Electron-hole liquid and the metal-dielectric phase transition in layered systems

E. A. Andryushin, L. V. Keldysh, and A. P. Silin

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

(Submitted April 13, 1977)
Zh. Eksp. Teor. Fiz. **73**, 1163–1173 (September 1977)

A phase transition of the gas-liquid type in a system of electrons and holes in a quasi-two-dimensional (layered) semiconductor is considered. The phase diagram, critical temperature and density for the transition are obtained. It is shown that near the critical point the transition is of a purely plasma nature. In other words, bound states of the exciton type are absent in both (gas and liquid) phases.

PACS numbers: 71.35.+z, 71.30.+h

It was shown in a previous work[1] that the role of correlation effects in electron-hole plasma (EHP) turns out to be anomalously large in semiconductors and semimetals possessing extremely strong anisotropies of the electron spectrum, in particular, layered systems. A significant decrease in the energy associated with interelectron correlations creates a tendency to "self-compression" of such a plasma, i.e., the formation of an electron-hole liquid (EHL) with particular density $n \gg a_x^{-3}$ and binding energy per electron-hole pair $|E_{\min}| \gg E_x$, where E_x and a_x are the binding energy and the effective radius of the hydrogen-like exciton. Formation of an EHL at low temperatures takes place through a first-order phase transition, in which the concentration of the carriers, free and bound in excitons, reaches some critical value $n_{\text{gas}}(T)$ that depends on the temperature—the density of the saturated vapor. This situation can be realized both at thermodynamic equilibrium for a semiconductor with a sufficiently narrow forbidden band E_g upon increase in the temperature or decrease in E_g, and under essentially nonequilibrium conditions, in which the critical concentration is achieved by introduction of selected carriers in intense excitation of the semiconductor.

In the first case, the formation of the EHL means a discontinuous change in the width of the forbidden band E_g to some negative value (overlapping of the bands), i.e., a transition of the initial semiconductor to a semimetal state. The second case corresponds to the so-called condensation of nonequilibrium carriers (or excitons) into EHL drops. With the same accuracy with which the nonequilibrium carriers can be assumed to be thermalized, the thermodynamics of both these transitions is identical and we shall consider them here using as an example one of the types of systems previously described[1]—layered systems. Just as in Ref. 1, we shall mean by a layered semiconductor or semimetal an idealized model in which the motion of the electrons and holes is two-dimensional, i.e., it takes place only in the plane of the layers without transitions between them. For simplicity, the dispersion laws of the electrons $\varepsilon_e(\mathbf{p})$ and holes $\varepsilon_h(\mathbf{p})$ will be assumed to be quadratic and isotropic in the plane of the layer: $\varepsilon_{e,h}(\mathbf{p}) = p^2/2m_{e,h}$. We shall also assume the permittivity tensor of the crystal to be isotropic in the plane of the layers. This tensor is characterized by two principal values $\varepsilon_\perp$ and $\varepsilon_\parallel$ for the directions along and perpendicular to the layers, respectively.

We emphasize that we are speaking of a permittivity without any contribution of free carriers. We shall use a system of units defined by the relations

$$e^2/(\varepsilon_\parallel\varepsilon_\perp)^{1/2} = \hbar = 2m = 1,$$

where e and $\hbar$ are the charge on the electron and Planck's constant, and m is the reduced mass, $m^{-1} = m_e^{-1} + m_h^{-1}$. In this work, along with the parameter of the ratio of the effective masses of the electron and the hole, $\sigma = m_e/m_h$ that is usually employed, we will find it convenient to use another parameter $s = (1-\sigma)/(1+\sigma)$.

The system of electrons and holes in the considered model is described by the Hamiltonian

$$H - \mu_e n_e - \mu_h n_h = \sum_{\mathbf{p}s_1l}\left[\left(\frac{1+s}{2}p^2 - \mu_e\right)a^+_{\mathbf{p}s_1l}a_{\mathbf{p}s_1l} + \left(\frac{1-s}{2}p^2 - \mu_h\right)b^+_{\mathbf{p}s_1l}b_{\mathbf{p}s_1l}\right]$$
$$+\frac{1}{2}\sum_{\substack{\mathbf{p}_1\mathbf{p}_2\mathbf{q}\\ s_1s_2l_1l_2}} V_{q,l_1-l_2}[a^+_{\mathbf{p}_1s_1l_1}a^+_{\mathbf{p}_2s_2l_2}a_{\mathbf{p}_2+\mathbf{q}s_2l_2}a_{\mathbf{p}_1-\mathbf{q}s_1l_1}$$
$$+ b^+_{\mathbf{p}_1s_1l_1}b^+_{\mathbf{p}_2s_2l_2}b_{\mathbf{p}_2+\mathbf{q}s_2l_2}b_{\mathbf{p}_1-\mathbf{q}s_1l_1} - 2a^+_{\mathbf{p}_1s_1l_1}b^+_{\mathbf{p}_2s_2l_2}b_{\mathbf{p}_2+\mathbf{q}s_2l_2}a_{\mathbf{p}_1-\mathbf{q}s_1l_1}], \qquad (1)$$

where $a^+_{\mathbf{p}s_{1,2}l}$ and $b^+_{\mathbf{p}s_{1,2}l}$ are the electron and hole creation operators in the layer l ($l = 0, \pm1, \pm2, \ldots$), and spin projections $s_{1,2}$ and momentum $\mathbf{p}$ in the plane of the layer; $\mathbf{p}_{1,2}$ and $\mathbf{q}$ are two-dimensional vectors; the normalized area of the layer is set equal to unity:

$$V_{q,l} = 2\pi q^{-1}\exp(-q|l|c^*); \qquad (2)$$

$c^* = c(\varepsilon_\perp/\varepsilon_\parallel)^{1/2}$; c is the distance between the neighboring planes; μ_e and μ_h are the chemical potentials of the electrons and holes, calculated from the bottom of the respective bands and connected by the condition of electric neutrality

$$\sum_{\mathbf{p}s_1}\langle a^+_{\mathbf{p}s_1l}a_{\mathbf{p}s_1l}\rangle = \sum_{\mathbf{p}s_2}\langle b^+_{\mathbf{p}s_2l}b_{\mathbf{p}s_2l}\rangle = nc; \qquad (3)$$

n is the volume density of the number of electron-hole pairs; nc is the surface density in a single layer.

The introduction of the chemical potentials μ_e and μ_h is necessary for the description of nonequilibrium systems, when the electrons and holes assumed to be dif-

0038-5646/77/4603-0616$02.40

ferent types of particles and their concentration n is specified arbitrarily. In a thermodynamically equilibrium situation, the chemical potential of the electrons in both bands should be the same, which adds an additional condition for μ_e and μ_h:

$$\mu_e + \mu_h = -E_g^{(0)}, \tag{4}$$

where $E_g^{(0)}$ is the width of the forbidden band, i.e., the distance between the initial values of the energies of the electrons and holes in (1). It should be noted that $E_g^{(0)}$ is the unrenormalized width of the forbidden band, a width that enters into the bare dispersion law. The account given below of the interaction of electrons and holes leads to a significant difference between $E_g^{(0)}$ and the real width E_g of the forbidden band right up to the transition from the semiconductor unrenormalized continuous spectrum ($E_g^{(0)}>0$) to the semimetal ($E_g<0$) after renormalization.

The investigation of the thermodynamics of the considered model is most conveniently carried out by starting from the dependence of the particle number density on the temperature and the chemical potential

$$n_{e,h}(\mu_{e,h}, T) = \frac{2T}{c}\sum_k \int \frac{d^2p}{(2\pi)^2} G_{e,h}(\mathbf{p}, \varepsilon_k) e^{i\varepsilon_\kappa\tau}$$
$$= \frac{2T}{c}\sum_k \int \frac{d^2p}{(2\pi)^2} \frac{e^{i\varepsilon_\kappa\tau}}{i\varepsilon_k + \mu_{e,h} - \varepsilon_{e,h}(\mathbf{p}) - \Sigma_{e,h}(\mathbf{p}, \varepsilon_k)}, \tag{5}$$

where $G_{e,h}$ is the Green's function of the electrons and holes in the Matsubara technique[2] (see also Ref. 3), $\Sigma_{e,h}$ are the self-energy parts, $\varepsilon_k = \pi T(2k+1)$, and k is an integer and $\tau \to +0$.

Further consideration is based on the assumption that the distances between the layers c are small in comparison with the radius of the two-dimensional exciton, i.e., with account of the scales employed,

$$c \ll 1. \tag{6}$$

We shall show below (formulas (43)–(49)) that upon satisfaction of the condition (6), in the region of concentrations and temperatures defined by the inequalities

$$c^{-1} \ll n \ll c^{-2}, \tag{7}$$
$$T \ll n^{1/2}, \tag{8}$$

the self-energy parts $\Sigma_{e,h}$ do not depend in first approximation on either the temperature or the arguments $\mathbf{p}$ and ε_k, and have the form

$$\Sigma_{e,h} = -a(\pm s)(c/c^*)^{1/4} n^{1/4}, \tag{9}$$

where the plus sign corresponds to electrons, and the minus to holes,

$$a(s) = \frac{2^{13/2}\pi^{3/4}}{[\Gamma(1/4)]^2} \int_0^\infty dx \frac{1+s}{x^2+(1+s)^2}\left[\frac{1+s}{x^2+(1+s)^2} + \frac{1-s}{x^2+(1-s)^2}\right]^{1/4}. \tag{10}$$

Thus, in this region of concentrations and temperatures, the interaction in the electron-hole plasma leads to a narrowing of the forbidden band without any appreciable change in any other of the parameters of the electron spectrum. We shall also show that the entire region of existence of the electron-hole liquid, including the critical point, falls in the interval of concentrations and temperatures satisfying the conditions (7) and (8). In order to establish this, we consider first the thermodynamics of the electron-hole plasma, starting out from the formulas (9) and (10), and then give the basis of the formulas themselves. Carrying out the summation and integration in (5) with account of the independence of Σ of $\mathbf{p}$ and ε_k, we obtain

$$n_{e,h} = n = \frac{T}{\pi c(1 \pm s)} \ln\left(1 + \exp\frac{\mu_{e,h} - \Sigma_{e,h}}{T}\right). \tag{11}$$

Solving the relation (11) for $\mu_{e,h}$ and combining the equations thus obtained, we find the dependence of the chemical potential of a single electron-hole pair $\mu = \mu_e + \mu_h$ on n and T:

$$\mu(n, T) = \Sigma_e + \Sigma_h + T\ln\left\{\left(\exp\frac{\pi nc(1+s)}{T} - 1\right)\left(\exp\frac{\pi nc(1-s)}{T} - 1\right)\right\}$$
$$= -A\left(\frac{c}{c^*}\right)^{1/4} n^{1/4} + T\ln\left[2e^{\pi nc/T}\left(\operatorname{ch}\frac{\pi nc}{T} - \operatorname{ch}\frac{\pi ncs}{T}\right)\right], \tag{12}$$

where the coefficients $A = a(s) + a(-s)$ is connected in the following fashion with the function $f(\sigma)$ introduced previously[1]:

$$A = 2^{23/4}\pi^{9/4} f(\sigma)/[\Gamma(1/4)]^4. \tag{13}$$

The relation (12) is essentially the equation of state of the electron-hole plasma, written in the variables μ and n instead of the more customary pressure and volume. Several typical curves, described by these relations for the simplest case $\sigma = 1$ ($s = 0$), are given in Fig. 1. It is seen that they have the typical van der Waals character, i.e., at temperatures $T < T_c$ there are two stable branches $n_{\text{gas}}(\mu, T)$ and $n_{\text{liq}}(\mu, T)$ and one thermodynamically unstable branch $(\partial\mu/\partial n)_T < 0$ of the solutions. The lesser of the two stable solutions n_{gas} corresponds to the gaseous phase, the greater—n_{liq}—corresponds to the electron-hole liquid. The critical point (T_c, n_c) is determined by the two equations

$$(\partial\mu/\partial n)_{T_c, n_c} = (\partial^2\mu/\partial n^2)_{T_c, n_c} = 0, \tag{14}$$

which reduce to a single transcendental equation for the parameter $z_c = \pi c n_c / T_c$:

$$z_c = \frac{3(\operatorname{ch} z_c - \operatorname{ch}(sz_c))(\exp(z_c) - \operatorname{ch}(sz_c) - s\operatorname{sh}(sz_c))}{4[(1+s^2)(\operatorname{ch} z_c \operatorname{ch}(sz_c) - 1) - 2s\operatorname{sh} z_c \operatorname{sh}(sz_c)]}. \tag{15}$$

The critical parameters of the liquid are expressed in terms of z_c by the formulas

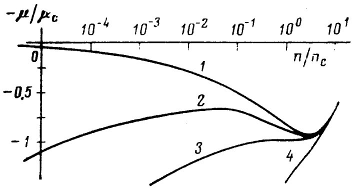

FIG. 1. Dependence of the chemical potential on the density in relative units at various temperatures: curves 1) $T=0$, 2) $T=0.5T_c$, 3) $T=T_c$, 4) $T=2T_c$. The semi-axis $n=0$, $\mu \leqslant 0$ also belongs to the isotherm 1 at $T=0$.

$$n_c c=\left[\frac{A}{4\pi}\frac{\operatorname{ch} z_c-\operatorname{ch}(z_c s)}{\exp(z_c)-\operatorname{ch}(z_c s)-s\operatorname{sh}(z_c s)}\right]^{4/3}\frac{1}{c^{*1/3}}, \tag{16}$$

$$T_c=\frac{\pi}{z_c}\left[\frac{A}{4\pi}\frac{\operatorname{ch} z_c-\operatorname{ch}(z_c s)}{\exp(z_c)-\operatorname{ch}(z_c s)-s\operatorname{sh}(z_c s)}\right]^{4/3}\frac{1}{c^{*1/3}}, \tag{17}$$

$$\mu_c=-T_c\left[\left(3+4\frac{\operatorname{sh} z_c-s\operatorname{sh}(sz_c)}{\operatorname{ch} z_c-\operatorname{ch}(sz_c)}\right)z_c-\ln 2-\ln(\operatorname{ch} z_c-\operatorname{ch}(sz_c))\right]. \tag{18}$$

By virtue of (15), z_c depends only on a single parameter—the ratio of the masses of the electron and hole σ (or s). The dependence of z_c on σ is very weak:

$$z_c(\sigma)=\begin{cases}0.5501, & \sigma=1\\ 0.4877, & \sigma=0.\end{cases} \tag{19}$$

With accuracy to one hundredth, the function $z_c(\sigma)$ can be obtained in explicit form by expansion of the right side of (15) in a series in z_c up to terms of third order:

$$z_c(\sigma)\approx\frac{3}{2(1+s^2)}\left[\left(1+\frac{8}{9}(1+s^2)\right)^{1/2}-1\right]. \tag{20}$$

With the same accuracy, the formulas (16)–(18) reduce to the form

$$n_c c\approx\left[\frac{3A}{32\pi}z_c\left(1+\frac{1+s^2}{12}z_c^2\right)\right]^{4/3}\frac{1}{c^{*1/3}}, \tag{21}$$

$$T_c=\frac{\pi}{z_c}\left[\frac{3A}{32\pi}z_c\left(1+\frac{1+s^2}{12}z_c^2\right)\right]^{4/3}\frac{1}{c^{*1/3}}, \tag{22}$$

$$\mu_c=-T_c\left[\frac{32}{3}\left(1-\frac{1+s^2}{12}z_c^2\right)-z_c-\ln z_c^2(1-s^2)\left(1+\frac{1+s^2}{12}z_c^2\right)\right]. \tag{23}$$

At temperatures less than critical, one of the two stable solutions $n_{\text{gas}}(\mu, T)$ and $n_{\text{liq}}(\mu, T)$ is generally metastable. Only for a single value of the chemical potential $\mu(T)$, determined for each temperature T, is the existence of both phases possible with balanced densities $n_{\text{gas}}(T)=n_{\text{gas}}(\mu(T), T)$ and $n_{\text{liq}}(T)=n_{\text{liq}}(\mu(T), T)$. The quantities $\mu(T)$, $n_{\text{gas}}(T)$ and $n_{\text{liq}}(T)$ are determined by the conditions that the chemical potentials and pressures be equal in both phases:

$$\mu(T)=\mu(n_{\text{gas}}(T), T)=\mu(n_{\text{liq}}(T), T), \tag{24}$$

$$\int\limits_{n_{\text{gas}}}^{n_{\text{liq}}} dn\,\mu(n, T)=\mu(T)(n_{\text{liq}}(T)-n_{\text{gas}}(T)), \tag{25}$$

with the accuracy shown above, near the critical point, $T_c-T\ll T_c$:

$$n_{\text{liq}}(T)-n_c=n_c-n_{\text{gas}}(T)\approx 2\cdot 6^{1/2}\left(1-\frac{1+s^2}{3}z_c^2\right)n_c\left(\frac{T_c-T}{T_c}\right)^{1/2}, \tag{26}$$

$$\mu(T)-\mu_c=\left\{2-\ln z_c^2(1-s^2)+\frac{1}{12}(1+s^2)z_c^2\right\}(T_c-T). \tag{27}$$

At low temperatures $T\ll T_c$:

$$\begin{aligned} n_{\text{gas}}c&=n_0 c\left\{1-\frac{25\pi^2}{2^{10}(1-s^2)}\left(\frac{15z_c}{4}\right)^{2/3}\left[1+\frac{2}{9}(1+s^2)z_c^2\right]\left(\frac{T}{T_c}\right)^2\right\}\\ &\approx 0.166f^{4/3}(\sigma)\left\{1+0.582\frac{z_c^{2/3}}{1-s^2}\left[1+\frac{2}{9}(1+s^2)z_c^2\right]\left(\frac{T}{T_c}\right)^2\right\}\frac{1}{c^{*1/3}}, \end{aligned} \tag{28}$$

$$\begin{aligned} \mu(T)&=\mu_0\left\{1+\frac{25\pi^2}{2^{12}(1-s^2)}\left(\frac{15z_c}{4}\right)^{2/3}\left[1+\frac{2}{9}(1+s^2)z_c^2\right]\left(\frac{T}{T_c}\right)^2\right\}\\ &\approx -1.57f^{4/3}(\sigma)\left\{1+0.146\frac{z_c^{2/3}}{1-s^2}\left[1+\frac{2}{9}(1+s^2)z_c^2\right]\left(\frac{T}{T_c}\right)^2\right\}\frac{1}{c^{*1/3}}, \end{aligned} \tag{29}$$

where

$$n_0 c=\left(\frac{A}{5\pi}\right)^{4/3}c^{*-1/3}=\frac{0.166f^{4/3}(\sigma)}{c^{*1/3}}, \tag{30}$$

$$\mu_0=E_{min}=-3\pi\left(\frac{A}{5\pi}\right)^{4/3}c^{*-1/3}=-\frac{1.57f^{4/3}(\sigma)}{c^{*1/3}} \tag{31}$$

are respectively the equilibrium density and minimum energy of the ground state of the EHL at $T=0$. In this limiting case, the expressions (28), (29) go over into the formula (10) of Ref. 1.[1)] It is seen from Eqs. (21)–(23) and (30), (31) that the region of existence of the EHL satisfies the inequalities (7) and (8).

It should be noted that the expansion used in (20) and subsequently assures an accuracy to within a few percent in all the formulas. This accuracy will obviously be quite sufficient when a comparison is made with the experimental data, the more so that, in the region of real values of the distance between the layers c, the expressions (21)–(23), (31) and (32) will be satisfied only qualitatively. The problems of the applicability of the present consideration are discussed at the end of the article.

Figure 2 shows a phase diagram in reduced units kn/n_c and T/T_c, constructed numerically according to (24), (25) for the case $\sigma=1$. The form of the phase diagram does not depend on the parameter c and depends weakly on the parameter σ. It follows from the calculation that for all values of the parameters, the following relation between the critical temperature T_c and the binding energy of the EHL $|E_{\min}|$ at $T=0$ is well satisfied:

$$T_c\approx 0.1|E_{min}|. \tag{32}$$

This relation turns out to be quite general. Thus, it holds approximately for the experimentally observed phase diagrams of the EHL in germanium and silicon, which are never layered materials. The experimental data are for germanium[4]: $T_c=6.5\pm0.1$ K, $|E_{\min}|=5.66\pm0.15$ meV; for silicon[5]: $T_c=28\pm2$ K, $|E_{\min}|=22.6\pm0.2$ meV. Calculations of the critical point in germanium and silicon were also carried out in other researches.[6] A similar relation evidently exists for liquid-gas transitions generally; see the note on page 288 of the book of Landau and Lifshitz.[7]

However, we now return to layered system. In the considered model, the gas phase is a nondegenerate plasma, and the excitons in it do not play any significant role. Even at very low temperatures $T\ll T_c$, their concentration $\propto\exp[(\mu_0+1)/T]$ is much less than the concentration of the free carriers, which is of the order of

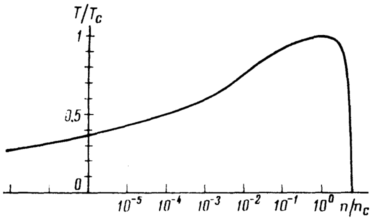

FIG. 2. Phase diagram of the EMP–EHL transition on the (n, T) plane in relative units, $\sigma=1$.

$\exp(\mu_0/2T)$, by virtue of the condition $|\mu_0|\sim c^{-1/2}\gg 2$. At temperatures that are close to critical, the density of the carriers in the gas phase is so large that the existence of excitons in it is generally impossible.

We also give the formulas for the symmetric case $\sigma=1$. The equation for z_c is materially simplified:

$$^4/_3 z_c=\exp(z_c)-1, \tag{33}$$

its solution is $z_c=0.5501$ and Eqs. (21)–(23) reduce to

$$n_c c=\left(\frac{A}{6\pi}z_c\exp(-z_c)\right)^{4/3}\frac{1}{c^{*1/3}}\approx\frac{0.028}{c^{*1/3}}\approx 0.17n_0c, \tag{34}$$

$$T_c=\frac{\pi}{z_c}\left(\frac{A}{6\pi}z_c\exp(-z_c)\right)^{4/3}\frac{1}{c^{*1/3}}\approx\frac{0.161}{c^{*1/3}}\approx 0.103|E_{min}|, \tag{35}$$

$$\mu_c=-2[3+4z_c-\ln(4z_c/3)]T_c\approx-11.0T_c. \tag{36}$$

Up to now we have had in mind a nonequilibrium situation, in which the total number of electrons and holes in the sample is given by the external source of the excitation, although in all the other parameters, these carriers are in equilibrium with the crystal lattice and with one another. In the case of complete thermodynamic equilibrium, i.e., in the absence of an excitation, the formulas (12)–(36) obtained above remain in force, but their meaning is modified somewhat. The condition that the chemical potentials of the electrons be the same in the valence band and the conduction band (4) leads to the replacement in Eq. (12) of μ by $E_g^{(0)}$—the unrenormalized width of the forbidden band:

$$E_g^{(0)}=A\left(\frac{c}{c^*}\right)^{1/4}n^{1/4}-T\ln 2\exp\left(\frac{\pi nc}{T}\right)\left(\operatorname{ch}\frac{\pi nc}{T}-\operatorname{ch}\frac{\pi ncs}{T}\right). \tag{37}$$

Equation (37) determines the dependence of the electron and hole concentration n on the temperature T and on $E_g^{(0)}$.

The actual (renormalized) width of the forbidden band E_g, by virtue of (9) and (10), is equal to

$$E_g(T)=E_g-A(c/c^*)^{1/4}n^{1/4}(T). \tag{38}$$

The gas-liquid phase transition means a discontinuous increase in $n(T)$ and therefore a decrease in $E_g(T)$, i.e., semiconductor—semimetal transition at sufficiently small values of the bare forbidden band. In the plane of the variables $(T, E_g^{(0)})$ (Fig. 3) the curve $E_g^{(0)}=-\mu(T)$, described by the formulas (22)–(25), (27), (29), (31), divides the regions of existence of the semimetal (SM) and semiconductor (SC) phases. In the region $T<T_c$, $|\mu_0|\leqslant E_g^{(0)}<|\mu_c|$ the carrier concentrations n and the width of the forbidden band E_g depend in discontinuous fashion on T and $E_g^{(0)}$:

$$n\leqslant n_{\text{gas}}(T),\ E_g(T)\geqslant|\mu(T)|-A(c/c^*)^{1/4}n_{\text{gas}}^{1/4}(T)$$

for the SC phase and

$$n\geqslant n_{\text{liq}}(T),\ E_g(T)\leqslant|\mu(T)|-A(c/c^*)^{1/4}n_{\text{liq}}^{1/4}(T)$$

for the SM phase. In the actual phase diagram, in place of $E_g^{(0)}$ (Fig. 3), we should have some thermodynamic parameter whose change causes $E_g^{(0)}$ to change, for example, the external pressure or the concentration of one of the components in the case of solid solutions.

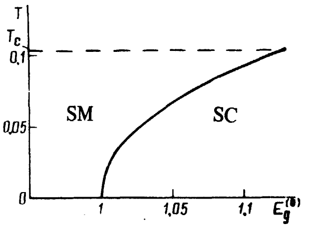

FIG. 3. Phase diagram of the semiconductor–semimetal transition on the plane of the variables $(T, E_g^{(0)})$, with both variables in units of $|E_{\min}|=|\mu_0|$. The continuous curve separates the semiconducting (SC) from the semimetal (SM) phase.

We now proceed to the derivation of Eqs. (9) and (10), on which all the previous consideration was based. The irreducible self-energy part Σ_e (Σ_h can be obtained from ε_e by the obvious substitution $s\to -s$ in all formulas), is expressed by the well known formula

$$\Sigma_e(\mathbf{p},\varepsilon_{k_1})=-T\sum_{k_2}\sum_{l_1}\int\frac{d^2q}{(2\pi)^2}\tilde{V}_{l-l_1}(\mathbf{q},\omega_{k_2})$$
$$\times G_l(\mathbf{p}+\mathbf{q},\varepsilon_{k_1}+\omega_{k_2})\Gamma^{(e)}_{l-l_1}(\mathbf{p},\mathbf{q};\varepsilon_{k_1},\omega_{k_2}). \tag{39}$$

where the screened interaction $\tilde{V}$ is determined by the equation

$$\tilde{V}_{l-l_1}(\mathbf{q},\omega_{k_2})=V_{q,l-l_1}+\sum_{l_3l_4}V_{q,l-l_3}\Pi_{l_3-l_4}(\mathbf{q},\omega_{k_2})\tilde{V}_{l_4-l_1}(\mathbf{q},\omega_{k_2}). \tag{40}$$

The polarization operator

$$\Pi_{l_3-l_4}(\mathbf{q},\omega_{k_2})=\sum_{\alpha=e,h}2T\sum_{k_1}\int\frac{d^2p}{(2\pi)^2}G_\alpha(\mathbf{p},\varepsilon_{k_1})G_\alpha(\mathbf{p}+\mathbf{q},\varepsilon_{k_1}+\omega_{k_2})$$
$$\times\Gamma^{(\alpha)}_{l_3-l_4}(\mathbf{p},\mathbf{q};\varepsilon_{k_1},\omega_{k_2}), \tag{41}$$

$\varepsilon_{k_1}=\pi T(2k_1+1)$, $\omega_{k_2}=2\pi k_2T$; $k_{1,2}$ are integers; $\Gamma^{(\alpha)}_{l-l_1}$ is the vertex part, defined as the sum of all the irreducible diagrams with a single incoming (into the layer l_1) interaction line and two out external electron or hole lines on the same layer l.

In the region of high concentrations $nc\gg 1$ of interest to us, classification of the terms of the series of perturbation theory in powers of the density n is possible. The set of principal diagrams forms, as is well known, the so-called diagrams of the random phase approximation (RPA), corresponding to the use for $\Gamma^{(\alpha)}_{l-l_1}$ and $\Pi_{l_3-l_4}$ of their first nonvanishing approximations:

$$\Gamma^{(\alpha)}_{l-l_1}(\mathbf{p};\mathbf{q};\varepsilon,\omega)=\delta_{l_1l},$$
$$\Pi_{l_3-l_4}(\mathbf{q},\omega_{k_2})=-4\delta_{l_3l_4}\sum_{\alpha=e,h}\int\frac{d^2p}{(2\pi)^2}f_\alpha(\mathbf{p})\frac{\varepsilon_\alpha(\mathbf{p}+\mathbf{q})-\varepsilon_\alpha(\mathbf{p})}{\omega_{k_2}^2+[\varepsilon_\alpha(\mathbf{p}+\mathbf{q})-\varepsilon_\alpha(\mathbf{p})]^2}, \tag{42}$$

where $f_{e,h}(\mathbf{p})$ is the Fermi distribution function. The form of the expression (42) is the usual one (see, for example, Ref. 3, p. 250); in the summation, however, it is now assumed that Σ_α is independent of its arguments, an assumption discussed below.

The expression (39) for Σ reduces in this approximation to

$$\Sigma_e(\mathbf{p},\varepsilon_{k_1})=-\int\frac{d^2q}{(2\pi)^2}\frac{2\pi}{q}f_e(\mathbf{p}+\mathbf{q})$$
$$+T\sum_{k_2}\int\frac{d^2q}{(2\pi)^2}\left[\tilde{V}_0(\mathbf{q},\omega_{k_2})-\frac{2\pi}{q}\right]G_e(\mathbf{p}+\mathbf{q},\varepsilon_{k_1}+\omega_{k_2}), \tag{43}$$

where V_0 is obtained by solution of the Eq. (40):

$$V_0(\mathbf{q}, \omega_{k_2}) = \frac{2\pi}{q}\left\{1 - \frac{4\pi}{q}\Pi(\mathbf{q}, \omega_{k_2}) \operatorname{cth} qc + \frac{4\pi^2}{q^2}\Pi^2(\mathbf{q}, \omega_{k_2})\right\}^{-1/2}. \quad (44)$$

The first term in the right side of (43) is the exchange correction to the dispersion of the electrons; the second is the correlation correction. The exchange part is obviously $\sim nc/\bar{p}$ and, as we shall see below, is small in comparison with the correlation part at concentrations satisfying the inequalities (7) and (8). For estimates of the correlation contribution Σ_{corr}, we consider (44) in more detail.

For the behavior of the polarization operation, a knowledge of the quantity

$$(q/\bar{p})^2 + (\omega_k/q\bar{p})^2$$

is essential, where $\bar{p}$ is the average momentum of the particles, $\bar{p} \sim \max[(2\pi nc)^{1/2}, T^{1/2}]$. In the two limiting cases, we have

$$\Pi(\mathbf{q}, \omega_k) \approx -\frac{1}{\pi}\left[\frac{f_e(0)}{1+s} + \frac{f_h(0)}{1-s}\right] \sim -\frac{nc}{\bar{p}^2}, \quad \left(\frac{q}{\bar{p}}\right)^2 + \left(\frac{\omega_k}{q\bar{p}}\right)^2 \ll 1, \quad (45)$$

$$\Pi(\mathbf{q}, \omega_k) \approx -4\left[\frac{1}{(1+s)(q^2 + (2\omega_k/(1+s)q)^2)} + \frac{1}{(1-s)(q^2 + (2\omega_k/(1-s)q)^2)}\right] nc, \quad \left(\frac{q}{\bar{p}}\right)^2 + \left(\frac{\omega_k}{q\bar{p}}\right)^2 \gg 1. \quad (46)$$

In the region of (7), (8) we have $\bar{p} \ll c^{-1/2}$. Therefore, the transferred momenta $q \gtrsim c^{-1}$ correspond to the asymptotic form (46) of the polarization operator,

$$q^{-1}\Pi \leqslant nc^4 \ll 1$$

and the contribution of the region $qc \gtrsim 1$ to the integral for the correlation part Σ_e is very small. The basic contribution is made by the region $qc \ll 1$. Using this inequality and the condition $|\Pi| \lesssim 1$, which is obvious by virtue of (45), (46) and the definition of p, we get

$$\Sigma_{e\,\text{corr}}(\mathbf{p}, \varepsilon_{k_1}) \approx -T\sum_{k_2}\int\frac{d^2q}{(2\pi)^2}\frac{2\pi}{q}\left[\left(1 - \frac{4\pi}{cq^2}\Pi(\mathbf{q}, \omega_{k_2})\right)^{-1/2} - 1\right] \times G_e(\mathbf{p}+\mathbf{q}, \varepsilon_{k_1} + \omega_{k_2}). \quad (47)$$

Substituting (45) and (46) successively in (47), we can easily establish the fact that at $\bar{p} \gg n^{1/4}$ the basic contribution to (47) is made by $q \lesssim (p/\bar{p})^{1/2} \ll \bar{p}$, to which corresponds the asymptotic form (45), and at $\bar{p} \ll n^{1/4}$ by the momenta $q \sim n^{1/4} \gg \bar{p}$, which corresponds to the asymptotic form (46). But the inequality $\bar{p} \ll n^{1/4}$, by virtue of the definition of $\bar{p}$, is equivalent to the inequalities (7) and (8), so that in the region of concentrations and temperatures of interest to us we should use the polarization operator in the form (46). Moreover, in (47) we can omit $\mathbf{p}$ in the argument of G_e because of the inequality $q \gg \bar{p}$ and, consequently, for real momenta of the particles Σ_{corr} does not depend on $\mathbf{p}$. In the summation over the frequencies, the basic contribution under these conditions is made by $|\omega_{k_2}| \lesssim n^{1/2}$, as is easily understood from (45) and (46). Therefore, by virtue of (8) and of the definition of $\omega_{k_2} = 2\pi k_2 T$, the summation over k_2 can be replaced by the integration

$$T\sum_{k_2} \to \int\frac{d\omega}{2\pi},$$

which is equivalent to neglect of the temperature dependence, i.e., the replacement of Σ_{corr} by its value at $T = 0$.

Finally, the dependence on ε_{k_1} in the argument G_e in (47) also turns out to be unimportant for $|\varepsilon_{k_1}| \lesssim n^{1/2}$, i.e., right up to energies significantly greater than the mean energy of the particles $\bar{p}^2$. With account of all these simplifications, (47) takes the form

$$\begin{aligned}\Sigma_{e\,\text{corr}} &= -\int\frac{d^2q}{(2\pi)^2}\int_{-\infty}^{+\infty}\frac{d\omega}{2\pi}\frac{2\pi}{q}\left[\left\{1 + 4\pi n\left(\frac{1+s}{\omega^2 + [(1+s)^2q^2/2]^2} + \frac{1-s}{\omega^2 + [(1-s)^2q^2/2]^2}\right)\right\}^{-1/2} - 1\right]\frac{1}{i\omega - q^2(1+s)/2 + \mu_e - \Sigma_e} \\ &= \int_0\frac{dq\,d\omega}{\pi}\left[\left\{1 + 4\pi n\left(\frac{1+s}{\omega^2 + [(1+s)^2q^2/2]^2} + \frac{1-s}{\omega^2 + [(1-s)^2q^2/2]^2}\right)\right\}^{-1/2} - 1\right]\frac{-q^2(1+s)/2 + \mu_e - \Sigma_e}{\omega^2 + [q^2(1+s)/2 + \Sigma_e - \mu_e]^2}\end{aligned} \quad (48)$$

or, after the introduction of the new integration variables

$$x = \frac{2\omega}{q^2}, \quad y = (16\pi n)^{-1/4}q,$$

$$\Sigma_{e\,\text{corr}} = \frac{(16\pi n)^{1/4}}{\pi}\int_0^\infty\int dx\,dy\left[\left\{1 + \frac{1}{y^4}\left(\frac{1+s}{x^2 + (1+s)^2} + \frac{1-s}{x^2 + (1-s)^2}\right)\right\}^{-1/2} - 1\right]\frac{1 + s - (\mu_e - \Sigma_e)/2y^2(\pi n)^{1/2}}{x^2 + [1 + s - (\mu_e - \Sigma_e)/2y^2(\pi n)^{1/2}]^2}. \quad (49)$$

The system of equations (11) and (49) together determines the dependence of $\mu_{e,h}$ and $\Sigma_{e,h}$ on n. With relative accuracy $\sim n^{1/4}$ the terms $(\mu - \Sigma)/n^{1/2}$ from the right side of (49) and, after integration over y, (49) reduces to (9) and (10). In connection with this solution, it is necessary to make the following remark. The approximation that we have used is, strictly speaking, not the random phase approximation, although it is similar to it. In the approach that is usual for the RPA, one must use in (5), (39) and (41) the zeroth approximation for the Green's function

$$G_{e,h}^{(0)}(\mathbf{p}, \varepsilon_k) = (i\varepsilon_k - \varepsilon_{e,h}(\mathbf{p}) + \mu_{e,h})^{-1}. \quad (50)$$

Neglect of $\Sigma_{e,h}$ in the integrand of (49) actually corresponds to this. However, the total Green's functions (in the principal approximation in n)

$$G_{e,h}(\mathbf{p}, \varepsilon_k) = (i\varepsilon_k - \varepsilon_{eh}(\mathbf{p}) + \mu_{e,h} - \Sigma_{e,h})^{-1} \quad (51)$$

must be used in (5), (11), (41) and (42). This difference is very important, since the Matsubara diagram technique is developed at fixed μ and not at fixed n. Therefore, substitution of $G^{(0)}$ in (5) and (41) would have yielded, at large negative values of μ of interest to us, exponentially small current carrier densities $n_0 \sim e^{\mu/T}$ and, correspondingly, exponentially small screening of the interaction $\Pi \sim n_0$. In other words, we would have found for each μ only a single solution corresponding at low temperatures to the gas phase and generally not suitable at $T \sim T_c$. The analysis given above shows the existence for the nonlinear system of Eqs. (39)–(41), (51) of a second solution of the form (41), (9), corresponding to high concentrations n at negative μ, i.e., to a liquid phase.

In conclusion, we discuss the limits of applicability of the results. The model used in the present work and described by the Hamiltonian (1) is itself extremely idealized. In real layered semiconductors and semimetals, the probability of an electron transition from one layer to another differs from zero; this leads to the dependence of the energies of the electrons and holes on the quasimomentum components p_z perpendicular to the planes of the layers. If the transition probability is sufficiently small, this dependence is of the form $W\cos(p_z c)$, where W is proportional to the overlap integral of the wave functions of the electrons on neighboring layers. The consideration given above, which does not take into account the dependence of $\varepsilon_{e,h}$ on p_z, is therefore valid only so long as all the energies entering into the calculations significantly exceed the width W of transverse-motion band, i.e., $W \ll c^{-1/2}$. In the case of opposite sign of this inequality, the motion of the electrons is essentially three-dimensional and for the description of the EHL we must use the approach based on the strong anisotropy of the effective masses for motion in the layer (m_t) and perpendicular to it (m_l).[1] The role of the small parameter of the theory is played in this case by $m_l/m_t \ll 1$ instead of c.

[1)] The coefficient A_{II} of Ref. 1, which determines the correlation contribution to the total energy, differs by the factor $\frac{4}{5}$ from the A coefficient used in the present work in the self-energy parts of the electrons and holes. Moreover, in formulas (10) from Ref. 1, there are errors. The correct form of these formulas agrees with formulas (30) and (31) of this paper.

[1] E. A. Andryushin, V. S. Babichenko, L. V. Keldysh, T. A. Opishchenko, and A. P. Silin, Pis'ma Zh. Eksp. Teor. Fiz. **24**, 210 (1976) [JETP Lett. **24**, 185 (1976)].

[2] T. Matsubara, Progr. Theor. Phys. **14**, 351 (1955).

[3] A. A. Abrikosov, L. P. Gor'kov, and I. E. Dzyaloshinskiĭ, Metody kvantovoĭ teorii polya v statisticheskoĭ fizike (Quantum Field Theory Methods in Statistical Physics) Fizmatgiz, 1962.

[4] T. K. Lo, Solid State Commun. **15**, 1231 (1974); C. A. Thomas, T. M. Rice, and J. C. Hensel, Phys. Rev. Lett. **33**, 219 (1974).

[5] A. F. Dite, V. D. Kulakovskiĭ, and V. B. Timofeev, Zh. Eksp. Teor. Fiz. **72**, 1156 (1977) [Sov. Phys. JETP **45**, 604 (1977)].

[6] R. N. Silver, Phys. Rev. **B8**, 2403 (1973); M. Combescot, Phys. Rev. Lett. **32**, 15 (1974); P. Vashishta, S. G. Das, and K. S. Singwi, Phys. Rev. Lett. **33**, 911 (1964); G. Mahler, Phys. Rev. **B11**, 4050 (1975); T. L. Reinecke and S. C. Ying, Phys. Rev. Lett. **35**, 311 (1975).

[7] L. D. Landau and E. M. Lifshitz, Statisticheskaya fizika (Statistical Physics) Ch. 1, Nauka, 1976.

Translated by R. T. Beyer

Phase transitions in the Potts three-component model

V. M. Zaprudskiĭ

All-Union Research Institute for Physicotechnical and Radio Measurements

(Submitted April 19, 1977)
Zh. Eksp. Teor. Fiz. **73**, 1174–1179 (September 1977)

It is shown by the renormalization-group method that the Potts three-component model has a second-order phase transition. The dependence of the critical exponent of the susceptibility on the dimensionality of space is determined.

PACS numbers: 05.70.Fh

A model introduced by Potts back in 1952[1] has recently become the subject of intensive study, with particular attention paid to the question of the order of the transition. The point is that the Hamiltonian of this model contains invariants of third order, and calculation within the framework of the self-consistent field[2] yields a first-order transition at a number $q \geqslant 3$ of the components of the order parameter of the model. On the other hand, an exact calculation of the two-dimensional q-component Potts model[3] yields at $q \leqslant 4$ a second-order phase transition.

We investigate here the three-component Potts model. We show that if the dimensionality d is less than the critical d_c, a second-order phase transition takes place. The obtained critical dimensionality $d_c \approx 4.58$ greatly exceeds 3, and in all probability a more exact calculation will not lead to $d_c < 3$.

The Potts model is a generalization of the Ising model, in which each site can be in one of the q states, the interaction energy assuming only two values, depending on whether the nearest sites are in identical or different states. The Potts model with various q describes different real systems: electric circuits at $q=0$,[3] percolation at $q=1$,[4] and nematic liquid crystals[5] or crystals with cubic symmetry[6–8] at $q=3$. Of course, at $q=2$ we obtain the usual Ising model.

In the q-component Potts model, a phase transition is described by a tensor order parameter Q_{ij}.[4] The tensor Q_{ij} has no trace, is symmetric, and has the dimensionality q. In particular, in the phase transition from a nematic liquid crystal into an isotropic liquid, the role of the order parameter can be played by the anisotropic part of the dielectric tensor.[5] The Hamiltonian of the model is of the form

$$H=AQ^2+BQ^3+CQ^4, \tag{1}$$

where Q^2, Q^3, and Q^4 denote respectively invariants of second, third and fourth order in Q.