Journal of Experimental and Theoretical Physics, Vol. 92, No. 2, 2001, pp. 195–202.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 119, No. 2, 2001, pp. 227–235.
Original Russian Text Copyright © 2001 by Gangrsky, Hradecný, Yermolajev, Zemlyanoĭ, Téthal.

NUCLEI, PARTICLES, AND THEIR INTERACTION

Optical Polarization of ^{24}Na Nuclei in a Trap Based on Photoinduced Drift

Yu. P. Gangrsky[a, *], C. Hradecný[b], I. M. Yermolajev[b], S. D. Zemlyanoĭ[a], and T. Téthal[b]
[a] *Joint Institute for Nuclear Research, Dubna, Moscow oblast, 141980 Russia*
[b] *Institute of Nuclear Physics, Rez, 25068 Czech Republic*
**e-mail: gang@cv.jinr.ru*

Received March 30, 2000

Abstract—The results of experiments on the polarization of nuclei of the ^{24}Na radioactive isotope in a gas trap are presented. The action of the trap is based on the effect of the photoinduced drift of atoms caused by laser radiation. The anisotropy of γ-radiation emitted by ^{24}Na during the illumination of sodium vapor in the trap by circularly polarized laser radiation is measured and its dependence on the temperature of the trap walls is analyzed. A degree of polarization of ^{24}Na nuclei of $62 \pm 2\%$ is attained at temperatures above 1300 K, for which a rapid desorption of Na atoms from the walls and the decomposition of Na-based compounds take place. *© 2001 MAIK "Nauka/Interperiodica".*

Since 1950, when the possibility of the optical spin orientation of atoms was predicted by Kastler [1], a large number of works devoted to the creation of ensembles of free polarized atoms and nuclei by the optical pumping method appeared (a detailed description of the method of optical orientation is given in the review by Happer [2]). Polarized gas ensembles of radioactive isotopes could be used for determining the nuclear electric and magnetic moments as well as the multipolarity of radiative transition [3] or for a detailed analysis of fundamental processes such as β-decay [4, 5]. However, the main limitation in the experiments with polarized radioactive isotopes in gas cells is the adsorption of the atoms under investigation at the cell walls and the depolarization of the atoms. For this reason, the experiments on optical polarization involving the measurement of the nuclear radiation anisotropy were carried out either on chemically inert atoms, or on isotopes with a lifetime shorter than or comparable with the time of diffusion of an atom under investigation to the cell walls [3, 6]. The effect of the walls can be reduced by choosing an appropriate coating and heating or can be eliminated almost completely by using magnetooptical traps [7]. However, the former method is applicable to a limited number of elements, while traps are complex and fine experimental setups with a small capacity for accumulating atoms (less than 10^8 as a rule) and a low efficiency of trapping [8, 9]. Besides, the electromagnetic field confining ions may cause the depolarization of nuclei.

An alternative method for reducing the effect of the walls of a gas cell is based on the effect of the photoinduced drift (PID) of atoms in a gas under the action of laser radiation [10]. This effect makes it possible to concentrate the atoms contained in the gas cell in a small volume (much smaller than the volume of the cell) and to suppress the diffusion of atoms to the walls. However, the optical orientation of atoms reduces the probability of the absorption of polarized laser radiation, thus decreasing the drift velocity and suppressing the PID effect. This negative influence of polarization on PID can be eliminated almost completely by the spatial separation of the regions in which the optical pumping and the PID take place. This idea was implemented in our earlier work [11] for the stable ^{23}Na isotope: 10^{11} atoms with a degree of polarization of 90% were confined with the help of PID in a tube sealed at one end (we measured the polarization of resonance-scattered optical radiation corresponding to the D_1 line of Na).

However, such an experimental setup could not reveal Na atoms adsorbed at the tube walls and hence depolarized since such atoms make zero contribution to the resonance-scattered radiation. In order to obtain a real pattern and to determine the degree of polarization affecting all atoms including those adsorbed at the walls, experiments with unstable nuclides are required. The degree of polarization in such experiments could be determined from the radioactive radiation. This work is devoted to such an experiment, which includes the polarization of the nuclei of the ^{24}Na radioactive isotope in a PID-based trap and the determination of the degree of polarization from the anisotropy of the γ-radiation as well as the influence of various conditions on the degree of polarization. The ^{24}Na isotope is a convenient object which is often used for such experiments (the polarization of its nuclei exposed to a short high-intensity pulse of laser radiation was observed by us earlier [12]). The sequence of spins and the multipolarity of radiative transitions in the daughter ^{24}Mg nucleus lead to a noticeable anisotropy of the γ-radiation emitted by polarized ^{24}Na nuclei.

The idea of a PID-based trap can be presented as follows. If the drift of atoms in the tube is limited by a partition, the atoms accumulate near it until the drift is compensated by the backward diffusion flow. In optimal experimental conditions, the photoinduced drift can create high density gradients, and most of the atoms are confined near the partition in a layer having a thickness of just a few millimeters.

An important property of a PID-based trap (which will be henceforth referred to as just a trap) is its selectivity. If atoms in a trap stop absorbing radiation after the formation of molecules, the PID stops acting on them, and the atoms leave the trap as a result of diffusion. Moreover, owing to the isotopic shift and (or) different modes of the hyperfine splitting for two different isotopes in the optical transition being excited, the wavelength of the exciting laser radiation can be tuned so that the PID of atoms of these two isotopes will occur in opposite directions. This peculiarity of PID was demonstrated in our earlier work [13] for two radioactive isotopes ^{22}Na and ^{24}Na. Thus, only a definite isotope in the atomic form is confined in the trap which can serve as an isotope separator.

It is well known that at temperatures near 400 K, the adsorption of alkali metals at the walls can be suppressed by a paraffin coating [2]. Unfortunately, the lifetime of atoms (relative to chemical binding) in cells with such a coating does not exceed a second, which sets a limit on their application. For this reason, in our experiments we suppressed the adsorption at the walls by heating the sapphire cell to a high temperature instead of coating the walls with paraffin [11]. In our subsequent experiments, we measured the energy of adsorption of sodium atoms on the sapphire surface, which proved to be equal to 2.48 ± 0.1 eV. This value is close to the binding energy of sodium with oxygen in the Na_2O molecule (2.55 eV [14]). For such a binding energy, the sapphire surface should be heated to a temperature above 1800 K to suppress the adsorption completely. If we coat the sapphire surface with a potassium layer, the adsorption energy for sodium atoms decreases to 1.76 ± 0.07 eV [14], and the temperature required for suppressing the adsorption decreases to 1300 K. This makes it possible to carry out experiments on the polarization of sodium at more accessible temperatures.

At high temperatures, the suppression of adsorption takes place along with the decomposition of molecules in which the captured atoms are bound chemically. The lifetime of a diatomic molecule with the binding energy E is defined by the well-known Arrhenius formula

$$\tau_{\rm mol} = \tau_0 \exp\left(\frac{E}{kT}\right), \qquad (1)$$

where τ_0 is the pre-exponential factor, k is Boltzmann's constant, and T is the absolute temperature. Substituting into expression (1) the typical values for sodium bound with oxygen ($\tau_0 \sim 10^{-13}$ s, $E = 2.55$ eV, and temperature $T = 1300$ K), we obtain $\tau_{\rm mol} \approx 0.5$ ms. This time is too short for the molecules to leave the trap due to diffusion. The atoms formed again drift to the closed end of the trap, where they are confined. Thus, the reversibility of chemical binding processes at high temperatures makes it possible to confine atoms in the trap without losses due to chemical binding. For the laser radiation intensity used in our experiments (exceeding 10^{19} photon/(s cm^2)), the polarization of ^{24}Na nuclei as a result of the optical pumping (including a number of consecutive acts of excitation and spontaneous de-excitation) occurs during a time shorter than a millisecond. For this reason, virtually all sodium atoms in the laser radiation zone are polarized.

Let us estimate the time of confinement of atoms in the trap, which is important for carrying out experiments. For the sake of simplicity, we assume that the adsorption at the cell walls is insignificant, i.e., the number of adsorbed sodium atoms is much smaller than the number of free atoms. We also neglect the losses associated with the radioactive decay and chemical binding. In this case, the losses of atoms in the trap are determined by their diffusion flow through the open end of the tube:

$$\frac{dN(t)}{dt} = -D\frac{dn(L,t)}{dx}, \qquad (2)$$

$$N(t) = \int_0^L n(x,t)\,dx, \qquad (3)$$

where $n(x, t)$ is the number density of atoms, D is the diffusion coefficient, and the values $x = 0$ and $x = L$ correspond to the closed and open ends of the trap. We assume that the time of confinement of atoms in the trap is much longer than the characteristic times of diffusion and drift:

$$\frac{\partial n(x,t)}{\partial t} \gg D\frac{\partial^2 n(x,t)}{\partial x^2}, \qquad (4)$$

$$\frac{\partial n(x,t)}{\partial t} \gg u\frac{\partial n(x,t)}{\partial x}, \qquad (5)$$

where u is the PID velocity. In this case, the number density of atoms is determined by the steady-state equation

$$D\frac{\partial^2 n(x,t)}{\partial x^2} - u\frac{\partial n(x,t)}{\partial x} = 0. \qquad (6)$$

Solving this equation with the boundary condition $n(L, t) = 0$, we find the following expression for the number density of atoms:

$$n(x, t) = n_0(t)\left[\exp\left(\frac{ux}{D}\right) - \exp\left(\frac{uL}{D}\right)\right], \quad 0 < x < L. \tag{7}$$

It should be recalled that $u < 0$. Assuming that the drift velocity is high ($|u|L \gg D$), we evaluate the integral in (3):

$$N(t) \approx -n_0(t)\frac{D}{u}. \tag{8}$$

Substituting (5) and (6) into Eq. (2), we find the solution in the form

$$N(t) \approx N_0 \exp\left(-\frac{t}{\tau}\right), \tag{9}$$

$$\tau = \frac{D}{u^2}\exp\left(\frac{|u|L}{D}\right). \tag{10}$$

Let us find a numerical estimate of τ, viz., the time of the confinement of atoms in the trap for the typical parameters of the experiments. For a pressure of the buffer gas (krypton) of 30 torr at 1300 K, the values of the diffusion coefficient and the drift velocity measured in [10, 11] are $D = 50$ cm^2/s and $u = 150$ cm/s. For the trap length $L = 10$ cm, the confinement time amounts approximately to 2×10^{10} s, or more than 600 years. It should be noted, however, that the situation changes radically when atoms are confined in the trap by circularly polarized light, and the effective optical pumping takes place. The optical orientation of atoms lowers the probability of radiation absorption and, as a result, decreases the PID velocity approximately by an order of magnitude [11]. The time of the confinement of atoms in the trap decreases to 4.5 s in view of its exponential dependence on the PID velocity. This can be avoided by placing the open end of the tube in a transverse magnetic field destroying the optical orientation of the atoms. In this way, the regions of effective influence of the PID and optical pumping are spatially separated: the drift is effective at the open end, confining the atoms in the trap, while the optical orientation of the confined atoms takes place at the sealed end. While estimating the confinement time for atoms on the basis of formula (10), we must take the length of the trap region in which the PID is operating actively (i.e., the larger part of the tube) instead of the entire length of the trap. This decreases the value of τ insignificantly.

It should be recalled that the optical polarization of radioactive ^{24}Na atoms ($T_{1/2} = 15.02$ h) by an optical radiation with circular polarization with a frequency tuned to the D_1 transition in the presence of a buffer gas has some peculiarities. The hyperfine splitting of the ground state of ^{24}Na is smaller than the Doppler's broadening so that both hyperfine sublevels interact with the optical radiation. It is essential for the polarization that in the course of the transition $S_{1/2} \longrightarrow P_{1/2}$, the sublevel corresponding to the maximum component of the atomic angular momentum $m_F = +9/2$ (nuclear spin $I = 4$) is unable to absorb a photon with the right polarization σ^+ by virtue of the momentum conservation law, while atoms from all the remaining sublevels go over to an excited state. The cross section of mixing of sublevels of the $P_{1/2}$ state as a result of collisions with particles of the buffer gas is of the order of 10^{-14} cm^2 [2], which exceeds the gas-kinetic collision cross section. Thus, for buffer gas pressures which are optimal for the PID effects (approximately 30 torr [15]), the sublevels of the $P_{1/2}$ state are mixed completely, and the excited state is depolarized. As a result of spontaneous relaxation, the sublevels of the ground state are populated uniformly, the sublevel with $m_F = +9/2$ being also populated with a probability of 1/18. The cross-section of the mixing of sublevels belonging to the $S_{1/2}$ state due to collisions with krypton atoms is much smaller (approximately equal to 2.0×10^{-21} cm^2 [2]). Thus, the nonuniformity in the population of the sublevels of the ground state emerging as a result of multiple excitation and spontaneous relaxation processes is accumulated. In this case, the atmosphere of sodium atoms with the maximal angular momentum component becomes completely transparent for the circularly polarized laser radiation tuned to D_1 transition.

The degree of the polarization of ^{24}Na nuclei was determined from the anisotropy of its γ-radiation emitted as a result of β-decay. This γ-radiation consisted of two γ-quanta with energies of 2.754 and 1.369 MeV in a cascade [16]. The sequence of spins and parities of the nuclear energy levels in the β- and γ-transitions was as follows:

$$4^+ \xrightarrow{\beta} 4^+ \xrightarrow{\gamma} 2^+ \xrightarrow{\gamma} 0^+.$$

The angular diagram of the emission of γ-quanta in the nuclear transition of the daughter ^{24}Mg nucleus is described by the relation [17]

$$W(\theta) = \sum_{\lambda = 0}^{2L} B_\lambda(I_0)U_\lambda A_\lambda(LLI_fI_i)P_\lambda(\cos\theta), \tag{11}$$

where θ is the angle between the directions of orientation of the nuclear spin and the emitted γ-quantum, $B_\lambda(I_0)$ are the parameters of the orientation of the parent nucleus, U_λ are the parameters of disorientation as a result of the transitions preceding the observed transition, $A_\lambda(LLI_fI_i)$ are the angular distribution coefficients, $P_\lambda(\xi)$ are the Legendre polynomials, L is the multipolarity of the transition, I_0 is the spin of the parent nucleus, and I_f and I_i are the final and initial values of the nuclear spin in the transition. The summation in (11) is carried out over only even values of λ.

The orientation parameters are defined by the relation

$$B_\lambda(I_0) = \sqrt{2\lambda+1}\sqrt{2I_0+1} \times \sum_{m=-I_0}^{I_0} (-1)^{I_0+m} \begin{pmatrix} I_0 & I_0 & \lambda \\ -m & m & 0 \end{pmatrix} p(m), \quad (12)$$

where the parentheses contain the 3*j* symbols, and $p(m)$ is the occupancy of a sublevel with the nuclear spin component m (m assumes the values from –4 to +4). The main depolarizing processes accompanying the optical pumping are the absorption of unpolarized spontaneous radiation and the relaxation at the walls. In the former process, all the m-sublevels of the ground state are populated with the same probability. We can assume that the same applies to the relaxation of the atomic angular momentum at the wall. Thus, concerning the distribution over m sublevels of the ground state, we can assume that the population of all the m_F-sublevels (except that with $m_F = +9/2$) is the same, while the elevated population of the sublevel with $m_F = +9/2$ indicates the polarization of atoms. In this case, the population of sublevels $p(m)_{m\neq 4} = p_0$, while $p(4) = p_0 + \Delta$, and the orientation parameters are

$$B_2(4) = 1.5954\Delta, \quad B_4(4) = 0.9387\Delta, \quad (13)$$

where Δ indicates the degree of polarization of ^{24}Na nuclei. The disorientation parameters are determined by all the previous transitions:

$$U_\lambda = \begin{cases} U_\lambda(\beta), & 4 \longrightarrow 2, \\ U_\lambda(\beta)U_\lambda(42), & 2 \longrightarrow 0. \end{cases} \quad (14)$$

It is natural to carry out the experimental observation of the anisotropy of the γ-radiation in directions where it has the maximum value, i.e., for the values of the angle $\theta = 0$ and 90°. We assume that the β-transition is a purely Gamow–Teller transition. Then the coefficients in formula (14) have the following values [17]:

$$\begin{gathered} U_2(\beta) = 0.8500, \quad U_2(42) = 0.7491, \\ U_4(\beta) = 0.5000, \quad U_4(42) = 0.2847, \\ A_2(2224) = -0.4477, \quad A_2(2202) = -0.5976, \\ A_4(2224) = -0.3044, \quad A_4(2202) = -1.0690, \\ P_2(\cos 0) = 1, \quad P_2(\cos(\pi/2)) = -1/2, \\ P_4(\cos 0) = 1, \quad P_4(\cos(\pi/2)) = -3/8. \end{gathered} \quad (15)$$

Substituting these parameters into formula (11), we obtain identical expressions for both transitions:

$$W(0) = 1 - 0.7500\Delta, \quad W(90) = 1 + 0.2500\Delta. \quad (16)$$

The block diagrams of the experimental setup are presented in Fig. 1. The main element of the trap was a sapphire capillary of length 20 cm and inner diameter 2.5 mm, which was sealed at one end. The capillary was placed in a tantalum coil and enclosed in another sapphire tube of a larger diameter. The coil ensured a high temperature of heating and created a magnetic field parallel to the direction of laser beam propagation. The strength of the magnetic field created by the coil in our experiments was 50–80 Oe. The cell was filled with the buffer gas (krypton) under a pressure of 30 torr. Before starting the experiment, the coil was calibrated: the temperature was measured with the help of a thermocouple located directly in the capillary at the sealed end for various values of the voltage across the coil under the experimental pressure of the buffer gas. A wire made of ultrapure aluminum, having a mass of 20 mg, and irradiated by fast neutrons in the LVR-15 nuclear reactor at the Institute of Nuclear Physics, Rez (Prague region) was placed on a tantalum foil in the capillary at a distance of 5 cm from the open end. As a result of the reaction ^{27}Al(n, α), ^{24}Na was obtained in an amount of 1.5×10^{11} atoms. The intensity of the γ-radiation emitted by it was 5.5×10^6 Bq. Since tantalum is wet by molten aluminum, the tantalum foil prevented the formation of a drop of liquid aluminum in the sapphire capillary, which could block the laser beam.

Permanent magnets creating a transverse magnetic field of a strength of approximately 100 Oe (which was sufficient for destroying the spin orientation of atoms in this half of the trap) were arranged in the vicinity of the cell at the open end. In the other half of the trap close to the sealed end, the strength of the field created by the permanent magnets rapidly decreased and was below 10 Oe at a distance of 3 cm from the magnets, which did not affect the orientation of atoms. The beam emitted by a dye laser (manufactured at the joint-stock firms “Inversiya” and “Tekhnoskan”, Novosibirsk) passed through the capillary from the open to the sealed end. The beam power was approximately 100 mW, and the radiation frequency was tuned to the center of the D_1 line emitted by stable sodium, which corresponds to an approximately 700-MHz detuning from the center of the line emitted by ^{24}Na to the “red” side (isotopic shift). This ensured a PID velocity in the direction of light propagation, which was close to its maximum value. The beam had a Gaussian profile, its diameter could be varied from 0.5 to 1.5 mm, and the polarization of the radiation was linear.

The larger tube contained an ampule with metallic potassium placed at a distance of 8 cm from the entrance to the trap. In the course of the experiment, the ampule was permanently heated so that the light of the resonance fluorescence of sodium contained in potassium as a small admixture could be seen. The temperature of the ampule with potassium was close to that at the sealed end of the tube and was measured by the thermocouple. Potassium vapor purified the buffer gas

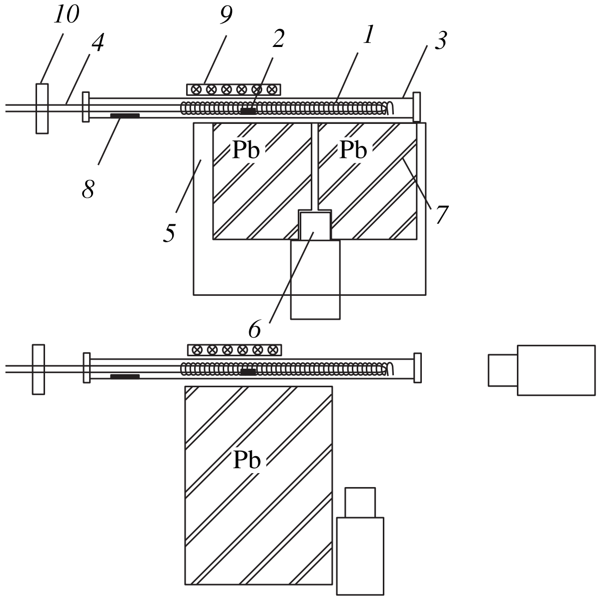


Fig. 1. Block diagrams of the experimental setups for measuring (top) the spatial distribution of ^{24}Na in the cell and (bottom) the anisotropy of the γ-radiation emitted by trapped ^{24}Na atoms: (*1*) sapphire capillary sealed at one end, (*2*) aluminum containing ^{24}Na atoms, (*3*) sapphire tube, (*4*) dye laser beam, (*5*) movable platform, (*6*) γ-detector, (*7*) lead shield with a slit, (*8*) ampule with metallic potassium; (*9*) permanent magnet creating a transverse magnetic field sufficient to destroy the optical polarization in a half of the trap, and (*10*) λ/4 polarization plate.

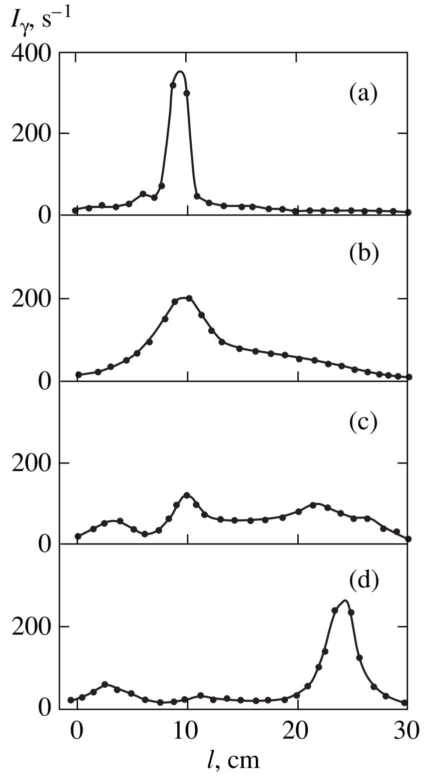


Fig. 2. Evolution of the ^{24}Na distribution over the length l of the cell during its preparation for the operation: (a) beginning of heating; (b) after 2.5 h; (c) after 4.5 h; (d) after 5.5 h. I_γ is the intensity of γ-radiation, $l = 0$ and $l = 30$ cm correspond to the open and sealed ends of the tube.

and decreased the adsorption of sodium at the sapphire surface which was gradually saturated with potassium.

The capillary was gradually heated, and sodium atoms started leaving the aluminum, which was heated and then melted. The spatial distribution of radioactive sodium during heating was determined with the help of a scintillation detector placed behind the sliding lead slit (see Fig. 1). The aluminum and the tantalum foil contained a considerable amount of stable sodium as an impurity. As a result, the medium became optically opaque at high temperatures, which hindered the experiments. For this reason, we first had to remove stable sodium from the trap, i.e., to separate the ^{23}Na and ^{24}Na isotopes in the trap. At this stage of the experiment, we selected a temperature at which the concentration of sodium being evaporated was not too high and the laser radiation penetrated to the sealed end of the cell. In due time, stable sodium left the capillary through diffusion, while ^{24}Na was concentrated at the closed end of the capillary owing to the PID effect.

The evolution of the spatial distribution of ^{24}Na upon a gradual increase in temperature is shown in Fig. 2. After 5.5 h, the capillary was absolutely free of stable sodium, and the process of separation of the ^{23}Na and ^{24}Na isotopes was completed. The glow of stable sodium in the tube disappeared, and only the luminescence of ^{24}Na was observed at the sealed end of the tube. The γ-radiation intensity measurements proved that approximately 10^{11} atoms of ^{24}Na were present in this region. It should be noted that this value is not the limiting trap capacity since the latter is determined by the trap volume which can be made quite large.

In order to calculate the enrichment factor (the variation of the isotope ratio) attained in the experiments, we must estimate the amount of stable sodium removed from the trap. The estimate can be obtained on the basis of formula (2) by substituting the experimental values of the parameters appearing in this formula, i.e., 5.5 h for the separation time and 10^{12} cm^{-3} for the sodium vapor density gradient. The diffusion coefficient of sodium under a krypton pressure of 30 torr taking into account the adsorption at the walls at a temperature near 1000 K is of the order of 0.1 cm^2/s[18, 19]. It follows from this estimate that the amount of stable sodium is 2×10^{14} atoms, i.e., three orders of magnitude larger than the amount of radioactive sodium. The absence of a glow from stable sodium suggests that its amount in the trap after the isotope separation is

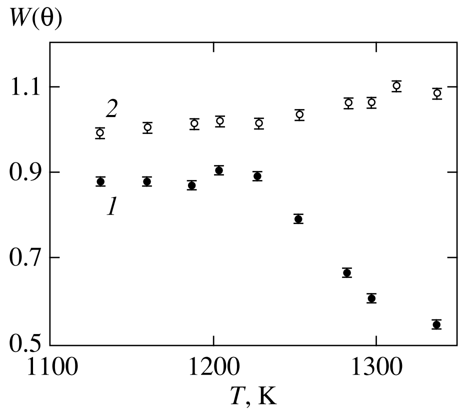


Fig. 3. Anisotropy of the γ-radiation at various temperatures of the cell: (*1*) readings of the detector arranged at 0°, (*2*) readings of the detector at 90°.

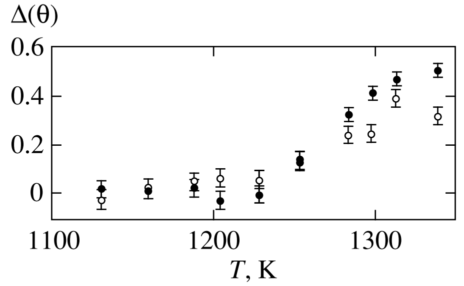


Fig. 4. Degree of polarization of ^{24}Na nuclei in the trap, obtained from the readings of the detector arranged at 0° (dark circles) and 90° (light circles) at various temperatures of the cell.

smaller than one tenth of the amount of ^{24}Na, i.e., less than 10^{10} atoms. This means that the enrichment factor attained in the experiment was larger than 10^4. In the course of isotopic separation, we lost only about 20% of ^{24}Na, which left the trap and was adsorbed at its open end. This determines the efficiency of the ^{24}Na trapping attained in the experiment, which was 80%.

After the spatial distribution of ^{24}Na was stabilized, we made experiments on the optical polarization of ^{24}Na in the trap. For this purpose, we measured the anisotropy of the γ-radiation emitted during the decay of ^{24}Na. The measurements were made with the help of two scintillation detectors having working volumes of 80 and 57 cm^3 and arranged at angles of 0 and 90° to the wave vector of the laser beam at 30 cm from the closed end of the trap. The polarization of light was changed from linear to circular polarization and back with the help of a λ/4 phase plate. The γ-radiation was detected at 2.754- and 1.369-MeV lines. The intensity of this radiation was measured by each detector alternately for the linear polarization of the laser radiation and for the circular polarization (in this case, ^{24}Na nuclei were polarized). Each act of measurement lasted for 10 min during which the detectors recorded at least 4×10^4 γ-quanta in the chosen spectral region. These measurements resulted in the ratios

$$K_0 = \frac{K_{0,c} - f_1}{K_{0,l} - f_1}, \quad K_{90} = \frac{K_{90,c}}{K_{90,l}},$$

where $K_{0,c}$ and $K_{0,l}$ are the numbers of γ-quanta recorded by a given detector for the circular and linear polarizations of the laser beam, respectively, and f_1 is the background from ^{24}Na escaping from the trap in the detector arranged at 0°. Figure 3 shows the dependence of these ratios on the cell temperature (for a laser beam diameter of 1.0 mm).The statistical errors of measurements are given with a 90% confidence level of the probability. It should be noted that at temperatures below 1230 K, the values of K_θ are independent of temperature, and their mean values are $K_0 = 0.89 \pm 0.02$ and $K_{90} = 1.01 \pm 0.02$. The constant values of K_0 at temperatures below 1230 K can be explained by the fact that the amount of ^{24}Na adsorbed at the trap walls at low temperatures is much smaller than the amount of ^{24}Na in the gas phase. For this reason, the anisotropy in the γ-radiation of polarized ^{24}Na in the gas phase does not exceed the statistical error for the isotropic radiation emitted by unpolarized ^{24}Na adsorbed at the walls. It was noted above that when linear polarization changes to circular polarization, the probability of light absorption decreases due to the optical pumping, and the PID velocity of ^{24}Na atoms decreases accordingly, leading to the atomic distribution blurring. As a result, the distance between the center of gravity of the distribution and the detector arranged at 0° increases, and this detector records a smaller number of γ-quanta. The value of the quantity $K_0 = f_2 = 0.89 \pm 0.02$ indicated above corresponds to a shift in the center of gravity of the distribution of trapped atoms by 1.8 ± 0.4 cm. This systematic error can be taken into account while calculating anisotropy $W(0)$, and $W(0) = K_0/f_2$. The value of anisotropy $W(90)$ directly corresponds to ratio K_{90}, i.e., $W(90) = K_{90}$.

As the temperature of the cell was increased further, the anisotropy of the γ-radiation became larger than the statistical error of measurements. Ratio K_0 started decreasing, and ratio K_{90} increased, both values attaining the saturation at temperatures above 1350 K. The degree of polarization Δ of nuclei, calculated using the data from Fig. 3 in accordance with formula (16), are presented in Fig. 4 as a function of temperature.

The degree of polarization of atoms in the trap is determined by the dynamic equilibrium between the polarization of atoms in the bulk and their depolarization at the wall. From this point of view, it was interesting to measure the anisotropy for various diameters of the laser beam since an increase in its diameter leads to an increase in the effective volume in which the optical pumping takes place and the PID effect is operating. The table contains the results of measurements of the

Degree of polarization of ^{24}Na nuclei for various temperatures and laser beam diameters

T, K	d = 0.5 mm		d = 0.8 mm		d = 1.4 mm	
	Δ_{90}, %	Δ_0, %	Δ_{90}, %	Δ_0, %	Δ_{90}, %	Δ_0, %
1255	14 ± 4	14 ± 3	–	–	–	–
1285	24 ± 4	33 ± 3	–	–	–	–
1300	25 ± 4	42 ± 3	–	–	–	–
1315	40 ± 4	48 ± 2	49 ± 4	56 ± 2	47 ± 4	58 ± 2
1340	33 ± 4	51 ± 2	51 ± 4	56 ± 2	60 ± 4	61 ± 2
1363	–	–	–	–	60 ± 4	63 ± 2

degree of polarization from the readings of both detectors for various values of the laser beam diameter d and of the sample temperature. It can be seen from the table that the γ-radiation anisotropy increases with the beam diameter, while the difference in the anisotropy readings of the detectors decreases. This is due to the fact that the maximum PID velocity averaged over the capillary cross section is attained when the radiation fills the capillary to the maximum possible extent. As a result, the distribution of trapped atoms in this case becomes more compact and is less blurred (in absolute measure) upon a transition from linear to circular polarization. Accordingly, the effect of random variations of the geometrical factor on the counting intensity of the detectors decreases. The maximum degree of polarization of ^{24}Na nuclei attained in our experiments (and averaged over the two detectors) was 62 ± 2%. One of the possible ways of increasing the degree of polarization is the addition of an inert molecular gas (e.g., nitrogen) to the buffer gas for quenching the depolarizing luminescence.

The experiments on the optical polarization of ^{24}Na and on the measurement of its degree lasted approximately 6 h. During this time, the losses of trapped ^{24}Na atoms corresponded only to their radioactive decay to within the error of measurements, i.e., the losses due to chemical binding were less than 2%. This means that if molecules were formed in the trap, they were rapidly decomposed into atoms at the experimental temperatures (1100–1400 K) before they could leave the trap as a result of diffusion. Such low losses correspond to a very long formal time of the confinement of ^{24}Na atoms in the trap due to the PID effect (exceeding 300 h).

In conclusion, we list the main results obtained in this work. It was demonstrated that a PID-based trap can operate as an isotopic separator with low losses and a large enrichment factor. In the course of the separation, the ratio of the ^{23}Na and ^{24}Na isotopes changed by a factor exceeding 10^4 for losses of 20%. The trap captured 10^{11} atoms of ^{24}Na, which were confined for 6 h in the form of vapor virtually without a leakage (losses were less than 2%). The trapped ^{24}Na atoms were polarized by laser radiation with circular polarization. The dependence of the degree of polarization on the capillary temperature was measured. It was found that at temperatures above 1300 K, at which the rapid desorption of Na atoms from the walls and the decomposition of the formed molecules took place, the degree of polarization of ^{24}Na atoms reached 62 ± 2%. The obtained results demonstrate the high efficiency of the developed method of the polarization of nuclei and raise hopes that this method will be applied in experiments with polarized atoms and nuclei.

ACKNOWLEDGMENTS

The authors are deeply indebted to A.M. Shalagin, E.V. Podivilov, D. Venos and E. Šimečkova for fruitful discussions, R. Mach and M.G. Itkis for their support, and also J. Franjo for the measuring equipment kindly provided for the experiments.

This work was financed by the Grant Agency of the Czech Republic (grants nos. 202/97/0369 and 202/97/0445).

REFERENCES

1. A. Kastler, J. Phys. Radium **11**, 225 (1950).
2. W. Happer, Rev. Mod. Phys. **44**, 170 (1972).
3. G. Huber, J. Bonn, H.-J. Kluge, and E. W. Otten, Z. Phys. A **276**, 203 (1976).
4. A. L. Hallin, F. P. Calaprice, and D. W. MacArthur, Z. Phys. A **276**, 187 (1976).
5. M. Stephens and C. Wieman, Phys. Rev. Lett. **72**, 3787 (1994).
6. G. Shimkaveg, W. W. Quivers, R. R. Dasari, *et al.*, Phys. Rev. Lett. **53**, 2230 (1984).
7. E. L. Raab, M. Prentiss, A. Cable, *et al.*, Phys. Rev. Lett. **59**, 2631 (1987).
8. K. Gibble, S. Kasapi, and S. Chu, Opt. Lett. **17**, 526 (1992).
9. W. Ketterle, K. B. Davis, M. A. Joffe, *et al.*, Phys. Rev. Lett. **70**, 2253 (1993).

10. F. Kh. Gel'mukhanov and A. M. Shalagin, Pis'ma Zh. Éksp. Teor. Fiz. **29**, 773 (1979) [JETP Lett. **29**, 711 (1979)].
11. F. Buric, C. Hradecný, J. Slovák, *et al.*, Phys. Rev. A **54**, 3250 (1996).
12. Yu. P. Gangrsky, C. Hradecný, D. F. Zaretskiĭ, *et al.*, Zh. Éksp. Teor. Fiz. **96**, 791 (1989) [Sov. Phys. JETP **69**, 449 (1989)].
13. Yu. P. Gangrsky, C. Hradecný, S. G. Zemlyanoĭ, *et al.*, Zh. Éksp. Teor. Fiz. **106**, 725 (1994) [JETP **79**, 399 (1994)].
14. *Thermodynamic Properties of Individual Substances. Reference Book*, Ed. by V. P. Glushko (Nauka, Moscow, 1982).
15. H. G. C. Werij and J. P. Woerdman, Phys. Rep. **169**, 145 (1988).
16. E. Brown and R. B. Firestone, *Table of Radioactive Isotopes* (Wiley, New York, 1986).
17. N. J. Stone and H. Postma, *Low-Temperature Nuclear Orientation* (North-Holland, Amsterdam, 1986).
18. C. Hradecný, J. Slovák, T. Téthal, *et al.*, Appl. Radiat. Isot. **43**, 1259 (1992).
19. C. Hradecný, J. Slovák, T. Téthal, *et al.*, Appl. Radiat. Isot. **45**, 257 (1994).

Translated by N. Wadhwa

Journal of Experimental and Theoretical Physics, Vol. 92, No. 2, 2001, pp. 203–209.
From Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 119, No. 2, 2001, pp. 236–242.
Original English Text Copyright © 2001 by Kuraev, Lipatov, Shishkina.

NUCLEI, PARTICLES, AND THEIR INTERACTION

QED Radiative Corrections to Impact Factors[¶]

E. A. Kuraev[a, *], L. N. Lipatov[b], and T. V. Shishkina[c]

[a]*Bogoliubov Laboratory of Theoretical Physics, Joint Institute for Nuclear Researches,*
Dubna, Moscow oblast, 141980 Russia
**e-mail: kuraev@thsun1.jinr.ru*
[b]*St. Petersburg Nuclear Physics Institute, Gatchina, Leningrad oblast, 188350 Russia*
[c]*Belarus State University, Minsk, 220040 Belarus*

Received August 16, 2000

Abstract—We consider radiative corrections to the electron and photon impact factors. The generalized eikonal representation for the e^+e^- scattering amplitude at high energies and fixed momentum transfers is violated by nonplanar diagrams. An additional contribution to the two-loop approximation appears from the Bethe–Heitler mechanism of fermion pair production with the identity of the fermions in the final state taken into account. The violation of the generalized eikonal representation is also related to the charge parity conservation in QED. A one-loop correction to the photon impact factor for small virtualities of the exchanged photon is obtained using the known results for the cross section of the e^+e^- production during photon–nuclei interactions. *© 2001 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

It is well known (see [1]) that the QED scattering amplitude for the process

$$a + b \longrightarrow a' + b'$$

in the Regge kinematics

$$\begin{gathered} A(p_A, a) + B(p_B, b) \longrightarrow A(p'_A, a') + B(p'_B, b'), \\ s = (p_A + p_B)^2 \gg -t = -(p_A - p'_A)^2 \propto m^2 \end{gathered} \tag{1}$$

has the impact factor representation

$$\begin{gathered} A(s, t) = \frac{is}{(2\pi)^2} \\ \times \int \frac{d^2k \tau^A(k, r) \tau^B(k, r)}{[(k + r)^2 + \lambda^2][(k - r)^2 + \lambda^2]} \left(1 + O\left(\frac{t}{s}\right)\right), \\ 4r^2 = -t > 0, \end{gathered} \tag{2}$$

that is valid in the first non-trivial order of perturbation theory. Here, λ is the photon mass; the two-dimensional vectors r and k are orthogonal to the initial particle momenta p_A and p_B. The impact factors τ describe the inner structure of colliding particles. For the electron, we have

$$|\tau^e| = 4\pi\alpha\delta_{ij},$$

where the indices i, j enumerate the electron polarization states. The expression for the impact factors of the photon on its mass shell can be written as [1]

$$\tau_{ij}^{\gamma} = 8\alpha^2 \int_0^1 dy \int_0^1 dx_+ dx_- \delta(x_+ + x_- - 1)(A_{ij} - B_{ij}), \tag{3}$$

with

$$\begin{gathered} A_{ij} = \frac{1}{4r^2 x_+^2 y(1 - y) + m^2} \Big[8x_+^3 x_- y(1 - y) r_i r_j \\ - x_+^2 r^2 \left(1 - 8x_+ x_- \left(y - \frac{1}{2}\right)^2\right) \delta_{ij} \Big], \end{gathered}$$

$$\begin{gathered} B_{ij} = \frac{1}{4Q^2 y(1 - y) + m^2} \Big[8x_+^3 x_- y(1 - y) Q_i Q_j \\ - Q^2 \left(1 - 8x_+ x_- \left(y - \frac{1}{2}\right)^2\right) \delta_{ij} \Big], \end{gathered}$$

$$Q = \frac{1}{2}(k + r) - x_+ r,$$

where i, j refer to the photon polarization states.

According to the Regge theory, the impact factor is proportional to the residue of the pole $\propto (j - 1)^{-1}$ of the t-channel partial wave f_j^+, with the positive signature describing the t-channel transition of two particles into a nonsense state of two virtual photons [1],

$$\tau^A = \lim_{j \to 1} (j - 1) \int_{s_{th}}^{\infty} \frac{ds'}{\pi} \mathrm{Im} A_{\mu_1\mu_2}^{\gamma^*}(q, k; s') \frac{p_B^{\mu_1}}{s} \frac{p_B^{\mu_2}}{s} Q_{j-2}(z'), \tag{4}$$

$$s' = -2kp_A,$$

[¶] This article was submitted by the authors in English.

where $z' = \cos\theta$ is the cosine of the scattering angle θ in the t-channel. This is a linear function of s'. Here and below s_{th} means the threshold value of s'. Higher orders of perturbation theory involve the poles $f_j \propto 1/(j-1)^n$ that must be subtracted from τ to provide sums of all the logarithmic contributions $\propto \log^n(s)$ using the Bethe–Salpeter equation [1].

For $t = 0$, the impact factor is proportional to the s' integral of the total cross section for the scattering of the photon with virtuality $-k^2$ off target a with mass m,

$$\tau = k^2 \int_{s_{th}}^{\infty} \frac{ds'}{\pi} \frac{\sigma_{a\gamma^*}(s', k^2)}{s'} f(s', k^2),$$
$$f(s', k^2) = \frac{\sqrt{s'^2 - 4k^2m^2}}{s'}, \tag{5}$$

where $f(s', k^2)$ accounts for the virtual photon flux factor. This multiplier equals unity in the limit as $k^2 \longrightarrow 0$, which corresponds to the Weizsäcker–Williams approximation:

$$\lim_{k^2 \to 0} \frac{\tau}{k^2} = \int_{s_{th}}^{\infty} \frac{ds'}{\pi} \frac{\sigma_{a\gamma}(s')}{s'}. \tag{6}$$

The motivation for our calculation of the radiative corrections to impact factors is the high-precision experiments performed on colliders where some interesting physical quantities (for example, the BFKL pomeron intercept) are measured [2]. In this case, one must know the impact factors of the virtual photon [2]. Generally, impact factors describe the coupling of particles with the pomeron in QED or in QCD. For colliders with electron (positron) beams, radiative corrections to impact factors can be used to calculate the QED part of cross sections with a good accuracy.

For small-angle e^+e^- scattering, the amplitude for the diagrams with the multi-photon exchange has the eikonal representation

$$\begin{aligned} A(s, t) &= A_0(s, t)e^{i\delta(t)}, \\ A_0(s, t) &= 4\pi\alpha\frac{2}{st}\bar{u}(p_1')\hat{p}_2 u(p_1) \\ \times \bar{v}(p_2)\hat{p}_1 v(p_2') &= 4\pi\alpha\frac{2s}{t}N_1N_2, \\ |N_i| = 1, \quad \delta(t) &= -i\alpha\ln\left(\frac{-t}{\lambda^2}\right), \end{aligned} \tag{7}$$

where we used the fact that only the longitudinal (nonsense) polarizations of the t-channel virtual photons are essential at high energies,

$$\bar{u}(p_1')\gamma_\mu u(p_1)\bar{v}(p_2)\gamma_\nu v(p_2')G^{\mu\nu}(q),$$
$$G^{\mu\nu}(q) = \frac{1}{q^2}\frac{2p_2^\mu p_1^\nu}{s}. \tag{8}$$

The radiative corrections to A_0 appear from the so-called "decorated boxes". These Feynman diagrams were assumed to lead to a generalized eikonal representation

$$A = A_0(s, t)[\Gamma_1(t)]^2 e^{i\delta(t)}, \tag{9}$$

where $\Gamma_1(t)$ is the Dirac form-factor of an electron,

$$V^\mu(t) = \gamma^\mu\Gamma_1(t) + \frac{\sigma^{\mu\nu}q_\nu}{2m}\Gamma_2(t), \quad q^2 = t,$$
$$\Gamma_1(t) = 1 + \gamma\Gamma_1^{(2)}(t) + \ldots, \quad \gamma = \frac{\alpha}{\pi}. \tag{10}$$

We note that $\delta(t)$ must also include corrections to the virtual photon Green function, leading in particular to the electric charge renormalization.

In the next section, we verify the generalized eikonal representation for the decorated boxes.

2. ONE-LOOP CORRECTION TO THE ELECTRON IMPACT FACTOR

Keeping in mind that the amplitude for near-forward scattering with two-photon exchange is purely imaginary (with corrections of the order m^2/s omitted), we can calculate its s-channel discontinuity. The radiative corrections to this discontinuity originate from the virtual photons and from the emission of a real photon in the intermediate state. We split the last contribution into two parts corresponding to the emission of soft and hard photons.

The virtual photon contribution contains the electron vertex function for the on-shell initial and final electrons,

$$\Delta\tau_e^{virt} = \frac{\alpha}{\pi}\tau_e(0)[F_1^{(2)}(-k^2) + F_1^{(2)}(-k'^2)],$$
$$F_1^{(2)}(t) = -G(t)\ln\frac{m}{\lambda} - G_1(t) - T(t),$$
$$G(t) = \frac{1+a^2}{2a}\ln b - 1, \quad G_1(t) = 1 - \frac{1+2a^2}{4a}\ln b, \tag{11}$$
$$T(t) = \frac{1+a^2}{2a}$$

$$\times\left[-\frac{1}{4}\ln^2 b + \ln b \ln(1+b) - \int_1^b \frac{dx}{x}\ln(1+x)\right],$$

$$a = \sqrt{1-\frac{4m^2}{t}}, \quad b = \frac{a+1}{a-1}, \quad t < 0.$$

The contribution from the emission of a soft photon has the classical form

$$-\frac{\alpha}{4\pi^2}\left(\frac{p_1}{p_1k_1} - \frac{p}{pk_1}\right)\left(\frac{p_1}{p_1k_1} - \frac{p'_1}{p'_1k_1}\right)\tau_e^{(0)}\frac{d^3k_1}{\omega_1}\bigg|_{\omega_1 < \delta E}, \qquad (12)$$

$$\delta E \ll E = \frac{\sqrt{s}}{2},$$

where p and p'_1 are the momenta of the initial and final electrons and p_1 is the electron momentum in the intermediate state. Because the energies of these particles are approximately equal (but large compared to the electron mass), we can use the relations

$$\frac{1}{2\pi}\int\frac{d^3k_1}{\omega_1}\frac{m^2}{(p_ik_1)^2} = 2L_e,$$

$$\frac{1}{2\pi}\int\frac{d^3k_1}{\omega_1}\frac{p_ip_j}{(p_ik_1)(p_jk_1)} = \frac{1+a^2}{a} \qquad (13)$$

$$\times\left[L_e\ln b - \frac{1}{4}\ln^2 b + \ln b\ln(1+b) - \int_1^b\frac{dx}{x}\ln(1+x)\right],$$

$$L_e = \ln\Delta + \ln\frac{m}{\lambda}, \quad t = (p_i - p_j)^2, \quad \Delta = \frac{\delta E}{E} \ll 1,$$

with quantities a and b defined above. Thus, we obtain

$$\Delta\tau_e^{\text{soft}} = \frac{\alpha}{\pi}[(G(-k^2) + G(-k'^2) - G(t))L_2$$
$$+ T(-k^2) + T(-k'^2) - T(t)], \qquad (14)$$

where $T(t)$ was defined above.

Now let us consider the hard photon emission. This contribution to the imaginary part of the electron–electron scattering amplitude can be written as

$$\text{Im}_sA(s,t) = -s\frac{\alpha^3}{2\pi^2}\int\frac{d^2k}{k^2k'^2}N_1N_2\frac{d^2k_1dx}{x(1-x)}I(x,k_1,k), \qquad (15)$$
$$\Delta < x < 1,$$

where x is the energy fraction of the hard photon. We obtain

$$I(x,k_1,k) = \frac{1}{d_1d_2}(-4m^2 + 2t_1z)$$
$$+\frac{1}{d_1d'_1}(-4m^2x^2(1-x) + 2tz(1-x)) \qquad (16)$$
$$+\frac{1}{d_2d'_1}(-4m^2 + 2t_2z) - 2z\frac{1}{d_1} - 4z\frac{1}{d_2} + \frac{8m^2}{d_2^2} - 2z\frac{1}{d'_1},$$

$$z = 1 + (1-x)^2,$$

where

$$d_1 = (p-k_1)^2 - m^2 = -\frac{1}{x}[m^2x^2 + \mathbf{k}_1^2],$$

$$d_2 = (p_1+k_1)^2 - m^2 = \frac{1}{x(1-x)}[m^2x^2 + (x\mathbf{k} - \mathbf{k}_1)^2],$$

$$d'_1 = (p'_1 - k_1)^2 - m^2 = -\frac{1}{x}[m^2x^2 + (x\mathbf{q} - \mathbf{k}_1)^2], \qquad (17)$$

$$2pp_1 = t_1 = \frac{1}{1-x}[m^2z + (\mathbf{k} - \mathbf{k}_1)^2],$$

$$2p_1p'_1 = t_2 = \frac{1}{1-x}[m^2z + (x\mathbf{q} + \mathbf{k}_1 - \mathbf{k})^2].$$

The subsequent integration is straightforward and gives the result

$$\Delta\tau_e^{\text{hard}} = \tau_e^{(0)}\frac{\alpha}{\pi}\Big[(G(-k^2) + G(-k'^2) - G(t))\ln\frac{1}{\Delta}$$
$$+ G_1(-k^2) + G_1(-k'^2) - G_1(t)\Big], \qquad (18)$$

where $G(t)$ and $G_1(t)$ were defined above.

The interference of two amplitudes with a photon emitted by two initial particles is small $\sim O(t/s)$. This fact is known in the literature as the up–down cancellation. The contribution of the diagrams with two-photon exchange is purely imaginary and, consequently, does not interfere with the real Born amplitude. Adding all the contributions, we obtain the final result for one-loop radiative corrections to the electron impact factor,

$$\Delta\tau_e = \frac{\alpha}{\pi}\tau_e^{(0)}F_1^{(2)}(t), \quad \tau_e^{(0)} = 4\pi\alpha. \qquad (19)$$

This result agrees with the generalized eikonal form of the small-angle scattering amplitude. However, in the higher orders, the eikonal representation is violated as shown below.

3. GENERALIZED EIKONAL REPRESENTATION

The above result for the radiative corrections to the electron impact factor can be obtained in a simple way. We consider again a decorated box with a positron block corresponding to the Born diagram and an electron block containing a set of four Feynman graphs with a virtual photon. We express the components of the exchanged photon momentum in terms of the squared invariant energies s_1 and s_2 for the electron and positron blocks using the Sudakov parameters

$$k = \alpha p_2 + \beta p_1 + k_\perp,$$

$$d^4k = \frac{1}{2s} ds_1 ds_2 d^2k_\perp, \quad k_\perp^2 = -\mathbf{k}^2,$$

$$s_1 = (k - p_1)^2 = -s\alpha - \mathbf{k}^2 + m^2,$$

$$s_2 = (k + p_2)^2 = s\beta - \mathbf{k}^2 + m^2.$$

Performing the s_2-integration by taking the residue of the intermediate positron propagator (which also takes the diagram with crossed photon lines into account), we obtain the total radiative corrections

$$\frac{4\alpha}{s(2\pi)^2}\int \frac{d^2\mathbf{k}}{(\mathbf{k}^2 + \lambda^2)((\mathbf{q} - \mathbf{k})^2 + \lambda^2)} \times \int_C ds_1 p_2^\mu p_2^\nu \bar{u}(p_1') A_{\mu\nu} u(p_1), \tag{20}$$

where $\bar{u}(p_1')A_{\mu\nu}u(p_1)$ is the Compton scattering amplitude corresponding to the Feynman diagrams with only the *s*-channel singularities and *C* is the contour situated above these singularities. The amplitude has a pole at $s_1 = m^2$, which corresponds to the electron intermediate state, and a right-hand cut starting from $s_1 = (m + \lambda)^2$, which corresponds to the one-electron and one-photon intermediate state.

Using the Sudakov parametrization for the photon momentum *k* and omitting the small contribution $\propto 1/s$ proportional to βp_1, we can represent p_2^μ as

$$p_2^\mu = \frac{1}{\alpha}(k - k_\perp - \beta p_1)^\mu \approx -\frac{s}{s_1 + \mathbf{k}^2}(k - k_\perp)^\mu. \tag{21}$$

We now consider the product of two terms in the right-hand side of this equation with the Compton amplitude $A_{\mu\nu}$. The contribution of the term $\propto k_\perp$ is zero,

$$|\mathbf{k}| s p_2^\nu \int_C \frac{ds_1 k_\perp^\mu}{(s_1 + \mathbf{k}^2)|\mathbf{k}|} \times \bar{u}(p_1') A_{\mu\nu}(s_1, k, k') u(p_1) = 0. \tag{22}$$

This follows from the convergence of the integral over the large circle in the s_1 plane and the absence of the left cut. The second property is valid for planar Feynman graphs. The integral converges because for the physical (transverse) polarizations of the virtual photon, the quantity $e^\mu p_2^\nu A_{\mu\nu}$, $\mathbf{e} = \mathbf{k}_\perp/|\mathbf{k}|$ behaves as m^2/s_1 at large s_1.

Applying the Ward identity for the first contribution $\propto k^\mu$, we obtain

$$p_2^\mu p_2^\nu \bar{u}(p_1') A_{\mu\nu}(s_1) u(p_1) = -\frac{se^2}{s_1} p_2^\nu \bar{u}(p_1') \Gamma^\nu(q) u(p_1), \quad s_1 \gg m^2. \tag{23}$$

The integral over the large semicircle gives the generalized eikonal result $\propto \Gamma_\nu$, which means, in particular, that the total contribution of the various intermediate states is not zero for physical $t < 0$. In particular, we see that radiative corrections to the impact factor of the electron contain infrared divergences cancelled only in the total cross section with the contribution of the inelastic process (the photon emission).

For the *n*-photon exchange, the eikonal result for the scattering amplitude corresponds to the classical picture where all the intermediate fermions are on their mass shell. This is so because the Born amplitude for the *t*-channel photon interactions with external particles tends to zero as $(p_A k_i)^{-2}$ for $(p_A k_i) \longrightarrow \infty$, which allows us to calculate all the integrals over $p_A k_i$ by taking residues. For the radiative corrections corresponding to the decorated diagrams with one additional virtual photon, we can use the arguments similar to those applicable in the two-photon case. The physical reason for the generalized eikonal result for the total contribution is that the integration over invariant s_i (corresponding to the virtuality of the inner fermion line to which the virtual gluon line is attached) gives zero because after the cancellation of the renormalization effects in accordance with the Ward identity, the amplitude behaves as $1/s_i^2$ at large s_i. The nonvanishing result is obtained only from the diagrams where the virtual gluon line is attached to the external fermion lines, but we then obtain the generalized eikonal result. This argument is not valid for nonplanar diagrams because they have left and right singularities in the s_i planes [6].

4. IMPACT FACTORS IN THE TWO-LOOP APPROXIMATION

In the radiative corrections to the photon impact factor, the infrared divergences are cancelled in the sum of contributions from the $e^+e^-\gamma$ and e^+e^- intermediate states. Using the crossing relations for $t = 0$ [7], one can express the contribution of the $e^+e^-\gamma$ intermediate state to τ^γ in terms of the contribution of the $e\gamma\gamma$ intermediate state to τ^e, which is investigated better (see [4–6]). We here estimate the radiative corrections for $t = 0$ only at small virtualities of the exchanged photon $\mathbf{k}^2$. Their value can be extracted from the results of [3], where the one-loop correction to the cross section of pair production by a photon on the Coulomb field of nuclei was calculated as

$$\sum_{i=1}^{i=2}[\tau+\Delta\tau]^{\gamma}_{ii}(k,0) = \frac{28\mathbf{k}^2\alpha^2}{9m^2}[1+\delta_p], \quad \mathbf{k}^2 \ll m^2, \tag{24}$$
$$\delta_p = \frac{\alpha}{\pi}\frac{9}{14}\left(\frac{1128}{35}\zeta(3)-\frac{6971}{210}\right) = 0.009.$$

The radiative corrections to the photon impact factor can also easily be found in the region $k^2 \gg m^2$, where one can use the DGLAP evolution equations [10].

Now let us consider the radiative corrections to the electron impact factor. The generalized eikonal hypothesis is violated in the two-loop approximation. (This fact was verified explicitly for $t = 0$ [6].) Indeed, if the generalized eikonal hypothesis were valid, the complete compensation of contributions from the transition of the initial electron to the intermediate states e, $e\gamma$, and $e\gamma\gamma$ would occur. However, it was shown that the total contribution is not zero and is equal to the interference term for the e^+e^- pair production amplitudes.

To clarify this result, we write the impact factor as

$$\tau^A = \int_C \frac{ds_1}{2\pi i}\frac{1}{s^2}J^{(A)}_{\mu\nu}p^{\mu}_B p^{\nu}_B, \tag{25}$$

where the quantity $(1/s^2)J^{(A)}_{\mu\nu}p^{\mu}_B p^{\nu}_B$ is expressed in terms of the amplitudes $J^{(A)}$ for the scattering of the virtual photon from the initial particles and does not depend on s as $s \longrightarrow \infty$.

In contrast to the planar amplitude $A_{\mu\nu}$ discussed in the previous section, $J^e_{\mu\nu}$ corresponds to contributions of all possible diagrams. The integration contour C is displaced in accordance with the Feynman prescription between the right- and left-hand side singularities of the amplitude. The right singularities are the poles at $s_1 = m^2$ and the cuts at

$$s_1 > (m+\lambda)^2, \quad s_1 > (m+2\lambda)^2, \quad s_1 > 9m^2.$$

There also exist left singularities at the same points for the crossing variable

$$u_1 = -s_1 - t - 2m^2 + \mathbf{k}^2 + (\mathbf{q}-\mathbf{k})^2.$$

The additional e^+e^- pair can be produced in accordance with the Bethe–Heitler or bremsstrahlung mechanisms. There also exist interference terms taking the identity of the final electrons into account. The most important contribution is from the Bethe–Heitler mechanism corresponding to the e^+e^- pair production by two virtual photons. The corresponding impact factor contains the divergence in s_1 related to the presence of two-photon intermediate states in the crossing channel. (For $t = 0$, this contribution was calculated in [11].) We write it here only in the Weizsäcker–Williams approximation, where it has the form of the sum rule for the Borselino formulas for the total cross section $\sigma(s_1)$ of the e^+e^- pair production in the electron–photon collisions through the Bethe–Heitler mechanism,

$$\tau^e_{BH_e} = k^2\int_{s_{th}}^{s}\frac{ds_1}{\pi}\frac{\sigma(s_1)}{s_1} = \frac{\alpha^3k^2}{\pi m^2}\left(a\ln^2\frac{s}{m^2} + b\ln\frac{s}{m^2} + c\right), \tag{26}$$

$$a = \frac{14}{9}, \quad b = -\frac{218}{27}, \quad c = \frac{418}{27} - \frac{13}{2}\zeta(2). \tag{27}$$

As discussed above, the logarithmic dependence on the upper limit s in the integral over s_1 must be subtracted in a self-consistent way to avoid double counting, because the logarithmic contributions are summed by the Bethe–Salpeter equation for the pomeron in QED (cf. a similar procedure for the BFKL pomeron in the next-to-leading approximation [12]). For muon production, we have

$$\tau^e_{BH_\mu} = \frac{\alpha^3k^2}{\pi M^2}\left(a\ln^2\frac{s}{M^2} + b\ln\frac{s}{M^2} + c\right), \tag{28}$$

$$a = \frac{14}{9}, \quad b = -\frac{218}{27} + \frac{28}{9}\ln\frac{M}{m},$$
$$c = \frac{3011}{324} - \frac{28}{9}\zeta(2) - \frac{107}{9}\ln\frac{M}{m}, \tag{29}$$

where m and M are the respective masses of the electron and muon.

The contribution of the bremsstrahlung mechanism to e^+e^- pair production must be added with the corresponding two-loop radiative corrections to the electron form-factor for the elastic intermediate state; the resulting expression corresponds to the generalized eikonal approximation because the corresponding diagrams are planar [6].

Among many Feynman graphs obtained from the interference between the various amplitudes for pair production, there are only four nonplanar diagrams corresponding to the identity of electrons in the final state in the Bethe–Heitler mechanism. Only these graphs give a nonvanishing result for τ^e at $t = 0$. In the

Weizsäcker–Williams approximation, the corresponding contribution was calculated in [7],

$$\tau_{int}^{(e)} \approx \frac{k^2}{m^2}\frac{\alpha^3}{\pi}\left(\frac{221}{315} + \frac{41549}{6300}\zeta(2) - \frac{216}{105}\zeta(3) - \frac{792}{105}\xi(2)\ln 2\right) \approx \frac{k^2}{m^2}\frac{\alpha^3}{\pi}(-3.57). \quad (30)$$

This leads to the sum rules for the integrals of the one- and two-photon bremsstrahlung cross sections and the slope of the Dirac form-factor at $t = 0$ [6].

Finally, the total two-loop contribution to the electron impact factor can be written as

$$\tau^e = \frac{\alpha^2}{\pi^2}\tau_e^0 F_1^{(4)} + \tau_{BH}^e, \quad (31)$$

where $F_1^{(4)}$ is the full two-loop correction to the Dirac form-factor (including the nonplanar diagrams and the diagrams with the inner fermion loop). The term τ_{BH}^e is the total contribution of the imaginary part corresponding to the Bethe–Heitler mechanism of the pair production including the interference effects related to the identity of the produced electrons ($\tau_{BH}^e = \tau_{BH_e}^e + \tau_{BH_\mu}^e + \tau_{int}^{(e)}$ for $t = k^2 = 0$).

The physical meaning of this formula is obvious: the non-trivial corrections to impact factors are related only to the charge particle production in the intermediate states.

5. CONCLUSION

In the three-loop approximation, the most important contribution to the photon impact factor corresponds to the diagram with two fermionic loops connected in the t-channel by two photons. It contains the logarithmic divergence $\propto \ln s$ because the imaginary part of the corresponding amplitude is proportional to s_1 for large s_1. In particular, for $t = \mathbf{k}^2 = 0$, the impact factor can be expressed as the integral of the cross section for the transition of two real photons into two e^+e^- pairs. Again, the ultraviolet divergence in s_1 is compensated by the infrared divergence in the relative rapidities of the produced pairs in the Bethe–Salpeter equation for the pomeron in QED. The virtual photon actually interacts with the electric dipoles inside the initial photon [13]. The growth of the impact factor $\propto \ln s$ is related to the logarithmic increase of the number of dipoles at large energies. The fermion identity effects in the intermediate state do not have any influence on this growth. The contribution of the diagrams with one e^+e^- pair and several photons gives a finite contribution to the photon impact factor.

We now consider three-loop corrections to the electron impact factor. The most important contribution $\propto \ln^2 s$ comes from the one-loop radiative corrections to the Bethe–Heitler mechanism of e^+e^- production. Other diagrams lead to finite terms. The generalized eikonal representation is violated by nonplanar diagrams related to e^+e^- pair production, but there is another reason for its violation. It is related to the charge parity conservation in QED. Indeed, two external photons with momenta k and $q - k$ cannot pass through the fermion loop to the three-photon intermediate state in the t-channel. Therefore, the generalized eikonal representation, containing in particular the form-factor corresponding to the transition of the external photon through the fermion loop into the three-photon state, cannot be valid in the three-loop approximation.

The methods developed above for QED can also be used in QCD, where we urgently need to calculate the radiative corrections to impact factors of the virtual photon and other particles to find the energy region of applicability of the BFKL theory in the next-to-leading approximation [11].

The work of E.A.K. was partially supported by RFBR (grant no. 99-02-17730) and HLP (grant no. 2000-02). The work of L.N.L. was supported by INTAS and CRDF (grants no. 97-31696 and RP1-2108). We thank A.E. Dorokhov for his help. We are also grateful to M.V. Galynski for collaboration in the initial stage of the work.

REFERENCES

1. H. Cheng and T. T. Wu, Phys. Rev. Lett. **22**, 666 (1969); Phys. Rev. D **1**, 2775 (1970); G. V. Frolov, V. N. Gribov, and L. N. Lipatov, Phys. Lett. B **31B**, 34 (1970); Yad. Fiz. **12**, 994 (1970) [Sov. J. Nucl. Phys. **12**, 543 (1971)]; L. N. Lipatov and G. V. Frolov, Yad. Fiz. **13**, 588 (1971) [Sov. J. Nucl. Phys. **13**, 333 (1971)].
2. L. N. Lipatov, Yad. Fiz. **23**, 642 (1976) [Sov. J. Nucl. Phys. **23**, 338 (1976)]; V. S. Fadin, E. A. Kuraev, and L. N. Lipatov, Phys. Lett. B **60B**, 50 (1975); Ya. Ya. Balitsky and L. N. Lipatov, Yad. Fiz. **28**, 1597 (1978) [Sov. J. Nucl. Phys. **28**, 822 (1978)].
3. E. Vinokurov, E. Kuraev, and N. Merenkov, Zh. Éksp. Teor. Fiz. **66**, 1916 (1974) [Sov. Phys. JETP **39**, 942 (1974)].
4. E. Kuraev, N. Merenkov, and V. Fadin, Yad. Fiz. **45**, 782 (1987) [Sov. J. Nucl. Phys. **45**, 486 (1987)].
5. E. Kuraev, L. Lipatov, N. Merenkov, and V. Fadin, Zh. Éksp. Teor. Fiz. **65**, 2155 (1973) [Sov. Phys. JETP **38**, 1076 (1973)].
6. E. Kuraev, L. Lipatov, and N. Merenkov, Phys. Lett. B **47B**, 33 (1973); S. J. Chang, Phys. Rev. D **1**, 2997 (1970); S. J. Yao, Phys. Rev. D **1**, 2971 (1970).
7. E. A. Kuraev, L. N. Lipatov, and M. I. Strikman, Yad. Fiz. **18**, 1270 (1973) [Sov. J. Nucl. Phys. **18**, 652 (1973)]; Zh.

Éksp. Teor. Fiz. **66**, 838 (1974) [Sov. Phys. JETP **39**, 405 (1974)].

8. A. Arbuzov *et al.*, Preprint CERN-TH/95-313; V. Fadin *et al.*, Preprint JINR E2-92-577 (Joint Institute for Nuclear Research, Dubna, 1992).
9. V. N. Baier, V. S. Fadin, V. A. Khoze, and E. A. Kuraev, Phys. Rep. **78**, 283 (1981).
10. V. N. Gribov and L. N. Lipatov, Yad. Fiz. **15**, 781 (1972) [Sov. J. Nucl. Phys. **15**, 438 (1972)]; Yad. Fiz. **15**, 1218 (1972) [Sov. J. Nucl. Phys. **15**, 675 (1972)]; L. N. Lipatov, Yad. Fiz. **20**, 181 (1974) [Sov. J. Nucl. Phys. **20**, 94 (1975)]; G. Altarelli and G. Parisi, Nucl. Phys. B **26**, 298 (1977); Yu. L. Dokshitzer, Zh. Éksp. Teor. Fiz. **73**, 1216 (1977) [Sov. Phys. JETP **46**, 641 (1977)].
11. E. A. Kuraev and L. N. Lipatov, Pis'ma Zh. Éksp. Teor. Fiz. **15**, 229 (1972) [JETP Lett. **15**, 159 (1972)]; Yad. Fiz. **16**, 1060 (1972) [Sov. J. Nucl. Phys. **16**, 584 (1973)].
12. V. S. Fadin and L. N. Lipatov, Phys. Lett. B **429**, 127 (1998).
13. A. H. Mueller, Nucl. Phys. B **415**, 373 (1994); Nucl. Phys. B **437**, 107 (1995); N. Nikolaev, B. Zakharov, and V. Zoller, Phys. Lett. B **328**, 486 (1994).

Journal of Experimental and Theoretical Physics, Vol. 92, No. 2, 2001, pp. 210–221.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 119, No. 2, 2001, pp. 243–256.
Original Russian Text Copyright © 2001 by Panteleev, Roerieh.

ATOMS, SPECTRA, RADIATION

The Influence of Quantum Interference Effects on the Resonance Fluorescence Spectra of a Degenerate Three-Level Atom

A. A. Panteleev and Vl. K. Roerieh*

State Research Center, Troitsk Institute for Innovation and Thermonuclear Research, Troitsk, Moscow oblast, 142092 Russia
**e-mail: vroerieh@fly.triniti.troitsk.ru*

Received August 18, 2000

Abstract—The resonance fluorescence spectra of a degenerate three-level atom of the *V*-type in the field of an intense monochromatic wave with an arbitrary polarization composition are investigated. Analytical expressions are derived for the resonance fluorescence spectra, and the angular distribution of spontaneous fluorescence of atoms is analyzed for the *D*-line emitted by vapors of alkali atoms. It is shown that the number of lines in the spectrum may decrease in the case of the linear polarization of spontaneous radiation. The radiation relaxation operator is obtained for the *D*-line of alkali metals in the case when an atom is near the metal surface. Interference effects for such systems are analyzed. *© 2001 MAIK “Nauka/Interperiodica”.*

1. INTRODUCTION

Quantum interference processes accompanying the spontaneous emission of atoms in two and more closely spaced states has attracted attention from researchers during recent years [1–10]. This is associated with numerous and diversified effects (such as the trapping of population [7], the decrease in the total intensity of spontaneous radiation [11], the narrowing of individual lines in spectra [1–5], occupancy beats [6, 12], and the dependence of the shape of the spectrum on the phase of exciting fields [2, 3, 10]) resulting from interference processes in such systems.

For the emergence of an interference pattern, at least two correlated paths are required. The interference processes in this case may affect the structure of the radiation relaxation operator only if the system arrives at the same final state. The simplest system of this type is a nondegenerate three-level *V*-type atom for which all the above-mentioned effects have already been manifested. It should be noted, however, that such a system is of the model kind and does not include many characteristics of real physical systems which are important for an analysis of interference processes. A disadvantage of such models is the disregard of the possible degeneracy of energy levels as well as the angular and polarization composition of the exciting and spontaneous radiations. Moreover, the dipole moment of transitions is assumed to be a scalar quantity. In our earlier publication [11], we proved that the inclusion of these factors renders the radiation relaxation operator for most physical systems in the form identical to the standard diagonal expression. This is due to the fact that the interference terms vanish either as a result of integration over the solid angle, or upon summation over polarizations. The interference terms appear only when the following additional condition is satisfied: the density of states of the electromagnetic field is anisotropic [11, 13]. The lifetime of the excited state for such systems has already been analyzed. Chew [14] studied the radiation emitted by a vibrating dipole in a small dielectric sphere. The effect of the variation of relaxation constants in a cavity is also well known. However, the influence of an anisotropic density of states on the interference process accompanying spontaneous emission has become an object of investigation only recently [13].

Another aspect which has been disregarded till now is the influence of interference effects on the angular distribution of spontaneous fluorescence of atoms. In [15], we analyzed the angular distribution of the resonance fluorescence intensity for degenerate three-level systems of the *V*-type in the field of an electromagnetic wave with an arbitrary polarization composition. We proved that the contribution of the interference processes must be taken into account in an analysis of the angular distribution of spontaneous radiation in spite of the fact that such processes do not affect the dynamics of occupancies and the total intensity of resonance fluorescence for systems under standard conditions.

This work is devoted to an analysis of the resonance fluorescence spectra for degenerate *V*-type three-level systems and their dependence on the direction of observation and the polarization composition of the pumping wave. We will analyze in detail the spectra of the *D*-line emitted by alkali metal vapors taking into account the influence of interference effects on the structure of the radiation relaxation in the case when an atom is near the metal surface.

2. BASIC EQUATIONS

Let us consider the interaction between an ensemble of degenerate three-level atoms of the *V*-type, characterized by the total angular momentum J_j of the levels ($j = b, c, d$), and a monochromatic wave which is in resonance with the transitions $b \longrightarrow d$ and $c \longrightarrow d$ (Fig. 1). Assuming that the electromagnetic wave has an arbitrary polarization composition and using the rotating wave approximation, we present it in the form of a sum:

$$\mathbf{E} = \sum_{\sigma = 0, \pm 1} \mathbf{e}^{\sigma} E_{\sigma}(e^{i\omega t} + e^{-i\omega t}), \qquad (1)$$

$$E_{\pm 1} = \mp\frac{1}{\sqrt{2}}(E_x \pm iE_y), \quad E_0 = E_z.$$

Here, $\mathbf{E}$ is the electromagnetic field vector, $\mathbf{e}^{\sigma}$ is the unit vector of σ-polarization, E_x, E_y, and E_z are the amplitudes of field oscillations along the corresponding axes, and $E_{\pm 1}$ and E_0 are the corresponding amplitudes of the polarized radiation. Using the dipole and resonance approximations, we present the Hamiltonian (in rad/s) of the system in the form of a sum:

$$\hat{H} = \hat{H}_0 + \hat{H}_f + \hat{V}_L + \hat{V}_q. \qquad (2)$$

Here, $\hat{H}_0$ describes the unperturbed atomic system:

$$\hat{H}_0 = \sum_{j = b, c} \sum_{M_j} \frac{\epsilon(J_j, M_j)}{\hbar} |jM_j\rangle\langle M_j j|, \qquad (3)$$

where $\epsilon(J_j, M)$ is the energy of the atomic state with the total angular momentum J_j and its component M_j along the quantization axis z, and $\langle M_j j|$ and $|jM_j\rangle$ are the bra and ket vectors of this state, respectively. Since we will henceforth disregard the effect of magnetic sublevel splitting as well as other effects leading to the dependence of the level energy on the magnetic moment component or other quantum numbers, the Hamiltonian of a free atom can be written in the form

$$\hat{H}_0 = \sum_{j = b, c} \sum_{M_j} \omega_{jd} |jM_j\rangle\langle M_j j|, \qquad (4)$$

where ω_{jd} is the frequency of the $j \longrightarrow d$ transition for any pair of sublevels.

The second term in (2) describes the Hamiltonian of the quantum field of radiation:

$$\hat{H}_f = \sum_{\mathbf{k}\sigma} \omega_k \hat{a}^{+}_{\mathbf{k}\sigma} \hat{a}_{\mathbf{k}\sigma}, \qquad (5)$$

where ω_k is the frequency of photons with the wave vector $\mathbf{k}$ and with σ polarization, and $\hat{a}^{+}_{\mathbf{k}\sigma}$ and $\hat{a}_{\mathbf{k}\sigma}$ are the creation and annihilation operators for the corresponding photons.

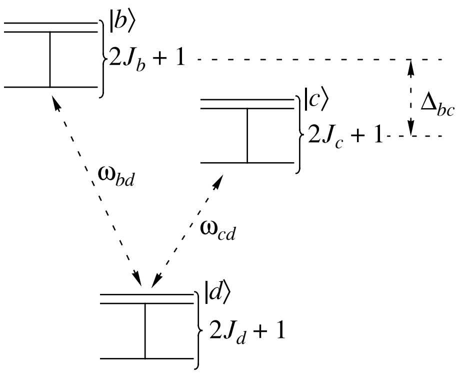


Fig. 1. Energy level diagram for a degenerate three-level atom of the *V*-type.

The interaction between a pumping wave and the atomic system can be described by the expression

$$\hat{V}_L = \sum_{j = b, c} \sum_{M_j M_d} [V_{jd}(M_j M_d) \exp(-i\omega_L t) |jM_j\rangle\langle M_d d| + V_{dj}(M_d M_j) \exp(i\omega_L t) |dM_d\rangle\langle M_j j|], \qquad (6)$$

$$V_{jd}(M_j M_d) = V^{*}_{dj}(M_d M_j) = \sum_{\sigma = 0, \pm 1} -\frac{\mu_{jd}(M_j M_d) E_{\sigma}}{2\hbar}, \qquad (7)$$

where $\mu_{jd}(M_j M_d)$ is the matrix element of the dipole moment. Using the Wigner–Eckart theorem [16], we can write expression (7) in the form

$$V_{jd}(M_j M_d) = \sum_{\sigma = 0, \pm 1} (-1)^{J_j - M_j} C^{1-\sigma}_{J_j M_j J_d - M_d} V_{jd}(\sigma), \qquad (8)$$

$$V_{jd}(\sigma) = -\frac{(-1)^{\sigma} \|\mu_{jd}\| E_{\sigma}}{2\sqrt{3}\hbar},$$

where $C^{1-\sigma}_{J_j M_j J_d - M_d}$ are the Clebsch–Gordan coefficients [16], and $\|\mu_{jd}\|$ is the irreducible value of the matrix element of the dipole moment.

The interaction of atoms with the quantum field of radiation is described by the expression

$$\hat{V}_q = \sum_{j = b, c} \sum_{\mathbf{k}\sigma} \sum_{M_j M_d} (g^{\mathbf{k}\sigma}_{jd}(M_j M_d) \hat{a}_{\mathbf{k}\sigma} |jM_j\rangle\langle M_d d| - g^{\mathbf{k}\sigma}_{dj}(M_d M_j) \hat{a}^{+}_{\mathbf{k}\sigma} |dM_d\rangle\langle M_j j|), \qquad (9)$$

where

$$g_{jd}^{\mathbf{k}\sigma}(M_jM_d) = \mu_{jd}(M_jM_d)\sqrt{\frac{2\pi\omega_k}{\hbar W}},$$

W being the quantization volume. Using the Wigner–Eckart theorem, we finally obtain

$$\begin{aligned} g_{dj}^{\mathbf{k}\sigma}(M_dM_j) &= g_{jd}^{\mathbf{k}\sigma}(M_jM_d) \\ &= (-1)^{\sigma}C_{J_jM_jJ_d-M_d}^{1-\sigma}\|\mu_{jd}\|\sqrt{\frac{2\pi\omega_k}{\hbar W}}. \end{aligned} \tag{10}$$

In order to describe the system under consideration, we will use the atomic density matrix formalism [17]. The equation for the atomic density matrix assumes the form

$$i\rho = [\hat{H}_0 + \hat{V}_L, \rho] + i\hat{\Gamma}^{at}(\rho), \tag{11}$$

$\hat{\Gamma}^{at}$ describing the radiation relaxation processes. It was proved by us earlier [15] that disregarding the doublet splitting, we can present operator $\hat{\Gamma}^{at}$ in the form

$$\hat{\Gamma}^{at}(\rho) = \sum_{\mathbf{k}}(2\hat{G}^k\rho\hat{G}^{k*} - \hat{G}^{k*}\hat{G}^k\rho - \rho\hat{G}^{k*}\hat{G}^k), \tag{12}$$

where the nonzero elements of operator $\hat{G}^k$ have the form

$$\hat{G}_{dj}^k(M_d, M_j) = g_{dj}^{\mathbf{k}\sigma}(M_dM_j), \quad j = b, c.$$

All the remaining elements of the operator are equal to zero. The asterisk in Eq. (12) indicates Hermitian conjugation. The last two terms in this equation describe the radiation relaxation of the density matrix and transitions to lower-lying levels, while the first term corresponds to the arrivals to the lower energy levels. Applying the Wigner–Weisskopf procedure to the sums in Eq. (12), we obtain the following expression for the constants of the radiation relaxation of magnetic sublevels, including the nondiagonal elements:

$$\begin{aligned} \gamma_{j_1j_2}(M_{j_1}M_{j_2}) &= \int\sum_{M_d\sigma\sigma'} g_{j_1d}^{\mathbf{k}\sigma}(M_{j_1}M_d) \\ &\times g_{dj_2}^{\mathbf{k}\sigma'}(M_dM_{j_2})\varsigma(\omega_k - \omega_p, \sigma, \sigma')\frac{d\mathbf{k}}{(2\pi)^3} \\ &= \frac{2\pi\|\mu_{j_1d}\|\|\mu_{j_2d}\|}{\hbar\sqrt{(2J_{j_1}+1)(2J_{j_2}+1)}}\sum_{M_d\sigma,\sigma'} C_{J_dM_d1\sigma}^{J_{j_1}M_{j_1}}C_{J_dM_d1\sigma'}^{J_{j_2}M_{j_2}} \\ &\times\int\sum_{\lambda=\pm1}(D_{\lambda\sigma'}^1)^*D_{\lambda\sigma}^1\omega_k\varsigma(\omega_k - \omega_p)\frac{d\mathbf{k}}{(2\pi)^3}. \end{aligned} \tag{13}$$

Applying a similar procedure, we can obtain the following relation for the arrivals to lower energy levels:

$$\begin{aligned} \gamma_{dd}(M_dM_d) &= \int 2\sum_{M_{j_1}M_{j_2}\sigma\sigma'} g_{dj_2}^{\mathbf{k}\sigma'}(M_dM_{j_2})g_{j_1d}^{\mathbf{k}\sigma}(M_{j_1}M_d) \\ &\times\rho_{j_1j_2}(M_{j_1}M_{j_2})\varsigma(\omega_k - \omega_p, \sigma, \sigma')\frac{d\mathbf{k}}{(2\pi^3)} \\ &= \frac{4\pi\|\mu_{j_1d}\|\|\mu_{j_2d}\|}{\hbar\sqrt{(2J_{j_1}+1)(2J_{j_2}+1)}} \\ &\times\sum_{j_1, j_2, M_{j_1}, M_{j_2}, \sigma, \sigma'} C_{J_dM_d1\sigma}^{J_{j_1}M_{j_1}}C_{J_dM_d1\sigma'}^{J_{j_2}M_{j_2}}\rho_{j_1j_2}(M_{j_1}M_{j_2}) \\ &\times\int\sum_{\lambda=\pm1}(D_{\lambda\sigma'}^1)^*D_{\lambda\sigma}^1\omega_k\varsigma(\omega_k - \omega_p)\frac{d\mathbf{k}}{(2\pi)^3}. \end{aligned} \tag{14}$$

Here, $\varsigma(\omega_k - \omega_p, \sigma, \sigma')$ is the spectral contour of the line, $\omega_p = (\omega_{bd} + \omega_{cd})/2$, and $D_{\lambda\sigma}^1 = D_{\lambda\sigma}^1(\alpha = \phi, \beta = \theta, \gamma = 0)$ is Wigner's function [18], where angles ϕ and θ define the direction of the wave vector $\mathbf{k}$ of an emitted photon. The method enabling us to determine the structure (12) of the radiation relaxation operator and the expression for its elements (13) taking into account the interference processes accompanying spontaneous emission for a degenerate three-level system of the V-type was considered in detail in [15]. Expressions (12)–(14) completely determine the radiation relaxation processes for the model under investigation. It was proved by us earlier [11] that the operator of radiative departures for most real systems has a diagonal structure. This follows from the orthogonality of Wigner's functions [18],

$$\int D_{\lambda_1\sigma_1}^{J_1}D_{\lambda_2\sigma_2}^{J_2}\frac{dO}{4\pi} = \frac{1}{2J+1}\delta_{J_1J_2}\delta_{\sigma_1\sigma_2}, \tag{15}$$

if we take into account the properties of the sums [16]

$$\sum_{M_d} C_{J_dM_d1\sigma}^{J_{j_1}M}C_{J_dM_d1\sigma}^{J_{j_2}M} = \delta_{J_{j_1}, J_{j_2}}. \tag{16}$$

Formulas (15) and (16) immediately show that the values of $\gamma_{j_1j_2}(M_{j_1}M_{j_2})$ differ from zero only for $j_1 = j_2$ and $M_1 = M_2$.

It should be emphasized that the application of equality (15) in expression (13) for the relaxation constants implicitly presumes the isotropy of the density of states of the electromagnetic field. Another situation takes place, for example, in the vicinity of the metal surface. In this case, the property of normalization and orthogonality (15) is not observed in the general case, and the operator $\gamma_{j_1j_2}(M_{j_1}M_{j_2})$ may not have a diagonal structure. Besides, the explicit form of the operator and the values of its elements may depend on the choice of the direction of the *z*-axis relative to the surface. By way of an example, the values of $\gamma_{j_1j_2}(M_{j_1}M_{j_2})$ near the surface of a metal are given in the Appendix for various orientations of the *z*-axis relative to the surface.

The dynamics of the system under investigation in the rotating wave approximation is described by the following equations:

$$\begin{aligned}
i\dot{\rho}_{j_1j_2}(M_{j_1}M_{j_2}) &= \Delta_{j_1j_2}\rho_{j_1j_2}(M_{j_1}M_{j_2}) \\
&+ \sum_{jM_j}(-i)[\gamma_{j_1j}(M_{j_1}M_j)\rho_{jj_2}(M_jM_{j_2}) \\
&+ \rho_{j_1j}(M_{j_1}M_j)\gamma_{jj_2}(M_jM_{j_2})] \\
&+ \sum_{M_d}[\tilde{V}_{j_1d}(M_{j_1}M_d)\tilde{\rho}_{dj_2}(M_dM_{j_2}) \\
&- \tilde{\rho}_{j_1d}(M_{j_1}M_d)\tilde{V}_{dj_2}(M_dM_{j_2})], \\
i\dot{\tilde{\rho}}_{j_1d}(M_{j_1}M_d) &= \Delta_L^{j_1}\tilde{\rho}_{j_1d}(M_{j_1}M_d) \\
&- i\sum_{jM_j}\gamma_{j_1j}(M_{j_1}M_j)\tilde{\rho}_{jd}(M_jM_d) \\
&+ \sum_{M'_d}\tilde{V}_{j_1d}(M_{j_1}M'_d)\tilde{\rho}_{dd}(M'_dM_d) \\
&- \sum_{j=b,c}\sum_{M_j}\rho_{j_1j}(M_{j_1}M_j)\tilde{V}_{jd}(M_jM_d), \\
i\dot{\rho}_{dd}(M_d^1M_d^2) &= i\gamma_{dd}(M_d^1M_d^2) \\
&+ \sum_{j=b,c}\sum_{M_d}[\tilde{V}_{dj}(M_d^1M_j)\tilde{\rho}_{jd}(M_jM_d^2) \\
&- \tilde{\rho}_{j_1d}(M_d^1M_j)\tilde{V}_{jd}(M_jM_d^2)], \\
\rho_{j_1d}(M_{j_1}M_d) &= \rho^*_{dj_1}(M_dM_{j_1}),
\end{aligned} \tag{17}$$

where

$$\begin{aligned}
\tilde{V}_{j_1d}(M_{j_1}M_d) &= \tilde{V}^*_{dj_1}(M_{j_d}M_{j_1}) \\
&= V_{j_1d}(M_{j_1}M_d)\exp(-i\omega_Lt),
\end{aligned}$$

$$\tilde{\rho}_{j_1d}(M_{j_1}M_d) = \rho_{j_1d}(M_{j_1}M_d)\exp(-i\omega_Lt),$$

$$\Delta_L^{j_1} = \omega_{j_1d} - \omega_L, \quad \Delta_{j_1j_2} = \Delta_L^{j_1} - \Delta_L^{j_2}.$$

We disregard the effect of relaxation processes associated with collisions, assuming that the density of the medium is quite low. Obviously, the effect of collisions for dense media is significant and may lead not only to a depolarization between states, but also to a displacement of magnetic sublevels. Since the effect of other (outer) energy levels is neglected, Eqs. (17) must be supplemented with a normalization condition:

$$\sum_{j=b,c,d}\sum_{M_j}\rho_{jj}(M_jM_j) = 1. \tag{18}$$

3. RESONANCE FLUORESCENCE SPECTRA

Let us derive the general expressions for the resonance fluorescence spectrum when the observation is carried out in an arbitrary direction. In order to construct the spectrum, we use the apparatus of the atom–photon density matrix [19]. We define the operation Cnv(…) which puts an $n \times n$ matrix, where

$$n = \sum_{j=b,c,d}(2J_j + 1),$$

in correspondence with a vector of length $N = n \times n$ according to the following rule:

$$\begin{aligned}
\mathrm{Cnv}(h) = (&h_{bb}(J_bJ_b), h_{bb}(J_bJ_b - 1), \ldots, \\
&h_{j_1j_2}(M_{j_1}M_{j_2}), \ldots, h_{dd}(-J_d - J_d)).
\end{aligned}$$

Using this transformation, we can write Eqs. (17) for the density matrix in the form

$$i\frac{d\mathrm{Cnv}(\rho)}{dt} = M\mathrm{Cnv}(\rho). \tag{19}$$

Using matrix *M*, we can write the equations for the atom–photon density matrix in the form

$$\begin{aligned}
&i\frac{d\mathrm{Cnv}(\rho_\nu^{\mathbf{k}_z\sigma})}{dt} \\
&= (M - \nu I)\mathrm{Cnv}(\rho_\nu^{\mathbf{k}_z\sigma}) + \mathrm{Cnv}(\rho^{\mathbf{k}_z}T^{\mathbf{k}_z\sigma}), \\
&\nu = \omega_{k_z} - \omega_L, \quad \mathbf{k}_z = |k|\mathbf{e}_z,
\end{aligned} \tag{20}$$

where $\rho_\nu^{\mathbf{k}\lambda}$ is the operator of the atom–photon density matrix and $\rho^{\mathbf{k}}$ is the operator of the atomic density matrix, which are obtained from Eqs. (17) and (20) in the case when the quantization axis *z* is directed along the wave vector of the emitted photon. The elements of operator $T^{\mathbf{k}_z\sigma}$ are defined by the relation

$$(T^{\mathbf{k}\sigma})_{j_1j_2}(M_{j_1}M_{j_2}) = [(G^{\mathbf{k}})_{j_1j_2}(M_{j_1}M_{j_2})a_{\mathbf{k}\sigma'}, \overset{+}{a}_{\mathbf{k}\sigma}].$$

The expression for the polarization components of the resonance fluorescence spectrum in the case of the observation along vector **k** has the form

$$A^{\mathbf{k}}_{\lambda\lambda'}(\nu)d\nu dO = i\sum_{M_d}[T^{k\lambda}\rho^{\mathbf{k}\lambda'}_{\nu} + (\rho^{\mathbf{k}\lambda'}_{\nu})^*(T^{k\lambda})]_{dd} \times (M_d M_d)\frac{\omega_k^2}{c^3}d\nu dO, \quad \lambda, \lambda = \pm 1. \tag{21}$$

In order to derive the expressions for the polarization components in the unified system of coordinates, we must express the elements of the atom–photon matrix $(\rho^{\mathbf{k}\lambda}_{\nu})_{j_1 j_2}(M_{j_1}M_{j_2})$ in terms of the elements of matrix $\rho^{\mathbf{k}_z\lambda}_{\nu}$ obtained from Eqs. (17), (20) in the unified system of coordinates. This transformation can be carried out in the standard form [18]:

$$(\rho^{\mathbf{k}\lambda}_{\nu})_{j_1 j_2}(M_{j_1}M_{j_2}) = \sum_{\sigma = 0, \pm 1}(D^1_{\sigma\lambda})^* \times \sum_{M'_{j_1}M'_{j_2}}\left(D^{J_{j_1}}_{M'_{j_1}M_{j_1}}\right)^* D^{J_{j_2}}_{M'_{j_2}M_{j_2}}(\rho^{\mathbf{k}_z\lambda}_{\nu})_{j_1 j_2}(M'_{j_1}M'_{j_2}). \tag{22}$$

Expressions (21) and (22) completely define the spectral, angular, and polarization characteristics of the spontaneous radiation emitted by the system under investigation. It should be noted that the transformation of the unified system of coordinates to that in which the observation is carried out cannot be applied in the general case to the atomic density matrix or directly to the vector **E** of the exciting field while calculating the quantities $A^{\mathbf{k}}_{\lambda\lambda'}(\nu)$ at the stage of the solution of Eq. (20). This is possible only in an analysis of systems for which the density of states of the electromagnetic field is isotropic. Otherwise, the radiation relaxation operator and, hence, matrix M depend on the choice of the direction of the z-axis.

The expression for the spontaneous fluorescence intensity integrated over the spectrum in the direction of vector **k** was derived by us earlier [15]:

$$S^{\mathbf{k}}_{\lambda\lambda'} = \int A^{\mathbf{k}}_{\lambda\lambda'}(\nu)d\nu = \sum_{j_1, j_2 = b, c}\sum_{M_{j_1}M_{j_2}}\sum_{M_d} g^{\mathbf{k}\lambda}_{dj_1}(M_{j_1}M_d) \times g^{\mathbf{k}\lambda'}_{j_2 d}(M_d M_{j_2})\rho^k_{j_1 j_2}(M_{j_1}M_{j_2}). \tag{23}$$

A transition to the unified system of coordinates for the quantities $\rho^k_{j_1 j_2}(M_{j_1}M_{j_2})$ is carried out in analogy with transformation (22):

$$\rho^{\mathbf{k}}_{j_1 j_2}(M_{j_1}M_{j_2}) = \sum_{M'_{j_1}M'_{j_2}}\left(D^{J_{j_1}}_{M'_{j_1}M_{j_1}}\right)^* D^{J_{j_2}}_{M'_{j_2}M_{j_2}}\rho^{\mathbf{k}_z}_{j_1 j_2}(M'_{j_1}M'_{j_2}). \tag{24}$$

The total intensity of the spontaneous radiation is defined by

$$I_{\text{total}} = \int\sum_{\lambda = \pm 1} S_{\lambda\lambda}dO = \sum_{j_1, j_2 = b, c}\sum_{M_{j_1}M_{j_2}}[\gamma_{j_1}\gamma_{j_2}(M_{j_1}M_{j_2})\rho^{\mathbf{k}_z}_{j_2 j_1}(M_{j_2}M_{j_1}) + \rho^{\mathbf{k}_z}_{j_1 j_2}(M_{j_1}M_{j_2})\gamma_{j_2 j_1}(M_{j_2}M_{j_1})]. \tag{25}$$

In fact, the right-hand side of Eq. (25) is the trace of operator $\sum_{\mathbf{k}}(\hat{G}^{k*}\hat{G}^k\rho^{\mathbf{k}_z} + \rho^{\mathbf{k}_z}\hat{G}^{k*}\hat{G}^k)$ in the variables pertaining to upper energy levels. Since (22) and (24) are unitary transformations, the total resonance fluorescence intensity does not depend on the choice of the direction of the z-axis as expected. In the case when the density of states of the electromagnetic field is isotropic, expression (25) can be reduced to the classical form

$$I_{\text{total}} = 2\sum_{j = a, b}\sum_{M_j}\gamma_{jj}(M_j M_j)\rho^{\mathbf{k}_z}_{jj}(M_j M_j). \tag{26}$$

It can be seen from formulas (26) and (13) that the interference effects accompanying the spontaneous emission of radiation in real physical systems do not affect the dynamics of an atom in the laser field and the total resonance fluorescence intensity. However, interference effects should be taken into account in an analysis of the angular distribution of spontaneous radiation (see [15]). Interference effects also make their contribution to the spectral characteristics of spontaneous radiation. It should be noted that interference processes are naturally taken into account for the polarization components of spectrum (21) and in Eq. (20) through the structure of operators $T^{\mathbf{k}\sigma}$, which contain the binding constants for two upper energy levels simultaneously. The final expression for the spectrum in this case contains nonzero terms with nondiagonal elements of the atomic density matrix, which reflect the contribution from interference processes.

4. AN ANALYSIS OF THE RESONANCE FLUORESCENCE SPECTRA FOR THE *D*-LINE OF ALKALI METAL VAPORS

Using the formalism developed in the previous sections, we consider the resonance fluorescence of the D-line emitted by vapors of alkali metals in the stationary state. We introduce the following notation for energy levels: $b \longrightarrow P_{3/2}$, $c \longrightarrow P_{1/2}$, and $d \longrightarrow S_{1/2}$, so that $J_b = 3/2$, $J_c = 1/2$, $J_d = 1/2$. The energy level diagram and the structure of transitions for an atom under ordinary conditions are presented in Fig. 2. In the subsequent analysis of the spectra, we will be using the well-known fact that the relaxation constants of magnetic sublevels for a free atom are identical: $\gamma_{j_1j_2}(M_{j_1}M_{j_2}) = \gamma$ (see Appendix).

Since the steady-state solution of Eq. (20) is a rational function of ν, we can present the resonance fluorescence spectra in the form of the following expansion:

$$A^{\mathbf{k}}_{\lambda\lambda'}(\nu) = \sum_{1}^{N}\frac{f_m}{\gamma_m + i(\Omega_m - \nu)} + \text{c.c.}, \qquad (27)$$

where $N = n \times n$, $n = 8$ is the number of energy levels in the system under investigation, and γ_m and Ω_m are the imaginary and real components of the mth eigenvalue of matrix M, which characterize the width and frequency of the spectral line, respectively. In the stationary case, f_m can be expressed in terms of the eigenvectors of matrix M, atomic density matrix $\rho^{\mathbf{k}}$, and operators $T^{\mathbf{k}\sigma}$. Expansion (27) can be used to determine all the characteristics pertaining to an individual line. The energy of the spontaneous radiation of an individual line is a parameter of practical importance. It can be expressed in the form

$$F^{\mathbf{k}m}_{\lambda\lambda'} = \int_{-\infty}^{\infty}\frac{f_m}{\gamma_m + i(\Omega_m - \nu)}d\nu + \text{c.c.} = 2\pi\text{Re}(f_m). \qquad (28)$$

The imaginary component of f_m corresponds to the dispersion-type component of the spectrum and makes zero contribution to the fluorescence energy of the line. It should be noted that the relation between the real and imaginary components of constants f_m is such that quantities $A^{\mathbf{k}}_{\lambda\lambda'}(\nu)$ in the "wings" of the spectrum decrease in proportion to $1/\nu^4$ as in the case of the Mollow spectrum [20].

Figures 3a and 4a show the dependences of the frequencies of spectral lines with a nonzero intensity on the field strength of the pumping wave. For convenience of computation and presentation of the results, ratio $\Delta_{bc}/\gamma = 120$ is set much smaller than for real systems (e.g., $\Delta^{\text{Na}}_{bc}/\gamma^{\text{Na}} \approx 6 \times 10^4$). However, the results obtained in this section for strong fields ($V_m \gg \gamma$) remain unchanged. The quantity

$$V_m = \sqrt{\sum_{\sigma}\frac{(V_{jd}(\sigma))^2}{2J_j + 1}}$$

defines the Rabi frequency for the system under investigation. It should be noted that since $\|\mu_{bd}\|/\sqrt{2J_b + 1} = \|\mu_{cd}\|/\sqrt{2J_c + 1}$, V_m does not depend on j.

Figures 3 and 4 correspond to the resonance fluorescence observation for the linear and circular polarization of the laser field, respectively. It was shown by us earlier [17] that the maximum possible number of lines in the spectrum is determined by the formula $N_\Lambda = n(n - 1) + 1$ ($N_\Lambda = 57$ in the case under investigation). The much smaller number of peaks in Figs. 3a and 4a is apparently due to our assumption that the magnetic sublevels are degenerate in energy. It is interesting to note, however, that the numbers of quasi-energy levels formed during the excitation differ considerably for different polarizations of the laser wave. For example, three quasi-energy levels are formed at the top and at the bottom, while five energy levels are formed at the top and at the bottom for the circular polarization of the exciting wave. The number of peaks in these cases is 7 and 21, respectively.

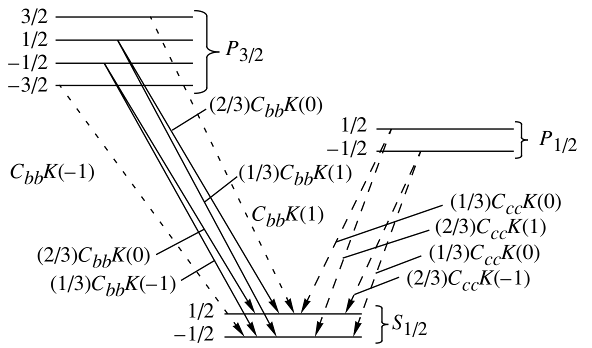


Fig. 2. Diagram of spontaneous transitions for the *D*-line emitted by vapors of alkali atoms.

Apart from the polarization composition of the pumping wave, the shape of the spectrum depends on the direction of the resonance fluorescence observation. The polarization composition of the spontaneous radiation varies along with the intensity of individual lines in the spectrum. Figure 5 shows the variation of the spontaneous fluorescence energy of spectral lines excited by linearly polarized radiation. The direction of observation coincides with the vector of field strength of the laser wave. According to our calculations, the nondiagonal elements $A^{\mathbf{k}}_{1-1}(\nu)$ and $A^{\mathbf{k}}_{-11}(\nu)$ are identically equal to zero, and the radiation is absolutely unpolarized in any spectral region. It should be noted that the fluorescence energy of the lines decreases with increasing field strength of the pumping wave. This fact agrees with a similar result obtained for the total resonance fluorescence intensity obtained by us earlier [15].

Another pattern of spontaneous radiation is observed in the perpendicular direction, i.e., when the observation axis is orthogonal to the field strength vector of the pumping wave. In this case (see Fig. 3d), the nondiagonal elements $A^{\mathbf{k}}_{1-1}(\nu)$ and $A^{\mathbf{k}}_{-11}(\nu)$ differ from zero, and their contribution significantly changes the shape

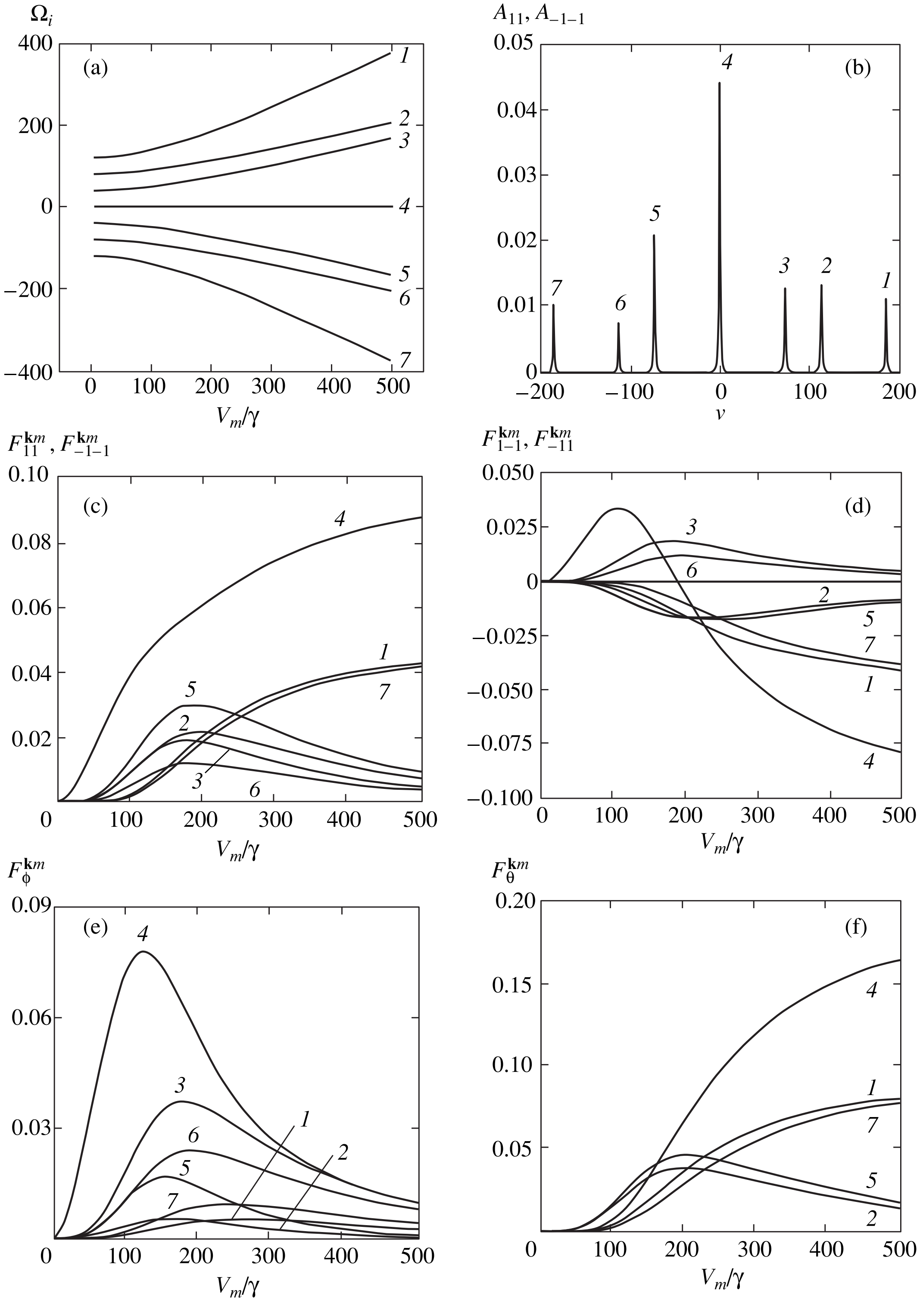


Fig. 3. Characteristics of the resonance fluorescence spectra as functions of the field strength of a linearly polarized pimping wave: (a) line frequencies; (b) characteristic resonance fluorescence spectrum for circular polarizations; (c) energies of lines for the observation of circular polarizations; $F_{11}^{\mathbf{k}m}$, $F_{-1-1}^{\mathbf{k}m}$; (d) nondiagonal elements $F_{-11}^{\mathbf{k}m}$, $F_{1-1}^{\mathbf{k}m}$; (e), (f) energies of lines for a linear polarized radiation. The observation is carried out at right angles to the field strength vector of the pumping wave. The values of the parameters: $E_x = E$, $E_y = 0$, $E_z = 0$, $\Delta_L^b = 80$, $\Delta_L^c = -40$, and $\gamma = 1$. The figures on the curves correspond to spectral lines in order of decreasing frequency.

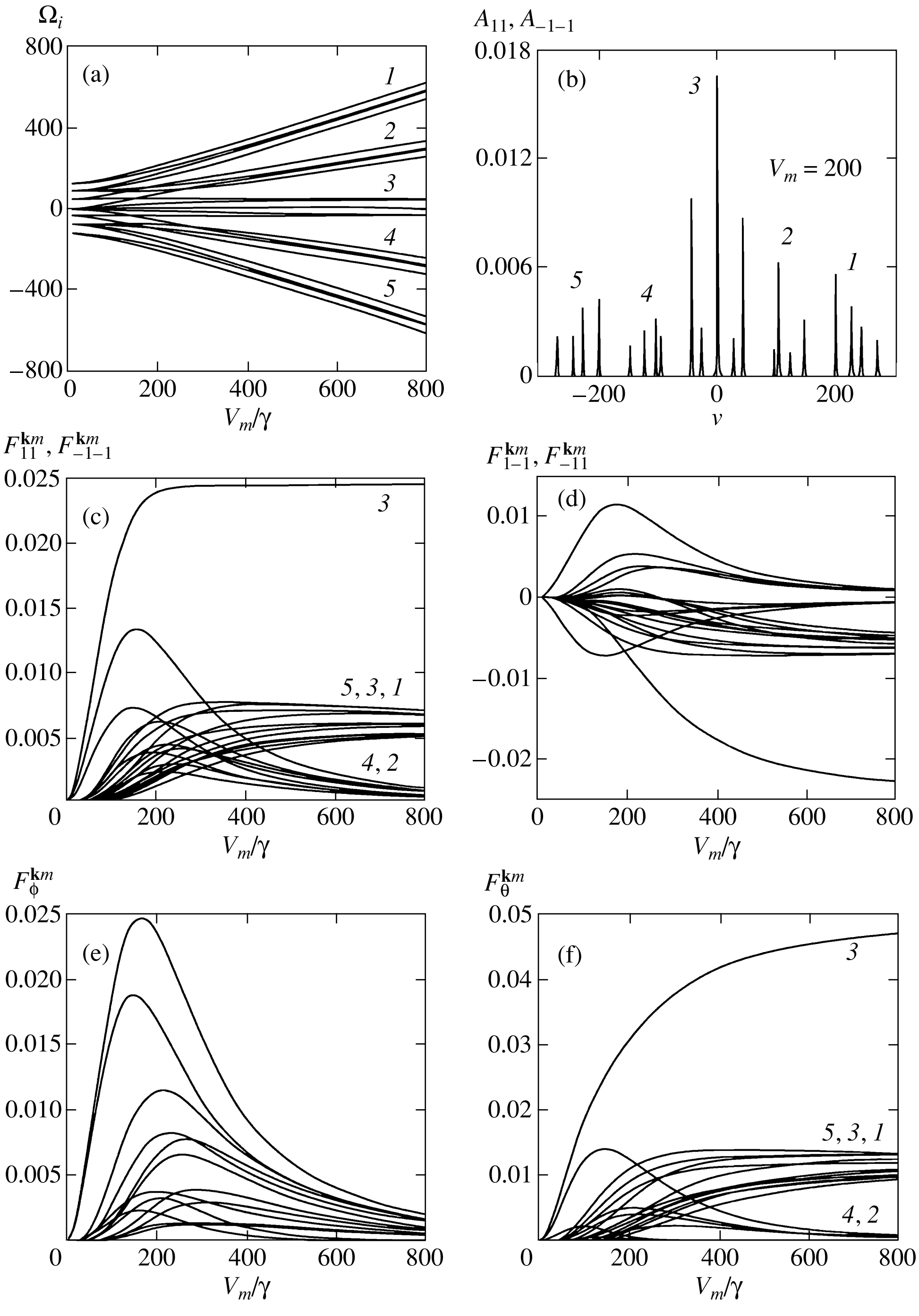


Fig. 4. Characteristics of the resonance fluorescence spectra as functions of the field strength of a circularly polarized pimping wave (the same curves as in Fig. 3). The observation is carried out at right angles to the field strength vector of the pumping wave. The values of the parameters: $E_x = E/\sqrt{2}$, $E_y = 0$, $E_z = iE/\sqrt{2}$, $\Delta^b_L = 80$, $\Delta^c_L = -40$, and $\gamma = 1$.

of the spectra for observations of a linearly polarized radiation along axes $\mathbf{e}_\phi$ and $\mathbf{e}_\theta$ (see Fig. 6):

$$\begin{aligned} F^{\mathbf{k}m}_{\theta} &= (F^{\mathbf{k}m}_{11} + F^{\mathbf{k}m}_{-1-1} - F^{\mathbf{k}m}_{1-1} - F^{\mathbf{k}m}_{-11})/2, \\ F^{\mathbf{k}m}_{\phi} &= (F^{\mathbf{k}m}_{11} + F^{\mathbf{k}m}_{-1-1} + F^{\mathbf{k}m}_{1-1} + F^{\mathbf{k}m}_{-11})/2. \end{aligned} \tag{29}$$

In the case when polarized radiation is observed in the $\mathbf{e}_\phi$ direction (see Fig. 3e), the energy of lines in the spectrum attains its maximum value in the range $\Delta_{bc} < V_m < 2\Delta_{bc}$ and then decreases to zero with increasing field strength. A spontaneous radiation polarized along the θ axis differs significantly. It contains only five peaks (see Fig. 3f). The spectral lines positioned symmetrically to the transition lines relative to the frequency of the exciting radiation are degenerate. As the laser field strength increases, the heights of the central

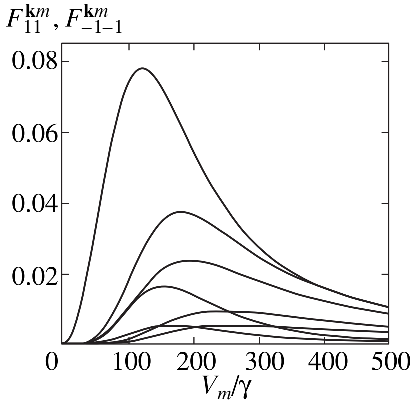


Fig. 5. Energies corresponding to lines of the resonance fluorescence spectra as functions of the field strength of a circularly polarized pumping wave. The observation is carried out at right angles to the field strength vector of the pumping wave. The values of the parameters: $E_x = E$, $E_y = 0$, $E_z = 0$, $\Delta_L^b = 80$, $\Delta_L^c = -40$, and $\gamma = 1$.

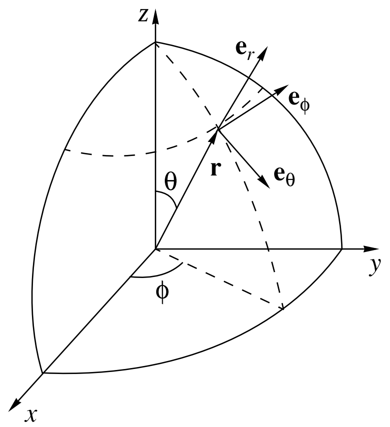


Fig. 6. Directions of linearly polarized radiation.

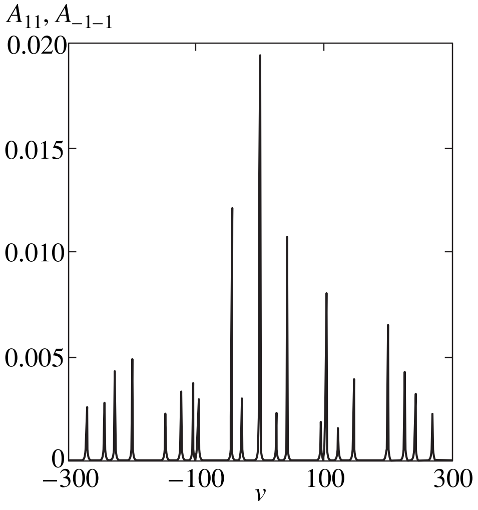


Fig. 7. Resonance fluorescence spectrum for atoms near the metal surface. The observation is carried out along the normal to the surface. The values of the parameters: $E_x = E/\sqrt{2}$, $E_y = 0$, $E_z = iE/\sqrt{2}$, $\Delta_L^b = 80$, $\Delta_L^c = -40$; the radiation relaxation constant for a free atom is taken as unity.

and extreme peaks increase, while the heights of the peaks corresponding to the transition frequencies decrease. At high powers ($V_m \gg \Delta_{bc}$), the shape of the spectrum corresponds to a two-level atom in a laser field with the transition frequency detuning Δ_{bc}. It should be noted that the polarization of the spontaneous radiation varies upon an increase in the strength of the laser field: the radiation in a weak field is mainly polarized in the $\mathbf{e}_\phi$ direction, and in a strong field, in the $\mathbf{e}_\theta$ direction.

An analysis of the resonance fluorescence in the case of the circular polarization of the exciting field proved that the behavior of the spontaneous radiation upon an increase in the laser field strength corresponds to the case discussed above (see Fig. 4). The shape of the spectra is obviously independent of the chosen direction of observation in a plane perpendicular to the direction of wave propagation. As in the case of the linear polarization of the exciting field, the number of lines $N_\theta = 17$ in the spectra for the observation in the $\mathbf{e}_\theta$ direction (see Fig. 4f) is smaller than the number of lines $N_{11} = N_{-1-1} = 21$ for the observation of the circular polarization (see Fig. 4c). However, in contrast to the previous case, the number of peaks ($N_\phi = 13$) is also smaller for the radiation observed in the $\mathbf{e}_\phi$ direction (see Fig. 4f). Another interesting fact is that for a large field strength, the groups of lines contain several peaks belonging to adjacent components for a small strength of the exciting field (see Fig. 4a).

Finally, we analyzed the resonance fluorescence spectra in the case when an atom is near the metal surface (see version (a) in Appendix). As expected, the change in the radiation relaxation operator does not produce any significant effect on the fluorescence energy of the lines. This is due to the fact that the constants f_m are functions of the eigenvectors of matrix M and of the values of the density matrix elements only. The variation of the relaxation constants for $V_m \gg \gamma_{j_1 j_2}(M_{j_1} M_{j_2})$ does not affect any of these quantities seriously. However, the intensity of lines in the spectrum changes significantly (see Figs. 7 and 4b). Indeed, the fluorescence intensity of the center of the line is determined by ratio f_m/γ_m, while linewidth γ_m strongly depends on the relaxation constants. The variation of the values of γ_m is not uniform, which leads to a change in the relative peak intensities.

5. CONCLUSION

The resonance fluorescence spectra of a degenerate three-level atom of the V-type in the field of a high-intensity monochromatic wave of an arbitrary polarization are studied. The general form of the radiation relaxation operator is obtained for the systems under investigation. The influence of interference effects accompanying the spontaneous emission on the structure of the radiation relaxation is analyzed. It is shown that for most real physical systems, the interference processes do not affect the dynamics of occupancies of atomic sublevels. The necessary condition under which the influence of the interference effects should be taken into account is a variation of the density of states of the electromagnetic field. The expressions for relaxation constants near the metal surface are derived in the Appendix. It is shown that in this case, the form of the radiation relaxation operator depends considerably on the choice of the direction of the quantization axis. The expressions for the resonance fluorescence spectra for the observation in a chosen direction are derived using the apparatus of the atom–photon matrix.

The steady-state resonance fluorescence spectra are analyzed in detail for the *D*-line emitted by the vapors of alkali metals. The angular and polarization characteristics of the spontaneous fluorescence of atoms are investigated. It is shown that the number of peaks in the spectrum may decrease depending on the choice of the direction of observation and the polarization of the spontaneous radiation studied.

ACKNOWLEDGMENTS

The authors are grateful to A.N. Starostin, A.G. Leonov, D.I. Chekhov, and A.A. Rudenko for fruitful discussions and valuable remarks on the subject.

This work was financed by the Russian Foundation for Basic Research (projects nos. 99-02-17063 and 99-15-96539).

APPENDIX

The Structure of the Radiation Relaxation Operator for a Degenerate Three-Level Atom of the V-Type Near the Metal Surface

We will derive and analyze the structure of the radiation relaxation operator for the transitions $P_{3/2} \longrightarrow S_{1/2}$, $P_{1/2} \longrightarrow S_{1/2}$ of the *D*-line emitted by atoms of alkali metals near the surface in the following two cases: (a) the *z*-axis is directed along the normal to the surface, and (b) the *z*-axis is parallel to the surface. The radiation relaxation operator in these cases differs from that for an atom in free space since the atom cannot emit in directions along which the distance to the metal surface is smaller than the wavelength of the emitted wave. For this reason, the quantities

$$K(\sigma, \sigma') = \int\limits_{\Phi} D^{J_1}_{\lambda_1\sigma_1} D^{J_2}_{\lambda_2\sigma_2} \frac{dO}{4\pi} \tag{A.1}$$

appearing in the expression (13) for the radiation relaxation operator are integrated not over the total solid angle, but over a certain angle Φ. In the case of spontaneous emission near the metal surface, the solid angle into which radiation cannot be emitted has the shape of a cone (see Fig. 8). In the space of the spherical angles ϕ, θ, this cone occupies different regions depending on the choice of the *z*-axis, which determines the properties of quantities $K(\sigma, \sigma')$ and the form of the radiation relaxation operator.

For our subsequent calculations, the explicit form of functions $s^1_{m'm}(\beta)$ determining the transformation of the wave functions upon a rotation of the reference frame about the nodal line [18] is required:

$$\begin{gathered} s^1_{10} = s^1_{0-1} = -s^1_{01} = -s^1_{-10} = \frac{1}{\sqrt{2}}\sin\beta, \\ s^1_{11} = s^1_{-1-1} = \frac{1}{2}(1+\cos\beta), \\ s^1_{1-1} = s^1_{-11} = \frac{1}{2}(1-\cos\beta). \end{gathered} \tag{A.2}$$

We recall the values of constants $K(\sigma, \sigma')$ and the form of nonzero elements of the radiation relaxation operator for a free atom:

$$K(\sigma, \sigma') = K_p\delta_{\sigma\sigma'} = \frac{2}{3}\delta_{\sigma\sigma'},$$

$$\begin{gathered} \gamma_{bb}\left(\frac{3}{2}, \frac{3}{2}\right) = C_{bb}K(1, 1), \\ \gamma_{bb}\left(-\frac{3}{2}, -\frac{3}{2}\right) = C_{bb}K(-1, -1), \\ \gamma_{bb}\left(\frac{1}{2}, \frac{1}{2}\right) = C_{bb}\frac{K(1, 1) + 2K(0, 0)}{3}, \\ \gamma_{bb}\left(-\frac{1}{2}, -\frac{1}{2}\right) = C_{bb}\frac{K(-1, -1) + 2K(0, 0)}{3}, \\ \gamma_{cc}\left(\frac{1}{2}, \frac{1}{2}\right) = C_{cc}\frac{2K(1, 1) + K(0, 0)}{3}, \\ \gamma_{cc}\left(-\frac{1}{2}, -\frac{1}{2}\right) = C_{cc}\frac{2K(-1, -1) + K(0, 0)}{3}, \end{gathered} \tag{A.3}$$

where

$$C_{j_1j_2} = \frac{2\pi\|\mu_{j_1d}\|\|\mu_{j_2d}\|\omega^3_{jd}}{\hbar\sqrt{(2J_{j_1}+1)(2J_{j_2}+1)}c^3}, \quad j_1, j_2 = b, c.$$

In view of the well-known relation

$$\frac{\|\mu_{bd}\|}{\sqrt{2J_b+1}} = \frac{\|\mu_{cd}\|}{\sqrt{2J_c+1}}, \tag{A.4}$$

all the relaxation constants for magnetic sublevels are identical.

(a) If the *z*-axis is directed along the normal to the surface (see Fig. 8a), we have

$$K_a(\sigma, \sigma') = K_p\delta_{\sigma\sigma'} - \frac{1}{2\pi}\int\limits_0^{2\pi}\exp[i(\sigma-\sigma')\alpha]d\alpha \times \int\limits_{(\pi/2+\beta_0)}^{\pi}\sum_{\lambda=\pm1} s^1_{\lambda\sigma}s^1_{\lambda\sigma'}\frac{\sin\beta}{2}d\beta. \tag{A.5}$$

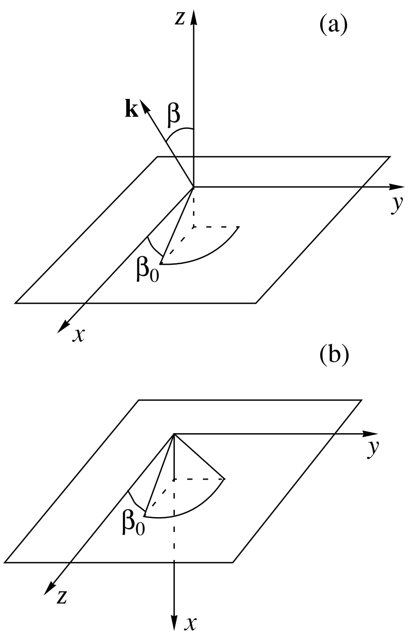


Fig. 8. Reference frame for an atom near the metal surface: (a) z axis is perpendicular to the surface, (b) z axis is parallel to the surface.

It follows immediately from this expression that $K_a(\sigma, \sigma') = 0$ for $\sigma \neq \sigma'$, and

$$K_a(1, 1) = K_a(-1, -1) = K_p\delta_{\sigma\sigma'}$$

$$-\frac{1}{2\pi}\int_0^{2\pi} d\alpha \int_{\frac{\pi}{2}+\beta_0}^{\pi} \frac{1}{2}(1+\cos^2\beta)\frac{\sin\beta}{2}d\beta \tag{A.6}$$

$$= \frac{1}{3} + \frac{1}{4}\left(\sin\beta_0 + \frac{\sin^3\beta_0}{3}\right),$$

$$K_a(0, 0) = K_p\delta_{\sigma\sigma'} - \frac{1}{2\pi}\int_0^{2\pi} d\alpha \int_{\pi/2+\beta_0}^{\pi} \sin^2\beta\frac{\sin\beta}{2}d\beta \tag{A.7}$$

$$= \frac{1}{3} + \frac{1}{2}\left(\sin\beta_0 - \frac{\sin^3\beta_0}{3}\right).$$

The nonzero elements of the radiation relaxation operator supplementing (A.3) have the form

$$\gamma_{bc}\left(\frac{1}{2}, \frac{1}{2}\right) = \gamma_{cb}\left(\frac{1}{2}, \frac{1}{2}\right) = C_{bc}\frac{\sqrt{2}}{3}[K(0, 0) - K(1, 1)],$$

$$\gamma_{bc}\left(-\frac{1}{2}, -\frac{1}{2}\right) = \gamma_{bc}\left(-\frac{1}{2}, -\frac{1}{2}\right) \tag{A.8}$$

$$= C_{bc}\frac{\sqrt{2}}{3}[K(-1, -1) - K(0, 0)].$$

It follows from formulas (A.6)–(A.8) that in contrast to the radiation relaxation operator for an atom in free space, the corresponding operator for an atom near the metal surface has nonzero nondiagonal elements. Constants $\gamma_{bc}(1/2,\ 1/2)$ and $\gamma_{bc}(-1/2,\ -1/2)$ describe the spontaneous emission of the atom associated with the induced interference between the corresponding magnetic sublevels for levels b and c.

(b) If the z-axis is parallel to the surface (see Fig. 8b), the general form of constants $K_b(\sigma, \sigma')$ is as follows:

$$K_b(\sigma, \sigma') = K_p\delta_{\sigma\sigma'}$$

$$-\frac{1}{2\pi}\int_{\beta_0}^{\pi-\beta_0} \sum_{\lambda=\pm1} s^1_{\lambda\sigma}s^1_{\lambda\sigma'}\frac{\sin\beta}{2}d\beta \int_{-\alpha'}^{\alpha'} \exp[i(\sigma-\sigma')\alpha]d\alpha, \tag{A.9}$$

where

$$\alpha' = \arctan\sqrt{\frac{1}{\sin^2\beta_0} - \frac{1}{\sin^2\beta}}.$$

Since the functions

$$\sum_{\lambda=\pm1} s^1_{\lambda\sigma}s^1_{\lambda\sigma'}\frac{\sin\beta}{2}$$

are odd relative to the angle $\beta = \pi/2$ for $\sigma - \sigma' = \pm1$, we have $K_b(\sigma, \sigma') = 0$, and

$$K_b(1, 1) = K_p - \int_{\beta_0}^{\pi-\beta_0} \frac{1}{2}(1+\cos^2\beta)\frac{\sin\beta}{2}d\beta \int_{-\alpha'}^{\alpha'} \frac{1}{2\pi}d\alpha$$

$$= \frac{2}{3} - \frac{1}{\pi}\int_{\beta_0}^{\pi/2} \arctan\left(\sqrt{\frac{1}{\sin^2\beta_0} - \frac{1}{\sin^2\beta}}\right)$$

$$\times(1+\cos^2\beta)\frac{\sin\beta}{2}d\beta = \frac{2}{3} - \frac{1}{6\pi u(1+u)^2\sqrt{u-1}}$$

$$\times\left[(u^2-1)E\left(\frac{\pi}{2} - \beta_0, 1 + \tan^2\beta_0\right)\right. \tag{A.10}$$

$$+ (1 - 3u - 5u^2)F\left(\frac{\pi}{2} - \beta_0, 1 + \tan^2\beta_0\right)$$

$$\left. + u(3+4u)\pi\left(\cot^2\beta_0, \frac{\pi}{2} - \beta_0, 1 + \tan^2\beta_0\right)\right].$$

Here, $F(\varsigma\ k)$, $E(\varsigma\ k)$, and $\pi(c, \varsigma\ k)$ are Legendre normal elliptic integrals of the first, second, and third kind, respectively, and $u = 1/\sin^2\beta_0$. Similarly, for $K_b(0, 0)$ we have

$$K_b(0, 0) = \frac{2}{3} - \frac{2}{\pi}\int_{\beta_0}^{\pi/2} \arctan\left(\sqrt{\frac{1}{\sin^2\beta_0} - \frac{1}{\sin^2\beta}}\right)$$

$$\times(1-\cos^2\beta)\frac{\sin\beta}{2}d\beta = \frac{2}{3}\frac{1}{6\pi u(1+u)^2\sqrt{u-1}}$$

$$\times\left[(u^2-1)E\left(\frac{\pi}{2}-\beta_0, 1+\tan^2\beta_0\right) + (1+3u+u^2)F\left(\frac{\pi}{2}-\beta_0, 1+\tan^2\beta_0\right) - u(3+2u)\pi\left(\cot^2\beta_0, \frac{\pi}{2}-\beta_0, 1+\tan^2\beta_0\right)\right]. \tag{A.11}$$

The expression for the nondiagonal element $K_b(1, -1)$ differs from those listed above in that the exponential term $\exp[i(\sigma-\sigma')\alpha]$ is not equal identically to unity:

$$\begin{aligned} K_b(1,-1) &= K_b^*(-1,1) = -\frac{1}{4\pi} \\ &\times\int_{\beta_0}^{\pi/2}\left\{i\sin\left(\left[2\arctan\left(\sqrt{\frac{1}{\sin^2\beta_0}-\frac{1}{\sin^2\beta}}\right)\right]-1\right)\right\} \\ &\times(1-\cos^2\beta)\frac{\sin\beta}{2}d\beta = -\frac{2}{3\pi(1+u)^3\sqrt{u}} \\ &\times\left\{(2u^3+6u^2+3u-1)\left[E(\arcsin(\sqrt{u}), \sin^2\beta_0) - E\left(\frac{\pi}{2}, \sin^2\beta_0\right)\right] + (2u^3+5u^2+u-2)\right. \\ &\times\left[F(\beta_0, 1+\tan^2\beta_0) - F\left(\frac{\pi}{2}, 1+\tan^2\beta_0\right)\right] \\ &- 3u\left[\pi\left(\frac{1}{1+\sin^2\beta_0}, \beta_0, 1+\cot^2\beta_0\right)\right. \\ &\left.\left. - \pi\left(\frac{1}{1+\sin^2\beta_0}, \frac{\pi}{2}, 1+\cot^2\beta_0\right)\right]\right\}. \end{aligned} \tag{A.12}$$

It should be emphasized that in spite of its appearance, the above expression is purely imaginary.

The elements of the radiation relaxation operator supplementing (A.8) can be written in the form

$$\begin{aligned} \gamma_{bb}\left(\frac{3}{2}, -\frac{1}{2}\right) &= \frac{1}{\sqrt{3}}C_{bb}K(1,-1), \\ \gamma_{bb}\left(-\frac{1}{2}, \frac{3}{2}\right) &= \frac{1}{\sqrt{3}}C_{bb}K(-1,1), \\ \gamma_{bb}\left(\frac{1}{2}, -\frac{3}{2}\right) &= \frac{1}{\sqrt{3}}C_{bb}K(1,-1), \\ \gamma_{bb}\left(-\frac{3}{2}, \frac{1}{2}\right) &= \frac{1}{\sqrt{3}}C_{bb}K(-1,1), \\ \gamma_{bc}\left(\frac{3}{2}, -\frac{1}{2}\right) &= C_{bc}K(1,-1), \\ \gamma_{cb}\left(-\frac{1}{2}, \frac{3}{2}\right) &= C_{cb}K(-1,1), \\ \gamma_{bc}\left(-\frac{3}{2}, \frac{1}{2}\right) &= -C_{bc}K(1,-1), \\ \gamma_{cb}\left(\frac{1}{2}, -\frac{3}{2}\right) &= -C_{cb}K(-1,1). \end{aligned} \tag{A.13}$$

These results differ from those obtained for version (a) in that the structure of the radiation relaxation operator acquires elements connecting not only the levels with identical values of the magnetic moment component, but also the levels with $M_{j_1} - M_{j_2} = \pm 2$. Besides, the interference effects will take place between the sublevels of $P_{3/2}$ even in the absence of excitation of level $P_{1/2}$.

REFERENCES

1. P. Zhou and S. Swain, Phys. Rev. Lett. **77**, 3995 (1996).
2. E. Paspalakis and P. L. Knight, Phys. Rev. Lett. **81**, 293 (1998).
3. E. Paspalakis, N. J. Kylstra, and P. L. Knight, Phys. Rev. Lett. **82**, 2079 (1999).
4. S. Y. Zhu and M. O. Scully, Phys. Rev. Lett. **76**, 388 (1996).
5. H. Lee, P. Polynkin, M. O. Scully, and S. Y. Zhu, Phys. Rev. A **55**, 4454 (1997).
6. G. C. Hegerfledt and M. B. Plenio, Phys. Rev. A **46**, 373 (1992).
7. S. Menon and G. S. Agarwal, Phys. Rev. A **61**, 013807 (2000).
8. M. O. Scully and S. Y. Zhu, Science **281**, 1973 (1998).
9. E. Paspalakis, S. Q. Gong, and P. L. Knight, Opt. Commun. **152**, 293 (1998).
10. S. Menon and G. S. Agarwal, Phys. Rev. A **57**, 4014 (1998).
11. A. A. Panteleev, Vl. K. Roerieh, and A. N. Starostin, Zh. Éksp. Teor. Fiz. **117**, 57 (2000) [JETP **90**, 50 (2000)].
12. A. K. Patnaik and G. S. Agarwal, J. Mod. Opt. **45**, 2131 (1998).
13. G. S. Agarwal, quant-ph/0005093.
14. H. Chew, Phys. Rev. A **38**, 3410 (1988).
15. A. A. Panteleev and Vl. K. Roerieh, Zh. Éksp. Teor. Fiz. **118**, 312 (2000) [JETP **91**, 273 (2000)].
16. D. A. Varshalovich, A. N. Moskalev, and V. K. Khersonskiĭ, *Quantum Theory of Angular Momentum* (Nauka, Leningrad, 1975; World Scientific, Singapore, 1988).
17. A. A. Panteleev, Zh. Éksp. Teor. Fiz. **111**, 440 (1997) [JETP **84**, 241 (1997)].
18. L. D. Landau and E. M. Lifshitz, *Course of Theoretical Physics,* Vol. 3: *Quantum Mechanics: Non-Relativistic Theory* (Nauka, Moscow, 1989, 4th ed.; Pergamon, New York, 1977, 3rd ed.).
19. S. Stenholm, *Foundations of Laser Spectroscopy* (Wiley, New York, 1984; Mir, Moscow, 1987).
20. B. R. Mollow, Phys. Rev. **188**, 1969 (1969).

Translated by N. Wadhwa

Journal of Experimental and Theoretical Physics, Vol. 92, No. 2, 2001, pp. 222–227.
Translated from Zhurnal Éksperimental’noĭ i Teoreticheskoĭ Fiziki, Vol. 119, No. 2, 2001, pp. 257–263.
Original Russian Text Copyright © 2001 by Saĭko.

On the Theory of Lasing at Vibronic Transitions in Impurity Crystals

A. P. Saĭko
Institute of Solid State and Semiconductor Physics, National Academy of Belarus, Minsk, 220072 Belarus
e-mail: saiko@ifttp.bas-net.by

Received September 5, 2000

Abstract—Some aspects of lasing at vibronic transitions in impurity crystals are theoretically studied. The threshold conditions for a vibronic laser are shown to be dependent on the strength of the interaction of optical centers with a local vibration, which forms the vibronic spectrum, and the crystal lattice temperature. The theory can easily be generalized to the spectrum containing a structureless phonon sideband and well agrees with the experimental temperature dependence of the output power of a Mg_2SiO_4:Cr^{4+} forsterite laser. © *2001 MAIK “Nauka/Interperiodica”.*

1. INTRODUCTION

The development of tunable solid-state lasers is one of the most urgent scientific and practical problems of laser physics, which has been actively elaborated in the last years.

The lasing conditions in solids are significantly complicated because of a strong adiabatic vibronic interaction, which determines the structure of vibronic levels, the main lasing parameters being dependent on the temperature of the laser medium and the strength of vibronic interaction (see, for example, [1–4]). In this case, a phenomenological description of lasing based on the use of rate equations for level populations and the polarization of the resonance medium, when the effect of the lattice and intramolecular vibrations is reduced only to the broadening of energy levels, cannot be considered adequate. A more fundamental, microscopic description of the laser action in vibronic systems, which explicitly takes the electron–vibrational interaction into account, will not only elucidate the fundamental problems of lasing at vibronic transitions in impurity crystals but also will be helpful in the search for new, more efficient laser materials.

The vibronic structure of optical spectra has been explicitly considered in the theoretical studies [5, 6] of threshold conditions and the radiation field dynamics of a vibronic laser performed within the framework of semiclassical equations of motion for field amplitudes and generalized operators of the transition between vibronic levels.

Below, the following aspects of the theory of a vibronic laser are discussed: The microscopic approach to the derivation of laser equations is realized, the linear Fokker–Planck equation is obtained for the distribution function of the laser field amplitudes, the threshold lasing conditions are discussed [7, 8], and the theory developed in this paper is applied to the description of lasing in a chromium-doped forsterite single crystal.

2. THE LINEAR FOKKER–PLANCK EQUATION FOR THE DISTRIBUTION FUNCTION OF LASER FIELD AMPLITUDES

We assume that impurities in a crystal, which exhibit lasing, can be treated as two-level quantum objects, whose ground and excited states are adiabatically coupled with phonon modes of the crystal lattice and local (for example, intramolecular) vibrations, whose frequencies exceed the maximum frequency of lattice phonons. The density matrix ρ of an ensemble of impurities interacting with vibrational modes and the electromagnetic field in a cavity is described by the master equation

$$\frac{\partial\rho(t)}{\partial t} = -i(L_f + L_a + L_{af} + L_l + L_{al} + L_L + L_{aL} + i\Lambda_f + i\Lambda_a + i\Lambda_l)\rho(t), \quad (1)$$

where

$$L_m X \equiv [H_m, X],$$

H_m and L_m are Hamiltonians and corresponding Liouvillians, respectively, $m = a, f, l$, etc.,

$$H_f = \omega a^+ a, \quad H_a = \varepsilon\sum_j R_j^z$$

are Hamiltonians of the single-mode electromagnetic field (radiation is collinear to an elongated pencil-like sample) and optical two-level centers;

$$H_{af} = g\sum_j (aR_j^+ + \text{H.c.})$$

is the operator of interaction of the optical centers with the radiation field; $H_l = \nu b^+ b$ is the Hamiltonian of the intramolecular vibration;

$$H_{al} = \sum_j \xi \nu R_j^z (b + b^+)$$

is the electron–vibrational interaction that produces the vibronic structure in the optical spectrum;

$$H_L = \sum_k \omega_k c_k^+ c_k$$

is the lattice Hamiltonian;

$$H_{aL} = \sum_k \lambda_k R_j^z (c_k + c_k^+)$$

is the electron–phonon interaction operator; a^+, b^+, c_k^+, and ω, ν, ω_k are the creation operators for photons, intramolecular vibrational modes, kth phonon mode and the corresponding frequencies, respectively; g, ξ, and λ_k are the interaction coefficients; $R_j^{\pm, z}$ are operators of the energy spin, which describe the jth two-level optical center ($j = 1, 2, \ldots, N$) and are identical to the Pauli spin matrices; and Planck's constant $\hbar$ is assumed equal to unity. Liouvillians Λ_f and Λ_a take into account incoherent interactions that result in the energy dissipation from the radiation field and excited optical impurities, respectively:

$$\Lambda_f X = \kappa([a(X, a^+)] + \text{H.c.}) + 2\kappa \bar{N}[a, [X, a^+]], \quad (2)$$

$$\begin{aligned} \Lambda_a X = \frac{1}{2}\sum_j \{ & \gamma_{12}([R_j^- X, R_j^+] + \text{H.c.}) \\ & + \gamma_{21}([R_j^+ X, R_j^-] + \text{H.c.})\} \qquad (3) \\ & + \Gamma(T)\sum_j ([R_j^z \rho, R_j^z] + \text{H.c.}), \end{aligned}$$

where κ is the decay constant of the radiation field caused by the irreversible escape of photons outside an elongated sample of length l (we assume that $\kappa = c/l$, where c is the speed of light); $\bar{N} = [\exp(\omega/k_B T) - 1]^{-1}$; T is the sample temperature; and $\gamma_{12}(\gamma_{21})$ is the rate of transition from the ground (excited) to the excited (ground) electronic state of the impurity. The latter, purely dephasing term in (3) describes the electron–phonon interaction, which is bilinear in the phonon variables. For example, if dephasing is predominantly determined by a narrow part of the phonon spectrum, then, in the second-order of perturbation theory in the electron–phonon coupling constant χ, we have [9]

$$\Gamma(T) = \chi^2 \bar{n}(\omega_0)[\bar{n}(\omega_0) + 1], \quad (4)$$

where ω_0 is an effective Einstein frequency of the crystal lattice and $\bar{n}(\omega_0) = [\exp(\omega_0/k_B T) - 1]^{-1}$. The Liouvillian Λ_l describes the decay of the intramolecular vibrational mode

$$\Lambda_l X = \alpha([bX, b^+] + \text{H.c.}) + 2\alpha\bar{n}[b, [X, b^+]], \quad (5)$$

where α is the decay constant and $\bar{n} = [\exp(\nu/k_B T) - 1]^{-1}$.

The form and derivation of Liouvillians (2), (3), and (5) that take the dissipative processes into account are presented, for example, in papers [9, 10].

The model of an ensemble of resonance optical centers interacting with the radiation field in a crystal considered here is quite general. In particular, a practically important and often encountered case of the impurity systems whose optical spectrum consists only of one zero-phonon line and a structureless phonon sideband, which can produce lasing, is a simplified version of this model.

To derive the master equation describing the laser action, one should adiabatically exclude the variables related to the polarization ($\propto R_j^\pm$), inverse population ($\propto R_j^z$), amplitudes of local vibrations (b, b^+), and lattice modes (c_k, c_k^+), because these quantities rapidly vary in time, whereas the amplitudes a and a^+ of the light field slowly vary in time. Using the methods of nonequilibrium statistical mechanics [11, 12] and taking into account that the orders of magnitude of the operators for a high-Q cavity satisfy the inequalities $O(\Lambda_f) \ll O(L_{af}) \ll \mathrm{O}(\Lambda_\mathrm{a})$, we obtain, in the second-order approximation in L_{af} (or H_{af}), the master equation for the density matrix $\sigma = \mathrm{Sp}_a \mathrm{Sp}_l \mathrm{Sp}_L(\rho)$ of the field in the cavity

$$\begin{aligned} \dot{\sigma}(t) = \Lambda_f \sigma(t) - \int_0^t d\tau \mathrm{Sp}_a \mathrm{Sp}_l \mathrm{Sp}_L \qquad (6) \\ \times \{\tilde{L}_{af}(t) \exp[\Lambda_a(t - \tau)] \tilde{L}_{af}(\tau) \rho_a \rho_l \rho_L \sigma(\tau)\}, \end{aligned}$$

where

$$\tilde{L}_{af}(t) = \exp(i\tilde{L}t) L_{af} \exp(-i\tilde{L}t),$$

$$\tilde{L} = L_a + L_f + L_l + L_{al} + L_L + L_{aL} + i\Lambda_l,$$

$$\rho_{l,L} = \frac{\exp(-H_{l,L}/k_B T)}{\mathrm{Sp}_{l,L}[\exp(-H_{l,L}/k_B T)]},$$

$$\rho_a = \Pi_j \left(\frac{1}{2} + \sigma_0 R_j^z\right),$$

$\sigma_0 = (\gamma_{12} - \gamma_{21})/(\gamma_{12} + \gamma_{21})$ is the inverse population of the levels ($-1 \le \sigma_0 \le 1$), and $\gamma_{12}(\gamma_{21})$ is the rate of transition from the ground (excited) electronic state to the excited (ground) state.

After some transformations, Eq. (6) can be rewritten in the form

$$\begin{gathered}\dot{\rho}(t) = \Lambda_f\sigma(t) - g^2\int_0^t d\tau\exp[-\gamma_\perp(t-\tau)] \\ \times\{\exp[i\Delta(t-\tau)]\langle U^+(t)U(\tau)\rangle_{l,L} \qquad (7) \\ \times[\mathrm{Sp}_a(R^+R^-\rho_a)(aa^+\sigma(\tau) - a^+\sigma(\tau)a) \\ + \mathrm{Sp}_a(R^-R^+\rho_a)(\sigma(\tau)a^+a - a\sigma(\tau)a^+)] + \text{H.c.}\},\end{gathered}$$

where $\Delta = \varepsilon - \omega$, $\gamma_\perp = (\gamma_{12} + \gamma_{21})/2$,

$$R^{\pm,z} = \sum_j R_j^{\pm,z}, \quad \langle\ldots\rangle_{l,L} = \mathrm{Sp}_l\mathrm{Sp}_L\{\ldots\rho_l\rho_L\},$$

$$\begin{gathered}\langle U^+(t)U(\tau)\rangle_{l,L} = \left\langle \overset{(+)}{T}\left(\exp\left[i\int_0^t dt_1F(t_1)\right]\right)\right\rangle \qquad (8) \\ \times\left\langle \overset{(-)}{T}\left(\exp\left[-i\int_0^\tau dt_2F(t_2)\right]\right)\right\rangle_{l,L},\end{gathered}$$

$$F(t) = V(t)\left[\xi\nu(b+b^+) + \sum_k\lambda_k(c_k + c_k^+)\right]V^+(t), \quad (9)$$

$$V(t) = \exp[i(H_l + H_L + H_{al} + H_{aL} + i\Lambda_l)t]. \quad (10)$$

The correlation function (8) can be calculated as in [9, 10]. Then, we have

$$\begin{gathered}ln\langle U^+(t)U(t')\rangle_{l,L} \approx \xi^2\{-2(2\bar{n}+1) \\ -[(2\bar{n}+1)\alpha + i\nu](t-t') \\ + [\bar{n}\exp[i\nu(t-t')] + (\bar{n}+1)\exp[-i\nu(t-t')]] \\ \times\exp(-\alpha|t-t'|)\} + \sum_k\left(\frac{\lambda_k}{\omega_k}\right)^2 \qquad (11) \\ \times[-2(\bar{n}_k+1) - i\omega_k(t-t') + \bar{n}_k\exp[i\omega_k(t-t')] \\ + (\bar{n}_k+1)\exp[-i\omega_k(t-t')]],\end{gathered}$$

where $\bar{n}_k = [\exp(\omega_k/k_BT) - 1]^{-1}$. The coefficients $i\xi^2\nu$ and $i\sum_k\lambda_k^2/\omega_k$ in terms linear in $(t - t')$ in the right-hand side of expression (11) will be omitted below, assuming that they have already been included into the renormalized values of energies of the ground and excited electronic states.

By passing in (7) to the P-representation of Glauber–Sudarshan for σ,

$$\sigma(t) = \int d^2\alpha P(\alpha,\alpha^*,t)|\alpha\rangle\langle\alpha|,$$

where $a|\alpha\rangle = \alpha|\alpha\rangle$, we obtain, in the Markov approximation, instead of (7) the Fokker–Planck equation

$$\begin{gathered}\dot{P}(\alpha,\alpha^*,t) = \left\{\kappa\left(\frac{\partial}{\partial\alpha}\alpha + \frac{\partial}{\partial\alpha^*}\alpha^*\right) - \frac{g^2N\sigma_0}{\gamma_\perp}\right. \\ \times\left(\int_0^\infty d\tau\gamma_\perp\exp[(i\Delta-\gamma_\perp)\tau]\right. \\ \left.\times\langle U^+(t)U(\tau)\rangle_{l,L}\frac{\partial}{\partial\alpha^*}\alpha^* + \text{c.c.}\right) \qquad (12) \\ + \left[2\kappa\bar{N} + \frac{g^2N(1+\sigma_0)}{2\gamma_\perp}\right. \\ \left.\times\left(\int_0^\infty d\tau\gamma_\perp\exp[(i\Delta-\gamma_\perp)\tau]\langle U^+(t)U(\tau)\rangle_{l,L} + \text{c.c.}\right)\right] \\ \left.\times\frac{\partial^2}{\partial\alpha\partial\alpha^*}\right\}P(\alpha,\alpha^*,t).\end{gathered}$$

Equation (12), which is linear in the field amplitude, describes the laser action below the threshold and allows us to determine the lasing threshold by equating the loss rate κ to the gain G:

$$\begin{gathered}\kappa = G, \\ G \equiv \frac{1}{2}\frac{g^2N\sigma_0}{\gamma_\perp} \qquad (13) \\ \times\left(\int_0^\infty d\tau\gamma_\perp\exp[(i\Delta-\gamma_\perp)\tau]\langle U^+(t)U(\tau)\rangle_{l,L} + \text{c.c.}\right).\end{gathered}$$

The coefficient at the second-order derivative with respect to the amplitude variable in (12)

$$\begin{gathered}2\kappa\bar{N} + \frac{g^2N(1+\sigma_0)}{2\gamma_\perp}\left(\int_0^\infty d\tau\gamma_\perp\right. \qquad (14) \\ \left.\times\exp[(i\Delta-\gamma_\perp)\tau]\langle U^+(t)U(\tau)\rangle_{l,L} + \text{c.c.}\right) \equiv D\end{gathered}$$

is the diffusion constant.

Recall that $\bar{N} = [\exp(\omega/k_BT) - 1]^{-1}$ in (14) is the occupation number of the photon mode at temperature T and N is the number of impurity particles.

3. THRESHOLD CONDITIONS OF LASING AT THE VIBRONIC REPLICAS OF THE PURELY ELECTRONIC LINE

By using expression (11) for the correlation function $\langle U^{+}(t)U(\tau)\rangle_{l,\,L}$, the gain can be represented in the explicit form

$$\begin{aligned} G &= \frac{g^2 N\sigma_0}{\gamma_\perp}\exp[-\xi^2(2\bar{n}+1)] \\ &\times \sum_{p=-\infty}^{\infty}\sum_{r,\,s=0}^{\infty}\prod_k \xi^{2(r+s)}(\bar{n}+1)^r\bar{n}^s\gamma_\perp \\ &\times[\gamma_\perp+(r+s)\alpha+\xi^2(2\bar{n}+1)\alpha] \\ &\times\{r!s![\varepsilon-\omega-(r-s)\nu-p\omega_k]^2 \\ &+[\gamma_\perp+(r+s)\alpha+\xi^2(2\bar{n}+1)\alpha]^2\}^{-1} \\ &\times\exp[-w(T)]I_p\left(2\frac{\lambda_k^2}{\omega_k^2}\sqrt{\bar{n}_k(\bar{n}_k+1)}\right)\left(\frac{\bar{n}_k+1}{\bar{n}_k}\right)^{p/2}, \end{aligned} \tag{15}$$

where exp[$-w(T)$] is the Debye–Waller factor, which is related to the total Stokes loss

$$\sum_k\left(\frac{\lambda_k}{\omega_k}\right)^2(2\bar{n}_k+1)$$

caused by phonons, and exp[$-\xi^2(2\bar{n}+1)$] is the Debye–Waller factor related to local vibrations; I_p is the Bessel function of the imaginary argument of the first kind; ε is the purely electronic transition energy; ω and κ are the frequency and the loss rate of a photon mode in the cavity, respectively; ν and α are the frequency and the decay of a local (intramolecular) vibration; ω_k is the frequency of the kth phonon; and

$$\gamma_\perp = \gamma_\parallel/2+\Gamma(T) \tag{16}$$

is the relaxation rate of the induced polarization, where $\gamma_\parallel = \gamma_{12}+\gamma_{21}$.

Consider first of all several particular cases concerning the threshold condition (13), (15). First, in the absence of the adiabatic interaction of the ground and excited electronic states of the optical centers with local vibrations and phonons ($\xi \approx 0$, $\lambda_k \approx 0$), the threshold condition (13), (15) coincides with the usual one [13],

$$\kappa = \frac{g^2 N\sigma_0}{\gamma_\perp}. \tag{17}$$

Second, in the case of a very strong coupling of optical centers with the vibrational modes ($\xi^2 \gg 1$, $\lambda_k^2/\omega_k^2 \gg 1$), as well as at high temperatures, gain G becomes so small that equality (13) cannot be satisfied, and lasing is impossible. Third, for the purely electronic transition ($r = s = p = 0$), in the case of exact resonance ($\varepsilon - \omega = 0$), we have the threshold condition [14]

$$\kappa = \frac{g^2 N\sigma_0}{\tilde{\gamma}_\perp}\exp[-\tilde{w}(T)], \tag{18}$$

where $\tilde{w}(T) = w(T)+\xi^2(2\bar{n}+1)$ and $\tilde{\gamma}_\perp = \gamma_\perp = \gamma_\perp + \xi^2(2\bar{n}+1)\alpha$. The threshold condition (18) corresponds to the case when lasing occurs within the zero-phonon line, i.e., photons interact with the impurity crystal in such a way that the vibrational states of both local modes and phonons do not change.

More typical is the situation when the interaction of light with an impurity crystal does not change the vibrational state of the crystal lattice ($p = 0$), but the resonance condition $\omega = \varepsilon - (r-s)\nu$ is simultaneously satisfied, so that one of the vibrational replicas of the purely electronic line is observed, which corresponds to the vibronic transition from the sth vibrational level in the excited electronic state (when s vibrational quanta are annihilated) to the rth vibrational level of the ground electronic state (when r vibrational quanta are created). The lasing threshold for such an individual vibronic transition can be written in a rather simple form

$$\begin{aligned} \kappa &= \frac{g^2 N\sigma_0}{\gamma_\perp}\exp[-\xi^2(2\bar{n}+1)] \\ &\times\frac{\gamma_\perp\xi^{2(r+s)}(\bar{n}+1)^r\bar{n}^s}{r!s![\gamma_\perp+\xi^2(2\bar{n}+1)\alpha+(r+s)\alpha]} \\ &\times\exp[-w(T)]\prod_k I_0\left(2\frac{\lambda_k^2}{\omega_k^2}\sqrt{\bar{n}_k(\bar{n}_k+1)}\right) \end{aligned} \tag{19}$$

which is convenient for estimates.

Expression (19) can be used for determining the lasing thresholds for any of the vibronic transitions, neglecting their interrelation (in the general case, expressions (13) and (15) should be used). If sufficiently high gains are simultaneously achieved for a whole series of vibronic transitions, so that the corresponding threshold conditions are realized, then, due to a distribution of the field modes in the cavity (because the emission line has a finite width), the resonance can be achieved at several transitions rather than at one transition, and tunable lasing becomes possible.

The product of the zero-order Bessel functions in expression (19) changes with temperature (or $\bar{n}_k$) slower than the Debye–Waller factor exp[$-w(T)$], and for $(\lambda_k/\omega_k)^2\sqrt{\bar{n}_k(\bar{n}_k+1)} \ll 1$, this product can be set equal to unity because $I_0(x) \approx 1$ for $x \ll 1$. Since the intramolecular frequency is much higher than the phonon frequency, the temperature dependence of the gain for different vibronic transitions [see the right-hand side of Eq. (19)] will mainly be determined by the Debye–Waller factor exp[$-w(T)$]. It also follows from expression (19) that as the strength of interaction (i.e., the Stokes loss ξ^2) of the optical centers with intramolecular vibrations increases, the maximum gain is achieved for the longer-wavelength vibronic transitions.

4. LASING IN THE REGION OF THE STRUCTURELESS PHONON SIDEBAND

If the optical spectrum of a crystal consists of only a zero-phonon line and a structureless phonon sideband, i.e., the coupling of electronic states with high-frequency local vibrations is negligible (or the latter are absent at all), then we should set $\xi = \nu = \alpha = 0$ in the above expressions and, in particular, in expression (15) for the gain. For the estimates that will be performed below, it is convenient to replace the phonon spectrum of the crystal by an effective frequency ω_0 (the Einstein approximation). Moreover, if lasing takes place within a narrow region of the phonon sideband, the condition of the exact resonance

$$\varepsilon - \omega - p\omega_0 = 0 \tag{20}$$

can be reasonably imposed. Then, the corresponding gain, taking into account expression (15), can be rewritten in the form

$$G(\omega = \varepsilon - p\omega_0;\, \lambda_0;\, T) = \frac{g^2 N\sigma_0}{\gamma_\perp(T)} f(\lambda_0;\, T), \tag{21}$$

$$\begin{aligned} f(\lambda_0;\, T) &= \exp\left\{-\frac{\lambda_0^2}{\omega_0^2}[2\bar{n}(\omega_0) + 1]\right\} \\ &\times I_p\left(2\frac{\lambda_0^2}{\omega_0^2}\sqrt{\bar{n}(\omega_0)(\bar{n}(\omega_0) + 1)}\right)\left(\frac{\bar{n}(\omega_0) + 1}{\bar{n}(\omega_0)}\right)^{p/2}. \end{aligned} \tag{22}$$

The arguments of the function G (21) show that the gain depends on the electron–phonon coupling constant λ_0 and the temperature T of the crystal lattice, as well as that an elementary event of emission of a quantum of the electromagnetic field of frequency ω during lasing is accompanied by the creation of p phonons with frequency ω_0 due to the resonance condition (20).

The linear Fokker–Planck equation (12) is sufficient for the determination of the lasing threshold and the diffusion constant, whereas the real amplitude of the laser-field oscillations can be found only with the help of nonlinear theory, i.e., taking into account in the derivation of the master equation for the density matrix the terms at least up to the fourth order inclusive in the interaction L_{af} (or H_{af}). However, if the resonance condition (20) is satisfied, the field amplitude can be found by modifying Eqs. (14)–(16) from paper [14] by making the substitution $g\exp[-S(T)/2] \longrightarrow gf^{1/2}(\lambda_0;\, T)$. As a result, polarization P, inversion Δ, amplitude a, and the average number of phonons $n = \langle a^+ a\rangle = a^2$ will be determined, in the average field approximation, by the set of equations

$$\begin{aligned} \dot{P} &= 2ga\Delta f^{1/2} - \gamma_\perp P, \\ \dot{\Delta} &= -2gaPf^{1/2} - \gamma_\parallel(\Delta - \sigma_0 N/2), \\ \dot{a} &= 2g\Delta f^{1/2} - \kappa a, \\ \dot{n} &= 2gaPf^{1/2}. \end{aligned} \tag{23}$$

Assuming $\dot{P} = \dot{\Delta} = 0$, we find the evolution equation for the field amplitude generated in the region of the phonon sideband at the frequency $\omega = \varepsilon - p\omega_0$:

$$\dot{a} = (G - \kappa)a - \frac{4\kappa}{N\gamma_\parallel\sigma_0}Ga^3. \tag{24}$$

For $G > \kappa$, we can find from (24) the average number of photons in the stationary regime ($\dot{a} = 0$):

$$n = a^2 = \frac{N\gamma_\parallel}{4\kappa}(\sigma_0 - \sigma_0\kappa G^{-1}). \tag{25}$$

In this case, the laser radiation intensity is

$$I = \omega\dot{n} = 2g\omega aPf^{1/2} = 2\kappa n, \tag{26}$$

or, after the substitution of (25) into (26),

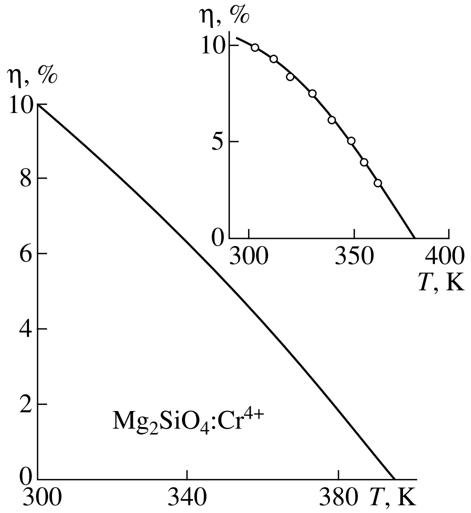
$$I = \frac{N\gamma_\parallel}{2}\left[\sigma_0 - \frac{\kappa\gamma_\perp(T)}{g^2 N}f^{-1}(\lambda_0;\, T)\right], \tag{27}$$

where the relation (21) between G and f is taken into account.

5. APPLICATION OF THE THEORY TO THE DESCRIPTION OF LASING IN A $Mg_2SiO_4:Cr^{4+}$ SINGLE CRYSTAL

The theory of a vibronic laser developed above can be compared with experiments on lasing observed in the phonon sideband in the optical spectrum of a $Mg_2SiO_4:Cr^{4+}$ forsterite single crystal. This laser produces stable quasi-continuous emission in the near IR region with an average output power above 1 W [3, 4]. The pump frequency, the purely electronic transition frequency, and the lasing frequency were 9276, 9158, and 8097 cm^{-1}, respectively. Because no decrease in the quantum yield of luminescence was observed in the temperature range between 300 and 400 K of interest for applications, i.e., nonradiative relaxation was absent, this means that the vibrational potential curves of the atoms in the configuration space in the ground and metastable electronic states should not intersect (which was noted in [4]).

In [4], the temperature dependence of the energy efficiency η (the ratio I/I_p between the output and pump powers) of a $Mg_2SiO_4:Cr^{4+}$ laser was measured. Expression (27) can be transformed to describe the experimental situation. Indeed, let the excess over the lasing



Temperature dependence of the energy efficiency of lasing η in the region of the phonon sideband. The insert shows the experimental temperature dependence of the efficiency of a chromium-doped forsterite laser [4].

threshold at some initial temperature T_0 be Δ and η equal η_0 (in percent), then, using (27), we can write

$$\eta(T) = \frac{I(T)}{I_p} = \eta_0\Delta \times\left[1-\frac{1}{\Delta}\frac{\bar{n}(\omega_0, T)(\bar{n}(\omega_0, T)+1)}{\bar{n}(\omega_0, T_0)(\bar{n}(\omega_0, T_0)+1)}\frac{f(\lambda_0, \omega_0; T_0)}{f(\lambda_0, \omega_0; T)}\right]. \tag{28}$$

Expression (28) takes into account that, at sufficiently high temperatures, for example, room temperature, the rate $\gamma_\parallel$ of energy relaxation is much lower than the dephasing contribution $\Gamma(T)$ into (16), so that in fact we have

$$\gamma_\perp(T) \approx \chi^2\bar{n}(\omega_0, T)[\bar{n}(\omega_0, T)+1] \tag{29}$$

(see expressions (4) and (16)). The function $f(\lambda_0, \omega_0; T)$ is defined in (22). Thus, given the reference points T_0, Δ, and η_0, we can express the energy efficiency of laser (28) only in terms of the crystal-lattice parameters, such as the effective frequency ω_0, temperature T, and the electron–phonon coupling strength λ_0. However, note that the effective frequency ω_0 in expression (29) for $\gamma_\perp$ or in expression (28) for $\bar{n}$ can be in the general case different from this quantity appearing as an argument of function f in (28).

The excess over the lasing threshold at $T_0 = 300$ K in the experiment [4] was $\Delta = 2.2$ and the value of η_0 was 10%. The temperature was measured with an external heater. The experimental results are presented in the insert in the figure. One can see that the energy efficiency of the laser decreases by 25% at $T < 330$ K and lasing disappears at 380 K.

In the calculation of $\eta(T)$ from expression (28), we will assume that ω_0 is equal to the difference between the purely electronic transition frequency 9158 cm^{-1} and the lasing frequency 8097 cm^{-1}, i.e., to 1061 cm^{-1}, and, because in this case only one-quantum transitions occur, we will set $p = 1$ in expression (22) for f and $\lambda_0^2/\omega_0^2 = 1$. The result of the calculations is presented in the figure. The theory adequately describes the temperature dependence $\eta(T)$, both qualitatively and quantitatively, despite the fact that the phonon spectrum of the crystal was represented by a very simplified one-oscillator model.

6. CONCLUSIONS

In this paper, some aspects of the operation of a vibronic laser below the threshold have been studied. It has been shown that the threshold conditions of lasing at vibronic transitions depend on the strength of the interaction of optical centers with a local (intramolecular) vibration that forms the vibronic spectrum—a series of vibrational replicas of the purely electronic line. These threshold conditions also depend on the crystal lattice temperature. As the Stokes loss per local vibration increases, the maximum value of the gain is achieved at the longer wavelength vibronic transitions, while the increase in the crystal temperature reduces the gain at all the vibronic transitions virtually in the same degree. The theory can readily be generalized to the spectrum containing a structureless phonon sideband and well agrees with the experimental temperature dependence of the output of power of a $Mg_2SiO_4:Cr^{4+}$ forsterite laser.

REFERENCES

1. *Tunable Solid State Lasers for Remote Sensing*, Ed. by R. L. Byer, E. K. Gustafson, and R. Trebino (Springer-Verlag, Berlin, 1985).
2. *Tunable Lasers*, Ed. by L. F. Mollenauer and J. C. White (Springer-Verlag, Berlin, 1987).
3. V. Petricevic, S. K. Gayen, and R. R. Alfano, Opt. Lett. **14**, 612 (1989).
4. M. G. Livshits, Ya. I. Mishkel', and A. A. Tarasov, Kvantovaya Élektron. (Moscow) **19**, 496 (1992).
5. E. Sigmund and P. Schwendimann, Opt. Acta **32**, 281 (1985).
6. P. Schwendimann, E. Sigmund, and K. Zeile, Phys. Rev. A **37**, 3018 (1988).
7. A. P. Saĭko and V. S. Kuz'min, Dokl. Akad. Nauk BSSR **34**, 998 (1990).
8. A. P. Saĭko and V. S. Kuz'min, Laser Phys. **7**, 384 (1997).
9. V. S. Kuz'min and A. P. Saĭko, Teor. Mat. Fiz. **83**, 291 (1990).
10. A. P. Saĭko, *Cooperative Nonlinear Optical Phenomena* (Belarussk. Gos. Univ., Minsk, 1999).
11. F. Haake, Springer Tracts Mod. Phys. **66**, 98 (1973).
12. F. Shibata and T. Arimitsu, J. Phys. Soc. Jpn. **49**, 891 (1980).
13. H. Haken, *Light*, Vol. 2: *Laser Light Dynamics* (North-Holland, Amsterdam, 1985; Mir, Moscow, 1988).
14. A. P. Saĭko and V. S. Kuz'min, Kvantovaya Élektron. (Moscow) **15**, 1249 (1988).

Translated by M. Sapozhnikov

Journal of Experimental and Theoretical Physics, Vol. 92, No. 2, 2001, pp. 228–234.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 119, No. 2, 2001, pp. 264–272.
Original Russian Text Copyright © 2001 by Vaulina, Khrapak.

PLASMA, GASES

Simulation of the Dynamics of Strongly Interacting Macroparticles in a Weakly Ionized Plasma

O. S. Vaulina and S. A. Khrapak

*Institute of High Energy Density, IVTAN (Institute of High Temperatures) Scientific Association,
Russian Academy of Sciences, Moscow, 127412 Russia*
e-mail: skhrapak@mail.ru

Received August 7, 2000

Abstract—The method of molecular dynamics is used to study the dynamic behavior of a nonideal system of particles interacting through screened Coulomb potential. The behavior of the self-diffusion coefficient of particles as a function of the nonideality parameter is investigated. The conditions of the crystallization of such a system are discussed, as well as the possibility of using the crystallization criterion, based on the dynamic characteristics of the system. *© 2001 MAIK “Nauka/Interperiodica”.*

1. INTRODUCTION

A dusty plasma consists of an ionized gas containing particles of condensed matter (dust particles) that either form spontaneously in the plasma as a result of various processes or are introduced into the plasma from the outside. This plasma is sometimes referred to as colloidal plasma or plasma with a condensed dispersed phase. Dust particles acquire an electric charge in the plasma and interact with the electric and magnetic fields, and the Coulomb interaction between particles may bring about a strong nonideality of a system of dust particles.

One of the reasons for the current interest in the physics of dusty plasma lies in the experimentally supported possibility of the crystallization of a strongly nonideal system of dust particles in plasma of different types. By now, the formation of ordered structures of dust particles has been observed in an rf discharge plasma [1–4], in the positive column of a dc glow discharge [5, 6], and in an atmospheric-pressure thermal plasma [7, 8]. Plasma-dust crystals exhibit a variety of unique properties and provide an indispensable tool in studying both the properties of strongly nonideal plasma and the fundamental properties of crystals.

It is generally agreed that dust particles in plasma interact with one another through screened Coulomb potential (Yukawa potential),

$$U(r) = Z_d^2 e^2 \exp(-r/\lambda_D)/r, \tag{1}$$

where Z_d is the charge number of the dust particles, and λ_D is the screening distance (the respective Debye length). The nonideality of dusty plasma is usually characterized by the nonideality parameter Γ_d equal to the ratio of the potential energy of Coulomb interaction between neighboring particles to their kinetic energy characterized by the particle temperature T_d,

$$\Gamma_d = Z_d^2 e^2 / bT_d, \tag{2}$$

where $b = n_d^{-1/3}$ characterizes the mean distance between particles. The dust particle charge Z_d in different plasmas may be very high. For example, in a low-pressure gas-discharge plasma, the charge may be estimated at

$$Z_d \sim -aT_e/e^2,$$

which, for the particle radius $a \sim 1$ μm and electron temperature $T_e \sim 1$ eV, gives $Z_d \sim -10^3$ elementary charges. Therefore, it is much easier to attain the nonideality of a subsystem of dust particles than the nonideality of an electron-ion subsystem, in spite of the fact that the particle concentration is usually much lower than the electron and ion concentrations.

The properties of nonideal systems of particles interacting through potential (1) are studied intensively using numerical simulation. Considerable attention is given to the phase diagrams of such systems (including the crystallization conditions) [9–12], to the behavior of particles in the field of various external forces [13–15], to the dynamics of the formation of ordered structures [8, 16], and to the dispersion properties of systems [17]. Two methods are usually employed in numerical calculations of this type, namely, the method of molecular dynamics and the method of Brownian dynamics. The method of molecular dynamics is based on the integration of invertible equations of motion of particles in view of interactions between them. The method of Brownian dynamics is based on the solution of stochastic Langevin equations for the evolution of the positions of particles. The former method enables one to adequately simulate the processes in atomic systems. On the other hand, in systems of dusty plasma (as well as in colloidal solutions), macroscopic particles are in a viscous medium, collisions with whose atoms or molecules play an important part.

The currently available results of the investigation of the dynamic behavior of particles interacting through Yukawa potential [9, 18, 19] fail to give a full picture of the dynamics of dust particles in plasma. For example, Robbins *et al.* [9] used the method of molecular dynamics and, thereby, ignored the interaction with the medium. Lowen *et al.* [19] studied the dynamic properties of a system in the vicinity of the crystallization point, and Lowen [18] studied the dynamics and structural properties of a two-dimensional system. In addition, in the last two studies, interest was given to particles suspended in a liquid (colloidal solutions) whose viscosity exceeded that of buffer gas in a dusty plasma by several orders of magnitude. The absence of quantitative results in theory makes difficult the analysis of recent results [20, 21] obtained during the investigation of nonideal dusty plasma systems.

This study is devoted to the systematic investigation of three-dimensional dynamics of strongly interacting macroparticles in a weakly ionized plasma using the method of molecular dynamics with due regard for the interaction between particles and medium. The dependence of the self-diffusion coefficient of particles on the force of interaction between them and on the viscosity of the medium is studied. The problem is treated of the possibility of using the dynamic criterion of melting suggested by Lowen *et al.* [19]. The conditions of crystallization in a dusty plasma are discussed.

2. MODEL

In order to include the interaction between particles and the surrounding medium, we used the approach previously employed to investigate the dynamic processes in a dusty plasma [8, 22, 23]. The interaction with the medium is included by way of introducing into the equation of motion the so-called Langevin force represented in the form of two terms, of which one describes systematic friction on the part of the medium, and the other term describes the momentum transfer upon individual collisions with atoms or molecules of the medium.

The noninvertible normalized equations of motion of dust particles, projected onto the Cartesian coordinate axis x, have the form

$$\dot{X}_k = (4\pi\Gamma_d)^{-1/2} V_k, \tag{3}$$

$$\dot{V}_k = \sum_{j \neq k} (\Gamma_d/4\pi)^{1/2} \frac{1 + KR_{kj}}{R_{kj}^3} \times \exp(-KR_{kj})(\mathbf{R}_{jk} \cdot \mathbf{e}_x) - \theta V_k + \sqrt{2\theta}\,\xi(\tau), \tag{4}$$

where X_k and V_k denote the dimensionless coordinate and velocity of the kth dust particle, respectively; τ is the dimensionless time; $\mathbf{R}_{kj} = \mathbf{R}_k - \mathbf{R}_j$ ($R_{kj} = |\mathbf{R}_{kj}|$); $\mathbf{e}_x$ is the unit vector in the direction of the x axis; $K = b/\lambda_D$ is the structure parameter; $\theta = \eta/\omega_{pd}$ is the ratio between the frequency of friction of the medium and the dusty plasma frequency; and $\xi(\tau)$ is the delta-correlated Gaussian white noise,

$$\langle \xi(\tau) \rangle = 0, \quad \langle \xi(\tau)\xi(\tau') \rangle = \delta(\tau - \tau'). \tag{5}$$

The unit of distance is provided by the mean distance between particles b, and the units of time and velocity are provided by the inverse dusty plasma frequency

$$\omega_{pd}^{-1} = (4\pi Z_d^2 e^2 n_d/m_d)^{-1/2}$$

and the thermal velocity of dust particles

$$v_{Td} = (T_d/m_d)^{1/2},$$

respectively.

The summation in the right-hand part of Eq. (4) is performed over all dust particles except for that being treated.

In the general case, according to Eqs. (3) and (4), the behavior of a system of dust particles is defined by three dimensionless parameters Γ_d, K, and θ. In the limit of $\theta = 0$, the standard method of molecular dynamics is realized. In this case, the system is characterized only by the nonideality parameter and by the structure parameter. Note that, in the standard method of molecular dynamics, because of free exchange between the potential and kinetic energies, a periodic renormalization of the particle velocities must be used in order to maintain constant the system temperature. At the same time, in the modified method of molecular dynamics, the system temperature defined by the parameters of Langevin force is maintained constant without requiring correction in the course of calculation.

The method of molecular dynamics employed by us consisted in solving Eqs. (3) and (4) in three dimensions for each particle and in analyzing the trajectories of particle motion. The calculation region was a cube of size $L = 5b$. In order to simulate a homogeneous spatially extended plasma, periodic boundary conditions were imposed. For reducing the calculation time, the potential of interaction between particles was cut off at distances exceeding $3.75b$, which is a standard procedure in the investigation of systems that interact through potential (1) with not too small a value of the structure parameter K. The cutoff of the interaction potential does not lead to a considerable error at $K \gtrsim 1$. At the same time, in simulating systems with $K \lesssim 1$, longer-range interactions must be included, which may be done with the aid of an appropriate algorithm [12].

The procedure of numerical experiment was as follows: at the initial moment of time, the particles were arranged in a random way within the calculation region; then, owing to interaction between them, the process of self-organization started; after reaching the configuration of the system of particles that was equilibrium for the preassigned values of Γ_d, K, and θ, the data on succes-

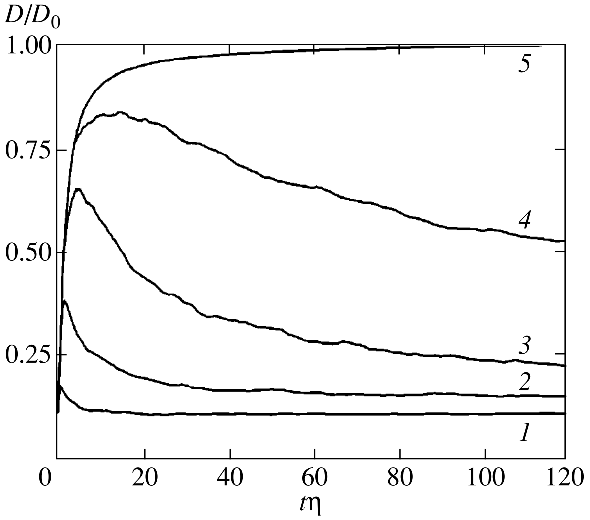


Fig. 1. The ratio of the diffusion coefficient D of interacting particles to the diffusion coefficient D_0 of noninteracting Brownian particles as a function of time (in terms of inverse deceleration time η^{-1}) for $K = 4.84$, $\Gamma_d = 554.7$, and different values of parameter θ equal to (*1*) 0.05, (*2*) 0.15, (*3*) 0.45), and (*4*) 1.33. Curve *5* represents the exact solution for noninteracting Brownian particles.

sive positions of particles were stored in the computer memory with a view to subsequent analysis. The step with respect to time was varied for different parameters of the system, but it did not exceed $\eta/20$ for a correct simulation of Brownian motion. The total count time was at least 2×10^5 steps with respect to time.

The coefficient of self-diffusion of particles is often used as a quantity characterizing the dynamic behavior of a system; this coefficient is defined as

$$D(t) = \langle \mathbf{r}(t) - \mathbf{r}(0) \rangle^2 / 6t,$$

where $\mathbf{r}(t)$ is the time-dependent trajectory of a single particle, and $\langle \ldots \rangle$ implies averaging over the ensemble. It is common practice to introduce two coefficients for interacting particles [19], namely, the short-term coefficient D_S and the long-term coefficient D_L, which are defined as

$$D_S = \lim_{t \to 0} D(t), \quad D_L = \lim_{t \to \infty} D(t).$$

Note that the diffusion behavior of particle motion shows up over periods at time that are longer compared with the inverse frequency of friction of the medium η^{-1}, and the limit of $t \longrightarrow 0$ must be taken to mean that the time is sufficiently short and, at the same time, long compared with η^{-1}. In the case of not too great a volume fraction of particles $b \gg a$, it is customary to assume [19] that the short-term coefficient of diffusion coincides with the regular coefficient of diffusion of noninteracting Brownian particles,

$$D_S = D_0 = v_{Td}^2 / \eta.$$

Because of interaction between particles, the value of D_L proves to be less than that of D_S. In the limiting case of crystalline structure, $D_L \longrightarrow 0$, because the displacement of particles located at the lattice sites is limited. In view of this, the ratio D_L/D_0 appears to be an adequate quantity for the investigation of the effect of nonideality on the dynamic behavior of a system of strongly interacting particles, up to the point of its crystallization.

3. SIMULATION RESULTS

The calculations were performed for the plasma and particle parameters typical of conditions of experiments in a gas-discharge dusty plasma. We will treat, for example, an argon plasma with the electron temperature $T_e \sim 1$ eV and room temperature of ions and neutrals (~0.03 eV) and with a dust component having the following characteristics: particle radius $a = 5$ μm, particle concentration $n_d \approx 5 \times 10^4$, density of the particle material $\rho = 1.5$ g/cm^3. Under these conditions and with the characteristic concentration of plasma $n_e \sim 10^8$ to 10^9 cm^{-3}, the value of the structure parameter is in the range of $K \sim 2$ to 7; in the pressure range of 3 to 300 Pa, parameter θ is in the range $\theta \sim 0.02$ to 2. Based on these estimates, we performed calculations for two values of the structure parameter $K = 2.42$ and $K = 4.84$ and for a set of values of parameter $\theta = 0.03$, 0.1, 0.3, 0.83, and 2.5 (for $K = 2.42$) and $\theta = 0.017$, 0.05, 0.15, and 0.45 (for $K = 4.84$). The nonideality parameter Γ_d was varied by varying the temperature of system T_d.

3.1. Time Dependence of the Diffusion Coefficient

Figure 1 gives the time dependence of the ratio of the diffusion coefficient D of interacting particles to the diffusion coefficient D_0 of noninteracting Brownian particles (in terms of inverse deceleration time η^{-1}) for $K = 4.84$, $\Gamma_d = 554.7$, and different values of parameter θ. Curve *5* represents the exact solution of a Langevin equation in the absence of interaction between particles,

$$\frac{D(t)}{D_0} = 1 - \frac{1 - \exp(-\eta t)}{\eta t}, \tag{6}$$

so that, for a long time compared with the inverse frequency of friction ($\eta t \gg 1$), $D(t) = D_0$, while in the case of short time ($\eta t \ll 1$) the ballistic behavior of particle motion shows up $\langle \Delta r^2(t) \rangle \approx 3 v_{Td} t^2$ and $D(t) \propto t$. In the presence of interaction, the behavior of $D(t)$ over short periods of time remains the same. Then it reaches a maximum; it is sound practice to use the value of the diffusion coefficient at this maximum in determining the short-term coefficient of self-diffusion D_S. Note that coefficient D_S proves to be less than D_0 and tends to the latter as the viscosity of the medium increases. With time, the diffusion coefficient tends to a constant value $D_L < D_S$ corresponding to the standard determination of the coefficient of self-diffusion of interacting

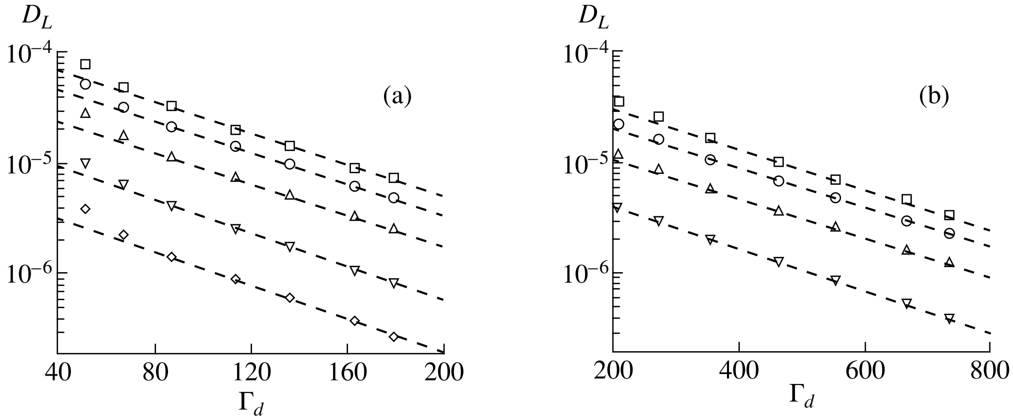


Fig. 2. The self-diffusion coefficient D_L (arbitrary units) as a function of the nonideality parameter Γ_d for (a) $K = 2.42$ and (b) $K = 4.84$: $\theta = 0.03$ (□), 0.1 (○), 0.3 (△), 0.83 (▽), and 2.5(◇) (for $K = 2.42$) and $\theta = 0.017$ (□), 0.05 (○), 0.15 (△), and 0.45 (▽) (for $K = 4.84$). Broken lines indicate the approximation results.

particles. In what follows, it is to the behavior of coefficient D_L that most attention is given, because it is this coefficient that may be determined experimentally during investigations of dusty plasma [20, 21]. Furthermore, its determination is not as arbitrary as the determination of D_S.

3.2. The Effect of Nonideality on the Behavior of D_L

Figures 2a and 2b give the coefficient D_L of self-diffusion of interacting particles as a function of the nonideality parameter Γ_d for two values of the structure parameter K and for different values of parameter θ. We will discuss some singularities of the behavior of D_L. First of all, note that, as the nonideality of the system increases, coefficient D_L decreases monotonically and, at some value of Γ_d, decreases abruptly by several orders of magnitude. This jump is apparently an indicator of the liquid–crystal phase transition in a system of dust particles. For $K = 2.4$, the diffusion coefficient experiences a jump at $\Gamma_d \sim 180$ and, for $K = 4.8$, at $\Gamma_d \sim 750$. In so doing, it turns out that, with a preassigned value of K and different values of θ, the diffusion coefficient experiences a jump at the same values of the nonideality parameter. Therefore, the conditions of crystallization of a system are defined by only two parameters Γ_d and K, and are independent of parameter θ.

Note that the behavior of the curve of the Γ_d—dependence of D_L, given in Fig. 2 in logarithmic coordinates, is virtually straight when approaching the crystallization point. This fact may be explained if, by analogy with molecular liquids, we represent the diffusion coefficient in the form

$$D_L = \frac{\Delta^2}{6\tau_0}\exp\left(-\frac{W}{T_d}\right),$$

where, in the case being treated, Δ characterizes the mean distance between particles, τ_0 is the characteristic period of particle vibrations in a "settled" state, and W is the energy barrier cleared by a particle upon transition between two neighboring "settled" states. It is reasonable to assume that $W/T_d \propto \Gamma_d$ and approximate the self-diffusion coefficient by a relation of the type of $D_L = A\exp(-B\Gamma_d)$. Here, A may also depend on Γ_d (as well as on K and θ). The approximation results are given in the table, and are shown by broken lines in Fig. 2. Note that, for a preassigned value of K, the values of parameter B are close to one another irrespective of θ.

3.3. Crystallization Curve in the Calculations by the Method of Molecular Dynamics

One of the fundamental problems arising during investigations of the properties of dusty plasma is that associated with the conditions of crystallization of the subsystem of dust particles. It is known from the simplest and best-studied model of nonideal plasma, i.e., the model of one-component plasma, that, at $\Gamma \gtrsim 1$, short-range order appears in a system, and, at $\Gamma \approx 106$, the one-component plasma crystallizes [24]. One can-

Results of the approximation of the Γ_d dependence of D_L by the relation $D_L = A\exp(-B\Gamma_d)$

K	θ	$A \times 10^{-5}$	$B \times 10^{-2}$	$B^* \times 10^{-2}$
2.4	0.03	12.9	1.6	2.9
	0.1	8.6	1.6	2.8
	0.3	4.6	1.6	2.9
	0.83	1.9	1.7	3.1
	2.5	0.65	1.7	3.1
4.8	0.017	7.1	0.42	3.0
	0.05	4.5	0.40	2.9
	0.15	2.4	0.40	2.9
	0.45	0.96	0.43	3.1

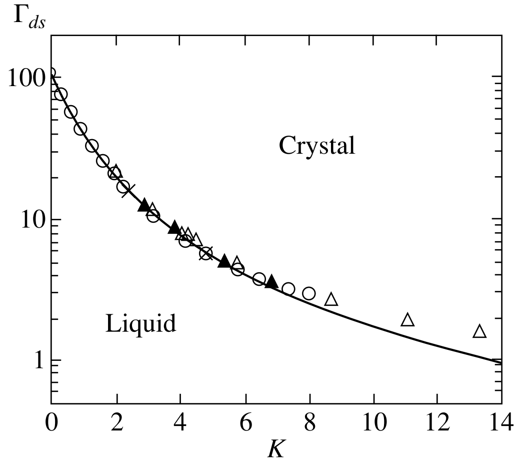


Fig. 3. The data on the calculation of the crystallization curve in Yukawa's model in the (K, Γ_{ds}) coordinates using the model of molecular dynamics: ▲ [10], △ [11], ○ [12]. The crosses correspond to the values of Γ_{ds}, at which a jump of the value of diffusion coefficient occurs in our calculations.

not claim the full adequacy of description of the properties of dusty plasma based on the model of one-component plasma, primarily, because of disregard of the screening effects. Nevertheless, it was reasoning based on the qualitative results of the model of one-component plasma that led Ikezi [25] to the conclusion about the possibility of the crystallization of the dust subsystem in a nonequilibrium gas-discharge plasma. He suggested that the conditions of the crystallization of dust particles in plasma may be provided by a condition of the form

$$\Gamma_{ds} = \frac{Z_d^2 e^2}{bT_d}\exp\left(-\frac{b}{\lambda_D}\right) \gtrsim 106. \tag{7}$$

The problem of the conditions of crystallization of a system of particles with Yukawa interaction was studied using numerical simulation by the method of molecular dynamics [10–12]. The simulation results demonstrate that the phase transition of the dust component from the liquid to crystalline state is described in terms of two dimensionless parameters Γ_d and K. In so doing, the value of parameter Γ_{ds} required for crystallization is not defined by a simple condition of the type of (7), but exhibits a more complex dependence on K.

Figure 3 gives the data on the calculation of the crystallization curve in Yukawa's model in the (K, Γ_{ds}) coordinates [10–12]. One can see that criterion (7) is invalid. At the same time, it proves possible to introduce the empirical condition of crystallization [26]

$$\Gamma_{ds}(1 + K + K^2/2) \approx 106, \tag{8}$$

which agrees, within the calculation error, with the numerical simulation results up to at least $K \lesssim 6$. The curve shown in Fig. 3 corresponds to exact equality in (8).

Marked with crosses in Fig. 3 are also the values of Γ_{ds}, at which the value of D_L in our calculations for $K = 2.4$ and $K = 4.8$ varies jumpwise. The obtained data correlate well with the results of other calculations. Thereby, an abrupt jump of the diffusion coefficient indeed serves as an indicator of the crystallization of a system. In addition, the calculation results do confirm that the conditions of crystallization are independent of parameter θ (viscosity of the medium). However, the boundary value of D_L/D_0, at which an abrupt jump occurs, depends on θ.

Note that normalization of a different type is often employed in determining Γ_d and K, namely, the Wigner–Seitz radius is used instead of the characteristic distance $n_d^{-1/3}$ between dust particles,

$$\rho = (4\pi n_d/3)^{-1/3}.$$

The thus determined nonideality parameter Γ'_d and parameter K' are related to Γ_d and K by simple relations

$$K' \approx K/1.612, \quad \Gamma'_d \approx 1.612\Gamma_d.$$

For example, the condition of crystallization of a one-component plasma in this notation will be provided by $\Gamma' \gtrsim 172$.

3.4. Normalization of the Nonideality Parameter and of the Dusty Plasma Frequency

Note that condition (8) may be obtained using a simple, though nonrigorous, approach [26]. Namely, we will treat a chain of particles interacting through potential (1). The characteristic frequency of the thermal vibrations of particles in such a chain may be written as [27]

$$\omega_0 \sim \omega_{pd}^* = \omega_{pd}\left(1 + K + \frac{K^2}{2}\right)^{1/2}\exp\left(-\frac{K}{2}\right). \tag{9}$$

The thermal displacement of particles relative to the position of equilibrium

$$\langle\delta u^2\rangle \sim T_d/m_d\omega_0^2$$

on the crystallization curve must satisfy the Lindemann condition,

$$\langle\delta u^2\rangle/b^2 = \text{const},$$

from which we derive that

$$\Gamma_d\left(1 + K + \frac{K^2}{2}\right)\exp(-K) = \Gamma_d^* = \text{const} \tag{10}$$

on the crystallization curve. We use the value, known from the model of one-component plasma ($K \longrightarrow 0$), as the value of constant to exactly derive condition (8).

The foregoing reasoning leads one to assume that parameters Γ_d^* and ω_{pd}^* may serve as similarity parameters for strongly nonideal systems interacting through potential (1). The simulation results support this assumption. Figure 4 gives the ratio D_L/D_0 as a function of the modified nonideality parameter Γ_d^* for a set of identical values of $\theta^* = \eta/\omega_{pd}^*$ for two different values of the structure parameter. It follows from this figure that the dependences of D_L/D_0 on Γ_d^* for a strongly nonideal system virtually coincide with the same values of parameter θ^* but with substantially different values of K. Therefore, the dynamic behavior of a system of strongly interacting particles is fully defined by the set of two parameters Γ_d^* and θ^*. Moreover, note that the approximation of the Γ_d^* dependence of D_L by the formula

$$D_L = A\exp(-B^*\Gamma_d^*)$$

produces almost identical values of B^* both for different values of the structure parameter and for different values of θ (see the table).

3.5. Crystallization Criteria

Different phenomenological criteria exist of the crystallization of a system of interacting particles, which found their application in the physics of dusty plasma as well. The best known is the Lindemann criterion [28], according to which the solid phase melts if the ratio of root-mean-square displacement of the particle-to-mean-interparticle distance reaches ~0.1 (this number may vary from ~0.05 to ~0.2 for different physical systems). Another criterion is provided by the constancy of the value of first maximum of the liquid structure factor [29], which reaches on the crystallization line a value of ~3.0 (this number likewise varies from 2.85 to 3.2 in different calculations). Treated as a simple crystallization criterion in terms of a binary correlation function is the ratio between the minimum and maximum values of the latter. The crystallization occurs with the value of this criterion is equal to 0.2. A simple dynamic criterion of melting, close in spirit to the Lindemann criterion, was suggested by Lowen *et al.* [19]. According to this criterion, crystallization occurs when the ratio D_L/D_0 of the diffusion coefficients of a dust particle decreases to a value of ~0.1.

The simulation results demonstrate that the boundary value of D_L/D_0, at which an abrupt jump of the self-diffusion coefficient occurs (crystallization), increases with θ^*. In so doing, note the fact that in all calculations, starting with $\theta^* \sim 0.5$ (and up to at least $\theta^* = 3.6$), the ratio D_L/D_0 in the vicinity of the crystallization point is constant and close to 0.1.

Thereby, the dynamic criterion of crystallization proves to be valid only for the case of not too low values of θ ($\theta^* \geqslant 0.5$). At the same time, for lower values of θ^*, the ratio D_L/D_0 at the crystallization point may be much lower and, therefore, the criterion is violated. In view of this, note that the experiments and calculations, performed by Lowen *et al.* [19] to substantiate the criterion being treated, were performed to fit the conditions of colloidal solutions, the viscosity (and, consequently, the value of θ^*) which is, as a rule, many orders of magnitude higher than the viscosity of buffer gas under conditions characteristic of experiments with dusty plasma. In view of this, the application of the dynamic criterion of melting to dusty plasma is restricted to the region of not too low pressures. Because, in investigating a dusty plasma under laboratory conditions, one often has to deal with the situation of $\omega_{pd} \gg \eta$, i.e., $\theta \ll 1$, the crystallization criterion treated herein must be used with caution.

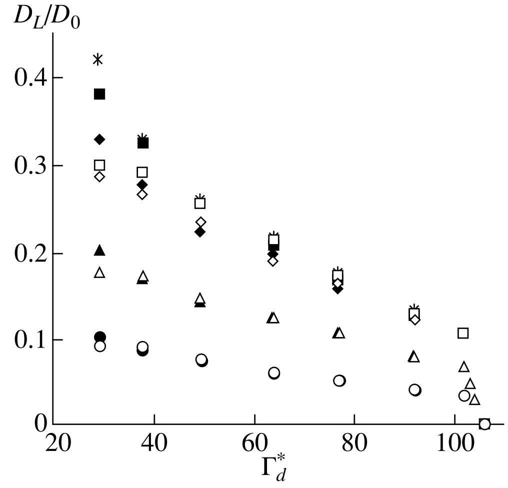


Fig. 4. The ratio D_L/D_0 as a function of the normalized nonideality parameter Γ_d^* : $\theta^* = 0.044$ (○), 0.13 (△), 0.4 (◇), 1.2 (□), and 3.6 (*) for $K = 2.4$; $\theta^* = 0.044$ (●), 0.13 (▲), 0.4 (◆), and 1.2 (■) for $K = 4.8$.

4. CONCLUSION

We used the modified method of molecular dynamics to investigate the dynamic behavior of dust particles interacting through Yukawa's potential. Main attention was given to the study of the dependence of the coefficient D_L of self-diffusion of dust particles on the force of the interaction between the particles under conditions of the plasma parameters typical of present-day laboratory experiments in a weakly ionized gas-discharge plasma.

The basic results may be formulated as follows. It has been found that the self-diffusion coefficient D_L experiences an abrupt jump (decreases by several orders of magnitude) at a certain value of the nonideality parameter Γ_d. This jump may serve as an indicator of phase transition in a system of dusty plasma. The value of the nonideality parameter, at which the jump (phase transition) occurs, depends substantially on the structure parameter K, but it is independent (within the accuracy of numerical experiment) of the value of parameter θ. It proves possible to use three independent dimensionless parameters Γ_d, K, and θ to construct two parameters Γ_d^* and θ^* which fully define the dynamic behavior of a strongly nonideal system of dust particles. It has been demonstrated that the Γ_d^* dependence of D_L may be well approximated by the expression

$$D_L = A\exp(-B^*\Gamma_d^*),$$

where $B^* \approx 3.0 \times 10^{-2}$. It is further indicated that the validity of the dynamic criterion of crystallization suggested by Lowen *et al.* [19] is, in reality, restricted to the region of not too low values of θ^* and, consequently, of not too low pressures ($\theta^* \gtrsim 0.5$).

The obtained results may be used both in analyzing the recent experimental results [20, 21] and in developing methods of the diagnostics of dusty plasma, based on determining the dynamic behavior of a system of dust particles.

REFERENCES

1. J. H. Chu and Lin I, Phys. Rev. Lett. **72**, 4009 (1994).
2. H. Thomas, G. E. Morfill, V. Demmel, *et al.*, Phys. Rev. Lett. **73**, 652 (1994).
3. A. Melzer, T. Trottenberg, and A. Piel, Phys. Lett. A **191**, 301 (1994).
4. Y. Hayashi and K. Tachibana, Jpn. J. Appl. Phys. **33**, L804 (1994).
5. V. E. Fortov, A. P. Nefedov, V. M. Torchinskiĭ, *et al.*, Pis'ma Zh. Éksp. Teor. Fiz. **64**, 86 (1996) [JETP Lett. **64**, 92 (1996)].
6. A. M. Lipaev, V. I. Molotkov, A. P. Nefedov, *et al.*, Zh. Éksp. Teor. Fiz. **112**, 2030 (1997) [JETP **85**, 1110 (1997)].
7. V. E. Fortov, A. P. Nefedov, O. F. Petrov, *et al.*, Pis'ma Zh. Éksp. Teor. Fiz. **63**, 176 (1996) [JETP Lett. **63**, 187 (1996)].
8. Y. K. Khodataev, S. A. Khrapak, A. P. Nefedov, and O. F. Petrov, Phys. Rev. E **57**, 7086 (1998).
9. M. O. Robbins, K. Kremer, and G. S. Grest, J. Chem. Phys. **88**, 3286 (1988).
10. E. J. Meijer and D. Frenkel, J. Chem. Phys. **94**, 2269 (1991).
11. M. J. Stevens and M. O. Robbins, J. Chem. Phys. **98**, 2319 (1993).
12. S. Hamaguchi, R. T. Farouki, and D. H. E. Dubin, Phys. Rev. E **56**, 4671 (1997).
13. H. Totsuji, T. Kishimoto, Y. Inoue, *et al.*, Phys. Lett. A **221**, 215 (1996).
14. V. V. Zhakhovskiĭ, V. I. Molotkov, A. P. Nefedov, *et al.*, Pis'ma Zh. Éksp. Teor. Fiz. **66**, 392 (1997) [JETP Lett. **66**, 419 (1997)].
15. O. M. Belotserkovskiĭ, I. E. Zakharov, A. P. Nefedov, *et al.*, Zh. Éksp. Teor. Fiz. **115**, 819 (1999) [JETP **88**, 449 (1999)].
16. A. P. Nefedov, O. F. Petrov, Ya. K. Khodataev, and S. A. Khrapak, Zh. Éksp. Teor. Fiz. **115**, 837 (1999) [JETP **88**, 460 (1999)].
17. D. Winske, M. S. Murillo, and M. Rosenberg, Phys. Rev. E **59**, 2263 (1999).
18. H. Lowen, J. Phys.: Condens. Matter **4**, 10105 (1992).
19. H. Lowen, T. Palberg, and R. Simon, Phys. Rev. Lett. **70**, 1557 (1993).
20. A. P. Nefedov, O. F. Petrov, S. A. Khrapak, *et al.*, Teplofiz. Vys. Temp. **36**, 141 (1998).
21. G. E. Morfill, H. M. Thomas, U. Konopka, and M. Zuzic, Phys. Plasmas **6**, 1769 (1999).
22. O. S. Vaulina, A. P. Nefedov, O. F. Petrov, and S. A. Khrapak, Zh. Éksp. Teor. Fiz. **115**, 2067 (1999) [JETP **88**, 1130 (1999)].
23. O. S. Vaulina, S. A. Khrapak, A. P. Nefedov, and O. F. Petrov, Phys. Rev. E **60**, 5959 (1999).
24. S. Ishimaru, Rev. Mod. Phys. **54**, 1017 (1982).
25. H. Ikezi, Phys. Fluids **29**, 1764 (1986).
26. O. S. Vaulina and S. A. Khrapak, Zh. Éksp. Teor. Fiz. **117**, 326 (2000) [JETP **90**, 287 (2000)].
27. F. Melandso, Phys. Plasmas **3**, 3890 (1996).
28. F. A. Lindemann, Z. Phys. **11**, 609 (1910).
29. J. P. Hansen and L. Verlet, Phys. Rev. **184**, 151 (1969).

Translated by H. Bronstein

Journal of Experimental and Theoretical Physics, Vol. 92, No. 2, 2001, pp. 235–245.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 119, No. 2, 2001, pp. 272–285.
Original Russian Text Copyright © 2001 by Pal', Serov, Starostin, Filippov, Fortov.

Non-Self-Sustained Discharge in Nitrogen with a Condensed Dispersed Phase

A. F. Pal'[a], A. O. Serov[a], A. N. Starostin[a], A. V. Filippov[a]*, and V. E. Fortov[b]

[a]*Russian Federation State Scientific Center Troitsk Institute for Innovation and Nuclear Research,*
Troitsk, Moscow oblast, 142190 Russia

[b]*Institute of Thermal Physics of Extremal States, Russian Academy of Sciences, Moscow, 127412 Russia*

**e-mail: fav@triniti.ru*

Received September 15, 2000

Abstract—A non-self-sustained discharge in nitrogen with a condensed dispersed phase is studied experimentally for the first time under atmospheric pressure at room temperature. It is shown that macroparticles strongly affect the current–voltage characteristics as well as the stability of the discharge process. A numerical simulation of dust particle charging in nitrogen is carried out at room temperature and cryogenic temperatures under continuous medium conditions. It is shown that a considerable charge is accumulated at macroparticles in the nitrogen beam plasma. As the gas temperature decreases, the charge of macroparticles in nitrogen increases, while in argon their charge decreases. For this reason, the Coulomb interaction parameter for dust particles in nitrogen increases strongly upon a transition from room to cryogenic temperature, while in argon this parameter decreases. It is also shown that the characteristic time of dust particle charging is shorter than 1 μs for a beam current density of 90 μA/cm^2, while the neutralization of the charge takes milliseconds. Possible mechanisms of the influence of the dust component on the characteristics of non-self-sustained discharge are considered. *© 2001 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

The interest in plasmas with a condensed dispersed phase (CDP) has increased during the last decade in connection with considerable advances in microtechnology and the progress in obtaining new materials [1, 2]. This interest is also due to the fact that such plasmas are the simplest real objects for studying self-organization processes which are of fundamental importance. Such a medium makes it possible to carry out experiments on convenient time and space scales and practically with the naked eye. The study of a non-self-sustained discharge in helium with a CDP revealed [3] that in a constant electric field, the current density in the discharge decreases upon an increase in the concentration of dust particles. The rate of the current density drop increases with the field. The theoretical model of a non-self-sustained discharge with a CDP constructed in [3] was based on the approximation of the orbital motion of electrons. The necessary condition for such an approximation to be applicable is that the molecular mode of electron motion is realized in the vicinity of a dust particle, where the quasi-neutrality of the plasma is violated. Under atmospheric pressure, the opposite case is realized in molecular gases, where the regime of a continuous medium takes place for the transport of charged particles. In this case, the description of the charging of macroparticles in the plasma of a non-self-sustained discharge is simplified since the joint solution of Boltzmann's equation for the electron energy distribution function (EDF) plus the continuity equations for charged particles and Poisson's equation for the electric field is not required. Consequently, it becomes possible to construct a consistent and comprehensive mathematical model of dust particle charging, which has been verified on other objects. For this reason, the study of a non-self-sustained discharge in molecular gases with a CDP is of considerable interest.

A non-self-sustained discharge in molecular gases is widely used for pumping CO$_2$, CO, and other high-power gas lasers. In such lasers, the erosion of electrodes or the polymerization of particles of the original gas or radicals formed in the discharge (e.g., the "laser snow" effect in the active medium of the XeCl laser [4]) leads to the formation of macroscopic particles in the working volume. Such particles can be responsible for a change in the characteristics of the discharge process and in the operation of the device proper (see [5] and the literature cited therein). For example, aerosol particles may lower the discharge stability. It was noted in [5] that the presence of dust particles in a high-pressure CO$_2$ laser and in excimer lasers correlated with instabilities such as discharge contraction or multiple streamers.

The present work is devoted to experimental and theoretical analyses of the effect of the dust component on the characteristics of a non-self-sustained discharge in nitrogen under atmospheric pressure, which is controlled by a fast electron beam.

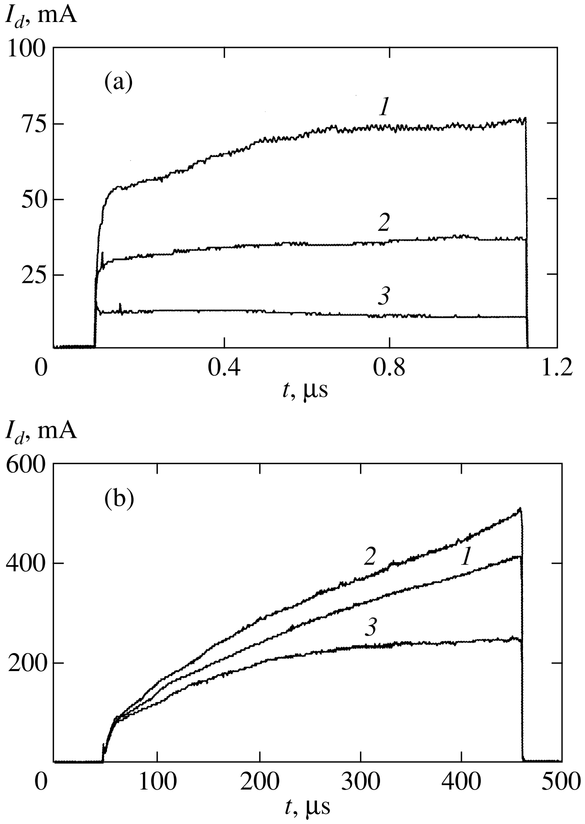


Fig. 1. The current of a non-self-sustained discharge in (a) helium and (b) nitrogen at j_b = 90 μA/cm^2 for various concentrations of macroparticles: (a) U = 780 V, n_d = 0 (curve *1*), 3.1×10^4 (curve *2*), and 4.8×10^5 cm^{-3} (curve *3*); (b) U = 2.1 kV, n_d = 0 (curve *1*), 1.4×10^4 (curve *2*), and 2.8×10^5 cm^{-3} (curve *3*).

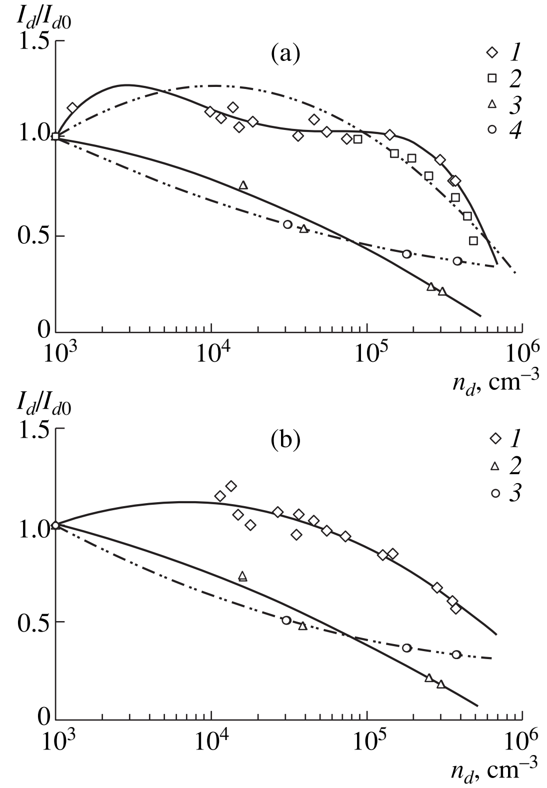


Fig. 2. The current of a non-self-sustained discharge in helium and nitrogen, reduced to the value without a CDP, as a function of the macroparticle concentration after (a) 150 and (b) 400 μs for j_b = 90 μA/cm^2: (a) N_2: U = 2.1 (curve *1*) and 2.9 kV (curve *2*); He: U = 480 (curve *3*) and 780 V (curve *4*). (b) N_2: U = 2.1 kV (curve *1*); He: U = 480 (curve *2*) and 780 V (curve *3*). The curves are plotted by the least squares technique using second-degree polynomials (except curve *1* in (a) for which a fourth-degree polynomial was used).

2. EXPERIMENT

The experiments were carried out in the plasma of a pulsed non-self-sustained discharge in nitrogen on the setup described in [3]. We used nitrogen with an impurity concentration below 0.005%. As in [3], the dust component was in the form of carbon glass balls 24 ± 5 μm in diameter. The experimental conditions were chosen so that, first, a comparison with the results of similar experiments in helium was possible, and second, the widest possible range of electric fields in the nitrogen plasma was covered. A non-self-sustained discharge was initiated under atmospheric pressure. The current density in the fast electron beam with an energy of 125 keV was 90 μA/cm^2, the length of the discharge gap was 0.9 cm, and the area of electrodes was 1 cm^2. The range of working voltages across the discharge gap was limited from below by the fact that during a certain uncontrollable time interval after the initiation, the discharge glows in the Thomson regime with a negligibly small energy contribution of no interest for the experiments, and from above, by a rapid evolution of instability leading to discharge contraction (in our conditions, the working range of voltages was from 2 to 3 kV for a stable discharge time of at least 150 μs).

Typical oscillograms for the discharge current are shown in Figs. 1a (helium) and 1b (nitrogen). For discharge voltages up to 800 V, the discharge current in helium virtually does not change in a certain time interval (400–1000 μs after the initiation of the discharge, the current can only increase by 10% for zero or low concentrations n_d of microparticles and decrease by the same value for $n_d > 10^5$ cm^{-3}). In contrast to experiments with helium, the quasi-stationary mode of the discharge in nitrogen was not observed under our conditions (the discharge current in Fig. 1b increases monotonically with time). For this reason, we compared the values of discharge current at 150 μs, when we can assume that the formation of the cathode layer is completed and the current becomes a linear function of time, and also at 400 μs.

Figure 2 shows the discharge current in nitrogen and helium, reduced to the values of a non-self-sustained discharge in the absence of dust particles, as a function of the concentration of dust particles 150 and 400 μs after the initiation of the discharge for various applied voltages. In nitrogen, as well as in helium plasma [3], a stronger dependence is observed for larger values of the electric field. It can be seen from Fig. 2 that, first, the decrease in the discharge current in nitrogen for high concentrations of macroparticles ($n_d > 10^5$ cm^{-3}) is slower than in helium. Second, the introduction of small amounts of dust ($n_d < 10^5$ cm^{-3}) leads to an increase in the discharge current in nitrogen as compared to the case when the dust component is absent.

A peculiarity of a non-self-sustained discharge in nitrogen with a CDP is a transition to a contracted state under voltages insufficient for a breakdown of the discharge gap in pure nitrogen (Fig. 3). Instability develops for all concentrations of macroparticles studied for discharge currents smaller and larger than in a non-self-sustained discharge without a CDP.

3. CHARGING OF DUST PARTICLES IN ZERO EXTERNAL ELECTRIC FIELD

The charge of dust particles is the most important parameter of a plasma with a CDP, which determines the self-organization ability of such a system. For not very high concentrations of macroparticles in the plasma, the Wigner–Seitz cell method can be used for describing the charging of dust particles in zero external electric field. This method considerably simplifies the situation, transforming the multidimensional problem into a one-dimensional problem and providing the opportunity of a more exact simulation of the dust particle charging.

According to estimates, the conditions for the applicability of the hydrodynamic description of the charged particle transport are satisfied for macroparticles having a radius of the order of 1 μm and above in nitrogen under atmospheric pressure [6–8]:

$$\lambda_e \ll r_0 + d, \quad \lambda_i \ll r_0 + d, \tag{1}$$

where λ_e and λ_i are the mean free paths for electrons and ions, respectively, r_0 is the radius of a dust particle, and d is the characteristic size of the region in which the quasi-neutrality of the plasma is violated.[1] Estimates also show that for macroparticles having a size of 10 μm and above in nitrogen under atmospheric pressure, a more stringent condition for the applicability of the local approximation for determining electronic parameters such as the mobility, the diffusion coefficient, and the rate constants for the creation and loss of electrons is satisfied [6]:

$$\lambda_u \ll r_0 + d, \tag{2}$$

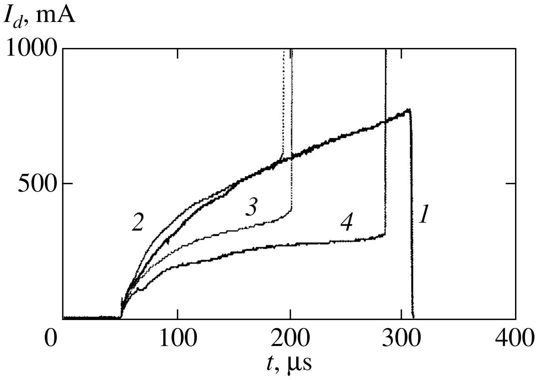


Fig. 3. The current of a non-self-sustained discharge in helium and nitrogen for $j_b = 90$ μA/cm^2 and $U = 2.9$ kV for various concentrations of macroparticles, $n_d = 0$ (curve *1*), 8.8×10^4 (curve *2*), 3.6×10^5 (curve *3*), and 4.8×10^5 cm^{-3} (curve *4*).

where λ_u is the energy relaxation length for electrons. This quantity can be defined as

$$\lambda_u \approx \sqrt{D_e/\nu_u},$$

where ν_u is the frequency of inelastic collisions. If condition (2) is satisfied, the transport coefficients for electrons and the kinetic electronic constants are determined by the local value of the electric field. As regards the transport coefficients for ions, they can be approximately assumed to be constant for not very high fields. It should be noted that for atomic gases the condition opposite to (2) usually holds:

$$\lambda_u \gg r_0 + d. \tag{3}$$

If this condition is satisfied, the electron distribution function is practically unperturbed by the field of a charged dust particle and, hence, we can use the electronic parameters for a plasma unperturbed by the dust component as the first approximation.

Following the Wigner–Seitz cell method, we will consider the charging of a spherical macroparticle placed at the center of a spherically symmetric cell of radius a_d, which is determined by the concentration n_d of dust particles:

$$a_d = \left(\frac{4}{3}\pi n_d\right)^{-1/3}. \tag{4}$$

It should be emphasized here that each unit cell is assumed to be electrically neutral. Consequently, the processes of interaction between dust particles cannot be considered in the given method. For this purpose, the solution of at least two-dimensional problem is required, which is beyond the scope of the present article. In the nitrogen beam plasma under atmospheric pressure, simple ions are rapidly converted into complex ions, and N_4^+ becomes the main positive ion (the fraction of such ions exceeds 80%). Therefore, we can

---

[1] Under atmospheric pressure, the mean free path of ions is usually of the order of 10^{-5}–10^{-6} cm ($\lambda_i = 0.053$ μm for N_4^+ ions in nitrogen at $T = 300$ K), while the same for electrons in nitrogen does not exceed 10^{-4} cm (the electron mean free path $\lambda_e = 0.72$ μm for the reduced field $E/N = 0.01 \times 10^{-16}$ V cm^2). According to calculations, the characteristic size of the region where the quasi-neutrality of plasma is violated is of the order of 10^{-3} cm for particles having a radius of 10^{-4} cm.

assume that we are dealing with only one species of positive ions. In this case, we have the self-consistent system of continuity equations and Poisson's equation for determining the charge of macroparticles in the nitrogen beam plasma:

$$\begin{aligned}\frac{\partial n_e}{\partial t} + \operatorname{div}\mathbf{j}_e &= Q + k_{\text{ion}} n_e N - \beta_{ei} n_e n_i,\\ \frac{\partial n_i}{\partial t} + \operatorname{div}\mathbf{j}_i &= Q + k_{\text{ion}} n_i N - \beta_{ei} n_e n_i,\\ \operatorname{div}\mathbf{E} &= 4\pi e(n_i - n_e),\end{aligned} \tag{5}$$

where n_e and n_i are the concentrations of electrons and ions, respectively, N is the concentration of neutral particles, Q is the rate of gas ionization by an external ionization source, k_{ion} is the rate constant for the gas ionization by intrinsic electrons in the plasma, β_{ei} is the coefficient of the dissociative electron–ion recombination, $\mathbf{j}_e$ and $\mathbf{j}_i$ are the flux densities of electrons and ions, respectively, $\mathbf{E}$ is the electric field strength, and e is the absolute value of the electron charge.

The current densities of electrons and ions in the diffusion-drift approximation are defined by the following expressions:

$$\begin{aligned}\mathbf{j}_e &= -n_e k_e \mathbf{E} - \operatorname{grad}(D_e n_e),\\ \mathbf{j}_i &= n_i k_i \mathbf{E} - D_i \operatorname{grad} n_i,\end{aligned} \tag{6}$$

where k_e and k_i are the mobilities, and D_e and D_i are the diffusion coefficients for electrons and ions, respectively (electronic transport coefficients are functions of the field).

The formulation of the boundary conditions on the surface of a macroparticle for the problem under investigation is a complicated problem in view of the inapplicability of the hydrodynamic approximation in the Knudsen layer, which necessitates the transition to an analysis of the molecular regime of charged particle transport. This problem was thoroughly discussed in connection with the development of theories for radiation and neutron transport [9]. We will assume that the charge of electrons and ions reaching the surface of a dust particle is absorbed with the unit probability. In this case, taking into account the spherical symmetry of the problem, we can introduce the effective boundary conditions [10, 11] for the system of equations (5):

$$\begin{aligned}\left.\left(n_e - \gamma_{0e}\frac{\lambda_e}{r_0}\frac{\partial(rn_e)}{\partial r}\right)\right|_{r=r_0} &= 0,\\ \left.\left(n_i - \gamma_{0i}\frac{\lambda_i}{r_0}\frac{\partial(rn_i)}{\partial r}\right)\right|_{r=r_0} &= 0,\end{aligned} \tag{7a}$$

where r_0 is the radius of a dust particle, and corrections γ_{0e} and γ_{0i} weakly depend on the ratios of the mean free paths to the macroparticle radius. In the limit $\lambda_e/r_0 \longrightarrow 0$ or $\lambda_i/r_0 \longrightarrow 0$, these corrections are equal to the Hopf constant $\gamma_0 \approx 0.71$. The boundary conditions (7a) are correct to within e^{-2r_0/λ_e} and e^{-2r_0/λ_i} [10], and hence can be replaced by the following conditions when $\lambda_e/r_0 \ll 1$ for electrons and $\lambda_i/r_0 \ll 1$ for ions:

$$n_e|_{r=r_0} = 0, \quad n_i|_{r=r_0} = 0. \tag{7b}$$

We will require that the fluxes of charged particles through the boundary of a unit cell be equal to zero at the right boundary, while the electroneutrality condition for the cell implies that the field is equal to zero at the boundary of the unit cell:

$$j_e|_{r=a_d} = 0, \quad j_i|_{r=a_d} = 0, \quad E|_{r=a_d} = 0. \tag{7c}$$

It should be noted that the charge of a dust particle in the one-dimensional approximation under investigation can be determined from the value of the electric field at the interface between the particle and the plasma from the "boundary" condition

$$qe = r_0^2 E|_{r=r_0}, \tag{8}$$

which is a consequence of the charge conservation law.

The boundary-value problem (5) with the boundary conditions (7) will be solved by using the finite-difference method. We construct the difference scheme on the basis of the integral identity method [12] ensuring the continuity of spherical flows on account of the spherical symmetry of the problem:

$$\begin{aligned}J_e &= 4\pi r^2\left(-n_e k_e E - \frac{\partial(D_e n_e)}{\partial r}\right),\\ J_i &= 4\pi r^2\left(n_i k_i E - D_i\frac{\partial n_i}{\partial r}\right).\end{aligned} \tag{9}$$

The charge of dust particles is determined through the equation

$$\frac{dq}{dt} = -(J_{i,0} - J_{e,0}), \tag{10}$$

where $J_{e,0}$ and $J_{i,0}$ are the fluxes of charged particles for $r = r_0$. A comparison of the values of the charge obtained by integrating (10) with those determined by using relation (8) allows us to judge the accuracy of the numerical solution of the problem. We used an inhomogeneous spatial mesh condensing in the direction to the surface of a macroparticle.

It should be noted that the system of equations (5) is stringent and therefore quite complicated for a numerical solution. For example, if we use the explicit scheme, the size of the time step is bounded by the condition $\Delta t < \min(h^2/2D_e)$, while the size of the step along the radius is limited from above by the required accuracy of the solution as well as by the Gibbs instabilities, which weakly affect the accuracy of determining the charge, but are nevertheless undesirable. On the other hand, the characteristic time of stabilization of the charge and concentration of charged particles is determined by much slower ions. For not too low concentrations of macroparticles, the order of magnitude of this time can be estimated as

$$\tau_D \sim a_d^2/D_i.$$

Consequently, the number of time steps required for obtaining a steady-state solution of the problem is of the order of $(a_d/h)^2(D_e/D_i)$ (for $a_d \sim 100$ μm and $h = 0.1$ μm, the number of steps in nitrogen can be as large as 10^7 and even more). For this reason, the application of stable semi-explicit or implicit integration schemes is the only possibility for solving the given boundary-value problem with reasonable computer time expenditures.

In the present work, the integration with respect to time was carried out according to the unconditionally stable semi-explicit Krank–Nicholson algorithm of the second order of accuracy. The solution of finite-difference equations on each time layer was sought by the matrix factorization technique [13] using the Newton–Kantorovich linearization method. The electronic parameters for nitrogen required for the calculations were borrowed from [14–16]. The tabulated data presented in [14, 15] were used for deriving approximate expressions ensuring the continuity of the function itself and (whenever possible) its first derivative depending on the reduced field. It is well known that the diffusion coefficient in an electric field depends on direction [7, 8]. In the course of charging a dust particle in zero external field, electrons diffuse in the direction opposite to the radial electric field, and hence we must take for D_e in (6) the value of the longitudinal diffusion coefficient D_L for electrons. The data on D_L in nitrogen at cryogenic temperatures are not available. For this reason, we used in our calculations the data on the transverse diffusion coefficient D_T for the gas temperature $T = 77$ K. The coefficient of dissociative recombination of electrons and N_4^+ ions was calculated from the functional dependence on the electronic temperature T_e from [16] at $T = 300$ K using the approximate dependence of T_e on the reduced field E/N obtained from the data borrowed from [15], while at $T = 77$ K, the approximate dependence of the characteristic electron energy on E/N was used instead of T_e. The results on electronic parameters in argon required for deriving approximate dependences were obtained by numbering the graphs presented in [17] and were supplemented with the data from [14]. The mobility of N_4^+ ions at various gas temperatures in nitrogen was determined in [18], while the data for Ar_2^+ ions at room and cryogenic temperatures were borrowed from [19].

4. DISCUSSION OF RESULTS OF NUMERICAL SIMULATION

Figure 4 shows the curves illustrating the evolution of the charge on macroparticles having a radius of 12 μm at room and cryogenic temperatures for various gas ionization rates, while Fig. 5 presents the dependences of the macroparticle charge on the gas ionization rate for particles of different sizes. Calculations were carried out until the charge attained a steady-state level (for low ionization rates, almost up to 1 ms). Note that the current density $j_b = 90$ μA/cm^2 in the fast electron beam corresponds, according to Cason *et al.* [20], to the ionization rate in nitrogen $Q = 1.5 \times 10^{17}$ cm^{-3} s^{-1}. It can be seen from Fig. 4 that the magnitude of the charge as well as the characteristic time of charge stabilization depend on the ionization rate of the gas. Calculations made with fine meshes proved that the values of the charge obtained by the method of the integration of the total current (10) per macroparticle and by formula (8) usually differ only starting from the third or fourth decimal place (note that the value $E|_{r = r_0}$ is determined from the solution of system (5) using the finite-difference technique described above). As we go over to coarser meshes, Gibbs instabilities emerge in the vicinity of a macroparticle (which is manifested in an irregular behavior of the fluxes), invalidating the method of integration for determining the charge, while formula (8) gives the value of the charge with an admissible accuracy.

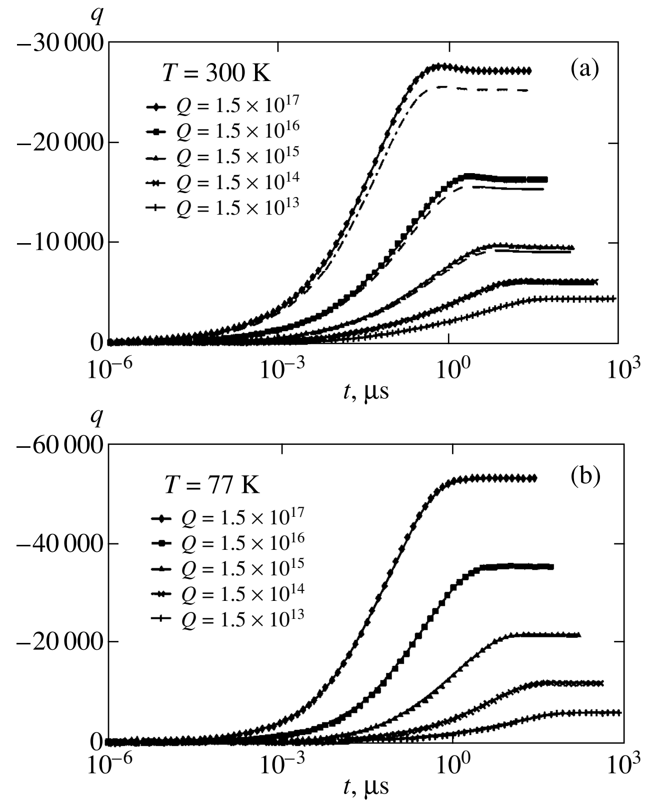


Fig. 4. Curves illustrating the evolution of the charge on dust particles of radius 12 μm for various ionization rates at the nitrogen temperature (a) 300 and (b) 77 K, $n_d = 10^4$ cm^{-3}. Dashed curves are calculated with the boundary conditions (7a).

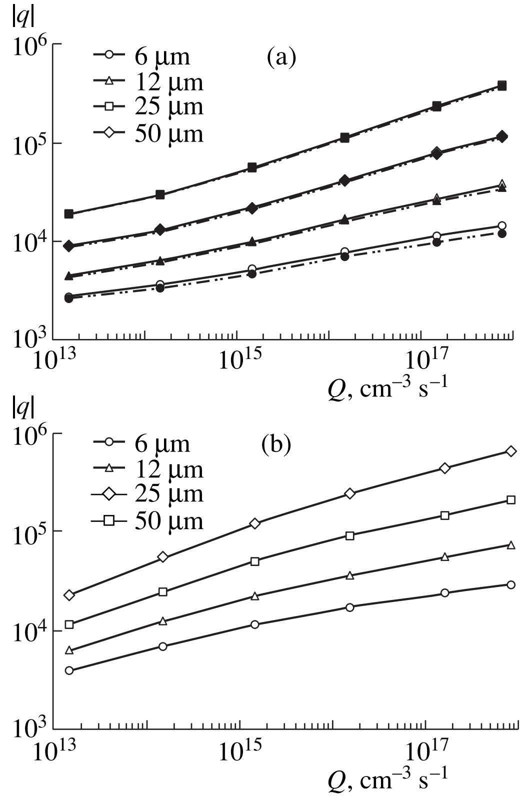


Fig. 5. Dependence of the charge on dust particles of various radii on the ionization rate for $n_d = 10^4$ cm^{-3} at $T = 300$ (a) and 77 K (b). Solid curves are calculated with the zero boundary conditions (7b) and dot-and-dash curves, with the boundary conditions (7a).

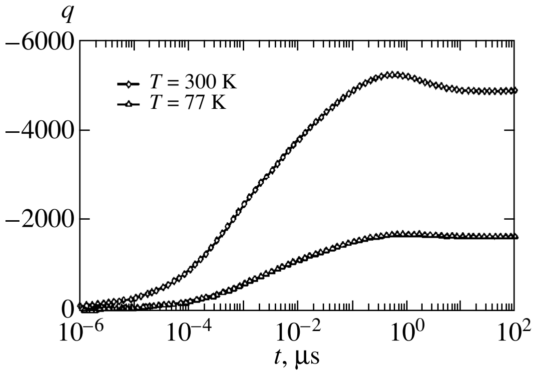


Fig. 6. Curves illustrating the evolution of the charge on dust particles of radius 12 μm for $Q = 10^{16}$ cm^{-3} s^{-1} in argon at room and cryogenic temperatures ($n_d = 10^4$ cm^{-3}).

For this reason, we will henceforth give the values of the charge determined by using formula (8).

Figure 4 shows that the charge of a macroparticle at room temperature passes through a peak, while at cryogenic temperatures, no peak is observed, the charge attaining the steady-state value monotonically. The same situation also takes place in argon (see Fig. 6). It can also be seen from Figs. 4 and 5 that the charge of macroparticles at cryogenic temperatures is noticeably larger than at room temperature. A comparison of the results obtained with the boundary conditions (7a) and (7b) shows that they differ insignificantly as expected. For this reason, we will henceforth operate with the data calculated with the zero boundary conditions (7b). The curves in Fig. 6 calculated with frozen electronic parameters (we used the parameters for zero field) illustrate the evolution of the charge of macroparticles in argon at room and cryogenic temperatures. A comparison of Figs. 4, 5, and 6 shows that as the gas temperature decreases in argon, the charge of macroparticles noticeably decreases in contrast to nitrogen. For this reason, the Coulomb parameter

$$\Gamma = e^2 q^2 / a_d T \qquad (11)$$

of the interaction between macroparticles in nitrogen increases considerably upon cooling (see Fig. 7a), while in argon, on the contrary, it decreases noticeably. This is due to the fact that the electron mobility in nitrogen increases upon cooling from 300 to 77 K by a factor exceeding three, while the mobility of ions decreases from 2.3 to 1.9 cm^2/(V s) [18]. A decrease of temperature in argon results in a decrease in the electron mobility [17], while the mobility of ions remains virtually unchanged [19]. Consequently, the gain in the value of Γ in (11) in argon due to a decrease in the denominator upon cooling is suppressed by a stronger decrease in the numerator.

The interaction of charged particles in the physics of low-temperature plasma is usually described either by the model of one-component plasma, or the model with a screened Debye potential. In the first model, it is assumed that the interaction potential is of Coulomb's type, and the value of the nonideality parameter Γ determines the phase state of the system [1]. For $\Gamma > 171$, the dusty plasma is transformed into a state with a crystalline structure, i.e., a Coulomb crystal is formed. According to the calculations, the value of this parameter under our experimental conditions is considerably larger. It can be seen from Fig. 7a illustrating the variation of the nonideality parameter upon a transition from cryogenic to room temperature that a decrease in the gas temperature in nitrogen leads to a considerable increase in the nonideality parameter of dusty plasma, which may facilitate the formation of ordered dust-plasma structures in nitrogen.

In the second model, the description of the plasma thermodynamics is complicated, and the conditions of phase transitions are determined in this case by two parameters instead of one, i.e., the ratio

$$\kappa = a_d / R_d$$

and the parameter

$$\Gamma_s = \Gamma \exp(-a_d / R_d).$$

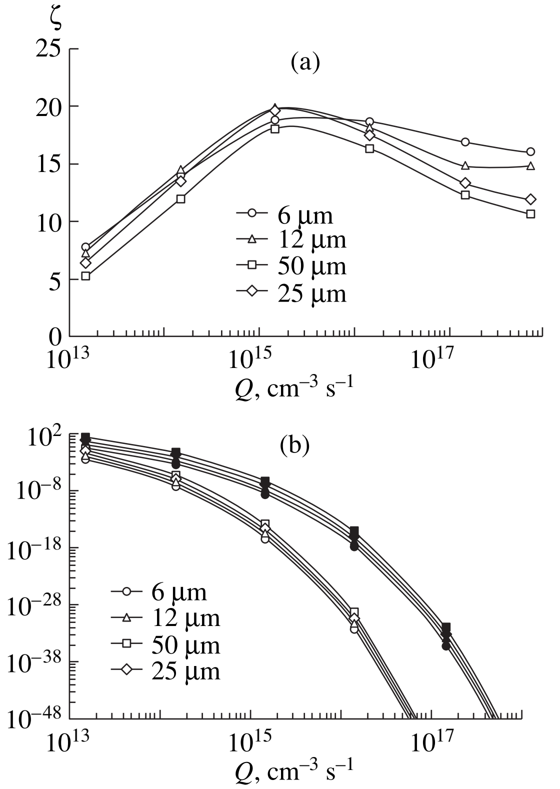


Fig. 7. Ratio of (a) the nonideality parameters $\zeta = \Gamma_{77}/\Gamma_{300}$ and (b) parameter Γ_s at $T = 77$ and 300 K as functions of the ionization rate for $n_d = 10^5$ cm^{-3} in nitrogen. Light symbols in (b) correspond to $T = 77$ K, and dark symbols, to $T = 300$ K.

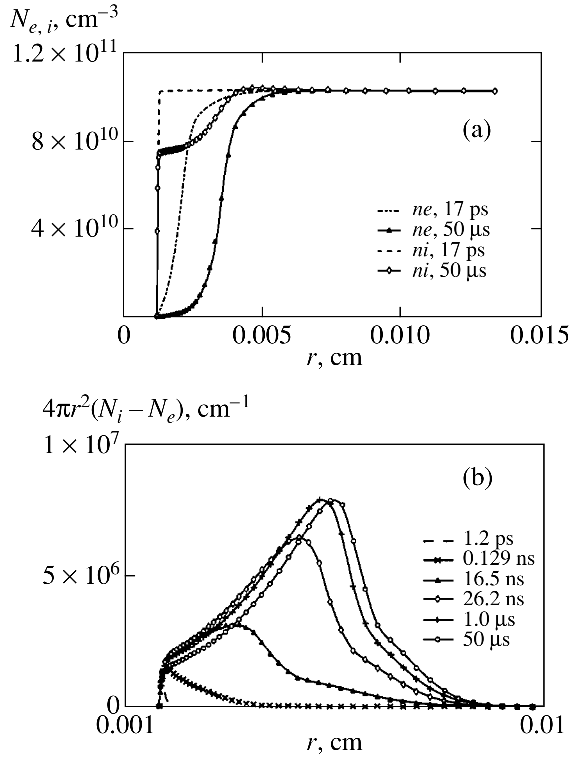


Fig. 8. Radial distributions of the concentrations of (a) electrons and ions and (b) charge in the nitrogen beam plasma at various instants of time under atmospheric pressure (r_0 = 12 μm, $n_d = 10^5$ cm^{-3}, $Q = 1.5 \times 10^{16}$ cm^{-3} s^{-1}).

Here, R_d is the Debye radius of plasma screening, which in the case of an isothermal plasma is defined as

$$R_d = \sqrt{T/8\pi e^2 n_\infty},$$

where n_∞ is the concentration of charged particles away from a dust particle (we assume here that the concentration of ions is equal to the electron concentration):

$$n_\infty = \sqrt{Q/\beta_{ei}}.$$

Figure 7b gives the values of parameter Γ_s as a function of the gas ionization rate for a dust particle concentration of 10^5 cm^{-3}. It can be seen from the figure that the value of parameter Γ_s decreases noticeably with the gas temperature, but the dependence on the gas ionization rate is even stronger. This is due to a decrease in the Debye radius of plasma screening (parameter Γ_s is an exponential function of the Debye radius) upon an increase in the gas ionization rate and, accordingly, upon an increase in the concentration of charged particles. The value of R_d also decreases with the gas temperature, leading to an increase in parameter κ and a decrease in parameter Γ_s upon a transition from room to cryogenic temperatures. According to the result of numerical calculations, the crystallization of dusty plasma in the model with the Debye interaction potential for $\kappa \sim 10$ occurs for values of the parameter $\Gamma_s \sim 1$ (see [21]). It can be seen from Fig. 7b that at low ionization rates, when the value of R_d is quite large and the value of parameter κ is close to 10, the charge on dust particles is still significant, and the value of parameter Γ_s is close to unity or exceeds it. This leads to the conclusion that the Debye model also indicates the possibility of creating an ordered dust structure in the nitrogen beam plasma.

Figure 8 shows that the steady-state radial distribution of ion density has a small peak whose height increases with the nitrogen ionization rate according to the results of calculations. The peak and the region of the slow decrease in the ion concentration behind it appearing as a macroparticle moves to the surface are associated with the drift motion of ions in the attractive field of the macroparticle, which increases with decreasing distance to it, while the next region of the sharp decrease in n_i in the vicinity of the macroparticle is associated with a diffusive departure of ions to the macroparticle. It follows from Fig. 8 that for $Q \approx 10^{16}$ cm^{-3} s^{-1}, the plasma becomes electrically neutral even at a distance of the order of 100 μm (this distance increases with decreasing rate of ionization).

Figure 9 shows the dependence of the characteristic time τ_q of the charge stabilization on the parameters of

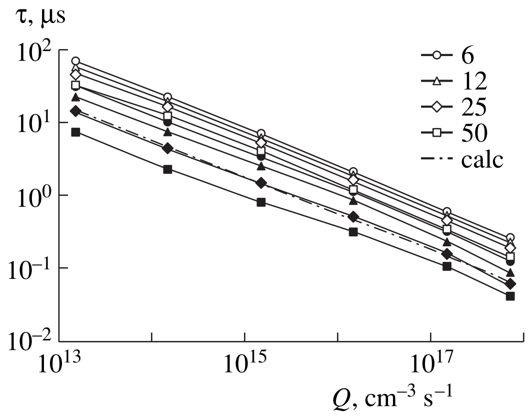


Fig. 9. Characteristic time of charge stabilization on dust particles of various radii (in micrometers) as a function of the ionization rate for $n_d = 10^4$ cm^{-3}in nitrogen. Light symbols correspond to $T = 77$ K and dark symbols, to $T = 300$ K. The dot-and-dash curve describes the diffusion time during which an ion traverses a distance equal to the Debye radius.

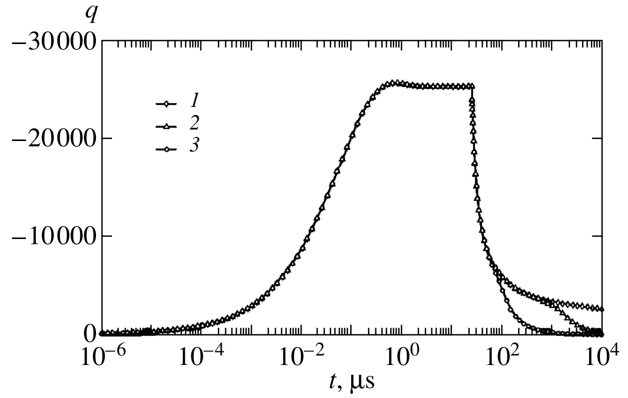


Fig. 10. Curves describing the evolution of the charge on dust particles of radius 12 μm for $j_b = 90$ μA/cm^2 in nitrogen at 300 K for various concentrations of macroparticles n_d = 10^4 (curve *1*), 10^5 (curve *2*), and 10^6 cm^{-3} (curve *3*) (calculated with the boundary conditions (7a)).

the dusty plasma. This quantity was defined as the time required for attaining the value

$$q = (1 - \exp(-3))q_\infty \approx 0.95q_\infty$$

(q_∞ is the charge under steady-state conditions). In our problem, τ_q is determined by slower ions, and we can assume for our estimates that τ_q is the diffusion time over which an ion crosses the region where quasi-neutrality is violated with a characteristic size (and not radius!) d:

$$\tau_q \approx \frac{4}{D_i}\frac{d^2}{\pi^2}. \qquad (12)$$

For our estimates, we assume that d is equal to the Debye radius. Figure 9 presents the curve calculated on the basis of (12). The behavior of this curve upon a change in the ionization rate correlates with the behavior of the characteristic time of charge stabilization as a function of the ionization rate.

It can be seen from Fig. 9 that for the current density $j_b = 90$ μA/cm^2 in the fast electron beam ($Q = 1.5 \times 10^{17}$ cm^{-3} s^{-1}), the characteristic time of the macroparticle charge stabilization amounts to less than 1 μs. An analysis of the evolution of the charge on macroparticles after switching off the fast electron beam following the attainment of the quasi-stationary level (see Fig. 10) shows that the neutralization of the charge is a much slower process as compared to charging (for $n_d = 10^4$ cm^{-3}, the charge still amounts to 10% of the steady-state value even after 10 ms). This is due to the fact that lighter electrons are accumulated with time from larger and larger distances until the region of electron density perturbation reaches the boundary of a unit cell. After this, there are virtually no electrons in the bulk. In order to completely neutralize the charge of a macroparticle, we must gather ions from the entire unit cell, which requires a time of the order of $(4/D_i)(a_d/\pi)^2$; for $n_d = 10^4$ cm^{-3} and $T = 300$ K, this time is ~5.6 ms. As the concentration of macroparticles increases tenfold to 10^5 cm^{-3}, this time decreases to 1.2 ms (see Fig. 10). Large values of the time of macroparticle charge neutralization indicate that the formation of ordered dust structures can be studied in systems with a pulse-periodic switching-on of the beam (with a large relative pulse duration). The time of the formation of ordered structures can be roughly estimated from above as the time during which a macroparticle possessing a thermal energy moves over the average distance between particles. For particles of diameter 24 μm in nitrogen under atmospheric pressure, this time is of the order of a few seconds.[2] Consequently, for a dust-plasma structure to have time to form, the beam plasma should be maintained during this time interval; this can be done easily in pulse-periodic systems. The electron gun used in our experiments could operate only in the pulse mode (with a pulse duration up to 1 ms). For this reason, no ordered dust-plasma structures were observed although the value of the nonideality parameter was much larger than the critical value according to the estimates.

In order to determine the extent to which the dust component affects the average concentrations of charged particles, we calculated the fraction of positive ions and, accordingly, electrons are lost as a result of recombination on the surface of dust particles under steady-state conditions. This quantity is defined as

$$\eta = J_{i,0}n_d/Q.$$

Figure 11a shows the curves describing η as a function of dusty plasma parameters. According to calculations,

---

[2] The steady-state velocity with which carbon glass macroparticles of radius 12 μm in nitrogen under atmospheric pressure at room temperature fall in the gravity field is 4.4 cm/s.

the ratio η/n_d as well as the charge of macroparticles are virtually independent of the concentration of macroparticles in the range of ionization rates under investigation for $n_d \leq 10^5$ cm^{-3}. At higher concentrations, the dependence on n_d appears first for low ionization rates and large macroparticles, and then for large values of Q and small values of r_0 also. However, the condition for the applicability of Wigner–Seitz cells starts being violated, and such calculations are of no practical value. Figure 11a demonstrates that the perturbation exerted by the dust component on the nitrogen plasma for $Q = 1.5 \times 10^{17}$ cm^{-3} s^{-1} is quite small under our experimental conditions even for $n_d = 10^5$ cm^{-3}.

The values of the effective coefficient of the recombination of charged particles on the surface of macroparticles were determined in [3]. The calculations made by us here indicate that the effective coefficient of the recombination of positive ions on a macroparticle is approximately equal to the quantity defined by the Langevin theory: $\beta_{id} = 4\pi e k_i$ [19], but is always smaller than this quantity in the steady state (see Fig. 11b). This is a consequence of the fact that the electric field decreases with increasing distance at a much higher rate as compared to Coulomb's law. It can be seen from Fig. 11b showing the dependence of the quantity $\xi = J_{i,0}/|q|\beta_{id} n_{i,M}$ ($n_{i,M}$ is the concentration of ions at the boundary of a unit cell) on the ionization rate that the ratio of the rate of ion lost at the surface of a macroparticle to the Langevin rate weakly depends on the ionization rate, and its behavior upon a change in Q changes significantly upon a transition from room to cryogenic temperatures.

Concluding this section, we consider the fulfillment of the similitude relations established by Belov *et al.* [22] on the basis of a simplified model of macroparticle charging. It was found in [22] that

$$q/r_0 = f(n_d r_0, Q),$$

where f is a certain function. An analysis of the results of a numerical solution shows that the similitude relation is quite rough and can be used only for obtaining estimates. This is due to the fact that the two main assumptions underlying the simplified model [22, 23], i.e., (1) the independence of the fluxes on the radial coordinate, and (2) the Coulomb dependence of the electric field, are quite approximate according to the numerical calculations. It should also be noted that the dependence of the charge of dust particles on the ionization rate is close to $Q^{1/4}$, while the charge is virtually independent of n_d in the range under investigation.

5. CHARGING OF DUST PARTICLES IN AN EXTERNAL ELECTRIC FIELD

Let us now consider the charging of macroparticles in the presence of an external electric field. In this case, the symmetry of the problem is lowered to the axial symmetry, and a two-dimensional analysis is required, which should be carried out separately. We will estimate the charge of macroparticles in such a case, following Reist [24]. In our experiments, macroparticles were made of carbon glass which is (although a poor) conductor, its conductivity being five to six orders of magnitude higher than the conductivity of the beam plasma in our experiments. For this reason, the electric field does not penetrate into a macroparticle, and the potential φ outside it is described by the formula [25]

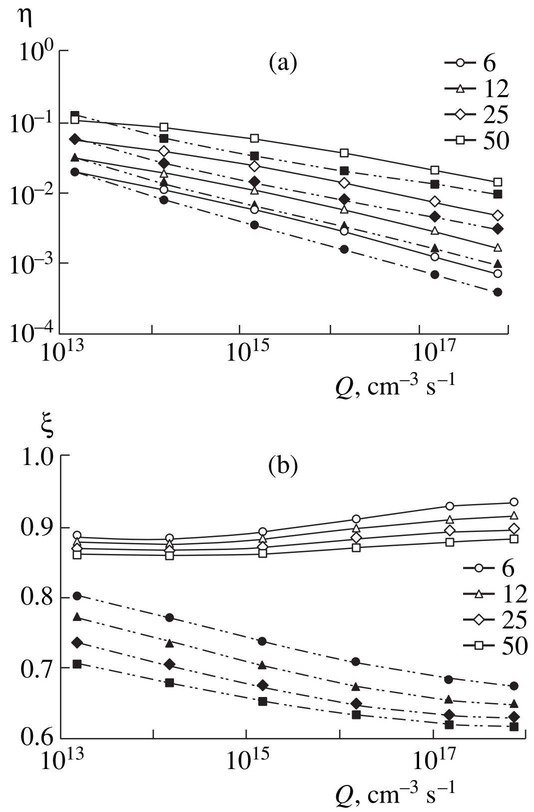


Fig. 11. (a) The fraction of ions perishing as a result of recombination on the surface of macroparticles of various radii (in micrometers) and (b) the ratio of the surface recombination rate to the Langevin rate in nitrogen under steady-state conditions for $n_d = 10^4$ cm^{-3} as functions of the ionization rate. Light symbols correspond to $T = 77$ K and dark symbols, to $T = 300$ K.

$$\varphi = \frac{qe}{r} - \mathbf{E} \cdot \mathbf{r}\left(1 - \frac{r_0^3}{r^3}\right), \qquad (13)$$

where $\mathbf{E}$ is the vector of the applied electric field strength, $\mathbf{r}$ is the radius vector with the origin at the center of the macroparticle, and q is the macroparticle charge. We assume that the potential on the surface of an uncharged macroparticle is equal to zero. Under conditions (1), charged particles move along the electric field lines, and the radial component of the field is given by

$$E_r = -\frac{\partial \varphi}{\partial r} = \frac{qe}{r^2} + E\cos\theta\left(1 + \frac{2r_0^3}{r^3}\right), \qquad (14)$$

where θ is the angle measured from vector $\mathbf{E}$. As the negative charge increases, the number of electrons reaching the macroparticle decreases, and the number of ions increases until the electron and ion fluxes balance each other for a certain value of the charge. Considering that $k_e \gg k_i$, we can assume that this takes place when the radial field at the surface of the macroparticle becomes negative everywhere. This leads to the following expression for estimating the charge:

$$q = -\frac{3Er_0^2}{e}, \tag{15}$$

whence $q \approx -9 \times 10^4$ for particles of radius $r_0 = 12$ μm in the field $\mathbf{E} = 3$ kV/cm. This value is almost thrice as large as the macroparticle charge in zero field, and hence the critical condition for a transition to the state with an ordered dust structure is satisfied in a non-self-sustained discharge in nitrogen also.

Let us now consider the effects associated with the influence of the CDP on the properties of a non-self-sustained discharge in nitrogen, which were observed in our experiments. It was proved above that a significant negative charge is accumulated on macroparticles. This charge can produce a focusing effect on electrons in the plasma (which is similar to focusing of the drift current of ions by negatively charged particles [26]). This leads to the formation of regions with an enhanced liberation of energy, which in turn leads to an increase in the extent of the vibrational excitation of nitrogen in these regions. It was proved in our earlier publication [27] that a transition to the contracted state in nitrogen occurs when the excitation wave reaches approximately the 15th or 16th vibrational level. Consequently, an increase in the extent of vibrational excitation may lead to a decrease in the stable-discharge time.

An increase in the extent of vibrational excitation may also lead to an increase in the discharge current for low concentrations of dust particles. A decrease in the non-self-sustained discharge current for high concentrations of macroparticles is due to the fact that the specific mass of macroparticles increases with n_d and becomes larger than the specific mass of nitrogen for $n_d > 9 \times 10^4$ cm^{-3}. This increases ineffective losses of the energy of fast electrons in the beam and reduces the gas ionization rate. Naturally, the discharge current decreases as a result.

A large negative charge of macroparticles may also lead to another mechanism of increase in the current. The large charge may cause an increase in the coefficient of the ion–electron emission of carbon glass and ultimately to the short-circuiting of the discharge current through macroparticles whose conductivity is much higher than that of the plasma (see above). This effect takes place when a minicathode layer is formed on the side of a macroparticle facing the anode. Such an event does not appear to be improbable if we take into account the fact that the electric field on this side ($\theta = \pi$) is strongly enhanced (by a factor of six on the surface of the macroparticle!). Under the above conditions in which the charge of a macroparticle was estimated, this field is 18 kV/cm, which is only half the value of the breakdown voltage in nitrogen under atmospheric pressure [28]. According to estimates, the charge of macroparticles in helium must be much smaller than in nitrogen, and hence such a mechanism should not operate in helium. Probably, the discharge current in helium only decreases upon an increase in the dust concentration due to a decrease in the gas ionization rate in this case.

6. CONCLUSION

The experimental investigations of a non-self-sustained discharge in nitrogen with a condensed disperse phase under atmospheric pressure carried out in this work demonstrated a strong effect of macroparticles on the current–voltage characteristics as well as on the stability of the discharge process. The numerical simulation of the charging of dust particles occurring in the continuous medium mode under our experimental conditions proved that a considerable charge is accumulated at macroparticles in the nitrogen beam plasma. It was found that a decrease in the gas temperature noticeably increases the charge of macroparticles in nitrogen and, on the contrary, decreases this charge in argon. As a result, the Coulomb parameter of interaction between dust particles strongly increases upon a transition from room to cryogenic temperature in nitrogen and noticeably decreases in argon. The calculations also proved that the charge attains the quasi-stationary level over time periods of the order of a microsecond at high ionization rates, while the charge neutralization takes time of the order of a millisecond. For this reason, the formation of ordered dust-plasma structures can be studied experimentally in systems with pulse-periodic ionization of the gas. An analysis of possible mechanisms of the effect of the dust component on the characteristics of a non-self-sustained discharge proved that it may occur through the focusing effect of the electric field of charged macroparticles on thermal electrons in the plasma. This leads to an increase in the discharge current and to a decrease in the stable discharge time through an increase in the extent of vibrational excitation of nitrogen molecules. The mechanism of the initiation of microscopic discharges between macroparticles may lead to similar effects. Additional studies including a multidimensional simulation of the problem under investigation are required to clarify completely these mechanisms of the effect of the CDP on the non-self-sustained discharge.

ACKNOWLEDGMENT

This work was partially supported by the scientific program Universities of Russia (grant no. 015.01.02.06).

REFERENCES

1. A. P. Nefedov, O. F. Petrov, and V. E. Fortov, Usp. Fiz. Nauk **167**, 1215 (1997) [Phys. Usp. **40**, 1163 (1997)].
2. V. N. Tsytovich, Usp. Fiz. Nauk **167**, 57 (1997) [Phys. Usp. **40**, 53 (1997)].
3. V. V. Ivanov, A. F. Pal’, T. V. Rakhimova, *et al.*, Zh. Éksp. Teor. Fiz. **115**, 2020 (1999) [JETP **88**, 1105 (1999)].
4. V. I. Donin and Yu. I. Khapov, Kvantovaya Élektron. (Moscow) **13**, 1583 (1986).
5. M. J. McCaughey and M. J. Kushner, J. Appl. Phys. **69**, 6952 (1991).
6. L. D. Tsendin, Zh. Éksp. Teor. Fiz. **66**, 1638 (1974) [Sov. Phys. JETP **39**, 815 (1974)].
7. N. L. Aleksandrov, A. P. Napartovich, and A. N. Starostin, Fiz. Plazmy **6**, 1123 (1980) [Sov. J. Plasma Phys. **6**, 618 (1980)].
8. N. L. Aleksandrov, A. M. Konchakov, A. P. Napartovich, and A. N. Starostin, Khim. Plazmy **11**, 3 (1984).
9. F. M. Morse and H. Feshbach, *Methods of Theoretical Physics* (McGraw-Hill, New York, 1953; Inostrannaya Literatura, Moscow, 1958, 1960), Parts 1, 2.
10. B. Davison, *Neutron Transport Theory*, with the collaboration of J. B. Sykes (Clarendon Press, Oxford, 1957; Atomizdat, Moscow, 1960).
11. G. I. Marchuk, *Methods of Analysis of Nuclear Reactors* (Atomizdat, Moscow, 1961).
12. G. I. Marchuk, *Methods of Computational Mathematics* (Nauka, Moscow, 1989).
13. A. A. Samarskiĭ and A. V. Gulin, *Numerical Methods* (Nauka, Moscow, 1989).
14. J. Dutton, J. Chem. Phys. Ref. Data **4**, 577 (1975).
15. A. V. Phelps, FTP://JILA.COLORADO.EDU/COLLISION–DATA/ELETRANS.TXT.
16. J. B. A. Mitchell, Phys. Rep. **186**, 215 (1990).
17. J. L. Pack, R. E. Voshall, A. V. Phelps, and L. E. Kline, J. Appl. Phys. **71**, 5363 (1992).
18. H. Bohringer and F. Arnold, Int. J. Mass Spectrom. Ion Phys. **49**, 61 (1983).
19. B. M. Smirnov, *Ions and Excited Atoms in Plasma* (Atomizdat, Moscow, 1974); *Complex Ions* (Nauka, Moscow, 1983).
20. C. Cason, J. F. Perkins, A. H. Werkheizer, and J. Dudersdadt, AIAA J. **15**, 1079 (1977).
21. M. J. Stevens and M. O. Robbins, J. Chem. Phys. **98**, 2319 (1993).
22. I. A. Belov, A. S. Ivanov, D. A. Ivanov, *et al.*, Zh. Éksp. Teor. Fiz. **117**, 105 (2000) [JETP **90**, 93 (2000)].
23. B. M. Smirnov, *Aerosols in Gas and Plasma* (Inst. Vys. Temp. Akad. Nauk, Moscow, 1990); Usp. Fiz. Nauk **170**, 495 (2000).
24. P. C. Reist, *Introduction to Aerosol Science* (Macmillian, London, 1984; Mir, Moscow, 1987).
25. V. V. Batygin and I. N. Toptygin, *Problems in Electrodynamics* (Nauka, Moscow, 1970; Academic, London, 1964).
26. O. M. Belotserkovskiĭ, I. E. Zakharov, A. P. Nefedov, *et al.*, Zh. Éksp. Teor. Fiz. **115**, 819 (1999) [JETP **88**, 449 (1999)].
27. A. F. Pal’ and A. V. Filippov, Preprint TRINITI No. 0019-A (TsNIIAtomInform, 1995).
28. Yu. P. Raĭzer, *The Physics of Gas Discharge* (Nauka, Moscow, 1987).

Translated by N. Wadhwa

Journal of Experimental and Theoretical Physics, Vol. 92, No. 2, 2001, pp. 246–259.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 119, No. 2, 2001, pp. 285–300.
Original Russian Text Copyright © 2001 by Misochko.

Coherent Phonons and Their Properties

O. V. Misochko
Institute of Solid State Physics, Russian Academy of Sciences, p/o Chernogolovka, Moscow oblast, 142432 Russia
e-mail: misochko@issp.ac.ru

Received September 19, 2000

Abstract—The optical phonons in semimetals, semiconductors, and superconductors were studied by the light reflection techniques with femtosecond time resolution and by the method of spontaneous Raman scattering. During measurements in the time domain, the phonon system is converted into a coherent state by the first ultrashort laser pulse and then probed at a variable delay by the second pulse. In this case, the phonons are shown to occur in a nonclassical state in which their fluctuational properties, different in various quadratures, are described by periodic functions of time. A comparison of the results obtained in the time and frequency domains gives evidence that the energies of thermal and coherent phonons coincide, while their dephasing and energy relaxation times are different. *© 2001 MAIK “Nauka/Interperiodica”.*

1. INTRODUCTION

Phonons, as well as all other elementary and collective excitations in solids, should be treated as essentially quantum objects. Repeated changes of any variable involved in the description of the phonon field provide for an average value with an accuracy limited by the noise. The quantum noise is a fundamental property of any system and it was commonly believed for a long time that the accuracy of measurements cannot be better than that determined by the noise level. However, despite being the inseparable property of any system as stipulated by the Heisenberg uncertainty principle [1], the quantum fluctuations possess no immunity against possible manipulations: the Heisenberg principle possesses an internal degree of freedom and poses restrictions on a system only with respect to a pair of conjugated variables. The nature does not forbid the noise being reduced for one of the conjugated variables, provided that the other would accordingly begin to fluctuate more strongly.

Until recently, optical phonons were usually investigated in the frequency domain by spectroscopic methods, which usually provide information on equilibrium thermal phonons [2, 3]. The main spectroscopically measurable quantities are the frequency (i.e., energy), which determines the energy separation of the quantized levels, and the damping (decrement) assigned to the interaction between the system studied and a thermal reservoir (unaccounted degrees of freedom). The progress in laser technologies made it possible to decrease significantly the laser pulse duration, thus opening new prospects for investigations by allowing the crystal lattice dynamics to be studied on a real time scale. Phonons excited by the ultrashort light pulses are conventionally referred to as coherent, because the excitation laser pulse duration is considerably smaller than a characteristic lifetime of the phonon state.

In order to elucidate the equivalence and possible distinctions of the information provided by investigations in the time and frequency domains, we have carried out a comparative study of the phonon characteristics of several solids—typical representatives of the classes of semimetals, semiconductors, and superconductors—by two different optical methods: (i) Raman light scattering and (ii) optical pumping followed by probing with femtosecond pulses. During investigations in the time domain, special attention was paid to the statistical properties of coherent phonons. The structure of this paper is as follows. Section 2 focuses on theoretical principles of the description of various states in the phonon field and explains the essence of experimental methods employed. The experimental results are presented in Section 3, separated according to the types of samples studied. Section 4 is devoted to the discussion and possible interpretations of the results. The main results are summarized in the Conclusion.

2. THEORETICAL PRINCIPLES AND METHODS OF MEASUREMENTS

As noted above, most of the information concerning optical phonons was obtained until recently by investigations in the frequency domain. At present, the availability and wide application of femtosecond lasers make it possible to study the phonons in real time and, moreover, allow some nontrivial experiments. Investigations in the time domain, with a typical procedure based on the phonon system excitation with subsequent probing by two time-separated laser pulses, are now extensively carried out in many condensed systems. There are several good reviews explaining the physics of measurements in the time domain and analyzing the common and distinctive features of these studies in comparison with experiments in the frequency domain [4–6]. However, there are two points, not considered in sufficient detail in these reviews, which are worth special attention.

First, the experiments with condensed media in the time domain usually provide information on the Raman-active phonons. The results obtained by applying the Fourier transform to the time-dependent response are frequently compared to the spontaneous Raman scattering spectra. However, despite the fact that information on the phonon subsystem in both methods is extracted from the results of measurements performed on the photon field, different correlation functions are measured in the frequency and time domains. The Raman scattering is related to correlators of the $\langle a_i^+ a_i \rangle$ type, whereas the optical response in the time domain depends on correlators of the type $\langle a_i^+ a_i a_j^+ a_j \rangle$ (where a_i^+ and a_i are the phonon creation and annihilation operators [7]. Measurements in the time domain provide data on the intensity correlations, the dependence of which on the first-order correlations measured in the frequency domain is often determined by the statistics of the scatterers (in our case, phonons) [8, 9].

The second important aspect, which is also frequently ignored during interpretation of the results obtained in the time domain, is that even performing investigations in the frequency and time domains on the same crystal by no means implies that the results of these measurements would refer to the same state of the system [8]. During spontaneous Raman scattering measurements in the frequency domain, phonons occur in thermal equilibrium in the state with an undetermined phase. In the time domain, the measurements are performed over a coherent system where the phase is a well determined quantity [2, 4, 10]. There are no grounds to ascertain *a priori* that all characteristics of the two systems are identical. It should be emphasized that measurements in the time domain provide information on the nondiagonal elements of the density matrix, whereas measurements in the frequency domain allow us to determine the diagonal elements.

The existence and properties of excitations with nonthermal statistics were (and still are) the subjects in many investigations performed in the photon field. The concept of the deformation (squeezing) of the uncertainty contour was originally theoretically introduced [11] and then experimentally realized [12] for photons. Investigations of the squeezed photon states are reported in a vast number of papers, all of which cannot be cited here; we restrict this list of references to several textbooks and reviews, which are widely known and readily available [13]. It was not until considerably later (in the end of the 1980s) that the first theoretical works appeared in which the analogy with the photon field was used to study the possibility of obtaining squeezed phonon states and their specific properties.

These (rather few) investigations can be divided into three groups. In two of these, phonons were considered as component elements of a composite object: polariton [14] and polaron [15]. In the third group, phonons were treated as independent excitations and their squeezing was derived from the phonon–phonon or phonon–photon interactions [16–21]. It is the theoretical investigations of the third group that provided a necessary basis for and facilitated the realization of the squeezed vibrational and phonon states in optical experiments [20, 22–27].

Historically, the squeezed states were discovered upon reaching a certain level of understanding in the physics of coherent states, which can be determined as eigenstates of the annihilation operator, using the displacement operators, or as the states with minimum uncertainty. Each method of determination of the coherent phonons elucidates various properties of the squeezed states.

In order to rationalize the physical pattern of squeezed phonon states, let us determine, for each separate mode of the phonon field with the annihilation operator a, a coherent state $|\alpha\rangle$ that is an eigenstate of this annihilation operator:

$$a|\alpha\rangle = \alpha|\alpha\rangle.$$

The coherent state can be obtained by applying a unitary displacement operator $D(\alpha) = \exp(\alpha a^+ - \alpha^* a)$ to the vacuum state $|0\rangle$:

$$|\alpha\rangle = D(\alpha)|0\rangle.$$

If a phonon field mode varies with time according to a harmonic law with frequency Ω, the mode can be described in terms of the canonical variables q and p defined as

$$q = \sqrt{\frac{\hbar}{2\Omega}}(a^+ + a), \quad p = i\sqrt{\frac{\hbar\Omega}{2}}(a^+ - a).$$

However, since the canonical variables q and p possess different dimensionalities, it is more convenient to introduce the dimensionless quadrature operators:

$$X = \frac{a^+ + a}{2} = \sqrt{\frac{\Omega}{2\hbar}}q, \quad Y = \frac{a - a^+}{2i} = \frac{1}{\sqrt{2\hbar\Omega}}p.$$

Using these operators, we may write the phonon field in the form of $A = A_c[X\cos(\Omega t) + Y\sin(\Omega t)]$. The coherent states represented as $\alpha = \alpha_1 + i\alpha_2$, where α_1 and α_2 are real quantities, obey the relationships $\langle a \rangle = \alpha$; $\langle N \rangle = |\alpha|^2$; $\langle \Delta X^2 \rangle = \langle \Delta Y^2 \rangle = 1/4$. Defining the dispersion of an arbitrary operator O as $\langle \Delta O^2 \rangle = \langle (O - \langle O \rangle)^2 \rangle$, we obtain for the quadrature operators by virtue of the Heisenberg principle $\langle \Delta X^2 \rangle \langle \Delta Y^2 \rangle \geq 1/16$. In addition, $\langle \Delta X^2 \rangle + \langle \Delta Y^2 \rangle \geq 1/2$. Therefore, the displacement operator shifts the vacuum uncertainty contour from the origin to a point with the coordinates (Reα, Imα), changing neither the shape nor area of the contour.

These relationships indicate that a coherent state can be represented as a classical state with superimposed vacuum state noise. This result explains why the coher-

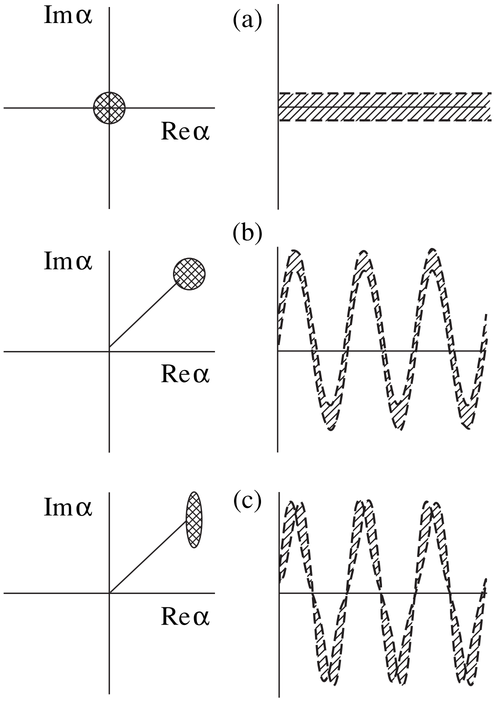


Fig. 1. Schematic representation of the (a) vacuum, (b) coherent, and (c) squeezed states (left diagrams). Diagrams on the right show the evolution of the coherent amplitude, the cross-hatched region between dashed curves corresponds to mean-square fluctuations.

ent state is sometimes referred to as the displaced state (see Fig.1). In the coherent state, only the vacuum is noncoherent: any oscillator in the vacuum state has a zero-point amplitude and a time-independent noise determined by zero oscillations. The amplitude of the coherent state is nonzero, while the noise is identical to the vacuum noise with respect to both magnitude and time variation. The circular shape of the uncertainty contour is evidence that the coherent state noise is the same in both quadratures.

In a more general case, we may create a phonon field in which the dispersion in one quadrature differs from that in the other. For a field with the phonon annihilation operator a, this can be achieved by introducing the operator $b = \mu a + \nu a^{+}$, where μ and ν are the complex values satisfying the condition $|\mu|^2 - |\nu|^2 = 1$. The latter operator transforms the uncertainty contour of the coherent state amplitude to an elliptical shape, whereby the noise in the two quadratures becomes different. Moreover, the noise in one of the quadratures can be reduced below the vacuum noise level at the expense of the corresponding increase in the other quadrature noise. From the standpoint of an observer immobile relative to the coherent amplitude, the noise becomes time-dependent, with the frequency of the noise variation being equal to doubled frequency Ω of the coherent phonon (Fig. 1).

The coherent phonons are usually investigated by the method of phonon system excitation with subsequent probing by subpicosecond light pulses. One of the simple variants of this method, which is most widely used for the study of nontransparent media, is based on the reflection measurements [2, 10]. In this technique, a train of laser pulses is divided into two beams with mutually orthogonal polarizations, the probing beam power being 10–100 times smaller than that of the pumping beam. A relative delay between the two beams is varied by changing the optical path for one of the beams with the aid of a step motor driven system. The pump beam is modulated by a chopper or a shaker. The detection is performed by mixing the signals from photodiodes at the input of a lock-in detector. The output data are obtained in the form of a normalized differential reflection coefficient

$$\frac{\Delta R}{R_0} = \frac{R - R_0}{R_0},$$

representing a change in the reflection of the probing pulse induced by the pumping pulse divided by the probing pulse reflection in the absence of pumping. If the modulation is provided by the shaker (which implies the integration with respect to time), the measured quantity is $\partial(\Delta R/R_0)/\partial t$.

The phenomenon of Raman scattering can be considered as the interaction of light with a crystal, whereby the initial radiation with frequency ω_i is converted into (scattered) radiation with a new frequency ω_s as a result of the medium (crystal) excitation at a characteristic frequency Ω [3, 10]. The laws of the conservation of energy and momentum determine two scattering channels—Stokes and anti-Stokes, differing by the direction of energy transfer: from the photon field to the medium and vice versa. The appearance of a relationship between ω_i and ω_s was explained by the Placzek theory assuming that the optical polarizability is a function of the phonon coordinate (coherent amplitude) $\alpha = \alpha_0 + Q\partial\alpha/\partial Q$, where $\partial\alpha/\partial Q$ is a constant value. A linear relationship $P \propto QE\partial\alpha/\partial Q$ between the photon field and the medium leads to a change in the field frequency. A nonlinear mixing of ω_i and ω_s taking place in the case of ultrashort laser pulses changes the amplitude Q as well, so that the process description requires three conjugated equations for the light field E, coherent amplitude Q, and the differential occupancy n_k of the phonon levels involved in the scattering. The excited phonon state depends on two values, n_k and Q, which are determined by the diagonal and nondiagonal components of the density matrix, respectively [28].

All measurements in this work were performed at room temperature. The Raman scattering was studied in the backscattering geometry, using a triple spectrometer equipped with a multichannel detector. The Raman spectra were excited by a radiation with λ = 780 nm from a Ti : sapphire laser or by a line of an Ar^{+} laser.

Instrumentation used for the Raman scattering measurements is described in detail in [29]. Experiments in the time domain were performed using a lock-in detection scheme and a Ti : sapphire laser operated in the pulsed mode with a pulse duration of 25–50 fs and a pulse repetition rate of 78 kHz. The pumping beam modulation was provided by a 2-kHz chopper or by a shaker operated at 0.5 kHz. In the latter case, the modulation amplitude did not exceed the light pulse duration [27]. The duration and sometimes the initial phase (determined by the time delay) of the laser pulses were measured with the aid of a two-photon detector or a nonlinear crystal [26].

3. EXPERIMENTAL RESULTS

At present, coherent phonons have been studied in various condensed media [4–6]. Our investigation is restricted to nontransparent media including semimetals, semiconductors, and superconductors.

3.1. Semimetals

Semimetals were among the first nontransparent crystalline solids in which oscillations interpreted as coherent phonons were observed [30]. This was partly favored by the fact that semimetals are characterized by a maximum level of the photoinduced optical response to ultrashort laser pulses. We will also use this advantage in studying the statistical properties of the phonons created by femtosecond pulses.

The investigation of coherent phonons was performed in bismuth and antimony—two typical representatives of the group of semimetals. Antimony is known to form crystals with two atoms per unit cell, belonging to the D_{3d} point symmetry group [31]. Among a total of six phonons existing in antimony, three optical phonons of the $A_{1g} + E_g$ type, which are Raman-active, have been registered by now in the time domain [30, 32]. The absence of nondiagonal phonon modes in the first experiments [30] led to the identification and description of the excitation in antimony as proceeding by the displacive excitation of coherent phonons. In this case, the vibrational coherence is created in the excited electron state with a potential minimum displaced relative to the minimum of the initial equilibrium state [33]. Only the fully symmetric phonons can be excited by the displacement mechanism, with the initial phases ±π/2 for the oscillations described by a decaying sinusoidal function. However, the observation of a nondiagonal E_g phonon in the time-dependent response [32] has brought doubt concerning the adequacy of the displacement mechanism. This observation stimulated the development of a mechanism based on the concept of inelastic light scattering (induced Raman scattering).

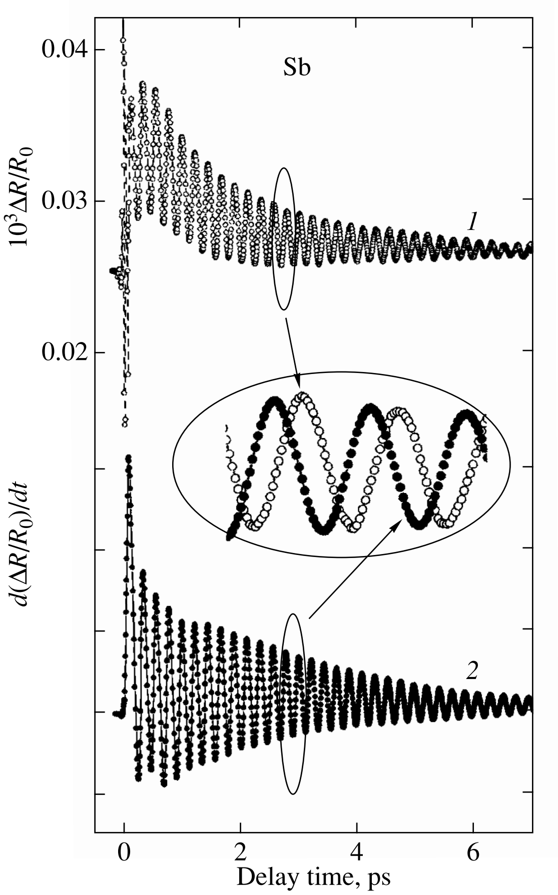


Fig. 2. Time-resolved normalized reflection signal in a Sb sample excited with 30-fs light pulses (the pumping and probing power is 50 and 1.2 mW, respectively). Curves *1* and *2* correspond to different methods of the pumping beam modulation (see the text for explanation). The inset shows a phase shift related to different modulation methods on a greater scale.

Figure 2 shows typical shapes of the optical response to excitation and subsequent probing of a polycrystalline Sb film by femtosecond pulses obtained by modulating the pumping beam with the aid of a mechanical chopper or a shaker. In the former case, the excited electron state relaxes to equilibrium within a time on the order of 2–3 ps. Superimposed onto this electron relaxation are the rapid oscillations related to the excitation of coherent phonons; the purpose of this work was to study the latter oscillations in detail. In order to separate the oscillatory phonon response, the measurements can be performed using the excitation beam modulated with the aid of a shaker. Because, as noted above, these measurements give a derivative of the time-dependent response, the relaxation process gives no significant contribution to the signal. Since the derivative of a harmonic function represents another harmonic function (cos $\longleftrightarrow$ sin), the initial phase of a signal obtained with the tilting retroreflector exhibits a π/2 shift (see the inset in Fig. 2).

In order to study the relationship between oscillations observed by the method of the phonon system

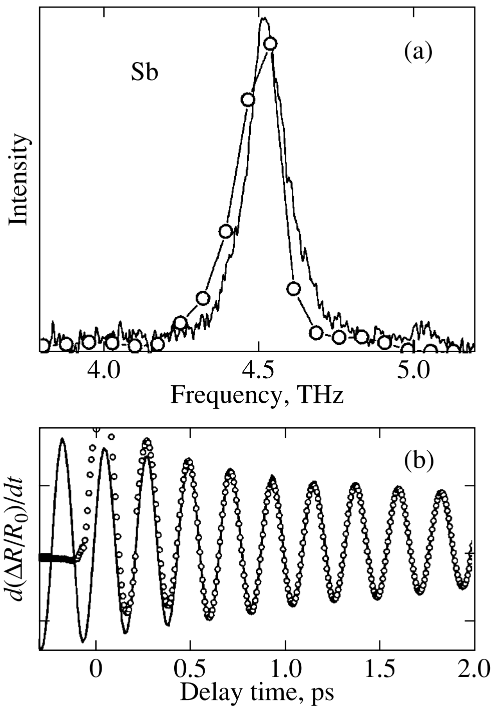


Fig. 3. (a) Coherent phonon oscillations in Sb: (open circles) Fourier transform of the time-resolved response; (solid curve) spontaneous Fourier scattering spectrum. (b) the results of fitting the time-resolved response (solid curve) illustrating the initial phase.

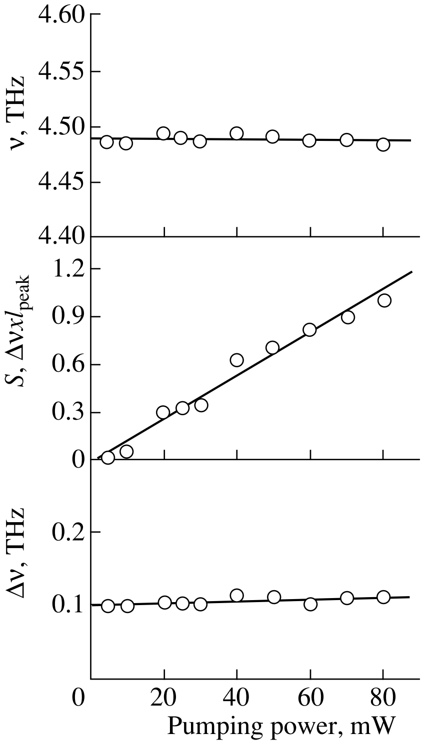


Fig. 4. Plots of the parameters of coherent oscillations versus pumping power for Sb (solid lines show the best linear approximation).

excitation with subsequent probing by delayed laser pulses and the Raman-active modes in Sb, we have used the Fourier transform for numerically converting the data obtained in the time domain into the frequency domain. A typical power spectrum (periodogram) obtained upon such a conversion is presented in Fig. 3a together with a spontaneous Raman spectrum of Sb. The peak observed at 4.49 THz (149.7 cm^{-1}) corresponds to a fully symmetric phonon mode. A comparative analysis of these results indicates that the positions of the phonon peaks obtained from measurements in the time and frequency domains coincide to within the accuracy of measurements. The width (FWHM) of the spectral line $\Delta\nu = 0.1$ THz obtained from measurements of the time-dependent response upon the Fourier transformation and the corresponding mode quality factor $Q = \nu/\Delta\nu$ ($Q = 45$), coincide with the values determined from the spontaneous Raman spectra ($\Delta\nu = 0.1$ THz, $Q = 45$) [30, 31]. The procedure of fitting in the time domain, using a decaying sinusoidal function, gives the same value of the oscillation frequency and the same initial phase $-\pi/2$ (or 0 for the chopper modulation; see Fig. 3).

An increase in the pumping pulse power leads to a linear growth in the amplitude of the coherent oscillations, while the frequency and the rate of dephasing remain unchanged. This is illustrated by Fig. 4 showing the corresponding characteristics determined upon the Fourier transformation of the time-dependent response. The frequencies and widths of the spectral lines were obtained by fitting to the Lorentz function, while the oscillation amplitudes were evaluated by integral intensities of the spectral lines in the periodogram. For convenience, the integral intensity corresponding to a maximum power is taken to be equal to unity. It should be noted that the direct measurement of the amplitude of coherent oscillations at a fixed delay time provides analogous results, but at a greater error.

In order to study the statistical (fluctuational) properties of phonons generated by ultrashort pulses, it is necessary to use a modified scheme of measurements [24]. The traditional scheme used for the coherent phonon probing ensures access only to the principal moment (mean value) of the phonon distribution, providing for an average value of the coherent amplitude as a function of the delay time. The study of fluctuational properties implies the knowledge of how the coherent amplitude fluctuates with time at a fixed phase determined by the delay time. This is illustrated in Fig. 5 showing essentially a squeezed state in the Heisenberg (Fig. 5a) and Schrödinger (Fig. 5b) representations. In the former case, we vary the detection angle, while in the latter case, the state exhibits evolution with time at a constant detection angle.

For studying the fluctuational properties of phonons, the measurements have to be conducted so as to retain information concerning the particular state realizations. Although each state realization cannot be measured separately, we may compare the statistical sample sets obtained using about 10^5 probing pulses at a fixed phase (delay time). For this purpose, we have repeatedly measured a part of the oscillating response S to calculate $\mu = \langle S\rangle$ and the second-order moment $\sigma^2 = \langle S^2\rangle - \mu^2$ in addition to the first-order moment for each delay time. Figure 6 shows the results of these measurements for Sb. As expected, the mean amplitude determined by this method is identical to the amplitude observed in a traditional experiment. Nevertheless, the dispersion of the coherent amplitude calculated for each time instant provides additional (new) information. In the case when the system studied is prepared in the coherent state, the dispersion characterizing the noise is time-dependent. Note that the noise in the system occurring in the equilibrium state is independent of the time, as is seen for the dispersion at negative delay times (phase angles).

In order to elucidate and explain the time variation of the noise, we have numerically converted the dispersion into the frequency domain with the aid of the Fourier transformation. Figure 6c shows a typical noise power spectrum obtained upon such conversion, in comparison to the power spectrum of the coherent amplitude (Fig. 6b). The two spectra are obviously different, the noise spectrum displaying a component with a frequency equal to the doubled frequency of the coherent phonon. As the pumping power grows, the noise amplitude shows a linear increase, while the spectrum remains virtually unchanged. However, the intensity of the spectral component in the noise spectrum exhibits slight redistribution, whereby the ratio of integral intensities of the high- and low-frequency components tends to increase.

Figures 7–9 show a similar set of data for a polycrystalline Bi film. Bismuth is crystallized in a rhombohedral unit cell, with phonons of the $A_{1g} + E_g$ symmetry allowed. Similar to the case of Sb, these modes were observed in Bi samples measured in the time domain [30, 34]. The intensity of the fully symmetric mode was also markedly higher as compared to that of the nondiagonal modes [34]. The Raman scattering spectra of polycrystalline bismuth films exhibit, similarly to the spectra of Sb, a dominating A_{1g} mode corresponding 2.9 THz (97 cm^{-1}) [31]. However, the amplitude of oscillations in Bi is somewhat lower than that in Sb observed under identical experimental conditions. The time-dependent response of Bi upon the Fourier transformation gives a frequency of the fully symmetric mode equal to 2.95 THz (98 cm^{-1}) that is analogous to the value observed in the Raman spectrum. However, the spectral line width ($\Delta\nu = 0.09$ THz) obtained from measurements in the time domain is smaller, while parameter

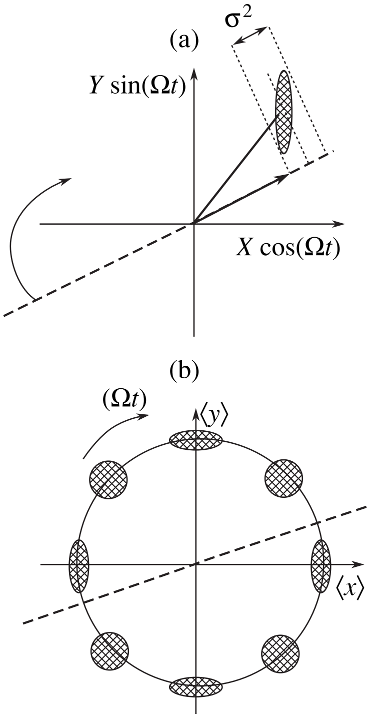


Fig. 5. Schematic diagram illustrating the measurement of coherent phonon noise.

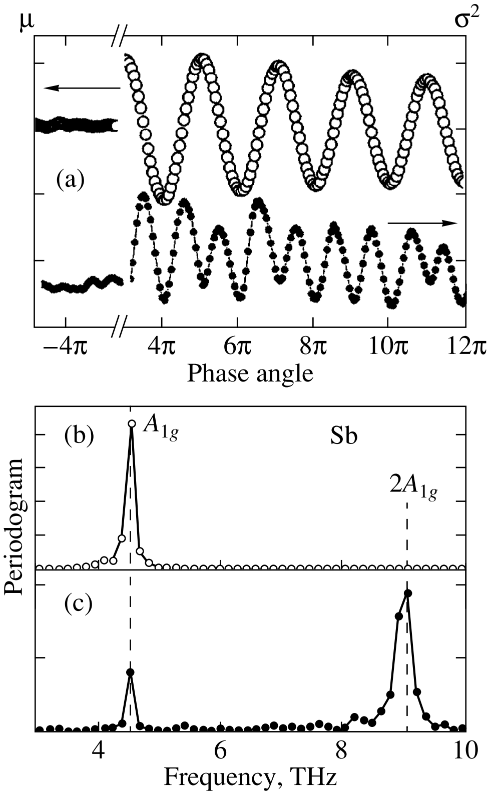


Fig. 6. Coherent phonon oscillations in Sb: the plots of (a) coherent amplitude and its dispersion versus phase angle; (b) coherent amplitude power spectrum; (c) coherent amplitude dispersion spectrum.

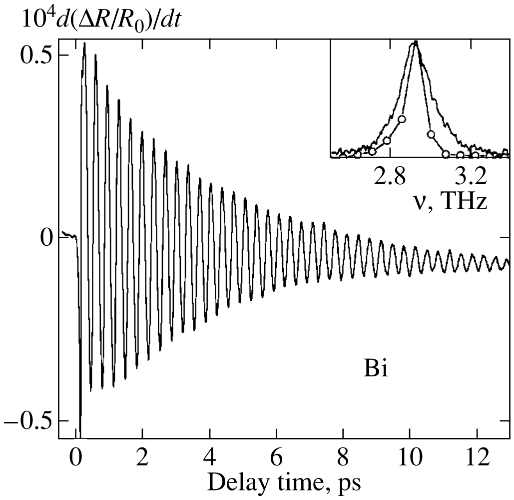


Fig. 7. Time-resolved normalized reflection signal in a Bi sample excited with 30-fs light pulses (the pumping and probing power is 60 and 1.2 mW, respectively). The pumping bean was modulated by shaking retroreflector. The inset shows (open circles) the Fourier transform of the time-resolved response and (solid curve) the spontaneous Fourier scattering spectrum.

Q is accordingly greater ($Q = 32$) than the value extracted from the Raman scattering data ($Q = 17$) [33].

The dependence of the coherent phonon parameters on the pumping power in Bi is identical to that in Sb: the oscillation amplitude linearly varies with the pumping power, while the frequency and the rate of dephasing remain unchanged. It should be noted that, despite a nonstationary character of the observed effect, the spectrum of coherent amplitude in semimetals is stationary. This fact was established by varying the upper and/or lower time limit during the Fourier transformation. The periodograms obtained for small delay times were identical to those observed for large times, which is indicative of the equivalence of the "fast" and "slow" dynamics of the phonon subsystem in semimetals. Nevertheless, it should be noted that the spectrum of coherent oscillations obtained in the case when more than one mode is excited can be time-dependent [35, 36].

3.2. Semiconductors

Besides semimetals crystallizing in a centrosymmetric crystal lattice, coherent phonons may be generated in polar crystals typically represented by semiconductors of the $A^{III}B^{V}$ group, which crystallize in a structure of the zinc blend type [4, 5].

Among semiconductors of the above group, the one most thoroughly investigated is GaAs, in which the time-dependent oscillations assigned to the coherent phonons and/or coupled phonon–plasmon modes were studied depending on the dopant concentration, temperature, and wavelength of exciting laser radiation [5, 37]. The major mechanism responsible for the excitation of

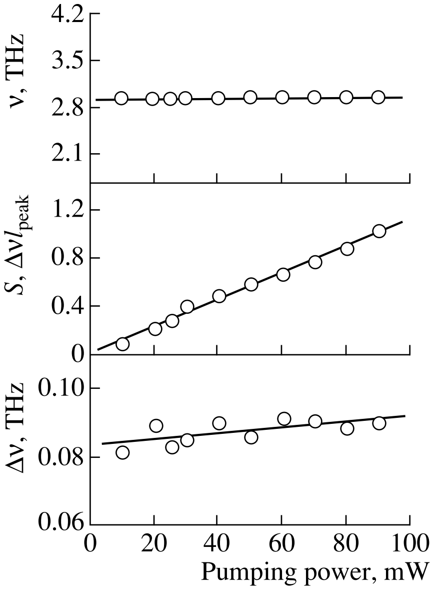


Fig. 8. Plots of the parameters of coherent oscillations versus pumping power for Bi (solid lines show the best linear approximation).

coherent oscillations is believed to consist in the interaction of photons with the electric field generated in a near-surface depleted layer of the semiconductor [37, 38]. Note that this mechanism is capable of initiating only excita-

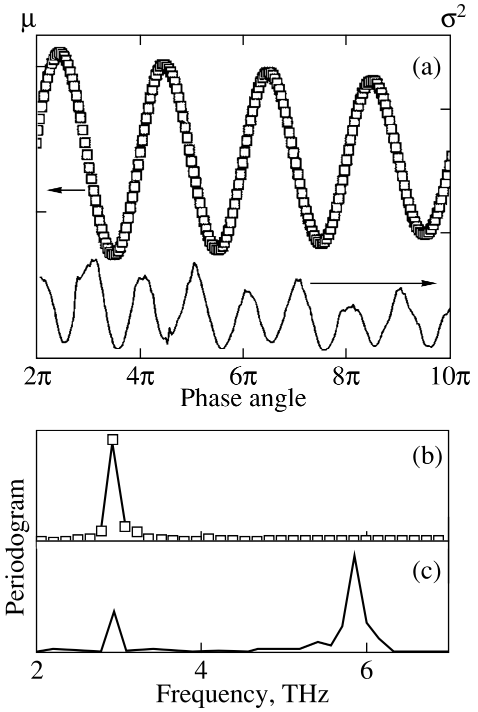


Fig. 9. Coherent phonon oscillations in Bi: the plots of (a) coherent amplitude and its dispersion versus phase angle (calculated using the delay time between exciting and probing pulses); (b) Fourier-transformed coherent amplitude spectrum; (c) coherent amplitude dispersion spectrum.

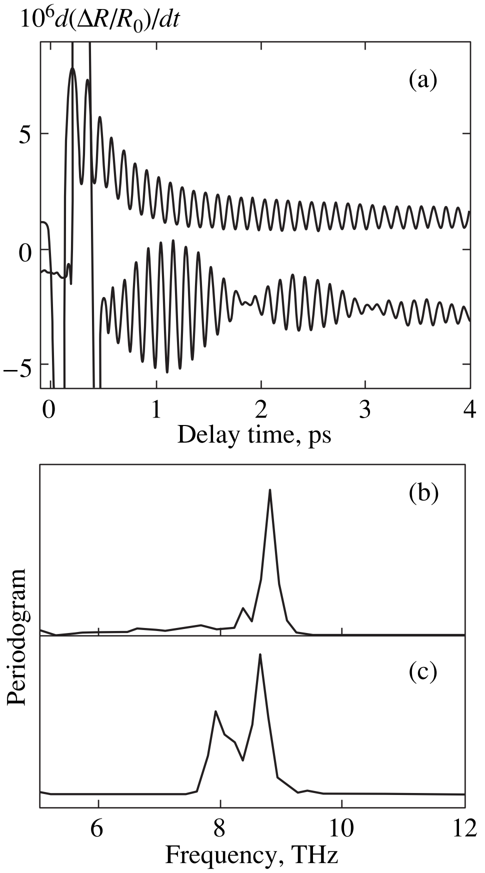


Fig. 10. Coherent phonon oscillations in GaAs: (a) Time-resolved normalized reflection signal in GaAs samples excited with 30-fs light pulses (the pumping and probing power is 80 and 1.1 mW, respectively). The upper and lower curves correspond to the samples doped to $n_e = 1.5 \times 10^{17}$ and 8×10^{17} cm^{-3}, respectively; (b, c) periodograms corresponding to the upper and lower curves in (a).

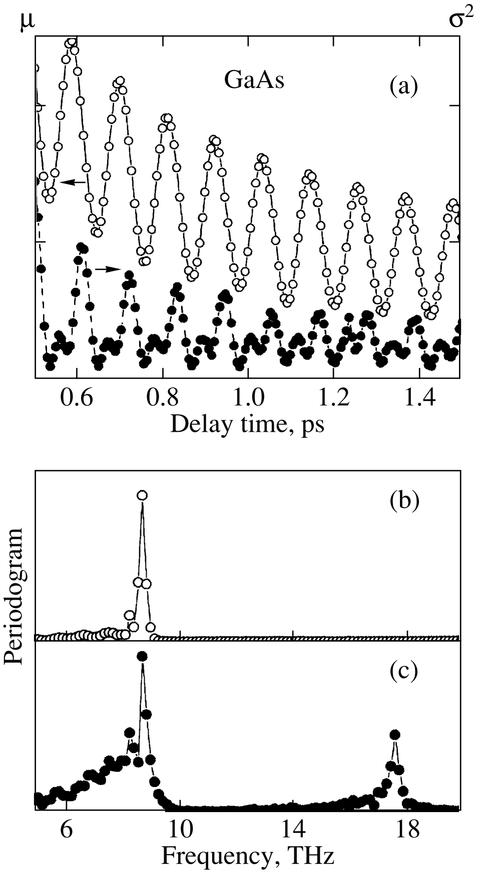


Fig. 11. Coherent phonon oscillations in a GaAs sample with $n_e = 1.5 \times 10^{17}$ cm^{-3}: (a) coherent amplitude and its dispersion versus time delay between exciting and probing pulses; (b, c) periodograms of the coherent amplitude and its dispersion, respectively.

tions carrying the electric field (longitudinal phonons and plasmons), but not the transverse phonons. It should be also noted that the Raman scattering in semiconductors is well studied with respect to variation of the crystal properties under the action of external factors leading to a decrease in the crystal symmetry (the morphic effect) [39].

Figure 10 shows the results obtained for two *n*-GaAs samples with charge carrier concentrations 1.5×10^{17} and 8×10^{17} cm^{-3}. In the former case, the coherent oscillations are observed at a frequency of 8.54 THz (284.6 cm^{-1}) and correspond to the longitudinal optical *LO* phonon [5]. The corresponding coherent oscillation noise spectrum presented in Fig. 10 contains a component with the doubled phonon frequency [26]. The coherent amplitude dispersion reaches maximum at the points of zero coherent amplitude, while the dispersion minima approximately coincide with the amplitude maxima.

The shape of the time-dependent response of GaAs significantly changes with an increase in the level of doping. Even measurements in the time domain show that oscillations in the highly doped sample are related to more than one mode, since the time-dependent response reveals an additional modulation caused by beats between the modes. The Fourier-transformed time-dependent response (Fig. 10c) displays two modes having the frequencies approximately coinciding with the *LO* and *TO* phonon frequencies. It should be noted that both frequencies remain unchanged upon increasing the pumping power. Note also that the *LO* and *TO* modes are mutually orthogonal and, hence, are not coupled. The two modes can be coupled by zero-point (vacuum) oscillations and/or an external field (e.g., plasmons).

In the $A^{III}B^{V}$ group of semiconductors, InSb is characterized by a minimum bandgap width and by a relatively small effective mass of charge carriers. The coherent oscillations in InSb decay rather rapidly and cannot be detected after 4 ps. Figure 12a shows a typical time-resolved optical response and the corresponding dispersion for InSb. The frequency of the coherent oscilla-

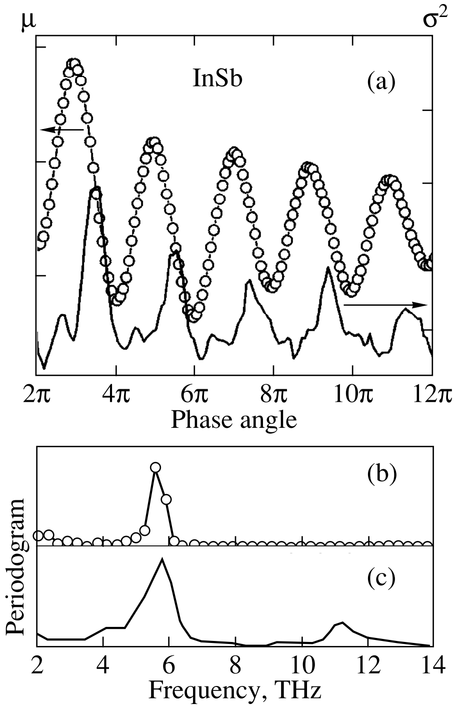


Fig. 12. Coherent phonon oscillations in InSb: (a) coherent amplitude and its dispersion versus time delay between exciting and probing pulses; (b, c) periodograms of the coherent amplitude and its dispersion, respectively.

tions (5.6 THz or 186 cm^{-1}), determined upon the Fourier transformation of the time-dependent response presented in Fig. 12b, coincides with the *LO* phonon frequency [40]. The fluctuational properties of *LO* phonons in InSb are illustrated in Fig. 12c, which shows that the noise of the coherent amplitude is time-dependent. The periodogram displays a peak at a frequency equal to the doubled *LO* phonon frequency.

Indium phosphide (InP) belongs to the least studied semiconductors of the $\text{A}^{\text{III}}\text{B}^{\text{V}}$ group with respect to coherent phonons. At present, the coherent oscillations in InP have been detected only in samples of the *p*-type and identified at large delay times as related to the *LO* phonon [41]. Figure 13 shows the time-resolved optical response for a series of InP samples of both the *p*- and *n*-type. As is seen, the coherent oscillations are most pronounced in samples of the *p*-type. Let us consider in more detail the data obtained for a *p*-InP sample with $n_p = 10^{18}$ cm^{-3} and for an *n*-InP sample with $n_e = 7 \times 10^{16}$ cm^{-3}. Although the signal polarity in most experiments was such as depicted in Fig. 13, some experimental runs performed under identical conditions showed signals of inverted polarity (Fig.14). This fact indicates that the initial phase is bistable and may acquire fixed values shifted by π. This behavior of the initial phase is hardly probable for the displacive mechanism or for the coherent phonon generation by an electric field in the near-surface depleted layer. However, a mechanism based on induced Raman light scattering allows this phenomenon to be explained by relating the initial phase to the Stokes or anti-Stokes scattering channels. The phase bistability was previously observed in YBa_2CuO_8 crystals at large delay times [42].

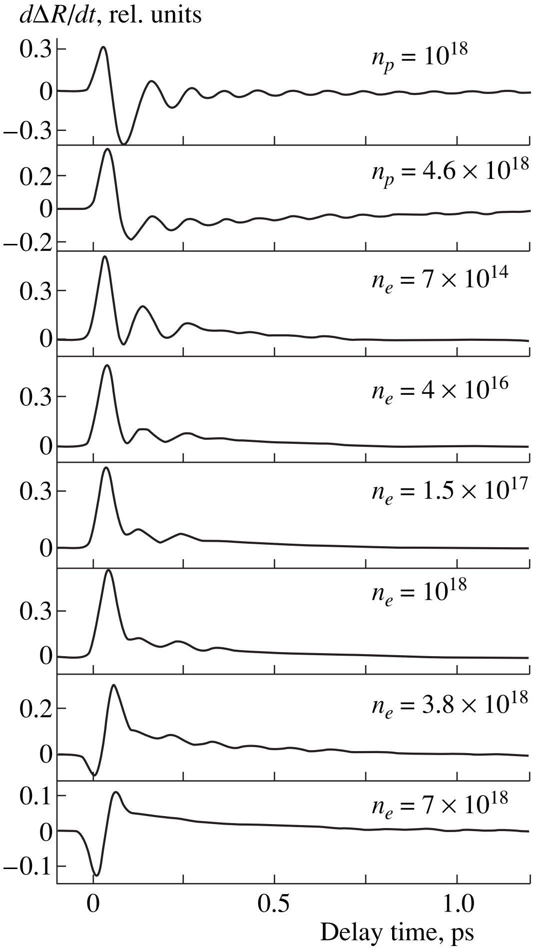


Fig. 13. Time-resolved reflection signal in a series of InP samples excited with 35-fs light pulses (the pumping and probing power is 40 and 1.2 mW, respectively). The type and concentration of dopants are indicated at the curves.

3.3. Superconductors

At present, coherent phonons have been detected in superconductors belonging to all principal classes [27, 42, 44– 46]. The $\text{YBa}_2\text{Cu}_3\text{O}_{7-\delta}$ system, where the phenomenon was observed for the first time, still remains the most thoroughly studied in this respect [43–46]. Therefore, it would be most expedient to consider the phonon characteristics reported for the phonon subsystem of $\text{YBa}_2\text{Cu}_3\text{O}_{7-\delta}$ and compare the results obtained in the time and frequency domains. The crystal lattice symmetry of $\text{YBa}_2\text{Cu}_3\text{O}_{7-\delta}$, belonging to the D_{2h} point symmetry group, allows five fully symmetric A_g phonons at point Γ of the Brillouin zone [47]. Two

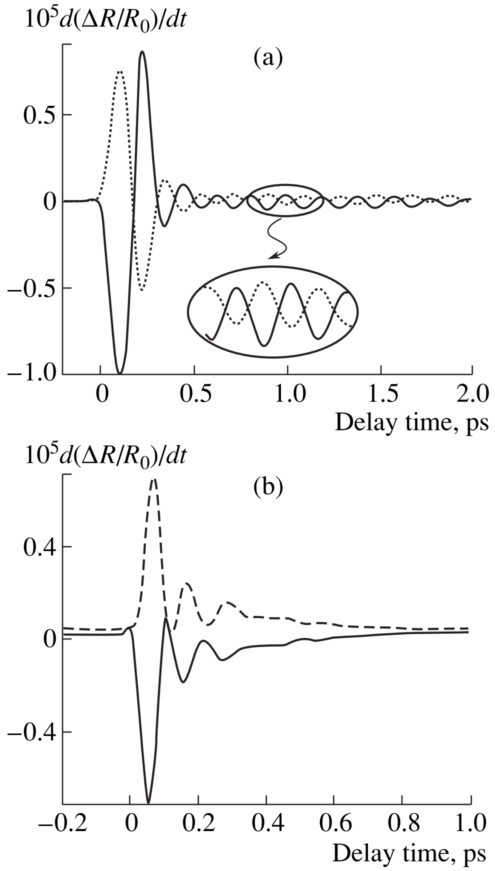


Fig. 14. Schematic diagrams illustrating the initial phase instability for InP with $n_k = 10^{18}$ cm^{-3} (a) and $n_e = 7 \times 10^{16}$ cm^{-3} (b).

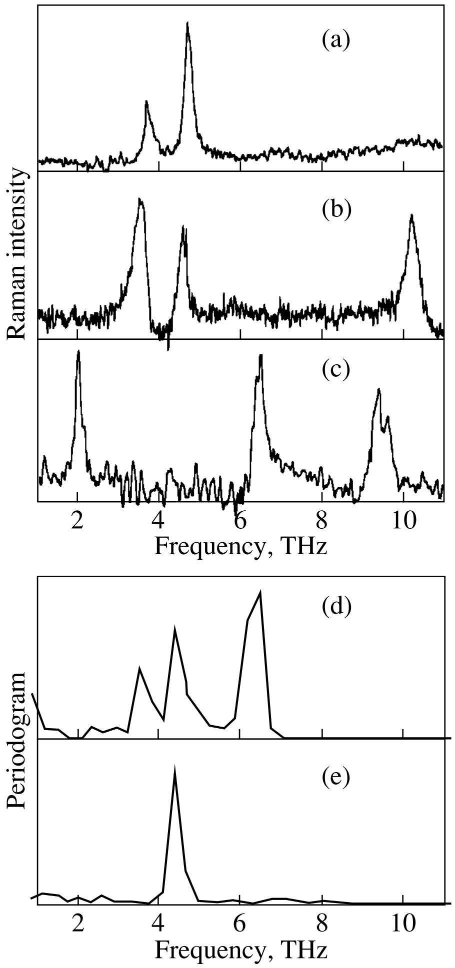


Fig. 15. Coherent phonon oscillations in a $YBa_2Cu_3O_{7-\delta}$ single crystal with T_e = 92 K: (a–c) Raman spectra of (a) A_g phonons (*zz* polarization), (b) A_g phonons ($x'x'$ polarization), and (c) $B_{3g} + B_{2g}$ phonons ($zy + zx$ polarization); (d, e) time-resolved response from (d) *ac* and (d) *ab* (basal) single crystal faces.

of these A_g phonons, detected using measurements in the time domain, represent the low-frequency phonons generated by *z*-displacements of Ba and Cu ions at frequencies of 3.6 and 4.5 THz (120 and 150 cm^{-1}), respectively. The copper phonon dominates in the femtosecond response at temperatures above T_c, whereas the barium phonon predominantly contributes at lower temperatures [44, 45].

Similarly to the case of semimetals, the initial hypothesis for the explanation of the pattern of coherent phonon excitation in $YBa_2Cu_3O_{7-\delta}$ was the displacive mechanism [44]. Using this model, modified and refined so as to apply to the superconducting state [48], it is possible to explain both the observed frequencies of the phonon modes and the change in magnitude and polarity of the response upon the material transition to the superconducting state. However, the results of measurements performed for various single crystal faces showed that the ultrashort pulses produce excitation of both the fully symmetric and nondiagonal modes. This is demonstrated in Fig. 15 showing the periodograms of time-resolved optical response from *ab* and *ac* faces. A comparison of these patterns to the spontaneous Raman scattering spectra of A_g and B_{2g} (B_{3g}) symmetry suggests the presence of a mode with frequency 6.3 THz (210 cm^{-1}). Based on this mode frequency and the comparison with a spontaneous Raman spectrum of the untwinned single crystal [47], this phonon is readily attributed to displacements of the bridging oxygen atom and assigned the B_{3g} symmetry. It is more difficult to explain the presence of nondiagonal modes within the framework of the displacive excitation of coherent phonons; however, this task is readily solved proceeding from the alternative mechanism based on the stimulated Raman light scattering.

Our analysis of the fluctuational properties of coherent phonons in $YBa_2Cu_3O_{7-\delta}$ will be restricted to the fully symmetric phonon mode. Figure 16 shows a typical time-resolved response obtained from the basal *ab* plane of an $YBa_2Cu_3O_{7-\delta}$ crystal at room temperature. The presence of oscillations with a frequency of 4.52 THz (150.6 cm^{-1}) is evidence that the coherent

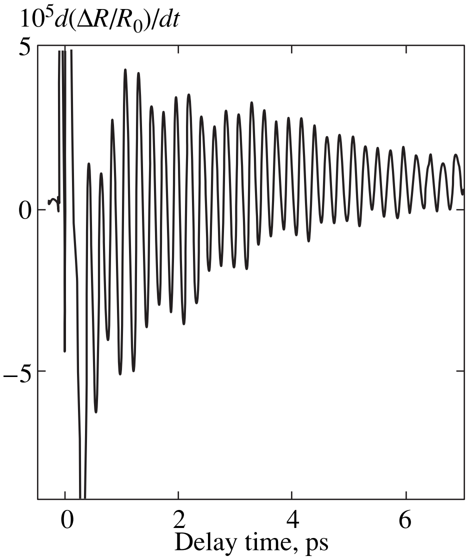


Fig. 16. Time-resolved normalized reflection signal in an YBa$_2$Cu$_3$O$_{7-\delta}$ sample excited with 30-fs light pulses (the pumping and probing power is 80 and 1 mW, respectively). The pumping bean was modulated by shaking retroreflector.

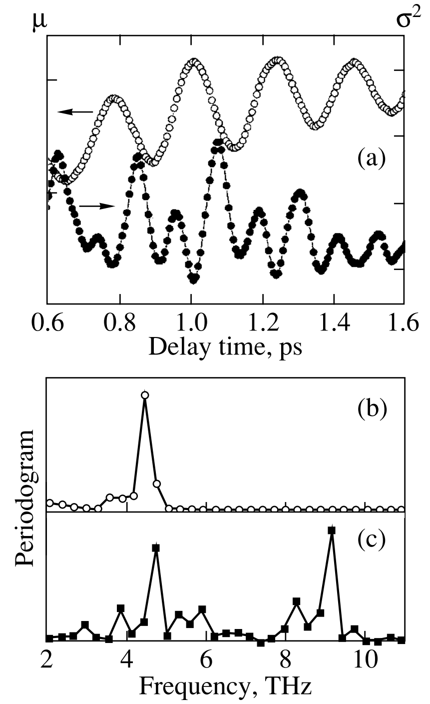


Fig. 17. Coherent phonon oscillations in YBa$_2$Cu$_3$O$_{7-\delta}$: (a) coherent amplitude and its dispersion versus time delay between exciting and probing pulses; (b, c) periodograms of the coherent amplitude and its dispersion, respectively.

background corresponds to a Raman-active mode with the A_g symmetry, generated by the displacements of Cu ions. A comparison to the Raman scattering spectrum shows that the characteristic time of dephasing at room temperature is greater than that of the energy relaxation [27, 44]. The experimental geometry and a relatively high temperature minimize the contribution of the other fully symmetric mode (3.6 THz or 120 cm^{-1}), in which the dominating role belongs to the displacements of Ba ions. Note that the room-temperature Raman spectrum displays both modes with approximately equal intensities (the relative Raman peak intensity depends on the particular polarization, but none of these leads to the absence of the low-energy mode) [35].

In order to obtain information on the statistical properties of coherent phonons in superconductors, we have repeatedly measured the part of the oscillatory response depicted in Fig. 16. Using this set of data, we calculated the mean values and dispersions for each delay time n the interval from –500 to 2000 fs. These data (for a limited range of positive delay times) are presented in Fig. 17. It is obvious that, similarly to the case of semimetals and semiconductors, the coherent phonon fluctuations in the superconductor studied depend on the phase determined by the delay between exciting and probing pulses. The dispersion of coherent phonons is maximum at the points of the extrema of the coherent amplitude [27]. The noise periodogram displays a peak at the frequency equal to the doubled coherent phonon frequency (Figs. 17b and 17c).

4. DISCUSSION OF EXPERIMENTAL RESULTS

A number of phenomenological models were suggested in order to describe the coherence arising in a crystal lattice as a result of the interaction between short laser pulses and the crystal [4, 5, 33]. These models explained the main experimental facts observed in numerous investigations. However, the physical meaning of the term “coherent amplitude” employed in these models cannot be explained without recourse to a microscopic theory. Developed by Kuznetsov and Stanton [49], such a theory showed that, within the framework of the microscopic description, the coherent oscillations are related to a macroscopic occupation of the phonon mode with $\mathbf{q} = 0$, this circumstance making the situation much like that observed for the Bose condensation.

Note that the Bogolyubov transformation used in the description of the superfluidity phenomenon can be also applied to squeezed states [50]. Moreover, the coherent phonon excitation, considered as a realization of a phase transition, may help us qualitatively understand the phenomenon of squeezing. As is known, the squeezing increases in systems occurring at the points of bifurcation between various dynamic regimes because of the quantum fluctuations growing for the variable losing stability at such a point [51]. As a result, fluctuations of the conjugated variable decrease by virtue of the phase volume conservation.

An analysis of the results of measurements in the time and frequency domains, allows us to conclude that the frequencies (energies) of thermal and coherent phonons coincide to within the experimental error. The error in determining the frequency is greater than the error of measurements in the time domain. The width of spectral lines in the spontaneous Raman spectra is determined by the phonon decay, which depends on the system interaction with the thermal reservoir. As a rule, a dominating contribution to this decay is due to the anharmonic decay of the optical phonon into two phonons with a total momentum of zero [3]. The spectral line width in the Fourier-transformed time-dependent response is determined by a characteristic time of the phase memory; for any classical state, the time of dephasing must be shorter than or equal to the energy relaxation time. Indeed, the dephasing may proceed by both inelastic and elastic collisions in the phonon gas, and the total number of inelastic and elastic collisions cannot be smaller than the number of elastic events alone.

A difference in the spectral line shape parameters revealed by the data of measurements in the time and frequency domain is by no means unexpected, since these methods measure essentially different states of the phonon system. In the case of femtosecond pulses, the initial system is prepared by the pumping pulse in the coherent state, for which the phase is a well defined quantity; the spontaneous Raman scattering probes the system in a thermal state, in which the phase is uncertain. It should be noted that use of the amplitude Fourier transformation (instead of the power spectrum) for the time-dependent response apparently cannot eliminate the difference in shape of the spectral lines measured in the time and frequency domains, since various crystal systems exhibit both wider and narrower lines as compared to those obtained in the frequency spectra.

An analysis of the dependence of the coherent amplitude on the pumping pulse power leads to an important conclusion concerning a nonclassical character of the state created by ultrashort laser pulses. Indeed, the decrement of a classical state is a function of the intensity of excitation. A typical example is offered by the thermal state where the excitation lifetime is a function of the temperature (the excitation factor). In optics, this is manifested by the temperature dependence of the phonon line width measured in the Raman spectra. In contrast, coherent phonons exhibit independence of the dephasing time of the degree of excitation, which is related to the fact that only the Poisson distribution is characterized by a constant ratio of the mathematical expectation to dispersion [13].

A remarkable and rather surprising property of coherent phonons is the noise, which depends on the phase of the coherent oscillations. A possible interpretation of the phase-dependent noise employs the concept of squeezed phonon states created during the interaction of ultrashort laser pulses with crystal. This interpretation is based on the fact that the phase dependence of the noise is a unique indication of the squeezing [13], since any other state (quantum or classical) is characterized by the phase-independent noise and, hence, the uncertainty profile can be described by a circle.

The conclusion that measurements in the time domain provide information on the quantum fluctuations (i.e., the phonon behaves as a quantum, rather than classical, object) is confirmed by the following considerations. Classical behavior dominates in the energy range determined by the inequality $k_BT \le \hbar\omega$, which stipulates that the thermal fluctuations are small as compared to the quantum fluctuations. However, this condition is valid only for the results of measurements averaged over time τ exceeding the characteristic time τ^* of the system relaxation ($\tau \gg \tau^*$). However, the use of femtosecond probing pulses implies that $\tau \ll \tau^*$. During such a short period of time, the system energy exchange with the reservoir is markedly smaller than $k_{\rm B}T$. In this case, the condition for the quantum character manifestation should be modified and expressed as $k_{\rm B}T \le \hbar\omega\tau^*/\tau$ [52]. This condition is satisfied in the experiments described above.

At present, a particular mechanism responsible for the squeezing is unclear. Apparently, establishing this mechanism would require elucidating the process of the coherent phonon excitation. If the coherent phonons are created by a mechanism related to the induced Raman scattering (or some modification) [4], the phonon field squeezing can in fact be realized under certain conditions, since the inelastic light scattering is a parametric process. Note that the noise in most experiments is maximum at the points of zero amplitude (and maximum, at the points of extrema) of the measured signal. In the case of modulation with the aid of a shaker, the zero points correspond to the points of rotation of the oscillatory wave packet trajectory.

Investigation of the noise phase behavior depending on the pulse duration may allow us to determine the relative contributions of the topological and dynamic squeezing components [17]. At present, in addition to explaining the phase-dependent noise as resulting from the phonon field squeezing, we cannot exclude the possibility that such noise may also be related to the bistability of the initial phase of coherent phonons, which can acquire fixed values shifted by π. Should this be the case, the phase-dependent noise has nothing in common with the squeezed phonon states, but can still shed light on the nature of the initial phase of the phonon condensate. If the initial phase bistability is not the major source of the noise, the squeezed phonon states are in fact registered in the experiment.

5. CONCLUSION

We have compared the thermal and coherent phonons created in several typical representatives of semimetals, semiconductors, and superconductors using the results of optical measurements in the frequency and time domains. We have demonstrated that the energies of thermal and coherent phonons coincide, whereas the phase information in the coherent state is lost within the time different from that of the energy relaxation. It was established that coherent phonons exhibit phase-dependent fluctuational properties. An analysis of the obtained experimental data indicates that the final interpretation of the phase-dependent noise requires determining the mechanism of the coherent phonon excitation. Additional information necessary for establishing the squeezing mechanism and its particular realizations can be obtained by measuring the time-resolved response, determining the frequency-resolved optical field components at the Stokes and anti-Stokes frequencies, and studying their statistical properties.

ACKNOWLEDGMENTS

The author is grateful to Kansai Advanced Research Center (Japan) and Alexander-von-Humboldt Foundation (Germany) for support of this work.

REFERENCES

1. W. Heisenberg, Z. Phys. **4**, 879 (1925).
2. J. Shah, *Ultrafast Spectroscopy of Semiconductors and Semiconductors Nanostructures* (Springer-Verlag, Berlin, 1996).
3. B. J. Berne and R. Pecora, *Dynamic Light Scattering* (Wiley, New York, 1976).
4. R. Merlin, Solid State Commun. **102**, 207 (1997).
5. T. Dekorsy, G. C. Cho, and H. Kurz, in *Light Scattering in Solids VIII*, Ed. by M. Cardona and G. Guenterodt (Springer-Verlag, Berlin, 2000), p. 169.
6. S. Ruhman, A. G. Joly, and K. A. Nelson, IEEE J. Quantum Electron. **24**, 460 (1988); L. Dhar, J. A. Rogers, and K. A. Nelson, Chem. Rev. **94**, 157 (1994).
7. H. Z. Cummins, in *Proceedings of the International School of Physics "Enrico Fermi"*, Ed. by R. J. Glauber (Academic, New York, 1969), p. 247.
8. O. V. Misochko, M. Tani, K. Sakai, *et al.*, Phys. Rev. B **58**, 12789 (1998).
9. A. V. Belinskiĭ and M. V. Chekhova, Zh. Éksp. Teor. Fiz. **108**, 1956 (1995) [JETP **81**, 1067 (1995)].
10. *Ultrashort Light Pulses*, Ed. by S. Shapiro (Springer-Verlag, Heidelberg, 1977; Mir, Moscow, 1981).
11. D. Stoler, Phys. Rev. D **1**, 3217 (1970).
12. R. E. Slusher, L. W. Hollberg, B. Yourke, *et al.*, Phys. Rev. Lett. **55**, 2409 (1985).
13. R. Loudon and P. L. Knight, J. Mod. Opt. **34**, 709 (1987); R. Meystre and M. Sargent III, *Elements of Quantum Optics* (Springer-Verlag, Berlin, 1991); V. P. Bykov, Usp. Fiz. Nauk **161** (10), 145 (1991) [Sov. Phys. Usp. **34**, 910 (1991)]; B. J. Dalton, Z. Ficek, and S. Swain, J. Mod. Opt. **46**, 379 (1999).
14. M. Artoni and J. F. Birman, Opt. Commun. **104**, 319 (1994); S. Ghoshal and A. Chatterjee, Phys. Lett. A **223**, 195 (1996).
15. B. K. Chakraverty, D. Feinberg, Z. Hang, and M. Avignon, Solid State Commun. **64**, 1147 (1987); D. L. Lin and H. Zheng, J. Appl. Phys. **64**, 5905 (1988); A. N. Das and S. Sil, Physica C (Amsterdam) **161**, 325 (1989); C. F. Lo, E. Manousakis, R. Sollie, and Y. L. Wang, Phys. Rev. B **50**, 418 (1994); T. Hakioglu, V. A. Ivanov, A. S. Shumovsky, and B. Tanatar, Phys. Rev. B **51**, 15363 (1995).
16. K. Wodkievicz, P. L. Knight, S. J. Buckle, and S. Barnett, Phys. Rev. A **35**, 2567 (1987).
17. J. Janszky and An. V. Vinogradov, Phys. Rev. Lett. **64**, 2771 (1990); J. Janszky, An. V. Vinogradov, I. A. Walmsley, and J. Mostovski, Phys. Rev. A **50**, 732 (1994).
18. B. Yurke, Phys. Rev. Lett. **60**, 2476 (1988).
19. A. S. Shumovsky and B. Tanatar, Phys. Lett. A **182**, 411 (1993).
20. G. A. Garrett, A. G. Rojo, A. K. Sood, *et al.*, Science **275**, 1638 (1997).
21. X. Hu and F. Nori, Phys. Rev. B **53**, 2419 (1996); Phys. Rev. Lett. **76**, 2294 (1996).
22. T. J. Dunn, J. N. Sweetser, I. A. Walmsley, and C. Radzewicz, Phys. Rev. Lett. **70**, 3388 (1993).
23. G. A. Garrett, J. F. Whitaker, A. K. Sood, and R. Merlin, Opt. Express **1**, 385 (1997).
24. O. V. Misochko, K. Kisoda, K. Sakai, and S. Nakashima, Appl. Phys. Lett. **76**, 961 (2000).
25. A. Bartels, T. Dekorsy, and H. Kurz, Phys. Rev. Lett. **84**, 2981 (2000).
26. O. V. Misochko, K. Sakai, and S. Nakashima, Phys. Rev. B **61**, 11225 (2000).
27. O. V. Misochko, Phys. Lett. A **269**, 97 (2000).
28. Y. R. Shen and N. Bloembergen, Phys. Rev. **137**, A1786 (1965).
29. O. V. Misochko, Fiz. Tverd. Tela (St. Petersburg) **40**, 998 (1998) [Phys. Solid State **40**, 914 (1998)].
30. T. K. Cheng, S. D. Brorson, S. Kazeroonian, *et al.*, Appl. Phys. Lett. **57**, 1004 (1990); T. K. Cheng, J. Vidal, H. J. Zeiger, *et al.*, Appl. Phys. Lett. **59**, 1923 (1991).
31. J. S. Lanin, J. M. Colleja, and M. Cardona, Phys. Rev. B **12**, 585 (1975).
32. G. A. Garrett, T. F. Albrecht, J. F. Whitaker, and R. Merlin, Phys. Rev. Lett. **77**, 3661 (1996).
33. H. J. Zeiger, J. Vidal, T. K. Cheng, *et al.*, Phys. Rev. B **45**, 768 (1992).
34. M. Hase, K. Mizoguchi, H. Harima, *et al.*, Appl. Phys. Lett. **69**, 2474 (1996).
35. O. V. Misochko, K. Kisoda, K. Sakai, and S. Nakashima, Phys. Rev. B **61**, 4305 (2000).
36. S. Hunsche, K. Wienecke, T. Dekorsy, and H. Kurz, Phys. Rev. Lett. **75**, 1815 (1995).
37. G. C. Cho, W. Kutt, and H. Kurz, Phys. Rev. Lett. **65**, 764 (1990); T. Dekorsy, H. Kurz, X. Q. Zhou, and K. Ploog, Appl. Phys. Lett. **63**, 2899 (1993).
38. A. V. Kuznetsov and C. J. Stanton, Phys. Rev. B **51**, 7555 (1995).
39. G. Abstreiter, M. Cardona, and A. Pinczuk, in *Light Scattering in Solids*, Ed. by M. Cardona and G. Guntherodt (Mir, Moscow, 1986), Iss. IV.
40. V. P. Gnezdilov, D. J. Lockwood, and J. W. Webb, Phys. Rev. B **48**, 11234 (1993).

41. S. Nakashima, K. Mizoguchi, H. Harima, and K. Sakai, J. Lumin. **76/77**, 6 (1998).
42. O. V. Misochko, K. Sakai, and S. Nakashima, Physica C (Amsterdam) **329**, 12 (2000).
43. J. M. Chwalek, C. Uher, J. F. Whitaker, *et al.*, Appl. Phys. Lett. **58**, 980 (1991).
44. W. Albrecht, Th. Kruze, and H. Kurz, Phys. Rev. Lett. **69**, 1451 (1992).
45. O. V. Misochko, K. Kisoda, H. Harima, *et al.*, Physica C (Amsterdam) **320**, 213 (1999).
46. O. V. Misochko, Fiz. Tverd. Tela (St. Petersburg) **42**, 1169 (2000) [Phys. Solid State **42**, 1204 (2000)].
47. V. D. Kulakovskiĭ, O. V. Misochko, and V. B. Timofeev, Fiz. Tverd. Tela (Leningrad) **31** (9), 220 (1989) [Sov. Phys. Solid State **31**, 1599 (1989)].
48. I. I. Mazin, A. I. Liechtenstein, O. Jepsen, *et al.*, Phys. Rev. B **49**, 9210 (1994).
49. A. V. Kuznetzov and C. J. Stanton, Phys. Rev. Lett. **73**, 3243 (1994).
50. M. M. Nieto, in *Frontiers of Nonequilibrium Statistical Physics*, Ed. by G. T. Moore and M. O. Scully (Plenum, New York, 1986), p. 287.
51. K. N. Alekseev and D. S. Priĭmak, Zh. Éksp. Teor. Fiz. **113**, 111 (1998) [JETP **86**, 61 (1998)].
52. V. B. Braginsky and F. Ya. Khalili, *Quantum Measurement*, Ed. by K. S. Thorne (Cambridge Univ. Press, Cambridge, 1992).

Translated by P. Pozdeev

Journal of Experimental and Theoretical Physics, Vol. 92, No. 2, 2001, pp. 260–266.
From Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 119, No. 2, 2001, pp. 301–308.

SOLIDS
Electronic Properties

Coulomb Effects in Spatially Separated Electron and Hole Layers in Coupled Quantum Wells[¶]

L. V. Butov[a, b], A. Imamoglu[a, c], K. L. Campman[a], and A. C. Gossard[a]

[a]*Department of Electrical and Computer Engineering and Center for Quantized Electronic Structures (QUEST), University of California, Santa Barbara, CA 93106, USA*

[b]*Institute of Solid State Physics, Russian Academy of Sciences, Chernogolovka, 142432 Moscow oblast, Russia*

[c]*Department of Physics, University of California, Santa Barbara, CA 93106, USA*

Received June 7, 2000

Abstract—We report on the (magneto-) optical study of many-body effects in spatially separated electron and hole layers in GaAs/Al$_x$Ga$_{1-x}$As coupled quantum wells (CQWs) at low temperatures (T = 1.4 K) for a broad range of electron–hole (*e–h*) densities. Coulomb effects were found to result in an enhancement of the indirect (interwell) photoluminescence (PL) energy with increasing the *e–h* density both for a zero magnetic field and at high fields for all Landau level transitions; this is in contrast to the electron–hole systems in single QWs where the main features are explained by the band-gap renormalization resulting in a reduction of the PL energy. The observed enhancement of the ground state energy of the system of the spatially separated electron and hole layers with increasing the *e–h* density indicates that the real space condensation to droplets is energetically unfavorable. At high densities of separated electrons and holes, a new direct (intrawell) PL line has been observed: its relative intensity increased both in PL and in absorption (measured by indirect PL excitation) with increasing density; its energy separation from the direct exciton line fits well to the X^- and X^+ binding energies previously measured in single QWs. The line is therefore attributed to direct multiparticle complexes. *© 2001 MAIK “Nauka/Interperiodica”.*

1. INTRODUCTION

Many-body interactions in neutral electron–hole (*e–h*) systems in semiconductor quantum wells (QWs) lead to renormalization effects [1] that were extensively studied in single quantum wells (SQWs) in the past. In particular, the main experimental features were explained by the band-gap renormalization that results in a reduction of the ground state energy with increasing the *e–h* density (see [2] and references therein).

In this paper, we study many-body effects in a system of spatially separated electron and hole layers at zero and finite magnetic fields perpendicular to the QW plane. Due to the long radiative recombination times, the *e–h* temperatures can be much lower in this system than those achieved in single-layer *e–h* systems. In particular, this unique property may enable the observation of a number of predicted collective phenomena [3–10]. The major difference of many-body effects in spatially separated electron and hole layers compared to single-layer *e–h* systems is the asymmetry between the *e–e*, and *e–h* interactions. In a set of papers, this asymmetry has been predicted to result in the instability of the uniform exciton phase at low temperatures [5, 6, 10]. In particular, condensation to an exciton liquid has been predicted for small interlayer separation: for $1.1a_2 < d < 1.9a_2$, the liquid was predicted to be metastable, while for $d < 1.1a_2$, the liquid was predicted to be in the ground state ($a_2 = \hbar^2\varepsilon/2me^2$ is the 2D exciton radius, ε is the dielectric constant, m is the reduced exciton mass, and $a_2 \sim 6.5$ nm for GaAs QWs) [10]. On the contrary, in another set of papers, the repulsive interaction between the indirect (interwell) excitons at low densities and the electrostatic term originating from the electric field between the separated electron and hole layers at high densities was predicted to stabilize the uniform phase in the system of separated electron and hole layers [7–9].

The spatially separated *e–h* system with the photoexcitation-controlled *e–h* density is realized in electric field tunable coupled quantum wells (CQWs) (see [11] and references therein). The effects of exciton–exciton interactions at low exciton densities ($\lesssim 10^{10}$ cm^{-2}) were studied earlier [12]: an enhancement of the exciton energy with density both at zero and finite magnetic fields has been observed and interpreted in terms of the net repulsive interaction between indirect excitons (which are dipoles oriented in the *z*-direction).

In this paper, we report on the experimental study of the system of spatially separated electron and hole layers in GaAs/Al$_x$Ga$_{1-x}$As CQW in the broad range of *e–h* densities, up to the maximum possible *e–h* densities corresponding to the complete screening of the external electric field in the *z*-direction (this maximum density depends on the applied electric field and reaches more than 2×10^{11} cm^{-2} for the present experiments). The *e–h* density was controlled by the excita-

---

[¶] This article was submitted by the authors in English.

tion density and by the excitation energy through the absorption variation. The maximum density *e–h* system with the minimum effective temperature was achieved at the excitation resonant to the direct (intrawell) exciton states. The experimental data suggest that the system of separated electrons and holes is uniform with the ground state energy increasing with the *e–h* density; these data, therefore, do not support the predicted condensation to the liquid state in the entire range of *e–h* densities that we studied.

The electric-field-tunable n^+–i–n^+ GaAs/Al$_x$Ga$_{1-x}$As CQW structures were grown by molecular beam epitaxy. A sketch of the band diagram of the structures is shown in the inset to Fig. 1. The *i*-region consists of two 8-nm GaAs QWs separated by a 4-nm Al$_{0.33}$Ga$_{0.67}$As barrier and surrounded by two 200-nm Al$_{0.33}$Ga$_{0.67}$As barrier layers. The n^+-layers are Si-doped GaAs with $N_{\mathrm{Si}} = 5 \times 10^{17}$ cm^{-3}. The second sample has the same design, except for the QW widths that are equal to 15 nm. The data throughout the paper refers to the 8–4–8-nm CQW sample if not specified. The electric field in the *z*-direction is monitored by the external gate voltage V_g applied between n^+-layers (see [12] for details).

Because the electron Fermi level in the n^+-GaAs layers is considerably below the electron energies in the GaAs QWs, the QWs are nominally empty in the absence of photoexcitation (the concentration of the residual impurities in the QW region is unknown; however, it is certainly below the Mott density to provide free electron or hole gases in the QWs and below the density of photoexcited carriers in the CQWs studied). In most of the experiments, carriers were photoexcited by a tunable cw Ti : Sapphire laser with photon energy considerably below the Al$_{0.33}$Ga$_{0.67}$As barrier energy. Possible deviations from the charge neutrality occurring in the CQW electron–hole system because of different collections of electrons and holes photoexcited in the barrier layers are minimized.

To minimize the effect of the mesa heating, we worked with the mesa area 200 × 200 μm^2, which was much smaller than the sample area of about 4 mm^2. In addition, the bottom of the sample was soldered to a metal plate. The excitation was modulated with a dark-to-light ratio of about 15. The measurements were performed in a Spectromag cryostat with the bath temperature T_{bath} = 1.4 K. The PL spectrum was measured using a charge-coupled-device camera.

2. COULOMB EFFECTS IN DIRECT AND INDIRECT PHOTOLUMINESCENCE

The separation of electrons and holes in different QWs (the indirect regime) is realized by applying a finite gate voltage that fixes the external electric field in the *z*-direction $F = V_g/d_0$, where d_0 is the *i*-layer width. The excitation density dependence of the PL spectrum in the indirect regime is shown in Fig. 1. The excitation energy E_{ex} = 1615 meV is sufficiently below the barrier energy to ensure the *e–h* photoexcitation directly in QWs. The direct and indirect transitions are identified by the PL kinetics and gate voltage dependence: the direct PL line has a short PL decay time and its position is practically independent of V_g, while the indirect PL line has a long PL decay time and shifts to lower energies with increasing V_g (in the low-density exciton regime, the shift magnitude is given by eFd, where $d \approx$ 11.5 nm is close to the mean separation between the electron and hole layers) [11, 12].

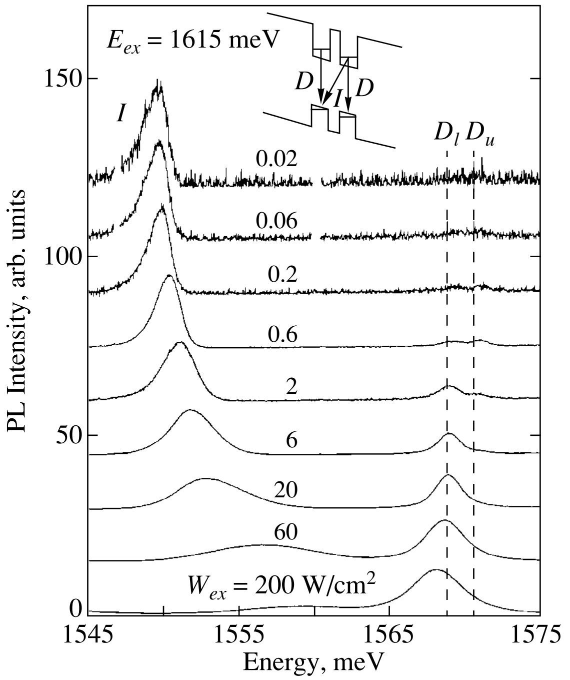


Fig. 1. Excitation density dependence of the PL spectrum at T_{bath} = 1.4 K, V_g = 1 V, and E_{ex} = 1615 meV. The dashed lines are a guide for the eyes. Inset: schematic band diagram of the GaAs/Al$_x$Ga$_{1-x}$As CQW structure under applied gate voltage; the direct (*D*) and indirect (*I*) transitions are shown by arrows.

Figure 1 shows that the indirect PL line monotonically shifts to higher energies with increasing the *e–h* density; this corresponds to an increase of the ground-state energy of the spatially separated *e–h* system. At high *e–h* densities, the energy shift is determined (1) by the exchange and correlation energies, which results in a reduction of the energy [1], and (2) by the electric field between the separated electron and hole layers, which partially compensates for the external electric field and thereby results in an increase of the energy [7, 9]. The latter contribution to the nonlinear energy shift is a unique feature of the system of spatially separated electron and hole layers and can be estimated using the plate capacitor formula $\delta E = 4\pi n_{eh} e^2 d/\varepsilon$, where n_{eh} is the *e–h* density.

The observed increase of the ground state energy of the spatially separated *e–h* system is opposite to the case of *e–h* plasma in SQWs, where exchange and cor-

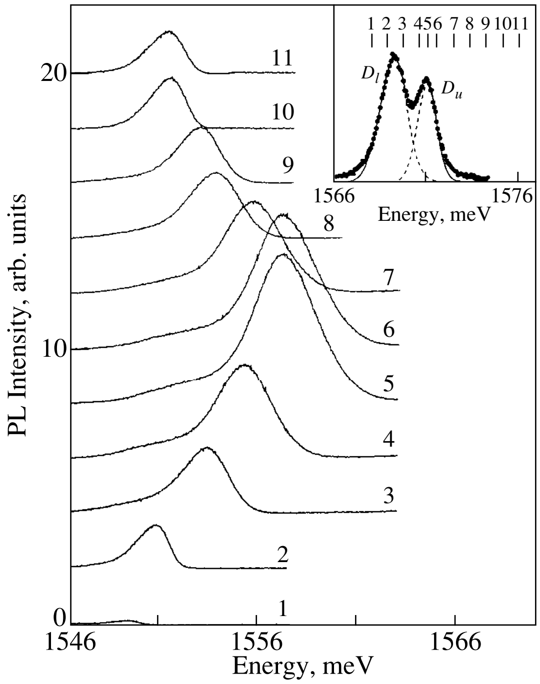


Fig. 2. Excitation energy dependence of the indirect PL at T_{bath} = 1.4 K, V_g = 1 V, and W_{ex} = 20 W/cm^2. The corresponding excitation energies are shown in the inset. Inset: the direct PL at T_{bath} = 1.4 K, V_g = 1 V, W_{ex} = 0.5 W/cm^2, and E_{ex} = 1.96 eV (points); for separation of the lines, the direct PL is fitted by two Gaussians (dashed lines, the sum is the solid line).

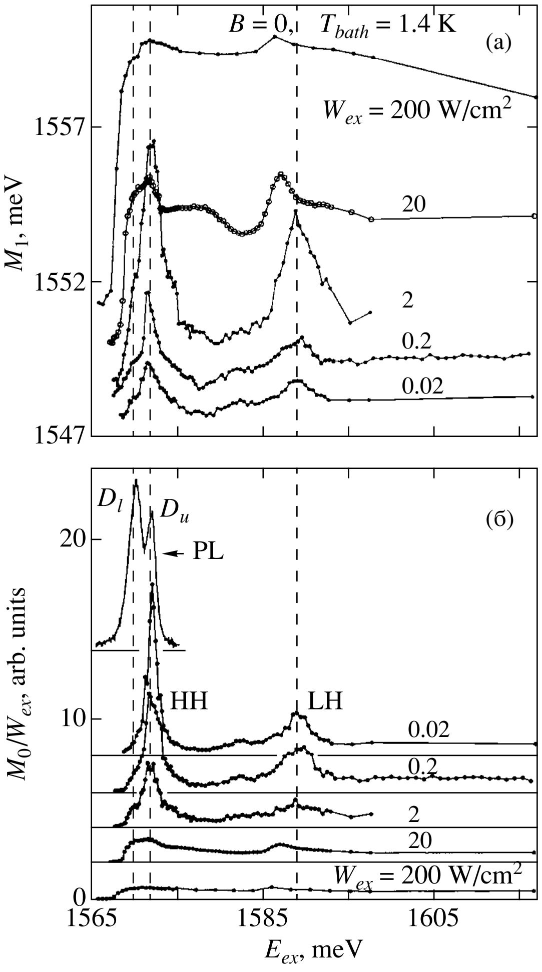


Fig. 3. Excitation energy dependence of (a) the indirect PL line position M_1 and (b) the integrated indirect PL intensity M_0 normalized by the excitation density vs. excitation density at T_{bath} = 1.4 K and V_g = 1 V. The direct PL at T_{bath} = 1.4 K, V_g = 1 V, W_{ex} = 0.5 W/cm^2, and E_{ex} = 1.96 eV is shown above (b). The dashed lines are a guide for the eyes.

relation terms result in the reduction of the energy (an effect known as the band-gap renormalization [1, 2]). Therefore, in CQWs studied here, the electrostatic term dominates over the exchange and correlation terms. In particular, the observed enhancement of the ground-state energy of the system of the spatially separated electron and hole layers with increasing the *e–h* density indicates that the real space condensation to droplets is energetically unfavorable and corresponds to the theoretical predictions of [7, 9]. On the contrary, the condensation to the exciton liquid predicted in [10] is not supported by the present experiment (we note that for the CQW studied, we have $d \approx 1.77a_2$, which must correspond to the metastable exciton liquid phase according to [10]). Indeed, if the exciton liquid were the ground state, the *e–h* density and hence, the PL shape and energy should be fixed and independent of the excitation density; this does not correspond to the experimental data (Fig. 1).

The lowest estimate of the *e–h* density can be obtained from the experimental shift of the indirect PL line to higher energies using the plate capacitor formula. This estimate does not include exchange and correlation terms, and the resulting value of the density is therefore lower than the actual one. In particular, the estimate for the maximum possible *e–h* density corresponding to the complete screening of the external electric field is ~2×10^{11} cm^{-2} for V_g = 1.8 V.

Two direct PL transitions, the upper (D_u) and the lower (D_l), are observed in the indirect regime (Fig. 1). Their excitation density dependence is opposite to that of the indirect PL line. The D_u line position is practically independent of W_{ex} and the D_l line shifts to lower energies (Fig. 1). The relative intensity of the D_l line is increased with W_{ex} (Fig. 1).

Figure 2 presents the excitation energy dependence of the indirect PL for the excitation energies in the range of direct PL (see the inset to Fig. 2). The integrated indirect PL intensity $M_0 = \int I(E)dE$ and the PL line position given by the line gravity center $M_1 = M_0^{-1}\int EI(E)dE$ are presented in Fig. 3 as a function of excitation energies. The set of $M_0(E_{ex})$ and $M_1(E_{ex})$ dependences is a more comprehensive analog of the indirect PL excitation (PLE) spectra: in addition to the latter, it contains information on a variation of the PL line position. In the case where the indirect PL dominates the *e–h* recombination, the $M_0(E_{ex})$ dependence coincides with the absorption spectrum.

Figures 2 and 3b show that the indirect PL line intensity is maximum at the excitation energies corresponding to the D_u line. At the same time, no absorption is observed at the energy of the D_l line at low excitation densities; the absorption at the D_l line energy appears and its relative intensity increases with increasing W_{ex} (Fig. 3b).

The nature of the D_u and D_l lines is discussed in what follows. The carrier distribution scheme in CQW in the indirect regime is shown in Fig. 4 (left). The ratio of the densities of the minor carriers (electrons in the left QW and holes in the right QW) to the densities of the dominant carriers (holes in the left QW and electrons in the right QW) is proportional to the ratio between the direct and indirect PL line intensities multiplied by the ratio between the direct and indirect radiative decay times and is small.[1] The scheme of possible direct and indirect PL transitions is evident from Fig. 4 (left) and is shown in Fig. 4 (right). We attribute the D_l line to direct multiparticle complexes because its relative intensity increases with increasing the electron–hole density in both PL and absorption (Figs. 1 and 3), which indicates that more than two particles (electrons and holes) are involved in the complex. The simplest charged complexes are X^- and X^+.

The D_u line corresponds to the direct heavy hole (HH) 1s exciton, X. In CQWs, the formation of charged complexes is promoted in the indirect regime because of the interwell charge separation and the corresponding realization of the charge configuration, where the electron (hole) is surrounded by a dominant number of holes (electrons) in the left (right) QW, see the scheme of Fig. 4. Indeed, the D_l line vanishes for the symmetric charge distribution at $V_g = 0$. We note that two direct lines could alternatively be ascribed to recombination from two QWs of slightly different widths (see the inset to Fig. 1). This alternative interpretation is discarded because the lower direct line is absent in the absorption at low excitation densities (Fig. 3).

With increasing the *e–h* density, the D_l line shifts to lower energies. This behavior of the intrawell optical transition corresponds to the band-gap renormalization in SQWs [1, 2] and is qualitatively discussed in what follows (for simplicity, we discuss the intrawell PL transitions in the right QW with excess electrons, the transitions in the left QW are characterized by a similar density dependence). At high *e–h* densities, more than one excess electron is in the vicinity of the photoexcited or recombining exciton. In the dense limit, the direct PL is described by the correlation effects in the 2D electron gas and the hole in the right QW and the 2D hole gas in the left QW. The PL energy of the 2D electron gas is reduced with increasing the electron density because of the band-gap renormalization [1]. The presence of the separated 2D hole gas in the left QW must further increase the PL energy reduction due to the exchange interaction with the hole in the right QW. In CQWs therefore, similarly to the SQW case, the intrawell PL energy must be reduced with increasing the density due to the band-gap renormalization. This corresponds to the experimental data (Fig. 1). We note that the X^- complex is the low-density limiting case of the correlations of 2D electron gas with a hole.

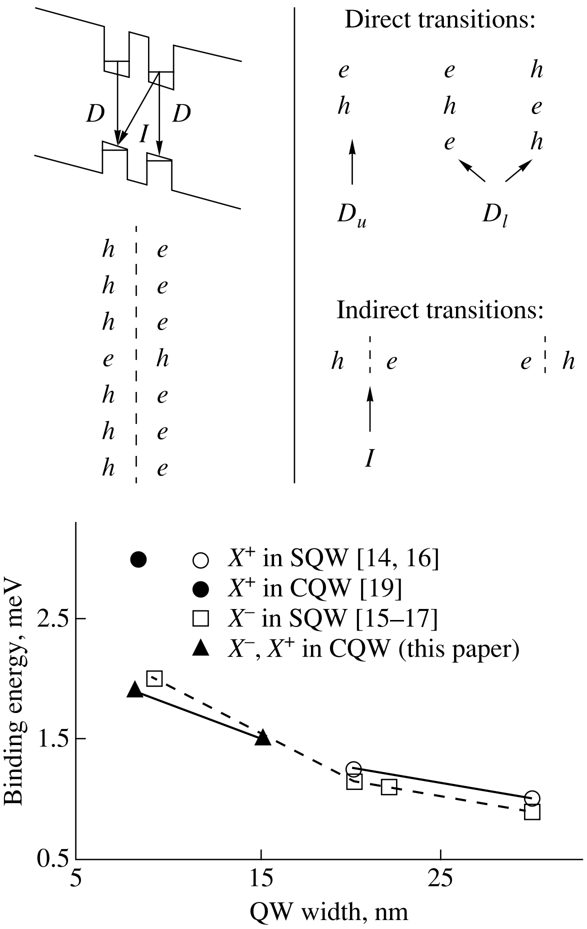


Fig. 4. Scheme of the carrier distribution in CQW in the indirect regime (left). Scheme of possible direct (intrawell) and indirect (interwell) PL transitions (right). The transition assignment (D_u, D_l, I) corresponds to that of the PL lines in Figs. 1–3. QW width dependence of X^- and X^+ binding energies in SQWs and CQWs (the lower plot).

The correct determination of the X^- and X^+ complex binding energies must be done at the lowest *e–h* densities to avoid the effect of extra (more than one) excess carriers occurring in the vicinity of the photoexcited or recombining exciton (see above). The binding energy

---

[1] The indirect radiative decay times are in the range of tens and hundreds of ns, while the direct radiative decay time is below our system resolution, 0.2 ns [11].

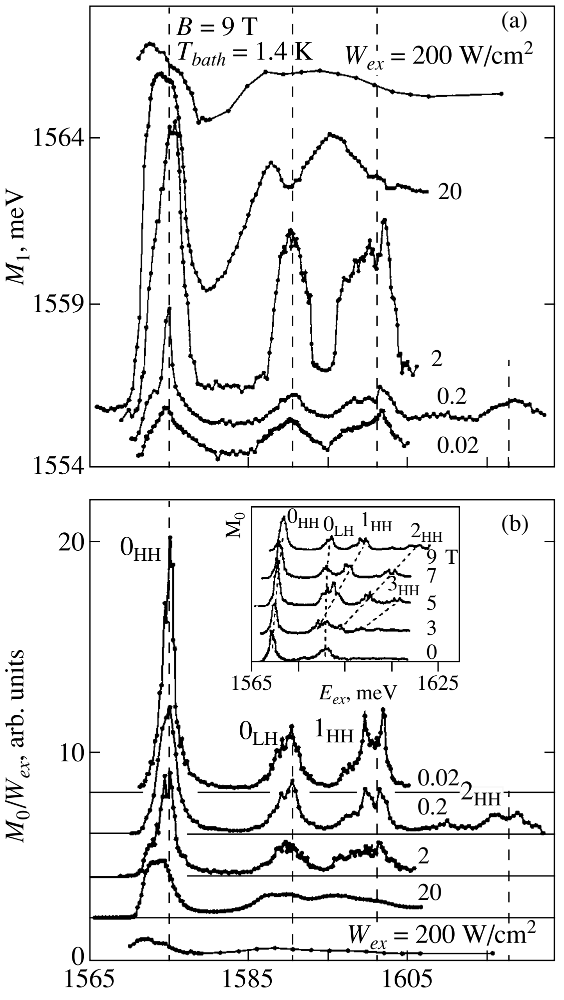


Fig. 5. Excitation energy dependence of (a) the indirect PL line position M_1 and (b) the integrated indirect PL intensity M_0 normalized by the excitation density vs. excitation density at $B = 9$ T, $T_{bath} = 1.4$ K, and $V_g = 1$ V. Inset: excitation energy dependence of the integrated indirect PL intensity M_0 vs. magnetic field at $T_{bath} = 1.4$ K, $V_g = 1$ V, and $W_{ex} = 0.2$ W/cm^2. The dashed lines are a guide for the eyes.

of the complexes was determined from the splitting between the D_u and D_l lines at low excitation density $W_{ex} = 0.6$ W/cm^2 fitting the direct PL by two Gaussians (see the inset to Fig. 2). The evaluation of the data at the lower excitation densities results in the close values but is less accurate due to noise. The obtained binding energy of the complexes is 1.9 meV for the 8–4–8-nm CQW and 1.5 meV for the 15–4–15-nm CQW. These values are consistent with the earlier reported X^- and X^+ binding energies in the 2D electron (hole) gas in modulation doped GaAs SQWs: 1 meV for X^+ in 30-nm QW [13]; 1.1 and 0.9 meV for X^- in 22- and 30-nm QWs, respectively [14]; 1.15 meV for X^- and 1.25 meV for X^+ in 20-nm QW [15]; and 2 meV for X^- in 9.1-nm QW [16]. The X^- and X^+ binding energies increase with reducing the QW thickness because of the enhanced Coulomb correlations, which is consistent with the theoretically predicted increase of the complex stability with reduced dimensionality [17]. The binding energies of X^- and X^+ for the same QW width are close to each other [13–15, 17]. The measured complex binding energy values of 1.9 and 13 meV for the respective 8- and 15-nm QWs fit well to the X^- and X^+ binding energy dependence on the QW width [13–17] (see the lower plot of Fig. 4). Recently, the binding energy of X^+ in the CQW structure with 8-nm QWs was reported to be 3 meV [18]: this value is also presented in Fig. 4. The difference in the binding energies observed in the CQWs with similar well and barrier widths (compare [18] and the present paper) is likely to be related to the larger amplitude of the in-plane random potential in the sample studied in [18], which is revealed in the larger PL linewidth: similarly to the case of SQWs [13–16], trions are most likely localized in CQWs by in-plane potential fluctuations; the larger amplitude of potential fluctuations results in a reduction of the carrier localization area. This additional carrier confinement must enhance the Coulomb correlations and, therefore, the complex stability.

The absorption line at ≈17 meV above the direct 1*s* HH exciton corresponds to the direct light hole (LH) exciton (see Section 3) while the shoulder at ≈10 meV above the direct 1*s* HH exciton corresponds to the onset of the direct excited HH exciton states 2*s*, 3*s*, …, and the HH free carrier absorption edge. This indicates a direct HH exciton binding energy about 10 meV.

Figures 2 and 3 show that the energy of the indirect PL line is locked to its intensity: $M_1(E_{ex})$ varies in phase with $M_0(E_{ex})$. This is consistent with the excitation density dependence of the indirect PL line (Fig. 1 and the excitation density dependence in Fig. 3a) and corresponds to the enhancement of the ground-state energy of spatially separated electrons and holes with increasing the *e*–*h* density (see above). The density depends on the excitation energy due to the absorption variation.

3. SPATIALLY SEPARATED ELECTRON AND HOLE LAYERS AT HIGH MAGNETIC FIELDS

Coulomb correlation effects in the PL spectra of spatially separated electron and hole layers at a perpendicular magnetic field are considered in this section. Figure 5 presents the excitation energy dependence of the indirect PL line position and the integrated indirect PL intensity vs excitation density at $B = 9$ T. These dependences are analogous to those at $B = 0$ presented in Fig. 3. At $B = 9$ T, the direct magnetoexciton lines dominate the absorption. They are identified by their magnetic field dependence presented in the inset to Fig. 5. The energy of the direct magnetoexciton is $\mathscr{E}_D(N, B) = E_g + (N + 1/2)\hbar\omega_c - E_D(N, B)$, where E_g is the energy gap including the electron and hole confinement energies in the CQW, $\hbar\omega_c$ is the sum of the electron and hole cyclotron energies, N is the Landau level number, and E_D is the direct exciton binding energy; at high magnetic fields, $E_D \sim 1/l_B$ and reduces with N, where $l_B = \sqrt{\hbar c/eB}$ is the magnetic length [19]. At high magnetic fields, similarly to the zero-field case, the energy of the indirect PL increases with increasing the *e–h* density in the entire density range (compare Figs. 3a and 5a). This is observed both with increasing excitation density and with increasing absorption (see the W_{ex} and E_{ex} dependences of Fig. 5a).

At high excitations, a neutral dense magnetoplasma of spatially separated electrons and holes, indirect magnetoplasma, is realized: several optical transitions between occupied Landau levels of spatially separated electrons and holes (LL indirect transitions) are observed in the PL spectra and identified by their magnetic field dependences (the left inset to Fig. 6). Hot direct PL is also observed (the left inset to Fig. 6). Higher Landau level transitions appear in PL spectra with increasing excitation density because of the consequent occupation of the higher Landau levels (the right inset to Fig. 6).

The energies of all the indirect LL transitions in indirect magnetoplasma monotonically increase with increasing the *e–h* density (Fig. 6).[2] This is opposite to the density dependence of direct LL transitions in SQWs where the uppermost occupied (Nth) LL transition energy is independent of the *e–h* density in the range of filling factors $N < \nu/2 < N + 1$ and is reduced with increasing the *e–h* density everywhere outside this range because of the band gap renormalization [20]. The density dependence of direct LL transitions in SQWs is quantitatively well explained: for the uppermost occupied LLs, electrons and holes bind into magnetoexcitons[3] that are noninteracting in the high-magnetic-field limit because of the compensation between repulsion at small distances caused by the Pauli exclusion principle and attraction at large distances caused by the exchange interaction [21, 22]; for the filled *e* and *h* LLs ($\nu/2 > N + 1$) or empty LLs ($\nu/2 < N$), the transition energy is reduced because the exchange interaction is not compensated [20, 23].

The energy independence of the direct magnetoexciton energy from density in SQWs originates from the symmetry of the *e–e* and *e–h* interaction [21, 22]. In CQWs, the interaction is strongly asymmetric. The asymmetry is the basis of two opposite theoretical models. According to the first model, the uniform magnetoexciton phase is unstable at low temperatures [5, 6, 10] and, in particular, the condensation to the exciton liquid is expected [10]; according to the second model, the repulsive interaction between the indirect magnetoexcitons stabilizes the uniform gas phase and the indirect magnetoexciton energy monotonically increases with density [7]. The experimentally observed enhancement of the magnetoexciton energy with density supports the second model, and, in particular, indicates the dominance of the electrostatic term over the exchange term for the spatially separated electron and hole layers (as in the zero-field case, see Section 2).

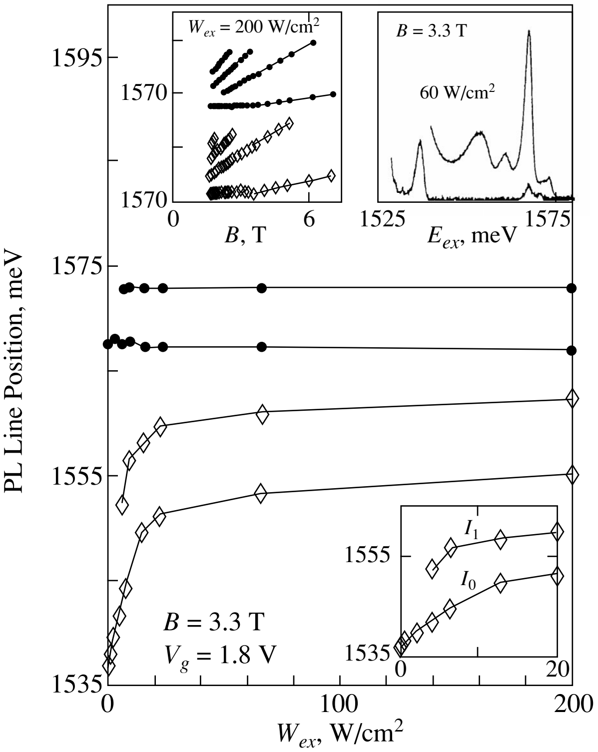


Fig. 6. Excitation density dependence of the main direct (D) and indirect (I) PL line positions at T_{bath} = 1.4 K, B = 3.3 T, V_g = 1.8 V, and E_{ex} = 1615 meV. The low W_{ex} region is expanded in the lower inset. Left inset: Magnetic field dependence of the main D and I PL line positions at T_{bath} = 1.4 K, V_g = 1.8 V, and W_{ex} = 20 W/cm^2. Upper right inset: PL spectra at W_{ex} = 0.2 and 60 W/cm^2, T_{bath} = 1.4 K, B = 3.3 T, and V_g = 1.8 V.

4. CONCLUSIONS

Many-body effects in spatially separated electron and hole layers in coupled quantum wells were found to result in an enhancement of the indirect PL energy with increasing the *e–h* density both for a zero magnetic field and at high fields for all Landau level transitions in the entire range of *e–h* densities. This behavior is

---

[2] The rate of the indirect PL energy enhancement reduces with increasing the excitation density (Fig. 6). The origin of this is likely to be related to a sublinear increase of n_{eh} with W_{ex}. At a higher excitation density, the internal electric field between the electron and hole layers is smaller because of the carrier screening of the external electric field (see above); at the smaller electric field, the recombination time of interwell PL becomes shorter [11], which results in a smaller density of photoexcited carriers for the fixed excitation density (in addition, at high $W_{ex} \gtrsim 20$ W/cm^2, the *e–h* recombination time is reduced also due to the enhancement of the direct recombination.

[3] Note that the screening of the magnetoexciton in a dense low-temperature magnetoplasma is suppressed compared to the zero field case: in particular, in the high-magnetic-field limit, the carriers at the completely filled LLs do not participate in the screening of magnetoexcitons at the uppermost occupied LL [20–23].

opposite to the case of electron–hole systems in single QWs, where the main features are explained by the band gap renormalization resulting in a reduction of the PL energy. The density dependence of the indirect PL energy is explained by the dominance of the electrostatic term originating from the electric field between the separated electron and hole layers; this reduces the net local electric field and results in an increase of the energy. The observed enhancement of the ground-state energy of the system of the spatially separated electron and hole layers with increasing the *e–h* density indicates that the real space condensation to droplets is energetically unfavorable.

At high densities of separated electrons and holes, a new direct (intrawell) PL line has been observed; its relative intensity increased both in PL and in absorption (measured by indirect PL excitation) with increasing the *e–h* density. It is therefore attributed to direct multiparticle complexes. The measured complex binding energy values 1.9 and 1.5 meV for the studied 8 and 15 nm QWs, respectively, fit well to the previously reported X^- and X^+ binding energy dependence on the QW width.

We acknowledge support from NSF Center for Quantized Electronic Structures (QUEST), INTAS, the Russian Foundation for Basic Research. and the Program “Physics of Solid State Nanostructures” from the Russian Ministry of Sciences. A.I. acknowledges the support of a David and Lucile Packard Fellowship.

REFERENCES

1. S. Schmitt-Rink, D. S. Chemla, and D. A. B. Miller, Adv. Phys. **38**, 89 (1989).
2. V. D. Kulakovskiĭ, E. Lach, A. Forchel, and D. Grützmacher, Phys. Rev. B **40**, 8087 (1989); L. V. Butov, V. D. Kulakovskiĭ, E. Lach, *et al.*, Phys. Rev. B **44**, 10680 (1991); L. V. Butov, V. D. Egorov, V. D. Kulakovskiĭ, and T. G. Andersson, Phys. Rev. B **46**, 15156 (1992); L. V. Butov, V. D. Kulakovskiĭ, and A. Forchel, Phys. Rev. B **48**, 17933 (1993).
3. Yu. E. Lozovik and V. I. Yudson, Zh. Éksp. Teor. Fiz. **71**, 738 (1976) [Sov. Phys. JETP **44**, 389 (1976)].
4. Y. Kuramoto and C. Horie, Solid State Commun. **25**, 713 (1978).
5. I. V. Lerner, Yu. E. Lozovik, and D. R. Musin, J. Phys. C **14**, L311 (1981).
6. Yu. A. Bychkov and E. I. Rashba, Solid State Commun. **48**, 399 (1983); Zh. Éksp. Teor. Fiz. **85**, 1826 (1983) [Sov. Phys. JETP **58**, 1062 (1983)].
7. D. Yoshioka and A. H. MacDonald, J. Phys. Soc. Jpn. **59**, 4211 (1990).
8. X. M. Chen and J. J. Quinn, Phys. Rev. Lett. **67**, 895 (1991).
9. X. Zhu, P. B. Littlewood, M. S. Hybertsen, and T. M. Rice, Phys. Rev. Lett. **74**, 1633 (1995).
10. Yu. E. Lozovik and O. L. Berman, Pis’ma Zh. Éksp. Teor. Fiz. **64**, 526 (1996) [JETP Lett. **64**, 573 (1996)]; Zh. Éksp. Teor. Fiz. **111**, 1879 (1997) [JETP **84**, 1027 (1997)].
11. L. V. Butov, A. Imamoglu, A. V. Mintsev, *et al.*, Phys. Rev. B **59**, 1625 (1999).
12. L. V. Butov, A. A. Shashkin, V. T. Dolgopolov, *et al.*, Phys. Rev. B **60**, 8753 (1999).
13. A. J. Shields, J. L. Osborne, M. Y. Simmons, *et al.*, Phys. Rev. B **52**, R5523 (1995).
14. A. J. Shields, M. Pepper, D. A. Ritchie, *et al.*, Phys. Rev. B **51**, 18049 (1995).
15. G. Finkelstein, H. Shtrikman, and I. Bar-Joseph, Phys. Rev. Lett. **74**, 976 (1995); Phys. Rev. B **53**, R1709 (1996).
16. H. Buhmann, L. Mansouri, J. Wang, *et al.*, Phys. Rev. B **51**, 7969 (1995).
17. B. Stébé and A. Ainane, Superlattices Microstruct. **5**, 545 (1989).
18. V. B. Timofeev, A. V. Larionov, M. Grassi Alessi, *et al.*, Phys. Rev. B **60**, 8897 (1999).
19. I. V. Lerner and Yu. E. Lozovik, Zh. Éksp. Teor. Fiz. **78**, 1167 (1980) [Sov. Phys. JETP **51**, 588 (1980)].
20. L. V. Butov, V. D. Kulakovskii, G. E. W. Bauer, *et al.*, Phys. Rev. B **46**, 12765 (1992).
21. I. V. Lerne and Yu. E. Lozovik, Zh. Éksp. Teor. Fiz. **80**, 1488 (1981) [Sov. Phys. JETP **53**, 763 (1981)].
22. D. Paquet, T. M. Rice, and K. Ueda, Phys. Rev. B **32**, 5208 (1985).
23. Yu. A. Bychkov and E. I. Rashba, Pis’ma Zh. Éksp. Teor. Fiz. **52**, 1209 (1990) [JETP Lett. **52**, 624 (1990)]; L. V. Butov, V. D. Kulakovskiĭ, and E. I. Rashba, Pis’ma Zh. Éksp. Teor. Fiz. **53**, 104 (1991) [JETP Lett. **53**, 109 (1991)].

Journal of Experimental and Theoretical Physics, Vol. 92, No. 2, 2001, pp. 267–276.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 119, No. 2, 2001, pp. 309–320.
Original Russian Text Copyright © 2001 by Dobryakov, Kovalenko, Lozovik, Merkulova, Farztdinov, Ernsting.

SOLIDS
Electronic Properties

Femtosecond Spectroscopy of Relaxation Processes in Metals and High-T_c Superconductors

A. L. Dobryakov[a], **S. A. Kovalenko**[b], **Yu. E. Lozovik**[a, *], **S. P. Merkulova**[a],
V. M. Farztdinov[b], **and N. P. Ernsting**[b]

[a]*Institute of Spectroscopy, Russian Academy of Sciences, Troitsk, Moscow oblast, 142092 Russia*
[b]*Humboldt Universität zu Berlin, Berlin, Deutschland*
**e-mail: lozovik@isan.troitsk.ru*

Received July 7, 2000

Abstract—Spectral dependences of charge carrier relaxation rates, $\gamma_{e-e}(\hbar\omega)$ and $\gamma_{e-ph}(\hbar\omega)$, were observed in Au and Cu films and $YBa_2Cu_3O_{7-\delta}$ high-T_c superconductor films. The relaxation rates decreased substantially in the spectral region corresponding to interband transitions to the Fermi level region ($\hbar\omega_{Au} = 2.45$ eV, $\hbar\omega_{Cu} = 2.15$ eV, and $\hbar\omega_1 = 1.89$ eV and $\hbar\omega_2 = 2.08$ eV for $YBa_2Cu_3O_{7-\delta}$). This relaxation deceleration opens up possibilities for developing a new method, based on the spectral dependences of relaxation rates, for the determination of the Fermi level position and the parameters of electron–electron and electron–phonon interactions on the one hand and for studying deviations from the Fermi-liquid behavior in strongly correlated electronic systems. The linear $\gamma_{e-e}(\hbar\omega) \propto |\hbar\omega - E_F|$ spectral dependence was observed for a $YBa_2Cu_3O_{7-\delta}$ film near $\hbar\omega_2 = 2.08$ eV, which may be evidence of a non-Fermi-liquid behavior of the electronic subsystem. © *2001 MAIK “Nauka/Interperiodica”.*

1. INTRODUCTION

The interaction of ultrashort laser pulses with solids is characterized by the disturbance of equilibrium between the electronic and ionic subsystems of the solid on the femtosecond time scale [1–4]. The temperature of radiation-excited electrons may exceed the temperature of the sublattice by several orders of magnitude [1], because the pulse width ($\tau_p \sim 10^{-14}$–10^{-13} s) and the time of electronic subsystem thermalization ($\tau_{e-e} = \gamma_{e-e}^{-1} \sim 10^{-15}$–$10^{-14}$ s) are much smaller than the characteristic time of energy exchange between electrons and the lattice ($\tau_{e-ph} = \gamma_{e-ph}^{-1} \sim 10^{-12}$ s). Equilibrium in the initially disturbed electron–lattice system is reestablished with time as a result of electron–phonon interaction. Data on the dynamics of attaining equilibrium in the femtosecond time range yield information about the Eliashberg parameter $\lambda\langle\Omega^2\rangle$, where λ is the electron–phonon interactions constant and Ω are phonon frequencies, for electron–phonon interaction in the material under study [2]. The value of this parameter determined experimentally can serve as evidence in favor of one or another mechanism of high-temperature superconductivity.

This method was used in several experimental studies of optical excitation relaxation in metals [3–9] and high-T_c superconductors [10–12]. Compared with traditional probing usually performed at the wavelength of excitation [3, 4, 10], probing with a femtosecond time resolution in a wide spectral range [13, 14] yields important additional information about the response of a system excited by ultrashort laser pulses. For instance, a detailed study of the dependences of relaxation rates on the energy of photoexcited charge carriers has made it possible to determine the spectral dependences of relaxation times and, through this, electron–electron and electron–phonon interaction parameters [8, 9, 15]. The existence of a substantial spectral dependence of electron–electron and electron–phonon relaxation rates has led the authors of [16, 17] to suggest a method for determining the position of the Fermi level based on a sharp increase in the relaxation time near this level caused by a decrease in the accessible phase volume. In addition, identifying certain special features in the spectral dependences of excited carrier relaxation times allows us to determine deviations from the Fermi-liquid behavior of electronic subsystems [17].

The purpose of this work was to study and compare the special features of the superfast dynamics of nonequilibrium charge carriers in Au and Cu and $YBa_2Cu_3O_{7-\delta}$ high-T_c superconductors in the normal phase (in the femtosecond time region) by the “excitation–probing” method with probing in a wide spectral range [13, 14]. Our goal was also to obtain detailed data on the spectral dependences of electron–electron and electron–phonon relaxation rates. The material is arranged as follows. Section 2 describes the basic concepts of the model for studying the relaxation of excited charge carriers, its shortcomings, and its possible generalization.

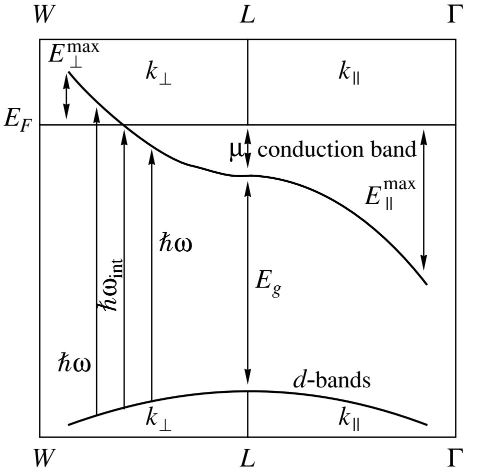


Fig. 1. Band structure of a Au film near the *L* point of the Brillouin zone in the parabolic approximation [19].

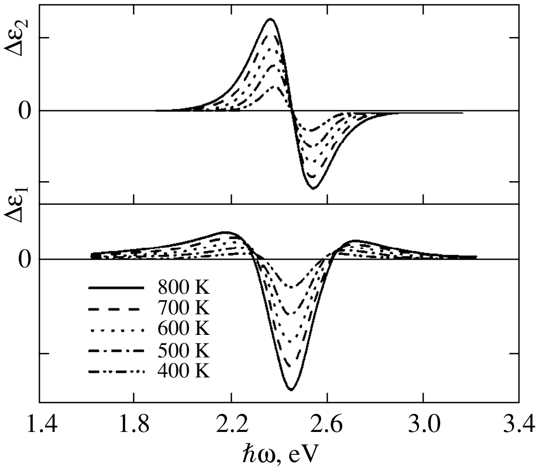


Fig. 2. Variations in difference permittivity components $\Delta\epsilon_1$ and $\Delta\epsilon_2$ for a Au film according to the thermomodulation model [18, 19] at various temperatures.

In Section 3, we analyze the spectral dependences of both electron–electron and electron–phonon relaxations characterized by substantial deceleration near the Fermi level. This allows the position of the Fermi level to be determined from the spectral dependences of relaxation rates obtained in femtosecond experiments. Section 4 contains the experimental femtosecond spectroscopy data on the Au and Cu. Similar data on the $YBa_2Cu_3O_{7-\delta}$ high-T_c superconductor are reported in Section 5. In Section 6, we consider the excitation of coherent phonons in $YBa_2Cu_3O_{7-\delta}$ films. Lastly, Section 7 summarizes the results and discusses possible applications of the observed effects to femtosecond studies of metals and high-T_c superconductors.

2. THERMOMODULATION MODEL

The traditional approach to studying the dynamics of nonequilibrium charge carriers in metals and high-T_c superconductors by femtosecond spectroscopy techniques uses the thermomodulation model [1, 18]. This model has been extensively employed to describe the experimental femtosecond spectroscopy data on both metals [3–9] and high-T_c superconductors [9–12, 16]. Usually, the following assumptions are made within the framework of the thermomodulation model.

(1) A narrow laser radiation pulse of width τ_p is absorbed by free electrons in the conduction band, which creates a nonequilibrium electronic subsystem distribution.

(2) Electron–electron interactions control the attainment of a quasi-equilibrium electronic subsystem distribution at a characteristic relaxation rate $\gamma_{e-e} = \tau_{e-e}^{-1}$. It is assumed that the exciting pulse width is related to τ_{e-e} as $\tau_p \gg \tau_{e-e}$; that is, at time $t \sim \tau_p$, the electron distribution function is independent of the form of the perturbing nonequilibrium distribution function and is characterized by the effective temperature of electrons, $T_e(t)$, which exceeds lattice temperature T_L. At longer times $t > \tau_p$, electron–phonon interactions equalize the electronic subsystem and lattice temperatures at a characteristic relaxation rate $\gamma_{e-ph} = \tau_{e-ph}^{-1}$.

In experiments of the type under consideration, a probing laser pulse is used to study absorption caused by transitions from the deep-lying filled *d* band to the Fermi level E_F region in the conduction band (Fig. 1). An excitation pulse increases the effective temperature of electrons, and the population of electronic states decreases below and increases above the Fermi level. The difference in responses of excited and nonexcited samples, which are measured in experiments and which are proportional to the imaginary part of permittivity ($\Delta\epsilon_2$), change their sign at the frequency corresponding to the transitions to the Fermi level region. This results in the formation of different absorption spectra with alternating signs, which pass through zero (Fig. 2) at the $\hbar\omega_{\text{int}}$ point [19] corresponding to interband transitions to the E_F region,

$$\Delta\epsilon_2(\omega, T_e(t), T_0) = \frac{2A}{\omega^2(1-g_\perp)} \times \int_0^{\sqrt{\hbar\omega_{\max}}} dx \left\{ \tanh\frac{\hbar\omega - \hbar\omega_{\text{int}} - x^2(g_\perp + g_\parallel)}{2k_B T_e(t)(1+g_\perp)} - \tanh\frac{\hbar\omega - \hbar\omega_{\text{int}} - x^2(g_\perp + g_\parallel)}{2k_B T_0(1+g_\perp)} \right\}. \quad (1)$$

Here, the notation is as follows (see [19] and Fig. 1):

$$\hbar\omega_{\rm int} = E_g + \mu(1+g_\perp), \quad g_\perp = \left|\frac{m_{c\perp}}{m_{d\perp}}\right|, \quad g_\parallel = \left|\frac{m_{c\parallel}}{m_{d\parallel}}\right|,$$

$$\sqrt{\hbar\omega_{\max}} = \frac{\beta}{\alpha} - \sqrt{\frac{1}{\alpha}\left(\hbar\omega + \frac{\beta^2}{\alpha} - \delta - E_g\right)},$$

$$\alpha = (1+g_\perp)\frac{E_\perp^{\max}}{E_\parallel^{\max}} - 1 + g_\parallel,$$

$$\beta = (1+g_\perp)\frac{E_\perp^{\max}}{\sqrt{E_\parallel^{\max}}}, \quad \delta = (1+g_\perp)E_\perp^{\max},$$

$$A = \frac{e^2|\lambda|^2|m_{c\perp}|\sqrt{2|m_{c\parallel}|}}{m^2\Omega^2},$$

and $k_{\rm B}$ is the Boltzmann constant.

Note that the determination of the position of the Fermi level by traditional femtosecond spectroscopy techniques requires performing independent reflectance (ΔD_R) and transmission (ΔD_T) measurements. Changes in the imaginary permittivity part can then be unambiguously reproduced by

$$\Delta D_T = \frac{\partial D_T}{\partial \epsilon_1}\Delta\epsilon_1 + \frac{\partial D_T}{\partial \epsilon_2}\Delta\epsilon_2,$$
$$\Delta D_R = \frac{\partial D_R}{\partial \epsilon_1}\Delta\epsilon_1 + \frac{\partial D_R}{\partial \epsilon_2}\Delta\epsilon_2, \tag{2}$$

which allows the position of the Fermi level to be determined [18, 19] ($\Delta\epsilon_{1,\,2}$ is the photoinduced change in permittivity). We show below that the position of the Fermi level can even more accurately be found from the deceleration of relaxation near the Fermi level measured in a single “excitation–wide-band probing” experiment performed in either the transmission or the reflectance mode.

In terms of the two-temperature model ignoring the spectral dependence of relaxation [2], the γ_T rate of cooling of the electronic subsystem under femtosecond experiment conditions can be related to the $\lambda\langle\Omega^2\rangle$ Eliashberg parameter for electron–phonon interactions as

$$\frac{\partial T_e(t)}{\partial t} = \gamma_T(T_L - T_e), \quad \gamma_T = \frac{3\hbar\lambda\langle\Omega^2\rangle}{\pi k_B T_e}. \tag{3}$$

The $\lambda\langle\Omega^2\rangle$ electron–phonon interaction parameter plays a very important role in the theory of superconductivity, and its determination can provide insight into the nature of interactions responsible for electron pairing. Studying the kinetics of photoinduced reflection (or transmission) by femtosecond spectroscopy techniques allows the evolution of the temperatures of electrons and the lattice to be reproduced and this parameter to be experimentally determined. This approach has been extensively used to measure $\lambda\langle\Omega^2\rangle$ for both metals (Cu, Au, and Nb [2, 10]) and high-T_c superconductors ($YBa_2Cu_3O_{7-\delta}$ and $BiSr_2CaCu_2O_{8+x}$ [2, 10–12]).

The use of the thermomodulation model has allowed the principal physical processes of the dynamics of charge carriers in metals and high-T_c superconductors to be studied and the evolution of the effective temperatures of the electronic and phonon subsystems to be reproduced. When picosecond and subpicosecond laser pulses are used, the application of the thermomodulation model is doubtless warranted. Indeed, the time of electron–electron relaxation in metals can be estimated as $\tau_{e\text{–}e} \sim 1/\omega_{pl}$, where ω_{pl} is the plasma frequency. For typical metals with $\omega_{pl} \sim 10$ eV, the time of attaining quasi-equilibrium in the electronic subsystem is of the order of $\tau_{e\text{–}e} \sim 10$ fs. As the laser excitation pulse width is $\tau_p \sim 100$ fs, there is sufficient time for electronic temperature T_e to be established, because $\tau_p \gg \tau_{e\text{–}e}$. The use of narrower excitation pulses ($\tau_p \sim \tau_{e\text{–}e}$), however, raises the question of whether or not the $\tau_{e\text{–}e}$ time of establishing a quasi-equilibrium distribution can safely be excluded from consideration [1].

Indeed, the existence of a nonequilibrium distribution of the electronic subsystem at times up to 600 fs was observed in femtosecond time resolution experiments on the photoemission of electrons from Au films [5]. In [5], the suggestion was made that electron–electron interaction processes determining the attainment of a quasi-equilibrium distribution of electrons played an important role when femtosecond laser pulses were used, and the principal concepts of the thermomodulation model should therefore be revised to take into account the process of establishing electronic temperature. This suggestion was substantiated in femtosecond experiments on the photoinduced absorption and reflection of Au films [6] and in studies of Ag and Au films by surface plasmon–polariton resonance techniques [7]. These experiments showed that describing the femtosecond responses of a metal required generalizing the simple thermomodulation model, which treated metals in terms of the two-temperature model comprising the electronic subsystem with temperature $T_e(t)$ and the lattice with temperature $T_L(t)$. The use of ultrashort pulses required taking into account electron–electron interaction during the “thermalization of the electronic subsystem;” that is, during the establishment of the effective temperature of the electronic subsystem [1].

There is another fundamental reason for revising the thermomodulation model. This model does not take into account the spectral dependence of electron–phonon interactions. At the same time, theory [15, 20, 21] predicts the spectral dependence of the rates of both electron–electron and electron–phonon relaxations to be substantial. Both relaxation rates noticeably decrease near the Fermi level because of a decrease in the phase volume accessible to relaxation. A decrease in the rate of relaxation near the Fermi level was indeed observed in femtosecond experiments conducted as excitation–wide-band probing [8, 9, 16] and by the two-photon photoemission method [22, 23].

For this reason, the use of ultrashort pulses of widths comparable with the characteristic electron–electron relaxation time requires generalizing the thermomodulation model to take into account establishing equilibrium in the electronic subsystem, and wide-band probing makes it possible to thoroughly study the spectral dependence of relaxation rates near the Fermi level and to directly determine the electron–phonon interaction parameter [15] from the spectral dependence of the rate of electron–phonon relaxation.

3. THE SPECTRAL DEPENDENCES OF ELECTRON–ELECTRON AND ELECTRON–PHONON RELAXATIONS

The most important mechanisms of the scattering of charge carriers in metals, semimetals, and semiconductors (such as electron–electron and electron–phonon interactions), the modern advances in the kinetics of charge carriers, and the principal experiments in which the scattering manifests itself most obviously are described in detail in the monograph by Gantmakher and Levinson [20]. According to the $\gamma_{e-e}(\hbar\omega)$ spectral dependence of the rate of electron–electron relaxation in the Fermi liquid [20, 21], this rate should decrease near the Fermi level,

$$\gamma_{e-e} = a_{e-e}T_e^2 + b_{e-e}(\hbar\omega - E_F)^2. \tag{4}$$

This decrease in the $\gamma_{e-e}(\hbar\omega)$ relaxation rate, which reaches a minimum at the Fermi level, is caused by a decrease in the phase volume accessible to relaxation, and the a_{e-e} and b_{e-e} coefficients of (4) do not depend on $\hbar\omega$ and T_e. Although typical electron–electron relaxation times in metals at room temperature are of about 10 fs [21], interelectron interactions clearly manifest themselves in the frequency dependence of the relaxation rate in excitation–wide-band probing femtosecond experiments [8, 9].

The spectral dependence of the γ_{e-ph} electron–phonon relaxation rate near the Fermi level in the low-temperature limit (that is, at $k_BT_e \ll \hbar\omega_{ph}$, where $\hbar\omega_{ph}$ is the characteristic phonon energy) is well known (see [20] for details) and has the form $\gamma_{e-e} \propto (\hbar\omega - E_F)^3$. The proportionality factor is not directly related to the $\lambda\langle\Omega^2\rangle$ electron–phonon interaction parameter. In contrast, under excitation–wide-band probing femtosecond experiment conditions [8, 9, 13] and intense pumping, we have the other (high-temperature) limiting situation; that is, the effective T_e temperature of the electronic subsystem then substantially exceeds characteristic phonon frequencies $\hbar\omega_{ph}$. In this situation (that is, when $\hbar\omega_{ph}/k_BT_e \ll 1$), the rate of electron–phonon relaxation near the Fermi level is proportional to the electron–phonon interaction parameter [15],

$$\gamma_{e-ph}(\omega, T_L, T_e) = a_{e-ph} + b_{e-ph}(\hbar\omega - E_F)^2, \tag{5}$$

where $b_{e-ph} = \pi\hbar\lambda\langle\Omega^2\rangle/(2k_BT_e)^3$. It follows from (5) that, near the Fermi level, we must observe relaxation deceleration. Approximating the experimental $\gamma_{e-ph}(\hbar\omega)$ dependence by (5) then allows us to determine the electron–phonon interaction parameter, $\lambda\langle\Omega^2\rangle = b_{e-ph}(2k_BT_e)^3/\pi\hbar$.

It should be stressed that the use of wide-band probing [13, 14] allows the position of the Fermi level (from relaxation deceleration), electron–electron and electron–phonon interaction parameters [15], and possible deviations [17] of the spectral dependence from theoretical predictions to be determined.

4. FEMTOSECOND SPECTROSCOPY OF GOLD AND COPPER FILMS

In this Section, we describe the results obtained for Au and Cu metal films by femtosecond laser spectroscopy techniques with the use of the excitation–wide-band probing method [13, 14].

The samples were Cu and Au films 20 nm thick deposited on a quartz substrate of thickness 200 μm. Excitation was effected by optical pulses 50 fs wide with excitation photon energies $\hbar\omega_p$ = 2.34 and 2.75 eV and 70 fs wide with excitation photon energy $\hbar\omega_p$ = 5.5 eV. The intensity of excitation pulses was ~10^{11} W cm^2, and the diameter of the excitation spot was about 100 μm. The difference in optical density of the samples in transmission and reflection experiments was measured by wide-band probing in the spectral range $\hbar\omega$ = 1.6–3.2 eV, where ω is the frequency of probing. The diameter of the probing spot was approximately 80 μm. The frequency of excitation and probing pulses was 3 Hz, and the time lag was varied in steps of 7 fs. A maximum time lag amounted to 4 ps. All measurements were taken at room temperature.

The dependences of the optical densities ΔD_T of Au and Cu films on the time lag between excitation ($\hbar\omega_p$ = 2.75 eV) and probing pulses are shown in Figs. 3a and 3b. Figures 4a and 4b show the time variation of the difference transmission ΔD_T and reflection ΔD_R spectra measured for an Au film at various delay times. Figure 5 presents the analogous data for a Cu film. Similar dependences were observed for the other excitation pulse energies ($\hbar\omega_p$ = 2.34 and 5.5 eV).

The stages of response time evolution were similar for Au (Fig. 4) and Cu (Fig. 5) films at excitation pulse energies used in our experiments. At the initial time, a strongly nonequilibrium distribution of electrons was formed; electron–electron interactions determined the relaxation of this strongly nonequilibrium distribution to a quasi-equilibrium distribution of electrons with the effective electronic temperature $T_e(t) > T_0$, where T_0 is the initial equilibrium temperature of electrons and the

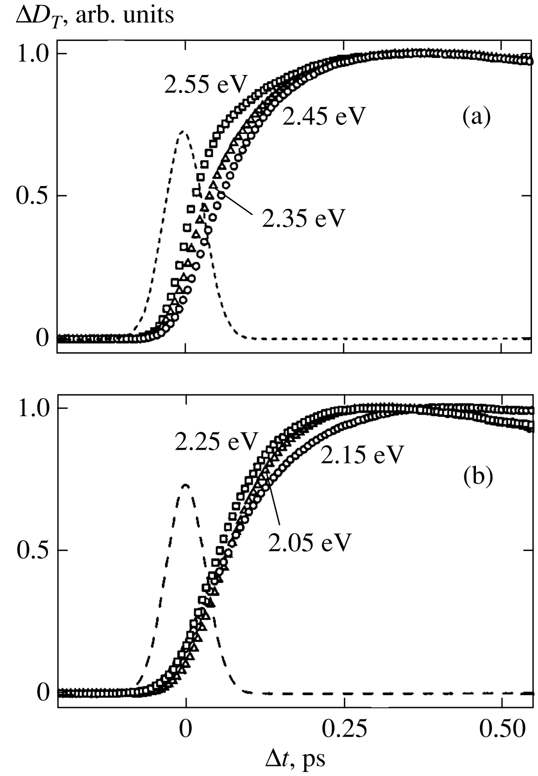


Fig. 3. Normalized time dependences of photoinduced response $\Delta D_T(\hbar\omega)/\Delta D_{\max}(\hbar\omega)$ for various probing energies $\hbar\omega$ near Fermi level E_F (a) for a Au film, $\hbar\omega_{\text{Au}} \approx 2.45$ eV, and (b) for a Cu film, $\hbar\omega_{\text{Cu}} \approx 2.15$ eV; dashed curves are pumping pulses, and Δt is the delay time.

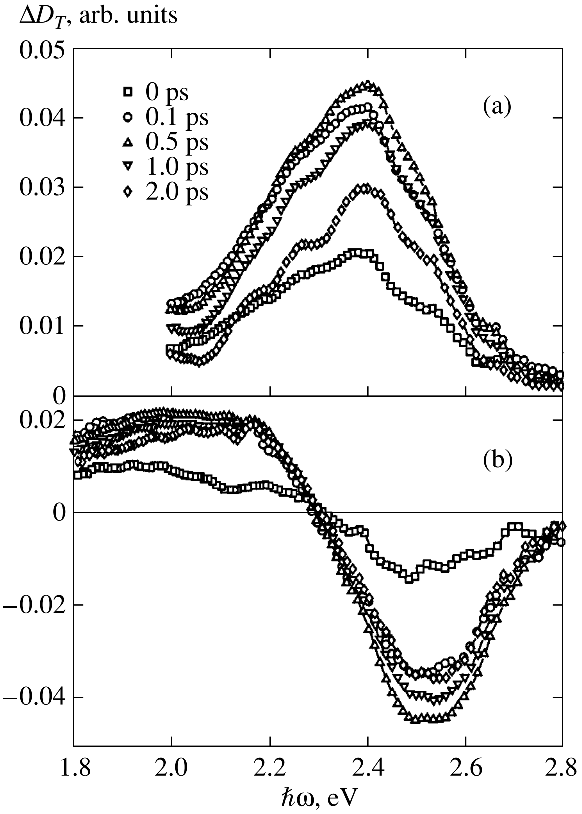


Fig. 4. Spectra of (a) ΔD_T and (b) ΔD_R for a Au film (20 nm) at various delay times Δt .

lattice. This stage of relaxation could be characterized by the $\gamma_{e\text{–}e}(\hbar\omega)$ rate of electron–electron relaxation, see (4). A maximum ΔD_T signal amplitude [for $\Delta D_T > 0$ at $\hbar\omega \approx 2.4$ eV (Au film, Fig. 4a) and $\Delta D_T > 0$ at $\hbar\omega \approx 2.15$ eV (Cu film, Fig. 5a)] was attained at times of about 500 fs and could be related to the contribution of the nonequilibrium electronic subsystem over this time interval. After attaining a quasi-equilibrium state of the electronic subsystem, further relaxation to the equilibrium state of the lattice–electrons system occurred as a result of electron–phonon interactions between quasi-equilibrium electrons and phonons. According to Figs. 4 and 5, the characteristic time of this relaxation equaled several picoseconds, and this relaxation stage could be characterized by the $\gamma_{e\text{–}ph}(\hbar\omega)$ electron–phonon relaxation rate, see (5).

Note that the $\Delta D_R(\hbar\omega)$ dependence is essentially nonmonotonic in the probing spectral region. Figures 4b and 5b show that $\Delta D_R > 0$ at $\hbar\omega < 2.15$ eV and $\Delta D_R < 0$ at $\hbar\omega > 2.15$ eV for the copper film. For the gold film, $\Delta D_R > 0$ at $\hbar\omega < 2.3$ eV and $\Delta D_R < 0$ at $\hbar\omega > 2.3$ eV. As mentioned above, this nonmonotonic dependence is related to probing transitions close to the Fermi level, and it can be used to determine the position of the Fermi level from the condition $\Delta\epsilon_2(\hbar\omega_{\text{int}}) = 0$. For the copper film, $\Delta D_R \approx \Delta\epsilon_2$ (in the spectral region close to 2.15 eV) and, therefore, $\hbar\omega_{\text{Cu}} \approx 2.15$ eV. In contrast, for the gold film, $\Delta D_R \neq \Delta\epsilon_2$ near $\hbar\omega \approx 2.4$ eV. Taking into account (2) yields $\hbar\omega_{\text{Au}} \approx 2.45$ eV [18, 19]. Note that a substantial relaxation deceleration is observed close to 2.15 eV for the copper film and close to 2.45 eV for the gold film, as is clearly shown in Fig. 3, which contains normalized time dependences of the photoinduced response at various probing energies close to the Fermi level.

To study the observed relaxation deceleration in the spectral region of transitions to the Fermi level region (see Fig. 3), the experimental dependences were approximated by two-exponential response functions including the rates of both electron–electron and electron–phonon relaxations [6, 8, 9]:

$$\begin{aligned}
\frac{dy_1}{dt} &= -y_1(\gamma_{e-e} + \gamma_{e-ph}),\\
\frac{dy_2}{dt} &= -y_2\gamma_{e-ph} + y_1\gamma_{e-e},\\
\frac{dy_3}{dt} &= (y_1 + y_2)\gamma_{e-ph}.
\end{aligned} \tag{6}$$

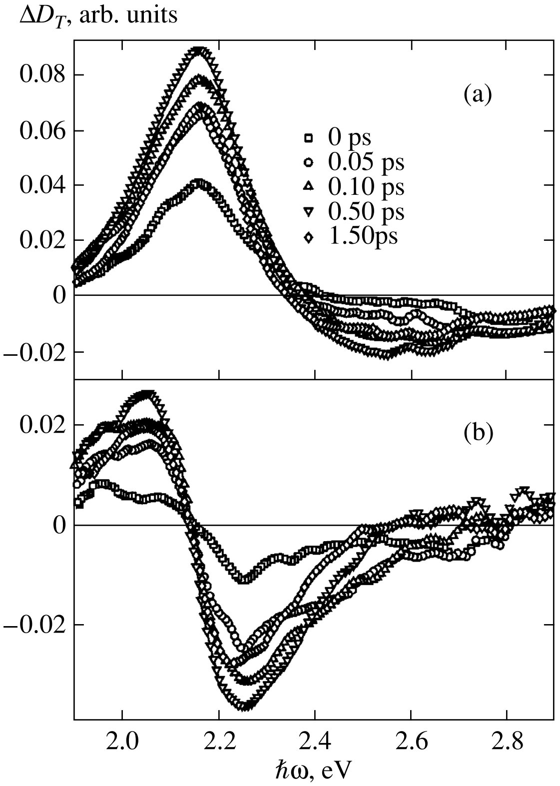


Fig. 5. Spectra of (a) ΔD_T and (b) ΔD_R for a Cu film (20 nm) at various delay times Δt.

The $\Delta D_{\text{exp}}(\hbar\omega, t)$ time dependence of the difference in optical density was modeled by the convolution of the $F(t) = \sum_{i=1}^{3} C_i y_i$ linear combination of functions obtained as solutions to system (6) with the $S_{cc}(t)$ mutual correlation function between the exciting and probing pulses,

$$\Delta D_{fit}(\hbar\omega, t) = \int_{-\infty}^{\infty} dt' F(t') S_{cc}(t - t'). \quad (7)$$

In the approximation that we use, the y_1 function is related to the process that determines the superfast response of the nonequilibrium electronic subsystem and effects quasi-equilibration of electrons at a characteristic relaxation rate of γ_{e-e}. The y_2 function may be related to the contribution of the relaxation of thermalized electrons, and the y_3 function describes lattice heating [3, 4, 8].

The purpose of approximating the experimental dependences was to obtain the best fit to the experimental data (over the whole spectral range of probing 1.6–3.2 eV) and to determine the $\gamma_{e-e}(\hbar\omega)$ and $\gamma_{e-ph}(\hbar\omega)$ spectral dependences of relaxation rates. The $\gamma_{e-e}(\hbar\omega)$ and $\gamma_{e-ph}(\hbar\omega)$ relaxation rates exhibited a strong spectral dependence (see Fig. 6), namely, the relaxation rates decreased substantially near $\hbar\omega_{\text{Cu}} \approx 2.15$ eV for the Cu film and near $\hbar\omega_{\text{Au}} \approx 2.45$ eV for the Au film. Recall that the positions of the Fermi levels determined from the spectra of ΔD_T and ΔD_R were the same. This coincidence allows us to assert with confidence that the position of the Fermi level can be found by the new method described above, namely, by determining the position of the minimum in the spectral dependence of the relaxation rate of excited charge carriers (also see [16]). Note that the dependence of relaxation rates on energy is fairly sharp, which allows the Fermi level to be determined more accurately than from the vanishing of $\Delta\epsilon_2(\hbar\omega)$ [3, 4, 10], where $\hbar\omega$ is the probing photon energy. Also recall that determining the $\Delta\epsilon_2(\hbar\omega)$ dependence requires the use of additional data [3], which are not necessary when the method under consideration is applied. The $\gamma_{e-e}(\hbar\omega)$ and $\gamma_{e-ph}(\hbar\omega)$ spectral dependences measured for the other excitation energies ($\hbar\omega_p = 2.34$ and 5.5 eV) gave results identical to those described above.

The approximation of experimental data by (4) allows us to check theoretical predictions concerning electron–electron interactions. It was found that the parameters obtained in this way noticeably differed from the values predicted for the one-component Fermi liquid [21] in the region $\hbar\omega_{\text{int}} \pm 0.2$ eV. For instance, the following parameters were obtained for the Cu and Au films: $b_{e-e}^{\text{Cu}} = 265$ ps^{-1} eV^{-2} and $b_{e-e}^{\text{Au}} = 260$ ps^{-1} eV^{-2}. Deviations from the predictions of the theory of the one-component Fermi liquid were also observed in femtosecond two-photon emission experiments [22, 23]. It follows that additional screening by valence *d* band electrons should be taken into account [24, 25].

As mentioned above, the spectral dependence of electron–phonon relaxation can be used to directly determine the electron–phonon interaction parameter (also see [15]) by approximating the experimental data according to (5). This gives $\lambda\langle\Omega^2\rangle_{\text{Cu}} \approx 35$ meV^2 and $\lambda\langle\Omega^2\rangle_{\text{Au}} \approx 28$ meV^2, in close agreement with the parameter values determined earlier from the rate of electronic temperature variations [10]. Note that the determination of the electron–phonon interaction parameter from the spectral dependence of the rate of electron–phonon relaxation [15] in excitation–probing femtosecond spectroscopy experiments is more warranted than the approach based on the Allen theory [2], which neglects the $\gamma_{e-ph}(\hbar\omega)$ spectral dependence.

5. FEMTOSECOND SPECTROSCOPY OF YBCO FILMS

The excitation–wide-band probing method of femtosecond spectroscopy was used to study the excited state of the $YBa_2Cu_3O_{7-\delta}$ high-T_c superconductor and to obtain detailed information about the relaxation of nonequilibrium charge carriers in the femtosecond time region.

The sample was a $YBa_2Cu_3O_{7-\delta}$ film ($T_c = 89$ K) 0.05 μm thick on a $SrTiO_3$ substrate of thickness 0.5 mm. Excitation was effected by optical pulses 50 fs

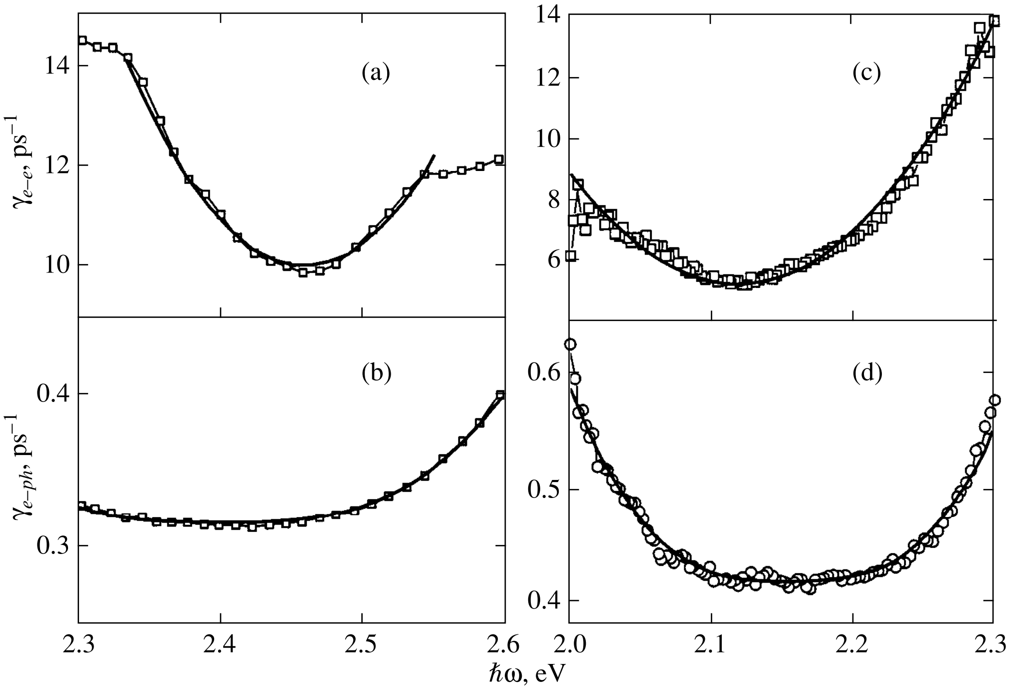


Fig. 6. Spectral dependences of relaxation rates for (a, b) Au and (c, d) Cu films 20 nm thick at an excitation pulse photon energy $\hbar\omega_p$ = 2.75 eV: (a, b) electron–electron relaxation rate $\gamma_{e-e}(\hbar\omega)$ and (c, d) electron–phonon relaxation rate $\gamma_{e-ph}(\hbar\omega)$; squares are experimental data, and solid lines are approximations by (a, c) (4) and (b, d) (5).

wide of a 2.5×10^{11} W/cm^2 intensity with photon energy $\hbar\omega_p$ = 2.34 eV. The diameter of the excitation spot was 100 μm. Reflectance variations ΔD_R were studied by 50-fs probing pulses in the energy range 1.6–3.2 eV. The diameter of the probing spot was 80 μm. The frequency of excitation and probing pulses was 2 Hz. The delay time was varied in steps of 7 fs. A maximum delay time was 4 ps. Measurements were taken at room temperature.

The method for determining the position of the Fermi level and the electron–electron and electron–phonon interaction parameters from the $\gamma_{e-e}(\hbar\omega)$ and $\gamma_{e-ph}(\hbar\omega)$ spectral dependences is described above for metals (Au and Cu). Below, this approach is applied to study $YBa_2Cu_3O_{7-\delta}$ films.

The difference in the reflectance spectra ΔD_R for various delay times in the spectral range 1.7–2.2 eV are shown in Fig. 7. Time evolution stages of these spectra are similar to those characteristic of metals (see Section 4). Note that the $\Delta D_R > 0$ signal in the studied spectral range at times of about several hundred femtoseconds may be related to the contribution of the strongly nonequilibrium electronic subsystem. At delay times exceeding 1 ps, the difference reflectance signal exhibits an essentially nonmonotonic dependence. We found that $\Delta D_R(t > 1 \text{ ps}) \approx 0$ at $\hbar\omega_1 \approx 1.9$ eV and $\hbar\omega_2 \approx 2.08$ eV (see Fig. 7) and $\Delta D_R(t > 1 \text{ ps}) < 0$ in the intermediate spec-

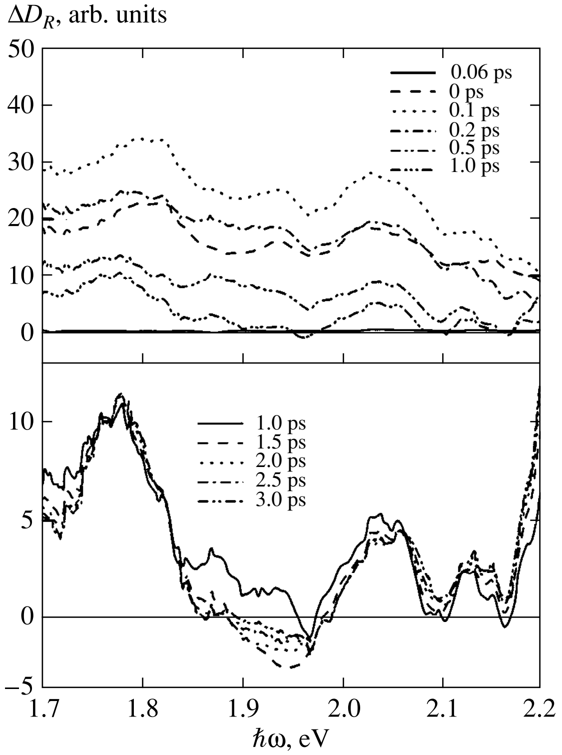


Fig. 7. Photoinduced changes in difference reflectance ΔD_R of a $YBa_2Cu_3O_{7-\delta}$ film at various delay times shown in the insets.

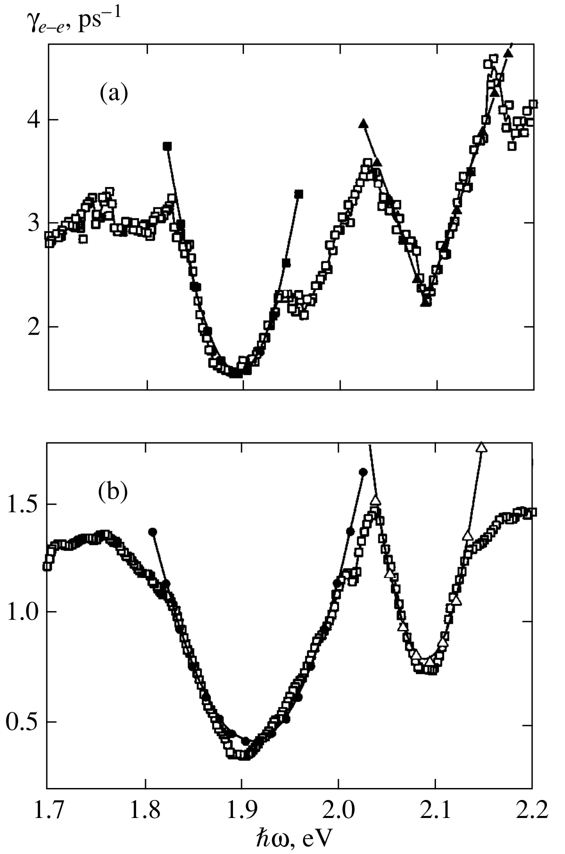


Fig. 8. Spectral dependences of (a) electron–electron γ_{e-e} and (b) electron–phonon γ_{e-ph} relaxation rates: □, experimental data; ■, $\gamma_{e-e} \propto |E - 1.89 \text{ eV}|^2$; ▲, $\gamma_{e-e} \propto |E - 2.08 \text{ eV}|$; ●, $\gamma_{e-ph} \propto |E - 1.90 \text{ eV}|^2$; and △, $\gamma_{e-ph} \propto |E - 2.09 \text{ eV}|^2$.

tral region. As mentioned above, this nonmonotonic dependence is related to probing electronic states near the Fermi level. The situation is similar to that with metals (Cu and Au, see Section 4), the only difference being that two interband transitions to the Fermi level region are observed at the $\hbar\omega_1 \approx 1.9$ eV and $\hbar\omega_2 \approx 2.08$ eV energies in the ΔD_R difference spectra.

Unlike the metals considered above, $YBa_2Cu_3O_{7-\delta}$ has a fairly complex band structure. For instance, the Fermi level passes across several bands with different effective masses of electrons and different dispersion signs [26]. Consider two types of possible interband transitions near the $\hbar\omega \approx 2$ eV energy [27, 28].

(1) Interband transitions can occur near the Γ point of the Brillouin zone from low-lying valence bands with negative effective masses of electrons to the conduction band formed by the Cu1(*dxy*)–O1(*px*)–O4(*py*) orbitals; this band contains the Fermi level. The ΔD_R value then has a nonmonotonic spectral dependence with $\Delta D_R \approx 0$ at $\hbar\omega_1 \approx 1.89$ eV, $\Delta D_R > 0$ in the spectral range $1.7 \text{ eV} < \hbar\omega < 1.9$ eV, and $\Delta D_R < 0$ in the range $1.9 \text{ eV} < \hbar\omega < 2$ eV (see Fig. 7).

(2) Interband transitions from low-lying valence bands with negative effective masses of electrons can occur to the same conduction band as with transitions (1) but with a negative effective mass of electrons near the *S* point of the Brillouin zone. The ΔD_R value then has a nonmonotonic dependence with $\Delta D_R \approx 0$ at $\hbar\omega_2 \approx 2$ eV, $\Delta D_R < 0$ in the spectral range $1.9 \text{ eV} < \hbar\omega < 2.0$ eV, and $\Delta D_R > 0$ in the spectral range $2 \text{ eV} < \hbar\omega < 2.1$ eV (see Fig. 7).

These interband transitions can be observed in the difference in the reflectance spectra of $YBa_2Cu_3O_{7-\delta}$ films [27, 28]. As shown above for Cu and Au, the position of the Fermi level for interband transitions from the valence band to the Fermi level region can be determined from the spectral dependences of the γ_{e-e} and γ_{e-ph} relaxation rates. The $\gamma_{e-e}(\hbar\omega)$ and $\gamma_{e-ph}(\hbar\omega)$ dependences were obtained following the procedure used with Cu and Au. The spectral dependences of the γ_{e-e} and γ_{e-ph} relaxation rates (Fig. 8) were found to be substantially different from those observed for the metals.

Note that both relaxation rates have minima at energies of $\hbar\omega_1 \approx 1.89$ eV and $\hbar\omega_2 \approx 2.08$ eV. These minima may be related to interband transitions of types (1) and (2) to the Fermi level region within the conduction band formed by the Cu1(*dxy*)–O1(*px*)–O4(*py*) orbitals [26]. An important difference from the metals (Cu and Au) is the closeness of the mean electron–electron and electron–phonon relaxation rates, which equal $\langle\gamma_{e-e}\rangle \approx 3$ ps^{-1} and $\langle\gamma_{e-ph}\rangle \approx 1$ ps^{-1} in the studied spectral region. For metals, these rates differ by an order of magnitude. The relaxation rates decrease near $\hbar\omega_1 \approx 1.89$ eV and $\hbar\omega_2 \approx 2.08$ eV, in agreement with the passage of the $\Delta D_R(\hbar\omega) \propto \Delta\epsilon_2$ curve through zero at these points (see Fig. 7). This agreement is another argument in favor of the conclusion that the position of the Fermi level can be determined from the deceleration of relaxation near this level. In addition, the obtained Fermi level position is close to the value found in independent experiments on the difference in the transmission response ΔD_T of $YBa_2Cu_3O_{7-\delta}$ films [10–12, 16].

The obtained spectral dependences were used to check predictions (4) of the theory of the Fermi liquids, as with Cu and Au. The electron–electron interaction parameters for the $YBa_2Cu_3O_{7-\delta}$ film were found to be $b_{e-e}(\hbar\omega_1 = 1.98 \text{ eV}) \approx 420$ and $b_{e-e}(\hbar\omega_2 = 2.08 \text{ eV}) \approx 375$ ps^{-1} eV^{-2}. Note that the b_{e-e} parameter is of the order of the values obtained for Au and Cu. At the same time, it should be stressed that substantial deviations from the theory are observed for transitions to the Fermi level region near $\hbar\omega_2 \approx 2.08$ eV. Data on $\gamma_{e-e}(\hbar\omega)$ variations can be used to study these deviations [28] by constructing dependences of the type $\gamma_{e-e}(\hbar\omega) \propto (\hbar\omega - E_F)^\alpha$ (for the Fermi liquid, $\alpha = 2$ [21]). This opens the way to the development of a new method for studying [17] manifestations of non-Fermi-liquid behaviors of strongly correlated electronic subsystems [30, 31]. The spectral dependence of the relaxation rate near $\hbar\omega_2 \approx$ 2.08 eV can well be described by the law $\gamma_{e-e}(\hbar\omega) \approx b_{e-e}|\hbar\omega - E_F|$ (see Fig. 8) with $b_{e-e}(\hbar\omega_2 = 2.08 \text{ eV}) \approx$ 27 ps^{-1} eV^{-1}. Note that a linear dependence of the electron–electron relaxation rate was also observed in experiments on the two-photon femtosecond photoemission of graphite; this dependence was interpreted in terms of electron–plasmon interactions [32]. In addition, the b_{e-e} parameter for the $YBa_2Cu_3O_{7-\delta}$ film has virtually the same value as for graphite ($b_{e-e} \approx 29$ ps^{-1} eV^{-1} [32]). Such a striking coincidence of the parameters may be explained by the two-dimensional behavior of the electronic subsystem characteristics of both $YBa_2Cu_3O_{7-\delta}$ and graphite. An alternative explanation of the linear $\gamma_{e-e}(\hbar\omega)$ dependences is based on the one-dimensional character of the conduction band formed by the Cu1$d(xy)$–O1$p(x)$–O4$p(y)$ bonding chain π orbital near the S point of the Brillouin zone [26], which makes the quasi-one-dimensional electronic liquid Luttinger in character [30, 31].

The spectral dependence of the γ_{e-ph} electron–phonon relaxation rate with deceleration at $\hbar\omega_1 \approx 1.89$ eV and $\hbar\omega_2 \approx 2.08$ eV leads us to conclude that electron–phonon interactions play an important role in $YBa_2Cu_3O_{7-\delta}$ [15]. Indeed, we found that $b_{e-ph} \approx 90$ ps^{-1} eV^{-2} at $\hbar\omega_1 \approx 1.9$ eV and $b_{e-ph} \approx 290$ ps^{-1} eV^{-2} at $\hbar\omega_2 \approx 2.09$ eV. These values were substantially larger than those determined for the metals (for Cu, $b_{e-ph} \approx 45$ ps^{-1} eV^{-2}; for Au, $b_{e-ph} \approx$ 15 ps^{-1} eV^{-2}). Let us estimate the maximum temperature of the electronic subsystem on the assumption that the excitation pulse energy is fully transferred to it. For the $YBa_2Cu_3O_{7-\delta}$ film, this gives $k_BT_e \approx 0.26$ eV, which allows the $\lambda\langle\Omega^2\rangle$ electron–phonon interaction parameter to be estimated at about 840 meV^2. This value differs from $\lambda\langle\Omega^2\rangle \approx 500$ meV^2 obtained earlier from the rate of electronic temperature variations. The theoretical estimation [33] of the electron–phonon interaction parameter gives $\lambda\langle\Omega^2\rangle \approx 2200$ meV^2. As mentioned above, the determination of $\lambda\langle\Omega^2\rangle$ from the rate of electronic temperature variations [2] ignores the spectral dependence of the electron–phonon relaxation rate, and the corresponding value can only be treated as a lower bound estimate. In addition, determining the parameter by this method requires additional experimental data [10] on the characteristics of the sample under study. Additional experimental data are not necessary for determining the parameter from the spectral dependence of the electron–phonon relaxation rate, and the resulting value is therefore more reliable [15].

6. COHERENT PHONONS IN YBCO FILMS

The use of femtosecond pulses opens up fundamentally new possibilities of studying elementary excitations in molecules and condensed media. These possibilities include the direct oscillography of molecular vibrations and probing not only the amplitudes but also the phases of molecular vibrations by ultrashort pulses of widths much smaller than the period of molecular vibrations, $\tau_p \ll T_m = \pi/\Omega_m$. The use of such laser pulses has allowed the excitation of coherent oscillations in various media to be experimentally observed [34–38]. In experiments of this type, ultrashort laser excitation pulses are used to cause impact excitation of the medium. Coherent lattice or molecule vibrations (that is, in-phase vibrations with a classical population number of the excited phonon mode) excited by such ultrashort pulses modulate medium permittivity variations, which causes changes in the intensity of transmitted or reflected light. A probing pulse with a time lag with respect to the excitation pulse allows the dynamics of excited vibrations to be directly observed.

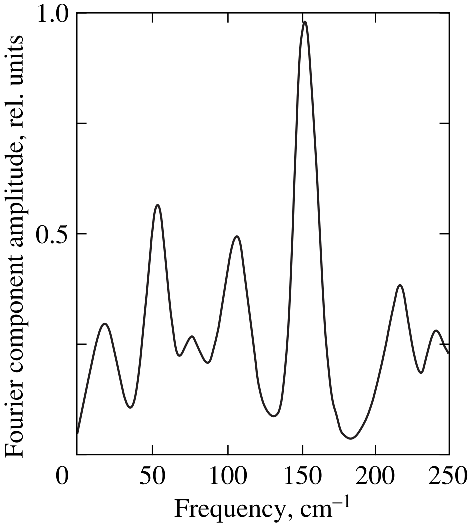


Fig. 9. Spectrum of coherent phonons excited in a $YBa_2Cu_3O_{7-\delta}$ superconductor film at room temperature.

The use of ultrashort optical pulses in the femtosecond spectroscopy of $YBa_2Cu_3O_{7-\delta}$ films allowed us to observe coherent lattice vibrations excited by them. A Fourier analysis of the time dependence of the difference in optical density variations, $\Delta D_{\exp}(t) - \Delta D_{fit}(t)$ [14, 16, 35], was used to determine the characteristic lattice vibrations in the frequency range 10–250 cm^{-1}. The excitation of coherent phonons was found to occur in the whole spectral range of probing, $\hbar\omega = 1.6$–3.0 eV. The frequencies of coherent phonons, which were effectively excited in the whole spectral range of probing, were determined from the product of all Fourier spectra; that is, from their geometric mean (also see [14, 16, 35]). A 152 cm^{-1} mode was found to dominate in this spectral range (see Fig. 9); vibrations with frequencies of about 55, 110, and 215 cm^{-1} and several other modes had lower amplitudes. The excitation of modes at about 116 and 150 cm^{-1} was earlier observed in [36–38]. For instance, femtosecond pulses about 100 fs wide with the energy of exciting and probing photons $\hbar\omega \approx 2$ eV were used in [36, 37] to study a $YBa_2Cu_3O_{7-\delta}$ film. As in this work, the excitation of the 150 cm^{-1} mode, most intense at room temperature, was observed.

7. CONCLUSIONS

In this work, we studied the dynamics of nonequilibrium processes in Au, Cu, and $YBa_2Cu_3O_{7-\delta}$ films. We analyzed the spectral dependences of both electron–electron and electron–phonon relaxations. Relaxation deceleration near the Fermi level was observed experimentally, which offers the possibility of developing a new method for determining the position of the Fermi level and the electron–electron and electron–phonon interaction parameters. A method for determining the electron–phonon interaction parameter from the spectral dependence of the rate of electron–phonon relaxation was suggested. Studies of the spectral dependences of relaxation rates open up new possibilities for determining possible deviations from the Fermi liquid behavior. We observed a linear dependence of the relaxation rate, $\gamma_{e-e}(\hbar\omega) \approx b_{e-e}|\hbar\omega - E_F|$, for the interband transition at $\hbar\omega \approx 2.08$ eV in the $YBa_2Cu_3O_{7-\delta}$ high-T_c superconductor. The excitation of coherent phonons with predominant vibrations at an about 152 cm^{-1} frequency was observed for the $YBa_2Cu_3O_{7-\delta}$ film.

ACKNOWLEDGMENTS

This work was financially supported by the Russian Foundation for Basic Research and a DFG grant.

REFERENCES

1. S. I. Anisimov, B. L. Kapeliovich, and T. L. Perel’man, Zh. Éksp. Teor. Fiz. **66**, 776 (1974) [Sov. Phys. JETP **39**, 375 (1974)].
2. P. B. Allen, Phys. Rev. Lett. **59**, 1460 (1987).
3. S. D. Brorson, J. G. Fujimoto, and E. P. Ippen, Phys. Rev. Lett. **59**, 1962 (1987).
4. T. Juhasz, H. E. Elsayed-Ali, G. O. Smith, *et al.*, Phys. Rev. B **48**, 15488 (1993).
5. W. S. Fann, R. Storz, H. W. Tom, *et al.*, Phys. Rev. Lett. **68**, 2834 (1992); Phys. Rev. B **46**, 13592 (1992).
6. C.-K. Sun, F. Vallee, L. Acioli, *et al.*, Phys. Rev. B **48**, 12365 (1993); **50**, 15337 (1994).
7. R. H. M. Groeneveld, R. Sprik, and A. Lagendijk, Phys. Rev. B **45**, 5079 (1992); **51**, 11433 (1995).
8. V. M. Farztdinov, A. L. Dobryakov, S. A. Kovalenko, *et al.*, Phys. Scr. **60**, 579 (1999).
9. Yu. E. Lozovik, S. A. Kovalenko, A. L. Dobryakov, *et al.*, Laser Phys. **9**, 557 (1999).
10. S. D. Brorson, A. Kazeroonian, J. S. Moodera, *et al.*, Phys. Rev. Lett. **64**, 2172 (1990).
11. S. V. Chekalin, V. M. Farztdinov, Yu. E. Lozovik, *et al.*, Phys. Rev. Lett. **67**, 3860 (1991).
12. A. L. Dobryakov, V. V. Golovlev, Yu. E. Lozovik, *et al.*, Opt. Spektrosk. **76**, 975 (1994) [Opt. Spectrosc. **76**, 871 (1994)].
13. S. A. Kovalenko, A. L. Dobryakov, J. Ruthmann, *et al.*, Phys. Rev. A **59**, 2369 (1999).
14. V. M. Farztdinov, A. L. Dobryakov, N. P. Ernsting, *et al.*, Phys. Rev. B **56**, 4176 (1997).
15. A. L. Dobryakov and Yu. E. Lozovik, Pis’ma Zh. Éksp. Teor. Fiz. **70**, 329 (1999) [JETP Lett. **70**, 329 (1999)].
16. I. I. Vengrus, A. L. Dobryakov, S. A. Kovalenko, *et al.*, Pis’ma Zh. Éksp. Teor. Fiz. **62**, 739 (1995) [JETP Lett. **62**, 758 (1995)].
17. A. L. Dobryakov, Yu. E. Lozovik, S. A. Kovalenko, *et al.*, Phys. Lett. A **223**, 303 (1996).
18. R. Rosei, F. Antonangeli, and U. M. Grassano, Surf. Sci. **37**, 689 (1973).
19. A. L. Dobryakov, V. M. Farztdinov, Yu. E. Lozovik, *et al.*, Phys. Scr. **60**, 572 (1999).
20. V. F. Gantmakher and I. B. Levinson, *Scattering of Carriers in Metals and Semiconductors* (Nauka, Moscow, 1984).
21. D. Pines and P. Noziéres, *Theory of Quantum Liquids* (Benjamin, New York, 1966; Mir, Moscow, 1967).
22. T. Hertel, E. Knoesel, M. Wolf, *et al.*, Phys. Rev. Lett. **76**, 535 (1996).
23. J. Cao, Y. Gao, R. J. Miller, *et al.*, Phys. Rev. B **56**, 1099 (1997).
24. S. Ogawa, H. Nagano, and H. Petek, Phys. Rev. B **55**, 10869 (1997).
25. A. V. Lugovskoy and I. Bray, J. Phys. D **31**, L78 (1998); Phys. Rev. B **60**, 3279 (1999).
26. W. E. Pickett, Rev. Mod. Phys. **61**, 433 (1989).
27. A. L. Dobryakov, V. M. Farztdinov, and Yu. E. Lozovik, Phys. Rev. B **47**, 11515 (1993).
28. A. L. Dobryakov, V. M. Farztdinov, Yu. E. Lozovik, *et al.*, Opt. Commun. **105**, 309 (1994).
29. P. W. Anderson, T. V. Ramakrishnan, S. Strong, *et al.*, Phys. Rev. Lett. **77**, 4241 (1996); Science **261**, 337 (1993).
30. P. W. Anderson, Science **235**, 1196 (1987); **256**, 1526 (1992).
31. S. Xu, J. Cao, C. C. Miller, *et al.*, Phys. Rev. Lett. **76**, 483 (1996).
32. G. L. Zhao and J. Callaway, Phys. Rev. B **49**, 6424 (1994).
33. R. Merlin, Solid State Commun. **102**, 207 (1997).
34. A. L. Dobryakov, S. A. Kovalenko, Yu. E. Lozovik, *et al.*, Pis’ma Zh. Éksp. Teor. Fiz. **61**, 957 (1995) [JETP Lett. **61**, 985 (1995)].
35. W. A. Kutt, W. Albrecht, and H. Kurz, IEEE J. Quantum Electron. **28**, 2434 (1992).
36. W. Albrecht, Th. Kruse, and H. Kurz, Phys. Rev. Lett. **69**, 1451 (1990).
37. J. M. Chwalek, C. Uher, J. F. Whitaker, *et al.*, Appl. Phys. Lett. **58**, 980 (1991).
38. S. J. Nettel and R. K. MacCrone, Phys. Rev. B **49**, 6395 (1994).

Translated by V. Sipachev

Journal of Experimental and Theoretical Physics, Vol. 92, No. 2, 2001, pp. 277–286.
From Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 119, No. 2, 2001, pp. 321–330.
Original English Text Copyright © 2001 by González de la Cruz, Gurevich, Kucherenko, Ramírez de Arellano.

SOLIDS
Electronic Properties

Contact Effects on the Magnetoresistance of Finite Semiconductors[¶]

G. González de la Cruz[a], Yu. G. Gurevich[a, *], V. V. Kucherenko[b], and E. Ramírez de Arellano[b]

[a]*Departamento de Física, Centro de Investigacion y Estudios Avanzados del Instituto Politécnico Nacional, Apartado Postal 14-740 México, D.F.*

**e-mail: gurevich@fis.cinvestav.mx*

[b]*Departamento de Matemáticas, ESFM del IPN; Departamento de Matemáticas, Centro de Investigacion y Estudios Avanzados del Instituto Politécnico Nacional, Apartado Postal 14-740 México, D.F.*

Received February 8, 2000

Abstract—We propose a new theoretical method to study galvanomagnetic effects in bounded semiconductors. The general idea of this method is as follows. We consider the electron temperature distribution and the electric potential as consisting of two terms, one of which represents the regular solution of the energy balance equation obtained from the Boltzmann transport equation at steady-state conditions and the Maxwell equation, while the other is the effect of the specimen size that is significant near the contacts (the boundary layer function). With the distribution of the electric potential at the contacts and the electron temperature distribution at the surface of the sample taken into account, we find that the magnetoresistance is different from the one in the standard theory of galvanomagnetic effects in boundless media. We show that, besides the usual quadratic dependence on the applied magnetic field B, the magnetoresistance can exhibit a linear dependence on B under certain conditions. We obtain new formulas for the linear and quadratic terms of the magnetoresistance in bounded semiconductors. This linear contribution of the magnetic field to the magnetoresistance is essentially due to the spatial dependence of the potential at the electric contacts. We also discuss the possibility of obtaining the distribution of the electric potential at the contacts from standard magnetoresistance experiments. Because the applied magnetic field acts differently on carriers with different mobilities, a redistribution of the electron energy occurs in the sample and thus, the Ettingshausen effect on the magnetoresistance must be considered in bounded semiconductors. *© 2001 MAIK “Nauka/Interperiodica”.*

1. INTRODUCTION

Physically, the magnetoresistance phenomenon consists in an increase of the electric resistance of a metal or semiconductor subject to an external magnetic field applied transversally to the electric field direction. We obtain a complete formula for the magnetoresistance in a bounded semiconductor involving several previously unknown terms. Using the expression for the magnetoresistance in bounded semiconductors, it is possible to obtain some information about the electron energy relaxation, the carrier density, and the electron temperature distribution in the semiconductor. Currently, the innovation of some sensitive magnetic field detectors is based on the magnetoresistance effect in semiconductors. This means that the linear contribution of the magnetic field to the magnetoresistance established in this paper, which arises due to the spatial dependence of the potential at electric contacts, can improve the sensitivity of the devices. Furthermore, the experimental measurements of magnetoresistance allow one to describe the homogeneity of the electric potential at the contacts and therefore also the homogeneity of the current density in the sample, which is very important for semiconductor devices.

Most of the theoretical works, as far as galvanomagnetic effects in bulk semiconductors are concerned, have been addressed to boundless media where the electric field is constant in all directions and the only contribution to the magnetoresistance is related to the dependence of the electric conductivity on the magnetic field [1, 2]. However, this assumption implicitly involves the effect of the sample surface, because the electrostatic Hall field, and thus the magnetoresistance, cannot be found otherwise. It is worth mentioning that, in reality, it is a usual practice to fix some specific boundary conditions at the surface of the sample; as a consequence, in general, magnetoresistance depends on the electric potential, which is a linear function of the coordinates [3]. Moreover, this linear term can only be calculated if the surface effects on the electric potential are considered through an additional function of coordinates. The coefficients characterizing the potential also depend strongly on these boundaries and, as the result, they are different from the coefficients obtained in the standard magnetoresistance theory.

[¶] This article was submitted by the authors in English.

Size-dependent contributions to the magnetoresistance of an isotropic semiconductor in a uniform electric field E_x and a transverse magnetic field B (in the y-direction) have been discussed in [4–7]. The discussion is given for systems bounded along only one direction (the z-axis) and boundless in the direction of the electric field. The current density is taken to vanish at the surface of the sample, which is viewed as a boundary condition (i.e., $j_z = 0$ at $z = \pm b$) in contrast with standard magnetoresistance theory, where $j_z = 0$ in the semiconductor sample. In this case, the electron temperature gradient $\partial T_e/\partial z$ arises because the magnetic field acts in a different way on carriers of different mobilities (the Ettingshausen effect) [8], which leads to a linear dependence of the electron temperature distribution on the electric field. The experimental evidence of these theoretical results has shown a strong influence of the semiconductor thickness on the magnetoresistance. When the Ettingshausen effect in bounded semiconductors is taken into account, a size-dependent term appears in the magnetoresistance. However, when the transverse dimensions of the semiconductor are very large compared to the electron–phonon energy relaxation length (k^{-1}) [9], the usual result of conventional magnetoresistance theory is recovered, with the Ettingshausen effect being important if $kb \le 1$. On the other hand, the size-dependent contribution to the magnetoresistance does not disappear in the limit as $kb \longrightarrow 0$ [10] and is in fact of the same order as the physical magnetoresistance term in the standard theory.

As can be seen, the surfaces of the sample play an important role in the theory of magnetoresistance in thin-film semiconductors. However, in real physical experiments on magnetoresistance, besides the effect of the size, the effects due to the inhomogeneity of the potentials at the contacts must be considered.

Magnetoresistance and the electric potential distribution in a bounded metal (degenerate electron gas) have been investigated in [11, 12]; in [12], in particular, it was studied using a conformal transformation in the complex plane. This approach is only valid when the electric potential is constant at the contacts, i.e., is independent of the coordinates; the approach cannot be applied to semiconductors where the current depends on the potential and the temperature and satisfies the Helmholtz equation.

In the limit of small electric and magnetic fields, size-dependent contributions of the magnetoresistance of an isotropic semiconductor have been considered in [13, 14] using a perturbative method. The relevant discussion is given for systems bounded in all directions, with the current density vanishing at $z = \pm b$. It is found that magnetoresistance exists even if the relaxation time is independent of the electron energy. However, when the distance between the contacts is very large, the perturbative approach of [13, 15] looses its applicability.

Recently, magnetoresistance in bulk semiconductors that are bounded in all directions was investigated within a new mathematical approach [3] for a degenerate electron gas, the result being a simple analytical expression. Moreover, it was shown in [7] that the carrier temperature distribution for a nondegenerate semiconductor (the Ettingshausen effect) plays an important role in the study of galvanomagnetic effects.

In this work, we analyze the magnetoresistance in bounded isotropic nondegenerate semiconductors and consider the effect of the inhomogeneous electric potentials at the contacts, thickness b, and length a of a thin-film semiconductor. This analysis is based on representing the potential and the temperature as the sum of a term that is regular (analytical) in the small parameters b/a and $\omega_H\tau_0$ and a term involving the boundary layer functions corresponding to vortex currents. The boundary layer functions are essential near the contacts. They vanish as the magnetic field $B \longrightarrow 0$ for a constant potential at the contacts, are regular in the small parameter $\omega_H\tau_0$, and decay exponentially along the sample. The analysis shows that it should be possible to observe an interesting electronic transport phenomenon caused by the electric field and the electron temperature distributions; moreover, the magnetoresistance that we find is different from the one in the standard theory.

2. THEORETICAL MODEL

We assume that a semiconductor sample has the shape of a parallelepiped bounded by the $x = 0, a$; $y = 0, c$; and $z = 0, b$ planes and the electric contacts with the distributions $\varphi^0(y, z)$ and $\varphi^a(y, z)$ are in the $x = 0$ and $x = a$ planes, respectively, while the applied uniform magnetic field is directed along the y-axis. The normal components of the current density vanish at the $y = 0$, c and $z = 0$, b planes of the sample (open circuit at these surfaces). If the potential distributions $\varphi^0(y, z)$ and $\varphi^a(y, z)$ are only functions of z, the transport problem is obviously two-dimensional (all the physical parameters only depend on x and z). We consider the effect that the redistribution of carriers according to their energy across the sample has on the magnetoresistance (the Ettingshausen effect). Assuming that the electric and magnetic fields are weak (and therefore, $T_e - T_0 \approx jB$, where T_0 is the ambient temperature), we can use the Maxwell and the thermal balance equations to find the electron temperature distribution and the electrostatic potential in the sample as functions of coordinates and the magnetic field. Under steady-state conditions, the equations for the coupled electron temperature and the electric potential can be written as [7, 14]

$$\begin{gathered}\nabla^2\varphi(x, z) + \frac{q+1}{e}\nabla^2 T_e(x, z) = 0. \\ \nabla^2 T_e(x, z) + \frac{e}{q+2}\nabla^2\varphi(x, z) = k^2(T_e(x, z) - T_0),\end{gathered} \qquad (1)$$

where k^{-1} is the scale length of the electron–phonon energy relaxation, referred to as the cooling length ($k^{-1} \approx 10^{-3}$–10^{-4} cm for nondegenerate semiconductors), and q is a parameter characterizing the dependence of the momentum relaxation time τ on energy ε via $\tau(\varepsilon) = \tau_0(\varepsilon/T_0)^q$. The values of q for various momentum relaxation mechanisms are given in [16] (it is important that $|q| < 3/2$). In this work, we assume that the temperature of the phonon system is equal to the ambient temperature T_0.

To arrive at Eqs. (1), we have assumed that the electron gas is nondegenerate (satisfies the Maxwell statistics), the energy–momentum relation is quadratic and isotropic, and the current density is sufficiently small for the nonlinear effects to be negligible, i.e., the kinetic coefficients do not depend on the electric field. We also consider a weak magnetic field such that $\omega_H\tau_0 \ll 1$, where ω_H is the cyclotron frequency.

The continuity and the energy balance equations for the potential $\varphi(x, z)$ and the electron temperature $T_e(x, z)$ must be supplemented by boundary conditions describing the distribution of the potential at the electric contacts and the normal components of the current density at the lateral surfaces:

$$\varphi(x,z)\big|_{x=0} = \varphi^0(z), \quad \varphi(x,z)\big|_{x=a} = \varphi^0(z), \qquad j_z\big|_{z=0,b} = 0. \tag{2}$$

The coupled equations for the potential and the electron temperature must be supplemented by boundary conditions describing the absorption of the carrier energy at the surface of the sample. These conditions can be written as [17]

$$Q_n\big|_s = \eta_s(T_e - T_0)\big|_s, \tag{3}$$

where Q_n is the electron normal component of the heat flux at the surface of the sample and parameter η_s represents the inelastic scattering of electrons at the boundaries (surface heat conductivity), with $\eta_s = 0$ corresponding to the absence of surface mechanisms, that is,

$$Q_z\big|_{z=0,b} = 0 \tag{4}$$

in our geometry, and with infinite η_s corresponding to a good thermal conductivity across the surface. We consider this latter boundary condition for the electron temperature at the contacts; i.e.,

$$T_e\big|_{x=0,a} = T_0. \tag{5}$$

Under the above assumptions, we see from the expressions for j and Q given in [18] that the potential and the temperature distributions satisfy the following equations at the surface of the sample, where $j_z|_{z=0,b} = Q_z|_{z=0,b} = 0$

$$\begin{aligned}
&\frac{\partial\varphi}{\partial z} + \frac{q+1}{e}\frac{\partial T_e}{\partial z} + \frac{\Gamma(2q+5/2)}{\Gamma(q+5/2)} \\
&\times(\omega_H\tau_0)\left(\frac{\partial\varphi}{\partial x} + \frac{2q+1}{e}\frac{\partial T_e}{\partial x}\right)\bigg|_{z=0,b} = 0, \\
&\frac{\partial\varphi}{\partial z} + \frac{q+2}{e}\frac{\partial T_e}{\partial z} + \frac{\Gamma(2q+7/2)}{\Gamma(q+7/2)} \\
&\times(\omega_H\tau_0)\left(\frac{\partial\varphi}{\partial x} + \frac{2q+2}{e}\frac{\partial T_e}{\partial x}\right)\bigg|_{z=0,b} = 0,
\end{aligned} \tag{6}$$

with $\Gamma(x)$ being the gamma function.

Assuming the difference in potential at the contacts to be small, which means restricting ourselves to the transport effects that are linear in the electric field, we see from [14, 19] that in the theory of galvanomagnetic phenomena with the electron temperature distribution taken into account, the x component of the current density is given by

$$\begin{aligned}
j_x &= -\sigma_0\frac{\partial\varphi}{\partial x} - \frac{(q+1)\sigma_0}{e}\frac{\partial T_e}{\partial x} + \sigma_0(\omega_H\tau_0) \\
&\times\frac{\Gamma(2q+5/2)}{\Gamma(q+5/2)}\left[\frac{\partial\varphi}{\partial z} + \frac{(2q+1)}{e}\frac{\partial T_e}{\partial z}\right] \\
&+ \sigma_0(\omega_H\tau_0)^2\frac{\Gamma(3q+5/2)}{\Gamma(q+5/2)}\left[\frac{\partial\varphi}{\partial z} + \frac{(3q+1)}{e}\frac{\partial T_e}{\partial z}\right],
\end{aligned} \tag{7}$$

where

$$\sigma_0 = \frac{4\Gamma(q+5/2)}{3\sqrt{\pi}}\frac{ne^2\tau_0}{m}.$$

The first term in Eq. (7) corresponds to the usual current; the second term corresponds to the thermoelectric current; the third term corresponds to the Hall effect and the transverse Nernst-Ettingshausen effect. The last term in Eq. (7) describes the longitudinal Nernst–Ettingshausen effect.

3. ASYMPTOTIC APPROXIMATION FOR MAGNETORESISTANCE

For small magnetic fields such that $(\omega_H\tau_0)^2 \ll 1$, we naturally seek solutions of Eqs. (1) in the form

$$\begin{aligned}
\varphi(x,z) &= \varphi_0(x,z) + \varphi_1(x,z)(\omega_H\tau_0) \\
&+ \varphi_2(x,z)(\omega_H\tau_0)^2 + \ldots, \\
T_e(x,z) &= T_0 + T_1(x,z)(\omega_H\tau_0) \\
&+ T_2(x,z)(\omega_H\tau_0)^2 + \ldots.
\end{aligned} \tag{8}$$

To calculate the terms $\varphi_j(x, z)$ and $T_j(x, z)$, we propose a new nonstandard perturbation theory with respect to the small magnetic field. This theory is uniform with respect to the small parameter b/a. Inserting Eqs. (8) in Eq. (7), we can write the x component of the current density to the second order of the magnetic field as

$$\begin{aligned} j_x(x, z) &= j_0(x, z) + j_1(x, z)(\omega_H\tau_0) \\ &\quad + j_2(x, z)(\omega_H\tau_0)^2 + \ldots, \end{aligned} \tag{9}$$

where

$$\begin{aligned} j_0(x, z) &= -\sigma_0\frac{\partial\varphi_0}{\partial x}, \\ j_1(x, z) &= -\sigma_0\left(\frac{\partial\varphi_1}{\partial x} + \frac{q+1}{e}\frac{\partial T_1}{\partial x}\right)\frac{\Gamma(2q+5/2)}{\Gamma(q+5/2)}, \\ j_2(x, z) &= -\sigma_0\left(\frac{\partial\varphi_2}{\partial x} + \frac{q+1}{e}\frac{\partial T_2}{\partial x} - \frac{\Gamma(2q+5/2)}{\Gamma(q+5/2)}\right. \\ &\quad \left.\times\left\{\frac{\partial\varphi_1}{\partial z} + \frac{2q+1}{e}\frac{\partial T_1}{\partial z}\right\} - \frac{\Gamma(3q+5/2)}{\Gamma(q+5/2)}\frac{\partial\varphi_0}{\partial x}\right). \end{aligned} \tag{10}$$

The average value of the current density over the semiconductor cross section that is significant for the magnetoresistance is given by

$$\bar{j} = \frac{1}{b}\int_0^b j_x(x, z)dz. \tag{11}$$

Because $\operatorname{div}\mathbf{j} = 0$, j is x-independent.

It is clear from the above that a detailed analysis of $\bar{j}$ is a very complicated problem. As we see in what follows, however, an analytical expression for the average current density can be obtained in the limit where $b/a \ll 1$. This condition allows us to study galvanomagnetic effects in semiconductors; depending on the results, we can decide whether it is possible to talk about the effects of the finite dimension of the sample on the magnetoresistance.

We now restrict ourselves to thin-film semiconductors with $a \gg b$. Because the cooling length is of the order 1 μm, we can use the relation

$$a \gg k^{-1}, b. \tag{12}$$

Alternatively, if the geometry of the sample is such that $a \ll b$, the distribution of the current density j_z corresponds to the closed Hall contacts [19].

We introduce the average potential at the contacts $x = 0$ and $x = a$ as

$$\bar{\varphi}^0 = \frac{1}{b}\int_0^b \varphi^0(z)dz, \quad \bar{\varphi}^a = \frac{1}{b}\int_0^b \varphi^a(z)dz. \tag{13}$$

Note that, if the distribution of the potential is constant at the contacts of the sample, we have $\varphi^0(z) = \bar{\varphi}^0$ and $\varphi^a(z) = \bar{\varphi}^a$, otherwise it depends on the z-coordinate.

For a constant potential at the contacts and in the presence of a weak magnetic field, the magnetoresistance can be defined as

$$\delta = \frac{(\bar{j} - \bar{j}_0)a}{(\bar{\varphi}^a - \bar{\varphi}^0)\sigma_0}.$$

In the case where $\varphi^0(z) = \bar{\varphi}^0$ and $\varphi^a(z) = \bar{\varphi}^a$, the magnetoresistance is given by

$$\delta = \left[\delta_0 - \frac{b}{a}(K + F(kb))\right](\omega_H\tau_0)^2 \tag{14}$$

(the proof of this formula is given in Section 6). It follows that

$$\begin{aligned} \delta_0 &= \frac{\Gamma(5/2+q)\Gamma(5/2+3q) - \Gamma^2(5/2+2q)}{\Gamma^2(5/2+q)} \\ &\quad - \frac{\dfrac{q^2}{5/2+q}\Gamma^2(2q+5/2)}{\dfrac{kb}{2}(q+2)^{1/2}\Gamma^2(q+5/2)}\tanh[(q+2)^{1/2}kb/2] \end{aligned}$$

is the magnetoresistance for samples such that the dimension along the x direction is infinite ($a \longrightarrow \infty$) and the transverse dimension b is finite. The formulas for coefficient K and function $F(kb)$ have not been known previously. We obtain that

$$K = \frac{\Gamma^2(2q+5/2)}{\Gamma^2(q+5/2)}\frac{16}{\pi^3}\sum_{i=0}^{\infty}\frac{1}{(2l+1)^3}, \tag{15}$$

$$\begin{aligned} F(kb) &= \frac{8q^2}{q+5/2}\frac{\Gamma^2(2q+5/2)}{\Gamma^2(q+5/2)} \\ &\quad \times \sum_{i=0}^{\infty}[\pi^2(2l+1)^2 + (kb)^2(q+2)]^{-3/2}. \end{aligned} \tag{16}$$

It follows from Eqs. (14)–(16) that, when the distribution of the potential is uniform at the contacts, the correction term to the magnetoresistance depends on the ratio $b/a \ll 1$ linearly rather than exponentially via $\exp(-a/b)$, as is assumed in the standard theory of galvanomagnetic effects in semiconductors. On the other hand, if the electric potential is inhomogeneous at the contacts, the magnetoresistance is given by

$$\delta = -\frac{4}{\pi b(\bar{\varphi}^a - \bar{\varphi}^0)}\left\{\int_0^b [\varphi^0(z) - \varphi^a(z) - \bar{\varphi}^0 - \bar{\varphi}^a] \times \sum_{l=0}^{\infty} \frac{\cos[(2l+1)\pi z/b]}{2l+1} dz\right\} \frac{\Gamma^2(2q+5/2)}{\Gamma^2(q+5/2)}(\omega_H\tau_0) \tag{17}$$

(the proof of this formula is given in Section 5). In this case, the magnetoresistance depends on the magnetic field linearly rather than quadratically as in the usual theory of galvanomagnetic effects in semiconductors. In addition, it changes sign when the magnetic field is reversed. Thus, the resistance in the sample decreases with the magnetic field before reversing its sign. We note that the sign in Eq. (17) strongly depends on the potential distribution at the contacts and is independent of the length a of the sample in the first approximation with respect to the magnetic field. Size effects on the magnetoresistance occur in the second-order approximation with respect to B. For example, if $\varphi^0(z) + \varphi^a(z) - \bar{\varphi}^0 - \bar{\varphi}^a = C(z - b/2)$, it follows from Eq. (17) that

$$\delta = \frac{8C\omega_H\tau_0}{\pi^3(\bar{\varphi}^a - \bar{\varphi}^0)} \frac{\Gamma(2q+5/2)}{\Gamma(q+5/2)} \sum_{l=0}^{\infty} (2l+1)^{-3}.$$

We note that Eq. (14) gives the magnetoresistance with the precision $[(\omega_H\tau_0)^3 + e^{(-\pi a/2b)}(\omega_H\tau_0)^2]$, and Eq. (17), with the precision $(\omega_H\tau_0)^2$. Therefore, Eq. (14) gives the correct results in case where $(\omega_H\tau_0$ 1 and $b/a \ll e^{(-\pi a/2b)}$; this does not necessarily imply the constraint $b/a \ll 1$. Equation (17) is applicable in the cases where $(\omega_H\tau_0) \ll 1$ and $(\omega_H\tau_0) \ll |\delta|$. We see that for the potential that is homogeneous at the contacts, we have $\delta_0 = 0$ for the degenerate electron gas, that is, for $q = 0$. This implies that the standard mechanisms of creating magnetoresistance do not work and the magnetoresistance is the result only of the mechanism proposed in this paper. However, if the linear part of magnetoresistance in the magnetic field does not vanish, it does not vanish for all values of q. This means that inhomogeneity of the potential at the contact plane is a new mechanism of creating magnetoresistance. The linear dependence coefficient in Eq. (17) is a product of two factors. The first factor depends only on the potential distributions at the contact planes. The second factor results from the Ettingshausen effect and is independent of the potential distribution. It follows from Eq. (17) that if we know the potential distributions at the contacts, we can calculate the parameter q of the relaxation mechanism using the magnetoresistance.

It is worth mentioning that if the magnetoresistance is calculated in all orders in the magnetic field, the potential distribution at the contacts can be evaluated explicitly. The solution in the form of a Taylor expansion has been exactly obtained only for the degenerate electron gas (metals) [15, 20]. Thus, experimental measurements of magnetoresistance allow one to shed some light on the distribution of the potential at the contacts.

4. MAGNETORESISTANCE CALCULATION FOR RECTANGULAR SAMPLES

We now proceed to describe a method of solving the problem concerning the two-dimensional potential and electron temperature distribution for magnetoresistance in the presence of a weak magnetic field. The geometry considered is again that of a rectangular semiconductor. We introduce a new function Φ depending on the potential and the electron temperature distribution such that current J_x is expressed through this function up to the order $(\omega_H\tau_0)^2$ (see Eq. (7)) as

$$\Phi = \varphi + \frac{q+1}{e}T_e, \quad T = T_e - T_0 \tag{18}$$

and the dimensionless variables $x' = x/b$ and $z' = z/b$ are such that $0 < x' < \beta^{-1}$ and $0 < z' < 1$, where $\beta = b/a$. With these new functions, Eq. (1) can be written as (we omit the prime on the variables)

$$\nabla^2\Phi = 0, \quad \nabla^2 T - (q+2)(kb)^2 T = 0 \tag{19}$$

and the boundary conditions in Eqs. (4)–(6) become

$$\Phi\big|_{x=\beta^{-1},\,0} = \varphi^a(z) + \frac{q+1}{e}T_0, \quad T\big|_{x=0,\,\beta^{-1}} = 0,$$

$$\left.\frac{\partial\Phi}{\partial z} + \alpha(\omega_H\tau_0)\frac{\partial\Phi}{\partial x} + \frac{\alpha q}{e}(\omega_H\tau_0)\frac{\partial T}{\partial x}\right|_{z=0,\,1} = 0, \tag{20}$$

$$\left.\frac{\partial T}{\partial z} + \chi(\omega_H\tau_0)\frac{\partial T}{\partial x} + \gamma(\omega_H\tau_0)\frac{\partial\Phi}{\partial x}\right|_{z=0,\,1} = 0.$$

$$\alpha = \frac{\Gamma(2q+5/2)}{\Gamma(q+5/2)},$$

$$\chi = (q+1)\frac{\Gamma(2q+5/2)}{\Gamma(q+5/2)} - q\frac{\Gamma(2q+5/2)}{\Gamma(q+5/2)}. \tag{21}$$

$$\gamma = e\left[\frac{\Gamma(2q+7/2)}{\Gamma(q+7/2)} - \frac{\Gamma(2q+5/2)}{\Gamma(q+5/2)}\right].$$

In most of the theoretical works related to galvanomagnetic effects in bulk semiconductors, solutions of Eqs. (19) are represented as infinite series in $\omega_H\tau_0$ for weak magnetic fields; to obtain approximations for the coefficients Φ_k and T_k of the orders $k = 0, 1, \ldots.$ the authors neglect the terms $(\omega_H\tau_0)\partial\Phi_k/\partial x$ and $(\omega_H\tau_0)\partial T_k/\partial x$ in boundary conditions (20). However, the exact solutions for the degenerate electron gas [15] demonstrate that this series diverges for large samples, i.e., for $a \gg b$. For this reason, we now seek solutions of Eqs. (19) in the form

$$\begin{aligned}\Phi &= \Phi_0(x, z, \omega_H\tau_0) + \Phi_1(x, z, \omega_H\tau_0)(\omega_H\tau_0)^1 \\ &+ \Phi_2(x, z, \omega_H\tau_0)(\omega_H\tau_0)^2 + O((\omega_H\tau_0)^3), \\ T &= T_1(x, z, \omega_H\tau_0)(\omega_H\tau_0)^1 \\ &+ T_2(x, z, \omega_H\tau_0)(\omega_H\tau_0)^2 + O((\omega_H\tau_0)^3).\end{aligned} \tag{22}$$

Functions Φ_j and T_j with $j = 0, 1, \ldots$ satisfy Eqs. (19). The boundary conditions for Φ_0 and T_0 in the planes $x = 0, \beta^{-1}$ are the same as for functions Φ and T, and we have $\Phi_j\big|_{x = 0,\, \beta^{-1}} = 0$, $T_j\big|_{x = 0,\, \beta^{-1}} = 0$ for $j \ge 1$. The boundary conditions for $\Phi_j(x, z, \omega_H\tau_0)$ and $T_j(x, z, \omega_H\tau_0)$ on the planes $z = 1, 0$ were obtained from boundary conditions (20) using perturbation theory with one exception. For Φ_j, we keep the term $(\omega_H\tau_0)\partial\Phi_j/\partial x$ in boundary condition (20) and omit the term $(\omega_H\tau_0)\partial T_j/\partial x$. For T_j, on the contrary, we keep the term $(\omega_H\tau_0)\partial T_j/\partial x$ in boundary condition (20) and omit the term $(\omega_H\tau_0)\partial\Phi_j/\partial x$. The terms $\partial T_{j-1}/\partial x$ and $\partial\Phi_{j-1}/\partial x$ enter the boundary conditions for the respective functions T_j and Φ_j and make them heterogeneous. We then see that the zero-order term T_0 satisfies Eq. (19) and zero boundary conditions in the planes $x = 0, \beta^{-1}$; $z = 0, 1$. Therefore, $T_0 = 0$, which is why we started with the term T_1 in Eq. (22). Functions Φ_j and T_j are analytical in $\omega_H\tau_0$ and can also be expressed in terms of the natural low-field expansion for $\omega_H\tau_0 \ll 1$. Within this approximation, we can obtain the solution of Eq. (19) and, thus, the magnetoresistance. The equations and boundary conditions for the coefficients of Eq. (22) are formulated in what follows. Since the average current in Eqs. (7), (9), and (10) depends on Φ_0, Φ_1, Φ_2, and T_1 and is independent of T_2 with the accuracy up to the $(\omega_H\tau_0)^3$ terms, it is not necessary to calculate it. We then consider the boundary problems for Φ_0, Φ_1, Φ_2, and T_1. Similarly to the above, we obtain the following boundary problem for Φ_0 and T_1:

$$\begin{aligned}&\nabla\Phi_0 = 0, \quad \Phi_0\big|_{x = 0} = \varphi^0(z) + \frac{q+1}{e}T_0, \\ &\Phi_0\big|_{x = \beta^{-1}} = \varphi^a(z) + \frac{q+1}{e}T_0, \\ &\frac{\partial\Phi_0}{\partial z} + \alpha\omega_H\tau_0\frac{\partial\Phi_0}{\partial x}\bigg|_{z = 0, 1} = 0,\end{aligned} \tag{23}$$

$$\begin{aligned}&\nabla^2 T_1 - (q+2)(kb)^2 T_1 = 0, \quad T_1\big|_{x = 0,\, \beta^{-1}} = 0, \\ &\frac{\partial T_1}{\partial z} + \chi\omega_H\tau_0\frac{\partial T_1}{\partial x} + \gamma\frac{\partial\Phi_0}{\partial x}\bigg|_{z = 0, 1} = 0.\end{aligned} \tag{24}$$

With $T_0 = 0$, function Φ_1 satisfies Eq. (19) with zero boundary conditions, and hence, $\Phi_1 = 0$. Function Φ_2 satisfies the boundary problem

$$\begin{aligned}&\nabla^2\Phi_2 = 0, \quad \Phi_2\big|_{x = 0,\, \beta^{-1}} = 0, \\ &\frac{\partial\Phi_2}{\partial z} + \alpha\omega_H\tau_0\frac{\partial\Phi_2}{\partial x} + \frac{\alpha q}{e}\frac{\partial T_1}{\partial x}\bigg|_{z = 0, 1} = 0.\end{aligned} \tag{25}$$

5. THE WEAK-FIELD Φ_0 SOLUTION

To derive the first term of the expansion of (23) for a weak magnetic field, we represent solution Φ_0 with the precision $O(e^{-\pi/\beta})$ as the sum of a regular and a boundary layer function

$$\Phi_0 = \Phi_{reg} + \Pi_0 + \Pi_1 + O(e^{-\pi\beta^{-1}}), \tag{26}$$

where

$$\Phi_{reg} = C_0 + (x - \alpha\omega_H\tau_0 z)C_1 \tag{27}$$

satisfies the boundary condition

$$\frac{\partial\Phi_{reg}}{\partial z} + \alpha\omega_H\tau_0\frac{\partial\Phi_{reg}}{\partial x}\bigg|_{z = 0, 1} = 0;$$

and Π_i (with $i = 0, 1$) are two boundary layer functions that are exact solutions of the problem

$$\nabla^2\Pi_i = 0 \text{ and } \frac{\partial\Pi_i}{\partial z} + \alpha\omega_H\tau_0\frac{\partial\Pi_i}{\partial x}\bigg|_{z = 0, 1} = 0$$

such that Π_0 and Π_i exponentially decrease as $x \longrightarrow \infty$ and $x \longrightarrow -\infty$, respectively. Separating the variables, we can write solutions for the last equations as

$$\begin{aligned}\Pi_0 &= \sqrt{2}\sum_{n=1}^{\infty} A_n(\cos\pi nz + \alpha\omega_H\tau_0\sin\pi nz)e^{-\pi nx}, \\ \Pi_0 &= \sqrt{2}\sum_{n=1}^{\infty} B_n(\cos\pi nz - \alpha\omega_H\tau_0\sin\pi nz)e^{-\pi n(\beta^{-1} - x)}.\end{aligned} \tag{28}$$

As noted above, the boundary layer functions Π_0 and Π_i correspond to the vortex current, and, therefore, do not contribute to the magnetoresistance. Now we will demonstrate this. Note that the average current (11) is x-independent. Therefore, we can calculate it at the point $x = \beta^{-1}/2$. However, the exponentials in the boundary layer functions (28) are less than or equal to $e^{-\pi/2\beta}$ at that point. We also have

$$\int_0^b \cos(\pi n z)dz = 0, \quad n = 1, 2, \ldots .$$

Hence, the boundary layer contributions to average current (7) and to the magnetoresistance have the order $\omega_H\tau_0 e^{-\pi/2\beta}$. We can refine this estimate and demonstrate that this contribution is smaller and has the order $(\omega_H\tau_0)^2 e^{-\pi/2\beta}$. Indeed, it follows from (7) that the contribution of Π_i (with $i = 0, 1$) to the average current with the precision $\omega_H\tau_0 e^{-\pi/2\beta}$ is equal to the integral

$$\sigma_0\int_0^b\left\{-\frac{\partial}{\partial x}\Pi_i(x, z) + \alpha\omega_H\tau_0\frac{\partial}{\partial z}\Pi_i(x, z)\right\}dz\Big|_{x = 1/2\beta}.$$

This is easy to verify for the functions

$$e^{\pm\pi x/b}\left(\cos\frac{\pi n z}{b} \mp \alpha\omega_H\tau_0\sin\frac{\pi n z}{b}\right);$$

in view of the expansions (28) for Π_i, the above integral is zero for all x. The boundary layer contributions to the average current and the magnetoresistance is therefore of the order $(\omega_H\tau_0)^2 e^{-\pi/2\beta}$. Inserting Eqs. (26)–(28) in boundary conditions (23) and neglecting terms of the order $\exp(-\pi\beta^{-1})$, we obtain

$$\begin{aligned}\sqrt{2}\sum_{n=1}^{\infty}A_n(\cos\pi n z + \alpha\omega_H\tau_0\sin\pi n z) &= \varphi^0(z) + \alpha\omega_H\tau_0 z C_1 - C_0,\\ \sqrt{2}\sum_{n=1}^{\infty}B_n(\cos\pi n z + \alpha\omega_H\tau_0\sin\pi n z) &= \varphi^a(z) - (\beta^{-1} - \alpha\omega_H\tau_0 z)C_1 - C_0.\end{aligned} \tag{29}$$

Equation (29) can be solved using the expansion in $\omega_H\tau_0 \ll 1$. A solution in the zero- and first-order approximation for A_n and B_n exists only if both C_0 and C_1, which depend on $\omega_H\tau_0$, satisfy special conditions with respect to the potential distribution at the contacts. We, thus, assume that

$$A_n = A_n^0 + A_n^1\omega_H\tau_0 + \ldots; \quad B_n = B_n^0 + B_n^1\omega_H\tau_0 + \ldots;$$
$$C_0 = C_0^0 + C_0^1\omega_H\tau_0 + \ldots; \quad C_1 = C_1^0 + C_1^1\omega_H\tau_0 + \ldots .$$

Inserting these series in Eq. (29) and keeping the terms of the zero order in $\omega_H\tau_0$, we obtain

$$\sqrt{2}\sum_{n=1}^{\infty}A_n^0\cos\pi n z = \varphi^0(z) - C_0^0;$$
$$\sqrt{2}\sum_{n=1}^{\infty}B_n^0\cos\pi n z = \varphi^1(z) - (C_0^0 + \beta^{-1}C_1^0).$$

It is well known that the system of functions 1, $\sqrt{2}\cos\pi n z$, $n = 1, 2, \ldots$ is complete and orthogonal on the segment [0, 1]. Therefore, every function that is orthogonal to the constant on [0, 1] can be uniquely expanded into a Fourier series with respect to the functions $\sqrt{2}\cos\pi n z$, $n = 1, 2, \ldots$. Hence, to solve the above system for A_n^0 and B_n^0, it is necessary and sufficient that

$$C_0^0 = \int_0^1\varphi^0(z)dz = \bar{\varphi}^0,$$
$$\beta^{-1}C_1^0 + C_0^0 = \int_0^1\varphi^a(z)dz = \bar{\varphi}^a.$$

That is, $C_0^0 = \bar{\varphi}^0$ and $C_1^0 = \beta(\bar{\varphi}^a - \bar{\varphi}^0)$, and therefore,

$$\begin{aligned}A_n^0 &= \sqrt{2}\int_0^1(\varphi^0(z) - \bar{\varphi}^0)\cos(\pi n z)dz;\\ B_n^0 &= \int_0^1(\varphi^a(z) - \bar{\varphi}^a)\cos(\pi n z)dz.\end{aligned} \tag{30}$$

Keeping the first-order terms in the magnetic field $\omega_H\tau_0$ in Eq. (28), we then obtain the equations for coefficients A_n^1 and B_n^1,

$$\begin{aligned}\sqrt{2}\sum_{n=1}^{\infty}A_n^1\cos\pi n z &= -\sqrt{2}\alpha\sum_{n=1}^{\infty}A_n^0\sin\pi n z - C_0^1 + \alpha z C_1^0,\\ \sqrt{2}\sum_{n=1}^{\infty}B_n^1\cos\pi n z &\\ = \sqrt{2}\alpha\sum_{n=1}^{\infty}B_n^1\sin\pi n z &+ \alpha z C_1^0 - (C_0^1 + \beta^{-1}C_1^1).\end{aligned} \tag{31}$$

It follows that system (31) has a solution if and only if the average of its right-hand side on [0, 1] is equal to zero. These conditions give C_0^1, C_1^1 and A_n^1, B_n^1. As can be seen from Eqs. (9)–(11), the magnetoresistance depends only on parameter C_1 given by

$$C_1 = C_1^0 + C_1^1\alpha\omega_H\tau_0 + C_1^2(\alpha\omega_H\tau_0)^2 + O((\alpha\omega_H\tau_0)^3).$$

where

$$\begin{aligned} C_1^0 &= \beta(\bar{\varphi}^a - \bar{\varphi}^0), \\ C_1^1 &= \beta\int_0^1\{\varphi^0(z) + \varphi^a(z) - \bar{\varphi}^0 - \bar{\varphi}^a\}I_1(z)(dz), \\ I_1(z) &= \frac{4}{\pi}\sum_{m=0}^{\infty}\frac{\cos[(2m+1)\pi z]}{(2m+1)\pi}, \\ I_2(z) &= 2\sum_{n=1}^{\infty}\cos\pi nz\int_0^1\sin(\pi n\xi)I_1(\xi)d\xi, \\ C_1^2 &= -\beta\int_0^1\{\varphi^0(z) - \varphi^a(z) - \bar{\varphi}^0 - \bar{\varphi}^a\}I_2(z)dz \\ &\quad - \frac{16\beta^2}{\pi^3}\sum_{m=0}^{\infty}\frac{1}{(2m+1)^3}. \end{aligned} \tag{32}$$

Using Eqs. (26)–(28) and recalling Eqs. (32), we now write the solution as a power series expansion in $\omega_H\tau_0 \ll 1$; i.e.,

$$\begin{aligned} \Phi_0 &= \Phi_0^0 + \Phi_0^1\alpha\omega_H\tau_0 \\ &\quad + \Phi_0^2(\alpha\omega_H\tau_0)^2 + O((\alpha\omega_H\tau_0)^3), \end{aligned} \tag{33}$$

with

$$\Phi_0^0 = \frac{q+1}{e}T_0 + \bar{\varphi}^0 + \beta(\bar{\varphi}^a - \bar{\varphi}^0)x + \Pi_0^0 + \Pi_1^0, \tag{34}$$

where Π_0^0 and Π_1^0 are the zero-order approximations in $\omega_H\tau_0$ of the respective functions Π_0 and Π_1 given by Eqs. (28) and

$$\begin{aligned} \Phi_0^1 &= [C_1^1x + \beta(\bar{\varphi}^a - \bar{\varphi}^0)(z - 1/2)] \\ &\quad + \Pi_0^1 + \Pi_1^1 + cte, \end{aligned} \tag{35}$$

where Π_0^1 and Π_1^1 are the first-order approximations of Π_0 and Π_1. The corresponding solutions are not given here because the magnetoresistance equations do not depend on them. Finally, the coefficient in the second-order approximation of Φ_0 in the magnetic field is written as

$$\Phi_0^2 = C_1^2(x - z + 1/2) + \Pi_0^2 + \Pi_1^2 + cte,$$

where Π_0^2 and Π_1^2 represent the second-order approximations of the functions in Eqs. (28) in the magnetic field; in this case, the magnetoresistance is also independent of them. Using all these approximations in Eqs. (10), we obtain the magnetoresistance given by Eq. (17), which depends linearly on the magnetic field as a consequence of the *z*-dependence of the potential at the contacts. It is important to note that, when the potential distribution at the contacts is constant, the linear term vanishes. In this case, the second-order contribution in the magnetic field must be considered in δ (see Eq. (14))

6. MAGNETORESISTANCE AND THE HOMOGENEOUS POTENTIAL DISTRIBUTION AT THE CONTACTS

Proceeding to the calculation of the coefficient $T_1(x, z, \omega_H\tau_0)$, we begin with the explicit equations that determine this quantity in the approximation of a constant potential at the contacts, i.e., for $\varphi^0(z) = \bar{\varphi}^0$ and $\varphi^a(z) = \bar{\varphi}^a$. As can be seen, Eqs. (24) depend on the magnetic field, and hence, $T_1(x, z, \omega_H\tau_0)$ is also a function of this parameter. It follows from Eqs. (10) that the magnetoresistance depends only on $T_1(x, z, 0)$, which implies that it is only necessary to consider $T_1(x, z, \omega_H\tau_0)$ in the zero-order approximation in the magnetic field in Eqs. (25) and (10). With these approximations, we write the zero-order term of the potential Φ_0^0 instead of Φ_0 in Eqs. (24). We can then write $T_1(x, z, \omega_H\tau_0)$ as a regular term and two boundary layer terms similar to Φ_0 in Eq. (26). In this specific case, it is possible to obtain the exact expression for $T_1(x, z, \omega_H\tau_0)$ if the term $\chi\omega_H\tau_0\partial T_1/\partial x$ is taken into account in the boundary conditions. We can then express $T_1(x, z, \omega_H\tau_0)$ as a series in $\omega_H\tau_0 \ll 1$; however, the only significant term is T_1^0 (the zero-order approximation) that is given by

$$\begin{aligned} T_1^0 &= \frac{\beta\gamma(\bar{\varphi}^a - \bar{\varphi}^0)}{kb\sqrt{q+2}}\frac{\sinh[kb\sqrt{q+2}(z-1/2)]}{\cosh\left[\frac{1}{2}kb\sqrt{q+2}\right]} \\ &\quad + \sum_{n=1}^{\infty}A_n^0\cos\pi nz[\exp\{-\sqrt{\pi^2n^2 + k^2b^2(q+2)}x\} \\ &\quad + \exp\{-\sqrt{\pi^2n^2 + k^2b^2(q+2)}(\beta^{-1} - x)\}], \end{aligned} \tag{36}$$

where

$$A_n^0 = \begin{cases} \dfrac{-2\beta\gamma(\bar{\varphi}^a - \bar{\varphi}^0)}{\pi^2n^2 + k^2b^2(q+2)} & \text{if } n = 2m+1, \\ 0 \text{ if } n = 2m, \end{cases}$$

$$m = 1, 2, 3, \ldots$$

We now derive the second-order approximation in the magnetic field for Φ_2, see Eqs. (25). We set $\Phi_2 = \psi_1 + \psi_2$, where function Ψ_2 satisfies the heterogeneous boundary conditions

$$\frac{\partial\psi_2}{\partial z} + \alpha\omega_H\tau_0\frac{\partial\psi_2}{\partial x} = -\frac{\alpha q}{e}\frac{\partial T_1^0}{\partial x}\bigg|_{z\,=\,0,\,1}$$

and $\Delta\psi_2 = 0$. It is therefore equal to

$$\begin{aligned}\psi_2 = {}& \frac{\alpha q}{e}\sum_{n=1}^{\infty}[\exp\{-\sqrt{\pi^2n^2+k^2b^2(q+2)}x\}\\ &-\exp\{-\sqrt{\pi^2n^2+k^2b^2(q+2)}(\beta^{-1}-x)\}]\\ &\times(D_n^1\sin\{[\pi^2n^2+k^2b^2(q+2)]^{1/2}z\}\\ &+D_n^2\cos\{[\pi^2n^2+k^2b^2(q+2)]^{1/2}z\}),\end{aligned} \tag{37}$$

where

$$D_n^1 = \begin{cases}\dfrac{-2\beta\gamma(\bar{\varphi}^a-\bar{\varphi}^0)}{\pi^2n^2+k^2b^2(q+2)} & \text{if } n = 2m+1,\\ 0 \text{ if } n = 2m,\end{cases}$$

$$m = 1, 2, 3, \ldots,$$

$$D_n^2 = \begin{cases}D_n^1\dfrac{1+\cos[\pi^2n^2+k^2b^2(q+2)]^{1/2}}{\sin[\pi^2n^2+k^2b^2(q+2)]^{1/2}}\\ \text{if } n = 2m+1,\\ 0 \text{ if } n = 2m,\end{cases}$$

$$m = 1, 2, 3\ldots.$$

For ψ_1, we obtain

$$\begin{gathered}\nabla^2\psi_1 = 0, \quad \psi_1\big|_{x\,=\,0,\,\beta^{-1}} = -\psi_2\big|_{x\,=\,0,\,\beta^{-1}},\\ \frac{\partial\psi_1}{\partial z}+\alpha\omega_H\tau_0\frac{\partial\psi_1}{\partial x}\bigg|_{z\,=\,0,\,1} = 0.\end{gathered} \tag{38}$$

The latter system of equations can be solved in the zero-order approximation in the magnetic field similarly to what was done in Section 5. The solution for Φ_2 in the zero-order approximation in the magnetic field is then

$$\begin{aligned}\Phi_2^0 = {}& -8\frac{\alpha q\gamma\beta^2x}{e}(\bar{\varphi}^a-\bar{\varphi}^0)\\ &\times\left\{\sum_{l=0}^{\infty}[\pi^2(2l+1)^2+k^2b^2(q+2)]^{-3/2}\right\}\\ &+\Pi_0^2+\Pi_1^0+cte,\end{aligned}$$

where functions Π_0^0 and Π_1^0 are the decreasing exponential functions of the distance ~1 from the contacts at $x = 0$ and $x = \beta^{-1} \gg 1$. It is important to note that the sum $\Pi_0^0 + \Pi_1^0 + cte$ gives a negligible contribution to the magnetoresistance of the order $e^{-\pi/2\beta}(\omega_H\tau_0)^2$. However, these functions must be considered, otherwise the regular function in Φ_2^0 cannot be calculated. Inserting Φ_2^0 into Eq. (10) and taking Eqs. (21) for α and γ into account, we obtain expression (16).

7. CONCLUSIONS

We have shown that, when the electric potential is inhomogeneous at the contacts, the magnetoresistance exhibits a linear dependence on the magnetic field and it is also possible to mathematically derive the electric potential distribution on the contacts from the experimental measurements of the magnetoresistance. The magnetoresistance changes its sign when the magnetic field is reversed; i.e., the resistance in the sample decreases with the magnetic field before it changes its direction. It is important to note that the sign in Eq. (17) strongly depends on the potential distribution at the contacts and is independent of the length of the sample in the first-order approximation in the magnetic field.

We emphasize that the correct evaluation of the current contacts for the constant potentials at the contacts leads to the effects of the order b/a but not to the exponential terms $e^{-\pi a/b}$ as was expected from the traditional theory of magnetoresistance.

Finally, it is worth mentioning that the solution of the problems in Eqs. (19) and (20) studied in this paper gives a finite total energy for the system under consideration. These problems can also have a nonphysical solution with an infinite total energy.

This work is partially supported by CONACYT, IPN, and CINVESTAV.

REFERENCES

1. A. I. Anselm, *Introduction to Semiconductor Theory* (Nauka, Moscow, 1978; Prentice-Hall, Englewood Cliffs, 1981).
2. O. Madelung, *Physics of III–V Compounds* (Wiley, New York, 1964; Mir, Moscow, 1967).
3. Yu. G. Gurevich and V. V. Kucherenko, Dokl. Akad. Nauk **362**, 181 (1998) [Dokl. Phys. **43**, 534 (1998)].
4. P. P. Vilms, V. S. Sardaryan, P. P. Dobrovolskii, and S. V. Kapylova, Pis'ma Zh. Éksp. Teor. Fiz. **10**, 377 (1969) [JETP Lett. **10**, 240 (1969)].
5. V. S. Bochkov and Yu. G. Gurevich, Fiz. Tverd. Tela (Leningrad) **11**, 714 (1969) [Sov. Phys. Solid State **11**, 570 (1969)].

6. A. I. Klimovskaya, O. V. Snitko, and S. I. Kirilova, Pis'ma Zh. Éksp. Teor. Fiz. **11**, 119 (1970) [JETP Lett. **11**, 73 (1970)].
7. E. I. Rashbaz, Z. S. Gribnikov, and V. Y. Kravchenko, Usp. Fiz. Nauk **119** (1), 3 (1976) [Sov. Phys. Usp. **19**, 361 (1976)].
8. F. G. Bass, V. S. Bochkov, and Yu. G. Gurevich, Fiz. Tekh. Poluprovodn. (Leningrad) **7**, 3 (1973) [Sov. Phys. Semicond. **7**, 1 (1973)].
9. Z. S. Gribnikov, V. I. Melnikov, and T. S. Sorokina, Fiz. Tverd. Tela (Leningrad) **8**, 3379 (1967) [Sov. Phys. Solid State **8**, 2699 (1967)].
10. Z. S. Gribnikov and N. A. Prima, Fiz. Tekh. Poluprovodn. (Leningrad) **5**, 1274 (1971) [Sov. Phys. Semicond. **5**, 1126 (1972)].
11. R. F. Wick, J. Appl. Phys. **25**, 741 (1954).
12. H. Y. Lippman and F. Z. Kahrt, Naturforscher **13**, 462 (1958).
13. G. González de la Cruz, Yu. G. Gurevich, and V. V. Prosentov, Solid State Commun. **97**, 1069 (1996).
14. G. González de la Cruz, Yu. G. Gurevich, and V. V. Prosentov, Phys. Rev. B **56**, 3863 (1997).
15. Yu. G. Gurevich, V. V. Kucherenko, and E. Ramírez de Arellano, Mathematical Notice **65**, 520 (1999).
16. V. F. Gantmakher and Y. B. Levinson, *Carrier Scattering in Metals and Semiconductors* (Nauka, Moscow, 1984; North-Hooland, New York, 1987).
17. Yu. G. Gurevich and M. I. Kaganov, Zh. Éksp. Teor. Fiz. **75**, 2330 (1978) [Sov. Phys. JETP **48**, 1176 (1978)].
18. Yu. G. Gurevich and O. L. Mashkevich, Phys. Rep. **181**, 328 (1989).
19. I. T. Akhiezer, Yu. G. Gurevich, and N. Zakirov, Fiz. Tekh. Poluprovodn. (St. Petersburg) **27**, 628 (1993) [Semiconductors **27**, 349 (1993)].
20. V. V. Kucherenko and N. L. Vasilevsky, Matematical Notice **67**, 539 (2000).

Journal of Experimental and Theoretical Physics, Vol. 92, No. 2, 2001, pp. 287–296.
From Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 119, No. 2, 2001, pp. 331–341.

SOLIDS
Electronic Properties

Quasiparticles in a Strongly Correlated Liquid with the Fermion Condensate: Applications to High-Temperature Superconductors[¶]

S. A. Artamonov and V. R. Shaginyan*
St. Petersburg Nuclear Physics Institute, Russian Academy of Sciences, Gatchina, Leningrad region, 188350 Russia
*e-mail: vrshag@thd.pnpi.spb.ru

Received June 9, 2000

Abstract—A model of a strongly correlated electron liquid based on fermion condensation (FC) is extended to high-temperature superconductors. Within our model, the appearance of FC presents a boundary separating the region of a strongly interacting electron liquid from the region of a strongly correlated electron liquid. We study the superconductivity of a strongly correlated liquid and show that, under certain conditions, the superconductivity vanishes at temperatures $T > T_c \approx T_{node}$, with the superconducting gap being smoothly transformed into a pseudogap. As a result, the pseudogap occupies only a part of the Fermi surface. The gapped area shrinks with increasing the temperature and vanishes at $T = T^*$. The single-particle excitation width is also studied. The quasiparticle dispersion in systems with FC can be represented by two straight lines, characterized by the effective masses M^*_{FC} and M^*_L, intersecting near the binding energy that is on the order of the superconducting gap. It is argued that this strong change of the quasiparticle dispersion upon binding can be enhanced in underdoped samples because of strengthening the FC influence. The FC phase transition in the presence of the superconductivity is examined, and it is shown that this phase transition can be considered as driven by the kinetic energy. *© 2001 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

The unusual properties of the normal state of high-temperature superconductors have been attracting attention for a long time. In describing these properties, which are well beyond the standard Fermi liquid theory, the notion of a strongly correlated liquid has emerged (see, e.g., [1, 2]). Later on, angle-resolved photoemission studies revealed unusual properties observed in underdoped samples, with the leading edge gap discovered up to the temperature $T^* > T_c$. This behavior is interpreted as coming from the pseudogap formation; it was observed in a number of underdoped compounds such as $YBa_2Cu_3O_{6+x}$, $Bi_2Sr_2CaCu_2O_{8+\delta}$, etc. As T increases above T^*, the pseudogap closes, leading to a large Fermi surface and an extremely flat dispersion in electronic spectra, which is called the extended Van Hove singularity [3–7]. A break in the quasiparticle dispersion observed near 50 meV results in a drastic change in the quasiparticle velocity [8–10]. This behavior is definitely different from what one would expect from a normal Fermi liquid.

A correlated liquid can be described in conventional terms, assuming that the correlated regime is related with the noninteracting Fermi gas by adiabatic continuity. This is done in the well-known Landau theory of the normal Fermi liquid, but the question arising at this point is whether this is possible. Most likely, the answer is negative. To tackle the above-mentioned problems, we consider a model where a strongly correlated liquid is separated from the conventional Fermi liquid by a phase transition related to the onset of FC [11, 12]. The purpose of our paper is to show that, without any adjustable parameters, a number of fundamental problems of strongly correlated systems are naturally explained within the model. The paper is organized as follows. In Section 2, we consider the general features of Fermi systems with FC. In Section 3, we show that the pseudogap behavior can be understood within the standard BCS superconductivity mechanism provided the appearance of FC is taken into account. In Section 4, we analyze the condensation energy that is liberated when the system in question undergoes the superconducting phase transition superimposing on the FC phase transition. In Section 5, we describe the quasiparticle dispersion and line shape. Finally, in Section 6, we summarize our main results.

2. THE MAIN FEATURES OF LIQUIDS WITH FC

We first consider the key points of FC theory. FC is related to a new class of solutions of the Fermi liquid theory equation [13]

[¶] This article was submitted by the authors in English.

$$\frac{\delta(F-\mu N)}{\delta n(p,T)} = \varepsilon(p,T)-\mu(T) - T\ln\frac{1-n(p,T)}{n(p,T)} = 0 \qquad (1)$$

for the quasiparticle distribution function $n(p, T)$ depending on momentum p and temperature T. Here F is the free energy, μ is the chemical potential, and $\varepsilon(p, T) = \delta E/\delta n(p, T)$ is the quasiparticle energy, which is a functional of $n(p, T)$ just like energy E and the other thermodynamic functions. Equation (1) is usually represented as the Fermi–Dirac distribution

$$n(p,T) = \left\{1+\exp\left[\frac{(\varepsilon(p,T)-\mu)}{T}\right]\right\}^{-1}. \qquad (2)$$

In a homogeneous matter and at $T = 0$, one obtains from Eq. (2) the standard solution $n_F(p, T = 0) = \theta(p_F - p)$, with $\varepsilon(p \approx p_F) - \mu = p_F(p - p_F)/M^*_L$, where p_F is the Fermi momentum and M^*_L is the commonly used effective mass [13],

$$\frac{1}{M^*_L} = \frac{1}{p}\frac{d\varepsilon(p,T=0)}{dp}\bigg|_{p=p_F}. \qquad (3)$$

It is assumed to be positive and finite at the Fermi momentum p_F. This implies the T-dependent corrections to M^*_L, the quasiparticle energy $\varepsilon(p)$, and the other quantities start with T^2-terms.

However, this solution of Eq. (1) is not the only one possible. There exist “anomalous” solutions of Eq. (1) associated with so-called fermion condensation [11, 14, 15]. Being continuous and satisfying the inequality $0 < n(p) < 1$ within some region in p, such a solution $n(p)$ admits a finite limit for the logarithm in Eq. (1) as $T \longrightarrow 0$, yielding

$$\varepsilon(p) = \frac{\delta E[n(p)]}{\delta n(p)} = \mu, \quad p_i \le p \le p_f. \qquad (4)$$

Equation (4) is used in searching for the minimum value of E as a functional of $n(p)$ under the assumption that a strong rearrangement of the single-particle spectrum can occur. We see from Eq. (4) that the occupation numbers $n(p)$ become variational parameters: the solution $n(p)$ exists if energy E is decreased by alteration of the occupation numbers. Thus, within the region $p_i < p < p_f$, the solution $n(p)$ deviates from the Fermi step function $n_F(p)$ such that the energy $\varepsilon(p, T)$ stays constant, while $n(p)$ coincides with $n_F(p)$ outside this region. As a result, the standard Kohn–Sham scheme for single-particle equations is no longer valid beyond the FC phase transition point [16].

This behavior of systems with FC is clearly different from what one expects from the well known local density calculations; therefore, these calculations are not applicable to systems with FC. On the other hand, the quasiparticle formalism is applicable to this problem, because as we see in what follows, the damping of single-particle excitations is not large compared to their energy [15]. It is also seen from Eq. (4) that a system with FC has a well-defined Fermi surface.

It follows from Eq. (1) that at low T, new solutions within the interval occupied by the fermion condensate have the spectrum $\varepsilon(p, T)$ that is linear in T [15, 17],

$$\varepsilon(p,T)-\mu(T) \approx \frac{(p-p_F)p_F}{M^*_{FC}} \approx T[1-2n(p)] \ll T_f. \qquad (5)$$

Here, T_f is the quasi-FC phase transition temperature above which FC effects become insignificant [15],

$$\frac{T_f}{\varepsilon_F} \sim \frac{p_f^2-p_i^2}{2M\varepsilon_F} \sim \frac{\Omega_{FC}}{\Omega_F}, \qquad (6)$$

where M is the bare electron mass, Ω_{FC} is the condensate volume, ε_F is the Fermi energy, and Ω_F is the volume of the Fermi sphere. One can imagine that the dispersionless plateau $\varepsilon(p) = \mu$ given by Eq. (4) is slightly tilted counter-clockwise about μ and rounded off at the end points. If $T \ll T_f$, it follows from Eqs. (1) and (5) that the effective mass M^*_{FC} related to FC is temperature dependent,

$$\frac{M^*_{FC}}{M} \sim \frac{N(0)}{N_0(0)} \sim \frac{T_f}{T}, \qquad (7)$$

where $N_0(0)$ is the density of states of the noninteracting electron gas, and $N(0)$ is the density of states at the Fermi level. We note that, outside the FC region, the single-particle spectrum is not distinctly affected by temperature, being determined by the effective mass M^*_L given by Eq. (3), which is now evaluated at $p \le p_i$. Thus, we are led to the conclusion that systems with an FC must be characterized by two effective masses: M^*_{FC} related to the single-particle spectrum of a low-energy scale and M^*_L related to the spectrum of a higher energy scale. The existence of these two effective masses can be observed as a break in the quasiparticle dispersion. This break is observed at temperatures $T \ll T_f$, and also when the superconducting state is superimposed on the FC state. In the former case, the occupation numbers over the area occupied by the fermion condensate are slightly disturbed by the pairing correlations such that the effective mass M^*_{FC} becomes large but finite. We remark that at comparatively low temperatures, FC and superconductivity go together because of the remarkable peculiarities of the FC phase transition. This transition is related to a spontaneous gauge symmetry breaking: the superconductivity order parameter

$$\kappa(p) = \sqrt{n(p)[1-n(p)]}$$

has a nonzero value over the region occupied by the fermion condensate, while gap Δ can vanish [15, 16].

It is seen from Eq. (4) that at the FC phase transition point, $p_f \longrightarrow p_i \longrightarrow p_F$, while the effective mass and the density of states tend to infinity as follows from Eqs. (4) and (7). One can conclude that the beginning of the FC phase transition is related to the absolute growth of M^*_{FC}. The onset of the charge density wave instability in an electron system, which occurs as soon as the effective electron–electron interaction constant r_s reaches its critical value r_{cdw}, must be preceded by the unbounded growth of the effective mass [18]. For a simple electron liquid, the effective constant is proportional to the dimensionless average distance $r_s \sim r_0/a_B$ between particles of the system in question, with r_0 being the average distance and a_B the Bohr radius. The physical reason for this growth is the contribution of the virtual charge density fluctuations to the effective mass. The excitation energy of these fluctuations becomes very small if $r_s \approx r_{cdw}$. Thus, an FC can occur when $r_s \sim r_{cdw}$. The standard Fermi liquid behavior can therefore be broken by strong charge fluctuations when the insulator regime is approached in a continuous fashion. We recall that the charge density wave instability occurs in three-dimensional [19] and two-dimensional (2D) electron liquids [20] at a sufficiently high r_s. As soon as r_s reaches its critical value $r_{FC} < r_{cdw}$, the FC phase transition occurs. Thereafter, the condensate volume is proportional to $r_s - r_{FC}$ and also $T_f/\varepsilon_F \sim r_s - r_{FC}$ [15, 18]. In fact, the effective coupling constant r_s increases with decreasing doping. It is assumed that both T_f and condensate volume Ω_{FC} build up with decreasing doping. The FC then serves as a stimulating source of new phase transitions lifting the degeneracy of the spectrum. FC can produce, for instance, the spin density wave (SDW) phase transition or the antiferromagnetic one, thereby promoting a variety of the system properties. We note that the SDW phase transition, the antiferromagnetic transition, and the charge density one also depend on r_s and occur at a sufficiently large value of r_s even if FC is absent. The superconducting phase transition is also aided by FC. We analyze the situation where the superconductivity wins the competition with the other phase transitions up to a temperature T_c. Above the temperature $T^* \ll T_f$, the system under consideration is in its anomalous normal state, Eq. (7) is valid, and one can observe smooth nondispersive segments of the spectra at the Fermi surface [6].

3. SUPERCONDUCTIVITY IN THE PRESENCE OF FC

We focus our attention on investigating the pseudogap that is formed above T_c in underdoped (UD) high-temperature superconductors [4–8]. As we see in what follows, the existence of the pseudogap is closely allied with the presence of FC characterized by a sufficiently high temperature T_c given by Eq. (6). Thus, the pseudogap is peculiar to UD samples, while optimally doped (OP) and overdoped (OD) samples may not exhibit this feature. We consider a 2D liquid on a simple square lattice that has a superconducting state with a d-wave symmetry of the order parameter κ. We assume that the long-range component $V_{lr}(\mathbf{q})$ of the particle-particle interaction $V_{pp}(\mathbf{q})$ is repulsive and has radius q_{lr} in the momentum space such that $p_F/q_{lr} \le 1$. The short-range component $V_{sr}(\mathbf{q})$ is relatively large and attractive, with its radius $p_F/q_{sr} \gg 1$. In agreement with the d-symmetry requirements, the low temperature, gap Δ is then given by the expression [21–23]

$$\Delta(\phi) = 2\kappa E(\phi) \approx \Delta_1 \cos(2\phi) = \Delta_1(x^2 - y^2),$$

where $E(\phi) = \sqrt{\varepsilon^2(\phi) + \Delta^2(\phi)}$ and Δ_1 is the maximal gap. At finite temperatures, the equation for the gap can be written as

$$\Delta(p, \phi) = -\int_0^{2\pi}\int V_{pp}(p, \phi, p_1, \phi_1)\kappa(p_1, \phi_1) \times \tanh\frac{E(p_1, \phi_1)}{2T}\frac{p_1 dp_1 d\phi_1}{4\pi^2}, \tag{8}$$

where p is the absolute value of the momentum and ϕ is the angle. It is also assumed that FC arises near the Van Hove singularities, leading to a large density of states at these points in accordance with Eq. (7). We note that the different FC areas overlap only slightly [17]. $\Delta(\phi)$ obeys the following equation that is determined by the chosen interaction V_{pp},

$$\Delta\left(\frac{\pi}{4} + \phi\right) = -\Delta\left(\frac{\pi}{4} - \phi\right). \tag{9}$$

It vanishes at $\pi/4$ and can therefore be expanded in the Taylor series around $\pi/4$, with $p \approx p_F$:

$$\Delta(p, \theta) = \theta a - \theta^3 b + \ldots, \tag{10}$$

where $\theta = \phi - \pi/4$. Hereafter. we consider solutions of Eq. (8) on the interval $0 < \theta < \pi/4$. We transform Eq. (8) by setting $p \approx p_F$ and separating the contribution I_{lr} coming from V_{lr}, with the contribution related to V_{sr} denoted by I_{sr}. At small angles, I_{lr} can be approximated in accordance with (10) by $I_{lr} = \theta A + \theta^3 B$, with parameters A and B independent of T if $T \le T^* \ll T_f$, because they are defined by the integral over the regions occupied by FC. This theoretical observation is consistent with the experimental results showing that Δ_1 is essentially T-independent at the temperatures $T < T^*$ [6]. The coefficients of the expansion of I_{sr} in powers of θ depend on T. It is therefore more convenient to use the

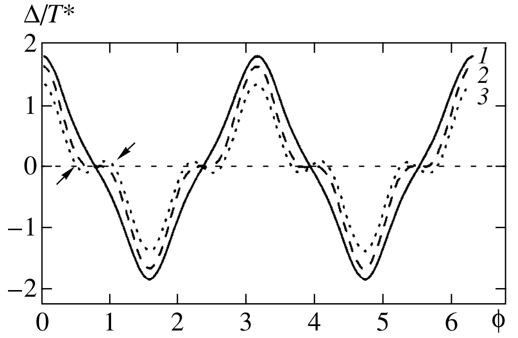


Fig. 1. Gap Δ as a function of φ calculated at three different temperatures expressed in terms of $T_{node} \approx T_c$, while Δ is presented in terms of T^*. Curve *1*, solid line, shows the gap calculated at temperature $0.9T_{node}$. In curve *2*, dashed line, the gap is given at T_{node}. Note the important difference in curve *2* compared with curve *1* due to a flattening of the curve *2* over the region Ω_n. Calculated $\Delta(\phi)$ at $1.2T_{node}$ is shown by curve *3*, dotted line. The arrows indicate the two nodes restricting area Ω_n and emerged at T_{node}.

integral representation for I_{sr} following from Eq. (8). We, thus, have

$$\Delta(\theta) = I_{sr} + I_{lr} = -\int_0^{2\pi}\int V_{sr}(\theta, p_1, \phi_1)\kappa(p_1, \phi_1) \times \tanh\frac{E(p_1, \phi_1)}{2T}\frac{p_1 dp_1 d\phi_1}{4\pi^2} + \theta A + \theta^3 B. \qquad (11)$$

In Eq. (11), variable p was omitted since $p \approx p_F$. It is seen from this equation that FC produces the free term $\theta A + \theta^3 B$. In what follows, we show that at $T \geq T_{node}$, the solution of Eq. (11) has a second node at $\theta_c(T)$ in the vicinity of the first node at $\pi/4$. We also demonstrate that temperature T_{node} has the meaning of temperature T_c at which the superconductivity vanishes. To show this, we simplify Eq. (11) to an algebraic equation. We have $I_{sr} \sim (V_0/T)\theta$ because $\tanh(E/2T) \approx E/2T$ for $E \ll T$ and $T \approx T_{node}$, as is the case in the vicinity of the gap node at $\theta = 0$. The integration in Eq. (11) runs over a small area located at the gap node because of the small radius of V_{sr}. Dividing both parts of Eq. (11) by $\kappa(\theta)$, we obtain

$$E(\theta) = -\left(\frac{V_0}{T} - A_1 - \theta^2 B_1\right)|\theta|, \qquad (12)$$

where A_1 and B_1 are new constants and $V_0 \sim V_{sr}(0)$ is a constant. Imposing the condition that Eq. (8) only has the solution $\Delta \equiv 0$ when $V_{sr} = 0$, we see that A_1 is negative and B_1 is positive. The factor in the brackets on the right-hand side of Eq. (12) changes its sign at some temperature $T_{node} \approx V_0/A_1$; on the other hand, the excitation energy must be $E(\theta) > 0$. Therefore, we have two possibilities [24, 25]. The first follows from the assumption that $\Delta(\theta) \equiv 0$ if θ belongs to the interval Ω_n $[0 < \theta < \theta_c]$. In this case, for $T > T_{node}$ we must solve Eq. (8) with the condition

$$\Delta(\theta) \equiv 0, \quad 0 < \theta < \theta_c, \quad T_{node} < T.$$

This resembles Eq. (4) with the parameter μ being equal to zero. The similarity is not coincidental, because we are searching for new solutions in both cases. Such solutions do exist because the points $\theta = 0$ and $\theta = \theta_c$ represent the branching points of the solution. The second possibility can occur if the above solution does not lead to a minimum value of the free energy. Because the excitation energy must be positive for a stable state, the sign of Δ must be reversed at the point $\theta = \theta_c$. Then gap $\Delta(\theta)$ has the same sign within interval Ω_n and changes its sign once more at the point $\theta = 0$, with $\Delta(\theta_c) = \Delta(0) = 0$. Thus we conclude that gap Δ possesses new nodes at $T > T_{node}$ [25], see Fig. 1. It can be seen from Eq. (12) that angle θ_c is related to $T > T_{node}$ by

$$T \approx \frac{V_0}{A_1 + B_1\theta_c^2}. \qquad (13)$$

It follows from the above consideration and Eq. (12) that even below T_{node}, the order parameter Δ cannot be approximated by a simple *d*-wave form; a more sophisticated expression must be used to fit the flattening of gap Δ around the node. The following expression can be used for this purpose,

$$\Delta(\phi) = \Delta_1[B\cos(2\phi) + (1 - B)\cos(6\phi)]. \qquad (14)$$

Here $0 < B < 1$ in accordance with the experimental results [7] and the term involving $\cos(6\phi)$ is the next compatible with the *d*-symmetry of the gap. It also follows from Eq. (12) that parameter B is a decreasing function of the temperature. At the temperatures $T > T_{node}$, the value of $1 - B$ is sufficiently large to produce new nodes of Δ given by Eq. (14).

As an example of the solutions of Eqs. (8) and (11), we show, in Fig. 1 gap $\Delta(\phi)$ calculated at three different temperatures: $0.9T_{node}$, T_{node}, and $1.2T_{node}$. An important difference between curves *2* and *1* is the flattening of curve *2* at the nodes localized within region Ω_n containing the interval $-\theta_c \leq \theta \leq \theta_c$. As seen from Fig. 1, the flattening occurs as a result of the new nodes restricting area Ω_n. It is also seen from Fig. 1 that gap Δ is extremely small over range Ω_n. It was recently shown in a number of papers (see, e.g., [26, 27]) that there exists an interplay between the magnetism and the superconductivity order parameters, leading to the damping of the magnetism order parameter below T_c. Conversely, one can anticipate the damping of the superconductivity order parameter by magnetism. Thus, we conclude that the gap in range Ω_n can be destroyed by strong antiferromagnetic correlations (or by spin density waves) existing in underdoped superconductors [28, 29]. It is believed that impurities can easily destroy

Δ in the considered area. As a result, one is led to the conclusion that $T_c \approx T_{node}$, with the exact value of T_c defined by the competition between the antiferromagnetic correlations (or spin density waves) and the superconducting correlations over range Ω_n.

We now consider the possibility of two quite different properties, the superconductivity and static spin density wave (SDW), coexisting. We start by briefly outlining the main features of the SDW [30]. A simple example is given by the linear SDW, with the net spin polarization $\mathbf{P}(\mathbf{r})$

$$\mathbf{P}(\mathbf{r}) = P_0 \mathbf{e} \cos(\widehat{Qx}), \tag{15}$$

where $\widehat{Qx}$ is the angle between vectors $\mathbf{Q}$ and $\mathbf{x}$. For convenience, the direction of the SDW is taken along the x-axis, and $\mathbf{e}$ is the unit polarization vector, which in general can have any orientation with respect to $\mathbf{Q}$. In contrast to the superconductivity, the SDW can occupy only a part of the Fermi sphere with the volume $\delta S \approx p_F \delta\phi \delta k$, where $\delta\phi$ is the Fermi surface angle and δk is the "penetration depth" of the SDW into the Fermi sphere. At $T = 0$, the energy gain δW due to the onset of the SDW is given by

$$\delta W \approx g^2 N(0) \delta\phi, \tag{16}$$

where g is the SDW gap determined by the formula [30]

$$g \approx \frac{p_F \delta k}{N(0)} \exp\left(-\frac{4}{N(0)\gamma_0 \delta\phi}\right), \tag{17}$$

where γ_0 is the coupling constant. As seen from Eq. (8), the variation of the gap within some area produces a variation of the gap over the entire occupied area with the same order of magnitude. Therefore, the elimination of Δ over a segment $\delta\phi$ requires the energy $\delta E_1 \sim N(0)\Delta^2(\phi)$. We conclude that at $T < T_{node}$, the destruction of the gap on the interval $\delta\phi$ eliminates Δ over the entire region, because δE_1 is comparable with gain δE due to the superconducting state. A different situation occurs at the temperatures $T > T_{node}$, when Δ is extremely small in Ω_n and the corresponding destruction energy satisfies inequality $\delta E_1 \ll \delta E$. Equations (16) and (17) are very similar to the corresponding BCS equations and this similarity also remains at finite temperatures [30]. Thus, gain δW and gap g vary with the temperature similarly to the superconducting gain δE and gap Δ. We also assume that the SDW transition temperature T_n is sufficiently high, namely, $T_n \geq T_c$. We then come to the conclusion that $\delta E_1 < \delta W$, and region Ω_n is therefore occupied by the SDW at temperatures $T \geq T_{node}$, resulting in the destruction of the superconductivity [24, 25]. We note that the Fermi surface angle $\delta\phi$ must be sufficiently large, because gap g depends exponentially on $\delta\phi$ in accordance with Eq. (17). On the other hand, because we are dealing with an SDW, we have $\delta\phi/\pi \sim 10^{-2}$ [30]. We thus conclude that a strong variation of the superconductivity characteristics may be observed in the vicinity of T_{node}.

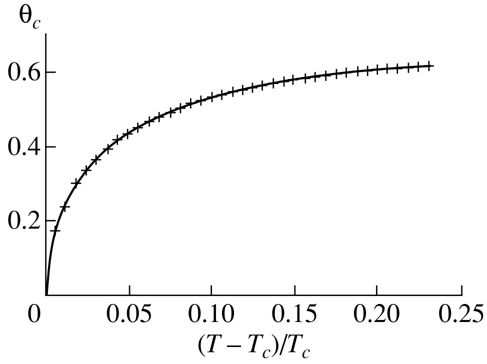


Fig. 2. Calculated angle θ_c, pulling apart the two nodes, as a function of $(T - T_c)/T_c$.

It follows from the above considerations that $\Delta(\theta)$ can be destroyed only locally within region Ω_n because of different reasons. It also follows that T_{node} is the temperature at which the superconductivity vanishes, that is, $T_c \approx T_{node}$. As to the gap at $T > T_c$, or more precisely, the pseudogap, it persists outside the Ω_n region. In accordance with [4, 7], we see that the superconducting gap $\Delta(\theta)$ smoothly transforms into the pseudogap at $T > T_c$. We can therefore expect a dramatic reduction in the difference between the free energy of the normal and the superconducting state at $T = T_c$ (the so-called condensation energy, which we consider in some detail in the next section). It can then be concluded that temperature T^* has the physical meaning of the BCS transition temperature between the state with the order parameter $\kappa \neq 0$ and the normal state. Because $T_c \approx V_0/A_1$, we find from Eq. (13) that $\theta_c \propto \sqrt{(T - T_c)/T_c}$. This result is in harmony with our calculations of the function $\theta_c([T - T_c]/T_c)$ plotted in Fig. 2. Thus, we conclude that the pseudogap "dies out" in UD samples as temperature T^* is approached. Quite naturally, one has to recognize that Δ_1 must scale with T^*.

A few remarks are in order at this point. On the basis of the previous consideration, we conclude that the BCS approach is fruitful in considering OD, OP, and UD samples in the weak coupling regime. With more underdoping, the antiferromagnetic correlations become stronger, breaking down the gap over range Ω_n at lower temperatures. Thus, one observes the decrease of T_c with the decrease of doping. On the other hand, the condensate volume Ω_{FC} becomes larger with the decrease of doping, leading to an increase of gap Δ_1 which is proportional to the volume and interaction V_{pp} [11]. Consequently, temperature T^* becomes higher with decreasing doping. All these results are in agreement with the experimental findings [4, 7]. A peak was observed at 41 meV $\approx 2\Delta_1$ in inelastic neutron scattering from single crystals of the OD, OP, and UD samples $YBa_2Cu_3O_{6+x}$ and $Bi_2Sr_2CaCu_2O_{8+\delta}$ at temperatures below T_c, while a broad maximum above T_c exists in underdoped samples only [31, 32]. The explanation of this peak given in [33] was based on the ideas of BCS theory. From the above discussion, it appears that the same explanation holds for the broad maximum in underdoped samples above T_c because the physics of the process is essentially the same.

4. CONDENSATION ENERGY

We now consider the energy gain or condensation energy E_{cond} liberated when the system in question undergoes the superconducting phase transition involved in the FC phase transition. We set $T = 0$ for simplicity. Energy E_{cond} can be schematically broken into two parts related to the kinetic and the potential energy. The condensation energy was considered in [34], where it was argued that the main contribution to the condensation energy comes from the kinetic energy, i.e., the superconducting phase transition of high-temperature superconductors is kinetic-energy driven. Here, we give a possible interpretation of the situation. It is known [35] that in the superconducting phase transition, the positive contribution comes from the potential energy, while the gain in the kinetic energy is negative. In the other words, the superconducting phase transition is driven by the gain in the potential energy. This result is rather obvious because the ground state energy E_{gs} is given by

$$E_{gs}[\kappa(\mathbf{p})] = E[n(\mathbf{p})] + E_{sc}[\kappa(\mathbf{p})], \tag{18}$$

with the occupation numbers $n(\mathbf{p})$ determined by $\kappa(\mathbf{p}) = \sqrt{n(\mathbf{p})[1-n(\mathbf{p})]}$. The second term $E_{sc}[\kappa(\mathbf{p})]$ on the right-hand side of Eq. (18) is defined by the superconducting contribution, which in the simplest case is of the form

$$[E_{sc}\kappa(\mathbf{p})] = g_2\int V_{pp}(\mathbf{p}_1,\mathbf{p}_2)\kappa(\mathbf{p}_1)\kappa(\mathbf{p}_2)\frac{d\mathbf{p}_1 d\mathbf{p}_2}{(2\pi)^4}. \tag{19}$$

The first term $E[n(\mathbf{p})]$ can be taken as

$$\begin{aligned} E[n(\mathbf{p})] &= \int \frac{p^2}{2M} n(\mathbf{p})\frac{d\mathbf{p}}{4\pi^2} \\ &+ \frac{g_1}{2}\int V(\mathbf{p}_1,\mathbf{p}_2)n(\mathbf{p}_1)n(\mathbf{p}_2)\frac{d\mathbf{p}_1 d\mathbf{p}_2}{(2\pi)^4}, \end{aligned} \tag{20}$$

with the second integral playing the role of the exchange-correlation contribution to the ground state energy. If the effective mass M_L^* given by Eq. (3) is positive and finite, $E[n(\mathbf{p})]$ reaches its minimum at $n(p) = n_F(p)$ and increases with the deviation of $n(p)$ from the Fermi distribution, as it occurs in the presence of superconducting correlations. Thus, the standard situation is that the superconducting phase transition is driven by a decrease of the potential energy [35]. The situation can be different if the system undergoes the FC phase transition. To see this we temporarily assume that $g_2 \longrightarrow 0$ and rewrite Eq. (20) as

$$\begin{aligned} E[n(\mathbf{p})] &= \int \varepsilon(\mathbf{p})n(\mathbf{p})\frac{d\mathbf{p}}{4\pi^2} \\ &- \frac{g_1}{2}\int V(\mathbf{p}_1,\mathbf{p}_2)n(\mathbf{p}_1)n(\mathbf{p}_2)\frac{d\mathbf{p}_1 d\mathbf{p}_2}{(2\pi)^4}, \end{aligned} \tag{21}$$

with the single particle energy

$$\varepsilon(\mathbf{p}) = \frac{\delta E[n(\mathbf{p})]}{\delta n(\mathbf{p})}. \tag{22}$$

The energy $E[n(\mathbf{p})]$ can be lowered by the alteration of $n(\mathbf{p})$ if Eq. (4) has solutions. As a result, we can write the inequality [11]

$$\begin{aligned} E_{cond} &= E_N - E_{FC} \\ &\ge \int [\varepsilon(\mathbf{p}) - \mu]\delta n(\mathbf{p})\frac{d\mathbf{p}}{4\pi^2} \ge 0, \end{aligned} \tag{23}$$

with E_N being the energy of the system in its normal state, E_{FC} the energy with FC, and the integral taken over the region occupied by FC. The chemical potential μ preserves the conservation of the particle number under the variation $\delta n(\mathbf{p})$. We assume that the kinetic energy is given by the first term on the right-hand side of Eq. (21). It then follows from Eq. (23) that the kinetic energy can be lowered, and this lowering is driven by the FC phase transition. It is instructive to illustrate this by a simple example. We take $V(\mathbf{p}_1, \mathbf{p}_2) = g_1\delta(\mathbf{p}_1 - \mathbf{p}_2)$, then E_{cond} given by Eq. (23) becomes

$$\begin{aligned} E_{cond} &= \int [\varepsilon_0(\mathbf{p})n_F(p) - \varepsilon(\mathbf{p})n(p)]\frac{d\mathbf{p}}{4\pi^2} \\ &+ \frac{g_1}{2}\int [n^2(p) - n_F^2(p)]\frac{d\mathbf{p}}{4\pi^2}, \end{aligned} \tag{24}$$

with $\varepsilon_0(\mathbf{p})$ being the single particle energy of the normal ground state. It is easily verified that the second term on the right-hand side of Eq. (24), which is related to the potential energy gain, is negative. This term can be written as

$$\begin{aligned} &\frac{g_1}{2}\int [n^2(p) - n_F^2(p)]\frac{d\mathbf{p}}{4\pi^2} \\ &= \frac{g_1}{2}\int [n(p) - n_F(p)][n(p) + n_F(p)]\frac{d\mathbf{p}}{4\pi^2}. \end{aligned}$$

Observing that

$$\int [n(p) - n_F(p)]\frac{d\mathbf{p}}{4\pi^2} = 0$$

because of the particle number conservation and taking into account that

$$[n(p) + n_F(p)]_{p \le p_F} > [n(p) + n_F(p)]_{p_F \le p},$$

we arrive at the conclusion. The first term is positive because of inequality (23). Thus, we are led to the conclusion that the FC phase transition can be considered as driven by the kinetic energy. We now let the coupling constant g_2 be small, then gap Δ is proportional to g_2 [11]. The optimum values of the occupation numbers given by Eq. (4) are disturbed, leading to an increase of the energy $E[n(\mathbf{p})]$. The positive gain in the potential energy given by Eq. (19) is driving the formation of the superconducting ground state. Because the coupling constant g_2 is sufficiently small, the structure of the system ground state is defined by the FC, and the superconducting state is a "shadow" of the FC under these conditions [15]. Then, the main contribution to E_{cond} comes from the FC phase transition, and the complex transition (FC plus superconductivity) is kinetic-energy driven [36]. On the other hand, in the case where FC is weak compared to the superconductivity (or is absent), we are dealing with a pure superconducting phase transition, which is obviously potential-energy driven.

5. QUASIPARTICLE DISPERSION AND LINE SHAPE

We now discuss the origin of two effective masses M_L^* and M_{FC}^* occurring in the superconducting state and leading to a nontrivial quasiparticle dispersion and a change of the quasiparticle velocity. As we see in what follows, our results are in reasonably good agreement with the experimentally deduced data [8–10]. For simplicity, we set $T = 0$. Varying E_{gs} given by Eq. (18) with respect to $\alpha_{\mathbf{p}}$, we find

$$\frac{E_{gs}[\alpha_{\mathbf{p}}]}{\delta\alpha_{\mathbf{p}}} = [\varepsilon(\mathbf{p}) - \mu]\tanh(2\alpha_{\mathbf{p}}) - \Delta(\mathbf{p}) = 0, \quad (25)$$

with $n(\mathbf{p}) = \cos^2\alpha_{\mathbf{p}}$, $\kappa(\mathbf{p}) = \sin\alpha_{\mathbf{p}}\cos\alpha_{\mathbf{p}}$, and $\varepsilon(\mathbf{p})$ defined by Eq. (22). As $g_2 \longrightarrow 0$, we have that $\Delta(\mathbf{p}) \longrightarrow 0$, and Eq. (25) becomes

$$[\varepsilon(\mathbf{p} - \mu)]\tanh(2\alpha_{\mathbf{p}}) = 0. \quad (26)$$

Equation (26) requires that

$$\varepsilon(\mathbf{p}) - \mu = 0, \text{ if } \tanh(2\alpha_{\mathbf{p}}) \ne 0 \quad (0 < n(\mathbf{p}) < 1), (27)$$

which leads to the FC solutions defined by Eq. (4) [16, 25]. As soon as the coupling constant g_2 becomes finite but small, such that $g_2/g_1 \ll 1$, the plateau $\varepsilon(\mathbf{p}) - \mu = 0$ is slightly tilted and rounded off at the end points. This implies that

$$\varepsilon(\mathbf{p}) - \mu \sim \Delta_1, \quad (28)$$

which allows us to estimate the effective mass as

$$\frac{M_{FC}^*}{M} \sim \frac{T_f}{\Delta_1}. \quad (29)$$

Outside the condensate area, the quasiparticle dispersion is determined by the effective mass M_L^* given by Eq. (3). We note that calculations in the context of a simple model support the above consideration [15]. In that case, putting $V(\mathbf{p}_1, \mathbf{p}_2) = \delta(\mathbf{p}_1, \mathbf{p}_2)$ and $V_{pp}(\mathbf{p}_1, \mathbf{p}_2) = \delta(\mathbf{p}_1, \mathbf{p}_2)$ in Eqs. (19) and (20) and carrying out direct calculations, we obtain at $T = 0$

$$E_0 = \varepsilon(p_f) - \varepsilon(p_i) \approx \frac{(p_f - p_i)p_F}{M_{FC}^*} \approx 2\Delta_1. \quad (30)$$

On the other hand, at $T \ge T_c$, taking into account that $n(p_i) \approx 1$ and $n(p_f) \approx 0$, we obtain from Eq. (5) with the same accuracy,

$$E_0 \approx \frac{(p_f - p_i)p_F}{M_{FC}^*} \approx 2T. \quad (31)$$

Equations (30) and (31) allow us to estimate the effective mass M_{FC}^* related to the region occupied by FC at temperatures $T \ll T_f$. Outside the region, the effective mass is M_L^*. When Eqs. (28) and (29) are compared with Eqs. (5) and (7), it is apparent that gap Δ_1 plays the role of the effective temperature that defines the slope of the plateau. On the other hand, at $T = T_c$ in OD or OP samples, the gap vanishes and Eqs. (5) and (31) define the quasiparticle dispersion and the effective mass. Taking into account that $\Delta_1 \sim T_c$, we are led to the conclusion that Eqs. (28) and (29) derived at $T = 0$ match Eqs. (5) and (7) at T_c. Thus, Eqs. (28) and (29) are approximately valid over the range $0 \le T \le T_c$. It follows from Eq. (30) that at $T \le T_c$, the quasiparticle dispersion can be presented with two straight lines characterized by the respective effective masses M_{FC}^* and M_L^* and intersecting near the binding energy $E_0 \sim 2\Delta_1$. Equation (31) implies that above T_c, the lines intersect near the binding energy $\sim 2T$. The break separating the faster dispersing high-energy part related to M_L^* from the slower dispersing low-energy part defined by M_{FC}^* is likely to be enhanced in UD samples at least because of the rise of temperature T_f, which grows with the decrease of doping. We recall that in accordance with our assumption, the condensate volume Ω_{FC} and T_f grow with underdoping, see Eq. (6) and Section 3. It we also suggested that FC arises near the Van Hove singularities, while the different FC areas overlap only slightly. Therefore, as one moves from (0, 0) towards $(\pi, 0)$ the ratio M_{FC}^*/M_L^* grows in magnitude, developing into the distinct break. In fact, assuming that temperature T_f depends on angle ϕ along the Fermi surface

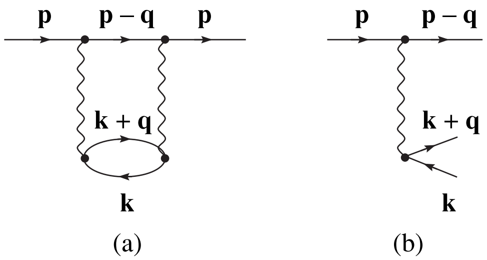


Fig. 3. Diagram (a) depicts a process contributing to the imaginary part. Diagram (b) shows a real process contributing to the imaginary part, observe that quasiparticles **p** – **q**, **k** + **q**, and **k** are on the mass shell.

and taking Eq. (29) into account, one can arrive at the same conclusion. The dispersions above T_c exhibit the same structure except that the effective mass M^*_{FC} is governed by Eq. (31) rather than (30) and both the dispersion and the break are partly "covered" by the quasiparticle width. Thus, one concludes that there also exists a new energy scale at $T \ll T_f$ defined by E_0 and intimately related to T_f [36].

We turn to the quasiparticle excitation with the energy $E(\phi) = \sqrt{\varepsilon^2(\phi) + \Delta^2(\phi)}$. At temperatures $T < T_c$, they are typical excitations of the superconducting state. We now qualitatively analyze the processes contributing to width γ. Within the limits of the analysis, we can take $\Delta \approx 0$, which corresponds to considering excitations at the node. Our treatment is then valid for both $T \le T_c$ and $T_c \le T$. For definiteness, we consider the decay of a particle with the momentum $p > p_F$. Then $\gamma(p, \omega)$ is given by the imaginary part of the diagram shown in Fig. 3a, where the wiggly lines stand for the effective interaction. Because of the unitarity, diagram 3b (which represents real events) can be used to calculate the width [37] as

$$\gamma(p, \omega) = 2\pi\int\left|\frac{V(q)}{\epsilon(q, -\omega_{pq})}\right|^2 \times n(\mathbf{k})[1 - n(\mathbf{k}+\mathbf{q})]\delta(\omega_{pq} + \omega_{kq})\frac{d\mathbf{q}d\mathbf{k}}{(2\pi)^4}, \qquad (32)$$

with $\epsilon(q, -\omega_{pq})$ being the complex dielectric constant and $V(q)/\epsilon$ the effective interaction. Here, $\mathbf{q}$ and $\omega_{kq} = \varepsilon(\mathbf{k}+\mathbf{q}) - \varepsilon(\mathbf{k})$ are the transferred momentum and energy, respectively, and $\omega_{pq} = \omega - \varepsilon(\mathbf{p}-\mathbf{q})$ is the decrease in the quasiparticle energy as the result of the rescattering processes: the quasiparticle with the energy ω decays into a quasihole $\varepsilon(\mathbf{k})$ and two quasiparticles $\varepsilon(\mathbf{p}-\mathbf{q})$ and $\varepsilon(\mathbf{k}+\mathbf{q})$. The transferred momentum $\mathbf{q}$ must satisfy the condition

$$p > |\mathbf{p} - \mathbf{q}| > p_F. \qquad (33)$$

Equation (32) gives the width as a function of p and ω; the width of a quasiparticle with energy $\varepsilon(p)$ is given by $\gamma(p, \omega = \varepsilon(p))$. Estimating the width in Eq. (32) with the constraint (33) and $\omega_{pq} \sim T$, we find that

$$\gamma(p, \omega = \varepsilon(p)) \sim (M_L^*)^3 T^2, \qquad (34)$$

for normal Fermi liquids. In the case of FC one could estimate $\gamma \sim 1/T$ upon using Eqs. (9) and (34). This estimate is correct if the dielectric constant is small, but $\epsilon \sim M^*_{FC}$. As the result, for the FC we have

$$\gamma(p, \omega = \varepsilon(p)) \sim \frac{(M^*_{FC})^3 T^2}{(M^*_{FC})^2} \sim T\frac{T_f}{\varepsilon_F}, \qquad (35)$$

where ε_F is the Fermi energy [38]. Calculating $\gamma(p, \omega)$ as a function of p at constant ω, we obtain the same result for the width given by Eq. (35) when $\omega = \varepsilon(p)$. The calculated function can be fitted with a simple Lorentzian form, because quasiparticles and quasiholes involved in the process are also located in the vicinity of the Fermi level, provided $\omega - \varepsilon_F \sim T$. It then follows from Eq. (35) that the well defined excitations exist at the Fermi surface even in the normal state [38]. This result is in line with the experimental findings determined from the scans at a constant binding energy (momentum distribution curves or MDCs) [8, 39]. On the other hand, considering $\gamma(p, \omega)$ as a function of ω at constant p, we can check that the quasiparticles and quasiholes contributing to the function can have an energy of the same order of the magnitude. For $\omega - \varepsilon_F \sim T$, the function is of the same Lorentzian form, otherwise the shape of the function is disturbed at high ω by high-energy excitations. In that case the special form of the quasiparticle dispersion characterized by the two effective masses must be taken into account. As the result, the lineshape of the quasiparticle peak as a function of the binding energy possesses a complex peak-dip-hump structure [9, 10, 40] directly defined by the existence of the effective masses M^*_{FC} and M^*_L. Our consideration shows that it is the spectral peak obtained from MDCs that provides important information on the existence of well defined excitations at the Fermi level and their width [36]. The detailed numerical results will be presented elsewhere.

At $T > T_c$, the gap is absent in OD or OP samples, and the width γ of excitations close to the Fermi surface is given by Eq. (35). For UD samples, $\Delta(\theta) \equiv 0$ in the range Ω_n and we have normal quasiparticle excitations with width γ. Outside range Ω_n, the Fermi level is occupied by BCS-type excitations with finite excitation energy given by gap $\Delta(\theta)$. Both types of excitations have widths of the same order of magnitude. We now estimate γ. For the entire Fermi level occupied by the normal state, the width is equal to $\gamma \approx N^3(0)T^2/\beta^2$, with the density of states $N(0) \sim 1/T$ and the dielectric constant $\beta \sim N(0)$. Thus, $\gamma \sim T$ [15]. In our case, however, only a part of the Fermi level within Ω_n belongs to the normal excitations. Therefore, the number of states allowed for quasiparticles and for quasiholes is proportional to θ_c, the factor T^2 is therefore replaced by $T^2\theta_c^2$. Taking these factors into account, we obtain $\gamma \sim \theta_c^2 T \sim T(T - T_c)/T_c \sim T - T_c$, because only small angles are considered. Here, we have omitted the small contribution coming from the BCS-type excitations. That is why width γ vanishes at $T = T_c$. Thus, the foregoing analysis shows that in UD samples at $T > T_c$, the superconducting gap smoothly transforms into the pseudogap. The excitations of the gapped area of the Fermi surface have the same width $\gamma \sim T - T_c$ and the region occupied by the pseudogap is shrinking with increasing temperature. These results are in good qualitative agreement with the experimental facts [4–7].

6. CONCLUDING REMARKS

We have discussed the model of a strongly correlated electron liquid based on the FC phase transition and extended it to high-temperature superconductors. The FC transition plays the role of a boundary separating the region of a strongly interacting electron liquid from the region of a strongly correlated electron liquid. On the basis of the BCS theory ideas we have also considered the superconductivity with the *d*-wave symmetry of the order parameter in the presence of FC. We can conclude that the BCS-type approach is fruitful for OD, OP, and UD samples. We have shown that in UD samples, the gap becomes flatter near the nodes at temperatures $T < T_c$, and the superconducting gap smoothly transforms into a pseudogap above T_c. The pseudogap occupies only a part of the Fermi surface, which eventually shrinks with increasing temperature, vanishing at $T = T^*$, and the maximum gap Δ_1 scales with the temperature T^*. We have also shown that the general dependence of T_c, T^*, and Δ_1 on the underdoping level fits naturally into the considered model. At temperatures $T^* > T > T_c$, the single-particle excitations of the gapped area of the Fermi surface have the width $\gamma \sim T - T_c$. The quasiparticle dispersion in systems with FC can be represented by two straight lines characterized by the respective effective masses M^*_{FC} and M^*_L. At $T < T_c$, these lines intersect near the point $E_0 \sim 2\Delta_1$, while above T_c, we have $E_0 \sim 2T$. It is argued that this strong change of the quasiparticle dispersion at E_0 can be enhanced in UD samples because of strengthening the FC influence. The single-particle excitations and their width γ are also studied. We have shown that well-defined excitations with $\gamma \sim T$ exist at the Fermi level even in the normal state. This result is in line with the experimental findings determined from the scans at a constant binding energy, or MDCs. We have also treated the FC phase transition in the presence of the superconductivity and shown that this phase transition can be considered as kinetic-energy driven. Thus, without any adjustable parameters, a number of the fundamental problems of strongly correlated systems are naturally explained within the proposed model.

This research was supported in part by the Russian Foundation for Basic Research under Grant no. 98-02-16170.

REFERENCES

1. Z.-X. Shen and D. S. Dessau, Phys. Rep. **253**, 1 (1995).
2. M. Imada, A. Fujimori, and Y. Tokura, Rev. Mod. Phys. **70**, 1039 (1998).
3. H. Ding, J. C. Campuzano, M. R. Norman, *et al.*, Nature **382**, 51 (1996).
4. H. Ding, J. C. Campuzano, M. R. Norman, *et al.*, cond-mat/9712100 (1997).
5. M. R. Norman, H. Ding, M. Randeria, *et al.*, cond-mat/9710163 (1997).
6. M. R. Norman, M. Randeria, H. Ding, and J. Campuzano, cond-mat/9711232 (1997).
7. J. Mesot, M. R. Norman, H. Ding, *et al.*, cond-mat/9812377 (1998).
8. T. Valla, A. V. Fedorov, P. D. Johnson, *et al.*, Science **285**, 2110 (1999).
9. P. V. Bogdanov, A. Lanzara, S. A. Kellar, *et al.*, Phys. Rev. Lett. **85**, 2581 (2000).
10. A. Kaminski, M. Randeria, J. C. Campuzano, *et al.*, cond-mat/0004482 (2000).
11. V. A. Khodel and V. R. Shaginyan, Pis'ma Zh. Éksp. Teor. Fiz. **51**, 488 (1990) [JETP Lett. **51**, 553 (1990)].
12. G. E. Volovik, Pis'ma Zh. Éksp. Teor. Fiz. **53**, 208 (1991) [JETP Lett. **53**, 222 (1991)].
13. L. D. Landau, Zh. Éksp. Teor. Fiz. **30**, 1058 (1956) [Sov. Phys. JETP **3**, 920 (1956)].
14. V. A. Khodel, V. R. Shaginyan, and V. V. Khodel, Phys. Rep. **249**, 1 (1994).
15. J. Dukelsky, V. A. Khodel, P. Schuck, and V. R. Shaginyan, Z. Phys. **102**, 245 (1997).
16. V. R. Shaginyan, Phys. Lett. A **249**, 237 (1998).
17. V. A. Khodel, J. W. Clark, and V. R. Shaginyan, Solid State Commun. **96**, 353 (1995).
18. V. A. Khodel', V. R. Shaginyan, and M. V. Zverev, Pis'ma Zh. Éksp. Teor. Fiz. **65**, 242 (1997) [JETP Lett. **65**, 253 (1997)].
19. M. Levy and J. P. Perdew, Phys. Rev. B **48**, 11638 (1993).
20. L. Swierkowski, D. Neilson, and J. Szymański, Phys. Rev. Lett. **67**, 240 (1991).
21. D.J. Scalapino, E. Loh, Jr., and J. E. Hirsch, Phys. Rev. B **34**, 8190 (1986).
22. D. J. Scalapino, Phys. Rep. **250**, 329 (1995).
23. A. A. Abrikosov, Physica C (Amsterdam) **222**, 191 (1994); Phys. Rev. B **52**, R15738 (1995); cond-mat/9912394 (1999).
24. M. Ya. Amusia and V. R. Shaginyan, Phys. Lett. A **259**, 460 (1999).
25. V. R. Shaginyan, Pis'ma Zh. Éksp. Teor. Fiz. **68**, 491 (1998) [JETP Lett. **68**, 527 (1998)].
26. N. Metoki, Y. Haga, Y. Koike, and Y. Onuki, Phys. Rev. Lett. **80**, 5417 (1998).

27. T. Honma, Y. Haga, E. Yamamoto, *et al.*, J. Phys. Soc. Jpn. **68**, 338 (1999).
28. J. Schmalian, D. Pines, and B. Stojkovic, Phys. Rev. Lett. **80**, 3839 (1998).
29. I. A. Privorotsky, Zh. Éksp. Teor. Fiz. **43**, 2255 (1962) [Sov. Phys. JETP **16**, 1593 (1963)].
30. A. W. Overhauser, Phys. Rev. **128**, 1437 (1962).
31. H. F. Fong, P. Bourges, Y. Sidis, *et al.*, Nature **398**, 588 (1999).
32. H. He, Y. Sidis, P. Bourges, *et al.*, cond-mat/0002013 (2000).
33. A. A. Abrikosov, Phys. Rev. B **57**, 8656 (1998).
34. M. R. Norman, M. Randeria, B. Janko, and J. C. Campuzano, cond-mat/9912043 (1999); cond-mat/0003406 (2000).
35. G. V. Chester, Phys. Rev. **103**, 1693 (1956).
36. S. A. Artamonov and V. R. Shaginyan, cond-mat/0006013 (2000).
37. R. N. Ritchie, Phys. Rev. **114**, 644 (1959).
38. V. A. Khodel’, V. R. Shaginyan, and P. Schuck, Pis’ma Zh. Éksp. Teor. Fiz. **63**, 719 (1996) [JETP Lett. **63**, 752 (1996)].
39. T. Valla, A. Fedorov, P. D. Jonson, *et al.*, Phys. Rev. Lett. **85**, 828 (2000).
40. A. Kaminski, J. Mesot, H. Fertwell, *et al.*, Phys. Rev. Lett. **84**, 1788 (2000).

Journal of Experimental and Theoretical Physics, Vol. 92, No. 2, 2001, pp. 297–303.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 119, No. 2, 2001, pp. 342–349.
Original Russian Text Copyright © 2001 by Benemanskaya, Daĭneka, Frank-Kamenetskaya.

SOLIDS
Electronic Properties

The Electronic Properties of the Cs/GaAs(100) Interface and the Formation of Metastable Cs Clusters

G. V. Benemanskaya[a], *, D. V. Daĭneka[a], and G. É. Frank-Kamenetskaya[b]

[a]*Ioffe Physicotechnical Institute, Russian Academy of Sciences, Politekhnicheskaya ul. 26, St. Petersburg, 194021 Russia*
[b]*St. Petersburg State Institute of Technology, St. Petersburg, 198013 Russia*
**e-mail: benem@optun.ioffe.rssi.ru*

Received June 30, 2000

Abstract—The method of threshold photoemission spectroscopy is used to investigate the electronic properties of the ultrafine gallium-enriched Cs/GaAs(100) interface. The rearrangement of the spectrum of surface photoemission as a function of Cs coating, as well as the temperature dependence of the spectrum, enable one to identify two phases of adsorption with strong (Cs-Ga) and weak (Cs-Cs) bonds. In the first phase of adsorption with the coating of approximately 0.3 monolayers, two surface bands are detected which are due to the local interaction of cesium adatoms with gallium dimers. It is found that the transition from the first to the second phase of adsorption occurs with the Cs coating of approximately 0.7 monolayers, which corresponds to the saturation of all dangling bonds of gallium on the gallium-enriched GaAs(100) surface. In the second phase of adsorption with the coating of more than 0.7 monolayers, a number of additional photoemission singularities are observed in the spectra, whose emergence is associated with the formation of metastable Cs formations. Photoemission peaks at 1.9 and 2.17 eV may be associated with the excitation of quasi-two- and/or quasi-three-dimensional Cs clusters, and the peaks at 2.05, 2.4, and 2.78 eV may be associated with the excitation of an interface plasmon and of surface and bulk Cs plasmons, respectively. *© 2001 MAIK “Nauka/Interperiodica”.*

1. INTRODUCTION

Great interest in the investigations of interfaces between metal and GaAs is associated both with the fundamental problems of the physics of interface phenomena and with numerous engineering applications of such systems. Of special significance is the GaAs(100) surface, because most GaAs-based semiconductor devices are developed on this particular face using molecular-beam epitaxy [1]. A large number of reconstructions may be observed on the GaAs(100) surface, which differ from one another by the stoichiometric composition. Heating at temperatures below 450°C leads to the formation of arsenic-enriched surfaces, and heating at higher temperatures results in the formation of gallium-enriched surfaces. The amorphization of a surface, associated with the formation of droplets of free gallium, occurs at a temperature of approximately 650°C [2–4]. Both arsenic- and gallium-enriched GaAs(100) surfaces are dimer-reconstructed. The dimerization results in a considerable reduction of the number of dangling bonds of gallium or arsenic. A GaAs(100) surface, obtained at an annealing temperature of about 560°C, is characterized by a $(4 \times 2)/c(8 \times 2)$ structure with gallium dimers in the top layer [5, 6]. According to the adopted model with a single absent Ga dimer [6], the concentration of gallium atoms in the top layer of the $(4 \times 2)/c(8 \times 2)$ GaAs(100) surface is approximately 3/4 of the monolayer, and a group of three Ga dimers may be identified in a unit cell. Unlike the fairly complete data about the structure of the GaAs surface, the available information about its electronic properties and, especially, about its surface states is rather limited. In spite of intensive theoretical and experimental investigations, the nature of the surface states of GaAs is still a subject for discussions. The investigation of the surface electronic properties involves, on the one hand, detecting localized surface and interface states and ascertaining their nature and, on the other hand, searching for the correlation between the structural and electronic properties. In this case, the adsorption of metals serves as an atomic probe for studying various interactions on the surface.

During adsorption, the dimer structure of the surface is preserved, as a rule, and the adatoms interacting with the dangling bonds of the substrate may take different positions, the so-called adsorption sites. For example, for the process of the adsorption of cesium on a dimer-reconstructed Si(100)2×1 surface, it has been found that the adatoms successively take adsorption sites above the dimers and between series of dimers [7]. In this case, local interactions of adatoms with dangling bonds of silicon of two types are observed, this leading to the formation of characteristic surface bands [8, 9]. Unlike numerous investigations of the electronic structure of interfaces between metal and Si(100)2×1, such investigations for interfaces on GaAs(100) substrates have just been started [10–13].

The results of investigations of (Cs, K, Na)/GaAs(100) systems by photoreflection spectroscopy [14], as well as of (Cs, K)/GaAs(110) systems by characteristic electron loss spectroscopy [15], have demonstrated that the process of interface formation is accompanied both by local interactions of adatoms with dangling bonds and, possibly, by the formation of quasi-two- and quasi-three-dimensional clusters of adsorbed atoms. It has been found that clusters of alkali metals, formed at a low temperature on the GaAs(100) surface, are metastable and disintegrate readily as a result of an insignificant increase in temperature.

It is known that the atoms of alkali metals are adsorbed in layers and feature a high mobility on the surface. At room temperature, only one monolayer (ML) of cesium may be deposited on the surface of metals and semiconductors, because the heat of adsorption after the deposition of one ML is low [16]. This means that the cesium atoms from the second layer have a short lifetime on the surface. Akhter and Venables [17] have demonstrated that, upon the deposition of cesium on the W(110) surface, the lifetime of atoms in the second layer at room temperature does not exceed 2000 s. The fact that atoms have a finite lifetime on the surface explains the difference between the adsorption processes under conditions of step-by-step deposition, when the measurements are performed after depositing a certain amount of alkali metal, and under conditions of dynamic deposition, when the measurements are performed directly in the process of deposition. In particular, Akhter and Venables [17] have found that the structure of a saturating Cs coating on the W(110) surface under conditions of step-by-step deposition differs from that under conditions of dynamic deposition. Note that most of the papers known to us fail to give proper attention to the problem of the stability of adsorption systems consisting of an alkali metal and a semiconductor.

We performed, for the first time ever, detailed photoemission investigations of the gallium-enriched Cs/GaAs(100) interface in the range of submonolayer coatings from 0.2 to 0.9 ML. Thanks to the use of different modes of deposition, metastable Cs coatings exceeding a monolayer were also investigated. The stability of the adsorption system was studied for different coatings and at different temperatures. Studies were made into the variation of ionization energy and the evolution of spectra of surface photoemission as functions of Cs coating. The resultant data point to the presence of two phases of cesium adsorption. In the initial phase of adsorption, local surface Cs bands were found, whose formation was completed on accomplishing a saturating Cs coating. It was found that a considerable rearrangement of the spectrum of surface photoemission occurred in the second phase of adsorption. In so doing, the emergence of photoemission peaks was observed, which may be associated with the excitation of metastable Cs clusters and of interface and surface plasmons. It was found that the Cs/GaAs(100) interface exhibited semiconductor behavior in the entire investigated range of coatings.

2. EXPERIMENTAL PROCEDURE

The measurements were performed *in situ* under conditions of an ultrahigh vacuum of $P \approx 5 \times 10^{-11}$ torr at room temperature. A pure GaAs(100) surface (*n*-type, 0.9×10^{18} cm^{-3}) was obtained after the thermal (at a temperature ~580°C) removal of a thin protective oxide layer which was preformed in a growth chamber. As was mentioned above, this temperature mode leads to the formation of a gallium-enriched dimer-reconstructed surface, as a rule, $(4 \times 2)/c(8 \times 2)$. The high surface quality of this sample and the presence of a dimer series was demonstrated in [10] by atomic-force microscopy.

Atomically pure cesium was deposited onto the sample surface from a standard source. A unique procedure [9, 10] was used to determine the intensity of cesium flux from the source. This made possible the determination, within 10%, of the dose D of cesium deposited onto the GaAs surface. In the step-by-step mode, the cesium source was switched off after deposition, the sample was held in darkness for a period of about 600 s, and photoemission spectra were then recorded. The photoemission current was measured in the range from 10^{-13} to 10^{-8} A. In the dynamic mode, the spectrum was recorded directly in the process of the deposition of cesium.

The electronic properties were investigated by threshold photoemission spectroscopy using *s*- and *p*-polarized excitation [18, 19]. This method is based on the separation of bulk and surface photoemission, as well as on the effect of the threshold amplification of photoemission from surface states.

In the case of *s*-polarization, only the bulk states of the substrate are excited, with the photoemission threshold $h\nu_s$ corresponding to the position of the top of the valence band, i.e., $h\nu_s = \phi$, where ϕ is the ionization energy. Two cases must be treated. In the first case, the width of the region of band bends during emission from weakly alloyed samples exceeds considerably the escape depth of photoelectrons, and threshold $h\nu_s$ is the ionization energy. In the second case, the width of the region of band bends is comparable to the escape depth of photoelectrons. Then, during photoemission from a *p*-type semiconductor, the photoemission threshold for electrons excited at some distance from the surface may differ from that for electrons excited directly at the surface, which may lead to errors in determining the value of ϕ. In our case, for an *n*-type semiconductor, threshold $h\nu_s$ is always the ionization energy.

In the case of *p*-polarization, the surface bands are excited owing to interaction with the normal component of an electric vector of light. If the surface states are located in the forbidden band of the semiconductor, the photoemission thresholds may differ, $h\nu_s > h\nu_p$. In so doing, threshold $h\nu_p$ is defined either by the position of the Fermi level E_F or by the long-wave edge of the surface band.

The method of threshold photoemission spectroscopy is characterized by a high surface sensitivity and optical resolution ($\Delta E \leq 0.02$ eV). Yet another advantage of this method is that it is fundamentally nondestructive. The method enables one to obtain detailed information about the structure of surface bands from the spectra of surface photoemission $I_p(h\nu)/I_s(h\nu)$, where I_p and I_s denote the photoemission current upon excitation by *s*- and *p*-polarized light, respectively; in particular, in the absence of collective excitations, the quantity I_p/I_s is proportional to the density of surface states [19].

3. EXPERIMENTAL RESULTS AND DISCUSSION

3.1. The First Phase of the Adsorption of Cesium

Shown in the top part of Fig. 1 is the variation of the ionization energy ϕ during the adsorption of cesium on the gallium-enriched GaAs(100) surface. One can see that, with a dose $D_{\mathrm{Cs}} \approx 6.1 \times 10^{14}$ atom/cm^2, the ionization energy reaches a minimum of $\phi_{\min} = 1.45 \pm 0.02$ eV. In order to estimate the degree of coating $\theta_{\min}$ corresponding to $\phi_{\min}$, one must know, in addition to the dose, the coefficient of the adhesion of cesium to the given surface. The results of Auger spectroscopy demonstrate that, at room temperature, the adhesion coefficient for cesium is equal to unity at $\theta < 0.5$ ML and decreases by a factor of two–three at $\theta > 0.5$ ML [20]. Note that one monolayer equal to 6.3×10^{14} atom/cm^2 corresponds to the coating at which the concentration of adsorbed atoms is equal to the concentration of atoms on the nonreconstructed GaAs(100) 1×1 surface. The scale of coatings is given in the top part of Fig. 1. The inadequate data on the adhesion coefficient define the error which is ± 0.05 ML for $\theta \leq 0.7$ ML and ± 0.1 ML for $\theta > 0.7$ ML. In view of the dose determined by us and the data on the adhesion coefficient, one can estimate $\theta_{\min} \approx 0.7$ ML. The coating of $\theta_{\min} \approx 0.7$ ML is the saturating coating of θ_{sat}, because it corresponds to the concentration of dangling bonds of gallium on the gallium-enriched GaAs(100) surface [3, 5]. The data given in Fig. 1 were obtained using the step-by-step mode of deposition. According to our estimates, the maximum coating that may be obtained in this mode is approximately 0.9 ML.

The bottom part of Fig. 1 gives the photoemission threshold difference $\Delta = h\nu_s - h\nu_p$ as a function of cesium coating. The photoemission thresholds almost coincide in the case of coatings of $\theta < 0.5$ ML, which points to the absence of surface states in the forbidden band. It is known that pure GaAs(100) surfaces exhibit a high density of surface states whose nature is associated with the defects on the surface [21, 22]. Therefore, the adsorption of cesium results in a reduction of the electron density of surface states in the forbidden band with coatings of about 0.2 ML. At $\theta > 0.5$ ML, a difference between photoemission thresholds shows up, this implying the emergence of surface states in the forbidden band. At $\theta \geq 0.7$ ML, $\Delta \approx 0.1$ eV.

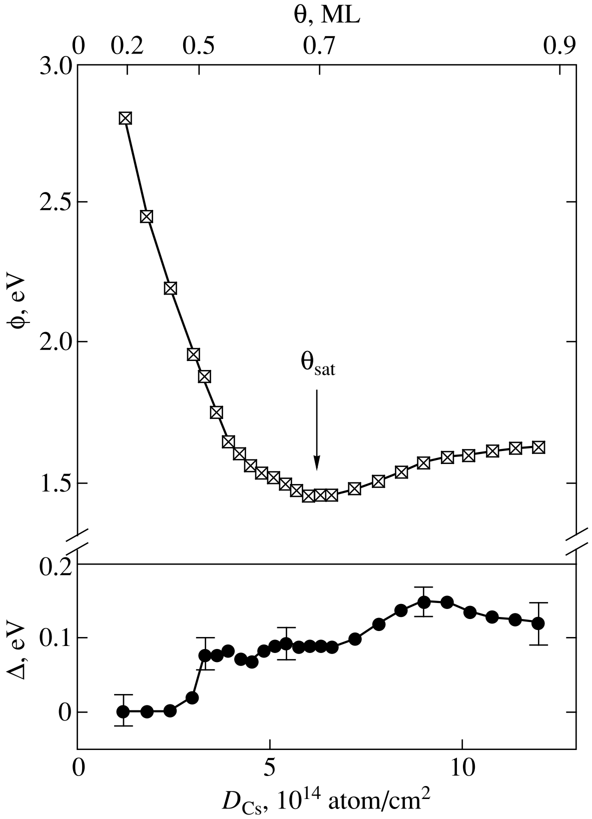


Fig. 1. The variations of the ionization energy (top) and of the photoemission threshold difference $\Delta = h\nu_s - h\nu_p$ (bottom) during the adsorption of cesium on the GaAs(100) surface. Step-by-step mode of deposition.

Figure 2 gives, for different coatings of $\theta < \theta_{\mathrm{sat}}$, the density ρ of surface states induced by the adsorption of cesium on the GaAs(100) surface, which was determined from the spectra of surface photoemission $I_p(h\nu)/I_s(h\nu)$ [19]. Two cesium-induced bands were found, the formation of which began at $\theta \approx 0.5$ ML. At $\theta \approx 0.55$ ML, two bands are clearly distinguishable in the spectra, namely, A_1 and A_2, with their bond energies differing by approximately 0.17 eV. The band intensity increases with the Cs coating; in so doing, the difference between the bond energies decreases (to approximately 0.1 eV for $\theta \approx 0.65$ ML), and the bands shift towards higher bond energies while demonstrating the behavior which is characteristic of bands induced by adsorption [8, 9].

At $\theta_{\mathrm{sat}} \approx 0.7$ ML, bands A_1 and A_2 are not resolved in the spectrum and form a single band A located 0.35 eV below the top of the valence band. Note that, for the case of the Cs/Si(100) 2×1 interface with a saturating coating of $\theta_{\mathrm{sat}} = 1$ ML, both induced bands are well resolved, which indicates that the differences in adsorption sites on Si dimers are retained. In the case of the

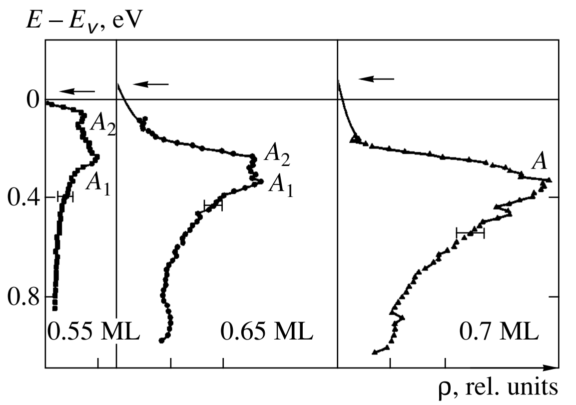


Fig. 2. The spectrum of the density of surface states induced by the adsorption of cesium for different coatings. The energy is reckoned from the top of the valence band E_V; the arrows indicate the thresholds $h\nu_p$ for p-polarized light that correspond to the position of the edge of the surface band.

gallium-enriched GaAs(100) surface, one can assume that with a saturating coating, when all dangling bonds of gallium atoms are filled, the differences between adsorption sites associated with Ga dimers disappear. The difference in the behavior of induced bands on the GaAs and Si surfaces with saturating coating is apparently due to the differences in the behavior of local interactions of dangling bonds of silicon and gallium with adsorbed cesium atoms. One can assume that the charge transfer to the substrate in the case of a Cs-Si bond exceeds that in the case of a Cs-Ga bond. Therefore, the interaction between cesium atoms is weaker on the Si surface than on the GaAs surface. It is possible that even with a saturating coating, conditions are developed for interaction between two cesium adatoms bound on a Ga dimer.

3.2. The Second Phase of Adsorption of Cesium

Figure 3 illustrates the development of photoemission spectra obtained using the step-by-step mode of deposition with coatings in the range 0.7 ML < $\theta \le 0.9$ ML. One can see that the shape of spectra varies cardinally compared with the spectra for $\theta \le 0.7$ ML. In addition to maximum *A*, the emergence of three photoemission peaks was observed, namely, *B*, *C*, and P_1, with energies of approximately 1.9, 2.05, and 2.4 eV, respectively. Figure 4 gives the spectra obtained using the dynamic mode which enables one to produce coatings of $\theta > 0.9$ ML. Even more significant changes in the intensity of photoemission peaks were observed, as well as the emergence of two new singularities, namely, *D* and P_2, with energies of approximately 2.17 and 2.78 eV, respectively.

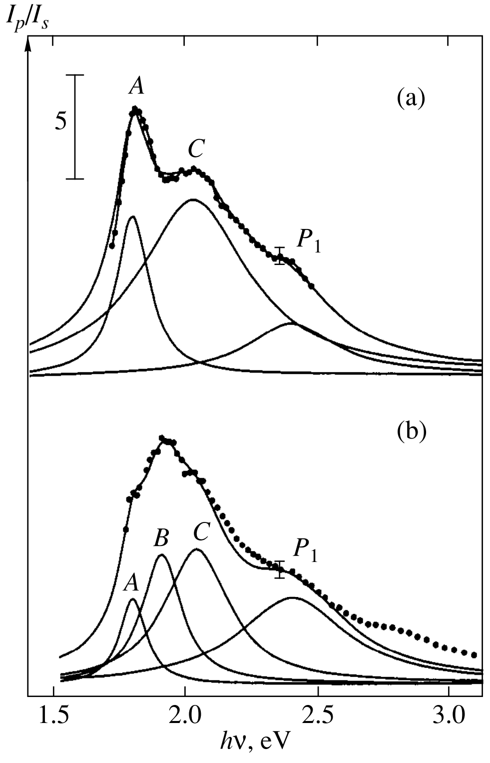


Fig. 3. Spectra of surface photoemission I_p/I_s for Cs coatings of $\theta > \theta_{sat}$: $\theta \approx$ (a) 0.8 and (b) 0.9 ML. Step-by-step mode of deposition.

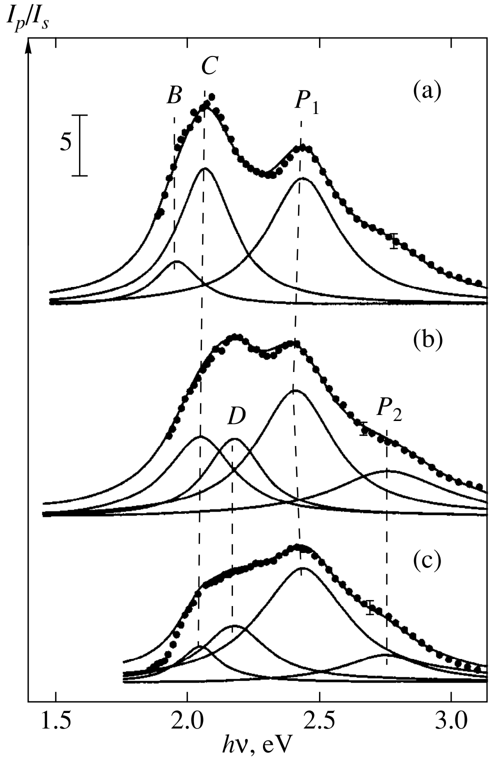


Fig. 4. Spectra of surface photoemission I_p/I_s for different doses of Cs: (a) 1.6×10^{15}, (b) 2.3×10^{15}, and (c) 2.6×10^{15} atom/cm^2. Dynamic mode of deposition.

The stability of the adsorption system of Cs/GaAs(100) for different Cs coatings was investigated using a modification of the spectra of surface photoemission after holding the sample for approximately one hour after deposition, as well as after a short-term heating of the sample. It was found that, for $\theta \leq 0.7$ ML, neither the short term heating at a temperature below 400°C, nor holding the sample in darkness for about one hour, resulted in changes in the spectrum. The results of experiments in thermal desorption [23] demonstrate the absence of the desorption of cesium for the given temperatures and coatings. Temperatures above 500°C are required for the desorption of cesium and, accordingly, for the disappearance of cesium-induced surface states A_1 and A_2. With a saturating coating of $\theta_{sat} \approx 0.7$ ML, the adsorption system is also stable. Even holding the sample in darkness for many hours after deposition fails to result in variations of photoemission spectra.

With $\theta > 0.7$ ML, the adsorption system becomes unstable. Both the intensity and shape of the spectra experience considerable changes after exposure of the sample. A short-term (60 s) and low (at 100°C) heating of the sample results in a spectrum that corresponds to saturating coating. One can conclude that, with coatings of $\theta > \theta_{sat}$, cesium atoms are adsorbed to a weakly bound state. Therefore, after reaching a saturating coating, i.e., when the concentration of cesium atoms almost coincides with that of gallium atoms in the top layer, the behavior of adsorption on the gallium-enriched GaAs(100) surface varies.

In the second phase of adsorption, the predominating process defining the shape of photoemission spectra is that of the generation of metastable cesium formations; in order to clarify the nature of these formations, we analyzed in detail the behavior of photoemission peaks in the spectra obtained at $\theta > 0.7$ ML. All of the photoemission peaks at the moment of emergence have a considerable half-width (0.3 to 0.5 eV), which distinguishes them from the peaks emerging in the first phase of adsorption. Figure 5 gives the data about the variation of the intensity of the peaks as a function of the dose of cesium, which were obtained using two modes of deposition. All peaks may be divided into three groups from the standpoint of their emergence and variation of their intensity.

(1) The maxima of A and B exhibit the same behavior of intensity variation. Both of them emerge in the case of the step-by-step mode of deposition, with the B peak emerging in the spectrum for approximately the coatings in the case of which the A peak disappears.

(2) The maxima of C and P_1 emerge in the step-by-step mode of deposition; however, they both exhibit an insignificant intensity in the case of coatings of $\theta = 0.8$–0.9 ML. Their intensity increases considerably with the dose of cesium in the dynamic mode. This is especially characteristic of the P_1 maximum, whose intensity increases jumpwise at $D_{Cs} \approx 1.5 \times 10^{15}$ atom/cm^2.

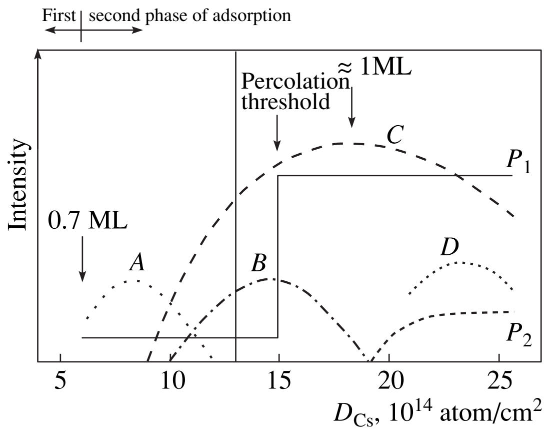


Fig. 5. The scheme of variation of the intensity of photoemission peaks in the second phase of adsorption. To the right of the thin vertical line, the region of the dynamic mode of deposition is located.

(3) The maxima of D and P_2 emerge in the spectrum only in the case of deposition in the dynamic mode.

As was pointed out above, the A peak emerges at $\theta_{sat} \approx 0.7$ ML, when two induced Cs bands A_1 and A_2 in the spectrum become forbidden. The A peak reaches the maximum intensity at $D_{Cs} \approx 8 \times 10^{14}$ atom/cm^2, which corresponds to $\theta \approx 0.8$ ML, and almost disappears at $D_{Cs} \approx 1.2 \times 10^{15}$ atom/cm^2. One can assume that the nature of the A peak is associated with the emergence, in the case of a saturating coating, of two cesium adatoms localized on a Ga dimer. We assume that this is accompanied by the formation of a pair of cesium adatoms, i.e., a minimal Cs cluster. It is to be emphasized that the correlation between such a formation with the Ga dimer is fairly strong, and the adsorption system is stable. The possibility of the existence of such minimal clusters, i.e., local surface clusters, was demonstrated experimentally for silver atoms by Gomoyunova *et al.* [24].

The B peak appears in the spectra of surface photoemission for approximately the same coatings in the case in which the A peak disappears. Simple geometric reasoning enables one to make some assumptions about the subsequent development of Cs clusters. For a model of reconstruction of $(4 \times 2)/c(8 \times 2)$ with one absent Ga dimer [3, 5], one can assume that, in the case of a saturating coating, a quasi-two-dimensional cluster of six (2×3) cesium atoms is formed in each unit cell on the surface. In so doing, the Cs-Cs interaction becomes predominant, and the bond with Ga dimers weakens considerably. We believe that the B peak is associated with local plasma excitations in such clusters. One can expect that the photoemission singularities caused by excitations in these clusters will emerge in the case of a coating of 6/8 ML. Then, when the coating increases to one monolayer, the interaction between clusters must bring about the formation of a "solid" film of cesium on the surface, i.e., to the "destruction" of hexatomic clusters and of the photoemission singularities associated with them. Basically, the *B* peak behaves (see Fig. 5) in accordance with the described model. The quantitative differences may be attributed to inaccuracies in determining the coating, especially, in the dynamic mode, and to the absence of defects on the surface. Hexatomic Cs clusters are metastable, because even a low heating leads to their destruction (disappearance of the *B* peak). In so doing, an inverse transition occurs to three pairs of adatoms of cesium weakly interacting with one another (emergence of the *A* peak).

The P_1 peak with an energy of 2.4 eV is characterized by an abrupt rise of intensity during the transition to the dynamic mode (see Fig. 5). The maximum intensity is attained by a jump at $D_{Cs} \approx 1.2 \times 10^{15}$ atom/cm^2, and, as the dose continues to increase, the peak intensity almost does not vary. As was demonstrated by the results of investigations of the spectra of differential reflection of a multilayer cesium film on the silver surface [25], a surface Cs plasmon is observed in the case of the *p*-polarized excitation of $h\nu \approx 2.4$ eV. Therefore, the nature of the P_1 peak may be associated with the excitation of a surface Cs plasmon. The coating, in the case of which a jumpwise increase in the intensity of this maximum is observed, apparently corresponds to the percolation threshold after which the excitation of a surface plasmon becomes possible.

The P_2 peak (2.78 eV) appears in the spectra only in the dynamic mode at $D_{Cs} > 1.9 \times 10^{15}$ atom/cm^2 (see Figs. 4 and 5). The available literature data [25, 26] enable one to interpret this singularity as the excitation of a bulk plasmon.

Revealing the nature of the *C* and *D* peaks causes the most difficulties. Only general assumptions may be made, based on the behavior of those peaks as the Cs dose increases (see Figs. 4 and 5). In view of the fact that the behavior of the *C* peak is suggestive, by and large, of the behavior of the P_1 peak associated with a surface Cs plasmon, one can assume that the *C* peak is caused by the so-called interface plasmon. Reaching the percolation threshold is not a necessary condition for the excitation of an interface plasmon, because, in this case, the main part is played by the distribution of electron density in the interface layer. The *C* peak reaches the maximum intensity at $D_{Cs} \approx 1.8 \times 10^{15}$ atom/cm^2, which apparently corresponds to a coating close to one monolayer, i.e., to the formation of a “solid” film of cesium. The subsequent adsorption of cesium and formation of the second layer of cesium brings about the reduction of the peak intensity. The photoemission peak *D* (see Figs. 4 and 5) emerges at $D_{Cs} > 2 \times 10^{15}$ atom/cm^2, when, according to our estimates, a monolayer coating has been formed. The nature of this peak may be associated with Cs formations of the type of three-dimensional clusters in the second layer. Because, in the case of an ideal surface and, accordingly, of an ideal first monolayer of cesium, the formation of such clusters is unlikely due to electrostatic repulsion, it is the defects on the surface that appear to be the centers of cluster formation. As the second layer of cesium is formed, these clusters “disappear”, much like the hexatomic clusters in the first layer.

4. CONCLUSION

The results of the investigation of the electronic properties of the Cs/GaAs(100) interface demonstrate the existence of two phases of adsorption, with the transition from the first to second phase occurring with a coating of approximately 0.7 ML, which corresponds to the saturation of all dangling bonds of gallium on the surface. In the first phase, local interactions of adsorbed cesium atoms on adsorption sites of two types are predominant, which lead to the formation of two induced bands. Fundamental changes of the photoemission spectra in the second phase of adsorption are defined by the processes of the formation of metastable quasi-two- and quasi-three-dimensional clusters of cesium on the surface, as well as by the formation, in the dynamic mode, of a monolayer (and probably close to two monolayers) Cs coating. A qualitative model has been suggested for interpreting the observed photoemission singularities, which takes into account the structure of the gallium-enriched GaAs(100) surface and the singularities of the formation of Cs clusters and excitations of an interface plasmon, as well as of surface and bulk Cs plasmons.

ACKNOWLEDGMENTS

We are grateful to V.S. Vikhnin for valuable discussions.

This study received support from the program on Surface Atomic Structures of the Ministry of Science of the Russian Federation (grant no. 99-2.14), from the Russian Foundation for Basic Research, and from the Intergovernmental Russian-Ukrainian program on Nanophysics and Nanoelectronics (grant no. 2000-2D).

REFERENCES

1. *Molecular Beam Epitaxy*, Ed. by A. Cho (American Inst. of Physics Press, New York, 1994).
2. M. Vitomirov, A. D. Raisanen, A. C. Finnefrock, *et al.*, J. Vac. Sci. Technol. B **10**, 1898 (1992).
3. W. Chen, M. Dumas, D. Mao, *et al.*, J. Vac. Sci. Technol. B **10**, 1886 (1992).
4. I. Chizhov, G. Lee, R. F. Willis, *et al.*, Surf. Sci. **419**, 11 (1998).
5. D. K. Biegelsen, R. D. Bringans, J. E. Northrup, *et al.*, Phys. Rev. B **41**, 5701 (1990).
6. D. J. Chadi, J. Vac. Sci. Technol. A **5**, 834 (1987).
7. I. Batra, Phys. Rev. B **43**, 12322 (1991).
8. G. V. Benemanskaya, D. V. Daĭneka, and G. É. Frank-Kamenetskaya, Pis’ma Zh. Éksp. Teor. Fiz. **65**, 699 (1997) [JETP Lett. **65**, 729 (1997)].

9. G. V. Benemanskaya, D. V. Daineka, and G. É. Frank-Kamanetskaya, Surf. Rev. Lett. **5**, 91 (1998).
10. G. V. Benemanskaya, V. P. Evtikhiev, and G. É. Frank-Kamenetskaya, Fiz. Tverd. Tela (St. Petersburg) **42**, 356 (2000) [Phys. Solid State **42**, 366 (2000)].
11. A. O. Gusev, D. Paget, V. Yu. Aristov, *et al.*, J. Vac. Sci. Technol. A **15**, 192 (1997).
12. K. Yamada, J. Asanari, M. Naitoh, *et al.*, Surf. Sci. **402–404**, 683 (1998).
13. O. E. Tereshchenko, A. N. Litvinov, V. L. Al’perovich, *et al.*, Pis’ma Zh. Éksp. Teor. Fiz. **70**, 537 (1999) [JETP Lett. **70**, 550 (1999)].
14. D. Paget, B. Kierren, and R. Houdre, J. Vac. Sci. Technol. A **16**, 2350 (1998).
15. U. del Pennino, R. Compano, B. Salvarani, *et al.*, Surf. Sci. **409**, 258 (1998).
16. A. G. Fedorus and A. G. Naumovets, Surf. Sci. **21**, 426 (1970).
17. P. Akhter and J. A. Venables, Surf. Sci. **103**, 301 (1981).
18. A. Liebsch, G. V. Benemanskaya, and M. N. Lapushkin, Surf. Sci. **302**, 303 (1994).
19. G. V. Benemanskaya, M. N. Lapushkin, and M. I. Urbakh, Zh. Éksp. Teor. Fiz. **102**, 1664 (1992) [Sov. Phys. JETP **75**, 899 (1992)].
20. G. Vergara, L. J. Gómez, J. Campany, *et al.*, Surf. Sci. **278**, 131 (1992).
21. M. D. Pashley, K. W. Haberern, R. M. Feenstra, *et al.*, Phys. Rev. B **48**, 4612 (1993).
22. G. Le Lay, D. Mao, A. Kahn, *et al.*, Phys. Rev. B **43**, 14301 (1991).
23. B. Goldstein and D. Szostak, Appl. Phys. Lett. **26**, 111 (1975).
24. M. V. Gomoyunova, Yu. S. Gordeev, V. M. Mikoushkin, *et al.*, Phys. Low-Dimens. Struct. **4/5**, 11 (1996).
25. A. Liebsch, G. Hincelin, and T. López-Ros, Phys. Rev. B **41**, 10463 (1990).
26. Y.-S. Chao, L. S. O. Johansson, and R. I. G. Uhrberg, Phys. Rev. B **56**, 15446 (1997).

Translated by H. Bronstein

Journal of Experimental and Theoretical Physics, Vol. 92, No. 2, 2001, pp. 304–311.
From Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 119, No. 2, 2001, pp. 350–358.
Original English Text Copyright © 2001 by Pogosov, Kurbatsky.

SOLIDS
Electronic Properties

Density-Functional Theory of Elastically Deformed Finite Metallic System: Work Function and Surface Stress[¶]

V. V. Pogosov[a], * and V. P. Kurbatsky[a, b]

[a]*Department of Microelectronics, Zaporozhye State Technical University, Zaporozhye, 69063 Ukraine*
[b]*Department of Physics, Zaporozhye State Technical University, Zaporozhye, 69063 Ukraine*
**e-mail: vpogosov@zstu.edu.ua*

Received August 3, 2000

Abstract—We study the external strain effect on the surface properties of simple metals within the framework of a modified stabilized jellium model. We derive the equations for the stabilization energy of the deformed Wigner–Seitz cells considered as a function of the bulk electron density and the given deformation. The results for the surface stress and the work function of aluminum calculated using the self-consistent Kohn–Sham method are also given. The problem of the anisotropy of the work function of a finite system is discussed. A clear explanation of independent experiments on the stress-induced contact potential difference at metal surfaces is presented. *© 2001 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

The early experimental investigations of the force acting on electrons and positrons inside a metal tube in the gravitational field of the Earth [1, 2] raised a question about the influence of metal deformation on the electron work function. Direct measurements using the Kelvin method showed a decrease/increase of the contact potential difference (CPD) of the stretched/compressed metal samples [3–5]. Similarly, the experiment with a high-speed spinning metal rotor nonuniformly deformed over the length demonstrated that the CPD changes between areas of the surface subjected to different deformations [6] (see also the discussion of the earlier experiments by Harrison [7]). The influence of the deformation on the electron emission from a thin metal film has also been investigated [8]. Recently, a similar effect on the CPD was observed at the surface of a sample with a nonuniform distribution of the residual mechanical stress [9]. These at first sight surprising results imply the respective increase/decrease of the work function with the uniaxial tension/compression of the metal sample. All these experiments raise two important questions that must be answered by microscopic theory: (i) Does the change of the CPD correspond to a change in the work function? (ii) What is the sign of the deformation gradient of the surface energy and the work function for a metal subject to tension (or compression) along some direction?

The first question is related to the violation of the local electroneutrality of the metal and hence, to nonequipotentiality of its geometric surface. The second question stems from the general statement of elasticity theory: the change in the total energy of a solid is proportional to the square of the relative deformation. Therefore, the energy must increase for compression as it does for tension. On the other hand, it was found experimentally that, in the elastic deformation range, a uniaxial deformation of a metal sample leads to a linear change in the CPD [4, 5]. This implies that classical elasticity theory is not completely correct in determining the elastic characteristics of surfaces. This question is also important in determining the surface tension or the surface stress for macroscopic samples [10] and small metal particles [11].

The measurements of the derivative of the surface tension of a solid with respect to the electrical variable (the so-called "estans" [12]) indirectly show a small difference between the surface stress and the surface energy. On the other hand, different calculations [13–15], including the ones based on the first principles [16], show an appreciable difference between these two quantities. A rough estimation of the difference between the surface energy and the surface stress can also be done using the cohesive energy and the vacancy formation energy. In the continuum approximation, the cohesive energy (or the atomic "work function") ε_{coh} and the vacancy formation energy ε_{vac} give respectively the irreversible and reversible work required for the creation of a new spherical surface of the Wigner–Seitz cell with radius r_0. Following [17], we have

$$\varepsilon_{coh} \approx 4\pi r_0^2 \gamma_0 (1 + \delta/r_0),$$

where γ_0 is the surface energy per unit area of the flat surface and δ/r_0 is the size correction for the surface of a positive curvature.

[¶]This article was submitted by the authors in English.

The reversible work for the creation of a vacancy (which can be defined as the work needed for blowing a small bubble) is given by [18]

$$\varepsilon_{vac} \approx \int_0^{r_0} dr 4\pi r^2 [2\tau_0(1-\delta/2r)/r] = 4\pi r_0^2 \tau_0(1-\delta/r_0),$$

where we introduce a well defined physical quantity—the surface stress of the flat surface τ_0—to describe a tensed curved surface [19, 20]. Combining the expressions for ε_{coh} and ε_{vac}, we obtain

$$\tau_0 \approx \gamma_0 \left(\frac{1+\delta/r_0}{1-\delta/r_0}\right) \frac{\varepsilon_{vac}}{\varepsilon_{coh}}.$$

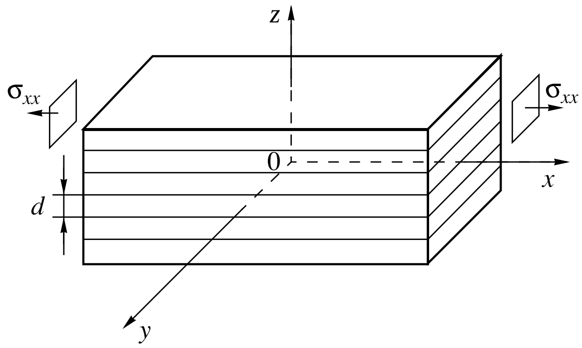
The Kohn–Sham calculations in [21, 22] give $\delta/r_0 \approx$ 0.40 and 0.52 for Na and Al, and the ratio of the experimental values $\varepsilon_{vac}/\varepsilon_{coh}$ is approximately equal to 1/2 and 1/3, respectively. These values agree very well with $\delta/r_0 \approx 1/2$ obtained in [18], which follows from the Langmuir semiempirical rule [23]. From this simple estimation, it follows that τ_0 is approximately equal to or less than γ_0.

In this work, we investigate theoretically the surface energy, stress, and work function of an elastically deformed metal. A uniaxial strain applied to the surface introduces anisotropy to the metal by changing the density (or separation) of the atomic planes and the electron gas concentration and contributes to an extra surface dipole barrier. A rigorous study of this problem from first principles is tedious and requires cumbersome numerical computations. On the other hand, the calculations based on the isotropic models of metal, i.e., on the jellium model [24] (which ignores the discrete nature of ions) or the stabilized jellium model (in which interparticle interactions are averaged over volumes of the spherical Wigner–Seitz cells), do not allow one to properly account for the inhomogeneous strain effects.

We develop a modification of the stabilized jellium model in order to describe the metal deformed by the strain [25–27]. In this modification, the metal energy is expressed as a function of the density parameter r_s and the given deformation. In Section 2, we give a general discussion of the effect of the deformation-induced anisotropy on the work function, which is one of the most important electron surface characteristics. In Section 3, we present equations for the stabilized jellium model accounting for the elastic deformation. In Section 4, the modified stabilized-jellium model is applied to calculate, by the Kohn–Sham method, the effect of the uniaxial strain on the electron surface characteristics of single-crystal aluminum.

2. THE DESCRIPTION OF DEFORMATION

It is important to note that in all experiments we deal with finite samples. Different reticular electron densities at particular faces of a single crystal (crystallite) of an irregular shape lead to different electrostatic potentials for these faces. A similar situation can occur in the deformed metal.



A qualitative sketch of the sample deformation.

We consider a hypothetical crystal having the shape of a rectangular parallelepiped (see figure). We assume the equivalence of all crystal faces in the undeformed state. This picture breaks down upon the crystal deformation. The four side faces remain equivalent to each other, but not to the two base faces. The electroneutrality condition for the metal sample that is stretched or compressed along the x-axis can be written as

$$\int dx \int dy \int dz [n(x,y,z) - \rho(x,y,z)] = 0, \qquad (1)$$

where the electron charge density distribution $n(\mathbf{r})$ attains magnitude n_0 in the metal bulk. The ion charge distribution can be modeled by the step function,

$$\rho(\mathbf{r}) = \bar{\rho}\theta(\mathbf{r}-\mathbf{r}'),$$

where $\mathbf{r}'$ is the radius vector of the surface, $\bar{\rho} = \bar{n}_0/Z$, and Z is the valence. We use atomic units ($e = m = \hbar = 1$) throughout.

By definition [14], with the electrostatic potential set equal to zero in the vacuum, the electron work function for a face of the semi-infinite crystal is

$$W_{face} = -\bar{\phi}_0 - \frac{d}{dn}(\bar{n}_0\varepsilon_J) - \langle \delta v \rangle_{face}, \qquad (2)$$

where $\bar{\phi}_0 < 0$ denotes the electrostatic potential in the metal bulk and $\varepsilon_J \equiv \varepsilon_J(\bar{n}_0)$ is the average energy per electron in the uniform electron gas. The last term represents the difference $\delta v(\mathbf{r})$ between the pseudopotential of the lattice of ions and the electrostatic potential of the positive background averaged over the Wigner–Seitz cell; this term allows us to distinguish between different faces of the crystal (cf. Section 3).

For a deformed sample, we assume that the y- and z-directions are equivalent. Deformation along the x-axis induces an artificial homogeneous anisotropy. The work functions along the x- and z-directions seem to be different for a finite sample, but this conclusion is not correct. This notion is related to the widely spread point of view (see [28] and references therein) that the work function "anisotropy" is determined by the reticular electron density of a given crystal face. However, the electron work function is defined as the difference between the electron energy level in a vacuum and at the Fermi surface. This difference is independent of space directions and coordinates and is constant for a metal sample. The work function (or the ionization potential) is a scalar quantity.

From the standpoint of a finite-size sample, the considerations presented by Smoluchowski [28] and by Lang and Kohn [29] are correct in the case where all faces of the finite sample posses the same atomic packing density. For the cubic crystals, it is a parallelepiped with all its sides having equivalent Miller indices. For a sample of an arbitrary form, the work function in the general case depends on the orientation of all parts of the surface.[1]

We note that the "spurious" difference $W_x - W_y$ of the work functions along the x- and z-directions defined using the standard form (2) vanishes. This leads to an important inequality

$$\bar{\phi}_x - \bar{\phi}_z = -\langle \delta v \rangle_x + \langle \delta v \rangle_z \neq 0 \tag{3}$$

that means that the values $\bar{\phi}_x$ and $\bar{\phi}_z$ of the electrostatic potential in the bulk of the metal can be treated as if they corresponded to different semi-infinite crystals. This inequality does not allow us to unambiguously define the work function of a finite macroscopic sample because the surface electrostatic barrier is different for different directions.

To simplify the analysis, we express the electron profile of the sample as

$$n(\mathbf{r}) = n_0(\mathbf{r}) + \delta n(\mathbf{r}) \tag{4}$$

and

$$\bar{\phi} = \bar{\phi}_0 + \delta\bar{\phi}, \tag{5}$$

where $n_0(\mathbf{r})$ and $\bar{\phi}_0$ are the values corresponding to a semi-infinite metal. The "surplus" density $\delta n(\mathbf{r})$ originates from the electron transfer from one crystal side to another [31] and differs from zero only in the near-surface layer. Condition (1) along each direction then takes the trivial form

$$A_i \int_{-\infty}^{\infty} dx [n_0(\mathbf{r}) - \rho(\mathbf{r})] = 0, \tag{6}$$

where $A_i \equiv A_x, A_y, A_z$ are the areas of faces of a macroscopic sample and $A_y = A_z$.

Taking Eq. (4) into account, Eq. (6) can be written in the "cross-directional" form

$$A_x \int_{-\infty}^{\infty} dx \delta n(\mathbf{r}) + A_y \int_{-\infty}^{\infty} dy \delta n(\mathbf{r}) + A_z \int_{-\infty}^{\infty} dz \delta n(\mathbf{r}) = 0, \tag{7}$$

where the surplus charge at each side is proportional to its area. Here, for simplicity of illustration, we assume that $\delta n(\mathbf{r})$ is constant on each side of the sample. It follows from Eq. (7) that

$$\frac{\int_{-\infty}^{\infty} dz \delta n(\mathbf{r})}{\int_{-\infty}^{\infty} dx \delta n(\mathbf{r})} = -\frac{A_x}{2A_z}, \tag{8}$$

which means that the charges on these sides have opposite signs. The entire sample must be neutral.[2]

The corresponding changes of the electrostatic potential are determined by the Poisson equation, which yields relations for the x- and z-components. These relations have the same form

$$\delta\bar{\phi}_x = -4\pi \int_{-\infty}^{\infty} dx x \delta n(\mathbf{r}) = -C_x x_0, \tag{9}$$

where x_0 and accordingly z_0 are the positions of self-induced charge density at the lateral and base sides and C_x and C_z are constants. This allows us to speak about the appearance of an additional, three-dimensional surface dipole barrier. Since (see Eq. (8))

$$A_x \ll A_z, A_y, \tag{10}$$

we have

$$|C_x/C_z| \propto A_z/A_x$$

for the weight coefficients and

$$|\delta\bar{\phi}_y| = |\delta\bar{\phi}_z| \ll |\delta\bar{\phi}_x|$$

for the additional potentials. Using (5), we can rewrite Eq. (3) as

$$\delta\bar{\phi}_x \approx \langle \delta v \rangle_z - \langle \delta v \rangle_x \quad \text{and} \quad \delta\bar{\phi}_z = \delta\bar{\phi}_y \approx 0. \tag{11}$$

Condition (10) means that the work function is weakly dependent on the electron transfer between the faces perpendicular to y- and z-directions, and the measure-

---

[1] In the special case of a nonzero quadrupole moment of the charge distribution in the elementary cell, the effective potential in the bulk depends on the shape of the sample [30].

[2] We note that the phase shift η_k of the single-particle wave function along each direction depends on the potential shape in the vicinity of the surface and the Sugiyama–Langreth neutrality sum rule [32] must be rewritten with the anisotropy (i.e., the self-charging) taken into account [33].

ment of the work function at these faces can therefore be replaced by the measurement for a semi-infinite metal. The true work function can be measured by the Kelvin method in the areas near the edges. These areas correspond to sign changes of the density, $\delta n(\mathbf{r} \approx 0)$. For the photoemission method of measuring the work function, conditions (10) and (11) imply that the registration of electrons must be performed at distances much greater than the linear dimensions of the sample. Otherwise, if the photon energy is not sufficiently high, an electron escaping from the metal does not reach "infinity" but may transit from one face into the other.

The surplus charge Q_x transferred from one face to the other (see Eq. (9)) can roughly be estimated with the help of the standard electrostatic relation

$$\delta\bar{\phi}_x \approx Q_x/\sqrt{A_x}.$$

Taking into account that

$$A_x \approx N_x 2\pi r_0^2,$$

where N_x is the number of the surface Wigner–Seitz cells of the radius r_0, we obtain

$$Q_x \approx 3r_0\sqrt{N_x}\delta\bar{\phi}_x.$$

The condition $Q_x > 0$ means that Q_x electrons are transfered from the base faces to the lateral ones. The surface energy per unit area therefore changes by $-W_xQ_x/A_x$ and $+W_zQ_x/2A_z$ at the base and the lateral sides, respectively. The ratio of these values corresponds to (7). Here, W_xQ_x is equal to the work needed to remove Q_x electrons from the base side of the metal sample to infinity and W_i is the work function of a given side i. Self-charging of the surface can therefore affect the surface energy anisotropy of the single crystal. For example, for an aluminum sample with $\delta\bar{\phi}_z \approx 0.5$ eV and $N_x = 10^2, 10^4$, the respective electronic charges are $Q_x \approx 1, 10$. It is worth noting that this charge can be very significant for a small crystal (cluster) [34]. Therefore, the elasticity and self-charging effects can play an important role in explaining the recently observed force and conductance fluctuations in stretched metal nanowires [35, 36].

On the ground of the above discussion, and owing to Eq. (11), the properties of a large surface plane of a deformed metal crystal can be calculated in the standard manner.

3. THE MODEL OF A UNIFORMLY DEFORMED METAL

The dependence of the CPD on the uniaxial deformation u_{xx} was measured for polycrystalline stretched samples [4, 5]. We assume that the deformation is a measured quantity and the polycrystal is considered as being assembled from a number of simple crystallites. Qualitatively, the problem can therefore be reduced to the consideration of the tension or compression applied to a single crystal.

We first express the average electron density in the metal as a function of the deformation. For this purpose, we consider an undeformed cubic cell of the side length a_0 and the volume

$$\Omega_0 = a_0^3 = \frac{4}{3}\pi r_0^3, \tag{12}$$

where $r_0 = Z^{1/3}r_s$ is the radius of the spherical Wigner–Seitz cell. For a uniaxially deformed cell elongated or compressed along the x-axis, we can write

$$\Omega = a_x a_y^2 = \frac{4}{3}\pi ab^2, \tag{13}$$

where a_x and $a_y = a_z$ are the sides of the elementary parallelogram and a and b are the half-axes of the equivalent prolate or oblate spheroid of revolution around the x-axis. We also have

$$\begin{aligned} a_x &= a_0(1+u_{xx}) \\ \text{and}\quad a_z &= a_0(1+u_{zz}) = a_0(1-\nu u_{xx}), \end{aligned} \tag{14}$$

where ν is the Poisson coefficient for the polycrystal, and

$$\Omega/\Omega_0 - 1 = u_{xx} + u_{yy} + u_{zz}.$$

It follows from Eqs. (12)–(14) that

$$a = r_0(1+u_{xx}) \text{ and } b = r_0(1-\nu u_{xx}). \tag{15}$$

Similarly, the spacing between the lattice planes perpendicular to the y- or z-direction is

$$d_u = d_0(1-\nu u_{xx}), \tag{16}$$

where d_0 is the interplanar spacing in the undeformed crystal. It then follows from (12)–(15) that the average electron density in the deformed metal is given by

$$\bar{n} = \bar{n}_0\Omega_0/\Omega = \bar{n}_0[1-(1-2\nu)u_{xx}] + O(u_{xx}^2) \tag{17}$$

and the corresponding density parameter is

$$r_{su} = r_s[1+(1-2\nu)u_{xx}]^{1/3}. \tag{18}$$

Proceeding similarly to the derivation of the equations for the original stabilized jellium model [25], we consider a metal assembled from Wigner–Seitz cells. The average energy per valence electron in the bulk is

$$\varepsilon = \varepsilon_J(\bar{n}) + \varepsilon_M + \bar{w}_R, \tag{19}$$

where the first term gives the jellium energy

$$\varepsilon_J(\bar{n}) = \frac{3k_F^2(\bar{n})}{10} - \frac{3}{4\pi}k_F(\bar{n}) + \varepsilon_{cor}(\bar{n}) \tag{20}$$

consisting of the average kinetic and exchange-correlation energy per electron,

$$k_F = (3\pi^2\bar{n})^{1/3}.$$

The remaining two terms in (19) represent the average of the repulsive part of the Ashcroft model potential, the Madelung energy. A small band-structure energy term [25, 37] in (19) is neglected.

By transforming the ordinary jellium into the stabilized one, the Coulomb interactions were averaged over the Wigner–Seitz cells, as is usual for an isotropic medium. The uniaxial strain applied to the crystal deforms the spherical Wigner–Seitz cells into ellipsoidal ones. This affects the Madelung energy ε_M that now must be averaged over the volume of the deformed cell. This energy can be expressed similarly to the gravitational energy of the uniform spheroid [38] as

$$\varepsilon_M(\bar{n}) = \frac{1}{Z}\int\limits_{spheroid} d\Omega\bar{n}\left[-\frac{Z}{r}\right] + \frac{1}{2Z}\int\limits_{spheroid} d\Omega\bar{n}V(r)$$
$$= \begin{cases} -\dfrac{9Z}{10a}\dfrac{1}{2p}\log\dfrac{1+p}{1-p}, & a > b, \\ -\dfrac{9Z}{10a}\dfrac{1}{p}\arctan p, & b > a, \end{cases} \qquad (21)$$

where $V(r)$ is the electrostatic potential inside the uniformly charged spheroid, $p = \sqrt{|1-b^2/a^2|}$ determines the spheroid eccentricity, and the upper/lower case corresponds to a prolate/oblate spheroid, respectively. This expression has the correct limit

$$\varepsilon_M(\bar{n}) \longrightarrow 0.9Z/r_0 \text{ as } u_{xx} \longrightarrow 0.$$

We assume that the shape of ionic cores is not influenced by the deformation and remains spherical; therefore,

$$\bar{w}_R = 2\pi\bar{n}r_c^2.$$

For the potential difference $\delta v(\mathbf{r})$ averaged over the Wigner–Seitz cell [25], we have the same relationship as that for the undisturbed crystal:

$$\langle\delta v\rangle w_S = \tilde{\varepsilon} + \varepsilon_M + \bar{w}_R, \qquad (22)$$

where the electrostatic self-energy of the uniform negative background inside the spheroid is

$$\tilde{\varepsilon} = -\frac{2}{3}\varepsilon_M. \qquad (23)$$

The pseudopotential core radius can be found from the mechanical equilibrium condition depending on the mechanical stress induced in the volume of the cell. To determine the core radius r_c, we note that, for the strained metal, the intrinsic pressure $P = -dE/d\Omega = \bar{n}^2 d\varepsilon/d\bar{n}$ in the bulk of a metal sample is compensated by the pressure exerted by external forces,

$$P = -(\sigma_{xx} + \sigma_{yy} + \sigma_{zz}) = -Yu_{xx}(1-2\nu), \qquad (24)$$

where σ_{ii} are the mechanical stress tensor components and Y is the Young modulus.

For a strained metal, the averaged energy per electron in the bulk is therefore given by

$$\varepsilon = \varepsilon_J(\bar{n}) + \varepsilon_M + \bar{w}_R + P/\bar{n}. \qquad (25)$$

For an ideal metal, $\nu = 1/2$ and $P = 0$. This means that the external force changes not the volume but the shape of a cell or a sample. In the linear approximation, the Madelung energy (21) is well approximated by

$$\varepsilon_M(\bar{n}) \longrightarrow -0.9Z/r_{0u}.$$

Inserting the explicit expressions for (20), (21), and (24) in (25), we have from the minimum condition that

$$r_c = \left\{-\frac{2}{15}\left(\frac{9\pi}{4}\right)^{2/3} r_s + \frac{1}{6\pi}\left(\frac{9\pi}{4}\right)^{1/3} r_s^2 + \frac{1}{5}Z^{2/3}r_s^2 + \frac{2}{9}r_s^4\frac{d\varepsilon_{cor}}{dr_s} + \frac{8}{9}\pi r_s^6 P\right\}^{1/2}_{r_s = r_{su}}, \qquad (26)$$

where r_{su} is the equilibrium density parameter of the strained metal. Here, we assume that the volume of the spheroid is equal to the volume of the equivalent sphere of the radius $r_{0u} = Z^{1/3}r_{su}$.

Taking into account that

$$\langle\delta v\rangle_{WS} = \bar{n}\frac{d}{d\bar{n}}(\varepsilon_M + \bar{w}_R), \qquad (27)$$

we obtain for the strained metal with the equilibrium density

$$\langle\delta v\rangle_{WS} = -\bar{n}\frac{d}{d\bar{n}}\left[\varepsilon_J(\bar{n}) + \frac{P}{\bar{n}}\right]. \qquad (28)$$

Subsequently, similarly to Perdew *et al.* [25], we can introduce the face dependence of the stabilization potential as

$$\langle\delta v\rangle_{face} = \langle\delta v\rangle_{WS} - \left(\frac{\varepsilon_M}{3} + \frac{\pi\bar{n}}{6}d_u^2\right). \qquad (29)$$

The total energy of a finite crystal can be written as the sum of the bulk E^b and the surface E^s energies, where

$$E^s = \gamma_y 4A_y + \gamma_x 2A_x, \qquad (30)$$

with γ_y and γ_x being the respective surface energies per unit area of the lateral and base sides. In the undeformed state, where $\gamma_x = \gamma_y = \gamma_z \equiv \gamma$, surface energy (30) changes by

$$dE^s = 4A_y\left(\gamma\delta_{\alpha\beta} + \frac{d\gamma}{du_{\alpha\beta}}\right)du_{\alpha\beta} + 2A_x\left(\gamma\delta_{\alpha\beta} + \frac{d\gamma}{du_{\alpha\beta}}\right)du_{\alpha\beta}, \qquad (31)$$

Table 1. The calculated surface energies γ, the work function W, the strain derivative $d\gamma/du_{xx}$, and the surface stress τ_{xx}, for elastically deformed Al ($r_s = 2.06$) samples

Metal	Face	γ, erg/cm^2	W, eV	u_{xx}	$d\gamma/du_{xx}$, erg/cm^2	τ, erg/cm^2	ΔW, eV
Al	(111)	946	4.096	(+)	460	1406	–0.032
				(–)	400	1346	+0.033
	(100)	1097	3.780	(+)	833	1930	–0.025
				(–)	810	1907	+0.016

Note: $u_{xx} = \pm 0.03$, positive and negative deformations are labeled with (+) or (–). ΔW is the work function difference. The value of Young's modulus for Al is 70 GPa [39].

where α and β denote directions in the plane of the lateral and base sides and $\delta_{\alpha\beta}$ is the Kronecker symbol. In our model, we calculate only

$$\tau_{xx} = \gamma + \frac{d\gamma}{du_{xx}}. \tag{32}$$

The work function is calculated using the displaced-profile-change-in-a-self-consistent-field (DPΔSCF) expression instead of Eq. (2).

To discuss our results, it is useful to rewrite Eq. (2) as

$$W_{face} = -v_{\rm eff} - \varepsilon_F, \tag{33}$$

where

$$\bar{v}_{eff} = \bar{\phi} + \bar{v}_{xc} + \langle \delta v \rangle_{face}$$

is the effective potential in the bulk giving the total barrier height at the metal–vacuum interface and $\bar{v}_{xc}$ is the exchange-correlation potential in the bulk ($\bar{v}_{xc} = v_{xc}(-\infty)$).

4. RESULTS AND DISCUSSION

To verify the theory presented in Section 3, we solved the Kohn–Sham equations for the two most densely packed surfaces of Al represented by the stabilized jellium model. In terms of our model, we consider two regular single crystals of Al such that all their sides are equivalent in the undeformed state. Under the crystal deformation, the four side faces remain equivalent to each other, but not to the two base faces (see figure). The $\langle\delta v\rangle_{face}$ term included into the effective potential allows us to generate the face-dependent density profiles used in calculating the surface characteristics: work function, surface energy, and surface stress. All calculations were carried out for the upper side of the sample (see figure) assuming the polycrystalline value of the Poisson coefficient $\nu = 0.36$ for the elastic properties of Al [39].

Within the applied range of deformations $-0.03 \le u_{xx} \le +0.03$, the changes in surface quantities remain linear. The positive/negative deformation u_{xx} implies the tension/compression of the side of the sample, i.e., the decrease/increase of the atomic packing density at this side, and the decrease/increase of the mean electron concentration $\bar{n}$ and the interplanar spacing in the direction perpendicular to the chosen crystal side. For a better understanding the crystal effects, we have also performed calculations for the special case of an "ideal" metal with $\nu = 1/2$. In this case, the deformation does not change $\bar{n}$, however, the second term (the corrugation dipole barrier) in the face-dependent potential (29) is changed.

The results of calculations are summarized in Table 1. As can be seen, the surface energy increases linearly with the applied positive deformation u_{xx} and decreases with the negative one. This means that $d\gamma/du_{xx}$ is positive for either $u_{xx} > 0$ or $u_{xx} < 0$. Accordingly, Eq. (32) gives the values of the surface stress component τ_{xx}, larger than the surface energy. For $u_{xx} > 0$, the surface stress is somewhat larger than for $u_{xx} < 0$. We now consider the "ideal" metal with $\nu = 1/2$. It seems that the ideal metal fits better to the classical definition of the surface stress [19, 20]. This is related to the fact that in an ideal metal subjected to deformation, only the surface area is changed, while the electron concentration in the bulk remains unchanged.

Calculations performed for the Al (111) surface yield the respective strain derivatives $d\gamma/du_{xx} = 247$ and 213 erg/cm^2 for $u_{xx} > 0$ and $u_{xx} < 0$. These values are much smaller than the ones reported in Table 1. In this case (with $\nu = 1/2$), we can also evaluate the other components of the surface stress as

$$\tau_{zz} = \tau_{yy} = \gamma + d\gamma/du_{yy}.$$

Inserting

$$du_{zz} = du_{yy} = -\nu du_{xx},$$

we obtain

$$\tau_{zz} = \tau_{yy} = \gamma - 2d\gamma/du_{xx} < \gamma.$$

We can make two observations at this point. First, the latter result agrees with our estimation ($\tau < \gamma$) in Section 1 and with the results derived on the basis of elasticity theory [40], where the τ/γ ratio expressed in terms of the Poisson coefficient ν is given by $(3\nu - 1)/(1 - \nu)$. For $\nu = 1/2$, this formula gives $\tau/\gamma = 1$ and $\nu < 1/2$ for $\tau/\gamma < 1/2$. Second, in order to calculate τ_{zz} and τ_{yy} for a sample stretched along the x-axis, we

Table 2. The calculated change in the effective potential for elastically deformed surfaces of Al single crystal

Metal	Face	u_{xx}	$\Delta v_{eff}\,(z_0, u_{xx})$, eV
Al	(111)	(+)	–0.103
		(–)	+0.106
	(100)	(+)	–0.064
		(–)	+0.069

must use $d\gamma/du_{xx}$ for $u_{xx} < 0$, whereas for a compressed sample, we use the corresponding value for $u_{xx} > 0$. This is because the tension applied along the *x*-direction leads to compressing the sample along the orthogonal (*y* and *z*) axes. The calculated surface stress for Al(111) is in very good agreement with the values resulting from the available *ab initio* calculations: 1441 erg/cm^2 in [15], and 1249 erg/cm^2 in [41]. This also improves the results obtained for the ordinary jellium model [24, 41] and the previous direct application of the stabilized-jellium model [13].

The work function decreases linearly with u_{xx}, but the relative change is less than 1% (see Table l) for the considered strains. A similar behavior is observed for $\nu = 1/2$. The dominating component leading to a decrease of *W* with u_{xx} is a change in the $\langle\delta v\rangle_{face}$ term. Thus, the change of the work function under the deformation conditions is determined by the competition of negative changes in the exchange-correlation (v_{xc}) and the electostatic (ϕ_s) components of the effective potential v_{eff} and the positive change in the face-dependent component $\langle\delta v\rangle_{face}$. A dominant role is played by the change of $\langle\delta v\rangle_{face}$, while the change in the Fermi energy is negligibly small. An overall decrease/increase of the work function *W* is determined by a positive/negative shift of the electrostatic potential in the metal interior.

The calculated change of the work function with strain seems to contradict the experimental results [3–6] where the work function was found to increase/decrease with the elongation/compression of the sample. This conclusion was based on the analysis of the measured CPD [3–7, 9, 27]. In what follows, we demonstrate that this contradiction is spurious. The point is that the measurement by the Kelvin method fixes the change of the surface potential. The experimental observations can therefore be explained not as a change of the work function but as the change of the effective potential v_{eff} upon deformation. The Kelvin method gives the value of the potential difference at the surface of a sample, which can be defined as the position of the image plane $z = z_0$ [26]. In distinction to the work function, to which $\langle\delta v\rangle_{face}$ contributes directly (Eq. (2)), at the image-plane position located outside the geometric surface, the effective potential feels the change in $\langle\delta v\rangle_{face}$ by means of the self-consistent procedure for solving the Kohn–Sham equations (even though $\langle\delta v\rangle_{face}$ is nonzero inside the sample only). The calculations performed for Al(111) demonstrate that the ratio of the effective potential differences Δv_{eff} of the strained ($u_{xx} = \pm 0.03$) and strain-free samples at the surface and in the bulk is

$$\Delta v_{eff}(z = z_0)/\Delta \bar{v}_{eff} \approx -3.$$

Here, v_{eff} denotes the respective difference in the metal bulk.

The results for $\Delta v_{eff}(z_0;\, u_{xx})$ are shown in Table 2. The potential difference outside the sample is more negative as the deformation increases. The calculated changes in the effective potential have the same sign as the CPD measured for Al. For a polycrystalline Al sample subject to deformation with $u_{xx} = 0.03$, the CPD amounts to –0.025 ± 0.002 V [5]. Because a polycrystalline sample can be considered as being assembled from arbitrarily oriented single crystals, the values obtained by us must be averaged in order to compare them with experiment. Thus, both the experiment and the calculations give a negative change of the surface potential,

$$\mathrm{CPD} = \Delta v_{eff}(z = z_0) < 0.$$

For the conventional method of measuring the work function changes upon strain [4, 5, 9], this implies that

$$W(u_{xx}) = W(0) - \mathrm{CPD}(u_{xx}) > W(0),$$

i.e., the work function increases for a tensed sample. In general, therefore, our results agree with the independent experiments for both stretched [4–6] and compressed [1, 3] metal samples. The results for $\Delta v_{eff}(z_0, u_{xx})$ correspond to a direct observation of the stress-induced shift in the measured contact potential: the effective potential outside the open faces of the sample is more negative/positive when tensile/compressive force is applied. However, unlike the effective potential at the surface, the value of the potential in the metal bulk is more positive/negative for an expanded/compressed sample because of the different effect of the $\langle\delta v\rangle_{face}$ term. Thus, for the Al sample, the work function change vs. strain shows the opposite trend compared to that of the contact potential (the behavior of which also differs from that predicted by non-self-consistent calculations [27]). Accordingly, the results in Table 1 demonstrate that the work function decreases with u_{xx}. In other words, our results show that the measurements by the Kelvin method give not the variation of the work function upon strain but the variation of the surface potential.

In summary, the stabilized-jellium model has been extended to encompass the elastic strain effects on the surface properties of simple metals. By imposing a uniaxial strain to the metal surface and limiting ourselves to linear terms in the deformation, we have obtained a realistic description of the strain dependence of surface quantities: surface energy, surface stress, and work function. We have presented a consistent explanation of experiments on the stress-induced contact potential difference at metal surfaces.

The authors are grateful to A. Kiejna for his help in numerical calculations. One of the authors (V.V.P.) would like to thank the Institute of Experimental Physics at the University of Wroclaw for kind hospitality. This work was partly supported by the NATO "Science for Peace" Program (project SfP-974109).

REFERENCES

1. F. C. Witteborn and W. N. Fairbank, Phys. Rev. Lett. **19**, 1049 (1967).
2. Sh. M. Kogan, Usp. Fiz. Nauk **105**, 157 (1972) [Sov. Phys. Usp. **14**, 658 (1972)].
3. P. P. Craig, Phys. Rev. Lett. **22**, 700 (1969).
4. P. I. Mints, V. P. Melekhin, and M. B. Partensky, Fiz. Tverd. Tela (Leningrad) **16**, 3584 (1974) [Sov. Phys. Solid State **16**, 2330 (1974)].
5. S. V. Loskutov, Fiz. Met. Metalloved. **86**, 149 (1998).
6. J. W. Beams, Phys. Rev. Lett. **21**, 1093 (1968).
7. W. A. Harrison, Phys. Rev. **180**, 1606 (1969).
8. Yu. A. Kulyupin and S. A. Nepijko, Fiz. Tverd. Tela (Leningrad) **17**, 2747 (1975) [Sov. Phys. Solid State **17**, 1822 (1975)].
9. V. V. Levitin, S. V. Loskutov, M. I. Pravda, and B. A. Serpetzky, Solid State Commun. **92**, 973 (1994).
10. A. Kiejna and K. F. Wojciechowski, *Metal Surface Electron Physics* (Pergamon, Oxford, 1996).
11. E. L. Nagaev, Usp. Fiz. Nauk **162** (9), 49 (1992) [Sov. Phys. Usp. **35**, 747 (1992)].
12. A. Ya. Gohshtein, Usp. Khim. **44**, 1956 (1975).
13. A. Kiejna and P. Ziesche, Solid State Commun. **88**, 143 (1993).
14. J. P. Perdew, Prog. Surf. Sci. **48**, 245 (1995).
15. P. J. Feibelman, Phys. Rev. B **50**, 1908 (1994).
16. H. L. Skriver and N. M. Rosengaard, Phys. Rev. B **46**, 7157 (1992).
17. J. P. Perdew, Y. Wang, and E. Engel, Phys. Rev. Lett. **66**, 508 (1991).
18. V. V. Pogosov, Solid State Commun. **89**, 1017 (1994).
19. R. Shuttleworth, Proc. Phys. Soc. London, Sect. A **63**, 444 (1950).
20. R. C. Cammarata, Prog. Surf. Sci. **46**, 1 (1994).
21. P. Ziesche, J. P. Perdew, and C. Fiolhais, Phys. Rev. B **49**, 7916 (1994).
22. P. Ziesche, in *Density Functional Theory*, Ed. by E. K. U. Gross and R. M. Dreizler (Plenum, New York, 1995), p. 559.
23. I. Langmuir, Chem. Rev. **13**, 147 (1933).
24. W. A. Tiller, S. Ciraci, and I. P. Batra, Surf. Sci. **65**, 173 (1977).
25. J. P. Perdew, H. Q. Tran, and E. D. Smith, Phys. Rev. B **42**, 11627 (1990).
26. A. Kiejna, Prog. Surf. Sci. **61**, 85 (1999).
27. V. V. Pogosov, Solid State Commun. **81**, 129 (1992).
28. R. Smoluchowski, Phys. Rev. **60**, 661 (1941).
29. N. D. Lang and W. Kohn, Phys. Rev. B **3**, 1215 (1971).
30. I. I. Tupizin and I. V. Abarenkov, Phys. Status Solidi B **82**, 99 (1977).
31. N. W. Ashcroft and N. D. Mermin, *Solid State Physics* (Holt, Rinehart and Winston, New York, 1976).
32. A. Kiejna and P. Ziesche, Phys. Rev. B **56**, 1095 (1997).
33. A. K. Theophilou and A. Modinos, Phys. Rev. B **6**, 81 (1972).
34. A. Kiejna and V. V. Pogosov, J. Phys.: Condens. Matter **8**, 4245 (1996).
35. J. M. Krans, J. M. van Ruitenbeek, V. V. Fisun, *et al.*, Nature **375**, 767 (1995).
36. G. Rubio, N. Agrait, and S. Viera, Phys. Rev. Lett. **76**, 2302 (1996).
37. E. G. Brovman and Yu. Kagan, Usp. Fiz. Nauk **112**, 369 (1974) [Sov. Phys. Usp. **17**, 125 (1974)].
38. L. D. Landau and E. M. Lifshitz, *The Classical Theory of Fields* (Nauka, Moscow, 1973; Pergamon, Oxford, 1975).
39. L. V. Tikhonov, V. A. Kononenko, G. I. Prokopenko, and V. A. Rafalovsky, *Structure and Properties of Metals and Alloys* (Naukova Dumka, Kiev, 1986).
40. R. G. Linford, L. A. Mitchel, C. Osgood, and M. P. Williams, Surf. Sci. **219**, 341 (1989).
41. R. J. Needs and M. J. Godfrey, Phys. Rev. B **42**, 10933 (1990).

Journal of Experimental and Theoretical Physics, Vol. 92, No. 2, 2001, pp. 312–325.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 119, No. 2, 2001, pp. 359–373.

SOLIDS
Electronic Properties

Ground State Formation in a Strong Hubbard Correlation Regime in Iron Monosilicide

N. E. Sluchanko[a, *], V. V. Glushkov[a], S. V. Demishev[a], M. V. Kondrin[a], V. Yu. Ivanov[a], K. M. Petukhov[a], N. A. Samarin[a], A. A. Menovsky[b], and V. V. Moshchalkov[c]
[a]*Institute of General Physics, Russian Academy of Sciences, Moscow, 117492 Russia*
**e-mail: nes@lt.gpi.ru*
[b]*Van der Waals–Zeeman Laboratory, University of Amsterdam, 1018 XE Amsterdam, Netherlands*
[c]*Laboratory voor Vaste–Stoffysica en Magnetisme, Kuleuven B-3001 Leuven, Belgium*
Received July 4, 2000

Abstract—Low-temperature anomalies in the physical properties of iron monosilicide are analyzed based on the results of thorough measurements of the conductivity, Hall coefficients, thermo emf, and magnetic characteristics of high-quality single-crystal FeSi samples at liquid helium (LHe) and intermediate temperatures. It is demonstrated that the most adequate and consistent interpretation of the experimental magnetic, transport, and optical characteristics can be given within the framework of the Hubbard model. The model parameters are determined and the arguments are presented which provide evidence of the spin polaron formation and the density of state (DOS) renormalization taking place in FeSi in the vicinity of the Fermi energy at intermediate temperatures. It was found that a decrease in the sample temperature in the region of $T < T_c \approx 15$ K is accompanied by a transition to a coherent regime of the spin density fluctuations. As a result, the ferromagnetic character of the interaction leads to the formation of magnetic microdomains with a characteristic size ~10 Å. The exchange-induced magnetization enhancement in the vicinity of charge carriers in these microdomains probably accounts for the anomalous components in the Hall coefficient and the magnetization hysteresis observed in FeSi at LHe temperatures. The nature of the low-temperature transition at $T_m \approx 7$ K in the system of interacting magnetic microparticles in iron monosilicide is discussed. *© 2001 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

In recent years, considerable attention from researchers has been devoted to the class of narrow-band-gap semiconductors based on rare-earth elements, where strong quasiparticle interactions lead to the formation of a gap in the band spectrum in the vicinity of the Fermi energy. The cubic compound FeSi [1] is also conventionally included into this class of condo insulators, primarily because the behavior of the physical characteristics of iron monosilicide [1, 3] is similar to that of the classical semiconductor SmB_6 [2].

However, some researchers (see, e.g., [4]) seriously doubt the validity of a condo lattice model in cases when the Fermi level occurs immediately within a narrow 4*f* (3*d*) band, which gives rise to rapid charge density fluctuations in samarium hexaboride (with the mean samarium valence $v_{Sm} \approx 2.6$ [5]) and spin density fluctuations in iron monosilicide [6]. Recently [7, 8], it was demonstrated that the most adequate description of the anomalous behavior of SmB_6 can be given within the framework of a model [9] assuming the formation of exciton–polaron complexes in the samarium centers at low temperatures ($T \le 15$ K) due to rapid (~10^{-12} s) fluctuations of the samarium valence.

Previously [10, 11], we presented analogous experimental data that provide evidence against applying the condo insulator model to the interpretation of the physical characteristics of FeSi. It was demonstrated that FeSi apparently represents a spin-polaron dielectric of the Mott type with strong Hubbard's correlations. At the same time, the ground state formation in iron monosilicide at liquid helium (LHe) temperatures is accompanied by the appearance of an anomalous component in the Hall coefficient [10–12] and some features in the temperature dependence of the SHF conductivity, thermo emf, and capacitance [13]. We suggested [10, 13] that the hysteresis of these anomalies in the region of $T = T_m \approx 7$ K might be explained in terms of a phase transition taking place in the electron subsystem of FeSi, but the nature and character of the electron structure rearrangement still remain unclear.

In this context, the purpose of this work was to study in detail the behavior of the Hall coefficient of FeSi at LHe and intermediate temperatures in a wide region around T_m and to perform thorough measurements of the magnetization and magnetoresistance on high-quality single-crystal FeSi samples.

2. EXPERIMENTAL METHODS

The experiments were performed on single-crystal samples made of the same FeSi ingot as that used previously [10, 13]. Special attention was paid to preparing the sample surface and making contacts for the resistance measurements. The magnetization measurements were performed using a setup employing commercial (PARC Model M155, USA) and laboratory vibration magnetometers with cryostats operating in a wide range of working temperatures (2–300 K) and with an electromagnet ($H \le 12.5$ kOe).

The Hall coefficients were measured in an automated experimental setup of an original design, with a sample rotated by a step (3.6°) electric motor in a fixed magnetic field ($H \le 80$ kOe) of a superconducting solenoid. In the range of small signals, the accuracy of measurements was increased by using a Keithley Model 2182 nanovoltmeter. The process of data acquisition, sample rotation, and temperature variation was controlled by a personal computer with a microprocessor-based interface of special design.

3. EXPERIMENTAL RESULTS

Figure 1a shows typical experimental curves obtained by measuring the field dependence of the Hall resistance $R(H)$ of an FeSi sample at LHe temperatures. A characteristic feature of the $R(H)$ curves is the hysteresis observed in the interval of magnetic fields $H < 5$ kOe, which reflects the appearance of an anomalous component in the Hall response (see also [10, 12]). It should be noted that the anomalous Hall signal amplitude in the region of $H \le 10$ kOe markedly exceeds the normal Hall component $U^H(H)$.

In contrast, the sample magnetization behavior at LHe temperatures (Fig. 1b) is characterized by a dominating paramagnetic contribution, linearly depending on the magnetic field strength in the range of $H \le 12.5$ kOe, and a small “ferromagnetic” contribution $M_s(H)$. Small absolute values [$M_s(H) < 5 \times 10^{-3}$ (G cm^3)/g] and the irreproducibility of the magnetization, accompanying variation of the sample cooling conditions and repeated magnetization cycles in a single-crystal FeSi sample (memory effects), are indicative of the magnetic moment “freezing” in the iron monosilicide matrix and the FeSi transition to a spin glass state at LHe temperatures see, e.g., [14]).

In order to elucidate the features of the magnetic moment formation in FeSi low temperatures, we have thoroughly studied the angular dependence of the Hall resistance. Figures 2–4 show the results of the $R(\varphi)$ measurements (where φ is the angle between the [111] normal to the sample surface and the magnetic field vector $\mathbf{H}$) for various constant values of $H < 80$ kOe and temperatures in the interval from 1.6 to 20 K. An analysis of the angular variation of the Hall resistance at LHe temperatures (Figs. 2 and 3) confirms dominating contribution of the anomalous component for the field strengths $H \le 20$ kOe. An increase in the magnetic field strength is accompanied by a sharp narrowing of the hysteresis loop: the loop width at $H > 10$ kOe becomes comparable to the experimental error. Simultaneously, the increase in H leads to an increase in the normal Hall voltage component (Figs. 2 and 3). Note also that the Hall resistance versus angle curves measured in FeSi at LHe temperatures in the region of field strengths $H \ge 35$ kOe reveal the appearance of an additional double-frequency component (indicated by the arrows in Fig. 2b).

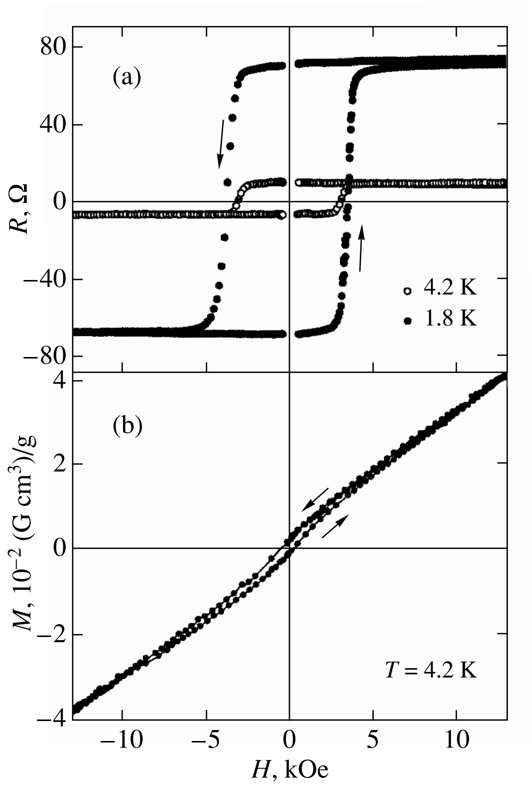


Fig. 1. Variation of (a) the Hall resistance $R(H)$ and (b) the magnetization $M(H)$ of an FeSi sample with a magnetic field strength at LHe temperatures (arrows indicate the direction of angle variation during the sample rotation).

As the temperature increases above the LHe level, the anomalous Hall component amplitude exhibits a sharp drop (Figs. 4 and 5), which is accompanied by a decrease in the hysteresis loop width $\Delta\varphi_{\text{hyst}}$. Figure 5 shows the angular dependence of $\Delta R_{\text{hyst}}(\varphi) = (R_+ - R_-)/2$ representing a half-difference between the $R(\varphi)$ curves (depicted in Fig. 4) measured for a sample rotated in opposite directions. The amplitude of the anomalous component of the Hall resistance determined by this method is used in what follows for separating the normal and anomalous contributions to the Hall coefficient.

The data presented in Fig. 5 can also be used to study the temperature variation of the hysteresis width φ_{hyst} measured at half maximum of $\Delta R_{\text{hyst}}(\varphi)$. Figure 6 shows the temperature dependence of the $\Delta\varphi_{\text{hyst}}$ value determined from the results of measurements per-

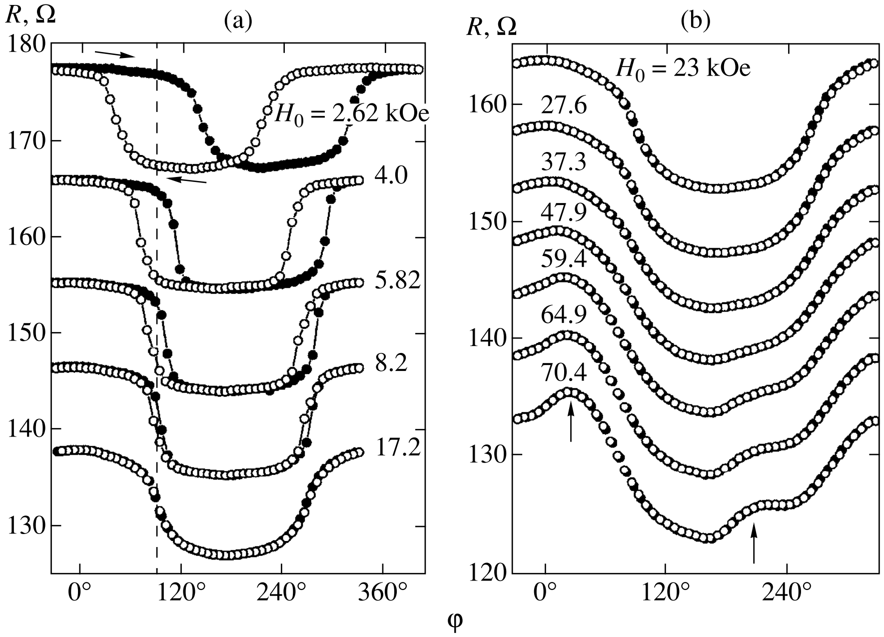


Fig. 2. The angular profiles $R(\varphi)$ of the Hall resistance measured for an FeSi sample at $T = 4.2$ K at various field strengths H_0 (indicated at the curves). Arrows indicate (a) the direction of angle variation during the sample rotation and (b) the additional double-frequency Hall signal components.

formed for various values of the magnetic field strength in the interval of $H < 4$ kOe. Note that the relatively slow decrease in the hysteresis loop width $\Delta\varphi_{\text{hyst}}$ observed at $T \le 11$ K is followed by a sharp drop in $\Delta\varphi_{\text{hyst}}$ upon further increase in the temperature (Fig. 6): the $R(\varphi)$ curves for $T > 15$ K measured in the entire range of the magnetic field strength show only the normal (sinusoidal) Hall signal component related to

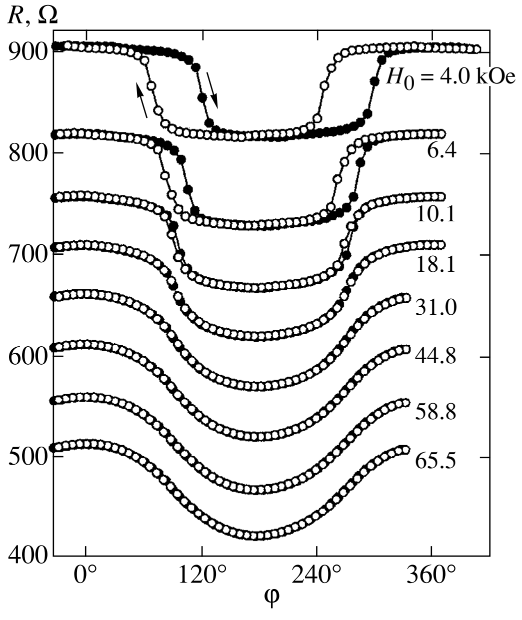


Fig. 3. The angular profiles $R(\varphi)$ of the Hall resistance measured for an FeSi sample at $T = 1.73$ K at various field strengths H_0 (indicated at the curves). Arrows indicate the direction of angle variation during the sample rotation.

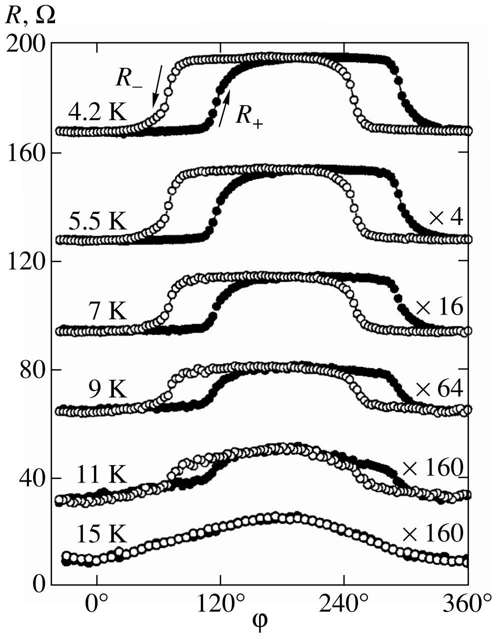


Fig. 4. The angular profiles $R(\varphi)$ of the Hall resistance measured for an FeSi sample at $H_0 = 1.63$ kOe and various temperatures (indicated at the curves with the corresponding scaling factors). Arrows indicate the direction of angle variation during the sample rotation.

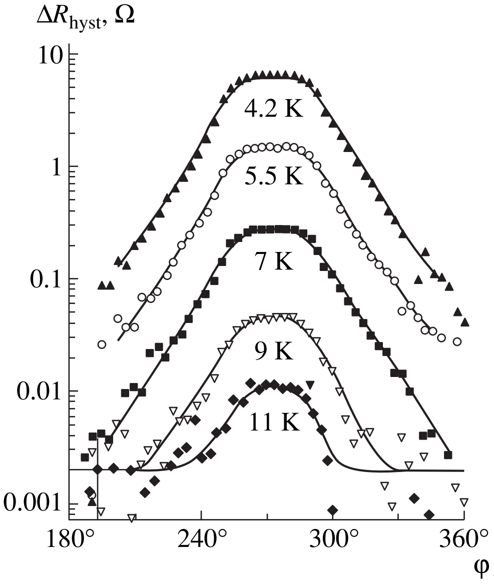


Fig. 5. Variation of the anomalous Hall resistance component $\Delta R_{\text{hyst}}(\varphi) = (R_+ - R_-)/2$ of an FeSi sample measured at H_0 = 1.63 kOe and various temperatures (indicated at the curves). The curves are constructed by data of Fig. 4.

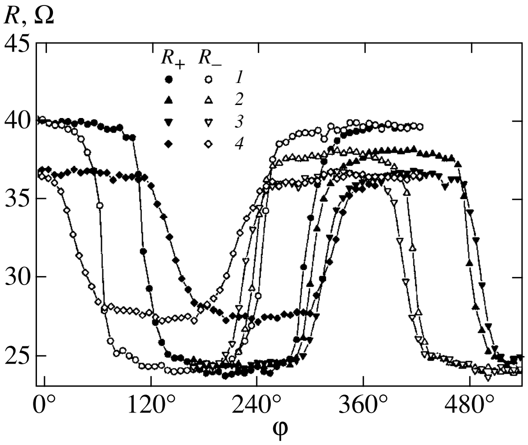


Fig. 7. A series of the angular profiles of the Hall resistance $R(\varphi)$ measured at T = 4.2 K and the field strength H_0 changed between fixed values H_0 for various orientations of an FeSi sample relative to the magnetic field vector **H**: φ_0 = 180, (1) ⟶ (2) ⟶ (3); φ_0 = 360°, (3) ⟶ (4); H_0 = 1.7 (*1*), 1.28 (*2*), 1.06 (*3*), and 0.89 kOe (*4*).

charge carriers of the *p*-type in the FeSi matrix (see, e.g., Fig. 4). It should also be emphasized that the signs of the Hall coefficient $R_H(T)$ and the thermo emf $S(T)$ coincide in the temperature interval from 15 to 70 K. As the temperature decreases below 15 K, the appearance and growth of the anomalous (negative) Hall signal component is accompanied by a rapid decrease in the positive $S(T)$ value [10].

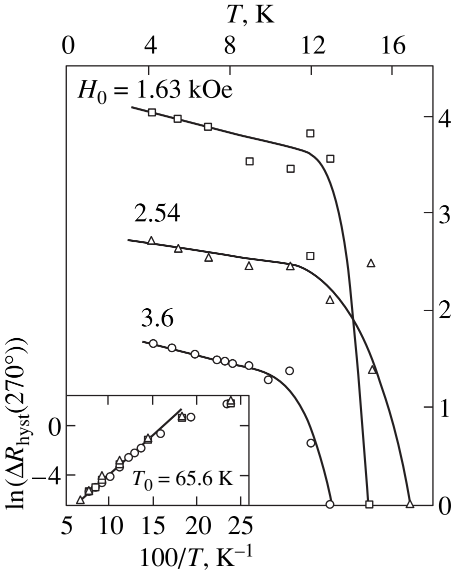


Fig. 6. Temperature variation of the hysteresis loop width $\Delta\varphi_{\text{hyst}}$ for the Hall resistance of an FeSi sample measured for three values of the magnetic field strength (indicated at the curves). The inset shows a plot of the activation parameter $\Delta R_{\text{hyst}}(270°)$ versus reciprocal temperature constructed by data for the three H_0 values.

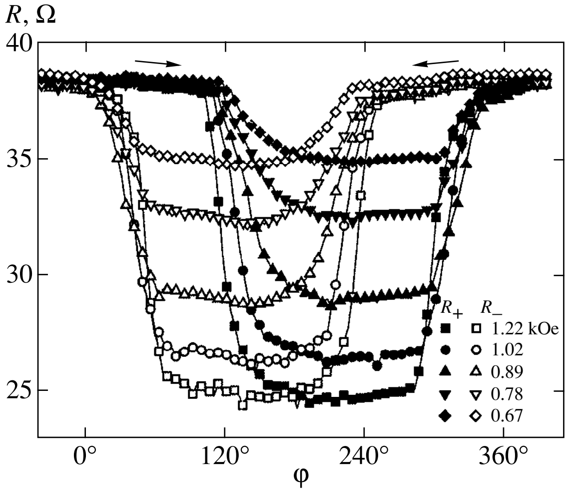


Fig. 8. A series of angular profiles of the Hall resistance $R(\varphi)$ for an FeSi sample measured at T = 4.2 K and the field strength H_0 changed from H_0 = 1.22 to 0.67 kOe, showing variation of the hysteresis contribution to $R(\varphi)$. The H_0 value was changed for a fixed sample orientation (φ_0 = 180°) relative to the magnetic field vector **H**.

The aforementioned memory effects related to a dependence of the experimental parameters of FeSi on the sample prehistory (i.e., on the cycles of temperature and magnetic field variation in the sample) were also manifested in the angular dependence of the Hall resistance. Figures 7 and 8 show two families of the experimental curves $R(\varphi)$ measured at LHe temperatures for fixed H_0 values in the 0.5–2 kOe interval. In the exper-

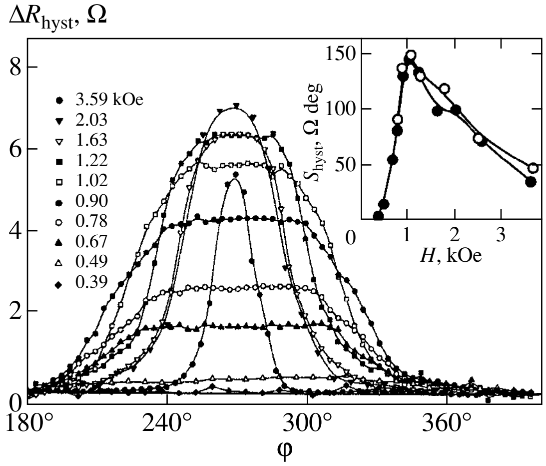


Fig. 9. A series of angular profiles of the anomalous component of the Hall resistance $\Delta R_{\mathrm{hyst}}(\varphi) = (R_+ - R_-)/2$ for an FeSi sample measured at $T = 4.2$ K and various values of the magnetic field strength H_0. The inset shows a plot of the hysteresis loop area S_{hyst} versus field strength H constructed by the data of Fig. 9.

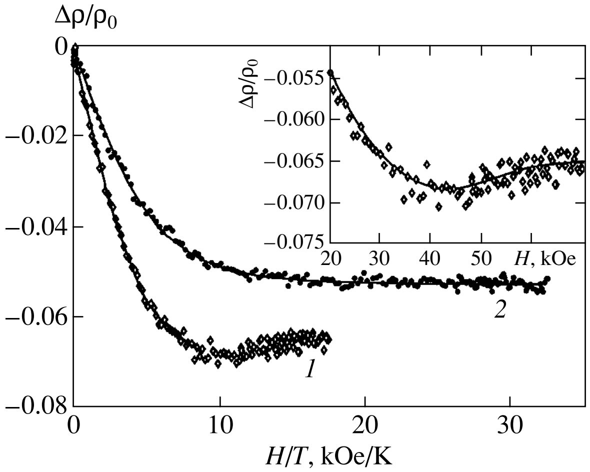


Fig. 10. Plots of the magnetoresistance versus field strength for an FeSi sample measured at $T = 4.2$ (*1*) and 1.73 K (*2*). The inset shows a part of the $\Delta\rho/\rho_0(H)$ curve measured at $T = 4.2$ with the magnetic field strengths above 20 kOe.

iment presented in Fig. 7, the values of $R(\varphi = 0)$ and $R(\varphi = 180°)$ (the difference of which gives the double total Hall resistance) were found to depend both on the applied magnetic field amplitude and on the magnitude and direction of the **H** vector during the H_0 variation between the fixed values. Moreover, the H_0 variation for a strictly determined sample orientation relative to the magnetic field (for **H** perpendicular to the sample surface) is accompanied by one of the boundary values ($R(\varphi = 0)$ in Fig. (8)) remaining constant, whereat a decrease in the Hall resistance with decreasing H_0 proceeds virtually entirely at the expense of the variation of the $R(\varphi = 180°)$ value. An analysis of the curves presented in Fig. 7 and 8 allows us to estimate a change in the hysteresis amplitude for the Hall resistance in the magnetic field (see, Fig. 9). The inset in Fig. 9 shows a field dependence of the hysteresis loop area $S_{\mathrm{hyst}}(H)$ determined by integrating the experimental $\Delta R_{\mathrm{hyst}}(\varphi)$ curves depicted in Fig. 9.

In order to determine the effect of the magnetoresistance contribution on the results of measurements of the Hall signal component, we have also studied the dependence of the FeSi sample resistivity on the field strength, angle, and temperature in the $T \leq 10$ K interval. Figure 10 shows typical plots of the negative magnetoresistance versus field strength measured for FeSi in the LHe temperature range. Note that the behavior of the negative magnetoresistance (as well as of the Hall resistance considered above) in FeSi significantly differs from that observed for the $\mathrm{SmB_6}$ condo insulator (considered as the analog of FeSi), where the $\Delta\rho(H)$ obeys a quadratic law for the field strength of up to 500 kOe [15]. The $\Delta\rho(H, T = 4.2\ \mathrm{K})$ curve of FeSi in the range of $H \leq 35$ kOe is approximated to a sufficiently good precision by the Brillouin function (Fig. 10):

$$\begin{aligned}\Delta\rho^{\mathrm{I}} &= \rho_0 B_J(\alpha)\\ &= \rho_0\left(\frac{2J+1}{2J}\coth\frac{2J+1}{2J}\alpha - \frac{1}{2J}\cot\frac{\alpha}{2J}\right),\end{aligned} \tag{1}$$

where $\alpha = g\mu_B JH/k_B T$, μ_B is the Bohr magneton, and k_B is the Boltzmann constant. An additional contribution ($\Delta\rho^{\mathrm{II}}$) appearing in $\Delta\rho(H)$ at LHe temperatures gives rise to a broad minimum in the region of $H \approx 35$ kOe, which is followed by a tendency to saturation for $H \geq 60$ keV (Fig. 10, curve *1* and inset). As the temperature decreases from 4.2 to 1.7 K, additional contribution is not observed for the fields below 70 kOe and the magnetoresistance is described by the Brillouin function with good precision in the entire range of field strengths studied in this work (Fig. 10, curve *2*).

Based on the results of measurements of the angular variation of the magnetoresistance (Fig. 11), we may exclude any significant dependence of the above-described (see Figs. 1–5) anomalies in the Hall resistance on the negative magnetoresistance related to a “nonequipotential” arrangement of the Hall contacts on the sample surface. Our investigations showed that the $\Delta\rho/\rho(\varphi)$ value in the indicated interval of temperatures and the range of fields (up to 70 kOe) does not exceed 0.5%. The anomalous angular dependence of the negative magnetoresistance (Fig. 11) is related to the aforementioned $\Delta\rho^{\mathrm{II}}(H)$ component (contributing to the magnetoresistance at $H \geq 35$ kOe) and vanishes when the temperature increases to a level of $T = T_m \approx 7$ K.

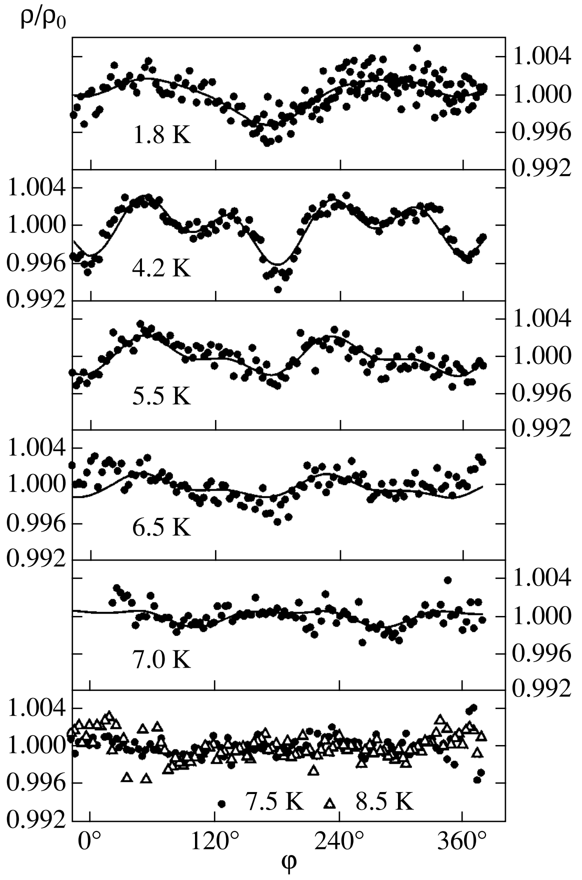


Fig. 11. Angular variation of the magnetoresistance $\rho(H = 70.2\ \text{kOe}, \varphi)/\rho(\varphi_0)$ observed at various temperatures for an FeSi sample rotated in a constant magnetic field $H = 70.2$ kOe.

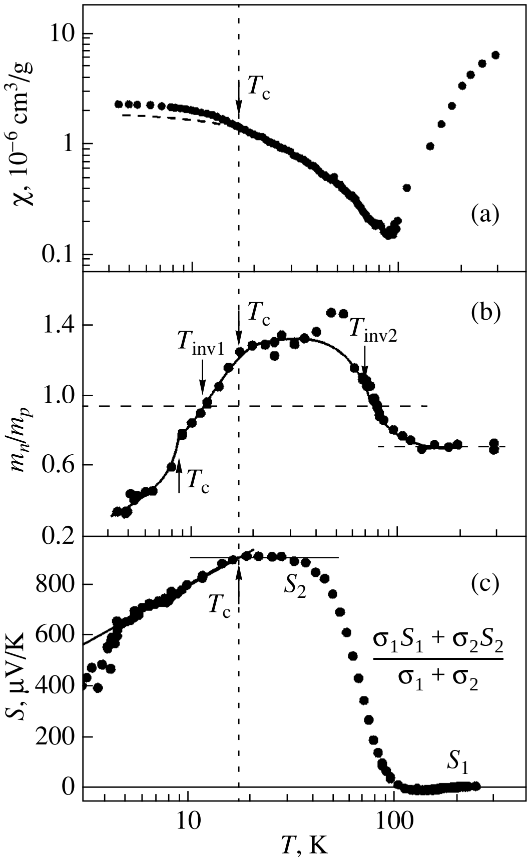


Fig. 12. Temperature variation of (a) the magnetic susceptibility $\chi(T)$, (b) the effective mass ratio m_n/m_p of charge carriers, and (c) thermo emf coefficient $S(T)$ in FeSi.

In concluding this section, prior to proceeding with the discussion and analysis of the physical properties of iron monosilicide, we will briefly consider the results of magnetization measurements performed in the region of intermediate temperatures. Note that a restricted precision of the available magnetometers allowed us to perform reliable measurements of $M(H, T_0)$ only in the regions of $T \leq 50$ K and $T \geq 120$ K. In the vicinity of a minimum observed in the magnetic susceptibility of FeSi at $T \approx 90$ K (Fig. 12a), the magnetic response exhibited a sharp drop in amplitude and the experimental data were poorly reproducible.

Figure 13 shows the results of magnetization measurements in FeSi. As is seen from Fig. 13a, the $M(H)$ curves obtained at $T \geq 120$ K and $H \leq 12.5$ kOe are well approximated by the initial portion of the Brillouin function. The temperature dependence of the magnetic susceptibility $M(T)$ of FeSi measured at $T \leq 50$ K shows a weakly pronounced inflection at $T_c \approx 13$–15 K (Fig. 13b) correlated with the appearance of low-temperature anomalies in the Hall resistance (Fig. 2–5) and the negative thermo emf contribution [10] (see also Fig. 12c) in FeSi. The interpolation of the results of magnetic measurements (Figs. 12a and 13b) to the temperature interval below the inflection point ($T < T_c$) reveals an additional magnetic contribution to $M(T, H_0)$ (Fig. 13b) and $\chi(T)$ (Fig. 12a). As the temperature decreases to $T \leq 7$ K, this contribution apparently converts into the aforementioned "ferromagnetic" magnetization component $M_s(H)$ in FeSi (see Fig. 1b).

Finally, it should be noted that a generally similar behavior of iron monosilicide was observed in [16] where single-crystal FeSi samples were studied using a SQUID magnetometer. However, the measurements in [16] were performed only at 5 and 300 K. This restriction did not allow the results to be interpreted with an allowance for various contributions to $M(H, T_0)$ in the passage from intrinsic to impurity conductivity in FeSi.

4. DISCUSSION OF RESULTS

An analysis of the experimental results presented above reveals a sharp quantitative difference between magnitudes of the components related to the normal and anomalous (magnetic) effects in the values of magnetization and those in the Hall coefficient. Indeed, the

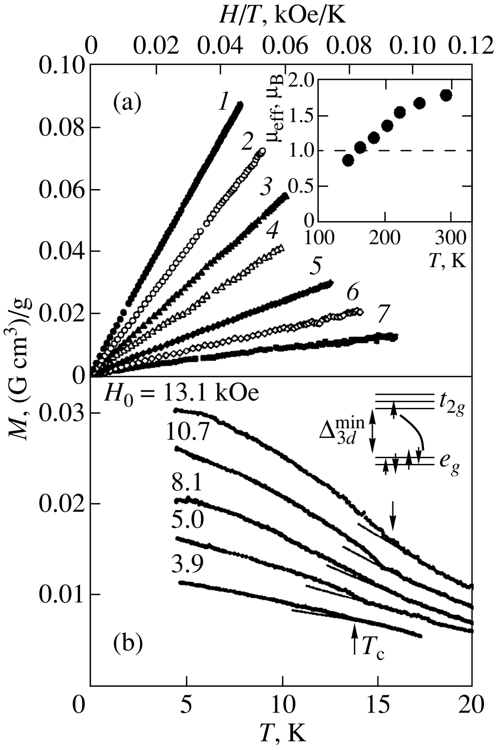


Fig. 13. (a) The plots of magnetization M versus H/T measured at fixed temperatures above 120 K: T = 290 (*1*), 250 (*2*), 220.2 (*3*), 200.7 (*4*), 179.5 (*5*), 159.5 (*6*), 141 K (*7*). The inset shows the temperature variation of the effective magnetic moment μ_{eff} determined upon approximating the experimental data by formula (3). (b) The plots of magnetization M versus T measured in the region of $T_c \approx 15$ K at fixed values of the magnetic field strength H_0 (indicated at the curves). The inset shows a schematic diagram illustrating the Fe 3*d* band splitting in the FeSi crystal field.

Hall coefficient determined at LHe temperatures and the field strengths below 10 kOe is characterized by a dominating anomalous contribution. At the same time, the description of magnetization shows a dominating role of the paramagnetic contribution determining both the signal amplitude and the character of the $M(H)$ value variation in FeSi. Since the experimental data were obtained for single-crystal FeSi samples of high quality, where the presence of magnetically-ordered impurity phase in the bulk is hardly probable, the enhanced anomalous contribution to the Hall coefficient has to be explained by assuming a local increase in magnetization in the immediate vicinity of charge carriers.

We believe that, in this situation, a highly promising approach to the interpretation of the anomalous transport and magnetic characteristics of iron monosilicide is offered by the method proposed previously [10, 11]. According to this approach, the properties of FeSi are described within the framework of the Hubbard model [17] in the critical range of parameters $2 \le U/D < 3$, where U is the Hubbard repulsion and 2D is the conduction band width (see also [18, 19]. In accordance with the results of the investigation of the transport properties and thermo emf [10, 11], we must distinguish several temperature intervals featuring different variations of the physical properties of FeSi.

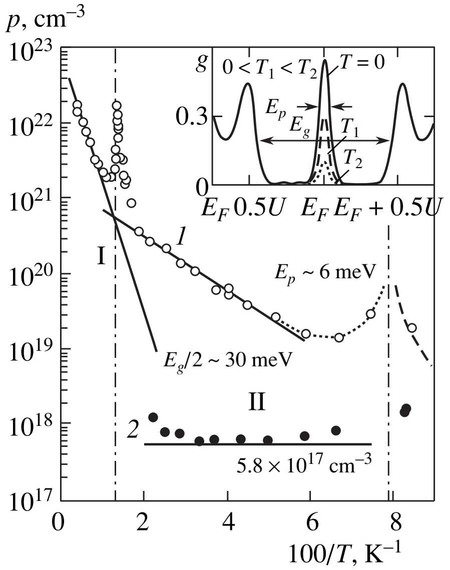


Fig. 14. Activation asymptotics of the parameters (*1*) $p = \frac{\mu_n/\mu_p - 1}{\mu_n/\mu_p + 1}\frac{1}{R_H|e|}$ with $\mu_n/\mu_p \approx 1.5$ [10] and (*2*) $p = N_{\text{Fe}}\exp(|e|S_2/k_B)$ for FeSi. The inset shows the density of states g in the Hubbard model with $2 \le U/D \le 3$, plotted by the results of calculations [17] (see also [18, 19]).

4.1. Intrinsic Conductivity in FeSi (T > 100 K)

According to the conclusions made in [10, 11], the intrinsic conductivity is observed in FeSi at $T > 100$ K. This regime corresponds to the charge carrier activation via the indirect gap ($E_g/2 \approx 30$ meV) separating the Hubbard bands composed predominantly of Fe 3*d* states. Figure 14 (curve *1*, region I) illustrates the behavior of charge carriers (for example, of the *p*-type) in the region of intrinsic conductivity determined from the results of measurements of the galvanomagnetic characteristics of FeSi [10].

The activation asymptotics determined for the intrinsic carrier concentration in the temperature interval $T > 100$ K can be used to evaluate the effective magnetic moment μ_{eff} of charge carriers in the upper 3*d* band (conduction band) from the results of our magnetization measurements (see Fig. 13a). Indeed, the results of the magnetic susceptibility measurements in a broad range of temperatures (77–800 K) [20] are indicative of a nonmagnetic character of the Fe 3*d* states (singlet $S = 0$) in the lower (valence) band, thus suggesting that the magnetic properties of FeSi must be related to 3*d* states in the conduction band. Approximating the $\chi(T)$ function by an expression of the type (see, e.g. [20])

$$\chi(T) = \frac{Ng^2\mu_B^2}{3k_BT}\frac{S(S+1)(2S+1)}{2S+1+\exp(E_g/k_BT)}, \qquad (2)$$

we obtain the values $S = 1$ and $g \approx 2.17$ for the parameters characterizing the upper 3*d* band of FeSi. Using the results of the magnetic susceptibility calculation performed in [20] for the two-level system of singlet and triplet states separated by a gap with a width of $E_g \approx$ 60–70 meV, we may describe with sufficiently high precision the behavior of the $\chi(T)$ curve in FeSi in the temperature range above room temperature.

We described the experimental data for the temperature interval from 120 to 300 K (presented in Fig. 13a) within the framework of the Curie–Weiss relationship

$$M = \frac{n_0\exp(-E_g/2k_BT)\mu_{\rm eff}^2(T)}{3k_BT}H. \qquad (3)$$

The values of $\mu_{\rm eff}(T)$ for the Fe centers determined from this analysis vary from $2\mu_B$ (at room temperature) to $\mu_{\rm eff} \approx \mu_B$ at $T \approx$ 140–160 K (see the inset in Fig. 13a). It should be emphasized that the latter value ($\mu_{\rm eff} \approx \mu_B$) corresponds to the free carrier with $s = 1/2$ and $g = 2$. Note also that the value $\mu_{\rm eff} \approx 2\mu_B$ was previously reported in [21], where the magnetic susceptibility of FeSi was measured at temperatures above 700 K; therefore, this value can actually be considered as characteristic of the Fe 3*d* states of charge carriers in the conduction band of FeSi.

In a tetrahedral environment of the paramagnetic Fe ion occurring in a cubic B20 crystal structure of FeSi, we may additionally expect that the Fe 3*d* bands would be split into e_g and t_{2g} subbands. We believe that the value $E_g \approx 60$ meV obtained for FeSi can be related to the minimum energy $\Delta_{3d}^{\rm min}$ of splitting in the crystal field of FeSi. The most probable structure of 3*d* bands corresponds to a tetravalent state of iron in this compound with silicon and accounts for the thermoactivated transitions between e_g (nonmagnetic singlet) and t_{2g} states (see the inset in Fig. 13b). In this situation, an electron–hole pair, which appears at the Fe center as a result of the thermal activation of a carrier from the completely filled e_g subband, has the spin $S = 1$ and the effective magnetic moment $\mu_{\rm eff} \approx 2\mu_B$ (see the inset in Fig. 13a).

4.2. Spin–Polaron Transport in FeSi (T < 100 K)

As the temperature decreases below 100 K, the regime of intrinsic conductivity is replaced by the carrier transport via spin-polaron states on the Fermi level [10, 11]. This regime is characterized by the activation behavior of the conductivity and the Hall coefficient (Fig. 14, region II) in the temperature interval from 15 to 50 K (activation energy $E_p \approx 6$ meV [10]) and by a nearly constant value of the thermo emf $S \approx 900 \pm 20$ μV/K in this interval [10, 11] (see also Fig. 12c). In this situation, the most probable explanation of the features in the behavior of the physical characteristics of FeSi can be given assuming that this crystal features strong Coulomb (or Hubbard) correlations at intermediate temperatures. The conduction and valence bands of FeSi (composed predominantly of the Fe 3*d* states) correspond to the upper and lower Hubbard bands. Within the framework of our approach, the valence band is essentially a Hubbard band of doubly occupied states (singlet $S = 0$, see the inset in Fig. 13b).

Under the conditions of strong hubbard correlations, the concentration of charge carriers is usually estimated using the Hicks formula [22]:

$$S(T) = -\frac{k_B}{|e|}\ln\left(\frac{1-\nu}{\nu}\right). \qquad (4)$$

Here, parameter ν is the reduced concentration of carriers (*p*-type) in the lower Hubbard band ($\nu = p/N_{\rm Fe} \approx 1.3 \times 10^{-5}$). Using this value, we may estimate the hole concentration as $p \approx 6 \times 10^{17}$ cm^{-3} [11]. Thus, the spin-polaron states (corresponding [23] to the regime of strong Hubbard correlations) in FeSi are characterized by a small (~10^{17} cm^{-3}) concentration of carriers (holes) in the e_g and t_{2g} subbands (see the inset in Fig. 13b). These states may appear as a result of rapid spin fluctuations between Fe 3*d* states in the conduction band and at the Fe centers situated in the nearest environment of a charge carrier. The activation energy of the low-temperature transport $E_p \approx 6$ meV (Fig. 14, curve *1*, region II) must correspond to the potential well depth of the spin-polaron states. The actual carrier concentration in the Hubbard bands of a given single-crystal FeSi sample is fixed on a level of 6×10^{17} cm^{-3} (Fig. 14, curve *2*).

It should be emphasized that, according to the conclusions made in [17–19], the appearance of spin-polaron states at the E_F level is related to the electron spectrum renormalization in the Hubbard model. The low-temperature multiparticle resonance at E_F (see the inset in Fig. 14) accounts for the "metallization" of the spectrum related to the transition to a narrow-band conductivity regime corresponding to charge carriers possessing a large effective mass. Apparently, the aforementioned decrease in the effective mass to $\mu_{\rm eff} \approx \mu_B$ (see the inset in Fig. 13a) with decreasing temperature should be related to a transition to the regime of charge carrier transport via a narrow band situated in the vicinity of E_F.

In order to estimate the role of the effective mass renormalization in the bands of FeSi, we can treat the results of our galvanomagnetic measurements [10] using the relationships

$$R_H\sigma = \mu_p - \mu_n, \quad (5)$$

$$\mu_{n,\,p} = e\tau/m_{n,\,p}, \quad (6)$$

which are applicable to semiconductors in the region of intrinsic conductivity. Taking into account the room-temperature (300 K) carrier mobility ratio $\mu_n/\mu_p = 1.5$ and the relationship $\mu_p - \mu_n \approx A/\sqrt{T}$ valid in the 150–300 K range, we may evaluate the variation of the m_n/m_p ratio in FeSi with the aid of formulas (5) and (6). Figure 12b shows the results of the calculation of the value $m_n/m_p \approx f(T)$ for iron monosilicide in a wide temperature range (4–300 K). A twofold increase in the m_n/m_p ratio observed for FeSi in the temperature interval from 50 to 150 K corresponds to the transition from electron states in the upper Hubbard band (conduction band with the density of states ~13 states/eV per unit cell; see, e.g., [24]) to the states of the multiparticle resonance at E_F (see the inset in Fig. 14). As is seen in Fig. 12b, the effective "weight gain" of electrons upon the formation of spin-polaron states at the Fermi level is just what accounts for the inversion of sign of the Hall coefficient at $T = T_{\text{inv2}} \approx 70$ K [10] corresponding to the condition $m_n \approx m_p$.

Another, quite correct estimate of the increase in amplitude of the multiparticle resonance at E_F can be obtained from the results of direct measurements of the magnetic susceptibility of iron monosilicide (Fig. 12a). The transition (accompanying a decrease in the temperature) from the exponential $\chi(T)$ asymptotics observed in the region of $T \sim 150$–500 K to a higher susceptibility level at $T < 90$ K (with allowance for the density of states depicted in the inset in Fig. 14) must be related to an increase in the paramagnetic Pauli's contribution to $\chi(T)$. In the approximation of a narrow rectangular potential well (with a width of $\delta = E_p \approx 6$ meV) at the Fermi level, the $\chi_p(T)$ value can be estimated as

$$\chi_p(T) = \frac{1}{2}g^2\mu_B^2 N^*(E_F)\int_{-\delta/2}^{\delta/2}\left(-\frac{\partial f}{\partial E}\right)dE. \quad (7)$$

Taking into account the $\chi_p(T)$ renormalization contribution due to the integral in the right-hand part of (7), which allows for a 4- to 5-fold increase in $\chi(T)$ with the temperature (for details, see, e.g., [25]), the gain factor directly describing the DOS renormalization at the Fermi level can be evaluated as $N^*(E_F, T = 4.2\text{ K})/N^*(E_F, T = 100\text{ K}) \geq 20$.

Thus, the formation of spin polarons at low temperatures ($T < 200$ K) in FeSi is accompanied (within the framework of the approach adopted), by the appearance of a narrow multiparticle resonance at the Fermi level. The resonance is characterized by the DOS value $N^*(E_F) \geq 20$ states/eV per unit cell. It should be noted that the conclusion concerning a significant DOS renormalization in the energy gap region was independently made in [26] based on the results of experiments with electron tunneling in FeSi.

Within the framework of the approach adopted in this study, we will estimate the parameters of the Hubbard model and the microscopic characteristics of spin polarons in iron monosilicide. For this purpose we will use the results of calculations of the optical conductivity of FeSi [17] in comparison with the optical spectra $\sigma(\omega)$ measured for iron monosilicide in a broad frequency interval in the IR range [27]. A comparison of data on the position and width of the $\sigma(\omega)$ maximum at $U/2$ [11] leads to the following values of the model parameters: $U \approx 270$ meV; $2D \approx 210$ meV; $U/D \approx 2.6$; and $E_g = U - 2D \approx 60$ meV; note that both the energy gap E_g and the model parameter $U/D \approx 2.6 \leq 3$ (determining the critical behavior of the system) are in good agreement with the values predicted in [17].

Using a criterion analogous to the Stoner product $UN(E_F) > 1$ for a band structure featuring multiparticle spin-polaron resonance at the Fermi level (see the inset in Fig. 14), we obtain $UN^*(E_F) > 5$. Thus, within the framework of the approach adopted for iron monosilicide, we may expect that the spin-polaron states must lead to the formation of ferromagnetic microdomains in FeSi at low temperatures.

Under the conditions of a strong electron–phonon interaction in FeSi [27], the width γ_i^{ph} of the bands of optical phonons observed in the 180–400 cm^{-1} interval can be used for estimating the charge carrier relaxation time in the Hubbard bands. Using the values of $\gamma_i^{ph} \approx 10$ cm^{-1} (for the bands at 198, 318, and 338 cm^{-1}), we obtain an estimate of the relaxation time $\langle\tau_{e-ph}\rangle \approx 1/3\gamma_i^{ph} \approx 1.7 \times 10^{-13}$ s. Using formula (6) and the values of $\mu_{n,\,p} \approx 4$–6 cm^2/(V s) [10], we may calculate the effective mass of the charge carrier for the upper and lower Hubbard bands: $m_n^*(T \approx 200\text{ K}) \approx 50m_0$ and $m_p^*(T \approx 200\text{ K}) \approx 75m_0$. It should be noted that a similar value of $m^* \approx 50m_0$ was obtained [28] from the results of measurements of the optical conductivity $\sigma(\omega)$ for FeSi in a broad (50–25000 cm^{-1}) spectral range.

Using these values of parameters m_n^* and m_p^* and taking into account the behavior of their ratio $m_n^*/m_p^* = f(T)$ (see Fig. 12b), we obtain an estimate from above for the effective mass of spin polarons at low temperatures, corresponding to the states of the multiparticle resonance at the Fermi level (see the inset in Fig. 14): $m_{sp}^* \leq 100m_0$. It should be noted that the relaxation times $\langle\tau_{e-ph}\rangle$ analogous to those calculated above can be obtained by estimates based on the width of the line of polarized neutron scattering in FeSi: $\Gamma \approx 1/\langle\tau_{e-ph}\rangle \approx$ 3–4 meV [29].

The localization radius a_p of the spin-polaron state can be estimated using the known expression [30]

$$a_p = \hbar/\sqrt{2E_p m_{sp}^*}. \quad (8)$$

For the values of the polaron potential well depth ($E_p \approx$ 6 meV) and the spin polaron mass ($m_{sp}^* \approx 100m_0$) determined above, we obtain an estimate $a_p \approx 5$ Å. At the same time, the a_p value in the region of the metal–insulator transition can be estimated using the relationship (see, e.g., [22])

$$a_p \approx a(2\pi U/2D)^{1/5} \approx 1.5a \approx 4.2\text{ Å}, \quad (9)$$

where parameter a was taken equal to $r_{\text{Fe–Fe}} \approx 2.75$ Å corresponding to the shortest distance between iron atoms in the FeSi lattice. The good agreement between two independent estimates of the spin polaron localization radius in iron monosilicide is additional evidence for the validity of the proposed interpretation.

4.3. Ground-State Formation in the Spin Polaron System of FeSi

As the temperature decreases to within the interval $T \le 15$ K, we may expect (taking into account the ferromagnetic character of exchange interactions due to the spin density fluctuations in FeSi [29]) an exchange-induced magnetization enhancement as a result of the polarization of the Fe centers in the vicinity of all charge carriers occurring in the upper Hubbard band. In FeSi at low temperatures, we may also expect the formation of “ferromagnetic microdomains” with dimensions on the order of 10 Å, possessing a nonzero total magnetic moment. The concentration of these magnetic particles in FeSi is directly related to the concentration of spin polarons and amounts (for the single crystal studied) to 10^{17} cm^{-3}.

Under these conditions, we may readily explain the aforementioned difference between the magnitudes of the effects related to the normal and anomalous (magnetic) components in the values of low-temperature magnetization and those in the Hall coefficient (Figs. 1a and 1b). Indeed, the exchange-induced magnetization enhancement in the vicinity of charge carriers occurring in the upper Hubbard band must be accompanied by the appearance and enhancement of the anomalous component of the Hall coefficient, whereas the integral magnetization component due to the “ferromagnetic microdomains” with a size of about 10 Å and a concentration of 10^{17} cm^{-3} is much less significant.

Thus, the formation of magnetic microdomains in the iron monosilicide matrix at low temperatures ($T < T_c \approx 15$ K) can be expected to lead to the appearance of localized magnetic moments (LMMs) in FeSi. Note that a rather similar situation takes place in micromagnetic systems (spin glasses) [14] in a superparamagnetic state. This state is characterized by the formation of a total magnetic moment for each cluster in a nonmagnetic matrix and is realized immediately before the spin glass phase transition. On the other hand, the nature of the total LMM of microdomains in the two cases may be significantly different.

We believe that the formation of magnetic particles (microdomains) with LMMs in FeSi may account for the appearance of an additional contribution to the magnetization, which is manifested by the inflection points in the $M(T)$ curves at $T \approx 15$ K (Fig. 13b). In addition, the appearance of correlations in the spin density fluctuations and a change in the characteristics of spin polarons in the vicinity of $T_c \approx 15$ K must be related to a sharp drop in the constant positive values of the Hall coefficient and thermo emf (see, e.g. Figs. 12b and 12c) [10, 11] in the temperature interval indicated. Additional evidence to confirm the hypothesis concerning the formation of ferromagnetic particles directly from spin polarons may be provided by the coincidence between the activation energy of the process of magnetization reversal ($T_0 \approx 65.6$ K ≈ 6 meV) determined for the system of magnetic microdomains in the FeSi matrix from data on the temperature dependence of parameter ΔR_{hyst} (see the inset in Fig. 6) and the value $E_p \approx 6$ meV (Fig. 14).

For LHe temperatures, the μ_{eff} value of LMMs in FeSi can be determined, for example, by analysis of the magnetoresistance curves (Fig. 10). Applying relationship (1) to the curves depicted in Fig. 10, we have obtained the following values of the effective magnetic moment for ferromagnetic microdomains formed in the immediate vicinity of charge carriers: $\mu_{\text{eff}}(4.2\text{ K}) \approx 5.3\mu_B$, $\mu_{\text{eff}}(1.73\text{ K}) \approx 3.9\mu_B$. Apparently, these μ_{eff} values should be considered as first estimates. At the same time, these μ_{eff} values, together with the results of preliminary analysis based on the separation of the Pauli and Brillouin low-temperature contributions to the magnetization $M(H, T)$ (see Fig. 13b and data in [12] and [16]), allow us to conclude that the effective magnetic moment of the magnetic particles $\mu_{\text{eff}}(T)$ varies in a nonmonotonic manner with a maximum in the vicinity of $T \approx T_m \approx 7$ K.

The effective magnetic moment of the ferromagnetic microdomains formed in the FeSi matrix grows up to ~(6–8)μ_B when the temperature decreases below $T \le T_c \approx 15$ K. The μ_{eff} value passes through a maximum and begins to decrease again when the temperature drops below $T \le T_m \approx 7$ K, which apparently indicates that the interaction between spin polarons is “switched on.” We believe that the character of variation of the $\mu_{\text{eff}}(T)$ value, together with the other features of low-temperature behavior (the shape of the magnetization hysteresis curve, memory effects, relaxation phenomena, etc.), may be evidence that a micromagnetic state

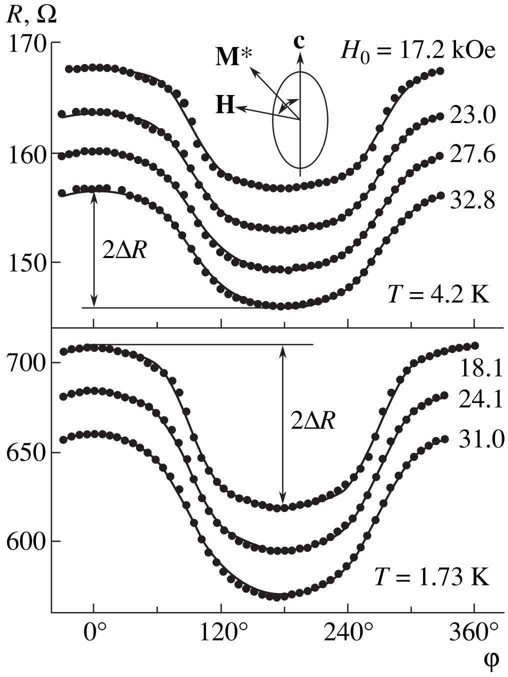


Fig. 15. The angular profiles of the Hall resistance $R(\varphi)$ at LHe temperatures approximated using formula (12) derived within the framework of the uniaxial ferromagnet model. The top inset shows the mutual arrangement of vectors **c**, **H**, and **M*** for an ellipsoidal magnetic particle.

of the spin glass type is realized in the LMM system of magnetic microdomains formed in the FeSi matrix at $T \le T_m \approx 7$ K.

At the same time, we may expect that the physical characteristics of iron monosilicide would exhibit a much more complicated behavior because the magnetic microdomains do not occupy fixed positions in the FeSi matrix. On the contrary, these microdomains play a significant role in the process of charge transfer at low temperatures. Under these conditions, we cannot exclude that, besides the anomalous physical properties characteristic of the systems with micromagnetic states, the system studied would exhibit the formation of new spatially organized magnetic structures, magnetic orientation transitions, etc. In our opinion, it is such a magnetic transition in the system of magnetic spin polarons that accounts for the features observed in the magnetic fields $H \ge 35$ kOe on the angular profiles of the Hall resistance (Fig. 2b) and on the field and angular dependences of the magnetoresistance (Figs. 10 and 11) in FeSi at LHe temperatures.

Thus, within the framework of the proposed interpretation, the low-temperature anomalies in variation of the Hall resistance (Figs. 1–9) and magnetoresistance (Figs. 10 and 11) should be related to the process of remagnetization in the system of ferromagnetic microdomains—magnetic spin polarons of low concentration (~10^{17} cm^{-3}) formed in the regime of coherent Hubbard correlations in the crystalline structure of iron monosilicide. We believe that an adequate quantitative description of the Hall effect, with the separation of the $R_H(H, T)$ components, in FeSi is provided (to the first approximation) by a model considering two groups of charge carriers and taking into account both the usual contributions to the Hall coefficient and the effect of the exchange-enhanced electron contribution.

Under these conditions, the normal component of the Hall resistance corresponds to the electron and hole contributions linear in the magnetic field, whereas the ferromagnetic character of the microdomains formed in the vicinity of electrons accounts for the anomalous contribution proportional to the local magnetization M^* of these microdomains. As a result, the Hall resistance can be represented in the following form (for $\omega\tau \ll 1$):

$$Rd \approx -R_H\frac{\sigma_p-\sigma_n}{\sigma}H + 4\pi R_H\left(\frac{\sigma_n}{\sigma}\right)^2 M^*, \tag{10}$$

where $R_H = -R_{Hn} = R_{Hp} = (pe)^{-1}$, $\sigma = \sigma_n + \sigma_p$ is the total conductivity, p, e are the hole and electron concentrations, and d is the sample thickness. At noted above, this relationship is approximate and essentially represents the usual form of description of the normal and anomalous components $R_\Sigma(H, T) = R_H^n H + R_H^a M(H, T)$ written taking into account some special features of the model adopted.

For a system of single-domain ellipsoidal ferromagnetic particles exposed to a magnetic field with the strength $H \gg H_a = 2K_u/M^*$ (H_a is the anisotropic field, K_u is the magnetic anisotropy constant), we may expect a jumplike change of the local magnetization as a function of the angle of rotation of the magnetic field vector **H**. If the rotation axis coincides with the a-axis, while c-axis is the easy magnetization direction (see the inset in Fig. 15), we may use the well known relationships for the director cosines of **H** and **M*** vectors (see, e.g., [31]) to rewrite the above relationship as follows:

$$\begin{aligned} R \approx {} & -\frac{R_H}{d}\frac{\sigma_p-\sigma_n}{\sigma}H\cos\varphi + 4\pi\frac{R_H}{d}\left(\frac{\sigma_n}{\sigma}\right)^2 M_0^* \\ & \times\left[(1+k+k^2)\cos\varphi - \left(k-\frac{7}{2}k^2\right)\cos^3\varphi + \frac{5}{2}k^2\cos^5\varphi\right], \end{aligned} \tag{11}$$

where $k = 2K_u/M^*H < 1$.

Taking into account the structure of Eq. (11), we may separately analyze the terms with odd powers of $\cos\varphi$:

$$\begin{aligned} R(\varphi) \approx {} & A_1(H, T)\cos\varphi \\ & + A_3(H, T)\cos^3\varphi + A_5(H, T)\cos^5\varphi. \end{aligned} \tag{12}$$

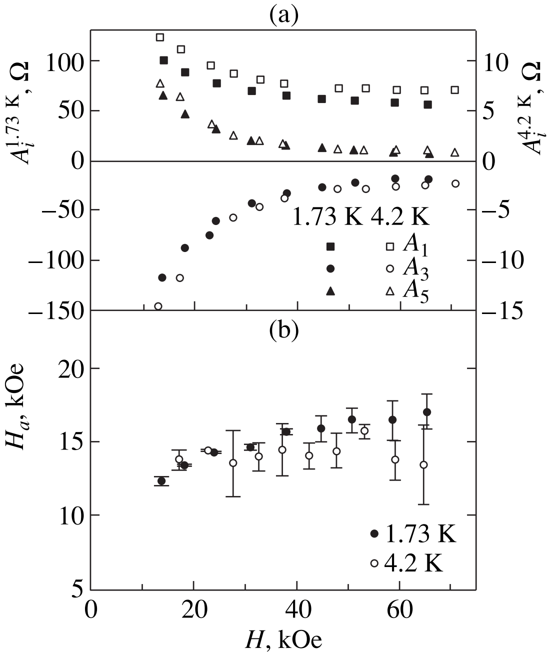


Fig. 16. Plots (a) showing coefficients A_i versus magnetic field strength calculated according to formulas (12), (13) and (b) illustrating the determination of the anisotropic field $H_a = 2K_u/M_0^*$ in FeSi at LHe temperatures.

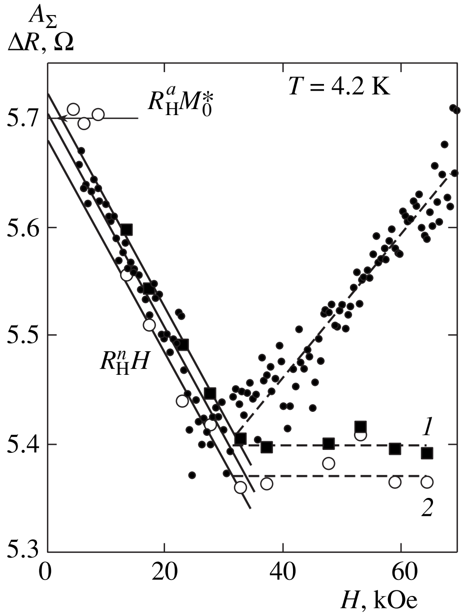


Fig. 17. Plots illustrating the procedure of the separation of the normal and anomalous contributions to the Hall coefficient in FeSi at LHe temperatures: (*1*) the sum of coefficients $A_1 + A_3 + A_5$ versus field strength H according to formulas (12) and (15); (*2*, *3*) total amplitude ΔR versus H plotted using data on the (*2*) angular and (*3*) field dependence of the Hall resistance.

Figure 15 shows an example of an approximation constructed within the framework of relationship (12) for a set of experimental curves of the Hall resistance in FeSi (Figs. 2 and 3). These curves represent analytical $R(\varphi)$ expressions calculated for a series of H values in the 15–35 kOe interval. The plots of $A_i(H, T)$ coefficients obtained in the course of the computational procedure are presented in Fig. 16a.

Comparing the expressions in (11) and (12), we readily derive the formulas for $A_i(H, T)$:

$$\begin{aligned} A_1(H, T) &= -\frac{R_\mathrm{H}}{d}\frac{\sigma_p - \sigma_n}{\sigma}H \\ &+ 4\pi\frac{R_\mathrm{H}}{d}\left(\frac{\sigma_n}{d}\right)^2 M_0^*(1 + k + k^2), \\ A_3(H, T) &= -4\pi\frac{R_\mathrm{H}}{d}\left(\frac{\sigma_n}{\sigma}\right)^2 M_0^*\left(k + \frac{7}{2}k^2\right), \\ A_5(H, T) &= 4\pi\frac{R_\mathrm{H}}{d}\left(\frac{\sigma_n}{\sigma}\right)^2 M_0^*\frac{5}{2}k^2. \end{aligned} \tag{13}$$

As is seen, the ratio of A_3 and A_5 given by formulas (13) can be used to estimate the anisotropic field $H_a(H)$ for the ferromagnetic microdomains formed in FeSi:

$$H_a(H) = -H/\left(\frac{5}{2}\frac{A_3}{A_5} + \frac{7}{2}\right). \tag{14}$$

At temperatures in the LHe interval (1.7–4.2 K), the value of $H_a \approx 15$ kOe (see Fig. 16b) remains virtually constant within a broad interval of magnetic fields ($H \le$ 70 kOe).

Since the sum of coefficients A_1, A_3, and A_5 given by formulas (13) appears as

$$\begin{aligned} &A_1 + A_3 + A_5 \\ &= 4\pi\frac{R_\mathrm{H}}{d}\left(\frac{\sigma_n}{\sigma}\right)^2 M_0^* - \frac{R_\mathrm{H}}{d}\frac{\sigma_p - \sigma_n}{\sigma}H, \end{aligned} \tag{15}$$

the procedure of separating the normal and anomalous components in the Hall resistance within the framework of the proposed model reduces to determining a constant term and a coefficient at the linear term in the expression $A_\Sigma = A_1 + A_3 + A_5 = f(H)$. The plot of parameter A as a function of the field strength is presented in Fig. 17. For comparison, Fig. 17 (curve *2*) shows the estimates of the total contribution to the Hall resistance obtained directly from the angular profiles of $R(\varphi)$. Parameter ΔR is calculated as the difference $\Delta R = (1/2)[R(0) - R(180°)]$ (see the notation in Fig. 15). In addition, Fig. 17 (curve *3*) shows a plot of the Hall resistance versus the field strength H obtained by measuring the output voltage on the Hall contacts in the course of the magnetic field sweep for two opposite orientations of the magnetic field vector $\mathbf{H} \parallel \mathbf{c}$.

As can readily be seen, the estimates of the normal and anomalous contributions to the Hall resistance of

FeSi obtained within the framework of our model using Eqs. (11)–(15) in the interval $H \leq 35$ kOe coincide (at a good precision) with the $R_{\rm H}^{a}H$ and $R_{\rm H}^{a}M_0^*$ values determined by the usual methods of separating the corresponding components in $R_{\Sigma}(H, T)$. Taking equal contributions in (15), we may roughly estimate the exchange field (to within a factor on the order of $(\sigma_p/\sigma_n)^2 - 1 \sim 1$:

$$H_0^* = 4\pi M_0^* \approx H(R_{\rm H} = 0) \approx 500 \text{ kOe}.$$

The estimate of H_0^* obtained for the ferromagnetic microdomains in FeSi is considerably smaller as compared to the typical values of exchange fields in ferromagnetic metals. For iron, estimates of the molecular field give the value H_0^* (Fe) $\approx 1.1 \times 10^4$ kOe [32]. A significant decrease in the H_0^* value for the ferromagnetic microdomains in FeSi (as compared to Fe) is quite natural, taking into account a relationship between the size of monodomain regions and the correlation radius of spin density fluctuations in these materials [6].

5. CONCLUSION

In concluding, it should be noted that the proposed approach to the interpretation of the low-temperature anomalies in the physical properties of iron monosilicide within the framework of the Hubbard model apparently provides for a certain level of understanding of the physical processes and phenomena taking place in this narrow-band-gap semiconductor featuring strong quasiparticle interactions. In the regime of strong Hubbard correlations, the spin-polaron description is quite effective both in elucidating the reasons for a considerable increase in the effective mass of charge carriers and for the DOS renormalization at the Fermi level in FeSi, on the one hand, and in determining the mechanisms of the formation of the weakly-magnetic ground state in iron monosilicide at LHe temperatures.

From the standpoint of the model adopted, the transition to a coherent regime of spin density fluctuations, involving the formation of magnetic microdomains with a size of ~10 Å on decreasing the temperature to $T < T_{\rm c} \approx 15$ K, is generally analogous to the superparamagnetic behavior of micromagnetic metal alloys (spin glasses). The appearance of localized magnetic moments of the ferromagnetic microdomains is apparently followed by the transition to the state with frozen moments at $T_{\rm m} \approx 7$ K. It should be emphasized that the study of coherent effects in the system of spin polarons with a low density ($\sim 10^{17}$ cm^{-3}) is possible only in high-quality single-crystal FeSi samples.

The quantitative estimates of the model parameters and the microscopic characteristics of spin polarons presented above are quite realistic. However, for elucidating the nature of the ground state formation in FeSi in more detail, it is necessary to measure the low-temperature quasioptical spectra of FeSi in the far IR range (≤ 50 cm^{-1}) in combination with high-precision measurements of the magnetic properties of single-crystal iron monosilicide at LHe and intermediate temperatures in strong magnetic fields (up to 500 kOe).

ACKNOWLEDGMENTS

The authors are grateful to A.A. Volkov, B.P. Gorshunov, and A.A. Pronin for their help in experimental measurements. The work was supported by the Russian Foundation for Basic Research (project nos. 98-02-17163 and 99-02-16067), and Ministry of Science of the Russian Federation (“Fundamental Spectroscopy” and “Microwave Physics” programs), by the INTAS Foundation (grant no. 96-451), and by the Copernicus EBR IC15 CT98 0812 international project.

REFERENCES

1. G. Aeppli and Z. Fisk, Comments Condens. Matter Phys. **16**, 155 (1992).
2. D. I. Khomskiĭ, Usp. Fiz. Nauk **129**, 443 (1979) [Sov. Phys. Usp. **22**, 879 (1979)].
3. Z. Fisk, J. L. Sarrao, J. D. Thompson, *et al.*, Physica B (Amsterdam) **206-207**, 798 (1995).
4. C. M. Varma, Phys. Rev. B **50**, 9952 (1994).
5. P. Wachter, in *Handbook on the Physics and Chemistry of Rare Earths*, Vol. 19: *Lanthanides/Actinides Physics*, Ed. by K. A. Gschneidner, Jr., L. Eyring, G. H. Lander, and G. R. Choppin (North-Holland, Amsterdam, 1994).
6. T. Moriya, *Spin Fluctuation in Itinerant Electron Magnetism* (Springer-Verlag, Heidelberg, 1985; Mir, Moscow, 1988).
7. N. E. Sluchanko, A. A. Volkov, V. V. Glushkov, *et al.*, Zh. Éksp. Teor. Fiz. **115**, 970 (1999) [JETP **88**, 533 (1999)].
8. N. E. Sluchanko, V. V. Glushkov, B. P. Gorshunov, *et al.*, Phys. Rev. B **61**, 9906 (2000).
9. K. A. Kikoin and A. S. Mishchenko, J. Phys.: Condens. Matter **7**, 307 (1995).
10. N. E. Sluchanko, V. V. Glushkov, S. V. Demishev, *et al.*, Pis’ma Zh. Éksp. Teor. Fiz. **68**, 774 (1998) [JETP Lett. **68**, 817 (1998)].
11. N. E. Sluchanko, V. V. Glushkov, S. V. Demishev, *et al.*, Europhys. Lett. **51**, 557 (2000).
12. S. Paschen, E. Felder, M. A. Chernikov, *et al.*, Phys. Rev. B **56**, 12916 (1997).
13. N. E. Sluchanko, V. V. Glushkov, S. V. Demishev, *et al.*, Physica B (Amsterdam) **284-288**, 1179 (2000).
14. P. A. Beck, Prog. Mater. Sci. **23**, 1 (1978).
15. J. C. Cooley, M. C. Aronson, A. Lacerda, *et al.*, Phys. Rev. B **52**, 7322 (1995).
16. E. Arushanov, M. Respaud, J. M. Broto, *et al.*, Phys. Rev. B **55**, 8056 (1997).
17. M. J. Rozenberg, G. Kotliar, and H. Kajueter, Phys. Rev. B **54**, 8452 (1996).
18. Ph. Nozieres, Eur. Phys. J. B **6**, 447 (1998).
19. G. Kotliar, Eur. Phys. J. B **11**, 27 (1999).
20. V. Jaccarino, G. K. Wertheim, J. H. Wernick, *et al.*, Phys. Rev. **160**, 476 (1967).

21. S. Takagi, H. Yasuoka, S. Ogawa, and J. H. Wernick, J. Phys. Soc. Jpn. **50**, 2539 (1981).
22. P. M. Chaikin, *Organic Superconductivity*, Ed. by V. Z. Kresin and W. A. Little (Plenum, New York, 1990).
23. N. F. Mott, *Metal-Insulator Transitions* (Taylor & Francis, London, 1974; Nauka, Moscow, 1979).
24. L. Matteiss and D. Hamann, Phys. Rev. B **47**, 13114 (1993).
25. V. V. Moshchalkov, Physica B (Amsterdam) **163**, 59 (1990).
26. M. Fath, J. Aarts, A. A. Menovsky, *et al.*, Phys. Rev. B **58**, 15483 (1998).
27. A. Damascelli, K. Schulte, D. van der Marel, *et al.*, Phys. Rev. B **55**, R4863 (1997).
28. Z. Schlesinger, Z. Fisk, Hai-Tao Zhang, *et al.*, Phys. Rev. Lett. **71**, 1748 (1993).
29. K. Tajima, Y. Endoh, J. E. Fischer, and G. Shirane, Phys. Rev. B **38**, 6954 (1988).
30. V. L. Bonch-Bruevich, I. P. Zvyagin, R. Kaiper, *et al.*, *Electron Theory of Disordered Semiconductors* (Nauka, Moscow, 1981).
31. S. Chikasumi, *The Physics of Ferromagnetism. Magnetic Characteristics and Engineering Applications* (Syokabo, Tokyo, 1984; Mir, Moscow, 1987), Part 2, translated from Japanese.
32. S. Chikasumi, *The Physics of Ferromagnetism* (Syokabo, Tokyo, 1984; Mir, Moscow, 1987), Part 1, translated from Japanese.

Translated by P. Pozdeev

Journal of Experimental and Theoretical Physics, Vol. 92, No. 2, 2001, pp. 326–337.
Translated from Zhurnal Éksperimental’noĭ i Teoreticheskoĭ Fiziki, Vol. 119, No. 2, 2001, pp. 374–387.
Original Russian Text Copyright © 2001 by Tarasenko.

SOLIDS
Electronic Properties

Local Geometry of the Isofrequency Phonon Surface and the Spectral Structure of Generalized Lamb Waves on a Bounded Ferroelastic Substance

S. V. Tarasenko

Donetsk Physicotechnical Institute, National Academy of Sciences of Ukraine, Donetsk, 340114 Ukraine
e-mail: tarasen@host.dipt.donetsk.ua

Received August 22, 2000

Abstract—The singularities in the spectrum of bulk acoustic phonons polarized in the sagittal plane are investigated on a plate made of a uniaxial ferroelastic material undergoing a proper ferroelastic transition from the paraelectric to the ferroelectric phase. The singularities are induced by anomalies in the reflection of this type of normal elastic vibrations at the crystal boundary. © *2001 MAIK “Nauka/Interperiodica”.*

1. INTRODUCTION

One of the most intriguing features of the interaction of an elastic wave polarized in the plane of incidence with the crystal surface is the possibility of the simultaneous excitation (under definite conditions) of four normal elastic waves characterized by the same excitation frequency ω and the same component of the wave vector $\mathbf{k}$ on the crystal surface [1]. Two of these waves belong to the spectrum of a quasi-transverse wave polarized in the plane of incidence (the acoustic birefringence effect without a change in the branch).

The theory of this effect developed by Balakirev and Gilinskiĭ [2] for a cubic crystal shows, among other things, that if the medium in which phonons are propagating possesses elastic anisotropy, the reflection and refraction of a normal elastic wave polarized in the plane of incidence from the crystal boundary are unambiguously determined by the shape of its surface of reciprocal phase velocities. For example, if the section of such an isofrequency surface by the plane of incidence of the wave is a convex curve, the reflection (refraction) of a quasi-longitudinal as well as a quasi-transverse elastic wave from the crystal boundary is qualitatively the same as in the case of an elastically isotropic medium. If, however, the anisotropy of elastic properties is high enough for the formation of segments with a negative Gaussian curvature on the section of the isofrequency surface by a quasi-transverse elastic wave (the surface of the wave vectors of a quasi-longitudinal wave is always convex), the birefringence effect may take place. In other words, two quasi-transverse elastic waves identically polarized in the plane of incidence can be formed. These waves possess the same frequency and the wave vector component in the plane of the interface between two media, but have different values of the wave vector component along the normal to the interface.

A similar effect can also be realized in piezoelectric crystals possessing a definite symmetry for a transverse elastic wave polarized at right angles to the plane of incidence. A comparison of these two cases shows that the necessary condition for the formation of such anomalies on the surface of the wave vectors of a quasi-transverse elastic wave is the presence of an accompanying quasi-longitudinal elastic surface wave that does not belong to natural vibrations of the system, but is formed only in the presence of a bulk elastic quasi-transverse wave incident on the crystal surface. In the case of piezoelectric crystals, the formation of the birefringence effect without a change in the branch for a shear-type elastic *SH* wave is associated with the presence of an accompanying surface vibration of the electrostatic or magnetostatic type. In all cases, the inclusion of the accompanying surface vibration is of paramount importance for an analysis of the interaction of an acoustic wave with the crystal surface. The presence of the surface vibration makes possible the formation of parabolic points (segments with zero curvature) as well as concave segments (possessing a negative Gaussian curvature) on the curve appearing as a result of the section of the refraction surface by the sagittal plane. From the viewpoint of the reflection (refraction) of a bulk elastic wave from the crystal boundary, the presence of such a segment may lead, for example, to the vanishing of the accompanying surface vibration and to the formation, along with the normal reflected wave, of an auxiliary bulk elastic wave with the same polarization (the effect of the multibeam reflection of waves without a change of the branch) [2].

Naturally, the local geometry of the wave vector surface of the type under investigation for normal bulk vibrations of an unbounded crystal must be manifested in the spectral structure for this type of normal bulk vibrations in a bounded crystal since the spatial distribution of the amplitude of the bulk vibrations is the result of the interference of the bulk waves incident on the sample surface and reflected from it.

However, the conditions required for the implementation of this effect (the formation of a segment with a negative Gaussian curvature on the surface of the wave vectors of a normal elastic wave) impose quite stringent constraints on the relative value of elastic interactions in a crystal. Consequently, we can expect that such conditions are observed primarily in the vicinity of nonisomorphic structural phase transitions of the soft mode type.

It is well known [3] that there exists a wide class of continuous structural phase transitions associated with a change in the crystal symmetry, for which one of nondiagonal (shear) components of the elastic deformation tensor u_{ik} can be chosen as the order parameter. In the model of an infinitely large crystal, this leads to the possibility of an abrupt deceleration (softening) of the phase velocity of the shear wave near the stability boundary of the given crystalline state provided that its wave number differs from zero only slightly, while the polarization and the direction of propagation are connected in a certain way with the symmetry properties of the order parameter [4–6]. In the vicinity of a shear-type structural phase transition in a finite sample, the interference effects for such a softening acoustic wave multiply reflected from the sample surface considerably modify the spectrum of surface as well as bulk elastic modes of the corresponding types [7–9]. Similar singularities in the phonon dynamics of a bounded crystal may also be observed in the course of a phase transition in polarized media if it is a proper ferroelastic transition. In other words, the (generally) multicomponent order parameter for such a transition is transformed according to the same representation as for a certain (generally) linear combination of the deformation tensor components u_{ik} [8, 10, 11]. A detailed analysis of the principal anomalies emerging in the phonon spectrum near the stability boundary of a given crystalline state is of not only academic, but also of practical importance since, first, it allows us to study the critical dynamics of the crystal using the well developed methods of acoustic and optical spectroscopy, and second, there exists a number of crystals for which a similar structure of the phonon spectrum can be formed even far away from the region of a structural phase transition in view of the strongly anisotropic nature of intermolecular interactions in the medium (quasi-low-dimensional crystals), or due to an artificially created additional translational symmetry (superlattices, etc.).

It is especially interesting to analyze the rearrangement in the vicinity of a continuous structural shear-type transition of the fraction of phonon vibrations for which the displacement vector $\mathbf{u}$ of the polarization lattice lies in the sagittal plane of the crystal. If the sample is a mechanically free plate, these waves are referred to as Lamb waves, and if the plate is a part of an acoustically continuous structure of the layer + half-space or half-space + layer + half-space type, the waves are called the generalized Lamb waves [1].

Most of the publications devoted to the effect of the anisotropy of elastic moduli on the dispersion properties and the propagation conditions for Lamb waves are associated with the study of a Rayleigh wave in a semibounded crystal. It is well known that as regards elastic vibrations of the plate, this corresponds to the short-range approximation for the spectrum of surface Lamb waves, which consists of two branches [13, 14]. The main effects obtained by using this model and associated with the influence of the elastic anisotropy on the spectral structure of a surface Rayleigh wave can be formulated as follows (we assume that the propagating elastic wave is not dipole-active, i.e., is not accompanied by a magnetostatic or electrostatic field):[1]

(1) the phase velocity decreases [14];

(2) the polarization tends to the transverse orientation [14];

(3) the penetration depth of the wave increases (the extent of wave localization near the crystal surface decreases) [14];

(4) a smooth transition from a generalized to a surface Rayleigh wave is possible [7, 8];

(5) a one-to-one correspondence exists between the parameters of the surface Rayleigh wave and the structure of the surface of reciprocal phase velocities for the corresponding type of normal elastic waves in an unbounded crystal [7, 8].

As regards the analysis of the effect of elastic anisotropy on the spectrum of Lamb waves in a plate in the long-wave limit, the main results obtained in this directions can be formulated as the following conclusions (for nondipole-active waves):

(1) all the modes belonging to the spectrum of quasi-transverse Lamb waves in the vicinity of a continuous shear-type structural phase transition are slowed down [7, 11];

(2) for a given value of the wave number and the ratio of elastic moduli in a mechanically free plate, one or two surface Lamb waves can propagate;

(3) as the wave number $k_\perp$ decreases, the dispersion curve for a symmetric surface Lamb wave may smoothly go over to the dispersion curve of a bulk longitudinal Lamb wave with $\nu = 1$ for $k_\perp \neq 0$;

(4) for a certain $\nu = 1$, the point of degeneracy of a Lamb wave and a transverse elastic wave of the *SH* type can be formed.

However, till now

(1) the lower (longitudinal and bending) modes of the Lamb spectrum of an anisotropic plate were mainly considered;

---

[1] This can be observed not only in nonpolar media, but also, for example, for one-component dipole-active ferroelastic phase transitions for a definite orientation of the plane of propagation of an elastic wave.

(2) the effects associated with the influence of spatial dispersion, which are known to affect significantly the critical dynamics of the order parameter under phase transitions of the soft mode type were ignored (the class of continuous shear-type structural phase transitions are just an example of such transitions);

(3) only anisotropic plates with mechanically free boundary conditions were considered in spite of the fact that an analysis of the peculiarities in the rearrangement of the spectrum of generalized Lamb waves is undoubtedly important in connection with considerable advances in the physics of composite materials;

(4) the relation between the spectral structure of bulk Lamb waves and peculiarities of reflection of normal elastic waves polarized in the plane of incidence at the sample boundary was not investigated. At the same time, this type of elastic vibration of an anisotropic plate is the result of the interference of quasi-transverse and quasi-longitudinal phonons which can be transformed into one another as a result of multiple reflections from the plate boundaries. Both components of the lattice displacement vector **u** lie in the wave incidence plane and have a multipartite structure.

An additional argument in favor of a considerable effect of the configuration of the refraction surface on the spectral structure of bulk elastic vibrations is also the fact that the presence of a segment with a negative Gaussian curvature on the surface is a sufficient condition for the formation of a semi-bounded hexagonal crystal of a generalized Rayleigh wave for a mechanically free surface [7, 8]. At the same time, it was shown [12, 13] that the dispersion relation for the symmetric branch in the spectrum of surface Rayleigh waves (in the short-wave limit) for a plate made of the same crystal smoothly goes over (for $k_\perp \neq 0$) into the dispersion curve for a bulk longitudinal Lamb wave upon a decrease in the wave number $k_\perp$.

However, the relation between the local geometry of the isofrequency surface for the normal elastic vibrations of an unbounded crystal and the spectral structure of Lamb phonons in a crystal plate undergoing a continuous shear-type structural phase transition has not been investigated as yet.

In connection with what has been said above, the present work aims at determining the relation between the configuration of the wave vector surface for normal elastic vibrations polarized in the plane of incidence in an unbounded crystal and the anomalies in the spectrum of bulk Lamb phonons, using as an example a one-component proper ferroelastic phase transition in a ferroelectric plate without a center of symmetry (i.e., displaying the piezoelectric effect) in the paraelectric phase [3]. The subject matter of the paper is as follows. Section 2 contains basic relations as well as the formulation of the boundary-value ferroelastic problem. An analysis of the spectral structure (in the absence of spatial dispersion) of bulk acoustic waves polarized in the sagittal plane of a crystalline ferroelectric film undergoing a structural phase transition is presented in Section 3. The results of the analysis of the shape of the section of the isofrequency surface for normal elastic vibrations in the ferroelastic material under investigation by the sagittal plane and the relation between this shape and the observed anomalies in the phonon spectrum are considered in Section 4. In Section 5, the singularities in the spectrum of generalized Lamb waves induced by the presence of a rhombic anisotropy in the sagittal plane are analyzed. Additional anomalies in the phonon spectrum associated with correlation effects in the ferroelectric subsystem of the crystal (spatial dispersion), which must be taken into account in the vicinity of the phase transition, are considered in Section 6. The main results obtained in this work are formulated in the Conclusion.

2. BASIC RELATIONS

Following [4], we consider the transition from the paraelectric to the ferroelectric state, which takes place, for example, in KDP crystals at $T \longrightarrow T_C$ (T_C is the Curie temperature) as an example of a one-component ferroelastic phase transition. The structure of the corresponding thermodynamic potential can be presented in the form [3, 14]

$$W = \frac{\delta}{2}(\nabla P_z)^2 + \frac{b}{2}P_z^2 + \gamma P_z u_{xy} + \frac{1}{2}c_{11}(u_{xx}^2 + u_{yy}^2) + \frac{1}{2}c_{33}u_{zz}^2 + c_{12}u_{xx}u_{yy} + c_{13}(u_{xx}u_{zz} + u_{zz}u_{yy}) + 2c_{44}(u_{zx}^2 + u_{zy}^2) + 2c_{66}u_{xy}^2. \quad (1)$$

Here, δ, $b > 0$, and γ are the constants characterizing spatial dispersion, anisotropy, and the piezoelectric interaction, respectively, and c_{ik} are the elastic interaction constants.

Since we are interested here in the nondipole-active dynamics of the model of the ferroelectric crystal under investigation, we must put, in accordance with (1), $\mathbf{k} \in xy$ ($k_z = 0$). In this case, the corresponding closed system of equations connecting only the polarization vector P_z and the vector **u** of elastic displacements of the lattice [15] has the form

$$\frac{\delta W}{\delta P_z} = f\frac{\partial^2 P_z}{\partial t^2}, \quad \rho\frac{\partial^2 u_i}{\partial t^2} = \frac{\partial^2 W}{\partial x_k \partial u_{ik}} \quad (2)$$

(ρ is the density). In the case of a bounded ferroelectric plate, the system of dynamic equations (2) must be supplemented with the required boundary conditions. The boundary condition for the polarization vector **P** is chosen in the form of the relation [15]

$$\delta\frac{\partial P_z}{\partial \zeta} + aP_z = 0, \quad \zeta = \pm d, \quad (3)$$

corresponding to a partial ($a \neq 0$) pinning of the polarization vector on both surfaces of the ferroelectric film under investigation. As regards the elastic boundary conditions, we will henceforth assume that the relations corresponding to a boundary with tangential slip hold on both surfaces of the plate [1]:

$$\varepsilon_{ijk}\frac{\partial W}{\partial u_{ik}}n_k n_l = 0, \quad \mathbf{u}\cdot\mathbf{n} = 0 \tag{4}$$

($\hat{\varepsilon}$ is a unit antisymmetric tensor). It is well known that this type of boundary conditions is realized on the interface between two solids with completely incoherent conjugation [16]. In this case, we assume in (4) that the elastic medium bordering the plate along the normal to the interface can be regarded as an absolutely rigid medium ($\mathbf{u}\cdot\mathbf{n} = 0$). Thus, we are dealing here with a layered structure of the half-space + layer + half-space type.

Calculations show that the spectrum of coupled ferroelastic vibrations in the model of the unbounded ferroelectric material under investigation with $\mathbf{u}, \mathbf{k} \perp [001]$ and $P_z \neq 0$ in the paraelectric tetragonal phase can be presented in the form

$$(\Lambda_{11} - \omega^2)(\Lambda_{22} - \omega^2) = \Lambda_{12}^2 \tag{5}$$

($\hat{\Lambda}$ is the Christoffel tensor). While deriving relation (5), we assumed that in the corresponding components of the Christoffel tensor, the elastic modulus c_{66} is multiplied by the parameter

$$\theta = \frac{\omega_0^2 + c^2k^2 - \omega^2}{\omega_0^2 + \omega_{pe}^2 + c^2k^2 - \omega^2}.$$

Here, ω_{pe} is the ferroelastic gap, ω_0 is the activation of the spectrum of polarization-induced vibrations, which is associated with the uniaxial anisotropy b, $c^2 = \delta/f$, and $k^2 = k_x^2 + k_y^2$. It can easily be proved that if we disregard the piezoelectric interaction ($\gamma \longrightarrow 0$), Eqs. (5) can be factorized, and the obtained dispersion relations describe two physically different sets of natural vibrations of the dynamic system under investigation, each of which describes, respectively, the spectrum of phonons in a nonpolar tetragonal or cubic crystal, which are polarized in the plane of incidence, and the spectrum of bulk vibrations of the z component of the electric polarization vector $\mathbf{P}$ disregarding the effect of the lattice. For $\gamma \neq 0$, expression (5) makes it possible to determine the relation between the wave vector component normal to the film surface, ($\mathbf{n} \perp [001]$), characterized by frequency ω and the wave number $k_\perp$ of a propagating bulk elastic wave; i.e., it can be regarded as the characteristic equation for the solution of the boundary-value problem for $\mathbf{n}, \mathbf{u}, \mathbf{k} \in xy$. It follows from relation (5) that the structure of the component of the lattice elastic displacement vector $\mathbf{u}$ normal to $\mathbf{n}$ for $\mathbf{n} \parallel [100]$ as well as $\mathbf{n} \parallel [110]$ can be presented in the form

$$u = \sum_{\nu=1}^{3} A_\nu \exp(-q_\nu\zeta)\exp(i\omega t - ik_\perp\tau) \tag{6}$$

(τ is the running coordinate along the direction of propagation of the wave $\boldsymbol{\tau} \perp \mathbf{n}$; $q^2 \equiv -(\mathbf{k}\cdot\mathbf{n})^2$). Using relations (5) and (6), we can classify the possible types of propagating elastic waves depending on their localization near the surface of a magnet ($\zeta \geq 0$) defined by $q_{1,2}$. In order to simplify subsequent calculations, we will formulate a number of simplifying assumptions complying with the requirements of the present work.

(1) Since the softening of an acoustic surface wave is known to occur in the long-wave region of the spectrum of ferroelastic vibrations, we will henceforth assume that the frequency and the wave number of the phonons under investigation are such that the following relation holds to a high degree of accuracy:

$$\omega^2 \ll \omega_{pe}^2 + \omega_0^2 \tag{7}$$

(the dynamics of the polarization subsystem of the crystal will be analyzed in the quasi-static limit). As a result, the boundary-value problem for the z component of the electric polarization vector $\mathbf{P}$ assumes the form

$$\begin{gathered} c^2\left(\frac{\partial^2}{\partial\zeta^2} + \frac{\partial^2}{\partial\tau^2}\right)P_z - (\omega_0^2 + \omega_{pe}^2)P_z = \gamma f^{-1}u_{xy}, \\ \delta\frac{\partial P_z}{\partial\zeta} + aP_z = 0, \quad \zeta = \pm d. \end{gathered} \tag{8}$$

If we use the apparatus of Fredholm's equations, we can use Green's function $G(\zeta, t)$ [17] of the form

$$G(t, \zeta) = \begin{cases} \Delta^{-1}\sinh(p(t-\zeta_-))\sinh(p(t-\zeta_+)), & -d \leq \zeta < t, \\ \Delta^{-1}\sinh(p(t-\zeta_+))\sinh(p(t-\zeta_-)), & t \leq \zeta < d, \end{cases}$$

$$\Delta \equiv p\sinh(p(t-\zeta_+)), \quad p^2 \equiv \omega_0^2 + \omega_{pe}^2 + c^2k_\perp^2, \tag{9}$$

$$a\delta^{-1} \equiv p\coth(p(\zeta_+ - d)), \quad a\delta^{-1} = p\coth(p(\zeta_- + d))$$

to eliminate $\mathbf{P}_z$ from the equations for the lattice displacement vector $\mathbf{u}$ both for $\mathbf{n} \parallel [100]$ and for $\mathbf{n} \parallel [110]$. As a result, the boundary-value problem under investigation can be reduced to an analysis of only the elastic boundary-value problem for the components of vector $\mathbf{u}$.

(2) Since the analysis of the conditions for the reflection (refraction) of elastic waves polarized in the plane of incidence shows that a change in the local curvature of the reciprocal phase velocity surface is possible only for the quasi-transverse branch of the spectrum of normal elastic vibrations of an unbounded crystal, we will assume that the velocities s_l and s_t of quasi-lon-

gitudinal and quasi-transverse acoustic waves are connected through the condition

$$s_t/s_l \ll 1. \qquad (10)$$

Subsequently, we will confine our analysis to the singularities in the phonon spectrum of a ferroelectric plate taking into account only the quasi-transverse branch in the spectrum of normal elastic vibrations of phonons polarized in the plane of incidence ($\mathbf{k}^2 = k_x^2 + k_y^2$, $\tan\vartheta = k_x/k_y$). Thus, the corresponding dispersion relation for the spectrum of quasi-transverse phonons with $\mathbf{u} \in xy$ in an unbounded crystal can be presented in the form

$$\omega^2 \approx s_0^2 k^2 (1 + r\sin^2 2\vartheta), \quad r \equiv \frac{(1-\eta)(\eta-\xi)}{\eta^2}, \qquad (11)$$

where $\eta \equiv 2c_{66}\theta/(c_{11} - c_{12})$, $\xi \equiv c_{66}\theta/c_{11}$, and $s_0^2 = c_{66}\theta/\rho$ for $\mathbf{n} \parallel [100]$. On account of the above approximations, relations (1), (2), and (6) show that the following types of propagating normal elastic vibrations ($c \longrightarrow 0$) are possible in the frequency range (11) depending on the magnitude and sign of parameters η and r both for $\mathbf{n} \parallel [100]$ and for $\mathbf{n} \parallel [110]$.

I. Bulk wave of the first type ($q_1^2 > 0$; $q_2^2 < 0$):

$$\begin{gathered} \omega^2 > s_0^2 k_\perp^2, \quad \mathbf{n} \parallel [100], \\ \omega^2 > s_0^2 (1+r) k_\perp^2, \quad \mathbf{n} \parallel [110]. \end{gathered} \qquad (12)$$

II. Generalized surface waves ($q_1^2 = (q_2^2)^*$):

$$\begin{gathered} \omega^2 < \omega_+^2 \equiv 4 s_t^2 k_\perp^2 [r^{1/2}(1+r)^{1/2} - r], \\ r > 0, \quad \mathbf{n} \parallel [110], \\ \omega^2 < \omega_+^2 \equiv 4 s_t^2 k_\perp^2 [|r|^{1/2} - |r|], \\ r < 0, \quad \mathbf{n} \parallel [100]. \end{gathered} \qquad (13)$$

III. Surface waves ($q_{1,2}^2 > 0$):

$$\begin{gathered} \omega_+^2 < \omega^2 < s_t^2 k_\perp^2, \quad r < 0, \quad |r| < 1/4, \quad \mathbf{n} \parallel [110], \\ \omega_+^2 < \omega^2 < s_t^2 (1+r) k_\perp^2, \\ 0 < r < 1/3, \quad \mathbf{n} \parallel [100]. \end{gathered} \qquad (14)$$

IV. Bulk waves of the second type ($q_{1,2}^2 < 0$):

$$\begin{gathered} \omega_+^2 < \omega^2 < s_t^2 k_\perp^2, \quad |r| < 1/4, \quad r < 0, \quad \mathbf{n} \parallel [100], \\ \omega_+^2 < \omega^2 < s_t^2 (1+r) k_\perp^2, \quad r > 1/3, \quad \mathbf{n} \parallel [110]. \end{gathered} \qquad (15)$$

If, however, $r > 0$ for $\mathbf{n} \parallel [100]$ or $r < 0$ for $\mathbf{n} \parallel [110]$, an analysis shows that all the relations obtained above remain valid except that now we must put $\omega_+^2 \equiv 0$ for any $k_\perp$. In other words, the formation of generalized surface waves with $\mathbf{k} \in xy$ is impossible for $\mathbf{n} \parallel [100]$ as well as for $\mathbf{n} \parallel [110]$. Thus, the necessary condition for the propagation of bulk phonons corresponding to the quasi-transverse branch in the spectrum of normal elastic waves along the ferroelectric plate under investigation is that their frequency ω and the wave number $k_\perp$ must lie in region I or IV. In this case, a tripartite elastic wave having the polarization in question and propagating along the plate has at least one bulk component.

Multiple reflection of this type of phonons from the plate boundary and their interference lead to the formation of normal elastic modes polarized in the sagittal plane of the ferroelectric film under investigation (Lamb waves). The next section will be devoted to an analysis of the main singularities in the phonon spectrum of a plate made of a uniaxial ferroelectric material in the paraelectric phase, undergoing the one-component ferroelastic phase transition $42m \longrightarrow mm2$ disregarding the correlation effects (we assume that $\delta \longrightarrow 0$ and $\theta = \omega_0^2/(\omega_0^2 + \omega_{pe}^2)$).

3. ELASTIC VIBRATIONS POLARIZED IN THE SAGITTAL PLANE

It follows from the boundary-value problem (4) that the spectrum of bulk ferroelastic vibrations with $\mathbf{k} \in xy$ ($\mathbf{u} \perp [001]$) can be presented in this case in the form

$$\begin{gathered} \Omega_\nu^2(k_\perp) = \frac{N_1}{2} \pm \left(\left(\frac{N_1}{2}\right)^2 - N_2\right)^{1/2}, \\ N_1(k_\perp) = \Lambda_{11}^* + \Lambda_{22}^*, \\ N_2(k_\perp) = \Lambda_{11}^*\Lambda_{22}^* - (\Lambda_{12}^*)^2, \\ \Lambda_{ik}^* \equiv \Lambda_{ik}, \quad k_x = m_\nu, \quad k_y = k_\perp, \quad k_z = 0, \end{gathered} \qquad (16)$$

where Λ_{ik} is the Christoffel tensor, $\mathbf{n} \parallel [100]$, $m_\nu = \pi\nu/2d$, $\nu = 1, 2, \ldots$ Since the corresponding characteristic equation (4) is biquadratic in $\mathbf{k} \cdot \mathbf{n}$, we will simplify our analytical calculations by assuming henceforth that the following conditions are satisfied:

$$c_{11}, c_{12} \gg c_{66}. \qquad (17)$$

In this limit, relations (11) and (16) can be presented in the form [2]

$$\begin{gathered} \omega^2 \approx s_0^2 k^2 (1 + r\cos^2 2\vartheta), \\ \Omega_\nu^2(k_\perp) \approx s_0^2 \left(k_\perp^2 + m_\nu^2 + r\frac{(k_\perp^2 - m_\nu^2)^2}{k_\perp^2 + m_\nu^2}\right), \end{gathered} \qquad (18)$$

for $\mathbf{n} \parallel [110]$ and

$$\begin{gathered} \omega^2 \approx s_0^2 k^2 (1 + r\sin^2 2\vartheta), \\ \Omega_\nu^2(k_\perp) \approx s_0^2 \left(k_\perp^2 + m_\nu^2 + 4r\frac{k_\perp^2 m_\nu^2}{k_\perp^2 + m_\nu^2}\right). \end{gathered} \qquad (19)$$

for $\mathbf{n} \parallel [100]$. An analysis of relations (18) and (19) shows that the shape of the dispersion curve of the bulk mode with number ν in a crystal with a preset ratio of elastic moduli (disregarding the piezoelectric interaction) changes qualitatively depending on the orientation of normal **n** to the film surface (**n** || [100] or **n** || [110]) in the plane (001) of wave propagation. An analysis shows, in particular, that the phonon spectrum (18), (19) undergoes the most significant transformation when the following condition is satisfied for the elastic moduli of the crystal even if we disregard the piezoelectric interaction ($\gamma = 0$):

$$\eta > 1, \quad r < 0 \tag{20}$$

(the condition $\xi < 1$ is observed everywhere). In this case, an analysis of relations (18) and (19) shows that the dispersion curves describing the spectrum of bulk quasi-transverse elastic vibrations for **n** || [110] correspond to forward-type waves ($\partial\Omega_\nu/\partial k_\perp > 0$) irrespective of the mode number ν and the magnitude of the wave number $k_\perp$ for $|r| < 1/3$. In this case, two points of inflection $k_{\nu 1,2} \neq 0$, which are real-valued roots of the equation

$$\partial\Omega_\nu^2/\partial k_\perp^2 = 0,$$

may appear on these curves. However, as we approach the region of a ferroelastic phase transition, the points of inflection on the dispersion curve with a fixed value of the mode number ν start converging and merge into one point for $\eta = 1$ ($r = 0$). As the crystal approaches the phase transition point (the value of $|r|$ increases, $r > 0$), the dispersion curve under investigation becomes more and more gently sloping and may ultimately acquire a minimum for $r > 1/3$ and for $k_\perp = k_{*\nu}$, where

$$k_{*\nu}^2 = m_\nu^2[(4r/(1+r))^2 - 1]. \tag{21}$$

As a result, for $k_\perp > k_{*\nu}$, the corresponding segment of the dispersion curve (18) describes a forward-type wave ($\partial\Omega_\nu/\partial k_\perp > 0$), while for $k_\perp < k_{*\nu}$, a backward-type wave ($\partial\Omega_\nu/\partial k_\perp < 0$) propagates in the film under investigation. Moreover, for $r > 1/3$, a point of crossover (intersection) of the corresponding dispersion curves may appear for the modes with numbers ν and ρ of spectrum (18) for $k_\perp = k_{\nu\rho} \neq 0$:

$$\Omega_\nu(k_{\nu\rho}) = \Omega_\rho(k_{\nu\rho}).$$

In the case when the condition $\eta > 1$ ($r < 0$) is satisfied in a film with the normal **n** || [100], the transformation of spectrum (19) for the bulk phonon type under investigation is qualitatively different. An analysis shows that in contrast to the case (18) with **n** || [110] considered above, the long-wave asymptotic form of the dispersion curve describing a mode with a fixed ν for the given geometry of propagation of an elastic wave may depend considerably on the value of $|r|$ even for $k_\perp \longrightarrow 0$ and far away from the structural phase transition region. For example, a backward-type wave ($\partial\Omega_\nu/\partial k_\perp < 0$) is formed irrespective of the mode number ν for $|r| > 1/4$, while for $|r| < 1/4$, the corresponding dispersion curve describes a forward-type wave for any value of the wave number $k_\perp$. As the regards the shape of the dispersion curve for a mode with number ν for $|r| > 1/4$, calculations show that for $k_\perp = k_{*\nu}$, where

$$k_{*\nu}^2 = m_\nu^2(2|r|^{1/2} - 1), \tag{22}$$

a minimum is formed on the curve. Besides, for modes with numbers ν and ρ ($\nu \neq \rho$), a point of degeneracy of the dispersion curves $\Omega_\nu(k_\perp)$ and $\Omega_\rho(k_\perp)$ belonging to spectrum (19) can be formed for $k_\perp = k_{\nu\rho}^{\pm}$.

As the crystal under investigation approaches the point of the ferroelastic phase transition, $|r|$ starts decreasing, quantity r remaining negative ($r \longrightarrow -1/4$, $r > -1/4$). As a result, in accordance with (19), the wave numbers $k_{*\nu}$ and $k_{\nu\rho}$ corresponding respectively to the point of the minimum on the dispersion curve with number ν and the point of degeneracy of the dispersion curves $\Omega_\nu(k_\perp)$ and $\Omega_\rho(k_\perp)$ belonging to the spectrum decrease in absolute value. As we approach the point of the shear-type structural phase transition under investigation, the absolute value of r becomes smaller and smaller, and for any r from the region $-1/4 < r$, all the dispersion curves described by relations (19) and belonging to the spectrum of Lamb waves of the type in question are forward-type wave ($\partial\Omega_\nu/\partial k_\perp > 0$) irrespective of the mode number ν and the wave number $k_\perp$. For any mode number ν and ρ, the following inequality is observed:

$$\Omega_\nu(k_\perp) \neq \Omega_\rho(k_\perp).$$

An analysis of relations (18) and (19) shows that the rearrangement of the spectrum for the bulk phonons under investigation in the vicinity of the continuous shear-type structural phase transition also takes place when the crystal has the properties of an elastically isotropic medium ($\eta = 1$ or $r = 0$) for $\gamma = 0$. In this case, in accordance with (18) and (19), the shape of the dispersion curves belonging to the spectrum of acoustic phonons of the given type is determined by the inequalities $\partial\Omega_\nu/\partial k_\perp > 0$ and $\partial^2\Omega_\nu/\partial k_\perp^2 > 0$ both for **n** || [100] and for **n** || [110] irrespective of the value of $k_\perp$ and the mode number ν ($\partial\Omega_\nu/\partial k_\perp \longrightarrow 0$ for $k_\perp \longrightarrow 0$). The existence of the piezoelectric interaction in the crystal ($\gamma \neq 0$ in (1)) leads to different transformations of the spectrum of Lamb phonons of the type (18), (19) under investigation for different orientations of **n** in the (001) plane in the vicinity of the ferroelastic phase transition in question ($T \longrightarrow T_C$). For example, for **n** || [100], the shape of the dispersion curve of the mode $\Omega_\nu(k_\perp)$ with a fixed ν is the same as for $\gamma = 0$: it corresponds to a forward-type wave $\partial\Omega_\nu/\partial k_\perp > 0$. However, as the value of r increases, two points of inflection ($\partial^2\Omega_\nu/\partial k_\perp^2 > 0$) can be formed for $k_{*\nu} \neq 0$. On the other hand, if the normal **n** to the surface of an elastically isotropic (for $\gamma = 0$) plate coincides with the direction [110] for $\gamma \neq 0$, an analysis of relations (18) shows that, first, for $r > 1/3$, a minimum may appear on the dispersion curve from (18) with number ν for $k_\perp \neq 0$ and, second, a point of crossover for the dispersion curves with numbers ν and ρ belonging to spectrum (18) can be formed for $k_{\nu\rho} \neq 0$ [$k_{\nu\rho}$ is the real-valued root of the equation $\Omega_\nu(k_\perp) = \Omega_\rho(k_\perp)$]. For $r > 0$, the pattern of rearrangement of the spectrum of bulk quasi-transverse phonons polarized in the sagittal plane of an anisotropic plate with **n** || [100] or **n** || [110] may also be qualitatively similar to that considered above for $r = 0$ and $\gamma = 0$. To this end, the condition $r < 1/3$ for **n** || [110] or $r < 1/4$ for **n** || [100] must be observed in the limit $\gamma = 0$.

Finally, an analysis of relations (18) and (19) shows that if the elastic properties of the crystal type under investigation satisfy the conditions $r > 1/3$ for **n** || [110] or $r < 1/4$ for **n** || [100] even in the absence of a piezoelectric interaction ($\gamma = 0$), the shape of the dispersion curves of modes (18) and (19) describing the spectrum of bulk quasi-transverse elastic vibrations does not display any qualitative changes as we approach the stability boundary for the given crystalline state irrespective of the direction of normal **n** in the sagittal plane of the film (**n** || [100] or **n** || [110]). In particular, for **n** || [100], the dispersion curves of spectrum (19) correspond to forward-type waves ($\partial\Omega_\nu/\partial k_\perp > 0$) for arbitrary ν and $k_\perp$. If the relation $\partial\Omega_\nu^2/\partial k_\perp^2 = 0$ holds for $k_\perp \neq 0$, two points of inflection can be formed on the corresponding curve describing the mode with a given number ν. For **n** || [110], the dispersion curve describing the mode with a number ν from spectrum (18) has a minimum for $k_\perp \neq 0$. In other words, it describes a backward-type wave ($\partial\Omega_\nu/\partial k_\perp < 0$) in the long-wave limit $k_\perp \longrightarrow 0$. Besides, if the condition $\Omega_\nu(k_\perp) = \Omega_\rho(k_\perp)$ is satisfied for the modes with numbers ν and ρ, a point of degeneracy of the phonon spectrum under investigation will be formed.

The above analysis shows that the physical reason behind the presence of a minimum on a dispersion curve belonging to the spectrum of bulk Lamb phonons of the type under investigation is the presence of an accompanying elastic vibration of the quasi-longitudinal type in the given region of frequencies and wave numbers. If this accompanying vibration is transformed into an auxiliary bulk wave for certain values of ω and $k_\perp$, the formation of a crossover point on dispersion curves belonging to the spectrum of the Lamb phonons in question becomes possible for $k_\perp \neq 0$.

Since the presence of an accompanying surface vibration in a crystal is essentially determined by the configuration of the refraction surface for the corresponding normal elastic wave in an unbounded crystal [1, 2] (see Introduction), we devote the next section to an analysis of the relation between the local geometry of the refraction surface for a normal quasi-transverse elastic wave in an unbounded cubic crystal, which is polarized in the plane of incidence, and the anomalies in the spectrum of bulk Lamb phonons determined above for a ferroelastic plate in the paraphase with **n** || [100] or **n** || [110] and $\mathbf{k} \in xy$.

4. RELATION WITH THE SHAPE OF THE REFRACTION SURFACE

Since the wave vector of the wave described by relations (18) and (19) lies in the (001) plane, we must analyze, using relation (11), the shape of the cross section of the surface of reciprocal phase velocities of the quasi-transverse ($\mathbf{u} \perp [001]$) elastic wave ($\omega = \text{const}$) by the (001) plane in the k space. Taking into account approximations (7) and (11), we can write the corresponding expression in the form

$$\mathbf{k}^2 = \frac{\omega^2}{s_0^2(1 + r\sin^2 2\vartheta)} \tag{23}$$

($k_x^2/\mathbf{k}^2 \equiv \sin^2\vartheta$, $\mathbf{k}^2 \equiv k_x^2 + k_y^2$, $\delta \longrightarrow 0$). An analysis of the extremum points of the curve described by (23) [1] and their comparison with the singularities in the shape of the dispersion curves (18) and (19) discussed above show that the presence of a local minimum on the dispersion curve of the waveguide phonon (18), (19) under investigation is connected with the formation of a segment with the maximum negative curvature (at $\vartheta = \pi/4$ for $r > 1/3$ and at $\vartheta = 0$ for $r < -1/4$) on the corresponding cross section of the refraction surface for a normal quasi-transverse wave of the same polarization (23) in an unbounded crystal. The position of this segment on curve (23) in the k space is determined by the condition $\vartheta = 0$ and is unambiguously connected with frequency ω, the mode number ν, the film thickness $2d$, and the wave number $k_\perp$ of the waveguide phonon (18), (19) under investigation.

If we consider the section of curve (23) by the straight lines determined by the conditions $k_x = \text{const}$ and $k_y = \text{const}$, an analysis of the common points of such a straight line and the refraction surface for (23) makes it possible to obtain information on the spectral structure of the corresponding waveguide phonon for a given wave number $k_\perp$, frequency ω, as well as a mode number ν (curves (18) and (19) in the present case). In particular, if the direction of normal **n** to the film surface coincides with the ordinate axis in the plane of the wave vectors k_x, k_y (**n** || [100]), the number of common points of the straight line $k_z = k_\perp$ and curve (23) determines the numbers ν of modes with the same wave number $k_\perp$ and frequency ω in the spectrum of the bulk Lamb phonons under investigation, which can propagate along the [010] axis of the given crystalline plane of thickness $2d$ (i.e., crossover points). In the same geometry, the presence of common points of curve (23) and the straight line $k_x = m_\nu$ allows us to determine the wave numbers $k_\perp$ for the given type of a Lamb waveguide phonon with a fixed mode number ν and frequency ω, which can propagate along the plate having a thickness $2d$ and made of the ferroelectric under investigation. Since the outward normal to the refraction surface coincides with the direction of the group velocity of the wave [1], the joint analysis of relations (18), (19), and (23) shows that the study of the local geometry of the cross section of the isofrequency surface (23) makes it possible to judge which type of the wave (forward or backward) corresponds to the relevant segment on the dispersion curve of a waveguide phonon, which can be determined from (18) and (19) by specifying ω, m_ν, and $k_\perp$. For example, in the case under consideration, when $\mathbf{k} \parallel [010]$ ($\mathbf{n} \parallel [100]$), the bulk Lamb wave (19) propagating along the film is a backward wave if the projection of the outward normal to the refraction surface on the direction [100] has negative sign at the point of intersection of this surface with the straight line $k_y = m_\nu$. If, however, the projection is positive, the corresponding wave for given k, ω, and m_ν is a forward wave. If the component along the x axis is equal to zero for a certain $k_\perp \neq 0$, such a situation is possible when the dispersion curve for the mode having number ν and belonging to the spectrum of bulk vibrations propagating along the surface of the film of thickness $2d$ ($\mathbf{n} \parallel [100]$) has an extremum for this value of the wave number $k_\perp$. The sign of the local Gaussian curvature of curve (23) at this point determines whether the point is a maximum or a minimum.

Till now, we carried out the analysis for a crystalline plate for which the elastic properties along the normal to the surface and along the direction of propagation of an elastic wave with a preset polarization were identical. In this connection, it would be interesting to find out how the elastic anisotropy in the sagittal plane affects the above necessary conditions for the existence of the point of crossover $k_{\nu\rho}$ as well as the point of minimum $k_{*\nu}$ (the formation of a backward wave $\partial\Omega_\nu(k_\perp)/\partial k_\perp < 0$) on the dispersion curve for the mode $\Omega_\nu(k_\perp)$ belonging to the spectrum of bulk Lamb phonons in an anisotropic plate. For this purpose, we will consider in the next section the effect of rhombic anisotropy in the sagittal plane of an anisotropic plate on the conditions of the formation of the above anomalies in the spectrum of the bulk Lamb phonons of the type under investigation. By way of an example, we will analyze a one-component ferroelastic phase transition from the paraphase to the ferroelectric state for a crystalline plate displaying the piezoelectric effect $222 \longrightarrow 2$ in the paraphase (Rochelle salt) [3].

5. EFFECTS OF RHOMBIC ANISOTROPY

For the convenience of comparison with the results of the above calculations, we assume that the z axis ([001]) is the polar axis for the one-component ferroelastic phase transition. In this case, the thermodynamic potential density corresponding to the given model can be presented in the form

$$\begin{aligned} W = {} & \frac{\delta}{2}(\nabla P)^2 + \frac{b}{2}P_z^2 + \gamma P_z u_{xy} + \frac{1}{2}c_{11}u_{xx}^2 \\ & + \frac{1}{2}c_{22}u_{yy}^2 + \frac{1}{2}c_{33}u_{zz}^2 + c_{12}u_{xx}u_{yy} + c_{13}u_{xx}u_{zz} \\ & + c_{23}u_{zz}u_{yy} + 2c_{55}u_{zx}^2 + 2c_{44}u_{zy}^2 + 2c_{66}u_{xy}^2. \end{aligned} \tag{24}$$

As before, we assume that vectors $\mathbf{k}$ and $\mathbf{n}$ lie in the (001) plane and confine our analysis to the spectrum of bulk elastic nonpiezoactive vibrations localized in the sagittal plane and propagating along the film. We will consider only the cases when $\mathbf{n} \parallel [100]$ and $\mathbf{n} \parallel [010]$ provided that relations (3) and (4) are satisfied on the surfaces of the plate.

Calculations show that in limit (7), the spectrum of coupled ferroelastic vibrations with $\mathbf{u}, \mathbf{k} \perp [001]$ and $P_z \neq 0$ in the paraelectric tetragonal phase in the model of an unbounded ferroelectric medium can be presented in the form

$$\begin{aligned} \Omega_\nu^2(k_\perp) = {} & 0.5(\Lambda_{11} + \Lambda_{22}) \\ & \pm 0.5((\Lambda_{11} - \Lambda_{22})^2 + 4\Lambda_{12}^2)^{1/2}, \end{aligned} \tag{25}$$

where $k^2 = k_x^2 + k_y^2$, ϑ is measured in the sagittal plane and is equal to zero for $\mathbf{k} \parallel [010]$, and Λ_{ik} are the Cristoffel tensor components. Expressions (24) and (25) make it possible to find the relation between the component of the wave vector $\mathbf{k}$ normal to the film surface with $\mathbf{n} \perp [001]$ on the one hand and the frequency and the wave number $k_\perp$ of a propagating bulk elastic wave on the other, i.e., to write the characteristic equation for the solution of the boundary-value problem for $\mathbf{n}$, $\mathbf{u}$, and $\mathbf{k} \in xy$. It follows from (24) and (25) that the structure of the component of the vector $\mathbf{u}$ of the lattice elastic displacements, which is normal to $\mathbf{n}$, can be presented in form (6) both for $\mathbf{n} \parallel [100]$ and for $\mathbf{n} \parallel [010]$. Using relations (6) and (25), we can classify the possible types of propagating elastic waves depending on the form of their localization near the surface of a magnet ($\zeta \geq 0$) defined by $q_{1,2}$. If we confine our subsequent analysis to the range of frequencies ω and wave numbers $k_\perp$ such that

$$\begin{aligned} & c_{11}, c_{12} \gg c_{66}, \quad \omega^2 \ll c_{11}/\rho \quad \text{for} \quad c_{11} \ll c_{22}, \\ & c_{22}, c_{12} \gg c_{66}, \quad \omega^2 \ll c_{22}/\rho \quad \text{for} \quad c_{11} \gg c_{22}, \end{aligned} \tag{26}$$

we can analyze the effect of rhombic anisotropy on the spectral structure of propagating bulk elastic waves polarized in the sagittal plane of a ferroelectric film taking into account only the quasi-transverse branch in the spectrum of normal elastic vibrations of an unbounded crystal (25), which is polarized in the plane of incidence (001) ($\chi \equiv c_{22}/c_{11}$):

$$\omega^2 \approx s_0^2 k^2 (1 + r_* 4k_x^2 k_y^2/(k_x^2 + \chi k_y^2)),$$
$$r_* \equiv \frac{c_{11}^2\chi + c_{66}^2\theta^2 - (c_{12} + c_{66}\theta)^2 - (1+\chi)c_{11}c_{66}\theta}{c_{11}c_{66}\theta}. \quad (27)$$

It follows from relations (6) and (27) that, depending on the form of localization of the film plane $\zeta = 0$ near the crystal surface, the following types of propagating normal quasi-transverse elastic waves are possible for **n** || [001] as well as for **n** || [010] for the chosen orientation of normal **n** to the film surface (**n** || [100] or **n** || [010]) and of the plane (001) of wave propagation.

I. Bulk wave of the first type ($q_1^2 > 0$; $q_2^2 < 0$):

$$\begin{aligned} &\omega^2 > s_t^2 k_\perp^2, \quad \mathbf{n} \parallel [100], \\ &\omega^2 > s_t^2 k_\perp^2, \quad \mathbf{n} \parallel [010]. \end{aligned} \quad (28)$$

II. Generalized surface waves ($q_1^2 = (q_2^2)^*$):

$$\begin{aligned} &\omega^2 < \omega_+^2 \equiv s_t^2 k_\perp^2 [1 - \chi + 4(\chi^{1/2}|r_*|^{1/2} - |r_*|)], \\ &2|r|^{1/2} > \chi^{1/2}, \quad r < 0, \quad \mathbf{n} \parallel [100], \\ &\omega^2 < \omega_+^2 \equiv s_t^2 k_\perp^2 [1 - 1/\chi + 4(|r_*|^{1/2} - |r_*|)/\chi], \\ &|r_*| > 1/4, \quad r_* < 0, \quad \mathbf{n} \parallel [010]. \end{aligned} \quad (29)$$

III. Surface waves ($q_{1,2}^2 > 0$):

$$\begin{aligned} &\omega_+^2 < \omega^2 < s_t^2 k_\perp^2, \quad 2|r_*|^{1/2} < \chi^{1/2}, \\ &r_* < 0, \quad \mathbf{n} \parallel [100], \\ &\omega_+^2 < \omega^2 < s_t^2 k_\perp^2, \quad |r_*| < 1/4, \\ &r < 0, \quad \mathbf{n} \parallel [010]. \end{aligned} \quad (30)$$

IV. Bulk waves of the second type ($q_{1,2}^2 < 0$):

$$\begin{aligned} &\omega_+^2 < \omega^2 < s_t^2 k_\perp^2, \quad 2|r_*|^{1/2} > \chi^{1/2}, \\ &r_* < 0, \quad \mathbf{n} \parallel [100], \\ &\omega_+^2 < \omega^2 < s_t^2 k_\perp^2, \quad |r_*| > 1/4, \quad \mathbf{n} \parallel [010]. \end{aligned} \quad (31)$$

If, however, $r_* > 0$, an analysis shows that all the relations derived above remain valid except that we must assume now that $\omega_+^2 \equiv 0$ for any $k_\perp$; i.e., the formation of generalized surface wave is ruled out.

Thus, the necessary condition for the propagation of the bulk Lamb phonons (corresponding to the quasi-transverse branch in the spectrum of normal elastic waves, which is polarized in the plane of incidence) along the ferroelectric plate under investigation is the requirement that their frequency and wave number $k_\perp$ belong to region I or IV. In this case, the bipartite elastic wave having the appropriate polarization and propagating along the film has at least one bulk component.

Multiple reflection of this type of phonons from the plate boundary and their interference lead to the formation of normal elastic modes polarized in the sagittal plane of the ferroelectric film under investigation. In our geometry (**k** ∈ *xy*, **n** || [100] or **n** || [010]), the spectrum of these bulk phonons can be determined in the explicit form:

$$\Omega_\nu^2(k_\perp) \approx s_0^2 k^2 \left(1 + \frac{r_* 4m_\nu^2 k_\perp^2}{m_\nu^2 + \chi k_\perp^2}\right), \quad \mathbf{n} \parallel [100], \quad (32)$$

$$\Omega_\nu^2(k_\perp) \approx s_0^2 k^2 \left(1 + \frac{r_* 4\chi k_\perp^2 m_\nu^2}{k_\perp^2 + \chi m_\nu^2}\right), \quad \mathbf{n} \parallel [010]. \quad (33)$$

Calculations show that in view of the above assumptions, the section of the isofrequency surface for the quasi-transverse branch in the spectrum of normal elastic waves polarized in the plane of incidence by the sagittal plane (001) can be described by the following relation ($\tan\vartheta = k_y/k_x$):

$$k^2 \approx \frac{\omega^2}{s_0^2}\left(1 + \frac{r_* \sin^2 2\vartheta}{\cos^2\vartheta + \chi\sin^2\vartheta}\right)^{-1}. \quad (34)$$

As in the case considered above, segments with a negative Gaussian curvature may be formed on the isofrequency surface of quasi-transverse normal elastic vibrations for definite relations between elastic moduli. The maxima on these segments in the *xy* plane coincide with the [100] and [010] directions. However, in view of the rhombic symmetry of the crystal, a concave segment can appear on curve (34) for **k** || [100] if

$$4|r_*| > \chi, \quad (35)$$

while the region with a negative curvature is formed on the slow-mode surface (34) for **k** || [010] when

$$4|r_*| > 1. \quad (36)$$

Thus, the conditions for the existence of a region with a negative Gaussian curvature on the refraction surface along normal **n** to the plate surface and along the direction of propagation of a bulk elastic wave, $\mathbf{k}_\perp \perp \mathbf{n}$, can be satisfied simultaneously in the given case. Let us find out how this circumstance is connected with the spectral structure of propagating bulk quasi-transverse elastic vibrations polarized in the sagittal plane of the crystal plate under investigation.

An analysis of relations (32) and (33) shows that the plate in question now has anisotropic elastic properties along the directions associated with vectors **n** and $\mathbf{k}_\perp$. For a given parameter of crystal rhombicity ($\chi \neq 1$), the necessary condition for the formation of a minimum on the dispersion curve for the phonon mode with a preset number ν is the fulfillment of the inequalities

$$\begin{aligned} &4|r_*| > 1, \quad \mathbf{n} \parallel [010], \\ &2|r_*| > \chi^{1/2}, \quad \mathbf{n} \parallel [100]. \end{aligned} \tag{37}$$

As regards the necessary condition for the formation of a point of degeneracy of the dispersion curves for modes with numbers ν and ρ, belonging to the spectrum of the bulk Lamb phonons under investigation, for the given directions of wave propagation and normal **n** to the film surface ($\mathbf{k}, \mathbf{n} \in xy$) it can be presented in the form

$$\begin{aligned} &2|r_*| > \chi^{1/2}, \quad \mathbf{n} \parallel [010], \\ &4|r_*| > 1, \quad \mathbf{n} \parallel [110]. \end{aligned} \tag{38}$$

The joint analysis of relations (18), (19) and (32), (33) shows that additional (as compared to those considered above) singularities appear in the spectrum of the bulk Lamb phonons in an anisotropic plate with $\mathbf{n} \parallel [100]$ (or $\mathbf{n} \parallel [010]$) if the following relations hold for $r_* < 0$:

$$\begin{aligned} &\chi > 4|r_*| > 1, \\ &\chi < 4|r_*| < 1. \end{aligned} \tag{39}$$

In this case, two additional versions (besides those analyzed above for $\chi = 1$) of the spectral rearrangement for phonons with a given polarization can be realized:

(1) for any mode number ν, the corresponding dispersion curve has a minimum at $k_\perp = k_{*\nu}$; however, there are no points of crossover for any $k_\perp$ and mode numbers ν and ρ if $\mathbf{n} \parallel [100]$, $\chi > 1$, or $\mathbf{n} \parallel [010]$, $\chi < 1$;

(2) for any $k_\perp$ and mode number ν, there are no points of minima on the dispersion curve $\Omega_\nu(k_\perp)$ ($\partial\Omega_\nu/\partial k_\perp > 0$); however, for any ν and ρ, there exists a point of degeneracy on the relevant dispersion curves $\Omega_\nu(k_\perp) = \Omega_\rho(k_\perp)$; for $\mathbf{n} \parallel [100]$, it is necessary that $\chi < 1$, while for $\mathbf{n} \parallel [010]$, the condition $\chi > 1$ must be observed.

Thus, the above analysis shows that the formation of these additional versions of the spectral rearrangement for bulk Lamb phonons is associated with a change in the local geometry of the isofrequency surface for the normal quasi-transverse phonons polarized in the plane of incidence in an unbounded crystal. Till now, the analysis of the effect of the local geometry of the isofrequency surface for a normal elastic wave in an unbounded crystal on the spectral structure of the corresponding type of bulk phonons propagating along the crystal plate in the vicinity of a shear-type structural phase transition was carried out in the limit $\delta \longrightarrow 0$, i.e., disregarding the correlation effects in the ferroelectric subsystem of the crystal. Consequently, the above results are valid primarily for plates of strongly anisotropic crystals away from the region of stability of the given crystalline state. As regards the spectral rearrangements for bulk phonons in the immediate vicinity of a ferroelastic phase transition, we must also take into account the correlation effects apart from the effects of the shape of the isofrequency surface for the normal acoustic vibrations of an unbounded crystal. In the model under investigation, this means that the approximation $\delta = 0$ becomes invalid. The results of such an analysis for the model of a rhombic crystal will be presented in the next section.

6. EFFECTS OF SPATIAL DISPERSION

An analysis of the boundary-value problem taking into account both elastic (4) and additional (3) boundary conditions shows that for any value of χ compatible with the conditions of elastic stability of the crystal, the spectrum of the Lamb waves of the type in question can be determined explicitly, as before, by using the same approximations, for any value of the wave number $k_\perp$ for $a = \infty$ both for $\mathbf{n} \parallel [100]$ and for $\mathbf{n} \parallel [010]$.

If we consider a rhombic crystal by way of an example, the spectrum of the bulk phonons under investigation in our geometry ($\mathbf{k} \in xy$, $\mathbf{n} \parallel [100]$ or $\mathbf{n} \parallel [010]$) can be determined in the explicit form ($\kappa_\nu \equiv \pi\nu/2d$, $\nu = 1, 2, \ldots$):

$$\begin{gathered} \Omega_\nu^2(k_\perp) \approx s_0^2 k^2 (1 + r_* 4\kappa_\nu^2 k_\perp^2/(\kappa_\nu^2 + \chi k_\perp^2)), \\ \mathbf{n} \parallel [100], \\ \Omega_\nu^2(k_\perp) \approx s_0^2 k^2 (1 + r_* 4k_\perp^2 \kappa_\nu^2/(k_\perp^2 + \chi \kappa_\nu^2)), \\ \mathbf{n} \parallel [010], \end{gathered} \tag{40}$$

$$r_* \equiv \frac{c_{11}^2\gamma + c_{66}^2\theta^2 - (c_{12} + c_{66}\theta)^2 - (1+\gamma)c_{11}c_{66}\theta}{c_{11}c_{66}\theta}. \tag{41}$$

A comparative analysis of relations (32), (33) and (40), (41) shows that the main auxiliary effects associated with the influence of the spatial dispersion of elastic moduli ($\delta \neq 0$) on the spectrum of the bulk Lamb phonons of the type in question are as follows.

(1) For $\omega_0^2 + c^2(\pi\nu/d)^2 \ll c^2 k_\perp^2 \ll \omega_{pe}^2$, the phase velocity of a mode with number ν becomes a function of the wave number $k_\perp$.

(2) The phase velocity of Lamb modes with numbers $\nu \neq 0$ does not vanish at the stability boundary of the given crystalline state.

(3) The spatial dispersion of the elastic modulus c_{66} leads to the disappearance of the points of crossover in the spectrum of the bulk Lamb phonons of the given type for modes with numbers ν and ρ (due to partial trapping at the surface) and to the divergence of the corresponding dispersion curves (dissipation is absent). For example, for the case of a rhombic crystal considered above, for $0 \neq a \ll 1$, the structure of the dispersion curves for modes with numbers ν and ρ in the vicinity of the point of degeneracy of the modes $\Omega_\nu(k_\perp)$ and $\Omega_\rho(k_\perp)$ existing at $\delta = 0$ can be approximately pre-

sented, taking into account relations (40) and (41), in the form

$$(\omega^2 - \Omega_\nu^2(k_\perp))(\omega^2 - \Omega_\rho^2(k_\perp)) - W_{\nu\rho}^2\Omega_\nu^2(k_\perp)\Omega_\rho^2(k_\perp) \approx 0, \tag{42}$$

where $||W_{\nu\rho}^2| < 1$, $W_{\nu\rho}^2 \approx a^2$. In the case of the intersection of the dispersion curves corresponding to the forward and backward types of the waves at the point of degeneracy $k_\perp = k_{\nu\rho}$ ($W_{\nu\rho}^2 = 0$), the removal of degeneracy ($W_{\nu\rho}^2 \neq 0$) leads to the disappearance of the degeneracy point and to the formation of two points of extrema in the vicinity of this point: a minimum for the branch whose frequency $\omega > \Omega_\nu(k_{\nu\rho})$, and a maximum for $\omega < \Omega_\nu(k_{\nu\rho})$.

(4) If the relation $c^2(\pi\nu/d)^2 \gg \omega_{pe}^2$ is satisfied for a phonon mode with number ν, the spectral structure of such an elastic mode taking into account the spatial dispersion in the ferroelectric subsystem of the crystal ($\delta \neq 0$) is determined only by the relation between elastic moduli in the limit $\gamma \longrightarrow 0$ not only far away from the ferroelastic phase transition, but also in its vicinity.

7. CONCLUSION

The main conclusions drawn in the present work can be formulated as follows.

(1) There exists a one-to-one correspondence between the presence of regions with a negative Gaussian curvature on the section of the isofrequency surface for normal quasi-transverse vibrations polarized in the incidence plane by the sagittal plane and certain anomalies in the spectrum (type of the wave and the presence of the points of extremum and crossover) of the dispersion curves of the modes belonging to the relevant type of generalized Lamb waves in an anisotropic plate.

(2) The form of elastic anisotropy can considerably affect the spectral structure of the Lamb phonons of the type under investigation.

(3) A consistent inclusion of the effects of spatial dispersion in the ferroelectric subsystem of the crystal may be of utmost importance for a correct description of spectral transformations for the bulk Lamb phonons under investigation in the vicinity of a proper ferroelastic phase transition.

In this paper, we analyzed the peculiarities in the transformation of the spectrum of bulk Lamb waves under the condition that their sagittal plane coincides with the symmetry plane of the crystal under investigation. It is known that in this case, *SH* waves and Lamb waves propagate independently. If, however, the crystal is hexagonal and the sixth-fold axis coincides with the direction of normal **n** to the surface of the film for which conditions (4) are satisfied on both surfaces as before, it can easily be verified that in view of elastic anisotropy, the anomalies observed by us (for an appropriate relation between bulk moduli) take place for any direction of propagation $\mathbf{k}_\perp/|\mathbf{k}_\perp|$ of a bulk Lamb wave in the plane of the film. In particular, the characteristic crossover and extremum points mentioned above will form closed curves in the $k_\perp$-space for $k_\perp \neq 0$, which lie in a plane with the normal along **n**.

It is well known that the presence in the dispersion relation for normal vibrations $\Omega_\nu(k_\perp)$ of points at which one or several components of the group velocity ($\partial\Omega_\nu/\partial k_\perp$) are equal to zero leads to the formation of singularities in the density of states of the relevant type of quasiparticles (phonons in the present case). Such points are known as critical and are associated with peculiarities in the thermodynamic, kinetic, and optical properties of the crystal. The formation of critical points can be associated with the symmetry properties of the crystal (symmetric critical points) [18]. Moreover, critical points whose presence is not connected in any way with the crystal symmetry may also exist. They are known as dynamic critical points [18]. An analysis of the conditions for the existence of critical points in the magnon spectrum was carried out in [19, 20], but only in the model of an unbounded crystal. In the present work, it is demonstrated for the first time that the consistent inclusion of the anisotropy of elastic moduli makes it possible to determine the mechanisms of the formation of a number of new critical points in the vicinity of a continuous shear-type structural phase transition in the phonon spectrum of a crystalline plate, which are absent in the model of an unbounded crystal.

It is well known [21] that the structure of normal bulk vibrational modes whose amplitude has nodes over the plate thickness weakly depends on the type of boundary conditions. Consequently, we can expect that the effects observed by us in the spectrum of bulk Lamb waves for a special case of elastic and auxiliary boundary conditions will take place for this type of bulk acoustic vibrations for other types of boundary conditions also (probably, except that the mode degeneracy will be removed and the dispersion curves for modes with different numbers will diverge in the region of crossover points).

In the present work, we considered by way of an example a proper ferroelastic phase transition in a crystalline plate. It can naturally be concluded, however, that the basic effects observed by us here can also be realized for plates of magnetic crystals undergoing a proper ferroelastic phase transition.

It should be emphasized that we considered only the transformation of the spectrum in the vicinity of a ferroelastic phase transition for nondipole-active bulk phonons polarized in the sagittal plane of a crystalline plate. The singularities in the spectrum of bulk phonons in the case of dipole-active elastic vibrations in an anisotropic plate will be analyzed in a separate publication.

ACKNOWLEDGMENTS

The author is grateful to E.P. Stefanovskiĭ and T.N. Tarasenko for the support of the idea of this work and for fruitful discussions.

REFERENCES

1. V. A. Krasil'nikov and V. V. Krylov, *Introduction to Physical Acoustics* (Nauka, Moscow, 1984).
2. M. K. Balakirev and I. A. Gilinskiĭ, *Waves in Piezoelectric Crystals* (Nauka, Novosibirsk, 1982).
3. B. A. Strukov and A. P. Levanyuk, *Physical Principles of Ferroelectric Phenomena in Crystals* (Nauka, Moscow, 1983).
4. R. Blinc and B. Zeks, *Soft Modes in Ferroelectrics and Antiferroelectrics* (North-Holland, Amsterdam, 1974; Mir, Moscow, 1975).
5. V. G. Bar'yakhtar, I. M. Vitebskiĭ, Yu. G. Pashkevich, *et al.*, Zh. Éksp. Teor. Fiz. **89**, 189 (1985) [Sov. Phys. JETP **62**, 108 (1985)].
6. I. E. Dikshteĭn, V. V. Tarasenko, and V. G. Shavrov, Fiz. Tverd. Tela (Leningrad) **19**, 1107 (1977) [Sov. Phys. Solid State **19**, 644 (1977)].
7. Yu. A. Kosevich and E. S. Syrkin, Fiz. Tverd. Tela (Leningrad) **26**, 2927 (1984) [Sov. Phys. Solid State **26**, 1769 (1984)].
8. Yu. A. Kosevich and E. S. Syrkin, Zh. Éksp. Teor. Fiz. **89**, 2221 (1985) [Sov. Phys. JETP **62**, 1282 (1985)].
9. A. M. Kosevich, Yu. A. Kosevich, and E. S. Syrkin, Zh. Éksp. Teor. Fiz. **88**, 1089 (1985) [Sov. Phys. JETP **61**, 639 (1985)].
10. Yu. V. Gulyaev and P. E. Zil'berman, Izv. Vyssh. Uchebn. Zaved., Fiz. **31**, 6 (1988).
11. Yu. V. Gulyaev, I. E. Dikshteĭn, and V. G. Shavrov, Usp. Fiz. Nauk **167**, 735 (1997) [Phys. Usp. **40**, 701 (1997)].
12. V. N. Lyubimov, Kristallografiya **25**, 675 (1980) [Sov. Phys. Crystallogr. **25**, 389 (1980)].
13. V. I. Al'shits, J. Lothe, and V. N. Lyubimov, Kristallografiya **28**, 635 (1983) [Sov. Phys. Crystallogr. **28**, 374 (1983)].
14. S. V. Gerus, I. E. Dikshteĭn, V. V. Tarasenko, and V. D. Kharitonov, Fiz. Tverd. Tela (Leningrad) **19**, 218 (1977) [Sov. Phys. Solid State **19**, 124 (1977)].
15. M. G. Cottam, D. R. Tilley, and B. Zeks, J. Phys. C **17**, 1793 (1984).
16. A. G. Khachaturyan, *The Theory of Phase Transformations and the Structure of Solids Solutions* (Nauka, Moscow, 1974).
17. F. Tricomi, *Differential Equations* (Hafner, New York, 1961; Inostrannaya Literatura, Moscow, 1962).
18. J. Birman, *Theory of Crystal Space Groups and Infra-Red Raman Lattice Processes of Insulating Crystals* (Springer-Verlag, Heidelberg, 1974; Mir, Moscow, 1978), vol. 1.
19. V. G. Bar'yakhtar, A. G. Kvirikadze, and V. A. Popov, Zh. Éksp. Teor. Fiz. **59**, 898 (1970) [Sov. Phys. JETP **32**, 488 (1971)].
20. V. V. Eremenko, S. A. Zvyagin, Yu. G. Pashkevich, *et al.*, Zh. Éksp. Teor. Fiz. **93**, 2075 (1987) [Sov. Phys. JETP **66**, 1184 (1987)].
21. L. M. Brekhovskikh, *Waves in Layered Media* (Nauka, Moscow, 1973; Academic, New York, 1980).

Translated by N. Wadhwa

Journal of Experimental and Theoretical Physics, Vol. 92, No. 2, 2001, pp. 338–344.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 119, No. 2, 2001, pp. 388–395.
Original Russian Text Copyright © 2001 by Skobov, Chernov.

SOLIDS
Electronic Properties

Nonlinear Doppler-Shifted Cyclotron Resonance in Aluminum

V. G. Skobov[a] and A. S. Chernov[b], *

[a]*Ioffe Physicotechnical Institute, Russian Academy of Sciences, ul. Politekhnicheskaya 26, St. Petersburg, 194021 Russia*
[b]*Moscow State Engineering-Physics Institute, Kashirskoe sh. 31, Moscow, 115409 Russia*
**e-mail: chernov@theor.mephi.msk.su*

Received September 18, 2000

Abstract—The Doppler-shifted cyclotron resonance in an aluminum plate in the geometry when constant magnetic field **H** is directed along the [100] crystallographic axis oriented normally to the surface of the plate is studied theoretically. The analysis is performed for a simple model Fermi surface possessing fourth-order symmetry. Capture of holes by the magnetic field of a radio-frequency wave is shown to considerably decrease the effectiveness of cyclotron absorption at large exciting field amplitudes. This suppresses the collisionless damping of dopplerons (propagating modes related to odd cyclotron resonance harmonics). As a result, the sample becomes more transparent to radio-frequency radiation. © *2001 MAIK "Nauka/Interperiodica".*

In [1], we studied the penetration of radio-frequency radiation through noble metals in a nonlinear regime. It was shown that the "capture" of carriers by the magnetic field of a large-amplitude wave could substantially decrease collisionless absorption. For instance, when constant magnetic field **H** was directed along the [110] axis and there were open orbits on the Fermi surface, damping caused by them was suppressed to the extent that helicon propagation became possible (in the linear regime, carriers in open orbits hindered helicon propagation). In the region below the helicon threshold, the suppression of the cyclotron wave absorption by holes could be so strong that the propagation of a new wave having no analogs in the linear regime became possible.

Substantial nonlinear effects in the propagation of radio-frequency waves can exist not only in noble metals but also in other metals with unequal concentrations of electrons and holes, in particular, in aluminum. The Fermi surface of aluminum bears no resemblance to the Fermi surfaces of noble metals. Collisionless absorption in aluminum is determined by the form of the $\partial S/\partial p_z$ function, where $S(p_z)$ is the area of the cross section of the Fermi surface by the p_z = const plane and p_z is the momentum component of the conduction electron along the z axis (the field **H** direction). The properties of collisionless absorption in aluminum are different from those of collisionless absorption in noble metals. There are no open orbits in aluminum, and the $\partial S/\partial p_z$ derivative does not become infinite, as it does in noble metals, at the boundary cross section separating hole orbits of the "dog's bone" type from electron orbits.

The hole Fermi surface of aluminum is such that, in the $\mathbf{H} \parallel [100]$ geometry, the $S(p_z)$ area is maximum at the central cross section ($p_z = 0$) and monotonically decreases as p_z increases. The $|\partial S/\partial p_z|$ function experiences complex nonmonotonic variations. It sharply increases near the central cross section, reaches a maximum, slightly decreases and passes a minimum, again increases and passes another maximum equal in height to the first one, and then monotonically decreases. Accordingly, holes can be divided into three groups. The first group comprises holes for which $|\partial S/\partial p_z|$ changes from zero to a maximum value. For holes of the second group, $|\partial S/\partial p_z|$ changes from the first maximum to the second, and the third group includes holes for which $|\partial S/\partial p_z|$ decreases after the second maximum. Although the second group is the largest, these holes do not contribute to cyclotron absorption in the short-wave region. Short-wave cyclotron absorption is caused by first- and third-group holes, which make up a comparatively small fraction of the total number of holes. For this reason, collisionless absorption in aluminum is considerably weakened compared with the metals in which cyclotron absorption is determined by majority carriers. It is therefore of interest to study the propagation of radio-frequency waves in aluminum in a nonlinear regime, in which collisionless absorption should be weakened to a still greater extent. It should be borne in mind that the situation with aluminum is essentially different from that with noble metals. In noble metals, the orbits of the electrons that determine the spectrum of a nonlinear wave are virtually circular; for this reason, only fundamental Doppler-shifted cyclotron resonance (DSCR) is observed for these metals. The Fermi surface of aluminum has fourth-order symmetry, and hole orbits of the second group, which are majority carriers, are closer to squares than to circles. For this reason, there occur multiple DSCRs, which should result in the existence of the corresponding propagating modes. This work is concerned precisely with these effects in aluminum in a nonlinear regime.

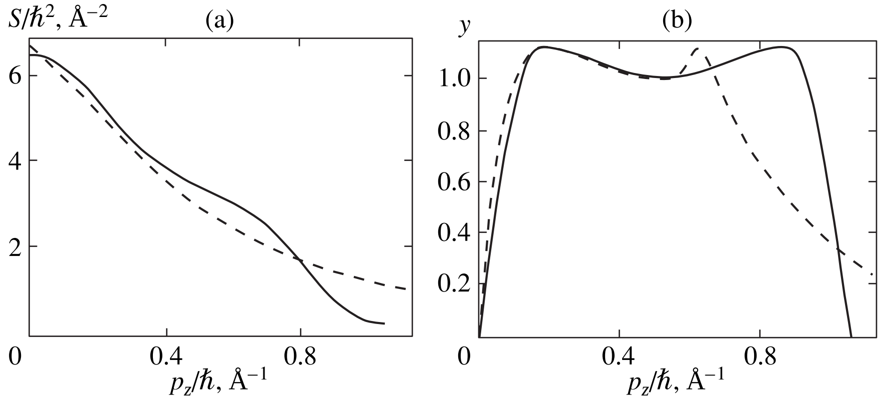


Fig. 1. (a) Cross section area and (b) the derivative of the cross section area of the Fermi surface of aluminum in the (**H** || [100] || z) geometry.

1. A FERMI SURFACE MODEL AND NONLOCAL CONDUCTIVITY

Consider the propagation of a radio-frequency wave in aluminum in the (**H** || **k** || [100] || z) geometry, where **k** is the wave vector. The concentration of electrons in aluminum is less than 3% of the concentration of holes, and the contribution of electrons to conductivity can therefore be ignored. The cross-section area $S(p_z)$ and its derivative $\partial S/\partial p_z$ are plotted as functions of p_z in Fig. 1 (dashed lines). These plots were obtained in the calculations performed by Larsen and Greisen [2]. We consider a model in which the $S(p_z)$ dependence is given by the formulas

$$\frac{1}{2\pi}\frac{\partial S}{\partial p_z} = -py(p_z)\,\mathrm{sgn}\,p_z, \tag{1}$$

$$y(p_z) = \alpha\sin\frac{\pi p_z}{2p_1}, \quad |p_z| \le p_1, \tag{2}$$

$$y(p_z) = \frac{1}{2}\left[1+\alpha+(\alpha-1)\cos\left(\pi\frac{|p_z|-p_1}{p_2}\right)\right], \quad p_1 \le |p_z| \le p_1+p_2, \tag{3}$$

where α is a dimensionless constant and p, p_1, and p_2 are momentum dimension parameters. In the interval $p_1 + p_2 \le |p_z| \le 2(p_1 + p_2)$, the y function should be continued symmetrically with respect to the $p_z = p_1 + p_2$ point. The $y(p_z)$ function at $\alpha = 1.1$ and $p_2/\hbar = 2p_1/\hbar = 0.35$ Å^{-1} is shown in Fig. 1b (solid line). Its symmetry with respect to the minimum considerably simplifies nonlocal conductivity calculations. The left part of the $y(p_z)$ plot is similar to the left part of the Larsen–Greisen calculation curve, whereas the difference between the right parts is of no consequence because the majority of holes occur in the $|p_z| \le p_1 + p_2$ region. The p parameter should be adjusted to provide coincidence between the period of Gantmakher–Kaner oscillations and the experimental value; this condition is met at $p/\hbar = 1$ Å^{-1}. Note that, at such a p value, the $\partial S/\partial p_z$ derivative at a minimum somewhat exceeds the value obtained by Larsen and Greisen [2]. It is also necessary to specify the area of the central cross section $S(p_z = 0)$. This value is determined by the condition of the equality of the concentration of holes in the model under consideration to the true concentration of holes in aluminum [3],

$$\frac{2V}{(2\pi\hbar)^3} = N, \tag{4}$$

where

$$V = 4(p_1+p_2)S(p_1+p_2) \tag{5}$$

and V is the volume enclosed by the hole Fermi surface. Equation (5) follows from $\partial S/\partial p_z$ function symmetry with respect to the $p_z = p_1 + p_2$ point. As

$$S(p_1+p_2) = S(0) - 2\pi p\int_0^{p_1+p_2} y(p')dp', \tag{6}$$

$S(0)$ is easy to express through N. The plot of the $S(p_z)$ function for our model at $N = 0.06$ Å^{-3} is shown in Fig. 1a (solid line).

Let the hole orbits be square,

$$|p_x| + |p_y| = \sqrt{\frac{1}{2}\tilde{S}(\varepsilon, p_z)}, \tag{7}$$

where $\tilde{S}(\varepsilon, p_z)$ is the area of the cross section of a constant-energy surface (ε = const) by a p_z = const plane. The attractive feature of square orbits is their simplicity, required symmetry, and some similarity to real hole orbits in aluminum. We in addition assume that the cyclotron masses of all holes equal m. The $\tilde{S}(\varepsilon, p_z)$ function can then be written in the form

$$\tilde{S}(\varepsilon, p_z) = 2\pi m(\varepsilon - \varepsilon_F) + S(p_z), \tag{8}$$

where ε_F is the Fermi energy. It follows from (8) that the dispersion of holes obeys the law

$$\varepsilon(p) - \varepsilon_F = \frac{1}{\pi m}(|p_x| + |p_y|)^2 - \frac{S(p_z)}{2\pi m}. \qquad (9)$$

The components of the velocity vector of a hole are determined by the formulas

$$v_\alpha \equiv \frac{\partial \varepsilon}{\partial p_\alpha} = \frac{2}{\pi m}(|p_x| + |p_y|)\,\mathrm{sgn}\,p_\alpha, \quad \alpha = x, y, \qquad (10)$$

$$v_z \equiv \frac{\partial \varepsilon}{\partial p_z} = -\frac{1}{2\pi m}\frac{\partial S}{\partial p_z}. \qquad (11)$$

In the geometry under consideration, when the **k** and **H** vectors are directed along the symmetry axis of the Fermi surface, the system of equations for circularly polarized wave field components decomposes into two independent equations containing the $\sigma_\pm(k, H) = \sigma_{xx} \pm i\sigma_{yx}$ conductivity elements. According to [4], these elements are determined by the formulas

$$\sigma_\pm(k, H) = \frac{2\pi e^2 m}{(2\pi\hbar)^3} \times \int dp_z \sum_{n=-\infty}^{\infty} \frac{v_n^\pm(v_n^+ + v_n^-)}{\nu + i(n\omega_c - k v_z - \omega)}, \qquad (12)$$

where

$$v_n^\pm = \frac{1}{2\pi}\int_0^{2\pi} [v_x(\tau) \pm i v_y(\tau)] e^{in\tau} d\tau, \qquad (13)$$

e is the absolute value of the charge of the electron, c is the velocity of light, $\omega_c = eH/mc$ is the cyclotron frequency of holes, ν is the frequency of hole collisions with the lattice, and τ is the dimensionless time characterizing the motion of a particle in the orbit. Further, we will be interested in strong magnetic fields when $\omega_c \gg |\nu - i\omega|$.

According to (7), $|p_z| + |p_y| = \text{const}$ at fixed ε and p_z. For this reason, at each square side, the v_x and v_y velocity components given by (10) remain constant. This considerably simplifies the calculations of integral (13), because over each quarter of the hole revolution period, the expression in square brackets remains constant. As a result, $v_n^\pm$ is given by

$$v_n^\pm = \pm\frac{4\sqrt{2}}{\pi^2 n}\frac{\sqrt{S(p_z)}}{m}\delta_{n,\,4s\pm1}, \qquad (14)$$

where s runs over all integers from $-\infty$ to ∞. It follows from (14) that DSCR harmonics with numbers $n = \ldots, -7, -3, 1, 5, \ldots$ and polarization (+) and with numbers $n = \ldots, -5, -1, 3, 7, \ldots$ and polarization (–) are present in conductivity (12). This is a consequence of the fourth-order symmetry of the Fermi surface.

Substituting (14) and (11) into (12) transforms the equation for the nonlocal conductivity into

$$\sigma_\pm = -i\frac{Nec}{H}s_\pm(q), \qquad (15)$$

$$s_\pm(q) = \frac{32}{\pi^3}\sum_{n=-\infty}^{\infty}\frac{1}{(4s\pm1)^2}F_s^\pm(q), \qquad (16)$$

$$F_s^\pm(q) = \frac{1}{2(p_1+p_2)} \times \int_0^{p_1+p_2} dp_z\left[\frac{1}{I_s^\pm - qy(p_z)} + \frac{1}{I_s^\pm - qy(p_z)}\right], \qquad (17)$$

$$q = \frac{kcp}{eH}, \quad I_s^\pm = 4s \pm 1 - i\gamma, \quad \gamma = \frac{\nu}{\omega_c}. \qquad (18)$$

The q value is the ratio between the shift of holes with $p_z = p_1 + p_2$, to which a minimum of $|\partial S/\partial p_z|$ corresponds, and the radio-frequency wave length.

Note that, in the local collisionless limit ($q \longrightarrow 0$, $\gamma \longrightarrow 0$),

$$F_s^\pm = 1/(4s\pm1),$$

$$\sum_{n=-\infty}^{\infty}\frac{1}{(4s\pm1)^3} = \frac{\pi^3}{32}$$

and, therefore, $s_\pm(0) = \pm1$; that is, we obtain the local Hall conductivity.

Next, let us substitute (2) and (3) into (17) and perform the integration over p_z. This yields

$$F_s^\pm(q) = \frac{4s\pm1}{|4s\pm1|} \times \left\{\frac{\eta_1}{\sqrt{(I_s^\pm)^2 - \alpha^2 q^2}} + \frac{\eta_2}{\sqrt{(I_s^\pm - \alpha q)(I_s^\pm - q)}} + \frac{\eta_2}{\sqrt{(I_s^\pm + \alpha q)(I_s^\pm + q)}}\right\}, \qquad (19)$$

for $q^2 < (4s \pm 1)^2/\alpha^2$, and

$$F_s^\pm(q) = \frac{i\eta_1}{\sqrt{(I_s^\pm)^2 - \alpha^2 q^2}} - \frac{\eta_2\,\mathrm{sgn}\,q}{\sqrt{(I_s^\pm - \alpha q)(I_s^\pm - q)}} + \frac{\eta_2\,\mathrm{sgn}\,q}{\sqrt{(I_s^\pm + \alpha q)(I_s^\pm + q)}}, \qquad (20)$$

for $q^2 > (4s \pm 1)^2$, where the coefficients

$$\eta_1 = \frac{p_1}{p_1 + p_2}, \quad \eta_2 = \frac{p_2}{2(p_1 + p_2)} \qquad (21)$$

are $\eta_1 = \eta_2 = 1/3$ for the model under consideration ($p_2 = 2p_1$). We do not give expressions for $F_s^{\pm}$ in intermediate regions $(4s \pm 1)^2/\alpha^2 < q^2 < (4s \pm 1)^2$, where the imaginary part of $F_s^{\pm}$ caused by cyclotron absorption is larger than its real part, and the propagation of modes of any kind is therefore impossible. Further, we will not be interested in these q value intervals.

2. FUNDAMENTAL DOPPLERON IN LINEAR AND NONLINEAR REGIMES

The properties of eigenmodes in metals are determined by solutions to the dispersion equation

$$k^2c^2 = 4\pi i\omega\sigma_{\pm}. \qquad (22)$$

Using the dimensionless wave vector q allows this equation to be conveniently written in the form

$$\Phi_{\pm}(q) = \frac{1}{\xi}, \qquad (23)$$

$$\Phi_{\pm}(q) = \frac{s_{\pm}(q)}{q^2}, \qquad (24)$$

$$\xi = \frac{4\pi\omega Ncp^2}{eH^3}. \qquad (25)$$

The character of solutions to (23) can be studied by considering the behavior of the $\Phi_{\pm}$ functions. The imaginary parts of the Φ_+ and Φ_- functions coincide, and their real parts only differ in sign. The $\Phi'_+ = \mathrm{Re}\Phi_+$ function in the region $q > 1/\alpha$ for $\gamma \longrightarrow 0$ is plotted in Fig. 2 (in this region, $\Phi''_+ = \mathrm{Im}\Phi_+ = 0$). At small q values, we have $\Phi'_+ \approx 1/q^2$. This dispersion curve branch corresponds to the helicon whose field rotates in the same direction as the holes do. At $q = q_H$, the Φ'_+ function reaches the minimum value $\Phi_H^{\min} \approx 4$, and, as $q \longrightarrow 1/\alpha$, it tends to infinity, which is caused by the resonance of the holes that are characterized by the largest shift during the cyclotron period (holes with $p_z = p_1$). It follows that, above the helicon threshold $H = H_L$, at which $\xi = 1/4$, a DSCR mode, or doppleron, does exist. This doppleron is, however, virtually unobservable because it has the same polarization as the helicon but its amplitude is much smaller.

The $\Phi'_- = \mathrm{Re}\Phi_-$ and $\Phi''_- = \mathrm{Im}\Phi_-$ functions in the region $1 < q < 3/\alpha$, where $\Phi''_- > 0$, are plotted in Fig. 3. At the boundaries of this interval, Φ''_- becomes infinite, and within this interval, it decreases to its minimum value $\Phi_1^{\min} \approx 0.04$. This means that, in the "–" polarization at fields $H > H_L/10^{2/3}$, two DSCR modes can exist. One of them is caused by the fundamental DSCR of holes with $p_z = p_1 + p_2$, for which $\partial S/\partial p_z$ and, therefore, the shift along field **H** during the cyclotron period has minima. The q value for this dispersion curve branch tends to one as H increases. The second mode is related to the third DSCR harmonic of holes with the largest shift during the cyclotron period. In strong fields, the q value for this doppleron tends to $3/\alpha$. Both dopplerons exist above field H_1, whose value is approximately five times lower than the threshold helicon field H_L. In reality, the doppleron caused by holes with the largest shift is virtually unobservable, because the amplitude of the third DSCR harmonic is much (approximately 30 times) smaller than the fundamental resonance amplitude.

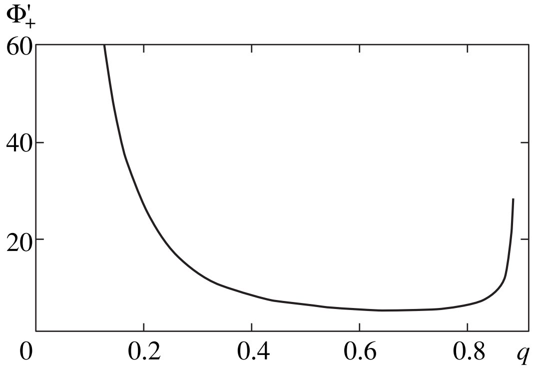


Fig. 2.

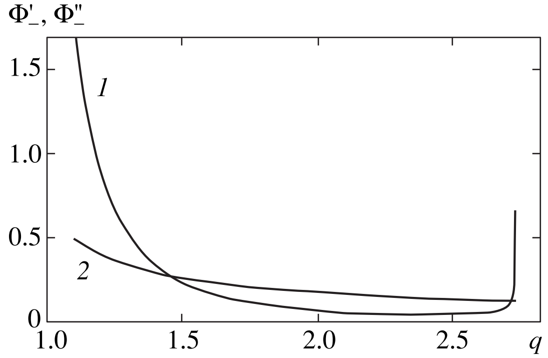


Fig. 3. Functions (*1*) Φ'_- and (*2*) Φ''_-.

The spectrum and damping of the fundamental doppleron can be determined by solving dispersion equation (23) with the $\Phi_-(q)$ complex function. At strong fields, in which $\xi \ll 1$, the dispersion equation can be solved analytically,

$$q_1 \approx 1 + \xi^2 + i\gamma. \qquad (26)$$

In this region, collisionless cyclotron absorption virtually does not contribute to doppleron damping, and doppleron damping is caused by collisions of holes. In a wide range of magnetic field values, the dispersion

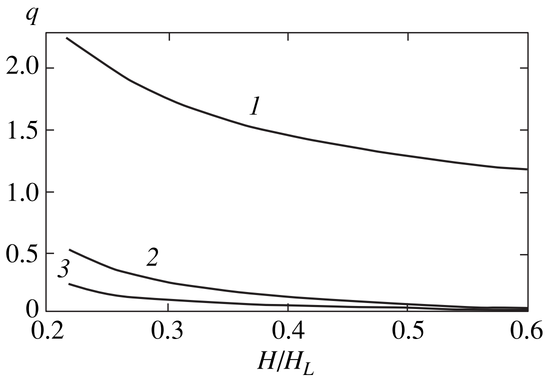


Fig. 4. Functions (*1*) $q'(H)$ and (*2*, *3*) $q''(H)$ in (*2*) linear and (*3*) nonlinear regimes.

equation can only be solved numerically. The $q' = \mathrm{Re}\,q_1$ and $q'' = \mathrm{Im}\,q_1$ curves (curves *1* and *2*, respectively) are plotted in Fig. 4 versus magnetic field **H**. The damping is seen to increase as the doppleron threshold is approached. The major contribution to the damping is made by cyclotron absorption caused by first- and third-group holes.

This is how matters stand with the linear regime. At large radio-frequency field amplitudes, the magnetic field of the wave can "capture" holes responsible for cyclotron absorption [5]. As a result, the translational motion of these holes along the **H** vector is superimposed on vibrational motions with the frequency

$$\omega_0 \approx \omega_c \sqrt{\frac{H_a}{H}}, \tag{27}$$

where H_a is the amplitude of the magnetic field of the wave. If the frequency of oscillations of captured holes is high, and $\omega_0 \gg \nu$, the effectiveness of wave absorption by such holes decreases ω_0/ν times. The equation for the $F_1^-(q)$ function describing the fundamental DSCR then takes the form [cf. (20)]

$$F_{1nl}^-(q) = \frac{i\eta_1\nu}{\omega_0\sqrt{\alpha^2 q^2 - 1}} + \frac{\eta_2}{\sqrt{(\alpha q - 1)(q - 1)}} - \frac{\eta_2}{\sqrt{(\alpha q + 1)(q + 1)}}, \tag{28}$$

and the other terms of series (16) remain unchanged. As a result, fundamental doppleron damping in comparatively weak magnetic fields can be substantially weaker than in the linear regime. For instance, at $\omega/2\pi$ = 50 kHz, $\nu = 4 \times 10^8$ s^{-1}, and an H_a = 100 Oe, the ω_0/ν ratio at the doppleron threshold ($H = H_1$) is close to five. The $q''(H)$ dependence then has the form of curve *3* in Fig. 4, which shows that the capture of holes by the wave field in a nonlinear regime substantially weakens doppleron damping.

In addition to the propagating mode (doppleron), there is one more mode (damped mode). The corresponding root q_2 of the dispersion equation in the vicinity of the doppleron threshold $H > H_1$ can be determined from the asymptotic behavior of conductivity $\sigma_-(q)$ at $q^2 \gg 1$, because $\xi \ll 1$ in this region. We can then disregard multiple resonances with small amplitudes and only take into account the term with $s = 0$ in (16). Neglecting the difference between the $32/\pi^3$ coefficient and one, we find

$$s_-(q) \approx \frac{i}{3\alpha|q|} + \frac{2(1+\alpha)}{3\alpha^{3/2}q^2}. \tag{29}$$

To this conductivity there corresponds the well-known anomalous skin effect.

If nonlinearity is strong ($\omega_0 \gg \nu$) and collisionless absorption is suppressed, the first term in (29) decreases ω_0/ν times; that is, the nonlocal conductivity takes the form

$$s_-^{nl}(q) \approx \frac{i\nu}{3\alpha|q|\omega_0} + \frac{2(1+\alpha)}{3\alpha^{3/2}q^2}. \tag{30}$$

If the ω_0/ν ratio is so large that the first term in (30) becomes smaller than the second, the dispersion equation becomes

$$q^4 \approx \xi\left[\frac{i\nu}{3\alpha|q|\omega_0} + \frac{2(1+\alpha)}{3\alpha^{3/2}q^2}\right]. \tag{31}$$

The first, almost real, root of this equation q_1 approximately describes the doppleron in the $\xi \gg 1$ region. The second, almost imaginary, root q_2 refers to the damped component,

$$q_2 \approx \left[\frac{2(1+\alpha)}{3\alpha^{3/2}}\right]^{1/4}. \tag{32}$$

Next, consider the surface impedance of an aluminum plate under antisymmetric excitation with respect to the electric field. When the reflection of carriers from the surface of a plate is diffuse and the field within the plate is a superposition of two eigenmodes, the impedance is described by the formula given in [6, Eq. (29)]. If the second component is damped as in the case under consideration, this formula becomes

$$Z_- = \frac{8\pi\omega p}{ceH} \times \frac{(1-q_1^2)(1+q_1^{-1}) - (1-q_2^2)q_2^{-1}(1+t)}{q_2^2 - q_1^2 + (1-q_1^2)q_2^{-1} - (1-q_2^2)t}, \tag{33}$$

$$t = \frac{1}{q_1}\frac{1-\exp(iq_1L)}{1+\exp(iq_1L)}, \quad L = \frac{eH}{pc}d, \tag{34}$$

where d is the thickness of the plate.

The results obtained in calculating $R_- = \mathrm{Re}Z_-$ at $\omega/2\pi = 50$ kHz, $\nu = 4 \times 10^8\ \mathrm{s}^{-1}$, and $d = 0.8$ mm are shown in Fig. 5. Curve *1* was obtained for the linear regime, and curve *2*, for the nonlinear regime at an $H_a = 100$ Oe amplitude of the wave magnetic field. At these parameter values, the threshold helicon field was $H_L \approx 8.3$ kOe, and the ω_0/ν ratio at the fundamental doppleron threshold equaled five. Figure 5 shows that the surface impedance in the linear regime is a monotonic function of H. In the nonlinear regime, R experiences strong oscillations because collisionless doppleron damping is suppressed and the plate becomes more transparent to the radio-frequency wave.

3. THE THIRD DSCR HARMONIC

As in the $1/\alpha < q < 1$ interval, strong cyclotron absorption caused by second-group holes occurs in the $3/\alpha < q < 3$ interval, and there can be no wave propagation. At large q values, $3 < q < 5/\alpha$, first- and third-group holes only contribute to cyclotron absorption, and this absorption is much less effective. Close to the boundaries of this interval, the Φ'_+ function sharply increases and becomes much larger than Φ''_+ (Fig. 6). Within the interval, the Φ'_+ function has a minimum and decreases to $\Phi_3^{\min} \approx 0.0006$. This means that two modes exist in magnetic fields $H > H_3 \equiv H_L(\Phi_3^{\min}/\Phi_H^{\min}) \approx 0.05H_L$. One of these is related to the third DSCR harmonic of holes with the largest shift during the cyclotron period. The second mode has a much smaller amplitude and manifests itself much less strongly. The ratio between these two modes is similar to the ratio between two modes in the $1 < q < 3/\alpha$ interval.

The dispersion equation determining the spectrum and damping of the doppleron with the "+" polarization can be written as

$$q^2 = \xi\left[F_0^+(q) + \frac{1}{9}F_{-1}^+(q) + \frac{1}{25}F_1^+(q)\right], \qquad (35)$$

$$F_0^+(q) = \frac{i\eta_1}{(1+\omega_0/\nu)\sqrt{\alpha^2q^2-1}} - \frac{\eta_2}{\sqrt{(\alpha q-1)(q-1)}} + \frac{\eta_2}{\sqrt{(\alpha q+1)(q+1)}}, \qquad (36)$$

$$F_{-1}^+(q) = \frac{i\eta_1}{(1+\omega_0/\nu)\sqrt{\alpha^2q^2-(3+i\gamma)^2}} + \frac{\eta_2}{\sqrt{(\alpha q-3-i\gamma)(q-3-i\gamma)}} - \frac{\eta_2}{\sqrt{(\alpha q+3-i\gamma)(q+3-i\gamma)}}, \qquad (37)$$

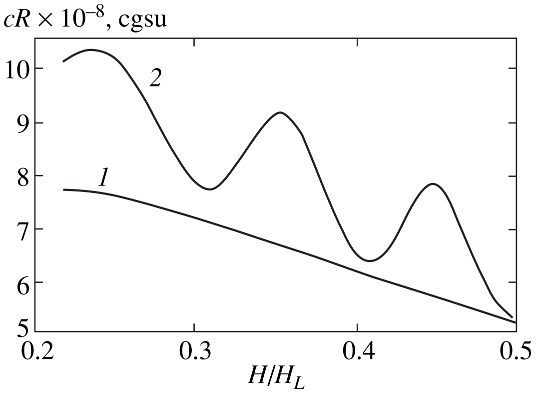


Fig. 5. Aluminum plate surface impedance R_- as a function of H in (*1*) linear and (*2*) nonlinear regimes.

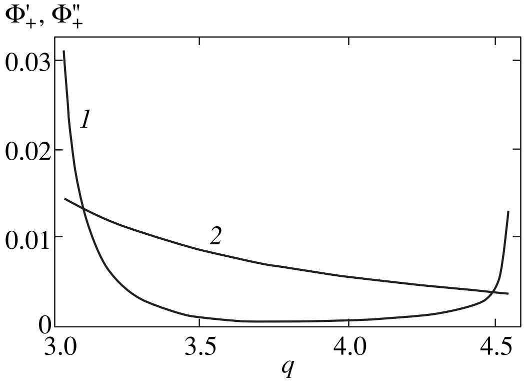


Fig. 6. Functions (*1*) Φ'_+ and (*2*) Φ''_+.

$$F_1^+(q) = \frac{\eta_1}{\sqrt{(5-i\gamma)^2-\alpha^2q^2}} + \frac{\eta_2}{\sqrt{(5-i\gamma-\alpha q)(5-i\gamma-q)}} + \frac{\eta_2}{\sqrt{(5-i\gamma+\alpha q)(5-i\gamma+q)}}. \qquad (38)$$

Equations (36) and (37) are interpolation equations; their first terms are written in such a way that they remain valid in both linear ($\omega_0 \ll \nu$) and nonlinear ($\omega_0 \gg \nu$) regimes. We omitted the terms for $s_+(q)$ (Eq. (16)) that described the seventh, ninth, and higher DSCR harmonics, because harmonic amplitudes rapidly decreased as the harmonic number increased, and the denominators in the q range under consideration are far from resonance.

The solution to the dispersion equation (35) determines the spectrum and damping of the doppleron related to the third DSCR harmonic (Fig. 7). Curves *1* and *2* in Fig. 7 correspond to the $q'_1 = \mathrm{Re}q_1(H)$ and

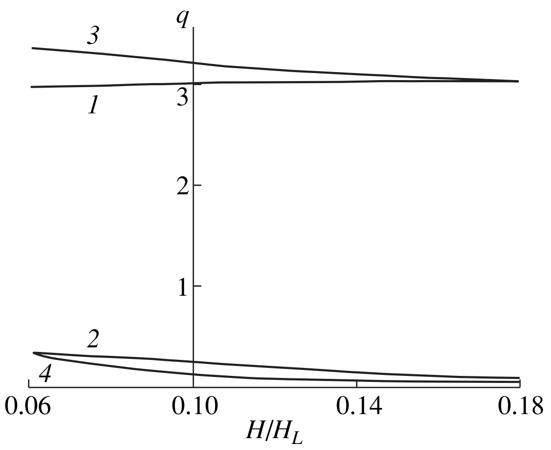


Fig. 7. Functions (*1*, *3*) $q'(H)$ and (*2*, *4*) $q''(H)$ in (*1*, *2*) linear and (*3*, *4*) nonlinear regimes for the third mode.

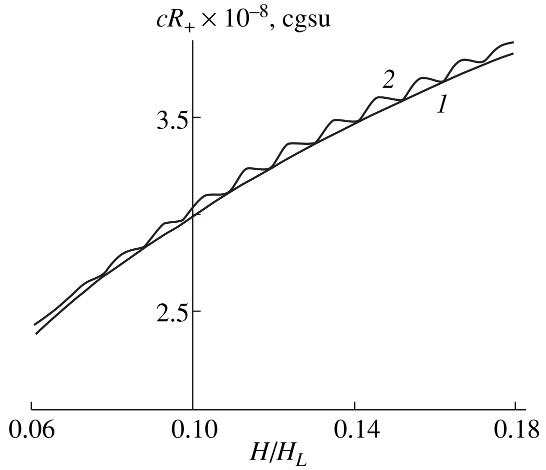


Fig. 8. Functions $R_+(H)$ in the region of weak magnetic fields in (*1*) linear and (*2*) nonlinear regimes.

$q_1'' = \mathrm{Im}\, q_1(H)$ functions, respectively, and curves *3* and *4* describe these functions at 50 kHz and $H_a = 200$ Oe.

Apart from the doppleron root q_1, Eq. (35) has two essentially complex roots, q_2 and q_3, which correspond to the damped wave field component. To determine these roots in the nonlinear regime, it suffices to take into account only the second and third terms of (36) in Eq. (16) for $s_+(q)$ and assume that $q \gg 1$ in these terms. This gives

$$s_+(q) \approx -\frac{1+\alpha}{3\alpha^{3/2}q^2} \approx -\frac{2}{3q^2}, \tag{39}$$

and we obtain

$$q_2 = (1+i)\left(\frac{\xi}{6}\right)^{1/4}, \quad q_3 = (-1+i)\left(\frac{\xi}{6}\right)^{1/4}. \tag{40}$$

If three different exponential wave field components exist in a metal, the impedance of the plate can be obtained with the use of a method similar to that applied in [6] to derive a formula for Z in the presence of two exponential components. The corresponding equation is very cumbersome. It can be substantially simplified taking into account that the q_2 and q_3 roots have large imaginary parts (therefore, $|\exp(iq_{2,3}L)| \ll 1$) and that the q_1 doppleron root is close to three, $(q_1 - 3) \ll 1$. This allows the equation for Z to be reduced to

$$Z_+ = \frac{8\pi\omega p}{ceH}[(q_2-3)(q_3-3) - 2(q_1-3)e^{iq_1L}] \tag{41}$$
$$\times [(q_2-3)(q_3-3)(q_2+q_3+1) + 2(q_1-3)q_1e^{iq_1L}]^{-1}.$$

The terms in the numerator and denominator of (41) that contain the $\exp(iq_1L)$ factor describe the penetration of the mode related to the third DSCR harmonic through the plate. These terms are proportional to the small $q_1 - 3$ value, which means that this mode is excited with a noticeably smaller amplitude than the fundamental doppleron. The results obtained in calculating the $R_+ = \mathrm{Re}Z_+$ surface impedance of the plate as a function of H in the region of weak fields are shown in Fig. 8. The calculations were performed at a 50-kHz exciting field frequency and a plate $d = 1.6$ mm thick. Curve *1* describes the $R_+(H)$ function in the linear regime, and curve *2* was obtained for an $H_a = 200$ Oe exciting field amplitude. Curve *1* corresponds to a monotonic function of H, whereas curve *2* contains well-defined oscillations caused by the passage of the doppleron related to the third DSCR harmonic through the plate. It follows that nonlinearity again plays an important role. The capture of holes by the wave field suppresses cyclotron damping of the doppleron and makes the aluminum plate more “transparent” to radio-frequency radiation.

ACKNOWLEDGMENTS

This work was financially supported by the Russian Foundation for Basic Research (project no. 98-02-18393).

REFERENCES

1. V. G. Skobov and A. S. Chernov, Zh. Éksp. Teor. Fiz. **109**, 992 (1996) [JETP **82**, 535 (1996)].
2. P. K. Larsen and F. C. Greisen, Phys. Status Solidi B **45**, 363 (1971).
3. S. Balibar, B. Perrin, and A. Libchaber, J. Phys. F **2**, 4629 (1972).
4. O. V. Konstantinov, V. G. Skobov, V. V. Lavrova, *et al.*, Zh. Éksp. Teor. Fiz. **63**, 224 (1972) [Sov. Phys. JETP **36**, 118 (1973)].
5. I. F. Voloshin, G. A. Vugal’ter, V. Ya. Demikhovskiĭ, *et al.*, Zh. Éksp. Teor. Fiz. **73**, 1503 (1977) [Sov. Phys. JETP **46**, 790 (1977)].
6. I. F. Voloshin, S. V. Medvedev, V. G. Skobov, *et al.*, Zh. Éksp. Teor. Fiz. **71**, 1555 (1976) [Sov. Phys. JETP **44**, 814 (1976)].

Translated by V. Sipachev

*Journal of Experimental and Theoretical Physics, Vol. 92, No. 2, 2001, pp. 345–351.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 119, No. 2, 2001, pp. 396–402.
Original Russian Text Copyright © 2001 by Ovchinnikov.*

SOLIDS
Electronic Properties

The Structure of the Vortex Lattice in Anisotropic Superconductors

Yu. N. Ovchinnikov
Max-Planck-Institute for Physics of Complex Systems, D-01187, Dresden, Germany
Landau Institute for Theoretical Physics, Russian Academy of Sciences, Moscow, 117940 Russia
e-mail: ovchin@labs.polycnrs-gre.fr

Received September 21, 2000

Abstract—The structure of the vortex lattice in anisotropic superconductors at an arbitrary temperature in magnetic fields close to critical was studied. Generally, a rhombohedral structure with a vertex angle depending on temperature, magnetic field, and material constants is formed. An important factor is the small (2%) difference of the free energies of the triangular and square lattices in the Ginzburg–Landau approximation. This factor also persists in anisotropic superconductors. *© 2001 MAIK “Nauka/Interperiodica”.*

1. INTRODUCTION

In type II isotropic superconductors, the transition to the superconducting state in a magnetic field occurs as a second-order phase transition with the formation of a vortex lattice [1]. It is clear from symmetry considerations that a vortex lattice can either be square or triangular. It has been found that, near the T_c transition temperature, the triangular lattice with one flux quantum per unit cell corresponds to a free energy minimum, whereas the square lattice is a saddle point [2]. This property of isotropic superconductors persists at arbitrary temperatures except in a narrow region of superconductor physical parameters within which the transition to the superconducting state may be more complex in character [3].

For numerical reasons, the difference of the free energies of the triangular and square lattices amounts to about two percent of the energy of the superconducting transition [1, 2]. It would therefore be of interest to study the influence of anisotropy on the structure of the vortex lattice, even if this influence is weak. Clearly, the vortex lattice of an anisotropic superconductor should be rhombohedral and have parameters depending on the temperature and magnetic field strength.

Below, we show how vortex lattice parameters depend on temperature and superconductor material parameters. We consider the most physically interesting case of a “dirty” superconductor to show that vortex lattice parameters should be determined taking into account free energy terms of higher orders with respect to the electron mean free path. In an anisotropic superconductor, the contribution of these correction terms to the equation for the angle made by the unit cell vectors may be not small.

2. A SYSTEM OF EQUATIONS FOR GREEN FUNCTIONS IN AN ANISOTROPIC SUPERCONDUCTOR

In the quasi-classical approximation, the system of equations for Green functions in an anisotropic superconductor has the same form as in an isotropic superconductor [4–6],

$$\left(\mathbf{v}\frac{\partial}{\partial \mathbf{r}}\right)\hat{G} + \hat{\omega}\hat{G} - \hat{G}\hat{\omega} = 0,$$

$$\hat{G} = \begin{pmatrix} \alpha_p & -i\beta_p \\ i\tilde{\beta}_p & -\alpha_p \end{pmatrix}, \quad \hat{G}^2 = 1, \qquad (1)$$

where

$$\hat{\omega} = \omega\tau_z - ie\mathbf{v}\cdot\mathbf{A}\tau_z - i\hat{\Delta}$$

$$- in\hat{\sum_{pp}} + \frac{1}{2\tau_s}\int\frac{d\Omega_{p_1}}{4\pi}\tau_z G_{p_1}\tau_z, \qquad (2)$$

$$\hat{\Delta} = \begin{pmatrix} 0 & \Delta \\ -\Delta^* & 0 \end{pmatrix},$$

$$\sum_{pp} = -\frac{i\nu}{2}\int\sigma_{pp_1}\hat{G}_{p_1}d\Omega_{p_1}. \qquad (3)$$

In (1)–(3), $\mathbf{v} = \partial\varepsilon_{\mathbf{p}}/\partial\mathbf{p}$ is the velocity on the Fermi surface, and the $\mathbf{p}$, $\mathbf{p}_1$ momentum values also depend on the direction. A solution to (1) is sought in the form

$$\hat{G} = \hat{G}_0 + \hat{G}_1 + \hat{G}_2 + \hat{G}_3 + \dots \qquad (4)$$

The $\hat{G}_0$ Green function in (4) is independent of vector $\mathbf{p}$ angles, and the $G_{1,2,3}$ values do not contain the mean value and are approximations of the first, second, and third order in the electron mean free path, respectively. Equations (1)–(4) give

$$\begin{aligned}
\hat{G}_1 &= -\tau_{tr}\hat{G}_0\mathbf{v}\cdot\mathbf{D}\hat{G}_0 \\
&\equiv -\tau_{tr}\hat{G}_0\left\{\left(\mathbf{v}\frac{\partial}{\partial\mathbf{r}}\right)\hat{G}_0 - ie\mathbf{v}\cdot\mathbf{A}(\tau_z\hat{G}_0-\hat{G}_0\tau_z)\right\}, \\
\hat{G}_2 &= \tau_{tr}\tilde{\tau}_2\hat{G}_0\{\mathbf{v}\cdot\mathbf{D}(\hat{G}_0\mathbf{v}\cdot\mathbf{D}\hat{G}_0) \\
&\quad -\langle\mathbf{v}\cdot\mathbf{D}(\hat{G}_0\mathbf{v}\cdot\mathbf{D}\hat{G}_0)\rangle\}, \\
\hat{G}_3 &= -\tau_{tr}\{\hat{G}_0\mathbf{v}\cdot\mathbf{D}\hat{G}_2\}_1-\tilde{\tau}_3\{\hat{G}_0\mathbf{v}\cdot\mathbf{D}\hat{G}_2\}_3, \\
\tau_{tr}^{-1} &= \frac{1}{\tau}-\frac{1}{\tau_1}, \quad \tilde{\tau}_2^{-1} = \frac{1}{\tau}-\frac{1}{\tau_2}, \quad \tilde{\tau}_3^{-1} = \frac{1}{\tau}-\frac{1}{\tau_3}.
\end{aligned} \tag{5}$$

The $\mathbf{D}$ operator in (5) is defined by the upper formula, and τ, τ_1, and τ_2 are the collision times for the first, second, etc. harmonics [3]. Under isotropic scattering conditions, all these times except τ are zero. The $\{\}_{1,3}$ symbol denotes the projection of the expression in braces onto the first and third harmonics, respectively. Substituting (5) into (1) yields a closed equation only for the $\hat{G}_0$ function. As usual, $\hat{G}_0$ will be written in the form

$$\hat{G}_0 = \begin{pmatrix} \alpha & -i\beta \\ i\beta^* & -\alpha \end{pmatrix}, \quad \alpha^2+|\beta|^2 = 1. \tag{6}$$

Equation (6) reduces the equation for the $\hat{G}_0$ Green function to one differential equation for the α and β functions and the normalization condition (6),

$$\begin{aligned}
&\alpha\Delta-\beta\omega+\frac{1}{2}\tau_{tr}\Big[\alpha\langle(\mathbf{v}\partial_-)^2\rangle\beta-\beta\Big(\mathbf{v}\frac{\partial}{\partial\mathbf{r}}\Big)^2\alpha\Big] \\
&+\frac{1}{2}\tau_{tr}^2\tilde{\tau}_2\Big\langle(\mathbf{v}\partial_-)^2\Big[\Big(\alpha(\mathbf{v}\partial_-)^2\beta-\beta\Big(\mathbf{v}\frac{\partial}{\partial\mathbf{r}}\Big)^2\alpha\Big) \\
&-\Big\langle\alpha(\mathbf{v}\partial_-)^2\beta-\beta\Big(\mathbf{v}\frac{\partial}{\partial\mathbf{r}}\Big)^2\alpha\Big\rangle\Big] \\
&+\Big[\Big(\mathbf{v}\frac{\partial\alpha}{\partial\mathbf{r}}\Big)^2+(\mathbf{v}\partial_-\beta)(\mathbf{v}\partial_+\beta^*)\Big] \\
&\times\Big[\Big(2\alpha(\mathbf{v}\partial_-)^2\beta-2\beta\Big(\mathbf{v}\frac{\partial}{\partial\mathbf{r}}\Big)^2\alpha\Big) \\
&-\Big\langle\alpha(\mathbf{v}\partial_-)^2\beta-\beta\Big(\mathbf{v}\frac{\partial}{\partial\mathbf{r}}\Big)^2\alpha\Big\rangle\Big] \\
&-2\Big((\mathbf{v}\partial_-\beta)(\mathbf{v}\partial_+\beta^*)+\Big(\mathbf{v}\frac{\partial\alpha}{\partial\mathbf{r}}\Big)^2\Big) \\
&\times\Big(\alpha(\mathbf{v}\partial_-)^2\beta-\beta\Big(\mathbf{v}\frac{\partial}{\partial\mathbf{r}}\Big)^2\alpha\Big) \\
&-\Big\langle\Big(\mathbf{v}\frac{\partial\alpha}{\partial\mathbf{r}}\Big)^2+(\mathbf{v}\partial_-\beta)(\mathbf{v}\partial_+\beta^*)\Big\rangle \\
&\times\Big(\alpha(\mathbf{v}\partial_-)^2\beta-\beta\Big(\mathbf{v}\frac{\partial}{\partial\mathbf{r}}\Big)^2\alpha\Big) \\
&-\Big[2\alpha\Big(\mathbf{v}\frac{\partial}{\partial\mathbf{r}}\Big)^2\alpha+\beta(\mathbf{v}\partial_+)^2\beta^*+\beta^*(\mathbf{v}\partial_-)^2\beta\Big] \\
&\times\Big\langle\alpha(\mathbf{v}\partial_-)^2\beta-\beta\Big(\mathbf{v}\frac{\partial}{\partial\mathbf{r}}\Big)^2\alpha\Big\rangle \\
&+\Big(\alpha(\mathbf{v}\partial_-\beta)-\beta\Big(\mathbf{v}\frac{\partial\alpha}{\partial\mathbf{r}}\Big)\Big)\Big(\mathbf{v}\frac{\partial}{\partial\mathbf{r}}\Big) \\
&\times\Big[\Big(\mathbf{v}\frac{\partial\alpha}{\partial\mathbf{r}}\Big)^2+(\mathbf{v}\partial_-\beta)(\mathbf{v}\partial_+\beta^*) \\
&-\Big\langle\Big(\mathbf{v}\frac{\partial\alpha}{\partial\mathbf{r}}\Big)^2+(\mathbf{v}\partial_-\beta)(\mathbf{v}\partial_+\beta^*)\Big\rangle\Big] \\
&+\Big(\alpha\Big(\mathbf{v}\frac{\partial}{\partial\mathbf{r}}\Big)\alpha+\beta(\mathbf{v}\partial_+)\beta^*\Big)(\mathbf{v}\partial_-)\Big[\alpha(\mathbf{v}\partial_-)^2\beta \\
&-\beta\Big(\mathbf{v}\frac{\partial}{\partial\mathbf{r}}\Big)^2\alpha-\Big\langle\alpha(\mathbf{v}\partial_-)^2\beta-\beta\Big(\mathbf{v}\frac{\partial}{\partial\mathbf{r}}\Big)^2\alpha\Big\rangle\Big] \\
&+\Big(\alpha(\mathbf{v}\partial_-)\beta-\beta\Big(\mathbf{v}\frac{\partial\alpha}{\partial\mathbf{r}}\Big)\Big)\Big(\mathbf{v}\frac{\partial}{\partial\mathbf{r}}\Big) \\
&\times\Big[\beta^*(\mathbf{v}\partial_-)^2\beta+\alpha\Big(\mathbf{v}\frac{\partial}{\partial\mathbf{r}}\Big)^2\alpha \\
&-\Big\langle\alpha\Big(\mathbf{v}\frac{\partial}{\partial\mathbf{r}}\Big)^2\alpha+\beta^*(\mathbf{v}\partial_-)^2\beta\Big\rangle\Big]\Big\rangle = \alpha\beta\Gamma,
\end{aligned} \tag{7}$$

where $\Gamma = \tau_s^{-1}$ and $\partial_\pm = \partial/\partial\mathbf{r} \pm 2ie\mathbf{A}$. Equation (7) contains the mean products of two and four velocities. Let us use the notation

$$D_{ij} = \tau_{tr}\langle v_i v_j\rangle, \quad \Gamma_{ijkl} = \tau_{tr}^2\tilde{\tau}_2\langle v_i v_j v_k v_l\rangle. \tag{8}$$

Our task is to obtain an expression for the free energy from (7). Near the transition temperature, this task can be fulfilled at an arbitrary external magnetic field value [3]. At an arbitrary temperature, the problem can only be solved for fields close to critical. The α and β Green functions can then be expanded in powers of the order parameter Δ. This yields

$$F_s - F_N = \nu \int d^3 r \left\{ \frac{|\Delta|^2}{|\lambda|} \right.$$
$$- 2\pi\tau \sum_{\omega > 0} \left[\Delta^* \left(\Gamma + \omega - \frac{1}{2} D_{ij} (\partial_-^0)_i (\partial_-^0)_j \right. \right.$$
$$- \frac{1}{2} \Gamma_{ijkl} (\partial_-^0)_i (\partial_-^0)_j (\partial_-^0)_k (\partial_-^0)_l$$
$$\left. + \frac{1}{2} \tilde{\tau}_2 D_{ij} D_{kl} (\partial_-^0)_i (\partial_-^0)_j (\partial_-^0)_k (\partial_-^0)_l \right)^{-1} \Delta$$
$$- \frac{\omega |\Delta|^4}{4(\omega + \Gamma + \hat{\lambda})^4} - \frac{D_{ij}}{8(\omega + \Gamma + \hat{\lambda})^4} \frac{\partial |\Delta|^2}{\partial r_i} \frac{\partial |\Delta|^2}{\partial r_j}$$
$$- \frac{ie(A_1)_i D_{ij}}{(\omega + \Gamma + \hat{\lambda})^2} (\Delta^* (\partial_-^0)_j \Delta - \Delta (\partial_+^0) j \Delta^*) \tag{9}$$
$$+ \frac{\Gamma_{ijkl}}{(\omega + \Gamma + \hat{\lambda})^4} \left(\frac{1}{8} \frac{\partial^2 |\Delta|^2}{\partial r_i \partial r_j} \frac{\partial^2 |\Delta|^2}{\partial r_k \partial r_l} \right.$$
$$- \frac{1}{2} ((\partial^0)_i \Delta)((\partial_-^0)_j \Delta)((\partial_+^0)_k \Delta^*)((\partial_+^0)_l \Delta^*)$$
$$+ \frac{1}{8} \frac{\partial^2 |\Delta|^2}{\partial r_i \partial r_j} ((\partial_-^0)_k \Delta)((\partial_+^0)_l \Delta^*)$$
$$\left. \left. \left. - \frac{|\Delta|^2}{4} ((\partial_-^0)_i (\partial_-^0)_j \Delta)((\partial_+^0)_k (\partial_+^0)_l \Delta^*) \right) \right] \right\}$$
$$+ \frac{1}{8\pi} \int d^3 r (\operatorname{rot} \mathbf{A} - H_0)^2.$$

Here, |λ| is the effective interaction constant. The vector potential A is selected in the form

$$A = A_0 + A_1 + \ldots, \quad A_0 = (0, Bx, 0), \tag{10}$$

where $A_{1,2,\ldots}$ are bounded functions. The $\hat{L}$ operator is defined as $\partial_-^0 = \partial/\partial\mathbf{r} - 2ie\mathbf{A}_0$, and $\hat{\lambda}$ is the smallest eigenvalue of the operator

$$\hat{L} = -\frac{1}{2} D_{ij} (\partial_-^0)_i (\partial_-^0)_j$$
$$- \frac{1}{2} (\Gamma_{ijkl} - \tilde{\tau}_2 D_{ij} D_{kl}) (\partial_-^0)_i (\partial_-^0)_j (\partial_-^0)_k (\partial_-^0)_l. \tag{11}$$

3. THE STRUCTURE
OF THE VORTEX LATTICE NEAR H_{c2}

Further, it will be convenient to pass to the principal axes of the D_{ij} tensor. These axes will (arbitrarily for the present) be denoted by a, b, and c. The magnetic field is assumed to be directed along one of the principal axes, suppose, axis c. Let the order parameter (Δ) be written as a series in powers of $H_{c2} - B$:

$$\Delta = \Delta_0 + \Delta_1 + \Delta_2 + \ldots, \tag{12}$$

where Δ_1, Δ_2, … are orthogonal to Δ_0.

In the main approximation with respect to the electron free path,

$$\Delta_0 = \sum_N C_N \exp\left\{ 2ieBNx_1 y - B \sqrt{\frac{D_b}{D_a}} (x - Nx_1)^2 \right\}. \tag{13}$$

Suppose that the modulus of the order parameter forms a periodic lattice with unit cell vectors $\mathbf{a}_{1,2}$; that is,

$$|\Delta(r - n\mathbf{a}_1 - m\mathbf{a}_2)|^2 = |\Delta(r)|^2, \tag{14}$$

where n and m are integers.

The $\mathbf{a}_{1,2}$ vectors will be selected in the form

$$\mathbf{a}_1 = a(\sin\varphi, \cos\varphi), \quad \mathbf{a}_2 = a(-\sin\varphi, \cos\varphi). \tag{15}$$

Suppose that there is one flux quantum in the unit cell. Periodicity condition (14) then yields

$$eBa^2 \sin(2\varphi) = \pi, \quad a\sin\varphi = x_1,$$
$$C_N = C_0 \exp\left(\frac{i\pi}{2} N^2\right). \tag{16}$$

The $\mathbf{K}_{1,2}$ reciprocal lattice vectors will be selected as

$$\mathbf{K}_1 = \frac{2\pi}{a\sin 2\varphi} (\cos\varphi, \sin\varphi),$$
$$\mathbf{K}_2 = \frac{2\pi}{a\sin 2\varphi} (\cos\varphi; -\sin\varphi). \tag{17}$$

In the main approximation with respect to the free path and the $(H_{c2} - B)$ parameter, the current density j is given by the equation

$$j = -ie\nu \hat{D} 2\pi T \sum_{\omega > 0} \frac{\Delta^* \partial_-^0 \Delta - \Delta \partial_+^0 \Delta^*}{(\omega + \Gamma + \hat{\lambda})^2}. \tag{18}$$

Using (13), we obtain the important relation

$$\operatorname{rot}(0, 0, |\Delta|^2) = \frac{i}{\sqrt{D_a D_b}} \{ \hat{D} (\Delta^* \partial_- - \Delta \partial_+ \Delta^*) \}. \tag{19}$$

In the geometry under consideration, the magnetic field is directed along the z axis. Equation (19) enables us to obtain an explicit expression for correction H_1 to the magnetic field. It has the form

$$\operatorname{rot} \mathbf{A}_1 = H_1 (0, 0, 1),$$
$$H_1 = -4\pi e\nu \left[\frac{1}{2\pi T} \psi' \left(\frac{1}{2} + \frac{\Gamma(T) + \hat{\lambda}}{2\pi T} \right) \right] \tag{20}$$
$$\times \sqrt{D_a D_b} (|\Delta|^2 - \langle |\Delta|^2 \rangle),$$

where $\psi(x)$ is the Euler psi function.

Taking (20) into account, (9) can be rewritten as

$$
\begin{aligned}
\frac{F_s - F_N}{V} &= \nu\Bigg\{\langle|\Delta|^2\rangle\Bigg[-\ln\left(\frac{T_c^0}{T}\right) \\
&+ \psi\left(\frac{1}{2} + \frac{\Gamma(T)+\hat{\lambda}}{2\pi T}\right) - \psi(1/2)\Bigg] \\
&- \beta_A \frac{\langle|\Delta|^2\rangle^2}{8(2\pi T)^2}\Bigg[\psi''\left(\frac{1}{2} + \frac{\Gamma(T)+\hat{\lambda}}{2\pi T}\right) \\
&+ \frac{\Gamma(T)+\hat{\lambda}}{6\pi T}\psi'''\left(\frac{1}{2} + \frac{\Gamma(T)+\hat{\lambda}}{2\pi T}\right)\Bigg] \\
&- (\beta_A - 1)D_a D_b \langle|\Delta|^2\rangle^2 \frac{e^2\nu}{2\pi T^2}\left(\psi'\left(\frac{1}{2} + \frac{\Gamma(T)+\hat{\lambda}}{2\pi T}\right)\right)^2\Bigg\} \\
&- \frac{H_0 B}{4\pi} + \frac{B^2 + H_0^2}{8\pi} \\
&+ \frac{\nu}{48(2\pi T)^3}\psi'''\left(\frac{1}{2} + \frac{\Gamma(T)+\hat{\lambda}}{2\pi T}\right) D_{ij}\left\langle \frac{\partial|\Delta|^2}{\partial r_i}\frac{\partial|\Delta|^2}{\partial r_j}\right\rangle \\
&- \frac{\nu}{6(2\pi T)^3}\psi'''\left(\frac{1}{2} + \frac{\Gamma(T)+\hat{\lambda}}{2\pi T}\right)\Gamma_{ijkl}\Bigg\langle \frac{1}{8}\frac{\partial^2|\Delta|^2}{\partial r_i \partial r_j}\frac{\partial^2|\Delta|^2}{\partial r_k \partial r_l} \\
&- \frac{1}{2}((\partial_-^0)_i\Delta)((\partial_-^0)_j\Delta)((\partial_+^0)_k\Delta^*)((\partial_+^0)_l\Delta^*) \\
&+ \frac{1}{8}\frac{\partial^2|\Delta|^2}{\partial r_i \partial r_j}((\partial_-^0)_k\Delta)((\partial_+^0)_l\Delta^*) \\
&- \frac{|\Delta|^2}{4}((\partial_-^0)_i(\partial_-^0)_j\Delta)((\partial_+^0)_k(\partial_+^0)_l\Delta^*)\Bigg\rangle.
\end{aligned}
\tag{21}
$$

Here, V is the volume of the superconductor, T_c^0 is the temperature of the superconducting transition in a superconductor free of paramagnetic admixtures, and $\beta_A = \langle|\Delta|^4\rangle/\langle|\Delta|^2\rangle^2$ is the Abrikosov structure constant.

The mean values present in (21) will be calculated with the help of the relations

$$
\begin{aligned}
(\partial_-^0)_y\Delta &= i\sqrt{\frac{D_a}{D_b}}(\partial_-^0)_x\Delta, \\
|\Delta|^2 &= \sum_{N,M} C_{NM}\exp\{i(N\mathbf{K}_1 + M\mathbf{K}_2)\cdot\mathbf{r}\}, \\
&((\partial_-^0)_x\Delta)(\partial_+^0)_x\Delta^* \\
&= \sum_{N,M}\tilde{C}_{NM}\exp\{i(N\mathbf{K}_1 + M\mathbf{K}_2)\cdot\mathbf{r}\}, \\
&((\partial_-^0)_x^2\Delta)(\partial_+^0)_x^2\Delta^* \\
&= \sum_{N,M} A_{NM}\exp\{i(N\mathbf{K}_1 + M\mathbf{K}_2)\cdot\mathbf{r}\}.
\end{aligned}
\tag{22}
$$

The C_{NM}, $\tilde{C}_{NM}$, and A_{NM} coefficients can be found from use of the explicit expression (13) for the order parameter (Δ). This yields

$$
\begin{aligned}
C_{NM} &= \langle|\Delta_0|^2\rangle(-)^{M(M-N)}\exp(-\pi Q_{NM}), \\
\tilde{C}_{NM} &= \langle|\Delta_0|^2\rangle(-)^{M(M-N)}eB\sqrt{\frac{D_b}{D_a}} \\
&\times[1 - 2\pi Q_{NM}]\exp(-\pi Q_{NM}), \\
A_{NM} &= \langle|\Delta_0|^2\rangle(-)^{M(M-N)}e^2B^2\frac{D_b}{D_a} \\
&\times\Bigg\{3 + 8\pi^2\left(\frac{M^2 - N^2}{4}\right)^2 + 4\pi^2\left(\frac{D_b}{D_a}\left(\frac{M-N}{2}\right)^4\tan^2\varphi \right.\\
&\left.+ \frac{D_a}{D_b}\left(\frac{M+N}{2}\right)^4\right)\cot^2\varphi - 4\pi\sqrt{\frac{D_a}{D_b}}\left(\frac{M+N}{2}\right)^2\cot\varphi \\
&- 12\pi\sqrt{\frac{D_b}{D_a}}\left(\frac{M-N}{2}\right)^2\tan\varphi\Bigg\},
\end{aligned}
\tag{23}
$$

where

$$
\begin{aligned}
Q_{NM} &= \sqrt{\frac{D_b}{D_a}}\left(\frac{M-N}{2}\right)^2\tan\varphi \\
&+ \sqrt{\frac{D_a}{D_b}}\left(\frac{M+N}{2}\right)^2\cot\varphi.
\end{aligned}
\tag{24}
$$

Using (22) and (23), we eventually recast Eq. (21) for the free energy density into the form

$$
\begin{aligned}
\frac{F_S - F_N}{V} &= \frac{(B - H_0)^2}{8\pi} \\
&+ \nu\Bigg\{-\langle|\Delta|^2\rangle\psi'\left(\frac{1}{2} + \frac{\Gamma+\hat{\lambda}}{2\pi T}\right)\frac{\hat{\lambda}_{cr} - \hat{\lambda}}{2\pi T} \\
&- \beta_A\frac{\langle|\Delta|^2\rangle^2}{8(2\pi T)^2}\Bigg[\psi''\left(\frac{1}{2} + \frac{\Gamma(T)+\hat{\lambda}}{2\pi T}\right) \\
&+ \frac{\Gamma(T)+\hat{\lambda}}{6\pi T}\psi'''\left(\frac{1}{2} + \frac{\Gamma(T)+\hat{\lambda}}{2\pi T}\right)\Bigg]
\end{aligned}
$$

$$-(\beta_A-1)D_aD_b\langle|\Delta|^2\rangle^2\frac{e^2\nu}{2\pi T^2}\left(\psi'\left(\frac{1}{2}+\frac{\Gamma(T)+\hat{\lambda}}{2\pi T}\right)\right)^2\Bigg\} \quad (25)$$

$$+\frac{\nu\hat{\lambda}\langle|\Delta|^2\rangle^2}{48\pi^2T^3}\psi'''\left(\frac{1}{2}+\frac{\Gamma+\hat{\lambda}}{2\pi T}\right)S$$

$$-\frac{\nu\langle|\Delta|^2\rangle^2\hat{\lambda}^2}{6(2\pi T)^3}\psi'''\left(\frac{1}{2}+\frac{\Gamma+\hat{\lambda}}{2\pi T}\right)\Bigg\{\left(\frac{\Gamma_{aaaa}}{D_a^2}-\frac{\Gamma_{bbbb}}{D_b^2}\right)S_1$$

$$-\left(\frac{\Gamma_{aaaa}}{D_a^2}+\frac{\Gamma_{bbbb}}{D_b^2}-6\frac{\Gamma_{aabb}}{D_aD_b}\right)S_2$$

$$+\left(\frac{\Gamma_{aaaa}}{D_a^2}+\frac{\Gamma_{bbbb}}{D_b^2}-2\frac{\Gamma_{aabb}}{D_aD_b}\right)S_3\Bigg\},$$

where

$$S = \sum_{N,M}Q_{NM}\exp(-2\pi Q_{NM}),$$

$$\beta_A = \sum_{N,M}\exp(-2\pi Q_{NM}),$$

$$S_1 = \sum_{N,M}\Bigg\{5\pi^2\left(\frac{D_a}{D_b}\left(\frac{N+M}{2}\right)^4\cot^2\varphi\right.$$

$$-\frac{D_b}{D_a}\left(\frac{M-N}{2}\right)^4\tan^2\varphi-\frac{3\pi}{2}\left(\sqrt{\frac{D_a}{D_b}}\left(\frac{N+M}{2}\right)^2\cot\varphi\right.$$

$$\left.\left.-\sqrt{\frac{D_b}{D_a}}\left(\frac{M-N}{2}\right)^2\tan\varphi\right)\right\}\exp(-2\pi Q_{NM}), \quad (26)$$

$$S_2 = 8\pi^2\sum_{N,M}\left(\frac{M^2-N^2}{4}\right)^2\exp(-2\pi Q_{NM}),$$

$$S_3 = 2\pi^2\sum_{N,M}Q_{NM}^2\exp(-2\pi Q_{NM}),$$

$$\hat{\lambda}_{cr}-\hat{\lambda} = e(H_{c2}-B)\sqrt{D_aD_b},$$

$$\hat{\lambda} = eB\sqrt{D_aD_b}.$$

In (26), we omit the terms responsible for the small renormalization of the coefficients of the β_A and S structure sums. All structure constants (β_A, S, $S_{1,2,3}$) present in Eq. (25) for the free energy and determining the vortex lattice structure only depend on the parameter

$$z = \sqrt{\frac{D_b}{D_a}}\tan\varphi. \quad (27)$$

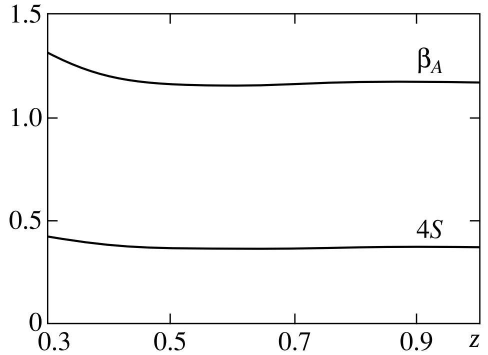


Fig. 1. Dependences of the β_A and $4S$ functions on the $z = \sqrt{D_b/D_a}\tan\varphi$ parameter.

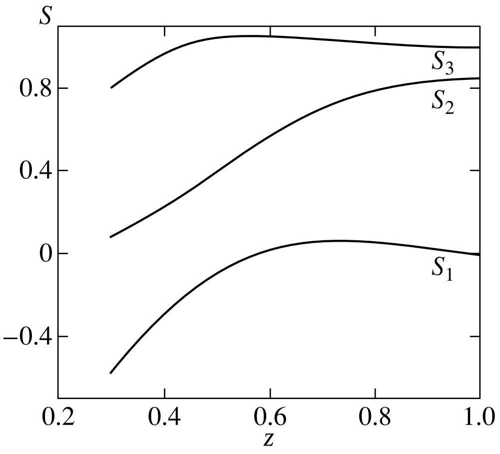


Fig. 2. Dependences of the S_1, S_2, and S_3 functions on the $z = \sqrt{D_b/D_a}\tan\varphi$ parameter.

The $z \longrightarrow 1/z$ replacement transforms the β_A, S, S_2, and S_3 functions into themselves and changes the sign of S_1. The β_A, S, S_1, S_2, and S_3 functions in the z range 0.3–1 are plotted in Figs. 1 and 2. It follows from symmetry considerations that the β_A, S, and S_3 structure sums have extrema at $z = 1$ and $z = 1$, $z = \sqrt{3}$, $1/\sqrt{3}$. In isotropic superconductors, the coefficients of S_1 and S_2 vanish.

4. EQUATIONS FOR THE $\langle|\Delta|^2\rangle$ AND z PARAMETERS

The $\langle|\Delta|^2\rangle$, z, and B values are free parameters in Eq. (25) for the free energy. They are determined from the free energy extremum conditions with respect to these parameters; that is,

$$\partial\left(\frac{F_S-F_N}{V}\right)\partial\langle|\Delta|^2\rangle = 0,$$

Table

	β_A	$4S$	S_1	S_2	S_3
1	1.18034	0.375714	0	0.85416	1.001706
$1/\sqrt{3}$	1.159595	0.3691106	0	0.5281606	1.056321

$$\partial\left(\frac{F_S - F_N}{V}\right)/\partial z = 0, \qquad (28)$$

$$\partial\left(\frac{F_S - F_N}{V}\right)/\partial B = 0.$$

Suppose that the temperature is close to T_c to the extent that only the highest terms in $1 - T/T_c$ may be retained in (25). The β_A value should then be extremal [1]; that is,

$$\sqrt{\frac{D_b}{D_a}}\tan\varphi = \{1, \sqrt{3}, 1/\sqrt{3}\}. \qquad (29)$$

Point 1 (square lattice in the isotropic case) corresponds to the saddle point (maximum with respect to z). It follows that, in this simple situation, a rhombohedral structure is formed with the φ angle given by

$$\sqrt{\frac{D_b}{D_a}}\tan\varphi = \{\sqrt{3}, 1/\sqrt{3}\}. \qquad (30)$$

As mentioned above, the β_A and S values vary very slowly over the whole interval (29). For this reason, the correction term in Eq. (25) for the free energy may become significant in the second equation in (28):

$$\begin{aligned}
&-\frac{\partial\beta_A}{\partial z}\left\{\frac{1}{8}\left[\psi''(x) + \frac{x-0.5}{3}\psi'''(x)\right]\right.\\
&\left.+ 2\pi e^2\nu D_a D_b(\psi'(x))^2\right\} + \frac{\partial S}{\partial z}\frac{\hat{\lambda}}{12T}\psi'''(x)\\
&-\frac{\hat{\lambda}^2}{12\pi T}\psi'''(x)\left\{\left(\frac{\Gamma_{aaaa}}{D_a^2} - \frac{\Gamma_{bbbb}}{D_b^2}\right)\frac{\partial S_1}{\partial z}\right. \qquad (31)\\
&-\left(\frac{\Gamma_{aaaa}}{D_a^2} + \frac{\Gamma_{bbbb}}{D_b^2} - 6\frac{\Gamma_{aabb}}{D_a^2}\right)\frac{\partial S_2}{\partial z}\\
&\left.+\left(\frac{\Gamma_{aaaa}}{D_a^2} + \frac{\Gamma_{bbbb}}{D_b^2} - 2\frac{\Gamma_{aabb}}{D_a D_b}\right)\frac{\partial S_3}{\partial z}\right\} = 0,
\end{aligned}$$

where $x = 1/2 + (\Gamma(T) + \hat{\lambda})/2\pi T$.

The first equation in (28) does not contain numerical smallness and can therefore be written in the main approximation with respect to the free path. This yields

$$\begin{aligned}
\langle|\Delta|^2\rangle &= 2\pi Te(H_{c2} - B)\sqrt{D_a D_b}\psi'(x)\\
&\times\left\{-\frac{\beta_A}{4}\left[\psi''(x) + \frac{x-1/2}{3}\psi'''(x)\right]\right. \qquad (32)\\
&\left.-4\pi(\beta_A - 1)D_a D_b e^2\nu + \frac{\hat{\lambda}}{6T}\psi'''(x)S\right\}^{-1}.
\end{aligned}$$

Equations (31) and (32) determine the z and $\langle|\Delta|^2\rangle$ parameter values and, through this, the free energy (25) value. The β_A and S functions have maxima at $z = 1$ and minima at $z = \{1/\sqrt{3}, \sqrt{3}\}$. The S_3 function has a minimum at $z = 1$ and maxima at $z = \{1/\sqrt{3}, \sqrt{3}\}$. It follows from (26) that the β_A, S, and S_3 functions are proportional to $z^{-1/2}$ when z tends to zero. The S_3 function therefore has two more minima. According to (25), taking into account correction free path terms only deepens the minimum corresponding to the triangular lattice in isotropic superconductors. Generally, the $z = 1$ point is not extremal in an anisotropic superconductor. The positions of all three minima change. This removes degeneracy, and one of the minima becomes a local minimum. More than three solutions to (31) may exist.

The β_A, S, S_1, S_2, and S_3 function values at the $\{1, 1/\sqrt{3}\}$ points are listed in the table.

5. CONCLUSION

We studied the structure of the vortex lattice in anisotropic superconductors at arbitrary temperatures in magnetic fields close to critical. If the magnetic field is oriented along one of the principal axes, the determining factor is anisotropy in the plane perpendicular to the magnetic field. Even weak anisotropy can substantially change the angle of the rhombohedral unit cell. This angle is a function of temperature, field, and material constants. Indeed, in the Ginzburg–Landau approximation, the difference of the free energies of the square and triangular sublattices in isotropic superconductors only amounts to 2% of the superconducting transition energy. Nothing significant occurs in isotropic superconductors as temperature lowers. In anisotropic superconductors, the degeneracy related to the $z \longrightarrow 1/z$ transformation is removed, and two identical states become different. One of these corresponds to only a local minimum, and the other, to the ground state. These states are separated by a barrier, and the transition between them is a first-order transition. It is quite possible that a change in the field can cause a transition between these two states [7]. Note also that Eq. (31) for the vertex angle of the rhombohedral structure can have more than three solutions even in fields close to H_{c2}. Studies of weak fields $H \ll H_{c2}$ involve serious difficulties, because the energy only depends on the density of vortices in the main approximation [2]. In our view, a simplified approach with a nonlocal generalization of the London approximation [8] is then unsatisfactory. It gives results that are at variance with experiment [7, 9].

ACKNOWLEDGMENTS

The author thanks L.Ya. Vinnikova for discussions. This work was financially supported by CRDF (USA) (Grant RP1-2251) and the Russian Foundation for Basic Research.

REFERENCES

1. A. A. Abrikosov, Zh. Éksp. Teor. Fiz. **32**, 1442 (1957) [Sov. Phys. JETP **5**, 1174 (1957)].
2. P. G. de Gennes, *Superconductivity of Metals and Alloys* (Benjamin, New York, 1966; Mir, Moscow, 1968).
3. Yu. N. Ovchinnikov, Zh. Éksp. Teor. Fiz. **115**, 726 (1999) [JETP **88**, 398 (1999)].
4. G. Eilanberger, J. Phys. **214**, 195 (1968).
5. A. I. Larkin and Yu. N. Ovchinnikov, Zh. Éksp. Teor. Fiz. **55**, 2262 (1968) [Sov. Phys. JETP **28**, 1200 (1968)].
6. Yu. N. Ovchinnikov and E. H. Brandt, Phys. Status Solidi B **67**, 301 (1975).
7. M. R. Eskildsen *et al.*, Physica B (Amsterdam) **284–288** (2000).
8. V. G. Kogan *et al.*, Phys. Rev. B **55**, R8693 (1997).
9. L. Ya. Vinnikov *et al.*, Physica B (Amsterdam) **284–288**, 813 (2000).

Translated by V. Sipachev

Journal of Experimental and Theoretical Physics, Vol. 92, No. 2, 2001, pp. 352–356.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 119, No. 2, 2001, pp. 403–408.
Original Russian Text Copyright © 2001 by Andreev, Savishchev.

SOLIDS
Electronic Properties

Motion of Charges and Singularities in the Vacancy Spectrum in Crystalline Helium

A. F. Andreev* and A. D. Savishchev
Kapitsa Institute for Physical Problems, Russian Academy of Sciences, ul. Kosygina 2, Moscow, 117334 Russia
**e-mail: andreev@kapitza.ras.ru*

Received October 9, 2000

Abstract—The Van Hove singularities in the vacancy spectrum of ^{4}He crystals are manifested as singularities in the field dependence of the drift velocity of charges. In ^{3}He, the field dependence of the drift velocity is determined by the vacancy spectrum in the completely spin-polarized state of the crystal. *© 2001 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

The quantum nature of ^{4}He crystals makes it possible to obtain a quite general theoretical description of the motion of charges in these crystals [1, 2] without using any model concepts of the structure of charge complexes. The description is essentially based on a single assumption concerning the vacancion mechanism of the motion of charges, which has received convincing experimental confirmation [3]. Owing to the quantum delocalization of vacancies in ^{4}He crystals and their transformation into freely moving quasiparticles, the vacancion mechanism of motion can be presented as successive acts of the inelastic scattering of vacancion quasiparticles at a charged complex, which are accompanied by the displacement of a complex from one localized state to other states crystallographically equivalent to it. The application of the quantum theory of scattering makes it possible to express the drift velocity of charges in terms of the general symmetry parameters of complexes as well as through the energy spectrum of vacancies. The experimental study of the motion of charges in this case is a sort of the spectroscopy of vacancion quasiparticles. On the other hand, it leads to definite conclusions concerning the structure and (which is especially important) symmetry of charge complexes.

Recent direct measurements of anisotropy in the drift velocity of charges in ^{4}He [4, 5] confirm the quantum-mechanical approach [1, 2] to the problem. For example, it is difficult to use the macroscopic description of the motion of charges [6] to explain the experimentally observed strong anisotropy in the drift velocity [4] as a diffusion–viscous vacancion flow of the crystal [7, 8]; however, this phenomenon can be interpreted in a simple and natural manner using the quantum-mechanical approach (see below). A quantitative interpretation of the results of experimental observations [4] requires that our previous results [2] be supplemented with a number of new concepts. The main goal of the present paper is to formulate these concepts and to compare them with the experimental data [4]. Preliminary results were reported by us earlier [5].

In Section 4, we will consider a peculiar pattern of the motion of charges in ^{3}He crystals. Keshishev [9] proved in his early works that a cubic field dependence of the velocity of the type $(E + E_0)^3$ (E_0 is a function of temperature and pressure) is observed both for ^{4}He (under not very low pressures if we are dealing with positive charges) and for ^{3}He. Our explanation [2] of this dependence was based on the assumption of the existence of vacancy quasimomentum as a good quantum number, which is valid for ^{4}He but is not correct for ^{3}He in view of its disordering in the values of nuclear spins of atoms. It will be proved below that the observed drift velocity of charges in ^{3}He in a wide range of electric field values coincides with the drift velocity in the completely spin-polarized state of the crystal (nuclear spins have the same direction), in which the vacancies behave in the same manner as in ^{4}He. The reason behind this interesting fact is the formation of macroscopic ferromagnetic polarons around vacancies in ^{3}He crystals (see [10]).

2. VAN HOVE SINGULARITIES

As in our previous publications [1, 2], we consider the simplest case when the main role is played by phonon-free one-vacancion processes in which the charge complex is displaced by vector $\mathbf{u}$ from its initially localized state as a result of the inelastic scattering of a vacancy. Probability W of the displacement of a charge per unit time is given by

$$W = \int w(\mathbf{k}, \mathbf{k}'; \mathbf{u}) n(\varepsilon) \frac{d^3k}{(2\pi)^3} \frac{d^3k'}{(2\pi)^3}, \tag{1}$$

where $\mathbf{k}$ and $\mathbf{k}'$ are the values of quasimomentum of the vacancy in the initial and final states, respectively;

$\varepsilon = \varepsilon(\mathbf{k})$ is the energy spectrum of vacancies; and $n(\varepsilon) = \exp(-\varepsilon/T)$ is the equilibrium function of vacancy distribution. The differential probability $w(\mathbf{k}, \mathbf{k}'; \mathbf{u})$ can be expressed in terms of the amplitude $f(\mathbf{k}, \mathbf{k}'; \mathbf{u})$ of the process through the familiar formula

$$w(\mathbf{k}, \mathbf{k}'; \mathbf{u}) = 2\pi|f(\mathbf{k}, \mathbf{k}'; \mathbf{u})|^2\delta(\varepsilon' - \varepsilon - e\mathbf{E}\cdot\mathbf{u}), \quad (2)$$

where $e\mathbf{E}\cdot\mathbf{u}$ is the change in the energy of the charge complex having charge e upon its displacement by vector $\mathbf{u}$ in the electric field $\mathbf{E}$, $\varepsilon' = \varepsilon(\mathbf{k}')$.

At low temperatures $T \ll \Delta$ and $T \ll e\mathbf{E}\mathbf{u}$, where $\Delta = \max(\varepsilon' - \varepsilon)$ is the width of the vacancy energy band, the major role is played by vacancies with $\mathbf{k} \approx \mathbf{k}_0$, where $\mathbf{k}_0$ is the value of the quasimomentum corresponding to the bottom of the band $\varepsilon_0 \equiv \varepsilon(\mathbf{k}_0) = \min\varepsilon(\mathbf{k})$, and by transitions with $e\mathbf{E}\cdot\mathbf{u} > 0$. The probabilities of reverse processes accompanied by a decrease in the energy of vacancies ($\varepsilon' < \varepsilon$) are exponentially small. In the expression for w in formulas (1) and (2), we can put $\mathbf{k} = \mathbf{k}_0$ and $\varepsilon = \varepsilon_0$. This gives

$$W = \frac{N}{(2\pi)^2}\int\frac{dS}{v}|f(\mathbf{k}_0, \mathbf{k}; \mathbf{u})|^2, \quad (3)$$

where $N = \int n(\varepsilon)\, d^3k/(2\pi)^3 \propto \exp(-\varepsilon_0/T)$ is the equilibrium number of vacancies per unit volume of the crystal. Integration in formula (3) is carried out over a constant-energy surface in the $\mathbf{k}$ space (for the sake of brevity, we denote the momentum $\mathbf{k}'$ of final states just by $\mathbf{k}$) with the energy

$$\varepsilon = \varepsilon_0 + e\mathbf{E}\cdot\mathbf{u}, \quad (4)$$

dS is the surface element and $v(\mathbf{k}) = |\partial\varepsilon/\partial\mathbf{k}|$ is the absolute value of the vacancy velocity.

The probability $W = W(E)$ as a function of the absolute value of the electric field E has Van Hove singularities (see [11]) for the values $E = E_s$ corresponding to the vanishing of the vacancy velocity $\mathbf{v}(\mathbf{k}') = \partial\varepsilon/\partial\mathbf{k}'$ in the final state. This is observed for certain values of $\mathbf{k}' = \mathbf{k}_s$ corresponding to the extrema (i.e., local minima, local maxima, or saddle points) of the energy spectrum $\varepsilon = \varepsilon(\mathbf{k}')$. For each such $\mathbf{k}_s$, the vacancy density of states $g(\varepsilon) = \int dS/v(\mathbf{k})$ has a Van Hove root singularity for $\varepsilon = \varepsilon_s = \varepsilon(\mathbf{k}_s)$, where integration is carried out over the constant-energy surface $\varepsilon = \varepsilon(\mathbf{k})$ (see [11]). Function $W(E)$ defined by formula (3) has similar root singularities for $E = E_s$, where quantity E_s is connected with ε_s through the relation

$$\varepsilon_s = \varepsilon_0 + euE_s\cos\theta. \quad (5)$$

Here, θ is the angle between the directions of u and E (it was mentioned above that $\cos\theta > 0$). By virtue of relation (3), the singular part $W_s(E)$ of function $W(E)$ for $E \approx E_s$ is connected with the singular part $g_s(\varepsilon)$ of the density of states $g(\varepsilon)$ for $\varepsilon \approx \varepsilon_s$ through the relation

$$W_s(E) = \frac{N}{(2\pi)^2}|f(\mathbf{k}_0, \mathbf{k}_s; \mathbf{u})|^2 g_s(\varepsilon), \quad (6)$$

while E and ε are connected through formula (4).

Thus, the type of singularities of functions $W(E)$ and $g(\varepsilon)$ is essentially the same. Since the drift velocity $\mathbf{V}$ of charges is equal (in the simplest case) to $\mathbf{u}W$, the field dependence $V = V(E)$ of the drift velocity must have the same type of singularities for $E = E_s$. There exist only four types of Van Hove singularities for drift velocity, which are exactly the same as the four known types of singularities in the density of states (see [11]).

1. The function $\varepsilon = \varepsilon(\mathbf{k})$ has a local minimum at the point $\mathbf{k} = \mathbf{k}_s$. In this case,

$$V(E) = V_R(E) + \begin{cases} 0 \quad \text{at} \quad E < E_s, \\ \text{const}\cdot(E - E_s)^{1/2} \quad \text{at} \quad E > E_s, \end{cases}$$

where $V_R(E)$ is a certain function regular for $E = E_s$ and const stands for a positive constant.

2. If $\varepsilon = \varepsilon(\mathbf{k})$ has a local maximum at the point $\mathbf{k} = \mathbf{k}_s$, then

$$V(E) = V_R(E) + \begin{cases} \text{const}\cdot(E_s - E)^{1/2} \quad \text{at} \quad E < E_s, \\ 0 \quad \text{at} \quad E > E_s. \end{cases}$$

3. If $\mathbf{k} = \mathbf{k}_s$ is a saddle point of the function $\varepsilon = \varepsilon(\mathbf{k})$, the following two cases are possible.

(a) The surface $\varepsilon = \varepsilon(\mathbf{k})$ for $\varepsilon \approx \varepsilon_s$ is a one-sheet hyperboloid. In this case,

$$V(E) = V_R(E) - \begin{cases} \text{const}\cdot(E_s - E)^{1/2} \quad \text{for} \quad E < E_s, \\ 0 \quad \text{at} \quad E > E_s. \end{cases}$$

(b) The surface $\varepsilon = \varepsilon(\mathbf{k})$ for $\varepsilon \approx \varepsilon_s$ is a two-sheet hyperboloid. In this case,

$$V(E) = V_R(E) - \begin{cases} 0 \quad \text{at} \quad E < E_s, \\ \text{const}\cdot(E - E_s)^{1/2} \quad \text{for} \quad E > E_s. \end{cases}$$

A special case of a first-type singularity is the root dependence of $W(E)$ in weak fields $e\mathbf{E}\cdot\mathbf{u} \ll \Delta$ determined by us earlier (see formula (19) in [2]) at low temperatures $T \ll e\mathbf{E}\cdot\mathbf{u}$. In this case, we are speaking of the absolute minimum of the function $\varepsilon = \varepsilon(\mathbf{k})$, and $E_s = 0$, $V_R(E) = 0$. It was shown in [2] that the root dependence in weak fields is confirmed by the experi-

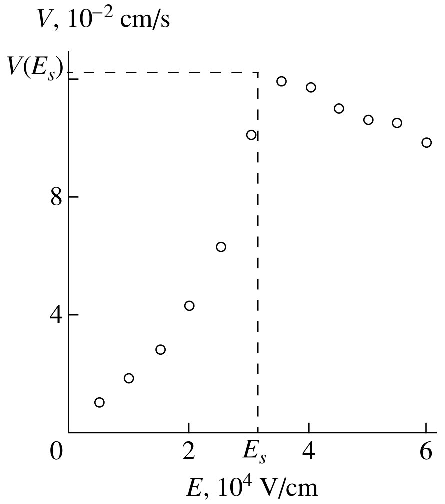


Fig. 1. Field dependence of the drift velocity of positive charges in hcp crystals of ^{4}He in the direction of the C_6 axis.

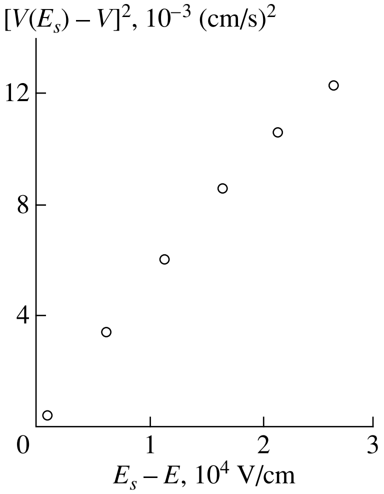


Fig. 2. Processed experimental data obtained in [4].

mental data [9] obtained in the low-pressure range for positive charges.

A singularity of the second type was also considered by us earlier for the case when we are dealing with an absolute maximum (see formula (21) in [2]). In the region $T \ll (\Delta - e\mathbf{E} \cdot \mathbf{u}) \ll \Delta$, probability W of a one-vacancion phonon-free process vanishes as we approach the energy threshold of the process in proportion to $(\Delta - e\mathbf{E} \cdot \mathbf{u})^{1/2}$ in complete accord with the general result for a second-type singularity. In this case, $V_R(E) = 0$ and $E_s = \Delta/(eu\cos\theta)$.

Figure 1 presents a typical dependence of the drift velocity $V(E)$ of positive charges in the hcp lattice of ^{4}He in the direction of the principal axis C_6 of the crystal as a function of field E, which was obtained by Andreeva *et al.* [4]. In Fig. 2, the same results for $E < E_s$ are presented in $[V(E_s) - V]^2$ vs. $E_s - E$ coordinates. It can be seen that the experimental data completely correspond to the type of Van Hove singularities in case 3a. Here, $E_s \approx 3.1 \times 10^4$ V/cm. According to the Van Hove theorem (see [11]), the vacancy spectrum necessarily contains at least one singular point of type 3b, and accordingly, one more singularity of type 3b must be observed on the $V(E)$ dependence upon a further increase in the field. In still stronger fields, the above-mentioned singularity of type 2 corresponding to the threshold of one-vacancion phonon-free processes must be observed.

The nonmonotonic field dependence of the drift velocity with a Van Hove-type singularity presented in Fig. 1 was observed [4] for the motion of positive charges in the direction of the principal axis C_6 of the crystal. The drift velocities of positive charges moving in directions perpendicular to C_6 and of negative charges moving in all directions are much smaller in absolute value, and their field dependences are of a regular monotonic type. The absence of singularities indicates the many-particle nature of the corresponding nonelastic processes. The expressions for W of type (1) for such processes acquire additional integrals, which strongly suppresses the singularity in any case.

Apart from one-vacancion processes, we considered in [2] the scattering of vacancies at a charge complex, accompanied by the emission (or absorption) of phonons. It is these processes that are responsible for a cubic dependence of the type $(E + E_0)^3$ that can be observed in strong fields. It should be noted that there exists the possibility of many-vacancion processes (which was not considered in [2], but was noted in [5]).

The displacement of the charge complex as a result of the simultaneous scattering of two vacancies from it is obviously characterized by a considerably higher differential probability w than that corresponding to a one-vacancion process. For this reason, the two-vacancion process can compete with the one-vacancion process. However, the expression of type (1) for the transition frequency W acquires the product of two distribution functions for the impinging vacancies. The temperature dependence of the drift velocity is characterized by the factor $\exp(-2\varepsilon_0/T)$, and the field dependence must have no Van Hove singularities (see above). In the experiments [4], the activation energy for the motion of positive charges in directions perpendicular to C_6 is (to a high degree of accuracy) twice as high as for motion along C_6 (11 K and 5.3 K), and the field dependences are regular. This is a strong argument in favor of the two-vacancion nature of motion in directions perpendicular to C_6.

3. ANISOTROPY OF MOBILITY AND THE STRUCTURE OF CHARGES

The presence of an anomalously high anisotropy in the mobility of positive charges in the hcp lattice of ^{4}He (the mobilities along and across the C_6 axis differ by a factor of 200(!)) indicates [4] (see also the end of [2]) that one possible vector of the displacement of a charge is directed strictly along the C_6 axis of the crystal, and the probability W of a displacement by this vector is much higher than the probability of a displacement by all other vectors. The large difference in these probabilities can be explained by different types of the corresponding charge-transfer processes (one-vacancion or two-vacancion process), and the fact that one of the displacement vectors is directed along the C_6 axis leads to a conclusion on the spatial position of the "center of gravity" of a positive charge in the hcp lattice of ^{4}He. Indeed, we assume that the "center of gravity" of a positive charge in the hcp lattice coincides with the center of the octahedron $ABCDEF$ depicted in Fig. 3. The dots indicate the positions of the atoms of the hcp lattice in the basal plane (coinciding with the plane of the figure) perpendicular to the C_6 axis. The crosses mark the projections of atoms lying in the plane displaced upwards from the plane of the figure by half the lattice period onto this plane. Such a position of an ion is crystallographically equivalent, for example, to the position of an ion at the centers of the octahedra $AOGcbD$ and $A^*B^*C^*DEF$, where sites A^*, B^*, and C^* are obtained from A, B, and C by the reflection in the plane of crosses (their position in Fig. 3 coincides with the positions of sites A, B, and C). The vector of translation from position $ABCDEF$ to position $AOGcbD$ lies in a plane perpendicular to the C_6 axis and is equal in magnitude to the atomic spacing a. The vector of translation from $ABCDEF$ to $A^*B^*C^*DEF$ is directed along the C_6 axis and has a length of $\sqrt{2/3}a$; i.e., it is noticeably shorter than the translation vector in the transverse direction. On the other hand, it can be seen easily that there are no other positions of an ion in the hcp lattice (except the octahedral positions mentioned above) which contain, among the set of vectors of translation to neighboring crystallographically equivalent positions, the vectors directed along the C_6 axis that are sufficient for displacing a charge over large distances (see example 2 from [2]). Knowing the density of helium crystals in the experiments [4], the length of the displacement vector ($\sqrt{2/3}a$) along the C_6 axis, and the value E_s of the field for which a singularity is observed on the field dependence ($E_s \approx 3.1 \times 10^4$ V/cm), we can easily calculate the position of the saddle point in the vacancy spectrum relative to the bottom of the vacancy band using formula (5). This value is found to be approximately equal to 10 K. The field dependence of the drift velocity must also contain at least one more singularity, but it is not observed on the experimental curve in fields up to 6×10^4 V/cm. This can mean that the width of the vacancy band in the hcp crystals of ^{4}He on the melting curve exceeds 18 K.

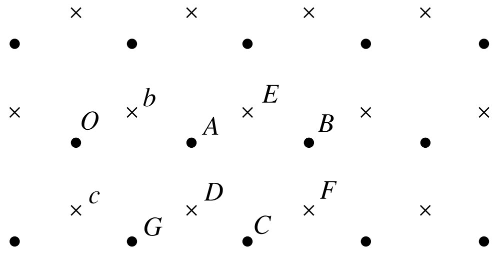


Fig. 3. Arrangement of atoms in hcp crystals.

4. MOTION OF CHARGES IN ^{3}He

The behavior of vacancies in ^{3}He crystals is determined by the configuration of the nuclear spins of the atoms (see [3]). If the spins are completely polarized, the crystal is periodic and the states of vacancies are characterized by definite values of quasimomentum as in the case of ^{4}He crystals. All the conclusions drawn above and in [2] concerning the drift of charges in ^{4}He crystals can be completely extended to this case (naturally, taking into account the difference in the symmetry of the crystal lattices of ^{4}He and ^{3}He).

It was shown by Nagaoka [12] that the energy of vacancies corresponding to the bottom of the band in the completely polarized state is the absolute minimum (for bcc crystals) of the energy of vacancies relative to its possible values for any other spin configuration. Under such conditions, at temperatures T smaller than the vacancy band width Δ in the completely polarized state, a ferromagnetic region of a completely polarized crystal must be formed around a vacancy (see [3, 10]). The order of magnitude of the radius R of such a spin polaron at temperatures much higher than the temperature T_c of antiferromagnetic ordering is determined by the relation $R \sim a(\Delta/T)^{1/5}$, where a is the lattice constant. In the antiferromagnetic state, we have $R \sim a(\Delta/T_c)^{1/5}$. In this case, a vacancy itself is localized and is on the lower energy level in the spherical potential well of radius R, which is formed by a spin polaron. Highly excited vacancy states within the polaron are virtually identical to the vacancy states in the homogeneous completely polarized state of the crystal; i.e., they are characterized by a definite quasimomentum. We can visualize the vacancy as a wave packet within a certain large volume. Such a representation is valid if the packet does not lose its shape during its flight across the polaron. The packet is deformed by $\sqrt{ta^2\Delta/\hbar}$ during time t. Substituting the time of flight R/v (v is the vacancy velocity) into this formula for t and requiring that the packet size does not exceed the polaron radius R, we obtain the condition $v \gg \Delta a^2/R\hbar$. This condition holds for the entire vacancy energy interval except for regions of a width on the order of $\Delta a^2/R^2 \sim \Delta^{3/5}T^{2/5}$ directly adjoining stationary points (including the bottom and top) of the vacancy band in a completely polarized crystal. The width of these regions for $T \ll \Delta$ is always much larger than the temperature, but much smaller than the band width Δ.

In the course of diffusive motion of vacancies together with ferromagnetic polarons surrounding them, a localized charge complex may get into one of the polarons. In this case, the charge can go over into another crystallographically equivalent localized state as a result of its interaction with a vacancy. The energy $e\mathbf{E} \cdot \mathbf{u}$ liberated in this case is transferred to the vacancy. The conclusions drawn here and in our earlier publication [2] concerning the field dependence of the drift velocity of charges in strong fields are obviously valid in the given case also. However, in the present case the constraint imposed on the field is more stringent: $e\mathbf{E} \cdot \mathbf{u} \gg \Delta^{3/5}T^{2/5}$. A unique property of charges under these conditions is that their drift in a unpolarized crystal is the same as in the completely polarized state of the crystal. In particular, the field dependence of the drift velocity must have Van Hove singularities the same as in the vacancy spectrum considered above. Such a possibility of studying the vacancy spectrum in the completely polarized state of ^{3}He crystals is of special interest in connection with the problem of the existence of zero-point vacancies in a polarized ^{3}He crystal, which was discussed earlier (see [3]).

ACKNOWLEDGMENTS

The authors are grateful to K.O. Keshishev and A.O. Andreeva for fruitful discussions of the results and for valuable remarks.

This research was partly supported by the Russian Foundation for Basic Research, the INTAS, and NWO.

REFERENCES

1. A. F. Andreev and A. É. Meĭerovich, Zh. Éksp. Teor. Fiz. **67**, 1559 (1974) [Sov. Phys. JETP **40**, 776 (1975)].
2. A. F. Andreev and A. D. Savishchev, Zh. Éksp. Teor. Fiz. **96**, 1109 (1989) [Sov. Phys. JETP **69**, 630 (1989)].
3. A. F. Andreev, in *Progress in Low Temperature Physics*, Ed. by D. F. Brewer (North-Holland, Amsterdam, 1982), Vol. VIII.
4. O. Andreeva, K. Keshishev, D. Kholin, and P. Kossyrev, J. Low Temp. Phys. **110**, 89 (1998).
5. O. A. Andreeva, K. O. Keshishev, and A. D. Savischev, Physica B (Amsterdam) **284-288**, 343 (2000).
6. V. B. Shikin, Pis'ma Zh. Éksp. Teor. Fiz. **13**, 65 (1971) [JETP Lett. **13**, 44 (1971)]; Zh. Éksp. Teor. Fiz. **61**, 2053 (1971) [Sov. Phys. JETP **34**, 1095 (1971)].
7. C. Herring, J. Appl. Phys. **21**, 437 (1950).
8. I. M. Lifshits, Zh. Éksp. Teor. Fiz. **44**, 1349 (1963) [Sov. Phys. JETP **17**, 909 (1963)].
9. K. O. Keshishev, Zh. Éksp. Teor. Fiz. **72**, 521 (1977) [Sov. Phys. JETP **45**, 273 (1977)].
10. A. F. Andreev, Pis'ma Zh. Éksp. Teor. Fiz. **24**, 608 (1976) [JETP Lett. **24**, 564 (1976)].
11. L. D. Landau and E. M. Lifshitz, *Statistical Physics* (Nauka, Moscow, 1976; Pergamon, Oxford, 1980), Part 1, Section 70.
12. Y. Nagaoka, Phys. Rev. **147**, 392 (1966).

Translated by N. Wadhwa

Journal of Experimental and Theoretical Physics, Vol. 92, No. 2, 2001, pp. 357–360.
From Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 119, No. 2, 2001, pp. 409–412.
Original English Text Copyright © 2001 by Cheremisin.

MISCELLANEOUS

Peltier-Effect-Induced Correction to Ohmic Resistance[¶]

M. V. Cheremisin[a, b, *]
[a]*Ioffe Physicotechnical Institute, Russian Academy of Sciences, St. Petersburg, 194021 Russia*
[b]*GES, University Montpellier-II, 34095, Montpellier, France*
**e-mail: maksim.vip1@pop.ioffe.rssi.ru*

Received August 23, 2000

Abstract—The standard ohmic measurements by means of two extra leads contain an additional thermal correction to resistance. The current results in heating (cooling) at the first (second) sample contact because of the Peltier effect. The contact temperatures are different. The measured voltage is the sum of the ohmic voltage swing and the Peltier-effect-induced thermoelectromotive force that is linear in the current. As a result, the thermal correction to the resistance measured exists as $I \longrightarrow 0$. The correction could be comparable with the ohmic resistance. Above some critical frequency depending on thermal inertial effects, the thermal correction disappears. *© 2001 MAIK "Nauka/Interperiodica".*

It is well known that ohmic measurements (see Fig. 1) are carried out at low current density in order to prevent sample heating. Usually, only the Joule heat is considered to be important. In contrast to the Joule heat, the Peltier and Thomson effects are linear in the current. The crucial point of the present paper is that the Peltier effect, which is linear in the current, influences the ohmic measurements and results in a correction to the resistance measured. Under the current carrying conditions, one of the sample contacts is heated and the other is cooled because of the Peltier effect. The temperature gradient established is proportional to the current. The Thomson heat is then proportional to the square of the current and can therefore be neglected. Finally, the voltage swing across the circuit includes the thermoelectromotive force induced by the Peltier effect, which is linear in the current. Accordingly, there exists a thermal correction to the ohmic resistance of the sample.

First we consider an isotropic (or possessing a cubic symmetry) conductor that can be in thermodynamic nonequilibrium with respect to conducting electrons. In general, the current density **j** and the energy flux density **q** of the inhomogeneous conductor are given by [1]

$$\mathbf{j} = \sigma(\mathbf{E} - \alpha\nabla T), \tag{1}$$

$$\mathbf{q} = (\phi + \alpha T)\mathbf{j} - \kappa\nabla T, \tag{2}$$

where σ is the conductivity, α is the thermopower, and κ is the thermal conductivity. For an inhomogeneous conductor, the potential $\phi = \varphi + \mu/e$ is the sum of the electric potential φ and the chemical potential μ of conducting electrons. For a homogeneous conductor, the above definition of the potential differs from φ by a constant, and the average microscopic electric field $-\nabla\varphi$ therefore coincides with $\mathbf{E} = -\nabla\varphi$.

The first term in Eq. (1) corresponds to the conventional Ohm's law and the second term describes thermoelectric phenomena. In the steady state,

$$\operatorname{div}\mathbf{j} = 0, \tag{3}$$

$$Q = -\operatorname{div}\mathbf{q} = \operatorname{div}(\kappa\nabla T) + j^2/\sigma - jT\nabla\alpha = 0, \tag{4}$$

where Q is the total amount of heat evolved per unit time and unit volume of the conductor. The current flow is accompanied by both the Joule and Thomson heats that are proportional to the second and first power of the current, respectively. Using Eqs. (1)–(4), one can find the potential $\phi(\mathbf{r})$ and the temperature $T(\mathbf{r})$ for the conductor under given boundary conditions.

Now we consider the thermal effects in connection with ohmic measurements of the conductor resistance (Fig. 1). The conductor is connected by means of two identical extra leads to the current source (not shown). Both contacts are assumed to be ohmic; α, σ, κ, length l, and the conductor cross section S are different for the leads and the sample. The voltage is measured between the open ends c and d that are kept at the temperature T_0 of the external thermal reservoir. In general, the contacts a and b could be at different temperatures T_a and T_b, respectively.

It is well known that Peltier heat is generated by the current crossing the contact of two different conductors. At the contact (for example, a in Fig. 1), temperature T_a, the electrochemical potential ϕ, the normal components of the current $I = jS$, and the total energy flux qS are continuous. There exists the difference of thermopowers $\Delta\alpha = \alpha_1 - \alpha_2$. For $\Delta\alpha > 0$, the charge intersecting contact a gains the energy $e\Delta\alpha T_a$. Consequently, $Q_a = I\Delta\alpha T_a$ is the amount of the Peltier heat evolved per unit time in contact a. We emphasize that

---

[¶]This article was submitted by the authors in English.

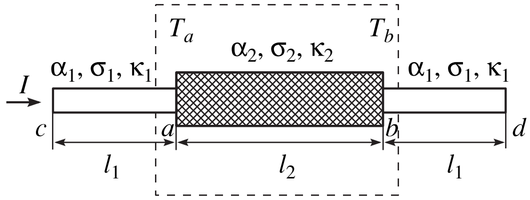


Fig. 1. The circuit for standard ohmic measurements. The dashed square represents the sample chamber.

Q_a can be calculated directly through the Thomson term in Eq. (4):

$$Q_a = \int -IT\nabla\alpha dx,$$

where the integration is taken over the contact length. In fact, the Peltier effect is equivalent to the Thomson effect established at the contact.

For $\Delta\alpha > 0$ and the current direction shown in Fig. 1, contact *a* is heated and contact *b* is cooled. Thus, the contacts are at different temperatures and $T_a - T_b = \Delta T > 0$. Now we will show that the standard ohmic measurements always result in a thermal correction to the resistance measured. Using Eq. (1), we express the voltage swing *U* between ends *c* and *d* as

$$U = \int_c^d \left(\frac{j}{\sigma} + \alpha\nabla T\right)dx = RI + \varepsilon_T, \tag{5}$$

where

$$R = 2R_1 + R_2 = \frac{2l_1}{S_1\sigma_1} + \frac{l_2}{S_2\sigma_2}$$

is the total resistance of the circuit. The first term in Eq. (5) corresponds to Ohm's law.[1] The second term,

$$\varepsilon_T = \int_c^d \alpha dT,$$

coincides with the expression for the conventional thermoelectromotive force under zero current conditions [1]. We notice that ε_T is a universal value because it only depends on the contact temperatures for arbitrary cooling conditions.

There exists a correlation between the thermoelectromotive force and the Peltier and Thomson heats. The total power evolved in the circuit, *UI*, is the sum of the Joule heat RI^2 and the power $\varepsilon_T I$ related to the thermal effects. The product $\varepsilon_T I$ is then exactly the sum of the Peltier heat

$$Q_P = Q_a - Q_b = I\Delta\alpha\Delta T$$

evolved at both contacts and the Thomson heat

$$Q_T = -\int_c^d IT\nabla\alpha dx$$

in the conductor bulk:

$$\varepsilon_T I = Q_P + Q_T. \tag{6}$$

It follows from Eq. (6) that for an arbitrary circuit under the same contact temperatures (T_a, T_b, and T_0), the zero current measurements of the thermoelectromotive force allow one to find the total amount of both the Peltier and Thomson heats at $I \neq 0$.

We recall that the sample contacts are always extra heated (or cooled) because of the Peltier effect. The difference of the contact temperatures ΔT is linear in current, and therefore, there exists a thermal correction to the ohmic resistance:

$$\Delta R = \varepsilon_T/I = U/I - R.$$

For simplicity, we assume further that conductivity σ, thermopower κ, and the thermal conductivity α are temperature independent. The thermoelectromotive force is then given by $\varepsilon_T = \Delta\alpha\Delta T$.

Using Eqs. (4) and (5), one can easily find the voltage swing *U* and, thus, the thermal correction ΔR for an arbitrary circuit. We emphasize that the real cooling conditions strongly influence ΔR. Now we specify the cooling conditions of the circuit shown in Fig. 1. We consider the adiabatic conditions with the sample being thermally isolated from the environment. For example, the sample can be placed into a vacuum chamber (see Fig. 1) surrounded by a thermal reservoir kept at T_0. We neglect the heat transfer within the leads considering a self-isolated sample. We emphasize that under the above conditions, the sample is not heated. In fact, at small current we have $T_a \approx T_b \approx T_0$, and hence, the Peltier heat evolved at contact *a* is equal to the heat absorbed at contact *b*. The energy flux *qS* is continuous at each contact, and therefore,

$$Q_a = -Q_b = I\Delta\alpha T_0 = -\kappa_2\frac{dT}{dx}S_2. \tag{7}$$

Using Eq. (7), we find the thermal correction to resistivity as

$$\Delta R = \frac{T_0(\Delta\alpha)^2 l_2}{S_2\kappa_2}. \tag{8}$$

According to Eq. (8), ΔR depends on the reservoir temperature and on the geometry and heat conductivity of the sample. We emphasize that the thermal correction is always positive, because the total amount of Peltier heat $Q_P = \Delta RI^2 > 0$.

---

[1] Ohm's law: "The amount of current flowing in a circuit made up of pure resistances is directly proportional to the electromotive forces (voltages) impressed on the circuit and inversely proportional to the total resistance of the circuit" (1827).

Now we estimate the magnitude of the thermal correction ΔR to resistivity in the case where both the conductor and leads are metals. At room temperature, the electron heat conductivity and thermopower of the electron gas are given by

$$\kappa = L\sigma T, \quad \alpha = \frac{\pi^2 k}{2e}\xi,$$

where $L = \pi^2 k^2/3e^2$ is the Lorentz number and $\xi = kT/E_F \ll 1$ is the degeneracy parameter. Difference $\Delta\alpha$ is of the order $k\xi/e$. It follows from the above assumptions that $\Delta R/R \sim \xi^2 \ll 1$. Thus, the thermal correction is small compared with the ohmic resistance because the electron gas is degenerate. For semimetals (bismuth, $E_F \approx 35$ meV), the thermal correction can be greater.

In contrast to the case of a strongly degenerate electron gas, the thermal correction must be greater for a non-degenerate semiconductor because

$$\Delta\alpha \approx \alpha_2 = \frac{k}{e}\left(\frac{5}{2} + r - \xi^{-1}\right) \sim \frac{k}{e},$$

where r is the parameter related to the mechanism of the electron scattering. For example, we consider the nondegenerate n-InSb at $T = 0.3$ K. The Fermi energy lies between the conduction band and the shallow donor impurity level $\Delta E_d \approx 7$ K. For the photon scattering ($r = 3/2$), we obtain $\Delta\alpha = 11k/e$. At low temperatures, the electron heat conductivity is smaller than the photon-related Debye one ($\kappa_{ph} = 0.05T^3$ W/(cm K)).

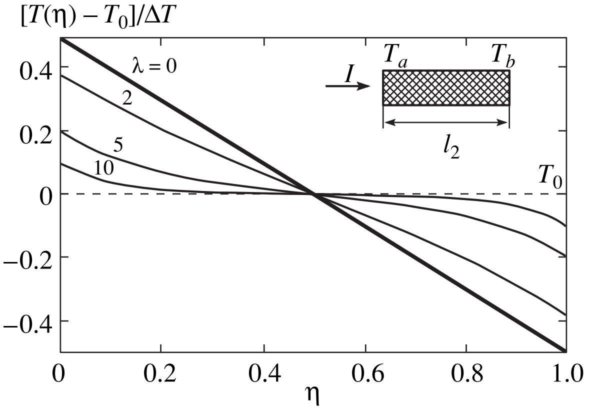


Fig. 2. The dimensionless $T(\eta)$ relation given by Eq. (9) for fixed current, $\lambda = 0, 2, 5, 10$, and the contact temperature difference ΔT found at $\lambda = 0$.

For n-InSb with the electron concentration $n = 10^{13}$ cm^{-3} and mobility $\mu = 5 \times 10^6$ cm^2/(V s), we obtain the thermal correction to resistivity $\Delta R/R \sim 0.01$.

In reality, the cooling conditions can be different from those assumed above. Now we consider a more realistic case where the local cooling of the sample is important, for example, with the sample chamber containing the gas. One can take the cooling effects into account using Eq. (4) with the linear term $-\beta(T - T_0)$ included, where β denotes the strength of the sample-to-gas thermal exchange. Under small current the spatial temperature distribution (Fig. 2) is given by

$$T(\eta) = \frac{(T_a - T_0)\sinh[\lambda(1-\eta)] + (T_b - T_0)\sinh[\lambda\eta]}{\sinh\lambda} + T_0, \tag{9}$$

where $\eta = x/l_2$ is the dimensionless coordinate. The sample local cooling is therefore governed by the dimensionless parameter $\lambda = \sqrt{\beta/\kappa_2}\, l_2$. Actually, λ is the ratio of the outgoing and internal (within the sample) heat fluxes. When $\lambda \ll 1$, the local cooling can be neglected, and hence, $T(\eta)$ is linear (Fig. 2). In the opposite intensive cooling case where $\lambda \gg 1$, the $T(\eta)$ dependence is sharp near the contacts.

The above results allow us to calculate the thermal correction to the sample resistance. Using Eqs. (5), (7), and (9) and omitting cumbersome algebraic calculations, we calculate the thermal correction to the resistance as

$$\Delta R = \frac{T_0(\Delta\alpha)^2 l_2}{S_2\kappa_2}\frac{\tanh(\lambda/2)}{\lambda/2}. \tag{10}$$

For small cooling as $\lambda \longrightarrow 0$, Eqs. (8) and (10) coincide. In the opposite strong-cooling case, where $\lambda \longrightarrow \infty$, difference ΔT and, thus, the thermal correction decreases (see Fig. 2).

Now we estimate ΔR given by Eq. (10) for natural air convection cooling. For a sample with the typical dimension $d \sim \sqrt{S}$, the outgoing thermal flux is given by $\kappa_{gas}(T - T_0)Nu/S$, where $Nu \sim 10$ is the Nusselt number. For a n-InSb sample ($0.5 \times 0.5 \times 0.5$ cm), the heat conductivity is $\kappa_{ph} = 0.15$ W/(cm K) ($T = 293$ K). Assuming the air heat conductivity $\kappa_{gas} = 2.6$ W/(cm K), we find $\lambda = 14$. Thus, the thermal correction to resistivity is approximately seven times smaller than that in the absence of convection.

We emphasize that both dc and ac ohmic measurements lead to a thermal correction. However, ΔR diminishes at high frequencies because of the thermal inertial effects. In fact, Eq. (8) is valid below some critical frequency $f_{cr} = \chi/d^2$, where χ is the temperature diffusive coefficient of the sample. For example, in a metal conductor at room temperature $\chi = \kappa/C \approx 10^2$ cm^2/s, where C is the heat of the electron gas. For a typical metal conductor with $d = 1$ mm, we obtain the critical frequency $f_{cr} = 10^4$ Hz. We suggest that the spectral dependence of the thermal correction can be used to estimate the magnitude of the thermal correction.

In conclusion, the ohmic measurements of a conductor resistance contain the thermal correction caused by the Peltier effect. The thermal correction always exists, while its magnitude depends on the actual cooling conditions of the circuit. Above some critical frequency depending on thermal inertial effects, the thermal correction disappears.

The author is grateful to M.I. Dyakonov and V.I. Perel for fruitful discussions.

REFERENCES

1. L. D. Landau and E. M. Lifshitz, *Course of Theoretical Physics*, Vol. 8: *Electrodynamics of Continuous Media* (Nauka, Moscow, 1982; Pergamon, New York, 1984).