Journal of Experimental and Theoretical Physics, Vol. 100, No. 2, 2005, pp. 229–241.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 127, No. 2, 2005, pp. 259–272.

ATOMS, MOLECULES, OPTICS

Dynamics of Moving Interacting Atoms in a Laser Radiation Field and Optical Size Resonances

O. N. Gadomskiĭ* and A. G. Glukhov
Ul'yanovsk State University, Ul'yanovsk, 432700 Russia
**e-mail: qed_group@mail.ru*
Received March 1, 2004

Abstract—The forces acting on interacting moving atoms exposed to resonant laser radiation are calculated. It is shown that the forces acting on the atoms include the radiation pressure forces as well as the external and internal bias forces. The dependences of the forces on the atomic spacing, polarization, and laser radiation frequency are given. It is found that the internal bias force associated with the interaction of atomic dipoles via the reemitted field may play an important role in the dynamics of dense atomic ensembles in a light field. It is shown that optical size resonances appear in the system of interacting atoms at frequencies differing substantially from transition frequencies in the spectrum of atoms. It is noted that optical size resonances as well as the Doppler frequency shift in the spectrum of interacting atoms play a significant role in the processes of laser-radiation-controlled motion of the atoms. © *2005 Pleiades Publishing, Inc.*

1. INTRODUCTION

The action of laser radiation on atoms, molecules, or dielectric microparticles makes it possible to control their motion owing to the action of radiation forces [1, 2]. Radiation forces that have been considered in the literature so far are associated with laser radiation proper and act over distances between particles much longer than the radiation wavelength (i.e., in the radiation wave zone). A fundamentally different situation takes place when particles are located at distances comparable to or much smaller than the radiation wavelength. In this case, an additional field associated with mutual effect of atoms in the external radiation field in the near-field zone is coupled with the external radiation field. It was shown in [3] that the additional field leads to a noticeable change in the amplitude, phase, and frequency characteristics of dipole radiation of interacting atoms. This property of the mutual effect of closely spaced atoms was interpreted in [3] as the near field effect. This effect can be manifested in optical properties of nanostructural objects (dimers, atomic chains, and fullerenes), in the Brewster reflection of light from the surface of nonabsorbing insulators [3], and in optical probing of the surfaces of solids [3].

In addition of the above-mentioned effects, the mutual influence of particles leads to the emergence of a force exerted by one particle on another. From the physical point of view, the force acting between two atomic particles emerges as a result of interaction of induced dipole moments. From the quantum-mechanical point of view, induced dipole moments of atoms appear as a result of atomic coherence induced by a laser field. In the case of dielectric particles, the force is due to macroscopic polarization of the medium.

This force may play an important role in various fields of application. The force of interaction between atoms and molecules in a laser radiation field is important in applications such as laser cooling of atoms in traps (including magnetooptical traps), in experiments on Bose–Einstein condensation, and atomic force microscopy. The force of interaction between dielectric particles in a radiation field is important in such applications as the control of motion of microparticles, including control of the spatial position of viruses and bacteria.

In this study, we present an analysis of the forces of dipole interaction of atoms exposed to stationary quasi-resonant laser radiation. The results of this study show that the forces of interaction depend to a considerable extent on the atomic spacing, polarization, and laser field frequency, as well as the Doppler frequency shift.

A large number of publications are devoted to the interaction of atoms in dense ensembles exposed to a laser field [3–9]. In contrast to these publications, we will show here that optical size resonances play an important role in controlling the motion of atoms in dense atomic ensembles exposed to laser radiation. Optical size resonances in a system of stationary atoms were considered for the first time in [10–12]. The existence of such resonances was confirmed experimentally in [13], where typical peaks were observed in the spectra of anisotropic reflection of light from a gallium arsenide surface reconstructed by antimony. In our opinion [14], such effects indicate the presence of optical size resonances in antimony dimers. Optical size resonances should apparently be manifested in various systems consisting of a small number of atoms forming nanostructural objects. It should be noted that the effect of atomic dipoles in polyatomic systems, leading to the

so-called local field effect, were considered by many authors (see, for example, [15–17]). A distinguishing feature of the effect of atomic dipoles in nanostructural objects containing a small number of atoms is that the intrinsic field in such systems can be calculated exactly without using statistical averaging of the dipole fields.

In this paper, we consider optical size resonances in systems of moving interacting atoms with allowance for the Doppler frequency shift. Moreover, in contrast to [3–9], we consider here the motion of atoms in laser fields without a noticeable change in the populations of atomic states and disregarding spontaneous atomic transitions of atoms from excited states to the ground state. It will be shown that the motion of atoms in dense ensembles can be effectively controlled in low-intensity laser fields by tuning the frequency of laser radiation to the size resonance frequency, which depends on the atomic spacing.

2. INTERACTION OF TWO ATOMS IN A RADIATION FIELD: EQUATIONS OF MOTION

We will henceforth assume that atoms are two-level systems and that one two-level atom is at the origin $\mathbf{r}_1 = 0$ of the coordinate system, while the other atom is at point $\mathbf{r}_2$ with coordinates $(0, R, 0)$. The atoms are irradiated by a running laser wave with constant amplitude $\mathbf{E}_{0L}$ and frequency ω, which is close to atomic transition frequencies ω_{01} and ω_{02}. The electric field of this wave has the form

$$\mathbf{E}_L(\mathbf{r}, t) = \mathbf{E}_{0L}\exp(i\mathbf{k}_0 \cdot \mathbf{r} - i\omega t), \quad (1)$$

where $\mathbf{k}_0$ is the wavevector of the external wave.

The electric field formed by the laser wave and by the dipole moments of the atoms at point of observation $\mathbf{r}$ at instant t can be represented in the form [3]

$$\mathbf{E}(\mathbf{r}, t) = \mathbf{E}_L(\mathbf{r}, t) + \sum_{j=1}^{2} \mathrm{curl}\,\mathrm{curl}\frac{\langle \mathbf{d}\rangle_j(t - R_j/c)}{R_j}, \quad (2)$$

where $\langle \mathbf{d}\rangle_j$ is the induced dipole moment of the jth atom, which depends on the field value at the location of this atom; $R_j = |\mathbf{r} - \mathbf{r}_j|$; and c is the velocity of light in vacuum. Observation point $\mathbf{r}$ in Eq. (2) may either coincide with points $\mathbf{r}_1$ and $\mathbf{r}_2$ or lie outside of these points. Differentiation in Eq. (2) is carried out over the coordinates of the observation point. It should be noted that the values of induced dipole moments $\langle \mathbf{d}\rangle_1$ and $\langle \mathbf{d}\rangle_2$ in the near-field zone are determined only by the electric vector of the electromagnetic field. The effect of the magnetic vector of the field becomes comparable to the effect of the electric vector only in the wave zone when $k_0 R_j \gg 1$.

For a system of two atoms, the Hamiltonian can be written in the form

$$H = \frac{1}{2}\hbar\sum_{j=1}^{2}\omega_{0j}(I + \sigma_{3j}) - \frac{1}{2}\sum_{j=1}^{2}(\mathbf{d}_{0j}^{*}\sigma_{+j}\mathbf{E}_j + \mathbf{d}_{0j}\sigma_{-j}\mathbf{E}_j). \quad (3)$$

Here, $\mathbf{E}_j$ are the electric fields at the points of locations of atoms 1 and 2, $\mathbf{d}_{0j}$ is the matrix element of the dipole moment of the jth atom, I is the unit operator, and $\sigma_{\alpha j}$ ($\alpha = 1, 2, 3$, $j = 1, 2$) and $\sigma_{\pm j} = \sigma_{1j} \pm i\sigma_{2j}$ are the effective spin operators for the jth atom, which satisfy the commutation relations

$$[\sigma_{1j}, \sigma_{2k}] = 2i\sigma_{3j}\delta_{jk}, \quad [\sigma_{2j}, \sigma_{3k}] = 2i\sigma_{1j}\delta_{jk},$$
$$[\sigma_{3j}, \sigma_{1k}] = 2i\sigma_{2j}\delta_{jk}, \quad [\sigma_{+j}, \sigma_{-k}] = 4i\sigma_{3j}\delta_{jk}, \quad (4)$$
$$[\sigma_{+j}, \sigma_{3k}] = -2i\sigma_{+j}\delta_{jk}, \quad [\sigma_{-j}, \sigma_{3k}] = 2i\sigma_{-j}\delta_{jk},$$

where δ_{jk} is the Kronecker delta. It should be noted that, in the notation used here, the dipole moment operators of the atoms have the form

$$\mathbf{d}_j = \frac{1}{2}(\mathbf{d}_{0j}^{*}\sigma_{+j} + \mathbf{d}_{0j}\sigma_{-j}).$$

Using Hamiltonian (3) and commutation relations (4), we can now derive the Heisenberg equations of motion for two parts of the dipole moment operators $\mathbf{d}_{+j} = \mathbf{d}_{0j}^{*}\sigma_{+j}$ and $\mathbf{d}_{-j} = \mathbf{d}_{0j}\sigma_{-j}$ and σ_{3j}:

$$\dot{\mathbf{d}}_{+j} = i\mathbf{d}_{+j}\omega_{0j} + \frac{2i}{\hbar}\mathbf{d}_{0j}^{*}\sigma_{3j}(\mathbf{d}_{0j}\cdot\mathbf{E}_j),$$
$$\dot{\mathbf{d}}_{-j} = -i\mathbf{d}_{-j}\omega_{0j} - \frac{2i}{\hbar}\mathbf{d}_{0j}\sigma_{3j}(\mathbf{d}_{0j}^{*}\cdot\mathbf{E}_j), \quad (5)$$
$$\dot{\sigma}_{3j} = \frac{i}{\hbar}(\mathbf{d}_{+j} - \mathbf{d}_{-j})\cdot\mathbf{E}_j.$$

After this, a transition can be made in Eqs. (5) to mean values $s_{\alpha j} = \langle\sigma\rangle_{\alpha j}$ and $\mathbf{s}_{\pm j} = \langle\mathbf{d}\rangle_{\pm j}$; in the case of a classical field, the form of Eqs. (5) remains unchanged.

It is convenient for subsequent analysis to separate explicitly the field oscillations with optical frequency, assuming that

$$\mathbf{E}_j = \mathbf{E}_{0j}\exp(-i\omega t), \quad \mathbf{E}_{0j} = \mathbf{E}'_{0j} - i\mathbf{E}''_{0j},$$

where $\mathbf{E}'_{0j}$ and $\mathbf{E}''_{0j}$ are the real and imaginary parts of the fields acting at the points of location of atoms 1

and 2. In stationary fields, these quantities do not depend on time. We introduce the transformation

$$\begin{aligned} s_{1j} &= u_j\cos(\omega t) - v_j\sin(\omega t),\\ s_{2j} &= u_j\sin(\omega t) + v_j\cos(\omega t), \quad s_{3j} = w_j,\\ s_{\pm j} &= (u_j \pm iv_j)\exp(\pm i\omega t),\\ \mathbf{s}_{\pm j} &= \mathbf{d}_{0j}(u_j \pm iv_j)\exp(\pm i\omega t), \end{aligned} \tag{6}$$

as well as the functions

$$\mathbf{X}_j = \mathbf{d}_{0j}(u_j - iv_j), \quad \mathbf{X}_j^* = \mathbf{d}_{0j}^*(u_j + iv_j),$$

defining the induced dipole moments of the atoms. These functions satisfy the equations for coupled quantum dipoles, which should be supplemented with relaxation terms taking into account the natural widths of atomic transitions. Assuming that the total rate of spontaneous decay of the upper levels is equal to $2\gamma_j$, we can write the equations of motion for atomic variables in final form:

$$\dot{\mathbf{X}}_j = i\delta_j\mathbf{X}_j - \frac{2i}{\hbar}|\mathbf{d}_{0j}|^2 w_j\mathbf{E}_{0j} - \gamma_j\mathbf{X}_j, \tag{7a}$$

$$\dot{\mathbf{X}}_j^* = -i\delta_j\mathbf{X}_j^* + \frac{2i}{\hbar}|\mathbf{d}_{0j}|^2 w_j\mathbf{E}_{0j}^* - \gamma_j\mathbf{X}_j^*, \tag{7b}$$

$$\dot{w}_j = \frac{i}{\hbar}(\mathbf{X}_j^*\cdot\mathbf{E}_{0j} - \mathbf{X}_j\cdot\mathbf{E}_{0j}^*) - 2\gamma_j(w_j - w_{0j}), \tag{7c}$$

where $\delta_j = \omega - \omega_{0j}$ are detunings of the field frequency relative to optical atomic transition frequencies in the jth atom and w_{0j} is the equilibrium value w_j. In deriving Eqs. (7), we took into account the fact that relation $\mathbf{d}_{0j}(\mathbf{d}_{0j}\cdot\mathbf{E}_0) = |\mathbf{d}_{0j}|^2\mathbf{E}_0$ holds for $\mathbf{d}_{0j} \parallel \mathbf{E}_{0j}$.

Noting further that the induced dipole moments of the atoms at instants $t - R_j/c$ are defined by the relations

$$\langle\mathbf{d}\rangle_j = \frac{1}{2}\mathbf{X}_j\exp(ik_0R_j - i\omega t),$$

and using expressions (1), (2), and (6), we can represent quantities $\mathbf{E}_{0j}$ in the form

$$\begin{aligned} \mathbf{E}_{01} &= \mathbf{E}_{0L}\exp(i\mathbf{k}_0\cdot\mathbf{r}_1) + \frac{1}{2}\hat{G}\mathbf{X}_2\exp(ik_0R),\\ \mathbf{E}_{02} &= \mathbf{E}_{0L}\exp(i\mathbf{k}_0\cdot\mathbf{r}_2) + \frac{1}{2}\hat{G}\mathbf{X}_1\exp(ik_0R), \end{aligned} \tag{8}$$

where tensor $\hat{G}$ is defined as

$$\hat{G} = \begin{pmatrix} -A & 0 & 0\\ 0 & 2B & 0\\ 0 & 0 & -A \end{pmatrix}, \quad B = \frac{1}{R^3} - i\frac{k_0}{R^2},$$

$$A = B - \frac{k_0^2}{R}, \quad k_0 = \frac{\omega}{c}.$$

Equations (7) and (8) form a coupled system of equations, which makes it possible to calculate the fields at the points of location of the atoms as well as induced dipole moments of the atoms taking into account the mutual influence of the atoms on each other. In this case, the velocities $\mathbf{v}_1$ and $\mathbf{v}_2$ of the atoms depend in the general case on acting fields (8), which leads to dynamic variation of transition frequencies in the spectrum of the interacting moving atoms. In subsequent analysis, we will use the coordinate system associated with the center of mass of the two atoms, introducing the atomic spacing vector $\mathbf{r} = \mathbf{r}_1 - \mathbf{r}_2$. Placing the origin at the center of mass, we obtain

$$\begin{aligned} \mathbf{r}_1 &= \frac{m_2}{m_1+m_2}\mathbf{r}, \quad \mathbf{r}_2 = -\frac{m_1}{m_1+m_2}\mathbf{r},\\ \mathbf{v}_1 &= -\mathbf{v}_2, \quad \mathbf{v}_1 = \frac{m_2}{m_1+m_2}\mathbf{v}, \end{aligned} \tag{9}$$

where m_1 and m_2 are the masses of atoms 1 and 2, respectively, and $\mathbf{v} = \dot{\mathbf{r}}$ is the velocity of relative motion of the atoms.

3. LINEAR OPTICAL SIZE RESONANCES IN THE SYSTEM OF TWO INTERACTING ATOMS

Let us consider the most interesting case from the standpoint of physics, when the atoms interact with radiation over time periods much longer than the spontaneous relaxation time $\tau_{sp} = 1/2\gamma$. For such periods of time, in Eqs. (7) we must take into account the fact that

$$\dot{\mathbf{X}}_j = 0, \quad \dot{w}_j = 0, \quad \dot{\mathbf{v}} = 0. \tag{10}$$

This means that relaxation processes and the processes of inducing the dipole moments in the atoms are mutually compensated. Condition $\dot{\mathbf{v}} = 0$ indicates that the motion of the center of mass of the atoms is uniform. The fulfillment of condition $\dot{\mathbf{v}} = 0$ in the system of atoms in the state of thermal equilibrium at a certain temperature T is connected with the number of collisions of an atom with the remaining atoms of an ideal gas per unit time,

$$\nu = \frac{N}{V}16r_0^2\sqrt{\frac{\pi k_B T}{m_A}},$$

where r_0 is the radius of the atom, k_B is the Boltzmann constant, m_A is the atomic mass, and N/V is the concentration of atoms. It is necessary that the time ν^{-1} between two consecutive collisions be longer than time τ_{sp}. In this case, if conditions (10) are satisfied, Eqs. (7) can be reduced to a system of nonlinear algebraic equations.

We introduce the quantum polarizabilities of the atoms in the region of isolated resonances ω_{0j}:

$$\alpha_j = -\frac{2|\mathbf{d}_{0j}|^2}{\hbar}\frac{1}{\delta_j + i\gamma_j}. \tag{11}$$

Then the steady-state solution to Eqs. (7a) and (7b) can be represented in the form

$$
\begin{aligned}
X_1^y &= -E_{0L}^y w_1\alpha_1 \frac{\exp(i\mathbf{k}_0\cdot\mathbf{r}_1) - w_2\alpha_2 B\exp(i\mathbf{k}_0\cdot\mathbf{r}_2)\exp(ik_0R)}{1 - w_1w_2B^2\alpha_1\alpha_2\exp(2ik_0R)} = \alpha_{\text{eff}}^y(1)E_{0L}^y,\\
X_1^{x,z} &= -E_{0L}^{x,z} w_1\alpha_1 \frac{\exp(i\mathbf{k}_0\cdot\mathbf{r}_1) + \frac{1}{2}w_2\alpha_2 A\exp(i\mathbf{k}_0\cdot\mathbf{r}_2)\exp(ik_0R)}{1 - \frac{1}{4}w_1w_2A^2\alpha_1\alpha_2\exp(2ik_0R)} = \alpha_{\text{eff}}^{x,z}(1)E_{0L}^{x,z},\\
X_2^y &= -E_{0L}^y w_2\alpha_2 \frac{\exp(i\mathbf{k}_0\cdot\mathbf{r}_2) - w_1\alpha_1 B\exp(i\mathbf{k}_0\cdot\mathbf{r}_1)\exp(ik_0R)}{1 - w_1w_2B^2\alpha_1\alpha_2\exp(2ik_0R)} = \alpha_{\text{eff}}^y(2)E_{0L}^y,\\
X_2^{x,z} &= -E_{0L}^{x,z} w_2\alpha_2 \frac{\exp(i\mathbf{k}_0\cdot\mathbf{r}_2) + \frac{1}{2}w_1\alpha_1 A\exp(i\mathbf{k}_0\cdot\mathbf{r}_1)\exp(ik_0R)}{1 - \frac{1}{4}w_1w_2A^2\alpha_1\alpha_2\exp(2ik_0R)} = \alpha_{\text{eff}}^{x,z}(2)E_{0L}^{x,z},
\end{aligned}
\tag{12}
$$

where $\hat{\alpha}_{\text{eff}}(j)$ is the nonlinear effective polarizability of the jth atom.

Induced dipole moments (12) of the atoms depend on quantities w_1 and w_2, i.e., on the differences in the probabilities of detecting the atoms in the ground and excited states. In the general case, quantities w_1 and w_2 are functions of fields $\mathbf{E}_{01}$ and $\mathbf{E}_{02}$; consequently, the effective polarizabilities are nonlinear functions of external field $\mathbf{E}_{0L}$. It will be shown below, however, that in a certain range of $\mathbf{E}_{0L}$ values, the effective polarizabilities are quantities that are practically independent of laser field $\mathbf{E}_{0L}$. This case corresponds to a linear approximation. Indeed, Eqs. (7a) and (7b) lead to the expressions

$$X_j^\gamma = \alpha_{\text{eff}}^\gamma(j)E_{0L}^\gamma = -w_j\alpha_j E_{0j}^\gamma,$$

where $\gamma = x, y, z$. Substituting these expressions into Eqs. (7c), we arrive at the equality

$$w_j = \frac{w_{0j}}{(2|\mathbf{d}_{0j}|^2/\hbar^2)|\mathbf{E}_{0j}|^2[1/(\delta_j^2 + \gamma_j^2) + 1]}.$$

If the condition

$$\frac{2|\mathbf{d}_{0j}|^2}{\hbar^2}|\mathbf{E}_{0j}|^2\frac{1}{\delta_i^2 + \gamma_j^2} \ll 1$$

is satisfied, we can assume that quantities w_j in our analysis coincide with their equilibrium value $w_{0j} = -1$. Taking into account the relation between laser field $\mathbf{E}_{0L}$ and acting fields $\mathbf{E}_{0j}$, we obtain the linear approximation condition in the form

$$\frac{|\alpha_{\text{eff}}^\gamma(j)|}{\sqrt{2}|\mathbf{d}_{0j}|}E_{0L}^\gamma \ll 1. \tag{13}$$

In the following analysis, we will use only this approximation, which allows us to treat the motion of the atoms in the laser field as a coherent process, disregarding stochastic processes associated with spontaneous decay of the excited states of the atoms, leading to random displacements of the atoms due to the recoil during the emission of spontaneous photons [1, 2].

Analysis of formulas (12) shows that the effective polarizabilities possess size resonances whose frequencies are functions of the atomic spacing. In the linear approximation, these size resonances are referred to as steady-state optical size resonances. These resonances can be described theoretically using Eqs. (7a) and (7b) and assuming that $w_j \longrightarrow -1$.

For two identical atoms (when $\alpha_1 = \alpha_2$), we can derive from Eqs. (12) the simple formulas for frequencies w_1' and w_2' of size resonance,

$$w_1' = \omega_0 - \frac{2d_0^2}{\hbar}B, \tag{14a}$$

$$w_2' = \omega_0 + \frac{d_0^2}{\hbar}A, \tag{14b}$$

where it is assumed that $w_1 = w_2 = -1$, $d_{01} = d_{02} = d_0$, $\omega_{01} = \omega_{02} = \omega_0$, and $k_0R \ll 1$. The condition $\alpha_1 = \alpha_2$ can be satisfied for stationary atoms. For $k_0R \ll 1$, we can disregard the time lag in the dipole–dipole interaction of the atoms and set $A = B = 1/R^3$. The equality of the natural frequencies of the atoms, $\omega_{01} = \omega_{02}$, can be observed if the interacting atoms are stationary. Thus, the system of two identical atoms ($\alpha_1 = \alpha_2$) possesses two linear stationary optical size resonances depending on the direction of polarization of the laser wave relative to the **R** axis of the diatomic system. When the polarization vector of the laser wave is directed along the **R** axis of the system, the size resonance frequency is equal to ω_1' and is defined by formula (14a). If, however, the polarization vector of the laser wave is perpendicular to the **R** axis of the diatomic system of identical atoms, the size resonance frequency is defined by formula (14b) and is equal to ω_2'. At frequencies ω_1' and ω_2', the induced dipole moments and, hence, effective

polarizabilities (12) attain their maximal values. The widths of these resonances are determined by the natural width of energy levels in isolated atoms constituting the system. These atoms are isotropic in our analysis. The mutual influence of the atoms in the laser radiation field leads to anisotropy of the diatomic system since $\alpha_{\text{eff}}^{y}(j) \neq \alpha_{\text{eff}}^{x,z}(j)$.

As a result of their motion, atoms with the same values of ω_0 and $|\mathbf{d}_0|$ cease to be identical. Indeed, the linear Doppler effect causes a phonon frequency shift upon the transition from the stationary coordinate system to the system of coordinates of the atom having velocity $\mathbf{v}_j$. Consequently, $\alpha_1 \neq \alpha_2$ and

$$\begin{aligned}\alpha_1 &= -\frac{2|\mathbf{d}_0|^2}{\hbar}\frac{1}{\omega - \mathbf{k}_0\cdot\mathbf{v}_1 - \omega_0 + i\gamma},\\ \alpha_2 &= -\frac{2|\mathbf{d}_0|^2}{\hbar}\frac{1}{\omega + \mathbf{k}_0\cdot\mathbf{v}_1 - \omega_0 + i\gamma}.\end{aligned} \tag{15}$$

Substituting these expressions into Eqs. (12), we can find the frequencies of size resonances, for which the effective polarizability of the atoms acquires its maximal values. At small distances such that $k_0R \ll 1$, we assume that polarizabilities (15) are real quantities and derive the following formulas for optical size resonance frequencies at such atomic spacings:

$$\begin{aligned}\omega_{1,2} &= \omega_0 \pm \left[(\mathbf{k}_0\cdot\mathbf{v}_1)^2 + w_{10}w_{20}B^2\left(\frac{2|\mathbf{d}_0|^2}{\hbar}\right)^2\right]^{1/2},\\ \omega_{3,4} &= \omega_0 \pm \left[(\mathbf{k}_0\cdot\mathbf{v}_1)^2 + \frac{1}{4}w_{10}w_{20}A^2\left(\frac{2|\mathbf{d}_0|^2}{\hbar}\right)^2\right]^{1/2}.\end{aligned} \tag{16}$$

Thus, in contrast to the case of stationary atoms, allowance for the motion of identical atoms gives rise to four linear optical steady-state size resonances. The limiting transition from formula (16) to formulas (14) for $(\mathbf{k}_0\cdot\mathbf{v}_1) \longrightarrow 0$ can be carried out if we analyze dispersion properties of the numerators in effective polarizabilities (12). This enables us to choose an appropriate sign in expressions (16) and leads to two size resonances whose frequencies are defined by formula (14).

Let us consider a numerical example of two sodium atoms, singling out in a sodium atom the quantum transition $3S \longrightarrow 3P$ with wavelength $\lambda = 5890$ Å (yellow line of the Na atom). The change in the velocity of a Na atom as a result of absorption (emission) of a photon with the given wavelength is $v_r = h/m_A\lambda \approx 3$ cm/s, where m_A is the mass of the Na atom. The recoil energy is 25 kHz. The mean thermal velocity of a Na atom at room temperature $T = 300$ K is $\bar{v} = 5 \times 10^4$ cm/s. The Doppler frequency shift of a photon for such a velocity is $\Delta\nu = k_0\bar{v}/2\pi = 850$ MHz. The natural width of the $3S \longrightarrow 3P$ transition line is equal to 10 MHz, which allows us to determine the dipole moment of the transition, $d_0 = 6.239 \times 10^{-18}$ CGSE units. It follows from formulas (16) that the Doppler shift and the frequency shift due to the dipole–dipole interaction of the atoms are comparable for atomic spacings $R \approx 5$ nm. For Na atoms in an ideal gas in thermal equilibrium at $T = 300$ K, the number of collisions of a sodium atom with the remaining atoms is $\nu = (N/V) \times 2.5 \times 10^{-11}$ s^{-1}, where N/V is the concentration, and the radius of the sodium atom is assumed to be equal to the Bohr radius. For $N/V = 10^{18}$ cm^{-3}, the average time ν^{-1} between collisions is approximately equal to 10^{-6} s; consequently, the steady-state condition for process (10) is satisfied.

Let us consider the results of numerical analysis of formulas (12) for two moving sodium atoms in a laser radiation field. In these formulas, inversions w_1 and w_2 correspond to the equilibrium value; for this reason, the effective polarizabilities of moving interacting atoms are independent of the laser field $\mathbf{E}_{0L}$. In accordance with formulas (16), dispersion relations $\text{Im}\,\alpha_{\text{eff}}^{\gamma}(j)$ indicate the presence of four linear steady-state size resonances in the diatomic system of moving atoms; two of these resonances correspond to the y polarization of the laser field, while the remaining two resonances correspond to the x and z laser field polarizations. Quantum polarizabilities (15) of stationary isolated atoms in the case of exact resonance are identical, $\alpha_1 = \alpha_2 = 3i/k_0^3$. At size resonance frequencies, the effective polarizabilities of moving interacting atoms approximately coincide with quantum polarizabilities of stationary atoms over small distances as compared to the wavelength.

Figure 1 illustrates the behavior of one of the size resonances for the y component of the effective polarizability depending on the atomic spacing and a fixed relative velocity of atoms. It can be seen from the figure that, for large atomic spacings, the optical resonance corresponds to a dimensional detuning $v\omega_0/c\gamma$ from resonance, which is equal to 1.697×10^2. With decreasing atomic spacing, the optical size resonance gradually expands, the effective polarizability gradually increasing thereby to a value comparable to the quantum polarizability of an isolated atom in the case of exact resonance. For a fixed velocity of relative motion of the atoms, there exists a limiting value of atomic spacing $R/\lambda \approx 0.05$ for which an optical size resonance is formed. The frequency dependence of the effective polarizability of the first sodium atom for positive detunings from resonance in accordance with formula (16) can be analyzed analogously. For such detunings from resonance, in a wide range of variation of R/λ, the value of $\text{Im}\,\alpha_{\text{eff}}^{\gamma}(1)$ remains virtually unchanged; the frequency of this size resonance slightly shifts to the right only in the vicinity of $R/\lambda \approx 0.05$. The dependence of $\text{Im}\,\alpha_{\text{eff}}^{\gamma}(2)$ on R/λ and δ/γ for the second sodium atom in the given numerical example coincides with the dependence presented in Fig. 1, the only difference being that the sign of detun-

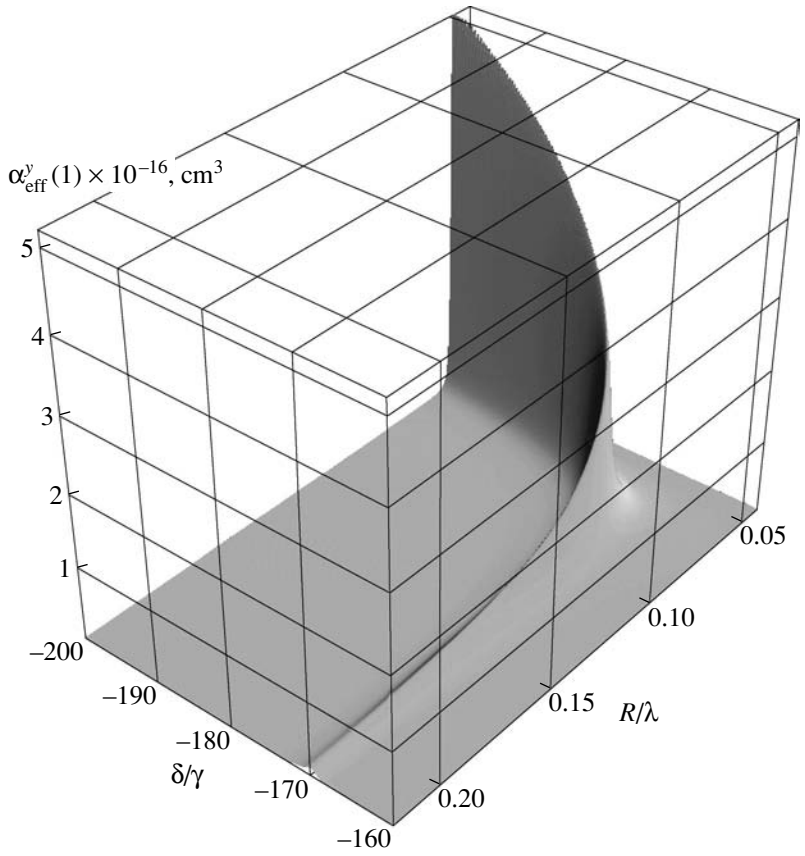

Fig. 1. Effective polarizability of moving sodium atom 1 interacting with moving sodium atom 2 for various atomic spacings; $\delta = \omega - \omega_0$, $\mathbf{k}_0 \parallel \mathbf{v}$, where $\mathbf{v}$ is the relative velocity of the atoms.

ing δ/γ is reversed. In this case, the value of $R/\lambda \approx 0.05$ corresponds to atomic spacings for which the dipole–dipole interaction parameter $2d_0^2/\gamma\hbar R^3$ becomes comparable to parameter $v\omega_0/c\gamma$ determining the Doppler frequency shift in the spectrum of moving atoms.

4. DIPOLE FORCES IN THE SYSTEM OF INTERACTING ATOMS IN A LASER RADIATION FIELD

The forces acting on the atoms of a diatomic object in a laser field are defined as [1, 2]

$$\mathbf{F}_1 = -\left\langle \frac{\partial V_1}{\partial \mathbf{r}_1} \right\rangle, \quad \mathbf{F}_2 = -\left\langle \frac{\partial V_2}{\partial \mathbf{r}_2} \right\rangle, \tag{17}$$

where the angle brackets imply quantum-mechanical averaging and V_1 and V_2 are the operators of interaction between the atoms and the electric field. In accordance with formula (3), for two-level atoms of the object, we can write

$$V_j = -\frac{1}{2}\sigma_{+j}\mathbf{d}_{0j}^* \cdot \mathbf{E}_{0j} - \frac{1}{2}\sigma_{-j}\mathbf{d}_{0j} \cdot \mathbf{E}_{0j}^*, \tag{18}$$

where acting fields $\mathbf{E}_{0j}$ are defined by formulas (8), in which quantities $\mathbf{X}_j$ can be calculated using effective polarizabilities (12) that are functions of coordinates $\mathbf{r}_1$ and $\mathbf{r}_2$ of the points of observation.

Substituting Eq. (18) into expressions (17), we obtain

$$\begin{aligned} \mathbf{F}_1 &= \frac{1}{2}\sum_{m=x,y,z} X_{1m}^* \frac{\partial E_{01}^m}{\partial \mathbf{r}_1} + \text{c.c.}, \\ \mathbf{F}_2 &= \frac{1}{2}\sum_{m=x,y,z} X_{2m}^* \frac{\partial E_{02}^m}{\partial \mathbf{r}_2} + \text{c.c.}, \end{aligned} \tag{19}$$

where

$$X_{jm} = \alpha_{\text{eff}}^m(j) E_{0L}^m. \tag{20}$$

Evaluating relations (19), we can present forces $\mathbf{F}_1$ and $\mathbf{F}_2$ as the sum of three partial forces,

$$\mathbf{F}_j = \mathbf{F}_j^{(1)} + \mathbf{F}_j^{(2)} + \mathbf{F}_j^{(3)}, \tag{21}$$

where $\mathbf{F}_j^{(1)}$ is directed along wavevector $\mathbf{k}_0$ of the laser wave, while forces $\mathbf{F}_j^{(2)}$ and $\mathbf{F}_j^{(3)}$ are directed parallel or

antiparallel to unit vector $\mathbf{n} = (\mathbf{r}_1 - \mathbf{r}_2)/|\mathbf{r}_1 - \mathbf{r}_2|$. We will refer to forces $\mathbf{F}_j^{(2)}$ and $\mathbf{F}_j^{(3)}$ as the forces of external and internal displacements of the atoms in the diatomic object, respectively, depending on the sign of vector $\mathbf{n}$ in the corresponding terms in general expressions (19).

In the absence of the dipole–dipole interaction, the atoms do not affect each other; in formulas (19) and (21), only radiation pressure forces $\mathbf{F}_1^{(1)}$ and $\mathbf{F}_2^{(1)}$, which are determined by quantum polarizabilities (15), differ from zero. If the dipole–dipole interaction is taken into account, forces $\mathbf{F}_1^{(1)}$ and $\mathbf{F}_2^{(1)}$ can be referred to as radiation pressure forces as before, although we must now use the effective polarizabilities defined by formulas (12) in the expressions for the induced dipole moments.

The directions of the radiation pressure forces coincide with wavevector $\mathbf{k}_0$ of external radiation. The dispersion relation for radiation pressure forces is determined by the dispersion relations for the effective polarizabilities of the atoms (see Fig. 1). This means that radiation pressure forces substantially increase for fixed atomic spacings R if the external field frequency coincides with the size resonance frequency. The radiation pressure forces are proportional to the square of the electric field of the external wave. Using the values of effective polarizabilities (see Fig. 1) for sodium atoms, we can determine dimensionless radiation pressure force $F_1^{(1)}/2\hbar k_0\gamma$ at the frequency of one of the size resonances. It can be seen that this force can be on the order of unity for laser fields $E_{0L} = 10^{-3}$ CGSE units, which satisfy the linear approximation condition (13).

Let us consider the behavior of two interacting atoms moving in a laser field for various directions of propagation of the laser wave relative to the axis of the diatomic object (Fig. 2). Such a situation can apparently take place in laser beams. Figures 3 and 4 show the dependences of the y components F_{1y} and F_{2y} of the resultant forces acting on atoms 1 and 2 on dimensionless atomic spacing R/λ and dimensionless detuning δ/γ from resonance. It can be seen from Figs. 3 and 4 that the dimensionless forces acting on the first and second atoms exceed the radiation pressure force approximately by a factor of two, which indicates a significant role of other partial forces (namely, the forces of internal and external displacement of the atoms) in the dynamics of the atoms. Figures 3 and 4 also show that the forces acting on the first and second atoms differ in absolute value since the moving atoms are not identical and their effective polarizabilities exhibit, in accordance with relations (12), different dependences on atomic spacings in view of the difference in quantum polarizabilities (15). We can estimate the values of accelerations of the atoms in the laser field. For example, for $F_{1y}/2\hbar k\gamma = 1$, where $k = \omega_0/c$, we find that the value of force F_{1y} is equal to 1.33×10^{-14} dyne and the accelerations is equal to 0.32×10^9 cm/s^2. For a fixed

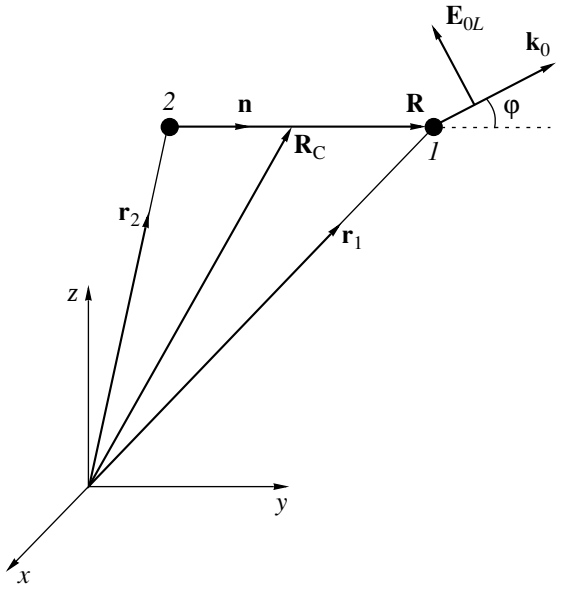

Fig. 2. Vector diagram of a diatomic system: $\mathbf{R} = \mathbf{r} = \mathbf{r}_1 - \mathbf{r}_2$; φ is the angle between vectors $\mathbf{k}_0$ and $\mathbf{R}$ lying in the zy plane; $\mathbf{R}_C$ is the radius vector of the center of mass of the system of two identical atoms of mass $m_1 = m_2 = m_A$; the diatomic object axis $\mathbf{R} \parallel \mathbf{y}_0$, where $\mathbf{y}_0$ is the unit vector directed along the y axis; and $\mathbf{n}$ is the unit vector directed from atom 2 to atom 1.

velocity of relative motion of the atoms, resultant forces F_{1y} and F_{2y} acting on the atoms in the laser field strongly change upon a variation of the atomic spacing. The dependence of these forces on the laser field frequency also changes in this case. This means that the motion of the atoms can be effectively controlled only for the relevant variation of the laser field frequency.

5. POTENTIAL ENERGY OF THE INTERACTION OF ATOMS IN A LASER RADIATION FIELD

The potential energy of an atom in a diatomic object exposed to a laser radiation field is defined by the formulas

$$\langle V_1\rangle = -\frac{1}{2}\mathbf{X}_1^* \times\left(\mathbf{E}_{0L}\exp(i\mathbf{k}_0\cdot\mathbf{r}_1) + \frac{1}{2}\hat{G}\mathbf{X}_2\exp(ik_0R)\right) + \text{c.c.},$$
$$\langle V_2\rangle = -\frac{1}{2}\mathbf{X}_2^* \times\left(\mathbf{E}_{0L}\exp(i\mathbf{k}_0\cdot\mathbf{r}_2) + \frac{1}{2}\hat{G}\mathbf{X}_1\exp(ik_0R)\right) + \text{c.c.} \quad (22)$$

Let us consider the dependence of the potential energy of the atoms on the laser radiation frequency and atomic spacing.

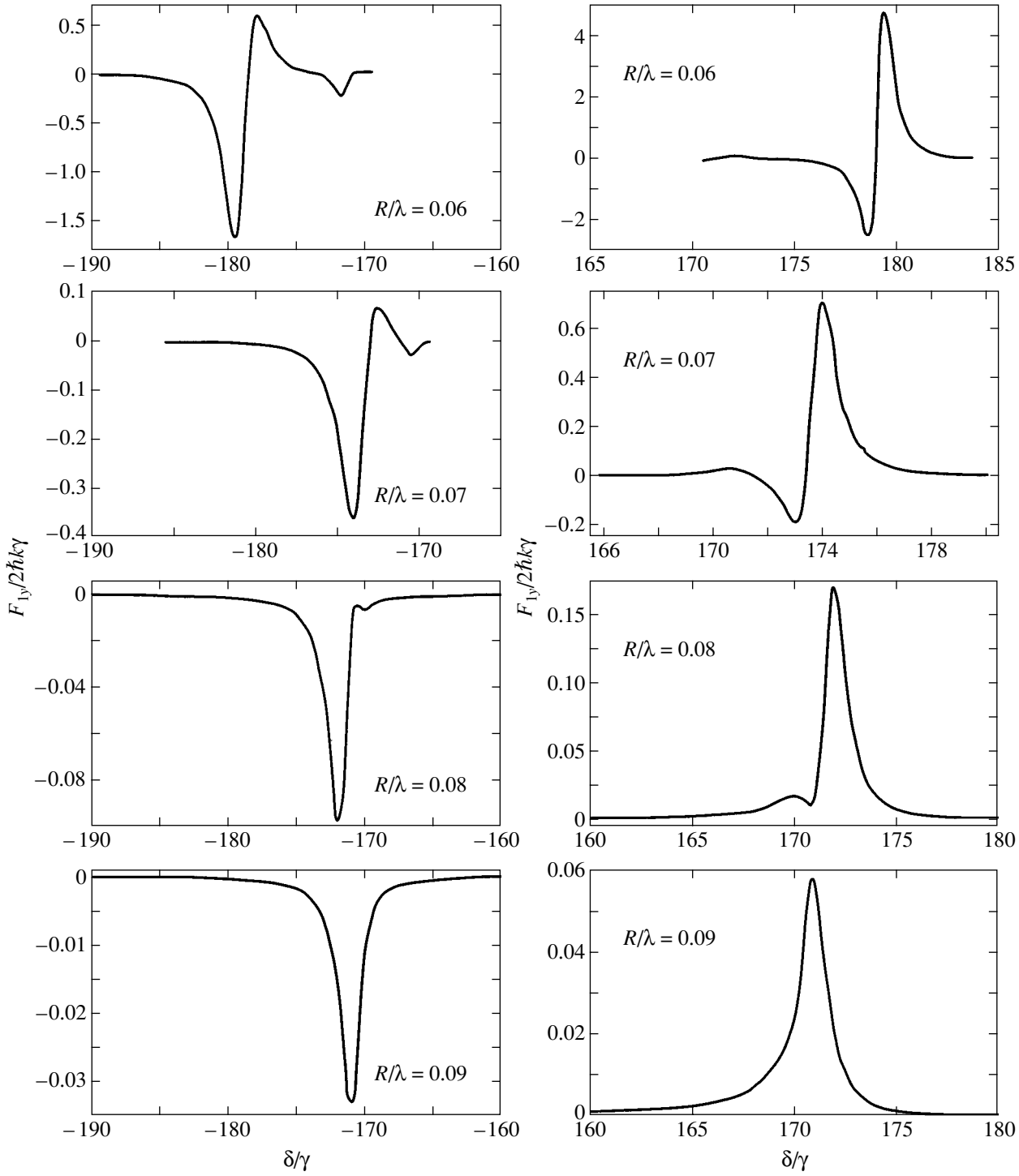

Fig. 3. Dependence of the y component F_{1y} of the resultant force acting on the first Na atom in laser field $E_{0L} = 5 \times 10^{-3}$ CGSE units on the atomic spacings and detunings from the resonance frequency of the $3S \longrightarrow 3P$ transition in the sodium atom. The velocity v of the relative motion of the atoms is fixed and equal to 5×10^4 cm/s; vector $\mathbf{k}_0$ coincides with the axis of the Na–Na object.

Figure 5 shows the dispersion curves of the potential energy of an atom in a diatomic Na–Na object exposed to laser radiation for fixed distances between the atoms. It can be seen that, for large distances between the atoms comparable to the laser radiation wavelength, the potential energy of an atom in the object may assume either positive or negative values depending on the detuning from resonance. In other words, by varying

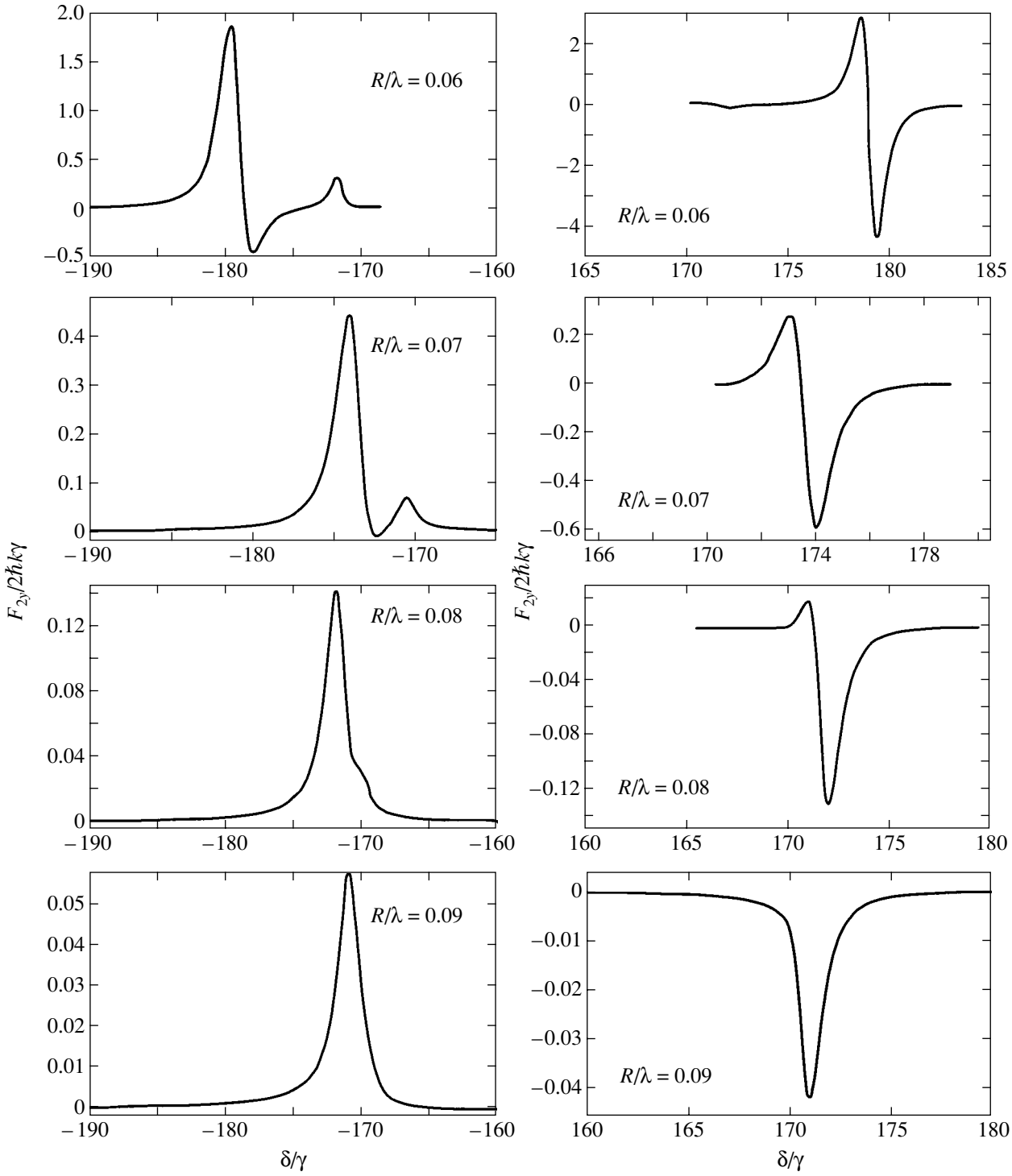

Fig. 4. Dependence of the y component F_{2y} of the resultant force acting on the second Na atom in a laser field on dimensionless distances R/λ and on dimensionless detunings δ/γ from resonance. The numerical values of physical quantities are the same as in Fig. 3.

the frequency of external radiation, we can attract or repel the atoms by laser radiation. In this case, the laser radiation intensity should not be high and the differences in probabilities w for the two atoms must almost coincide with their equilibrium values.

Figure 6 shows the typical dependences of the potential energy of an atom in the diatomic object on the value of R/λ in the field of laser radiation with fixed frequencies. It can be seen that the system of Na atoms can pass to a state with the minimal potential energy of

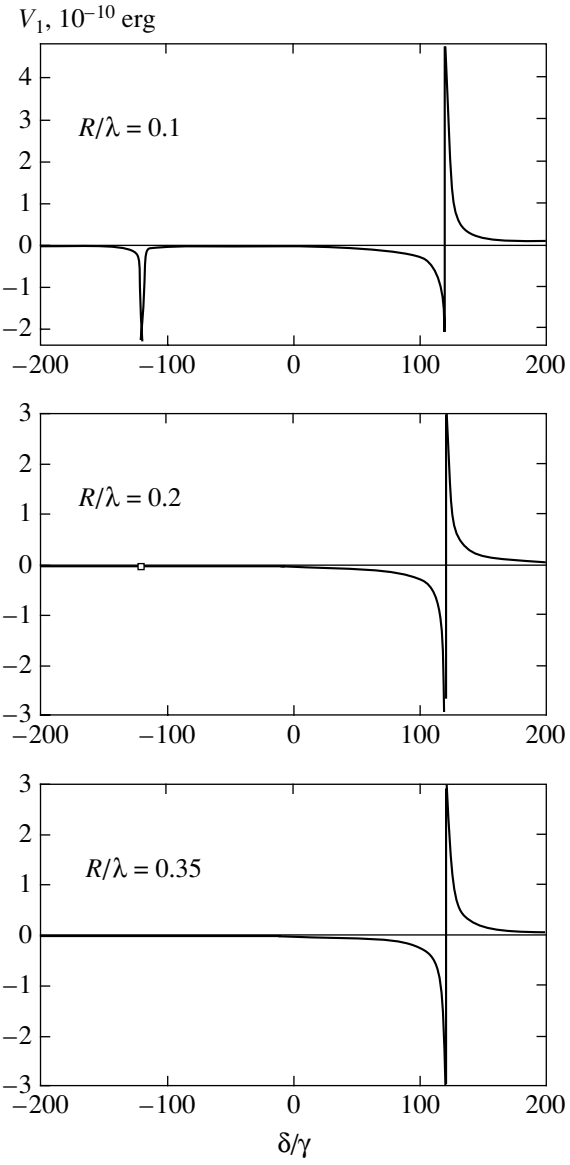

Fig. 5. Potential energy of a sodium atom in the Na–Na object in a laser radiation field for large atomic spacings. The numerical values of physical quantities are the same as in Fig. 3.

about $\langle V_1 \rangle = -5 \times 10^{-12}$ erg for small atomic spacings ($R \approx 0.045\lambda$).

Let us consider the behavior of potential energy (22) of the atoms for small atomic spacings. Substituting expression (8) into (22), we obtain the following formula for the polarization of laser radiation directed along the *y* axis:

$$\langle V_1 \rangle = -\frac{1}{2}\alpha_{\text{eff}}^{y*}(1)(\mathbf{E}_{0L}^{y})^2\left[\exp(i\mathbf{k}_0 \cdot \mathbf{r}_1) + \frac{1}{2}\alpha_{\text{eff}}^{y}(1) \times \exp(ik_0R)\left(\frac{1}{R^3} - i\frac{k_0}{R^2}\right)\right] + \text{c.c.} \quad (23)$$

It can be seen from Fig. 1 that the absolute value of effective polarizability $\alpha_{\text{eff}}^{y}(1)$ in size resonances for various atomic spacings is on the order of 10^{-15} cm^3. For weak fields, the internal field is much stronger than the external field for small atomic spacings (on the order of 1 nm). For this reason, the value of $\langle V_1 \rangle$ for such atomic spacings may reach 10^{-12} erg.

5.1. Energy and Momentum Conservation Laws in a System of Two Moving Interacting Atoms in a Laser Radiation Field

The interaction of two hydrogen-like atoms with stationary centers at arbitrary distances was treated in [3] as a third-order effect in quantum electrodynamics, in which three photons (two of which are virtual and the third is real) interact. In such an analysis, the emission or absorption of a real photon in the electric dipole approximation is represented by using the electron polarizability of the atom producing the polarizing field at the point of location of the second atom. In the nonrelativistic theory, the problem of interaction of two identical atoms in the ground state in the absence of a radiation field is solved in the second order of perturbation theory; the corresponding potential energy of interaction between two atoms without retardation is defined as $V_{12}^{(0)} = A_0/R^6$, where A_0 is a constant quantity [18]. The interaction of two identical atoms, one of which is in an excited state at a distance $R \ll 2\pi c/\omega_0$, where ω_0 is the transition frequency, is considered in the first order of perturbation theory; the corresponding potential energy is $V_{12}^{(\pm)} = \pm B_0/R^3$, where B_0 is a constant and the plus and minus signs correspond to symmetric and antisymmetric wave functions of the states of the two atoms.

Taking into account these results, we can write the energy conservation law for two moving atoms with the emission of a real photon in the form

$$\frac{1}{2}m_1\mathbf{v}_1^2 + \frac{1}{2}m_2\mathbf{v}_2^2 + V_{12}^{(\pm)} + \hbar\omega_0 = \frac{1}{2}m_1(\mathbf{v}_1')^2 + \frac{1}{2}m_2(\mathbf{v}_2')^2 + V_{12}^{(0)} + \hbar\omega, \quad (24)$$

where $\mathbf{v}_1$ and $\mathbf{v}_2$ are the initial velocities of the atoms and $\mathbf{v}_1'$ and $\mathbf{v}_2'$ are their final velocities after the emission of a real photon and a transition of the two atoms to the ground state. An analogous equality can also be written for the reverse process involving the absorption of a real photon.

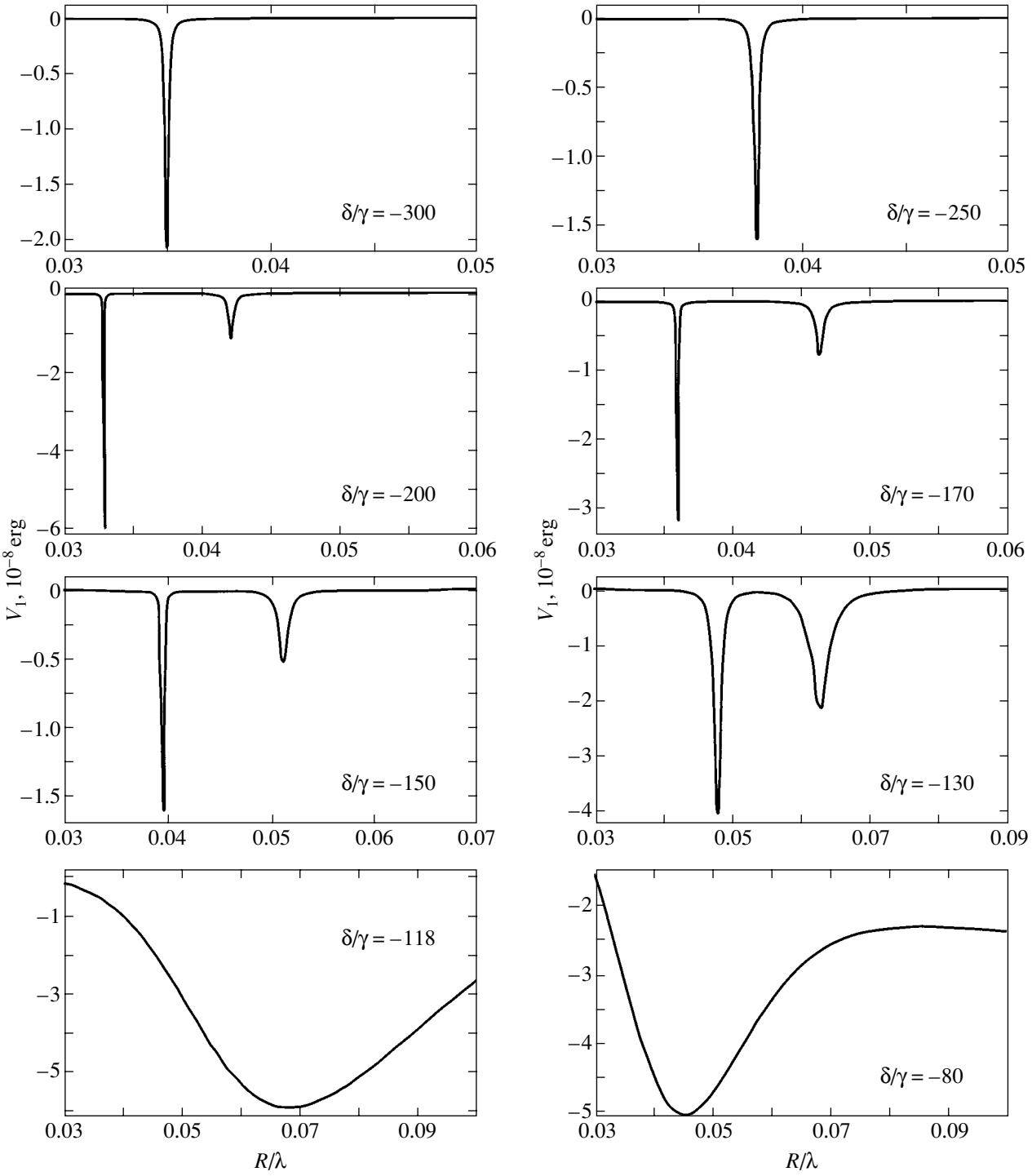

Fig. 6. Potential energy of an atom in the Na–Na object for small atomic spacings. The numerical values of physical quantities are the same as in Fig. 3.

To derive the momentum conservation law, we represent the velocities of the atoms in the form

$$\mathbf{v}_j = \mathbf{V}_j + \mathbf{v}_{j\mathrm{C}}, \tag{25}$$

where $\mathbf{V}_j$ are the velocities of translatory motion of the pair of atoms, associated with radiation pressure forces $\mathbf{F}_j^{(1)}$ and external displacement forces $\mathbf{F}_j^{(2)}$, and $\mathbf{v}_{j\mathrm{C}}$ are the velocities of the atoms, which are associated with internal atomic displacement forces $\mathbf{F}_j^{(3)}$ in the laser radiation field. For two identical atoms with

mass $m_1 = m_2 = m_A$, the radius vector of the center of mass $\mathbf{R}_C = (\mathbf{r}_1 + \mathbf{r}_2)/2$, where $\mathbf{r}_1$ and $\mathbf{r}_2$ are the radius vectors of atoms 1 and 2 relative to the laboratory system of coordinates. Introducing vector $\mathbf{r} = \mathbf{r}_1 - \mathbf{r}_2$ of the atomic spacing and placing the origin at the center of mass moving with velocity $\mathbf{V}_C$, we arrive at relation (9), where $\mathbf{v}$ is the relative velocity of the atoms in the diatomic object. Using the obvious equalities $\mathbf{v} = \mathbf{v}_1 - \mathbf{v}_2$ and $\mathbf{v}_1 + \mathbf{v}_2 = 0$, we derive the following relations:

$$\begin{aligned} \mathbf{v} &= \mathbf{v}_{1C} - \mathbf{v}_{2C}, \quad \mathbf{V}_1 = \mathbf{V}_2, \\ 2\mathbf{V}_1 &+ \mathbf{v}_{1C} + \mathbf{v}_{2C} = 0. \end{aligned} \tag{26}$$

Consequently, in the center-of-mass system, in the absence of real photons, we obtain

$$\mathbf{P}_1 + \mathbf{p}_{1C} + \mathbf{p}_{2C} = 0, \tag{27}$$

where $\mathbf{P}_1$ is the total nonrelativistic momentum of translatory motion of the pair of atoms and $\mathbf{p}_{1C} = m_A\mathbf{v}_{1C}$ and $\mathbf{p}_{2C} = m_A\mathbf{v}_{2C}$ are the momenta of the internal motion of the atoms. Thus, the momentum conservation law in the diatomic object in the case of absorption (emission) of a real photon with momentum $\hbar\mathbf{k}_0'$ has the form

$$\begin{aligned} 2m_A\mathbf{V}_1 &+ m_A\mathbf{v}_{1C} + m_A\mathbf{v}_{2C} \pm \hbar\mathbf{k}_0' \\ &= 2m_A\mathbf{V}_1' + m_A\mathbf{v}_{1C}' + m_A\mathbf{v}_{2C}', \end{aligned} \tag{28}$$

where $\mathbf{V}_1'$, $\mathbf{v}_{1C}'$, and $\mathbf{v}_{2C}'$ are the velocities of translatory and internal motion of the atoms after absorption of a real photon and the wavevector $\mathbf{k}_0'$ (on account of the linear Doppler effect) is given by

$$\mathbf{k}_0' = \frac{\mathbf{k}_0}{|\mathbf{k}_0|}\frac{\omega'}{c}, \quad \omega' = \omega - \mathbf{k}_0 \cdot \mathbf{V}_C. \tag{29}$$

Using the energy and momentum conservation laws (24) and (28), we can calculate the variations of the corresponding quantities in single events of emission or absorption of photons.

Thus, we have demonstrated here that the motion of atoms can be effectively controlled by laser radiation with allowance for the dipole–dipole interaction of atoms and the Doppler shift in the frequency of the moving atoms. The dipole–dipole interaction of atoms is taken into account without resorting to perturbation theory by using effective polarizabilities (12). We have proved that moving atoms of the same species are not identical due to the Doppler frequency shift affecting the quantum polarizabilities of two-level atoms. It has also been shown that the forces acting on the atoms in a laser field are quadratic functions of the electric field of the laser wave. This means that the motion of the atoms is analyzed in a linear approximation, in which the effective polarizabilities of the atoms are independent of the field. Such a situation is defined by condition (13), under which the inversions of the atoms differ from their equilibrium values only slightly. It can be seen from condition (13) that this approximation holds for low-intensity laser fields. However, owing to optical size resonances in the system of interacting moving atoms, it is possible to attain large values of forces acting on the atoms by smoothly varying the laser radiation frequency if the laser field frequency coincides with the frequency of one of size resonances. It has been shown that the application of the linear approximation in which the interacting atoms are mainly in the ground state makes it possible to disregard spontaneous processes of emission of photons, which lead to stochastization of the motion of the atoms due to photon recoil. Thus, the process of controlling the motion of atoms in dense ensembles considered here is coherent. The laser field intensity for which such a coherent process can be accomplished is defined by condition (13). In the numerical example considered above, this condition can be satisfied in fields $E_{0L} = 10^{-2}$–10^{-3} CGSE units, which corresponds to intensities $I_L = 1.19$–1.19×10^{-2} mW/cm^2.

In our opinion, the approach proposed here for a theoretical description of the dynamics of moving interacting atoms will make it possible to analyze various processes in atomic ensembles both in dense atomic beams as in gases with various atomic spacings. It would be interesting to analyze the trajectory of motion of atoms in a laser field taking into account the variation of the velocities of the atoms upon a decrease or an increase in the distance between them.

ACKNOWLEDGMENTS

The authors are grateful to V.G. Minogin for fruitful discussion of certain results of this study.

REFERENCES

1. V. G. Minogin and V. S. Letokhov, *Laser Light Pressure on Atoms* (Nauka, Moscow, 1986; Gordon and Breach, New York, 1987).
2. V. I. Balykin, V. G. Minogin, and V. S. Letokhov, Rep. Prog. Phys. **63**, 1429 (2000).
3. O. N. Gadomskiĭ, Usp. Fiz. Nauk **170**, 1145 (2000) [Phys. Usp. **43**, 1071 (2000)].
4. H. T. Dung, L. Knöl, and D.-G. Welsch, Phys. Rev. A **66**, 063 810 (2002).
5. V. I. Yukalov, Eur. Phys. J. D **13**, 83 (2001).
6. R. H. Lemberg, Phys. Rev. A **3**, 889 (1970).
7. M. J. Stephen, J. Chem. Phys. **40**, 569 (1964).

8. O. N. Gadomskiĭ, Teor. Mat. Fiz. **106**, 145 (1996).
9. O. N. Gadomsky and K. V. Krutitsky, J. Eur. Opt. Soc. **9**, 343 (1997).
10. O. N. Gadomskiĭ and Yu. Yu. Voronov, Pis'ma Zh. Éksp. Teor. Fiz. **69**, 750 (1999) [JETP Lett. **69**, 804 (1999)].
11. O. N. Gadomskiĭ and A. S. Kunitsyn, Zh. Prikl. Spektrosk. **67**, 777 (2000).
12. V. V. Klimov and V. S. Letokhov, Phys. Rev. B **62**, 1639 (2000).
13. V. L. Berkovits, A. B. Gordeeva, and V. A. Kosobukin, Fiz. Tverd. Tela (St. Petersburg) **43**, 985 (2001) [Phys. Solid State **43**, 1018 (2001)].
14. O. N. Gadomskiĭ and K. Yu. Moiseev, Opt. Spektrosk. **92**, 613 (2002) [Opt. Spectrosc. **92**, 560 (2002)].
15. M. G. Benedict, V. A. Malyshev, E. D. Trifonov, and A. I. Zaitsev, Phys. Rev. A **43**, 3845 (1991).
16. M. E. Crenshaw and C. M. Bowden, Phys. Rev. A **53**, 1139 (1996).
17. O. N. Gadomskiĭ and S. V. Sukhov, Kvantovaya Élektron. (Moscow) **25**, 529 (1998).
18. A. S. Davydov, *Quantum Mechanics* (Fizmatgiz, Moscow, 1963; Pergamon, Oxford, 1976).

Translated by N. Wadhwa

Journal of Experimental and Theoretical Physics, Vol. 100, No. 2, 2005, pp. 242–255.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 127, No. 2, 2005, pp. 273–288.

ATOMS, MOLECULES, OPTICS

Coherent Momentum Transfer due to Interaction between Three-Level Atoms and Counterpropagating Laser Pulses

V. I. Romanenko* and L. P. Yatsenko

Institute of Physics, National Academy of Sciences of Ukraine, Kiev, 03028 Ukraine
**e-mail: vr@iop.kiev.ua*

Received March 22, 2004

Abstract—A theoretical analysis is presented of the change in the momentum of a three-level atom due to its interaction with counterpropagating laser pulses that overlap in time. The two lower energy states of the atom are metastable; i.e., a lambda-type configuration of atomic levels is considered. The cases of two and four counterpropagating pulses having different carrier frequencies are considered. In the case of adiabatic atom–field interaction, it is shown that the atom's momentum can change by an integer multiple of the photon momentum and the corresponding standard deviation is small as compared to the photon momentum squared. © *2005 Pleiades Publishing, Inc.*

1. INTRODUCTION

Radiation pressure on atoms leads both to a change in the average momentum of an atom proportional to the atom–field interaction time and to momentum diffusion characterized by a standard deviation from the average momentum proportional to the square root of the interaction time (e.g., see [1–4] and references therein). As a result, a finite-width distribution of the atomic velocities parallel to the light propagation direction develops even if these velocity components are equal for all atoms at the initial moment of the interaction. Moreover, since momentum diffusion involves spontaneous radiative decay of an excited state, the atomic ensemble created as a result of the interaction is an incoherent superposition of atomic states having different momenta. This lack of coherence in the atomic ensemble impedes the development of atomic optical elements that must ensure the possibility of interference of atom waves. It is clear that, in order to preserve coherence, the atom–field interaction must be organized so that the resulting ensemble of atoms is unaffected by spontaneous emission.

One possibility is provided by stimulated Raman adiabatic passage (STIRAP), which makes use of coherent population trapping [5–7]. In the simplest case of a three-level atom interacting with two temporally overlapping pulses that couple the metastable states to the excited one, STIRAP makes it possible to transfer a population from one metastable state to the other almost completely without significantly populating the excited state [8]. When the atom interacts with counterpropagating pulses, the change in its momentum due to absorption of a photon from one wave and stimulated emission of a photon of the other wave is the sum of the photons' momenta. This method can be used to transfer a population between magnetic sublevels of an atom [9–11] or molecule [12] and simultaneously change its momentum without decoherence.

An alternative method for substantially changing the momentum of an atom without loss of coherence makes use of the interaction between a two-level atom and two counterpropagating laser pulses having different carrier frequencies [13]. It was shown in [13] that, if the variation of the field amplitudes is sufficiently slow, then the average momentum gained by the atom can be an integer multiple of photon momentum and the corresponding standard deviation is small as compared to the photon momentum squared, while the electronic state of the atom either remains unchanged or changes to the excited state. This momentum transfer is physically possible when the atom remains in an adiabatic state (eigenstate of the time-dependent Hamiltonian) throughout the atom–field interaction. The adiabatic state is described by the eigenfunctions corresponding to different momentum values at the initial and final moments of the interaction. In the experimental scheme suggested in [13] as an implementation of momentum transfer to an atom, the pulse widths must be smaller than the excited-state lifetime, because the excited state of the interacting atom is substantially populated. The scheme cannot be used in experiments on atomic beams using continuous-wave lasers, where a short-lived effect of light on the atom is produced as its trajectory crosses a laser beam, because the typical atom—field interaction time (about 10^{-5} s) is much longer than the typical excited-state lifetime (10^{-7} to 10^{-8} s). Note that the effect is mitigated as the pulse frequency is detuned from resonance in order to reduce spontaneous radiative decay of the excited state.

1063-7761/05/10002-0242$26.00 © 2005 Pleiades Publishing, Inc.

In this paper, we analyze the transfer of momentum to a three-level atom or molecule (with a lambda-type configuration of atomic levels) due to its interaction with counterpropagating laser pulses. Under certain conditions, the existence of the third level makes it possible to substantially reduce the population of the excited state and, therefore, spontaneous radiative loss of population. The phenomenon under analysis can be observed in experiments on the same transitions in the atoms and molecules for which STIRAP is observed, such as Ne*, NO, and SO_2 [8]. Depending on the energy difference between the metastable states and on the laser polarization, each laser pulse can couple either one metastable state or both to the excited one.

1.1. The Case of Large Energy Difference between the Metastable States

Consider an atom interacting with two pairs of laser pulses propagating in opposite directions along the *z* axis. The field of each pulse is assumed to couple only one of the metastable states to the excited one. This can be due either to a large energy difference between the metastable levels or, in some cases, to a specific laser polarization. The four pulses are grouped into two pairs in accordance with their carrier frequencies. The difference in carrier frequency between the pulses in each pair is comparable to the frequency of the transition between the metastable states. Therefore, the two-photon process of absorption at one carrier frequency and emission at the other can occur in a pair. However, two-photon transitions involving absorption and stimulated emission in different pairs are impossible because of a large detuning from two-photon resonance for these processes. When the carrier frequencies are detuned sufficiently far from the one-photon transition, the excited state is weakly populated even if the atom–field interaction time is longer than the excited-state lifetime. In the resulting effective two-level system, momentum transfer can be implemented with a small standard deviation as compared to the photon momentum squared, while the electronic state of the atom can either remain unchanged or change to the other metastable state.

1.2. The Case of Small Energy Difference between the Metastable States

In this case, each pulse couples both metastable states to the excited one. Population transfer involving a change in the atom's momentum can be induced by two counterpropagating pulses. (In the case of copropagating pulses considered in [14], momentum cannot be transferred from the field to the atom.) As in the case of a two-state atom interacting with counterpropagating pulses, the atom–field interaction time must much shorter than the excited-state lifetime to ensure that coherence is preserved.

2. THE CASE OF LARGE ENERGY DIFFERENCE BETWEEN THE METASTABLE STATES

2.1. Model

Consider a three-level system with metastable states $|1\rangle$ and $|3\rangle$ and an excited state $|2\rangle$ whose energy is higher than those of the metastable states. (Hereinafter, the system called *atom* can be either an atom or a molecule.) The state $|2\rangle$ can spontaneously decay into states other than $|1\rangle$ and $|3\rangle$ at a rate γ; i.e., the mean excited-state lifetime is $1/\gamma$. Spontaneous transitions to states $|1\rangle$ and $|3\rangle$ are neglected. Assume that only the state $|1\rangle$ is populated prior to the atom–field interaction. The atom–field interaction is described in a semiclassical approximation: the field is treated as classical, while the atom is modeled as a quantum-mechanical system. The atom interacts with four laser pulses with carrier frequencies ω_1, ω_2, ω_3, and ω_4 propagating along the *z* axis (see Fig. 1). The energy difference between $|1\rangle$ and $|3\rangle$,

$$\hbar S = W_3 - W_1,$$

is assumed to be so large that each pulse couples the

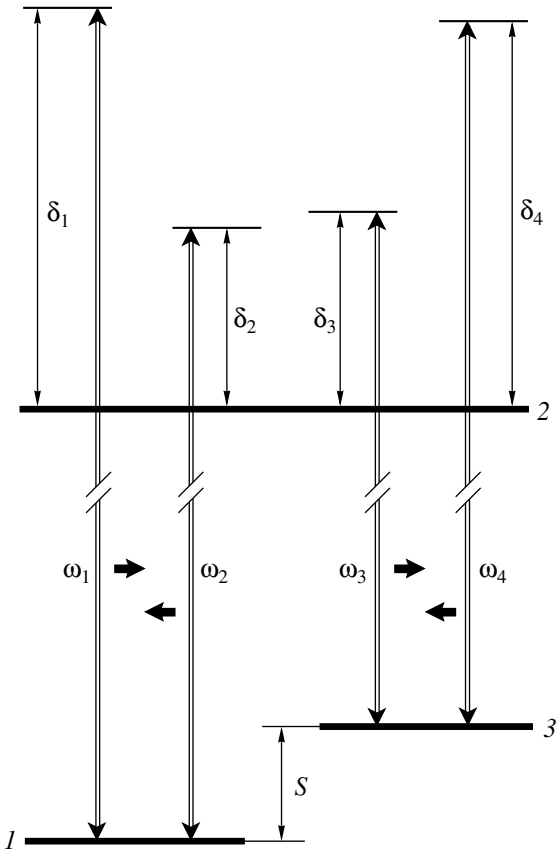

Fig. 1. Schematic diagram of interaction between three-level atom and four laser pulses. Short horizontal arrows indicate pulse propagation directions; ω_j and δ_j ($j = 1, \ldots, 4$) are carrier frequencies and detunings from corresponding transition frequencies; $\hbar S = W_3 - W_1$ is energy difference between metastable states.

excited state to one metastable state only. Pulses 1 and 2 couple the state $|1\rangle$ to the state $|2\rangle$. Pulses 3 and 4 propagate in the respective opposite directions and couple $|3\rangle$ to $|2\rangle$. The electric field interacting with the atom is

$$\begin{aligned}\mathbf{E}(t) &= \mathbf{E}_1(t)\cos(\omega_1 t - k_1 z + \varphi_1)\\ &+ \mathbf{E}_2(t)\cos(\omega_2 t + k_2 z + \varphi_2)\\ &+ \mathbf{E}_3(t)\cos(\omega_3 t - k_3 z + \varphi_3)\\ &+ \mathbf{E}_4(t)\cos(\omega_4 t + k_4 z + \varphi_4),\end{aligned} \tag{1}$$

where k_i are the wavevector magnitudes and φ_i are the initial phases ($i = 1, \ldots, 4$).

In the Raman–Nath approximation (when kinetic energy is neglected), the Hamiltonian of the atom is expressed as

$$\begin{aligned}H(t) &= H_0 - \mathbf{d}\cdot\mathbf{E}_0(t) - i\hbar\frac{\gamma}{2}|2\rangle\langle 2|,\\ H_0 &= W_1|1\rangle\langle 1| + W_2|2\rangle\langle 2| + W_3|3\rangle\langle 3|,\\ \mathbf{d} &= d_{12}|1\rangle\langle 2| + d_{21}|2\rangle\langle 1| + d_{32}|3\rangle\langle 2| + d_{23}|2\rangle\langle 3|.\end{aligned} \tag{2}$$

Here, H_0 is the Hamiltonian of the free atom, $|n\rangle$ and W_n ($n = 1, \ldots, 3$) denote its eigenstates and the corresponding energies, and $\mathbf{d}$ is the dipole moment operator. The decay term is introduced into (2) to describe the evolution of the atom by the Schrödinger equation without using an equivalent, but more complicated, equation for the density matrix. This description in dipole approximation is made possible by neglecting spontaneous radiative decay of the state $|2\rangle$ into $|1\rangle$ and $|3\rangle$, which can be done if the probability of transition to other states is much higher (e.g., for molecules). Changes in kinetic energy are negligible if the mass of the atom is sufficiently large; i.e., both Doppler effect and recoil can be ignored. Our previous study has shown that allowance for kinetic energy does not qualitatively change the behavior of a two-level atom interacting with two counterpropagating pulses [13] (at least, if the transferred momentum is about ten times the photon momentum, as in the case considered here). However, an analysis is much simpler to perform in the Raman–Nath approximation, because the atom's coordinate can be treated as a parameter in the Schrödinger equation.

In the rotating-wave approximation, the Hamiltonian is written as

$$\begin{aligned}H_{12}^{\mathrm{RWA}} &= \frac{\hbar}{2}[\Omega_1\exp(i(\delta_1 t - k_1 z + \varphi_1))\\ &+ \Omega_2\exp(i(\delta_2 t + k_2 z + \varphi_2))],\\ H_{21}^{\mathrm{RWA}} &= \frac{\hbar}{2}[\Omega_1^*\exp(-i(\delta_1 t - k_1 z + \varphi_1))\\ &+ \Omega_2^*\exp(-i(\delta_2 t + k_2 z + \varphi_2))],\\ H_{22}^{\mathrm{RWA}} &= -i\frac{\hbar}{2}\gamma,\\ H_{23}^{\mathrm{RWA}} &= \frac{\hbar}{2}[\Omega_3^*\exp(-i(\delta_3 t - k_3 z + \varphi_3))\\ &+ \Omega_4^*\exp(-i(\delta_4 t + k_4 z + \varphi_4))],\\ H_{32}^{\mathrm{RWA}} &= \frac{\hbar}{2}[\Omega_3\exp(i(\delta_3 t - k_3 z + \varphi_3))\\ &+ \Omega_4\exp(i(\delta_4 t + k_4 z + \varphi_4))].\end{aligned} \tag{3}$$

The remaining matrix elements are zero. In (3), we introduce carrier-frequency detunings from resonances,

$$\begin{aligned}\delta_1 &= \omega_1 - \frac{1}{\hbar}(W_2 - W_1),\\ \delta_2 &= \omega_2 - \frac{1}{\hbar}(W_2 - W_1),\\ \delta_3 &= \omega_3 - \frac{1}{\hbar}(W_2 - W_3),\\ \delta_4 &= \omega_4 - \frac{1}{\hbar}(W_2 - W_3),\end{aligned} \tag{4}$$

and the Rabi frequencies

$$\Omega_1 = -\langle 1|\mathbf{d}\cdot\mathbf{E}_1(t)|2\rangle/\hbar, \quad \Omega_2 = -\langle 1|\mathbf{d}\cdot\mathbf{E}_2(t)|2\rangle/\hbar,$$

$$\Omega_3 = -\langle 3|\mathbf{d}\cdot\mathbf{E}_3(t)|2\rangle/\hbar, \quad \Omega_4 = -\langle 3|\mathbf{d}\cdot\mathbf{E}_4(t)|2\rangle/\hbar,$$

characterizing the strengths of respective atom–field interactions.

2.2. Adiabatic Elimination of the Excited-State Population

Suppose that the detuning is much larger than the corresponding Rabi frequency for each pulse, but the difference between δ_1 and δ_4, as well as that between δ_2 and δ_3, is comparable to the Rabi frequencies. In addition, assume that

$$\begin{aligned}|\Delta_{14} + i\gamma| &\gg \max(|\Omega_1(t)|, |\Omega_2(t)|, |\Omega_3(t)|, |\Omega_4(t)|),\\ |\Delta_{23} + i\gamma| &\gg \max(|\Omega_1(t)|, |\Omega_2(t)|, |\Omega_3(t)|, |\Omega_4(t)|),\\ |\Delta_{14} - \Delta_{23}| &\gg \max(|\Omega_1(t)|, |\Omega_2(t)|, |\Omega_3(t)|, |\Omega_4(t)|),\end{aligned} \tag{5}$$

where

$$\Delta_{14} = \frac{1}{2}(\delta_1 + \delta_4),$$
$$\Delta_{23} = \frac{1}{2}(\delta_2 + \delta_3), \tag{6}$$

and the atom–field interaction time τ_{int} is such that

$$\tau_{int}\max(\Omega_n(t)) \gg 1$$

for each n. Adiabatically eliminating the excited-state population amplitude, we find the matrix elements of the Hamiltonian of the effective two-level system in the basis of the states $|1\rangle$ and $|3\rangle$:

$$H_{1,1}^{TL} = \frac{\hbar}{2}\left(\frac{|\Omega_1(t)|^2}{2\Delta_{14} + i\gamma} + \frac{|\Omega_2(t)|^2}{2\Delta_{23} + i\gamma}\right),$$

$$H_{1,3}^{TL} = \frac{\hbar}{2}\left(\frac{\Omega_1(t)\Omega_4^*(t)}{2\Delta_{14} + i\gamma}e^{i\Phi_{14}} + \frac{\Omega_2(t)\Omega_3^*(t)}{2\Delta_{23} + i\gamma}e^{i\Phi_{23}}\right),$$
$$H_{3,1}^{TL} = \frac{\hbar}{2}\left(\frac{\Omega_1^*(t)\Omega_4(t)}{2\Delta_{14} + i\gamma}e^{-i\Phi_{14}} + \frac{\Omega_2^*(t)\Omega_3(t)}{2\Delta_{23} + i\gamma}e^{-i\Phi_{23}}\right), \tag{7}$$

$$H_{3,3}^{TL} = \frac{\hbar}{2}\left(\frac{|\Omega_4(t)|^2}{2\Delta_{14} + i\gamma} + \frac{|\Omega_3(t)|^2}{2\Delta_{23} + i\gamma}\right),$$

where

$$\Phi_{14} = \varphi_1 - \varphi_4 + (\delta_1 - \delta_4)t - (k_1 + k_4)z,$$

$$\Phi_{23} = \varphi_2 - \varphi_3 + (\delta_2 - \delta_3)t + (k_2 + k_3)z.$$

Since the detunings are small as compared to the corresponding transition frequencies, the relation

$$k_1 + k_4 = k_2 + k_3 = 2k_{av},$$

where k_{av} is the average photon momentum, holds to a high accuracy.

Non-Hermitian Hamiltonian (7) with complex Rabi frequencies describes a two-level system moving in the field defined by a complex potential. An analysis of scattering in an analogous field was performed in [15] for a two-level atom interacting with a standing wave with allowance for relaxation. In particular, it was shown that an extremely narrow distribution of atoms along the light-propagation direction can evolve in the field of two standing waves, with a temperature much smaller than the recoil energy.

2.3. Schrödinger Equation in the Momentum Representation

To be specific, we represent the atom as a plane wave and assume that the z component of its velocity is zero (the Doppler shift due to finite velocity can be subsumed under the detunings). Since Hamiltonian (7) is a periodic function z with period π/k_{av}, the probability amplitudes of $|1\rangle$ and $|3\rangle$ should be sought as series expansions in the eigenfunctions

$$\langle z|2n\rangle = \exp(2ink_{av}z)$$

of the z-component of the momentum operator in the coordinate representation, where n is an arbitrary integer and $2n\hbar k_{av}$ is the corresponding eigenvalue. The resulting expansion shows that the probability amplitude of the state $|1\rangle$ with the momentum z-component $2n\hbar k_{av}$ is related to the analogous probability amplitude for $|3\rangle$ with $(2n \pm 2)\hbar k_{av}$. In other words, if the initial state is $|1\rangle$ with zero z-component of momentum, then the only nonzero probability amplitudes are those of the state $|1\rangle$ with the z-component $4n\hbar k_{av}$ and the state $|3\rangle$ with the z-component $2(2n + 1)\hbar k_{av}$. Then, the state of the atom can be represented by a vector $\mathbf{B}$ whose $2n$th component is the probability amplitude of the state $|1\rangle$ with the z-component $4n\hbar k_{av}$ and $(2n + 1)$th component is the probability amplitude of the state $|3\rangle$ with the z-component $2(2n + 1)\hbar k_{av}$. The corresponding probability that the atom located at a point z is in the state $|1\rangle$ or $|3\rangle$ is

$$C_1(z, t) = \sum_{n=-\infty}^{\infty} B_{2n}(t)e^{i\Phi_{2n}}\langle z|4n\rangle,$$
$$C_3(z, t) = \sum_{n=-\infty}^{\infty} B_{2n+1}(t)e^{i\Phi_{2n+1}}\langle z|2(2n + 1)\rangle, \tag{8}$$

respectively. It is obvious that the phases in (8) can be arbitrary. Setting

$$\Phi_{2n} = n(\phi_2 - \phi_1 + \phi_4 - \phi_3) + \frac{1}{2}(\delta_0 - 2n\delta_d)t,$$
$$\Phi_{2n+1} = \Phi_{2n} - \frac{1}{2}(\phi_2 + \phi_1 - \phi_4 - \phi_3) - \delta_0 t, \tag{9}$$

where

$$\delta_0 = \frac{1}{2}(\delta_1 - \delta_4 + \delta_2 - \delta_3),$$
$$\delta_d = \delta_1 - \delta_4 - \delta_2 + \delta_3, \tag{10}$$

in (7), we obtain the Schrödinger equation in the momentum representation:

$$i\hbar\frac{\partial}{\partial t}\mathbf{B}(t) = H^M\mathbf{B}(t), \tag{11}$$

where the nonzero matrix elements H^{M} are

$$H^{\mathrm{M}}_{2n,\,2n-1} = \frac{\hbar}{2}\frac{\Omega_2\Omega_3^*}{2\Delta_{23}+i\gamma},$$

$$H^{\mathrm{M}}_{2n,\,2n} = \frac{\hbar}{2}\left(\frac{|\Omega_1|^2}{2\Delta_{14}+i\gamma}+\frac{|\Omega_2|^2}{2\Delta_{23}+i\gamma}+\delta_0-n\delta_{\mathrm{d}}\right),$$

$$H^{\mathrm{M}}_{2n,\,2n+1} = \frac{\hbar}{2}\frac{\Omega_1\Omega_4^*}{2\Delta_{14}+i\gamma},$$

$$H^{\mathrm{M}}_{2n+1,\,2n} = \frac{\hbar}{2}\frac{\Omega_1^*\Omega_4}{2\Delta_{14}+i\gamma}, \qquad (12)$$

$$H^{\mathrm{M}}_{2n+1,\,2n+1} = \frac{\hbar}{2}\left(\frac{|\Omega_3|^2}{2\Delta_{23}+i\gamma}+\frac{|\Omega_4|^2}{2\Delta_{14}+i\gamma}\right.$$

$$\left.-\delta_0-\frac{1}{2}(2n+1)\delta_{\mathrm{d}}\right),$$

$$H^{\mathrm{M}}_{2n+1,\,2n+2} = \frac{\hbar}{2}\frac{\Omega_2^*\Omega_3}{2\Delta_{23}+i\gamma}.$$

Note that the equation is independent of the initial phases of the fields. This is a consequence of the assumed carrier-frequency detuning from resonance.

According to (11) and (12), the case when γ is negligible and condition (5) holds corresponds to an effective two-level atom interacting with two laser pulses without losses. The corresponding Rabi frequencies are

$$\Omega_{14} = \frac{1}{4}\frac{\Omega_1\Omega_4}{\Delta_{14}}, \quad \Omega_{23} = \frac{1}{4}\frac{\Omega_2\Omega_3}{\Delta_{23}}.$$

Each pulse induces a transition between the states $|1\rangle$ and $|3\rangle$ and a change of $2\hbar k_{\mathrm{av}}$ in the atom's momentum. If

$$\max(\Omega_{14}, \Omega_{23})\tau_{\mathrm{int}} \gg 1, \qquad (13)$$

then, by analogy with the scheme of interaction between a two-level atom and the bichromatic field of counterpropagating pulses considered in [13], the momentum equal to an integer multiple of $2\hbar k_{\mathrm{av}}$ can be transferred from the field to the atom with or without change in the internal state of the atom. It should also be noted that Eqs. (11) and (12) are qualitatively different from those obtained in [13] in that they involve Stark shifts quadratic in the Rabi frequencies.

The spontaneous decay rate γ is contained in both diagonal and off-diagonal matrix elements in (12) and is responsible for their imaginary parts, which describe the decrease in the metastable-state populations due to transitions to states other than $|1\rangle$, $|2\rangle$, and $|3\rangle$. These transitions can be ignored if

$$\max(\gamma_1, \gamma_3)\tau_{\mathrm{int}} \ll 1, \qquad (14)$$

where

$$\gamma_1 = \frac{\gamma}{4}\left(\frac{|\Omega_1|^2}{\Delta_{14}^2+\frac{1}{4}\gamma^2}+\frac{|\Omega_2|^2}{\Delta_{23}^2+\frac{1}{4}\gamma^2}\right),$$

$$\gamma_3 = \frac{\gamma}{4}\left(\frac{|\Omega_3|^2}{\Delta_{23}^2+\frac{1}{4}\gamma^2}+\frac{|\Omega_4|^2}{\Delta_{14}^2+\frac{1}{4}\gamma^2}\right). \qquad (15)$$

Criteria (13) and (14) are satisfied simultaneously only if

$$\max(|\Delta_{14}|, |\Delta_{23}|) \gg \gamma. \qquad (16)$$

This necessary, but not sufficient, condition is assumed to hold in what follows.

With increasing field strength, adiabaticity criterion (13) is satisfied more reliably, whereas condition (14) becomes less reliable. For condition (14) to hold, an increase in the field strength must be combined with a proportional increase in the detunings Δ_{14} and Δ_{23}.

Since Eqs. (11) and (12) cannot be solved analytically, we discuss their numerical solutions and demonstrate that a substantial momentum can be transferred to the atom in the field of counterpropagating pulses with a small standard deviation.

2.4. Numerical Simulation

We set $B_0 = 1$, and the remaining B_n are set to zero, since the atom is modeled by a plane wave and the z component of its velocity is zero at the initial moment. We consider pulse shapes of the form

$$\Omega_1 = \Omega_{\max}F\left(\frac{t-t_1}{\tau}\right),$$

$$\Omega_2 = \Omega_{\max}F\left(\frac{t-t_2}{\tau}\right),$$

$$\Omega_3 = \Omega_{\max}F\left(\frac{t-t_3}{\tau}\right), \qquad (17)$$

$$\Omega_4 = \Omega_{\max}F\left(\frac{t-t_4}{\tau}\right),$$

where

$$F(x) = \begin{cases}\cos^2(\pi x), & -\frac{1}{2}\le x\le\frac{1}{2}\\ 0, & x\le-\frac{1}{2},\quad x\ge\frac{1}{2}.\end{cases} \qquad (18)$$

In this case (e.g., in contrast to $\cos(\pi x)$), the time derivatives vanish at the initial and final instants of the atom–field interaction, which makes it easier to simulate adiabatic passage. The Gaussian pulse shape commonly

used in simulations is not considered in this study since a real pulse cannot have an infinite length. Moreover, the atom tends to return to its initial state without any change in its momentum as the pulse length (or field strength) is increased, because the time-dependent eigenvalues corresponding to infinite-length pulses never cross.

To be specific, we assume

$$\Delta_{14} = -\Delta_{23} = \Delta_0. \tag{19}$$

Under this condition, criterion (5) for adiabatic elimination of the excited-state amplitude is satisfied if

$$\Delta_0 \gg \Omega_{\max}.$$

The parameters Δ_0 and $\Omega_{\max}$ determine the Rabi frequencies Ω_{14} and Ω_{23} for two-photon processes, with maximum value

$$\Omega_0 = \frac{1}{4}\frac{\Omega_{\max}^2}{\Delta_0}$$

and the Stark shifts

$$\Delta_{S1} = \frac{1}{4}\frac{|\Omega_1|^2}{\Delta_{14}} + \frac{1}{4}\frac{|\Omega_2|^2}{\Delta_{23}},$$

$$\Delta_{S3} = \frac{1}{4}\frac{|\Omega_3|^2}{\Delta_{23}} + \frac{1}{4}\frac{|\Omega_4|^2}{\Delta_{14}}$$

of $|1\rangle$ and $|3\rangle$, respectively.

Population transfer between atomic states depends both on time delay between Ω_{14} and Ω_{23} and on dynamic Stark splitting. The latter effect was studied in [16] for a two-level atom interacting with a pulse field in the regime of laser-induced Stark splitting. Here, these effects are analyzed separately.

2.4.1. Delayed two-photon processes. First, we consider a pulse model closest to the two-level atom interacting with counterpropagating pulses analyzed in [13]. We set

$$t_1 = t_4 = t_d/2, \quad t_2 = t_3 = -t_d/2; \tag{20}$$

i.e., the counterpropagating pulse pairs 1, 4 and 2, 3 give rise to Rabi frequencies of similar form with delay t_d. Under these conditions, the Stark shift $\Delta_{S1} = \Delta_{S3}$ cannot affect population and momentum transfer, which depend only on the difference between the Stark shifts of $|1\rangle$ and $|3\rangle$.

Figure 2 shows the average z-component of momentum in the dressed state $|1\rangle$,

$$p_1 = \frac{2\hbar k_{av}}{n_1}\sum_{j=-\infty}^{\infty} 2j|B_{2j}|^2, \tag{21}$$

and the corresponding standard deviation,

$$\Delta p_1 = \sqrt{\frac{4\hbar^2 k_{av}^2}{n_1}\sum_{j=-\infty}^{\infty}(2j)^2|B_{2j}|^2 - p_1^2}, \tag{22}$$

versus Ω_0/δ_d. Here,

$$n_1 = \sum_{j=-\infty}^{\infty}|B_{2j}|^2$$

is the population of the state $|1\rangle$. It is clear that $p_1(\Omega_0/\delta_d)$ is a steplike function and Δp_1 drops in the plateau regions. With increasing duration of the atom–field interaction, the degree of nonadiabaticity decreases, which leads both to decrease in Δp_1 in the plateau regions and to increase in the plateau width. The fact that $\Delta p_1 \ll \hbar k_{av}$ in the plateau regions implies that the z-components of the momenta of almost all atoms change by $p_1 = 0, 4\hbar k_{av}, 8\hbar k_{av}, \ldots$; i.e., the transferred momentum is an integer multiple of $4\hbar k_{av}$. In the jump regions, which become narrower with increasing pulse width, $|B_{2j}|^2 \sim 1$ for at least two values of j, and Δp_1 reaches a maximum value.

The figure demonstrates that both transferred-momentum magnitude and the corresponding standard deviation are independent of spontaneous emission: the curves corresponding to zero and finite decay rates are indistinguishable in Figs. 2a and 2b. However, the ground-state population n_1 can significantly decrease when the decay is taken into account. This result is clear from the following argumentation. In the present model, which does not allow for spontaneous transitions to the states $|1\rangle$ and $|3\rangle$, spontaneous decay removes an atom from the ensemble interacting with the field without changing the momentum distributions of atoms in the states $|1\rangle$ and $|3\rangle$. As a result, both average momenta of atoms in the states $|1\rangle$ and $|3\rangle$ and the corresponding standard deviations do not change even though the populations of these states can change substantially.

In the interval of Ω_0/δ_d considered here, the average momentum is

$$p_3 = \frac{2\hbar k_{av}}{n_3}\sum_{j=-\infty}^{\infty}(2j+1)|B_{2j+1}|^2, \tag{23}$$

and the corresponding standard deviation is

$$\Delta p_3 = \sqrt{\frac{4\hbar^2 k_{av}^2}{n_3}\sum_{j=-\infty}^{\infty}(2j+1)^2|B_{2j+1}|^2 - p_3^2}, \tag{24}$$

where

$$n_3 = \sum_{j=-\infty}^{\infty}|B_{2j+1}|^2$$

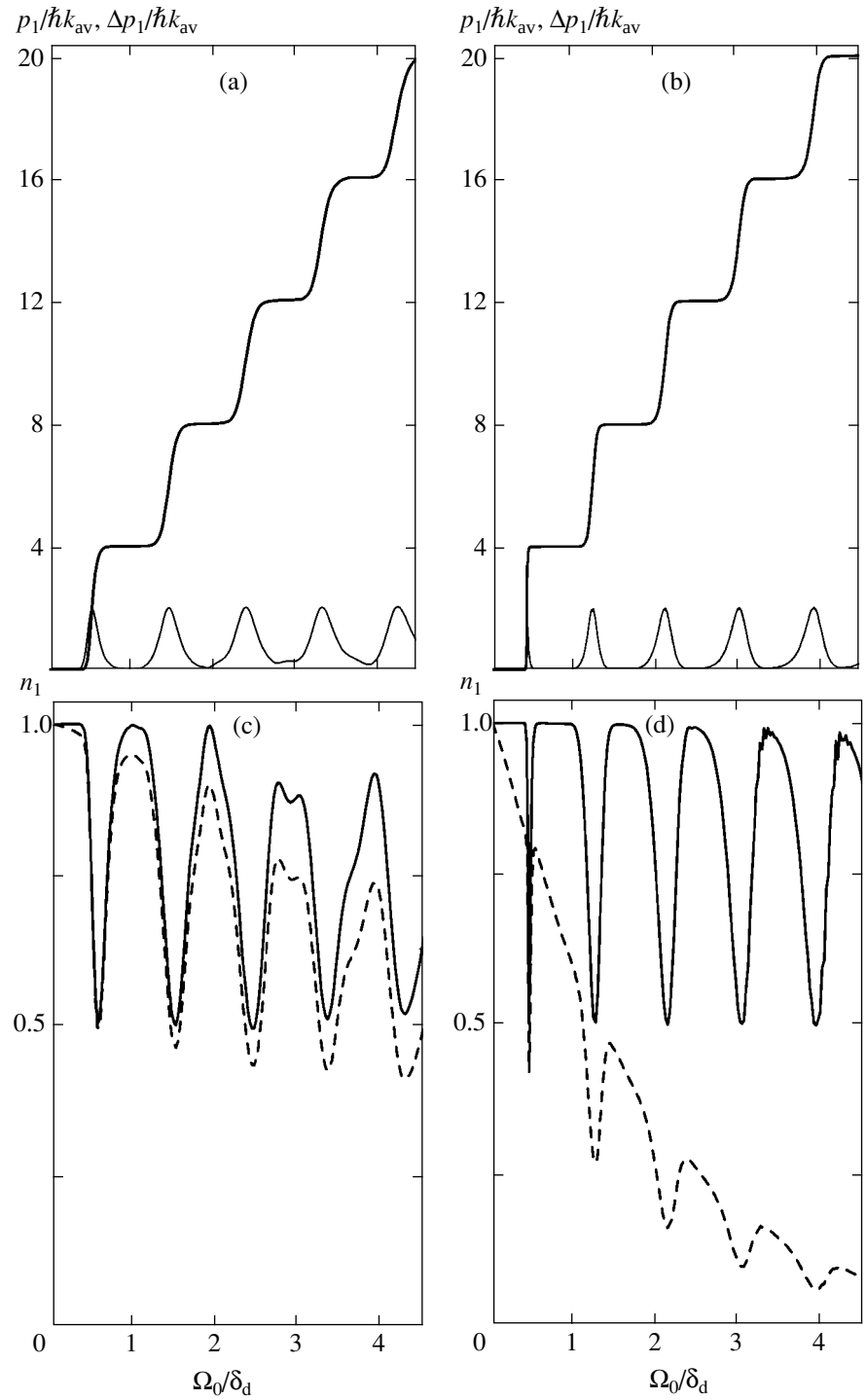

Fig. 2. (a, b) Mean z-component (thick curve) and standard deviation (thin curve) of transferred momentum measured in units of $\hbar k_{av}$ for atoms in state $|1\rangle$ and (c, d) population of state $|1\rangle$ after interaction with pulses defined by (17), (20) vs. ratio of the highest two-photon Rabi frequency to difference in two-photon detuning between the pulse pairs 1, 4 and 2, 3: $t_d = 0.4\tau$; $\delta_0 = 0$; $\delta_d\tau = 200$ (a, c), 2000 (b, d). Solid and dashed curves correspond to $\gamma = 0$ and $\gamma/\Delta_0 = 0.0002$, respectively.

is the population of the state $|3\rangle$, are zero and $2\hbar k$, respectively, to a high accuracy. For the parameter values specified in the figure caption, the momentum distributions of atoms in the state $|3\rangle$ have sharp peaks at $+2\hbar k$ and $-2\hbar k$.

When Ω_0/δ_d is large, the values of the left-hand side of (14) corresponding to the dashed curves in Figs. 2c and 2d are 0.2 and 2, respectively, if we set $\tau_{int} = \tau$. These estimates are consistent with the weak effect of spontaneous radiative decay on population transfer in

the former case and the substantial decrease in the population of the state $|1\rangle$ in the latter.

When the probability of spontaneous transitions to the states $|1\rangle$ and $|3\rangle$ are higher than or comparable to that of transitions to other states, relaxation can strongly affect the momentum distribution function, because spontaneous emission increases the momentum spread. However, this effect can be virtually eliminated by tuning the carrier frequencies farther off resonance and simultaneously increasing the pulse energies so that criteria (13) and (14) hold.

The steplike behavior of transferred momentum as a function of Rabi frequency and other parameters (e.g., δ_0 and t_d) can be explained by analyzing the adiabatic-state quasienergies (eigenvalues of Hamiltonian (12)). Both before and after the atom–field interaction (when all Rabi frequencies are zero), the quasienergies of the adiabatic states corresponding to $|1\rangle$ are

$$\hbar\varpi_{2n} = \frac{1}{2}\hbar\delta_0 - \frac{n}{2}\hbar\delta_d,$$

whereas the quasienergies of the adiabatic states corresponding to $|3\rangle$ are

$$\hbar\varpi_{2n+1} = -\frac{1}{2}\hbar\delta_0 - \frac{2n+1}{4}\hbar\delta_d.$$

If the difference between the adiabatic-state energy initially equal to $\hbar\varpi_0$ and other adiabatic-state energies remains large throughout the interaction and equals $\hbar\varpi_n$ after the interaction, then the expected change in the z-component of momentum is $2n\hbar k_{av}$. When the atom–field interaction parameters (field strength, detuning, and the delay between the pulses) are varied within certain limits, the configuration of the quasienergy curves does not change, and the momentum transferred to the atom is virtually independent of these parameters. After the varied parameter values pass a domain where the quasienergies are comparable and the atomic state cannot be described by a single adiabatic state, the configuration of the curves is different, and a different amount of momentum is transferred. This can be explained by discussing the following example.

Figure 3a illustrates the evolution of adiabatic-state quasienergies $\hbar\varpi_n$ measured in units of $\hbar\delta_d$ for parameter values that ensure the transfer of $4\hbar k_{av}$ to the atom. The quasienergy of the adiabatic state corresponding to the ground state with zero initial momentum projection on the pulse propagation direction is represented by the thick curve. The vertical lines mark the endpoints of the time interval of simultaneous interaction between the atom and all pulses. Since the atom initially interacts with pulses 2 and 3, only the pairs of amplitudes B_{2n} and B_{2n-1} can be coupled by the field during the time interval on the left of the vertical lines (see (12)). These

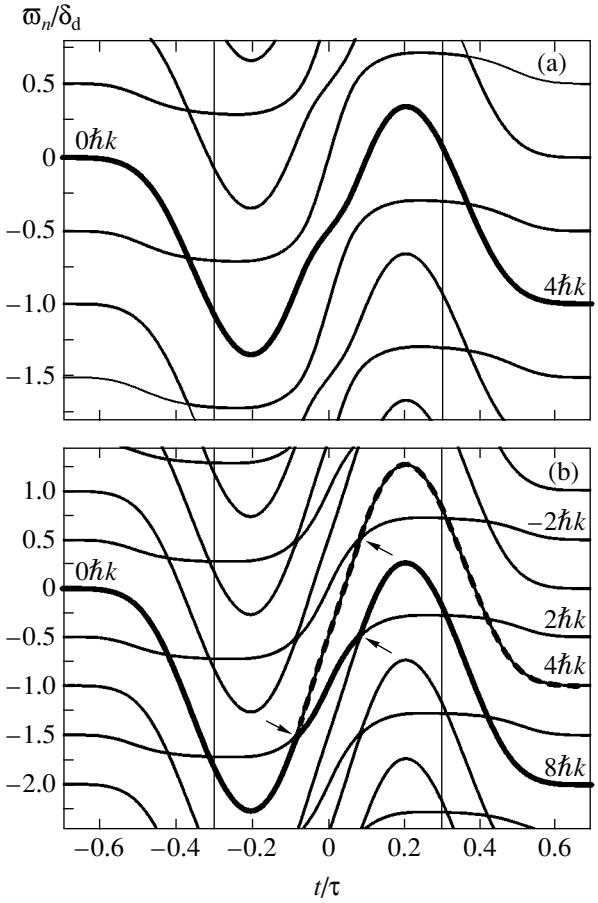

Fig. 3. Adiabatic-state quasienergies for Hamiltonian (12) in units of $\hbar\delta_d$ vs. time for pulses defined by (17), (20): $t_d = 0.4\tau$; $\gamma = 0$; $\delta_0 = 0$; $\Omega_0/\delta_d = 0.78125$ (a), 1.25 (b). Thin vertical lines mark the endpoints of the time interval of simultaneous interaction between the atom and all pulses.

pairs are associated with the adiabatic states whose initial energies are

$$\hbar\varpi_{2n} = \frac{1}{2}\hbar\delta_0 - \frac{n}{2}\hbar\delta_d$$

and

$$\hbar\varpi_{2n-1} = -\frac{1}{2}\hbar\delta_0 - \frac{2n-1}{2}\hbar\delta_d,$$

and the state of the atom can change only within such a pair. Since the difference of the corresponding quasienergies is much greater than $\hbar/\tau$, the atom remains in the initial adiabatic state represented by the thick curve. At $t = -0.35\tau$, this curve crosses one that corresponds to another pair of adiabatic states coupled by the field, but the state of the atom does not change. Similarly, the atomic state does not change at $t = +0.35\tau$, another intersection of quasienergy curves. According to the figure, the atomic state after the interaction corresponds to the energy $-\hbar\delta_d$; i.e., the z-component of momentum changes by $4\hbar k_{av}$.

The configuration of curves representing time-dependent quasienergies changes with increasing field strength: a new intersection appears before the left vertical line is reached, and the quasienergy configuration takes the form illustrated by Fig. 3b ($\Omega_0/\delta_d = 1.25$). This value of Ω_0/δ_d corresponds to the change $4\hbar k_{av}$ in the mean momentum z-component in the state $|1\rangle$ when $\delta_d\tau = 200$ (Fig. 2a). As $\delta_d\tau$ is increased to 2000, the mean z-component of the momentum transferred to the atom changes to an intermediate value between $4\hbar k_{av}$ and $8\hbar k_{av}$ (Fig. 2b), and the corresponding standard deviation approaches its maximum value $2\hbar k_{av}$. With further increase in the duration of atom–field interaction, the change in the z-component of momentum of an atom in the state $|1\rangle$ becomes $8\hbar k_{av}$, while the standard deviation decreases to a value much smaller than $\hbar k_{av}$. This behavior is illustrated by the curves of time-dependent quasienergies presented in Fig. 3b. Unlike Fig. 3a, Fig. 3b shows examples of avoided crossings of converging quasienergy curves. In the neighborhoods of the avoided crossings (indicated by arrows), Landau–Zener transitions can occur between the adiabatic states represented by the converging curves. These transitions are characterized by a probability P, which exponentially depends on the time of passage through the avoided crossings and varies from unity to zero with increasing pulse width. If condition (13) holds and the pulse is not very wide, the state of the atom follows the thick curve until the point indicated by an arrow is reached. After a Landau–Zener transition occurs at this point (with a probability close to unity), the state follows the dashed curve. The metastable state $|3\rangle$ is populated with a small probability $1 - P$, with the momentum z-component $+2\hbar k_{av}$. After the next Landau–Zener transition, the atomic state either follows (with a high probability) the dashed curve or changes (with a small probability) to the metastable state $|3\rangle$, with the momentum z-component $-2\hbar k_{av}$. The final state of the atom is $|1\rangle$ with the momentum z-component $+4\hbar k_{av}$. With increasing atom–field interaction time, the Landau–Zener transition probability tends to zero, and the atom remains in the adiabatic state represented by the thick curve. The final state of the atom is $|1\rangle$ with the momentum z-component $+8\hbar k_{av}$. Thus, an increase in the duration of the atom–field interaction must shift the steps in p_1 leftwards, which is clear from comparison of Figs. 2a and 2b. Taking into account the Landau–Zener transitions between converging quasienergy curves, we find that the relative population of a final atomic state with momentum $\pm 2\hbar k_{av}$ is $P(1 - P)$. Therefore, the mean value of the z-component of transferred momentum for an atom in the state $|3\rangle$ is zero, and the corresponding standard deviation is $+2\hbar k_{av}$, in agreement with numerical results.

Figure 4 shows the mean z-component of the momentum transferred to the atom, the corresponding standard deviation for $|1\rangle$ and $|3\rangle$, and the populations of these states as functions of δ_0/δ_d. If the mean two-photon detuning is such that

$$-\frac{1}{2}\delta_d < \delta_0 < \frac{1}{2}\delta_d$$

then the atom gains a momentum of $8\hbar k_{av}$, the final state $|3\rangle$ is weakly populated, and the population of $|1\rangle$ is close to unity. Otherwise, the atom gains the momentum $10\hbar k_{av}$ over a wide detuning range (in various directions), and the state $|3\rangle$ is mainly populated. This behavior, as well as that illustrated by Fig. 2, is analogous to the results obtained in [13] for a two-level atom in the bichromatic field of counterpropagating pulses. When $\delta_0 = (\pm 1/2)\delta_d$, the plane wave representing the atom splits into two waves of equal intensity having the momenta $8\hbar k_{av}$ (state $|1\rangle$) and $\mp 10\hbar k_{av}$ (state $|3\rangle$). The physical explanation of the splitting lies in the fact that this condition corresponds to a two-photon resonance between the pairs of pulses 1, 4 and 2, 3. Accordingly, the initial state of the atom is a linear superposition of two adiabatic states. These states adiabatically evolve into states $|1\rangle$ and $|3\rangle$ with different momenta. Note also that an atom can also change its state and simultaneously gain momentum when $\delta_0 = 0$ and the pulses have different amplitudes (see [17]).

2.4.2. Simultaneous two-photon processes. Now, we set

$$t_1 = t_3 = t_d/2, \quad t_2 = t_4 = -t_d/2, \tag{25}$$

so that $\Omega_{14} = \Omega_{23}$. This case is qualitatively different from the two-level system considered in [13], where a delay between the pulses is required to ensure that the standard deviation of the momentum transferred to the atom is small. However, the difference in the Stark shift between the states $|1\rangle$ and $|3\rangle$ makes it possible to transfer a substantial momentum to the atom with a small standard deviation in this case as well.

Suppose that criterion (14) holds, i.e., spontaneous radiative decay of the excited state does not affect momentum and population transfer. Figure 5 shows the change in the mean z-component of momentum, the corresponding standard deviation in state $|1\rangle$, and the population of this state as functions of Ω_0/δ_d calculated under condition (25). Analogous results obtained for $|3\rangle$ are not shown here. (It suffices to note that the mean transferred momentum and the corresponding standard deviation are zero and $2\hbar k_{av}$, respectively.) As in Fig. 2b (Eq. (20)), the function $p_1(\Omega_0/\delta_d)$ plotted in Fig. 5 behaves as a steplike function and Δp_1 drops in the plateau regions. Thus, the effect of difference in dynamic Stark shift between $|1\rangle$ and $|3\rangle$ is similar to the role played by the delay between pulses. This is quite natural, since the detunings of the pairs of pulses 1, 4 and 2, 3 from two-photon resonance are different, and the gradual variation of the difference of the Stark shifts

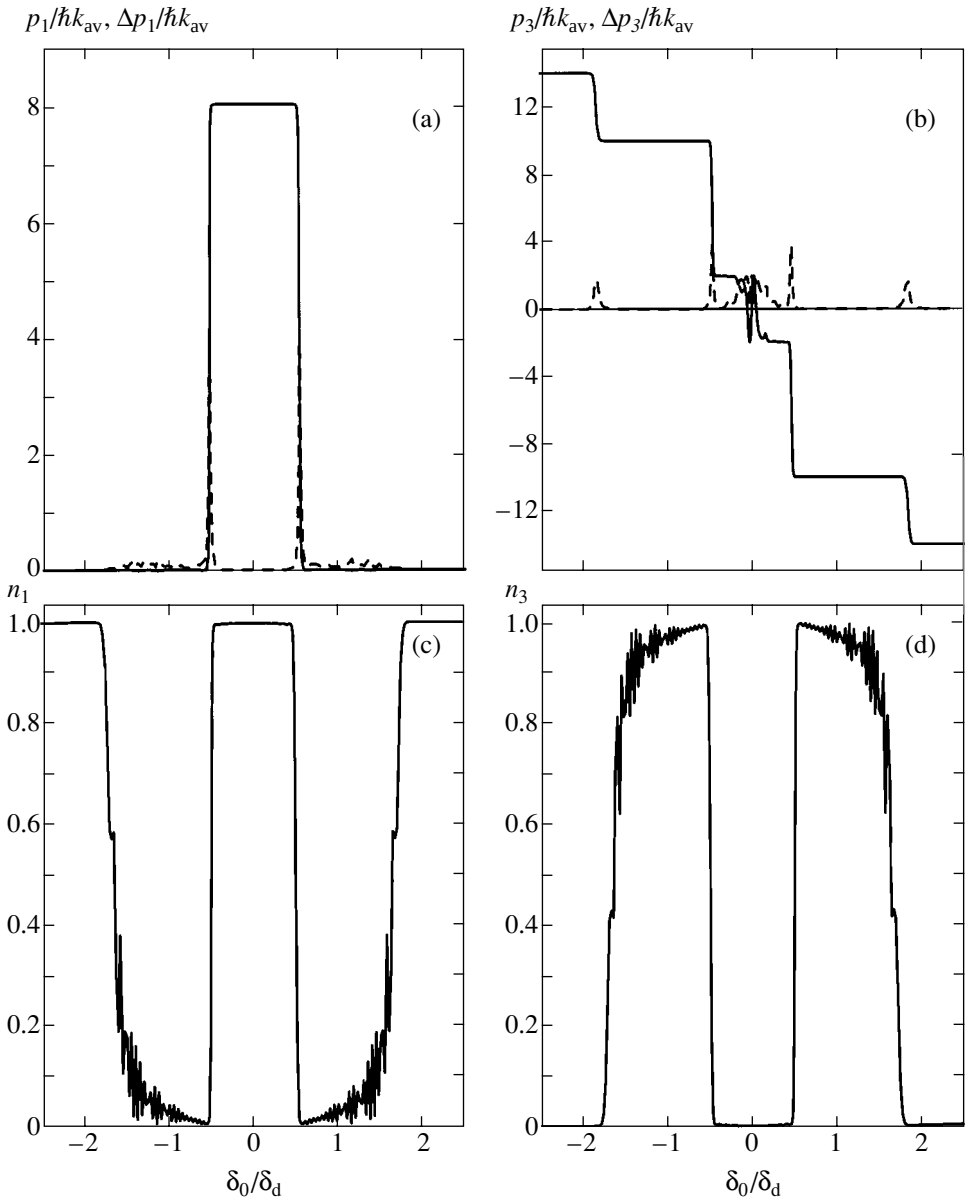

Fig. 4. Mean z-component (thick curve) and standard deviation (thin curve) of transferred momentum measured in units of $\hbar k_{av}$ for states $|1\rangle$ (a) and $|3\rangle$ (b) and populations of $|1\rangle$ (c) and $|3\rangle$ (d) after interaction with pulses defined by (17) vs. ratio of mean two-photon Rabi frequency to two-photon detuning difference δ_0/δ_d: $\Omega_0/\delta_d = 1.55$; $t_d = 0.4\tau$; $\delta_d\tau = 2000$.

of $|1\rangle$ and $|3\rangle$ (see Fig. 6) leads to difference between the corresponding atom–field interaction conditions. The interaction is stronger for one pair at an early stage and for the other pair at a late stage. The resulting process is qualitatively similar to the interaction under condition (20).

3. THE CASE OF SMALL ENERGY DIFFERENCE BETWEEN THE METASTABLE STATES

In this section, we consider the case when the energy difference between the metastable states (measured in frequency units) is close to the Rabi frequencies that characterize the atom–field interaction. In this case, the atom in either metastable case interacts with each pulse (unless this is forbidden by selection rules). Therefore, as in the case of a two-level atom [13], analysis of momentum transfer can be restricted to counterpropagating pulses 1 and 2 (by setting $\omega_3 = \omega_1$ and $\omega_4 = \omega_2$ in Fig. 1). Note that the case of copropagating laser pulses, when momentum cannot be transferred from the field to the atom, was considered in [14].

The pulse propagating in the positive z direction induces $1 \longleftrightarrow 2$ transitions with a Rabi frequency Ω_1 and $2 \longleftrightarrow 3$ transitions with a Rabi frequency Ω_3. The Rabi frequencies Ω_2 and Ω_4 are associated with the

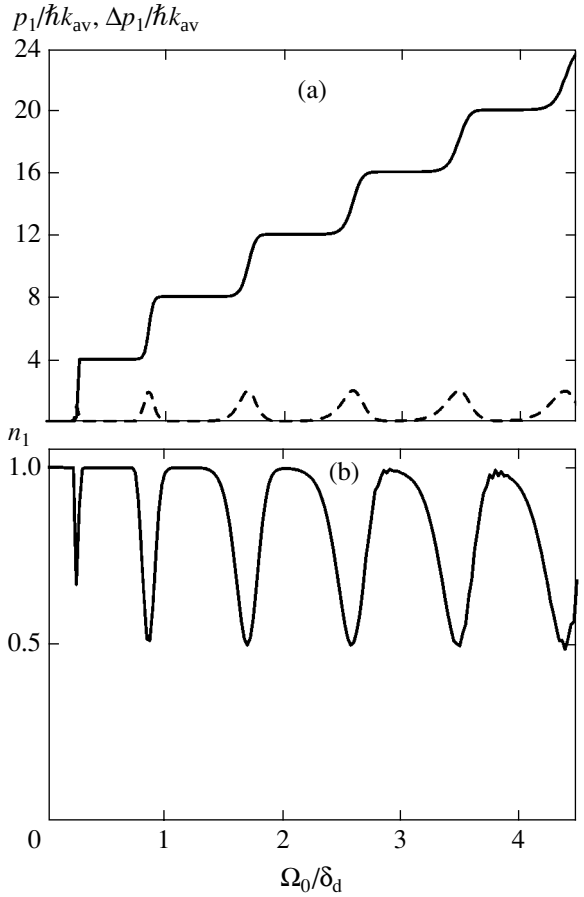

Fig. 5. (a) Mean (solid curve) and standard deviation (thin curve) of transferred momentum measured in units of $\hbar k_{av}$ for atoms in state $|1\rangle$ and (b) population of state $|1\rangle$ after interaction with pulses defined by (17), (25) vs. ratio of the highest two-photon Rabi frequency to difference in two-photon detuning between the pulse pairs 1, 4 and 2, 3: $\delta_d\tau = 2000$; $t_d = 0.4\tau$; $\delta_0 = 0$; $\gamma = 0$.

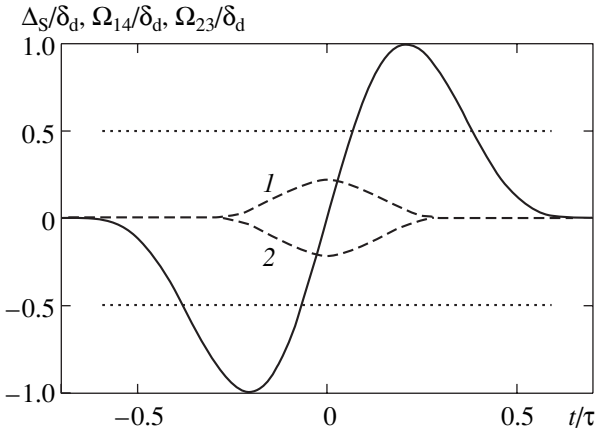

Fig. 6. Evolution of Stark shift difference $\Delta_S = \Delta_{S1} - \Delta_{S3}$ between $|1\rangle$ and $|3\rangle$ (solid curve) and two-photon Rabi frequencies Ω_{14} (curve *1*) and Ω_{23} (curve *2*) measured in units of δ_d: $t_d = 0.4\tau$; $\delta_0 = 0$; $\Omega_0/\delta_d = 0.5$. Intersections of solid curve with dotted lines correspond to resonances with pulse pair 1, 4 or 2, 3.

$1 \longleftrightarrow 2$ and $2 \longleftrightarrow 3$ transitions induced by the pulse propagating in the negative z direction. The atom–field interaction is described by the Schrödinger equation (3) subject to the following additional conditions: (i) the time dependence of Ω_3 is similar to that of Ω_1 up to a constant factor, and analogous similarity conditions holds both for Ω_2 and Ω_4 and for $\phi_1 = \phi_3$ and $\phi_2 = \phi_4$; (ii) the detunings δ_1, δ_2, δ_3, and δ_4 satisfy the relation

$$\delta_3 - \delta_1 = \delta_4 - \delta_2 = S, \tag{26}$$

where $S = (W_3 - W_1)/\hbar$; (iii) since $W_3 - W_1 \ll W_2 - W_1$, the difference between the wavevectors of the counter-propagating pulses can by neglected by setting $k_i = k$ ($i = 1, \ldots, 4$).

Changing to the momentum representation, we seek the probability amplitude for each state $|j\rangle$ in the form

$$C_j(z, t) = \sum_{n=-\infty}^{\infty} A_{j,n}(t)\exp(i\Phi_{j,n}(t))\langle z|n\rangle, \tag{27}$$

where

$$\langle z|n\rangle = \exp(inkz), \quad j = 1, 2, 3.$$

It is obvious that the phases in (8) can be arbitrary. Setting

$$\Phi_{2,n} = \frac{n}{2}(\phi_2 - \phi_1) + \frac{1}{2}(\delta_2 - \delta_1)nt,$$

$$\Phi_{1,n} = \Phi_{2,n+1}(t) + \phi_1 + \delta_1 t, \tag{28}$$

$$\Phi_{3,n} = \Phi_{2,n+1}(t) + \phi_1 + (\delta_1 + S)t,$$

we obtain a set of equations for the probability amplitudes corresponding to the states $|j\rangle$ with momentum $n\hbar k$:

$$\frac{\partial A_{1,n}}{\partial t} = \frac{i}{2}[\delta_1(n-1) - \delta_2(n+1)]A_{1,n} - \frac{i}{2}\Omega_1 A_{2,n+1} - \frac{i}{2}\Omega_2 A_{2,n-1},$$

$$\frac{\partial A_{2,n}}{\partial t} = \frac{i}{2}[\delta_1 n - \delta_2 n + i\gamma]A_{2,n} - \frac{i}{2}\Omega_1^* A_{1,n-1} - \frac{i}{2}\Omega_2^* A_{1,n+1} - \frac{i}{2}\Omega_3^* A_{3,n-1} - \frac{i}{2}\Omega_4^* A_{3,n+1}, \tag{29}$$

$$\frac{\partial A_{3,n}}{\partial t} = \frac{i}{2}[\delta_1(n-1) - \delta_2(n+1) - 2S]A_{3,n} - \frac{i}{2}\Omega_3 A_{2,n+1} - \frac{i}{2}\Omega_4 A_{2,n-1}.$$

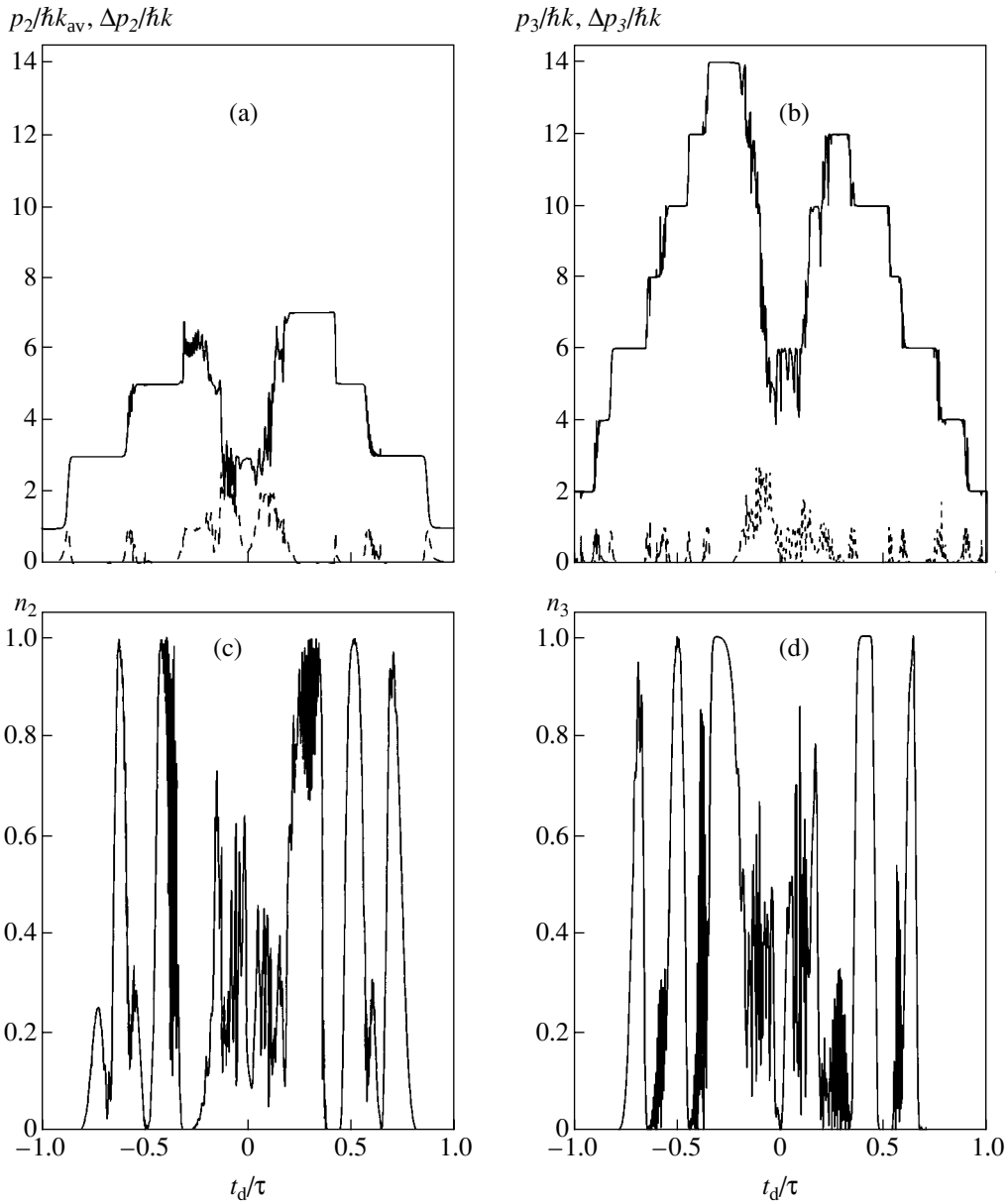

Fig. 7. Mean z-component and standard deviation of transferred momentum measured in units of photon momentum $\hbar k$ for atoms in states $|2\rangle$ (a) and $|3\rangle$ (b) and populations of these states (c, d) after interaction with pulses defined by (17), (25) vs. delay between pulses in the absence of spontaneous radiative decay of the excited state during interaction. Parameters of pulse and detuning are given in text.

One example of the scheme of atom–field interaction considered in this section is the atom whose metastable state is characterized by the total angular momentum $J = 1$ (i.e., the angular-momentum projections $M = -1$, 0, and 1 on the z axis) and excited state has the angular momentum $J = 0$. When a magnetic field parallel to the z axis is applied, the degeneracy with respect to M is removed, and the frequencies of the transitions between the state with $M = 0$ (upper level) and those with $M = 1$ and $M = -1$ differ by a quantity S proportional to the magnetic induction. Since the $J = 0$, $M = 0 \longrightarrow J = 1$, $M = 0$ transitions are forbidden, we have a three-level scheme of atom–field interaction. If the pulses are linearly polarized, then each pulse induces both admissible transitions with the same Rabi frequency.

Omitting both physical explanation of the possibility of transfer of a substantial momentum to the atom with a small standard deviation without changing its internal state and numerical results analogous to those presented in the preceding section, we discuss here the dependence of the transferred momentum on the delay between the pulses. A numerical analysis was performed for Rabi frequencies of the form defined by (17)

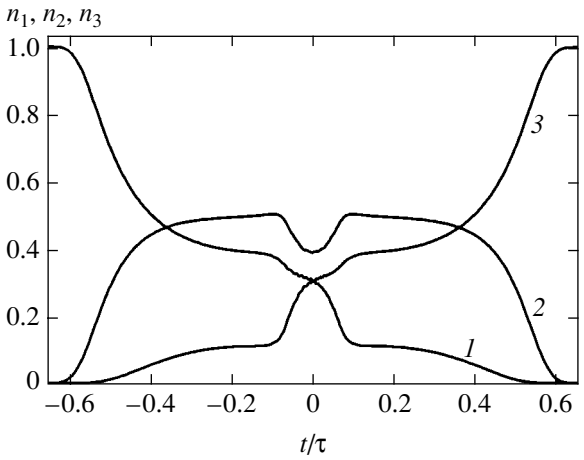

Fig. 8. Evolution of populations of states $|1\rangle$ (curve *1*), $|2\rangle$ (curve *2*), and $|3\rangle$ (curve *3*) during atom–field interaction. The final state is $|3\rangle$ with $p = 14\hbar k$. Detuning, pulse width, and Ω_0 are as in Fig. 7; $t_d = -0.31\tau$.

and (25) with $\Omega_{max} = 2S$ and for the relatively small detunings

$$\delta_1 = -\frac{1}{4}S, \quad \delta_2 = -\frac{3}{4}S,$$

which correspond to

$$\delta_3 = \frac{3}{4}S, \quad \delta_4 = \frac{1}{4}S$$

by virtue of (26). The pulse width was sufficiently large, $S\tau = 2000$, to ensure adiabatic atom–field interaction. Figures 7a and 7b demonstrate that the time-dependent mean changes in momentum *z*-components p_j ($j = 2, 3$) as functions of the delay between the pulses have plateaulike portions, where the corresponding standard deviation is virtually independent of the delay and $\Delta p \ll \hbar k$. Figures 7c and 7d show the populations of states $|2\rangle$ and $|3\rangle$ as functions of the delay between the pulses. Comparing them with Figs. 7a and 7b, one can determine the number of atoms in the states $|2\rangle$ and $|3\rangle$ and the momenta transferred to them during the atom–field interaction. Note that the steplike behavior of $p_j(t_d/\tau)$ predicted for the particular set of parameter values is observed only for atoms that are in states $|2\rangle$ and $|3\rangle$ after the interaction.

Thus, when the energy difference between the metastable states is relatively small as compared to $\hbar\Omega_0$, a substantial momentum ($\sim 10\hbar k$) can be transferred simultaneously with a change in the state of an atom interacting with two counterpropagating pulses, and the corresponding standard deviation of momentum is smaller than $\hbar^2 k^2$. However, the model of atom–field interaction discussed here is applicable only when the pulses are relatively short, because the excited state of the atom is increasingly populated in the course of its interaction with radiation (see Fig. 8). In this case, the excited state cannot be adiabatically eliminated even if the fields are tuned farther off resonance, because the resulting equations are characterized by identical time variation of two-photon Rabi frequencies and by equal Stark shifts for the states $|1\rangle$ and $|3\rangle$.

4. CONCLUSIONS

The possibility of momentum transfer to a three-level atom with metastable lower states due to its interaction with laser pulses is analyzed. It is shown that the mean momentum gained by the atom interacting with counterpropagating pulses can be an integer multiple of the photon momentum if the pulses overlap in time.

When the energy difference between the metastable states is so small that the atom in either state interacts with each laser pulse, the momentum of the atom interacting with two counterpropagating pulses can change by a substantial amount as compared to the photon momentum. In this case, the excited-state population is large during the atom–field interaction, and spontaneous radiative loss of atoms can be avoided only if the pulses have widths shorter than the spontaneous decay time.

When the energy difference between the metastable states is so large that each pulse couples only one metastable state to the excited state, the transfer of a substantial momentum, as compared to the photon momentum, requires the use of two pairs of counterpropagating laser pulses. If the carrier frequencies are tuned sufficiently far off resonance, but the pulse pairs induce two-photon transitions between the metastable states, then the momentum transfer is insensitive to spontaneous decay of the excited state, which remains virtually unpopulated in this case. This is important in view of possible applications of the proposed method of momentum transfer in experiments on atomic beams using continuous-wave lasers instead of pulsed lasers, when the effect of a light pulse on a moving atom is obtained when its trajectory crosses a laser beam and a typical atom—the field interaction time is much longer than the typical excited-state lifetime.

Physically, the possibility of momentum transfer in the field of counterpropagating laser pulses is due to the existence of adiabatic states that are identical to the atomic state with a certain momentum at the initial instant and to the same or different atomic state with a different momentum at the final instant of the atom–field interaction. If the adiabatic-passage condition is satisfied, then the transferred momentum is virtually insensitive to the values of the interaction parameters, such as the delay between the pulses or their peak intensities.

ACKNOWLEDGEMENTS

This work was supported by the National Academy of Sciences of Ukraine, project nos. VTs 93/24 and V/112.

REFERENCES

1. V. G. Minogin and V. S. Letokhov, *Laser Light Pressure on Atoms* (Nauka, Moscow, 1986; Gordon and Breach, New York, 1987).
2. A. P. Kazantsev, G. I. Surdutovich, and V. P. Yakovlev, *The Mechanical Action of Light on Atoms* (Nauka, Moscow, 1991) [in Russian].
3. B. D. Pavlik, *Ultracold and Cold Atoms* (Naukova Dumka, Kiev, 1993) [in Russian].
4. H. J. Metcalf and P. van der Stratten, *Laser Cooling and Trapping* (Springer, New York, 1999).
5. G. Alzetta, A. Gozzini, L. Moi, and G. Orriols, Nuovo Cimento B **36**, 5 (1976).
6. H. R. Gray, R. W. Whitley, and C. R. Stroud, Jr., Opt. Lett. **3**, 218 (1978).
7. B. D. Agap'ev, M. B. Gornyĭ, B. G. Matisov, and Yu. V. Rozhdestvenskiĭ, Usp. Fiz. Nauk **163** (9), 1 (1993) [Phys. Usp. **36**, 763 (1993)].
8. K. Bergmann, H. Theuer, and B. W. Shore, Rev. Mod. Phys. **70**, 1003 (1998).
9. P. Marte, P. Zoller, and J. L. Hall, Phys. Rev. A **44**, R4118 (1991).
10. B. W. Shore, K. Bergmann, A. Kuhn, *et al.*, Phys. Rev. A **44**, 7442 (1991).
11. L. S. Goldner, C. Gerz, R. J. C. Spreeuw, *et al.*, Phys. Rev. Lett. **72**, 997 (1994).
12. M. V. Danileĭko, A. M. Negriĭko, V. I. Romanenko, and L. P. Yatsenko, Ukr. Fiz. Zh. **40**, 793 (1995).
13. V. I. Romanenko and L. P. Yatsenko, Zh. Éksp. Teor. Fiz. **117**, 467 (2000) [JETP **90**, 407 (2000)].
14. R. Unanyan, S. Guérin, B. W. Shore, and K. Bergmann, Eur. Phys. J. D **8**, 443 (2000).
15. D. O. Chudesnikov and V. P. Yakovlev, Laser Phys. **1**, 110 (1991).
16. T. Rickes, L. P. Yatsenko, S. Steuerwald, *et al.*, J. Chem. Phys. **113**, 534 (2000).
17. S. Guérin, L. P. Yatsenko, and H. R. Jauslin, Phys. Rev. A **63**, 031403 (2001).

Translated by A. Betev

Journal of Experimental and Theoretical Physics, Vol. 100, No. 2, 2005, pp. 256–271.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 127, No. 2, 2005, pp. 289–307.

ATOMS, MOLECULES, OPTICS

Resonant Transparency Regimes under Conditions of Long/Short-Wave Coupling

S. V. Sazonov* and N. V. Ustinov
Kaliningrad State University, Kaliningrad, 236041 Russia
**e-mail: nst@alg.kaliningrad.ru*
Received February 13, 2004

Abstract—Nonlinear two-component electromagnetic pulse propagation through a resonant axially symmetric anisotropic medium having a permanent dipole moment is analyzed under conditions of strong coupling between the ordinary (short-wavelength) and extraordinary (long-wavelength) pulse components. It is shown that a pulse can propagate through the medium in regimes different from self-induced transparency if its ordinary component is detuned off resonance. In particular, a pulse propagating in the regime of self-induced supertransparency substantially changes quantum-level populations, but its group velocity remains almost equal to the linear velocity. If a pulse propagates in the extraordinary transparency regime and the carrier-frequency detuning from resonance is small, then its group velocity is substantially lower, while the level populations remain virtually invariant. Regimes of propagation through weakly excited media under quasi-resonance conditions are also identified. © *2005 Pleiades Publishing, Inc.*

1. INTRODUCTION

Owing to progress in experimental techniques, coherent optical phenomena in nonlinear media have remained a subject of permanent interest over the decades since the discovery of self-induced transparency (SIT) [1] (see reviews in [2, 3] and references therein). Substantial progress in theoretical analysis of these phenomena have been made by applying powerful mathematical tools, such as the inverse scattering method, to light–matter interaction models [4–6]. Various types of solutions to equations solvable by this method, including the so-called SIT equations [7] and the reduced Maxwell–Bloch equations [8], can be used to describe numerous processes in nonlinear optics. In particular, pulse propagation in integrable models can be associated with soliton solutions [7–14].

Coherent phenomena in those nonlinear media where quantum particles have a permanent dipole moment are an important subject of current studies [15–23]. If a medium is anisotropic, then the parity of quantum states is not well defined. Therefore, diagonal dipole matrix elements may not vanish. This is also possible for polar molecules. However, the most diverse phenomena can be observed in experimental studies of asymmetric quantum wells and quantum wires [24]. Modern technologies make it possible to grow semiconductor crystals with widely varying properties. In particular, for $GaAs/Al_{0.14}Ga_{0.86}$ crystals with $Al_{0.3}Ga_{0.3}As$ barriers, the absolute value of the ratio of diagonal dipole matrix elements to the corresponding off-diagonal element varies between 0.15 and 7.1, depending on quantum-transition frequency [18].

The optical anisotropy induced by the internal electric fields generated by low-dimensional quantum objects (wells, wires, and dots) manifests itself by birefringence. Propagation of electromagnetic pulses having ordinary and extraordinary components through a resonant anisotropic medium was analyzed in [23]. The operator version of the WKB method [25, 26] proposed in [27] was applied to show that a permanent dipole moment is responsible for a strong nonlinear coupling between pulse components having nearly equal linear velocities. The short-wavelength (ordinary) component induces quantum transitions and effectively gives rise to the long-wavelength (extraordinary) component, which dynamically shifts the transition frequency and chirps the ordinary component. This regime is henceforth referred to as long/short-wave coupling (LSWC). When the extraordinary component is dominant, the regime called extraordinary transparency (EOT) in [23] takes place. In this regime, the two-component pulse decelerates as in the SIT regime and has no appreciable dynamic effect on the populations of quantum states (population trapping).

Analysis of the LSWC equations for two-component pulses propagating in a resonant axially anisotropic medium is continued in this paper. Behavior of solutions to these equations was examined in [23] by assuming that the ordinary component is resonant with quantum transitions. In this case, the only effect of the extraordinary component, which has no well-defined carrier frequency, is to detune the ordinary component from resonance, thus weakening the excitation of the medium. In the more interesting case of an input pulse tuned off resonance, the effect of the extraordinary component varies. The dynamic shift in the ordinary-

wave frequency induced by this component can either detune the pulse from resonance with an anisotropic medium or brings it into resonance. Moreover, owing to the existence of a permanent dipole moment, the extraordinary component can affect the pulse velocity, which leads to transparency regimes in resonant media that substantially differ from SIT. In this paper, we identify and classify these regimes.

The paper is organized as follows. In Section 2, we formulate the system of material and wave equations for electromagnetic pulse propagation perpendicular to the optical axis of an axially symmetric anisotropic medium containing resonant two-level particles with a doubly degenerate level. By applying a unitary transformation, these equations are rewritten as a system of equations for a two-component pulse interacting with a nondegenerate two-level system that describes extends the SIT equations to the case of a permanent dipole moment. The system can be solved by the inverse scattering method if the LSWC condition is exactly satisfied. In Section 3, we develop a general analysis of the one-soliton solution to the LSWC system, which is used as a basis for the discussion of various regimes of two-component pulse propagation in anisotropic media presented in the sections that follow. We modify the soliton solution to describe the case when the ordinary and extraordinary waves have different velocities and the model equations are written without assuming that the concentration of resonant transitions is low. When a two-component pulse propagates through an anisotropic medium, the anisotropy of the medium depends on pulse parameters. In Section 4, we consider pulse propagation through a strongly excited resonant medium, in which case electromagnetic pulses can propagate not only in the SIT regime, but also in the self-induced supertransparency (SIST) regime. These regimes are essentially different in that the pulse velocity is not substantially reduced in the SIST regime. The duration of the pulses that cause the largest change in level populations depends on the amount of detuning. When the components have equal linear-wave velocities, the ordinary-wave carrier frequency is lower than the resonance frequency. Furthermore, the SIST regime is characterized by substantial chirping of the ordinary wave. In Section 5, we analyze the regimes of two-component pulse propagation through a weakly excited medium. Apart from the EOT regime, whose existence under exact resonance conditions was substantiated in [23], we consider the positive and negative nonresonant transparency (PNT and NNT) regimes that take place under conditions of quasi-resonance approximation. In Sections 6 and 7, we discuss manifestations of the identified transparency regimes in high-density media and positively and negatively birefringent media. In the latter case, coupled propagation of a high-frequency ordinary wave and an extraordinary video pulse is analyzed in detail under the Zakharov–Benney resonance condition, i.e., when the group velocity of the high-frequency wave and the phase velocity of the low-frequency waves are equal.

2. BASIC EQUATIONS

Consider two-component electromagnetic pulse propagation through an axially anisotropic medium. For example, suppose that the medium contains quantum wires [24]. An electron moving in a quantum wire is characterized by anisotropic mobility: its motion perpendicular to the wire axis is much more restricted as compared to its axial motion. Since the anisotropy is induced by a strong internal electric field, the energy levels are degenerate in the absolute value of the projection of electron angular momentum.

Consider a laser pulse propagating in the positive direction of the y axis and suppose that the quantum wires are parallel to the z axis. Then, the ordinary and extraordinary electric field components in the pulse, E_o and E_e, are parallel to the x and z axes, respectively. Assume that the ordinary-wave carrier frequency ω is nearly equal to the frequency ω_0 of a σ-transition between the levels, one of which is degenerate in the absolute value of the projection of total angular momentum. In this case, it can readily be shown that the ordinary and extraordinary pulse components play totally different roles: the ordinary wave induces a resonant transition, whereas the extraordinary one induces a dynamic shift in the transition frequency [23]. Under these conditions, the time-dependent quantum states obey the following operator equation in the slowly varying envelope approximation [23]:

$$\frac{\partial \bar{R}}{\partial t} = i[\tilde{\Omega}, \tilde{R}]. \tag{1}$$

Here,

$$\tilde{R} = \begin{pmatrix} \rho_{11} & R_{12} & R_{13} \\ R_{12}^* & \rho_{22} & \rho_{23} \\ R_{13}^* & \rho_{23}^* & \rho_{33} \end{pmatrix},$$

$$\tilde{\Omega} = \begin{pmatrix} \Delta + \frac{D_{11}}{\hbar}E_e & \frac{d_{12}}{\hbar}\mathscr{E}_o & \frac{d_{13}}{\hbar}\mathscr{E}_o \\ \frac{d_{12}}{\hbar}\mathscr{E}_o^* & \frac{D_{22}}{\hbar}E_e & 0 \\ \frac{d_{13}}{\hbar}\mathscr{E}_o^* & 0 & \frac{D_{33}}{\hbar}E_e \end{pmatrix},$$

R_{12} and R_{13} are the slowly varying envelopes in

$$\rho_{12} = R_{12}\exp\left(i\omega\left(t - \frac{n_o y}{c}\right)\right),$$

$$\rho_{13} = R_{13}\exp\left(i\omega\left(t - \frac{n_o y}{c}\right)\right),$$

ρ_{jk} ($j, k = 1, 2, 3$) are the density-matrix elements; $\hbar$ is Planck's constant; $\Delta = \omega_0 - \omega$ is the detuning from resonance ($|\Delta| \ll \omega_0$); D_{11}, D_{22}, and D_{33} are the nonzero elements of the projection of the dipole matrix on the z axis; and $\mathscr{E}_o$ is the ordinary-wave electric field envelope:

$$E_o = \mathscr{E}_o \exp\left(i\omega\left(t - \frac{n_o y}{c}\right)\right) + \text{c.c.}$$

The Maxwell equations yield the following system [23]:

$$\frac{\partial \mathscr{E}_o}{\partial y} + \frac{n_o}{c}\frac{\partial \mathscr{E}_o}{\partial t} = -2i\frac{\pi N\omega}{n_o c}(d_{12}R_{12} + d_{13}R_{13}), \quad (2)$$

$$\frac{\partial^2 E_e}{\partial y^2} - \frac{n_e^2}{c^2}\frac{\partial^2 E_e}{\partial t^2} = \frac{4\pi N}{c^2}\frac{\partial^2}{\partial t^2}(D_{11}\rho_{11} + D_{22}\rho_{22} + D_{33}\rho_{33}), \quad (3)$$

where c is the speed of light in free space; n_o and n_e denote the ordinary and extraordinary refractive indices, respectively; and N is the concentration of resonant σ-transitions.

According to Eq. (3), the extraordinary component

$$P_e = N(D_{11}\rho_{11} + D_{22}\rho_{22} + D_{33}\rho_{33})$$

of the polarization of anisotropic resonant quantum states is determined by the quantum-level population and the diagonal dipole matrix elements. Since the population is changed by the effect of the resonant ordinary component of the pulse, the two-component pulse propagation in the medium under study is modeled self-consistently.

If $2\pi N D_m^2 \ll \hbar\omega_0$ (the concentration of resonant transitions is low), where

$$D_m^2 = \max\{|d_{12}(D_{11} - D_{22})|, |d_{13}(D_{11} - D_{33})|\},$$

then the order of Eq. (3) can be reduced by applying the unidirectional approximation [8]:

$$\frac{\partial E_e}{\partial y} + \frac{n_e}{c}\frac{\partial E_e}{\partial t} = -\frac{2\pi N}{n_e c}\frac{\partial}{\partial t}(D_{11}\rho_{11} + D_{22}\rho_{22} + D_{33}\rho_{33}). \quad (4)$$

Assuming that $D_{33} = D_{22} \neq D_{11}$ [21, 23], we can change to new variables by applying the unitary transformation

$$\tilde{R}' = T\tilde{R}T^{-1}, \quad (5)$$

where

$$\tilde{R}' = \begin{pmatrix} \rho'_{11} & R'_{12} & R_{13} \\ R'^{*}_{12} & \rho'_{22} & \rho'_{23} \\ R'^{*}_{13} & \rho'^{*}_{23} & \rho'_{33} \end{pmatrix},$$

$$T = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\sigma & \sin\sigma \\ 0 & -\sin\sigma & \cos\sigma \end{pmatrix},$$

with $\sigma = \arctan(d_{13}/d_{12})$. The new variables are expressed as

$$\rho'_{11} = \rho_{11}, \quad \rho'_{22} = \cos^2\sigma\rho_{22} + \cos\sigma\sin\sigma(\rho_{23} + \rho^*_{23}) + \sin^2\sigma\rho_{33},$$

$$\rho'_{33} = \sin^2\sigma\rho_{22} - \cos\sigma\sin\sigma(\rho_{23} + \rho^*_{23}) + \cos^2\sigma\rho_{33}, \quad (6)$$

$$R'_{12} = \cos\sigma R_{12} + \sin\sigma R_{13},$$

$$R'_{13} = \cos\sigma R_{13} - \sin\sigma R_{12},$$

$$\rho'_{23} = \cos^2\sigma\rho_{23} + \cos\sigma\sin\sigma(\rho_{33} - \rho_{22}) - \sin^2\sigma\rho^*_{23}.$$

Then, Eqs. (2) and (4) are rewritten as

$$\frac{\partial \mathscr{E}_o}{\partial y} + \frac{n_o}{c}\frac{\partial \mathscr{E}_o}{\partial t} = -2i\frac{\pi N\omega}{n_o c}dR'_{12}, \quad (7)$$

$$\frac{\partial E_e}{\partial y} + \frac{n_e}{c}\frac{\partial E_e}{\partial t} = -\frac{\pi N}{n_e c}D\frac{\partial}{\partial t}(\rho'_{11} - \rho'_{22}), \quad (8)$$

where $d = \sqrt{d_{12}^2 + d_{13}^2}$, and

$$D = D_{11} - D_{22}$$

is a permanent dipole moment. Combining (1) with (5), we obtain

$$\frac{\partial \tilde{R}'}{\partial t} = i[\tilde{\Omega}', \tilde{R}'], \quad (9)$$

where

$$\tilde{\Omega}' \equiv T\tilde{\Omega}T^{-1} = \begin{pmatrix} \Delta + \frac{D_{11}}{\hbar}E_e & \frac{d}{\hbar}\mathscr{E}_o & 0 \\ \frac{d}{\hbar}\mathscr{E}_o^* & \frac{D_{22}}{\hbar}E_e & 0 \\ 0 & 0 & \frac{D_{22}}{\hbar}E_e \end{pmatrix}.$$

Equation (9) can be expanded into equations for matrix

elements:

$$\frac{\partial \rho'_{11}}{\partial t} = i\frac{d}{\hbar}(\mathscr{E}_o R'_{21} - \mathscr{E}_o^* R'_{12}), \quad (10)$$

$$\frac{\partial \rho'_{22}}{\partial t} = i\frac{d}{\hbar}(\mathscr{E}_o^* R'_{12} - \mathscr{E}_o R'_{21}), \quad (11)$$

$$\frac{\partial R'_{12}}{\partial t} = i\left(\Delta + \frac{D}{\hbar}E_e\right)R'_{12} + i\frac{d}{\hbar}\mathscr{E}_o(\rho'_{22} - \rho'_{11}), \quad (12)$$

$$\frac{\partial \rho'_{33}}{\partial t} = 0, \qquad \frac{\partial \rho'_{23}}{\partial t} = i\frac{d}{\hbar}\mathscr{E}_o^* R'_{13}, \quad (13)$$

$$\frac{\partial R'_{13}}{\partial t} = i\left(\Delta + \frac{D}{\hbar}E_e\right)R'_{13} + i\frac{d}{\hbar}\mathscr{E}_o\rho'_{23}. \quad (14)$$

Since the last two equations make up an independent system, we can analyze them separately. In terms of physics, the most interesting case corresponds to the trivial solution to Eqs. (13) and (14). Indeed, suppose that only diagonal elements of $\tilde{R}$ are nonzero and the populations of states with opposite signs of the projection of total angular momentum are equal, as in the case of Boltzmann distribution ($\rho_{22} = \rho_{23}$). Then, it follows from Eqs. (6), (13), and (14) that $\rho'_{23} = R'_{13} = 0$ at any instant, and the original variables are expressed as

$$\rho_{11} = \rho'_{11}, \quad \rho_{22} = \cos^2\sigma\rho'_{22} + \sin^2\sigma\rho'_{33},$$

$$\rho_{33} = \sin^2\sigma\rho'_{22} + \cos^2\sigma\rho'_{33},$$

$$R_{12} = \cos\sigma R'_{12}, \quad R_{13} = \sin\sigma R'_{12},$$

$$\rho_{23} = \cos\sigma\sin\sigma(\rho'_{22} - \rho'_{33}).$$

Thus, Eqs. (7)–(12) provide a comprehensive description of a two-component electromagnetic pulse propagating through a resonant axially symmetric anisotropic medium in the case of the Boltzmann distribution of particles over energy levels. This extension of the SIT equations to the case of a permanent dipole moment was derived in [23] under the additional assumption that $d_{12} = d_{13}$. If the initial energy distribution differs from the Boltzmann distribution, then Eqs. (13) and (14) must be solved under nonzero initial conditions.

Changing to the variables

$$W = \frac{\rho'_{22} - \rho'_{11}}{2}, \quad R = R'_{12}, \qquad \Omega_o = \frac{2d}{\hbar}\mathscr{E}_o, \quad \Omega_e = \frac{D}{\hbar}E_e, \quad (15)$$

we rewrite Eqs. (7)–(12) as

$$\frac{\partial W}{\partial t} = \frac{i}{2}(\Omega_o^* R - \Omega_o R^*), \quad (16)$$

$$\frac{\partial R}{\partial t} = i(\Delta + \Omega_e)R + i\Omega_o W, \quad (17)$$

$$\frac{\partial \Omega_o}{\partial y} + \frac{n_o}{c}\frac{\partial \Omega_o}{\partial t} = -i\beta_o R, \quad (18)$$

$$\frac{\partial \Omega_e}{\partial y} + \frac{n_e}{c}\frac{\partial \Omega_e}{\partial t} = \beta_e\frac{\partial W}{\partial t}, \quad (19)$$

where

$$\beta_o = 4\frac{\pi N\omega}{\hbar c n_o}d^2, \quad \beta_e = 2\frac{\pi N}{\hbar c n_e}D^2.$$

Note that the Rabi frequency Ω_o associated with the ordinary component is the field envelope, whereas Ω_e is proportional to the electric field component E_e, which has no well-defined carrier frequency.

In [23], an operator version of the WKB was used to show that the pulse components are strongly coupled in the case of exact resonance ($\Delta = 0$) if their linear velocities are nearly equal. When the LSWC condition is satisfied exactly, i.e.,

$$n_e = n_o. \quad (20)$$

Equations (18) and (19) can be integrated to obtain the following expression for the "Rabi frequency" of the extraordinary wave:

$$\Omega_e = -\tilde{D}\frac{|\Omega_o|^2}{4d^2\omega}, \quad (21)$$

where $\tilde{D} = D^2$. It should be noted here that this nonlinear system, as well as its gauge-invariant equivalents, arise in various problems in nonlinear optics and belong to the class of systems that can be integrated by the inverse scattering method (see [28] and references therein). Studies of the nonlinear evolution of soliton solutions to this system have mainly been focused on the characteristics of electromagnetic pulses, whereas the behavior of resonant media has remained unclear. In this paper, we eliminate this shortcoming.

If the concentration of resonant transitions is not low, then the unidirectional approximation cannot be applied and we perform a gauge transformation and change to the variables defined by (15) to obtain the following equation instead of (19):

$$\frac{\partial^2 \Omega_e}{\partial y^2} - \frac{n_e^2}{c^2}\frac{\partial^2 \Omega_e}{\partial t^2} = -2\beta_e\frac{n_e}{c}\frac{\partial^2 W}{\partial t^2}. \quad (22)$$

Equations differing from (16)–(18) and (22) only by notation were also obtained in [27] for longitudinal–transverse acoustic pulse propagation parallel to an

external magnetic field in a system of resonant paramagnetic impurities with effective spin $S = 1/2$.

3. GENERAL ANALYSIS OF THE SOLITON SOLUTION

As noted above, system (16)–(19) subject to condition (20) belongs to the class of systems that can be solved by the inverse scattering method. Its is well known that particular solutions to such equations can be found by algebraic methods [4–6]. These methods can be used to obtain soliton solutions to integrable equations, which are of primary physical interest in many cases.

It is convenient to express the ordinary component of the one-soliton solution to (16)–(19) in exponential form:

$$\Omega_o = |\Omega_o|\exp(i\Phi), \tag{23}$$

where

$$|\Omega_o| = \frac{2}{\tau_p}\left|\frac{2g}{g-\alpha+\operatorname{sgn} g\sqrt{1+(g-\alpha)^2}\cosh(2\zeta)}\right|^{1/2}, \tag{24}$$

$$\Phi = W_\infty\frac{\beta_o\alpha\tau_p}{1+\alpha^2}y - \arctan\frac{\tanh\zeta}{s} + \text{const}, \tag{25}$$

$$g = \frac{2\omega_0\tau_p d^2}{\tilde{D}}, \quad \zeta = \frac{1}{\tau_p}\left(t-\frac{y}{v_g}\right),$$

$$s = d-\alpha+\operatorname{sgn} g\sqrt{1+(g-\alpha)^2},$$

$\alpha = \tau_p\Delta$, and W_∞ is the initial population of the medium. The phase and group velocities of the pulse in the laboratory frame are

$$v_{ph} = \frac{c}{n_o}\left(1+2A\frac{\alpha\omega_0\tau_p}{1+\alpha^2}\right)^{-1},$$

$$v_g = \frac{c}{n_o}\left(1+2A\frac{\omega_0^2\tau_p^2}{1+\alpha^2}\right)^{-1}, \tag{26}$$

where

$$A = -\frac{c\beta_o W_\infty}{2n_o\omega_0^2}.$$

Substituting (24) into (21), we obtain

$$\Omega_e = -4\{\tau_p(g-\alpha + \operatorname{sgn} g\sqrt{1+(g-\alpha)^2}\cosh(2\zeta))\}^{-1}. \tag{27}$$

The change in the population of the medium is described by the expression

$$W = W_\infty\left(1-\frac{\tau_p^2|\Omega_o|^2}{2(1+\alpha^2)}\right). \tag{28}$$

Expressions (23) and (25) show that the ordinary-component phase is modulated, which leads to a local nonlinear chirping of the carrier frequency, $\omega \longrightarrow \omega_{loc} = \omega + \delta\omega_{non}$, where

$$\delta\omega_{non} = \frac{\partial\Phi}{\partial t} = -\frac{s}{\tau_p}\frac{\operatorname{sech}^2\zeta}{s^2+\tanh^2\zeta}. \tag{29}$$

In an isotropic resonant medium ($D = 0$, $g \longrightarrow \infty$), phase modulation is absent and Eqs. (24), (26), and (28) reduce to the respective expressions for a SIT 2π-pulse propagation.

Comparing (29) with (24), we find that the nonlinear shift in the local carrier frequency decreases toward the pulse edges faster than the ordinary component. Thus, Δ can be interpreted as the detuning of the ordinary-wave carrier frequency from resonance at the pulse edges.

According to Eq. (17), the ordinary wave dynamically shifts the quantum-transition frequency: $\omega_0 \longrightarrow \omega_0^{ef} = \omega_0 + \Omega_e$. Therefore, it is convenient to introduce an effective ordinary-wave detuning:

$$\Delta_{ef} \equiv \omega_0^{ef} - \omega_{loc} = \Delta + \Omega_e - \delta\omega_{non}. \tag{30}$$

If the pulse width T_p is defined as twice the distance from the point of zero $t - y/v_g$ at which $|\Omega_o|$ is half its maximum value, then expression (24) yields

$$T_p = \tau_p\operatorname{arccosh}\left(4+3\operatorname{sgn} g\frac{g-\alpha}{\sqrt{1+(g-\alpha)^2}}\right). \tag{31}$$

In the slowly varying envelope approximation, both the pulse length and nonlinear shift of the carrier frequency must obviously satisfy the conditions $\omega_0 T_p \gg 1$ and $|\delta\omega_{non}| \ll \omega_0$. It can easily be shown that these inequalities are valid if $\omega_0\tau_p \gg 1$. Moreover, for g and α such that $|g| < |\alpha|$, $g\alpha > 0$, and $|g-\alpha| \gg 1$, the condition $\omega_0\tau_p \gg |g-\alpha|$ must also be satisfied, which is the case only when $D^2 \gg 4d^2$. This condition can be met within the scope of modern semiconductor crystal growth technology. It is clear from the inequality $\omega_0\tau_p \gg 1$ that the ordinary-wave phase velocity is much closer to the linear velocity, as compared to the group velocity.

The expressions presented above can be modified to describe the case when condition (20) is not satisfied and/or the unidirectional approximation is not employed. Indeed, if the extraordinary component is sought as a traveling-wave solution $\Omega_e = \Omega_e(t - y/v_g)$, then Eqs. (22), (16)–(18), and (23)–(26) can be used to obtain a relation between Ω_e and Ω_o similar to (21) with

$$\tilde{D} = \frac{4A_e d^2}{\left(1-A_o\frac{1+\alpha^2}{\omega_0^2\tau_p^2}\right)\left(1+\tilde{A}\frac{\omega_0^2\tau_p^2}{1+\alpha^2}\right)}, \tag{32}$$

where

$$A_e = \frac{n_o D^2}{2(n_o + n_e)d^2}, \quad A_o = \frac{n_e - n_o}{2n_o A}, \quad \tilde{A} = \frac{2n_o A}{n_o + n_e}.$$

It is obvious that the modified expressions (23)–(28) provide a solution to system (16)–(18), (22). By virtue of (26), expression (32) can be rewritten as

$$\tilde{D} = \frac{2n_o v_g (n_o v_g - c)}{n_e^2 v_g^2 - c^2} D^2.$$

If

$$\tilde{D} = \frac{n_e}{n_o} \frac{D^2}{1 - A_o \dfrac{1 + \alpha^2}{\omega_0^2 \tau_p^2}}, \tag{33}$$

then (23)–(28) define a soliton solution to (16)–(19) in the case of low concentration of quantum transitions.

Note that (33) reduces to (32) as $A \longrightarrow 0$ only if $n_e = n_o$. The parameter $|\tilde{D}|$, which does not vanish only for a medium having a permanent dipole moment, is called here the effective anisotropy. Thus, we emphasize its dependence not only on the properties of the medium, but also on the pulse parameters. Comparing (32) with (33), we see that, for a medium with equilibrium initial population, the effect of the extraordinary wave on pulse formation (due to anisotropy) must increase with decreasing concentration of resonant transitions.

In what follows, we use expressions (24) and (26)–(29) to identify several regimes of two-component pulse propagation through an anisotropic medium. Unless stated otherwise, it is assumed that the particles are in thermodynamic equilibrium prior to the pulse propagation, i.e., the initial population W_∞ is negative ($-1/2 \le W_\infty < 0$), and both A and $\tilde{A}$ are positive. In the next two sections, we consider the case of $\tilde{D} = D^2$, which corresponds to a low-density medium with $n_e = n_o$. A more general analysis is presented in Sections 6 and 7.

4. SOLITONS IN A STRONGLY EXCITED MEDIUM

Strong excitation is interpreted here as the largest possible change in the population of quantum levels induced by a two-component pulse. By virtue of (24) and (28), this condition can be formulated as

$$\tau_p^2 |\Omega_o(\zeta = 0)|^2 = 4(1 + \alpha^2).$$

Performing some simple algebra, we obtain

$$(\alpha^2 - 2g\alpha + 1)^2 = 0,$$

which yields

$$g = \frac{1}{2}\left(\alpha + \frac{1}{\alpha}\right). \tag{34}$$

This condition can also be rewritten as a relation between the detuning Δ and the parameter τ_p:

$$\frac{1}{\tau_p} = \sqrt{4\Delta\omega_0 \frac{d^2}{D^2} - \Delta^2}.$$

Setting the right-hand side to zero, we find the interval of admissible values of the detuning: $0 < \Delta < \Delta_m$, where $\Delta_m = 4\omega_0 d^2/D^2$.

For a finite detuning, τ_p vanishes in the isotropic limit ($D \longrightarrow 0$); i.e., the field amplitude tends to infinity. This agrees with the well-known fact that the largest change in population in an isotropic medium corresponds to exact resonance ($\Delta = 0$). Strong excitation can then be attained only for a positive detuning ($\Delta > 0$); i.e., the carrier frequency must be lower than the resonant transition frequency ($\omega < \omega_0$). The positive detuning from resonance required for a pulse to cause complete excitation in a medium having a permanent dipole moment is consistent with the fact that the extraordinary electric-field component generated in such a medium induces a red shift of the effective transition frequency ω_0^{ef}, since $\Omega_e < 0$ according to (21).

Using condition (34), we rewrite expressions (24) and (26)–(29) as follows:

$$\begin{aligned}
|\Omega_o| &= \frac{\Omega_{om}}{\sqrt{1 + (1 + \alpha^2)\sinh^2\zeta}}, \\
\Omega_e &= -\operatorname{sgn}\alpha \frac{\Omega_{em}}{1 + (1 + \alpha^2)\sinh^2\zeta}, \\
v_{ph} &= \frac{c}{n_o}\left(1 + \frac{8A\alpha\omega_0}{\Omega_{om}^2 \tau_p}\right)^{-1}, \\
v_g &= \frac{c}{n_o}\left(1 + \frac{8A\omega_0^2}{\Omega_{om}^2}\right)^{-1}, \\
W &= W_\infty\left(1 - \frac{2}{1 + (1 + \alpha^2)\sinh^2\zeta}\right), \\
\delta\omega_{non} &= -\operatorname{sgn}\alpha\, \delta\omega_{non}^m \frac{\operatorname{sech}^2\zeta}{1 + \alpha^2 \tanh^2\zeta},
\end{aligned} \tag{35}$$

where the amplitudes are

$$\Omega_{om} = \frac{2}{\tau_p}\sqrt{1+\alpha^2},$$
$$\Omega_{em} = \frac{4|\alpha|}{\tau_p}, \quad (36)$$
$$\delta\omega_{non}^{m} = \frac{|\alpha|}{\tau_p}.$$

The corresponding pulse width is

$$T_p = 2\tau_p \operatorname{arcsinh}\sqrt{\frac{3}{1+\alpha^2}}. \quad (37)$$

According to Eqs. (35) and (36), the effective dynamic detuning of the ordinary component from resonance with the medium, Δ_{ef}, equals Δ at the edges of the pulse and -2Δ at its center. Thus, owing to the effect of the extraordinary component, the ordinary pulse component is resonant with quantum transitions on average over the pulse duration. As a result, the largest possible change in population is achieved.

When the detuning is very small ($\alpha \ll 1$, $g \gg 1$), the expressions in (35) can be represented as

$$|\Omega_o| = \frac{2}{\tau_p}\operatorname{sech}\zeta, \quad \Omega_e = -\frac{4\alpha}{\tau_p}\operatorname{sech}^2\zeta,$$
$$v_g = \frac{c}{n_o}(1+2A\omega_0^2\tau_p^2)^{-1}, \quad (38)$$
$$W = W_\infty(1-2\operatorname{sech}^2\zeta),$$
$$\delta\omega_{non} = -\frac{\alpha}{\tau_p}\operatorname{sech}^2\zeta.$$

Here, $|\Omega_e| \ll |\Omega_o|$ and the phase-modulation depth for the ordinary component is much smaller than its spectral width in the input pulse: $|\delta\omega_{non}| \ll 1/\tau_p$. Under the exact resonance conditions ($\alpha = 0$, $g \longrightarrow \infty$), expressions (38) reduce to well-known expressions of the SIT theory for isotropic media ($D = 0$), where neither extraordinary-wave video pulse nor ordinary-wave phase modulation are observed ($\Omega_e = \delta\omega_{non} = 0$).

According to (36), an increase in detuning (and therefore α) leads to larger amplitudes of both ordinary and extraordinary components and a deeper chirping of the ordinary component toward lower frequencies. Since the pulse length decreases, the profiles of both components become sharper.

It is also important that group velocity approaches the linear-wave velocity c/n_o as detuning increases. The medium remains strongly excited: the largest possible change in population is reached at the center of the soliton. This is obviously explained by the fact that an increase in pulse amplitude combined with a decrease in pulse width increases its power. The ensuing higher rate of excitation/de-excitation processes leads to a higher soliton propagation velocity.

It is clear from (34) that each particular value of g corresponds to two distinct values of α. Regimes with relatively small detuning ($\alpha < 1$) can be associated with the SIT regime, which is implemented when $\alpha = 0$. When $\alpha > 1$, we say that the solutions given by (35) describe the regime of self-induced supertransparency, thus emphasizing the fact that the propagation velocity is higher as compared to the SIT regime while excitation is equally strong. Since $\Delta \ll \omega_0$ in the slowly varying envelope approximation, the SIST regime must be most strongly manifested in media with $D^2 \gg 4d^2$.

When $g \gg 1$ and $\alpha \gg 1$, both ordinary and extraordinary wave amplitudes and the phase-modulation depth for the ordinary component of the SIST pulse reach their limits $2\Delta_m$, $4\Delta_m$, and Δ_m, respectively. In this case, (37) yields the following estimate for the pulse duration:

$$T_p \approx \frac{2\sqrt{3}\tau_p}{\alpha} \approx \frac{D^2}{\omega_0 d^2}.$$

This time scale corresponds to the time scale of phase-modulation localization. When the condition $\omega_0 T_p \gg 1$ is not satisfied, the slowly varying envelope approximation cannot be employed. However, a tendency toward distortion of SIST solitons in an anisotropic medium is observed as the detuning increases. A deeper phase modulation combined with a decrease in the corresponding localization time scale and a shorter pulse width can be interpreted as an effect due to generation of a supercontinuum. Indeed, the Fourier transform of solution (23)–(25), (27) defined as

$$F_{o,e}(\nu) = \int_{-\infty}^{\infty} e^{i\nu t}\Omega_{o,e}\,dt,$$

yields

$$|F_o(\nu)| = 2\pi\frac{\sqrt{|g|}}{\sqrt[4]{1+(g-\alpha)^2}}\frac{\exp(\theta\tau_p\nu/2)}{\cosh(\pi\tau_p\nu/2)},$$
$$|F_e(\nu)| = 4\pi\frac{\sinh(\theta\tau_p\nu/2)}{\sinh(\pi\tau_p\nu/2)},$$

where $\theta = \operatorname{arccoth}(g-\alpha)$. (It is assumed here that $0 < \operatorname{sgn} g\theta < \pi$.) The absolute values of the Fourier transforms of the extraordinary and ordinary components reach maximum values at $\nu = 0$ and $\nu = \nu_0$, respectively, where

$$\nu_0 = \frac{2}{\pi\tau_p}\ln\frac{\pi+\theta}{\pi-\theta}.$$

When $g - \alpha \gg 1$ and $\theta \longrightarrow \pi$, the expression for $|F_o(\nu)|$ shows that the spectral pulse width is $\delta\omega \sim 1/(\pi-\theta)\tau_p$ and maximum of the spectral energy distribution is

reached at the frequency $\omega - \nu_0$, which is much lower than the carrier frequency ω. In this case, despite a large linear detuning from resonance ($\tau_p\Delta \gg 1$), a substantial nonlinear spectral broadening ($\tau_p\delta\omega \gg 1$) leads to generation of resonant Fourier components (photons), which stimulate quantum transitions.

The present analysis shows that an ordinary component initially tuned off resonance can be brought into resonance via generation of an extraordinary component. Therefore, strong excitation by a pulse with a nonzero initial detuning can occur only in a medium having a permanent dipole moment and never takes place in isotropic media.

We should note that Eqs. (16)–(19) differ from the systems considered in [20–22], where one-component pulse propagation was analyzed. The evolution of both field and medium was described by the system of reduced Maxwell–Bloch equations extended to the case of a permanent dipole moment:

$$\begin{aligned}\frac{\partial U}{\partial t} &= -(\omega_0 + \mu\tilde{\Omega})U,\\ \frac{\partial V}{\partial t} &= (\omega_0 + \mu\tilde{\Omega})V + \tilde{\Omega}W,\\ \frac{\partial W}{\partial t} &= -\tilde{\Omega}V,\\ \frac{\partial\tilde{\Omega}}{\partial y} + \frac{n}{c}\frac{\tilde{\Omega}}{\partial t} &= \beta V,\end{aligned} \tag{39}$$

where U, V, and W are the Bloch vector components; $\tilde{\Omega}$ is a time-dependent parameter proportional to the wave field; μ and β are parameters proportional to the permanent dipole transition moment and the concentration of quantum objects; and n is refractive index.

If we set $U + iV \propto \exp(i\omega t)$, where ω is the input carrier frequency, then the terms proportional to $\tilde{\Omega}U$ and $\tilde{\Omega}V$ in the material equations in (39) describe second harmonic generation when $\omega \approx \omega_0$. As applied to the present model of axially symmetric quantum objects, system (39) describes a regime in which the pulse has only an extraordinary component ($\tilde{\Omega} = \Omega_e$), which stimulates a resonant π-transition and simultaneously shifts the transition frequency. This regime can be implemented in the system of anisotropic quantum objects considered here. However, in contrast to the present model, the input pulse must be polarized in the principal plane of the anisotropic medium (i.e., in the extraordinary-wave polarization plane), and its carrier frequency must be close to the frequency of a nondegenerate π-transition. In the case considered here, the strict distinction between the respective effects of the ordinary and extraordinary waves indicated above leads to relation (21) between them, which is equivalent to the replacement of $\Delta + \Omega_e$ with $\Delta - (\tilde{D}/4\omega_0 d^2)|\Omega_o|^2$ in Eq. (17). Since the dynamic coefficient $|\Omega_o|^2$ has no carrier frequency, no higher harmonics are generated. The high-frequency input ordinary-wave pulse generates an extraordinary-wave (zeroth-harmonic) video pulse, which leads to phase modulation of the ordinary component and, as a consequence, to its spectral broadening.

5. TRANSPARENCY REGIMES IN A WEAKLY EXCITED MEDIUM

Let us now consider electromagnetic pulse propagation in anisotropic media in the case when $g \ll 1$. Since $\omega_0\tau_p \gg 1$ in the slowly varying envelope approximation, the value of $|D/d|$ must sufficiently large to ensure that $\omega\tau_p(d/D)^2 \ll 1$. If $\omega_0\tau_p \sim 10^2$, then $|D/d| \geq 20$. To date, GaAs/$Al_{0.14}Ga_{0.86}$ semiconductor crystals with $Al_{0.3}Ga_{0.3}As$ barriers are available, with $|D/d| \approx 7$ [18]. This is not sufficient to ensure that both $g \ll 1$ and the slowly varying envelope approximation is applicable, but $|D/d|$ is expected to increase with further progress in semiconductor crystal growth technology.

First, we consider the case of large detuning from resonance ($|\alpha| \gg 1$). If $\alpha < 0$ (or $\omega > \omega_0$), then (24) and (26)–(29) reduce to

$$\begin{aligned}|\Omega_o| &= \frac{2}{\tau_p}\sqrt{\frac{g}{|\alpha|}}\operatorname{sech}\zeta, \quad \Omega_e = -\frac{2}{|\alpha|\tau_p}\operatorname{sech}^2\zeta,\\ v_{ph} &= \frac{c}{n_o}\left(1 - \frac{2A\omega_0\tau_p}{|\alpha|}\right)^{-1},\\ v_g &= \frac{c}{n_o}\left(1 + \frac{2A\omega_0^2\tau_p^2}{\alpha^2}\right)^{-1},\\ W &= W_\infty\left(1 - \frac{2g}{|\alpha|^3}\operatorname{sech}^2\zeta\right),\\ \delta\omega_{non} &= -\frac{1}{2|\alpha|\tau_p}\operatorname{sech}^2\zeta.\end{aligned} \tag{40}$$

Comparing these expressions with (38), we see that the ordinary-wave amplitude is much smaller as compared to that in the SIT regime in an anisotropic medium, whereas the respective extraordinary components are comparable. We also note that

$$\frac{\Omega_o(\zeta = 0)}{\Omega_e(\zeta = 0)} = \sqrt{g|\alpha|};$$

i.e., the amplitude ratio can have an arbitrary value. The medium remains almost unexcited as the soliton described by (40) propagates through it, and the soliton velocity decreases only very slightly. The phase-modulation depth is also small ($|\delta\omega_{non}| \ll 1/\tau_p$), and the effective detuning Δ_{ef} only increases as the extraordinary wave is generated, which leads to an even weaker exci-

tation of the medium as compared to that induced by the input pulse.

If $g \ll 1$ and $\alpha \gg 1$, then

$$|\Omega_o| = \frac{\Omega_{om}}{\sqrt{1 + 4\alpha^2 \sinh^2 \zeta}}, \quad \Omega_e = -\frac{\Omega_{em}}{1 + 4\alpha^2 \sinh^2 \zeta},$$
$$W = W_\infty \left(1 - \frac{8g}{\alpha(1 + 4\alpha^2 \sinh^2 \zeta)}\right), \tag{41}$$
$$\delta\omega_{non} = -\delta\omega_{non}^m \frac{\operatorname{sech}^2 \zeta}{1 + 4\alpha^2 \tanh^2 \zeta},$$

where

$$\Omega_{om} = \frac{4}{\tau_p}\sqrt{g\alpha}, \quad \Omega_{em} = \frac{8\alpha}{\tau_p}, \quad \delta\omega_{non}^m = \frac{2\alpha}{\tau_p}. \tag{42}$$

The expressions for the phase and group velocities are obtained from those in (40) by replacing $|\alpha|$ with $-\alpha$.

Note that expressions (35), (36) and (41), (42) are somewhat similar: in both case, the solitons are sharply peaked, and their propagation velocities are nearly equal to c/n_0. However, these regimes are essentially different in terms of behavior of the medium. Whereas the medium is strongly excited as the soliton described by (35) propagates through it, no significant excitation is caused in the case of (41). Indeed, since the effective detuning Δ_{ef} of the soliton described by (41) is -5Δ at its center, it is not resonant with the medium. However, according to (40) and (41), the excitation of the medium at $\omega < \omega_0$, being relatively weak, is still stronger than that at $\omega > \omega_0$. The reason is that the effective detuning decreases toward the pulse edges when $\omega < \omega_0$, owing to the extraordinary component, and increases when $\omega > \omega_0$. Thus, a comparison of (40) with (41), (42) demonstrates obvious asymmetry with respect to ordinary-wave detuning. Since $\alpha < 0$ for the solitons described by (40) and $\alpha > 0$ for the solitons described by (41) and (42), we call the corresponding regimes negative and negative nonresonant transparency, respectively.

The inequality $|\alpha| \gg 1$ combined with the applicability conditions for the two-level and rotating-wave approximations ($|\Delta| \ll \omega, \omega_0$) completely characterize the conditions of quasi-resonant interaction between the pulse and the medium [29, 30]. Following [30], we eliminate the material variables R and W from (16)–(19) by using expansions in the small parameter $|\alpha|^{-1}$. We also assume that $|\Omega_o/\Delta| \ll 1$ and $(\Omega_e/\Delta) \ll 1$. First, we rewrite Eq. (17) as

$$R = -\frac{\Omega_o W}{\Delta + \Omega_e} - \frac{i}{\Delta + \Omega_e}\frac{\partial R}{\partial t}. \tag{43}$$

It is obvious that the ratio of the second term on the right-hand side to the left-hand side (and therefore, to the first term on the right-hand side) is α^{-1}. Therefore,

$$R = -\frac{\Omega_o W}{\Delta + \Omega_e}$$

in the zeroth approximation. Substituting this expression into the right-hand side of (43), we obtain the first approximation. Repeating these operations two more times, we find that

$$R = -\frac{\Omega_o W}{\Delta + \Omega_e} - \frac{i}{\Delta + \Omega_e}\frac{\partial}{\partial t}\left(\frac{\Omega_o W}{\Delta + \Omega_e}\right) + \frac{W_\infty}{\Delta^3}\frac{\partial^2 \Omega_o}{\partial t^2} - i\frac{W_\infty}{\Delta^4}\frac{\partial^3 \Omega_o}{\partial t^3} \tag{44}$$

in the third approximation. (Here, we set $\Omega_e = 0$ and $W = W_\infty$ in the last two terms, which are responsible for dispersion, in view of the conditions imposed above.) By virtue of the inequality $\Omega_e/\Delta \ll 1$, we expand the first two terms, which are responsible for zero-delay linear and nonlinear response and linear and nonlinear dispersion of polarization response, in powers of Ω_e/Δ. Substituting (44) with $W \approx W_\infty$ into Eq. (16) and integrating the result, we obtain

$$W = W_\infty \left[1 - \frac{|\Omega_o|^2}{2(\Delta + \Omega_e)^2} + \frac{i}{2\Delta^3}\left(\Omega_o^* \frac{\partial \Omega_o}{\partial t} - \Omega_o \frac{\partial \Omega_o^*}{\partial t}\right)\right]. \tag{45}$$

Here, we dropped the last term in (44), because it represents second-order dispersion.

Substituting (45) into (44); retaining zero-delay linear and nonlinear response, first-order nonlinear dispersion, and linear dispersion to the third order; and restricting the analysis to cubic nonlinearity, we have

$$R = \frac{W_\infty}{\Delta}\left[\Omega_o - \frac{i}{\Delta}\frac{\partial \Omega_o}{\partial t} - \frac{i}{\Delta^2}\frac{\partial^2 \Omega_o}{\partial t^2} + \frac{i}{\Delta^3}\frac{\partial^3 \Omega_o}{\partial t^3} - \frac{\Omega_e \Omega_o}{\Delta} - \frac{|\Omega_o|^2 \Omega_o}{2\Delta^2} + \frac{i}{\Delta^2}\left(2\Omega_e \frac{\partial \Omega_o}{\partial t} + \Omega_o \frac{\partial \Omega_e}{\partial t}\right) + \frac{3i}{2\Delta^3}|\Omega_o|^2 \frac{\partial \Omega_o}{\partial t}\right]. \tag{46}$$

Now, we can substitute (45) and (46) into the right-hand sides of Eqs. (18) and (19) to obtain a set of two nonlin-

ear wave equations. Recalling (21) and performing some simple algebra, we obtain the equation

$$i\frac{\partial\Psi}{\partial y} = q_2\frac{\partial^2\Psi}{\partial\tau^2} + iq_3\frac{\partial^3\Psi}{\partial\tau^3} - ib_1|\Psi|^2\frac{\partial\Psi}{\partial\tau} - ib_2\Psi\frac{\partial}{\partial\tau}|\Psi|^2 \quad (47)$$

with $\tau = t - y/v_g$, where v_g is the group velocity defined by the relation

$$\frac{1}{v_g} = \frac{n_o}{c} - \frac{\beta_o W_\infty}{\Delta^2}\left(1 - 2\frac{\gamma}{\Delta} - 3\frac{\gamma^2}{\Delta^2}\right),$$

$$\gamma = -\frac{\Delta(\alpha - g)}{2\alpha - 3g}, \quad q_2 = \frac{\beta_o W_\infty}{\Delta^2}\frac{1}{3\gamma + \Delta},$$

$$q_3 = -\frac{\beta_o W_\infty}{\Delta^4}, \quad b_1 = -\frac{\beta_o W_\infty}{\Delta^4}\left(\frac{\alpha}{\gamma} - \frac{3}{2}\right)^{-1},$$

$$b_2 = -\frac{\beta_o W_\infty \alpha}{2g\Delta^4}.$$

The terms proportional to Ω_o and $|\Omega_o|^2\Omega_o$ are eliminated by the phase transformation

$$\Omega_o = \Psi\exp[i(\gamma t - \varepsilon y)], \quad (48)$$

where

$$\varepsilon = \frac{n_o}{c}\gamma - \frac{\beta_o W_\infty}{\Delta}\left(1 + \frac{\gamma}{\Delta} + \frac{\gamma^2}{\Delta^2} + \frac{\gamma^3}{\Delta^3}\right).$$

We consider two extreme cases: $g/\alpha \gg 1$ and $g/\alpha \ll 1$. In the former limit, we have $\gamma = -\Delta/3$, $q_2 = b_2 = 0$, and $b_1 = -3q_1/2$ and (47) is rewritten as the modified complex Korteweg–de Vries equation

$$\frac{\partial\Psi}{\partial y} = \frac{3}{2}q_3|\Psi|^2\frac{\partial\Psi}{\partial\tau} + q_3\frac{\partial^3\Psi}{\partial\tau^3}. \quad (49)$$

Note that the parameter q_3 does not change sign under the change $\Delta \longrightarrow -\Delta$; i.e., the quasi-resonant soliton regime in an isotropic medium is independent of the detuning sign. This is obvious since $\Omega_e = 0$.

In the other limit,

$$\gamma = -\frac{\Delta}{2}, \quad q_2 = -\frac{\beta_o W_\infty}{2\Delta^3}, \quad b_1 = 2b_2 = -\frac{\beta_o W_\infty \alpha}{g\Delta^4}.$$

Dropping the third-order dispersion terms, we obtain

$$i\frac{\partial\Psi}{\partial y} = q_2\frac{\partial^2\Psi}{\partial\tau^2} + 2ib_1|\Psi|^2\frac{\partial\Psi}{\partial\tau} + ib_2\Psi\frac{\partial}{\partial\tau}|\Psi|^2. \quad (50)$$

It can readily be shown by using (50), (48), and (45) that the soliton solution to Eq. (50) corresponds to the expressions (40) describing the NNT regime.

The PNT regime cannot be described by applying this approach, because the value of $\Delta + \Omega_e$ passes through zero and changes sign as the pulse propagates through the medium, owing to the extraordinary component, which makes expansion in terms of $((\Delta + \Omega_e)\tau_p)^{-1}$ inapplicable.

If $g \ll 1$ and $|\alpha| \ll 1$ (detuning is small), then (24) and (26)–(29) lead to expressions identical to those found in [23]:

$$|\Omega_o| = \Omega_{om}\,\mathrm{sech}^{1/2}(2\zeta), \quad \Omega_e = -\Omega_{em}\,\mathrm{sech}(2\zeta),$$
$$W = W_\infty(1 - 4g\,\mathrm{sech}(2\zeta)), \quad (51)$$
$$\delta\omega_{non} = -\delta\omega_{non}^m\,\mathrm{sech}(2\zeta),$$

where

$$\Omega_{om} = \frac{2}{\tau_p}\sqrt{2g}, \quad \Omega_{em} = \frac{4}{\tau_p}, \quad \delta\omega_{non}^m = \frac{1}{\tau_p}.$$

The phase velocity v_{ph} is close to the linear-wave velocity, and the expression for group velocity is identical to that in (38) corresponding to the SIT regime.

In this case, the extraordinary component exceeds the ordinary one: $(\Omega_{em}/\Omega_{om})^2 \gg 1$. The medium is not excited, since the effective detuning is large, $3/\tau_p$. However, the propagation velocity decreases as in the case of strong excitation at $\alpha \sim 1$. Pulse deceleration in the EOT regime is explained by the dispersion of the medium within the ordinary-wave bandwidth. This effect was analyzed in detail in [23].

Thus, the transparency regimes corresponding to a high degree of anisotropy ($g \ll 1$) can be not only quantitatively, but also qualitatively, different. As shown above, this difference is explained by the detuning of the ordinary wave from resonance with quantum objects in anisotropic medium. One common feature of the three regimes is virtual absence of any excitation of quantum levels.

6. THE CASE OF HIGH-DENSITY MEDIUM

Now, suppose that the parameter $\tilde{D}$ in (23)–(28) is defined by (32). This means that the refractive indices are not equal and the quantum-transition density is not small, yet expressions (23)–(28) provide solutions to Eqs. (16)–(18), (22).

In this section, we consider the case when the condition $\tilde{A}\omega_0^2\tau_p^2 \ll 1$ is not satisfied. Since the refractive indices of a typical anisotropic medium are such that $|n_e - n_o| \ll 1$, we can assume that the left multiplicand in the denominator on the right-hand side of (32) is close to unity. Then, (32) can be rewritten as

$$\tilde{D} = 4A_e d^2\left(1 + \tilde{A}\frac{\omega_0^2\tau_p^2}{1 + \alpha^2}\right)^{-1}, \quad (52)$$

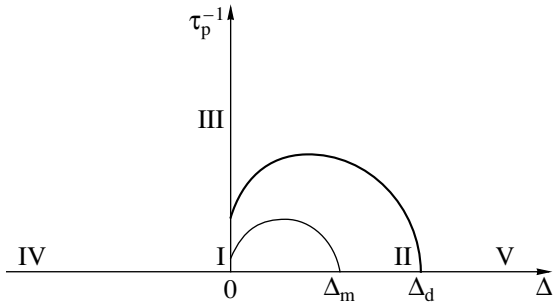

Fig. 1. Curves of strong excitation for high-density medium. Parameter domains for $\Delta_d \ll \omega_0$: (I) SIT; (II) SIST; (III) EOT; (IV) NNT; (V) PNT. Thin curve: $\tilde{A} = 0$.

i.e., the parameter $\tilde{D}$ is independent of the optical anisotropy A_o for a high-density medium.

It can be shown that the transparency regimes identified in the preceding sections exist in this case as well. If $n_e < n_o$ ($A_o < 0$), then $\tilde{D}$ can exceed D^2 in the general case. It follows from (52) that $\tilde{D} > D^2$ if

$$\frac{1}{\tau_p^2} + \Delta^2 > \frac{\omega_0^2}{|A_o|}.$$

Under this condition, (32) cannot be replaced by (52). Therefore, the effective anisotropy cannot exceed D^2 for a high-density medium, and the strong-excitation regimes with $g > 1$ must be more pronounced.

Using (34) and (52), we find that the pulse parameters τ_p and Δ are related as follows in the case of strong excitation:

$$\frac{1}{\tau_p} = \sqrt{F_{1,2}}, \tag{53}$$

where

$$F_{1,2} = -\Delta^2 + (\Delta \pm \sqrt{\Delta(\Delta + 4\tilde{A}A_e\omega_0)})\frac{\omega_0}{2A_e}. \tag{54}$$

It is obvious that

$$F_1F_2 = (A_e\Delta^3 - \omega_0\Delta^2 - \tilde{A}\omega_0^3)\frac{\Delta}{A_e}. \tag{55}$$

It follows from (53) that the detuning of the pulses that cause the largest change in quantum-level populations can only be such that at least one of the quantities F_1 and F_2 is positive. It is easy to see that this is possible only if $\Delta > 0$. Since the coefficients of Δ^3 and Δ on the right-hand side of (55) are positive and the coefficient of Δ^2 is zero, the product F_1F_2 changes sign only once at $\Delta > 0$. Therefore, F_2 is negative in this domain, and F_1 is positive if $0 < \Delta < \Delta_d$, where Δ_d is the positive root of the equation $F_1F_2 = 0$. Thus, for a given detuning satisfying the condition $0 < \Delta < \Delta_d$, only pulses having a certain unique duration can strongly excite the medium. A similar result was obtained in the case when the unidirectional approximation was used and the refractive indices were assumed to be equal. The carrier frequency of the ordinary component of a pulse that strongly excites the medium was also lower than the resonance frequency. Note that the parameter τ_p cannot be smaller than a certain minimum for pulses of this kind. Furthermore, in accordance with the classification adopted here, both regimes of strong excitation can be identified in this case: the SIT ($\alpha \longrightarrow 0$) and SIST ($\alpha \longrightarrow \infty$) regimes correspond to $\Delta \longrightarrow 0$ and $\Delta \longrightarrow \Delta_d$, respectively (see Fig. 1).

The formulas corresponding to the unidirectional approximation under condition (20) are obtained from (53) and (54) by setting $\tilde{A} = 0$, and the corresponding maximum detuning is $\Delta_m = 4\omega_0 d^2/D^2$. It can readily be shown that the interval of admissible detuning is wider for a high-density medium ($\Delta_d > \Delta_m$). In a low-density medium, the effective anisotropy is more pronounced, and a pulse can be resonant with the medium on average and thus cause strong excitation within a narrower interval of detuning. Note also that the medium effectively becomes more isotropic with increasing $\tilde{A}$, and the slope of the curve described by (53) increases with decreasing Δ. This is explained by the fact that the largest change in quantum-level population in an isotropic medium can occur only under exact resonance conditions.

The analysis developed above implies that a decrease in the effective anisotropy of a high-density medium does enhance characteristic properties of pulses in the SIT and SIST regimes. Conversely, the NNT, PNT, and EOT regimes (with $g \ll 1$) become less pronounced. In these regimes, the difference between the refractive indices of an anisotropic medium is not essential. The domains of regimes corresponding to (53) are schematized in Fig. 1.

7. MEDIA WITH PRONOUNCED POSITIVE AND NEGATIVE BIREFRINGENCE

Suppose that the concentration of quantum objects is so small that

$$\tilde{A}\omega_0^2\tau_p^2 \ll 1. \tag{56}$$

Then (32) can be rewritten as (33), as in the case when the unidirectional approximation was used. In contrast to the case of high-density medium considered in the preceding section, the parameter $\tilde{D}$ now depends on the optical anisotropy A_o. Therefore, pulses propagat-

ing through media characterized by positive and negative birefringence have different characteristics. In what follows, we show that the difference in refractive index manifests itself in the behavior of two-component pulses interacting with the medium.

First, we assume that the medium is positively birefringent ($n_e > n_o$, $A_o > 0$). In this case, $\tilde{D}$ is negative if the pulse parameters τ_p and Δ satisfy the inequality

$$\frac{1}{\tau_p^2} + \Delta^2 > \frac{\omega_0^2}{A_o}.$$

The polarity of the extraordinary component changes accordingly (see (27)). This leads to important consequences concerning strong excitation of the medium.

As $\tau_p \longrightarrow \tilde{\tau}_p$ and $\Delta \longrightarrow \tilde{\Delta}$, where $\tilde{\tau}_p$ and $\tilde{\Delta}$ satisfy the relation

$$\frac{1}{\tilde{\tau}_p^2} + \tilde{\Delta}^2 = \frac{\omega_0^2}{A_o},$$

the effective anisotropy increases indefinitely ($|\tilde{D}| \longrightarrow \infty$), and the group velocity of the pulse tends to the extraordinary-wave velocity ($v_g \longrightarrow c/n_e$). A pulse with τ_p and Δ close to $\tilde{\tau}_p$ and $\tilde{\Delta}$, respectively (see (2)), propagates in the EOT regime, because its extraordinary component is stronger than the ordinary one. Since the effective anisotropy is higher in the parameter domain in question, the EOT regime is more pronounced, especially when Δ is small.

If

$$\frac{1}{\tau_p^2} + \Delta^2 > \left(1 + \frac{2n_o}{n_o + n_e}\right)\frac{\omega_0^2}{A_0}, \qquad (57)$$

then the effective anisotropy cannot exceed D^2. Therefore, the NNT and PNT regimes ($g \ll 1$, $|\alpha| \gg 1$) are less pronounced in this parameter domain. Since $\tilde{D} < 0$, the detuning corresponding to these regimes changes sign; i.e., the former regime exists at $\alpha < 0$; the latter, at $\alpha > 0$.

Consider the case of strong excitation in a positively birefringent medium. It follows from (33) and (34) that the largest change in quantum-level population occurs if τ_p and Δ are such that

$$\frac{1}{\tau_p} = \sqrt{F}, \qquad (58)$$

where

$$F = \frac{\omega_0\Delta}{A_e + A_o\dfrac{\Delta}{\omega_0}} - \Delta^2. \qquad (59)$$

The condition $F > 0$ determines the range of admissible detuning. Using (58), we rewrite (33) as

$$\tilde{D} = 4d^2\left(A_e + A_o\frac{\Delta}{\omega_0}\right). \qquad (60)$$

We denote by

$$\Delta_0 = -\frac{A_e}{A_o}\omega_0$$

and

$$\Delta_{1,2} = (-A_e \pm \sqrt{A_e^2 + 4A_o})\frac{\omega_0}{2A_o}, \qquad (61)$$

the root of the equation $\tilde{D} = 0$ and the nonzero roots of the equation $F = 0$. Since $A_o > 0$ in the case in question, it holds that $\Delta_2 < \Delta_0 < \Delta_1$.

It is clear that the largest change in quantum-level population can be achieved for both positive and negative values of detuning ($0 < \Delta < \Delta_1$ and $\Delta_2 < \Delta < \Delta_0$). The existence of the latter interval of admissible detuning is due to the possibility of change in the sign of $\tilde{D}$ for positively birefringent media. Indeed, if $\Delta_2 < \Delta < \Delta_0$, then $\tilde{D} < 0$ and therefore $\Omega_e > 0$. This means that a negatively detuned pulse can strongly excite the medium, being resonant with it on average because of the dynamic shift induced by the extraordinary component.

In the case under consideration, the transparency regimes involving strong excitation of the medium are similar to those predicted under condition (20), but they can exist in two parameter domains. If $\Delta \longrightarrow 0$ or $\Delta \longrightarrow \Delta_0$, then $\alpha \longrightarrow 0$, and the pulse propagates in the SIT regime. In the latter limit case, both components have larger amplitudes and, therefore, higher velocities. If $\Delta \longrightarrow \Delta_1$ or $\Delta \longrightarrow \Delta_2$, then $|\alpha| \longrightarrow \infty$ and the pulse propagates in the SIST regime; i.e., its shape is sharper and velocity is higher as compared to the case when $\Delta \longrightarrow 0$. The corresponding transparency domains and the curve of strong excitation are shown in Fig. 2.

If $A_o \longrightarrow 0$, then $\Delta_1 \longrightarrow \omega_o/A_e$ and both Δ_2 and Δ_0 increase indefinitely. The corresponding curve of strong excitation is represented in Fig. 1 by a thin curve.

It should be noted that positively birefringent media, as well as isotropic media, are strongly excited by pulses with arbitrary τ_p. Since the present analysis relies on condition (56), the effective anisotropy exceeds

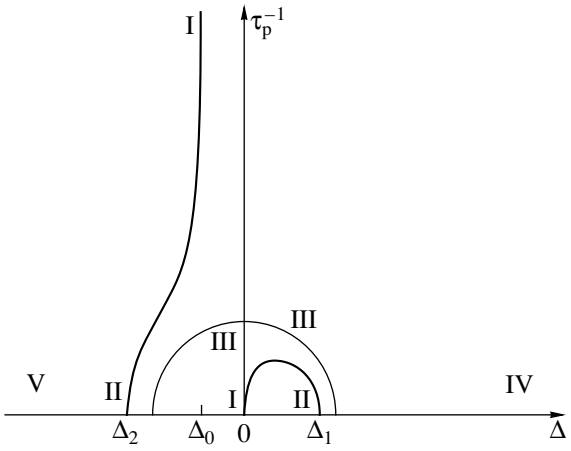

Fig. 2. Curve of strong excitation for positively birefringent medium. Parameter domains are shown for $|\Delta_2| \ll \omega_0$ and labeled as in Fig. 1. Thin curve: $v_g = c/n_e$.

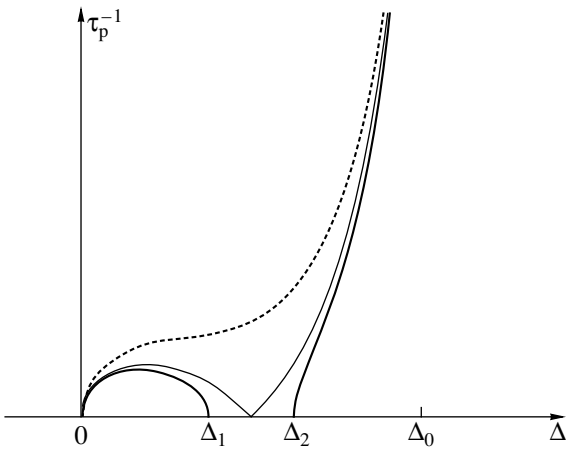

Fig. 3. Curve of strong excitation for negatively birefringent medium for $-A_o < A_e^2/4$, $-A_o = A_e^2/4$ (thin curve), and $-A_o > A_e^2/4$ (dotted curve).

D^2 in the parameter domain where inequality (57) is not valid. Therefore, both SIT and SIST regimes are less pronounced in the case of positive detuning, whereas the SIT regime corresponding to $\Delta \longrightarrow \Delta_0$ must be more pronounced because the effective anisotropy tends to zero.

Next, we consider negatively birefringent media ($n_o > n_e$, $A_o < 0$). It can readily be shown by using condition (56) that $\tilde{D} < D^2$; i.e., the effective anisotropy is weaker for such systems. The reason is that the linear extraordinary-wave velocity is higher as compared to the ordinary one and, therefore, is always higher than the pulse velocity. In this case, the extraordinary component plays a minor role in pulse formation.

Since negative birefringence reduces the effective anisotropy, regimes characterized by strong excitation are more pronounced. Equations (58)–(61) corresponding to these regimes remain valid. Note that condition (56) imposes a constraint on the parameter τ_p and, therefore, on detuning (see (58)). Accordingly, the detuning in (60) must be such that the parameter $\tilde{D}$ is less than D^2 even though the first summand on the right-hand side is greater than $\tilde{D}$.

According to (59), $F < 0$ for negative Δ; i.e., the pulses that cause the largest change in quantum-level population can be only positively detuned. This agrees with the fact that $\tilde{D}$ does not change sign in the case under study, i.e., the extraordinary component induces a dynamic red shift. However, as in the case of positive detuning, there exist two intervals of admissible detuning corresponding to strong excitation of the medium. Such conditions can be implemented if the medium is such that $0 < -A_o < A_e^2/4$. In this case, Δ_1 and Δ_2 are real and distinct; moreover, and $\Delta_1 < \Delta_2 < \Delta_0$. The SIT regime is implemented when either $\Delta \longrightarrow 0$ or $\Delta \longrightarrow \Delta_0$; the SIST regime, when either $\Delta \longrightarrow \Delta_1$ or $\Delta \longrightarrow \Delta_2$. If $-A_o = A_e^2/4$, then $\Delta_1 = \Delta_2 = \Delta_0$ and the parameter domain corresponding to SIST lies in the neighborhood of $\Delta = \Delta_0/2$. If $-A_o > A_e^2/4$, then the roots $\Delta_{1,2}$ are complex quantities, and the curve of strong excitation consists of a single branch. Under this condition, the SIST regime may not exist. As in the case of an isotropic medium, the parameter τ_p can have arbitrary values for pulses that cause the largest change in population in a negatively birefringent medium. Figure 3 shows the curves of strong excitation corresponding to several values of parameters of the medium.

The NNT, PNT, and EOT regimes must be less pronounced in negatively birefringent media because of weaker effective anisotropy. This behavior resembles that of high-density media, except that the effective anisotropy decreases with increasing absolute value of detuning in the present case.

The quasi-resonance approximation can be used to identify yet another transparency regime of two-component soliton propagation with nonlinear coupling between the components for $n_e \neq n_o$. In this case, system (16)–(19) cannot be reduced to a single wave equation. Neglecting the third-order linear dispersion in (46), as well as the nonlinearity $|\Omega_o|^2\Omega_o$, on account of the quasi-resonance condition $\Delta\tau_p \gg 1$ (i.e., minimizing the allowance for nonlinearity and dispersion), we combine Eqs. (45), (46), (18), and (19) to obtain

$$\begin{aligned} i\frac{\partial\tilde{\Omega}_o}{\partial y} + 2p\frac{\partial^2\tilde{\Omega}_o}{\partial\tau^2} &= -\mu_o\Omega_e\tilde{\Omega}_o, \\ \frac{\partial\Omega_e}{\partial y} + q\frac{\partial\Omega_e}{\partial\tau} &= \mu_e\frac{\partial}{\partial\tau}|\tilde{\Omega}_o|^2, \end{aligned} \qquad (62)$$

where

$$\tilde{\Omega}_o = \Omega_o \exp\left(-i\frac{\beta_o W_\infty}{\Delta} y\right), \quad \tau = t - \frac{y}{v_g},$$

$$\frac{1}{v_g} = \frac{n_o}{c} - \frac{\beta_o W_\infty}{\Delta^2},$$

$$q = \frac{n_e}{c} - \frac{1}{v_g}, \quad \mu_o = -\frac{\beta_o W_\infty}{\Delta^2},$$

$$\mu_e = -\frac{\beta_e W_\infty}{2\Delta^2}, \quad p = -\frac{\beta_o W_\infty}{2\Delta^3}.$$

When $q = 0$, i.e.,

$$v_g = c/n_e \tag{63}$$

Eqs. (62) are identical to the Yajima–Oikawa model [31], which describes interaction between long- and short-wavelength waves in the Zakharov–Benney resonance (ZBR) regime [32]. In this case, according to (63), the group velocity of the short-wavelength (ordinary) component equals the phase velocity of the long-wavelength (extraordinary) one. A system of this form was derived in [33, 34] by assuming that the ordinary-wave carrier frequency is tuned far off resonance with a multilevel system: $\omega \ll \omega_0$ or $\Delta \sim \omega_0$. The analysis performed here shows that a ZBR regime can also be implemented under quasi-resonance conditions. In the present case of pulse propagation through a medium in thermodynamic equilibrium, it holds that $v_g < c/n_o$. Therefore, condition (63) can be satisfied when $n_e > n_o$; i.e., the host crystal that contains resonant objects must be positively birefringent. Note that conditions (20) and (63) can be combined in one medium, for example, by using a constant electric field parallel to the anisotropy axis. The ensuing Stark effect will shift the frequency ω_0 of degenerate σ-transitions. By virtue of the frequency dependence of n_o, one can then switch between conditions (20) and (63). It was shown in [33, 34] that the solitons described by the Yajima–Oikawa system are stable in the region of normal dispersion, $p > 0$ or $\Delta > 0$.

The one-soliton solution to system (62) under condition (63) is

$$\tilde{\Omega}_o = \frac{8}{\Delta\tau_p}\frac{d}{D}\sqrt{\frac{n_e}{n_o}\omega\Omega} \times \exp[-i(\Omega t - ry)]\,\mathrm{sech}\frac{t - y/v_t}{\tau_p}, \tag{64}$$

$$\Omega_e = \frac{2}{\Delta\tau_p^2}\mathrm{sech}^2\frac{t - y/v_t}{\tau_p}, \tag{65}$$

where the free parameters τ_p and Ω can be interpreted as pulse width and carrier-frequency shift, respectively; $r = n_e\Omega/c + 2p(1/\tau_p^2 - \Omega^2)$; and the propagation velocity is defined by the relation

$$\frac{1}{v_t} = \frac{n_e}{c} - 4p\Omega. \tag{66}$$

If $p, \Omega > 0$, then $v_t > c/n_e = v_g$. Thus, the pulse velocity in the ZBR regime exceeds both linear phase velocity of the long-wavelength component and linear group velocity of the short-wavelength component. This phenomenon can be explained by the following argumentation. Suppose that a pulse containing only an ordinary (high-frequency) component enters a medium. According to the second equation in (62), since $\mu_e \propto D^2 \neq 0$, an extraordinary-wave video pulse is then generated in the medium under condition (63). In terms of photons, this means that each ordinary-wave photon will be red-shifted after a portion $\hbar\Omega$ of its energy has been transferred to an extraordinary wave: $\omega \longrightarrow \omega - \Omega$ ($\Omega \ll \Delta$). The expression for v_g becomes

$$\frac{1}{v_g} = \frac{n_o}{c} - \frac{\beta_o W_\infty}{(\Delta + \Omega)^2} \approx \frac{n_o}{c} - \frac{\beta_o W_\infty}{\Delta^2} - 4p\Omega = \frac{n_e}{c} - 4p\Omega,$$

which is equivalent to (66). As the ordinary component is detuned from resonance by the red shift, the medium is weakly excited.

Combining (64) with (65), we find

$$\frac{|\Omega_{om}|^2}{\Omega_{em}^2} = \frac{2g\Omega\tau_p n_e}{n_o}.$$

Since $\Omega \ll \Delta$ while $\Delta\tau_p \gg 1$, the product $\Omega\tau_p$ cannot be very large. Otherwise, both two-level and rotating-wave approximations should be abandoned. Assuming that $\Omega\tau_p \sim 1$ and $n_e/n_o \sim 1$ and taking into consideration the condition $g \ll 1$, we obtain $|\Omega_{om}|^2/\Omega_{em}^2 \ll 1$. Thus,

Characteristics of two-component pulse propagation regimes

	Δ_{ef}^m	$\lvert\alpha\rvert$	$\lvert g\rvert$	v_g	$\lvert\Omega_o/\Omega_e\rvert$	Excitation
SIT	-2Δ	$\ll 1$	$\gg 1$	v_{SIT}	$\gg 1$	strong
SIST	-2Δ	$\gg 1$	$\gg 1$	$< c/n_o$	~ 1	strong
NNT	$\approx\Delta$	$\gg 1$	$\ll 1$	c/n_o	arbitrary	weak
PNT	-5Δ	$\gg 1$	$\ll 1$	c/n_o	$\ll 1$	weak
EOT	$\gg\lvert\Delta\rvert$	$\ll 1$	$\ll 1$	$> v_{SIT}$	$\ll 1$	weak
ZBR	$\approx\Delta$	$\gg 1$	$\ll 1$	$> c/n_e$	$\ll 1$	weak

Note: Δ_{ef}^m is the effective detuning at the center of a pulse; Δ_{ef}^m is compared with Δ in the EOT in terms of absolute value.

the extraordinary component plays a dominant role in the ZBR regime. Moreover, the ordinary component is not phase modulated in this regime. However, as explained above, this component is modified by a uniform nonlinear red shift, $\delta\omega_{non} = -\Omega$.

8. CONCLUSIONS

The analysis presented in this paper leads to the conclusion that axial anisotropy greatly expands the variety of transparency regimes. Owing to the anisotropy, an input quasi-monochromatic pulse tuned off resonance can generate an extraordinary-wave video pulse, which changes the mechanism of its interaction with quantum objects.

In this study, we identify six regimes of two-component pulse propagation through a resonant medium, which can be classified with regard to propagation velocity and degree of excitation caused in the medium (see table). Self-induced transparency is characterized by strong excitation and substantial decrease in pulse propagation velocity relative to linear velocity. Self-induced supertransparency differs from SIT in that the decrease in velocity is small, but the medium is strongly excited as well (the propagating soliton causes total inversion of population followed by return to the initial state). In this regime, the ordinary-wave carrier frequency is lower than the resonance frequency if the refractive indices associated with the components are equal. The pulse width corresponding to the largest change in quantum-level population is determined by the amount of detuning. SIST solitons have larger amplitudes and smaller lengths as compared to SIT solitons, and their ordinary components are strongly frequency modulated. There also exist regimes that involve population trapping. The pulses propagating in the extraordinary transparency regime are characterized by small detuning and dominant role of the extraordinary component. Their group velocity substantially changes and may become comparable to that of pulses causing strong excitation. In the positive and negative nonresonant transparency regimes, the pulse velocity changes insignificantly, while the absolute value of detuning is large. The former regime is dominated by the extraordinary pulse component, whereas the ordinary-to-extraordinary amplitude ratio can be arbitrary in the latter. The most substantial difference between these regimes concerns the behavior of detuning. In the negative nonresonant transparency regime, it remains virtually constant. If a pulse propagates in the positive nonresonant transparency regime, then the effective detuning changes sign, passing through a resonance, owing to the effect of the extraordinary component. In the latter case, a slightly stronger excitation is obtained, and the pulses are sharply peaked, as in the SIST regime.

If the concentration of resonant objects is high and/or the ordinary and extraordinary refractive indices are different, then the anisotropy is determined not only by the properties of the medium, but also by pulse parameters. In a high-density medium, the effective anisotropy decreases for all pulses. Accordingly, transparency regimes characterized by strong excitation of quantum objects are more pronounced in such a medium. In the most interesting case of a strongly positively birefringent medium, the effective anisotropy can change sign. For this reason, SIT and SIST regimes are implemented not only when the carrier frequency is lower than the resonance one, but in the opposite case as well. In the PNT and NNT regimes, detuning changes sign. Furthermore, the effective anisotropy of the medium increases indefinitely as the group velocity of the pulse approaches the extraordinary-wave velocity. Under these conditions, pulses propagate in the extraordinary transparency regime, which is more pronounced, especially when the detuning is small. Moreover, a positively birefringent medium containing resonant particles can sustain propagation of a soliton-like bound state of a quasi-monochromatic ordinary wave and an extraordinary-wave video pulse in the Zakharov–Benney regime, owing to an essential role played by microscopic quantum objects having a permanent dipole moment. If the medium is strongly negatively birefringent, then its effective anisotropy decreases. As in high-density media, the transparency regimes involving strong excitation are more pronounced in this case. On the other hand, as in positively birefringent media, there can exist two intervals of ordinary-wave carrier frequency where the largest possible change in quantum-level population can occur. However, the carrier frequency remains lower than the resonance frequency, because the sign of the effective anisotropy does not change.

It is well known that there exists correspondence between coherent optical and acoustic phenomena. A correspondence of this kind is also observed between propagation of two-component electromagnetic and acoustical pulses through resonant media. In particular, an effect analogous to EOT was predicted for longitudinal–transverse acoustic pulse propagating parallel to an external magnetic field in a system of resonant paramagnetic impurities with effective spin $S = 1/2$ when the longitudinal amplitude of the pulse exceeds the transverse one [27]. Moreover, analogs of all regimes considered above exist in acoustic pulse propagation, because the corresponding governing equations are identical to system (16)–(18), (22) [35]. In these regimes, the role played by the high-frequency (transverse) component of an acoustic pulse is analogous to the role of the ordinary pulse component, and the longitudinal component is the counterpart of the extraordinary component.

As noted at the end of Section 2, both the system of LSWC equations and their gauge equivalents are well known in the theory of inverse scattering method and arise in other physical problems. In particular, systems of this type describe two-photon propagation of electromagnetic waves in a two-level medium [36, 37], inter-

action between two polarized electromagnetic waves in a two-level medium with cubic nonlinearity [38], four-wave mixing, and single-photon pulse propagation in a resonant medium involving the Stark effect [28]. The transparency regimes of soliton propagation analyzed here must also be manifested in these processes. The regimes of electromagnetic pulse propagation involving the Stark effect differ only by the phase modulation of the high-frequency field. In other cases, additional analysis is required to elucidate the physical meaning of particular differences, because the equations that are gauge equivalent to the LSWC system were derived by using nonlinear change of independent variables.

In this study, we ignored inhomogeneous spectral line broadening. A study allowing for this effect may throw light on the pulse area theorem for anisotropic media, providing a basis for analysis of their stability.

REFERENCES

1. S. L. McCall and E. L. Hahn, Phys. Rev. Lett. **18**, 908 (1967); Phys. Rev. **183**, 457 (1969).
2. A. I. Maimistov and A. M. Basharov, *Nonlinear Optical Waves* (Kluwer Academic, Dortrecht, 1999).
3. A. I. Maimistov, Kvantovaya Élektron. (Moscow) **30**, 287 (2000).
4. V. E. Zakharov, S. V. Manakov, S. P. Novikov, and L. P. Pitaevskiĭ, *Theory of Solitons: the Inverse Scattering Method* (Nauka, Moscow, 1980; Consultants Bureau, New York, 1984).
5. G. L. Lamb, Jr., *Elements of Soliton Theory* (Wiley, New York, 1980; Mir, Moscow, 1983).
6. *Solitons*, Ed. by R. K. Bullough and P. J. Caudrey (Springer, Berlin, 1980; Mir, Moscow, 1983).
7. G. L. Lamb, Rev. Mod. Phys. **43**, 99 (1971).
8. J. C. Eilbeck, J. D. Gibbon, P. J. Caudrey, and R. K. Bullough, J. Phys. A: Math. Gen. **6**, 1337 (1973).
9. L. A. Bol'shov, V. V. Likhanskiĭ, and M. I. Persiantsev, Zh. Éksp. Teor. Fiz. **84**, 903 (1983) [Sov. Phys. JETP **57**, 524 (1983)].
10. A. M. Basharov and A. I. Maĭmistov, Zh. Éksp. Teor. Fiz. **87**, 1594 (1984) [Sov. Phys. JETP **60**, 913 (1984)].
11. L. A. Bol'shov, N. N. Elkin, V. V. Likhanskiĭ, and M. I. Persiantsev, Zh. Éksp. Teor. Fiz. **94** (10), 101 (1988) [Sov. Phys. JETP **67**, 2013 (1988)].
12. A. M. Basharov and A. I. Maĭmistov, Zh. Éksp. Teor. Fiz. **97**, 1530 (1990) [Sov. Phys. JETP **70**, 864 (1990)].
13. A. Yu. Parkhomenko and S. V. Sazonov, Zh. Éksp. Teor. Fiz. **114**, 1595 (1998) [JETP **87**, 864 (1998)].
14. S. B. Leble and N. V. Ustinov, Chaos, Solitons & Fractals **11**, 1763 (2000).
15. R. Bavli and Y. B. Band, Phys. Rev. A **43**, 5039 (1991).
16. J. P. Lavoine, C. Hoerner, and A. A. Villaeys, Phys. Rev. A **44**, 5947 (1991).
17. G. B. Hadjichristov, M. D. Stamova, and P. P. Kircheva, J. Phys. B: Mol. Opt. **28**, 3441 (1995).
18. S. Kŏcinac, Z. Ikonić, and V. Milanović, Opt. Commun. **140**, 89 (1997).
19. L. W. Casperson, Phys. Rev. A **57**, 609 (1998).
20. M. Agrotis, N. M. Ercolani, S. A. Glasgow, and J. V. Moloney, Physica D (Amsterdam) **138**, 134 (2000).
21. A. A. Zabolotskiĭ, Zh. Éksp. Teor. Fiz. **121**, 1012 (2002) [JETP **94**, 869 (2002)].
22. A. I. Maĭmistov and J.-G. Caputo, Opt. Spektrosk. **94**, 275 (2003) [Opt. Spectrosc. **94**, 245 (2003)].
23. S. V. Sazonov, Zh. Éksp. Teor. Fiz. **124**, 803 (2003) [JETP **97**, 722 (2003)].
24. V. Ya. Demikhovskiĭ and G. A. Vugal'ter, *Physics of Quantum Low-Dimensional Structures* (Logos, Moscow, 2000) [in Russian].
25. N. N. Moiseev, *Asymptotic Methods of Nonlinear Mechanics* (Nauka, Moscow, 1981) [in Russian].
26. A. H. Nayfeh, *Introduction to Perturbation Techniques* (Wiley, New York, 1981; Mir, Moscow, 1984).
27. S. V. Voronkov and S. V. Sazonov, Zh. Éksp. Teor. Fiz. **120**, 269 (2001) [JETP **93**, 236 (2001)].
28. A. A. Zabolotskiĭ, Zh. Éksp. Teor. Fiz. **94** (11), 33 (1988) [Sov. Phys. JETP **67**, 2195 (1988)].
29. M. D. Crisp, Phys. Rev. A **8**, 2128 (1973).
30. A. M. Basharov and A. I. Maĭmistov, Opt. Spektrosk. **88**, 428 (2000) [Opt. Spectrosc. **88**, 380 (2000)].
31. N. Yajima and M. Oikawa, Prog. Theor. Phys. **56**, 1719 (1976).
32. R. K. Dodd, J. C. Eilbeck, J. Gibbon, and H. C. Morris, *Solitons and Nonlinear Wave Equations* (Academic, New York, 1982; Mir, Moscow, 1988).
33. S. V. Sazonov and A. F. Sobolevskiĭ, Pis'ma Zh. Éksp. Teor. Fiz. **75**, 746 (2002) [JETP Lett. **75**, 621 (2002)].
34. S. V. Sazonov and A. F. Sobolevskiĭ, Zh. Éksp. Teor. Fiz. **123**, 1160 (2003) [JETP **96**, 1019 (2003)].
35. A. A. Zabolotskiĭ, Zh. Éksp. Teor. Fiz. **123**, 560 (2003) [JETP **96**, 496 (2003)].
36. H. Steudel, Physica D (Amsterdam) **6**, 155 (1983).
37. D. J. Kaup, Physica D (Amsterdam) **6**, 143 (1983).
38. V. E. Zakharov and A. V. Mikhaĭlov, Pis'ma Zh. Éksp. Teor. Fiz. **45**, 279 (1987) [JETP Lett. **45**, 349 (1987)].

Translated by A. Betev

Journal of Experimental and Theoretical Physics, Vol. 100, No. 2, 2005, pp. 272–282.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 127, No. 2, 2005, pp. 308–319.

ATOMS, MOLECULES, OPTICS

On the Nature of "Coherent Artifact"

M. V. Lebedev[a], O. V. Misochko[a], T. Dekorsy[b], and N. Georgiev[b]

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

[b]*Institute for Ion Beam Physics and Materials Research, Forschungszentrum Rossendorf, P.O. Box 510119, D-01314 Dresden, Germany*

e-mail: lebedev@issp.ac.ru

Received June 18, 2004

Abstract—The coherent interaction of femtosecond laser pulses in the pump–probe regime has been experimentally studied in the time domain by monitoring light reflection from a tellurium single crystal. The optical response of the probed medium exhibits periodic variations at a frequency equal to that of the exciting laser radiation. Experimental dependences of the observed "coherent artifact" on the pump/probe intensity ratio, the number of accumulated pulses, and the mutual orientation of the polarization vectors of electromagnetic fields and the crystallographic axes are well described by the proposed phenomenological model. © *2005 Pleiades Publishing, Inc.*

1. INTRODUCTION

Recent decades have seen considerably growing interest in the study of response of condensed media to the action of ultrashort laser pulses, in which the energy is concentrated within a short time, small space, and narrow spectral intervals [1]. This research activity has been stimulated to a considerable extent by the progress in laser technologies in reducing the laser pulse duration down to the subpicosecond range, thus opening the way to real-time investigation of many processes related to the properties of solids.

One of the main approaches to investigations of fast dynamics is based on the pump–probe method. According to this approach, a medium is first excited by a short powerful laser pulse and then the optical properties of this excited medium are studied using much weaker probing pulses, typically obtained by separating part of the main laser beam (with the aid of a beam splitter) and attenuating it to a considerable degree. The first (pumping) pulse prepares a certain nonequilibrium state of the medium, whereas the second (probing) pulse of very low intensity (practically incapable of modifying the state of the medium) is used to study the process of system relaxation to the equilibrium state. However, these notions are completely inconsistent with the experimental situation in the region of zero delay of the probing pulse relative to the pumping pulse, where the two pulses are superimposed in the sample. In this case, we can no longer consider one pulse as pumping and the other merely as probing, since the excitation of the medium substantially depends on the coherent superposition of pulses, and the smallness of the probing pulse by no means implies that its effect is negligibly small.

Experiments in the time domain in the region of zero delays frequently show unusual and poorly reproduced deviations from the curves measured at greater delays and then extrapolated to a zero delay (Fig. 1). In the literature devoted to coherent phonons [2, 3], all these phenomena have been given the name "coherent artifact." This term (in our opinion, rather inadequate) has to reflect, first, the connection of the observed effects to coherent superposition of laser pulses and, second, the irregular and poorly controlled character of these phenomena. Phonons excited by ultrashort laser pulses are usually called coherent, since the exciting pulse duration is shorter than the period of lattice oscillations. The excitation of oscillations in nontransparent media can

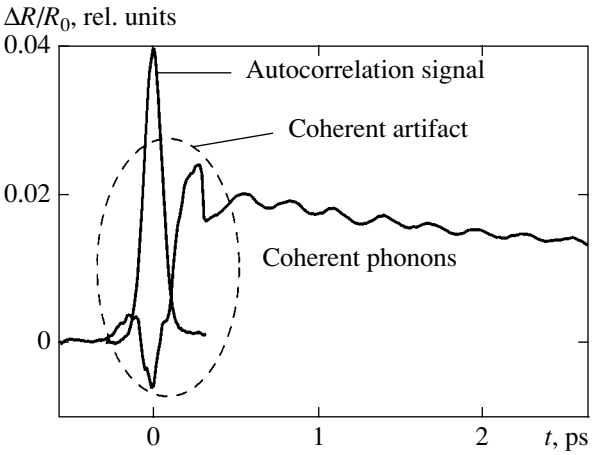

Fig. 1. Normalized time-resolved differential reflection signal $\Delta R/R_0$ from tellurium single crystal measured at room temperature for a time resolution of 6 fs. Laser pulse intensity, 0.1 mJ/cm^2.

1063-7761/05/10002-0272$26.00

be successfully described using phenomenological models based on the displacive excitation of coherent phonons and a mechanism close to the inelastic light scattering [2, 3]. It is a common practice to minimize the coherent artifact signal using orthogonal polarizations of the pumping and probing radiation, nondegenerate pump–probe scheme, etc. [2, 3]. There are several points of view on the nature of the coherent artifact. The formation of this signal is usually related to nonlinear optical effects (for collinear polarizations of the pumping and probing radiation) of the four-wave mixing [4]. According to this, the pumping pulse is scattered toward the probing pulse as a result of the phase grating formation.

This paper presents the results of thorough experimental and theoretical investigation of the coherent artifact in the time domain on a subfemtosecond resolution level. It is experimentally demonstrated that the "artifact" observed in the case of orthogonal polarizations of the pumping and probing radiation possesses a regular periodic structure, is perfectly reproduced, and can be described within the framework of a simple phenomenological model of the light–medium interaction. This analysis opens ways to investigating the response of a medium within ultimately short periods of time and, in addition, establishes the natural zero point for time count in the pump–probe method.

2. DESCRIPTION OF EXPERIMENT

Figure 2 shows a schematic diagram of the experimental arrangement. The experiments were performed with a tunable femtosecond Ti-sapphire laser of the Tsunami type (Spectra Physics) pumped at 532 nm from a Millenia Xs laser (Spectra Physics). The laser pulse shape was optimized using a double prism compressor and monitored with the aid of an autocorrelator. By shifting one of the two prisms, it was possible to adjust the phase modulation (chirp) so as to provide the optimum compensation of the group velocity dispersion in beam splitters, lenses, and other transparent elements in the pathways of both pumping and probing beams, thus ensuring the minimum pulse duration in the sample plane. This duration, as determined by monitoring the nonlinear transmission of a GaP crystal, was 90 fs. The spectral composition of the pulse was determined using a monochromator with a multichannel detection system. The pulse shape and spectrum were displayed in real time on the autocorrelator and computer displays. This setup allowed the pulse parameters to be continuously monitored and adjusted.

The measurements were performed using a fast scan system (AIXscan) comprising a shaker (angular retroreflector mounted on an electromagnetic vibrator driven by a high-precision oscillator), a delay line with a step motor, high-frequency twelve-digit analog-to-

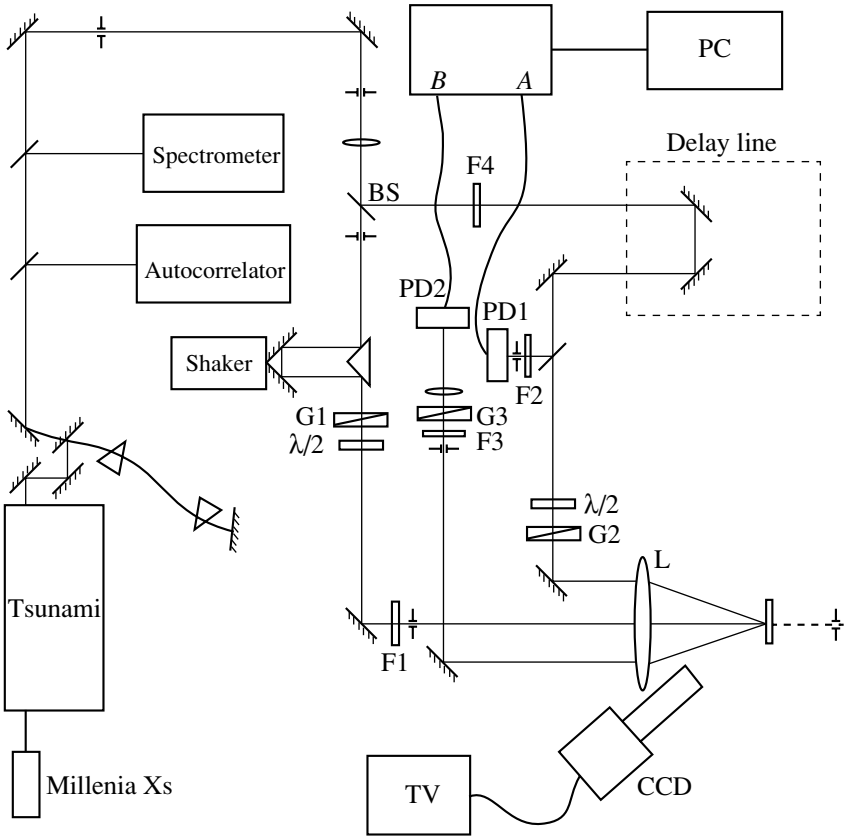

Fig. 2. Schematic diagram of the experimental arrangement (see the text for explanations).

digital amplitude converter (operating at 0.5 and 1.0 MHz), and a PC with special control software. The shaker made it possible to scan over and display a range of delay times with the initial value set by a delay line. The signal was digitized only during the forward run of the shaker. The measurements were performed at room temperature.

A sample crystal was placed at the focus of lens L (Fig. 2) with a focal distance of 90 mm, through which the pumping and probing pulses were transmitted parallel to the optical axis. The parallelism of laser beams entering the lens, which is a necessary prerequisite for the optimum spatial superposition of pumping and probing beams, was thoroughly checked at the system adjustment stage. The two beams were focused by the lens into one spot on the crystal surface. The imaging system, comprising a zoom objective, a CCD camera, and a TV monitor provided sharp beam focusing on the crystal surface and controlled matching of the pumping and probing beam spots. The pumping pulse was transmitted to the crystal via shaker, followed by a Glan prism G1 (eliminating parasitic polarization), a half-wave plate (rotating the polarization plate), and a neutral filter F1 of variable optical density (adjusting the required pump intensity).

The probing beam was obtained by separating a small part of the pumping beam with the aid of beam splitter BS. This beam passed through a delay line, a half-wave plate, and a Glan prism G2. The latter Glan prism was used to control polarization of the probing beam, while the intensity of this beam was adjusted by rotating the half-wave plate (changing the projection of the electric field vector onto the given polarization direction). The beams of pumping and probing pulses were separated by means of spatial filtration performed by diaphragm F3 mounted in front of a neutral light filter. Oscillations of the vibrator in the shaker ensured scanning of a certain interval of the time delay between pumping and probing pulses. The data were digitized in the high-frequency analog-to-digital amplitude converter and stored in a computer memory. The delay line was used for a rough leveling of the optical paths of the pumping and probing pulses and for calibration of the shaker sweep.

In order to increase the data accumulation rate and expand the working range of delay times provided by the shaker, it is necessary to increase the amplitude of the retroreflector and use a relatively high working frequency of the shaker. Evidently, just the opposite requirements have to be satisfied for reaching maximum time resolution of the system: this corresponds to the minimum velocity of the angular retroreflector that can be achieved by decreasing the amplitude and frequency of shaker vibrations. In our experiments, the shaker was operated with amplitude two to three times smaller than maximum and at a frequency of 40 Hz. The frequency was stable to within $\pm 10^{-3}$ Hz, which allowed the delay times in a range of 4 ps to be digitized at a step of 0.1 fs per point. The error caused by the frequency instability was comparable with the discreteness of digitization. A decrease in the range of scanned delay times and the frequency of shaker operation could increase the accuracy of digitization (under otherwise equal conditions). However, in our case this would detrimentally influence the conditions of operation of the program of automated time scale calibration. This program set a series of fixed time delays between the pumping and probing pulses, compared these delays to the shift of a pulse reflected from the shaker, and determined the coefficients of time scale calibration, thus taking into account nonlinearity of the sweep. The minimum parameters at which this calibration program could successfully operate corresponded to a shaker frequency of 34.76 Hz and a calibration range of about 4 ps. In a standard probe–pulse experiment with the response signal measured using a lock-in detector, the discreteness of digitization is determined by the delay line step and usually does not exceed 6.6 fs per point [2, 3].

We have measured the relaxation of changes in the reflection coefficient (or in the transmission coefficient, in the case of a transparent GaP crystal used for the system calibration) caused by the pumping pulse. Since these changes are relatively small, we have used a special differential technique in order to increase the accuracy of measurements. According to this, a part of the probing beam was separated (before striking the sample crystal) by the beam splitter and directed to photodiode PD1 (Fig. 2). The probing beam reflected from the crystal was detected by photodiode PD2. The signals from both detectors were fed to the inputs (*A* and *B*) of a differential amplifier, where the difference *A*–*B* was amplified and measured using a high-frequency analog-to-digital converter of the AIXscan system. The *A* and *B* signal channels were thoroughly balanced (whereby the signals from photodiodes were equilibrated using a neutral filter F2 with variable optical density mounted in front of photodiode PD1) in the absence of pumping pulses. As a result, we monitored the signal proportional to a change in the reflection coefficient, which was on a level from 10^{-3} to 10^{-5} of the absolute value of this coefficient. The differential amplifier was equipped with a built-in frequency filter, which allowed the working signal to be measured in a preset frequency band and thus increased the signal to noise ratio. For the study of coherent phonons, it was sufficient to use a frequency band from 1 Hz to 3 kHz, where virtually no distortions were introduced into the corresponding oscillations (with a period of about 300 fs). In investigations of the coherent artifact, the period of signal variation was about 3 fs and this required the working frequency band of the amplifier to be increased (in our measurements, the frequency band was typically from 1 Hz to 300 kHz).

All measurements were performed for a basal plane of a tellurium single crystal crystallized at room

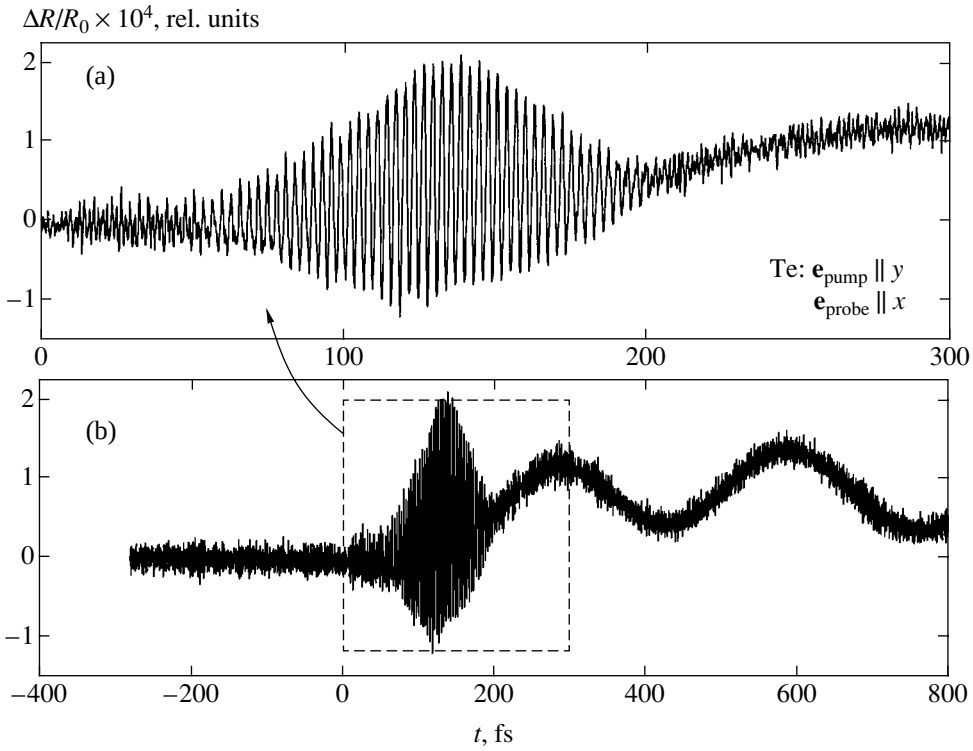

Fig. 3. (a) The shape of a coherent artifact signal with well-pronounced regular structure measured at a high resolution; (b) the coherent artifact on the background of first two oscillations of the reflection coefficient, related to the excitation of fully symmetric (A_1) coherent phonons in a tellurium single crystal. Pumping laser beam power, 37 mW; probing laser beam power, 3.2 mW; number of accumulated runs, 100; polarization: $\mathbf{e}_{pump} \parallel y$, $\mathbf{e}_{probe} \parallel x$.

temperature in a hexagonal lattice belonging to space group D_3^6.

3. EXPERIMENTAL RESULTS

In the measurements of coherent phonons, the coherent artifact is usually manifested by a special region in the initial part of relaxation curves oscillating at a frequency of the phonon mode [2, 3]. This region is clearly seen in Fig. 3b together with two initial oscillations of the reflection coefficient caused by excitation of the coherent fully symmetric phonons in the tellurium crystal [5]. A comparison of Figs. 1 and 3 shows that an increase in the time resolution reveals a regular periodic structure of the artifact (see Fig. 3a). It is seen at first glance that the rapid oscillations are performed not relative to the zero line, but relative to a certain slowly varying (on the given time scale) function. This function can be readily separated from the oscillating curve by averaging over the oscillation period. The results of such separation for various pump/probe intensity ratios are presented in Fig. 4. At equal intensities of the pumping and probing beams, rapid oscillations are virtually symmetric relative to the zero level. If the pumping pulse intensity is significantly higher than that of the probing pulse, the artifact exhibits a slowly varying positive additive to the reflection coefficient. The magnitude of this additive also significantly depends on the orientation of light polarization relative to the crystallographic axes. The ratio of the slowly varying and rapidly oscillating components of the response also changes when the beams change their roles, whereby the probing beam is used for pumping and vice versa (in experiment, this is achieved by exchanging photodiodes between the probing and pumping beams, while retaining the same beam polarizations).

Figure 5 shows a change in the artifact in the course of a gradual increase in the probing beam intensity in the case when the polarizations of both beams were rotated by 90° as compared to their orientations in the experiment presented in Fig. 4. As can be seen, a change in the ratio of beam intensities leads to a change in the ratio of slowly varying and rapidly oscillating components of the response (analogous to that shown in Fig. 4). However, in this geometry, the contribution of the slowly varying component is greater than in the case presented in Fig. 4. This was specially checked for the same number of accumulated data, which is an important circumstance since it was found that the shape of the artifact depends on the number of accumulated runs. This dependence is illustrated in Fig. 6, which shows that an increase in this number is accompanied by a decrease in the relative contribution of the oscillating part of the response. The effect is related to

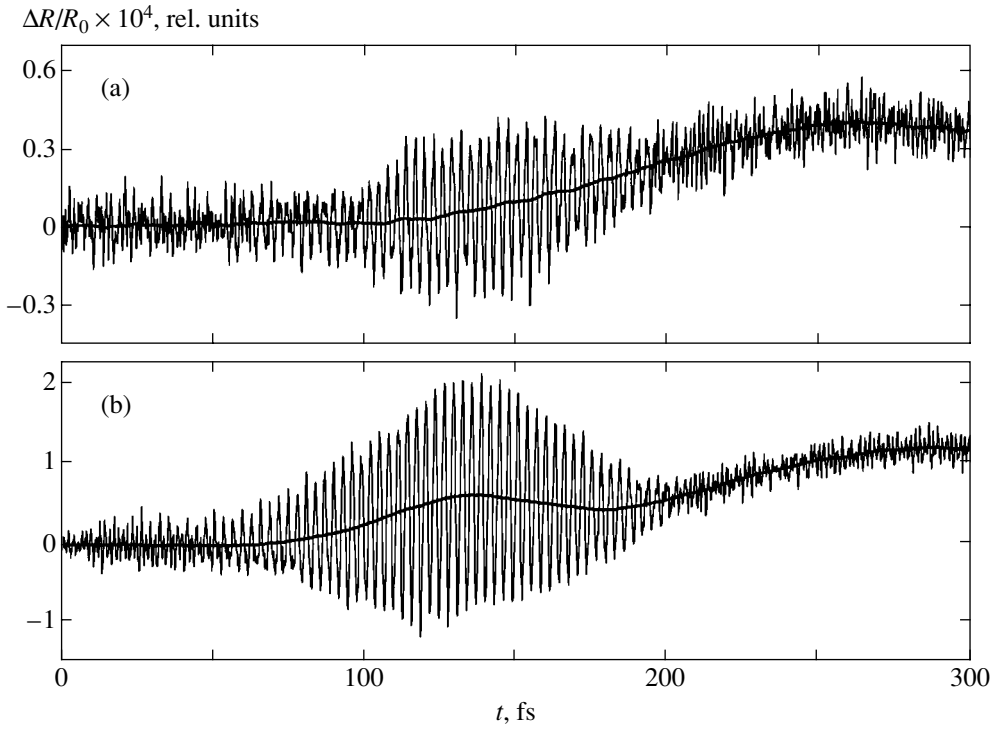

Fig. 4. The influence of the ratio of the pumping and probing beam intensities on the shape of the coherent artifact. An increase in the pumping beam intensity leads to the appearance of a slowly varying positive component. Slowly varying background curves were obtained by smoothening over the period of the artifact. Pump/probe intensity ratio (mW): (a) 7.2/7; (b) 37/3.2; number of accumulated runs, 1000; polarization: $\mathbf{e}_{\text{pump}} \parallel y$, $\mathbf{e}_{\text{probe}} \parallel x$.

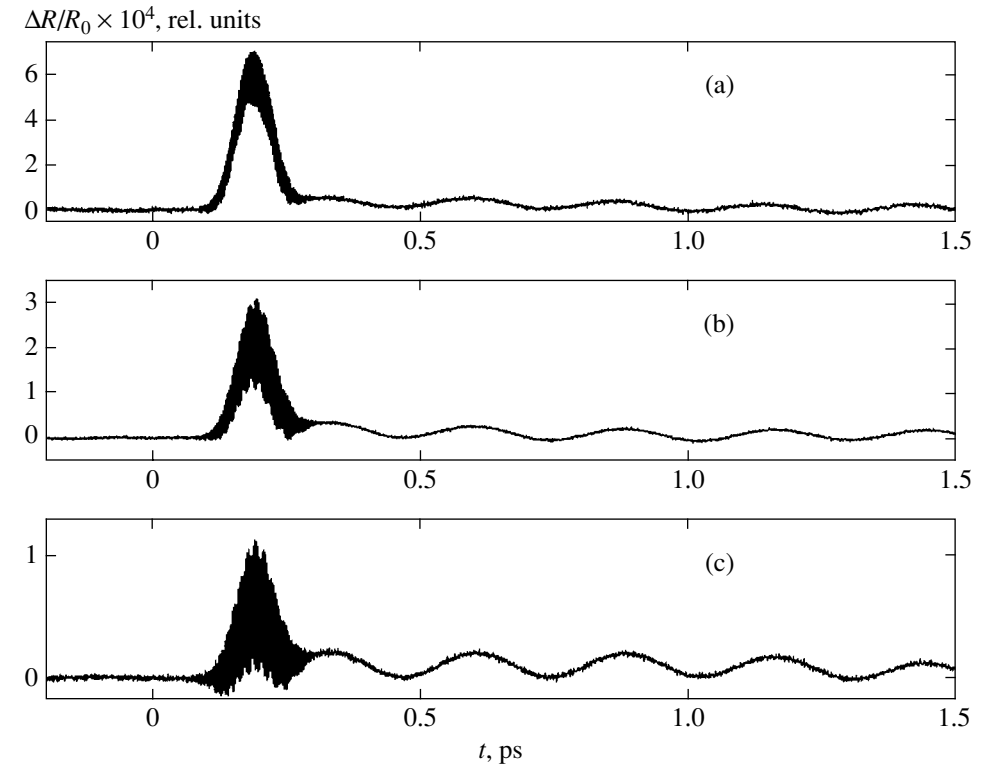

Fig. 5. The influence of the ratio of the pumping and probing beam intensities on the shape of the coherent artifact measured for a laser beam polarization other than that in Fig. 4: $\mathbf{e}_{\text{pump}} \parallel x$, $\mathbf{e}_{\text{probe}} \parallel y$; number of accumulated runs, 10000; pumping laser beam power, 28.4 mW; probing laser beam power (mW): (a) 7.2; (b) 18.2; (c) 28.7.

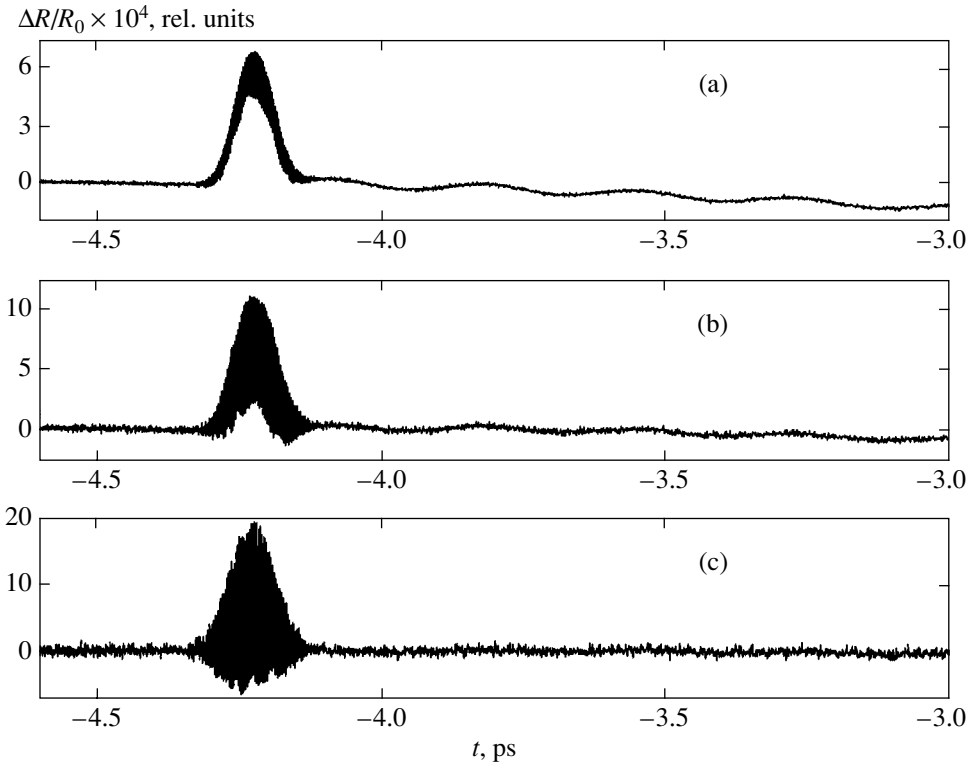

Fig. 6. Variation of the coherent artifact depending on the number of accumulated runs: (a) 10000; (b) 1000; (c) 100. An increase in this number leads to a decrease in the relative contribution of the rapidly oscillating component. Polarization: $\mathbf{e}_{\text{pump}} \parallel x$, $\mathbf{e}_{\text{probe}} \parallel y$; pumping laser beam power, 28.4 mW; probing laser beam power, 7.2 mW.

averaging of the oscillating curve with time as a result of uncontrolled phase fluctuations appearing, probably, because of imperfection (the lack of interferometric precision) of the experimental setup.

Figure 7 shows the results of two sequential experiments performed using the same number of accumulated runs and equal values of all controlled experimental parameters. As can be seen, one of the curves exhibits clearly pronounced oscillations with a single period of 2.67 fs. The amplitude of these oscillations in the other curve is significantly decreased and additional small peaks are observed between the main maxima. The change is most probably caused by a spontaneous fluctuation in the phase of shaker oscillations. The period of the oscillating component in the response (2.67 fs) coincides to within 4.8% with the period of the carrier wave of the light pulse (2.80 fs). The latter is determined to within 1.3% (for an excitation laser wavelength of 841 ± 11 nm).

The above experimental data elucidate the reasons for which the artifact was previously observed in the form of an irregular and poorly reproducible behavior of the relaxation curves in the region of zero delay times. First, the standard time resolution (not better than 6.6 fs) did not allow detecting oscillations with a period shorter than 3.0 fs. Second, the goal of increasing the signal to noise ratio in the investigation of features of the relaxation kinetics required using large signal accumulation times, selecting sufficiently high shaker frequency, and employing optimum signal filtration regimes. All these factors lead to effective averaging of the oscillating component of the artifact and strongly distort its shape, the more so that it is highly sensitive to small uncontrollable phase fluctuations. The sensitivity of the artifact to intensities of the beams and their polarization relative to the crystal axis still complicates the pattern.

All the experimental results considered above referred to differential measurements of the relaxation of a change in the reflection coefficient of tellurium single crystal exposed to a high-power femtosecond laser pulse. It should be emphasized that very similar results were obtained when the metal crystal was replaced by a crystal of high-temperature superconductor or manganite. However, for the sake of brevity, the consideration is restricted to the analysis of data obtained for tellurium.

In the experiments with all nontransparent (strongly absorbing) crystals, the time sweep of the shaker was calibrated using a standard procedure, whereby a sample crystal in the lens focus was replaced by a transparent (in the laser wavelength range employed) crystal of GaP (an indirect-band semiconductor with a bandgap width of 2.8 eV at the Γ point; $n = 3.24$ at 850 nm [6]). The photodiode was moved from the reflected beam to the transmitted beam and the same differential tech-

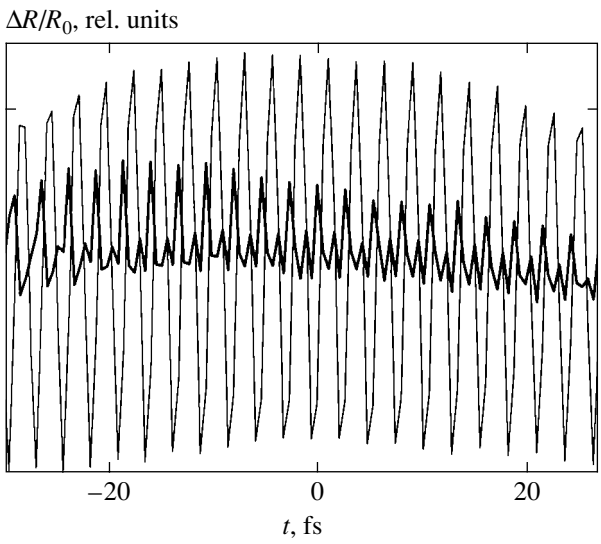

Fig. 7. Two time series of the coherent artifact recorded for the same fixed experimental parameters. A sharp decrease in the magnitude of oscillations and the appearance of additional peaks in one of the curves is probably related to a spontaneous fluctuation in the phase of shaker vibrations. Number of accumulated runs, 300; pumping laser beam power, 28 mW; probing laser beam power, 18 mW; polarization: $\mathbf{e}_{\text{pump}} \parallel x$, $\mathbf{e}_{\text{probe}} \parallel y$.

nique was used to measure a change in the transmission of the probing beam after excitation of the GaP crystal by the pumping pulse. It is commonly accepted that (i) a decrease in the probing beam transmission observed at the moment of superposition of the pumping and probing pulses is related to the two-photon absorption and (ii) the shape of the detected signal is proportional to the autocorrelation function of laser pulse intensities. This provides a simple and convenient method of optimization of the excitation pulse duration immediately at the site where the sample will be placed, while rather large signal amplitude makes it possible to use this method for calibration of the shaker sweep. The two-photon absorption is insensitive to the phase of the light wave. The interference between the pumping and probing beams was excluded by using orthogonal polarizations of the two beams. Despite this, our study of the nonlinear transmission of a GaP crystal at a high time resolution revealed the presence of oscillating component in the response signal, which had a shape quite similar to that presented in Fig. 4, with a period close to that of the carrier wave. Further investigation showed that, under certain conditions, it is possible to observe a change in the sign of the differential transmission, whereby the pumping pulse induces an increase (rather than a decrease) in transmission of the crystal. Moreover, the response of a GaP crystal has proved to be sensitive not only to the phase difference between the pumping and probing beams, but also to the phase modulation of the laser pulse that could be varied with the aid of a prism compressor. All these features are difficult to explain within the framework of a simple two-photon absorption model, although this process can still play a certain role in the phenomena under consideration.

Consistent theoretical analysis of the propagation of a high-power femtosecond laser pulse in a nonlinear medium is a highly complicated problem and there are a huge number of publications on this subject. As for the task of description of the nonlinear reflection of a short high-power laser pulse, this problem admits a simple solution within the framework of a simple phenomenological model considered below.

4. DISCUSSION OF RESULTS

It should be noted that the observed phenomena cannot be explained by simple interference between the pumping and probing beams. Indeed, polarizations of the two beams in all experiments were orthogonal and the reflected signal was detected at a site distant from the region of beam overlap, so that no superposition of light waves took place in the detector.

The coherent phonons excited in a tellurium crystal by the pumping laser pulse are manifested in the experimental curves by a periodic time variation of the reflection coefficient of the sample (Fig. 3). We naturally assumed that the observed periodic temporal structure of the coherent artifact is also related to a periodic time variation of the reflection coefficient under the action of electromagnetic field of the pumping pulse. Based on this assumption, it is possible to provide a phenomenological description of the properties of a coherent artifact observed in experiment.

4.1. Phenomenological Model

Let us describe a laser pulse at the crystal surface by a wave packet,

$$\Phi(t) = A(t)\exp[-i(\omega_0 t + \psi(t))], \qquad (1)$$

where ω_0 is the carrier frequency and $A(t)$ and $\psi(t)$ are slowly varying amplitude and phase. Since only a surface layer of the crystal with a thickness much smaller than the wavelength contributes to the reflection coefficient, we may decline from analysis of the character of light propagation in the crystal and solve an essentially local problem. This circumstance makes possible a simple description of the experiment. In order to simplify the formulas, we will not take into account the vector character of electromagnetic fields. If necessary, the form of $A(t)$ and $\psi(t)$ functions can be specified. These functions are quite rigidly set by the condition of femtosecond laser pulse generation using synchronized modes of a Ti-sapphire laser. In particular, the carrier frequency is a higher harmonic of the fundamental mode of the laser cavity and the amplitude $A(t)$ is directly related to the number of synchronized modes. The phase $\psi(t)$ includes the linear and quadratic terms, the latter determining the phase modulation (chirp) of

the pulse. The carrier frequency and the spectral width of a pulse can be directly measured in experiment. The chirp determining (together with the slowly varying amplitude) the spectral width can be varied with the aid of the prism compressor.

We assume that a laser pulse produces a change in the coefficient of reflection of the crystal surface, which varies with time as

$$R(t) = R_0 + r(t), \tag{2}$$

where R_0 is the background reflection coefficient. Accordingly, the amplitude of the reflected pulse can be written as

$$\begin{aligned}\tilde{\Phi}(t) &= \Phi(t)R(t)\\ &= R(t)A(t)\exp[-i(\omega_0 t + \psi(t))].\end{aligned} \tag{3}$$

Let us express the variable increment of the reflection coefficient as

$$r(t) = a(\mathbf{e})\Phi(t), \tag{4}$$

where $a(\mathbf{e})$ is a coefficient taking into account the dependence of the increment on the orientation of the polarization of light relative to the crystallographic axes. In the general case, this coefficient is a complex quantity whose modulus determines the magnitude of the increment and the argument reflects the possible phase shift of the laser-induced change in the reflection coefficient relative to the phase of the electric field vector of the pulse:

$$a(\mathbf{e}) = |a|e^{i\varphi}. \tag{5}$$

Under the conditions of superposition of the pumping and probing laser pulses, the increment of the reflection coefficient can be written as

$$r(t) = a_{\text{pump}}\Phi_{\text{pump}}(t+\tau) + a_{\text{probe}}\Phi_{\text{probe}}(t), \tag{6}$$

where τ is the time delay between the pumping (Φ_{pump}) and probing (Φ_{probe}) pulses. Then, the amplitude of the reflected probing pulse is

$$\begin{aligned}\tilde{\Phi}_{\text{probe}}(t) &= \Phi_{\text{probe}}(t)\\ &\times (R_0 + a_{\text{pump}}\Phi_{\text{pump}}(t+\tau) + a_{\text{probe}}\Phi_{\text{probe}}(t)).\end{aligned} \tag{7}$$

Since the photodiodes measure the light intensity averaged over the response time, the measured differential signal S can be determined by multiplying the amplitude (7) by its complex conjugate, averaging the product over the diode time constant, and rejecting the terms not containing Φ_{pump} (which are compensated in a balanced differential scheme). As a result, we obtain

$$\begin{aligned}S \propto\ & \langle a_{\text{pump}}^2 A_{\text{pump}}^2(t+\tau)A_{\text{probe}}^2(t)\rangle\\ &+ 2\langle a_{\text{pump}}A_{\text{pump}}(t+\tau)a_{\text{probe}}A_{\text{probe}}^3(t)\\ &\times \cos(\omega_0\tau + \Delta\psi + \Delta\varphi)\rangle,\end{aligned} \tag{8}$$

where the angle brackets denote averaging with respect to time, the a_i and A_i are assumed to be positive (all signs are included into the phase factors), and the terms $\Delta\psi$ and $\Delta\varphi$ are given by the formulas

$$\begin{aligned}\Delta\psi &= \psi_{\text{pump}}(t+\tau) - \psi_{\text{probe}}(t),\\ \Delta\varphi &= \varphi_{\text{pump}}(t+\tau) - \varphi_{\text{probe}}(t).\end{aligned} \tag{9}$$

The linear term in the pulse phase increment, which is related to a difference in optical paths between the pumping and probing beams, is explicitly taken into account by the delay time τ. Then, $\Delta\psi$ is determined entirely by the quadratic term, that is, by the chirp. For the sake of simplicity, we will assume the phase modulation (chirp) in the sample plane to be zero. The second formula (9) implies that the delay of the induced increment in the reflection coefficient depends on the slowly varying amplitude and on the pulse phase (chirp). It is natural to assume that this dependence is weak and can be ignored to the first approximation. Under these assumptions, expression (8) can be rewritten as

$$\begin{aligned}S \propto\ & \langle a_{\text{pump}}^2 A_{\text{pump}}^2(t+\tau)A_{\text{probe}}^2(t)\rangle\\ &+ 2\langle a_{\text{pump}}A_{\text{pump}}(t+\tau)a_{\text{probe}}A_{\text{probe}}^3(t)\rangle\cos\omega_0\tau.\end{aligned} \tag{10}$$

This expression qualitatively describes all the main features of the coherent artifact observed in experiment. First, it contains the rapidly oscillating component of the response varying with the period of the carrier wave. The slowly varying component is always positive (as in experiment) and coincides to within a factor (a_{pump}^2) with the correlation function of the laser pulse intensities. The presence of the factor a_{pump}^2 explains why the slowly varying contribution depends on the orientation of the polarization of the pumping pulse relative to the crystal axes. The ratio of the second term to the first term in expression (10) is proportional to

$$2\frac{a_{\text{probe}}A_{\text{probe}}}{a_{\text{pump}}A_{\text{pump}}},$$

which implies that the contribution of the rapidly oscillating component increases with the probing pulse amplitude. Not that this ratio also depends on the orien-

tation of the polarization of the pumping pulse relative to the crystallographic axes.

Although the proposed model quite satisfactorily describes the trends observed in experiment, it is necessary to develop a consistent microscopic description of these phenomena, since the phenomenological model is self-contradictory. Indeed, the reflection coefficient is a single-valued function of the susceptibility $\chi(\omega)$ of the medium. The susceptibility, characterizing the optical properties of the medium, is defined as the coefficient of proportionality between the Fourier components of the electric field and the polarization of the medium at the same frequency ω. Therefore, the susceptibility $\chi(\omega)$ and, hence, the reflection coefficient cannot explicitly depend on the time, as it was assumed in proposed phenomenological model. It should be noted that, in the case of coherent phonon excitation, the susceptibility implicitly depends on the time because of nonstationary excitation of lattice vibrations in the medium [2, 3].

4.2. Analysis of the Adequacy of the Phenomenological Model

We believe that the physical reason ensuring the adequacy of the proposed phenomenological model is the conservation of the total probability of finding the electron subsystem of a crystal either in the ground or in the excited state in the time scale on the order of laser pulse duration. This, in turn, is a consequence of the smallness of the laser pulse duration as compared to the characteristic times of the energy and phase relaxation of the electron subsystem. This can be illustrated by a simple example of a two-level model of the medium, despite the fact that this model is oversimplified and does not aspire to provide a detailed description of real experiments.

Consider a medium consisting of identical two-level systems with a transition frequency ω_{21}, which are distributed in the volume with a density of N. As is known, the susceptibility of such a medium initially occurring in a ground (nonexcited) state can be expressed in a quasi-stationary case as [7]

$$\chi'(\omega) = \frac{d^2N}{\hbar}\frac{\Delta}{\Delta^2 + \Gamma_2^2 + \frac{\Gamma_2}{\Gamma_1}\Omega^2},$$
$$\chi''(\omega) = \frac{d^2N}{\hbar}\frac{\Gamma_2}{\Delta^2 + \Gamma_2^2 + \frac{\Gamma_2}{\Gamma_1}\Omega^2}, \tag{11}$$

where $\chi'(\omega)$ and $\chi''(\omega)$ are the real and imaginary parts of the susceptibility, respectively; d is the matrix element of the dipole moment of the transition; $\Delta = \omega_{21} - \omega$ is the frequency detuning; Γ_2 and Γ_1 are the transverse and longitudinal relaxation constants, respectively; and Ω is the Rabi frequency. These formulas for the real and imaginary parts of the susceptibility were obtained assuming that a dipole transition is allowed between levels of the two-level system. For $\Gamma_2 = \Gamma_1 \longrightarrow 0$, formulas (11) give correct expressions for the susceptibility in the absence of relaxation:

$$\chi'(\omega) = \frac{d^2N}{\hbar}\frac{\Delta}{\Delta^2 + \Omega^2},$$
$$\chi''(\omega) = 0. \tag{12}$$

These expressions can be also obtained directly, by solving a nonstationary Schrödinger equation with harmonic time-dependent perturbation:

$$i\hbar\frac{\partial\Psi}{\partial t} = [H_0 + V(t)]\Psi,$$
$$V(t) = V\cos\omega t. \tag{13}$$

Let us consider a somewhat more general case of this problem with a perturbation of the following type:

$$V(t) = V_1\cos\omega t + V_2\cos(\omega t + \phi). \tag{14}$$

As can be readily seen, perturbation (14) with real V_1 and V_2 values can be reduced to the form (13) by substitution

$$V = \sqrt{V_1^2 + 2V_1V_2\cos\phi + V_2^2},$$
$$\omega t \longrightarrow \omega t + \theta, \tag{15}$$
$$\sin\theta = -\frac{V_2\sin\phi}{\sqrt{V_1^2 + 2V_1V_2\cos\phi + V_2^2}}.$$

This implies that a solution of the problem with perturbation (14) reduces to solution of the problem (13) with then effective Rabi frequency

$$\tilde{\Omega} = \sqrt{\Omega_1^2 + 2\Omega_1\Omega_2\cos\phi + \Omega_2^2}, \tag{16}$$

where $\Omega_i = V_i/\hbar$. Not that, since the form of the operator of interaction between the electromagnetic field and the two-level system was not specified, formulas (14)–(16) are valid not only for the electric dipole interaction, but for the other (e.g., quadrupole) interactions as well. In the case of the electric dipole interaction, the susceptibility in the absence of relaxation is

$$\chi'(\omega) = \frac{d^2N}{h}\frac{\Delta}{\Delta^2 + \Omega_1^2 + 2\Omega_1\Omega_2\cos\phi + \Omega_2^2},$$
$$\chi''(\omega) = 0. \tag{17}$$

The fact that the susceptibility is real allows us to use a relation between the susceptibility and refractive index, which is valid for transparent media:

$$n = \sqrt{1 + 4\pi\chi'(\omega)} = \sqrt{1 + 4\pi\frac{d^2N}{\hbar}\frac{\Delta}{\Delta^2 + \tilde{\Omega}^2}}. \quad (18)$$

Accordingly, for $\tilde{\Omega} \ll \Delta$, the reflection coefficient is

$$R(\omega) = \frac{1-n}{1+n} = \frac{\pi d^2 N}{\hbar} \times\left[-\frac{1}{\Delta} + \frac{\Omega_1^2}{\Delta^3} + \frac{\Omega_2^2}{\Delta^3} + \frac{2\Omega_1\Omega_2\cos\phi}{\Delta^3}\right]. \quad (19)$$

A comparison of expression (19) to the phenomenological formulas (2) and (6) shows their close similarity. In the case of a quasi-stationary excitation, Ω_1 and Ω_2 depend on the slowly varying amplitude and phase, that is, are also slowly varying functions of time. If the phased modulation is absent, we have

$$\Omega_i = \frac{V_i}{\hbar} = \frac{d}{\hbar}A_i(t),$$

where $A_i(t)$ are slowly varying pulse amplitudes. The first term in square brackets of expression (19) corresponds to the background reflection coefficient R_0, while the second and third terms correspond to changes in the reflection coefficients induced separately by the pumping and probing pulses. The last term (which is absent in the phenomenological formula) describes the mutual influence of the two pulses. Formula (19) does not contain explicit dependence of the reflection coefficient on the time with a period of the carrier wave, but it naturally reflects the influence of the relative phase of the two waves, $\phi = \omega_0\tau$, which removes the inconsistency of the phenomenological approach. Expression (19) shows that, during the coherent interaction of two laser pulses with the medium. The reflection coefficient cannot be independently determined for each of the waves and has a physical meaning only for their combination. Even a small change in amplitude of one of the waves will lead to a change in the reflection coefficient for the other wave as well. Using relation (19), we readily obtain an expression for the reflected signal measured using our differential technique:

$$S = -\frac{2\pi^2 d^6 N^2}{\hbar^4\Delta^4}[\langle A_{\text{pump}}^2(t+\tau)A_{\text{probe}}^2(t)\rangle + 2\langle A_{\text{pump}}(t+\tau)A_{\text{probe}}^3(t)\rangle\cos\omega_0\tau]. \quad (20)$$

This expression coincides in form with relation (10) obtained within the framework of the phenomenological model. It should be noted that the dependence of the signal on the orientation of the polarization of laser beams relative to the crystal axes can be also determined taking into account the tensor character of the dipole moment of the transition. We did not perform this analysis for the sake of simplicity. The only significant difference between formula (20) and relation (10) is related to the sign. According to the two-level model, a nonlinear additive to the reflection coefficient must always lead to a decrease in reflection of the probing beam, whereas experiment shows the opposite trend.[1]

Thus, a simple two-level model unexpectedly provides a good qualitative description of the observed features of the coherent artifact. The physical reason of the mutual influence of laser pulses in this model is absolutely clear, being related to the conservation of the total probability. As is known, the interaction of a two-level system with an external harmonic field leads to periodic oscillations in the probability to find the system in an excited state. Therefore, a delayed probing pulse finds the system in the state of a superposition with a certain phase whose value determines the interaction of the probing pulse with the excited system.

5. CONCLUSIONS

The results presented in this paper show that the so-called coherent artifact possesses a regular structure, has a reproducible character, and appears as a result of conservation of the total probability of finding the electron system in one of the two possible (ground or excited) states in the course of the interaction of the medium with spatially superimposed pumping and probing pulses. A good qualitative description of the experimentally observed properties of the coherent artifact, provided by the simple phenomenological model, makes it possible to use this artifact as a natural marker determining the zero point on the relaxation curve. This approach to determining the zero point offers considerable advantages to the other known methods (e.g., reversal of the time axis by means of changing the role of the beams with the transfer of photodiodes from probing to pumping beam channel or use of an auxiliary crystal replacing the sample [2]). Indeed, the method employing the artifact allows the position of the zero delay time to be determined immediately in the course of measurements, without introducing additional changes into the experimental setup. The slow component of the coherent artifact is proportional to the correlation function of the laser pulse intensities and can be used for optimization of these intensities in the course of measurements.

[1] The authors are grateful to the referee for drawing their attention to the fact that the discrepancy between the two-level model and experiment in the sign of the effect is quite reasonable, since a field-induced decrease of the average population difference in the two-level system always leads to a decrease in the resonance additive to the permittivity and, hence, in the refection coefficient.

ACKNOWLEDGMENTS

This study was supported in part by the Russian Foundation for Basic Research (project nos. 02-02-17074 and 04-02-97204), by the German Physical Society (grant no. DE 567/9), and NATO program for joint scientific research (grant no. PST.GLG.978935).

REFERENCES

1. J. Shah, *Ultrafast Spectroscopy of Semiconductors and Semiconductor Nanostructures* (Springer, New York, 1996), Springer Series in Solid State Sciences, Vol. 115.
2. T. Dekorsy, G. C. Cho, and H. Kurz, in *Light Scattering in Solids VIII*, Ed. by M. Cardona and G. Güntherodt (Springer, Berlin, 2000), p. 169.
3. O. V. Misochko, Zh. Éksp. Teor. Fiz. **119**, 285 (2001) [JETP **92**, 246 (2001)].
4. P. Borri, F. Romstad, W. Langbein, *et al.*, Opt. Express **7**, 107 (2000).
5. H. Z. Zeiger, J. Vidal, T. K. Cheng, *et al.*, Phys. Rev. B **45**, 768 (1992).
6. *Optical Properties of Semiconductors: a Handbook*, Ed. by V. I. Gavrilenko, A. M. Grekhov, D. V. Korbutyak, and V. G. Litovchenko (Naukova Dumka, Kiev, 1987) [in Russian].
7. D. N. Klyshko, *Physical Principles of Quantum Electronics* (Nauka, Moscow, 1986) [in Russian].

Translated by P. Pozdeev

Journal of Experimental and Theoretical Physics, Vol. 100, No. 2, 2005, pp. 283–293.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 127, No. 2, 2005, pp. 320–330.

ATOMS, MOLECULES, OPTICS

Probe Field Spectroscopy in Two-Level Systems with Arbitrary Phase Memory in Collisions

A. I. Parkhomenko* and A. M. Shalagin**

Institute of Automation and Electrometry, Siberian Division, Russian Academy of Sciences, Universitetskiĭ pr. 1, Novosibirsk, 630090 Russia

**e-mail: par@iae.nsk.su*

***e-mail: shalagin@iae.nsk.su*

Received June 21, 2004

Abstract—The spectrum of weak probe field absorption (amplification) by two-level atoms experiencing collisions with buffer gas atoms in a strong resonance laser field is studied theoretically. Analysis is carried out for systems with a weak Doppler broadening under relatively mild constraints on the strong field intensity for the general case of an arbitrary change in the phase of the radiation-induced dipole moment in elastic collisions of gas particles. It is shown that, in spite of uniform broadening of the absorption line, the probe field spectrum exhibits a clearly manifested anisotropy to mutual orientation of the wavevectors of strong and probe radiation. It is found that the width of resonances in the probe field spectrum under definite conditions (that can easily be created in experiments) is proportional to the diffusion coefficient for atoms interacting with radiation. This fact can form the basis of the spectroscopic method for measuring the transport frequencies of collisions between particles absorbing radiation and buffer particles. It is shown that phase memory effects in collisions strongly modify the probe field spectrum both qualitatively and quantitatively. Simple operative formulas proposed for the probe field spectrum are convenient for experimental data processing. © *2005 Pleiades Publishing, Inc.*

1. INTRODUCTION

The probe field method developed in the 1970s is one of the most widely used and effective methods in nonlinear spectroscopy [1–3]. The essence of the method is that a weak (probe) field "sounds" the structure of atomic states perturbed by a strong field. The possibility of varying the frequency, polarization, and direction of propagation of the probe field renders this method a powerful tool for studying various relaxation processes.

The study of the interaction between the weak field and atoms in the strong resonance laser field is also important due to the possibility of noninversion amplification of radiation. The possibility of noninversion amplification of radiation was predicted by Rautian and Sobel'man [4], who considered the interaction between a probe field and two-level atoms perturbed by a strong laser field. It turned out that the probe field is not absorbed but amplified by the noninversion medium in a certain frequency range. In this case, nonlinear interference effects[1] are responsible for noninversion amplification [1]. The noninversion amplification effect by two-level atoms, which was predicted in [4], was subsequently studied in many theoretical works [5–11] and was observed experimentally in the rf range [12, 13] and in the optical spectral region [14, 15].

The interaction between the probe field and atoms was investigated by theorists under certain assumptions and using some approximations. Calculations were made either in the framework of perturbation theory in the strong laser radiation intensity or by disregarding atomic motion and collisions. At the same time, the results of contemporary experiments cannot be interpreted in the framework of perturbation theory in the strong field intensity, and atomic collisions should be consistently taken into account.

In this paper, we present a theoretical analysis of the spectrum of absorption (amplification) of a probe field by two-level atoms located in a strong resonance laser field and experiencing collisions with buffer gas atoms. The analysis is carried out under a relatively mild constraint imposed on the strong field intensity in the general case of random variation of the phase of the radiation-induced dipole moment (from complete breakdown to complete conservation) in elastic collisions of gas particles (phase-preserving collisions are known to lead to the Dicke effect of spectral line narrowing [1, 2]). We consider the case of interaction between atoms and radiation with a low rate selectivity, which corresponds to a weak Doppler broadening as compared to the sum of the collision rate and the radiation relaxation rate. This enabled us to apply a simple computational method based on the velocity-averaging procedure directly in the kinetic equations for the density matrix. This led to equations coinciding with the corresponding system of equations for stationary particles to within

[1] Coherent nonlinear phenomena in atoms associated with correlation of quantum states, which is caused by an external field, are known as nonlinear interference effects [1].

the redefinition of relaxation constants, which reflects the motion of particles in accordance with the diffusion law.

It may appear that allowance for the motion of atoms in the case of a weak Doppler broadening should only introduce small corrections to the probe field spectrum. It turns out, however, that the motion of atoms must necessarily be taken into account even in the case of a weak Doppler broadening since it leads to an unexpectedly sharp qualitative and quantitative (by several orders of magnitude) transformation of the probe field spectrum. The present paper is devoted to theoretical analysis of this unexpectedly sharp transformation of the probe field spectrum.

2. BASIC EQUATIONS

Let a two-level atom be in an external electromagnetic field $\mathscr{E}$ represented by a strong and a weak running wave which are in resonance with the *m*–*n* transition between the ground (*n*) and excited (*m*) energy levels:

$$\mathscr{E} = \frac{1}{2}\{\mathbf{E}\exp(i\mathbf{k}\cdot\mathbf{r} - i\omega t) + \mathbf{E}_\mu\exp(i\mathbf{k}_\mu\cdot\mathbf{r} - i\omega_\mu t - \varphi_0)\} + \text{c.c.} \quad (1)$$

The quantities with subscript μ characterize a weak (probe) field; ω, ω_μ and $\mathbf{k}$, $\mathbf{k}_\mu$ are the frequencies and wavevectors of the corresponding waves; and φ_0 is the phase shift of the probe field relative to the strong field. In the resonance approximation, the interaction between radiation and two-level particles is described by the following kinetic equations for density matrix [1]:

$$\left[\frac{d}{dt} + \Gamma_m\right]\rho_m(\mathbf{v}) = S[\rho_m(\mathbf{v})] - 2\mathrm{Re}\{i[G^* + G_\mu^*\exp(-i\varphi)]\rho_{mn}(\mathbf{v})\},$$
$$\frac{d}{dt}\rho_n(\mathbf{v}) = \Gamma_m\rho_m(\mathbf{v}) + S[\rho_n(\mathbf{v})] + 2\mathrm{Re}\{i[G^* + G_\mu^*\exp(-i\varphi)]\rho_{mn}(\mathbf{v})\}, \quad (2)$$
$$\left[\frac{d}{dt} + \frac{\Gamma_m}{2} - i(\Omega_0 - \mathbf{k}\cdot\mathbf{v})\right]\rho_{mn}(\mathbf{v}) = S[\rho_{mn}(\mathbf{v})] + i[G + G_\mu\exp(i\varphi)][\rho_n(\mathbf{v}) - \rho_m(\mathbf{v})],$$

where

$$\varphi = \mathbf{q}\cdot\mathbf{r} - \varepsilon t - \varphi_0, \quad \mathbf{q} = \mathbf{k}_\mu - \mathbf{k},$$
$$\varepsilon = \omega_\mu - \omega, \quad \Omega_0 = \omega - \omega_{mn}, \quad (3)$$
$$G = \frac{d_{mn}E}{2\hbar}, \quad G_\mu = \frac{d_{mn}E_\mu}{2\hbar}.$$

Here, $\rho_i(\mathbf{v})$ is the velocity distribution of the particles at the *i*th level ($i = m, n$); $S[\rho_i(\mathbf{v})]$ and $S[\rho_{mn}(\mathbf{v})]$ are collision integrals; Γ_m is the spontaneous decay rate for the excited level *m*; d_{mn} is the matrix element of the dipole moment for the *m*–*n* transition; and ω_{mn} is the *m*–*n* transition frequency.

For the collision integrals in Eqs. (2), we will use the strong collision model [1]

$$S[\rho_i(\mathbf{v})] = -\nu_i\rho_i(\mathbf{v}) + \nu_i\rho_i W(\mathbf{v}),$$
$$\rho_i \equiv \int\rho_i(\mathbf{v})d\mathbf{v}, \quad i = m, n,$$
$$S[\rho_{mn}(\mathbf{v})] = -\nu\rho_{mn}(\mathbf{v}) + \tilde{\nu}\rho_{mn}W(\mathbf{v}), \quad (4)$$
$$\rho_{mn} \equiv \int\rho_{mn}(\mathbf{v})d\mathbf{v},$$

where $W(\mathbf{v})$ is the Maxwellian velocity distribution, ν_m and ν_n are the frequencies of collisions of absorbing particles in the excited and ground states with buffer particles, and ν and $\tilde{\nu}$ are "nondiagonal" frequencies of "departure" and "arrival", which are complex-valued quantities in the general case. Quantities ν_i ($i = m, n$) are connected with diffusion coefficient D_i of particles in the *i*th state via the relation $D_i = v_T^2/2\nu_i$, where v_T is the most probable velocity of absorbing particles [16]. The case when phase memory is not preserved during collisions corresponds to

$$\tilde{\nu} = 0. \quad (5)$$

In the case of completely preserved phase memory during collisions, the frequencies ν and $\tilde{\nu}$ of departure and arrival are real-valued and identical [1]:

$$\tilde{\nu} = \nu = \nu_m = \nu_n. \quad (6)$$

To simplify the problem, we will confine our analysis to the case when the transport collision frequencies differ insignificantly,

$$\frac{|\nu_m - \nu_n|}{\nu_n} \ll 1. \quad (7)$$

Under this condition, collision frequencies ν_m and ν_n can be replaced by the average transport frequency

$$\bar{\nu} = \frac{\nu_m + \nu_n}{2}. \quad (8)$$

It should be noted that condition (7) imposed on collisions of any type (collisions preserving the phase of the radiation-induced dipole moment as well as those causing a strong phase mismatching). Indeed, the estimates obtained using the rigid sphere model show (see Conclusions) that phase memory effects can be manifested when the difference between transport collision fre-

quencies is small (under a few percent). On the other hand, condition (7) is satisfied even when collision frequencies differ by 10–15%; in this case, collisions are responsible for a fortiori strong phase mismatching in an atomic oscillator.

We seek a solution to Eqs. (2) in the form

$$\begin{aligned}\rho_i(\mathbf{v}) &= R_i(\mathbf{v}) + 2\mathrm{Re}[r_i(\mathbf{v})\exp(i\varphi)], \quad i = m, n,\\ \rho_{mn}(\mathbf{v}) &= R(\mathbf{v}) + r(\mathbf{v})\exp(i\varphi) + \tilde{r}(\mathbf{v})\exp(-i\varphi).\end{aligned} \tag{9}$$

Matrix elements $R_i(\mathbf{v})$ and $R(\mathbf{v})$ correspond to the solution with strong field G alone. Small corrections $r_i(\mathbf{v})$, $r(\mathbf{v})$, and $\tilde{r}(\mathbf{v})$ are due to the presence of probe field G_μ.

Under steady-state and spatially homogeneous conditions, in the first approximation in G_μ, system of equations (2) after substituting relation (9) and taking into account condition (7) splits into the following two subsystems:

$$\begin{aligned}(\Gamma_m + \bar{\nu})R_m(\mathbf{v}) &= \bar{\nu}R_m W(\mathbf{v}) - 2\mathrm{Re}[iG^*R(\mathbf{v})],\\ R_m(\mathbf{v}) + R_n(\mathbf{v}) &= (R_m + R_n)W(\mathbf{v}),\\ [\Gamma - i(\Omega - \mathbf{k}\cdot\mathbf{v})]R(\mathbf{v}) &= \tilde{\nu}RW(\mathbf{v}) + iG[R_n(\mathbf{v}) - R_m(\mathbf{v})]\end{aligned} \tag{10}$$

and

$$\begin{aligned}&[\Gamma_m + \bar{\nu} - i(\varepsilon - \mathbf{q}\cdot\mathbf{v})]r_m(\mathbf{v})\\ &= \bar{\nu}r_m W(\mathbf{v}) + i[G\tilde{r}^*(\mathbf{v}) + G_\mu R^*(\mathbf{v}) - G^*r(\mathbf{v})],\\ &[\Gamma - i(\Omega_\mu - \mathbf{k}_\mu\cdot\mathbf{v})]r(\mathbf{v}) = \tilde{\nu}rW(\mathbf{v})\\ &- 2iGr_m(\mathbf{v}) + iG_\mu[R_n(\mathbf{v}) - R_m(\mathbf{v})],\\ &\{\Gamma - i[\varepsilon - \Omega - (\mathbf{q} - \mathbf{k})\cdot\mathbf{v}]\}\tilde{r}^*(\mathbf{v})\\ &= \tilde{\nu}^*\tilde{r}^*W(\mathbf{v}) + 2iG^*r_m(\mathbf{v}).\end{aligned} \tag{11}$$

Here, the following notation has been introduced:

$$\begin{aligned}\Gamma &= \frac{\Gamma_m}{2} + \nu', \quad \Omega = \Omega_0 - \nu'',\\ \Omega_\mu &= \omega_\mu - \omega_{mn} - \nu''.\end{aligned} \tag{12}$$

Prime and double prime indicate the real and imaginary parts of a complex number, respectively. Quantities R, R_i, r, $\tilde{r}^*$, and r_i appearing in Eqs. (10) and (11) are integrals over velocities of quantities $R(\mathbf{v})$, $R_i(\mathbf{v})$, $r(\mathbf{v})$, $\tilde{r}^*(\mathbf{v})$, and $r_i(\mathbf{v})$, respectively ($R \equiv \int R(\mathbf{v})d\mathbf{v}$, etc.). In deriving system of equations (11), we used the relation $r_m(\mathbf{v}) + r_n(\mathbf{v}) = 0$, which is valid under condition (7).

In accordance with generally accepted rules, the probability of probe field absorption at frequency ω_μ (the number of radiation absorption events per unit time per absorbing atom) is defined as

$$\begin{aligned}P_\mu &\equiv -\frac{2}{N}\mathrm{Re}[iG_\mu^*\exp(-i\varphi)\rho_{mn}] = -\frac{2}{N}\\ &\times \mathrm{Re}\{iG_\mu^*[r + \tilde{r}\exp(-2i\varphi) + R\exp(-i\varphi)]\},\end{aligned} \tag{13}$$

where N is the concentration of absorbing particles. For $\omega_\mu \neq \omega$, oscillating terms $\tilde{r}\exp(-2i\varphi)$ and $R\exp(-i\varphi)$ on the right-hand side of Eq. (13) can be disregarded and the formula for the probe field absorption probability assumes the form

$$P_\mu = -\frac{2}{N}\mathrm{Re}[iG_\mu^* r]. \tag{14}$$

This formula can also be used instead of Eq. (13) for $\omega_\mu = \omega$ if we take into account the fact that phase shift φ_0 between the probe and strong waves fluctuates under real experimental conditions.

Thus, in compliance with the problem formulated above, we must determine quantity r from system of equations (10), (11).

3. THE METHOD OF PRELIMINARY VELOCITY AVERAGING

In the general case of an arbitrary relation between the uniform and Doppler widths of the absorption line, the solution of system of equations (10), (11) leads to a too complicated and cumbersome expression for the probe field absorption probability, which can be analyzed only using numerical methods. It turns out, however, that an approximate but very simple and effective approach to solving these equations in velocity-integrated quantities R, R_i, r, $\tilde{r}^*$, and r_i exists in the case of uniform broadening of absorption line, when the Doppler width kv_T is much smaller than the value of Γ,

$$\Gamma \gg kv_T. \tag{15}$$

This approach can be called the preliminary velocity-averaging method.

The proposed method can be described as follows. Equations (11) and the last equation in system (10) have the same structure and each of these equations can be written in the form

$$[\Gamma_a - i(\Omega_a - \mathbf{k}_a\cdot\mathbf{v})]f_a(\mathbf{v}) = \nu_a f_a W(\mathbf{v}) + Q_a(\mathbf{v}), \tag{16}$$

where $f_a \equiv \int f_a(\mathbf{v})d\mathbf{v}$; the quantities with subscript a are the corresponding parameters and functions (for example, bearing in mind the last equation in system (11), we must set in Eq. (16) $\Gamma_a = \Gamma$, $\Omega_a = \varepsilon - \Omega$, $\mathbf{k}_a = \mathbf{q} - \mathbf{k}$, $f_a(\mathbf{v}) = \tilde{r}^*(\mathbf{v})$, $\nu_a = \tilde{\nu}^*$, and $Q_a(\mathbf{v}) = 2iG^*r_m(\mathbf{v})$).

For integrated quantity f_a appearing in Eq. (16), we obtain the equation

$$(1-\nu_a I_a) f_a = J_a,$$

$$I_a = \int \frac{W(\mathbf{v})d\mathbf{v}}{\Gamma_a - i(\Omega_a - \mathbf{k}_a\cdot\mathbf{v})}, \qquad (17)$$

$$J_a = \int \frac{Q_a(\mathbf{v})d\mathbf{v}}{\Gamma_a - i(\Omega_a - \mathbf{k}_a\cdot\mathbf{v})}.$$

Taking into account the fact that the relation $\Gamma_a \gg k_a v_T$ holds under uniform broadening condition (15), we can calculate integral I_a using the well-known asymptotic expansion [17]

$$\int_{-\infty}^{\infty} \frac{\exp(-t^2)}{z-t}dt = \frac{\sqrt{\pi}z}{z^2-1/2}, \quad |z| \gg 1, \qquad (18)$$

which leads to

$$I_a = \frac{\Gamma_a - i\Omega_a}{(\Gamma_a - i\Omega_a)^2 + \dfrac{(k_a v_T)^2}{2}}. \qquad (19)$$

Let us now calculate integral J_a. The integrand $Q_a(\mathbf{v})$ rapidly decreases for $|\mathbf{v}| > v_T$; consequently, by virtue of condition $\Gamma_a \gg k_a v_T$, we disregard the Doppler shift $\mathbf{k}_a\cdot\mathbf{v}$ in the integrand of the expression for J_a, which immediately leads to $J_a = Q_a/(\Gamma_a - i\Omega_a)$. Considering that function $Q_a(\mathbf{v})$ exhibits a weak asymmetry (under uniform broadening conditions (15) for the absorption line, the interaction of atoms with radiation exhibits a low velocity selectivity; consequently, the velocity distribution for atoms in states m and n differs from the Maxwellian distribution only slightly), we conclude that small corrections on the order of $(k_a v_T)^2/|\Gamma_a - i\Omega_a|^2 \ll 1$ are discarded in the expression derived for integral J_a. It should be noted that, in spite of the identical structure of integrals I_a and J_a, we cannot disregard the Doppler shift $\mathbf{k}_a\cdot\mathbf{v}$ in the integrand of the expression for I_a in view of the strong sensitivity of factor $1-\nu_a I_a$ in Eq. (17) to the value of $k_a v_T$ (for $\Gamma_a \approx \nu_a$).

Ultimately, we derive from Eqs. (10) and (11) combined with Eqs. (17) and (19) the following equations, which are valid to within small corrections on the order of $(k_a v_T)^2/|\Gamma_a - i\Omega_a|^2 \ll 1$:

$$\Gamma_m R_m = -2\mathrm{Re}[iG^*R],$$

$$(\Gamma_k - i\Omega_1)R = iG(R_n - R_m),$$

$$(\Gamma_q - i\varepsilon)r_m = i(G\tilde{r}^* + G_\mu R^* - G^* r), \qquad (20)$$

$$(\Gamma_\mu - i\Omega_{1\mu})r = -2iGr_m + iG_\mu(R_n - R_m),$$

$$[\Gamma_p - i(\varepsilon - \Omega_1)]\tilde{r}^* = 2iG^* r_m.$$

Here, the following notation is introduced:

$$\Gamma_q = \Gamma_m + \frac{(qv_T)^2}{2(\Gamma_m + \bar{\nu} - i\varepsilon)},$$

$$\Gamma_k = \Gamma_1 + \frac{(kv_T)^2}{2(\Gamma - i\Omega)},$$

$$\Gamma_\mu = \Gamma_1 + \frac{(k_\mu v_T)^2}{2(\Gamma - i\Omega_\mu)},$$

$$\Gamma_p = \Gamma_1 + \frac{(k_1 v_T)^2}{2[\Gamma - i(\varepsilon - \Omega)]}, \qquad (21)$$

$$\Gamma_1 = \frac{\Gamma_m}{2} + (\nu - \tilde{\nu})',$$

$$k_1 = |2\mathbf{k} - \mathbf{k}_\mu|, \quad \Omega_1 = \Omega + \tilde{\nu}'',$$

$$\Omega_{1\mu} = \Omega_\mu + \tilde{\nu}''.$$

System of equations (20) should be supplemented with the normalization condition $R_m + R_n = N$ (R_m and R_n are the partial concentrations of absorbing particles at levels m and n and N is the total concentration of absorbing particles).

System of algebraic equations (20) formally coincides with that for particles at rest (see, for example, [1–3]). The only difference is that the relaxation constants are generalized to take into account the motion of particles in accordance with the diffusion law. Solving Eqs. (20) in standard way, we arrive at the following expression for probability P_μ of probe field absorption:

$$P_\mu = 2|G_\mu|^2 \frac{R_n - R_m}{N} B_\mu,$$

$$B_\mu = \mathrm{Re}\frac{(\Gamma_q - i\varepsilon)[\Gamma_p - i(\varepsilon - \Omega_1)] - 2|G|^2 \dfrac{\Gamma_p - \Gamma_k^* - i\varepsilon}{\Gamma_k^* + i\Omega_1}}{(\Gamma_q - i\varepsilon)(\Gamma_\mu - i\Omega_{1\mu})[\Gamma_p - i(\varepsilon - \Omega_1)] + 2|G|^2(\Gamma_\mu + \Gamma_p - 2i\varepsilon)}. \qquad (22)$$

Quantity $(R_n - R_m)/N$ appearing in this equation is the relative difference between the populations of the ground and excited states,

$$\frac{R_n - R_m}{N} = \left(1 + \frac{4|G|^2}{\Gamma_m}\mathrm{Re}\frac{1}{\Gamma_k - i\Omega_1}\right)^{-1}. \quad (23)$$

Factor B_μ in formula (22) describes the shape of the probe field spectrum.

Formula (22) for the absorption probability is a natural generalization of the corresponding formula for stationary particles (see, for example, problem 11 in [1]) and coincides with this formula after the substitution $\Gamma_k, \Gamma_\mu, \Gamma_p \longrightarrow \Gamma_1$ and $\Gamma_q \longrightarrow \Gamma_m$. Thus, the probe field spectrum both for stationary and for moving particles is described by the same general formula (22). The motion of atoms is manifested only in a change in relaxation constants Γ_i, $i = k, q, \mu, p$, due to corrections reflecting the diffusion law of motion.

Direct comparison of the results of numerical calculations of dependence $P_\mu(\Omega_\mu)$ based on formula (22) and on the exact formula derived by solving the system of equations (10), (11) (we do not write here the exact formula since it is too cumbersome) shows that formula (22) successfully describes the probe field spectrum for any radiation intensity in the case of a strong phase mismatching of an atomic oscillator as a result of collisions. In the case of phase memory preserving collisions, formula (22) is valid for not very high radiation intensity, such that

$$\frac{2|G|^2}{|(\Gamma_m + \bar{\nu} - i\varepsilon)(\Gamma - i\Omega_\mu)|} \ll 1,$$
$$\frac{2|G|^2}{|(\Gamma_m + \bar{\nu} - i\varepsilon)[\Gamma - i(\varepsilon - \Omega)]|} \ll 1. \quad (24)$$

It should be noted that these conditions depend on detuning Ω of the radiation frequency and can be satisfied even for $|G| \gtrsim \Gamma$ for large values of detuning ($|\Omega| \gg \Gamma$).

4. ANALYSIS OF THE PROBE FIELD SPECTRUM

Let us analyze the effect of collisions on the probe field spectrum. Let us first investigate the behavior of the quantities introduced in Eqs. (12) and (21) and appearing in the final expression for the probe field absorption probability. In the case of complete conservation of phase memory during collisions, in accordance with relations (6) and (8), we have

$$\Gamma = \frac{\Gamma_m}{2} + \bar{\nu}, \quad \Gamma_1 = \frac{\Gamma_m}{2},$$
$$\Omega_1 = \Omega, \quad \Omega_{1\mu} = \Omega_\mu. \quad (25)$$

Thus, the following relation holds for completely preserved phase memory and a fairly high collision frequency (for $\bar{\nu} \gg \Gamma_m/2$):

$$\Gamma_1 \ll \Gamma. \quad (26)$$

In the absence of phase memory, it follows from relations (5), (12), and (21) that

$$\Gamma_1 = \Gamma, \quad \Omega_1 = \Omega, \quad \Omega_{1\mu} = \Omega_\mu. \quad (27)$$

For partial phase memory, the relation $\Gamma_1 < \Gamma$ holds and it may even happen that $\Gamma_1 \ll \Gamma$. Radiation frequency detunings Ω, Ω_1 and Ω_μ, $\Omega_{1\mu}$ may be different on account of collision-induced shifts. However, we will disregard this difference in subsequent analysis and assume that conditions (25) and (27) hold.

Under the same conditions, by virtue of relation (15), quantities Γ_i (21) appearing in formula (22) for the probe field absorption probability can be treated as real-valued. In subsequent analysis, it will be implied that the corresponding substitution

$$\Gamma_i \longrightarrow \mathrm{Re}\Gamma_i, \quad (28)$$

$i = k, q, \mu, p$, has already been carried out.

Let us now consider the probe field spectrum for various values of the parameters of the problem and write simple analytic relations for some limiting cases.

4.1. Probe Field Spectrum under Conditions of Its Splitting into Three Components

It is well known (see, for example, [1–3]) that the number and position of the absorption (amplification) spectrum components of the probe field are determined by the effect of field-induced level splitting (the Autler–Townes, or dynamic Stark effect). In the problem considered here, three spectral components must exist in the general case: for $\Omega_\mu = \Omega \pm \sqrt{\Omega^2 + 4|G|^2}$ and for $\Omega_\mu = \Omega$.

We begin our analysis with the case when these components are spectrally resolved. This occurs when the following condition is satisfied:

$$\Omega_R \equiv \sqrt{4|G|^2 + \Omega^2} \gg \Gamma_\mu, \Gamma_p, \Gamma_k. \quad (29)$$

In this case, formula (22) can be transformed to the following expression, in which the above-mentioned three spectral components appear in explicit form:

$$B_\mu = \mathrm{Re}\left\{\frac{1}{\Gamma_k + i\Omega}\left[\frac{iA^{(-)}}{\Gamma_{\mathrm{ef}}^{(-)} - i(\Omega_\mu - \Omega_\mu^{(-)})} + \frac{iA^{(+)}}{\Gamma_{\mathrm{ef}}^{(+)} - i(\Omega_\mu - \Omega_\mu^{(+)})} + \frac{A^{(0)}}{\Gamma_{\mathrm{ef}}^{(0)} - i(\Omega_\mu - \Omega)}\right]\right\}; \quad (30)$$

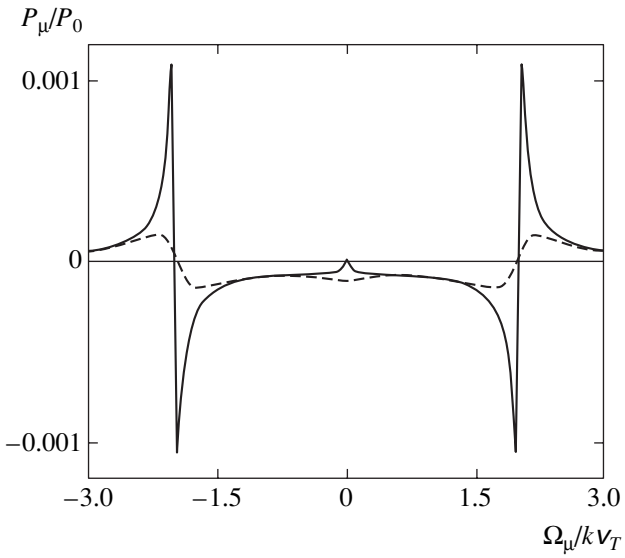

Fig. 1. Dependence of probe field absorption probability P_μ on frequency detuning Ω_μ for collisions preserving the phase of radiation-induced dipole moment; $\Omega = 0$, $|G|/kv_T = 1$, $\bar{\nu}/kv_T = 10$, $\Gamma_m/kv_T = 10^{-2}$; the solid and dashed curves correspond to $\mathbf{k} \uparrow\uparrow \mathbf{k}_\mu$ and $\mathbf{k} \uparrow\downarrow \mathbf{k}_\mu$, respectively.

here, we have

$$\Gamma_{ef}^{(\pm)} = \frac{\Gamma_\mu + \Gamma_p}{2} \pm \frac{\Omega(\Gamma_p - \Gamma_\mu)}{2\Omega_R} - \frac{|G|^2(\Gamma_\mu + \Gamma_p - 2\Gamma_q)}{\Omega_R^2},$$

$$\Gamma_{ef}^{(0)} = \frac{2|G|^2(\Gamma_\mu + \Gamma_p) + \Omega^2\Gamma_q}{\Omega_R^2},$$

$$A^{(\pm)} = \frac{\Omega}{2\Omega_R}\left(\Omega_R \mp \Omega \mp \frac{2|G|^2}{\Omega}\right), \qquad (31)$$

$$A^{(0)} = \frac{2|G|^2}{\Omega_R^4}[2|G|^2(2\Gamma_k + \Gamma_\mu - \Gamma_p) + \Omega^2(\Gamma_\mu + \Gamma_k - \Gamma_q)],$$

$$\Omega_\mu^{(\pm)} = \Omega \pm \Omega_R.$$

The central spectral component lies in the vicinity of $\Omega_\mu = \Omega$. Its half-width is determined by quantity $\Gamma_{ef}^{(0)}$. Two side component with half-widths $\Gamma_{ef}^{(\pm)}$ lie in the vicinity of $\Omega_\mu = \Omega_\mu^{(\pm)}$ so that the separation between them is equal to $2\Omega_R$. Quantities $A^{(0)}$ and $A^{(\pm)}$ characterize the amplitudes of the corresponding components. Both resonance widths and amplitudes in the probe field spectrum are functions of intensity and radiation frequency detuning as well as mutual orientation of wavevectors $\mathbf{k}$ and $\mathbf{k}_\mu$, collision frequency, and the extent of conservation of phase memory during collisions. Let us consider various special cases for expression (30).

4.1.1. High radiation intensity. For a high radiation intensity ($2|G| \gg |\Omega|, \Gamma_\mu, \Gamma_p, \Gamma_k$), formula (30) implies that the probe field spectrum is described by the formula

$$B_\mu = \mathrm{Re}\left\{\frac{1}{2(\Gamma_k + i\Omega)}\left[\frac{i|G|}{\Gamma_{ef} - i(\Omega_\mu + 2|G|)} - \frac{i|G|}{\Gamma_{ef} - i(\Omega_\mu - 2|G|)} + \frac{A_0}{\Gamma_{ef}^{(0)} - i(\Omega_\mu - \Omega)}\right]\right\}, \qquad (32)$$

where

$$\Gamma_{ef} = \frac{2\Gamma_q + \Gamma_\mu + \Gamma_p}{4}, \quad \Gamma_{ef}^{(0)} = \frac{\Gamma_\mu + \Gamma_p}{2}, \quad A_0 = \Gamma_k + \frac{\Gamma_\mu - \Gamma_p}{2}. \qquad (33)$$

The shape of the spectrum substantially depends on the relation between $|\Omega|$ and Γ_k. Let us consider two different cases for expression (32): exact resonance of the strong field ($\Omega = 0$) and substantial detuning of the strong field frequency from the absorption line center ($|\Omega| \gg \Gamma_k$).

Exact resonance for the strong field. Let us first consider the case when $\Omega = 0$. In accordance with relation (32), profile $P_\mu(\Omega_\mu)$ in this case is symmetric relative to $\Omega_\mu = 0$ and has two disperse side components located near $\Omega_\mu = \pm 2|G|$ (Fig. 1). The central Lorentzian component is located near $\Omega_\mu = 0$ and its amplitude is many time smaller than that of the side components. The widths of the components depend on the extent of conservation of phase memory during collisions. As the unit of measurements in all figures, we use the quantity

$$P_0 = \frac{2|G_\mu|^2}{\Gamma_1 + \gamma}, \quad \gamma = \frac{(kv_T)^2}{2\Gamma},$$

which is equal to the absorption probability for probe radiation at the line center in the absence of the strong field.

In the absence of phase memory (it should be recalled that $\Gamma_1 = \Gamma$ in this case), the broadening of the side components is determined by quantity $\Gamma_{ef} = \Gamma/2$, while the half-width of the central component is $\Gamma_{ef}^{(0)} = \Gamma$.

In the case when phase memory is completely preserved during collisions, in expression (32) we have[2]

$$\Gamma_{\text{ef}} = \frac{3\Gamma_m}{4} + (5k^2 - 4\mathbf{k}\cdot\mathbf{k}_\mu)\frac{D}{2},$$
$$\Gamma_{\text{ef}}^{(0)} = \frac{\Gamma_m}{2} + (3k^2 - 2\mathbf{k}\cdot\mathbf{k}_\mu)D, \quad (34)$$
$$A_0 = \frac{\Gamma_m}{2} + (2\mathbf{k}\cdot\mathbf{k}_\mu - k^2)D,$$
$$\Gamma_k = \frac{\Gamma_m}{2} + k^2 D,$$

where $D = v_T^2/2\bar{\nu}$ is the diffusion coefficient for particles interacting with radiation. In accordance with relations (34), the widths and amplitudes of resonances in spectrum $P_\mu(\Omega_\mu)$ depend on the diffusion coefficient of particles and are anisotropic to mutual orientation of the wavevectors of the strong and probe radiation (see Fig. 1).

Anisotropy is manifested most clearly for $\Gamma_m \ll k^2 D$. In the case of unidirectional waves (we will henceforth use the notation $\mathbf{k} \uparrow\uparrow \mathbf{k}_\mu$), for quantities (34) we have $\Gamma_{\text{ef}} = k^2 D/2$, $\Gamma_{\text{ef}}^{(0)} = A^0 = k^2 D$. In the case of counterpropagating waves ($\mathbf{k} \Uparrow \mathbf{k}_\mu$), we have $\Gamma_{\text{ef}} = 9k^2 D/2$, $\Gamma_{\text{ef}}^{(0)} = 5k^2 D$, and $A_0 = -3k^2 D$. Thus, for $\mathbf{k} \Uparrow \mathbf{k}_\mu$, the width of the side dispersive resonances increase by a factor of nine as compared to the case when $\mathbf{k} \uparrow\uparrow \mathbf{k}_\mu$, while their amplitudes decrease by a factor of nine and the width of the central Lorentzian resonance decreases by a factor of five, its amplitude "reversing its sign" (quantity A_0 changes its sign).

$|\Omega| \gg \Gamma_k$. In this case (but still for $|\Omega| \ll 2|G|$), it follows from Eq. (32) that the line profile $P_\mu(\Omega_\mu)$ is asymmetric relative to point $\Omega_\mu = \Omega$. The side components are Lorentzian in shape, while the central component is dispersive (Fig. 2). Near the side component located in the vicinity of $\Omega_\mu = 2|G|\,\text{sgn}\,\Omega$, the absorption probability $P_\mu(\Omega_\mu) < 0$ (B_μ is negative), which indicates the amplification of the probe field. Near the other side component located in the vicinity of $\Omega_\mu = -2|G|\,\text{sgn}\,\Omega$, the probe field is absorbed ($P_\mu(\Omega_\mu) > 0$). In the case of completely preserved phase memory, the probe field spectrum is strongly anisotropic relative to mutual orientation of wavevectors $\mathbf{k}$ and $\mathbf{k}_\mu$ (see Fig. 2). Upon variation of the angle between vectors $\mathbf{k}$ and $\mathbf{k}_\mu$, the widths and amplitudes of all the three resonances

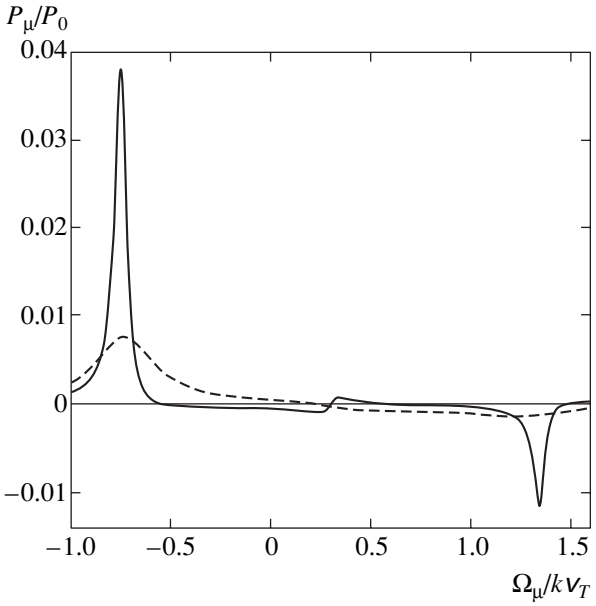

Fig. 2. Dependence $P_\mu(\Omega_\mu)$ for collisions preserving phase memory; $\Omega/kv_T = 0.3$, $|G|/kv_T = 0.5$, $\bar{\nu}/kv_T = 10$, $\Gamma_m/kv_T = 10^{-2}$; the solid and dashed curves correspond to $\mathbf{k} \uparrow\uparrow \mathbf{k}_\mu$ and $\mathbf{k} \uparrow\downarrow \mathbf{k}_\mu$, respectively.

behave exactly in the same way as in the case $\Omega = 0$ considered above.

4.1.2. Large detuning of strong field frequency. In the limiting case of large detuning of strong field frequency ($|\Omega| \gg 2|G|, \Gamma_\mu, \Gamma_p, \Gamma_k$), formula (30) implies that the probe field spectrum is described by the formula

$$B_\mu = \text{Re}\left\{\frac{1}{\Gamma_\mu - i\Omega_\mu} - \frac{|G|^4}{\Omega^4[\Gamma_p - i(\Omega_\mu - 2\Omega)]} - \frac{2i|G|^2(\Gamma_\mu + \Gamma_k - \Gamma_q)}{\Omega^3[\Gamma_q - i(\Omega_\mu - \Omega)]}\right\}. \quad (35)$$

The spectrum is asymmetric relative to point $\Omega_\mu = \Omega$ and has two side components with half-widths Γ_μ and Γ_p as well as the dispersive central component whose width is determined by quantity Γ_q.

Half-width Γ_μ of the resonance in the vicinity of $\Omega_\mu = 0$ does not depend on mutual orientation of wavevectors $\mathbf{k}$ and $\mathbf{k}_\mu$. In the case of collisions leading to a strong phase mismatch in an atomic oscillator (for $\Gamma_1 \gg (k_\mu v_T)^2/2\Gamma$), half-width $\Gamma_1 \approx \Gamma$. In the case of collisions preserving the atomic oscillator phase ($\Gamma_1 = \Gamma_m/2$), the resonance half-width is given by

$$\Gamma_\mu = \frac{\Gamma_m}{2} + k^2 D. \quad (36)$$

[2] The difference between the moduli of wavevectors $\mathbf{k}_\mu$ and $\mathbf{k}$ is small as compared to k: $|\mathbf{k}_\mu - \mathbf{k}| = k|\varepsilon|/\omega \ll k$.

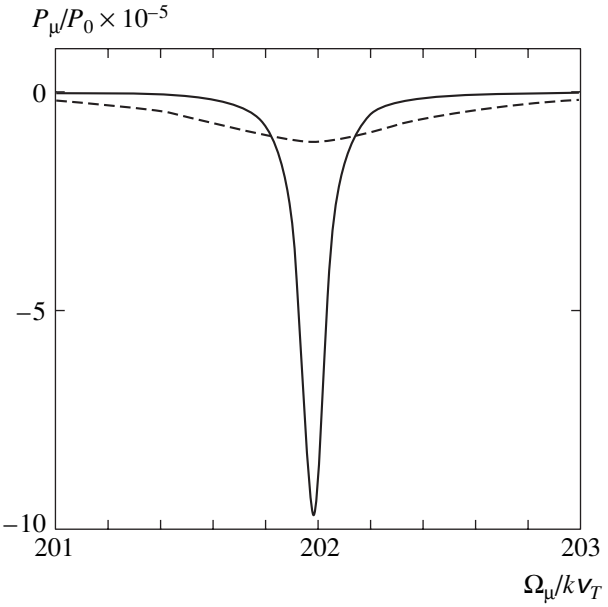

Fig. 3. Dependence $P_\mu(\Omega_\mu)$ in the vicinity of $\Omega_\mu = 2\Omega$ for collisions preserving phase memory; $\Omega/kv_T = 100$, $|G|/kv_T = 10$, $\bar{\nu}/kv_T = 10$, $\Gamma_m/kv_T = 10^{-2}$; the solid and dashed curves correspond to $\mathbf{k} \uparrow\uparrow \mathbf{k}_\mu$ and $\mathbf{k} \uparrow\downarrow \mathbf{k}_\mu$, respectively.

If, in addition, $\Gamma_m \ll 2k^2D$, the resonance half-width is determined only by the diffusion coefficient ($\Gamma_\mu = k^2D$).

The resonance in the vicinity of $\Omega_\mu = 2\Omega$ has half-width Γ_p. In the absence of phase memory in collisions, we have $\Gamma_p = \Gamma_1 = \Gamma$. When phase memory is completely preserved during collisions, the resonance half-width is determined by mutual orientation of wavevectors $\mathbf{k}$ and $\mathbf{k}_\mu$ and diffusion coefficient D of particles interacting with radiation:

$$\Gamma_p = \frac{\Gamma_m}{2} + (5k^2 - 4\mathbf{k}\cdot\mathbf{k}_\mu)D. \qquad (37)$$

If $\Gamma_m \ll 2k^2D$, we have $\Gamma_p = k^2D$ for unidirectional waves; in the case of counterpropagating waves, the resonance width increases by a factor of nine ($\Gamma_p = 9k^2D$; Fig. 3).

In the vicinity of resonance $\Omega_\mu = 2\Omega$, amplification of the probe field can take place. It turns out that a considerable contribution to the operation of the probe field in this region comes from the wing of the resonance line, which is described by the first term in expression (35). This contribution is equal to $\Gamma_\mu/4\Omega^2$ (the contribution from the third term is small). Consequently, in this region, we have

$$B_\mu = \frac{\Gamma_\mu}{4\Omega^2}\left\{1 - \frac{4|G|^4\Gamma_p}{\Omega^2\Gamma_\mu[\Gamma_p^2 + (\Omega_\mu - 2\Omega)^2]}\right\},$$

$$\Gamma_p = \Gamma_1 + \frac{(k_1 v_T)^2}{2\Gamma}, \quad \Gamma_\mu = \Gamma_1 + \frac{\gamma\Gamma^2}{\Gamma^2 + 4\Omega^2}, \qquad (38)$$

$$\gamma = \frac{(kv_T)^2}{2\Gamma}.$$

For a high radiation intensity ($4|G|^4 > \Omega^2\Gamma_\mu\Gamma_p$), the value of B_μ becomes negative, which corresponds to amplification of probe radiation in the vicinity of $\Omega_\mu = 2\Omega$.

In the case of weak phase mismatching of the atomic oscillator during collisions and for a small rate of spontaneous decay of the excited level, the condition $\Gamma_1 < \gamma\Gamma^2/4\Omega^2$ can be satisfied. In this case, an interesting situation takes place: the condition for the emergence of negative values of B_μ does not depend on strong field frequency detuning Ω in the region

$$kv_T\sqrt{\frac{\Gamma}{8\Gamma_1}} \gg |\Omega| \gg \frac{\Gamma}{2}.$$

Figure 3 illustrates the possibility of amplification of probe radiation in the far wing of the line in the vicinity of $\Omega_\mu = 2\Omega$.

The third term in formula (35) describes the central dispersive resonance in the vicinity of $\Omega_\mu = \Omega$. The resonance width is determined by quantity Γ_q, which is defined as

$$\Gamma_q = \Gamma_m + q^2D = \Gamma_m + 2(k^2 - \mathbf{k}\cdot\mathbf{k}_\mu)D \qquad (39)$$

on account of expression (28). The value of Γ_q depends on the diffusion coefficient of particles and is anisotropic to mutual orientation of the wavevectors of strong and probe radiation.

The following important circumstance is worth noting. Even in the absence of phase memory for an atomic oscillator, the spectrum of the probe field contains a line with "diffusion" broadening typical of the Dicke effect, which is narrowed upon an increase in the buffer gas pressure. This is due to the fact that the strong and probe fields induce oscillations (beats) of population densities (see formula (9) and Eqs. (11)). On the one hand, these beats are sensitive to the Doppler effect; on the other hand, their phase is not seriously affected by collisions since the populations of the levels remain unchanged during collisions. Thus, we are dealing here with an almost complete analog of the Dicke effect leading to the emergence of the corresponding spectral structure. This structure exhibits a clearly manifested anisotropy (Fig. 4).

The resonance amplitude is proportional to the quantity

$$\Gamma_\mu + \Gamma_k - \Gamma_q = 2\Gamma_1 - \Gamma_m + 2\mathbf{k}\cdot\mathbf{k}_\mu D. \qquad (40)$$

This expression implies that the resonance is of a purely collision origin (the radiation relaxation constants in

Eq. (40) are compensated). Under uniform broadening conditions (15), the relation $\Gamma \gg \mathbf{k} \cdot \mathbf{k}_\mu D$ always holds; hence, it can be concluded that phase memory effects "quench" the central resonance.

4.2. Unsplit Probe Field Spectrum for $\Omega = 0$

Let us consider the probe field spectrum at exact resonance for the strong field ($\Omega = 0$) in the case when all three components overlap (i.e., for a not very high radiation intensity).

The most interesting case is that when collisions cause a considerable phase breakdown in the atomic oscillator so that $\Gamma_1 \gg (k_1 v_T)^2/2\Gamma$ and we can assume that $\Gamma_k = \Gamma_\mu = \Gamma_p = \Gamma_1$. In this case, for a moderate radiation intensity (for $2|G| \ll \Gamma_1$), it follows from relation (22) that the probe field spectrum is described by the formula

$$\beta_\mu = \mathrm{Re}\left\{\frac{1}{\Gamma_1 - i\Omega_\mu} - \frac{4|G|^2}{\Gamma_1^2\left(\Gamma_q + \dfrac{4|G|^2}{\Gamma_1} - i\Omega_\mu\right)}\right\}, \quad (41)$$

where Γ_q satisfies relation (39) under high pressures of the buffer gas ($\bar{\nu} \gg \Gamma_m$).

In accordance with formula (41), the probe field spectrum contains two Lorentzian profiles summed with opposite signs. The first term in the braces describes a Lorentzian absorption line with half-width Γ_1, which is typical in the absence of a strong field. The second term is due to the above-mentioned beats in population densities and described a dip with half-width $\Gamma_q + 4|G|^2/\Gamma_1$ against the background of the Lorentzian profile with half-width Γ_1, which is associated with the first term. The dip width is anisotropic to mutual orientation of the wavevectors of the strong and probe radiation. The anisotropy is manifested especially clearly for a low radiation intensity ($4|G|^2 \ll \Gamma_1\Gamma_q$), when the dip half-width is equal to Γ_q. In this case, the dip width for unidirectional waves is determined only by the radiative decay rate of the excited state: $\Gamma_q = \Gamma_m$. In the case of counterpropagating waves, the dip half-width is determined by diffusion coefficient D of absorbing particles: $\Gamma_q = \Gamma_m + 4k^2D$. The relative depth of the dip is $4|G|^2/\Gamma_1\Gamma_q$ for a low radiation intensity.

Figure 5 illustrates anisotropy in the dip width relative to mutual orientation of wavevectors $\mathbf{k}_\mu$ and $\mathbf{k}$.

The special case corresponding to formula (41) is the most promising for studying the effects of population beats.

With increasing radiation intensity, the dip widths for unidirectional and counterpropagating waves virtually level out, while their depths may still differ substantially. As the intensity increases further (for $4|G|^2 \gg \Gamma_1\Gamma_q$), anisotropy in the probe field spectrum vanishes.

Analysis shows that, in the case of completely preserved phase memory, the probe field spectrum exhibits clearly manifested anisotropy to mutual orientation of wavevectors $\mathbf{k}_\mu$ and $\mathbf{k}$ for any radiation intensity.

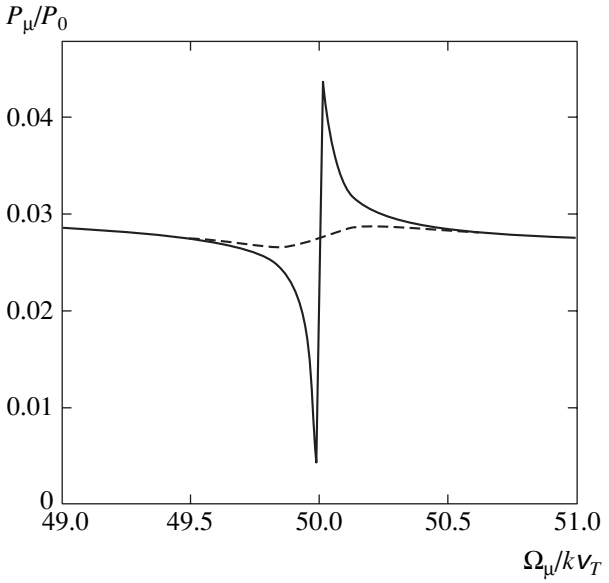

Fig. 4. Dependence $P_\mu(\Omega_\mu)$ for collisions that do not preserve phase memory; $\Omega/kv_T = 50$, $|G|/kv_T = 0.5$, $\bar{\nu}/kv_T = 10$, $\Gamma_m/kv_T = 10^{-2}$; the solid and dashed curves correspond to $\mathbf{k} \uparrow\uparrow \mathbf{k}_\mu$ and $\mathbf{k} \uparrow\downarrow \mathbf{k}_\mu$, respectively.

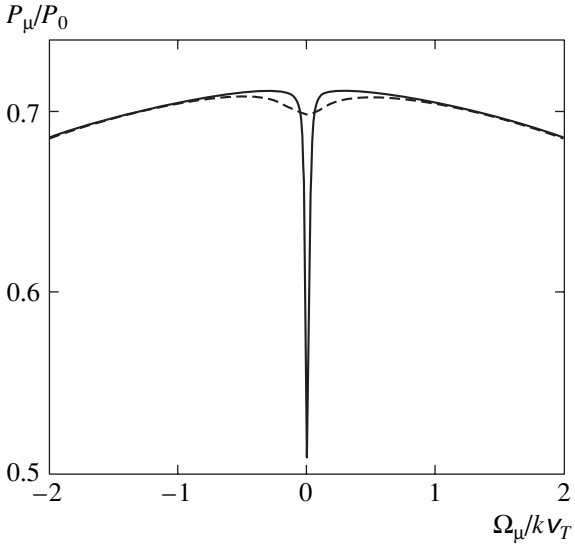

Fig. 5. Dependence $P_\mu(\Omega_\mu)$ for collisions that do not preserve phase memory; $\Omega = 0$, $|G|/kv_T = 0.1$, $\bar{\nu}/kv_T = 10$, $\Gamma_m/kv_T = 10^{-2}$; the solid and dashed curves correspond to $\mathbf{k} \uparrow\uparrow \mathbf{k}_\mu$ and $\mathbf{k} \uparrow\downarrow \mathbf{k}_\mu$, respectively.

5. CONCLUSIONS

We have theoretically investigated the spectrum of absorption (amplification) of a weak probing field by two-level atoms located in a strong resonant laser field and colliding with buffer gas atoms. The analysis was carried out for uniform broadening of the absorption line ($\Gamma \gg kv_T$) for high buffer gas pressures, when the collision frequency is high as compared to the Doppler width of the absorption line. The problem has been solved under a relatively mild restriction imposed on the strong field intensity and under the conditions of an arbitrary variation of the radiation-induced dipole moment phase (from complete breakdown to complete conservation) during elastic collisions of particles.

The fact that the probe field spectrum exhibits a clearly manifested anisotropy to mutual orientation of the wavevectors of the strong and probe fields in spite of uniform broadening of the absorption line appears as unexpected. In addition, the width and amplitude of resonances in the probe field spectrum may depend on diffusion coefficient $D = v_T^2/2\bar{\nu}$ of radiation-absorbing particles (both in the absence of phase memory and in the case of its conservation). This circumstance may serve as the basis of a spectroscopic method for measuring the diffusion coefficient (or transport frequency $\bar{\nu}$ of collisions between absorbing and buffer particles).

It is found that the phase memory effects lead to a very strong qualitative and quantitative change in the probe field spectrum.

It should be noted that the phase memory effects cannot be treated as exotic. Let us consider the conditions for their emergence.

The phase memory effects strongly affect the probe field spectrum in the case when the relation $\Gamma_1/\Gamma \ll 1$ holds for the parameters introduced in Eqs. (12) and (21). This is equivalent to the condition of weak variation of the potential of interaction between excited and unexcited resonance particles and buffer particles. In the hard sphere model, this condition can be presented in the form of a constraint imposed on the relative difference between the transport frequencies of collisions between particles in excited states and the ground state [18]:

$$\frac{|\nu_m - \nu_n|}{\nu_n} \ll \frac{2\lambda_D}{a}, \qquad (42)$$

where $\lambda_D = \hbar/\mu u_T$ is the de Broglie wavelength for the most probable velocity u_T of relative motion, μ is he reduced mass of colliding particle, and a is the radius of a spherical particle. For $\mu = 20$ amu, $T = 500$ K, and $a = 5 \times 10^{-8}$ cm, condition (42) leads to

$$\frac{|\nu_m - \nu_n|}{\nu_n} \ll 2 \times 10^{-2}. \qquad (43)$$

The typical value of parameter $|\nu_m - \nu_n|/\nu_n = 10$–$50\%$ for atoms with an excited outer shell [19, 20], and the conditions for observing the phase memory effects do not hold. However, in transitions associated with the excitation of electrons of inner shells (atoms of rare-earth elements), the phase memory effects can be manifested in view of the screening action of outer filled shells [21, 22]. Phase memory is preserved almost completely during collisions involving microwave transitions between energy levels of the hyperfine structure of the ground state of alkali metal atoms in the atmosphere of inert buffer gases [23, 24]. For this reason, inert buffer gases are used in alkali metal vapor masers for eliminating Doppler broadening [24], and one of the first observations of the Dicke narrowing was carried out on the transition between the states in the hyperfine structure of the cesium atom [23]. In some molecules, the phase memory effects can also be manifested in vibration–rotation transitions within the same electron state (see, for example, experimental work [25]). A strong manifestation of the phase memory effects should be expected for ions colliding with charged buffer particles since scattering of particles in this case is due to the Coulomb interaction and is weakly sensitive to a change in the atomic state of ions during their excitation.

ACKNOWLEDGMENTS

This study was financed by the Russian Foundation for Basic Research (project no. 04-02-16771), the program “Optical Spectroscopy and Frequency Standards” of the Physics Department, Russian Academy of Sciences, and the program “Universities of Russia.”

REFERENCES

1. S. G. Rautian, G. I. Smirnov, and A. M. Shalagin, *Nonlinear Resonances in Atomic and Molecular Spectra* (Nauka, Novosibirsk, 1979) [in Russian].
2. A. K. Popov, *Introduction in Nonlinear Spectroscopy* (Nauka, Novosibirsk, 1983) [in Russian].
3. V. S. Letokhov and V. P. Chebotaev, *High-Resolution Nonlinear Laser Spectroscopy* (Nauka, Moscow, 1990) [in Russian].
4. S. G. Rautian and I. I. Sobel’man, Zh. Éksp. Teor. Fiz. **41**, 456 (1961) [Sov. Phys. JETP **14**, 328 (1962)].
5. B. R. Mollow, Phys. Rev. A **5**, 2217 (1972).
6. S. Haroche and F. Hartmann, Phys. Rev. A **6**, 1280 (1972).
7. G. S. Agarwal, Phys. Rev. A **19**, 923 (1979).
8. R. W. Boyd, M. G. Raymer, P. Narum, and D. J. Harter, Phys. Rev. A **24**, 411 (1981).
9. D. S. Bakaev, Yu. A. Vdovin, V. M. Ermachenko, and S. I. Yakovlenko, Kvantovaya Élektron. (Moscow) **12**, 126 (1985).
10. G. Grynberg and C. Cohen-Tannoudji, Opt. Commun. **96**, 150 (1993).

11. C. Szymanowski, C. H. Keitel, B. J. Dalton, and P. L. Knight, J. Mod. Opt. **42**, 985 (1995).
12. A. M. Bonch-Bruevich, V. A. Khodovoĭ, and N. A. Chigir', Zh. Éksp. Teor. Fiz. **67**, 2069 (1974) [Sov. Phys. JETP **40**, 1027 (1975)].
13. C. Wei and N. B. Manson, Phys. Rev. A **49**, 4751 (1994).
14. F. Y. Wu, S. Ezekiel, M. Ducloy, and B. R. Mollow, Phys. Rev. Lett. **38**, 1077 (1977).
15. L. S. Gaĭda, I. S. Zeĭlikovich, S. A. Pul'kin, and É. E. Fradkin, Opt. Spektrosk. **65**, 802 (1988) [Opt. Spectrosc. **65**, 474 (1988)].
16. A. Yu. Parkhomenko and A. M. Shalagin, Zh. Éksp. Teor. Fiz. **120**, 830 (2001) [JETP **93**, 723 (2001)].
17. V. N. Faddeeva and N. M. Terent'ev, *Tables of the Values of the Probability Integral in Terms of Complex Argument*, Ed. by V. A. Fok (Gostekhizdat, Moscow, 1954) [in Russian].
18. F. Kh. Gel'mukhanov and A. I. Parkhomenko, Phys. Scr. **44**, 477 (1991).
19. S. N. Atutov, I. M. Ermolaev, and A. M. Shalagin, Zh. Éksp. Teor. Fiz. **92**, 1215 (1987) [Sov. Phys. JETP **65**, 679 (1987)].
20. S. N. Atutov, B. V. Bondarev, S. M. Kobtzev, *et al.*, Opt. Commun. **115**, 276 (1995).
21. F. Sh. Ganikhanov, I. G. Konovalov, V. N. Kulyasov, *et al.*, Pis'ma Zh. Éksp. Teor. Fiz. **54**, 433 (1991) [JETP Lett. **54**, 433 (1991)].
22. I. G. Konovalov, V. N. Kulyasov, V. B. Morozov, and V. G. Tunkin, Opt. Spektrosk. **77**, 329 (1994) [Opt. Spectrosc. **77**, 291 (1994)].
23. E. C. Beaty, P. L. Bender, and A. R. Chi, Phys. Rev. **112**, 450 (1958).
24. M. E. Zhabotinskiĭ, in *Quantum Electronics: Little Encyclopedia*, Ed. by M. E. Zhabotinskiĭ (Sovetskaya Éntsiklopediya, Moscow, 1969), p. 35 [in Russian].
25. R. S. Eng, A. R. Calawa, T. C. Harman, *et al.*, Appl. Phys. Lett. **21**, 303 (1972).

Translated by N. Wadhwa

Journal of Experimental and Theoretical Physics, Vol. 100, No. 2, 2005, pp. 294–304.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 127, No. 2, 2005, pp. 331–342.

NUCLEI, PARTICLES, FIELDS, GRAVITATION, AND ASTROPHYSICS

Growth of Black Holes at the Centers of Galaxies: Absorption of Stars and Activity of Galactic Nuclei

V. A. Sirota, A. S. Il'in*, K. P. Zybin, and A. V. Gurevich

Lebedev Physical Institute, Russian Academy of Sciences, Leninskii pr. 53, Moscow, 119991 Russia

**e-mail: asi@lpi.ru*

Received July 13, 2004

Abstract—We study the influence of supermassive black holes on the distribution of stars at the centers of galaxies. We analyze relaxation processes associated with encounters between stars and their absorption by black holes. For an isothermal distribution of stars, we obtain the growth law and estimate the current masses of black holes. The tidal disruption of stars near black holes is considered as a possible cause of the activity of galactic nuclei. © *2005 Pleiades Publishing, Inc.*

1. INTRODUCTION

At present, the fact that the compact mass concentrations discovered at the centers of most galaxies are supermassive black holes with masses of 10^6–10^9 solar masses ($M_\odot$) is almost beyond question [1].

There are two scenarios for the origin of supermassive black holes. They can form simultaneously with their host galaxies (or even earlier) [2] or grow gradually, starting from a small seed black hole. In this paper, we consider the possibilities for the growth of black holes in terms of the latter scenario.

A black hole can affect the surrounding matter in two ways. First, it is a source of a strong Coulomb potential that directly affects the motion of the nearest stars. The corresponding kinetic problem was first considered in [3]. A black hole can also absorb baryonic matter (stars, interstellar gas) and cold dark matter from the central part of the galactic nucleus (bulge), thereby affecting the distribution of stars and dark matter. For the nearest stars (whose orbits lie in the region of influence of the gravitational Coulomb potential of a black hole), this problem was discussed previously [4–8]. In particular, it was shown in [7, 8] that the flow of stellar matter onto a black hole arises from the energy and angular momentum diffusion of stellar orbits. However, as was first pointed out in [7], a supermassive black hole at the center of a typical galaxy must absorb stars from a region that is much larger than the region of their own self-consistent Coulomb potential. In this region, the gravitational potential is determined by the distribution of stars, interstellar gas, and dark matter. In addition, we show below that, in contrast to previously considered stationary solutions, the real process is essentially nonstationary.

The kinetic theory that describes the dynamics of nonbaryonic cold dark matter alone (without baryons and a black hole) was developed in [9]. Dark matter was shown to break up under its own gravitational forces into spherically symmetric objects with a singular density distribution at their centers.[1] These objects have long been observed as giant galactic halos. As baryonic matter cools down, it sinks to the bottom of the potential wells produced by dark matter and forms galaxies with seed black holes [10]. In particular, this scenario well explains why black holes are located exactly at the dynamical centers of galaxies [11].

The interaction between baryonic matter and dark matter at the center of a galaxy and, in particular, the gravitational scattering of dark matter particles by stars followed by their capture by a black hole were investigated in [12]. It was shown that this process could give rise to black holes with masses of $(10^7$–$10^8)M_\odot$.

In this paper, we analyze the combined dynamics of cold dark matter and stars in galactic nuclei. We show that distant encounters of stars with one another and with dark matter particles play an important role. These encounters are responsible for the relaxation processes in a medium of stars and dark matter that eventually lead to the growth of a black hole. In this case, the role of a black hole reduces not so much to a direct gravitational influence as to the creation of a boundary condition that the distribution function satisfies.

This paper is structured as follows. In Section 2, we write out the kinetic equation that describes the evolution of the distribution function for stars and dark matter particles in a closed form. In Section 3, we discuss the boundary conditions that the distribution function must satisfy in the presence of a black hole. In Section 4, we find an approximate solution of the kinetic equation for stars in the absence of dark matter and interstellar gas and show that the mass of the black hole as it absorbs stars increases with time as $t^{1/2}$. In conclusion, we

[1] However, the degree of this singularity is not enough for a black hole to be formed.

obtain a rough estimate of the energy release during the absorption of stellar matter and compare it with the average observed activity of galactic nuclei.

2. THE KINETIC EQUATION

Let us denote the distribution functions of stars and dark matter particles by $f_*(\mathbf{r}, \mathbf{v})$ and $f_d(\mathbf{r}, \mathbf{v})$, respectively. The mass of the type-α particles per phase volume element d^3rd^3v is $f_\alpha d^3rd^3v$.

The combined dynamics of the system is described by a kinetic equation with the collision term in Landau's form [13]:

$$\frac{df_\alpha}{dt} = \mathrm{St}[f],$$

or

$$\frac{\partial f_\alpha}{\partial t} + \mathbf{v}\cdot\nabla f_\alpha + \nabla\Psi\cdot\frac{\partial f_\alpha}{\partial \mathbf{v}} = 2\pi G^2\Lambda\frac{\partial}{\partial v_k}\sum_\beta\int d^3v' w_{kp} \times\left(M_\beta f_\beta\frac{\partial f_\alpha}{\partial v_p} - M_\alpha\frac{\partial f_\beta}{\partial v'_p}f_\alpha\right). \quad (1)$$

Here, the Greek subscript denotes the type of particles, M_α is the mass of the type-α particles,

$$w_{kp} = (u^2\delta_{kp} - u_k u_p)/u^3,$$
$$\mathbf{u} = \mathbf{v}' - \mathbf{v}$$

is the relative collision velocity, $\Lambda \sim \ln(N/2)$ is the gravitational Coulomb logarithm, and N is the number of particles in the system (e.g., $N \sim 10^9$ in the bulge of a typical galaxy). The gravitational potential $\Psi(\mathbf{r})$ satisfies the Poisson equation

$$\Delta\Psi = 4\pi G\sum_\alpha\rho_\alpha(\mathbf{r}) = 4\pi G\sum_\alpha\int f_\alpha(\mathbf{r}, \mathbf{v})d^3v,$$

where $\rho_\alpha(\mathbf{r})$ is the mass density of the type-α particles.

Discarding the terms of the order of M_d/M_*, we rewrite Eq. (1) as

$$\frac{df_d}{dt} = 2\pi G^2 M_*\Lambda\frac{\partial}{\partial v_k}W_{kp}\frac{\partial f_d}{\partial v_p}, \quad (2)$$

$$\frac{df_*}{dt} = 2\pi G^2 M_*\Lambda\left(\frac{\partial}{\partial v_k}W_{kp}\frac{\partial f_*}{\partial v_p} - \frac{\partial}{\partial v_k}W_k f_*\right), \quad (3)$$

where

$$W_k = \int\frac{\partial}{\partial v'_p}(f_* + f_d)w_{kp}d^3v',$$
$$W_{kp} = \int f_* w_{kp}d^3v'. \quad (4)$$

Thus, the evolution of the distribution function for dark matter particles is determined by the right-hand side of Eq. (2), which describes their scattering by stars. For the stars themselves, the corresponding term in Eq. (3) has a more complex form and contains the term that corresponds to dynamical friction (against stars and dark matter particles).

An important feature of the motion of gravitating particles in real galactic systems is that the collisions are rare and, therefore, the collision term in the kinetic equation is small compared to the remaining terms. This implies that the changes in the orbital parameters of an individual particle during the orbital period are also small. This fact is easiest to take into account by passing to the action–angle variables $(\mathbf{I}, \phi)$, where $\mathbf{I} = (I, m, m_z)$ are the action variables (m is the magnitude of the angular momentum, m_z is its component along the z axis, and I is the adiabatic invariant; all of the quantities were normalized to unit mass):

$$I = \frac{1}{2\pi}\oint v_r dr = \frac{1}{\pi}\int_{r_-(E,m)}^{r_+(E,m)}\sqrt{2\left(E - \Psi(r) - \frac{m^2}{2r^2}\right)}dr, \quad (5)$$
$$E = \frac{v^2}{2} + \Psi(r), \quad \mathbf{m} = \mathbf{r}\times\mathbf{v}, \quad E - \Psi(r_\pm) - \frac{m^2}{2r_\pm^2} = 0,$$

and $\boldsymbol{\phi}$ are the corresponding angle variables. In the new variables, the left-hand side of the kinetic equation contains no derivatives with respect to I_k and is

$$\frac{\partial f_\alpha}{\partial t} + \omega_k\frac{\partial f_\alpha}{\partial\phi_k} = \mathrm{St}[f],$$

where ω_k are the orbital frequencies that correspond to I_k. Let us now use the fact that the collisions are rare and, hence, the collision frequency is much lower than the orbital frequencies. In this case, we can perform averaging over the angle variables ϕ_k [12]. Equations (2) and (3) then take the form [12, 14]

$$\frac{\partial f_*}{\partial t} = \frac{\partial}{\partial I_k}R_{kp}\frac{\partial f_*}{\partial I_p} - \frac{\partial}{\partial I_k}(R_k f_*), \quad (6)$$

$$\frac{\partial f_d}{\partial t} = \frac{\partial}{\partial I_k}R_{kp}\frac{\partial f_d}{\partial I_p}, \quad (7)$$

where

$$R_k = 2\pi G^2 M_*\Lambda\frac{1}{(2\pi)^3}\int d^3\phi\frac{\partial I_k}{\partial v_p}W_p,$$
$$R_{kp} = 2\pi G^2 M_*\Lambda\frac{1}{(2\pi)^3}\int d^3\phi\frac{\partial I_k}{\partial v_m}\frac{\partial I_p}{\partial v_n}W_{mn}. \quad (8)$$

Near the center, the distribution function is spherically symmetric and does not depend on the direction of the

angular momentum. It can be shown that, in this case,

$$R_3 = \frac{m_z}{m}R_2, \quad R_{31} = \frac{m_z}{m}R_{21}, \quad R_{32} = \frac{m_z}{m}R_{22},$$

and we can eliminate m_z from Eq. (6). Finally, the equations take the form

$$\begin{aligned}\frac{\partial f_*}{\partial t} &= \frac{\partial}{\partial I}\left(R_{11}\frac{\partial f_*}{\partial I} + R_{12}\frac{\partial f_*}{\partial m} - R_1 f_*\right)\\ &+ \frac{1}{m}\frac{\partial}{\partial m}m\left(R_{21}\frac{\partial f_*}{\partial I} + R_{22}\frac{\partial f_*}{\partial m} - R_2 f_*\right),\end{aligned} \tag{9}$$

$$\begin{aligned}\frac{\partial f_d}{\partial t} &= \frac{\partial}{\partial I}\left(R_{11}\frac{\partial f_d}{\partial I} + R_{12}\frac{\partial f_d}{\partial m}\right)\\ &+ \frac{1}{m}\frac{\partial}{\partial m}m\left(R_{21}\frac{\partial f_d}{\partial I} + R_{22}\frac{\partial f_d}{\partial m}\right).\end{aligned} \tag{10}$$

Thus, the kinetic equation (1) reduces to a divergence equation,

$$\frac{df_\alpha}{dt} = \mathrm{div}\mathbf{S}_\alpha,$$

in some effective three-dimensional *I*–*m* space with a natural cylindrical structure; the radial action *I* is along the *z* axis, while the magnitude of the angular momentum *m* acts as the polar radius. The distribution function, which depends only on *I* and *m*, is also cylindrically symmetric.

The flux of matter through a surface in *I*–*m* space can be determined by integrating **S** over this surface. Since the $I = 0$ plane is a nonphysical boundary, the flux through it must be identically equal to zero. Thus, we have obtained the first boundary condition

$$S_I|_{I=0} = \left(R_{11}\frac{\partial f_\alpha}{\partial I} + R_{21}\frac{\partial f_\alpha}{\partial m} - R_1 f_\alpha\right)\Big|_{I=0} = 0. \tag{11}$$

The second boundary condition arises from the presence of a black hole and is discussed in the next section.

In the absence of dark matter, the simplest (but important) solution of Eq. (9) is an isothermal distribution function:

$$\begin{aligned}f_{is} &= f_0 e^{-E/\sigma^2}, \quad \Psi_{is} = 2\sigma^2\ln r,\\ f_0 &= [(2\pi)^{5/2}G\sigma]^{-1},\end{aligned} \tag{12}$$

which is specified by only one parameter, the stellar velocity dispersion σ. This is an equilibrium distribution with a zero flux through an arbitrary surface in *I*–*m* space. Two relations that the coefficients R_k and R_{kp} must satisfy for an isothermal function follow from this:

$$\begin{aligned}R_{11}\frac{\partial f_{is}}{\partial I} + R_{12}\frac{\partial f_{is}}{\partial m} - R_1 f_{is} &= 0,\\ R_{21}\frac{\partial f_{is}}{\partial I} + R_{22}\frac{\partial f_{is}}{\partial m} - R_2 f_{is} &= 0.\end{aligned} \tag{13}$$

The coefficients R_k and R_{kp} themselves are calculated for this case in the Appendix. Here, we present the result:

$$\begin{aligned}R_{22} &\approx 0.5GM_*\Lambda\sigma \equiv R,\\ R_{12} &\approx -0.6R(1 - \mu\Phi_2(\mu)),\\ R_{11} &\approx R(0.4\Phi_1(\mu) - 0.8\mu\Phi_2(\mu) + 0.4),\\ R_1 &\approx -\frac{R}{I + 0.6m}(0.7\Phi_1(\mu) - 0.8\mu\Phi_2(\mu)),\\ R_2 &\approx -1.3\frac{R}{I + 0.6m}\mu\Phi_2(\mu),\\ \mu &\approx 0.3\frac{m}{I + 0.6m},\end{aligned} \tag{14}$$

where

$$\Phi_2(\mu) \approx 1 + 0.35\mu^{-0.6},$$

$$\Phi_1(\mu) \approx 0.5 - 0.8\ln\mu(\Phi_2(\mu) - 1) \approx 0.5 + 0.33\mu^{-0.9}.$$

The isothermal distribution function written in the action variables is

$$f_{is}(I, m) = f_0 e^{-E/\sigma^2} \approx \frac{0.3 f_0\sigma^2}{(I + 2m/\pi)^2}. \tag{15}$$

3. A BLACK HOLE: FORMULATION OF THE PROBLEM

The collisions of dark matter particles with stars and their capture by a black hole were considered in detail in [12]. Here, we assume that dark matter and interstellar gas are present in the region under consideration in insignificant amounts and discuss only the dynamics of stars. Therefore, below, we omit the subscript "*" for f_*.

A black hole with a mass M_{bh} disrupts (and then partially absorbs) all of the stars that pass near the tidal radius

$$r_t = \left(\frac{6}{\pi}\frac{M_{\mathrm{bh}}}{\rho_*}\right)^{1/3},$$

where ρ_* is the mean stellar density (we assume it to be comparable to the solar density $\rho_\odot$). In terms of the variables (I, m), this implies that all of the stars within the loss cone

$$m \le m_t = \sqrt{2GM_{\rm bh}r_t}$$
$$\approx 6\times 10^{18}\left(\frac{M_{\rm bh}}{M_\odot}\right)^{2/3}\left(\frac{\rho_\odot}{\rho_*}\right)^{1/6} \text{cm}^2\ \text{s}^{-1}, \quad (16)$$

will be disrupted during one orbital period. On the other hand, if the angular momentum is lower than its limiting value

$$m \le m_g = 4GM_{\rm bh}/c, \quad (17)$$

then the star will be captured by the black hole, irrespective of m_t [15]. (If the black hole mass $M_{\rm bh}$ is larger than $\sim 3\times 10^8 M_\odot$, the tidal radius is smaller than the gravitational radius r_g and all stars are absorbed without disruption. For dark matter particles, the loss cone is determined only by m_g.) For simplicity, we denote the maximum of the two values, m_g and m_t, by m_t. Thus, the capture of a star depends only on its angular momentum.

It is important to note that, since the gravitational potential of the black hole is spherically symmetric, it does not change the angular momenta of stars. Therefore, if the collisions are ignored, then the number of stars in the loss cone can only decrease (through the capture by the black hole). The maximum value of this quantity is determined by the initial distribution function (before the formation of a black hole). For example, for the isothermal function (15), the total mass of the stars in the loss cone of a black hole [12] is

$$\Delta M = (2\pi)^{1/2}\sigma m_t/G.$$

Simple estimates [12] indicate that the relative increase in black hole mass, $\Delta M/M_{\rm bh}$, caused by the capture of these stars is small. Thus, a significant growth of the black hole is possible only if the encounters between stars are taken into account, since this is the only process that changes their angular momenta.[2]

The disruption of stars at $m \le m_t$ yields the boundary condition to Eq. (9):

$$f|_{m=m_t} = 0. \quad (18)$$

This affects the gradient of the distribution function $\partial f/\partial m$ near the boundary of the loss cone and gives rise to a flux of stars in momentum space.

[2] Here, we do not consider another effect related to the collective interaction between stars that can also lead to the filling of the loss cone (the so-called cone instability).

It follows from Eqs. (6), (9), and Gauss's theorem that the black hole growth rate $dM_{\rm bh}/dt$ is equal to the flux of stars through the boundary m_t:

$$\dot M_{\rm bh} = S = \int d^3\phi \int_0^\infty dI \int_{-m_t}^{+m_t} dm_z S_m \Big|_{m=m_t}$$
$$= (2\pi)^3 2m_t \int_0^\infty dI\left(R_{22}\frac{\partial f}{\partial m} + R_{12}\frac{\partial f}{\partial I} - R_2 f\right)\Big|_{m=m_t}.$$

Since the last two terms in this equation vanish due to the boundary condition (18), the expression for $\dot M_{\rm bh}$ takes the form

$$\dot M_{\rm bh} = (2\pi)^3 2m_t \int_0^\infty dI\left(R_{22}\frac{\partial f}{\partial m}\right)\Big|_{m=m_t}. \quad (19)$$

Thus, to determine the current masses of black holes and the distribution of stars around them, we must solve Eq. (9) with the boundary conditions (11) and (18). The choice of the isothermal function (15) as the initial condition seems most natural, because this is a stationary distribution with a self-consistent potential and without any fluxes.

However, the disruption and capture of stars by a black hole disturb the equilibrium and leads to the evolution of the initial isothermal function. Below, we assume that the change in the distribution function affects only slightly the coefficients R_k and R_{kp} (14) calculated for the isothermal function. This is a natural assumption, since the distribution function appears in them only via integrals. We also disregard the gravitational potential of the black hole, which changes the velocities of the particles, thereby affecting the collision processes, without directly changing their momenta. In Section 4, we discuss the validity of these assumptions.

4. AN APPROXIMATE SOLUTION

It appears that Eq. (9) cannot be solved exactly. In this section, we obtain only some of the integrated characteristics that allow the consumption rate S of stars and its time evolution to be estimated.

Let us first consider a region near the boundary, $m_t \le m \ll \sqrt{R_{22}t}$ and $I > m$. It follows from a comparison of the various terms in Eq. (9) that, for example, the term with R_{22} is on the order of

$$\frac{R_{22}f}{m^2} \gg \frac{\partial f}{\partial t}.$$

This implies that the pattern near the boundary is quasi-stationary. All of the other terms, except two terms (with R_{11} and R_1, which are singular for $m \longrightarrow 0$ (see (14)), are on the order of $R_{22}f/Im$; i.e., they are smaller. To get rid of the two singular terms, let us integrate Eq. (9) over I and introduce

$$F(m) = \int f(I, m)dI.$$

Thus, we obtain[3]

$$\frac{1}{m}\frac{\partial}{\partial m}mR_{22}\frac{\partial F}{\partial m} \approx 0. \quad (20)$$

Using the boundary condition at $m = m_t$ yields

$$F(m) = C(t)\ln\frac{m}{m_t}, \quad m \ll \sqrt{Rt}. \quad (21)$$

The isothermal function (15) remains the solution far from the boundary,

$$F(m) = 0.5 f_0\sigma^2\frac{1}{m}, \quad m \geq \sqrt{Rt}.$$

Substituting (21) into (19), we see that

$$S = (2\pi)^3 2R_{22}C(t). \quad (22)$$

The factor $C(t)$ can be determined from the conservation condition for the total mass

$$S = -\frac{d}{dt}\int d^3\phi\int dI\int dm\int_{-m}^{+m} dm_z f$$
$$= -(2\pi)^3\frac{d}{dt}\int 2mF(m)dm.$$

Comparing this equality with (22) and calculating the integral (joining the two asymptotics at $m = m_D$), we obtain two equations:

$$C(t)m_D\ln\frac{m_D}{m_t} = 0.5 f_0\sigma^2, \quad (23)$$

$$-\frac{\dot{C}}{C}m_D^2\left(\ln\frac{m_D}{m_t} - \frac{1}{2}\right) = 2R. \quad (24)$$

Thus, m_D is the characteristic scale that defines the region where the distribution function changes significantly. If the boundary does not move, m_t = const, then the approximate solution is

$$m_D(t) \approx 2\sqrt{Rt}\ln^{-1/2}\left(\frac{2\sqrt{Rt}}{m_t}\right), \quad (25)$$

$$C(t) \approx \frac{f_0\sigma^2}{4\sqrt{Rt}}\ln^{-1/2}\left(\frac{2\sqrt{Rt}}{m_t}\right). \quad (26)$$

Knowing the time dependence of the flux $S = \dot{M}_{\rm bh}$, we can find the time dependence of the black hole mass

$$M_{\rm bh} \approx \frac{(2\pi)^3 f_0\sigma^2\sqrt{Rt}}{\sqrt{\ln(2\sqrt{Rt}/m_t)}}.$$

Substituting f_0 yields

$$M_{\rm bh} \approx \frac{2.6\sigma\sqrt{Rt}}{G\sqrt{\ln(2\sqrt{Rt}/m_t)}}. \quad (27)$$

Remarkably, this result depends weakly on m_t. Therefore, solution (25)–(27) also remains valid if we take into account the increase in m_t with time ((16) or (17)) in accordance with the growth of the black hole mass (27).

Taking $\Lambda = 15$ and assuming the typical stellar mass M_* to be close to the solar mass, we obtain

$$M_{\rm bh} \approx 1.2\times 10^7 M_\odot\left(\frac{\sigma}{200\ {\rm km/s}}\right)^{3/2}\sqrt{\frac{t}{3\times 10^{17}\ {\rm s}}}. \quad (28)$$

The flux of stars onto the black hole (22) is

$$S \approx \frac{M_{\rm bh}}{2t} \approx (6\times 10^{-4} M_\odot/{\rm yr})$$
$$\times\left(\frac{\sigma}{200\ {\rm km/s}}\right)^{3/2}\left(\frac{t}{3\times 10^{17}\ {\rm s}}\right)^{-1/2}. \quad (29)$$

It decreases with time as $t^{-1/2}$ and is equal to 3.6×10^{43} erg s^{-1} at present if $\sigma = 200$ km s^{-1}. The value of m_D that corresponds to mass (28) is

$$m_D \approx 0.8\sqrt{Rt}.$$

Note that at least half of the initial mass always remains in the region $m \leq m_D$. Therefore, the change in the distribution function caused by the particle escape into the loss cone will not significantly affect the coefficients R_{kp}.

Let us now discuss the influence of the black hole potential. The following region of space corresponds to the characteristic angular momentum m_D at time t:

$$r_D \approx \frac{m_D}{\sigma} \approx 1\ {\rm pc}\left(\frac{\sigma}{200\ {\rm km/s}}\right)^{-1/2}\left(\frac{t}{3\times 10^{17}\ {\rm s}}\right)^{1/2}. \quad (30)$$

The evolution of the distribution function of stars in this region is determined by their absorption by the black

[3] Here, we disregard the low values of I for which the terms with R_{12} and R_2 can also be large (but no larger than the term with R_{22}). In any case, it can be easily verified that the adiabatic invariant near the boundary (and, hence, everywhere) can also increase with time (since $S_I > 0$). Therefore, the contribution of stars with low I to the flux S and the function F is small.

hole. Since this region is much smaller than the bulge size (~1 kpc), we may disregard the outer boundary of the distribution. On the other hand, the black hole potential is significant at distances of

$$r \le r_a \approx \frac{GM_{\rm bh}}{\sigma^2} \approx 1\ {\rm pc}\left(\frac{M_{\rm bh}}{10^8 M_\odot}\right)\left(\frac{\sigma}{200\ {\rm km/s}}\right)^{-2}.$$

It thus follows that for black holes whose current masses do not exceed

$$M_{\rm bh} \le 10^8 M_\odot \left(\frac{\sigma}{200\ {\rm km/s}}\right)^{3/2}, \qquad (31)$$

the region of influence of the black hole is always smaller than the characteristic region r_D (30). Therefore, the stars of interest move outside the region of gravitational influence of the black hole most of the time. In general, this conclusion agrees with the estimates from [7].

Let us now briefly discuss the validity range of the boundary condition (18). Only the quantities averaged over the orbital period appear in all of the equations starting from (6). The boundary condition (18) also implies that all of the particles with the mean angular momentum lower than m_t are captured by the black hole during one period. However, for fairly elongated orbits, the change in the particle angular momentum over the orbital period is comparable to the boundary angular momentum m_t. Thus, the condition for the absorption of a star by a black hole during one orbital period is determined not only by its mean angular momentum, but also by its small fluctuations about the mean. (It can be shown that the loss cone is smeared at large distances.)[4] In our terms, this condition can be formulated as follows.

The boundary condition (18) ceases to be valid at large $I > I_{\rm crit}$ (i.e., for elongated orbits). The rms deviation of the angular momentum Δm over the period $T(I, m)$ is

$$\Delta m \sim \sqrt{RT(I, m)}.$$

$I_{\rm crit}$ can then be defined by the relation

$$\sqrt{RT(I_{\rm crit}, m_t)} = m_t.$$

Stars with $I \ll I_{\rm crit}$ and $m < m_t$ will be captured by the black hole during one period, while those with $I \gg I_{\rm crit}$ will definitely not be captured. On the other hand, as we saw in Section 3, the main changes in the distribution function that contribute to the flux occur on scales

$$I \le \sqrt{Rt},$$

where t is the age of the black hole. Hence, the loss-cone smearing effect under discussion does not affect the consumption rate of stars by the black hole if

$$I \sim \sqrt{Rt} \ll I_{\rm crit}. \qquad (32)$$

This estimate can be obtained if we restrict ourselves to the integration from I to $I_{\rm crit}$ when determining the function $F(m)$. Relation (23) will then remain valid if $I_{\rm crit} \gg m_D$. Using relation (25), we obtain condition (32).

For large $I \gg m_t$, the orbital period is

$$T(I, m_t) \sim \pi I/\sigma^2$$

(see Appendix). Given relation (16), it may be concluded that, if the black hole mass at time t satisfies the inequality

$$M_{\rm bh} \ge M_{\rm crit} = \left(\frac{1}{50}\frac{t}{3\times 10^{17}\ {\rm s}}\frac{200\ {\rm km/s}}{\sigma}\right)^{3/8} \times 10^7 M_\odot, \qquad (33)$$

then the existence of $I_{\rm crit}$ may be disregarded. $M_{\rm crit} \approx 2 \times 10^6 M_\odot$ corresponds to the current time. Comparing the time dependence (33) with (28), we can easily see that (only if σ does not change greatly) the masses of the black holes that grow through the described relaxation mechanism are almost always larger than $M_{\rm crit}$.

5. DISCUSSION

In this paper, we derived an equation that describes the kinetic relaxation of a two-component system (stars and dark matter) in the vicinity of a supermassive black hole. We analyzed the possibilities for the growth of a black hole via the consumption of stars without including nonbaryonic cold dark matter and interstellar gas. We showed that for black holes with current masses less than $10^6 M_\odot$, (33), individual collisions between stars near the pericenter can play an important role. At masses larger than $10^8 M_\odot$, (31), the black hole potential can strongly affect the consumption of stars. Nevertheless, most of the discovered black holes [11] have masses between $10^6 M_\odot$ and $10^8 M_\odot$ for which the approximations used above are valid. Thus, the current masses that correspond to solution (28) are in reasonable agreement with observational data.

Solution (28) also indicates that $M_{\rm bh}$ depends weakly on time,

$$M_{\rm bh} \sim \sqrt{t} \sim (1+z)^{-3/4}.$$

Therefore, the differences between the black hole masses in nearby and distant galaxies are small and invisible against the background of other effects. On the

[4] The importance of this effect was first pointed out in [5, 7], where the concepts of critical energy and critical radius were introduced.

other hand, the observed relations between the black hole masses and the stellar velocity dispersion in galactic bulges (the M–σ relations: $\propto\sigma^{3.7}$ in [16] and $\propto\sigma^{5.3}$ in [17]) differ from solution (28). However, this discrepancy is not catastrophic for the described theory, since the measurement errors of the dispersion in the range $10^6 M_\odot < M_{bh} < 10^8 M_\odot$ under discussion is too large to make the final choice between the theoretical (28) and empirical M–σ relations. It should also be noted that, apart from the mentioned M–σ relation, the M_{bh}–bulge mass [11] and M_{bh}–dark matter halo mass [12, 18] are also discussed in the literature.

Let us now discuss the contribution of nonbaryonic dark matter, which can increase the black hole mass compared to (28) in galaxies where dark matter is present in the central part in significant amounts. Previously [12], we considered the possibilities for the growth of a black hole via the consumption of dark matter. We showed that, if the dark matter mass in the bulge is comparable to the baryonic mass of the bulge, then its contribution to the black hole is about an order of magnitude larger than the contribution from stars:

$$M_{bh(dark)} = 8\times 10^7 M_\odot \left(\frac{M_H}{10^{12} M_\odot}\right)^{1/2} \times \left(\frac{R_H}{100\ \text{kpc}}\right)^{-9/14} \left(\frac{\sigma}{200\ \text{km/s}}\right)^{4/7} \left(\frac{t}{3\times 10^{17}\ \text{s}}\right)^{4/7}, \quad (34)$$

where M_H and R_H are the mass and radius of the dark matter halo, respectively.

This relation and (28) correspond to two limiting cases where the black hole grows via the consumption of mainly dark or baryonic matter. Given the natural spread in parameters, solutions (28) and (34) cover the range

$$5\times 10^6 M_\odot \le M_{bh} \le 2\times 10^8 M_\odot, \quad (35)$$

which includes most of the observed black holes.

It should be emphasized that accretion of dark and stellar matter onto black holes is attributed to scattering by stars. Accordingly, a black hole begins to grow at a moment t_0 when both a seed black hole and its stellar environment already exist. Thus, it is the starting moment for all processes allowed for in the theory. Both in Eqs. (1) and their solutions, the time t is measured from t_0. It is very important the flow of matter to a black hole behaves as $(t - t_0)^{-1/2}$ at $t \longrightarrow t_0$ (see (28)). This means that the most intense accretion takes place during the formation of black hole and stellar bulge.

It is commonly assumed that the consumption of stars by a black hole is mainly responsible for the activity of galactic nuclei [19, 20]. Based on this assumption, we analyzed the observed X-ray activity of the galactic centers using data on 46 galaxies [21]. The data were obtained mainly from the ROSAT and EINSTEIN satellites, which recorded X-ray emission in the energy ranges 0.1 to 2.4 keV and 0.2 to 4.0 keV, respectively. Because of the high concentration of interstellar gas and dust, the galactic centers in the low-energy part of the energy range are usually invisible. The results are presented in the table. Their comparison with theory leads us to the following conclusions:

(1) The theory developed here is valid for black holes with masses in the range $5\times 10^6 M_\odot \le M_{bh} \le 2\times 10^8 M_\odot$ (see (35)). Most of the black holes listed in the table (more than 60%) also lie within this range.

(2) In accordance with (29), the total power released during the disruption of stars within the tidal radius of a black hole is, on average, $\langle S\rangle \approx 3.6\times 10^{43}$ erg s^{-1}. According to the table, the luminosity averaged over $N = 46$ sources is

$$\langle L\rangle = \frac{1}{N}\sum_{k=1}^{N}\langle L_k\rangle \approx 4\times 10^{43}\ \text{erg/s}.$$

We can assume that the derived close values of $\langle S\rangle$ and $\langle L\rangle$ indicate that the estimate of the stellar flux into the tidal radius that follows from the presented theory and the explanation of the activity of galactic nuclei in terms of the tidal disruption of these stars are quite plausible.

(3) We see from the table that there is no correlation between the masses of black holes and the activity of galactic nuclei. Moreover, the differences in the luminosities of nuclei containing close-mass black holes can be significant (three or four orders of magnitude). As an example, we can cite the galaxies 3C 120 ($M_{bh} = 2.3\times 10^7 M_\odot$, $L = 9.8\times 10^{43}$ erg s^{-1}) and NGC 1068 ($M_{bh} = 1.6\times 10^7 M_\odot$, $L = 1.3\times 10^{41}$ erg s^{-1}), NGC 4459 ($M_{bh} = 6.5\times 10^7 M_\odot$, $L = 4.2\times 10^{39}$ erg s^{-1}) and PG 0026+129 ($M_{bh} = 4.5\times 10^7 M_\odot$, $L = 2.8\times 10^{44}$ erg s^{-1}). It would be natural to assume that these differences stem from the fact that the duration τ_e of the energy emission due to the absorption of a star is much shorter than the mean time between two absorption events, $T \approx M_\odot/S \approx 2000$ years (29). An estimate of $\tau_e \le T/10 \approx 200$ years follows from the observed difference between the luminosities (three or four orders of magnitude) (assuming the luminosity to decrease exponentially). Therefore, we may conclude that most of the galaxies with black holes presently radiate very weakly and are not active at all (as, e.g., our Galaxy). On the other hand, the maximum luminosity must be much higher than the mean S from (29). For example, the highest luminosity observed in PG 0052+251 is 5×10^{44} erg s^{-1}, which is an order of magnitude higher than the mean luminosity.

(4) We also see from the table that there is a large group of galaxies (NGC 4374, NGC 4594, NGC 4649, NGC 4251 and others) with large black holes (>3 ×

X-ray luminosities and masses of supermassive black holes in galactic nuclei

Object	L, erg s^{-1}	M_{bh}, $M_\odot$	Object	L, erg s^{-1}	M_{bh}, $M_\odot$
3C 120 (Mrk 1506)	9.772×10^{43}	2.3×10^{7}	NGC 4203	7.471×10^{40}	$<1.2\times10^{7}$
	1.343×10^{44}			6.424×10^{40}	
Ark 120 (Mrk 1095)	7.805×10^{43}	1.84×10^{8}	NGC 4258 (M106)	2.914×10^{40}	4.1×10^{7}
Circinus	2.47×10^{40}	1.3×10^{6}	NGC 4261 (3C 270)	1.575×10^{41}	5.2×10^{8}
Fairall 9	6.688×10^{43}	8.0×10^{7}		1.434×10^{41}	
IC 1459	8.398×10^{40}	3.7×10^{8}	NGC 4291	6.128×10^{40}	1.5×10^{8}
	1.051×10^{41}			6.769×10^{40}	
IC 4329A	3.748×10^{43}	5×10^{6}		4.28×10^{40}	
	2.541×10^{43}		NGC 4342	2.077×10^{39}	3.4×10^{8}
UGC 3973 (Mrk 79)	2.594×10^{43}	5.2×10^{7}	NGC 4374 (M84)	6.317×10^{40}	1.6×10^{9}
Mrk 110	1.625×10^{44}	5.6×10^{6}		6.657×10^{40}	
Mrk 335	2.04×10^{43}	6.3×10^{6}		6.138×10^{40}	
	4.916×10^{43}		NGC 4459	4.169×10^{39}	6.5×10^{7}
Mrk 509	1.188×10^{44}	5.78×10^{7}	NGC 4473	1.109×10^{40}	
Mrk 590 (NGC 863)	3.853×10^{43}	1.78×10^{7}		5.089×10^{39}	1.0×10^{8}
	8.884×10^{43}		NGC 4486 (M87)	3.257×10^{42}	3.4×10^{9}
NGC 205 (M110)	$<1.172\times10^{38}$	$<9.3\times10^{4}$	NGC 4593	6.535×10^{42}	8.1×10^{6}
NGC 598 (M33)	1.46×10^{39}	$<1.5\times10^{3}$	NGC 4594 (M104)	3.361×10^{40}	1.1×10^{9}
	1.902×10^{39}		NGC 4649	1.002×10^{41}	2.0×10^{9}
NGC 1068 (M77)	1.315×10^{41}	1.6×10^{7}		1.587×10^{41}	
	5.124×10^{41}		NGC 4697	9.141×10^{39}	1.2×10^{8}
	8.81×10^{41}		NGC 4945	5.447×10^{39}	1.1×10^{6}
NGC 3115	1.773×10^{39}	9.1×10^{8}	NGC 5548	2.182×10^{43}	1.23×10^{8}
NGC 3227	1.371×10^{42}	3.9×10^{7}		2.154×10^{43}	
	7.127×10^{41}			2.778×10^{43}	
NGC 3516	5.578×10^{41}	2.3×10^{7}	NGC 6251	1.71×10^{42}	5.4×10^{8}
	1.084×10^{43}		NGC 7469	1.699×10^{43}	6.5×10^{6}
NGC 3608	1.174×10^{40}	1.1×10^{8}		3.34×10^{43}	
NGC 3783	8.515×10^{42}	9.4×10^{6}		2.071×10^{43}	
	7.384×10^{42}		PG 0026+129	2.798×10^{44}	4.5×10^{7}
NGC 3998	1.425×10^{41}	5.6×10^{8}	PG 0052+251	4.766×10^{44}	2.2×10^{8}
NGC 4051	7.609×10^{41}	1.3×10^{6}	PG 1211+143	1.098×10^{44}	4.05×10^{7}
	7.624×10^{41}		PG 1411+442	5.049×10^{42}	8.0×10^{7}
NGC 4151	9.025×10^{42}	1.53×10^{7}	PG 1426+015 (Mrk 1383)	8.507×10^{43}	4.7×10^{8}
	5.578×10^{41}		PG 1613+658 (Mrk 876)	1.652×10^{44}	2.41×10^{8}
	4.784×10^{41}		PG 1617+175 (Mrk 877)	6.188×10^{43}	2.73×10^{8}

Note: The black hole masses were taken from [11].

$10^8\ M_\odot$) whose activity is low. This corresponds to the fact mentioned in Section 3 that the gravitational radius for massive black holes is larger than the tidal radius, and stars are swallowed entirely without being disrupted and, hence, without emitting energy. The exceptions (e.g., galaxy NGC 4486) probably correspond to the cases where there is much interstellar gas at the galactic center whose absorption can lead to substantial energy release [22].

ACKNOWLEDGMENTS

We thank Prof. A.M. Cherepashchuk and Prof. A.V. Zasov for helpful discussions. This work was supported by the Basic Research Program "Mathematical Methods in Nonlinear Dynamics" of the Presidium of the Russian Academy of Sciences, a grant from the President of Russia for support of leading scientific Schools (no. NSh-1603.2003.2), and INTAS grant no. 03-51-4286.

APPENDIX

Calculating the Diffusion Coefficients R_k and R_{kp} in (8) and the Adiabatic Invariant for an Isothermal Potential

Let us consider the coefficients W_{kp} and W_k defined in formula (4). If the distribution function f_* is isotropic, i.e., depending only on energy, then the tensor W_{kp} depends only on the vector **v** and, therefore, takes the form

$$W_{kp} = A(E, r)\delta_{kp} - B(E, r)\frac{v_k v_p}{v^2}; \qquad (36)$$

hence, calculating the convolutions W_{kk} and $W_{kp}v_p$, we obtain

$$A = \frac{8\pi}{3}\int\limits_{\Psi(r)}^{\infty} dE' f(E')\begin{cases}\frac{3}{2}\frac{v'}{v}\left(1-\frac{v'^2}{3v^2}\right), & E' < E\\ 1, & E' > E,\end{cases}$$

$$A - B = \frac{8\pi}{3}\int\limits_{\Psi(r)}^{\infty} dE' f(E')\begin{cases}\frac{v'^3}{v^3}, & E' < E\\ 1, & E' > E,\end{cases}$$

$$v = \sqrt{2(E-\Psi(r))}, \quad v' = \sqrt{2(E'-\Psi(r))}.$$

Similarly,

$$W_k = \frac{v_k}{v}D(E, r), \qquad (37)$$

$$D(E, r) = \frac{8\pi}{3}\int\limits_{\Psi(r)}^{\infty} dE'\frac{\partial f}{\partial E'}v\begin{cases}(v'/v)^3, & E' < E\\ 1, & E' > E.\end{cases}$$

To calculate the coefficients R_k and R_{kp} in (8), it is convenient to use the new variables

$$\xi = \{\xi_1, \xi_2\} = \{E, m\}.$$

Thus, $I_i = I_i(E, m)$, and we derive from (8)

$$R_{ij} = \frac{\partial I_i}{\partial \xi_a}\frac{\partial I_j}{\partial \xi_b}R'_{ab}, \quad R_i = \frac{\partial I_i}{\partial \xi_a}R'_a.$$

Denoting

$$\Lambda_0 = 2\pi G^2 M_*\Lambda, \quad \langle\ldots\rangle = \frac{1}{(2\pi)^3}\int d^3\phi$$

and using (36) and (37), we obtain for R'_a and R'_{ab}

$$R'_1 = \Lambda_0\langle vD\rangle, \quad R'_2 = \Lambda_0 m\langle D/v\rangle,$$

$$R'_{11} = \Lambda_0\langle v^2(A-B)\rangle, \quad R'_{12} = \Lambda_0 m\langle A-B\rangle, \quad (38)$$

$$R'_{22} = \Lambda_0\langle Ar^2 - Bm^2/v^2\rangle,$$

where

$$\langle\ldots\rangle = \frac{2}{T(E, m)}\int\frac{dr}{v_r}, \quad T(E, m) = \oint\frac{dr}{v_r},$$

$$v_r = \sqrt{2\left(E-\Psi(r)-\frac{m^2}{2r^2}\right)}.$$

All of the above results are applicable to an arbitrary isotropic distribution function. Let us now consider the special case of an isothermal function (12). It is convenient to use the dimensionless variables

$$\mu = \frac{m}{2\sigma e^{E/2\sigma^2}}, \quad x = r\exp\left(-\frac{E}{2\sigma^2}\right).$$

It follows from the condition

$$v_r^2 = 2(E - \Psi - m^2/2r^2) > 0$$

that

$$0 \le x \le 1, \quad 0 \le \mu \le 1/\sqrt{2e}.$$

Thus, for example, for the orbital period, we obtain

$$T(E, m) = \sqrt{2}\int\frac{dr\theta(E-\Psi-m^2/2r^2)}{\sqrt{E-\Psi-m^2/2r^2}}$$

$$= \frac{1}{\sigma}\exp(E/2\sigma^2)\tilde{T}(\mu),$$

$$\tilde{T}(\mu) = \int\limits_0^1\frac{dx\theta(-\ln x-\mu^2/x^2)}{\sqrt{-\ln x-\mu^2/x^2}}.$$

It can be verified that $\tilde{T}(\mu)$ is an almost constant quantity that slowly increases from 1.77 at $\mu = 0$ to 1.9 at $\mu \approx 0.4$, the maximum possible μ.

The averaging procedure can now be written as

$$\langle \ldots \rangle = \frac{2}{T(E,m)}\int_{r_-}^{r_+}\frac{dr}{v_r}(\ldots)$$

$$= \frac{1}{\tilde{T}(\mu)}\int_0^1 \frac{dx\theta(-\ln x - \mu^2/x^2)}{\sqrt{-\ln x - \mu^2/x^2}}(\ldots).$$

Calculating A, B, and D for the isothermal distribution function yields

$$A = \frac{8\pi}{3}\frac{f_0}{r^2}\sigma^2(-2\ln x(2H + x^2) - H + x^2),$$

$$A - B = \frac{8\pi}{3}\frac{f_0}{r^2}\sigma^2(2H + x^2),$$

$$\frac{D}{v} = -\frac{8\pi}{3}\frac{f_0}{r^2}(2H + x^2), \quad v^2 = -4\sigma^2\ln x,$$

$$H(x) = \int_x^1 y dy\left(\sqrt{\frac{\ln y}{\ln x}}\right)^3.$$

Note that

$$A - B = -\sigma^2 D/v,$$

which leads to expression (13). Using (38), we obtain

$$R'_{11} = -\sigma^2 R'_1, \quad R'_{12} = -\sigma^2 R'_2,$$

$$R'_1(E,m) = -4\sigma^2\exp\left(-\frac{E}{\sigma^2}\right)\Lambda_0\frac{8\pi}{3}f_0\Phi_1(\mu),$$

$$R'_2(E,m) = -m\exp\left(-\frac{E}{\sigma^2}\right)\Lambda_0\frac{8\pi}{3}f_0\Phi_2(\mu),$$

$$R'_{22}(E,m) = \sigma^2\Lambda_0\frac{8\pi}{3}f_0\Phi_{22}(\mu),$$

where

$$\Phi_2(\mu) = \langle 2H(x)/x^2\rangle + 1,$$

$$\Phi_1(\mu) = \langle -\ln x \times 2H(x)/x^2\rangle + \langle -\ln x\rangle,$$

$$\Phi_{22}(\mu) = \langle -2\ln x(2H + x^2) - H + x^2\rangle - \mu^2\left\langle\frac{4H}{x^2} + 1 + \frac{3H}{x^2\ln x}\right\rangle.$$

The following approximate equality holds for $\mu \le 10^{-4}$:

$$\Phi_2(\mu) \approx 1 + 0.35\mu^{-0.6},$$

$$\Phi_1(\mu) \approx 0.5 - 0.8\ln\mu(\Phi_2(\mu) - 1) \approx 0.5 - 0.3\ln\mu\mu^{-0.6},$$

$$\Phi_{22}(\mu) \approx 0.87.$$

Finally, let us calculate the adiabatic invariant I defined by formula (5):

$$I(E,m) = \frac{2\sigma}{\pi}\exp\left(\frac{E}{2\sigma^2}\right) \times \int_0^1 \sqrt{-\ln x - \frac{\mu^2}{x^2}}\,\theta\left(-\ln x - \frac{\mu^2}{x^2}\right)dx.$$

Calculating the integral yields

$$I(E,m) \approx \frac{2\sigma}{\pi}\exp\left(\frac{E}{2\sigma^2}\right)(0.9 - 2\mu)$$

$$= \frac{2}{\pi}\left(0.9\sigma\exp\left(\frac{E}{2\sigma^2}\right) - m\right),$$

therefore,

$$I'_m \approx -\frac{2}{\pi}, \quad I'_E \approx \frac{I + 2m/\pi}{2\sigma^2}.$$

The isothermal function in the variables I and m is

$$f(I,m) = f_0\exp\left(-\frac{E}{\sigma^2}\right) \approx f_0\sigma^2\left(\frac{2\times 0.9}{\pi}\right)^2\frac{1}{(I + 2m/\pi)^2}.$$

Another useful formula, the dependence of the orbital period T on I and m, is

$$T(I,m) = \frac{\tilde{T}(\mu)}{\sigma}\exp\left(\frac{E}{2\sigma^2}\right) \approx \frac{2}{\sigma}\frac{(\pi/2)I + 2m}{0.9\sigma}.$$

REFERENCES

1. S. K. Chakrabarti, in *Astrophysics and Space Science Library* (Kluwer Academic, Dordrecht, 1999), Vol. 234; L. Kaper, E. P. J. van den Heuvel, and P. A. Woudt, in *Proceedings of the ESO Workshop, Garching, Germany, 1999, in Honor of Riccardo Giacconi (ESO Astrophysical Symposia)* (Springer, Berlin, 2001).
2. V. I. Dokuchaev and Yu. N. Eroshenko, astro-ph/0209324v1.
3. A. V. Gurevich, Geomagn. Aeron. **4**, 247 (1964).
4. P. J. E. Peebles, Astrophys. J. **178**, 371 (1972).
5. J. N. Bahcall and R. A. Wolf, Astrophys. J. **209**, 214 (1976).

6. J. Frank and M. J. Rees, Mon. Not. R. Astron. Soc. **176**, 633 (1976).
7. A. P. Lightman and S. L. Shapiro, Astrophys. J. **211**, 244 (1977).
8. V. I. Dokuchaev and L. M. Ozernoĭ, Zh. Éksp. Teor. Fiz. **73**, 1587 (1977) [Sov. Phys. JETP **46**, 834 (1977)].
9. A. V. Gurevich and K. P. Zybin, Usp. Fiz. Nauk **165**, 723 (1995) [Phys. Usp. **38**, 687 (1995)]; Zh. Éksp. Teor. Fiz. **94** (10), 5 (1988) [Sov. Phys. JETP **67**, 1957 (1988)].
10. A. V. Gurevich and K. P. Zybin, Zh. Éksp. Teor. Fiz. **97**, 20 (1990) [Sov. Phys. JETP **70**, 10 (1990)].
11. A. M. Cherepashchuk, Usp. Fiz. Nauk **173**, 345 (2003) [Phys. Usp. **46**, 335 (2003)].
12. A. S. Il'in, K. P. Zybin, and A. V. Gurevich, Zh. Éksp. Teor. Fiz. **125**, 5 (2004) [JETP **98**, 1 (2004)]; astro-ph/0306490.
13. L. D. Landau, Zh. Éksp. Teor. Fiz. **7**, 203 (1937).
14. G. I. Budker and S. T. Belyaev, in *Plasma Physics and Problems of Controlled Thermonuclear Reactions* (Gostekhizdat, Moscow, 1958), Vol. 2.
15. L. D. Landau and E. M. Lifshitz, *The Classical Theory of Fields*, 6th ed. (Nauka, Moscow, 1973; Pergamon, Oxford, 1975).
16. K. Gebhardt, R. Bender, G. Bower, *et al.*, Astron. Astrophys. **539**, L13 (2000).
17. L. Ferrarese and D. Merritt, Astrophys. J. **539**, L9 (2000).
18. M. Baes, P. Buyle, G. K. T. Hau, and H. Dejonghe, Mon. Not. R. Astron. Soc. **341**, L44 (2003).
19. S. L. Shapiro, *Carnegie Observatories Astrophysical Series* (2003), Vol. 1; astro-ph/0304202v1.
20. D. Merritt and L. Ferrarese, astro-ph/0107134v2.
21. NASA/IPAC Extragalactic Database (NED), http://ned-www.ipac.caltech.edu.
22. M. Vestergaard, astro-ph/0401430.

Translated by V. Astakhov

Journal of Experimental and Theoretical Physics, Vol. 100, No. 2, 2005, pp. 305–310.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 127, No. 2, 2005, pp. 343–349.

ORDER, DISORDER, AND PHASE TRANSITIONS IN CONDENSED SYSTEMS

The Possible Mechanisms of Phase Transitions in RMn_2O_5 Oxide Family

A. M. Kadomtseva*, S. S. Krotov, Yu. F. Popov, G. P. Vorob'ev, and M. M. Lukina

Department of Physics, Moscow State University, Vorob'evy gory, Moscow, 119899 Russia

**e-mail: kadomts@plms.phys.msu.su*

Received June 15, 2004

Abstract—The magnetic, magnetoelectric, and magnetoelastic properties of $ErMn_2O_5$ single crystals have been studied at low temperatures and strong magnetic fields (up to 250 kOe) and compared to the analogous results obtained previously for YMn_2O_5. Based on these data, the possible mechanisms of various spontaneous and magnetic-field-induced phase transitions in these compounds are considered within the framework of the theory of representations of the *Pbam*-D_{2h}^9 space group. It is shown that a biquadratic exchange plays an important role in the formation (and mutual transformation) of magnetic structures revealed by neutron diffraction in the RMn_2O_5 oxide family. © *2005 Pleiades Publishing, Inc.*

1. INTRODUCTION

The family of antiferromagnetic (AFM) ferroelectrics representing RMn_2O_5 rare earth manganates (where R^{3+} is the ion of rare earth element from Nd to Lu, Y, or Bi) has received much attention of researchers. This interest is caused by special features of the magnetoelectric interactions in these compounds characterized by closely related magnetic and phase transitions.

According to neutron diffraction data [1], Mn^{3+} and Mn^{4+} ions situated in the *ab* plane of RMn_2O_5 exhibit spin ordering at $T_N \sim 40$ K and form a spatially modulated spin structure with the wavevector $\mathbf{k} = \{1/2, 0, \tau\}$. At low temperatures, an ordered spin structure in these compounds can be also formed by rare earth ions. Previously, we demonstrated [2] that YMn_2O_5 exhibits a field-induced electrical ordering at temperatures below the AFM phase transition temperature.

According to [3, 4], a number of RMn_2O_5 manganates (including YMn_2O_5) exhibit a ferroelectric phase transition at $T_C = 20$ K, whereby the spontaneous electric polarization is directed along the *b* axis of the crystal structure. In [2], we also observed new phase transitions induced by a strong magnetic field, but the nature of these transitions remains incompletely clear.

The mechanisms of the new phase transitions have also been studied in $ErMn_2O_5$, which is of interest both independently and in comparison with the data for YMn_2O_5 [2] for establishing the general laws in behavior of the RMn_2O_5 systems. Published data on the magnetic structure of manganese ions in $ErMn_2O_5$ are rather contradictory. According to Buisson [1], the spins of manganese ions in this compound form a helicoidal spin structure in the *ab* plane. Gardner *et al.* [5] observed an amplitude-modulated structure of these spins in the *ac* plane, while the spins of Er^{3+} ions were ordered along the *c* axis of the crystal. $ErMn_2O_5$ (in contrast to YMn_2O_5) exhibited no ferroelectric phase transition near 20 K: this transition was shifted toward higher temperatures.

In order to refine the character of the magnetic and electrical ordering in $ErMn_2O_5$ and to elucidate the general mechanisms of phase transitions in RMn_2O_5 manganates, it was of interest to study the magnetic, magnetoelectric, and magnetoelastic properties of $ErMn_2O_5$ single crystals in strong pulsed magnetic fields (up to 250 kOe) at low temperatures (10–50 K) and compare the results to the analogous data obtained previously for YMn_2O_5 [2].

As will be demonstrated below, many features of the interrelated magnetic and ferroelectric phase transitions in the compounds studied can be explained taking into account both the isotropic exchange interactions and a special antisymmetric exchange in Mn^{3+}–Mn^{3+}, Mn^{4+}–Mn^{4+}, and Mn^{3+}–Mn^{4+} ion pairs.

The intrinsic low-temperature ordering in the subsystem of rare earth ions R^{3+} (in cases where these ions possess magnetic moments) leads, by virtue of the anisotropy of the R–Mn exchange [6], to additional interesting features in the behavior of compounds under consideration.

2. THEORETICAL MODEL

To our knowledge, despite quite long and extensive investigation into the properties of the RMn_2O_5 systems [7, 8], there have been no attempts to construct the general picture of behavior of these systems featuring

a variety of phase transitions both in the presence and in the absence of applied fields. For this reason, it would be expedient, in addition to presenting the results of a new series of experiments, to consider in some detail the underlying theory as well. This consideration will be restricted to an approximation sufficient for explaining most of the experimental data reported so far.

Since the elementary cell of RMn_2O_5 contains four Mn^{3+} ions (4*h* position) and four Mn^{4+} ions (4*f* position), the magnetic properties of systems under consideration are conveniently described in terms of linear combinations of the spins ($\mathbf{S}_1$, $\mathbf{S}_2$, $\mathbf{S}_3$, $\mathbf{S}_4$) of Mn^{3+} ions and the corresponding magnetic modes of the spins ($\mathbf{S}_5$, $\mathbf{S}_6$, $\mathbf{S}_7$, $\mathbf{S}_8$) of Mn^{4+} ions, which are defined as

$$\begin{aligned}
&\mathbf{A} = \mathbf{S}_1 - \mathbf{S}_2 - \mathbf{S}_3 + \mathbf{S}_4, \quad \mathbf{A}' = \mathbf{S}_5 - \mathbf{S}_6 - \mathbf{S}_7 + \mathbf{S}_8,\\
&\mathbf{G} = \mathbf{S}_1 - \mathbf{S}_2 + \mathbf{S}_3 - \mathbf{S}_4, \quad \mathbf{G}' = \mathbf{S}_5 - \mathbf{S}_6 + \mathbf{S}_7 - \mathbf{S}_8,\\
&\mathbf{C} = \mathbf{S}_1 + \mathbf{S}_2 - \mathbf{S}_3 - \mathbf{S}_4, \quad \mathbf{C}' = \mathbf{S}_5 + \mathbf{S}_6 - \mathbf{S}_7 - \mathbf{S}_8,\\
&\mathbf{F} = \mathbf{S}_1 + \mathbf{S}_2 + \mathbf{S}_3 + \mathbf{S}_4, \quad \mathbf{F}' = \mathbf{S}_5 + \mathbf{S}_6 + \mathbf{S}_7 + \mathbf{S}_8.
\end{aligned} \tag{1}$$

The numbering of ions adopted in expressions (1) follows that used in the original study of Bertaut *et al.* [9] and then in [1]. In determining (within the framework of the theory of symmetry) the corresponding magnetic contributions to the total energy of the system, the symmetry and multiplicity of the lattice sites occupied by R^{3+} ions (4*g* position) show evidence in favor of introducing symmetrized combinations of the type (1) for the magnetically ordered R^{3+} ions as well. As will be shown below (by analogy with the manganese subsystem), the corresponding AFM order parameter is also characterized by a two-dimensional (2D) representation. The electric polarization response induced by virtue of the magnetoelectric interaction is quadratic with respect to the AFM order parameter [6]. The aforementioned symmetrized combinations, intended for the description of ordering in the spin system of R^{3+} ions, are also expediently used in consideration of the effect of exchange energies of the Mn^{3+}–R^{3+} and Mn^{4+}–R^{3+} ion pairs (*f–d* exchange) on the properties of RMn_2O_5 systems at temperatures above that of the intrinsic ordering of the rare earth ion subsystem (it can be readily shown that, in our case, this is related to equivalence of the corresponding irreducible representations).

In order to describe a change in the symmetry of the system experimentally observed [1] upon a phase transition at the Neél temperature T_N, we will use (according to the Landau approach [10]) the theory of space group representations (here, the *Pbam* space group) and the theory of exchange symmetry [11]. Using the method developed in [9, 11], the AFM structure with the wavevector $k = \{1/2, 0, \tau\}$ (where $\tau \approx 0.26$–0.50 according to the results of high-precision experiments on determining the magnetic structure in RMn_2O_5 [8]) appearing at the Neél point (T_N) will be described in the exchange approximation in terms of the 2D representation Γ_1 [9] constructed using two AFM vectors (**A** and **G**) implementing (according to the experimental data [1]) the main 2D AFM order parameter. One can readily check that the vector pairs {**C**, –**F**} and {**C**', **F**'} also correspond to the same 2D representation Γ_1 as that for the order parameter, while the vector pair {**G**', –**A**'} corresponds to the second 2D representation Γ_2 [9].

Now we will briefly outline the construction of a thermodynamic potential of RMn_2O_5 systems (for more detail, see [2, Ref. 13]). Taking into account bilinear exchange invariants, the "diagonal" contribution $\Delta\Phi_{ind}^{(2)}$ can be written (here, for the manganese subsystem only) as

$$\begin{aligned}
\Delta\Phi_{ind}^{(2)} = \frac{1}{2}\Lambda_1(A^2 + G^2) + \frac{1}{2}\Lambda_2(C'^2 + F'^2)\\
+ \frac{1}{2}\Lambda_3(A'^2 + G'^2) + \frac{1}{2}\Lambda_4(C^2 + F^2).
\end{aligned} \tag{2}$$

Mixed exchange invariants of the second order, responsible for the interaction between the spin subsystems, are as follows:

$$\begin{aligned}
\Delta\Phi_{ind}^{(2)} = \Lambda_{12}(\mathbf{A}\cdot\mathbf{C}' + \mathbf{G}\cdot\mathbf{F}') + \Lambda_{13}(\mathbf{A}\cdot\mathbf{C} - \mathbf{G}\cdot\mathbf{F})\\
+ \Lambda_{23}(\mathbf{C}\cdot\mathbf{C}' - \mathbf{F}\cdot\mathbf{F}').
\end{aligned} \tag{3}$$

Taking into account the experimental data [1] and using the main ideas of the Landau theory [10], we conclude that the following relations are valid at the phase transition temperature (T_N):

$$\begin{aligned}
\Lambda_1(T_N) = 0, \quad \Lambda_2(T_N) > 0,\\
\Lambda_3(T_N) > 0, \quad \Lambda_4(T_N) > 0.
\end{aligned} \tag{4}$$

In a "zeroth" approximation, these relations lead to two consequences immediately below T_N. First, the linear combinations of spins obey the conditions

$$\mathbf{C}' = \mathbf{F}' = \mathbf{A}' = \mathbf{G}' = \mathbf{F} = \mathbf{C} = 0. \tag{5a}$$

Second, taking into account that the spins must be unimodular (i.e., possess equal absolute values) [12], we obtain the condition

$$(\mathbf{A}\cdot\mathbf{G}) = 0. \tag{5b}$$

In application to the $\mathbf{S}_1$, $\mathbf{S}_2$, $\mathbf{S}_3$, $\mathbf{S}_4$ spins, this implies (provided that $A^2 \neq 0$ and $G^2 \neq 0$) that $\mathbf{S}_1 = -\mathbf{S}_2$, $\mathbf{S}_3 = -\mathbf{S}_4$, and $(\mathbf{S}_1 - \mathbf{S}_2) \perp (\mathbf{S}_3 - \mathbf{S}_4)$. Therefore, relations (5a) and (5b) lead to the conclusion that the AFM phase transition in the orthorhombic system studied result in the formation of a noncollinear exchange AFM structure (cross-like). The presence of interaction terms in the system energy (3) (and a more thorough analysis of the experimental data [1]) require using a more accurate approach.

Taking into account the above considerations, the subsequent analysis will be performed for the quadratic exchange contribution to the thermodynamic potential involving the following terms:

$$\Delta\tilde{\Phi}^{(2)} = \frac{1}{2}\Lambda_1(A^2 + G^2) + \Lambda_{12}(\mathbf{A}\cdot\mathbf{C}' + \mathbf{G}\cdot\mathbf{F}') + \frac{1}{2}\Lambda_2(C'^2 + F'^2). \quad (6)$$

In order to take into account the effects of inhomogeneous exchange, it was possible to include gradient exchange terms into expression (6) and consider the vector pairs {**A**, **G**} and {**C**', **F**'} as dependent on the *z* coordinate. However, as will become clear below, most of the features in the macroscopic behavior of RMn_2O_5 systems observed in the temperature interval under consideration can be explained within the framework of the main approximation. Thus, at this stage, we may neglect explicit contributions due to the inhomogeneous exchange. These contributions can be readily taken into account using perturbation theory, but this paper does not contain such an analysis.

In order to separate the "true," independent magnetic modes in the system under consideration (via rotation in the spin space), the above expression (6) has to be diagonalized. As a result, the ground state of the system will represent a certain combination of simultaneously ordering spins of Mn^{3+} and Mn^{4+} ions, where the spin subsystem of Mn^{3+} ions plays a seeding role. In what follows, we assume that the corresponding trivial procedures are accomplished and the resulting second-order exchange contribution $\Delta\Phi^{(2)}$ in the new variables is represented as

$$\Delta\Phi^{(2)} = \frac{1}{2}\tilde{\Lambda}_2(\tilde{A}^2 + \tilde{G}^2), \quad (7)$$

where the variable vector pair $\{\tilde{\mathbf{A}}, \tilde{\mathbf{G}}\}$ also forms a basis of the 2D representation Γ_1 (the consideration is restricted to the exchange contribution due to the main ordering mode).

Besides the obvious (stabilizing) fourth-order exchange contribution

$$\Delta\Phi_1^{(4)} = \frac{1}{2}\Lambda_4^{(1)}(\tilde{A}^2 + \tilde{G}^2)^2, \quad (8)$$

we will also include (taking into account close energies of various magnetic phases of the RMn_2O_5 system [1]) the specific biquadratic homogeneous exchange [12]

$$\Delta\Phi_2^{(4)} = \frac{1}{4}\Lambda_4^{(2)}(\tilde{A}^2 - \tilde{G}^2)^2. \quad (9)$$

The origin of the latter term is related to the fact that the $\tilde{A}^2 - \tilde{G}^2$ function is a basis of the 1D representation of the *Pbam* space group [9], in which the P_y component of the electric polarization vector is transformed (see below). The term (9) plays a significant role in explaining the neutron diffraction data [1] (including the presence of magnetic phases of three types) and, which is especially important, in interpreting the experimental data concerning the specific spin reorientation (at $T = T^*$, where $\Lambda_4^{(2)}(T^*) = 0$) with the formation of a two-domain magnetic state (either $\tilde{\mathbf{G}} \neq 0$, $\tilde{\mathbf{A}} = 0$ or $\tilde{\mathbf{G}} = 0$, $\tilde{\mathbf{A}} \neq 0$).

As can be readily checked, the aforementioned magnetic transformation at T^* involving the isotropic-exchange magnetoelectric contribution

$$\Delta\Phi_{ME}^{(ex)} = \alpha_1 P_y(\tilde{A}^2 - \tilde{G}^2) \quad (10)$$

is accompanied by spontaneous appearance of the electric polarization vector (induced ferroelectric transition), with the opposite directions of the electric polarization components P_y in various magnetic domains. Apparently, the contribution (10) vanishes at $T > T^*$, where $\Lambda_4^{(2)} > 0$ and $\tilde{\mathbf{A}} = \tilde{\mathbf{G}}$. It should be emphasized that the contribution (10) is substantially related to the 2D character of the order parameter and, in particular, the inversion transformation $\bar{1}$ leads to the following transformations of the basis set functions: $\bar{1}\tilde{\mathbf{A}} = -\tilde{\mathbf{G}}$ and $\bar{1}\tilde{\mathbf{G}} = -\tilde{\mathbf{A}}$, so that $\bar{1}(\tilde{A}^2 - \tilde{G}^2) = \tilde{G}^2 - \tilde{A}^2$. The invariance of expression (10) can be also readily checked for any other two generators of the *Pbam* space group (see [9]).

It should be also noted that, below the Neél temperature T_N (i.e., at $T^* < T < T_N$), the relativistic interactions in the presence of the antisymmetric exchange contribution

$$\Delta\Phi_{ME}^{(rel)} = \alpha_2 P_y[\tilde{\mathbf{G}}\times\tilde{\mathbf{A}}]_z \quad (11)$$

also lead to the induced spontaneous polarization P_y. Thus, it is a general result that RMn_2O_5 systems exhibit an intrinsic AFM transition and, by virtue of contribution (1), an induced ferroelectric phase transition at the Neél point. In accordance with experiment, relation (11) takes into account that the "AFM cross" of $\mathbf{S}_1$, $\mathbf{S}_2$, $\mathbf{S}_3$, and $\mathbf{S}_4$ spin occurs (due to anisotropy) in the *ab* plane.

In the case where spins of the manganese ion subsystem at the Neél point occur in the *ac* plane, the electric polarization vector cannot appear via mechanism (11), but then the spontaneous electric polarization will appear by virtue of mechanism (10) at a lower temperature corresponding to magnetic reorientation. As can be readily seen from the form of the magnetoelectric contributions $\Delta\Phi_{ME}^{(rel)}$ and $\Delta\Phi_{ME}^{(ex)}$, these terms are retained in cases where inhomogeneity of the magnetic structure along the *c* axis has a helicoidal character (with spins occurring in the *ab* plane). The exchange contribution also retains its form upon the formation of

an amplitude-modulated magnetic structure (with spins occurring in the *ac* plane).

From the standpoint of the general pattern of spin rearrangement, the aforementioned spin reorientation transition at T^* apparently consists in the following [12]. When an antiferromagnetic state appears at T_N, where

$$G^2 = A^2, \quad \mathbf{G} \perp \mathbf{A} \tag{12}$$

(which corresponds to phase Ia according to [1]), the pairs of AFM-ordered spins $\mathbf{S}_1$, $\mathbf{S}_2$ and $\mathbf{S}_3$, $\mathbf{S}_4$ are oriented mutually perpendicularly to form the exchange cross: $(\mathbf{S}_1 - \mathbf{S}_2) \perp (\mathbf{S}_3 - \mathbf{S}_4)$, $|\mathbf{S}_1 - \mathbf{S}_2| = |\mathbf{S}_3 - \mathbf{S}_4|$. As can be readily shown, the magnetic susceptibility in this (equiarmed cross) phase according to the general thermodynamic considerations is isotropic, which implies the thermodynamic relation $\chi_\perp = \chi_\parallel$.

As the temperature decreases and the $\Lambda_4^{(2)}$ value passes through zero at $T = T^*$, the system apparently exhibits intrinsic spin flop, whereby the AFM-coupled $\mathbf{S}_3$, $\mathbf{S}_4$ pair aligns in the direction of the $\mathbf{S}_1$, $\mathbf{S}_2$ pair. If $\mathbf{S}_3 \sim \mathbf{S}_1$ and $\mathbf{S}_4 \sim \mathbf{S}_2$, we obtain the **G** domain state (corresponding to phase IIa according to [1]); by the same token, for $\mathbf{S}_3 \sim -\mathbf{S}_1$ and $\mathbf{S}_4 \sim -\mathbf{S}_2$, we obtain the **A** domain state (corresponding to phase IIb according to [1]). From this it follows, in particular, that the system under consideration will exhibit sharply pronounced magnetoelectric properties near the temperature $T \approx T^*$, where $\Lambda_4^{(2)} \geq 0$. On the other hand, this explains the role of the electric annealing field $\mathbf{E}_b$ observed below T^* [3]. This field, converting the sample (by means of magnetoelectric interactions) into a monodomain state, reveals the presence of two ferroelectric domains with $P_Y > 0$ and $P_Y < 0$ [3].

In the case of an RMn_2O_5 system with R = Er, the strongly anisotropic behavior of Er^{3+} rare earth ions whose spins are aligned in the *c* axis makes possible the orientation transition for spins of the Mn subsystem. Indeed, the spins of Mn^{3+} and Mn^{4+} ions can rearrange via the anisotropic *f–d* exchange from the initial *ab* plane to the *ac* plane. The stronger the *f–d* exchange in the system, the higher the temperature of this transition (which may even exceed the temperature of intrinsic ordering in the system of rare earth ions).

Taking into account the above considerations, it is also possible to explain features observed in the behavior of the magnetoelectric susceptibility in the region of a low-temperature ($T_{N2} \sim 10$ K) intrinsic magnetic ordering in a rare earth ion system [13], where the additional electric polarization varied with the temperature as the square of the magnetic order parameter.

3. EXERIMENTAL RESULTS AND DISCUSSION

The samples of RMn_2O_5 single crystals were grown by method of spontaneous crystallization [14]. The magnetization curves in strong pulsed magnetic fields were measured using the induction technique. The electric polarization and magnetostriction in strong magnetic fields (up to 250 kOe) were studied as described in [15]. We have studied the magnetic, magnetoelectric, and magnetoelastic properties of $ErMn_2O_5$ single crystals and compared the results to the analogous data obtained previously for YMn_2O_5 single crystals [2].

First, let us consider the magnetic-field-induced phase transition in YMn_2O_5 with nonmagnetic rare earth ions. YMn_2O_5 single crystals studied in [2] showed some new low-temperature phase transitions induced by a strong magnetic field $H \parallel b$. The nature of these transitions has remained insufficiently clear. It was pointed out that the magnetic-field-induced longitudinal electric polarization $P_b(H_b)$ along the *b* axis of the crystal exhibited a significantly different character in the temperature intervals $20\ \text{K} \leq T \leq T_N = 40\ \text{K}$ and $10\ \text{K} \leq T \leq 20\ \text{K}$. In the 10–20 K interval, the $P_b(H_b)$ value exhibited a jumplike increase at a certain critical field H_{cr}. At higher temperatures within the 20–40 K interval, the $P_b(H_b)$ value was small and exhibited a quadratic dependence on H_b up to 250 kOe. The observed change in behavior of the electric polarization as a function of the magnetic field at high ($T > 20$ K) and low ($T < 20$ K) temperatures suggests that the magnetic structures of YMn_2O_5 in these temperature intervals are different.

According to the neutron diffraction data [1], the magnetic structure of RMn_2O_5 admits the existence of magnetic phases of three types and features phase transitions between these phases depending on the temperature. Below T_N, the magnetic structure of YMn_2O_5 is modulated in the *ab* plane [1] and, according to the theory outlined above, can be characterized in the exchange approximation by a 2D order parameter constructed using two AFM vectors, **A** and **G**, obeying the relations $|\mathbf{G}| = |\mathbf{A}|$ and $(\mathbf{G} \cdot \mathbf{A}) = 0$ (this structure corresponds to type I according to [1]). Since this phase has the isotropic magnetic susceptibility χ ($\chi_\parallel = \chi_\perp$), no magnetic-field-induced phase transitions are possible, which is confirmed by the results of our experiments at 20–40 K. As the temperature is decreased to $T^* = 20$ K, a term corresponding to the spin reorientation transition to phase II probably appears in the expression for the biquadratic exchange $\Lambda_4^{(2)}(T^*) = 0$ [1, Eq. (4)]. In this case, $\mathbf{A} = 0$ and $\mathbf{G} \neq 0$, so that the exchange cross is formed with **G** and **F** vectors having different absolute values ($|\mathbf{G}| \gg |\mathbf{F}|$, $(\mathbf{F} \cdot \mathbf{G}) = 0$). This creates prerequisites for the spin reorientation transitions induced in the magnetic field $H \parallel b$, which are accompanied by jumps in the electric polarization observed in [2].

In the case of $ErMn_2O_5$, it is necessary to consider, in addition to the above mechanisms leading to rearrangement of the spin structure of manganese ions, an additional anisotropy due to the Er–Mn interaction. The role of this interaction especially increases at low tem-

peratures. Orientation of the magnetic moment of Er^{3+} ions in the $I_{15/2}(4f)$ ground state is determined by the crystal field, which aligns the spins of Er^{3+} in the *c* axis direction. The exchange field acting upon the rare earth ions in $ErMn_2O_5$ is significantly weaker than the anisotropic field. This makes possible the metamagnetic transitions of the spin flop type in the AFM sublattices of rare earth ions. During these metamagnetic transitions, the spins of manganese ions are also reoriented due to the Er–Mn interaction.

The magnetization curves of $ErMn_2O_5$ measured along the *c* axis (the direction of alignment for the spins of Er^{3+} ions) showed a sharp increase in the magnetic moment (up to 100 G cm^3/g) at $T = 10$ K in strong magnetic fields (Fig. 1). This behavior is associated with a metamagnetic transition of the ordered spins of Er^{3+} ions and the accompanying rearrangement of the manganese subsystem. The magnetization of $ErMn_2O_5$ along the *a* and *b* axes was much smaller and exhibited a linear dependence on the magnetic field strength (Fig. 1). In the field $H \parallel b$ at $T < T_N = 41$ K, the electric polarization $P_b(H_b)$ exhibited a quadratic dependence on the magnetic field and showed no evidence of field-induced anomalies. These results probably indicate that the magnetic subsystem of $ErMn_2O_5$ (as well as that of YMn_2O_5) is ordered at high temperatures according to type I [1]. By virtue of the isotropic magnetic susceptibility ($\chi_\parallel = \chi_\perp$), the equiarmed cross of the spins of manganese ions does not favor magnetic-field-induced phase transitions for $H \parallel b$. When the $ErMn_2O_5$ crystal was magnetized along the *c* axis, it led to the appearance of electric polarization in various crystallographic directions. This polarization was maximum along the *b* axis and exhibited complicated dependence on field and temperature (Fig. 2). In the region of low temperatures ($10 \le T < 20$ K) and weak magnetic fields (about 20 kOe), we observed a negative anomaly in the polarization manifested by a sharp change in the sign with increasing field (Fig. 2a). This behavior is probably related to rearrangement of the magnetic structures of both erbium and manganese subsystems, as also evidenced by the character of the field dependence of the magnetostriction in this temperature interval (Fig. 3). As the field was further increased, the polarization changed sign again (see the curves for $T = 15$ and 16 K in Fig. 2a), which probably reflects variations in the relative contributions of different subsystems.

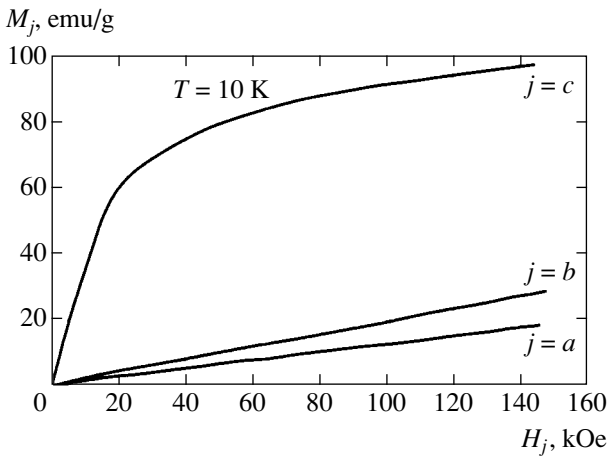

Fig. 1. Magnetization curves measured along the *a*, *b*, and *c* axes of an $ErMn_2O_5$ single crystal at $T = 10$ K.

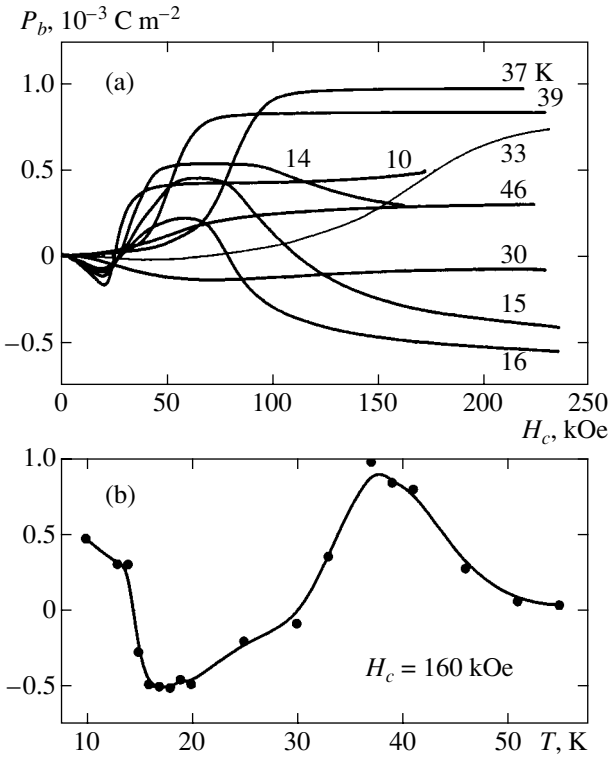

Fig. 2. Plots of the electric polarization P_b along the *b* axis of an $ErMn_2O_5$ single crystal versus (a) the magnetic field H_c oriented along the *c* axis and (b) the temperature at a fixed field strength $H_c = 160$ kOe.

It should be noted that anomalies were observed not only in the region of low temperatures, but at $30 < T < 40$ K as well. In this interval, the polarization exhibited jumps at $H = H_{cr}$, which were probably related to spin reorientation (Fig. 2a). The presence of field-induced anomalies in the case of $H \parallel c$ at temperatures below T_N apparently indicates that spin reorientation in the manganese subsystem under the action of Er–Mn exchange is accompanied by deviation of spins from the *ab* plane toward the *c* axis. It is natural to assume that the spins of manganese ions are reoriented with decreasing temperature near $T_R = 30$ K from the *ab* to the *ac* plane as a result of the Er–Mn interaction. This behavior is consistent with the neutron diffraction data for $ErMn_2O_5$ single crystals [5], according to which the spins of Mn^{3+} and Mn^{4+} ions in the temperature range 4.2–25 K form an amplitude-modulated structure in the *ac* plane, where the spin orientation angle decreases with increasing temperature.

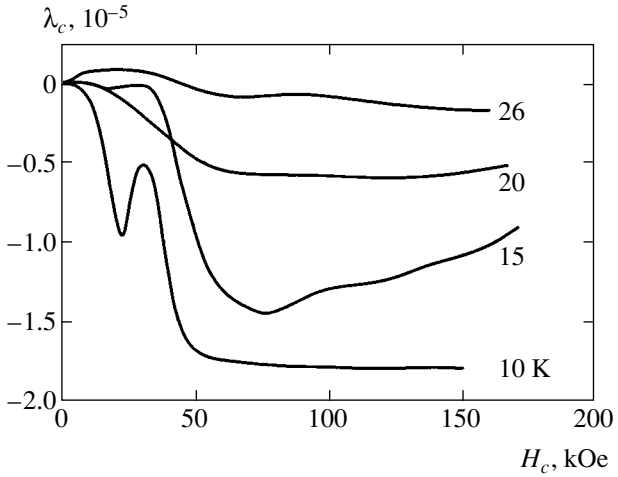

Fig. 3. Plots of the longitudinal magnetostriction λ_c versus magnetic field along the *c* axis in an $ErMn_2O_5$ single crystal at various temperatures.

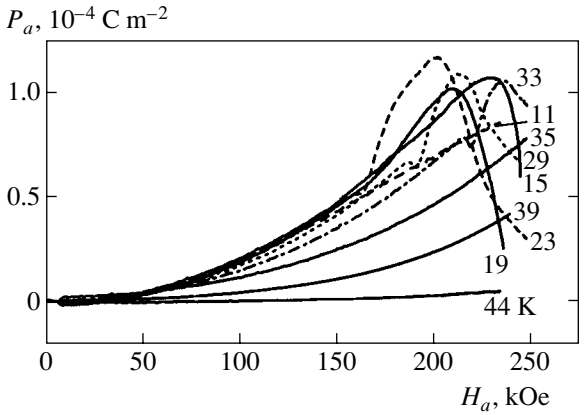

Fig. 4. Plots of the longitudinal electric polarization P_a along the *a* axis of an $ErMn_2O_5$ single crystal versus the magnetic field H_a along this axis at various temperatures.

Application of magnetic field along the *c* and *a* axes changes the orientation of the spins of manganese ions and may lead to anomalies in the field dependence of the electric polarization. It should be noted that such anomalies in $P_a(H_a)$ for $H \parallel a$ were observed only in the temperature interval of $T = 20$–33 K in strong magnetic fields of about 200 kOe (Fig. 4). Correctness of the assumptions concerning changes in the magnetic structure of $ErMn_2O_5$ as dependent on field and temperature is confirmed by the behavior of $P_b(T)$ measured at a fixed magnetic field of 160 kOe (Fig. 2b). As can be seen, the electric polarization exhibits sharp variation at low temperatures (near 15 K), which is related to the rearrangement of both erbium and manganese subsystems. Then, polarization increases with temperature up to $T_C \approx 37$ K and eventually decreases again due to the transition to a paramagnetic state. As can be seen from Fig. 2b, the electric polarization changes sign near $T_R \approx 30$ K, which is probably related to reorientation of the spins of Mn ions from the *ab* plane to the *ac* plane (similar to that observed for YMn_2O_5 in [2]). Note that the above conclusion concerning spin reorientation from the *ab* plane to the *ac* plane with decreasing temperature in $ErMn_2O_5$ does not remove the discrepancy between [1] and [5], according to which a modulated structure in the manganese subsystem of this compound is formed in different planes.

To summarize, the results of our experimental and theoretical investigation have revealed the mechanisms responsible for the formation of various spontaneous and magnetic-field-induced phase transitions in AFM ferroelectrics and provided an explanation of the magnetoelectric nature of these phenomena.

ACKNOWLEDGMENTS

This study was supported in part by the Russian Foundation for Basic Research, project no. 04-02-16592a.

REFERENCES

1. G. Buisson, Phys. Status Solidi A **16**, 533 (1973); **17**, 191 (1973).
2. Yu. F. Popov, A. M. Kadomtseva, S. S. Krotov, *et al.*, Zh. Éksp. Teor. Fiz. **123**, 1090 (2003) [JETP **96**, 961 (2003)].
3. A. Ikeda and K. Kohn, Ferroelectrics **169**, 75 (1995).
4. A. Inomata and K. Kohn, J. Phys.: Condens. Matter **8**, 2673 (1996).
5. P. Gardner, C. Wilkinson, J. B. Forsyth, *et al.*, J. Phys. C: Solid State Phys. **21**, 5653 (1988).
6. Y. Koyata and K. Kohn, Ferroelectrics **204**, 115 (1997).
7. I. Kagomiya, M. Yoshina, S. Matsumoto, *et al.*, in *Proceedings of 8th International Conference on Ferrites* (2001), p. 229; I. Kagomiya, K. Kohn, and T. Uchiyama, Ferroelectrics **280**, 131 (2002).
8. S. Kobayashi, T. Osawa, H. Kimura, *et al.*, J. Phys. Soc. Jpn. **73**, 1031 (2004).
9. E. F. Bertaut, G. Buisson, S. Quezel-Ambrunaz, and G. Quezel, Solid State Commun. **5**, 25 (1967).
10. J.-C. Toledano and P. Toledano, *The Landau Theory of Phase Transitions* (World Sci., Singapore, 1987; Mir, Moscow, 1994).
11. A. F. Andreev and V. I. Marchenko, Usp. Fiz. Nauk **130**, 39 (1980) [Sov. Phys. Usp. **23**, 21 (1980)].
12. I. M. Vitebsky, N. M. Lavrinenko, and V. L. Sobolev, J. Magn. Magn. Mater. **97**, 263 (1991).
13. I. Kagomiya, H. Kimura, Y. Noda, and K. Kohn, J. Phys. Soc. Jpn. A **70**, 145 (2001).
14. V. A. Sanina, A. M. Sapozhnikova, E. I. Golovenchits, and N. V. Morozov, Fiz. Tverd. Tela (Leningrad) **30**, 3015 (1988) [Sov. Phys. Solid State **30**, 1736 (1988)].
15. Yu. Popov, A. M. Kadomtseva, D. V. Belov, *et al.*, Pis'ma Zh. Éksp. Teor. Fiz. **69**, 302 (1999) [JETP Lett. **69**, 330 (1999)].

Translated by P. Pozdeev

Journal of Experimental and Theoretical Physics, Vol. 100, No. 2, 2005, pp. 311–313.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 127, No. 2, 2005, pp. 350–352.

ORDER, DISORDER, AND PHASE TRANSITIONS IN CONDENSED SYSTEMS

On Wetting Transition in Superconductivity

V. I. Marchenko and E. R. Podolyak

Kapitza Institute for Physical Problems, Russian Academy of Sciences, Moscow, 119334 Russia
e-mail: mar@kapitza.ras.ru

Received September 27, 2004

Abstract—The existence of two solutions for proximity-induced superconductivity described by the Ginzburg–Landau theory is established in the general case. A first-order phase transition can occur between these states, resulting in superconductive wetting. © *2005 Pleiades Publishing, Inc.*

It has been commonly believed that a critical wetting transition must occur when two superconductors characterized by different critical temperatures are brought into contact: as the critical magnetic field strength is approached, the proximity-induced superconducting layer gradually expands into the bulk of the weaker superconductor [1]. However, it was found that the normal state of aluminum brought into contact with tin or tantalum can be substantially overcooled [2].

To resolve the controversy, we have analyzed the proximity effect by applying the Ginzburg–Landau theory. We have found that a first-order phase transition can occur in the proximity-induced superconducting layer, with a jump in the wetting-layer thickness, and that wetting can take place only in the phase characterized by the larger thickness.

The possibility of diverse junction behavior can be demonstrated under quite general assumptions. Suppose that the wetting-layer thickness L increases indefinitely as the magnetic field H approaches the critical value H_c. Then, the system can be described by a macroscopic model.

The order parameter has the equilibrium bulk value almost everywhere in the thicker layer, deviating from it only within a distance on the order of the coherence length ξ from the layer boundaries. Owing to exponential decay of these deviations into the bulk of the layer, the NS interface interacts with the junction. Thus, the energy of the proximity-induced superconducting layer can be represented as (cf. [3])

$$\sigma_{NS} + \sigma_{SS} + \frac{H^2 - H_c^2}{8\pi}L + \beta e^{-L/\tilde{\xi}}, \qquad (1)$$

where $\tilde{\xi} = \sqrt{2}\,\xi$ (in conventional notation [4]), σ_{NS} is the energy of the NS interface, and σ_{SS} is the junction energy. Note that, if the Ginzburg–Landau parameter κ is close to $1/\sqrt{2}$, then the term proportional to $\exp(-L/\delta)$ must also be included to allow for magnetic field penetration.

When the NS interface is repelled from the junction ($\beta > 0$), the following wetting law is obtained by minimizing the energy given by (1):

$$L \propto \xi \ln \frac{H_c}{H - H_c}. \qquad (2)$$

When the NS interface is attracted to the junction ($\beta < 0$), the wetting-layer thickness remains finite as the field strength approaches the critical value. This state can be described only by microscopic theory.

Suppose that, in addition to the "long" state, there exists a locally stable proximity-induced superconductive state characterized by a finite thickness. A first-order phase transition between these states of the proximity-induced superconducting layer can occur across a curve $H_s(T) > H_c(T)$. Since both phases have the same symmetry, the curve $H_s(T)$ can terminate at some critical point, as is common for first-order transitions. However, when the curves $H_s(T)$ and $H_c(T)$ meet at some point (wetting point) (H_{WT}, T_{WT}), the system exhibits an uncommon behavior that can also be analyzed in the framework of a macroscopic model.

On the phase-transition curve, the energy of the "long" (wetting) solution given by (1) equals the energy σ of the "short" (nonwetting) one. In the vicinity of the wetting transition point, the linear magnetic-field dependence of σ can be neglected as compared to the stronger dependence expressed by (1). The temperature dependence can also be neglected by setting $T = T_{WT}$ everywhere except for

$$\sigma - \sigma_{NS} - \sigma_{SS} = C(T - T_{WT}), \qquad (3)$$

where the constant $C > 0$, in agreement with the phase diagram obtained by computing the Ginzburg–Landau equation (see below). Minimizing (1) with respect to L

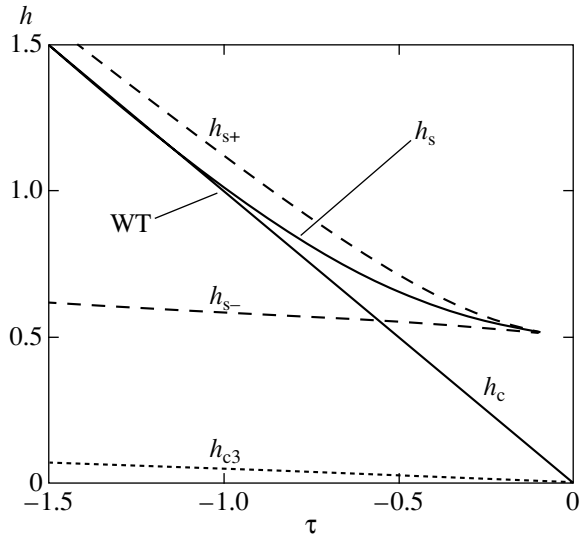

Fig. 1.

and matching the energy of both states, we obtain the following expression describing the behavior of the phase equilibrium curve in the neighborhood of T_{WT}:

$$H - H_c \propto -\frac{T - T_{WT}}{\ln(T - T_{WT})}. \quad (4)$$

Now, we analyze the proximity effect and explore the applicability of the Ginzburg–Landau theory to the phase transition. We consider the simplest case of states that are uniform in the junction plane. Then, the Ginzburg–Landau equations reduce to two equations for the ψ-function and the vector potential A (Eqs. (46.8) and (46.9) in [4]).

To formulate boundary conditions at the junction, we must take into account only the term linear in the order parameter ψ of the weaker superconductor,

$$\lambda(\psi^* e^{i\alpha} + \psi e^{-i\alpha}). \quad (5)$$

Here, λ represents the coupling between the superconductors, and α is the value of the order-parameter phase for the stronger superconductor near the junction.

When $\lambda < 0$, the order-parameter phases are equal at the boundary; when $\lambda > 0$, they differ by π (this case is known as π-junction [5]). We note here that the Ginzburg–Landau theory does not make any qualitative distinction between these states. The only observable distinction is the presence of a ubiquitous vortex carrying one-half of the magnetic-flux quantum near the curve where λ changes sign in inhomogeneous structures.

By virtue of gauge invariance at field strengths that are low for the stronger superconductor, we can set $\alpha = 0$. Furthermore, near the bulk phase transition point, where all characteristic lengths in the weaker superconductor increase, the finiteness of the magnetic penetration depth can be ignored. Then, the boundary condition naturally derived by varying the energy reduces to

$$\frac{\hbar^2}{4m}\partial_x\psi = \lambda. \quad (6)$$

The x axis is directed along the normal to the boundary into the weaker superconductor. Note that the ψ-function and its derivative at the boundary have opposite signs. In dimensionless quantities (see [4]), the boundary condition is rewritten as

$$\psi' = \kappa\Lambda, \quad \Lambda = 2\lambda\frac{\sqrt{mb}}{\hbar|a|}. \quad (7)$$

The structure of the boundary between the superconductors is readily calculated in the absence of magnetic field:

$$\psi = -\coth\frac{\kappa(x + c)}{\sqrt{2}} \quad (8)$$

(it is assumed that $\lambda > 0$). The constant c is determined by using the boundary condition. The corresponding SS-junction energy contained in (1) is

$$\sigma_{SS} = \frac{H_c^2\xi}{3\sqrt{2}\pi}(1 - |\psi_0|^3), \quad (9)$$

where ψ_0 is the value of the ψ-function at the boundary.

We have performed a numerical analysis of the proximity effect and determined the domain of stability for the solutions obtained. Depending on κ, magnetic field, and temperature, there exist either one or two (locally) stable solutions, and a phase transition between them can occur.

Figure 1 shows the phase diagram for proximity-induced superconductivity in aluminum calculated for $\kappa = 0.02$ in the case when the wetting transition occurs in the neighborhood of the critical point T_c. The phase diagram is universal in the coordinates $h = H/H_{WT}$ and $\tau = (T - T_c)/(T_c - T_{WT})$. The equilibrium solution below the transition curve $h_s(\tau)$ corresponds to wetting.

Figure 2 shows the effective thickness of the surface solution

$$L = \int_0^\infty |\psi|^2 dx$$

plotted versus magnetic field strength at a temperature above the wetting transition point T_{WT} $\tau = -0.38$.

At field strengths above the curve $h_s(\tau)$, the equilibrium solution remains its finite thickness. When the stability boundary for the "short" solution h_{s-} lies below h_c, superconductivity can penetrate into the bulk of the weaker superconductor only after a critical nucleus forms either in the bulk or at the junction under the so-called incomplete wetting conditions.

When the Ginzburg–Landau parameter is small, both short and long solutions obtained in the neighborhood of the transition point admit analytical description. In particular, the coordinates of the point (H_{WT}, T_{WT}) can be found. The short solution is characterized by the thickness $\sqrt{\xi\delta}$, and the corresponding ψ-function amplitude is small ($\propto\sqrt{\kappa}$, see the considerations that follow after Eq. (46.14) in [4]). The energy of the short solution is small as compared to that of the SS-junction energy σ_{SS} given by (9) and the NS-interface energy given by (9) with $\psi_0 = 0$ (see [4, Eq. (46.14)]). Therefore, the wetting transition occurs at the point defined by the condition $\sigma_{SS} + \sigma_{NS} = 0$, which yields the value of the ψ-function at the SS junction for the long solution at the wetting transition point, $|\psi_0| = 2^{1/3}$, and $\Lambda = 2^{1/6} - 2^{-1/2}$.

The parameter Λ increases as $(T_c - T)^{-1}$ toward the critical temperature, and wetting must always be observed in the neighborhood of the superconducting transition, in agreement with [1]. The wetting transition can occur in the close neighborhood of the critical temperature, where it can be described by the Ginzburg–Landau theory, only if the parameter λ is small, i.e., when the superconductors are weakly coupled. If λ is too small, (the domain of equilibrium wetting lies too close to T_c and $\xi(T)$ is on the order of ξ_0^2/a, where a is an atomic distance), then the term linear in the ψ-function must be retained in boundary condition (6) (see [6]). In this case, the phase diagram may not contain the domain of wetting solution.

A numerical analysis shows that the wetting transition discussed here can occur only in type I superconductors, whereas there exists only one surface solution for type II superconductors, which has a finite thickness until H_{c2} is reached.

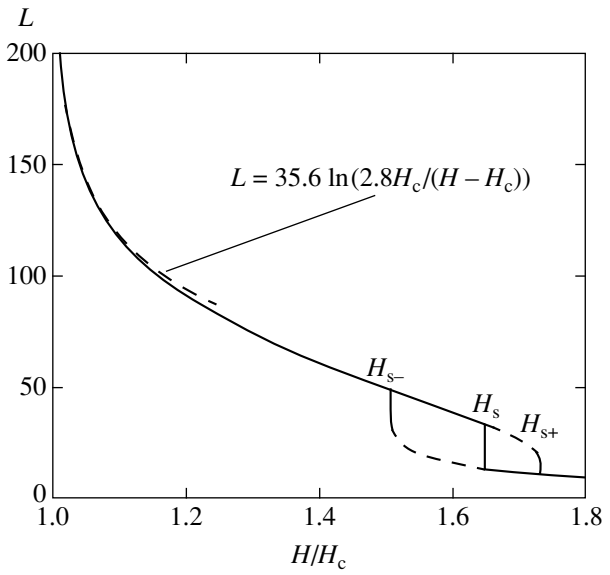

Fig. 2.

ACKNOWLEDGMENTS

We thank I.N. Khlyustikov for helpful discussion. This work was supported by the Russian Foundation for Basic Research, project no. 03-02-16958.

REFERENCES

1. R. A. Buhrman and W. P. Halperin, J. Low Temp. Phys. **16**, 409 (1974).
2. I. N. Khlyustikov, Zh. Éksp. Teor. Fiz. **112**, 1119 (1997) [JETP **85**, 609 (1997)].
3. V. I. Marchenko and E. R. Podolyak, Zh. Éksp. Teor. Fiz. **124**, 172 (2003) [JETP **97**, 154 (2003)].
4. E. M. Lifshitz and L. P. Pitaevskiĭ, *Statistical Physics* (Fizmatlit, Moscow, 2000; Butterworth, Oxford, 1998), Part 2.
5. A. F. Andreev, Pis'ma Zh. Éksp. Teor. Fiz. **46**, 463 (1987) [JETP Lett. **46**, 584 (1987)].
6. C. Caroli, P. G. De Gennes, and J. Marticon, J. Phys. Radium **23**, 707 (1962).

Translated by A. Betev

Journal of Experimental and Theoretical Physics, Vol. 100, No. 2, 2005, pp. 314–321.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 127, No. 2, 2005, pp. 353–361.

ELECTRONIC PROPERTIES OF SOLIDS

Hopping Transport in Two-Dimensional Systems with Spin–Orbit Interaction in External Magnetic Field

V. V. Bryksin

Ioffe Physicotechnical Institute, Russian Academy of Sciences, St. Petersburg, 194021 Russia
e-mail: vvb@mail.ioffe.ru

Received May 11, 2004

Abstract—The theory of spin rotation waves (SRWs), representing excitations of a new type arising in two-dimensional systems with spin–orbit interaction in an external electric field, has been developed. These intrinsic modes correspond to rotation of the magnetic moment vector in the plane formed by the electric field vector and the normal to the sample plate surface. An experimental method is proposed for detecting SRWs by measuring the frequency dependence of the magnetic susceptibility, which exhibits a resonance at the intrinsic mode frequency. A particular calculation is performed for a hopping conductivity model (for small-size polarons), but it is likely that intrinsic oscillations of the SRW type also take place for the band transport, since their appearance is related to the symmetry of the system. © *2005 Pleiades Publishing, Inc.*

1. INTRODUCTION

The possibility of providing effective control of a spin flow by means of an electric field [1, 2], by optical methods [3, 4], and using contacts with a ferromagnet [5, 6] has attracted considerable attention to the problem of spin transport in thin films in the presence of spin–orbit interaction. Two mechanisms of interactions in such systems were developed previously, one by Dresselhaus [7] and the other by Rashba and Bychkov [8, 9]. Investigations in this direction have been mainly devoted to materials with a broad electron band and a weak interaction with phonons [10–15]. An alternative, the hopping transport mechanism for localized electrons in the presence of spin–orbit interaction, was recently described in [16, 17].

In the presence of spin–orbit coupling, both the band and hopping transport mechanisms lead to fundamentally new phenomena related to symmetry of the system under consideration. These include the anomalous Hall effect (without an external magnetic field) [18] and the spin Hall effect (with a transverse spin current) [19–21]. In the presence of electric field **E**, the macroscopic magnetic moment begins to rotate in the plane formed by the **E** vector and the normal to the sample plate surface. While originally predicted for the hopping transport only [16, 17], this effect must, for symmetry considerations, exist for the band transport as well. It should be noted that, under usual conditions, both the anomalous Hall effect and the magnetic moment rotation exhibit a relaxation character. Indeed, in the absence of an external magnetic field, the magnetic moment appears in the initial instant and then decays as a result of the scattering of carriers on impurities or phonons. For this reason, these effects can only be observed either in the nonstationary regime or in samples of small dimensions (below the diffusion length) with spin-polarized carriers injected via the contact with a ferromagnet.

The situation changes in the presence of an external magnetic field. In this case, the sample has a magnetic moment induced by the field and, hence, the rotation of the magnetization vector can be observed in a stationary regime as well. This paper presents the results of a theoretical investigation of spin transport in the presence of spin–orbit interaction in an external magnetic field, in a model of hopping transport (for example, of small polarons). It will be demonstrated that the presence of spin–orbit coupling gives rise to the intrinsic oscillation modes referred to below as spin rotation waves (SRWs). These intrinsic oscillations can be excited by means of an external alternating magnetic field. It is established that this excitation has a resonance character, provided that the external magnetic field frequency coincides with that of the intrinsic SRWs, which is manifested in the character of the frequency dependence of the magnetic susceptibility.

2. BASIC EQUATIONS

The model of spin–orbit interaction in a two-dimensional electron gas, which was developed in [8, 9], can be described (to within an insignificant constant $-\hbar^2K^2/4m$) using the Hamiltonian

$$\frac{\hbar^2}{2m}(\mathbf{k} - \boldsymbol{\sigma} \times \mathbf{K})^2, \tag{1}$$

where m is the electron mass, $\mathbf{k}$ is the electron wavevector, $\boldsymbol{\sigma}$ is the spin operator, and $\mathbf{K}$ is the vector constant of spin–orbit coupling. The latter quantity has a dimension of the wavevector and is directed along the normal

to the plate surface (along the z axis). This study is devoted to materials with a narrow electron band and, hence, needs generalization. It should be noted that expression (1) formally resembles a Hamiltonian describing an electron in an external magnetic field. Therefore, the required generalization can obtained using methods developed for the description of electrons in a narrow allowed band in the presence of magnetic field. This approach was previously used in the theory of small polarons [22], where the generalization was provided by writing Hamiltonian (1) in the form of $\varepsilon(\mathbf{k} - \boldsymbol{\sigma} \times \mathbf{K})$, where $\varepsilon(\mathbf{k})$ is the band energy.

It should be noted that this generalization is formally valid only to within a correction linear in the parameter $\boldsymbol{\sigma} \times \mathbf{K}$. However, it is convenient (for technical reasons) to use the above expression in the general form and perform linearization in the final stage.

Within the framework of such generalization, the Hamiltonian of a system occurring in the external electric ($\mathbf{E}$) and magnetic ($\mathbf{B}$) fields in the site representation can be written as

$$\begin{aligned} H &= \sum_{m, m', \lambda, \lambda'} J^{\lambda', \lambda}_{m', m} a^{\dagger}_{m', \lambda'} a_{m, \lambda} \\ &- e\mathbf{E} \cdot \sum_{m, \lambda} \mathbf{R}_m a^{\dagger}_{m, \lambda} a_{m, \lambda} \\ &- 2\mu_B \mathbf{B} \cdot \sum_{m, \lambda', \lambda} \boldsymbol{\sigma}_{\lambda', \lambda} a^{\dagger}_{m, \lambda'} a_{m, \lambda} + H_{\text{int}} + H_{\text{ph}}, \end{aligned} \tag{2}$$

where $a^{\dagger}$ (a) is the electron creation (annihilation) operator; $\lambda = 1$ or 2; μ_B is the Bohr magneton; m is the lattice site number; $\mathbf{R}_m$ is the radius vector of the mth site;

$$J^{\lambda', \lambda}_{m', m} = J_{m', m} \langle \lambda' | \exp(i\boldsymbol{\sigma} \cdot [\mathbf{K} \times \mathbf{R}_{m'm}]) | \lambda \rangle; \tag{3}$$

$\mathbf{R}_{m'm} = \mathbf{R}_{m'} - \mathbf{R}_m$; $J_{m', m}$ is the usual resonance integral between m'th and mth sites; H_{ph} is the phonon Hamiltonian; and H_{int} is the Hamiltonian of the electron–phonon interaction. Below, the electron–phonon interaction is modeled by the standard Fröhlich Hamiltonian and does not take into account the electron spin rotation upon collisions with phonons.

In the study of hopping transport, it is convenient to use a polaron canonical transformation [23, 24] converting Hamiltonian (2) into

$$\begin{aligned} H &= \sum_{m, m', \lambda, \lambda'} J^{\lambda', \lambda}_{m', m} \Phi_{m'm} a^{\dagger}_{m', \lambda'} a_{m, \lambda} \\ &- e\mathbf{E} \cdot \sum_{m, \lambda} \mathbf{R}_m a^{\dagger}_{m, \lambda} a_{m, \lambda} \\ &- 2\mu_B \mathbf{B} \cdot \sum_{m, \lambda', \lambda} \boldsymbol{\sigma}_{\lambda', \lambda} a^{\dagger}_{m, \lambda'} a_{m, \lambda} + H_{\text{ph}}, \end{aligned} \tag{4}$$

where

$$\begin{aligned} \Phi_{m', m} &= \exp\left\{ \frac{1}{\sqrt{2N}} \right. \\ &\left. \times \sum_{\mathbf{q}} b^{\dagger}_{\mathbf{q}} \gamma_{\mathbf{q}} ([\exp(i\mathbf{q} \cdot \mathbf{R}_m) - \exp(i\mathbf{q} \cdot \mathbf{R}_{m'})] - \text{H.c.}) \right\} \end{aligned} \tag{5}$$

is a multiphonon operator, $b^{\dagger}$ is the operator of phonon creation, N is the total number of sites in the crystal lattice, $\mathbf{q}$ is the phonon wavevector, and $\gamma_{\mathbf{q}}$ is the electron–phonon interaction constant. As can be seen from expression (4), the spin–orbit interaction is only manifested by substitution of the matrix element $J^{\lambda', \lambda}_{m', m}$ for the resonance integral $J_{m', m}$. for this reason, the theory of hopping transport with allowance for the spin–orbit interaction can be constructed using (naturally generalized) methods developed in the theory of small polarons.

Let us introduce the density matrix in the site representation as

$$\begin{aligned} \rho^{m, \lambda}_{m', \lambda'}(t) &= \text{Sp}\left\{ \exp(-\beta H_0) \exp\left(\frac{i}{\hbar} Ht\right) \right. \\ &\left. \times a^{\dagger}_{m', \lambda'} a_{m, \lambda} \exp\left(-\frac{i}{\hbar} Ht\right) \right\}, \end{aligned} \tag{6}$$

where H_0 is the Hamiltonian corresponding to a zero electric field and $\beta = (k_B T)^{-1}$. Below we will assume that the frequency of the alternating field is sufficiently low, so that the explicit time dependence of the Hamiltonian has an adiabatic character. In what follows, we retain (which is usual for the hopping transport) only diagonal (with respect to lattice sites) components of the density matrix, $\rho^{\lambda}_{\lambda'}(m) \equiv \rho^{m, \lambda}_{m, \lambda'}$. For these components, we obtain the following balance equation (the calculation technique is described elsewhere [24]):

$$\begin{aligned} &\frac{d\rho^{\lambda}_{\lambda'}(m)}{dt} \\ &- \frac{2i}{\hbar} \mu_B \mathbf{B} \cdot \sum_{\lambda_1} [\boldsymbol{\sigma}_{\lambda_1 \lambda'} \rho^{\lambda}_{\lambda_1}(m) - \boldsymbol{\sigma}_{\lambda \lambda_1} \rho^{\lambda_1}_{\lambda'}(m)] \\ &= \sum_{m', \lambda_1, \lambda_2} \rho^{\lambda_1}_{\lambda_2}(m') W^{\lambda_1, \lambda}_{\lambda_2, \lambda'}(m', m), \end{aligned} \tag{7}$$

where $W(m', m)$ is the probability of hopping from m'th to mth site in the presence of external electric and magnetic fields and the spin–orbit interaction. Note that Eq. (7) is written in the Markov form, which is admitted provided that the magnetic field frequency ω is not very high, in particular, provided that $\beta\hbar\omega \ll 1$. By vir-

tue of the conservation of the number of particles, the hopping probabilities satisfy the sum rule,

$$\frac{1}{S}\sum_{m,\lambda}\rho_{\lambda}^{\lambda}(m) = n, \quad \sum_{m,\lambda} W_{\lambda_1\lambda}^{\lambda_2\lambda}(m', m) = 0, \tag{8}$$

where S is the plate surface area and n is the electron density per unit surface area; in addition, the probabilities obey the following relation:

$$W_{\lambda_1\lambda'}^{\lambda_2\lambda}(m', m) = [W_{\lambda_2\lambda}^{\lambda_1\lambda'}(m', m)]^*. \tag{9}$$

Equation (7) describes hopping transport in the presence of external electric and magnetic fields and the spin–orbit interaction in both ordered and disordered systems. For disordered systems, this equation can hardly be studied by analytical methods. The quantitative analysis requires calculations performed by numerical methods, while qualitative conclusions can be derived using ideas from percolation theory. In crystalline samples (e.g., for the model of small polarons) with $W(m', m) = W(m' - m)$, Eq. (7) is strongly simplified on going to the momentum representation,

$$\begin{aligned}\rho_{\lambda'}^{\lambda}(\boldsymbol{\kappa}) &= \frac{1}{S}\sum_{m}\rho_{\lambda'}^{\lambda}(m)\exp(i\boldsymbol{\kappa}\cdot\mathbf{R}_m),\\ W_{\lambda_2\lambda'}^{\lambda_1\lambda}(\boldsymbol{\kappa}) &= \sum_{m} W_{\lambda_2\lambda'}^{\lambda_1\lambda}(m)\exp(i\boldsymbol{\kappa}\cdot\mathbf{R}_m),\end{aligned} \tag{10}$$

where it takes the form

$$\begin{aligned}&\frac{d\rho_{\lambda'}^{\lambda}(\boldsymbol{\kappa})}{dt} - \frac{2i}{\hbar}\mu_B\mathbf{B}\cdot\sum_{\lambda_1}[\boldsymbol{\sigma}_{\lambda_1\lambda'}\rho_{\lambda_1}^{\lambda}(\boldsymbol{\kappa}) - \boldsymbol{\sigma}_{\lambda\lambda_1}\rho_{\lambda'}^{\lambda_1}(\boldsymbol{\kappa})]\\ &\qquad = \sum_{\lambda_1,\lambda_2}\rho_{\lambda_2}^{\lambda_1}(\boldsymbol{\kappa})W_{\lambda_2,\lambda'}^{\lambda_1,\lambda}(\boldsymbol{\kappa}).\end{aligned} \tag{11}$$

For studying Eqs. (7) or (11), it is convenient to pass from the matrix form of $\rho_{\lambda'}^{\lambda}$ to a different form of writing the four-component density matrix. The first component,

$$f(m) = \sum_{\lambda}\rho_{\lambda}^{\lambda}(m), \tag{12}$$

is the particle density matrix, and the other three components,

$$\mathbf{s}(m) = \sum_{\lambda,\lambda'}\boldsymbol{\sigma}_{\lambda,\lambda'}\rho_{\lambda'}^{\lambda}(m), \tag{13}$$

represent the magnetic moment density matrix.

3. RESPONSE TO ALTERNATING MAGNETIC FIELD

Let the external magnetic field $\mathbf{B}$ be the sum of a constant component $\mathbf{B}_0$ and a time-dependent component with the amplitude $\mathbf{b}$. Taking into account only the response linear in $\mathbf{b}$, we can write the expression for the field $\mathbf{B}(t)$ in the complex form as

$$\mathbf{B}(t) = \mathbf{B}_0 + \mathbf{b}e^{-i\omega t}.$$

Then we perform linearization of the transport equation (11) with respect to the alternating field component amplitude $\mathbf{b}$. In a zero approximation, Eq. (11) is the equation of transport in a constant electric field. Considering only a stationary regime, we can omit the time derivative $d\rho_{\lambda'}^{\lambda}(\boldsymbol{\kappa})/dt$. Restricting the consideration further to a spatially homogeneous situation, we may set $\boldsymbol{\kappa} = 0$ in Eq. (11), after which the equation for the matrix $f(\boldsymbol{\kappa})$ becomes trivial (because $f \equiv f(\boldsymbol{\kappa} = 0) = n$ by virtue of the sum rule (10)) and depends neither on the character of interactions in the system nor on the electric and magnetic fields, provided that the number of electrons in this system is retained.

As for the stationary value of $\mathbf{s}_0 \equiv \mathbf{s}(\boldsymbol{\kappa} = 0)$ for $\mathbf{b} = 0$, it can be determined without solving Eq. (11). This quantity determines the thermodynamically equilibrium paramagnetic moment $\mathbf{M}(\mathbf{B}_0) = 2\mu_B\mathbf{s}_0$ induced by the external field $\mathbf{B}_0$. It is assumed, that the electric field does not influence the magnetic moment $\mathbf{M}$.

Now we will proceed to analysis of the linear response to the alternating magnetic field $\mathbf{b}$ and determine the correction $\mathbf{s}_1$ linear in $\mathbf{b}$ to the spin density matrix $\mathbf{s}$. After some transformations in Eq. (11), an equation for this correction can be written as

$$\begin{aligned}&-i\omega\mathbf{s}_1 + \boldsymbol{\omega}_c\times\mathbf{s}_1 - \sum_{\{\lambda\}}{}^{(0)}\rho_{\lambda_2}^{\lambda_1(1)}W_{\lambda_2\lambda'}^{\lambda_1\lambda}\boldsymbol{\sigma}_{\lambda'\lambda}\\ &\qquad = \frac{1}{\hbar}\mathbf{b}\times\mathbf{M} + \sum_{\{\lambda\}}{}^{(1)}\rho_{\lambda_2}^{\lambda_1(0)}W_{\lambda_2\lambda'}^{\lambda_1\lambda}\boldsymbol{\sigma}_{\lambda'\lambda},\end{aligned} \tag{14}$$

where ${}^{(0,1)}\rho$ and ${}^{(0,1)}W$ are the corresponding values in the zero and first approximations with respect to the alternating field amplitude $\mathbf{b}$; $\boldsymbol{\omega}_c = 2\mu_B\mathbf{B}_0/\hbar$ is the cyclotron frequency. The quantities ${}^{(0,1)}\rho_{\lambda'}^{\lambda}$ are related to $f_{0,1}$ and $\mathbf{s}_{0,1}$ by relations (12) and (13) (with $f_0 = n$, $f_1 = 0$).

Let us restrict the consideration, as is usual in hopping transport theory, to the lowest (quadratic) approximation with respect to the resonance integral J for the hopping probabilities (two-site model). However, it should be borne in mind that the two-site model overlooks important features such as the anomalous Hall effect and the spin Hall effect, which can be described only using hopping probabilities incorporating the three-site terms [17]. This situation is analogous to that in the theory of small polarons, where the Hall effect is also described using a three-site model [22]. The absence of the anomalous Hall effect and the spin Hall effect is related to the fact that the two-site model involves no entanglement between equations for the electron (f) and spin (s) density matrices in the absence

of an external magnetic field. From the standpoint of physics, the disappearance of anomalous Hall effects reflects the absence of quantum interference in the two-site hopping model [22, 24]. It should be noted that the electron and spin degrees of freedom also exhibit no entanglement for the band conduction mechanism in a spatially homogeneous situation with neglect of scattering [10]. Moreover, it can be shown that no such entanglement under weak coupling to phonons is present even with allowance for scattering if the effect of the spin–orbit coupling is ignored.

However, in the presence of magnetic field, equations for the electron and spin degrees of freedom are interrelated and Eq. (14) contains both a contribution proportional to $\mathbf{s}_0$ and proportional to the particle concentration n (in the terms proportional to the hopping probabilities). The latter contribution describes a trivial variation of the paramagnetic dipole moment $\mathbf{b}(d\mathbf{M}/d\mathbf{B}_0)$ due to a small increment $\mathbf{b}$ of the magnetic field. Below we omit this standard renormalization of the magnetic moment and, assuming $\mathbf{B}_0$ to be sufficiently small, also neglect the influence of a magnetic field on the hopping probabilities in the right-hand side of Eq. (14).

The right-hand side of Eq. (14) plays the role of a driving force in the transport equation. Under conditions of a rather small hopping probability (i.e., a small value of the resonance integral J), the second term in the right-hand side of Eq. (14) is small compared to the first term and can be omitted.

Now let us consider the third term in the left-hand side of Eq. (14). Note that, even with allowance for the influence of the magnetic field $\mathbf{B}_0$ on the hopping probability, this term does not link equations for the electron and spin components of the density matrix since, according to our assumptions, the concentration of particle is independent of the magnetic field. Below we assume the magnetic field to be sufficiently weak ($\beta\mu_B B_0 \ll 1$) and ignore its influence on the hopping probability. In this approximation, it is expedient to restrict the analysis to a linear relation between the magnetic field and the moment, $\mathbf{M} = \chi\mathbf{B}_0$, where the paramagnetic susceptibility has a standard form $\chi = \beta\mu_B^2 n$. The two-site hopping probabilities in Eq. (14) at $\mathbf{B}_0 = 0$ can be written as

$$\begin{aligned}\sum_{\lambda,\lambda'} W^{\lambda_1,\lambda}_{\lambda_2,\lambda'}(m', m)\boldsymbol{\sigma}_{\lambda'\lambda} &= W(m', m)\\ &\times \langle\lambda_2|\exp(i\boldsymbol{\sigma}\cdot[\mathbf{K}\times\mathbf{R}_{mm'}])\boldsymbol{\sigma}\\ &\times \exp(-i\boldsymbol{\sigma}\cdot[\mathbf{K}\times\mathbf{R}_{mm'}])|\lambda_1\rangle\\ &- \delta_{m',m}\sum_{m_1} W(m, m_1)\langle\lambda_2|\boldsymbol{\sigma}|\lambda_1\rangle.\end{aligned} \tag{15}$$

Here, $W(m', m)$ are the usual two-site hopping probabilities according to the theory of small polarons (with neglect of the spin–orbit coupling) in the presence of an electric field $\mathbf{E}$. For a not very strong field E, these two-site hopping probabilities can be represented as [24]

$$\begin{aligned}W(m', m) &= \exp\left(\frac{1}{2}\beta e\mathbf{E}\cdot\mathbf{R}_{m'm}\right)\\ &\times \frac{J^2_{m'm}\sqrt{\pi\beta}}{2\hbar\sqrt{E_a}}\exp(-\beta E_a),\end{aligned} \tag{16}$$

where E_a is the activation energy for the hopping between sites and $J_{m'm}$ is the usual resonance integral for the hoping between m'th and mth sites

In order to transform expression (15), let us find the operator

$$\boldsymbol{\sigma}(\tau) = \exp(i\tau\boldsymbol{\sigma}\cdot\mathbf{A})\boldsymbol{\sigma}\exp(-i\tau\boldsymbol{\sigma}\cdot\mathbf{A}),$$

where $\mathbf{A} = \mathbf{K}\times\mathbf{R}_{mm'}$. Differentiation with respect to τ yields

$$\frac{d\boldsymbol{\sigma}(\tau)}{dt} = i\exp(i\tau\boldsymbol{\sigma}\cdot\mathbf{A})[(\boldsymbol{\sigma}\cdot\mathbf{A}), \boldsymbol{\sigma}]\exp(-i\tau\boldsymbol{\sigma}\cdot\mathbf{A}) = \mathbf{A}\times\boldsymbol{\sigma}.$$

Solving this equation with the initial condition $\boldsymbol{\sigma}(0) = \boldsymbol{\sigma}$, we obtain

$$\begin{aligned}\boldsymbol{\sigma}(\tau) &= \frac{\mathbf{A}(\mathbf{A}\cdot\boldsymbol{\sigma})}{A^2}[1-\cos(A\tau)]\\ &+ \boldsymbol{\sigma}\cos(A\tau) + \frac{\mathbf{A}\times\boldsymbol{\sigma}}{A}\sin(A\tau).\end{aligned} \tag{17}$$

Substituting expression (17) into Eq. (15), we set $\tau = 1$ and consider the limit $A \ll 1$ (which corresponds to $Ka \ll 1$, where a is the hopping length, a is the lattice parameter for small polarons in the case of hopping between nearest neighbors). Then, to within the terms on the order of A^2,

$$\begin{aligned}&\sum_{\lambda,\lambda'} W^{\lambda_1,\lambda}_{\lambda_2,\lambda'}(m', m)\boldsymbol{\sigma}_{\lambda'\lambda}\\ &= \boldsymbol{\sigma}_{\lambda_2\lambda_1}\left\{W(m', m) - \delta_{m'm}\sum_{m_1} W(m, m_1)\right\}\\ &+ W(m', m)\left\{[\mathbf{K}\times\mathbf{R}_{mm'}]\times\boldsymbol{\sigma}_{\lambda_2\lambda_1}\right.\\ &\left.+\frac{1}{2}[\mathbf{K}\times\mathbf{R}_{mm'}]\times[[\mathbf{K}\times\mathbf{R}_{mm'}]\times\boldsymbol{\sigma}_{\lambda_2\lambda_1}]\right\}.\end{aligned} \tag{18}$$

Finally, substituting expression (18) into Eq. (14) and

making transformations, we arrive at

$$-i\omega s_{1j} + [[\mathbf{V}\times\mathbf{K}]\times\mathbf{s}_1]_j + [\boldsymbol{\omega}_c\times\mathbf{s}_1]_j + \Omega_{jk}s_{1k} = -\frac{1}{\hbar}[\mathbf{b}\times\mathbf{M}]_j, \tag{19}$$

where

$$\mathbf{V} = \sum_m \mathbf{R}_m W(m) = \frac{\mathbf{E}}{E}\sinh\left(\frac{1}{2}\beta eEa\right) \times \frac{J^2 a\sqrt{\pi\beta}}{\hbar\sqrt{E_a}}\exp(-\beta E_a) \tag{20}$$

is the drift velocity in the model of small polarons [24, Eq. (16)]. Expression (20) takes into account only one variant of hopping—that between nearest neighbors via the sites of a square lattice with the parameter a (J is the resonance integral between the nearest neighbors). For the square lattice, the relaxation tensor $\Omega_{jk} = \delta_{jk}\Omega_j$ is diagonal ($\Omega_{ij} = \delta_{ij}\Omega_j$) and

$$\Omega_x = K^2 D_{xx}, \quad \Omega_y = K^2 D_{yy}, \quad \Omega_z = K^2(D_{xx} + D_{yy}), \tag{21}$$

where

$$D_{jj} = \frac{1}{2}\sum_m R_{mj}^2 W(m) \tag{22}$$

is the tensor of diffusion coefficients. Orienting the x axis along the electric field and the z axis along the normal to the plate surface, we have

$$D_{xx} = \cosh\left(\frac{1}{2}\beta eEa\right)\frac{J^2 a^2\sqrt{\pi\beta}}{2\hbar\sqrt{E_a}}\exp(-\beta E_a), \tag{23}$$

$$D_{yy} = \frac{J^2 a^2\sqrt{\pi\beta}}{2\hbar\sqrt{E_a}}\exp(-\beta E_a) \equiv D. \tag{24}$$

In the limiting case of a weak electric field ($\beta eEa \ll 1$), the transverse and longitudinal components of the diffusion coefficient are the same ($D_{xx} = D_{yy} = D$). In this region of electric field strengths, the Ohm law is valid and the drift velocity $\mathbf{V}$ is proportional to the field $\mathbf{V} = u\mathbf{E}$, where u is the drift mobility related to the diffusion coefficient (24) by the Einstein relation: $u = \beta eD$.

In order to solve Eq. (19), let us assume that the constant component of the magnetic field is oriented along the y axis, perpendicular to the electric field, and occurs in the plane of the plate, while the alternating component occurs in the xz plane and is perpendicular to the constant component. In this geometry, $s_{1y} = 0$ and

$$(-i\omega + K^2 D_{xx})s_{1x} + (KV - \omega_c)s_{1z} = -\frac{1}{\hbar}b_z M, \tag{25}$$

$$[-i\omega + K^2(D_{xx} + D_{yy})]s_{1z} - (KV - \omega_c)s_{1x} = \frac{1}{\hbar}b_x M. \tag{26}$$

4. SPIN ROTATION WAVES

It can be demonstrated that the frequency dependence of the linear response to the alternating magnetic field (i.e., of the magnetic susceptibility) determined by Eqs. (25) and (26) has a resonance character with a maximum at the frequency of intrinsic spin rotation modes. The solution of these equations has the following form:

$$s_{1x} = -\frac{M}{\hbar}\frac{b_z[-i\omega + K^2(D_{xx} + D_{yy})] + b_x(KV - \omega_c)}{[\omega - \nu + iK^2(D_{xx} + D_{yy}/2)][\omega + \nu + iK^2(D_{xx} + D_{yy}/2)]}, \tag{27}$$

$$s_{1z} = -\frac{M}{\hbar}\frac{b_x(-i\omega + K^2 D_{xx}) + b_z(KV - \omega_c)}{[\omega - \nu + iK^2(D_{xx} + D_{yy}/2)][\omega + \nu + iK^2(D_{xx} + D_{yy}/2)]}, \tag{28}$$

where

$$\nu = \sqrt{(KV - \omega_c)^2 - K^4 D_{yy}^2/4}. \tag{29}$$

This solution (27), (28) corresponds to rotation of the vector of induced magnetic moment $\mathbf{s}_1$ along an elliptically polarized trajectory in the xz plane. Expressions for the polarization ellipticity parameters are not presented here because they are very cumbersome. The most important feature of this rotation is the presence of a resonance, which corresponds to coincidence of the magnetic field frequency ω to the characteristic frequency ν. The resonance is caused by the excitation of intrinsic oscillations representing the SRWs. Recently, it was demonstrated that the spin–orbit coupling in the presence of an external electric field leads to rotation of the magnetic moment in the plane formed by the electric field vector and the normal to the sample plate surface [17]. In the absence of a constant magnetic field (for $\omega_c = 0$), the frequency ω_r of rotation of the magnetic moment is described by the formula

$$\omega_r = \frac{1}{2}K^2 D\sqrt{\frac{E^2}{E_c^2} - 1}, \quad E_c = \frac{K}{2e\beta}. \tag{30}$$

This relation corresponds to the case of a weak electric field, whereby $V = uE$, $D_{xx} = D_{yy} = D$, and $u = \beta eD$. The quantity E_c is a minimum electric field strength at which the magnetic moment begins to rotate.

In the presence of a constant magnetic field, the resonance frequency shifts by a value corresponding to the cyclotron frequency: $\omega = |\omega_r \pm \omega_c|$, where selection of the sign depends on the directions of the electric field and the constant magnetic field. For determining ω_r, it is necessary to perform two measurements for the opposite directions of the applied electric field. Then ω_r is calculated as half the sum of two measured resonance frequencies (it is assumed that $|\omega_r| > |\omega_c|$).

The positive role of the constant magnetic field component $\mathbf{B}_0$ in these experiments reduces to polarization of the spin moment, which makes the stationary magnetic moment M in formulas (27) and (28) nonzero. However, the presence of BW0 also leads to an additional "parasitic" (Larmor-type) precession of vector $\mathbf{s}_1$. For this reason, it is expedient to perform such experiments using magnetic fields without a constant component and creating a polarized spin state by some other means. Such a polarized state with $\mathbf{M} \neq 0$ can be provided, for example, by injection of polarized spins via contact of the sample with a ferromagnetic material. The experiments should be performed with sufficiently thin samples, in which a distance between the opposite sample–ferromagnet contacts is shorter than the spin diffusion length. In this case, a homogeneous magnetic moment is created in the sample. Note that, for the hopping transport, the spin diffusion length is K^{-1} [17]. The results of such experiments can be also described in terms of Eqs. (27)–(28) with $\omega_c = 0$.

Now it will be demonstrated that the aforementioned resonance is related to the excitation of intrinsic oscillation modes in a system with the spin-orbit interaction and has common features with a ferromagnetic resonance in ferromagnets. In order to determine the spectrum of intrinsic oscillations, let us return to Eq. (11) for the density matrix in the absence of an external magnetic field. Taking into account relation (18) and performing simple transformations, we arrive at an expression for the spin vector component $\boldsymbol{\rho}(\boldsymbol{\kappa})$ in the time-dependent Fourier representation:

$$\begin{aligned} -i\omega\boldsymbol{\rho}(\boldsymbol{\kappa}) &= [W_E(\boldsymbol{\kappa}) - W_E(0)]\boldsymbol{\rho}(\boldsymbol{\kappa}) \\ &+ i\boldsymbol{\rho}(\boldsymbol{\kappa}) \times [\mathbf{K} \times \nabla_{\boldsymbol{\kappa}}] W_E(\boldsymbol{\kappa}) \\ &+ \frac{1}{2}[\boldsymbol{\rho}(\boldsymbol{\kappa}) \times [\mathbf{k} \times \nabla_{\boldsymbol{\kappa}}]] \times [\mathbf{K} \times \nabla_{\boldsymbol{\kappa}}] W_E(\boldsymbol{\kappa}), \end{aligned} \tag{31}$$

where $W_E(\boldsymbol{\kappa})$ is the Fourier component of the two-site hopping probability (16) in an external electric field. Using Eqs. (10) and (16), we obtain

$$W_E(\boldsymbol{\kappa}) = W(\boldsymbol{\kappa} + ie\mathbf{E}\beta/2),$$

where $W(\boldsymbol{\kappa})$ is the Fourier component of the hopping probability at $\mathbf{E} = 0$ such that $W(\boldsymbol{\kappa}) = W(-\boldsymbol{\kappa})$. In the long-wavelength limit of the approximation linear in the electric field strength, we obtain to within the terms proportional to κ^2:

$$\begin{aligned} -i\omega\boldsymbol{\rho}(\boldsymbol{\kappa}) &= -[D\kappa^2 + i(\boldsymbol{\kappa} \cdot \mathbf{E})u]\boldsymbol{\rho}(\boldsymbol{\kappa}) \\ &- \boldsymbol{\rho}(\boldsymbol{\kappa}) \times [\mathbf{K} \times (u\mathbf{E} - 2iD\boldsymbol{\kappa})] - \hat{\Omega}\boldsymbol{\rho}(\boldsymbol{\kappa}). \end{aligned} \tag{32}$$

Here, $2D = -(\partial^2/\partial\kappa_x^2)W(\boldsymbol{\kappa})|_{\kappa=0}$ is the diffusion coefficient, $u = \beta eD$ is the mobility, $\hat{\Omega}$ is the diagonal relaxation tensor determined by formulas (21); in a weak electric field, $\Omega_{xx} = \Omega_{yy} = K^2D$, $\Omega_{zz} = 2K^2D$. Equating to zero the determinant of the system of equations (32), we obtain three dispersion relations for the intrinsic modes:

$$\begin{aligned} \omega_{1,2}(\boldsymbol{\kappa}) &= -uE\kappa_x - iD\left(\kappa^2 + \frac{3}{2}K^2\right) \\ &\pm \frac{1}{2}DK^2\sqrt{\frac{E^2}{E_c^2} - 1 - 16\frac{\kappa^2}{K^2} - 8i\frac{E\kappa_x}{E_cK}}, \end{aligned} \tag{33}$$

$$\omega_3(\mathbf{k}) = -uEk_x - iD(k^2 + K^2), \tag{34}$$

where E_c is the critical electric field strength given be formula (30).

The $\omega_3(\boldsymbol{\kappa})$ mode, which describes a diffusion-relaxation dynamics of the magnetic moment along the y axis, is not of interest for the problem under consideration. The $\omega_{1,2}(\boldsymbol{\kappa})$ modes just represent the SRWs corresponding to rotation of the magnetic moment vector in the xz plane. In the limit of $\boldsymbol{\kappa} \longrightarrow 0$, the expression for the SRW frequency simplifies to

$$\omega_{1,2}(0) = \pm\frac{1}{2}DK^2\sqrt{\frac{E^2}{E_c^2} - 1} - \frac{3}{2}iDK^2. \tag{35}$$

The real part of this relation determines the resonance frequency ω_r (30) for the frequency dependence of the magnetic susceptibility.

In order to measure the dispersion of the intrinsic modes (33), it is necessary to create a spatially inhomogeneous polarized spin state $\mathbf{M}(\mathbf{r})$, for example, by irradiating the sample with an interference pattern of polarized light. Then, the equation of motion under the action of an external driving force can be written as (cf. Eqs. (32) and (25)–(26))

$$\begin{aligned} -i\omega\boldsymbol{\rho}(\mathbf{q}) &+ [Dq^2 + i(\mathbf{q} \cdot \mathbf{E})u]\boldsymbol{\rho}(\mathbf{q}) \\ &- \boldsymbol{\rho}(\mathbf{q}) \times [\mathbf{K} \times (u\mathbf{E} - 2iD\mathbf{q})] \\ &+ \hat{\Omega}\boldsymbol{\rho}(\mathbf{q}) = \frac{1}{\hbar}\mathbf{M}(\mathbf{q}) \times \mathbf{b}, \end{aligned} \tag{36}$$

where $\mathbf{q}$ is the wave vector of the interference pattern. The solution of Eq. (36) exhibits a resonance at $\omega =$

$\omega_{1,2}(\mathbf{q})$ which makes it possible, in principle, to measure the dispersion relation of the SRW by varying the wavelength of the interference pattern. It should be noted that such a dispersion relation is described according to Eq. (33) by a dimensionless parameter κ/K. Note also that, under the conditions of spatial dispersion, the mode frequency (33) is shifted by $-uE\kappa_x$, which indicates that the mode drifts at a rate of uE in the external field [25, 26].

In conclusion, let us consider numerical estimates of the physical parameters describing the SRWs. At room temperature and $K \approx 10^5$ cm^{-1} [14, 17, 27], the critical field according to formula (30) is $E_c \approx 10^4$ V/cm. Estimation of the resonance frequency at $E \approx 10^5$ V/cm (i.e., for $E \gg E_c$, whereby $\omega_r \approx KuE$) and $u \approx 1$ cm^2/(V s) yields $\omega_r \approx 10^{10}$ s^{-1}. It should be emphasized that this estimate is closely related to the model of small polarons used in this study, according to which the charge carriers possess a small mobility and, accordingly, a low value of the intrinsic SRW frequency. The resonance peak is well pronounced in cases when the Q value, defined as the ratio of the resonance frequency to the inverse lifetime of intrinsic oscillations, is sufficiently large. According to formulas (27) and (28), the Q value can be estimated for $E \gg E_c$ as

$$Q = \frac{\omega_r}{DK^2} \approx \frac{eE\beta}{K} = \frac{E}{2E_c}, \tag{37}$$

which shows that the condition of high Q is valid in the region of sufficiently strong electric fields.

In this study, the SRWs were described using the model of hopping conductivity. However, it is likely that such elementary excitations in the presence of spin–orbit coupling possess a universal character and exist for any charge transport mechanism. This is related to the presence of the $\mathbf{K} \times \mathbf{V}$ pseudo-vector ($\mathbf{V}$ is the drift velocity), which transforms into $\mathbf{K} \times \mathbf{E}$ in the region of applicability of the Ohm law. The presence of this quantity in the system studied is reflected on the phenomenological level by the form of the equation of motion for the magnetic moment. Based on the general considerations, the equation of motion at a nonzero average velocity of particles can be written in the κ–ω representation as

$$\begin{aligned} -i\omega\boldsymbol{\rho}(\boldsymbol{\kappa}, \omega) &= -i(\boldsymbol{\kappa} \cdot \mathbf{U}(\boldsymbol{\kappa}))\boldsymbol{\rho}(\boldsymbol{\kappa}, \omega) \\ -\boldsymbol{\rho}(\boldsymbol{\kappa}, \omega) &\times [\mathbf{K} \times \nabla_{\boldsymbol{\kappa}}(\boldsymbol{\kappa} \cdot \mathbf{U}(\boldsymbol{\kappa}))] - \hat{\Omega}\boldsymbol{\rho}(\boldsymbol{\kappa}, \omega), \end{aligned} \tag{38}$$

where $\mathbf{U}(\boldsymbol{\kappa}) = \mathbf{V} - iD\boldsymbol{\kappa}$ is the drift velocity in the presence of a spatial dispersion and $\hat{\Omega}$ is the aforementioned tensor of the inverse relaxation time. For simplicity, Eq. (38) is written here for a plate isotropic in the xy plane and is valid only in the region of not very strong electric fields (where D and Ω are independent of E). Otherwise, the diffusion coefficient is a tensor and $\hat{\Omega}$ is a biaxial tensor (rather than uniaxial as described above for a weak electric field).

A nontrivial contribution due to the spin–orbit interaction is described by the second term in the right-hand part of Eq. (38). A particular form of this term can be restored using Hamiltonian (1) or (2), by substituting the drift velocity $\mathbf{U}(\boldsymbol{\kappa})$ for the random electron velocity $\hbar\mathbf{k}/m$ in the equation of motion for the free spin. The form of this additional contribution to the equation of motion indicates that the spin–orbit interaction gives rise to an effective magnetic field proportional to $\mathbf{K} \times \nabla_{\boldsymbol{\kappa}}(\boldsymbol{\kappa} \cdot \mathbf{U}(\boldsymbol{\kappa}))$, which causes precession of the magnetic moment. Direct comparison of the phenomenological equation (38) with microscopic equation (32) shows their complete identity and suggests that Eq. (32) has a universal character. In particular, it can be expected that Eq. (38) is also applicable in the case of band transport (with properly redetermined drift velocity and diffusion coefficient).

In a microscopic description of the magnetic moment dynamics for the band transport, a certain technical difficulty is encountered as a result of the twofold role of the electric field in the phenomena nonlinear with respect to the field strength. The field produces (i) the heating of carriers in a broad electron band and (ii) the rotation of magnetic moment in the presence of a directed flow of electrons. For this reason, a simpler approach is provided by the phenomenological equation (38), which should be used with certain care. The problem consists in description of the band transport leading to the appearance of an additional dimensionless parameter $\Delta\tau/\hbar = Kl$, where $\Delta = \hbar^2 Kk/m$ is the splitting of the spin subbands (see, e.g., [10]) and $l = \hbar k\tau/m$ is the electron mean free path. According to the results of preliminary analysis, Eq. (38) is valid only for $Kl < 1$ and requires modification in the case of strong spin–orbit coupling ($Kl > 1$). This problem is still waiting to be solved.

It can be expected that the intrinsic SRW frequency in materials with high mobility at a given electric field strength will be much greater than the estimate obtained for this frequency above for the hopping transport (at low mobility).

REFERENCES

1. H. J. Zhu, M. Ramsteiner, H. Kostial, *et al.*, Phys. Rev. Lett. **87**, 016601 (2001).
2. E. I. Rashba and A. L. Efros, Phys. Rev. Lett. **91**, 126405 (2003).
3. J. M. Kikkawa and D. D. Awschalom, Phys. Rev. Lett. **80**, 4313 (1998).
4. S. D. Ganichev, E. L. Ivchenko, V. V. Belkov, *et al.*, Lett. Nature **417**, 153 (2002).
5. P. R. Hammar, B. R. Bennett, M. J. Yang, and M. Johnson, Phys. Rev. Lett. **83**, 203 (1999).

6. A. Brataas, Y. Tserkovnyak, G. E. W. Bauer, and B. I. Halperin, Phys. Rev. B **66**, 060404 (2002).
7. G. Dresselhaus, Phys. Rev. **100**, 580 (1955).
8. E. I. Rashba, Fiz. Tverd. Tela (Leningrad) **2**, 1224 (1960) [Sov. Phys. Solid State **2**, 1109 (1960)].
9. Yu. L. Bychkov and E. I. Rashba, Pis'ma Zh. Éksp. Teor. Fiz. **39** (2), 66 (1984) [JETP Lett. **39**, 78 (1984)].
10. E. G. Mishchenko and B. I. Halperin, Phys. Rev. B **68**, 045317 (2003).
11. T. P. Pareek and P. Bruno, Phys. Rev. B **65**, 241305 (2002).
12. Jun-ichiro Inoue, G. E. W. Bauer, and L. W. Molenkamp, Phys. Rev. B **67**, 033104 (2003).
13. J. Schliemann, J. C. Egues, and D. Loss, Phys. Rev. Lett. **90**, 146 801 (2003).
14. T. V. Shahbazyan and M. E. Raikh, Phys. Rev. Lett. **73**, 1408 (1994).
15. D. Culcer, J. Sinova, N. A. Sinitsyn, *et al.*, cond-mat/0309475.
16. T. Damker, V. V. Bryksin, and H. Boettger, Phys. Status Solidi C **1**, 84 (2004).
17. T. Damker, H. Boettger, and V. V. Bryksin, Phys. Rev. B **69**, 205327 (2004).
18. A. Crepieux and P. Bruno, Phys. Rev. B **64**, 014416 (2001).
19. J. E. Hirsch, Phys. Rev. Lett. **83**, 1834 (1999).
20. S. Zhang, Phys. Rev. Lett. **85**, 393 (2000).
21. Liangbin Hu, Ju Gao, and Shu-Quin Shen, Phys. Rev. B **68**, 115302 (2003).
22. T. Holstein, Phys. Rev. **124**, 1329 (1961).
23. I. G. Lang and Yu. A. Firsov, Zh. Éksp. Teor. Fiz. **43**, 1843 (1962) [Sov. Phys. JETP **16**, 1301 (1963)].
24. H. Boettger and V. V. Bryksin, *Hopping Conduction in Solids* (Akademie, Berlin, 1985).
25. A. F. Volkov and Sh. M. Kogan, Usp. Fiz. Nauk **96**, 633 (1968) [Sov. Phys. Usp. **11**, 881 (1968)].
26. V. V. Bryksin, P. Kleinert, and M. P. Petrov, Fiz. Tverd. Tela (St. Petersburg) **45**, 1946 (2003) [Phys. Solid State **45**, 2044 (2003)].
27. L. W. Molenkamp, G. Schmidt, and G. T. W. Bauer, Phys. Rev. B **64**, 121202 (2001).

Translated by P. Pozdeev

Journal of Experimental and Theoretical Physics, Vol. 100, No. 2, 2005, pp. 322–330.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 127, No. 2, 2005, pp. 362–371.

ELECTRONIC PROPERTIES OF SOLIDS

The Crystal Field and Crossover Effects in $ErVO_4$

Z. A. Kazei and V. V. Snegirev
Moscow State University, Vorob'evy gory, Moscow, 119899 Russia
e-mail: kazei@plms.phys.msu.su
Received June 23, 2004

Abstract—The Zeeman effect, magnetization $M(H)$, and differential magnetic susceptibility dM/dH of $ErVO_4$ crystals in a pulsed magnetic field have been experimentally and theoretically studied. In magnetic fields $H \parallel [001]$ and $H \parallel [100]$, the energy levels of Er^{3+} ions exhibit mutual approach and crossing (the crossover effect), which results in the peaks in dM/dH and the jumps in $M(H)$ curves at low temperatures. The anomalies in the magnetic properties related to the crossover in $ErVO_4$ for $H \parallel [001]$ are highly sensitive to the electronic structure of Er^{3+} ion, which allows this effect to be used for refining the crystal field parameters. The influence of the temperature, field misorientation from the symmetry axis, parameters of pair interactions, and other factors on the magnitude and character of magnetic anomalies in $ErVO_4$ crystals is considered. © *2005 Pleiades Publishing, Inc.*

1. INTRODUCTION

The crystal field forming the electronic structure of rare earth (RE) element ions in crystals also largely determines their magnetic and magnetoelastic properties. Knowledge of the crystal field parameters, along with the parameters of pair interactions in various RE compounds, allows the magnetic behavior of the RE crystals to be analyzed and adequately described. The development of methods for determining and refining the crystal field parameters in hexagonal and tetragonal crystals is among currently important tasks in solid state physics. In order to solve the problem, it is a common practice to use, in addition to the spectroscopic data, the results of magnetic measurements (in particular, of the initial magnetic susceptibility) in weak magnetic fields. It is naturally expected that the properties of materials in strong magnetic fields—including the magnetization anomalies accompanying the level crossing—are also determined to a considerable extent by the crystal field parameters.

Indeed, the results of investigations of the phenomenon of energy level crossing (the crossover effect) in RE compounds provide valuable information about the electronic structure of RE ions as formed by the crystal field. The crossover effect was observed and studied in much detail in a series of RE zircons RXO_4 (see, e.g., [1, 2]) and was also predicted based on the results of numerical calculations for RE compounds of the $RBa_2Cu_3O_{7-\delta}$ system [3]. The dependence of the magnetic anomalies observed in the crossover region on the magnetic field and the temperature is determined to a considerable extent by the character of intersecting levels and their wavefunctions. The most pronounced anomalies have been observed in van Vleck singlet paramagnets, where the "true" (gapless) crossover of the ground-state singlet level and an excited level in the critical field H_c results in the intense narrow peak in dM/dH. The magnetic-field-induced mixing of the wavefunctions of intersecting levels leads to a finite gap and smoothens the anomalies. Less pronounced anomalies in the crossover region can be expected for the Kramers ions, whereby the lower component of the lower Kramers doublet exhibits crossing with another component of an excited or the same doublet.

This paper presents the results of theoretical and experimental investigations of the magnetic anomalies related to the crossing of levels of Er^{3+} ions in $ErVO_4$, a paramagnet with a zircon structure, observed in pulsed magnetic fields $H \parallel [001]$ and $H \parallel [100]$. A comparison of experimental data obtained in the pulsed magnetic fields to the results of numerical calculations of the magnetic properties allowed us to refine the crystal field parameters and the pair interaction parameters for the compound studied.

2. EXERIMENTAL RESULTS

The magnetization $M(H)$ of an $ErVO_4$ single crystal was measured at helium temperatures for a magnetic field H parallel to the hard magnetization axis [001]. The $ErVO_4$ crystals were grown by method of spontaneous crystallization from solution in a $PbO–PbF_2$ based melt and had an average size of $2 \times 2 \times 3$ mm^3. The magnetization measurements were performed using the induction technique in a pulsed magnetic field with a pulse amplitude of up to 280 kOe and a rise time of the field in the pulse of about 3 ms. The $M(t)$ and $H(t)$ signals from the measuring and field-pick up coils, respectively, were recorder during the field pulse at an 0.02 ms interval (a total of about 500 points). A decom-

pensation signal from the measuring coils was recorded under analogous conditions in the absence of the sample and then subtracted from the total signal in the stage of computer data processing, which gave smoothened curves of the magnetization $M(H)$ and its derivative $dM(H)/dH$.

In addition, the magnetization curves of an $ErVO_4$ crystal were measured at various temperatures in the range from 1.5 to 300 K in stationary magnetic fields with a strength of up to 80 kOe oriented parallel to the symmetric crystallographic directions [100], [110], and [001]. The obtained magnetic data were plotted in the coordinates of $H/M = \chi_M^{-1} + bM^2$ versus M^2 for determining the magnetic susceptibilities of the first and third orders. For sufficiently weak magnetic field, the $H/M(M^2)$ curves are linear, and the H/M value at $M^2 = 0$ gives the usual magnetic susceptibility χ_M^{-1}, while the slope b is related to the nonlinear susceptibility as $\chi_M^{(3)} = -b/(\chi_M^{-1})^4$.

The magnetic susceptibility of $ErVO_4$ is anisotropic in the directions parallel and perpendicular to the tetragonal axis (Fig. 1). However, owing to the symmetry, it is isotropic with respect to the [100] and [110] axes in the basal plane. According to the variation of the Stevens parameter α_J, the hard magnetization axis at high temperatures coincides with the tetragonal axis in the series of vanadates with RE ions from Tb to Ho, and occurs in the basal plane for the RE vanadates with ions of erbium (Er), thulium (Tm), and ytterbium (Yb). As can be seen in Fig. 1, the χ_M^{-1} curve for $ErVO_4$ at $T >$ 150 K obeys the Curie–Weiss law, although the slopes of the linear parts of these plots for the *c* and *a* axes are different. This behavior suggests that the crystal field effects at temperatures about 300 K still play a significant role and the magnetic moment does not reach the value $\mu_{eff} = 9.59\mu_B$ characteristic of the free Er^{3+} ion.

The low-temperature magnetization of $ErVO_4$ along the [001] axis in the fields $H < 200$ kOe has a shape characteristic of paramagnets, exhibits a jumplike change in the vicinity of a critical field ($H_c \sim 235$ kOe), and approaches saturation in higher fields (Fig. 2). The critical field H_c is more reliably determined from the curves of the differential magnetic susceptibility dM/dH (Fig. 3). We believe that a small hysteresis observed in the $M(H)$ and $dM(H)/dH$ curves at low temperatures is probably related to incomplete correction for the phase decompensation of measuring coils. As the temperature increases above 10 K, the jumps in the $M(H)$ curves exhibit rapid smearing, while the critical field remains almost unchanged.

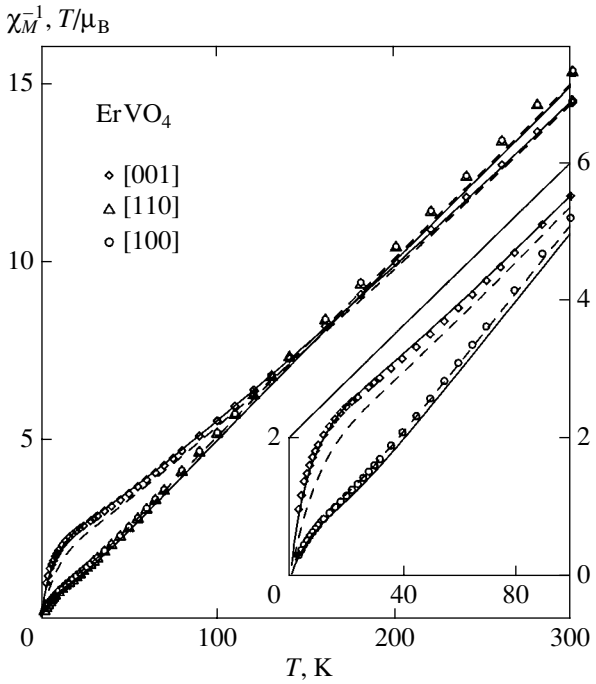

Fig. 1. Temperature dependences of the reciprocal magnetic susceptibility χ_M^{-1} of an $ErVO_4$ crystal measured along the [100], [110], and [001] directions. Solid and dashed curves show the results of calculations with neglect of pair interactions for two sets of the crystal field parameters (1 and 2, respectively, in Table 1).

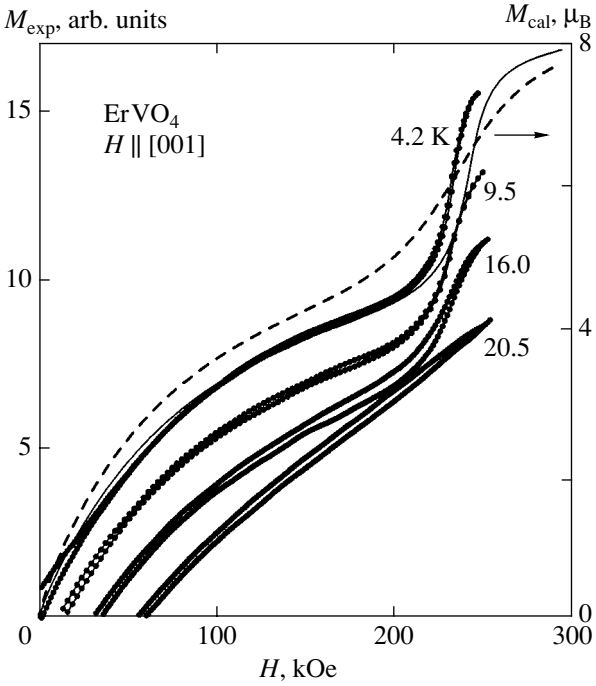

Fig. 2. Magnetization curves for an $ErVO_4$ crystal measured along the [001] direction (for clarity, the curves for $T = 9.5$, 16.0, and 20.5 K are shifted along the *y* axis by 1.5 units relative to each other). Solid and dashed curves show the results of calculations with neglect of pair interactions for two sets of the crystal field parameters (1 and 2, respectively, in Table 1).

3. THEORETICAL ANALYSIS

The complete Hamiltonian for a single 4*f* ion comprises a sum of the crystal field Hamiltonian H_{CF}, the

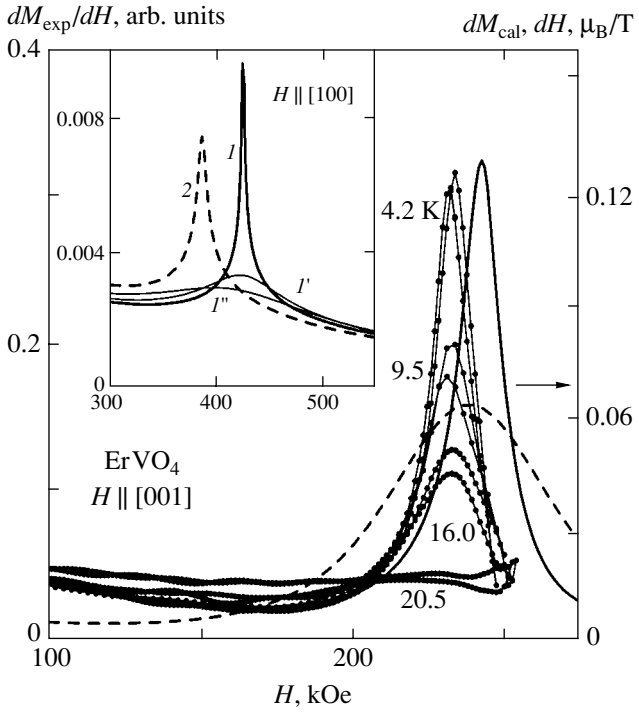

Fig. 3. Differential magnetic susceptibility curves for an $ErVO_4$ crystal measured along the [001] direction. Solid and dashed curves show the adiabatic dependences calculated with neglect of pair interactions for two sets of the crystal field parameters (1 and 2, respectively, in Table 1). The inset shows the curves of dM/dH for the [100] axis at T = 2.1 K calculated using sets 1 (curve *1*) and 2 (curve *2*) and illustrates the effects of the temperature (T = 8 K, curve *1*') and the misorientation ($\Delta\varphi = 2°$, curve *1*'') for set 1.

Zeeman term H_Z (describing the interaction between the angular momentum **J** and the external field **H**), and the Hamiltonians of bilinear (H_B) and quadrupole (H_{QT}) interactions:

$$H = H_{CF} + H_Z + H_B + H_{QT}. \quad (1)$$

Using the method of operator-equivalents and the molecular-field approximation for the bilinear and quadrupole pair interactions, these terms can be expressed as follows (for detail see, e.g., [4]):

$$H_{CF} = \alpha_J B_2^0 O_2^0 + \beta_J (B_4^0 O_4^0 + B_4^4 O_4^4) + \gamma_J (B_6^0 O_6^0 + B_6^4 O_6^4), \quad (2)$$

$$H_Z = -g_J \mu_B \mathbf{H} \cdot \mathbf{J}, \quad (3)$$

$$H_B = -g_J \mu_B \mathbf{H}_B \cdot \mathbf{J}, \quad \mathbf{H}_B = n g_J \mu_B \langle \mathbf{J} \rangle, \quad (4)$$

$$H_{QT} = -G^\alpha \langle O_2^0 \rangle O_2^0 - G^\gamma \langle O_2^2 \rangle O_2^2 - G^\delta \langle P_{xy} \rangle P_{xy}, \quad P_{xy} = \frac{1}{2}(J_x J_y + J_y J_x). \quad (5)$$

Here, O_n^m and B_n^m (n = 2, 4, 6; m = 0, 4; $m \le n$) are the Stevens operators and the crystal field parameters; α_J, β_J, and γ_J are the Stevens parameters; g_J is the Lande factor; μ_B is the Bohr magneton; n is the bilinear exchange parameter, and G^μ are the quadrupole interaction constants.

The quadrupole constants entering into the generalized quadrupole Hamiltonian H_{QT} can be expressed as

$$G^\mu = G^\mu_{ME} + K^\mu = (B^\mu)^2/C_0^\mu + K^\mu \quad (\mu = \gamma, \delta),$$

where G^μ_{ME} is the contribution due to the single-ion magnetoelastic interaction B^μ (C_0^μ is the background elastic constant in the absence of any interaction) and K^μ is the quadrupole pair interaction. The magnetoelastic and quadrupole interactions (distortions) are expressed in terms of symmetrized notations. The magnetoelastic modes of the γ and δ symmetry in the quadrupole Hamiltonian describe the orthorhombic deformation of a tetragonal crystal along the [100] and [110] axes, respectively, arising in the external magnetic field or in a qadrupole-ordered phase. The α mode corresponds to a distortion of the crystal retaining the tetragonal symmetry, that is, to the bulk ($\alpha 1$) and tetragonal ($\alpha 2$) deformations. We have omitted terms of the ε symmetry (corresponding to the monoclinic deformation in the *ac* plane) in the expression for H_{QT}, because these terms do not contribute to the Hamiltonian for the case of the magnetic field orientation in the basal plane and along the tetragonal axis.

In order to estimate the role of quadrupole interactions of the α, γ, and δ symmetry in the formation of magnetic anomalies observed in the crossover region, we used the parameters of interactions derived for the nearest-neighbor vanadate $HoVO_4$, corrected for the trends in the RE series [5]. The magnetoelastic coefficients for the tetragonal modes $\alpha 1$ and $\alpha 2$ in $HoVO_4$ are comparable in magnitude ($B^{\alpha 1}/\alpha_J = -4.05 \times 10^3$ K, $B^{\alpha 2}/\alpha_J = 6.75 \times 10^3$ K), which yields $G^\alpha \sim 0.2$ mK for $ErVO_4$. The magnetoelastic coefficients for the orthorhombic modes in $HoVO_4$ ($B^\gamma/\alpha_J = 13.5 \times 10^3$ K, $B^\delta/\alpha_J = -24.3 \times 10^3$ K) yields $G^\gamma \sim 1$ mK and $G^\delta \sim 20$ mK for $ErVO_4$. Since bilinear interactions in $ErVO_4$ are weak and the ground state in the crystal field is weakly magnetic, the magnetic ordering in this vanadate can be expected only at temperatures significantly below 1 K [6]. However, even weak bilinear interactions for the degenerate ground state in the case of a gapless crossover may lead to the appearance of a transverse component of the magnetic moment. Using the temperature of ordering known for the isomorphous compound $DyVO_4$ (T_N = 3 K), we have estimated the bilinear exchange constant $n = \theta/C$ (θ = –1.4 K) for this compound. Taking into account the trends in θ for the RE series, the de Gennes empirical rule yields $\theta \approx -0.5$ K for $ErVO_4$. The eigenvalues and eigenfunctions of Er^{3+} ion necessary for the calculations of magnetiza-

tion and various susceptibilities were determined by numerical diagonalization of the complete Hamiltonian, in which the bilinear and quadrupole interactions of the α, γ, and δ symmetry (dependent on the electronic structure) were taken into account in a self-consistent manner.

4. DISCUSSION OF RESULTS

4.1. First- and Third-Order Magnetic Susceptibilities

Data on the crystal field parameters in pure vanadate $ErVO_4$ are not available and were only reported for a doped compound $Er : YVO_4$ (Table 1, set 3) [7]. In addition, Guo *et al.* [8] determined the crystal field parameters for $ErVO_4$ using the Morris–Levit extrapolation procedure (Table 1, set 4). We have also attempted to describe the obtained experimental results using the crystal field parameters available for the nearest-neighbor vanadate ($HoVO_4$). The crystal field parameters of this compound, determined on the ground-state multiplet, are commonly accepted as reliably established based on numerous experimental data: $B_2^0 = -126$ K, $B_4^0 = 55$ K, $B_4^4 = 1105$ K, $B_6^0 = -62$ K, $B_6^4 = -112$ K (Table 1, set 2) [5]. However, an analysis showed that the curves calculated using the aforementioned sets of crystal field parameters do not provide sufficiently accurate description of either the initial susceptibility or the magnetic anomalies in the crossover region (see below). The best description of experimental data was provided by the curve calculated using the set of parameters for $HoVO_4$.

Using experimental data for the first-order magnetic susceptibility (Fig. 1), it is possible to refine the crystal field parameters of $ErVO_4$. The curves calculated using set 2 (Table 1) correctly describe the main features in the behavior of $\chi_M^{-1}(T)$, namely, a small anisotropy at high temperatures and the change of easy axis at $T \sim 150$ K. However, the curve of $\chi_M^{-1}(T)$ calculated with the same set of parameters for the hard axis [001] significantly deviates from experiment at temperatures below 80 K (see the inset to Fig. 1). The discrepancy can be removed by slightly (to within 20%) varying the crystal field parameters. The calculated $\chi_M^{-1}(T)$ curve has proved to be most sensitive to the B_4^4 and B_6^0 values. However a single set of the crystal field parameters cannot be determined using only data on the initial magnetic susceptibility. The parameters can further be refined using the magnetic anomalies in the crossover region (see below). The final optimized set of the crystal field parameters (Table 1, set 1) provides an adequate (to within the experimental error) description of the experimental $\chi_M^{-1}(T)$ curves for the [001] and [100] axes (Fig. 1, solid curves).

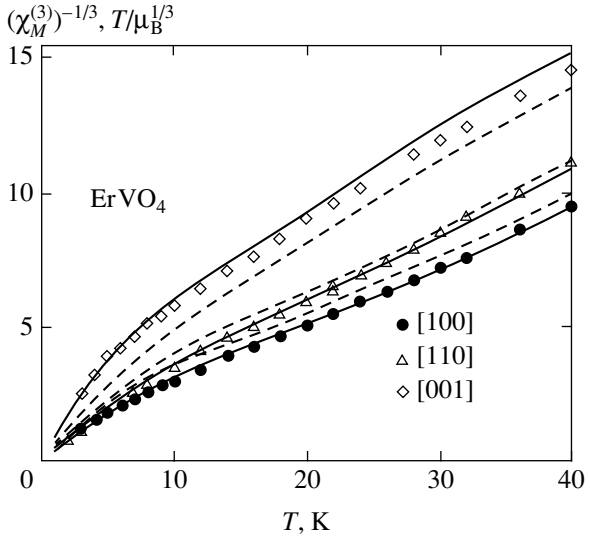

Fig. 4. Temperature dependences of the reciprocal third-order magnetic susceptibility $(\chi_M^{(3)})^{-1/3}$ of an $ErVO_4$ crystal measured along the [100], [110], and [001] directions. Solid and dashed curves show the results of calculations with neglect of pair interactions for two sets of the crystal field parameters (1 and 2, respectively, in Table 1).

The third-order magnetic susceptibility $\chi_M^{(3)}(T)$ of $ErVO_4$ is weakly anisotropic and follows the Curie–Weiss law, whereby $(\chi_M^{(3)})^{-1/3} \propto (T - \theta^*)$ in the temperature interval 10–40 K (Fig. 4). It should be noted that bilinear and quadrupole interactions in $ErVO_4$ do not lead to any significant renormalization of $\chi_M^{(3)}$. A comparison of the curves calculated for various sets of parameters (cf. solid and dashed curves in Fig. 4) shows that the optimized crystal field provides for the best description of these experimental data as well.

4.2. Crystal Field and the Zeeman Effect

Figure 5b shows the splitting of energy levels of Er^{3+} ion in the magnetic field (Zeeman effect) oriented along the [001] and [100] axes calculated with the optimized

Table 1. Crystal field parameters B_n^m of Er^{3+} ion in RE vanadate structure

RXO_4	Set	Ref.	B_2^0, K	B_4^0, K	B_4^4, K	B_6^0, K	B_6^4, K
$ErVO_4$	1	This work	–126	55	1215	–56	–90
$HoVO_4$	2	[5]	–126	55	1105	–62	–112
$Er : YVO_4$	3	[7]	–148	65.5	1394	–62	32
$ErVO_4$	4	[8]	–113	56.3	1187	–59	–122

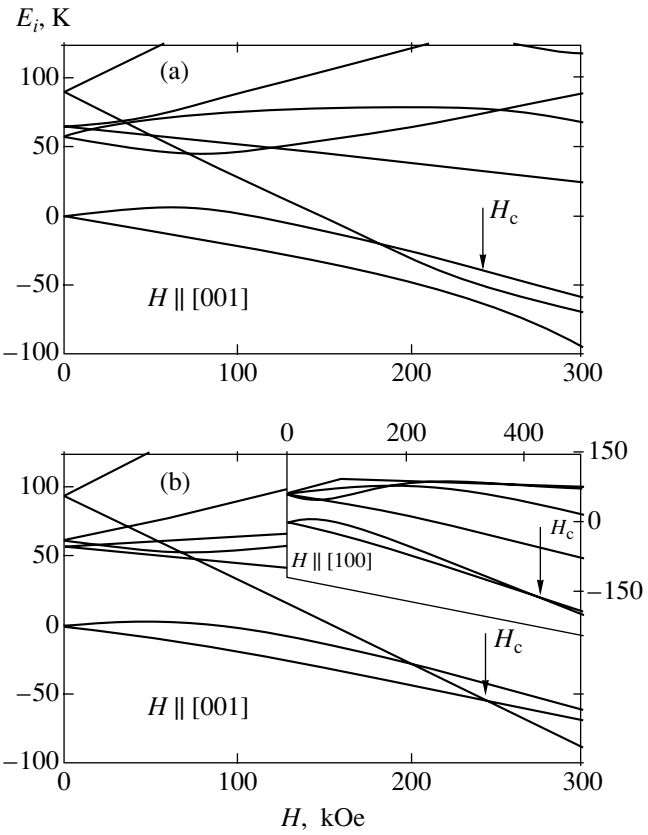

Fig. 5. Zeeman effect in an $ErVO_4$ crystal for the field oriented along the tetragonal [001] axis, as calculated with neglect of pair interactions for (a) set 2 and (b) set 1 of the crystal field parameters (Table 1). The inset shows the Zeeman effect for the three lowest doublets calculated using the optimized set 1 for the field oriented along the [100] axis.

crystal field. The total splitting of the multiplet at $H = 0$ is about 420 K. The data are presented only for the four lowest doublets whose interaction in the case of $H \parallel [001]$ determines the main features of the crossover. In the crystal field of vanadate, the ground-state multiplet $^4I_{15/2}$ of Er^{3+} ion exhibits splitting such that the ground state corresponds to the "weakly" magnetic Kramers doublet with the wavefunction $\{-0.82|\pm 7/2\rangle - 0.42|\mp 9/2\rangle + \ldots\}$ with a small g-tensor component g_z^{gr} separated by a gap (on the order of 60 K) from the first and second excited doublets forming a quasi-quartet (also with small $g_z^{ex1,2}$ values, see Table 2). In contrast, the third excited doublet has a maximum g-tensor component along the z axis ($g_z^{ex3} \gg g_x^{ex3}$), with the wavefunction in the J-, J_z-representation having the form $\pm 0.99|\pm 15/2\rangle$. It is this feature of the spectrum and the wavefunctions of Er^{3+} ion in vanadate that accounts for the levels crossing in the case when the magnetic field is oriented along the tetragonal [001] axis. It is interesting to note that two sets of the crystal field parameters—the optimized (set 1) and that for $HoVO_4$ (set 2)—give very similar spectra of Er^{3+} ion (Fig. 5), differing only by inversion of the first and second excited doublets (Table 2).

The ground-state doublet, as well as the first and second excited doublets, exhibit not very large splitting in the field $H \parallel [001]$. The crossover in $ErVO_4$ is due to the third excited doublet, the lower level of which rapidly decreases in the field $H \parallel [001]$ and becomes the ground level when the field strength exceeds $H_c \sim 235$ kOe. As a result, this sublevel of the third excited doublet with a large projection $\langle M_z \rangle$ approaches the ground level with a small projection $\langle M_z \rangle$, which leads to a jumplike increase in the magnetic moment projection.

Since the magnetic field $H \parallel [001]$ produces mixing of the wavefunctions of interacting levels (sublevels of the ground-state and third excited doublets), the spectrum retains a finite gap at H_c. The width of this gap significantly influences the character of magnetic anomalies in the crossover region. For the optimized crystal field, the gap value is relatively small (~2 K), which accounts for a sharp change in $M(H)$ and a pronounced maximum in $dM(H)/dH$. For the set of parameters corresponding to the crystal field in $HoVO_4$, the gap value is much greater (~13 K). For this reason, the magnetic anomalies in the crossover region are lower in amplitude and more smeared. The gap value is determined by the ground-state wavefunction component $|+15/2\rangle$ before the crossover. For the set of parameters of the crystal field of $HoVO_4$, this component is rather significant because the wavefunction of the excited doublet closest to the ground state contains such a component. Thus, the character of wavefunctions of the first excited doublet significantly influences the magnetic anomalies in the crossover region.

In addition, $ErVO_4$ is expected to exhibit a rather unusual crossover for the field oriented in the basal plane, $H \parallel [100]$, so that the lower component of the ground-state doublet is crossed by the upper component (see the inset in Fig. 5). The crossover is gapless, since the wavefunctions of the interacting levels are not mixed by the field, and is accompanied by small but sharp magnetic anomalies at low temperatures. These anomalies are rapidly smeared due to the field misorientation relative to the symmetry axis and due to the temperature increase caused by the magnetocaloric effect in a pulsed magnetic field and by some other factors. The character of anomalies is virtually the same for the two sets of crystal field parameters (1 and 2, Table 1), while the critical field H_c is somewhat higher for the optimized set (see below).

4.3. Magnetization Curves in Strong Fields

The crossing or mutual approach of the energy levels of Er^{3+} ion in a strong magnetic field, as well as a change in the ground-state in the fields above H_c,

Table 2. Energies, wavefunctions, and g factors of the lower levels of Er^{3+} ion in the vanadate structure calculated for the crystal field parameters according to sets 1 and 2 in Table 1

E_i, K	g_z	$\|i\rangle$
		Set 1, $H = 0$
0	3.85	$\{-0.84\|\pm 7/2> -0.42\|\mp 9/2\rangle + 0.35\|\mp 1/2\rangle + \ldots\}$
56.4	2.84	$\{0.83\|\pm 5/2> -0.53\|\mp 3/2\rangle + 0.16\|\mp 11/2\rangle + \ldots\}$
62.0	6.52	$\{\pm 0.88\|\pm 9/2\rangle \mp 0.47\|\mp 7/2\rangle + \ldots\}$
93.8	18.00	$\{\pm 0.99\|\pm 15/2\rangle + \ldots\}$
		Set 1, $H \approx H_c$
–54.4		$\{0.66\|+15/2\rangle - 0.72\|+7/2\rangle + 0.19\|-1/2\rangle + \ldots\}$
–52.1		$\{-0.75\|+15/2\rangle - 0.64\|+7/2\rangle + 0.17\|-1/2\rangle + \ldots\}$
		Set 2, $H = 0$
0	5.12	$\{-0.89\|\pm 7/2\rangle - 0.36\|\mp 9/2\rangle + 0.30\|\mp 1/2\rangle + \ldots\}$
57.5	7.62	$\{\pm 0.91\|\pm 9/2\rangle \mp 0.39\|\mp 7/2\rangle + \ldots\}$
64.3	3.30	$\{0.85\|\pm 5/2\rangle -0.51\|\mp 3/2\rangle + 0.12\|\mp 11/2\rangle + \ldots\}$
89.0	17.83	$\{\pm 0.99\|\pm 15/2\rangle + \ldots\}$
		Set 2, $H \approx H_c$
–62.9		$\{-0.72\|+15/2\rangle - 0.69\|+7/2\rangle + 0.16\|-1/2\rangle + \ldots\}$
–50.3		$\{0.70\|+15/2\rangle - 0.69\|+7/2\rangle + 0.17\|-1/2\rangle + \ldots\}$

results in the anomalies in both the magnetization $M(H)$ and its derivative $dM(H)/dH$. Figure 6a shows the calculated isothermal (continuous) and adiabatic (dashed) magnetization curves for the three symmetric crystallographic directions [100], [110], and [001]. Figure 6b presents the corresponding magnetocaloric effects ΔT for the initial temperatures T = 4.2 and 15 K. The mutual approach of levels for Er^{3+} ion in the field $H \parallel [001]$ (Fig. 5) leads to a sharp and significant jump $\Delta M \sim 7\mu_B$ in the magnetization isotherms for the critical field. After the crossover, the magnetization almost reaches saturation level, $M_s = 9\mu_B$. For $H \parallel [100]$, the critical field is higher ($H_c \sim 430$ kOe) and the magnetization jump is much smaller ($\Delta M \sim 1.5\mu_B$), so that the magnetization upon crossover does not reach complete saturation. Calculations for still stronger fields show that the magnetization reaches saturation in almost a jumplike manner ($\Delta M \sim 1.5\mu_B$) during crossover taking pace at 3600 kOe. The magnetization along the easy axis [110] exhibits no anomalies, but reaches saturation only in a rather strong field (above 600 kOe). The gap value in the spectrum for $H \parallel [001]$ significantly influences the character of anomalies on the $M(H)$ and $dM(H)/dH$ curves. For the optimized set of crystal field parameters (corresponding to a gap value of about 2 K), the anomalies are very sharp, while an increase in the gap value up to about 13 K (for the crystal field of $HoVO_4$, set 2) leads to a significant decrease and smearing of the anomalies (cf. solid and dashed curves in Figs. 2 and 3).

The jumps in the $M(H)$ curves observed for the fields $H \parallel [100]$ and $H \parallel [001]$ depend differently on temperature. Indeed, for $H \parallel [001]$, the jumps at T = 4.2 K in the isothermal regime remain rather sharp up to a temperature of about 12 K. In the case of a gapless crossover for $H \parallel [100]$, the magnetization jump and the differential susceptibility peak are rapidly smeared as the temperature increases to about 6 K. This smearing is related primarily to close values of the g_x factor for the intersecting levels. In contrast, on approaching zero temperature for a field orientated strictly along the [100] axis, the anomalies calculated using the adopted Hamiltonian become infinitely sharp (see thin solid curves in Fig. 6a for T = 0.5 K).

In order to interpret the magnetic properties of samples in a pulsed magnetic field at rather short pulse duration, it is necessary to perform calculations for the adiabatic magnetization processes. The rate of field rise in our experiments was close to the estimate of the upper bound for the adiabaticity condition in [9]. These estimates have been confirmed by the results of numerous experiments for paramagnetic RXO_4 zircons.

Calculations of the magnetic characteristics in the adiabatic regime for each value of the field strength from 0 to 600 kOe with a step ΔH = 1 kOe involved

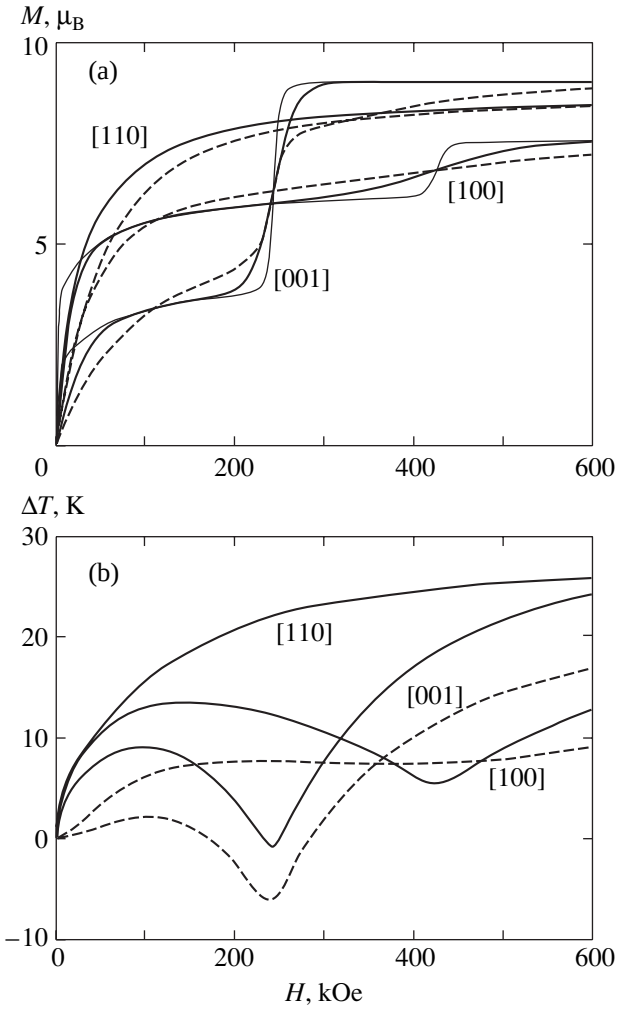

Fig. 6. (a) The isothermal (thick solid lines, $T = 4.2$ K; thin solid lines, $T = 0.5$ K) and adiabatic (dashed lines, $T_{st} = 4.2$ K) magnetization curves for an $ErVO_4$ crystal along the three symmetric crystallographic directions [100], [110], and [001] (calculated with neglect of pair interactions using set 1 of the crystal field parameters); (b) magnetocaloric effects ΔT for the initial temperatures $T_{st} = 4.2$ K (solid lines) and 15 K (dashed lines).

numerical diagonalization of the complete Hamiltonian for determining the energy spectrum and wavefunctions of Er^{3+} ion. Accordingly, the "elementary" magnetocaloric effect ΔT corresponding to the field changing from H to $H + \Delta H$ was calculated as

$$\Delta T = -(\partial M/\partial T)_H T \Delta H/C_H, \qquad (6)$$

where C_H is the total heat capacity of the crystal, including the lattice contribution $C_{lat} = (12\pi^4 k_B \nu/5)(T/T_D)^3$ ($T_D = 275$ K is the Debye temperature for the vanadate lattice [10]; $\nu = 6$) and the magnetic contribution C_{mag} (calculated using the RE ion spectrum for each given field and temperature). Using these data, it is possible to determine the temperature and adiabatic magnetization of $ErVO_4$ crystals for the given field directions.

The jumps in the adiabatic magnetization curves (dashed curves in Fig. 6) are strongly smeared because of a significant heating of the sample crystal. The maximum magnetocaloric effect $\Delta T \sim 25$ K is observed for the field oriented along the easy axis [110] (Fig. 6b). For the fields oriented along the hard [001] and moderate [100] axes, changes in the sample crystal temperature in the fields below crossover were lower, since the magnetization was smaller and the sample began to cool on approaching the critical field H_c. In the fields above H_c for the same orientations, the sample heating started again leading to a strong smearing of anomalies in the crossover region. Here, cooling of the sample in this region is more significant for the hard [001] axis and less pronounced for a higher start temperature T_{st} (see dashed curves for $T_{st} = 15$ K in Fig. 6).

In the case of a finite-gap crossover in $ErVO_4$ for $H \parallel [001]$, the magnetization jump in the isothermal regime is less dependent on temperature. For this reason, the jump is less smeared in the adiabatic cycle with a 20 K increase in the final temperature [cf. isothermal (solid) and adiabatic (dashed) $M(H)$ curves in Fig. 6]. According to the results of calculations, the anomalies on the adiabatic curves of $M(H)$ and $dM(H)/dH$ curves weakly change on heating the sample up to 10 K, then rapidly smear and decrease as the temperature increases above 10 K, and practically vanish at $T >$ 20 K. This behavior is related primarily to the character of the magnetocaloric effect near the crossover. For a starting temperature $T_{st} < 10$ K, an $ErVO_4$ crystal exhibits significant cooling near the crossover: at the maximum of $dM(H)/dH$, the temperature decreases to about 3 K. For $T_{st} > 10$ K, the crystal is cooled less effectively, the temperature at this maximum is higher, and the anomalies start smearing.

The curves of low-temperature adiabatic magnetization for the [001] axis, calculated for the optimized set of crystal field parameters describe the main experimental features, including a decrease in the magnetization growth rate for the fields about 150 kOe, the jumplike increase in $M(H)$ near the critical field $H_c \sim$ 240 kOe, and the approach to saturation for the field strength above the critical level (see solid curves in Fig. 2). The small difference between the calculated and experimental H_c values decreases upon taking into account the quadrupole interactions. Calculations using the other sets of parameters (e.g., those for $HoVO_4$) give anomalies in the crossover region, which are more smeared as compared to experiment (dashed curves in Fig. 2).

The degenerate ground state in the crossover region is unstable with respect to various pair interactions capable of lifting the degeneracy and leading to a finite gap appearing in the fields close to H_c. The results of numerical calculations for various compounds belong-

ing to the family of RE zircons show that both bilinear and quadrupole interactions may remove the degeneracy and lead to the appearance of a transverse (relative to the magnetic field) component of the magnetic moment, that is, to its deviation from the magnetic field direction on approaching the crossover. Since these effects may, in principle, take place near the gapless crossover in $ErVO_4$ for $H \parallel [100]$, we have studied the effect of various pair interactions on the magnetic anomalies near H_c.

Taking into account bilinear interactions in $ErVO_4$ leads to spontaneous magnetic ordering at $H = 0$ and the field-induced ordering at $H = H_c$. According to these calculations, the spontaneous ordering of the magnetic moments of Er^{3+} ions occurs along the [100] axis at $T < 0.2$ K ($M_s \sim 3.5\ \mu_B$). Figure 7 shows the isotherms of longitudinal magnetization $M_\parallel$ and the field dependences of the order parameter $M_\perp$ for $H \parallel [100]$ near the crossover field, as calculated for an isotropic antiferromagnetic exchange with $n_\parallel = n_\perp = \theta/C$ ($\theta = -0.5$ K). Bilinear interactions along the magnetic field are equivalent to an effective field H_B^x added to the external field so as to change the jump width in the magnetization curve (cf. curves *1* and *3* in Fig. 7).

The magnetic moment component $M_\perp$ (perpendicular to the field $H \parallel [100]$) is absent in a weak field but may appear in the region of H_c at a sufficiently low temperature (Fig. 7). The ordered phase with $M_\perp \neq 0$ at $T = 0.1$ K is stable in the interval of magnetic fields from 423 to 433 kOe (Fig. 7, curve *1*'). As the temperature increases, this intervals exhibits narrowing and vanishes at $T > 0.22$ K. Orientation of the $M_\perp$ component along the [001] axis in the plane perpendicular to the magnetic field is determined by the anisotropy. For the [010] axis (Fig. 7, curves *1a*, *1*'*a*), this component has a much smaller value and exists in a narrower interval of field strengths.

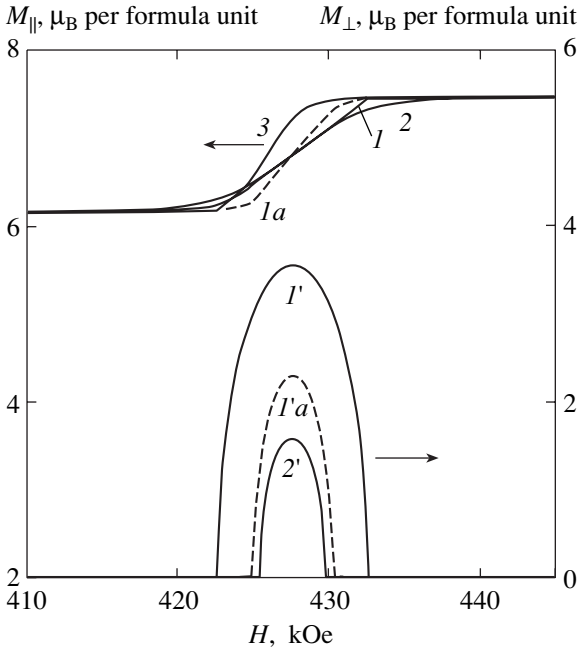

Fig. 7. The isotherms of longitudinal ($M_\parallel$) and transverse ($M_\perp$) magnetization in $ErVO_4$ for $H \parallel [100]$, calculated (*1*, *1*', *2*, *2*') with and (*3*) without allowance for bilinear interactions for $T = 0.1$ K (*1*, *1*', *3*) and 0.2 K (*2*, *2*'); $\theta_\parallel = \theta_\perp = -0.5$ K; curves *1a* and *1*'*a* refer to the case of $M_\perp$ oriented along the [010] axis.

4.4. Differential Susceptibility in Strong Fields

By studying the differential magnetic susceptibility dM/dH, it is possible to trace the dependence of the critical parameters of the crossover on various factors, including the magnetic field misorientation relative to the crystal symmetry axis, the temperature, the pair interaction parameters, etc. For $H \parallel [001]$, the magnetic anomalies vary but slightly for small misorientations of the magnetic field because the energy spectrum of Er^{3+} retains a finite gap at the crossover. Allowance for the quadrupole interactions with a parameter of $G^\alpha = 0.2$ mK in the calculations leads to an almost twofold increase in the $dM(H)/dH$ peak height and decreases the critical field for the crossover by approximately 10 kOe.

For the gapless crossover observed at $H \parallel [100]$, the results of calculations predict a narrow peak in dM/dH for $H_c \sim 430$ kOe, which is small as compared to the initial susceptibility at $H = 0$ (solid curves in the inset to Fig. 3). For the set of crystal field parameters in $HoVO_4$, both the critical field and the dM/dH peak height somewhat decrease (dashed curve *2* in the inset to Fig. 3). Even a small misorientation of the magnetic field relative to the crystal symmetry axis leads to a significant decrease in magnetic anomalies, hindering their experimental observation. In particular, a deviation of the field from the [100] axis in the basal plane by $\Delta\varphi = 2°$ (curve *1*'' in the inset to Fig. 3), as well as an increase in the initial temperature up to 8 K (curve *1*'), virtually smear the peak into a plateau. The measurements of the dM/dH curve at $T = 2.1$ K for an $ErVO_4$ crystal in a pulsed magnetic field $H \parallel [100]$ with an amplitude of 600 kOe reveal a bending point near 450 kOe, which can be interpreted as a plateau or a smeared peak (traces of anomaly) [11].

It should be noted that, in the case of $H \parallel [100]$, cooling of a sample in the crossover region as a result of the magnetocaloric effect significantly influences the peak height at H_c. Should the magnetization process deviate from the adiabatic regime and the magnetic subsystem fail to cool down to about 3 K near H_c, no magnetic anomalies would be observed at the crossover field. A strong increase in the spin–lattice relaxation time at low temperatures may lead to deviation of the magnetization process from the adiabatic to isolated regime,

resulting in a significant smoothening of the magnetic anomalies.

5. CONCLUSIONS

The main results of our investigation of the crossover effect in $ErVO_4$ are as follows. Significant advantages of the system studied are that it is relatively simple and the data are available on the parameters of interactions (at least for the nearest-neighbor compounds), which makes possible a quantitative comparison of theory and experiment. For an $ErVO_4$ crystal, the crossover effects of various types have been predicted and observed for two special orientations of the external magnetic field. A special feature of an $ErVO_4$ crystal is a finite gap value (about 2 K) between the interacting levels for the crossover at $H \parallel [001]$, while the crossover at $H \parallel [100]$ is gapless. For this reason, the magnetic anomalies observed for these orientations have different sensitivity to the magnetic field deviation from the axis of symmetry and to the temperature variations. In particular, for the gapless crossover at $H \parallel [100]$, factors such as an increase in the initial temperature, a small misorientation of the field relative to the axis of symmetry, an increase in the spin–lattice relaxation time (isolated regime) significantly decrease and smoothen the magnetic anomalies, making them difficult for experimental observation. In contrast, allowance for the quadrupole interactions of α symmetry at $H \parallel [001]$ makes the magnetic anomalies at H_c more sharply pronounced.

Experiments show that the crossover with a finite gap value in $ErVO_4$ at $H \parallel [001]$ is accompanied by pronounced magnetic anomalies. These anomalies are determined to a considerable extent by the electronic structure of the RE ion formed by the crystal field and, hence, can be used for refining the crystal field parameters. Since no direct methods are available for determining these parameters, it is necessary to solve the inverse problem. According to the adopted approach, a minimum of the function of deviation of the experimental data from the values calculated for various sets of crystal field parameters is determined. The existence of a single-valued, reliable solution to this problem significantly depends on the crystal field symmetry, or the number of independent parameters, and the set of experimental data used for the determination of these parameters. Even for crystals of hexagonal and tetragonal symmetry described by four or five crystal field parameters, the inverse problem frequently has no single-valued solution and yields several local minima differing within the experimental error. The development of methods for determining and refining the crystal field parameters based on experimental data for such crystals is currently an important task. This study showed that the behavior of crystals in strong magnetic fields, in particular, magnetization anomalies in the crossover region determined to a considerable extent by the excited levels, may be highly informative for the refinement of crystal field parameters.

ACKNOWLEDGMENTS

This study was supported in part by the Russian Foundation for Basic Research (project no. 03-02-16809) and the International Scientific-Technological Center (grant no. 2029).

REFERENCES

1. Z. A. Kazei, N. P. Kolmakova, V. V. Platonov, *et al.*, Physica B (Amsterdam) **284**–**288**, 1483 (2000).
2. A. Kirste, N. Puhlmann, I. Stolpe, *et al.*, Physica B (Amsterdam) **294**–**295**, 132 (2001).
3. A. A. Demidov, Z. A. Kazeĭ, and N. P. Kolmakova, Vestn. Mosk. Univ., Ser. 3: Fiz., Astron., No. 3, 53 (2002).
4. P. Morin, J. Rouchy, and D. Schmitt, Phys. Rev. B **37**, 5401 (1988).
5. P. Morin, J. Rouchy, and Z. Kazei, Phys. Rev. B **51**, 15103 (1995).
6. W. Kockelmann, W. Schafer, and G. Will, Eur. J. Solid State Inorg. Chem. **25**, 515 (1991).
7. D. Kuse, Z. Phys. **203**, 49 (1967).
8. M.-D. Guo, A. T. Aldred, and S.-K. Chan, J. Phys. Chem. Solids **48**, 229 (1987).
9. R. Z. Levitin, V. V. Snegirev, A. V. Kopylov, *et al.*, J. Magn. Magn. Mater. **170**, 223 (1997).
10. A. Kasten, H. G. Kahle, P. Klofer, and D. Schafer-Siebert, Phys. Status Solidi B **144**, 423 (1987).
11. Z. A. Kazei, A. Kirste, M. von Ortenberg, *et al.*, Physica B (Amsterdam) **346**–**347**, 241 (2004).

Translated by P. Pozdeev

Journal of Experimental and Theoretical Physics, Vol. 100, No. 2, 2005, pp. 331–369.
From Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 127, No. 2, 2005, pp. 372–411.

ELECTRONIC PROPERTIES OF SOLIDS

Interaction Corrections to Thermal Transport Coefficients in Disordered Metals: The Quantum Kinetic Equation Approach¶

G. Catelani and I. L. Aleiner

Physics Department, Columbia University, New York, NY 10027, USA
e-mail: aleiner@phys.columbia.edu

Received August 30, 2004

Abstract—We consider singular electron–electron interaction corrections to transport coefficients in disordered metals to test the validity of the Wiedemann–Franz law. We develop a local, quantum kinetic equation approach in which the charge and energy conservation laws are explicitly satisfied. To obtain the local description, we introduce bosonic distribution functions for the neutral low-energy collective modes (electron–hole pairs). The resulting system of kinetic equations enables us to distinguish between the different physical processes involved in charge and energy transport: elastic electron scattering affects both, while the inelastic processes influence only the latter. Moreover, neutral bosons, although incapable of transporting charge, contribute significantly to energy transport. In our approach, we calculate on equal footing the electric and thermal conductivities and the specific heat in each dimension. We find that the Wiedemann–Franz law is always violated by the interaction corrections; the violation is larger for one- and two-dimensional systems in the diffusive regime $T\tau \ll \hbar$ and is due to the energy transported by neutral bosons. For two-dimensional systems in the quasi-ballistic regime $T\tau \gg \hbar$, the inelastic scattering of the electron on the bosons also contributes to the violation. © *2005 Pleiades Publishing, Inc.*

1. INTRODUCTION

It is well-known that measurement of the thermal transport coefficient may provide additional information about the scattering processes in disordered metals. In particular, the Wiedemann–Franz [1] law holds as long as elastic scattering dominates in the system,[1]

$$L = \frac{\kappa}{\sigma T} = \frac{\pi^2}{3e^2}, \qquad (1.1)$$

where κ and σ are the respective thermal and electric conductivities in the system, T is the temperature in energy units ($k_B = 1$), and e is the electron charge. On the other hand, for the deep inelastic forward scattering, the Wiedemann–Franz law is violated [2], and hence the Lorentz number L is smaller than the universal value, $L < \pi^2/3e^2$.

Historically, the transport (in particular, thermal transport) coefficients were first calculated using the Boltzmann equation (BE) [3]. The advantage of this approach is that it allows a clear separation of the scales in the problem: a particle moves freely most of the time and rarely scatters on other particles or impurities. The BE is applicable on a time scale much larger than the time that it takes for the scattering to happen, and hence all the scattering events are encoded into the local collision integral. All the quantum mechanical part of the calculation is then reduced to the solving of the scattering problems for the relevant physical processes. This gives the precise form of the collision integral but does not affect the general structure of the BE. The great advantage of the BE is that its structure illuminates the relevant conservation laws.

In the late 1950s, an alternative approach became popular—the so-called Kubo formulas [4]. In this approach, the transport equation is not derived but rather the connection of the transport coefficient to the equilibrium correlation function of certain current operators is used. (The Kubo approach to thermal transport was claimed to be put on rigorous footing by Luttinger [5] based on the assumption that there exists some spatial scale in the system such that the gradient expansion is possible for perturbations smooth at that scale.) Being exact, the Kubo formulas are formally applicable even in the regime where the transport equation cannot be justified (the evolution cannot be separated into free motion and rare collisions).

However, in practice, the possibility of explicit calculations within the Kubo formula is somewhat limited. The most spectacular results of the Kubo-formula cal-

¶ This article was submitted by authors in English.

[1] It was shown by G.V. Chester and A. Thellung [Proc. Phys. Soc. London **77**, 1005 (1961)] that Eq. (1.1) remains valid for arbitrary scattering strength as long as the scattering rates and the density of states are smooth ($\mathcal{C}_2$) functions of energy near the Fermi level.

culations—such as the Maki–Thompson [6, 7], Aslamazov–Larkin [8], and weak localization [9] corrections to the electrical conductivity—require a small parameter, which is the same parameter that determines the applicability of the BE. This means that all these effects can also be described in terms of quantum corrections to the collision integral (for weak localization, this was done in [13]). The most relevant effect for this paper, the Altshuler–Aronov [10] interaction correction to the electrical conductivity in two dimensions [11, 14]

$$\delta\sigma_{AA} = -\frac{e^2}{2\pi^2\hbar}\ln\left(\frac{\hbar}{T\tau}\right) \times\left[1+3\left(1-\frac{1}{F_0^\sigma}\ln(1+F_0^\sigma)\right)\right], \quad (1.2)$$

originates from the elastic scattering of electrons on a self-consistent potential (Friedel oscillation) [15, 16] and can be once again obtained from the correction to the collision integral [17].

The success of the Kubo formulas in describing the quantum and interaction effects in thermal transport is by far more modest and controversial. Particularly, despite a 20-year history, there is no consensus on the answer to a natural question: how does the logarithmic correction to conductivity Eq. (1.2) translate into a correction to Wiedemann–Franz law Eq. (1.1)?

The first attempt to answer this question was made by Castellani *et al.* [19] by analyzing Ward identities for a disordered Fermi liquid; they found that the Wiedemann–Franz law should hold for interacting disordered electrons. Their claim was later disputed by Livanov *et al.* [20]: in a "quantum kinetic equation" approach,[2] a logarithmic divergence for the thermal conductivity in two dimensions was even found to have a sign opposite to the Wiedemann–Franz law. More recently, Niven and Smith [22] applied the Kubo formula and again found a logarithmically divergent contribution (for the Coulomb, but not a short-range interaction) in addition to the one that follows from the Wiedemann–Franz law.

The reason for this confusion in the literature is twofold. Technically, identification of the correct form of current operator is complicated by the presence of electron–electron interaction (the energy current operator in the form defined by Luttinger [5] is cumbersome for use due to the presence of additional disorder and interaction potentials in it, whereas the superficially more elegant expression in the Matsubara frequency representation does not in fact correspond to any conservation law for the interacting system and violates gauge invariance, see Appendix B). Physically, the use of the diagram calculation within the Kubo formula prevents one from clearly identifying the relevant scattering processes, because each diagram taken separately describes some mixture of such processes and does not have a physical meaning individually.

This situation calls for the development of the kinetic equation description, which takes the interaction correction of the Altshuler–Aronov type into account for both the electric and thermal transport. The advantage of this approach is that it allows keeping track of the conservation laws explicitly and thus excludes any ambiguity in defining the currents. This paper is devoted to the development and application of this method.

We use units with $\hbar = 1$ throughout the paper and restore the Planck constant only in the final results. This paper is organized as follows: in Section 2, we discuss some general features of the kinetic equation approach using a simple "toy model." In Section 3, we present our final expression for the kinetic equation describing interacting electrons in disordered metals. Section 4 summarizes the results for the thermal conductivity and the specific heat obtained by solving the kinetic equation. The derivation of the kinetic equation is presented in Section 5, and calculation of the transport coefficients and the specific heat is given in Section 6. Some mathematical details are relegated to the Appendices.

2. STRUCTURE OF THE KINETIC EQUATION: CURRENTS AND SPECIFIC HEAT

The purpose of this section is to show how the structure of the kinetic equation permits the proper identification of the relevant currents. We first recall how to calculate the specific heat from the kinetic equation once the conservation laws are obtained (this makes possible a direct check against the much simpler thermodynamic calculation). We then discuss the locality requirement for a proper kinetic equation. This requirement determines the number of the necessary degrees of freedom (i.e., independent distribution functions) that must be introduced into the kinetic description.

2.1. Kinetic Equation and Conservation Laws

As a specific example, we here consider electronlike and holelike excitations coupled to neutral bosons in the presence of an external electric field **E**. (As we see later, the system of interacting electrons can be effectively described at low temperatures by such a coupled system for the scattering at small momentum transfer in the particle-hole channel.) The kinetic equations for electrons and bosons have the form

$$\left[\frac{\partial}{\partial t}+v_F\mathbf{n}\cdot\nabla+ev_F\mathbf{n}\cdot\mathbf{E}\frac{\partial}{\partial\varepsilon}\right]f = \mathrm{St}_e\{f,N\}, \quad (2.1a)$$

$$\left[\frac{\partial}{\partial t}+v(\omega)\mathbf{n}\cdot\nabla\right]N = \mathrm{St}_b\{f,N\}, \quad (2.1b)$$

[2] The quantum kinetic equation with the necessary conservation laws was not actually derived in [20], and we are therefore unable to compare their approach with ours.

where $f = f(\varepsilon, \mathbf{n}; t, \mathbf{r})$ is the distribution function for the electrons with charge e, v_F is the Fermi velocity, and $\mathbf{n}$ is the direction of the momentum. The energy ε is counted from the Fermi level such that $f(\varepsilon > 0)$ describes electron-like excitations and $1 - f(-\varepsilon)$, $\varepsilon > 0$ corresponds to holelike excitations. Concentrating only on the corrections that are singular in T, we neglect the dependence of the electron velocity on the energy (the electron–hole asymmetry) because it does not introduce anything but a small correction that is regular in powers of T^2.

The bosonic function $N = N(\omega, \mathbf{n}; t, \mathbf{r})$ is the distribution function for the bosons with the velocity $\mathbf{v}(\omega)$. All the interaction effects are included into the collision integrals St_e and St_b; for example, an electron-like excitation can decay into a less energetic electron and a neutral boson, or an electron and a hole can be annihilated into a neutral bosons, etc. By locality, the collision integrals depend on the same variables as the distribution functions, i.e., $\mathrm{St}_e = \mathrm{St}_e(\varepsilon, \mathbf{n}; t, \mathbf{r})$ and $\mathrm{St}_b = \mathrm{St}_b(\omega, \mathbf{n}; t, \mathbf{r})$.

In thermodynamic equilibrium with $\mathbf{E} = 0$, the Fermi function for fermions and the Planck function for the neutral bosons,

$$f_F(\varepsilon) = \frac{1}{\exp(\varepsilon/T) + 1}, \qquad N_P(\omega) = \frac{1}{\exp(\omega/T) - 1}, \tag{2.2}$$

solve the kinetic equation. The temperature here T is a constant determined by the initial conditions for the kinetic equation.

Being an effective description for the slow dynamics of the original quantum mechanical system, the kinetic equation must respect the conservation laws of the original system: (i) the total charge conservation and (ii) the total energy conservation. These two conditions are enforced by the requirements

$$\int d\varepsilon \nu \langle \mathrm{St}_e\{f, N\}\rangle_{\mathbf{n}} = 0 \tag{2.3a}$$

and

$$\int d\varepsilon\varepsilon\nu \langle \mathrm{St}_e\{f, N\}\rangle_{\mathbf{n}} + \int d\omega\omega b(\omega)\langle \mathrm{St}_b\{f, N\}\rangle_{\mathbf{n}} = 0 \tag{2.3b}$$

for the collision integrals; here, ν is the density of states (DOS) of the electrons (we neglect its energy dependence) and $b(\omega)$ is the density of states of the bosons. We also introduce the short notation for the angular integral

$$\langle \ldots \rangle_{\mathbf{n}} \equiv \int \frac{d\mathbf{n}}{\Omega_d} \ldots, \tag{2.4}$$

where Ω_d is the total solid angle in d dimensions.

Let the electron density ρ be given by

$$\rho(t, \mathbf{r}) = e\nu \int d\varepsilon \langle f(\varepsilon, \mathbf{n}; t, \mathbf{r})\rangle_{\mathbf{n}}. \tag{2.5}$$

Integrating Eq. (2.1a) over the energy and the direction of the momentum and using Eq. (2.3a), we arrive at the continuity equation

$$\frac{\partial \rho}{\partial t} + \nabla \cdot \mathbf{j} = 0, \tag{2.6}$$

with the electron current density defined as

$$\mathbf{j}(t, \mathbf{r}) = e\nu v_F \int d\varepsilon \langle \mathbf{n} f(\varepsilon, \mathbf{n}; t, \mathbf{r})\rangle_{\mathbf{n}}. \tag{2.7}$$

(Strictly speaking, i.e., Eq. (2.6) fixes only the longitudinal component of the current—an arbitrary curl may be added to Eq. (2.7). We do not consider the effect of the magnetic field here and therefore disregard such magnetization currents.)

We now turn to the analysis of the energy conservation. We multiply Eq. (2.1a) by $\nu\varepsilon$ and integrate over ε and $\mathbf{n}$. Next, we multiply Eq. (2.1b) by $\omega b(\omega)$ and integrate over ω and $\mathbf{n}$. Adding the two results together and using Eqs. (2.3b) and (2.7), we obtain

$$\frac{\partial u_{\mathrm{tot}}}{\partial t} + \nabla \cdot \mathbf{j}^{\varepsilon}_{\mathrm{tot}} = \mathbf{j} \cdot \mathbf{E}, \tag{2.8}$$

where

$$u_{\mathrm{tot}} = u_e(t, \mathbf{r}) + u_b(t, \mathbf{r}), \tag{2.9a}$$

$$u_e(t, \mathbf{r}) = \nu \int d\varepsilon\varepsilon \langle f(\varepsilon, \mathbf{n}; t, \mathbf{r})\rangle_{\mathbf{n}}, \tag{2.9b}$$

$$u_b(t, \mathbf{r}) = \int d\omega\omega b(\omega)\langle N(\omega, \mathbf{n}; t, \mathbf{r})\rangle_{\mathbf{n}}, \tag{2.9c}$$

and

$$\mathbf{j}^{\varepsilon}_{\mathrm{tot}} = \mathbf{j}^{\varepsilon}_e + \mathbf{j}^{\varepsilon}_b, \tag{2.10a}$$

$$\mathbf{j}^{\varepsilon}_e(t, \mathbf{r}) = v_F \nu \int d\varepsilon\varepsilon \langle \mathbf{n} f(\varepsilon, \mathbf{n}; t, \mathbf{r})\rangle_{\mathbf{n}}, \tag{2.10b}$$

$$\mathbf{j}^{\varepsilon}_b(t, \mathbf{r}) = \int d\omega\omega b(\omega) v(\omega)\langle \mathbf{n} N(\omega, \mathbf{n}; t, \mathbf{r})\rangle_{\mathbf{n}}. \tag{2.10c}$$

The right-hand side of Eq. (2.8) is nothing but the Joule heat. For a homogeneous system, the gradient term in the left-hand side vanishes, and by virtue of the energy conservation, expression (2.9) must be identified with the total energy density of the system. On the other hand, for $\mathbf{E} = 0$, Eq. (2.8) has the form of the continuity equation for the energy density; therefore, Eqs. (2.10) must be identified with the total energy current density. This statement is not entirely trivial. One could imagine that for an interacting system, the DOS entering the respective expressions (2.5) and (2.9) for

the charge and the energy density are renormalized differently. Energy conservation Eq. (2.8) eliminates such a possibility.

The conservation of energy, Eq. (2.8), is valid for any rate of the energy flow into and out of the system. On the other hand, the collision integrals in Eqs. (2.1) define a certain time scale τ_{in}: the dynamics slow at the scale of τ_{in} can be characterized by distribution functions (2.2) with a time-dependent temperature $T(t)$ (corrections to such an adiabatic description are of the order of $\tau_{in}\partial_t \ln T$). Substituting this form of the distribution function in Eq. (2.9) and then using the result in Eq. (2.8), we find, for a homogeneous system,

$$c_V(T)\frac{\partial T}{\partial t} = \mathbf{j}\cdot\mathbf{E}, \tag{2.11}$$

where

$$c_V = \frac{\partial}{\partial T}[\nu\int d\varepsilon\varepsilon f_F(\varepsilon) + \int d\omega\omega b(\omega)N_P(\omega)] \tag{2.12}$$

is nothing but the specific heat of the system. The latter quantity may be calculated independently by applying the standard diagram technique for equilibrium systems. The agreement of such a calculation with the structure of the kinetic equation result in (2.12) is the most important check of the consistency of our description of the thermal transport.

2.2. Locality of the Kinetic Equation and the Number of the Degrees of Freedom

The form of the collision integrals local in space and time is clearly a simplified description. Actually, the collision integral may be nonlocal at the time scale of the order of $\hbar/T$ and at the spatial scale of the order of $\hbar v_F/T$. We say that such a description is *local* and the description where the nonlocality is involved at larger spatial and time scales is *nonlocal.*

The number of distribution functions to be introduced into the description is governed by the locality of the kinetic equation. We use the model in Eqs. (2.1) to illustrate the point. We had a local description in terms of the fermionic and bosonic distribution functions. But we can try to eliminate the bosonic distribution function and obtain a description in terms of the electronic degrees of freedom only.

Assuming that the deviation of the distribution function from its equilibrium value is small, we can linearize the bosonic collision integral to the form

$$\mathrm{St}_b\{f, N\} = -\hat{I}(N - \tilde{N}\{f\}), \tag{2.13}$$

where $\hat{I}$ is some positive definite integral operator and $\tilde{N}\{f\}$ is the functional of the fermionic distribution function $f(\varepsilon)$ such that for $f(\varepsilon) = f_F(\varepsilon)$, $\tilde{N}(\omega) = N_P(\omega)$. Using Eq. (2.13), we can formally solve Eq. (2.1b) as

$$N = \frac{1}{\partial/\partial t + v(\omega)\mathbf{n}\cdot\nabla + \hat{I}}\hat{I}\tilde{N}\{f\}. \tag{2.14}$$

Substituting Eq. (2.14) in Eq. (2.1a), we apparently obtain the kinetic equation in terms of the electron distribution function only,

$$\left[\frac{\partial}{\partial t} + v_F\mathbf{n}\cdot\nabla + e v_F\mathbf{n}\cdot\mathbf{E}\frac{\partial}{\partial\varepsilon}\right]f = \mathrm{St}_e^{?}\{f\},$$
$$\mathrm{St}_e^{?}\{f\} \equiv \mathrm{St}_e\left\{f, \frac{1}{\partial/\partial t + v(\omega)\mathbf{n}\cdot\nabla + \hat{I}}\hat{I}\tilde{N}\{f\}\right\}. \tag{2.15}$$

If we are interested in the linear response to a weak and smooth external perturbation, the description in terms of this single kinetic equation is completely equivalent to the original coupled system (2.1). However, there are clear drawbacks: the presence of the integral operator $\hat{I}$ in the collision integral makes it nonlocal on the scale determined by the kinetic equation itself rather than by temperature. Moreover, though it is still easy to identify the continuity equation for the electron charge using Eq. (2.3a), there is no longer a relation similar to Eq. (2.3b). This is why the analysis of the energy conservation law becomes cumbersome: calculation of the specific heat and energy current requires the time expansion of the collision integral, which in turns seems to require knowledge of the concrete form of the inelastic collision integral.

The example we have just considered is somewhat trivial because the system was separated into fermionic and bosonic modes from the very beginning. The problem that we consider in this paper is how to include the collective modes of the interacting electron system into the kinetic equation. Indeed, in this case, any calculation gives the result in terms of the electronic distribution function only, and it is not clear a priori how to introduce the occupation numbers for the collective modes into the description.

As we show in what follows, it may be possible to reverse our previous argument. We consider a system of interacting electrons find that the interactions are described by a *nonlocal* collision integral. We therefore introduce bosonic degrees of freedom that allow rewriting of the nonlocal kinetic equation in terms of coupled, local kinetic equations. This then makes it possible to identify the energy density and energy current density as sums of the fermionic and bosonic contributions.

A specific example is briefly discussed in the next subsection.

2.3. Degrees of Freedom for the Kinetics of a Disordered Fermi Liquid

We now focus on the disordered, interacting Fermi liquid. For simplicity, we consider the interaction in the singlet channel only. Our goal is to show that the thermodynamic result for the interaction correction to the specific heat has indeed the kinetic equation structure (2.12). As a result, we are then able to determine the necessary number of the bosonic degrees of freedom for the local kinetic equation. For the paper to be self-contained, we briefly review the thermodynamic approach, referring the reader to the literature for further details.[3]

The thermodynamic calculation of the specific heat c_V is based on the relation between c_V and the thermodynamic potential Ω:

$$c_V = -T\frac{\partial^2\Omega}{\partial T^2}. \tag{2.16}$$

The thermodynamic potential can be written as the sum of the thermodynamic potential Ω_0 for noninteracting quasiparticles and a correction $\delta\Omega$ associated with soft modes in the system. Keeping such a correction is legitimate because it turns out to be a more singular function of the temperature than the T^3 correction due to the electron–hole asymmetry.

The correction $\delta\Omega$ is given by the sum of the so-called ring diagrams, see Fig. 1. The Matsubara representation for this diagram is

$$\delta\Omega = \frac{T}{2}\sum_{\omega_n}\int\frac{d^dq}{(2\pi)^d}\ln\left(1+\frac{F}{\nu}\Pi(i|\omega_n|,\mathbf{q})\right), \tag{2.17}$$

where F is the coupling constant, $\omega_n = 2\pi Tn$ are the bosonic Matsubara frequencies, and Π is the polarization operator. The explicit expression for this operator is not important for the present discussion and is given later, see Eq. (5.43a).

A straightforward calculation, relegated to Appendix A, enables us to rewrite Eq. (2.17) as

$$\delta\Omega = -\int\frac{d\omega}{2\pi}\left(\frac{1}{2}\coth\frac{\omega}{2T}\right) \times\int\frac{d^dq}{(2\pi)^d}\mathrm{Im}\,\mathrm{Tr}[\ln\hat{\mathscr{L}}^\rho - \ln\hat{\mathscr{L}}^g]. \tag{2.18}$$

The explicit expressions for the bosonic propagators $\mathscr{L}^\rho$ and $\mathscr{L}^g$ are not relevant (they can be found from Eq. (3.11), with the trace understood as the sum or integration over all variables other than ω, $\mathbf{q}$); we just mention here that $\mathscr{L}^\rho = \mathscr{L}^g$ in the absence of interaction ($F = 0$). Substituting Eq. (2.18) in Eq. (2.16) and integrating over ω by parts, we find

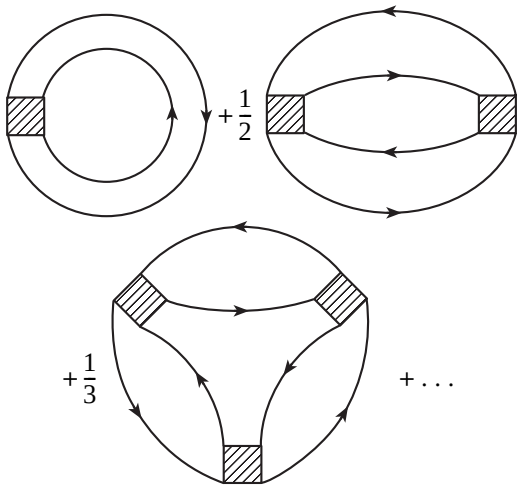

Fig. 1. Leading singular contribution to the thermodynamic potential for the clean system. The shaded box corresponds to F/ν, defined via the two-particle vertex Γ^w, see [26]; the solid lines are coherent parts of the electron Green functions. For the disordered system, the polarization bubbles should be dressed by impurity scattering [12].

$$\delta c_V = \frac{\partial}{\partial T}\int_0^\infty d\omega\omega N_P(\omega)[b^\rho(\omega) - b^g(\omega)], \tag{2.19a}$$

where the densities of states are defined as

$$b^\rho(\omega) = \frac{1}{\pi}\mathrm{Im}\int\frac{d^dq}{(2\pi)^d}\partial_\omega\mathrm{Tr}\ln\mathscr{L}^\rho, \tag{2.19b}$$

$$b^g(\omega) = \frac{1}{\pi}\mathrm{Im}\int\frac{d^dq}{(2\pi)^d}\partial_\omega\mathrm{Tr}\ln\mathscr{L}^g. \tag{2.19c}$$

The function $b^\rho(\omega)$ has the physical meaning of the density of states (DOS) of the bosonic degrees of freedom in the system (soft electron–hole pairs). The function $b^g(\omega)$ has the meaning of the DOS of fictitious bosons (we call them "ghosts") that describe soft electron–hole pairs in the absence of interaction. The physical meaning of the minus sign in front of $b^g(\omega)$ is that with the formation of collective modes, some degrees of freedom are removed from the description of the noninteracting system; the ghost bosons in the last term in Eq. (2.19a) take this reduction into account.

Comparison of Eq. (2.19a) with Eqs. (2.12) and (2.9) suggests the following expression for the contri-

[3] See, e.g., [12, 26] and references therein.

bution of the collective modes to the energy density in the nonequilibrium case:

$$u_b = \int_0^\infty d\omega\omega[N^\rho(\omega)b^\rho(\omega) - N^g(\omega)b^g(\omega)]; \quad (2.20)$$

here, $N^\rho = N^g = N_P$ in the equilibrium and must be found from some kinetic equation otherwise (this definition requires that Eq. (2.8) hold for an arbitrary distribution function). A similar expression can be obtained for the contribution due to the interaction in the triplet channel by introducing an additional propagator $\mathcal{L}^\sigma$ and the distribution function N^σ. This means that the proper local kinetic equation must include four distribution functions: one for the fermions, $f(\varepsilon)$, and three for the bosons, $N^{\rho,\sigma,g}(\omega)$. We derive such a description in the subsequent sections.

3. FINAL FORM OF THE KINETIC EQUATION AND SCATTERING PROCESSES

In this section, we summarize the final form of the quantum kinetic equation, the conservation laws, and the corresponding currents. The explicit expressions of the collision integrals are given in Subsection 2. The detailed derivation of these results is presented in Section 5.

In accord with the previous section, the kinetics of the system is described by the electronic distribution function $f(\varepsilon, \mathbf{n}; t, \mathbf{r})$, the "distribution functions" $\hat{N}^\rho$ and $\hat{N}^\sigma$ of the bosonic singlet and triplet excitations, and the "distribution function" $\hat{N}^g$ of the ghost excitation.

The electron distribution function $f(\varepsilon, \mathbf{n}; t, \mathbf{r})$ is diagonal in the space of momentum directions. On the contrary, bosonic excitations are characterized by the density matrices $N^\alpha(\omega, \mathbf{q}; \mathbf{n}_i, \mathbf{n}_j; t, \mathbf{r})$ ($\alpha = \rho, \sigma, g$), which may not be diagonal in the space of momentum direction $\mathbf{n}$. Only in the thermal equilibrium, with

$$\begin{aligned} f_{\text{eq}}(\varepsilon, \mathbf{n}; t, \mathbf{r}) &= f_{\text{F}}(\varepsilon), \\ N^\alpha_{\text{eq}}(\omega, \mathbf{q}; \mathbf{n}_i, \mathbf{n}_j; t, \mathbf{r}) &= \Omega_d \delta(\widehat{\mathbf{n}_i, \mathbf{n}_j}) N_P(\omega) \end{aligned} \quad (3.1)$$

and with the Fermi and Planck distribution functions given by Eq. (2.2), the matrices $N^\alpha(\omega, \mathbf{q}; \mathbf{n}_i, \mathbf{n}_j; t, \mathbf{r})$ acquire the diagonal form.[4] But even outside the equilibrium, these matrices have the property

$$\begin{aligned} &N^\alpha(\omega, \mathbf{q}; \mathbf{n}_i, \mathbf{n}_j) \\ &= -[N^\alpha(-\omega, -\mathbf{q}; \mathbf{n}_j, \mathbf{n}_i) + \Omega_d(\widehat{\mathbf{n}_i, \mathbf{n}_j})] \end{aligned} \quad (3.2)$$

[4] As given in Eq. (3.1), the equilibrium distribution functions $N^\alpha_{\text{eq}}(\omega)$ are defined only for $\omega > 0$; for $\omega < 0$, they are found using the property in Eq. (3.2).

(hereafter, the spectator t, $\mathbf{r}$ variables may be suppressed).

Strictly speaking, $f(\varepsilon, \mathbf{n}; t, \mathbf{r})$ is a 2×2 density matrix in the spin space and N^σ is a 3×3 density matrix in the angular momentum $L = 1$ space; however, this is not important in the calculations below and we write the equations for the diagonal components only. To account for the threefold degeneracy of the triplet mode, we explicitly introduce factors of 3 in the corresponding collision integrals and currents.

For compactness, we use the operator notation for matrices in the space of momentum directions, such that, for example, $\hat{N}$ should be understood as an operator acting on a function $a(\mathbf{n}_i)$ as

$$[\hat{N}a](\mathbf{n}_i) \equiv \int \frac{d\mathbf{n}_j}{\Omega_d} N(\mathbf{n}_i, \mathbf{n}_j) a(\mathbf{n}_j). \quad (3.3)$$

3.1. Kinetic Equations and Conserved Currents

The kinetic equation for the electrons in the electric field $\mathbf{E}$ (we do not consider the magnetic field effects) has the canonical form

$$\begin{aligned} &\left[\partial_t + \mathbf{v}\cdot\nabla + e\mathbf{v}\cdot\mathbf{E}\frac{\partial}{\partial\varepsilon}\right] f(\varepsilon, \mathbf{n}; t, \mathbf{r}) = \hat{\text{St}}_e(\varepsilon, \mathbf{n}; t, \mathbf{r}), \\ &\hat{\text{St}}_e = \hat{\text{St}}_\tau f + \hat{\text{St}}^{e-\rho}\{f, N^\rho\} + 3\hat{\text{St}}^{e-\sigma}\{f, N^\sigma\} \\ &\qquad - 4\hat{\text{St}}^{e-g}\{f, N^g\} + \hat{\text{St}}^{e-e}\{f\}, \end{aligned} \quad (3.4)$$

where the first term in the right-hand side is the "bare" collision integral

$$\text{St}_\tau(\mathbf{n}_i, \mathbf{n}_j) = \frac{1}{\tau(\theta_{ij})} - (\widehat{\mathbf{n}_i, \mathbf{n}_j})\int \frac{d\mathbf{n}_k}{\tau(\theta_{ik})}, \quad (3.5)$$

with $\theta_{ij} = \widehat{\mathbf{n}_i\mathbf{n}_j}$, and the other terms, which will be written shortly, take the interaction effects into account.

The bosonic distributions, for $\alpha = \rho, \sigma, g$, are governed by

$$\begin{aligned} &\omega\left[\left\{\frac{1}{1+\hat{F}^\alpha}; \partial_t\hat{N}^\alpha\right\} + \{\hat{\mathbf{s}}^\alpha(\omega, \mathbf{q}); \nabla\hat{N}^\alpha\}\right. \\ &\qquad \left. + i[\hat{H}^\alpha_{e-h}(\omega, \mathbf{q}); \hat{N}^\alpha]\right] \\ &\qquad = \hat{\text{St}}^{\alpha-e}\{N^\alpha, f\}(\omega, \mathbf{q}; \mathbf{n}_i, \mathbf{n}_j; t, \mathbf{r}), \end{aligned} \quad (3.6)$$

where the commutator and anticommutator are defined as

$$\{\hat{A}; \hat{B}\} \equiv \frac{1}{2}(\hat{A}\hat{B} + \hat{B}\hat{A}), \quad [\hat{A}; \hat{B}] \equiv \hat{A}\hat{B} - \hat{B}\hat{A}. \quad (3.7)$$

The operators $\hat{H}^\alpha_{e-h}$ acting in the angular (momen-

tum direction) space are defined as

$$\hat{H}^{\alpha}_{e-h}(\omega, \mathbf{q}) = \mathbf{v}\cdot\mathbf{q} - \frac{\omega}{1+\hat{F}^{\alpha}} \tag{3.8}$$

and the velocity operator is

$$\hat{\mathbf{s}}^{\alpha}(\omega, \mathbf{q}) = \frac{\partial \hat{H}^{\alpha}_{e-h}(\omega, \mathbf{q})}{\partial \mathbf{q}} = \mathbf{v} + \omega\frac{\partial}{\partial \mathbf{q}}\left(\frac{\hat{F}^{\alpha}}{1+\hat{F}^{\alpha}}\right). \tag{3.9}$$

The action of the operators $\hat{F}^{\alpha}$ in the angular space is the same as in Eq. (3.3); these operators are given by

$$\begin{aligned} \hat{F}^{g} &= 0, \\ [\hat{F}^{\sigma}](\mathbf{n}_i, \mathbf{n}_j) &= F^{\sigma}(\theta_{ij}), \\ [\hat{F}^{\rho}](\mathbf{n}_i, \mathbf{n}_j) &= \nu V(\mathbf{q}) + F^{\rho}(\theta_{ij}), \end{aligned} \tag{3.10}$$

where $F^{\rho,\sigma}(\theta)$ are the Landau Fermi-liquid interaction parameters. The angular-independent term $\nu V(\mathbf{q})$ takes the long-range part of the Coulomb density–density interaction into account.

To characterize the density of states for the bosonic excitations, we introduce the propagators $\hat{\mathscr{L}}^{\alpha}(\omega, \mathbf{q}; \mathbf{n}_i, \mathbf{n}_j)$, $\alpha = \rho, \sigma, g$ as

$$[i\hat{H}^{\alpha}_{e-h}(\omega, \mathbf{q}) - \hat{\mathrm{St}}_{\tau}]\hat{\mathscr{L}}^{\alpha} = \hat{1}. \tag{3.11}$$

They describe the propagation of an electron–hole pair scattered by the disorder potential. This propagation is affected by the corresponding interactions for $\alpha = \rho, \sigma$, and it reduces to the usual diffusion for the ghosts.

We are now prepared to write the conservation laws that must be satisfied by the collision integrals independently of their explicit form or the particular shape of the distribution functions. The conservation of the number of particles is ensured by the condition

$$\begin{aligned} \int \hat{\mathrm{St}}^{e-\alpha}\{f, N^{\alpha}\}(\varepsilon, \mathbf{n}; t, \mathbf{r})d\mathbf{n}d\varepsilon &= 0, \\ \alpha &= g, \rho, \sigma, \\ \int \hat{\mathrm{St}}^{e-e}\{f\}(\varepsilon, \mathbf{n}; t, \mathbf{r})d\mathbf{n}d\varepsilon &= 0, \end{aligned} \tag{3.12a}$$

and the impurity collision integral (3.5) preserves the number of particles on each energy shell,

$$\int \hat{\mathrm{St}}_{\tau} f(\varepsilon, \mathbf{n}; t, \mathbf{r})d\mathbf{n} = 0. \tag{3.12b}$$

The conservation of energy during purely electron–electron collisions is ensured by

$$\int \varepsilon\hat{\mathrm{St}}^{e-e}\{f\}(\varepsilon, \mathbf{n}; t, \mathbf{r})d\mathbf{n}d\varepsilon = 0. \tag{3.12c}$$

Finally, the conservation of energy during the electron–boson collision is guaranteed by the conditions

$$\begin{aligned} &\nu\int \varepsilon\hat{\mathrm{St}}^{e-\alpha}\{f, N^{\alpha}\}(\varepsilon, \mathbf{n}; t, \mathbf{r})\frac{d\mathbf{n}d\varepsilon}{\Omega_d} \\ &+ \int \mathrm{Tr}[\hat{\mathscr{L}}^{\alpha}(\omega)\hat{\mathrm{St}}^{\alpha-e}\{f, N^{\alpha}\}(\omega; t, \mathbf{r})]\frac{d\omega}{2\pi} \\ &= -i\int \mathrm{Tr}[[\hat{H}^{\alpha}_{e-h}(\omega); \hat{\mathscr{L}}^{\alpha}(\omega)]\hat{N}^{\alpha}(\omega; t, \mathbf{r})]\frac{\omega d\omega}{2\pi}; \end{aligned} \tag{3.12d}$$

for $\alpha = g, \rho, \sigma$, where the trace is defined as

$$\begin{aligned} \mathrm{Tr}\hat{A}\hat{B} &= \int\frac{d\mathbf{n}_1 d\mathbf{n}_2}{\Omega_d^2} \\ &\times \int\frac{d^d q}{(2\pi)^d}A(\mathbf{q}; \mathbf{n}_1, \mathbf{n}_2)B(\mathbf{q}; \mathbf{n}_2, \mathbf{n}_1). \end{aligned} \tag{3.13}$$

The existence of conservation laws (3.12) immediately enables us to establish the expressions for the conserved currents in the spirit of Section 2.1. By integrating both sides of Eq. (3.4) over ε and $\mathbf{n}$, we find the relations

$$\begin{aligned} \frac{\partial \rho}{\partial t} + \nabla\cdot\mathbf{j} &= 0, \\ \rho(t, \mathbf{r}) &= e\nu\int f(\varepsilon, \mathbf{n}; t, \mathbf{r})\frac{d\varepsilon d\mathbf{n}}{\Omega_d}, \\ \mathbf{j}(t, \mathbf{r}) &= e\nu v_{\mathrm{F}}\int \mathbf{n} f(\varepsilon, \mathbf{n}; t, \mathbf{r})\frac{d\varepsilon d\mathbf{n}}{\Omega_d}, \end{aligned} \tag{3.14}$$

which express the conservation of charge in terms of the usual charge density and electric current density, cf. Eqs. (2.5)–(2.7).

Turning to the energy conservation, we multiply Eq. (3.4) by ε and then integrate over $\mathbf{n}$ and ε. Similarly, we multiply Eq. (3.6) by $\hat{\mathscr{L}}^{\alpha}$, take the trace (3.13), and integrate over ω. Adding the results together, we find

$$\begin{aligned} \frac{\partial u_{\mathrm{tot}}}{\partial t} + \nabla\cdot\mathbf{j}^{\varepsilon}_{\mathrm{tot}} &= \mathbf{j}\cdot\mathbf{E}, \\ u_{\mathrm{tot}} &= u_e + u_{\rho} + 3u_{\sigma} - 4u_g, \\ \mathbf{j}^{\varepsilon}_{\mathrm{tot}} &= \mathbf{j}^{\varepsilon}_e + \mathbf{j}^{\varepsilon}_{\rho} + 3\mathbf{j}^{\varepsilon}_{\sigma} - 4\mathbf{j}^{\varepsilon}_g. \end{aligned} \tag{3.15a}$$

The electronic contributions to the energy density and current density are given by

$$\begin{aligned} u_e(t, \mathbf{r}) &= \nu\int\frac{d\varepsilon d\mathbf{n}}{\Omega_d}\varepsilon f(\varepsilon, \mathbf{n}; t, \mathbf{r}), \\ \mathbf{j}^{\varepsilon}_e(t, \mathbf{r}) &= \nu v_F\int\frac{d\varepsilon d\mathbf{n}}{\Omega_d}\varepsilon\mathbf{n} f(\varepsilon, \mathbf{n}; t, \mathbf{r}). \end{aligned} \tag{3.15b}$$

The contributions of the bosonic neutral excitations are

$$u_\alpha(t, \mathbf{r}) = \int \mathrm{Tr}\left\{\frac{1}{1+\hat{F}^\alpha}\hat{\mathcal{L}}^\alpha(\omega)\hat{N}^\alpha(\omega; t, \mathbf{r})\right\}\frac{\omega d\omega}{2\pi}, \quad (3.15c)$$

$$j^\varepsilon_\alpha(t, \mathbf{r}) = \int \mathrm{Tr}\{\hat{s}^\alpha(\omega)\hat{\mathcal{L}}^\alpha(\omega)\hat{N}^\alpha(\omega; t, \mathbf{r})\}\frac{\omega d\omega}{2\pi}$$

for $\alpha = g, \rho, \sigma$.

Equations (3.14)–(3.15) constitute our main results: the conserved currents are defined in terms of the distribution functions of the quasiparticles that describe low-energy excitations of the interacting electron gas for interaction in the particle–hole channel. In contrast with previous calculations [20–23], we explicitly show the validity of the continuity equation for energy transport; no such proof has been presented before in the quantum kinetic equation approach.[5] Moreover, we believe that the form of the energy current in those references is not correct, because it is not gauge invariant, see Appendix B for more details. As an additional benefit, our approach enables us to clearly identify the contributions of the collective modes and the scattering processes involved (this last task is accomplished by analyzing the explicit form of the collision integrals, which is also needed to calculate the transport coefficients). The derivation of the collision integrals can be found in Section 5; in the next section, we summarize the results and give them a physical interpretation.

3.2. The Collision Integrals

In this section, we give the explicit form of various collision integrals and verify that they satisfy conservation laws (3.12). To shorten the formulas, we introduce the combinations of the distribution functions

$$\Upsilon^\alpha_{ij;\,kl}(\varepsilon, \omega, \mathbf{q}; t, \mathbf{r}) \equiv N^\alpha(\omega, \mathbf{q}; \mathbf{n}_i, \mathbf{n}_j; t, \mathbf{r}) \times \{f(\varepsilon, \mathbf{n}_k; t, \mathbf{r}) - f(\varepsilon-\omega, \mathbf{n}_k; t, \mathbf{r})\} \quad (3.16a)$$
$$+ \Omega_d\delta(\widehat{\mathbf{n}_i\mathbf{n}_j})\{f(\varepsilon, \mathbf{n}_l; t, \mathbf{r})[1 - f(\varepsilon-\omega, \mathbf{n}_k; t, \mathbf{r})]\}$$

and

$$\Psi_{ij;\,kl}(\varepsilon, \varepsilon_1; \omega) \equiv f(\varepsilon-\omega, \mathbf{n}_i)[1 - f(\varepsilon, \mathbf{n}_j)] \times f(\varepsilon_1, \mathbf{n}_k)[1 - f(\varepsilon_1-\omega, \mathbf{n}_l)] \quad (3.16b)$$
$$- f(\varepsilon, \mathbf{n}_i)[1 - f(\varepsilon-\omega, \mathbf{n}_j)]\,f(\varepsilon_1-\omega, \mathbf{n}_k)[1 - f(\varepsilon_1, \mathbf{n}_l)].$$

It follows from Eq. (3.2) that $\Upsilon^\alpha_{ij;\,kl}$ has the property

$$\int d\varepsilon \Upsilon^\alpha_{ij;\,kl}(\varepsilon, \omega, \mathbf{q}) = \int d\varepsilon \Upsilon^\alpha_{ji;\,lk}(\varepsilon, -\omega, -\mathbf{q}). \quad (3.16c)$$

[5] The current operator used in [21] does not satisfy the continuity equation for the long-range interaction potential.

It is easy to check that $\Upsilon = \Psi = 0$ in the thermal equilibrium (Eq. (3.1)). The combination in (3.16b) enters the collision integral in the symmetric form

$$\Psi^s_{ij} \equiv \frac{1}{4}[\Psi_{ij;\,ij} + \Psi_{ji;\,ij} + \Psi_{ij;\,ji} + \Psi_{ji;\,ji}]. \quad (3.16d)$$

It is worth noting that the term's involving four distribution functions f are in fact cancelled from Ψ^s; besides, it has the properties

$$\int \Psi^s_{ij}(\varepsilon, \varepsilon_1; \omega)d\varepsilon d\varepsilon_1 = 0,$$
$$\int a(\omega)\int \varepsilon\Psi^\varepsilon_{ij}(\varepsilon, \varepsilon_1; \omega)d\varepsilon d\varepsilon_1 d\omega = 0 \quad (3.16e)$$

for any even function $a(\omega)$. Finally, we introduce the vertex γ for the impurity scattering

$$\gamma^k_{ij} \equiv \frac{1}{\tau(\theta_{ij})}\Omega_d[\delta(\widehat{\mathbf{n}_j\mathbf{n}_k}) - \delta(\widehat{\mathbf{n}_i\mathbf{n}_k})] \quad (3.17)$$

and the short notation

$$\mathcal{L}^\alpha_{ij} \equiv \mathcal{L}^\alpha(\omega, \mathbf{q}; \mathbf{n}_i, \mathbf{n}_j).$$

The boson–electron collision integral is then explicitly expressed as

$$\hat{\mathrm{St}}^{\alpha-e}(\omega, \mathbf{q}; \mathbf{n}_1, \mathbf{n}_2; t, \mathbf{r}) = -\int d\varepsilon \int \frac{d\mathbf{n}_3 d\mathbf{n}_4}{\Omega_d^2} \quad (3.18)$$
$$\times \{\gamma^3_{14}\Upsilon^\alpha_{32;\,41}(\varepsilon, \omega, \mathbf{q}; t, \mathbf{r}) + \gamma^2_{34}\Upsilon^\alpha_{13;\,14}(\varepsilon, \omega, \mathbf{q}; t, \mathbf{r})\},$$

for $\alpha = g, \rho, \sigma$. The formula for the electron–boson collision integral can be conveniently decomposed into local (l) and nonlocal (n) (in the sense of Section 2) parts

$$\hat{\mathrm{St}}^{e-\alpha} = \hat{\mathrm{St}}^{e-\alpha}_l + \hat{\mathrm{St}}^{e-\alpha}_n. \quad (3.19a)$$

The local part of the collision integral is

$$\hat{\mathrm{St}}^{e-\alpha}_l(\varepsilon, \mathbf{n}_1) = \frac{1}{\nu}\int\frac{d\omega}{2\pi}\frac{1}{\omega}\int\frac{d^dq}{(2\pi)^d}\int\frac{d\mathbf{n}_2 d\mathbf{n}_3 d\mathbf{n}_4}{\Omega_d^3} \quad (3.19b)$$
$$\times \{\gamma^3_{12}[\mathcal{L}^\alpha_{34}\Upsilon^\alpha_{41;\,21}(\varepsilon, \omega, \mathbf{q}) + \Upsilon^\alpha_{34;\,21}(\varepsilon, \omega, \mathbf{q})\overline{\mathcal{L}}^\alpha_{41}]$$
$$+ \gamma^3_{21}[\mathcal{L}^\alpha_{34}\Upsilon^\alpha_{42;\,21}(\varepsilon, \omega, \mathbf{q}) + \Upsilon^\alpha_{34;\,21}(\varepsilon, \omega, \mathbf{q})\overline{\mathcal{L}}^\alpha_{42}]\},$$

where the bar denotes Hermitian conjugation,

$$\overline{\mathcal{L}}^\alpha(\omega, \mathbf{q}; \mathbf{n}_i, \mathbf{n}_j) = \mathcal{L}^\alpha(-\omega, -\mathbf{q}; \mathbf{n}_j, \mathbf{n}_i). \quad (3.19c)$$

Using Eq. (3.11) and definitions (3.13) and (3.17), we can verify that the pair of equations (3.18), (3.19b) satisfy the energy conservation law Eq. (3.12d) on its own.

The function in Eq. (3.19b) also satisfies the particle number conservation law (3.12a). To verify this, we

change the variables as $(\omega, \mathbf{q}) \longrightarrow (-\omega, -\mathbf{q})$ in the terms containing $\overline{\mathcal{L}}^{\alpha}$ and then use Eqs. (3.16b) and (3.19c) to rewrite the integral in Eq. (3.19b) in terms of $\mathcal{L}^{\alpha}$ only:

$$\int \hat{\mathrm{St}}_{l}^{e-\alpha}(\varepsilon, \mathbf{n}_1) d\varepsilon d\mathbf{n}_1 = \ldots \int \frac{d\mathbf{n}_1 \ldots d\mathbf{n}_4}{\Omega_d^4}$$

$$\times \{\gamma_{12}^3 [\mathcal{L}_{34}^{\alpha} \Upsilon_{41;\,21}^{\alpha}(\varepsilon, \omega, \mathbf{q}) - \mathcal{L}_{14}^{\alpha} \Upsilon_{43;\,12}^{\alpha}(\varepsilon, \omega, \mathbf{q})]$$

$$+ \gamma_{21}^3 [\mathcal{L}_{34}^{\alpha} \Upsilon_{42;\,21}^{\alpha}(\varepsilon, \omega, \mathbf{q}) - \mathcal{L}_{24}^{\alpha} \Upsilon_{43;\,12}^{\alpha}(\varepsilon, \omega, \mathbf{q})]\}.$$

Performing the $\mathbf{n}_3$-integration using the delta functions in Eq. (3.17), we obtain a result that is antisymmetric under the $\mathbf{n}_1 \rightleftharpoons \mathbf{n}_2$ permutation. Hence, the above expression vanishes after the $\mathbf{n}_{1,\,2}$-integrations.

The physical meaning of collision integrals (3.18) and (3.19b) is as follows. In the absence of disorder, the electron–hole pair propagates for an infinitely long time. Due to the impurity potential, the decay of the pair into an electron and a hole moving in different directions as shown in Fig. 2 is allowed. Equations (3.18) and (3.19b) are the probabilities for such a decay. (See also Section 5 after Eq. (5.67) for further discussion.)

By construction, the nonlocal contribution to the collision integral

$$\hat{\mathrm{St}}_{n}^{e-\alpha}(\varepsilon, \mathbf{n}_1) = \frac{2}{\nu} \int \frac{d\omega}{2\pi} \frac{1}{\omega^2} \int \frac{d^d q}{(2\pi)^d} \int \frac{d\mathbf{n}_2 \ldots d\mathbf{n}_7}{\Omega_d^6}$$

$$\times \gamma_{13}^2 \gamma_{46}^5 [\mathcal{L}_{14}^{g} - \mathcal{L}_{34}^{g}] \sinh^2\left(\frac{\omega \partial}{2\partial \varepsilon}\right) [f(\varepsilon, \mathbf{n}_6) - f(\varepsilon, \mathbf{n}_4)] \quad (3.19d)$$

$$\times \int d\varepsilon_1 \{[\Upsilon_{57;\,64}^{\alpha}(\varepsilon_1, \omega, \mathbf{q}) + \Upsilon_{57;\,46}^{\alpha}(\varepsilon_1, \omega, \mathbf{q})] \overline{\mathcal{L}}_{72}^{\alpha}$$

$$+ \mathcal{L}_{57}^{\alpha} [\Upsilon_{72;\,13}^{\alpha}(\varepsilon_1, \omega, \mathbf{q}) + \Upsilon_{72;\,31}^{\alpha}(\varepsilon_1, \omega, \mathbf{q})]\}$$

satisfies its own conservation law

$$\int \varepsilon^m \hat{\mathrm{St}}_{n}^{e-\alpha}(\varepsilon, \mathbf{n}_1) d\varepsilon = 0, \quad m = 0, 1, \qquad (3.19e)$$

i.e., preserves the energy and the number of electrons moving along a given momentum direction $\mathbf{n}$. Moreover, one can see that the collision integral (3.19d) does not contribute to the linear response at all, because $\Upsilon = 0$ and f is independent of the angle in thermodynamic equilibrium. The nonlocality of this collision integral indicates that the task formulated in Section 2 has not been fully accomplished. Technically, this nonlocality can be decoupled by introducing a density matrix that is nondiagonal in the boson–ghost space. We choose not to pursue this line because the term in Eq. (3.19d) does not contribute to any observable quantity of interest and does not affect any conservation laws.

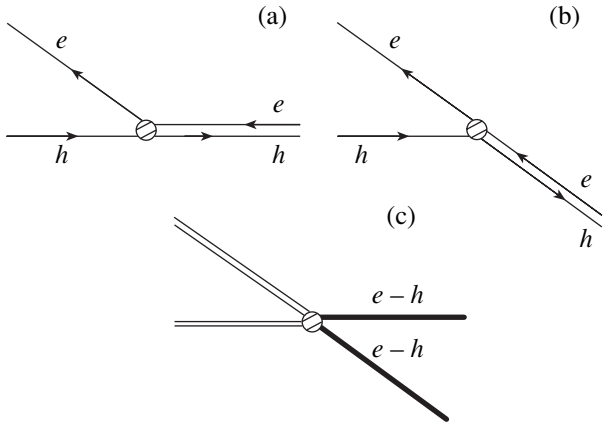

Fig. 2. The scattering amplitudes leading to the creation of the same electron and hole out of different electron–hole pairs (double lines) (a, b) and their interference contribution (c). The impurity is denoted by filled circles.

The electron–electron collision integral can be split into elastic, nonlocal, and local parts:

$$\hat{\mathrm{St}}^{e-e}(\varepsilon, \mathbf{n}_1) = \hat{\mathrm{St}}_{el}^{e-e}(\varepsilon, \mathbf{n}_1) + \hat{\mathrm{St}}_{n}^{e-e}(\varepsilon, \mathbf{n}_1) + \hat{\mathrm{St}}_{l}^{e-e}(\varepsilon, \mathbf{n}_1). \qquad (3.20a)$$

The elastic term describes the scattering of the electron on the static self-consistent potential created by all the other electrons,

$$\hat{\mathrm{St}}_{el}^{e-e} = \frac{2}{\nu} \mathrm{Re} \int \frac{d\omega}{2\pi} \frac{1}{\omega} \int \frac{d^d q}{(2\pi)^d} \int \frac{d\mathbf{n}_2 \ldots d\mathbf{n}_6}{\Omega_d^5} \gamma_{13}^2 \gamma_{46}^5 \quad (3.20b)$$

$$\times [\mathcal{L}^{\rho} + 3\mathcal{L}^{\sigma} - 4\mathcal{L}^{g}]_{52} [f(\varepsilon - \omega, \mathbf{n}_6) - f(\varepsilon - \omega, \mathbf{n}_4)]$$

$$\times [\mathcal{L}_{14}^{g} f(\varepsilon, \mathbf{n}_3) + \mathcal{L}_{34}^{g} f(\varepsilon, \mathbf{n}_1)].$$

Its physical origin is discussed in detail in [17]. Being elastic, it preserves the number of particles, for each energy shell,

$$\int \hat{\mathrm{St}}_{el}^{e-e}(\varepsilon, \mathbf{n}_1) d\mathbf{n}_1 = 0, \qquad (3.20c)$$

as can be seen from the property $\gamma_{13}^2 = -\gamma_{31}^2$ of vertex (3.17).

The nonlocal term

$$\hat{\mathrm{St}}_{n}^{e-e} = -\frac{4}{\nu} \mathrm{Re} \int \frac{d\omega}{2\pi} \frac{1}{\omega^2}$$

$$\times \int \frac{d^d q}{(2\pi)^q} \int \frac{d\mathbf{n}_2 \ldots d\mathbf{n}_6}{\Omega_d^5} \gamma_{13}^2 \gamma_{46}^5 \int d\varepsilon_1 \{[\mathcal{L}_{14}^{g} - \mathcal{L}_{34}^{g}] \quad (3.20d)$$

$$\times [\overline{\mathcal{L}}^{\rho} + 3\overline{\mathcal{L}}^{\sigma} - 4\overline{\mathcal{L}}^{g}]_{52} \Psi_{46}^{s}(\varepsilon, \varepsilon_1; \omega)\}$$

describes the inelastic electron–electron collisions during which the bosons and ghosts act as virtual states.

(The function Ψ^s was introduced in Eq. (3.16d).) The real part being an even function, we can use Eq. (3.16e) to verify that Eq. (3.20d) satisfies the conservation law

$$\int \varepsilon^m \hat{\mathrm{St}}^{e-e}_{n,l}(\varepsilon, \mathbf{n}_1) d\varepsilon = 0, \quad m = 0, 1. \tag{3.20e}$$

As indicated, the same law is satisfied by the local (and elastic) term:

$$\begin{aligned}\hat{\mathrm{St}}^{e-e}_l &= \frac{2}{\nu}\int\frac{d\omega}{2\pi}\frac{1}{\omega^2}\int\frac{d^d q}{(2\pi)^d}\int\frac{d\mathbf{n}_2 \ldots d\mathbf{n}_6}{\Omega_d^5} \\ &\times \int d\varepsilon_1 \gamma_{13}^2 \gamma_{46}^5 [\mathscr{L}^g_{14} + \mathscr{L}^g_{34}][\mathscr{L}^\rho + 3\mathscr{L}^\sigma - 4\mathscr{L}^g]_{52} \\ &\times \sinh^2\left(\frac{\omega\partial}{2\partial\varepsilon}\right)[f(\varepsilon, \mathbf{n}_6) - f(\varepsilon, \mathbf{n}_4)] \\ &\times [f(\varepsilon_1, \mathbf{n}_1)[1 - f(\varepsilon_1 - \omega, \mathbf{n}_3)] + (\mathbf{n}_1 \rightleftharpoons \mathbf{n}_3)].\end{aligned} \tag{3.20f}$$

Therefore, Eq. (3.20e) enables us to conclude that both collision integrals (3.20d) and (3.20f) do not affect the transport coefficients (in the case where they can be considered perturbations in comparison to the bare impurity collision integral).

We note that although it might not be evident, the present form of the kinetic equation permits the proper identification of the inelastic kernel that determines the phase relaxation time; further details can be found in Appendix G.

4. SUMMARY OF THE RESULTS FOR THE THERMAL TRANSPORT AND SPECIFIC HEAT

In this section, we present our final answers for the interaction corrections to the thermal conductivity and the specific heat. They are obtained by solving the kinetic equations and then substituting the solutions in definitions (3.15) of the energy and energy current densities. The explicit calculations are performed in Section 6. We consider short-range impurities for which the scattering time τ is independent of the scattering angle, $\tau(\theta) = \tau$. We report our results for quasi-one-dimensional and three-dimensional systems in the diffusive limit $T\tau \ll \hbar$; for two-dimensional systems, we do not put such a restriction on the temperature range. However, common to all dimensionalities is the zeroth harmonic approximation for the Fermi-liquid constants (see Eq. (6.14)).

4.1. Thermal Conductivity

In the absence of magnetic field, the thermal conductivity tensor is diagonal, $\kappa_{\mu\nu} = \kappa\delta_{\mu\nu}$, and we write the expression for the diagonal components simply as

$$\kappa = \kappa_{\mathrm{WF}} + \Delta\kappa. \tag{4.1}$$

The first term is given by the Wiedemann–Franz law $\kappa_{\mathrm{WF}} = L\sigma T$ with the inclusion of the interaction corrections to the conductivity and the Lorentz number given in Eq. (1.1). The second term causes a violation of the Wiedemann–Franz law. In the diffusive limit and for low dimensionality, the main contribution to $\Delta\kappa$ is due to the long-range nature of the bosonic energy transport, which originates from the long-range part of the interaction in the singlet channel. In the quasiballistic case, large contribution also is made by the inelastic scattering of the electron on the bosons. Smaller corrections arise due to the triplet channel bosonic transport and to the energy dependence of the elastic scattering.

For quasi-one-dimensional and three-dimensional systems in the diffusive limit, we write

$$\Delta\kappa = \delta\kappa^\rho + 3\delta\kappa^\sigma + \delta\kappa_{el},$$

where the bosonic corrections include the ghost contributions

$$\delta\kappa^\alpha = \kappa^\alpha - \kappa^g, \quad \alpha = \rho, \sigma$$

(see Eq. (6.12) for the definition of κ^α) and we neglect the inelastic contributions $\delta\kappa_{\mathrm{in}}$, which are smaller by the parameter $T\tau/\hbar$.

For quasi-one-dimensional systems, the explicit expressions are

$$\begin{aligned}\delta\kappa_{\mathrm{el}} &= \frac{1}{8\sqrt{2\pi}}\zeta\left(\frac{3}{2}\right)\sqrt{\frac{DT}{\hbar}} \\ &\times\left\{-1 + 3\left[1 - \frac{2}{F_0^\sigma}(1 + F_0^\sigma - \sqrt{1 + F_0^\sigma})\right]\right\},\end{aligned} \tag{4.2a}$$

$$\delta\kappa^\rho = \frac{3}{8\sqrt{2\pi}}\zeta\left(\frac{3}{2}\right)\sqrt{\frac{DT}{\hbar}}ak\ln^2\left(\frac{\hbar Dk^2}{T}\right), \tag{4.2b}$$

$$\delta\kappa^\sigma = \frac{3}{8\sqrt{2\pi}}\zeta\left(\frac{3}{2}\right)\sqrt{\frac{DT}{\hbar}}[\sqrt{1 + F_0^\sigma} - 1], \tag{4.2c}$$

where a is a length on the order of the wire width, $k = \sqrt{4\pi e^2 \nu}$ is the inverse screening length in the bulk, and $D = \tau v_{\mathrm{F}}^2/d$ is the diffusion constant (in d dimensions).

For three-dimensional systems, the results are

$$\begin{aligned}\delta\kappa_{\mathrm{el}} &= \frac{5}{48\sqrt{2\pi^3}}\zeta\left(\frac{5}{2}\right)\sqrt{\frac{T^3}{\hbar^3 D}} \\ &\times\left\{1 + 3\left[1 - \frac{2}{F_0^\sigma}\left(1 - \frac{1}{\sqrt{1 + F_0^\sigma}}\right)\right]\right\},\end{aligned} \tag{4.3a}$$

$$\delta\kappa^\rho = \frac{15}{32\sqrt{2\pi^3}}\zeta\left(\frac{5}{2}\right)\sqrt{\frac{T^3}{\hbar^3 D}}, \tag{4.3b}$$

$$\delta\kappa^{\sigma} = \frac{15}{32\sqrt{2\pi^3}}\zeta\left(\frac{5}{2}\right)\sqrt{\frac{T^3}{\hbar^3 D}}\left[1 - \frac{1}{\sqrt{1+F_0^{\sigma}}}\right]. \quad (4.3c)$$

In these expressions, $\zeta(x)$ is the Riemann zeta function, with $\zeta(3/2) \approx 2.612$ and $\zeta(5/2) \approx 1.341$.

For two-dimensional systems, we separate the corrections due to the singlet and the triplet channel interactions:

$$\Delta\kappa = \Delta\kappa_s + 3\Delta\kappa_t. \quad (4.4a)$$

With logarithmic accuracy, the singlet channel contribution is

$$\Delta\kappa_s = \frac{T}{6\hbar}g_1\left(2\pi\frac{T\tau}{\hbar}\right)\ln\left(\frac{\hbar v_F k}{T}\right) - \frac{T}{24\hbar}g_2\left(\pi\frac{T\tau}{\hbar}\right)\ln\left(1+\frac{\hbar^2}{(T\tau)^2}\right) - \frac{\pi^2}{15}\frac{T}{\hbar}\left(\frac{T\tau}{\hbar}\right)^2\ln\left(\frac{E_F}{T}\right), \quad (4.4b)$$

where $k = 2\pi e^2\nu$ is the two-dimensional inverse screening length. The crossover functions g_1 and g_2 are given in Eq. (6.42). Here we note that $g_1(x)$, $g_2(x) \approx 1$ for $x \ll 1$ and $g_1(x) \approx 3/x$, $g_2(x) \approx 14x^2/15$ for $x \gg 1$. For the triplet channel, we have

$$\Delta\kappa_t = \begin{cases} -\frac{T}{18\hbar}\left[1 - \frac{1}{F_0^{\sigma}}\ln(1+F_0^{\sigma})\right] + \frac{T}{12\hbar}\ln(1+F_0^{\sigma}), \\ T\tau \ll \hbar \\ -\frac{\pi^2}{15}\frac{T}{\hbar}\left(\frac{T\tau}{\hbar}\right)^2\ln\left(\frac{E_F}{T}\right)\left(\frac{F_0^{\sigma}}{1+F_0^{\sigma}}\right)^2, \quad T\tau \gg \hbar. \end{cases} \quad (4.4c)$$

In the diffusive limit $T\tau \ll \hbar$, our results are consistent with those in [22], even though the form of the energy current operator used in this reference is, in our opinion, incorrect, see Appendix B.

4.2. Specific Heat

The specific heat is given by

$$c_V = \frac{\pi^2}{3}\nu T + \delta c_V, \quad (4.5)$$

where the first term is the usual noninteracting electronic contribution and the second term is the bosonic interaction correction.

For quasi-one-dimensional and three-dimensional systems,

$$\delta c_V = (1+3\Lambda_d)\left(\frac{T}{\hbar D}\right)^{d/2} a_d. \quad (4.6)$$

The two terms in the first brackets are respectively the singlet and triplet channel contributions. The singlet channel term is considered in the unitary limit and is therefore independent of any interaction parameter. On the other hand, the Fermi-liquid parameter for the interaction in the triplet channel enters Eq. (4.6) as

$$\Lambda_d = 1 - \frac{1}{(1+F_0^{\sigma})^{d/2}} \quad (4.7)$$

and the numerical factors $a_{1,3}$ are

$$a_3 = \frac{15}{32\pi\sqrt{2\pi}}\zeta\left(\frac{5}{2}\right),$$

$$a_1 = -\frac{3}{8\sqrt{2\pi}}\zeta\left(\frac{3}{2}\right).$$

For two-dimensional systems, with logarithmic accuracy, the result is

$$\delta c_V = -\left(\frac{T}{\hbar D}\right) \times \left[(1+3\Lambda_2)\frac{1}{12}\ln\left(\frac{E_F}{T}\right) + (1+3\Lambda_2^2)\frac{3}{4\pi}\zeta(3)T\tau\right], \quad (4.8)$$

where $\zeta(3) \approx 1.202$. The first term in the right-hand side extends the logarithmic behavior known in the diffusive limit to higher temperatures (the upper cutoff is on the order of the Fermi energy E_F and not $\hbar/\tau$); the second term becomes relevant in the quasiballistic limit and coincides with the correction calculated in [29] for the clean Fermi liquid. In the diffusive limit, our results are the same as those obtained in [12] by an explicit thermodynamic calculation.

5. DERIVATION OF THE KINETIC EQUATION

This section is devoted to the derivation of the local kinetic equation. We first introduce the Eilenberger equation and some basic notation. Next, we perform a (generalized) gauge transformation: this is the crucial step that enables us to obtain the local description. Then we introduce the bosonic degrees of freedom and derive the collision integrals.

5.1. Eilenberger Equation

Our starting point for the derivation of the kinetic equation is the Eilenberger equation (Eq. (5.7)). For disordered metals, the derivation of this equation, starting from the action for the interacting electron gas in the presence of disorder, can be found in [17]. We briefly summarize it here to introduce some notation and a straightforward generalization to the angle-dependent impurity scattering rate and Fermi-liquid parameter (see Eqs. (5.8) and (5.15)).

The interaction with small momentum and energy transfer in the singlet channel (the triplet channel is to be discussed in Section 5.8) is decoupled using the two Hubbard–Stratonovich fields $\phi_{\pm}(t, \mathbf{r}, \mathbf{n})$. For the purpose

of the one-loop approximation that we use, these fields can be considered Gaussian with the propagators

$$\begin{aligned}\langle\langle\phi_+(1)\phi_+(2)\rangle\rangle &= -\frac{i}{2}D^K(1,2),\\ \langle\langle\phi_+(1)\phi_-(2)\rangle\rangle &= -\frac{i}{2}D^R(1,2),\\ \langle\langle\phi_-(1)\phi_+(2)\rangle\rangle &= -\frac{i}{2}D^A(1,2),\\ \langle\langle\phi_-(1)\phi_-(2)\rangle\rangle &= 0,\end{aligned} \tag{5.1}$$

where $\langle\langle\ldots\rangle\rangle$ denotes averaging over the fields $\phi_\pm$. We use the short notation

$$\begin{aligned}(i)\equiv(t_i,\mathbf{r}_i,\mathbf{n}_i),\quad \int di\equiv\int dt_i d\mathbf{r}_i\int\frac{d\mathbf{n}_i}{\Omega_d},\\ (i^*)\equiv(\mathbf{r}_i,\mathbf{n}_i),\quad \int di^*\equiv\int d\mathbf{r}_i\int\frac{d\mathbf{n}_i}{\Omega_d},\end{aligned} \tag{5.2}$$

where $i = 1, 2, \ldots$ and Ω_d is the total solid angle.

We introduce the disorder-averaged Green function of the electron in the field $\phi_\pm$ in its matrix form in the Keldysh space,

$$\hat G(1,2|\phi) = \begin{pmatrix} G^R(1,2|\phi) & G^K(1,2|\phi)\\ G^Z(1,2|\phi) & G^A(1,2|\phi)\end{pmatrix}_K, \tag{5.3}$$

such that its average over the fluctuating field $\phi_\pm$ gives the usual expressions for the physical propagators:

$$\begin{aligned}\langle\langle G^R(1,2)\rangle\rangle &= -i\theta(t_1-t_2)\\ &\times\langle\psi(1)\psi^\dagger(2)+\psi^\dagger(2)\psi(1)\rangle,\\ \langle\langle G^A(1,2)\rangle\rangle &= i\theta(t_2-t_1)\\ &\times\langle\psi(1)\psi^\dagger(2)+\psi^\dagger(2)\psi(1)\rangle,\\ \langle\langle G^K(1,2)\rangle\rangle &= -i\langle\psi(1)\psi^\dagger(2)-\psi^\dagger(2)\psi(1)\rangle,\\ \langle\langle G^Z(1,2)\rangle\rangle &= 0.\end{aligned} \tag{5.4}$$

Here, $\theta(t)$ is the Heaviside step function, $\phi^\dagger$ and ψ are the fermionic creation/annihilation operators in the Heisenberg representation, and quantum mechanical averaging $\langle\ldots\rangle$ is performed with an arbitrary distribution function to be found from the solution of the kinetic equation.

For the disorder-averaged Green function, the semiclassical approximation is obtained by integrating the Wigner transform of $\hat G(1, 2|\phi)$ over the distance from the Fermi surface:

$$\hat G(t_1,t_2,\mathbf{p},\mathbf{R}) = \int d^d r e^{i\mathbf{P}\cdot\mathbf{r}}\hat G(1,2|\phi), \tag{5.5}$$

where

$$\begin{aligned}\mathbf{r} &= \mathbf{r}_1-\mathbf{r}_2,\quad \mathbf{R} = \frac{1}{2}(\mathbf{r}_1+\mathbf{r}_2),\\ \mathbf{P} &= \mathbf{p}-\frac{1}{2}[\mathbf{A}(t_1,\mathbf{R})+\mathbf{A}(t_2,\mathbf{R})],\\ \hat g(t_1,t_2,\mathbf{n},\mathbf{r}) &= \frac{i}{\pi\nu}\int_{-\infty}^{\infty}d\xi\hat G\left(t_1,t_2,\mathbf{n}\left[p_F+\frac{\xi}{v_F}\right],\mathbf{r}\right),\end{aligned} \tag{5.6}$$

where $\mathbf{A}$ is the vector potential of an external electromagnetic field, p_F is the Fermi momentum, v_F is the Fermi velocity and ν is the density of states on the Fermi level. The dynamics of the semiclassical Green function $\hat g$ in the matrix form is governed by the Eilenberger equation [24]

$$\begin{aligned}\left[\tilde\partial_t+\mathbf{v}\cdot\tilde\nabla+\boldsymbol{\omega}_c\cdot\left(\mathbf{n}\times\frac{\partial}{\partial\mathbf{n}}\right)\right]\hat g\\ +i[\hat\phi;\hat g] = \frac{[\hat g\circ\hat{\mathrm{St}}_\tau\hat g]}{2},\end{aligned} \tag{5.7}$$

where $\mathbf{v} = v_F\mathbf{n}$, the action of the "bare" collision integral on any function $a(\mathbf{n})$ is defined as

$$\begin{aligned}[\hat{\mathrm{St}}_\tau a](\mathbf{n}) &= \int\frac{d\mathbf{n}_1}{\Omega_d}\mathrm{St}_\tau(\mathbf{n},\mathbf{n}_1)a(\mathbf{n}_1),\\ \hat{\mathrm{St}}_\tau(\mathbf{n}_1,\mathbf{n}_2) &= \frac{1}{\tau(\theta_{12})}-\delta(\widehat{\mathbf{n}_1\mathbf{n}_2})\int\frac{d\mathbf{n}_2}{\tau(\theta_{12})},\end{aligned} \tag{5.8}$$

and $\theta_{12} = \widehat{\mathbf{n}_1\mathbf{n}_2}$ (for the short-range impurity, $\tau(\theta)$ is independent of θ; however, the formulas derived here are valid for an arbitrary impurity scattering). The time convolution of two matrices $\hat a(t_1, t_2)$ and $\hat b(t_1, t_2)$ is given by

$$\begin{aligned}\hat a\circ\hat b &= \int dt_3\hat a(t_1,t_3)\hat b(t_3,t_2),\\ [\hat a\circ\hat b] &= \hat a\circ\hat b-\hat b\circ\hat a.\end{aligned} \tag{5.9}$$

Defining the commutator between a matrix $\hat c(t, \mathbf{r}, \mathbf{n})$ and $\hat g$ as

$$\begin{aligned}[\hat c;\hat g] &= \hat c(t_1,\mathbf{r},\mathbf{n})\hat g(t_1,t_2,\mathbf{n},\mathbf{r})\\ &-\hat g(t_1,t_2,\mathbf{n},\mathbf{r})\hat c(t_2,\mathbf{r},\mathbf{n}),\end{aligned} \tag{5.10}$$

we express the covariant derivatives in Eq. (5.7) as

$$\tilde\partial_t\hat g = \partial_{t_1}\hat g+\partial_{t_2}\hat g+i[\hat\varphi;\hat g], \tag{5.11a}$$

$$\tilde\nabla\hat g = \nabla\hat g+i[\hat{\mathbf{A}};\hat g] \tag{5.11b}$$

with $\hat{\mathbf{A}} = A\hat{\mathbb{1}}_K$ and $\hat\varphi = \varphi\hat{\mathbb{1}}_K$. Here, $\hat{\mathbb{1}}_K$ denotes the unit matrix in the Keldysh space and φ is the scalar potential

for an external electromagnetic field such that

$$e\mathbf{E} = -\nabla\varphi + \partial_t\mathbf{A}, \quad e\mathbf{B} = -c\nabla\times\mathbf{A}.$$

The vector $\boldsymbol{\omega}_c = e\mathbf{B}/(mc)$ has the magnitude of the cyclotron frequency and the direction of the magnetic field **B**. Finally, $\hat{\phi}$ is the matrix in the Keldysh space:

$$\hat{\phi} = \begin{pmatrix} \phi_+ & \phi_- \\ \phi_- & \phi_+ \end{pmatrix}_K. \tag{5.12}$$

The matrix Green function $\hat{g}$ is subject to the constraints

$$\hat{g}(\mathbf{n}, \mathbf{r}) \circ \hat{g}(\mathbf{n}, \mathbf{r}) = \delta(t_1 - t_2)\hat{1\!\!1}_K, \tag{5.13a}$$

$$\mathrm{Tr}\hat{g}(t, t, \mathbf{n}, \mathbf{r}) = 0. \tag{5.13b}$$

In thermal equilibrium, the relation

$$\begin{aligned} g^K(t_1, t_2) &= [g^R \circ n - n \circ g^A](t_1, t_2), \\ n(t_1, t_2) &= \int\frac{d\varepsilon}{2\pi}\exp(i\varepsilon(t_2 - t_1))n(\varepsilon), \\ n(\varepsilon) &= 1 - 2f_F(\varepsilon) = 2\tanh\frac{\varepsilon}{2T} \end{aligned} \tag{5.14}$$

must hold independently of the form of the spectral functions $g^{R,A}$.

In what follows, we assume that there is no magnetic field, $\mathbf{B} = 0$ and $\boldsymbol{\omega}_c = 0$, but no gauge choice is made: although one could set $\mathbf{A} = 0$ by a gauge transformation, both the scalar and vector external potentials are left arbitrary in order to keep track of the gauge invariance of the equations.

The propagators defined by Eq. (5.1) satisfy the matrix Dyson equation

$$\begin{aligned} \hat{D}(1, 2) &= \hat{D}_0(1, 2) \\ &+ \int d3\int d4\hat{D}_0(1, 3)\hat{\Pi}(3, 4)\hat{D}(4, 2), \\ \hat{D}_0(1, 2) &= -\left[V(\mathbf{r}_{12}) + \frac{F^\rho(\theta_{12})\delta(\mathbf{r}_{12})}{\nu}\right]\delta(t_{12})\hat{1\!\!1}_K, \end{aligned} \tag{5.15}$$

where $V(\mathbf{r})$ is the long-range part of the interaction (for the Coulomb interaction, $V(\mathbf{r}) = e^2/|\mathbf{r}|$, $\theta_{12} = \widehat{\mathbf{n}_1\mathbf{n}_2}$, $\mathbf{r}_{12} = \mathbf{r}_1 - \mathbf{r}_2$, and $t_{12} = t_1 - t_2$). The matrix propagator is denoted by $\hat{D}$ and $\hat{\Pi}$ is the matrix polarization operator. They have a structure similar to the Green function one:

$$\hat{D} = \begin{pmatrix} D^R & D^K \\ 0 & D^A \end{pmatrix}_K, \quad \hat{\Pi} = \begin{pmatrix} \Pi^R & \Pi^K \\ 0 & \Pi^A \end{pmatrix}_K. \tag{5.16}$$

The polarization operators are given by variational derivatives of the solutions of Eilenberger Eq. (5.7),

$$\begin{aligned} \Pi^R(1, 2) &= \Pi^A(2, 1) \\ &= \nu\left[\delta_{12} + \frac{\pi}{2}\frac{\delta g^K(t_1, t_1, \mathbf{n}_1, \mathbf{r}_1)}{\delta\phi_+(t_1, \mathbf{r}_2, \mathbf{n}_2)}\right], \end{aligned} \tag{5.17a}$$

$$\Pi^K(1, 2) = \frac{\pi\nu}{2}\left[\frac{\delta(g^K + g^Z)(t_1, t_1, \mathbf{n}_1, \mathbf{r}_1)}{\delta\phi_-(t_2, \mathbf{r}_2, \mathbf{n}_2)}\right], \tag{5.17b}$$

where

$$\delta_{12} = \Omega_d\delta(\widehat{\mathbf{n}_1\mathbf{n}_2})\delta(\mathbf{r}_1 - \mathbf{r}_2)\delta(t_1 - t_2), \tag{5.18}$$

with Ω_d being the total solid angle.

5.2. The Gauge Transformation

With Eilenberger Eq. (5.7) at hand, one could proceed as in [17] in order to derive an equation for the distribution function. But the resulting inelastic part of the collision integral, expressed in terms of the electron distribution function only, is nonlocal and the evaluation, e.g., of the thermal conductivity would require the time and spatial gradient expansion of this term in the spirit of Eq. (2.15). As we already discussed, such a route makes the energy conservation in the kinetic equation obscure. Here, we follow a different approach, inspired by the following considerations [25]: if the fluctuating fields were uniform, they would be eliminated from Eq. (5.7) by the gauge transformation

$$\hat{g} \longrightarrow \exp\left\{-i\int^{t_1}\hat{\phi}(t)dt\right\}\hat{g}\exp\left\{i\int^{t_2}\hat{\phi}(t)dt\right\}. \tag{5.19}$$

In other words, the position-independent fluctuations of the ϕ fields define a time-dependent position of the energy levels, but the occupation numbers for such levels do not change. Therefore, such fluctuations affect neither the electric transport nor the electron contribution to the thermal transport in the system. Moreover, if the path of the electron were a straight line, all smooth fluctuating fields would still be eliminated in the eikonal approximation and, once again, they should not affect the electronic contribution to the transport. To eliminate such spurious contributions, we use the gauge transformation described below.

We introduce a new matrix field $\hat{K}(t, \mathbf{n}, \mathbf{r})$,

$$\hat{K} = \begin{pmatrix} K_+ & K_- \\ K_- & K_+ \end{pmatrix}, \tag{5.20}$$

which is a functional of the field $\hat{\phi}$ and is used to perform the "generalized" gauge transformation

$$\hat{g} \longrightarrow e^{-i\hat{K}(t_1, \mathbf{n}, \mathbf{r})}\hat{g}e^{i\hat{K}(t_2, \mathbf{n}, \mathbf{r})}. \tag{5.21}$$

This transformation is unitary and therefore preserves constraints Eq. (5.13). As we see in what follows, it leads to the local kinetic equations. Applying the transformation to Eilenberger Eq. (5.7), we obtain

$$[\tilde{\partial}_t + \mathbf{v}\cdot\tilde{\nabla}]\hat{g} - i[(\partial_t + \mathbf{v}\cdot\nabla)\hat{K} - \hat{\phi}, \hat{g}] = \frac{1}{2}[\hat{g} \circ \hat{\mathrm{St}}^{\phi}_{\tau}\hat{g}], \tag{5.22}$$

where

$$[\hat{\mathrm{St}}^{\phi}_{\tau}\hat{g}](t_1, t_2, \mathbf{n}) \equiv \int \frac{d\mathbf{n}_1}{\Omega_d}\hat{\mathrm{St}}_{\tau}(\mathbf{n}, \mathbf{n}_1) \times e^{i\hat{K}(t_1, \mathbf{n})}e^{-i\hat{K}(t_1, \mathbf{n}_1)}\hat{g}(t_1, t_2, \mathbf{n}_1)e^{i\hat{K}(t_2, \mathbf{n}_1)}e^{-i\hat{K}(t_2, \mathbf{n})}.$$

The "bare" impurity collision integral and the derivatives are defined respectively in Eq. (5.8) and (5.11). We suppress the argument **r**, which is the same in all functions.

We seek a perturbative solution of the Eilenberger equation in the form of (5.22) in the one-loop approximation; for this, it suffices to retain only the terms at most quadratic in the K fields in the collision integral. In the lowest order, $\hat{g}$ has the form

$$\hat{g} = \begin{pmatrix} \delta(t_1 - t_2) & g^K \\ 0 & -\delta(t_1 - t_2) \end{pmatrix}. \tag{5.23}$$

We require that this form be preserved even in the first and second order in K; i.e., the corrections to the spectrum (described by $g^{R,A}$) be indeed eliminated by the gauge transformation.

In the linear order, the retarded, advanced, and "Z" components of Eq. (5.22) vanish if K_- satisfies the equation

$$(\partial_t + \mathbf{v}\cdot\nabla)K_- + \hat{\mathrm{St}}_{\tau}K_- = \phi_-. \tag{5.24}$$

The solution of integro-differential Eq. (5.24) can be written in terms of the diffusion propagator $\mathscr{L}^g(t_1, t_2; \mathbf{n}_1, \mathbf{n}_2; \mathbf{r}_1, \mathbf{r}_2)$, the retarded solution of the classical kinetic equation

$$(\partial_{t_1} + \mathbf{v}_1\cdot\nabla_{\mathbf{r}_1} - \hat{\mathrm{St}}_{\tau})\mathscr{L}^g = \delta_{12}, \tag{5.25}$$

where δ_{12} is defined in Eq. (5.18). Using Eq. (5.25), we find

$$K_-(1) = -\int d2\,\overline{\mathscr{L}}^g(1, 2)\phi_-(2), \qquad \overline{\mathscr{L}}^g(1, 2) = \mathscr{L}^g(2, 1), \tag{5.26}$$

where we use the short notation in (5.2). In the operator notation, Eq. (5.26) can be rewritten as

$$K_- = -\hat{\overline{\mathscr{L}}}^g\phi_-. \tag{5.26'}$$

To simplify further manipulations, we introduce the following function of three angular variables:

$$\gamma\begin{pmatrix} \mathbf{n}_k \\ \mathbf{n}_i;\ \mathbf{n}_j \end{pmatrix} \equiv \frac{1}{\tau(\theta_{ij})} \times \Omega_d[\delta(\widehat{\mathbf{n}_j\mathbf{n}_k}) - \delta(\widehat{\mathbf{n}_i\mathbf{n}_k})] \equiv \gamma^k_{ij}. \tag{5.27}$$

This function is related to the impurity collision integral (5.8) by

$$\int\frac{d\mathbf{n}_2}{\Omega_d}\gamma\begin{pmatrix} \mathbf{n}_3 \\ \mathbf{n}_1;\ \mathbf{n}_2 \end{pmatrix} = [\hat{\mathrm{St}}_{\tau}](\mathbf{n}_1, \mathbf{n}_3),$$

$$\int\frac{d\mathbf{n}_1}{\Omega_d}\gamma\begin{pmatrix} \mathbf{n}_3 \\ \mathbf{n}_1;\ \mathbf{n}_2 \end{pmatrix} = -[\hat{\mathrm{St}}_{\tau}](\mathbf{n}_2, \mathbf{n}_3), \tag{5.28}$$

$$\int\frac{d\mathbf{n}_3}{\Omega_d}\gamma\begin{pmatrix} \mathbf{n}_3 \\ \mathbf{n}_1;\ \mathbf{n}_2 \end{pmatrix} = 0.$$

With δg^K denoting the first-order correction to g^K, the Keldysh component of the Eilenberger equations in the linear order is

$$i(\tilde{\partial}_t + \mathbf{v}\cdot\tilde{\nabla} - \hat{\mathrm{St}}_{\tau})\delta g^K + [(\partial_t + \mathbf{v}\cdot\nabla - \hat{\mathrm{St}}_{\tau})K_+ - \phi_+;\ g^K] = \hat{Q}K_- + \int\frac{d\mathbf{n}_2 d\mathbf{n}_3}{\Omega_d^2}\gamma\begin{pmatrix} \mathbf{n}_3 \\ \mathbf{n};\ \mathbf{n}_2 \end{pmatrix}[K_+(\mathbf{n}_3);\ g^K(\mathbf{n}_2) - g^K(\mathbf{n})] \tag{5.29}$$

with (cf. Eq. (5.11))

$$\tilde{\partial}_t\delta g^K = (\partial_{t_1} + \partial_{t_2})\delta g^K + i[\varphi;\ \delta g^K],$$

$$\tilde{\nabla}\delta g^K = \nabla\delta g^K + i[\mathbf{A};\ \delta g^K].$$

Here, the operator $\hat{Q}$ is local in space,

$$\hat{Q}K_- = \int dt_3\int\frac{d\mathbf{n}_1}{\Omega_d}Q(t_1, t_2, \mathbf{n};\ t_3, \mathbf{n}_1;\ \mathbf{r})K_-(t_3, \mathbf{n}_1, \mathbf{r})$$

with the kernel

$$Q(t_1, t_2, \mathbf{n}_1;\ t_3, \mathbf{n}_2) = \frac{1}{2}\int\frac{d\mathbf{n}_3}{\Omega_d}\gamma\begin{pmatrix} \mathbf{n}_2 \\ \mathbf{n}_1;\ \mathbf{n}_3 \end{pmatrix} \times [Pg^K(t_1, t_3, \mathbf{n}_1)g^K(t_3, t_2, \mathbf{n}_3) + (\mathbf{n}_1 \rightleftarrows \mathbf{n}_3)]. \tag{5.30}$$

We suppress the spectator argument **r**, which is the same in each term of the equation; the last term means that terms similar to the ones shown, but with the angular arguments of the Green function switched, must be

added. The principal value sign P in Eq. (5.30) means that the part of the product of the Green functions that are divergent as $t_1 \longrightarrow t_2$,

$$g^K(t_1, t_2, \mathbf{n}, \mathbf{r})\big|_{t_1 \to t_2} = -\frac{2i}{\pi(t_1 - t_2)} + \text{regular}, \quad (5.31)$$

must be eliminated,

$$Pg^K(t_1, t_3)g^K(t_3, t_2)$$
$$\equiv g^K(t_1, t_3)g^K(t_3, t_2) - 4\delta(t_1 - t_3)\delta(t_3 - t_2),$$

or, equivalently,

$$Pg^K(t_1, t_3)g^K(t_3, t_2)$$
$$\equiv \frac{1}{2}\sum_{\sigma = \pm 1} g^K(t_1, t_3 + \sigma i0)g^K(t_3 + \sigma i0, t_2).$$

It is worth noting that all nonequilibrium effects contribute to the regular part in Eq. (5.31) but not to the singular part; the states deep in the Fermi sea, which are not perturbed, contribute to it.

To solve Eq. (5.59), we define a new field $\tilde{K}_-$ by the relation

$$\tilde{K}_-(t, \mathbf{n}, \mathbf{r}) = (i\partial_t)^{-1}\hat{M}K_-, \quad (5.32)$$

where the operator $\hat{M}$ is shown below to be related to certain products of the Green functions g^K, see Eq. (5.38). The operator $\hat{M}$ is Hermitian and local in space but not in the momentum direction and time. We again use the operator notation

$$\hat{M}K_- \equiv \int dt_1 \int \frac{d\mathbf{n}_1}{\Omega_d} M(t, \mathbf{n}; t_1, \mathbf{n}_1; \mathbf{r})K_-(t_1, \mathbf{n}_1, \mathbf{r}). \quad (5.33)$$

We require K_+ to satisfy the equation

$$(\partial_t + \mathbf{v}\cdot\nabla - \hat{\mathrm{St}}_\tau)K_+ = \phi_+ - 2\tilde{K}_-, \quad (5.34)$$

whose solution is

$$K_+ = \hat{\mathcal{L}}^g\phi_+ - 2\hat{\mathcal{L}}^g\tilde{K}_-. \quad (5.35)$$

The operator notation here is the same as in Eq. (5.26).

The next task is to choose the "best" form for the operator $\hat{M}$ to maximally simplify the further perturbative expansion. Writing $\delta g^K = \delta g_+^K + \delta g_-^K$, we obtain the equations

$$i\hat{L}\delta g_-^K = \hat{Q}K_- + 2[\tilde{K}_-, g^K], \quad (5.36a)$$

$$i\hat{L}dg_+^K = \int \frac{d\mathbf{n}_2 d\mathbf{n}_3}{\Omega_d^2}\gamma\begin{pmatrix}\mathbf{n}_3\\ \mathbf{n};\ \mathbf{n}_2\end{pmatrix} \quad (5.36b)$$
$$\times [K_+(\mathbf{n}_3), (g^K(\mathbf{n}_2) - g^K(\mathbf{n}))],$$

where $\hat{L} \equiv (\tilde{\partial}_t + \mathbf{v}\cdot\tilde{\nabla} - \hat{\mathrm{St}}_\tau)$.

We note that the right-hand side of Eq. (5.36b) vanishes in equilibrium, because $g^K = \langle g^K\rangle_{\mathbf{n}}$. Therefore, δg_+^K also vanishes in equilibrium and cannot contribute to the equilibrium properties such as the specific heat. Moreover, even outside the equilibrium, $\delta g_+^K(t_1, t_1, \mathbf{n}, \mathbf{r}) = 0$, because the right-hand side of Eq. (5.36b) vanishes, see the remark after Eq. (5.31). This means that δg_+^K does not contribute to the electron density or current.

We are now ready to choose the operator $\hat{M}$. We require that $\delta g_-^K(t_1, t_1, \mathbf{n}, \mathbf{r}) = 0$; i.e., δg_-^K does not contribute to the electron density or current either. This means that the right-hand side of Eq. (5.36a) must vanish for $t_1 = t_2$ for any field K_-. Imposing this requirement, we obtain

$$\hat{M}(t_1, t_2, \mathbf{n}, \tilde{\mathbf{n}}, \mathbf{r}) = \frac{\pi}{4}\hat{Q}(t_1, t_1, \mathbf{n}; t_2, \tilde{\mathbf{n}}; \mathbf{r}). \quad (5.37)$$

Together with Eq. (5.30), it yields

$$M(t_1, \mathbf{n}_1; t_2, \mathbf{n}_2; \mathbf{r}) = \frac{\pi}{8}\int \frac{d\mathbf{n}_3}{\Omega_d}\gamma\begin{pmatrix}\mathbf{n}_2\\ \mathbf{n}_1;\ \mathbf{n}_3\end{pmatrix} \quad (5.38)$$
$$\times [Pg^K(t_1, t_2, \mathbf{n}_1)g^K(t_2, t_1, \mathbf{n}_3) + (\mathbf{n}_1 \rightleftharpoons \mathbf{n}_3)].$$

Expression (5.27) for the vertex γ enables us to establish the following properties of kernel (5.38):

$$M(t_1, \mathbf{n}_1; t_2, \mathbf{n}_2) = M(t_1, \mathbf{n}_2; t_2, \mathbf{n}_1) = M(t_2, \mathbf{n}_2; t_1, \mathbf{n}_1); \quad (5.38')$$

the operator $\hat{M}$ is therefore Hermitian.

It is instructive to find $\hat{M}$ in the thermal equilibrium. With Eq. (5.14), because Eq. (5.24) implies that the retarded and advanced components of $\hat{g}$ are still given by Eq. (5.23), it follows from Eqs. (5.28) and (5.38) that

$$M_{eq}(t_1, \mathbf{n}_1; t_2, \mathbf{n}_2; \mathbf{r}) = \int \frac{d\omega}{2\pi} e^{i\omega(t_2 - t_1)}\hat{M}_{eq}(\omega; \mathbf{n}_1, \mathbf{n}_2), \quad (5.39)$$
$$M_{eq}(\omega; \mathbf{n}_1, \mathbf{n}_2) = -\omega\coth\left(\frac{\omega}{2T}\right)[\hat{\mathrm{St}}_\tau](\mathbf{n}_1, \mathbf{n}_2).$$

Equation (5.39) is useful in checking the fluctuation–dissipation theorem.

5.3. Polarization Operators and Propagators

Knowledge of the linear-order corrections to the Green function permits the calculation of the polarization operators as variational derivatives of the original Green functions (i.e., before the gauge transformation)

in the limit $t_2 \longrightarrow t_1$, see Eq. (5.17). In the linear order, the corrections to the original Green functions are given by the relations (cf. Eq. (5.21))

$$\delta g^K \longrightarrow \delta g^K - i[K_+, g^K] - 2K_-\delta(t_1 - t_2), \quad (5.40a)$$

$$\delta g^Z \longrightarrow 2K_-\delta(t_1 - t_2). \quad (5.40b)$$

By construction in the previous subsection,

$$\lim_{t_2 \to t_1} \delta g^K(t_1, t_2, \mathbf{n}, \mathbf{r}) = 0, \quad (5.41)$$

and using Eq. (5.31), we have

$$\lim_{t_2 \to t_1} -i[K_+, g^K] = -\frac{2}{\pi}\partial_t K_+(t, \mathbf{n}, \mathbf{r}). \quad (5.42)$$

Substituting these results in Eq. (5.17) and using Eqs. (5.26), (5.32), and (5.35), we obtain

$$\begin{aligned}\Pi^R(1, 2) &= \nu[\delta_{12} - \partial_{t_1}\mathcal{L}^g(1, 2)],\\ \Pi^A(2, 1) &= \nu[\delta_{12} - \partial_{t_2}\overline{\mathcal{L}}^g(1, 2)];\end{aligned} \quad (5.43a)$$

we use notation (5.2) throughout this subsection. The result for the Keldysh component is

$$\Pi^K(1, 2) = 2i\nu[\hat{\mathcal{L}}^g\hat{M}\hat{\overline{\mathcal{L}}}^g](1, 2). \quad (5.43b)$$

The actions of the operators $\hat{M}$ and $\hat{D}$ are defined in Eqs. (5.26) and (5.33).

It is easy to check that the fluctuation–dissipation relation between the polarization operators holds in the thermal equilibrium. As follows from Eqs. (5.25) and (5.26),

$$-2\hat{\mathcal{L}}^g\hat{\mathrm{St}}_\tau\hat{\overline{\mathcal{L}}}^g = \hat{\mathcal{L}}^g + \hat{\overline{\mathcal{L}}}^g. \quad (5.44)$$

We perform the time Fourier transformation for all the propagators and the polarization operators in thermodynamic equilibrium,

$$A(1, 2) = \int\frac{d\omega}{2\pi}e^{i\omega(t_2 - t_1)}A(\omega; 1^*, 2^*). \quad (5.45)$$

Substituting Eqs. (5.39) and (5.44) in Eq. (5.43b), we obtain that in equilibrium,

$$\begin{aligned}&\Pi^K_{eq}(\omega; 1^*, 2^*)\\ &= [\Pi^R(\omega; 1^*, 2^*) - \Pi^A(\omega; 1^*, 2^*)]\coth\frac{\omega}{2T}.\end{aligned} \quad (5.46)$$

With the help of Eq. (5.15), we then derive the fluctuation–dissipation relation

$$\begin{aligned}&D^K_{eq}(\omega; 1^*, 2^*)\\ &= [D^R(\omega; 1^*, 2^*) - D^A(\omega; 1^*, 2^*)]\coth\frac{\omega}{2T}.\end{aligned} \quad (5.47)$$

With the expressions for the polarization operators obtained above, we can solve Dyson Eq. (5.15) and obtain the explicit expressions for the interaction propagators. In the operator notation,

$$\nu\hat{D}^R = -\frac{1}{1 + \hat{F} - \partial_t\hat{F}\hat{\mathcal{L}}^g}\hat{F}, \quad (5.48a)$$

$$\nu\hat{D}^A = -\hat{F}\frac{1}{1 + \hat{F} + \partial_t\hat{\overline{\mathcal{L}}}^g\hat{F}}, \quad (5.48b)$$

$$\hat{D}^K = 2i\nu\hat{D}^R\hat{\mathcal{L}}^g\hat{M}\hat{\overline{\mathcal{L}}}^g\hat{D}^A, \quad (5.48c)$$

where the action of the operator $\hat{F}$ on any function $a(t, \mathbf{n}, \mathbf{r})$ is defined by

$$\begin{aligned}[\hat{F}a](t; \mathbf{n}, \mathbf{r}) &\equiv \int\frac{d\mathbf{n}_1}{\Omega_d}\\ &\times [F^\rho(\widehat{\mathbf{n}\mathbf{n}}_1)a(t, \mathbf{n}_1, \mathbf{r}) + \int d\mathbf{r}_1\nu V(\mathbf{r} - \mathbf{r}_1)a(t, \mathbf{n}_1, \mathbf{r}_1)],\end{aligned} \quad (5.49)$$

see also the text after Eq. (5.15).

To find the propagators for the fields $K_\pm$ given in Eqs. (5.26') and (5.35), defined as

$$\begin{aligned}\langle\langle K_+(1)K_+(2)\rangle\rangle &= \frac{i}{2}\mathcal{H}^K(1, 2),\\ \langle\langle K_+(1)K_-(2)\rangle\rangle &= \frac{i}{2}\mathcal{H}^R(1, 2),\\ \langle\langle K_-(1)K_+(2)\rangle\rangle &= \frac{i}{2}\mathcal{H}^A(1, 2),\\ \langle\langle K_-(1)K_-(2)\rangle\rangle &= 0,\end{aligned} \quad (5.50)$$

we use Eqs. (5.1) and (5.48) and obtain the retarded and advanced propagators

$$\hat{\mathcal{H}}^R = \hat{\mathcal{L}}^g\hat{D}^R\hat{\mathcal{L}}^g, \quad \hat{\mathcal{H}}^A = \hat{\overline{\mathcal{L}}}^g\hat{D}^A\hat{\overline{\mathcal{L}}}^g, \quad (5.51a)$$

whereas the result for the Keldysh propagator is

$$\begin{aligned}\hat{\mathcal{H}}^K &= -\hat{\mathcal{L}}^g\hat{D}^K\hat{\overline{\mathcal{L}}}^g\\ &+ 2i[\hat{\mathcal{L}}^g(\partial_t)^{-1}\hat{M}\hat{\overline{\mathcal{L}}}^g\hat{D}^A\hat{\overline{\mathcal{L}}}^g - \hat{\mathcal{L}}^g\hat{D}^R\hat{\mathcal{L}}^g\hat{M}(\partial_t)^{-1}\hat{\overline{\mathcal{L}}}^g].\end{aligned} \quad (5.51b)$$

The fluctuation–dissipation relation between the D propagators in Eq. (5.47), the equilibrium form for $\hat{M}$ in Eq. (5.39), and identity (5.44) allow us to verify the fluctuation–dissipation relation for the $\mathcal{H}$ propagators:

$$\begin{aligned}&\mathcal{H}^K_{eq}(\omega; 1^*, 2^*)\\ &= [\mathcal{H}^R(\omega; 1^*, 2^*) - \mathcal{H}^A(\omega; 1^*, 2^*)]\coth\frac{\omega}{2T}.\end{aligned} \quad (5.52)$$

5.4. Additional Bosonic Fields

Equation (5.51b), together with Eqs. (5.51a) and (5.38), allows expressing the Keldysh propagator $\mathcal{H}^K$ in terms of the electron distribution function. This relation, however, is nonlocal on a spatial scale much larger than the temperature length

$$L_T \approx \min\left[\frac{\hbar v_F}{T}, v_F\sqrt{\frac{\hbar\tau}{T}}\right], \tag{5.53}$$

recall the discussion in Section 2.2. Indeed, the collision integral and all physical quantities are then given by integrals of the type

$$\mathcal{I}^\alpha = \int d\omega\, f(\omega)\mathcal{H}^\alpha(\omega), \quad \alpha = R, A, K,$$

where the function $f(\omega)$ depends on its argument at the characteristic scale of T. A retarded function is an analytic function of ω at $\mathrm{Im}\,\omega > 0$, which implies that for $\alpha = R$, the integral is determined only by singularities of $f(\omega)$, i.e., $\mathcal{I}^R \approx \mathcal{H}^R(\omega = iT)$. This immediately restricts the spatial scales to L_T. The same argument applies to the advanced case, because of the analyticity at $\mathrm{Im}\,\omega < 0$. But the function $\mathcal{H}^K(\omega)$ is not analytic. Moreover, according to Eq. (5.51b), it contains overlapping singularities of the retarded and advanced propagators. This means that the characteristic frequencies entering $\mathcal{I}^K$ are determined by the poles of the propagator rather than by the width of the function f; i.e., the spatial scale may far exceed L_T and any expression of type $\mathcal{I}^K$ is therefore nonlocal.

To overcome this difficulty, the standard parameterization of the Keldysh function $D^K = D^R \circ N - N \circ D^A$ is usually introduced and the kinetic equation for the distribution function N is then derived. All nonlocality in the problem is then contained in the partial solution of the kinetic equation, to be compared with Eq. (2.14), whereas the kinetic equation itself is *local*. In what follows, we adopt this program in a slightly modified form. We introduce a new retarded propagator $\mathcal{L}^\rho(1, 2)$ (to be compared with Eq. (5.25))

$$[i\hat{H}_{e-h}(i\partial_{t_1}, -i\nabla_1) - \hat{\mathrm{St}}_\tau]\mathcal{L}^\rho = \delta_{12},$$
$$\hat{H}_{e-h}(\omega, \mathbf{q}) = \mathbf{v}\cdot\mathbf{q} - \frac{\omega}{1+\hat{F}} \tag{5.54}$$

and its advanced counterpart $\overline{\mathcal{L}}^\rho(1, 2) = \mathcal{L}^\rho(2, 1)$. The multiplications in Eq. (5.54) are to be understood in the operator sense, and the action of the operator $\hat{F}$ on a function $a(t, \mathbf{n}, \mathbf{r})$ is defined by Eq. (5.49).

For $\hat{F} = 0$, $\mathcal{L}^\rho$ ($\overline{\mathcal{L}}^\rho$) reduces to the usual diffusion $\mathcal{L}^g$ ($\overline{\mathcal{L}}^g$). Physically, $\mathcal{L}^\rho$ describes the spectrum of a propagating electron–hole pair and the operator in the left-hand side of Eq. (5.54) corresponds to the kinetic equation for the collective mode in the Fermi liquid theory [26]. The operator $\hat{H}_{e-h}(\omega, \mathbf{q})$ can be interpreted as a "Hamiltonian" (cf. Appendix C) of the interacting electron–hole pair.

In terms of $\mathcal{L}^\rho$ and $\mathcal{L}^g$, Eqs. (5.51) acquire the form

$$\nu\hat{\mathcal{H}}^R = (\partial_t)^{-1}[\hat{\mathcal{L}}^g - \hat{\mathcal{L}}^\rho],$$
$$\nu\hat{\mathcal{H}}^A = -[\hat{\overline{\mathcal{L}}}^g - \hat{\overline{\mathcal{L}}}^\rho](\partial_t)^{-1}, \tag{5.55a}$$

$$\nu\hat{\mathcal{H}}^K = -2i(\partial_t)^{-1}$$
$$\times [\hat{\mathcal{L}}^g\hat{M}\hat{\overline{\mathcal{L}}}^g - \hat{\mathcal{L}}^\rho\hat{M}\hat{\overline{\mathcal{L}}}^\rho](\partial_t)^{-1}. \tag{5.55b}$$

We introduce two bosonic "distribution functions" (the density matrices, to be more precise) $\hat{\mathcal{N}}^g$ and $\hat{\mathcal{N}}^\rho$ that satisfy the equations

$$(\hat{\mathcal{L}}^g)^{-1}\hat{\mathcal{N}}^g + \hat{\mathcal{N}}^g(\hat{\overline{\mathcal{L}}}^g)^{-1} = 2\hat{M}, \tag{5.56a}$$

$$(\hat{\mathcal{L}}^\rho)^{-1}\hat{\mathcal{N}}^\rho + \hat{\mathcal{N}}^\rho(\hat{\overline{\mathcal{L}}}^\rho)^{-1} = 2\hat{M}. \tag{5.56b}$$

The operator $\hat{M}$ is defined in Eqs. (5.33) and (5.38) and, in a more explicit notation, the action of the operators $\hat{\mathcal{N}}^{\rho, g}$ on any function $a(t, \mathbf{n}, \mathbf{r})$ is understood as

$$[\hat{\mathcal{N}}^{\rho, g}a](1) = \int d2\,\mathcal{N}^{\rho, g}(1, 2)a(2),$$

where short notation (5.2) is used. We note that the above equations imply that the bosonic functions $\mathcal{N}^{\rho, g}$ are symmetric:

$$\mathcal{N}^{\rho, g}(1, 2) = \mathcal{N}^{\rho, g}(2, 1). \tag{5.57}$$

Equations (5.56) allow us to rewrite Eq. (5.55b) as

$$\nu\mathcal{H}^K = -i(\partial_t)^{-1}[\hat{\mathcal{L}}^g\hat{\mathcal{N}}^g + \hat{\mathcal{N}}^g\hat{\overline{\mathcal{L}}}^g](\partial_t)^{-1}$$
$$+ i(\partial_t)^{-1}[\hat{\mathcal{L}}^\rho\hat{\mathcal{N}}^\rho + \hat{\mathcal{N}}^\rho\hat{\overline{\mathcal{L}}}^\rho](\partial_t)^{-1}. \tag{5.58}$$

This expression is local in the sense discussed above and is used in the construction of the conserved energy current. Obtaining the local expression, however, requires the introduction of two additional bosonic distribution functions: $\mathcal{N}^\rho$, describing the interacting electron–hole pairs, and the ghost field distribution $\mathcal{N}^g$, subtracting the contribution of the electron–hole pairs in the absence of interactions.

Closing this subsection, we rewrite Eq. (5.56) in a form resembling the kinetic equation in Section 2. We substitute Eqs. (5.25) and (5.54) in Eq. (5.56) and obtain

$$[\partial_t + \mathbf{v}\cdot\nabla; \hat{\mathcal{N}}^g] = \hat{\mathrm{St}}^b\{\mathcal{N}^g, g^K\}, \tag{5.59a}$$

$$[i\hat{H}_{e-h}(i\partial_{t_1}, -i\nabla_1); \hat{\mathcal{N}}^\rho] = \hat{\mathrm{St}}^b\{\mathcal{N}^\rho, g^K\}, \tag{5.59b}$$

where the collision integrals are

$$\hat{\mathrm{St}}^b\{\mathcal{N}^g, g^K\} \equiv 2\{\hat{\mathrm{St}}_\tau; \hat{\mathcal{N}}^\alpha\} + 2\hat{M} \tag{5.59c}$$

for $\alpha = g, \rho$. They depend on g^K via $\hat{M}$ and we use the notation

$$\{\hat{A}; \hat{B}\} \equiv \frac{1}{2}(\hat{A}\hat{B} + \hat{B}\hat{A}),$$
$$[\hat{A}; \hat{B}] \equiv \hat{A}\hat{B} - \hat{B}\hat{A}. \tag{5.60}$$

We perform the time and space Wigner transformations of Eqs. (5.59) to introduce the bosonic distribution functions $N^{g,\rho}$,

$$\hat{\mathcal{N}}^{g,\rho}(1,2) = \int\frac{d\omega}{2\pi}e^{-i\omega(t_1-t_2)}\int\frac{d^dq}{(2\pi)^d}e^{i\mathbf{q}\cdot(\mathbf{r}_1-\mathbf{r}_2)} \times \omega[2N^{g,\rho}(\omega, \mathbf{q}; \mathbf{n}_1, \mathbf{n}_2; t, \mathbf{r}) + \Omega_d\delta(\widehat{\mathbf{n}_1, \mathbf{n}_2})], \tag{5.61}$$

where $t = (t_1 + t_2)/2$, $\mathbf{r} = (\mathbf{r}_1 + \mathbf{r}_2)/2$. Symmetry relation (5.57) transform into the condition

$$N^{g,\rho}(\omega, \mathbf{q}; \mathbf{n}_1, \mathbf{n}_2) = -[N^{g,\rho}(-\omega, -\mathbf{q}; \mathbf{n}_2, \mathbf{n}_1) + \Omega_d\delta(\widehat{\mathbf{n}_1, \mathbf{n}_2})]. \tag{5.62}$$

The physical meaning of this relation is the Bose statistics: at $\omega > 0$, $N^{g,\rho}$ corresponds to the occupation numbers entering the probability of the absorption of the bosons, whereas the $\omega < 0$ part describes the boson emission.

The fermionic distribution function f is obtained in two steps: (a) we introduce the gauge-invariant Green function g (also see the next subsection) and (b) we perform the time Wigner transformation:

$$g^K(t_1, t_2, \mathbf{n}, \mathbf{r}) \equiv \exp\left(-i\int_{t_2}^{t_1}d\tilde{t}\,\varphi(\tilde{t}, \mathbf{r})\right)g(t_1, t_2, \mathbf{n}, \mathbf{r}), \tag{5.63a}$$

$$g(t_1, t_2, \mathbf{n}, \mathbf{r}) = 2\int\frac{d\varepsilon}{2\pi}e^{-i\varepsilon(t_1-t_2)}[1 - 2f(\varepsilon, \mathbf{n}; t, \mathbf{r})]. \tag{5.63b}$$

Performing such Wigner transformations of Eqs. (5.56) and (5.38), we find

$$\omega[\partial_t\hat{N}^g + \{\mathbf{v}; \partial\hat{N}^g\} + i[\mathbf{v}\cdot\mathbf{q}; \hat{N}^g]] = \hat{\mathrm{St}}^b\{N^g, f\}, \tag{5.64a}$$

$$\omega\left[\left\{\frac{1}{1+\hat{F}}; \partial_t\hat{N}^\rho\right\} + \{\hat{\mathbf{s}}(\omega, \mathbf{q}); \nabla\hat{N}^\rho\} + i[\hat{H}_{e-h}(\omega, \mathbf{q}), \hat{N}^\rho]\right] = \hat{\mathrm{St}}^b\{N^\rho, f\}, \tag{5.64b}$$

where the collective mode velocity operator is

$$\hat{\mathbf{s}}(\omega, \mathbf{q}) = \frac{\partial\hat{H}_{e-h}(\omega, \mathbf{q})}{\partial\mathbf{q}} = \mathbf{v} + \omega\frac{\partial}{\partial\mathbf{q}}\left(\frac{\hat{F}}{1+\hat{F}}\right). \tag{5.65}$$

In the left-hand side of Eq. (5.64b), we limited ourselves to the leading Poisson brackets (the equation becomes exact for a short-range interaction because $\partial_{\mathbf{q}}\hat{F} = 0$, and in the unitary limit, $\hat{F} \longrightarrow \infty$). However, no Poisson brackets arise in the right-hand sides of Eqs. (5.64) as a consequence of the locality of the kinetic equations.

The right-hand sides of Eq. (5.64) describe the decay of an electron–hole pair into an electron and a hole moving in different directions. To write the expression for this collision term, it is convenient to introduce the following object:

$$\Upsilon^{g,\rho}_{ij;\,kl}(\varepsilon, \omega, \mathbf{q}; t, \mathbf{r}) \equiv N^{g,\rho}(\omega, \mathbf{q}; \mathbf{n}_i, \mathbf{n}_j; t, \mathbf{r}) \times\{f(\varepsilon, \mathbf{n}_k; t, \mathbf{r}) - f(\varepsilon - \omega, \mathbf{n}_k; t, \mathbf{r})\} + \Omega_d\delta(\widehat{\mathbf{n}_i\mathbf{n}_j})\{f(\varepsilon, \mathbf{n}_l; t, \mathbf{r})[1 - f(\varepsilon - \omega, \mathbf{n}_k; t, \mathbf{r})]\}. \tag{5.66}$$

It is easy to see that $\Upsilon^{g,\rho} = 0$ in thermal equilibrium, Eq. (3.1).

In terms of this object and vertex (5.27), we have

$$\hat{\mathrm{St}}^b\{N^{g,\rho}, f\}(\omega, \mathbf{q}; \mathbf{n}_1, \mathbf{n}_2; t, \mathbf{r}) = -\int d\varepsilon\int\frac{d\mathbf{n}_3 d\mathbf{n}_4}{\Omega_d^2} \times\{\gamma_{14}^3\Upsilon^{g,\rho}_{32;\,41}(\varepsilon, \omega) + \gamma_{34}^2\Upsilon^{g,\rho}_{13;\,14}(\varepsilon, \omega)\}, \tag{5.67}$$

where we suppress the spectator arguments t, $\mathbf{r}$, and $\mathbf{q}$ in the right-hand side of the equation. In deriving Eq. (5.67), we used Eqs. (5.28) and (5.38') and the property

$$\int d\varepsilon[f(\varepsilon) - f(\varepsilon - \omega)] = -\omega.$$

To understand the physical meaning of the processes described by collision integral (5.67), we use the explicit form of the vertex γ (Eq. (5.27)) for the isotro-

pic impurity scattering $\tau(\theta_{12}) = \tau$. Then the collision integrals can be decomposed into the sum of two contributions:

$$\hat{\mathrm{St}}^b\{N^{g,\rho}, f\} = \hat{\mathrm{St}}^b_{cl}\{N^{g,\rho}, f\} + \hat{\mathrm{St}}^b_q\{N^{g,\rho}, f\}.$$

The first term in the right-hand side can be obtained from a simple counting of the probabilities of the processes depicted in Figs. 2a and 2b:

$$\hat{\mathrm{St}}^b_{cl}\{N^{g,\rho}, f\}\{\omega; \mathbf{n}_1, \mathbf{n}_2\} = \frac{1}{\tau}\int d\varepsilon \int \frac{d\mathbf{n}_3}{\Omega_d}\{\Upsilon^{g,\rho}_{12;\,32}(\varepsilon, \omega) + \Upsilon^{g,\rho}_{12;\,13}(\varepsilon, \omega)\}.$$

The second term in the right-hand side originates from the interference of two scattering processes, see Fig. 2c. It therefore makes contributions to $\hat{N}$ that are not diagonal in the momentum directions:

$$\hat{\mathrm{St}}^b_q\{N^{g,\rho}, f\}(\omega; \mathbf{n}_1, \mathbf{n}_2) = -\frac{1}{\tau}\int d\varepsilon \int \frac{d\mathbf{n}_3}{\Omega_d}\{\Upsilon^{g,\rho}_{13;\,12}(\varepsilon, \omega) + \Upsilon^{g,\rho}_{32;\,21}(\varepsilon, \omega)\}.$$

5.5. The Collision Integral for Electrons

With the bosonic propagators $\mathscr{K}$ at hand, we can proceed with the calculation and include the second-order contributions in the fluctuating fields $K_\pm$ to the collision term of Eilenberger Eq. (5.22). With the fluctuating fields $K_\pm$ given by Eqs. (5.26') and (5.35), the Eilenberger equation becomes

$$[\tilde{\partial}_t + \mathbf{v}\cdot\tilde{\nabla}]\hat{g} = \left[\hat{g} \circ \frac{1}{2}\hat{\mathrm{St}}^{\phi}_{\tau}\hat{g} - i(\hat{\mathrm{St}}_{\tau}K_+ - 2\tilde{K}_-)\hat{\mathbb{1}}_K + i\hat{\mathrm{St}}_{\tau}K_-\hat{\sigma}^x_K\right], \quad (5.68)$$

where we use notation (5.11) for the derivatives, $\hat{\mathrm{St}}^{\phi}_{\tau}$ is defined after Eq. (5.22), $\tilde{K}_-$ is defined in Eq. (5.32), and $\hat{\sigma}^x_K$ is the Pauli matrix.

We expand the right-hand side of Eq. (5.68) up to the second order in $\hat{K}$, see Fig. 3a; then we average it to obtain Fig. 3b. The resulting second-order contributions can have two different origins: (1) they can arise from the expansion of the exponentials truncated at the second order, the term $\hat{\mathrm{St}}_1$ in Fig. 3c, or (2) they are obtained as products of the linear correction δg^K of Eq. (5.36a) and the first-order expansion of the exponentials, the term $\hat{\mathrm{St}}_2$ in Fig. 3c.

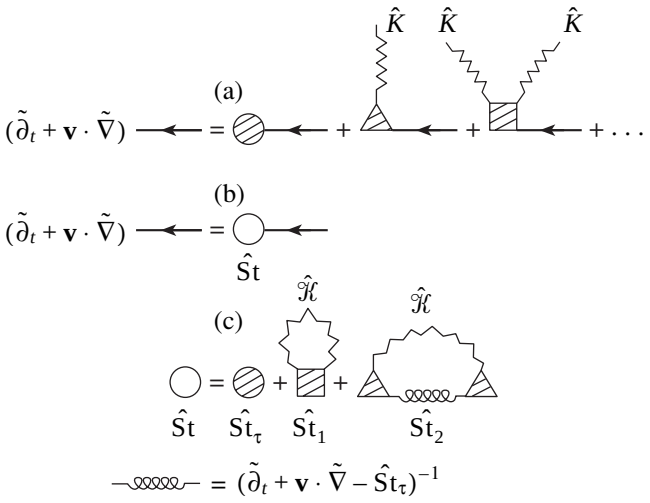

Fig. 3. Schematic representation of averaging over the fluctuation fields $\hat{K}$. (a) Expansion of Eilenberger Eq. (5.68) before averaging. (b) The equation for the Green function averaged over $\hat{K}$, see Eq. (5.69). (c) The contributions to the total collision integral in the one-loop approximation.

The Eilenberger equation for the averaged Green function takes the form

$$[\tilde{\partial}_t + \mathbf{v}\cdot\tilde{\nabla}]\hat{g} = \hat{\hat{\mathrm{St}}}\{\hat{g}, N^{\rho}, N^g\}, \quad (5.69)$$

where $\hat{\hat{\mathrm{St}}}$ contains both zeroth and second-order contributions. We find (see Appendix D for details on the cancellation of second-order corrections in the R, A, Z sectors)

$$\hat{\hat{\mathrm{St}}}\{g^R_0, g^A_0, g^K\} = \begin{pmatrix} 0 & \hat{\mathrm{St}}\{g^K, N^{\rho}, N^g\} \\ 0 & 0 \end{pmatrix}, \quad (5.70)$$

where $g^R_0 = -g^A_0 = \delta(t_1 - t_2)$. This means that the matrix Green function of form (5.23) is still a solution of the Eilenberger equation—the main gain of gauge transformation (5.21)—provided that the kinetic equation for the Keldysh component is satisfied; accordingly, we concentrate on this component only.

Performing gauge transformation (5.63a) of the Keldysh component of the Eilenberger equation, we arrive at the explicitly gauge-invariant form of the kinetic equation

$$\left[\partial_{t_1} + \partial_{t_2} + \mathbf{v}\cdot\nabla + i\mathbf{v}\cdot\int_{t_2}^{t_1} d\tilde{t}\, e\mathbf{E}(\tilde{t}, \mathbf{r})\right] g = \hat{\mathrm{St}}\{g, N^{\rho}, N^g\}. \quad (5.71)$$

In the collision integral $\hat{\mathrm{St}}\{g, N^{\rho}, N^g\}$, we use $\hat{\mathrm{St}}_1\{g, N^{\rho}, N^g\}$ to denote contributions of type (1) and

$\hat{\mathrm{St}}_2\{g, N^\rho, N^g\}$ for those of type (2);[6]

$$\hat{\mathrm{St}}\{g, N^\rho, N^g\} = \hat{\mathrm{St}}_\tau g - 4\hat{\mathrm{St}}_1\{g, N^\rho, N^g\} - 4\hat{\mathrm{St}}_2\{g, N^\rho, N^g\}, \tag{5.72}$$

where $\hat{\mathrm{St}}_\tau$ is defined in Eq. (5.8). The numerical factors in front of the last two terms are introduced to facilitate the transformation to the canonical form of the kinetic equation in subsequent sections.

The expression for $\hat{\mathrm{St}}_1$ written in terms of the $\mathcal{K}$ propagators (5.50) and the γ vertex (5.27) is

$$\begin{aligned}[\hat{\mathrm{St}}_1](t_1, t_2; \mathbf{n}_1) &= -\frac{i}{16}\int\frac{d\mathbf{n}_2 d\mathbf{n}_3}{\Omega_d^2}\gamma_{12}^3 \\ &\times\{g(t_1, t_2, \mathbf{n}_2)[\tilde{\mathcal{K}}_{32}^K(t_1, t_2) - \tilde{\mathcal{K}}_{31}^K(t_1, t_2)] \\ &+\int dt_3 g(t_1, t_3, \mathbf{n}_1)g(t_3, t_2, \mathbf{n}_2)[\mathcal{K}_{32}^A(t_3, t_2) - \mathcal{K}_{31}^A(t_3, t_2)] \\ &- g(t_1, t_3, \mathbf{n}_2)g(t_3, t_2, \mathbf{n}_1)[\mathcal{K}_{23}^R(t_1, t_3) - \mathcal{K}_{13}^R(t_1, t_3)]\}.\end{aligned} \tag{5.73}$$

We introduce the short notation

$$\mathcal{K}_{ij}(t_1, t_2) \equiv \mathcal{K}(t_1, \mathbf{n}_i; t_2, \mathbf{n}_j) \tag{5.74}$$

and

$$\tilde{\mathcal{K}}^K(t_1, t_2) \equiv 2\mathcal{K}^K(t_1, t_2) - \mathcal{K}^K(t_1, t_1) - \mathcal{K}^K(t_2, t_2).$$

We omit the variable $\mathbf{r}$, which always appears in the distribution function as $g(t_1, t_2, \mathbf{n}, \mathbf{r})$ and in the propagators as $\mathcal{K}(t_1, \mathbf{n}_1, \mathbf{r}; t_2, \mathbf{n}_2, \mathbf{r})$. The dependence on the electron distribution function g is explicit, whereas the dependence on the bosonic distribution functions is hidden into the propagators, see Eq. (5.58). For reasons that will become clear in the next subsection, we split Eq. (5.73) into two parts:

$$\hat{\mathrm{St}}_1 = \hat{\mathrm{St}}_1^{in} + \hat{\mathrm{St}}_1^{el}; \tag{5.75a}$$

$$\begin{aligned}[\hat{\mathrm{St}}_1^{\mathrm{in}}](t_1, t_2; \mathbf{n}_1) &= -\frac{i}{32}\int\frac{d\mathbf{n}_2 d\mathbf{n}_3}{\Omega_d^2}\gamma_{12}^3 \\ &\times\{2g(t_1, t_2, \mathbf{n}_2)[\tilde{\mathcal{K}}_{32}^K(t_1, t_2) - \tilde{\mathcal{K}}_{31}^K(t_1, t_2)] \\ &+\int dt_3[g(t_1, t_3, \mathbf{n}_1)g(t_3, t_2, \mathbf{n}_2) \\ &+ g(t_1, t_3, \mathbf{n}_2)g(t_3, t_2, \mathbf{n}_1)] \\ &\times[\mathcal{K}_{32}^A(t_3, t_2) - \mathcal{K}_{23}^R(t_1, t_3) - \mathcal{K}_{31}^A(t_3, t_2) + \mathcal{K}_{13}^R(t_1, t_3)]\},\end{aligned} \tag{5.75b}$$

$$\begin{aligned}[\hat{\mathrm{St}}_1^{\mathrm{el}}](t_1, t_2; \mathbf{n}_1) &= -\frac{i}{32}\int dt_3\int\frac{d\mathbf{n}_2 d\mathbf{n}_3}{\Omega_d^2}\gamma_{12}^3 \\ &\{[g(t_1, t_3, \mathbf{n}_1)g(t_3, t_2, \mathbf{n}_2) - g(t_1, t_3, \mathbf{n}_2)g(t_3, t_2, \mathbf{n}_1)] \\ &\times[\mathcal{K}_{32}^A(t_3, t_2) + \mathcal{K}_{23}^R(t_1, t_3) - \mathcal{K}_{31}^A(t_3, t_2) - \mathcal{K}_{13}^R(t_1, t_3)]\}.\end{aligned} \tag{5.75c}$$

As regards $\hat{\mathrm{St}}_2$, it is convenient to separate it into two parts, depending on which field, K_+ or K_-, we retain in the expansion,

$$\hat{\mathrm{St}}_2 = \langle\langle\hat{\mathrm{St}}_+\rangle\rangle + \langle\langle\hat{\mathrm{St}}_-\rangle\rangle, \tag{5.76a}$$

where $\langle\langle\ldots\rangle\rangle$ denotes averaging over the fluctuating fields $K_\pm$ with propagators (5.50) and

$$\begin{aligned}\hat{\mathrm{St}}_-(t_1, t_2; \mathbf{n}, \mathbf{r}) &= \frac{i}{4}\int dt_3\int\frac{d\mathbf{n}_1}{\Omega_d}K_-(t_3, \mathbf{n}_1, \mathbf{r}) \\ &\times\left[\delta Q(t_1, t_2, \mathbf{n}; t_3, \mathbf{n}_1; \mathbf{r}) - \frac{i\pi}{2}\delta g_+(t_1, t_2; \mathbf{n}, \mathbf{r})\right. \\ &\left.\times\int_{t_2}^{t_1} dt_5 Q(t_5, t_5, \mathbf{n}; t_3, \mathbf{n}_1; \mathbf{r})\right],\end{aligned} \tag{5.76b}$$

$$\begin{aligned}\hat{\mathrm{St}}_+(t_1, t_2; \mathbf{n}) &= \frac{i}{4}\int\frac{d\mathbf{n}_2 d\mathbf{n}_3}{\Omega_d^2}\begin{pmatrix}\mathbf{n}_3 \\ \mathbf{n}; \mathbf{n}_2\end{pmatrix} \\ &\times[K_+(\mathbf{n}_3); (\delta g(\mathbf{n}_2) - \delta g(\mathbf{n}))](t_1, t_2).\end{aligned} \tag{5.76c}$$

The commutator is defined in Eq. (5.10) and the kernel δQ is the first variation of operator (5.30) with respect to the Keldysh component of the electron Green function,

$$\begin{aligned}\delta Q(t_1, t_2, \mathbf{n}; t_3, \mathbf{n}_1; \mathbf{r}) &= \frac{1}{2}\int\frac{d\mathbf{n}_2}{\Omega_d}\gamma\begin{pmatrix}\mathbf{n}_1 \\ \mathbf{n}; \mathbf{n}_2\end{pmatrix} \\ &\times\{[\delta g_+(t_1, t_3, \mathbf{n}, \mathbf{r})g(t_3, t_2, \mathbf{n}_2, \mathbf{r}) \\ &+ g(t_1, t_3, \mathbf{n}, \mathbf{r})\delta g_+(t_3, t_2, \mathbf{n}_2, \mathbf{r})] + (\mathbf{n}_2 \rightleftharpoons \mathbf{n})\}.\end{aligned} \tag{5.76d}$$

Finally, the functions $\delta g_\pm$ and $\delta g = \delta g_+ + \delta g_-$ are obtained by solving Eqs. (5.36a) and (5.36b) (after transformation (5.63a)); with the help of Eq. (5.37), we have

$$\begin{aligned}\delta g_+(t_1, t_2; \mathbf{n}, \mathbf{r}) &= -i\int d3\int\frac{d\mathbf{n}_2 d\mathbf{n}_4}{\Omega_d^2}\gamma_{32}^4\mathcal{L}^g(t_1, \mathbf{n}, \mathbf{r}; 3) \\ &\times[K_+(t_3, \mathbf{n}_4, \mathbf{r}_3) - K_+(t_3 - t_{12}, \mathbf{n}_4, \mathbf{r}_3)] \\ &\times[g^K(t_3, t_3 - t_{12}; \mathbf{n}_2, \mathbf{r}_3) - g^K(t_3, t_3 - t_{12}; \mathbf{n}_3, \mathbf{r}_3)],\end{aligned} \tag{5.77}$$

[6] This separation has no particular physical meaning; it is just a matter of practicality in the calculations; we return to the physical aspects when we analyze the conservation laws in the next subsection.

$$\delta g_-(t_1, t_2; \mathbf{n}, \mathbf{r}) = -i\int d3\int dt_4\int\frac{d\mathbf{n}_4}{\Omega_d}\mathcal{L}^g(t_1, \mathbf{n}, \mathbf{r}; 3)$$

$$\times K_-(t_4, \mathbf{n}_4, \mathbf{r}_3)\Bigg[Q(t_3, t_3 - t_{12}, \mathbf{n}_3; t_4, \mathbf{n}_4; \mathbf{r}_3) \tag{5.78}$$

$$-\frac{i\pi}{2}g^K(t_3, t_3 - t_{12}; \mathbf{n}_3, \mathbf{r}_3)\int\limits_{t_3 - t_{12}}^{t_3} dt_5 Q(t_5, t_5, \mathbf{n}_3; t_4, \mathbf{n}_4; \mathbf{r}_3)\Bigg],$$

where $t_{12} = t_1 - t_2$ and notation (5.2) is used.[7]

For future use, we note the properties (see also the discussion following Eq. (5.36))

$$\delta g_-(t, t) = \delta g_+(t, t) = 0,$$
$$\int d\mathbf{n}\delta Q(t_1, t_2, \mathbf{n}; t_3, \mathbf{n}_1; \mathbf{r}) = 0, \tag{5.79}$$

which follow from Eqs. (5.77)–(5.78) and from definitions (5.76d) and (5.27).

The canonical form of the kinetic equation is obtained by performing the time Wigner transformation (5.63b) of both sides of Eq. (5.71). It is clear from the structure of the collision integrals that this procedure leads to the appearance of Poisson brackets in the right-hand side of the kinetic equation. We choose another route, however: we prove the existence of the conservation laws before the Wigner transformation. This then allows us to argue that these Poisson brackets (in our formulation of the kinetic equation) give only small contributions, which can be neglected within the accuracy of the kinetic equation.

5.6. Conservation Laws

The derivation of the conservation laws is based on the following properties of the collision integrals in the previous subsection:

$$\langle\hat{\mathrm{St}}_\tau g\rangle_{\mathbf{n}} = 0, \tag{5.80a}$$

$$\langle\hat{\mathrm{St}}_-\rangle_{\mathbf{n}} = 0, \tag{5.80b}$$

$$\langle\hat{\mathrm{St}}_1^{\mathrm{el}}\rangle_{\mathbf{n}} = 0. \tag{5.80c}$$

The physical meaning of conditions (5.80) is that the corresponding terms in the collision integral preserve the number of particles within the energy shell. Equations (5.80) follow immediately from definitions (5.8), (5.75b), (5.76b), (5.76d), and (5.79).

The two remaining contributions to the collision integral have the properties

$$\lim_{t_1 \to t_2}\hat{\mathrm{St}}_+(t_1, t_2) = 0, \tag{5.81a}$$

$$\lim_{t_1 \to t_2}\hat{\mathrm{St}}_1^{\mathrm{in}}(t_1, t_2) = 0, \tag{5.81b}$$

and

$$\lim_{t_1 \to t_2}(\partial_{t_1} - \partial_{t_2})\hat{\mathrm{St}}_+(t_1, t_2) = 0, \tag{5.82a}$$

$$\lim_{t_1 \to t_2}(\partial_{t_1} - \partial_{t_2})\hat{\mathrm{St}}_1^{\mathrm{in}}(t_1, t_2)$$
$$= \frac{1}{4\pi}\int\frac{d\mathbf{n}_2 d\mathbf{n}_3}{\Omega_d^2}\gamma_{12}^3\Big\{[\partial_t\mathcal{H}^K\partial_t]_{31}(t_1, t_1) \tag{5.82b}$$
$$-\frac{i\pi}{4}\int dt_3[\partial_t\mathcal{H}^R]_{13}(t_1, t_3)[g(t_1, t_3, \mathbf{n}_1)g(t_3, t_1, \mathbf{n}_2)$$
$$+ g(t_1, t_3, \mathbf{n}_2)g(t_3, t_1, \mathbf{n}_1)] - (\mathbf{n}_1 \rightleftarrows \mathbf{n}_2)\Big\},$$

where we use notation (5.74) and the vertex is defined in Eq. (5.28). Equations (5.81a) and (5.82a) immediately follow from definition (5.76c) and condition (5.79). Derivations of Eqs. (5.81b) and (5.82b) are given in Appendix E.

Expressions (5.81) mean that, while not preserving the number of particles for a given energy shell, the terms $\hat{\mathrm{St}}_+$ and $\hat{\mathrm{St}}_1^{\mathrm{in}}$ preserve the total number of particles for a given direction (small-angle inelastic scattering). Equation (5.82a) means that the inelastic $\hat{\mathrm{St}}_+$ term preserves not only the number of particles but also the energy for a given direction. Equation (5.82b) means that the $\hat{\mathrm{St}}_1^{\mathrm{in}}$ term does not preserve the energy for a given direction, thus describing the energy exchange between quasiparticles and electron–hole pairs discussed in Section 5.3.

The possibility of finding the conserved energy current is based on a certain relation between Eq. (5.82b) and the collision integral for electron–hole pairs. We now turn to the discussion of this relation.

We substitute Eqs. (5.55a) and (5.58) in Eq. (5.82b) and average the result over **n**; then using the analytical property

$$\mathcal{H}^R(t, t) = \int\frac{d\omega}{2\pi}\mathcal{H}^R(\omega) = 0 \tag{5.83}$$

together with $\mathcal{L}^\rho(1, 2) = \overline{\mathcal{L}}^\rho(2, 1)$, $\mathcal{L}^g(1, 2) = \overline{\mathcal{L}}^g(2, 1)$,

[7] These solutions are exact only in the absence of the electric field, because in its gauge invariant form, the operator acting on $\delta g_\pm$ is the same that appears in the left-hand side of Eq. (5.71). We could perturbatively include field-dependent corrections into our expressions, which would be of the first order in **E** for δg_- and of the second order for δg_+ (because δg_+ vanishes in equilibrium). However, as noted above, the first property in Eq. (5.79) implies that these corrections cannot contribute to the physical quantities in which we are interested, and therefore we do not include them in our calculations.

$\mathcal{N}^{\rho,g}(1,2) = \mathcal{N}^{\rho,g}(2,1)$, and Eq. (5.38), we find

$$\lim_{t_1\to t_2}(\partial_{t_1}-\partial_{t_2})\langle \hat{\mathrm{St}}_1^{\mathrm{in}}\rangle_{\mathbf{n}} = -\frac{i}{\pi\nu}\mathrm{Tr}_{\mathbf{n}}[\hat{\mathcal{L}}^{\rho}(\hat{\mathcal{N}}^{g}\hat{\mathrm{St}}_\tau + \hat{M})] + \frac{i}{\pi\nu}\mathrm{Tr}_{\mathbf{n}}[\hat{\mathcal{L}}^{\rho}(\hat{\mathcal{N}}^{\rho}\hat{\mathrm{St}}_\tau + \hat{M})], \tag{5.84}$$

where $\hat{\mathrm{St}}_\tau$ is defined in Eq. (5.8) and $\mathrm{Tr}_{\mathbf{n}}$ acts as

$$[\mathrm{Tr}_{\mathbf{n}}\hat{A}](t,\mathbf{r}) \equiv \int\frac{d\mathbf{n}}{\Omega_d}A(t,\mathbf{n},\mathbf{r};\,t,\mathbf{n},\mathbf{r}). \tag{5.85}$$

The corresponding traces of the collision integrals for the electron–hole pairs, Eq. (5.59c),[8] are

$$\mathrm{Tr}_{\mathbf{n}}[\hat{\mathcal{L}}^{g}\hat{\mathrm{St}}^{g}\{\mathcal{N}^{g},g\}] = \mathrm{Tr}_{\mathbf{n}}[[\hat{\mathcal{L}}^{g};\,\hat{\mathrm{St}}_\tau].\hat{\mathcal{N}}^{g} + 2\hat{\mathcal{L}}^{g}(\hat{\mathcal{N}}^{g}\hat{\mathrm{St}}_\tau + \hat{M})], \tag{5.86a}$$

$$\mathrm{Tr}_{\mathbf{n}}[\hat{\mathcal{L}}^{\rho}\hat{\mathrm{St}}^{\rho}\{\mathcal{N}^{\rho},g\}] = \mathrm{Tr}_{\mathbf{n}}[[\hat{\mathcal{L}}^{\rho};\,\hat{\mathrm{St}}_\tau].\hat{\mathcal{N}}^{\rho} + 2\hat{\mathcal{L}}^{\rho}(\hat{\mathcal{N}}^{\rho}\hat{\mathrm{St}}_\tau + \hat{M})]. \tag{5.86b}$$

Comparing Eqs. (5.84) and (5.86) and using Eqs. (5.25) and (5.54) once again, we obtain the desired relation between the collision integrals:

$$2i\pi\nu\lim_{t_1\to t_2}(\partial_{t_1}-\partial_{t_2})\langle\hat{\mathrm{St}}\{g,\mathcal{N}^{\rho},\mathcal{N}^{g}\}\rangle_{\mathbf{n}} + \mathrm{Tr}_{\mathbf{n}}[\hat{\mathcal{L}}^{\rho}\hat{\mathrm{St}}^{\rho}\{\mathcal{N}^{\rho},g\}] - \mathrm{Tr}_{\mathbf{n}}[\hat{\mathcal{L}}^{g}\hat{\mathrm{St}}^{g}\{\mathcal{N}^{g},g\}] = \mathrm{Tr}_{\mathbf{n}}[[\mathbf{v}\cdot\nabla;\,\hat{\mathcal{L}}^{g}]\hat{\mathcal{N}}^{g} - i[\hat{H}_{e-h}(i\partial_t,-i\nabla);\,\hat{\mathcal{L}}^{\rho}]\hat{\mathcal{N}}^{\rho}]. \tag{5.87}$$

The left-hand side of Eq. (5.87) is the quantum counterpart of relation (2.3b), and Eqs. (5.80)–(5.81) are related to Eq. (2.3a); we now derive the expressions for the electric and energy currents in the spirit of our discussion in Section 2.1.

We begin with the conservation of electric charge. According to Eqs. (5.4) and (5.6), the charge density is given by

$$\rho(t,\mathbf{r}) = -\frac{e\nu\pi}{2}\lim_{t_1\to t_2\to t}\langle g(t_1,t_2,\mathbf{n},\mathbf{r})\rangle_{\mathbf{n}}. \tag{5.88}$$

Taking the limit $t_1 \longrightarrow t_2 \longrightarrow t$ in both sides of Eq. (5.71) and using Eqs. (5.72), (5.80), and (5.81), we obtain the continuity equation

$$\partial_t\rho + \nabla\cdot\mathbf{j} = 0, \tag{5.89}$$

where

$$\mathbf{j}(t,\mathbf{r}) = -\frac{e v_{\mathrm{F}}\nu\pi}{2}\lim_{t_1\to t_2\to t}\langle\mathbf{n}g(t_1,t_2,\mathbf{n},\mathbf{r})\rangle_{\mathbf{n}}, \tag{5.90}$$

to be compared with Eqs. (2.5)–(2.7).

[8] The operator $\hat{M}$ is gauge-invariant and has the same form in terms of g or g^K.

Having found the usual equation for the electric currents, we turn to the energy conservation. Acting with the operator $(\partial_{t_1} - \partial_{t_2})$ on both sides of Eq. (5.71) and introducing the quantities

$$u_e(t,\mathbf{r}) = -\frac{i\pi\nu}{4}\lim_{t_1\to t_2\to t}(\partial_{t_1}-\partial_{t_2})\langle g(t_1,t_2,\mathbf{n},\mathbf{r})\rangle_{\mathbf{n}},$$
$$\mathbf{j}_e^{\varepsilon}(t,\mathbf{r}) = -\frac{i\pi\nu v_{\mathrm{F}}}{4}\lim_{t_1\to t_2\to t}(\partial_{t_1}-\partial_{t_2})\langle\mathbf{n}g(t_1,t_2,\mathbf{n},\mathbf{r})\rangle_{\mathbf{n}}, \tag{5.91}$$

we find

$$\partial_t u_e + \nabla\cdot\mathbf{j}_e^{\varepsilon} = \mathbf{j}\cdot\mathbf{E} + i\nu\pi\lim_{t_1\to t_2\to t}(\partial_{t_1}-\partial_{t_2})\langle\hat{\mathrm{St}}\{g,\mathcal{N}^{\rho},\mathcal{N}^{g}\}\rangle_{\mathbf{n}}. \tag{5.92}$$

The expression in the left-hand side of Eq. (5.92) has the form of a continuity equation for the energy current of electrons: the first term in the right-hand side is the Joule heat acting as an energy source. The last term in the right-hand side indicates that the electron system is open by itself, due to the energy exchange with electron–hole pairs. As we discussed in Section 2.1, this means that the contribution of these degrees of freedom must be taken into account in the definition of the conserved energy and energy current densities. For this, we multiply Eqs. (5.59a) and (5.59b) by $\hat{\mathcal{L}}^{g}$ and $\hat{\mathcal{L}}^{\rho}$, respectively. Using Eqs. (5.25), (5.54), and (5.59c) and taking the trace $\mathrm{Tr}_{\mathbf{n}}$ (see Eq. (5.85)) of both sides, we obtain

$$\partial_t u_g + \nabla\cdot\mathbf{j}_g^{\varepsilon} - \frac{1}{2}\mathrm{Tr}_{\mathbf{n}}[\mathbf{v}\cdot\nabla;\,\hat{\mathcal{L}}^{g}]\hat{\mathcal{N}}^{g} = \frac{1}{2}\mathrm{Tr}_{\mathbf{n}}\hat{\mathcal{L}}^{g}\hat{\mathrm{St}}^{g}, \tag{5.93a}$$

$$\partial_t u_\rho + \nabla\cdot\mathbf{j}_\rho^{\varepsilon} - \frac{i}{2}\mathrm{Tr}_{\mathbf{n}}[\hat{H}_{e-h}(i\partial_t,-i\nabla);\,\hat{\mathcal{L}}^{\rho}]\hat{\mathcal{N}}^{\rho} = \frac{1}{2}\mathrm{Tr}_{\mathbf{n}}\hat{\mathcal{L}}^{\rho}\hat{\mathrm{St}}^{\rho}, \tag{5.93b}$$

where the energy densities $u_{\rho,g}$ and currents $\mathbf{j}_{g,\rho}^{\varepsilon}$ are defined as

$$u_g(t,\mathbf{r}) = \frac{1}{2}\mathrm{Tr}_{\mathbf{n}}\hat{\mathcal{L}}^{g}\hat{\mathcal{N}}^{g},$$
$$u_\rho(t,\mathbf{r}) = \frac{1}{2}\mathrm{Tr}_{\mathbf{n}}\left\{\frac{1}{1+\hat{F}};\,\hat{\mathcal{L}}^{\rho}\hat{\mathcal{N}}^{\rho}\right\}, \tag{5.93c}$$

$$\mathbf{j}_g^{\varepsilon} = \frac{1}{2}\mathrm{Tr}_{\mathbf{n}}\mathbf{v}\hat{\mathcal{L}}^{g}\hat{\mathcal{N}}^{g},\quad \mathbf{j}_\rho^{\varepsilon} = \frac{1}{2}\mathrm{Tr}_{\mathbf{n}}\{\hat{\mathbf{s}};\,\hat{\mathcal{L}}^{\rho}\hat{\mathcal{N}}^{\rho}\}. \tag{5.93d}$$

The velocity operator $\hat{\mathbf{s}}$ is defined in Eq. (5.65), and notation (5.60) is used.

We now add Eq. (5.93a) to Eq. (5.92) and subtract Eq. (5.93b). According to Eq. (5.87), all the collision terms and the commutators cancel, and we obtain the energy balance equation (compare with Eq. (2.8)):

$$\partial_t u_{\rm tot} + \nabla \cdot \mathbf{j}^{\varepsilon}_{\rm tot} = \mathbf{j} \cdot \mathbf{E}, \tag{5.94a}$$

$$u_{\rm tot}(t, \mathbf{r}) = u_e(t, \mathbf{r}) + u_\rho(t, \mathbf{r}) - u_g(t, \mathbf{r}), \tag{5.94b}$$

$$\mathbf{j}^{\varepsilon}_{\rm tot}(t, \mathbf{r}) = \mathbf{j}^{\varepsilon}_e(t, \mathbf{r}) + \mathbf{j}^{\varepsilon}_\rho(t, \mathbf{r}) - \mathbf{j}^{\varepsilon}_g(t, \mathbf{r}). \tag{5.94c}$$

Equations (5.94), (5.91), (5.93c), and (5.93d) constitute the main result in this subsection. They define the conserved currents in terms of quantities to be found from the kinetic equations. We emphasize that the conservation laws thus found are exact (at one loop) in the sense that no approximation has been made beyond the usual Fermi-liquid theory: specifically, no gradient or harmonic expansion has been made and no time or space Poisson brackets have been neglected yet (except those suppressed by the factor $q/p_{\rm F}$).

Within the same accuracy with which kinetic equations (5.64) were derived, performing Wigner transformations (5.61)–(5.63) of Eq. (5.90), (5.91), and (5.93d), we find

$$\mathbf{j} = e v_{\rm F} \nu \int d\varepsilon \langle \mathbf{n} f(\varepsilon, \mathbf{n}; t, \mathbf{r}) \rangle_{\mathbf{n}} \tag{5.95}$$

for the electric current density and

$$\mathbf{j}^{\varepsilon}_e = v_{\rm F} \nu \int d\varepsilon \varepsilon \langle \mathbf{n} f(\varepsilon, \mathbf{n}; t, \mathbf{r}) \rangle_{\mathbf{n}}, \tag{5.96a}$$

$$\mathbf{j}^{\varepsilon}_\rho = \int \frac{d\omega}{2\pi} \omega \int \frac{d^d q}{(2\pi)^d} \times \left\langle \{\hat{\mathbf{s}}(\omega, \mathbf{q}); \hat{\mathscr{L}}^\rho(\omega, \mathbf{q}) \hat{N}^\rho(\omega, \mathbf{q}; t, \mathbf{r})\} \right\rangle_{\mathbf{n}}, \tag{5.96b}$$

$$\mathbf{j}^{\varepsilon}_g = \int \frac{d\omega}{2\pi} \omega \int \frac{d^d q}{(2\pi)^d} \times \left\langle \mathbf{v} \hat{\mathscr{L}}^g(\omega, \mathbf{q}) \hat{N}^g(\omega, \mathbf{q}; t, \mathbf{r}) \right\rangle_{\mathbf{n}}, \tag{5.96c}$$

for the energy current densities, in agreement with Eqs. (3.14)–(3.15).

It remains to reduce the found expressions to the usual form of the quantum BE. This is the subject of the next subsection.

5.7. The Quantum Kinetic Equation

After Wigner transformations (5.61)–(5.63b), Eq. (5.71) becomes

$$\left[\partial_t + \mathbf{v} \cdot \nabla + e\mathbf{v} \cdot \mathbf{E} \frac{\partial}{\partial \varepsilon}\right] f(\varepsilon, \mathbf{n}; t, \mathbf{r}) = \hat{\rm St}\{f, N^\rho, N^g\}. \tag{5.97}$$

The collision integral is the sum of the inelastic and elastic parts,

$$\hat{\rm St}\{f, N^\rho, N^g\} = \hat{\rm St}_{\rm in}\{f, N^\rho, N^g\} + \hat{\rm St}_{\rm el}\{f\}. \tag{5.98}$$

The elastic part is obtained by adding the transform of the "bare" collision integral (the first term in the right-hand side of Eq. (5.72)) to the transforms of $\hat{\rm St}_1^{\rm el}$, Eq. (5.75c), and $\hat{\rm St}_-$, Eq. (5.76b). The inelastic part is given by the transform of $\hat{\rm St}_1^{in}$, Eq. (5.75b), plus the transform of $\hat{\rm St}_+$, Eq. (5.76c). However, for the sake of compactness, we do make no such distinction between elastic and inelastic contributions and, using a notation resembling that in Section 5.5, write the collision integral in the form

$$\hat{\rm St}\{f, N^\rho, N^g\} = \hat{\rm St}_\tau f + \hat{\rm St}_1\{f, N^\rho, N^g\} + \hat{\rm St}_-\{f\} + \hat{\rm St}_+\{f, N^\rho, N^g\}, \tag{5.99}$$

where the first term in the right-hand side is the transform of the bare collision integral and the other terms are given below.

With the elastic and inelastic parts of collision integral (5.73) kept in a single formula, the corresponding contribution is obtained by first substituting Eqs. (5.58) and then performing Wigner transformations (5.61)–(5.63) (and using their property (5.83)). We decompose the result into distinct contributions due to the two bosonic degrees of freedom:

$$\hat{\rm St}_1\{\varepsilon, \mathbf{n}; t, \mathbf{r}\} = \hat{\rm St}_1^{e-\rho}(\varepsilon, \mathbf{n}; t, \mathbf{r}) - \hat{\rm St}_1^{e-g}(\varepsilon, \mathbf{n}; t, \mathbf{r}). \tag{5.100}$$

As usual, the collisions with the ghost particles enter with the opposite signs. In terms of combination (5.66) of distribution functions denoted by Υ and vertex (5.28), these contributions are (we suppress the spectator arguments t and $\mathbf{r}$ in both sides of the equations)

$$\hat{\rm St}_1^{e-\rho}(\varepsilon, \mathbf{n}_1) = \frac{1}{\nu} \int \frac{d\omega}{2\pi} \frac{1}{\omega} \int \frac{d^d q}{(2\pi)^d} \int \frac{d\mathbf{n}_2 d\mathbf{n}_3 d\mathbf{n}_4}{\Omega_d^3} \times \{\gamma_{12}^3 [\mathscr{L}^\rho_{34} \Upsilon^\rho_{41;\,21}(\varepsilon, \omega, \mathbf{q}) + \Upsilon^\rho_{34;\,21}(\varepsilon, \omega, \mathbf{q}) \overline{\mathscr{L}}^\rho_{41}] + \gamma_{21}^3 [\mathscr{L}^\rho_{34} \Upsilon^\rho_{42;\,21}(\varepsilon, \omega, \mathbf{q}) + \Upsilon^\rho_{34;\,21}(\varepsilon, \omega, \mathbf{q}) \overline{\mathscr{L}}^\rho_{42}]\}, \tag{5.101a}$$

$$\hat{\rm St}_1^{e-g}(\varepsilon, \mathbf{n}_1) = \frac{1}{\nu} \int \frac{d\omega}{2\pi} \frac{1}{\omega} \int \frac{d^d q}{(2\pi)^d} \int \frac{d\mathbf{n}_2 d\mathbf{n}_3 d\mathbf{n}_4}{\Omega_d^3} \times \{\gamma_{12}^3 [\mathscr{L}^g_{34} \Upsilon^g_{41;\,21}(\varepsilon, \omega, \mathbf{q}) + \Upsilon^g_{34;\,21}(\varepsilon, \omega, \mathbf{q}) \overline{\mathscr{L}}^g_{41}] + \gamma_{21}^3 [\mathscr{L}^\rho_{34} \Upsilon^g_{42;\,21}(\varepsilon, \omega, \mathbf{q}) + \Upsilon^g_{34;\,21}(\varepsilon, \omega, \mathbf{q}) \overline{\mathscr{L}}^g_{42}]\}. \tag{5.101b}$$

Here and below, the short notation

$$\mathscr{L}^\alpha_{ij} = \mathscr{L}^\alpha(\omega, \mathbf{q}; \mathbf{n}_i, \mathbf{n}_j), \quad \alpha = g, \rho$$

is used. It is readily seen that these contributions coincide with the local electron–boson collision integral in Eq. (3.19b).

Proceeding as above, we obtain the transform of Eq. (5.76b) as

$$\hat{\mathrm{St}}_-(\varepsilon, \mathbf{n}_1) = \hat{\mathrm{St}}_{el}(\varepsilon, \mathbf{n}_1) + \hat{\mathrm{St}}_{-,I}(\varepsilon, \mathbf{n}_1), \tag{5.102a}$$

$$\begin{aligned}\hat{\mathrm{St}}_{el} &= \frac{2}{\nu}\mathrm{Re}\int\frac{d\omega}{2\pi}\frac{1}{\omega}\int\frac{d^dq}{(2\pi)^d}\int\frac{d\mathbf{n}_2\ldots d\mathbf{n}_6}{\Omega_d^5}\gamma_{13}^2\gamma_{46}^5\\ &\times[\mathscr{L}^\rho - \mathscr{L}^g]_{52}[f(\varepsilon-\omega,\mathbf{n}_6) - f(\varepsilon-\omega,\mathbf{n}_4)]\\ &\times[\mathscr{L}^g_{14}f(\varepsilon,\mathbf{n}_3) + \mathscr{L}^g_{34}f(\varepsilon,\mathbf{n}_1)],\end{aligned} \tag{5.102b}$$

$$\begin{aligned}\hat{\mathrm{St}}_{-,I} &= \frac{1}{\nu}\int\frac{d\omega}{2\pi}\frac{1}{\omega^2}\int\frac{d^dq}{(2\pi)^d}\int\frac{d\mathbf{n}_2\ldots d\mathbf{n}_6}{\Omega_d^5}\gamma_{13}^2\gamma_{46}^5\\ &\times[2f(\varepsilon,\mathbf{n}_4) - f(\varepsilon-\omega,\mathbf{n}_4)\\ &- f(\varepsilon+\omega,\mathbf{n}_4) - (\mathbf{n}_4 \longrightarrow \mathbf{n}_6)]\\ &\times\int d\varepsilon_1[f(\varepsilon_1,\mathbf{n}_1)[1-f(\varepsilon_1-\omega,\mathbf{n}_3)] + (\mathbf{n}_1 \rightleftharpoons \mathbf{n}_3)]\\ &\times[\mathscr{L}^\rho - \mathscr{L}^g]_{52}\mathscr{L}^g_{14} + \frac{1}{2}[\mathscr{L}^\rho - \mathscr{L}^g]_{52}[\mathscr{L}^g_{14} - \mathscr{L}^g_{34}].\end{aligned} \tag{5.102c}$$

Equation (5.102b) is (the singlet part of) the elastic electron–electron collision integral, Eq. (3.20b). To obtain Eq. (5.102c) in the given form, we used the analytic properties of the propagators and changed the variable as $\varepsilon_1 \longrightarrow \varepsilon_1 + \omega$ in some of the terms.

We finally transform Eq. (5.76c) and obtain

$$\begin{aligned}\hat{\mathrm{St}}_+(\varepsilon,\mathbf{n}_1) &= \hat{\mathrm{St}}_n^{e-e}(\varepsilon,\mathbf{n}_1)\\ &+ \hat{\mathrm{St}}_{+,n}(\varepsilon,\mathbf{n}_1) + \hat{\mathrm{St}}_{+,I}(\varepsilon,\mathbf{n}_1).\end{aligned} \tag{5.103a}$$

The first term is given by (the singlet part of) Eq. (3.20d). The second term is

$$\hat{\mathrm{St}}_{+,n} = \hat{\mathrm{St}}_{+,n}^{e-\rho} - \hat{\mathrm{St}}_{+,n}^{e-g} \tag{5.103b}$$

with $\hat{\mathrm{St}}_{+,n}^{e-\alpha}$ given by Eq. (3.19d) excluding the last line. The third term is

$$\begin{aligned}\hat{\mathrm{St}}_{+,I} &= \frac{1}{2\nu}\int\frac{d\omega}{2\pi}\frac{1}{\omega}\int\frac{d^dq}{(2\pi)^d}\int\frac{d\mathbf{n}_2\ldots d\mathbf{n}_6}{\Omega_d^5}\gamma_{13}^2\gamma_{46}^5\\ &\times[2f(\varepsilon,\mathbf{n}_4) - f(\varepsilon-\omega,\mathbf{n}_4)\\ &- f(\varepsilon+\omega,\mathbf{n}_4) - (\mathbf{n}_4 \longrightarrow \mathbf{n}_6)]\\ &\times[\mathscr{L}^\rho(2N^\rho+1) - \mathscr{L}^\rho(2N^g+1)]_{52}[\mathscr{L}^g_{34} - \mathscr{L}^g_{14}].\end{aligned} \tag{5.103c}$$

Adding Eqs. (5.103c) and (5.102c), we recover the last line in Eq. (3.19d) and (the singlet part of) the local electron–electron collision integral, Eq. (3.20f). This concludes the derivation of the quantum kinetic equation for the singlet channel.

5.8. The Triplet Channel

Inclusion of the interaction in the triplet channel is straightforward; in Eilenberger Eq. (5.7), we add the term

$$i[\hat{\boldsymbol{\phi}}\cdot\boldsymbol{\sigma};\hat{g}] \tag{5.104}$$

to the left-hand side, where σ_i are the Pauli matrices ($i = x, y, z$) and the fluctuating field $\hat{\phi}_i$ is a 3-component vector in the $L = 1$ angular momentum space. Therefore, all the triplet channel propagators, polarization operators, and density matrices should be considered 3×3 matrices; for example, we have

$$[D_0]_{ij}(1,2) = -\frac{F^\sigma(\theta_{12})\delta(\mathbf{r}_{12})}{\nu}\delta(t_{12})\delta_{ij}$$

(cf. Eq. (5.15)) and the retarded polarization operator is given by

$$\Pi^R_{ij}(1,2) = \nu\left\{\delta_{12} + \frac{\pi}{4}\mathrm{Tr}\left[\sigma_i\frac{\delta g^K(t_1,t_1,\mathbf{n}_1,\mathbf{r}_1)}{\delta\phi_{+j}(t_2,\mathbf{r}_2,\mathbf{n}_2)}\right]\right\},$$

where the trace is over spin indices.

The trace of triplet channel operators includes the sum over the indices i, j. In the absence of the magnetic field, all the operators are diagonal, e.g.,

$$[\mathscr{L}^\sigma]_{ij} = \mathscr{L}^\sigma\delta_{ij},$$

and the trace results in extra factors of 3 in comparison to the singlet channel. The derivation can therefore be repeated with simple modifications and it gives the quantum kinetic equation presented in Section 3. We only note one main difference in the derivation for the triplet channel: the gauge transformation, which has the form

$$\hat{g} \longrightarrow e^{-i\hat{\mathbf{K}}\cdot\boldsymbol{\sigma}}\hat{g}e^{i\hat{\mathbf{K}}\cdot\boldsymbol{\sigma}},$$

does not commute with interaction term (5.104). Additional second-order terms arise due to commutators of the Pauli matrices; however, these terms vanish in the one-loop approximation and we can neglect them.

In the next section, we use the quantum kinetic equation to calculate the interaction corrections to the transport coefficients and specific heat.

6. DERIVATION OF TRANSPORT COEFFICIENTS AND SPECIFIC HEAT

In this section, we calculate the transport coefficients for quasi one-dimensional, two-dimensional, and three-dimensional systems; the evaluation of the interaction correction to the specific heat is in the final subsection.

To calculate the currents in the presence of an external field (electric field **E** or temperature gradient ∇T), we need to solve the kinetic equations. We assume that the external fields are weak, i.e.,

$$eEL_T \ll T, \quad \nabla T L_T \ll T,$$

with the temperature length defined in Eq. (5.53). These conditions ensure that the deviations from the equilibrium distribution functions are small and we can solve the equations by iteration.

In the lowest order, the distribution functions should turn the collision integrals in Eqs. (3.4) and (3.6) to zero, and the sought corrections δf, δN^α are linear in the electric field or in the gradients of the distribution functions. In other words, we seek a solution of the kinetic equations of the form

$$\begin{aligned} f(\varepsilon, \mathbf{n}; \mathbf{r}) &= f_F(\varepsilon; \mathbf{r}) + \delta f(\varepsilon, \mathbf{n}; \mathbf{r}), \\ \hat{N}^\alpha(\omega, \mathbf{q}; \mathbf{r}) &= N_P(\omega; \mathbf{r})\hat{\mathbb{1}} + \delta\hat{N}^\alpha(\omega, \mathbf{q}; \mathbf{r}), \end{aligned} \tag{6.1}$$

where the Fermi and Planck distribution functions (2.2) depend on the spatial coordinate only through the temperature $T(\mathbf{r})$. For compactness, we consider only the singlet channel explicitly and indicate how to include the triplet channel.

We start from the electron part of the kinetic equation. The bare impurity collision part $\hat{\mathrm{St}}_\tau$ is larger than the other terms, and it therefore suffices to calculate the latter in the first order of perturbation theory. Considering short-range impurities, such that τ is independent of the scattering angle, we find

$$\begin{aligned} \delta f &= \delta f_0 + \delta f_1, \\ \delta f_0 &= \tau \mathbf{v} \cdot \left(e\mathbf{E} - \frac{\varepsilon \nabla T}{T} \right) \left(-\frac{\partial f_F(\varepsilon)}{\partial \varepsilon} \right), \\ \delta f_1 &= \tau \delta\hat{\mathrm{St}}\{ f_F + \delta f_0, N_P + \delta N^p, N_P + \delta N^g \}, \end{aligned} \tag{6.2}$$

where $\delta\hat{\mathrm{St}}$ is the linearized collision integral. We note that according to the discussion of the conservation laws in Section 3, we need to consider only the local electron–boson contribution, Eq. (3.19b), and the elastic electron–electron one, Eq. (3.20b).

Expression (6.2) is to be substituted in Eqs. (3.14) and (3.15b) to find the electric current and the electron component of the energy current. Integration of the δf_0 term is straightforward. Due to the structure of collision integrals (3.19)–(3.20), the integration over ε can be performed before the ω and $\mathbf{q}$ integrations in the δf_1 term. For the combination of distribution functions entering Eq. (3.16a), we find

$$\begin{aligned} \delta\Upsilon_{ij,kl}(\varepsilon, \omega) &= \delta N(\omega, \mathbf{q}; \mathbf{n}_i, \mathbf{n}_j)\{ f_F(\varepsilon) - f_F(\varepsilon - \omega) \} \\ &+ \Omega_d \delta(\widehat{\mathbf{n}_i \mathbf{n}_j}) N_P(\omega)\{ \delta f(\varepsilon, \mathbf{n}_k) - \delta f(\varepsilon - \omega, \mathbf{n}_k) \} \\ &+ \Omega_d \delta(\widehat{\mathbf{n}_i \mathbf{n}_j})[\delta f(\varepsilon, \mathbf{n}_l)(1 - f_F(\varepsilon - \omega)) \\ &- f_F(\varepsilon)\delta f(\varepsilon - \omega, \mathbf{n}_k)], \end{aligned} \tag{6.3}$$

where we next use the identities

$$\begin{aligned} \int d\varepsilon\, f_F(\varepsilon) \frac{\partial f_F(\varepsilon - \omega)}{\partial \varepsilon} &= \frac{\partial}{\partial \omega}[\omega N_P(\omega)], \\ \int d\varepsilon \frac{\partial f_F(\varepsilon)}{\partial \varepsilon}[1 - f_F(\varepsilon - \omega)] &= \frac{\partial}{\partial \omega}[\omega N_P(\omega)], \end{aligned} \tag{6.4a}$$

$$\begin{aligned} \int d\varepsilon \varepsilon \frac{\partial f_F(\varepsilon)}{\partial \varepsilon}[1 - f_F(\varepsilon - \omega)] &= \frac{1}{2}\omega^2 \frac{\partial N_P(\omega)}{\partial \omega}, \\ \int d\varepsilon (\varepsilon - \omega) \frac{\partial f_F(\varepsilon - \omega)}{\partial \varepsilon} f_F(\varepsilon) &= -\frac{1}{2}\omega^2 \frac{\partial N_P(\omega)}{\partial \omega}, \end{aligned}$$

$$\begin{aligned} &\int d\varepsilon \varepsilon^2 \frac{\partial f_F(\varepsilon)}{\partial \varepsilon}[1 - f_F(\varepsilon - \omega)] \\ &= \frac{\pi^2}{3} T^2 \frac{\partial}{\partial \omega}[\omega N_P(\omega)] + \frac{1}{3}\omega^3 \frac{\partial N_P(\omega)}{\partial \omega}, \\ &\int d\varepsilon \varepsilon (\varepsilon - \omega) f_F(\varepsilon) \frac{\partial f_F(\varepsilon - \omega)}{\partial \varepsilon} \\ &= \frac{\pi^2}{3} T^2 \frac{\partial}{\partial \omega}[\omega N_P(\omega)] - \frac{1}{6}\omega^3 \frac{\partial N_P(\omega)}{\partial \omega} \end{aligned} \tag{6.4b}$$

to obtain

$$\begin{aligned} &\frac{1}{2}\int d\varepsilon [\delta\Upsilon_{ij,kl}(\varepsilon, \omega, \mathbf{q}) - \delta\Upsilon_{ij,kl}(\varepsilon, -\omega, -\mathbf{q})] \\ &= e v_F \tau \mathbf{E} \cdot (\mathbf{n}_k - \mathbf{n}_l) \frac{\partial}{\partial \omega}(\omega N_P^{ij}), \end{aligned} \tag{6.5a}$$

$$\begin{aligned} &\frac{1}{2}\int d\varepsilon \varepsilon [\delta\Upsilon_{ij,kl}(\varepsilon, \omega, \mathbf{q}) - \delta\Upsilon_{ij,kl}(\varepsilon, -\omega, -\mathbf{q})] \\ &= -\frac{\omega^2}{2}\delta N(\omega, \mathbf{q}; \mathbf{n}_i, \mathbf{n}_j) - \frac{v_F \tau \nabla T}{4T} \cdot (\mathbf{n}_l + \mathbf{n}_k) \\ &\times \omega^3 \frac{\partial N_P^{ij}}{\partial \omega} + \frac{v_F \tau \nabla T}{T} \cdot (\mathbf{n}_l - \mathbf{n}_k) \\ &\times \left[\frac{\pi^2 T^2}{3} \frac{\partial(\omega N_P^{ij})}{\partial \omega} + \frac{\omega^3}{12} \frac{\partial N_P^{ij}}{\partial \omega} \right], \end{aligned} \tag{6.5b}$$

where

$$N_P^{ij} \equiv [\hat{N}_P](\mathbf{n}_i, \mathbf{n}_j) \equiv N_P(\omega)\Omega_d \delta(\widehat{\mathbf{n}_i \mathbf{n}_j}). \tag{6.6}$$

We here retained only the contribution odd in ω because the even part vanishes after the ω integration in the relevant collision integral, see Eq. (3.19b).

The combination of the distribution functions entering the elastic collision part, Eq. (3.20b), gives

$$\frac{1}{2}\int d\varepsilon\, f_F(\varepsilon)[\delta f(\varepsilon-\omega,\mathbf{n})-\delta f(\varepsilon+\omega,\mathbf{n})] = -e v_F \tau \mathbf{E}\cdot\mathbf{n}\frac{\partial}{\partial\omega}[\omega N_P(\omega)], \tag{6.7a}$$

$$\frac{1}{2}\int d\varepsilon\varepsilon\, f_F(\varepsilon)[\delta f(\varepsilon-\omega,\mathbf{n})-\delta f(\varepsilon+\omega,\mathbf{n})] = \frac{v_F\tau\mathbf{n}\cdot\nabla T}{T} \times\left[\frac{\pi^2T^2}{3}\frac{\partial(\omega N_P(\omega))}{\partial\omega}-\frac{\omega^3}{6}\frac{\partial N_P(\omega)}{\partial\omega}\right], \tag{6.7b}$$

where we again retained only the part that is odd in ω, nonvanishing after the ω integration in Eq. (3.20b).

Using Eq. (3.14), we find the electric current $\mathbf{j}=\hat\sigma\mathbf{E}$ with the conductivity tensor $\hat\sigma$ given by

$$\hat\sigma=\sigma_D\left\{\hat{\mathbb{1}}+\int d\omega[\hat{\mathcal{S}}^{\rm el}(\omega)+\hat{\mathcal{E}}(\omega)]\frac{\partial}{\partial\omega}[\omega N_P(\omega)]\right\}, \tag{6.8}$$

where $\sigma_D=\tau v_F^2e^2\nu/d$ is the Drude conductivity. With the spatial indices denoted by $\mu,\nu=1,\ldots,d$, the elastic kernels $\mathcal{S}^{\rm el}_{\mu\nu}$ and $\mathcal{E}_{\mu\nu}$, which originate from Eq. (3.19b) and Eq. (3.20b) respectively, are given by

$$\mathcal{S}^{\rm el}_{\mu\nu}(\omega)=\mathcal{S}^{11}_{\mu\nu}(\omega)+\mathcal{S}^{12}_{\mu\nu}(\omega),$$

$$\mathcal{S}^{11}_{\mu\nu}(\omega)=\frac{d}{\pi\omega\nu}\int\frac{d^dq}{(2\pi)^d}\int\frac{d\mathbf{n}_1d\mathbf{n}_2}{(\Omega_d)^2}n_{1\mu}n_{1\nu} \times{\rm Re}[\mathcal{L}^\rho_{11}-\mathcal{L}^g_{11}]+{\rm Re}[\mathcal{L}^\rho_{22}-\mathcal{L}^g_{22}]-2{\rm Re}[\mathcal{L}^\rho_{12}-\mathcal{L}^g_{12}],$$

$$\mathcal{S}^{12}_{\mu\nu}(\omega)=\frac{2d}{\pi\omega\nu}\int\frac{d^dq}{(2\pi)^d} \times\int\frac{d\mathbf{n}_1d\mathbf{n}_2}{(\Omega_d)^2}n_{1\mu}n_{2\nu}{\rm Re}[\mathcal{L}^\rho_{12}-\mathcal{L}^g_{12}], \tag{6.9}$$

$$\mathcal{E}_{\mu\nu}(\omega)=-\frac{d\tau}{\pi\omega\nu}\int\frac{d^dq}{(2\pi)^d}\int\frac{d\mathbf{n}_1\ldots d\mathbf{n}_6}{(\Omega_d)^6}\gamma^6_{12}\gamma^5_{43} \times(n_{1\mu}n_{3\nu}-n_{2\mu}n_{3\nu}){\rm Re}\{[\mathcal{L}^\rho_{56}-\mathcal{L}^g_{56}][\mathcal{L}^g_{13}+\mathcal{L}^g_{14}]\},$$

where we keep only the singlet channel correction for compactness; inclusion of the triplet channel contribution is straightforward.[9] We show in Appendix F that our expression for the conductivity coincides with the one in [17]. It is natural that the conductivity does not involve any bosonic distribution function (cf. Eq. (6.5a)), because the inelastic electron collision with such bosons changes the energy of the electron but not the direction of its motion.

In contrast, even the electron contribution to the thermal conductivity tensor $\hat\kappa$, such that $\mathbf{j}^\varepsilon_{\rm tot}=-\hat\kappa\nabla T$, is sensitive to the bosonic distribution functions. We represent the total thermal conductivity as

$$\hat\kappa=\hat\kappa_{\rm WF}+\delta\hat\kappa+\hat\kappa^\rho-\hat\kappa^g. \tag{6.10}$$

The first term in this expression obeys the Wiedemann–Franz law with the interaction corrections to the electric conductivity included, i.e., $\hat\kappa_{\rm WF}=L\hat\sigma T$, with the Lorentz number given by Eq. (1.1). The second term represents the (electronic) correction to the Wiedemann–Franz law due to the energy dependence of the elastic scattering and due to the inelastic electron scattering on bosons. Finally, the third and the fourth terms represent the contribution of the ρ and g bosons to the thermal transport. These additional contributions are given by

$$\delta\hat\kappa=\delta\hat\kappa_{\rm el}+\delta\hat\kappa_{\rm in}, \tag{6.11a}$$

$$[\delta\kappa_{\rm el}]_{\mu\nu}=\frac{\sigma_D}{e^2T} \times\int d\omega[\mathcal{S}^{\rm el}_{\mu\nu}(\omega)-2\mathcal{E}_{\mu\nu}(\omega)]\left[\frac{\omega^3}{12}\frac{\partial N_P}{\partial\omega}\right], \tag{6.11b}$$

$$[\delta\kappa_{\rm in}]_{\mu\nu}=\frac{\sigma_D}{e^2T}\int d\omega[\mathcal{S}^{12}_{\mu\nu}(\omega)-\mathcal{S}^{11}_{\mu\nu}(\omega)]\left[\frac{\omega^3}{4}\frac{\partial N_P}{\partial\omega}\right] +v_F\int\frac{d\omega}{2\pi}\omega\int\frac{d^dq}{(2\pi)^d}\int\frac{d\mathbf{n}_1d\mathbf{n}_2d\mathbf{n}_3}{(\Omega_d)^3}n_{1\mu} \times\{{\rm Re}[\mathcal{L}^\rho_{12}\delta_\nu N^\rho_{23}+\mathcal{L}^\rho_{32}\delta_\nu N^\rho_{21}-\mathcal{L}^\rho_{12}\delta_\nu N^\rho_{21}] -{\rm Re}[\mathcal{L}^g_{12}\delta_\nu N^g_{23}+\mathcal{L}^g_{23}\delta_\nu N^g_{21}-\mathcal{L}^g_{12}\delta_\nu N^g_{21}]\}, \tag{6.11c}$$

$$\kappa^\alpha_{\mu\nu}=-\int\frac{d\omega}{2\pi}\omega\int\frac{d^dq}{(2\pi)^d}\int\frac{d\mathbf{n}_1d\mathbf{n}_2}{(\Omega_d)^2}\{\hat s^\alpha_\mu;\,\mathcal{L}^\alpha_{12}\delta_\nu N^\alpha_{21}\}, \tag{6.12}$$

where

$$\delta_\mu N^\alpha_{ij}=\frac{\delta}{\delta(\nabla_\mu T)}[\delta N^\alpha(\omega,\mathbf{q};\mathbf{n}_i,\mathbf{n}_j)]. \tag{6.13}$$

Equations (6.8)–(6.12) are the complete expressions for the electric and thermal transport coefficients. To obtain an explicit result, we must solve Eqs. (3.6) to find the distribution functions δN^α. We do this by

[9] By the simple substitution $\mathcal{L}^\rho-\mathcal{L}^g\longrightarrow\mathcal{L}^\rho+3\mathcal{L}^\sigma-4\mathcal{L}^g$ in the kernels.

restricting ourselves to the diffusive $T\tau \ll 1$ regime, except for two-dimensional systems for which we consider an arbitrary temperature range.[10] Moreover, for the Coulomb interaction, we consider the unitary limit (for infrared-finite momentum integrals), which enables us to drop all the terms that depend on $\partial_{\mathbf{q}}\hat{\mathbf{F}}$.

From now on, we retain only the zeroth harmonic of the Fermi-liquid constants, for which we use the notation F_0^α. For the singlet channel, this means

$$F_0^\rho \equiv \int \frac{d\theta}{2\pi} F^\rho(\theta) + \nu V(\mathbf{q}) \tag{6.14a}$$

and for the triplet channel,

$$F_0^\sigma \equiv \int \frac{d\theta}{2\pi} F^\sigma(\theta). \tag{6.14b}$$

We recall that $F^g = 0$ and that the Coulomb interaction potential is given by

$$V(\mathbf{q}) = \begin{cases} \dfrac{4\pi e^2}{q^2}, & d = 3 \\ \dfrac{2\pi e^2}{q}, & d = 2 \\ e^2 \ln \dfrac{1}{(qa)^2}, & d = 1, \end{cases} \tag{6.14c}$$

where a is a length of the order of the quasi one-dimensional wire width.

6.1. Diffusive Regime

We first consider the distribution function N^g; substituting expression (6.1) in Eq. (3.6), in the linear order in ∇T, we obtain

$$-\frac{\omega}{T}\mathbf{v}_1 \cdot \nabla T \frac{\partial N_P^{12}(\omega)}{\partial \omega} + i[\mathbf{v}\cdot\mathbf{q};\, \delta N^g] = 2\{\hat{\mathrm{St}}_\tau;\, \delta N^g\} + \nu_F \frac{\omega}{T}\frac{\partial N_P(\omega)}{\partial \omega}\nabla T \cdot (\Omega_d \delta(\widehat{\mathbf{n}_1, \mathbf{n}_2})\mathbf{n}_1 - \mathbf{n}_1 - \mathbf{n}_2), \tag{6.15}$$

with N_P^{ij} defined in Eq. (6.6). The (exact at this order) solution for $\delta N^g(\omega; \mathbf{n}_1, \mathbf{n}_2)$ is

$$\delta N^g = \delta N^0 \equiv \nu_F \tau \mathbf{n}_1 \cdot \nabla T \frac{\omega}{T}\frac{\partial N_P^{12}(\omega)}{\partial \omega}. \tag{6.16}$$

[10] The Boltzmann equation description of strictly one-dimensional systems is not applicable and considering the quasi-one-dimensional ballistic case within our scheme is meaningless because of the effects of boundary scattering. The ballistic regime in three dimensions also cannot be considered within our scheme because the main effect on the thermal conductivity is due to the inelastic scattering processes with momentum transfer of the order of k_F.

For the distribution function N^ρ, the above is only the starting point for the iterative solution:

$$\delta N^\rho = \delta N^0 + \delta N^1. \tag{6.17}$$

The equation for δN^1 is

$$\left(-\frac{\omega}{T}\frac{\partial \hat{N}_P(\omega)}{\partial \omega}\nabla T\right)\left(\omega \frac{\partial}{\partial \mathbf{q}}\frac{\hat{F}^\rho}{1+\hat{F}^\rho}\right) + i\left[\omega \frac{\hat{F}^\rho}{1+\hat{F}^\rho};\, \delta\hat{N}^g\right] + i[\hat{H}_{e-h};\, \delta\hat{N}^1] = 2\{\hat{\mathrm{St}}_\tau;\, \delta\hat{N}^1\}. \tag{6.18}$$

In the diffusive limit $T\tau \ll 1$, the (first iteration) solution would be of the form

$$\delta N^1 \approx \omega\tau\left(v_F \tau \frac{\omega}{T}\frac{\partial N_P(\omega)}{\partial \omega}\nabla T\right)\cdot \mathbf{V}$$

for a vector $\mathbf{V}$ with a magnitude on the order of one. However, contributions from frequencies ω larger than the temperature T are exponentially suppressed, i.e., $\omega\tau \lesssim T\tau \ll 1$; therefore, δN^1 can be neglected in comparison to δN^g. Thus, in the diffusive limit,

$$\delta_\mu N_{ij}^\alpha = v_F \tau [\mathbf{n}_{i\mu}] \frac{\omega}{T}\frac{\partial N_P^{ij}(\omega)}{\partial \omega}. \tag{6.19}$$

For the propagators $\mathcal{L}^\alpha$ in Eq. (3.11), the diffusive approximation amounts to the substitution

$$n_\mu n_\nu \longrightarrow \frac{\delta_{\mu\nu}}{d},$$

which leads to the expression

$$\mathcal{L}^\alpha(\mathbf{n}_1, \mathbf{n}_2) = \tau(\delta(\widehat{\mathbf{n}_1, \mathbf{n}_2}) - 1) + L_0^\alpha + (n_1 + n_2)_\mu L_{1\mu}^\alpha + n_{1\mu} n_{2\nu} L_{2\mu\nu}^\alpha, \tag{6.20}$$

where the functions L_i^α depend on ω, $\mathbf{q}$ only and are explicitly given by

$$L_0^\alpha = \frac{1}{\dfrac{-i\omega}{1+F_0^\alpha} + Dq^2}, \tag{6.21a}$$

$$L_{1\mu}^\alpha = -i\tau v_F q_\mu L_0^\alpha, \tag{6.21b}$$

$$L_{2\mu\nu}^\alpha = -d\tau D q_\mu q_\nu L_0^\alpha. \tag{6.21c}$$

Here, $D = \tau v_F^2/d$ is the diffusion constant. These formulas are valid whenever $\omega, Dq^2 \ll 1/\tau$.

Within this approximation, we have

$$\mathcal{S}_{\mu\nu}^{12}(\omega) \pm \mathcal{S}_{\mu\nu}^{11}(\omega) \propto L_{2\mu\nu}^\rho - L_{2\mu\nu}^g.$$

This means that in both $\hat{\sigma}$ and $\delta\hat{\kappa}$ (see Eqs. (6.8) and (6.11)), we can neglect the contributions of the $\hat{\mathcal{S}}$ kernels (we note that by inserting solution (6.19) in Eq. (6.11c), $\delta\hat{\kappa}_{in}$ is given by twice the first line of that equation). Indeed, the leading contribution is given by the kernel $\hat{\mathcal{E}}$ in Eq. (6.11b). This kernel has the approximate form

$$\mathcal{E}_{\mu\nu}(\omega) = \frac{4}{\pi\nu\tau}\frac{1}{d\omega} \times \int\frac{d^d q}{(2\pi)^d}\mathrm{Re}[L_0^g(L_{2\mu\nu}^\rho - L_{2\mu\nu}^g)]. \tag{6.22}$$

Finally, the bosonic contributions (cf. Eq. (6.12)) to thermal conductivity (6.10) can be written as[11]

$$\hat{\kappa}^\rho - \hat{\kappa}^g = \frac{\sigma_D}{e^2 T}\int d\omega\hat{\mathcal{B}}(\omega)\left[\frac{\omega^3}{4}\frac{\partial N_P}{\partial\omega}\right] \tag{6.23}$$

with

$$\mathcal{B}_{\mu\nu}(\omega) = -\frac{2\delta_{\mu\nu}}{\pi\omega\nu}\int\frac{d^d q}{(2\pi)^d}\mathrm{Re}[L_0^\rho - L_0^g]. \tag{6.24}$$

The next step is to evaluate the momentum integrals; we first give the results for the short-range interaction described by (the zeroth harmonic of) the Fermi-liquid constant F_0, and then we indicate the modifications needed to account for the long-range part of the Coulomb interaction in the singlet channel. The triplet channel contributions are obtained by multiplying the obtained results by three and identifying F_0 with F_0^σ.

For the elastic kernel, we find

$$\mathcal{E}_{\mu\nu}(\omega) = \frac{\delta_{\mu\nu}}{d}\frac{e^2}{\sigma_D}\frac{\Omega_d}{(2\pi)^d}\left(\frac{|\omega|}{D}\right)^{d/2-1}\frac{1}{\omega}\frac{1}{\cos\dfrac{\pi d}{4}} \times\left[\frac{d}{2} - \frac{1}{F_0}(1+F_0-(1+F_0)^{1-d/2})\right]. \tag{6.25}$$

The expression for the Coulomb interaction is obtained by taking the unitary limit $F_0 \longrightarrow +\infty$.

The result for the bosonic kernel is

$$\mathcal{B}_{\mu\nu}(\omega) = \frac{\delta_{\mu\nu}}{2}\frac{e^2}{\sigma_D}\frac{\Omega_d}{(2\pi)^d}\left(\frac{|\omega|}{D}\right)^{d/2-1}\frac{1}{\omega} \times\frac{1}{\cos\dfrac{\pi d}{4}}[1-(1+F_0)^{1-d/2}]. \tag{6.26}$$

For $d = 3$, the limit $F_0 \longrightarrow +\infty$ gives the correct formula for the long-range contribution, but for $d = 1$, the limit diverges. However, with the full form of the interaction potential retained, this infrared divergence is cut off at the inverse screening radius. With logarithmic accuracy, the result for the Coulomb interaction is found by substituting $[\ldots] \longrightarrow -ak\ln^{1/2}(Dk^2/|\omega|)$, where a is a length on the order of the wire width and $k^2 = 4\pi e^2\nu$ is the square of the inverse screening radius (in the bulk).

We can now proceed with the calculation of the integrals over ω in Eqs. (6.8), (6.11), and (6.23) using the identity

$$\int d\omega\omega^m\partial_m N_P(\omega) = -2T^m\zeta(m)\Gamma(m+1). \tag{6.27}$$

Here, $\zeta(x)$ is the Riemann zeta function, whose values at the points relevant for our discussion are

$$\zeta\left(-\frac{1}{2}\right)\approx -0.208, \quad \zeta(0) = -\frac{1}{2}, \quad \zeta\left(\frac{1}{2}\right)\approx -1.460,$$

$$\zeta\left(\frac{3}{2}\right)\approx 2.612, \quad \zeta(2) = \frac{\pi^2}{6}\approx 1.645, \quad \zeta\left(\frac{5}{2}\right)\approx 1.341,$$

and $\Gamma(x)$ is the Euler gamma function, with the values

$$\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}, \quad \Gamma(1) = 1, \quad \Gamma\left(\frac{3}{2}\right) = \frac{1}{2}\sqrt{\pi},$$

$$\Gamma\left(\frac{5}{2}\right) = \frac{3}{4}\sqrt{\pi}, \quad \Gamma(3) = 2, \quad \Gamma\left(\frac{7}{2}\right) = \frac{15}{8}\sqrt{\pi}.$$

Performing the final ω integrations, we obtain

$$\sigma = \sigma_D + e^2\frac{\Omega_d}{(2\pi)^d}\frac{2}{d}\zeta\left(\frac{d}{2}-1\right)\left(\frac{T}{D}\right)^{d/2-1}\Gamma\left(\frac{d}{2}\right)\frac{d-4}{2-d} \times\frac{1}{\cos\dfrac{\pi d}{4}}\left[\frac{d}{2} - \frac{1}{F_0}(1+F_0-(1+F_0)^{1-d/2})\right], \tag{6.28}$$

$$\delta\kappa = \frac{1}{3d}\frac{\Omega_d}{(2\pi)^d}\left(\frac{T}{D}\right)^{d/2} D\zeta\left(\frac{d}{2}+1\right)\Gamma\left(\frac{d}{2}+2\right) \times\frac{1}{\cos\dfrac{\pi d}{4}}\left[\frac{d}{2} - \frac{1}{F_0}(1+F_0-(1+F_0)^{1-d/2})\right], \tag{6.29}$$

$$\kappa^\rho - \kappa^g = -\frac{1}{4}\frac{\Omega_d}{(2\pi)^d}\left(\frac{T}{D}\right)^{d/2} D\zeta\left(\frac{d}{2}+1\right) \times\Gamma\left(\frac{d}{2}+2\right)\frac{1}{\cos\dfrac{\pi d}{4}}[1-(1+F_0)^{1-d/2}], \tag{6.30}$$

where in the absence of magnetic field $\sigma_{\mu\nu} = \sigma\delta_{\mu\nu}$ and a similar relation holds for the thermal conductivity.

[11] We choose to collect the bosonic contributions into a single kernel such that the resulting momentum integral is convergent.

According to the previous discussion, for the Coulomb interaction, the correct expressions are given by the limit $F_0 \longrightarrow +\infty$, with the exception of the term $\kappa^\rho - \kappa^g$ in the case where $d = 1$ for which, with logarithmic accuracy, the result is obtained by substituting $[\ldots] \longrightarrow -ak\ln^{1/2}(Dk^2/T)$. The final answers for the corrections to the thermal conductivity are given in Eqs. (4.2)–(4.3).

We note that for $d = 3$, the ω integration in $\hat{\sigma}$ and $\hat{\kappa}_{WF}$ is actually ultraviolet-divergent. This divergence can be incorporated as a renormalization into the Drude results; this renormalization, however, does not invalidate the Wiedemann–Franz law.

6.2. Two-Dimensional System

To evaluate the interaction corrections for the entire temperature range, we need the exact form of the propagators. In two dimensions, they are given by

$$\mathcal{L}^\alpha(\omega, \mathbf{q}; \mathbf{n}_1, \mathbf{n}_2) = \Omega_d \delta(\widehat{\mathbf{n}_1, \mathbf{n}_2})\mathcal{L}_0(\omega, \mathbf{q}; \mathbf{n}_1) + \mathcal{L}_0(\omega, \mathbf{q}; \mathbf{n}_1)\mathcal{L}_0(\omega, \mathbf{q}; \mathbf{n}_2)\frac{\left(-i\omega\frac{F_0^\alpha}{1+F_0^\alpha} + \frac{1}{\tau}\right)\mathscr{C}}{\mathscr{C} - \left(-i\omega\frac{F_0^\alpha}{1+F_0^\alpha} + \frac{1}{\tau}\right)}, \quad (6.31)$$

where

$$\mathcal{L}_0(\omega, \mathbf{q}; \mathbf{n}) = \frac{1}{-i\omega + iv\cdot\mathbf{q} + 1/\tau}, \qquad \mathscr{C} = \sqrt{(-i\omega + 1/\tau)^2 + (v_F q)^2}. \quad (6.32)$$

We note that the variables ω and $v_F q$ are now bounded only by the Fermi energy E_F.

As before, we need to find the nonequilibrium corrections $\delta N^{\rho, g}$ to the bosonic distribution functions. These are again given by Eqs. (6.16), (6.17), but to obtain the thermal conductivity in an arbitrary temperature range, we calculate the solution of Eq. (6.18) exactly (in the linear order in ∇T):

$$\delta N^1 = \tilde{N}\left\{\frac{\mathbf{q}\cdot\nabla T}{v_F q^2} + \frac{i}{iv_F\mathbf{q}\cdot(\mathbf{n}_1 - \mathbf{n}_2) + 2/\tau} \times [\bar{\lambda}(\mathbf{n}_1)\mathbf{n}_1^\perp - \lambda(\mathbf{n}_2)\mathbf{n}_2^\perp]\cdot\nabla T\right\}, \quad (6.33)$$

where the bar denotes complex conjugation and we introduce the quantities

$$\tilde{N} = v_F\tau\frac{\omega^2}{T}\frac{\partial N_P}{\partial\omega}\frac{F_0}{1+F_0}, \qquad \mathbf{n}^\perp = \mathbf{n} - \frac{(\mathbf{n}\cdot\mathbf{q})\mathbf{q}}{q^2}, \qquad \lambda(\mathbf{n}) = \frac{a(\mathbf{n})}{a(\mathbf{n}) - b}, \quad b = \frac{-i\omega F_0}{1+F_0} + \frac{1}{\tau}, \qquad a(\mathbf{n}) = \sqrt{(v_F q)^2 + \left(-i\mathbf{q}\cdot\mathbf{v} + \frac{2}{\tau}\right)^2}. \quad (6.34)$$

We do not calculate the corrections to the electric conductivity, which would reproduce the results in [17], as shown in Appendix F (see also [18] for the generalization to arbitrary disorder). For convenience in the calculations, we separate the contributions due to δN^0 and δN^1 in $\delta\hat{\kappa}_{in}$ and $\hat{\kappa}^\rho$ (cf. Eqs. (6.11c) and (6.23)):

$$\delta\hat{\kappa}_{in} = \delta\hat{\kappa}_{in}^0 + \delta\hat{\kappa}_{in}^1, \quad (6.35a)$$

$$\delta\hat{\kappa}_{in}^0 = \frac{2\sigma_D}{e^2T}\int d\omega[\hat{\mathcal{P}}^{12}(\omega) - \hat{\mathcal{P}}^{11}(\omega)]\left[\frac{\omega^3}{4}\frac{\partial N_P}{\partial\omega}\right], \quad (6.35b)$$

$$[\delta\kappa_{in}^1]_{\mu\nu} = v_F\int\frac{d\omega}{2\pi}\omega\int\frac{d^2q}{(2\pi)^2}\int\frac{d\mathbf{n}_1 d\mathbf{n}_2 d\mathbf{n}_3}{(\Omega_2)^3}n_{1\mu} \times \mathrm{Re}\{\mathcal{L}_{12}^\rho\delta_\nu N_{23}^1 + \mathcal{L}_{32}^\rho\delta_\nu N_{21}^1 - \mathcal{L}_{12}^\rho\delta_\nu N_{21}^1\}, \quad (6.35c)$$

$$\hat{\kappa}^\rho = \hat{\kappa}_0^\rho + \hat{\kappa}_1^\rho, \quad (6.36a)$$

$$\hat{\kappa}_0^\rho - \hat{\kappa}^g = \frac{\sigma_D}{e^2T}\int d\omega\hat{\mathcal{B}}^0(\omega)\left[\frac{\omega^3}{4}\frac{\partial N_P}{\partial\omega}\right], \quad (6.36b)$$

$$\hat{\kappa}_1^\rho = -\int\frac{d\omega}{2\pi}\omega\int\frac{d^2q}{(2\pi)^2}\int\frac{d\mathbf{n}_1 d\mathbf{n}_2}{(\Omega_2)^2}\{\hat{s}_\mu;\ \mathcal{L}_{12}^\rho\delta_\nu N_{21}^1\}, \quad (6.36c)$$

where, as in Eq. (6.13), we use $\delta_\nu N^1$ to denote the variational derivative with respect to the temperature gradient and

$$\mathcal{B}_{\mu\nu}^0(\omega) = -\frac{4}{\pi\omega\nu}\int\frac{d^2q}{(2\pi)^2} \times \int\frac{d\mathbf{n}_1}{\Omega_2}n_{1\mu}n_{1\nu}\mathrm{Re}[\mathcal{L}_{11}^\rho - \mathcal{L}_{11}^g]. \quad (6.36d)$$

Expressions (6.8) for the electric conductivity and (6.11b) for the elastic correction to the thermal conductivity and definition (6.9) for the kernels remain unchanged. The momentum and angular integrals

in (6.9) can be calculated exactly; for the singlet channel in the unitary limit, we find

$$\mathscr{C}_{\mu\nu}(\omega) = -\frac{e^2\delta_{\mu\nu}}{\sigma_D 2\pi^2}\frac{1}{2\omega}\Big[2 - 2\omega\tau H(\omega\tau)\arctan\frac{1}{\omega\tau} + \frac{(\omega\tau)^2}{2}\Big(\frac{1}{2} - H(\omega\tau)\Big)\ln\Big(1 + \frac{1}{(\omega\tau)^2}\Big) - (\omega\tau)^2 H(\omega\tau)\ln 2\Big], \tag{6.37a}$$

$$\mathscr{S}^{11}_{\mu\nu}(\omega) = -\frac{e^2\delta_{\mu\nu}}{\sigma_D 2\pi^2}\tau\frac{\pi}{2}\operatorname{sgn}\omega, \tag{6.37b}$$

$$\mathscr{S}^{12}_{\mu\nu}(\omega) = -\frac{e^2\delta_{\mu\nu}}{\sigma_D 2\pi^2}\tau\Big\{(2H(\omega\tau) - 1)\arctan\frac{1}{\omega\tau} + \frac{\pi}{2}\operatorname{sgn}\omega + \omega\tau H(\omega\tau)\Big[\frac{1}{2}\ln\Big(1 + \frac{1}{(\omega\tau)^2}\Big) + \ln 2\Big]\Big\} \tag{6.37c}$$

with the function H defined as

$$H(x) = \frac{1}{4 + x^2}. \tag{6.37d}$$

To perform the momentum integral in Eq. (6.36d), we must keep the full form of the propagator in order to avoid the infrared divergence that we would obtain in the unitary limit,

$$\mathscr{B}^0_{\mu\nu}(\omega) = -\frac{e^2\delta_{\mu\nu}}{\sigma_D 2\pi^2}\frac{1}{\omega}\Big\{\ln\Big(\frac{Dk^2}{|\omega|}\Big) + \pi|\omega|\tau - \frac{1}{2}\ln[1 + (\omega\tau)^2] - \omega\tau\arctan\Big(\frac{1}{\omega\tau}\Big)\Big\}, \tag{6.37e}$$

where $k = 2\pi e^2\nu$ is the inverse screening radius. Next, we calculate the angular integrals in Eqs. (6.35c) and (6.36c) as well as the angular part of the momentum integrals, with the result

$$\int\frac{d\theta_q}{2\pi}\int\frac{d\mathbf{n}_1 d\mathbf{n}_2 d\mathbf{n}_3}{(\Omega_2)^3}n_{1\mu}\operatorname{Re}\{\mathscr{L}^\rho_{12}\delta_\nu N^1_{23} + \mathscr{L}^\rho_{32}\delta_\nu N^1_{21}\} = \frac{i}{4}\tilde{N}\delta_{\mu\nu}\Big\{\frac{1}{(vq)^2}\frac{1}{\mathscr{C}\bar{\mathscr{C}}}\Big[\frac{\mathscr{C}\bar{b}'}{\mathscr{C} - b} + \frac{\bar{\mathscr{C}}b'}{\bar{\mathscr{C}} - \bar{b}}\Big] \times[\bar{\mathscr{C}} - \bar{b}' - \mathscr{C} + b'] + \frac{1}{(vq)^2}\Big[\frac{\mathscr{C} - b'}{\mathscr{C}} - \frac{\bar{\mathscr{C}} - \bar{b}'}{\bar{\mathscr{C}}}\Big] - \frac{1}{\mathscr{C}\bar{\mathscr{C}}}\Big[\frac{\mathscr{C}}{\mathscr{C} - b} - \frac{\bar{\mathscr{C}}}{\bar{\mathscr{C}} - \bar{b}}\Big]\Big\}, \tag{6.38a}$$

$$\int\frac{d\theta_q}{2\pi}\int\frac{d\mathbf{n}_1 d\mathbf{n}_2}{(\Omega_2)^2}n_{1\mu}\operatorname{Re}\{\mathscr{L}^\rho_{12}\delta_\nu N^1_{21}\} = \frac{i}{4}\tilde{N}\delta_{\mu\nu}\frac{\tau}{2}\Big\{\Big[\frac{\bar{\mathscr{C}}}{\bar{\mathscr{C}} - \bar{b}} - \frac{\mathscr{C}}{\mathscr{C} - b}\Big]\Big[\frac{1}{\mathscr{C}} + \frac{1}{\bar{\mathscr{C}}}\Big] - \frac{1}{(vq)^2}\Big[\frac{\mathscr{C}\bar{b}'}{\mathscr{C} - b} + \frac{\bar{\mathscr{C}}b'}{\bar{\mathscr{C}} - \bar{b}}\Big]\Big[\frac{\mathscr{C} - b'}{\mathscr{C}} - \frac{\bar{\mathscr{C}} - \bar{b}'}{\bar{\mathscr{C}}}\Big]\Big\}. \tag{6.38b}$$

The function $\mathscr{C}$ is defined in Eq. (6.32) and b' is given by the $F_0 \longrightarrow +\infty$ limit of b in Eq. (6.34), where $\tilde{N}$ is also defined. The remaining integrals over the magnitude of the momentum can be evaluated approximately; the result can be written as

$$\delta\hat{\kappa}^1_{\text{in}} + \hat{\kappa}^\rho_1 = \frac{\sigma_D}{e^2T}\int d\omega\hat{\mathscr{B}}^1(\omega)\Big[\frac{\omega^3}{4}\frac{\partial N_P}{\partial\omega}\Big], \tag{6.39a}$$

with

$$\mathscr{B}^1_{\mu\nu}(\omega) = \frac{e^2\delta_{\mu\nu}}{\sigma_D 2\pi^2}\tau\Big\{\frac{2\omega\tau}{1 + (\omega\tau)^2}\ln\Big(\frac{v_F k}{2|\omega|}\Big) - \arctan\omega\tau - \omega\tau\ln\Big(\frac{E_F}{\sqrt{\omega^2 + \tau^{-2}}}\Big)\Big\}. \tag{6.39b}$$

In the above kernel, the first term in the curly brackets originates from $\hat{\kappa}^\rho_1$ only: as discussed in Section 5.2, no long-range terms can be present in the electron contribution to the thermal conductivity. We note that the second term in the above expression is beyond the logarithmic accuracy of our approximate calculation and must be dropped. Similarly, most of the terms in the other kernels can be neglected, and collecting the logarithmic contributions, we obtain

$$\Delta\kappa_s = -\frac{1}{2\pi^2 T}\int d\omega\Big[\frac{\omega^2}{4}\frac{\partial N_P}{\partial\omega}\Big] \times\Big\{(\omega\tau)^2\ln\Big(\frac{E_F}{|\omega|}\Big) + \frac{2}{1 + (\omega\tau)^2}\ln\Big(\frac{v_F k}{|\omega|}\Big) + \Big[\frac{7}{12}(\omega\tau)^2 - \frac{5}{6} + \frac{16}{3}\frac{1}{4 + (\omega\tau)^2}\Big]\ln\Big(1 + \frac{1}{(\omega\tau)^2}\Big)\Big\}, \tag{6.40}$$

where we define the singlet and triplet channel corrections as

$$[\kappa - \kappa_{WF}]_{\mu\nu} = (\Delta\kappa_s + 3\Delta\kappa_t)\delta_{\mu\nu}.$$

The final integration can now be performed within the

logarithmic accuracy; we find (cf. Eq. (4.4b))

$$\Delta\kappa_s = -\frac{\pi^2}{15}T(T\tau)^2\ln\left(\frac{E_F}{T}\right) + \frac{T}{6}g_1(2\pi T\tau)\ln\left(\frac{v_F k}{T}\right) - \frac{T}{24}g_2(\pi T\tau)\ln\left(1+\frac{1}{(T\tau)^2}\right), \quad (6.41)$$

where

$$g_1(x) = \frac{3}{x^2}\left\{\frac{1}{x}\left[2\psi'\left(\frac{1}{x}\right) - x^2\right] - 2\right\}, \quad (6.42a)$$

$$g_2(x) = \frac{14}{15}x^2 + \frac{8}{3}g_1(x) - \frac{5}{3}, \quad (6.42b)$$

and ψ' is the derivative of the digamma function. Because the asymptotic behavior of $g_1(x)$ is

$$g_1(x) = \begin{cases} 1 - \frac{1}{5}x^2 + \frac{1}{7}x^4 + \ldots, & x \ll 1 \\ \frac{3}{x} - \frac{6}{x^2} + \frac{\pi^2}{x^3} + \ldots, & x \gg 1, \end{cases} \quad (6.42c)$$

both these functions tend to 1 as $T\tau \longrightarrow 0$; therefore, in the diffusive limit, the main contribution is $T\ln(Dk^2/T)/12$. On the other hand, for $T\tau \gg 1$, the first term in Eq. (6.41) is the dominant one.

Turning to the triplet channel, we restrict ourselves to the limiting diffusive and quasiballistic cases for simplicity, although one can extend the calculation to the entire temperature range, as is done in [17] for the electric conductivity.

In the diffusive limit $T\tau \ll 1$, we know from our previous analysis that we can discard the $\mathcal{B}^1$ term as well as the $\mathcal{S}$ terms. The relevant kernels are then[12]

$$\mathcal{E}_{\mu\nu}(\omega) = -\frac{e^2\delta_{\mu\nu}}{\sigma_D 2\pi^2\omega}\left[1 - \frac{1}{F_0^\sigma}\ln(1+F_0^\sigma)\right], \quad (6.43)$$

$$\mathcal{B}^0_{\mu\nu}(\omega) = -\frac{e^2\delta_{\mu\nu}}{\sigma_D 2\pi^2\omega}\ln(1+F_0^\sigma). \quad (6.44)$$

Their substitution in Eqs. (6.11)–(6.36b) gives

$$\Delta\kappa_t = -\frac{T}{18}\left[1 - \frac{1}{F_0^\sigma}\ln(1+F_0^\sigma)\right] + \frac{T}{12}\ln(1+F_0^\sigma). \quad (6.45)$$

In the opposite limit $T\tau \gg 1$, the main contribution comes, as for the singlet channel, from the logarithmic divergence at large momenta in the kernel $\mathcal{B}^1$:

$$\mathcal{B}^1_{\mu\nu}(\omega) = -\frac{e^2\delta_{\mu\nu}}{\sigma_D 2\pi^2} \times \tau\left\{\omega\tau\ln\left(\frac{E_F}{\sqrt{\omega^2+\tau^{-2}}}\right)\right\}\left(\frac{F_0^\sigma}{1+F_0^\sigma}\right)^2. \quad (6.46)$$

Then the correction to the thermal conductivity is

$$\Delta\kappa_t = -\frac{\pi^2}{15}T(T\tau)^2\ln\left(\frac{E_F}{T}\right)\left(\frac{F_0^\sigma}{1+F_0^\sigma}\right)^2, \quad (6.47)$$

which concludes the derivation of Eq. (4.4c).

This correction to the thermal conductivity (and the corresponding one in the singlet channel) is the contribution of inelastic processes to the energy relaxation rate.[13] In a clean system, such inelastic processes cannot relax momentum (because of the Galilean invariance), and hence they do not affect the electric conductivity, but they can contribute to the energy relaxation rate Γ_ε. In the kinetic theory, the thermal conductivity can be written, up to a numerical coefficient, as

$$\kappa \propto TE_F/\Gamma_\varepsilon,$$

and the rate is given by the sum of the rates for the relevant processes, namely, the electron–impurity and electron–electron scattering rates,

$$\Gamma_\varepsilon = \Gamma_{\text{imp}} + \Gamma_{e-e}$$

with $\Gamma_{\text{imp}} = 1/\tau$ and

$$\Gamma_{e-e} = a\frac{T^2}{E_F}\ln\left(\frac{E_F}{T}\right).$$

Here, a is a constant whose exact value is irrelevant for our argument. In the limit $(T^2/E_F)\ln(E_F/T) \ll 1/\tau$, we can expand the expression for the total rate, substitute the result into the above formula for κ and obtain

$$\kappa \propto T\tau E_F - aT(T\tau)^2\ln\left(\frac{E_F}{T}\right).$$

The first term in the right-hand side is the usual Drude result for the thermal conductivity and the second term has the form of correction (6.47). We note that in the opposite limit (clean system), the result is

$$\kappa \propto \frac{E_F^2}{T\ln\left(\frac{E_F}{T}\right)},$$

in agreement with the result in [30].

[12] They can be obtained from Eqs. (6.25), (6.26) in the limit $d \longrightarrow 2$.

[13] A similar argument is presented in [21].

6.3. Specific Heat

Recalling our discussion on the structure of the kinetic equation in Section 2, we write the total specific heat as the sum of the electronic and bosonic contributions,

$$c_V = c_V^0 + \delta c_V, \tag{6.48a}$$

$$c_V^0 = \frac{\partial u_e}{\partial T} = \frac{\pi^2}{3}\nu T, \tag{6.48b}$$

$$\delta c_V = \frac{\partial}{\partial T}(u_\rho - u_g), \tag{6.48c}$$

where, in accordance with Eq. (5.93c), the bosonic energy densities (in the equilibrium (3.1)) are

$$u_\alpha = \int d\omega \omega b^\alpha(\omega) N_P(\omega) \tag{6.49}$$

with[14]

$$b^\alpha(\omega) = \frac{\mathrm{Re}}{2\pi}\int \frac{d^d q}{(2\pi)^d}\mathrm{Tr}_{\mathbf{n}}\left\{\frac{1}{1+\hat{F}^\alpha}; \mathscr{L}^\alpha(\omega, \mathbf{q})\right\}. \tag{6.50}$$

As before, we explicitly consider the singlet-channel, short-range interaction in the zeroth harmonic approximation for the Fermi-liquid constant (denoted by F_0). The results for the long-range interaction in the unitary limit are obtained by letting $F_0^\rho \longrightarrow +\infty$. For the triplet channel, we must substitute F_0^ρ with F_0^σ and multiply by an overall factor of 3. The final answer with the correct coefficients is given in Section 4.2.

In the diffusive limit, to which we restrict our attention for $d = 1, 3$, we have

$$b^\rho(\omega) - b^g(\omega) = \frac{\mathrm{Re}}{2\pi} \times \int \frac{d^d q}{(2\pi)^d}\left[\frac{1}{1+F_0^\rho} L_0^\rho(\omega, \mathbf{q}) - L_0^g(\omega, \mathbf{q})\right], \tag{6.51}$$

with the functions L_0^α defined in Eqs. (6.20) and (6.21). After the momentum integration, we find

$$b^\rho(\omega) - b^g(\omega) = \frac{\Omega_d}{(2\pi)^d}\left(\frac{|\omega|}{D}\right)^{d/2}\frac{1}{4\omega} \times \frac{\cos\frac{\pi}{4}(d-2)}{\sin\frac{\pi}{2}(d-2)}\left[1 - \frac{1}{(1+F_0^\rho)^{d/2}}\right]. \tag{6.52}$$

[14]This definition of the density of states is half of that in Eqs. (2.19) because of the different limits for the ω integration in the energy density and the specific heat.

Inserting this in Eq. (6.48c) gives

$$\delta c_V = \frac{1}{2}\frac{\Omega_d}{(2\pi)^d}\left(\frac{T}{D}\right)^{d/2}\zeta\left(\frac{d}{2}+1\right)\Gamma\left(\frac{d}{2}+2\right) \times \frac{\cos\frac{\pi}{4}(d-2)}{\sin\frac{\pi}{2}(d-2)}\left[1 - \frac{1}{(1+F_0^\rho)^{d/2}}\right]. \tag{6.53}$$

The relevant numerical values for the zeta and gamma functions are given after Eq. (6.27), which has been used to evaluate the ω integral.

For $d = 2$, we can keep the full form of the propagators to find the singular contribution to the specific heat at an arbitrary value of $T\tau$,

$$b^\rho(\omega) - b^g(\omega) = -\frac{\mathrm{Re}}{2\pi}\int \frac{d^2 q}{(2\pi)^2} \times \left[\frac{F_0^\rho}{1+F_0^\rho}\frac{1}{\mathscr{C}-b} - \frac{(-i\omega + 1/\tau)}{\mathscr{C}}\left(\frac{1}{\mathscr{C}-b} - \frac{1}{\mathscr{C}-1/\tau}\right)\right], \tag{6.54}$$

with $\mathscr{C}$ and b defined respectively in Eqs. (6.32) and (6.34). The first term in the integral is formally divergent as $|q| \longrightarrow \infty$; this divergence gives a linear-in-T contribution to the specific heat that does not depend on disorder. This term must be disregarded because all the linear terms are included in the definition of the effective electron mass, and taking it into account would lead to a double counting. To regularize the integral, we replace $1/(\mathscr{C}-b) \longrightarrow 1/(\mathscr{C}-b) - 1/\mathscr{C}$ in the first line.

Evaluating the momentum integral, we obtain

$$b^\rho(\omega) - b^g(\omega) = -\frac{1}{8\pi^2 D}\left[\frac{F_0^\rho}{1+F_0^\rho}\ln\left(\frac{E_\mathrm{F}}{|\omega|}\right) + \left(\frac{F_0^\rho}{1+F_0^\rho}\right)^2\frac{\pi}{2}\tau|\omega| - \frac{1}{1+F_0^\rho}\ln(1+F_0^\rho)\right]. \tag{6.55}$$

The final answer for the correction to the specific heat is then

$$\delta c_V = -\frac{1}{12}\frac{T}{D}\frac{1}{1+F_0^\rho}\left[F_0^\rho \ln\left(\frac{E_\mathrm{F}}{T}\right) - \ln(1+F_0^\rho)\right] - \frac{1}{4\pi^2}[(2\gamma-3)\zeta(2) - 2\zeta'(2)]\frac{T}{D}\frac{F_0^\rho}{1+F_0^\rho} - \frac{3}{4\pi}\zeta(3)\frac{T}{D}(T\tau)\left(\frac{F_0^\rho}{1+F_0^\rho}\right)^2, \tag{6.56}$$

where $\zeta(2) \approx 1.645$, $\zeta'(2) \approx -0.938$, and $\zeta(3) \approx 1.202$. In the quasiballistic limit $\tau \longrightarrow +\infty$, only the last line is relevant:

$$\delta c_V = -\frac{3}{2\pi}\zeta(3)\left(\frac{F_0^\rho}{1+F_0^\rho}\right)^2\frac{T^2}{v_F^2}.$$

This T^2 correction to the specific heat has the same form as the correction found for two-dimensional Fermi liquids [28] and agrees (in the appropriate limit) with the result in [29].

As discussed before, the long-range interaction can be taken into account by passing to the limit as $F_0^\rho \longrightarrow +\infty$, while the triplet channel contribution is three times as large (cf. Section 4.2).

7. CONCLUSIONS

Locality at the scale determined by the temperature and the validity of the conservation laws are two main requirements for a proper kinetic description of any system. In the present paper, we derived such a description for the interaction effects in disordered metals (assuming that the clean counterpart of the system is a stable Fermi liquid).

We showed that this description requires the introduction of bosonic distribution functions in addition to the usual fermionic quasiparticle distribution function. These neutral bosons are of two types: (i) the ones describing oscillations in charge density (singlet) or spin density (triplet) and (ii) fictitious (ghost) bosons that prevent overcounting the degrees of freedom (electron–hole pairs) already included in the fermionic part. The conservation laws obtained together with gauge invariance allow an unambiguous definition of the corresponding electric and energy currents.

For the electric transport, the neutral bosons are not important and our description reproduces the known results for the correction to the conductivity obtained in [11] for the diffusive regime and in [17] in the ballistic and crossover regimes.

The neutral bosons, however, are crucial for the thermal properties of the system. Namely, their contributions to the energy density are responsible for nonanalytic corrections to the specific heat, see Eqs. (4.6) and (4.8). Our kinetic equation approach reproduces the results for the interaction corrections to the specific heat previously calculated within the equilibrium diagram technique [12]. Moreover, the neutral boson contributions to the energy current violate the Wiedemann–Franz law, see Eq. (4.1) and the discussion that follows it. The violation is stronger for lower dimensional systems ($d = 1, 2$) in the diffusive regime, see Eqs. (4.2) and (4.4b). Other effects contributing to the violation of the Wiedemann–Franz law are the energy dependence of the electron elastic scattering and the inelastic scattering of the electrons on the neutral bosons. The latter effect was found to be relevant in the quasiballistic regime $T\tau \gg \hbar$ for two-dimensional systems, see Eqs. (4.4).

The violation of the Wiedemann–Franz law was investigated before in the diffusive regime in [20] and [23] within the "quantum kinetic equation" approach and by the Kubo formula in [22]. Ironically, even though the forms of the energy current operator used in those references are wrong, the final results for the thermal conductivity are consistent with our Eqs. (4.2)–(4.4). We think that this agreement is accidental.

ACKNOWLEDGMENTS

We are grateful to B.L. Altshuler for initiating this work. Interesting conversations with M.Yu. Reizer are acknowledged. We thank A.I. Larkin, A.J. Millis, and B.N. Narozhny for discussions of our results. I. A. was supported by the Packard foundation.

APPENDIX A

Correction to the Thermodynamic Potential

A standard analytic continuation of Eq. (2.17) gives

$$\delta\Omega = \int\frac{d\omega}{4\pi}\coth\left(\frac{\omega}{2T}\right)\int\frac{d^dq}{(2\pi)^d} \times \mathrm{Im}\ln\left(1+\frac{F}{\nu}\Pi^R(\omega,\mathbf{q})\right) \quad (A.1)$$

$$= \int\frac{d\omega}{2\pi}\frac{1}{2}\coth\left(\frac{\omega}{2T}\right)\mathrm{Im}\,\mathrm{Tr}\ln\left(1+\frac{\hat F}{\nu}\hat\Pi^R\right).$$

In the second line, we use the operator notation (see Eqs. (3.3) and (3.13)), which gives the correct generalization for the momentum-dependent Fermi-liquid parameter. Substituting the transform of the explicit expression (5.43a) for the polarization operator, we rewrite the argument of the logarithm as

$$(1+\hat F)\left[(\hat{\mathscr{L}}^g)^{-1}+\frac{\hat F}{1+\hat F}i\omega\right]\hat{\mathscr{L}}^g.$$

According to definition (3.11), the term in the square brackets is $(\hat{\mathscr{L}}^\rho)^{-1}$. Using the property

$$\mathrm{Tr}\ln(\hat A\hat B) = \mathrm{Tr}\ln\hat A + \mathrm{Tr}\ln\hat B$$

and the fact that $\ln(1+\hat F)$ does not contribute to the imaginary part, we conclude that

$$\mathrm{Im}\,\mathrm{Tr}\ln\left(1+\frac{\hat F}{\nu}\hat\Pi^R\right) = -\mathrm{Im}\,\mathrm{Tr}[\ln\hat{\mathscr{L}}^\rho - \ln\hat{\mathscr{L}}^g].$$

Substituting this identity in Eq. (A.1), we finally obtain Eq. (2.18).

APPENDIX B

The Microscopic Form of the Energy Current Operator

The action entering the partition function that describes the electron gas in the presence of an external electric field is given by

$$S = \int dt d^d r [i\psi^\dagger \partial_t \psi - \psi^\dagger \hat{H}_{gi} \psi - \psi^\dagger \varphi \psi] \tag{B.1}$$

with the condition $\nabla \times \mathbf{A} = 0$ that ensures the absence of the magnetic field. The variables $(t, \mathbf{r})$ on which all the fields depend are suppressed. The gauge-invariant part of the Hamiltonian for the noninteracting system has the usual form

$$\begin{aligned}\psi^\dagger \hat{H}_{gi} \psi &= \frac{1}{2m}(i\nabla + \mathbf{A})\psi^\dagger(-i\nabla + \mathbf{A})\psi \\ &\quad + \psi^\dagger V_{\text{imp}} \psi,\end{aligned} \tag{B.2}$$

where V_{imp} is the impurity potential and the potentials φ and $\mathbf{A}$ describe the external electric field:

$$e\mathbf{E}_{\text{ext}} = -\nabla\varphi + \partial_t \mathbf{A}.$$

As usual, the charge conservation law

$$\partial_t \rho + \nabla \cdot \mathbf{j} = 0 \tag{B.3}$$

follows from the requirement of gauge invariance, with the charge and current densities given by

$$\begin{aligned}\rho &= e\psi^\dagger \psi, \\ \mathbf{j} &= \frac{e}{2m}[\psi^\dagger(-i\nabla + \mathbf{A})\psi + (i\nabla + \mathbf{A})\psi^\dagger \psi].\end{aligned} \tag{B.4}$$

The invariance of the action under the replacement

$$\psi(t, \mathbf{r}) \longrightarrow \psi(t + \alpha(t, \mathbf{r}), \mathbf{r})$$

(and a similar replacement for $\psi^\dagger$) underlies for the derivation of the energy conservation law. A straightforward calculation gives

$$\begin{aligned}\frac{\delta S}{\delta \alpha} &= 0 = \partial_t'(\psi^\dagger \hat{H}_{gi} \psi) + \varphi \partial_t(\psi^\dagger \psi) \\ &\quad - \frac{i}{2m}\nabla \cdot [\partial_t \psi^\dagger(-i\nabla + \mathbf{A})\psi - (i\nabla + \mathbf{A})\psi^\dagger \partial_t \psi],\end{aligned} \tag{B.5}$$

where the prime means that the derivative acts on ψ, $\psi^\dagger$ only. By adding and subtracting terms proportional to φ in the last bracket and to $\partial_t \mathbf{A}$ in the first term, we find the energy conservation law

$$\partial_t u_0 + \nabla \cdot \mathbf{j}_0^\varepsilon = \mathbf{j} \cdot \mathbf{E}_{\text{ext}} - \frac{1}{e}\varphi[\partial_t \rho + \nabla \cdot \mathbf{j}], \tag{B.6}$$

where

$$u_0 = \psi^\dagger \hat{H}_{gi} \psi, \tag{B.7a}$$

$$\begin{aligned}\mathbf{j}_0^\varepsilon &= -\frac{1}{2m}[(i\partial_t + \varphi)\psi^\dagger(-i\nabla + \mathbf{A})\psi \\ &\quad - (i\nabla + \mathbf{A})\psi^\dagger(i\partial_t - \varphi)\psi].\end{aligned} \tag{B.7b}$$

The first term in the right-hand side of Eq. (B.6) is the usual Joule heat; the last term in that equation is not gauge invariant, but it vanishes because it is proportional to the left-hand side of continuity Eq. (B.3).

We now consider the generalization to the interacting case. The Hamiltonian then contains the additional term

$$\frac{1}{2}\int d^d r_1 \psi^\dagger \psi(\mathbf{r}) V(\mathbf{r} - \mathbf{r}_1) \psi^\dagger \psi(\mathbf{r}_1),$$

where $V(\mathbf{r}) = e^2/|\mathbf{r}|$ describes the density–density Coulomb interaction, which can be decoupled by the Hubbard—Stratonovich transformation. This amounts to introducing of the quantum fields ϕ and $\mathbf{A}$ in the action by adding the term

$$(-\psi)^\dagger \phi \psi + \frac{1}{2}\mathbf{E}_{fl}^2, \quad e\mathbf{E}_{fl} = -\nabla\phi + \partial_t \mathbf{A} \tag{B.8}$$

and redefining the vector potential as the sum of the external and fluctuating ones:

$$\mathbf{A} \longrightarrow \mathbf{A}_{\text{ext}} + \mathbf{A}. \tag{B.9}$$

The variation of the action with respect to the fluctuating potentials results in the first and fourth Maxwell equations relating the fluctuating electric field to the charge and current densities,

$$\nabla \cdot \mathbf{E}_{fl} = \rho, \quad 0 = -\mathbf{j} + \partial_t \mathbf{E}_{fl}, \tag{B.10}$$

where the electric current is defined in Eq. (B.4), but with substitution (B.9) performed.

To obtain the energy conservation law, we must consider the further transformation

$$\phi(t, \mathbf{r}) \longrightarrow \phi(t + \alpha(t, \mathbf{r}), \mathbf{r})$$

and a similar one for $\mathbf{A}$. Proceeding as before, we find the conservation law

$$\begin{aligned}\frac{\partial u_?}{\partial t} + \nabla \cdot \mathbf{j}_?^\varepsilon &= \mathbf{j} \cdot \mathbf{E}_{\text{ext}}, \\ u_? &= u_0 - \frac{1}{2}\mathbf{E}_{fl}^2 + \frac{1}{e}\left[\rho\phi + \frac{\partial \mathbf{A}}{\partial t} \cdot \mathbf{E}_{fl}\right], \\ \mathbf{j}_?^\varepsilon &= \mathbf{j}_0^\varepsilon - \frac{1}{e}\phi\mathbf{j} + \frac{1}{e}\left[\phi\mathbf{j} - \frac{\partial \phi}{\partial t}\mathbf{E}_{fl}\right],\end{aligned} \tag{B.11}$$

where u_0, $\mathbf{j}_0^\varepsilon$ are defined in Eq. (B.7) (with substitution (B.9)). Given the form of Eq. (B.11), one might be

tempted to call $u_?$ and $\mathbf{j}^{\varepsilon}_?$ the energy and energy current densities, in terms of which the conservation law takes exactly the same form as in the noninteracting case. But such a redefinition would result in gauge-dependent expressions for the densities, because the terms in the square brackets taken separately are not gauge invariant. Hence, this naive redefinition of the conserved quantities is nonphysical, because any physical perturbation can be coupled only to gauge-invariant quantities. To find gauge-invariant definitions, we rewrite the contribution of those gauge-noninvariant terms as

$$\frac{\partial}{\partial t}\left[\rho\phi + \frac{\partial \mathbf{A}}{\partial t}\cdot\mathbf{E}_{fl}\right] + \nabla\left[\phi\mathbf{j} - \frac{\partial\phi}{\partial t}\mathbf{E}_{fl}\right]$$

$$= \frac{1}{e}\phi[\partial_t\rho + \nabla\cdot\mathbf{j}] + \frac{1}{e}\frac{\partial\phi}{\partial t}[\rho - \nabla\cdot\mathbf{E}_{fl}]$$

$$+ \frac{1}{e}\frac{\partial}{\partial t}(\partial_t\mathbf{A} - \nabla\phi)\cdot\mathbf{E}_{fl} + \frac{1}{e}[\nabla\phi\cdot\mathbf{j} + \partial_t\mathbf{A}\partial_t\mathbf{E}_{fl}].$$

Here, the second line vanishes because of charge conservation, Eq. (B.3), and because of the first Maxwell Eq. (B.10). In the third line, we use the second Maxwell equation to eliminate the current; in the result, we substitute the definition of the fluctuating field given in Eq. (B.8) and obtain that the third line of the above equation is equal to $\partial_t\mathbf{E}^2_{fl}$. This enables us to conclude that the correct gauge-invariant expressions for the energy and energy current densities are

$$u = \psi^\dagger\hat{H}_{gi}\psi + \frac{1}{2}\mathbf{E}^2_{fl}, \tag{B.12}$$

$$\mathbf{j}^{\varepsilon} = -\frac{1}{2m}[(i\partial_t + \varphi)\psi^\dagger(-i\nabla + \mathbf{A})\psi - (i\nabla + \mathbf{A})\psi^\dagger(i\partial_t - \varphi)\psi], \tag{B.13}$$

where the potentials are total ones:

$$\mathbf{A} = \mathbf{A}_{\text{ext}} + \mathbf{A}, \quad \varphi = \varphi_{\text{ext}} + \phi.$$

We note that these expressions are gauge-invariant with respect to gauge transformations of both the external and fluctuating potentials. We believe that only such quantities can be coupled to the "gravitational field" in the Luttinger scheme for the calculation of the thermal conductivity [5].

The same final answer is obtained if the interaction is decoupled in the "gauge-fixed" form $\mathbf{A} = 0$. In this case, which is the most widely used in the literature, there are two contributions to the energy current vertex in the diagram approach, see Fig. 4: in addition to the usual vertex of the noninteracting case, which arises from the terms $\partial_\tau\psi^\dagger\nabla\psi$, there is a vertex from the $\phi\psi^\dagger\nabla\psi$ terms. These vertices were not taken into account in [20, 22, 23]. However, analogous vertices were previously considered in the calculations of the thermoelectric coefficient[15] with the inclusion of the electron–electron interaction in the particle–hole channel [31] and in the Cooper channel [32] and for the electron–phonon interaction [33].

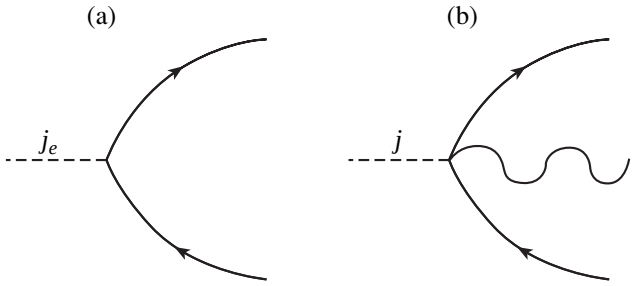

Fig. 4. (a) The energy current vertex for the noninteracting case; (b) the additional vertex induced by the interaction. The solid lines with arrows are the electron Green functions, the wavy line is the interaction propagator, the dashed lines are the "standard" (noninteracting case) energy and electric current operators defined respectively in Eqs. (B.7b) and (B.4).

APPENDIX C

Alternative Parameterization

The operator $\hat{H}_{e-h}$ defined in Eq. (5.54) is clearly not a standard Hamiltonian. However, we can introduce a different definition of the propagator $\mathcal{L}^\rho$,

$$\left[\frac{\partial}{\partial t_1} + i\hat{H}_{e-h}(-i\nabla_1) - \hat{\text{St}}^{\text{out}}_s\right]\mathcal{L}^\rho = \delta_{12}, \tag{C.1}$$

such that the (new) $\hat{H}_{e-h}$ operator is indeed a Hamiltonian:

$$\hat{H}_{e-h}(\mathbf{q}) \equiv \hat{\mathcal{F}}^\rho\mathbf{v}\cdot\mathrm{q}\hat{\mathcal{F}}^\rho, \quad \hat{\text{St}}^{\text{out}}_s \equiv \hat{\mathcal{F}}^\rho\hat{\text{St}}_\tau\hat{\mathcal{F}}^\rho, \quad \hat{\mathcal{F}}^\rho \equiv (1+\hat{F})^{1/2} \tag{C.2}$$

(the action of the operator $\hat{F}$ is defined in Eq. (5.49)). Proceeding as in Section 5.4, we obtain the following expressions for the $\mathcal{K}$ propagators:

$$\nu\hat{\mathcal{K}}^R = (\partial_t)^{-1}[\hat{\mathcal{L}}^g - \hat{\mathcal{F}}^\rho\hat{\mathcal{L}}^\rho\hat{\mathcal{F}}^\rho],$$
$$\nu\hat{\mathcal{K}}^A = -[\hat{\bar{\mathcal{L}}}^g - \hat{\mathcal{F}}^\rho\hat{\bar{\mathcal{L}}}^\rho\hat{\mathcal{F}}^\rho](\partial_t)^{-1}, \tag{C.3a}$$

$$\nu\hat{\mathcal{K}}^K = -i(\partial_t)^{-1}[\hat{\mathcal{L}}^g\hat{\mathcal{N}}^g + \hat{\mathcal{N}}^g\hat{\bar{\mathcal{L}}}^g](\partial_t)^{-1} + i(\partial_t)^{-1}\hat{\mathcal{F}}^\rho[\hat{\mathcal{L}}^\rho\hat{\mathcal{N}}^\rho + \hat{\mathcal{N}}^\rho\hat{\bar{\mathcal{L}}}^\rho]\hat{\mathcal{F}}^\rho(\partial_t)^{-1}. \tag{C.3b}$$

[15] A derivation similar to that in this appendix was performed independently by M.Yu. Reizer (private communication).

The "kinetic equation" for $\mathcal{N}^g$, Eq. (5.64a), remains unchanged, while $\mathcal{N}^\rho$ now satisfies the equation

$$[\partial_t + \hat{H}_{e-h}(-i\nabla);\, \hat{\mathcal{N}}^\rho] = \hat{\mathrm{St}}^\rho\{\mathcal{N}^\rho, g^K\}, \tag{C.4}$$

where

$$\hat{\mathrm{St}}^\rho\{\mathcal{N}^\rho, g^K\} \equiv 2\{\hat{\mathrm{St}}_s^{\mathrm{out}};\, \hat{\mathcal{N}}^\rho\} + 2\hat{\mathcal{F}}\hat{M}\hat{\mathcal{F}} \tag{C.5}$$

or, after Wigner transformations (5.61)–(5.63),

$$\omega[\partial_t \hat{N}^\rho + \{\hat{s}, \nabla\hat{N}^\rho\} + i[\hat{H}_{e-h}(\mathbf{q}), \hat{N}^\rho]] = \hat{\mathrm{St}}^\rho\{N^\rho, f\}, \tag{C.6}$$

$$\hat{\mathrm{St}}^\rho\{N^\rho, f\}(\omega, \mathbf{n}_1, \mathbf{n}_2) = \int d\varepsilon \int \frac{d\mathbf{n}_3 \ldots d\mathbf{n}_6}{\Omega_d^4} \tag{C.7}$$

$$\times [\mathcal{F}_{16}\gamma^6_{34}\Upsilon^\rho_{52;\,43}(\varepsilon, \omega)\mathcal{F}_{45} + \mathcal{F}_{56}\gamma^6_{34}\Upsilon^\rho_{15;\,34}(\varepsilon, \omega)\mathcal{F}_{42}],$$

with definitions (5.27) and (5.66) for γ and Υ.

We can then proceed as in Section 5.6 and obtain conservation laws (5.89) and (5.94); the only formal difference is in the definition of the bosonic energy density, which is now

$$u_\alpha(t, \mathbf{r}) = \frac{1}{2}\mathrm{Tr}_{\mathbf{n}}[\hat{\mathcal{L}}^\alpha\hat{\mathcal{N}}^\alpha].$$

In the alternative parameterization, the formalism can be developed with not more difficulties than in the original one. However, the evaluation of the thermal conductivity becomes cumbersome. In the original parameterization, it is also easier to include (at least perturbatively) the effects due to higher harmonics of the Fermi-liquid parameters.

APPENDIX D

Derivation of the Electron Collision Integral

The calculation of the matrix collision integral is simplified by the introduction of two functions $A(t, \mathbf{r}, \mathbf{n}, \tilde{\mathbf{n}})$ and $B(t, \mathbf{r}, \mathbf{n}, \tilde{\mathbf{n}})$ such that

$$e^{i\hat{K}(t, \mathbf{n}, \mathbf{r})}e^{-i\hat{K}(t, \tilde{\mathbf{n}}, \mathbf{r})} = (A\hat{\mathbb{1}}_K + B\hat{\sigma}^x_K)_{\mathbf{n}, \tilde{\mathbf{n}}}. \tag{D.1}$$

We recall that $\hat{K} = K_+\hat{\mathbb{1}}_K + K_-\hat{\sigma}^x_K$ and

$$\hat{\sigma}^x_K = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

The collision integral (i.e., the right-hand side of Eq. (5.22)) in the matrix notation is then

$$\frac{1}{2\tau}\langle[\hat{g}(\mathbf{n}) \circ (A\hat{\mathbb{1}}_K + B\hat{\sigma}^x_K)_{\mathbf{n}, \tilde{\mathbf{n}}}\hat{g}(\tilde{\mathbf{n}})(A\hat{\mathbb{1}}_K + B\hat{\sigma}^x_K)_{\tilde{\mathbf{n}}, \mathbf{n}} - (A\hat{\mathbb{1}}_K + B\hat{\sigma}^x_K)_{\mathbf{n}, \tilde{\mathbf{n}}}\hat{g}(\tilde{\mathbf{n}})(A\hat{\mathbb{1}}_K + B\hat{\sigma}^x_K)_{\tilde{\mathbf{n}}, \mathbf{n}} \circ \hat{g}(\mathbf{n})]\rangle_{\tilde{\mathbf{n}}}, \tag{D.2}$$

where the open dot indicates the time convolution (cf. Eq. (5.9)) and the time argument of the functions A and B is the first (second) time argument of the Green function on their right (left), e.g.,

$$BgB \circ g \equiv \int dt_3 B(t_1)g^K(t_1, t_3)B(t_3)g^K(t_3, t_2).$$

Substituting the matrix Green function of form (5.23), we find that the collision integral becomes

$$\left\langle \begin{pmatrix} \hat{\mathrm{St}}^R & \hat{\mathrm{St}}^K \\ \hat{\mathrm{St}}^Z & \hat{\mathrm{St}}^A \end{pmatrix} \right\rangle_{\tilde{\mathbf{n}}}. \tag{D.3}$$

The explicit expressions for the retarded, advanced, and "Z" components are

$$\hat{\mathrm{St}}^R = g^K(\mathbf{n})[B;\, A] + g^K(\mathbf{n}) \circ B(\mathbf{n}, \tilde{\mathbf{n}})g^K(\tilde{\mathbf{n}})B(\tilde{\mathbf{n}}, \mathbf{n}),$$

$$\hat{\mathrm{St}}^A = [A;\, B]g^K(\mathbf{n}) - B(\mathbf{n}, \tilde{\mathbf{n}})g^K(\tilde{\mathbf{n}})B(\tilde{\mathbf{n}}, \mathbf{n}) \circ g^K(\mathbf{n}),$$

$$\hat{\mathrm{St}}^Z = 2\delta(t_1 - t_2)[A;\, B] - 2B(\mathbf{n}, \tilde{\mathbf{n}})g^K(\tilde{\mathbf{n}})B(\tilde{\mathbf{n}}, \mathbf{n}),$$

where the (equal-time) commutator is

$$[A;\, B] \equiv A(\mathbf{n}, \tilde{\mathbf{n}})B(\tilde{\mathbf{n}}, \mathbf{n}) - B(\mathbf{n}, \tilde{\mathbf{n}})A(\tilde{\mathbf{n}}, \mathbf{n}).$$

The calculations performed so far are exact. But at the one-loop level, we are interested in terms up to the second order in the fluctuating fields. Then expansion of the exponentials in Eq. (D.1) shows that any product of two functions B is proportional to terms of the form $K\_K\_$, which vanish after averaging over the fluctuating fields; accordingly, we drop such terms. The remaining terms are all commutators, whose explicit (approximate) form is

$$[A;\, B] = 2i(K_-(\tilde{\mathbf{n}}) - K_-(\mathbf{n})). \tag{D.4}$$

We note that the second-order terms cancel each other exactly. The surviving first-order terms lead to Eq. (5.24) for K_-.

For reference, we present the expression for $\hat{\mathrm{St}}^K$ from which Eq. (5.73) and Eq. (5.76) are derived (with

the exception of the last line in Eq. (5.76b), which is a consequence of requirement (5.34) for K_+):

$$\begin{aligned}\hat{\mathrm{St}}^K &= 2\delta(t_1-t_2)[A;\,B] + 2A(\mathbf{n},\tilde{\mathbf{n}})g^K(\tilde{\mathbf{n}})A(\tilde{\mathbf{n}},\mathbf{n})\\ &- g^K(\mathbf{n})A(\mathbf{n},\tilde{\mathbf{n}})A(\tilde{\mathbf{n}},\mathbf{n}) - A(\mathbf{n},\tilde{\mathbf{n}})A(\tilde{\mathbf{n}},\mathbf{n})g^K(\mathbf{n})\\ &+ (g^K(\mathbf{n})B(\mathbf{n},\tilde{\mathbf{n}})B(\tilde{\mathbf{n}},\mathbf{n}) + B(\mathbf{n},\tilde{\mathbf{n}})B(\tilde{\mathbf{n}},\mathbf{n})g^K(\mathbf{n}))\\ &+ g^K(\mathbf{n})\circ B(\mathbf{n},\tilde{\mathbf{n}})g^K(\tilde{\mathbf{n}})A(\tilde{\mathbf{n}},\mathbf{n})\\ &- A(\mathbf{n},\tilde{\mathbf{n}})g^K(\tilde{\mathbf{n}})B(\tilde{\mathbf{n}},\mathbf{n})\circ g^K(\mathbf{n}).\end{aligned} \tag{D.5}$$

APPENDIX E

Derivation of Equations (5.81b) and (5.82b)

We start the derivation by separating the contributions of the Keldysh and retarded/advanced propagators to $\hat{\mathrm{St}}_1^{\mathrm{in}}$ in Eq. (5.75b):

$$\begin{aligned}\hat{\mathrm{St}}_1^K &= -\frac{i}{16}\int\frac{d\mathbf{n}_2 d\mathbf{n}_3}{(\Omega_d)^2}\gamma_{12}^3 g(\bar{t},\delta t,\mathbf{n}_2)\\ &\times\left[2\mathcal{H}^K(\bar{t},\delta t,\mathbf{n}_3,\mathbf{n}_2) - \mathcal{H}^K\left(\bar{t}+\frac{\delta t}{2},0,\mathbf{n}_3,\mathbf{n}_2\right)\right.\\ &\left.- \mathcal{H}^K\left(\bar{t}-\frac{\delta t}{2},0,\mathbf{n}_3,\mathbf{n}_2\right) - (\mathbf{n}_2\longrightarrow\mathbf{n}_1)\right],\end{aligned} \tag{E.1}$$

$$\begin{aligned}\hat{\mathrm{St}}_1^{RA} &= -\frac{i}{32}\int dt_3\int\frac{d\mathbf{n}_2 d\mathbf{n}_3}{(\Omega_d)^2}\gamma_{12}^3\\ &\times[\mathcal{H}^A(t_3,t_2,\mathbf{n}_3,\mathbf{n}_2) - \mathcal{H}^R(t_1,t_3,\mathbf{n}_2,\mathbf{n}_3) - (\mathbf{n}_2\longrightarrow\mathbf{n}_1)]\\ &\times[g(t_1,t_3,\mathbf{n}_1)g(t_3,t_2,\mathbf{n}_2) + g(t_1,t_3,\mathbf{n}_2)g(t_3,t_2,\mathbf{n}_1)].\end{aligned} \tag{E.2}$$

For convenience, we rewrite the Keldysh part in terms of the new time variables $\bar{t}$, δt:

$$\begin{aligned}\bar{t} &= \frac{t_1+t_2}{2},\quad \delta t = t_1 - t_2,\\ g(t_1,t_2) &\longrightarrow g(\bar{t},\delta t),\\ \mathcal{H}^K(t_1,t_1) &\longrightarrow \mathcal{H}^K(\bar{t}+\delta t/2,0),\ldots.\end{aligned} \tag{E.3}$$

We then consider the limit of Eq. (E.1) as $t_2\longrightarrow t_1$; clearly, as $\delta t\longrightarrow 0$, the square bracket vanishes. But we know that in this limit, $g\longrightarrow -2i/\pi\delta t$ (cf. Eq. (5.31)): in principle, there could be a nonvanishing contribution from the first-order expansion of the propagators in δt. The last two Keldysh propagators depend on δt in their first variable, but with opposite signs, and hence their respective first-order terms cancel each other. As for the first propagator, the property $\mathcal{H}^K(1,2) = \mathcal{H}^K(2,1)$ translates into $\mathcal{H}^K(t,\delta t) = \mathcal{H}^K(t,-\delta t)$, which ensures the absence of first-order terms. We conclude that in the limit as $\delta t\longrightarrow 0$, the Keldysh propagator terms vanish. Similarly, from the property $\mathcal{H}^A(1,2) = \mathcal{H}^R(2,1)$, it follows that Eq. (E.2) vanishes for $t_2 = t_1$; this concludes the proof of Eq. (5.81b).

We now turn to Eq. (5.82b). Because $\partial_{t_1} - \partial_{t_2} = 2\partial_{\delta t}$, we must expand the Keldysh propagators to the second order in δt. At this order, the square bracket in Eq. (E.1) is (up to the proper combination of angular variables)

$$\begin{aligned}&\delta t^2\left(\partial_{\delta t}^2 - \frac{1}{4}\partial_{\bar{t}}^2\right)\mathcal{H}^K(\bar{t},0)\\ &= -\delta t^2\lim_{t_2\to t_1}\partial_{t_1}\partial_{t_2}\mathcal{H}^K(t_1,t_2),\end{aligned}$$

where we restored the original time variables. In the operator notation, this is

$$\delta t^2[\partial_t\mathcal{H}^K\partial_t]_{t_2=t_1}.$$

Therefore,

$$\begin{aligned}\lim_{t_2\to t_1}(\partial_{t_1}-\partial_{t_2})\hat{\mathrm{St}}_1^K(t_1,t_2) &= \frac{1}{4\pi}\int\frac{d\mathbf{n}_2 d\mathbf{n}_3}{\Omega_d^2}\gamma_{12}^3\\ &\times\{[\partial_t\mathcal{H}^K\partial_t](t_1,t_1,\mathbf{n}_3,\mathbf{n}_1) - [\partial_t\mathcal{H}^K\partial_t](t_1,t_1,\mathbf{n}_3,\mathbf{n}_2)\},\end{aligned}$$

which proves the first part of Eq. (5.82b).

As regards Eq. (E.2), using the analytic property $\mathcal{H}^A(1,2) = \mathcal{H}^R(2,1)$ again, we conclude that when the derivatives ∂_{t_1}, ∂_{t_2} act on the distribution functions g, the terms in the second line cancel each other. However, there are nonvanishing contributions when a derivative acts on the propagators, such as

$$\begin{aligned}&\int dt_3\int\frac{d\mathbf{n}_2 d\mathbf{n}_3}{\Omega_d^2}\gamma_{12}^3 g(t_1,t_3,\mathbf{n}_1)g(t_3,t_1,\mathbf{n}_2)\\ &\times\{[\partial_t\mathcal{H}^R](t_1,t_3,\mathbf{n}_1,\mathbf{n}_3) - [\partial_t\mathcal{H}^R](t_1,t_3,\mathbf{n}_2,\mathbf{n}_3)\}.\end{aligned}$$

Collecting all the terms, we arrive at

$$\begin{aligned}\lim_{t_2\to t_1}(\partial_{t_1}-\partial_{t_2})\hat{\mathrm{St}}_1^{RA}(t_1,t_2) &= \frac{i}{16}\int dt_3\int\frac{d\mathbf{n}_3 d\mathbf{n}_2}{\Omega_d^2}\gamma_{12}^3\\ &\times\{[\partial_t\mathcal{H}^R](t_1,t_3,\mathbf{n}_2,\mathbf{n}_3) - [\partial_t\mathcal{H}^R](t_1,t_3,\mathbf{n}_1,\mathbf{n}_3)\}\\ &\times[g(t_1,t_3,\mathbf{n}_1)g(t_3,t_1,\mathbf{n}_2) + g(t_1,t_3,\mathbf{n}_2)g(t_3,t_1,\mathbf{n}_1)],\end{aligned}$$

which concludes the derivation of Eq. (5.82b).

APPENDIX F

Elastic Kernels in Terms of the Interaction Propagator D

To compare the kernels in Eq. (6.9) with the corresponding expressions in [17], we use the Fourier transforms of Eqs. (5.51) and (5.55a) to obtain

$$\mathrm{Re}[\hat{\mathscr{L}}^{\rho} - \hat{\mathscr{L}}^{g}] = -\nu\omega\mathrm{Im}[\hat{\mathscr{L}}^{g}\hat{D}^{R}\hat{\mathscr{L}}^{g}]. \tag{F.1}$$

If we assume, as is done in [17], that the Fermi-liquid parameters are independent of the momentum direction, then the interaction propagators $D^{R,A}$ are also independent of it and the above equation becomes

$$\mathrm{Re}[\mathscr{L}^{\rho} - \mathscr{L}^{g}] = -\nu\omega\mathrm{Im}[\mathscr{L}^{g}\rangle D^{R}\langle\mathscr{L}^{g}], \tag{F.2}$$

where we generalized the angular integral notation such that

$$\langle\mathscr{L}^{g} = \int\frac{d\mathbf{n}_1}{\Omega_d}\mathscr{L}^{g}(\mathbf{n}_1, \mathbf{n}_2), \quad \mathscr{L}^{g}\rangle = \int\frac{d\mathbf{n}_2}{\Omega_d}\mathscr{L}^{g}(\mathbf{n}_1, \mathbf{n}_2).$$

We recall that our ghost propagator $\mathscr{L}^{g}$ coincides with the diffusion propagator D in [17].

By substituting Eq. (F.2), we rewrite kernels (6.9)

$$\mathscr{S}^{11}_{\mu\nu}(\omega) = \frac{2}{\pi}\delta_{\mu\nu} \times\int\frac{d^dq}{(2\pi)^d}(\langle\mathscr{L}^{g}\rangle\langle\mathscr{L}^{g}\rangle - \langle\mathscr{L}^{g}\mathscr{L}^{g}\rangle)D^{R}, \tag{F.3}$$

$$\mathscr{S}^{12}_{\mu\nu}(\omega) = -\frac{2d}{\pi}\int\frac{d^dq}{(2\pi)^d}\langle n_\alpha\mathscr{L}^{g}\rangle\langle\mathscr{L}^{g}n_\beta\rangle D^{R}, \tag{F.4}$$

$$\begin{aligned}\mathscr{C}_{\mu\nu}(\omega) = \frac{d}{\pi\tau}\mathrm{Im}\int\frac{d^dq}{(2\pi)^d}D^{R}[&\langle\mathscr{L}^{g}n_\alpha\mathscr{L}^{g}n_\beta\rangle\langle\mathscr{L}^{g}\rangle \\ &- \langle\mathscr{L}^{g}n_\alpha\mathscr{L}^{g}\rangle\langle n_\beta\mathscr{L}^{g}\rangle + \langle\mathscr{L}^{g}n_\alpha\rangle\langle\mathscr{L}^{g}n_\beta\rangle\langle\mathscr{L}^{g}\rangle \\ &- \langle\mathscr{L}^{g}\rangle\langle n_\alpha\mathscr{L}^{g}n_\beta\rangle\langle\mathscr{L}^{g}\rangle] + D^{R}[\langle\mathscr{L}^{g}\rangle\langle n_\alpha\mathscr{L}^{g}n_\beta\mathscr{L}^{g}\rangle \\ &- \langle\mathscr{L}^{g}n_\alpha\rangle\langle\mathscr{L}^{g}n_\beta\mathscr{L}^{g}\rangle - \langle\mathscr{L}^{g}n_\alpha\mathscr{L}^{g}n_\beta\mathscr{L}^{g}\rangle].\end{aligned} \tag{F.5}$$

The first square bracket in the kernel $\mathscr{C}$ can be expressed as

$$\tau[\langle\mathscr{L}^{g}n_\alpha\mathscr{L}^{g}[n_\beta; \hat{\mathrm{St}}_\tau]\mathscr{L}^{g}\rangle + \langle\mathscr{L}^{g}[n_\alpha; \hat{\mathrm{St}}_\tau]\mathscr{L}^{g}n_\beta\rangle\langle\mathscr{L}^{g}\rangle]. \tag{F.6}$$

Using the identity

$$\mathscr{L}^{g}[\mathbf{n}; \hat{\mathrm{St}}_\tau]\mathscr{L}^{g} = [\mathbf{n}; \mathscr{L}^{g}], \tag{F.7}$$

we rewrite it as

$$\tau\Big[\frac{\delta_{\mu\nu}}{d}(\langle\mathscr{L}^{g}\mathscr{L}^{g}\rangle - \langle\mathscr{L}^{g}\rangle\langle\mathscr{L}^{g}\rangle) + \langle\mathscr{L}^{g}\rangle\langle n_\alpha\mathscr{L}^{g}n_\beta\rangle - \langle\mathscr{L}^{g}n_\alpha\mathscr{L}^{g}n_\beta\rangle\Big]. \tag{F.8}$$

In the second square bracket, we use the identity

$$\mathscr{L}^{g}\mathbf{n}\mathscr{L}^{g} = \frac{i}{v_{\mathrm{F}}}\partial_{\mathbf{q}}\mathscr{L}^{g} \tag{F.9}$$

to obtain

$$\frac{i}{v_{\mathrm{F}}}[\langle\mathscr{L}^{g}\rangle\partial_{q_\beta}\langle n_\alpha\mathscr{L}^{g}\rangle - \langle\mathscr{L}^{g}n_\alpha\rangle\partial_{q_\beta}\langle\mathscr{L}^{g}\rangle - \langle\mathscr{L}^{g}n_\alpha\partial_{q_\beta}\mathscr{L}^{g}\rangle]. \tag{F.10}$$

Finally, the identity

$$\mathscr{L}^{g}\mathscr{L}^{g} = -i\partial_\omega\mathscr{L}^{g} \tag{F.11}$$

enables us to conclude that the sum of the three kernels

$$\mathscr{S}^{11} + \mathscr{S}^{12} + \mathscr{C}$$

that determines the correction to the conductivity (Eq. (6.8)) coincides with the combination $(K_0 - K_1 - L_0/v_{\mathrm{F}}\tau)$ in the expression for the conductivity in [17].

APPENDIX G

Inelastic Kernel for the Phase Relaxation Time

We consider a uniform system in which the bosons are assumed to be in equilibrium with the electrons. In other words, the distribution function f is independent of **r**, **n** and the boson–electron collision integral (3.18) must vanish. The latter condition enables us to express the bosonic distributions N^α in terms of f and obtain

$$\Upsilon_{ij}(\varepsilon, \omega) = -\Omega_d\delta(\widehat{\mathbf{n}_i\mathbf{n}_j})\frac{1}{\omega}\int d\varepsilon_1\Psi(\varepsilon, \varepsilon_1; \omega) \tag{G.1}$$

(from now on, irrelevant angular and momentum variables are omitted; all relevant definitions can be found in Section 3). The former condition implies that collision integrals (3.19d), (3.20b), and (3.20f) vanish, and therefore the kinetic equation for (the zeroth harmonic of) f reduces to

$$\partial_t f(\varepsilon; t) = \int d\omega\int d\varepsilon_1 A(\omega)\Psi(\varepsilon, \varepsilon_1; \omega), \tag{G.2}$$

where

$$A(\omega) = \frac{-2}{\nu\pi\omega^2} \times \mathrm{Re}\mathrm{Tr}\left\{[2\hat{\mathrm{St}}_\tau\hat{\bar{\mathscr{L}}}^g + \hat{1}]\hat{\mathrm{St}}_\tau[\hat{\mathscr{L}}^\rho + 3\hat{\mathscr{L}}^\sigma - 4\hat{\mathscr{L}}^g]\right\}. \quad \text{(G.3)}$$

We substitute Eq. (F.1) and a similar relation for the triplet channel (D_T^R being the triplet channel propagator) in the expression for $A(\omega)$; we then use identity (5.44) and obtain

$$A(\omega) = -\frac{2}{\pi\omega}\mathrm{Im}\mathrm{Tr}[\hat{\mathrm{St}}_\tau\hat{\bar{\mathscr{L}}}^g(\hat{D}^R + \hat{D}_T^R)\hat{\mathscr{L}}^g]. \quad \text{(G.4)}$$

Using Eq. (5.44) again, we immediately recover the form of the inelastic kernel $A(\omega)$ given in [27].

REFERENCES

1. G. Wiedemann and R. Franz, Ann. Phys. (Leipzig) **89**, 497 (1853).
2. A. A. Abrikosov, *Fundamentals of the Theory of Metals* (Nauka, Moscow, 1987; North-Holland, New York, 1988).
3. E. M. Lifshitz and L. P. Pitaevskiĭ, *Physical Kinetics* (Nauka, Moscow, 1979; Pergamon, Oxford, 1981).
4. R. Kubo, J. Phys. Soc. Jpn. **12**, 570 (1957).
5. J. M. Luttinger, Phys. Rev. **135**, 1505 (1964).
6. K. Maki, Prog. Theor. Phys. **39**, 897 (1968).
7. R. S. Thompson, Phys. Rev. B **1**, 327 (1970).
8. L. G. Aslamazov and A. I. Larkin, Sov. Phys. Solid State **10**, 875 (1968).
9. L. P. Gorkov, A. L. Larkin, and D. E. Khmelnitskii, JETP Lett. **30**, 228 (1979).
10. B. L. Altshuler and A. G. Aronov, Zh. Éksp. Teor. Fiz. **77**, 2028 (1979) [Sov. Phys. JETP **50**, 968 (1979)].
11. B. L. Altshuler, A. G. Aronov, and P. A. Lee, Phys. Rev. Lett. **44**, 1288 (1980).
12. B. L. Altshuler and A. G. Aronov, in *Electron–Electron Interactions in Disordered Systems*, Ed. by A. L. Efros and M. Pollak (North-Holland, Amsterdam, 1985).
13. I. L. Aleiner and A. I. Larkin, Phys. Rev. B **54**, 14423 (1996).
14. A. M. Finkelstein, Zh. Éksp. Teor. Fiz. **84**, 168 (1983) [Sov. Phys. JETP **57**, 97 (1983)]; Z. Phys. B: Condens. Matter **56**, 189 (1984).
15. A. M. Rudin, I. L. Aleiner, and L. I. Glazman, Phys. Rev. B **55**, 9322 (1997).
16. I. L. Aleiner, B. L. Altshuler, and M. E. Gershenson, Waves Random Media **9**, 201 (1999).
17. G. Zala, B. N. Narozhny, and I. L. Aleiner, Phys. Rev. B **64**, 214204 (2001).
18. I. V. Gornyi and A. D. Mirlin, Phys. Rev. B **69**, 045313 (2004).
19. C. Castellani, C. DiCastro, G. Kotliar, *et al.*, Phys. Rev. B **37**, 9046 (1988).
20. D. V. Livanov, M. Yu. Reizer, and A. V. Sergeev, Zh. Éksp. Teor. Fiz. **99**, 1230 (1991) [Sov. Phys. JETP **72**, 760 (1991)].
21. J. S. Langer, Phys. Rev. **128**, 110 (1962).
22. D. R. Nieven and R. A. Smith, cond-mat/0301451.
23. R. Raimondi, G. Savona, P. Schwab, and T. Lück, cond-mat/0402245.
24. G. Eilenberger, Z. Phys. B **214**, 195 (1968); A. I. Larkin and Y. N. Ovchinnikov, Zh. Éksp. Teor. Fiz. **55**, 2262 (1968) [Sov. Phys. JETP **28**, 1200 (1969)]; Zh. Éksp. Teor. Fiz. **73**, 299 (1977) [Sov. Phys. JETP **46**, 155 (1977)].
25. A. Kamenev and A. Andreev, Phys. Rev. B **60**, 2218 (1999).
26. A. A. Abrikosov, L. P. Gorkov, and I. E. Dzyaloshinskii, *Methods of Quantum Field Theory in Statistical Physics* (Fizmatgiz, Moscow, 1962; Prentice-Hall, Englewood Cliffs, N.J., 1963).
27. B. N. Narozhny, G. Zala, and I. L. Aleiner, Phys. Rev. B **65**, 180202(R) (2002).
28. D. Coffey and K. S. Bedell, Phys. Rev. Lett. **71**, 1043 (1993); G. Y. Chitov and A. J. Millis, Phys. Rev. B **64**, 054414 (2001).
29. A. V. Chubukov and D. L. Maslov, Phys. Rev. B **69**, 121102(R) (2004).
30. A. O. Lyakhov and E. G. Mishchenko, Phys. Rev. B **67**, 041304 (2003).
31. J. W. P. Hsu, A. Kapitulnik, and M. Yu. Reizer, Phys. Rev. B **40**, 7513 (1989).
32. M. Yu. Reizer and A. Sergeev, Phys. Rev. B **50**, 9344 (1994).
33. A. Vilenkin and P. L. Taylor, Phys. Rev. B **18**, 5280 (1978); M. Reizer, A. Sergeev, J. W. Wilkins, and D. V. Livanov, Ann. Phys. (N.Y.) **257**, 44 (1997).

Journal of Experimental and Theoretical Physics, Vol. 100, No. 2, 2005, pp. 370–384.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 127, No. 2, 2005, pp. 412–430.

STATISTICAL, NONLINEAR, AND SOFT MATTER PHYSICS

Collisions and Langmuir Waves in Nonideal Plasmas

I. V. Morozov* and G. E. Norman**

Institute for Thermophysics of Extreme States, Joint Institute for High Temperatures, Russian Academy of Sciences, Moscow, 125412 Russia

*e-mail: *bogous@orc.ru; **henry_n@orc.ru*

Received April 16, 2004

Abstract—Langmuir waves and particle–particle collisions in an equilibrium two-component plasma consisting of nondegenerate electrons and singly charged ions are studied by molecular dynamics simulation. Dispersion of frequency and damping rate is determined for Langmuir waves. A method is outlined for extending the theory of Langmuir waves developed for ideal plasmas to nonideal plasmas. Conductivity and dynamic collision frequency are determined as functions of perturbation frequency. Relaxation of electron and ion energy in nonequilibrium plasmas is analyzed. © *2005 Pleiades Publishing, Inc.*

1. INTRODUCTION

Nonideal plasmas arise in experimental studies of shock waves in gases [1–3] and solids [4–6], Coulomb explosions in semiconductors [7, 8], and other phenomena. However, the development of the kinetic theory of nonideal plasmas is impeded by substantial difficulties, in particular, because screening in nonideal plasmas differs from that in an ideal plasma. Theoretical results obtained for ideal plasmas rely on the assumption that the Debye sphere contains many particles and particle–particle collisions are so weak that the scattering angles are small. In a nonideal plasma, the screening length is comparable to the mean distance between particles, and the Debye sphere formally contains less than one particle (tenths or even hundredths of a particle). Therefore, straightforward extension of theoretical results obtained for an ideal plasma leads to incorrect results in various problems and to substantial disagreement with experimental observations. An adequate model of collisions in nonideal plasmas is necessary to investigate the following problems:

(i) static and dynamic conductivity;

(ii) existence conditions and damping rate for Langmuir waves;

(iii) relaxation mechanisms and equilibration times for nonequilibrium nonideal plasmas;

(iv) absorption of electromagnetic energy by plasmas.

Among the items enumerated above, Langmuir waves are of special importance. An extension of the Landau theory [9] predicts that Langmuir waves cannot propagate in a nonideal plasma [10]. However, both Langmuir and ion-acoustic waves have been revealed in theoretical [11–14] and numerical [15–19] studies of nonideal plasmas. These results were supported by experimental observations reported in [20, 21]. Nevertheless, these findings have never been fully acknowledged. For this reason, the properties of Langmuir waves, such as dispersion and damping rate, remain relatively poorly studied.

Since the theory of ideal plasmas [9, 10, 22–25] is well developed and can be applied to various processes involving Langmuir waves, one should find parameters that could be modified so as to extend the scope of the theory to problems concerning nonideal plasmas. In this paper, the complex effective collision frequency evaluated by means of molecular dynamics (MD) simulation is treated as a free parameter. The existence of Langmuir waves in nonideal plasmas, being supported by MD simulation results, offers an opportunity for applying the wide variety of theoretical methods developed for ideal plasmas.

Even though the Landau formula for energy relaxation rate in a two-temperature system [9] was derived in the approximation of weak collisional damping, it is widely used to evaluate relaxation rates for nonideal plasmas. However, it was found in [4, 6] that discrepancy between experimental data and theoretical predictions amounts to several orders of magnitude. Furthermore, an analysis of conductivity measurements in nonideal plasmas [20, 21] suggests that nonequilibrium states can persist over times much longer than the relaxation time calculated by using formulas derived for ideal plasmas. For this reason, we use MD simulation to analyze the characteristics and durations of relaxation processes in fully ionized, highly nonequilibrium, nonideal plasmas.

In Section 2, we specify the parameter region to be explored, describe the plasma model, and outline the simulation method. In Section 3, we discuss dispersion and damping rate for Langmuir waves in nonideal plasmas. We briefly review the theory of ideal plasmas, demonstrate how it can be extended to describe non-

ideal plasmas, and discuss the scope of the extension. In Section 4, we present numerical results on conductivity and effective collision frequency, which is related to conductivity. Both static case and dependence on perturbation frequency are considered. In Section 5, we analyze energy relaxation in a plasma, starting from various nonequilibrium states. Dependence of collision frequency on the electron-to-ion mass ratio m/M for various degrees of nonideality is discussed in subsections 2.1, 4.3, and 5.3. The principal results of this study are summarized in Section 6.

2. PLASMA MODEL

2.1. Physical Modeling

Physical modeling of a nonideal plasma begins with choice of the scope of a study, including degree of ionization, quantum-mechanical effects, number of particles per cell, and possible dependence of results on M/m.

In this study, we consider a nondegenerate two-component system of singly charged particles, electrons and ions, with masses m and M, respectively. The basic plasma parameters include the nonideality parameter Γ, the number N_D contained in the Debye sphere, the plasma frequency ω_p and the period τ_e of electron plasma oscillations, and the Debye length r_D:

$$\Gamma = \left(\frac{4\pi n_e}{3}\right)^{1/3}\frac{e^2}{k_B T}, \quad N_D = \frac{4\pi r_D^3 n_e}{3},$$

$$\omega_p = \frac{2\pi}{\tau_e} = \sqrt{\frac{4\pi n_e e^2}{m}}, \quad r_D = \sqrt{\frac{k_B T}{4\pi n_e e^2}},$$

where n_e is the electron concentration, e is the electron charge, T is temperature, and k_B is Boltzmann's constant. The nonideality parameter is varied over the interval $0.1 < \Gamma < 4$ ($4 > N_D > 0.03$). The plasma under analysis is fully ionized, which means that $n_e \sim 10^{17}$–10^{24} cm^{-3} and $T \sim 10^4$–10^6 K (Fig. 1) in the case of an equilibrium plasma. In Fig. 1, the corresponding region is hatched. Its left vertical boundary is drawn arbitrarily, because its location depends on the chemical element. A plasma is treated as fully ionized if the degree of ionization is at least 10^{-1}, in which case collisions with neutral particles can be neglected.

A nonequilibrium nonideal plasma can exist at temperatures below 10^4 K, even below room temperature. Supercooled metastable nonideal plasmas were analyzed theoretically in [26, 27]. The results presented in [28, 29] were interpreted in [27] as experimental evidence of the existence of such plasmas. Ultracold plasmas at temperatures of about 1 K have been the subject of recent experimental investigations [30]. The results obtained in the present study apply to low-temperature nonequilibrium nonideal plasmas as well.

Even though the study deals with a nondegenerate plasma, quantum-mechanical effects on pair interactions between charged particles must be taken into account, because only these effects can ensure stability of a nonideal plasma. The key role played by the Fermi–Dirac statistics in preventing a Coulomb system from collapsing was demonstrated in [31]. However, allowance for quantum-mechanical effects in pair interactions is sufficient to ensure thermodynamic stability of a nonideal nondegenerate plasma [32]. In the present study, this is done by using pseudopotentials [33]. The contribution of quantum-mechanical effects to scattering weakens repulsion between electrons and attraction between electrons and ions.

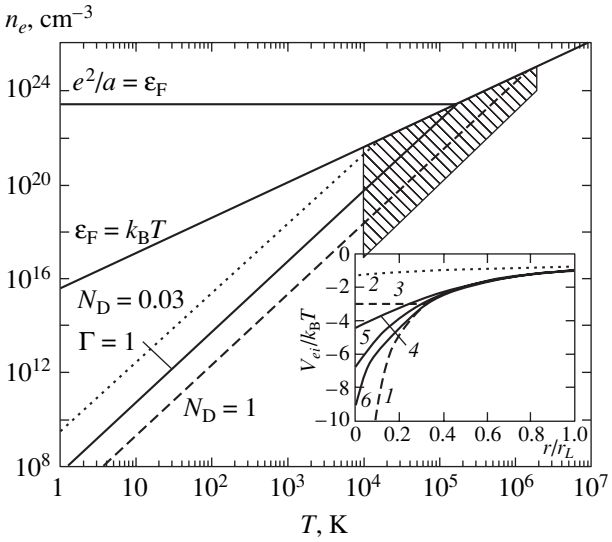

Fig. 1. Temperature–concentration diagram for electrons. Region between $N_D = 1$ and $\varepsilon_F = k_B T$: nondegenerate nonideal plasmas. region between $\varepsilon_F = k_B T$ and $e^2/a = \varepsilon_F$: degenerate nonideal plasmas (ε_F is Fermi energy, $a = (4\pi n_e/3)^{1/3}$ is mean interparticle distance). Hatched region is explored in this study. Electron–ion interaction potentials (inset): (*1*) Coulomb; (*2*) Deutsch [35, 36]; (*3*) [33]; Kelbg [34] for T = (*4*) 10^5, (*5*) 3×10^4, and (*6*) 10^4 K. Landau length $r_L = e^2/k_B T$.

Since we focus on the subsystem of free charged particles, the possibility of bound states forming must be ruled out in the model. This is done by truncating the electron–ion pseudopotential at short distances [33]. We make use of the Kelbg semiclassical pseudopotential [34]. The inset to Fig. 1 shows several electron–ion interaction potentials $V_{ei}(r)$ used in nonideal plasma models. The electron–electron interaction potential $V_{ee}(r)$ corresponding to curve *3* is the Coulomb potential, and the remaining potentials are characterized by less repulsive potentials $V_{ee}(r)$.

The Kelbg potential is advantageous as compared to the truncated potential from [33] and the Deutsch high-temperature potential [35, 36] in that both its value and its first derivative at the origin agree with the analytical solution to the quantum-mechanical two-body scattering problem. At the same time, since the relative vol-

ume occupied by bound states is small in the parameter region considered here [33], the specific form of the potential at small distances is not essential. In the computations reported below, we used the Kelbg potential for $T = 30000$ K (curve 5 in Fig. 1). The accuracy of this approximation deteriorates with decreasing T, and additional computations using different pseudopotentials are required for ultracold plasmas.

For physical reasons, the number of particles N in a cell must be such that the cell size is much larger than r_D for an ideal plasma or much larger than the long-range correlation (screening) radius for a strongly nonideal plasma. If $\Gamma \gtrsim 1$, then $N \approx 100$ is sufficient for computing both thermodynamics of equilibrium plasmas and electron–ion equilibration in nonequilibrium plasmas (see [37]). In studies of Langmuir wave dispersion, the minimal value of N depends on the wavevector k. Therefore, computations with different N should be performed for the same k.

With regard to the value of ion mass, we note that the time complexity of a computation increases with ion mass, because time step and MD run length are determined by electron motion and the mixing time for ion paths, respectively. Therefore, the largest M depends on computer performance (see subsection 2.2). To reduce computational costs, interpolation and extrapolation formulas should be found instead of performing simulations for each value of M.

According to elementary formulas of mechanics of collisions between particles of different mass, it should be expected that the characteristics of collisions would fall into two groups. In one group, the dominant role is played by electron–electron collisions, while collisions with heavy ions can be treated as background processes, and the dependence on M should be neglected for M/m above a certain value. This value is found by performing MD simulations: $M/m = 10^2$ (see subsection 4.3). Another group of collisions is responsible for energy transfer between electrons and ions, and the corresponding characteristics must depend on M/m. The dependence is well known for ideal plasmas, and MD simulations required to find an analogous dependence for nonideal plasmas (see Section 5).

2.2. Numerical Simulation

The apparatus of an MD simulation study includes a integration scheme, a method for calculating forces, an averaging algorithm with an acceptable statistical error, and various diagnostic tools required to calculate observable quantities by using computed electron and ion trajectories.

Equations of motion were computed by using the second-order leapfrog scheme. The time step was such that the total energy was conserved up to 0.1% (see [38] for detailed discussion of energy conservation). The forces acting on particles in a cell were calculated by using periodic boundary conditions and the nearest-image convention.

The total number of particles per cell N varied from 128 to 5000. The thermodynamic variables and correlation functions calculated by using MD simulation results obtained for small N are characterized by considerable statistical errors. Since the time complexity increases as N^2, we averaged the results over an ensemble of I independent initial states. In this approach, the time complexity increases as I, and the statistical error is on the order of $(NI)^{-1/2}$.

According to [38], when small errors of numerical integrations are combined with Lyapunov divergence, an MD trajectory departs from a Newtonian trajectory in a dynamic memory time t_{me} and thus "forgets" its history, while the macroscopic plasma characteristics remain constant. This natural stochastization of MD systems facilitates the generation of a statistical ensemble, because particle configurations computed in the same run at instants separated by time intervals longer than the dynamic memory time are statistically independent. Thus, a single MD run of length $(10^3–10^5)t_{me}$ can be used to generate an ensemble of independent states.

When computing dynamics of electrons over a time interval equal to several periods τ_e of electron plasma oscillation, one can ignore the motion of ions. However, a path with immobile ions cannot be used in averaging over ion configurations. For this reason, $M/m = 100$ is used in Sections 3 and 4. Even though the ion motion does not reflect real plasma dynamics in this case, it weakly affects the electron dynamics over short time intervals and, at the same time, can be used to perform long-time averaging over ion configurations (see subsection 4.3).

In Section 5, we examine energy relaxation, in which the ion mass plays an essential role. To generate an ensemble of initial states, we compute an auxiliary equilibrium path with $M/m = 1$ and use the configurations taken at points separated by t_{me} to compute energy relaxation and perform averaging as described in subsection 5.1. While the ensemble of equilibrium configurations is independent of the value of M/m, the time required to compute an equilibrium run is substantially reduced by using $M/m = 1$. Then, relaxation is computed by using the required mass ratio and appropriately renormalized ion velocities.

Diagnostic tools are described below in the context of the discussion of each particular problem.

3. LANGMUIR WAVES IN EQUILIBRIUM PLASMAS

3.1. Langmuir Waves in Ideal Plasmas

Let us recall the basic expressions for the dielectric constant of an ideal plasma [9, 22], writing them in dimensionless form. All times and frequencies are nor-

malized to the plasma frequency ω_p: $\omega/\omega_p \longrightarrow \omega$, $\nu/\omega_p \longrightarrow \nu$, $\sigma/\omega_p \longrightarrow \sigma$, $\delta/\omega_p \longrightarrow \delta$; all lengths and wavevectors, to the Debye length r_D: $kr_D \longrightarrow k$. Solving the Landau dispersion equation

$$\varepsilon(\omega - i\delta, k) = 0 \tag{1}$$

for the dielelectric constant [9], we find the dispersion and damping rate for Langmuir waves:

$$\omega^2 = 1 + 3k^2, \tag{2}$$

$$\delta = \delta_c + \delta_L = \frac{\nu}{2} + \sqrt{\frac{\pi}{8}}\frac{1}{k^3}\exp\left(-\frac{1}{2k^2} - \frac{3}{2}\right), \tag{3}$$

where $\delta_c = \nu/2$ and δ_L are the collisional and Landau damping rates, respectively.

The effective electron–ion collision frequency ν is expressed as follows [9]:

$$\nu_{ei} = \sqrt{3}Z\Gamma^{3/2}L_e. \tag{4}$$

This formula is not valid when $\Gamma > 1$ because of the divergence of the expression for L_e. When this formula is extended to a nonideal plasma with the constant Coulomb logarithm $L_e = 3.2$ [39], the value of δ_c exceeds unity at $\Gamma > 1.2$, which sets a limit for the domain of existence of Langmuir waves. However, an extrapolation of this kind goes far beyond the scope of the Landau theory. More accurate estimates [14] show that the value of ν reaches a maximum with increasing Γ, remaining less than unity. Thus, we can expect that Langmuir waves exist in nonideal plasmas, and their Landau damping and dispersion characteristics change insignificantly [14].

Since expressions (2) and (3) are obtained in the long-wave approximation, we should use the model proposed in [22] at large k:

$$\varepsilon(\omega, k) = 1 + \frac{1}{k^2}\frac{1 - J_+\left(\frac{\omega + i\nu}{k}\right)}{1 - \frac{i\nu}{\omega + i\nu}J_+\left(\frac{\omega + i\nu}{k}\right)}, \tag{5}$$

where

$$J_+(x) = x\exp\left(-\frac{x^2}{2}\right)\int_{i\infty}^{x}\exp\left(\frac{\tau^2}{2}\right)d\tau. \tag{6}$$

In this case, dispersion equation (1) is solved numerically.

3.2. Dynamic Structure Factor

Langmuir waves in a nonideal plasma can be analyzed by computing the charge–charge dynamic structure factor (DSF). Its real part is related to the imaginary part of dielectric constant as follows [23, 40]:

$$S(\omega, k) = -\frac{k_B T}{4\pi^2 e^2}\frac{k^2}{\omega}\mathrm{Im}\frac{1}{\varepsilon(\omega, k)}. \tag{7}$$

In this study, DSF was determined from MD simulation results. First, we used an equilibrium MD trajectory to calculate density autocorrelation functions for electrons and ions,

$$F_{cd}(k, t) = \frac{1}{N_c N_d} \times\left\langle\sum_{j=1}^{N_c}\sum_{k=1}^{N_d}\exp\{-i\mathbf{k}\cdot(\mathbf{r}_j(t) - \mathbf{r}_k(0))\}\right\rangle, \tag{8}$$

where $\mathbf{r}_j(t)$ is the coordinate of the jth particle; $c, d = e, i$ refer to particle species. Averaging was performed over the initial configurations $\mathbf{r}(0)$ and the directions of the wavevector $\mathbf{k}$ (for isotropic plasma). Since MD cells are finite-sized, the corresponding spectra of $k = 2\pi l/L$ are discrete and the minimal wavevector value is $k = 2\pi/L$, where L is the cell edge and l is an integer. To approach the limit of $k \longrightarrow 0$, MD cells must be enlarged. This leads to an increase in N and, therefore, in time complexity.

The desired DSF is found by performing the Fourier transform:

$$S(\omega, k) = \frac{1}{2\pi}\int_0^\infty F_{zz}(k, t)e^{i\omega t}dt, \tag{9}$$

$$F_{zz} = \frac{1}{2}(F_{ii} + F_{ee} - 2F_{ei}). \tag{10}$$

The DSF calculated for a weakly nonideal plasma in the neighborhood of $\omega = \omega_p$ (see Fig. 2) are in good agreement with the theoretical curve obtained by using expressions (5) and (7).[1]

The high-frequency portion of the DSF can be approximated by a power law: $S(\omega, k) \sim \omega^{-9/2}$. Generally, this asymptotic behavior is determined by the short-range part of the interaction potential. Since it is consistent with a theoretical calculation for the Coulomb potential [18], we conclude that the short-range part of the Kelbg potential used in [34] does not significantly contribute to the DSF.

3.3. Dispersion of Frequency and Damping Rate

Dispersion and damping rate can be evaluated by analyzing the DSF peak in the neighborhood of the

[1] The DSF is calculated up to a normalization factor, because we need only width and location of the peak in the neighborhood of the plasma frequency.

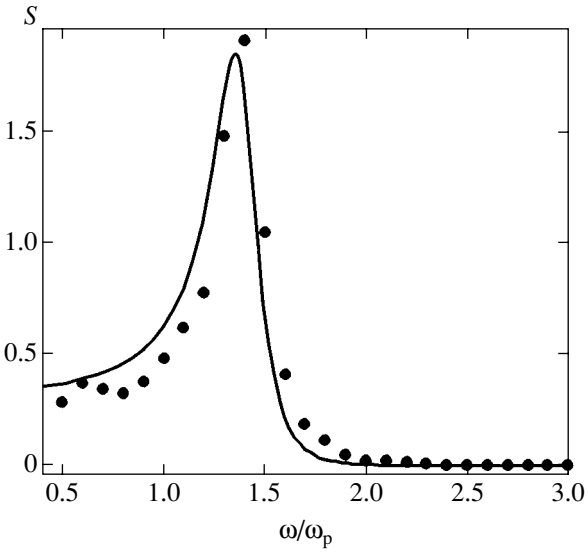

Fig. 2. Dynamic structure factor: MD simulation (circles); theory of nonideal plasmas (5)–(7) (curve). $\Gamma = 0.26$, $N_D = 1.5$, $kr_D = 0.48$.

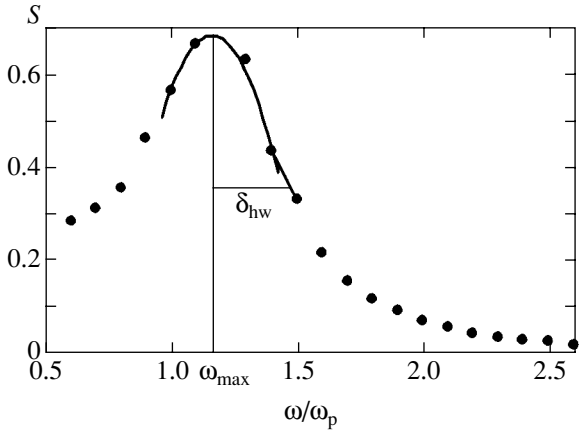

Fig. 3. Determination of DSF peak location ω_{max} and half-width δ_{hw} from MD simulation results (circles).

plasma frequency. This is done by applying the procedure illustrated by Fig. 3, where data points around the peak are approximated by a parabola. The ensuing error depends on the approximation procedure.

The peak location ω_{max} and height $S_{max} = S(\omega_{max}, k)$ obtained by interpolation are used to find the half-width at half-maximum (HWHM) δ_{hw} (see Fig. 3). This is done for the right half-width of the peak, because the curve on its left is distorted by the ion-acoustic maximum located at $\Omega_i = \omega_p\sqrt{m/M} \ll \omega_p$ (not shown). Additional linear interpolation is used to find the intersection of the horizontal line $S = S_{max}/2$ with the MD curve.

Note that the HWHM δ_{hw} for the DSF obtained by using (5) and (7) with $\nu = 2\delta$ is in good agreement with the value of δ at $\delta < 0.8$. For this reason, we also used δ_{hw} to calculate damping rates for nonideal plasmas.

Figures 4 and 5 show ω_{max} and δ_{hw} as functions of wavevector. Since the spectra of k obtained in MD simulations are discrete and bounded from below for a specific N, we performed computations for several values of N. Agreement between numerical results obtained for the same k and different N means that numerical results are independent of boundary conditions; i.e., the MD cell size L was correctly chosen with regard to screening radius.

The behavior of $\delta(k)$ and $\delta_{hw}(k)$ shown in Figs. 4a and 5a at large k demonstrates that Landau damping does not change substantially as we extend our analysis to a nonideal plasma. As $k \longrightarrow 0$, the Landau damping rate vanishes, and the damping rate is determined only by the contribution due to collisions, $\delta_c = \nu/2$. To calculate an effective collision frequency ν, we used model (5) with ν treated as an adjustable parameter. The figures show that theoretical curves accurately approximate MD simulation results under an appropriate choice of ν.

Now, consider the DSF peak location ω_{max} plotted versus k in Figs. 4b and 5b. The dispersion relation found for an ideal plasma ($\nu = 0$) by using formulas (5)–(7) is represented by dashed curves. For collisionless plasmas, $\omega_0(k)$ is localized at the point of the DSF maximum $\omega_{max}(k)$. At the values of ν mentioned above, the DSF maximum may be shifted toward lower frequencies by a significant amount; i.e., the value of $\omega_{max}(k)$ may differ from the corresponding root of the equation $\varepsilon(\omega_0 - i\delta, k) = 0$. Figure 6 shows an example calculated for $\nu = 0.8$ in which the zero of the function $\varepsilon(\omega_0 - i\delta, k)$ is substantially shifted off the real axis. Since the DSF is defined on the real axis, it follows that $\omega_{max} < \omega_0$.

Using values of ν corresponding to the graph presented in Figs. 4a and 5a, one can analytically calculate the shift of the DSF peak relative to $\omega(k)$ and plot curves of $\omega_{max}(k)$ allowing for this effect. These are the solid curves in Figs. 4b and 5b. Calculations show that theoretical curves are in good agreement with MD simulation results when $\Gamma < 3$ ($N_D > 0.04$) (see Fig. 4b). This implies that dispersion in a nonideal plasma is similar to $\omega_0(k)$, and negative dispersion is characteristic only of the location of the DSF peak $\omega_{max}(k)$ obtained for relatively large ν.

The calculated curve and MD simulation results for $\Gamma = 3.84$ are in poor agreement. This implies that dispersion in strongly nonideal plasmas with $\Gamma > 3$ ($N_D < 0.04$) is qualitatively different from that predicted by model (5).

Thus, MD simulation results show that the effective collision frequency ν can be used to extend standard formulas of the theory of ideal plasmas [9, 22, 24] to nonideal plasmas with $\Gamma < 3$. Note also that Langmuir waves were directly observed in a recent experiment with an ultracold nondegenerate nonideal plasma [30].

3.4. Collective Degrees of Freedom

The existence of Langmuir waves in ideal plasmas was used by Bohm [41] to introduce collective variables, or collective degrees of freedom. The relative number of such degrees of freedom in ideal plasmas is on the order of $1/r_D$, i.e., very small. Estimates [14] show that the relative number of collective degrees of freedom increases with nonideality and levels off, reaching a maximum value of about 1/3 at $N_D \approx 1$. The approach developed in this study makes it possible to substantiate the description of nonideal plasmas in terms of collective variables, which requires evaluation of a limit k for Langmuir wave dispersion, rather than the point k_{max} of maximum DSF considered in this study. The shift of the DSF maximum toward lower frequencies leads to the disappearance of the maximum at $\omega_{max}(k_{max}) = 0$ (see Figs. 4b and 5b). In a nonideal plasma, k_{max} is smaller than the maximum admissible k determined by the condition $\delta(k) = \omega(k)$. Thus, Lang-

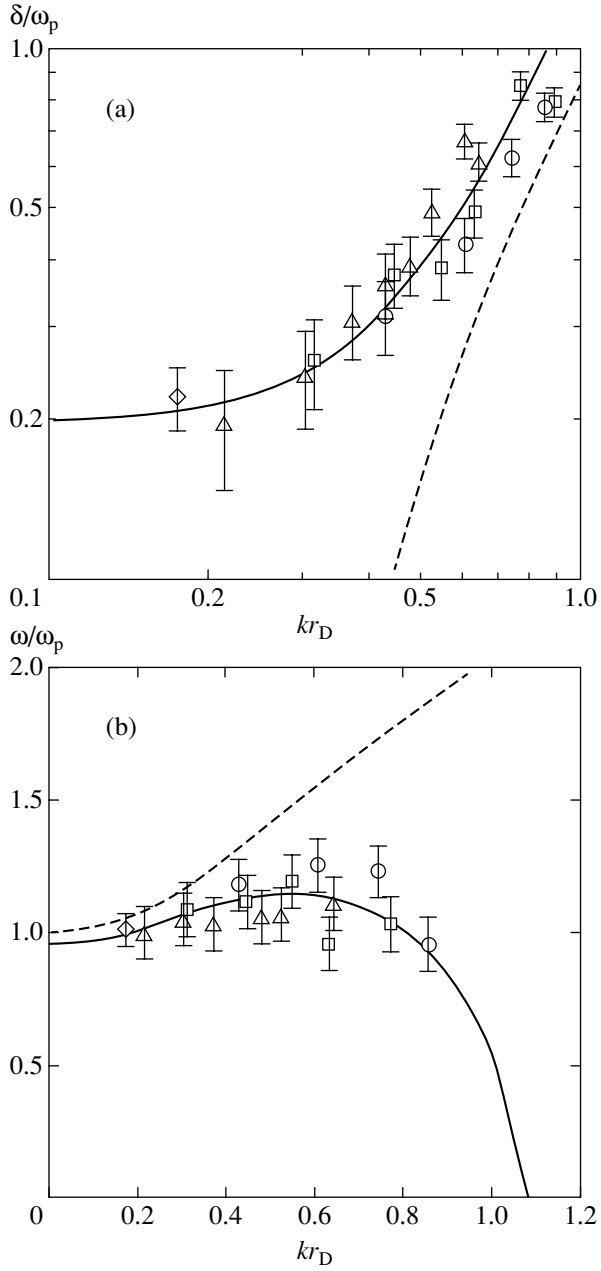

Fig. 4. DSF peak half-width and location for $\Gamma = 1.28$ ($N_D = 0.13$). Dashed curves: (a) Landau damping δ_L (3); (b) Eqs. (5)–(7) for $\nu = 0$. Solid curves: DSF for $\nu = 0.42$, Eqs. (5)–(7). MD results: $N = 3000$ (diamonds), 800 (triangles), 500 (squares), 200 (circles).

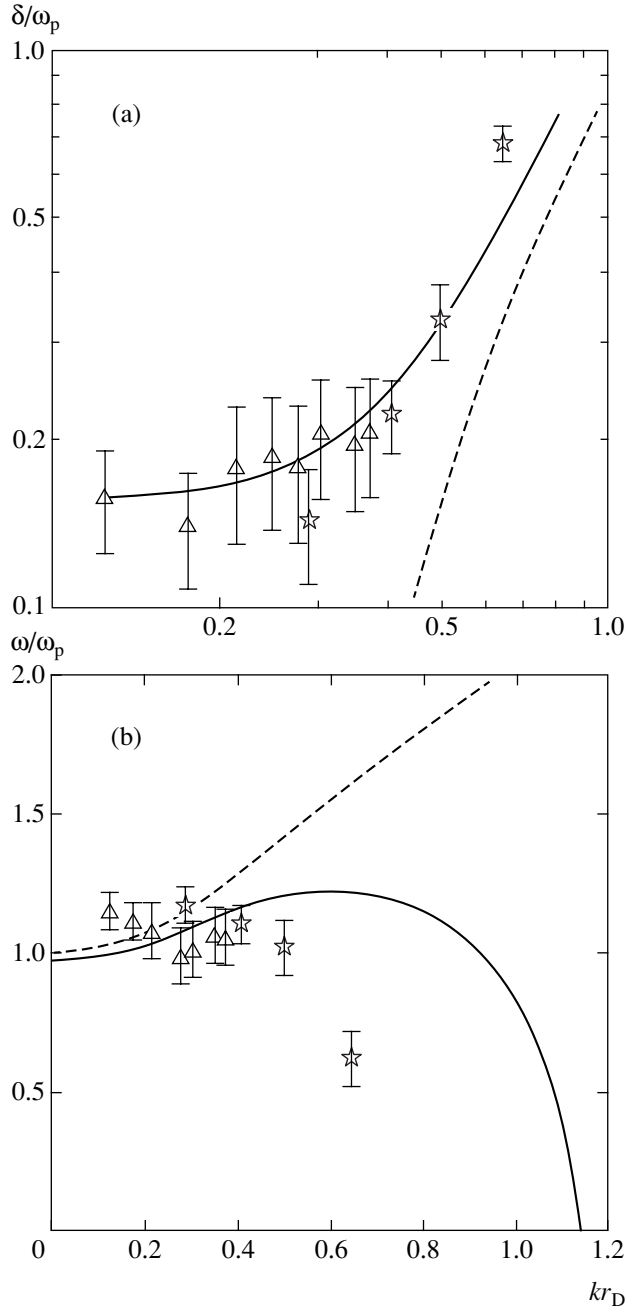

Fig. 5. DSF peak half-width and location for $\Gamma = 3.84$ ($N_D = 0.026$). Curves: $\nu = 0.32$, see caption to Fig. 4. MD results: $N = 800$ (triangles), 128 (stars).

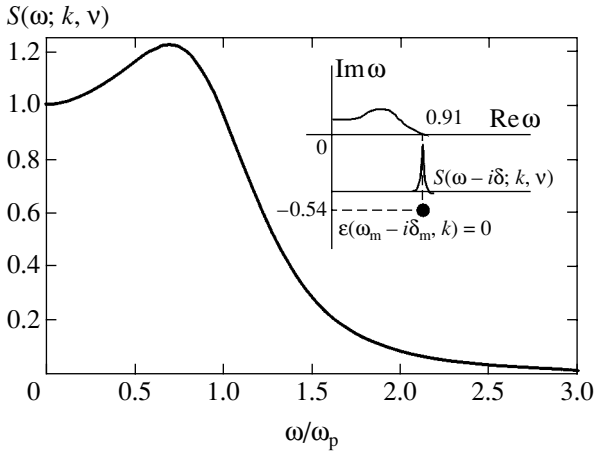

Fig. 6. Dynamic structure factor by Eq. (5): $\nu/\omega_p = 0.8$, $kr_D = 0.2$. Inset illustrates shift of DSF peak relative to zero dielectric constant due to strong damping.

muir waves can exist at $k > k_{max}$. However, their frequency and damping rate cannot be found by the method used here.

As the nonideality parameter increases to $\Gamma \gtrsim 2$, both collisional damping rate and the domain of Landau damping begin to decrease, in agreement with estimates presented in [14]. Thus, with further increase in Γ, collective degrees of freedom become better defined. This tendency suggests that collective degrees of freedom can provide an adequate basis for developing a theory of nonideal plasmas.

4. COLLISIONS IN EQUILIBRIUM PLASMAS

In contrast to ideal plasmas, several definitions of effective collision frequency can be given for nonideal plasmas.

4.1. Collisional Damping of Langmuir Waves

The dependence of $\nu = 2\delta_c$ on the nonideality parameter is illustrated by Fig. 10 (see subsection 4.5). The open circles reach a maximum at $\Gamma \approx 2$ ($N_D \approx 0.07$), remaining smaller than the plasma frequency. This behavior reflects a change in collisional processes. In nonideal plasmas, screening length is comparable to interparticle distance, which decreases with increasing Γ at constant temperature. The effective cross sections of collisional processes decrease accordingly. When $\Gamma \ll 1$, MD simulation results agree with expression (4) given by the Landau theory.

4.2. Autocorrelation Functions

MD simulations are performed to evaluate the autocorrelation functions for various variables, because these functions can be used to determine the properties of equilibrium systems, such as conductivity of an equilibrium plasma, by applying linear response theory. The wavenumber dependence of conductivity is important only if the wavelength of the radiation interacting with the plasma is comparable to the screening radius r_D. Since $\lambda \gg r_D$ in most experimental studies, the present analysis is restricted to the long-wavelength limit $k \longrightarrow 0$, and the dependence on k is not specified. Treating the internal electric field in the plasma as perturbation and the current density **J** as response, we can express the "internal" plasma conductivity σ in terms of current fluctuations at equilibrium [23]:

$$\sigma(\omega) = \beta\Omega_0 \langle J^z; J^z\rangle_{\omega + i\eta}, \tag{11}$$

where Ω_0 is the volume of the system, $\beta = 1/k_B T$, $\langle J^z; J^z\rangle_{\omega + i\eta}$ is the power spectral density for the z component of current,

$$\langle \mathbf{J}; \mathbf{J}\rangle_{\omega + i\eta} = \int_0^\infty \exp[i(\omega + i\eta)t]\langle \mathbf{J}(t)\mathbf{J}(0)\rangle dt. \tag{12}$$

Here, η is introduced to eliminate singularity and is set to zero after taking the thermodynamic limit. These difficulties do not arise in MD computations. The mean current density for an MD system can be calculated directly by using the electron and ion velocities:

$$\mathbf{J}(t) = \frac{1}{\Omega_0}\sum_{i=1}^{N} e_i \mathbf{v}_i(t). \tag{13}$$

To facilitate further analysis, we define the dimensionless autocorrelation function and its Fourier transform as follows:

$$K(t) = \frac{\langle \mathbf{J}(t), \mathbf{J}(0)\rangle}{\langle J^2\rangle}, \tag{14}$$

$$\begin{aligned}\tilde{K}(\omega) &= \frac{4\pi\beta\Omega_0}{\omega_p}\langle J^z; J^z\rangle_\omega \\ &= \omega_p \int_0^\infty \exp[i(\omega)t]K(t)dt.\end{aligned} \tag{15}$$

Here, the coefficients are found by invoking expressions for characteristics of fluctuations in an equilibrium plasma:

$$\begin{aligned}\langle J\rangle &= \frac{e^2}{\Omega_0^2} N\langle v^2\rangle = \frac{3\omega_p^2}{4\pi\Omega_0\beta}, \\ \langle J^z; J^z\rangle_\omega &= \frac{1}{3}\langle \mathbf{J}; \mathbf{J}\rangle_\omega.\end{aligned} \tag{16}$$

In MD simulations, the autocorrelation function $K(t)$ is calculated by using the values of (13) for a set of points in time. The averaging is performed over I = 2000–50000 initial configurations of which 200 to 5000 are statistically independent. The choice of these numbers is dictated by the computing performance.

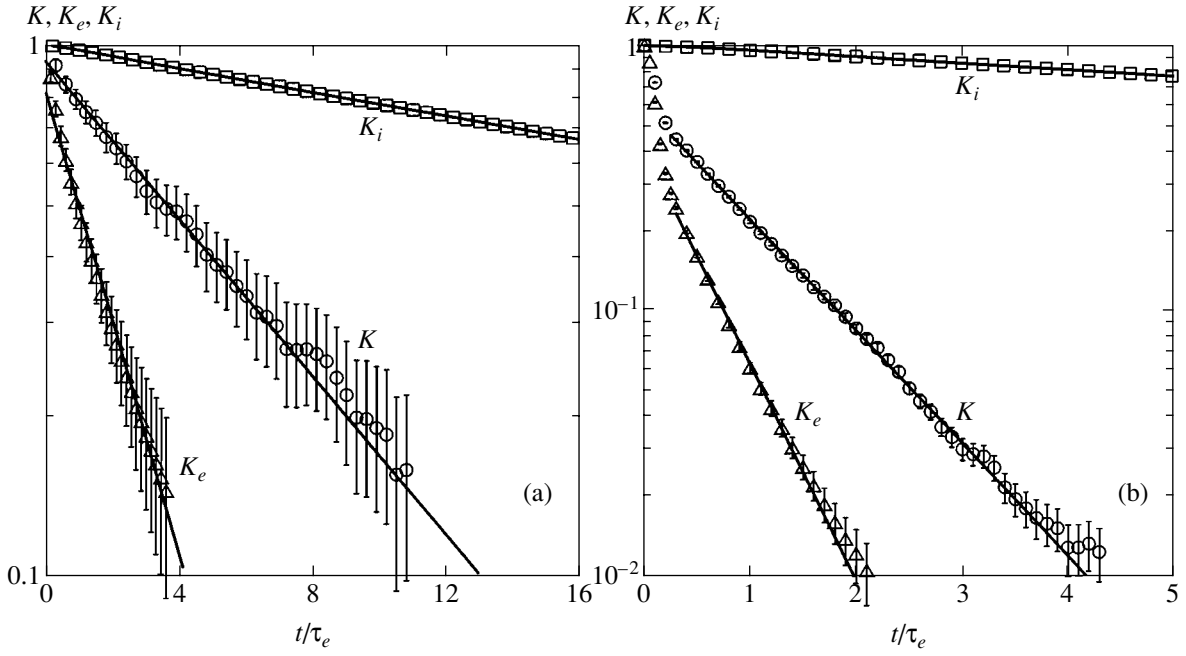

Fig. 7. Velocity autocorrelation functions for electrons, ions, and current: (a) $\Gamma = 0.13$ ($N_D = 4.2$); (b) $\Gamma = 1.28$ ($N_D = 0.13$); $M/m = 100$.

The Fourier transform in (15) is performed numerically. Figure 7b shows the computed current autocorrelation function $K(t)$ and electron and ion velocity autocorrelation functions

$$K_c(t) = \frac{\langle \mathbf{v}_c(t)\mathbf{v}_c(0)\rangle}{\langle \mathbf{v}_c^2\rangle}, \quad c = e, i. \tag{17}$$

Each autocorrelation function shown in Fig. 7b exhibits nonexponential decay over $t \approx 0.2\tau_e$ and exponential decay at longer times. With decreasing degree of nonideality, the decay of $K(t)$ approaches an exponential. At $\Gamma = 0.13$, it can be accurately approximated by $K(t) = \exp(-\nu_j t)$. Thus, the nonexponential decay reflects plasma nonideality.

4.3. Dependence on Ion Mass

Figure 8 shows the exponents in the autocorrelation functions $K_e(t) = \exp(-\nu_e t)$ (electrons), and $K(t) = \exp(-\nu_j t)$ (current) as functions of M/m. Since ν_e and ν_j are virtually independent of the mass ratio, $M/m = 100$ can be used as a good approximation in a study of electron-gas dynamics in a real plasma. The effective collision frequency $\nu(0)$ discussed below is also independent of M at $M/m > 100$.

Computations have shown that the exponent of the ion velocity autocorrelation function $K_i(t) = \exp(-\nu_i t)$ can be approximated by a power law: $\nu_i \propto (M/m)^{-\alpha}$. The value of α may differ from the expected $\alpha = 0.5$, which follows from an expression for the ion thermal velocity $v_{T,i} \propto 1/\sqrt{M}$. MD simulation results yield $\alpha = 0.6 \pm 0.04$ for $\Gamma = 1.28$ and $\alpha = 0.5 \pm 0.04$ for $\Gamma = 3.84$.

4.4. Conductivity

Conductivity of nonideal plasmas has been studied both theoretically [12, 16, 20, 42–46] and experimen-

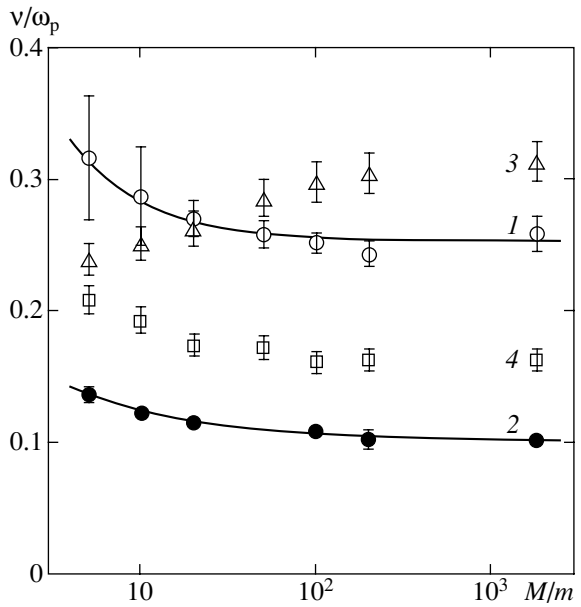

Fig. 8. Dependence on ion mass: effective collision frequency $\nu(0)$ for (*1*) $\Gamma = 1.28$ ($N_D = 0.13$), (*2*) $\Gamma = 3.84$ ($N_D = 0.026$). Exponents in autocorrelation functions for $\Gamma = 1.28$: (*3*) electrons, ν_e; (*4*) current, ν_j. Solid curves: approximation $\nu(M) = \nu(\infty)(1 + B(M/m)^{-\xi})$.

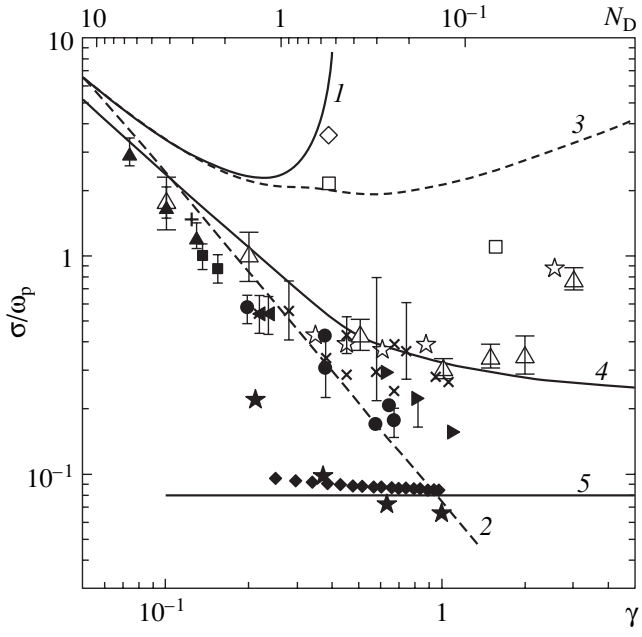

Fig. 9. Static conductivity vs. nonideality parameter $\gamma = e^2(n_e + n_i)^{1/3}/k_BT \approx \Gamma/1.28$. Closed symbols: experiment, including [47] (diamonds). Open symbols: MD results, including present study (triangles) and [16] (diamond and squares); analytical calculations [48] (stars). Calculated curves: (*1*) standard Landau theory, (*2*) $L_e = 3$, (*3*) $L_e = \sqrt{\Lambda^2 + 1}/2$ ($\Lambda = r_D/r_L$ [14]; (*4*) [43]; (*5*) $\sigma = \omega_p/4\pi$.

tally (e.g., see reviews in [12, 21, 39]). However, certain difficulties have yet to be resolved. In [20, 42], a vast amount of experimental data concerning conductivity was examined for various materials over ranges of density and temperature. The contribution of Coulomb interactions to conductivity is shown in Fig. 9 as a function of the nonideality parameter

$$\gamma = \frac{e^2(n_e + n_i)^{1/3}}{k_BT} \approx \frac{\Gamma}{1.28}.$$

The highest conductivity values correspond to curves *1* and *3*, which were calculated by using formulas for an ideal plasma. Curve *4* in Fig. 9 represents the conductivity evaluated for an equilibrium plasma in [12, 43].

The figure illustrates the wide scatter of results observed in experiments (closed symbols), whereas theory predicts a unique dependence $\sigma(\gamma)$. It was hypothesized in [44] that an important role in nonideal plasma dynamics is played by the electron scattering by collective modes. In laboratory plasmas, these modes can be heated to a suprathermal level. In this case, the conductivity of a nonequilibrium nonideal plasma is lower than that of an equilibrium plasma. The fact that almost all experimental data lie below curve *4* indicates that there is a certain deviation from nonequilibrium in any experiment. Horizontal line *5* corresponds to the highest attainable degree of excitation of Langmuir waves [43], in agreement with experimental results reported in [47].

The results obtained in the present study are represented by open triangles in Fig. 9. Conductivity was calculated by substituting autocorrelation function (13) into (11). It is clear that these results are in good agreement with the estimates from [12, 43] (curve *4*) at $\gamma < 2$ ($\Gamma < 2.5$, $N_D > 0.03$). Thus, the theory presented in [43] and the MD results obtained in this study are mutually consistent. Good agreement is also achieved with the results of independent computations based on a liquid-metal model [48].

4.5. Effective Collision Frequency

The effective collision frequency $\nu(\omega)$ is defined by the generalized Drude formula for σ or the corresponding expression for the dielectric constant:

$$\sigma(\omega) = \frac{\omega_p^2}{4\pi}\frac{1}{\nu(\omega) - i\omega}, \quad (18)$$

$$\varepsilon(\omega) = 1 + \frac{4\pi i\sigma(\omega)}{\omega} = 1 - \frac{\omega_p^2}{\omega(\omega + i\nu(\omega))}. \quad (19)$$

The term *collision frequency* should not be interpreted literally. It is just a convenient characteristic of collisional processes in plasmas. In the theory of ideal plasmas, ν is given by (4) and is independent of perturbation frequency. It is shown below that this approximation is not valid for nonideal plasmas. However, if $\nu(\omega)$ is defined as a complex quantity depending on the perturbing-field frequency ω [11, 49], then Eqs. (18) and (19) only express the algebraic relationship between $\nu(\omega)$, $\sigma(\omega)$, and $\varepsilon(\omega)$ without imposing any restrictions associated with the Drude's theory.

First, we consider the static limit. The zero-frequency value $\nu(0)$ is inversely proportional to (15). In the case of an ideal plasma, it is equal to the exponent ν_j in the current autocorrelation function. In the case of a nonideal plasma, its value is higher than ν_j because of the nonexponential part of $K(t)$.

Figure 10 compares the dependence of $\nu(0)$ on the nonideality parameter with the prediction of the Landau theory [9]. At low degrees of nonideality, the MD simulation results are in good agreement with the formulas for ideal plasmas. However, when extrapolated into the domain of nonideality, even with a constant Coulomb logarithm, these formulas substantially overestimate the effective collision frequency. The existence of a maximum of $\nu(\Gamma)$ agrees with the asymptotic analysis presented in [14]. Its steep decrease can be explained by inaccuracy of a pseudopotential model. The MD simulation results obtained for $\nu(0)$ are in good agreement with independent calculations performed in [50] for $\Gamma < 3$.

The inset to Fig. 11 shows the real part of the dynamic collision frequency ν as a function of pertur-

bation frequency. In ideal plasmas, $\nu(\omega)$ is virtually constant. This is clear from Fig. 10, where the difference between $\nu(\omega_p)$ and the static collision frequency $\nu(0)$ increases with Γ. The increase in the maximum value of $\nu(\omega)$ with Γ is followed by a decrease because it is related to $\nu(0)$.

The numerical values of $\nu(\omega_p)$ are in good agreement with $\nu = 2\delta_c$, where δ_c is the collisional damping rate for Langmuir waves (see subsection 4.1). Thus, the model is self-consistent.

4.6. Behavior of $\nu(\omega_p)$

Turning back to $\nu(\omega_p)$, we should note that the existence of a maximum of the effective collision frequency (plotted versus degree of nonideality in Fig. 11) can be explained as follows. Define a collision time as the time of electron–ion interaction inside a "screening" sphere of radius r_s:

$$\tau_{col} = \frac{2r_s}{v_T}, \tag{20}$$

where $v_T = \sqrt{3k_BT/m}$ is the electron thermal velocity.

In an ideal plasma, the value of r_s is determined by the Debye length r_D. When $\Gamma > 1/3$, the latter parameter is meaningless since it is smaller than the mean interparticle distance $a = \sqrt{3\Gamma}\, r_D$, and the average screening length should be a rather than r_D (see [33]). This conjecture is supported by the behavior of the electron–ion pair correlation functions shown in Fig. 12, which were calculated over MD paths with $T = (200–1000)\tau_e$ by using 5000 to 20000 particle configurations.

The frequency $\omega_{col} = \tau_{col}^{-1}$ does not play any significant role in ideal plasmas, where close collisions are rare. As their role increases, one should expect a resonant absorption peak at $\omega = \omega_{col}$.

Substituting the screening radius

$$r_s = \begin{cases} -r_D, & \Gamma < 1/3 \\ -a, & \Gamma \geq 1/3 \end{cases} \tag{21}$$

into (20), we find that $\omega_{col}/\omega_p \sim \Gamma^{-1/2}$ for nonideal plasmas ($\Gamma \geq 1/3$) (Fig. 11), which is qualitatively consistent with MD simulation results.

To evaluate ω_{col} more accurately, equations of motion for an electron moving inside a screening sphere in the centrally symmetric field generated by an ion should be integrated [34]:

$$\tau_{col} = 2\int_{r_{min}}^{r_s} \frac{dr}{\sqrt{\frac{2}{m}\left(E - \frac{mv_T^2r^2}{2\rho^2} - V_{ei}(r)\right)}}, \tag{22}$$

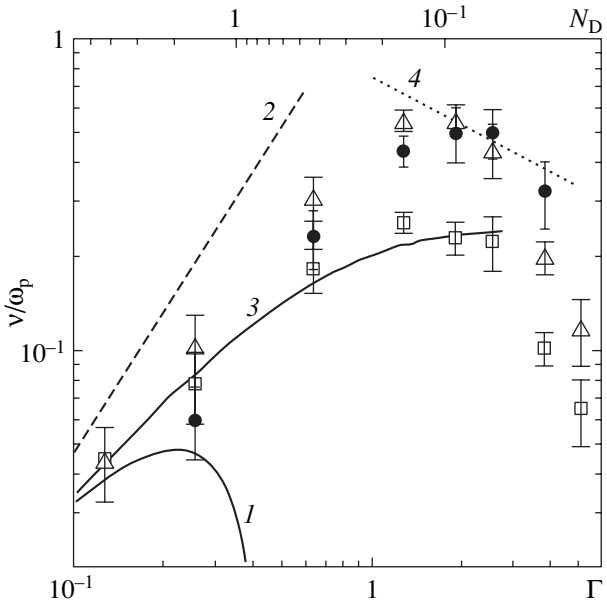

Fig. 10. Effective collision frequencies: from current autocorrelation function at $\omega = 0$ (squares) and $\omega = \omega_p$ (triangles); Langmuir wave damping rate $2\delta_c$ (circles); (*1*) theory of ideal plasmas, (*2*) $L_e = 3.2$: (*3*) [50]; (*4*) asymptotic result from [14] for $\Gamma > 1$.

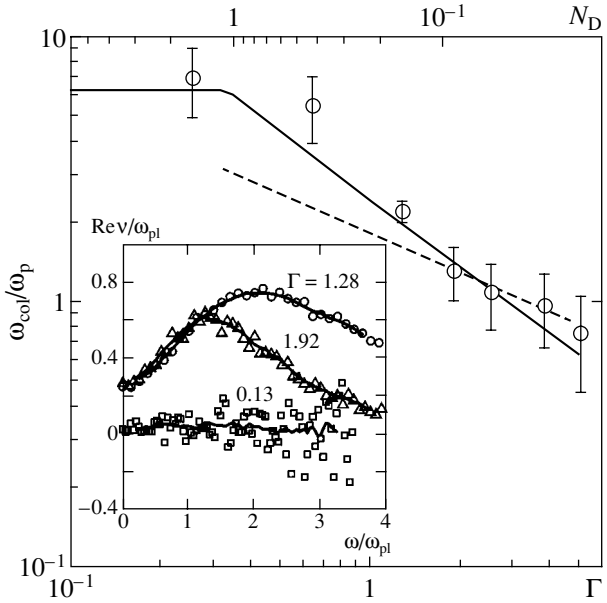

Fig. 11. Shift of maximum $\mathrm{Re}\nu(\omega)$ toward low frequencies vs. nonideality parameter: MD results for maximum $\mathrm{Re}\nu(\omega)$ (symbols); qualitative estimates for ω_{col} by (20) (dashed curve) and (22) (solid curve). Inset: real part of effective collision frequency vs. perturbation frequency for Γ shown at curves. Solid curves: averaged MD results.

where $E = mv_T^2/2$ is the electron energy at infinity, ρ is the impact parameter corresponding to the maximum value of τ_{col}, and r_{min} is the minimal distance between the electron and the ion. Figure 11 shows the calculated

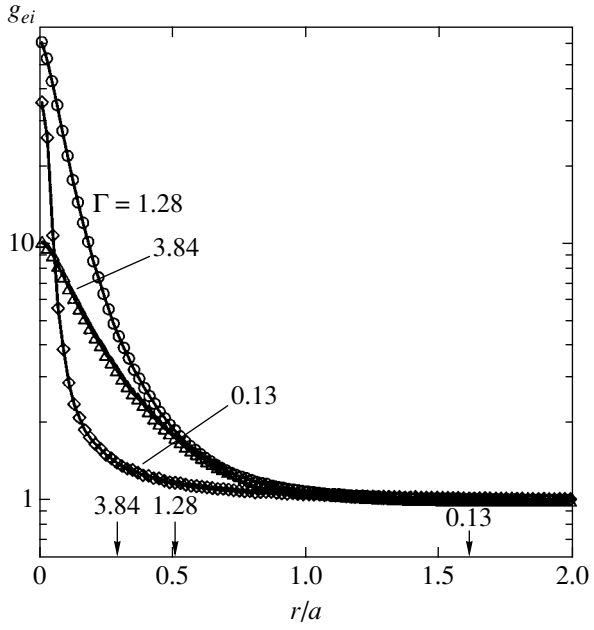

Fig. 12. Electron–ion pair correlation function (a = interparticle distance). Arrows: Debye length r_D for Γ shown at curves.

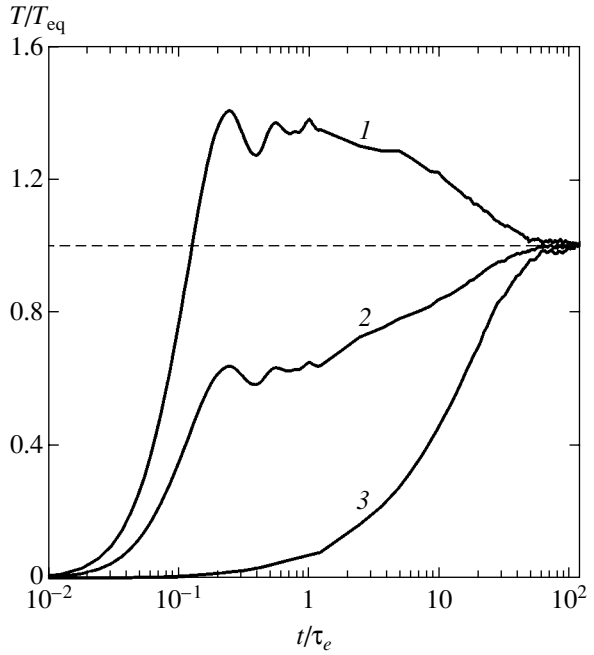

Fig. 13. Electron energy relaxation to equilibrium: (*1*) T_e (electrons), (*2*) T (mean kinetic energy), (*3*) T_i (ions) for Γ = 3.3, M/m = 100.

$\omega_{col} = \tau_{col}^{-1}$. Good agreement with simulation results corroborates the proposed interpretation.

The discussion above provides only a qualitative explanation of the behavior of ν(ω). A more rigorous theoretical analysis can be found in [49].

5. ENERGY RELAXATION IN NONEQUILIBRIUM PLASMAS

5.1. General Mechanism of Relaxation

In most experiments, the plasma produced by ionization is far from equilibrium at the initial stage. The details of a nonequilibrium state depend on experimental conditions. In the plasma produced in a shock front, the ion temperature is much higher than the electron temperature. Conversely, target ionization by a short laser pulse creates a high-temperature electron gas, while ions are relatively cool. According to available experimental data, the theory of ideal plasmas fails when $\Gamma \approx 0.2$ ($N_D \approx 2$).

In this study, we examine energy relaxation for several types of initial plasma nonequilibrium: $T_e \ll T_i$ (plasma behind a shock front), $T_e \gg T_i$ (laser heating of plasmas), and $T_e = T_i = 0$ (a model example). For $T_e \gg T_i$, we consider two cases of ion distribution: a crystal lattice and a distribution corresponding to an equilibrium state.

To create an ensemble of nonequilibrium states, we use the procedure described in [37]. At the first stage, we compute a sufficiently long equilibrium MD run. Then, we take 50 to 200 statistically independent states on the equilibrium trajectory and kick the system out of equilibrium by stopping the motion of electrons, ions, or all particles. When a crystal lattice is used, the equilibrium trajectory is computed for immobile ions.

As an illustration of relaxation process, consider the case when both electrons and ions are stopped at the initial moment (see Fig. 13). This example demonstrates that, in addition to collisional relaxation characteristic of ideal plasmas, relaxation in a nonideal plasma includes evolution toward equilibrium between the mean kinetic energy $T = (T_e + T_i)/2$ of electrons and ions (curve *2*) and the potential energy U of their interaction ($U = 0$ for ideal plasmas). In the course of a computation, we hold $2T + U$ = const. Note that oscillatory behavior may be observed at the initial stage of the process.

At an early stage, the electron energy rapidly increases (see Fig. 13), and a Maxwellian energy distribution is approached, with temperature equal to two-thirds of the mean kinetic energy (see Fig. 14). For example, when $M/m = 100$ and $\Gamma = 3.3$ ($N_D = 0.03$), relaxation to the Maxwellian velocity distribution with temperature corresponding to the mean kinetic energy of particles requires a time of about τ_e for electrons (Fig. 14a) and $15\tau_e$ for ions. At the next stage, electron–ion relaxation is observed in a two-temperature system. In parallel with this process, conversion of potential energy into kinetic energy continues, as manifested by the continuing growth of kinetic energy (curve *2* in Fig. 13). Note that the ion velocity distributions slightly exceeds the Maxwellian distribution in the tail region ($E > 2k_BT$).

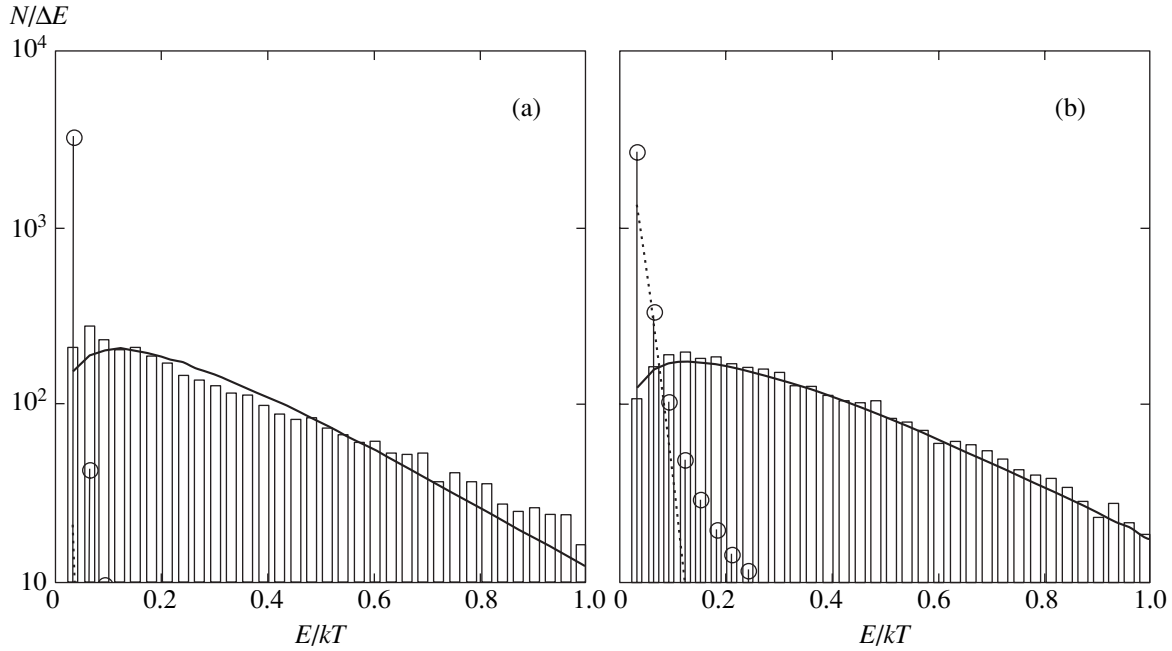

Fig. 14. Electron (columns) and ion (circles) energy distributions at t = (a) $0.15\tau_e$ and (b) $0.58\tau_e$. Curves: Maxwell distributions for electrons (solid) and ions (dashed) at temperatures corresponding to respective mean kinetic energies. $T_e(0) = T_i(0) = 0$, $\Gamma = 3.3$, $M/m = 100$.

Thus, the process can be characterized by two time scales: the time $\tau_{n\mathrm{B}}$ of initial (nonexponential) relaxation of $\Delta T = |T_e - T_i|$ and the time τ_B of exponential decay $\Delta T \propto \exp(-t/\tau_\mathrm{B})$. The latter process requires a time of about $5\tau_e$ and begins before the Maxwellian distribution for ions is approached in the example illustrated by Fig. 13.

Other examples of time evolution of the electron and ion energies were discussed in [37, 51]. Here, we focus on the dependence of $\tau_{n\mathrm{B}}$ and τ_B on Γ, which was not considered in [37, 51].

5.2. Nonexponential Relaxation Time

Figure 15 shows $\tau_{n\mathrm{B}}$ as a function of the nonideality parameter for three sets of initial conditions: $T_i(0) = 0$, $T_e(0) = 0$, and $T_e(0) = T_i(0) = 0$. The difference in $\tau_{n\mathrm{B}}$ between these cases does not exceed an order of magnitude. Irrespective of the initial conditions, $\tau_{n\mathrm{B}}$ increases with Γ.

To elucidate the physical meaning of $\tau_{n\mathrm{B}}$, we recall an equation describing relaxation of the difference between electron and ion temperatures in an ideal plasma (e.g., see [25]):

$$\frac{d\Delta T}{dt} = -\nu_{ei}^{\varepsilon}\Delta T. \qquad (23)$$

Here, the relaxation time $\tau_{ei} = (\nu_{ei}^{\varepsilon})^{-1}$ corresponds to electron–ion collisions:

$$\tau_{ei} = \frac{3}{8}\frac{(mT_i + MT_e)}{\sqrt{2\pi mM}}\frac{1}{Z(Z+1)e^4 n_e L_e}. \qquad (24)$$

The initial nonexponential relaxation cannot be described by solving Eq. (23). Since (23) is derived from a kinetic equation by assuming that the system is stochastic and individual particle–particle collisions are statistically independent, it should be valid only at $t > t_{\mathrm{me}}$, where t_{me} is the dynamic memory time for the electron subsystem [38]. Indeed, according to Fig. 15, the nonexponential relaxation time is comparable to t_{me} in all cases. By definition, τ_{me} is the upper limit for the validity of solutions to Cauchy problems computed by MD simulation [38, 52]. Therefore, the nonexponential and exponential relaxation stages correspond to dynamic and stochastic regimes. No appreciable dependence of $\tau_{n\mathrm{B}}$ or t_{me} on M has been found for $M/m > 100$. The nonexponential relaxation stage vanishes in an ideal plasma as $\Gamma \longrightarrow 0$.

5.3. Exponential Relaxation Time

The dependence of τ_B on M/m is not linear, as it would be in the case of an ideal plasma. However, it can be approximated by a power law: $\tau_\mathrm{B} \propto (M/m)^\alpha$ (see Fig. 16). Both τ_B and α are virtually independent of the type of initial deviation from equilibrium. As

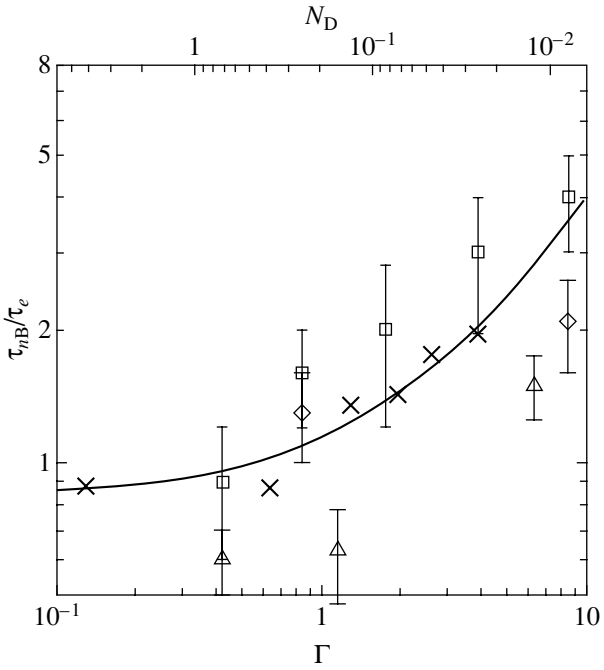

Fig. 15. Nonexponential relaxation time vs. nonideality parameter: $T_i(0) = 0$ (squares), $T_e(0) = 0$ (diamonds), $T_e(0) = T_i(0) = 0$ (triangles). Dynamic memory time t_{me} for equilibrium plasma is shown by crosses. Curve: linear approximation for t_{me}. $M/m = 100$.

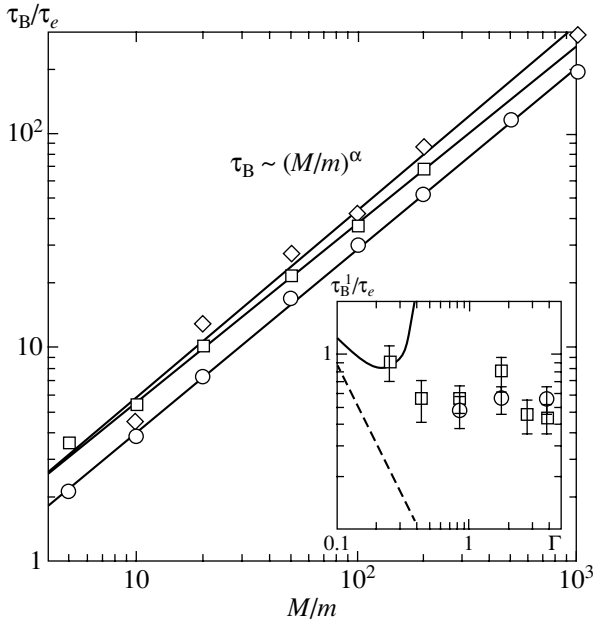

Fig. 16. Exponential relaxation time vs. ion mass for Γ = 3.3: $T_e(0) = 0$ (diamonds); $T_i(0) = 0$ (squares); $T_i(0) = 0$, crystal (circles); power-law approximations (lines). Curves (inset): Landau theory (solid); $L_e = 3.2$ (dashed).

$\Gamma \longrightarrow 0$, $\alpha(\Gamma)$ approaches the theoretical value $\alpha = 1$. For a nonideal plasma, it can be approximated by the expression

$$\alpha = 1 - 0.15\Gamma + 0.035\Gamma^2, \quad \Gamma < 4. \tag{25}$$

Its minimal value, $\alpha = 0.84$, is reached at $\Gamma = 2.1$. An effective electron mass can be introduced tentatively for a nonideal plasma.

To analyze the dependence of τ_B on Γ, we represent it as

$$\tau_B(\Gamma, M) = \tau_B^1(\Gamma)\left(\frac{M}{m}\right)^{\alpha(\Gamma)}. \tag{26}$$

The dependence $\tau_B^1(\Gamma)$ is shown in the inset to Fig. 16. Using (26), one extend it to any M/m. The error in α does not exceed 5%, which entails an error of about 40% in τ_B calculated for aluminum. This accuracy is sufficient for order-of-magnitude relaxation-time estimates.

Figure 16 shows that the relaxation rate evaluated by extrapolating the results obtained for an ideal plasma [25] with a constant Coulomb logarithm (dashed curve) may exceed that obtained for a nonideal plasma by several orders of magnitude. This agrees with the results of experiments on relaxation in shock-compressed aluminum and silicon [4, 6]. Quantitative comparison is impossible because of the role played by degeneracy under the conditions considered in [4, 6].

The results of our analysis of energy relaxation rates can be compared with those concerning momentum relaxation. For an ideal plasma, the energy and momentum relaxation times, τ_B^{id} and ν^{-1}, are related as follows [9, 25] (in reduced units):

$$\tau_B^{id} = \frac{M}{m}\frac{3\Gamma^{-3/2}}{8\sqrt{6\pi}L_e} = \frac{M}{m}\frac{1}{\sqrt{8\pi}}\nu^{-1}. \tag{27}$$

MD simulation results show that relation (27) is not valid for nonideal plasmas,

$$\tau_B^1 > \tau_B^* = \frac{1}{\sqrt{8\pi}}\nu^{-1}(0), \tag{28}$$

where $\nu(0)$ is the effective collision frequency found in Section 4. In the range of nonideality parameter examined in this study, τ_B^1 and τ_B^* differ by a factor of 2 to 3.

Thus, relaxation in nonideal plasmas is characterized by several effective time scales.

6. CONCLUSIONS

MD simulations of equilibrium and nonequilibrium, highly ionized, nondegenerate, nonideal plasmas have been performed for $0.1 < \Gamma < 4$ ($4 > N_D > 0.03$) with the use of the Kelbg pseudopotential. An analysis of the results obtained leads to the following conclusions.

It is shown that the peak in the dynamic structure factor corresponding to Langmuir waves persists over the entire range of the nonideality parameter examined in this study. Dispersion of the frequency and damping rate is determined for Langmuir waves. It is found that the damping rate has a maximum at $\Gamma \approx 2$ ($N_D \approx 0.07$), remaining lower than the plasma frequency by a factor of four. Thus, extrapolation of the expressions describing collisional damping in the theory of ideal plasmas to $\Gamma > 0.1$ ($N_D < 10$) with a constant L_e leads to incorrect results. However, no significant difference in characteristics of Landau damping is found between ideal and nonideal plasmas in the parameter range under study.

It is shown that expression (5) for the dielectric constant of an ideal plasma [22] can be applied, with a modified collision frequency, to describe Langmuir waves in nonideal plasmas for $\Gamma < 3$ ($N_D > 0.04$). It is noted that the negative dispersion of the dynamic structure factor with respect to frequency is due to a shift of the zero of the dielectric constant into the complex plane, which is explained by collisional damping in nonideal plasmas. The MD simulation results obtained for $\Gamma > 3$ have yet to be given a theoretical interpretation.

It is shown that, in contrast to ideal plasmas, particle–particle collisions cannot be described in terms of a single effective collision frequency in the framework of the theoretical model employed here. In particular, equilibrium plasmas are found to be characterized by a static collision frequency (at $\omega = 0$) and a collision frequency for a perturbing field with $\omega = \omega_p$. The latter is determined by two independent methods: from the damping rate for Langmuir waves with small k and from the current autocorrelation function. Agreement between the respective results corroborates the approach developed in this study. Collision frequency is found as a function of ω for several values of the nonideality parameter, and a qualitative explanation of its behavior is presented. The results concerning the static collision frequency are consistent with the theory of ideal plasmas at $\Gamma < 0.3$ ($N_D > 1.2$), with asymptotic estimates presented in [14] at $\Gamma > 2$ ($N_D < 0.07$), and with the theoretical calculations performed in [50] at intermediate $\Gamma \lesssim 2$.

Model computations performed for ions having a small mass have shown that collision frequency is independent of the ion-to-electron mass ratio M/m for $M/m > 10^2$. Thus, both collision frequencies and dispersion laws obtained here are valid for any realistic value of M/m. The static conductivity of a nonideal plasma is calculated for $4 > N_D > 0.03$. The results are compared with those obtained in experimental data and other theoretical studies.

Relaxation of electron and ion energy in nonequilibrium plasmas is analyzed for three types of initial plasma nonequilibrium: $T_e \ll T_i$, $T_e \gg T_i$, and $T_e = T_i = 0$. It is shown that evolution toward equilibrium between the total kinetic energy of electrons and ions and the energy of their interaction continues during the entire relaxation period (this process does not take place in ideal plasmas). Another distinction of ideal plasmas is the existence of an initial relaxation stage characterized by a nonmonotonic time dependence of $\Delta T = |T_e - T_i|$. This dependence cannot be described by a kinetic equation with an effective collision frequency. The duration τ_{nB} of this stage is related to the dynamic memory time t_{me} for electrons. At $t > \tau_{nB}$, the variation of ΔT follows an exponential law, $\Delta T \propto \exp(-t/\tau_B)$; i.e., stochastic relaxation is observed. The relaxation time τ_B is calculated, and interpolation formula (26) is proposed for a nonideal plasma with ions of arbitrary mass. The proposed formula, $\tau_B(M) \sim (M/m)^{\alpha(\Gamma)}$ differs from the linear scaling characteristic of ideal plasmas. The exponent $\alpha(\Gamma)$ is determined. The factor $\tau_B^1(\Gamma)$ independent of M/m correlates with collision frequency ν for equilibrium plasmas, but the relation between τ_B^1 and ν differs from that characteristic of ideal plasmas.

An analysis of the MD simulation results concerning collision frequencies and relaxation times shows that both momentum and energy relaxation times exceed those obtained by extrapolating the Landau theory by several orders of magnitude. Momentum relaxation has a direct effect on microscopic properties, such as conductivity, dielectric constant, and reflectivity. The relatively large values of the estimated relaxation time implies that effects due to nonequilibrium, such as suprathermal excitation of Langmuir waves [20], may substantially contribute to measurement results.

ACKNOWLEDGMENTS

We thank participants of the seminars moderated by A.S. Kingsepp, A.A. Rukhadze, and V.D. Shafranov for discussions and helpful remarks. This work was supported by Presidium of the RAS under the multidisciplinary research program "Thermophysics and Mechanics of Intensive Energy Impacts," by the Division of Power Engineering, Mechanical Engineering, and Control Process of the RAS under the program "Information and Computer Modeling of Basic Problems of Nonequilibrium Media," and by the Russian Foundation for Basic Research, project no. 03-07-90272v. Computations were performed on a cluster computer procured under a grant from Deutscher Akademischer Austauschdienst (German Academic Exchange Service). One of us (I.V.M.) gratefully acknowledges the support provided by the Dynasty Foundation affiliated with the International Center for Fundamental Physics in Moscow. We thank the referee for remarks that helped to improve presentation in this paper.

REFERENCES

1. B. N. Lomakin and V. E. Fortov, Zh. Éksp. Teor. Fiz. **63**, 92 (1972) [Sov. Phys. JETP **36**, 48 (1973)].
2. V. K. Gryaznov, M. V. Zhernokletov, V. N. Zubarev, *et al.*, Zh. Éksp. Teor. Fiz. **78**, 573 (1980) [Sov. Phys. JETP **51**, 288 (1980)].
3. V. B. Mintsev and Yu. B. Zaporoghets, Contrib. Plasma Phys. **29**, 493 (1989).
4. A. Ng, P. Celliers, G. Hu, and A. Forsman, Phys. Rev. E **52**, 4299 (1995).
5. V. K. Gryaznov, V. E. Fortov, M. V. Zhernokletov, *et al.*, Zh. Éksp. Teor. Fiz. **114**, 1242 (1998) [JETP **87**, 678 (1998)].
6. D. Riley, N. C. Woolsey, D. McSherry, *et al.*, Phys. Rev. Lett. **84**, 1704 (2000).
7. A. W. DeSilva and J. D. Katsouros, J. Phys. IV **10** (5), 209 (2000).
8. V. N. Korobenko, A. D. Rakhel, A. I. Savvatimskiy, and V. E. Fortov, Plasma Phys. Rep. **28**, 1008 (2002).
9. E. M. Lifshitz and L. P. Pitaevskiĭ, *Physical Kinetics*, 2nd ed. (Fizmatlit, Moscow, 2001; Pergamon, Oxford, 1981).
10. L. A. Artsimovich and R. Z. Sagdeev, *Plasma Physics for Physicists* (Atomizdat, Moscow, 1979) [in Russian].
11. A. A. Valuev, Teplofiz. Vys. Temp. **18**, 422 (1980).
12. Yu. K. Kurilenkov and A. A. Valuev, Beitr. Plasmaphys. **24**, 161 (1984).
13. M. A. Berkovsky and Yu. K. Kurilenkov, J. Phys. B **24**, 5043 (1991).
14. A. A. Valuev, A. S. Kaklyugin, and G. É. Norman, Zh. Éksp. Teor. Fiz. **114**, 880 (1998) [JETP **86**, 480 (1998)].
15. A. A. Valuev, Teplofiz. Vys. Temp. **15**, 1143 (1977).
16. J. P. Hansen and I. R. McDonald, Phys. Rev. A **23**, 2041 (1981).
17. A. A. Valuev, I. V. Morozov, and G. É. Norman, Dokl. Akad. Nauk **362**, 752 (1998) [Dokl. Phys. **43**, 608 (1998)].
18. A. Selchow, G. Röpke, A. Wierling, *et al.*, Phys. Rev. E **64**, 056410 (2001).
19. V. Golubnychiy, M. Bonitz, D. Kremp, and M. Schlanges, Phys. Rev. E **64**, 016409 (2001).
20. G. E. Norman and A. A. Valuev, in *Strongly Coupled Coulomb Systems*, Ed. by G. Kalman, M. Rommel, and K. Blagoev (Plenum, New York, 1998), p. 103.
21. A. S. Kaklyugin, G. E. Norman, and A. A. Valuev, J. Tech. Phys. **41**, 65 (2000).
22. A. F. Aleksandrov, L. S. Bogdankevich, and A. A. Rukhadze, *Oscillations and Waves in Plasma Media* (Mosk. Gos. Univ., Moscow, 1990) [in Russian].
23. D. N. Zubarev, V. G. Morozov, and G. Röpke, *Statistical Mechanics of Nonequilibrium Processes* (Akademie, Berlin, 1996; Fizmatlit, Moscow, 2002), Vol. 1.
24. A. S. Kingsep, *Introduction to Nonlinear Plasma Physics* (Mosk. Fiz.–Tekh. Inst., Moscow, 1996) [in Russian].
25. R. R. Ramazashvili, A. A. Rukhadze, and V. P. Silin, Zh. Éksp. Teor. Fiz. **43**, 1323 (1962) [Sov. Phys. JETP **16**, 939 (1963)].
26. L. M. Biberman and G. É. Norman, Teplofiz. Vys. Temp. **7**, 822 (1969).
27. G. É. Norman, Pis'ma Zh. Éksp. Teor. Fiz. **73**, 13 (2001) [JETP Lett. **73**, 10 (2001)].
28. C. Aman, J. B. C. Pettersson, L. Lindroth, and L. Holmlid, J. Mater. Res. **7**, 100 (1992).
29. L. Holmlid, J. Phys. B: At. Mol. Opt. Phys. **37**, 357 (2004).
30. T. C. Killian, M. J. Lim, S. Kulin, *et al.*, Phys. Rev. Lett. **86**, 3759 (2001).
31. F. J. Dyson, J. Math. Phys. **8**, 1538 (1967).
32. G. É. Norman and A. N. Starostin, Teplofiz. Vys. Temp. **6**, 410 (1968).
33. B. V. Zelener, G. E. Norman, and V. S. Filinov, *Perturbation Theory and Pseudopotential in Statistical Thermodynamics* (Nauka, Moscow, 1981) [in Russian].
34. A. V. Filinov, M. Bonitz, and W. Ebeling, J. Phys. A **36**, 5957 (2003).
35. C. Deutsch, Phys. Lett. A **60A**, 317 (1977).
36. C. Deutsch, M. M. Gombert, and H. Minoo, Phys. Lett. A **66A**, 381 (1978); Phys. Lett. A **72A**, 481 (1979).
37. I. V. Morozov and G. E. Norman, J. Phys. A **36**, 6005 (2003).
38. I. V. Morozov, G. E. Norman, and A. A. Valuev, Phys. Rev. E **63**, 36 405 (2001).
39. S. I. Andreev and N. F. Ivasenko, *Fundamentals of Calculation of Pulse Xenon Lamps* (Tomsk. Gos. Univ., Tomsk, 1982) [in Russian].
40. H. Reinholz, I. Morozov, G. Röpke, and Th. Millat, Phys. Rev. E **69**, 066412 (2004).
41. D. Bohm, *General Theory of Collective Coordinates* (London, 1959; Mir, Moscow, 1964).
42. G. E. Norman and A. A. Valuev, Plasma Phys. **21**, 531 (1979).
43. A. A. Valuev and Yu. K. Kurilenkov, Teplofiz. Vys. Temp. **21**, 591 (1983).
44. Yu. K. Kurilenkov and A. A. Valuev, Beitr. Plasmaphys. **24**, 529 (1984).
45. H. Reinholz, R. Redmer, G. Röpke, and A. Wierling, Phys. Rev. E **62**, 5648 (2000).
46. A. Esser, R. Redmer, and G. Röpke, Contrib. Plasma Phys. **43**, 33 (2003).
47. I. Ya. Dikhter and V. A. Zeĭgarnik, Dokl. Akad. Nauk SSSR **227**, 656 (1976).
48. E. M. Apfel'baum and M. F. Ivanov, Fiz. Plazmy **27**, 79 (2001) [Plasma Phys. Rep. **27**, 76 (2001)].
49. M. A. Berkovsky, D. Djordjevic, Yu. K. Kurilenkov, and H. M. Milchberg, J. Phys. B **24**, 5043 (1991).
50. M. Schlanges, Th. Bornath, D. Kremp, and P. Hilse, Contrib. Plasma Phys. **43**, 360 (2003).
51. I. V. Morozov, G. E. Norman, A. A. Valuev, and I. A. Valuev, J. Phys. A **36**, 8723 (2003).
52. G. É. Norman and V. V. Stegaĭlov, Zh. Éksp. Teor. Fiz. **119**, 1011 (2001) [JETP **92**, 879 (2001)].

Translated by A. Betev

Journal of Experimental and Theoretical Physics, Vol. 100, No. 2, 2005, pp. 385–397.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 127, No. 2, 2005, pp. 431–444.

STATISTICAL, NONLINEAR, AND SOFT MATTER PHYSICS

The Percolation Transition in Filling a Nanoporous Body by a Nonwetting Liquid

V. D. Borman[a], A. A. Belogorlov[a], A. M. Grekhov[a], G. V. Lisichkin[b], V. N. Tronin[a], and V. I. Troyan[a]

[a]*Moscow State Institute of Engineering Physics (Technical University), Kashirskoe sh. 31, Moscow, 115409 Russia*
[b]*Moscow State University, Vorob'evy gory, Moscow, 119992 Russia*
e-mail: a_belogorlov@mail333.com
Received May 21, 2004

Abstract—The paper presents the results of an experimental study of the percolation transition in filling by nonwetting liquids of nanoporous bodies of various natures with different specific surface areas and mean pore and granule sizes. The liquid that we used was an aqueous solution of ethylene glycol. The hysteresis and nonoutflow phenomena observed in this transition at various (known) surface energies of liquids were studied by varying the concentration of ethylene glycol. This helped us explain the mechanism of the percolation transition in filling nanoporous bodies with nonwetting liquids. It was shown that, to quantitatively describe the observed dependences in terms of percolation theory taking into account energy barriers to filling, we must use a nonscaling distribution function of clusters of accessible and filled pores that admits the formation of pore clusters of arbitrary dimensions. *© 2005 Pleiades Publishing, Inc.*

1. INTRODUCTION

In spite of a long history of studies of phenomena related to filling porous bodies with nonwetting liquids, many questions still remain unanswered. These include capillary phenomena in superthin channels, the dynamics of filling of disordered porous solids, and the problem of unwettability (wettability) of porous bodies [1, 2]. The corresponding phenomena are related to transport of nonwetting liquids in porous solids with nanoscale pores, which is currently of special interest both for basic science and in view of numerous nano and membrane technology applications [1–5]. One of the promising directions of basic research in this field is the development of devices for the absorption and accumulation of mechanical energy [6, 7]. Such devices can be created by exploiting the phenomenon of filling nanoporous solids with nonwetting liquids.

The common approach to describing the filling of porous solids by nonwetting liquids is based on percolation theory [8, 9]. The filling process is then described as a hydrodynamic process that occurs at the Laplace pressure p_L. If a porous body has a pore-size distribution, percolation through an infinite cluster of filled pores takes place at the pressure at which the number of filled pores reaches the percolation threshold. Hysteresis and nonoutflow phenomena are observed in filling nanoporous solids with nonwetting liquids [10, 11]. These phenomena are related either to the special structure of pores [12–14] or to changes in the wetting angle during inflow–outflow [10, 11, 15]. For instance, it is assumed that nanoporous solids contain blind pores that a liquid cannot escape from when excess pressure decreases to zero. The hysteresis phenomenon is related to the existence of narrow-necked ("bottle-shaped") pores [9]. An alternative explanation of the hysteresis and nonoutflow phenomena is based on the introduction of a distribution of clusters in percolation theory [16] and an energy barrier to filling nanoporous solids with nonwetting liquids [17]. The surface energy of the formation of liquid nanoparticles is then comparable to their volume energy. Near the critical filling pressure ($p_c \sim p_L$), the energy of fluctuation formation of nanoparticles of a liquid changes from several electron volts to an energy comparable to the temperature [17], which results in the fluctuation formation of pore clusters filled with a liquid at $p \sim p_c$. The formation of similar clusters in the vicinity of the percolation transition reduces to the problem considered in [16]. In the theory suggested in [17], both hysteresis and nonoutflow phenomena are described taking into account the energy barrier to filling without additional assumptions concerning the special geometry of pores and wetting angle changes. The volume of the liquid that remains in a porous body should then strongly depend on the surface energy of the nonwetting liquid.

In this work, we experimentally studied the filling of nanoporous solids of various natures with a nonwetting liquid. The solids had different specific surface areas and mean pore and granule sizes. The liquid was an aqueous solution of ethylene glycol, for which the concentration dependence of surface energy is well known [18]. We varied the concentration of ethylene glycol to study the hysteresis and nonoutflow phenomena at various liquid surface energies σ. Performing

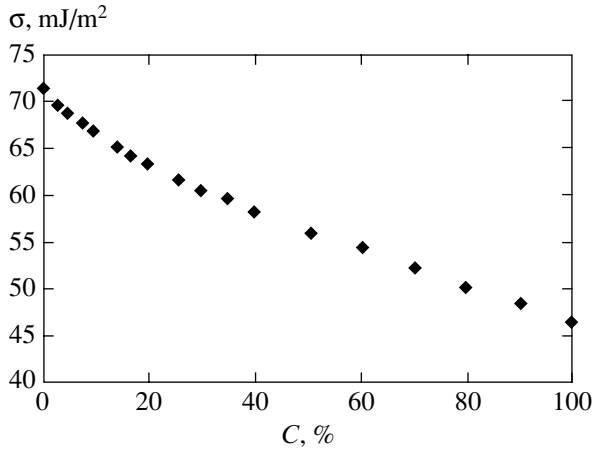

Fig. 1. Surface energy of water–ethylene glycol solutions as a function of the concentration of ethylene glycol in wt %.

such experiments makes it possible in principle to answer the question of the mechanism of the percolation transition in filling nanoporous solids by nonwetting liquids.

A comparison of our experimental dependences with theoretical calculations, however, showed that the theory suggested in [17] did not describe the totality of the experimental data. We found this to be caused by the use of a scaling distribution function for filled pore clusters in [17]. Such a function only takes into account the formation of fractal clusters close to the percolation threshold of dimensions not larger than the correlation length [8]. At the same time, an important role in inflow–outflow processes can be played by clusters whose size is comparable to the characteristic size of porous bodies (granule size). Such clusters are formed as a result of the interaction (coalescence) of small clusters. Their formation can broaden the percolation transition region.

In the second section, we describe the procedure for measurements and experimental results. It was found that the inflow and outflow pressures dropped as σ decreased for all porous media that we used. The fraction of the liquid that remained in porous solids simultaneously increased even to complete nonoutflow. We show in the third section that the observed dependences can be described quantitatively in terms of percolation theory taking into account the energy barrier only if we use a nonscaling distribution function for clusters [16] of accessible and filled pores. Such a function should admit the formation of pore clusters of arbitrary sizes.

2. PROCEDURE AND MEASUREMENT RESULTS

The pores of hydrophobic nanoporous solids were filled with a liquid as follows: A porous solid of a fixed mass (m = 1–4 g) was placed into a hermetic high-pressure chamber of volume 65 cm^3. The remaining free volume of the chamber was filled with the liquid. A rod of length 17 cm and cross section area 0.8 cm^2 was inserted into the chamber through a gasket. The assembled chamber was mounted on a test bench that allowed the rod inside the chamber to move under the action of an applied force, which was measured by a strain gauge with a working measurement range from 1 to 2000 kgf. Changes in the volume of the liquid–porous solid system were recorded using a displacement pickup by measuring the length of the rod introduced into the chamber. The error of pressure and volume change data did not exceed 10%. The rate of pressure changes was no more than 1 atm/s. The estimates given in Section 3 show that the process can be considered quasi-static under these conditions. The experimental dependences could therefore be analyzed without introducing corrections for liquid viscosity. Signals from the displacement pickup and strain gauge were fed into an analog-to-digital converter, stored in a computer, and processed to construct the pressure dependence of volume changes for the porous solid–liquid system under study.

Measurements were performed for four hydrophobic porous solids. These were silasorb S18, silasorb S8, polysorb-1, and libersorb 2U-8. The first two solids were modified sorbents with a SiO_2 framework, their specific surface areas of pores were 300 m^2/g, the mean pore radius was about 5 nm, and the size of granules was about 7.5–10 μm [19]. Porous polymeric sorbent polysorb-1 had specific surface area of about 220 m^2/g, mean pore radius of about 6.5 nm, and granule size of about 250–500 μm [20]. Libersorb 2U-8 was KSK-G silica gel with a SiO_2 framework (specific surface area about 400 m^2/g, mean pore radius about 4 nm, and granule size 10–63 μm). The surface of KSK-G was subjected to hydrophobization by chemical modification with octyldimethylchlorosilane and additional silanization with trimethylchlorosilane [21]. Aqueous solutions of ethylene glycol were used as nonwetting liquids. The surface energy of water and ethylene glycol at 25°C was 72 [22] and 49 mJ/m^2 [23], respectively. According to [18], the surface energy of water–ethylene glycol solutions decreases as the concentration of ethylene glycol C grows (see Fig. 1). The concentration of ethylene glycol was varied from 0 to 60% in our experiments. The compressibility of the chamber and liquid was determined in additional experiments in the absence of porous solids.

Typical plots describing a decrease in volume ($-\delta V$) as a function of pressure p excess over atmospheric pressure are shown in Fig. 2 for the libersorb 2U-8–water system. These dependences were obtained by subtracting the compressibility of the chamber and liquid from the results of measurements. The linear segment (curve I) from $p_0 = 0$ to $p_1 = 200$ atm (point *1*) corresponds to volume changes related to the elastic strain

of the unfilled porous solid. The data on the compressibility of the unfilled porous solid determined as

$$\chi = V_s^{-1} V_1 / p_1,$$

where V_s is the volume of the porous solid sample, are listed in the table. At pressures $p > p_1$ (starting with point *1*), we observe a substantial decrease in the volume of the system caused by pore filling as the pressure increases. Note that the change in the system volume on account of elastic porous body compression as pressure increases from 200 to 400 atm (points *1–3)* does not exceed 3%. The compressibility of the system monotonically increases over the region $p_1 - p_2 \approx 220$ atm (points *1*, *2*) and reaches a maximum at point *2*. According to [17], the susceptibility (compressibility) of the system

$$\lambda = \frac{d(\delta V(p))}{dp}$$

takes on the largest value at this point (see Fig. 3) because of the filling of the porous body. Pressures p_2 and susceptibilities λ for the other systems studied in this work are listed in the table.

The system volume increases and the liquid flows out of the porous solid pores as pressure decreases (curve II). We found that the major part of the liquid (>70%) flows out at pressure $p < 10$ atm. When pressure is again increased (curve III, points *0'–3*), the system volume begins to decrease at point *0'* (Fig. 2). The *0–0'* segment corresponds to the volume $V_0 = \delta V_{0'} - \delta V_0$ of the liquid that remains in pores when pressure excess is zero. The $\delta V(p)$ dependence for the second liquid

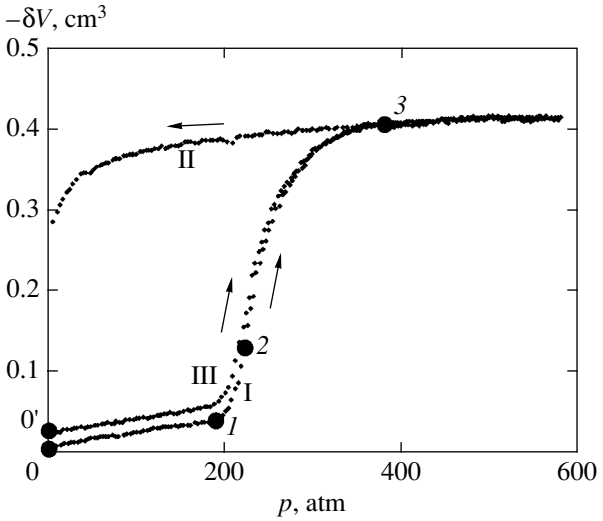

Fig. 2. Pressure dependence of volume changes for the libersorb 2U-8–water system (curve I, pressure rise; curve II, pressure drop; and curve III, repeated pressure rise).

Table

Ethylene glycol concentration, %	σ, mJ/m²	λ, 10^{-3} cm³/atm	p_2, atm	ψ	δσ, mJ/m²	σ/δσ	R_0, nm
Libersorb 2U-8							
$\bar{R}$ = 3.6 nm, δR = 0.4 nm, V_{por} = 0.38 ± 0.03 cm³/g, φ = 0.33, χ = (0.8 ± 0.1) × 10^{-3} atm^{-1}							
0	72.0	6.4	220	0.05	23.0	3.1	3.2
6	68.5	6.1	200	0.24	21.0	3.3	3.3
11	66.3	6.3	190	0.66	19.0	3.5	3.5
16	64.4	6.1	175	0.76	17.0	3.8	3.8
60	55.6	5.5	125	1.00	12.5	4.5	4.1
Silasorb S8							
$\bar{R}$ = 5.0 nm, δR = 0.5 nm, V_{por} = 0.40 ± 0.04 cm³/g, φ = 0.45, χ = (1.2 ± 0.2) × 10^{-3} atm^{-1}							
0	72.0	7.7	135	1.00	11.0	6.6	7.6
60	55.6	10.8	90		6.0	9.3	8.8
Silasorb S18							
$\bar{R}$ = 5.0 nm, δR = 0.5 nm, V_{por} = 0.48 ± 0.05 cm³/g, φ = 0.45, χ = (1.2 ± 0.2) × 10^{-3} atm^{-1}							
0	72.0	17.0	100	1.00	9.0	83.0	7.1
60	55.6	17.4	55		2.5	22.2	11.4
Polysorb-1							
$\bar{R}$ = 6.5 nm, δR = 0.9 nm, V_{por} = 0.95 ± 0.07 cm³/g, φ = 0.55, χ > 10 × 10^{-3} atm^{-1}							
0	72.0	52.8	30	1.00	4.0	18.0	8.9
60	55.6	120.0	15		1.0	55.6	11.0

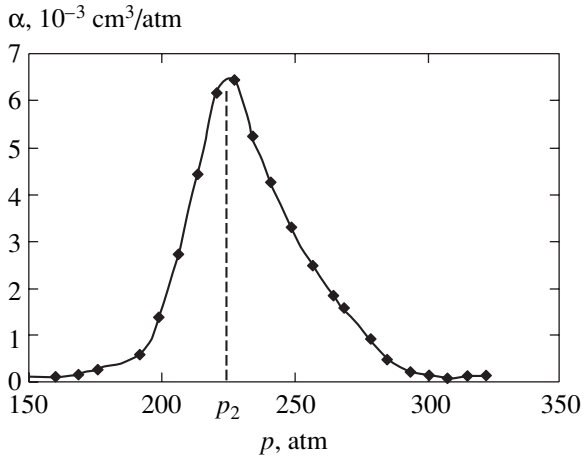

Fig. 3. Susceptibility of the libersorb 2U-8–water system as a function of pressure.

outflow repeats dependence II obtained in the first cycle. After the second inflow–outflow cycle, the volume of the liquid that remains in the porous solid does not change.

The experimental $\delta V(p)$ dependences were used to determine the specific volume of pores

$$V_{por} = \delta V_3 - \delta V_1$$

and the relative volume of the liquid remaining in the porous body,

$$\psi = V_0/V_{por},$$

after the first inflow–outflow cycle.

To visually represent the results, the experimental relative volume $\delta V/V_{por}$ values at various pressures are

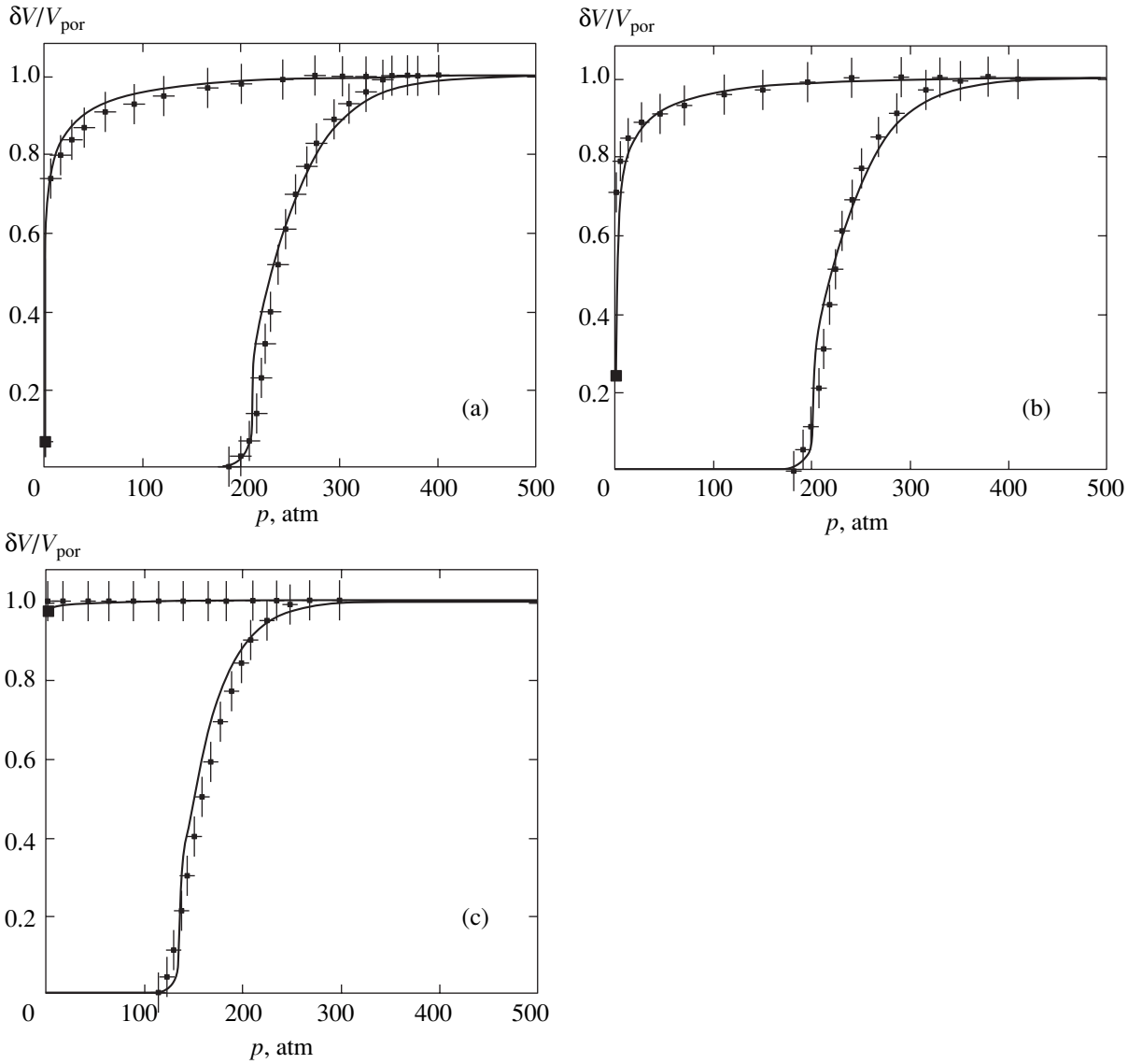

Fig. 4. Pressure dependences of changes in the relative liquid volume in a porous body for the systems (a) libersorb 2U-8–water, (b) libersorb 2U-8–aqueous ethylene glycol (6%), and (c) libersorb 2U-8–aqueous ethylene glycol (60%) [solid circles correspond to experimental data and solid lines to calculations by (6), (7), (14), and (16)].

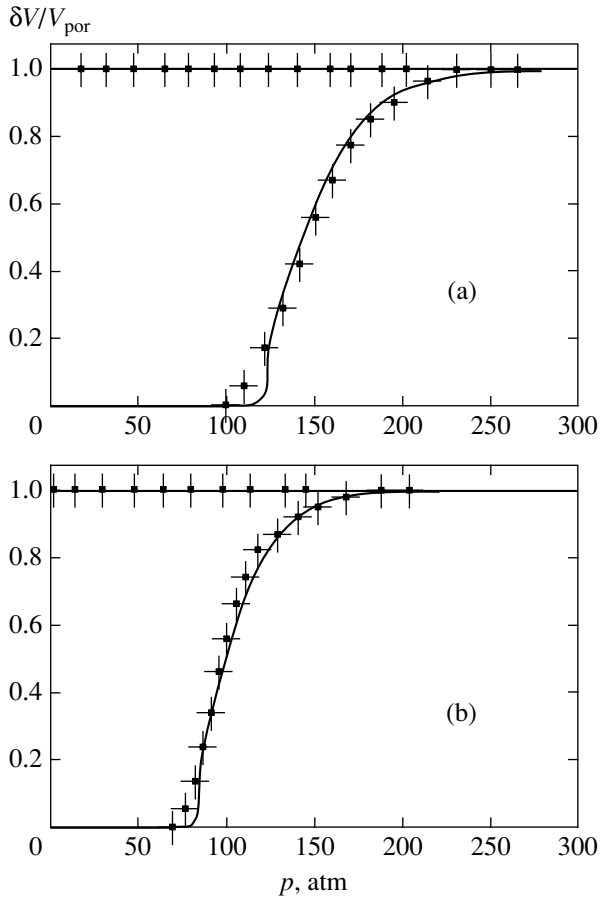

Fig. 5. Pressure dependences of changes in the relative liquid volume in a porous body for the systems (a) silasorb S8–water and (b) silasorb S8–aqueous ethylene glycol (60%) [solid circles correspond to experimental data and solid lines to calculations by (6), (7), (14), and (16)].

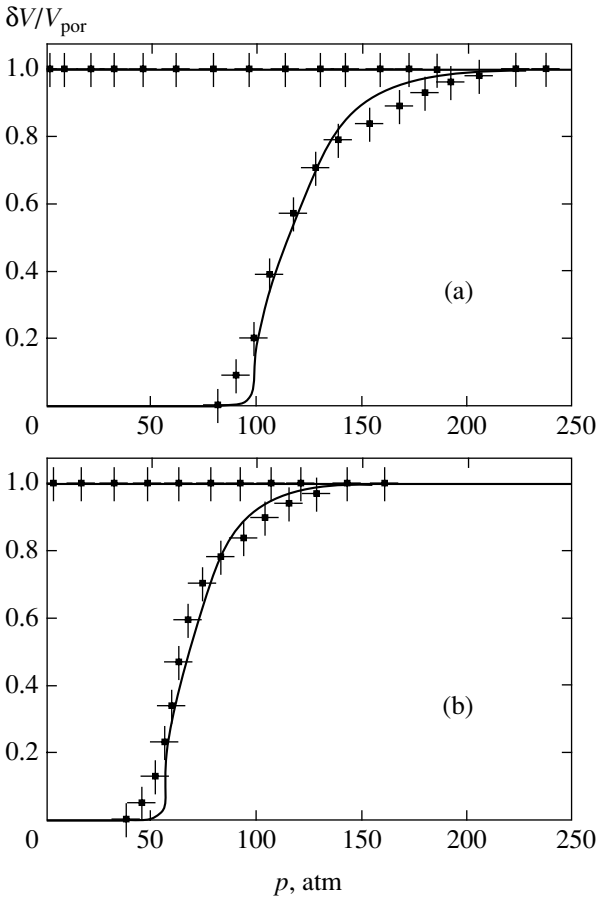

Fig. 6. Pressure dependences of changes in the relative liquid volume in a porous body for the systems (a) silasorb S18–water and (b) silasorb S18–aqueous ethylene glycol (60%) [solid circles correspond to experimental data and solid lines to calculations by (6), (7), (14), and (16)].

shown in Fig. 4 (solid circles) for the system libersorb 2U-8–aqueous ethylene glycol. Compared with Fig. 2, the dependences were corrected for porous body deformation. Hysteresis and incomplete liquid outflow were observed at concentrations of $C = 0$ and 11% (Figs. 4a, 4b). The relative volume of the liquid that remained in pores increased from $\psi = 0.05 \pm 0.01$ to 0.66 ± 0.06 as the concentration increased to 11% (see table). At a $C =$ 60% ethylene glycol concentration, no liquid flowed out of the pores ($\psi \approx 1.0$). The pressure of pore filling then decreased from $p_2 = 220$ atm for pure water to $p_2 =$ 125 atm at $C = 60\%$ (see table).

Figure 4 shows that an increase in the concentration of ethylene glycol from 0 to 60% shifts the inflow and outflow curves toward lower pressures and causes an increase in the volume of the liquid that remains in the pores after decreasing the pressure to atmospheric, even to complete nonoutflow of a 60% solution of ethylene glycol (see table). Note that a 25% relative change in the surface energy of the solution decreases the filling pressure p_2 by more than 50%. This is evidence that the surface energy dependence of the filling pressure is not described by the Laplace equation ($p_L \approx \sigma/R$, where R is the radius of the pores).

The pressure dependences of the specific volume $\delta V/V_{por}$ for the silasorb S8–water, silasorb S8–aqueous ethylene glycol (60%) and silasorb S18–water, silasorb S18–aqueous ethylene glycol (60%) systems are shown in Figs. 5 and 6. In these systems, pore filling was observed at $p_2 = 135$ and 90 atm, respectively (see table). These pressures are different although the mean radii of pores are equal in these porous media and the surface energies of the liquids are also equal. The figures show that, as with libersorb 2U-8, a decrease in the surface energy of the liquid decreases the pressure p_2 of porous body filling by the liquid (see table). Complete nonoutflow of the liquid from the porous bodies ($\psi = 1$) was also observed for these systems.

The pressure dependences of $\delta V/V_{por}$ for the polysorb-1–aqueous ethylene glycol (0 and 60%) systems are shown in Fig. 7. For these systems also, we observe a decrease in the pressure of pore filling from $p_2 = 30$ atm (at $C = 0$) to $p_2 = 15$ atm (at $C = 60\%$) and

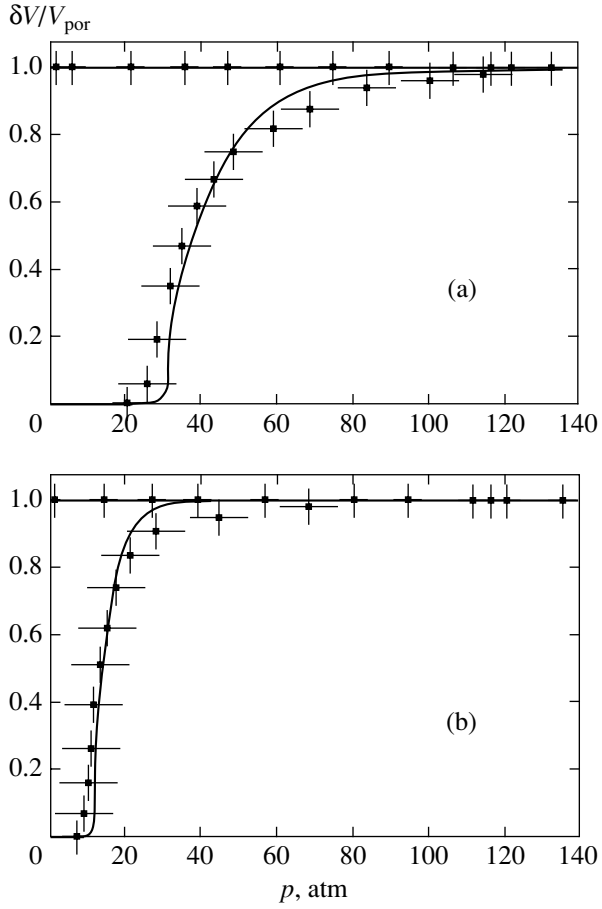

Fig. 7. Pressure dependences of changes in the relative liquid volume in a porous body for the systems (a) polysorb-1–water, (b) polysorb-1–aqueous ethylene glycol (6%), and (c) polysorb-1–aqueous ethylene glycol (60%) [solid circles correspond to experimental data and solid lines to calculations by (6), (7), (14), and (16)].

complete liquid nonoutflow at both ethylene glycol concentrations.

3. RESULTS AND DISCUSSION

Our experiments show that the hysteresis and nonoutflow phenomena are observed in the aqueous ethylene glycol–hydrophobic nanoporous medium systems studied in this work. The assumption of a special structure of pores, when large-sized pores are surrounded by small-sized ones (bottle-shaped pores) leads us to conclude that a decrease in pressure after complete filling should cause the liquid first to flow out of small-sized pores under Laplace pressure action irrespective of its surface energy. The liquid then remains in large-sized pores. However according to the results obtained in this work, the volume of the liquid that remains in pores depends on σ. For this reason, the nonoutflow phenomenon cannot be explained by the special structure of the pores of the porous media that we studied. Our experiments also show that the pressure of the second filling, when part of the pores are filled in the initial state, coincides to within errors with the pressure of the first filling of the initially empty porous body. This is evidence that the liquid does not flow out of small-radius pores; otherwise, the pressure of the second filling would be higher than that of the first filling.

The experimental data on nonwetting liquid flow in–flow out of the nanoporous bodies and the hysteresis and nonwetting liquid nonoutflow phenomena will be described following [17] in terms of percolation theory taking into account energy barriers to nonwetting liquid inflow into and outflow from the pores.

Let us consider a porous solid immersed into a nonwetting liquid. The applied pressure is p, and this pressure does work when the porous solid is filled. Let $\delta A(p, R)$ be the work done to perform the fluctuation filling of one pore of radius R in the porous body. As the pore can be either filled (probability $w \approx 1$, $\delta A(p, R) < 0$) or empty ($w = 0$, $\delta A(p, R) > 0$), the normalized probability can be written as

$$w_i(p, R) = \left[1 + \exp\left(\frac{\delta A(p, R)}{T}\right)\right]^{-1}. \quad (1)$$

We ignore the elastic interaction of the pores with each other. The work $\delta A(p, R)$ done to fill the pore in a porous body then consists of the work of liquid expansion into the volume v of the empty pore $-pv$, the energy per unit area of the liquid–porous body interface $\delta\sigma(1-\eta)$,

$$\delta\sigma = \sigma^{sl} - \sigma^{sg},$$

where σ^{sl} is the surface energy of the solid–liquid interface and σ^{sg} is the surface energy of the solid–gas interface, and the energy per unit area $\sigma\eta$ of menisci of the liquid in the throats of the filled pore (η is the ratio between the area of the menisci and the area of the surface of the pore). For simplicity, let the pore be spherical and have radius R. The work is then written as

$$\delta A = A_i(p, R)\frac{4}{3}\pi R^3,$$
$$A_i(p, R) = -p + \frac{3\delta\sigma}{R}\left[1 + \eta\left(\frac{\sigma}{\delta\sigma} - 1\right)\right]. \quad (2)$$

It follows from (2) that the $A_i(p, R)$ value changes sign at the pressure

$$p_i = \frac{3\delta\sigma}{R}\left[1 + \eta\left(\frac{\sigma}{\delta\sigma} - 1\right)\right]. \quad (3)$$

For two communicating pores with different radii, the ratio between the meniscus surface area and the sur-

face area of the pore is larger for the pore with a smaller diameter. The η value can therefore be written as

$$\eta = \varphi^{2/3}\left(\frac{R_{\min}}{R}\right)^2. \quad (4)$$

Here, φ is the porosity of the porous body equal to the volume ratio between the pores and the body and $R_{\min}$ is the smallest observed radius of the pores. At pressure $p = 0.9p_i$, $\sigma = 0.1$ J/m^2, $\delta\sigma \approx 0.1$ J/m^2, $\eta = 0.5$, and $\bar{R} = 5$ nm, we have $\delta A_i \approx 1$ eV. Therefore, according to (1)–(4),

$$\frac{\delta A_i}{T} \gg 1$$

at $T = 300$ K, and the probability changes from zero to one at pressures in a small neighborhood of p_i.

For a nonwetting liquid, $\delta\sigma > 0$. It therefore follows from (3) and (4) that $\delta A_i > 0$ at low pressures $p < p_i$. According to (1), the probability of filling pores of radius R is then small at $\delta A_i \gg T$ ($w_i \ll 1$). Conversely, $\delta A_i < 0$ at $p > p_i$, and, at $|\delta A_i| \gg T$, the probability w_i approaches one. At a slow (quasi-static) increase in pressure, porous body filling begins when the relative volume ϑ of the pores accessible to filling becomes close to the percolation threshold ϑ_c. For a nonwetting liquid, the fraction of the accessible pores is determined by the external pressure, from condition (3) of the development of filling fluctuations. It follows from (1) and (3) that increasing pressure increases the number of pores accessible to filling in porous media containing differently-sized pores. If the pore-size distribution function $f(R)$ is known for a porous body, the relative volume of all accessible pores can be written in the form

$$\vartheta_{\text{in}}(p) = \int_0^\infty \frac{4}{3}\pi R^3 f(R) w_i(p, R)\, dR, \qquad \int_0^\infty f(R)\, dR = 1. \quad (5)$$

According to percolation theory, percolation in a disordered three-dimensional medium occurs at $\vartheta = \vartheta_c = 0.16$ [8]. This means that an infinite fractal cluster of accessible pores is formed in a porous medium at the pressure that satisfies condition (3), as the $\vartheta_{\text{in}}(p_{\text{in}}) = \vartheta_c = 0.16$ value is attained with increasing the pressure, and porous body filling should occur close to the percolation threshold.

It follows from (1) that, during porous body filling, the pores in the system under consideration can at each time instant be divided into three types: inaccessible [for which $\delta A(p, R) > 0$] and therefore empty, accessible [$\delta A(p, R) < 0$] but empty (accessible pores in what follows), and filled. Porous body filling under slow pressure variations involves the formation of clusters of accessible and filled pores followed by filling clusters of accessible pores by a liquid as a result of its flow from filled to accessible pores and growth of filled pore clusters.

The transformation of accessible into filled pore clusters can be treated as their interaction. The problem of describing inflow–outflow processes then reduces to the problem of calculating the distribution of clusters $F(n)$ according to the number of pores in them n; this problem was considered in [16]. According to [16], under the conditions of slow (quasi-stationary) porous body filling, this function, which takes into account the interaction of accessible and filled pore clusters, has the form

$$F(n, p) = cn^{-\tau}\exp\left(\frac{-|\vartheta_{\text{in}}(p) - \vartheta_c| n^\sigma}{\vartheta_c}\right),$$

$$c = \left(\int_1^\infty F(n)\, dn\right)^{-1}, \quad (6)$$

$$\frac{|\vartheta_{\text{in}}(p) - \vartheta_c|}{\vartheta_c} < 1, \quad n \gg 1$$

for three-dimensional systems. Here, $\vartheta_{\text{in}}(p)$ is the fraction of pore volume filled at the given pressure [Eq. (5)]. The exponent τ determines the dependence of the number of clusters on the number of pores in them close to the percolation threshold at

$$\frac{|\vartheta - \vartheta_c|}{\vartheta_c} \ll 1$$

(see [8]), and the index σ characterizes the number of pores in a cluster whose size equals the correlation length $\xi \sim |\vartheta - \vartheta_c|^{-\nu}$ close to the percolation threshold [8]. For three-dimensional systems, we have $\tau = 2.2$, $\sigma = 0.44$, and $\nu = 0.89$ [8]. Note that the inflow–outflow percolation transition was described in [17] using the scaling distribution function

$$F = \begin{cases} cn^{-\tau}, & 1 < n < n(\xi) \\ 0, & n \geq n(\xi), \end{cases} \qquad c = \left(\int_1^{n(\xi)} n^{-\tau}\, dn\right)^{-1}. \quad (7)$$

Here, $n(\xi) = |\vartheta - \vartheta_c|^{-1/\sigma}$ is the number of pores in a cluster whose size equals the correlation length ξ [8, 16]. It follows from (6) and (7) that the use of a scaling distribution function in form (7) exaggerates the number of

clusters containing n pores at $1 < n < n(\xi)$ and ignores the formation of clusters whose size exceeds the correlation length. In contrast, nonscaling function (6) takes into account the formation of fractal clusters of larger and smaller dimensions compared with the correlation length. We show below that this difference is essential to the quantitative description of liquid outflow from a porous body.

Filled pores are only formed from accessible pores. According to (5) and (6), the equation describing the filled volume of the pore space can therefore be written in the form

$$V_{in}(p) = \int_1^\infty nF(n, p)\vartheta_{in}(p)dn$$
$$= \vartheta_{in}(p)\varepsilon^{-1/\sigma}\frac{\gamma\left(-\frac{\tau-2}{\sigma}, \varepsilon\right)}{\left(-\frac{\tau-1}{\sigma}, \varepsilon\right)}, \quad \varepsilon = \frac{|\vartheta_{in}(p) - \vartheta_c|}{\vartheta_c}. \tag{8}$$

Here, $\gamma(x, y)$ is the incomplete gamma-function. The filled volume value [Eq. (8)] is determined by the probability of formation of pores accessible at the given pressure and depends on $|\vartheta_{in}(p) - \vartheta_c|$,

$$V_{in}(p) = V_{in}(p, |\vartheta_{in}(p) - \vartheta_c|).$$

It follows from (5) that the ε parameter decreases as the pressure increases. Close to the percolation threshold, we have

$$\varepsilon = \frac{|\vartheta_{in}(p) - \vartheta_c|}{\vartheta_c} \longrightarrow 0$$

and the filled volume value remains finite at $\tau > 2$,

$$V_{in} \sim A(\tau, \sigma, \vartheta_c) - B(\tau, \sigma, \vartheta_c)\varepsilon^{(\tau-2)/\sigma}$$

[$A(\tau, \sigma, \vartheta_c)$ and $B(\tau, \sigma, \vartheta_c)$ are positive values depending on the indices τ and σ and percolation threshold ϑ_c].

It follows from (1) and (5) that the relative volume of the accessible pores can be written as

$$\vartheta_{in}(p) \sim \int_{R(p)}^\infty \frac{4}{3}\pi R^3 f(R)dR. \tag{9}$$

Here, $R(p)$ is determined from the condition of the vanishing of the work necessary for the formation of a pore accessible at the given pressure [Eq. (2)]. It follows from (9) that, in conformity with (2), $R(p)$ decreases and the relative volume of the accessible pores $\vartheta_{in}(p)$ increases as the pressure grows. Simultaneously, the size of the cluster of accessible pores monotonically increases because the relative volume of the accessible pores [Eq. (5)] approaches the percolation threshold ϑ_c. As a porous body can only be filled by sequentially filling communicating pores, the filling of the whole porous body volume is a result of the formation of an infinite fractal cluster of filled pores. Such a cluster is formed when the relative volume of the accessible pores [Eq. (5)] is close to the percolation threshold ϑ_c. Calculations show that, although the filling process is controlled by the number-of-pores distribution of clusters $F(n, p)$ given by (6), the difference between function (6) and scaling function (7) is then inessential to filled volume calculations. The differentiation of (8) allows the behavior of the susceptibility of the liquid–porous body system to be calculated close to the filling percolation transition,

$$\lambda \approx \frac{\partial V_{in}}{\partial p} \approx |\vartheta_{in}(p) - \vartheta_c|^{(\tau-2-2\sigma)/\sigma}.$$

The well-known dependences of the τ and σ indices on the basic critical indices ν and β,

$$\tau = \frac{2\nu d - \beta}{\nu d - \beta}, \quad \sigma = \frac{1}{\nu d - \beta}$$

(d is the space dimensionality), which determine the dependence of the correlation length and the probability of infinite cluster formation close to the percolation threshold, can be used to obtain

$$\lambda \approx |\vartheta_{in}(p) - \vartheta_c|^{\beta-2}.$$

For three-dimensional systems, the critical infinite cluster index is $\beta = 0.42$ [8]. It follows that the critical susceptibility index is determined only by the critical index of the probability of infinite cluster formation.

The pressure p_i in (3) depends on the surface energy of the liquid σ and the energy $\delta\sigma$, and the pore-size distribution $f(R)$ can be characterized by the mean pore radius $\bar{R}$ and the distribution half-width δR. The filled volume $V_{in}(p)$ at the given pressure in (7) is therefore also a function of these parameters.

Equation (8) can only be used to describe filling when the pressure increases slowly and viscous energy dissipation during liquid movement in the space of pores can be ignored. The liquid moves under the action of the pressure $\delta p = p - p_i$, that is, when pressure p rises above the p_i value. The characteristic time τ of filling N pores with a liquid with the viscosity coefficient μ can be estimated as

$$\tau \approx \mu\frac{dN}{dp}.$$

For a porous body with the granule size $\bar{R}_g$, pore size $\bar{R}$, porosity φ, and sample compressibility $\lambda = dV/dp$,

the number of pores δN filled when the pressure increases by δp is given by the equation

$$\delta N = \varphi\lambda\left(\frac{\bar{R}_g}{\bar{R}}\right)^3 \frac{\delta p}{V_{\text{obr}}} \quad (10)$$

(V_{obr} is the volume of the pores in the sample of mass m).

For the porous bodies under investigation, for instance, for libersorb 2U-8, we have $\bar{R}_g \approx 20$ μm, $\bar{R} \approx 3.6$ nm, $\lambda \approx 5 \times 10^{-3}$ cm^3/atm, $\varphi \approx 0.5$, $\mu = 0.8 \times 10^{-3}$ Pa s, and $V_{\text{obr}} \approx 1.6$ cm^3 ($m = 4$ g). The number of pores filled when the pressure increases by $\delta p = 1$ atm is therefore $\delta N/\delta p \approx 10^6$ 1/atm, and the characteristic time of filling is $\tau \approx 10^{-2}$ s. The time of increasing the pressure by 1.0 atm was approximately 1 s in our experiments. It follows that the condition of quasi-static filling was fulfilled, and Eqs. (1)–(8) could be used to describe the experimental data on filling.

Next, let us consider the stability of the state of the system "a nanoporous body immersed into a nonwetting liquid with initially filled pores at $p > p_{\text{in}}$." The nonwetting liquid can flow out of the porous body as the pressure decreases. The normalized probability of empty pore formation is

$$w_0(p, R) = \left[1 + \exp\left(\frac{\delta A_0}{T}\right)\right]^{-1}. \quad (11)$$

Here, δA_0 is the work done to form an empty pore. It consists of the work done by the liquid in outflow from the pore (pV), the energy of formation of liquid menisci in the pores connected with the emptied pore ($\sigma\eta$), and the energy of formation of the liquid–porous body interface [$-\delta\sigma(1 - \eta)$]. For a spherical pore of radius R, we have

$$\begin{aligned} \delta A_0 &= A_0(p, R)\frac{4}{3}\pi R^3, \\ A_0(p, R) &= p - \frac{3\delta\sigma}{R}\left[1 - \eta\left(\frac{\sigma}{\delta\sigma} + 1\right)\right]. \end{aligned} \quad (12)$$

According to (12), δA_0 changes sign at the pressure

$$p_0 = \frac{3\delta\sigma}{R}\left[1 - \eta\left(\frac{\sigma}{\delta\sigma} + 1\right)\right]. \quad (13)$$

For a nonwetting liquid, $\delta\sigma > 0$, and it follows from (13) that $\delta A_0 > 0$ for pores of radius R at a high pressure $p > p_0$. At $\delta A_0 \gg T$, the probability of empty pore formation is close to zero ($w_0 \approx 0$). At pressure $p < p_0$, work $\delta A_0 < 0$, and, at $|\delta A_0| \gg T$, the nonwetting liquid can flow out of the pores ($w_0 \approx 1$). It follows from (4) and (13) that the pressure p_0 is maximum and equals

$$p_{0\,\max} = 2\frac{\delta\sigma}{R_n}$$

for pores of radius

$$R_n = R_{\min}(3\varphi^{2/3})^{1/2}\left(\frac{\sigma}{\delta\sigma} + 1\right)^{1/2}. \quad (14)$$

This means that, after complete filling followed by a decrease in the pressure, the nonwetting liquid flows out first at $p_{0\,\max}$ from pores of radius R_n. The liquid, however, begins to flow out of the porous body at a pressure lower than $p_{0\,\max}$, when the relative volume of the pores capable of releasing the liquid is close to the percolation threshold ($\vartheta_c = 0.16$),

$$\vartheta_{\text{out}}(p_{\text{out}}) = \int_0^\infty \frac{4}{3}\pi R^3 f(R) w_0(p_{\text{out}}, R) dR = \vartheta_c. \quad (15)$$

Equation (15) and the distribution function of clusters according to the number of pores in them [Eq. (6), we must substitute $\vartheta_{\text{out}}(p)$ into it for $\vartheta_{\text{in}}(p)$] can be used to calculate the volume of the liquid–porous body system with the initially completely filled pore volume when the pressure decreases to a p value close to p_{out},

$$\begin{aligned} V_{\text{out}}(p) &= \int_1^\infty nF(n, p)\vartheta_{\text{out}}(p) dn \\ &= \vartheta_{\text{out}}(p)\varepsilon^{-1/\sigma}\frac{\gamma\left(-\dfrac{\tau - 2}{\sigma}, \varepsilon\right)}{\gamma\left(-\dfrac{\tau - 1}{\sigma}, \varepsilon\right)}, \\ \varepsilon &= \frac{|\vartheta_{\text{out}}(p) - \vartheta_c|}{\vartheta_c}. \end{aligned} \quad (16)$$

The volume $V_{\text{out}}(p)$ [Eq. (16)] is determined by the probability of empty pore formation at the given pressure and depends on $|\vartheta_{\text{out}}(p) - \vartheta_c|$,

$$V_{\text{out}}(p) = V_{\text{out}}(p, |\vartheta_{\text{out}}(p) - \vartheta_c|).$$

It follows from (16) that the ε parameter decreases as the pressure lowers, and, close to the percolation threshold,

$$\varepsilon = \frac{|\vartheta_{\text{out}}(p) - \vartheta_c|}{\vartheta_c} \longrightarrow 0.$$

Like the filled volume value [Eq. (7)], V_{out} remains finite at $\tau > 2$,

$$V_{\text{out}} \approx A(\tau, \sigma, \vartheta_c) - B(\tau, \sigma, \vartheta_c)\varepsilon^{(\tau - 2)/\sigma}.$$

An analysis shows that the relative volume of the pores capable of releasing the liquid [Eq. (15)] can be

represented in the form

$$\vartheta_{\text{out}}(p) \approx \int_{R_1(p)}^{R_2(p)} \frac{4}{3}\pi R^3 f(R)dR. \quad (17)$$

Here, $R_1(p)$ and $R_2(p)$ are found from condition (12) of the vanishing of the work ($\delta A_0 = 0$) that should be done to form an empty pore [17]. The probability of empty pore formation with radius $R_1(p) < R < R_2(p)$ is then nonzero. The outflow of a liquid from a porous body occurs as a result of the formation of clusters of empty pores in it. On the other hand, an analysis of the external pressure p dependences $R_1(p)$ and $R_2(p)$ shows that a small decrease in external pressure at

$$p < p_{0\,\max} = \frac{2\delta\sigma}{R_n}$$

corresponds to a substantial change in the relative volume of the pores capable of releasing the liquid, see (17). As a result, the difference between the nonscaling and scaling distribution functions [Eqs. (6) and (7), respectively] becomes essential to the description of the outflow of a liquid from a porous body.

The differentiation of (16) allows us to calculate the behavior of the susceptibility of the liquid–porous body system close to the outflow percolation threshold. In this case also, the critical susceptibility index is only determined by the critical index of the probability of infinite cluster formation,

$$\lambda = \frac{\partial V_{\text{out}}}{\partial p} \approx |\vartheta_{\text{in}}(p) - \vartheta_c|^{\beta-2}.$$

We stress that the work δA_0 [Eq. (12)] of empty pore formation depends on the difference of the surface energies of the liquid and the interface rather than their sum. For this reason, the p_0 pressure is lower than p_i. Like inflow, outflow occurs when condition (15) of the percolation transition is met, but at the p_{out} pressure, which is lower than p_{in}.

It follows that the hysteresis phenomenon in the system under consideration is related to different energy barriers to the development of inflow fluctuations and nonwetting liquid outflow; this difference is caused by the difference in the initial system states.

According to (13) and (4), the pressure p_0 changes sign and becomes negative for pores whose radius is smaller than R_0,

$$R_0 = R_{\min}\varphi^{1/3}\left(\frac{\sigma}{\delta\sigma} + 1\right)^{1/2}. \quad (18)$$

This means that energy barrier (12) to empty pore formation and outflow probability (11) can be close to one only at a pressure lower than atmospheric, that is, when the system is evacuated. For this reason, in experiments in which pressure is changed from atmospheric to $p > p_{\text{in}}$ and then decreased to atmospheric, the pores of radii $R < R_0$ should remain filled. In conformity with (18), the relative volume of the remaining liquid can be calculated by the equation

$$V_0 = \int_0^{R_0} \frac{4}{3}\pi R^3 f(R)dR. \quad (19)$$

A nonwetting liquid does not flow out of a porous body because energy expenditures for the formation of menisci when small-sized ($R < R_0$) (Eq. (12)) pores are emptied are not compensated for by the energy of gas–porous body interface formation in the pores.

Equations (8), (16), and (19) are used below to describe the experimental data obtained in this work (see Figs. 4–7). The volume of the system after liquid outflow and the volume of the liquid remaining in a porous body, as well as the volume of the system when the porous body is filled, depend on the parameters σ, $\delta\sigma$, $\bar{R}$, and δR and porosity φ. In our experiments, the surface energy of the liquid σ was varied (decreased) by adding ethylene glycol to water. The $\bar{R}$ and φ values for the porous media were taken from [19, 20].

The experimental data on the libersorb 2U-8–water system (surface energy $\sigma = 72$ mJ/m^2) are compared in Fig. 4a with the dependences calculated by (8), (16), and (19). Calculations of dependences (8) and (16) and volume (19) requires selecting the δR and $\delta\sigma$ parameters. The specific volume of libersorb 2U-8 pores is 0.38 cm^3/g (see table), whereas the specific pore volume of initial KSK-G silica gel, whose surface was modified to prepare libersorb 2U-8, is 0.73 cm^3/g. After the modification of KSK-G, the radius of pores decreases, and the mean radius of libersorb 2U-8 pores is therefore smaller than the $\bar{R} = 4$ nm value tabulated for KSK-G. Based on the measured specific volume of libersorb 2U-8 pores, their mean radius was estimated at $\bar{R} \approx 3.3$ nm. At the same time, small-sized libersorb 2U-8 pores can be filled by modifier molecules. For this reason, we selected the mean value $\bar{R} = 3.6$ nm for use in calculations. The pore-size distribution functions of KSK-G and libersorb 2U-8 are not known. We therefore used a Gauss distribution with the mean radius $\bar{R} = 3.6$ nm and the distribution half-width $\delta R = 0.4$ nm characteristic of porous media of the type of silica gels and silochroms ($2\delta R/\bar{R} \approx 20\%$). The $\delta\sigma$ parameter was adjusted to describe both $V_{\text{in}}(p)$ and $V_{\text{out}}(p)$ dependences and the volume V_0 of the liquid remaining in the porous body fairly accurately, to within measurement errors

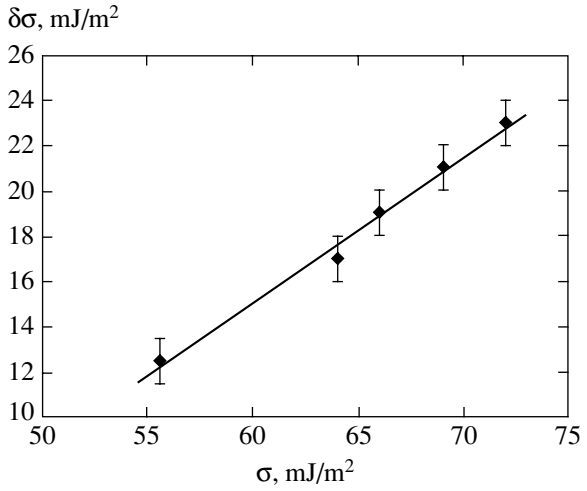

Fig. 8. Energy of liquid–solid interface formation as a function of the surface energy of the liquid for libersorb 2U-8 as a porous body.

(see Fig. 4a). The $\delta\sigma$ value answering this requirement was 23 mJ/m^2.

The experimental data on libersorb 2U-8 and aqueous ethylene glycol were described by setting the porous body parameters $\bar{R}$, δR, V_{por}, and φ equal to their values determined in experiments with pure water. The surface energy values of water–ethylene glycol mixtures were taken from the dependence shown in Fig. 1. It was found (see Figs. 4b, 4c) that the theoretical pressure dependences of system volume $V_{in}(p)$ (pressure increase) and $V_{out}(p)$ (pressure decrease) and the calculated volume of the solution remaining in the porous body coincided to within measurement errors with the experimental data for all ethylene glycol concentrations at the $\delta\sigma$ values listed in the table. The dependence of $\delta\sigma$ on σ (therefore, on the concentration of ethylene glycol) obtained this way is plotted in Fig. 8. The figure shows that $\delta\sigma$ monotonically decreases as the concentration of ethylene glycol increases and σ lowers. This dependence explains why the characteristic pressures p_2 of porous body filling decrease and the volume V_0 of the liquid remaining in the porous body increases as the concentration of ethylene glycol increases.

In agreement with (5), the pressure p_i can be written in the form

$$p_i = \frac{3\delta\sigma}{R}(1-\eta) + \frac{3\sigma}{R}\eta. \qquad (20)$$

According to (20), a decrease in σ and $\delta\sigma$ causes a decrease in p_i and, according to (5), the critical pressure p_{in}. It follows that a decrease in the inflow pressure as the concentration of ethylene glycol grows is caused by a decrease in the work δA spent for fluctuation pore filling.

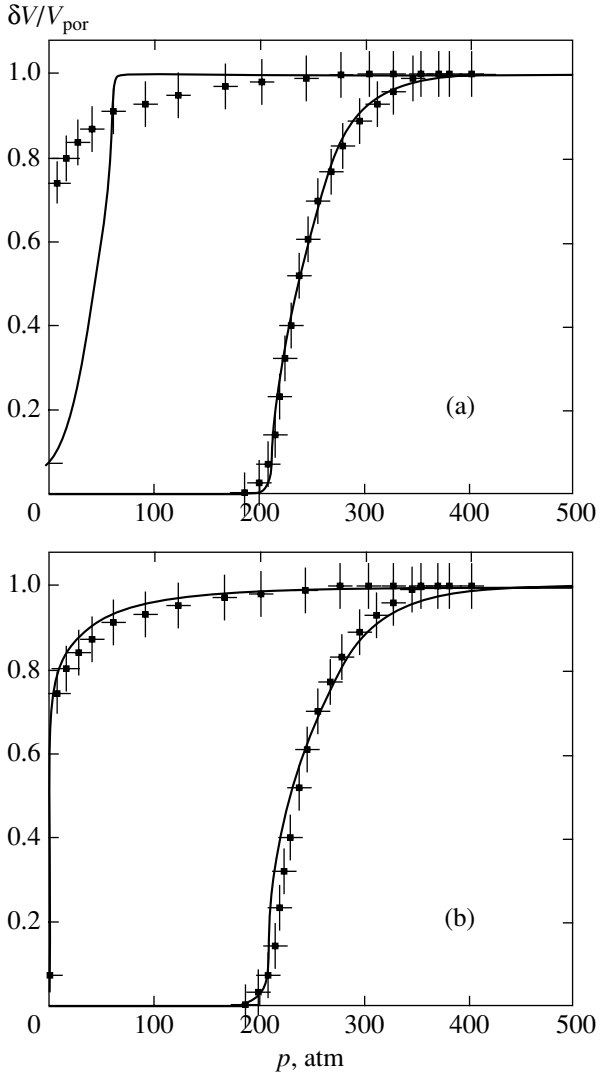

Fig. 9. Pressure dependences of changes in the relative liquid volume in a porous body for the libersorb 2U-8–water system [solid circles correspond to experimental data and solid lines to calculations by (14) and (16) with (a) scaling (7) and (b) nonscaling (6) functions for the distribution of clusters].

At the same time, an increase in the concentration of ethylene glycol increases the $\sigma/\delta\sigma$ ratio from 3.1 at C = 0 to 4.5 at C = 60%. According to (18), the radius R_0 simultaneously increases from 3.2 to 4.1 nm. In conformity with the Gauss pore-radius distribution, the pores largely have radii from $\bar{R} - \delta R$ to $\bar{R} + \delta R$. In addition, for the porous body under consideration,

$$R_{max} = \bar{R} + \delta R = 4.0 \text{ nm}.$$

We therefore have $R_0 > R_{max}$ at the concentration C = 60%. This means that the pressure at which the work of fluctuation empty pore formation is close to zero is lower than atmospheric. For this reason, almost all of the nonwetting liquid should remain in the porous body

($V_0 \approx 1$) at the concentration $C = 60\%$. Just this is observed for the system libersorb 2U-8–aqueous ethylene glycol (60%).

The theoretical dependences $V_{in}(p)$ and $V_{out}(p)$ and the calculated V_0 values are compared with the experimental data on silasorb S8, silasorb S18, and polysorb-1 as porous media and water as the liquid in Figs. 6 and 7. The $\bar{R}$, δR, and φ values used in the calculations are listed in the table, they correspond to the literature data on these systems [19, 20]. Dependences (8) and (17) describe the experimental data to within measurement errors. The $\delta\sigma$ values that we obtained are listed in the table. The special feature of silasorb S8–, silasorb S18–, and polysorb-1–aqueous ethylene glycol systems is complete liquid outflow after filling the porous bodies followed by pressure lowering irrespective of ethylene glycol concentration. For water and silasorb S8, silasorb S18, and polysorb-1 porous bodies, the R_0 radii (R_0 is the upper limit of the size of pores from which the liquid can flow out) were larger than the maximum pore radii of the corresponding pore-size distributions (see table). For instance, for the silasorb S8–water system, we have

$$R_{max} = \bar{R} + \delta R = 5.5 \text{ nm},$$

$\sigma/\delta\sigma = 6.9$, and, according to (13), $R_0 = 7.6$ nm.

4. CONCLUSIONS

To summarize, our analysis of the experimental data shows that nonwetting liquid inflow–outflow processes can be quantitatively described for the porous bodies studied on the basis of percolation theory taking into account energy barriers to fluctuation filling of pores with a liquid and empty pore formation as the pressure decreases. Our experiments showed that porous body filling occurred in a narrow neighborhood of the percolation threshold $\varepsilon \ll 1$. As distinct from what was done in [17], the experiments were described using nonscaling distribution function (6) for filled pore clusters, $F(n, p)$, which took into account the interaction of clusters with each other and the formation of fractal clusters of arbitrary dimensions. We found that taking such clusters into account close to the percolation inflow transition had no substantial (exceeding measurement errors) effect on the pressure dependence of the filled volume $V_{in}(p)$ for the porous bodies studied, which consisted of finite-size granules (Fig. 9). Indeed, the probability of porous body filling [Eq. (1)] is noticeably different from one only in a narrow neighborhood of the critical pressure p_{in} determined by the equation $\vartheta_{in}(p_{in}) = \vartheta_c$. Therefore, as follows from (8) and (9), porous body filling occurs only in a narrow neighborhood of the percolation threshold. Figure 9 shows that the difference between the nonscaling function that we used [Eq. (6)] and the scaling function [Eq. (7)] is of importance for describing liquid outflow from a porous body. Indeed, a small decrease in external pressure at

$$p < p_{0\,max} = \frac{2\delta\sigma}{R_n}$$

corresponds to a substantial change in the relative volume of the pores capable of releasing liquid, which requires taking into account fractal clusters of empty pores of arbitrary dimensions when the liquid flows out. The distribution of such clusters is described by nonscaling function (6).

It follows that, whereas the scaling function is capable of only qualitatively describing Wood alloy inflow into and outflow from silochroms as porous bodies [17], the nonscaling function gives a quantitative description of the experimental data on the liquid–porous body systems studied in this work. This is substantiated by the results shown in Fig. 9, from which it follows that the scaling function of the distribution of filled pore clusters incorrectly describes inflow–outflow processes and hysteresis phenomena.

Percolation theory taking into account the energy barriers to the development of inflow–outflow fluctuations is capable of describing the phenomenon of the nonoutflow of a nonwetting liquid from a porous body without assumptions on the special geometry of the space of pores such as were made in [9, 12–14]. It turns out that to remain filled is energetically favorable for small-sized pores for which the work δA_0 changes sign when the pressure of liquid outflow is lower than atmospheric. The hysteresis of nonwetting liquid inflow into and outflow from a porous body is also related to the energy conditions of the attainment of the percolation transition $\vartheta_{in,\,out} = \vartheta_c = 0.16$.

ACKNOWLEDGMENTS

The authors thank L.A. Maksimov for discussions.

REFERENCES

1. V. N. Bogomolov, Phys. Rev. B **51**, 17040 (1995).
2. P. S. Grinchuk, Phys. Rev. E **66**, 016124 (2002).
3. R. Cafiero, G. Caldarelh, and A. Gabrielli, Phys. Rev. E **56**, 1291 (1997).
4. T. Naheiri, K. A. Ludwig, M. Anand, *et al.*, Sep. Sci. Technol. **32**, 1589 (1997).
5. M. B. Rao and S. Sircar, J. Membr. Sci. **85**, 253 (1993).
6. V. N. Bogomolov, Usp. Fiz. Nauk **124**, 171 (1978) [Sov. Phys. Usp. **21**, 77 (1978)].
7. V. N. Bogomolov, Poverkhnost **9**, 136 (1992).
8. M. B. Isichenko, Rev. Mod. Phys. **64**, 961 (1992).
9. M. Sahimi, Rev. Mod. Phys. **65**, 1393 (1993).
10. A. Yu. Fadeev and V. A. Eroshenko, Ross. Khim. Zh. **39**, 93 (1995).

11. V. A. Eroshenko and A. Yu. Fadeev, Zh. Fiz. Khim. **70**, 1482 (1996).
12. G. P. Matthews, C. J. Ridgway, and M. C. Spearing, J. Colloid Interface Sci. **171**, 8 (1995).
13. J. Kloubek, J. Colloid Interface Sci. **163**, 10 (1994).
14. C.-Y. Park and S.-K. Ihm, AIChE J. **36**, 1641 (1990).
15. P. de Gennes, Usp. Fiz. Nauk **151**, 620 (1987).
16. A. A. Abrikosov, Pis'ma Zh. Éksp. Teor. Fiz. **29**, 72 (1979) [JETP Lett. **29**, 65 (1979)].
17. V. D. Borman, A. M. Grekhov, and V. I. Troyan, Zh. Éksp. Teor. Fiz. **118**, 193 (2000) [JETP **91**, 170 (2000)].
18. K. Nakanishi, T. Matsumoto, and M. Hayatsu, J. Chem. Eng. Data **16**, 1 (1971).
19. *The Hemapol Firm Catalog: Materials for Chromatography* (Czech Rep., 1997), p. 43.
20. K. I. Sakodynskiĭ and L. I. Panina, *Polymer Sorbents for Molecular Chromatography* (Nauka, Moscow, 1977) [in Russian].
21. *Chemistry of Grafted Surface Compounds*, Ed. by G. V. Lisichkin *et al.* (Fizmatlit, Moscow, 2003) [in Russian].
22. *Physical Values: Handbook*, Ed. by I. S. Grigor'ev and E. Z. Meĭlikhov (Énergoatomizdat, Moscow, 1991) [in Russian].
23. D. R. Lide, *Handbook of Chemistry and Physics* (CRC, Boca Raton, 1994).

Translated by V. Sipachev

Journal of Experimental and Theoretical Physics, Vol. 100, No. 2, 2005, pp. 398–409.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 127, No. 2, 2005, pp. 445–457.

STATISTICAL, NONLINEAR, AND SOFT MATTER PHYSICS

Non-Markovian Theory of Electron Paramagnetic Resonance in Localized and Quasi-Localized Electron Spins via an Example of Manganites with Colossal Magnetoresistance

E. Kh. Khalvashi

Batumi Polytechnical Institute, Georgian Technical University, Batumi, 384500 (6000) Georgia
e-mail: bpi@mail.ge; omaric@rambler.ru

Received June 3, 2004

Abstract—An approach based on the memory functions formalism is applied to derive non-Markovian equations of motion for the magnetization components of localized and quasi-localized electron spins under electron paramagnetic resonance (EPR) conditions using the example of manganites with colossal magnetoresistance. General Hasegawa–Bloch-type equations are applied to describe certain experimental data concerning the shape and the width of EPR lines and the longitudinal and transverse relaxation rates. Particular cases of these equations reproduce well-known theoretical results concerning EPR in manganites with colossal magnetoresistance. The results obtained explain certain well-known experimental phenomena and may stimulate further research. © *2005 Pleiades Publishing, Inc.*

1. INTRODUCTION

The formalism of memory functions is a powerful and simple method in nonequilibrium statistical physics. It is associated with the name of Zwanzig, who was the first to introduce a projection operator P into nonequilibrium statistical mechanics. He also obtained a basic kinetic equation for a statistical operator $\rho(t)$ and a non-Markovian basic kinetic equation that describes the evolution of a macroscopic state of a physical system. In addition, Zwanzig proposed a procedure for deriving kinetic transfer kernels (integral kernels of integro-differential equations—memory functions) in the most general form, which saved one the trouble of solving a rather difficult problem of constructing a nonequilibrium statistical operator $\rho(t)$. The formalism of memory functions is also associated with the name of Mori, who proposed a method for constructing a projection operator P (a superoperator) and thereby laid the foundation for modified nonequilibrium dynamics in its present form [1–3].

In the theory of magnetic resonance, the formalism of memory functions was first employed by Lado, Memory, and Parker [4–6].

The non-Markovian theory of magnetic resonance in solids for a system of localized spins with dominant dipole–dipole interaction was constructed in [7, 8].

In the present paper, we follow [7] and derive general expressions for the equations describing the trivial and nontrivial (relaxation) non-Markovian dynamics of magnetization for three components of localized electron spins and three components of quasi-localized electron spins (for *s* and *e* electrons, respectively) under EPR conditions. The general equations obtained in this paper will be used to determine the relaxation rates, the EPR line shape, and the temperature dependence of the EPR linewidth, as well as to investigate a relaxation bottleneck. For definiteness, we consider the spin system of manganite compounds.

Manganites with the perovskite structure that are doped with alkaline earth metals and described by the general formula $A_{1-x}A'_xMnO_3$ (where A = La, Pr, … and A' = Ca, Sr, Ba, …) are of much interest because of the colossal magnetoresistance effects observed in these materials. EPR investigations have been carried out on various samples of manganites with colossal magnetoresistance in a wide range of temperatures. In particular, the EPR linewidth, the transverse (T_2) and longitudinal (T_1) electron spin relaxation times, the electron spin susceptibility, and the behavior of the spin system as a function of temperature and the concentration of divalent ions, etc., have been studied (see, for example, [9–17]).

The formalism of memory functions was first applied to study manganites with colossal magnetoresistance in [9], where the authors analyzed and interpreted data concerning the EPR linewidth, which was identified with the relaxation rate of the transverse (with respect to the external dc magnetic field) component of the total spin. Before starting to solve the problem, i.e., to apply the memory functions formalism to construct non-Markovian dynamics for the spin system of a manganite, we make a few general remarks.

The drawback of this method (which, in our opinion, is insubstantial) is characteristic of all modern theories of nonequilibrium processes: this is the absence of a criterion for choosing a set of relevant operators. This

choice is made intuitively, with regard to previous experimental and theoretical data, and, as a rule, proves quite successful. If the set of relevant operators turns out to be incomplete, then the characteristics of a nonequilibrium process—kinetic coefficients (relaxation rates, transfer coefficients, etc.)—are determined only approximately. If this set contains more operators than necessary to describe a nonequilibrium state, then this has no effect on the calculation of kinetic coefficients [2]. Among the advantages of the formalism of memory functions are the following. The equations of dynamics obtained by this method are valid for any type of interaction and are exact, because no approximation has been used when deriving the basic kinetic equation of Zwanzig [1, 5, 6]. Moreover, as pointed out above, there is no need to apply a complicated procedure of constructing and using a nonequilibrium density matrix; this, which is quite important, allows one to avoid the introduction of the concept of spin temperature [6]. Finally, this method is simple in application because almost everything reduces to the calculation of commutators or (and) to operations with the derivatives of relevant operators (fluxes) under the sign of Sp in memory functions (in appropriate correlation functions composed of flux operators), as well as to the application of (Gaussian (most frequently), Lorentzian, etc.) approximations to the above-mentioned correlation functions [1, 18].

2. HAMILTONIAN

Consider a magnetic system as an ensemble of two subsystems of localized and quasi-localized electron spins with the Hamiltonian

$$H = H_{ex}^{is} + H_{ex}^{doub} + H^{anis} + H^{z} + H_1. \quad (1)$$

Here,

$$H_{ex}^{is} = \sum_{ij} \lambda_{ij}^{is} \mathbf{M}_{si} \mathbf{M}_{sj} \quad (2)$$

is the isotropic Heisenberg superexchange interaction between manganese ions at sites i and j (it may consist of two components with the exchange constants in the ac plane and between planes ac [9]),

$$H_{ex}^{doub} = H_H + H_t \quad (3)$$

is the double-exchange Hamiltonian [19],

$$H_H = -\lambda_H \sum_i \mathbf{M}_{si} \mathbf{M}_{ei} \quad (4)$$

is the Hund Hamiltonian,

$$H_t = -t \sum_{ij\sigma} (c_{i\sigma}^{\dagger} c_{j\sigma} + c_{i\sigma} c_{j\sigma}^{\dagger}) \quad (5)$$

is the Hamiltonian of hopping electrons, and

$$H^{anis} = H_{CF} + H_{DM}, \quad (6)$$

$$H_{CF} = \lambda_{CF}^{x} (M^x)^2 - \lambda_{CF}^{y} (M^y)^2 - \lambda_{CF}^{z} (M^z)^2 \quad (7)$$

is the Hamiltonian of the interaction of manganese ions with the noncubic crystalline field produced by neighboring oxygen ions [9]. The term

$$H_{DM} = \sum_{i>j} \frac{d_{ij}}{\gamma_s^2} \mathbf{M}_{si} \times \mathbf{M}_{sj} = \sum_{i>j} \sum_{\alpha\beta} \lambda_{ij} M_i^{\alpha} M_i^{\beta} = \sum_{i>j} (G_0^{DM} + G_1^{DM} + G_{-1}^{DM}) \quad (8)$$

describes the Dzialoshinsky–Moriya antisymmetric exchange interaction between localized spins [15, 20], where

$$G_0^{DM} = \frac{1}{2} i \lambda_{ij}^{xy} (M_{si}^{+} M_{sj}^{-} - M_{si}^{-} M_{sj}^{+}),$$

$$G_1^{DM} = \frac{1}{2} (\lambda_{ij}^{xz} - i\lambda_{ij}^{yz}) (M_{si}^{+} M_{sj}^{z} - M_{si}^{z} M_{sj}^{+}), \quad (9)$$

$$G_{-1}^{DM} = \frac{1}{2} (\lambda_{ij}^{xz} + i\lambda_{ij}^{yz}) (M_{si}^{-} M_{sj}^{z} - M_{si}^{z} M_{sj}^{-}),$$

$$\lambda_{ij}^{xy} = \frac{1}{2} (d_{ij}^{X} \sin\theta\cos\varphi - d_{ij}^{Y} \sin\theta\sin\varphi + d_{ij}^{Z} \cos\theta),$$

$$\lambda_{ij}^{xz} = d_{ij}^{X} \sin\varphi - d_{ij}^{Y} \cos\varphi, \quad (10)$$

$$\lambda_{ij}^{yz} = d_{ij}^{X} \cos\theta\cos\varphi - d_{ij}^{Y} \cos\theta\sin\varphi + d_{ij}^{Z} \sin\theta;$$

the axes X, Y, and Z are fixed crystallographic axes a, b, and c; the axis Z is parallel to the axis c; and the external magnetic field $\mathbf{H}_0$ is parallel to the axis z and is directed at polar (θ) and azimuthal (φ) angles with respect to the axis c,

$$H^{z} = -\mathbf{H}_0 \cdot (\gamma_s \mathbf{M}_s + \gamma_e \mathbf{M}_e) \quad (11)$$

is the Zeeman interaction between spins and the field $\mathbf{H}_0$, and H_1 is the interaction between spins and an external ac (radio-frequency (RF)) field,

$$M_s = \gamma_s \sum_i S_i, \quad M_e = \gamma_e \sum_i s_i = \gamma_e \sum_i c_{i\sigma}^{\dagger} \tau_{\sigma\nu} c_{i\nu}$$

are the magnetizations of localized M_s and quasi-localized M_e spins; $c_{i\sigma}^{\dagger}(c_{i\sigma})$ are the creation (annihilation) operators of a quasi-localized electron with the spin orientation $\boldsymbol{\sigma}(\nu)$; and $\tau_{\sigma\nu}$ are the Pauli spin matrices,

$$\lambda_{ij}^{is} = \frac{2J_{ij}}{\gamma_s^2}, \quad \lambda_{CF}^{x} = \lambda_{CF}^{y} = \frac{E}{\gamma_s^2},$$

$$\lambda_{CF}^{z} = \frac{D}{\gamma_s^2}, \quad \lambda_H = \frac{2J_H}{g_s g_e \mu_B^2},$$

where J_{ij}, J_H, and d_{ij} are the constants of superexchange, Hund, and Dzialoshinsky–Moriya exchange interactions; E and D are the crystalline-field constants; t is the hopping integral; $\gamma_{s,e}$ and $g_{s,e}$ are the gyromagnetic ratio and the g factor of the spins S and s, respectively; and μ_B is the Bohr magneton.

3. EQUATIONS OF MOTION

In this case, the choice of a relevant set is obvious: these are the x, y, and z components of the magnetization operators M_k^α ($\alpha = x, y, z$, $k = s, e$) of localized and quasi-localized electron spins. Using these components, we will describe the dynamics of the spin system of a manganite. In contrast to [7], where the secular part of the dipole–dipole interaction takes part in the nontrivial spin dynamics by broadening the magnetic resonance linewidth [21], we do not take, say H_{ex}^{is}, as a relevant operator, because it commutes with the x, y, and z components of magnetization.

In such a statement, the Mori projection operator is expressed as

$$P = \sum_{\alpha k} |M_k^\alpha\rangle\langle M_k^\alpha|/\langle M_k^\alpha|M_k^\alpha\rangle, \tag{12}$$

where $\langle Q|Q\rangle = \mathrm{Sp}(Q)^2$. Moreover, the identity $P^2 = P$ holds, which is characteristic of this superoperator.

Following [5–7], we can obtain equations of motion for the magnetizations M_k^α:

$$\begin{aligned}\sum_{k\alpha}\frac{d\langle M_k^\alpha(t)\rangle}{dt} \\ = \sum_{k\alpha}\left[\left(\frac{d\langle M_k^\alpha(t)\rangle}{dt}\right)_{TD} + \left(\frac{d\langle M_k^\alpha(t)\rangle}{dt}\right)_{NTD}\right],\end{aligned} \tag{13}$$

where the terms

$$\begin{aligned}\sum_{k\alpha}\left(\frac{d\langle M_k^\alpha(t)\rangle}{dt}\right)_{TD} \\ = -\mathrm{Sp}\left\{\sum_{k\alpha}[M_k^\alpha, H]\sum_{l\beta}\frac{M_l^\beta}{\mathrm{Sp}(M_l^\beta)^2}\right\}\langle M_l^\beta(t)\rangle\end{aligned} \tag{14}$$

and

$$\begin{aligned}\sum_{k\alpha}\left(\frac{d\langle M_k^\alpha(t)\rangle}{dt}\right)_{NTD} \\ = -\sum_{k\alpha}\int_0^t dt' K_k^\alpha(\tau)\langle M_k^\alpha(t')\rangle\end{aligned} \tag{15}$$

describe trivial and nontrivial (relaxation) dynamics, respectively; the functions

$$K_k^\alpha(\tau) = \mathrm{Sp}\left\{[M_k^\alpha, H]\sum_{l,\beta}\frac{[M_l^\beta, H(\tau)]}{\mathrm{Sp}(M_l^\beta)^2}\right\} \tag{16}$$

represent memory functions, where $\tau = t - t'$, $\alpha \neq \beta = x, y, z$, $k \neq l = s, e$; and

$$\langle Q(t)\rangle = \mathrm{Sp}\{Q\rho(t)\}, \quad Q(\tau) = e^{-iH\tau}Q(0)e^{iH\tau}.$$

Note that, to simplify the problem, we made the following assumptions when deriving the integrals (15) and (16) of nontrivial dynamics. In the exponential multipliers of the memory functions, we took into account that

$$(1-P)H = H_{ex}^{is} + H^{doub} + H_{ex}^{anis},$$

since

$$P(H_{ex}^{is} + H^{doub} + H_{ex}^{anis}) = 0, \quad P(H^z + H_1) = H^z + H_1.$$

In the Hamiltonians contained in the commutators, we dropped the Zeeman and RF terms (the latter terms were dropped in (14) as well) and then took into account that

$$(1-P)[M_k^\alpha, H(\tau)] = [M_k^\alpha, H(\tau)],$$

because

$$\langle M_k^\alpha|H\rangle = \langle M_k^\alpha|[M_k^\alpha, H(\tau)]\rangle = 0.$$

If we calculate the commutators in (14) and (15) (see also (16)) and drop, for short, the angular brackets and the brackets containing functions of time t, then from Eqs. (13) we obtain

$$\begin{aligned}\frac{dM_s^\alpha}{dt} = \left(\frac{dM_s^\alpha}{dt}\right)_{HB} + \left(\frac{dM_s^\alpha}{dt}\right)_{CF} \\ + \left(\frac{dM_s^\alpha}{dt}\right)_{DM}^{dir} + \left(\frac{dM_s^\alpha}{dt}\right)_{DM}^{cr},\end{aligned} \tag{17}$$

$$\frac{dM_e^\alpha}{dt} = \left(\frac{dM_e^\alpha}{dt}\right)_{HB}. \tag{18}$$

Here, the terms

$$\left(\frac{dM_s^\alpha}{dt}\right)_{HB} = \left(\frac{dM_s^\alpha}{dt}\right)_{HB}^{TD} + \left(\frac{dM_s^\alpha}{dt}\right)_{HB}^{NTD} - \frac{M_s^\alpha - M_{s0}^\alpha}{T_s^\alpha}, \tag{19}$$

and

$$\left(\frac{dM_e^\alpha}{dt}\right)_{HB} = \left(\frac{dM_e^\alpha}{dt}\right)_{HB}^{TD} - \frac{\gamma_e}{\gamma_s}\left(\frac{dM_s^\alpha}{dt}\right)_{HB}^{NTD} - \frac{M_e^\alpha - M_{e0}^\alpha}{T_e^\alpha} \tag{20}$$

represent analogs of the Hasegawa–Bloch equations, which were obtained earlier phenomenologically to describe EPR in metals (one can even say that for-

mulas (17) and (18) are Hasegawa–Bloch-type equations) [22, 23]. The terms

$$\left(\frac{d\mathbf{M}^{\alpha}_{\mathrm{s,e}}}{dt}\right)^{\mathrm{TD}}_{\mathrm{HB}} = \gamma_{\mathrm{s,e}}\mathbf{M}_{\mathrm{s,e}} \times (\mathbf{H}_0 + \lambda_{\mathrm{H}}\mathbf{M}_{\mathrm{e,s}}) \qquad (21)$$

describe the trivial dynamics of the operators M^{α}_{s} and M^{α}_{e} (here, as pointed out above, we neglected the RF field). The terms $(M^{\alpha}_{\mathrm{s,e}} - M^{\alpha}_{\mathrm{s,0}})/T^{\alpha}_{\mathrm{s,eL}}$, which are added to the right-hand side of Eqs. (19) and (20) phenomenologically, describe the spin–lattice relaxation of *s* and *e* spins; for clarity, they are extracted from the nontrivial dynamics (these terms can also be obtained by the memory functions formalism by adding appropriate interactions to the Hamiltonian of the problem). The term

$$\left(\frac{dM^{\alpha}_{\mathrm{s}}}{dt}\right)^{\mathrm{NTD}}_{\mathrm{HB}} = -\int_0^t dt' K^{\alpha}_{\mathrm{H}}(\tau) \times \left[M^{\alpha}_{\mathrm{s}}(t') - \frac{\gamma_{\mathrm{e}}}{\gamma_{\mathrm{s}}}\frac{\mathrm{Sp}(M^{\alpha}_{\mathrm{s}})^2}{\mathrm{Sp}(M^{\alpha}_{\mathrm{e}})^2}M^{\alpha}_{\mathrm{e}}(t')\right] \qquad (22)$$

represents an explicit expression for the nontrivial dynamical part of Eqs. (19) and (20) (without spin–lattice dynamics). The terms

$$\left(\frac{dM^{\alpha}_{\mathrm{s}}}{dt}\right)_{\mathrm{CF}} = -\int_0^t dt' K^{\alpha}_{\mathrm{CF}}(\tau)M^{\alpha}_{\mathrm{s}}(t') \qquad (23)$$

and

$$\left(\frac{dM^{\alpha}_{\mathrm{s}}}{dt}\right)^{\mathrm{dir}}_{\mathrm{DM}} = -\int_0^t dt' K^{\alpha}_{\mathrm{DM}}(\tau)M^{\alpha}_{\mathrm{s}}(t') \qquad (24)$$

describe the relaxation of the components of M^{α}_{s} due to the interaction with the crystalline field and due to the Dzialoshinsky–Moriya interaction, respectively. The term

$$\left(\frac{dM^{\alpha}_{\mathrm{s}}}{dt}\right)^{\mathrm{cr}}_{\mathrm{DM}} = -\int_0^t dt' [K^{\alpha\beta}_{\mathrm{DM}}(\tau)M^{\beta}_{\mathrm{s}}(t') + K^{\alpha\gamma}_{\mathrm{DM}}(\tau)M^{\gamma}_{\mathrm{s}}(t')] \qquad (25)$$

is a cross contribution of two components of M^{α}_{s} to the relaxation of the third component. The quantities

$$K^{\alpha}_{\mathrm{H,CF}}(\tau) = K^{\alpha}_{\mathrm{H,CF}}(0)G^{\beta\gamma}_{\mathrm{H,CF}}(\tau), \qquad (26)$$

$$K^{\alpha}_{\mathrm{DM}}(\tau) = K^{\alpha\beta\mathrm{dir}}_{\mathrm{DM}}(0)G^{\alpha\gamma}_{\mathrm{DM}}(\tau) + K^{\alpha\gamma\mathrm{dir}}_{\mathrm{DM}}(0)G^{\alpha\beta}_{\mathrm{DM}}(\tau), \qquad (27)$$

$$K^{\alpha\beta,\gamma}_{\mathrm{DM}}(\tau) = K^{\alpha\beta,\gamma\mathrm{cr}}_{\mathrm{DM}}(0)G^{\alpha\beta}_{\mathrm{DM}}(\tau) \qquad (28)$$

are memory functions obtained from (16);

$$K^{\alpha}_{\mathrm{H,CF}}(0) = (\gamma_{\mathrm{s}}\lambda^{\alpha}_{\mathrm{H,CF}})^2\sum_i \frac{\mathrm{Sp}(F^{\beta\gamma}_{i\mathrm{H,CF}})^2}{\mathrm{Sp}(M^{\alpha}_{\mathrm{s}})^2}, \qquad (29)$$

$$K^{\alpha\beta,\gamma\mathrm{dir}}_{\mathrm{DM}}(0) = \sum_{i>j}(\gamma_{\mathrm{s}}\lambda^{\alpha\beta,\gamma}_{ij})^2\frac{\mathrm{Sp}(F^{\alpha\beta,\gamma}_{ij})^2}{\mathrm{Sp}(M^{\alpha}_{\mathrm{s}})^2}, \qquad (30)$$

$$K^{\alpha\beta,\gamma\mathrm{cr}}_{\mathrm{DM}}(0) = \gamma^2_{\mathrm{s}}\sum_{i>j}A^{\alpha\beta,\gamma}_{ij}\frac{\mathrm{Sp}(F^{\alpha\beta,\gamma}_{ij})^2}{\mathrm{Sp}(M^{\beta,\gamma}_{\mathrm{s}})^2} \qquad (31)$$

are the contributions of the Hund and anisotropic interactions to the second moment M_2 of the resonance line because $M_2 = K(0)$ [12, 13]; and

$$G^{\beta\gamma}_{\mathrm{H,CF}}(\tau) = \frac{\mathrm{Sp}F^{\beta\gamma}_{i\mathrm{H,CF}}(0)F^{\beta\gamma}_{i\mathrm{H,CF}}(\tau)}{\mathrm{Sp}(F^{\beta\gamma}_{i\mathrm{H,CF}})^2}, \qquad (32)$$

$$G^{\alpha\beta,\gamma}_{\mathrm{DM}}(\tau) = \frac{\mathrm{Sp}F^{\alpha\beta,\gamma}_{ij}(0)F^{\alpha\beta,\gamma}_{ij}(\tau)}{\mathrm{Sp}(F^{\alpha\beta,\gamma}_{ij})^2}, \qquad (33)$$

are four-spin correlation functions. Here, we also used the notation

$$A^{xy}_{ij} = \lambda^{xz}_{ij}\lambda^{yz}_{ij}, \quad A^{zx}_{ij} = \lambda^{zx}_{ij}\lambda^{xy}_{ij},$$

$$A^{yz}_{ij} = \lambda^{xy}_{ij}\lambda^{yz}_{ij},$$

$$F^{\beta\gamma}_{i\mathrm{H,CF}} = M^{\beta}_{si}M^{\gamma}_{\mathrm{e}i,si} - M^{\gamma}_{si}M^{\beta}_{\mathrm{e}i,si}, \qquad (34)$$

$$F^{\alpha\beta,\gamma}_{ij} = M^{\alpha}_{si}M^{\beta,\gamma}_{sj} - M^{\beta,\gamma}_{si}M^{\alpha}_{si},$$

$$\alpha \neq \beta \neq \gamma = x, y, z.$$

Note that we used a molecular-field approximation [24] when deriving Eqs. (21) of trivial dynamics. The trivial part of Eqs. (21) shows that localized and quasi-localized spins shift the resonance frequencies of each other.

Note also that, after switching off the RF field, the quantities $\langle M^{x,y}(t)\rangle$ tend to zero, while $\langle M^{z}(t)\rangle$ tends to its equilibrium value $\langle M^{z}\rangle_0 = M^{z}_0$. Therefore, in all expressions for $\langle M^{z}(t)\rangle$ in Eqs. (22)–(25) (recall that $\langle Q(t)\rangle = \mathrm{Sp}\{Q\rho(t)\}$), one should change $\rho \longrightarrow \rho - \rho_0$ (ρ_0 is the equilibrium density matrix), as is done in [25]; i.e., one should take $\langle M^{z}(t)\rangle - M^{z}_0$ instead of $\langle M^{z}(t)\rangle$.

Note also that, in contrast to [9], relaxation in our analysis is of tensorial nature. Moreover, one can see that, due to the presence of products of different

constants of the Dzialoshinsky–Moriya interaction (see (25), (28), (31), and (33)), the contribution of cross relaxation rates to the relaxation of localized spins depends on the structure of a manganite with colossal magnetoresistance and on experimental conditions (the direction of the magnetic field, etc.). For example, if we assume that the vector $\mathbf{d}_{ij}$ is directed along axis c and the field $\mathbf{H}_0$ is parallel (i.e., $\theta = 0$) or perpendicular (i.e., $\theta = \pi/2$) to axis c in Eqs. (8)–(10), then, according to (34), the contribution of cross relaxation (25) vanishes. The contribution of the cross relaxation also vanishes, for example, when $\theta = \pi/2$, $\varphi = \pi/4$, $d_{ij}^X = d_{ij}^Y$, and $d_{ij}^Z \neq 0$. If d_{ij}^X, d_{ij}^Y, $d_{ij}^Z \neq 0$, then one may face different, more complicated, situations. In particular, for an arbitrary direction of $\mathbf{H}_0$ (for arbitrary polar θ and azimuthal φ angles), the contribution of the cross relaxation (26) is different from zero in most cases. Moreover, when, for example, $\theta = \varphi = 0$, i.e.,

$$\lambda_{ij}^{xy} = d_{ij}^Z, \quad \lambda_{ij}^{yz} = d_{ij}^X, \quad \lambda_{ij}^{xz} = -d_{ij}^Y,$$

the contribution of localized s spins to relaxation may be either positive or negative (the EPR linewidth may either increase or decrease). Note that the tensorial nature of relaxation and, possibly, a significant contribution of the cross terms to the relaxation of localized spins, which is revealed in our theoretical analysis, may also manifest themselves in the variation of, say the EPR linewidth, under the variation of experimental conditions (magnetic field direction, etc.) or when various samples of manganite with colossal magnetoresistance are used in the experiment.

One can obtain the following well-known results from Eqs. (17) and (18).

1. In the equilibrium state, one can easily derive the following relations from Eqs. (17) and (18) (more precisely, from Eqs. (19)–(21)) [23, 26]:

$$\chi_s^\alpha = \chi_{s0}^\alpha \frac{1 + \lambda_H \chi_{e0}^\alpha}{1 - \lambda_H^2 \chi_{s0}^\alpha \chi_{e0}^\alpha},$$
$$\chi_e^\alpha = \chi_{e0}^\alpha \frac{1 + \lambda_H \chi_{s0}^\alpha}{1 - \lambda_H^2 \chi_{s0}^\alpha \chi_{e0}^\alpha}. \tag{35}$$

It should be noted that the classical expressions, with which we compare relations (35), contain individual susceptibilities for Mn^{3+} and Mn^{4+} ions. According to our model, one of the susceptibilities in (35) is associated with the total magnetization of localized s spins (of Mn^{3+} and Mn^{4+} ions), and the other is associated with the magnetization of quasi-localized electron spins e. When the isotropic exchange interaction is dominant, this model seems to be more general and adequate.

2. Let us take $M_s^\alpha(t)$ and $M_e^\alpha(t)$ outside the sign of integral in Eqs. (23) and replace the upper limit of integration by infinity; this procedure is incorrect as applied to $M_e^\alpha(t)$, but it allows one to qualitatively estimate the behavior of the spin system of quasi-localized electrons e. Then, in the equilibrium state, we obtain

$$T_{es}^\alpha / T_{se}^\alpha = g_s \chi_e^\alpha / g_e \chi_s^\alpha. \tag{36}$$

This is a well-known result (see, for example, [27]).

Quantitative estimates for the transverse and longitudinal relaxation rates; the shape and width of the resonance line; and their dependence on the impurity concentration, temperature, and sample structure can be obtained from a rigorous solution and a detailed analysis of Eqs. (17) and (18). However, such as analysis is beyond the scope of the present paper.

4. THE WIDTH AND THE SHAPE OF EPR LINE: APPROXIMATIONS AND CONCLUSIONS

Let us consider several approximations for the correlation functions (32) and (33) and the second moments (30) and (31), which constitute the memory functions (26)–(28). This might allow us to assess the shape and the width of a resonance line and its temperature dependence, as well as the "type" of a spectral EPR line.

4.1. Gaussian Approximation for Correlation Functions: Shape and Width of a Resonance Line (Relation between the Second Moments and Correlation Times of Memory Functions with the Second and Fourth Moments of EPR Lines)

For the memory functions (16) (see also their explicit expressions (26)–(28)), or, more precisely, for the correlation functions (32) and (33) appearing in them, we can take the well-known Gaussian approximation [5, 6, 8, 24]

$$G_{H,CF}^{\beta\gamma}(\tau) \propto \exp(-N_{2H,CF}^\alpha \tau^2/2), \tag{37}$$

$$G_{DM}^{\alpha\beta,\gamma}(\tau) \propto \exp(-N_{DM}^\alpha \tau^2/2), \tag{38}$$

where

$$N_2 = M_2(\mu - 1), \quad \mu = M_4/M_2^2, \tag{39}$$

N_2 is the second moment of a memory function, and M_2 and M_4 are the second and fourth spectral moments of the EPR line.

Note that, if we use the Gaussian approximation

$$G(\tau) \propto \exp(-\tau^2/\tau_{INT}^2),$$

defined by the correlation time τ_{INT} (where "INT" = "H," "CF," "DM"), for the correlation functions (32)

and (33), then we can easily establish a relation between the corresponding quantities N_2 and τ_{INT}:

$$\tau_{INT} = \left(\frac{N_{2INT}^{\beta\gamma}}{2}\right)^{-1/2} = \left[\frac{M_{2INT}^2}{2}\left(\frac{M_{4INT}}{2M_{2INT}^2} - 1\right)\right]^{-1/2}. \quad (40)$$

According to [5, 6], we can argue that

$$\text{(a)} \quad \tau_{INT} = \left(\frac{N_{2INT}^{\beta\gamma}}{2}\right)^{-1/2} = \left(\frac{M_{4INT}}{2M_{2INT}^2}\right)^{-1/2},$$

under exchange narrowing, because $M_4/M_2^2 \gg 1$ in this case and the shape of the EPR line is close to the Lorentzian shape;

(b) in the absence of exchange narrowing, $\tau_{INT} = M_{2INT}^{-1/2}$, because $M_4/M_2^2 \approx 3$ in this case and the shape of the EPR line is of Gaussian form.

Note that the superexchange interaction H_{ex}^{is} makes a contribution to the resonance line of localized spins only via N_{2CF} and N_{2DM} (via M_{4CF} and M_{4DM}); therefore, only localized spins can contribute to the exchange-narrowed EPR line with a Lorentz-type shape, whereas spins of both types contribute to the EPR line of Gaussian form. It should also be noted that, if there is a single EPR line, then it certainly contains contributions of spins of both types; however, since one of the interactions H^{doub} or H_{ex}^{is} is dominant, either the Gaussian line dominates the Lorentzian line or vice versa (the case of a single Lorentz-type line is considered below in Section 6 and in the Appendix). If the resonance frequencies of localized and quasi-localized spins are sufficiently well separated so that there are two EPR lines, then each of them can be identified with a certain type of spins, provided that the line shape is purely Gaussian or purely Lorentzian (the Lorentzian line corresponds to localized and Gaussian, to quasi-localized spins). When both lines are Gaussian (in this case, $H^{doub} > H_{ex}^{is}$), the identification can be made only by resonance frequencies. Thus, the shape and the resonance frequency of an experimental EPR line may help one to determine which of the spin–spin interactions dominates and to which type of electrons the resonance line corresponds. Here, one must keep in mind that the reality may be more complicated because interactions that are not taken into account by the present theoretical approach (perturbations of the electronic structure of manganese ions and the crystalline field that are associated with hopping conductivity, inhomogeneity of a magnetic field, the presence of unaccounted impurities, etc.) may also contribute to the width and the shape of the EPR line.

4.2. Approximation from the Bloch–Vangsness–Redfield Theory: Spin in a Fluctuating Field

In this approximation [25], we can represent, for example, the product $M_{si}^{\beta}M_{ei}^{\gamma}$ taken from the four-spin correlation functions (32) and (33) as

$$M_{si}^{\beta}M_{ei}^{\gamma} \approx \frac{1}{2\lambda_H}(M_{si}^{\beta}H_{ei}^{\gamma} + M_{ei}^{\gamma}H_{si}^{\beta}), \quad (41)$$

where $H_{s,ei}^{\alpha} = \sqrt{\lambda_H^2\langle(\delta M_{s,ei}^{\alpha})^2\rangle}$ are the rms fluctuations of the fields produced by the spin $s(e)$ and acting on the spin $e(s)$ and $\delta M_{s,ei}^{\alpha} = M_{s,ei}^{\alpha} - \langle M_{s,ei}^{\alpha}\rangle$. Moreover, the theory of [25] uses the approximation

$$\overline{H_{s,ei}^{\alpha}H_{s,ei}^{\alpha}(t)} = (H_{s,ei}^{\alpha})^2\exp(-t/\tau_{e,s}), \quad (42)$$

where the upper bar denotes averaging over an ensemble and $\tau_{s,e}$ are the correlation times of the fluctuating fields created by the localized s and quasi-localized e electron spins, respectively.

Now, using Eqs. (17) and (18) and the approximations of the theory of [25] given above, we obtain, for example, expressions for the rates $(T_{se}^{\alpha})^{-1}$ of the Hund relaxation of localized spins s to quasi-localized spins e. To this end, we write the expression for $(T_{se}^{\alpha})^{-1}$ in explicit form, using relations (23), (29), and (34):

$$(T_{se}^{\alpha})^{-1} \approx \int_0^{\infty} dt K_H^{\alpha}(t) = \frac{(K_H^{\alpha}(0))^2}{\mathrm{Sp}(F_{iH}^{\beta\gamma})^2}\int_0^{\infty} dt \times \mathrm{Sp}(M_{si}^{\beta}M_{ei}^{\gamma} - M_{si}^{\gamma}M_{ei}^{\beta})(M_{si}^{\beta}M_{ei}^{\gamma} - M_{si}^{\gamma}M_{ei}^{\beta})(t). \quad (43)$$

Note that the time dependence of the four-spin correlation functions under the sign of integral in (43) is determined by exponential multipliers of the form $\exp[i(H_{DM} + H_{CF} + H_H + H_s^z + H_e^z)t]$ with noncommutative operators (the operator $H_{DM} + H_{CF} + H_H$ does not commute with H_s^z, and H_H does not commute with H_e^z).

Thus, using formula (43) and taking into account that, for example,

$$\exp(-i\omega_s S^z t)M_{si}^y\exp(i\omega_s S^z t) = M_{si}^y\cos\omega_s t - M_{si}^x\sin\omega_s t,$$

we can easily derive

$$(T_{se}^{\perp})^{-1} = \left(\frac{\gamma_s}{2}\right)^2 \times \sum_i \left\{ (H_{ei}^{\perp})^2 \tau_e + (H_{ei}^z)^2 \frac{\tau_c}{1+\omega_s^2\tau_e^2} + \frac{\mathrm{Sp}(M_{ei}^{\alpha})^2}{\mathrm{Sp}(M_{si}^{\alpha})^2} \times \left[(H_{si}^{\perp})^2 \tau_s + (H_{si}^z)^2 \frac{\tau_s}{1+\omega_e^2\tau_s^2} \right] \right\}, \quad (44)$$

$$(T_{se}^{z})^{-1} = \left(\frac{\gamma_s}{2}\right)^2 \sum_i \left\{ (H_{ei}^{y})^2 \frac{\tau_e}{1+\omega_s^2\tau_e^2} + \frac{\mathrm{Sp}(M_{ei}^{\alpha})^2}{\mathrm{Sp}(M_{si}^{\alpha})^2} (H_{si}^{x})^2 \frac{\tau_s}{1+\omega_e^2\tau_s^2} \right\}, \quad (45)$$

where $\perp = x, y$, and $\omega_{s,e}$ are the Zeeman frequencies of s and e spins. Note that thermal averaging allows one to make the following substitution in (44) and (45):

$$\mathrm{Sp}(M_e^{\alpha})^2/\mathrm{Sp}(M_s^{\alpha})^2 = \chi_e^{\alpha}(T)/\chi_s^{\alpha}(T).$$

One can see that the expressions obtained resemble those used in the Bloch–Vangsness–Redfield theory for T_2^{-1} and T_1^{-1} (see formulas (5.210) from [25]). An analysis of similar expressions can be found in [25]. Recall that the superexchange interaction H_{ex}^{is} does not contribute to the relaxation rates $(T_{se}^{\alpha})^{-1}$ and, just as in [25], these rates are proportional to $(H_{s,e}^{\alpha})^2$. Note that τ_e is different from τ_s because $[M_e^{\alpha}, H^{anis}] = 0$ and $[M_s^{\alpha}, H^{anis}] \neq 0$; however, this difference may be inconsequential since $M_{s,e}^{\alpha}$ does not commute with the interaction H^{doub} (more precisely, with its Hund component H_H), which may prove dominant in manganites with colossal magnetoresistance in a certain range of impurity concentrations and temperatures. Moreover, due to the difference between the approximation of fluctuating fields and the approximation associated with the Gaussian approximation in the memory functions formalism, the quantities $\tau_{e,s}$ and τ_{INT} differ from each other: the former characterize the correlation between $H_{s,e}^{\alpha}(0)$ and $H_{s,e}^{\alpha}(t)$, i.e., actually between $M_{s,e}^{\alpha}(0)$ and $M_{s,e}^{\alpha}(t)$, whereas the latter characterize the correlation between $\dot{M}_{s,e}^{\alpha}(0)$ and $\dot{M}_{s,e}^{\alpha}(t)$, i.e., between $[M_{s,e}^{\alpha}, H](0)$ and $[M_{s,e}, H](t)$ from (16). In our opinion, the advantage of the memory functions formalism consists in the fact that the times $\tau_{H,CF,DM}$ can be calculated quantitatively, in terms of the second moment of an appropriate memory function (38), i.e., in terms of the corresponding second and fourth moments of the EPR line (in terms of M_2 and M_4). This formalism also allows one to take into account higher order moments of the EPR line [5, 6].

Finally, note that in the fast-motion limit, we have $T_{se}^{\perp} = T_{se}^{z}$, whereas, in the opposite case, when $\omega_{s,e}\tau_e \ll 1$, the ratio $T_{se}^{z}/T_{se}^{\perp}$ is large [10–12].

4.3. Temperature Approximation: Relaxation and Resonance Linewidth

Following [9], we can obtain the following general expression for the kinetic coefficients of Eqs. (22)–(25) of nontrivial dynamics, more precisely, for their non-Markovian version (49)–(52), i.e., for the spin relaxation rates:

$$(T_{s,e}^{\alpha})^{-1} = \frac{\chi_{0s,e}}{\chi_{s,e}(T)} \frac{1}{4kC\gamma_{s,e}^2} \times \int_{-\infty}^{\infty} dt \mathrm{Sp} \frac{dM_{s,e}^{\alpha}(t)}{dt} \frac{dM_{s,e}^{\alpha}(0)}{dt}. \quad (46)$$

Here, the expressions $\mathrm{Sp}\{[dM_{s,e}^{\alpha}(t)/dt][dM_{s,e}^{\alpha}(0)/dt]\}$ represent the numerators of the memory functions from (16) in the general form, $\chi_{s,e}(T)$ is the temperature-dependent susceptibility, $\chi_{0s,e} = C/T$ is the Curie susceptibility, C is the Curie constant, k is the Boltzmann constant, and T is temperature.

Taking into account (46), we can easily obtain the Huber law for the EPR linewidth [9]:

$$\Delta H_{s,e}(T) = \frac{2(T_{s,e}^{\perp})^{-1}}{\sqrt{3}\gamma_{s,e}k} = \frac{\chi_{0s,e}}{\chi_{s,e}(T)} \Delta H_{s,e}(\infty), \quad (47)$$

where

$$\Delta H_{s,e}(\infty) = \frac{2}{\sqrt{3}\gamma_{s,e}^3 kC} \times \int_{-\infty}^{\infty} dt \mathrm{Sp} \frac{dM_{s,e}^{\alpha}(t)}{dt} \frac{dM_{s,e}^{\alpha}(0)}{dt}. \quad (48)$$

In the case of a paramagnetic state considered here, the most reasonable approximation for the susceptibility is the Curie–Weiss law [10–12],

$$\chi_{s,e}(T) = C/(T-\Theta),$$

where Θ is the Curie–Weiss temperature. The substitution of the Curie–Weiss law into (47) yields a good agreement with the set of experimental data for the temperature dependence of the susceptibility, the EPR line-

width, and the transverse and longitudinal relaxation rates in a wide range of temperatures and impurity concentrations [10–12].

5. RELAXATION BOTTLENECK

Now, we demonstrate the potential of Eqs. (17) and (18) as applied to investigating relaxation and relaxation bottleneck in a coupled system of localized *s* and quasi-localized *e* spins in the cases of weak and strong coupling between them.

For simplicity, we neglect the contribution of cross terms, $(dM_s^\alpha/dt)_{DM}^{cr} = 0$ (see the remark in Section 3). In addition, to simplify the nontrivial part of Eqs. (17) and (18), we notice the following. The commutation relations

$$\sum_i [M_{si}^\alpha M_{ei}^\beta, H_{ex}^{is}] = 0, \quad \sum_i [M_{si}^\alpha M_{si}^\beta, H_{ex}^{is}] \neq 0,$$

$$\sum_{i>j} [M_{si}^\alpha M_{sj}^\beta, H_{ex}^{is}] \neq 0$$

imply that $\exp(-iH_{ex}^{is}t)M_s^\alpha \exp(iH_{ex}^{is}t) = M_s^\alpha$, $[K_H^\alpha, \exp(iH_{ex}^{is}t)] = 0$, and $\exp(-iH_{ex}^{is}t)K_{CF,DM}^\alpha \exp(iH_{ex}^{is}t) = K_{CF,DM}^\alpha(t)$, i.e., the operator H_{ex}^{is} is missing in the exponential multipliers of $M_s^\alpha(t)$ and $K_H^\alpha(t)$ but appears in the exponential multipliers of $K_{CF,DM}^\alpha(t)$. Therefore, if the interaction H_{ex}^{is} dominates H^{doub} and other interactions, then the function $K_{CF,DM}^\alpha(t)$ varies (decreases) rapidly, but the functions $M_s^\alpha(t)$ and $K_H^\alpha(t)$ vary slowly. Hence, the magnetizations $M_s^\alpha(t)$ in Eqs. (23)–(25) decrease slower than the corresponding memory functions; therefore, as pointed out in subsection 1 (Section 3), we can apply a Markov approximation: take $M_s^\alpha(t)$ outside the integral sign and change the upper limit of integration to infinity. In this case, along with the functions $M_s^\alpha(t)$ and $K_H^\alpha(t)$, the function $M_e^\alpha(t)$ also varies slowly; therefore, this approximation is inapplicable to the integrals of Eqs. (19) and (20) (see also (22)). When $H_{ex}^{is} < H^{doub}$, a similar Markov approximation is totally inapplicable, because all variables under the integral sign vary equally rapidly (the corresponding operators do not commute with H^{doub}). Thus, to solve system of equations (17) and (18) rigorously, one should apply the Laplace transform to this system, as was done, for example, in [7]. However, to qualitatively estimate the behavior of a spin system, we apply a Markov approximation to the equations of motion (19)–(25). Then, the nontrivial terms of these equations (see (22)–(25)) are rewritten as

$$\left(\frac{dM_s^\alpha}{dt}\right)_{HB}^{NTD} = -\sqrt{\frac{\pi}{2N_{2H}^\alpha}}K_H^\alpha(0) \times \left[M_s^\alpha(t) - \frac{\gamma_e \mathrm{Sp}(M_s^\alpha)^2}{\gamma_s \mathrm{Sp}(M_e^\alpha)^2}M_e^\alpha(t)\right], \tag{49}$$

$$\left(\frac{dM_s^\alpha}{dt}\right)_{CF} = -\sqrt{\frac{\pi}{2N_{2CF}^\alpha}}K_{CF}^\alpha(0)M_s^\alpha(t), \tag{50}$$

$$\left(\frac{dM_s^\alpha}{dt}\right)_{DM}^{dir} = -\sqrt{\frac{\pi}{2N_{2DM}^\alpha}} \times [K_{DM}^{\alpha\beta dir}(0) + K_{DM}^{\alpha\gamma dir}(0)]M_s^\alpha(t), \tag{51}$$

$$\left(\frac{dM_s^\alpha}{dt}\right)_{DM}^{cr} = -\sqrt{\frac{\pi}{2N_{2DM}^\alpha}} \times [K_{DM}^{\alpha\beta cr}(0)M_s^\beta(t) + K_{DM}^{\alpha\gamma cr}(0)M_s^\gamma(t)]. \tag{52}$$

Thus, taking into account Eqs. (49)–(52), we obtain the following relations for the nontrivial part of Eqs. (17) and (18):

$$\left(\frac{dM_s^\alpha}{dt}\right)^{NTD} = -[(T_{DM}^\alpha)^{-1} + (T_{CF}^\alpha)^{-1} + (T_{sL}^\alpha)^{-1} + (T_{se}^\alpha)^{-1}] \times (M_s^\alpha - M_{s0}^\alpha) + (T_{es}^\alpha)^{-1}(M_e^\alpha - M_{e0}^\alpha), \tag{53}$$

$$\left(\frac{dM_e^\alpha}{dt}\right)^{NTD} = -[(T_{eL}^\alpha)^{-1} + (T_{es}^\alpha)^{-1}] \times (M_e^\alpha - M_{e0}^\alpha) + (T_{se}^\alpha)^{-1}(M_s^\alpha - M_{s0}^\alpha), \tag{54}$$

where $(T_{DM}^\alpha)^{-1}$ and $(T_{CF}^\alpha)^{-1}$ are represented by kinetic coefficients in front of the terms $M_s^\alpha(t)$ on the right-hand sides of Eqs. (50) and (51). The relaxation rates $(T_{se}^\alpha)^{-1}$ and $(T_{es}^\alpha)^{-1}$ can be represented by kinetic coefficients either from Eqs. (49) or from (44) and (45), depending on the choice of the approximation to the memory functions. Recall that $M_{s,e0}^{x,y} = 0$ in Eqs. (53) and (54).

Thus, to study relaxation (relaxation bottleneck) in a system of coupled *s* and *e* spins, we have Eqs. (17) and (18) with the trivial part in the form (21) (these expressions contain terms associated with the interaction of spins with the RF field) and the nontrivial part in the form (53) and (54). Here, we assumed that the *g*-factors of *s* and *e* spins are equal because, to solve this

system of equations, we use the results of [28, 29], where, for simplicity, the authors considered precisely this case when solving phenomenological Hasegawa–Bloch equations by the theory of coupled oscillators. Moreover, it was assumed in those papers that both localized and delocalized spins are equal to 1/2, whereas the structure of manganese ions in manganites with colossal magnetoresistance has a spin of 3/2. However, this difference does not affect the estimated results of the present section. Note also that, in those papers, the role of $(T_{se}^{\alpha})^{-1}$ and $(T_{es}^{\alpha})^{-1}$ is played by the Corring and Overhauser relaxations, respectively. Now, let us consider a solution to Eqs. (53) and (54) together with (21) (a simplified version of Eqs. (17) and (18)) in the cases of strong and weak coupling between localized *s* and quasi-localized *e* spins mentioned above. Everywhere below, we will borrow notation from [28, 29] and assume that $\alpha = x, y$.

5.1. Strong Coupling of s and e Spins (Relaxation Bottleneck)

In this situation, the condition

$$\sigma^{\alpha} = \frac{(T_{se}^{\alpha})^{-1} + (T_{es}^{\alpha})^{-1}}{\left|\delta^{*\alpha} + (T_{sL}^{\alpha})^{-1} - (T_{eL}^{\alpha})^{-1}\right|} \gg 1$$

holds, where σ^{α} is the coupling parameter of the *s* and *e* spin subsystems,

$$\delta^{*\alpha} = \frac{(\delta^{0\alpha})^2}{(T_{se}^{\alpha})^{-1}}, \quad \delta^{0\alpha} = (T_{CF}^{\alpha})^{-1} + (T_{DM}^{\alpha})^{-1}.$$

Note that, in [28, 29], the authors used a dipole–dipole interaction of localized *s* spins as a broadening interaction that makes the main contribution to $\delta^{0\alpha}$. This interaction is insignificant in manganites with colossal magnetoresistance [9]. Therefore, for the normal relaxation of the transverse magnetization components of *s* and *e* spins, i.e., as solutions to Eqs. (53) and (54), respectively, we can use expressions (19) from [29]:

$$\begin{aligned}
\omega''_{(t=+)} &\approx (T_{se}^{\alpha})^{-1} + (T_{es}^{\alpha})^{-1} \\
&+ \frac{\chi_e^{\alpha}(T_{sL}^{\alpha})^{-1} + \chi_s^{\alpha}(T_{eL}^{\alpha})^{-1}}{\chi_e^{\alpha} + \chi_s^{\alpha}} + \frac{\chi_e^{\alpha}}{\chi_e^{\alpha} + \chi_s^{\alpha}}\delta^{*\alpha}, \\
\omega''_{(-t=-)} &\approx \frac{\chi_s^{\alpha}(T_{sL}^{\alpha})^{-1} + \chi_e^{\alpha}(T_{eL}^{\alpha})^{-1}}{\chi_e^{\alpha} + \chi_s^{\alpha}} + \frac{\chi_s^{\alpha}}{\chi_e^{\alpha} + \chi_s^{\alpha}}\delta^{0\alpha},
\end{aligned} \tag{55}$$

where ω''_t and ω''_{-t} are normal relaxation rates (observable EPR linewidths), *t* indicates a mode and takes values "+" or "–"; the *t* mode is always *s*-like, while the (–*t*) mode is *e*-like.

5.2. Weak Coupling between the s and e Spin Subsystems ($\sigma^{\alpha} \ll 1$)

Here, we should use expressions (11) from [29] as a solution to Eqs. (53) and (54):

$$\begin{aligned}
\omega''_{(t=+)} &\approx (T_{se}^{\alpha})^{-1} + (T_{sL}^{\alpha})^{-1} + \delta_s^{0\alpha} \\
&+ \frac{\sigma^{\alpha}}{4}[(T_{se}^{\alpha})^{-1} + (T_{es}^{\alpha})^{-1}],
\end{aligned} \tag{56}$$

$$\begin{aligned}
\omega''_{(-t=-)} &\approx (T_{se}^{\alpha})^{-1} + (T_{eL}^{\alpha})^{-1} \\
&- \frac{\sigma^{\alpha}}{4}[(T_{se}^{\alpha})^{-1} + (T_{es}^{\alpha})^{-1}].
\end{aligned} \tag{57}$$

When deriving Eqs. (56) and (57), we used the condition

$$(T_{es}^{\alpha})^{-1} + (T_{sL}^{\alpha})^{-1} < \delta_s^{0\alpha} + (T_{se}^{\alpha})^{-1} + (T_{sL}^{\alpha})^{-1},$$

which can also be realized in manganites with colossal magnetoresistance.

Note that the situation considered in [28, 29] implies that a secular dipole–dipole interaction between localized *s* spins contributes only to the transverse relaxation (to T_2). In our case, the crystalline-field and Dzialoshinsky–Moriya interactions, and thereby the superexchange interaction, contribute both to T_2 and to the longitudinal relaxation T_1.

A detailed analysis of the expressions that entail formulas (54)–(57) can be found in [28, 29].

Thus, the application of Eqs. (17) and (18) in this section shows that they represent a quantum-statistical analog of Hasegawa–Bloch-type equations; each phenomenological term in the latter equations corresponds to a certain term in Eqs. (17) and (18), which is naturally derived by a method of statistical physics of nonequilibrium processes—the formalism of memory functions.

6. BLOCH EQUATIONS

Now, let us apply Eqs. (17) and (18) to describe a situation when there is a single Lorentz-type EPR line. Such a situation may arise when a manganite with colossal magnetoresistance is in paramagnetic and insulating states.[1]

In this case, as pointed out above, H_{ex}^{is} is the dominant interaction within the model proposed, and the EPR line—a Lorentz-type line associated with localized spins—overlaps the Gaussian-type resonance line

[1] For example, in the phase diagrams for $La_{1-x}Sr_xMnO_3$ and $La_{1-x}Ca_xMnO_3$ in [15, 30], such domains are indicated as O/I, PM; O'/I, PM; and PI. Similar domains of impurity concentrations and temperatures can be found in [10–13].

associated with the spins of both types. Therefore, we can neglect the role of quasi-localized spins and set $\mathbf{M}_e = 0$.

As a result, we obtain Bloch-type equations from (17) and (18) (see Appendix):

$$\frac{d\mathbf{M}_s}{dt} = \gamma_s \mathbf{M}_s \times \mathbf{H}_0 - \mathbf{i}\frac{M_s^x}{T^x} - \mathbf{j}\frac{M_s^y}{T^y} - \mathbf{k}\frac{M_s^z - M_{s0}^z}{T^z} + \mathbf{i}\left(\frac{\alpha M_s^x}{dt}\right)_{DM}^{cr} + \mathbf{j}\left(\frac{\alpha M_s^y}{dt}\right)_{DM}^{cr} + \mathbf{k}\left(\frac{\alpha M_s^z}{dt}\right)_{DM}^{cr}, \quad (58)$$

where

$$(T^{x,y})^{-1} = (T_{CF}^{x,y})^{-1} + (T_{DM}^{x,y})^{-1},$$

$$(T^z)^{-1} = (T_{CF}^z)^{-1} + (T_{DM}^z)^{-1} + (T_{sL}^z)^{-1},$$

$(T_{CF,DM}^{\alpha})^{-1}$ are kinetic (relaxation) coefficients in Eqs. (50) and (51) in front of appropriate components of magnetization $\mathbf{M}_s$; the term $(dM_s^{\alpha}(t)/dt)_{DM}^{cr}$ is given by expressions (25) and (52); and **i**, **j**, and **k** are unit vectors directed along coordinate axes *x*, *y*, and *z*, respectively. This case corresponds to the situation considered in [9, 31] because the magnetization $\mathbf{M}_s$ is a sum of magnetizations of the localized spins of Mn^{3+} and Mn^{4+} ions. Moreover, due to the exchange narrowing [5, 6], the inequality $\mu \gg 1$ holds. Therefore, from (39) we obtain

$$N_{2CF} \approx \frac{M_{4CF}}{M_{2CF}}, \quad N_{2DM} \approx \frac{M_{4DM}}{M_{2DM}}.$$

Thus, using the kinetic (relaxation) coefficients from Eqs. (49)–(52), we can draw the following qualitative conclusions from the aforesaid.

1. Taking into account that $\mu \gg 1$ ($N_2 \approx M_4/M_2$) and $K(0) = M_2$, we easily derive the following expression from the relaxation coefficient in (50):

$$T_2^{-1} = \sqrt{\frac{\pi}{2}\frac{M_2^3}{M_4}} \quad (59)$$

which is in agreement with the results of [9, 32]. If we apply the second and fourth moments from [9], then, according to (59), we obtain

$$(T_{CF}^{x,y})^{-1} = (T_{CF}^{\perp})^{-1} \sim T_2^{-1} \sim D^2/\langle J \rangle,$$

where $\langle J \rangle$ is the superexchange constant for nearest neighbors. The contribution of the Dzialoshinsky–Moriya interaction to the transverse relaxation (59) is of the same order of magnitude.

2. Following [9], we obtain the following expression for the linewidth:

$$\Delta H_{p,p}(\infty) = \frac{2\sqrt{3}}{\gamma_s T_2} = \frac{1}{\gamma_s}\sqrt{\frac{2\pi}{3}\frac{M_2^3}{M_4}}. \quad (60)$$

Naturally, the contributions of the crystalline-field and Dzialoshinsky–Moriya interactions to the linewidths can be estimated in the same way as it was done in the previous subsection of this section. As pointed out above, in this case, the isotropic superexchange interaction (via the contribution to M_4) is responsible for the exchange narrowing of the resonance line.

3. If we neglect anisotropy in this range of temperatures and impurity concentrations, which (as is shown in the Appendix) is quite realistic due to the presence of the dominant contribution H_{ex}^{is}, then we can assume that all second and fourth moments for the directions *x*, *y*, and *z* are equal, i.e., $M_{2,4}^x = M_{2,4}^y = M_{2,4}^z$. Hence, the corresponding relaxation times from Eqs. (58) associated with these moments are also equal (see formulas (50)–(52)): $T_x = T_y = T_z = T_1 = T_2$. This result is in agreement with the experimental results of [10–12].

7. CONCLUSIONS

We have obtained the following original results.

1. Using the formalism of memory functions, we have obtained a system of equations in the general form that describes trivial and nontrivial (relaxational) dynamics for the *x*, *y*, and *z* components of spin magnetization for localized and quasi-localized electrons in a manganite with colossal magnetoresistance under EPR conditions.

2. From the equations of dynamics pointed out above, we have obtained, as particular cases, Hasegawa–Bloch-type equations, Bloch-type equations, and expressions similar to those derived from the Bloch–Vangsness–Redfield theory in which each kinetic coefficient (relaxation rate) in front of a dynamical variable is quantitatively connected with a certain interaction in the spin system of a manganite. The equations obtained reproduce many well-known theoretical results and describe certain experimental data concerning the transverse and longitudinal relaxation rates in manganites with colossal magnetoresistance.

3. The application of the approximation connected with the method of moments to the correlation functions of kinetic coefficients allows one to determine, by the shape of EPR line, which of the spin–spin interactions is dominant and to which of the two types of spins the EPR line corresponds. On the other hand, the application of the approximation of nearest neighbors to the above correlation functions yields the well-known temperature dependence of the EPR linewidth—the Huber law—which also makes it possible to identify EPR

lines with a certain type of spins in the spin system of a manganite.

4. A tensorial character of relaxation in a system of manganites with colossal magnetoresistance has been revealed for the first time. Expressions have been obtained for the so-called cross relaxation rates, and the possibility of experimental observation of their effect on the relaxation and the EPR linewidth has been pointed out.

Finally, note that a more detailed investigation of the spin system of manganites with colossal magnetoresistance and the specification of the qualitative results obtained in this paper, as well as obtaining new results (for example, the dependence of the shape of the EPR line on the direction of external magnetic fields, etc.), require the use of information on the structure of a specific manganite with colossal magnetoresistance and on the experimental conditions. The theoretical approach developed makes it possible to investigate the dynamics of the spin system of a manganite with colossal magnetoresistance under conditions of ferromagnetic resonance (in the range of temperatures and impurity concentrations where the material is in a ferromagnetic state).

ACKNOWLEDGEMENTS

We are grateful to L.L. Buishvili† for his attention. We also thank N.P. Fokina for drawing our attention to the subject and for useful remarks and recommendations, as well as to V.A. Atsarkin and V.V. Demidov for providing the manuscript of [11] before publication and for stimulating remarks.

APPENDIX

Here, we demonstrate the application of, say, Eqs. (58). Let us introduce a system of coordinates that rotates with angular frequency ω, restore the terms of trivial dynamics associated with the RF field in Eqs. (58), and, for simplicity, restrict ourselves to the case, considered in Section 5, when the cross relaxation terms vanish. Then, Eqs. (58) yield the classical Bloch equations [32] with T^x and T^y in place of T_2 and T^z in place of T_1:

$$\frac{dM_s^x}{dt} = \Delta M_s^y - \frac{M_s^y}{T^x}, \tag{A.1}$$

$$\frac{dM_s^y}{dt} = -\Delta M_s^x - \frac{M_s^y}{T^y} - \omega_1 M_s^z, \tag{A.2}$$

$$\frac{dM_s^z}{dt} = -\frac{M_s^z - M_{s0}^z}{T^z} + \omega_1 M_s^y. \tag{A.3}$$

† Deceased.

Here, $\Delta = \omega - \omega_s$, ω_s is the Zeeman frequency of localized spins, ω_1 is the amplitude of the RF field,

$$(T^x)^{-1} = (T_{CF}^x)^{-1} + (T_{DM}^x)^{-1} = \sqrt{\frac{\pi}{2}}\left[\frac{K_{CF}^x(0)}{\sqrt{N_{2CF}^x}} + \frac{K_{DM}^{yx\mathrm{dir}}(0) + K_{DM}^{xz\mathrm{dir}}(0)}{\sqrt{N_{2DM}^x}}\right], \tag{A.4}$$

$$(T^y)^{-1} = (T_{DM}^y)^{-1} = \sqrt{\frac{\pi}{2N_{2DM}^y}}[K_{DM}^{yz\mathrm{dir}}(0) + K_{DM}^{yx\mathrm{dir}}(0)], \tag{A.5}$$

$$(T^z)^{-1} = (T_{DM}^z)^{-1} = \sqrt{\frac{\pi}{2N_{2DM}^z}}[K_{DM}^{zx\mathrm{dir}}(0) + K_{DM}^{zy\mathrm{dir}}(0)] \tag{A.6}$$

(see (50) and (51)). Taking into account formulas (49), the equality $K(0) = M_2$, and the fact that $\mu \gg 1$ under exchange narrowing, we can easily derive expressions similar to (59) from (A.4)–(A.6). Choosing the constants of crystalline-field and Dzialoshinsky–Moriya interactions and calculating M_2 and M_4, we can obtain relaxation rates for a specific material with colossal magnetoresistance.

Note that, in contrast to the assumption about total isotropy in the general remarks in Section 4, here T^x is different from T^y and T^z. However, this difference is insignificant and cannot spoil the general picture due to the presence of the constant of superexchange interaction H_{ex}^{is} in the expression for M_4.

Setting to zero the right-hand sides of expressions (A.1)–(A.3), we can easily obtain steady-state values of M_s^α and, hence, expressions for the shape of resonant absorption and dispersion. In particular, for a steady-state value of M_s^y, we have

$$|M_s^y| = M_{s0}^z \omega_1 T^y (1 + T^y T^x \Delta^2 + T^y T^z \omega_1^2)^{-1}, \tag{A.7}$$

which is exactly the same as in [32], except that T^x and T^y are replaced by T_2 and T^z is replaced by T_1 as above. By analogy with [32], we can conclude from (A.3) that the line of resonant absorption in the absence of appreciable saturation ($\omega_1^2 T^y T^z \ll 1$) has a Lorentzian shape with the half-width at half maximum equal to $(T^y)^{-1}$ from (A.4). Moreover, closer to the line center, where $\Delta \approx 0$, the linewidth is determined solely by $(T^y)^{-1}$, i.e., by expression (60), in which the moments M_2 and M_4 are associated with the Dzialoshinsky–Moriya interaction without any contribution from the crystalline-field interaction; as pointed out above, this is inessential due to the contribution of the interaction H_{ex}^{is} to M_4.

Finally, note that, substituting the second and fourth moments, say, from [9], to the expressions for T^x, T^y, and T^z by using (A5)–(A7), one can obtain a graphic picture of the shape of the resonant absorption line [8].

REFERENCES

1. D. N. Zubarev, in *Modern Problems of Mathematics* (VINITI, Moscow, 1980), Vol. 15 [in Russian].
2. G. Röpke, *Statistische Mechanik für das Nichtgleichgewicht* (Wissenschaften, Berlin, 1986; Mir, Moscow, 1990).
3. R. Balescu, *Equilibrium and Nonequilibrium Statistical Mechanics* (Wiley, Chichester, 1975; Mir, Moscow, 1978), Vol. 2.
4. F. Lado, J. D. Memory, and J. W. Parker, Phys. Rev. B **4**, 1406 (1971).
5. A. Abragam and M. Goldman, *Nuclear Magnetism: Order and Disorder* (Clarendon Press, Oxford, 1982; Mir, Moscow, 1984).
6. M. Mehring, *High Resolution NMR Spectroscopy in Solids* (Springer, Berlin, 1976; Mir, Moscow, 1980).
7. É. Kh. Khalvashi, Zh. Éksp. Teor. Fiz. **110**, 703 (1996) [JETP **83**, 379 (1996)].
8. É. Kh. Khalvashi and M. V. Chkhartishvili, Fiz. Tverd. Tela (St. Petersburg) **40**, 1036 (1998) [Phys. Solid State **40**, 946 (1998)].
9. D. L. Huber, G. Alejandro, A. Caneiro, *et al.*, Phys. Rev. B **60**, 12155 (1999).
10. V. A. Atsarkin, V. V. Demidov, G. A. Vasneva, *et al.*, Phys. Rev. B **63**, 092405 (2001).
11. F. Simon, V. A. Atsarkin, V. V. Demidov, *et al.*, Phys. Rev. B **67**, 224433 (2003).
12. V. A. Atsarkin, V. V. Demidov, G. A. Vasneva, *et al.*, Appl. Magn. Reson. **21**, 147 (2001).
13. O. Chauvet, G. Goglio, P. Molinie, *et al.*, Phys. Rev. Lett. **81**, 1102 (1998).
14. M. T. Causa, M. Tovar, A. Caneiero, *et al.*, Phys. Rev. B **58**, 3233 (1998).
15. V. A. Ivanshin, J. Deisenhofer, H.-A. Krug von Nidda, *et al.*, Phys. Rev. B **61**, 6213 (2000).
16. A. Shengelaya, Guo-meng Zhao, H. Keller, *et al.*, Phys. Rev. Lett. **77**, 5296 (1996).
17. A. Shengelaya, Guo-meng Zhao, H. Keller, *et al.*, Phys. Rev. B **61**, 5888 (2000).
18. L. L. Buishvili and É. Kh. Khalvashi, *Radiospectroscopy* (Perm. Gos. Univ., Perm, 1987), p. 58 [in Russian].
19. C. Zener, Phys. Rev. **81**, 440 (1951); Phys. Rev. **82**, 403 (1951).
20. I. Yamada, H. Fujii, and M. Hidaka, J. Phys.: Condens. Matter **1**, 3397 (1989).
21. B. N. Provotorov, Zh. Éksp. Teor. Fiz. **41**, 1582 (1961) [Sov. Phys. JETP **14**, 1126 (1962)].
22. H. Hasegawa, Prog. Theor. Phys. **23**, 483 (1959).
23. H. Hasegawa and A. M. Stewart, Prog. Theor. Phys. **74**, 943 (1985).
24. A. G. Gurevich, *Magnetic Resonance in Ferrites and Antiferromagnets* (Nauka, Moscow, 1973) [in Russian].
25. C. P. Slichter, *Principles of Magnetic Resonance*, 2nd ed. (Springer, Berlin, 1980; Mir, Moscow, 1967).
26. S. E. Barnes, Adv. Phys. **30**, 801 (1981).
27. N. P. Fokina and K. O. Khutsishvili, Appl. Magn. Reson. **17**, 503 (1999).
28. N. P. Fokina and K. O. Khutsishvili, Zh. Éksp. Teor. Fiz. **123**, 98 (2003) [JETP **96**, 83 (2003)].
29. N. P. Fokina, M. O. Élizbarashvili, V. A. Atsarkin, *et al.*, Fiz. Tverd. Tela (St. Petersburg) **45**, 1921 (2003) [Phys. Solid State **45**, 2017 (2003)].
30. V. Cataudella, G. De Filippis, and G. Ladonisi, Phys. Rev. B **63**, 052406 (2001).
31. D. L. Huber, Phys. Rev. B **12**, 31 (1975).
32. A. Abragam, *The Principles of Nuclear Magnetism* (Clarendon Press, Oxford, 1961; Inostrannaya Literatura, Moscow, 1963), pp. 435–440.

Translated by I. Nikitin

Journal of Experimental and Theoretical Physics, Vol. 100, No. 2, 2005, pp. 410–421.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 127, No. 2, 2005, pp. 458–470.

STATISTICAL, NONLINEAR, AND SOFT MATTER PHYSICS

Phase Transition in a Plasma with Anisotropic Electron Distribution in an External Magnetic Field

L. I. Men'shikov

Russian Research Center Kurchatov Institute, Moscow, 123182 Russia
e-mail: ieronimovich@mtu-net.ru

Received June 28, 2004

Abstract—A phase transition is discussed that can occur in a plasma with substantially different transverse and longitudinal temperatures of electrons moving in a magnetic field, $T_\perp/T_\parallel \gg 1$. The Debye cloud surrounding an ion sharply contracts as $T_\parallel$ decreases or the magnetic field increases. The effect of larger radiative electron–ion recombination cross sections compared with their theoretical values is explained; this effect is observed in storage rings with electron cooling systems (coolers). The role played by the phase transition in the crystallization of ion beams is discussed. © *2005 Pleiades Publishing, Inc.*

1. INTRODUCTION

Because of kinematic cooling of the accelerated beam, the distribution of electrons in a cooler is strongly compressed in the longitudinal direction [1–4]:

$$f(v) = G(v_\perp)g(v_\parallel), \quad \int f(v)d^3v = 1, \tag{1}$$

$$G(v_\perp) = \frac{1}{2\pi\Delta_\perp^2}\exp\left(-\frac{v_\perp^2}{2\Delta_\perp^2}\right),$$
$$g(v_\parallel) = (\Delta_\parallel\sqrt{2\pi})^{-1}\exp\left(-\frac{v_\parallel^2}{2\Delta_\parallel^2}\right), \tag{2}$$

where

$$\Delta_\perp = \sqrt{T_\perp/m}, \quad \Delta_\parallel = \sqrt{T_\parallel/m},$$

and $v_\perp$ and $v_\parallel$ are the transverse and longitudinal electron velocities with respect to the magnetic field direction, respectively. Two electrons emitted from a cathode heated to $T_0 \approx 0.1$ eV typically have initial longitudinal velocities of $v_0 \approx \sqrt{T_0/m}$, which differ by $\Delta v_0 \sim v_0$. The difference in their energies does not change when the electrons are accelerated. Therefore, after gaining a speed v and an energy $E = mv^2/2$, they have longitudinal velocities differing by

$$\Delta v_\parallel \sim v_0\Delta v_0/v.$$

Taking into account electron–electron interaction, we conclude that

$$T_\perp \sim T_0, \quad T_\parallel \sim T_1 + \bar{U} \tag{3}$$

in an accelerated electron beam [2–4]. Here, $T_\perp$ and $T_\parallel$ are the transverse and longitudinal temperatures of electrons moving in the magnetic field, respectively; $T_1 = T_0^2/E$; and $\bar{U} = e^2/\bar{R}$ is the mean energy of Coulomb repulsion between electrons in the beam ($\bar{R} = n^{-1/3}$ is the mean distance between electrons when their concentration is n). As an example, consider the experiments on fully stripped uranium ions U^{+Z_0} ($Z_0 = 92$) reported in [5], with

$$n \approx 10^6\ \text{cm}^{-3}, \quad T_\perp \approx 0.12\ \text{eV}, \quad E \approx 160\ \text{keV},$$
$$T_1 \approx 10^{-3}\ \text{K}, \quad \bar{U} \approx 0.15\ \text{K}, \quad H \approx 1000\ \text{G}. \tag{4}$$

Under these conditions,

$$T_1 \ll \bar{U}, \quad T_\parallel \sim \bar{U} \approx 0.15\ \text{K}, \quad T_\perp \approx 1300\ \text{K};$$

i.e.,

$$T_\perp/T_\parallel \gg 1. \tag{5}$$

In [5], the electron plasma was ideal with respect to transverse motion ($\xi_\perp = T_\perp/\bar{U} \sim 10^4$) and nonideal with respect to longitudinal motion,

$$\xi_\parallel = T_\parallel/\bar{U} \sim 1. \tag{6}$$

According to a theory taking into account both collective and collisional processes [6–11], the electron–

electron equilibration time exceeds 1 μs for all coolers currently in use. This time scale corresponds to a distance of more than 200 m traveled by an electron in the laboratory frame. Since the length of a cooler is $L_c \approx 3$ m, distribution function (1) can be treated as invariant.

2. LINEAR SCREENING MODEL

The Fourier component of the charge density in the Debye cloud surrounding an ion that is immobile in the laboratory frame is [12–14]

$$\Delta\rho(\mathbf{k}) = eZ_0\left(\frac{1}{\varepsilon(\mathbf{k}, \omega = 0)} - 1\right), \tag{7}$$

where

$$\varepsilon(\mathbf{k}, \omega) = 1 - \frac{\omega_p^2}{k^2} m \sum \int d^3v J_s^2(k_\perp r_H) \times f(\mathbf{v})\left(\frac{k_\parallel v_\parallel}{T_\parallel} + \frac{s\omega_H}{T_\perp}\right)(\omega + i\cdot 0 - s\omega_H - k_\parallel v_\parallel)^{-1} \tag{8}$$

is the longitudinal part of the plasma dielectric tensor $\varepsilon_{\alpha\beta}(\mathbf{k}, \omega)$,

$$\varepsilon(\mathbf{k}, \omega) = \varepsilon_{\alpha\beta}(\mathbf{k}, \omega) q_\alpha q_\beta;$$

$$\omega_p = \sqrt{\frac{4\pi n e^2}{m}}$$

is the plasma frequency;

$$\omega_H = \frac{eH}{mc}$$

is the Larmor frequency;

$$r_H = v_\perp/\omega_H$$

is the Larmor radius; and J_s is a Bessel function, $q_\alpha = k_\alpha/k$. The sum in (8) runs over all integer s in the interval $(-\infty, \infty)$. It can be replaced by an integral as $H \longrightarrow 0$. (However, it is simpler to use the expression for ε at $H = 0$ from the outset; e.g., see [15].) This gives

$$\Delta\rho(\mathbf{k}) = -\frac{eZ_0}{1 + k_\parallel^2 R_\parallel^2 + k_\perp^2 R_\perp^2}, \tag{9}$$

$$\Delta\rho(\mathbf{r}) = -\frac{eZ_0}{4\pi R_\parallel R_\perp^2}\frac{e^{-L}}{L}, \tag{10}$$

where

$$L = \sqrt{(z/R_\parallel)^2 + (r_\perp/R_\perp)^2},$$

$$R_\parallel = \sqrt{\frac{T_\parallel}{4\pi n e^2}}, \quad R_\perp = \sqrt{\frac{T_\perp}{4\pi n e^2}},$$

and the z axis is directed along the magnetic field **H**. It follows that the Debye cloud at $H = 0$ is an oblate spheroid with the semiaxes

$$l_\parallel \sim R_\parallel \ll l_\perp \sim R_\perp.$$

Next, consider the case of a strong magnetic field (subject to a condition given below), assuming that the Debye length is small,

$$r_D \ll \bar{r}_H. \tag{11}$$

This assumption is confirmed by further calculations and substantiated by physical considerations. As

$$k_\perp \sim |k_\parallel| \sim \frac{1}{r_D} \ll \frac{1}{\bar{r}_H}, \tag{12}$$

the Bessel function can be approximated by the asymptotic expression

$$J_s^2(y) \approx \frac{\theta(y - s)}{\pi\sqrt{y^2 - s^2}}, \tag{13}$$

which is obtained by replacing a squared cosine with mean value 1/2. Since the integrand in (8) is invariant under the inversion

$$s \longrightarrow -s, \quad v_\parallel \longrightarrow -v_\parallel,$$

the singular factor can be replaced by its principal value

$$\frac{1}{i\cdot 0 - s\omega_H - k_\parallel v_\parallel} \longrightarrow P\left(-\frac{1}{s\omega_H + k_\parallel v_\parallel}\right). \tag{14}$$

Equations (12)–(14) lead us to conclude that, if

$$\beta = \frac{|k_\parallel v_\parallel|}{\omega_H} \sim \frac{\Delta_\parallel}{r_D \omega_H} \ll 1 \tag{15}$$

then (8) is dominated by the term with $s = 0$:

$$\varepsilon(\mathbf{k}, \omega) \approx 1 + \frac{q_0^3}{k^2 k_\perp}, \tag{16}$$

$$q_0 = (\sqrt{2\pi} R_\parallel^2 \bar{r}_H)^{-1/3}, \tag{17}$$

$$\Delta\rho(\mathbf{k}) = -\frac{eZ_0}{1 + k^2 k_\perp/q_0^3}, \tag{18}$$

$$\Delta\rho(\mathbf{r}) = -eZ_0 q_0^3 (4\pi)^{-1} \int_0^\infty dk_\perp J_0(k_\perp r_\perp) e(-k|z|), \tag{19}$$

where

$$k = \sqrt{k_\perp^2 + q_0^3/k_\perp}.$$

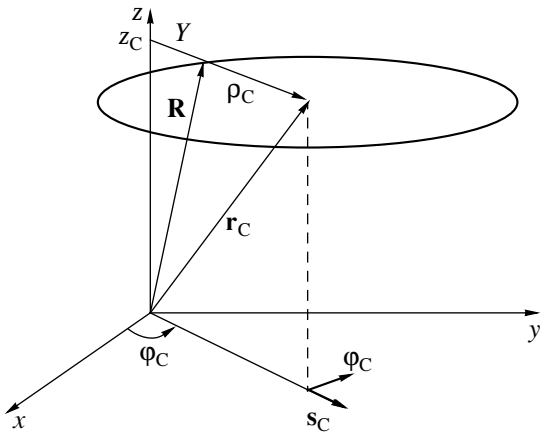

Fig. 1. Electron ring coordinates relative to the ion located at the origin: $\mathbf{r}_C$ is the radius vector of the ring center; $\boldsymbol{\varphi}_C$ and $\mathbf{s}_C$ are the unit vectors, $\mathbf{s}_C \parallel \boldsymbol{\rho}_C$, $\boldsymbol{\varphi}_C \perp \mathbf{s}_C$.

Expression (19) can be used to determine the longitudinal and transverse Debye lengths:

$$l_\perp \sim l_\parallel \sim r_D = \frac{1}{q_0} \sim (R_\parallel^2 \bar{r}_H)^{1/3}. \quad (20)$$

It follows from (20) that (15) is equivalent to the condition

$$\bar{r}_H \ll R_\parallel \sqrt{\frac{\Delta_\perp}{\Delta_\parallel}}. \quad (21)$$

However, if

$$\bar{r}_H \gg R_\perp \sqrt{\frac{\Delta_\perp}{\Delta_\parallel}}, \quad (22)$$

then the sum over s in (8) can be replaced by an integral, which corresponds to the limit $\omega_H \longrightarrow 0$ ($H \longrightarrow 0$). It is therefore clear that (9) and (10) are valid under condition (22), rather than the seemingly obvious condition $\bar{r}_H \gg R_\perp$, which pertains to case (5).

It follows from (20) that the initial assumption [Eq. (11)] holds under the condition

$$R_\parallel \ll \bar{r}_H, \quad (23)$$

combined with (21). The latter inequality sets a limit for magnetic field magnitudes at which (20) is valid. We always have then

$$R_\parallel \ll r_D \ll \bar{r}_H. \quad (24)$$

Now, consider the case of an extremely strong magnetic field defined by the condition

$$\bar{r}_H \ll R_\parallel, \quad (25)$$

also called the case of magnetized electron gas [2, 3]. In this limit, we can set the Larmor radius to zero, i.e., consider electrons that move only along magnetic field lines like small beads threaded on stretched strings. Their motion is characterized by the temperature $T_\parallel$. If (25) holds, then

$$l_\perp = l_\parallel = r_D = R_\parallel, \quad (26)$$

and we have the spherically symmetric Debye cloud described by (10) with $R_\perp$ replaced by $R_\parallel$. Indeed, by virtue of (25), estimate (12) can be rewritten as

$$k_\perp \sim |k_\parallel| \sim \frac{1}{r_D} \ll \frac{1}{\bar{r}_H},$$

and we therefore have $k_\perp r_H \ll 1$ in (8). This allows us to discard the terms with $s \neq 0$ and assume that

$$J_0(k_\perp r_H) \approx 1.$$

We finally obtain

$$\varepsilon(\mathbf{k}, 0) \approx 1 + \frac{1}{k^2 R_\parallel^2}, \quad \Delta\rho(\mathbf{k}) = -\frac{eZ_0}{1 + k^2 R_\parallel^2}, \quad (27)$$

which substantiates the above comments concerning (26).

Relation (25) is the condition of electron-gas magnetization with respect to collective interaction. It must be distinguished from the magnetization condition with respect to pair interactions, $\bar{r}_H \ll \bar{R}$, when an electron colliding with an ion behaves as a bead that only moves along a magnetic field line.

3. ELECTRON-RING SCREENING MODEL

Both starting equations (7) and (8) and all results presented in the preceding section are obtained by using the Maxwell equations and the linearized Vlasov equation [12–15]. However, certain phenomena observed in cooler plasmas require an analysis going beyond the scope of a linear approximation.

It is clear from (5) that averaging can be performed over the fast Larmor rotation of electrons. Every electron is then represented by a ring of radius r_H with the linear charge density $e/2\pi r_H$. The potential energy of the ring in the field of an ion is (see Fig. 1)

$$U(\mathbf{r}_C) \approx \begin{cases} -\dfrac{Z_0 e^2}{2\pi r_H} \ln\left(\dfrac{r_H}{R}\right), & R \ll r_H \\ -\dfrac{Z_0 e^2}{r_C}, & r_C \gg r_H, \end{cases} \quad (28)$$

where R is the distance from the ion to the nearest point on the ring. Expression (28) for $R \ll r_H$ describes the interaction of the ion with a rectilinear charged thread.

The ring executes a one-dimensional motion along the z axis characterized by the temperature $T_\parallel$. The concentration of the ring centers is described by the Boltzmann distribution

$$n_C(\mathbf{r}_C) = n\exp\left(-\frac{U(\mathbf{r}_C)}{T_\parallel}\right). \quad (29)$$

If $R \ll r_H$, then we have

$$n_C(\mathbf{r}_C) = n\left(\frac{r_H}{R}\right)^\gamma, \quad (30)$$

where

$$\gamma = \frac{Z_0 e^2}{\pi r_H T_\parallel}. \quad (31)$$

The combination of an exponential and a logarithm in (29) and (28), respectively, and the power law in (30) suggest the possibility of a phase transition that is mathematically (but not physically) analogous to the Berezinskii–Kosterlitz–Thouless transition [16, 17]. Prior to considering this effect, we must substantiate the electron-ring model. This can be done by showing that, in the linear approximation, when

$$n_C(\mathbf{r}_C) \approx n + \Delta n_C(\mathbf{r}_C),$$
$$\Delta n_C(\mathbf{r}_C) = -\frac{nU(\mathbf{r}_C)}{T_\parallel}, \quad (32)$$

the ring model yields the results of the linear theory developed in the preceding section. Qualitatively, this is clear from the following estimates. The rings are grouped around an ion: their centers are separated from the ion by a distance of about r_H; the nearest ring boundaries, by distances $R \ll r_H$. Thus, the ion is surrounded by a Debye cloud of size $r_D \ll r_H$. According to (32) and (28),

$$\Delta n_C \sim \frac{Z_0 e^2 n}{r_H T_\parallel}.$$

The fraction of the ring charge contained in the Debye sphere is estimated as r_D/r_H. The ring centers are grouped in the annulus between coaxial cylinders of radii,

$$\rho_1 \approx r_H, \quad \rho_2 \approx r_H + r_D.$$

Its thickness in the z direction is estimated as r_D, and its volume is

$$V_C \sim r_H r_D^2.$$

The number of excess electrons in the Debye cloud is

$$\Delta N_e \sim V_C \Delta n_C \left(\frac{r_D}{r_H}\right). \quad (33)$$

Substituting this into the screening condition

$$\Delta N_e = Z_0 \quad (34)$$

yields (20).

When a ring is translated, the shift vectors of all its points are equal. The ring charge density at an arbitrary point $\mathbf{r}$ therefore depends only on $\mathbf{r} - \mathbf{r}_C$:

$$\rho_0(\mathbf{r}, \mathbf{r}_C) = \rho_0(\mathbf{r} - \mathbf{r}_C).$$

The excess charge density created by all rings at a point r is

$$\Delta\rho(\mathbf{r}) = \int \Delta n_C(\mathbf{r}_C)\rho_0(\mathbf{r} - \mathbf{r}_C)d^3 r_C. \quad (35)$$

Here, $\Delta n_C(\mathbf{r}_C)$ is given by (32), and the potential energy of the ring is

$$U(\mathbf{r}_C) = \int \phi(\mathbf{r})\rho_0(\mathbf{r} - \mathbf{r}_C)d^3 r, \quad (36)$$

where $\phi(\mathbf{r})$ is the total electric potential at point r. Using the Poisson equation

$$\Delta\phi(\mathbf{r}) = -4\pi Z_0 e\delta(\mathbf{r}) - 4\pi\Delta\rho(\mathbf{r}),$$

we obtain

$$k^2\phi(\mathbf{k}) = 4\pi Z_0 e + 4\pi\Delta\rho(\mathbf{k}). \quad (37)$$

It follows from (35), (36), and (32) that

$$\Delta\rho(\mathbf{k}) = \Delta n_C(\mathbf{k})\rho_0(\mathbf{k})$$
$$= -\frac{n\rho_0(\mathbf{k})U(\mathbf{k})}{T_\parallel} = -\frac{n|\rho_0(\mathbf{k})|^2\phi(\mathbf{k})}{T_\parallel}. \quad (38)$$

For a ring in the xy plane (see Fig. 2), we have

$$\rho_0(\mathbf{r}) = -e\int_0^{2\pi}\frac{d\phi}{2\pi}\delta(\mathbf{r} - r_H\mathbf{s}),$$

where $\mathbf{s} = \boldsymbol{\rho}/\rho$. Therefore, the form factor of the ring

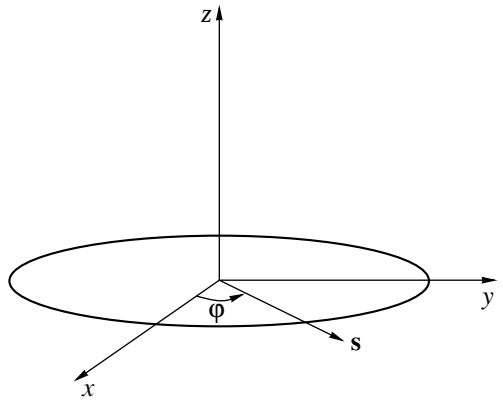

Fig. 2. Explanation of form factor calculations for the electron ring charge.

charge is

$$\rho_0(\mathbf{k}) = \int d^3r \exp(-i\mathbf{k}\cdot\mathbf{r})\rho_0(\mathbf{r}) = -e\int_0^{2\pi}\frac{d\phi}{2\pi}\exp(-i\mathbf{k}\cdot\mathbf{s}r_H) = -eJ_0(k_\perp r_H). \quad (39)$$

Averaging over distribution (1) and oscillations of a squared cosine for $k_\perp r_H \gg 1$ (compare Section 2), we obtain (18) from (39), (38), and (37). Thus, the electron-ring model is applicable when the Larmor radius satisfies (23) and (21). (Under condition (25) of electron-gas magnetization, the applicability of the ring model follows from the discussion at the end of Section 2.) The physical reason for the validity of the ring model is clear from (5): according to (28), the characteristic ion–ring interaction energy is low as compared to the kinetic energy of an electron:

$$T_\perp \gg |U| \sim Z_0 e^2/r_H.$$

4. PHASE TRANSITION

Recall a simple method for estimating the Debye length r_D in an equilibrium plasma at temperature T. A linearized Boltzmann distribution yields an estimate for the excess electron density:

$$\Delta n \approx n(e\phi/T).$$

Inside the Debye cloud,

$$\phi \sim \phi_0 = \frac{Z_0 e}{r} \sim \frac{Z_0 e}{r_D}, \quad (40)$$

at $r \sim r_D$. By analogy with (33), we have

$$\Delta N_e \sim \Delta n r_D^3 \sim Z_0 n e^2 r_D^2/T.$$

Substituting this into (34), we obtain the usual formula

$$r_D \sim \sqrt{T/ne^2}.$$

According to (40), we can use the energy of unscreened ion–electron interaction instead of the true potential energy $U = -e\phi$ of an electron to estimate the Debye length. Similarly, we can use (30) in the ring model to elucidate the nature of physical phenomena outside the scope of the linear approximation considered in the preceding sections. For simplicity, let us first assume that the Larmor radii are equal for all electrons:

$$r_H = \text{const}. \quad (41)$$

The electron charge density in the vicinity of an ion is (see Fig. 1)

$$\rho(\mathbf{r}) \sim n\int d^3r_C\left(\frac{r_H}{R}\right)^\gamma \rho_0(\mathbf{r}-\mathbf{r}_C). \quad (42)$$

For

$$r \ll r_H, \quad (43)$$

the rim of the ring can be approximated by a line. Accordingly,

$$\rho_0(\mathbf{r}-\mathbf{r}_C) \approx -e(2\pi r_H)^{-1}\delta(z-z_C)\delta[(\mathbf{r}-\mathbf{R})\cdot\mathbf{s}_C], \quad (44)$$

where $\mathbf{s}_C = \boldsymbol{\rho}_C/\rho_C$. Under condition (43), we can write

$$\rho_C \approx r_H,$$

$$d^3r_C \approx dz_C\rho_C d\rho_C d\phi_C \approx r_H dz_C dY d\phi_C,$$

where ρ_C is replaced by the new coordinate $Y = \rho_C - r_H$ with values in the interval $(-\infty, \infty)$. By virtue of axial symmetry,

$$\rho(\mathbf{r}) = \rho(r_\perp, z), \quad r_\perp = \sqrt{x^2+y^2}.$$

The integral in (42) can therefore be performed by assuming that $\mathbf{r} = (r_\perp, 0, z)$ (i.e., for $\mathbf{r}$ lying in the xz plane):

$$\rho(r_\perp, z) \approx -en\int_0^{2\pi}\frac{d\phi_C}{2\pi}\int_{-\infty}^{\infty}dY\left[\frac{r_H}{\sqrt{Y^2+z^2}}\right]^\gamma \times \delta(r_\perp\cos\phi_C - Y) = -en\int_0^{2\pi}\frac{d\phi_C}{2\pi}\left[\frac{r_H^2}{r_\perp^2\cos^2\phi_C + z^2}\right]^{\gamma/2}. \quad (45)$$

This yields the limit expressions

$$\rho(r_\perp, z) \approx \begin{cases} -\dfrac{en}{r_\perp}r_H^\gamma|z|^{-\gamma+1}, & |z| \ll r_\perp \\ -\dfrac{en}{|z|}r_H^\gamma|z|^{-\gamma+1}, & |z| \gg r_\perp, \end{cases} \quad (46)$$

which can be unified into the following formula by interpolation:

$$\rho(r_\perp, z) \approx -\frac{en}{r} r_H^\gamma (z^2 + z_1^2)^{(-\gamma+1)/2}. \quad (47)$$

Here, z_1 is the value of R at which the electron-ring approximation fails. For an equilibrium Debye cloud,

$$|z_1| \approx R_T = \frac{Z_0 e^2}{T_\perp}. \quad (48)$$

(It is shown below that Debye clouds in coolers cannot reach complete equilibrium, and the actual longitudinal cloud size z_1 is therefore much larger than R_T.) The charge of the Debye cloud can be calculated as

$$Q_D \approx \int d^3 r \rho(\mathbf{r}) \approx -en \int r_\perp dr_\perp dz r^{-1} r_H^\gamma (z^2 + z_1^2)^{(-\gamma+1)/2} \approx enr_H^\gamma l_\perp J, \quad (49)$$

$$J = \int_0^\infty dz (z^2 + z_1^2)^{(-\gamma+1)/2}. \quad (50)$$

Here, $l_\perp$ is the cross size of the cloud determined by the equation given below.

It follows from (49) and (50) that there is a critical γ value,

$$\gamma_C = 2. \quad (51)$$

If

$$\gamma < \gamma_C \quad (52)$$

then the major contribution to the integral J is due to large $|z|$. Physically, this means that the linear theory developed in Section 2 is applicable when (52) is satisfied. If

$$\gamma > \gamma_C \quad (53)$$

then the major contribution to J is due to small $|z|$,

$$|z| \sim z_1. \quad (54)$$

It follows that the Debye cloud contracts when the temperature $T_\parallel$ decreases to

$$T_C = \frac{Z_0 e^2}{2\pi r_H} \quad (55)$$

(see table). The Debye cloud also contracts at $T_\parallel$ = const, when the magnetic field increases to H_C and the Debye cloud size at various temperatures $T_\parallel$ in intermediate [conditions (21) and (23)] and high [condition (25)] magnetic fields

Radius r_H	Cloud size denotation	$T_\parallel > T_C$	$T_\parallel < T_C$
(21), (23)	$l_\perp$	$(R_\parallel^2 r_H)^{1/3}$	$(R_\parallel^2 r_H)^{1/3}$
	$l_\parallel$	$(R_\parallel^2 r_H)^{1/3}$	R_T
(25)	$l_\perp$	$R_\parallel$	r_H
	$l_\parallel$	$R_\parallel$	R_T

Larmor radius therefore decreases to

$$r_{cr} = Z_0 e^2 / 2\pi T_\parallel.$$

Since the contraction of the Debye cloud changes the correlation energy of electrons (e.g., see [18]), it can be interpreted as a first-order phase transition analogous to crystallization of a liquid. However, this conclusion is valid only if (41) holds. In reality, r_H varies from one electron to another, and the phase transition spans a temperature interval $\Delta t_\parallel \sim T_C$, i.e., is rather similar to liquid solidification into an amorphous state. Another reason is explained as follows. The value of γ decreases to $\gamma_C = 2$ as $T_\parallel$ approaches T_C. As $T_\parallel$ decreases below T_C, the number of electrons in the cloud surrounding the ion increases. As a result, the total charge Z of the ion and cloud decreases, but the condition

$$\gamma = \frac{Ze^2}{\pi r_H T_\parallel} = 2$$

holds. Therefore, the decrease in total charge follows the law

$$Z = Z_0 \left(\frac{T_\parallel}{T_C} \right).$$

Since a charge is screened in a plasma, the total charge of the Debye cloud is equal in magnitude and opposite in sign to the ion charge:

$$Q_D = -Z_0 e. \quad (56)$$

In an equilibrium ideal plasma at a temperature T, the number of electrons contained in the Debye sphere is

$$N_D \approx n r_D^3 \approx \xi^{3/2},$$

where the parameter

$$\xi = T/\bar{U} = T\bar{R}/e^2 \gg 1$$

characterizes the degree of ideality of the plasma. The small excess of electrons created in the Debye sphere by the field of the ion makes the total charge of the cloud equal to (56). For the nonequilibrium anisotropic plasma considered here (see (5)), in the domain of linear behavior at $T_\parallel > T_C$, this follows from (18) and (27) and the relation $Q_D = \Delta\rho(\mathbf{k} = 0)$. At $T_\parallel < T_C$, the Debye cloud consists of two parts: a bunch near the ion with small longitudinal size (54) and charge

$$-(Z_0 - Z)e = -eZ_0\left(1 - \frac{T_\parallel}{T_C}\right)$$

and an outer part with size r_H and charge

$$-eZ_0(T_\parallel/T_C).$$

When conditions (21) and (23) hold, the number of electrons in the Debye cloud can be estimated as the number of ring centers in the volume $V_C \approx r_H r_D^2$ introduced in deriving (33):

$$N_D \approx V_C n \approx \left(\frac{r_H}{r_D}\right)\left(\frac{T_\parallel}{T_C}\right)Z_0 \approx \left(\frac{r_H}{R_\parallel}\right)^{2/3}\left(\frac{t_\parallel}{T_C}\right)Z_0 \gg Z_0. \quad (57)$$

Such a cloud is therefore a collective formation. It quickly forms in the time $\tau_p \approx 1/\omega_p \approx 3$ ns characteristic of collective interactions.

Under condition (25) of electron-gas magnetization,

$$N_D \approx nR_\parallel^3 \approx \xi_\parallel^{3/2} \quad (58)$$

at $T_\parallel > T_C$. According to (6), the Debye cloud is a natural shell surrounding an ion and consisting of weakly bound electron rings. The relative fluctuations of the shape and size of this cloud are on the order of unity. At $T_\parallel > T_C$, it forms in the time $\tau_p \approx 1/\omega_p$. The kinetics of cloud formation at $T_\parallel > T_C$ requires a detailed analysis, which is outside the scope of the present study. A crude estimate for the formation time is given in the next section. When condition (25) holds, the phase transition temperature is substantially different from (55). Indeed, according to (49) and (50), the number of electrons in the contracted cloud is

$$N_D \approx nr_H^3(\gamma - 2)^{-1}\left(\frac{r_H}{R_T}\right)^{\gamma - 2}. \quad (59)$$

Combining this with the screening condition $N_D = Z_0$ (neutrality of the ion–shell system), we obtain

$$\gamma_C \approx 2 + \frac{\ln(Z_0/nr_H^3)}{\ln(r_H/R_T)}. \quad (60)$$

The phase transition discussed here occurs if condition (5) is satisfied, i.e., when the electron energy is characterized by a highly anisotropic and nonequilibrium distribution. Probably, "the formation of a nonequilibrium dissipative structure" would be a more appropriate term for this transition.

5. COMPARISON WITH EXPERIMENTS AND DISCUSSION

There is no direct experimental evidence of the effect under consideration. Indirect support to its existence is provided by the observation that the experimental rates of radiative recombination of electrons with fully stripped ions exceed their theoretical values [5, 19–21]. An increase in recombination is easy to explain by assuming that a decrease in $T_\parallel$ or the ion velocity relative to an electron cloud whose temperature is $T_\parallel < T_C$ leads to an increase in electron density around the ion. Estimates made by using (55) and (60) agree with the observed energies of electron motion relative to ions at which the recombination cross section begins to increase. In particular, for uranium ions and a plasma with parameters (4) and (5), it follows from (60) that $T_C = 0.001$ eV, which agrees with results obtained in [5]. For fully stripped carbon ions at $T_\perp = 0.008$ eV (obtained by adiabatic transverse expansion of an electron beam diverted into a low magnetic field region), $H = 400$ G, and $n = 10^7$ cm^{-3}, we find that $T_C = 5 \times 10^{-5}$ eV, in agreement with results obtained in [21]. Let us discuss the possibility of explaining beam crystallization [22–26] by the phase transition].

In [22, 23], it was found that the proton momentum spread in the NAP-M proton storage ring began to decrease rapidly as the number of ions N_i in the ring decreased to $N_i \approx 3 \times 10^7$. This corresponded to a decrease in the proton temperatures $\tau_\parallel$ and $\tau_\perp$ to the lowest attainable value, $\tau_\parallel \approx \tau_\perp \approx T_\parallel \approx 10^{-4}$ eV [2]. The Schottky noise level in the beam dropped simultaneously, manifesting the onset of correlation between proton locations in the beam. For the number of protons in the ring specified above and the ring length $C_{ring} = 65$ m, the mean distance between protons is $l_\parallel \approx 3 \times 10^{-4}$ cm, and the mean energy of their repulsion is $\bar{U} \approx e^2/l_\parallel \approx 5$ K. These observations suggest that the protons in the beam are lined up in an equally spaced arrangement [24, 25]. A similar effect was observed in experiments with heavy ions, such as uranium ions [26], where it was found that the momentum spread narrowed by two orders of magnitude as the number of ions in the ring decreased to $N_i \approx 10^3$, which corresponded to $\bar{U} \sim Z_0^2 e^2/l_\parallel \sim 5$ K at $C_{ring} = 108$ m. The following possible explanation of this effect was given in [27]. The time-dependent ion temperatures $\tau_\parallel$ and $\tau_\perp$ in a storage ring are described by the approximate equations

$$\begin{aligned} \frac{d\tau_\perp}{dt} &= q - \lambda(\tau_\perp - \tau_\parallel) - \lambda_C\tau_\perp, \\ \frac{d\tau_\parallel}{dt} &= q - \lambda(\tau_\perp - \tau_\parallel) - \lambda_C\tau_\parallel, \end{aligned} \quad (61)$$

where the transverse heating of ions in the beam due to machine resonances [2] is represented by q; λ is the rate

constant for the transverse–longitudinal relaxation of ions through the so-called in-beam ion–ion scattering, which results in energy exchange between transverse and longitudinal ion motions; and λ_C is the rate constant for electron cooling of ions. (For simplicity, it is assumed in (61) that the values of λ_C for transverse and longitudinal motions are equal). Under steady beam conditions, Eq. (61) reduces to

$$\tau_\parallel = \Lambda q, \tag{62}$$

$$\frac{\tau_\parallel}{\tau_\perp} = \frac{\lambda}{\lambda + \lambda_C} \approx \frac{\lambda}{\lambda_C}, \tag{63}$$

where

$$\Lambda = \frac{\lambda}{\lambda_C(2\lambda + \lambda_C)} \approx \frac{\lambda}{\lambda_C^2}. \tag{64}$$

Here, we use the fact that, typically, $\lambda \ll \lambda_C$ and therefore $\tau_\parallel \ll \tau_\perp$; more precisely (see [28]),

$$\frac{\tau_\parallel}{\tau_\perp} \approx 0.001. \tag{65}$$

If $\tau_\parallel \ll \tau_\perp$, then λ_C can be treated as independent of $\tau_\parallel$. This approximation is, however, inapplicable to λ: according to calculations [29], the in-beam scattering constant λ reaches a maximum at $\tau_\parallel \approx 1$ K. The physical nature of this maximum is explained as follows. At $\tau_\parallel \gg 1$ K, the beam is a gas of freely moving ions, and their transverse–longitudinal relaxation is controlled by their interactions at scattering distances estimated as

$$R_{Ti} \approx Z_0^2 e^2/\tau_\perp \approx Z_0^2 e^2/\tau_\parallel$$

(following [29], we consider the case when $\tau_\parallel \approx \tau_\perp \equiv \tau$). Since R_{Ti} decreases as τ increases, λ also decreases (as $\tau^{1/2}$). At $\tau < 1$ K, the ions line up into a stable one-dimensional chain. The amplitude of their longitudinal oscillation scales with $\tau^{1/2}$ as τ grows, and λ increases accordingly. Therefore, the function $\lambda(\tau)$ has a maximum (Fig. 3). As $\tau \longrightarrow 0$, the transverse–longitudinal relaxation due to accidental (e.g., magnetic field) inhomogeneities becomes essential, and the right-hand side of (62) tends to a nonzero constant (see Fig. 3). The distance $l_\parallel$ between ions in the storage ring increases as their number N_i decreases. Therefore, $\lambda = \lambda(\tau, l_\parallel)$ also decreases. Thus, Eq. (62) has three roots $\tau_\parallel \equiv \tau$ if $N_i > N_C$ ($\tau_1 < \tau_2 < \tau_3$) and only one root if $N_i < N_C$. The root τ_2 corresponds to an unstable solution to (61): if the initial ion temperature τ_0 lies in the interval $\tau_1 < \tau_0 < \tau_2$, then cooling is faster than the heating via in-beam scattering, and therefore $\tau \longrightarrow \tau_1$ with time elapsed. Similarly, if $\tau_2 < \tau_0 < \tau_3$, then $\tau \longrightarrow \tau_3$. When $N_i = N_C$, the beam state bifurcates: the ion temperature drops from

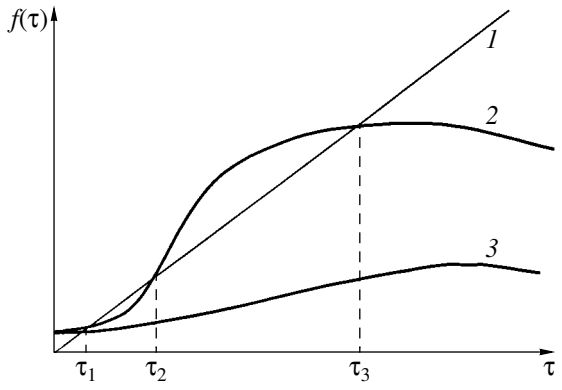

Fig. 3. Graphic solution of (62) at $\tau_\parallel = \tau_\perp = \tau$: (*1*) $f(\tau) = \tau$ (left-hand side of (62)): (*2*, *3*) $f(\tau) = \Lambda q$ (right-hand side of (62)) at $N_i > N_C$ and $N_i < N_C$, respectively.

τ_3 (which coincides with τ_2) to τ_1, in agreement with experiment. In reality [see Eq. (65)], λ begins to decrease with increasing $\tau_\parallel$ only at $\tau_\parallel > \tau_\perp \approx 2000$ K; that is, the temperatures τ_2 and τ_3 are two to three orders of magnitude higher than τ_1. Therefore, the validity of the bifurcation mechanism proposed in [27] should be proved by additional calculations.

Let us now analyze another mechanism of beam crystallization that can be inferred from (62) and (64): a sharp increase in λ_C due to the phase transition considered in here (this possibility was brought to my attention by A.O. Sidorin). When an ion enters a cooler, an electron cloud (coat) immediately forms around it (see estimates below), with an electron binding energy $\varepsilon \approx (2–3)T_\parallel$. An ion leaving the cooler passes through the region of transverse magnetic field $H_1 \approx 50$ G used to divert the electron beam into a collector. In the ion rest frame, this field acts as the electric field $E = H_1 v/c$, where v is the ion velocity in the laboratory frame. The potential difference across the coat ($\sim r_H$) is sufficiently large ($eEr_H \sim 30$ eV $\gg \varepsilon$) to strip the ion as it leaves the cooler. (This process is sometimes called Lorentz ionization, because it is the difference between the accelerations gained by the electrons and the ion under the Lorentz force acting in the laboratory frame, due to the difference in their respective e/m, that destroys the coat.) Thus, every ion that enters the cooler captures Z_0 electrons and carries them along as it passes through the cooler. All electrons contained in a in-beam cylindrical region of length $l \sim Z_0/nS$ are captured by each ion, where S is the beam cross section. If

$$N_i > N_C' = \frac{C_{\text{ring}}}{l}, \tag{66}$$

then the number of available electrons is not sufficient for all ion to be coated. Conversely, if $N_i > N_C'$, then all ions are coated. If the frictional force exerted by elec-

trons on the ion (calculated below) increases, then so does λ_C. In a pioneering experiment with protons [22], the proton energy was $E_I = 65$ MeV and the remaining parameters were $n = 2.4 \times 10^8$ cm^{-3}, $S \approx 0.8 \times 10^{-4}$ cm^2, and $Z_0 = 1$. Therefore, $l \approx 0.5 \times 10^{-4}$ cm, and (66) yields $N'_C \approx 10 \times 10^7$. In the experiment reported in [26], E_i = 400 MeV/nucleon, $n = 2 \times 10^6$ cm^{-3}, $S \approx 0.5 \times 10^{-5}$ cm^2, and $Z_0 = 92$; therefore, $l \approx 10$ cm and $N'_C \approx 10^3$. These estimates agree with the measured critical numbers of ions given above, $N'_C \sim N_C$, suggesting that beam crystallization can be attributed to the formation of ion coats.

To show that the coats do form, let us estimate the probability for an electron ring to be captured into a bound state by an ion under condition (25). The energy of ions in the electron-cloud rest frame is on the order of $T_\parallel$; they can therefore be treated as immobile. According to (28), the energy of ion–ring interaction in a typical collision with an impact parameter on the order of r_H is

$$U(z) \approx -\left(\frac{Z_0 e^2}{r_H}\right)\theta(r_H - |z|). \tag{67}$$

As the ring enters the region $|z| < r_H$, it accelerates and therefore emits plasmons via the bremsstrahlung effect. Then, the ring is decelerated by the arising frictional force **F**, losing its total energy until it ends up in a bound state. Estimates show that this is the most effective mechanism of capturing. The ring is treated as a point charge $q = -e$ moving along a trajectory $\mathbf{r} = \mathbf{r}_C(t)$. According to [12–14],

$$\mathbf{F} = q\int \frac{d^3k d\omega}{(2\pi)^4}\left(-\frac{4\pi i\mathbf{k}}{k^2}\right)$$
$$\times \exp(i\mathbf{k}\cdot\mathbf{r}_C(t) - i\omega t)\rho(\mathbf{k},\omega)\left(\frac{1}{\varepsilon(\mathbf{k},\omega)-1}\right), \tag{68}$$
$$\rho(\mathbf{k},\omega) = q(\exp(i\mathbf{k}\cdot\mathbf{r}_C(t))_\omega,$$
$$(f(t))_\omega = \int_{-\infty}^{\infty} dt f(t)\exp(-i\omega t).$$

As noted at the end of Section 2, we can ignore the terms with $s \neq 0$ in (8) and assume that $k_\perp r_H \ll 1$ if condition (25) holds. For estimation purposes, it is sufficient to consider the cold-plasma limit $T_\parallel \longrightarrow 0$ (which should be taken after calculating the integral with respect to $dv_\parallel$ in (8) by parts):

$$\varepsilon(\mathbf{k},\omega) \approx 1 - \frac{\omega^2(\mathbf{k})}{\omega^2}, \tag{69}$$

where $\omega(\mathbf{k}) = \omega_p|k_z|/k$ is the dispersion relation for Langmuir waves in a magnetized electron cloud.

Using (68), we find the longitudinal momentum lost by the ring through emission of plasmons:

$$\Delta p = \int_{-\infty}^{\infty} dt F(t) = -(2\pi)^{-2}\omega_p^2 \tag{70}$$
$$\times \int d^3k d\omega |\rho(\mathbf{k},\omega)|^2 k_z^3 k^{-4}\operatorname{sgn}\omega\,\delta(\omega^2 - \omega^2(\mathbf{k})).$$

According to (67),

$$\rho(\mathbf{k},\omega) = -e\int_{-\tau}^{\tau} dt\exp(i(\omega - k_z V)t) \approx -2e\tau, \tag{71}$$

where $V \approx \Delta_\parallel$ is the longitudinal ring velocity in the region of its interaction with an ion and $\tau = r_H/V$. In expression (70) use is made of the fact that, by virtue of (25),

$$|(\omega - k_Z V)t| \approx \frac{\omega_p r_H}{\Delta_\parallel} \approx \frac{r_H}{R_\parallel} \ll 1. \tag{72}$$

Combining Eqs. (70) and (71), we obtain the estimate

$$\Delta p = -\pi^{-2}e^2\tau^2\omega_p\int d^2k_\perp\int_0^\infty dk_Z k_Z^2 k^{-3} \tag{73}$$
$$\approx -e^2\tau^2\omega_p r_H^{-2} \sim -e^2\omega_p V^{-2} \approx -e^2\omega_p\Delta_\parallel^{-2},$$

where the divergence at large k arises as the ring is replaced by a point particle; in reality, $k < k_{max} \approx 1/r_H$. Furthermore, since R_T is small, the most probable process is the ring–ion recombination into a state with $\varepsilon \approx T_\parallel$ and

$$l_\parallel \approx l_\perp \approx r_H, \tag{74}$$

followed, with a probability of about 1/2, either by liberation of the ring or its relaxation into a strongly bound state with

$$l_\parallel \approx R_T, \quad \varepsilon \approx T_\parallel \ln(r_H/R_T).$$

Estimates show that, since R_T/r_H is small, the ion residence time in the cooler is not sufficient for complete relaxation. Therefore, if $T_\parallel < T_C$ and condition (25) holds, then the dimensions of the electron cloud are given by (74) rather than the tabulated values. Combining (73) with (6), we obtain

$$|\Delta p|/p \approx (\xi_\parallel)^{-3/2} \sim 1,$$

which implies that the probability for a ring to be captured by an ion in a collision with an impact parameter on the order of r_H is

$$W_C \sim 1. \tag{75}$$

An ion captures one ring from a cylindrical region of length estimated as

$$\frac{1}{nr_H^2} \approx 0.01\text{–}0.1 \text{ cm}. \quad (76)$$

The fact that this length is small quantitatively substantiates the above-described qualitative model of coat formation around an ion.

Since the increase in λ_C necessary for beam crystallization requires a frictional force that increases in the course of coat formation, let us estimate the frictional force F that acts on an ion. This can be done by the simple method illustrated here by estimating F for an equilibrium plasma at a temperature T and $H = 0$ (the corresponding F value was rigorously calculated, e.g., in [12]). An estimate can be obtained by retaining only the collective term in F (68) [2, 3] and ignoring the contribution due to pair interactions [30]. If $V = 0$, then the Debye cloud is spherically symmetric and the mean force acting on an ion is zero. If

$$V < \Delta = \sqrt{T/m},$$

then the Debye cloud is centered at a distance $x \approx r_D V/\Delta$ behind the ion. Therefore, a nonzero electric field E_D is generated by the Debye cloud at the ion, location, and the ion is decelerated by the force

$$F = qE_D \approx \frac{x}{r_D}\frac{q^2}{r_D^2} \approx \frac{Z_0^2 e^2 V}{r_D^2 \Delta} \approx \frac{nZ_0^2 e^4 V}{m\Delta^3}. \quad (77)$$

If $V > \Delta$, then the ion is at the periphery of a cloud of size

$$L \approx V\tau_p \approx \frac{V}{\omega_p}, \quad (78)$$

because the response of electrons to the field of the ion is delayed by $\tau_p \approx 1/\omega_p$ and the corresponding distance traveled by the ion is L. Hence,

$$F \approx \frac{q^2 e^2}{L^2} \approx \frac{nZ_0^2 e^4}{mV^2}.$$

If conditions (21), (23), and $T_\parallel > T_C$ are met, then, by virtue of (57), the electron density in the Debye cloud is comparable to n, and the rearrangement of charge-rings that move along axis z with velocities on the order of $\Delta_\parallel$ takes a time $\tau_p \approx r_D/\Delta_\parallel$. Therefore, the cloud moves at a distance $x \approx V\tau_p \approx r_D V/\omega_p$ behind the ion (for simplicity, it is assumed hereinafter that the ion moves along **H**). By analogy with (77), if $V < \Delta_\parallel$, then

$$F \approx \frac{q^2}{r_D^2}\frac{V}{\Delta_\parallel} \approx Z_0^2 e^2 (r_H R_\parallel^2)^{-2/3}\frac{V}{\Delta_\parallel}. \quad (79)$$

When $V \sim \Delta_\parallel$, the ion is at the periphery of an approximately spherical Debye cloud, and the frictional force reaches a maximum,

$$F \approx \frac{q^2}{r_D^2} \approx Z_0^2 e^2 (r_H R_\parallel^2)^{-2/3}. \quad (80)$$

If

$$\Delta_\parallel < V < V_0,$$

where

$$V_0 = \Delta_\parallel r_H / R_\parallel,$$

then the charge moves faster than the rings, and the rearrangement of ring positions takes a time on the order of $1/\omega_p$. For this reason, $R_\parallel = \Delta_\parallel/\omega_p$ should be replaced by $L = V/\omega_p$ in (80), by analogy with (78). This yields

$$F \approx Z_0^2 e^2 (r_H R_\parallel^2)^{-2/3}\left(\frac{\Delta_\parallel}{V}\right)^{4/3}. \quad (81)$$

(Note that result can be derived from rigorous equations (68) and (8).) In this case, the Debye cloud is approximately spherical, its radius is

$$r_D' \approx (r_H L^2)^{1/3},$$

and the ion is at its periphery. For

$$V \approx V_0, \quad r_D' \approx r_H,$$

we have

$$F \approx \frac{Z_0^2 e^2}{r_D'^2} \approx \frac{nz_0^2 e^4}{mV_0^2}.$$

Finally, if $V > V_0$, then Larmor orbits can be treated as point charges moving along H. The major contribution to the frictional force is then due to the pole in (68) (i.e., to $\omega = \mathbf{k} \cdot \mathbf{V}$ at which $\varepsilon(\mathbf{k}, \omega) = 0$, which corresponds to emission of Cherenkov plasmons by the ion [31],

$$F \approx \frac{2\pi nZ_0^2 e^4}{mV^2}.$$

Of all velocity intervals considered above, the case of $V \ll \Delta_\parallel$ is of primary interest for practice. According to the table, a spherical Debye cloud of radius on the order of r_D transforms into a thin disk having the same cross radius at $T_\parallel < T_C$. The electric field at the disk boundary is 2 to 3 times stronger than on the boundary of the sphere that has transformed into the disk. Being attracted by the ion, electron rings move chaotically inside the disk with velocity Δ higher than $\Delta_\parallel$,

$$\Delta \approx (2\text{–}3)\Delta_\parallel.$$

Substituting it into (79) instead of $\Delta_\parallel$, we find that the frictional force does not change substantially as a result of the phase transition if conditions (21), (23) are met.

Now, consider the case when (25) holds, as in the beam crystallization experiments considered here. According to (75), for $T_\parallel < T_C$ and a fully developed coat, the frictional force is

$$F \approx m\Delta_\parallel n r_H^2 \Delta_\parallel \left(\frac{V}{\Delta_\parallel}\right) = mnr_H^2 \Delta_\parallel V. \quad (82)$$

Expression (82) reflects the transfer of a momentum on the order of $m\Delta_\parallel$ from every captured electron ring to the ion and the dynamic equilibrium between the number of rings captured by an ion per unit time and the number of rings lost in collisions with electrons. According to (25) and (26), the frictional force at $T_\parallel > T_C$ can be estimated as follows:

$$F \approx Z_0^2 e^2 R_\parallel^{-2} (V/\Delta_\parallel).$$

In view of (55), this result coincides with (82).

It follows that, both in the case when (21) and (23) and under condition (25), the frictional force does not change substantially in the phase transition. Thus, ion beam crystallization can currently be explained only within the framework of the bifurcation model [27].

6. CONCLUSIONS

The phase transition discussed in this paper explains why the radiative electron–ion recombination cross sections observed experimentally [5, 19–21] exceed their calculated values. According to the estimates given above, this transition is not related to ion beam crystallization [22–26]. Indeed, the formation of a small electron cloud around an ion can not cause a sharp increase in the frictional force that acts on the ion.

A key role in the phase transition under consideration is played by the sharp anisotropy of electron distributions in coolers characterized by condition (5). As the magnetic field strength H increases, this anisotropy changes the character of ion–electron interaction. When $H = 0$, an electron interacts as a point particle, whereas, it can be considered a ring of radius r_H with a uniformly distributed charge under condition (21). This change leads to a phase transition as $T_\parallel$ decreases. Mathematically (but not physically), the phase transition is analogous to the Berezinskii–Kosterlitz–Thouless transition [16, 17]. At $T_\parallel \approx T_C$, the Debye cloud surrounding an ion contracts as demonstrated in the table. However, the actual ion residence time in a cooler is not sufficiently long for the cloud to contract to the longitudinal size R_T. The actual size is substantially larger. This also follows from the logarithmic dependence on z_1 obtained by substituting $\gamma = 2$ into (49) and (50) when, as explained at the end of Section 4. Rigorous longitudinal size calculations can be performed by solving the Bogolyubov–Born–Green–Kirkwood–Yvon chain of equations [34, 35]. These calculations can also be used to determine the type of phase transition. According to the argumentation at the end of Section 4, it should be a second-order phase transition. This, however, requires a separate study.

The phase transition is totally unrelated to nonideality of the initial plasma. In particular, for a plasma magnetized with respect to pair collisions,

$$\bar{r}_H \ll \bar{R}, \quad (83)$$

the relation

$$T_\parallel \gg \frac{Z_0 e^2}{\bar{R}} \geq \frac{e^2}{\bar{R}} \quad (84)$$

follows from phase transition condition (55). Therefore, if (83) holds, then the phase transition can occur only in a plasma that is ideal with respect to longitudinal motion (84). (A fortiori, by virtue of inequality (5), this plasma is ideal with respect to transverse motion.) Note also the essential difference between the phase transition in question and the nonlinear Debye screening studied in [32, 33].

ACKNOWLEDGMENTS

I thank I.N. Meshkov, I. Seleznev, A.O. Sidorin, A.V. Smirnov, G.V. Trubnikov, and the participants of the seminar moderated by V.D. Shafranov for helpful discussions and critical remarks.

REFERENCES

1. G. I. Budker, At. Énerg. **22**, 346 (1967).
2. V. V. Parkhomchuk and A. N. Skrinskiĭ, Usp. Fiz. Nauk **170**, 473 (2000) [Phys. Usp. **43**, 433 (2000)].
3. I. N. Meshkov, Fiz. Élem. Chastits At. Yadra **25**, 1487 (1994) [Phys. Part. Nucl. **25**, 631 (1994)].
4. G. I. Budker, F. F. Bulyshev, and N. S. Dikanskiĭ, in *Proceedings of 5th All-Union Meeting on Charged Particle Accelerators* (Nauka, Moscow, 1977), Vol. 1; Preprint No. 76-92, IYaF SO AN SSSR (Inst. of Nuclear Physics, Siberian Division, USSR Academy of Sciences, Novosibirsk, 1976).
5. W. Shi, S. Bohm, C. Bohme, *et al.*, Eur. Phys. J. D **15**, 145 (2001).
6. A. V. Gaponov, Izv. Vyssh. Uchebn. Zaved., Radiofiz. **2**, 450 (1959).
7. V. Yu. Trakhtengerts, in *Basic Plasma Physics*, Ed. by A. A. Galeev and R. N. Sudan (Énergoatomizdat, Moscow, 1984; North-Holland, Amsterdam, 1984), Vol. 2.
8. S. Ishimaru and M. N. Rosenbluth, Phys. Fluids **13**, 2778 (1970).
9. D. Montgomery, G. Joyce, and L. Turner, Phys. Fluids **17**, 954 (1974); Phys. Fluids **17**, 2201 (1974).

10. A. W. Hyatt, C. F. Driscoll, and J. H. Malmberg, Phys. Rev. Lett. **59**, 2975 (1987).
11. B. R. Beck, J. Fajans, and J. H. Malmberg, Phys. Rev. Lett. **68**, 317 (1992).
12. V. D. Shafranov, in *Problems of the Plasma Theory* (Gosatomizdat, Moscow, 1963), No. 3 [in Russian].
13. V. V. Zheleznyakov, *Radiation in Astrophysical Plasma* (Yanus-K, Moscow, 1997) [in Russian].
14. A. F. Aleksandrov and A. A. Rukhadze, *Lectures on Electrodynamics of Plasma-Like Media* (Mosk. Gos. Univ., Moscow, 1999) [in Russian].
15. E. M. Lifshitz and L. P. Pitaevskiĭ, *Course of Theoretical Physics*, Vol. 10: *Physical Kinetics* (Nauka, Moscow, 1979; Pergamon, Oxford, 1981).
16. V. L. Berezinskiĭ, Zh. Éksp. Teor. Fiz. **59**, 907 (1970) [Sov. Phys. JETP **32**, 493 (1970)]; Zh. Éksp. Teor. Fiz. **61**, 1144 (1971) [Sov. Phys. JETP **34**, 610 (1972)].
17. J. M. Kosterlitz and D. J. Thouless, J. Phys. C **6**, 1181 (1973).
18. L. D. Landau and E. M. Lifshitz, *Course of Theoretical Physics*, Vol. 5: *Statistical Physics* (Nauka, Moscow, 1976; Pergamon, Oxford, 1980), Part 1, Sect. 8.
19. A. Muller, S. Schennach, M. Wagner, *et al.*, Phys. Scr. **37**, 62 (1991).
20. H. Gao, D. R. DeWitt, R. Schuch, *et al.*, Phys. Rev. Lett. **75**, 4381 (1995).
21. G. Gwinner, A. Hoffknecht, T. Bartsch, *et al.*, Phys. Rev. Lett. **84**, 4822 (2000).
22. G. I. Budker, N. S. Dikansky, V. I. Kudelainen, *et al.*, in *Proceedings of 4th All-Union Conference on Charged Particle Accelerators* (Nauka, Moscow, 1975), Vol. 2, p. 309.
23. E. N. Dement'ev, N. S. Dikanskiĭ, A. S. Medvedko, *et al.*, Preprint No. 70-79, IYaF SO AN SSSR (Inst. of Nuclear Physics, Siberian Division, USSR Academy of Sciences, Novosibirsk, 1979).
24. V. V. Parkhomchuk, in *Proceedings of Workshop on Electron Cooling and Related Applications* (Kernforschungszentrum, Karlsruhe, 1984).
25. V. V. Parkhomchuk and A. N. Skrinsky, Rep. Prog. Phys. **54**, 919 (1991).
26. M. Steck, K. Beckert, H. Eickhoff, *et al.*, Phys. Rev. Lett. **77**, 3803 (1996).
27. I. Meshkov, A. Sidorin, A. Smirnov, *et al.*, Preprint RIKEN-AF-AC-34 (RIKEN, Tokyo, 2002).
28. B. Hochadel, F. Albrecht, M. Grieser, *et al.*, in *Proceedings of Workshop on Beam Cooling and Related Topics* (CERN, Geneva, 1994), p. 198.
29. J. Wei, H. Okamoto, and A. Sessler, Phys. Rev. Lett. **80**, 2606 (1998).
30. B. A. Trubnikov, *The Plasma Theory* (Énergoatomizdat, Moscow, 1996) [in Russian].
31. A. S. Artamonov and Ya. Derbenev, Nucl. Instrum. Methods Phys. Res. A **284**, 248 (1989).
32. I. Hofmann, in *Proceedings of Workshop on Beam Cooling and Related Topics* (CERN, Geneva, 1994), p. 330.
33. A. Wolf, C. Ellert, M. Grieser, *et al.*, in *Proceedings of Workshop on Beam Cooling and Related Topics* (CERN, Geneva, 1994), p. 416.
34. N. N. Bogolyubov, *Dynamical Theory Problems in Statistical Physics* (Gostekhizdat, Moscow, 1946) [in Russian].
35. T. L. Hill, *Statistical Mechanics: Principles and Selected Applications* (McGraw-Hill, New York, 1956; Inostrannaya Literatura, Moscow, 1960).

Translated by V. Sipachev

Journal of Experimental and Theoretical Physics, Vol. 100, No. 2, 2005, pp. 422–427.
From Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 127, No. 2, 2005, pp. 471–477.

STATISTICAL, NONLINEAR, AND SOFT MATTER PHYSICS

Pyroelectric and Electromechanical Properties of the Antiferroelectric Liquid Crystal MHPOBC¶

S. V. Yablonskii*[a,b], K. Nakano[a], M. Ozaki[a], and K. Yoshino[a]

[a]*Department of Electronic Engineering, Graduate School of Engineering, Osaka University 2-1 Yamada-Oka, Suita, Osaka 565-0871, Japan*

**e-mail: lbf@ns.crys.ras.ru*

[b]*Institute of Crystallography, Russian Academy of Sciences, Moscow, 117333 Russia*

Received June 30, 2004

Abstract—We study pyroelectric and electromechanical effects in the prototype antiferroelectric liquid crystal 4-(1-methylheptyloxycarbonyl-phenyl)4'-octylbiphenil-4-carboxylate (MHPOBC). The linear electromechanical effect in the freely suspended liquid crystal films of MHPOBC has been detected in the broad temperature range inclusive of the antiferroelectric SmC_A^* as well as paraelectric Sm*A*. The anomalous behavior of the hysteresis loop of SmC_β^* in the (pyroelectric coefficient, dc bias electric voltage) coordinates has been found.

1. INTRODUCTION

Antiferroelectricity of a liquid crystal was first discovered in 4-(1-methylheptyloxycarbonyl-phenyl)4'-octylbiphenil-4-carboxylate (MHPOBC) [1, 2]. This material has three chiral smectic subphases (SmC_α^*, SmC_β^*, and SmC_γ^*) between paraelectric Sm*A* and antiferroelectric SmC_A^*. The arrangements of the molecules in these phases are shown in Fig. 1. Different experimental techniques, such as resonant X-ray scattering [4], dielectric spectroscopy [5, 6], optical rotation [7], conoscopic observation [8], ellipsometry on freely suspended films [9], differential scanning calorimetry [10], and electric current and optical transmittance responses [1, 11] have been used to identify and characterize the structure of these phases.

In chiral mesophases, a spontaneous polarization arises as a secondary order parameter due to a molecular tilt with respect to the smectic layer normal; the $\mathbf{P}_s$ vector lies in the tilt plane, reflecting the polar properties of a liquid crystal [12]. Chirality plays a crucial role for the emergence of polarization in the direction of the tilt plane normal. In smectic phases, where rod-like molecules are organized into fluid stacks of planar layers, weak chiral twisting forces induce a helical order with the helical axis parallel to the layer normal.

The resulting helical pitch is in the optical wavelength range. If the pitch is large, it is easy to unwind the helix by application of a weak strain, for example, imposed by a mechanical shear or by a weak applied electric field. Then the sample has a ferroelectric response under moderate applied electric and mechanical fields.

In this paper, the polar properties of chiral smectic subphases of the classic antiferroelectric liquid crystal MHPOBC are investigated by recording its pyroelectric coefficient and studying the vibration of freely suspended films induced by alternating electric voltage (the so-called linear electromechanical effect) [13]. The pyroelectric activity of the "unwound" chiral liquid crystal confirms the polar structure of mesophases, whereas the thermal behavior of the linear electromechanical response reveals specific features related to phase transitions of smectic subphases and their structure.

2. EXPERIMENTAL

2.1. Pyroelectric Set-Up and Cell

To study the antiferroelectric and ferrielectric behavior of our material, we used a pulse pyroelectric setup shown in Fig. 2 [14]. A 100 μs pulse of a Nd^{3+} YAG laser was used to provide a small local temperature change in the sample. Laser radiation (λ = 1.06 μm) was partly absorbed in indium-tin-oxide (ITO) layers. The pyroelectric response was measured as a pulse voltage across the load resistor R_L = 100 kΩ with a wideband amplifier and a storage oscilloscope. A dc field of various strengths was applied to the sample in order to measure hysteresis loops in the coordinates (pyroelectric response, dc bias voltage).

The temperature dependence of the spontaneous polarization (on an arbitrary scale) was calculated by

¶ This article was submitted by authors in English.

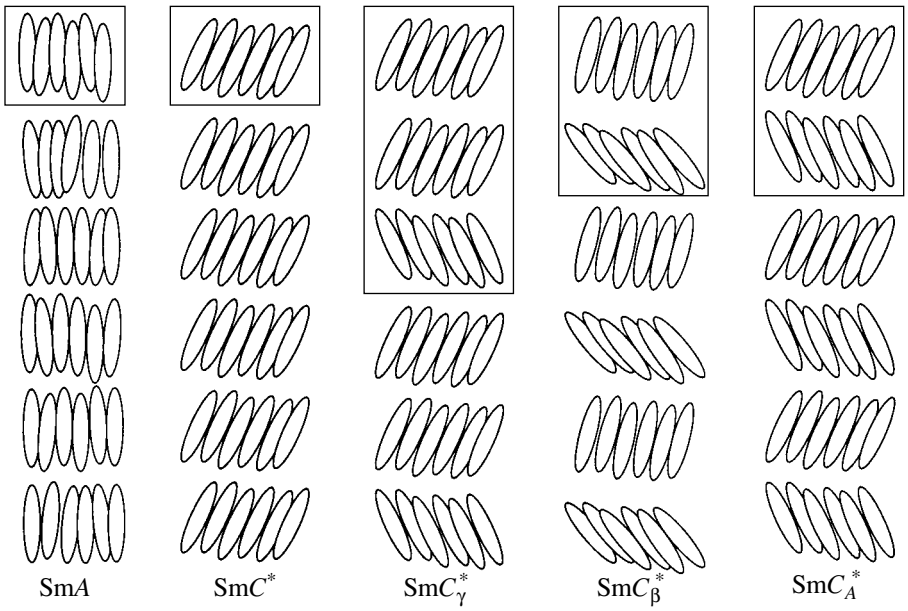

Fig. 1. Model structures of the paraelectric Sm*A*, ferroelectric Sm*C**, ferrielectric $\mathrm{Sm}C^*_\gamma$, $\mathrm{Sm}C^*_\beta$, and antiferroelectric $\mathrm{Sm}C^*_A$ phases. The unit cells are marked by frames. In $\mathrm{Sm}C^*_\alpha$ (not shown in the picture), the superlattice incommensurate period (corresponding to the helical pitch) ranges between 5 and 8 smectic layers [3].

integrating the pyroelectric coefficient γ in accordance with the expression [14]

$$P_s(T) = \int_{T_i}^{T} \gamma(T)dT, \qquad (1)$$

where T_i is the temperature of the transition to the paraelectric phase. Then the correct scale for γ and P_s was introduced by comparison of the pyroelectric response at a certain temperature with the value measured for a well-known ferroelectric liquid crystal. Strictly speaking, Eq. (1) is valid only for the field-off regime. When an external dc field is applied to prepare a ferroelectric monodomain or an "unwound" antiferroelectric (or ferrielectric), the actually measured quantity is the total polarization $\mathbf{P} = \mathbf{P}_s + \mathbf{P}_i$, where $\mathbf{P}_i$ is the field-induced contribution, for example, observed in the isotropic phase. Because $\mathbf{P}_i$ is much smaller than $\mathbf{P}_s$, we can use Eq. (1) for the measurements of $\mathbf{P}_s$ in the "unwound" antiferroelectric phase.

The shape of the hysteresis loop is one of the major criteria to refer the nonlinear dielectrics to ferro-, ferri-, or antiferroelectric types. Hysteresis loops are usually represented in coordinates of the electric displacement against the external electric field. But it is possible to demonstrate that the representation of hysteresis loops in the coordinates given by the pyroelectric coefficient vs the external electric field differs only quantitatively. Indeed, using the definition of the volume polarization $P_s = (N\mu/V)\langle\cos\varphi\rangle$ and the pyroelectric coefficient

$$\gamma = \frac{d\mu}{dT}\frac{N}{V}\langle\cos\varphi\rangle,$$

where N is the number of dipoles of moment μ within the volume V and $\langle\cos\varphi\rangle$ is the average value of the cosine of the angle between the dipole direction and the net polarization direction, we can deduce that the ratio P_s/γ is constant at a given temperature.

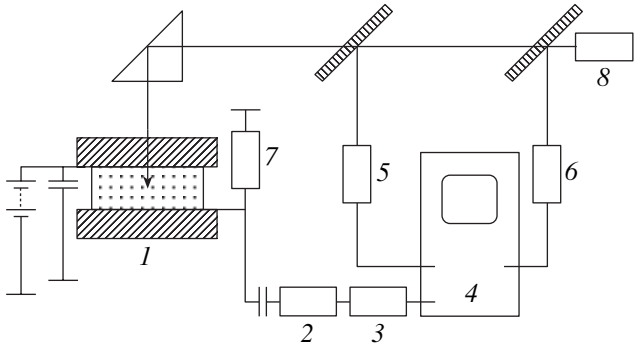

Fig. 2. Pulse pyroelectric set-up: *1*—liquid crystal cell, *2*—amplifier, *3*—time-delay line, *4*—oscilloscope, *5*—photomultiplier to control the form of the laser pulse, *6*—photomultiplier to start-up the line scanning, *7*—load resistor (100 kΩ), *8*—YAG laser.

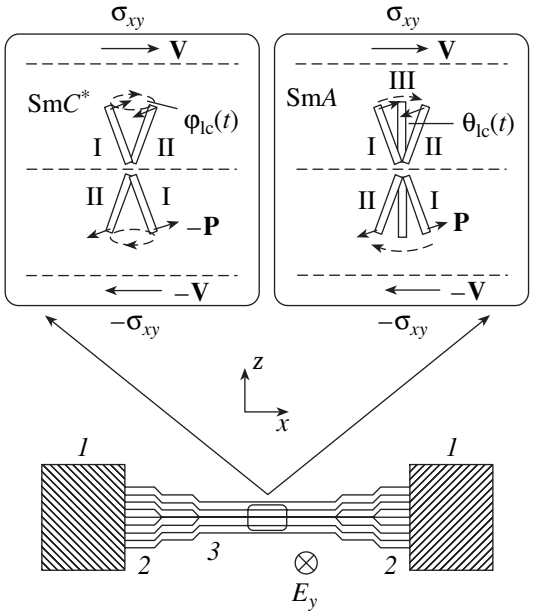

Fig. 3. Schematic representation of a smectic liquid crystal freely suspended film: *1*—frame, *2*—meniscus, *3*—uniform part of the freely suspended film. The insets illustrate a mechanism responsible for the development of a viscous stress σ_{xy} in both ferroelectric and paraelectric phases. $\varphi_{lc}(t)$ and $\theta_{lc}(t)$ are the respective variable azimuthal and zenithal angles. **V** is the liquid crystal velocity field, inhomogeneous along the *z* axis, **P** is the spontaneous polarization. Roman numerals I and II correspond to the respective positions of liquid crystal molecules for the positive and negative electric field.

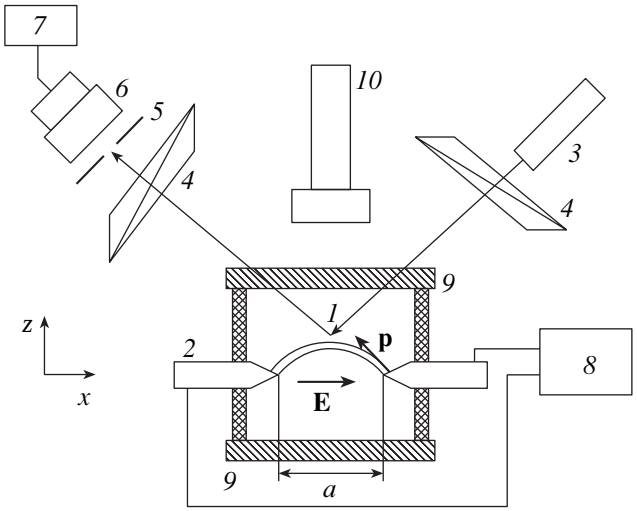

Fig. 4. Experimental set-up for the study of the linear electromechanical effect in freely suspended films. *1*—ferroelectric freely suspended liquid crystal film (MHPOBC), *2*—two Au electrodes, *3*—He–Ne laser (*s*-polarization), *4*—two polarizers, *5*—slit diaphragm, *6*—photodiode, *7*—lock-in amplifier, *8*—audio frequency generator, *9*—two glasses placed near the film surfaces to prevent the influence draughts, *10*—polarization microscope.

2.2. Linear Electromechanical Effect in Ferroelectric Freely Suspended Film

Unlike in the standard case of a liquid crystal sample in a confined geometry, the surface of a freely suspended liquid crystal film can easily be deformed under the action of weak acoustic or electric fields [15, 16]. Such a film is considered a perfect membrane, whose vibrational motion depends only on the film geometry, the isotropic surface tension σ, and the homogeneous two-dimensional density ρ_s (including the inertia of the air moving with the film) [17]. For excitation of the transverse film vibrations, a linear coupling of the lateral electric field to the spontaneous polarization of a ferroelectric liquid crystal has been exploited [13]. In the experiment, the azimuthal motion of the liquid crystal director is accompanied by the so-called back-flow, which induces viscous stress acting on the film surface as shown in Fig. 3. This mechanism assumes a velocity gradient along the normal to the film surface, which implies the presence of an internal structure in freely suspended films. The resulting film deformation strongly depends on the surface viscosity of the liquid crystal and also on the value and sign of spontaneous polarization. This combination of properties makes the measurements of freely suspended film oscillations a useful tool for studying the spontaneous polarization phenomenon in liquid crystals [13].

The periodical displacement of the film surface results in deflection of the probing beam of a low-power He–Ne laser (Fig. 4). By passing the beam through an iris diaphragm, the deflection of the beam is converted into an amplitude modulation of the laser intensity, detected by a photodiode. The photodiode response current is analyzed by a lock-in amplifier tuned to the first harmonic of the control voltage. Microscope observations were carried out by inserting the sample holder on the turntable of the polarization microscope.

2.3. Materials and Samples

The antiferroelectric liquid crystal used in our experiment is S-MHPOBC with a moderate spontaneous polarization (about 70 nC/cm^2 at 112°C) and the following sequence of phase transitions:

$$\text{Iso}–(149.8^\circ\text{C})–\text{Sm}A–(122^\circ\text{C})–\text{Sm}C^*_\alpha–(120.9^\circ\text{C})$$
$$–\text{Sm}C^*_\beta–(119.2^\circ\text{C})–\text{Sm}C^*_\gamma–(118.4^\circ\text{C})–\text{Sm}C^*_A.$$

For pyroelectric measurements, the liquid crystal was introduced in a flat capillary cell made up of ITO-covered, nontreated glass plates with 10-μm thick Teflon spacers. ITO surfaces were cleaned with acetone and used without any orienting layers. The cell with the area between electrodes $A = 5 \times 5$ mm^2 was filled with the liquid crystal in the isotropic phase. Cells were placed in a thermal jacket with optical windows.

The electromechanical effect was studied in freely suspended films fabricated by the standard procedure described in [18]. The experiment was performed with a glass frame of fixed geometry (a rectangular slit with an area of 2×10 mm^2 and a thickness of 1 mm). A frame with a spanned film of MHPOBC was mounted in a heating stage and the film could be stabilized at a given temperature to ±0.3°C.

3. RESULTS AND DISCUSSIONS

Figure 5 presents the results of the measurements of the pyroelectric coefficient γ and spontaneous polarization P_s as functions of the temperature obtained in the cooling process. In Fig. 5, from an analysis of the pyroelectric curve, we can clearly distinguish three ferrielectric phases (SmC^*_α, SmC^*_β, and SmC^*_γ) from Sm*A* and SmC^*_A. According to this plot, the pyroelectric signal is detectable not only in polar ferrielectric phases but also in antiferroelectric SmC^*_A. The pyroelectric signal in these phases appears to be due to a distortion of the helical structure induced by the external electric field. The maximum of the pyroelectric effect occurs in ferrielectric SmC^*_β (which some authors identified as Sm*C** [19]) at $T = 120.3$°C. It is interesting to note that the temperature dependence of the real part of the dielectric constant of MHPOBC peaked at the same temperature [20]. A similar behavior has also been detected in other antiferroelectric systems [21].

The polarization shown in Fig. 5 is obtained by integrating the pyroelectric coefficient over temperature, starting from the temperature $T_0 = 126$°C, about 4°C above the transition from paraelectric Sm*A* to ferrielectric SmC^*_α: a small pyroelectric signal induced by the field in Sm*A* (about 0.04 nC/(cm^2 K)) was subtracted as a background value. The maximum value of polarization in SmC^*_A of 60 nC/cm^2 is comparable with that obtained by the repolarization current technique [22].

The bias dependence of the sum of the induced and spontaneous pyroelectric coefficients measured at $T = 120$°C in SmC^*_β is shown in Fig. 6. This plot drastically differs from the standard hysteresis loop of solid ferroelectrics [23]. First of all, the hysteresis loop shrinks to a thin line, which is typical of ferroelectric liquid crystal materials, where dipoles are ordered in a helical fashion [24]. Second, the nonmonotonic behavior of the pyroelectric signal, shown in this plot, is not common for solid and liquid ferroelectrics. The initial linear growth corresponding to the helix distortion is followed by the decrease of the pyroelectric response at elevated bias field. A decrease in the pyroelectric response with a further increase in the field can be a consequence of two reasons. One of them is explained in [21] as the effect of the competition between the induced and spontaneous polarizations having the opposite signs of the pyroelectric coefficients. The other cause can be connected with the electroclinic effect, promoting a decrease in pyroelectric response with increasing dc bias voltage [25].

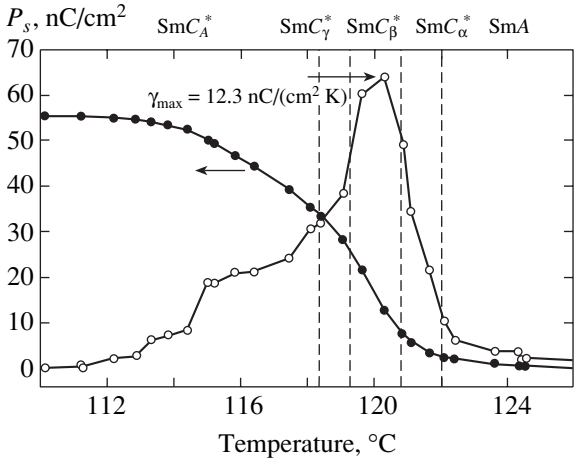

Fig. 5. The pyroelectric coefficient γ (open circles) and the spontaneous polarization P_s (filled circles) of 10 μm-thick cell measured as functions of temperature for the bias dc voltage 10 V.

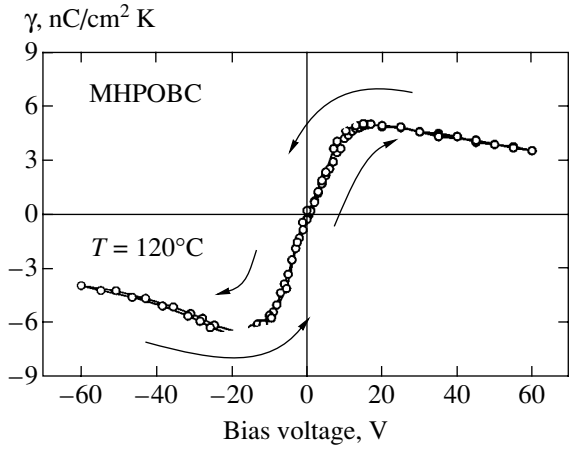

Fig. 6. Hysteresis loop in the coordinates (pyroelectric coefficient, bias dc voltage) for MHPOBC measured at the fixed temperature $T = 120$°C corresponding to the SmC^*_β phase. Arrows show the direction of the round.

Figure 7 demonstrates the temperature dependence of the linear and quadratic electromechanical effects in a relatively thick, about 800 nm, freely suspended film made from MHPOBC. Such a thickness corresponds approximately to two complete turns of the natural smectic helix in long-pitch phases. According to these plots, the linear and quadratic electromechanical effects are present in all phases (with the exception of SmC^*_α in the case of the linear electromechanical effect). The linear effect is rather strong in SmC^*_β, SmC^*_γ, and (unexpectedly) SmC^*_A. The temperature dependence of

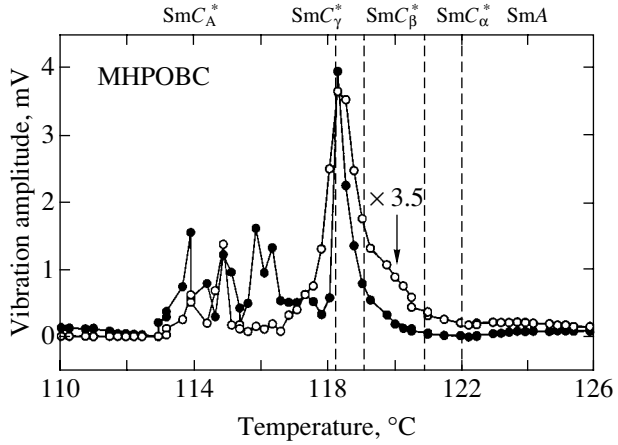

Fig. 7. Temperature dependence of the vibration amplitude in a freely suspended film in MHPOBC of the linear (filled circles) and quadratic (open circles) electromechanical effects. Quadratic amplitudes were multiplied by the factor 3.5. The cooling rate is 5°C/min with the temperature resolution 0.3°C. The rectangular film size is 2×10 mm^2 (a = 2 mm, b = 10 mm). The number of layers is N = 205. Sinusoidal voltage with the amplitude U = 115 V and frequency ν = 2200 Hz was applied to electrodes. A photovoltage of 10 mV corresponds to displacement of the film surface by 800 nm.

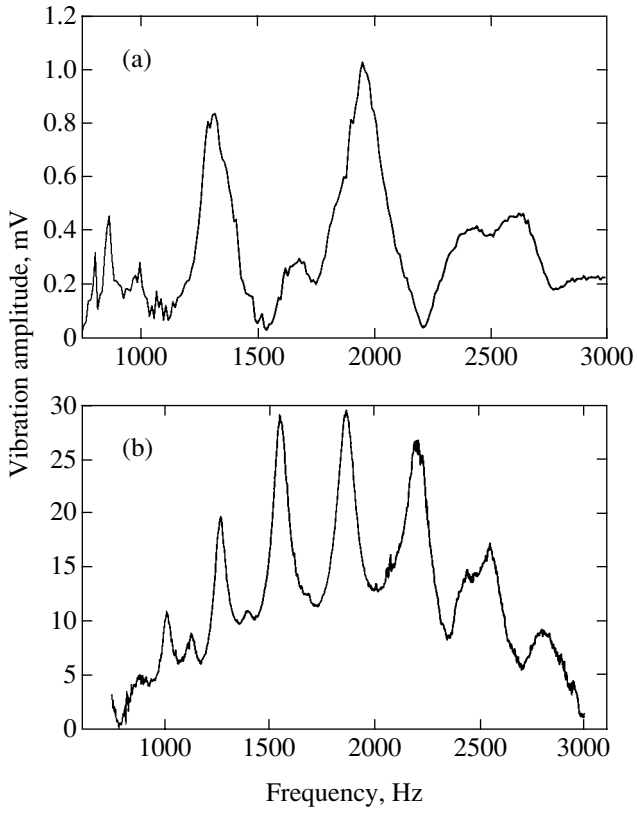

Fig. 8. Spectra of the linear electromechanical effect in a freely suspended film due to: the electroclinic effect (a), T = 135°C (SmA) and the ferroelectric switching (b), T = 114.5°C (SmC_A^*). MHPOBC, N =205, U = 115 V.

the quadratic electromechanical effect correlates closely with the temperature dependence of the real part of the dielectric permittivity [20]. This condition is not random; it simply reflects the influence of the dielectric torque, quadratic in the electric field, on the shape of the film surface. Thus, the cause of the quadratic effect in a freely suspended film is electrostriction, whereas the dominant ferroelectric torque is responsible for the linear effect in unwound ferroelectrics. In the general case of chiral smectics, one should also take flexoelectric and "electroclinic" torques into account. (We use "electroclinic" in quotation marks because the electric field exerts no torque on the director in the electroclinic action but influences only the whole medium, shifting the direction of the equilibrium in space [24].) Obviously, the linear electromechanical effects, shown in Fig. 8, in antiferroelectric SmC_A^* and paraelectric SmA are due to the influence of the sum of the flexoelectric and "electroclinic" torques. The external torques induce a "back-flow," which is linearly coupled with a mechanical stress tensor [26]. Accordingly, the films vibrate with the fundamental and double frequencies of the applied ac electric field.

As one can see in Fig. 7, the linear electromechanical effect is absent in incommensurate tilted SmC_α^*. This phase has an extremely short helical period that typically extends over ten smectic layers and, consequently, requires large power expenses for its disturbance, which evidently could not be realized by relatively weak lateral electric fields used in the experiment (of the order 0.05 V/μm). The difficulty in generating the linear effect is also favored by smallness of the spontaneous polarization in SmC_α^*, as one can see in Fig. 5, and space averaging of the flexoelectric and spontaneous polarization over the chiral structure.

4. CONCLUSIONS

By pyroelectric and electromechanical methods sensitive to both the polar and non-central-symmetric ordering of molecules in liquid crystalline media, we investigated the prototype antiferroelectric liquid crystal MHPOBC. Our observations confirm the polar properties of MHPOBC in ferrielectric phases and in the "unwound" antiferroelectric state. We also found that MHPOBC manifests a linear electromechanical effect in unipolar non-central-symmetric SmA and SmC_A^*. This phenomenon can be interpreted as the effect of electroclinic and flexoelectric torques. In addition, we did not succeed in observing the linear electromechanical effect in noncommensurate SmC_α^*, which is seemingly connected with its nanoscale orientational order.

ACKNOWLEDGMENTS

We acknowledge financial support from a Grant-in-Aid for Scientific Research from the Ministry of Education, Science, Sports, and Culture. One of S.V.Ya. thanks the Japanese Society for Promotion of Science for a Fellowship at Osaka University and the Russian Foundation for Basic Research (project nos. 03-02-17288 and 04-02-16466).

REFERENCES

1. A. D. L. Chandani, T. Hagiwara, Y. Suzuzki, *et al.*, Jpn. J. Appl. Phys. **27**, L729 (1988).
2. A. D. L. Chandani, E. Gorecka, Y. Ouchi, *et al.*, Jpn. J. Appl. Phys. **35**, 5054 (1996).
3. V. Laux, N. Isaert, V. Faye, and H. T. Nguen, Liq. Cryst. **27**, 81 (2000).
4. P. Mach, R. Pindak, A. M. Levelut, *et al.*, Phys. Rev. Lett. **81**, 1015 (1998).
5. K. Hiraoka, A. Taguchi, Y. Ouchi, *et al.*, Jpn. J. Appl. Phys., Part 2 **29**, L103 (1990).
6. K. Hiraoka, Y. Takanishi, H. Takezoe, *et al.*, Jpn. J. Appl. Phys. **31**, 3394 (1992).
7. I. Musevic and M. Skarabot, Phys. Rev. E **64**, 051706 (2001).
8. H. Takezoe, J. Lee, A. D. L. Chandani, *et al.*, Ferroelectrics **114**, 187 (1991).
9. D. Schlauf, Ch. Bahr, and H. T. Nguen, Phys. Rev. E **60**, 6816 (1999).
10. M. Fukui, H. Orihara, Y. Yamada, *et al.*, Jpn. J. Appl. Phys. **28**, L849 (1989).
11. M. Kimura, D. Kang, and C. Rosenblatt, Phys. Rev. E **60**, 1867 (1999).
12. R. B. Meyer, L. Liebert, L. Strzelecki, and P. Keller, J. Phys. Lett. **36**, 69 (1975).
13. S. V. Yablonskiĭ, A. S. Mikhaĭlov, K. Nakano, *et al.*, Zh. Éksp. Teor. Fiz. **120**, 109 (2001) [JETP **93**, 94 (2001)].
14. L. A. Beresnev and L. M. Blinov, Ferroelectrics **33**, 129 (1981).
15. S. V. Yablonskii, N. Oue, H. Nambu, *et al.*, Appl. Phys. Lett. **75**, 64 (1999).
16. S. Uto, M. Ozaki, and K. Yoshino, Appl. Phys. Lett. **74**, 117 (1999).
17. I. Kraus, Ch. Bahr, I. V. Chikina, and P. Pieranski, Phys. Rev. E **58**, 610 (1998).
18. P. Pieranski, L. Beliard, J.-Ph. Tournellec, *et al.*, Physica A (Amsterdam) **194**, 364 (1993).
19. H. F. Gleeson, L. Baylis, W. K. Robinson, *et al.*, Liq. Cryst. **26**, 1415 (1999).
20. Y. Takanishi, K. Hiraoka, V. K. Agrawal, *et al.*, Jpn. J. Appl. Phys. **30**, 2023 (1991).
21. N. M. Shtykov, J. K. Vij, R. A. Lewis, *et al.*, Phys. Rev. E **62**, 2279 (2000).
22. J. Hou, J. Schacht, F. Giesselmann, and P. Zugenmaier, Liq. Cryst. **22**, 401 (1997).
23. L. V. Blinov, in *Advances in Liquid Crystals: A Special Volume of Advances in Chemical Physics*, Ed. by J. Vij (Wiley, New York, 2000), Vol. 113, p. 77.
24. S. T. Lagerwall, in *Handbook of Liquid Crystals. Low Molecular Weight Liquid Crystals II*, Ed. by D. Demus *et al.* (Wiley–VCH, New York, 1998), Vol. 2B, p. 515.
25. L. M. Blinov, L. A. Beresnev, and W. Haase, Ferroelectrics **174**, 221 (1995).
26. T. Carlson, F. M. Leslie, and N. A. Clark, Phys. Rev. E **51**, 4509 (1995).

Journal of Experimental and Theoretical Physics, Vol. 100, No. 2, 2005, pp. 428–439.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 127, No. 2, 2005, pp. 478–490.

STATISTICAL, NONLINEAR, AND SOFT MATTER PHYSICS

Dipole Resonances of an Ionized Cluster

A. M. Bystrov and V. B. Gildenburg

Institute of Applied Physics, Russian Academy of Sciences, Nizhni Novgorod, 603950 Russia
e-mail: bystrov@appl.sci-nnov.ru

Received July 8, 2004

Abstract—We study the resonance properties of an ionized spherical cluster that interacts with an optical radiation field in terms of linear hot-plasma polarizability models. Based on a generalization that includes spatial dispersion in the well-known Mie problem of the diffraction of a plane wave by a small-size plasma sphere, we calculate the eigenfrequencies, the radiative and collisional damping constants, and the resonance amplitudes of the cluster surface and bulk plasmon fields. The role of collisionless dissipation processes is analyzed in terms of a one-dimensional kinetic model. The latter allows the corresponding damping constants for both types of plasmons to be determined as functions of the characteristic electron–plasma boundary collision frequency. We show that both types of plasmons in certain domains of cluster and external-radiation parameters can undergo a strong resonance that causes both the amplitude of the scattered wave and the absorbed power and the field inside the cluster to increase significantly. © *2005 Pleiades Publishing, Inc.*

1. INTRODUCTION

A number of new interesting effects observed during the breakdown of atomic clusters by intense laser radiation (anomalously strong absorption, the generation of X-ray radiation and multiply charged ions [1–8]) are probably related to resonance phenomena in a bounded plasma [4–6, 8–11]. The theoretical studies of these phenomena in a hot laser-cluster plasma have been based so far mainly on the simplest model of the electrostatic dipole resonance of a dielectric sphere (also known as the geometrical, Mie, or surface-plasmon resonance) [4–6, 8, 12]. This model, which disregards the collisionless and radiative damping of oscillations and the possibility of the resonant excitation of bulk plasmons,[1] is inadequate at low electron collision frequencies or when the cluster radius a is not too small compared to the wavelength of the electromagnetic wave λ_0 (e.g., at the ratios $a/\lambda_0 = 0.03$–0.10 even at the initial cluster breakdown stage in experiments [3, 10]). The shortcomings of this model cannot be properly compensated for, even based on the well-known results of the works in which the bulk and surface plasmons of cold metal clusters were studied (see, e.g., [15, 16] and references therein), because the approaches used are inapplicable to the hot plasma of an optical breakdown. In this paper, we present the results of our analysis that demonstrates the role of basic dissipation processes at the resonances of both types in terms of a linear classical model that includes the spatial dispersion due to the thermal motion of cluster electrons and the radiative terms in the relations defining the cluster polarization response. However, this model, which is valid only in moderately intense fields, predicts the conditions under which a resonant growth of the field in plasma begins and where the appearance of strong nonlinear effects might be primarily expected.

Setting aside the problems of cluster ionization and dynamics of the forming plasma, we will focus our attention on analyzing the linear resonance parameters (eigenfrequencies, damping constants, resonance dipole moments, field amplitude, and absorbed power) of an ionized spherical cluster with given radius and plasma density. Based on the results of [17, 18] that generalize the solution of the well-known Mie problem [12, 19] of the diffraction of a plane wave by a homogeneous dielectric sphere by including spatial dispersion, we initially (Section 2) describe the complex eigenfrequency spectrum for plasmons of different types in the hydrodynamic approximation that disregards the Landau damping. Subsequently, based on the kinetic one-dimensional (capacitor) model described in Section 3, whose resonance spectrum qualitatively corresponds to the complete dipole resonance spectrum of the real three-dimensional object, we roughly estimate the damping constants determined by collisionless absorption, which can be interpreted here as the absorption due to the collisions of electrons with the plasma boundaries. In Section 4, we use the Lorenz approximation of the resonance line shape, which describes the response of any linear oscillator near the resonance for small losses, to analyze the resonance properties of a plasma sphere by taking into account all of the types of losses considered in Sections 2 and 3. We determine the domain of parameters (the sphere radius, the wavelength of the electromagnetic wave, and the thermal electron velocity) where the hitherto disregarded resonances of bulk plasmons cause a larger growth of the

[1] The role of bulk and surface plasmons in forming the one-dimensional nanostructure of an optical breakdown was studied in [13, 14].

field and the absorbed power than does the Mie resonance.

2. SURFACE AND BULK PLASMON RESONANCES: HYDRODYNAMIC APPROXIMATION

The solution of the Mie problem [12, 19] of the diffraction of a plane electromagnetic wave by a dielectric sphere was generalized in [17, 18] to the case where the dielectric is a warm nondegenerate plasma, i.e., a medium with spatial dispersion taken into account in the hydrodynamic approximation.[2] The electromagnetic field in such a medium acquires an additional degree of freedom related to the possibility of the excitation of longitudinal waves (bulk plasmons), which requires setting an additional boundary condition at the plasma boundary. The continuity condition for the normal electric field component that corresponds to the model of the mirror reflection of electrons at the boundary was taken as a condition that complements the well-known continuity conditions for the tangential electric and magnetic field components. This generalization was used in [18] to study (in the collisionless approximation) the multipole electric resonances of a small-size plasma sphere related to the excitation of both surface and bulk plasmons. Below, based on the results of this work and including electron collisions, we analyze the resonance properties of a sphere in the dipole approximation by disregarding the excitation of high electric multipoles and magnetic oscillations. Although the former are capable of resonating in the same frequency range as the dipole oscillations, they prove to be less significant due to their stronger suppression by various internal dissipation mechanisms. The latter actually do not exhibit any resonance properties at all, since their eigenfrequencies correspond to wavelengths shorter than or on the order of the sphere radius and, therefore, are comparable to the radiative damping constants. The fields of the magnetic oscillations inside and outside a plasma in the frequency range of interest are negligible compared to the field found in the dipole approximation.

In the dipole approximation, the solution to the problem of the diffraction (scattering) of a plane electromagnetic wave

$$\mathbf{E}_0 = \mathbf{e}_x E_0 \exp(ik_0 z - i\omega t), \quad \mathbf{H}_0 = \mathbf{e}_z \times \mathbf{E}_0 \qquad (1)$$

($k_0 = \omega/c$, $\mathbf{e}_x$ and $\mathbf{e}_z$ are the unit vectors along the x and z axes) by a homogeneous plasma sphere with radius a centered at the coordinate origin can be expressed in terms of the first-order vector spherical functions $\mathbf{n}_{e11}^{(1,3)}$ and $\mathbf{l}_{e11}^{(1)}$ [19]. In a spherical r, ϑ, φ coordinate system with the polar z axis and the azimuthal angle φ measured in the xy plane from the x axis, we have for the electric field of the scattered wave $\mathbf{E}^{(r)}$ and the field inside the sphere $\mathbf{E}^{(i)}$

$$\mathbf{E}^{(r)} = \frac{3E_0}{2} b^r \mathbf{n}_{e11}^{(3)} \exp(-i\omega t), \qquad (2)$$

$$\mathbf{E}^{(i)} = \frac{3E_0}{2} (b^t \mathbf{n}_{e11}^{(1)} + b^p \mathbf{l}_{e11}^{(1)}) \exp(-i\omega t), \qquad (3)$$

$$\mathbf{n}_{e11}^{(3)} = \frac{1}{k} \nabla \times [\nabla \times \mathbf{r} h_1^{(1)}(k_0 r) f(\vartheta, \varphi)], \qquad (4)$$

$$\mathbf{n}_{e11}^{(1)} = \frac{1}{k} \nabla \times [\nabla \times \mathbf{r} j_1(kr) f(\vartheta, \varphi)], \qquad (5)$$

$$\mathbf{l}_{e11}^{(1)} = \frac{1}{k} \nabla [j_1(k_p r) f(\vartheta, \varphi)]. \qquad (6)$$

Here, $j_1(\xi)$, $n_1(\xi)$, and $h_1^{(1)}(\xi) = j_1(\xi) + in_1(\xi)$ are the first-order Bessel, Neumann, and Hankel spherical functions, respectively; $f(\vartheta, \varphi) = \sin\vartheta\cos\varphi$; $\mathbf{r}$ is the radius vector of the point; $k = k_0\sqrt{\varepsilon}$ and $k_p = k_0\sqrt{\varepsilon}/\beta_T$ are the wavenumbers of the transverse and longitudinal waves, respectively; $\varepsilon = 1 - \omega_p^2/[\omega(\omega + i\nu)]$ is the complex permittivity of a cold plasma; $\omega_p = \sqrt{4\pi e^2 N/m}$ is the plasma (Langmuir) frequency; e and m are the electron charge and mass, respectively; N is the electron (plasma) density; ν is the effective collision frequency between electrons and heavy particles; $\beta_T = \sqrt{3}\, V_T/c$; $V_T = \sqrt{k_B T/m}$ is the thermal velocity; T is the electron temperature; and k_B is the Boltzmann constant. The condition $\beta_T \ll k_0 a$, which is equivalent to the requirement that the Debye length $r_D = V_T/\omega_p$ be small compared to the sphere radius at frequencies $\omega \sim \omega_p$, is assumed to be satisfied. Outside the sphere ($r > a$), the total field is the sum of the fields of the incident (1) and scattered (2) waves; the field $\mathbf{E}^{(i)}$ inside the sphere ($r < a$) is a superposition of the vortex (transverse) field $\mathbf{E}_t \propto \mathbf{n}_{e11}^{(1)}$ and the potential (longitudinal) field $\mathbf{E}_p \propto \mathbf{l}_{e11}^{(1)}$ defined by the equations

$$\nabla \cdot \mathbf{E}_t = 0, \quad \Delta \mathbf{E}_t + k^2 \mathbf{E}_t = 0, \qquad (7)$$

$$\mathbf{E}_p = -\nabla \varphi_p, \quad \Delta \varphi_p + k_p^2 \varphi_p = 0. \qquad (8)$$

The continuity conditions for the field components H_φ, E_ϑ, and E_r at $r = a$ lead to the following expressions

[2] A similar problem for the cold degenerate plasma of a metal cluster was considered in [20].

for the coefficients b^r, b^t, and b^p:

$$b^r = -\frac{F}{F+iG}, \quad b^t = \frac{j_1(k_0a) + b^r h_1^{(1)}(k_0a)}{\sqrt{\varepsilon} j_1(ka)}, \quad b^p = \frac{2(\varepsilon-1)b^t j_1(ka)}{ka j_1'(k_pa)}, \tag{9}$$

where

$$F = (1-\varepsilon) j_1(k_pa) j_1(ka) j_1(k_0a) + \frac{k_pa}{2} j_1'(k_pa) \times \{\varepsilon j_1(ka)[k_0a j_1(k_0a)]' - j_1(k_0a)[ka j_1(ka)]'\}, \tag{10}$$

the primes denote the derivatives with respect to the arguments of the Bessel functions; the expression for G can be derived from the expression for F by substituting $n_1(k_0a)$ for $j_1(k_0a)$:

$$G = F\{j_1(k_0a) \longrightarrow n_1(k_0a)\}. \tag{11}$$

The coefficients b^t and b^p determine, respectively, the amplitudes of the transverse and longitudinal fields in a plasma, while the coefficient b^r determines the dipole moment of the sphere,

$$p = -3ib^r E_0/2k_0^3, \tag{12}$$

and the total dipole scattering cross section,

$$\sigma = 6\pi|b^r|^2/k_0^2. \tag{13}$$

In the absence of collisions, the dipole resonance frequency spectrum for the sphere is determined by the condition

$$G(k_0a, k_pa, \varepsilon) = 0. \tag{14}$$

If this condition is satisfied, then the coefficient b^r, the dipole moment, and the scattering cross section reach their maximum absolute (resonance) values,

$$b^r = -1, \quad p = p_{\text{res}} = \frac{3iE_0}{2k_0^3}, \quad \sigma = \frac{6\pi}{k_0^2}, \tag{15}$$

which are the same for all types of dipole resonances and that do not depend on the radius a. For $k_0a \ll 1$, the resonance condition can be approximately written as the equation

$$(1-\varepsilon) j_1(k_pa) - \frac{k_pa}{2} j_1'(k_pa)(\varepsilon+2) = 0, \tag{16}$$

that defines the resonance frequencies ω_{res} of the whole series of bulk plasmons and one surface plasmon. The bulk plasmon resonances (i.e., the plasma resonances proper, the resonances of standing plasma waves) occur for certain relationships between the wavelength of the plasma wave and the sphere radius; their frequencies ω_n lies at $\varepsilon > 0$:

$$\omega_n^2 = \omega_p^2 + 3\alpha_n^2 V_T^2/a^2, \tag{17}$$

where $\alpha_n = k_p(\omega_n)a$ is a coefficient that depends on the resonance number $n = 1, 2, 3, \ldots$ ($\alpha_1 = 5.76$, $\alpha_2 = 9.09$, $\alpha_n(n \gg 1) \approx \pi(n+1)$). At fixed field frequency, Eq. (17) defines a set of resonance values ω_{pn}^2 (plasma density) whose number in the hydrodynamic approximation used, $\beta_T/k_0a = \sqrt{3}V_T\omega a \ll 1$, can be fairly large, but, in any case, it does not exceed $k_0a/\beta_T - 1$. The charge density ρ in a bulk plasmon has the same spatial distribution as does the potential φ_p:

$$\rho(r, \psi) = -\frac{\Delta\varphi_p}{4\pi} = \frac{k_p^2\varphi_p}{4\pi}, \quad \varphi_p(r, \psi) = -\frac{3}{2} b_p j_1(k_pr) E_0 \cos\psi. \tag{18}$$

Here, $\psi = \arccos(\sin\vartheta\cos\varphi)$ is the angle between the radius vector and the vector of the external electric field. Distributions (18) for the first two bulk plasmons ($n = 1, 2$) are shown in Fig. 1.

The resonance of a surface plasmon (known as the Mie resonance in optics and as the geometrical resonance in the microwave and radio frequency ranges, or the first of the Tonks–Dattner resonances [21]) lies at $\varepsilon < 0$; its frequency ω_s can be approximately determined by solving the corresponding electrostatic problem and depends on the geometrical shape of the objects ($\varepsilon(\omega) = -2$ and -1 for a sphere and a cylinder, respectively). The electric field inside the sphere at this resonance (when $\beta_T \ll k_0a \ll 1$) is nearly uniform, and the charge is concentrated mainly near the boundary. The general resonance condition $G = 0$ allows us to refine the electrostatic formula by applying the wave ($\sim(k_0a)^2 \ll 1$) and thermal ($\sim\beta_T/k_0a \ll 1$) corrections:

$$\varepsilon(\omega_s) = -2 + \delta, \quad \omega_s = \omega_p/\sqrt{3-\delta}, \quad \delta = \frac{3\sqrt{2}\beta_T}{k_0a} - \frac{12}{5}(k_0a)^2 \tag{19}$$

(only the thermal part of the total correction δ can be found from Eq. (16) in which the terms on the order of $(k_0a)^2$ were discarded).

The radiative width γ^r of each resonance line can be calculated as the difference between the frequencies that correspond to the conditions $G = 0$ and F. For the bulk and surface plasmon resonances (below, these are

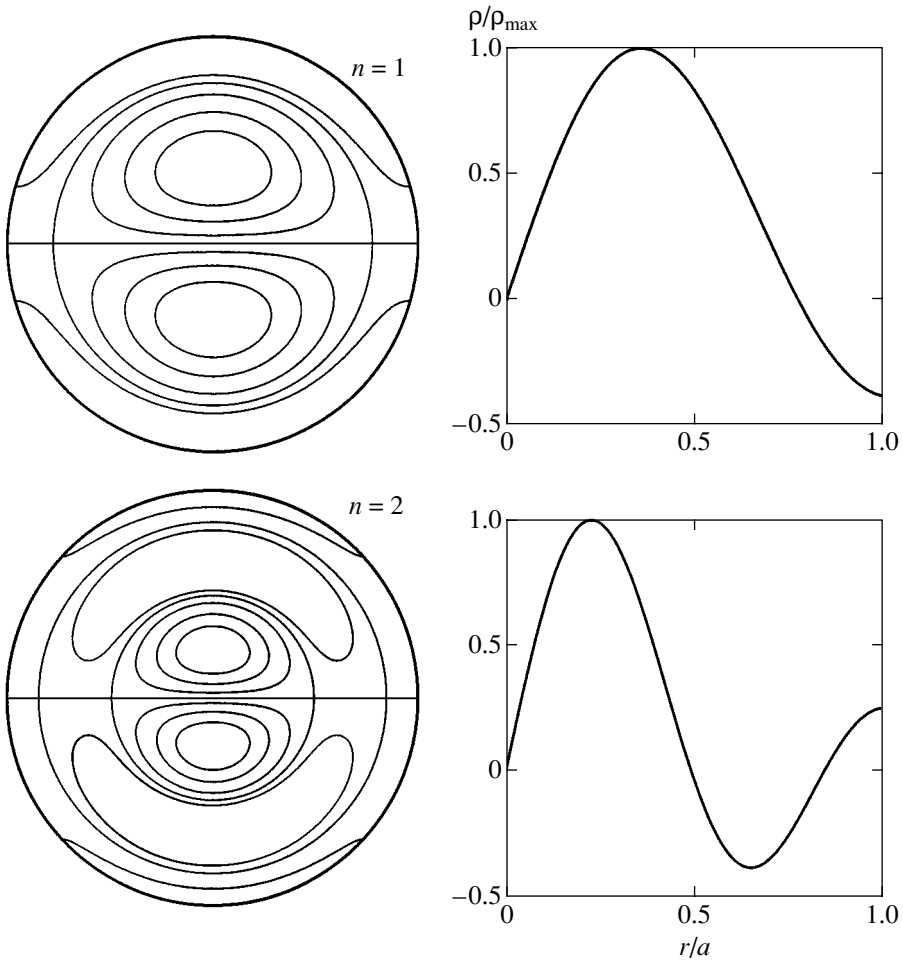

Fig. 1. Isolines in the meridional section and radial charge density profiles in the first two bulk plasmons.

called the bulk and surface resonances for short), we have

$$\frac{\gamma_n^r}{\omega} = \alpha_n^2 \frac{\beta_T^4}{2k_0 a}, \quad \frac{\gamma_s^r}{\omega} = \frac{(k_0 a)^3}{3}. \tag{20}$$

In view of the condition $\beta_T \ll k_0 a \ll 1$, the bulk resonances in the absence of internal losses are well resolved ($\gamma_n^r/(\omega_{n+1} - \omega_n) \sim \beta_T^2 k_0 a \ll 1$) and have a much smaller radiative width than does the surface resonance ($\gamma_n^r/\gamma_s^r \sim \alpha_n^2 (\beta_T/k_0 a)^4 \ll 1$).

3. THE COLLISIONLESS DAMPING OF BULK AND SURFACE PLASMONS: THE PARALLEL-PLATE CAPACITOR MODEL

Generalizing the results obtained in the previous section based on a kinetic description involves significant difficulties attributable to the necessity of satisfying the boundary conditions for the electron velocity distribution function on a spherical (not plane) surface. However, to get a qualitative idea of the role of the kinetic effects disregarded in the hydrodynamic approximation, we can restrict ourselves to an analysis of the resonances of a bounded plasma in terms of the so-called capacitor model by considering the oscillations of a plane plasma layer that partially fills the space between the plates of a parallel-plate capacitor. This one-dimensional model, which was previously studied only in the hydrodynamic approximation [21], can be kinetically described by representing the solution as an expansion in terms of the normal longitudinal waves of a homogeneous plasma.

Let us consider a system (Fig. 2) that is a plane plasma layer of thickness L located inside a parallel-plate capacitor parallel to its plates with the spacing $L_c = L + L_0$ (L_0 is the width of the vacuum gap that is not occupied by plasma). An extraneous source maintains a variable voltage of fixed frequency and amplitude between the plates, $U = U_0 \exp(-i\omega t)$, that generates a variable field $\mathbf{E} = \mathbf{e}_x E(x)\exp(-i\omega t)$ (x axis is perpendicular to the plates) inside the capacitor. Since the sizes of the plates are large compared to their spacing, we may consider the problem as one-dimensional. At the $x = 0$ and L boundaries of the plasma layer, the electron

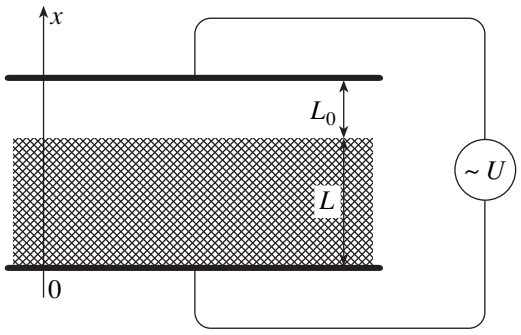

Fig. 2. The one-dimensional (capacitor) model used for the kinetic description of the surface and bulk plasmons of a spherical cluster.

velocity distribution function satisfies the mirror reflection condition

$$f(V_x, V_y, V_z) = f(-V_x, V_y, V_z). \quad (21)$$

Since this boundary condition is automatically satisfied in all nodal planes of the standing longitudinal wave in a homogeneous unbounded plasma, the fields of the normal (free) layer oscillations are functions of the form $\sin k_n x$, where $k_n = n\pi/L$ ($n = 1, 2, 3, \ldots$), in the interval $0 < x < L$. The amplitudes of the x components of the electric induction D and the electric field in plasma $E(x)$ can be represented as Fourier expansions in terms of these normal oscillations:

$$D = \sum D_n \sin k_n x, \quad E = \sum E_n \sin k_n x, \quad k_n = n\pi/L, \quad n = 1, 2, \ldots. \quad (22)$$

Since $D = \text{const} = E_c$, where E_c is the uniform electric field in the vacuum gap, in the one-dimensional problem under consideration and since the amplitudes of the Fourier components of the induction and the field are related by the standard relations of the kinetic theory of waves in a homogeneous plasma, $D_n = \varepsilon_\parallel(\omega, k_n)E_n$, we obtain

$$D_{2n} = 0, \quad D_{2n-1} = \frac{4E_c}{k_{2n-1}L}, \quad (23)$$

$$E = \frac{4}{\pi}E_c \sum_{n=1}^{\infty} \frac{\sin k_{2n-1}x}{(2n-1)\varepsilon_\parallel(\omega, k_{2n-1})}. \quad (24)$$

Here, $\varepsilon_\parallel$ is the longitudinal permittivity, which is defined for the Maxwellian electron velocity distribution by the expression [22]

$$\varepsilon_\parallel(\omega, k) = 1 + \frac{\omega_p^2}{k^2 V_T^2}\left[1 - J_+\left(\frac{\omega}{kV_T}\right)\right], \quad (25)$$

where

$$J_+(x) = x\exp\left(-\frac{x^2}{2}\right)\int_{i\infty}^{x} \exp\left(\frac{\tau^2}{2}\right)d\tau.$$

At small wavenumbers ($kV_T \ll \omega$),

$$\mathrm{Re}\varepsilon_\parallel = 1 - \frac{\omega_p^2}{\omega^2}\left(1 + 3\frac{k^2V_T^2}{\omega^2}\right), \quad (26)$$

$$\mathrm{Im}\varepsilon_\parallel = \sqrt{\frac{\pi}{2}}\frac{\omega\omega_p^2}{(kV_T)^3}\exp\left(-\frac{\omega^2}{2k^2V_T^2}\right). \quad (27)$$

Note that series (24) converges uniformly in the entire interval $(0, L)$, except for its boundary at which the field may be taken to be equal to the limit of (24), $\lim E(x) = E_c$, for $x \longrightarrow +0$ and $x \longrightarrow L - 0$.

The constant E_c in (24) can be expressed using the relation

$$\int_0^L E dx + E_c L_0 = U \quad (28)$$

in terms of the fixed voltage on the capacitor plates:

$$E_c = \frac{U}{L_0}\left[1 + \frac{8L}{\pi^2 L_0}\sum_{n=1}^{\infty}\frac{1}{(2n-1)^2\varepsilon_\parallel(\omega, k_{2n-1})}\right]^{-1}. \quad (29)$$

The condition

$$K(\omega) \equiv 1 + \frac{8L}{\pi^2 L_0}\sum_{n=1}^{\infty}\frac{1}{(2n-1)^2\varepsilon_\parallel(\omega, k_{2n-1})} = 0 \quad (30)$$

determines the complex eigenfrequency spectrum for the system under consideration. Let us analyze this spectrum in the case of weak spatial dispersion ($V_T/\omega L \ll 1$). Disregarding $\mathrm{Im}\varepsilon_\parallel$ and taking Eq. (26) for $\mathrm{Re}\varepsilon_\parallel$, we obtain an equation from (30) using the well-known identity

$$\sum_{n=1}^{\infty}\frac{1}{(2n-1)^2 - x^2} = \frac{\pi}{4x}\tan\frac{\pi x}{2} \quad (31)$$

that defines the real parts of the eigenfrequencies:

$$1 + \frac{L}{\varepsilon L_0}\left(1 - \frac{2}{\pi x}\tan\frac{\pi x}{2}\right) = 0, \quad (32)$$

where $x = \sqrt{\varepsilon/3}\,\omega^2 L/\pi V_T\omega_p$. At $L \sim L_0$, this equation has a set of roots ω_n in the range $0 < \varepsilon \ll 1$, where it takes

the form $\tan(\pi x/2) = \pi x/2$, and one root ω_s defined by the condition $\varepsilon = -L/L_0$ in the range $\varepsilon < 0$, where $\tan x = i\tanh|x|$, $|x| \gg 1$. For the real eigenfrequencies, we have

$$\omega_n^2 = \omega_p^2 + 3\pi^2\left[(2n+1)^2 - \frac{8}{\pi^2}\right]\frac{V_T^2}{L^2}, \quad n = 1, 2, 3, \ldots, \tag{33}$$

$$\omega_s^2 = \frac{\omega_p^2}{1 + L/L_0}. \tag{34}$$

Equation (33) defines the resonance spectrum of one-dimensional bulk plasmons at fixed voltage U. It is approximately valid only for a limited number of oscillation types ($\varepsilon(\omega_n) \ll 1$). Equation (34) defines the surface plasmon frequency in our one-dimensional model (it is lower than the plasma frequency and depends on the geometry of the object, i.e., on the relationship between the lengths L and L_0). When the frequency ω of the extraneous source coincides with one of the eigenfrequencies (33) and (34), the real part of the function $K(\omega)$ becomes equal to zero, while the field amplitude E_c reaches its resonance maximum

$$E_c^{\max} = \frac{U}{L_0 \mathrm{Im}K(\omega_{\mathrm{res}})}.$$

The line width γ^k of each resonance determined by kinetic losses can be calculated as the difference between the nearest real frequencies for which $\mathrm{Re}K(\omega) = 0$ and $\mathrm{Re}K(\omega) = \mathrm{Im}K(\omega)$. However, this calculation can no longer be performed using the simplified expression (26), because the high-order terms for which $k_{2n-1}V_T/\omega \sim 1$ make a major contribution to the infinite sum that, according to (30) determines $\mathrm{Re}K(\omega)$. An analysis based on the general expression (25) for $\varepsilon_\parallel(\omega, k)$ yields the following results in the case under consideration ($V_T \ll \omega L$, $L \sim L_0$):

(a) for the bulk plasmons of the lowest types ($\varepsilon(\omega_n) \ll 1$)

$$\mathrm{Im}K(\omega_n) = \frac{8gV_T}{\pi\omega_n L_0},$$

$$\frac{\gamma_{n0}^k}{\omega_0} = \frac{12g}{\pi}\varepsilon(\omega_n)\left(\frac{2V_T}{\omega_n L}\right)^3 \approx 9g\pi\left[(2n+1)^2 - \frac{8}{\pi^2}\right]\left(\frac{2V_T}{\omega_n L}\right)^5; \tag{35}$$

(b) for the surface plasmon

$$\mathrm{Im}K(\omega_s) = \frac{2gV_T}{\pi\omega_s L_0}, \quad \frac{\gamma_s^r}{\omega_s} = \frac{gV_T L}{\pi\omega_s L_0(L + L_0)}. \tag{36}$$

Here, $g \approx 1.07$ is a coefficient that is calculated numerically. It should be noted that the Landau damping constant [22]

$$\gamma_L(k) = \sqrt{\frac{\pi}{8}}\frac{\omega^4}{(kV_T)^3}\exp\left(-\frac{\omega^2}{2(kV_T)^2}\right) \tag{37}$$

for the wavenumbers $k(\omega_n) \approx (2n+1)\pi/L$ that correspond to the resonances of the bulk plasmons of the lowest types is exponentially small; the line width (35) for them is determined by the damping at the high-order spatial harmonics of expansion (24). The high-order resonances broaden with increasing number n and parameter $V_T/\omega L$ faster than that given by Eq. (35) and are completely suppressed under the condition $k(\omega)V_T/\omega > 0.2$, i.e., starting from $nV_T/\omega L \approx 0.03$ for which the Landau damping ceases to be exponentially small and becomes comparable to the separation between the neighboring lines. The line width γ_n^k for the nth-order bulk plasmon in the entire domain of its existence can be approximated by the sum

$$\gamma_n^k \approx \gamma_{n0}^k + \gamma_L(k_{2n+1}). \tag{38}$$

Note also that the damping described by Eqs. (35) and (36) may be considered to be the result of electron collisions with the plasma boundaries. In this approach, the damping rate is determined by the energy conversion rate of the ordered (oscillatory) electron motion at the time of collision with the boundary into thermal energy. In this case, the line width γ_s^k for the surface plasmon at $L_0 \sim L$ is on the order of the characteristic electron–boundary collision frequency $\nu_b = V_T/L$. The small ratio of the field amplitude at the boundary to the field amplitude at the center of the plasma layer $\gamma_{n0}^k \sim \omega(V_T/\omega L)^5$ is responsible for the much smaller line width of the bulk plasmons of the lowest types.

Figure 3 shows the results of our numerical calculations of the resonance line widths for the surface plasmon and several first bulk plasmons as a function of the parameter $V_T/\omega L$. The calculations were performed using general formulas (25), (29), and (30) for $L/L_0 = 2$, which corresponds the same surface plasmon resonance frequency, $\omega_s/\omega_p = 1/\sqrt{3}$, as that for a spherical cluster. Curves *1–3* for the bulk plasmons are in good agreement with approximation (38). Their initial (gently sloping) parts are described by Eq. (35); as the thermal velocity increases, the passage to a faster increase in damping ($\gamma_n^k \approx \gamma_L(k_{2n+1})$) initially takes place, and, subsequently, the resonance is completely suppressed. The line width for the surface plasmon (curve *S*) fol-

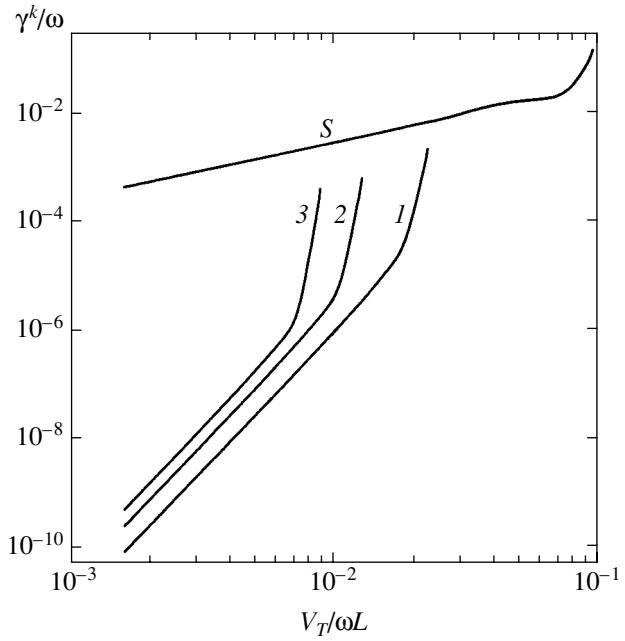

Fig. 3. Kinetic (collisionless) damping constant γ^k versus $V_T/\omega L$ for the first three bulk (curves *1–3*) plasmons and the surface (curve *S*) plasmon at $L/L_0 = 2$.

lows Eq. (36) up to $V_T/\omega L \sim 0.1$, starting from which its increase also becomes faster.

4. THE TOTAL LINE WIDTH: THE RESONANCES OF THE FIELD, THE DIPOLE MOMENT, AND THE ABSORBED POWER

The total line width (damping constant) γ^t of each dipole resonance of a spherical cluster is the sum of the radiative width γ^r found in Section 2 and the width γ^i determined by the internal losses that consist primarily of the losses due to the collisions of electrons with heavy particles and collisionless (kinetic) absorption:

$$\gamma^t = \gamma^r + \gamma^i, \quad \gamma^i = \gamma^\nu + \gamma^k. \tag{39}$$

In the strong-resonance case of interest ($\gamma^i \ll \Delta\omega$, where $\Delta\omega$ is the interval between the neighboring lines), each of the terms γ^r, γ^ν, and γ^k can be calculated independently, i.e., by assuming that the corresponding loss mechanism is unique. The term γ^ν, which includes the collisional losses, can be easily calculated from the general characteristic equation that defines the complex oscillation eigenfrequency. In the absence of other types of losses, this equation for any of the plasmons considered can be written as

$$\varepsilon(\omega) = 1 - \omega_p^2/\omega(\omega + i\nu) = \varepsilon_{\text{res}},$$

where ε_{res} is a real constant (that depends on the cluster parameters and the type of oscillations). It thus follows that the imaginary part of γ^ν for all resonances is the same and equal to $\gamma^\nu = \nu/2$.

The kinetic damping constant γ^k can be roughly estimated using the results of the previous section by taking a plasma layer of thickness L equal to the cluster diameter $2a$ as the spherical cluster model. For this choice of the thickness L, the locations of the bulk plasmon resonance lines in both (plane and spherical) models prove to be close: the frequency differences are about 20% for the first plasmon ($n = 1$), rapidly decrease with increasing number n, and completely disappear in the limit $n \longrightarrow \infty$. As was noted above, the surface plasmon frequencies closely coincide ($\varepsilon = -2$, $\omega_s = \omega_p/\sqrt{3}$) under the condition $L = 2L_0$. The constants γ_s^k and γ_n^k are calculated in this model using Eqs. (35)–(38) and can be expressed in terms of the characteristic electron–cluster boundary collision frequency $\nu_b = V_T/2a$. As we noted above, with regard to the surface plasmon damping, the collisions with the boundary play the same role as the collisions with heavy particles: the damping constant is $\gamma_s^k \approx (4/3\pi)\nu_b$, so the internal losses for this plasmon are roughly determined by the total collision frequency: $\gamma^i \approx (\nu + \nu_b)/2$ (see also [16, 23]). However, the kinetic damping constant for the bulk plasmons proves to be of a higher order of smallness:

$$\frac{\gamma_n^k}{\omega} \approx 9\pi(2n+1)^2\left(\frac{\nu_b}{\omega}\right)^5.$$

The ratios of γ^k defined by Eq. (35) and (36) to the corresponding radiative damping constants γ^r for both types of plasmons differ only by a constant factor and are completely determined by the parameter $\beta_T/(k_0a)^4$:

$$\frac{\gamma_s^k}{\gamma_s^r} = \frac{2}{\pi\sqrt{3}}\frac{\beta_T}{(k_0a)^4}, \quad \frac{\gamma_n^k}{\gamma_n^r} \approx \frac{\gamma_s^k}{8\gamma_s^r}. \tag{40}$$

The resonance curves $b^r(\omega)$ that characterize the frequency dependence of the linear polarization response of a cluster in an external field have a Lorenz shape in the immediate vicinity of each resonance ω_{res} in which, apart from the radiative losses, the internal losses can also be taken into account in an obvious way under the condition $\gamma^t \ll \Delta\omega$:

$$b^r(\omega) = -\frac{\gamma^r}{i(\omega - \omega_{\text{res}}) + \gamma^r + \gamma^i}. \tag{41}$$

The excitation coefficient in this formula (the numerator of the fraction, or the oscillator strength in optical

terminology) coincides with the radiative width γ^r, so $b^r(\omega_{res})$ reaches its maximum absolute value (12) determined by radiative losses only under the condition $\gamma^i \ll \gamma^r$. For the surface resonance, this condition imposes a constraint on the total collision frequency of electrons with heavy particles and the cluster boundary: $(\nu + \nu_b)/\omega \ll (k_0a)^3$. At relatively large k_0a, in particular, at $a/\lambda_0 = k_0a/2\pi = 0.1$, the radiative losses for this resonance are significant (and $|b_r(\omega_s)| \sim 1$) up to $(\nu + \nu_b)/\omega \sim 0.2$–$0.3$. However, for $\beta_T \gg (k_0a)^4$, the radiative losses play a minor role (even at low ν), and the amplitude of the dipole moment at the surface resonance decreases to $|b^r(\omega_s)| = \gamma^r/(\nu + \nu_b) \ll 1$. The bulk resonances with a very small radiative width, $\gamma_n^r \ll \gamma_s^r$, are suppressed more strongly by internal losses and play a lesser role in the scattered radiation: in the range $\gamma_n^r \ll \gamma^i \ll \Delta\omega_n$, the corresponding peaks of the dipole moment at the same collision frequencies are much lower than those at the surface resonance. In particular, at $k_0a/2\pi = 0.1$ and $\beta_T = 0.03$, the lower and upper boundaries of this γ^i range correspond to the collision frequencies $\nu/\omega \sim 3 \times 10^{-5}$ and $\sim 10^{-1}$, respectively (the kinetic losses for the lowest bulk plasmons are negligible in this case: $\gamma_n^k \ll \gamma_n^r$). For $\nu > \Delta\omega_n$, the bulk resonances are virtually suppressed completely.

There is a different relationship between the peak field amplitudes inside the cluster at resonances of different types. In contrast to the resonance of the dipole moment, the field resonance in a certain range of collision frequencies proves to be strongest precisely in the case of bulk plasmon excitation. This is attributable to the oscillating pattern of the radial distribution of the potential field of the plasma oscillations $\mathbf{E}_p$, which causes the dipole moment to decrease greatly at the same field maximum. Determining the electric field amplitudes at the cluster center near the bulk (E_n) and surface (E_s) resonances using the general relations (9)–(11) and generalizing the derived Lorenz dependences of type (41) by taking into account all types of losses (for $\gamma^i \ll \Delta\omega$), we obtain

$$E_n \approx \frac{E_0}{k_0a\beta_T^2}\frac{\gamma_n^r}{\gamma_n^r + \nu/2 + \gamma_n^k}, \tag{42}$$

$$E_s \approx \frac{3E_0}{2(k_0a)^3}\frac{\gamma_s^r}{\gamma_s^r + \nu/2 + \gamma_s^k}. \tag{43}$$

As follows from these expressions, the bulk resonance for negligible internal losses (in the range $\nu/2 + \gamma_n^k \ll \gamma_n^r$) causes a much larger increase in the field amplitude than does the surface plasmon:

$$\frac{E_n}{E_s} = \frac{2(k_0a)^2}{3\beta_T^2} \gg 1.$$

The inequality $E_n > E_s$ holds up to values of the sum $\nu/2 + \gamma_n^k \sim \omega\alpha_n^2 k_0a\beta_T^2$ that greatly exceed the radiative width γ_n^r, i.e., those that lie in the range where the relationship between the maxima of the cluster dipole moment at the resonances of the two types is inverse ($|b_s^r| \gg |b_n^r|$). In the case where $a/\lambda_0 = 0.1$ and $\beta_T = 3 \times 10^{-2}$ considered above, the field at the resonance of the first bulk plasmon is stronger than that at the surface plasmon resonance up to $\nu/\omega \sim 10^{-2}$.

Figure 4 presents the resonance curves that show the dependences of the normalized dipole moment $|b^r| = 2k_0^3|p|/3E_0$ (Figs. 4a and 4b) and field amplitude at the cluster center (Figs. 4c and 4d) on the plasma density N normalized to the critical value of $N_{cr} = m\omega^2/4\pi e^2$. The curves were calculated using the general expressions (9)–(11) at various values of the parameter ν/ω for $k_0a = 0.5$ and $\beta_T = 10^{-2}$.

The kinetic losses in our calculations were taken into account using the results of this and previous sections by substituting $\nu + 2\gamma^k$ for the collision frequency ν in the expression for the complex permittivity ε of a cold plasma. The dependence of the kinetic damping constant γ^k on the ratio $\omega_p^2/\omega^2 = N/N_{cr}$ was determined using Eqs. (35)–(38). Our results show the presence of many resonance peaks related to the excitation of a surface plasmon at $N \approx 3N_{cr}$ ($\omega \approx \omega_p/\sqrt{3}$) and a series of bulk plasmons at $N < N_{cr}$ ($\omega > \omega_p$). The differences in the behavior of both types of resonances described above for the field in plasma and for the dipole moment with increasing electron collision frequency are clearly traceable. The bulk resonances of the dipole moment (Figs. 4a and 4b) are strongly suppressed even at relatively low collision frequencies ($\nu/\omega \sim 10^{-4}$), while the bulk resonances dominate in the pattern of field resonances (Figs. 4c and 4d) up to $\nu/\omega \sim 10^{-2}$. The latter is also illustrated by the curves in Fig. 5, which show the dependence of the field amplitudes at the cluster center at resonances of different types on the electron collision frequency.

In a real experiment, the parameters V_T and ν that determine the internal losses depend significantly on the cluster breakdown conditions. In particular, under tunneling or above-threshold ionization conditions at laser intensities $I \sim 10^{14}$–10^{16} W cm^{-2}, the model in which the thermal electron energy is assumed to be equal in order of magnitude to the energy of the elec-

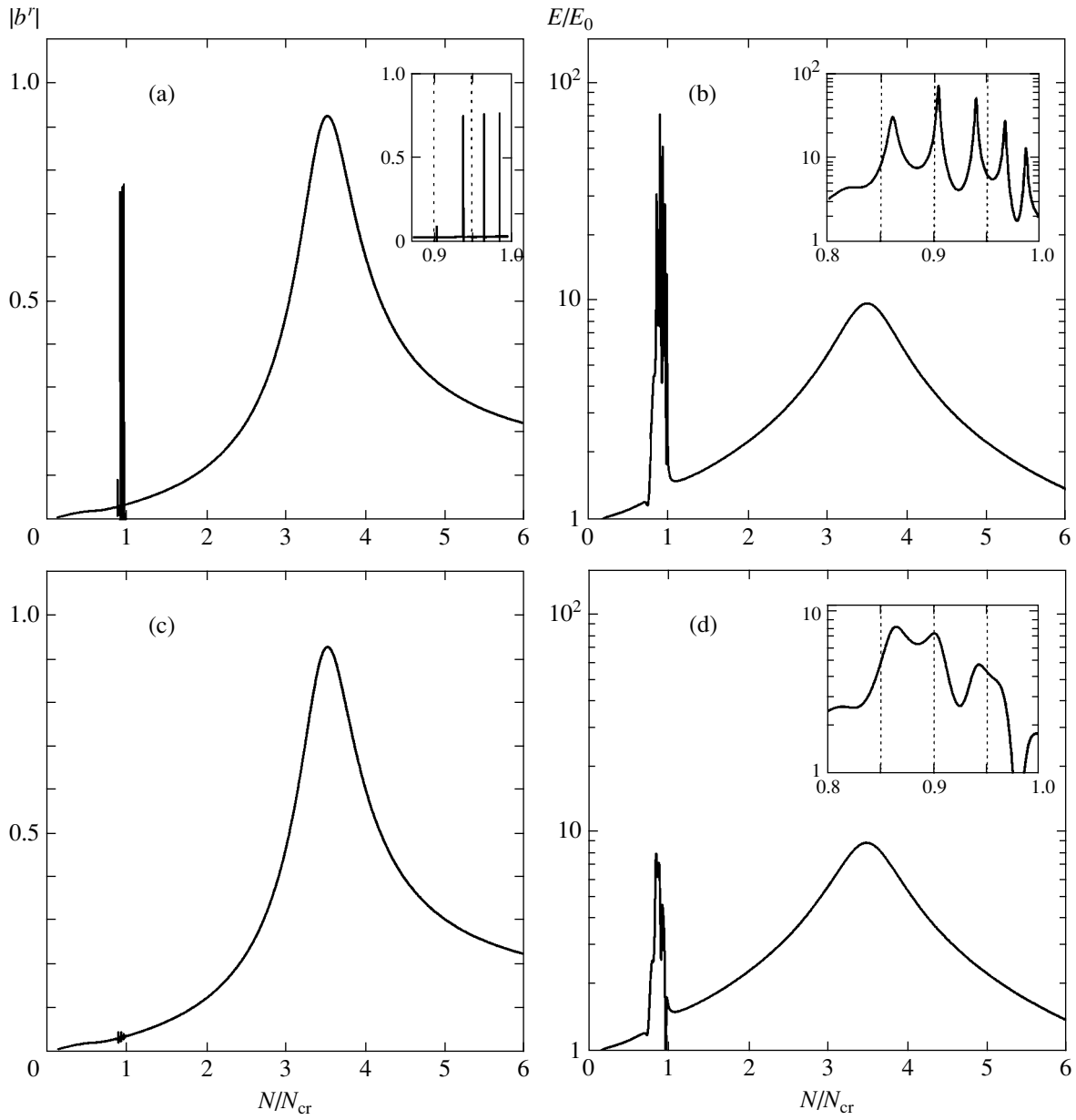

Fig. 4. Normalized dipole moment $|b^r| = 2k_0^3|p|/3E_0$ of the cluster (a, b) and field E/E_0 at the cluster center (c, d) versus plasma density for $k_0a = 0.5$ and $\beta_T = 10^{-2}$ at $\nu/\omega = 0$ (a), 10^{-4} (b), 10^{-3} (c), and 10^{-2} (d).

tron oscillatory motion in an optical field and the collision frequency ν is determined by the Coulomb electron–ion interaction seems realistic enough. The results of our calculations based on this model (under the assumption that $V_T = eE_0/\sqrt{2}\,m\omega$) presented in Fig. 6 for $\lambda_0 = 400$ nm, $a = 30$ nm, and two intensities also show the possibility of resonant field enhancement in plasma at the frequencies of both the surface and bulk plasmons. Since the parameter $k_0a = 0.47$ is relatively large, the line width and the field maximum for the surface resonance in both cases are determined by the radiative losses ($\gamma_s^r \approx 3.5 \times 10^{-2}\omega$), which exceed the losses due to the collisions of electrons with ions and the cluster boundaries by about an order of magnitude. However, the losses at the bulk resonances are attributable mainly to electron–ion collisions, while their number is determined by the kinetic losses (the Landau damping that admits the existence of two bulk plasmons at $I \sim 10^{15}$ W cm^{-2} and only one bulk plasmon at $I = 2 \times 10^{15}$ W cm^{-2}). Interestingly, the field amplitude of the second bulk resonance (corresponding to lower plasma densities) is approximately twice the amplitude of the first one at $I \sim 10^{15}$ W cm^{-2}. This is because its excita-

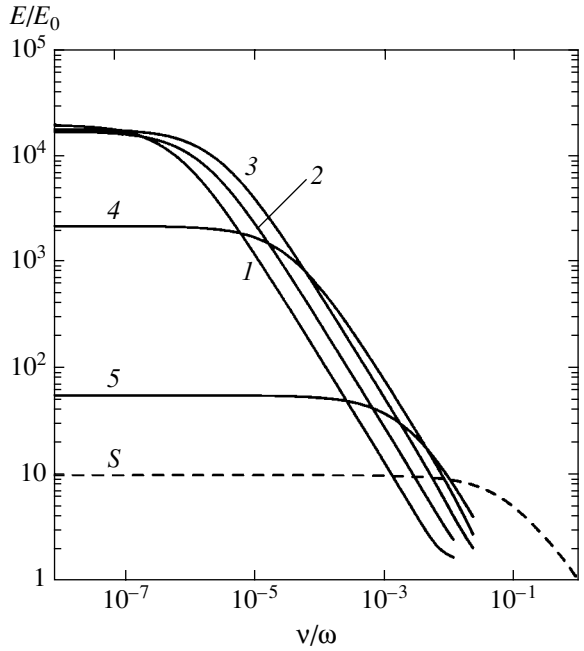

Fig. 5. Field at the cluster center at bulk plasmon resonances (curves *1*–*5*) and at surface plasmon resonance (curve *S*) versus electron collision frequency; $k_0a = 0.5$ and $\beta_T = 10^{-2}$.

tion coefficient (oscillator strength) in the Lorenz formula (42) increases with resonance number n, $\gamma_n^r \sim \alpha_n^2 \approx \pi^2(n+1)^2$, while the internal losses for the first two resonances, which are determined in this case, as was noted above, mainly by electron–ion collisions, are approximately the same.

The mean power (the power of the internal losses) absorbed in plasma, Q^i, can be calculated in the dipole approximation as the difference between the total power

$$Q_\Sigma = \left\langle \frac{d\mathbf{p}}{dt} \cdot \mathbf{E}_0 \right\rangle = \frac{1}{2}\mathrm{Re}(-i\omega \mathbf{p}_0 \cdot \mathbf{E}_0^*),$$

given up by the incident wave field and the power of the dipole radiation

$$Q^r = \frac{2}{3c^3}\left\langle \left(\frac{d^2\mathbf{p}}{dt^2}\right)^2 \right\rangle = \frac{1}{3c^3}\omega^4 \mathbf{p}_0^2$$

(the angular brackets denote time averaging). Based on Eq. (41) that defines the resonance dipole moments, we obtain the resonant absorption Q^i_{res} at $\gamma^i \ll \Delta\omega$:

$$Q^i_{\mathrm{res}} = Q_\Sigma - Q^r = \frac{4Q_0\gamma^r\gamma^i}{(\gamma^r + \gamma^i)^2}, \quad Q_0 = \frac{3\omega E_0^2}{16k_0^3}. \quad (44)$$

Here, Q_0 is the maximum absorbed power that is reached when the radiative and total internal losses are equal ($\gamma^i = \gamma^k + \nu/2 = \gamma^r$) and that depends only on the parameters of the incident wave. In Fig. 7, the power absorbed by the cluster at the resonances of the surface and several first bulk plasmons is plotted against the collision frequency for $k_0a = 0.5$ and $\beta_T = 10^{-2}$. At low

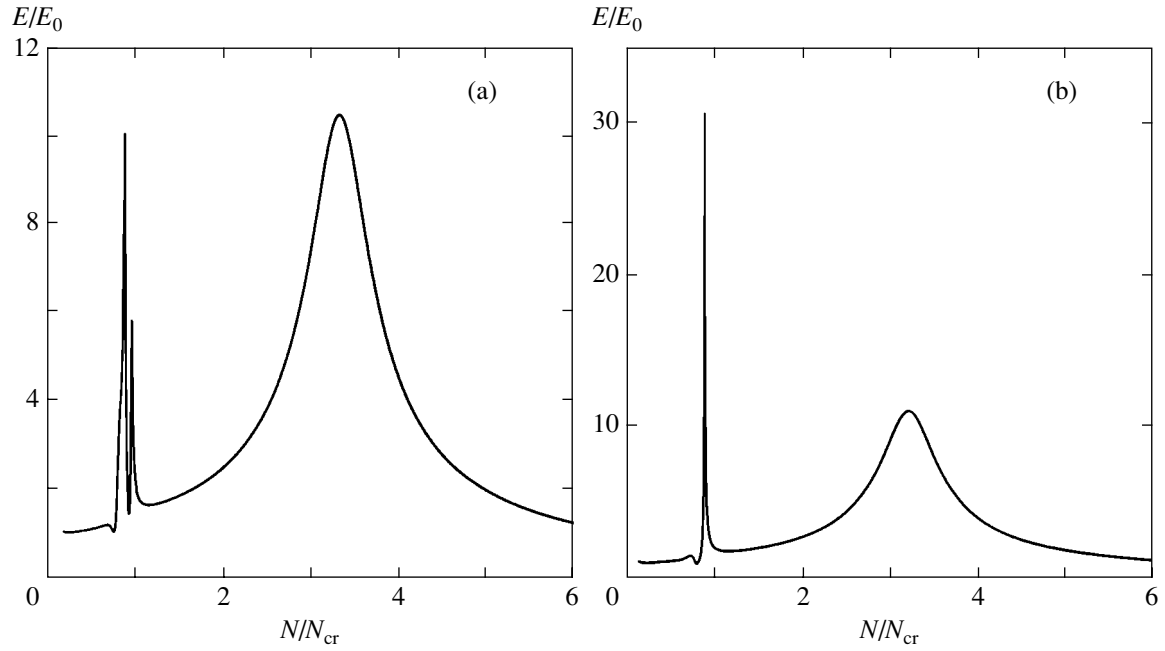

Fig. 6. Normalized field E/E_0 at the cluster center versus plasma density for two laser intensities, (a) $I = 10^{15}$ W cm^{-2} and (b) $I = 2 \times 10^{15}$ W cm^{-2}, under the assumption that the collision frequency and thermal velocity of the electrons are determined by the velocity of their oscillatory motion; the wavelength is $\lambda_0 = 400$ nm, and the cluster radius is $a = 30$ nm.

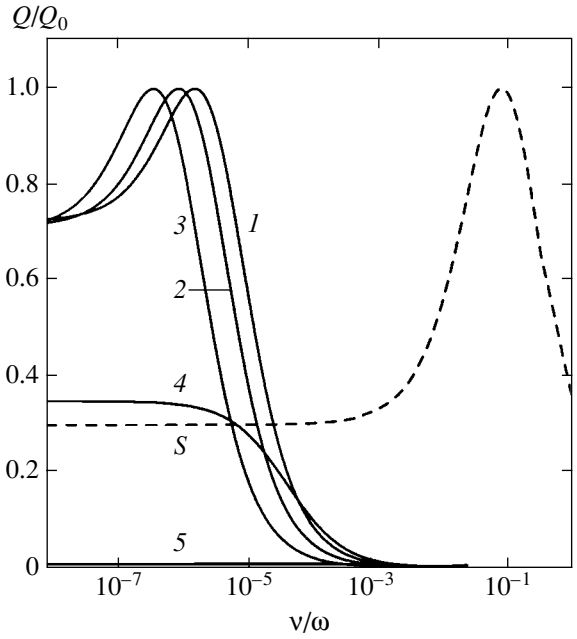

Fig. 7. Power absorbed by the cluster at bulk plasmon (curves *1–5*) and surface plasmon (curve *S*) resonances versus electron collision frequency; $k_0a = 0.5$ and $\beta_T = 10^{-2}$.

ν/ω, the absorption at bulk resonances is stronger than that at surface resonance. As follows from (20), (35), (36), and (39), this relationship is reversed at $\nu/\omega \approx \alpha_n \beta_T^2 k_0 a/\sqrt{6}$.

5. CONCLUSIONS

The hydrodynamic and kinetic models that we considered allowed us to analyze the dipole resonance spectra for a small-size homogeneous spherical cluster ($(k_0a)^3 \ll 1$, $k_0 = \omega/c$, a is the cluster radius) with the inclusion of radiative, collisional, and kinetic (collisionless) losses. Although the cluster plasma in these models was assumed to be nondegenerate, the results obtained also remain qualitatively valid (with the substitution of the Fermi velocity V_F for the mean thermal velocity V_T) for a degenerate plasma (in particular, for metal clusters at normal temperature).

In general, the cluster resonance spectra contain the lines that correspond to the excitation of both a surface plasmon (at a frequency $\omega \approx \omega_p\sqrt{3}$) and a series of bulk plasmons (at frequencies $\omega_n \approx \omega_p$). At a small Debye length, $r_D \sim V_T/\omega_p \ll a$, the radiative ($\gamma_s^r \sim \omega(k_0a)^3$) and kinetic ($\gamma_s^k \sim V_T/a$) damping constants for the surface plasmon (the Mie resonance) greatly exceeds the corresponding values of γ_n^r and γ_n^k for the bulk plasmons:

$$\frac{\gamma_n^r}{\gamma_s^r} \sim \frac{\gamma_n^k}{\gamma_s^k} \sim \left(\frac{r_D}{a}\right)^4 \ll 1.$$

Whereas the inclusion of kinetic losses for the surface plasmon causes the resonance line to broaden by a value on the order of the characteristic electron–cluster boundary collision frequency ($\gamma_s^k \sim \nu_b \sim V_T/a$), the corresponding line broadening for the bulk plasmons is proportional to the fifth power of this frequency ($\gamma_n^k \sim \omega(\nu_b/\omega)^5$). The ratio of the kinetic (γ^k) and radiative (γ^r) damping constants is the same in order of magnitude for both types of resonances:

$$\frac{\gamma^k}{\gamma^r} \sim \frac{\beta_T}{(k_0a)^4} \quad \left(\beta_T \sim \frac{V_T}{c}\right).$$

The maximum intensity of the dipole wave scattering by the cluster reached when the internal losses are negligible compared to the radiative losses is the same for all types of resonances: the total scattering cross section at resonance under these conditions is

$$\sigma_{\max}^r = \frac{8\pi|p_{res}|^2\omega^4}{3cE_0^2} = \frac{6\pi}{k_0^2}.$$

The cross section reaches its maximum,

$$\sigma_{\max}^i = \frac{8\pi Q_0}{cE_0^2} = \frac{3\pi}{2k_0^2},$$

which does not depend on the type of resonance and the cluster size either, when the radiative and internal losses are equal, $\gamma^r = \gamma^k + \nu/2$. The maximum amplitude of the field E at the cluster center, which determines the possibility of the development of various nonlinear processes in plasma, is an important resonance parameter. For negligible internal losses, this parameter is largest for the bulk resonances ($E_n/E_s \sim (a/r_D)^2 \gg 1$).

The ratio of the radiative damping constant γ^r (which also acts as the plasmon–external field coupling parameter) to the total line width, $\gamma^t = \gamma^r + \gamma^k + \nu/2$, essentially determines whether resonances of various types can appear under specific conditions and what their relative role is. The resonances for which this ratio is small are poorly represented in the general spectrum. This can explain the absence of bulk resonances in the absorption and scattering spectra for metal clusters with relatively high electron collision frequencies ($\nu/\omega > 0.1$) at normal temperature that were studied in detail (see, e.g., [16]). However, the bulk resonances in the hot cluster plasma produced by intense laser pulses can be fairly strong. The thermal electron velocities in

such a plasma are higher than or on the order of the ordered velocity of their oscillatory motion. For this reason, at energy flux densities $\sim 10^{15}$ W cm^{-2}, the corresponding Coulomb collision frequencies decrease to $\nu \sim (10^{-2}$–$10^{-3})\omega$, which, as we showed here, admit effective resonant excitation of bulk plasmons. The related resonances of the powers scattered and absorbed by the cluster at given ν are still indistinct (which, of course, makes it difficult to directly observe them in an experiment). However, at comparatively large cluster sizes ($k_0 a \sim 0.2$–0.5), the bulk plasmon resonance causes the electric field amplitude in plasma to increase greatly. Although we calculated this field enhancement in terms of the linear theory, it suggests that strong nonlinear processes (the generation of fast electrons, the electron detachment from deep atomic levels, and their escape from the cluster) can arise even at densities $N \approx N_{cr}$, and not only for the Mie resonance at a factor of three higher values of N. In particular, we may assume that precisely this bulk plasmon excitation effect is responsible for the nonlinear phenomena observed during the X-ray breakdown of clusters in the experiments [3] where the electron density for single atomic ionization, i.e., at the initial stage of the nonlinear process, could not exceed its critical value. The high energy transferred to the cluster plasma in these experiments can probably be explained by the bulk resonance when the internal and radiative losses are approximately equal.

ACKNOWLEDGMENTS

This work was supported by the Russian Foundation for Basic Research (project nos. 02-02-17271 and 04-02-16684).

REFERENCES

1. A. McPherson, T. S. Luk, B. D. Thompson, *et al.*, Phys. Rev. Lett. **72**, 1810 (1994).
2. T. D. Donnelly, T. Ditmire, K. Newman, *et al.*, Phys. Rev. Lett. **76**, 2472 (1996).
3. H. Wabnitz, L. Bittner, A. R. B. de Castro, *et al.*, Nature **420**, 482 (2002).
4. T. Ditmire, T. Donnelly, A. M. Rubenchik, *et al.*, Phys. Rev. A **53**, 3379 (1996).
5. V. P. Kraĭnov and M. B. Smirnov, Usp. Fiz. Nauk **170**, 969 (2000) [Phys. Usp. **43**, 901 (2000)].
6. V. P. Krainov and M. B. Smirnov, Phys. Rep. **370**, 237 (2002).
7. M. B. Smirnov and V. P. Krainov, Phys. Rev. A **69**, 043201 (2004).
8. J. Zweiback, T. Ditmire, and M. D. Perry, Phys. Rev. A **59**, R3166 (1999).
9. H. M. Milchberg, S. J. McNaught, and E. Parra, Phys. Rev. E **64**, 056402 (2001).
10. K. Y. Kim, I. Alexeev, V. Kumarappan, *et al.*, Phys. Plasmas **11**, 2882 (2004).
11. S. V. Fomichev, S. V. Popruzhenko, D. F. Zaretsky, and W. Becker, J. Phys. B **36**, 3817 (2003).
12. G. Mie, Ann. Phys. (Leipzig) **25**, 377 (1908).
13. V. B. Gildenburg, N. A. Zharova, and M. I. Bakunov, Phys. Rev. E **63**, 066402 (2001).
14. N. V. Vvedenskiĭ and V. B. Gil'denburg, Pis'ma Zh. Éksp. Teor. Fiz. **76**, 440 (2002) [JETP Lett. **76**, 380 (2002)].
15. V. V. Kresin, Phys. Rep. **220**, 1 (1992).
16. U. Kreibig and M. Vollmer, *Optical Properties of Metal Clusters* (Springer, Berlin, 1995).
17. A. Yildiz, Nuovo Cimento **30**, 1182 (1963).
18. V. B. Gil'denburg and I. G. Kondrat'ev, Radiotekh. Élektron. (Moscow) **10**, 658 (1965).
19. J. A. Stratton, *Electromagnetic Theory* (McGraw-Hill, New York, 1941; Gostekhizdat, Moscow, 1948).
20. R. Ruppin, Phys. Rev. B **11**, 2871 (1975).
21. P. E. Vandenplas and R. W. Gould, Physica (Amsterdam) **28**, 357 (1962).
22. E. M. Lifshitz and L. P. Pitaevskiĭ, *Course of Theoretical Physics*, Vol. 10: *Physical Kinetics* (Nauka, Moscow, 1979; Pergamon, Oxford, 1981).
23. A. Kawabata and R. Kubo, J. Phys. Soc. Jpn. **21**, 1765 (1966).

Translated by V. Astakhov

Journal of Experimental and Theoretical Physics, Vol. 100, No. 2, 2005, pp. 440–448.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 127, No. 2, 2005, pp. 491–501.

STATISTICAL, NONLINEAR, AND SOFT MATTER PHYSICS

Dynamics of Ionization Processes during the Interaction of a Microwave Pulse with Plasma near Hybrid Resonance

V. I. Arkhipenko[a], V. A. Pisarev[a], L. V. Simonchik[a,*], and E. Z. Gusakov[b]

[a]*Institute of Molecular and Atomic Physics, National Academy of Sciences of Belarus, Minsk, 220072 Belarus*
[b]*Ioffe Physicotechnical Institute, Russian Academy of Sciences, St. Petersburg, 194021 Russia*
**e-mail: simon@imaph.bas-net.by*

Received July 12, 2004

Abstract—We experimentally investigated the dynamics of ionization processes and the formation of a plasma wave channel during the interaction of intense microwave pulses with a magnetized inhomogeneous plasma in the presence of hybrid resonance. The competition between fast electron and slow ionization plasma nonlinearities was studied under conditions when the electron oscillation energy in the wave field was much higher than the atomic ionization energy of the working gas. © *2005 Pleiades Publishing, Inc.*

1. INTRODUCTION

The phenomenon of plasma resonance plays an important role in the ionization formation of plasma under intense microwave and laser emission. According to [1, 2], the significant increase in the projection of the electric field of the electromagnetic wave onto the plasma density gradient direction that takes place at a critical plasma density and the generation of Langmuir waves in this region determine both the ionization front localization and the density growth dynamics. At the same time, the excitation of slow plasma waves in this region also significantly facilitates the conditions for the manifestation of fast electron plasma nonlinearities. In particular, the condition for the equality between the wave phase velocity and the electron oscillation velocity in the wave field is relatively easily satisfied near plasma resonance, which makes the plasma wave breaking possible [3–5]. The breaking effect, which was disregarded in the theoretical analysis [1, 2], is accompanied by the wave energy transfer to very fast electrons. This can change significantly the plasma ionization pattern.

In this paper, we investigate the influence of the competition between fast electron and relatively slow ionization nonlinearities on the formation of a plasma waveguide channel in model experiments on the absorption of an intense microwave pulse near the point of hybrid resonance in a two-dimensional inhomogeneous magnetoactive plasma. The first results of these experiments were published previously [6]. However, no detailed studies were carried out at that time, and the observed phenomena were not explained.

This paper is structured as follows. The experimental setup and the pattern of wave propagation are described in Section 2. The observational results on the dynamics of the fast electron generation, the plasma density growth, and the behavior of the plasma emission after a microwave pulse was switched on are presented in Section 3. These results are then discussed in Section 4, where the increase in the frequency of the wave reflected from the region of hybrid resonance is interpreted.

2. THE EXPERIMENTAL SITUATION

The experiments were carried out at the Granit linear plasma facility [7]. Plasma was produced by electron cyclotron breakdown in a quartz tube *1* (Fig. 1) with an inner diameter of $2r_0 = 1.8$ cm and a length of about 1 m filled with argon at a pressure of 3 Pa and placed in a longitudinal magnetic field about 3 kG in strength. A monotonically decreasing (along the magnetic field) distribution of the plasma electron density

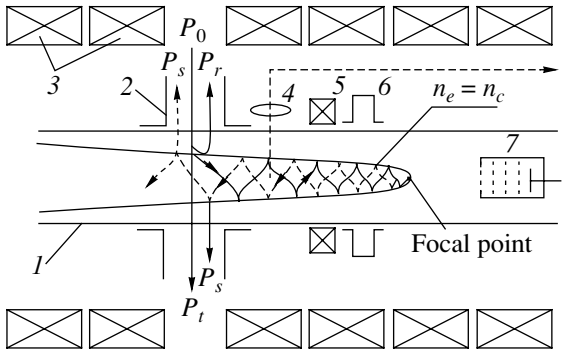

Fig. 1. The experimental setup: P_0, P_r, P_s, and P_t are the incident, reflected, scattered and transmitted waves, respectively; (*1*) quartz tube; (*2*) waveguide; (*3*) magnet coil; (*4*) optical system; (*5*) Rogowski loop; (*6*) cavity; (*7*) analyzer.

n_e was established in the middle part of the tube. This distribution can be fitted by

$$n_e(r, z) = n_c \exp\left(-\frac{z}{l}\right)\left(1 - \frac{r^2}{r_0^2}\right)^{\beta}, \quad (1)$$

where $l = 5$ cm is the plasma inhomogeneity scale length along the magnetic field, $\beta = 4$, and $n_c = \pi m f_0^2/e^2$ is the critical electron density (f is the microwave frequency).

A microwave was applied to the plasma from one side through a 7.2×3.4 cm^2 waveguide *2*; the electric field of the microwave was parallel to the external magnetic field. The typical plasma parameters at the entrance were $n_e < 10^{12}$ cm^{-3} and $T_e \approx 2$ eV. When the density on the tube axis exceeded appreciably its critical value for the input microwave frequency f_0, an oblique Langmuir wave was excited in the plasma predominantly in the form of a fundamental radial Trievelpiece–Gould mode. The dispersion relation for this mode in an inhomogeneous plasma is

$$k_\perp^2 = \left(\frac{\omega_{pe}^2(r, z)}{\omega_0^2} - 1\right)k_\parallel^2,$$

where $k_\parallel$ and $k_\perp$ are the components of the wavevector parallel and perpendicular to the magnetic field. The transparency region for this wave is a dense plasma with a density higher than the critical value, $n_e > n_c$. The near-axis plasma region (Fig. 1) is a plasma waveguide for it with weak axial inhomogeneity. While propagating through this waveguide toward the lower densities, the wave slows down. At the point at which the external magnetic field lines are perpendicular to the surface of the critical density $n_e = n_c$ (focal point), the wave linearly transforms into a "warm" plasma wave, while its field reaches the largest strengths given by

$$E_0 = \sqrt{\frac{2P_0'}{\omega_0}\frac{k_0^3}{3r_D^2 b k_0^3 + 1}} \times \exp\left\{i\int_{-z}^{z}(k_0 + ik_0'')dz' - \frac{k_0}{2b}r^2 - i\omega_0 t\right\} + \text{c.c.}, \quad (2)$$

where $P_0' = \kappa P_0$ is the fraction of the power applied to the plasma P_0 that goes into the excitation of the fundamental radial Trievelpiece–Gould mode ($\kappa \approx 0.2$ [7]), r_D is the Debye length, and k_0 is the component of the wavevector **k** along the external magnetic field. The latter can be determined near the focal point from the equation

$$3r_D^2(k_0 + ik_0'')^2 - \frac{z}{a} - \frac{2}{(k_0 + ik_0'')b} + i\eta'' = 0, \quad (3)$$

where $a \approx l = 5$ cm and $b \approx r_0\beta^{-0.5} = 0.4$ cm are the experimentally determined parameters of the plasma density distribution near the focal point. Thus, the expression for the longitudinal plasma permittivity is

$$\eta = 1 - \frac{\omega_{pe}^2(r, z)}{\omega_0^2}(1 + 3r_D^2k_0^2) + i\eta'' \approx \frac{z}{a} + \left(\frac{r}{b}\right)^2 - 3r_D^2k_0^2 + i\eta''.$$

Using Eq. (2), the electron vibrational energy at the focus can be written as

$$W_\sim = \frac{P_0'}{\pi\omega_0 n_e}\frac{k_0^3 \exp(-2ak_0\nu_{ea}/\omega_0)}{3r_D^2 b k_0^3 + 1}. \quad (4)$$

For $k_0 \approx 40$ cm^{-1} (a typical value for the wavenumber at the focus), we can derive the relation

$$W\ [\text{eV}] \approx 3.2P_0'\ [\text{W}].$$

For $P_0' = 5$ W, the electron vibrational energy, $W_\sim = 16$ eV, is higher than the ionization energy for argon atoms, $E_i = 15.76$ eV. In this case, one might expect very fast ionization in the region of a strong microwave field and, as a result, a displacement of the point of hybrid resonance from the input region. At the same time, according to [8], the plasma wave breaking accompanied by the acceleration of electrons to an energy of 1 keV occurs at these powers.

Various diagnostic techniques were used in the experiments: a 10-cm-band cavity controlled the plasma density distribution; optical diagnostics provided information about the change in plasma emission both along the axis of the quartz tube and along the radius and allowed the microwave absorption region to be determined; a multigrid analyzer controlled the electron distribution function both in the unperturbed plasma and under exposure to a pumping wave; a Rogowski loop was used to change the component of the current through the plasma along the external magnetic field; information about the wave processes in the plasma was extracted by analyzing the spectra of the signals from the waveguide. The experiments were carried out at the following pumping parameters: a frequency of $f_0 = 2840$ MHz, a pulse power of $P_0 = 40$–2000 W, a pulse duration of $t = 0.4$–2.5 μs, a pulse front duration of $t_f \approx 40$ ns, and a repetition frequency of 300 Hz.

3. EXPERIMENTAL RESULTS

The shape of the microwave pumping pulse with a duration of about 2.4 μs and a power of ~50 W in the absence of plasma is shown in Fig. 2a. In the presence of plasma, low-frequency oscillations with a frequency of 20–30 MHz are observed in the discharge tube on the

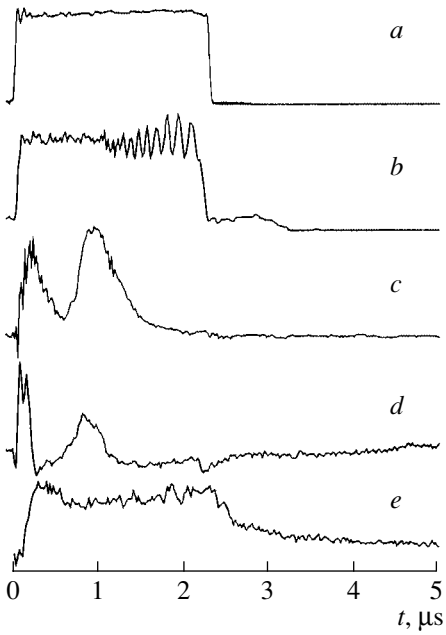

Fig. 2. Oscillograms for the incident (*a*) and scattered (*b*) microwave pulses, the current of the charged particle analyzer (*c*), the signal of the Rogowski loop (*d*), and the intensity of the light at the focus (*e*).

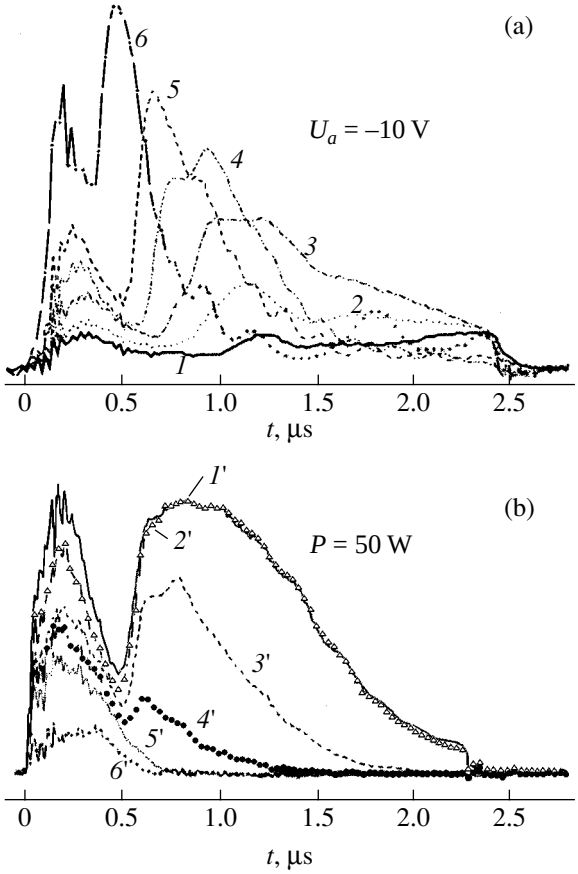

Fig. 3. Oscillograms for the current of the charged particle analyzer at various microwave pulse powers P (a) and retarding potentials U_a (b): P = 15 (*1*), 20 (*2*), 30 (*3*), 50 (*4*), 160 (*5*), 220 W (*6*); U_a = 0 (*1'*), –10 (*2'*), –50 (*3'*), –100 (*4'*), –200 (*5'*), –500 V (*6'*).

microwave signal envelope after the homodyne detection at $t \approx 1$ μs from its beginning (Fig. 2b). Their frequency decreases to 3–5 MHz at the end of the pulse. The low-frequency oscillations are the beats between two waves: the transmitted wave P_t and the wave P_s scattered into the waveguide duct from the plasma (see the wave propagation scheme in Fig. 1).

The multigrid analyzer located behind the focus at a distance of about 25 cm from it on the side of low densities records the electron current (Fig. 2c). The oscillogram of the current pulse has two peaks: one at the beginning of the pulse and the other near the first microsecond. The first peak increases in amplitude with power and is slightly displaced to the beginning of the pulse. The nature of this peak was studied in detail in [8]. Electrons with energies of several hundred electronvolts make a major contribution to it. The second peak also increases with power, narrows, and is displaced to the beginning of the pulse; this displacement is much faster than that for the first peak. The time at which the low-frequency oscillations appear on the oscillogram of the homodyne detector signal is also displaced almost simultaneously with the displacement of the second peak. In general, the low-frequency oscillations develop after the second peak of the electron current is reached.

As we see from Fig. 2d, the signal of the Rogowski loop has two positive peaks. Their displacement dynamics as the microwave pulse power changes is the same as that for the current pulses of the analyzer. Since the time constant of the Rogowski loop is about 40 ns, its signal reflects the derivative of the current pulse of accelerated electrons.

As we see from Fig. 2e, which shows an oscillogram for the current signal of the photomultiplier, the plasma emission increases in intensity near the focus almost simultaneously with the application of a microwave pulse. It should be noted that the light pulses are characterized by a slower decay (several microseconds) after the microwave signal termination than the decay of the electron current (less than 1 μm).

The electron energy. The amplitude and shape of the current pulse of the charged particle analyzer depend on the power of the applied microwave pulse and the retarding potential of the charged particle analyzer. In the experiment, the analyzer was located at a distance of about 25 cm from the focal point on the side of low densities. Figure 3a shows the current pulses of the analyzer at a retarding potential of U_a = –10 V for various powers of the applied microwave pulses. We see that the second peak of the current is rapidly displaced to the beginning of the pulse as the power increases. If the pulse power increases further, then these two peaks merge together at $P \approx 1$ kW. The peaks

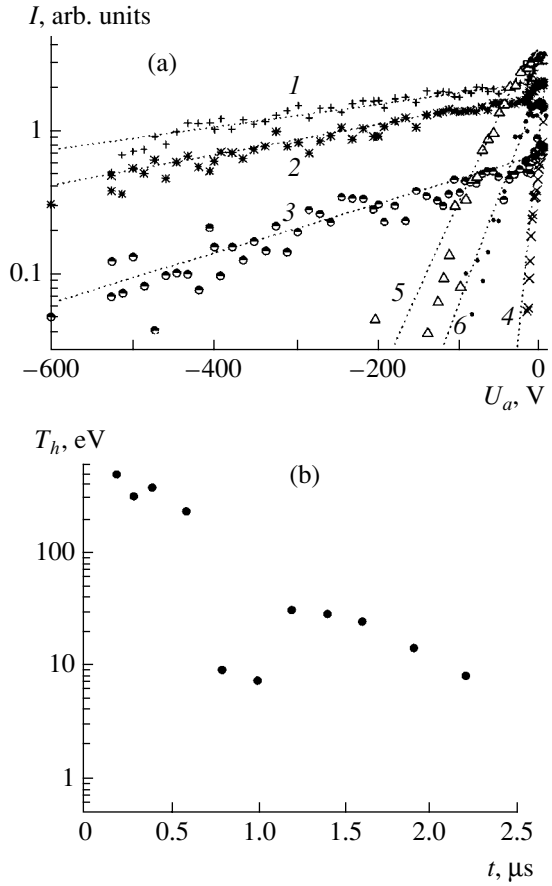

Fig. 4. Current–voltage characteristics (a) and accelerated electron energies (b) at various times: t = 0.2 (*1*), 0.4 (*2*), 0.6 (*3*), 0.8 (*4*), 1.2 (*5*), 1.6 μs (*6*).

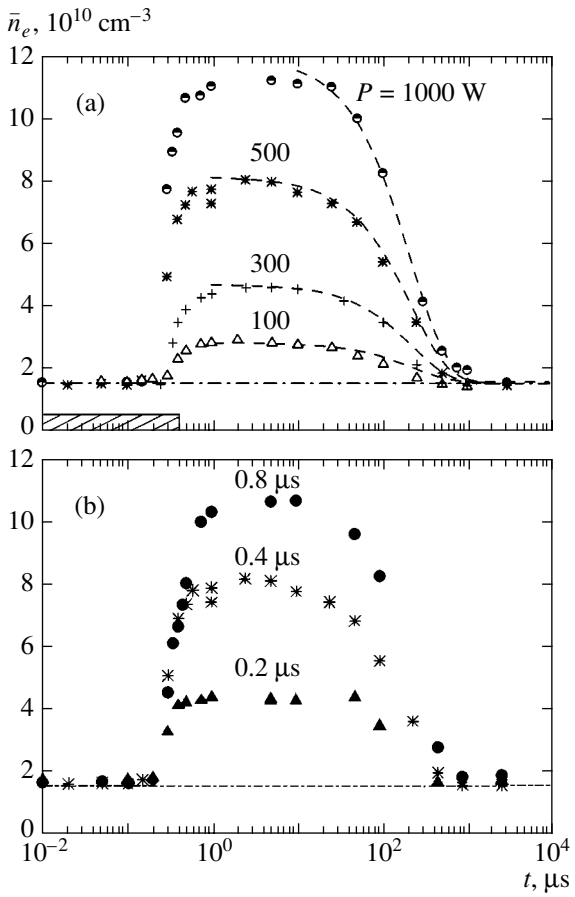

Fig. 5. Time dependences of the mean electron density at the focus at various pumping powers P (a) and various pulse durations (b) for a power of 50 W.

behave differently at various retarding potentials and fixed power, as demonstrated in Fig. 3b for a pulse with P = 50 W. The first peak produced by electrons with energies of several hundred electronvolts decreases by a factor of about 5 at a retarding potential of $U_a \approx -500$ V. The second peak virtually disappears even at $U_a \approx -200$ V. This indicates that the energies of the electrons that contribute to these current peaks differ significantly.

Figure 4a shows the current–voltage characteristics for the charged particle analyzer on a semi-logarithmic scale at various times from the beginning of the pulse at P = 50 W. They have nearly linear segments (dotted lines) at energies $W \gg T_e$ in which the electron energy can be characterized by the effective temperature T_h. We see that the slope of the characteristics from the beginning to the end of the pulse changes nonmonotonically. This can also be seen in the corresponding time dependence of the effective temperature T_h shown in Fig. 4b. It follows from this figure that the first peak of the current is produced by high-energy electrons with an effective temperature of $T_h \approx 400$ eV. The generation of such electrons lasts for 0.5 μs and is then abruptly interrupted; the electron energy at the end of the first microsecond is about 10 eV. The effective electron temperature during the second peak is much lower than that during the first peak, being about 40 eV, and again decreases to 10 eV at the end of the pulse (2 μs). After pulse termination, the suprathermal electrons disappear rapidly, in less than one microsecond. Their disappearance is probably responsible for the presence of a small minimum in the signal of the Rogowski loop (Fig. 2d) after the microwave pulse termination.

The effective temperature of the electrons accelerated near the focus is much higher than the electron energy in the initial plasma, $T_h \gg T_e$, during both the first and second fast electron generation peaks. This allows both peaks to be associated with the Landau damping of the plasma wave, which is definitely [8] and presumably nonlinear in the former and latter cases, respectively (plasma wave breaking).

Electron density dynamics. The electron density was controlled with a 10-cm-band cavity, which allows the density averaged over the plasma volume in the cavity to be determined. Since the experiment was carried out in the regime of repetitive pulses, the densities at various times could be obtained by using a stroboscopic technique.

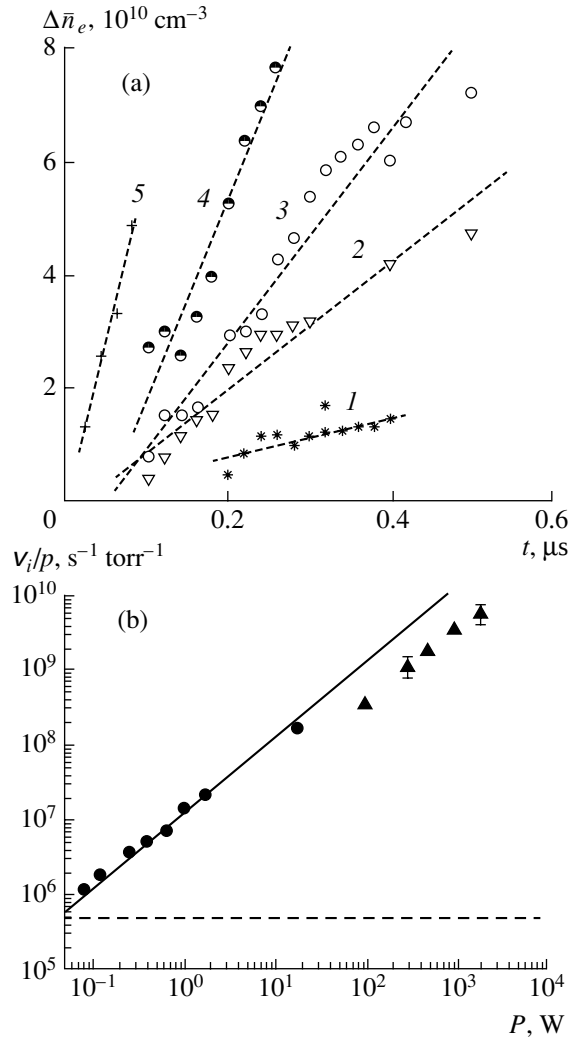

Fig. 6. (a) The increase in density at the focus at the initial time for various pumping powers: 100 (*1*), 300 (*2*), 500 (*3*) W, 1 (*4*), and 2 (*5*) kW; (b) argon ionization rate versus microwave signal power: the filled circles represent the data from [9].

Figure 5 shows time dependences of the electron density at the focus for various microwave pulse powers and durations. We see from Fig. 5a that, depending on the pulse power, an appreciable increase in density begins 200–300 ns after the beginning of the pulse. The pulse duration in this experiment was 0.4 μs and is marked by the hatched rectangle in Fig. 5a. The electron density reaches its maximum in all cases after the pulse termination between the first and tenths microseconds. Subsequently, the plasma slowly relaxes to its initial state. The critical density is indicated in Figs. 5a and 5b by the dash–dotted horizontal lines. As can be inferred from Fig. 5a, the plasma decay finishes at the first microsecond for almost all powers. The dashed lines in Fig. 5a represent the exponential dependences that reflect the decrease in electron density:

$$\bar{n}_e = n_m \exp[-(t-t')\tau_d],$$

where n_m is the maximum electron density for each power, t' is the time at which the density decay begins, and τ_d is the decay constant. In all cases, the density decay constant proved to be the same, about 200 μs.

The maximum density after the pulse depends on the pulse duration (Fig. 5b): the density increases with pulse duration. The duration of the plasma decay phase changes only slightly: from 0.6–0.7 to 1 ms.

Let us consider in more detail the change in density during the initial period of the pulse shown in Fig. 6a for various pumping powers. We see that the density growth rate increases with pulse power. Since these dependences are satisfactorily fitted by a linear function (dashed lines), we can calculate the ionization rate ν_i at the focal point normalized to the gas pressure p in the discharge tube by using them:

$$\frac{\nu_i}{p} \approx \frac{1}{pn_c}\frac{\partial n}{\partial t}.$$

The ionization rates calculated in this way are shown in Fig. 6b as a function of the microwave pulse power. This figure also shows the data from [9] obtained with the same experimental setup for pulse powers up to 10 W. The horizontal dashed line in this figure indicates the ionization rate in the initial plasma, and the solid line represents the linear power dependence, $\nu_i \propto P$.

As we see from Fig. 7, which shows the longitudinal distribution of the mean electron density at various times for an incident microwave pulse power of about 100 W, the density increases not only near the initial position of the resonance, but also far from it. Indeed, the increase in density at the initial time ($t < 0.5$ μs) is localized near the resonance at a distance of z = 20–25 cm from the pumping input point. However, at $t > 0.5$ μs, the density increases mainly in low-density regions of the initial plasma ($z > 25$ cm). This effect can be explained in terms of both the ionization produced by accelerated electrons at the initial phase of the pulse and the plasma channel burning, and the displacement of the resonance region at later phases of the pulse.

Optical emission. Plasma emission spectra near the resonance in the wavelength range 400–500 nm were recorded in the experiment. Using a stroboscopic technique to record them allowed us to trace their change with time since the application of a microwave pulse. At the beginning of the pulse, when electrons with energies of several hundred electronvolts are generated, the neutral argon lines in this part of the spectrum are much weaker than the ion lines, and the spectrum is represented mainly by Ar II lines. The transverse intensity distributions of the Ar II (454.5 nm) spectral line were recorded in various cross sections along the length of the plasma column. The transverse distributions were also recorded at various times by using a stroboscopic technique. These data were used to construct the spatial intensity distributions of the Ar II 454.5-nm line shown in Fig. 8. We see that the peak of the line inten-

sity at the initial time ($t < 0.2$ μs) corresponds to the position of the resonance region. The intensity increases with time, with the emission growing more strongly on the side of low densities; an extended channel of enhanced emission is formed behind the focus at $t > 1$ μs. The size of this channel (the separation between the points at which the intensity accounts for half of the intensity on the channel axis) is 1.5–2 mm across. For comparison, the size of the initial plasma emission region at half intensity reaches 4–6 mm across.

Since the degree of plasma ionization behind the focus does not exceed 0.01%, it is justifiable to use a coronal model to interpret the results of the plasma emission measurements. In this model, atoms and ions are excited from the ground atomic state through collisions with electrons and deexcited through spontaneous emission. According to [10], the excitation functions of argon atomic and ionic lines are at a maximum near their ionization potentials. At electron energies on the order of several hundred electronvolts, they decrease severalfold. That is why the accelerated electrons produce only weak emission far from the focus (at $z >$ 25 cm) at the initial phase of the pulse, which corresponds to the first peak of the current at $t < 0.5$ μs (see Fig. 2c). The growth of the emission in these regions begins with the increase in density and with the displacement of the plasma resonance to this region. Since the electrons at this time have a much lower energy close to the maximum of the line excitation function, they are capable of more effectively exciting and ionizing the argon atoms.

Low-frequency oscillations. Low-frequency oscillations appear on the microwave pulse envelope at a power of about 50 W in the middle of the pulse ($t \approx$ 1 μs, see Fig. 2b). The variations of their frequency with time at various powers of the pumping pulse are shown in Fig. 9. The oscillation frequency is seen to rapidly decrease from 20–25 to 5 MHz. This time evolution of the oscillation frequency is preserved as the power increases, but the onset of the development of oscillations is displaced to the beginning of the pulse.

Oscillations are also observed on the signal of the sounding wave that was applied to the plasma through the same waveguide as the pumping at $f_0 = 2840$ MHz. The frequency of the sounding wave was $f_\zeta =$ 2400 MHz and differed significantly from the pumping frequency, $f_\zeta < f_0$, while its power was several orders of magnitude lower than the pumping power. The signal from the waveguide duct was split into two measuring channels. Rejection filters at the pumping and sounding wave frequencies were placed in the first and second channels, respectively. The homodyne detection signals from both channels were recorded by a digital oscillograph.

Figure 10a shows oscillograms for the signals from the microwave detectors placed in the pumping and sounding wave ducts at frequencies of 2840 and 2400 MHz, respectively. An analysis of these oscillograms indicates that the oscillations begin almost simultaneously, irrespective of the wave frequency. However, the oscillation period for the sounding wave at each given time is a factor of 1.2 to 1.5 longer than that for the pumping wave. This difference increases as

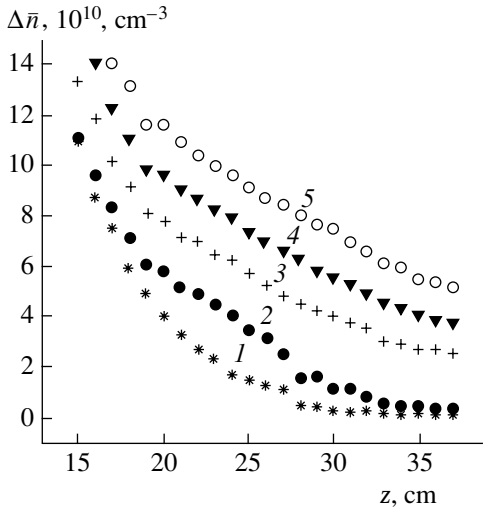

Fig. 7. The axial distribution of the mean electron concentration at various times: initial plasma (*1*), $t = 0.5$ (*2*), 1.1 (*3*), 1.7 (*4*), 2.4 μs (*5*).

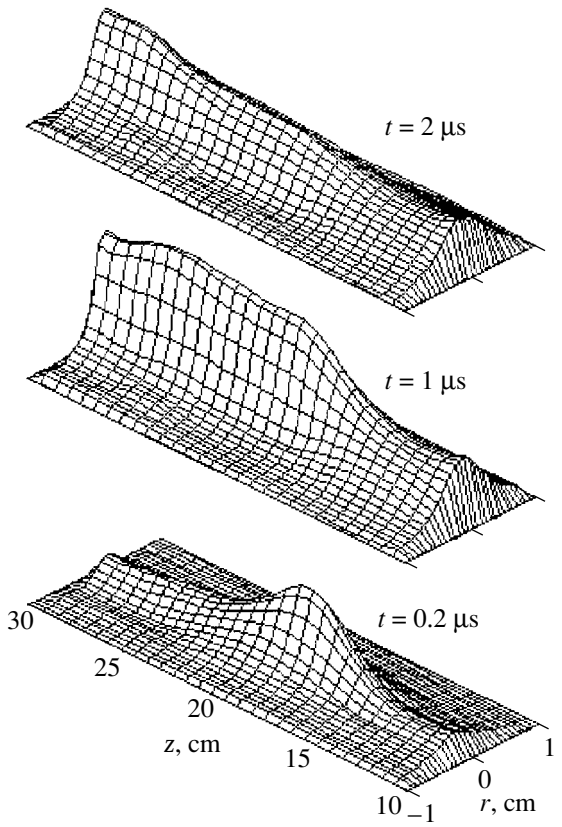

Fig. 8. The spatial intensity distribution of the ArII 454.5-nm line at various times.

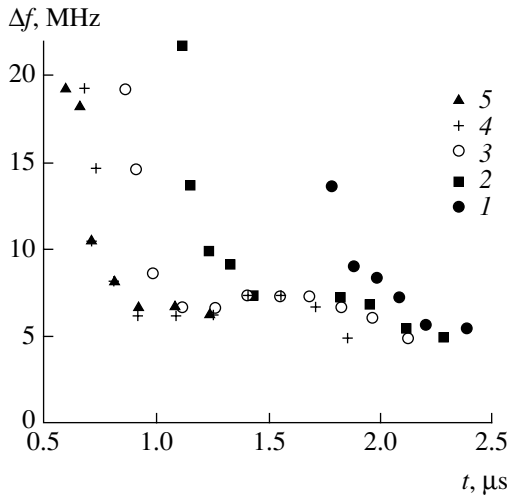

Fig. 9. Variations of the oscillation frequency with time at various microwave pulse powers: 40 (*1*), 60 (*2*), 87 (*3*), 125 (*4*), and 155 W (*5*).

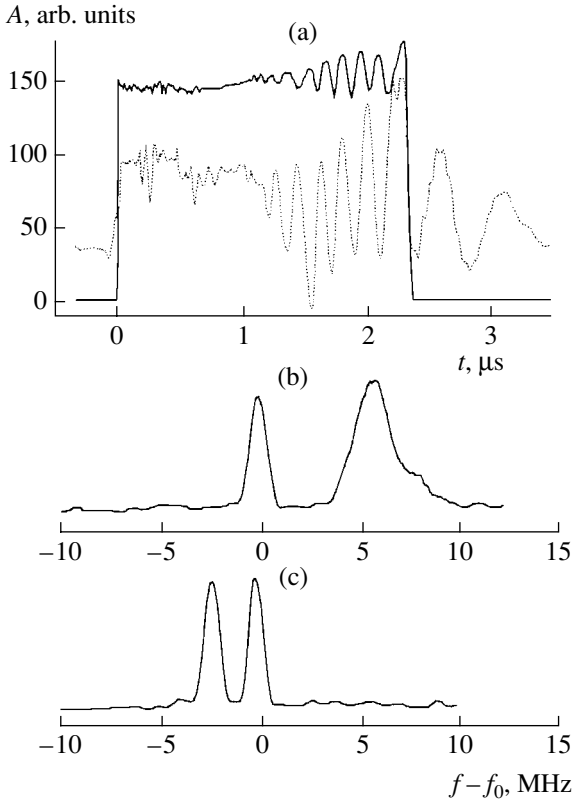

Fig. 10. Oscillograms of the detector signals for frequencies of 2840 (solid line) and 2400 (dashed line) MHz (a) and spectra of the scattered signal during the pulse (b) and after its termination (c).

the frequency difference between the pumping and sounding waves increases. It should be noted that the oscillations on the detector signal in the sounding wave duct are also observed for 1 to 2 μs after the pumping pulse termination, with their period being longer than that during pulse. Like the duration of the oscillations, it depends on the power and duration of the pumping pulse. For example, the oscillation frequency changes from 4 to 1 MHz as the power increases from 50 to 150 W and from 6 to 1 MHz as the pulse duration changes from 1 to 3 μs at a constant power of 100 W.

We studied the spectrum of the signal scattered into the waveguide duct by using an S4-80 pulsed signal spectrum analyzer for a window duration of 0.4 μs. The plasma-scattered pumping and sounding wave signals were found to have higher frequencies than the corresponding frequencies of the applied waves. During a pumping pulse, the scattered signals show up in the form of satellites in the anti-Stokes spectral region of both waves (Fig. 10b). At the same time, the sounding wave spectrum after the pumping pulse termination changes sharply: a satellite with a smaller frequency shift appears in the Stokes region (Fig. 10c).

4. DISCUSSION

As we noted above, the electron densities measured with a 10-cm-band cavity are the values averaged over the cavity volume. To construct the spatial density distribution over the plasma column, the data for the axial distributions of the mean density (see Fig. 7) must be supplemented with the radial electron density distributions in various axial cross sections. The radial intensity distributions of the plasma emission (see Fig. 8) were used as such distributions. We used this method for constructing the spatial electron density distributions previously [8].

Figure 11 shows lines of equal density at various times. Ten shades of gray in the figure are used as the density scale. The lightest shade of gray corresponds to an electron density of 5×10^{11} cm^{-3}; the darkest shade, to the minimum electron density. The derived pictures of the lines of equal density clearly illustrate the formation of an extended narrow plasma channel that is a plasma waveguide for an oblique Langmuir wave. The region of propagation for this wave is the region with $n_e > n_c$ whose boundary advances along the magnetic field in the direction opposite to the density gradient. At a distance of about 45 cm from the input waveguide, which corresponds to the edge of the electromagnet, the external magnetic field weakens sharply. This appears to result in a sharp change of the propagation conditions for the wave at this location and, hence, to its reflection. When the multigrid analyzer was placed in the tube closer to the initial position of the focal point than the edge of the magnet, the wave was reflected from its surface. The reflected wave propagated in the opposite direction in the same nonstationary plasma channel with an increasing plasma density, which must cause its frequency to shift toward the higher frequencies. To some extent, this upward frequency conversion mechanism for the reflected wave is similar to that discussed in [2, 11] and related to the Langmuir wave conversion frequency during ionization at plasma pumping reso-

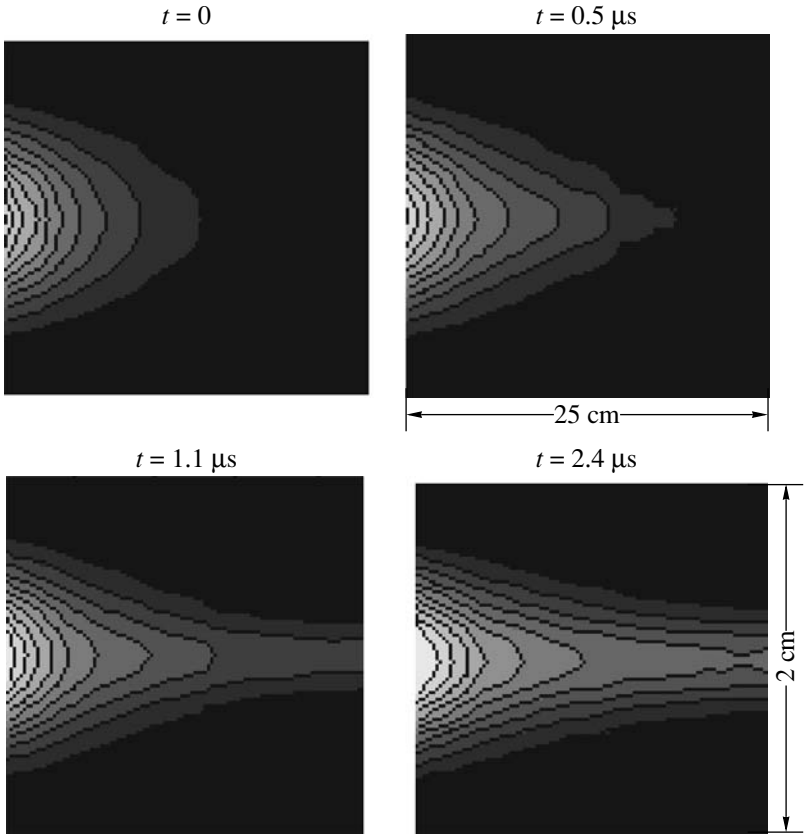

Fig. 11. Lines of equal electron density at various times.

nance. It should be emphasized that the frequency shift of the reflected wave in our experiment reaches 20 MHz and is more than twice the ion plasma frequency near the resonance. This makes it difficult to explain the observed effects in terms of plasma acceleration under a ponderomotive force [12].

Let us consider some of the corollaries of the suggested frequency conversion mechanism. As a wave propagates to and fro in a plasma waveguide with length L, the wave phase shift is $\delta\Phi = -2k_0L$. Since the waveguide is nonstationary, the wavenumber k_0 depends not only on the longitudinal coordinate in accordance with Eq. (3), but also on time. Taking into account this dependence, we can obtain

$$\delta\Phi = -\frac{2L}{b}\frac{n_c(f_c)}{n_e(z,t)-n_c(f_\zeta)}.$$

The expression for the phase factor is

$$\Phi = \omega_0 t - 2\int k_0 dz,$$

hence, we can write for the wave frequency

$$\omega = \frac{d\Phi}{dt} = \omega_p + \frac{d\delta\Phi}{dt},$$

and derive an expression for the frequency difference in the form

$$\delta f_\zeta = \frac{\delta\omega_\zeta}{2\pi} = \frac{L}{\pi b}\frac{n_c(f_\zeta)}{[n_e(z,t)-n_c(f_\zeta)]^2}\frac{dn_e}{dt}. \qquad (5)$$

The change in density at the focus can be represented as

$$n_c = [1+\alpha(t)]n_c(f_0),$$

where the function $\alpha(t)$ describes the change in density with time. Then,

$$\delta f_\zeta = \frac{L}{\pi b}\frac{dn_e/dt}{\left[\alpha + \dfrac{n_c(f_0)-n_c(f_\zeta)}{n_c(f_0)}\right]^2 n_c(f_0)}. \qquad (6)$$

Let us make certain assumptions about the function $\alpha(t)$. We assume (in accordance with Fig. 5a) that the rate of change in electron density at the focus is constant, i.e.,

$$\frac{dn_e}{dt} = \frac{d\alpha}{dt}n_c(f_0) = \text{const}.$$

Then,

$$\alpha(t) \approx \frac{dn_e}{dt}\frac{t-t_*}{n_c(f_0)},$$

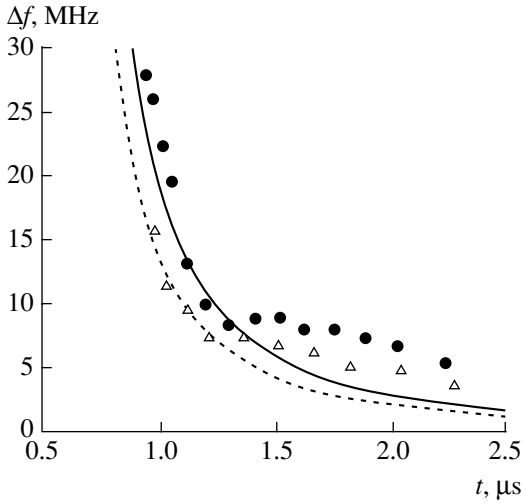

Fig. 12. Time dependences of the frequency shift: experimental (dots) and calculated (curves) data for f_0 = 2840 MHz (circles and solid line) and f_0 = 2400 MHz (triangles and dotted line).

where t_* is the plasma channel burning time. Substituting this expression for $\alpha(t)$ into Eq. (6) yields

$$\delta f_\zeta = \frac{L}{\pi b}\frac{n_c(f_0)}{\left[(t-t_*)\dfrac{dn_e}{dt} + n_c(f_0) - n_c(f_\zeta)\right]^2}\frac{dn_e}{dt}. \quad (7)$$

It follows from Eq. (7) that the oscillation frequency on the sounding duct detector must decrease both with time as $1/(t - t_*)^2$ as with increasing frequency difference $f_0 - f_\zeta$. The results of our calculations of the oscillation frequency performed using Eq. (7) at a power of about 100 W are presented in Fig. 12. We see that there is good agreement between the calculated and experimental values for both the pumping and sounding waves.

5. CONCLUSIONS

Thus, the competition between fast (electron) and slow (ionization) nonlinear mechanisms is observed near hybrid resonance during the interaction of an intense electromagnetic pulse with an inhomogeneous magnetized weakly ionized plasma. Fast electron nonlinearities, which cause wave breaking and strong electron acceleration near the point of resonance at the initial phase of the pulse, dominate at the first stage. Thus, for example, this occurs over a period of less then 0.5 μs at a power of about 50 W. At longer times ($t >$ 0.5 μs), a narrow homogeneous plasma channel is formed through violent ionization attributable to the electron oscillations in the wave field. In this channel, the wave amplitude is below the breaking threshold due to the increase in group velocity and collisional damping; as a result, the electron acceleration effect is temporarily suppressed. When the plasma channel is burned up to the boundary of the system, the conditions for the formation of a resonant field structure, wave slowdown, increase in wave amplitude, and, hence, intense interaction with electrons again arise for a short time. The reflection of the intense wave is observed immediately after the recurrent peaks of accelerated electrons, at $t > 1$ μs. The propagation of the incident and reflected waves in a plasma waveguide with an increasing electron density causes a rapid change in the phase shift of the reflected wave and, as a result, in its significant upward displacement in frequency. Analysis of the wave reflection and propagation using our simple model confirms this picture.

ACKNOWLEDGMENTS

This work was supported in part by the Belarussian and Russian Foundations for Basic Research (project nos. F02R-092, 02-02-81033 Bel 2002_a), INTAS (grant no. 01-0233) and under the Program for Support of Leading Scientific Schools (grant 2159.2003.2).

REFERENCES

1. V. B. Gildenburg and N. V. Vvedenskiy, Phys. Plasmas **8**, 1953 (2001).
2. N. V. Vvedenskiĭ and V. B. Gil'denburg, Pis'ma Zh. Éksp. Teor. Fiz. **76**, 440 (2002) [JETP Lett. **76**, 380 (2002)].
3. S. V. Bulanov, L. M. Kovrizhnykh, and A. S. Sakharov, Zh. Éksp. Teor. Fiz. **72**, 1809 (1977) [Sov. Phys. JETP **45**, 949 (1977)].
4. S. V. Bulanov, L. M. Kovrizhnykh, and A. S. Sakharov, Phys. Rep. **186**, 1 (1990).
5. S. V. Bulanov, F. Pegoraro, and J. I. Sakai, Nucl. Instrum. Methods Phys. Res. A **410**, 497 (1998).
6. V. I. Arkhipenko, V. N. Budnikov, E. Z. Gusakov, *et al.*, in *Proceedings of the III International Workshop on Strong Microwaves in Plasmas*, Ed. by A. G. Litvak (Nizhni Novgorod, Russia, 1997), Vol. 2, p. 516.
7. V. I. Arkhipenko, V. N. Budnikov, I. A. Romanchuk, and L. V. Simonchik, Fiz. Plazmy **7**, 396 (1981) [Sov. J. Plasma Phys. **7**, 216 (1981)].
8. V. I. Arkhipenko, V. N. Budnikov, E. Z. Gusakov, *et al.*, Fiz. Plazmy **13**, 693 (1987) [Sov. J. Plasma Phys. **13**, 398 (1987)].
9. V. I. Arkhipenko, V. N. Budnikov, E. Z. Gusakov, *et al.*, Zh. Tekh. Fiz. **54**, 2136 (1984) [Sov. Phys. Tech. Phys. **29**, 1254 (1984)].
10. P. V. Fel'tsman, I. P. Zapesochnyĭ, *et al.*, Ukr. Fiz. Zh. **12**, 633 (1967).
11. M. I. Bakunov, A. M. Bystrov, and V. B. Gildenburg, Phys. Plasmas **9**, 2803 (2002).
12. X. Xu, Y. Nishida, and N. Yugami, Phys. Rev. E **55**, 3328 (1997).

Translated by V. Astakhov

Journal of Experimental and Theoretical Physics, Vol. 100, No. 2, 2005, pp. 449–454.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 127, No. 2, 2005, pp. 502–508.

STATISTICAL, NONLINEAR, AND SOFT MATTER PHYSICS

On the Initial Stage of Spinodal Decomposition

Yu. V. Shikina[a] and V. B. Shikin[b,*]

[a]*Institute of Microelectronic Technology and Ultra-High-Purity Materials, Russian Academy of Sciences, Chernogolovka, Moscow oblast, 142432 Russia*

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

**e-mail: shikin@issp.ac.ru*

Received August 13, 2004

Abstract—A consistent theory describing the initial stage of spinodal decomposition of a two-component system is proposed. It is shown that the structure factor $S(q, t)$ has two maxima as a function of wavenumber at this stage. The main maximum point q_+ varies with time, first moving from q_{saddle} to q_{m} (given by expressions (22) and (12), respectively) and then back after the turning point in time given by (26) is passed. The other maximum point is localized at $q \approx 0$, and the corresponding peak amplitude is virtually independent of time. The characteristics of the main maximum are sensitive to the existence of the "zero" peak. Available experimental observations support the predictions of the theory. *© 2005 Pleiades Publishing, Inc.*

Detailed experimental studies of the behavior of the structure factor $S(q, t)$ at the early stages of spinodal decomposition in various systems, when they span a wide range of wavenumber q, reveal that it has two maxima characterized by essentially different properties (see [1–4]). The main maximum point is separated from the origin in the q space, and its location varies with time as the instability evolves. The other one, being less distinct, is located in the neighborhood of $q = 0$ and is virtually time-independent.

The time-independent maximum is either ignored [1, 2] or attributed to instrumental noise [3]. However, a spurious peak in the neighborhood of $q = 0$ must not affect the characteristics of the main maximum of $S(q, t)$, which disagrees with observations.

In this paper, we propose a scenario that attributes the existence of a time-independent "zero" maximum of $S(q, t)$ to the influence of initial conditions, which cannot be ignored in any analysis of spinodal decomposition. The persistence of this maximum as a distinct entity against the background exponential growth of the main peak can be explained, first, by their localization in separate intervals of the q axis and, second, by special properties of relaxation time, which is anomalously long for small wavenumbers, $\tau^{-1}(q \longrightarrow 0) \propto q^2$.

In conventional theories of spinodal decomposition, analysis of effects due to initial conditions is avoided when the properties of $S(q, t)$ are determined (see [5–10]). It is assumed that decomposition in the domain of linear behavior with respect to the $S(q, t)$ amplitude is "triggered" by a Langevin force. Correct definition of these forces for the spinodal portion of the phase diagram requires a special analysis (see [5–7]). A formal analysis leads to an equation that is similar to (6), but is supplemented with an inhomogeneous term $L(t)$ representing a Langevin force:

$$\frac{\partial S(q)}{\partial t} = 2R(q)S(q) + L(t).$$

The solution to this equation obtained without using any initial conditions,

$$S(q, t) = \exp\left(-\int_0^t 2R dp\right)\int_0^t ds L(s)\exp\left(+\int_0^s 2R d\sigma\right),$$

is not "strictly" exponential (e.g., see Eq. (12)), unlike the function

$$S(q, t) = S_0(q)\exp\left(+\int_0^t 2R dp\right),$$

where $S_0(q)$ is an initial distribution. However, strictly exponential behavior is believed to be a qualitative indication of the spinodal nature of decomposition and is regularly demonstrated by experimentalists [1–4] (see comments to Eq. (8) below). These considerations are additional motivation an analysis of initial conditions as a factor that ensures exponential behavior of the linear part of spinodal decomposition and suggest specific experimental schemes that make it possible to completely expose their role, as in the measurements of "mechanically triggered" spinodal decomposition reported in [1, 2]. Consider a liquid solution whose equilibrium characteristics lie in the spinodal instability region. Mechanical mixing brings the solution into an unstable quasi-homogeneous state, which begins to decompose as soon as the mixer is switched off. The

kinetics of the ensuing decomposition was measured in [1, 2]. Since the Langevin force is characterized by certain specific scales due to its thermal nature, it is clear that they cannot have any substantial effect on the kinetics of a mechanically stimulated decomposition at its initial stage. A detailed description of the scenario discussed in [1, 2] is given below.

1. The starting nonlinear equation for the Fourier component $S(q, t)$ of the structure factor proposed in [7] is written as

$$\frac{\partial S(q)}{\partial t} = -2Mq^2\left\{\left(Kq^2 + \frac{\partial^2 f}{\partial c_0^2}\right)S(q) + \frac{1}{2}\frac{\partial^3 f}{\partial c_0^3}S_3(q) + \frac{1}{6}\frac{\partial^4 f}{\partial c_0^4}S_4(q) + \ldots\right\}, \quad (1)$$

$$S(q, t) \equiv \int d\mathbf{r} S(r, t)\exp(i\mathbf{q}\cdot\mathbf{r}), \quad (2)$$

$$S(|\mathbf{r} - \mathbf{r}_0|) \equiv \langle u(\mathbf{r}, t)u(\mathbf{r}_0, t)\rangle, \quad u(\mathbf{r}, t) = c(\mathbf{r}, t) - c_0, \quad (3)$$

where c_0 is the mean solution concentration, angle brackets denote averaging over a distribution function ϱ preserving translational symmetry, and M is a phenomenological quantity proportional to the mobility of the separating components.

Both function $f(c)$ and kinetic coefficient K arise when Eq. (1) is derived by invoking the Flory–Higgins free energy for a binary system,

$$F\{c\} = \int dr\{0.5K(\nabla c)^2 + f(c)\},$$
$$f(c) = \frac{c}{N_1}\ln c + \frac{1-c}{N_2}\ln(1-c) + \chi c(1-c), \quad (4)$$

where χ is a coupling parameter and $R_g^2 = a^2N/6$ is the gyration radius. It is assumed in (1) that $N_1 = N_2 \equiv N$ and $a_1 = a_2 \equiv a$ (a is an elementary length).

The quantities denoted by S_n are the Fourier components of the higher order correlation functions

$$S_n(|\mathbf{r} - \mathbf{r}_0|) \equiv \langle u^{n-1}(\mathbf{r})u(r_0)\rangle. \quad (5)$$

Equation (1) is the first in the hierarchy of equations of motion for correlation functions that arises when higher order moments of the corresponding master equation are calculated (see [7]). Dropping all nonlinear terms in (1), we obtain Cahn's linear equation for S [8]:

$$\frac{\partial S(q)}{\partial t} = 2R(q)S(q), \quad (6)$$

where

$$R(q) = -Mq^2\left(Kq^2 + \frac{\partial^2 f}{\partial c_0^2}\right). \quad (7)$$

For negative $\partial^2 f/\partial c_0^2$, i.e., in the neighborhood of c_0 corresponding to spinodal instability, the factor $R(q)$ is positive at $q < q_c$, where

$$q_c = \left(\frac{1}{q}\left|\frac{\partial^2 f}{\partial c_0^2}\right|\right)^{1/2} \quad (8)$$

and has a maximum at $q = q_m = q_c/\sqrt{2}$. In this approximation, when q is close to q_m, fluctuations grow exponentially into a quasiperiodic structure with period $\lambda_m = 2\pi/q_m$. A standard analysis of scattering data reduces to finding the logarithmic time derivative of scattering intensity as a function of q^2. If the result is a line, then the assumption about exponential behavior at the initial decomposition stage is valid, and the corresponding graph provides information about $R(q)/q^2$.

The nonlinear form of (1) can be simplified by applying the mean-field approximation developed in [7]. In this model, all odd correlations vanish, while the even ones (restricted to S_4 here) take the form

$$S_4(q) \approx 3\langle\delta u^2\rangle S(q), \quad \langle\delta u^2(t)\rangle = \langle u^2(t)\rangle - \langle u^2(0)\rangle \quad (9)$$

with

$$\langle u^2\rangle = \frac{1}{(2\pi^3)}\int d\mathbf{q}S(q, t). \quad (10)$$

The resulting equation for $S(q, t)$ (up to S_4), again (cf. (6)), is linear, but the constant $\partial^2 f/\partial c_0^2$ is replaced here by the time-dependent quantity

$$\frac{\partial^2 f}{\partial c_0^2}\left(1 + \frac{1}{2}\frac{\partial f^4/\partial c_0^4}{\partial f^2/\partial c_0^2}\langle\delta u^2(t)\rangle\right). \quad (11)$$

2. It is convenient to start an analysis of Eq. (1) from its linear form (6). Using the initial conditions and introducing appropriate dimensionless variables, we obtain

$$S(x, t) = \frac{S_0}{1 + x^2/2x_0^2}\exp\left\{\frac{t}{\tau}x^2\left(1 - \frac{x^2}{2x_m^2}\right)\right\},$$
$$x^2 = q^2R^2, \quad x_m^2 = q_m^2R^2, \quad x_0^2 = q_0^2R^2, \quad (12)$$
$$q_m^2 = \epsilon/2R^2, \quad \epsilon = (\chi - \chi_s)/\chi_s,$$

where x is the dimensionless wavenumber, x_m is the "linear" maximum point, x_0 is the standard deviation in the initial distribution, and τ is a relaxation time proportional to the coefficient M in (1). The form of the initial conditions in (12) corresponds either to the Ornstein–Zernike structure function in thermal transition to the spinodal region, as was done in [3, 4], or to mechanical preparation of the initial state (approximately described by the Lorentzian distribution) starting from the spinodal instability zone [1, 2]. The quantities χ and χ_s correspond to Flory–Higgins interaction in the transition and equilibrium regions, respectively. (Calculations are performed here in general form, but the available measurements discussed below are taken from polymer studies.) The characteristic size R, for example, in the case of mechanical preparation of the initial state [1], is

$$R^{-1} = q_0, \tag{12a}$$

where q_0 is the characteristic wavenumber for a mechanically prepared initial state. According to [1], $q_0 \approx 10^{-3}$ nm^{-1}.

When $S(x, t)$ in (12) is known, one can readily find the extrema of this function by solving the equation

$$dS(x, t)/dx = 0, \tag{13}$$

which has three roots. The root

$$x_1 = 0 \tag{14}$$

is time-independent.

The remaining roots are expressed as

$$x_\pm^2 = (0.5x_m^2 - x_0^2) \pm \{(0.5x_m^2 - x_0^2)^2 - (x_m^2\tau/t - 2x_0^2x_m^2)\}^{0.5}. \tag{15}$$

Roots (15) are meaningful if two conditions are satisfied:

$$(0.5x_m^2 - x_0^2) > 0, \quad (0.5x_m^2 - x_0^2)^2 > (x_m^2\tau/t - 2x_0^2x_m^2). \tag{16}$$

The second inequality in (16) determines the minimal time t_{min} required for a shifted maximum of $S(q, t)$ to develop:

$$t \geq t_{min}, \quad \frac{\tau}{t_{min}} = 2x_0^2 + \frac{(x_0^2 - 0.5x_m^2)^2}{x_m^2}. \tag{17}$$

We can also find the time t_{max} defined by the condition

$$x_m^2\tau/t_{max} - 2x_0^2x_m^2 = 0 \tag{18}$$

for disappearance of the "zero" maximum.

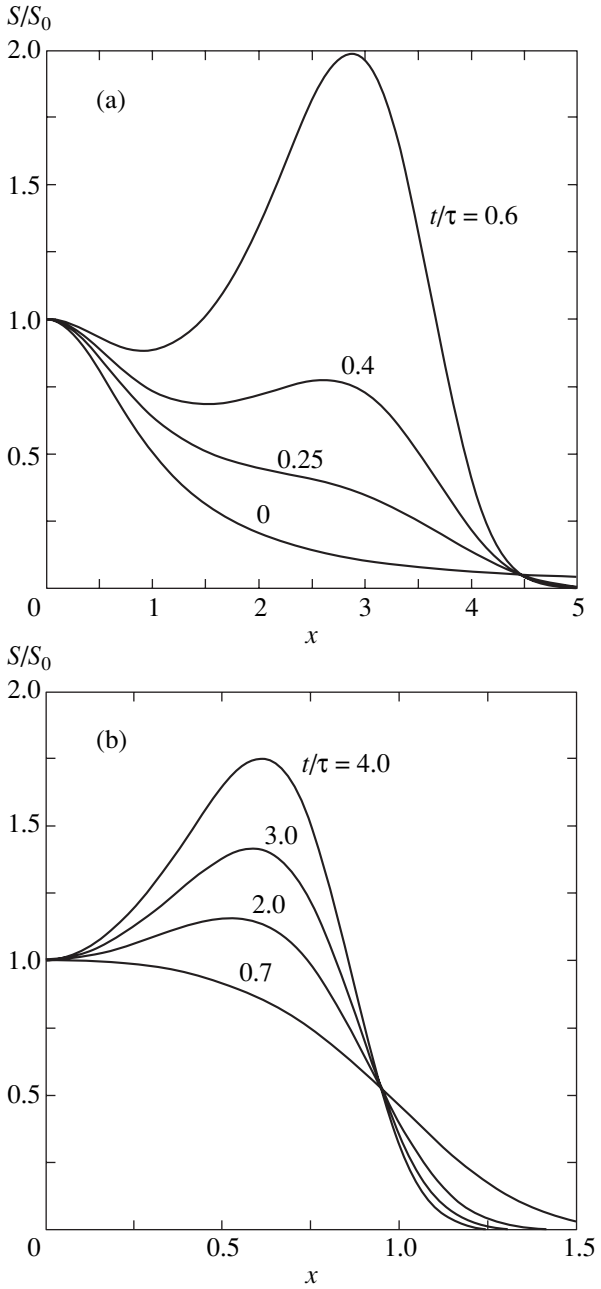

Fig. 1. Typical behavior of $S(x, t)$: (a) two maxima when condition (16) is satisfied, $x_0 = 0.7$, $x_m = 3.16$; (b) alternative possibility, $x_0 = 0.7$, $x_m = 0.67$; t/τ values are shown at curves.

Thus, the function $S(x, t)$ has two maxima, x_1 and x_+, in the interval $t_{min} \leq t \leq t_{max}$, where t_{min} and t_{max} are defined by (17) and (18), respectively. At $t > t_{max}$, the equation $dS/dq = 0$ has only two roots: a minimum at $x = 0$ and a maximum at the point defined as

$$x_+^2 = (0.5x_m^2 - x_0^2) + \{(0.5x_m^2 - x_0^2)^2 - (x_m^2\tau/t - 2x_0^2x_m^2)\}^{0.5}. \tag{19}$$

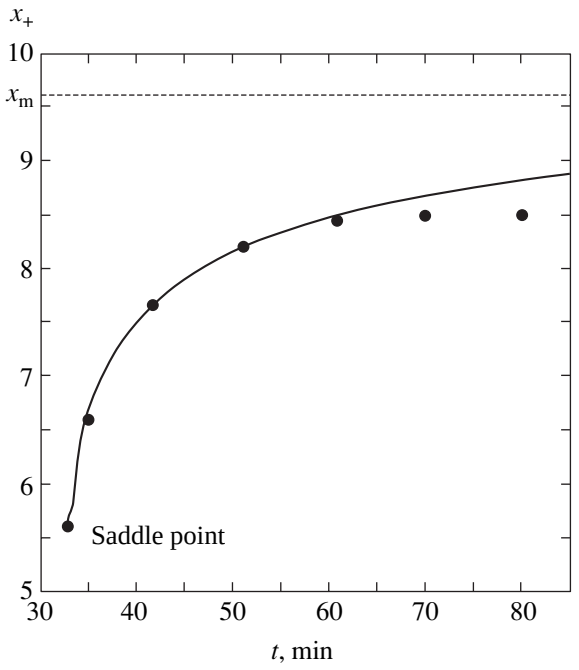

Fig. 2. Trajectory $x_+(t)$ (solid curve) for data from [1] (symbols) with parameters $S(x, t)$ in (12) adjusted at the initial stage of decomposition. Nonlinear correction of $x_+(t)$ is necessary after $t \geq 70$–80 min.

It is clear from (19) that $x_+ \neq x_m$ even at $t > t_{max}$. Only in the limit of $\tau/t \longrightarrow 0$, we have $x_+ \longrightarrow x_m$.

Figure 1 shows typical shapes of $S(x, t)$ corresponding to several values of x_m and x_0.

3. The results obtained here can be used to interpret the data reported in [1, 2], where a liquid symmetric polymer solution was prepared in an unstable state (the parameter ϵ in (12) was positive). The solution was thoroughly stirred to obtain as homogeneous a state as possible at the initial moment. After the mixer was switched off, spinodal decomposition was observed. The process could be monitored by using light-scattering data. It was shown that the main spinodal extremum $S(q, t)$ grew exponentially during an early decomposition stage (see Fig. 4 in [1]). Therefore, there exists a finite time interval, $0 \leq t < t_*$, where the exponential representation (12) corresponding to a linear approximation is reasonable.

Now, we note that the measured profile of $S(q, t)$ has a maximum at $q \longrightarrow 0$ as $t \longrightarrow 0$ (see Fig. 4 in [1]). This behavior can be approximated by using $S(x, 0)$ from (12) with R from (12a). An adjustment leads to the following estimate for x_0 in $S(x, 0)$ (1):

$$x_0 \approx 3.8. \tag{20}$$

We can also estimate x_m:

$$x_m = 9.6, \quad x = q/q_0, \quad q_0 = 10^{-3}\ \text{nm}^{-1}. \tag{21}$$

In calculating (20) and (21), a new normalization of variables is used. Instead of the relations in (12), which are well suited for general analysis, define a new scale in the q space: $q_0 = 10^3\ \text{nm}^{-1}$.

Using (20) and (21), we find that the first condition in (16) is satisfied,

$$0.5x_m^2 - x_0^2 > 0.$$

Therefore, $S(q, t)$ can have two maxima. The location q of the "saddle" point that corresponds to the appearance of a peak at x_+ is

$$x_{saddle}^2 = 0.5x_m^2 - x_0^2 = 31.64, \quad x_{saddle} \approx 5.62. \tag{22}$$

This result is consistent with the saddle-point location for the data reported in [1] (see Fig. 4 therein).

Using (17), we can also estimate τ/t_{saddle}:

$$\frac{\tau}{t_{min}} \equiv \frac{\tau}{t_{saddle}} = 2x_0^2 + \frac{(0.5x_m^2 - x_0^2)^2}{x_m^2} = 39.7. \tag{23}$$

In the experiments in [1], $\tau = x_m^2/4R_m$, where $R_m = R(q_m) = 0.5q_m^2 D_{eff} \approx 3 \times 10^{-4}\ \text{s}^{-1}$; i.e.,

$$\tau \approx 7.9 \times 10^4\ \text{s}. \tag{24}$$

Furthermore, according Fig. 3 in [1], $t_{saddle} \approx 33$ min. Using data from [1], we obtain

$$\tau/t_{sdaddle} \approx 39.9. \tag{23a}$$

Agreement between (23a) and (23) is obvious.

Let us discuss the behavior of the time-dependent maximum point x_+ given by (19). In the linear representation of (19), this point moves from x_{saddle} defined by (22) toward the asymptotic location x_m as $t \longrightarrow \infty$. Figure 2 provides an illustration based on data from [1]. In particular, this behavior explains why the root x_m obtained by analyzing the exponential part of $R(x)/x^2$, where $R(x) \sim x^2(1 - 0.5x^2/x_m^2)$ ($x_m = 9.6$ under the conditions of [1]), is distinct from the root x_+ given by (19) under the condition $dS/dx = 0$. The answer is obvious: x_+ is sensitive both to x_m and to x_0.

Figure 2 also illustrates the early nonlinear behavior in the development of spinodal decomposition: at $t >$ 70 min, the experimental data points representing $x_+(t)$ fall below the solid curve predicted by the linear theory. This phenomenon is explained below.

Figure 3 shows the general behavior of $S(q, t)$ under the conditions of [1]. The parameters S_0 and x_0 are adjusted to match the data from [1] at $t = 2.97$ and 22.2 min with $q = 4\ \text{nm}^{-1}$. Thus, we have $S_0 = 66$ and $x_0 = 3.8$. Under this normalization, expressions (12) for $S(x, t)$ and (12a) for R accurately reproduce the data from [1]. Unfortunately, the no data concerning small wavenumbers were presented in [1] (the initial data correspond to $q \geq 4 \times 10^3\ \text{nm}^{-1}$); i.e., no information is

available about the time-independent extremum $S(q \longrightarrow 0, t)$. Nevertheless, Fig. 3 adequately illustrates the situation in the neighborhood of the saddle point and the evolution of the main maximum point x_+. It is obvious that initial conditions play a significant role in the kinetics of the main maximum point at an early decomposition stage.

4. Now, consider the turning point in Fig. 2. Here, two processes begin to compete: one of them is responsible for the upward movement of $x_+(t)$; the other (of nonlinear nature) moves $x_+(t)$ to the downwards. The competition determines the evolution of the moving root near the turning point. To describe it, we solve the general equation (11). Instead of (15), we obtain

$$x_+^2(t) = \frac{1}{2}x_m^2 - x_0^2 - \frac{x_m^2}{2}\frac{\beta}{t}\int_0^t \langle \delta u_0^2(t)\rangle dt + \left\{\left(\frac{1}{2}x_m^2 - x_0^2 - \frac{x_m^2}{2}\frac{\beta}{t}\int_0^t \langle \delta u_0^2(t)\rangle dt\right)^2 - \frac{x_m^2}{t} + 2x_0^2 x_m^2\left(1 - \frac{\beta}{t}\int_0^t \langle \delta u_0^2(t)\rangle dt\right)\right\}^{0.5}, \quad (25)$$

$$\langle \delta u_0^2(t)\rangle = \int_0^{\sqrt{2}x_m} \frac{x^2 dx}{1 + \frac{x^2}{2x_0^2}}\left\{\exp\left(tx\left(1 - \frac{x^2}{2x_m^2}\right)\right) - 1\right\}.$$

Here, $\langle \delta u_0^2(t)\rangle$ is calculated by using (9) and (10) with linear expression (12) for $S(x, t)$ in the integrand. Note that the data concerning $S(q, t)$ in [1] are relative: the value of S_0 cannot be extracted from them. Therefore, the numerical value of $\langle \delta u_0^2(t)\rangle$ in $x_+(t)$ remains undetermined, and we have to introduce a factor β and adjust it to match the data from [1] in the neighborhood of the turning point t_{turn}. The quality of the adjustment is then checked by comparing the theoretical location x_{turn}^{calcul} with experimental result x_{turn}^{exp} from [1]:

$$x_0 = 3.8, \quad x_m = 9.6, \quad \frac{t_{turn}}{\tau} = 0.061, \quad x_{turn}^{exp} = 8.5. \quad (26)$$

This yields

$$\beta = 0.00013. \quad (27)$$

Accordingly,

$$x_{turn}^{calcul} = 8.35. \quad (28)$$

It is obvious that x_{turn}^{calcul} agrees with x_{turn}^{exp}.

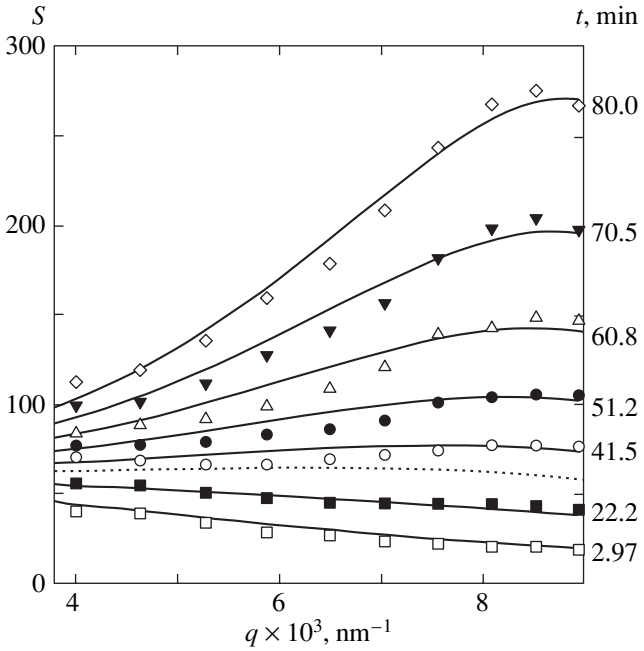

Fig. 3. Structure of $S(q, t)$ (12) at several instants normalized to data from [1] (symbols) at $t = 2.97$ and 22.2 min for $q = 4\ \text{nm}^{-1}$. Solid curves: normalized structure, $S_0 = 66$, $x_0 = 3.8$, $x = q/q_0$ with q_0 defined by (12a). Dotted curve corresponds to the saddle line with $t_{saddle} \sim 33$ min and $x_{saddle} \sim 5.62$.

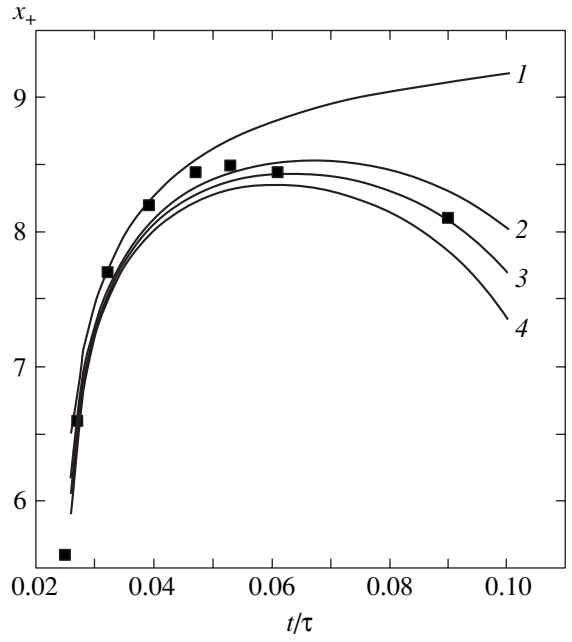

Fig. 4. Evolution of $x_+(t)$ (25) for β = (*1*) 0, (*2*) 0.0001, (*3*) 0.000125, and (*4*) 0.00015.

Figure 4 illustrates the behavior of $x_+(t)$ near t_{turn} corresponding to the parameter values in (26) and (27). The data points are borrowed from [1]. Calculations are in fair agreement with experiment.

According to Figs. 2 and 4, the nonlinearity of the problem begins to manifest itself at $t/\tau \geq 0.06$ under the conditions of [1]. (Nonlinear terms in (1) must be small as compared to the linear part of the equation.)

We conclude that initial conditions play an appreciable role in description of spinodal decomposition kinetics at an initial stage. First of all, the present analysis explains the "two-hump" form of the function $S(x, t)$ and the stability of the "zero" extremum ($\tau^{-1}(q \longrightarrow 0) \propto q^2$). Furthermore, an analysis of the data presented in [1] gives an answer to the following question: Why is the main extremum point x_+ in the linear treatment of spinodal decomposition does not coincide with the point x_m obtained in a standard analysis of exponential behavior of $S(x, t)$ at this stage. It turns out that these are different quantities. In the general case, $x_+(t) < x_m$, and x_+ is time-dependent. Figure 2 illustrates the difference between x_m (dashed curve) and $x_+(t)$ (solid curve) for the system studied in [1]. Also of interest is the existence of a turning point in the evolution of $x_+(t)$. Details of this nonlinear phenomenon are illustrated by Fig. 4.

REFERENCES

1. T. Izumitani and T. Hashimoto, J. Chem. Phys. **83**, 3694 (1985).
2. M. Okada and Ch. C. Han, J. Chem. Phys. **85**, 5307 (1986).
3. P. Wiltzius, F. S. Bates, and W. R. Hefner, Phys. Rev. Lett. **60**, 1538 (1988).
4. F. S. Bates and P. Wiltzius, J. Chem. Phys. **91**, 3258 (1989).
5. I. M. Lifshitz and V. V. Slyozov, J. Phys. Chem. Solids **19**, 35 (1961).
6. K. Binder, J. Chem. Phys. **79**, 6387 (1983).
7. J. S. Langer, M. Baron, and H. D. Miller, Phys. Rev. A **11**, 1417 (1975).
8. J. W. Cahn, Acta Metall. **9**, 795 (1961); Acta Metall. **10**, 179 (1962).
9. H. E. Cook, Acta Metall. **18**, 297 (1970).
10. A. Olemskoi and I. Koplyk, Usp. Fiz. Nauk **165**, 1105 (1995) [Phys. Usp. **38**, 1061 (1995)].

Translated by A. Betev