546.87+538.1

PHYSICAL PROPERTIES OF BISMUTH

L. A. FAL'KOVSKIĬ

Institute of Theoretical Physics, U.S.S.R. Academy of Sciences

Usp. Fiz. Nauk **94**, 3−41 (January, 1968)

CONTENTS

INTRODUCTION

WE consider in this review the properties of solid bismuth. Bismuth has been the subject of a large number of experimental and theoretical papers. This is principally due to the fact that it is easiest to observe in bismuth the phenomena that are inherent in all metals. However, bismuth is frequently called a semimetal since it occupies from the point of view of the electronic properties, a position intermediate between a metal and a semiconductor. The number of conduction electrons in it is 10^{-5} per atom, their effective mass if of the order of 0.1−0.001 of the electron mass, and the Fermi energy is several hundredths of an electron volt.

We shall not try to describe everything known concerning bismuth* and will confine ourselves essentially to those phenomena which have been quantitatively explained. Thus, for example, we barely touch upon the temperature dependence of the different kinetic coeffi-

_____
*Reference to certain papers not mentioned here can be found in the review articles [1,2].

1

cients. An elucidation of this dependence calls for a de-
tailed analysis of the mechanism of interaction between
the electrons and thermal lattice vibrations. At the
same time, it is necessary to take into account the
small value of the degeneracy temperature (on the order
of 100°K), the increase of the number of carriers as a
result of the valence band,[3] (the forbidden gap is also
of the order of 100°K) and the transitions of the elec-
trons from one part of the Fermi surface to another (in-
tervalley transitions). No corresponding calculations
have been performed as yet. For this reason we are re-
ferring here essentially to low temperatures. At such
temperatures the electrons in the bismuth constitute a
degenerate Fermi gas.

We do not consider here the important practical
question of the methods used to obtain pure bismuth.
These methods are common to all metals (see, for ex-
ample,[11]). However, the requirements with respect to
the purity of the bismuth are more stringent. Inasmuch
as the number of conduction electrons in bismuth is
small, their number changes noticeably when a small
number of impurities is introduced. The state of the art
of obtaining pure metals is at present such that experi-
menters in different countries have samples which
hardly differ in their properties.

I. SPACE LATTICE

The space lattice of bismuth is of the rhombohedral
system with two atoms per cell. It can be obtained from
the primitive cubic lattice by small displacements of
the atoms. We denote the periods of the primitive cubic
lattice by a_i^0 (Fig. 1). We separate two face-centered
sublattices with periods

$$a_1 = a_2^0 + a_3^0, \quad a_2 = a_3^0 + a_1^0,$$
$$a_3 = a_1^0 + a_2^0.$$

In the figure the atoms of one of the sublattices are
marked by crosses. We now displace the two sublattices
relative to each other along the space diagonal of the
cube $a_1^0 + a_2^0 + a_3^0$. The symmetry of the lattice is re-
duced—it becomes rhombohedral. But the most import-
ant fact is that there are now two atoms per unit cell.
In order to obtain the bismuth lattice, it remains to
stretch each sublattice slightly along the same diagonal.
The angle between the vectors a_1 and a_2, which together
with a_3 (they are shown by solid lines) are the periods
of the bismuth lattice, turns out to be 57° 19'.[4] In a
cubic lattice this angle is 60°. The value of a_i is
4.73 Å. The point symmetry group of the bismuth lattice

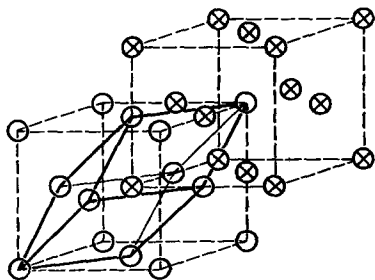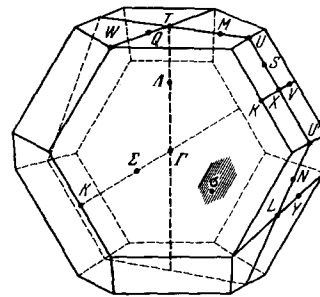

FIG. 1. Space lattice.

FIG. 2. First Brillouin zone.

D_{3d} includes a sixfold mirror-rotation axis, usually
called the trigonal axis and denoted C_3, three twofold
axes, (binary axis C_2) perpendicular to the trigonal axis,
and three symmetry planes perpendicular to the binary
axis (bisector planes). The origin is usually placed at
the inversion center—on the trigonal axis halfway be-
tween the two atoms of the cell, whose coordinates with
respect to the translational periods are (u, u, u) and
(ü, ü, ü). In bismuth u = 0.234 and in a cube u = 0.25. It
can be noted that it is possible to obtain the lattices of
arsenic and antimony in the same manner.

II. DOUBLING OF THE PERIOD

We shall henceforth frequently use the concepts of
band theory. In band theory the interaction of each elec-
tron with the remaining ones is replaced by a certain
average field. This field has the periodicity of the space
lattice. The wave functions of the electron, the well-
known Bloch functions

$$\psi_{n\mathbf{k}}(\mathbf{r}) = u_{n\mathbf{k}}(\mathbf{r}) e^{i\mathbf{k}\mathbf{r}}, \tag{1}$$

are numbered by integer indices—the number of the
band n—and by a continuous three-dimensional index
—the quasimomentum $\mathbf{k}(k_x, k_y, k_z)$ (see, for example,
[5]). The functions $u_{n\mathbf{k}}$ are periodic in $\mathbf{r}$ with a lat-
tice period a_i. The Bloch functions with $\mathbf{k}$ and
$\mathbf{k} + 2\pi\mathbf{b}$ have the same translational properties if
$\mathbf{b} \cdot a_i = r$, where r is an integer. We can therefore as-
sume that the vectors $\mathbf{k}$ lie within a certain bounded
region. The choice of this region is not unique. Thus,
for example, it is possible to construct a parallele-
piped of the vectors $2\pi\mathbf{b}_i$ defined in the following
manner:

$$b_1 = \frac{[a_2 a_3]}{a_1 a_2 a_3}, \quad b_2 = \frac{[a_3 a_1]}{a_1 a_2 a_3}, \quad b_3 = \frac{[a_1 a_2]}{a_1 a_2 a_3}. \tag{2}*$$

This parallelepiped is called the unit cell of the recip-
rocal lattice. The term "first Brillouin zone" is more
frequently used. In order to obtain it it is necessary to
take the vectors $2\pi\mathbf{b}_i$ and $-2\pi\mathbf{b}_i$, and also all possible
sums of these vectors, and draw a perpendicular plane
through the middle of each of these vectors. The smal-
lest polyhedron bounded by these planes is the first
Brillouin zone. Figure 2 shows the first Brillouin zone
of bismuth.

The energy levels of the electrons $\varepsilon_n(\mathbf{k})$ depend on
the quantum numbers n and $\mathbf{k}$, with $\varepsilon_n(\mathbf{k}) = \varepsilon_n(\mathbf{k} + 2\mathbf{b})$.

---

*$[a_2 a_3] \equiv a_2 \times a_3$.

In a crystal, just as in an atom, the spin is not a "good" quantum number, owing to the presence of the spin-orbit coupling. However, in a crystal with an inversion center in the absence of a magnetic field all the levels are doubly degenerate,[6] just as the levels of the free electrons are spin-degenerate. The corresponding quantum number is called for brevity spin, as before (a deeper justification can be found in the transformation properties of the wave functions, which can be described, as before, with the aid of Pauli matrices).

At zero temperature, all the levels, up to the Fermi level, are occupied, and there can be only one electron in each state, by virtue of the Pauli principle. An exceedingly important role in the theory of metals is played by the Fermi surface, defined by the equation $\varepsilon(\mathbf{k}) = \varepsilon_F$.

The main difficulty of the band theory is that the interaction of the electrons cannot be described in terms of an average field. The concept of individual quasiparticles can therefore be used to describe only weakly excited states that include quasiparticles located in the immediate vicinity of the Fermi surface (the word "quasiparticles" is used to emphasize that the corresponding excitations are quite remotely reminiscent of the initial particles, in this case the electrons). Inasmuch as the lifetime of an electron located far (in energy) from the Fermi surface, is comparable with $\hbar/\varepsilon$ the very classification of the levels used in the band theory loses its meaning. However, the intensity of the interaction of the electrons with one another and with the lattice oscillations (phonons) depends in significant fashion on the dielectric constant. As shown by experiment[7] and confirmed by calculation[8], the dielectric constant of bismuth is quite large in the frequency and wave-vector region under consideration. This makes it possible to regard the electron interaction as small not only near the Fermi surface but in several bands closest to the conduction band, up to energies on the order of several electron volts.

We now turn to the deformation with the aid of which the bismuth lattice can be obtained from the primitive cubic lattice. In the latter, there is one atom unit cell. Each bismuth atom has an odd number of electrons which in this case cannot occupy fully a certain number of the bands, since each band contains an even number of places (it is known that the number of places in the band per cell, with allowance for the double spin degeneracy, is two). Thus, for example, if the bands do not overlap, then half of the places in the conduction band remain empty. Cubic bismuth would certainly be a typical metal. Since the deformation makes the number of electrons per unit cell even, the substance can in principle become a dielectric. If some band contains a certain number of electrons, then an equal number of places should become free in another band—holes appear in the latter. The transition of a good metal into a dielectric or a poor metal can be easily understood by using a one-dimensional model. We consider a linear chain of atoms situated at different distances a from one another. The unit cell of the reciprocal lattice has a dimension $2\pi/a$. If each atom has an odd number of electrons, then half of the places from $-\pi/2a$ to $\pi/2a$ will be occupied in the case of the spectrum shown dashed in Fig. 3. Let us double the translation period, shifting
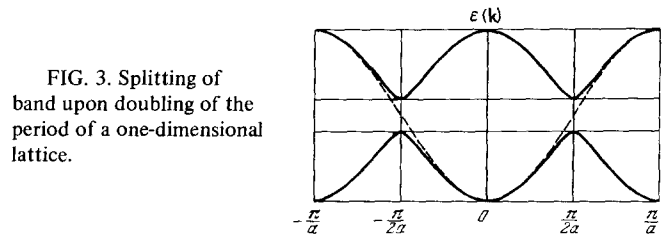

FIG. 3. Splitting of band upon doubling of the period of a one-dimensional lattice.

slightly every other atom to the right. The size of the reciprocal lattice cell will be decreased by one-half and the energy band will split. Using the property $\varepsilon(\mathbf{k}) = \varepsilon(\mathbf{k} + 2\pi b)$ we can transfer the upper band to within the limits of the new cell (Fig. 3, solid line). As a result of the deformation the lower band turns out to be fully occupied, and the other one is free, and we get a dielectric.

In the one-dimensional example, the Fermi surface is actually a point, and an infinitesimally small deformation can transform the metal into a dielectric. In the three-dimensional case, the transition into a dielectric should be accompanied by a significant change in the area of the Fermi surface. It is clear that a small deformation, generally speaking, cannot change this surface greatly. However, as shown in [9], the energy spectrum of bismuth has the following singularity: A slight displacement from the Fermi level of real bismuth along the energy variable ε results in an equal-energy surface whose dimensions are in any case of the order of the dimensions of the reciprocal lattice. This circumstance explains the presence of metallic modifications of solid bismuth. Figure 4, which is taken from [10] (see also [11]) shows the phase diagram. The phases II and III exhibit typical metallic properties. In particular, they reveal superconductivity.[12, 13] Unfortunately, the crystal structure of all phases except ordinary bismuth I is unknown. Phase I is a state which is stable under normal conditions and in which the bismuth is almost a dielectric. It has been established that with the aid of relatively small pressures it is possible to change the number of carriers in it.[14] Under these conditions, when the bismuth lattice comes close to primitive cubic, that is, when $\gamma \sim |0.25 - u| \to 0$, the metallic state is apparently stable. At the point $\gamma = 0$ itself, the "dielectric" phase becomes absolutely unstable.
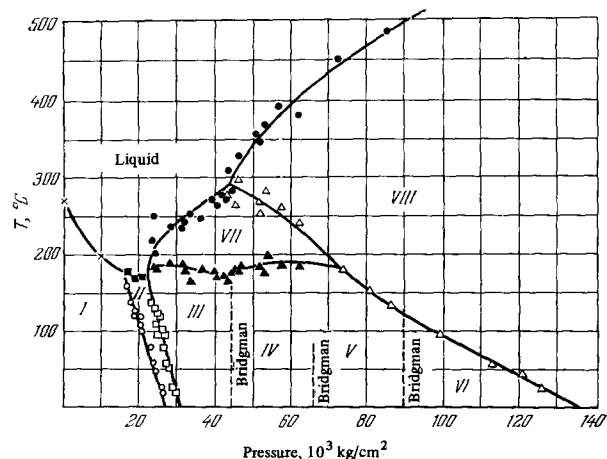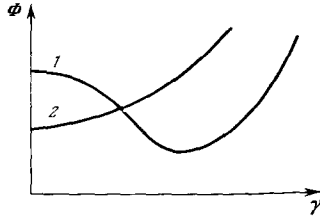

FIG. 4. Phase diagram.

FIG. 5. Dependence of the thermodynamic potentials of the "dielectric" (1) and metallic (2) phases on the deformation γ [9].

If we consider the thermodynamic potentials of the "dielectric" phase I and of the metallic phase (phase II or III) as functions of the deformation γ, then the corresponding curves should have the form shown in Fig. 5. In [9] it is shown that the thermodynamic potential of phase I is determined by the expression

$$\Phi^{(I)} = \Phi_0 + \sigma\gamma^2 - \chi\gamma^2 \ln \rho/\gamma$$

with positive σ, χ, and ρ. Some finite deformation occurs spontaneously if the minimum of $\Phi^{(I)}$ lies below the energy of the metallic phase at $\gamma = 0$.

III. ELECTRON ENERGY SPECTRUM

1. Expansion Near Extremum

The first to investigate the spectrum of bismuth intensively was Schoenberg.[15] He investigated the oscillations of the magnetic moment of bismuth at temperatures of liquid helium (the de Haas–van Alphen effect). Using the calculations of Landau[16] and of Blackman,[17] Shoenberg found that the Fermi surface can be regarded as consisting of three unconnected ellipsoids which go over one into the other by rotation around the trigonal axis. Each ellipsoid is slightly inclined, approximately 6°, to the basal plane (the plane perpendicular to the trigonal axis) and is strongly elongated in one direction, being approximately an ellipsoid of revolution around this direction. The ratio of the largest and smallest axes is approximately 15. The ellipsoid can be described by the equation

$$\varepsilon = \frac{1}{2m_0}(\alpha_{xx}p_x^2 + \alpha_{yy}p_y^2 + 2\alpha_{yz}p_yp_z + \alpha_{zz}p_z^2) \equiv p\alpha p/2m_0, \quad (3)$$

which represents simply an expansion of the function $\varepsilon(\mathbf{k})$ in powers of $\mathbf{p} = \hbar(\mathbf{k} - \mathbf{k_i})$ near the minimum. The x axis is here the binary axis, the z axis the trigonal axis, and m_0 is the mass of the free electron. The symmetrical tensor α_{ik} can be reduced to the principal axes by rotating the coordinate system in the (y, z) plane. The principal axes of the ellipsoid are customarily denoted by the indices 1, 2 and 3 (with respect to projections p_1, p_2 and p_3); the axis 1 coincides with the binary axis C_2, the axis 2 is taken to be the direction of elongation, and the axis 3 is perpendicular to 1 and 2.

It was subsequently established that the surface (3) actually corresponds to electrons, that is, the energy inside the ellipsoid is smaller than on the Fermi surface. From theoretical considerations[9] (see also [18]) it follows that the centers of the ellipsoids are located at points of type L. It was observed, however, that the shape of the Fermi surface of the electrons differs from ellipsoidal,[19,20] and the extremal cyclotron masses and masses at the limiting points do not coincide noticeably,[20] whereas according to formula (3) the cyclotron mass should not depend on the projection of the momentum on the direction of the magnetic field.
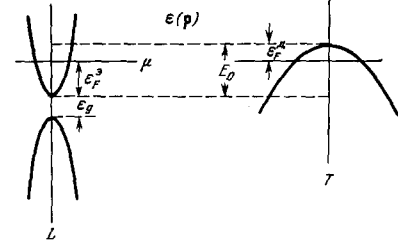

FIG. 6. Diagram of bands near the Fermi level μ.

In addition, it became known[20] that near the minimum of the conduction band there is a valence band at a distance of the order of the Fermi energy reckoned from the bottom of the band. For this reason, the $\varepsilon(\mathbf{p})$ dependence for energies on the order of ε_F cannot be quadratic. In this energy region the expansion (3) is not valid and to determine the form of $\varepsilon(\mathbf{p})$ it is necessary to take into account both bands. The two-band model was investigated theoretically in [22,23,24]. The purpose of [9] was to clarify the origin of the narrow gap and to relate the singularities of the space lattice with the character of the energy spectrum. These papers will be discussed in detail later. We merely note here that the electrons and holes appear in bismuth in a consistent manner.[9] The consistency is manifest not so much in the quality of the number of electrons and holes (this equality is the consequence of electroneutrality), as in the fact that their spectra are very similar.

The hole part of the Fermi surface was observed in [25]. This surface can be approximated by an ellipsoid of revolution with principal axis along the trigonal axis:

$$E_0 - \varepsilon = \frac{1}{2m_0}(\beta_\perp p_\perp^2 + \beta_{||} p_{||}^2). \quad (4)$$

In the experiment of [26], a small deviation of the hole spectrum from quadratic was observed. The center of the surface (4) is located at points of the type T.[9,18]

Thus the first Brillouin zone has one hole section and three electron sections of the Fermi surface. The level scheme is shown in Fig. 6. The figure does not show the far zones (see [27]) for which there is still no reliable experimental data.

A hypothesis is sometimes advanced that other groups of carriers are present in bismuth. The possibility of doubling of the number of electron and hole ellipsoids is also discussed. This is usually stimulated by some discrepancy between certain experimental data and the described model. However, the reasons for the discrepancies, which we shall consider later, lie apparently in the interpretation of the experiment.

2. The Two-band Model

Experiments revealed[21] that there is a valence band at the point L, at a small distance ε_g from the minimum of the conduction band. The appearance of a narrow gap is closely connected with the small deviation of the bismuth lattice from primitive cubic.[9] This circumstance makes it possible to estimate the order of magnitude of the effective mass at the bottom of the conduction band. Near the bottom itself, the expansion (3) is valid. For values of ε larger than ε_g, the $\varepsilon(\mathbf{p})$ dependence should go over into the corresponding de-

pendence for the undeformed lattice, which can be approximated by the linear relation

$$\varepsilon(\mathbf{p}) = \mathbf{vp}, \tag{5}$$

where v is of the order of the ordinary electron velocities, that is, 10^8 cm/sec. Intermediate between (3) and (5) is $p \sim \varepsilon_g/v$. Comparing (3) with (5) we obtain the order of magnitude of the effective mass:

$$m^* \equiv m_0/\alpha \sim \varepsilon_g/v^2. \tag{6}$$

Of course, this is a rather rough estimate, since it does not take into account the crystal anisotropy. However, the main law is correctly represented: small masses, as a rule, are connected with small energy gaps.

The two-band model was investigated independently of the singularities of the space lattice in [22-24]. For an analysis of the spectrum near a certain point $\mathbf{k_i}$, in which there are several closely lying bands, it is convenient to use the effective-mass method, which is widely used in semiconductor theory.[28] The field acting on the electron is described by the Hamiltonian

$$\mathcal{H} = \frac{\hat{\mathbf{p}}^2}{2m_0} + U + \frac{\hbar^2}{8m_0^2c^2}\Delta U + \frac{\hbar}{4m_0^2c^2}[\sigma\nabla U]\hat{\mathbf{p}}, \tag{7}$$

where $\hat{\mathbf{p}} = -i\hbar\nabla$ is the momentum operator and U is a periodic potential that takes into account in a self-consistent manner the influence of the remaining electrons. The last term of (7), which contains the Pauli matrices σ, is the spin-orbit coupling energy. The wave function is expanded in terms of the functions $\chi_{nk} = [\exp{(i\mathbf{k}\cdot\mathbf{r})}]u_{nk_i}$ $(\chi_{nk_i} = [\exp{(i\mathbf{k_i}\cdot\mathbf{r})}]u_{nk_i}(\mathbf{r})$ are the eigenfunctions of the Hamiltonian and correspond to the point $\mathbf{k_i}$):

$$\psi(\mathbf{r}) = \sum_n \int d^3k A_n(\mathbf{k})\chi_{nk}(\mathbf{r}). \tag{8}$$

If r bands are close at the point $\mathbf{k_i}$, and the remaining bands can be taken into account by perturbation theory, then we obtain for the determination of $A_n(\mathbf{k})$ the following system of equations:

$$\sum_{m=1}^{r}\left(\varepsilon_n\delta_{nm} + V_{nm} - \sum_{p=r+1}^{\infty} V_{np}V_{pm}/\varepsilon_p\right)A_m(\mathbf{k}) = \varepsilon A_n(\mathbf{k}). \tag{9}$$

The role of the perturbation is played here by $\hat{V} = \mathbf{p}\cdot\hat{\mathbf{v}}$, where

$$\hat{\mathbf{v}} = \left(-i\hbar\nabla + \frac{\hbar}{4m_0^2c^2}[\sigma\nabla U]\right)\Big/ m_0$$

is the velocity operator, $\mathbf{p} = \hbar(\mathbf{k} - \mathbf{k_i})$, the matrix elements are taken in terms of the functions χ_{nk_i}, and ε_n is the energy at p = 0. It is practically impossible to calculate V_{nm}, and it becomes necessary to use symmetry considerations. The operator in the left side of (9) in front of $A_m(\mathbf{k})$ is usually called the effective Hamiltonian.

We are interested in the case of two close spin-degenerate bands. If the symmetry of the point $\mathbf{k_i}$ is sufficiently high, then the functions χ_{nk_i} can differ also in parity. In bismuth, the conduction and valence bands have apparently different parities. Then the effective Hamiltonian, accurate to terms of first order in V/ε_p, can be written in the form[22, 24]

$$\widetilde{\mathcal{H}} = \begin{pmatrix} -\varepsilon_g I & H_1 \\ H_1^+ & 0 \end{pmatrix}, \tag{10}$$

and the matrix elements themselves are 2 × 2 matrices. Thus, I is a unit matrix and

$$H_1 = \mathbf{p}\begin{pmatrix} t & u \\ -u^* & t^* \end{pmatrix}, \tag{11}$$

where $t = \langle 01|v|a1\rangle$ and $u = \langle 01|v|a2\rangle$. The indices 1 and 2 designate functions pertaining to different spin projections, and 0 and a designate the valence and the conduction bands. The relations

$$\langle a2|\hat{\mathbf{v}}|02\rangle = t, \quad \langle a1|\hat{\mathbf{v}}|02\rangle = -u$$

are the consequence of invariance to time reversal. The energy in (10) is reckoned from the bottom of the conduction band, and $\varepsilon_g = \varepsilon_a - \varepsilon_0$.

The spectrum $\varepsilon(\mathbf{p})$ is obtained by solving the equation

$$|\widetilde{\mathcal{H}} - \varepsilon I| = 0. \tag{12}$$

The corresponding calculation is best carried out in general form. The matrix H_1 is expanded in terms of four linearly-independent matrices σ_r and I:

$$H_1 = i\sum_{r=1}^{3}\mathbf{p}W(r)\sigma_r = -H_1^+,$$

and the coefficient preceding the unit matrix turns out to be equal to zero if the basis functions are suitably chosen. The vectors $\mathbf{W}(r)$ are expressed in terms of u and t:

$$\mathbf{W}(1) = \text{Im}\,u, \quad \mathbf{W}(2) = \text{Re}\,u, \quad \mathbf{W}(3) = \text{Im}\,t.$$

In accordance with (t), the four-element column $A_n(\mathbf{k})$ is best written in the form

$$A = \begin{pmatrix} \varphi_u \\ \varphi_l \end{pmatrix},$$

with φ_u and φ_i each containing two elements. Equations (12) take the form

$$-\varepsilon_g\varphi_u + H_1\varphi_l = \varepsilon\varphi_u,$$
$$H_1^+\varphi_u = \varepsilon\varphi_l.$$

Obtaining φ_i from the second equation and substituting it in the first, we get

$$\frac{1}{\varepsilon_g}H_1 H_1^+\varphi_u = \varepsilon(1 + \varepsilon/\varepsilon_g)\varphi_u. \tag{13}$$

We can now reduce the $\varepsilon(\mathbf{p})$ dependence to the form[29]

$$E(\varepsilon) \equiv \varepsilon(1 + \varepsilon/\varepsilon_g) = \frac{1}{2m_0}\sum_{i,k=1}^{3}\widetilde{\alpha}_{ik}p_ip_k, \tag{14}$$

where

$$\widetilde{\alpha}_{ik} = \frac{2m_0}{\varepsilon_g}\sum_{r=1}^{3}W_i(r)W_k(r). \tag{15}$$

Although the $\varepsilon(\mathbf{p})$ dependence (14) is not quadratic, the surfaces ε = const have the form of an ellipsoid, just as the equal-energy surfaces of (3). It is possible to measure in experiment directly the area of the intersection $S(\varepsilon, \widetilde{p}_z)$ of the surface $\varepsilon(\mathbf{p}) = \varepsilon$ = const with a certain plane $\widetilde{p}_z$ = const and the cyclotron mass

$$m(\varepsilon, \widetilde{p}_z) = \frac{1}{2\pi}\frac{\partial S(\varepsilon, p_z)}{\partial\varepsilon},$$

wherein in the former case $\widetilde{p}_z$ corresponds to an extremum of $S(\varepsilon, p_z)$ and in the latter to an extremum of

$\partial S(\varepsilon, p_z)/\partial\varepsilon$. For the spectra (13) and (14), the ratio S/m does not depend on the direction of the normal to the secant plane. According to the experimental data the ratio $S(\varepsilon, \tilde{p}_z)/2\pi\, m(\varepsilon, \tilde{p}_z) = 0.0176$ eV and does not depend on the direction of the normal within the limits of errors. We shall return later to a comparison of the ellipsoidal model with experiment.

In accordance with the estimate (6), the denominator of (15) contains a small gap ε_g. Experiment has shown however, that only two out of the three principal values of the tensor α_{ik} are large, and the third, α_{22}, is of the order of unity. The cause of the anisotropy of the electron Fermi surface is at present regarded as accidental.* It is necessary to take into account the second-order terms in V/ε_p, at least for directions closed to the elongation of the ellipsoid. For the central part of the ellipsoid, the approximation (14) can be regarded as sufficiently good.

The second-order terms were considered in [23], where the following equation was obtained for the electrons;

$$\varepsilon^2 - (Q + R - \varepsilon_g)\varepsilon + R(Q - \varepsilon_g) - (|pt|^2 + |pu|^2) = 0. \quad (16)$$

We have introduced here the notation

$$Q = \frac{p^2}{2m_0} - p\sum_{n,\rho}\frac{\langle 01|\hat{v}|n\rho\rangle\langle n\rho|\hat{v}|01\rangle}{\varepsilon_n}p,$$

$$R = \frac{p^2}{2m_0} - p\sum_{n,\rho}\frac{\langle a1|\hat{v}|n\rho\rangle\langle n\rho|\hat{v}|a1\rangle}{\varepsilon_n}p.$$

At the very bottom of the conduction band, the spectrum (16) is quadratic:

$$\varepsilon_a(p) = R + (|pt|^2 + |pu|^2)/\varepsilon_g.$$

If we consider values of p such that Q, $Q + R \ll \varepsilon_g$ and $|pt|^2 + |p\cdot u|^2 \sim \varepsilon_g$, then Eq. (16) leads to the model (14). To explain the existing anisotropy of the Fermi surface it is necessary to assume that the coefficient preceding p_2 in $|pt|^2 + |p\cdot u|^2$ is small. It can be assumed that the dependence of ε on p_2 is determined primarily by Q and R. Then the spectrum takes the form

$$\frac{p_1^2}{2m_1} + \frac{p_3^2}{2m_3} = \left\{\left(\varepsilon - \frac{p_2^2}{2m_2}\right)\left(\varepsilon + \varepsilon_g + \frac{p_2^2}{2m_2'}\right)\right\}\Big/\varepsilon_g. \quad (17)$$

It is quite difficult to compare Eqs. (16) and (17) with experiment. No definite conclusions can be drawn with respect to this matter at present.

3. Deformation Theory

We proceed to describe the concepts developed in [9]. To obtain the energy spectrum, the lattice considered in [9] was infinitesimally close to a cubic one, and was then deformed into the true lattice of bismuth. In order that the change of the spectrum occur without a phase transition, the "dielectric" phase (see p. 3) is con-

sidered throughout, although this is not the equilibrium phase at small deformations.

The deformation will be effective if it leads to a lifting of the degeneracy (a similar instability of molecules was considered by Jahn and Teller). We have already seen, with the one-dimensional model as an example, that the bands split at the points $\pi/2a$ and $-\pi/2a$, which corresponded prior to the deformation to one value of the energy. If the equality of the energies is not accidental also in the three-dimensional case, then such points should enter in a single star, that is, they should satisfy the condition

$$G^0k = k', \quad (18)$$

where G^0 is a certain symmetry transformation of the initial cubic lattice. In the one-dimensional case, k goes over into k' upon inversion. One more condition relating the vectors k and k' can be obtained by recognizing that there are new translation periods a_i in the deformed lattice. Indeed, for the splitting of the bands it is necessary that the following matrix elements be different from zero

$$V_{k'k} = \int\psi_{k'}^*(r)V(r)\psi_k(r)\,d^3r,$$

where $\psi_{k'}(r)$ and $\psi_k(r)$ are the wave functions in the undeformed lattice, and $V(r)$ is the perturbation describing the deformation. $V(r)$ is periodic and has periods a_i. Since these periods are sums of elementary vectors of the cubic lattice, we can use the translation properties of the wave functions (1). Following translation through the vector a_i, the matrix element $V_{k'k}$ is multiplied by $\exp[i(k - k')\cdot a_i]$, and on the other hand the integral over the entire volume should not change upon translation of the coordinates. Therefore,

$$k - k' = 2\pi b,$$

where b is some vector of the reciprocal lattice vector (2) and, of course, differs from the reciprocal-lattice vectors of the initial cube. We note that this condition is also satisfied in the one-dimensional model. Combining both conditions, we obtain an equation for determining the vectors k:

$$G^0k - k = 2\pi b.$$

Taking for G^0 the inversion operation, we obtain eight vectors within the limits of the first Brillouin zone:

$$k_\alpha = \pi b_\alpha, \quad k_\alpha^{(1)} = -k_\alpha \quad (\alpha = 1, 2, 3),$$
$$k_0 = \pi(b_1 + b_2 + b_3), \quad k_0^{(1)} = -k_0.$$

These points lie on the boundaries of the zone.

We note that in the free-electron model the band splitting also takes place on the boundaries of the Brillouin zone.[30] The first six points are of type L and the last two of type T. All constitute a single star in the initial lattice and are pairwise equivalent in the deformed lattice. Prior to the deformation, the small group of each vector was C_{3v}. After the deformation, the star turns out to be reducible. The vector k_0 goes over into itself under all transformations of the symmetry group of bismuth, since k_0 and minus $-k_0$ are now equivalent. Its small group is $D_{3d} = C_{3v}C_i$. Another star is made up of the vectors k_α and $k_\alpha^{(1)}$. Since only one sixfold mirror-rotation axis remains in the crystal, leading to the transformations $k_1 \to -k_3 \to k_2 \to -k_1 \to k_3 \to -k_2$, it follows that k_α now has the less symmetrical small

---

*The opinion is advanced in [83] that the very existence of bismuth as a semimetal is possibly due to the anisotropy of its electron spectrum. Using a simple model as an example, it is shown in [83] that an isotropic semimetal should go over into a dielectric as a result of the attraction between the electrons and the holes. However, this instability takes place only at very low carrier densities. As noted by Mott [84], an isotropic semimetal is stable if the condition $n^{1/3} > 0.25me^2/\hbar^2\varepsilon_0$ (the condition for the absence of bound states) is satisfied, where ε_0 is the dielectric constant and m is the reduced mass of the electron and the hole.

group C_{2h}. For the point L (see Fig. 2) it is obtained by adding inversion to one of the cube symmetry planes σ_d remaining in the deformed lattice.

When determining the $\varepsilon(p)$ dependence near the points under consideration, we assume that the deformation is small. Since the spin-orbit coupling is much smaller than the ordinary atomic energies, it can be reasonably accounted for by perturbation theory. Therefore, we choose as the zeroth-approximation functions the products of the spinors φ_ρ ($\rho = 1, 2$) by the functions $\psi_\alpha(k_i)$, which transform in accordance with the ordinary representations of the small group C_{3v} of the vector k_i (i = 1, 2, 3). If the deformation is sufficiently small, then the electron and hole bands originate in the same bands of the initial lattice. Therefore $\varepsilon(p)$ should have minima at L and a maximum at T. The one-dimensional representations of the group C_{3v} do not satisfy this requirement. This leaves the two-dimensional representation.[31] To obtain the spectrum prior to the deformation, it is sufficient to choose as the basis the four functions $\varphi_\rho\psi_\alpha(k_i)$. It is then necessary to take the deformation into account. The procedure for doing so is known from semiconductor theory.[32] The only singularity in the case of bismuth is that the points k_i and $-k_i$ become equivalent upon deformation, and the corresponding basis functions should be unified. The rank of the effective Hamiltonian is then doubled. To determine $\varepsilon(p)$ it is necessary to diagonalize an 8 × 8 matrix. However, owing to the double degeneracy, the problem reduces to a determination of the eigenvalues of a 4 × 4 matrix:

	$\varphi_1\psi_1(k_i)$	$\varphi_1\psi_2(k_i)$	$\varphi_1\psi_1(-k_i)$	$\varphi_1\psi_2(-k_i)$
$\varphi_1\psi_1(k_i)$	$ap_z + \Delta + f$	$bp_+ + \beta$	γ	δ
$\varphi_1\psi_2(k_i)$	$bp_- + \beta$	$ap_z - \Delta + f$	δ	γ
$\varphi_1\psi_1(-k_i)$	γ	δ	$-ap_z + \Delta + f$	$-bp_+ + \beta$
$\varphi_1\psi_2(-k_i)$	δ	γ	$-bp_- + \beta$	$-ap_z - \Delta + f$

The basis functions are indicated at the left and at the top, and $p_\pm = p_y \pm ip_x$. The matrix corresponding to the second spin projection is obtained by reversing the sign of the constant Δ, which characterizes the spin-orbit coupling. The parameters a and b, which have the order of magnitude of ordinary electron velocities, and also Δ, are the same for all points k_i. The remaining parameters and the directions of the coordinate axes at each point are different. For the point L, the z axis (the direction ΓL in Fig. 7) lies in the σ_d plane and makes an angle $\cos^{-1} 1/3$ with the trigonal axis, the x axis is the binary axis, and γ, δ, β, and f are proportional to the deformation, γ and δ being connected only with the displacement of the sublattices. For the point
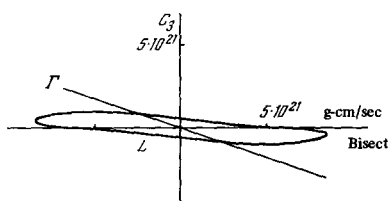

FIG. 7. Section of electron Fermi surface in the bisector plane [27].

T, the z axis is the trigonal axis, the x axis the binary axis, $\gamma_0 = -3\gamma$, $f_0 = -3f$, and $\delta_0 = \beta_0 = 0$.

Thus, the spectrum near the point L is given by the equation

$$\varkappa_1^4 + 2\varkappa_1^2(D + \varkappa_2\varkappa_3) + (\varkappa_2^2 + B)(\varkappa_3^2 + A) + 2\Delta^2\varkappa_2\varkappa_3 - C = 0. \quad (19)$$

For brevity, we have introduced here the notation:

$$A = \gamma_-^2 - \varepsilon_+^2, \quad B = \gamma_+^2 - \varepsilon_-^2, \quad D = \gamma_+\gamma_- - \varepsilon_+\varepsilon_- + \Delta^2,$$
$$C = 2\Delta^2(\gamma_+\gamma_- + \varepsilon_+\varepsilon_-) - \Delta^4, \quad \gamma_\pm = \gamma \pm \delta, \quad \varepsilon_\pm = \varepsilon \pm \beta - f,$$
$$\varkappa_1 = bp_x, \quad \varkappa_{2,3} = bp_y \pm ap_z.$$

Equation (19) can be quickly solved with respect to ε. Different experimentally-observed quantities such as the areas of the sections, the cyclotron masses, the density of states, are expressed by rather complicated integrals. They are given in [27]. In Eq. (9) account was taken of four close bands. If only two are particularly close, ε_0 and ε_a, then for values $\varepsilon \sim \varepsilon_F \sim \varepsilon_a$ the free term in (19), namely $C - AB \equiv -(\varepsilon - \varepsilon_1)(\varepsilon - \varepsilon_0)(\varepsilon - \varepsilon_a)$ × $(\varepsilon - \varepsilon_4)$* can be written in the form $C - AB \approx \varepsilon_1\varepsilon_4$ × $(\varepsilon - \varepsilon_0)(\varepsilon - \varepsilon_a)$. For those sections of the Fermi surface which are not too close to the direction of elongation, it is possible to omit from (19) the fourth-order terms in p, and A, B, D can be taken on the bottom of the band ε_a. In this case we obtain an equation of the type (14). A singularity of the surface (19) is that when ε is sufficiently large, it ceases to be closed. For such values of ε, the values of p can be sufficiently large and it is necessary to take into account in the matrix (18) terms of higher order in p. It is impossible to ascertain whether the surface (19) remains open in this case.

The spectrum near T is of the form

$$\varepsilon = f_0 \pm \{\Delta^2 + \gamma_0^2 + (ap_{||})^2 + (bp_\perp)^2 \pm 2[(\gamma_0\Delta)^2 + (abp_{||}p_\perp)^2 + (a\Delta p_{||})^2]^{1/2}\}^{1/2}. \quad (20)$$

The hole part of the Fermi surface is a figure of revolution. Its symmetry turns out to be higher than the symmetry of the point T. The reason is that small terms of higher order in p_x and p_y were not taken into account in (18).

If (20) is expanded in powers of the square root of p and second-order terms retained, then expression (4) is obtained. However, experiment[20,26] revealed a difference between the extremal cyclotron masses and the masses at the limiting points. In the case when the magnetic field is perpendicular to the trigonal axis, the extremal mass (on the central section) is given by the formula

$$m_{\perp, \text{extr}} = \frac{-\Delta}{\pi ab \, ch \, \psi} [K(q^2) + (cth \, \psi - 1) E(q^2)],$$

where

$$ch \, \psi = (\gamma_0 - \tilde{\varepsilon})/2\sqrt{\gamma_0\tilde{\varepsilon}}, \quad q^2 = \left[1 - \left(\frac{\gamma_0 - \tilde{\varepsilon}}{\Delta}\right)^2\right] ch^2 \, \psi, \quad \tilde{\varepsilon} = |\varepsilon - f_0|,$$

$K(q^2)$ and $E(q^2)$ are the complete elliptic integrals of the first and second kind, and the mass at the limiting point is

$$m_{\perp, \text{lim}} = -\frac{\tilde{\varepsilon}}{ab}\sqrt{\frac{\gamma_0\Delta}{\gamma_0^2 - \tilde{\varepsilon}^2 - \gamma_0\Delta}}$$

(we assume a, b, γ_0, $\Delta > 0$, omitting for brevity the absolute value sign). If the deviation of the spectrum (20)

*) $\varepsilon_1 < \varepsilon_0 < \varepsilon_a < \varepsilon_4$ —values of ε at $p = 0$

from quadratic is not large ($\tilde{\varepsilon} \rightarrow \gamma_0 - \Delta$), then

$$\frac{m_{\perp,\,\lim} - m_{\perp,\,\text{extr}}}{m_{\perp,\,\text{extr}}} = \frac{3}{8}\left(\frac{\Delta}{\gamma_0 - \tilde{\varepsilon}} - 1\right)\left(\frac{\gamma_0}{\tilde{\varepsilon}} + \frac{\tilde{\varepsilon}}{\gamma_0} - \frac{2}{3}\right).$$

For a magnetic field directed along the trigonal axis, the dependence of the mass on $p_\parallel$ is determined by the expression

$$m_\parallel = -\frac{\tilde{\varepsilon}}{b^2}\left(1 + \frac{(ap_\parallel)^2}{\sqrt{\gamma_0^2\Delta^2 - (\gamma_0^2 - \tilde{\varepsilon}^2)(ap_\parallel)^2}}\right).$$

We note that in either case the mass at the limiting point is larger (in absolute magnitude) than the mass on the central section.

The parameters of (19) and (20) were determined in [27] by comparing the calculated and measured values of the cyclotron masses and the cross-section areas. In addition, we needed the value of the gap ε_g. Unfortunately, ε_g was not measured directly in the experiments. In [21,33], the two-band model (14) was used to determine the gap. Since the two-band model can be regarded as sufficiently good for magnetic-field directions close to the direction of the elongation, the value of the gap apparently differs from that obtained in [21,33]. Thus, we obtained in [27] $|a| = 0.390 \times 10^8$ cm/sec, $|b| = 1.068 \times 10^8$ cm/sec, $|\Delta| = 0.1242$ eV, $|\gamma| = 0.1635$ eV, $|\delta| = 0.0853$ eV, $|\beta| = 0.0516$ eV, $f = -0.0973$ eV, and $\varepsilon = -0.0875$ eV. The value of ε given here corresponds to the Fermi level.

For holes, the agreement between the values calculated from these data and the observed ones turned out to be good. At the point T the distance from the maximum of the hole band to the other bands turned out to be much larger than the Fermi energy measured from the top of the band. Therefore the hole spectrum for energies on the order of the Fermi energy is quite close to quadratic. The hole density turned out to be equal to 3.15×10^{17} cm^{-3}. For electrons the situation is more complicated. Their Fermi surface differs from an ellipsoid. Figure 7 shows the intersection of the bisector plane and the electron part of the Fermi surface in the vicinity of the point L (see Fig. 2). To be able to compare Figs. 2 and 7 directly we assume that $a < 0$ and $b, \gamma, \Delta, \delta > 0$. The direction ΓL shown in Fig. 7 makes an angle $\cos^{-1} 1/3$ with the trigonal axis. This direction is very close to the line actually joining the points Γ and L. We shall reckon the angles from the trigonal axis towards the line ΓX. An acute angle is taken to be positive. The smallest central section corresponds to an angle $-83°15'$ between the trigonal axis and the secant plane. The experimentally measured[20] value is $-83°40' \pm 15'$. Regardless of the choice of the system of coordinates, the direction of the elongation of the "ellipsoid" lies between the bisector axis and the direction of ΓL. We note that in [27] it was impossible to choose the parameters of (19) and (20) such as to obtain approximately 84° for this angle (in this case the ellipsoid is also inclined 6° to the basal plane, but its bisector axis and the direction of ΓL were on the same side of the elongation direction). The calculated electron density per ellipsoid is $n_e^{(1)} = 1.15 \times 10^{17}$ cm^{-3}. The small difference between the number of holes and the total electron density $3n_e^{(1)}$ can be attributed to the limited accuracy of the theory. It must be stated that for directions corresponding to large cross sections and

large cyclotron masses, different small combinations of the parameters play an important role. Strictly speaking, in this case it would be necessary to take into account in (18) terms of higher order in the deformation. Such an allowance would have little effect on the values of the parameters but would greatly complicate the comparison with experiment.

IV. ELECTRON SPECTRUM IN A CONSTANT MAGNETIC FIELD

The greater part of the information concerning the spectrum of bismuth was obtained with the aid of constant magnetic fields. The influence of the field on the spectrum can be taken into account by making in the Hamiltonian (8) the substitution $\hat{p} \rightarrow \hat{P} = p - e\mathbf{A}/c$ and adding $-\mu_0\sigma H$ where $\mathbf{A}$ is the vector potential of the field H and μ_0 is the Bohr magneton. In the effective-mass theory it is shown that a similar change takes place also in the effective Hamiltonian. As a rule, the eigenvalues of the operator obtained in this manner can not be determined.

1. Classical Limit

If the magnetic field is sufficiently weak, then the motion of the electrons and holes can be considered on the basis of the classical equations of motion:

$$\dot{p} = \frac{e}{c}[\mathbf{v}H], \qquad \dot{\mathbf{r}} = \mathbf{v} = \frac{\partial \varepsilon}{\partial \mathbf{p}}. \tag{21}$$

The equations of (21) are valid for an arbitrary $\varepsilon(p)$ dependence. With the aid of (21) it is possible to obtain the important conservation laws $\dot{\varepsilon} = \mathbf{v} \cdot \dot{p} = 0$, $\dot{p}_z = 0$ (the axis is chosen along the direction of the magnetic field). This means that the trajectory of the electron in momentum space is determined by the intersection of the surface $\varepsilon(p) = \varepsilon = $ const and the plane $p_z = $ const. Depending on the character of the $\varepsilon(p)$ spectrum, the trajectories can be closed or open. In the former case knowing the $\varepsilon(p)$ dependence, it is possible to calculate the area $S(\varepsilon, p_z)$ bounded by the trajectory and calculate the frequency of revolution along the trajectory

$$\Omega = eH/cm(\varepsilon, p_z),$$

where

$$m(\varepsilon, p_z) = \frac{1}{2\pi}\frac{\partial S}{\partial \varepsilon}$$

is the cyclotron mass. For a free electron $m(\varepsilon, p_z)$ coincides with its mass m_0. On the bottom or on the top of the bands (3) and (4) the mass does not depend on ε or p_z but depends on the orientation of the magnetic field:

$$m = m_0[\det \alpha \cdot \lambda \alpha^{-1}\lambda]^{-1/2}, \tag{22}$$

where λ is a unit vector in the direction of the field. In the two-band model $S(\varepsilon, p_z) \sim E(\varepsilon) - ap_z^2$. Therefore the cyclotron mass does not depend on p_z but depends on ε:

$$m(\varepsilon, p_z) = \tilde{m}E'(\varepsilon) = \tilde{m}(1 + 2\varepsilon/\varepsilon_g), \tag{23}$$

where m is the mass at the bottom of the band as $\varepsilon \rightarrow 0$. The dependence of $m(\varepsilon, p_z)$ on $\tilde{\alpha}_{ik}$ is given by formula (22), in which α_{ik} should be replaced by $\tilde{\alpha}_{ik}/E'(\varepsilon)$.

In coordinate space, the projection of the trajectory on the plane perpendicular to λ is obtained from the trajectory in the p-space by rotation through an angle $\pi/2$ and by a change of scale. In particular, the diameters of the trajectories in the r- and p-spaces are connected by the relation

$$p_x = \frac{eH}{c} r_y. \tag{24}$$

2. Two-band Model and Extremum of Band

The spectrum of the electrons in a magnetic field can be obtained in the approximation linear in $\hat{V}$ by making in (10) the substitution

$$\mathbf{p} \rightarrow \hat{\mathbf{p}} \equiv \mathbf{p} - \frac{e}{c} \mathbf{A} \left(i\hbar \frac{\partial}{\partial \mathbf{p}} \right).$$

Repeating the calculations shown in page 5 we arrive at Eq. (13), in which

$$\frac{1}{\varepsilon_g} H_i H_i^+ = \frac{\tilde{\alpha}_{ik}}{4m_0} (\hat{p}_i \hat{p}_k + \hat{p}_k \hat{p}_i) - \mu_0 \sum_{r=1}^{3} HG(r) \sigma_r. \tag{25}$$

Here

$$G_h(r) = \frac{m_0}{\varepsilon_g} \sum_{i,j} C_{ij}(r) e_{ijh},$$

and the quantities $C_{ij}(r)$ are connected with the previously introduced $W_i(p)$ by the relation

$$C_{ij}(r) = \sum_{p,q} W_i(p) W_j(q) e_{pqr},$$

where e_{ijk} is a unit antisymmetrical tensor. The eigenvalues of the operator (25) were obtained in [22]:

$$\varepsilon (1 + \varepsilon/\varepsilon_g) = \hbar \tilde{\Omega} (r + 1/2 \pm 1/2) + p_z^2/2\tilde{m}_z. \tag{26}$$

The frequency $\tilde{\Omega}$ is determined by the cyclotron mass on the bottom of the band, $\tilde{\Omega} = eH/\tilde{m}c$. The connection between the mass and the inverse tensor $\tilde{\alpha}^{-1}$ is given by a formula similar to (22), and

$$\tilde{m}_z = m_0 (\lambda \tilde{\alpha}^{-1} \lambda). \tag{22'}$$

It is seen from (26) that the distance between the levels with values of r differing by unity is the same as the distance between the levels with different spin, corresponding to the ± signs. Because of this, all the levels except the lowest one in the conduction band and the highest one in the valence band, for which r = 0, turn out to be doubly degenerate. Inclusion of the terms of second order in $\hat{V}$ in the effective Hamiltonian[34] greatly complicates the form of the spectrum, since 16 new parameters appear. So large a number of unknowns can apparently not be determined experimentally.

The situation simplifies near the very bottom of the band. In this case

$$\varepsilon = \hbar \tilde{\Omega} (r + 1/2) + p_z^2/2\tilde{m}_z \pm \frac{\tilde{g}}{2} \mu_0 H. \tag{27}$$

The masses $\tilde{m}$ and $\tilde{m}_z$ are expressed in terms of $\tilde{\alpha}$ by means of the same formulas (22) and (22'), but account is taken of all the bands, including the remote ones, in the calculation of α and of the $\tilde{g}$-factor. The following expression is obtained for the $\tilde{g}$-factor:[22]

$$\tilde{g} = (\lambda G \lambda)^{1/2}, \tag{28}$$

where G is a certain symmetrical tensor.

In place of the $\tilde{g}$-factor, it is more convenient to use the effective g-factor, characterizing the ratio of the spin splitting to the cyclotron-splitting:

$$g = 2 \frac{\varepsilon^+(r) - \varepsilon^-(r)}{\varepsilon^+(r) - \varepsilon^+(r-1)}. \tag{29}$$

For the spectrum (26), just as for the free electron, g = 2. Near the bottom of the band (27) $g = \tilde{g}\tilde{m}/m_0$.

Inasmuch as for electrons the Fermi energy is of the same order as the gap ε_g, formula (27) can be used only for holes.

3. Quasiclassical Quantization

Some general information concerning the rather complicated spectrum of electrons in a magnetic field can be obtained for large quantum numbers with the aid of quasiclassical quantization. Accurate to terms of order $1/r$, the spectrum is given by the relation

$$\frac{c}{e\hbar H} S(\varepsilon, p_z) \pm \frac{\pi}{2} g(\varepsilon, p_z) = 2\pi (r + 1/2) + O(1/r). \tag{30}$$

The form of the principal terms of the expansion (30) (the first term on the left and $2\pi r$ on the right) is determined by the commutation relations satisfied by the operators $\hat{p}_\alpha$ in the magnetic field.[35] In the next higher order in r^{-1} there appears in (30) besides π (on the right) also a g-factor which takes into account the spin splitting of the levels in the magnetic field.[36] In the derivation of (30) it is assumed that the classical trajectory of the electron in p-space has no self-intersections, and the probability of jumping from one trajectory to another (magnetic breakdown) is small. The function $g(\varepsilon, p_z)$ is expressed by the integral over the trajectory:

$$g(\varepsilon, p_z) = \frac{1}{\pi} \operatorname{Im} \oint \frac{dp}{v} \psi_i^* \left[V_{ik} \frac{d}{d\mathbf{p}} \right]_z \psi_k.$$

The matrix element is taken here over the normalized eigenfunctions of the effective Hamiltonian in the absence of a magnetic field, and

$$V_{ik} = \frac{\partial \mathcal{H}_{ik}}{\partial \mathbf{p}}, \qquad v = \left| \frac{\partial \varepsilon}{\partial \mathbf{p}} \right|.$$

The function $g(\varepsilon_F, 0)$ was calculated in [36] for both electrons and holes using the spectrum parameters obtained in [27]. The dependence of $g(\varepsilon_F, 0)$ on the field direction is shown for holes in Fig. 8a, and for electrons in Fig. 8b. The dashed lines in the same figure show the variation of $g(\varepsilon_F, 0)$. Actually $g(\varepsilon_F, 0)$ does not vanish. At the corresponding directions of the magnetic field, it is necessary to take into account the term $-\mu_0 \sigma H$, which was omitted from the calculation. This leads to a splitting of the dashed and solid curves by a small distance of the order of m/m_0.

From (30), with accuracy of the order of r^{-1}, we obtain the distance between the cyclotron and a spin levels:

$$\Omega = \varepsilon^+(r) - \varepsilon^+(r-1) = \varepsilon^-(r) - \varepsilon^-(r-1) = e\hbar H/cm(\varepsilon, p_z),$$
$$\Omega_c = \varepsilon^+(r) - \varepsilon^-(r) = \Omega g(\varepsilon, p_z)/2.$$

Besides the effective g-factor, one introduces the "spin" mass

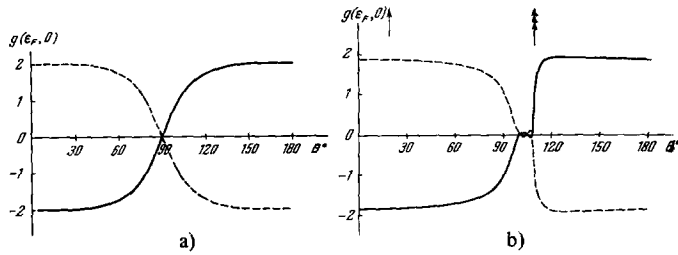$$m_s(\varepsilon, p_z) = \frac{2}{|g(\varepsilon, p_z)|} m(\varepsilon, p_z).$$

FIG. 8. a) dependence of the hole g-factor on the direction of the magnetic field. The angle θ is measured from the trigonal axis [36].
b) dependence of the electron g-factor on the direction of the magnetic field. The directions of the bisector and trigonal axes are shown by the arrows. The angle is measured from the "former" trigonal axis [36].

in the approximation (26) $m_S = m$. This relation is also satisfied quite well in the deformation model in a wide interval of magnetic-field directions (see Fig. 8).

The g-factor can be determined experimentally by observing paramagnetic or combined resonance. However, neither has been observed in bismuth so far. Individual communications published at times were not confirmed in later investigations.

V. SPECIFIC HEAT

The most precise measurements of the specific heat at low temperatures were made by Strelkov and Kalinkina[37] in the interval 0.3–4.4°K and by Phillips[38] at 0.1–2°K. The temperature dependence of the specific heat in this region should be described by the relation

$$C = c_1 T^{-2} + c_2 T^3 + c_3 T.$$

The first term (which was not observed in [37]) is connected with the quadrupole levels of the nucleus in the inhomogeneous crystal-lattice electric field. The coefficient c_3 in the term linear in the temperature is determined by the density of the states on the Fermi surface:

$$c_3 = \frac{\pi^2 k^2}{3} \frac{dn}{d\varepsilon_F}.$$

The cubic term describes the contribution of the phonons, and the coefficient c_2 can be described in terms of the Debye temperature θ:

$$c_2 = \frac{12}{5} \pi^4 k / \theta^3.$$

According to the data of [37] $\theta = 118.5°$K, and according to [38] $\theta = 120.4°$K. The density of states is listed in Table I, which also indicates the density of states obtained in the ellipsoidal approximation using the experimental data of [19, 20]. It is easy to verify that the

Table I. Density of states on the Fermi surface.

$dn/d\varepsilon$, 10^{19} eV^{-1}cm^{-3}	80 *)	25 **)	2.5$_e$+3.7$_h$***)	3.51$_e$+3.75$_h$****)

The indices "e" and "h" denote the contributions of the electrons and holes, respectively.

     *) Measured in [37]
     **) Measured in [38]
     ***) Calculated in the ellipsoidal approximation using the data of [19,20]
     ****) Calculated in [27]

$E(\varepsilon)$ dependence (see (14)) is immaterial here if one uses the experimentally determined extremal cross sections and cyclotron masses. In this sense, (3) and (14) lead to the same result. The contribution of one ellipsoid (electron or hole) is equal to

$$\frac{dn^{(1)}}{d\varepsilon_F} = 3\pi n^{(1)} m / S,$$

where m and S are the cyclotron mass and the area of the central section at the same direction of the magnetic field, and $n^{(1)}$ is the number of electrons or holes in one ellipsoid (with allowance for the spin degeneracy):

$$n^{(1)} = (S_1 S_2 S_3)^{1/2} / 3h^3 \pi^{7/2}.$$

Apparently the principal cause of the discrepancies in Table I lies in the smallness of the electronic contribution, which does not exceed 15% of the total specific heat in the investigated temperature interval. Therefore the inaccuracies in the determination of the temperature, sample defects, and the presence of the quadrupole term unaccounted for in [37] can lead to a large error. Another possible cause of the difference between the measured density of states and the calculated one is connected with allowance for the Fermi-liquid interaction. However, as already noted, in bismuth this interaction must be regarded as small.

VI. MAGNETIC SUSCEPTIBILITY

Let us consider first the constant field-independent part of the susceptibility

$$\chi_{ik} = -\frac{\partial M_i}{\partial H_k}\bigg|_{H \to 0}.$$

In the case of rhombohedral symmetry the symmetrical tensor has two independent components, $\chi_{||}$ and $\chi_\perp$ (the polar axis is taken to be the trigonal axis of the crystal). Bismuth is diamagnetic—both components $\chi_{||}$ and $\chi_\perp$ are negative. A noticeable anisotropy of the susceptibility, that is, the difference between $\chi_{||}$ and $\chi_\perp$, was observed already in the earlier investigations. A small amount of impurities changes the susceptibility strongly. The temperature dependence of $\chi_{||}$ and $\chi_\perp$ in the interval 14–400°K was measured in [39]. At helium temperatures $\chi_\perp = -1.7 \times 10^{-6}$ and $\chi_{||} = -1.2 \times 10^{-6}$.

The question of the magnetic susceptibility is one of the most difficult ones in the theory. The gas of free electrons is paramagnetic, since the diamagnetic part of the susceptibility amounts to $\frac{1}{3}$ of the paramagnetic part. One might assume that the diamagnetism of crystalling bismuth is due to the internal shells of the ions. However, this assumption is contradicted by measurements of χ in samples with small amounts of impurities. Another argument in favor of the connection between the susceptibility and the crystal structure was advanced by Jones.[30] He called attention to the fact that the susceptibility of molten bismuth is only -0.08×10^{-6}. In the same paper, Jones attempted to explain the large value of the susceptibility on the basis of the Landau-Peierls formula, which describes the susceptibility of electrons in the conduction band in analogy with free electrons. This attempt was not successful—the calculated value (using modern values of the spectral parameters) is much smaller than the ex-

perimental one. The sign of the change of $(\chi_\perp - \chi_{||})/\chi_\perp$ following addition of small amounts of lead impurities acting as acceptors[40] is also incorrect. On the other hand, since the spin splitting of the levels is close to the cyclotron splitting, the conduction electrons should lead to paramagnetism, just like the free electrons. The main difficulty is connected here with the impossibility of obtaining a general expression for the levels in a magnetic field. Since a weak field is sufficient for the calculation of χ, one could hope to use the quasiclassical expression (30) (the necessary allowance for the term of order H^2 in the left side of (30) entails no difficulty). However, the expression obtained in this case for χ is determined by levels with $r \sim 1$, for which the quasiclassical theory is not applicable as a rule. However, the quasiclassical calculation[35] leads to one important result: contributions to the constant part of the susceptibility are made not only by the electrons on the Fermi surface, but also by the deep electrons, including those in the valence band. The order of the effective mass of the valence-band electrons is roughly speaking the same as that of the conduction electrons (see the estimate (6)) but their number is much larger. This circumstance was taken into consideration by Adams[41] even before the publication of [35]. His result agrees qualitatively with experiment.[42, 40]

VII. QUANTUM OSCILLATIONS OF THE SUSCEPTIBILITY AND OF OTHER THERMODYNAMIC AND KINETIC QUANTITIES

The nonmonotonic dependence of the magnetic moment of the bismuth crystal on the magnitude of the magnetic field was observed by de Haas and van Alphen.[43] A detailed investigation of the de Haas–van Alphen effect in bismuth was made by Schoenberg.[15] The origin of the oscillations is connected with the successive coincidence of the Fermi level with the levels (30) corresponding to different r and to a certain preferred value of $\tilde{p}_z$. Inasmuch as the corresponding values of r turn out to be large, it is possible to use quasiclassical quantization. The expression for the oscillating part of the thermodynamic potential, obtained by Lifshitz and Kosevich,[35] is of the form

$$F = \frac{2VkT}{h^3}\left(\frac{e\hbar H}{c}\right)^{3/2}\left(\frac{\partial^2 S}{\partial p_z^2}\right)^{-1/2}\sum_{r=1}^{\infty}\frac{(-1)^r\cos\left(\frac{rcS}{e\hbar H}-\pi/4\right)\cos(\pi r g/2)}{r^{3/2}\,\mathrm{sh}\,(2\pi^2 kTr/\hbar\Omega)}. \quad (31)$$

The quantities S, m, $\dfrac{\partial^2 S}{\partial p_z^2}$, and g, which depend on ε and p_z, are taken here at $\varepsilon = \varepsilon_F$ and at the value of $\tilde{p}_z$ for which S has a maximum as a function of p_z (in the case of a minimum, the sign in front of $\pi/4$ in the argument of the cosine is reversed). Since the Fermi surface can have several extremal sections, it is necessary to sum in (31) over these sections. Formula (31) is valid when $\varepsilon_F \gg kT$ and $cS/e\hbar H \gg 1$. It does not take into account the level broadening due to collisions. It is shown in [44, 45] that electron scattering by impurities leads to an additional factor $\exp(-\pi^2 r/\Omega\tau)$ in (31), where τ is the collision time.

The main reason why the de Haas–van Alphen effect became a powerful method for studying the Fermi surface is that the observation of the oscillations makes it possible to determine, with high degree of accuracy, the
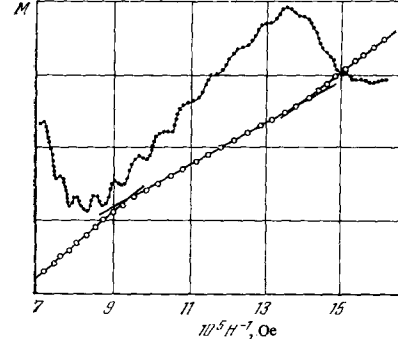
FIG. 9. The de Haas-van Alphen effect. The lower curve shows the dependence of the numbers of the maxima and minima on H^{-1} [48].

area S of the extremal section. In experiment, as a rule, one measures the torque $J = -\partial F/\partial\psi$ that twists the sample, and the magnetic field can usually be rotated in a plane perpendicular to the suspension of the sample, making it possible to measure the extremal sections at different positions of the secant surface (ψ—angle between the direction of the field H and a certain fixed direction in the plane perpendicular to the suspension). A formula for J is obtained from (31) by replacing cos $(rcS/e\hbar H - \pi/4)$ by sin $(rcS/e\hbar H - \pi/4)$, and by multiplying by $-(rc/e\hbar H)\partial S/\partial\psi$ (in the differentiation with respect to ψ it is sufficient to differentiate the larger quantity $cS/e\hbar H$). One usually records the dependence of J on H^{-1} at a given field direction (Fig. 9). This dependence is oscillating with a period $\Delta(H^{-1}) = e\hbar/cS$. The period of the oscillations is obtained by the plotting on the graph the values of $H^{-1}_{max}(n)$ corresponding to the maxima or minima. The period can be determined from the slope of the linear section. This procedure is sufficiently simple for those magnetic field orientations at which the oscillations from different ellipsoids can be distinctly separated. The oscillation amplitude changes when the temperatures change, making it possible to determine the cyclotron mass on the extremal sections. However, these measurements call for a vigorous separation of the harmonics and their accuracy is low compared with the cyclotron resonance.

With the aid of the de Haas–van Alphen effect it is possible to draw certain conclusions concerning the magnitude of $g(\varepsilon_F, \tilde{p}_z)$. In particular, the first harmonic (with $r = 1$) is not observed in the case $g(\varepsilon_F, \tilde{p}_z) = 1$, and the series (31) begins with $r = 2$. The factor with the g-factor (31) is the result of summation over two spin directions. However, the splitting itself is usually observed only in a sufficiently strong magnetic field, where the quasiclassical expression (31) is not valid.

The conditions for observation are favorable at those field directions for which the sections S are small. On such sections the cyclotron mass is also small and therefore the amplitude of the oscillations is relatively large. Oscillations on large sections are much more difficult to observe. One of the two principal major sections of the electron ellipsoid was measured only recently,[19] and the other one, in the (y, z) plane has not yet been measured. The reason for these difficulties lies both in the large mass and in the condition $\partial S/\partial\psi = 0$ for those directions for which the cross section has

an extremum (as a function of the angle ψ). The accuracy with which the large cross section is determined becomes worse as a result of one more specific circumstance. In the derivation of (31) it was assumed that the change of the chemical potential in a magnetic field can be neglected. In the case $cS/e\hbar H \gg 1$, this assumption is perfectly valid. However, to determine the largest electronic cross section the field must be directed along the binary axis. Then the areas of the sections of the two other electron ellipsoids will be smaller by approximately a factor of 10. In order for the amplitude for the oscillations on the large section to be observable, the magnetic field must be chosen such that the ratio $cS/e\hbar H$ for small cross sections be of the order of unity. Therefore, the chemical potential determined from the condition of the equality of the total number of electrons to the number of holes cannot be regarded as constant. The "spilling" of electrons from one ellipsoid to the other, and also into the hole band,[46, 48] leads to a variable period of the oscillations (see Fig. 9).

Figure 10[19] shows the angle of dependence of the cross sections for the three principal crystal planes. The points not lying on the curves constructed in accordance with the ellipsoidal model can be attributed to higher harmonics or regarded as the consequence of the spilling (crosses). We present the values of the principal sections:[19, 48]

holes:

H along the trigonal axis:
$$S_3 = (6.75 \pm 0.25) \times 10^{-42} \ (\text{g-cm/sec})^2,$$
H perpendicular to the trigonal axis:
$$S_1 \approx S_2 \approx (20.5 \pm 1) \times 10^{-42} \ (\text{g-cm/sec})^2;$$

electrons:

H along the binary axis:
$$S_1 = (19 \pm 1) \times 10^{-42} \ (\text{g-cm/sec})^2,$$
H along the elongation direction:
$$S_2 = (1.34 \pm 0.05) \times 10^{-42} \ (\text{g-cm/sec})^2.$$

With the aid of these data it is possible to estimate the densities n_h of the holes and $n_e^{(1)}$ of the electrons (per ellipsoid), approximating their Fermi surfaces by ellipsoids. The unmeasured principal electronic section S_3 is obtained by assuming that the ratio S/m does not depend on the direction, and by using the corresponding value of the cyclotron mass.[20] We get $n_h = 2.76 \times 10^{17} \ \text{cm}^{-3}$ and $n_e^{(1)} = 0.99 \times 10^{17} \ \text{cm}^{-3}$. It is difficult to estimate the error in this case, but the agreement between the number of holes and electrons (under the condition that there are three electron ellipsoids per hole ellipsoid) is surprisingly good. An opinion was advanced in [19] that the plane drawn through the axes 1 and 2 is not a symmetry plane of the electronic section of the Fermi surface, since the $S(\psi)$ dependence cannot be sufficiently well reconciled with the ellipsoidal model in the entire angle interval.

Oscillations of the type (31) should be experienced by all thermodynamic quantities. Oscillations of the temperature of a thermally-insulated sample were observed in [49]. With increasing magnetic field, the oscillations took the form of peaks, and the last observed peak was split. When the magnetic field was directed along the
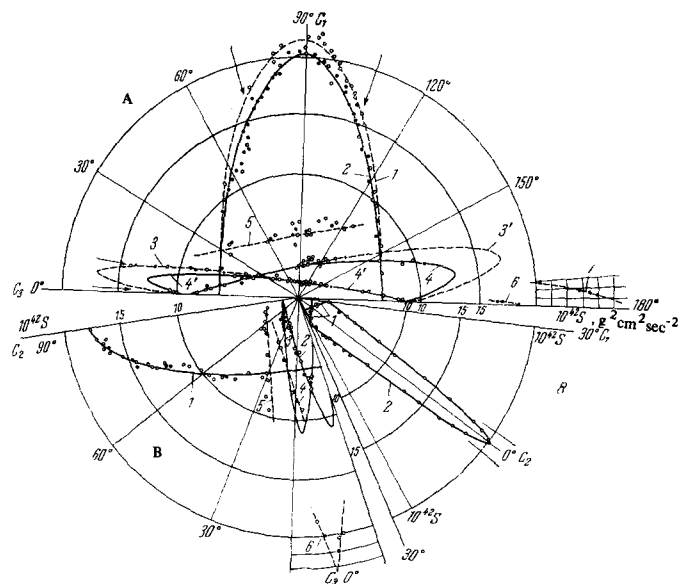
FIG. 10. Polar diagrams of the areas of the extremal sections in the planes perpendicular to the direction of the magnetic field [19]. A) (H)—in the trigonal-bisector plane C_3C_1. Curves 1 and 2 illustrate the possible anisotropy of the hole part of the Fermi surface in the planes C_3C_2 and C_3C_1; curves 3, 4 and 3', 4' are constructed in accordance with the ellipsoidal model with different principal sections. B) H is in the trigonal-binary plane C_3C_2: 1—holes, 2 and 3-electrons, 4, 5, 6-observed second and third harmonics of the fundamental frequency. C) H is in the basal plane C_1C_2: 1 and 2-electrons. The arrows indicate the directions of the field at which the frequency doubles as a result of the spin splitting of the Landau levels.

binary axis, the split peak corresponded to a field of 15 kOe, and the difference between the two components of the peak was much smaller than the distance to the nearest unsplit peak. The authors of [49] attributed this splitting to spin-splitting of the levels and used the formula

$$\varepsilon = \left(r + \frac{1}{2} \pm g/4\right)\frac{e\hbar H}{mc}.$$

If we assume that both components corresponded to different r that differ by unity, we obtain for the g-factor the value 1.84. This figure pertains to those two ellipsoids for which the indicated axis is not a symmetry axis.

Oscillations having the same nature as the de Haas–van Alphen effect were observed also in kinematic phenomena. Oscillations of the resistance were first observed by Shubnikov and de Haas.[50] The earlier investigations consisted essentially of observations of the phenomenon itself. A single-crystal sample was placed in a magnetic field. Direct current was made to flow in a direction perpendicular to the field through the sample. At low temperatures in fields on the order 10 kOe, oscillations appeared against the background of smooth variation of the resistance. It became clear subsequently that the Hall coefficient and the longitudinal resistance also oscillate (at coinciding field and current directions). The experimental procedure was also improved. Modulation of the constant field, amounting in magnitude to several Oe, made it possible to employ a sensitive electronic measurement scheme, wherein the first or second derivative of a definite component of the resistance tensor with respect to the magnetic field

were measured. The results of the measurement are presented in the form of a plot of the resistance or of the corresponding derivative against the magnetic field (the field itself increases smoothly during the measurement from zero to a certain maximum value).

The theoretical interpretation of the Shubnikov–de Haas effect is quite complicated. For this reason, the reduction of the experimental data consists only of determining the period of the oscillations, that is, the extremal sections of the Fermi surface. The procedure is the same as in the de Haas–van Alphen effect. It is surprising that in spite of the obvious advantages of its observation, the Shubnikov–de Haas effect was not successful in yielding additional information. In particular, the major sections of the electronic surface were not measured. Incidentally, Lerner[51] reported existence of "heavy" electrons and "heavy" holes. He measured dR/dH with the aid of a modulation procedure and observed on the $H_{max}^{-1}(n)$ plot small linear sections, which he connected with the heavy carriers.

A detailed analysis of the data obtained with the aid of the Shubnikov–de Haas effect is contained.[52] Using the ellipsoidal model, its authors constructed for certain magnetic field directions a plot of $H_{max}^{-1}(n)$ with allowance for all four ellipsoids—three electronic and one hole. The maxima which cannot be distinguished as a result of temperature and relaxation broadening were combined. It turned out that it is actually possible to separate linear sections on the curve, and this is what suggested to Lerner the idea of the existence of heavy carriers. The same paper[52] analyzes also the oscillation of the chemical potential with variation of the magnetic field, a fact likewise not taken into account by Lerner. Nor was the existence of heavy carriers supported by the data of [53].

So far we referred to oscillations of the dc resistance. If alternating current is excited in the metal, then the properties of the sample are described by its surface impedance. However, in the case of bismuth the corresponding values of the magnetic field turned out to be so large, that the presence of the skin layer has essentially no effect on the character of the phenomenon. It was investigated in detail in [47] at a frequency 5 Mhz with the aid of an NMR spectrometer. The interpretation of the results was greatly facilitated in this case by the fact that the amplitude of the oscillations from the different sections of the Fermi surface depends on the direction of the high-frequency currents. The method has fairly high sensitivity. However, even here it was impossible to trace the entire angular dependence of the extremal section of the electron ellipsoid. The polar diagram $S(\psi)$ of the hole ellipsoid is shown in Fig. 11. The arrows denote the directions in which the amplitude of the oscillation is decreased sharply, and something that looked like a doubling of the period is observed at $\psi = 73°$. (Doubling of the frequency of both the hole and electron oscillations at definite field directions was observed earlier in [19].) A unique method was used in [33]. The conductivity of a "sandwich" made of a bismuth plate, a dielectric layer, and an aluminum foil was measured at 100 kHz. Rather strong magnetic fields were used, up to 88 kOe. The data reduction was by means of formula (27) for holes and formula (26) for electrons, and the g-factor of the electrons was as-
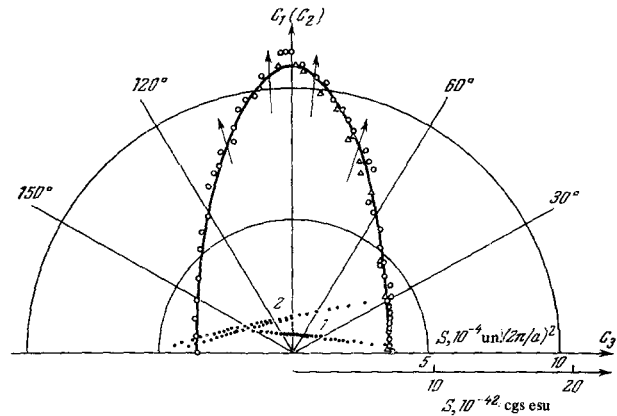
FIG. 11. Anisotropy of extremal sections in the $C_1 C_3$ plane. ○—holes; ●—electrons; △—holes in the $C_2 C_3$ plane [47].

sumed to be different from 2. From the theoretical point of view, this assumption is not justified. The difference between the g-factor and 2 is connected with allowance for other bands within a framework of the two-band model. However, the cyclotron masses also change in this case.[34] From the methodological point of view, the work in [33] is of definite interest. The state density was calculated with a computer in the approximation of (26) and (27). The dependence of the peak position on the field direction was traced by assuming that the peaks of the state density correspond to resistance minima. Best agreement between the calculated and experimental values was obtained by choosing the spectrum parameters. The main difficulty consisted in determining the quantum numbers corresponding to the peaks. We present a few values obtained in [33]:
$\varepsilon_g = 0.0153$ eV, $\varepsilon_F^{(e)} = 0.0276$ eV, $E_0 = 0.0385$ eV, $n_e = n_h = 2.75 \times 10^{17}$ cm^{-3}. The g-factor of the holes is discussed in [19, 33, 52, 53]. The first harmonic of the hole oscillations was not seen in [53] when at an angle 70° or 85° between the direction of the magnetic field and the trigonal axis; the corresponding values of [52] are 67.5° and 85.5°. As already noted, this can be attributed to the equality of the g-factor to an odd number in the corresponding directions. If it is assumed that the g-factor decreases monotonically, being small when $H \perp C_3$, then $g_{70°} = 3$ and $g_{85°} = 1$. According to these data, it is possible to estimate with the aid of (26) the g-factor for $H \parallel C_3$. The value obtained is close to 4. The fact that the g-factor of the holes is much larger than 2 is extremely surprising. It contradicts the theory of [36]. According to [22], g = 2 at the top of the hole band if there is only one close band (at a distance smaller than or of the order of the spin-orbit energy). Inclusion of more remote bands, and also of the spin of the free electron, leads to small corrections. It is not clear how to eliminate this contradiction between theory and experiment. The decisive result will be obtained by experiments in a very strong magnetic field. When g = 2, all the levels except one last one are doubly degenerate, and when g = 4 the last two levels are non-degenerate. Violation of the quasiclassical conditions in the strong field is not an obstacle in this case, for if the theory of [36] is valid, then the quasiclassical results will coincide with the exact one when $H \parallel C_3$.

VIII. ELECTRIC CONDUCTIVITY

The classical theory of electric conductivity is constructed on the basis of a compatible solution of the Maxwell equations and of the kinetic equations

$$\frac{\partial f}{\partial t} + \mathbf{v}\,\frac{\partial f}{\partial \mathbf{r}} + \Omega\,\frac{\partial f}{\partial \varphi} + \left(\frac{\partial f}{\partial t}\right)_{\text{col}} = -e\mathbf{v}\mathcal{E}\,\frac{df_0}{d\varepsilon}\,. \tag{32}$$

Here f_0 is the equilibrium distribution function, and f is a small non-equilibrium addition proportional to the electric field $\mathcal{E}$. The term $\Omega(\partial f/\partial\varphi)$ takes into account the constant magnetic field. The influence of the alternating magnetic field connected with the field $\mathcal{E}$ can be neglected in most cases. Equation (32) linearized in terms of the field $\mathcal{E}$ is valid if $\mathcal{E}$ is sufficiently weak. From (32) we can express f in terms of $\mathcal{E}$ and then determine the current

$$\mathbf{j} = \frac{2e}{h^3}\int f\mathbf{v}\,d^3p,$$

after which Maxwell's equations can be used to determine the field. Besides purely technical difficulties encountered in this case, there are two fundamental difficulties, namely determination of the form of the collision integral $(\partial f/\partial t)_{\text{col}}$ and the question of the boundary conditions in (32). In the calculation of $(\partial f/\partial t)_{\text{col}}$ is is customary to use some model of the scattering mechanism. It is sometimes possible to prove in general form[54] the validity of introducing a collision time $\tau(\mathbf{p})$:

$$\left(\frac{\partial f}{\partial t}\right)_{\text{col}} = \frac{f}{\tau(\mathbf{p})}\,.$$

Special interest attaches, however, to those cases when the collisions do not play an important role.

1. Static Conductivity in a Constant Magnetic Field

If we disregard the surface effect,[55] which exists also in the case of direct current in a strong magnetic field, then the electric field and the current are connected by the local relation $j_i = \sigma_{ik}(H)\mathcal{E}_k$. In the experiment one measures the resistivity tensor $\rho_{ik} = \sigma_{ik}^{-1}$. The form of $\rho_{ik}(H)$ in a strong field $(\Omega\tau \gg 1)$ can be obtained without any assumptions regarding the scattering mechanism.[56] In the case of closed trajectories

$$\rho_{\alpha\beta} \sim H^2, \quad \rho_{\alpha z} \sim \rho_{z\alpha} \sim H, \quad \rho_{zz} \sim H^0 \sim \text{const.} \tag{33}$$

The z axis is chosen here along the magnetic field and $\alpha,\beta = x, y$. The proportionality coefficients in (33) depend on the orientation of the magnetic field relative to the crystal axes.

The resistance of bismuth in a strong magnetic field was measured many times. The dependence on the magnetic field agrees with (33). In one of the latest investigations[57] it was found that $\rho_{\alpha\alpha} \sim H^{1.86}$. In the same paper, an attempt is made to compare the results of the measurements with a calculation in which the scattering of the electrons by phonons is described with the aid of deformation potential. Several simplifications are made in the calculation, but these simplifying assumptions are not well founded. The agreement between such a theory and experiment is qualitative.

Although in the case of weak fields $(\Omega\tau \ll 1)$ the collisions play a decisive role and it is impossible to calculate the resistance in general form, it is possible to use a phenomenological expansion in powers of the projections of the magnetic field

$$\rho_{ij}(H) = \rho_{ij}^0 + \rho_{ij,\,k}H_k + \rho_{ij,\,kl}H_kH_l + O(H^3). \tag{34}$$

The symmetry of the coefficients of this expansion is determined by the symmetry of the lattice:[58] the tensors ρ_{ij}^0 and $\rho_{ij,\,k}$ have two independent components each, while $\rho_{ij,\,kl}$ have eight components. Thus, galvanomagnetic phenomena in bismuth in a weak magnetic field are described by 12 parameters.

Thorough measurements in weak fields at helium temperatures are reported in [59]. All 12 components were measured at 4.2°K. If it is assumed that the collision time is isotropic,[60] that is, dependent only on the energy, then it is easy to obtain from the kinetic equation (32), in the ellipsoidal approximation, the relation

$$j = \sigma^0\{\mathcal{E} + (n^{(1)}ec)^{-1}[jH]\}, \tag{35}$$

between the current density and the electric and magnetic fields. Here $n^{(1)}$ is the number of carriers of any particular ellipsoid (electron or hole) and $\sigma_{ik}^0 = n^{(1)}e^2\tau\alpha_{ik}/m_0$. Summarizing the contributions of all the ellipsoids and expanding the resistance in powers of H, we can express the components of the tensor (34) in terms of the mobilities $\mu = e\tau\alpha/m_0$ of the electrons and holes and their densities. By comparison with experiment, Zitter[59] determined four components of the electron mobility tensor, (in a coordinate system fixed in the crystal), two components of the hole mobility, and the total number of electrons, which was found to agree with the total number of holes, $2.5 \times 10^{17}\text{ cm}^{-3}$ ($\pm$ 10%). Returning then to τ and using the values of the reciprocal mass tensor determined in [36], Zitter obtained $\tau_1 = 2.2 \times 10^{-10}$ sec, and $\tau_2 = 1.9 \times 10^{-10}$ sec, and $\tau_3 = 1.6 \times 10^{-10}$ sec for the principal values of the electron mobility tensor, and $\tau_1 = 4.3 \times 10^{-10}$ and $\tau_2 = \tau_3 = 5 \times 10^{-10}$ sec for the holes. We see now that the anisotropy of τ does not exceed 15% and this justifies the initial assumption. The values of the tensor (34) calculated from these components agree quite well with the measured values.

2. Conductivity in the Absence of a Constant Magnetic Field

For a sufficiently pure sample and for low temperatures, the frequency interval can be divided into three essentially different regions:

(a) $\omega \gg v\omega_n/c$—high frequency normal skin effect;

(b) $v\omega_n/c \gg (c/v\tau\omega_n)^2/\tau$—anomalous skin effect;

(c) $(c/v\tau\omega_n)^2/\tau \gg \omega$—low frequency normal skin effect.

In regions (a) and (c) the electric field ($\mathcal{E} \sim \exp[i(\mathbf{k}\cdot\mathbf{r} - \omega t)]$) and the current are related by the local connection $j_i(\omega) = \sigma_{ik}(\omega)\mathcal{E}_k(\omega)$, and the dependence of the conductivity on the wave vector (spatial dispersion) is not significant. The expression for $\sigma_{ik}(\omega)$ in the region (a) is obtained from (32):

$$\sigma_{ik}(\omega) = \frac{2e^2i}{\omega h^3}\int v_iv_k\,\frac{dS}{v}\,. \tag{36}$$

The integration is carried out here over the Fermi surface. Besides σ_{ik}, one frequently introduces the dielec-

tric constant

$$\varepsilon_{ih} = 4\pi i\sigma_{ik}/\omega = -\frac{(\omega_n^2)_{ih}}{\omega^2},$$

where

$$(\omega_n^2)_{ik} = \frac{8\pi e^2}{h^3} \int v_i v_k \frac{dS}{v}. \qquad (37)$$

The symmetrical tensor $(\omega_n^2)_{ik}$ in the expression for ε_{ik} plays the role of the square of the plasma frequency.* The frequency-independent contribution ε_{ik}^0 in ε_{ik} can be neglected in this frequency region, but the region of applicability of (36) is actually determined by the condition $\omega \ll \omega_n\varepsilon_0^{-1/2}$. The tensor $(\omega_n^2)_{ik}$ has two independent components. The order of magnitude of ω_n is 3×10^{14} sec^{-1}, and $\varepsilon_0 \sim 100$. In the region (c) the frequency dependence of the conductivity becomes negligible. The principal role is played here by scattering processes. The order of magnitude of σ is obtained from (36) by making the substitution $i/\omega \to \tau$. The boundaries of the normal skin effect are determined by comparing the values of kv and ω for the high frequency region and of kv and τ^{-1} for the low-frequency region (see (32)). The value of k, determined with the aid of Maxwell's equations, is of the order of $(2\pi(2\pi\sigma\omega)^{1/2}/c$.

Regions (a) and (b) do not overlap for pure samples and low temperatures: $l = v\tau \gg c/\omega_n$. In this case there appears a region in which the spatial dispersion is significant. The current density at any given point of space is determined here by the value of the field in the entire vicinity of this point. The properties of the metals are best described in terms of the surface impedance, which connects the tangential projections (relative to the surface) of the alternating electric and magnetic fields on the surface of the metal:

$$\frac{4\pi i\omega}{c^2}\mathcal{E}_\mu(0) = \sum_{\nu=1}^{2} Z_{\mu\nu}\mathcal{E}_\nu'(0).$$

(The derivatives are expressed in terms of the magnetic field on the surface.) The real and the imaginary parts of the impedance $Z_{\mu\nu} = R_{\mu\nu} + iX_{\mu\nu}$ can be measured directly. To calculate $Z_{\mu\nu}$ it is necessary to know the conditions for the reflection of the electrons from the surface of the metal. A distinction is made between diffuse and specular reflections.† A simple expression for the impedance exists[62] in the limit of the highly anomalous skin effect, $kv \gg |\omega + i/\tau|$. In this case $R_{\mu\nu}$ and $X_{\mu\nu}$ are simultaneously reduced to the principal axes

$$Z_{\mu\mu} = p\left(\frac{\sqrt{3}\,\pi\omega^2}{c^2 B_{\mu\mu}}\right)^{1/3}(1 - i\sqrt{3}), \qquad (38)$$

where the quantity

$$B_{\mu\nu} = \frac{8e^2}{3h^3}\int_0^{2\pi} \frac{n_\mu n_\nu}{K(\varphi)}\, d\varphi$$

is determined by the characteristics of the electron spectrum. The integration in $B_{\mu\nu}$ is along a line lying on the Fermi surface and on which the normal projection (with respect to the metal surface) of the electrons is zero; n is a unit vector in the velocity direction, $K(\varphi)$ is the Gaussian curvature of the Fermi surface, $p = 8/9$ for specular reflection, and $p = 1$ for diffused reflection. The depth of the skin layer can be expressed in this case in terms of the impedance δ
~ $Zc^3(3\sqrt{3}\,\pi^2)^{-1/3}/\omega$.

The surface impedance of bismuth was measured in [63] at 23.5 GHz and in [64] at 9 GHz. These measurements do not agree if the difference in the frequency is accounted for with the aid of (38). An estimate of the total area of the Fermi surface was made in [64] on the basis of the experimental data. Assuming specular reflection, it turned out to be $(17.6 \pm 5.1) \times 10^{13}$ cm^{-2}; if the reflection is assumed diffuse, then the value is $(25.0 \pm 7.3) \times 10^{13}$ cm^{-2}. This agrees with the estimate on the basis of the ellipsoidal model, $S = (3 \times 4.2 + 5.6) \times 10^{13}$ cm^{-2} = 18.2×10^{13} cm^{-2}. According to more accurate data,[19] $S = (3 \times 4.8 + 6.6) \times 10^{13}$ cm^{-2} = 21×10^{13} cm^{-2}. To reconcile the data with the results of measurements de Haas–van Alphen effect, it was necessary in [63] to double the number of the ellipsoids. A more likely explanation of the noted contradictions is found in [64]: formula (30) is apparently not valid at the frequency used in [63], and it is necessary to take into account the finite mean free path of the carriers.

3. Cyclotron Resonance

The character of the dependence of the surface impedance on the constant magnetic field in the centimeter band is shown in Fig. 12, which is taken from [65]. The quantum oscillations (Q.O.) were already discussed. The magnetoplasma oscillations (M.O.) are connected
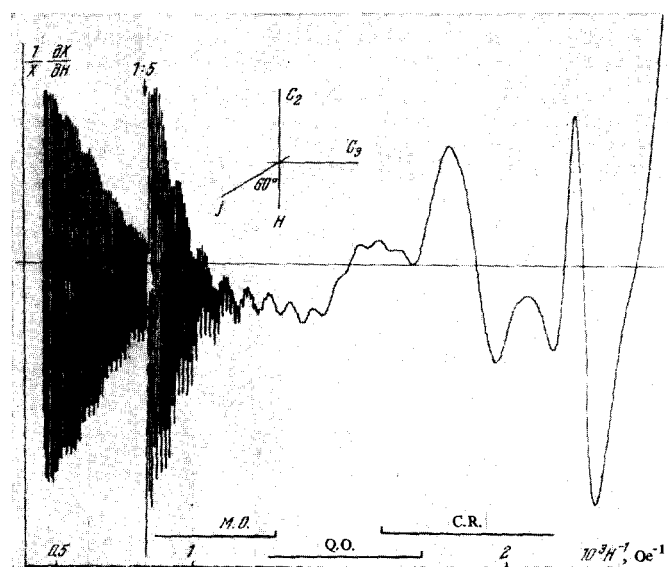
FIG. 12. Surface impedance (logarithmic derivative of the imaginary part) of a bismuth place [65]. Magnetic field; binary and trigonal axes lie in the plane of the surface of the sample.

---

*To avoid misunderstandings we note that the propagation of a field in a metal is possible at frequencies $\omega > \omega_n/\sqrt{\epsilon^0}$ and not $\omega > \omega_n$.

† It is quite frequently assumed that the reflection in bismuth is specular, on the basis of the fact that the wave length of the electron or the hole is much larger than the characteristic inhomogeneities of the surface. But sight is lost in this case of the intervalley transitions, which are accompanied by a large (of the order of atomic) change of the momentum.

Table II. Cyclotron masses m/m_0

Electrons			Holes		
H ∥ C_1	H ∥ C_2	$\not\approx$ H, $C_3 \approx 6°$ *)	H ∥ C_3	H ⊥ C_3	
0.009+10%	0.13+10%		0.068+10%	0.25+10%	61
0.0078	0.119		0.04	0.15	64
0.0091±0.0009	0.14±0.02	0.11±0.01	0.067±0.007	0.226±0.02	67
0.0081±0.0001	0.120±0.003	0.086±0.002	0.063±0.001	0.203±0.004 **)	20
	0.137±0.003	0.117±0.004	0.064±0.003	0.210±0.004 ***)	20
				0.220±0.002 ***)	26

*) Maximum value of mass in the plane H $\perp C_2$
**) Extremal masses
***) Limiting masses

with standing magnetoplasma waves. They will be considered in the next subsection. We turn now to cyclotron resonance (C.R.). Cyclotron resonance yielded the most complete and the most accurate information on the Fermi surface in bismuth. It was possible to trace with its aid the dependence of the masses on the direction of the magnetic field in all the angle regions under consideration. The observations are usually carried out in the centimeter band, using a sample that serves as part of a resonator. The constant magnetic field is varied and one measures the Q or the shift of the natural frequency, which are respectively proportional to the real or to the imaginary parts of the impedance.

Cyclotron resonance in a magnetic field perpendicular to the surface was observed in [61]. Figure 13, which is taken from [61], shows the dependence of the absorption coefficient on the intensity of the magnetic field. The use of a circularly-polarized high frequency field has made it possible to distinguish between three types of carrier. An exact determination of the values of the magnetic field corresponding to resonance (they are indicated by arrows in Fig. 13) is quite difficult. In the ideal case ($\omega\tau \to \infty$), one group of carriers) the absorption appears when $\Omega \geq \omega$. Inasmuch as bismuth has four groups of carriers, and $\omega\tau$ ranged from 3 to 30 in this experiment, the resonant field was determined by extrapolation, and its value was then made more precise by using a theory in which τ was assumed to be isotropic and the spectrum was assumed to be quadratic. The corresponding theoretical curve is shown also in Fig. 13. Among the disadvantages of the method is that it cannot be used to measure the angular dependence of the masses, since each field direction calls for a specially grown sample. The values of the cyclotron

masses obtained in [61] and in other investigations of cyclotron resonance are listed in Table II.

Cyclotron resonance in a parallel field was predicted theoretically by Azbel' and Kaner.[54] To be able to observe it, it is necessary, first, to satisfy the conditions for the anomalous skin effect and second, to have the diameter of the trajectory larger than the skin layer. In this case, an electron interacts with the external electromagnetic field as it moves along the trajectory only in a narrow surface layer. Since the time of the complete revolution can span an arbitrary number of periods of the operating field, the impedance oscillates as a function of the constant magnetic field. The resonance occurs at $\omega = r\Omega$. In the general case, the cyclotron mass $m(\varepsilon, p_z)$, which determines the cyclotron frequency Ω, depends on ε and p_z. Since the only electrons capable of absorbing energy are those located on the Fermi surface, the value of ε is determined by the Fermi level ε_F. As to the dependence on p_z, in the presence of a continuous spectrum of resonant frequencies, resonance is observed only on the boundaries of the spectrum, that is, at the extremal and limiting masses. The latter correspond to the limits of the admissible interval of p_z. The requirement with respect to $\Omega\tau$ is less stringent here —oscillations are observed when $\Omega\tau \lesssim 1$. It is convenient to record experimentally the dependence on H^{-1} of the real or imaginary part of the impedance, or of their derivatives with respect to H, since the distance between the two succeeding minima (or maxima)

$$\Delta H^{-1} \equiv H_r^{-1} - H_{r-1}^{-1} = \frac{e}{mc\omega}$$

does not depend on the magnetic field (Fig. 14). The change of the polarization of the alternating surface currents relative to the direction of the constant magnetic field makes it possible to distinguish between the extremal (j ⊥ H) and limiting masses (j ∥ H).[20] No comparison with the theory[54] has been made so far, and the analysis is usually confined to a determination of the masses.*

Cyclotron resonance in a parallel field was investigated in [64, 20, 67]. Figure 15 shows the angular dependence of the extremal and limiting cyclotron masses. As a rule, the experimental points fit quite well the curves plotted in accordance with the ellipsoidal model. In [67] an attempt was made to determine the deviations from the ellipsoidal model by means of the angular dependence of the masses. Although this deviation did not ex-
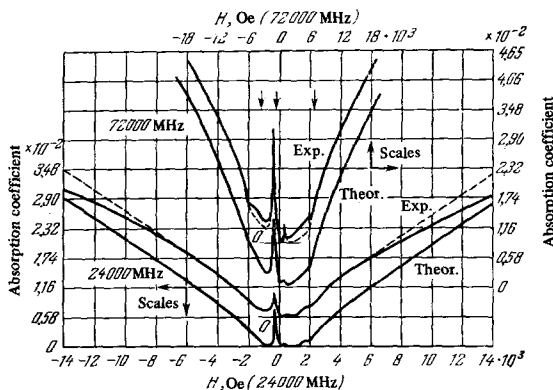

FIG. 13. Cyclotron resonance in a perpendicular field.

---

*See [66] concerning the cyclotron-resonance line shape.

FIG. 14. Cyclotron resonance in parallel field [20].

FIG. 15. Polar diagram of the dependence of the extremal and limiting cyclotron masses on the direction of the magnetic field H [20]. A) H is in the C_2C_1 plane; B') H is in the plane perpendicular to the direction of the elongation of one of the electron "ellipsoids"; C) H is in the C_1C_2 plane. The masses of the electrons on the central section—○, at the limiting point—●. Hole masses on the central section—△, at the limiting point—▲. The curves were plotted in the ellipsoid approximation from the principal values of the masses.

ceed the experimental error, the author of [67] favored the model of [23]. However, the deviation of the values measured in [67] from the ellipsoidal model is more likely connected with an inaccurate determination of the tensor α_{ik}. The most reliable proof of the deviation of the electron spectrum from the ellipsoidal form (14) was obtained in [20] and is shown in Fig. 15. This is shown, first, by the difference between the extremal masses and the limiting masses, and second, by the deviation of the symmetry of the curve from that corresponding to an ellipse (see also Fig. 10). From the measured values of the cyclotron masses it is possible to determine the tensor $\tilde{\alpha}_{ik}/E'$. To this end it is convenient to use the formulas

$$m_1 = m_0 E' (\tilde{\alpha}_{22}\tilde{\alpha}_{33})^{-1/2}, \quad m_2 = m_0 E' (\tilde{\alpha}_{33}\tilde{\alpha}_{11})^{-1/2}, \quad m_3 = m_0 E' (\tilde{\alpha}_{11}\tilde{\alpha}_{22})^{-1/2},$$

which connect the values of the tensor α_{ik} at the principal axes of the ellipsoid with the principal values of the cyclotron masses. According to the data of [20], in the case when the magnetic field is perpendicular to the binary axis and to the direction of elongation of the ellipsoid, the extremal value of the mass is $m_{3,e} = (0.086 \pm 0.002)m_0$ and the mass at the limiting point is $m_{3,l} = (0.117 \pm 0.004)m_0$. In the case of m_2 it is possible to take the value for **H** lying in the bisector direction, since the inclination of the ellipsoid, 6°20', does not come into play here. No difference between the extremal and the limiting masses is observed for this direction. Thus, we get

$$\tilde{\alpha}_{11}^e/E' = 172, \quad \tilde{\alpha}_{22}^e/E' = 0.785, \quad \tilde{\alpha}_{33}^e/E' = 88.5,$$
$$\tilde{\alpha}_{11}^l/E' = 145, \quad \tilde{\alpha}_{22}^l/E' = 0.505, \quad \tilde{\alpha}_{33}^l/E' = 105.$$

The values designated "e" and "l" were calculated respectively from the extremal and limiting masses. The difference between them can serve as a measure of the deviation of the Fermi surface of ellipsoidal. On the crystal axes we have

$$\tilde{\alpha}_{xx}^e/E' = 172, \quad \tilde{\alpha}_{yy}^e/E' = 1.85, \quad \tilde{\alpha}_{yz}^e/E' = +9.61, \quad \tilde{\alpha}_{zz}^e/E' = 87.4,$$
$$\tilde{\alpha}_{xx}^l/E' = 145, \quad \tilde{\alpha}_{yy}^l/E' = 1.78, \quad \tilde{\alpha}_{yz}^l/E' = +11.4, \quad \tilde{\alpha}_{zz}^l/E' = 104.$$

There is some confusion in the literature concerning the sign of $\tilde{\alpha}_{yz}$. To be specific, it is necessary to indicate the direction of the trigonal axis which is taken to be positive. A positive value of α_{yz} corresponds to the choice of the coordinates shown in Fig. 7.

4. Magnetoplasma Waves

Inasmuch as the transverse components of the resistance increase with the magnetic field, the depth of penetration of the alternating field into the metal also increases. Because of this, when $\omega < \Omega$ there can propagate in bismuth electromagnetic waves that attenuate weakly if $\omega\tau \gg 1$, something which is possible in the absence of the magnetic field only when $\omega > \omega_n \varepsilon_0^{-1/2}$. It is simplest to explain the properties of these waves in the limiting case of strong magnetic fields $\Omega \gg \omega_n v/c$. The form of the conductivity tensor can be obtained here without making any assumptions concerning the spectrum of the electrons, and the spatial dispersion is insignificant.[68] If the direction of the constant magnetic field is chosen to be along the z axis, then

$$\sigma_{\alpha\beta} = \frac{e^2\omega}{2\pi^2\hbar^3 i} \int \frac{dp_z}{m\Omega^2} \begin{cases} \overline{(p_y - \bar{p}_y)^2}, & \alpha = \beta = x, \\ \overline{(p_x - \bar{p}_x)^2}, & \alpha = \beta = y, \\ -\overline{(p_x - \bar{p}_x)(p_y - p_y)}, & \alpha = x, \\ & \beta = y; \end{cases}$$

$$\sigma_{z\alpha} = \frac{e^2}{2\pi^2\hbar^3} \int \frac{dp_z}{\Omega} \begin{cases} -\overline{v_z(p_y - \bar{p}_y)}, & \alpha = x, \\ \overline{v_z(p_x - \bar{p}_x)}, & \alpha = y, \end{cases}$$

$$\sigma_{zz} = \frac{e^2 i}{2\pi^2\hbar^3\omega} \int dp_z m\bar{v}_z^2,$$

$$\sigma_{\alpha z} = -\sigma_{z\alpha}, \quad \sigma_{yx} = \sigma_{xy}.$$

(39)

The bar denotes averaging over the trajectories, and the integration with respect to p_z is over the Fermi surface. We then use Maxwell's equations

$$\mathbf{j} = \frac{ic^2}{4\pi\omega} [\mathbf{k} [\mathbf{k}\mathscr{E}]],$$

in which we neglect the displacement current. Assuming that $\Omega \ll \omega_n \varepsilon_0^{-1/2}$, we find that two electromagnetic

Table III. Velocities of magnetoplasma waves

	S-waves (42), uH^{-1}, 10^4 cm/sec-Oe		P-waves (41), $uH^{-1}/\cos\theta$, 10^4 cm/sec-Oe	
	Experiment *) [68,26]	Theory **) [68]	Experiment [68,26]	Theory [68]
H || C_2, k || C_3	1.85	2.2$_h$	2.6	2.7$_e$
H || C_2, k || C_1	2.3	2.7$_e$		
H || C_1, k || C_3	2.0	2.2$_h$	5.55	6.7$_h$
H || C_1, k || C_2	6.0	6.7$_h$		

*) Measurement error approximately 20%
**) Calculated with the ellipsoidal model, "e" ("h")—the main contribution is made by electrons (holes).

waves can propagate in the metal. The connection between the frequency and the wave vector in each wave is linear:

$$\omega = ku_{\pm}(H). \tag{40}$$

The phase velocities of the waves $u_{\pm}(H)$ are proportional to the constant magnetic field and depend on its orientation, and also on the direction of propagation relative to the crystallographic axes. In order of magnitude, $u \sim c\Omega/\omega_n$. In the general case, both waves are excited in a coupled fashion. For each wave, the projection of the electric wave parallel to the constant field is smaller by a factor ω/Ω than the two other projections. If $k \perp H$ one of the waves cannot propagate. If the angle θ between k and H is close to a right angle (the region of applicability of the written equations below is actually broader, since for most orientations the off-diagonal components of (39) are small), then the spectrum of the waves is given for the formulas:

$$k^2 c^2 = 4\pi i\omega \,|\det\sigma\,|/(\sigma_{yy}\sigma_{zz} + \sigma_{yz}^2), \tag{41}$$

$$(kc\cos\theta)^2 = 4\pi i\omega\,(\sigma_{yy} + \sigma_{yz}^2/\sigma_{zz}). \tag{42}$$

In this case the electric field in the wave (41) is directed predominantly perpendicular to k and H, and in the wave (42) it lies in a direction perpendicular to H in the plane of k and H. The attenuation of the waves as a result of collisions of electrons and holes can be taken into account qualitatively by adding $i/2\tau$ in the spectrum (40). Since the complex refractive index kc/ω is proportional to H^{-1}, the absorption of energy in the bulk metal increases in proportion to H (see Fig. 14). In a thin plate, standing waves are produced. Their observation makes it possible to determine the velocities $u_{\pm}$ and to compare them with the calculated values. Such a comparison was made in [68] (see also [69,70]) for different propagation directions. The components of the conductivity tensor (39) were calculated in the ellipsoidal model from the measured values of the masses and cross sections (Table III).

With decreasing magnetic field, at a specified frequency, the wave vector of the wave increases, and when $\Omega \approx \omega_n v/c$ it becomes necessary to take the spatial dispersion into account. One of the most interesting is the question of the limiting value of the magnetic field, below which the propagation of the waves is impossible. The damping mechanism is analogous to the well-known Landau damping for the case of a plasma: intensive energy transfer from the wave to the electrons occurs when the phase velocity of the wave coincides with the drift velocity of the electrons. Since the electrons drift in the direction of the magnetic field and the wave velocity decreases with decreasing field, the projection of the phase velocity of the wave on the direction of the field

coincides at a certain value of the magnetic field with the maximum possible electron velocity projection $v_{z\,max}$. As a rule $v_{z\,max}$ is the velocity at the limiting point of the Fermi surface. The condition $\omega = k_z v_{z\,max}$ determines the limiting value of the magnetic field.

It is possible to distinguish in an experiment between the Landau damping[26] and the damping due to cyclotron resonance.[69,26] The latter turns out to be shifted as a result of the Doppler effect.

Electron drift causes cyclotron resonance in the field of the wave to occur at a frequency ω satisfying the condition $\Omega = \omega - k_z v_z$. Since the resonance condition is satisfied in the entire interval of the velocity projections, from $v_{z\,min}$ to $v_{z\,max}$, the waves cannot propagate in the corresponding interval of magnetic field values. In the experiment, owing the large anisotropy of the Fermi surface, it is possible to separate the contributions made to the conductivity by the electrons and the holes, and to observe all types of damping. By measuring the frequency, the limiting values of the field, and the wave vector of the wave (from the standing waves in a thin plate) it is possible to determine the velocity on the Fermi surface, and also the cyclotron mass at the limiting point. In [26] they obtained $v_F = (2.35 \pm 0.15) \times 10^7$ cm/sec and $m_{lim} = (0.220 \pm 0.002)m_0$ for holes at $H \perp C_3$, and $v_F = (1.13 \pm 0.05) \times 10^8$ cm/sec for electrons at $H \parallel C_2$. We note that the cyclotron mass of the holes at the limiting point differs from the mass on the central section at the same magnetic field direction, $m_{exp} = (0.203 \pm 0.004)m_0$,[20] thus evidencing a certain deviation of the hole spectrum from quadratic. This difference agrees quantitatively with the theory[9] (see the estimate on page 7). There is neither Landau damping nor damping due to the cyclotron resonance when $k \perp H$. This case was investigated experimentally in [71]. If $k \perp H$, as already noted, one wave can propagate, and its electric field lies in a plane perpendicular to the constant magnetic field. For the same reason, the absorption coefficient increases with the magnetic field only in the case of a polarization $\mathcal{E} \perp H$. The absorption coefficient measured in [71] was compared with a theory in which the electron spectrum was assumed to be quadratic, the relaxation time was assumed isotropic, and no account was taken of spatial dispersion. The most doubtful is the latter assumption, since $k \to \infty$ at the limits of the region of the wave existence. On the whole, the agreement with the theory is satisfactory. The limit of the wave existing at large values of H lies at the frequency of the so-called hybrid electron–hole resonance. In weaker magnetic fields, we have $k^2 < 0$, corresponding to strong damping of the wave. With further decrease of H, k^2

vanishes. This is the "dielectric anomaly." Subsequently k^2 increases and tends to ∞ on approaching the electron–electron hybrid resonance, after which it reverses sign. Wave propagation in the field interval between the electron–electron hybrid resonance and the dielectric anomaly was observed in [72].

5. Optical Properties in the Infrared Region

The dielectric constant of bismuth in the optical frequency region was investigated in [8] on the basis of the theory of [9]. For frequencies higher than $\omega_n v/c$, the spatial dispersion is negligible and

$$\text{Re } \varepsilon_{ik} = \varepsilon_{ik}^0 - \frac{(\omega_n^2)_{ik}}{\omega^2}. \tag{43}$$

The value of $(\omega_n^2)_{ik}$ is given by (37), and the frequency-independent part of ε_{ik}^0 is determined by the virtual transitions of the electrons between close bands and can be expressed in terms of the spectrum parameters (19) and (20). A weakly damped electromagnetic wave can propagate in the region where Re ε_{ik} is positive. The wave damping is determined by the electron mean free path, that is, by the electron and electron–phonon scattering. The corresponding Im ε_{ik} is smaller than Re ε_{ik} by a factor $\omega\tau$, where $\tau \sim \hbar\varepsilon^0/\gamma$. The wave propagating on the surface of the metal is absorbed in the metal within a distance

$$c \sqrt{\varepsilon^0}/\omega \text{ Im } \varepsilon \sim \hbar c \sqrt{\varepsilon^0}/\gamma \sim 10^{-3} \text{ cm.}$$

At higher frequencies, one more absorption mechanism comes into play, namely, direct transitions of the electrons from the valence band $\varepsilon_0(p)$ into the conduction band $\varepsilon_\alpha(p)$ with absorption of a photon. Since the photon momentum is small compared with the electron momentum, the electron momentum is conserved. The corresponding absorption limit is determined by the lowest frequency for which

$$\hbar\omega = \varepsilon_a(p) - \varepsilon_0(p). \tag{44}$$

In the two-band model (14) $\hbar\omega_{\min} = \varepsilon_g + 2\varepsilon_F$. In this case the condition (44) is satisfied at $\omega = \omega_{\min}$ for all p lying on the Fermi surface, and Im ε_{ik} changes jumpwise. In the general case, the contribution made to Im ε_{ik} by the interband transitions increases smoothly, starting with $\omega_{\min}$. However, for the central part of the Fermi surface, the approximation (14) is good enough, therefore the growth of Im ε_{ik} is rapid and the condition $\hbar\omega_{\min} = \varepsilon_g + 2\varepsilon_F$ determines satisfactorily the limiting value of the frequency. At large ω the behavior of Im ε_{ik} depends on the peculiarities of the band structure.

Figure 16, which is taken from [73], shows the transmission coefficient for a plate of 4×10^{-3} cm thick as a function of the wavelength λ_0 in vacuum. The start of
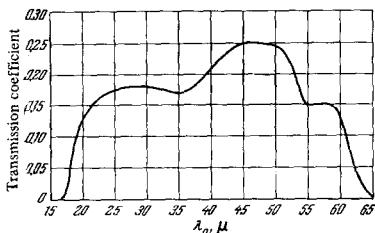

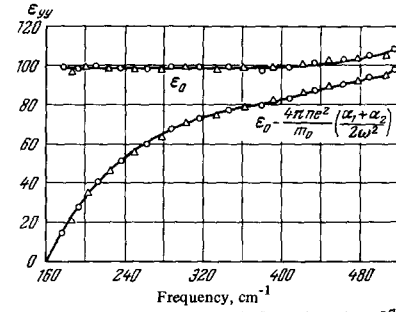FIG. 16. Coefficient of transmission in the infrared region along the binary axis of a bismuth plate [73].

FIG. 17. Dielectric constant in the infrared region [7]. The figure shows an expression for ε_{yy} (in the ellipsoidal approximation) determined essentially by electrons. By a_1 and a_2 we mean here a_{xx} and a_{yy} (3) or $\tilde{a}_{xx}/E'$ and $\tilde{a}_{yy}/E'$ (14).

the transparency at 65 μ and its growth at 55 μ are apparently connected with the vanishing of Re ε_{zz} and Re ε_{yy}. Starting with these frequencies, waves can propagate in which the electric field is directed along the trigonal and bisector axes, respectively. The connection between the frequency and the wave vector inside the plate is given by the relations

$$(ck/\omega)^2 = \varepsilon_{zz}, \quad (ck/\omega)^2 = \varepsilon_{yy}.$$

The frequencies corresponding to the dielectric anomalies can be determined also from the minima of reflection of the radiation of the surface of a bulky sample.[74] To be sure, the positions of the minima are determined by the condition $(ck/\omega)^2 = 1$, but owing to the small value of ε_{ik}^0 the values obtained are close: $\lambda_0 = 53.6$–53.5 μ for Re ε_{zz} and $\lambda_0 = 63.3$–63.2 μ for Re ε_{xx}.[7] In [7] there were observed standing waves in a plate of thickness 14.7 μ, making it possible to measure the $k(\omega)$ dependence for a wave propagating in the direction of the trigonal axis, and to calculate ε_{yy}^0 and $(\omega_n^2)_{yy}$ (Fig. 17). The abrupt drop of the transparency at $\lambda_0 \sim 20$ μ (see Fig. 16) is due to the beginning of direct electron transitions.

Oscillations of the reflection coefficient in a magnetic field were observed in [21]. To explain this phenomenon, the authors of [21] considered the transitions of the electron from the valence band to the conduction band within the framework of the model (26). If we assume the selection rules $\Delta r = 0$ and $\Delta m = \pm 1$ for the spin quantum number (the reflection coefficient was measured at almost normal incidence in a magnetic field perpendicular to the surface), then we obtain for the frequency of the observed photon the expression

$$\hbar\omega = \frac{1}{2}[\varepsilon_g^2 + 4\varepsilon_g(r+1)\hbar\tilde{\Omega}]^{1/2} + \frac{1}{2}[\varepsilon_g^2 + 4\varepsilon_g r\hbar\tilde{\Omega}]^{1/2}. \tag{45}$$

At a specified frequency, the condition (45) determines the values of the magnetic field (corresponding to the integer values of r) at which the reflection coefficient has maxima. Analyzing the dependence of the resonance frequencies (45) on the magnetic field, it is possible to obtain ε_g^2 and $\varepsilon_g/\tilde{m}$. This procedure was used to determine the gap $\varepsilon_g = 0.015 \pm 0.002$ eV and the values of the cyclotron masses $\tilde{m}$ on the bottom of the band at magnetic-field directions parallel to the binary and bisector axes. From these values, and also from the cyclotron masses on the Fermi surface (determined in [61]) the Fermi energy, reckoned from the bottom of the

conduction band, was found with the aid of (23) to be $\varepsilon_F = 0.025 \pm 0.005$ eV. This value yields $\varepsilon_g + 2\varepsilon_F = 0.065$ eV (19 μ) for the limit of the direct transitions, which agrees with the value given in [73].

IX. MEASUREMENT OF THE FERMI MOMENTA IN EXPERIMENTS WITH ULTRASOUND AND BY DETERMINING THE CUTOFF OF THE CYCLOTRON RESONANCE

The Fermi momenta in bismuth were determined by two methods: by studying the absorption of ultrasound[75, 76], and with the aid of cyclotron resonance.[77] These measurements are based on formula (20), which connects the diameters of the trajectories in the r- and p-spaces. In the former case, the comparison standard is the wavelength of sound. In measurements of the magnetic field, different numbers of waves are spanned by the diameter of the orbit, and this leads to oscillations of the absorption ("geometric resonance"). By using cyclotron resonance (in a parallel field), the diameter of the orbit is compared with the thickness of the sample. No cyclotron resonance is observed in fields in which the orbit is not fully contained in the sample. In both cases it is necessary to have $\Omega\tau \gg 1$ for the observations. We present the principal dimensions of the ellipsoids measured by the above methods:

electrons.

$$p_1 = (5.4 \pm 0.15)\,10^{-22}\text{ g-cm/sec},^{[77]}$$

$$(5.5 \pm 0.2)\cdot 10^{-22}\text{ g-cm/sec},^{[76]}$$

$$p_2 = (76 \pm 3 \cdot 10^{-22})\text{ g-cm/sec},^{[76]}$$

$$p_3 = (7.3 \pm 0.3)\cdot 10^{-22}\text{ g-cm/sec},^{[76]}$$

holes.

$$p_y = (13.8 + 0.5)\cdot 10^{-22}\text{ g-cm/sec}.^{[76]}$$

All these values agree with those calculated in [27], although only $p_\perp$ was used to determine the parameters of the theory.[9]

We note also an investigation[78] in which ultrasound was used, and which reports the giant oscillations of ultrasound absorption first considered in [79].

X. TUNNEL EFFECT

A study of the tunnel effect makes it possible to determine the position of the singularities in the state density. In interpreting the experimental data of [80], they took into account the singularities at the bottom or top of the bands, although this does not exhaust all the possibilities. The numbers given in [80] do not agree with the data of [21, 33]. This contradiction is eliminated if account is taken of the singularities noted in [9] and called there "open surface." In spite of the fact that there can be no real open surface when account is taken of the higher-order terms in p, nevertheless the rapid variation of the density of states near the corresponding values of ε should remain in force. In this connection, it becomes unnecessary to assume[80] the existence of a narrow gap (of width on the order of the Fermi energy) near the hole maximum.

XI. PHONON SPECTRUM

The elastic properties of a bismuth crystal are described by six elastic moduli (see, for example, [5]).

They were determined from measurements from the speed of sound.[81] A sufficiently complete investigation of the phonon-spectrum of bismuth was carried out with the aid of slow neutrons.[82] It is curious that at the point of the type T, on the boundary of the Brillouin zone, the gaps between the acoustic and optical branches are quite large, on the order of the Debye frequency. At the same time, repeating the reasoning used in connection with the electron spectrum on page 3, we can arrive at the conclusion that these gaps should be small. Actually, the realignment of the electron spectrum due to the deformation leads to a renormalization of the phonon spectrum; this renormalization is large for frequencies on the order of the new Fermi energy (the Debye frequency and the Fermi energy in bismuth are of the same order).

Thus, summarizing on the foregoing, we note that the properties of bismuth are essentially correctly explained by modern theory of metals. At the same time, the information at our disposal can, of course, not be regarded as exhaustive. There is no doubt that in the future the study of phenomena in strong magnetic fields, started by Kapitza more than forty years ago, will continue, as will the investigations of the specific heat and the mechanism of the impurity influence. The number of different applications of bismuth in practice will also increase.

[1] W. S. Boyle, G. E. Smith, Progr. Semicond., vol. 7, London, 1963.

[2] A. N. Kahn, H. P. R. Frederiske, Sol. Stat. Phys., vol. 9, New York, 1959.

[3] A. A. Abrikosov, Zh. Eksp. Teor. Fiz. 44, 1632 (1963) [Sov. Phys.—JETP 17, 1099 (1963)].

[4] P. Cucka, C. S. Barrett, Acta Cryst. 15, 865 (1962).

[5] L. D. Landau and E. M. Lifshitz, Statisticheskaya fizika, Nauka, 1964 [Statistical Physics, Addison-Wesley, 1958].

[6] R. E. Elliott, Phys. Rev. 96, 280 (1954).

[7] W. S. Boyle, A. D. Brailsford, Phys. Rev. 120, 1943 (1960).

[8] A. A. Abrikosov, Zh. Eksp. Teor. Fiz. 44, 2039 (1963) [Sov. Phys.—JETP 17, 1372 (1963)].

[9] A. A. Abrikosov, and L. A. Fal'kovskiĭ, ibid. 43, 1089 (1962) [16, 769 (1963)].

[10] E. P. Bundy, Phys. Rev. 110, 314 (1958).

[11] G. C. Kennedy and R. Newton, in Solids Under Pressure, McGraw-Hill.

[12] P. F. Chester, G. O. Jones, Phil. Mag. 44, 1281 (1953).

[13] N. B. Brandt and N. I. Ginzburg, Zh. Eksp. Teor. Fiz. 39, 1554 (1960) [Sov. Phys.—JETP 12, 1082 (1961)].

[14] N. B. Brandt, Yu. P. Gaĭdukov, E. S. Itskevich, and N. Ya. Minina, ibid. 47, 455 (1964) [20, 301 (1965)].

[15] D. Shoenberg, Proc. Roy. Soc. A170, 341 (1939); Phil. Trans. Roy. Soc. (London) A245, 1 (1952).

[16] L. D. Landau, Addendum to article [15] (1939).

[17] M. B. Blackman, Proc. Roy. Soc. A166, 1 (1938).

[18] S. J. Mase, Phys. Soc. Japan 13, 434 (1958); 14, 584 (1959).

[19] N. B. Brandt, T. F. Dologolenko, N. N. Stupochenko, Zh. Eksp. Teor. Fiz. 45, 1319 (1963) [Sov. Phys.—JETP 18, 908 (1964)].

[20] V. S. Édel'man, M. S. Khaĭkin, ibid. 49, 107 (1965) [22, 77 (1966)]; M. S. Khaĭkin, R. T. Mina, and V. S. Édel'man, ibid. 43, 205 (1962) [16, 146 (1963)].

[21] R. N. Brown, J. G. Mavroides, B. Lax, Phys. Rev. 129, 2055 (1963).

[22] M. H. Cohen, E. I. Blount, Phil. Mag. 5, 115 (1960).

[23] M. H. Cohen, Phys. Rev. 121, 387 (1961).

[24] P. A. Wolff, J. Phys. Chem. Solids 25, 1057 (1964).

[25] N. B. Brandt, A. E. Dubrovskaya, and G. A. Kytin, Zh. Eksp. Teor. Fiz. 37, 572 (1959) [Sov. Phys.—JETP 10, 405 (1960)].

[26] M. S. Khaĭkin and V. S. Édel'man, ibid. 49, 1695 (1965) [22, 1159 (1966)].

[27] L. A. Fal'kovskiĭ and G. S. Razina, ibid. 49, 265 (1965) [22, 187 (1966)].

[28] J. M. Luttinger, Phys. Rev. 119, 1153 (1960).

[29] B. Lax, Bull. Amer. Phys. Soc. 5, 167 (1960).

[30] H. Jones, Proc. Roy. Soc. A147, 396 (1934); A155, 635 (1936).

[31] L. D. Landau and E. M. Lifshitz, Kvantovaya me-khanika, Fizmatgiz, 1963 [Quantum Mechanics, Addison-Wesley, 1958].

[32] G. E. Pikus, Zh. Eksp. Teor. Fiz. 41, 1258 and 1507 (1961) [Sov. Phys.—JETP 14, 898 and 1075 (1962)].

[33] G. E. Smith, G. A. Baraff, J. M. Rowell, Phys. Rev. A135, 1118 (1964).

[34] G. A. Barraff, Phys. Rev. A137, 842 (1965).

[35] I. M. Lifshitz and A. I. Kosevich, Zh. Eksp. Teor. Fiz. 29, 730 (1955) [Sov. Phys.—JETP 2, 636 (1956)].

[36] L. A. Fal'kovskiĭ, ibid. 49, 609 (1965) [22, 423 (1966)].

[37] I. N. Kalinkina and P. G. Strelkov, ibid. 34, 616 (1958) [7, 426 (1958)].

[38] N. E. Phillips, Phys. Rev. 118, 644 (1960).

[39] D. Shoenberg, L. Z. Uddin, Proc. Roy. Soc. A156, 687 (1936).

[40] N. B. Brandt and M. V. Razumeenko, Zh. Eksp. Teor. Fiz. 39, 276 (1960) [Sov. Phys.—JETP 12, 198 (1961)].

[41] E. N. Adams, Phys. Rev. 89, 633 (1953).

[42] V. Heine, Proc. Phys. Soc. A69, 513 (1956).

[43] W. J. de Haas, P. M. van Alphen, Proc. Acad. Sci. Amst. 33, 1106 (1930).

[44] R. B. Dingle, Proc. Roy. Soc. A211, 517 (1952).

[45] Yu. Bychkov, Zh. Eksp. Teor. Fiz. 39, 1401 (1960) [Sov. Phys.—JETP 12, 971 (1961)].

[46] G. G. Grenier, J. M. Reinolds, J. R. Sybert, Phys. Rev. 132, 58 (1963).

[47] E. P. Vol'skiĭ, Zh. Eksp. Teor. Fiz. 46, 2035 (1963) [Sov. Phys.—JETP 19, 137 (1963)]

[48] N. B. Brandt and L. G. Lyubutina, ibid. 47, 1711 (1964) [20, 1150 (1965)].

[49] W. S. Boyle, F. S. L. Hsu, J. E. Kunzler, Phys. Rev. 128, 1084 (1964).

[50] L. Schubnikow, W. J. de Haas, Commun. Phys. Lab. Univ. Leiden., 207a, 207c, 207d, 210a (1930).

[51] L. S. Lerner, Phys. Rev. 127, 1480 (1962); 130, 605 (1963).

[52] R. J. Balcombe, A. M. Forrest, Phys. Rev. 151, 550 (1966).

[53] J. Eckstein, J. B. Ketterson, Phys. Rev. 137, A1777 (1965).

[54] M. Ya. Azbel' and É. A. Kaner, Zh. Eksp. Teor. Fiz. 32, 896 (1957) and 39, 80 (1960) [Sov. Phys.—JETP 5, 730 (1957) and 12, 58 (1961)].

[55] M. Ya. Azbel', ibid. 44, 983 (1963) [17, 667 (1963)].

[56] I. M. Lifshitz, M. Ya. Azbel', and M. I. Kaganov, ibid. 31, 63 (1956) [4, 41 (1957)].

[57] S. Mase, A. W. Lawson, Phys. Rev. 127, 1030 (1962).

[58] H. J. Luretsckke, Acta Cryst. 8, 716 (1955); T. Okada, J. Phys. Soc. Japan 11, 89 (1956).

[59] R. N. Zitter, Phys. Rev. 127, 1471 (1962).

[60] B. Abeles, S. Meiboom, Phys. Rev. 101, 544 (1956).

[61] J. K. Galt, W. A. Yager, F. P. Merritt, B. B. Cetlin, A. D. Brailsford, Phys. Rev. 114, 1396 (1959).

[62] G. E. H. Reuter, E. H. Sondheimer, Proc. Roy. Soc. A195, 336 (1949); A. B. Pippard, Proc. Roy. Soc. A224, 273 (1954); M. I. Kaganov and M. Ya. Azbel', Dokl. Akad. Nauk SSSR 102, 49 (1950).

[63] G. E. Smith, Phys. Rev. 115, 1561 (1959).

[64] J. E. Aubrey, J. Phys. Chem. Solids 19, 321 (1961).

[65] M. S. Khaĭkin, V. S. Edel'man, and R. T. Mina, Zh. Eksp. Teor. Fiz. 45, 826 (1963) [Sov. Phys.—JETP 18, 566 (1964)].

[66] R. G. Chambers, Proc. Phys. Soc. 86, 305 (1965).

[67] Yi Han Kao, Phys. Rev. 129, 1122 (1963).

[68] M. S. Khaĭkin, L. A. Fal'kovskiĭ, V. S. Édel'man and R. T. Mina, Zh. Eksp. Teor. Fiz. 45, 1704 (1963) [Sov. Phys.—JETP 18, 1167 (1964)].

[69] J. Kirsch, Phys. Rev. A133, 1390 (1964).

[70] G. A. Williams, Phys. Rev. A139, 771 (1965).

[71] G. E. Smith, L. C. Hebel, S. J. Buchsbaum, Phys. Rev. 129, 154 (1963).

[72] V. S. Édel'man and M. S. Khaĭkin, Zh. Eksp. Teor. Fiz. 45, 826 (1963) [Sov. Phys.—JETP 18, 566 (1963)].

[73] W. S. Boyle, K. F. Rodgers, Phys. Rev. Lett. 2, 338 (1959).

[74] W. S. Boyle, A. D. Brailsford, J. K. Galt, Phys. Rev. 109, 1396 (1958).

[75] D. H. Reneker, Phys. Rev. 115, 303 (1959).

[76] A. P. Korolyuk, Zh. Eksp. Teor. Fiz. 49, 1009 (1965) [Sov. Phys.—JETP 22, 701 (1966)].

[77] M. S. Khaĭkin and V. S. Édel'man, ibid. 47, 878 (1964) [20, 587 (1965)].

[78] A. M. Toxen, S. Tansal, Phys. Rev. 137, A211 (1965).

[79] V. L. Gurevich, V. G. Skobov, and Yu. A. Firsov, Zh. Eksp. Teor. Fiz. 40, 786 (1961) [Sov. Phys.—JETP 13, 552 (1961)].

[80] L. Esaki, P. J. Stiles, Phys. Rev. Lett. 14, 902 (1965).

[81] Y. Eckstein, A. W. Lawson, D. H. Reneker, J. Appl. Phys. 31, 1534 (1960).

[82] J. L. Yarnell, J. L. Warren, R. G. Wenzel, S. H. Koenig, IBM, J. Res. Dev. 8, 234 (1964).

[83] Yu. V. Kopaev, Fiz. Tverd. Tela 8, 223 (1966) [Sov. Phys. Solid State 8, 175 (1966)].

[84] N. F. Mott, Phil. Mag. 6, 287 (1961).

Translated by J. G. Adashko