

Resonances of interband impact ionization in $\text{Bi}_{1-x}\text{Sb}_x$ semiconductors subjected to quantizing magnetic fields

S. D. Beneslavskii, E. V. Bogdanov, V. M. Manankov, V. L. Popov, and L. S. Fleishman

M. V. Lomonosov State University, Moscow

(Submitted 7 June 1990)

Zh. Eksp. Teor. Fiz. **98**, 1790–1803 (November 1990)

Experimental and theoretical investigations were made of quantum resonances of interband impact ionization in narrow-gap $\text{Bi}_{1-x}\text{Sb}_x$ semiconductors subjected to magnetic fields in which the separation between the Landau subbands $\sim \hbar\Omega$ became comparable with the width of the band gap ε_g . It was found that resonances in $\text{Bi}_{1-x}\text{Sb}_x$ alloys, characterized by nearly quasirelativistic energy spectrum of carriers, were due to lifting (in the applied magnetic field) of the forbiddenness imposed on the intravalley interband Auger transitions by the laws of conservation of energy and momentum. The amplitudes of the resonances and their positions were calculated and determined as a function of ε_g and of the magnetic field orientation. The impact ionization rate began to rise rapidly in fields determined by the condition $N\hbar\Omega = 2\varepsilon_g$, whereas the amplitude g_{max}^N of the resonances decreased rapidly with the number N : $g_{\text{max}}^N \propto 2^{-N}N^{-2}$. The theoretical results were compared with the experimental data.

1. INTRODUCTION

A narrow band gap and small effective masses of charge carriers in $\text{Bi}_{1-x}\text{Sb}_x$ alloys have the effect that even in magnetic fields of the order of $H = 10$ kOe the separation between the Landau levels becomes comparable with the energy gap in the spectrum. We naturally then expect some manifestation of quantum anomalies of the interband transition rates. We shall report theoretical and experimental investigations of resonances of the interband impact ionization rate in longitudinal quantizing magnetic fields.

Resonances of the interband Auger transition rates in a strong magnetic field were first predicted by Takeshima¹ for semiconductors with very different effective masses of electrons and holes. He showed that a series of resonances of the interband Auger recombination rate, associated with oscillations of the density of states at the impact ionization threshold, should be observed in quantizing magnetic fields. A few years later the existence of this effect was confirmed experimentally for narrow-gap $\text{Hg}_{1-x}\text{Cd}_x\text{Te}$ alloys.²

Similar anomalies of the interband impact ionization rate were detected later in $\text{Bi}_{1-x}\text{Sb}_x$ semiconductor alloys.³ The energy spectrum of carriers in these semiconductors is however different from the spectrum of $\text{Hg}_{1-x}\text{Cd}_x\text{Te}$ and that postulated by Takeshima. In the range of antimony concentrations $0.085 < x < 0.17$ the minimum gap ε_g in the spectrum is governed by the extrema L_a and L_s (see Fig. 1 in Ref. 4) and the dispersion law in the vicinity of the L points is a symmetric quasirelativistic type with practically identical electron and hole masses:

$$\varepsilon = \pm \left[\left(\frac{\varepsilon_g}{2} \right)^2 + \varepsilon_g \left(\frac{p_1^2}{2m_1} + \frac{p_2^2}{2m_2} + \frac{p_3^2}{2m_3} \right) \right]^{1/2}, \quad (1)$$

where m_1 , m_2 , and m_3 are the principal values of the tensor of the effective masses near the band edges; p_1 , p_2 , and p_3 are the corresponding projections of the quasimomentum.

A different spectrum leads, as shown below, to substantially different relationships governing resonances of the interband Auger transitions (shown schematically in Fig. 1) in a magnetic field.

The nature of these differences becomes clear even

when we analyze just the laws of conservation governing such transitions. In the case of the energy band structure of the type found in $\text{Hg}_{1-x}\text{Cd}_x\text{Te}$ the intravalley interband transitions are allowed even in the absence of a magnetic field. In the case of the spectrum described by Eq. (1) the intravalley interband transitions occurring in the absence of a magnetic field are forbidden by the laws of conservation of energy and momentum, and breakdown occurs not because of single-valley transitions, but because of the Emtage mechanism⁵ when a carrier from one L valley creates an electron—hole pair in another L valley. This forbiddenness is

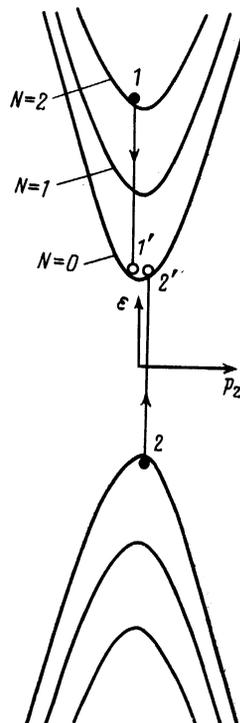


FIG. 1. Interband Auger transitions in a quantizing magnetic field.

lifted by a sufficiently strong magnetic field so that transitions of the type shown in Fig. 1 become allowed.

We shall show that in the case of a quasirelativistic spectrum in a magnetic field the forbiddenness of the intravalley transitions from any N th Landau subband is retained until the bottom of the subband reaches an energy ε_g counting from the bottom of the conduction band. When this happens, the forbiddenness is lifted immediately for transitions from states with any initial value of the momentum. Consequently, the contribution to the Auger ionization rate governed by the intravalley transitions from the N th Landau subband is zero up to a certain critical value of the magnetic field H_N and then it rises rapidly [approximately as $(H - H_N)^{1/2}$]. A further increase of the magnetic field continues to shift the bottom of the relevant Landau subband in the upward direction on the energy scale, and its contribution to impact ionization decreases because of an exponential fall of the populations of the states in the N th subband.

Limitations imposed on the interband transitions by the laws of conservation of energy and momentum are discussed in Sec. 2. Calculations of the matrix elements of the transitions and of the impact ionization rate are reported in Sec. 3 and the amplitude of this rate is described by a simple analytic expression. The results of experimental investigations of semiconducting $\text{Bi}_{1-x}\text{Sb}_x$ alloys are presented in Sec. 4, whereas discussion of the results and a comparison of the theory and experiment are given in Sec. 5.

2. THRESHOLD CONDITIONS FOR THE INTRAVALLEY INTERBAND AUGER TRANSITIONS IN THE CASE OF A QUASIRELATIVISTIC SPECTRUM

We shall consider the impact ionization processes in which an ionizing particle 1, which can be an electron (in view of the symmetry of the spectrum, the analysis for holes is identical), is initially in the N th Landau subband (Fig. 1) and the field is usually directed along the z axis. The lowest energy which this particle must have in order to cause impact ionization (i.e., the impact ionization threshold) is reached in a situation when the final momenta p'_{z1} and p'_{z2} , and the momentum p_{z2} are equal in the absolute sense and directed in such a way as to ensure the minimum absolute value. Since

$$p_{z1} = p'_{z1} + p'_{z2} - p_{z2},$$

this happens when

$$p'_{z1} = p'_{z2} = -p_{z2} = p_{z1}/3. \quad (2)$$

The law of conservation of energy ε for an isotropic ($m_1 = m_2 = m_3 = m$) quasirelativistic spectrum of Eq. (1) in a magnetic field can then be reduced to⁶

$$\left[\left(\frac{\varepsilon_g}{2} \right)^2 + \varepsilon_g \left(N\hbar\Omega + \frac{p_{z1}^2}{2m} \right) \right]^{1/2} = 3 \left[\left(\frac{\varepsilon_g}{2} \right)^2 + \varepsilon_g \frac{(p_{z1}/3)^2}{2m} \right]^{1/2}, \quad (3)$$

where

$$\Omega = eH/mc, \quad (4)$$

e is the electron charge, c is the velocity of light, \hbar is the Planck constant; this law of conservation determines the threshold value of the momentum p_{z1} .

Squaring of Eq. (3) cancels out the terms with p_{z1}^2 ,

which demonstrates that for any value of N there is a range of magnetic fields in which Eq. (3) is disobeyed. This is the forbiddenness imposed on the intravalley Auger transitions in a relativistic spectrum when $H = 0$ by the laws of conservation of energy and momentum and this forbiddenness applies equally in the presence of a magnetic field until the energy gap between the bottom of the lowest ($N = 0$) and of the N th Landau subbands becomes equal to ε_g , which happens when

$$N\hbar\Omega_N = 2\varepsilon_g. \quad (5)$$

However, if the magnetic field satisfies the condition

$$N\hbar\Omega > 2\varepsilon_g,$$

the transitions from the states with any initial value of p_{z1} are allowed.

In the case of an anisotropic spectrum we can show that the form of the condition (3) is retained, but we now have to assume that

$$\Omega = eH/m_{\perp}c, \quad (6)$$

where

$$\frac{1}{m_{\perp}^2} = \frac{n_1^2}{m_2 m_3} + \frac{n_2^2}{m_1 m_3} + \frac{n_3^2}{m_1 m_2},$$

$n_1, n_2,$ and n_3 are the direction cosines of the magnetic field vector relative to the principal axes of the effective mass tensor, whereas m should be replaced with

$$m_{\parallel} = m_1 n_1^2 + m_2 n_2^2 + m_3 n_3^2. \quad (7)$$

The threshold condition of Eq. (5) does not change them, but we have to use the relationship (6) and (4) in order to determine Ω .

It is found that in a magnetic field H_N governed by the condition (5), the transitions are allowed, but the volume of the region of allowed initial momenta is zero. A further increase in H increases this volume in accordance with the law $(H - H_N)^{1/2}$.

We shall consider a transition in which an electron from the N th Landau subband knocks out an electron from the upper hole Landau subband and we shall assume that after this transition both electrons are in the lowest Landau subband of the conduction band (Fig. 1). The laws of conservation of energy and momentum then become

$$\left[\left(\frac{\varepsilon_g}{2} \right)^2 + \varepsilon_g \left(N\hbar\Omega + \frac{p_{z1}^2}{2m_{\parallel}} \right) \right]^{1/2} - \left[\left(\frac{\varepsilon_g}{2} \right)^2 + \varepsilon_g \frac{p_{z2}^2}{2m_{\parallel}} \right]^{1/2} = \left[\left(\frac{\varepsilon_g}{2} \right)^2 + \varepsilon_g \frac{(p'_{z1})^2}{2m_{\parallel}} \right]^{1/2} + \left[\left(\frac{\varepsilon_g}{2} \right)^2 + \varepsilon_g \frac{(p'_{z2})^2}{2m_{\parallel}} \right]^{1/2}, \quad (8)$$

$$p_{z1} + p_{z2} = p'_{z1} + p'_{z2}. \quad (9)$$

When the applied magnetic field exceeds slightly the critical value H_N , i.e., when

$$\Delta\varepsilon = N\hbar\Omega - N\hbar\Omega_N = N\hbar\Omega - 2\varepsilon_g \ll \varepsilon_g,$$

and in the range of small momenta

$$p_i^2 \ll 2m_{\parallel}\varepsilon_g$$

these equations have the following solution:

$$p'_{z1(z2)} = \frac{p_{z1} + p_{z2}}{2} \pm \left[\frac{m_{\parallel}\Delta\varepsilon}{3} - \left(\frac{p_{z1}^2}{12} + \frac{p_{z1}p_{z2}}{2} + \frac{p_{z2}^2}{4} \right) \right]^{1/2}. \quad (10)$$

The allowed initial values of the momentum are those for which the radicand in Eq. (10) is positive. They occupy a region in the form of a strip in the phase plane (p_{z1}, p_{z2}) and this strip is bounded by straight lines

$$p_{z1} = \pm 2(m_{\parallel} \epsilon_g)^{1/2} - 3p_{z2}.$$

The characteristic transverse size of the strip is

$$\Delta p = (0, 4m_{\parallel} \Delta \epsilon)^{1/2}.$$

In the longitudinal direction the strip is in principle unlimited, but only the transitions with the initial momenta not exceeding the thermal values are realized and this sets a certain limit in this direction as well. Consequently, the volume of the region of allowed initial states and therefore the transition rate are both governed by the quantity

$$\Delta p \propto \Delta \epsilon^{1/2} \propto (N \hbar \Omega - N \hbar \Omega_N)^{1/2} \propto (H - H_N)^{1/2}.$$

3. CALCULATION OF THE INTERBAND TRANSITION RATE

The probabilities of transitions between the initial i and the final f states can be described by the usual expression derived using the first order of perturbation theory:

$$P_{if} = \frac{2\pi}{\hbar} |M_{if}|^2 \delta(\epsilon_i - \epsilon_f), \quad (11)$$

where $M_{if} = \langle f | V | i \rangle$ is the matrix element of a transition between the states i and f , whereas the δ function reflects the law of conservation of energy in the investigated transition. The perturbation potential is assumed to be represented by the screened Coulomb interaction:

$$V(r) = \frac{e^2}{\kappa r} \exp\left(-\frac{r}{r_D}\right),$$

where κ is the permittivity and r_D is the Debye screening radius.

The matrix element M_{ij} corresponds to the scattering diagram:



$$(12)$$

Here, $V(q)$ is the Fourier component of the interaction potential

$$V(q) = \frac{4\pi e^2}{\kappa} (q^2 + q_D^2)^{-1},$$

where $q_D = 1/r_D$.

The symbols 1, 2, 1', and 2' in Eq. (12) represent the states in a magnetic field (see the Appendix A) characterized by the following quantum numbers: the momenta p_z and p_x , the Landau subband number (equal to N for a generating particle and 0 for all the others), and the index distinguishing the valence (v) and conduction (c) bands. For the transition shown in Fig. 1 the states 1, 2, 1', and 2' are as follows:

$$\begin{aligned} |1\rangle &= (N, c, p_{z1}, p_{x1}), \\ |2\rangle &= (0, v, p_{z2}, p_{x2}), \\ |1'\rangle &= (0, c, p_{z1}', p_{x1}'), \\ |2'\rangle &= (0, c, p_{z2}', p_{x2}'). \end{aligned}$$

We must bear in mind that in the case under consideration all the Landau subbands, apart from the zeroth, are doubly degenerate. At low momenta, they can be classified approximately in accordance with the spin direction.⁶ The transitions from the states with the same direction of the spin

as in the zeroth subband are much more convenient than from the states with the opposite spin, since the perturbation Hamiltonian (Coulomb interaction) is independent of the spin and cannot give rise to spin-reversal transitions. Therefore, in calculating the interband transition rate it is natural to ignore spin-reversal transitions.

The required impact ionization rate per particle is given by

$$g^N = \frac{2\pi}{\hbar} \int |M_{if}|^2 \delta(\epsilon_i - \epsilon_f) f_N(p_{z1}) \times \frac{dp_{z1} dp_{x1} dp_{z2} dp_{x2} dp_{z1}' dp_{x1}' dp_{z2}' dp_{x2}'}{(2\pi\hbar)^8} \quad (13)$$

i.e., it is governed by the transition probability of Eq. (11) summed over all the occupied initial and vacant final states. It is assumed that

$$\sum_N \frac{1}{(2\pi\hbar)^2} \int f_N(p_{z1}) dp_{z1} dp_{x1} = 1,$$

where $f_N(p_{z1})$ is the distribution function for the N th Landau subband.

The Appendix B gives the full calculation of the ionization rate and shows that in the case of the Maxwellian distribution of the carrier energies with an electron temperature T_e the maximum contribution to the impact ionization rate by carriers from the N th Landau subband is

$$g_{max}^N = \frac{1}{2^N N^2} \frac{\pi^{3/2}}{\sqrt{6}} \frac{e^4 m_{\perp} (m_{\perp}/m_{\parallel})^{1/2}}{\kappa^2 \hbar^3} \left(\frac{\epsilon_g}{kT_e}\right)^{1/2} \exp\left(-\frac{\epsilon_g}{kT_e}\right). \quad (14)$$

Here, k is the Boltzmann constant.

Figure 2 shows graphically the dependence of g^N on the magnetic field. The dependence is strongly asymmetric and the characteristic scales of the fields corresponding to the rise and fall of g^N are

$$\Delta \epsilon = \Delta \hbar \Omega \approx (m_{\perp}/m_{\parallel}) \epsilon_g, \quad (15)$$

$$\Delta \epsilon = \Delta \hbar \Omega \approx 3kT_e. \quad (16)$$

The rise of g^N is due to the process of expansion of the region of the allowed initial states discussed above, whereas the fall is due to the upward shift of the Landau subbands on the energy scale on application of a magnetic field and due to corresponding exponential fall of the population of the investigated subband.

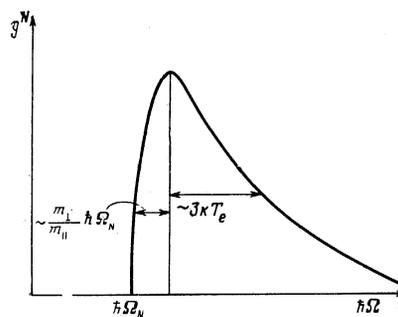


FIG. 2. Schematic representation of the dependence of the contribution of the N th Landau subband to the impact ionization process on the magnetic field (on $\hbar\Omega$).

4. EXPERIMENTAL TECHNIQUES AND RESULTS

We determined the dependences of the impact ionization rate, of the dynamic and static electrical conductivity, and of the resistivity of semiconducting n - and p -type $\text{Bi}_{1-x}\text{Sb}_x$ ($0.09 \leq x \leq 0.17$) alloys on the intensity of a longitudinal magnetic field $H \leq 40$ kOe at $T = 4.2$ K under normal conditions and under pressures p up to 7 kbar. Our samples were cut from bulk single crystals in the basal plane along the bisector C_1 and binary C_2 axes and along intermediate directions. The results of galvanomagnetic measurements carried out in weak electric fields E indicated that the carrier density and the mobility in the investigated alloys, measured at 4.2 K, were $10^{14} - 2 \times 10^{15} \text{ cm}^{-3}$, and $2 \times 10^4 - 6 \times 10^6 \text{ cm}^2 \cdot \text{V}^{-1} \cdot \text{s}^{-1}$, respectively. At these carrier densities the investigated $\text{Bi}_{1-x}\text{Sb}_x$ semiconductors are heavily doped. The antimony content x of the investigated alloys was determined by accurate weighing and with the aid of a microprobe analysis. The results of the latter indicated the absence of inhomogeneities in our samples. The direct energy gap was deduced from the known dependence of ϵ_g on x (Ref. 7) and also from the edges of the photomagnetic effect and of the intrinsic photoconductivity at helium temperatures. Measurements carried out on low-resistivity samples with typical dimensions $0.3 \times 0.5 \times 4$ mm were carried out under the given-current conditions,⁸ whereas in the case of high-resistivity thin samples, with typical dimensions $0.1 \times 0.05 \times 4$ mm, the measurements were carried out under the given-voltage conditions.⁹ Rectangular current voltage pulses of 30–450 ns duration and with a leading edge less than 1 ns were supplied by a mercury-relay generator. The low repetition frequency of the pulses ensured that there was no heating of our samples.

The dependences of the current on the magnetic field typical of $\text{Bi}_{1-x}\text{Sb}_x$ semiconductors were recorded under constant-voltage conditions at different moments of time relative to the beginning of a field pulse (Fig. 3a).

The impact ionization rate g was expressed in terms of the experimentally determined rate of rise of the current I through a sample on application of an electric field pulse causing breakdown:⁹

$$g = \frac{1}{1+b} \left. \frac{d \ln I(t)}{dt} \right|_{t \rightarrow 0, E = \text{const}}, \quad (17)$$

where $b = \mu_h / \mu_e$ is the ratio of the mobilities of electrons and holes (for n -type material) and t is the delay time of the moment of measurement relative to the beginning of a voltage pulse.

Consequently, the rise of the current with time shown in Fig. 3a was evidence of a strongly nonmonotonic dependence of the impact ionization rate on a magnetic field with maxima at $H = 10$ and 21 kOe. The existence of the latter maximum was supported also by the current pulses shown as an inset in Fig. 3a. Clearly, in a field of $H = 21$ kOe the rise of the current representing generation of carriers in the course of breakdown [see Eq. (17)] was much faster than in fields of $H = 17$ and 35 kOe.

The dependences of g on H were calculated on the assumption that $b = 1$ in Eq. (17) for two fixed values of the electric field and the same n -type $\text{Bi}_{0.86}\text{Sb}_{0.14}$ alloy (Fig. 3b). Both curves in this figure had asymmetric maxima of the resonance type at positions which were independent of the

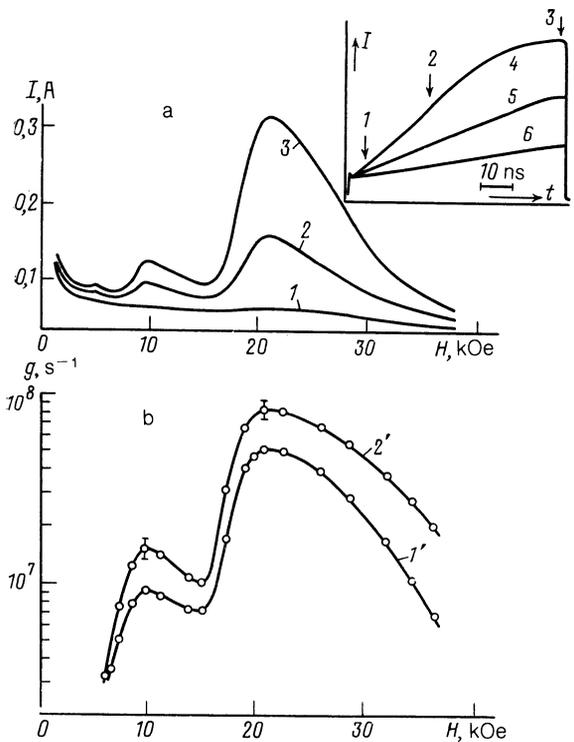


FIG. 3. Magnetic-field dependences of: a) the current I measured at different moments after the beginning of an electric field pulse; b) the impact ionization rate in a sample of n -type $\text{Bi}_{0.86}\text{Sb}_{0.14}$ in the orientation $\mathbf{H} \parallel \mathbf{I} \parallel C_1$ at 4.2 K in electric fields of the following intensities: 1), 2), 3) $E = 40$ V/cm; 1') 45 V/cm; 2') 50 V/cm. The inset shows oscillograms of the current pulses flowing through the same sample in the field $E = 40$ V/cm and in the magnetic fields: 4) $H = 21$ kOe; 5) 17 kOe; 6) 35 kOe. The arrows identify the moments at which the magnetic-field dependences of the current were determined: 1) $t = 5$ ns; 2) 25 ns; 3) 60 ns.

electric field. The impact ionization rate maxima were manifested also by the dependences of the steady-state current, i.e., of the conductivity under the constant-voltage conditions, on the magnetic field recorded using breakdown electric fields and $t > 100$ ns pulses (Fig. 4). In the prebreakdown range of the electric fields there were no oscillations of the conductivity in the same magnetic field. However, when measurements were carried out under the constant-current conditions, the impact ionization rate maxima were manifested by minima of the longitudinal magnetoresistance but

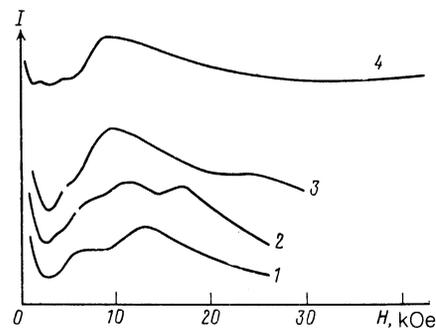


FIG. 4. Magnetic field dependences of the current in p -type $\text{Bi}_{0.88}\text{Sb}_{0.12}$ in breakdown electric fields, recorded in a longitudinal magnetic field for samples with their axis tilted in the basal plane away from the C_1 axis by the following angle: 1) $\varphi = 0^\circ$ ($\mathbf{H} \parallel C_1$); 2) 10° ; 3) 15° ; 4) 30° ($\mathbf{H} \parallel C_2$).

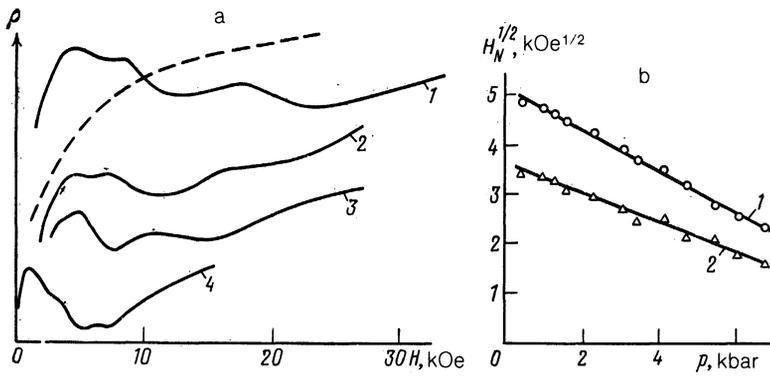


FIG. 5. a) Magnetic-field dependences of the resistivity ρ of a sample of n -type $\text{Bi}_{0.833}\text{Sb}_{0.167}$ ($\mathbf{H}\|\mathbf{I}\|C_1$) obtained using a measuring current of density $j = 5000 \text{ A/cm}^2$ at the following pressures: 1) $p = 0.001 \text{ kbar}$; 2) 1.3 kbar ; 3) 3.1 kbar ; 4) 4.7 kbar . The dashed curve represents the dependence obtained under normal pressure for $j = 10 \text{ A/cm}^2$. b) Dependences of $H^{1/2}$ (line 1) and $H_N^{1/2}$ (line 2) on the pressure applied to the same sample.

only at high values of the current density corresponding to breakdown conditions (Fig. 5a).

Similar results were obtained for all the investigated samples and the magnetic fields H_N corresponding to maxima of g or of the conductivity and minima of the magnetoresistance were independent of the type of conduction and of the carrier density, but fell on reduction in the antimony content in an alloy (compare curve 1 in Figs. 3 and 4) and also under a hydrostatic pressure (Fig. 5), i.e., this happened on reduction in the direct energy gap at the L points of the Brillouin zone. The fields H_N depended also on the orientation. For example, in the $\mathbf{H}\|\mathbf{I}\|C_2$ case they were approximately 1.7 times less than for an alloy with the same composition but in the $\mathbf{H}\|\mathbf{I}\|C_1$ configuration (curves 4 and 5 in Fig. 4), whereas for intermediate orientations the observed oscillation pattern was much more complex (Fig. 4).

5. DISCUSSION OF RESULTS

In an interpretation of the experimental data we must bear in mind that the energy spectrum of charge carriers in the investigated $\text{Bi}_{1-x}\text{Sb}_x$ alloys consisted of three electron or hole valleys at the L points of the Brillouin zone.⁴ The distribution of the relevant constant-energy surfaces is shown in Fig. 6. In the present study we shall be mainly concerned with two orientations of the magnetic field $\mathbf{H}\|C_1$ and $\mathbf{H}\|C_2$. In the $\mathbf{H}\|C_1$ case (Fig. 6) a section of the constant-energy surface A is characterized by the minimum possible mass m_A , whereas the cyclotron masses corresponding to sections B and C are practically equal to $2m_A$. In the case of the $\mathbf{H}\|C_2$ orientation (Fig. 6) the directions of the constant-energy surfaces B and C are characterized by a cyclotron mass $(2/\sqrt{3})m_A$, whereas in the case of the section A the cyclotron mass is approximately an order of magnitude larger.

Introducing the notation

$$\hbar\Omega_0 = \hbar eH / 2m_A c, \quad (18)$$

we find from Eq. (5) that, in the $\mathbf{H}\|C_1$ case, the impact ionization rate singularities should begin to appear for the valley A in fields displaying the following conditions:

$$\hbar\Omega_0 / 2\varepsilon_g = 1/2; 1/4; 1/6; \dots \quad (19)$$

In the case of the valleys B and C , the corresponding conditions are

$$\hbar\Omega_0 / 2\varepsilon_g = 1; 1/2; 1/3; 1/4; \dots \quad (20)$$

and the critical magnetic fields are twice as large. However, a set of resonances with both types of valleys should occur under the following conditions:

$$\hbar\Omega_0 / 2\varepsilon_g = 1; 1/2; 1/3; \dots 1/N; \dots \quad (21)$$

The experimentally observed anomalies occur in fields such that

$$\hbar\Omega_0 / 2\varepsilon_g = 1; 1/2; 1/4. \quad (22)$$

The absence of a peak at $\hbar\Omega_0 / 2\varepsilon_g = 1/3$ can probably be explained by the fact that—according to Eq. (14)—the amplitude of the singularities falls rapidly on increase in the number N . In fact, in the series described by Eq. (21) a peak at $\hbar\Omega_0 / 2\varepsilon_g = 1/3$ corresponds to the third resonance in the valleys B and C , whereas a peak at $\hbar\Omega_0 / 2\varepsilon_g = 1/4$ corresponds only to the second resonance in the valley A . In this case the amplitude of the fourth peak in Eq. (21) may be higher than that of the third peak. Then, in reasonable agreement with the experimental data on the relative amplitude of the resonances, we find from Eq. (14) (ignoring heating in a magnetic field) that

$$g_{\text{max}}^1 : g_{\text{max}}^{1/2} : g_{\text{max}}^{1/4} \approx 1 : 0.3 : 0.02. \quad (23)$$

In the $\mathbf{H}\|C_2$ orientation the effective masses in two

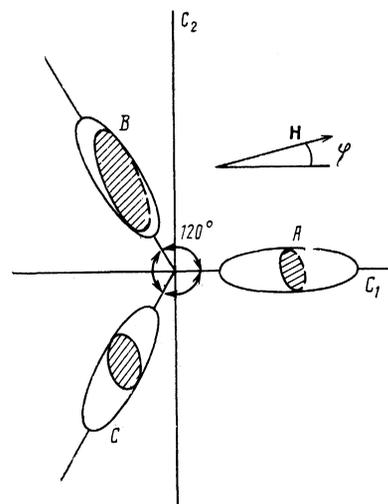


FIG. 6. Schematic diagram showing the distribution of the ellipsoidal constant-energy surfaces of $\text{Bi}_{1-x}\text{Sb}_x$ alloys at the L points of the Brillouin zone.

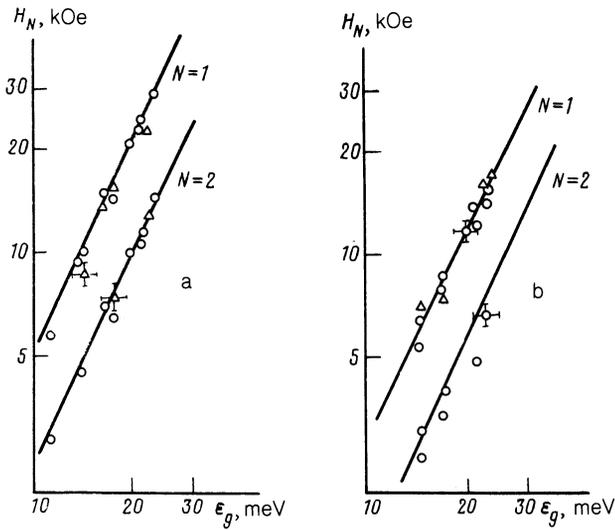


FIG. 7. Dependences of the resonance magnetic field H_N on the direct energy gap ε_g ($H=0$) obtained for the orientations $\mathbf{H} \parallel \mathbf{I} \parallel C_1$ (a) and $\mathbf{H} \parallel \mathbf{I} \parallel C_2$ (b) for n -type (O) and p -type (Δ) $\text{Bi}_{1-x}\text{Sb}_x$ alloys; $T = 4.2$ K. The continuous straight lines are plotted in accordance with Eq. (24) using the known values of the matrix elements¹⁰ and allowing for the change in the gap in a magnetic field¹¹ in the cases when $N = 1$ and $N = 2$.

equivalent valleys are, as pointed out already, $\sqrt{3} \approx 1.7$ times less than in the $\mathbf{H} \parallel C_1$ case. The fields in which the resonances are observed decrease in reality by the same factor (Fig. 4). However, the resonances for the third valley should then be observed in fields an order of magnitude higher, not reached in our experiments.

Apart from a small (approximately 1–3%) correction given by Eq. (15), the magnetic fields H_N at which reson-

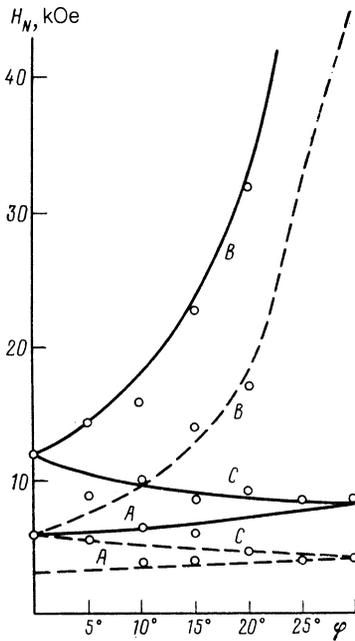


FIG. 8. Angular dependences of the resonance magnetic fields H_N obtained for p -type $\text{Bi}_{0.88}\text{Sb}_{0.12}$ in the basal plane. The continuous and dashed curves are the results of a theoretical calculation for $N = 1$ and $N = 2$, respectively. The points are the experimental values. The relative orientation of the magnetic field and the positions of the constant-energy surfaces are shown in Fig. 6.

ances are observed are governed by the relationship (5). It therefore follows that

$$H_N = \frac{1}{N} \frac{2cm_{\perp}}{e\hbar} \varepsilon_g.$$

In other words, if we bear in mind that m_{\perp} itself rises proportionally to ε_g in accordance with the law

$$m_{\perp} = \varepsilon_g / 2V^2,$$

where V is the interband matrix element of the velocity operator,¹⁰ we obtain

$$H_N = \frac{1}{N} \frac{c}{e\hbar V^2} \varepsilon_g^2, \quad (24)$$

i.e., H_N rises proportionally to the square of the width of the band gap. This is also found experimentally (Fig. 7). It should be pointed out that, because of the “mirror-like” nature of the spectrum, the results for n - and p -type alloys are described by the single dependence given by Eq. (24). Moreover, the observed linear dependence of $H_N^{1/2}$ on the applied pressure (Fig. 5b) follows directly from Eq. (24) if we allow for the familiar¹¹ linear reduction in the direct band gap of $\text{Bi}_{1-x}\text{Sb}_x$ semiconductors under pressure. The value $d\varepsilon_g/dp \approx -1.9$ meV/kbar deduced from the data in Fig. 5b agrees with the published results.¹¹ We can see from Fig. 8 that if we use the known values of the parameters of the energy spectrum,¹⁰ we can describe qualitatively also the angular dependences of the resonance fields. It is true, in the case of the branch B in the range $\varphi \gg 15^\circ$ we have to allow for the deviation of the real spectrum of $\text{Bi}_{1-x}\text{Sb}_x$ alloys from the quasirelativistic form and we cannot limit ourselves to the relationship (24).

A rigorous quantitative comparison is difficult to carry out in the case of the impact ionization rate, because we do not know the exact value of T_e . Estimates based on the dependences of the carrier mobility on the electric field in the prebreakdown region indicate that the electron temperature in this region amounts to about 20 K for $\text{Bi}_{1-x}\text{Sb}_x$ (Ref. 12). In the case of $\text{Bi}_{0.86}\text{Sb}_{0.14}$, for which the impact ionization data are plotted in Fig. 3b, the width of the band gap is $\varepsilon_g \approx 21.5$ meV. Using these values, as well as $\kappa \approx 100$, $m_{\parallel} \approx m_0$ and $m_{\perp} \approx 0.01m_0$ (m_0 is the mass of a free electron) we find that a calculation based on Eq. (14) predicts that the amplitude of the first peak for this alloy should be of the order of $g_{\text{max}}^1 \approx 10^8 \text{ s}^{-1}$. This is in agreement with the experimental results (Fig. 3b).

The mechanism of the impact ionization resonances considered here also accounts for the considerable asymmetry of the singularities of g (Fig. 3b) which demonstrate a steep rise and a slower fall. True, this asymmetry is always less than that predicted theoretically [according to Eq. (15) the interval of rise of the ionization rate to its maximum value should represent a fraction amounting to $m_{\perp}/m_{\parallel} \approx 1\%$ of the magnetic field at which a resonance is observed].

The authors are deeply grateful to A. D. Belaya, T. M. Lifshitz, and M. G. Shakhtakhtinskii for supplying high-quality samples, and also to V. A. Martynov and V. I. Trifonov for optical measurements.

APPENDIX A

The electron spectrum of $\text{Bi}_{1-x}\text{Sb}_x$ alloys can be described approximately by the two-band Luttinger–Kohn

model.¹³ In the isotropic case the Hamiltonian of this model is described by the following formally relativistic expression,

$$\hat{H} = \begin{pmatrix} \varepsilon_g/2 & V\sigma_{\mathbf{p}} \\ V\sigma_{\mathbf{p}} & -\varepsilon_g/2 \end{pmatrix}, \quad (\text{A1})$$

where $\sigma = (\sigma_x, \sigma_y, \sigma_z)$ is the matrix vector whose components are the Pauli matrices; V is the interband matrix element of the velocity operator; $\mathbf{p} = -i\hbar\partial/\partial\mathbf{r}$ is the momentum operator in the coordinate representation. The Hamiltonian \hat{H} is a 4×4 matrix and it acts in the space of four-dimensional wave functions.

Application of a magnetic field occurs in the usual way replacing \mathbf{p} with $\mathbf{P} = \mathbf{p} - (e/c)\mathbf{A}$, where \mathbf{A} is the vector potential. In the presence of a homogeneous magnetic field directed along the z axis, we can use the gauge

$$\mathbf{A} = (-yH, 0, 0). \quad (\text{A2})$$

The states which are then obtained are doubly degenerate, with the exception of $N = 0$, and the eigenfunctions of the Hamiltonian are given by

$$\Psi_I^{(N)} = A_N(p_z) \begin{pmatrix} \frac{(\varepsilon_g N \hbar \Omega)^{1/2}}{\varepsilon_N - \varepsilon_g/2} \varphi_N \\ -\frac{V p_z}{\varepsilon_N - \varepsilon_g/2} \varphi_{N-1} \\ 0 \\ \varphi_{N-1} \end{pmatrix} \exp(ik_z z + ik_x x) \quad (\text{A3})$$

and by similar expression for $\Psi_{II}^{(N)}$ where only the last two terms in the column are altered: $0 \rightarrow \varphi_N$; $\varphi_{N-1} \rightarrow 0$. Here, $k_i = p_i/\hbar$ are the wave vectors and $N = 0, 1, 2, \dots$. We then have

$$\varphi_N = \Phi_N(y - \lambda^2 k_x), \\ \Phi_N(x) = \left(\frac{1}{\pi \lambda^2}\right)^{1/2} \frac{1}{(2^N N!)^{1/2}} \exp\left(-\frac{x^2}{2\lambda^2}\right) P_N\left(\frac{x}{\lambda}\right), \\ \varphi_{N-1} = 0 \text{ when } N=0.$$

We then have $\lambda^2 = \hbar c/eH$ which is the square of the magnetic length, P_N represents the Hermite polynomials, and

$$A_N(p_z) = \frac{(\varepsilon_N - \varepsilon_g/2)^2}{(\varepsilon_N + \varepsilon_g/2) + \varepsilon_g(N\hbar\Omega + p_z^2/2m)}. \quad (\text{A4})$$

The corresponding eigenvalues of the Hamiltonian in a magnetic field are

$$\varepsilon_N = \pm \left[\left(\frac{\varepsilon_g}{2}\right)^2 + \varepsilon_g \left(N\hbar\Omega + \frac{p_z^2}{2m} \right) \right]^{1/2}. \quad (\text{A5})$$

Here the upper sign applies to the conduction band and the lower one to the valence band.

Quantization of an anisotropic spectrum of the quasirelativistic type involves a similar procedure.¹⁴ The expressions for the wave functions given by Eq. (A3) with a dispersion law given by Eq. (A5) remain the same, but we have to replace Eq. (6) with Eq. (7).

APPENDIX B

The diagram in Eq. (12) represents the matrix element of the transition M_{ij} , given by

$$M_{ij} = A_N(p_{z1}) A_0(p_{z1}') A_0(p_{z2}) A_0(p_{z2}') R(p_{z1}, p_{z1}', p_{z2}, p_{z2}')$$

$$\times \int \Phi_N(y_1 - \lambda^2 k_{x1}) \Phi_0(y_1 - \lambda^2 k_{x1}') \Phi_0(y_2 - \lambda^2 k_{x2}) \Phi_0(y_2 - \lambda^2 k_{x2}') \\ \times \exp\{i[(k_{x1} - k_{x1}')x_1 + (k_{z1} - k_{z1}')z_1 + (k_{x2} - k_{x2}')x_2 + (k_{z2} - k_{z2}')z_2]\} \\ \times V(\mathbf{r}_1 - \mathbf{r}_2) d\mathbf{r}_1 d\mathbf{r}_2. \quad (\text{B1})$$

Here, the quantities A_N and F_N are defined in the Appendix A and we also have

$$R(p_{z1}, p_{z1}', p_{z2}, p_{z2}') = \frac{\varepsilon_g(\varepsilon_g N \hbar \Omega)^{1/2}}{2m_{\parallel}[\varepsilon_N(p_{z1}) - \varepsilon_g/2]} \\ \times \frac{\hbar k_{z1}'}{\varepsilon_0(p_{z1}') - \varepsilon_g/2} \left[\frac{\hbar k_{z2}'}{\varepsilon_0(p_{z2}') - \varepsilon_g/2} - \frac{\hbar k_{z2}}{\varepsilon_0(p_{z2}) - \varepsilon_g/2} \right].$$

Integration with respect to spatial variables $d\mathbf{r}_1 d\mathbf{r}_2$ in Eq. (B1) presents no fundamental difficulties. After calculation of M_{ij} it has to be substituted in Eq. (13) and all the integrations indicated there have to be carried out (it is then convenient to go over from momenta to wave vectors in accordance with the substitution $dp_i/2\pi\hbar \rightarrow dk_i/2\pi$).

In the course of integration with respect to $dk'_{x1} dk'_{x2} dk_{x2}$, we obtain:¹⁴

$$\int |M_{ij}|^2 dk'_{x1} dk'_{x2} dk_{x2} = A_N^2(k_{z1}) A_0^2(k_{z1}') A_0^2(k_{z2}) A_0^2(k_{z2}') \\ \times R^2 \delta(k_z) \frac{4\pi^2}{2^N N! \lambda^4} \int_0^{\infty} \eta^{2N+1} V^2 \left[\left(\frac{\eta}{\lambda}\right)^2 + \Delta k_z \right] \exp(-\eta^2) d\eta, \quad (\text{B2})$$

where

$$\delta(k_z) = \delta(k_{z1} + k_{z2} - k_{z1}' - k_{z2}'), \\ \Delta k_z = k_{z1} - k_{z1}' = k_{z2} - k_{z2}'.$$

At low values of k'_{z1} , k_{z2} , k'_{z2} the factor in front of the δ function in Eq. (B2) is readily calculated:

$$A_N^2(k_{z1}) A_0^2(k_{z1}') A_0^2(k_{z2}) A_0^2(k_{z2}') R^2 \approx \frac{\hbar^2 \Delta k_z}{8m_{\parallel} \varepsilon_g}.$$

Further integration with respect to dk_{z2} eliminates $\delta(k_z)$ from Eq. (B2), and integration with respect to dk_{z2} eliminates also $\delta(\varepsilon_i - \varepsilon_f)$ [the integrand is then multiplied by $|\partial(\varepsilon_i - \varepsilon_f)/\partial k_{z2}|^{-1} = (m_{\parallel}/\hbar^2) |k_{z1} - k'_{z1}|^{-1}$]. Integration with respect to $dk_{z1} dk'_{z1}$ can be transformed, with its substitution of variables, to integration with respect to k_{z1} and Δk_z .

Consequently, the number of pairs created from the N -th Landau subband, and per unit time by electrons and per generating particle is given by

$$g^N = \frac{4e^4}{3\kappa^2 \varepsilon_g} \frac{1}{2^N N! \lambda^4} \int \frac{\eta^{2N+1} \exp(-\eta^2) (\Delta k_z)^2}{[\eta^2/\lambda^2 + (\Delta k_z)^2]^2} \\ \times \frac{f_N(k_{z1}) dk_{z1} dk_{z1}' d(\Delta k_z) d\eta}{\hbar \Delta k_z/2 + \{m_{\parallel} \Delta \varepsilon - \hbar^2 [(4/3)^{1/2} k_{z1} + (3/4)^{1/2} \Delta k_z]^2\}^{1/2}}. \quad (\text{B3})$$

Since the integrand in Eq. (B3) is independent of k_{x1} , the relevant integration can generally be avoided and the distribution function can be regarded as normalized by the condition

$$\sum_N \frac{1}{2\pi} \int f_N(k_{z1}) dk_{z1} = 1. \quad (\text{B4})$$

Allowing for the strong anisotropy of the constant-en-

ergy surfaces (which is true of bismuth-antimony alloys), we find that in a narrow interval near the ionization threshold $\Delta\varepsilon \approx (m_{\perp}/m_{\parallel})\varepsilon_g \ll \varepsilon_g$ we can simplify Eq. (B3) by ignoring the expression in the radicand. Then, after integration the impact ionization rate is found, as already pointed out, to rise from zero in accordance with the law $g^N \propto (\Delta\varepsilon)^{1/2}$. In the opposite limiting case when $\Delta\varepsilon \gg (m_{\perp}/m_{\parallel})\varepsilon_g$ we can, on the contrary, ignore the dependence of the denominator of the second fraction in the integrand in Eq. (B3) on Δk_z . Then, integration with respect to $d(\Delta k_z)$ gives

$$\int_{-\infty}^{\infty} \frac{(\Delta k_z)^2 d(\Delta k_z)}{[\eta^2/\lambda^2 + (\Delta k_z)^2]^2} = \frac{\pi\lambda}{2\eta}.$$

The integral with respect to $d\eta$ is then calculated quite simply

$$\int_0^{\infty} \eta^{2N} \exp(-\eta^2) d\eta \approx \frac{N!}{2(2N)^{1/2}}$$

and the impact ionization rate is given by the following final expression:

$$g^N = \frac{1}{2^N N^{1/2} \lambda^3 4 \cdot 3^{1/2} \kappa^2 \varepsilon_g (km_{\parallel} T_e)^{1/2}} \exp\left(-\frac{3\varepsilon_g + \Delta\varepsilon}{3kT_e}\right). \quad (\text{B5})$$

It therefore follows that in the case of a strongly anisotropic spectrum the dependence of g^N on $\Delta\varepsilon$ [and, consequently, on the magnetic field (Fig. 2)] is characterized by two scales: it rises steeply to a maximum in the interval $\Delta\varepsilon \approx (m_{\perp}/m_{\parallel})\varepsilon_g$ [Eq. (15)] and falls more smoothly in the interval $\Delta\varepsilon \approx 3kT_e$ [Eq. (16)].

If in Eq. (B5) we substitute $\Delta\varepsilon = 0$ and the value of λ_N , which corresponds to the N th peak

$$\lambda_N^{-2} = \frac{m_{\perp}}{\hbar^2} \frac{2\varepsilon_g}{N},$$

we obtain Eq. (14) for the maximum of g^N .

- ¹ M. Takeshima, J. Appl. Phys. **44**, 4717 (1973).
- ² R. Dornhaus, K.-H. Müller, G. Nimtz, and M. Schifferdecker, Phys. Rev. Lett. **37**, 710 (1976).
- ³ E. V. Bogdanov, N. B. Brandt, V. M. Manankov, and L. S. Fleishman, Pis'ma Zh. Eksp. Teor. Fiz. **35**, 75 (1982) [JETP Lett. **35**, 88 (1982)].
- ⁴ N. B. Brandt, S. M. Chudinov, and V. G. Karavaev, Zh. Eksp. Teor. Fiz. **61**, 689 (1971) [Sov. Phys. JETP **34**, 368 (1972)].
- ⁵ P. R. Emtage, J. Appl. Phys. **47**, 2565 (1976).
- ⁶ V. B. Berestetskii, E. M. Lifshitz, and L. P. Pitaevskii, *Quantum Electrodynamics*, 2nd ed., Pergamon Press, Oxford (1982).
- ⁷ G. A. Mironova, M. V. Sudakova, and Ya. G. Ponomarev, Fiz. Tverd. Tela (Leningrad) **22**, 3628 (1980) [Sov. Phys. Solid State **22**, 2124 (1980)].
- ⁸ N. B. Brandt, E. V. Bogdanov, V. M. Manankov, and G. D. Yakovlev, Fiz. Tekh. Poluprovodn. **15**, 813 (1981) [Sov. Phys. Semicond. **15**, 465 (1981)].
- ⁹ J. C. McGroddy and M. I. Nathan, Proc. Eighth Intern. Conf. on Physics of Semiconductors, Kyoto, 1966, in J. Phys. Soc. Jpn. **21**, Suppl., 437 (1966).
- ¹⁰ G. A. Mironova, M. V. Sudakova, and Ya. G. Ponomarev, Zh. Eksp. Teor. Fiz. **78**, 1830 (1980) [Sov. Phys. JETP **51**, 918 (1980)].
- ¹¹ N. B. Brandt, S. M. Chudinov, and V. G. Karavaev, Zh. Eksp. Teor. Fiz. **70**, 2296 (1976) [Sov. Phys. JETP **43**, 1198 (1976)].
- ¹² V. L. Popov, Fiz. Tverd. Tela (Leningrad) **25**, 2127 (1983) [Sov. Phys. Solid State **25**, 1223 (1983)].
- ¹³ J. M. Luttinger and W. Kohn, Phys. Rev. **97**, 869 (1955).
- ¹⁴ V. L. Popov, Thesis for Candidate's Degree [in Russian], Moscow State University (1985).

Translated by A. Tybulewicz