The editorial board of the journal Uspekhi Fizicheskikh Nauk announces with deep sorrow the sudden death, on October 9th, 1976, in his seventieth year, of the Editor in Chief of the journal, the outstanding Soviet physicist, Hero of Socialist Labor, Lenin prize winner and State prize winner, Academician

EVGENIĬ KONSTANTINOVICH ZAVOĬSKIĬ.

The electronic energy spectrum of zero-gap semiconductors

B. L. Gel'mont, V. I. Ivanov-Omskiĭ, and I. M. Tsidil'kovskiĭ

A. F. Ioffe Physico-technical Institute, USSR Academy of Sciences, Leningrad and Institute of Physics of Metals of the Urals Center of the USSR Academy of Sciences, Sverdlovsk

Usp. Fiz. Nauk **120**, 337–362 (November 1976)

PACS numbers: 71.30.Mw

CONTENTS

1. Introduction 879
2. Structure of the Energy Bands of Gapless Semiconductors 880
3. Experimental Data that Confirm the Inverted Band Structure of Gapless Semiconductors 881
4. The Energy Spectrum near the Center of the Brillouin Zone as Based on the Hamiltonian of Luttinger 882
5. Peculiarities of Gapless Semiconductors 884
6. Impurity States in Gapless Semiconductors 887
7. Statistics of Carriers in Gapless Semiconductors 890
8. Conclusion 892
Bibliography 892

1. INTRODUCTION

We currently know an entire set of semiconductors having a forbidden band that is identically equal to zero, and they are commonly called gapless semiconductors (GS). Gapless semiconductors form the natural boundary between metals[1)] and semiconductors. They are distinguished from typical semiconductors by the lack of a threshold for creation of electron-hole pairs, and from metals by the substantially lower density of the electron gas.

Two types of GS are possible. The first type includes substances in which the forbidden band proves to be zero owing to fortuitous degeneracy of the conduction band and the valence band. In these semiconductors, the gapless state can be destroyed by any perturbation, including those that don't alter the symmetry of the crystal. Examples of the first type of GS are the alloys Bi_xSb_{1-x},[2] $Pb_{1-x}Sn_xTe$,[3] and $Pb_{1-x}Sn_xSe$[4] when having a certain ratio of components.

GS of the second type show the greatest interest from the standpoint of the characteristics of the gapless state. Here this state arises from the crystal-lattice symmetry. Their conduction and valence bands belong to the same irreducible representation of the symmetry group. That is, the states in these bands have identical symmetry. Therefore the gapless state in GS of the second type can vanish only under perturbations that reduce the crystal-lattice symmetry. Examples are gray tin (α-Sn), which crystallizes in the diamond structure, and the chalcogenides of mercury (β-HgS, HgSe, and HgTe), which crystallize in the sphalerite structure.

This review is devoted to presenting the current ideas on the energy spectrum of the charge carriers in GS's of the second type.[2)] The first part of the review gives

[1)] They can be defined as metals having a point Fermi surface.[1]

[2)] For brevity, we shall henceforth take the term gapless semiconductor to mean only a GS of the second type.

Copyright © 1977 American Institute of Physics

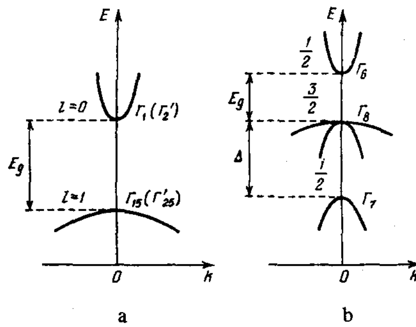

FIG. 1. Band diagram of a semiconductor having the diamond or sphalerite structure: a) without taking account of the spin-orbit interaction; b) with account taken of the spin-orbital interaction (the values $J=1/2$, $3/2$, and $1/2$ are shown).

the pattern of arrangement of the bands in GS's, and describes the energy spectrum within the framework of the local-potential approximation. The second part shows that this approximation is valid only far from the point of degeneracy; for electrons, at an energy above the Bohr energy of an electron, and for holes, at an energy above the Bohr energy of a hole.

2. STRUCTURE OF THE ENERGY BANDS OF GAPLESS SEMICONDUCTORS

As we have noted, the energy spectrum of a GS arises from the symmetry of the crystal. If we compare two semiconductors having the same crystal-lattice symmetry, the orders of arrangement of the energy bands in the "normal" semiconductor and in the GS prove to be mutually inverted. People say that the GS has an inverted band structure. Let us illustrate this with the example of the "normal" semiconductor InSb and the GS HgTe. Both crystals have the sphalerite structure, with similar parameters of the energy spectrum. Let us recall the structure of the bands of InSb near the center of the Brillouin zone (the vicinity of the point Γ), as is schematically shown in Fig. 1.

The conduction band is separated from the valence band by the energy spacing E_g (forbidden band). It is described in the vicinity of the minimum by wave functions having s-symmetry (irreducible representation Γ_6). The valence band near its vertex is described by functions of p-symmetry. Upon accounting for spin, the degeneracy of the conduction and the valence bands would be two- and sixfold, respectively. However, spin-orbit interaction partially removes the degeneracy of the valence band, and the states are no longer classified in terms of the orbital angular momentum **l**, but in terms of the total angular momentum $\mathbf{J}=\mathbf{l}+\mathbf{s}$ (**s** is the spin angular momentum). The conduction band remains doubly degenerate ($J=1/2$). J can adopt two values for states in the valence band: $J=3/2$ and $J=1/2$. The value $J=3/2$ corresponds to bands of light and heavy holes, each of which is doubly degenerate when $k\neq 0$ (k is the wave vector). When $k=0$, the bands with $J=3/2$ are fourfold degenerate. The bands of light and heavy holes belong to the same irreducible representation Γ_8. The value $J=1/2$ corresponds to a valence band that is separated at the point Γ from the two upper bands by the magnitude of the spin-orbital interaction energy Δ. The actual spectrum of the valence band of InSb is somewhat more complicated, owing to the lack of a center of symmetry in the sphalerite structure. Yet the terms arising hereby that are linear in the quasimomentum in the dispersion law for holes contribute insignificantly to the total energy for InSb.[5]

In the inverted band structure that is realized in HgTe, the band of s-symmetry lies below the bands having $J=3/2$, and it has a negative curvature. The inversion causes the curvature of one of the bands having $J=3/2$ to be positive, while the other is negative (Fig. 2). In this structure, the forbidden band is identically equal to zero, while $E_g<0$, if, as before, we measure energy from the top of the valence band, and denote by E_g the energy gap between the extremal points of the band having s-symmetry (irreducible representation Γ_6 in HgTe or Γ_7^- in α-Sn) and of that having $J=3/2$ (irreducible representation Γ_8 in HgTe or Γ_8^+ in α-Sn). The fact that the forbidden band equals zero for this band model is not fortuitous, since all four bands having $J=3/2$ belong to a single irreducible representation. Therefore a forbidden band can arise only from an external agent that lowers the crystal-lattice symmetry, e.g., uniaxial compression.

Groves and Paul[6] proposed an inverted band structure in 1963 for α-Sin in order to explain the contradictions that arose at that time between different experimental data. One group of experiments indicated that the width of the forbidden band of α-Sn is very small; according to other experiments, the effective mass of the electrons is of the same order of magnitude as in other diamondlike semiconductors. Yet we know from theory (see. e.g.[7]) that the effective mass is proportional to the width of the forbidden band if the transition between the valence band and the conduction band is direct and allowed. Here the proportionality coefficient varies but slightly within the group of diamondlike semiconductors. Therefore, with a normal band structure, a small width of the forbidden band entails a proportional decrease in the effective mass of the electrons. The inverted structure proposed by Groves and Paul eliminates this contradiction. Actually, the mass of the electrons is mainly determined by the interaction of the bands having s- and p-symmetries, which are separated by an energy gap of size E_g. This is positive for the

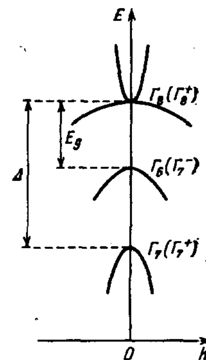

FIG. 2. Inverted band diagram.

normal band structure, and negative for the inverted band structure. This is just why the effective mass of the electrons differs from zero in semiconductors having a zero forbidden band.

What are the reasons for the inversion of the bands? If we trace the variation of the mutual arrangement of the bands in the series C (diamond)–Si–Ge–α-Sn, all elements of which have the same crystal-lattice symmetry, we note the following tendency. As the atomic number increases from diamond to tin, the direct energy gap at the point Γ, i.e., the spacing between the bands $E_g = E(\Gamma_7^-) - E(\Gamma_8^+)$ diminishes. Simultaneously, relativistic effects play a greater role, and they act in the same direction. The noted tendency for the bands Γ_7^- and Γ_8^+ to approach can ultimately invert them.

The calculations imply that the onset of the inverted band structure in α-Sn substantially involves the relativistic effects. Actually, it has been shown[8] that the order of the bands in α-Sn should be normal within the framework of the nonrelativistic self-consistent orthognalized plane-wave (OPW) method. Therefore corrections to the potential were introduced into the scheme of calculations that permitted them to adjust certain reference points of the energy spectrum to the experimental data.[9] The relativistic self-consistent OPW calculations[9,10] for α-Sn led to an inverted band structure. The calculated value $E_g = -0.416$ eV proved to be close to the experimental value -0.413 eV.[11] The relativistic OPW method was used in[10] to calculate the energies at the high-symmetry points of the Brillouin zone. The relationship between the energy and the wave vector at the intermediate points was determined by using the $\mathbf{k}\cdot\mathbf{p}$ method. Here the matrix elements of the transitions were the adjustment parameters. It turned out that the agreement between the experimental and theoretical values of the matrix elements of the transition between the Γ_7^- and Γ_8^+ bands is poorer than for the E_g values: the experimental value is 1.69 atomic units, and the theoretical is 1.13 atomic units.

The studies[8–10] are important in that they clearly imply that the inverted band structure in α-Sn, and perhaps in the other GS, arises from relativistic effects.

Most of the other calculations of the energy spectrum of a GS have been performed by using empirical methods: either the empirical pseudopotential method for α-Sn[12,13] and HgTe and HgSe,[14,15] or the Korringa-Kohn-Rostoker method for HgTe and HgSe.[16] The energy at the points of high symmetry in HgTe, HgSe, and β-HgS has also been calculated by the empirical OPW method.[17]

3. EXPERIMENTAL DATA THAT CONFIRM THE INVERTED BAND STRUCTURE OF GAPLESS SEMICONDUCTORS

Let us briefly take up the experimental data that attest to the inverted band structure of the GS. First of all, as we have noted, an inevitable consequence of the inverted band structure is that the forbidden band is of size zero. Even the first experimental studies of the conductivity and the Hall coefficient in α-Sn[18] and HgTe[19–22] indicated that their forbidden bands are small. As for HgTe and HgSe, hypotheses were later advanced of even a small overlap of the conduction band and the valence band.[23–25] All of the estimates of the size of the forbidden band and the depth of overlap were made in the cited studies under various assumptions. Yet the conclusion was general that the studied materials have an unusually small gap between the conduction and valence bands.

An important characteristic of a GS is the form of the temperature-dependence of the concentration of charge carriers. In an intrinsic GS at $T=0$, the Fermi level passes through the point of contact of the bands. At finite temperature, if $m_n \neq m_p$ (m_n amd m_p are the effective masses of the electrons and holes), the Fermi level must lie in one of the allowed bands. When $m_p \gg m_n$, the Fermi level lies in the conduction band, and the electrons are degenerate while the holes are nondegenerate. Then we can write the following expressions for the concentration n of electrons and p of holes:

$$n = \frac{(2m_n E_F)^{3/2}}{3\pi^2\hbar^3}, \quad p = 2\left(\frac{m_p T}{2\pi\hbar^2}\right)^{3/2} \exp\left(-\frac{E_F}{T}\right), \tag{1}$$

where E_F is the Fermi energy. Upon equating these expressions, we can find the Fermi energy as a function of the temperature for an intrinsic GS:

$$\frac{E_F}{T} = \ln\left[\frac{3\sqrt{\pi}}{4}\left(\frac{m_p}{m_n}\right)^{3/2}\right] + \frac{3}{2}\ln\frac{T}{E_F}. \tag{2}$$

Upon neglecting the second term on the right-hand side of (2), we get with logarithmic accuracy when $m_p \gg m_n$:

$$E_F = T\ln\left[\frac{3\sqrt{\pi}}{4}\left(\frac{m_p}{m_n}\right)^{3/2}\right]. \tag{3}$$

Thus the Fermi energy E_F is proportional to T. Then, according to (1), $n \propto T^{3/2}$. The purer the studied GS's are, the greater the accuracy of this law and the greater the temperature range over which it holds.[18,26–28] Moreover, the temperature independence of the reduced Fermi level that is implied by Eq. (3) has been found experimentally upon studying the thermo-emf in pure enough crystals of HgTe.[26]

Evidence favoring the inverted band structure of the GS's is furnished by some interesting experiments involving hydrostatic pressure and uniaxial compression. Hydrostatic pressure does not affect the nature of the temperature-dependence of the electron concentration,[29] since the zero forbidden band arises from the symmetry of the crystal. Omnidirectional pressure does not alter this symmetry, but only alters the effective mass of the electrons.[30] In contrast to omnidirectional pressure, uniaxial pressure lowers the symmetry of the crystal, and thus removes the degeneracy of the valence band and of the conduction band. That is, it gives rise to a forbidden band, as has been confirmed by measurements[31] of the Hall coefficient under pressure along the [100] and [111] axes in α-Sn.

The most convincing proofs of the inverted-band model have been furnished by magneto-optical studies of GS's, the results of which are presented in Chap. 4. Here we shall only mention that oscillations of the reflection coefficient in a magnetic field were observed experimentally[32] for the first time in HgTe, and they involve two types of interband transitions between the Landau subbands of the different bands: $\Gamma_8 \to \Gamma_6$ and $\Gamma_8 \to \Gamma_8$. They showed that the frequencies of the transitions of the first type also approach zero upon extrapolating the magnetic field to zero, in accord with the zero width of the forbidden band. We should note that the properties of semiconductors in a magnetic field are generally highly sensitive to the symmetry of the band states. For example, in InSb the electrons, whose band possesses s-symmetry, are diamagnetic, whereas in HgTe the electrons are paramagnetic.[33,34] Studies of the magnetic susceptibility have pictorially revealed[35] the generality of the nature of the conduction band of HgTe and the band of light holes in InSb and Ge.

A fuller and more detailed review of the experiments that have served to prove the inverted band structure of the GS can be found in the monograph[36].

4. THE ENERGY SPECTRUM NEAR THE CENTER OF THE BRILLOUIN ZONE AS BASED ON THE HAMILTONIAN OF LUTTINGER

The spectrum of the free charge carriers near the point of degeneracy of the bands of a GS can be determined from the Hamiltonian of Luttinger.[37] We shall be interested in small values of k, i.e., the effective-mass approximation. The Hamiltonian in the effective-mass approximation can be found by using the $\mathbf{k}\cdot\mathbf{p}$ method. Yet it is simpler to derive the Hamiltonian from symmetry considerations, as Luttinger[37] has done for the valence band of Ge, which is quadruply degenerate at the point $k=0$.

For a nondegenerate band, the only invariant that is quadratic in k is k^2. Therefore the Hamiltonian has the form $\mathcal{H}=\hbar^2k^2/(2m^*)$, where the coefficient m^* is the effective mass of an electron. Degenerate bands possess another vector $\mathbf{J}$ (the total angular momentum) that characterizes the state, in addition to the vector $\mathbf{k}$. We are interested in the value $J=3/2$. The components J_x, J_y, and J_z are 4×4 matrices. Therefore, in addition to the invariant k, which is the only possible one for a simple band, there is also a spherically-symmetrical invarant $(\mathbf{k}\cdot\mathbf{J})^2$ and the invariant $k_x^2J_x^2+k_y^2J_y^2+k_z^2J_z^2$, which satisfies cubic symmetry. Correspondingly, the Hamiltonian of Luttinger has the form

$$\mathcal{H}=\frac{\hbar^2}{m_0}\left[\frac{1}{2}\left(\gamma_1+\frac{5}{2}\gamma_2\right)k^2-\gamma_3(\mathbf{kJ})^2+(\gamma_3-\gamma_2)(k_x^2J_x^2+k_y^2J_y^2+k_z^2J_z^2)\right], \tag{4}$$

where m_0 is the mass of a free electron. The last term describes the corrugation of the isoenergetic surfaces. The constants γ_1, γ_2, and γ_3 determine the spectrum of free charge carriers, which can be found by diagonalizing the matrix of (4). The numerical matrices J_x, J_y, and J_z are given in the book[7]. After diagonalization, we obtain the dispersion law for $E(k)$:

$$E(k)=\frac{\hbar^2\gamma_1}{2m_0}k^2\pm\frac{\hbar^2}{m_0}\sqrt{\gamma_2^2k^4+3(\gamma_3^2-\gamma_2^2)(k_x^2k_y^2+k_y^2k_z^2+k_z^2k_x^2)}. \tag{5}$$

In semiconductors having a small $|E_g|$, such as InSb or HgTe, the major role is played by the interaction of bands having s- and p-symmetries. This interaction strongly diminishes the anisotropy of the isoenergetic surfaces. Since we are interested in the characteristic features involving the absence of a forbidden band, we shall henceforth neglect the corrugation (i.e. we shall assume that $\gamma_2=\gamma_3=\gamma$) as well as the small terms linear in k that should be added to the Hamiltonian of (4) for semiconductors lacking a center of symmetry.

As we see from Eq. (5), the Hamiltonian of Luttinger describes either an ordinary semiconductor having two types of holes, or a GS, depending on the relationship between γ_1 and γ. If $\gamma_1>2\gamma$, then the spectrum of (5) corresponds to light and heavy holes (with energy measured into the interior of the valence band), and the constants γ_1 and γ are related as follows to the masses of light and heavy holes:

$$m_l=m_0(\gamma_1+2\gamma)^{-1},\quad m_h=m_0(\gamma_1-2\gamma)^{-1}. \tag{6}$$

Yet if $2\gamma>\gamma_1$, then the Hamiltonian of (4) gives the spectrum of a GS having the following mass of electrons and holes:

$$m_n=m_0(\gamma_1+2\gamma)^{-1},\quad m_p=m_0(-\gamma_1+2\gamma)^{-1} \tag{7}$$

(we assume that the energy of the electrons is positive, while that of the holes is negative.)

In order to find the spectrum near the point of degeneracy in the magnetic field $\mathbf{H}$, we must add to the Hamiltonian of (4) two new invariants $(\mathbf{J}\cdot\mathbf{H})$ and $J_x^3H_x+J_y^3H_y+J_z^3H_z$, while we introduce the quantity $\mathbf{K}=-i\nabla+(e/c\hbar)\mathbf{A}$ in place of the wave vector $\mathbf{k}$. Here $\mathbf{A}=[\mathbf{H}\times\mathbf{r}]/2$ is the vector potential in the magnetic field. Consequently, the Hamiltonian of Luttinger in the magnetic field takes on the form

$$\mathcal{H}=\frac{\hbar^2}{2m_0}\left(\gamma_1+\frac{5}{2}\gamma_2\right)K^2-\frac{\hbar^2\gamma_2}{m_0}\sum_{\alpha=1}^{3}K_\alpha^2J_\alpha^2-\frac{\hbar^2\gamma_3}{m_0}\sum_{\alpha\neq\beta}\{K_\alpha K_\beta\}\{J_\alpha J_\beta\}+\frac{\hbar e}{m_0c}\varkappa(\mathbf{JH})+\frac{\hbar eq}{m_0c}\sum_{\alpha=1}^{3}J_\alpha^3H_\alpha, \tag{8}$$

where $\{AB\}=(AB+BA)/2$, and $\varkappa$ and q are Luttinger's new parameters. The values of the constants γ_1, γ_2, γ_3, and $\varkappa$ for various materials have been determined from light-absorption or -reflection experiments in a strong magnetic field. The peaks of the reflection or absorption coefficient as functions of the magnetic field intensity H were correlated with transitions between the levels, as determined from the Hamiltonian of (8). We give below in Table I the experimental values of the Luttinger parameters for the GS's. For comparison, we give the Luttinger parameters for Ge and InSb.

As we see from the table, the parameters obtained by different authors for HgTe differ rather strongly. The studies performed in[45] have shown that the combination-resonance lines correspond better to a net of levels

TABLE I.

	γ_1	γ_2	γ_3	$\varkappa$		γ_1	γ_2	γ_3	$\varkappa$
Ge[38]	13.2	4.10	5.62	3.29	HgTe[41]	12.8	8.4	8.4	10.5
InSb[39]	33.51	14.48	15.65	13.47	HgTe[42]	14.8	8.6	8.6	10.2
α-Sn[40]	15.14	11.40	8.02	{ 13.11	HgTe[43,44]	16.8	10.4	9.2	11.2
				11.84		14.8	10.9	8.7	11.7

having the Luttinger parameters from[43], whereas the cyclotron-resonance lines agree better with the Luttinger parameters from[41]. However, as the level diagram implies, the positions of the cyclotron-resonance peaks given in[45] are less sensitive to the choice of parameters than are the positions of the recombination-resonance peaks. One of the possible reasons for the large disagreement of the Luttinger parameters given in the literature is that the energy of the lower Landau levels is a small difference of large quantities.

In[41,42], the Luttinger parameters were determined while neglecting the corrugation of the isoenergetic surfaces. The parameter q is anomalously small in diamondlike semiconductors (e.g., $q = 0.4$ for InSb[5]), and we shall henceforth neglect it.

Luttinger[37] showed that if we neglect the corrugation of the isoenergetic surfaces ($\gamma_2 = \gamma_3 = \gamma$), there exists an operator

$$Q = \frac{e}{c\hbar}(\mathbf{JH}) + \frac{1}{2}K^2,$$

that commutes with the Hamiltonian of (8). This circumstance permits us to find the energy levels in a magnetic field. When $k_H = 0$, where k_H is the projection of the wave vector on the magnetic-field direction, there are two non-equidistant systems of levels. In line with[37], we shall introduce the quantum number $n \geqslant 0$, which is the eigenvalue of the operator Q. For the first series of levels a, the energy spectrum has the form

$$\left.\begin{aligned} &E_a(n) = \frac{\hbar eH}{m_0 c}\varepsilon_a(n),\\ &\varepsilon_a(0) = \frac{1}{2}(\gamma_1 - \gamma - \varkappa), \quad \varepsilon_a(1) = \frac{1}{2}(3\gamma_1 - 3\gamma - \varkappa),\\ &\varepsilon_a^{\pm}(n) = \gamma_1\left(n - \frac{1}{2}\right) - \gamma + \frac{1}{2}\varkappa \pm \sqrt{\left[\gamma\left(n - \frac{1}{2}\right) - \gamma_1 + \varkappa\right]^2 + 3\gamma^2 n(n-1)} \end{aligned}\right\} \tag{9}$$

for $n \geqslant 2$. The energy of the series of levels b is determined by the formulas

$$\left.\begin{aligned} &E_b(n) = \frac{\hbar eH}{m_0 c}\varepsilon_b(n),\\ &\varepsilon_b(0) = \frac{1}{2}(\gamma_1 + \gamma - 3\varkappa), \quad \varepsilon_b(1) = \frac{3}{2}(\gamma_1 + \gamma - \varkappa),\\ &\varepsilon_b^{\pm}(n) = \gamma_1\left(n - \frac{1}{2}\right) + \gamma - \frac{1}{2}\varkappa \pm\\ &\qquad \pm \sqrt{\left[\gamma\left(n - \frac{1}{2}\right) + \gamma_1 - \varkappa\right]^2 + 3\gamma^2 n(n-1)} \end{aligned}\right\} \tag{10}$$

for $n \geqslant 2$.

If the fundamental contribution to the parameters γ_1, γ, and $\varkappa$ comes from the interaction with the nearest s-band, then the relationship holds that $\gamma_1 = 2\gamma = 2\varkappa$. Then the energy spectrum takes on an especially simple form. The energy of the holes has an infinite-fold degeneracy at $E = 0$, which is not removed even when $k_H \neq 0$. The energy of the electron levels is

$$\begin{aligned} &E_a(n, k_H) = \frac{\hbar eH}{m_n c}\left(n + \frac{1}{4}\right) + \frac{\hbar^2 k_H^2}{2m_n},\\ &E_b(n, k_H) = \frac{\hbar eH}{m_n c}\left(n + \frac{3}{4}\right) + \frac{\hbar^2 k_H^2}{2m_n}, \end{aligned} \tag{11}$$

where $m_n = m_0/4\gamma$.

According to (9)–(11), an energy gap arises in a GS in a strong magnetic field between the levels of the conduction band and the valence band. In the case where $m_p \gg m_n$, it is equal to a good approximation to $\hbar eH/(4m_n c)$. The onset of a forbidden band in a magnetic field is illustrated by magnetoresistance experiments in HgTe.[46,47] Figure 3 shows the curves for the relationship of the longitudinal and transverse magnetoresistance to the magnetic field intensity for an HgTe specimen having a Hall coefficient in zero magnetic field $R_0 = -4\times10^3$ cm^3/C, and a Hall mobility of 7×10^5 cm^2/V·sec at 4.2 °K.[46] One observes a sharp growth that is close to exponential in the longitudinal magnetoresistance, which involves the formation of a forbidden band. Upon further increase in the magnetic field, it changes into a slower increase in the resistance. In this magnetic-field region, the longitudinal magnetoresistance is practically equal to the transverse value, as is characteristic of conduction by impurities whose levels have "frozen" out of the allowed bands into the forbidden band that is created by the magnetic field. While at low temperatures the electron concentration depends substantially on the concentration of impurities existing in the specimen, intrinsic conduction plays a greater role with increasing temperature, and one can estimate quantitatively the size of the forbidden band that arises in the magnetic field. Figure 4 shows the longitudinal magnetoresistance curves of HgTe at various temperatures from[47], where the authors estimated the size of the forbidden band at 64° and 77 °K. The size of the gap proved to depend on the temperature, being 7.6 $\hbar eH/(m_0 c)$ at 64 °K, and 8.2 $\hbar eH/(m_0 c)$ at 77 °K. The uncertainty of the parameters of the bands currently hinders an exact quantitative comparison. Let us compare the experimental value of 8.2 $\hbar eH/(m_0 c)$ at 77 °K with the value calculated from the data of[43,44], where they gave a value of E_g at 77 °K. If we assume that the corrections to the Luttinger parameters that account for the interaction with the higher bands depend weakly on the temperature, we find that the width of the forbidden band in a magnetic field lies in the range (6.4–8.3) $\hbar eH/(m_0 c)$.

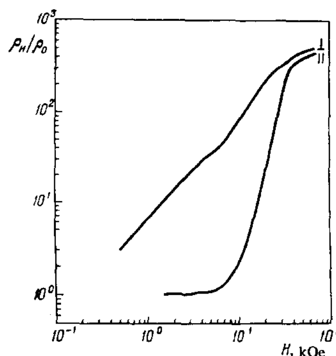

FIG. 3. Experimental curves of the longitudinal (∥) and transverse (⊥) magnetoresistance of HgTe as functions of the magnetic field strength at 1.7 °K.[46]

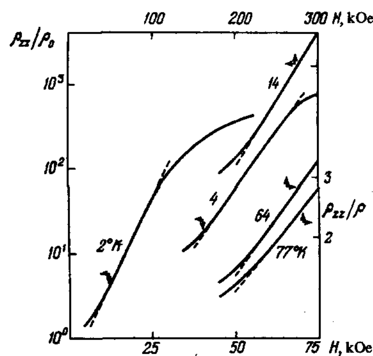

FIG. 4. The longitudinal resistance as a function of the magnetic field for HgTe at different temperatures.[47] The arrows indicate the scale for the given curve.

They calculated in[41,48] the curves for the relationship of the energy of the levels of HgTe and α-Sn in a magnetic field to k_H. The energy of the hole levels of series a for $n \geq 2$ is a non-monotonic function of k_H. The energy has a minimum at $k_H = 0$, and two symmetrical maxima.

5. PECULIARITIES OF GAPLESS SEMICONDUCTORS

a) The material in Chap. 4 can create the impression that the whole difference of a gapless semiconductor from an ordinary one having a quadruply degenerate valence band in terms of mathematical description consists only in the different relationship of the constants γ_1 and γ. This conclusion holds only to the extent that the spectrum can be described within the framework of the one-particle approximation using the local potential (we recall that the Luttinger Hamiltonian was derived by the $\mathbf{k} \cdot \mathbf{p}$ method for a particle moving in the field of the crystal potential). However, in contrast to an ordinary semiconductor, effects involving electron-electron interaction are very substantial in a GS. Energy is not required for creating electron-hole pairs, since the forbidden band is zero. The possibility of virtual creation of excitons leads both to altering the dielectric properties of the GS, and to a substantial renormalization of the spectrum of the electrons, and especially, that of the holes.

Let us first discuss the peculiarities of the dielectric constant of a GS. Sherrington and Kohn[49] have calculated the interband absorption coefficient and the corresponding contribution to the dielectric constant at $T = 0$. In an ordinary semiconductor, the contribution to the conductivity σ at frequencies somewhat above the threshold of interband absorption is proportional to $\sqrt{\hbar\omega - E_g}/\omega$ for allowed transitions.[50] In a GS, the conduction band and the valence band have the same parity at $\mathbf{k} = 0$. Therefore the transition is forbidden for $\mathbf{k} = 0$, and $\sigma \propto (\hbar\omega - E_g)^{3/2}/\omega$.[50] If we assume that $E_g = 0$, we find that $\sigma \propto \sqrt{\omega}$. This is Sherrington and Kohn's result. The only characteristic quantity having the dimensions of energy in a GS is the Bohr energy of an electron (or hole). The real component of the dielectric constant is of the same order of magnitude as the imaginary component. According to[49], when $m_p \gg m_n$,

$$\varepsilon(\omega) = \varepsilon_0 + \frac{e^2\sqrt{2m_n}}{\hbar^{3/2}\sqrt{\omega}}(1+i). \tag{12}$$

Here ε_0 is the contribution to the dielectric constant from all the bands except the Γ_8 bands. Equation (12) was derived in the random-phase approximation without accounting for the non-parabolicity and corrugation of the isoenergetic surfaces. We shall discuss the limits of applicability of this formula somewhat later.

The presence of free charge carriers (of impurity origin or arising from thermal excitation from the valence band) has the effect that $\mathrm{Im}\,\varepsilon(\omega)$ differs from zero only when $\omega > E_F/\hbar$, owing to the Moss-Burstein effect. When $\omega = E_F/\hbar$, $\mathrm{Re}\,\varepsilon(\omega)$ has a logarithmic singularity instead of a radical one.[51]

The discussed peculiarities of a GS prove to be essential in analyzing phenomena that involve interaction of radiation with free charge carriers[49,51] or with the crystal lattice.[52,53] The excitation of plasma oscillations or optical phonons by action of an electromagnetic wave occurs on a background of interband transitions that open up an additional decay channel for plasmons and polarized phonons, as described by the imaginary component of Eq. (12). We note that the contribution of interband transitions to the real component of the dielectric function can be treated as an additional mechanism of shielding of polarized optical phonons that must be taken into account in analyzing an electron-phonon interaction in a GS, e.g., scattering of electrons by optical phonons, magnetophonon resonance, etc. It is also evident that this situation must be taken into account in determining the effective charge of the ions in a GS.

The need of accounting for the dielectric anomaly of a GS in studying optical phenomena in the fundamental lattice absorption region was first shown[53] in analyzing ir reflection experiments in HgTe. This has later been demonstrated in greater detail in[54]. An effect of the frequency-dependence of the dielectric constant of a GS on the reflection spectrum in the region of the plasma minimum has been found experimentally in α-Sn.

The charge shielding is determined by the static dielectric constant $\varepsilon(q)$, where q is the wave vector. Account taken of virtual interband transitions leads[56] to an expression for $\varepsilon(q)$ that is analogous to (12), in which one must replace $\hbar\omega$ by $\hbar^2 q^2/2m_n$:

$$\varepsilon(q) = \varepsilon_0 + \frac{3\pi e^2 m_n}{2\hbar^2 q}. \tag{13}$$

The shielding radius is of the order of the Bohr radius of an electron, but the potential of the charge declines considerably more slowly at large distances than in ordinary Debye shielding. According to[56], the potential of a shielded charge center declines as $1/r^2$ at distances large in comparison with the Bohr radius of an electron.

If the GS contains donors, then the electrons are degenerate (we shall discuss the role of impurities in a GS in Chap. 6). Owing to the Moss-Burstein effect, transitions can occur only at energies larger than the Fermi energy E_F. Therefore, the static dielectric con-

stant $\varepsilon(0)$ is finite in a doped GS, and it increases with decreasing concentration of electrons[57]:

$$\varepsilon(0)=\varepsilon_0+\frac{8m_n e^2}{\pi\hbar^2 k_F}. \tag{14}$$

Here the Fermi momentum $\hbar k_F$ is related in the usual way to the electron concentration: $k_F^3=3\pi^2 n$.

We can easily note that Eqs. (12)–(14) have similar structures. They involve the ratio of the Bohr energy of an electron $E_B=m_n e^4/2\hbar^2\varepsilon_0^2$ to the characteristic energy. In the case of $\varepsilon(\omega)$, the characteristic energy is of the order of $\hbar\omega$; the characteristic energy for $\varepsilon(q)$ is $\hbar^2 q^2/m_n$; and in a doped semiconductor, the characteristic energy is the Fermi energy $E_F=\hbar^2 k_F^2/2m_n$. A gap of the order of $\hbar e H/(4m_n c)$ arises in an external magnetic field, and accordingly, the static dielectric constant becomes a finite quantity that is inversely proportional to $\sqrt{H}$.[58]

The unusual behavior of the dielectric constant and its sensitivity to external parameters must be reflected in the kinetic properties of GS's. If, for example, one calculates the mobility of electrons in HgTe at low temperatures by the Brooks-Herring formula for scattering by charged centers, then the calculated quantities prove to be smaller than the experimental.[59] In this case, the discrepancy of theory and experiment cannot be explained by an underestimate of any scattering mechanisms, since the theory strongly overestimates the scattering power of the charged centers, even when these mechanisms are not taken into account. An analogous but less marked discrepancy between theory and experiment occurs also in α-Sn[60] and in HgSe.[61] The experimental results were obtained with specimens that contained rather large amounts of impurities ($\gtrsim 10^{15}$ cm^{-3}). Hence the refined calculation of the scattering cross-section of the charged centers was performed by using Eq. (14). It turned out that taking account of the dependence of the dielectric constant on the electron concentration as well as of the angular dependence of the amplitudes of the Bloch wave functions arising from the degeneracy of the bands led to relatively good agreement of theory with experiment.[62]

Thus, taking account of electron-electron interaction alone in GS's permits one more or less successfully to explain the kinetic and optical properties. Yet we must not forget that all the presented formulas have been derived within the framework of the random-phase approximation (RPA). Hence the calculations of the mobility or the absorption coefficient that are performed within the framework of the RPA can serve only to indicate qualitatively the scale and the direction in which the electron-electron interaction affects the properties of the GS. It makes no sense to achieve quantitative agreement, since if the dielectric anomaly is small, it plays no substantial role. Yet if it is large, the RPA loses applicability. Actually Eq. (14), with which the mobility was calculated, is only the first term of an expansion in powers of $\sqrt{E_B/E_F}$. Therefore, whenever an additional term that involves virtual transitions becomes substantial, the expansion ceases to hold. The value of the concentration in the purest currently available specimens of HgTe is precisely that which makes $E_B\sim E_F$. At energies lower than the E_B energy that characterizes the interaction of an electron with a hole in a pure GS ($m_p\gg m_n$), many-particle effects determine the properties of the pure GS.

The lack of a parameter in terms of which one could reduce the calculation to a number compelled Abrikosov and Beneslavskiĭ[1,63] to use the scaling theory[95] that had been developed in the theory of phase transitions. The fundamental results of[1,63] can be formulated as follows. The automodeling solution for the Fourier component of the overall Coulomb interaction has the form

$$\Gamma=\frac{e^2}{\varepsilon_0 q_0^2}\left(\frac{q}{q_0}\right)^{\nu-3} d(x). \tag{15}$$

Here $q_0=m_n e^2/\hbar^2\varepsilon_0$, and $x=(q_0/q)^{\nu}\, m_n\omega/\hbar q_0^2$. The exponent ν satisfies the inequality $1<\nu<3$. The limiting values of the unknown function $d(x)$ are chosen from the condition that Γ depends only on q when $\omega\to 0$, and only on ω when $q\to 0$. Thus, $d(0)\sim 1$, and $d(x\to\infty)\propto x^{1-(3/\nu)}$. The charge-carrier spectrum is determined by the poles of the one-particle Green's function. According to[1], the automodeling behavior of the latter can be represented in the form

$$G=\frac{m_n}{\hbar^2 q_0^2}\left(\frac{q}{q_0}\right)^{-\alpha} g(x). \tag{16}$$

Here g is an unknown function having the boundary conditions $g(0)\sim 1$, $g(x\to\infty)\propto x^{-\alpha/\nu}$, and α is a constant that satisfies the inequalitites $\alpha<2$ and $\alpha<\nu$. Unfortunately, the scaling theory gives excessively broad limits for possible variation of the physical quantities. The constants α and ν must be determined from experiment. Subsequently Abrikosov[64] calculated the critical exponents α and ν by using two methods that are used in the theory of phase transitions (expansion in terms of dimensionality and expansion in terms of a "large number of components"). Both methods gave qualitatively similar results in a first approximation. Except for a very small neighborhood near the point of degeneracy (of the order of $2.5\times 10^{-4} E_B$), the values of the critical exponents proved to be equal: $\alpha=\nu=2$. Thus, according to[64], the RPA gives the correct dependence of the dielectric constant on ω and q, apart from a very narrow region near the point of degeneracy.

b) Now let us study the effect of the electron-electron interaction on the energy spectrum of the carriers in a GS. As we have noted, the Hamiltonian of Luttinger[37] was derived within the framework of the $\mathbf{k}\cdot\mathbf{p}$ method, which is valid for an electron that lies in the periodic field of the local potential. However, even in the one-particle approximation described by the Hartree-Fock equation, the exchange interaction cannot always be written in the form of a local potential. This is justified only the case of a completely filled valence band that is separated from the conduction band by a gap. Here we neglect interband transitions. In ordinary semiconductors, the local-potential approximation generally describes the energy spectrum well. In the case of a GS, we face a situation in which the nonlocal nature of the

potential becomes substantial. In essence, we are dealing in a pure GS with a half-filled band, since the conduction band and the valence band belong at the point Γ to a single irreducible representation. The exchange interaction is accounted for in the Schrödinger equation for an electron in the one-particle approximation by the additional term

$$\mathscr{H}'_{\alpha\beta}\chi_\beta(\lambda, \mathbf{k}) = -\frac{e^2}{(2\pi)^3\varepsilon_0}\int\frac{d^3r_1}{|\mathbf{r}-\mathbf{r}_1|}\sum_{\mu=\pm 3/2}\int d^3q\,[e^{i\mathbf{k}(\mathbf{r}_1-\mathbf{r})}-1]e^{i\mathbf{q}(\mathbf{r}-\mathbf{r}_1)} \times \chi^*_\beta(\mu, \mathbf{q})\,\chi_\alpha(\mu, \mathbf{q})\,\chi_\beta(\lambda, \mathbf{k}). \tag{17}$$

Here μ and λ give the numbers of the states, and α and β are the numbers of the columns of the eigenfunction χ of the Hamiltonian of (4). The summation is performed over all the occupied states. The second term in the brackets has been subtracted in order to make the integral converge. This term is a part of the local potential. We can assume that the same term has been added to the lattice potential. The functions $\chi(\lambda, \mathbf{k})$ are spinors that correspond to spin 3/2, and λ is the helicity of the particle.[65] The functions with helicity $\lambda = \pm 3/2$ describe hole states, while those with $\lambda = \pm 1/2$ describe electron states.

In a pure GS, we must sum in (17) over the states of the valence band. Upon using the relationship

$$\sum_{\mu=\pm 3/2}\chi_\alpha(\mu, \mathbf{q})\,\chi^*_\beta(\mu, \mathbf{q}) = \Lambda^{(p)}_{\alpha\beta}(\mathbf{q}) = \frac{1}{2}\left[\frac{(\mathbf{q}\mathbf{J})^2}{q^2}-\frac{1}{4}\right]_{\alpha\beta}, \tag{18}$$

where $\Lambda^{(p)}$ is the operator for projection on the states of the valence band, we obtain from (17)

$$\mathscr{H}'_{\alpha\beta}(\mathbf{k}) = \frac{3\pi e^2}{32\varepsilon_0}k\left[\frac{5}{4}-\frac{(\mathbf{k}\mathbf{J})^2}{k^2}\right]_{\alpha\beta}. \tag{19}$$

Thus the exchange interaction leads to additional terms that are linear in k in the spectrum of the charge carriers. The one-particle Hamiltonian of a pure GS with account taken of the exchange interaction has the form

$$\mathscr{H} = \frac{\gamma_1\hbar^2}{2m_0}k^2 + \left[\frac{5}{4}-\frac{(\mathbf{k}\mathbf{J})^2}{k^2}\right]\left(\frac{\gamma\hbar^2}{m_0}k^2+\frac{3\pi e^2}{32\varepsilon_0}k\right). \tag{20}$$

Equation (20) has been written in the spherical approximation. We can easily find from (20) the energy spectrum of the carriers:

$$E_n(k) = \frac{\hbar^2k^2}{2m_0}(\gamma_1+2\gamma)+\frac{3\pi e^2}{32\varepsilon_0}k, \quad E_p(k) = \frac{\hbar^2k^2}{2m_0}(2\gamma-\gamma_1)+\frac{3\pi e^2}{32\varepsilon_0}k. \tag{21}$$

The existence of terms linear in k in the electron and hole spectra was first demonstrated by Halperin and Rice.[66] Abrikosov[67] has made a detailed analysis of the spectrum for a large hole-electron mass ratio. For electrons, the renormalization of the spectrum that involves the terms linear in k is substantial only when $E < E_B$, i.e., in the region where many-electron effects become substantial. At the same time, the hole spectrum is renormalized owing to the nonlocal nature of the potential over a far wider range of energies up to the Bohr energy of a hole $(m_p/m_n)E_B$. Thus we conclude that the Hamiltonian of Luttinger, which was derived within the framework of the local-potential theory, and which usually describes well the valence bands of light and heavy holes in a "normal" semiconductor, cannot be used in its previous form of a GS, even in the one-particle approximation.

In the presence of donors, when the electron concentration differs from zero, the filling of the electron states affects the one-particle spectrum of the free charge carriers of the GS. Let us study the case of a degenerate gas that occupies the states up to the Fermi energy E_F. In the summation that takes part in (17), we must account for the filling of the electron states. Consequently an additional term appears:

$$\mathscr{H}''_{\alpha\beta}(\mathbf{k})\,\chi_\beta(\lambda, \mathbf{k}) = -\frac{e^2}{(2\pi)^3\varepsilon_0}\int\frac{d^3r_1}{r_1}\sum_{\mu=\pm 1/2}\int_0^{k_F}d^3q\,(e^{i\mathbf{k}\mathbf{r}_1}-1)e^{-i\mathbf{q}\mathbf{r}_1} \times \chi^*_\beta(\mu, \mathbf{q})\,\chi_\alpha(\mu, \mathbf{q})\,\chi_\beta(\lambda, \mathbf{k}). \tag{22}$$

The upper limit of the integral over q equals the Fermi momentum k_F. Upon using the relationship

$$\sum_{\mu=\pm 1/2}\chi_\alpha(\mu, \mathbf{q})\,\chi^*_\beta(\mu, \mathbf{q}) = \Lambda^{(n)}_{\alpha\beta}(\mathbf{q}) = \delta_{\alpha\beta}-\Lambda^{(p)}_{\alpha\beta}(\mathbf{q}), \tag{23}$$

where $\Lambda^{(n)}$ is the operator for projection on the electron states, we obtain from Eqs. (19) and (22):

$$\mathscr{H}'+\mathscr{H}'' = \frac{e^2}{\varepsilon_0}k_F\left\{f_1\left(\frac{k_F}{k}\right)+\left[\frac{5}{4}-\frac{(\mathbf{k}\mathbf{J})^2}{k^2}\right]f\left(\frac{k_F}{k}\right)\right\}. \tag{24}$$

The functions $f(x)$ and $f_1(x)$ have the form

$$f(x) = \frac{1}{8\pi}\left[\frac{5}{4}+\frac{3}{4}x^2-\frac{1}{2}\left(\frac{3\ln x}{x}+x-\frac{7}{4x}+\frac{3x^2}{4}\right)\ln\left|\frac{1+x}{1-x}\right|+\frac{3}{x}\int_x^\infty\frac{dy\ln y}{y^2-1}\right],$$
$$f_1(x) = \frac{1}{2\pi}\left[1-\frac{1}{2}\left(x-\frac{1}{x}\right)\ln\left|\frac{1+x}{1-x}\right|\right]. \tag{25}$$

The one-particle Hamiltonian with account taken of the filling of the electron states can be written as follows:

$$\mathscr{H} = \frac{\hbar^2\gamma_1}{2m_0}k^2+\frac{e^2}{\varepsilon_0}k_Ff_1\left(\frac{k_F}{k}\right)+\left[\frac{5}{4}-\frac{(\mathbf{k}\mathbf{J})^2}{k^2}\right]\left[\frac{\hbar^2\gamma}{m_0}k^2+\frac{e^2}{\varepsilon_0}k_Ff\left(\frac{k_F}{k}\right)\right]. \tag{26}$$

We find from (26):

$$E_n(k) = \frac{\hbar^2k^2}{2m_0}(\gamma_1+2\gamma)+\frac{e^2}{\varepsilon_0}k_F\left[f_1\left(\frac{k_F}{k}\right)+f\left(\frac{k_F}{k}\right)\right],$$
$$E_p(k) = \frac{\hbar^2k^2}{2m_0}(2\gamma-\gamma_1)+\frac{e^2}{\varepsilon_0}k_F\left[f\left(\frac{k_F}{k}\right)-f_1\left(\frac{k_F}{k}\right)\right]. \tag{27}$$

The spectrum of holes and electrons for $k \gg k_F$ can be found as before by Eq. (21). When $k_F \gg k$, the spectrum is quadratic, but the masses of the electrons and holes depend on the concentration of electrons:

$$\frac{1}{m_n} = \gamma_1+2\gamma+\frac{26e^2}{15\pi\varepsilon_0k_F}, \quad \frac{1}{m_p} = 2\gamma-\gamma_1+\frac{2e^2}{5\pi\varepsilon_0k_F}. \tag{28}$$

Interestingly, the linear terms that arise from the nonlocality of the potential can liquidate the overlap of the bands in a GS having the sphalerite structure.

All the formulas written above have pertained to a GS whose lattice contains a center of symmetry. Yet if a symmetry center is lacking, then the following additional terms[7] must appear in the Hamiltonian:

$$-\frac{4\hbar}{\sqrt{3}}\varkappa_0\{k_x\{J_x(J_y^2-J_z^2)\}+k_y\{J_y(J_z^2-J_x^2)\}+k_z\{J_z(J_x^2-J_y^2)\}\},\qquad(29)$$

where $\varkappa_0$ is a constant. Terms of this type lead to overlap of the valence band and the conduction band in a GS having the sphalerite structure. The energy of this overlap is usually small. For example, for InSb the energy involved in the linear terms of the type of (29) is of the order of 5×10^{-6} eV. [5] The maxima of the overlap energy lie in the [111] direction. The values of the energy and the momentum at the points of the maximum are

$$E_{0V}=4m_p\varkappa_0^2,\quad \hbar k_{0V}=2\sqrt{2}\,m_p\varkappa_0.\qquad(30)$$

However, the terms linear in k that involve the nonlocality of the potential were not taken into account in calculating the overlap energy. The overlap of the bands disappeares when we account for these terms if $\varkappa_0<3\pi e^2/64\sqrt{2}\,\varepsilon_0\hbar$. It is more convenient to rewrite this inequality in terms of the overlap energy that figures in the local-potential theory. In order that an overlap should exist, the condition must be satisfied that $E_{0V}>(3\pi/32)^2\,m_p e^4/2\varepsilon_0^2\hbar^2$.

The renormalization of the hole spectrum alters the statistics, e.g., that of the free charge carriers in the GS. Thus, in an intrinsic semiconductor, the electrons are degenerate when $m_p\gg m_n$. According to (3), the Fermi energy in the local-potential case depends linearly on the temperature, while the carrier concentration is

$$n=p=\frac{(2m_nT)^{3/2}}{3\pi^2\hbar^3}\ln^{3/2}\left[\frac{3\sqrt{\pi}}{4}\left(\frac{m_p}{m_n}\right)^{3/2}\right].\qquad(31)$$

When we account for the nonlocality of the potential, we get a different relationship when $(3\pi/16)^2(m_p/m_n)E_B \gg T\gg(3\pi/16)^2E_B\ln[(3/4)(32/3\pi)^3(T/E_B)^{3/2}]$:

$$E_F=T\ln\left[\frac{3}{4}\left(\frac{32}{3\pi}\right)^3\left(\frac{T}{E_B}\right)^{3/2}\right],\qquad(32)$$

$$n=p=\frac{(2m_nT)^{3/2}}{3\pi^2\hbar^3}\ln^{3/2}\left[\frac{3}{4}\left(\frac{32}{3\pi}\right)^3\left(\frac{T}{E_B}\right)^{3/2}\right].\qquad(33)$$

We must take account of the linear terms that arise from the nonlocality of the potential in calculating any of the kinetic coefficients. The nonlocality of the potential also leads to renormalization of the energy spectrum of semiconductors that have a narrow width of forbidden band. The effect was studied in[68] of the exchange interaction on the energy spectrum of the carriers in the case in which Kane's model[69] describes well the spectrum of the electrons and light holes. Here the spectrum of the heavy holes is determined within the framework of the local-potential theory by the interaction with higher bands. The nonlocality of the potential primarily renormalizes the spectrum of the heavy holes. The relationship of the energy E_h of a heavy hole to the wave vector k has the form

$$E_h(k)=\frac{\hbar^2k^2}{2m_h}+\frac{3\pi e^2}{64\varepsilon_0}k+\frac{3eE_g}{16\pi P\varepsilon_0}\int_0^\infty dy\left(1+\frac{8}{3}y^2\right)^{-1/2}\left[-\frac{5}{3}+\frac{y^2E_g^2}{P^2k^2}-\frac{(y^2E_g^2-P^2k^2)^2}{4yP^3k^3E_g}\ln\left(\frac{yE_g+Pk}{yE_g-Pk}\right)^2\right].\qquad(34)$$

Here m_h is the contribution to the mass of a heavy hole from the upper bands, $P=-(i\hbar/m_0)\langle S|\hat{p}_z|z\rangle$, $\hat{\mathbf{p}}$ is the momentum operator, $|S\rangle$ is the function for an electron having s-symmetry at the point Γ, and $|Z\rangle$ is the function for p-symmetry. The energy is measured in (34) into the interior of the valence band.

Equation (34) implies that the spectrum of the heavy holes, just like that of the light holes, is parabolic only in the region of relatively small values of the wave vector $kP\ll E_g$. The nonparabolicity becomes substantial when $kP\gtrsim E_g$, in the same region of values of the wave vector as for the bands of light holes and electrons.

When $kP\ll E_g$, we have

$$E_h(k)=\frac{k^2}{2}\left(\frac{\hbar^2}{m_h}+\frac{2e^2P}{5\pi\varepsilon_0E_g}\sqrt{\frac{2}{3}}\right).\qquad(35)$$

Just like the mass of a light hole, that of a heavy hole declines with decreasing E_g in the parabolic region of the spectrum. In InSb, the correction to the mass of a heavy hole involving nonlocality amounts to 15%. In solid substitutional solutions $Cd_xHg_{1-x}Te$, the mass of a heavy hole is mainly determined by the exchange interaction when $E_g\lesssim 50$ meV.

6. IMPURITY STATES IN GAPLESS SEMICONDUCTORS

a) The statistics of the carriers and the kinetic properties of semiconductors are determined in many ways by the nature and concentration of the impurities present. In an ordinary semiconductor having a forbidden band of finite width, the carriers can give rise to bound states at the impurity atoms. For example, this situation gives rise to the activational type of variation of the conductivity at low temperatures.

In a GS, the impurity states lie in a region of the continuous spectrum of the conduction band (acceptor levels) or of the valence band (donor levels), whereby discrete levels are absent. An electron acquires the capability of transferring into any of the states of the continuous spectrum having the same energy.

Moreover, in the case of a semiconductor having a zero forbidden band, the problem of the possible values of an electron at an impurity is substantially complicated by the fact that it ceases to be a one-electron problem. In a certain sense, it resembles the problem of the deep levels in semiconductors having a forbidden band of finite width. In the case of a shallow level in a semiconductor having a finite forbidden band, the many-electron effects are not substantial, since the correction that they engender is of the order of the ratio of the energy of the shallow level to the width of the forbidden band. If the level is deep, then this ratio becomes of the order of unity, and we must not neglect the many-electron effects. An analogous difficulty is known in quantum electrodynamics. The role of the forbidden band is played there by the energy that is required for creation of an electron-positron pair, $2m_0c^2$, where c is the speed of light. The corrections to the energy of an electron in the field of a nucleus of charge Ze that involve the polarization of the vaccum (i.e., creation of virtual electron-positron pairs), and which give rise

to the so-called Lamb shift of the levels, are of the order of the ratio of the Bohr energy to the energy $2m_0c^2$, i.e., $(Ze^2/\hbar c)^2 \approx (Z/137)^2$. This correction is small for the hydrogen atom ("shallow level"). Yet the solution cannot be restricted to the one-electron problem for large enough Z.

As we have noted, the many-electron effects for a GS become substantial near the point of degeneracy in an energy range of the order of the Bohr energy of an electron. Bearing this in mind, let us study first the one-electron problem of the impurity states, and then analyze which of the results continue to hold when we account for the interelectronic interaction.

If the masses of an electron and of a hole were of the same order, then there would be only one characteristic energy, the Bohr energy of an electron (or of a hole), which would determine both the positions and the diffuseness of the levels. Then the diffuseness of the levels would be of the order of the spacing between them, even in the one-electron approximation, and the density of states would be a smooth function of the energy. It is also evident that the many-electron effects must be substantial when the masses of an electron and a hole are comparable. The situation differs when there is a large difference between the masses of an electron and a hole. If the mass of a hole is much larger than that of an electron, as happens in all known GS's, then the density of states in the conduction band for an acceptor must show sharp peaks. As we shall now show, the relative width of these peaks is proportional to $(m_n/m_p)^{3/2}$. We can estimate the width of a level crudely from perturbation theory. The probability W of transition from the discrete spectrum to the continuous spectrum is proportional to the expression

$$W=\frac{2\pi}{\hbar}|V_a|^2 g(E_0), \tag{36}$$

where $g(E_0)$ is the density of states in the conduction band at the energy $E=E_0$, and V_a is the matrix element of the transition between the states of the discrete and the continuous spectra. In order of magnitude, the energy E_0 of a hole at the acceptor is equal to the Bohr energy of the hole $m_pe^4/2\varepsilon_0^2\hbar^2$. The characteristic radius of the acceptor state is of the order of the Bohr radius of the hole $a_p=\hbar^2\varepsilon_0/m_pe^2$. In the conduction band, the energy E_0 corresponds to the momentum $\hbar k\sim\sqrt{m_nE_0}\sim\hbar/\sqrt{a_na_p}$, where $a_n=\hbar^2\varepsilon_0/(m_ne^2)$ is the Bohr radius of an electron. Since the wavelength of an electron of energy E_0 is much larger than a_p $(a_p\sqrt{m_nE_0}/\hbar\sim(m_n/m_p)^{3/2})$, then in estimating V_a, we can neglect the wave vector of the electron in the oscillatory exponential of the Bloch wave function. Consequently we get the following estimate: $V_a\sim E_0a_p^{3/2}$. Thus the sharpness of the acceptor levels fully involves the density of states in the conduction band. For a parabolic dispersion law of the energy spectrum of the electrons,

$$g(E_0)=\frac{\sqrt{2E_0m_n^3}}{\pi^2\hbar^3}.$$

Equation (34) implies that $W\sim E_0\hbar^{-1}(m_n/m_p)^{3/2}$. The width of the level is

$$\Gamma=\hbar W\sim E_0\left(\frac{m_n}{m_p}\right)^{3/2}, \tag{37}$$

i.e., the relative width is $\Gamma/E_0\sim(m_n/m_p)^{3/2}$.

Sharp donor levels do not exist in a GS having $m_p \gg m_n$. An analogous estimate by perturbation theory shows that the diffuseness of the levels is of the order of their energy. Moreover, the donors alter the density of states at energies of the order of the Bohr energy of an electron. The many-electron effects are substantial at these energies, since the shielding radius in a GS is of the order of the Bohr radius of an electron, even in the absence of free carriers. At the same time, the one-particle approximation proves sufficient for calculating the position and width of an acceptor level since the energy of an acceptor state when $m_p \gg m_n$ is of the order of the Bohr energy of a hole, and it is hence far larger than the Bohr energy of an electron, while the characteristic radius a_p is small in comparison with the Bohr radius a_n of an electron. As the temperature or the donor concentration increases, the charge of the acceptors becomes shielded by free electrons: The shielding radius decreases with increasing electron concentration. However, it is comparable with the Bohr radius of a hole at an acceptor only when the Fermi level $E_F\sim(m_p/m_n)^3E_0$ lies far above the ground acceptor level.

b) Let us proceed to calculate the acceptor level numerically. First we shall treat for comparison a semiconductor having a forbidden band of finite width in which the mass m_h of a heavy hole is large in comparison with the mass m_l of a light hole. In order to derive the Schrödinger equation for a hole in the field of a shallow impurity center, we must add to the Hamiltonian of (4) the potential $V(r)$ of the center. In the spherical approximation, the equation has the form

$$\begin{gathered}[\mathscr{H}_0(\hat{\mathbf{p}})+V(r)]\psi=E\psi,\\ \mathscr{H}_0(\hat{\mathbf{p}})=\frac{1}{2m_0}\left(\gamma_1+\frac{5}{2}\gamma\right)\hat{p}^2-\frac{\gamma}{m_0}(\hat{\mathbf{p}}\mathbf{J})^2.\end{gathered} \tag{38}$$

Here $\hat{\mathbf{p}}$ is the momentum operator. The problem of the acceptor levels with an arbitrary ratio of masses has been solved in many studies, e.g., by using the variational method.[70–75] The system of equations (38), which consists of four equations in partial derivatives, can be reduced to two subsystems (each consisting of two equations) consisting of ordinary differential equations, and then the possible values of E can be determined. We are interested in the values of E for $m_l \ll m_h$. Therefore let us proceed as follows. We shall resolve the wave function into two terms $\psi=\psi^{(l)}+\psi^{(h)}$ in such a way that $\psi^{(l)}$ is composed of the wave functions of the light holes, while $\psi^{(h)}$ is a packet of wave functions of the heavy holes. The expression for the operator $\Lambda^{(h)}(\mathbf{k})$ for projection onto the heavy-hole states coincides with the expression for the operator for projection onto the hole states in a GS in (18), while the formula (23) holds for the operator for projection onto the light-hole states. Thus,

$$\psi^{(l)}(\mathbf{k})=\Lambda^{(l)}(\mathbf{k})\psi(\mathbf{k}),\quad \psi^{(h)}(\mathbf{k})=\Lambda^{(h)}(\mathbf{k})\psi(\mathbf{k}). \tag{39}$$

The Hamiltonian can be represented in the form

$$\mathscr{H}_0(\mathbf{k})=\frac{\hbar^2k^2}{2m_l}\Lambda^{(l)}(\mathbf{k})+\frac{\hbar^2k^2}{2m_h}\Lambda^{(h)}(\mathbf{k}). \tag{40}$$

We can derive the integral equations for $\psi^{(l)}$ and $\psi^{(h)}$ from (38) by multiplying on the left by $\Lambda^{(l)}$ and $\Lambda^{(h)}$. Eq. (40) implies that

$$\mathscr{H}_0(\mathbf{k})\psi^{(l)}(\mathbf{k})=\frac{\hbar^2k^2}{2m_l}\psi^{(l)}(\mathbf{k}),\quad \mathscr{H}_0(\mathbf{k})\psi^{(h)}(\mathbf{k})=\frac{\hbar^2k^2}{2m_h}\psi^{(h)}(\mathbf{k}). \tag{41}$$

As we see from (38) and (41), $\psi^{(l)}\to 0$ in the limit as $m_l\to 0$. Hence the positions of the levels can be found from the equations

$$\left(\frac{\hbar^2k^2}{2m_h}-E\right)\psi^{(h)}(\mathbf{k})-\frac{e^2}{2\pi^2\varepsilon_0}\Lambda^{(h)}(\mathbf{k})\int\frac{d^3q}{(\mathbf{k}-\mathbf{q})^2}\psi^{(h)}(\mathbf{q})=0. \tag{42}$$

As before, Eq. (42) consists in a system of four integral equations. However, we can reduce it to a single integral equation. The function for the ground state is of the form

$$\psi^{(h)}(\mathbf{k})=f(k)\Lambda^{(h)}(\mathbf{k})\chi(\mu). \tag{43}$$

Here $f(k)$ is a scalar function that depends only on the absolute magnitude of k, while $\chi(\mu)$ is a numerical spinor in which any of the four rows in the column differs from zero. The ground state is quadruply degenerate. [7] Upon using (43), we can derive from (42) the equation for $f(k)$:

$$\left(\frac{\hbar^2k^2}{2m_h}-E\right)f(k)-\frac{e^2}{8\pi^2\varepsilon_0}\int\frac{d^3qf(q)}{(\mathbf{k}-\mathbf{q})^2}\left[1+\frac{3(\mathbf{kq})^2}{k^2q^2}\right]=0. \tag{44}$$

Equation (44), when transformed to dimensionless form, contains no parameters of the material. The eigenvalues of this equation have been found, both with a computer (Eq. (44) having been transformed preliminarily into an ordinary differential equation), and by using the variation principle. In the latter case, the trial function was taken to be hydrogen-like with a single variational parameter n:

$$f(k)=\frac{2}{\pi}\left(\frac{\hbar^2\varepsilon_0 n}{m_he^2}\right)^{3/2}\left[1+\left(\frac{\hbar^2\varepsilon_0 nk}{m_he^2}\right)^2\right]^{-2}. \tag{45}$$

We get the following equation for n:

$$E=\frac{m_he^4}{\varepsilon_0^2\hbar^2}\left[\frac{1}{n}\left(6\ln 2-\frac{7}{2}\right)-\frac{1}{2n^2}\right]. \tag{46}$$

Minimization with respect to n gives the value $n=1.52$, which agrees well with the value $n\approx 1.5$ that was obtained in the computer calculation. When $m_h\gg m_l$, the energy of the acceptor ground state in a semiconductor having a forbidden band of finite width is $E_0=2m_he^4/9\varepsilon_0^2\hbar^2$. [76]

A detailed study of the relationship of the wave function to the distance from the impurity center shows that it first declines exponentially with a characteristic scale of the order of the Bohr radius $a_h=\hbar^2\varepsilon_0/m_he^2$ of a heavy hole. Then this exponential decline goes over into a power law, and the wave function declines as $1/r^3$. Finally, when $r\gg a_h$, the power-law decline is replaced by an exponential decline, though the characteristic scale of the decline is now the geometric-mean quantity $\hbar/\sqrt{m_lE_0}\sim\sqrt{a_la_h}$, where $a_l=\varepsilon_0\hbar^2/m_le^2$ is the Bohr radius of a light hole. The existence of the long exponential tail is not significant in solving the problem of the position of the ground-state level, and we can ignore it in the variational calculations. Such tails come to light in the kinetic properties of the semiconductors, in which the overlap of the wave functions of adjacent centers is essential, e.g., in solving the problem of hopping conduction. [77]

In the case of a GS, we must account for the terms linear in k in determining the energy of the quasidiscrete levels in the one-particle approximation. We shall write the equation for the wave function in the field of a Coulombic center in the momentum representation:

$$\mathscr{H}_0(\mathbf{k})\psi(\mathbf{k})+\frac{e^2}{2\pi^2\varepsilon_0}\int\frac{d^3q}{(\mathbf{k}-\mathbf{q})^2}\psi(\mathbf{q})=E\psi(\mathbf{k}). \tag{47}$$

The expression for $\mathscr{H}_0(\mathbf{k})$ has been derived above (see (20). Quasidiscrete states exist only when $m_p\gg m_n$. Analogously to what is done in the problem of the energy levels in a semiconductor having a forbidden band of finite width, we shall resolve the wave function into two parts $\psi=\psi^{(n)}+\psi^{(p)}$, where

$$\psi^{(n)}(\mathbf{k})=\Lambda^{(n)}(\mathbf{k})\psi(\mathbf{k}),\quad \psi^{(p)}(\mathbf{k})=\Lambda^{(p)}(\mathbf{k})\psi(\mathbf{k}).$$

We shall not write explictly the equations for $\psi^{(n)}$ and $\psi^{(p)}$, since the further course of the treatment fully copies the derivation of Eq. (42). Just in the same way, we can neglect the function $\psi^{(n)}$ in the studied case, owing to the smallness of m_n/m_p. The equation for $\psi^{(p)}$ has the form

$$\left(\frac{\hbar^2k^2}{2m_p}+\frac{3\pi e^2}{32\varepsilon_0}k+E\right)\psi^{(p)}(\mathbf{k})-\frac{e^2}{2\pi^2\varepsilon_0}\Lambda^{(p)}(k)\int\frac{d^3q\psi(\mathbf{q})}{(\mathbf{k}-\mathbf{q})^2}=0. \tag{48}$$

Equation (48) differs from (42) in the presence of the term linear in k, which arises from the nonlocality of the exchange interaction. At an energy of the resonance level that is of the order of the Bohr energy of a hole, this term is of the same order as all the rest of the terms of Eq. (48).

Now let us find the energy of the resonance level having maximum energy E. As above, we can seek the wave function in the form of (43). We get the following equation for the scalar function $f(k)$:

$$\left(\frac{\hbar^2k^2}{2m_p}+\frac{3\pi e^2k}{32\varepsilon_0}+E\right)f(k)-\frac{e^2}{8\pi^2\varepsilon_0}\int\frac{d^3qf(q)}{(\mathbf{k}-\mathbf{q})^2}\left[1+\frac{3(\mathbf{kq})^2}{k^2q^2}\right]=0. \tag{49}$$

As in the case of Eq. (44), the eigenvalues of Eq. (49) can be found by the variation principle. Upon choosing $f(k)$ in the form of (45), we obtain the following relationship between E and the variational parameter n:

$$E=\frac{m_pe^4}{\varepsilon_0^2\hbar^2}\left[\frac{1}{n}\left(6\ln 2-\frac{15}{4}\right)-\frac{1}{2n^2}\right]. \tag{50}$$

Minimization with respect to n gives the value $n\approx 2.44$ which corresponds to the energy $E_0=0.085m_pe^4/\varepsilon_0^2\hbar^2$. [68]

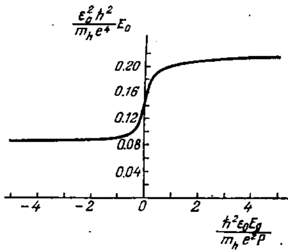

FIG. 5. The energy E_0 of the acceptor ground state as a function of E_g. The ground-state energy E_0 is in units of $m_h e^4/\varepsilon_0^2\hbar^4$, and E_g in units of $m_h e^2 P/\hbar^2\varepsilon_0$.

For the parameters of HgTe, where $\varepsilon_0 = 20$ and $m_p = 0.4m_0$, the energy of the resonance state is 2.3 meV, whereas a calculation performed within the framework of the local-potential theory gives $\varepsilon_0 = 6$ meV.[78] The experimental value of the energy E_0 found by cyclotron resonance measurements in HgTe[79] is 2.2 meV.

We have taken up in detail the solution of the problem of the levels in the field of a Coulombic impurity center. With decreasing $|E_g|$, we cannot restrict the treatment to terms proportional to k^2 in the electron and hole spectra. We must take account of nonparabolicity, which introduces an interaction with the nearest band having s-symmetry at the point Γ. The energy of the ground state as a function of $|E_g|$ increases with decreasing $|E_g|$, and when $E_g = 0$, it becomes equal to $m_h e^4/7\varepsilon_0^2\hbar^2$.[68] E_0 continues to grow further in the semiconductor region where $E_g > 0$, and it approaches the value $2m_h e^4/9\varepsilon_0^2\hbar^2$ at large E_g (Fig. 5).

The problem of the position of the impurity level in the case where the potential is strongly localized within a single cell has been treated in[80,81]. The coupling between the bands was assumed in[81] to be so strong that the interband matrix elements of the potential were equal to the intraband elements. As in the case of a Coulomb center, the acceptor levels are of resonance type, while the donor levels are highly diffuse. The trend of movement of the resonance levels is qualitatively the same as for a Coulomb center. The authors of[80] think that there are two types of centers in the solid solution $Cd_xHg_{1-x}Te$: those with a Coulomb potential are Cu ions, and those with a highly localized potential are Hg vacancies. The values of the matrix elements of the potential of the Hg vacancies have been determined in[81] from experimental data, whereupon they constructed the curve of the dependence of E_0 on x (in the appropriate units, this is the dependence on E_g). It agreed satisfactorily with the results of[80].

7. STATISTICS OF CARRIERS IN GAPLESS SEMICONDUCTORS

The existence of sharp acceptor levels on the background of the continuous spectrum of the conduction band gives rise to a number of interesting features in the thermodynamics and kinetics of the charge carriers in a GS. When the masses of electrons and holes differ greatly, donors and acceptors play differing roles in such a semiconductor.

Since there are no bound states of electrons at donors, the concentration n of free electrons equals the concentration N_D of donors, even at $T=0$. On the other hand, the acceptor levels are sharp, and hence the temperature must be of the same order as the energy of the acceptor ground state E_0 for an appreciable ionization of the acceptors to occur. (We are treating only the case of a low enough acceptor concentration $N_A a_p^3 \ll 1$, so that the wave functions of adjacent acceptors overlap but little. For HgTe this condition implies that $N_A \ll 10^{20}$ cm^{-3}). Since there is no forbidden band, free carriers arise at arbitrarily low temperatures, even in the absence of donors. Thus, when $N_D = 0$, intrinsic conduction arises in the GS at lower temperatures than does the impurity conduction involving ionization of acceptors.

If both donors and acceptors are present, then the GS will be n-type ($n > p$) at low temperature, regardless of the relationship between the concentrations N_A and N_D. Compensation does not occur at low temperatures (it is energetically unfavorable) if the Fermi level lies below the acceptor level. A sort of compensation sets in with increasing temperature, and at a high enough temperature, the semiconductor becomes p-type if $N_A > N_D$. In the course of this compensation, the concentration n of free electrons can decline with increasing temperature in a certain temperature range.

Let us find the concentration n of electrons as a function of the temperature. We shall consider the acceptor levels to be sharp. Upon assuming that all the donors are ionized, we can find the Fermi energy E_F from the equation of neutrality

$$N_D + p = n + N_A\left[1 + 4\exp\left(\frac{E_0 - E_F}{T}\right)\right]^{-1}. \tag{51}$$

The coefficient 4 accounts for the quadruple degeneracy of the ground state.

Let us treat the most interesting case in which $E_F > T$, so that the electrons are degenerate, while the holes are not degenerate. Then we get from (49):

$$\frac{3\sqrt{\pi}}{4}\left(\frac{m_p}{m_n}\right)^{3/2} e^{-x} - x^{3/2} = \left(\frac{E_0}{T}\right)^{3/2}\left\{n_A\left[1 + 4\exp\left(\frac{E_0}{T} - x\right)\right]^{-1} - n_D\right\}, \tag{52}$$

$$p = \frac{3\sqrt{\pi}}{4}\left(\frac{m_p T}{m_n E_0}\right)^{3/2} e^{-x} N_0, \quad n = \left(\frac{xT}{E_0}\right)^{3/2} N_0, \tag{53}$$

where

$$N_0 = \frac{(2m_n E_0)^{3/2}}{3\pi^2\hbar^3}, \quad x = \frac{E_F}{T}, \quad n_A = \frac{N_A}{N_0}, \quad n_D = \frac{N_D}{N_0}.$$

We have chosen the characteristic concentration N_0 such that the Fermi level coincides with the energy of the acceptor level when $n = N_0$. We can determine from (52) and (53) the concentrations of electrons and holes as functions of the temperature. We shall give as an example the curves calculated on a computer for $n(T)$ for the case in which $(3\sqrt{\pi}/4)(m_p/m_n)^{3/2} = 100$. This corresponds to the mass ratio $m_p/m_n = 18$, which is close to that in HgTe. Figures 6 and 7 show the results of calculations for different values of the dimensionless concentration n_A and n_D. The shape of the function $n(T)$ differs qualitatively, depending on whether $n_D < 1$ or $n_D > 1$.

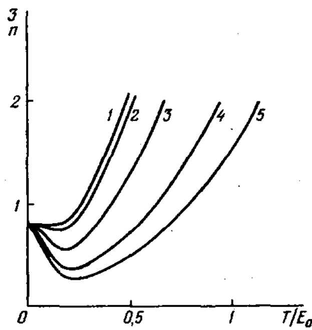

FIG. 6. Theoretical curves of the relationship of the concentration n of free electrons to the temperature T at a donor concentration $n_D = 0.8$ at different acceptor concentrations: 1—$n_A = 0$; 2—$n_A = 1$; 3—$n_A = 10$; 4—$n_A = 50$; 5—$n_A = 100$. All concentrations are expressed in units of N_0.

Figure 6 corresponds to the case $n_D = 0.8$, so that at $T=0$, the Fermi level lies somewhat below the ground-state acceptor energy E_0. In the absence of acceptors (curve 1) at low temperatures, the electron concentration remains constant, and then it begins to increase monotonically with increasing temperature, owing to formation of electron-hole pairs. A minimum appears on the $n(T)$ curve with increasing acceptor concentration.

The decline in the concentration of free electrons with increasing temperature in the low-temperature region is due to capture of electrons by acceptors. With further temperature increase, the role of thermal generation of pairs increases as the acceptors become occupied, and this increases the concentration of electrons.

Figure 7 corresponds to the case in which the Fermi level is above the acceptor level at $T=0$ ($n_D = 1.6$). In contrast to the previous case, an increase in N_A leads to a partial compensation even at $T=0$: as the acceptor concentration increases, the electron concentration n declines to the value $n = N_0$ when $N_A = N_D - N_0$. Further increase in N_A does not diminish the concentration of free electrons, since their capture by acceptors becomes energetically unfavorable. If $N_A > N_D - N_0$, then two competing processes occur at $T \neq 0$: thermal stoking of electrons onto the unoccupied acceptor levels, and thermal generation of electron-hole pairs. The former process dominates at low temperatures, and the latter at higher temperatures. Consequently the function $n(T)$ must have a minimum, just as in the case where $n_D < 1$.

Evidently, the temperature-dependence curve of the conductivity must also show a minimum, which will be more sharply marked because impurity scattering plays a greater role upon ionization of the acceptors.

We note that such an unusual temperature-dependence of the electron concentration and the conductivity allows us to expect that the volt-ampere characteristic of a GS will have a falling region. Actually, the heating of the carriers by the field can lead at appropriate donor and acceptor concentrations to a decline in the concentration of electrons, and a concomitant decline in the conductivity.

We must bear in mind that the electrons that have been captured by acceptors can participate in conduction, since the acceptor level is quasidiscrete and the lifetime of an electron in such a level is finite. Evidently the mobility of these electrons is small.

There is currently an entire set of experimental studies that confirm the existence of quasidiscrete acceptor levels.

A non-monotonic type of temperature-dependence of the conductivity has been observed in HgTe[82,83] and in the solid solutions $Cd_xHg_{1-x}Te$.[84,85] We can assume that this non-monotonic character involves the acceptor levels. They studied in[86] the pressure-dependence of the concentration of free carriers at $T = 4.2$ °K in p-type specimens of $Cd_xHg_{1-x}Te$ having $x = 0.14$–0.15. It turned out that the position of the Fermi level does not change with the pressure. This behavior of the Fermi level compelled the authors of[86] to conclude that an acceptor level exists on the background of the continuous spectrum of the conduction band. A detailed study of galvanomagnetic phenomena in $Cd_xHg_{1-x}Te$ p-type solid solutions having $0.1 < x < 0.15$ under omnilateral pressure up to 15 kiloatm has been performed in[87]. The acceptor concentration was large, and the quasidiscrete states became diffuse in the impurity band, which overlaps the conduction band at an energy of about 3 meV, according to the data of[87]. In[94] they studied the shift in the Fermi level with increasing temperature in p-type $Cd_xHg_{1-x}Te$ having $x = 0.1$.

A relationship between the frequencies of inter- and intraband transitions in HgTe and the magnetic field intensity was given in[52,88]. The measurements were performed in a Faraday-Voigt geometry in a range of wavelengths from 2 mm to 100 μm at 4.2 °K. They found extra absorption lines in addition to the frequencies that the authors identified with different transitions among the levels, as described by the Luttinger Hamiltonian. In the limit of weak magnetic field, the energy of the corresponding levels proved to differ from zero. The authors ascribed these lines to quasidiscrete acceptor levels. Analogous transitions have been observed in[42]. Manifestations of an impurity acceptor band in kinetic effects in $Cd_xHg_{1-x}Te$ solid solutions having $x = 1$ with $N_A > 5 \times 10^{17}$ cm^{-3} have found in[89]. A minimum on the $n(T)$ curve was firmly established in[89,90] in specimens of $Cd_xHg_{1-x}Te$ having $x = 0.1$ with a hole concentration $p = 8 \times 10^{17}$ cm^{-3}.

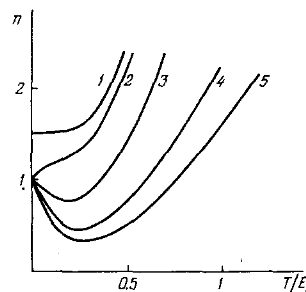

FIG. 7. The same as in Fig. 6, for $n_D = 1.6$.

8. CONCLUSION

Considerable progress has currently been attained in studying the energy spectrum of GS's. It has become clear that the latter is not a copy of the spectrum of a "normal" semiconductor having degenerate bands in which the sign of the curvature of one of them has been reversed, as was originally assumed in the model of Groves and Paul.[6] It has been shown that many-particle effects play a considerable role near the point of degeneracy of the bands, while the valence band is substantially restructured owing to the nonlocality of the lattice potential of the GS. Quasilocalized levels of impurity states occur on the background of the continuous spectrum of the conduction band.

Yet a whole set of unelucidated problems remains. On the theoretical level, this primarily concerns the regions of the spectrum near the point of degeneracy of the bands, where the many-electron effects are substantial. The problem remains open of the stability of the spectrum in this region. Instability can give rise to an exciton dielectric.[91–93] Much work looms ahead in studying the effect of a magnetic field on the electron and hole spectra of GS's with account taken of the nonlocality of the potential. As for the experimental studies, we need purer materials than we yet have. Studies of galvanomagnetic and resonance phenomena at low temperatures will lead to greater understanding of the features of GS's.

As yet the impurities have not been fully identified, even in the most studied substance, HgTe. In this sense, a large volume of studies still remains on the effect of controlled impurities on electronic phenomena in GS's.

Refinement is needed in the band parameters in HgTe (and in other GS's), which as yet differ greatly in the measurements made by different authors.

[1] A. A. Abrikosov and S. D. Beneslavskiĭ, Zh. Eksp. Teor. Fiz. 59, 1280 (1970) [Sov. Phys. JETP 32, 699 (1971)].
[2] N. B. Brandt and E. A. Svistova, J. Low Temp. Phys. 2, 1 (1970).
[3] J. C. Dimmock, J. Melngailis, and A. J. Strauss, Phys. Rev. Lett. 16, 1193 (1966).
[4] A. J. Strauss, Phys. Rev. 157, 608 (1967).
[5] C. R. Pidgeon and S. H. Groves, *ibid.* 186, 824 (1969).
[6] S. H. Groves and W. Paul, Phys. Lett. 11, 194 (1963).
[7] G. L. Bir and G. E. Pikus, Simmetriya i deformatsionnye éffekty v poluprovodnikakh (Symmetry and Strain-Induced Effects in Semiconductors), Nauka, M., 1972 (Engl. Transl., Wiley, New York, 1974).
[8] F. Herman, R. Kortum, C. Kuglin, and R. Short, in: Intern. Conference on Physics of Semiconductors, Kyoto, 1966, p. 7.
[9] F. Herman, R. Kortum, C. Kuglin, J. Van Dyke, and S. Skillman, in: Methods in Computational Physics, V. 8, N. Y., 1968.
[10] F. H. Pollak, M. Cardona, C. W. Higginbotham, F. Herman, and J. P. Van Dyke, Phys. Rev. B2, 352 (1970).
[11] S. H. Groves, C. R. Pidgeon, A. W. Ewald, and R. J. Wagner, J. Phys. and Chem. Sol. 31, 2031 (1970).
[12] M. Cohen, and T. Bergstresser, Phys. Rev. 141, 789 (1966).
[13] S. Bloom and T. Bergstresser, Sol. State Comm. 6, 465 (1968).
[14] S. Bloom and T. Bergstresser, Phys. Stat. Sol. 42, 191 (1970).
[15] S. Katsuki and M. Kunimure, J. Phys. Soc. Japan 31, 337, 415 (1970).
[16] H. Overhof, Phys. Stat. Sol. b43, 221; b45, 315 (1971).
[17] F. Herman, R. Kortum, C. Kuglin, and J. Shay, in: Proc. of Intern. Conference on II–VI Semiconductor Compounds, Providence, USA, 1967, p. 503.
[18] G. Bush and T. Kern, Sol. State Phys. 11, 1 (1960).
[19] N. P. Mokrovskiĭ and A. R. Regel', Zh. Tekh. Fiz. 22, 1281 (1952).
[20] I. M. Tsidil'kovskiĭ, Zh. Tekh. Fiz. 27, 1744 (1957) [Sov. Phys. Tech. Phys. 2, 1622 (1958)].
[21] T. C. Herman, M. J. Logan, and H. L. Goering, J. Phys. and Chem. Sol. 7, 228 (1958).
[22] R. O. Carlson, Phys. Rev. 111, 476 (1958).
[23] T. Harman and A. Strauss, J. Appl. Phys. 32, 2265 (1961).
[24] T. Harman, A. Strauss, J. Mavroides, D. Dickey, and M. Dresselhaus, in: Proc. of Intern. Conference on Physics of Semiconductors, Exeter, England, 1962, p. 703.
[25] V. I. Ivanov-Omskiĭ, B. T. Kolomiets, A. A. Mal'kova, V. K. Ogorodnikov, and K. P. Smekalova, Izv. Akad. Nauk SSSR, ser. fiz. 28, 1057 (1964); Phys. Stat. Sol. 8, 613 (1965).
[26] Z. Dzuiba and T. Zakrzewski, *ibid.* 7, 1019 (1964).
[27] V. I. Ivanov-Omskiĭ, B. T. Kolomiets, V. K. Ogorodnikov, and K. P. Smekalova, Fiz. Tekh. Poluprovodn. 4, 264 (1970) [Sov. Phys. Semicond. 4, 214 (1970)].
[28] C. Vérié, Phys. Stat. Sol. 17, 889 (1966).
[29] S. Porowski and T. Zakrzewski, *ibid.* 11, K39 (1965).
[30] R. Piotrzkowski, S. Porowski, Z. Dzuiba, J. Ginter, W. Giriat, and L. Sosnowski, *ibid.* 8, K135 (1965).
[31] B. J. Roman and A. W. Ewald, Phys. Rev. B5, 3914 (1972).
[32] S. H. Groves, R. N. Brown, and C. R. Pidgeon, *ibid.* 161, 779 (1967).
[33] B. L. Gel'mont, Fiz. Tverd. Tela 11, (Leningrad) 1096 (1969) [Sov. Phys. Solid State 11, 894 (1969)].
[34] V. I. Ivanov-Omskiĭ, B. T. Kolomiets, V. M. Mel'nik, and V. K. Ogorodnikov, *ibid.*, 11, 2563 (1969) [Sov. Phys. Solid State 11, 2067 (1970)].
[35] B. L. Gel'mont, V. I. Ivanov-Omskiĭ, B. T. Kolomiets, and V. M. Mel'nik, Fiz. Tekh. Poluprovodn. 4, 287 (1970) [Sov. Phys. Semicond. 4, 234 (1970)].
[36] I. M. Tsidil'kovskiĭ, Élektrony i dyrki v poluprovodnikakh (Electrons and Holes in Semiconductors), Nauka, M., 1972.
[37] J. M. Luttinger, Phys. Rev. 102, 1030 (1956).
[38] Opticheskie svoĭstva poluprovodnikov (Optical Properties of Semiconductors), Ed. N. G. Basov (Vol. 75 of Proc. P. N. Lebedev Phys. Inst.), (Engl. Transl., Consultants Bureau, 1975).
[39] C. R. Pidgeon and R. N. Brown, Phys. Rev. 146, 575 (1966).
[40] S. H. Groves, C. R. Pidgeon, A. W. Ewald, and R. J. Wagner, J. Phys. and Chem. Sol. 31, 2031 (1970).
[41] J. Guldner, C. Rigaux, M. Grynberg, and A. Mycielski, Phys. Rev. B8, 3875 (1973).
[42] S. Uchida, R. Yoshizaki, and S. Tanaka, in: Intern. Conference on Physics of Semimetals and Narrow-gap Semiconductors, Nice, France, 1973.
[43] S. H. Groves, C. R. Pidgeon, and R. N. Brown, Phys. Rev. 161, 779 (1967).
[44] C. R. Pidgeon and S. H. Groves, see Ref. 17, p. 1080.
[45] R. J. Wagner and L. H. Ngai, in: Proc. of 12th Intern. Conference on Physics of Semiconductors, Stuttgart, 1974, p. 1147.
[46] V. I. Ivanov-Omskiĭ, N. N. Konstantinova, R. V. Parfen'ev, V. V. Sologub, and I. G. Tagiev, Fiz. Tekh. Poluprovodn. 7, 715 (1973) [Sov. Phys. Semicond. 7, 496 (1973).
[47] V. Giriat, É. A. Neĭfel'd, and I. M. Tsidil'kovskiĭ, Fiz. Tekh. Poluprovodn. 9, 188 (1975) [Sov. Phys. Semicond. 9, 129 (1975)].
[48] J. Kowalskii and W. Zawadzki, Sol. State Comm. 13, 1433

(1973).
[49]D. Sherrington and W. Kohn, Phys. Rev. Lett. 21, 153 (1968).
[50]See Ref. 38, p. 182.
[51]J. G. Broerman, Phys. Rev. B5, 397 (1972).
[52]J. Tuchendler, M. Grynberg, Y. Gouder, H. Thomé, and R. Le Toullec, *ibid.* B8, 3884 (1973).
[53]V. I. Ivanov-Omskiĭ, B. T. Kolomiets, A. A. Mal'kova, Yu. F. Markov, and A. Sh. Mekhtiev, Fiz. Tekh. Poluprovodn. 4, 417 (1970) [Sov. Phys. Semicond. 4, 352 (1970)].
[54]M. Grynberg and R. Le Toullec, in Proc. of 11th Intern. Conference on Physics of Semiconductors, V 2, Warszawa, 1972, p. 931.
[55]R. J. Wagner and A. W. Ewald, J. Phys. and Chem. Sol. 32, 697 (1971).
[56]L. Liu and D. Brust, Phys. Rev. Lett. 20, 651 (1968).
[57]L. Liu and E. Tosatti, *ibid.* 23, 772 (1969).
[58]L. Liu and M. Tan, Phys. Rev. B9, 632 (1974).
[59]See Ref. 27.
[60]E. D. Hinkley and A. W. Evald, Phys. Rev. A4, 1260 (1964).
[61]J. G. Broerman, *ibid.* B2, 1818 (1970).
[62]B. L. Gel'mont, V. I. Ivanov-Omskiĭ, B. T. Kolomiets, V. K. Ogorodnikov, and K. P. Smekalova, Fiz. Tekh. Poluprovodn. 5, 266 (1972) [Sov. Phys. Semicond. 5, 228 (1972)].
[63]A. A. Abrikosov and S. D. Beneslavskii, J. Low Temp. Phys. 5, 141 (1971).
[64]A. A. Abrikosov, Zh. Eksp. Teor. Fiz. 66, 1443 (1974) [Sov. Phys. JETP 39, 709 (1974)].
[65]V. B. Berestetskiĭ, E. M. Lifshitz and L. P. Pitaevskiĭ, Relyativistakaya kvantovaya teoriya (Relativistic Quantum Theory), Part 1, Nauka, M., 1968 (Engl. Transl., Pergamon, Oxford, New York, 1971).
[66]B. J. Halperin and T. M. Rice, Rev. Mod. Phys. 40, 775 (1968).
[67]A. A. Abrikosov, J. Low. Temp. Phys. 18, 185 (1975).
[68]B. L. Gel'mont, Fiz. Tekh. Poluprovodn. 9, 1912 (1975) [Sov. Phys. Semicond. 9, 1257 (1975)].
[69]E. O. Kane, J. Phys. and Chem. Sol. 1, 249 (1957).
[70]B. L. Gel'mont, V. I. Ivanov-Omskiĭ, and V. K. Ogorodnikov, Phys. Stat. Sol. 69, 717 (1975).
[71]W. Kohn and D. Shechter, Phys. Rev. 99, 1903 (1955); W. Kohn, Sol. State Phys. 5, 258 (1957).
[72]D. Shechter, J. Phys. and Chem. Sol. 28, 237 (1962).
[73]K. S. Mendelsohn and H. M. James, *ibid.* 25, 729 (1964).
[74]V. I. Sheka and D. I. Sheka, Zh. Eksp. Teor. Fiz. 51, 1445 (1966) [Sov. Phys. JETP 24, 975 (1967)].
[75]K. S. Mendelsohn and J. A. Shultz, Phys. Stat. Sol. 31, 59 (1969).
[76]B. L. Gel'mont and M. I. D'yakonov, Fiz. Tekh. Poluprovodn. 5, 2191 (1971) [Sov. Phys. Semicond. 5, 1905 (1972)].
[77]B. I. Shklovskiĭ, Fiz. Tekh. Poluprovodn. 6, 1197 (1972) [Sov. Phys. Semicond. 6, 1053 (1973)].
[78]B. L. Gel'mont and M. I. D'yakonov, Zh. Eksp. Teor. Fiz. 62, 713 (1972) [Sov. Phys. JETP 35, 377 (1972)]; in Materialy 4-ĭ zimneĭ shkoly po fizike poluprovodnikov (Materials of the 4th Winter School on Semiconductor Physics), Leningrad, 1972, p. 124.
[79]G. Bastard, Y. Guldner, A. Mycielski, P. Nozieres, and C. Rigaux, see Ref. 45, p. 1162.
[80]A. Mauger, S. Otmezguine, C. Vérié, and J. Friedel, *ibid.*, p. 1166.
[81]A. Mauger and J. Friedel, Phys. Rev. 12, 2412 (1975).
[82]V. I. Ivanov-Omskiĭ, B. T. Kolomiets, V. K. Ogorodnikov, and K. P. Smekalova, Fiz. Tekh. Poluprovodn. 1, 289 (1967) [Sov. Phys. Semicond. 1, 232 (1967)]; see also Ref. 27.
[83]B. L. Gel'mont, M. I. Dyakonov, V. I. Ivanov-Omskiĭ, B. T. Kolomiets, V. K. Ogorodnikov, and K. P. Smekalova, see Ref. 54, Vol. 2, p. 938.
[84]C. Finck, S. Otmezguine, G. Weill, and C. Vérié, *ibid.*, p. 944.
[85]S. Otmezguine, F. Raymond, G. Weill, and G. Vérié, in: Proc. of 10th Intern. Conference on Physics of Semiconductors, Cambridge, 1970, p. 536.
[86]C. T. Elliott, J. Melngailis, T. C. Harman, J. A. Kafalas, and W. C. Kerman, Phys. Rev. B5, 2985 (1972).
[87]N. B. Brandt, O. N. Belousova, L. A. Bovina, V. I. Stafeev, and Ya. G. Ponomarev, Zh. Eksp. Teor. Fiz. 66, 330 (1974) [Sov. Phys. JETP 39, 157 (1974)].
[88]J. Tuchendler, M. Grynberg, Y. Gouder, and H. Thomé, see Ref. 54, Vol. 1, p. 346.
[89]I. M. Tsidilkovskii, W. Giriat, G. I. Kharus, and E. A. Neifeld, Phys. Stat. Sol. 64, 717 (1974).
[90]V. Giriat, E. A. Neĭfel'd, I. M. Tsidil'kovskiĭ, and N. A. Gorodilov, Fiz. Tekh. Poluprovodn. 10, 63 (1976) [Sov. Phys. Semicond. 10, 36 (1976)]; in IV Vsesoyuznyĭ simpozium po poluprovodnikam s uzkoĭ zapreshchennoĭ zonoĭ i polumetallam (4th All-Union Symposium on Semiconductors with a Narrow Forbidden Band and Semimetals), Part 1, 1975, p. 16.
[91]L. V. Keldysh and Yu. V. Kopaev, Fiz. Tverd. Tela (Leningrad) 6, 2791 (1964) [Sov. Phys. Solid State 6, 2219 (1965)].
[92]A. N. Kozlov and L. A. Maksimov, Zh. Eksp. Teor. Fiz. 48, 1184 (1965) [Sov. Phys. JETP 21, 790 (1965)].
[93]D. Sherrington and W. Kohn, Rev. Mod. Phys. 40, 767 (1968).
[94]N. B. Brandt, G. N. Belousova, and Ya. G. Ponomarev, Fiz. Tverd. Tela (Leningrad) 17, 3585 (1975) [Sov. Phys. Solid State 17, 2334 (1975)].
[95]Kvantovaya teoriya polya i fizika fazovykh perekhodov (Quantum Field Theory and the Physics of Phase Transitions), Mir, M., 1975.

Translated by M. V. King