/3 + (W + 2\Delta_{\pm})^2]^{1/2}. \quad (20)$$

The two signs of αk_{\parallel} in (19) correspond to the solutions of the two conjugate blocks (upper and middle) of the matrix (11). It can be seen from (20) and (15) that the shift of the band extrema $\omega_{1,3}^{(0)}$ from the point H (see Fig. 3) is due only to their interaction, via $\boldsymbol{\nu}\boldsymbol{\tau}_y$, with the "two-hump" bands $\omega_{2,4}(k_{\parallel})$ (19). This shift, however, hardly influences the effective mass, which can be determined directly from (20).

The dependence of the energy on k_{\perp} for all bands can be easily determined from (11) by perturbation theory.

4. It is known from experiment¹ that the dispersion of the valence band $\Omega_2(k_{\parallel})$ (16) has a characteristic double-hump form (Fig. 3). This calls for $\xi = |(\Omega_+ - \Omega_-)A/F^2| < 1$ (at $\xi > 1$ the band has a standard parabolic form). The experimentally determined value of ξ for tellurium¹⁰ is 0.764. From optical measurements we now also the energy gaps (at 4.2 K) $\omega_3 - \omega_2 \approx \varepsilon_g = 0.334$ eV, $\Omega_2 - \Omega_1 = \Omega_+ - \Omega_- = 0.126$ eV (Refs. 1 and 8) and $\Omega_1 - \omega_2 \approx 0.035$ eV, $\omega_2 - \omega_1 \approx 0.213$ eV (Ref. 11).

According to (12)–(15), the locations of the eight energy terms at the point H is determined by the six parameters Δ_{\pm} , $\delta_{R,I}$, W , and λ . Experience with the use of the p model or IV–VI semiconductors^{4,5} and for group-V semimetals⁶ shows that the spin-orbit constant λ is close to its value for the free atom. This constant can be gotten either from the calculation tables¹² or from an analysis of the fine structure of the atomic terms.¹³ For tellurium $\lambda = 0.18$ eV. The remaining five parameters can be determined from the value of ξ and from the four energy gaps. (We note that it follows from

(13), (17), and (18) that $\xi \approx \sqrt{2}\lambda / 3\delta_R$ if it is recognized that $|\Omega_+ - \Omega_-| \ll \Omega_{\pm}$. As a result we obtain the following set of parameters for tellurium (in eV):

$$\begin{aligned} \Delta_+ &= -0.35, & \delta_T &= 0.16, & \delta_R &= 0.04, \\ \Delta_- &= 0.26, & W &= -0.73, & \lambda &= 0.28. \end{aligned} \quad (21)$$

The sign of W cannot be obtained from the experimental data on the dispersion law. It determines the order in which the levels A_1 and B_2 follow. The situation shown in Fig. 3 corresponds to the case $W < 0$ (at $W > 0$ the level A_1 would be lower than B_2 and accordingly H_1 lower than H_2). Favoring the choice $W < 0$ are numerical calculations of the levels at the point H without allowance for the spin-orbit interaction, according to which the level H_1 lies above H_2 .¹⁴ If the absolute value of W is compared with the likewise numerically calculated¹⁵ distance $W_0 \approx 0.05$ eV between levels A_1 and B_2 in the cubic "parent phase" of tellurium, it becomes clear that the main contribution to W is made by acoustic deformation. The coefficient α of k_{\parallel} can be obtained from the experimentally known¹ parameters $A \approx 0.38 \times 10^{-14}$ eV·cm² or $F^2 \approx 0.6 \times 10^{-15}$ eV²·cm². It follows from (18) that $\alpha = (2|A|)^{1/2}(\Omega_+ + \Omega_-)^{1/2} \approx 0.64 \cdot 10^{-7}$ eV·cm. If instead of the variable k_{\parallel} we use the dimensionless quantity $k_{\parallel}a$, where $a \sim 3.24$ Å is the period of the cubic lattice of the "parent phase," the role of the coefficient α is assumed by the quantity $\alpha/a \approx 2$ eV. It is curious that if this quantity is calculated by using the tight-binding parameters $\xi_0 \approx 3.5$ eV and $\xi_1 \approx -0.9$ eV for the unrenormalized bands p_x and p_y of the "parent phase,"⁴⁾ we obtain $\alpha/a \approx p\sqrt{2}/\sqrt{3} \approx (\xi_0 - \xi_1)/2 \approx 2.2$ eV [(without allowance for the contribution of p_z (7)].

Consequently, the contribution of the acoustic deformation p_z to the matrix element of the momentum α is about 10%. Similar estimates for the coefficients of the matrix k_{\perp} yield

$$\alpha_{1,2}/a \approx -\sqrt{3}(\xi_0 + \xi_1)/4\sqrt{2} \approx -0.8 \text{ eV} \quad \text{and} \quad p/2\sqrt{3} \approx 0.8 \text{ eV}.$$

Knowing the numerical values of the parameters (21) we can calculate the energies $\Omega_{3,4}$ and ω_4 of the upper terms of the conduction bands. No experimental data on the location of these terms are known so far. With the aid of (13) and (15) we obtain $\Omega_3 - \omega_3 \approx 1.6$; $\Omega_4 - \Omega_3 = 0.126$; $\omega_4 - \Omega_3 = 0.08$ eV. It must be noted that these numbers, as well as the values given above for the constants (21), are approximate. The reason is that the experimental data used by us on the distances to the valence bands $\omega_{1,2}$ can not yet be regarded as fully reliable. In addition, the value of λ can differ somewhat from the spin-orbit constant for the free atom and should also be determined from a comparison of the calculated spectrum with experiment.

5. The developed deformation theory makes it possible to describe with a small set of parameters the dispersions of eight relatively close-lying bands. It suffices to use the experimental data on the dispersion of the upper valence band and several energy gaps.⁵⁾ We obtain simultaneously the connection between the wave functions of the different bands, which are expressed in terms of the two initial symmetry functions A_1 and B_2 . It is possible therefore, within the framework of the given scheme, to calculate the optical absorption and to determine the intensities of the transitions in

a group of eight bands. As seen from matrix (11), dipole transitions between levels of type ω_i (Fig. 3) are possible for both longitudinal and transverse polarization of the light (relative to the C_3 axis); the transitions $\Omega_i \rightarrow \Omega_j$ are allowed only for longitudinal polarization, while the transitions $\omega_i \leftrightarrow \omega_j$ are allowed for transverse. This agrees with the experimentally established¹¹ selection rules.

In the traditional $\mathbf{k}\cdot\mathbf{p}$ approach⁸ one considers usually only the group of terms $\Omega_{1,2}$ and ω_2 which appears as the result of a spin-orbit interaction λ of order ε_g and can in no way be regarded as small compared with the distance to the band ω_3 . If, however, the $\mathbf{k}\cdot\mathbf{p}$ Hamiltonian is constructed using the basis functions of representations of the binary group (the method of invariants¹⁰), the full leeway in the actual form of the functions makes the number of independent parameters very large. In essence, this is the results of an exaggeration of the accuracy since, as shown above, the wave functions of different terms are not independent but are constructed from functions of a cubic "parent phase."

Deformation theory can also be used to calculate the electronic structure of extended defects, e.g., dislocations. To this end it is necessary to introduce a coordinate dependence of the optic and acoustic distortions.

We note in conclusion that our calculation confirms the adequacy of the concept of the cubic "parent phase," which was demonstrated earlier with IV-VI semiconductors^{4,5} and bismuth⁶ as examples. This concept is apparently effective for a large group of substances in which the valence bonds are made up mainly of atomic orbitals. The parameters of the energy spectrum of the "parent phase" (ξ_0, ξ_1, W_0) for different substances turn out in this case to be close enough.

¹⁾The proximity of the structures of tellurium and selenium to cubic was noted already long ago (see, e.g., Ref. 1). It was characterized as a rule, however, by the geometry of a configuration of six atoms closest to a definite site (they form a distorted octahedron).² This description of the structure, although equivalent to the one given above, is extremely awkward for the determination of the electron spectrum.

²⁾The stability of α -Po is probably due to the tremendous spin-orbit interaction that destroys the quasicongruence of the Fermi surface of the "parent phase."

³⁾We have actually listed here the elements of the factor group of the space group D_3^4 of right-helix tellurium. The small group of the point H contains, besides these elements, translations by periods A_1 and A_2 in the (1, 1, 1) plane. But to find the substantially different representations of \bar{G}_H the factor group is sufficient.⁸ We therefore do not distinguish between G_H and \bar{G}_H .

⁴⁾The quantities ξ_0 and ξ_1 define the bands $\varepsilon_{x,y}(\mathbf{k}) = \xi_0 \cos k_{x,y}a + \xi_1 (\cos k_{y,z}a + \cos k_{z,x}a)$ (Ref. 5). The numerical values cited for them are practically universal for all IV-VI compounds⁵ and for semimetals of the bismuth group.⁶

⁵⁾Only the parameters of the upper valence band were determined reliably enough in experiment. The conduction band was little investigated, owing to the absence of n -type tellurium.

¹⁾I. M. Tsidil'kovskii, *Zonnaya struktura poluprovodnikov* (Band Structure of Semiconductors), Nauka, 1978, p. 142.

²⁾H. Krebs, *Fundamentals of Inorganic Crystal Chemistry*, McGraw, 1968.

³⁾A. A. Abrikosov and L. A. Fal'kovskii, *Zh. Eksp. Teor. Fiz.* **43**, 1989 (1962) [*Sov. Phys. JETP* **16**, 769 (1963)].

⁴⁾B. A. Volkov and O. A. Pankratov, *ibid.* **75**, 1362 (1978) [**48**, 687 (1978)].

⁵⁾B. A. Volkov, O. A. Pankratov, and A. V. Sazonov, *ibid.* **85**, 1239 1395 (1983) [**58**, 809 (1983)].

- ⁶B. A. Volkov and L. A. Fal'kovskii, *ibid.* **85**, 2135 (1983) [**58**, (1983)].
- ⁷L. D. Landau and E. M. Lifshitz, *Quantum Mechanics, Nonrelativistic Theory*, Pergamon, 1977.
- ⁸T. Doi, K. Nakao, and H. Mamimura, *J. Phys. Soc. Jpn.* **28**, 36 (1970).
- ⁹G. L. Bir and G. E. Pikus, *Symmetry and Strain-Induced Effects in Semiconductors*, Wiley, 1975.
- ¹⁰M. S. Bresler, V. G. Veselago, Yu. V. Kosichkin, G. E. Pikus, I. I. Farbshtein, and S. S. Shalyt, *Zh. Eksp. Teor. Fiz.* **57**, 1479 (1969) [*Sov. Phys. JETP* **30**, 799 (1970)].
- ¹¹R. Herwig, S. Stuke, and G. Weiser, *Phys. Stat. Sol. (b)* **61**, 521 (1974).
- ¹²F. Herman and S. Skilman, *Atomic Structure Calculations*, Prentice-hall, 1963.
- ¹³I. I. Sobel'man, *Introduction to the Theory of Atomic Spectra*, Pergamon, 1973.
- ¹⁴P. Grosse, in: *Springer Tracts in Modern Physics*, 1969, Vol. 48.
- ¹⁵C. Weigel, R. P. Messmer, and J. M. Corbett, *Phys. Stat. Sol. (b)* **57**, 455 (1973).

Translated by J. G. Adashko