SOVIET PHYSICS JETP VOLUME 24, NUMBER 1 JANUARY, 1967

TUNNEL CURRENT IN A TRANSVERSE MAGNETIC FIELD

A. G. ARONOV and G. E. PIKUS

Institute of Semiconductors, Academy of Sciences, U.S.S.R.

Submitted to JETP editor February 4, 1966

J. Exptl. Theoret. Phys. (U.S.S.R.) 51, 281–295 (July, 1966)

It is shown that in the case of a strong electric field a two-band equation should be employed to analyze the motion of electrons in crossed electric and magnetic fields E_x and H_z. In the simplest case this equation differs from the Dirac equation only in that the limiting velocity is not c but $s = (\epsilon_g/2m)^{1/2}$. In this case, the electron motion for $sH_z/cE_x < 1$ is infinite, just as for $H_z = 0$, and the decrease of the tunnel current is due to a decrease in the effective field $E = (E_x^2 - s^2H_z^2/c^2)^{1/2}$. When $sH_z/cE_x \geq 1$, the electron motion is finite and direct transitions in a homogeneous electric field are forbidden by the energy and momentum conservation laws. The effect of a magnetic field and of deformation on the current in crystals of the PbTe, PbSe and PbS type, where the constant-energy surfaces are ellipsoids, is considered. The results of the calculations are compared with the experimental data of Rediker and Calawa for PbTe.

1. INTRODUCTION

THE change of the tunnel current in a magnetic field was investigated in several experiments,[1-6] where it was established that the tunnel current decreases with increasing magnetic field—more rapidly in a transverse field and more slowly in a longitudinal one, when the magnetic and the electric fields are parallel. The theory of tunneling in a longitudinal magnetic field, developed in [1, 7-9], explains satisfactorily the experimentally observed regularities. This theory differs in fact from the tunneling theory developed by Keldysh[10] and refined by Franz[11] and Kane[12] only in that account is taken of the Landau quantization, which leads to formation of magnetic subbands.

Theoretical estimates of the current in a transverse magnetic field were made by Haering and Adams,[7] who proposed that Landau quantization always takes place when an electron moves in crossed fields, and described the electrons and holes by the ordinary single-band equations in the effective-mass approximation. Yet whereas such an approach is indeed valid for the motion of electrons in a longitudinal field, when the electron rotates in a plane perpendicular to the electric field, this approximation is not suitable in crossed fields when the electric field is strong. Indeed, in this case and when $E_x/H_z > c^{-1}(\epsilon_g/m)^{1/2}$, the energy corresponding to the drift velocity $v_y = cE_x/H_z$, which is equal to $(1/2)mc^2(E_x/H_z)^2$, as well as the energy acquired by the electron on the Larmor orbit as it rotates in the plane of the field E_x, which is equal to $mc^2(E_x/H_z)^2$, exceeds the width of the forbidden band ϵ_g. This estimate shows that when we consider the motion of electrons in crossed fields in a strong electric field, the deviation of the spectrum from parabolic plays the principal role. Therefore the calculation of the tunnel current in a transverse magnetic field cannot be carried out in the effective-mass approximation. At the same time, these calculations are not necessary to know the form of the spectrum in the entire allowed band, since an appreciable role is played only by the energies comparable with the width of the forbidden band. It would therefore be consistent to solve this problem by using the two-band model. In such a model, both bands are assumed to be unlimited, thus avoiding difficulties arising when exact account is taken of the quantization connected with the periodic dependence of the energy on the quasimomentum. We note that owing to the finite mean free path of the electron and of the hole, and also owing to the limited dimensions of the p–n junction,[9] this quantization is a physically unobservable effect.

2. THE TUNNEL CURRENT

The concrete form of the two-band model is determined by the symmetry of the crystal and by the position of the extremum point. The simplest two-band equation for nondegenerate bands coincides in form with the Dirac equation, differing from it only in that the speed of light c is replaced by the quantity

$$s = (\varepsilon_g / 2m)^{1/2}, \tag{1}$$

and the effective masses of the electron and of the hole are the same in this approximation and are equal to m. Such an equation, as noted in [13], would be valid for crystals of the Ge or InSb type with the sign of the spin-orbit splitting of the valence band reversed, provided, of course, we neglect the interaction with the other bands as well as the terms of order m/m_0, where m_0 is the mass of the free electron.

For several crystals in which the extrema of both bands are located on the edge of the Brillouin zone, for example PbSe, PbTe, and PbS, the nearest bands in each extremum are described in this approximation by an equation which differs from the Dirac equation only in the anisotropy of the effective masses. As shown in Appendix A, by suitably transforming the coordinates we can reduce this equation to the usual form

$$\{ms^2\rho_3 + \rho_1 s(\mathbf{P}'\boldsymbol{\sigma})\}\Psi = \{\varepsilon + e\Phi'(r')\}\Psi. \tag{2}$$

Here e is the charge of the electron

$$\rho_1 = \begin{vmatrix} 0 & I \\ I & 0 \end{vmatrix},$$

$$\rho_2 = \begin{vmatrix} 0 & -iI \\ iI & 0 \end{vmatrix}, \quad \rho_3 = \begin{vmatrix} I & 0 \\ 0 & -I \end{vmatrix}, \quad \sigma_i = \begin{vmatrix} \sigma_i & 0 \\ 0 & \sigma_i \end{vmatrix},$$

where σ_i are two-by-two Pauli matrices and I is a unit matrix. These matrices satisfy the following relations:

$$\rho_i\rho_k = \delta_{ik} + i\varepsilon_{ikl}\rho_l, \quad \sigma_i\sigma_k = \delta_{ik} + i\varepsilon_{ikl}\sigma_l, \quad \{\rho_i, \sigma_k\} = 0. \tag{3}$$

In this case s is determined by the state-density effective mass:

$$m = (m_x m_y m_z)^{1/3}. \tag{4}$$

In Eq. (2) it is convenient to introduce in lieu of the usual vector potential **A** the potential $\mathbf{A}' = (s/c)\mathbf{A}$, and accordingly replace the vector **H** by

$$\mathbf{H}' = \frac{s}{c}\mathbf{H}. \tag{5}$$

Then the momentum operator $\mathbf{P}'$ is defined by the expression

$$\mathbf{P}' = -i\hbar\frac{\partial}{\partial\mathbf{x}'} - \frac{e}{s}\mathbf{A}', \tag{6}$$

which differs from the usual one by replacement of c by s and of **A** by $\mathbf{A}'$. Equation (2) is invariant to a transformation that differs from the Lorentz transformation only in that c is replaced by s. The components E' and H' are transformed here like the components of the four-dimensional field tensor in relativity theory. Since this transformation leaves invariant the quantities $E'^2 - H'^2$ and $(\mathbf{E}'\cdot\mathbf{H}')$, we can, if $\mathbf{H}' \perp \mathbf{E}'$, eliminate the electric field when $H' > E'$ and eliminate the magnetic field when $E' > H'$, by choosing a suitable moving coordinate system.

For concreteness let the reduced electric field $\mathbf{E}'$, which in the usual case is equal to

$$E' = \left(E_x^2\frac{m}{m_x} + E_y^2\frac{m}{m_y} + E_z^2\frac{m}{m_z}\right)^{1/2}, \tag{7}$$

be directed along the x′ axis, and let the reduced magnetic field, which is equal to

$$H' = \frac{s}{c}\left(H_x^2\frac{m_x}{m} + H_y^2\frac{m_y}{m} + H_z^2\frac{m_z}{m}\right)^{1/2}, \tag{8}$$

be directed along the z′ axis. Here E_i and H_i are the components of the vectors along the principal axes of the corresponding ellipsoid. In the former case, i.e., when $\beta = H'/E' > 1$, we have in a coordinate system moving with velocity s/β along the y′ axis

$$E = 0, \quad H = H'(1 - \beta^{-2})^{1/2}. \tag{9}$$

In the second case, i.e., when $\beta < 1$, in a coordinate system moving with velocity βs, we have

$$H = 0, \quad E = E'(1 - \beta^2)^{1/2} = (E'^2 - H'^2)^{1/2}. \tag{10}$$

In this moving coordinate system Eq. (2) is written

$$\hat{\mathcal{H}}_0\Psi \equiv \left\{ms^2\rho_3 + \hbar s\rho_1(\mathbf{k}\boldsymbol{\sigma}) + ie\mathbf{E}\frac{\partial}{\partial\mathbf{k}}\right\}\Psi = \varepsilon\Psi. \tag{11}$$

We have gone over here directly to the k-representation, and to abbreviate the notation we tagged all the quantities in the moving coordinate system with unprimed indices. Although Eq. (11) can be solved exactly, it is more convenient for us to calculate the tunnel current, i.e., the current connected with transitions between bands, by using not this exact solution, but an approximate solution, which describes the quasistationary states of the electron in the valence band and in the conduction band. To this end we should go over to the representation in which the operator $\mathcal{H}_0$ in (11) is diagonal when $\mathbf{E} = 0$, and to calculate the interband matrix elements of the operator **x**.

The transition to this representation can be realized with the aid of a transformation similar to the Foldy-Wouthuysen transformation:[14] $T = e^{iS}$, where

$$S = \rho_2\frac{(\boldsymbol{\sigma}\mathbf{k})}{2k}\operatorname{arctg}\frac{\hbar k}{ms}, \tag{12*}$$

*arctg ≡ $\tan^{-1}$.

$$e^{iS}=\left(\frac{\eta+1}{2\eta}\right)^{1/2}+i\rho_2\frac{(\sigma\mathbf{k})}{k}\left(\frac{\eta-1}{2\eta}\right)^{1/2} \qquad (13)$$

Here

$$\eta=(1+\hbar^2k^2/m^2s^2)^{1/2}. \qquad (14)$$

In this representation

$$\tilde{\mathcal{H}}_0=e^{iS}\hat{\mathcal{H}}_0e^{-iS}=e^{2iS}\hat{\mathcal{H}}_0=\rho_3ms^2\eta=\rho_3\varepsilon(k), \qquad (15)$$

and the operator $\mathbf{x} = -i\partial/\partial\mathbf{k}$ is equal in the new representation, in accordance with (13), to

$$\mathbf{X}=e^{iS}\mathbf{x}e^{-iS}=-i\frac{\partial}{\partial\mathbf{k}}-\frac{\eta-1}{2\eta}\frac{[\sigma\mathbf{k}]}{k^2}$$

$$+\frac{\rho_2\hbar}{2ms\eta^2k^2}\{\eta[\mathbf{k}[\mathbf{k}\sigma]]-\mathbf{k}(\sigma\mathbf{k})\}. \qquad (16)^{**}$$

The operator $\mathbf{X}$ includes both the intraband term $\mathbf{X}_3$ and the intraband terms, including the second term $\mathbf{X}_2$, which describes the term splitting in the electric field, due to spin-orbit interaction. The role of the term $\mathbf{X}_2$, as will be shown below, is small, but since it has a singularity at the same point $\eta = 0$ as the interband term $\mathbf{X}_3$, we shall not omit it immediately and estimate its contribution more accurately.

When $\mathbf{E}\parallel\mathrm{x}$, we use a transformation $\tilde{S}$ similar to (12), namely

$$\tilde{S}=-\frac{\sigma_x}{2}\operatorname{arctg}\frac{k_z}{k_y}, \qquad (17)$$

to diagonalize the operator $\mathcal{H}$ with accuracy to the intraband terms, and obtain in this representation for the components of the spinor $\Psi_i^{\pm}(\mathrm{k})$ the equation

$$\left(\varepsilon_i^{\pm}+ieE\frac{\partial}{\partial k_x}\right)\Psi_i^{\pm}=\varepsilon\Psi_i^{\pm}. \qquad (18)$$

Here the index i denotes the number of the band (1—valence, 2—conduction), and the ± sign denotes the spin. At the same time

$$\varepsilon_1^{\pm}=-ms^2\eta\pm\Delta,\quad \varepsilon_2^{\pm}=+ms^2\eta\pm\Delta, \qquad (19)$$

where

$$\Delta=eE\frac{\eta-1}{2\eta}\frac{k_\perp}{k^2},\quad k_\perp^2=k_y^2+k_z^2. \qquad (20)$$

Recognizing that

$$e^{2i\tilde{S}}=(k_y+i\sigma_xk_z)/k_\perp, \qquad (21)$$

we find that in the same representation the interband operator $\hat{\mathcal{H}}'$ is equal to

$$\hat{\mathcal{H}}'=eEX_3=eE\rho_2\frac{\hbar}{2ms\eta^2k^2}$$

$$\times\{\eta\sigma_xk^2-[\sigma_xk_x^2+\sigma_yk_xk_\perp](\eta-1)\}. \qquad (22)$$

The solution of (18) is of the form

$$\Psi_i^{\pm}=(eEL_x)^{-1/2}\exp\left\{\frac{i}{eE}\int_0^{k_x}(\varepsilon-\varepsilon_i^{\pm}(k))\,dk_x\right\}\delta_{k_yk_y'}\delta_{k_zk_z'}. \qquad (23)$$

The functions $\Psi_i^{\pm}(k_x, k_y, \text{and } k_z)$ are characterized by four quantum numbers k_y', k_z', and ϵ, which we designate by a single index μ, and the spin. The normalization factor in (23) is chosen to normalize the functions $\Psi_i^{\pm}$ to energy δ-functions:

$$\sum_{k_xk_yk_z}\Psi_{i\mu'}^{*}(\mathbf{k})\,\Psi_{i\mu''}(\mathbf{k})$$

$$=\frac{L_x}{2\pi}\int dk_x\sum_{k_y,k_z}\Psi_{i\mu'}^{*}\Psi_{i\mu''}=\delta_{k_y'k_y''}\delta_{k_z'k_z''}(\varepsilon'-\varepsilon''). \qquad (24)$$

As seen from (22), in the chosen representation the operator $\hat{\mathcal{H}}'$ causes only transitions with spin flip, and the corresponding transition probabilities are

$$W_{12}^{\pm}=\frac{2\pi}{\hbar}\delta(\varepsilon'-\varepsilon'')\left|\sum_{k_xk_yk_z}^{(\pm)}\Psi_{2\mu''}^{+}\hat{\mathcal{H}}'^{(\mp)}\Psi_{1\mu'}\right|^2$$

$$=(2\pi\hbar)^{-1}\delta_{k_y'k_y''}\delta_{k_z'k_z''}\delta(\varepsilon'-\varepsilon'')\,|I_\pm|^2, \qquad (25)$$

where

$$I_\pm=\int_{-\infty}^{+\infty}dk_x\frac{[ik_x^2\mp k_xk_\perp](\eta-1)-i\eta k^2}{\eta^2k^2}$$

$$\exp\left\{\frac{2i}{eE}\int_0^{k_x}(ms^2\eta\pm\Delta)\,dk_x\right\}. \qquad (26)$$

To calculate the integral (26) we introduce the dimensionless variables

$$x=\frac{k_x}{q},\quad y=\frac{k_\perp}{q},\quad \alpha=\frac{ms}{\hbar q},\quad a=\frac{2\hbar s}{eE}q^2;$$

$$\hbar^2q^2=(ms)^2+(\hbar k_\perp)^2. \qquad (27)$$

In these variables the argument of the exponent in (26) is

$$\varphi(x)=i\int_0^x\left\{a(1+x^2)^{1/2}\pm[1-\alpha(1+x^2)^{-1/2}]\frac{y}{x^2+y^2}\right\}dx. \qquad (28)$$

**$[\mathbf{k}\sigma]\equiv\mathbf{k}\times\sigma$, $(\sigma\mathbf{k})\equiv(\sigma\cdot\mathbf{k})$.

The integral (26) along the arc in the upper half-plane vanishes, and we can deform the integration contour, as shown by the solid line in Fig. 1, and direct the cut in the upper half-plane upward from the point $x = i$ along the imaginary axis. When $\Delta = 0$ the saddle point is $\eta = 0$, i.e., $x_0 = \exp(i\pi/2)$. When $\Delta \neq 0$, the saddle points shift, but when $y \ll 1$ this shift is small and we can carry out the expansion about $x = x_0$. We put $x = x_0 + \xi$, and represent the integral $\varphi(x)$ in the form

$$\varphi(x)=\int_0^{x_0}+\int_{x_0}^{x_0+\xi}=\varphi(x_0)+\varphi(\xi).$$

Then, neglecting terms of higher order in ξ, we obtain

$$\varphi(x)=\varphi(x_0)+\frac{a}{3}(2\xi)^{3/2}e^{i3\pi/4}\pm\frac{y}{y-1}$$

$$\times[\xi-\alpha(2\xi)^{1/2}e^{-i\pi/4}]\,e^{i\pi/2}.$$

Since in the integral (26) the important values are $\xi \sim a^{-2/3}$, the last terms are of the order of $ya^{-1/3}$ and can be neglected. It is easy to verify in exactly the same way that the contribution made to the integral $\varphi(x_0)$ from the second term in (28) is of the order of y. As will be shown later, an appreciable contribution to the current is made by the values $y \sim a^{-1/2}$ and consequently, with accuracy to $y^2 \sim 1/a$, these terms can be neglected, as well as the third term in (26), which makes a contribution $\sim a^{-2/3}$. Then

$$\varphi(x_0)=ia\int_0^i\sqrt{1+x^2}\,dx=-\frac{\pi a}{4}.$$

Expanding the pre-exponential factor in (26) near x_0, neglecting terms of higher order in ξ, and also terms containing y, we obtain

$$I_\pm=-\frac{1}{4}e^{-\pi a/4}\int\frac{d\xi}{\xi}\exp\left\{\frac{a}{3}(2\xi)^{3/2}e^{i3\pi/4}\right\}. \qquad (29)$$

Putting $\xi = r\exp(i\theta)$ we find that the steepest-descent lines are the rays corresponding to $\theta = \pi/6$, and $\theta = -7\pi/6$, the angle between which is $4\pi/3$ (Fig. 1). The integrals in (29) along these rays cancel each other and all that remains is the integral along the contour, which is equal to $4\pi i/3$. Consequently

$$|I_\pm|^2=(\pi/3)^2e^{-\pi a/2}. \qquad (30)$$

In order to find the probability of the transition of the electron with a given energy ϵ', k_y', k_z', and spin, it is necessary to sum (25) over all values of k_y'' and k_z'' and integrate with respect to ϵ''. According to (25) and (30), we obtain

$$P_\pm=(\pi/18\hbar)\,e^{-\pi a/2}. \qquad (31)$$

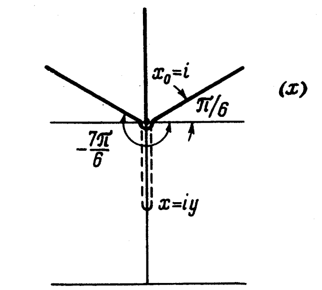


FIG. 1

To calculate this electron flux I in a time t and in a volume $V = L_xL_yL_z$, we sum (31) over all values of k_y', k_z', and the spin and integrate over the energy. Recognizing that

$$\int d\varepsilon'=eEL_x, \qquad (32)$$

we obtain from (27)

$$I=\frac{Vt}{36}\frac{eE}{\hbar}\int_0^\infty\exp\left\{-\frac{\pi\hbar s}{eE}\left(\frac{m^2s^2}{\hbar^2}+k_\perp^2\right)\right\}dk_\perp^2$$

$$=\frac{Vt(eE)^2}{36\pi\hbar^2s}\exp\left\{-\frac{\pi m^2s^3}{eE\hbar}\right\}. \qquad (33)$$

Since the four-dimensional volume Vt remains invariant under a Lorentz transformation, i.e., $Vt = V't'$, the number of transitions per cm^2 and per second is the same in the moving and stationary systems of coordinates and is equal to

$$\dot{J}=\frac{e^2(E'^2-H'^2)}{36\pi\hbar^2s}\exp\left\{-\frac{\pi m^2s^3}{e\hbar(E'^2-H'^2)^{1/2}}\right\} \qquad (34)$$

where E′ and H′ are defined in (7) and (8). Formula (34) with $H = 0$ coincides with the expression obtained by Keldysh (apart from a numerical factor), and coincides in the isotropic case with $H=0$ with the equation given in [11, 12].

We note that Kane[12] used in place of the exact two-band expression, which he derived in [15], a partially diagonalized equation, which in our notation takes the form

$$\mathcal{H}=ms^2\sigma_z+s\hbar k\sigma_x+ie\mathbf{E}\frac{\partial}{\partial\mathbf{k}}, \qquad (2a)$$

where $k = (k_x^2 + k_y^2 + k_z^2)^{1/2}$. No account was taken here of the fact that on going from (2) to (2a) the form of the operator $\mathbf{x}$ should change. In the exact calculation, the Hamiltonian (2a) yields for the interband matrix element X_3 a value which differs by a factor k_x/k from that given in [12]. In [12] this factor has been omitted and thus the correct result was obtained. Although at first glance the factor k_x/k is of little importance, it does lead to a sharp difference in the value of the integral. The point is that k_x/k has a branch point at $x = \pm iy$.

Therefore the contour should now encircle this point as shown dashed in Fig. 1. It turns out here that the main contribution to the integral is made by regions on both edges of the cut near the point iy. As a result, the expression for the transition probability differs noticeably from (31), leading to a principally different dependence of the current on the field. Thus, in the presence of the factor k_x/k, the number of transitions $\dot{J}$ would be determined not by (34) but by the formula $\dot{J} = (4\hbar/3\pi^2) \times (eE)^5/m\epsilon_g^5$. It is possible that in the case of more complicated structures and degenerate bands the matrix element X_3 will actually have singular points lying below the saddle point $x_0 = i$.

When $H'/E' \ll (8e\hbar E'/m^2s^3)^{1/4}$, the ratio of $\dot{J}(H)$ to the current $\dot{J}_0$ in the absence of a magnetic field is determined by the formula

$$\frac{J(H)}{j_0} = \exp\left\{-\frac{\pi m^2 s^3 H'^2}{2e\hbar E'^3}\right\} = \exp\left\{-\frac{\pi}{8\sqrt{2}}\frac{\varepsilon_g^{5/2}}{e\hbar c^2 m^3}\frac{\Sigma_i H_i^2 m_i}{(\Sigma_i E_i^2/m_i)^{3/2}}\right\} \quad (35)$$

We have substituted in (35) the explicit expressions for E′ and H′ from (7) and (8). We note that although the initial premises of Haering and Adams, as shown above, are debatable, the final formula (5.10) for the ratio $\dot{J}(H)/\dot{J}_0$, given in [7] for $m_1 = m_2$, differs from the formula obtainedf rom (35) for the isotropic case ($m_x = m_y = m_z$) and which is valid only for $J(H)/J_0 > 1/e$, only in a numerical factor in the exponential: in [7] this factor is equal to $4/15 = 0.266$ in place of $\pi/8\sqrt{2} = 0.277$ in (35).

3. CURRENT IN TUNNEL DIODE. COMPARISON WITH EXPERIMENT

To calculate the current in the tunnel diode, where account is taken of the final filling of both bands, it is necessary to know the current produced by the electrons with momenta and energies in the intervals dk_y, dk_z, and $d\epsilon$ with both spin orientations. According to (31), this current is equal to

$$dI = \frac{qt}{36\pi\hbar} e^{-\pi a/2}\, d\varepsilon\, dk_y\, dk_z. \quad (36)$$

Here the area is $q = L_yL_z$ and $qt = q't'$, i.e., this quantity is likewise invariant to Lorentz transformations. As is well known, [16, 17] in the quasiclassical approximation the current through a unit area is equal to

$$dJ = \frac{2}{(2\pi)^3} Dv_x\, dk_x\, dk_y\, dk_z = \frac{1}{4\pi^3\hbar} D\, d\varepsilon\, dk_y\, dk_z. \quad (37)$$

Comparing (32) and (31) we see that the quasiclassical coefficient of transparency D in a magnetic field is equal to

$$D = (\pi/3)^2 e^{-\pi a/2}. \quad (38)$$

When H = 0 and E′ = E, this expression also coincides with that given in [11, 16]. Consequently, the current through the p-n junction in a transverse magnetic field is described by formulas that differ from those given in [11, 16] only in that E is replaced by $(E'^2 - H'^2)^{1/2}$.

We emphasize that the current through the p-n junction in a magnetic field must be calculated just in a moving system of coordinates, where the average electron velocity in the absence of current through the junction is equal to zero, and consequently the distribution of the electrons in both bands is described by equilibrium Fermi functions. Here, of course, the total current along the p-n junction is zero (with the exception of the usual Hall current which is proportional to the current through the p-n junction). In the case of completely filled bands the current along the p-n junction, produced by electrons that are at the edges of the band and described by Eqs. (2), are compensated by other electrons of the band. In the p-n junction, where there are free electrons and holes, the field current along the field **E** is balanced by the diffusion current resulting from the presence of a carrier density gradient—∇n.

In exactly the same way, the "field" current connected with the motion of the electrons in a direction perpendicular to **E** and **H**, should be compensated by a diffusion current perpendicular to ∇n and **H**. Indeed, in the opposite case this current, which flows in a narrow region of the p-n junction, will be short-circuited by the body of the diode and this would lead to continuous energy dissipation.

As is well known, [16, 17] the current in a tunnel diode is determined principally by the dependence of the transparency coefficient of the dielectric field. Therefore, in accordance with (38), the current should decrease with increasing H, as was indeed observed in [1-4]. It is easy to show that, accurate to a pre-exponential factor, the relative change of the current at constant bias and its dependence on the direction of **E** and of **H** relative to the principal axes of the ellipsoids are determined by an expression similar to (34). The main contribution to the current is made by those ellipsoids for which the reduced field E′ is maximal.

Figures 2a–2c show the dependence of the tunnel current J(ϑ) on the direction of the magnetic field, given in [3] for different orientations of the

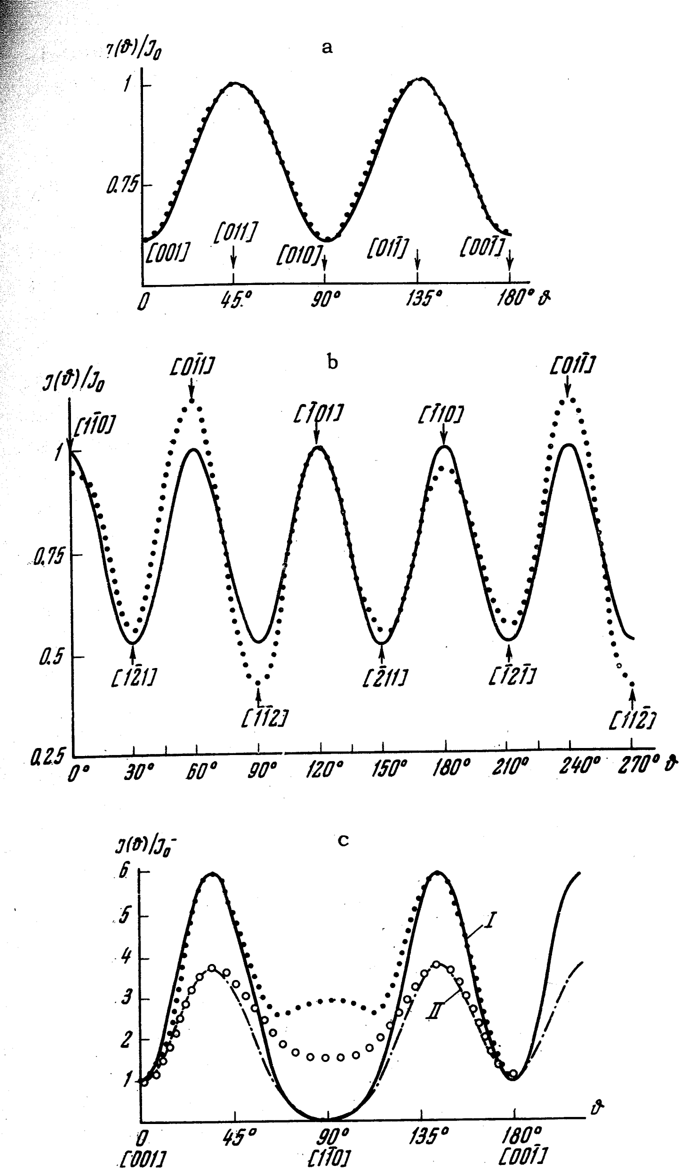


FIG. 2. Tunnel current vs. orientation of the magnetic field for PbTe diodes with different orientations of the p-n junction. The lines show the theoretical curves, the points the experimental data from [3]. The values of H are taken in accord with [3], and E was determined from the ratio of the maximum and minimum currents: a—p-n junction in (100) plane E = 114 kV/cm, H = 60 kOe; b—plane (111), E = 48 kV/cm, H = 53.2 kOe; c—plane (110), E = 45.7 kV/cm (curve 1) and E = 48.5 kV/cm (curve II), H = 53.2 kOe.

p-n junction relative to the crystallographic axes for PbTe diodes. The same figure shows plots of the corresponding theoretical relations, as given by Eq. (34). In agreement with [18], we have put here $\epsilon_g = 0.19$ eV, $m_\perp = 0.027\ m_0$, and $k = m_\parallel/m_\perp = 11$. The electric fields in the p-n junctions, indicated in the figures, were determined in such a way that the ratio of the maximum current to the minimum current coincided with the experimental value. The electric fields obtained in this manner

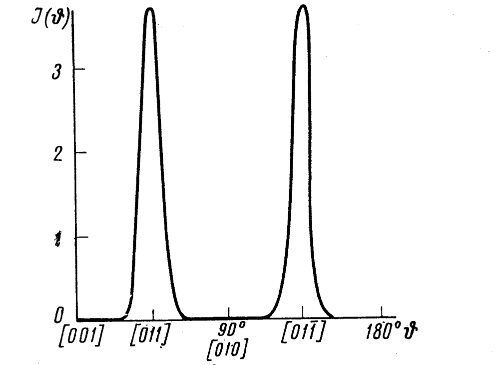
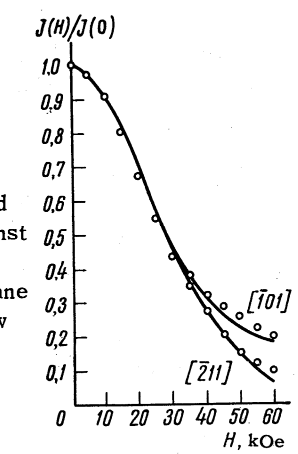


FIG. 3. Theoretical plot of the tunnel current in a PbTe diode with p-n junction in the (100) plane at E = 50.5 kV/cm and H = 60 kOe.

are in satisfactory agreement with those given in [3]. We see that the theory is in good agreement with the experimental data. A certain discrepancy in Fig. 2b is connected with the inexact orientation of the field relative to the crystallographic axis. As regards the curve on Fig. 2c, the large value of the current and even the presence of an additional maximum at ϑ = 90° cannot be attributed to the contribution of [11̄1] and [111] of the ellipsoids. Although their current should have a maximum at ϑ = 90°, its magnitude should be considerably smaller than that observed, because of the smallness of the effective electric field E′. It is possible that this effect is connected with the nonparallelism of the p-n junction.

Figure 3 shows the theoretical angular dependence of J(ϑ) for a diode in which the junction lies in the (100) plane, but the electric field is approximately half as large as for Fig. 2a. In this case there are regions of angles for which E′ < H′ and there is no tunnel current at all. As follows from the data of [3], the ratio J_{max}/J_{min} increases sharply with increasing H at corresponding orientations of H. By way of an example, Fig. 4 shows a plot of J(H)/J(0) against H for E constant and directed along [111], and for H directed along

FIG. 4. Ratio of maximum and minimum currents J(H)/J(0) against magnetic field for a PbTe diode with p-n junction in the (111) plane at E = 48 kV/cm. The lines show the theoretical curves, and the points the experimental data from [3]. The directions of H are indicated in the figure.



$[\bar{1}01]$ and $[\bar{2}11]$ respectively. We see that this plot agrees well with experiment. These calculations show that measurement of the changes of the tunnel current in a magnetic field can serve as a reliable method for measuring the characteristics of a substance and also, apparently, as a most exact method of determining the field in a p-n junction.

In concluding this section, let us discuss briefly the limits of applicability of the theory developed above. The electric field E, on the one hand, should be sufficiently weak to make the interband terms in (16) small enough to be treated as a perturbation. Consequently, the condition $a \gg 1$ should be satisfied, or

$$e(E'^2 - H'^2)^{1/2} \ll 2m^2 s^3 / \hbar. \tag{39}$$

On the other hand, the field E should be sufficiently large. Indeed, the uncertainty in the energy of the junction is $\Delta\epsilon = \hbar(1/\tau_n + 1/\tau_p)$, where τ_n and τ_p are the relaxation times of the electrons and the holes, i.e., of the particles produced past the junction. Putting in (26) $\epsilon' = \epsilon''$, we neglect terms of the order of $q\Delta\epsilon/eE = ms\Delta\epsilon/\hbar eE$ in the exponential. Consequently, the following inequality should be satisfied:

$$e(E'^2 - H'^2)^{1/2} \gg ms(1/\tau_n + 1/\tau_p). \tag{40}$$

From (39) and (40) we obtain

$$\varepsilon_g = 2ms^2 \gg \hbar(1/\tau_n + 1/\tau_p). \tag{41}$$

Condition (41) is much less stringent than the usual criteria for the applicability of the kinetic equation $kT \gg \hbar/\tau$ or $\mu \gg \hbar/\tau$ respectively for the nondegenerate and degenerate semiconductors. Of course, even a relatively weak interaction with the oscillations or with the impurity can greatly influence the characteristics of tunnel diodes if this interaction changes the density of the states at distances larger than or of the order of kT from the Fermi surface.[19]

4. TUNNEL CURRENT IN QUANTIZING MAGNETIC FIELD

From the theory developed above it follows that when $E' > H'$, when the motion of the electron along the field E' becomes infinite, the transverse magnetic field does not cause quantization. This explains the failure of attempts to observe oscillations in the tunnel current in a transverse magnetic field. Thus, the cause of the absence of oscillations is the absence of quantization, and not its smearing, as proposed in [2,7,8]. We note that in a longitudinal magnetic field, where the quantization actually takes place, such current oscillations, connected with oscillations of the chemical potentials, are clearly observed. In exactly the same way, the voltage V_{max}, corresponding to the maximum of the tunnel current and equal approximately to $(\mu_N + \mu_P)/2$, where μ_N and μ_P is the Fermi energy measured from the edges of the bands in the n and p regions, should not change appreciably in a transverse magnetic field, whereas in a longitudinal magnetic field V_{max} should increase by approximately $\hbar(\omega_n + \omega_p)/2$ times, where ω_n and ω_p are the cyclotron frequencies of the electrons and the holes. Such a difference was actually observed in [1].

It follows from (31), (34), and (38) that the tunnel current vanishes when $H' = E'$. The analysis given above shows that when $H' > E'$ there will be no tunnel current, since in this case $E = 0$ in the moving system of coordinates, and a magnetic field alone cannot lead to transitions between the bands. This property is not a feature of the chosen model. Indeed, in the case of an arbitrary band structure, the electron motion in semiconductors in crossed fields and in a sufficiently weak electric field remains finite and the Landau quantization is conserved. If **E** is directed along the x axis and **H** along the z axis, then the electrons drift in this case in the y direction with constant velocity $v_y = cE_x/H_z$,[20] and consequently $\partial\epsilon/\partial k_y = \hbar v_y$ = const, i.e., $\epsilon = \epsilon_0(E_x, k_z, n) + \hbar k_y v_y$.

Since energy is conserved in the tunnel transitions, the difference $\epsilon_{02} - \epsilon_{01}$ should be offset by the difference $\hbar(k_{y2} - k_{y1})v_y$, i.e., by the energy $eE(x_{02} - x_{01})$ acquired by the electrons as a result of the displacement of the position x_0 of the center of gravity of the oscillator on going from band to band. Here, as above, the index 1 denotes the valence band and 2 the conduction band. When $E = 0$ we have

$$(\varepsilon_{02} - \varepsilon_{01})_{min} = \varepsilon_g + {}^1\!/_2\hbar(\omega_n + \omega_p).$$

With increasing field E, the minimum energy difference $(\epsilon_{02} - \epsilon_{01})_{min}$ at fixed k_y decreases (and the energy difference at fixed x_0 increases). However, so long as $(\epsilon_{02} - \epsilon_{01})_{min}$ remains constant, the transitions can occur only if the wave vector changes, i.e., they will be proportional to the corresponding component $V_q = \int V(r)\exp(i\mathbf{q}\cdot\mathbf{r})\,dr$ of the perturbation potential $V(\mathbf{r})$. Since $V_q = 0$, when $V = eEx$, such transitions are possible in a homogeneous field only as a result of scattering by impurities or phonons. Therefore the tunnel current in a strong magnetic field should decrease rapidly. In a moving coordinate system where $E = 0$ when $H' > E'$, the transitions between the bands will be caused in this case by the motion of the scatterers in the opposite direction, i.e., with velocity $v = -s/\beta$. This case will be considered in a separate paper. We emphasize that we are dealing here with transitions between bands for which the extrema are situated in one point of k-space and where the transitions are direct in strong electric fields.

In conclusion the authors thank A. A. Ansel'm, A. I. Ansel'm, G. L. Bir, V. L. Gurevich, L. V. Keldysh, and B. D. Laĭkhtman for useful advice during the discussion and É. B. Sonin for the electronic computer calculations.

APPENDIX A

TWO-BAND EQUATION FOR TYPE PbS–PbTe CRYSTALS WITH ALLOWANCE FOR RELATIVISTIC EFFECTS

Recent experimental investigations[18] have shown that in PbS, PbSe, and PbTe crystals the extrema of both bands are situated at the point L, i.e., on the [111] axes on the edge of the Brillouin zone. In this case, as shown in [21], there are no representations for which the matrix elements of all three components of the momentum P_i in the nonrelativistic approximation would not vanish simultaneously. However, theoretical calculations of the bands[22] have shown that in these crystals the spin-orbit splitting, i.e., the distance between the terms $L_5 = L_4 + L_5$ and L_6, formed from L_3, can exceed the crystal splitting between levels of X, Y, and Z type, i.e., the distance between the representations L_2 (or L_1) and L_3, due to the splitting of the term Γ_{15} (or Γ_{25}). Therefore, as noted in [22], the wave functions of the representations $L_4 + L_5$ and L_6 contain x, y, and z components, and the relativistic terms are in fact not small.

The double-valued representations L and L* are complex and are equivalent, and k_0 is equivalent to $-k_0$. Therefore in accord with Sec. 3 of [23], if the nearest bands correspond to the representations A and B, then $\hat{\mathscr{H}}$ should contain functions that are even with respect to the time-reversal operation and transform in accordance with representations contained in the antisymmetrical product

$$\{(A + B)^2\} = \{A^2\} + \{B^2\} + AB, \tag{A.1}$$

and odd functions that transform in accordance with representations contained in the symmetrical product

$$[(A + B)^2] = [A^2] + [B^2] + AB. \tag{A.2}$$

For the single-valued representations A and B, the even functions should, to the contrary, be contained in the symmetrical product and the odd ones in the antisymmetrical. Consequently, in both cases the operator should contain even and odd interband terms transforming in accordance with the representations contained in AB. Since the group L includes the inversion operation, the interband terms can contain components P_i if one of the representations AB is even and the other odd. The components P_z transform in accord with L_2^-, and P_x and P_y transform in accord with L_3^-; recognizing that

$$L_6^+ L_6^- = L_1^- + L_2^- + L_3^-,\ \tilde{L}_5^+ \tilde{L}_5^- = \tilde{L}_5^\pm L_6^\mp = 2L_3^-. \tag{A.3}$$

we see that the two-band Hamiltonian can contain all three components P_i only for the representations L_6^+ and L_6^-, whereas in the remaining cases, regardless of the arrangement of the other bands, it contains only P_x and P_y.

The operator $\hat{\mathscr{H}}$ can be readily constructed in accordance with the rules indicated in Sec. 2 of [23], if account is taken of the fact that the representation $L_6^\pm$ can be obtained from $L_1^\pm$. In accord with (A.1) and (A.2), the interband parts of the Hamiltonian can include even and odd functions of **P**, **σ**, and **H**, which transform in accordance with the representations L_1^-, i.e., $\sigma_z P_z$ and $(\sigma_x P_x + \sigma_y P_y)$, and the intraband parts have only even terms that transform in accord with L_1^+, i.e., $\sigma_z H_z$ and $(\sigma_x H_x + \sigma_y H_y)$. The matrices that transform in accordance with these representations can be chosen respectively in the form ρ_1, ρ_2, I, and ρ_3. Then the Hamiltonian $\hat{\mathscr{H}}$ in accordance with the rules indicated in Sec. 4, Item 4 of [23], can be written in the form

$$\hat{\mathscr{H}} = ms^2\rho_3 + \rho_1 s \sum_i \left(\frac{m}{m_i}\right)^{1/2} (\hat{\sigma}_i P_i)$$

$$+ \frac{1}{2}\sum_i [(\mu_i'' + \mu_i') + \rho_3(\mu_i'' - \mu_i')]\hat{\sigma}_i H_i. \tag{A.4}$$

We have used here the notation corresponding to (2). Here $m_x = m_y = m_\perp$, $m_z = m_\parallel$, and $m = (m_\parallel m_\perp^2)^{1/2}$, just as for the magnetic components in the two bands we have $\mu_x' = \mu_y'$ and $\mu_x'' = \mu_y''$.

It can be shown that for the representations $L_5^\pm$ and $L_6^\mp$

$$\hat{\mathscr{H}} = ms^2\rho_3 + \rho_4 s[\alpha(\sigma_x P_x + \sigma_y P_y) + \beta(P_x + i\sigma_z P_z)] +$$

$$+\frac{1}{2}\sum_i [(\mu_i'' + \mu_i') + \rho_3(\mu_i'' - \mu_i')]\sigma_i H_i. \qquad (A.5)$$

In accord with (A.3), k_x and k_y enter in $\hat{\mathscr{H}}$ twice with two independent constants α and β. We note that the matrices ρ_i in (A.4) and (A.5) are not 4×4 but 2×2 matrices, i.e., Pauli matrices, and $\hat{\sigma}$ is the spin operator. If we now choose the basis functions in the form of a product of the coordinate and spin functions, then in the corresponding representation we obtain the 4×4 matrices ρ_i and $\rho_i\sigma_k$, which coincide with the matrices in Eq. (2).

We now change over in Eq. (A.4) to the new variables

$$x_i' = x_i\left(\frac{m_i}{m}\right)^{1/2}, \quad P_i' = P_i\left(\frac{m}{m_i}\right)^{1/2},$$

$$E_i' = E_i\left(\frac{m}{m_i}\right)^{1/2}. \qquad (A.6)$$

We also introduce the notation $A_i' = (s/c)A_i(m/m_i)^{1/2}$, and then*

$$H_x' = \mathrm{rot}_{x'} A' = \frac{s}{c}\mathrm{rot}_x A\left(\frac{m^2}{m_y m_z}\right)^{1/2} = \frac{s}{c}H_x\left(\frac{m_x}{m}\right)^{1/2},$$

i.e.,

$$H_i' = \frac{s}{c}H_i\left(\frac{m_i}{m}\right)^{1/2}. \qquad (A.7)$$

Since the Jacobian of the transformation (A.6) is $\partial(x', y', z')/\partial(x, y, z) = 1$, the volume in the old and in the new coordinate systems remains the same. Here $(\mathbf{E}'\cdot\mathbf{H}') = (s/c)(\mathbf{E}\cdot\mathbf{H})$, i.e., the vectors $\mathbf{E}'$ and $\mathbf{H}'$ remain orthogonal. In the new coordinates (A.4) differs from (2) only in the presence of additional terms $\sigma_i\mu_i' H_i'$. If the additional spin magnetic moment described by these terms were to vanish, then the corresponding g-factors would be connected with m_i by the relations $g_{\|} = 2(m_0/m_{\perp})$ and $g_{\perp} = 2(m_0^2/m_{\|}m_{\perp})^{1/2}$, where m_0 is the mass of the free electron, and the spin splitting would equal the cyclotron splitting, as is the case for free electrons.

According to [18], the spin splitting amounts in fact for PbSe to 0.5 and for PbS to 0.7 of the cyclotron splitting, whereas for PbTe the two are apparently equal.

It is easy to see that the additional magnetic moment described by the second term in (A.4) leads to an additional difference in the transition energies, of the order of $\pm e\hbar H/mc = \pm e\hbar H'/ms$, which in accord with (26)–(27) leads to the appearance in the expression for the current of an additional factor of the order of

$$\mathrm{ch}\left(\frac{q}{eE'}\frac{e\hbar H'}{ms}\right) \approx \cosh\left(\frac{H'}{E'}\right).$$

When $H' \ll E$, this factor is generally insignificant, and when H' is comparable with E' allowance for this factor in accord with (34) makes a contribution which is $\pi m^2 s^3/e\hbar E'$ times smaller than the term quadratic in H' in (34) or (35). We note also that allowance for the influence of the near-lying bands, the distances to which in PbSe and PbS are comparable with the width of the forbidden band, can lead to the appearance of additional factors of the same order as the additional spin splitting.

*rot x ≡ curl_x.

APPENDIX B

INFLUENCE OF DEFORMATIONS ON THE TUNNEL CURRENT IN A MAGNETIC FIELD

As is well known, the tunnel current is noticeably altered by deformation.[24] The main cause of the change in current is the change in the width of the forbidden band and the corresponding changes in the effective mass.

To take into account the influence of the deformation in the operator in (A.4) it is necessary to include terms that transform in accord with the representation L_1^+, i.e., ϵ_{zz} and $\epsilon_{xx} + \epsilon_{yy}$:

$$\hat{\mathscr{H}}_\varepsilon = \frac{1}{2}\left[\sum_i (C_{ii}'' - C_{ii}')\varepsilon_{ii} + \rho_3(C_{ii}'' + C_{ii}')\varepsilon_{ii}\right]. \qquad B.1)$$

Here C_{ii}' and C_{ii}'' are the constants of the deformation potential of the valence band and of the conduction band. In this case $C_{xx} = C_{yy}$. From this we see that the deformation leads only to renormalization of the reduced mass, i.e., to a change in m by an amount

$$\Delta m = \frac{1}{2s^2}\sum_i (C_{ii}'' + C_{ii}')\varepsilon_{ii}. \qquad (B.2)$$

The ratio (m_i/m) is not altered thereby. Accordingly, in agreement with (34) and (35), for sufficiently small deformations, we have

$$\Delta j/J = -\frac{\pi ms}{e\hbar(E'^2 - H'^2)}\sum_i (C_{ii}'' + C_{ii}')\varepsilon_{ii}, \qquad (B.3)$$

and when $H' \ll E'$

$$\Delta\left(\frac{j(H)}{j_0}\right) = -\frac{\pi msH'^2}{2e\hbar E'^3}\sum_i (C_{ii}'' + C_{ii}')\varepsilon_{ii}. \qquad (B.4)$$

Note added in proof (7 June 1966). We have recently learned of a paper [25] whose author also concluded the existence in crossed fields of two regions of solutions of the Schrödinger equation for the Bloch electron, depending on the ratio of the electric and magnetic fields, viz, a region with continuous energy spectrum and a region with discrete spectrum corresponding to Landau quantization.

1 A. R. Calawa, R. H. Rediker, B. Lax, and A. L. McWhorter, Phys. Rev. Lett. **5**, 55 (1960).

2 A. G. Chynoweth, R. A. Logan, and P. A. Wolff, Phys. Rev. Lett. **5**, 548 (1960).

3 R. H. Rediker and A. R. Calawa, J. Appl. Phys. Suppl. **32**, 2189 (1961).

4 L. Esaki and R. R. Haering, J. Appl. Phys. **33**, 2106 (1962).

5 P. H. Butcher, J. A. Hulbert, and K. F. Hulme, J. Phys. Chem. Solids **21**, 320 (1961).

6 B. Lax, Proc. of the Int. Conf. on the Physics of Semiconductors, Prague (1960), p. 331.

7 R. R. Haering and E. N. Adams, J. Phys. Chem. Solids **19**, 8 (1961).

8 R. R. Haering and P. B. Miller, Phys. Rev. Lett. **6**, 269 (1961).

9 P. N. Argyres, Phys. Rev. **126**, 1386 (1962).

10 L. V. Keldysh, JETP **33**, 994 (1957), Soviet Phys. JETP **6**, 763 (1958).

11 W. Franz, Z. Naturforsch. **14a**, 415 (1959).

12 E. O. Kane, J. Phys. Chem. Solids **12**, 181 (1959).

13 L. V. Keldysh, JETP **45**, 364 (1963), Soviet Phys. JETP **18**, 253 (1964).

14 L. L. Foldy and S. A. Wouthuysen, Phys. Rev. **78**, 29 (1950); see also A. S. Davydov, Kvantovaya mekhanika (Quantum Mechanics), Nauka, 1964, Sec. 67.

15 E. O. Kane, J. Phys. Chem. Solids **1**, 249 (1956).

16 E. O. Kane, J. Appl. Phys. **32**, 83 (1961).

17 G. E. Pikus, Osnovy teorii poluprovodnikovykh priborov (Principles of the Theory of Semiconductor Devices), Nauka, 1965, Sec. 32.

18 D. L. Mitchell, E. D. Palik, and J. N. Zemel, Proc. of the 7th Int. Conf. on the Physics of Semiconductors, Paris (1964), p. 325. K. F. Cuff, M. R. Ellett, C. D. Kuglin, and R. Williams, ibid., p. 677. D. L. Mitchell, E. D. Palik, and J. N. Zemel, Phys. Rev. **135**, A763 (1964). R. Nii, J. Phys. Soc. Japan **19**, 58 (1964).

19 R. N. Hall, J. H. Racette, and H. Ehrenreich, Phys. Rev. Lett. **4**, 456 (1960).

20 I. M. Lifshitz and M. I. Kaganov, UFN **69**, 419 (1959), Soviet Phys. Uspekhi **2**, 831 (1960). R. Kubo and H. Hasegawa, and N. Hashitsume, J. Phys. Soc. Japan **14**, 56 (1959). Yu. A. Firsov, ZhTF **28**, 1129 (1958), Soviet Phys. Tech. Phys. **3**, 1051 (1959).

21 G. E. Pikus and G. L. Bir, FTT **4**, 2090 (1962), Soviet Phys. Solid State **4**, 1530 (1963).

22 J. B. Conklin, L. E. Johnson, and J. W. Pratt, Phys. Rev. **137**, A1282 (1965). J. O. Dimmock and J. B. Wright, Phys. Rev. **135**, A821 (1964); Proc. of the 7th Int. Conf. on the Physics of Semiconductors, Paris (1964), p. 77. J. W. Pratt and L. J. Fereira, ibid., p. 69. L. Kleinman and P. J. Lin, ibid., p. 63.

23 G. E. Pikus, JETP **41**, 1258 and 1507 (1961), Soviet Phys. JETP **14**, 898 and 1075 (1962).

24 H. Fritzsche and J. J. Tiemann, Phys. Rev. **130**, 617 (1963); Phys. Rev. **132**, 2506 (1963).

25 H. C. Praddaude, Phys. Rev. **140**, A1292 (1965).

Translated by J. G. Adashko
34