Journal of Experimental and Theoretical Physics, Vol. 96, No. 3, 2003, pp. 383–401.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 123, No. 3, 2003, pp. 437–456.

ATOMS, SPECTRA, RADIATION

The Light Pressure Force and the Friction and Diffusion Coefficients for Atoms in a Resonant Nonuniformly Polarized Laser Field

A. V. Bezverbnyi[a], O. N. Prudnikov[b], A. V. Taichenachev[b,*], A. M. Tumaikin[b], and V. I. Yudin[b]

[a]*Maritime State University, Vladivostok, 690059 Russia*
[b]*Novosibirsk State University, Novosibirsk, 630090 Russia*
**e-mail: avt@admin.nsu.ru*

Received May 6, 2002

Abstract—The motion of slow atoms with degenerate energy levels in a resonant, nonuniformly polarized laser field is described by the Fokker–Planck equation for the atomic distribution function in phase space in terms of the semiclassical approach. Field gradient expansions are used for the spatially nonuniform coefficients of the equation. For closed atomic transitions $J_g = J \longrightarrow J_e = J + 1$ (J_g and J_e are the total angular momenta of the ground and excited states, respectively), new analytical results are presented for the light pressure force and the friction and diffusion coefficients in momentum space. These results allow the kinetic effects (laser cooling, localization in optical potential wells, etc.) in a field of arbitrary one-, two-, or three-dimensional configuration to be investigated. In several cases, the new contributions to the friction coefficient are interpreted qualitatively.

1. INTRODUCTION

Major progress in laser cooling and in trapping neutral atoms is known to be associated with the use of field configurations with spatial polarization gradients. Field polarization gradients underlie the operation of magnetooptical traps, allow the Doppler cooling limit to be overcome, and are used to produce optical lattices and in a number of other applications. Such configurations arise every time the field is produced by interfering laser beams whose propagation directions and polarization vectors do not coincide. The problem of theoretically describing the motion of an atom in a nonuniformly polarized laser field is complex and requires taking into account all of the possible recoil effects and the processes of optical orientation associated with the redistribution of atoms in magnetic sublevels. In particular, cooling below the Doppler limit ($T_D \sim 10^{-3}$ K) is achieved by using laser fields in which the orientation or eccentricity of the polarization ellipse varies at distances on the order of the light wavelength. In these cases, a peculiar correlation arises between the optical orientation of the ground state, the translational motion, and the momentum transfer from the field to atoms. This correlation is generally responsible for the formation of new kinetic effects.

Dalibard and Cohen-Tannoudji [1] laid the foundations of the sub-Doppler cooling theory. They considered simple one-dimensional field configurations in the semiclassical approximation: counterpropagating waves with orthogonal linear (lin $\perp$ lin) and circular (σ_+–σ_-) polarizations (in the former and latter cases, field configurations with spatial gradients in the ellipticity and rotation angle of the polarization ellipse are realized, respectively), and simple atomic transitions from the class $J \longrightarrow J + 1$ ($1/2 \longrightarrow 3/2$ and $1 \longrightarrow 2$, respectively). Sub-Doppler cooling can be also achieved on transitions from the class $J \longrightarrow J$ (J is a half-integer), which was shown in [2, 3] for the $1/2 \longrightarrow 1/2$ transition. Despite the large number of theoretical papers on this type of laser cooling, the analytical treatment of the friction and diffusion coefficients has been restricted so far to low angular momenta and simple field configurations. In particular, the dependence of cooling parameters on the angle between the linear polarizations of counterpropagating waves in a lin–θ–lin configuration [2, 4] and the effect of intensity imbalance between circularly polarized waves [5] were analyzed. Calculations related to real experiments, in which transitions with large angular momenta ($J = 2$ for ^{23}Na, ^{87}Rb; $J = 3$ for ^{85}Rb; $J = 4$ for ^{133}Cs) and complex (two- and three-dimensional) field configurations are commonly used, were performed by various numerical methods [6].

This situation can be partly explained by the fact that no analytical expressions for the stationary atomic density matrix in a field with an arbitrary elliptical polarization had been known before the series of papers [7–9] even without considering the translational motion and recoil effects. This problem, which forms the zero approximation in a semiclassical treatment of the kinetics of slow atoms, was completely solved for $J \longrightarrow J$

(J is a half-integer) [7] and $J \longrightarrow J + 1$ [8–11] transitions. The existence of polarization features of the kinetics, which had not been considered previously, followed even from the form of the solution itself. Thus, even in the simplest case of a uniformly polarized standing wave and the atomic $1/2 \longrightarrow 1/2$ transition, additional (compared to the model of nondegenerate states) contributions to the friction and diffusion attributable to field ellipticity were found [12]. Even more unexpected features were detected when considering the kinetics of atoms in nonuniformly polarized fields with spatial gradients simultaneously in several field parameters (the ellipticity and rotation angle of the ellipse, the field amplitude and phase) [13, 14].

The primary goal of this study is to develop analytical invariant methods for describing the motion of atoms in a resonant, nonuniformly polarized laser field of the most general configuration. In other words, we seek to obtain results valid for an arbitrary field without associating our treatment with any specific field configuration (with the exception of Section 8, where a special class of one-dimensional field configurations is considered). The main approximations that allow this program to be carried out are as follows. First, assuming the photon momentum $\hbar k$ to be small compared to the atomic momentum dispersion Δp, we use the semiclassical approximation for the atomic center-of mass motion. In this approximation, the evolution of an atomic ensemble at the kinetic stage is known [15, 16] to be described by the Fokker–Planck equation for the atomic distribution function in phase space. Therefore, the problem reduces to seeking the coefficients of this equation, i.e., the force and diffusion (in momentum space). Second, representing these coefficients as a power series of the atomic velocity, we restrict our analysis to the lowest orders (the zero and first orders for the force and the zero order for the diffusion). This corresponds to the slow-atom approximation, where an atom is displaced by a distance much smaller than the light wavelength in the relaxation time scale in internal degrees of freedom.

Of course, these approximations impose certain restrictions on the applicability of our results to practical problems of laser cooling and atomic trapping. In particular, as we know from theory [2, 6] and sub-Doppler cooling experiments [17], the atomic temperature first decreases (roughly linearly) and then increases sharply with decreasing field intensity (or with increasing detuning from resonance). The slow-atom approximation is definitely inapplicable near the minimum of this dependence, and the condition for semiclassical motion is violated as the intensity further decreases. Nevertheless, in the range where the temperature is a linear function of the intensity, our approximations are valid and the semiclassical kinetics of slow atoms satisfactorily describes the experimental results.

2. STATEMENT OF THE PROBLEM

Consider an atomic ensemble with total angular momenta J_g and J_e in the ground and excited states that resonantly interacts with a monochromatic light field

$$\mathbf{E}(\mathbf{r}, t) = e^{-i\omega t}\mathbf{E}(\mathbf{r}) + \text{c.c.} \tag{1}$$

The Hamiltonian of the interaction with the external field (1) in the dipole approximation in a basis rotating at the field frequency does not explicitly depend on time:

$$\hat{H}_{D-E}(\mathbf{r}) = \hbar\Omega(\mathbf{r})\hat{V}(\mathbf{r}) + \text{h.c.}, \tag{2}$$

where the Rabi frequency $\Omega = -\langle e||d||g\rangle E/\hbar$ specifies the scalar amplitude of the interaction and $\langle e||d||g\rangle$ is the reduced matrix element of the dipole moment operator. According to the Wigner–Eckart theorem, we represent the tensor part of the interaction operator as $\hat{V} = \hat{\mathbf{T}} \cdot \mathbf{E}/E$, where the covariant components of the vector operator $\hat{\mathbf{T}}$ can be expressed in terms of the Clebsch–Gordan coefficients:

$$\hat{T}_q = \sum_{(m)} |J_e, m_e\rangle C^{J_e, m_e}_{J_g, m_g;\, 1, q}\langle J_g, m_g|. \tag{3}$$

In this case, the Hamiltonian of a free atom in the center-of-mass frame can be written as

$$\hat{H}_0 = -\hbar\delta\hat{\Pi}_e, \tag{4}$$

where $\delta = \omega - \omega_0$ is the field frequency detuning from the atomic transition frequency and the projection operator

$$\hat{\Pi}_e = \sum_{\mu_e} |J_e, \mu_e\rangle\langle J_e, \mu_e|$$

is represented by a unit matrix in the basis of sublevel wave functions for the excited state ($\hat{\Pi}_g$, the projector onto the ground state, is defined in a similar way).

The quantum kinetic equation for the atomic density matrix in Wigner representation for translational degrees of freedom [18] is

$$\begin{aligned}\left(\frac{\partial}{\partial t} + \frac{\mathbf{p}\cdot\nabla_r}{M}\right)\hat{\rho}(\mathbf{r}, \mathbf{p}) &= -\frac{i}{\hbar}[\hat{H}_0, \hat{\rho}(\mathbf{r}, \mathbf{p})] \\ &- \frac{i}{\hbar}\left[\hat{H}_{D-E}\left(\mathbf{r} + \frac{i\hbar\nabla_p}{2}\right)\hat{\rho}(\mathbf{r}, \mathbf{p})\right. \\ &\left. - \hat{\rho}(\mathbf{r}, \mathbf{p})\hat{H}_{D-E}\left(\mathbf{r} - \frac{i\hbar\nabla_p}{2}\right)\right] - \gamma\frac{1}{2}\{\hat{\Pi}_e, \hat{\rho}(\mathbf{r}, \mathbf{p})\} \\ &+ \gamma\frac{3}{2}\left\langle \sum_{s=1,2} (\hat{\mathbf{T}}\cdot\mathbf{e}_s(\mathbf{k}))^{\dagger}\hat{\rho}(\mathbf{r}, \mathbf{p} + \hbar\mathbf{k})(\hat{\mathbf{T}}\cdot\mathbf{e}_s(\mathbf{k}))\right\rangle_{\Omega_k},\end{aligned} \tag{5}$$

where γ is the radiative damping rate for the excited level. The vector operators ∇_r and ∇_p act on the Wigner density matrix $\hat{\rho}(\mathbf{r}, \mathbf{p})$. The normalization condition for the latter can be written as

$$\int \mathrm{Tr}\{\hat{\rho}(\mathbf{r}, \mathbf{p})\} d^3r d^3p = 1. \quad (6)$$

The induced recoil effect is described by the operators ∇_p in Eq. (5) and is related to the spatial nonuniformity of the energy of interaction with the external field. The spontaneous recoil effect manifests itself in the term that describes the radiative arrival [the last term on the right-hand side of Eq. (5)] in the form of a shift in the argument of the Wigner density matrix by the photon momentum $\hbar\mathbf{k}$ along which the averaging ($\langle \ldots \rangle_{\Omega_k}$) is performed. The unit polarization vectors $\mathbf{e}_s(\mathbf{k})$ of the spontaneous photons are orthogonal to the escape direction.

It is well known [19–21] that the presence of a small parameter $\hbar k/\Delta p \ll 1$ (the ratio of resonant photon momentum to atomic momentum dispersion) allows the rapid processes of ordering in internal degrees of freedom to be separated from the slow processes associated with translational motion. Relaxation times to a stationary (in the zero order in recoil) distribution in internal degrees of freedom $t \gg \tau_{\max} = \max\{\gamma^{-1}, (\gamma S)^{-1}\}$, where the saturation parameter is defined as

$$S = \frac{|\Omega|^2}{\gamma^2/4 + \delta^2}, \quad (7)$$

correspond to the kinetic stage of the evolution.

In this case, the dynamics of the atomic ensemble is determined by the slow change of the distribution function in translational degrees of freedom, $w(\mathbf{r}, \mathbf{p}) = \mathrm{Tr}\{\hat{\rho}(\mathbf{r}, \mathbf{p})\}$. To within the second order in recoil parameter [19, 15], reducing the original quantum kinetic equation (5) to a closed equation for $w(\mathbf{r}, \mathbf{p})$ yields the Fokker–Planck equation

$$\frac{d}{dt} w(\mathbf{r}, \mathbf{p}) = \left[-\sum_i \frac{\partial}{\partial p_i} F_i(\mathbf{r}, \mathbf{p}) + \sum_{ij} \frac{\partial}{\partial p_i} \frac{\partial}{\partial p_j} D_{ij}(\mathbf{r}, \mathbf{p}) \right] w(\mathbf{r}, \mathbf{p}), \quad (8)$$

whose coefficients $F_i(\mathbf{r}, \mathbf{p})$ and $D_{ij}(\mathbf{r}, \mathbf{p})$ have the meaning of Cartesian components of the force vector and diffusion tensor (in momentum space), respectively for an atom in the light field; and $\frac{d}{dt} = \frac{\partial}{\partial t} + \frac{\mathbf{p}\nabla r}{M}$.

In deriving Eq. (8), we assumed that the Wigner density matrix as a function of the momentum changes only slightly on photon momentum scales $\hbar k$. This requirement leads to the condition

$$\min\{\gamma, \gamma S\} \gg \frac{\mathscr{E}_r}{\hbar}. \quad (9)$$

Otherwise, the density matrix elements are sharply oscillating (on $\hbar k$ scale) functions of the momentum [15, 22].

Let us discuss in more detail the condition for the emergence of two distinctly different time scales for the evolution in internal and translational degrees of freedom. As we see from Eq. (8), the rate of change in the distribution function is determined by the light pressure force and atomic momentum dispersion: $dw/dt \approx F/\Delta p w$. This rate must be much lower than the minimum relaxation rate in internal degrees of freedom:

$$\frac{F}{\Delta p} \ll \min\{\gamma, \gamma S\}. \quad (10)$$

If condition (10) is violated, then the atomic kinetics is not described by one equation (8); i.e., it has a multicomponent character (see, e.g., [16]). For the induced light pressure, the force is determined by the rate of photon rescattering from one external field mode to another. In this case, condition (10) can lead to more stringent constraints than $\hbar k/\Delta p \ll 1$. Thus, in a strong field $S \gg 1$, the induced force is proportional to the Rabi frequency, $F \approx \hbar k \Omega$, which leads to the condition [16]

$$\frac{\hbar k}{\Delta p} \ll \frac{\gamma}{\Omega} \ll 1.$$

In the opposite limit $S \ll 1$, we have an estimate $F \approx \hbar k \delta S$ for the force at large detunings, $\delta \gg \gamma$. As a result, the time-scale separation condition can be written as

$$\frac{\hbar k}{\Delta p} \ll \frac{\gamma}{\delta} \ll 1.$$

3. THE SLOW ATOM APPROXIMATION

The coefficients of the Fokker–Planck equation (8) can be analytically calculated for a field nonuniform in intensity and polarization only in various limiting situations. As applied to laser cooling problems, of certain interest (see the Introduction) is an analysis of slow atoms that are displaced by distances much smaller than the light wavelength in the characteristic times of ordering in internal degrees of freedom, i.e., $v\tau_{\max} \ll \lambda$. By the definition of $\tau_{\max}$, this leads to the inequality

$$kv \ll \max\{\gamma, \gamma S\}. \quad (11)$$

When this condition is satisfied, to properly take into account dissipative processes, it will suffice to restrict

our analysis to the linear approximation in velocity in the expression for the force:

$$F_i(\mathbf{r}, \mathbf{p}) \approx F_i(\mathbf{r}) + \sum_j \xi_{ij}(\mathbf{r}) v_j, \tag{12}$$

where $\mathbf{F}(\mathbf{r})$ is the light pressure force exerted on an atom at rest at point $\mathbf{r}$. The antisymmetric part of the tensor ξ_{ij} ($\propto \epsilon_{ijk} b_k$) corresponds to the effective Lorentz force ($\propto \mathbf{v} \times \mathbf{b}$), and its symmetric part defines the friction force. The zero approximation in velocity is commonly used for the diffusion:

$$D_{ij}(\mathbf{r}, \mathbf{p}) \approx D_{ij}(\mathbf{r}). \tag{13}$$

Note that caution should be exercised in using the slow-atom approximation for field configurations with nodes (e.g., in a standing wave [16]) or sharp gradients with characteristic spatial scales $z_{sc} \ll \lambda$ (e.g., in a lin–θ–lin configuration for $\theta \ll 1$ [2]). In both cases, the slowness condition becomes more stringent than (11). In a strong standing wave, $\Omega \gg \delta > \gamma$, the linear velocity dependence of the force is known [16] to be restricted by the condition

$$\frac{kv}{\gamma} \ll \frac{\delta}{\Omega} \ll 1.$$

For sharp gradients, the following condition should be used in place of (11):

$$\frac{v}{z_{sc}} \ll \min\{\gamma, \gamma S\}.$$

Below, we give general expressions for the kinetic coefficients $F_i(\mathbf{r})$, $\xi_{ij}(\mathbf{r})$, and $D_{ij}(\mathbf{r})$. The light pressure force exerted on an atom at rest is determined by the mean value of the force operator

$$\hat{\mathbf{F}}(\mathbf{r}) = -\nabla_r \hat{H}(\mathbf{r}), \tag{14}$$

where $\hat{H}(\mathbf{r}) = \hat{H}_0 + \hat{H}_{D-E}(\mathbf{r})$ is the total Hamiltonian of the atom at point $\mathbf{r}$:

$$\mathbf{F}(\mathbf{r}) = \mathrm{Tr}\{\hat{\mathbf{F}}(\mathbf{r})\hat{\sigma}(\mathbf{r})\}. \tag{15}$$

The density matrix $\hat{\sigma}(\mathbf{r})$ is the solution of the system of stationary Bloch equations:

$$-\frac{i}{\hbar}[\hat{H}(\mathbf{r}), \hat{\sigma}] - \hat{\Gamma}\{\hat{\sigma}\} = 0, \quad \mathrm{Tr}\{\hat{\sigma}\} = 1. \tag{16}$$

Here, the spatially uniform operator $\hat{\Gamma}$ describes the radiative relaxation:

$$\hat{\Gamma}(\hat{\sigma}) = \frac{\gamma}{2}\{\hat{\Pi}_e, \hat{\sigma}\} - \gamma \sum_q \hat{T}_q^\dagger \hat{\sigma} \hat{T}_q. \tag{17}$$

This solution, which is presented in explicit analytical form in [8–11], describes the stationary atomic distribution in magnetic sublevels of the ground and excited states in the zero approximation in the recoil parameter and atomic velocity. The tensor ξ_{ij} is proportional to the spatial gradient of $\hat{\sigma}(\mathbf{r})$:

$$\xi_{ij}(\mathbf{r}) = -\mathrm{Tr}\{\hat{\varphi}_i(\mathbf{r})\nabla_j \hat{\sigma}(\mathbf{r})\}, \tag{18}$$

where ∇_j are the Cartesian components of the vector operator ∇. Like the light pressure force, the spontaneous diffusion tensor component can be expressed by using $\hat{\sigma}(\mathbf{r})$:

$$D_{ij}^{(sp)}(\mathbf{r}) = \frac{\gamma(\hbar k)^2}{5} \times \mathrm{Tr}\left\{\left(\delta_{ij}\hat{\Pi}_e - \frac{1}{4}(\hat{T}_i\hat{T}_j^\dagger + \hat{T}_j\hat{T}_i^\dagger)\right)\hat{\sigma}(\mathbf{r})\right\}. \tag{19}$$

The induced diffusion tensor can be represented as

$$D_{ij}^{(\mathrm{ind})}(\mathbf{r}) = \frac{1}{2}\mathrm{Tr}\{(\hat{\varphi}_i(\mathbf{r})\widehat{\delta F}_j(\mathbf{r}) + \widehat{\delta F}_j(\mathbf{r})\hat{\varphi}_i(\mathbf{r}))\hat{\sigma}(\mathbf{r})\}, \tag{20}$$

where we denoted the fluctuation of the force operator by $\widehat{\delta\mathbf{F}} = \hat{\mathbf{F}} - \mathbf{F}$. Generalizing the method from [12] to three dimensions, we introduced auxiliary matrices $\hat{\varphi}_i$ in formulas (18) and (20), which allowed us to write the expressions for the friction and induced diffusion coefficients in a uniform way. The matrices $\hat{\varphi}_i$ are the solution of the inhomogeneous linear equation

$$\left[\left(\frac{\gamma}{2} + i\delta\right)\hat{\Pi}_e\hat{\varphi}_i(\mathbf{r}) + \left(\frac{\gamma}{2} - i\delta\right)\hat{\varphi}_i(\mathbf{r})\hat{\Pi}_e\right] - \frac{i}{\hbar}[\hat{H}_{D-E}(\mathbf{r}), \hat{\varphi}_i(\mathbf{r})] - \gamma\sum_q \hat{T}_q\hat{\varphi}_i(\mathbf{r})\hat{T}_q^\dagger = \widehat{\delta F}_i(\mathbf{r}), \tag{21}$$

in which the force operator fluctuation is the source. The induced diffusion tensor $D_{ij}^{(\mathrm{ind})}(\mathbf{r})$ is quadratic in $\widehat{\delta\mathbf{F}}$, which corresponds to the standard definition of diffusion via a two-time correlator of the force operator (see, e.g., [20]). As was noted in [12], the system of linear equations (21) is degenerate; i.e., the matrices $\hat{\varphi}_i$ are defined to within the gauge condition. For example, $\mathrm{Tr}\{\hat{\varphi}_i\} = 0$ can be used as this condition.

An alternative approach to calculating the tensors ξ_{ij} and $D_{ij}^{(\mathrm{ind})}$ is to solve the equations for the corrections of the first order in velocity and recoil to the Wigner density matrix. In this case, an invariant description based on the expansion of the atomic density matrix into a basis of bipolar harmonics [14] is possible. In several cases, these equations are more convenient in interpreting the various contributions to ξ_{ij} and $D_{ij}^{(\mathrm{ind})}$, and they are used below.

4. THE KINETIC COEFFICIENTS IN A FIELD OF ARBITRARY CONFIGURATION

To analyze the kinetics of atoms in a laser field of arbitrary configuration, it is convenient to represent the kinetic coefficients as an expansion in terms of spatial field gradients. In general, six real quantities completely specify the local complex amplitude of the field $\mathbf{E}(\mathbf{r})$. The real amplitude E, the phase Φ, and the unit complex polarization vector $\mathbf{e}$ can be separated in an invariant way:

$$\mathbf{E}(\mathbf{r}) = E(\mathbf{r})\exp(i\Phi(\mathbf{r}))\mathbf{e}(\mathbf{r}). \quad (22)$$

As usual, we assume that $\mathrm{Im}(\mathbf{e}\cdot\mathbf{e}) = 0$ and introduce the ellipticity angle $\varepsilon(\mathbf{r})$:

$$\cos(2\varepsilon) = \mathbf{e}\cdot\mathbf{e}. \quad (23)$$

Note that the ambiguity in determining the real amplitude, the phase, and the ellipticity disappears if we require that they be analytic functions of the coordinates. In addition to E, Φ, and ε, we need the following three angles that specify the orientation of the polarization ellipse: $\phi(\mathbf{r})$, the angle of rotation about the axis orthogonal to the polarization ellipse; and $\alpha(\mathbf{r})$ and $\beta(\mathbf{r})$, the angles of rotation about the principal axes of the polarization ellipse. For the angles, we cannot give an invariant definition via the polarization vector $\mathbf{e}$ and its conjugate vector $\mathbf{e}^*$ similar to (23). Nevertheless, this can be done for their spatial gradients as follows. For example, writing the expansion of the polarization vector in terms of circular unit vectors $\mathbf{e}_{\pm1}$ in the local coordinate system with the axis

$$\mathbf{e}_0 = \frac{i\mathbf{e}\times\mathbf{e}^*}{\sin(2\varepsilon)} \quad (24)$$

as

$$\mathbf{e} = \cos\left(\varepsilon-\frac{\pi}{4}\right)e^{-i\phi}\mathbf{e}_{+1} + \sin\left(\varepsilon-\frac{\pi}{4}\right)e^{i\phi}\mathbf{e}_{-1}$$

and calculating the scalar product $\mathbf{e}\cdot\nabla_i\mathbf{e}^*$, we find that

$$\sin(2\varepsilon)\nabla_i\phi = \mathrm{Im}(\mathbf{e}\cdot\nabla_i\mathbf{e}^*). \quad (25)$$

Similarly, considering the component of an infinitesimal increment of the polarization vector along the $\mathbf{e}_0$ axis, we obtain

$$\cos(\varepsilon)\nabla_i\alpha = \mathrm{Re}(\mathbf{e}_0\cdot\nabla_i\mathbf{e}), \quad (26)$$

$$\sin(\varepsilon)\nabla_i\beta = \mathrm{Im}(\mathbf{e}_0\cdot\nabla_i\mathbf{e}). \quad (27)$$

By definition (15), the light pressure force exerted on an atom at rest is linear in field gradients and it can be written as

$$\mathbf{F} = \hbar\sum_{\kappa=1}^{6}\mathscr{F}_\kappa\mathbf{g}^{(\kappa)}, \quad (28)$$

where

$$\mathbf{g}^{(1)} = \nabla\ln E, \quad \mathbf{g}^{(2)} = \nabla\Phi,$$
$$\mathbf{g}^{(3)} = \nabla\varepsilon, \quad \mathbf{g}^{(4)} = \nabla\phi, \quad (29)$$
$$\mathbf{g}^{(5)} = \nabla\alpha, \quad \mathbf{g}^{(6)} = \nabla\beta.$$

For the tensors ξ_{ij} and $D_{ij}^{(\mathrm{ind})}$ quadratic in field gradients, the expansions similar to (28) are

$$\xi_{ij} = \hbar\sum_{(\kappa)}\mathscr{K}_{\kappa\kappa'}g_i^{(\kappa)}g_j^{(\kappa')}, \quad (30)$$

$$D_{ij}^{(\mathrm{ind})} = \gamma\hbar^2\sum_{(\kappa)}\mathscr{D}_{\kappa\kappa'}g_i^{(\kappa)}g_j^{(\kappa')}, \quad (31)$$

where $g_i^{(\kappa)}$ are the Cartesian components of the vector $\mathbf{g}^{(\kappa)}$. These expansions have a number of remarkable properties:

(1) The coefficients $\mathscr{F}_\kappa$, $\mathscr{K}_{\kappa\kappa'}$, and $\mathscr{D}_{\kappa\kappa'}$ depend only on the local intensity $E^2(\mathbf{r})$, ellipticity $\varepsilon(\mathbf{r})$, and field detuning δ from resonance. They do not depend on the local phase and rotation angle, because we can always pass to a local basis where the phase and the rotation angles are zero when calculating the traces of the matrices in (15), (18), and (20).

(2) Each of the coefficients $\mathscr{F}_\kappa$, $\mathscr{K}_{\kappa\kappa'}$, and $\mathscr{D}_{\kappa\kappa'}$ has a certain parity relative to ε and δ. This statement is also quite general, because it can be proved by considering the symmetries of the kinetic equation for the matrix density relative to coordinate system inversion and time reversal [23]. We emphasize that the force vector $\mathbf{F}$ and the tensors ξ_{ij} and $D_{ij}^{(\mathrm{ind})}$ have no such parity.

(3) In general, the nondiagonal elements $\mathscr{K}_{\kappa\kappa'}$ and $\mathscr{D}_{\kappa\kappa'}$ at $\kappa\neq\kappa'$ are nonzero; i.e., the spatial field gradients of various types correlate with one another rather than act independently.

(4) Since the mean dipole moment of an atom at rest lies in the plane of the polarization ellipse and orthogonal to the small variations in polarization vector that bring it out of the polarization plane, the components $\mathscr{F}_5$ and $\mathscr{F}_6$ are generally zero. This reasoning leads us to conclude that the matrices $\mathscr{K}_{\kappa\kappa'}$ and $\mathscr{D}_{\kappa\kappa'}$ have a block structure—there are no correlations between sectors {1–4} and {5, 6}.

As regards the spontaneous diffusion tensor $D_{ij}^{(sp)}(\mathbf{r})$, it clearly also depends only on the ellipticity, intensity, and detuning.

As we will see below, our representation for the vector $\mathbf{E}(\mathbf{r})$ is convenient for writing the analytical results and the qualitative interpretation of various contributions. At the same time, it is clearly not the only possible one: any six real quantities that uniquely define the complex vector $\mathbf{E}(\mathbf{r})$ can be used. However, it is important to emphasize that the final results for the force vec-

tor **F** and the tensors ξ_{ij} and D_{ij} can always be expressed in terms of the complex vector amplitude of the field **E**(**r**) and its spatial gradients in an invariant and analytical way. The recipe for this transformation is described in Appendix A.

It should also be noted that the analytical expressions given below can be directly used to analyze the dependences of the force and diffusion on cooling-field parameters in the practically important cases of simple field configurations, where only one of the spatial field gradients is nonzero: a standing wave ($\mathbf{g}^{(1)} = \mathbf{k}$), a traveling wave ($\mathbf{g}^{(2)} = \mathbf{k}$), a lin $\perp$ lin configuration ($\mathbf{g}^{(3)} = \mathbf{k}$), and a σ_+–σ_- configuration ($\mathbf{g}^{(4)} = \mathbf{k}$). In all cases, the friction and induced diffusion coefficients are completely determined by one diagonal element $\mathscr{K}_{\kappa\kappa}$ and $\mathscr{D}_{\kappa\kappa}$ at $\kappa = 1, \ldots, 4$.

Thus, knowing the analytical expressions for the dependences of $\mathscr{F}_\kappa$, $\mathscr{K}_{\kappa\kappa'}$, $\mathscr{D}_{\kappa\kappa'}$, and $D_{ij}^{(sp)}$ on ellipticity, intensity, and detuning, we have general and complete information that allows the light pressure force, friction and diffusion coefficients to be determined for any specific field configuration.

To all appearances, expansions of the form (28) and (31) were first used by Gordon and Ashkin [24] to analyze the motion of an atom in a radiative trap. Subsequently, Kazantsev *et al.* [16] used an expansion similar to (30) in the problem of an effective Lorentz force in the field of a Gaussian beam. The authors of these papers used a simple model of a resonant atom in the form of two nondegenerate levels and, accordingly, considered the influence of only two gradients: $\mathbf{g}^{(1)}$ and $\mathbf{g}^{(2)}$, the field amplitude and phase, respectively. Expansions in terms of all the possible gradients, including the polarization ones, for the light pressure force exerted on an atom with a degenerate ground state in the zero and first approximations in velocity are given in [9, 14], where we also briefly discussed some of their properties.

Note that the results of this section are general. They remain valid for much more general atomic excitation and relaxation schemes than those considered in preceding sections. In particular, the ground and excited levels can have several hyperfine components with different angular momenta and energies; the isotropic relaxation operator can include extra terms that describe collisional relaxation, etc. Here, we restrict our analysis to closed dipole transitions $J_g \longrightarrow J_e$ in the approximation of purely radiative relaxation (as formulated in Section 2); the relaxation rate of the ground level is assumed to be exactly equal to zero. In this statement of the problem for the two classes of atomic transitions $J \longrightarrow J$ (J is an integer) and $J \longrightarrow J - 1$, where coherent population trapping in the ground state takes place [25], the light pressure force and the diffusion tensor for an atom at rest become zero. The tensor ξ_{ij} is antisymmetric; i.e., the linear (in velocity) force component reduces to the Lorentz force. In this case, the effective magnetic field is defined by the rotor of the vector geometric potential for which we derived analytical expressions in [26] for arbitrary angular momenta. As regards the corrections of a higher order in velocity (quadratic for the diffusion and cubic for the force), which determine the laser cooling effects in this case [27, 28], the approach developed here is clearly inapplicable to their calculation.

Transitions from the two remained classes $J \longrightarrow J$ (J is a half-integer) and $J \longrightarrow J + 1$ are considered in the laser cooling theory. However, laser cooling was experimentally achieved only for $J \longrightarrow J + 1$ transitions, because they are much more abundant than the $J \longrightarrow J$ (J is a half-integer) transitions. Therefore, we consider below $J \longrightarrow J + 1$ transitions, which are most important for applications, although we have similar results for $J \longrightarrow J$ (J is a half-integer) transitions.

5. THE LIGHT PRESSURE ON AN ATOM AT REST

The stationary solution found for the density matrix $\hat{\sigma}(\mathbf{r})$ in [7, 10, 11] allows the light pressure force exerted on an atom at rest to be analytically calculated with the same generality, i.e., for arbitrary angular momenta of the ground and excited states and for an arbitrary monochromatic field configuration. In general, the four nonzero force components $\mathscr{F}_\kappa$ can be written as

$$\mathscr{F}_1 = -\frac{2\delta S\alpha_1}{\alpha_0 + 2S\alpha_1}, \quad \mathscr{F}_2 = \frac{\gamma S\alpha_1}{\alpha_0 + 2S\alpha_1},$$
$$\mathscr{F}_3 = \tan(2\varepsilon)\frac{2\delta S(\alpha_1 - A)}{\alpha_0 + 2S\alpha_1}, \qquad (32)$$
$$\mathscr{F}_4 = -\sin(2\varepsilon)\frac{\gamma SA}{\alpha_0 + 2S\alpha_1}.$$

The even (in detuning) components $\mathscr{F}_2$ and $\mathscr{F}_4$ correspond to the spontaneous light pressure force associated with the induced absorption and the subsequent spontaneous emission. The odd components $\mathscr{F}_1$ and $\mathscr{F}_3$ contribute to the induced light pressure force associated with the coherent reemission of photons from one mode into another mode.

The coefficients α_0, α_1, and A depend only on $\cos(2\varepsilon)$. Their explicit form is determined by the type of optical transition. Thus, for $J \longrightarrow J + 1$ transitions, based on the results of [10], we can obtain

$$\alpha_0 = \frac{1}{(2J+1)(4J+1)!\cos(2\varepsilon)}\sum_{L=0}^{2J} C_L P_L\left(\frac{1}{\cos(2\varepsilon)}\right),$$
$$\alpha_1 = P_{2J+1}\left(\frac{1}{\cos(2\varepsilon)}\right), \qquad (33)$$
$$A = \frac{1}{(2J+1)\cos(2\varepsilon)}P'_{2J+1}\left(\frac{1}{\cos(2\varepsilon)}\right),$$

where

$$C_L = (2L+1)(2J-L)!(2J+L+1)!$$

In the expression for α_0 in (33), the summation is over even L (starting from 0) for integer angular momenta J and over odd L (starting from 1) for half-integer J, and the coefficient A is expressed in terms of the derivative of the Legendre polynomials

$$P'_L(x) = \frac{d}{dx}P_L(x).$$

When the field polarization approaches a circular one, $\varepsilon \longrightarrow \pm\pi/4$, the argument of the polynomials $1/\cos(2\varepsilon)$ and the polynomials themselves increase without limit. However, it is easy to see that the ratio of the polynomials in terms of which $\mathscr{F}_\kappa$ are expressed remain finite:

$$\frac{\alpha_1}{\alpha_0} \longrightarrow 1, \quad \frac{A}{\alpha_0} \longrightarrow 1,$$
$$\frac{\alpha_1 - A}{\alpha_0 \cos^2(2\varepsilon)} \longrightarrow \frac{2J}{4J+1}.$$

A remarkable feature that shows up the most clearly at large angular momenta J is the rapid decrease of C_L with decreasing L, which allows several terms in the sum with the largest L to be retained in approximate calculations. For example, the ratio

$$\frac{C_{2J}}{C_{2J-2}} = (4J+1)^2\frac{2J}{4J-3}$$

for $J = 4$ (a closed transition on the ^{133}Cs D_2 line) is 177.8. As a result, if we substitute the sum for α_0 with one term with $L = 2J$, then the error will be at a maximum for $\varepsilon = 0$ and will not exceed one percent.

6. SPONTANEOUS DIFFUSION

Since the diffusion tensor (19), which is attributable to fluctuations in the escape direction of spontaneous photons, can be expressed in terms of $\hat{\sigma}(\mathbf{r})$ like the light pressure force, we can also obtain an analytical result for it at arbitrary J. Like any symmetric second-rank tensor, D_{ij}^{sp} can be represented as an isotropic part ($\propto \delta_{ij}$) and an irreducible part with a zero trace:

$$D_{ij}^{(sp)} = \gamma(\hbar k)^2\left[\delta_{ij}\frac{\pi^{(e)}}{6} - (-1)^{J_e+J_g}\frac{2J_e+1}{10}\right.$$
$$\left.\times\begin{Bmatrix} 2 & 1 & 1 \\ J_g & J_e & J_e \end{Bmatrix}(\{\mathbf{e}_i\otimes\mathbf{e}_j\}_2\cdot\sigma_2^{(e)})\right], \quad (34)$$

where we use the standard notation [29] for the symbol $6j$ and the irreducible tensor product; $\mathbf{e}_i$ and $\mathbf{e}_j$ are the Cartesian unit vectors (irreducible first-rank tensors). The isotropic component in (34) is proportional to the total excited-level population,

$$\pi^{(e)} = \frac{\alpha_1 S}{\alpha_0 + 2\alpha_1 S}, \quad (35)$$

and the component with a zero trace corresponds to alignment in the excited state, which for $J \longrightarrow J+1$ transitions is

$$\sigma_2^{(e)} = -\frac{S}{\alpha_0+2\alpha_1 S}\frac{1}{4J+1}$$
$$\times\sqrt{\frac{30(J+1)}{(J+2)(2J+1)(2J+3)(2J+5)}}$$
$$\times\left[\frac{\{\mathbf{e}\otimes\mathbf{e}^*\}_2}{\cos(2\varepsilon)}P'_{2J+1}\left(\frac{1}{\cos(2\varepsilon)}\right)\right.$$
$$\left.+\frac{\{[\mathbf{e}\times\mathbf{e}^*]\otimes[\mathbf{e}\times\mathbf{e}^*]\}_2}{\cos^2(2\varepsilon)}P''_{2J+1}\left(\frac{1}{\cos(2\varepsilon)}\right)\right]. \quad (36)$$

We see from formula (36) that the spontaneous diffusion tensor is diagonalized in the local coordinate system associated with the polarization ellipse, where the z axis is orthogonal to the polarization plane (i.e., directed along $\mathbf{e}_0$) and the x and y axes are parallel to the principal axes of the ellipse (Re{$\mathbf{e}$} and Im{$\mathbf{e}$}). Since the Cartesian tensor $K_{ij} = (\{\mathbf{e}_i\otimes\mathbf{e}_j\}_2\cdot\{\mathbf{e}\otimes\mathbf{e}^*\}_2)$ in this coordinate system is

$$K_{ij} = \mathrm{diag}\left\{\frac{1+3\cos(2\varepsilon)}{6}, \frac{1-3\cos(2\varepsilon)}{6}, -\frac{1}{3}\right\},$$

in general, all eigenvalues of the spontaneous diffusion tensor are different and only in the special cases of linear and circular polarization does an axial symmetry take place.

7. THE FRICTION FORCE, THE LORENTZ FORCE, AND INDUCED DIFFUSION

Unfortunately, we failed to obtain an analytical solution for the matrices of the first-order corrections $\hat{\varphi}_i(\mathbf{r})$ at arbitrary angular momenta J [in contrast to $\hat{\sigma}(\mathbf{r})$]. Nevertheless, solving Eq. (21) for a given J involves no fundamental difficulties. The only problem is that the analytical expressions become increasingly cumbersome with increasing J. For this reason, below we present the coefficients $\mathscr{X}_{\kappa\kappa'}$ and $\mathscr{D}_{\kappa\kappa'}$ in explicit form for several transitions from the class $J \longrightarrow J+1$ with low values of J.

The linear (in velocity) friction force, which causes the mean atomic kinetic energy to change with time, is determined by the symmetric part of the tensor ξ_{ij}; therefore, the diagonal components $\mathscr{X}_{\kappa\kappa}$ and the sums of the nondiagonal components ($\mathscr{X}_{\kappa\kappa'} + \mathscr{X}_{\kappa'\kappa}$) contribute

to this force. The differences between the cross components ($\mathscr{K}_{\kappa\kappa'} - \mathscr{K}_{\kappa'\kappa}$) determine the Lorentz force, which does not change the kinetic energy.

In general, the induced diffusion tensor components $\mathscr{D}_{\kappa\kappa'}$ are not symmetric relative to the permutation of subscripts. However, as was noted above, only the symmetric components contribute to the Fokker–Planck equation. Therefore, we assume below that

$$\mathscr{D}_{\kappa\kappa'} = \mathscr{D}_{\kappa'\kappa} = (\mathscr{D}_{\kappa\kappa'} + \mathscr{D}_{\kappa'\kappa})/2.$$

7.1. An Arbitrary Saturation Parameter S. The 0 ⟶ 1 Transition

This simplest transition from the class $J \longrightarrow J + 1$ is of considerable interest from a practical point of view, because it is used for the laser cooling (to the Doppler limit) of calcium and strontium atoms. For arbitrary saturation parameters S, the coefficients $\mathscr{K}_{\kappa\kappa'}$ have the following form. The diagonal elements that are odd in detuning and even in ellipticity are

$$\mathscr{K}_{11} = \frac{2\tilde{\delta}S}{(1+2S)^3}\left[\frac{1}{1/4+\tilde{\delta}^2}(1-2S) - 8S^2\right], \quad (37)$$

$$\mathscr{K}_{22} = \frac{2\tilde{\delta}S}{(1/4+\tilde{\delta}^2)(1+2S)^2}, \quad (38)$$

$$\mathscr{K}_{33} = \frac{2\tilde{\delta}S}{(1/4+\tilde{\delta}^2)(1+2S)[1+S+(1/4+\tilde{\delta}^2)S^2]}, \quad (39)$$

$$\mathscr{K}_{44} = \frac{2\tilde{\delta}S}{(1/4+\tilde{\delta}^2)(1+2S)^2} \times \frac{1+(1+\cos^2(2\varepsilon))S+(1/4+\tilde{\delta}^2)(1-\cos^2(2\varepsilon))S^2}{1+S+(1/4+\tilde{\delta}^2)S^2}, \quad (40)$$

$$\mathscr{K}_{55} = \frac{1+\cos(2\varepsilon)}{2}\mathscr{K}_{33}, \quad \mathscr{K}_{66} = \frac{1-\cos(2\varepsilon)}{2}\mathscr{K}_{33}, \quad (41)$$

where $\tilde{\delta} = \delta/\gamma$ is the detuning, in units of the radiative width. Formulas (37) and (38) closely correspond to the results for the friction coefficient in the field of standing and traveling waves obtained in [16, 24] for a two-level model atom with nondegenerate states. Clearly, the scalar model is not enough to derive (39), (40), and (41); and the Zeeman structure of atomic energy levels and the vector nature of the electromagnetic field should be taken into account. The component $\mathscr{K}_{44}$ (40) for a zero ellipticity transforms to the well-known result for the $0 \longrightarrow 1$ transition in a σ_+–σ_- field configuration [30].

In addition, the linear (in velocity) force contains nonlinear contributions due to the combined action of various types of field gradients:

$$\mathscr{K}_{12} = \frac{2S[1-4\tilde{\delta}^2+2(1+4\tilde{\delta}^2)S]}{(1+2S)^2(1+4\tilde{\delta}^2)}, \quad (42)$$

$$\mathscr{K}_{21} = -\frac{2S(1-4\tilde{\delta}^2-4S)}{(1+2S)^3(1+4\tilde{\delta}^2)}, \quad (43)$$

$$\mathscr{K}_{14} = -\sin(2\varepsilon)\mathscr{K}_{12}, \quad \mathscr{K}_{41} = -\sin(2\varepsilon)\mathscr{K}_{21}, \quad (44)$$

$$\mathscr{K}_{24} = \mathscr{K}_{42} = -\sin(2\varepsilon)\mathscr{K}_{22}, \quad (45)$$

$$\mathscr{K}_{34} = -\mathscr{K}_{43} = -\frac{2S}{1+2S}\frac{\cos(2\varepsilon)}{1+S+(1/4+\tilde{\delta}^2)S^2} \times \left[\frac{1-4\tilde{\delta}^2}{1+4\tilde{\delta}^2} - \left(\frac{1}{4}+\tilde{\delta}^2\right)S^2\right], \quad (46)$$

$$\mathscr{K}_{56} = -\mathscr{K}_{65} = -\frac{S}{1+2S}\frac{\sin(2\varepsilon)}{1+S+(1/4+\tilde{\delta}^2)S^2} \times \left[\frac{1-4\tilde{\delta}^2}{1+4\tilde{\delta}^2} - \left(\frac{1}{4}+\tilde{\delta}^2\right)S^2\right], \quad (47)$$

which, as far as we know, are given for this transition for the first time.

As we see from (42)–(44), the contributions to the friction force from the correlation of amplitude–phase and amplitude–angle gradients have an even dependence on the detuning and do not vanish at $\tilde{\delta} = 0$. Consequently, in general, the friction coefficient has no definite parity in $\tilde{\delta}$ and laser cooling is possible even at exact resonance. We discussed the latter in a brief communication [13] for the $1/2 \longrightarrow 3/2$ transition in the field of a special one-dimensional configuration (see also Section 8).

Let us physically interpret some of the components $\mathscr{K}_{\kappa\kappa'}$. Specific field configurations (a traveling wave, a standing wave, the field produced by orthogonally polarized waves, etc.) are commonly considered to elucidate the physical friction mechanisms. However, this is not necessary, because all of the contributions from $\mathscr{K}_{\kappa\kappa'}$ can be interpreted locally, without associating the reasoning with any global field configuration. Thus, when considering the diagonal components $\mathscr{K}_{\kappa\kappa}$, it will suffice to assume that only one field gradient $\mathbf{g}^{(\kappa)}$ is nonzero at a given point of space, with the field amplitude and polarization being generally arbitrary. This approach leads to the mechanisms for $\mathscr{K}_{\kappa\kappa}$ that were previously known from the analysis of simple field configurations (Doppler, Sisyphus, and other friction mechanisms). To interpret the cross contributions from

$\mathscr{K}_{\kappa\kappa'}$, we must assume that two gradients, $\mathbf{g}^{(\kappa)}$ and $\mathbf{g}^{(\kappa')}$, are nonzero at a given point and consider their combined action.

For example, the nondiagonal element $\mathscr{K}_{12}$, which corresponds to the correlation of amplitude–phase gradients, can be interpreted as follows. The induced light pressure force $\hbar\mathscr{F}_1\mathbf{g}^{(1)}$, which depends on the field detuning from resonance [see (32)], acts on an atom in a field with an amplitude gradient. If, in addition, the atom moves in a field with a phase gradient, then the field frequency shifts $\delta \longrightarrow \delta - \mathbf{g}^{(2)} \cdot \mathbf{v}$ because of the Doppler effect. For the corrections of the first order in velocity, we have $-\hbar(\mathscr{F}_1)'_\delta \mathbf{g}^{(1)}(\mathbf{g}^{(2)} \cdot \mathbf{v})$, which exactly corresponds to (42). Thus, this contribution is attributable to the resonant frequency dependence of the induced light pressure force and the Doppler effect. In contrast, the other nondiagonal element $\mathscr{K}_{21}$ is due to the spontaneous light pressure force $\hbar\mathscr{F}_2(\mathbf{r}, \mathbf{p})\mathbf{g}^{(2)}$, where $\mathscr{F}_2(\mathbf{r}, \mathbf{p}) = \mathrm{Tr}\{\hat{\mathscr{F}}_2\hat{\sigma}(\mathbf{r}, \mathbf{p})\}$ is the mean value of the operator $\hat{\mathscr{F}}_2 = i[\Omega\hat{V} - \mathrm{H.c.}]$, which determines the absorption rate of external field photons. The first (in velocity) corrections are associated with the delayed (with respect to the slow field amplitude variations) part of the mean dipole moment for the moving atom. In particular, for the $0 \longrightarrow 1$ transition under consideration without saturation, the in-phase and delayed parts of the corresponding density matrix elements are

$$\hat{\sigma}_{eg}(\mathbf{r}, \mathbf{p}) \approx -i\frac{\Omega\hat{V}}{\gamma/2 - i\delta} + i\mathbf{g}^{(1)} \cdot \mathbf{v}\frac{\Omega\hat{V}}{(\gamma/2 - i\delta)^2},$$

$$\hat{\sigma}_{eg} = \hat{\Pi}_e\hat{\sigma}\hat{\Pi}_g,$$

which leads to (43), to within the first order in saturation parameter S.

Note that expressions (42) and (43) can be derived by using a two-level model atom. The Lorentz force due to the combined action of the amplitude and phase gradients in the field of a Gaussian beam was first considered by Kazantsev *et al.* [16]. Our result for the difference $(\mathscr{K}_{12} - \mathscr{K}_{21})$ matches that from [16] in the limit $S \ll 1$. Curiously, for a large saturation, $S \gg 1$, the component $\mathscr{K}_{21}$ decreases as $1/S$, while $\mathscr{K}_{12}$ ceases to depend on the parameters that characterize the intensity of the atom–field interaction and the corresponding contribution to the force $\hbar\mathbf{g}^{(1)}(\mathbf{g}^{(2)} \cdot \mathbf{v})$ is completely determined by the field geometry and the atomic velocity.

The components $\mathscr{K}_{\alpha\alpha'}$ $[(\alpha) = 5, 6]$ associated with the spatial rotation of the polarization ellipse were previously considered by one of the authors for $J_g > 0$ in the context of sub-Doppler cooling [31]. Here, we point out that the physical mechanisms responsible for the orientation contributions to the friction force ($\mathscr{K}_{44}$, on the one hand, and $\mathscr{K}_{\alpha\alpha'}$, on the other) are in many ways similar. Indeed, the variation of the total Hamiltonian for an atom in the case of an infinitesimal spatial displacement $\delta\mathbf{r}$ can be represented as a expansion in terms of field gradients:

$$\delta\hat{H}(\mathbf{r}) = -\hbar\sum_{\kappa=1}^{6}\mathbf{g}^{(\kappa)} \cdot \delta\mathbf{r}\hat{\mathscr{F}}_\kappa(\mathbf{r}),$$

where the variations related to rotation ($\kappa = 4, 5, 6$) are generated by the total angular momentum operator $\hat{\mathbf{J}}$:

$$\hat{\mathscr{F}}_\kappa(\mathbf{r}) = \frac{i}{\hbar}[\hat{J}_\kappa, \hat{H}(\mathbf{r})].$$

Here, $\hat{J}_\kappa$ are the corresponding components of the vector $\hat{\mathbf{J}}$ (along $\mathbf{e}_0$ for $\kappa = 4$ and along the principal axes of the ellipse for $\kappa = 5, 6$). According to the equations of motion, we can write the following relations for the mean values of the orientation force components $\mathscr{F}_\kappa(\mathbf{r}, \mathbf{p}) = \mathrm{Tr}\{\hat{\mathscr{F}}_\kappa(\mathbf{r})\,\hat{\sigma}(\mathbf{r}, \mathbf{p})\}$:

$$-\mathscr{F}_\kappa(\mathbf{r}, \mathbf{p}) = \mathrm{Tr}\{\hat{J}_\kappa\hat{\Gamma}\{\hat{\sigma}(\mathbf{r}, \mathbf{p})\}\} + \frac{d}{dt}J_\kappa(\mathbf{r}, \mathbf{p}). \quad (48)$$

These relations express the conservation of angular momentum: the absorption rate of external field angular momentum, $-\mathscr{F}_\kappa$, is the sum of the rate of angular momentum transfer to vacuum modes via spontaneous emission and the rate of change of the mean atomic angular momentum, $J_\kappa(\mathbf{r}, \mathbf{p}) = \mathrm{Tr}\{\hat{J}_\kappa\hat{\sigma}(\mathbf{r}, \mathbf{p})\}$. It is easy to show that the first term on the left-hand side of (48) is proportional to the excited-state angular momentum $J_\kappa^{(e)}(\mathbf{r}, \mathbf{p}) = \mathrm{Tr}\{\hat{J}_\kappa\hat{\sigma}_{ee}(\mathbf{r}, \mathbf{p})\}$:

$$\mathrm{Tr}\{\hat{J}_\kappa\hat{\Gamma}\{\hat{\sigma}(\mathbf{r}, \mathbf{p})\}\} = \gamma\mathscr{C}J_\kappa^{(e)}(\mathbf{r}, \mathbf{p}).$$

Here, the depolarization coefficient

$$\mathscr{C} = \frac{2 + J_e(J_e + 1) - J_g(J_g + 1)}{2J_e(J_e + 1)} = \frac{1}{J + 1}$$

is determined by the difference between the angular momentum lost by an atom as it spontaneously escapes from the excited state and the angular momentum gained by the atom through its spontaneous arrival to the ground state. For an atom at rest, $\mathbf{p} = 0$. Using the stationary solution $\hat{\sigma}(\mathbf{r})$ from [11], we obtain

$$\mathbf{J}^{(e)}(\mathbf{r}) = i\mathbf{e} \times \mathbf{e}^*\frac{SA(J + 1)}{\alpha_0 + 2\alpha_1 S},$$

in agreement with the result for $\mathscr{F}_4(\mathbf{r})$ (32) presented above. The first (in velocity) corrections in the equation for the atomic density matrix, which arise during the

motion in a field with orientation gradients are known [30, 1] to be described by the effective Hamiltonian

$$\hat{H}_{\text{eff}} = \sum_{\kappa=4}^{6} \mathbf{g}^{(\kappa)} \cdot \mathbf{v}\hat{J}_\kappa;$$

i.e., they are equivalent to the switching on of a fictitious magnetic field directed along the principal axes of the ellipse for $\mathbf{g}^{(5,6)}$ or orthogonally to the ellipse plane for $\mathbf{g}^{(4)}$. Thus, the orientation components $\mathscr{K}_{55}$ and $\mathscr{K}_{66}$, as well as $\mathscr{K}_{44}$, correspond to the magnetic susceptibility tensor components for optical pumping by elliptically polarized light. In addition, the second term on the left-hand side of (48) contributes to the cross components $\mathscr{K}_{56}$ and $\mathscr{K}_{65}$.

Let us now turn to the induced diffusion tensor. The diagonal components $\mathscr{D}_{\kappa\kappa'}$ are even functions of the detuning and ellipticity:

$$\mathscr{D}_{11} = \frac{S}{2(1+2S)^3}\times\left[1+\frac{6-8\tilde{\delta}^2}{1+4\tilde{\delta}^2}S+12S^2+8(1+4\tilde{\delta}^2)S^3\right],$$

$$\mathscr{D}_{22} = \frac{S}{2(1+2S)^3}\left[1-\frac{2-24\tilde{\delta}^2}{1+4\tilde{\delta}^2}S+4S^2\right],$$

$$\mathscr{D}_{33} = \frac{S}{2(1+2S)}, \tag{49}$$

$$\mathscr{D}_{44} = \frac{S}{2(1+2S)^3}\times\left[1-\frac{2-24\tilde{\delta}^2-(6-8\tilde{\delta}^2)\cos^2(2\varepsilon)}{1+4\tilde{\delta}^2}S+4S^2\right],$$

$$\mathscr{D}_{55} = \frac{1+\cos(2\varepsilon)}{2}\mathscr{D}_{33}, \quad \mathscr{D}_{66} = \frac{1-\cos(2\varepsilon)}{2}\mathscr{D}_{33}.$$

The nonzero nondiagonal elements break this symmetry:

$$\mathscr{D}_{12} = \frac{4\tilde{\delta}S^2}{(1+2S)^3}\left(\frac{2}{1+4\tilde{\delta}^2}+S\right),$$
$$\mathscr{D}_{14} = -\sin(2\varepsilon)\mathscr{D}_{12}, \tag{50}$$
$$\mathscr{D}_{24} = -\sin(2\varepsilon)\mathscr{D}_{22}.$$

The components $\mathscr{D}_{11}$, $\mathscr{D}_{22}$, and $\mathscr{D}_{12}$ attributable to the amplitude and phase gradients match those obtained previously in the scalar model of a two-level atom [24]. The orientation component $\mathscr{D}_{44}$ for the transition in question was considered in [30] for $\varepsilon = 0$. Without considering the physical interpretation of the various contributions to the induced diffusion tensor in detail, we only note that they are attributable to fluctuations of the induced and spontaneous light pressure forces, which generally correlate with each other.

The dependence of $\mathscr{K}_{\kappa\kappa'}$ and $\mathscr{D}_{\kappa\kappa'}$ on the saturation parameter S can be expressed as a quotient of two polynomials whose coefficients, in turn, depend on the detuning, ellipticity, and level angular momenta. The peculiarity of the $0 \longrightarrow 1$ transition associated with ground-state nondegeneracy is that the expansions in powers of S for $\mathscr{K}_{\kappa\kappa'}$, just as for $\mathscr{D}_{\kappa\kappa'}$, start from the first power. As we see from the above results, the contributions of higher orders in S are negligible in the low-saturation limit:

$$S \ll \min\{1, \gamma/\delta\}, \tag{51}$$

when the rate of atomic excitation from the ground state of the order of γS and the optical shifts of excited-state sublevels of the order of δS are much smaller than the rate of spontaneous decay γ. In this case, ξ_{ij} and $D_{ij}^{(\text{ind})}$ are linear in field intensity and we obtain the well-known results [32]:

$$\xi_{ij} \propto \text{Im}\left[\frac{\nabla_i \mathbf{E}^* \cdot \nabla_j \mathbf{E}}{(1/2 - i\tilde{\delta})^2}\right], \quad D_{ij}^{(\text{ind})} \propto \text{Re}\left[\frac{\nabla_i \mathbf{E}^* \cdot \nabla_j \mathbf{E}}{(1/2 - i\tilde{\delta})}\right].$$

Thus, the formulas given in this section generalize the previously obtained results for the $0 \longrightarrow 1$ transition to arbitrary field configurations and arbitrary saturation parameters.

7.2. The Low-Saturation Limit. The 1/2 ⟶ 3/2 Transition

For atomic transitions with a degenerate ground state $J_g > 0$ for low saturations (51), the first nonvanishing contributions to $\mathscr{K}_{\kappa\kappa'}$ have a zero order in S. This is known [1] to be fundamentally important for laser cooling below the Doppler limit. Several new peculiar properties of the coefficients $\mathscr{K}_{\kappa\kappa'}$ arise in this approximation. In particular, since the density matrix $\hat{\sigma}(\mathbf{r})$ in the zero order in saturation parameter S contains no phase and amplitude dependences, $\mathscr{K}_{\kappa 1} = \mathscr{K}_{\kappa 2} = 0$ for any $\kappa = 1, \ldots, 4$. In addition, the following relation holds between the coefficients:

$$\mathscr{K}_{1\kappa} = -2\tilde{\delta}\mathscr{K}_{2\kappa}. \tag{52}$$

It is a generalization of the relation for the force acting on an atom at rest (32): $\mathscr{F}_1(\mathbf{r}) = -2\tilde{\delta}\mathscr{F}_2(\mathbf{r})$. It can be shown that this relation between the induced and spon-

taneous light pressure forces also remains valid for a moving atom if we retain the terms of order $(v/S)^n$, the leading terms for $S \ll 1$, in each order of the expansion of the force in powers of the velocity. From a more general point of view, (52) is a corollary of the Kramers–Kronig relation for atomic susceptibility.

For the ground-state angular momentum $J_g = 1/2$, the matrix $\hat{\sigma}(\mathbf{r})$ in the zero order in S does not depend on the angle ϕ either. Therefore, in this special case, $\mathcal{X}_{\kappa 4} = 0$. The nonzero components $\mathcal{X}_{\kappa\kappa'}$ are

$$\mathcal{X}_{33} = 6\tilde{\delta}\cos^2(2\varepsilon), \quad \mathcal{X}_{13} = 6\tilde{\delta}\sin(2\varepsilon)\cos(2\varepsilon),$$
$$\mathcal{X}_{23} = -3\sin(2\varepsilon)\cos(2\varepsilon), \quad \mathcal{X}_{43} = 3\cos(2\varepsilon),$$
$$\mathcal{X}_{55} = 6\tilde{\delta}\frac{\sin^2(2\varepsilon)[3+\cos(2\varepsilon)]}{9-\cos^2(2\varepsilon)+36\tilde{\delta}^2\sin^2(2\varepsilon)}, \qquad (53)$$
$$\mathcal{X}_{66} = 6\tilde{\delta}\frac{\sin^2(2\varepsilon)[3-\cos(2\varepsilon)]}{9-\cos^2(2\varepsilon)+36\tilde{\delta}^2\sin^2(2\varepsilon)},$$

$$\mathcal{X}_{56} = -3\sin(2\varepsilon)[1+\cos(2\varepsilon)]$$
$$\times\frac{3+\cos(2\varepsilon)-12\tilde{\delta}^2[1-\cos(2\varepsilon)]}{2[9-\cos^2(2\varepsilon)+36\tilde{\delta}^2\sin^2(2\varepsilon)]}, \qquad (54)$$
$$\mathcal{X}_{65} = -3\sin(2\varepsilon)[1-\cos(2\varepsilon)]$$
$$\times\frac{3-\cos(2\varepsilon)-12\tilde{\delta}^2[1+\cos(2\varepsilon)]}{2[9-\cos^2(2\varepsilon)+36\tilde{\delta}^2\sin^2(2\varepsilon)]}.$$

The physical mechanism responsible for the components $\mathcal{X}_{13}$ and $\mathcal{X}_{33}$ is known from the analysis of a lin–θ–lin field configuration [2]—this is the so-called Sisyphus effect [1]. This effect is commonly associated with induced light pressure forces, which at low field intensities ($S \ll 1$) can be considered as resulting from the spatial nonuniformity of the optical ground-state sublevel shifts (the dynamic Stark effect). The optical shift is proportional to the field intensity and depends on ellipticity. In particular, for the $1/2 \longrightarrow 3/2$ transition in a local coordinate system with the z axis along $\mathbf{e}_0$,

$$U_{\pm 1/2} = \hbar\delta S\frac{2\pm\sin(2\varepsilon)}{3}.$$

It should be said that the presence of at least two distinct potentials is a necessary condition for the realization of the Sisyphus effect. The force components associated with the amplitude ($\mathbf{g}^{(1)}$) and ellipticity ($\mathbf{g}^{(3)}$) gradients are

$$\mathcal{F}_1(\mathbf{r},\mathbf{p}) = -\frac{4}{3}\delta S[1+\sin(2\varepsilon)\mathcal{J}(\mathbf{r},\mathbf{p})],$$

$$\mathcal{F}_3(\mathbf{r},\mathbf{p}) = -\frac{4}{3}\delta S\cos(2\varepsilon)\mathcal{J}(\mathbf{r},\mathbf{p}),$$

where $\mathcal{J}(\mathbf{r},\mathbf{p})$ is the mean z component of the atomic angular momentum (half the difference between the ground-state sublevel populations). Another necessary ingredient of the Sisyphus mechanism of sub-Doppler friction is a delay in the ground-state anisotropy for a slowly varying local ellipticity. In the case under consideration, the in-phase and delayed parts of the angular momentum, to within the first order in velocity, can be written as

$$\mathcal{J} \approx \frac{1}{2}\sin(2\varepsilon) - \frac{9}{2}\frac{\mathbf{g}^{(3)}\cdot\mathbf{v}}{\gamma S}\cos(2\varepsilon).$$

As we see, the characteristic delay is inversely proportional to the optical pumping rate γS, which leads to the intensity-independent contributions $\mathcal{X}_{13}$ and $\mathcal{X}_{33}$.

The interpretation of the $\tilde{\delta}$-independent components $\mathcal{X}_{23}$ and $\mathcal{X}_{43}$ is similar to that considered above, but with the significant distinction that the force action is provided by the spontaneous light pressure through phase and ϕ gradients:

$$\mathcal{F}_2(\mathbf{r},\mathbf{p}) = \frac{2}{3}\gamma S[1+\sin(2\varepsilon)\mathcal{J}(\mathbf{r},\mathbf{p})],$$

$$\mathcal{F}_4(\mathbf{r},\mathbf{p}) = \frac{1}{3}\gamma S[\sin(2\varepsilon)-2\mathcal{J}(\mathbf{r},\mathbf{p})].$$

Previously [13], using as an example the symmetric class of one-dimensional field configurations produced by counterpropagating, elliptically polarized waves of equal amplitudes, we showed that the detuning-independent friction coefficient $\mathcal{X}_{23}g_z^{(2)}g_z^{(3)} + \mathcal{X}_{43}g_z^{(4)}g_z^{(3)}$ does not vanish when averaged over the spatial period; i.e., there is a systematic effect of cooling (or heating, depending on the geometric parameters of the field configuration) at exact resonance $\delta = 0$. The spatial nonuniformity of the delay proportional to $(\gamma S)^{-1}$ was found to play a significant role.

The physical mechanism responsible for the orientation components (54) was discussed above and their influence on the sub-Doppler cooling was considered by one of the authors [31]. Here, we note only one significant point. In contrast to all the remaining components of the zero order in S, the components $\mathcal{X}_{55}$ and $\mathcal{X}_{66}$ do not become zero at points where the field polarization is circular; i.e., these components can dominate if the atoms are spatially localized near these points.

The coefficients $\mathcal{D}_{\kappa\kappa'}$ are linear in S in the low-saturation limit (51). For the $1/2 \longrightarrow 3/2$ transition, all 13 elements are nonzero. These include the four even

(in detuning) diagonal and two cross elements in sector {1–4},

$$\begin{aligned}
\mathcal{D}_{11} &= \frac{S}{6}[3 - \cos^2(2\varepsilon) + 12\cos^2(2\varepsilon)\sin^2(2\varepsilon)\tilde{\delta}^2],\\
\mathcal{D}_{22} &= \frac{S}{6}[3 + 2\cos^2(2\varepsilon) - 3\cos^4(2\varepsilon)],\\
\mathcal{D}_{33} &= \frac{S}{6}[1 + \cos^2(2\varepsilon) + 12\cos^4(2\varepsilon)\tilde{\delta}^2],\\
\mathcal{D}_{44} &= \frac{S}{6}[3 + 4\cos^2(2\varepsilon)],\\
\mathcal{D}_{13} &= \frac{S}{3}\sin(2\varepsilon)\cos(2\varepsilon)[1 + 12\cos^4(2\varepsilon)\tilde{\delta}^2],\\
\mathcal{D}_{24} &= -\frac{S}{3}\sin(2\varepsilon)[3 + 4\cos^2(2\varepsilon)];
\end{aligned} \tag{55}$$

the four odd (in $\tilde{\delta}$) nondiagonal elements in sector {1–4} attributable to the correlation between induced and spontaneous light pressure force fluctuations,

$$\begin{aligned}
\mathcal{D}_{12} &= -2\tilde{\delta}S\sin^2(2\varepsilon)\cos^2(2\varepsilon),\\
\mathcal{D}_{14} &= \frac{8\tilde{\delta}S}{3}\sin(2\varepsilon)\cos^2(2\varepsilon),\\
\mathcal{D}_{23} &= -2\tilde{\delta}S\sin(2\varepsilon)\cos^3(2\varepsilon),\\
\mathcal{D}_{34} &= \frac{8\tilde{\delta}S}{3}\cos^3(2\varepsilon);
\end{aligned} \tag{56}$$

and the three diagonal even and nondiagonal odd (in $\tilde{\delta}$) elements in sector {5, 6},

$$\begin{aligned}
\mathcal{D}_{55} &= \frac{S}{36}\left\{9 + 3\cos(2\varepsilon) - 2\cos^2(2\varepsilon)\right.\\
&\quad \left. + \frac{4\cos(2\varepsilon)[3 + \cos(2\varepsilon)]^3}{9 - \cos^2(2\varepsilon) + 36\tilde{\delta}^2\sin^2(2\varepsilon)}\right\},\\
\mathcal{D}_{66} &= \frac{S}{36}\left\{9 - 3\cos(2\varepsilon) - 2\cos^2(2\varepsilon)\right.\\
&\quad \left. - \frac{4\cos(2\varepsilon)[3 - \cos(2\varepsilon)]^3}{9 - \cos^2(2\varepsilon) + 36\tilde{\delta}^2\sin^2(2\varepsilon)}\right\},\\
\mathcal{D}_{56} &= -\frac{2\tilde{\delta}S}{3}\frac{\sin(2\varepsilon)\cos(2\varepsilon)[9 - \cos^2(2\varepsilon)]}{9 - \cos^2(2\varepsilon) + 36\tilde{\delta}^2\sin^2(2\varepsilon)}.
\end{aligned} \tag{57}$$

For a ground-state angular momentum $J_g \geq 1$, the density matrix $\hat{\sigma}(\mathbf{r})$ contains nondiagonal elements (Zeeman coherences) in the local basis with the quantization axis orthogonal to the polarization ellipse even in the zero order in S. Therefore, $\hat{\sigma}(\mathbf{r})$ also depends on ϕ and, in general, $\mathcal{K}_{\kappa 4} \neq 0$. The $\tilde{\delta}$ dependences of $\mathcal{K}_{\kappa\kappa'}$ and $\mathcal{D}_{\kappa\kappa'}$ have the form of a quotient of two polynomials whose coefficients depend on the ellipticity and angular momentum J in a complex way. The complete set of coefficients required to calculate the tensors ξ_{ij} and $D_{ij}^{(\mathrm{ind})}$ is given in Appendix B for the $1 \longrightarrow 2$ transition.

8. THE FIELD PRODUCED BY COUNTERPROPAGATING ELLIPTICALLY POLARIZED WAVES

Let us illustrate application of the results obtained above to the relatively simple but nontrivial case of one-dimensional field configurations simultaneously with two or more gradients (e.g., intensity and ellipticity gradients). These configurations are produced by counterpropagating waves with equal intensities and arbitrary elliptical polarizations $\mathbf{e}_1$ and $\mathbf{e}_2$; thus, they are a generalization of the (lin–θ–lin and σ_+–σ_-) field configurations widely used for laser cooling. Thus, we consider the ε_1–θ–ε_2 configuration: the principal axes of the polarization ellipses are oriented at an angle θ to each other, and ε_1 and ε_2 are the corresponding ellipticities. Denote the real amplitude of each wave by E_0. For counterpropagating waves of equal intensity, $\mathrm{Im}(\mathbf{E} \cdot \nabla_i \mathbf{E}^*) = 0$ and formula (A.5) give a rigid relation between the phase gradients and the rotation angle of the local polarization ellipse:

$$g^{(2)} = \sin(2\varepsilon)g^{(4)}, \tag{58}$$

where ε is the local field ellipticity and $g^{(\kappa)}$ is the gradient component along the direction of wave propagation (along the z axis). In this case, the coordinate dependence of the kinetic coefficients and their dependences on wave parameters are completely determined by two invariants, $\mathbf{E} \cdot \mathbf{E}$ and $\mathbf{E} \cdot \mathbf{E}^*$ [see (A.2)–(A.4)]:

$$\begin{aligned}
\mathbf{E} \cdot \mathbf{E}^* &= 2E_0^2[1 + \cos\theta\cos(\varepsilon_1 - \varepsilon_2)\cos(2kz)\\
&\quad - \sin\theta\sin(\varepsilon_1 + \varepsilon_2)\sin(2kz)],\\
\mathbf{E} \cdot \mathbf{E} &= E_0^2\{2\cos\theta\cos(\varepsilon_1 + \varepsilon_2) + [\cos(2\varepsilon_1)\\
&\quad + \cos(2\varepsilon_2)]\cos(2kz) + 2i\sin\theta\sin(\varepsilon_1 - \varepsilon_2)\\
&\quad + i[\cos(2\varepsilon_1) - \cos(2\varepsilon_2)]\sin(2kz)\}.
\end{aligned} \tag{59}$$

More specifically, we will consider one subclass of symmetric configurations for which the most interesting physical effects arise, namely, the ε–θ–$\bar{\varepsilon}$ field configuration where the ellipticity parameters of the coun-

terpropagating waves have opposite signs, $\varepsilon_1 = -\varepsilon_2 = \varepsilon_0$. According to (59), all of the possible field gradients are then nonzero:

$$\mathbf{E}\cdot\mathbf{E}^* = 2E_0^2[1+\cos(2\varepsilon_0)\cos\theta\cos(2kz)],$$
$$\sin(2\varepsilon) = -\frac{\sin\theta\cos(2\varepsilon_0)\sin(2kz)}{1+\cos(2\varepsilon_0)\cos\theta\cos(2kz)},$$
$$\tan(2\Phi) = \frac{\sin\theta\sin(2\varepsilon_0)}{\cos\theta+\cos(2\varepsilon_0)\cos(2kz)}.$$

8.1. The Light Pressure Force

For specific calculations, it is convenient to represent the light pressure force exerted on an atom at rest as

$$F = -\mathrm{Re}\left\{\frac{\hbar(\delta+i\gamma/2)S}{\alpha_0+2S\alpha_1}\right.$$
$$\left.\times[(\alpha_1-A)\nabla_z\ln(\mathbf{E}\cdot\mathbf{E})+A\nabla_z\ln(\mathbf{E}\cdot\mathbf{E}^*)]\right\}, \quad (60)$$

because the ratios of the coefficients α_0, α_1, and A depend only on even powers of $\cos(2\varepsilon)$. In particular, for the 1/2 ⟶ 3/2 transition in the low-saturation limit, we have

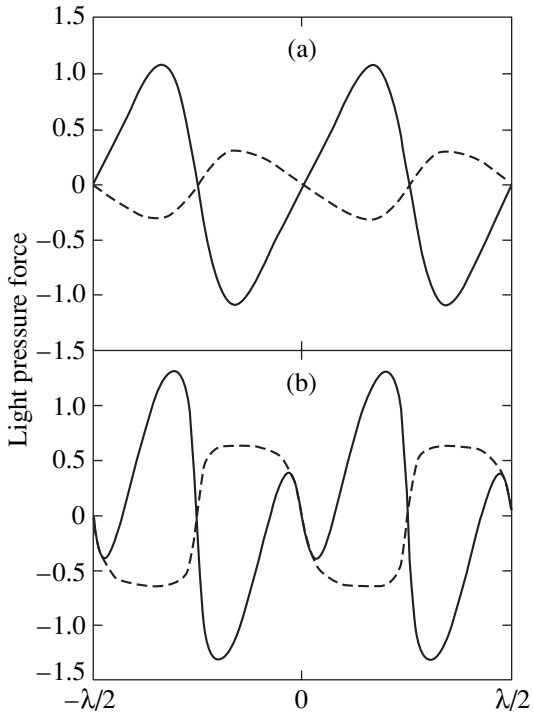

Fig. 1. Spatial dependence of the induced (solid lines) and spontaneous (dashed lines) light pressure forces in a weak ($S_0 \ll 1$) field of ε–θ–$\bar{\varepsilon}$ configuration for the (a) 1/2 ⟶ 3/2 and (b) 4 ⟶ 5 transitions. The spontaneous light pressure force $\hbar(\mathscr{F}_2 g^{(2)} + \mathscr{F}_4 g^{(4)})$ is in units of $\hbar k\gamma S_0$, and the induced light pressure force $\hbar(\mathscr{F}_1 g^{(1)} + \mathscr{F}_3 g^{(3)})$ is in units of $\hbar k\delta S_0$. The field configuration parameters are $\varepsilon_0 = \pi/8$ and $\theta = \pi/3$.

$$F = 2\hbar k S_0\cos(2\varepsilon_0)\sin(2kz)\frac{\delta[4\cos\theta+(1+3\cos\theta)\cos(2\varepsilon_0)\cos(2kz)]-\gamma\sin\theta\cos(2\varepsilon_0)}{3[1+\cos\theta\cos(2\varepsilon_0)\cos(2kz)]},$$

where the saturation parameter S_0 corresponds to the intensity of one wave.

It follows from symmetry relations [23] that the force (60) averaged over the spatial period becomes zero; i.e., the force corresponds to a periodic potential. A characteristic feature of this configuration is that a periodic potential is produced by both the induced and spontaneous light pressure forces (see Fig. 1). In principle, this behavior allows the spontaneous force to be used to produce optical gratings in the case of exact resonance where the induced force becomes zero and the spontaneous force reaches a maximum.

8.2. The Friction Coefficient

Consider the 1/2 ⟶ 3/2 transition in the low-saturation limit $S \ll 1$, where cooling below the Doppler limit becomes possible in a field with polarization gradients [1]. The friction coefficient is defined by four components (53). For the $\tilde{\delta}$-independent anomalous components to appear, the phase and ellipticity or rotation angle and ellipticity gradients must be simultaneously present. Both are realized in the ε–θ–$\bar{\varepsilon}$ configuration at $\theta \neq 0$. Using (59) and (A.2)–(A.4), we can derive explicit expressions for the normal, ξ_n, and anomalous, ξ_a, friction coefficients in terms of the ellipticity ε_0 of the forming waves and the angle θ:

$$\xi_n = \hbar[\mathscr{X}_{33}(g^{(3)})^2 + \mathscr{X}_{13}g^{(1)}g^{(3)}]$$
$$= \hbar\tilde{\delta}\left[\frac{3}{2[1-\cos^2(2\varepsilon)]}(\tilde{g}^{(3)})^2 + 3g^{(1)}\tilde{g}^{(3)}\right]$$
$$= \hbar k^2\tilde{\delta}6\sin^2\theta\cos^2(2\varepsilon_0)\cos(2kz)$$
$$\times[\cos(2kz)+\cos\theta\cos(2\varepsilon_0)]$$
$$\times[1+\cos\theta\cos(2\varepsilon_0)\cos(2kz)]^{-3}, \quad (61)$$

$$\begin{aligned}\xi_a &= \hbar[\mathcal{K}_{23}g^{(2)}g^{(3)} + \mathcal{K}_{43}g^{(4)}g^{(3)}]\\ &= \hbar\left[-\frac{3}{2}g^{(2)}\tilde{g}^{(3)} + \frac{3}{2[1-\cos^2(2\varepsilon)]}\tilde{g}^{(4)}\tilde{g}^{(3)}\right]\\ &= \hbar k^2 3\sin\theta\sin(2\varepsilon_0)\cos(2\varepsilon_0)\\ &\times[\cos(2kz) + \cos\theta\cos(2\varepsilon_0)]\\ &\times[1 + \cos\theta\cos(2\varepsilon_0)\cos(2kz)]^{-3}.\end{aligned} \tag{62}$$

In general, when averaged over the field period, the two components are nonzero:

$$\langle\xi_n\rangle = \hbar k^2\tilde{\delta}\frac{3\sin^2\theta\cos^2(2\varepsilon_0)}{[1-\cos^2\theta\cos^2(2\varepsilon_0)]^{3/2}}, \tag{63}$$

$$\langle\xi_a\rangle = -\hbar k^2\frac{3}{2}\frac{\sin\theta\cos\theta\sin(2\varepsilon_0)\cos^2(2\varepsilon_0)}{[1-\cos^2\theta\cos^2(2\varepsilon_0)]^{3/2}}. \tag{64}$$

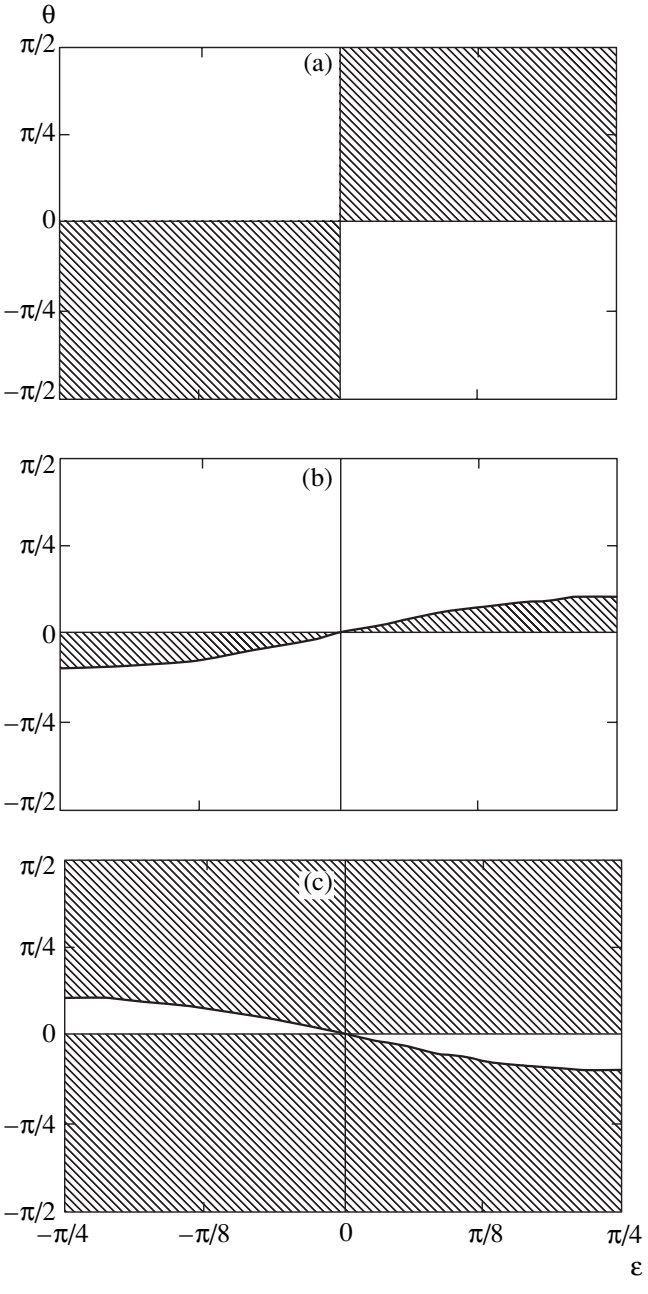

Fig. 2. The heating (blank) and cooling (hatched) regions for the detunings $\delta = 0$ (a), 3γ (b), and -3γ (c).

8.3. The Direction of the Kinetic Process

The anomalous friction force qualitatively manifests itself in changes of the direction (heating or cooling) and rate of the kinetic process. Perhaps the most glaring example is an exact resonance, $\delta = 0$, when the normal friction force becomes zero. Disregarding the anomalous component, one might expect a slow (linear in time) temperature rise due to atomic momentum diffusion. However, if $\varepsilon_0 \neq 0, \pm\pi/4$ and $\theta \neq 0, \pi/2$, then $\langle\xi_a\rangle \neq 0$ and an atom is, on average, under the action of nonconservative forces even at $\delta = 0$. The sign of $\langle\xi_a\rangle$ is determined by the product $\sin\theta\sin(2\varepsilon_0)$. As a result, we can distinguish the domains of θ and ε_0 (see Fig. 2) where the friction coefficient is negative and where cooling takes place and the domains where the friction coefficient is positive, which corresponds to a rapid (exponential) temperature rise.

The anomalous friction force can be also detected at nonzero detunings. Thus, for example, for $\delta < 0$, one might expect cooling to take place in the entire domain of θ and ε_0. However, because of the anomalous component, the friction coefficient reverses its sign when crossing the line

$$\frac{\delta}{\gamma}\tan\theta = -\sin(2\varepsilon_0) \tag{65}$$

and an exponentially rapid temperature rise takes place in the region bounded by this line and the $\theta = 0$ axis. In contrast, at negative detunings, the friction coefficient is negative in the region bounded by line (65) and the $\theta = 0$ axis; i.e., cooling and heating take place inside and outside this domain, respectively. This behavior can be explained by the fact that the local field polarization for $\theta = 0$ is linear everywhere and there is no sublevel population difference nor difference in the corresponding optical potentials necessary for the Sisyphus friction mechanism. As a result, the normal components tend to zero, $\langle\xi_n\rangle \propto \theta^2$, as $\theta \longrightarrow 0$ ($\varepsilon_0 \neq 0$). For the anomalous component, only the first of the two causes remains, which results in a slower decrease, $\langle\xi_a\rangle \propto \theta$.

Explicit expressions for the diffusion coefficients can be calculated in a similar but more cumbersome way. Here, we do not consider this issue.

9. CONCLUSION

Let us summarize our results. The kinetics of slow atoms in a nonuniformly polarized field is described in the semiclassical approximation by the Fokker–Planck equation for the atomic distribution function in phase space. The coefficients of this equation (the force in the zero and first orders in velocity and the diffusion in the zero order) were represented as expansions in terms of spatial field gradients. For closed atomic transitions from the class $J \longrightarrow J + 1$, we derived analytical expressions for the force and diffusion coefficient that allow the atomic motion in a field of arbitrary configuration to be considered. For the force exerted on an atom at rest and for the spontaneous diffusion, we obtained compact analytical results, which are valid for an arbitrary angular momentum J. Since the analytical expressions for the first (in velocity) corrections to the force and for the induced diffusion become cumbersome with increasing J, we restricted our analysis to a series of transitions with small angular momenta, $J = 0$, 1/2, and 1. The friction force and the induced diffusion tensor were shown to generally contain the cross contributions due to the combined action of the various types of gradients. As a result, the friction force and the diffusion tensor have no definite parity relative to the change of sign of the field detuning δ from resonance. We considered a qualitative interpretation of the physical mechanisms that give rise to cross contributions to the friction. We found that these mechanisms could be assigned to a mixed type—the Doppler friction through induced light pressure forces ($\mathscr{K}_{12}$) and the Sisyphus effect in the presence of spontaneous light pressure forces ($\mathscr{K}_{23}$ and $\mathscr{K}_{43}$).

In conclusion, note that our analytical results for the coefficients of the Fokker–Planck equation give a solution (in a general form that does not depend on the specific field configuration) of the first half of the problem of theoretically describing the motion of atoms in the semiclassical approximation. The second half must involve an analysis of the (stationary and dynamic) solutions to the Fokker–Planck equation. To all appearances, such an analysis, qualitative or quantitative, is possible only when the spatial field configuration is specified. The well-known [1, 15, 16, 24] analytical estimates of the cooling temperature based on fluctuation–dissipation relations using spatially averaged diffusion and friction coefficients constitute the simplest case. Clearly, more sophisticated methods must include the spatial localization of atoms (see, e.g., [16, 33]) and the complex pattern of motion in a nonuniformly polarized field.

ACKNOWLEDGMENTS

This study was supported by the Russian Foundation for Basic Research (project nos. 01-02-17036, 01-02-17744, 02-02-06513, and 03-02-16513) and the Ministry of Education of Russia (project nos. E00-3.2-153 and UR.01.01.062). One of us (O.N.P.) was supported by an INTAS Grant For Young Scientists (YSF 00-129).

APPENDIX A

The natural parametrization (29) that we used is convenient for writing the analytical results and for qualitatively interpreting the various contributions. However, the final results for the force vector $\mathbf{F}$ and the tensors ξ_{ij} and D_{ij} can always be expressed in an invariant and analytical way in terms of the complex vector amplitude of the field $\mathbf{E}(\mathbf{r})$ and its spatial gradients. In this case, we can completely avoid taking the nth root. It is well known that the latter operation can lead to ambiguities in the phase factors. In this sense, the degree of linear polarization, $\cos(2\varepsilon)$, defined as

$$\cos(2\varepsilon) = \frac{\sqrt{|\mathbf{E}\cdot\mathbf{E}|^2}}{\mathbf{E}\cdot\mathbf{E}^*}, \tag{A.1}$$

contains this ambiguity. For example, the nonnegative value of the root could be used everywhere, which is not quite convenient, because, in general, this function is not smooth. Such reasoning is also valid for $\sin(2\varepsilon)$, the degree of circular polarization. At the same time, the square of the degree of linear polarization

$$\cos^2(2\varepsilon) = \frac{|\mathbf{E}\cdot\mathbf{E}|^2}{(\mathbf{E}\cdot\mathbf{E}^*)^2} \tag{A.2}$$

does not have this drawback.

The transformation to the generally invariant expressions for $\mathbf{F}$, ξ_{ij}, and D_{ij} can be made as follows. The gradients $\mathbf{g}^{(1)}$ and $\mathbf{g}^{(2)}$ can be expressed as

$$\begin{aligned}\mathbf{g}^{(1)} &= \frac{1}{2}\frac{\nabla(\mathbf{E}\cdot\mathbf{E}^*)}{\mathbf{E}\cdot\mathbf{E}^*},\\ \mathbf{g}^{(2)} &= \frac{i}{4}\frac{\mathbf{E}\cdot\mathbf{E}}{(\mathbf{E}\cdot\mathbf{E})^*}\nabla\frac{(\mathbf{E}\cdot\mathbf{E})^*}{\mathbf{E}\cdot\mathbf{E}}.\end{aligned} \tag{A.3}$$

Next, we should use

$$\tilde{\mathbf{g}}^{(3)} = \nabla\cos^2(2\varepsilon) \tag{A.4}$$

in place of the ellipticity gradient $\mathbf{g}^{(3)}$ and the product $\tilde{\mathbf{g}}^{(4)} = \sin(2\varepsilon)\mathbf{g}^{(4)}$ in place of the angle gradient $\mathbf{g}^{(4)}$. For

the Cartesian coordinates of the latter, we have the expression

$$\tilde{g}_i^{(4)} = \frac{\mathrm{Im}(\mathbf{E}\cdot\nabla_i\mathbf{E}^*)}{\mathbf{E}\cdot\mathbf{E}^*} + g_i^{(2)}, \tag{A.5}$$

equivalent to (25). It is easy to verify that after changing to the new gradients, all of the coefficients $\mathcal{F}_\kappa$, $\mathcal{K}_{\kappa\kappa'}$, and $\mathcal{D}_{\kappa\kappa'}$ in sector {1–4} will have an ellipticity dependence only in the form of even powers of cos(2ε); i.e., the ambiguity associated with taking the root will be removed.

The situation in sector {5, 6} is more complex, because no unique analytical expressions can be obtained for the linear combinations of the gradients $\mathbf{g}^{(5)}$ and $\mathbf{g}^{(6)}$. Here, two circumstances play a significant role. First, since $\mathcal{F}_5 = \mathcal{F}_6 = 0$, we need only quadratic combinations of the gradients in the form $g_i^{(\alpha)} g_j^{(\alpha')}$. Second, the coefficients $\mathcal{K}_{\alpha\alpha'}$ and $\mathcal{D}_{\alpha\alpha'}$ in sector {5, 6} contain both even and odd powers of cos(2ε). In this case, the following properties of the symmetry relative to the change of sign of cos(2ε) hold:

$$\begin{aligned}\mathcal{K}_{55}(-\cos(2\varepsilon)) &= \mathcal{K}_{66}(\cos(2\varepsilon)),\\ \mathcal{K}_{56}(-\cos(2\varepsilon)) &= -\mathcal{K}_{65}(\cos(2\varepsilon))\end{aligned} \tag{A.6}$$

(similarly for $\mathcal{D}_{\alpha\alpha'}$), which predict that analytical invariant expressions can be derived for the combinations

$$\begin{aligned}G_{ij}^{(1)} &= g_i^{(5)}g_j^{(5)} + g_i^{(6)}g_j^{(6)},\\ G_{ij}^{(2)} &= \cos(2\varepsilon)(g_i^{(5)}g_j^{(5)} - g_i^{(6)}g_j^{(6)}),\\ G_{ij}^{(3)} &= \sin(2\varepsilon)(g_i^{(5)}g_j^{(6)} - g_i^{(6)}g_j^{(5)}),\\ G_{ij}^{(4)} &= \sin(2\varepsilon)\cos(2\varepsilon)(g_i^{(5)}g_j^{(6)} + g_i^{(6)}g_j^{(5)}).\end{aligned} \tag{A.7}$$

Indeed, consider the two tensors

$$\begin{aligned}P_{ij} &= 2(\mathbf{e}_0\cdot\nabla_i\mathbf{e})(\mathbf{e}_0\cdot\nabla_j\mathbf{e}^*)\\ &= G_{ij}^{(1)} + G_{ij}^{(2)} - iG_{ij}^{(3)},\\ R_{ij} &= 2(\tilde{\mathbf{e}}\cdot\nabla_i\mathbf{e}_0)(\tilde{\mathbf{e}}^*\cdot\nabla_j\mathbf{e}_0)\\ &= G_{ij}^{(1)} - G_{ij}^{(2)} + iG_{ij}^{(3)},\end{aligned} \tag{A.8}$$

where $\tilde{\mathbf{e}}$ is the third unit vector of the complex triad: $\tilde{\mathbf{e}}^* = \mathbf{e}_0 \times \mathbf{e}$. Clearly, each of the tensors P_{ij} and R_{ij} can be uniquely expressed in terms of $\mathbf{E}(\mathbf{r})$ and its derivatives (we do not give explicit expressions, because they are cumbersome), because it is composed of an even number of vectors $\mathbf{e}_0$ and an equal number of vectors $\mathbf{e}$ and $\mathbf{e}^*$; i.e., there is no need to use ambiguous expressions for sin(2ε) and exp($i\Phi$). The tensors $G_{ij}^{(1)}$ and $G_{ij}^{(2)}$ are equal to the half-sum and half-difference of the real parts of (A.8), and $G_{ij}^{(3)}$ corresponds to the imaginary part. To derive an analytical expression for the tensor $G_{ij}^{(4)}$, we can take the symmetric part of the tensor

$$\begin{aligned}S_{ij} &= -2\sin(2\varepsilon)\mathrm{Re}\{(\mathbf{e}_0\cdot\nabla_i\mathbf{e})(\tilde{\mathbf{e}}\cdot\nabla_j\mathbf{e}_0)\}\\ &= G_{ij}^{(3)} + G_{ij}^{(4)},\end{aligned} \tag{A.9}$$

where the factor sin(2ε) compensates for the odd number of the vectors $\mathbf{e}_0$ used to construct S_{ij}.

APPENDIX B

In this appendix, we give analytical expressions for the coefficients $\mathcal{K}_{\kappa\kappa'}$ and $\mathcal{D}_{\kappa\kappa'}$ for the $1 \longrightarrow 2$ transition in the low-saturation limit (51).

In the zero order in S, the nonzero components $\mathcal{K}_{\kappa\kappa'}$ in sector {1–4} are

$$\begin{aligned}\mathcal{K}_{33} &= 40\tilde{\delta}\frac{625 + 5200c^2 + 13728c^4 - 768c^6 + 128c^2(125 + 259c^2 - 384c^4)\tilde{\delta}^2}{(25 - 8c^2)^3[5 + 4(5 - 4c^2)\tilde{\delta}^2]},\\ \mathcal{K}_{44} &= \tilde{\delta}\frac{120c^2(15 + 2c^2)}{(25 - 8c^2)^2[5 + 4(5 - 4c^2)\tilde{\delta}^2]},\\ \mathcal{K}_{13} &= 160sc\tilde{\delta}\frac{5(275 + 684c^2 - 32c^4) + 8(625 + 1040c^2 - 1584c^4)\tilde{\delta}^2}{(25 - 8c^2)^3[5 + 4(5 - 4c^2)\tilde{\delta}^2]},\\ \mathcal{K}_{24} &= -\tilde{\delta}\frac{160sc^2(5 + 4c^2)}{(25 - 8c^2)^2[5 + 4(5 - 4c^2)\tilde{\delta}^2]},\end{aligned} \tag{B.1}$$

$$\mathcal{X}_{23} = -80sc\frac{5(275+684c^2-32c^4)+8(625+1040c^2-1584c^4)\tilde{\delta}^2}{(25-8c^2)^3[5+4(5-4c^2)\tilde{\delta}^2]},$$

$$\mathcal{X}_{43} = -10c\frac{11625+26260c^2-1760c^4+4(9375+14140c^2-25984c^4+1024c^6)\tilde{\delta}^2}{(25-8c^2)^3[5+4(5-4c^2)\tilde{\delta}^2]},$$

$$\mathcal{X}_{34} = -10c\frac{25-8c^2-4(25+28c^2+32c^4)\tilde{\delta}^2}{(25-8c^2)^2[5+4(5-4c^2)\tilde{\delta}^2]}, \tag{B.2}$$

$$\mathcal{X}_{14} = \frac{320sc^2(5+4c^2)\tilde{\delta}^2}{(25-8c^2)^2[5+4(5-4c^2)\tilde{\delta}^2]}.$$

Below, to save space, we use the notation $c = \cos(2\varepsilon)$ and $s = \sin(2\varepsilon)$. It is convenient to represent the orientation components $\mathcal{X}_{\alpha\alpha'}$ in sector {5, 6} in symmetrized form:

$$\mathcal{X}_{55}+\mathcal{X}_{66} = \frac{40\tilde{\delta}[45-34c^2-8c^4+16s^4(5-c^2)\tilde{\delta}^2]}{(25-8c^2)[9-4c^2+(52-48c^2)\tilde{\delta}^2+64s^4\tilde{\delta}^4]},$$

$$\mathcal{X}_{55}-\mathcal{X}_{66} = \frac{40c\tilde{\delta}[18-11c^2-4c^4+32s^4\tilde{\delta}^2]}{(25-8c^2)[9-4c^2+(52-48c^2)\tilde{\delta}^2+64s^4\tilde{\delta}^4]},$$

$$\mathcal{X}_{56}-\mathcal{X}_{65} = -10s\frac{45+10c^2+8c^4-20(5-4c^4)\tilde{\delta}^2-320s^4\tilde{\delta}^4}{(25-8c^2)[9-4c^2+(52-48c^2)\tilde{\delta}^2+64s^4\tilde{\delta}^4]}, \tag{B.3}$$

$$\mathcal{X}_{56}+\mathcal{X}_{65} = -20cs\frac{27+4c^2+16s^2(3+c^2)\tilde{\delta}^2}{(25-8c^2)[9-4c^2+(52-48c^2)\tilde{\delta}^2+64s^4\tilde{\delta}^4]}.$$

The components $D_{\kappa\kappa'}$, which define the induced diffusion tensor, can be written, in the first order in saturation parameter S, as

$$\mathcal{D}_{11} = \frac{S}{\mathcal{N}}[25(25-8c^2)^2(5-3c^2)+20(15625-26875c^2+29380c^4-18688c^6+1280c^8)\tilde{\delta}^2$$
$$+128c^2(3125+1025c^2-10660c^4+6528c^6)\tilde{\delta}^4],$$

$$\mathcal{D}_{22} = \frac{S}{\mathcal{N}}[5(15625-14375c^2+13880c^4-12608c^6+512c^8)$$
$$+4(78125-134375c^2+123700c^4-12480c^6+56064c^8)\tilde{\delta}^2],$$

$$\mathcal{D}_{33} = \frac{S}{\mathcal{N}}[(25-8c^2)^2(25+23c^2)+4(15625-3125c^2+8900c^4+64384c^6-4864c^8)\tilde{\delta}^2$$
$$+2048c^4(125+259c^2-384c^4)\tilde{\delta}^4],$$

$$\mathcal{D}_{44} = \frac{S}{\mathcal{N}}[25(3125-1875c^2+3966c^4-592c^6) \tag{B.4}$$
$$+4(78125-109375c^2+120450c^4-98440c^6+11552c^8)\tilde{\delta}^2],$$

$$\mathcal{D}_{13} = 4sc\frac{S}{\mathcal{N}}[25(25-8c^2)^2+4(15625-10500c^2+37440c^4-2816c^6)\tilde{\delta}^2$$
$$+256c^2(625+1040c^2-1584c^4)\tilde{\delta}^4],$$

$$\mathcal{D}_{24} = -2s\frac{S}{\mathcal{N}}[5(15625 - 5625c^2 + 15480c^4 - 1664c^6)$$

$$+ 4(78125 - 90625c^2 + 92700c^4 - 76960c^6 + 5888c^8)\tilde{\delta}^2],$$

where $\mathcal{N} = 2(25 - 8c^2)^3[5 + 4(5 - 4c^2)\tilde{\delta}^2]$,

$$\mathcal{D}_{12} = -\tilde{\delta}S\frac{32c^2[5(625 + 785c^2 - 1456c^4 + 64c^6) + 4(3125 + 1025c^2 - 10660c^4 + 6528c^6)\tilde{\delta}^2]}{(25 - 8c^2)^3[5 + 4(5 - 4c^2)\tilde{\delta}^2]},$$

$$\mathcal{D}_{14} = \tilde{\delta}S\frac{48sc^2[3125 + 4950c^2 - 560c^4 + 4(3125 + 1850c^2 - 5080c^4 + 384c^6)\tilde{\delta}^2]}{(25 - 8c^2)^3[5 + 4(5 - 4c^2)\tilde{\delta}^2]},$$

$$\mathcal{D}_{23} = -\tilde{\delta}S\frac{128sc^3[3(125 + 290c^2 - 16c^4) + 2(625 + 1040c^2 - 1584c^4)\tilde{\delta}^2]}{(25 - 8c^2)^3[5 + 4(5 - 4c^2)\tilde{\delta}^2]}, \quad \text{(B.5)}$$

$$\mathcal{D}_{34} = \tilde{\delta}S\frac{48c^3[1625 + 2420c^2 - 288c^4 + 8(625 + 535c^2 - 1256c^4 + 96c^6)\tilde{\delta}^2]}{(25 - 8c^2)^3[5 + 4(5 - 4c^2)\tilde{\delta}^2]},$$

$$\mathcal{D}_{55} + \mathcal{D}_{66} = \frac{S}{2(25 - 8c^2)[9 - 4c^2 + (52 - 48c^2)\tilde{\delta}^2 + 64s^4\tilde{\delta}^4]}$$

$$\times [225 + 143c^2 + 16c^4 + 16c^6 + 4(325 - 309c^2 - 80c^4 + 72c^6)\tilde{\delta}^2 + 64s^4(25 - 13c^2)\tilde{\delta}^4],$$

$$\mathcal{D}_{55} - \mathcal{D}_{66} = \frac{cS}{2(25 - 8c^2)[9 - 4c^2 + (52 - 48c^2)\tilde{\delta}^2 + 64s^4\tilde{\delta}^4]} \quad \text{(B.6)}$$

$$\times [405 - 57c^2 + 52c^4 + 4(225 - 301c^2 + 68c^4 + 16c^6)\tilde{\delta}^2 + 64s^4(5 - c^2)\tilde{\delta}^4],$$

$$\mathcal{D}_{56} = -\tilde{\delta}S\frac{4sc[45 - 6c^2 + 4(20 - 31c^2 + 11c^4)\tilde{\delta}^2]}{(25 - 8c^2)[9 - 4c^2 + (52 - 48c^2)\tilde{\delta}^2 + 64s^4\tilde{\delta}^4]}.$$

REFERENCES

1. J. Dalibard and C. Cohen-Tannoudji, J. Opt. Soc. Am. B **6**, 2023 (1989).
2. V. Finkelstein, P. Berman, and J. Guo, Phys. Rev. A **45**, 1829 (1992).
3. A. V. Bezverbnyi, A. M. Tumaikin, and N. L. Kosulin, Laser Phys. **2**, 1010 (1992).
4. J. Guo and P. Berman, Phys. Rev. A **48**, 3225 (1993).
5. J. Werner, H. Wallis, G. Hillenbrand, and A. Steane, J. Phys. B **26**, 3063 (1993).
6. Y. Castin and J. Dalibard, Europhys. Lett. **14**, 761 (1991); J. Dalibard, Y. Castin, and K. Mølmer, Phys. Rev. Lett. **68**, 580 (1992); R. Taïeb, P. Marte, R. Dum, and P. Zoller, Phys. Rev. A **47**, 4986 (1993); K. Mølmer, Y. Castin, and J. Dalibard, J. Opt. Soc. Am. B **10**, 524 (1993); K. Berg-Sørensen, Phys. Rev. A **49**, R4297 (1994); Y. Castin, K. Berg-Sørensen, J. Dalibard, and K. Mølmer, Phys. Rev. A **50**, 5092 (1994); K. Mølmer and Y. Castin, Quantum Semiclassic. Opt. **8**, 49 (1996).
7. A. V. Taichenachev, A. M. Tumaikin, V. I. Yudin, and G. Nienhuis, Zh. Éksp. Teor. Fiz. **108**, 415 (1995) [JETP **81**, 224 (1995)].
8. A. V. Taichenachev, A. M. Tumaikin, and V. I. Yudin, Pis'ma Zh. Éksp. Teor. Fiz. **64**, 8 (1996) [JETP Lett. **64**, 7 (1996)].
9. A. V. Taichenachev, A. M. Tumaikin, and V. I. Yudin, Zh. Éksp. Teor. Fiz. **110**, 1727 (1996) [JETP **83**, 949 (1996)].
10. A. V. Taichenachev, A. M. Tumaikin, V. I. Yudin, and G. Nienhuis, Zh. Éksp. Teor. Fiz. **114**, 125 (1998) [JETP **87**, 70 (1998)].
11. G. Nienhuis, A. V. Taichenachev, A. M. Tumaikin, and V. I. Yudin, Europhys. Lett. **44**, 20 (1998).
12. O. N. Prudnikov, A. V. Taichenachev, A. M. Tumaikin, and V. I. Yudin, Zh. Éksp. Teor. Fiz. **115**, 791 (1999) [JETP **88**, 433 (1999)].
13. O. N. Prudnikov, A. V. Taichenachev, A. M. Tumaikin, and V. I. Yudin, Pis'ma Zh. Éksp. Teor. Fiz. **70**, 439 (1999) [JETP Lett. **70**, 443 (1999)].
14. A. V. Bezverbnyi, Zh. Éksp. Teor. Fiz. **118**, 1066 (2000) [JETP **91**, 921 (2000)].
15. V. G. Minogin and V. S. Letokhov, *Laser Light Pressure on Atoms* (Nauka, Moscow, 1986; Gordon and Breach, New York, 1987).

16. A. P. Kazantsev, G. I. Surdutovich, and V. P. Yakovlev, *The Mechanical Action of Light on Atoms* (Nauka, Moscow, 1991; World Sci., Singapore, 1990).
17. P. D. Lett, R. N. Watts, C. I. Westbrook, *et al.*, Phys. Rev. Lett. **61**, 169 (1988); D. S. Weiss, E. Riis, Y. Shevy, *et al.*, J. Opt. Soc. Am. B **6**, 2072 (1989); C. Salomon, J. Dalibard, W. D. Phillips, *et al.*, Europhys. Lett. **12**, 683 (1990).
18. S. G. Rautian, G. I. Smirnov, and A. M. Shalagin, *Nonlinear Resonances in Atomic and Molecular Spectra* (Nauka, Novosibirsk, 1979).
19. V. G. Minogin, Zh. Éksp. Teor. Fiz. **79**, 2044 (1980) [Sov. Phys. JETP **52**, 1032 (1980)].
20. J. Dalibard and C. Cohen-Tannoudji, J. Phys. B **18**, 1661 (1985).
21. J. Javanainen, Phys. Rev. A **44**, 5857 (1991).
22. S. M. Yoo and J. Javanainen, J. Opt. Soc. Am. B **8**, 1341 (1991).
23. A. V. Taichenachev, A. M. Tumaikin, and V. I. Yudin, *Advance Program ICONO'01* (Minsk, 2001), p. 64.
24. J. P. Gordon and A. Ashkin, Phys. Rev. A **21**, 1606 (1980).
25. V. S. Smirnov, A. M. Tumaikin, and V. I. Yudin, Zh. Éksp. Teor. Fiz. **96**, 1613 (1989) [Sov. Phys. JETP **69**, 913 (1989)].
26. A. V. Taichenachev, A. M. Tumaikin, and V. I. Yudin, Zh. Éksp. Teor. Fiz. **118**, 77 (2000) [JETP **91**, 67 (2000)].
27. A. V. Taichenachev, A. M. Tumaikin, and V. I. Yudin, Laser Phys. **2**, 575 (1992).
28. P. Marte, R. Dum, R. Taïeb, *et al.*, Phys. Rev. A **49**, 4826 (1994).
29. D. A. Varshalovich, A. N. Moskalev, and V. K. Khersonskiĭ, *Quantum Theory of Angular Momentum* (Nauka, Leningrad, 1975; World Sci., Singapore, 1988).
30. J. Dalibard, S. Reynaud, and C. Cohen-Tannoudji, J. Phys. B **17**, 4577 (1984).
31. A. V. Bezverbnyi, Pis'ma Zh. Éksp. Teor. Fiz. **74**, 162 (2001) [JETP Lett. **74**, 144 (2001)].
32. M. Gajda and J. Mostowski, Phys. Rev. A **49**, 4864 (1994).
33. K. Mølmer and C. Westbrook, Laser Phys. **4**, 872 (1994).

Translated by V. Astakhov

Journal of Experimental and Theoretical Physics, Vol. 96, No. 3, 2003, pp. 402–406.
From Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 123, No. 3, 2003, pp. 457–461.

ATOMS, SPECTRA, RADIATION

Perturbative Approach to the Hydrogen Atom in a Strong Magnetic Field¶

V. A. Gani*, A. E. Kudryavtsev, V. A. Lensky, and V. M. Weinberg
State Scientific Center Institute of Theoretical and Experimental Physics, Moscow, 119259 Russia
**e-mail: gani@heron.itep.ru*
Received August 15, 2002

Abstract—We consider states of the hydrogen atom with the principal quantum number $n \le 3$ and zero magnetic quantum number in a constant homogeneous magnetic field $\mathcal{H}$. The perturbation theory series is summed using the Borel transformation and conformal mapping of the Borel variable. Convergence of the approximate energy eigenvalues and their agreement with the corresponding existing results are observed for external fields up to $n^3\mathcal{H}/\mathcal{H}_0 \sim 5$, where $\mathcal{H}_0$ is the atomic magnetic field. The possibility of restoring the asymptotic behavior of energy levels using perturbation theory coefficients is also discussed. © *2003 MAIK "Nauka/Interperiodica".*

The magnetic fields proper of some astrophysical objects reach very high values [1, 2]. If we are interested in the atomic spectra in these external fields, it is convenient to introduce the natural measure of field strength, the atomic magnetic field

$$\mathcal{H}_0 \equiv \frac{e^3 m^2 c}{\hbar^3} = 2.55 \times 10^9 \text{ G}.$$

The fields $\mathcal{H}$ up to one-half of $\mathcal{H}_0$ are detected in the vicinity of some white dwarves. Neutron stars possess fields up to the order of $10^4\mathcal{H}_0$. For the correct interpretation of the observation results, it is desirable to know the atomic hydrogen spectrum in this range of external fields. For this aim, computations based on the adiabatic approach with a Landau level as the initial approximation were performed [3]. In what follows, we show which part of the desired external field range can be covered using the standard expansion in powers of $\mathcal{H}$, starting from the Coulomb levels of the hydrogen atom. The computation involves many orders of perturbation theory (up to the 75th order). Summation of the series is performed using the Borel transformation supplemented by a conformal mapping of the Borel variable.

The Borel summation method was introduced into quantum field theory relatively long ago (see, e.g., [4]). It has been tested on some quantum-mechanical problems (one of numerous examples is described in [5]) and is still applied in modern works [6]. Great hopes of the possibility to advance into the strong-coupling region were related to the Borel summation of the perturbation series. Some rather simple problems where the details can be traced and compared with the corresponding exact results supported this optimism. For example, for the funnel potential

$$V(r) = -\frac{1}{r} + gr,$$

applying a conformal mapping of the Borel variable and Padé summation of the Borel transformant gives the ground state energy as $g \longrightarrow \infty$ in the form

$$E(g) = Cg^{\nu}$$

with a precision of about 0.2% for the index ν and about 5% for the coefficient C [5]. It became clear later that such a successful summation presents a special but not general case. We can surmise that this success is a consequence of the simplicity of this problem. In contrast, the asymptotic behavior of energy levels in the Stark and Zeeman effects is established at very large values of the external fields. For the Stark effect, it is now practically impossible to reach the region of the truly asymptotic behavior by perturbation series summation. An intermediate linear asymptotic behavior is observed instead [7, 8].

To introduce the notation and the scale, we write the Hamiltonian

$$H = -\frac{1}{2}\nabla^2 - \frac{1}{r} + \frac{1}{8}g(r^2 - z^2) \equiv \hat{H}_0 + g\hat{H}_1, \qquad (1)$$

where $g \equiv \mathcal{H}^2/c^2$, and we use the atomic units $\hbar = m = e = 1$ hereafter. In (1), we drop the elementary contribution of the electron spin and consider only states with the magnetic quantum number $m \equiv 0$. We can expand $E(g)$ as a formal power series in g,

$$E(g) = \sum_{k=0}^{\infty} E_k g^k. \qquad (2)$$

¶This article was submitted by the authors in English.

We must then obtain hypersusceptibilities E_k. We can use the moment method for this aim. This method is especially useful in the cases where variables cannot be separated in the Schrödinger equation. Obviously, the Zeeman effect is precisely such a problem. In the previous work [9], the moment method was applied to the recurrent evaluation of hypersusceptibilities. A somewhat different version of the moment method was introduced in [10].

For the lower four "isolated" hydrogen levels, we immediately use the results of [9]. Unfortunately, the computer code employed in [9] for the relatively more complicated case of degenerate 3*s* and 3*d* states contained a mistake.[1] We therefore performed a new computation of the 3*s* and 3*d* hypersusceptibilities. Results of the computation for some orders are presented in Table 1. These results are in agreement with the results in [11], where high-order hypersusceptibilities were obtained for the first time (but the method used in [11] is much more complicated than the moment method).

As the order k increases, hypersusceptibilities grow as a factorial [12],

$$E_k \longrightarrow \tilde{E}_k = (-1)^{k+1} C_{nl} a_n^k \Gamma(1k + \beta_{nl}), \qquad (3)$$

where

$$a_n = \left(\frac{n^2}{\pi}\right)^2, \quad \beta_{nl} = 2n - 1 + \frac{(-1)^l}{2},$$

and C_{nl} are not essential for us; their values can be found in [9] and references therein. Equation (3) implies that series (2) is asymptotical, and the formal sum of such a series is therefore ambiguous. However, the choice of the summation method is in fact restricted: from physical considerations, the function $E(g)$ must have analytical properties that are to be reproduced by the true sum of series (2). In the nonphysical region $g < 0$, the diamagnetic perturbation $g\hat{H}_1$ changes its sign, the total Hamiltonian becomes "open," and the possibility of a spontaneous ionization of the atom emerges. The energy eigenvalue must therefore have an imaginary part at $g < 0$, and the function $E(g)$ must have a cut along the negative real semiaxis in the g plane. Summation using the Borel transformation results in a function that has the left cut and the discontinuity on this cut is represented by a smooth function of g.

The Borel transformant $B(w)$ of $E(g)$ is a series,

$$B(w) = \sum_{k=0}^{\infty} B_k w^k \qquad (4)$$

with the coefficients

$$B_k = E_k/\Gamma(2k + b_0),$$

[1] We are thankful to Prof. V.D. Ovsyannikov for drawing our attention to this mistake.

Table 1. Hypersusceptibilities of degenerate states

k	E_k for 3*s* state
1	19.57851476711195477229924488394
2	–7992.55848864256699349104381687
3	9951240.46627684231026404630780 0
4	–20931559882.53444368634980579917
5	58826900682409.79349115290157121
25	$1.37932338518206094144637879132 15 \times 10^{94}$
50	$-9.32271326968896166177886769035 16 \times 10^{211}$
75	$2.80535339708117043265749308311 76 \times 10^{340}$
k	E_k for 3*d* state
1	5.17148523288804522770075511 6050
2	–1017.42588635743300665089561 8312
3	738127.82473878268973595369219 95
4	–923576528.554411294118944200 8231
5	1677908319019.727217770438272530
25	$1.04312177717586140118123118583 95 \times 10^{92}$
50	$-6.07219785614468843000727265530 11 \times 10^{209}$
75	$1.73025529950554326807310876350 37 \times 10^{338}$

where b_0 is an arbitrary constant. The choice of b_0 can, in principle, affect the numerical results, but because changing its value within an interval of about $0.5 \le b_0 \le 5$ has a weak effect, we fixed b_0 for convenience. The numerical calculations in this work were performed at $b_0 = 3$. Series (4) converges, as usual, inside the circle $|w| < 1/a_n$. Substituting asymptotic coefficients $\tilde{E}_k$ for E_k, it is easy to verify that the singularity of $B(w)$ is located at $w = -1/a_n$. The energy of the level is related to the function $B(w)$ by the integral transformation

$$E(g) = \int_0^{\infty} e^{-x} B(gx^2) x^{b_0 - 1} dx. \qquad (5)$$

For the numerical integration in the right-hand side to be successful, $B(w)$ must be analytically continued from its convergence disk to the domain containing the image of the entire positive real w semiaxis. For this, we performed a conformal mapping of the Borel variable w. Many sufficiently effective versions of this mapping are appropriate. The main point is that the nearest singularity of the Borel transformant must be removed to infinity. We used the mapping

$$y = \frac{a_n w}{1 + a_n w} \qquad (6)$$

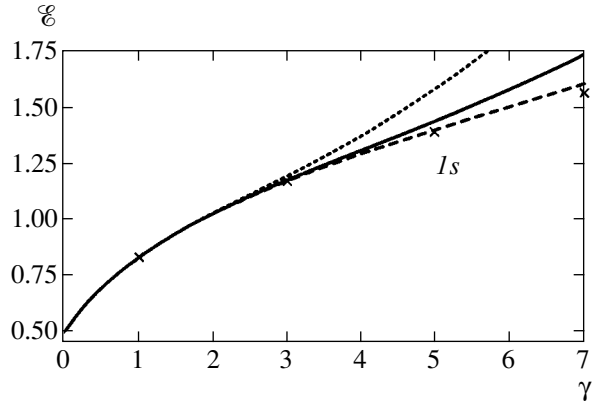

Fig. 1. Binding energy for the 1*s* state in at. units. The data is evaluated with double precision (using the Padé approximant [30/30], solid curve; and by straightforward summation, dotted curve) and with quadruple precision (using the Padé approximant [30/30], dashed curve). Crosses denote the data in [13].

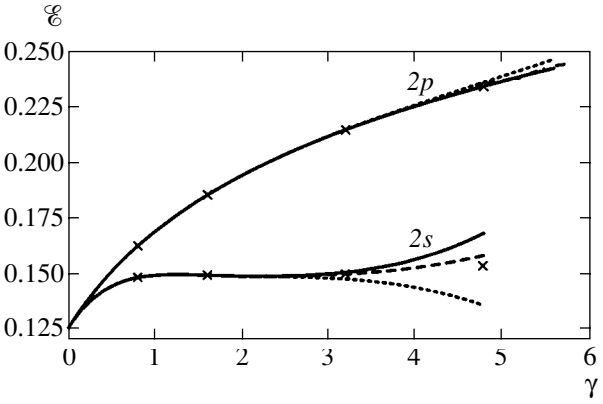

Fig. 2. Binding energy for the 2*s* and 2*p* states in at. units. Notation is the same as in Fig. 1.

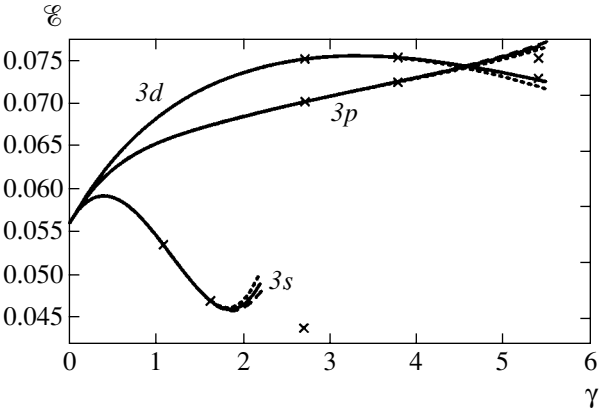

Fig. 3. Binding energy for the 3*s*, 3*p*, and 3*d* states in at. units. Notation is the same as in Fig. 1.

that was employed in [6]. As explained in [6], this transformation is optimal in the sense that it decreases the influence of all possible singularities of $B(w)$ from the nonphysical region. Transformation (6) is equivalent to the following series rearrangement:

$$B(w) = \sum_{m=0}^{\infty} D_m y^m, \quad D_0 = B_0,$$
$$D_m = \sum_{k=1}^{m} \frac{(m-1)!}{(k-1)!(m-k)!} \frac{B_k}{a^k}, \quad m \geq 1. \tag{7}$$

To improve convergence, we applied the Padé summation to rearranged series (7),

$$B(w) \approx [M/N](y) \approx \frac{P_M(y)}{Q_N(y)}, \tag{8}$$

where P_M and Q_N are polynomials of the respective degrees M and N.

We performed computations using various Padé approximants and a straightforward summation of rearranged series (7). To illustrate the effect of computational accuracy on summation results, we compared those made with double precision (16 decimal places) and quadruple precision (32 decimal places).

Some graphs of the obtained binding energy

$$\mathscr{E}(\gamma) = \frac{1}{2n^3}\gamma - E(\gamma^2)$$

as a function of the parameter $\gamma \equiv n^3\mathscr{H}/c$ are given in Figs. 1–3. Compared with [9], the region of external field values for which these eigenvalues are successfully recovered is extended by a factor of about 5. As usual, the precision of the sum considerably increases at lower $\mathscr{H}$ values. The accuracy of the summation method described above is illustrated in Table 2.

We note that in [9], Padé approximants were immediately applied to summation of divergent series (2). These approximants imitate the discontinuity on the cut $g < 0$ by a set of delta functions, which is a very rough approximation. On the other hand, the same discontinuity is represented by a smooth function of g as a result of the Borel summation. Our calculations confirmed that mapping (6) is indeed very efficient: after this mapping, Padé summation of the Borel transformant improves the convergence only slightly, and its straightforward summation appears to be sufficient in some cases (see Figs. 1–3).

One technical detail is of principal importance for the perturbation series summation by any method. The precision of the entire chain of computations must increase as the number of the involved successive terms increases. This is simply a consequence of the fact that the sum, being of the order of unity, is the result of a compensation of very large terms with alternating signs.

At a first glance, it seems that the high-precision requirement is not necessary for the Borel transformant because all the essential alternating sign coefficients B_k have approximately the same order. However, any

numerical procedure of analytic continuation usually requires a high precision. Turning to series rearrangement (7), we see that binomial coefficients entering the sum for D_k change by 20 orders of magnitude (in the present case). Obviously, an enormous loss of precision occurs in performing the sum for D_k in (7). Therefore, if we want to use all B_k up to the 75th order, the precision of the B_k coefficients must be better than about 10^{-20}. In our calculations, the precision of E_k, and consequently, the precision of B_k was about 10^{-30}, and the precision of D_k therefore decreased from 10^{-30} at $k = 0$ to approximately 10^{-10} at $k = 75$.

We now turn to the problem of restoring the $E(g)$ dependence at large values of g and focus on the ground state. We first note that an interpolation expression for the ground (tightly bound) state energy was obtained in [3]. In spite of multiple anticrossings at $\gamma \le 300$ and of the related computational complications, the fit in [3] provides precision within 10^{-3}–10^{-2} in the range $0.1 \le \gamma \le 10^4$.

The asymptotic form of the ground level energy at large g (equivalently, at large γ) is given by

$$E(\gamma) \longrightarrow \frac{1}{2}\gamma - \frac{1}{2}\ln^2(\lambda\gamma) + \ldots, \qquad (9)$$

where λ is a dimensionless constant (see, e.g., [14]). We first consider the possibility of restoring the leading term parameters in (9)—the power index and the constant multiplier—using the perturbation theory. Methods applicable to this problem are considered in [5, 6]. We note that for the asymptotic regime to establish, the leading term in (9) must be large compared with the correction term. We refer to the results in [13] (where the values of $\mathscr{E}(\gamma)$ were obtained by a variational procedure), which show that the binding energy is less than 20% of $(1/2)\gamma$ only if $\gamma > 10^2$. We can therefore speculate on restoring the asymptotic parameters only if we succeed in summing $E(g)$ in this region of external fields, but we failed to do this using only 75 coefficients E_k, and the linear asymptotic behavior could therefore not be restored. This was confirmed in our attempts to apply the methods proposed in [5, 6]: no plausible result was obtained. In the method in [6], parameters of the asymptotic form of $E(g)$ are related to the large-k behavior of the coefficients D_k. Namely, if $E(g) \longrightarrow Cg^\nu$ as $g \longrightarrow \infty$, we obtain similarly to [6] that

$$D_k \longrightarrow \frac{Ck^{\nu-1}}{a_n^k \Gamma(\nu)\Gamma(2\nu + b_0)}.$$

It was then suggested to perform the fit of C and ν using the known D_k coefficient and their errors by means of the χ^2 method. However, in our case, the value of χ^2 at its minimum is extremely large (about 10^8 even if we try to fit only five coefficients D_k at a statistical error of $\sigma = 10^{-10}$, and we have no reason to increase this value of σ). This result indicates that the asymptotic regime of D_k is established at values of k much larger than 75.

The power index in the asymptotic form of $E(g)$ could also be traced using the method described in [5]. This method consists in taking the limit of the expression $wB'(w)/B(w)$ as $w \longrightarrow \infty$ (or equivalently, the limit of $y(1-y)B'(y)/B(y)$ as $y \longrightarrow 1$), which gives the exact value of ν. However, numerical calculation showed that we did not obtain a reasonable precision for the limit value in the region where $B(y)$ was recovered (for y close to 1, the error must obviously increase because of a finite number of D_k used).

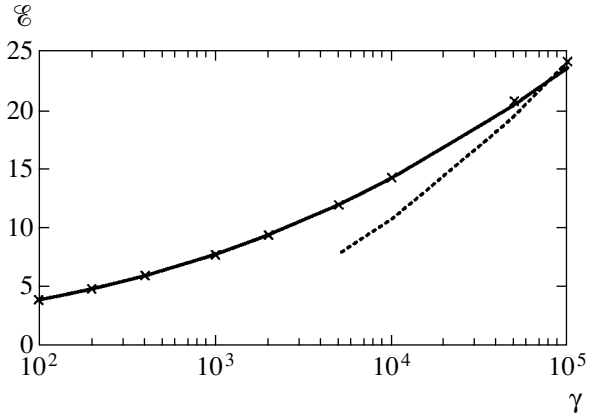

Fig. 4. Binding energy and its asymptotic form. The solid curve is plotted using Eq. (6) in [3]. Crosses denote the data in [13]. The dotted curve is the logarithmic asymptotic behavior with $\lambda = 0.010$.

Table 2. Values of the binding energy for $2p$ state

[M/N]	$\mathscr{E}$ at $\gamma = 1.12$	[M/N]	$\mathscr{E}$ at $\gamma = 3.20$	$\mathscr{E}$ at $\gamma = 4.80$	$\mathscr{E}$ at $\gamma = 8.00$
[28/28]	0.172618226340	[30/30]	0.214270	0.23396	0.2655
[29/29]	0.172618226340	[37/36]	0.214257	0.23371	0.2612
[30/30]	0.172618226339	[37/37]	0.214265	0.23370	0.2610
[31/31]	0.172618226343				
[32/32]	0.172618226340				
Ref. [13]	0.17261822	Ref. [13]	0.2142655	0.233675	0.260006

It thus appears to be impossible to obtain asymptotic parameters corresponding to a Landau level on the basis of all the known perturbation theory coefficients. Nevertheless, taking all possible information into account (including that contained in interpolation formula [3] and variational calculation results [13]), the question naturally arises whether it is possible to subtract the Landau level energy from the "exact" function $E(g)$ and trace the second term of the asymptotic expression. In other words, at which external field values does the logarithmic term in Eq. (9) become noticeable? An illustration to the answer is given in Fig. 4, where we plotted the binding energy and its logarithmic asymptotic form. The constant λ is chosen such that the value of the logarithmic term $(1/2)\ln^2(\lambda\gamma)$ in the asymptotic expression coincides with the data in [13] at $\gamma = 10^5$. This occurs at $\lambda = 0.01$, and the smallness of this constant indicates that the value of γ is too small to speak about the asymptotic regime. We can see that the asymptotic curve and the curve of exact data have considerably different slopes.

Thus, for the Zeeman effect, perturbation theory does not allow recovering even the linear part of the ground state energy asymptotic behavior in a strong field and the logarithmic term becomes essential at huge fields beyond the neutron stars range.

REFERENCES

1. *The Lives of the Neutron Stars*, Ed. by M. Alpar, Ü. Kiziloğlu, and J. van Paradijus (Kluver, Dordrecht, 1995).
2. P. Fassbinder and W. Schweizer, Astron. Astrophys. **314**, 700 (1996); Phys. Rev. A **53**, 2135 (1996).
3. A. Y. Potekhin, J. Phys. B **27**, 1073 (1994); J. Phys. B **31**, 49 (1998).
4. J. Zinn-Justin, Phys. Rep. **70**, 109 (1981).
5. V. S. Popov and V. M. Weinberg, Phys. Lett. A **90**, 107 (1982).
6. I. M. Suslov, Zh. Éksp. Teor. Fiz. **120**, 5 (2001) [JETP **93**, 1 (2001)].
7. D. I. Kazakov and V. S. Popov, Zh. Éksp. Teor. Fiz. **122**, 675 (2002) [JETP **95**, 581 (2002)].
8. V. M. Weinberg, V. D. Mur, V. S. Popov, and A. V. Sergeev, Zh. Éksp. Teor. Fiz. **93**, 450 (1987) [Sov. Phys. JETP **66**, 258 (1987)].
9. V. M. Weinberg, V. A. Gani, and A. E. Kudryavtsev, Zh. Éksp. Teor. Fiz. **113**, 550 (1998) [JETP **86**, 305 (1998)].
10. F. M. Fernandez and J. A. Morales, Phys. Rev. A **46**, 318 (1992).
11. H. J. Silverstone and R. K. Moats, Phys. Rev. A **23**, 1645 (1981).
12. J. Avron, Ann. Phys. (N.Y.) **131**, 73 (1981).
13. Jang-Huar Wang and Chen-Shiung Hsue, Phys. Rev. A **52**, 4508 (1995).
14. L. D. Landau and E. M. Lifshitz, *The Quantum Mechanics: Non-Relativistic Theory*, 3rd ed. (Nauka, Moscow, 1974; Pergamon Press, Oxford, 1976).

Journal of Experimental and Theoretical Physics, Vol. 96, No. 3, 2003, pp. 407–417.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 123, No. 3, 2003, pp. 462–473.

ATOMS, SPECTRA, RADIATION

Excitation of Rydberg States of Atoms and Diatomic Molecules by Short Laser Impulses

G. K. Ivanov* and V. L. Bodneva

Semenov Institute of Chemical Physics, Russian Academy of Sciences, Moscow, 119991 Russia
**e-mail: genivan@chph.ras.ru*

Received August 29, 2002

Abstract—Excitation of multilevel Rydberg states (atoms and diatomic molecules) is studied under an intense time-periodic perturbation that is instantaneously turned on and off. For a one-quantum excitation from a lower lying state, general expressions for the Laplace images of the population amplitudes of Rydberg states are obtained with regard to their decay characteristics. It is shown that the problem considered is reduced to the determination of the positions and widths of the energy levels of a quantum system in the field of monochromatic laser radiation of the same frequency and intensity as that in a pulse mode. To determine these quantities, an integral formulation is proposed for the eigenvalue problem for energies that is relatively simply solved with regard to the effect of the ionic core and its complex vibration–rotation structure on diatomic molecules. The specific features of the excitation of Rydberg states and the behavior of Rydberg wave packets are studied depending on the intensity and duration of laser radiation. A quantum phenomenon of rotational orientation of electron-excited diatomic molecules is considered. © *2003 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

In this paper, we study the specific features of excitation of Rydberg states of atoms and diatomic molecules by short laser impulses.

We assume that the field acting on a quantum system is linearly polarized and can be approximated by the expression

$$\mathbf{F}(\omega, t) = \mathbf{f}(t)\cos(\omega_f t), \quad (1)$$

where $f(t)$ is the field-strength amplitude and ω_f is the field frequency ($e = \hbar = m_e = 1$). In this study, we concentrate our attention on the population amplitudes of excited Rydberg states and on the structure of wave packets that are formed during and immediately after the transmission of a laser impulse.

When describing these processes, the shape $f(t)$ of the envelope of the exciting impulse plays an important role. This problem admits an analytic approach for short laser impulses. Such an approach is widely used in the theory of femtosecond spectroscopy of molecules and is based on perturbation theory for exciting and scanning impulses [1, 2].

The second efficient method for solving this problem is based on the assumption that the pulse envelope $f(t)$ is rectangular; i.e., the excitation is time periodic and is instantaneously turned on (at $t = 0$) and off (at $t = t_0$) [3–5]. During the last two decades, extensive investigations have been carried out based on the Laplace method used in these assumptions [4–8] (as a rule, a small number of excited states are taken into consideration in these investigations). A qualitative analysis for multilevel systems was carried out in [4] under certain constraints imposed on the characteristics of the exciting impulse. In [9], the authors considered transitions in a system of equidistant Rydberg levels that are resonantly coupled to a lower lying state under the assumption that the Rydberg states were initially occupied. We should also mention study [10], in which the authors considered certain specific features of resonant ionization through Rydberg states of atoms. Multilevel systems were also investigated via direct numerical solution of the Schrödinger equation [11]. It is important that the aforementioned analytic studies [4–10] did not explicitly take into account the effect of the ionic core on Rydberg atoms. The extension of the existing methods (in the form they were used in the aforementioned works) to molecular systems presents a rather complicated problem, especially when large groups of states are involved in the process and perturbation theory becomes inapplicable.

In Section 2 of the present paper, we describe a sufficiently general procedure for determining the Laplace images as applied to processes of the aforementioned type. In excited Rydberg states, one can take into account the effect of the ionic core on the Rydberg states of atoms and molecules, a nonadiabatic coupling between various types of electron and nuclear motions, and the interaction between various series of Rydberg states against the background of a continuous spectrum, including combined transitions in continua. The main problem for the systems considered is the determination of Laplace images because, in view of the fact that the parameters involved in the equations are actually independent of energy, the subsequent calculation of

the population amplitudes of the states is carried out by means of the residue theorem.

It would be natural to start the application of the approach proposed to the study of the class of problems in question from the simplest multilevel systems that are formed as a result of the absorption of a laser-radiation quantum. These are atoms of hydrogen and alkaline metals (the latter feature the effect of the ionic core on the excited states of Rydberg series). These atoms are considered in Section 3, where we discuss the possibility of controlling the reaction yield and the formation of Rydberg states with fixed values of the principal quantum number n. In this section, we also discuss the specific features of the behavior of Rydberg wave packets.

In Section 4, we investigate laser excitation of low-energy Rydberg electron-rotational states as applied to molecular systems. Within the framework of a two-level transition scheme, we consider the problem of rotational orientation of excited diatomic molecules under linearly polarized radiation from a pulse source. Earlier, such a problem was considered for the rotational orientation of molecular ions that are formed in the field of a monochromatic laser radiation [12] and, with regard to a finite duration of a laser impulse in the classical approximation, for the rotational motion of diatomic molecules [13].

2. MECHANISMS OF INSTANTANEOUS TURNING ON OF PERTURBATIONS IN MULTILEVEL SYSTEMS (GENERAL CONSIDERATION)

The description of the reaction of a quantum system to an instantaneous turning on of a time-periodic perturbation has been sufficiently well developed [3–5]. This procedure reduces to calculating the expansion coefficients $C_s(t)$ of a wave function Ψ over the set N of discrete states φ_s involved in the process,

$$\Psi = \sum C_s(t)\varphi_s, \tag{2}$$

which are determined by the Laplace transformation,

$$C_s(t) = \frac{1}{2\pi i}\int_c a_s(E)\exp(iEt)dE, \tag{3}$$

where the integration contour c lies below all the singularities of the function $a_s(E)$. The Laplace images $a_s(E)$ are obtained by eliminating the states of the continuous spectrum as well as other states of the quantum system inessential for the process considered. As a result, in the notation adopted in the present paper, we obtain the following system of inhomogeneous algebraic equations with the initial conditions $C_{s_0}(0) = 1$ and $C_{s\neq s_0} = 0$ ($t = 0$ is the moment of turning on the field) for the transition scheme considered:

$$(E - E_s)a_s = \sum_{s'} T'_{ss'}a_{s'} + i\delta_{s_0 s}. \tag{4}$$

Here, E_s is measured from the boundary of the spectrum and $E_{s_0} = E_0^0 + \omega_f$, where E_0^0 is the position of a lower lying state of the system in the absence of field.

The quantities a_s determined from Eq. (4) also allow us to obtain the Laplace images for the amplitudes a_p of the states $|p\rangle$ of the continuous spectrum with energies E_p [4, 5]:

$$a_p = \sum_s \frac{a_s V_{sp}}{E - E_p}, \tag{5}$$

where V_{sp} is the matrix element of the transition, $V = \mathbf{f}\cdot\mathbf{D}/2$, and $\mathbf{D}$ is the dipole-moment operator.

Recall that, when using Eqs. (4), it is assumed that only one of the states (s_0) is occupied with a probability of one at the moment $t = 0$. The expression for the matrix elements $T'_{ss'}$ includes a direct field interaction between the states s and s' with the functions $|s\rangle$ and $|s'\rangle$ normalized to unity, as well as transitions through continua. It is well known that, for moderately high field strength f such that

$$f\omega_f^{-5/3} \ll 1, \tag{6}$$

the operator T' is represented as follows:

$$T' = V - iV\sum_p |p\rangle\langle p|V. \tag{7}$$

The subscripts s in (4) and (5) and p in (7) include a natural (in the absence of external field) set of quantum numbers that characterize the system, as well as the number k—the variation of the number of quanta ω_f during the interaction with the radiation field. The summation over p presumes that there are various continua into which the quantum system may decompose under the influence of the field. These processes are responsible for the broadening of energy levels. Note that the electron wave functions $|p\rangle$ differ from those normalized in a conventional way by the factor $\sqrt{\pi}$. The first term in the matrix elements $T'_{ss'}$ in (7) describes a direct field interaction between the states s and s' with the wave functions $|s\rangle$ and $|s'\rangle$ normalized to unity, and the second term describes Raman-type transitions (the so-called Λ transitions through continua [9]).

Under condition (6), the strongest effects are associated with laser-induced direct transitions between the initial s_0 and excited discrete states. Just as in [9, 10], the field is assumed to be sufficiently strong:

$$|V_{0s}| \gg |E_s - E_{s\pm 1}|.$$

However, one cannot consider it as a superstrong field because, due to condition (6), the second term in (7) is small as compared with the first. At the same time, the second term in (7) should also be correctly taken into account when considering the decay of the excited states.

Now, note that there exists a more general representation for the operator T'. We derive this representation within the formalism of the stationary radiation-collision method, which was proposed in [14] and developed in [15, 16] (Eqs. (4) for the Laplace images are also stationary; the specific feature of these equations is displayed only by the free term in the right-hand side of (4)). To this end, in the Green function of a system with a turned-off interaction V, we single out the contribution of N states that are essential for the processes considered,

$$G = G^{\mathrm{d}} + G^0, \tag{8}$$

where

$$G^{\mathrm{d}} = \sum_{s=0}^{N} \frac{|s\rangle\langle s|}{E - E_s}. \tag{9}$$

Then, we introduce the operator

$$t^{\mathrm{f}} = V + V\tilde{G}t^{\mathrm{f}}, \tag{10}$$

whose definition involves $\tilde{G}$—the real part of the function

$$G^0 = \tilde{G} - i\sum_p |p\rangle\langle p|$$

that is smooth with respect to energy. Then, the operator T' defined by

$$T' = V + VG^0T' \tag{11}$$

assumes a more general (with respect to (7)) form [15, 16]:

$$T' = t^{\mathrm{f}} - it^{\mathrm{f}}\sum_p |p\rangle\langle p|T'. \tag{12}$$

This expression formally takes into account (through t^{f}) the Stark shifts of energy levels as well as cascade transitions in continua that affect the broadening of these levels.

Introducing the quantities

$$B_s = (E - E_s)a_s, \tag{13}$$

we can rewrite Eq. (4) as

$$B_s = \sum_{s'} T'_{ss'}\frac{B_{s'}}{E - E_{s'}} + i\delta_{s_0 s}. \tag{14}$$

This equation is equivalent to (4) and may serve as a basis for the further generalization of the theory. It can be obtained within the formalism of reconstructed integral equations that take into account the possibility of splitting the Green function (9) of the quantum system into the parts that strongly (G^{d}) and weakly (G^0) depend on energy. Including G^0 in the definition of the operator T' (11), we obtain the following expression for the operator B that forms the matrix B_{ss_0} in the general form:

$$B = i|s_0\rangle\langle s_0| + T'G^{\mathrm{d}}B. \tag{15}$$

In representation (9)–(12), Eq. (15) is naturally reduced to Eq. (14), which determines the required quantities B_s in (13). Here, one can also see a direct analogy with the formulation of the problem for the matrix T of radiation collisions [14–16] that occur in the field of a monochromatic laser radiation,

$$T = T' + T'G^{\mathrm{d}}T. \tag{16}$$

The operator equations (15) and (16) differ only by free terms.

The substitution of the explicit expression (9) for the Green function of the system into (15) (with a turned-off field interaction) allows one to easily take into consideration the contributions of large groups of Rydberg states for the values of the principal quantum number $n \gg 1$. In this case, one can also take into account the effect of the ionic core and various types of nonadiabatic coupling between electron and nuclear motions. In other words, to solve the problem, one can directly take advantage of the stationary formulation of the problem of radiation collisions in the field of a monochromatic laser radiation [14–16] in the theoretical analysis.

To solve the problem under a strong perturbation of a Rydberg state, one has to find representations for the functions a_0 that admit the application of the residue theorem. To this end, one should return to standard equations (4) for the Laplace images. In the case of one-quantum excitation of Rydberg states (which are henceforth denoted by subscript n) from the lower lying level E_0^0 (denoted by subscript 0), these equations are expressed as

$$\begin{aligned}(E - E_0 - T'_{00})a_0 &= \sum_n T'_{0n}a_n + i\delta_{0n},\\ (E - E_n - T'_{nn})a_n &= T'_{0n}a_0.\end{aligned} \tag{17}$$

Here, E_n are Rydberg levels that are not perturbed by the field, $E_0 = E_0^0 + \omega_{\mathrm{f}}$ is the position of the level E_0 after the absorption of the electromagnetic field quantum ω_{f}. Here, we can assume that $T'_{0n} = V_{0n}$ because this operator describes a two-quantum transition from the initial state to the continuum. At the same time, one must take into account the terms of the operator T' in (7) that are diagonal in n because these terms are

responsible for the broadening Γ_n of the Rydberg levels. The off-diagonal terms T'_{nn} ($n \neq n'$) make a negligible contribution due to condition (6) adopted in this work, which suggests that the transition probabilities in the continua are small (in the atomic scale). This condition leads to the inequality $\Gamma_n \ll 1/n^3$, which is well satisfied and implies that there is no efficient mixing of Rydberg states through continua.

Under the assumptions made, for a one-quantum excitation of Rydberg states from the lower lying level E_0, Eq. (17) admits the following compact analytic solution:

$$a_0 = \frac{\prod_n (E - E_n)}{(E - \tilde{E}_0)\prod_n (E - \tilde{E}_n)},$$
$$a_m = -iV_{0m}\frac{\prod_{n\neq m} (E - E_n)}{(E - \tilde{E}_0)\prod_n (E - \tilde{E}_n)}, \tag{18}$$

where $\tilde{E}_0$ and $\tilde{E}_n$ are the positions of the above-mentioned levels with regard to the field perturbation. As we have shown above (by comparing Eqs. (14) and (16)), these levels are determined in the same way as in the problem on a Rydberg system subject to a permanent (for $-\infty < t < \infty$) time-periodic perturbation.

The expressions, obtained from (18), for the population amplitudes of states for $0 < t < t_0$ are represented as

$$C_0(t) = \frac{\prod_n (\tilde{E}_0 - E_n)}{\prod_n (\tilde{E}_0 - \tilde{E}_n)}\exp(-i\tilde{E}_0 t)$$
$$+ \sum_r \frac{\prod_n (\tilde{E}_r - E_n)}{(\tilde{E}_r - \tilde{E}_0)\prod_{n\neq r} (\tilde{E}_r - \tilde{E}_n)}\exp(-i\tilde{E}_r t),$$
$$C_m(t) = -iV_{0m}\left[\frac{\prod_n (\tilde{E}_0 - E_n)}{\prod_n (\tilde{E}_0 - \tilde{E}_n)}\exp(-i\tilde{E}_0 t)\right. \tag{19}$$
$$\left. + \sum_r \frac{\prod_{n\neq m} (\tilde{E}_r - E_n)}{(\tilde{E}_r - \tilde{E}_0)\prod_{n\neq r} (\tilde{E}_r - \tilde{E}_n)}\exp(-i\tilde{E}_r t)\right].$$

Recall that $\tilde{E}_0$ and $\tilde{E}_n$ are complex quantities:

$$\tilde{E}_0 = \tilde{E}'_0 - i\frac{\Gamma_0}{2}, \quad \tilde{E}_n = \tilde{E}'_n - i\frac{\Gamma_n}{2}; \tag{20}$$

they contain information on the positions $\tilde{E}'_0$ and $\tilde{E}'_n$ and widths Γ_0 and Γ_n of the system levels under laser irradiation.

The expression for $C_0(t)$ in (19) corresponds to formula (5) in [10], where this quantity is represented as a sum of contributions of states with different quasi-energies ϵ_n of a dressed atom:

$$C_0(t) = \sum \alpha_n \exp(-i\epsilon_n t).$$

The coefficients α_n were not calculated in [10]; the authors only pointed out the functions from which these coefficients can be obtained by the residue theorem.

Expressions (19) obtained in the present study give an analytic representation for these coefficients and provide a direct method for their determination, provided that the positions of the quasilevels $\tilde{E}_n$ in the field of monochromatic radiation of frequency ω_f and strength **f** are known.

In the two-level approximation, which is realized in these systems when the quasilevel E_0 lies close to a certain Rydberg state E_m, so that $|\tilde{E}_0 - \tilde{E}_m| \ll |E_0 - \tilde{E}_n|$ ($n \neq m$), we have

$$C_m(t) = -iV_{0m}\left[\frac{1}{\tilde{E}_0 - \tilde{E}_m}\exp(-i\tilde{E}_0 t)\right.$$
$$\left. + \frac{1}{\tilde{E}_m - \tilde{E}_0}\exp(-i\tilde{E}_m t)\right]. \tag{21}$$

When $\Gamma_0, \Gamma_m = 0$, (21) implies formula (3.7) of [3].

Let us illustrate result (19) by an example of a three-level system when, together with the quasilevel $\tilde{E}_0$, the two closest Rydberg states with energies E_n and E_{n+1} ($E_n < E_0 < E_{n+1}$) are taken into consideration. In this case, we have the following expression for the population amplitude of the state n subject to a laser impulse for $t \leq t_0$:

$$C_n(t) = -iV_{0n}\left[\frac{\tilde{E}_0 - E_{n+1}}{(\tilde{E}_0 - \tilde{E}_n)(\tilde{E}_0 - \tilde{E}_{n+1})}\exp(-i\tilde{E}_0 t)\right.$$
$$+ \frac{\tilde{E}_n - E_{n+1}}{(\tilde{E}_n - \tilde{E}_0)(\tilde{E}_n - \tilde{E}_{n+1})}\exp(-i\tilde{E}_n t) \tag{22}$$
$$\left. + \frac{\tilde{E}_{n+1} - E_{n+1}}{(\tilde{E}_{n+1} - \tilde{E}_0)(\tilde{E}_{n+1} - \tilde{E}_n)}\exp(-i\tilde{E}_{n+1} t)\right].$$

The general analysis of the excitation of multilevel states under moderately high strength f of the external field, such that $|E_n - E_0| \gg V_{0n}$ for any n, also proves to

be simple. This analysis can be carried out analytically, without resorting to perturbation theory, which requires restrictions on the duration of a laser impulse.

In this case, we actually have the two-level approximation formula

$$W^{(n)} = \frac{V_{0n}^2}{(\tilde{E}_n - \tilde{E}_0)^2} \times \left| \exp\left(-iE_0 t_0 - \frac{\Gamma_0}{2} t_0\right) - \exp\left(-iE_n t_0 - \frac{\Gamma_n}{2} t_0\right) \right|^2 \quad (23)$$

for the probability of excitation of the nth state of series P after a laser impulse. One can easily verify that $C_n(t) \longrightarrow 0$ as $t \longrightarrow 0$ in formulas (19) and (22). The same property is incident to the multilevel amplitudes $C_n(t)$.

Using expressions (18) and (19), we have reduced the problem under consideration to the analysis of the behavior of Rydberg levels in the field of a monochromatic laser radiation. These levels can be determined by equating the determinant of system (17) to zero. This procedure is convenient when the levels E_n and the eigenfunctions ψ_n of the Rydberg system in the absence of field are well known. As applied to molecular systems, these quantities are very sensitive to specific features of the ionic-core structure and to the diversity of types of nonadiabatic coupling between electron and nuclear motions. Therefore, it is expedient to control these features in the equations themselves. Then, it is preferable to use eigenvalue equations for energies in the integral formulation:

$$T = T' G^d T. \quad (24)$$

Since the interactions that have not been taken into account in the Green function $G = (E - H_0)^{-1}$ are included in the operator T' defined by (12), Eq. (24) is readily rewritten in the form convenient for solving the problem considered. If one uses the Coulomb Green function as G, then the effect of the ionic core on Rydberg atoms X^{**} and molecules XY^{**} and on induced Rydberg electrons of the inelastic transitions $e^- - X^+$ or $e^- - XY^+$ can be explicitly manifested in Eq. (12) if one includes the electrostatic interaction t^e with the lattice in this equation. In this case, the operator t^f in (12) is replaced by the combination of two operators

$$t = t^e + t^f; \quad (25)$$

moreover, under condition (6), which presumes that the transition probabilities in the continua are small (in the atomic scale), we have

$$t^f = (1 + t^e G^0)(V + VG^0 V)(1 + G^0 t^e) \quad (26)$$

in linear and quadratic (in the field) approximations [15].

Formula (26) takes into account the effect of the ionic core and the associated distortions of the wave functions of the Coulomb center. Below, we introduce a subscript q assigned to a chosen series of Rydberg states. Based on the arguments presented and the formulations of Eqs. (16) and (24) developed earlier [15, 16] in the theory of radiation collisions, one can represent a generalized equation for the Laplace images in the following form (below, it is assumed that the group of interacting Rydberg states adjoining the boundary of the spectrum is coupled to the continuous spectrum and the lower lying level E_0 through the electromagnetic field):

$$B = i|0\rangle\langle 0| + T' \sum_{q,k} |q_k\rangle\langle q_k| B \cot(\pi \nu_{qk}) + T' \frac{|0\rangle\langle 0|}{E - E_0} B, \quad (27)$$

where $\nu_{qk} = \{1(\varepsilon_q + k\omega_f - E)\}^{-1/2}$ is the effective principal quantum number of the Rydberg series corresponding to different states of the ionic core with the excitation energy ε_q. The energy E in (27) is measured from the boundary between the spectra of the ground electron state and the vibration–rotation state of the lattice. Therefore, in what follows, we omit the subscript $k = 0$ when considering Rydberg states.

Equation (27) is constructed on the basis of the functions $|q\rangle$, which are normalized over the energy scale both in continuous and discrete spectra. Therefore, the quantities $T'_{0q} = V_{0q}$ arising below differ from the quantities V_{0n} introduced above by a factor of $\sqrt{\pi n^3}$:

$$V_{0n} = \frac{1}{\sqrt{\pi n^3}} V_{0q}.$$

Recall that the excited initial state of energy (after the absorption of a light quantum ω_f) is overlapped by the Rydberg states of an atom or a molecule.

3. LASER-INDUCED TRANSITIONS ASSOCIATED WITH THE EXCITATION OF RYDBERG STATES OF ALKALINE METAL ATOMS

In atoms of alkaline metals, only the P series of Rydberg resonances, denoted by subscript q (see Fig. 1), is excited during transitions from the ground state. Under condition (6), it is hardly probable that other Rydberg series will be involved in the process due to Λ transitions through the continua. In this case, for the eigenvalue problem for energies, we have

$$B_0 = T'_{0q} \cot(\pi\nu) B_q + T'_{00} \frac{1}{E - E_0} B_0,$$
$$B_q = T'_{qq} \cot(\pi\nu) B_q + T'_{q0} \frac{1}{E - E_0} B_0. \quad (28)$$

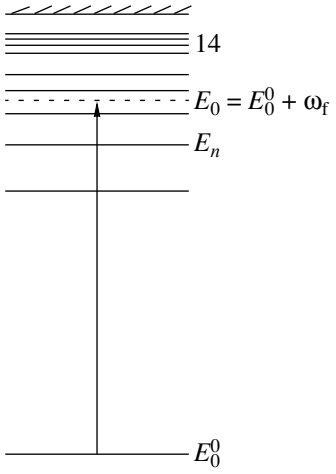

Fig. 1. Typical configuration of Rydberg levels of a hydrogen atom. The dashed line represents the position of the ground state after the absorption of a quantum ω_f of external radiation.

Here, $T'_{qq} = t^e_{qq} - i\gamma$, $t^e_{qq} = -\tan(\pi\mu_1)$ (μ_1 is the quantum defect of Rydberg levels of the P series), $T'_{0q} = \tilde{t}^f_{0q}$ (for T'_{00}, one can set $T'_{00} = 0$ assuming that t^e_{00} is included in E^0_0), $E_0 = E^0_0 + \omega_f$, and $\nu = (-2E)^{-1/2}$ (E is energy measured from the boundary of the spectrum).

The quantity B_q represents a large group of Rydberg states of an atom. Therefore, instead of system (4) involving a large number of states, here we have only two equations. The quantities a_0 and a_q corresponding to the Laplace images and determined according to (28) are given by the following expressions:

$$a_0 = \frac{\tan(\pi\nu) - t_{qq} + i\gamma}{D(E)}, \quad a_q = \frac{V_{0q}}{D(E)}, \qquad (29)$$
$$D(E) = (E - E_0)[\tan(\pi\nu) - t^e_{qq} + i\gamma] - V^2_{0q}.$$

It is interesting to note that expressions of type (29) for a_0 and a_q also arise in calculating the intense interaction between strongly coupled Rydberg states in radiation-collision processes [15].

Equations (28), which take into account the contribution of one Rydberg series only, represent the simplest realization of a multichannel (in q) version of the theory based on representation (27).

Note also that Eq. (29) for $D(E)$ can be rewritten as a sum of contributions of separate Rydberg states; i.e., it can be reduced to a characteristic equation used in [9] for solving the eigenvalue problem for energies.

Unfortunately, we cannot compare our results for the population amplitudes (19) with the results of [9] because these results have been obtained under different assumptions on the initial state of the system.

Now, the main problem is to find the roots of the equation

$$D(E) = 0,$$

which determine the poles of the functions a_0 and a_q. This problem proves to be analogous to that one has to solve in the theory of interaction of a quantum system with monochromatic laser radiation.

The advantage of representation (28) and of more general equations for the eigenvalue problem for energies that follow from (27) lies in the fact that they allow one to take into account the relation between the widths Γ_n, Γ_0 and the positions E'_n, E'_0 of the levels in the field of laser radiation in the analytic form. Returning to Eq. (29), we note that, in view of the smallness of γ ($\gamma \ll 1$), this equation can be solved in two steps. First, for $\gamma = 0$, we determine the real roots of Eq. (29). Then, using the values of E'_n and E'_0 obtained and passing to the energy scale $\tilde{E}_n = E'_n - i\Gamma_n/2$, $\tilde{E}_0 = E'_0 - i\Gamma_0/2$, we find that the field-induced shifts in the scale ν and the level widths behave as follows:

$$\tan(\pi\nu) = t^e_{qq} + \frac{V^2_{0q}}{E'_n - E'_0},$$
$$\Gamma_n = \frac{2\gamma}{\pi\nu_n^3}\frac{1}{1 + \left(t^e_{qq} + \frac{V^2_{0q}}{E'_n - E'_0}\right)^2}. \qquad (30)$$

The quantities E'_n introduced above are expressed as

$$E'_n = -\frac{1}{2\nu^2} = \frac{1}{2(n - \mu_f)^2},$$

where, as is shown in [17] by an example of a hydrogen atom, the field-induced quantum defects μ_f may reach values close to $\mu_f = \pm 1/2$ for field strength of $f \sim 10^{-2}$. This means that the difference between E'_n and E_n in the scale of distances between Rydberg levels, $\delta E_n \sim 1/n^3$, may reach considerable values.

When $V_{0q} = 0$, we derive the following expression for the level width from (30):

$$\Gamma_n = \frac{2\gamma}{\pi\nu^3}\cos^2(\pi\mu),$$

which takes into account the quantum defect μ in the basis of the Coulomb functions [16]. For $V_{0q} \neq 0$, we have

$$\Gamma_n = \frac{2\gamma}{\pi\nu^3}\cos^2(\pi\nu),$$

where ν is determined from Eq. (30) for $\tan(\pi\nu)$. Thus, it turns out that the width Γ_n depends on the level shift: the closer a perturbed Rydberg level to the center

of the Rydberg interval $\nu = n \pm 1/2$, the smaller its width and, hence, the greater the lifetime of the corresponding Rydberg state.

Earlier, formula (30) was obtained in [10] for a Rydberg state that is not perturbed by the ionic core, i.e., for $t_{qq}^{e} = -\tan(\pi\mu_l) = 0$. However, one can see that, when $t_{qq}^{e} \neq 0$, the ionic core can appreciably affect the decay characteristics of the states that are perturbed by an external field under the condition that $V_{0q}^2 n^3 \sim t_{qq}^{e}$. The quantum defects μ_l for small l may reach considerable values, $\mu_0 = 0.35$ and $\mu_1 = -0.14$ for Na, so that $t_{qq}^{e} = 0.47$ for the population of states of the P series. Thus, there exists a domain of the spectrum where the interaction between a Rydberg electron and the ionic core plays a significant role. Under strong fields such that $V_{0q}^2 n^3 \gg 1$, the effect of the ionic core is insignificant and the difference between formulas (30) and (11) of [10] vanishes.

4. DISTRIBUTION OF EXCITED RYDBERG STATES OVER QUANTUM NUMBERS n

Selective excitation conditions for the nth state are satisfied when

$$|E_0 - E_n| \ll |E_0 - E_{n'}|_{n' \neq n}$$

both under moderately high and high strength of the field. In the first case, the restriction $|\tilde{E}_0 - \tilde{E}_n| \gg |V_{0n}|$ is imposed on the excited Rydberg states, which implies that $\Gamma_0 \ll \Gamma_n$. In this case, the length of a laser impulse should be such that $\Gamma_0 t_0 \ll 1$ and $\Gamma_n t_0 > 1$. Then,

$$|C_n(t_0)|^2 = V_{0n}^2 \frac{1}{|E_0 - E_n|^2} \gg |C_{n'}(t_0)|_{n' \neq n}^2. \quad (31)$$

Under higher strength of the field,

$$|V_{0n}| \sim |E_0 - E_n| \ll |E_0 - E_{n'}|_{n' \neq n}.$$

The selectivity of the excitation of a chosen Rydberg state is preserved under the additional constraints

$$\Gamma_0 t_0 \ll 1, \quad \Gamma_n t_0 \ll 1, \quad (32)$$

since here $\Gamma_0 \sim \Gamma_n$. Then,

$$W_n = |C_n(t_0)|^2 = \frac{4V_{0n}^2}{|\tilde{E}_n - \tilde{E}_0|^2} \sin^2\left(\frac{E_n' - E_0'}{2} t_0\right). \quad (33)$$

Note that, for $|E_n' - E_0'| \gg \Gamma_0, \Gamma_n$, formula (33) reproduces the well-known result of the two-level approximation.

The expressions given above show that, by varying the intensity $I \sim V_{0n}^2$, frequency ω_f, and the duration t_0 of a laser impulse, one can obtain the most favorable conditions for the selective excitation of a Rydberg atom to fixed states.

Under nonresonance conditions $|E_0 - E_n| \sim |E_n - E_{n\pm1}|$, a whole group of states is excited. In the case of a strong field $|E_0 - E_n| \sim V_{0n}$, the highest intensity is attained for $\Gamma_0 t_0, \Gamma_n t_0 \ll 1$ and the distribution of the population probability is expressed as $W_n = |C_n(t_0)|^2$, in which $C_n(t_0)$ is given by an appropriate formula in (19), where $\tilde{E}_0$ and $\tilde{E}_n$ are replaced by real quantities E_0' and E_n' that characterize the positions of field-perturbed levels.

Figures 2 and 3 show typical diagrams of the Rydberg levels for hydrogen and sodium atoms and typical populations of the Rydberg levels versus the duration t_0 of a laser impulse in atomic units for $V_{0q}^2 = 10^{-3}$.

Now, let us discuss the structure of Rydberg wave packets that arise for $t > t_0$, i.e., after the transmission of a laser impulse,

$$\Psi(\mathbf{r}, t) = \sum_n C_n(t_0) \Phi_n^{(q)}(\mathbf{r}) \exp(-iE_n(t - t_0)) \quad (34)$$

for $r > r_0$. Here, $\Phi_n^{(q)}(\mathbf{r})$ is a stationary wave function, normalized to unity, of a Rydberg electron of the qth series, E_n is the position of the nth level in the absence of field, and $\mathbf{r}$ is the radius-vector of a Rydberg electron; it is assumed that the admixed lower lying E_0 state decays rapidly as r increases in the range r_0.

The function $\Psi(\mathbf{r}, t)$ in (34) reproduces a sufficiently complicated pattern in the distribution of the electron density $|\Psi(\mathbf{r}, t)|^2$ in space and time.

Notice the following interesting property of this function that is attributed to the equidistant character of the Rydberg levels $E_n - E_{n_0} \approx \Delta n / n_0^3$: the recovery of a wave packet for $n_0^{-3}(t - t_0) = 2\pi k$, where k is an integer and E_{n_0} is the level closest to E_0. This phenomenon occurs for sufficiently small evolution times of a Rydberg wave packet at $t - t_0 \sim 10^{-3}$–10^{-12} s for the principal quantum number n_0 ranging from 10 to 30 and is analogous to the vibrational excitation of diatomic molecules observed in femtosecond experiments [18].

In the case of a hydrogen atom, when the quasilevel E_0 lies at the midpoint between the levels n_0 and $n + 1$ and a group of states in the interval $\Delta n \ll n_0$ is excited by a laser impulse, the amplitude $C_n(t_0)$ for $\Delta n t_0 / n_0^3 \ll 1$ proves to be antisymmetric with respect to $n' = n_0 + 1/2$: $C_{n' + \Delta n} = -C_{n' - \Delta n}$. Then, for $\Delta n |t - t_0| / n_0^3 \ll 1$ and suffi-

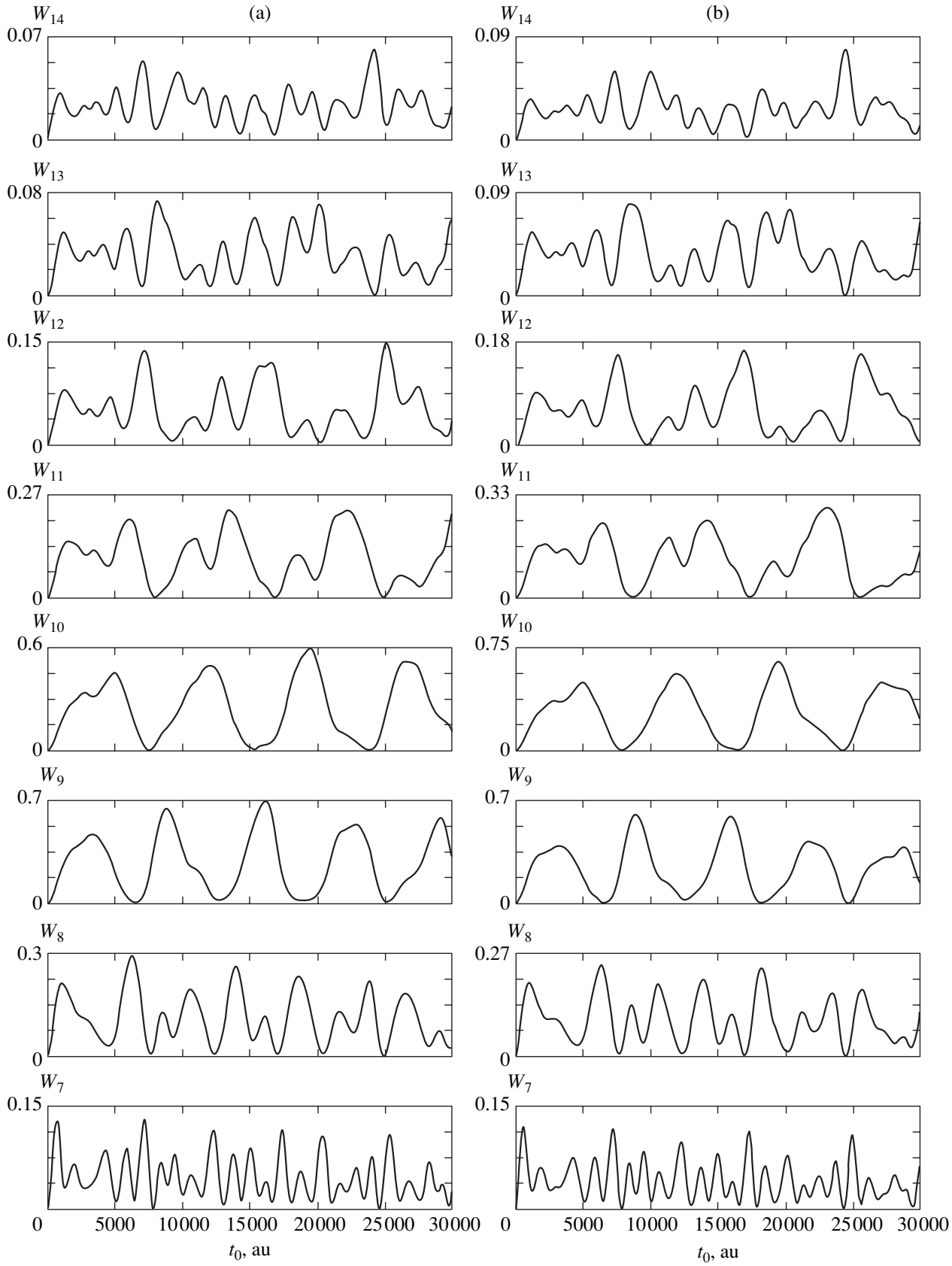

Fig. 2. Population probabilities of Rydberg states in the range of n = 7–14 for a hydrogen atom versus the duration t_0 of a laser impulse; (a) $\nu_0 = (-2(E_0^0 + \omega_f))^{-1/2} = 9.5$ and (b) $\nu_0 = 9.7$.

ciently small r such that $r^{3/2}|E| \ll 1$ ($E = -1/2n^2$), where the wave function $\Phi_n^q(\mathbf{r})$ of a Rydberg electron is virtually independent of n, an interference suppression of state (34) occurs. This means that, for a short period of time, the Rydberg electron is expelled to a large distance from the ionic core where its interaction with the radiation field is weaker. Note that this phenomenon occurs under special physical conditions and, unlike the phenomena considered in the Kramers–Henneberger approximation [19], is not associated with a variation in the form of the potential acting on the Rydberg electron.

5. ROTATIONAL ORIENTATION OF ELECTRON-EXCITED DIATOMIC MOLECULES

In this section, we briefly consider the excitation of low-energy electron-rotational Rydberg states. Here, we can restrict ourselves to a two-level approximation if the Raby frequency Ω is much less than the distance between adjacent levels ($\Omega \le 1/n^3$). Within this problem, which is formulated with regard to the decay of the states participating in the process, one can consider an interesting quantum phenomenon—a rotational orientation of excited diatomic molecules by a linearly polarized radiation from a pulse source. We will consider this phenomenon under the condition $\Gamma_n t \le 1$, when the probability of ionization is small. Then, the main analysis can be based on the dependence of the excitation probability W of a molecule on the total angular momentum J and its projection M onto the direction of the field-strength vector $\mathbf{f}$ according to the simple two-level approximation formula (3.7) of [3],

$$W_{JM} = \frac{V_{JM}^2}{\Omega^2}\sin^2(\Omega t_0),$$

where

$$\Omega = \frac{1}{2}(\Delta^2 + 4V_{JM}^2)^{1/2}, \quad \Delta = E_{JM}^0 + \omega_f - E^{J\pm1,M},$$

and

$$V_{JM} = \frac{1}{2}(\mathbf{f}\cdot\mathbf{D})_{JM,J\pm1,M}$$

is a field-induced interaction between the electron-rotational states under consideration. It is assumed that the excited state is characterized by a clear-cut vibration–rotation structure. The rotation constant B of low-energy excited states of most molecules (NO, CO, etc.) is commensurable to B in the ground electron state. Therefore, the ground and excited rotational states may be sufficiently well separated in energy ($BJ > \Delta$). The initial rotational state with energy E_{JM}^0 is assigned an angular momentum of $J \gg 1$ and projection M onto the direction of the polarization vector $\mathbf{f}$ of external radia-

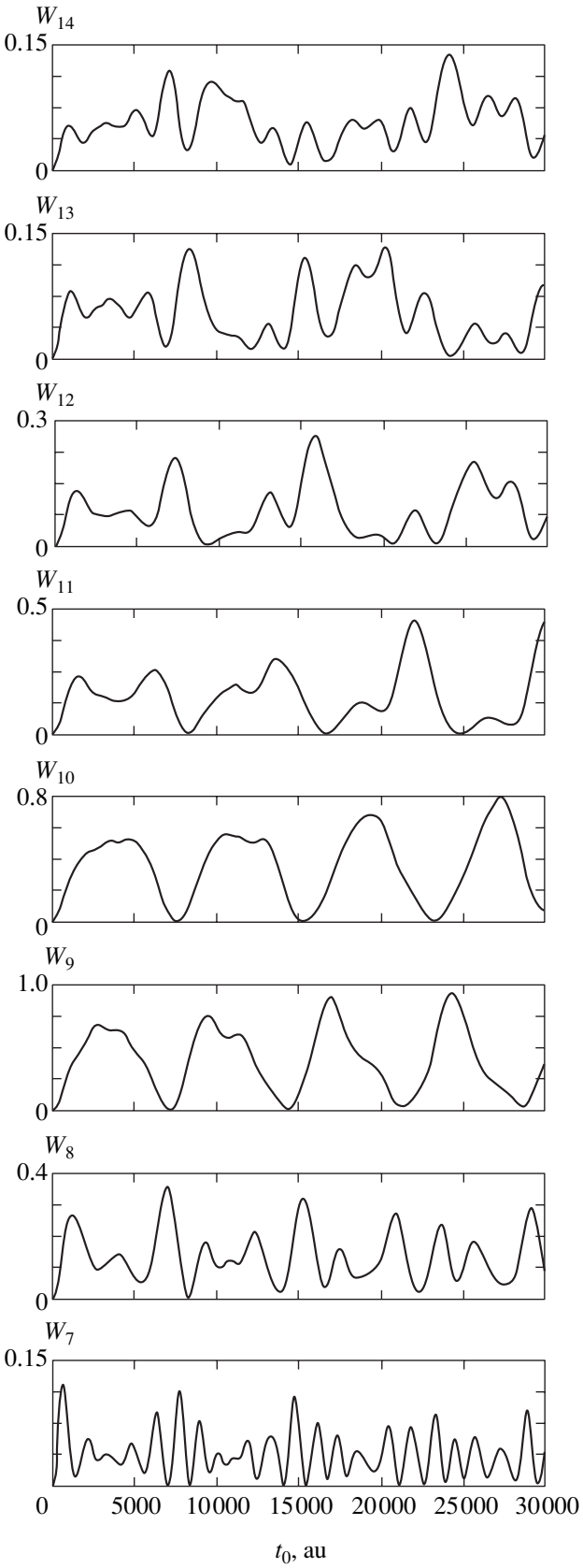

Fig. 3. Population probabilities of Rydberg states in the range of n = 7–14 for a sodium atom versus the duration t_0 of a laser impulse for ν_0 = 9.7.

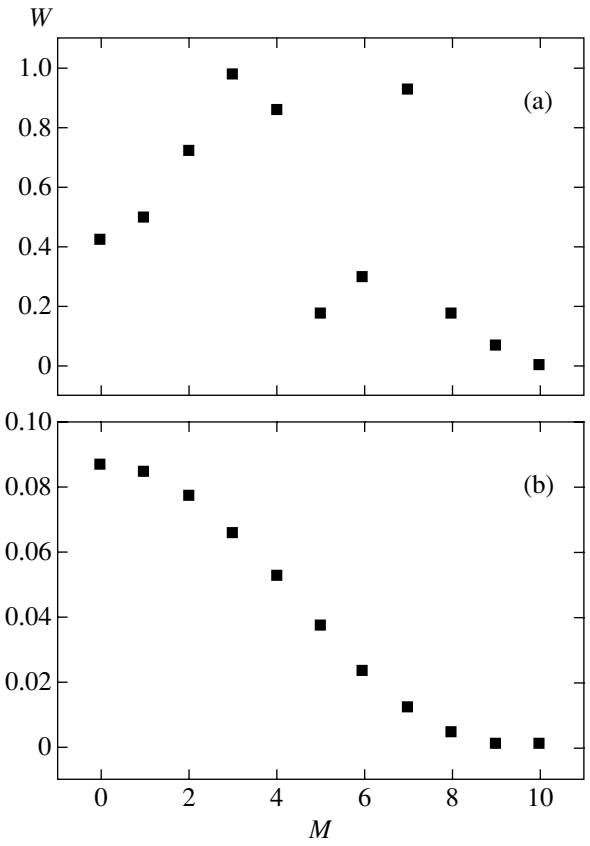

Fig. 4. Characteristic distributions of electron-excited NO molecules versus the projections M of the rotational angular momentum for various characteristics of the exciting laser impulse for $J = 10$; (a) a zero defect of a resonance (formula (35) for $V_{J0}^2 t_0 = 15$) and (b) $\Delta \gg V_{J0}$ (formula (36) for $2\pi k V_{J0}^2/\Delta^2 = 0.3$).

tion. During a direct transition $J, M \longrightarrow J \pm 1, M$, the quantity M is invariant. It is interesting to consider the dependence of excited (due to a dipole transition) states on M against the background of a uniform (in M) distribution of the initial rotational states of a molecule, taking into account that [20]

$$V_{JM,\,J\pm1,\,M} = V_{J0}\sqrt{1-\frac{M^2}{J^2}},$$

where

$$V_{J0} = \frac{1}{2}(\mathbf{f}\cdot\mathbf{D})_{J0,\,J\pm1,\,0}$$

is a field interaction of resonant states for $M = 0$.

For a zero defect of a resonance $\Delta = 0$, $W(J, M)$ satisfies the following relation:

$$W(J,M) = \sin^2\left(V_{J0}\sqrt{1-\frac{M^2}{J^2}}\,t_0\right). \qquad (35)$$

In the other limiting case $\Delta \ll |V_{J0}|$, representing Ω as

$$\Omega = \frac{1}{2}\Delta + \frac{V_{JM}^2}{\Delta},$$

we obtain the following expression for a special choice of the duration t_0 of a laser impulse such that $t_0 = 2\pi k/\Delta$, where k is an integer:

$$W = \frac{V_{JM}^2}{\Omega^2}\sin^2\frac{V_{JM}^2 t_0}{\Delta} \propto \sqrt{1-\frac{M^2}{J^2}}\sin^2\left[2\pi k\frac{V_{J0}^2}{\Delta^2}\left(1-\frac{M^2}{J^2}\right)\right]. \qquad (36)$$

Therefore, for various values of V_{J0} and t_0, we obtain different distributions of the excitation probability versus M; these distributions may vary over sufficiently wide ranges. It is essential that the processes considered above exhibit states that are indistinguishable in energy.

Figure 4 represents typical distributions of electron-excited molecules versus the projections M of the rotational angular momentum for various parameters of the exciting laser impulse.

6. CONCLUSION

This work has been carried out within the general program of investigations on the problem of forming highly excited Rydberg states of atoms and molecules under laser radiation. The analysis is based on sufficiently generally accepted simplifications of the impulse shape that presume that a time-periodic perturbation is turned on and off instantaneously. Nevertheless, this approximation requires new approaches to this important practical problem because of the complexity of the traditional analysis of multilevel systems.

We have demonstrated that the problem considered is reduced to the determination of the positions and widths of the energy levels of the system in the field of a monochromatic laser radiation of the same frequency and strength as those in a pulse mode. Therefore, instead of the direct solution of the nonstationary Schrödinger equation, we can apply a stationary method of a radiation-collision matrix [14–16]. Many investigations were carried out in a stationary formulation [21], and the results of these investigations can be directly applied to the problems of pulse excitation of complex (as compared with earlier studied) quantum systems. In this case, one can take into consideration the effect of the ionic core and its complex vibration–rotation structure in molecules, which leads to various forms (in different spectral domains) of nonadiabatic coupling between electron and nuclear motions. In the integral formulation, the method of radiation-collision matrices reduces the problem to transcendental equa-

tions, which are solved in two steps for moderately high field intensities (see condition (6)). First, assuming that the broadening of levels is negligible, one determines their positions. Then, from the same equations, one derives analytic formulas that establish a relation between the widths of the levels and their positions within characteristic Rydberg intervals.

The results obtained show that separate Rydberg states or groups of states can be stabilized according to the predictions of numerous studies devoted to this subject field in simpler cases. In the present work, the effect of stabilization is also demonstrated on the specific features of the behavior of Rydberg wave packets.

In conclusion, we note that the assumption used in this model that the rise and decay times of the impulse are negligible may lead to certain errors in the results obtained. However, if the characteristic rise–decay time τ satisfies the condition

$$|V_{0n}|\tau \ll 1,$$

this effect can be taken into account within perturbation theory and does not substantially influence the main results of the present work.

ACKNOWLEDGMENTS

We are grateful to G.V. Golubkov for useful discussions.

REFERENCES

1. M. Gruebele and A. H. Zewail, Phys. Today **43** (5), 24 (1990).
2. V. V. Lozovoĭ, O. M. Sarkisov, and S. Ya. Umanskiĭ, Khim. Fiz. **14**, 83 (1995).
3. N. B. Delone and V. P. Kraĭnov, *Atom in a Strong Light Field*, 2nd ed. (Énergoatomizdat, Moscow, 1984; Springer, Berlin, 1985).
4. A. N. Andryushin, A. E. Kazakov, and M. V. Fedorov, Zh. Éksp. Teor. Fiz. **76**, 1907 (1979) [Sov. Phys. JETP **49**, 966 (1979)].
5. A. N. Andryushin, A. E. Kazakov, and M. V. Fedorov, Zh. Éksp. Teor. Fiz. **82**, 91 (1982) [Sov. Phys. JETP **55**, 53 (1982)].
6. N. P. Poluéktov and M. V. Fedorov, Zh. Éksp. Teor. Fiz. **114**, 821 (1998) [JETP **87**, 445 (1998)].
7. S. H. Tersini, P. Gaspard, and S. A. Rice, J. Chem. Phys. **93**, 1670 (1990).
8. M. V. Fedorov, Kvantovaya Élektron. (Moscow) **28**, 19 (1999).
9. P. P. Poluéktov and M. V. Fedorov, Zh. Éksp. Teor. Fiz. **117**, 913 (2000) [JETP **90**, 794 (2000)].
10. M. Yu. Ivanov, Phys. Rev. A **49**, 1165 (1994).
11. E. A. Volkova, A. M. Popov, and O. V. Tikhonov, Zh. Éksp. Teor. Fiz. **116**, 1929 (1999) [JETP **89**, 1045 (1999)].
12. G. K. Ivanov, G. V. Golubkov, and S. V. Drygin, Zh. Éksp. Teor. Fiz. **107**, 1503 (1995) [JETP **80**, 840 (1995)].
13. M. E. Sukharev and V. P. Kraĭnov, Zh. Éksp. Teor. Fiz. **113**, 573 (1998) [JETP **86**, 318 (1998)].
14. G. K. Ivanov, Chem. Phys. Lett. **135**, 89 (1987).
15. G. K. Ivanov, A. S. Vartazaryan, and G. V. Golubkov, Dokl. Akad. Nauk **303**, 390 (1998).
16. G. K. Ivanov and G. V. Golubkov, Zh. Éksp. Teor. Fiz. **99**, 1404 (1991) [Sov. Phys. JETP **72**, 783 (1991)].
17. G. K. Ivanov, G. V. Golubkov, and D. M. Manakov, Zh. Éksp. Teor. Fiz. **106**, 1306 (1994) [JETP **79**, 707 (1994)].
18. O. M. Sarkisov and S. Ya. Umanskiĭ, Usp. Khim. **70**, 515 (2001).
19. W. C. Henneberger, Phys. Rev. Lett. **21**, 838 (1968).
20. K. V. Nikol'skiĭ, *Quantum Mechanics of Molecules* (GTTI, Moscow, 1934).
21. G. K. Ivanov and G. V. Golubkov, Zh. Éksp. Teor. Fiz. **115**, 1987 (1999) [JETP **88**, 1087 (1999)].

Translated by I. Nikitin

Journal of Experimental and Theoretical Physics, Vol. 96, No. 3, 2003, pp. 418–424.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 123, No. 3, 2003, pp. 474–481.

ATOMS, SPECTRA, RADIATION

Dependence of the Refractive Index on Crystal Thickness due to a Pekar Extra Light Wave

M. I. Strashnikova[a,*] and E. V. Mozdor[b,]**

[a] *Institute of Physics, National Academy of Sciences of Ukraine, Kiev, 03028 Ukraine*
[b] *Institute of Semiconductor Physics, National Academy of Sciences of Ukraine, Kiev, 03028 Ukraine*
**e-mail: strashnik@iop.kiev.ua*
***e-mail: sheka@class.semicond.ua*

Received November 2, 2002

Abstract—Transmission spectra of CdS crystals of different thickness are calculated in the region of $A_{n=1}$ exciton using the Pekar theory taking into account an extra light wave. It is shown that the dependence of the refractive index variance on the thickness of superthin crystals at low temperatures, which has been observed earlier, is associated with interference of the extra light wave, to which the crystal is opaque, with the fundamental wave transmitted by the crystal. © *2003 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

Extra light waves were predicted in a fundamental paper by Pekar published in 1957 [1], which was later classified as a discovery in Soviet science [2]. The results of many years of investigations in this field were generalized in the monograph [3]. Extra light waves were predicted as a result of inclusion of spatial dispersion of permittivity $\varepsilon(\omega, \mathbf{k})$, i.e., the dependence of ε not only on frequency, but also on the wave vector of a light wave. These waves possess the same frequency and polarization as the fundamental waves, but are characterized by different refractive indices and, hence, different velocities of propagation. Since the two waves are coherent, they experience interference, which must be manifested in complex dependences of the intensity and phase of transmitted and reflected light waves on the thickness d of the crystal and on the wavelength λ of the light wave.

Investigation of the effect of interference of Pekar waves on optical properties of crystals began soon after the theory of extra light waves had been developed. However, this phenomenon could be observed experimentally only many years later. For example, it was reported in a series of publications [4–6] that the Fabry–Perot interference (FPI) pattern in CdSe and CdS crystals at $T = 4.2$ and 1.8 K in the region of $A_{n=1}$ exciton for energies $\omega < \omega_L$, where ω_L is the longitudinal frequency, exhibits a usual single-mode character, while at $\omega > \omega_L$ it shows two modes: a densely periodic weak structure associated with interference with an extra wave is superimposed on deep extrema of the fundamental wave with a large period. The obtained spectral dependence was correctly described by the Pekar theory.

Simultaneously, it was shown for the first time in [7, 8] using a direct method for measuring the phase of an optical wave transmitted through a crystal that the dispersion of the refractive index $n(\omega)$ within the absorption band for the $A_{n=1}$ exciton has no classical "anomalous" region, but increases monotonically up to frequency ω_L. In Pekar's terminology, this branch is referred to as the plus wave (see Fig. 1 below). Starting from frequency ω_L, the main contribution to energy transfer through the crystal comes from the minus wave, and dispersion $n(\omega)$ measured at this frequency experiences a jump from the plus to the minus branch. Below frequency ω_L, the minus wave is not transmitted through the crystal; however, it produces a strong effect

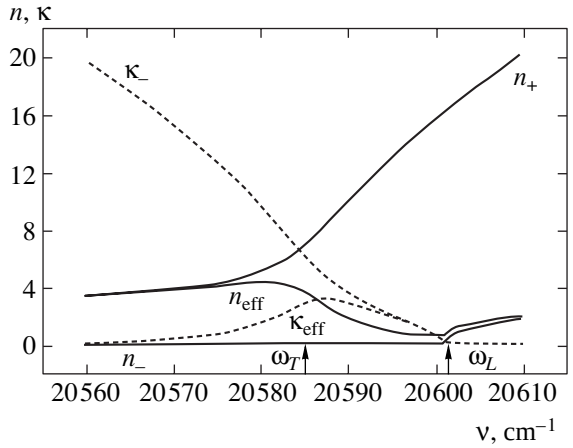

Fig. 1. Real (solid curves) and imaginary (dashed curves) parts of the complex refractive indices of the plus and minus waves as well as the effective refractive index for the reflectance spectrum. Energy values on the abscissa axis are given in units of the wave number $\nu = 1/\lambda$ taking into account the correction to the refractive index of air according to Kayzer's tables [24]; ω_L and $\omega_T \equiv \omega_0$ are the frequencies of longitudinal and transverse excitons.

on the reflection spectrum, which has a conventional classical form with an anomalous region within the absorption band halfwidth. It contradicts the reflection spectrum calculated using the Fresnel formula even qualitatively if we substitute into this formula the values of $n(\omega)$ measured in light transmitted through the crystal [9]. In order to match experiment with theory, we must introduce an effective refractive index n_{eff}, which is a complex function of the refractive indices of the plus and minus waves [10]. We can state that the emergence of n_{eff} itself is a result of interference of the plus and minus waves in reflected light. An extra wave also radically changes the spectral dependence of the phase of the reflected wave [11, 12].

Subsequently, the dispersion curves $n(\omega)$ were measured experimentally on a wedge-shaped CdS crystal from the deflection of a laser beam [13–15]. Only the plus wave branch increasing with frequency was detected in a spectral range below ω_L at $T = 1.8$ K, while both branches (fundamental minus wave and extra plus wave) could be measured in light transmitted through a crystal at frequencies above ω_L. In contrast to plane-parallel plates, two beams at the exit from a wedge-shaped crystal are separated in space and hence do not interfere with each other. The refractive indices and intensities are measured for each wave separately.

Thus, it has been reliably established on the basis of the results obtained in the above-mentioned publications that only one wave passes through a CdS crystal at low temperatures in the region of an $A_{n=1}$ exciton allowed in the $\mathbf{E} \perp \mathbf{c}$ polarization ($\mathbf{c}$ is the optical axis of the crystal) at frequencies lower than the longitudinal exciton frequency, while two identically polarized waves ($\mathbf{E} \perp \mathbf{c}$), but with different refractive indices, are transmitted through the crystal at frequencies exceeding this value.

Since all measurements of optical characteristics in transmitted light were made using classical methods and formulas of single-wave crystal optics (see below), it appeared that these methods can rightfully be used in other cases also (in particular, below frequency ω_L, when only one wave passes through the crystal). So, FPI patterns and polarized beam interference (PBI) patterns were analyzed in [16, 17] using relations of classical crystal optics and were used for calculating "experimental" dispersion curves $n(\omega)$. It was found that the thinner the crystal, the smaller its refractive index in the important resonance region of the spectrum and the lower the position of the $n(\omega)$ curve relative to that for a "thick" crystal. In this way, a dependence of the refractive index dispersion in superthin CdS crystals ($d < 1$ μm) on its thickness was observed. "Experimental" curves were approximated in accordance with the Pekar theory taking into account the exciton effective mass and damping constant, the oscillator strength of the transition, and upper excited states. The Myasnikov theory described in [18–21], according to which the exciton–phonon interaction (i.e., the transition oscillator strength) decreases with the crystal thickness, is in qualitative agreement with these experimental facts [18–21].

We have calculated the transparency spectrum of crystal plates taking into account an extra light wave. It was found that the observed dependence of the $n(\omega)$ curves on the thickness is associated with the effect of the extra wave on the phase of transmitted light. It turned out that the minus wave (although it is not transmitted through the crystal plate) affects not only the reflected wave, but also the resultant phase of transmitted light. In other words, it was found that the application of the formulas of classical crystal optics is not justified for very thin crystals and the refractive index of the plus wave calculated on the basis of these formulas is lower than that for a thick crystal. The results of calculations given below substantiate this statement.

2. FORMULAS OF CLASSICAL CRYSTAL OPTICS USED FOR CALCULATING DISPERSION CURVES

The dispersion of the refractive index in the range of lower exciton states in a CdS crystal at low temperatures was measured for the first time in [22] and then in [7, 8] using a Jamin interferometer crossed with a spectrograph. In these publications, experimental data were processed using formulas of classical single-wave crystal optics, according to which only one wave with refractive index n propagates in a crystal with birefringence in each polarization, and the phase incursion of a wave passing through a crystal of thickness d is $\Delta\varphi = nd(2\pi/\lambda)$. Consequently, the computation formula used in such measurements has the form

$$d(n-1) = \frac{\Delta y}{H}\lambda, \tag{1}$$

where $\Delta y/H$ is the ratio of the zero-order fringe shift upon the introduction of the crystal into one of the arms of the interferometer to the width of the interference fringe. In view of the fundamental importance of the result obtained in [7] (the absence of an anomalous region on the $n_\perp(\lambda)$ curve within the absorption band halfwidth), which contradicts basic concepts of the classical oscillator model of exciton absorption, in addition a method involving interference of polarized beams was employed. Calculations were based on the formula

$$\Delta n d = m\lambda, \tag{2}$$

where $\Delta n = n_\perp - n_\parallel$ is the difference in the refractive indices for light with polarizations $\mathbf{E} \perp \mathbf{c}$ and $\mathbf{E} \parallel \mathbf{c}$, m being an integer or a half-integer determining the interference order. In the component corresponding to polarization $\mathbf{E} \parallel \mathbf{c}$, we have $n_\parallel(\lambda) \approx \text{const}$ and $n_\perp(\lambda) = n_\parallel + \Delta n(\lambda)$. The correctness of the dependence obtained

was verified, and it was found that the two $n_\perp(\lambda)$curves coincide.

In [4–6], dispersion $n_\perp(\lambda)$ was measured from the Fabry–Perot interference pattern in polarization $\mathbf{E} \perp \mathbf{c}$. The pattern was interpreted using the Pekar theory as well as the formula

$$2dn_\perp = k\lambda, \tag{3}$$

where k is the interference order.

Analogously to these publications, formulas (2) and (3) were also used in [16, 17] for calculating the $n_\perp(\lambda)$ curves from the FPI and PBI. In order to reduce the error in measuring thickness and to eliminate the difference in the properties of different samples, measurements were made on the same wedge-shaped crystal. In the first series of measurements, the wedge was positioned vertically in the cell of a cryostat so that its thickness varied in the vertical plane. The crystal image was projected by a microscopic attachment to the slit of the spectrograph, which cut a variable-thickness cross section from the image. In this case, the interference pattern of polarized beams was photographed and the spectral $\Delta n(\lambda)$ dependences were calculated by formula (2) for wedge regions with multiple thicknesses. The curves were plotted on the basis of the assumption that n is independent of the crystal thickness; for this reason, multiple-thickness cross sections were chosen so that they had multiple interference orders m over a wavelength λ' corresponding to the nearest longitudinal exciton frequency. The curves plotted for four segments and having common point λ' according to the principle of their construction strongly diverge towards the long-wave region, indicating that the initial assumption concerning of n being independent of thickness is erroneous.

In the second series of measurements, the wedge was arranged horizontally and its thickness varied in the same direction. In view of a very small value of the angle of refraction of the crystal, we could assume that the slit cuts a constant-thickness cross section from the image. In this geometry, the FPI and PBI patterns were photographed for different thicknesses. On the basis of the spectral positions of the extrema in these patterns, the $\Delta n(\lambda)$ and $n_\perp(\lambda)$ curves were calculated using formulas (2) and (3). The spectral dependences $n_\perp(\lambda)$ obtained using these two formulas coincided. It was noted above that the curves measured in this way for thinner parts of the crystal were found to be displaced toward shorter waves as compared to those measured for a thicker part.

In connection with this result, we can mention publication [23], in which the measurements of the $n_\perp(\lambda)$ dependence for CdS and CdSe crystals with varying thickness were reported. The authors of this publication observed a short-wave shift of the $n(\lambda)$ curves upon a decrease in the crystal thickness, but attributed it to a change in the resonance frequency ω_0 of the exciton transition. In order to reach the coincidence of the dispersion curves measured for CdS crystals of different thickness, the curves had to be displaced over distances from 1 to 3 cm^{-1}. The curves obtained by us are also shifted by approximately the same distance. However, since we were tracing the spectral position of a longitudinal exciton, which practically did not depend on the thickness, we assigned the observed effect to a decrease in the longitudinal–transverse splitting $\Delta_{LT} = \omega_L - \omega_T$, which is proportional to the oscillator strength of the exciton transition.

It is interesting to note that the dispersion curve measured in [7], which was calculated by formula (1), could be approximated in [10] by the theoretical dependence for a value of Δ_{LT} slightly smaller than for thick crystals.

In order to clarify the situation, we calculated the transparency spectra of CdS plates of different thickness using the Pekar theory and compared the spectral positions of the extrema in the Farby–Perot interference pattern with those calculated using the formulas of single-wave crystal optics.

3. CALCULATION OF THE TRANSPARENCY OF CDS CRYSTAL PLATES ON THE BASIS OF THE PEKAR THEORY

The refractive indices for two waves emerging in a crystal due to spatial dispersion of permittivity can be determined, according to Pekar [3], from the relation

$$\tilde{n}_\pm^2 = \frac{1}{2}(\mu + \varepsilon_0) \pm \sqrt{\frac{1}{4}(\mu - \varepsilon_0)^2 + b}, \tag{4}$$

where

$$\mu = \mu' + i\mu'', \quad \mu' = \frac{2Mc^2}{\hbar\omega_0^2}(\omega - \omega_0),$$

$$\mu'' = \frac{2Mc^2}{\hbar\omega_0^2}\Gamma, \quad b = \frac{2Mc^2}{\hbar\omega_0^2}\Delta_{LT}\varepsilon_0,$$

ε_0 is the background permittivity, ω_0 is the resonance frequency of a transverse exciton, c is the velocity of light, M is the translational mass of an exciton, Δ_{LT} is the longitudinal-transverse splitting, and Γ is the damping constant. For the parameters of the theory, we used the values obtained by us earlier in [13, 14] for the lower exciton $A_{n=1}$ at 4.2 K: $\Delta_{LT} = 16$ cm^{-1}, $\varepsilon_0 = 7.4$, $\omega_0 = 20585$ cm^{-1}, $\Gamma = 0.16$ cm^{-1}, and $M = 0.8m_e$. Figure 1 shows the real and imaginary parts of the complex refractive indices $\tilde{n}_\pm = n_\pm + i\kappa_\pm$ for the plus and minus waves obtained for such values of the parameters. In order to simplify the figure, the κ_+ branch is not depicted since it virtually coincides with the abscissa axis for such a small value of Γ: the absorption coefficient for the plus wave does not exceed 0.1 in the frequency range under investigation.

Figure 1 also shows the real and imaginary parts of the effective refractive index $\tilde{n}_{\text{eff}}$, which must be substituted into the Fresnel formulas to match the experimentally measured spectra and the phase variation upon reflection from a semiinfinite medium with the theoretical spectra,

$$\tilde{n}_{\text{eff}} = \frac{\varepsilon_0 + \tilde{n}_+\tilde{n}_-}{\tilde{n}_+ + \tilde{n}_-}. \quad (5)$$

The spectra of transparency T and reflectance R for CdS crystals of different thickness, taking into account the extra light wave, were calculated from the relations given in [3]:

$$T = \frac{4|G|^2}{\left|(1+iF)^2 + G^2\right|^2}, \quad (6)$$

$$R = \frac{\left|1 + F^2 - G^2\right|^2}{\left|(1+iF)^2 + G^2\right|^2}, \quad (7)$$

where

$$F = \frac{\tilde{n}_+}{1-q}\cot(k_+ d) + \frac{\tilde{n}_-}{1-1/q}\cot(k_- d),$$

$$G = \frac{\tilde{n}_+}{(1-q)\sin(k_+ d)} + \frac{\tilde{n}_-}{(1-1/q)\sin(k_- d)},$$

$$k_\pm = \frac{\omega}{c}\tilde{n}_\pm, \quad q = \frac{\tilde{n}_+^2 - \varepsilon_0}{\tilde{n}_-^2 - \varepsilon_0}.$$

Relations (6) and (7), as well as the formula for the effective refractive index, were derived using the Pekar extra boundary condition; namely, the exciton component of polarizability is equal to zero on the crystal surface: $|P_{\text{ex}}| = 0$.

Figure 2 shows the transparency spectra for crystal plates of different thickness. It can be seen that, in spite of the fact that we took multiple thicknesses, the spectral positions of interference extrema with multiple orders do not coincide. The thinner the crystal, the larger the shift of the corresponding extrema towards the short-wave region. Figure 3 shows for comparison analogous spectra calculated on the basis of the single-wave theory, i.e., corresponding to formula (3), which was used for processing the experimental dependences. The calculations were made using the program proposed in [25] and taking into account multiple reflections within the crystal plate. Values of $n_+(\omega)$ and $\kappa_+(\omega)$ were used as optical parameters of the medium since only the plus wave passes through the crystal for $\omega < \omega_L$. As expected, the positions of the extrema with multiple orders coincide exactly.

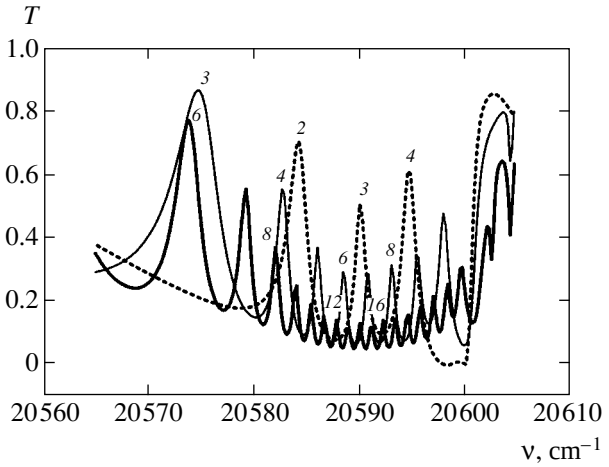

Fig. 2. Transparency spectra for three CdS crystals, calculated in accordance with the Pekar theory. The crystal plate thicknesses are equal to 0.09 (dashed curve), 0.18 (fine solid curve) and 0.36 μm (bold solid curve). The figures on the peaks indicate orders of interference.

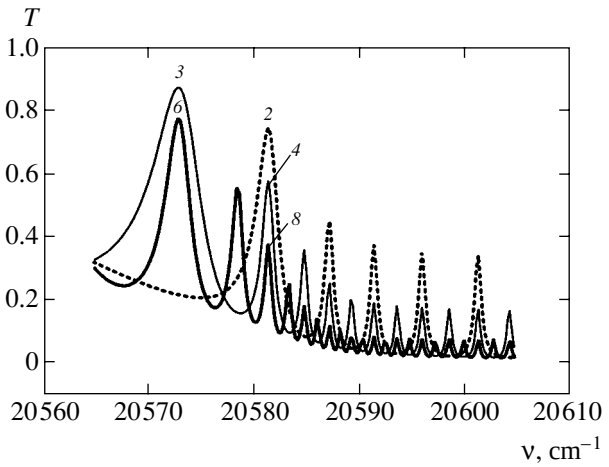

Fig. 3. Transparency spectra for the same crystals as in Fig. 2, calculated on the basis of the single-wave classical theory.

It follows from the above figures that the positions of interference extrema in the presence of an extra wave are determined not only by the values of n_+ and κ_+ of the fundamental wave transmitted through the crystal, but are also affected by the extra wave. The refractive index of the minus wave below the longitudinal frequency can be regarded as purely imaginary: $\kappa_- \neq 0$ and $n_- \approx 0$ (it does not exceed 0.2). In accordance with classical crystal optics, a medium with such optical parameters totally reflects light. It was noted above that the crystal is indeed opaque for the minus wave below ω_L. In accordance with the results of measurements from the laser beam deflection by a wedge-shaped crystal, which were made in [15] at different temperatures, the minus wave can pass through the crystal when its refractive index attains a value of $n_- \approx 1$ at $T = 7.7$ K due to an

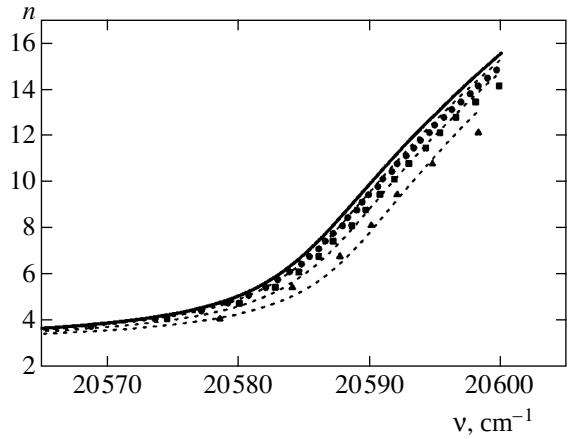

Fig. 4. Results of calculation of the dispersion curves $n(\omega)$ from the interference extrema shown in Fig. 2 by formula (3): d = 0.36 (•), 0.18 (▪), and 0.09 (▲) μm. The solid curve corresponds to $n_+(\omega)$ and the dashed curves are the result of approximation.

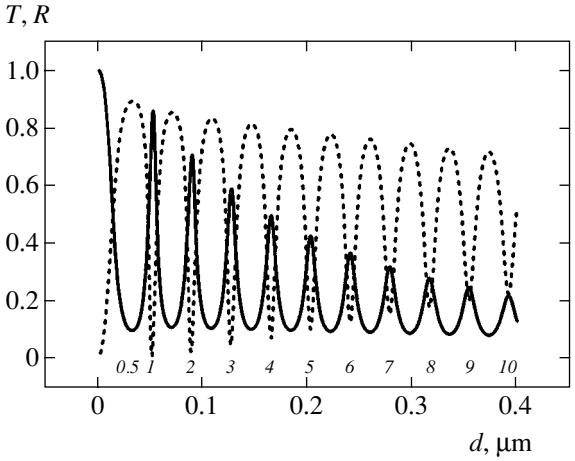

Fig. 5. Transparency T (solid curve) and reflectance R (dashed curve) as functions of the crystal thickness, calculated in accordance with the Pekar theory for ω = 20584.2 cm^{-1}. Figures indicate interference orders.

increase in Γ, the entire region of Δ_{LT} clearing up jumpwise. Thus, in spite of the fact that the crystal is opaque for the minus wave at T = 4.2 K, this wave changes the phase of transmitted light. For this reason, the values of refractive index calculated by formula (3) are averaged to a certain extent. This quantity is a sort of "effective" refractive index for transmitted light, which does not coincide, however, with the value of n_{eff} for reflection from a semiinfinite crystal (this can be seen from a comparison of Figs. 4 and 1).

Further, we tried to apply the procedure of evaluation of the refractive index, which was used for processing experimental data in [16, 17] (i.e., formula (3)) for the transparency spectra calculated taking into account the extra light wave (see Fig. 2). The results of these calculations are shown in Fig. 4. It can be seen from the figure that the thinner the crystal, the lower and more shifted to the right are the calculated values of $n(\omega)$ relative to the quantity $n_+(\omega)$, which was initially assumed in formula (6) and which was used for calculating the transparency spectrum. On the basis of the displaced $n(\omega)$ curves, we can choose new parameters of the theory and estimate the variation of Δ_{LT} in the case of coincidence of the new approximating curve with the dispersion curve obtained earlier. This procedure corresponds to that used by us in [16, 17] for determining the theoretical parameters for experimentally measured curves. Dashed curves in Fig. 4 approximate the "experimental" dependences (Δ_{LT} = 15.5, 14.5, and 13.1 cm^{-1} for a thickness of 0.36, 0.18, and 0.09 μm, respectively). The latter dependences can also be approximated by varying ω_0 alone, as was done in [23].

Thus, the shift of the dispersion curve observed in [16, 17, 23] is associated not with a decrease in Δ_{LT} or a shift of ω_0, but with the application of relations of single-wave crystal optics for experimental data processing. However, other relations are not available for experimenters.

It should be noted, by the way, that for an appreciable value of κ in an absorbing medium, the positions of extrema obtained in transmitted and reflected light do not coincide in view of an extra phase shift upon reflection from the absorbing medium. However, the damping constant of the $A_{n=1}$ exciton at low temperatures is very small and analysis of the $T(\omega)$ and $R(\omega)$ spectra according to the classical program, as well as the program taking into account the extra light wave, proved that the maxima of $T(\omega)$ virtually coincide with the $R(\omega)$ minima. Figure 5 also demonstrates a similar coincidence of extrema. This is another evidence of the fact that the shift of the $n(\omega)$ curves is not associated with absorption of crystals.

We have also carried out another series of calculations: the transparency and reflectance were calculated as functions of the crystal thickness for different wavelengths from the region of Δ_{LT}. An example of such calculations is shown in Fig. 5. It was found that the interference extrema are separated by different distances. The thinner the crystal, the larger the period of oscillations; in other words, in order to form the next interference peak, light in a thin crystal must traverse a larger distance than in a thick crystal. For example, the interference order k = 1 corresponds to $d \geq 0.05$ μm, while k = 10 corresponds to $d \leq 0.4$ μm. If we process the obtained dependences using relation (3), ascribing a certain interference order to each extremum, we obtain the curves depicted in Fig. 6. These curves clearly exhibit the crystal thickness dependence of the "effective" refractive index for the transmitted wave.

It should be noted that the above calculations do not claim exact quantitative agreement with experiment,

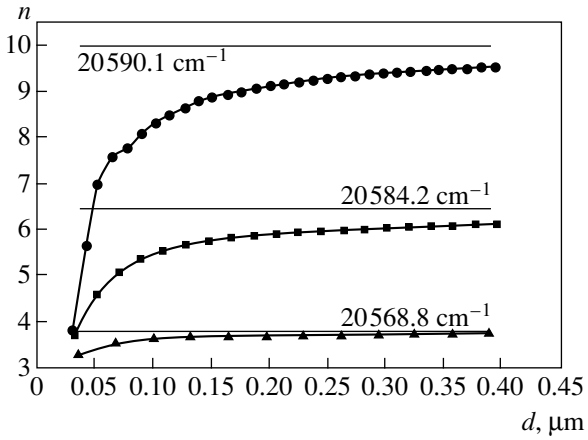

Fig. 6. Dependence of the "effective" refractive index for light transmitted through a crystal on its thickness for different frequencies from the resonance exciton region of the spectrum. Solid lines parallel to the abscissa axis indicate the values of n_+ for the same frequencies.

although the obtained shifts of dispersion curves are close to the shifts measured in experiments. First, it was noted above that formulas (6) and (7) were derived under the extra Pekar boundary condition: $|P_{ex}| = 0$ at the crystal surface. Indeed, it was proved in many publications, beginning with [26] (including those devoted to the measurement of the phase shift of the reflected wave [11, 12]), that there exists a "dead" layer on the crystal surface into which an exciton cannot penetrate. However, we are not aware of formulas for calculating T and R of a crystal plate taking into account the dead layer that effectively reduces the thickness. Second, the calculations were made for a constant value of Γ, which is not an exact approximation. The quantity Γ reciprocal of the lifetime of an exciton, which may be comparable to the mean free time between the crystal surfaces for very thin perfect crystals, can be a function of the sample thickness and frequency. For this reason, it is very difficult to attain quantitative agreement between the experimentally measured and theoretically calculated transparency spectra of crystals [10, 14].

4. CONCLUSIONS

We have calculated the spectral dependence of the transparency of CdS crystal plates of various thickness on the basis of the formulas of the Pekar' theory [3] taking into account an extra light wave. The calculations prove that the phase of the transmitted wave is a complex function of the refractive indices of both waves: the plus wave transmitted through the crystal and the minus wave reflected from it. Since all the formulas used by experimenters for processing spectrograms were derived under the assumption that a single wave passes through the crystal, the obtained results are averaged characteristics of the medium referred to unit thickness. In such calculations and with such an interpretation, the refractive index indeed becomes a function of sample thickness. However, in actual practice, the effect is a manifestation of the interference of the plus and minus waves, whose refractive indices remain unchanged. It is found that the effect of the minus wave is the stronger, the thinner the crystal.

Thus, the dependence of the dispersion of refractive index in the region of an $A_{n=1}$ exciton of a CdS crystal at low temperatures on its thickness observed in [16, 17] is associated with the effect of the Pekar extra wave on the phase of transmitted light.

It should be noted in conclusion that we do not deny the correctness of the theory developed in [18–21]. However, in order to verify this theory, use should be made of experimental data obtained that takes into account the Pekar extra wave rather than the results obtained on the basis of the formulas of the single-wave theory.

REFERENCES

1. S. I. Pekar, Zh. Éksp. Teor. Fiz. **33**, 1022 (1957) [Sov. Phys. JETP **6**, 785 (1957)].
2. S. I. Pekar, Certificate No. 323, OT-11003 (September 27, 1984); Otkrytiya, Izobret., No. 32, 3 (1987).
3. S. I. Pekar, *Crystal Optics and Additional Light Waves* (Naukova Dumka, Kiev, 1982).
4. V. A. Kiselev, B. S. Razbirin, and I. N. Ural'tsev, Pis'ma Zh. Éksp. Teor. Fiz. **18**, 504 (1973) [JETP Lett. **18**, 296 (1973)].
5. V. A. Kiselev, B. S. Razbirin, and I. N. Uraltsev, Phys. Status Solidi B **72**, 161 (1975).
6. I. V. Makarenko, I. N. Uraltsev, and V. A. Kiselev, Phys. Status Solidi B **98**, 773 (1980).
7. M. S. Brodin, N. A. Davydova, and M. I. Strashnikova, Pis'ma Zh. Éksp. Teor. Fiz. **19**, 567 (1974) [JETP Lett. **19**, 297 (1974)].
8. M. S. Brodin, N. A. Davydova, and M. I. Strashnikova, Phys. Status Solidi B **70**, 365 (1975).
9. M. I. Strashnikova, Fiz. Tverd. Tela (Leningrad) **17**, 729 (1975) [Sov. Phys. Solid State **17**, 467 (1975)].
10. S. I. Pekar and M. I. Strashnikova, Zh. Éksp. Teor. Fiz. **68**, 2047 (1975) [Sov. Phys. JETP **41**, 1024 (1975)].
11. L. E. Solov'ev and A. V. Babinskiĭ, Pis'ma Zh. Éksp. Teor. Fiz. **23**, 291 (1976) [JETP Lett. **23**, 263 (1976)].
12. A. V. Komarov, S. M. Ryabchenko, and M. I. Strashnikova, Zh. Éksp. Teor. Fiz. **74**, 251 (1978) [Sov. Phys. JETP **47**, 128 (1978)].
13. M. V. Lebedev, M. I. Strashnikova, V. V. Timofeev, and V. V. Chernyĭ, Pis'ma Zh. Éksp. Teor. Fiz. **39**, 366 (1984) [JETP Lett. **39**, 440 (1984)].
14. A. A. Demidenko, M. V. Lebedev, S. I. Pekar, *et al.*, Zh. Éksp. Teor. Fiz. **89**, 330 (1985) [Sov. Phys. JETP **62**, 185 (1985)].
15. M. V. Lebedev, M. I. Strashnikova, V. V. Timofeev, and V. V. Chernyĭ, Fiz. Tverd. Tela (Leningrad) **29**, 1948 (1987) [Sov. Phys. Solid State **29**, 1122 (1987)].

16. V. Ya. Reznichenko, M. I. Strashnikova, and V. V. Chernyĭ, Ukr. Fiz. Zh. **33**, 535 (1988).
17. V. Ya. Reznichenko, M. I. Strashnikova, and V. V. Cherny, Phys. Status Solidi B **152**, 675 (1989).
18. V. N. Myasnikov, S. V. Marisova, and A. N. Lipovchenko, Phys. Status Solidi B **117**, 109 (1983).
19. M. I. Strashnikova, V. Ya. Reznichenko, V. V. Cherny, and V. N. Myasnikov, Solid State Commun. **47**, 375 (1983).
20. V. N. Myasnikov, V. Ya. Reznichenko, M. I. Strashnikova, and V. V. Chernyĭ, Dokl. Akad. Nauk Ukr. SSR, Ser. A, No. 2, 61 (1984).
21. N. M. Ivanov and É. N. Myasnikov, Ukr. Fiz. Zh. **31**, 1648 (1986).
22. M. S. Brodin and M. I. Strashnikova, Fiz. Tverd. Tela (Leningrad) **4**, 2454 (1962) [Sov. Phys. Solid State **4**, 3352 (1962)].
23. V. A. Kiselev, I. N. Uraltsev, and I. V. Makarenko, Solid State Commun. **53**, 591 (1985).
24. H. Kayzer, *Tabelle der Schwingungszahlen* (Hirzelin, Leipzig, 1925).
25. A. Penzkofer, E. Drotleft, and W. Holzer, Opt. Commun. **158**, 221 (1998).
26. J. J. Hopfield and D. G. Thomas, Phys. Rev. **132**, 563 (1963).

Translated by N. Wadhwa

Journal of Experimental and Theoretical Physics, Vol. 96, No. 3, 2003, pp. 425–429.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 123, No. 3, 2003, pp. 482–486.

ATOMS, SPECTRA, RADIATION

Laser Detection of Optical Signals

V. P. Bykov

General Physics Institute, Russian Academy of Sciences, Moscow, 119991 Russia
e-mail: v.p.bykov@mtu-net.ru

Received October 14, 2002

Abstract—A scheme is proposed for detecting optical signals in which electrons bound in atoms, molecules, or ions are used as a sensitive element rather than free electrons, as in conventional detectors (photoelements, photomultipliers, etc.). It is shown that such a scheme has a high sensitivity and a reduced shot-noise level.

The problem of detecting optical signals by optical methods themselves was probably first formulated by Bloembergen [1]. However, this idea has not been adequately developed. It was shown in paper [2] that the signal-to-noise ratio could be improved by using nonlinear radiation detectors in which the signal transformation begins from an absorption transition. It was shown in papers [3–5] that photocounts appear in conventional photodetectors due to the strong Coulomb instability of a weak electron current produced in the detector by a signal being detected. This suggests that in detectors, instead of free electrons, the electrons that are bound in atoms, ions, or molecules are used, where they are well stabilized by a strong Coulomb field of nuclei.

In this paper, a possible scheme for detecting weak optical signals by laser means is described. At present microcavities are created which can be used in the Bloembergen scheme, providing the passage from spontaneous to stimulated effect, thereby substantially improving the scheme.

Our scheme is based on a system of three-level atoms (Fig. 1). It is assumed that the atoms are located in a specially designed optical (micro)cavity, in which there are a resonance signal mode at the frequency ω of the $|0\rangle \longleftrightarrow |1\rangle$ transition and two resonance modes at the frequency Ω of the $|1\rangle \longleftrightarrow |2\rangle$ transition. When levels $|1\rangle$ and $|2\rangle$ are not populated, the two last modes are degenerate and are not coupled with each other. One of these modes (pump mode) is excited; i.e., it contains a strong monochromatic pump field, which is determined by an external source. The second mode is intended for excitation of an output signal (hereafter, the output mode). In the initial state, i.e., before the arrival of a signal, this mode is not excited and contains no fields.

The principle of the detector operation is as follows. Before the arrival of a signal exciting the signal mode, an atom, not being in resonance with the pump field and not in fact interacting with the field, remains in the $|0\rangle$ state. After the arrival of a signal at the frequency ω, the $|1\rangle$ level becomes somewhat populated. In this case, a strong pump field induces transitions between levels $|1\rangle$ and $|2\rangle$, and an oscillating dipole moment appears at the frequency Ω, which excites a field in the output mode. The problem is to show that the output signal can be substantially greater than the input signal. In addition, it is necessary to find the characteristic rise time of the output signal. The study is performed in the so-called semiclassical approximation, when processes occurring in atoms are investigated quantum-mechanically, while all the fields are assumed classical.

Consider the case when the rise time of the output signal is shorter than the phase (transverse) relaxation time of active atoms in the medium. The state of a three-level atom changes under the action of three fields: the signal field, the pump field, and an initially unknown field of the output mode. The output mode field is excited by a polarization current of a system of atoms. Consideration of the evolution of atoms and excitation of the output mode leads to a self-consistent problem. The solution of this problem gives the field in the output cavity and the rise time of the output signal.

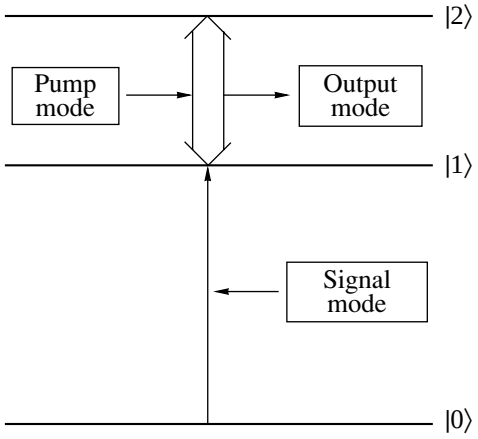

Fig. 1. Principal detection scheme.

Excitation of the field in the output mode of the cavity is described by the equation

$$\ddot{U}(t) + \Omega^2 U(t) = -4\pi\frac{\partial j}{\partial t}, \quad (1)$$

where

$$j(t) = \int d\mathbf{r}\mathbf{v}_o(\mathbf{r})\mathbf{j}(\mathbf{r}, t) \quad (2)$$

is the projection of the current $\mathbf{j}(t)$ on the normalized distribution of the mode field $\mathbf{v}_o(\mathbf{r})$ ($\int dV|\mathbf{v}_o(\mathbf{r})|^2 = 1$) of the eigenfrequency Ω, and $U(t)$ is the amplitude of this mode.

Below, we assume that the field amplitude $U(t)$ is a slowly varying function of time. Let us separate the negative-frequency part in it:

$$U^{(-)}(t) = u(t)e^{-i\Omega t}.$$

Then, by neglecting the second derivative of the amplitude $u(t)$ with respect to time, we transform the equation for excitation of the output mode to the form

$$\dot{u}(t) = -2\pi j'(t), \quad (3)$$

where $j'(t)$ is the amplitude of the negative-frequency part of the current $j(t)$,

$$j^{(-)}(t) = j'(t)e^{-i\Omega t}, \quad (4)$$

which also varies slowly in time.

The current exciting the output mode consists of elementary currents from individual atoms, which interact with three fields. The evolution of a three-level atom in the fields of three modes is described by the Schrödinger equation in the interaction representation

$$i\frac{\partial|\psi(t)\rangle}{\partial t} = W(t)|\psi(t)\rangle, \quad (5)$$

where

$$W(t) = [\alpha|2\rangle\langle 1| + \alpha^*|1\rangle\langle 2|] + [\beta|1\rangle\langle 0| + \beta^*|0\rangle\langle 1|] + [\gamma(t)|2\rangle\langle 1| + \gamma^*(t)|1\rangle\langle 2|] \quad (6)$$

and

$$\alpha = -\frac{erU}{\hbar}, \quad \beta = -\frac{er_iU_i}{\hbar}, \quad \gamma(t) = -\frac{er_oU_o(t)}{\hbar} \quad (7)$$

are the quantities describing the interaction of the atom with the field modes; U, U_i, and $U_o(t)$ are the negative-frequency field amplitudes of the pump, signal, and output modes, respectively; and

$$r = \mathbf{v}\cdot\mathbf{r}_{21}, \quad r_i = \mathbf{v}_i\cdot\mathbf{r}_{10}, \quad r_o = \mathbf{v}_o\cdot\mathbf{r}_{21} \quad (8)$$

are the projections of the matrix elements of the electron coordinate $\mathbf{r}$ on the amplitudes $\mathbf{v}$, $\mathbf{v}_i$, and $\mathbf{v}_o$ of the normalized pump, signal, and output modes at the atom's location (below, we assume in estimates that the average value of the normalized mode amplitude is $\bar{v} \approx \sqrt{V^2}$, where V is the mode volume). The quantities related to the pump mode have no indices, while the quantities related to the signal and output modes are indicated by indices i (in) and o (out).

The first term in the expression for $W(t)$ describes the interaction of the atom with the pump mode, taking into account that the pump frequency Ω coincides with the frequency Ω_0 of the $|1\rangle \longleftrightarrow |2\rangle$ transition. The second term describes the interaction of the atom with the output mode. Because the field amplitude of the output mode proportional to $\gamma(t)$ varies slowly in time, this field is not monochromatic. Its average frequency is taken equal to Ω. The third term in the expression for $W(t)$ described the interaction of the atom with the signal (detected) field at the frequency ω_0, which is equal to the resonance frequency ω of the $|0\rangle \longrightarrow |1\rangle$ transition.

Due to the action of the fields, the atom undergoes transition to the state

$$|\psi(t)\rangle = \psi_0|0\rangle + \psi_1|1\rangle + \psi_2|2\rangle, \quad (9)$$

where, according to perturbation theory,

$$\psi_0 \approx 1 - |\beta|^2t^2, \quad \psi_1 \approx -i\beta t, \quad \psi_2 \approx -\beta t\int_0^t dt_1 t_1(\alpha + \gamma(t)). \quad (10)$$

Because the current density operator of the atom is

$$\mathbf{J}(\mathbf{r}, \mathbf{q}) = \frac{e}{m}\mathbf{p}\delta(\mathbf{r} - \mathbf{q}), \quad (11)$$

we obtain the expression for the current at the $|1\rangle \longleftrightarrow |2\rangle$ transition in the interaction representation:

$$\mathbf{j}(\mathbf{r}, t) = \frac{e}{m}(\langle 2|\mathbf{p}\delta(\mathbf{r} - \mathbf{q})|1\rangle)|2\rangle\langle 1| + \text{H.c.} \quad (12)$$

The polarization current $j(t)$ exciting the output mode

has, in the dipole approximation, the form

$$\begin{aligned} j(t) &= n\int d\mathbf{r}\mathbf{v}_o(\mathbf{r})\mathbf{j}(\mathbf{r}, t) \\ &\approx \frac{e}{m}n\mathbf{v}_o(\langle 2|\mathbf{p}|1\rangle)|2\rangle\langle 1| + \text{H.c.} \\ &= ien\Omega r_0|2\rangle\langle 1| + \text{H.c.}, \end{aligned} \tag{13}$$

where n is the number of active atoms, $r_0 = \mathbf{v}_o(0)\mathbf{r}_{21}$, $\mathbf{r}_{21}$ is the matrix element of the electron coordinate for the $|2\rangle \longrightarrow |1\rangle$ transition. By averaging this expression over state (9), we obtain the current exciting the output mode of the cavity:

$$j(t) = -en\Omega r_0|\beta|^2 t\int_0^t dt_1 t_1(\alpha + \gamma(t_1)). \tag{14}$$

Therefore, according to (3) and (14), the output-mode field is determined by the equation

$$\dot{u}(t) = -2\pi iecnr_o^*|\beta|^2 t\int_0^t dt_1 t_1(\alpha + \gamma(t_1)). \tag{15}$$

In order to pass from $u(t)$ to $\gamma(t)$, we multiply this equation, according to (7), by $-er_o/\hbar$ to obtain

$$\dot{\gamma} = -\xi^2 t\int_0^t dt_1 t_1(\alpha + \gamma(t_1)), \tag{16}$$

where

$$\xi^2 = \frac{2\pi\varepsilon^2\Omega\omega^2 n|r_o|^2|r_i|^2|u_i|^2}{\hbar c}. \tag{17}$$

Taking into account that

$$|u_i|^2 = \frac{2\pi\hbar c^2 N}{\omega}, \tag{18}$$

where N is the number of photons in the signal mode, we obtain

$$\xi^2 = (2\pi)^2\varepsilon^2 v_o^2 v_i^2|r_{21}|^2|r_{10}|^2\omega\Omega c^2 nN. \tag{19}$$

By dividing expression (16) by t and differentiating the obtained equality with respect to t, we obtain the equation

$$t\ddot{\gamma} - \dot{\gamma} + \xi^2 t^3\gamma + \xi^2 t^3\alpha = 0. \tag{20}$$

The general solution of this equation has the form

$$\gamma(t) = -\alpha + A\sin\frac{\xi t^2}{2} + B\cos\frac{\xi t^2}{2}, \tag{21}$$

as can be verified by a direct substitution of this expression into (20). Because $\gamma(0) = 0$ and $\dot{\gamma}(t)/t|_{t=0} = 0$ at the moment of the signal arrival, as one can see, for example, from (16), then $A = 0$ and $B = \alpha$. Therefore, the response of the system to the input signal is described by the relation

$$\gamma(t) = -\alpha\left(1 - \cos\frac{\xi t^2}{2}\right). \tag{22}$$

Therefore, after the arrival of the external signal, i.e., the field u_i, a field rises in the output mode for the time $\tau = \sqrt{2\pi/\xi}$, which is approximately equal to the pump field, which is far greater than the signal field. This characteristic response time τ of the detector should be smaller than or equal to the phase relaxation time τ_0. The quantity ξ^2 can be written in the form

$$\xi^2 = \left(\frac{e^2}{\hbar c}\right)^2\frac{(2\pi c)^4|r_o|^2|r_i|^2}{V_i\lambda_i V_o\lambda_o}Nn, \tag{23}$$

where n is the number of atoms interacting with the modes, N is the number of photons in the signal mode, V_i and V_o are the volumes of the signal and output modes, and λ_i and λ_o are the wavelengths of the signal and output radiation. For $\tau \approx \tau_0$, this expression determines the minimum number of photons that should be present in the signal mode for the field amplitude in the output mode to achieve its maximum value during a detection time on the order of τ_0. Therefore, this parameter determines the sensitivity of the scheme to the value of a signal being detected. Below, we assume that the phase relaxation time is $\tau \approx 10^{-8}$ s. Such a transverse relaxation time can probably be obtained by cooling the active medium to liquid nitrogen temperature.

By assuming $\xi^2 = (2\pi)^2/\tau_0^4$, we find that, for the number of photons in the signal mode equal to

$$N = \frac{V_o\lambda_o V_i\lambda_i}{(2\pi)^2\varepsilon^2 c^4|r_{12}|^2|r_{01}|^2 n\tau_0^4}, \tag{24}$$

where ε is the fine structure constant, the field in the output mode increases for the time τ_0 up to its maximum value. For estimates, we assume that the transverse size of the mode is on the order of the wavelength λ and its longitudinal size is on the order of $10^2\lambda$. The sizes of both modes are assumed to be approximately identical. As mentioned above, the time $\tau_0 \approx 10^{-8}$ s. The concentration of active atoms $n_0 \approx 5 \times 10^{19}$ atom/cm^3 and their number is $n = n_0 V$. The matrix elements are $|r|^2 \approx 2 \times 10^{-19}$ cm^2 [6]. Then, according to (21), when the number of photons in the signal mode is $N \approx 10^{-8}$, the field in the output mode will increase to its maximum value for the time $\tau_0 = 10^{-8}$ s. As pointed out above, the maximum value of the output-mode field is

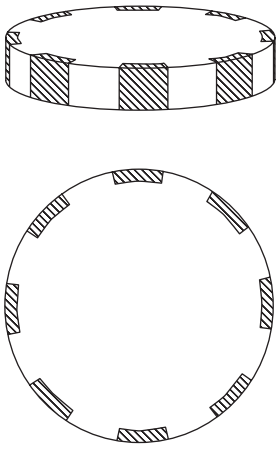

Fig. 2. Cylindrical cavity (eight sectors). Active ions are located in hatched regions.

approximately equal to the pump-mode field. If the pump mode contains, for example, 100 photons, the energy gain of a signal can be very high. Of course, to obtain such a high sensitivity, serious experimental efforts are required and the quantum nature of the detected signal should be theoretically considered. We considered above only the case of signals with a duration shorter than the phase relaxation time. However, our additional study showed that the phase relaxation does not drastically change the detection process, resulting only in an increase in the response time of the detector.

Our estimate showed that the number of parasitic photons in the signal mode at the frequency ω appearing, for example, due to nonresonance Rayleigh scattering is at least two orders of magnitude smaller than the value of N presented above.

Note an important circumstance. It was assumed implicitly in relation (13) that all the active ions are located at the same place in the cavity. However, modes in optical cavities are spatially distributed and, in particular, their phase changes from point to point. In this case, the output and pump modes can be in phase at some points and out of phase at other points. In the case of a homogeneous spatial distribution in the cavity, the active ions, which follow the phase of the pump mode, can excite the output mode in some regions and suppress excitation at other regions, making the overall interaction between the modes extremely weak. For this reason, the distribution of active ions should not be homogeneous. Active ions should be located only in the regions of the cavity where the output and pump modes are in phase.

Let us explain this by a simple example. Consider a dielectric cylindrical cavity (Fig. 2). As the output and pump modes, consider the so-called "whispering gallery" modes, i.e., the modes slipping over a circle along the cylindrical surface of the cavity. The fields of such modes are described by high-order Bessel functions.

We assume for definiteness that the output mode is described by the distribution

$$E_{mn}(r, \varphi) = J_m(\alpha_{mn} r)\sin(m\varphi),$$

where φ is the angular coordinate of the field point. Because of the boundary condition at the cylindrical surface, the condition

$$\alpha_{mn} r_c = \rho_{mn}$$

should be fulfilled, where ρ_{mn} is the nth root of the derivative of the Bessel function $J'_m(\rho_{mn}) = 0$. It is clear that the phase of this field depends on the angular coordinate. We also assume that the pump mode is described by a similar distribution with other values of m and n. The resonance frequencies of these modes can coincide for some selected values of m and n.

Assuming for definiteness that the cavity diameter is 0.01 cm (100 μm) and its height is 0.001 cm (10 μm), we can readily show that the output mode with indices $m = 314$, $n = 1$ and the pump mode with indices $m' = 325$ and $n' = 0$ are degenerate; i.e., they have the same resonance frequencies (the corresponding wavelengths are on the order of 1 μm).

In this case, the cavity volume adjacent to its cylindrical surface will be divided into $\Delta m = m' - m$ (11 = 325 – 314) sectors (Fig. 2). The modes in one-half of each sector will be in phase, and out of phase in the other half of the sector. The active ions should be distributed only where the output and pump modes are in phase, i.e., as shown in Fig. 2. It should be also taken into account that the output mode changes its sign depending on the radius, so that the depth of the distribution of ions should not exceed the value

$$\Delta = r_c - \frac{\rho_{m0}}{\alpha_{m1}} = r_c\left(1 - \frac{\rho_{m0}}{\rho_{m1}}\right).$$

For the parameters of the cavity specified above, this value is approximately 1.7 μm; at a larger depth, the output mode changes its sign and becomes out of phase with the pump mode. The distribution of active ions shown in Fig. 2 allows efficient interaction between the modes. Of course, this is only one possible variant of the cavity for laser detection of weak optical signals.

In conclusion, we summarize the results. It is shown in the paper that optical signals can be detected by laser means. Our estimates showed that such schemes have a high sensitivity.

The most striking feature of laser detection schemes considered above is that they are capable of detecting energy representing only small portions of an incident photon. If this feature is realized experimentally, the

situation in the quantum theory of measurements can change drastically.

The important feature of laser detectors of weak optical signals is that they are capable of operating without photocounts because the electronic excitation is distributed in them over many active atoms, in which electronic systems are well stabilized by a strong Coulomb field of nuclei. This means that the shot noise in laser detectors is strongly suppressed. The physical nature of noise in laser detectors substantially differs from that in conventional photodetectors and requires further study.

ACKNOWLEDGMENTS

The author thanks A.M. Prokhorov, F.V. Bunkin, E.M. Dianov, Yu.V. Kopaev, V.I. Konov, Zh. Loshak, M.V. Fedorov, A.N. Oraevsky, V.S. Zuev, B.A. Veklenko, M.A. Ananyan, V.V. Savranskiĭ, *et al*. for discussions of the results. This work was supported by the Russian Foundation for Basic Research (project no. 98-02-16671).

REFERENCES

1. N. Bloembergen, Phys. Rev. Lett. **2**, 84 (1959).
2. V. P. Bykov and V. K. Dubrovich, Kratk. Soobshch. Fiz., No. 9, 11 (1989).
3. V. P. Bykov and A. V. Gerasimov, Dokl. Akad. Nauk **38**, 50 (1993) [Phys.–Dokl. **38**, 35 (1993)].
4. V. P. Bykov, A. V. Gerasimov, and V. O. Turin, Usp. Fiz. Nauk **165**, 955 (1995) [Phys.–Usp. **38**, 911 (1995)].
5. V. P. Bykov, A. V. Gerasimov, and V. O. Turin, Ann. Fond. Louis Broglie **20**, 331 (1995).
6. G. M. Zverev and Yu. D. Golyaev, *Lasers on Crystals and Their Applications* (Radio i Svyaz', Moscow, 1994).

Translated by M. Sapozhnikov

Journal of Experimental and Theoretical Physics, Vol. 96, No. 3, 2003, pp. 430–435.
From Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 123, No. 3, 2003, pp. 487–492.

ATOMS, SPECTRA, RADIATION

Weibel Instability in Plasma Produced by a Superintense Femtosecond Laser Pulse¶

V. P. Krainov

Moscow Physicotechnical Institute, Dolgoprudnyĭ, Moscow oblast, 141700 Russia
e-mail: krainov@online.ru

Received October 24, 2002

Abstract—The Weibel instability increment is analytically derived for plasma produced at the barrier-suppression ionization of atoms and atomic ions by a superintense femtosecond laser pulse. The cases of linear and circular polarization are considered. Relativistic effects are discussed. It is found that the instability increment is larger for the circular polarization than for the linear polarization. This increment can attain the plasma frequency. Barrier-suppression ionization decreases the increment compared with the case of tunneling ionization. Relativistic effects also decrease the value of the increment. Estimates of the produced maximum quasistatic magnetic field are given. *© 2003 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

Erich Weibel [1] (see also textbook [2]) was the first to predict spontaneously growing transverse electromagnetic waves in plasma due to an anisotropic velocity distribution of electrons. The maximum increment of this instability for the wave frequency ω is (in the nonrelativistic approximation)

$$\mathrm{Im}\omega = \frac{u}{c}\omega_p,$$

where ω_p is the plasma frequency and u is the average velocity of electrons in the (longitudinal or transverse) direction along which this velocity has a maximum. This solution is valid under the condition of a strong anisotropy of the velocity distribution in longitudinal and transverse directions.

This approach has been applied in [3] to electrons produced in the tunneling ionization of atoms by a strong low-frequency linearly polarized laser field. The corresponding average velocities of electrons along the field strength polarization $u_\parallel$ and in the transverse plane $u_\perp$ strongly differ from each other. Their ratio is found in [4],

$$\frac{u_\perp}{u_\parallel} = \frac{\gamma}{\sqrt{3}},$$

where

$$\gamma = \frac{\omega_0\sqrt{2E_i}}{F}$$

is the Keldysh parameter (the atomic system of units with $e = m = \hbar = 1$ is used in this paper). Here, F is the field strength amplitude and ω_0 is the laser frequency. The quantity $E_i \gg \omega_0$ is the ionization potential of the atom (or atomic ion). In the case of tunneling ionization, we have $\gamma \ll 1$. It was found in [3] that the maximum instability increment is

$$\mathrm{Im}\omega = \frac{u_\parallel}{c}\omega_p,$$

where, according to [4],

$$u_\parallel = \frac{\sqrt{3F^3/2}}{\omega_0(2E_i)^{3/4}}.$$

In this paper, we consider the barrier-suppression ionization that occurs upon irradiation of atoms and atomic ions by the field of a superintense laser pulse with a peak intensity larger than 10^{16} W/cm^2. The corresponding anisotropic distribution of ejected electrons was obtained in [5]. We can neglect the collisions of strongly heated ejected electrons with analogous electrons and atomic ions (bearing in mind, e.g., the cluster plasma [6]) because the Weibel instability develops over a very short time, on the order of ω_p^{-1}. This process occurs at the peak of the superintense femtosecond laser pulse. We solve the problem in the linear regime only, when the perturbation of the velocity distribution function is smaller than the unperturbed distribution function. We find that the real part of the frequency of the Weibel electromagnetic field is much smaller than the laser frequency. We can therefore consider the Vlasov–Maxwell equations for the Weibel field independent of the Maxwell equation for the external laser field.

¶This article was submitted by the author in English.

2. LINEARLY POLARIZED FIELD

We first assume that the external laser radiation pulse is linearly polarized. With only linear terms of perturbation retained, the Boltzmann transport equation for the Weibel electromagnetic field is of the standard form:

$$\frac{\partial f}{\partial t} + \mathbf{v}\frac{\partial f}{\partial \mathbf{r}} = -\left(\mathbf{E} + \frac{1}{c}\mathbf{v}\times\mathbf{B}\right)\frac{\partial f_0}{\partial \mathbf{v}},$$

where $f_0(\mathbf{v})$ is a nonisotropic stationary distribution of electrons, f is a perturbation of the distribution function, and $\mathbf{E}$ and $\mathbf{B}$ represent a perturbation of the electromagnetic field (i.e., Weibel field).

Assuming that the first-order quantities $f(\mathbf{v}, \mathbf{r}, t)$, $\mathbf{E}(\mathbf{r}, t)$, and $\mathbf{B}(\mathbf{r}, t)$ depend on $\mathbf{r}$ and t only via the factor $\exp(i\omega t + i\mathbf{k}\cdot\mathbf{r})$, we obtain for the Weibel field with a frequency of ω and wave vector $\mathbf{k}$ that

$$i(\omega + \mathbf{k}\cdot\mathbf{v})f = -E\frac{\partial f_0}{\partial \mathbf{v}} + \frac{1}{c}\mathbf{v}\times\frac{\partial f_0}{\partial \mathbf{v}}\cdot\mathbf{B}. \quad (1)$$

The Maxwell equation

$$\mathrm{rot}\mathbf{E} = -\frac{1}{c}\frac{\partial \mathbf{B}}{\partial t}$$

implies a relationship between the electric and magnetic fields,

$$\mathbf{B} = -\frac{c}{\omega}\mathbf{k}\times\mathbf{E}.$$

Substituting this equation in Eq. (1), we find the equation containing only the electric field,

$$i(\omega + \mathbf{k}\cdot\mathbf{v})f = -\mathbf{E}\frac{\partial f_0}{\partial \mathbf{v}} - \frac{1}{\omega}\left[\mathbf{v}\times\frac{\partial f_0}{\partial \mathbf{v}}\right]\cdot[\mathbf{k}\times\mathbf{E}]. \quad (2)$$

We now assume that wave vector $\mathbf{k}$ is directed along the x axis and the electric field strength $\mathbf{E}$ is directed along the z axis. We then find the function f from Eq. (2):

$$f = \frac{iE}{\omega + kv_x}\left\{\frac{\partial f_0}{\partial v_z} - \frac{k}{\omega}\left(v_z\frac{\partial f_0}{\partial v_x} - v_x\frac{\partial f_0}{\partial v_z}\right)\right\}. \quad (3)$$

The second Maxwell equation is given by

$$\mathrm{rot}\mathbf{B} = \frac{4\pi}{c}\mathbf{j} + \frac{1}{c}\frac{\partial \mathbf{E}}{\partial t}, \quad (4)$$

where the electric current density is determined by the distribution function,

$$\mathbf{j}(\mathbf{r}, t) = \int \mathbf{v} f(\mathbf{v}, \mathbf{r}, t) d^3\mathbf{v}.$$

Substituting this expression and Eq. (3) into Eq. (4), we obtain the Vlasov equation in the form

$$i\mathbf{k}\times\mathbf{B} = \frac{ick^2}{\omega}\mathbf{E} = \frac{4\pi}{c}\int \mathbf{v}\frac{iE}{\omega + kv_x} \times\left\{\frac{\partial f_0}{\partial v_z} - \frac{k}{\omega}\left(v_z\frac{\partial f_0}{\partial v_x} - v_x\frac{\partial f_0}{\partial v_z}\right)\right\}d\mathbf{v} + \frac{i\omega}{c}\mathbf{E}. \quad (5)$$

Projection of this equation to the z axis gives a dispersion relation between the frequency ω and the wave number k,

$$k^2c^2 - \omega^2 = 4\pi \times\int_{-\infty}^{\infty} v_z dv_x dv_y dv_z\left\{\frac{\partial f_0}{\partial v_z} - \frac{kv_z}{\omega + kv_x}\frac{\partial f_0}{\partial v_x}\right\}. \quad (6)$$

We simplify the first term in the right-hand side of this equation by taking into account the normalization condition for the unperturbed distribution function:

$$\int f_0 d\mathbf{v} = n,$$

where n is the number density of free electrons; the above equation then becomes

$$k^2c^2 - \omega^2 + \omega_p^2 = -4\pi k\int_{-\infty}^{\infty} dv_x dv_y dv_z\frac{v_z^2}{\omega + kv_x}\frac{\partial f_0}{\partial v_x}, \quad (7)$$

where we define the plasma frequency as

$$\omega_p = \sqrt{4\pi n}.$$

The inequality $\omega \gg kv_x$ is valid in the tunneling and barrier-suppression ionization regimes. It corresponds to the condition that the longitudinal electron velocity v_z is much larger than the transverse electron velocity v_x. We can then expand the denominator in Eq. (7) as

$$\frac{1}{\omega + kv_x} = \frac{1}{\omega} - \frac{kv_x}{\omega^2} + \ldots$$

Substituting this expansion in Eq. (7), we integrate by parts,

$$k^2c^2 - \omega^2 + \omega_p^2 = \frac{4\pi k}{\omega^2}\int_{-\infty}^{\infty} dv_y\int_{-\infty}^{\infty} v_z^2 dv_z\int_{-\infty}^{\infty} v_x\frac{\partial f_0}{\partial v_x}dv_x = -\frac{4\pi k^2}{\omega^2}\int_{-\infty}^{\infty} dv_y\int_{-\infty}^{\infty} v_z^2 dv_z\int_{-\infty}^{\infty} f_0 dv_x,$$

and finally obtain

$$k^2c^2 - \omega^2 + \omega_p^2 = -\omega_p^2\frac{k^2}{\omega^2}\langle v_z^2\rangle, \quad (8)$$

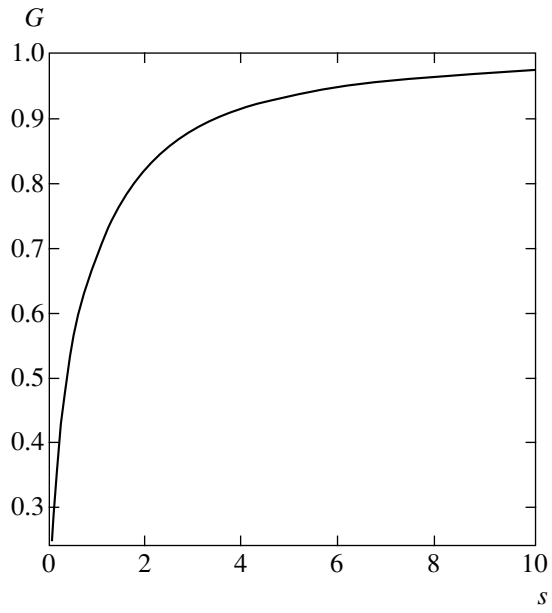

Figure.

where $\langle v_z^2 \rangle$ is the average square of the longitudinal electron velocity.

Because the most interesting case in dispersion relation (8) is $\omega \ll kc$ (see below), we can neglect the term $-\omega^2$ in the left-hand side of (8). It then follows from Eq. (8) that

$$\omega^2 = -\frac{k^2 \langle v_z^2 \rangle}{\omega_p^2 + k^2 c^2} \omega_p^2 \tag{9}$$

and, therefore, the frequency is a purely imaginary quantity that produces the Weibel plasma instability. The maximum value of this instability increment is achieved at $kc \gg \omega_p$ (the short wavelength limit),

$$\omega^2 = -\frac{\langle v_z^2 \rangle}{c^2} \omega_p^2. \tag{10}$$

In the case of tunneling ionization, the distribution function f_0 is of Gaussian form [4, 7]. Hence, $\langle v_z^2 \rangle = u_\parallel^2$, where

$$u_\parallel^2 = \frac{3\omega_0}{2\gamma^3}$$

(see the Introduction). Here, ω_0 is the laser frequency and γ is the Keldysh parameter. It therefore follows from Eq. (10) that [3]

$$\omega^2 = -\frac{u_\parallel^2}{c^2}\omega_p^2 = -\frac{3\omega_0}{2\gamma^3 c^2}\omega_p^2 = -\frac{3F^3}{2\omega_0^2 (2E_i)^{3/2} c^2}\omega_p^2.$$

We now consider the case of barrier-suppression ionization. According to [5], the anisotropic distribution is given by

$$f_0(v_z, v_\perp) \sim n \mathrm{Ai}^2 \left\{ \frac{2E_i + v_z^2 \gamma^2/3 + v_\perp^2}{(2F)^{2/3}} \right\}, \tag{11}$$

where $\mathrm{Ai}(x)$ is the Airy function. This distribution reduces to the tunneling limit [4] under the condition of weak field (compared with the atomic field),

$$F \ll (2E_i)^{3/2}.$$

Therefore, the deviation for the square of the instability increment from the tunneling limit, Eq. (10), is determined by the ratio

$$\frac{\langle v_z^2 \rangle}{u_\parallel^2} = G(s),$$

and hence,

$$\omega^2 = -G(s) \frac{3F^3}{2\omega_0^2 (2E_i)^{3/2} c^2} \omega_p^2, \tag{12}$$

where

$$s = \frac{2E_i}{(2F)^{2/3}}.$$

The universal function $G(s)$ is given by

$$G(s) = 4\sqrt{s} \frac{\int_0^\infty dt \int_0^\infty z^2 dz \mathrm{Ai}^2(s + t + z^2)}{\int_0^\infty dt \int_0^\infty dz \mathrm{Ai}^2(s + t + z^2)}. \tag{13}$$

We have $G(s) \longrightarrow 1$ at $s \gg 1$ (the tunneling limit). This function is shown in the figure.

It can be seen that in the case of barrier-suppression ionization, the increment increases more slowly with an increase in laser field F than in the case of tunneling ionization. We can therefore conclude that the electromagnetic field is generated in plasma with the same linear polarization as the initial laser radiation that produced these anisotropic plasma electrons. The frequency of this field (see Eq. (10)) does not contain the real component, and this field is therefore quasistationary, but with an exponentially growing amplitude of the electric and magnetic strengths.

We now briefly discuss the relativistic generalization of the tunneling results. According to [8], the energy distribution of ejected electrons along the field polarization is given by

$$f_0 \propto \exp\left(-\frac{v_z^2 \gamma^3}{3\omega_0} \left(1 + \frac{3v_z^2}{4\gamma^2 c^2} \right) \right).$$

The second term in the exponent is responsible for the relativistic effect. It diminishes the average longitudinal velocity,

$$u_\parallel^2(\mathrm{rel}) = \frac{u_\parallel^2}{1 + 3u_\parallel^2/(2\gamma c)^2},$$

where $u_\parallel^2 = 3\omega_0/2\gamma^3$ (see the Introduction). This decrease in the instability increment agrees with the relativistic results in [9],

$$\omega^2 = -\frac{u_\parallel^2(\mathrm{rel})}{c^2}\omega_p^2.$$

3. CIRCULARLY POLARIZED FIELD

In this section, we consider the Weibel instability produced in plasma during the tunneling and barrier-suppression ionization of atoms (or atomic ions) by a circularly polarized laser femtosecond pulse. We again direct the wave vector **k** of the laser field and of the produced electromagnetic perturbation field along the *x* axis. The perturbation electric field strength **E** is also circularly polarized and rotates in the *yz* plane. Hence,

$$\mathbf{E} = E(\mathbf{i}_z + i\mathbf{i}_y)\exp(i\omega t + i\mathbf{k}\cdot\mathbf{r}),$$

where $\mathbf{i}_z$ and $\mathbf{i}_y$ are unit basis vectors. Equation (2) then becomes

$$i(\omega + kv_x)f = -\frac{E}{v_\parallel}(v_z + iv_y)\frac{\partial f_0}{\partial v_\parallel} - \frac{kE}{\omega}(v_z + iv_y)\left(\frac{v_x}{v_\parallel}\frac{\partial f_0}{\partial v_\parallel} - \frac{\partial f_0}{\partial v_x}\right),$$

where $v_\parallel$ is the velocity in the polarization plane. We thus obtain the perturbation distribution function

$$f = iE(v_z + iv_y)(\omega + kv_x)^{-1} \times\left\{\frac{1}{v_\parallel}\left(1 + \frac{kv_x}{\omega}\right)\frac{\partial f_0}{\partial v_\parallel} - \frac{k}{\omega}\frac{\partial f_0}{\partial v_x}\right\}.$$

Instead of Eq. (6), we find the dispersion relation in the form

$$(k^2c^2 - \omega^2)(\mathbf{i}_z + i\mathbf{i}_y) = 4\pi\int_{-\infty}^{\infty}\mathbf{v}dv_x dv_y dv_z(v_z + iv_y) \times\left\{\frac{1}{v_\parallel}\frac{\partial f_0}{\partial v_\parallel} - \frac{k}{\omega + kv_x}\frac{\partial f_0}{\partial v_x}\right\},$$

or

$$(k^2c^2 - \omega^2) = 4\pi^2\int_0^\infty v_\parallel^3 dv_\parallel\int_{-\infty}^{\infty}dv_x \times\left\{\frac{1}{v_\parallel}\frac{\partial f_0}{\partial v_\parallel} - \frac{k}{\omega + kv_x}\frac{\partial f_0}{\partial v_x}\right\}. \quad (14)$$

The unperturbed electron energy distribution function for tunneling ionization is given by (see, e.g., [10])

$$f_0 = \frac{n}{4\pi^2u^2v_0}\exp\left\{-\frac{v_x^2 + (v_\parallel - v_0)^2}{2u^2}\right\}, \quad (15)$$

where $v_0 = F/\omega_0$ is the ponderomotive electron velocity and

$$u^2 = \frac{F}{2\sqrt{2E_i}}$$

(*F* is again the laser field strength amplitude and E_i is the ionization potential of an atom or an atomic ion). Unlike for the linear polarization, dispersions of the average longitudinal and transverse velocities are now equal to each other.

We note that $u \ll v_0$; i.e., the width of the distribution is small compared to its shift in the longitudinal direction. The first term in the right-hand side of Eq. (14) vanishes because the integrand is an odd function of the argument $(v_\parallel - v_0)$. Dispersion relation (14) then becomes

$$(k^2c^2 - \omega^2) = -4\pi^2\int_0^\infty v_\parallel^3 dv_\parallel\int_{-\infty}^{\infty}dv_x\frac{k}{\omega + kv_x}\frac{\partial f_0}{\partial v_x}.$$

We again assume that $\omega \gg kv_x$, i.e., $\omega \gg ku$, and expand the denominator in a Taylor series,

$$\frac{1}{\omega + kv_x} = \frac{1}{\omega} - \frac{kv_x}{\omega^2} + \ldots$$

Integrating by parts, we simplify the dispersion relation as

$$(k^2c^2 - \omega^2) = -\frac{4\pi^2k^2}{\omega^2}\int_0^\infty v_\parallel^3 dv_\parallel\int_{-\infty}^{\infty}f_0 dv_x. \quad (16)$$

Substituting Eq. (15) in Eq. (16), we obtain

$$\omega^4 - (kc\omega)^2 - \frac{(\omega_p kv_0)^2}{2} = 0.$$

The solution of this equation is

$$\omega^2 = -\frac{(kc)^2}{2}\left\{\sqrt{1 + 2\left(\frac{\omega_p v_0}{kc^2}\right)^2} - 1\right\} < 0.$$

The frequency ω is therefore a purely imaginary quantity that produces a circularly polarized exponentially increasing electromagnetic wave. Its real part is zero, and the wave is therefore quasistationary. In the short wave limit

$$k \gg \frac{\omega_p v_0}{c^2}, \tag{17}$$

we simplify this solution by taking into account that $v_0 = F/\omega_0$,

$$\omega^2 = -\frac{1}{2}\left(\frac{v_0}{c}\right)^2 \omega_p^2 = -\frac{1}{2}\left(\frac{F}{\omega_0 c}\right)^2 \omega_p^2. \tag{18}$$

The condition $\omega \gg ku$ bounds the wave number k from above,

$$k^2 \ll \frac{F\omega_p^2\sqrt{2E_i}}{\omega_0^2 c^2}. \tag{19}$$

Inequalities (17) and (18) do not contradict each other under the condition

$$F \ll c^2\sqrt{2E_i}$$

which is satisfied up to very high values of laser field intensities ($c = 137$ a.u.).

In the case of barrier-suppression ionization by a circularly polarized field, the unperturbed distribution function is given by [5]

$$f_0(v_x, v_\parallel) \sim n\mathrm{Ai}^2\left\{\frac{2E_i + (v_\parallel - v_0)^2 + v_x^2}{(2F)^{2/3}}\right\}.$$

Substituting this expression in Eq. (16), we obtain the same dispersion relation as in the case of tunneling ionization. Therefore, the maximum increment of the Weibel instability is again determined by Eq. (18) also for barrier-suppression ionization.

In the nonrelativistic limit, we have

$$v_0 = \frac{F}{\omega_0} \ll c.$$

Hence, the Weibel increment is small compared to the plasma frequency. The anisotropic relativistic distribution of ejected electrons was obtained in [10]. Most electrons are ejected not in the polarization plane of circularly polarized laser radiation, but at the angle θ with respect to this polarization plane determined from the relation [11]

$$\tan\theta = \frac{F}{2\omega_0 c}.$$

The normalized unperturbed relativistic distribution function is given by [10]

$$f_0(p_x, p_\parallel) = \frac{n}{4\pi^2 u_r^2 v_0} \times \exp\left\{-\frac{(p_x - v_0\tan\theta)^2 + (p_\parallel - v_0)^2}{2u_r^2}\right\}, \tag{20}$$

where p_x and $p_\parallel$ are the respective momentum components of the ejected electron along the wave vector and in the polarization plane. The relativistic width u_r of the distribution is given by [10]

$$u_r^2 = \frac{(1 + (F/\omega_0 c)^{1/2})^2}{1 + (F/\omega_0 c)^2} u^2,$$

where (see above)

$$u^2 = \frac{F}{2\sqrt{2E_i}}$$

is the nonrelativistic width.

Substituting Eq. (20) in Eq. (16), we find instead of Eq. (18) that

$$\omega^2 = -\frac{1}{2}\frac{\langle v_\parallel^2\rangle}{c^2}\omega_p^2,$$

where $\sqrt{\langle v_\parallel^2\rangle}$ is the average relativistic velocity in the polarization plane, to be compared with Eq. (10) for the linearly polarized field. This quantity can be expressed through the corresponding relativistic momentum,

$$p_\parallel = v_0 = F/\omega_0,$$

and the relativistic energy,

$$E_{\mathrm{rel}} = c\sqrt{p_\parallel^2 + p_x^2 + c^2},$$

due to narrow peaks in unperturbed distribution (20),

$$\frac{\sqrt{\langle v_\parallel^2\rangle}}{c} = \frac{p_\parallel c}{E_{\mathrm{rel}}} = \frac{v_0}{v_0^2/2c + c}.$$

Thus, we finally obtain

$$\omega = -\frac{i}{\sqrt{2}}\frac{v_0}{v_0^2/2c + c}\omega_p. \tag{21}$$

It follows that relativistic effects diminish the Weibel increment for a circularly polarized field similarly to the case of linear polarization (see above). The maximum value of the increment is achieved at $v_0 = F/\omega_0 = c\sqrt{2}$,

$$\omega_{\max} = -\frac{i}{2}\omega_p.$$

4. CONCLUSIONS

We have found that the plasma instability produces a quasistatic magnetic field B (the frequency does not contain the real part). The corresponding quasistatic electric field E is much smaller in the short-wave limit $kc \gg \omega$,

$$E = \frac{\omega}{kc} B \ll B.$$

We estimate the maximum value of this field for the circularly polarized field. Our derivation is valid in the linear approximation where $f \ll f_0$. In accordance with the results in the previous section, we rewrite this inequality for a circularly polarized field as

$$\frac{E}{\omega^3} k^2 v_0 \ll 1,$$

or

$$\frac{kBv_0}{\omega^2 c} \ll 1.$$

Substituting $\omega \sim \omega_p (v_0/c)$, we find the maximum magnetic field

$$B_{\max} \sim \omega_p c.$$

It follows that the magnetic field is determined only by the number density of plasma electrons and can be very large.

ACKNOWLEDGMENTS

This work was supported by the International Science and Engineering Center (project no. 2155) and by the Russian Foundation for Basic Research (project nos. 01-02-16056 and 02-02-16678).

REFERENCES

1. E. S. Weibel, Phys. Rev. Lett. **2**, 83 (1959).
2. A. F. Alexandrov, L. S. Bogdankevich, and A. A. Rukhadze, *Principles of Plasma Electrodynamics* (Springer, Berlin, 1984; Vysshaya Shkola, Moscow, 1988), p. 162.
3. V. I. Arefyev, V. P. Silin, and S. A. Uryupin, Phys. Lett. A **255**, 307 (1999).
4. N. B. Delone and V. P. Krainov, J. Opt. Soc. Am. B **8**, 1207 (1991).
5. V. P. Krainov, J. Opt. Soc. Am. B **14**, 425 (1997).
6. V. P. Krainov and M. B. Smirnov, Phys. Rep. **370**, 237 (2002).
7. N. B. Delone and V. P. Krainov, *Multiphoton Processes in Atoms* (Énergoatomizdat, Moscow, 1984; Springer, Berlin, 2000).
8. V. P. Krainov, Opt. Express **2**, 268 (1997).
9. P. A. Polyakov, Izv. Ross. Akad. Nauk, Ser. Fiz. **63**, 2374 (1999).
10. V. P. Krainov, J. Phys. B **32**, 1607 (1999).
11. H. R. Reiss, J. Opt. Soc. Am. B **7**, 574 (1990).

Journal of Experimental and Theoretical Physics, Vol. 96, No. 3, 2003, pp. 436–439.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 123, No. 3, 2003, pp. 493–497.

PLASMA, GASES

Dust Particles in a Thermophoretic Trap in a Plasma

L. M. Vasilyak*, S. P. Vetchinin, V. S. Zimnukhov, D. N. Polyakov, and V. E. Fortov

Institute for High Energy Densities, Associated Institute for High Temperatures, Russian Academy of Sciences, Moscow, 125412 Russia

**e-mail: lab852@ihed.ras.ru*

Received June 14, 2002

Abstract—The feasibility of confining dust particles in a plasma by thermophoretic forces was demonstrated. An extended dust structure in a positive glow discharge column was experimentally obtained at liquid nitrogen temperature. The dust structure was confined in an electrostatic-thermal trap, in which vertical stability was provided by the summed action of longitudinal electrostatic field and thermophoretic forces. Traps of this kind can be analyzed in terms of the general principles developed for confining particles in traps with the use of electric and magnetic multipole fields. We were able to change the shape and volume of the structure and even separate it into parts by varying temperature fields. © *2003 MAIK "Nauka/Interperiodica".*

Ordered structures comprising charged dust particles with strong Coulomb interactions have been observed in a nonequilibrium low-pressure plasma in charged electric layers, which are actually electrostatic traps [1]. Such structures are formed in strata in a glow discharge and in a charged layer close to electrodes in an RF discharge. When a charged particle is at equilibrium, its gravity force is balanced by a longitudinal electric field, which is much higher in layers than in the surrounding plasma and is longitudinally nonuniform. An ambipolar electric field confines particles in the radial direction. The number of confined particles depends on the size of the trap and can vary from several dozens to several thousands. The particles can experience ordering to produce a fairly regular crystal structure, or their behavior can resemble that of a liquid. As the charge of particles is sustained by the balance between electron and ion fluxes onto their surface, the arising structures are dissipative. Because of the high mobility of electrons in glow and RF discharges, the charge is always negative and amounts to $q = 10^4$–10^5 electron charge units.

An electrostatic trap can exist in a plasma if there are gradients of the concentration of charged particles. This in turn results in electric field nonuniformity in a layer caused by the arising of both longitudinal and radial ambipolar electric fields. The presence of heat sources in a bounded plasma causes the appearance of a temperature gradient and the arising of thermophoretic forces, which act on dust particles. According to [2, 3], the existence of various forms and structures of plasma-dust clouds confined in electrostatic traps primarily depends on equilibrium conditions in the radial direction, because the characteristic depth of the well that confines particles in the radial direction is much smaller. The radial ambipolar field that acts on charged dust particles is directed toward the axis of the discharge tube. Forces that push dust particles toward walls are ionic friction and thermophoretic forces. The thermophoretic force is proportional to the gradient of gas temperature and is directed similarly to the heat flux caused by thermal conductivity of the neutral gas. In most of the experiments in which dust particles were confined in charged layers, ionic friction forces were substantially weaker than thermophoretic forces. All the forces are therefore proportional to the corresponding gradients and, as shown in [2], the concept of potential energy can then be introduced for particles confined in an electrostatic–thermal trap. Generally, these forces also determine the equilibrium of dust particles in a nonuniform thermal field in the longitudinal direction.

The principle of confining charged particles in natural traps of the type of strata in a strongly nonuniform plasma is similar to the general principles of confining particles in special traps using electric or magnetic multipole fields. These principles have long been successfully used in the physics of molecular beams, in mass spectrometry, and in the physics of accelerators of charged particles [4]. Particles can be trapped or phased under the action of an elastic binding force that linearly increases as the distance grows,

$$F = -cr,$$

or, in other words, if particles move in the parabolic potential

$$\varphi \sim \alpha x^2 + \beta y^2 + \gamma z^2.$$

It follows from the Laplace equation

$$\Delta\varphi = 0$$

that

$$\alpha + \beta + \gamma = 0$$

for potential fields. For instance, at $\alpha = \beta = 1$ and $\gamma = -2$ in the cylindrical coordinates, this condition leads to a three-dimensional trap with the potential [4]

$$\varphi = \varphi_0 \frac{r^2 - 2z^2}{r_0^2 + 2z_0^2},$$

where $2z_0^2 = r_0^2$.

Note that precisely this potential configuration was used in the first experiments on confining charged micron aluminum particles in a quadrupole trap in 1959 [5]. In addition, dynamic stabilization of the orbits of separate particles with the use of a low-frequency (50 Hz) periodic electric field was used. Random motion of particles was suppressed by a buffer gas. It was found that the particles lined up into a regular structure to form a crystal. The pseudocrystallization of ions cooled by a laser beam was observed in a similar trap [6]. The ions moved to the positions where Coulomb repulsion forces were balanced by focusing forces in the trap and the energy of the whole ensemble was minimum.

For plasma-dust structures confined in a volume charge layer, for instance, in a near-electrode RF discharge layer, in which dust structures of particles were observed for the first time [7], the condition

$$\alpha + \beta + \gamma = 0$$

does not hold, because the Poisson equation applies to charged layers. At the same time, the Laplace equation is valid for thermal fields without heat sources, and the general principles can therefore be used for creating traps.

In a gas-discharge low-pressure plasma, thermophoretic forces strongly influence the motion of dust particles and their equilibrium arrangement in space [2, 3]. Thermophoretic forces can be used not only to change the position and shape of a dust cloud but also to remove dust particles from the discharge zone under the action of the temperature gradient in a neutral gas. At a low gas pressure, the thermophoretic force is

$$F_T = -4PLa^2 \nabla T/T,$$

where P is the neutral gas pressure, L is the free path of molecules, a is the radius of dust particles, and $L \gg a$. In a glow discharge plasma, the force of the radial ambipolar electric field that confines charged dust particles is

$$F_E = qkT_e \nabla n/n,$$

where T_e is the electron temperature and n is the concentration of electrons. In the ionization-diffusion model, the distribution of electrons in the positive column and in a stratum is described by the Bessel function $J_0(2.4r/R)$, where R is the radius of the tube. For this reason, the F_E force is proportional to the radius in the vicinity of the axis and grows more sharply as the particle approaches tube walls. If heat release in a plasma is axially symmetric, the thermophoretic force close to the axis is also proportional to the radius.

A dust of particles cannot be confined in the longitudinal direction in a uniform positive column, because electric field gradients are then absent, and particles, which are in the neutral equilibrium state, depart along the axis of the discharge tube. In addition, an analysis of the experimental data on glow discharge in various gases [8] shows that the strength of the longitudinal electric field is insufficient for balancing the gravity force of micron particles; this is why dust structures are observed in strata or near-electrode layers. Creating a trap in the longitudinal direction requires additional gradient forces. An electrostatic thermal trap in a plasma with a longitudinal temperature gradient in a neutral gas can be formed with the use of thermophoretic forces. Cooling or heating part of the discharge causes the appearance of additional longitudinal forces capable not only of balancing gravity forces but also of locking particles in a trap. Such forces arise if gas temperature along the symmetry axis changes by the quadratic or a sharper law. Estimates show that, for micron particles of mass 10^{-10} g in a glow discharge in air, longitudinal temperature gradients on the order of 10 K/cm are required.

The confinement of dust particles in an electrostatic-thermal trap was experimentally observed in a glow discharge plasma in air when the air was cooled to the boiling temperature of liquid nitrogen (77 K). The scheme of experiments was similar to that used in [9]. Dust structures of magnesium oxide particles 3–5 μm in diameter, which were introduced from above from the side of the anode, were observed in the positive column of a low-pressure glow discharge in a quartz discharge tube 2 cm in diameter immersed in a cryostat with liquid nitrogen. The dust structures were studied through a longitudinal optical window of the cryostat. An extended structure up to 20 cm long arose in the narrow pressure range 0.04–0.05 Torr; this structure consisted of long dust threads, which occupied almost the whole volume of the glow discharge positive column. A photograph of a dust structure 12 cm long in a positive glow discharge column in liquid nitrogen is shown in Fig. 1. Above the level of liquid nitrogen, the character of the discharge changes and the discharge column is divided into strata. Alongside the factors well known previously (the form of the electrostatic potential, gravity force, and ionic friction force), the key role in determining the stability of extended dust structures is played by longitudinal thermophoretic forces. These forces are proportional to temperature gradients and are almost an order of magnitude larger than the ionic friction force. They arise because the momentum transferred from molecules that arrive from the side of the

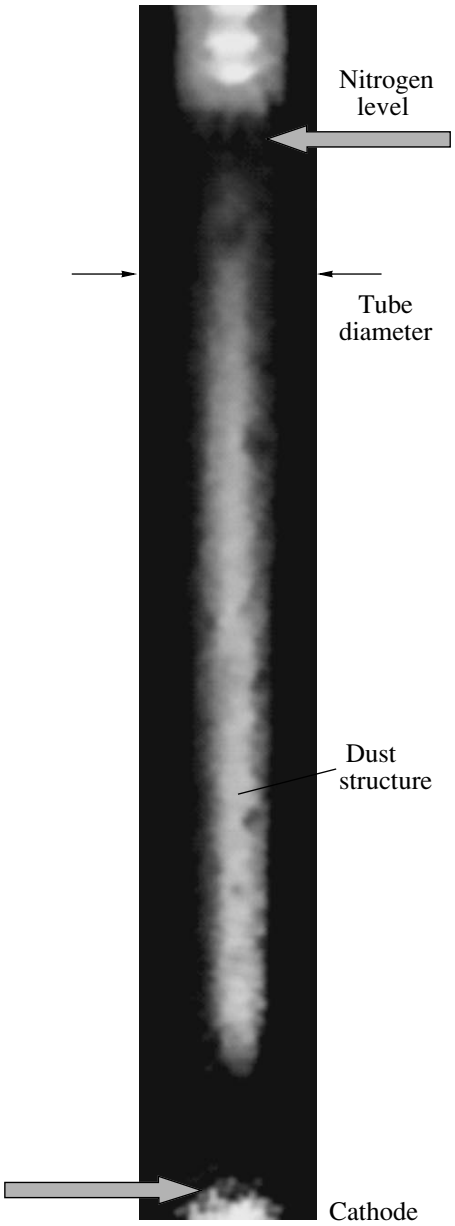

Fig. 1. Extended dust structure in a positive glow discharge column at cryogenic temperature. Seen at the bottom is cathode glow, and at the top, strata formed above the boundary between air and liquid nitrogen. The liquid nitrogen level and cathode glow are marked by arrows.

more heated gas to a dust particle is larger than the summed momentum of "cold" molecules. The dust structure occurs in an electrostatic-thermal trap, in which vertical stability is provided by the summed action of longitudinal electric field and thermophoretic forces. Thermophoretic forces lock the particles on two sides, namely, on the sides of the cathode and the boundary of liquid nitrogen. They act both downward, in the direction of the heat flux from above (from the zone free of liquid nitrogen), and upward, because of heat release in the cathode zone. Earlier, such

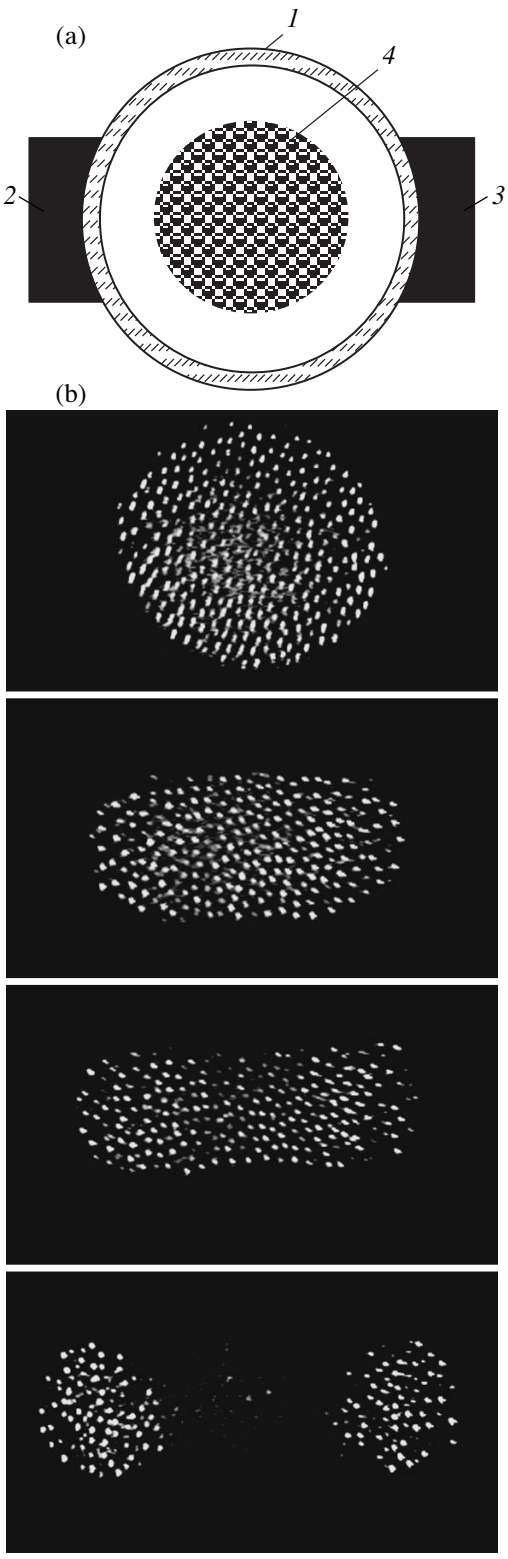

Fig. 2. Dynamics of dust structure partitioning under the conditions of cooling the discharge tube on two sides: (a) scheme of the arrangement of microrefrigerators; the discharge tube cross section is shown: (*1*) discharge tube, (*2, 3*) microrefrigerators, and (*4*) dust structure; and (b) sequential photographs of the dust structure in discharge tube cross section after switching on cooling; the diameter of the tube equals the horizontal frame size.

extended dust structures had never been observed experimentally.

The dust structure becomes separated into several parts 4–5 cm long each as pressure decreases. The resulting structure is, however, unstable with respect to longitudinal low-frequency perturbations. The whole extended dust structure divides into transverse layers about 0.7 mm thick, which have different densities. Most likely, this is caused by the arising of dust-acoustic instability. The appearance of dust-acoustic instability is a consequence of general plasma-flow instability, when ions move through a charged dust at drift velocities exceeding the velocity of sound in the dust itself. The development of such an instability is facilitated by a decrease in the temperature of heavy particles, because the velocity of sound in the dust then decreases [9].

We performed experiments in which a discharge tube 2 cm in diameter was cooled with the use of Peltier elements. The middle part of the discharge tube was cooled on both sides by two microrefrigerators 2.5 cm long, which tightly fitted the side surface of the tube. Two gaps 8 mm wide were situated between the microrefrigerators for observing dust structures. The dust structures were concentrated in strata close to the tube axis at air pressures of 0.1–0.5 Torr and currents of 0.25–1 mA. Cooling the discharge tube walls by 20 K stretched the dust structure of the stratum between the microrefrigerators in the radial direction under the action of thermophoretic forces. First, the dust structure cross section transformed into an ellipse prolate toward cold walls. Next, the cloud of dust particles divided into two circles, which experienced deformations and were attracted to cold walls (Fig. 2). Particles did not reach the walls under the action of the radial temperature gradient but were confined in a new equilibrium state, because the radial electric force directed toward the axis grew stronger as the radius increased. Dust clouds of various shapes can be created by varying the temperature field. If the walls of a discharge tube are locally cooled by microrefrigerators in the longitudinal direction, all charged dust particles are withdrawn by the longitudinal temperature gradient from the strata lying lower by 5–6 cm. They in part go to cold tube walls in the region of refrigerators. Particles remain in the strata that are situated above the refrigerators. Longitudinal thermophoretic forces are strong enough to draw dust particles upward from electrostatic traps, that is, from strata. As a consequence, a new trap arises; this trap is formed as a result of the superposition of thermophoretic and longitudinal electric field forces. Creating various thermal traps with the use of Peltier elements allows the shape of dust structures to be changed or even allows dust structures to be removed from the discharge zone.

To summarize, thermophoretic forces can be used to create thermal traps in a plasma for confining dust particles. At low temperatures, extended dust structures can be formed; such structures cannot be obtained in natural electrostatic traps in a discharge. Dust structures can change their shape and volume and even become divided into parts when Peltier elements are used. As thermophoretic forces are potential, the construction of traps can be performed with the use of the well developed theory of creating electromagnetic traps.

REFERENCES

1. V. E. Fortov *et al.*, Phys. Plasmas **7**, 1374 (2000).
2. V. V. Balabanov, L. M. Vasilyak, S. P. Vetchinin, *et al.*, Zh. Éksp. Teor. Fiz. **119**, 99 (2001) [JETP **92**, 86 (2001)].
3. G. M. Jellum, J. E. Daugherty, and D. B. Graves, J. Appl. Phys. **69**, 6923 (1991).
4. W. Paul, Usp. Fiz. Nauk **160**, 109 (1990).
5. R. F. Wuerker and R. V. Lengmuir, Appl. Phys. **30**, 342 (1959).
6. W. Neuhauser, M. N. Hohenstett, P. Toschek, and H. Dehmelt, Phys. Rev. A **22**, 1137 (1980).
7. V. N. Tsytovich, Usp. Fiz. Nauk **167**, 57 (1997) [Phys.–Usp. **40**, 53 (1997)].
8. L. M. Vasilyak, S. P. Vetchinin, A. P. Nefedov, and D. N. Polyakov, Teplofiz. Vys. Temp. **38**, 701 (2000).
9. V. E. Fortov, L. M. Vasilyak, S. P. Vetchinin, *et al.*, Dokl. Akad. Nauk **382**, 50 (2002) [Dokl. Phys. **47**, 21 (2002)].

Translated by V. Sipachev

Journal of Experimental and Theoretical Physics, Vol. 96, No. 3, 2003, pp. 440–443.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 123, No. 3, 2003, pp. 498–502.

PLASMA, GASES

The Action of an Electron Beam on Dust Structures in a Plasma

L. M. Vasilyak[a,*], M. N. Vasil'ev[b], S. P. Vetchinin[a], D. N. Polyakov[a], and V. E. Fortov[a]

[a] *Institute for High Energy Densities, Associated Institute for High Temperatures, Russian Academy of Sciences, Moscow, 125412 Russia*

[b] *Moscow Institute of Physics and Technology (Technical University), Dolgoprudnyĭ, Moscow oblast, 141700 Russia*

**e-mail: lab852@ihed.ras.ru*

Received June 14, 2002

Abstract—The action of an electron beam on ordered dust structures in glow and low-pressure RF discharges was studied experimentally. The electron beam produces destruction and dynamic displacement of the dust structure. In the center of a dust structure, an electron beam with a low electron energy (tens of eV) at currents up to 1 mA caused structural disordering and "melting" in the region of its action but did not excite external crystal regions. Local action of an electron beam with a high electron energy (25 keV) and a beam current above 10 mA caused deformation of the whole dust structure and shifted it in the horizontal direction so that it was carried away from the RF discharge zone. The effect of dust structure displacements can be used to locally remove particles from a plasma. © *2003 MAIK "Nauka/Interperiodica".*

A dust plasma in a high-frequency or glow discharge can be used to obtain high nonideality parameters

$$\gamma = Z^2 e^2 n^{1/3}/kT \sim 10^4–10^5$$

because of a large charge of dust particles $Z \sim 10^5$–10^6. Interparticle Coulomb interactions are then strong to the extent that they cause the formation of ordered dust structures, in which Coulomb interaction exceeds thermal movement energy by orders of magnitude. The reaction of an ordered structure to actions followed by its relaxation to the initial state allows new data on the properties of ordered structures to be obtained [1]. The relaxation of a dust structure after an action depends on the character of the action. Under slow weak actions, for instance, thermal [2] or electric [3] action, the structure moves in space, and some part of it can be deformed or destroyed. The action of high-power electric field pulses of nanosecond width causes the destruction of dust structures in a plasma [1]. Under such short-term actions, dust particles have no time to move and structural transformations occur as a result of rapid changes in plasma parameters. It is likely that the destruction of dust structures and the recession of particles under nanosecond actions are primarily caused by an increase in their charge, because a fair number of high-energy electrons with energies of hundreds of eV then appear in the plasma [4]. An electron beam can directly act on dust particles to increase their charge and transfer momentum to them. In [5, 6], charging of dust particles in a non-self-sustained discharge, which was sustained by an electron beam at atmospheric pressure, for creating a radioactive battery was analyzed theoretically. The purpose of this work was to study the direct action of an electron beam on dust particles in a low-pressure discharge.

Experiments were conducted in air at 0.01–0.20 Torr in cylindrical discharge tubes 2 cm in diameter under glow discharge conditions and in a large-volume chamber (Fig. 1) with plane electrodes 10 cm in diameter under high-frequency (13.56 MHz) discharge conditions. Structures of 5–20 μm MgO particles were studied by means of a video camera at transverse and longitudinal illumination by a laser knife (0.63 μm wavelength radiation) with a 150 μm caustic. A weak-current electron beam with 25 keV of electron energy was used to act on dust structures in an RF discharge plasma. The beam was formed in an accelerator and emitted horizontally into the discharge chamber, in which an ordered dust structure was preliminarily created. The diameter of the beam in the chamber was 2–2.5 mm. At pressures characteristic of our experiments, the beam introduced into the working chamber remained virtually unscattered in the gas medium. The beam current could be continuous or have the shape of rectangular pulses of 2–200 Hz. The ordered dust structure was confined in an RF discharge between two plane electrodes situated 5 cm apart. The pressure in the working chamber was 0.20 Torr. The structure was formed in the volume charge zone above the lower electrode, to which RF power was supplied. The upper electrode was grounded. To radially confine macroparticles in the structure, a ring 5 mm high made of organic glass

was used. Ring walls ran along the electrode perimeter and lay on the lower electrode. The ordered dust structure hung over the ring and consisted of several layers of charged particles. The electron beam fell on the structure from one side in the radial direction. The character of its action depended on the beam current and distance d between the axis of electron beam injection and the plane of the ordered plasma-dust structure. At $d > 10$ mm, the structure did not react to switching on the beam if the beam current was lower than 10 mA. Structural perturbations observed as the beam approached the plane of the plasma-dust structure were as follows. Switching on a beam with an $I < 5$ mA beam current set the plasma-dust structure in motion, which was accompanied by changes in interparticle distances and displacements of the external boundary of the structure. Switching the beam off caused the return of the structure to the initial state. When a low-frequency (some 2–5 Hz) pulsed periodic beam ($I < 5$ mA) was injected, plasma-dust structure particles began to oscillate. An increase in the frequency of beam current modulation accelerated particle vibrational motions, which then ceased to be discernible. Beams with beam currents above 10 mA slightly deformed the whole dust structure, displaced it in the horizontal direction, and carried it away from the discharge zone. Figure 2 shows a sequence of photographs illustrating displacement of a dust closed by a 25-keV electron beam (exposure, 0.04 s; time interval between shots, 0.16 s). The dust structure was not only carried away from the zone of beam action determined by the initial diameter of the beam, but it behaved similarly to an elastic body, which caved in under the beam action and then moved as a whole. The structure was not restored after switching off the beam. The destruction of the structure was primarily caused by the dynamic action of beam electrons on dust particles.

Our estimate shows that the force transferred to a separate particle from an electron beam at beam currents of about 10 mA exceeds the radial electric field force that confines the particle in a layer. This estimate is based on the assumption that electrons are absorbed by the particle because the free path of beam electrons in the material of particles is smaller than the size of particles and that they directly transfer their momentum to the particle. A beam acts on the surface of a dust particle and causes its acceleration. Note that the force estimated from the transfer of momentum from electrons coincides with the value estimated from the characteristic acceleration of dust particles as determined from our video records. As particles are not forced from the structure by the beam, the dynamics of beam actions is, in all probability, as follows. After switching on the beam, dust particles acquire a large additional charge, because the energy of beam electrons and the energy of electrons knocked out from the surface of particles (~100 eV) [7] substantially exceed the energy of plasma electrons. This was experimentally observed

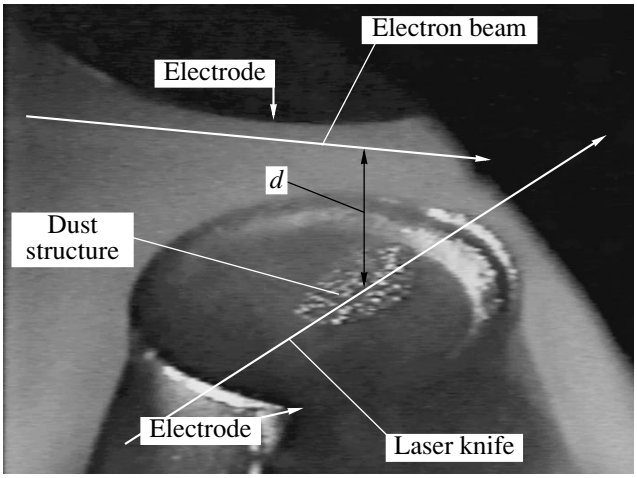

Fig. 1. Scheme of experiments and working cell. The diameter of the lower electrode is 10 cm, the angle between the electron beam direction and the laser beam is 90°, d is the distance between the dust structure plane and the electron beam. The video camera is situated at a 45° angle with respect to the horizontal plane of the electrodes.

under the conditions of electron beam action on separate particles in a low-pressure gas [8]. Particles flew away from the zone of beam action irrespective of the material of particles and their emission properties. The measured potential of macroparticles was about 100 V [8]. This caused intense scattering of beam electrons and secondary electrons by dust particles, because the cross section of Coulomb scattering increased proportionally to the squared charge of particles and became commensurate with the interparticle distance. The scattering of the beam and secondary electrons could cause charging of peripheral particles, which increases the effective area of beam action on the dust cloud. As particles in the cloud interact with each other, the dust cloud moves as a whole. To check this suggestion, we performed the following model experiment. A metallic ball 8 mm in diameter was placed in the way of the beam. When the beam acted on the ball, gas glow around the ball was observed, which was evidence of scattering of beam electrons and electrons emitted from the surface over a large area.

Dust structures form in strata when a glow discharge is ignited in a cylindrical discharge tube. The conditions for the formation of ordered structures of micron particles in air in a glow discharge at a current of 0.3–1 mA are approximately the same as with RF discharges. An increase in the glow discharge current at 0.2 Torr is accompanied by an increase in the size of the dust structure and in the distance between dust particles. The crystal experiences disordering and begins to resemble a liquid as the current grows to several mA. At a discharge current of 0.6–1 mA, ordered structures in the form of a cylinder, whose diameter equals one-third or one-fourth of the tube diameter, are observed. The

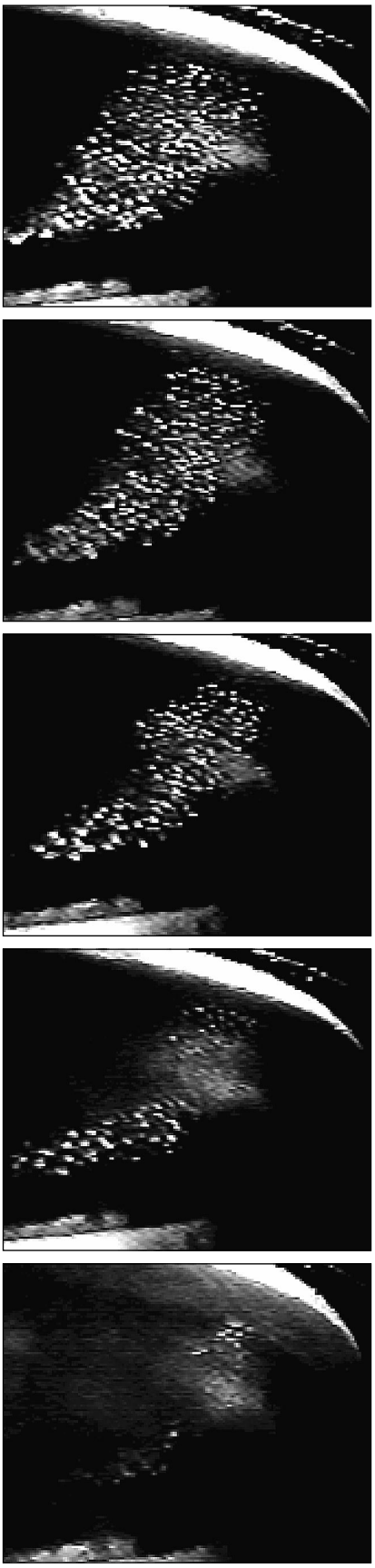

Fig. 2. Carrying away of a dust cloud.

diameter of the ordered structure grows as current increases and reaches two-thirds of the tube diameter at currents of 1–2 mA. The thickness of the structure decreases as the current grows, and, at 1.5–2 mA, plane structures comprising several (5–10) layers of particles are observed. At still stronger currents, a ring structure is formed; its diameter increases and thickness decreases as current grows. In the near-axis region, particles are absent. The currents at which transitions between different ordered structural forms are observed decrease as the size of dust particles increases.

A plane cathode was replaced by a hollow cathode to obtain an electron beam in the specified pressure range in a glow discharge. A cylindrical ordered structure was then obtained at the strata at a current of 0.3–0.6 mA. An increase in discharge tube voltage caused the formation of a narrow electron beam from the hollow cathode with a characteristic electron energy of several tens of eV. An increase in the total discharge current was largely determined by an increase in the electron beam current. The size of the dust structure and the shapes of glowing stratum regions remained unchanged as the total current increased. When the electron beam, which was formed in the near-axis region at a current of 0.6 mA, reached the center of the dust crystal, the beam acted on it and caused crystal "melting" and disordering only in the region of beam action, without exciting external crystal regions (Fig. 3). The diameter and the depth of beam action increased as the beam current grew to cause the excitation of the whole crystal at a current of 2 mA. The general picture of the action of an electron beam in a glow discharge with a hollow cathode corresponded to RF discharge experiments in which a high-energy electron beam with a weak electron beam current was used. The destruction of structures in both RF and glow discharges was primarily caused by an increase in the charge of dust particles. An electron beam rapidly increased the charge of dust particles, which was proportional to the energy of electrons. The neutralization of excess negative charge occurred at a much lower rate, because this process was determined by the flow of ions. The action of excess charge caused Coulomb repulsion of particles and pushed away particles from the plasma trap upward. Subsequently, the particles fell down, which resulted in structural disordering. Particles experienced various oscillations under the action of an electron beam. These oscillations could also be related to the excitation of plasma–beam instabilities of various kinds in the dust medium.

To summarize, the action of an electron beam on dust structures in a plasma is, in essence, as follows: it increases the charge of dust particles and causes momentum transfer from beam electrons in collisions. As a result, both the destruction and dynamic displacements of dust structures occur. Estimates show that the force transferred to a separate macroparticle from an electron beam and the force calculated from the charac-

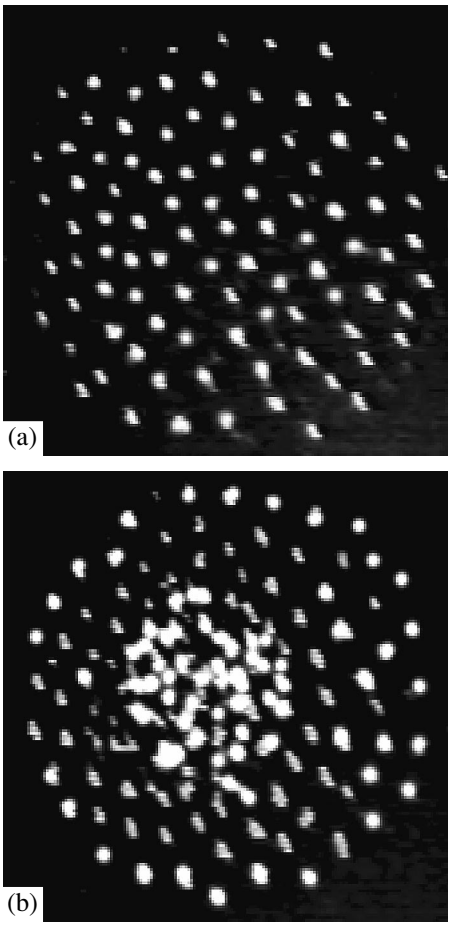

Fig. 3. Electron beam action on the central region of an ordered dust structure in a glow discharge from a hollow cathode: dust structure (cross section) (a) without a beam action and (b) with a beam action.

teristic acceleration of the macroparticle exceeds the radial force of the electric field that confines the particle in a layer. This is evidence that interparticle interaction forces in a plasma–dust crystal are substantially stronger than electric forces responsible for radial equilibrium of the crystal in a plasma trap. The effect of the dynamic action of an electron beam on a dust cloud can be used to locally remove particles from a plasma, for instance, in devices for growing particles or separating them according to size. Such a device for separating particles can be a glow discharge in a discharge tube with a variable cross section, for instance, a tube having the shape of a cone [9]. In such a discharge tube, longitudinal field varies along the tube axis and particles of different sizes are localized at different tube cross sections. The removal of particles of the required size in a certain cross section can be performed using a transverse electron beam.

REFERENCES

1. L. M. Vasilyak, S. P. Vetchinin, A. P. Nefedov, *et al.*, Zh. Éksp. Teor. Fiz. **121**, 609 (2002) [JETP **94**, 521 (2002)].
2. V. V. Balabanov, L. M. Vasilyak, S. P. Vetchinin, *et al.*, Zh. Éksp. Teor. Fiz. **119**, 99 (2001) [JETP **92**, 86 (2001)].
3. D. Samsonov, A. V. Ivlev, and G. E. Morfill, Phys. Rev. E **63**, 025401 (2001).
4. L. M. Vasilyak, S. V. Kostyuchenko, N. N. Kudryavtsev, and I. V. Filyugin, Usp. Fiz. Nauk **164**, 263 (1994) [Phys.–Usp. **37**, 247 (1994)].
5. A. F. Pal', A. O. Serov, A. N. Starostin, *et al.*, Zh. Éksp. Teor. Fiz. **119**, 272 (2001) [JETP **92**, 235 (2001)].
6. A. F. Pal', D. V. Sivokhin, A. O. Serov, *et al.*, Fiz. Plazmy (Moscow) **28**, 32 (2002) [Plasma Phys. Rep. **28**, 28 (2002)].
7. *An Encyclopedia of Low-Temperature Plasma*, Ed. by V. E. Fortov (Nauka, Moscow, 2000), Vol. 3, p. 84.
8. V. L. Bychkov and M. N. Vasil'ev, in *Proceedings of International Symposium on Theory and Application Aspects of Plasma Chemistry* (Riga, 1991), p. 40.
9. L. M. Vasilyak, S. P. Vetchinin, A. P. Nefedov, and D. N. Polyakov, Teplofiz. Vys. Temp. **38**, 701 (2000).

Translated by V. Sipachev

Journal of Experimental and Theoretical Physics, Vol. 96, No. 3, 2003, pp. 444–456.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 123, No. 3, 2003, pp. 503–517.

PLASMA, GASES

Effect of Dust Particles on the Properties of Low-Temperature Plasmas

M. A. Olevanov*, Yu. A. Mankelevich, and T. V. Rakhimova**

Skobel'tsyn Institute of Nuclear Physics, Moscow State University, Vorob'evy gory, Moscow, 119899 Russia
**e-mail: olevanov_m@mtu-net.ru*
***e-mail: ymankelevich@mics.msu.su*

Received July 15, 2002

Abstract—Parameters of a low-temperature plasma containing dust particles are calculated numerically with the help of a self-consistent solution of the balance equation for production and recombination of electrons and ions, combined with the molecular dynamics method for direct simulation of processes in the vicinity of macroparticles. The relation between the charges of macroparticles and the neutral gas pressure, as well as the dependence of the low-temperature plasma parameters on the volume concentration of dust particles, is analyzed. It is shown that the plasma characteristics and composition may change noticeably relative to the case unperturbed by dust even for comparatively low concentration of dust particles. *© 2003 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

Processes occurring in dusty plasmas provoke tremendous interest and have become objects of intense investigations in recent years. Among the most thoroughly studied phenomena, we can mention processes of self-organization occurring in dust–plasma systems; the formation of dust crystals [1–7] and gas-liquid phase transitions [8–13], development of instability in dust clouds [11, 14]; and interactions between macroparticles, which may lead to the formation of dust clusters. An understanding of the dynamics of the above-listed processes is of fundamental importance both for plasma physics and for physics of condensed systems. However, apart from purely theoretical interest, the study of these processes has become of practical significance. This is primarily due to rapid development of microtechnology as well as the methods for obtaining new materials in plasmochemical reactors.

Since the application of dusty plasma is regarded as one of the effective methods for synthesizing nanoparticles with unique physical properties, a grasp of the mechanisms governing the interaction of macroparticles is a decisive condition for creating the required technical basis. The main factor responsible for the emergence of a strong interaction between particles is the accumulation of considerable electric charges on macroparticles, which may attain values of $Z \sim 10^3$–10^4 electron charges depending on specific conditions. In view of the high mobility of electrons, this charge is usually negative; however, a number of effects such as secondary electron emission and photoemission [15] may lead to accumulation of a positive charge. Nevertheless, in spite of the fact that macroparticles accumulate considerable static charges of the same polarity, the presence of the plasma gives rise to additional attractive forces facilitating self-organization processes and the growth of clusters. A variety of models describing such processes have been proposed in the literature [16–18].

In connection with wide application of dc discharges in the technology of deposition of thin films with preset properties, the effect of macroparticles on the micro- and macroscopic parameters of the plasma, as well as the interaction between dust particles, is of special interest. The presence of particles in a plasma may significantly modify its properties due to the emergence of an extra source of recombination and (sometimes) production of electrons and ions. At the same time, dust may considerably affect the chemical and charge composition of a plasma as well as the rates of reactions occurring in it.

This study aims at numerical analysis of the effect of external factors (such as the volume concentration of dust, the neutral gas pressure, and the rate of ionization by an external source) on macroscopic and microscopic characteristics of the plasma, namely, ion and electron concentrations in the plasma, dust particle charges, recombination rate, and energy distribution function for plasma particles. An important factor that must be taken into account in constructing an adequate model of a plasma with dust particles is that the system components cannot be treated independently since all parameters of individual (ion, electron, and dust) subsystems, as well as processes occurring in these subsystems, strongly affect one another. Thus, for a given set of values of external factors, the system ultimately attains a state of dynamic equilibrium, which can be shifted in a certain way upon a change in external conditions.

1063-7761/03/9603-0444$24.00

In view of a large number of interdependent factors and processes, the problem formulated above is very complicated and difficult for analysis. The situation is complicated further still by the fact that the equations describing the dynamics of the system turn out to be strongly nonlinear. In this connection, the development of numerical methods for simulating such systems is of practical significance.

Numerical calculations have been widely used for a long time as a tool for studying the processes occurring in dusty plasmas [18–25]. The drift diffusion approximation, in which densities and fluxes of charged particles are determined by solving a system of equations (including the Poisson equation, continuity equations, and Boltzmann equations for determining the distribution function for plasma particles), is one of the most frequently used approaches. This approach was used in [19] for calculating the charge of dust particles, plasma fluxes to the surface of an individual dust particle, the concentration distribution of electrons and ions in the vicinity of a macroparticle, and other parameters in the plasma of a semi-self-maintained discharge in helium under atmospheric pressure. The same approach was employed in [20] for similar calculations in a nitrogen plasma under atmospheric pressure. The main criterion for applicability of this approach is the condition of the smallness of the mean free paths of ions and electrons as compared to the characteristic size of macroparticles and to size of the region of quasineutrality violation, whose size can be estimated by the Debye radius. A disadvantage of this approach is that additional assumptions are required for choosing the boundary conditions at the surface of a dust particle and at the boundary of the calculation cell.

Another widely used approach is the molecular dynamics method involving a direct solution of the equation of motion for ions and electrons in a self-consistent field produced by the plasma and the charge accumulated on dust particles. The advantage of this method is that the above-formulated problem is solved "from first principles" without using any additional a priori assumptions. However, the application of this approach is accompanied by other difficulties associated with a large number of particles for which calculations must be performed. These difficulties necessitate the introduction in actual computations of various simplified models of interaction, limiting the number of participating particles. This method was used for determining the type of screening of dust particles [21], for verifying the limits of applicability for the orbital motion limited (OML) approximation taking into account the interaction between ions and atoms of the neutral gas [22], and for calculating the plasma dynamics in the vicinity of a dust particle in the presence of an ion flow [23, 24]. In addition to the basic computational method mentioned above, some authors [25] also employed combined methods in which the hydrodynamic description was used for one of the subsystems and the Monte Carlo method was applied for describing another subsystem.

In this study, we carry out numerical simulation of processes in the plasma–dust system on the basis of the molecular dynamics method. Our object of investigation is the plasma of a semi-self-maintained discharge in helium in the presence of a certain amount of dust. The model adopted by us is oriented on studying the evolution of the system to a dynamic equilibrium state depending on external conditions. Consequently, the basic principle governing the construction of the model was that all the quantities characterizing the state of the system must attain their equilibrium values in a natural way. In particular, we have discarded any compulsory control of the number of particles in the calculation cell or the control of its electrical neutrality.

2. DESCRIPTION OF NUMERICAL MODEL

The initial state of the system was the homogeneous plasma of a semi-self-maintained discharge in helium. The mechanisms controlling the production and recombination of charged particles were ionization by an external source and the electron–ion recombination in the bulk of the plasma. It should be noted that we apply the term "semi-self-maintained discharge" not only to a classical high-pressure semi-self-maintained discharge controlled by an external electron beam with an energy of hundreds of kiloelectronvolts, but also to low-pressure quasi-semi-self-maintained discharges like those with runaway electrons.

The chosen ionization source ensured a constant rate $Q = 10^{15}$ cm^{-3} s^{-1} of plasma particle production per unit volume. Equilibrium values of ion and electron concentrations were formed by the balance between the production and recombination rates. In our model, the particle recombination constant was chosen so that the ion and electron concentrations in the absence of dust were $n_0 = 10^9$ cm^{-3}. The ions have a mass of 4 and are assumed to be singly charged.

The equilibrium temperature for ions is determined by the interaction with the neutral gas and is approximately equal to room temperature $T_i = 0.025$ eV. The electron temperature is much higher and is determined by the external electric field. Generally speaking, the electric field is one of the external factors determining the system dynamics. For this reason, the study of its effect on the process of dust particle charging, as well as on other processes occurring in the dusty plasma, may be of independent interest. Experience shows, however, that, since the electron temperature is one of the basic parameters characterizing the processes in the plasma–dust system, the main role of the electric field is associated with determining the value of this intrinsic parameter. The remaining influence of the electric field can be treated just as an additional correction to the temperature factor. In our model, the field remains constant and the electron temperature corresponding to its

value is $T_e = 2.5$ eV. We will use the Maxwell distribution as the equilibrium velocity distribution function for plasma particles.

Further, we bring a certain amount of dust into the plasma. In our model, we assume that all macroparticles are spherical in shape and have the same radius $a = 10$ μm. Such a size of particles is typical of many experiments and can be treated as an averaged value. In simulations, the dust particle concentration n_d varies in a wide range from 10^2 to 10^7 cm^{-3}, being one of the external parameters of the system. Plasma particles are initially neutral.

After the introduction of macroparticles, the system passes to a new state of dynamic equilibrium. The dynamics of transition to a new state can be described by the system of equations

$$\begin{aligned}\frac{dn_e}{dt} &= Q - \beta n_i n_e - \beta_{ed} n_d n_e,\\ \frac{dn_i}{dt} &= Q - \beta n_i n_e - \beta_{id} n_d n_i,\\ \frac{dZ_d}{dt} &= \beta_{id} n_d n_i - \beta_{ed} n_d n_e.\end{aligned} \tag{1}$$

Here, n_e, n_i, and n_d are the concentrations of electrons, ions, and macroparticles; Q is the ionization rate; β is the rate constant of the electron–ion recombination; Z_d is the charge of dust particles; and β_{ed} and β_{id} are the effective rate constants of electron and ion recombination on the surface of macroparticles. The following two remarks should be made as a comment on the above system of equations. First, all the parameters appearing in the system are quantities averaged over volume. This should be borne in mind in calculating the concentrations of ions and electrons, since their values experience very strong perturbations in the vicinity of dust particles and noticeably deviate from their mean values. Second, the rate constant of electron and ion recombination on the surface of macroparticles during the evolution of the system are not constant but are formed by the charge of dust particles and by the plasma fluxes to their surface.

Thus, in order to solve system (1), we must know the flows of charged particles to the surface of dust for any point of time. In our model, these flows are calculated using an algorithm based on the molecular dynamics method. We assume that all macroparticles have the same size and shape. Then, if we disregard fluctuations due to the presence of other dust particles in their neighborhood, the processes occurring in the vicinity of each particle (namely, its charging and the action on the surrounding plasma) can be treated as identical. Under these conditions, the system dynamics can be investigated by choosing a dust particle as a representative. In our model, a single dust particle is at the center of a calculation cell, which is spherical like the particle itself. The choice of such geometry is dictated by the symmetry properties of the problem.

The cell radius L is chosen equal to 60 μm, which is six times as large as the radius of the particle. However, in regard to establishing the boundary conditions, such a relation is not optimal and it would be better to increase the proportion. For the initial concentrations of ions and electrons $n_0 = 10^9$ cm^{-3}, a cell of the chosen size contains approximately 1800 particles; however, since the number of electrons and ions in the cell is not fixed and is determined by the ratios of the rates of their production and recombination, it can vary in the course of computations from hundreds to a few tens of thousands depending on the conditions.

Electrons and ions appear in the cell together with a flow through its external boundary or they can be produced directly in the cell volume. The production of particles in the cell is simulated as a random process whose probability per unit time is determined by the ionization rate. The generated particles are distributed uniformly over the cell volume, and their initial velocity distribution is assumed to be a Maxwell distribution. The conditions at the external boundary of the cell will be considered at a later stage. The number of plasma particles may decrease due to their absorption at the surface of the dust particle, due to recombination in the cell volume and also due to the departure of particles through the external boundary of the simulation region.

The trajectories of charged particles in the calculation cell were determined by direct integration of the system of equations of motion:

$$\begin{aligned}\mathbf{v}_k(t+\tau) &= \mathbf{v}_k(t) + \frac{\mathbf{F}_k(t)}{m_k}\tau,\\ \mathbf{r}_k(t+\tau) &= \mathbf{r}_k(t) + \mathbf{v}_k(t)\tau + \frac{\mathbf{F}_k(t)}{m_k}\frac{\tau^2}{2}.\end{aligned} \tag{2}$$

Here, $\mathbf{r}_k$ is the radius vector of the kth particle relative to the center of the cell, $\mathbf{v}_k$ is the particle velocity, m_k is the particle mass, $\mathbf{F}_k$ is the force acting on the particle, and τ is the time step. The value of the time step was recalculated on each iteration. The criterion for the recalculation was that the distance traversed by the fastest particle from the simulation region must be much smaller than the size of the dust particle, which plays the role of the characteristic scale of length in the present case: $\tau \ll a/v_{\max}$.

Let us consider the forces that cause the motion of particles. These forces can be written in the form

$$\mathbf{F}_k = \frac{Z_d q_k \mathbf{r}_k}{r_k^3} + \sum_{i \neq k}^{N} \frac{q_i q_k \mathbf{r}_{ik}}{r_{ik}^3} + \mathbf{f}_{k0}. \tag{3}$$

The first term in this expression describes the interaction of the kth particle with the charge accumulated by

the dust particle. Here, q_k is the charge of a microparticle, which is equal to $+e$ for ions and $-e$ for electrons, and $\mathbf{r}_k$ is the distance from the center of the dust particle. The second term describes the interaction between ions and electrons. Here, $\mathbf{r}_{ik}$ is the distance between the ith and kth particles. The last term in Eq. (3) describes the interaction of the kth microparticle with atoms of the neutral gas.

Since we must calculate forces taking into account pair interaction between electrons and ions, and the number of participants of this interaction is very large, we must use additional assumptions to reduce the computation time. A popular and widely used method is selection of plasma particles falling inside the cutoff interaction radius [26]. In other words, in calculating the force acting on a chosen particle, only the contributions from its nearest neighbors separated from it by a distance not exceeding the interaction length are taken into account. However, this approach is inapplicable in our case for certain reasons. The role of the characteristic cutoff length that determines the neighbor search region for plasma particles is played by the Debye radius, which has a value of $\lambda_d \approx 40$ μm for the chosen conditions. First, the size of the corresponding region is larger than the size of the dust particle and is comparable in order of magnitude to the size of the calculation cell. Thus, a certain gain can be obtained only for much higher concentrations of charged particles in the cell. Second, the application of the given approach is ruled out by the fact that the charge accumulated on the dust particle leads to such a strong redistribution of concentrations of electrons and ions in the vicinity of the macroparticle that disregarding the contribution of the volume charge of these regions according to the principle that it does not fall within the cutoff length would lead to a large computational error.

Taking into account the above considerations, we have chosen a different way based on the geometrical features of the problem. The number density distribution of particles in the cell is spherically symmetric; consequently, the electric field also possesses spherical symmetry and is directed, on the average, to the center of the cell at any point of it. If we draw a sphere with the center at the middle of the cell through the chosen point, the electric field will be determined, in accordance with the Gauss theorem, only by the charge concentrated in the selected sphere:

$$\tilde{Q} = (Z_d + N_i - N_e)e. \tag{4}$$

Here, Z_d is the charge of the dust particle and N_i and N_e are the total numbers of ions and electrons in the sphere.

Let us now consider in greater detail the mechanism of plasma particle interaction with neutral gas atoms. The data on the interaction cross sections were borrowed from [27].

For electrons in helium, the frequency of elastic collisions is approximated by the following dependence:

$$\frac{\nu_{e0}}{P}[10^9\ \text{Torr}^{-1}\ \text{s}^{-1}] = \begin{cases} 2.4\dfrac{T_e\ [\text{eV}]}{4}, & T_e < 4\ \text{eV}, \\ 2.4, & T_e \geq 4\ \text{eV}. \end{cases} \tag{5}$$

Here, ν_{e0} is the frequency of elastic collisions, P is the neutral gas pressure, and T_e is the electron temperature. In the case of electrons, the contribution from inelastic collisions can be neglected, since these collisions become significant at higher values of temperature.

In the description of the interaction between ions and neutral gas atoms in our model, we have taken into account the polarization mechanism of the interaction and resonant recharging. The polarization interaction cross section has the form

$$\sigma_{\text{pol}} = 2\sqrt{2}\pi a_0^2 \sqrt{\frac{\alpha}{a_0^3}\frac{2I_H}{m_i v^2}}. \tag{6}$$

Here, $a_0 = 0.529 \times 10^{-8}$ cm is the first Bohr radius, $I_H = 13.6$ eV is the ionization potential for a hydrogen atom, m_i is the mass of an ion, v is its velocity, and α is the polarizability of a neutral atom. For a helium atom in the ground state, we have $\alpha/a_0^3 = 1.39$.

The resonant recharging cross section is given by

$$\sigma_{\text{res}} = \pi a_0^2 \frac{I_H}{I} \ln^2\left(\frac{100 v_0}{v}\sqrt{\frac{I}{I_H}}\right). \tag{7}$$

Here, $I = 24.6$ eV is the ionization potential for a helium atom, $v_0 = 2.19 \times 10^8$ cm/s is the electron velocity on the first Bohr orbit, and v is the velocity of an ion.

Taking into account relations (6) and (7), we can write the expression for the frequency of collisions between ions and neutral gas atoms in the form

$$\nu_{i0} = (\sigma_{\text{pol}} + \sigma_{\text{res}}) n_0 v. \tag{8}$$

Here, $n_0 = P/kT_0$ is the concentration of neutral gas atoms, $kT_0 = 0.025$ eV is the neutral gas temperature, P is the pressure, and v is the velocity of an individual ion.

The interaction of plasma particles with the neutral gas is of the type of random collisions. In each collision, the velocity of an ordered plasma particle motion is lost and a part of energy is transferred to a neutral atom. In our model, we take into account this interaction by adding an effective frictional force, which was denoted earlier by $\mathbf{f}_{k0}$. In our case, this way is preferential as compared to the statistical method of description, i.e., the method of random generation of collisions in

the calculation cell. As a matter of fact, when we consider the interaction with neutral atoms, the main factor producing the strongest effect on the evolution of the system is the loss of ion energy (which was acquired in the field of the dust particle) during collisions. As a result, the ions get into a potential trap and their trajectories become finite. Over the course of time, every ion of this sort experiences new and new collisions, losing energy each time, and ultimately finds itself on the surface of the macroparticle. If we use the statistical description, we should trace the trajectories of all such ions (even those which must leave the simulation region). In addition, the times of their finite motion may turn out to be significant as compared to time periods over which the system dynamics can be simulated in actual practice. For this reason, we have chosen the model in which collisions are treated not as discrete events, but as an interaction distributed over the time. The formula for the effective frictional force can be written in the form

$$\mathbf{f}_{k0} = -\nu_{k0}\max\left(0, 1 - \frac{T_0}{\varepsilon_k}\right)\mu\mathbf{v}_k. \qquad (9)$$

Here, ν_{k0} is the frequency of collisions with neutral gas atoms, T_0 is the neutral gas temperature, ε_k is the kinetic energy of a particle, μ is the reduced mass, and $\mathbf{v}_k$ is the particle velocity. The second factor in Eq. (9) has been introduced to take into account the effectiveness of energy transfer in collisions. This is required since, in simulating the motion of an individual particle, it is impossible to separate the energy acquired in external fields from its thermal energy. At the same time, if the particle is thermalized, the energy transfer to the neutral gas must cease.

Taking into account relations (4) and (9), we can rewrite expression (3) for the force acting on a microparticle in the form

$$\mathbf{F}_k = \frac{\tilde{Q}q_k\mathbf{r}_k}{r_k^3} - \nu_{k0}\max\left(0, 1 - \frac{T_0}{\varepsilon_k}\right)\mu\mathbf{v}_k. \qquad (10)$$

To implement the algorithm of force computation, the calculation region was divided into spherical layers. The value of $\tilde{Q}$ and, accordingly, the value of the electric field in each layer were recalculated on each time step, and the force exerted on an individual particle was determined from its occurrence in a certain layer.

Let us now discuss the boundary conditions on the outer surface of the calculation cell. In order to maintain the balance of the number of particles in the simulation region, we must know the ion and electron flows through the outer surface to the bulk of the cell. Until the dust particle accumulates a considerable charge, these flows are the same as in the case of plasma unperturbed by the presence of dust. Thus, the number of particles getting into the cell per unit time is determined by the average concentrations n_e and n_i of electrons and ions as well as by the area of the outer surface of the cell. The velocity distribution function for particles is isotropic and assumed to be of the Maxwell type. In order to determine the flow unperturbed by the presence of dust, preliminary computation without a macroparticle in the simulation region is carried out. Under equilibrium conditions, the particle flows outside the cell and those directed to its bulk coincide. For this reason, instead of each departing particle, a new particle is created at random at an arbitrary site on the cell boundary, the velocity of this particle being directed to the bulk of the cell. Then, the numbers of electrons and ions falling on the external boundary are counted for some time. After collection of acceptable statistics and determination of plasma flows with sufficient accuracy, the electron and ion production rates on the outer boundary are preserved and are subsequently used in the algorithm simulating at random the inflow of electrons and ions into the simulation region through its outer surface.

While the charge of the dust particle increases, an appreciable potential is created at the boundary of the simulation region, which rules out the use of the electron and ion production rates, determined in the unperturbed state, in subsequent computations. This difficulty can be overcome simply by increasing the cell size; however, as mentioned above, a purely technical restriction exists that limits the maximal number of particles whose motion can be calculated using the molecular dynamics method. For this reason, we had to use auxiliary considerations for correcting the flows through the external boundary as well as the form of the velocity distribution function for particles falling into the calculation cell.

In order to solve the problem in our model, we use the following procedure. First of all, we find a sphere of radius $\tilde{L}$ for electrons and ions, beyond which the plasma can be treated as unperturbed. As a criterion, we use the assumption that the average kinetic energy of particles at the boundary of the sphere has the same order of magnitude as the potential energy of interaction with a macroparticle. This criterion can be written in the form

$$T_{i,e} \sim |e\varphi(L)|\exp\left(-\frac{\tilde{L} - L}{\tilde{\lambda}_d}\right). \qquad (11)$$

Here, $T_{i,e}$ are the temperatures of ions and electrons, which are the measures of their kinetic energy; $\varphi(L)$ is the potential at the boundary of the calculation cell; and $\tilde{\lambda}_d$ is the effective screening radius. It is necessary to use the effective screening radius instead of the conventional Debye radius, since screening becomes nonlinear

on account of large values of the charge accumulated by the macroparticle. In order to determine the value of $\tilde{\lambda}_d$, we make use of the equation

$$\varphi(L) = \frac{Z_d}{L}\exp\left(-\frac{L}{\tilde{\lambda}_d}\right). \quad (12)$$

Here, L is the calculation cell radius, Z_d is the macroparticle charge, and $\varphi(L)$ is the value of potential at the cell boundary, which is determined in the course of simulation. Calculations show that the value of $\tilde{\lambda}_d$, in contrast to the Debye radius, may be as high as several hundred micrometers. The values of $\tilde{\lambda}_d$ and $\tilde{L}$ are recalculated on each time step.

When the size of the region beyond which the plasma can be regarded as an equilibrium plasma exceeds the size of the calculation cell, the algorithm correcting the flows at the external boundary comes into play. Particles are generated at the boundary of sphere $\tilde{L}$ at a rate corresponding to the conditions in the unperturbed plasma. Their spatial positions, directions of motion, and the magnitudes of their velocity are set at random. For each added electron or ion, the calculation cell with all particles contained inside plays the role of a charged sphere in whose field particle motion takes place. If we disregard for the time being the interaction with the neutral gas, we can easily determine from the momentum and energy conservation laws whether a generated particle gets into the calculation cell and, if it does, its parameters. Denoting the velocity of the generated particle by v_0 and the angle between the direction of the velocity vector and the normal to the surface of sphere $\tilde{L}$ by θ_0 (see Fig. 1), we can write the condition for the particle getting into the calculation cell in the form

$$\frac{\tilde{L}v_0}{Lv}\sin\theta_0 < 1. \quad (13)$$

In this relation, v stands for the velocity with which the particle falls into the calculation cell:

$$v = \sqrt{v_0^2 \pm \frac{2e[\varphi(\tilde{L}) - \varphi(L)]}{m}}. \quad (14)$$

Here, the plus sign corresponds to ions, and the minus sign, to electrons. If the radicand is found to be negative (which may be the case for electrons), this also means that the particle does not reach the surface of the computational cell. Such verification is carried out for each new electron or ion, which can appear in the computational region only if condition (13) is satisfied.

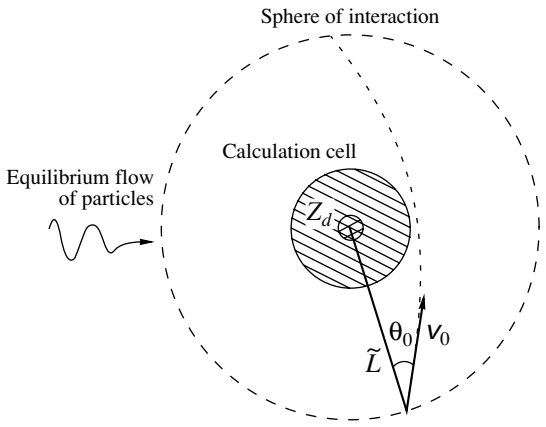

Fig. 1.

The proposed correction algorithm makes it possible to take into account the deviation of flows from their equilibrium values due to the fact that the calculation cell is not electrically neutral. However, in addition to this factor, there also exist auxiliary flows associated with the effect of the interaction of charged particles with the neutral gas. This concerns primarily the ions, which, while moving in sphere $\tilde{L}$, may experience an inelastic collision; this inevitably leads to their falling into the calculation cell after a certain time. This possibility is taken into account by additional coin tossing for each ion. The collision probability can be estimated as

$$p \sim \frac{2\tilde{L}\nu_{i0}\cos\theta_0}{v_0}. \quad (15)$$

An ion that has experienced a collision hits the cell at the thermal velocity.

Let us return to system of equations (1) and rewrite it in the form

$$\begin{aligned} n_e(t+\tau) &= n_e(t) + [Q - \beta n_i(t)n_e(t)]\tau - dN_e n_d, \\ n_i(t+\tau) &= n_i(t) + [Q - \beta n_i(t)n_e(t)]\tau - dN_i n_d, \\ Z_d(t+\tau) &= Z_d(t) + dN_i - dN_e. \end{aligned} \quad (16)$$

Here, dN_i and dN_e are the numbers of ions and electrons absorbed by the macroparticle over time interval τ. Thus, starting the computational procedure and counting directly the number of particles absorbed by the surface of the dust particle per unit time, we can trace the dynamics of the system evolution to the equilibrium state. In turn, a change in the averaged concentration of ions and electrons with time leads to a change in the flows of plasma particles through the external boundary to the bulk of the calculation cell, since these flows are proportional to n_e and n_i. Taking this factor into

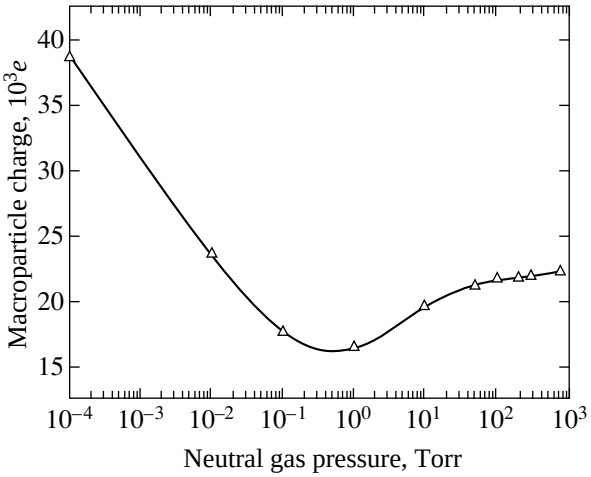

Fig. 2. Dependence of the macroparticle charge on the neutral gas pressure: $Q = 10^{15}$ cm^{-3} s^{-1}, $n_0 = 10^9$ cm^{-3}, $n_d = 10^3$ cm^{-3}, $T_e = 2.5$ eV, $T_i = 0.025$ eV, $a = 10$ μm.

account, we obtain a completely closed and self-consistent solution to the formulated problem.

3. RESULTS OF SIMULATION

Using the approach described above, we simulated the dynamics of processes occurring in the plasma–dust system upon a change in the neutral gas pressure and the dust concentration in the bulk for a constant value of parameter E/P.

The dependence of the charge accumulated by an individual dust particle on the neutral gas pressure is shown in Fig. 2. The parameters used in calculations are as follows: $Q = 10^{15}$ cm^{-3} s^{-1}; the ion and electron concentrations unperturbed by the presence of dust are n_i, $n_e = 10^9$ cm^{-3}; $n_d = 10^3$ cm^{-3}; $T_e = 2.5$ eV; and $T_i = 0.025$ eV. The macroparticle radius is $a = 10$ μm, and the size of the calculation cell is $L = 60$ μm. The pressure varied from 10^{-4} to 760 Torr. It can be seen from the curve that the absolute value of charge behaves nonmonotonically and decreases sharply to half the initial value upon a change in pressure approximately to 1 Torr; then the curve displays an insignificant increase, after which the charge attains saturation and remains approximately constant upon an increase in pressure up to atmospheric pressure.

Analyzing the obtained dependence, we conclude that the sharp decrease in the absolute value of charge upon an increase in pressure from 10^{-4} to 1 Torr is due to the fact that the number of inelastic collisions between ions and neutral gas atoms in the vicinity of the particle increases. As a result, ions find themselves in a potential well and, losing energy in subsequent collisions, ultimately fall on the surface of the macroparticle, thus increasing the flow of positive charge. The size of the region in which the ions experiencing a collision get into the potential trap can be estimated by equating the energy of interaction with the macroparticle to their average thermal energy. As the pressure increases further, the mean free path of ions decreases to such an extent that practically every ion falling into the sphere of interaction experiences a collision and is trapped by the dust particle. This means that the cross section of interaction between ions and the macroparticle stops increasing starting from a certain instant and is determined by the size of the interaction region. Thus, we can draw the conclusion that the charge of the macroparticle remains practically unchanged in a wide range of pressures. This can be clearly seen from Fig. 2.

Using the obtained result, we can estimate the size of the ion–macroparticle interaction region. For this purpose, we equate the electron flux obtained in the OML approximation to the gas-kinetic flux of ion to the surface of the sphere in which an ion gets trapped. It should be noted at the very outset that the application of the OML approximation for electrons is justified, since the cross section of their interaction with neutral gas atoms is small and the mean free path is much larger than the size of the region in which the interaction with the macroparticle noticeably affects the trajectories of motion. We write the final expression in the form

$$2\sqrt{2}n_e v_{T_e} a^2 \exp\left(-\frac{e|\phi_0|}{T_e}\right) = 2\sqrt{2}n_i v_{T_i} \tilde{\lambda}^2. \qquad (17)$$

Here, v_{T_i} and v_{T_e} are the thermal velocities of ions and electrons, n_i and n_e are the mean concentrations of charged particles, ϕ_0 is the potential of the macroparticle surface, a is the radius of the dust particle, and $\tilde{\lambda}$ is the size of the interaction region. Deriving $\tilde{\lambda}$ from Eq. (17), we obtain the following estimate:

$$\tilde{\lambda} = a\sqrt{\frac{1}{1+P}}\sqrt{\frac{m_i T_e}{m_e T_i}}\exp\left(-\frac{Z_d e^2}{2T_e a}\right). \qquad (18)$$

Here, $P = Z_d n_d / n_e$ is a dimensionless parameter characterizing the relative charge accumulated at dust particles under the dynamic equilibrium conditions. Substituting the values of parameters listed above and $Z_d \approx 20000$ into the expression obtained and considering that $P \ll 1$ for the dust concentration $n_d \approx 10^3$ cm^{-3}, we obtain $\tilde{\lambda} \approx 165$ μm. This value is much larger than the Debye screening radius in plasma, which is equal to $\lambda_d \approx 40$ μm under the above conditions.

The parameters of the system used for estimating $\tilde{\lambda}$ are typical of most experiments with plasma–dust crystals; according to the results of such experiments [1–7], the characteristic size of the crystal lattice constant is on the order of 100–300 μm. In respect to the theory of

crystal structure formation in a plasma, this fact leads to considerable difficulties, since the observed lattice constant is much larger than the predicted interaction length that can be estimated by the Debye radius. This discrepancy provoked considerable interest in the screening mechanisms in dusty plasmas.

The quantity $\tilde{\lambda}$ estimated by us has an important physical meaning. It is just the size of a macroatom with a dust particle as the nucleus, in which the role of the electron cloud is played by the cloud of ions in the potential trap. In this connection, we can view the process of dust crystal formation from a slightly different standpoint; namely, we can consider the above-mentioned atoms as the building blocks of the crystal structure. Such an approach differs from the traditional one. In the conventional treatment, plasma is regarded just as a factor responsible for the screening of the macroparticle interaction potential. The crystal obtained in this case is of the purely Coulomb type. In contrast to the traditional approach, the variant considered here envisages the existence of crystals with a free boundary. In addition, we can state that the interaction between ions and neutral gas atoms, which is completely disregarded in the traditional model, may in fact become the main factor in the description of the mechanisms of self-organization in a plasma–dust system. Returning to the results of experiments, we note that the value of $\tilde{\lambda}$ obtained by us is in full accordance with the observed distance between particles in dust crystals. Although crystals with a free boundary have not been obtained as yet, plasma–dust drops were indeed observed [16].

Continuing the analysis of Fig. 2, we note that the change in the equilibrium charge of macroparticles upon an increase in the neutral gas pressure is nonmonotonic. This can be explained by the fact that, upon an increase in the gas pressure, ions in the vicinity of the dust particle start experiencing multiple collisions, thereby losing the energy accumulated in the accelerating field of the particles. This leads to a decrease in the positive charge flow to its surface. As regards the description of the system, this necessitates a transition to the hydrodynamic approximation, while the OML approximation can be used for low pressures. However, the sharp change in the value of charge at the beginning of the curve indicates that, even in the case of low pressure, this approximation may lead to a significant error, since even individual collisions in the vicinity of the macroparticle lead to trapping of ions and their fall to its surface.

Let us now analyze the evolution of the system for various values of dust volume concentration. Figure 3 shows a family of curves illustrating the dynamics of the macroparticle charge evolution for different concentrations of dust particles. Simulation was carried out for the same values of parameters as in the previous case, but for a fixed value of neutral gas pressure of $P = 1$ Torr. The dust concentration varies from 10 to 10^7 cm^{-3}. Along the abscissa axis, we laid the time elapsed from the introduction of neutral dust particles into a volume filled with ionized gas. The value of pressure chosen for simulating the system dynamics was not accidental: it corresponds to the transition region between the OML approximation and the hydrodynamic regime. Numerical calculations for this range of parameters are of special interest, since analytical methods are the least suitable in this case.

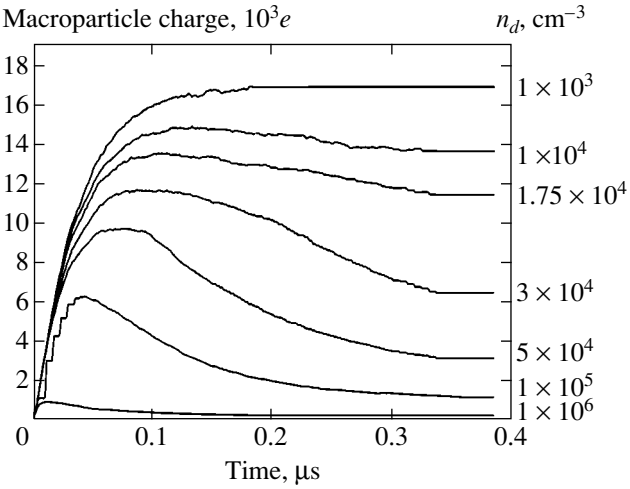

Fig. 3. Time evolution of the macroparticle charge for different values of dust concentration. At the initial instant, all particles are neutral: $Q = 10^{15}$ cm^{-3} s^{-1}, $n_0 = 10^9$ cm^{-3}, $P =$ 1 Torr, $T_e = 2.5$ eV, $T_i = 0.025$ eV, $a = 10$ μm.

Figure 3 shows that the processes leading to the establishment of a dynamic equilibrium state exhibit different modes of evolution for different numbers of macroparticles per unit volume. For low concentrations of dust, the charge attains a steady-state value via an ordinary relaxation process with a characteristic charging time on the order of 0.1 μs. However, the shape of the curve changes upon an increase in the macroparticle concentration. First, the absolute value of the charge attains its peak and then decreases to a certain steady-state value. The rate of the system transition to the stationary state increases with the volume concentration of dust.

The change in the dynamics of charging can be explained by analyzing the evolution of processes in the system from the viewpoint of the balance between the production and recombination of charged particles in the plasma. Since electrons possess higher mobilities and energies as compared to ions, the rate of their recombination at the surface of macroparticles for fairly high dust concentrations is much higher than the rate of destruction of ions. As a result, the equilibrium between the production and recombination of electrons in the bulk sets in much earlier than for ions, and their flow to the surface of macroparticles stops changing. This leads to the emergence of a characteristic maxi-

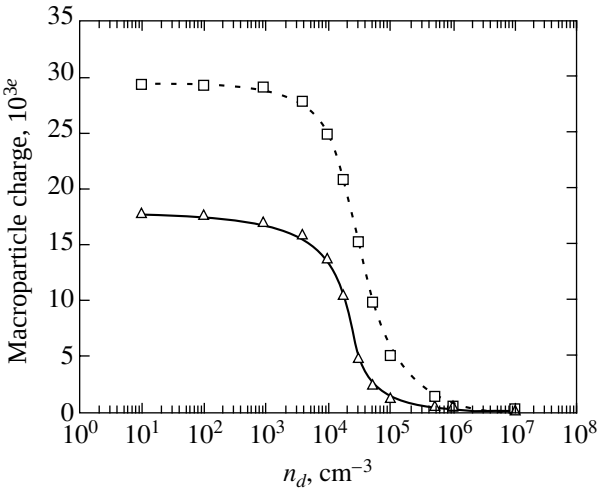

Fig. 4. Dependence of the macroparticle charge in equilibrium on the dust concentration. Solid curve is the result of simulation and dashed curve is the result of calculation in the orbital motion limited approximation.

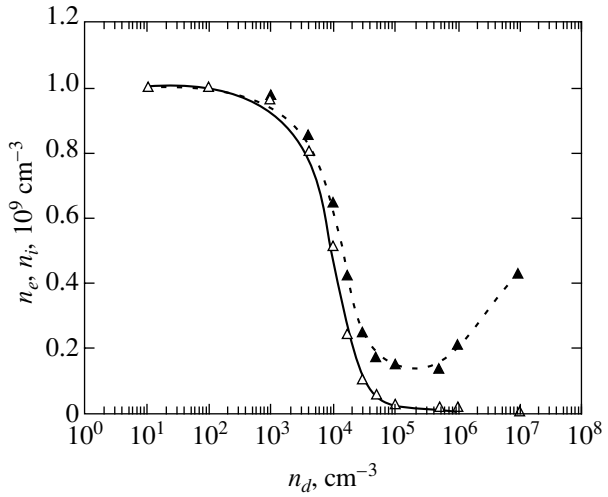

Fig. 5. Dependence of the average concentration of ions (solid curve) and electrons (dashed curve) in equilibrium on the dust concentration.

mum on the charging curve since the equilibrium for ions sets in more slowly, and the ion flow to the surface continues to increase for some more time.

Figures 4 and 5 show the dependences on the dust particle charge Z_d as well as average electron and ion concentrations in dynamic equilibrium on the concentration of macroparticles in the bulk. The ion and electron concentrations are expressed in relative units. As a unit, we have chosen the equilibrium concentrations n_{i0}, $n_{e0} = 10^9$ cm^{-3} in the absence of dust. It can be seen from the figures that the charge of macroparticles increases with the number of dust particles and that the steady-state values of n_e and n_i change almost synchronously. This confirms the thought that the states of the dust and plasma subsystems are strongly interrelated under real conditions. The effect of the large amount of dust in the bulk on the dynamics of the system is manifested starting from $n_d \sim 10^4$ cm^{-3}; in this case, a sharp decrease in the charge accumulated by dust particles takes place, and the degree of ionization also decreases. A similar behavior of the system was observed earlier in experiments on the dynamics of dust particle charging [28].

The observance of the quasineutrality condition $n_e + Z_d n_d = n_i$ was not specially controlled by the numerical algorithm; the results of calculation show, however, that the system passed spontaneously to a state in which this condition is satisfied almost exactly. A deviation from the quasineutrality condition expressed in the above form takes place only for very high dust concentrations $n_d \sim 10^7$ cm^{-3}. This is a consequence of the fact that the size of the calculation cell becomes comparable with mean distances between dust particles, and we find ourselves almost at the boundary of the proposed scheme applicability. The quasineutrality condition and the symmetry of the problem presume that the space filled with macroparticles can be divided into electrically neutral cells in which the balance between the production and recombination of plasma particles is maintained in the steady state. The volume of such a cell can be estimates at $V \sim n_d^{-1}$. Since the entire space is filled with such cells, the average flow of plasma particles through the boundary of each cell can be assumed to be equal to zero. The above estimate of V can be used as the measure of the calculation cell size, for which the application of periodic boundary conditions is justified in various numerical models. The use of the Debye radius for such an estimate may lead to large errors since the calculation of screening in the case of dust particles cannot be carried out in the framework of the linear approximation.

In our model, for low concentrations of dust particles, the size of a cell in which the quasineutrality conditions must be observed is much larger than the size of the calculation region. Under this condition, the balance between the particle production and recombination is controlled by flows from the external cell boundary and the sink to the surface of a macroparticle. However, as the concentration of macroparticles increases, the cell volume V becomes comparable with the size of the simulation region, and the role of flows from the external boundary becomes smaller. In this case, the processes of ionization and recombination in the computational region itself play the major role. Thus, Eqs. (1) describing the balance between the particle production and recombination in the system in the form averaged over volume should be replaced by a purely statistical description of these processes. In our model, the latter regime was implemented by generating the corresponding events with allowance made for their

probability. In the latter case, the quasineutrality condition is satisfied for the calculation region proper, but noticeable charge fluctuations are possible under these conditions and were actually observed in our simulations.

Let us return to analysis of the evolution of plasma particle concentration upon an increase in the amount of dust in the volume. Figure 5 shows that, starting from a value of $n_d \sim 10^5$ cm^{-3}, the concentration of positive ions increases. This is due to the fact that, under the given conditions, dust is the main mechanism of recombination for plasma particles; however, the cross section of ion recombination at macroparticles noticeably decreases in view of the decrease in the absolute value of negative charge accumulated by macroparticles. On the whole, the electron concentration decreases to such an extent that the system passes to the state of a purely ionic two-component plasma. Positively charged gas ions play the role of the light component, while negatively charged macroparticles form the heavy component. This consideration may be useful in constructing a kinetic description of the system in the case of a high dust concentration.

Figure 6 shows the rate constants of electron and ion recombination at macroparticles under dynamic equilibrium as functions of the dust concentration. For calculating β_{ed} and β_{id}, we have used Eqs. (1) with a zero left-hand side and with the values of n_e and n_i obtained as a result of simulation. In this case, we can write the following expressions for β_{ed} and β_{id}:

$$\beta_{ed} = \frac{Q - \beta n_i n_e}{n_d n_e}, \qquad \beta_{id} = \frac{Q - \beta n_i n_e}{n_d n_i}. \tag{19}$$

The dependence is nonmonotonic on the whole. This means that different modes of evolution to equilibrium take place depending on the relative role of various factors. For a low dust concentration in the volume, the equilibrium between the production and recombination of ions and electrons is maintained due to the electron–ion recombination. As a result, the concentrations of plasma particles are nearly identical; consequently, the rate constants of recombination at dust particles are also equal. This conclusion follows from the condition of equality of the ion and electron flows to the surface of macroparticles. The enhancement of the role of recombination at dust particles becomes significant starting from a concentration of $n_d \sim 10^4$ cm^{-3}. This is manifested in the divergence of the curves in Fig. 5.

In respect to the macroparticle charging dynamics, the competition of two processes with different relaxation times is possible, which determines the path of the system to equilibrium. The first process corresponds to the establishment of balance for ion and electron flows to the surface of a macroparticle, while the second process leads to equilibrium between the production and recombination of plasma particles in the bulk. The situation when equilibrium between the flows to the macroparticle surface sets in much later than the balance between the rates of particle production and recombination in the bulk is realized for dust concentrations $n_d \gtrsim 10^5$ cm^{-3}. Under these conditions, the absolute value of the charge of dust particles decreases, leading to a rapid increase in the electron recombination rate constant at macroparticles and to a decrease in the corresponding constant for ions. As the concentration of dust increases further, the electron recombination rate constant passes through a maximum and also starts decreasing. This is due to the fact that the existing source of ionization cannot maintain the increasing rate of electron departure from the bulk any longer. Thus, the electron flow to the surface of individual macroparticles decreases.

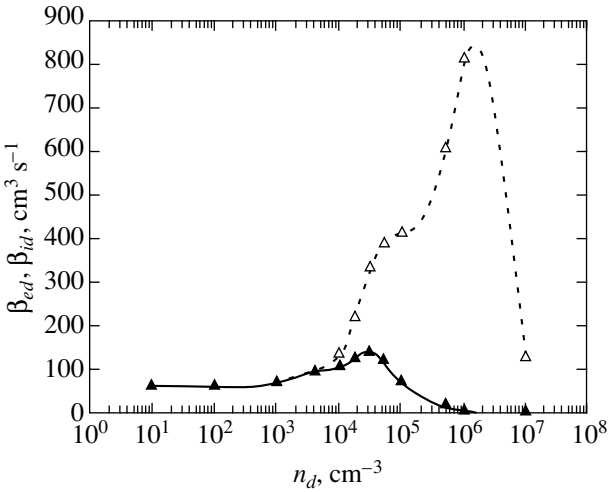

Fig. 6. Dependence of the rate constants β_{id} and β_{ed} of ion (solid curve) and electron (dashed curve) recombination at the surface of macroparticles on the dust concentration.

An analysis of the curves shows that the rate constants of recombination at macroparticles are not constant quantities depending only on the features of interaction between macroparticles and plasma particles, but are determined by the state of the system as a whole, which is given by a set of external factors such as the ionization rate or the concentration of macroparticles in the bulk of the plasma. This result is a consequence of the open nature of the system. It should also be noted that the rate constants of recombination at dust particles are many orders of magnitude higher than the rate constant of the electron–ion recombination. In our model calculations, the value of this rate constant was taken equal to $\beta = 10^{-3}$ cm^3 s^{-1}, which was generally dictated by computational considerations since the number of particles participating in simulation cannot be extremely large. Under real conditions in helium, the electron–ion recombination rate constant is on the order of $\beta = 10^{-8}$–10^{-9} cm^3 s^{-1}. Let us see how this

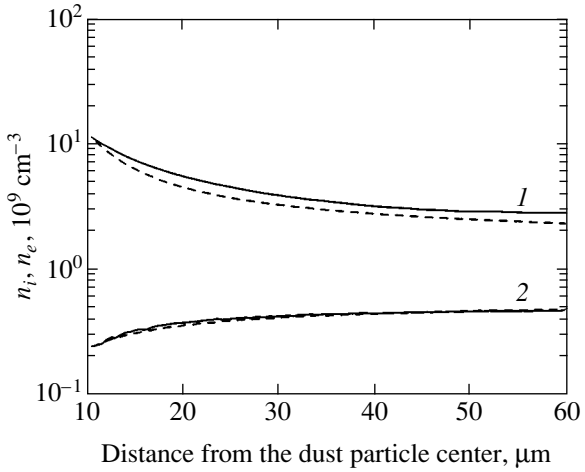

Fig. 7. Dependence of the ion (*1*) and electron (*2*) concentration in the vicinity of a dust particle on the distance from its center; $n_d = 10^4$ cm^{-3}, $P = 1$ Torr. Solid curves correspond to the density distribution obtained in numerical experiment and dashed curves correspond to the equilibrium concentration distribution.

affects the parameters of the system in dynamic equilibrium. For this purpose, we solve system of equations (1) in the stationary case for n_e and n_i and write the asymptotic expressions for the concentrations for small values of n_d:

$$n_e \approx n_0 - \frac{\beta_{ed}}{\beta} n_d, \qquad n_i \approx n_0 - \frac{\beta_{id}}{\beta} n_d. \tag{20}$$

For $\beta \sim 10^{-8}$ cm^3 s^{-1} and $\beta_{id} \approx \beta_{ed} \sim 10^2$ cm^3 s^{-1}, the ratio β_{ed}/β is on the order of 10^{10}. The equilibrium concentrations of charged particles under the same conditions are $n_0 \approx 10^{11}$ cm^{-3}. The estimates obtained above indicate that even a minimal concentration of dust in the volume of about 10^2 cm^{-3} must lead to a sharp decrease in the concentration of charged particles and the role of bulk recombination in the establishment of dynamic equilibrium must be small as compared to the processes of recombination at the macroparticle surface.

However, this conclusion strongly contradicts the results of a number of experiments. The dependence of the current in a semi-self-maintained discharge in helium on the dust concentration reported in [19] shows that a noticeable decrease in the number of plasma particles is observed only for dust concentrations on the order of 10^4 cm^{-3}. Consequently, the electron–ion recombination treated in our model as the main factor responsible for the balance of the production and recombination of charged particles is far from being the only way, and other possibilities can also be realized in actual practice. For example, the balance equations should be supplemented with diffusion and drift fluxes of charged particles. In this case, the problem ceases to be spatially homogeneous and geometrical features of specific experiments must be taken into account in calculations.

Summarizing the discussion of the system parameters in equilibrium, let us compare our results with the results of similar calculations made in the OML approximation. This may be interesting, since the latter approximation is widely used in applications. The dashed curve in Fig. 4 describes the charge of macroparticles obtained in the OML approximation. It can be seen from the figure that the values of charges calculated for a neutral gas pressure of 1 Torr differ almost by a factor of 2. This is another evidence of the fact that the interaction of plasma particles with neutral gas atoms is a significant factor determining the system dynamics, which cannot be disregarded while constructing models.

Let us now consider processes in the immediate vicinity of a macroparticle. Figure 7 shows the time-averaged distributions of the ion and electron concentrations as functions of the distance to the center of an individual dust particle. The neutral gas pressure P is equal to 1 Torr and the volume concentration of dust n_d is equal to 10^4 cm^{-3}. The same figure shows for comparison the dependences of the densities of charged particles on the distance, which are given by the Boltzmann distribution. It can be seen that the curves obtained statistically and the theoretical curves are in acceptable agreement. However, to match the curves for ions to a certain extent, we had to take as the distribution parameter the average kinetic energy of ions $\bar{\varepsilon}_i = 1.0$ eV in the immediate vicinity of the macroparticle surface, but not the equilibrium temperature $T_i = 0.025$ eV. This means that the heating of ions in the field of a macroparticle plays a significant role in describing the system dynamics and that the employment of the parameters of the plasma unperturbed by the presence of dust for this purpose is unjustified. In the case of electrons, good agreement between the theoretical curves and the results of numerical experiment was reached using the electron equilibrium temperature $T_e = 2.5$ eV.

One more feature of the plasma particle concentration distribution should also be noted. Comparison of Figs. 7 and 5 shows that the value of the electron concentration at the boundary of the computational region almost coincides with the volume-averaged value represented in Fig. 5, while the ion concentration exceeds the corresponding value in the region unperturbed by the field of the macroparticle approximately by a factor of 1.7. This is not accidental and is due to the fact that the distance from the center of the macroparticle beyond which the plasma can be regarded as unperturbed is much longer than the size of the calculation cell. In this respect, the observed deviation in ion con-

centrations may serve as an additional confirmation of the necessity of introducing a plasma flow correction procedure on the external boundary of the calculation cell into the computational algorithm and also illustrates the effectiveness of its operation.

Figure 8 shows the distributions of the average kinetic energies of ions and electrons under the same conditions. These data, as well as the concentration distributions, were collected statistically over a certain period of time during which the system had already been in dynamic equilibrium. In this procedure, the arithmetic mean values of the kinetic energy of particles located at the same distance from the center of the calculation cell were calculated for every time selection. The figure shows clearly that the energy of ions increases, while the energy of electrons decreases with decreasing distance to the macroparticle surface. We also obtained the values of energies of ions and electrons absorbed at the surface of macroparticles. For the same parameters of the system, these energies are $\bar{\varepsilon}_{ed}$ = 5.12 eV and $\bar{\varepsilon}_{id}$ = 1.05 eV. It can be seen that the value of $\bar{\varepsilon}_{ed}$ for electrons is several times larger than the value of the averaged kinetic energy directly at the macroparticle surface. This means that the absorption at the surface of macroparticles strongly affects the energy distribution function for particles that remain in the bulk. On the whole, the electron energy spectrum is noticeably "cooled" relative to the case of dust absence. Since the charge accumulated by dust decreases upon an increase in the amount of dust in the bulk, electrons with lower energies can overcome the potential barrier. This means that the effect of the electron distribution function cooling upon an increase in the dust concentration in the bulk must be manifested more strongly. This effect was indeed observed in our calculations and was represented in the gradual increase in the distance between the boundary of the region in which the averaged electron kinetic energy differs noticeably from its unperturbed value and the macroparticle surface upon an increase in n_d. Thus, the inclusion of this factor is actually required in the description of the plasma kinetics and treatment of the transport processes in the presence of dust particles even for comparatively low dust concentrations.

On the whole, the analysis of the results obtained here indicates that dynamic equilibrium is established in the system via the attainment of a complex balance between various processes, and the ways in which the system reaches equilibrium may differ significantly depending on the combination of external factors. The self-consistent approach developed by us here proves to be a suitable tool for solving this type of problems, but requires a further refinement for obtaining a more comprehensive and adequate description. For example, the evolution of the energy distribution for plasma particles must obviously be included along with other dynamic processes considered in this connection. This in turn might change the cross sections of interactions and reaction rates in a multicomponent plasma. Thus, further investigations in this field are of profound interest, and we plan to work in this direction.

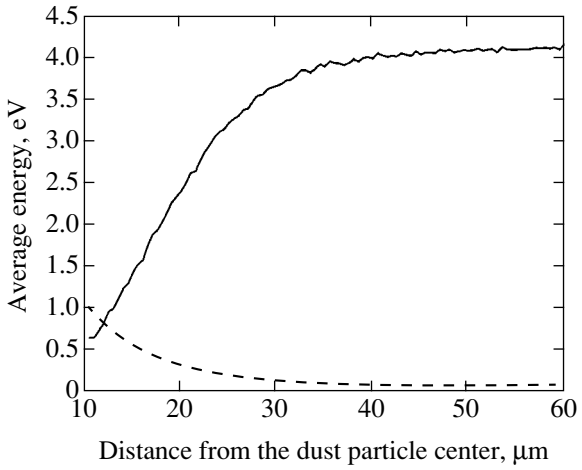

Fig. 8. Dependence of the average kinetic energies of electrons (solid curve) and ions (dashed curve) in the vicinity of a dust particle on the distance from its center; $n_d = 10^4$ cm^{-3}, P = 1 Torr.

ACKNOWLEDGMENTS

This study was partly supported by the Russian Foundation for Basic Research (project nos. 01-02-17726 and 00-15-96554) and by NATO SFP 974354.

REFERENCES

1. J. H. Chu, J. B. Du, and I. Lin, J. Phys. D **27**, 296 (1994).
2. J. H. Chu and I. Lin, Phys. Rev. Lett. **72**, 4009 (1994).
3. H. Thomas, G. E. Morfill, V. Demmel, *et al.*, Phys. Rev. Lett. **73**, 652 (1994).
4. A. Melzer, T. Trottenberg, and A. Piel, Phys. Lett. A **191**, 301 (1994).
5. J. B. Pieper, J. Goree, and R. A. Quinn, J. Vac. Sci. Technol. A **14** (2), 519 (1996).
6. A. P. Nefedov, O. F. Petrov, and V. E. Fortov, Usp. Fiz. Nauk **167**, 1215 (1997) [Phys.–Usp. **40**, 1163 (1997)].
7. M. Zuzic, A. V. Ivlev, J. Goree, *et al.*, Phys. Rev. Lett. **85**, 4064 (2000).
8. H. H. Hwang and M. J. Kushner, J. Appl. Phys. **82**, 2106 (1997).
9. G. Praburam and J. Goree, Astrophys. J. **441**, 830 (1995).
10. D. Samsonov and J. Goree, J. Vac. Sci. Technol. A **17**, 2835 (1999).
11. D. Samsonov and J. Goree, Phys. Rev. E **59**, 1047 (1999).
12. G. E. Morfill, H. M. Thomas, U. Konopka, *et al.*, Phys. Rev. Lett. **83**, 1598 (1999).
13. J. Goree, G. E. Morfill, V. N. Tsytovich, and S. V. Vladimirov, Phys. Rev. E **59**, 7055 (1999).

14. O. S. Vaulina, A. P. Nefedov, O. F. Petrov, and V. E. Fortov, Zh. Éksp. Teor. Fiz. **118**, 1325 (2000) [JETP **91**, 1147 (2000)].
15. J. Goree, Plasma Sources Sci. Technol. **3**, 400 (1994).
16. V. N. Tsytovich, Usp. Fiz. Nauk **167**, 57 (1997) [Phys.–Usp. **40**, 53 (1997)].
17. Yu. A. Mankelevich, M. A. Olevanov, and T. V. Rakhimova, Zh. Éksp. Teor. Fiz. **121**, 1288 (2002) [JETP **94**, 1106 (2002)].
18. V. A. Schweigert, I. V. Schweigert, A. Melzer, *et al.*, Phys. Rev. Lett. **80**, 5345 (1998).
19. V. V. Ivanov, A. F. Pal', T. V. Rakhimova, *et al.*, Zh. Éksp. Teor. Fiz. **115**, 2020 (1999) [JETP **88**, 1105 (1999)].
20. A. F. Pal', A. O. Serov, A. N. Starostin, *et al.*, Zh. Éksp. Teor. Fiz. **119**, 272 (2001) [JETP **92**, 235 (2001)].
21. Ya. K. Khodataev, R. Binghem, V. P. Tarakanov, *et al.*, Fiz. Plazmy (Moscow) **22**, 1028 (1996).
22. A. V. Zobnin, A. P. Nefedov, V. A. Sinel'shchikov, and V. E. Fortov, Zh. Éksp. Teor. Fiz. **118**, 554 (2000) [JETP **91**, 483 (2000)].
23. S. A. Maiorov, S. V. Vladimirov, and N. F. Cramer, Phys. Rev. E **63**, 017401 (2000).
24. S. A. Maiorov, S. V. Vladimirov, and N. F. Cramer, Phys. Rev. E **63**, 045401 (2001).
25. T. J. Sommer, M. S. Barnes, J. H. Keller, *et al.*, Appl. Phys. Lett. **59**, 638 (1991).
26. D. W. Heermann, *Computer Simulation Methods in Theoretical Physics*, 2nd ed. (Springer, New York, 1986; Nauka, Moscow, 1990).
27. Yu. P. Raĭzer, *The Physics of Gas Discharge* (Nauka, Moscow, 1987).
28. A. Barkan, N. D'Angelo, and R. I. Merlino, Phys. Rev. Lett. **73**, 3093 (1994).

Translated by N. Wadhwa

Journal of Experimental and Theoretical Physics, Vol. 96, No. 3, 2003, pp. 457–464.
From Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 123, No. 3, 2003, pp. 518–526.

PLASMA, GASES

Measurement of the Temperature Dependence of the *dd*μ Molecule Formation Rate in Dense Deuterium at Temperatures of 85–790 K¶

V. R. Bom[a], D. L. Demin[b], C. W. E. van Eijk[a], V. V. Filchenkov[b], N. N. Grafov[b,*], V. G. Grebinnik[b], K. I. Gritsaj[b], A. D. Konin[b], A. V. Kuryakin[c], V. A. Nazarov[c], V. V. Perevozchikov[c], A. I. Rudenko[b], S. M. Sadetsky[d], Yu. I. Vinogradov[c], A. A. Yukhimchuk[c], S. A. Yukhimchuk[b], V. G. Zinov[b], and S. V. Zlatoustovskii[c]

[a]*Delft University of Technology, 2629 JB Delft, the Netherlands*
[b]*Joint Institute for Nuclear Research, Dzhelepov Laboratory of Nuclear Problems, Dubna, Moscow oblast, 141980 Russia*
**e-mail: grafov@nu.jinr.ru*
[c]*Russian Federal Nuclear Center, All-Russia Research Institute of Experimental Physics, Sarov, Nizhni Novgorod oblast, 607200 Russia*
[d]*St. Petersburg Nuclear Physics Institute, Gatchina, Leningrad oblast, 188350 Russia*

Received July 29, 2002

Abstract—Muon catalyzed fusion (MCF) in deuterium was studied by the MCF collaboration on the Joint Institute for Nuclear Research Phasotron. The measurements were carried out with a high-pressure deuterium target in the temperature range 85–790 K at densities of about 0.5 and 0.8 of the liquid hydrogen density. The first experimental results for the *dd*μ molecule formation rate $\lambda_{dd\mu}$ in the temperature range 400–790 K with a deuterium density of about 0.5 of the liquid hydrogen density are presented. © *2003 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

The processes of muon catalyzed fusion (MCF) in deuterium have been studied at many laboratories [1–10]. Recent years have seen significant success both in theoretical consideration [11–14] and in the measurements of the *dd*μ molecule formation rate including strong spin effects. Recent data [9, 10] give an example of the progress achieved in the accuracy of measuring the fine effects of MCF in low-density deuterium gas at temperatures of 28–350 K. However, there are still no direct experimental data for the MCF processes at temperatures above 400 K. The only experimental (nondirect) result for $\lambda_{dd\mu}$ at temperatures up to 600 K [2] was obtained from the analysis of muon losses in an experiment with double deuterium–tritium mixtures and had large errors. It is important that at these temperatures, a large *dd*μ formation rate for the *d*μ atom spin state $F = 1/2$ should be expected to prevail over that for the spin $F = 3/2$ [14]. These measurements of $\lambda_{dd\mu}$ at high temperatures are also necessary for correcting the evaluation of the parameters of the *dt* fusion cycle derived from experiments [15].

Here, we present the results of the measurements in a dense deuterium gas in a wide temperature range, 85–790 K. The aim of the experiment was to measure the formation rate $\lambda_{dd\mu}$ in dense deuterium for the first time in a temperature range up to 800 K. The scheme of the process is presented in Fig. 1.

2. EXPERIMENT

The experimental method is similar to that used in [16]. We measured and analyzed the time and charge (deposited energy in a neutron detector) distributions of 2.5 MeV neutrons from the *dd* fusion reactions

$$d\mu + d \longrightarrow dd\mu \longrightarrow {}^3\mathrm{He} + n + \mu, \quad (1)$$

$$d\mu + d \longrightarrow dd\mu \longrightarrow {}^3\mathrm{He}\mu + n. \quad (2)$$

The simplified experimental layout is shown in Fig. 2. The essence of the experimental installation is described in [17]. The installation was mounted on the muon beam line of the Joint Institute for Nuclear Research Phasotron.

2.1. Target

The central part of the installation was a specially constructed deuterium high-pressure target (T) [18]

¶This article was submitted by the authors in English.

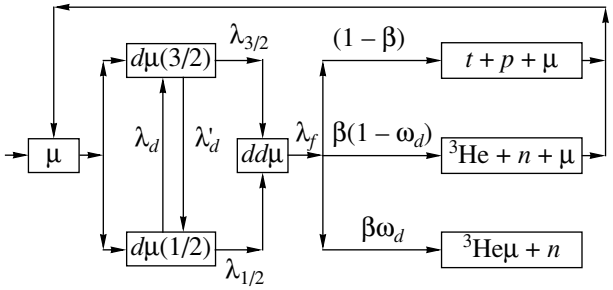

Fig. 1. Scheme of the *dd* cycle.

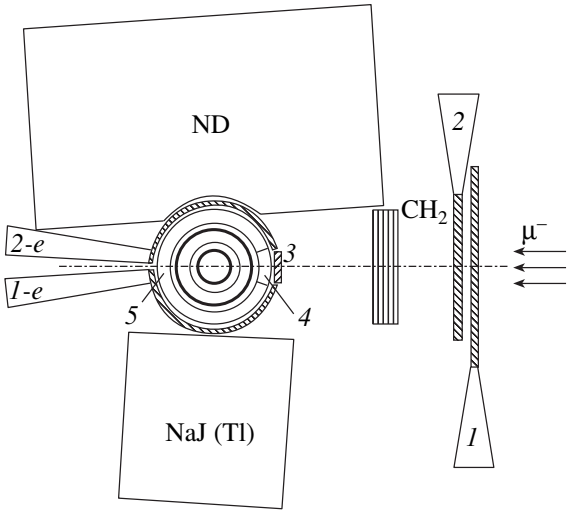

Fig. 2. Experimental layout.

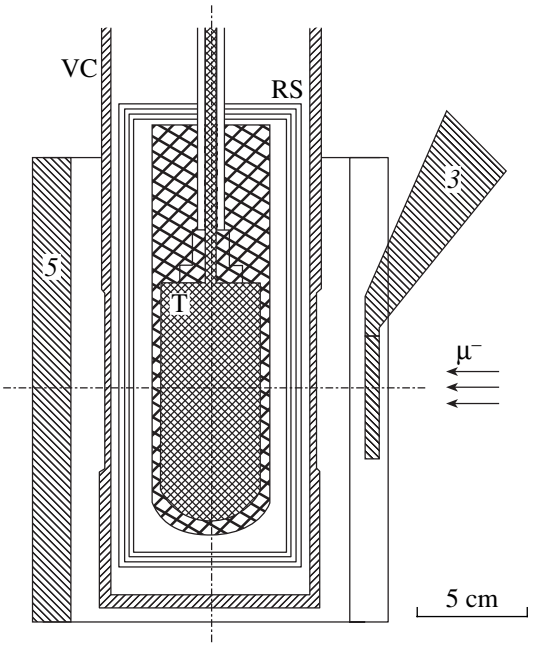

Fig. 3. Scheme of the deuterium high-pressure target. RSs are radiation screens, VC is the vacuum chamber, and 3 and 5 are the scintillation counters depicted in Fig. 2.

with a volume of 76 cm^3 (Fig. 3). The ampoule of the target was made of the special hydrogen-resistant alloy XH40MDTYu-ID. The ampoule was surrounded by a set of radiation screens (RSs) and is placed inside a vacuum chamber (VC) made of stainless steel. The total amount of wall matter on the muon beam path was about 11 mm. The target was able to withstand pressures up to 1500 bar at temperatures up to 800 K. A set of devices [18] was used for handling the deuterium gas. To achieve the deuterium purity of impurities with $Z > 1$ at a level of 0.1 ppm, the target was filled using a palladium filter and a vanadium–deuterium pressure source [19].

2.2. Detectors

The target was surrounded by a set of detectors. Scintillation counters *1–3* detected incoming muons. Cylinder-shaped scintillation counters *4* and *5* served to select muon stops in the target (signal $1 \times 2 \times 3 \times 4 \times \bar{5}$). Specially designed cylinder-shaped scintillation counters *1-e* and *2-e* were used to detect μ-decay electrons. A coincidence between the signals of counters *5* and *1-e, 2-e* was considered as a μ-decay electron. A large neutron detector (with the NE-213 volume 12.5 l) [20] was aimed to detect neutrons from reactions (1) and (2). To reduce the background, the *n*–γ separation was realized by comparing signals for the total light and the fast component light of the neutron detector signal. The γ-quantum discrimination efficiency was better than 10^{-3} for energies larger than 100 keV.

A NaI(Tl) crystal was used to search for the rare fusion channel $d(d, \gamma)^4$He in parallel with the main aim of the experiment (the subject matter of this paper).

The timing sequence of the *4*, *5*, *1-e*, *2-e*, and neutron detector signals was registered by flash ADC and was recorded on a PC. The trigger is described in [22].

2.3. Experimental Conditions

Two exposure runs were carried out. One was with the target filled with an equilibrium protium–deuterium mixture (21% of H) in the temperature range 85–300 K at a density of $\varphi \approx 0.84$ LHD (as usual, density is given in units of the liquid hydrogen density, LHD = 4.25×10^{22} nucl/cm^3). The exposure at $\varphi \approx 0.47$ LHD was also measured at 300 K.

The second exposure run was carried out with the target filled with pure deuterium (protium content was about 0.1%) in the temperature range 300–790 K at a density of $\varphi \approx 0.48$ LHD.

The experimental conditions and the electron statistics gathered for all runs are presented in Table 1. The

Table 1. Experimental conditions and statistics

Run	T, K	Content, %		φ, LHD	N_e
		H	D		
1	85 (1)	20.7 (0.1)	79.3 (0.5)	0.840 (0.025)	685200
2	110 (1)	20.7 (0.1)	79.3 (0.5)	0.841 (0.025)	456500
3	230 (1)	20.7 (0.1)	79.3 (0.5)	0.831 (0.025)	415900
4	301 (3)	20.7 (0.1)	79.3 (0.5)	0.831 (0.025)	427000
5	299 (3)	20.7 (0.1)	79.3 (0.5)	0.473 (0.014)	374900
6	290 (4)	0.1 (0.1)	99.9 (0.1)	0.480 (0.014)	355500
7	401 (10)	0.1 (0.1)	99.9 (0.1)	0.480 (0.014)	226900
8	530 (10)	0.1 (0.1)	99.9 (0.1)	0.480 (0.014)	194900
9	660 (10)	0.1 (0.1)	99.9 (0.1)	0.480 (0.014)	208700
10	791 (15)	0.1 (0.1)	99.9 (0.1)	0.483 (0.020)	301900

background exposure with an empty target was also carried out.

3. KINETICS OF THE $dd\mu$ FUSION CYCLE

The basic processes of the kinetics of the resonant formation of $dd\mu$ molecules are considered in [11, 12]. The $d\mu$ atoms are formed over a time of about $10^{-12}\varphi^{-1}$ s [23] with the energy of a few eV and are thermalized at a rate of about $10^9\varphi$ s^{-1} [24]. Two different hyperfine states of the $d\mu$ atoms $F = 3/2$ and $F = 1/2$ are statistically populated by the respective probabilities 2 : 3 and 1 : 3. After that, processes of resonant $dd\mu$ formation (3) and $d\mu$ atom spin-flip processes (4) occur:

$$d\mu + D_2 \longrightarrow [(dd\mu)dee]^*, \tag{3}$$

$$(d\mu)_{F=3/2} + D_2 \longleftrightarrow (d\mu)_{F=1/2} + D_2. \tag{4}$$

The rates of direct and inverse spin-flip processes (4) (λ_d and λ_d') are connected by the detailed balance relation

$$\lambda_d' = \gamma\lambda_d, \quad \gamma = 2\exp(-\Delta E/T), \quad \Delta E = 0.0485 \text{ eV}. \tag{5}$$

In the process of resonant $dd\mu$ formation (3), the released energy (the binding energy of the $dd\mu$ mesomolecule formed) is transmitted to the excitation of vibration–rotational states of the mesic molecular complex $[(dd\mu)dee]^*$. In the case of the usually nonresonant $dd\mu$ formation,

$$d\mu + D_2 \longrightarrow [(dd\mu)de]^+ + e, \tag{6}$$

the released energy is taken away by a conversion electron. The rate of this nonresonant mechanism (λ_{nr}) is sufficiently small and increases with temperature (energy) as [25]

$$\lambda_{nr} = \lambda_1 + \lambda_2(3/2)kT, \quad \lambda_1 = 0.04\ \mu\text{s}^{-1}, \quad \lambda_2 = 2.3\ \mu\text{s}^{-1}\ \text{eV}^{-1}. \tag{7}$$

In the $dd\mu$ complex formed (3), the processes of dd fusion (with rates of $\lambda_f \sim 10^9$ s^{-1} [26])

$$dd\mu \longrightarrow {}^3\text{He} + n + \mu, \tag{8}$$

$$dd\mu \longrightarrow t + p + \mu, \tag{9}$$

$$dd\mu \longrightarrow {}^3\text{He}\mu + n \tag{10}$$

compete with the processes of complex deexcitation and complex back decay. The $d\mu$ atom formation rate and the thermalization rate, as well as the dd fusion rate, are much higher than the effective hyperfine transition rates and the effective (experimentally observable) mesomolecule formation rate. At times larger than the lifetime (~0.5 ns) of the mesomolecular complex, the kinetics is therefore described by the scheme in Fig. 1 and depends on the effective rates of spin-flip processes (4) and the effective rates of $dd\mu$ formation from two hyperfine states ($\lambda_{1/2}$ and $\lambda_{3/2}$). The following system of differential equations corresponds to the kinetics in Fig. 1 [11, 27]:

$$\frac{dN_{3/2}}{dt} = -(\lambda_0 + \lambda_d + \lambda_{3/2})N_{3/2} + \frac{2}{3}(1-w)\lambda_{3/2}N_{3/2} + \left[\frac{2}{3}(1-w)\lambda_{1/2} + \lambda_d'\right]N_{1/2}, \tag{11}$$

$$\frac{dN_{1/2}}{dt} = -(\lambda_0 + \lambda_d' + \lambda_{1/2})N_{1/2} + \frac{1}{3}(1-w)\lambda_{1/2}N_{1/2} + \left[\frac{1}{3}(1-w)\lambda_{3/2} + \lambda_d\right]N_{3/2}, \quad (12)$$

$$\frac{dN_n}{dt} = \beta\varphi(\lambda_{3/2}N_{3/2} + \lambda_{1/2}N_{1/2}). \quad (13)$$

Here, $N_{3/2}$ and $N_{1/2}$ are populations of the $F = 3/2$ and $F = 1/2$ $d\mu$ atom hyperfine states, N_n is the number of fusion neutrons, $w = \beta w_d$, β is the probability of ^{3}He-reaction channel (8) and (10), $w_d = 0.13$ is the probability of muon sticking to ^{3}He [11], and λ_0 is the muon decay rate.

This system of differential equations (11)–(13) was considered, and its approximate solution was obtained [11, 27].

4. ANALYSIS

The first step of the analysis of registered events consists in separation of neutrons, γ-quanta, and μ-decay electrons. For each exposure, we then build and analyze the time and charge (deposited energy in a neutron detector) distributions of fusion neutrons and the time distributions of μ-decay electrons.

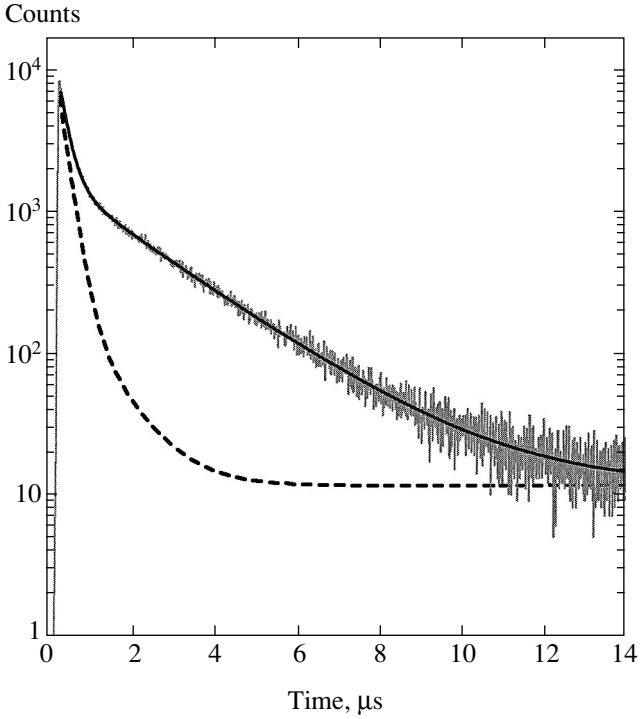

Fig. 4. Example of the experimental time spectrum of electrons from μ decay. Solid line is the fitting function. Dashed line corresponds to the empty target.

4.1. Electron Time Spectra

Time spectra of electrons from muons that stopped and decayed in the target were created and analyzed for each exposure, including the background run with an empty target. Fitting of the background electron time spectra from muons that stopped and decayed in target walls allows us to obtain the form of this distribution, $B_{\text{empty}}(t)$. For each exposure with the deuterium-filled target, we then fitted the electron time spectra, with the obtained form of the background spectra taken into account, using the formula

$$N_e^{\text{total}}(t) = kB_{\text{empty}}(t) + A_e\exp(-\lambda_e t) + F,$$

where λ_e is the muon disappearance rate and F is an accidental background. In this fit, k, A_e, λ_e, and F are parameters. As a result, the numbers of electrons $N_e = A_e/\lambda_e$ from the muons that stopped in deuterium were obtained. The numbers N_e were necessary for normalization. A typical example of the fitted time distributions of decay electrons for the deuterium-filled target is shown in Fig. 4. The dashed line corresponds to the electrons from decays of muons that stopped in the target walls (empty target).

The observed muon disappearance rates λ_e agree with the muon decay rate $\lambda_0 = 0.455\ \mu\text{s}^{-1}$ with an accuracy not worse than 1.5%. Analysis of electron time distributions showed that about 45% of all the incoming muons stop in deuterium. The remaining part of muons stop in the target walls.

4.2. Neutron Time and Charge Spectra

Only the first detected neutrons were selected for analysis. To reduce the background, we used the time selection criterion [28]

$$t_n + 0.5\ \mu s < t_e < t_n + 4.5\ \mu\text{s}, \quad (14)$$

where t_n and t_e are neutron and electron detection times measured from the moment a muon stops in the target. The time and charge (energy) distributions were plotted for these events. Examples of time distributions of fusion neutrons are shown in Fig. 5 for four temperatures.

According to [11, 27], the solution of the system of differential equations (11)–(13), time distributions of the first neutrons detected with the efficiency ϵ, is given by the sum of two exponentials,

$$N_n(t) = A_{\text{fast}}\exp(-\lambda_{\text{fast}}t) + A_{\text{slow}}\exp(-\lambda_{\text{slow}}t). \quad (15)$$

The parameters of this expression are functions of λ_d, $\lambda_{1/2}$, $\lambda_{3/2}$, and $\alpha \equiv \epsilon + \omega - \epsilon\omega$. They can be reconstructed from the amplitudes and slopes of the exponentials in (15). The amplitude A_{fast} of the first term in (15) is related to the $dd\mu$ formation from the $d\mu$ atom with a

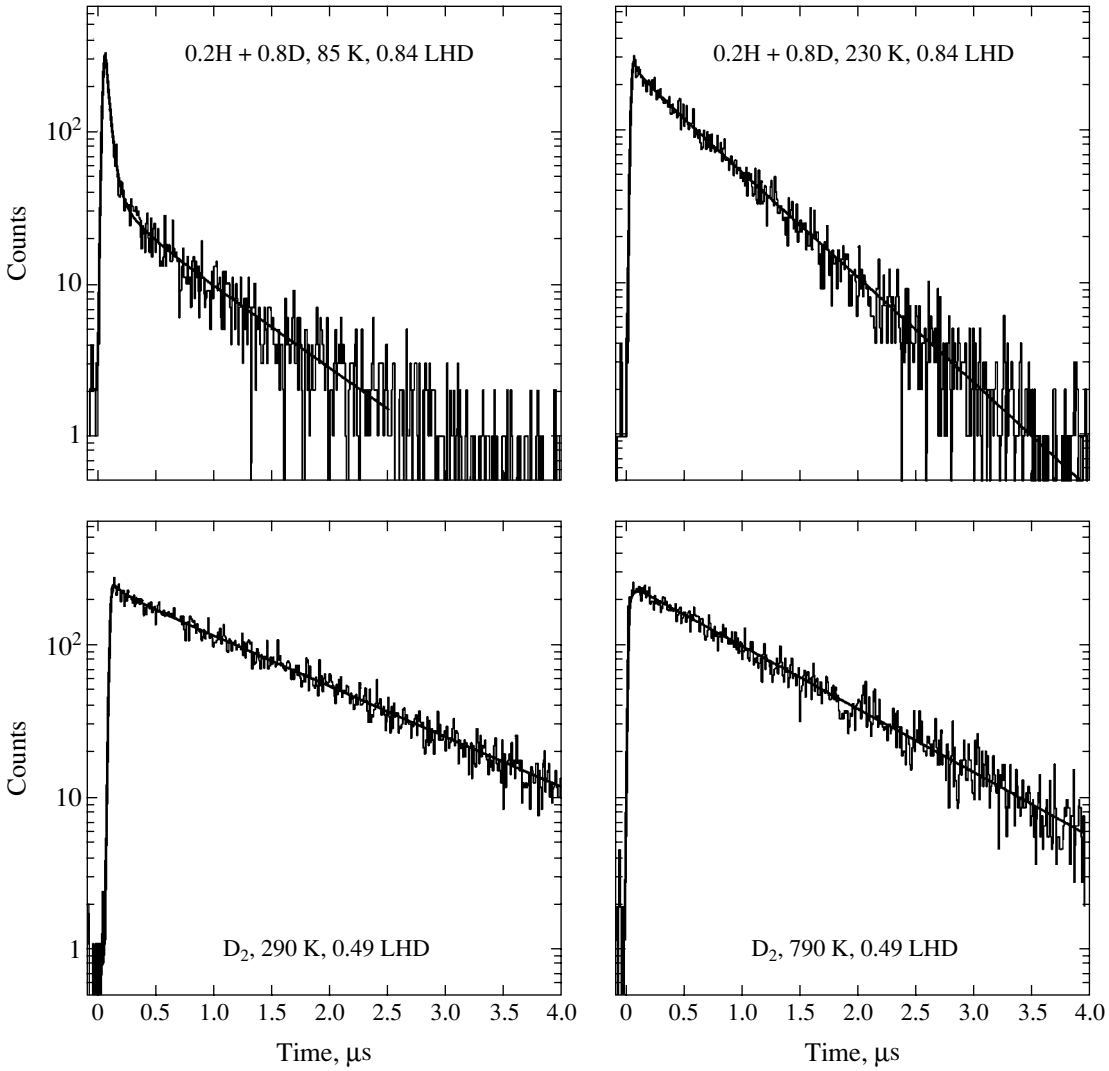

Fig. 5. Experimental time spectra of fusion neutrons measured for exposures at temperatures of 85, 230, 290, and 790 K. Solid lines are the fitting functions.

spin of $F = 3/2$, and the corresponding slope λ_{fast} is determined mainly by the value of λ_d. The amplitude of the second term (slow component) corresponds to the so-called steady state of the $dd\mu$ cycle. The steady-state $dd\mu$ formation rate $\lambda_{dd\mu}^{ss}$ presents the "mean" value of the $dd\mu$ formation rate from different $d\mu$ spin states for times $t > \lambda_{\text{fast}}^{-1}$. The ratio of the amplitudes $A_{\text{fast}}/A_{\text{slow}}$ is responsible for the ratio $\lambda_{3/2}/\lambda_{1/2}$.

The values of $\lambda_{dd\mu}^{ss}$ were obtained from

$$\lambda_{dd\mu}^{ss} = \frac{\lambda_n}{\varphi\beta}\frac{N_n}{N_e\epsilon f_t}. \quad (16)$$

Here, the second ratio gives the absolute neutron yield for the steady state of the dd cycle, λ_n is the slow exponent slope of (15), and N_n is the number of fusion neutrons. The time selection factor $f_t = 0.67$ is due to criterion (14).

Determination of the neutron detection efficiency ϵ is analogous to that in [29]. To determine the efficiency loss due to a finite energy threshold, the calculated recoil proton energy spectrum was reconciled with the experimental energy distribution. This procedure was done for each exposure; an example is given in Fig. 6. The line in the figure is the calculated response function of the neutron detector, and the histogram is the experimental energy distribution. The efficiency of the neutron detector was found to be approximately 12%, and its accuracy is about 8%.

The fitting strategy was as follows. At the first step, we fitted only the slow part of the neutron time spectra and found the steady-state formation rates ($\lambda_{dd\mu}^{ss}$) using expression (16). At the next step, we fitted the entire neutron time spectra using formula (15). In this fit, for-

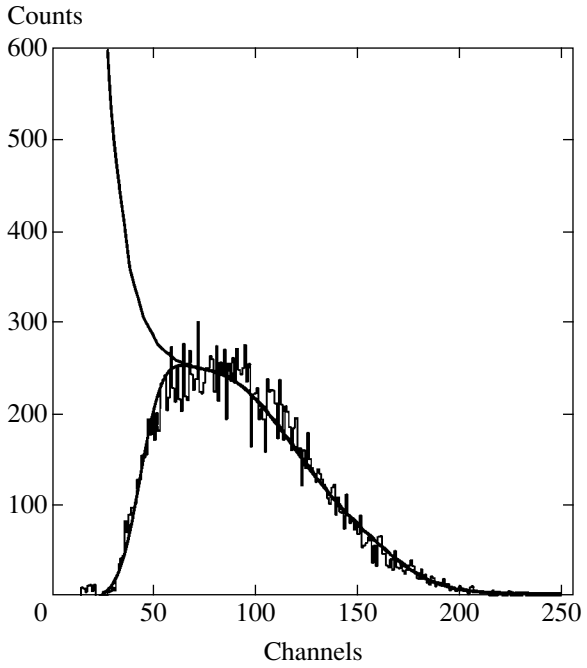

Fig. 6. Neutron energy distribution. The histogram is the experimental neutron energy distribution. Solid line is the calculated response function of the neutron detector.

mula (15) was convolved with a Gaussian resolution function to allow for a finite time resolution and to determine the time zero position. The analysis showed that the time zero stability during data collection was better than 1 ns. The background due to muons stopping in the target walls and due to accidental coincidence was taken into account in the fitting procedure. At the final step of the analysis, the values of $\lambda_{3/2}$, $\lambda_{1/2}$, and λ_d were obtained from the values of $\lambda_{dd\mu}^{ss}$ and the parameters of expression (15). The method developed in [27] was used for this purpose.

Table 2. Experimental results for the steady-state $dd\mu$ formation rate

T, K	$\lambda_{dd\mu}^{ss}$, μs^{-1}		
	Value	Stat. error	Syst. error
85	0.197	0.007	0.018
110	0.485	0.015	0.044
230	2.06	0.06	0.19
301	2.69	0.08	0.24
299	2.57	0.08	0.24
290	2.67	0.08	0.24
401	3.00	0.09	0.27
530	3.10	0.09	0.27
660	2.67	0.08	0.24
791	2.30	0.07	0.21

We made some corrections to the final values of the rates that are due to the nonresonant $dd\mu$ formation rate on molecules D_2 and HD [30] (in runs with double H/D mixtures).

It turned out that a reliable separation of fast and slow terms of (15) in fitting is possible only at low temperatures (in our case, not higher than 290 K), where the rates $\lambda_{3/2}$ and $\lambda_{1/2}$ are substantially different and λ_d is much larger than λ_d'. At high temperatures (higher than 300 K), the rates $\lambda_{3/2}$ and $\lambda_{1/2}$ are approximately equal [14], and it is therefore a large problem during fitting to obtain the fast component parameters and thus to distinguish the values of $\lambda_{3/2}$ and $\lambda_{1/2}$. At these temperatures, we obtained only the value of $\lambda_{dd\mu}^{ss}$. For λ_d, which is determined from the fast exponent slope, we obtained its value only at the temperatures 85 and 110 K because of the same reasons.

5. RESULTS AND DISCUSSION

The obtained values of $\lambda_{dd\mu}^{ss}$, λ_d, $\lambda_{3/2}$, and $\lambda_{1/2}$ are presented in Tables 2 and 3. All rates are normalized to the liquid hydrogen density. The statistical error is 2–3%. The main sources of systematic errors are the uncertainties of the neutron detector efficiency (about 8%) and gas density (about 3%).

In Fig. 7, the temperature dependence $\lambda_{dd\mu}^{ss}(T)$ is shown together with the data [2]. The main result of this work is the direct measurement of the steady-state $dd\mu$ formation rate at temperatures above 400 K. As can be seen, our data are in good agreement with theoretical predictions [14] at low temperatures (up to 400 K). However, the experimental data are slightly larger than the theoretical ones at high temperatures.

In Fig. 8, the temperature dependence of $\lambda_{1/2}$ and $\lambda_{3/2}$ is shown. The solid line represents theoretical calculations [14]. The values of λ_d, $\lambda_{1/2}$, and $\lambda_{3/2}$ are in good agreement with the theory and with other experimental data [9, 10, 17].

As mentioned above, reliably separating fast and slow terms in (15) in fitting, and hence the obtaining the values of λ_d, $\lambda_{1/2}$, and $\lambda_{3/2}$, is possible only at low temperatures under our experimental conditions. The lifetime of the fast component of neutron spectra (15) is

$$\tau = \frac{1}{\lambda_{\text{fast}}} \sim \frac{1}{\lambda_d \varphi}.$$

Under our conditions ($\varphi \approx 0.5$ LHD), this approximately equals 50 ns, or about five channels in the time histogram measured. This fact is also confirmed by our Monte Carlo calculations. It is therefore worthwhile to

Table 3. Experimental results for the effective $dd\mu$ formation rates from different spin states and the effective $d\mu$ spin-flip rate

T, K	$\lambda_{1/2}$, μs^{-1}			$\lambda_{3/2}$, μs^{-1}			λ_d, μs^{-1}		
	Value	Stat. error	Syst. error	Value	Stat. error	Syst. error	Value	Stat. error	Syst. error
85	0.170	0.006	0.015	4.63	0.13	0.42	32.8	1.0	3.2
110	0.403	0.014	0.036	4.79	0.14	0.43	31.4	1.2	3.2
230	1.72	0.06	0.16	3.92	0.12	0.35	–	–	–
290	2.40	0.08	0.22	3.52	0.10	0.32	–	–	–

carry out an experiment at high temperatures with low-density deuterium (about 0.05 LHD), where the spin-effect time range at neutron time distributions is about 0.5 μs. An experiment of this type could be conducted in a muon factory.

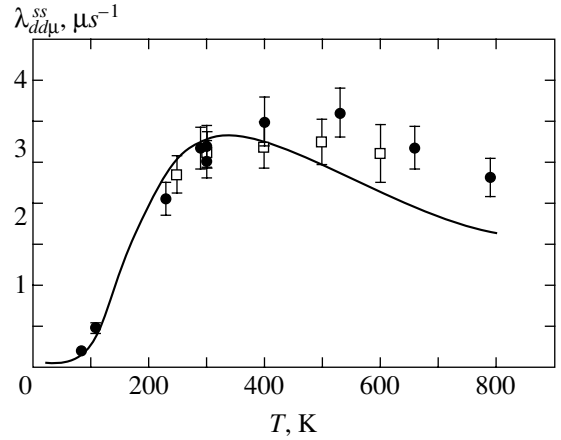

Fig. 7. Values of $\lambda^{ss}_{dd\mu}$ as a function of temperature. Circles are our data, squares are the LAMPF data [2]; the line represents theory [14].

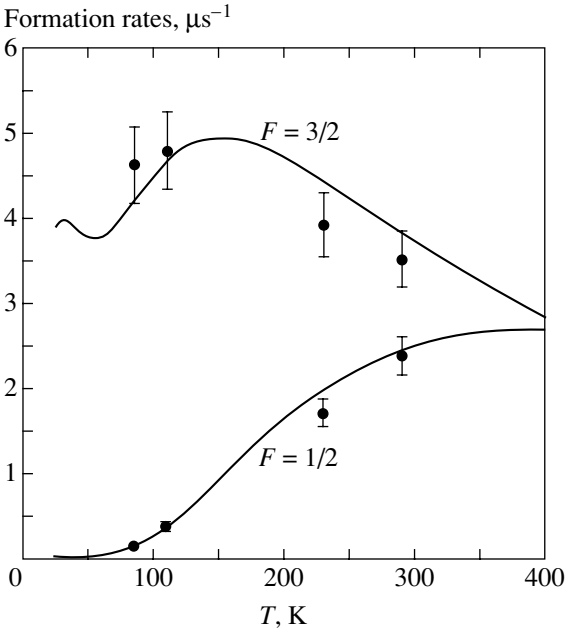

Fig. 8. Values of $\lambda_{3/2}$ and $\lambda_{1/2}$ as a function of temperature; the lines represent theory [14].

6. CONCLUSIONS

The first direct measurements of the muon catalyzed *dd* fusion cycle parameters at high temperatures (up to 790 K) are carried out. The values of the steady-state $dd\mu$ formation rate are obtained for the entire temperature range. Comparison with theory showed good agreement at low temperatures (up to 400 K) and some difference at high temperatures (above 400 K). At low temperatures, the values of the effective rates $\lambda_{3/2}$ and $\lambda_{1/2}$ are obtained; they are in agreement with the theory. Measuring these hyperfine parameters in the high-temperature range seems to require an experiment of a special type, i.e., at a low deuterium density (about (0.01–0.05) LHD) with a high-intensity muon beam.

ACKNOWLEDGMENTS

The authors are grateful to L.I. Ponomarev, M.P. Faifman, and N.I. Voropaev for discussions. We also wish to thank M.M. Petrovsky, A.P. Kustov, and the staff of the RFNC–VNIIEF group for their help in tests and runs. The work was supported by the Ministry of Atomic Energy of the Russian Federation (project no. 6.25.19.19.99.969), the Ministry of Science and Technology of the Russian Federation (project no. 103-7(00)-II), and the Russian Foundation for Basic Research (project no. 01-02-16425).

REFERENCES

1. V. M. Bystritsky *et al.*, Zh. Éksp. Teor. Fiz. **76**, 460 (1979) [Sov. Phys. JETP **49**, 232 (1979)].
2. S. E. Jones *et al.*, Phys. Rev. Lett. **56**, 588 (1986).
3. D. V. Balin *et al.*, Pis'ma Zh. Éksp. Teor. Fiz. **40**, 318 (1984) [JETP Lett. **40**, 1112 (1984)].
4. D. V. Balin *et al.*, Muon Catal. Fusion **2**, 241 (1988).
5. J. Zmeskal *et al.*, Muon Catal. Fusion **1**, 109 (1987).
6. N. Nagele *et al.*, Nucl. Phys. A **493**, 397 (1989).
7. V. M. Bystritsky *et al.*, Muon Catal. Fusion **5**/**6**, 141 (1990/1991).
8. D. V. Balin *et al.*, Muon Catal. Fusion **5**/**6**, 163 (1990/1991).

9. C. Petitjean, D. V. Balin, W. H. Breunlich, *et al.*, Hyperfine Interact. **118**, 127 (1999).
10. N. I. Voropaev *et al.*, Hyperfine Interact. **138**, 331 (2001).
11. M. P. Faifman *et al.*, Zh. Éksp. Teor. Fiz. **92**, 1173 (1987) [Sov. Phys. JETP **65**, 656 (1987)].
12. M. P. Faifman, Muon Catal. Fusion **2**, 247 (1988); M. P. Faifman, L. I. Menshikov, and T. A. Strizh, Muon Catal. Fusion **4**, 1 (1989).
13. A. Scrinzi, Muon Catal. Fusion **5**/**6**, 179 (1990).
14. M. P. Faifman, Hyperfine Interact. **101**/**102**, 179 (1996).
15. V. R. Bom *et al.*, Hyperfine Interact. **138**, 213 (2001).
16. Yu. P. Averin *et al.*, Hyperfine Interact. **118**, 111 (1999).
17. V. P. Dzhelepov *et al.*, Zh. Éksp. Teor. Fiz. **101**, 1105 (1992) [Sov. Phys. JETP **74**, 589 (1992)]; Muon Catal. Fusion **7**, 387 (1992).
18. V. V. Perevozchikov *et al.*, Prib. Tekh. Éksp., No. 4, 155 (2002); Hyperfine Interact. **138**, 417 (2001).
19. A. N. Golubkov and A. A. Yukhimchuk, Hyperfine Interact. **138**, 403 (2001).
20. V. P. Dzhelepov *et al.*, Nucl. Instrum. Methods Phys. Res. A **269**, 634 (1988).
21. L. N. Bogdanova *et al.*, Preprint No. E15–264, JINR (Joint Inst. for Nuclear Research, Dubna, 2001); Yad. Fiz. **65**, 1826 (2002) [Phys. At. Nucl. **65**, 1778 (2002)].
22. V. G. Zinov, A. I. Rudenko, and V. T. Sidorov, Preprint No. P13-96-439, JINR (Joint Inst. for Nuclear Research, Dubna, 1996).
23. J. S. Cohen, Phys. Rev. A **27**, 167 (1983).
24. A. Adamchak and V. S. Melezhik, Muon Catal. Fusion **4**, 303 (1989).
25. S. I. Vinitsky *et al.*, Zh. Éksp. Teor. Fiz. **74**, 849 (1978) [Sov. Phys. JETP **47**, 444 (1978)].
26. L. N. Bogdanova *et al.*, Phys. Lett. B **115**, 171 (1982).
27. V. V. Filchenkov, Preprint No. E1-89-57, JINR (Joint Inst. for Nuclear Research, Dubna, 1989).
28. V. R. Bom *et al.*, Zh. Éksp. Teor. Fiz. **111**, 1163 (1997) [JETP **84**, 641 (1997)].
29. V. R. Bom and V. V. Filchenkov, Hyperfine Interact. **118**, 365 (1999).
30. G. G. Semenchuk *et al.*, Hyperfine Interact. **118**, 141 (1999).

Journal of Experimental and Theoretical Physics, Vol. 96, No. 3, 2003, pp. 465–479.
From Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 123, No. 3, 2003, pp. 527–542.

PLASMA, GASES

Internal Energy of High-Density Hydrogen: Analytic Approximations Compared with Path Integral Monte Carlo Calculations¶

S. A. Trigger[a], W. Ebeling[a], V. S. Filinov[b,*], V. E. Fortov[b], and M. Bonitz[c,]**

[a]*Institut für Physik, Humboldt-Universität Berlin, D-10115 Berlin, Germany*
[b]*Institute for High Energy Density, Russian Academy of Sciences, Moscow, 127412 Russia*
**e-mail: filinov@ok.ru*
[c]*Fachbereich Physik, Universität Rostock, D-18051 Rostock, Germany*
***e-mail: bonitz@physik.uni-rostock.de*

Received September 30, 2002

Abstract—The internal energy of high-density hydrogen plasmas in the temperature range $T = 10000$–50000 K is calculated by two different analytic approximation schemes—the method of an effective ion–ion interaction potential and the Padé approach within the chemical picture—and are compared with direct path integral Monte Carlo results. A reasonable agreement between the results obtained from the three independent calculations is found and the reasons for still existing differences are investigated. Interesting high-density phenomena such as the onset of ion crystallization are discussed. © *2003 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

The thermodynamics of strongly correlated Fermi systems at high pressure is of growing importance in many fields, including shock and laser plasmas, astrophysics, solids, and nuclear matter (see [1–6] for an overview). In particular, thermodynamic properties of hot dense plasmas are essential for the description of plasmas generated by strong lasers [5]. Further, among the phenomena of current interest are the high-pressure compressibility of deuterium [7], metallization of hydrogen [8], plasma phase transition, etc., which occur in situations where both interaction and quantum effects are relevant. Among the early theoretical papers on dense hydrogen, we refer to Wigner and Huntington [9], Abrikosov [10], Ashcroft [11], and Brovman *et al.* [12]; concerning the plasma phase transition, see Norman and Starostin [13], Kremp *et al.* [14], Saumon and Chabrier [15], and Schlanges *et al.* [16], and also some earlier investigations of one of us [17–20]. Among the early simulation approaches, we refer to several Monte Carlo calculations, e.g., [21–23].

There has been a significant progress in recent years in studying these systems analytically and numerically (see, e.g., [1, 2, 4, 24–28] for an overview), but there remains an urgent need to test analytic models by an independent numerical approach. In addition to the molecular dynamics approach, e.g., [24, 26], the path integral Monte Carlo (PIMC) method is particularly well suited to describe thermodynamic properties in the high-density region. This is because it starts from the fundamental plasma particles—electrons and ions (physical picture)—and treats all interactions, including bound state formation, rigorously and self-consistently. We note a remarkable recent progress in applying these techniques to Fermi systems (see, e.g., [1, 2, 29, 30] for an overview).

Several methods have been developed to perform quantum Monte Carlo calculations. We first mention the restricted PIMC method (RPIMC) [31–34], where special assumptions on the density operator $\hat{\rho}$ are introduced in order to reduce the sum over permutations to even (positive) contributions only. It can be shown, however, that this method does not reproduce the correct ideal Fermi gas limit [35]. An alternative is given by direct fermionic PIMC simulations (DPIMC), which have occasionally been attempted by various groups (see, e.g., [36, 37] and references therein). However, these simulations have been very inefficient because of the fermionic sign problem. Recently, three of us proposed a new path integral representation for the N-particle density operator [38–41] that allows direct fermionic PIMC simulations of dense plasmas in a wide range of densities and temperatures. Using this concept, the pressure and energy of a degenerate strongly coupled hydrogen plasma [39–42] and the pair distribution functions in the partial ionization and dissociation region [40, 41] have been computed. This scheme is rather efficient when the number of time slices (beads) in the path integral is less than or equal to 50 and was found to work well for temperatures $k_BT > 0.1$ Ry.

¶This article was submitted by the authors in English.

One difficulty of PIMC simulations is that reliable error estimates are often not available, in particular for strongly coupled degenerate systems. Here, we make a comparison of two independent analytic methods. The first is the method of an effective ion–ion interaction potential (EIIP) that has previously been developed for application to simple solid and liquid metals [12, 24] and which is here adapted to dense hydrogen for the first time. The second is the method of Padé approximations in combination with Saha equations, i.e., the chemical picture (PACH) [3]. The Padé formulas are constructed on the basis of the known analytic low-density [3, 43] and high-density [3] limits and are exact up to quadratic terms in the density, interpolating between the virial expansions and the high-density asymptotic regime [19, 44, 45].

We show here that both methods, EIIP and PACH, provide results for the internal energy that agree well with each other at high densities where the electrons are strongly degenerate and no bound states exist, approximately for $n > 10^{24}$ cm^{-3}. In this region, there is also a good agreement with recent density functional results [46]. The agreement of the PACH and DPIMC results is good below 10^{22} cm^{-3}. For intermediate densities, where the ionization degree changes strongly, we observe deviations. Also, at high densities, the DPIMC results tend toward lower energies than the analytic approaches. Finally, they reveal several interesting effects, such as formation of clusters and the onset of ion crystallization.

2. PHYSICAL PARAMETERS AND BASIC EFFECTS

We study a hydrogen plasma consisting of N_e electrons and N_p protons ($N_e = N_p = N$). The total proton (atom) density is $n = N_p/V$. The average distance between the electrons is the Wigner–Seitz radius

$$d = \left(\frac{3}{4\pi n}\right)^{1/3},$$

and other characteristic lengths are the Bohr radius

$$a_B = \frac{\hbar^2}{me^2},$$

the Landau length

$$l = \frac{e^2}{kT},$$

and the De Broglie wavelength

$$\Lambda_e = \frac{h}{(2\pi m_e kT)^{1/2}}$$

of the electrons. The degeneracy parameter is $n\Lambda_e^3$. We define the dimensionless temperature $\tau = kT/\mathrm{Ry}$, which varies between $0.06 < \tau < 0.4$ in the temperature interval considered below. We also introduce the Wigner–Seitz parameter

$$r_s = \frac{d}{a_B}$$

and the dimensionless classical coupling strength

$$\Gamma = \frac{e^2}{kTd}.$$

Hydrogen is antisymmetric with respect to the charges ($e_- = -e_+$) and symmetric with respect to the densities ($n_+ = n_- = n$). Ions and electrons behave quite differently because of the big mass difference, $m_p = 1836m_e$. At the temperatures considered, the ions can be treated classically as long as $n \lesssim 10^{27}$ cm^{-3}. For these temperatures and densities, the proton coupling parameter is in the range $0 < \Gamma < 150$, and we can therefore expect strong coupling effects. We study internal energies of the fluid hydrogen system and start with providing some simple estimates for guidance. In what follows, we give all energies in Rydberg units.

First, at very low densities, the electrons and protons behave as an ideal Boltzmann gas. Therefore, the energy (of free electrons and protons) per proton is given by

$$\epsilon = E/N = 3\tau. \tag{1}$$

In other words, the low-density limit is, in our temperature interval, a positive number in the region $\epsilon \approx 0.2$–1.2. With increasing density, we expect a region where atoms and possibly also a few molecules are formed [17, 41]. In the region of atoms, the lower bound for the energy per proton is

$$\epsilon = \frac{3}{2}\tau - 1, \tag{2}$$

where the last term represents the binding energy 1 Ry of H atoms. If molecules are formed, the corresponding estimate per proton is lower,

$$\epsilon = \frac{3}{4}\tau - 1.17. \tag{3}$$

Generally, the existence of a lower bound for the energy per proton was proven by Dyson and Lenard [47] and Lieb and Thirring [48],

$$E/N > -C, \tag{4}$$

where the best estimate known to us (which is certainly much too large) is $C \approx 23$ [48]. We see that with increasing density, the energy per proton tends to negative values and can reach a finite minimum. Further density

increase causes the energy to increase again as a result of quantum degeneracy effects.

To understand this increase, we first consider the limit of a very high density (but in the region where the protons are classical). The first estimate of the energy is then

$$\epsilon = \frac{3}{2}\tau + \frac{2.21}{r_s^2}, \tag{5}$$

which is positive. The last term, representing the Fermi energy of the electrons, strongly increases with the density (as $n^{2/3}$). In the next approximation, according to Wigner's estimate [49],[1] we must take into account the Hartree contribution to the electron energy and the corresponding estimate for the proton energy. The latter is estimated under the assumption that the protons form a lattice. In this way, we find the estimate

$$\epsilon = \left(\frac{3}{2}\tau - \frac{0.8755}{r_s}\right) + \left(\frac{2.21}{r_s^2} - \frac{0.916}{r_s}\right). \tag{6}$$

The two corrections that were added to Eq. (5) are both negative and scale as $n^{1/3}$. In other words, these interaction terms might play a major role with decreasing density. At a critical density, the energy per proton can become negative. This density can be estimated from Eq. (6) by solving the quadratic equation

$$0 = \frac{3}{2}\tau r_s^2 - 1.7915 r_s + 2.21 \tag{7}$$

perturbatively, starting with the zero temperature limit, and adding the first correction (linear in τ),

$$r_s^0 \approx 1.234 + 2.283\tau + \ldots \tag{8}$$

As $\tau \longrightarrow 0$, this result coincides with Wigner's criterion for the existence of molecules: for $d < a_B$, molecules cannot exist because there is no room for forming bound state wave functions. According to Eq. (8), molecules exist at a finite temperature only for larger d as thermal fluctuations increase the wave function overlap. More generally, with increasing temperature, the energy becomes positive at lower density compared to the case where $T = 0$.

Summarizing the qualitative results obtained in this section, we can state that we expect the following general behavior of the internal energy per proton in the given temperature range: at zero density, the energy starts with an ideal gas expression that depends only on temperature. With increasing density, the energy per proton becomes negative because of correlation effects (bound states, electron correlations, and proton correlations). A minimum is formed, and at the density where the proton density is close to the inverse Bohr radius cubed, the energy per proton turns to positive values and is more and more determined by the ideal electron energy increasing with $n^{2/3}$, corrected by correlation contributions of the order $n^{1/3}$ determined by the Hartree term and by proton–proton coupling effects. In what follows, we show that this qualitative picture is supported by the results of our calculations.

3. THE METHOD OF AN EFFECTIVE ION–ION INTERACTION POTENTIAL

It is well known that in plasmas and plasmalike systems, in a broad parameter range, the interaction between the electron and ion subsystems is weak, whereas the interactions within the electron and ion subsystems can be strong. The corresponding small parameter is the ratio u_{ei}/E_F of the characteristic value of the electron–ion interaction u_{ei} to the electron Fermi energy E_F. Therefore, the approximation of a small u_{ei}/E_F ratio is valid for systems with degenerate electrons if $E_F \gg T_e \geq T_i$, where T_e and T_i are the electron and ion temperatures, respectively (below, we consider the case where $T_e = T_i$). Typical systems where this approximation is applicable are simple solid and liquid metals and nontransitional metals in general; this approximation serves as a basis for the computation of thermodynamic and electron kinetic properties (see, e.g., [24, 50]).

For simple metals, the Fermi energy is not very large compared with the characteristic electron–ion Coulomb interaction taken at the average interparticle distance. However, because the wave functions for the conduction electrons and the electrons bound in the ion shells are orthogonal, a partial compensation of the electron–ion Coulomb attraction occurs at small distances, which effectively weakens the electron–ion interaction. This fact is described in the theory of simple metals in the framework of the so-called pseudopotential theory. The calculation of the pseudopotential is a complicated problem in general, in particular due to its nonlocal structure [50, 51]. For practical applications, it can be represented approximately as a local interaction with one or two fitting parameters for each metal. On the basis of the pseudopotential theory, all thermodynamic properties and electron kinetic coefficients can be calculated with a sufficiently high accuracy for a wide range of temperatures and pressures. Naturally, these calculations require a reliable knowledge of the properties of the two quasi-independent subsystems: the degenerate electron liquid in the positive charge background and the classical ion subsystem with some effective strong interion interaction.

It is apparent that there is also a wide range of parameters for highly ionized, strongly compressed

[1] Wigner's original estimate for the lattice energy was corrected later, and we use an improved result. For a discussion of various estimates, see, e.g., Chapter V of G. D. Mahan, *Many-Particle Physics*, Plenum Press (1990).

hydrogen plasmas where the electron–ion interaction is weak. For these parameters, the complicated problem of calculating the properties of a strongly coupled quantum electron–proton system can be substantially simplified. In so doing, the results obtained for high compression (when no bound electron states—hydrogen atoms or molecules—exist) do not require any fitting, in contrast to the case of simple metals, because the interion potential for hydrogen is a purely Coulomb one. The data obtained with this analytic approximation can therefore be considered as a reliable basis for comparison with the results of alternative approaches, including analytic and simulation methods for degenerate quantum systems of Fermi particles. The results of this pseudopotential approach are especially important for conditions of extreme compression where the plasma is characterized by a strong interaction within the electron and especially the ion subsystem. For these difficult situations, experimental data are still missing and new accurate numerical methods for Fermi systems are only just emerging.

We consider the Hamiltonian of an electron–proton plasma, where the $q = 0$ infinite contributions to the potentials cancel because of quasineutrality (and we retain the charge number Z of the ions for generality),

$$H = \sum_k \epsilon_k a_k^\dagger a_k + \frac{1}{2V}\sum_{k,k',q\neq 0}\frac{4\pi e^2}{q^2}a_{k-q}^\dagger a_{k'+q}^\dagger a_{k'}a_k + \frac{1}{V}\sum_{k,q'\neq 0} u_{ei}(q)a_k^\dagger a_{k+q}\sum_{j=1}^{N_i}\exp(i\mathbf{q}\cdot\mathbf{R}_j) + \frac{1}{2V}\sum_{i\neq j}\sum_{q\neq 0}\frac{4\pi Z^2 e^2}{q^2}\exp(i\mathbf{q}\cdot(\mathbf{R}_i-\mathbf{R}_j)) + K_i. \quad (9)$$

Here, ϵ_k is the energy of the electron with the momentum $\hbar k$ and

$$u_{ei}(q) = -\frac{4\pi Z e^2}{q^2}$$

is the Fourier component of the electron–proton interaction potential. For the electron degrees of freedom in the Hamiltonian H, the secondary quantization representation is used, with $a_p^\dagger$ and a_p being the respective creation and annihilation operators of an electron with momentum p. For classical ions, the coordinate representation is more convenient, and R_i therefore denotes the coordinate of the ith ion in Eq. (9). As in the theory of simple metals [12, 24], two main approximations have to be used to calculate the plasma energy. The first is the adiabatic approximation for the ion motion, which is slow compared to the electron one. The second is the smallness of the ratio of the characteristic electron–proton Coulomb interaction to the Fermi energy E_F. The respective parameter is

$$\Gamma_{ei} = \frac{Ze^2}{dE_F} = Z\Gamma\frac{kT}{E_F} \propto n^{-1/3}.$$

Calculation of the electron energy in the external field of immobile ions (protons) leads to the energy of the plasma given as a function of the ion coordinates R_j. In general, perturbation theory in terms of the parameter Γ_{ei} gives rise not only to pair but also to higher order ion–ion interactions, which are quite complicated. To the second order of perturbation theory in the parameter Γ_{ei}, the energy per one electron of the plasma with a fixed proton configuration $\{R_j\}$ is easily written as

$$\frac{E(\{R_j\})}{N_i} = \frac{\langle H\rangle_e}{N_i} = \epsilon_e + \frac{3}{2}kT - \frac{1}{2}\int\frac{d^3q}{(2\pi)^3}\frac{u_{ei}^2(q)\Pi_e(q)}{\varepsilon_e(q)} - \frac{Z^2 n_i}{2\Pi_e(q=0)} + \frac{1}{2VN_i}\sum_q\sum_{i\neq j}\mathcal{V}^{\text{eff}}(q)\exp(i\mathbf{q}\cdot(\mathbf{R}_i-\mathbf{R}_j)), \quad (10)$$

where ϵ_e is the energy (per ion) of the correlated electron liquid in the homogeneous positive charge background. The respective functions $\Pi_e(q)$ and $\varepsilon_e(q)$ are the static polarization function and the static dielectric function of the correlated electron liquid. They are related by the usual equality

$$\varepsilon_e(q) = 1 + \frac{4\pi e^2}{q^2}\Pi_e(q). \quad (11)$$

The Fourier component of the effective pair interaction potential between the ions, $\mathcal{V}_{ii}^{\text{eff}}$, involved in (10) is given by

$$\mathcal{V}_{ii}^{\text{eff}}(q) = \frac{4\pi Z^2 e^2}{q^2} - u_{ei}^2(q)\frac{\Pi_e(q)}{\varepsilon_e(q)} = \frac{4\pi Z^2 e^2}{q^2\varepsilon_e(q)}. \quad (12)$$

In what follows, we concentrate on hydrogen and set $Z = 1$, which leads to the effective proton–proton interaction

$$\mathcal{V}_{pp}^{\text{eff}}(q) = \frac{4\pi e^2}{q^2\varepsilon_e(q)}. \quad (13)$$

It is clear that in contrast to liquid metals, where the presence of the pseudopotential leads to a more complicated structure of the effective potential, in a dense hydrogen plasma, the effective potential is determined only by the electron screening. As shown in [12] for liquid metals, the additional pair interaction arising from

third- and fourth-order terms in the expansion of the electron energy in the pseudopotential can play an important role in the effective interaction. A detailed analysis of the effective potential of a hydrogen plasma [52] revealed that these terms are essential only for sufficiently rarified plasma conditions ($r_s > 1.5$) and are practically negligible for higher densities, $r_s < 1.5$, which we consider in this paper. In fact, for $r_s > 1.6$, the structure of the effective ion–ion potential in hydrogen changes drastically and can be considered as a precursor of the appearance of molecular states. In this paper, we use the simplest version of the method of the effective ion–ion potential that includes the electron–proton interaction up to the second order, and we are therefore restricted to sufficiently high densities corresponding to $r_s < 1.5$.

Further progress can be made using the random phase approximation (RPA) for Π_e together with the long-wavelength and short-wavelength limits,

$$\Pi^{RPA}(q) = \begin{cases} \Pi^{RPA}(0)\left[1 - \frac{1}{12}\frac{q^2}{q_F^2}\right], & q \ll q_F, \\ \Pi^{RPA}(0)\frac{4}{3}\frac{q_F^2}{q^2}, & q \gg q_F, \end{cases} \tag{14}$$

where $\hbar q_F = \sqrt{2m\epsilon_F}$ is the Fermi momentum of the electrons. The analysis of this expression shows that the main contribution to energy (10) comes from small wave numbers.[2] With sufficient accuracy, we can therefore neglect the q dependence of Π_e in Eq. (10), and in particular, in effective potential (12), replacing

$$\Pi^{RPA}(q) \longrightarrow \Pi^{RPA}(0).$$

This implies that we also neglect the well-known small oscillations of the effective potential for large distances, which are the result of a logarithmic singularity of the derivative

$$\left.\frac{d\Pi^{RPA}}{dq}\right|_{q = 2q_F}.$$

For the densities under consideration (which are much higher than the usual metallic densities), these oscillations are not essential for the thermodynamic functions.

[2] As shown by [52], for hydrogen at $r_s < 1.6$ (to which we apply the EEIP method), the contribution of nonzero wavenumbers is comparatively small. For $r_s > 1.65$, however, the situation starts to change drastically.

On the other hand, it is crucial to calculate the polarization function $\Pi_e(0)$ fully self-consistently,

$$\Pi_e(0) = \left(\frac{\partial n}{\partial \mu_e}\right)_{T,V}, \quad \mu_e = \left(\frac{\partial n\epsilon_e}{\partial n}\right)_{T,V}, \tag{15}$$

where ϵ_e is determined by (10) and consequently takes the electron–electron exchange and correlations into account. In the case of degenerate electrons, we can use one of the analytic approximations for ϵ_e such as that of Nozières and Pines or Wigner (see, e.g., [53] for an overview). Below, we use Wigner's formula for the correlation energy, although the approximation of Nozierès and Pines is better for small r_s (in the region of $r_s < 1$, where the deviations between these approximations for the correlation energy become substantial, we can completely neglect correlations in comparison to the kinetic and exchange terms). Because

$$\Pi_e^{RPA}(0) = \frac{\kappa_{TF}^2}{4\pi e^2},$$

it is clear that Eq. (15) implies a renormalization $\Pi_e^{RPA} \longrightarrow \Pi_e$ due to the electron–electron interaction, and therefore, a renormalization of the momentum $\kappa_{TF} \longrightarrow \tilde{\kappa}_{TF}$,

$$\Pi_e(0) = \Pi_e^{RPA}(0)\gamma(r_s), \quad \tilde{\kappa}_{TF} \equiv \kappa_{TF}\sqrt{\gamma(r_s)},$$
$$\gamma(r_s) = \left(\frac{9\pi}{4}\right)^{2/3}\frac{6}{r_s^2}\left[r_s^2\frac{\partial^2\epsilon_e}{\partial r_s^2} - 2r_s\frac{\partial\epsilon_e}{\partial r_s}\right]^{-1}. \tag{16}$$

Because the effective proton–proton potential is described by the screened potential of the Thomas–Fermi type in the considered approximation (see Eqs. (12)–(16)),

$$\Phi_{pp}(r) = \frac{e^2}{r}\exp\left(-\frac{r}{\tilde{r}_{TF}}\right), \tag{17}$$

we conclude that a renormalization of the screening radius due to electronic correlations occurs,

$$\tilde{r}_{TF} = \frac{1}{\tilde{\kappa}_{TF}} \equiv \frac{r_{TF}}{\sqrt{\gamma(r_s)}}. \tag{18}$$

We now rewrite Eq. (10) for the considered approximation as

$$\epsilon = \epsilon_e + \epsilon_i, \tag{19}$$

$$\epsilon_i = \frac{3}{2}kT + \frac{1}{2N}\sum_{i \neq j}\Phi_{pp}(R_i - R_j) - \frac{e^2}{d}\left(\frac{\kappa}{2} + \frac{3}{2\kappa^2}\right), \tag{20}$$

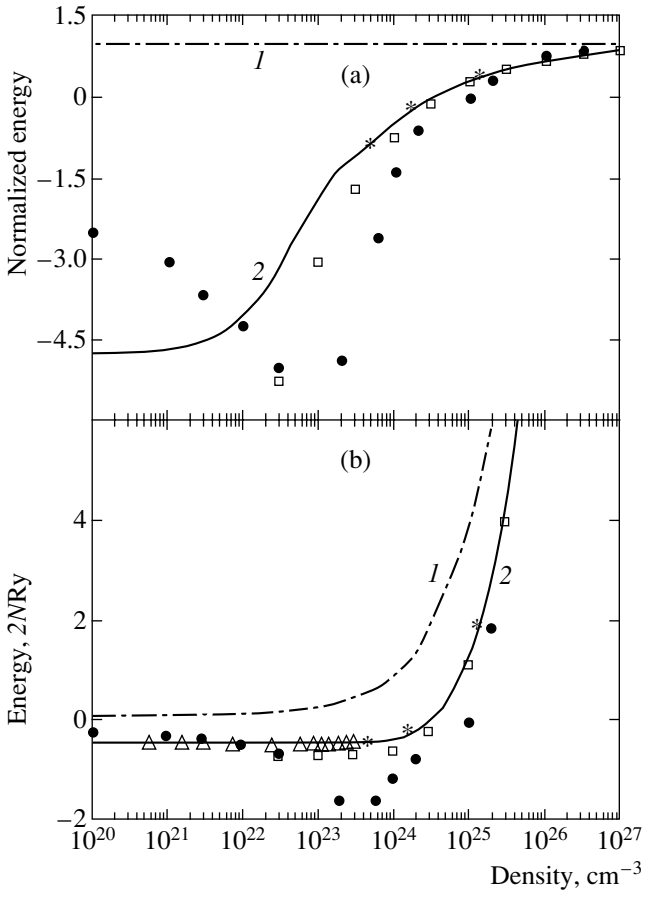

Fig. 1. Internal energy of hydrogen for $T = 10000$ K: (a) normalized to the energy of a noninteracting electron–proton system; (b) in the units of $2N$ Ry. The curves show the following results: ideal plasma (*1*), the PACH calculations (*2*), the EIIP model (□), our Monte Carlo simulations, DPIMC (•), density functional theory [46] (∗), and restricted PIMC data, RPIMC of Militzer *et al.* [34] (△).

where $\kappa \equiv d\tilde{\kappa}_{TF}$. After averaging over the proton positions with the Gibbs distribution (denoted by $\langle \ldots \rangle$), Eq. (19) can be represented as the sum of two terms: the energy ϵ_e of a degenerate electron liquid in the positive homogeneous charge background and the energy of screened classical charged protons interacting via the screened potential (18) and renormalized by the constant terms obtained above,

$$\epsilon_i = \left(u + \frac{3}{2}\right)kT, \tag{21}$$

with

$$u \equiv \Gamma\left\{\frac{d}{2Ne^2}\left\langle \sum_{i \neq j} \Phi_{pp}(R_i - R_j)\right\rangle - \frac{\kappa}{2} - \frac{3}{2\kappa^2}\right\}. \tag{22}$$

Here, u is the ionic interaction energy in kT units. With an accuracy of $(kT/E_F)^2$, energy (21) coincides with the usual thermodynamic energy determined from the free energy of the system because the electrons are degenerate in the considered parameter range (with the same accuracy). Expression (21) implies that as $\kappa \longrightarrow 0$ the energy of a classical one-component system of charged particles interacting via a screened (Debye or Yukawa) potential tends to infinity as $3k_BT\Gamma/2\kappa^2$ (i.e., the screening radius diverges). As a function of the two parameters, Γ and the dimensionless screening length κ, the function u/Γ has been tabulated in [54, 55] for the calculations of the phase diagram of a purely classical one-component Debye plasma (OCP), based on accurate MC calculations for the Debye system. In what follows, we use these numerical results to calculate the energy of a dense hydrogen plasma in the above approximations. Within the Wigner approximation for the electron energy,

$$\epsilon_e = \left(\frac{2.21}{r_s^2} - \frac{0.916}{r_s} + \epsilon_{\rm corr}\right)\text{Ry},$$
$$\epsilon_{\rm corr} = -\frac{0.88}{r_s + 7.8}, \tag{23}$$

we obtain from Eq. (16)

$$\gamma(r_s) = \frac{22.1}{r_s^2\varphi(r_s)},$$
$$\varphi(r_s) = \frac{22.1}{r_s^2} - \frac{3.664}{r_s} - \frac{1.76r_s}{(r_s + 7.8)^2} - \frac{1.76r_s^2}{(r_s + 7.8)^3}, \tag{24}$$

where $\gamma(r_s \longrightarrow 0) \longrightarrow 1$. The total internal energy in Eq. (21) can now be expressed in terms of the tabulated function u/Γ as

$$\epsilon = \left[\frac{2.21}{r_s^2} - \frac{0.916}{r_s} + \epsilon_{\rm corr} + \frac{2}{r_s}\left(\frac{u}{\Gamma} + \frac{3}{2\Gamma}\right)\right]\text{Ry}. \tag{25}$$

Numerical results computed from this approximation are included in Figs. 1–3 below.

Alternatively, we can use additional approximations for the computation of the internal energy of the plasma. This can be done by averaging Eq. (10) over the ion Gibbs distribution with the same effective Hamiltonian. We then immediately find the average energy per proton

$$\frac{\langle E\{R_i\}\rangle}{N_p} = \epsilon_e + \frac{3}{2}k_BT - \frac{1}{2}\int \frac{d^3q}{(2\pi)^3}\frac{u_{ei}^3(q)\Pi_e(q)}{\varepsilon_e(q)}$$
$$+ \frac{1}{2}\int \frac{d^3q}{(2\pi)^3}\tilde{V}_{ii}^{\rm eff}(q)[S_{ii}(q) - 1]$$
$$= \epsilon_e + \frac{3}{2}k_BT + \frac{1}{2}\int \frac{d^3q}{(2\pi)^3}u_{ii}(q)[S_{ii}(q) - 1]$$
$$- \frac{1}{2}\int \frac{d^3q}{(2\pi)^3}\frac{u_{ei}^2\Pi_e(q)}{\varepsilon_e(q)}S_{ii}(q), \tag{26}$$

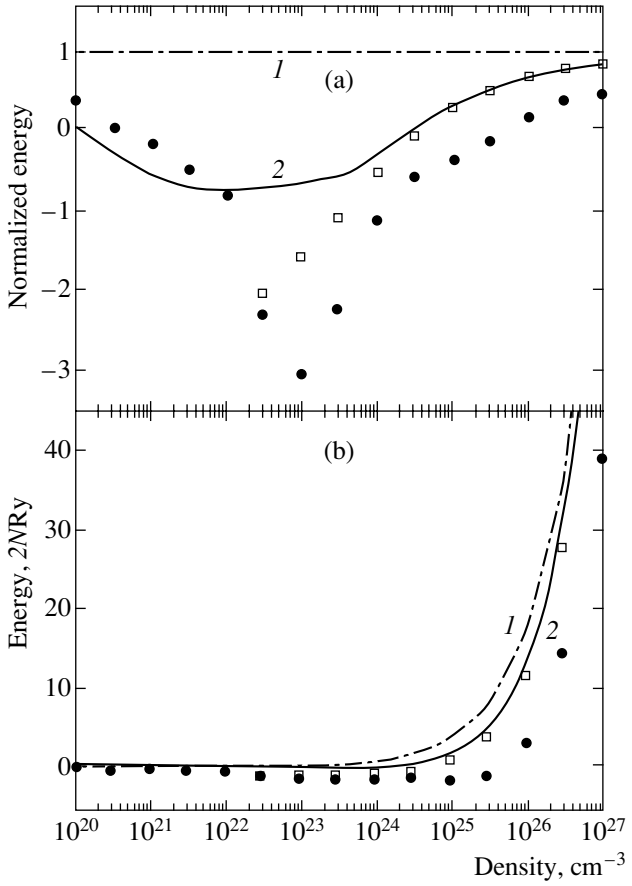

Fig. 2. Internal energy of hydrogen for $T = 30000$ K. The notation is the same as in Fig. 1.

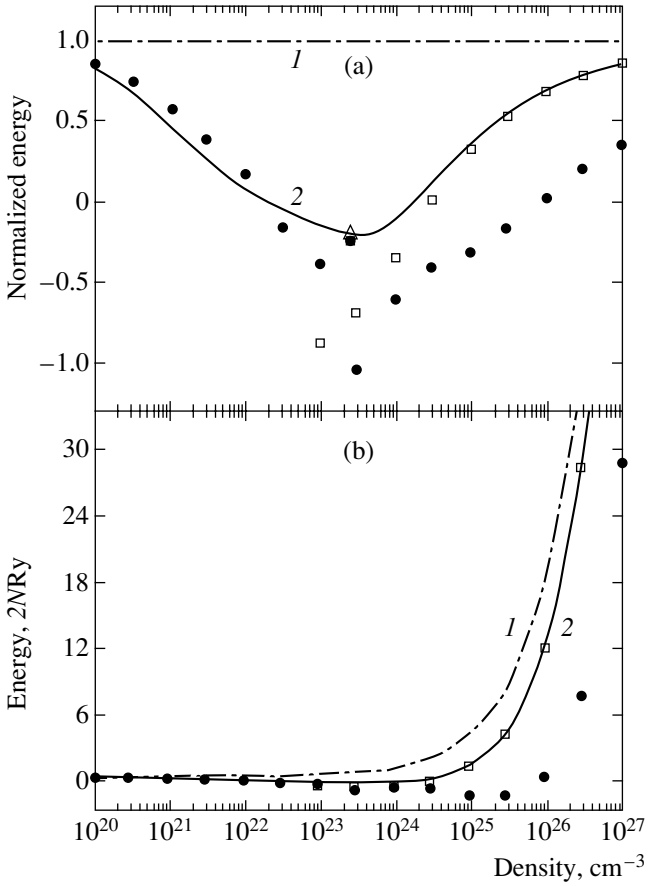

Fig. 3. Internal energy of hydrogen for $T = 50000$ K. The notation is the same as in Fig. 1.

where we introduced the ion–ion structure factor $S_{ii}(k)$, defined as

$$\langle \varrho_{\mathbf{k}_1} \varrho_{\mathbf{k}_2} \rangle = NS_{ii}(\mathbf{k}_1)\delta_{\mathbf{k}_1+\mathbf{k}_2,0} + N^2\delta_{\mathbf{k}_1,0}\delta_{\mathbf{k}_2,0},$$

$$\varrho_{\mathbf{k}} \equiv \sum_j \exp(-i\mathbf{k}\cdot\mathbf{R}_j), \quad \delta_{\mathbf{k},0} = \begin{cases} 1, & k = 0, \\ 0, & k \neq 0. \end{cases} \tag{27}$$

Equation (26) can be simplified by replacing, approximately, the full structure factor by the OCP structure factor S_{ii}^{OCP} computed with the effective ion–ion interaction. The total energy can then be written as the sum of three contributions: the first from the electron subsystem, the second from the classical ion OCP subsystem (each imbedded into a positive and negative charge background, respectively), and a third term ϵ_i^{POL} that describes a perturbation-theory approximation for the polarization of the electron liquid by the ions. The resulting formulas coincide with the perturbation approximations derived by Hansen, De Witt, and others [22, 23],

$$\frac{\langle E\{R_i\}\rangle}{N_p} = \epsilon_e + \epsilon_i^{OCP} + \delta\epsilon, \tag{28}$$

$$\delta\epsilon = \frac{e^2}{\pi}\int_0^\infty dq\left(\frac{1}{\varepsilon_e(q)} - 1\right)S_{ii}^{OCP}(q). \tag{29}$$

As is clear from the above derivations, Eqs. (28) and (29) are less accurate than the full EIIP model presented above.

4. PADÉ APPROXIMATIONS AND CHEMICAL PICTURE: THE PACH METHOD

In this section, we briefly explain the method of Padé approximations in combination with the chemical picture, i.e., Saha equations [3, 19, 44, 45] (PACH). On the basis of the PACH approximation, we calculate the internal energy for the three isotherms $T = 10000$, 30000, and 50000 K. This method works with only analytical formulas, which, however, are rather complicated; nevertheless, the calculation of one energy data point takes no more than a few seconds on a PC.

The Padé approximations were constructed in earlier works from the known analytic results in the limiting cases of low density [3, 43] and high density [3]. The structure of the Padé approximations was devised

such that they are analytically exact up to quadratic terms in the density (up to the second virial coefficient) and interpolate between the virial expansions and the high-density asymptotic expressions [19, 44, 45]. The formation of bound states was taken into account using the chemical picture.

We here follow these cited works in large part, only the contribution of the OCP-ion–ion interaction, which is the largest one in most cases, was substantially improved following [56]. With respect to the chemical picture, we restricted ourselves to the strong ionization region, where the number of atoms is still relatively low and no molecules are present. We here discuss only the general structure of the Padé formulas. The internal energy density of the plasma is given by

$$E = E_{\rm id} + E_{\rm int}, \tag{30}$$

where $E_{\rm id}$ is the internal energy of an ideal plasma consisting of Fermi electrons, classical protons, and classical atoms; and

$$E_{\rm int} = N_p(\epsilon_e + \epsilon_i + \epsilon_a) \tag{31}$$

is the interaction energy. The splitting of the interaction contribution to the internal energy corresponds largely to the previous section. The individual pieces are as follows.

(1) The electron–electron interaction ϵ_e. This term corresponds to the OCP energy of the electron subsystem. Instead of the simple expressions used in earlier works [19, 42, 44], we here used a more refined formula for the energy [57]. This formula is an interpolation between the Hartree limit with the Gell-Man–Brückner correction (already used in the previous section), the Wigner limit, and the Debye law including quantum corrections:

$$\epsilon_e = -\frac{(r_s^3+50)[a_H + a_W(r_s)] + 2\sqrt{6}d_0 r_s^{5.5}\tau^{2.5} + 24 d_H r_s^4 \tau^2}{(r_s^3+50)r_s + 2.3 r_s^4\tau^2 + 2\sqrt{6}d_1 r_s^{5.5}\tau^2 + r_s^7\tau^3}. \tag{32}$$

Here, a Wigner function has been introduced as

$$a_W(x) = 2b_0 x \times \ln\left(1 + \left[x^{0.5}\exp\left(-\frac{b_1}{2b_0}\right) + \frac{2b_0 x}{a_W}\right]^{-1}\right), \tag{33}$$

and the constants take the values

$$d_0 = 0.5, \quad d_1 = 0.6631, \quad d_H = 0.125,$$

$$a_H = 0.91633, \quad a_W = 0.87553,$$

$$b_0 = 0.06218, \quad b_1 = 0.0933.$$

We mention that similar formulas are also valid for other thermodynamic functions with the constants adjusted [57]. The formula for the OCP used here contains all the terms taken into account in the previous section and, in addition, also temperature-dependent corrections.

(2) The ion contribution to the internal energy ϵ_i. This term was calculated in the previous section. We here use a procedure based on approximation (28), (29). This enables us to use the results of the MC calculations of Hansen, De Witt, and others [23, 58]. In accordance with Eqs. (28) and (29), the ion contribution is split into two terms,

$$\epsilon_i = \epsilon_i^{OCP} + \epsilon_i^{POL}, \tag{34}$$

where the first is the OCP contribution of the protons and the second represents the polarization of the proton OCP by the electron gas. For the region of high densities, i.e., large Γ and small r_s, we use the Monte Carlo data parametrized by De Witt as [23]

$$\epsilon_i^{OCP} = -0.8946\Gamma + 0.8165\Gamma^{0.25} - 0.5012, \tag{35}$$

$$\epsilon_i^{POL} = -r_s(0.0543\Gamma + 0.1853\Gamma^{0.25} - 0.0659). \tag{36}$$

We note that the polarization term describes the correction due to screening of the proton–proton interaction by the electron fluid. In order to obtain these expressions, semiclassical Monte Carlo calculations were performed based on effective ion interactions that model the electrons as a responding background [22, 23]. We do not need to go into the details of this method because the procedure corresponds to Eq. (29) derived in the previous section.

In the low density limit, we used the Debye law with quantum corrections [3, 45],

$$\epsilon_i^{OCP} = -0.86603\tau d_0\Gamma^{1.5}[1 - B_1\Gamma^{1.5}], \tag{37}$$

$$\epsilon_i^{POL} = -0.71744\Gamma^{1.5}[1 - C_1\Gamma^{1.5}]. \tag{38}$$

Here, the temperature functions B_1 and C_1 describe rather complex quantum corrections, which are, however, explicitly known and are easily programmed [3]. The Padé approximations that connect the high- and the low-density limits are constructed by standard methods [19, 44, 45] and are not given here explicitly. For the OCP energy of the ions, we use the very accurate formulas proposed by Kahlbaum [56].

(3) The atomic contribution ϵ_a. In the region of densities and temperatures studied in this work, this contribution gives only a small correction (except for T = 10000 K). We calculate the number of atoms on the basis of a nonideal Saha equation described below. The formation of molecules is not taken into account. We restrict calculations to the region where the number density of atoms is smaller than that of the electrons. The contributions to the chemical potential that appear in the Saha equation are calculated in part from scaling

relations and in part by numerical differentiation of the free energy given earlier [19, 44]. For the partition function in the Saha equation, we use the Brillouin–Planck–Larkin expression [3, 45]. The nonideal Saha equation that determines the degree of ionization (the density of atoms) is solved by iterations, starting from the ideal Saha equation. Because of a high degree of ionization, the atomic interaction contributions can be approximated in the simplest way by the second virial contribution and by treating the atoms as small hard spheres and by neglecting the charged particle-neutral interaction.

The results of our Padé calculations for a broad density interval for three isotherms are included in Figs. 1–3.

5. SUMMARY OF THE PATH INTEGRAL MONTE CARLO SIMULATIONS

The analytic approximations discussed in the previous sections work very well at high densities if bound states are of minor importance. These conditions are not fulfilled for densities below a Mott point corresponding to $r_s > 1$. Here, recently developed DPIMC simulations can be used. Starting from the basic plasma particles, electrons and ions, they "automatically" account for bound state formation and ionization and dissociation. Furthermore, in contrast to the chemical picture, no restrictions on the type of chemical species are made and the appearance of complex aggregates such as molecular ions or clusters of several atoms are fully included. On the other hand, simulations are expected to become increasingly difficult at high densities where the electron degeneracy is large due to the fermion sign problem. It is therefore very interesting to compare results of the DPIMC approach with alternative theories that are expected to complement each other. This is done in the next section.

First, however, we briefly outline the idea of our DPIMC scheme. All thermodynamic properties of a two-component plasma are defined by the partition function Z; for N_e electrons and N_p protons, it is given by

$$Z(N_e, N_p, V, \beta) = \frac{Q(N_e, N_p, \beta)}{N_e! N_p!},$$
$$Q(N_e, N_p, \beta) = \sum_\sigma \int_V dq dr \rho(q, r, \sigma; \beta), \tag{39}$$

where $\beta = 1/k_B T$. For a quantum system, the exact density matrix is not known in general, but can be constructed using a path integral representation [21, 59–61],

$$\int_V dR^{(0)} \sum_\sigma \rho(R^{(0)}, \sigma; \beta)$$
$$= \int_V dR^{(0)} \ldots dR^{(n)} \rho^{(1)} \rho^{(2)} \ldots \rho^{(n)} \tag{40}$$
$$\times \sum_\sigma \sum_P (\pm 1)^{\kappa_P} \mathcal{S}(\sigma, \hat{P}\sigma') \hat{P} \rho^{(n+1)},$$

where

$$\rho^{(i)} \equiv \rho(R^{(i-1)}, R^{(i)}; \Delta\beta)$$
$$\equiv \left\langle R^{(i-1)} \middle| \exp(-\Delta\beta \hat{H}) \middle| R^{(i)} \right\rangle,$$

with

$$\Delta\beta \equiv \frac{\beta}{n+1}, \quad \Delta\lambda_a^2 = \frac{2\pi\hbar^2 \Delta\beta}{m_a}, \quad a = p, e.$$

Here, $\hat{H} = \hat{K} + \hat{U}_c$ is the Hamilton operator containing the kinetic and potential energy contributions, $\hat{K}$ and $\hat{U}_c$, respectively, with

$$\hat{U}_c = \hat{U}_c^p + \hat{U}_c^e + \hat{U}_c^{ep}$$

being the sum of the Coulomb potentials between protons (p), electrons (e), and electrons and protons (ep). Further, σ comprises all particle spins, and the particle coordinates are denoted by

$$R^{(i)} = (q^{(i)}, r^{(i)}) \equiv (R_p^{(i)}, R_e^{(i)}), \quad i = 1, \ldots, n+1,$$
$$R^{(0)} \equiv (q, r) \equiv (R_p^{(0)}, R_e^{(0)}), \quad R^{(n+1)} \equiv R^{(0)},$$

where q and r denote the electron and proton coordinates respectively and

$$\sigma' = \sigma.$$

The particles are then represented by fermionic loops with the coordinates (beads)

$$[R] \equiv [R^{(0)}; R^{(1)}; \ldots; R^{(n)}; R^{(n+1)}].$$

The spin gives rise to the spin part of the density matrix $\mathcal{S}$, and the exchange effects are taken into account by the permutation operator $\hat{P}$ that acts on the electron coordinates and spin projections and by the sum over permutations with the parity κ_P. In the fermionic case (minus sign), the sum contains $N_e!/2$ positive and negative terms, which leads to the notorious sign problem. Because of the large mass difference of electrons and ions, the exchange of the latter is not included.

Thermodynamic functions are given by derivatives of the logarithm of the partition function with respect to thermodynamic variables. In particular, the internal energy E follows from Q by

$$\beta E = -\beta\frac{\partial \ln Q}{\partial \beta}, \tag{41}$$

which gives (cf. [42] for details)

$$\begin{aligned}\beta E &= \frac{3}{2}(N_e + N_p) + \frac{1}{Q}\frac{1}{\lambda_p^{3N_p}\Delta\lambda_e^{3N_e}}\\ &\times \sum_{s=0}^{N_e}\int dq\,dr\,d\xi\,\rho_s(q,[r],\beta)\\ &\times\left\{\sum_{p<t}^{N_p}\frac{\beta e^2}{|q_{pt}|} + \sum_{l=0}^{n}\left[\sum_{p<t}^{N_e}\frac{\Delta\beta e^2}{|r_{pt}^l|} + \sum_{p=1}^{N_p}\sum_{t=1}^{N_e}\Psi_l^{ep}\right]\right.\\ &+\left.\sum_{l=1}^{n}\left[-\sum_{p<t}^{N_e}C_{pt}^l\frac{\Delta\beta e^2}{|r_{pt}^l|^2} + \sum_{p=1}^{N_p}\sum_{t=1}^{N_e}D_{pt}^l\frac{\partial\Delta\beta\Phi^{ep}}{\partial|x_{pt}^l|}\right]\right.\\ &\left.-\frac{1}{\det\|\psi_{ab}^{n,1}\|_s}\frac{\partial\det\|\psi_{ab}^{n,1}\|_s}{\partial\beta}\right\},\end{aligned} \tag{42}$$

$$C_{pt}^l = \frac{\langle r_{pt}^l | y_{pt}^l\rangle}{2|r_{pt}^l|}, \quad D_{pt}^l = \frac{\langle x_{pt}^l | y_p^l\rangle}{2|x_{pt}^l|},$$

where

$$\Psi_l^{ep} \equiv \Delta\beta\left.\frac{\partial[\beta'\Phi^{ep}(|x_{pt}^l|,\beta')]}{\partial\beta'}\right|_{\beta'=\Delta\beta}$$

contains the electron–proton Kelbg potential Φ^{ep} (cf. Eq. (45) below); $\langle\ldots|\ldots\rangle$ denotes the scalar product; and q_{pt}, r_{pt}, and x_{pt} are differences of the two coordinate vectors,

$$q_{pt} \equiv q_p - q_t, \quad r_{pt} \equiv r_p - r_t,$$

$$x_{pt} \equiv r_p - q_t, \quad r_{pt}^l = r_{pt} + y_{pt}^l,$$
$$x_{pt}^l \equiv x_{pt} + y_p^l, \quad y_{pt}^l \equiv y_p^l - y_t^l,$$

with

$$y_a^n = \Delta\lambda_e\sum_{k=1}^{n}\xi_a^{(k)}.$$

We introduced dimensionless distances between neighboring vertices on the loop, $\xi^{(1)}$, …, $\xi^{(n)}$, and thus, explicitly, $[r] \equiv [r; y_e^{(1)}; y_e^{(2)}; \ldots]$. The density matrix ρ_s in Eq. (42) is given by

$$\begin{aligned}\rho_s(q,[r],\beta) &= C_{N_e}^s\exp(-\beta U(q,[r],\beta))\\ &\times\prod_{l=1}^{n}\prod_{p=1}^{N_e}\phi_{pp}^l\det\|\psi_{ab}^{n,1}\|_s,\end{aligned} \tag{43}$$

where

$$\begin{aligned}U(q,[r],\beta) &= U_c^p(q)\\ &+\frac{\{U^e([r],\Delta\beta) + U^{ep}(q,[r],\Delta\beta)\}}{n+1},\end{aligned}$$

$$\phi_{pp}^l \equiv \exp[-\pi|\xi_p^{(l)}|^2].$$

Density matrix (43) does not involve an explicit sum over the permutations and hence does not involve the sum of terms with alternating signs. Instead, the entire exchange problem is contained in a single exchange matrix given by

$$\|\psi_{ab}^{n,1}\|_s \equiv \left\|\exp\left\{-\frac{\pi}{\Delta\lambda_e^2}|(r_a - r_b) + y_a^n|^2\right\}\right\|_s. \tag{44}$$

As a result of the spin summation, the matrix carries a subscript s indicating the number of electrons having the same spin projection.

The potential Φ^{ab} in Eq. (42) is an effective quantum pair interaction between two charged particles immersed into a weakly degenerate plasma. It was derived by Kelbg *et al.* [62, 63], who showed that it contains quantum effects exactly in the first order in the coupling parameter Γ,

$$\begin{aligned}\Phi^{ab}(|\mathbf{r}_{ab}|,\Delta\beta) &= \frac{e_a e_b}{\lambda_{ab}x_{ab}}\\ &\times\{1 - \exp(-x_{ab}^2) + \sqrt{\pi}x_{ab}[1 - \mathrm{erf}(x_{ab})]\},\end{aligned} \tag{45}$$

where $x_{ab} = |\mathbf{r}_{ab}|/\lambda_{ab}$; we emphasize that the Kelbg potential is finite at zero distance.

The structure of Eq. (42) is obvious: we have separated the classical ideal gas part (the first term). The ideal quantum part in excess of the classical one and the correlation contributions are contained in the integral term, where the second line results from the ionic correlations (the first term) and the *ee* and *ei* interactions at the first vertex (the second and the third term, respectively). Equation (42) therefore contains an important limit of the ideal quantum plasma in a natural way. The third and fourth lines are due to further electronic vertices and the explicit temperature dependence (Eq. (42)) and volume dependence (the corresponding equation of state) of the exchange matrix, respectively. The main

advantage of Eq. (42) is that the explicit sum over permutations has been converted into the spin determinant that can be very efficiently computed using standard linear algebra methods. Furthermore, each of the sums in curly brackets in Eq. (42) is bounded as the number of vertices increases, $n \longrightarrow \infty$. The error of the total expression is of the order $1/n$. Expression (42) and the analogous result for the equation of state are therefore well suited for numerical evaluation using the standard Monte Carlo techniques (see, e.g., [21, 29]).

In our Monte Carlo scheme, we used three types of steps, where either electron or proton coordinates, r_i or q_i, or individual electronic beads $\xi_i^{(k)}$ were moved until convergence of the calculated values was reached. Our procedure has been extensively tested. In particular, from comparison with the known analytic expressions for the pressure and energy of the ideal Fermi gas, we found that the Fermi statistics is very well reproduced with a limited number of particles ($N \lesssim 100$) and beads for degeneracy up to $n\lambda^3 \lesssim 10$ [40]. We also performed extensive tests for few-electron systems in a harmonic trap, where the analytically known limiting behavior (e.g., energies) is again reproduced well [64, 65]. For the present simulations of dense hydrogen, we varied both the number of particles and the number of time slices (beads). As a result of these tests, we found that to obtain convergent results for the thermodynamic properties of hydrogen in the density–temperature region of interest here, number of particles $N_e = N_p = 50$ and bead numbers in the range $n = 6$–20 are an acceptable compromise between accuracy and computational effort [39–41].

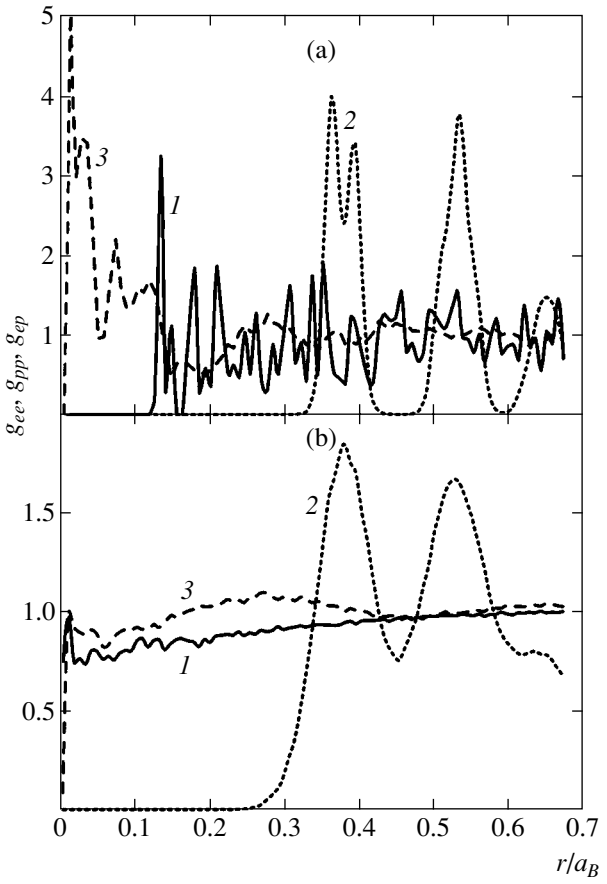

Fig. 4. Electron–electron (*1*), proton–proton (*2*), and electron–proton (*3*) pair distribution functions of hydrogen from the DPIMC simulations at $n = 10^{26}$ cm^{-3} for the temperatures 10000 K (a) and 50000 K (b). Note the different vertical scales.

6. NUMERICAL RESULTS: COMPARISON OF THE ANALYTIC AND SIMULATION DATA

We now discuss the numerical results. We have computed the internal energy of dense hydrogen using two analytic (EIIP and PACH) approaches and DPIMC simulations. The data are shown in Figs. 1–3 for three temperatures: 10000, 30000, and 50000 K, respectively.

We first consider the general behavior that is most clearly seen for the highest temperature (cf. Fig. 3a). The overall trend is an increase of the energy with density, which is particularly rapid at high densities because of electron degeneracy effects; this is clearly seen from the ideal plasma curve (dashed and dotted lines in the lower parts of Figs. 1–3). The nonideal plasma results show a prominent deviation from this trend, which is in full agreement with the discussion in Section 2, the formation of an energy minimum (where the energy can become negative) at intermediate densities. Our calculations for a nonideal hydrogen plasma asymptotically approach the ideal curve both at low density (the ideal classical plasma) and at high density (the ideal mixture of classical protons and quantum electrons). For intermediate densities, between 10^{21} and 10^{25} cm^{-3}, the nonideal plasma energy is significantly lower than the ideal energy because of strong correlations and formation of bound states. As the temperature decreases, this region broadens. In particular, we clearly see that the total energy indeed reaches negative values for the temperatures considered.

We now compare the results of the different methods. We consider three density regions, separately (A) the high-density limit, (B) the region around the minimum, and (C) the region below the minimum.

(A) The first observation from Figs. 1–3 is that for all temperatures (including temperatures above those shown), the PACH and EIIP approaches nearly coincide in the limit of high densities. It is also interesting to compare these approaches with another theoretical approach based on the density functional theory (DFT). Recently, Xu and Hansen [46] published data for $T =$ 10000 K and $r_s \leq 1.5$, which are also included in Fig. 1. Evidently, in the high-density limit, PACH and EIIP coincide with these DFT data (cf. Fig. 1). This good agreement of the three completely independent approaches—EIIP, PACH, and DFT—is a strong indi-

cation that they can yield reliable results for a fully ionized macroscopic hydrogen plasma at high densities. This asymptotic agreement is not surprising, because the ideal Fermi gas limit is "built into" each of these three approaches, but this gives no information about the lowest densities for which these results remain quantitatively correct. The comparison presented is therefore of great importance for giving a hint (although not a proof) that the value of that minimum density is above $n \approx 3 \times 10^{24}$ cm^{-3} (cf. Figs. 1–3).

We next observe that at higher densities, the DPIMC simulations yield lower energies and a shift of the increase in energy to higher density values compared with the analytic models. This tendency becomes stronger with increasing temperature, as can be seen in Figs. 1–3. In view of the asymptotic accuracy of the analytic results (see above), the total energy of macroscopic high-density hydrogen is certainly above the DPIMC results for densities exceeding 10^{25} cm^{-3}. There are two main factors tending to reduce our DPIMC results for energies at high densities. The first factor is given by degeneracy effects. Practical limitations that must be imposed on the number of beads and particles (see Section 5) necessarily make our results less reliable for densities exceeding 10^{25} cm^{-3}. The second factor is given by finite-size effects related to photon ordering. To understand the high-density results better, we analyze the electron–electron (*ee*), proton–proton (*pp*), and electron–proton (*ep*) pair distribution functions in Fig. 4. These functions exhibit features typical for strongly correlated systems. The most prominent effect is seen in the *pp* function that exhibits a periodic structure at $T = 50000$ K, which is even more pronounced at $T = 10000$ K. This proton ordering is typical of a strongly correlated ion fluid near the crystallization temperature.[3] Our simulations for even higher densities reveal the formation of an ionic lattice immersed into a delocalized sea of electrons, i.e., an ionic Wigner crystal [40], known to exist in high-density objects such as white or brown dwarf stars. Thus, qualitatively, the simulations show the correct behavior at high densities. However, because of the small size of the simulations (only 50 electrons and protons are presently feasible), the results are much closer to those for small, strongly correlated ionic clusters that are known to exhibit quite peculiar behavior, including a strong dependence of the energy on size, negative specific heat, etc. In order to obtain more accurate data for the internal energy of a macroscopic two-component plasma at ultrahigh compression, a significant increase in the simulation size is therefore desirable (it should become feasible in the near future).

(B) The energy minimum at intermediate densities is reproduced by all methods, but there are quantitative differences regarding its depth and width. The general observation made for all temperatures (cf. Figs. 1–3) is that the simulations yield a deeper minimum and shift of the increase in energy toward higher densities. We also observe that the EIIP method yields lower energies than the PACH results and is closer to the DPIMC results. Further, the PACH results nearly coincide with the DFT data [46] where available ($T = 10000$ K and $n \geq 5 \times 10^{23}$ cm^{-3}). However, atom and molecule formation becomes important at these densities, and the EIIP and DFT methods (in their present form) become increasingly unreliable. The PACH results presented include bound states approximately, whereas the DPIMC calculations have no restrictions with respect to atom and molecule formation.

We now analyze the DPIMC simulations around the energy minimum. Our data for $T = 10000$ K are also significantly lower than the RPIMC results of Militzer *et al.* [34] (cf. Fig. 1b), but we found excellent quantitative agreement between the two independent quantum Monte Carlo methods above $T = 50000$ K (see the point for $T = 62500$ K in Fig. 3a; a more extensive comparison is also given in [42]). A detailed analysis of the DPIMC simulation results at $T = 10000$ K and $10^{23} \lesssim n \lesssim 10^{24}$ cm^{-3} revealed that the homogeneous plasma state is unstable there: the plasma gains energy by forming higher density clusters or droplets that are embedded in a lower-density plasma. The droplets are clearly visible in the electron–proton configurations in the simulation box [68] and are interpreted as a direct indication of a first-order phase transition, as discussed in the Introduction [13–20]. These effects emerge in weakly ionized plasma (low density) and vanish above the Mott point, $r_s \sim 1$. We mention that the same effects are observed in our DPIMC simulations of electron–hole plasmas under similar conditions [69], for which droplet formation has been well established and was observed experimentally three decades ago [70]. Our conclusion is also indirectly supported by analytic methods. In the present variants of the PACH and EIIP methods, homogeneous density distributions are assumed,[4] but it is interesting that at $T = 10000$ K and 10^{23} cm$^{-3} \lesssim n \lesssim 10^{24}$ cm^{-3}, both methods yield unstable results for the thermodynamic functions, which is a clear indication of the existence of a first-order phase transition. Xu and Hansen [46] also observed strong fluctuations in their density functional calculations below $r_s = 1.5$, which they found to strongly resemble precursors of a phase transition.

Even if we accept the existence of a phase transition, the energy obtained in the DPIMC simulations appears to be unexpectedly low. In this region, we observe large

[3] In fact, the first minimum of the proton–proton function (around $r = 0.45a_B$) for $T = 10000$ K is far lower than the standard value of 0.35 typical for a liquid.

[4] Using a modified PACH approach, Beule *et al.* [20] predicted a first-order phase transition in hydrogen for $T = 10000$ K with the pressure and density of $p \approx 110$ GPa and $\rho \approx 0.8$ g cm^{-3}. Similar results were obtained by Schlanges *et al.* [16], where other references are also given.

fluctuations of pressure and energy related to the formation and decay of droplets. Furthermore, there are significant surface energy effects. Our simulations yield only a very small number of droplets (typically one to three), each containing 15–50 electron–proton pairs. Of these, almost all are on the surface.[5] We therefore cannot be certain that our simulations adequately represent the macroscopic system in the two-phase region and its energy per particle in particular. Finally, in such small systems, there exist additional specific factors that tend to lower the energy [12] and can therefore be responsible for our result. This can also be the reason for the low-energy minimum observed at higher temperatures (Figs. 2 and 3), where our analytic models do not show unstable behavior. Clarifying this interesting issue in more detail requires performing extensive simulations with a substantially larger number of particles; such calculations are presently under way.

(C) In the region to the left of the energy minimum, $n \lesssim 10^{22}$ cm^{-3}, simulations are of special importance. In this region, the plasma is strongly correlated and largely dominated by bound states. Analytic methods based on perturbation expansions in the coupling strength (Γ or r_s) do not work in this parameter range. While there exist many chemical models of various degrees of sophistication, rigorous theoretic results are very rare. In particular, as mentioned above, the EIIP model breaks down here (we recall that the present version is limited to $r_s \leq 1.5$, cf. Section 3). Also, the present version of the PACH approach uses a chemical picture with a nonideal Saha equation, but treats bound states in a very simple approximation (see above) and therefore yields only approximate results in situations with a low degree of ionization). In contrast, the DPIMC results have no such limitations and provide reliable results in this region in principle. On the other hand, there exist specific technical difficulties at low densities, where the extension of bound electron wave functions is many orders of magnitude smaller than the interparticle distance, which leads to very slow convergence of DPIMC simulations (and other quantum Monte Carlo methods as well). This explains the different energies in the DPIMC and Padé results at the lowest densities at $T = 10000$ K, where the plasma consists of atoms (see, e.g., [16, 20]), while good agreement is found at higher temperatures.

7. DISCUSSION

This work is devoted to the investigation of the total energy of warm dense plasmas in the temperature range between 10000 and 50000 K. We presented a new theoretical approach to high-density plasmas based on the theory of an effective ion–ion potential. This method was shown to be quite efficient for fully ionized, strongly correlated plasmas above the Mott density. Furthermore, a detailed comparison of several theoretical approaches and simulations was performed over a wide density range. The first included the EEIP and PACH analytic models, on the one hand, and the recent DFT data of Xu and Hansen [46], on the other. The second group of data consisted of several new data points based on DPIMC simulations of a correlated proton–electron system with degenerate electrons. From these comparisons, we conclude that the three theoretical approaches—PACH, EEIP, and DFT—are in a very good agreement with each other for a fully ionized hydrogen plasma in the high-density region where $r_s < 1$. We therefore expect these results to be reliable for densities above 3×10^{24} cm^{-3}. This agreement between the three independent analytic methods is highly interesting because the physical approximations involved are very different. On the other hand, our DPIMC simulations agree with the available RPIMC data for temperatures above 50000 K (cf. Figs. 1–3 and [42]). This agreement over a broad range of parameters is certainly remarkable because the plasma is far outside the perturbative regime: it is strongly correlated and the electrons are degenerate; and the two simulations are essentially independent.

The comparison of our DPIMC simulation results with the analytic data reveals an overall good agreement. In addition, existing deviations are a useful guide for future improvement and extension of the various approaches. Most importantly, the good quality of the quantum Monte Carlo data in the region of strong changes of the ionization degree fills a gap in the present variants of analytic methods. These data can be used to improve the treatment of a dense plasma via analytic methods in the theoretically very complicated region of strong correlations and strongly varying degrees of ionization and dissociation. Moreover, the high-density asymptotic results of the analytic methods may be useful for further improvement of the simulations.

Furthermore, our DPIMC simulations revealed an instability of the homogeneous plasma state around the minimum of the energy isotherm $T = 10000$ K for densities between 10^{23} and 10^{24} cm^{-3}. We have made arguments that this is related to the droplet formation, which is a strong indication of a first-order phase transition [68], which had been previously predicted by many authors on the basis of simple chemical models. The existence of a plasma phase transition would have drastic consequences for the transport properties of many astrophysical objects, such as giant planets, and its verification therefore remains an important theoretical issue. It would therefore be of great interest if independent first-principle simulations, in particular RPIMC, could reproduce this result. This, however, may require a particular choice of nodes of the density matrix that allow

[5] Imagine 27 (64) particles arranged, for simplicity, in a small cube. Evidently, 26 (56) of them are situated at the surface of the "droplet," which is in a striking contrast to the situation of a macroscopic system, where droplets are expected to contain many orders of magnitude more particles.

an inhomogeneous equilibrium plasma state. Finally, at very high densities, our DPIMC simulations revealed ordering of protons into a strongly correlated fluid and the onset of the formation of a proton Wigner crystal. These interesting physical effects in high-pressure hydrogen are of relevance for many astrophysical systems and many laboratory experiments, including ultracold degenerate trapped ions and laser plasmas.

In conclusion, we may state that the analytic methods and the DPIMC approach are already in a reasonable overall agreement. Both methods should be developed to further explore the equilibrium properties of dense hydrogen.

ACKNOWLEDGMENTS

We acknowledge stimulating discussions with H.E. DeWitt, W.D. Kraeft, D. Kremp, B. Militzer, R. Redmer, and M. Schlanges.

This work has been supported by the Deutsche Forschungsgemeinschaft (BO-1366) and by a grant for CPU time at NIC Jülich.

REFERENCES

1. *Strongly Coupled Coulomb Systems*, Ed. by G. Kalman (Pergamon, New York, 1998).
2. *Proceedings of International Conference on Strongly Coupled Plasmas*, Ed. by W. D. Kraeft and M. Schlanges (World Sci., Singapore, 1996).
3. W. D. Kraeft, D. Kremp, W. Ebeling, and G. Röpke, *Quantum Statistics of Charged Particle Systems* (Akademie, Berlin, 1986).
4. M. Bonitz, *Quantum Kinetic Theory* (Teubner, Stuttgart, 1998).
5. *Proceedings of the 10th International Workshop on the Physics of Nonideal Plasmas*, Ed. by H. Haberland, M. Schlanges, and W. Ebeling; Contrib. Plasma Phys. **41** (2–3) (2001).
6. *Progress in Nonequilibrium Green's Functions*, Ed. by M. Bonitz (World Sci., Singapore, 2000).
7. L. B. Da Silva *et al.*, Phys. Rev. Lett. **78**, 483 (1997).
8. S. T. Weir, A. C. Mitchell, and W. J. Nellis, Phys. Rev. Lett. **76**, 1860 (1996).
9. E. P. Wigner and H. B. Huntington, J. Chem. Phys. **3**, 764 (1935).
10. A. A. Abrikosov, Zh. Éksp. Teor. Fiz. **39**, 1798 (1960) [Sov. Phys. JETP **12**, 1254 (1961)]; Zh. Éksp. Teor. Fiz. **41**, 565 (1961) [Sov. Phys. JETP **14**, 408 (1962)]; Zh. Éksp. Teor. Fiz. **45**, 2038 (1963) [Sov. Phys. JETP **18**, 1397 (1963)].
11. N. W. Ashcroft and D. Stroud, in *Solid State Physics*, Ed. by H. Ehrenreich, F. Seitz, and D. Turnbull (Academic, New York, 1968), Vol. 33, p. 1; N. W. Ashcroft, Phys. Rev. Lett. **21**, 1748 (1968).
12. E. G. Brovman, Yu. Kagan, and A. Kholas, Zh. Éksp. Teor. Fiz. **61**, 2429 (1971) [Sov. Phys. JETP **34**, 1300 (1971)].
13. G. E. Norman and A. N. Starostin, Teplofiz. Vys. Temp. **6**, 410 (1968); Teplofiz. Vys. Temp. **8**, 413 (1970).
14. P. Haronska, D. Kremp, and M. Schlanges, Wiss. Z. Univ. Rostock **98**, 1 (1987).
15. D. Saumon and G. Chabrier, Phys. Rev. A **44**, 5122 (1991).
16. M. Schlanges, M. Bonitz, and A. Tschttschjan, Contrib. Plasma Phys. **35**, 109 (1995).
17. W. Ebeling, W. D. Kraeft, and D. Kremp, *Theory of Bound States and Ionization Equilibrium in Plasmas and Solids* (Akademie, Berlin, 1976; Mir, Moscow, 1979).
18. W. Ebeling, Physica A (Amsterdam) **130**, 587 (1985).
19. W. Ebeling and W. Richert, Phys. Lett. A **108**, 80 (1985); Phys. Status Solidi B **128**, 167 (1985).
20. D. Beule *et al.*, Phys. Rev. B **59**, 14177 (1999); Contrib. Plasma Phys. **39**, 21 (1999).
21. V. M. Zamalin, G. E. Norman, and V. S. Filinov, *The Monte Carlo Method in Statistical Thermodynamics* (Nauka, Moscow, 1977); B. V. Zelener, G. E. Norman, and V. S. Filinov, *Perturbation Theory and Pseudopotential in Statistical Thermodynamics* (Nauka, Moscow, 1981).
22. J.-P. Hansen and P. Viellefosse, Phys. Lett. A **53**, 187 (1975); S. Galam and J. P. Hansen, Phys. Rev. A **14**, 816 (1976).
23. H. E. De Witt, in *Strongly Coupled Plasmas*, Ed. by G. Kalman and P. Carini (Plenum, New York, 1978).
24. N. P. Kovalenko, Yu. P. Krasny, and S. A. Trigger, *Statistical Theory of Liquid Metals* (Nauka, Moscow, 1990).
25. M. Bonitz *et al.*, J. Phys. B **8**, 6057 (1996).
26. V. Golubnychiy, M. Bonitz, D. Kremp, and M. Schlanges, Phys. Rev. E **64**, 016409 (2001).
27. W. Ebeling, W. Stolzmann, A. Förster, and M. Kasch, Contrib. Plasma Phys. **39**, 287 (1999).
28. D. Klakow, C. Toepffer, and P.-G. Reinhard, Phys. Lett. A **192**, 55 (1994); J. Chem. Phys. **101**, 10766 (1994).
29. *The Monte Carlo and Molecular Dynamics of Condensed Matter Systems*, Ed. by K. Binder and G. Cicotti (SIF, Bologna, 1996).
30. *Classical and Quantum Dynamics of Condensed Phase Simulation*, Ed. by B. J. Berne, G. Ciccotti, and D. F. Coker (World Sci., Singapore, 1998).
31. D. M. Ceperley, in *The Monte Carlo and Molecular Dynamics of Condensed Matter Systems*, Ed. by K. Binder and G. Cicotti (SIF, Bologna, 1996), p. 447.
32. D. M. Ceperley, Rev. Mod. Phys. **67**, 279 (1995).
33. B. Militzer and E. L. Pollock, Phys. Rev. E **61**, 3470 (2000).
34. B. Militzer and D. M. Ceperley, Phys. Rev. Lett. **85**, 1890 (2000).
35. V. S. Filinov, J. Phys. A **34**, 1665 (2001).
36. M. Imada, J. Phys. Soc. Jpn. **53**, 2861 (1984).
37. M. F. Herman, E. J. Bruskin, and B. J. Berne, J. Chem. Phys. **76**, 5150 (1982).
38. V. S. Filinov, P. R. Levashov, V. E. Fortov, and M. Bonitz, in *Progress in Nonequilibrium Green's Functions*, Ed. by M. Bonitz (World Sci., Singapore, 2000), p. 513; cond-mat/9912055.
39. V. S. Filinov and M. Bonitz, cond-mat/9912049.
40. V. S. Filinov, M. Bonitz, and V. E. Fortov, Pis'ma Zh. Éksp. Teor. Fiz. **72**, 361 (2000) [JETP Lett. **72**, 245 (2000)].

41. V. S. Filinov, V. E. Fortov, M. Bonitz, and D. Kremp, Phys. Lett. A **274**, 228 (2000).
42. V. S. Filinov, M. Bonitz, W. Ebeling, and V. E. Fortov, Plasma Phys. Controlled Fusion **43**, 743 (2001).
43. W. Ebeling, Ann. Phys. (Leipzig) **21**, 315 (1968); Ann. Phys. (Leipzig) **22**, 33, 383, 392 (1969); Physica (Amsterdam) **38**, 378 (1968); Physica (Amsterdam) **40**, 290 (1968).
44. W. Ebeling, Contrib. Plasma Phys. **29**, 165 (1989); Contrib. Plasma Phys. **30**, 553 (1990).
45. W. Ebeling, A. Förster, V. Fortov, *et al.*, *Thermophysical Properties of Hot Dense Plasmas* (Teubner, Stuttgart, 1991).
46. H. Xu and J. P. Hansen, Phys. Rev. E **57**, 211 (1998).
47. F. J. Dyson and A. Lenard, J. Math. Phys. **8**, 423 (1967).
48. E. H. Lieb and W. Thirring, Phys. Rev. Lett. **35**, 1116 (1975).
49. E. Wigner, Phys. Rev. **46**, 1002 (1934).
50. W. A. Harrison, *Pseudopotentials in the Theory of Metals* (Benjamin, New York, 1966; Mir, Moscow, 1968).
51. V. B. Bobrov and S. A. Trigger, Solid State Commun. **56**, 21 (1985).
52. S. D. Kaim, N. P. Kovalenko, and E. V. Vasiliu, J. Phys. Stud. **1**, 589 (1977).
53. G. Mahan, *Many-Particle Physics* (Plenum, New York, 1981).
54. S. Hamaguchi, R. T. Farouki, and D. H. E. Dubin, Phys. Rev. E **56**, 4671 (1997).
55. R. T. Farouki and S. Hamaguchi, J. Chem. Phys. **101**, 9885 (1994).
56. T. Kahlbaum, in *Physics of Strongly Coupled Plasmas*, Ed. by W. D. Kraeft and M. Schlanges (World Sci., Singapore, 1996).
57. W. Ebeling, W. D. Kraeft, D. Kremp, and G. Röpke, Physica A (Amsterdam) **140**, 160 (1986).
58. W. L. Slattery, G. D. Doolen, and H. E. De Witt, Phys. Rev. A **21**, 2087 (1980).
59. R. P. Feynman and A. R. Hibbs, *Quantum Mechanics and Path Integrals* (McGraw-Hill, New York, 1965; Mir, Moscow, 1968).
60. V. S. Filinov, High Temp. **13**, 1065 (1975); High Temp. **14**, 225 (1976).
61. B. V. Zelener, G. E. Norman, and V. S. Filinov, High Temp. **13**, 650 (1975).
62. G. Kelbg, Ann. Phys. (Leipzig) **12**, 219 (1963); Ann. Phys. (Leipzig) **13**, 354 (1964); Ann. Phys. (Leipzig) **14**, 394 (1964).
63. W. Ebeling, H. J. Hoffmann, and G. Kelbg, Contrib. Plasma Phys. **7**, 233 (1967) and references therein.
64. A. V. Filinov, Yu. E. Lozovik, and M. Bonitz, Phys. Status Solidi B **221**, 231 (2000).
65. A. V. Filinov, M. Bonitz, and Yu. E. Lozovik, Phys. Rev. Lett. **86**, 3851 (2001).
66. S. A. Trigger and W. Ebeling, in *Proceedings of the II Capri Workshop on Dusty Plasmas* (2001).
67. N. I. Klyuchnikov and S. A. Trigger, Dokl. Akad. Nauk SSSR **238**, 565 (1978) [Sov. Phys. Dokl. **23**, 60 (1978)].
68. V. S. Filinov, V. E. Fortov, M. Bonitz, and P. R. Levashov, Pis'ma Zh. Éksp. Teor. Fiz. **74**, 422 (2001) [JETP Lett. **74**, 384 (2001)].
69. V. S. Filinov, M. Bonitz, P. R. Levashov, *et al.*, J. Phys. A: Math. General (accepted for publication).
70. *Electron–Hole Droplets in Semiconductors*, Ed. by C. D. Jeffries and L. V. Keldysh (Nauka, Moscow, 1977); J. C. Hensel, T. G. Phillips, and G. A. Thomas, Solid State Phys. **32**, 88 (1977).

Journal of Experimental and Theoretical Physics, Vol. 96, No. 3, 2003, pp. 480–488.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 123, No. 3, 2003, pp. 543–551.

FLUIDS

Features of Diffusion of an Incident Particle

E. Z. Gribova*, I. S. Zhukova, S. A. Lapinova, A. I. Saichev, and T. Elperin
Nizhnii Novgorod State University, Nizhnii Novgorod, 603950 Russia
**e-mail: gribova@rf.unn.ru*
Received July 29, 2002

Abstract—Mechanisms of turbulent diffusion of particles moving in a turbulent medium under the action of gravity and Stokes forces are studied. It is observed that the velocity variance of particles decreases upon an increase in the inertia of the motion. At the same time, the diffusion coefficients of a particle are independent of its inertia. It is shown that, as the velocity of incidence increases, diffusion in the vertical plane becomes stronger than in the horizontal plane, and the ratio of the horizontal and vertical diffusion coefficients decreases to 1 : 2. © *2003 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

Analysis of diffusion of impurity particles in a turbulent medium is important for solving ecological and meteorological problems. Many authors assume that an impurity is passive and moves at the velocity of the ambient medium (see, for example, [1–6]). In real situations, however, effects associated with the inertia of motion of particles and force of gravity play a significant role. A typical example is aerosol or raindrops in a turbulent atmosphere. The inertia of particles may considerably affect the laws of turbulent diffusion. For example, Maxey [7] proved, using numerical simulation, that the average velocity of particles in a turbulent medium is higher than the velocity of incidence in a stationary medium, while Csanady [8] noted that the turbulent diffusion coefficient in the vertical plane may be twice as large as the diffusion coefficient in the horizontal plane. Here, we derive an equation for particle diffusion taking into account the gravity and Stokes forces and consider its corollaries as applied to the statistics of coordinates and the velocity of particles. The following physical effect is discovered: although the variance of particle velocity decreases upon an increase in the inertia of particle motion, diffusion coefficients are virtually independent of inertia and are determined only by the free-fall velocity. The dependence of the particle velocity variance and of the transverse and longitudinal diffusion coefficients on the average velocity of particle fall is considered in detail. It is shown that as the fall velocity increases, the ratio of the transverse and longitudinal diffusion coefficients decreases gradually to 1 : 2. The physical reason for this effect is that the transverse correlation function of the velocity of vortex motion decreases at a higher rate than the longitudinal correlation function and contains by negative correlation regions.

2. LAWS OF MOTION OF PARTICLES

Let us consider the motion of impurity particles in an incompressible turbulent medium whose random velocity field is $\mathbf{u}(\mathbf{x}, t)$. We assume that the size of impurity particles is much smaller than the inner scale of turbulence and disregard the possibility of particle collisions. If, in addition, we can disregard the effect of particles on the velocity field of the medium and molecular diffusion, the motion of a particle can be correctly described by the equations

$$\frac{d\mathbf{X}}{dt} = \mathbf{V}, \quad \frac{d\mathbf{V}}{dt} + \lambda\mathbf{V} = \lambda\mathbf{u}(\mathbf{X}, t) + \mathbf{g}.$$

Here, $\mathbf{X}(t)$ and $\mathbf{V}(t)$ are the coordinates and velocity of an impurity particle. Coefficient λ takes into account viscous friction, and $\mathbf{g}$ stands for gravity and buoyancy forces. It is convenient to pass from these equations to relations describing the deviations of the coordinates and velocity of the particle from uniform motion:

$$\mathbf{Y}(t) = \mathbf{X}(t) - \mathbf{v}t, \quad \mathbf{W}(t) = \mathbf{V}(t) - \mathbf{v}.$$

Here, $\mathbf{v} = \mathbf{g}/\lambda$ is the steady-state velocity of the particle falling in a stationary medium. The equations for $\mathbf{Y}$ and $\mathbf{W}$ have the form

$$\frac{d\mathbf{Y}}{dt} = \mathbf{W}, \quad \frac{d\mathbf{W}}{dt} + \lambda\mathbf{W} = \lambda\mathbf{u}(\mathbf{Y} + \mathbf{v}t, t). \qquad (1)$$

To close the diffusion equation, we will require an approximate solution of Eq. (1) in the time interval θ, which will be determined at a later stage. We represent $\mathbf{Y}$ and $\mathbf{W}$ in the form

$$\mathbf{Y}(t) = \mathbf{Y}^0(t) + \tilde{\mathbf{Y}}(t), \quad \mathbf{W}(t) = \mathbf{W}^0(t) + \tilde{\mathbf{W}}(t),$$

where $\mathbf{Y}^0(t')$ and $\mathbf{W}^0(t')$ describe the motion of a parti-

cle in a stationary medium and satisfy the equations

$$\frac{d\mathbf{Y}^0}{dt'} = \mathbf{W}^0, \quad \frac{d\mathbf{W}^0}{dt'} + \lambda\mathbf{W}^0 = 0,$$
$$\mathbf{Y}^0(t-\theta) = \mathbf{Y}(t-\theta), \quad \mathbf{W}^0(t-\theta) = \mathbf{W}(t-\theta), \tag{2}$$

while $\tilde{\mathbf{Y}}(t')$ and $\tilde{\mathbf{W}}(t')$ are solutions to the equations

$$\frac{d\tilde{\mathbf{Y}}}{dt'} = \tilde{\mathbf{W}},$$
$$\frac{d\tilde{\mathbf{W}}}{dt'} + \lambda\tilde{\mathbf{W}} = \lambda\mathbf{u}(\mathbf{Y}^0(t') + \tilde{\mathbf{Y}}(t') + \mathbf{v}t', t'), \tag{3}$$
$$\tilde{\mathbf{Y}}(t-\theta) = 0, \quad \tilde{\mathbf{W}}(t-\theta) = 0.$$

These equations take into account the effect of turbulence on the motion of the particle in a time interval $t' \in [t-\theta, t]$. Let l_u be the external scale of the turbulent velocity field $\mathbf{u}(\mathbf{x}, t)$. We will henceforth assume that the inequality

$$|\tilde{Y}| \ll l_u \tag{4}$$

holds, which means that the "turbulent component" $\tilde{\mathbf{Y}}$ of the motion of the particle during time θ is much smaller than l_u. If inequality (4) is valid, Eqs. (3) can be replaced by the simpler approximate equations

$$\frac{d\tilde{\mathbf{Y}}}{dt'} = \tilde{\mathbf{W}}, \quad \frac{d\tilde{\mathbf{W}}}{dt'} + \lambda\tilde{\mathbf{W}} = \lambda\mathbf{u}(\mathbf{Y}^0(t') + \mathbf{v}t', t'), \tag{5}$$
$$\tilde{\mathbf{Y}}(t-\theta) = 0, \quad \tilde{\mathbf{W}}(t-\theta) = 0.$$

Let us derive the laws of motion of a particle in the time interval $t' \in [t-\theta, t]$. It follows from Eq. (2) that the motion of a particle in a stationary medium, which is expressed in terms of its coordinates and velocity at instant t, is described by the equality

$$\mathbf{Y}^0(t') = \mathbf{Y}^0(t) + \frac{1}{\lambda}\mathbf{W}^0(t)(1 - e^{\lambda(t-t')}).$$

Accordingly, the solutions to Eqs. (5), allowing for turbulence of the medium, have the form

$$\tilde{\mathbf{W}}(t) = \lambda\int_{t-\theta}^{t}\mathbf{u}(\mathbf{Y}^0(t') + \mathbf{v}t', t')e^{\lambda(t'-t)}dt',$$
$$\tilde{\mathbf{Y}}(t) = \int_{t-\theta}^{t}\mathbf{u}(\mathbf{Y}^0(t') + \mathbf{v}t', t')(1 - e^{\lambda(t'-t)})dt'. \tag{6}$$

3. DERIVATION OF THE DIFFUSION EQUATION

Let us derive the equation for the combined probability density

$$f(\mathbf{y}, \mathbf{w}; t) = \langle\delta(\mathbf{Y}(t) - \mathbf{y})\delta(\mathbf{W}(t) - \mathbf{w})\rangle \tag{7}$$

of coordinates and velocity of particles, proceeding from the approximate description (6) of the motion of particles in the time interval $t' \in [t-\theta, t]$. The angle brackets in Eq. (7) and below indicate statistical averaging over an ensemble of the random velocity field $\mathbf{u}(\mathbf{x}, t)$ of the medium.

We assume that $\theta = 2\tau_u$, where τ_u is the time of statistical dependence of field $\mathbf{u}(\mathbf{x}, t)$. Passing from Eqs. (1) to the equation for the average of function $\varphi(\mathbf{Y}(t), \mathbf{W}(t))$, where $\varphi(\mathbf{y}, \mathbf{w})$ is a smooth function whose form will be specified below, we can easily show that the sought average satisfies the following equation:

$$\frac{\partial\langle\varphi\rangle}{\partial t} - \langle\mathbf{W}\cdot\nabla_Y\varphi\rangle + \lambda\langle\mathbf{W}\cdot\nabla_W\varphi\rangle = \lambda\langle\mathbf{u}(\mathbf{Y} + \mathbf{v}t, t)\cdot\nabla_W\varphi\rangle. \tag{8}$$

All averages on the left-hand side of this equation are unambiguously determined by probability density $f(\mathbf{y}, \mathbf{w}; t)$ (7). In order to pass from Eq. (8) to a closed equation for distribution f, we must close the average on the right-hand side of Eq. (8) relative to this distribution. We can do that, taking function $\varphi(\mathbf{Y}(t), \mathbf{W}(t))$ in the form

$$\varphi(\mathbf{Y}, \mathbf{W}) \sim \exp\left(-\frac{(\mathbf{Y}-\mathbf{y})^2}{l_\varphi^2} - \frac{(\mathbf{W}-\mathbf{w})^2}{w_\varphi^2}\right) \tag{9}$$

and assuming that the following inequalities hold:

$$|\tilde{Y}| \ll l_\varphi, \quad |\tilde{W}| \ll w_\varphi. \tag{10}$$

It should be noted that function φ plays approximately the same role as a "physically infinitesimal volume" in the transition from integral conservation laws to macroscopic equations of hydrodynamics. This functions sets the resolution limits $l_f \gtrsim l_\varphi$ and $w_f \gtrsim w_\varphi$ for probability density (7) in the phase space $\{\mathbf{y}, \mathbf{w}\}$, in which the corollaries of diffusion equations (17) are valid.

When inequalities (4) and (10) are satisfied, the average on the right-hand side of Eq. (8) can be replaced by the approximate expression

$$\langle\mathbf{u}(\mathbf{Y}(t) + \mathbf{v}t, t)\cdot\nabla_W\varphi(\mathbf{W}(t), \mathbf{Y}(t))\rangle$$
$$= \langle\mathbf{u}(\mathbf{Y}^0(t) + \mathbf{v}t, t)\cdot\nabla_{W^0}\varphi(\mathbf{W}^0(t), \mathbf{Y}^0(t))\rangle$$
$$+\left\langle\tilde{Y}_j(t)\frac{\partial u_i(\mathbf{Y}^0(t) + \mathbf{v}t, t)}{\partial Y_j^0(t)}\frac{\partial\varphi(\mathbf{Y}^0(t), \mathbf{W}^0(t))}{\partial W_i^0(t)}\right\rangle \tag{11}$$
$$+\left\langle u_i(\mathbf{Y}^0(t) + \mathbf{v}t, t)\tilde{\mathbf{W}}_j(t)\frac{\partial^2\varphi(\mathbf{Y}^0(t), \mathbf{W}^0(t))}{\partial W_i^0(t)\partial W_j^0(t)}\right\rangle$$
$$+\left\langle u_i(\mathbf{Y}^0(t) + \mathbf{v}t, t)\tilde{Y}_j(t)\frac{\partial^2\varphi(\mathbf{Y}^0(t), \mathbf{W}^0(t))}{\partial W_i^0(t)\partial Y_j^0(t)}\right\rangle.$$

The first average on the right-hand side of Eq. (11) is equal to zero in view of the statistical independence of

coordinates $\mathbf{Y}^0(t)$ and velocity $\mathbf{W}^0(t)$ of field $\mathbf{u}(\mathbf{x}, t)$. The remaining averages can be calculated using the same algorithm; consequently, we will consider here only the first of these averages. We carry out averaging in two stages, writing the average in the form

$$\left\langle \tilde{Y}_j(t)\frac{\partial u_i(\mathbf{Y}^0(t)+\mathbf{v}t, t)}{\partial Y_j^0(t)}\frac{\partial \varphi(\mathbf{Y}^0(t), \mathbf{W}^0(t))}{\partial W_i^0(t)}\right\rangle$$
$$= \left\langle \left\langle \tilde{Y}_j(t)\frac{\partial u_i(\mathbf{Y}^0(t)+\mathbf{v}t, t)}{\partial Y_j^0(t)}\right\rangle_{\mathbf{Y}^0(t),\,\mathbf{W}^0(t)} \frac{\partial \varphi(\mathbf{Y}^0(t), \mathbf{W}^0(t))}{\partial W_i^0(t)}\right\rangle.$$

The inner average on the right-hand side can be calculated provided that $\mathbf{Y}^0(t)$ and $\mathbf{W}^0(t)$ are defined, and the outer angle brackets indicate averaging over random values of $\mathbf{Y}^0(t)$ and $\mathbf{W}^0(t)$. Let us calculate the conditional average, substituting in it explicit expression (6) for $\tilde{\mathbf{Y}}(t)$:

$$\left\langle \tilde{Y}_j(t)\frac{\partial u_i(\mathbf{Y}^0(t)+\mathbf{v}t, t)}{\partial Y_j^0(t)}\right\rangle_{\mathbf{Y}^0,\,\mathbf{W}^0}$$
$$= \int_{t-2\tau_u}^{t}\left\langle u_j(\mathbf{Y}^0(t')+\mathbf{v}t', t')\frac{\partial u_i(\mathbf{Y}^0(t)+\mathbf{v}t, t)}{\partial Y_j^0(t)}\right\rangle_{\mathbf{Y}^0,\,\mathbf{W}^0}$$
$$\times (1-e^{\lambda(t'-t)})dt'.$$

Noting that, for $t' < t-\tau_u$, the average in the integrand is equal to zero, while for $t' > t-\tau_u$, the values of field $\mathbf{u}(\mathbf{x}, t')$ are statistically independent of the conditions, we have

$$\left\langle \tilde{Y}_j(t)\frac{\partial u_i(\mathbf{Y}^0(t)+\mathbf{v}t, t)}{\partial Y_j^0(t)}\right\rangle_{\mathbf{Y}^0,\,\mathbf{W}^0} = C_i(\mathbf{W}^0(t))$$
$$= \int_0^\infty \frac{\partial}{\partial s_j}B_{ij}\left(\mathbf{s}+\mathbf{v}\tau+\frac{1}{\lambda}\mathbf{W}^0(t)(e^{\lambda\tau}-1), \tau\right)\Big|_{\mathbf{s}=0}$$
$$\times (1-e^{-\lambda\tau})d\tau. \qquad (12)$$

This expression contains the tensor of the velocity field $\mathbf{u}(\mathbf{x}, t)$, which will be regarded as statistically isotropic. The tensor has the form

$$B_{ij}(\mathbf{s}, \tau) = F(s, \tau)\delta_{ij} + [G(s, \tau) - F(s, \tau)]\frac{s_i s_j}{s^2}. \qquad (13)$$

Here, $F(s)$ and $G(s)$ describe the transverse and longitudinal correlations of the velocity field, respectively. Thus, we finally have

$$\left\langle \tilde{Y}_j(t)\frac{\partial u_i(\mathbf{Y}^0(t)+\mathbf{v}t, t)}{\partial Y_j^0(t)}\frac{\partial \varphi(\mathbf{Y}^0(t), \mathbf{W}^0(t))}{\partial W_i^0(t)}\right\rangle$$
$$\times \langle \mathbf{C}(\mathbf{W}^0(t))\cdot\nabla_{W^0(t)}\varphi(\mathbf{Y}^0(t), \mathbf{W}^0(t))\rangle, \qquad (14)$$

where $\mathbf{C}(\mathbf{W}^0)$ is defined by equality (12). Expressions similar to (14) and (12) can also be obtained easily for the last two averages on the right-hand side of Eq. (11). Namely,

$$\left\langle u_i(\mathbf{Y}^0(t)+\mathbf{v}t, t)\tilde{\mathbf{W}}_j(t)\frac{\partial^2\varphi(\mathbf{Y}^0(t), \mathbf{W}^0(t))}{\partial W_i^0(t)\partial W_j^0(t)}\right\rangle$$
$$= \left\langle \sigma_{ij}^2(\mathbf{W}^0(t))\frac{\partial^2\varphi(\mathbf{Y}^0(t), \mathbf{W}^0(t))}{\partial W_i^0(t)\partial W_j^0(t)}\right\rangle,$$
$$\left\langle u_i(\mathbf{Y}^0(t)+\mathbf{v}t, t)\tilde{Y}_j(t)\frac{\partial^2\varphi(\mathbf{Y}^0(t), \mathbf{W}^0(t))}{\partial W_i^0(t)\partial Y_j^0(t)}\right\rangle$$
$$= \left\langle \left[D_{ij}(\mathbf{W}^0(t)) - \frac{1}{\lambda}\sigma_{ij}^2(\mathbf{W}^0(t))\right]\frac{\partial^2\varphi(\mathbf{Y}^0(t), \mathbf{W}^0(t))}{\partial W_i^0(t)\partial Y_j^0(t)}\right\rangle. \qquad (15)$$

Here, we have

$$\sigma_{ij}^2(\mathbf{W}^0)$$
$$= \lambda\int_0^\infty B_{ij}\left(\mathbf{v}\tau+\frac{1}{\lambda}\mathbf{W}^0(t)(e^{\lambda\tau}-1), \tau\right)e^{-\lambda\tau}d\tau,$$
$$D_{ij}(\mathbf{W}^0)$$
$$= \int_0^\infty B_{ij}\left(\mathbf{v}\tau+\frac{1}{\lambda}\mathbf{W}^0(t)(e^{\lambda\tau}-1), \tau\right)d\tau. \qquad (16)$$

Substituting Eqs. (14) and (15) in Eq. (11) and Eq. (11) into Eq. (8), we arrive at the equation

$$\frac{\partial\langle\varphi\rangle}{\partial t} - \langle\mathbf{W}\cdot\nabla_Y\varphi\rangle + \lambda\langle[\mathbf{W}-\mathbf{C}(\mathbf{W})]\cdot\nabla_W\varphi\rangle$$
$$= \lambda\left\langle \sigma_{ij}^2(\mathbf{W})\frac{\partial^2\varphi}{\partial W_i\partial W_j}\right\rangle$$
$$+\left\langle [\lambda D_{ij}(\mathbf{W}) - \sigma_{ij}^2(\mathbf{W})]\frac{\partial^2\varphi}{\partial W_i\partial Y_j}\right\rangle.$$

Here, by virtue of inequalities (10), processes $\mathbf{Y}^0(t)$ and $\mathbf{W}^0(t)$ on the right-hand side have been replaced by processes $\mathbf{Y}(t)$ and $\mathbf{W}(t)$. It follows from the form of function $\varphi(\mathbf{Y}(t), \mathbf{W}(t))$ (9) and from the last equality that, if the probability density (7) has characteristic scales in $\mathbf{y}$ and $\mathbf{w}$ much larger than l_φ and w_φ, it obeys the diffusion equation

$$\frac{\partial f}{\partial t} + \mathbf{w}\nabla_y f = \lambda(\nabla_w\cdot[\mathbf{w}-\mathbf{C}(\mathbf{w})]f)$$
$$+\lambda\frac{\partial^2(\sigma_{ij}^2(\mathbf{w})f)}{\partial w_i\partial w_j} + \frac{\partial^2([\lambda D_{ij}(\mathbf{w}) - \sigma_{ij}^2(\mathbf{w})]f)}{\partial w_i\partial y_j}. \qquad (17)$$

It should also be noted that, in the case of the vortex velocity field $\mathbf{u}(\mathbf{x}, t)$ in question, when the longitudinal

and transverse components of the correlation tensor are connected through the Kármán relation [9]

$$F(s, \tau) = G(s, \tau) + \frac{s}{2}G'(s, \tau), \tag{18}$$

the identity $\mathbf{C}(\mathbf{W}) \equiv 0$ holds.

Let us consider the methods of measurement and the geometrical meaning of solutions to Eq. (17). It should be recalled that the probability density f (7) satisfying this equation is defined as the microscopic density of an impurity particle averaged over the ensemble of realizations of the random field $\mathbf{u}(\mathbf{x}, t)$. Consequently, generally speaking, we must trace the motion of a single particle to measure f. After quite a long time $T \gg \tau_u$, we must trace the motion of the next particle, which is independent of the previous particle, and repeat this experiment many times to gather the statistics of realizations sufficient for a more or less accurate determination of average (7). Another method of visualization of the solution to Eq. (17) is based on the fact that the average density of a cloud of impurity particles in the phase space $\{\mathbf{y}, \mathbf{w}\}$ also satisfies Eq. (17). In particular, the average density of particles in the ordinary space (and in the case of isotropic turbulence) is given by

$$\langle \rho(\mathbf{y}, t)\rangle = \int \rho_0(\mathbf{y}') f(\mathbf{y} - \mathbf{y}', \mathbf{w}; t) d^3w d^3y'.$$

Here, $\rho_0(\mathbf{x})$ is the initial density of the cloud of particles. However, the probability density can be measured as the average density of a cloud of particles at a given instant t provided that stringent spatial ergodicity conditions are satisfied. First of all, the size of the cloud must be much larger than the external scale of turbulence, so that the motion of an individual particle is virtually independent of the motion of most of the remaining particles in the cloud. In addition, the realizations of the density field of impurity particles are known to vary in space more and more strongly with time, exhibiting alternation and impurity clusterization effects (see, for example, [3–5]). For this reason, additional spatial averaging of a measured realization of density $\rho(\mathbf{y}, t)$ over regions with sizes on the order of the external scale of turbulence is required for measuring the average density $\langle \rho(\mathbf{y}, t)\rangle$, which varies in space more and more smoothly.

4. UNIFORM FALL APPROXIMATION

Let us consider in detail the corollaries of Eq. (17) in the uniform fall approximation. Namely, while calculating the coefficients in Eq. (17), we disregard deviation $\mathbf{W}$ of the velocity of a particle from $\mathbf{v}$. We carry out analysis in the frame of reference falling at velocity $\mathbf{v}$. Let $\mathbf{v}$ be directed along the $y_1 = y$ axis. Similarly, we omit the subscript on the relevant component of vector $\mathbf{w}$: $w_1 = w$. It can easily be proved that, in this reference frame and in the uniform fall approximation, Eq. (17) is transformed into

$$\begin{aligned}\frac{\partial f}{\partial t} + (\mathbf{w}_\perp \cdot \nabla_{\perp y}) f + w\frac{\partial f}{\partial y} &= \lambda(\nabla_{\perp w} \cdot \mathbf{w}_\perp f) \\ + \lambda\frac{\partial (wf)}{\partial w} + \lambda\sigma_\perp^2\nabla_{\perp w}^2 f + \lambda\sigma_\parallel^2\frac{\partial^2 f}{\partial w^2} \\ + [\lambda D_\perp - \sigma_\perp^2](\nabla_{\perp w} \cdot \nabla_{\perp y}) f + [\lambda D_\parallel - \sigma_\parallel^2]\frac{\partial^2 f}{\partial y \partial w}.\end{aligned} \tag{19}$$

Here, gradients are taken in the plane perpendicular to the vertical y axis, and the following notation is used:

$$\begin{aligned}\sigma_\parallel^2 &= \lambda\int_0^\infty G(v\tau, \tau)e^{-\lambda\tau}d\tau, \\ \sigma_\perp^2 &= \lambda\int_0^\infty F(v\tau, \tau)e^{-\lambda\tau}d\tau, \\ D_\parallel &= \int_0^\infty G(v\tau, \tau)d\tau, \\ D_\perp &= \int_0^\infty F(v\tau, \tau)d\tau.\end{aligned} \tag{20}$$

Formulas (19) and (20) show that the diffusion of a particle in the direction of its incidence and in the perpendicular plane is not the same: diffusion in the direction of incidence is determined by the correlation function of the longitudinal components of the velocity field of the medium, while transverse diffusion is determined by the correlation function of the transverse component. We will consider separately diffusion in different directions. Averaging Eq. (19) over transverse coordinates and velocity, we arrive at the longitudinal diffusion equation

$$\begin{aligned}\frac{\partial f_\parallel}{\partial t} + w\frac{\partial f_\parallel}{\partial y} &= \lambda\frac{\partial (w f_\parallel)}{\partial w} \\ + \lambda\sigma_\parallel^2\frac{\partial^2 f_\parallel}{\partial w^2} + [\lambda D_\parallel - \sigma_\parallel^2]\frac{\partial^2 f_\parallel}{\partial y \partial w}\end{aligned} \tag{21}$$

in the probability density

$$f_\parallel(y, w; t) = \langle \delta(Y(t) - y)\delta(W(t) - w)\rangle,$$

where $Y(t)$ and $W(t)$ are the components of the particle coordinates and velocity in the direction of incidence. Equation (21) can easily be solved in the general case. We confine our analysis to the physically most important case of $\lambda t \gg 1$, which corresponds to the steady-state average velocity of the particle and a linear law of

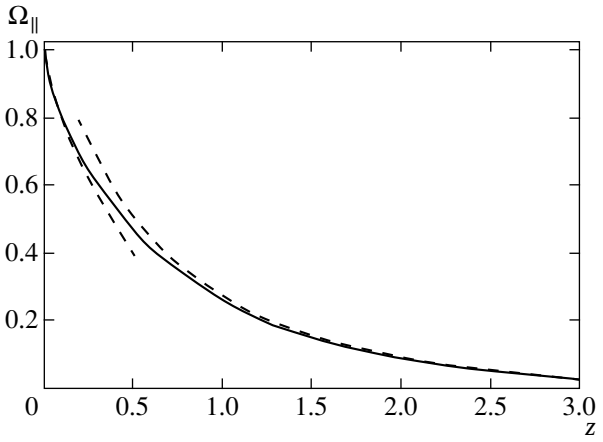

Fig. 1.

diffusion. Over such time intervals, the probability density of the longitudinal velocity component $W(t)$ coincides with the steady-state distribution $f_{st}(w)$ satisfying the equation

$$w f_{st}(w) + \sigma_\parallel^2(v, \varepsilon)\frac{d f_{st}(w)}{dw} = 0.$$

Here, the dependence of coefficient $\sigma_\parallel^2(v, \varepsilon)$ on the velocity of fall and on the inertia characterized by the dimensionless parameter $\varepsilon = \lambda\tau_u$ is indicated explicitly. The larger this parameter, the less inertial the particle motion. In particular, for $\varepsilon \gg 1$, the particle velocity is approximately equal to $\mathbf{V}(t) \approx \mathbf{v} + \mathbf{u}(\mathbf{X}(t), t)$.

It follows from the latter equation that the longitudinal velocity of the particle obeys Maxwell's distribution:

$$f_{st}(w) = \frac{1}{\sqrt{2\pi\sigma_\parallel^2(v, \varepsilon)}}\exp\left(-\frac{w^2}{2\sigma_\parallel^2(v, \varepsilon)}\right). \qquad (22)$$

We will get a more comprehensive idea of the statistics of the longitudinal velocity of a particle by studying the dependence of its variance $\sigma_\parallel^2$ on the average velocity v of fall and on parameter ε. As the first step, we choose the longitudinal correlation function $G(s, \tau)$ reflecting spatial properties inherent in strong turbulence. We take this function in the form

$$G(s, \tau) = \sigma_u^2 G(s)\exp\left(-\frac{|\tau|}{\tau_u}\right), \qquad (23)$$

where $G(s)$ is the spatial correlation function taking into account the presence of the inner scale of turbulence l_0, its external scale l_u ($l_u \gg l_0$), and the Kolmogorov–Obukhov two-thirds power law in the inertia interval $l_0 \ll s \ll l_u$. It is most convenient to take into account these features of the turbulence correlation function in terms of the corresponding spectral density,

$$\tilde{G}(\kappa) = \frac{1}{(2\pi)^3}\int_{-\infty}^{\infty} G(s)e^{-i(\boldsymbol{\kappa}\cdot\mathbf{s})}d^3s.$$

The spectral density corresponding to strong turbulence properties is defined by the Kármán spectrum

$$\tilde{G}(\kappa) = AL^3\frac{\exp(-\kappa^2/\kappa_0^2)}{(\kappa^2 l_u^2 + 1)^{11/6}}, \quad \kappa_0 = \frac{5.92}{l_0},$$

where A is chosen from the condition $G(0) = 1$. While analyzing absolute diffusion of particles treated by us here and in the case of strong turbulence ($l_0 \ll l_u$), we can apparently disregard the effect of inner scale. Consequently, we set $l_0 = 0$ and use a simplified Kármán spectrum:

$$\tilde{G}(\kappa) = AL^3(\kappa^2 l_u^2 + 1)^{-11/6}.$$

The correlation function corresponding to this spectrum is given by

$$G(s) = \Omega_\parallel\left(\frac{s}{l_u}\right),$$
$$\Omega_\parallel(z) = \frac{\sqrt{3}}{\pi}\Gamma\left(\frac{2}{3}\right)\left(\frac{z}{2}\right)^{1/3}K_{1/3}(z). \qquad (24)$$

Here,

$$A = \frac{\sqrt{3}}{2\pi^2\sqrt{\pi}}\Gamma\left(\frac{2}{3}\right)\Gamma\left(\frac{11}{6}\right) \approx 0.063.$$

For $z \longrightarrow 0$, correlation function (24) possesses the asymptotic form

$$\Omega_\parallel(z) \sim 1 - \frac{3\sqrt{\pi}}{\Gamma(1/6)}z^{2/3} \quad (z \longrightarrow 0),$$

corresponding to the Kolmogorov–Obukhov two-thirds power law, while for large values of z, this function decreases exponentially to zero:

$$\Omega_\parallel(z) \sim \sqrt{\frac{3}{4\pi}}\Gamma\left(\frac{2}{3}\right)\left(\frac{2}{z}\right)^{1/6}e^{-z} \quad (z \gtrsim 1).$$

Figure 1 shows the graph of function $\Omega_\parallel(z)$ (24) together with its asymptotic form.

From Eqs. (20), (23), and (24), we have

$$\sigma_\parallel^2(v, \varepsilon) = \sigma_u^2 \varepsilon \Phi(\gamma, \varepsilon),$$
$$\Phi(\gamma, \varepsilon) = \frac{1}{1+\varepsilon} F\left(\frac{1}{2}, 1, \frac{2}{3}, \frac{\gamma^2}{(1+\varepsilon)^2}\right) - \frac{\sqrt{3}\Gamma^3\left(\frac{2}{3}\right)}{\pi[(1+\varepsilon)^2 - \gamma^2]^{5/6}}\left(\frac{\gamma}{2}\right)^{2/3}. \quad (25)$$

These expressions contain the dimensionless parameter $\gamma = v/\tilde{u}$, where $\tilde{u} = l_u/\tau_u$. The curves describing variance $\sigma_\parallel^2$ of the longitudinal velocity of a particle, normalized to σ_u^2, as a function of ε for different values of γ are shown in Fig. 2. It can be seen that, for a low inertia ($\varepsilon \gg 1$), the variance of the particle velocity tends to the variance of components σ_u^2 of the velocity field of the medium for any γ. As inertia of the motion of a particle increases (for $\varepsilon \longrightarrow 0$), the variance of its velocity tends to zero owing to the averaging effect of inertia.

Let us derive the law governing longitudinal diffusion of a particle, passing from formula (21) to the equation for average $\langle Y^2(t)\rangle$. Relation (21) implies that

$$\frac{d\langle Y^2(t)\rangle}{dt} = 2\langle Y(t)W(t)\rangle. \quad (26)$$

The average on the right-hand side, in turn, obeys the equation

$$\frac{d\langle Y(t)W(t)\rangle}{dt} + \lambda\langle Y(t)W(t)\rangle = \langle W^2(t)\rangle + \lambda D_\parallel - \sigma_\parallel^2.$$

Considering that the mean square of the relative velocity tends to steady-state variance (25) in the case $\lambda t \gg 1$ of interest to us,

$$n\langle W^2(t)\rangle \longrightarrow \sigma_\parallel^2,$$

and the average $\langle Y(t)W(t)\rangle$ tends to the steady-state value

$$\langle YW\rangle_{\rm st} = D_\parallel(\gamma),$$

we obtain the following asymptotic law of diffusion from Eq. (26):

$$\langle Y^2(t)\rangle = 2D_\parallel(\gamma)t. \quad (27)$$

Here, we have explicitly taken into account the dependence of the longitudinal diffusion coefficient on the velocity of fall (parameter γ). In our opinion, the independence of the diffusion coefficient in Eq. (27) of the particle inertia (parameter ε) is even more remarkable. This appears as paradoxical since the variance of the particle velocity decreases upon an increase in inertia (with decreasing ε). Apparently, this should suppress diffusion. The paradox is removed by the fact that, with increasing inertia, the correlation time τ_v for the particle velocity increases in the same proportion. As a result, the product $\sigma_\parallel^2\tau_v$, which determines the diffusion coefficient, turns out to be independent of ε. It should be noted in addition that, for the longitudinal correlation function of turbulent velocity field (23), (24), we have

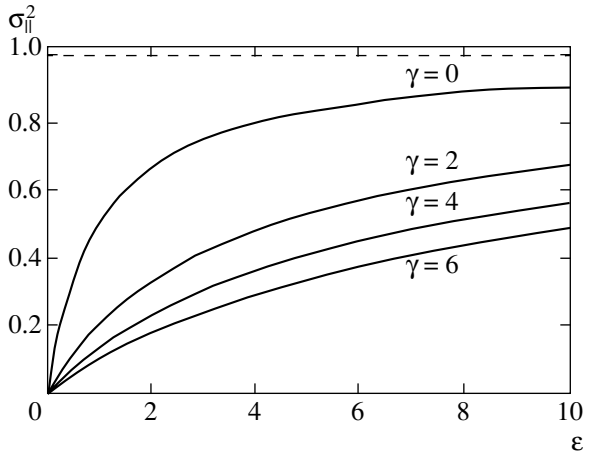

Fig. 2.

$$D_\parallel(\gamma) = D_0\Phi(\gamma), \quad D_0 = \sigma_u^2\tau_u, \quad \Phi(\gamma) = \Phi(\gamma, \varepsilon = 0), \quad (28)$$

while function $\Phi(\gamma, \varepsilon)$ is defined by formula (25).

5. TRANSVERSE DIFFUSION

The diffusion of a particle in a plane transverse to the direction of incidence differs from the longitudinal diffusion in the dependence of the velocity variance and the transverse diffusion coefficient on parameters γ and ε. This is due to the fact that transverse diffusion is determined by the transverse correlation function F of the medium and not by the longitudinal function G. In accordance with Eq. (18), function F in the case when Eqs. (23) and (24) hold is given by

$$F(s, \tau) = \sigma_u^2\Omega_\perp\left(\frac{s}{l_u}\right),$$
$$\Omega_\perp(z) = \frac{\Gamma\left(\frac{2}{3}\right)}{4\pi\sqrt{3}}\left(\frac{z}{2}\right)^{1/3} \times [14K_{1/3}(z) - 3zK_{2/3}(z) - 3zK_{4/3}(z)]. \quad (29)$$

In contrast to $\Omega_\parallel(z)$ (24), function $\Omega_\perp(z)$ decreases at a higher rate and becomes negative for large values of z.

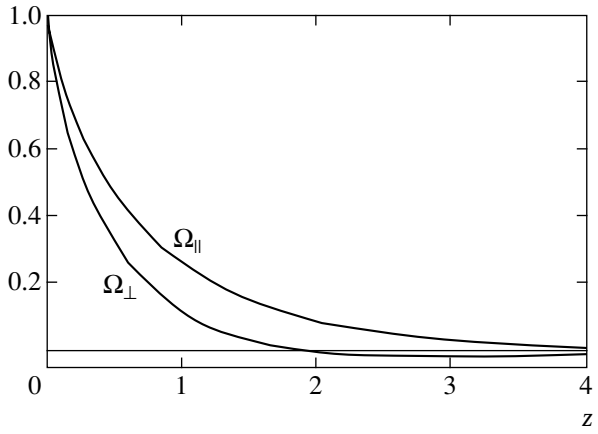

Fig. 3.

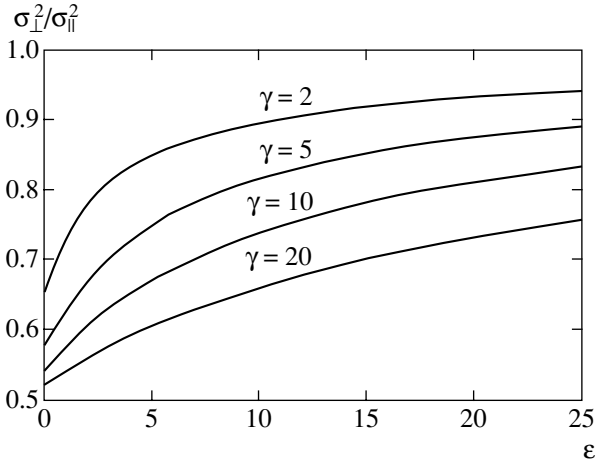

Fig. 4.

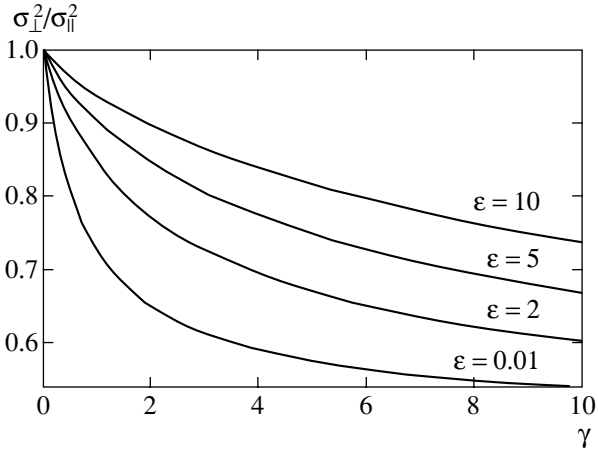

Fig. 5.

The $\Omega_\parallel(z)$ and $\Omega_\perp(z)$ curves illustrating the difference between the transverse and longitudinal correlation functions of the vortex velocity field are depicted in Fig. 3. The above-mentioned features of transverse correlations of the vortex velocity field are responsible for qualitative and even quantitative difference between transverse and longitudinal diffusion. Let us compare above all the transverse and longitudinal variances of the particle velocity over time periods of steady-state diffusion ($\lambda t \gg 1$). In this case, the steady-state longitudinal variance of the velocity is given by Eq. (25), while the corresponding variance of one of the transverse components of the particle velocity, which can be obtained by substituting formula (29) in (20) and evaluating the integral, is given by

$$\sigma_\perp^2(\nu, \varepsilon) = \sigma_u^2 \varepsilon \Psi(\gamma, \varepsilon),$$

$$\Psi(\gamma, \varepsilon) = \frac{1}{6(1+\varepsilon)}\left[7F\left(\frac{1}{2}, 1, \frac{2}{3}, \frac{\gamma^2}{(1+\varepsilon)^2}\right)\right.$$

$$\left. - F\left(\frac{1}{2}, 1, -\frac{1}{3}, \frac{\gamma^2}{(1+\varepsilon)^2}\right)\right] \qquad (30)$$

$$+ \frac{3\gamma^2}{8(1+\varepsilon)^3} F\left(\frac{3}{2}, 2, \frac{5}{3}, \frac{\gamma^2}{(1+\varepsilon)^2}\right)$$

$$- \frac{2^{1/3}\gamma^{2/3}}{4\pi\sqrt{3}}\Gamma^3\left(\frac{2}{3}\right)\frac{8(1+\varepsilon)^2 - 3\gamma^2}{[(1+\varepsilon)^2 - \gamma^2]^{11/6}}.$$

The curves in Fig. 4 describe the ratio of the transverse dispersion of the particle velocity to its longitudinal counterpart as a function of ε for different values of γ. It can be seen that the above features of the transverse correlation function (29) lead to a relative decrease in the transverse variance of the velocity, which is manifested most clearly for a strongly inertial motion ($\varepsilon \longrightarrow 0$).

An increase in the velocity of particle fall (parameter γ) also facilitates a decrease in the transverse velocity variance. This is illustrated in Fig. 5, showing ratio $\sigma_\perp^2/\sigma_\parallel^2$ as a function of γ for different values of ε.

It was noted above that inertia of motion of a particle does not affect its diffusion coefficient. However, the suppression of transverse diffusion as compared to longitudinal diffusion is observed in this case also. Let us illustrate this for the longitudinal correlation function (23) and the corresponding transverse correlation function (29) of the velocity field of the medium. In this case, calculations lead to the following expression for the mean square of transverse deviation of a particle from the trajectory of its incidence in a stationary medium:

$$\langle \mathbf{X}_\perp^2 \rangle = 2D_\perp(\gamma), \quad D_\perp(\gamma) = D_0 \Psi(\gamma), \qquad (31)$$
$$\Psi(\gamma) = \Psi(\gamma, \varepsilon = 0).$$

Figure 6 shows the longitudinal (28) and transverse (31) diffusion coefficients normalized to D_0 as functions

of γ. The dashed curve in the same figure shows the ratio of the transverse and longitudinal diffusion coefficients. It can be seen that, and the velocity of fall v increases, the ratio of the diffusion coefficients decreases smoothly from 1 : 1 to 1 : 2.

It should be noted that the limiting value of the ratio of the diffusion coefficients, which is equal to 1 : 2, is a universal constant independent of the spectral and correlation properties of the velocity field **u**(**x**, *t*). Indeed, in the limiting case of a high velocity of particle fall ($\gamma \gg 1$), we can disregard the time dependence of the correlation functions for the velocity field and use the asymptotic formulas

$$D_\parallel \approx \int_0^\infty G(v\tau, 0)d\tau,$$
$$D_\perp \approx \int_0^\infty F(v\tau, 0)d\tau \quad (\gamma \gg 1), \tag{32}$$

indicating that the particle "pierces" elementary vortices during a time over which the vortices virtually cannot shift. Substituting the Kármán formula (18) into the last relation from (32), we obtain

$$D_\perp \approx D_\parallel + \frac{1}{2}\int_0^\infty \tau \frac{dG(v\tau, 0)}{d\tau}d\tau;$$

after integrating by parts, we arrive at the following universal limiting relation:

$$D_\perp \approx \frac{1}{2}D_\parallel \quad (\gamma \gg 1).$$

6. CONDITIONS FOR APPLICABILITY OF DIFFUSION EQUATION

It should be recalled that, while deriving the diffusion equation (17), we required that the effect of turbulence on the motion of a diffusing particle could be neglected for the correlation time τ_u of the turbulent velocity field of the medium. Let us estimate the validity of this assumption, replacing expression (4) by the inequality for averages,

$$\langle Y^2(\tau_u)\rangle \ll l_u^2,$$

and substituting the mean square $\langle Y^2(\tau_u)\rangle$ of the increment of the longitudinal coordinate over time τ_u (27),

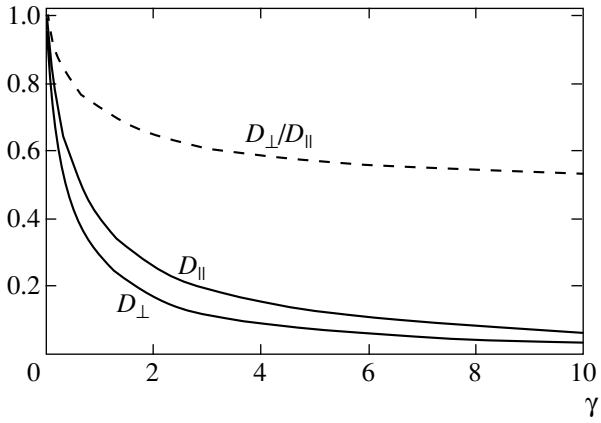

Fig. 6.

(28) in this inequality. This leads to the following condition of applicability of Eq. (17):

$$\sigma_u^2\Phi(\gamma) \ll \tilde{u}^2, \quad \tilde{u} = l_u/\tau_u. \tag{33}$$

The asymptotic form

$$\Phi(\gamma) \sim \frac{a}{\gamma} \quad (\gamma \gtrsim 1), \quad a = \frac{3}{2\pi 2^{2/3}}\Gamma\left(\frac{2}{3}\right) \approx 0.747,$$

shows that, in the case of $\sigma_u \approx \tilde{u}$ typical of turbulence, inequality (33) holds for quite large values of γ.

The freefall approximation used in the above analysis is valid if we can disregard the dependence of coefficients (16) on W^0. Let us see when this can be done as applied to the diffusion tensor D_{ij}. The characteristic value of W^0 is determined by the effective width of the Maxwell distribution (22): $W^0 \approx \sigma_\parallel$. In addition, we assume that the value of γ is large enough for the effective interval of integration in expression (16) for D_{ij} to be equal to $\tau^* \approx \tau_u/\gamma$. In this case, the dependence of D_{ij} (16) on W^0 can be neglected if

$$\sigma_u^2\Phi(\gamma, \varepsilon)\left[\exp\left(\frac{\varepsilon}{\gamma}\right) - 1\right]^2 \ll \varepsilon^2\tilde{u}^2.$$

In particular, for $\varepsilon \lesssim \gamma$, this condition can be reduced to the inequality

$$\sigma_u^2\Phi(\gamma, \varepsilon) \ll \tilde{u}^2\gamma^2,$$

which obviously holds for large values of γ.

ACKNOWLEDGMENTS

This study was supported financially by the Russian Foundation for Basic Research (project

nos. 00-02-16167 and 00-15-96619) and Ben Gurion University, Israel.

REFERENCES

1. G. T. Csanady, *Turbulent Diffusion in the Environment* (Reidel, Dordrecht, 1980).
2. M. B. Isichenko, Rev. Mod. Phys. **64**, 961 (1992).
3. V. S. L'vov, I. Procaccia, and A. L. Fairhall, Phys. Rev. E **50**, 4684 (1994).
4. V. I. Klyatskin and A. I. Saichev, Zh. Éksp. Teor. Fiz. **111**, 1297 (1997) [JETP **84**, 716 (1997)].
5. I. S. Zhukova and A. I. Saichev, Prikl. Mat. Mekh. **61**, 788 (1997).
6. A. I. Saichev and I. S. Zhukova, Lect. Notes Phys. **511**, 353 (1998).
7. M. R. Maxey, J. Fluid Mech. **174**, 441 (1987).
8. G. T. Csanady, J. Atmos. Sci. **20**, 201 (1963).
9. A. S. Monin and A. M. Yaglom, *Statistical Fluid Mechanics* (Nauka, Moscow, 1965; MIT Press, Cambridge, 1971), Part 1.

Translated by N. Wadhwa

Journal of Experimental and Theoretical Physics, Vol. 96, No. 3, 2003, pp. 489–495.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 123, No. 3, 2003, pp. 552–559.

FLUIDS

Phase Change of a Plane Wave Propagating through Polymer Dispersed Liquid Crystal Film with Nanosized Droplets

V. A. Loiko* and A. V. Konkolovich

Stepanov Institute of Physics, National Academy of Sciences of Belarus, Minsk, 220072 Belarus
**e-mail: loiko@dragon.bas-net.by*

Received October 3, 2002

Abstract—The coherent field propagation in a polymer dispersed liquid crystal layer is described using a method based on the Foldy–Twersky integral equation for the vector case. Expressions for a polarization-independent phase shift and the coherent transmission coefficient of such a layer containing nanodimensional nematic liquid crystal droplets are obtained. Theoretical results for the phase shift are compared to the available experimental data. © *2003 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

Polymer dispersed liquid crystal (PDLC) films, which appeared in the late 1980s, are promising materials for the development of electric-field-controlled radiation modulators. In recent years, extensive research has been carried out aimed at the creation of displays, diffraction gratings, lenses, polarizers, and other field-controlled optical elements based on PDLC films [1–7].

In a PDLC film system, particles of an optically anisotropic liquid crystal (nematic, ferroelectric, cholesteric, or a mixture of these [8–13]) are dispersed in an isotropic binding polymer matrix confined between two transparent substrates with transparent electrodes formed on their inner surfaces. Special features of such film systems are that they possess a disperse structure and require no additional polaroids for producing electrooptical effects. The dimensions of liquid crystal droplets may vary from 0.05 to 3.0 μm.

PDLC films with a particle size exceeding light wavelength exhibit strong light scattering. The transmission of such films depends on the applied field rotating optical axes (directors) of the LC droplets and/or changing the internal structure of these particles. In PDLC systems with a particle size much smaller than the light wavelength, the scattering is small. In this case, description of light propagation reduces in the first approximation to an analysis of the amplitude and phase of a coherent (guided) [14] component of the transmitted radiation. Such films are usually characterized by a high transmission, which is not strongly influenced by changes on the LC droplet director orientation. Here, the main field effect is a change in the effective refractive index of the film (related to the Kerr effect) under the action of an applied electric field [15, 16]. Recent interest in the investigation of films possessing a fine disperse structure is related to prospects for obtaining polarization-independent light phase modulators [17, 18]. Such films would be promising materials for field-controlled spectral filters [17] and telecommunication systems [18].

An important task encountered in the study of fine disperse films is the development of methods for the description of light propagation in such systems, relating characteristics of the transmitted light (wave amplitude and phase) to the material parameters. Solving this problem would allow us to predict new electrooptical effects—in particular, in ferroelectric PDLC films characterized by small switching times [1–4]—and develop methods for the obtaining of films with the desired properties.

This paper proposes a model for description of the propagation of a coherent light field in a PDLC film. The model is based on solving the problem of propagation of a plane polarized wave through such a layer in terms of the Foldy–Twersky integral equation [19]. Using this model, formulas for the coherent transmission coefficient, light attenuation coefficient, and field-controlled polarization-independent phase shift in a PDLC film with nanodimensional nematic LC droplets are obtained in the Rayleigh approximation. Theoretical results for the phase shift are compared to the experimental data of Lucchetta *et al.* [20].

2. COHERENT FIELD OF A LINEARLY POLARIZED PLANE WAVE PROPAGATING TROUGH A POLYMER ENCAPSULATED LIQUID CRYSTAL LAYER

Consider a linearly polarized plane light wave with amplitude $\mathbf{E}_0(|\mathbf{E}_0| = E_0)$ and wave vector $\mathbf{k}$ normally

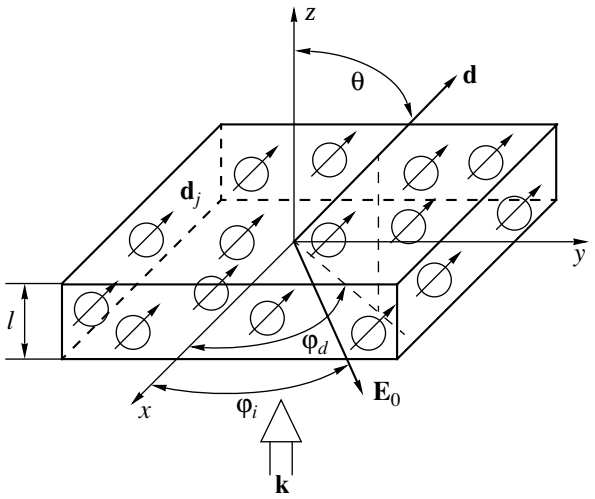

Fig. 1. Geometry of the plane wave propagating through a PDLC film: **k** is the wave vector of the incident wave; $\mathbf{d}_j$ is the director of an LC droplet; **d** is the common direction of orientation of the LC droplet directors; φ_d and φ_i are the angles determining orientation of the principal plane (**k**, **d**) and the polarization plane of the incident wave (**k**, $\mathbf{E}_0$) in the coordinate system (*x*, *y*, *z*); $\mathbf{E}_0$ is the electric vector of the incident wave, *l* is the PDLC film thickness, θ is the polar angle between the *z* axis and the LC director.

incident onto a plane-parallel PDLC film of thickness *l* (Fig. 1). Introduce a right-handed rectangular Cartesian coordinate system (*x*, *y*, *z*) with the *z* axis directed along the normal to the PDLC film, parallel to the wave vector **k** of the incident wave, and the (*x*, *y*) plane coinciding with the front surface of the layer.

Let us assume that (i) each LC droplet in the layer is characterized by the director vector $\mathbf{d}_j$ ($j = \overline{1, N}$, where *N* is the number of LC droplets in the PDLC film), which characterizes the orientation of long axes of the LC molecules averaged over a given particle, and (ii) that all particles in the layer possess identical shape and size and are oriented in the same direction parallel to a given vector **d** (i.e., $\mathbf{d}_j \parallel \mathbf{d}$, $j = \overline{1, N}$). Let *VV* and *VH* be the transmitted wave components with the polarizations parallel and perpendicular to the plane of polarization of the incident wave [21, 22]. In terms of the Foldy–Twersky integral equation [19], the *VV* and *VH* components of the electric vector of the coherent electromagnetic wave field transmitted through the PDLC film can be expressed as

$$\langle E_{VV}(z)\rangle = E_0\psi_{VV}(z)\big|_{z=l}, \tag{1}$$

$$\langle E_{VH}(z)\rangle = E_0\psi_{VH}(z)\big|_{z=l}, \tag{2}$$

where $\psi_{VV}(z)$ and $\psi_{VH}(z)$ are given by the formulas

$$\psi_{VV}(z) = \psi_1(z)\sin^2(\varphi_d - \varphi_i) + \psi_2(z)\cos^2(\varphi_d - \varphi_i) + \frac{1}{2}(\psi_3(z) + \psi_4(z))\sin 2(\varphi_d - \varphi_i), \tag{3}$$

$$\psi_{VH}(z) = \frac{1}{2}(\psi_2(z) - \psi_1(z))\sin 2(\varphi_d - \varphi_i) + \psi_3(z)\sin^2(\varphi_d - \varphi_i) - \psi_4(z)\cos^2(\varphi_d - \varphi_i). \tag{4}$$

Here, angular brackets denote statistical averaging over the ensemble, φ_i is the angle between the polarization plane of the wave (**k**, $\mathbf{E}_0$) and the vertical plane (*z*, *x*), φ_d is the angle between the principal plane (**k**, **d**) and the (*z*, *x*) plane, and functions $\psi_j(z)$ ($j = \overline{1, 4}$) are solutions of the system of equations

$$\psi_{1;2}(z) = \exp(ikz) \times \left\{1 - qS_{1;2}(0)\int_0^z \exp(-ikz_s)\psi_{1;2}(z_s)dz_s\right\}, \tag{5}$$

$$\psi_{3;4}(z) = -\exp(ikz)qS_{3;4}(0) \times \int_0^z \exp(-ikz_s)\psi_{3;4}(z_s)dz_s. \tag{6}$$

In these equations, $q = 2\pi k^{-2}N_v$; N_v is the number of LC droplets per unit layer volume; $k = 2\pi/\lambda_p$; λ_p is the light wavelength in the binding polymer matrix; and $\mathbf{S}_j(0)$ ($j = \overline{1, 4}$) are the elements of the amplitude scattering matrix [22] for the zero scattering angle. The latter matrix elements depend on the optical properties of the LC crystal and the shape, size, and structure of the LC droplets.

Equations (1)–(6) represent a generalization of the Foldy–Twersky integral equation [19] for the vector case. Using these equations, it is possible to analyze some effects related to the coherent component of the light transmitted through the PDLC film, including birefringence and the related transformation of the polarization state, rotation of the polarization plane, and linear and circular dichroism. This description requires knowledge of the amplitude scattering matrix elements $\mathbf{S}_j(0)$ ($j = \overline{1, 4}$) and must take into account statistical features in the distribution of LC droplet directors in the PDLC film.

Below we will consider a PDLC film with nanodimensional, nonabsorbing nematic LC droplets, assuming a cylindrical symmetry in the arrangement of LC molecules inside the particles and in the distribution of particle directors in the layer.

3. PHASE SHIFT RELATIONS AND COMPARISON TO EXPERIMENT

Let the configuration of molecular axes in the PC particles be cylindrically symmetric with respect to the average direction of orientation (i.e., relative to the particle directors). In the absence of an external electric field, the distribution of directors in the layer is random. Under the action of an applied field applied along the normal to the layer, directors of a liquid crystal with positive anisotropy will be aligned in the control field direction. The distribution of director orientations is assumed to be cylindrically symmetric relative to the electric vector of the control field (i.e., relative to the z axis). In the Rayleigh approximation for an anisotropic dipole [22] representing the LC droplet, the amplitude scattering matrix elements for polydisperse particles at a zero scattering angle are as follows:

$$S_3(0) = S_4(0) = 0, \tag{7}$$

$$S_1(0) = S_2(0) = S(0), \tag{8}$$

$$S(0) = -\frac{ik^3}{4\pi}\langle v\rangle\left[\frac{h_{do}^2}{n_p^2} - 1 + \frac{n_{de}^2 - n_{do}^2}{3n_p^2}(1 - S_f)\right]. \tag{9}$$

Here, $\langle v\rangle$ is the average volume of the LC droplet; $S_f = (3\langle\cos^2\theta\rangle - 1)/2$ is the order parameter of the PDLC film [2, 23]; θ is the angle between the LC droplet director and the z axis; n_p is the refractive index of the binding polymer matrix;

$$n_{do}^2 = n_{iso}^2 - \frac{1}{3}(n_e^2 - n_o^2)S_d S, \tag{10}$$

$$n_{de}^2 = n_{iso}^2 + \frac{2}{3}(n_e^2 - n_o^2)S_d S \tag{11}$$

are the average refractive indices of the LC droplets for the ordinary and extraordinary waves, respectively [2, 24]; S_d is the order parameter of the LC droplet; S is the molecular order parameter of the liquid crystal; n_o and n_e are the refractive indices of the liquid crystal for the ordinary and extraordinary waves, respectively; and

$$n_{iso}^2 = \frac{2n_o^2 + n_e^2}{3}. \tag{12}$$

Relations (5)–(8) yield

$$\psi_3(z) = \psi_4(z) = 0, \tag{13}$$

$$\psi_1(z) = \psi_2(z) = 0. \tag{14}$$

Averaging relations (3) and (4) with respect to angle φ_d [so that $\langle\cos^2(\varphi_d - \varphi_i)\rangle = \langle\sin^2(\varphi_d - \varphi_i)\rangle = 1/2$ and $\langle\sin 2(\varphi_d - \varphi_i)\rangle = 0$], using formulas (9), (13), and (14), and taking into account cylindrical symmetry of the droplet director distribution, we obtain

$$\psi_{VH}(z) = 0, \tag{15}$$

$$\psi_{VV}(z) = \exp(ikz) \times\left\{1 - qS(0)\int_0^z \exp(-ikz_s)\psi_{VV}(z_s)dz_s\right\}. \tag{16}$$

Relations (15) and (16) show that the problem of determining the amplitude and phase of a normally incident wave transmitted through a PDLC film can be solved by analogy to the scalar case [19, 25].

Solving integral equation (16) with respect to $\psi_{VV}(z)$, we obtain

$$\psi_{VV}(z) = \exp(iKz), \tag{17}$$

where K is the propagation constant defined as

$$K = k + iqS(0). \tag{18}$$

For a coherent field at the output ($z = l$) from the PDLC film, Eqs. (1) and (17) yield (the VV subscript is omitted by virtue of the polarization-independent character of the solution)

$$\langle E\rangle = E_0\exp(iKl). \tag{19}$$

This relation indicates that, in the case under consideration, the film behaves like a homogeneous plane-parallel plate with the complex refractive index

$$\tilde{n} = \frac{K}{k} = 1 + i\frac{2\pi}{k^3}N_v S(0). \tag{20}$$

Let us define a phase shift $\Delta\Phi$ as the phase difference for a wave transmitted with the control electric field switched on (whereby the film order parameter is $S_f > 0$) and off ($S_f = 0$) [20]. Using expressions (9), (12), and (20), we obtain

$$\Delta\Phi = kl[\mathrm{Re}\tilde{n}(S_f) - \mathrm{Re}\tilde{n}(S_f = 0)], \tag{21}$$

or

$$\Delta\Phi = -\frac{1}{6}klc_v\frac{n_e^2 - n_o^2}{n_p^2}SS_dS_f, \tag{22}$$

where $c_v = N_v\langle v\rangle$ is the number density of LC droplets in the PDLC film. In expression (21), symbol Re indicates that we take the real part of the effective refractive index $\tilde{n}$, while the sign minus for $\Delta\Phi$ in formula (22) indicates that the phase velocity of the wave is greater when particles of a liquid crystal with the positive anisotropy ($n_e > n_o$) are oriented along the control field. As the ordering of LC droplet directors (i.,e., the film order parameter S_f) increases, the phase velocity of the wave grows.

Let us compare the above theoretical results obtained for the phase shift to the experimental data

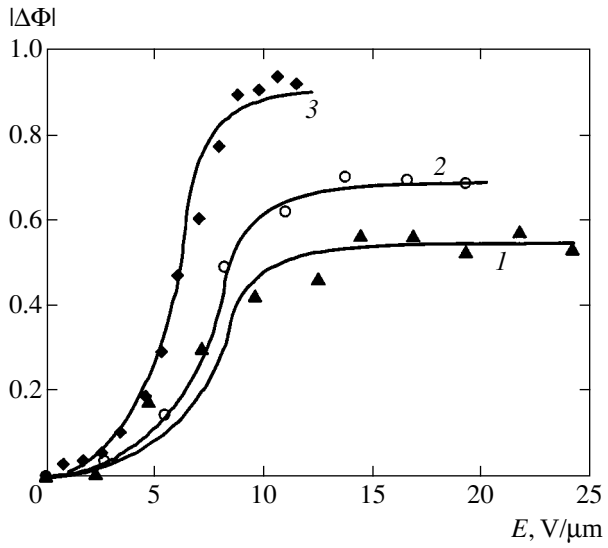

Fig. 2. Plots of the absolute phase shift $\Delta\Phi$ versus the applied electric field strength E calculated by formula (22) (solid curves) in comparison to the experimental data [20] (symbols) for various sets of parameters: (*1*) $l = 13$ μm, $c_v = 0.125$, $A = 0.12$ μm/V; (*2*) $l = 23$ μm, $c_v = 0.089$, $A = 0.124$ μm/V; (*3*) $l = 36$ μm, $c_v = 0.075$, $A = 0.16$ μm/V.

reported by Lucchetta *et al.* [20]. We assume that (i) the order parameter of the LC droplets S_d is independent of the applied field and that (ii) the order parameter of the PDLC film depends on the electric field strength E as [2, 20]

$$S_f = \frac{1}{4} + \frac{3(e^2+1)}{16e^2} + \frac{3(3e^2+1)(e^2-1)}{32e^3}\ln\left|\frac{e+1}{e-1}\right|, \quad (23)$$

$$e = AE, \quad (24)$$

where A is a parameter determined by the average LC droplet size, optical anisotropy $\Delta n = n_e - n_o$ of the liquid crystal, and the elastic constants. The results of numerical analysis showed that this parameter is weakly dependent on the order parameters S, S_d, and S_f.

Figure 2 presents the results of calculations of the absolute phase shift $|\Delta\Phi|$ as a function of the control electric field strength E in comparison to the experimental data [20]. The measurements were performed by an interference technique using the laser radiation with a wavelength of $\lambda = 0.6328$ μm. The films were prepared using a nematic liquid crystal of the E7 type (with the refractive indices $n_o = 1.511$ and $n_e = 1.74$) dispersed in a polymer matrix representing a mixture of DPHPA, NVP, and BP monomers. The refractive index of the polymer matrix, the molecular order parameter, and order parameter of LC droplets (determined assuming a bipolar configuration of LC molecules in the particles) were $n_p = 1.524$, $S = 0.6$, and $S_d = 0.7$ [20]. In comparison to the experimental data, taking the known values of the incident radiation wavelength λ, refractive indices (n_o, n_e, n_p), order parameters (S, S_d) and the PDLC film thickness (l), the particle number density c_v and the parameter A were selected so as to provide for the best fit between theory and experiment (the corresponding c_v and A values are indicated in the legend to Fig. 2).

A more detailed comparison must include an analysis of the morphological and structural features of the PDLC films employed and the dependence of the order parameters S_d and S_f on the applied electric field.

4. COHERENT TRANSMISSION COEFFICIENT

The average amplitude transmission coefficient $\langle T_a \rangle$ of the film determined using expressions (18) and (19) is

$$\langle T_a \rangle = \frac{\langle E \rangle}{E_0 \exp(ikl)} = \exp(-qS(0)l). \quad (25)$$

For the coherent transmission coefficient T_c, this yields

$$T_c = |\langle T_a \rangle|^2 = \exp(-\gamma l), \quad (26)$$

$$\gamma = 2q\,\mathrm{Re}S(0) = \sigma_{\mathrm{ext}} N_v, \quad (27)$$

$$\sigma_{\mathrm{ext}} = \frac{4\pi}{k^2}\mathrm{Re}S(0), \quad (28)$$

where γ is the attenuation coefficient and σ_{ext} is the extinction cross section of a single LC droplet [22].

It should be noted that, within the framework of the Rayleigh approximation for nonabsorbing LC droplets (with real n_o and n_e values), the cross section σ_{ext} cannot be determined based on the optical theorem [25] (formula (28)), since the real part of the amplitude scattering function at a zero scattering angle $S(0)$ is zero (see formulas (9)–(12). In this case, cross section σ_{ext} has to be determined by integrating the amplitude scattering matrix elements over the entire solid angle $\Omega = 4\pi$ [22, 25].

Taking into account the symmetry of the problem under consideration, the average scattering matrix S can represented as

$$\underline{S} = \begin{pmatrix} S(0)\cos\delta & 0 \\ 0 & S(0) \end{pmatrix}, \quad (29)$$

where δ is the scattering angle. For the extinction cross section, this yields

$$\sigma_{\mathrm{ext}} = \frac{|S(0)|^2}{k^2}\int_{4\pi}(1+\cos^2\delta)\,d\Omega. \quad (30)$$

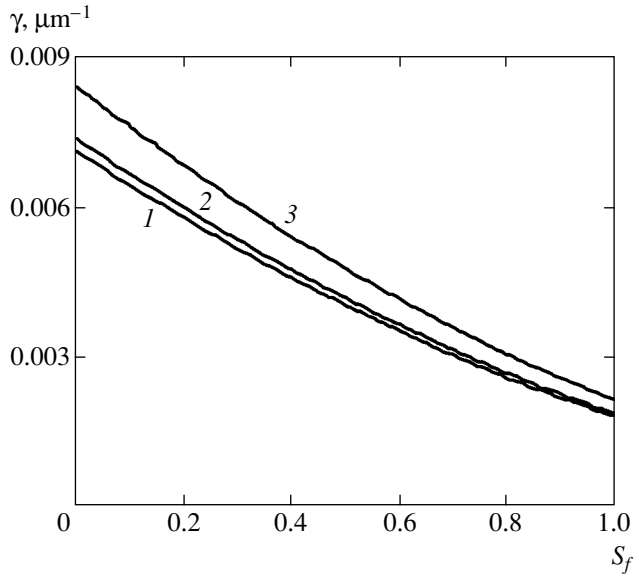

Fig. 3. Plots of the attenuation coefficient γ versus the film order parameter S_f, calculated for $\langle d\rangle$ = 150 nm, c_v = 0.089, S_d = 0.7, and various coefficients of variation of the LC droplet size distribution: $D_d/\langle d\rangle$ = 0.0625 (*1*), 0.125 (*2*), and 0.25 (*3*).

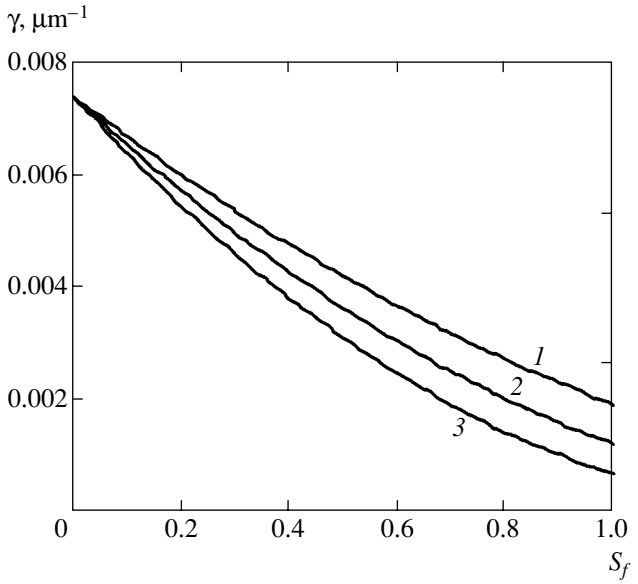

Fig. 4. Plots of the attenuation coefficient γ versus the film order parameter S_f, calculated for $D_d/\langle d\rangle$ = 0.125, $\langle d\rangle$ = 150 nm, c_v = 0.089, and various LC droplet order parameters: S_d = 0.7 (*1*), 0.85 (*2*), and 1.0 (*3*).

Accomplishing integration, we obtain

$$\sigma_{\text{ext}} = \frac{16\pi}{3}\frac{|S(0)|^2}{k^2}. \tag{31}$$

Using expressions (9)–(12) and (31), we obtain a formula for the attenuation coefficient γ, assuming the LC droplets to have the shape of spherical droplets,

$$\gamma = \frac{8}{9}\langle x\rangle^4 c_v f \langle d\rangle^{-1} \times\left(\frac{n_o^2}{n_p^2} - 1 + \frac{n_e^2 - n_o^2}{3n_p^2}(1 - SS_dS_f)\right)^2, \tag{32}$$

where $\langle x\rangle = \pi\langle d\rangle n_p/\lambda$ is the average parameter of diffraction, $\langle d\rangle$ is the average droplet diameter, λ is the incident radiation wavelength in vacuum, and $f = \langle d^3\rangle/(\langle d\rangle)^3$ is the ratio of the third moment of the droplet size distribution function to the cubic average diameter.

For a droplet size distribution described by the gamma function [26, 27],

$$P(d) = \frac{\mu^{\mu+1}}{\Gamma(\mu+1)}\frac{d^\mu}{(d_m)^{\mu+1}}\exp\left(-\frac{\mu d}{d_m}\right), \tag{33}$$

where d_m is the modal droplet size and μ is the distribution parameter, we obtain

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

The quantity μ is related to the variation coefficient $D_d/\langle d\rangle$ of the LC droplet size distribution, representing the ratio of the mean-square deviation D_d to the average droplet diameter $\langle d\rangle$, by the relation

$$\mu = 1/(D_d/\langle d\rangle)^2 - 1. \tag{35}$$

For monodisperse droplets, $\mu \longrightarrow \infty$ and $f \longrightarrow 1$.

Let us consider the behavior of the attenuation coefficient γ and the coherent transmission coefficient T_c for

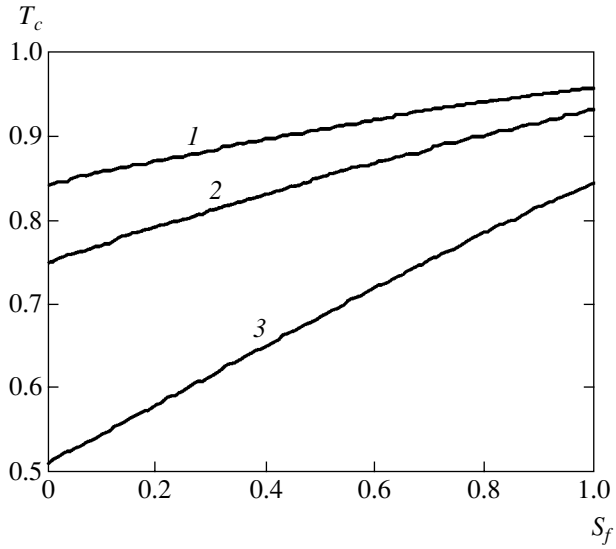

Fig. 5. Plots of the coherent transmission coefficient T_c versus the film order parameter S_f, calculated for $D_d/\langle d\rangle$ = 0.125, $\langle d\rangle$ = 150 nm, S_d = 0.7, l = 23 μm, and various number densities of LC droplets in the PDLC film: c_v = 0.089 (*1*), 0.15 (*2*), and 0.35 (*3*).

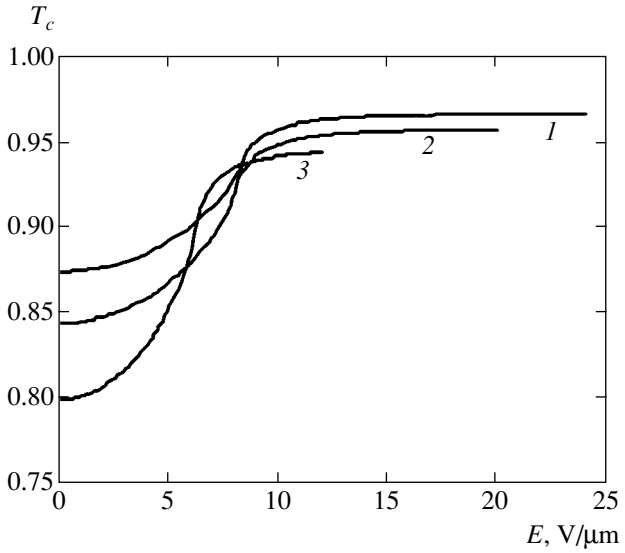

Fig. 6. Plots of the coherent transmission coefficient T_c versus applied electric field strength E, calculated for $D_d/\langle d\rangle$ = 0.125, $\langle d\rangle$ = 150 nm, S_d = 0.7 and various sets of parameters: (*1*) l = 13 μm, c_V = 0.125, A = 0.12 μm/V; (*2*) l = 23 μm, c_V = 0.089, A = 0.124 μm/V; (*3*) l = 36 μm, c_V = 0.075, A = 0.16 μm/V.

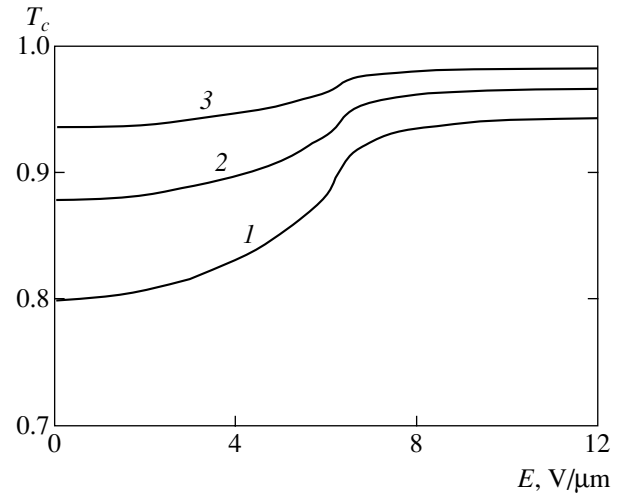

Fig. 7. Plots of the coherent transmission coefficient T_c versus applied electric field strength E, calculated for $D_d/\langle d\rangle$ = 0.125, S_d = 0.7, A = 0.16 μm/V, l = 36 μm, c_V = 0.075, and various average LC droplet dimensions: $\langle d\rangle$ = 150 (*1*), 125 (*2*), 100 nm (*3*).

a PDLC film. The influence of a polydisperse character and the LC droplet order parameter S_d on the value of γ is illustrated in Figs. 3 and 4. The curves depicted in these and other figures were calculated for n_o = 1.511, n_e = 1.74, n_p = 1.524, S = 0.6, and λ = 0.6328 μm; other values are specified in the figure legends. As can be seen in Fig. 3, growing polydispersity leads to an increase in the attenuation coefficient for all values of the film order parameter S_f ($0 \le S_f \le 1$). Increasing LC droplet order S_d leads to a decrease in γ (Fig. 4). In this case, an increase in the number density of droplets decreases the transmission coefficient T_c (Fig. 5).

Figure 6 shows variation of the coherent transmission coefficient T_c depending on the applied electric field strength for PDLC films with the same parameters (l, c_V, A) as those used in calculating the phase shift as a function of the field strength (Fig. 2). The vales of transmission in the absence of the control field (E = 0) correspond to the experimental values reported in [20].

Figure 7 illustrates the influence of the average LC droplet size on the PDLC film transmission. As can be seen, an decrease in the average LC droplet size leads to a growth in the transmission coefficient T_c, at a rater weak dependence of the transmission on the applied field strength.

5. CONCLUSION

Using the proposed model, describing propagation of a coherent radiation component through fine disperse PDLC films, it is possible to solve problems related to optimization of the morphology and structure of such films, aimed at (i) obtaining the required phase modulation effect, (ii) linearizing the field dependence of the phase shift, (iii) decreasing the level of control field strength, and (iv) increasing the transmission of PDLC films. The model is applicable to an analysis of the phase modulation and radiation attenuation in PDLC films with bipolar, axial and other cylindrically-symmetric structures of LC molecules in nanodimensional inclusions.

The results can be of importance for the development of telecommunication systems employing phase modulation in data processing and transmission.

ACKNOWLEDGMENTS

The authors are grateful to Prof. F. Simoni and other authors of [20] for kindly providing their experimental results prior to publication.

This study was supported by the Foundation for Basic Research of Belarus (project no. F01-042) and by the Fourth Framework Program of the European Commission INCO-Copernicus (contract No. ERB IC15- CT98-0806).

REFERENCES

1. P. S. Drzaic, *Liquid Crystal Dispersions* (World Sci., New York, 1995), p. 430.
2. F. Simoni, *Nonlinear Optical Properties of Liquid Crystals and Polymer Dispersed Liquid Crystals* (World Sci., Singapore, 1997).
3. *Liquid Crystals in Complex Geometries*, Ed. by G. P. Crawford and S. Zumer (Taylor and Francis, London, 1996), p. 279.

4. G. M. Zharkova and A. S. Sonin, *Mesomorphic Composites* (Nauka, Moscow, 1994), p. 213.
5. K. Takizawa, T. Fujii, H. Kikuchi, *et al.*, Appl. Opt. **38**, 5646 (1999).
6. V. Ya. Zyryanov, E. P. Pozhidaev, S. L. Smorgon, *et al.*, Liq. Cryst. **28**, 433 (2001).
7. V. P. Tondiglia, L. V. Natarajan, R. L. Sutherland, *et al.*, Opt. Lett. **20**, 1325 (1995).
8. J. Bajc and S. Zumer, Phys. Rev. E **55**, 2925 (1997).
9. J. Bajc, J. Bezic, and S. Zumer, Phys. Rev. E **51**, 2176 (1995).
10. F. Xu, H. S. Kitzerow, and P. P. Crooker, Phys. Rev. E **49**, 3061 (1994).
11. V. Vorflusev and S. Kumar, Science **283**, 1903 (1999).
12. V. A. Loiko and A. V. Konkolovich, J. Opt. B **3**, S155 (2001).
13. A. V. Konkolovich, V. V. Presnyakov, V. Ya. Zaryanov, *et al.*, Pis'ma Zh. Éksp. Teor. Fiz. **71**, 710 (2000) [JETP Lett. **71**, 486 (2000)].
14. A. P. Ivanov, V. A. Loiko, and V. P. Dik, *Light Propagation in Close-Packed Dispersed Media* (Nauka i Tekhnika, Minsk, 1988), p. 192.
15. S. Matsumoto, Y. Sugiyama, S. Sakata, *et al.*, Liq. Cryst. **27**, 649 (2000).
16. M. J. Sansone, G. Khanarian, and T. M. Leslie, J. Appl. Phys. **67**, 4253 (1990).
17. S. Matsumoto, K. Hirabayashi, S. Sakata, *et al.*, IEEE Photonics Technol. Lett. **11**, 442 (1999).
18. D. C. O'Brien, R. J. Mears, T. D. Wilkinson, *et al.*, Appl. Opt. **33**, 2795 (1994).
19. A. Isimaru, *Wave Propagation and Scattering in Random Media* (Academic, New York, 1978; Mir, Moscow, 1994), Vol. 2.
20. L. Lucchetta, R. Karopiran, A. Mann, and F. Simoni, J. Appl. Phys. **91**, 6060 (2002).
21. S. Zumer and J. W. Doane, Phys. Rev. A **34**, 3373 (1986).
22. C. F. Bohren and D. R. Huffman, *Absorption and Scattering of Light by Small Particles* (Wiley, New York, 1983; Mir, Moscow, 1986).
23. J. R. Kelly and P. Palffy-Muhoray, Mol. Cryst. Liq. Cryst. **243**, 11 (1994).
24. V. A. Loiko, A. V. Konkolovich, F. Simoni, *et al.*, in *Proceedings of the 10th SID Symposium on Advanced Display Technologies* (Minsk, 2001), p. 58.
25. H. C. van de Hulst, *Light Scattering by Small Particles* (Wiley, New York, 1957; Inostrannaya Literatura, Moscow, 1961).
26. O. A. Afonin, Yu. V. Panina, A. B. Pravdin, and D. A. Yakovlev, Liq. Cryst. **15**, 395 (1997).
27. V. A. Loiko and A. V. Konkolovich, J. Phys. D **33**, 2201 (2000).

Translated by P. Pozdeev

Journal of Experimental and Theoretical Physics, Vol. 96, No. 3, 2003, pp. 496–509.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 123, No. 3, 2003, pp. 560–574.

SOLIDS
Structure

Integrable Models for the Dynamics of a Longitudinal–Transverse Acoustic Wave in a Crystal with Paramagnetic Impurities

A. A. Zabolotskii

Institute of Automatics and Electrometry, Siberian Division, Russian Academy of Sciences, Universitetskii pr. 1, Novosibirsk, 630090 Russia
e-mail: zabolotskii@iae.nsk.su.

Received July 27, 2002

Abstract—We theoretically study the evolution of longitudinal–transverse acoustic pulses propagating parallel to an external magnetic field in a system of resonant paramagnetic impurities with an effective spin $S = 1/2$. For equal group velocities of the longitudinal and transverse waves, the pulse dynamics is shown to be described by evolution equations. In limiting cases, these equations reduce to equations integrable in terms of the inverse scattering transform method (ISTM). For the most general integrable system of equations that describes the dynamics of acoustic pulses outside the scope of the slow-envelope approximation, we derive the corresponding ISTM equations. These equations are used to find a soliton solution and a self-similar solution. The latter describes the leading edge of the packet of acoustic pulses generated when the initial unstable state of a spin system decays. Analysis of our solutions and models indicates that the presence of a longitudinal acoustic wave leads not only to a change in the amplitude and phase of the transverse wave but also to a qualitatively new dynamics of sound in such a medium. © *2003 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

At present, the nonlinear coherent optical phenomena that are associated with soliton and other self-similar solutions [1, 2] have been analytically studied in greatest detail in terms of integrable models [3]. When elastic waves propagate in paramagnetic crystals, soliton-like pulses can be produced by effects related to anharmonic oscillations and dispersion [4] and under conditions of nonlinear coherent interaction between acoustic waves and paramagnetic impurities in the medium, for acoustic self-induced transparency (ASIT) [5–7]. An analogy between optical and acoustic effects gave rise to magnetic quantum acoustics [8]. At the same time, the evolution of an acoustic pulse in a crystal with paramagnetic impurities has a number of qualitative distinctions from the dynamics of light waves in a medium, which are related, for example, to the fact that the acoustic wave in a crystal can be longitudinal–transverse. The characteristic length of the acoustic pulse is no less than 10^{-4} cm, i.e., much larger than the lattice constant. For a picosecond acoustic pulse of duration $\tau_a \sim 10$ ps, its length is 10^{-7}–10^{-6} cm. Therefore, when it propagates in a crystal, the spatial dispersion attributable to the lattice structure should be taken into account, but the discreteness of the crystal medium can be disregarded. Recently [4], the generation of picosecond acoustic solitons has been observed experimentally. The solitons formed at a distance of several millimeters due to the balance between the dispersion attributable to the positions of atoms in the crystal lattice and the nonlinearity that arises from the anharmonicity of interatomic forces. An acoustic resonant effect, an analog of optical self-induced transparency, was observed in low-temperature crystalline samples with paramagnetic impurities. This effect was observed on Fe^{2+} impurities in MgO [9] and $LiNbO_3$ [10] crystal lattices when a longitudinal acoustic pulse propagated at an angle to the external field.

In general, the group velocities of the longitudinal ($v_\parallel$) and transverse ($v_\perp$) components of an acoustic pulse in a solid are different. As a result, the local interaction between the pulses of these components of length l_a is limited by the time $t_{\text{int}} \sim l_a/(v_\parallel - v_\perp)$. The interaction of the pulses is most effective when the phase velocities are close, $v_\parallel \approx v_\perp$. This situation takes place in elastic-isotropic crystals, in which the velocities of the longitudinal and transverse elastic field components do not depend on the direction. These conditions are best satisfied for ion crystals of alkali metal galogenides with central forces of interatomic interaction [11]. For example, NaBr belongs to these crystals [11].

When propagating in a crystal, the longitudinal and transverse acoustic waves can affect the dynamics of each other through their interaction with paramagnetic impurities. The ASIT theory for a transverse pulse propagating parallel to the magnetic field in a system of $S = 1/2$ spins was developed by Denisenko [5]. This theory was generalized to longitudinal–transverse acoustic waves by Voronkov and Sazonov [6, 7], who derived equations describing the dynamics of acoustic pulses.

The authors used several approximations to solve these complex evolution equations. Voronkov and Sazonov [7] theoretically studied the ASIT of quasi-monochromatic pulses with a transverse–longitudinal structure propagating parallel to the vector of constant magnetic field in a system of paramagnetic impurities with an effective spin $S = 1/2$ under the condition $v_\perp = v_\parallel$ and assuming that the frequency of the transverse component of a quasi-monochromatic pulse was equal to the Zeeman splitting frequency of the Kramers doublet. The approximations of a slow envelope and an exact resonance were used for the transverse field, and the approximation of unidirectional acoustic wave propagation was used to simplify the equations for the longitudinal field. In [6], the authors considered acoustic pulses of extremely short duration similar to ultrashort optical pulses [2]. In addition, the authors of [6, 7] used the Wentzel–Brillouin–Kramers–Jeffery method, which requires a slow field variation, to simplify the Bloch equations. However, even when all of the corresponding conditions imposed on the field dynamics were satisfied, the equations derived by these authors were not integrable. The condition for the absence of a longitudinal wave, under which the system of simplified equations reduces to the well-studied sinus-Gordon equation, constitutes an exception. Therefore, the approximations used in [6, 7] do not allow the dynamics and mutual influence of the transverse and longitudinal acoustic waves to be studied in detail.

At the same time, the rich structure of these equations opens up the possibility of reducing them, for quite realistic approximations, to integrable models both when similar stringent conditions are imposed and without them. Here, we show that, to pass from the original evolution equations to an integrable system of equations, it will suffice to use the approximation of unidirectional wave propagation and the condition of equal group velocities for the longitudinal and transverse acoustic waves. The resulting new integrable system is related to the integrable system of equations that we derived previously in the theory of four-wave mixing of optical pulses in media with two-photon-induced Kerr nonlinearity [12, 13]. At the same time, the integrable system of equations derived here has the symmetry of the spectral problem, for which the apparatus of the inverse scattering transform method (ISTM) [3] has not been developed. The ISTM application to this model allows the various regimes of evolution of picosecond acoustic pulses to be studied outside the scope of the slow-envelope approximation.

Here, we develop the ISTM apparatus for our system of equations and find a soliton solution of the model that explicitly describes, in particular, the deformation and nonlinear rotation of the polarization of an acoustic pulse of a transverse wave in the presence of a pulse of a longitudinal wave. Apart from the soliton solutions associated with ASIT, other regimes of pulse evolution can also be analyzed in terms of the ISTM. Soliton solutions often require more stringent conditions for the generation and observation than do solutions of a different kind. For example, when the spins of impurity ions are initially directed along the magnetic field, an unstable state of the system arises and a weak seed acoustic wave is sufficient to remove it from this state. Here, for such a spin system, we use the ISTM to find the shape of the pulse that corresponds to the leading edge of the packet of generated pulses when the system approaches a stable state.

Apart from intense picosecond acoustic pulses, weak acoustic pulses that require lower powers for their generation can be formed. It is of interest to study the physical conditions for the generation of such low-amplitude pulses and their parameters. From this point of view, it is important to determine the type of interaction, the minimum nonlinearities, and other conditions under which stable nonlinear structures such as solitons and other similar solutions can be formed. It is also important to seek integrable reductions, because the most detailed information can be obtained by solving models integrable in terms of the ISTM and by developing perturbation theory for similar models. Therefore, we also derive other integrable versions of the original system of equations obtained when passing to low amplitudes and/or to a negligible change in the component of the effective spin along the magnetic field. We show that, when passing to quasi-monochromatic acoustic pulses, the original system of equations also reduces to the integrable system of equations that previously arose in connection with various applications in nonlinear optics [14]. Next, we show that, under the additional condition of a negligible change in Zeeman level populations, the derived system reduces to an integrable equation related to the Thirring model, whose application in nonlinear optics was described, for example, in [15]. Analysis of the derived equations for the dynamics of acoustic waves and comparison with the solutions of related equations obtained in nonlinear optics reveal several new physical observational properties of the effects associated with the presence of a longitudinal acoustic wave without a detailed solution of the evolution equations.

This paper has the following structure. The basic system of evolution equations that describes the dynamics of a longitudinal–transverse wave is derived in Section 2. The most general integrable reduction of the original system of equations for this system was found in Section 3.1. In Section 3.2, we give the corresponding Lax representations and develop the ISTM apparatus for this system. A one-soliton solution for this model was found in Section 3.3. In Section 3.4, we analyze the decay of an initially stable state and give an asymptotic solution for the first pulse from the packet of nonlinear oscillations that correspond to this solution. Section 4 is devoted to the derivation and discussion of other integrable model reductions. In Section 5 we discuss our results.

2. THE DERIVATION OF BASIC EQUATIONS

Below, we derive the equations that describe the dynamics of a longitudinal–transverse wave in a crystal with paramagnetic impurities following [6]. Assume that an external constant and uniform magnetic field **B** is directed along the *z* axis. The Zeeman interaction of the magnetic moment $\hat{\mu}^{(a)}$ at point *a* contributes $\hat{H}_a = -\hat{\mu}^{(a)}\mathbf{B}$ to the total Hamiltonian. The $\hat{\mu}^{(a)}$ components can be expressed in terms of the $\mathbf{S}^{(a)}(\mathbf{r}_a)$ spin components, where $\mathbf{r}_a$ is the radius vector of spin *a*, as

$$\hat{\mu}_j^{(a)} = -\sum_k \mu_B g_{jk} \hat{S}_k^{(a)}.$$

Here, μ_B is the Bohr magneton and g_{jk} are the Lande tensor components. Thus,

$$\hat{H}^z = \sum_{a=1}^{N} \hat{H}_a^z = \mu_B \sum_a \sum_{j,k} B_j g_{jk} \hat{S}_k^{(a)}, \quad (1)$$

where *N* is the total number of spins. Since the effective spin is 1/2, it can be decomposed into Pauli matrices:

$$\hat{S}_x^a = \frac{1}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \hat{S}_y^a = \frac{1}{2}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \hat{S}_z^a = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \quad (2)$$

Assume that the *x*, *y*, and *z* coordinates along the principal Lande tensor axes coincide with the crystal symmetry axes. The Lande tensor is then diagonal in a nondeformed unperturbed medium: $g_{jk} = g_{jk}^{(0)} = g_{jj}\delta_{jk}$, where δ_{jk} is the delta function. The deformation of the crystal by an acoustic wave is described by linear corrections to the Lande tensor:

$$g_{jk} = g_{jk}^{(0)} + \sum_{p,q} \left(\frac{\partial g_{jk}}{\partial \epsilon_{pq}}\right)_0 \epsilon_{pq} + \ldots, \quad (3)$$

where ϵ is the crystal elastic strain tensor at the spin location. The derivatives are taken at the point of zero deformation. The strain tensor components can be expressed in terms of the components of displacement vector $\mathbf{u} = (u_x, u_y, u_z)$ as

$$\epsilon_{pq} = \frac{1}{2}\left(\frac{\partial u_p}{\partial x_q} + \frac{\partial u_q}{\partial x_p}\right). \quad (4)$$

Hamiltonian (1) takes the form

$$\hat{H}_s = \sum_\alpha \sum_j \mu_B g_{jj} B_j \hat{S}_j^{(a)}. \quad (5)$$

The spin–phonon interaction is described by the Hamiltonian

$$\hat{H}_{\text{int}} = \sum_\alpha \sum_{j,k,p,q} \mu_B B_j F_{jkpq} \epsilon_{pq} \hat{S}_k^{(a)}. \quad (6)$$

Here, $F_{jkpq} = \partial q_{jk}/\partial \epsilon_{pq}$ are the spin–phonon coupling constants [10, 16].

The dynamics of an acoustic field in a crystal without anharmonicity is described by the Hamiltonian

$$H_a = \int \left(\frac{1}{2n_0}\sum_j p_j^2 + \frac{1}{2}\sum_{j,k,l,m} \lambda_{jklm} \frac{\partial u_j}{\partial x_k}\frac{\partial u_l}{\partial x_m}\right) d\mathbf{r}, \quad (7)$$

where n_0 is the mean crystal density, p_j ($j = x, y, z$) are the momentum density components that arise during dynamic displacements, and l_{jklm} is the elastic modulus tensor of the crystal [17]. The integral in (7) is taken over the crystal volume. We assume that the number of phonons is large and that the classical description of the acoustic field dynamics is valid. At the same time, a two-level spin system requires the quantum-mechanical description. For $S = 1/2$, the terms quadratic in spin operators can be disregarded (for more detail, see [18]). Here, an analogy with the interaction of a classical electromagnetic field with an optical quantum two-level medium holds [19].

As in the case of an optical medium, we can pass from the description of the spin dynamics to the evolution equations for the density matrix elements $\hat{\rho}^{(a)}$,

$$i\hbar \frac{\partial \hat{\rho}^{(a)}}{\partial t} = [\hat{H}, \hat{\rho}^{(a)}], \quad (8)$$

$$\frac{\partial \mathbf{u}}{\partial t} = \frac{\partial H}{\partial \mathbf{p}}, \quad \frac{\partial \mathbf{p}}{\partial t} = -\frac{\partial H}{\partial \mathbf{u}}. \quad (9)$$

Here, $H = H_a + \langle \hat{H}_{\text{int}} \rangle$, where the interaction between the spin and the field of an elastic pulse is described by the classical Hamiltonian equations for a continuous medium:

$$\langle \hat{H}_{\text{int}} \rangle = \sum_\alpha \sum_{j,k,p,pq} \mu_B B_j F_{jkpq} \int \epsilon_{pq}(\mathbf{r}) \langle \hat{S}_k^{(a)}(\mathbf{r}) \rangle d\mathbf{r}. \quad (10)$$

Here, *n* is the paramagnetic impurity density, the angular brackets denote an averaging over quantum states, and the summation over the ions with spins 1/2 uniformly distributed in the crystal is substituted with integration over the entire space.

We consider the field evolution only along the *z* axis parallel to the vector **B** and the fourth-order symmetry axis. In this case, the symmetry transformation includes the rotation through 90° around the *z* axis and the

reflections $x \longrightarrow -x$ and $y \longrightarrow -y$ in the $z = 0$ plane. Given these conditions, the expressions for $\hat{H}_s$ and $\hat{H}_{\text{int}}$ take the form [6]

$$\hat{H}_s = \sum_{\alpha} \hbar\omega_B \hat{S}_z^{(\alpha)}, \tag{11}$$

$$\hat{H}_{\text{int}} = \sum_{\alpha} \frac{\hbar\omega_B}{g} \times \{F_{11}\epsilon_{zz}\hat{S}_z^{(\alpha)} + F_{44}(\epsilon_{xz}\hat{S}_x^{(\alpha)} + \epsilon_{yz}\hat{S}_y^{(\alpha)})\}, \tag{12}$$

where $\omega_B = g\mu_B B/\hbar$ is the Zeeman splitting frequency of the Kramers doublet and $g = g_{xx} = g_{yy} = g_{zz}$.

Under these symmetry conditions, Hamiltonian H_a is

$$\hat{H}_a = \frac{1}{2}\int\left\{\frac{p_x^2 + p_y^2 + p_z^2}{n_0} + \lambda_{11}\left(\frac{\partial u_z}{\partial z}\right)^2 + \lambda_{44}\left[\left(\frac{\partial u_x}{\partial z}\right)^2 + \left(\frac{\partial u_y}{\partial z}\right)^2\right]\right\} d\mathbf{r}. \tag{13}$$

Here, the Vogt notation is used for the subscripts: $xx \longrightarrow 1$, $yy \longrightarrow 2$, $zz \longrightarrow 3$, $yz \longrightarrow 4$, $xz \longrightarrow 5$, $xy \longrightarrow 6$.

Using (8)–(10) and (11)–(13), we obtain the basic system of evolution equations (cf. [6])

$$\frac{\partial^2\epsilon_\perp}{\partial t^2} - v_\perp^2\frac{\partial^2\epsilon_\perp}{\partial z^2} = \frac{nG_\perp}{n_0}\frac{\partial^2 S_\perp}{\partial z^2}, \tag{14}$$

$$\frac{\partial^2\epsilon_\parallel}{\partial t^2} - v_\parallel^2\frac{\partial^2\epsilon_\perp}{\partial z^2} = -\frac{nG_\parallel}{n_0}\frac{\partial^2 S_3}{\partial z^2}, \tag{15}$$

$$\frac{\partial S_\perp}{\partial t} = i\left(\frac{G_\parallel\epsilon_\parallel}{\hbar} + \omega_B\right)S_\perp + i\frac{G_\perp\epsilon_\perp}{\hbar}S_3, \tag{16}$$

$$\frac{\partial S_3}{\partial t} = \frac{iG_\perp}{2\hbar}(\epsilon_\perp^* S_\perp - \epsilon_\perp S_\perp^*), \tag{17}$$

where $G_\parallel = \hbar\omega_B F_{11}/g$, $G_\perp = \hbar\omega_B F_{44}/g$, $\epsilon_\parallel = \epsilon_{zz}$, $\epsilon_\perp = \epsilon_{xz} + i\epsilon_{yz}$, $v_\parallel = \sqrt{\lambda_{11}/n_0}$, $v_\perp = \sqrt{\lambda_{44}/n_0}$, $S_3 = (\rho_{11} - \rho_{22})/2$, and $S_\perp = \rho_{21}$. The derived system of equations (14)–(17) describes the propagation of acoustic pulses in a two-level medium, with the longitudinal component in the Bloch equations (16) and (17) leading only to a nonlinear phase modulation. However, as we show below, the longitudinal field can lead to a qualitatively new dynamics of an acoustic pulse. Below, the group velocities of the longitudinal and transverse waves are assumed to be equal: $v_\perp = v_\parallel = v$.

3. USING THE ISTM TO SOLVE THE EVOLUTION MODEL FOR ACOUSTIC PULSES WITH DURATION OF THE ORDER OF ω_B^{-1}

3.1. The Derivation of an Integrable Model

Let us derive the most general integrable reduction of the basic system of equations (14)–(17). This reduction describes the dynamics of acoustic pulses on the order of or shorter than $\pi\omega_B^{-1}$ in duration. Under this condition, the slow-envelope approximation is inapplicable. In real media, picosecond acoustic pulses can correspond to this case [6]. The equations that describe the dynamics of such pulses are difficult to analyze, but system (14)–(17) can be simplified at low densities of paramagnetic impurities. Such a physical situation takes place in the overwhelming majority of cases. This approximation is similar to the approximation used in [20] to derive the reduced Maxwell–Bloch equations for a two-level optical medium and is called the condition of unidirectional wave propagation. The following approximate formal equality holds for these conditions:

$$\partial_z \approx -v^{-1}\partial_t + O(\varepsilon),$$

where ε is a small parameter. Physically, this means that the acoustic pulses propagate in a medium at a velocity close to the group velocity v. The normalized impurity density is of the same order of smallness as the derivative $\partial_{\tilde{\chi}} = \partial_z + v^{-1}\partial_t$ of the acoustic field amplitudes. We can now substitute $v^{-1}\partial_t$ for the derivative with respect to z on the right-hand sides of Eqs. (14) and (15) with an accuracy of $O(\varepsilon^2)$. Thus, when the condition of unidirectional acoustic pulse propagation is satisfied, system (14) and (15) reduces to

$$\frac{\partial\epsilon_\perp}{\partial\tilde{\chi}} = \frac{nG_\perp}{2v^2 n_0}\frac{\partial S_\perp}{\partial t}, \tag{18}$$

$$\frac{\partial\epsilon_\parallel}{\partial\tilde{\chi}} = -\frac{nG_\parallel}{\partial v^2 n_0}\frac{\partial S_3}{\partial t}. \tag{19}$$

It is now easy to find from Eqs. (16)–(19) that the amplitudes of the longitudinal and transverse fields are related by

$$|\epsilon_\perp^2| + \left(\epsilon_\parallel + \frac{2\omega_B\hbar}{G_\parallel}\right)^2 = U_0^2(t). \tag{20}$$

Here, the real function $U_0(t)$ is determined by boundary conditions. Below, we assume that $U_0(t) = \text{const}$.

Given (20), the new integrable system (16)–(19) takes the form

$$\begin{aligned}\partial_\chi E &= i\beta_0\sqrt{1-|E|^2}S_\perp + iES_3,\\ \partial_\tau S_\perp &= i\beta_0\sqrt{1-|E|^2}S_\perp + iES_3,\\ \partial_\tau S_3 &= -\frac{i}{2}(ES_\perp^* - E^*S_\perp),\end{aligned} \tag{21}$$

where

$$E = \frac{\epsilon_\perp}{U_0(t)}, \quad \chi = \tilde{\chi}\frac{nG_\perp^2}{2\hbar n_0 v^2},$$

$$\tau = t\frac{U_0 G_\perp}{\hbar}, \quad \beta_0 = \frac{G_\parallel}{G_\perp}.$$

3.2. The ISTM Apparatus for System (21)

We will solve the problem on the entire axis for $E(\tau) \longrightarrow 0$ and $\tau \longrightarrow \pm\infty$ by assuming that the spin system is in a stable ground state corresponding to minimum energy, i.e., $S_3(\tau, \chi) = 1$, $S_\perp(\tau, \chi) = 0$, $\tau \longrightarrow \pm\infty$, at the initial and final times. A pulse of acoustic field $E(\tau, 0)$ with an area large enough for the formation of solitons is assumed to be injected into the crystal. To describe the corresponding soliton dynamics, it is convenient to choose the following Lax representation for the system of equations (21):

$$\partial_\tau \Phi = \begin{pmatrix} -i\lambda F_3 & (\lambda+\beta)E \\ -\lambda E^* & i\lambda F_3 \end{pmatrix}\Phi \equiv \hat{L}_1\Phi, \tag{22}$$

$$\begin{aligned}\partial_\chi \Phi &= \frac{1}{2\lambda+\beta_0}\\ &\times\begin{pmatrix} i\lambda S_3 & \beta_0(\lambda+\beta)S_\perp \\ -\beta_0\lambda S_\perp^* & -i\lambda S_3 \end{pmatrix}\Phi \equiv \hat{A}_1\Phi,\end{aligned} \tag{23}$$

where λ is the spectral parameter,

$$F_3 = \sqrt{1-|E|^2}, \quad \beta = \frac{1}{2}\left(\beta_0 - \frac{1}{\beta_0}\right), \quad \beta_0 \neq 0.$$

Here, the case $\beta_0 = 0$ is not considered, although Eqs. (21) also admit the Lax representation and the ISTM application.

The spectral problem (22) differs from the studied related problems associated with the solution of the integrable Heisenberg and Landau–Lifshitz equations [21] or the equations of Raman scattering, four-wave mixing [13], by symmetry properties. Therefore, the ISTM apparatus must be developed for this model by taking into account its specifics.

The solutions of the spectral problem (22) have the involution

$$\Phi = \hat{M}\Phi(\lambda^*)^*\hat{M}^{-1}, \tag{24}$$

where

$$\hat{M} = \begin{pmatrix} 0 & (\lambda+\beta)/\lambda \\ -1 & 0 \end{pmatrix}. \tag{25}$$

In a standard way, we introduce the Jost functions $\Psi^\pm$, the solutions of Eq. (22) [for a potential $E(\tau)$ that rapidly vanishes when $\tau \longrightarrow \pm\infty$] with the asymptotics

$$\Psi^\pm = \exp(-i\lambda\sigma_3\tau), \quad \tau \longrightarrow \pm\infty. \tag{26}$$

The symmetry properties (24) and (25) correspond to the following matrix form of the Jost functions:

$$\Psi^\pm = \begin{pmatrix} \psi_1^\pm & -\psi_2^{\pm *}(\lambda+\beta)/\lambda \\ \psi_2^\pm & \psi_1^{\pm *} \end{pmatrix}.$$

These solutions are related by the scattering matrix $\hat{T}$,

$$\Psi^- = \Psi^+\hat{T}. \tag{27}$$

It follows from the symmetry property (24) and (25) that the scattering matrix can be chosen in the form

$$\hat{T} = \begin{pmatrix} a^* & b(\lambda+\beta)/\lambda \\ -b^* & a \end{pmatrix}. \tag{28}$$

The Jost functions have standard analytical properties (cf., e.g., [22]). The function $a(\lambda)$ is holomorphic in the upper half-plane λ, where its zeros correspond to the soliton solutions.

Let us represent the Jost functions as

$$\begin{aligned}\Psi^+(\tau) &= \exp(-i\lambda\sigma_3\tau)\\ &+ \int_\tau^\infty \begin{pmatrix} \lambda K_1(\tau, s) & (\lambda+\beta)K_2(\tau, s) \\ -\lambda K_2^*(\tau, s) & \lambda K_1^*(\tau, s) \end{pmatrix}\exp(-i\lambda\sigma_3 s)ds.\end{aligned} \tag{29}$$

It follows from system (27) that

$$\psi_1^{+*} = \frac{\psi_1^{-*}}{a} - \frac{b}{a}\frac{\lambda+\beta}{\lambda}\psi_2^+, \tag{30}$$

$$\psi_2^{+*} = -\frac{\psi_2^{-*}}{a} + \frac{b}{a}\psi_1^+. \tag{31}$$

We substitute the components of these functions from (29) into (30) and (31) and integrate the resulting equations over λ from $-\infty$ to ∞ with the weight

$\exp(-i\lambda y)(2\pi\lambda)^{-1}$. As a result, we obtain the Marchenko equations for the right end of the axis,

$$K_2^*(\tau, y) = F_0(t+y) + i\int_\tau^\infty K_1(\tau, s)\partial_y F_0(s+y)ds, \quad y \geq \tau, \tag{32}$$

$$K_1^*(\tau, y) = -\int_\tau^\infty K_2(\tau, s)(\beta + i\partial_y)F_0(s+y)ds, \quad y \geq \tau. \tag{33}$$

Here, we denoted

$$F_0(y) = \int_C \frac{b(\chi)}{a(\chi)}\frac{1}{2\pi\lambda}e^{-i\lambda y}d\lambda, \tag{34}$$

where C is the contour that includes the real axis and that passes above all poles in the upper half of the complex plane. Given the residues at poles λ_k in the upper half-plane, the kernel F_0 can be represented as

$$F_0(y) = \int_{-\infty}^\infty \frac{b(\chi)}{a(\chi)}\frac{1}{2\pi\lambda}e^{-i\lambda y}d\lambda - i\sum_k \frac{b(\chi)}{\partial_\lambda a(\chi;\ \lambda = \lambda_k)}\frac{1}{\lambda_k}e^{-i\lambda_k y}. \tag{35}$$

The Marchenko equations for the left end of the axis (for $y \leq \tau$) can be found in a similar way. Using the results obtained below, we can then easily show that the corresponding solutions are joined at $y = \tau$.

Substituting expression (29) for Ψ^+ into the spectral problem (22) and equating the expressions for different powers of λ yields

$$\begin{aligned} i\partial_\tau K_1(\tau, y) &= E\partial_y K_2(\tau, y) + i\beta E K_2(\tau, y) \\ &\quad - iF_3\partial_y K_1(\tau, y), \quad y = \tau, \\ i\partial_\tau\partial_y K_1(\tau, y) &= E\partial_y^2 K_2(\tau, y) + i\beta E\partial_y K_2(\tau, y) \\ &\quad - iF_3\partial_y^2 K_1(\tau, y), \\ \partial_\tau K_2(\tau, y) &= iE^*\partial_y K_1(\tau, y) + F_3\partial_y K_2(\tau, y), \\ &\quad y = \tau, \\ i\partial_\tau\partial_y K_2(\tau, y) &= iF_3\partial_y^2 K_2(\tau, y) - E^*\partial_y^2 K_1(\tau, y). \end{aligned} \tag{36}$$

The following relation is also valid:

$$K_2(\tau, \tau)[1 + F_3(\tau)] = E^*(\tau)[1 - iK_1(\tau, \tau)]. \tag{37}$$

Using this relation and the condition $F_3 + EE^* = 1$, we can easily find a relationship between the potentials E, F_3 and the kernels $K_{1,2}$ in the form

$$F_3(\tau) = \frac{[1 + iK_1^*(\tau, \tau)][1 - iK_1(\tau, \tau)] - |K_2(\tau, \tau)|^2}{[1 + iK_1^*(\tau, \tau)][1 - iK_1(\tau, \tau)] + |K_2(\tau, \tau)|^2}, \tag{38}$$

$$E(\tau) = \frac{2[1 - iK_1(\tau, \tau)]K_2^*(\tau, \tau)}{[1 + iK_1^*(\tau, \tau)][1 - iK_1(\tau, \tau)] + |K_2(\tau, \tau)|^2}. \tag{39}$$

3.3. A One-Soliton Solution for the Model

The ASIT-associated soliton solutions describe the propagation of acoustic pulses with no change in their shape against the background of a stable ground state. This state for system (22) is

$$E(\tau, \chi) = 0, \quad S_3(\tau, \chi) = 1, \quad S_\perp(\tau, \chi) = 0. \tag{40}$$

Let us find the one-soliton solution of the problem associated with the only eigenvalue λ. Substitute the kernel F that corresponds to this λ value in the form

$$F_0(y) = C_0(\chi)e^{-i\lambda y}. \tag{41}$$

The function $C_0(\chi)$,

$$C_0 = -\frac{i}{\lambda}\frac{b(\chi;\ \lambda)}{\partial_\eta a(\chi;\ \eta)}\bigg|_{\eta = \lambda},$$

is derived below.

To solve the Marchenko equations, we introduce new functions:

$$Q_1(\tau) = \int_\tau^\infty K_1(\tau, s)e^{-i\lambda s}ds,$$

$$Q_2(\tau) = \int_\tau^\infty K_2(\tau, s)e^{-i\lambda s}ds.$$

Substituting these functions into Marchenko equations (33) and (34) and integrating them over y yields

$$K_1^*(\tau, y) = -\gamma_0^*(\chi)\frac{i(\lambda - \lambda^*)}{\lambda^*} \times \frac{\omega(\tau)\exp(i\lambda^*\tau - i\lambda y)}{1 + \gamma_0^*(\chi)\omega^2(\tau)}, \tag{42}$$

$$K_2^*(\tau, y) = \frac{C_0\exp[-i\lambda(\tau + y)]}{1 + \gamma_0(\chi)\omega^2(\tau)}, \tag{43}$$

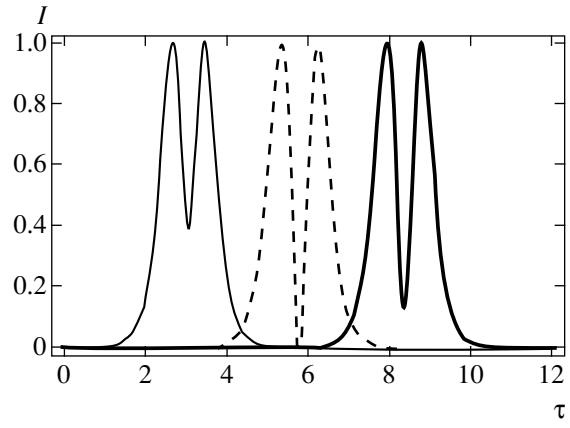

Fig. 1. Dependence (47) for the soliton intensity $I(\tau) = |E|^2$ for $\eta = 1$ and various β_0: $\beta_0 = 0.5$ (thin solid line), $\beta_0 = 1$ (dashed line), and $\beta_0 = 1.5$ (heavy line). The soliton position on the τ axis is arbitrary.

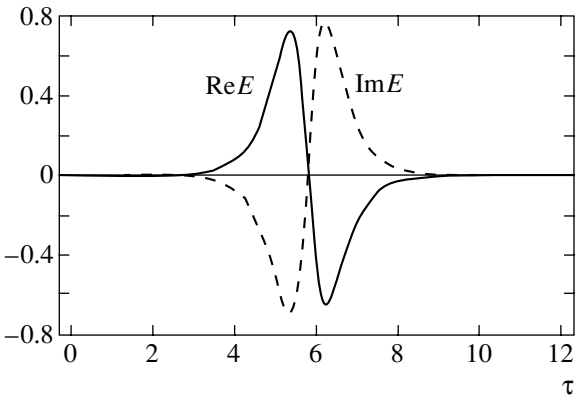

Fig. 2. Field components versus τ for $\beta_0 = 1.1$

where we denoted

$$\gamma_0(\chi) = -\frac{|C_0(\chi)|^2(\lambda^* + \beta)\lambda}{(\lambda - \lambda^*)^2},$$

$$\omega(\tau) = \exp[-i\tau(\lambda - \lambda^*)].$$

Next, it is necessary to find the function $C_0(\chi)$, which can be determined from system (23) for $S_3(-\infty, \chi) = S_3(\infty, \chi) \equiv 1$. This function can be found by using the formula (see [22] for its derivation)

$$\partial_\chi \hat{T} = -\hat{T}\exp(-i\sigma_3\lambda\tau)\hat{A}_s(\tau = -\infty, \chi)\exp(i\sigma_3\lambda\tau) + \exp(-i\sigma_3\lambda\tau)\hat{A}_s(\tau = \infty, \chi)\exp(i\sigma_3\lambda\tau)\hat{T}. \quad (44)$$

Hence, for the chosen initial and boundary conditions that correspond to soliton dynamics, we obtain

$$\frac{b(\lambda)}{\partial_\eta a(\eta)|_{\eta = \lambda}}(\chi) = -S_0\exp\left(\frac{2i\chi\lambda}{2\lambda + \beta_0}\right), \quad (45)$$

where S_0 is a constant. It follows from the derived formulas that the function $C_0(\chi)\exp(-i\lambda\tau) = \exp[\phi(\tau, \chi)]$ for $\lambda = i\eta + \xi$ is

$$\phi_s = i\phi_0 + \phi_1 - 2i\phi_2 + 2\phi_3, \quad (46)$$

where

$$\phi_0 = \arg\left(\frac{iS_0}{\lambda}\right), \quad \phi_1 = \ln\left|\frac{S_0}{\lambda}\right|,$$

$$\phi_2 = \xi\tau - \chi\frac{\beta_0\xi + 2(\xi^2 - \eta^2)}{(2\xi + \beta_0)^2 + 4\eta^2},$$

$$\phi_3 = \eta\tau - \chi\frac{2\beta_0\eta}{(2\xi + \beta_0)^2 + 4\eta^2}.$$

Next, we set $\xi = 0$ for simplicity. Using (39), (42), (43), and (45), we obtain a one-soliton solution of the model for $\lambda = i\eta$ in the form

$$E(\tau, \chi) = 2S_0\eta^{-1}\mu(\tau, \chi)[1 - \gamma_1|\mu(\tau, \chi)|^2] \times \{1 + |\mu(\tau, \chi)|^2[|S_0|^2\eta^{-2} - \gamma_1 - \gamma_1^*] + |\mu(\tau, \chi)|^4|\gamma_1|^2\}^{-1}\exp(i\phi_s), \quad (47)$$

where

$$\mu(\tau, \chi) = \exp\left(2\tau\eta - \frac{4\beta_0\eta\chi}{\beta_0^2 + 4\eta^2} - \frac{4i\eta^2\chi}{\beta_0^2 + 4\eta^2}\right),$$

$$\phi_s = \arg\left[\frac{1 + \gamma_1^*|\mu(\tau, \chi)|^2}{1 + \gamma_1|\mu(\tau, \chi)|^2}\right], \quad \gamma_1 = \frac{|S_0|^2(\eta^2 + i\eta\beta)}{4\eta^4}.$$

We see from expression (46) for the phase and from solution (47) that the soliton shape and velocity depend on the coefficient β_0, i.e., on the relative contribution of the longitudinal field. The soliton velocity decreases with increasing β_0, starting from the group velocity v in the medium. At $\beta_0 \sim |\lambda|$, the velocity reaches a minimum and then again tends to the phase velocity. In Fig. 1, the soliton intensity is plotted against τ for various values of β_0. We see from the figure that there is a dip to zero near $\beta_0 = 1$ at the line center. This dip disappears at $\beta_0 \ll 1$ or $\beta_0 \gg 1$. Figure 2 shows the imaginary and real parts of E, i.e., the y and x components of the transverse field, respectively. As we see from this figure, they are out of phase. At $\beta_0 = 1$, their positions relative to the τ coordinate axis are symmetric. It follows from formula (38) that the normalized soliton amplitude does not exceed unity. Therein lies an important difference between the acoustic field dynamics in the

model under consideration and the dynamics studied in [5]. The soliton solution of the model corresponds to a relationship between the amplitudes of the longitudinal and transverse acoustic waves in the form of (20). This condition implies that at large amplitudes such that $\epsilon_\parallel \gg 2\omega_B\hbar/G_\parallel$, the transverse field amplitude decreases with longitudinal field amplitude. Note that the coefficient $G_\perp$ does not enter equality (20); i.e., for the model under consideration in the approximation used, we cannot ignore the longitudinal field amplitude compared to the transverse field amplitude and vice versa. Otherwise, relation (20) yields trivial solutions: $\epsilon_\perp(z, t) \equiv \epsilon_\perp(0, t)$ or $\epsilon_\parallel(z, t) \equiv \epsilon_\parallel(0, t)$.

It also follows from solution (47) that the asymmetry due to β_0 differing from unity gives rise to a nonlinear addition ϕ_s to the soliton phase. This addition describes the nonlinear rotation of transverse field polarization in the $z = 0$ plane and disappears when $G_\perp = G_\parallel$.

3.4. The Decay of an Unstable Initial State of the Spin System

Strictly speaking, one-soliton model solutions describe the pulse evolution in a medium on the condition that a pulse with a certain area was injected into the medium at point $\chi = 0$. In the case of a deviation from this condition, apart from soliton solutions, other types of solutions, for example, the radiative solutions corresponding to the continuum spectrum of the problem (22) should also be taken into account. A nonsoliton field dynamics associated with the generation of weak and strong acoustic pulses is possible for some initial and boundary conditions. Thus, for example, for a weak seed pulse of acoustic field $|E| \ll 1$ and $S_3(0, \chi) = 1$, no soliton solutions emerge and the field dynamics is associated only with the continuum spectrum of the problem (22). For low amplitudes of the acoustic field and its derivatives, we can easily determine the scattering coefficient $\rho(\chi; \lambda) = b/a$ at $\chi = 0$:

$$\begin{aligned}\rho(0; \lambda) &= \rho_0 = \frac{b(0; \lambda)}{a(0; \lambda)} \\ &\approx -\frac{\lambda}{2}\int_{-\infty}^{\infty} E^*(\tau, 0)e^{2i\lambda\tau}d\tau.\end{aligned} \tag{48}$$

Let $E(\tau, 0) = \text{const}$. It is then easy to show that the scattering coefficient $\rho = b/a = \rho_0$ does not depend on λ.

For $|\rho_0| \ll 1$, Marchenko equations (31) and (33) have the following approximate solution:

$$\begin{aligned}K_2^*(\tau, y) &\approx F_0(\tau + y) + O(|\rho_0|^3), \\ K_1(\tau, y) &\sim O(|\rho_0|^2).\end{aligned} \tag{49}$$

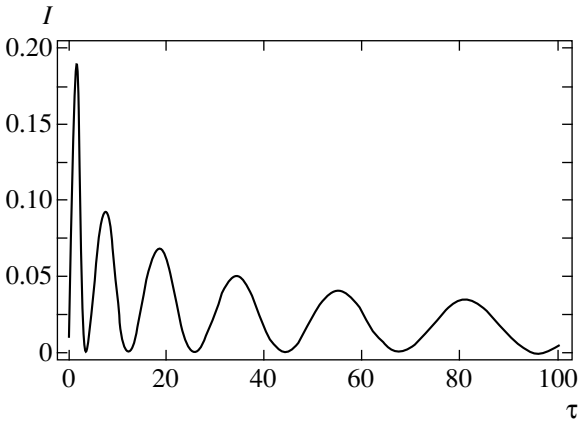

Fig. 3. Intensity $I = |E|^2$ versus τ for a small initial deviation from a stable ground state

The kernel F_0 roughly calculated for $\sqrt{\beta_0(\tau + y)/2\chi} \ll 1$ is

$$\begin{aligned}&F_0(\tau + y, \chi) \\ &= \frac{\rho_0}{2\pi}\int_{C_1} \frac{\exp\left[\frac{\zeta}{2}\left(\mu - \frac{1}{\mu}\right) - i\frac{(\tau + y)\beta_0}{2} - i\chi\right]}{i\mu - \sqrt{\beta_0(\tau + y)/2\chi}}d\mu \\ &\approx i\rho_0\left[J_0(\zeta) + \sqrt{\frac{\beta_0(\tau + y)}{2\chi}}J_1(\zeta) + \ldots\right] \\ &\quad\times \exp\left[i\frac{(\tau + y)\beta_0}{2} - i\chi\right].\end{aligned} \tag{50}$$

Here, $J_{0,1}(\zeta)$ is the Bessel function, $\zeta = \sqrt{2\beta_0\chi(y + \tau)}$.

Using expression (39), we obtain

$$\begin{aligned}E(\tau, \chi) &= i\rho_0\left[J_0(\theta) + \sqrt{\frac{\beta_0(\tau)}{\chi}}J_1(\theta)\right] \\ &\times \exp[i(\tau)\beta_0 - i\chi]\left[1 + O\left(\sqrt{\frac{\beta_0(\tau)}{\chi}}\right)\right] + O(|\rho_0|^3).\end{aligned} \tag{51}$$

Here, $\theta = 2\sqrt{\beta_0\chi\tau}$. It follows from this solution that the dynamics of a weak initial acoustic field is described by damped oscillations (see Fig. 3). The nonlinear effects that give a correction on the order of $|\rho_0|^3$ can be easily calculated by iterations using Marchenko equations (31) and (33).

Consider the initial conditions that correspond to an unstable state of the medium: when the field turns on, the impurity ion spins are aligned and directed along the magnetic field. Under the action of perturbations, the system escapes from this state and approaches a stable state in which the spins are aligned along the field but directed oppositely. This asymptotic state corresponds to minimum energy. Let us find a solution for

the shape of the leading edge of the general solution that describes the system transition to a stable state.

We will solve the problem for the initial and boundary conditions that correspond to an unstable state and a small (seed) acoustic deformation of the crystal:

$$|E(\tau, 0)| = \text{const} \ll 1, \quad S_3(-\infty, \chi) = -1, \quad S_\perp(-\infty, \chi) = 0. \tag{52}$$

The seed field $E(\tau, 0)$ causes the decay of this state and the system approaches stable state (40), which is reached when $\chi \longrightarrow \infty$.

As above, we find that $\rho_0(\chi)$ does not depend on λ. Let us calculated the kernel F (35) by taking into account the dependence of the scattering data on χ and conditions (52), i.e., with no allowance for the discrete spectrum. For the initial and boundary conditions (52), it will suffice to take into account only the contribution from the continuum spectrum of the problem. The dependence $\rho(\chi)$ can be determined by using expression (44):

$$\rho(\chi) = \frac{b(\chi; \lambda)}{a(\chi; \lambda)} = \rho_0 \exp\left(-\frac{2i\chi\lambda}{2\lambda + \beta_0}\right). \tag{53}$$

Let us now find an expression for the kernels $F_0(\tau + s)$ and $i\partial_y F_0(\tau + s)$ (35). To this end, we insert the new integration variable $\lambda = \sqrt{\beta_0\chi/2y}\,\mu - \beta_0/2$ in the right-hand side of expression (35). Deforming the integration contour in such a way that it bends around the point $\mu = i$ in the upper half-plane, we obtain

$$F_0(y, \chi) = \frac{\rho_0}{2\pi}\int_{C_1} \frac{\exp\left[\frac{\vartheta}{2}\left(\mu - \frac{1}{\mu}\right) + i\frac{y\beta_0}{2} - i\chi\right]}{\mu - \sqrt{\beta_0 y/2\chi}} d\mu \approx i\rho_0\left[I_0(\vartheta) + \sqrt{\frac{\beta_0 y}{2\chi}} I_1(\vartheta) + \ldots\right]\exp\left(i\frac{y\beta_0}{2} - i\chi\right), \tag{54}$$

where $I_{0,1}$ are the modified Bessel functions and $\vartheta = \sqrt{2\beta_0\chi\tau}$; it is also assumed that $\sqrt{\beta_0 y/2\chi} \ll 1$. Similarly, we find that

$$i\partial_y F_0(y, \chi) = \frac{i\rho_0}{2\pi}\sqrt{\frac{\beta_0\chi}{2y}} \times \int_{C_1} \exp\left[\frac{\vartheta}{2}\left(\mu - \frac{1}{\mu}\right) + i\frac{y\beta_0}{2} - i\chi\right] d\mu = i\rho_0\sqrt{\frac{\beta_0\chi}{2y}} I_1(\vartheta)\exp\left(i\frac{y\beta_0}{2} - i\chi\right). \tag{55}$$

It follows from this expression and from the Marchenko equations that the kernels $|K_{1,2}|$ exponentially increase at the initial stage with increasing τ; i.e., the solution for $|E|$ can reach values on the order of unity for arbitrarily low amplitudes of the seed field. Another conclusion that follows from the derived expressions for kernels (54) and (55) is that the solution for the acoustic field is concentrated at $\theta \gg 1$ for small ρ_0 such that $-\ln|\rho| \gg 1$. In this range, the integrals in Marchenko equations (31) and (33) can be roughly calculated by the saddle-point method. Under these conditions, we find the solution that describes the leading edge of the packet of generated pulses at the nonlinear stage of the process.

To approximately solve Eqs. (31) and (33), we substitute the following expressions for the kernels into them:

$$i\partial_y F_0(\tau + y; \chi) = \frac{\rho_0}{2\pi} \times \int_{C_1} \exp\left[-i\lambda(\tau + y) + i\frac{\chi\beta_0}{2\lambda} + i\frac{(\tau + y)\beta_0}{2} - i\chi\right] d\lambda, \tag{56}$$

$$F_0^*(\tau + y; \chi) = \frac{\rho_0}{2\pi}\int_{C_2} \exp\left[i\mu(\tau + y) + i\frac{\chi\beta_0}{2\mu} + i\frac{(\tau + y)\beta_0}{2} - i\chi\right]\frac{d\mu}{\mu - \beta_0/2}. \tag{57}$$

We substitute $K_1(\tau, y)$ from (33) into (31) and integrate over the variable s. We change to the following variables in the derived integral equation:

$$\lambda = \sqrt{\frac{\chi\beta_0}{2(\tau + y)}}\tilde{\lambda}, \quad \mu = \sqrt{\frac{\chi\beta_0}{2(\tau + y)}}\tilde{\mu}.$$

Subsequently, we deform the contours C_1 and C_2 in such a way that they bend around the points $\tilde{\lambda} = i$ and $\tilde{\mu} = -i$, respectively, in the positive and negative directions. We integrate over these variables by using the saddle-point method. Since the exponentials mainly contribute to the integrals, we can assume that $\sqrt{y + \tau} \approx \sqrt{2\tau}$ and $\sqrt{s + \tau} \approx \sqrt{2\tau}$ in the algebraic factor that emerges in front of the exponential after the integration. Similarly, substituting the expression for $K_2^*(t, y)$ from (32) into (33) and performing the procedure described above, we obtain the second integral equation. The approximation $1 + (\beta_0\chi\tau)^{-1} \approx 1$ was used to derive these equations.

The two derived equations can easily be solved for the new functions:

$$Q_1(\tau) = \int_\tau^\infty K_1(\tau, s) \times \exp\left[\sqrt{2\chi\beta_0(\tau+s)} + is\frac{\beta_0}{2}\right]ds, \quad (58)$$

$$Q_2^*(\tau) = \int_\tau^\infty K_2^*(\tau, s) \times \exp\left[\sqrt{2\chi\beta_0(\tau+s)} - is\frac{\beta_0}{2}\right]ds. \quad (59)$$

These solutions are

$$Q_1(\tau) = \frac{-i}{D^*}\frac{\left(1-\frac{2i\beta}{D}\right)\frac{|\rho_0|^2}{4\sqrt{2}}e^{2\theta}}{1+\left(1-\frac{2i\beta}{D}\right)\frac{|\rho_0|^2}{8\sqrt{2}}e^{2\theta}}, \quad (60)$$

$$Q_2^*(\tau) = \frac{i\rho_0}{2D}\frac{\exp\left(2\theta + i\frac{\beta_0\tau}{2}\right)}{1+\left(1-\frac{2i\beta}{D}\right)\frac{|\rho_0|^2}{8\sqrt{2}}e^{2\theta}}, \quad (61)$$

where $\theta = 2\sqrt{\beta_0\chi\tau}$ and $D = i\beta_0 - \sqrt{\chi\beta_0/\tau}$. Substituting solutions (60) and (61) into the original equations yields the kernels

$$K_1(\tau, \tau) = \frac{i}{D^*} \times \frac{\left(1-\frac{2i\beta}{D}\right)\frac{|\rho_0|^2\sqrt{\beta_0\chi}}{2\sqrt{\tau}}\exp\left(\theta - i\frac{\tau\beta_0}{2}\right)}{1+\left(1-\frac{2i\beta}{D}\right)\frac{|\rho_0|^2}{8\sqrt{2}}e^{2\theta}}, \quad (62)$$

$$K_2^*(\tau, \tau) = \frac{\rho_0}{2D}\frac{\exp\left(\theta + i\frac{\beta_0\tau}{2}\right)}{1+\left(1-\frac{2i\beta}{D}\right)\frac{|\rho_0|^2}{8\sqrt{2}}e^{2\theta}}. \quad (63)$$

Finally, we obtain the following solution for the leading edge of the packet of pulses:

$$E(\tau, \chi) = \frac{\tilde{Z}\frac{\rho_0}{D}\exp\left(\theta + i\frac{\tau\beta_0}{2}\right)}{\tilde{Z}\tilde{Z}^* + \frac{|\rho_0|^2}{4|D|^2}e^{2\theta}}\exp(i\psi_a). \quad (64)$$

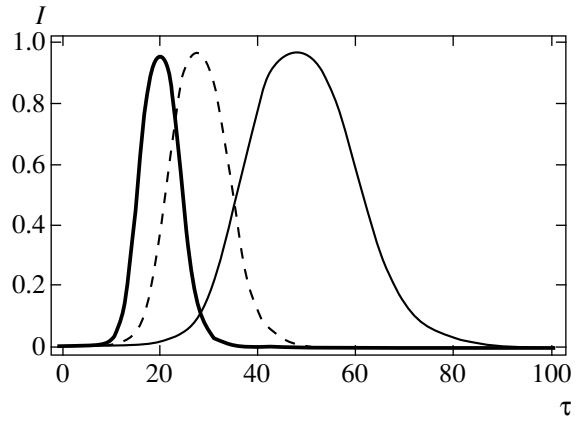

Fig. 4. Dependence $I(\tau) = |E|^2$ of the leading edge of solution (64) for various β_0: $\beta_0 = 0.5$ (thin solid line), $\beta_0 = 1$ (dashed line), and $\beta_0 = 1.5$ (heavy line).

Here,

$$\psi_a = \arg\left(\frac{Z^*}{Z}\right),$$

$$Z = 1 + \left(1-\frac{2i\beta}{D}\right)\frac{|\rho_0|^2}{8\sqrt{2}}e^{2\theta}, \quad (65)$$

$$\tilde{Z} = Z + \left(1-\frac{2i\beta}{D}\right)\frac{|\rho_0|^2\sqrt{\beta_0\chi}}{D^*2\sqrt{\tau}}\exp\left(\theta - i\frac{\tau\beta_0}{2}\right).$$

Solution (64) corresponds to the first nonlinear oscillation or the leading edge of the asymptotic solution that describes the evolution of the spin system from state (52) to stable state (40). In Fig. 4, the intensity $I = |E|^2$ is plotted against τ for various β_0. We found that the leading edge becomes steeper and the generation delay decreases with increasing β_0. The figure shows that, as the longitudinal field contribution increases, the pulse at the leading edge of the packet of pulses that describes the decay of an unstable state shortens. At the same time, the amplitude of this pulse does not change with increasing β_0.

A numerical analysis of the original evolution system (20) and analysis of the asymptotics for the solution indicate that it consists of a packet of damped nonlinear pulsations. In real media, relaxation and diffraction result in a significant relative suppression of the oscillation amplitudes compared to the leading edge. Therefore, in practice, it will often suffice to find an expression for the first nonlinear oscillation (64). The solution obtained shows that the asymptotics for large χ is nonsoliton and is characterized by the self-similar variable $\sqrt{\beta_0\chi\tau}$. As in the soliton case, the asymmetry

due to a nonzero β leads to a nonlinear phase modulation described by the phase ψ_a (65).

4. THE QUASI-MONOCHROMATIC APPROXIMATION

Here, we find a solution for the model that describes the dynamics of acoustic pulses on the order of or shorter than ω_B^{-1} in duration by using the ISTM. This model is the most general integrable reduction of the original system (14)–(17). However, it is also of interest to find other integrable reductions of this model that arise under additional assumptions. In general, these models are easier to solve and analyze. On the other hand, soliton solutions and other coherent structures arise in these models from the balance between dispersion, cross-modulation, nonlinear mixing, etc. The corresponding terms in the equations simulate the real physical effects that show up at various field amplitudes and degrees of spin reversal in our problem. Therefore, it is important to determine the conditions when these effects, while being mutually balanced, give rise to solitons and other coherent structures. Studying these models is also useful for solving similar nonintegrable models, because soliton and other stable solutions of integrable models can be used as a zero approximation in constructing the perturbation theory.

As above, we use the condition of equal group velocities: $v_\perp = v_\parallel$. Let us now pass to quasi-monochromatic fields in system (14)–(17):

$$\frac{\epsilon_\perp G_\perp}{\hbar\omega_B} = \tilde{E}\exp[i(\omega t - kz)], \tag{66}$$

$$S_\perp = R\exp[i(\omega_B t - kz)], \quad |\omega_B - \omega| \ll \omega_B. \tag{67}$$

We use the slow-envelope approximation, which requires the satisfaction of the inequalities

$$\begin{aligned} \left|\frac{\partial\tilde{E}}{\partial t}\right| &\ll \omega_B|\tilde{E}|, \quad \left|\frac{\partial\tilde{E}}{\partial z}\right| \ll k|\tilde{E}|, \\ \left|\frac{\partial R}{\partial t}\right| &\ll \omega_B|R|, \quad \left|\frac{\partial R}{\partial z}\right| \ll k|R|. \end{aligned} \tag{68}$$

Let us change to the variables

$$\tilde{z} = z\frac{knG_\perp^2}{v^2 n_0 \hbar\omega_B}, \quad \tilde{\tau} = \left(t - \frac{z}{v}\right)\omega_B$$

and denote

$$W = \frac{\epsilon_\parallel G_\perp}{\hbar\omega_B}, \quad \tilde{E} = \frac{\epsilon_\perp G_\perp}{\hbar\omega_B}, \quad \nu_0 = \frac{\omega_B - \omega}{2\omega_B}.$$

To simplify the description of the dynamics of the longitudinal field $\epsilon_\parallel$, we use the condition of unidirectional field propagation. Given these approximations and changes, system (14) and (15) takes the form

$$\frac{\partial\tilde{E}}{\partial\tilde{z}} = iR, \tag{69}$$

$$\frac{\partial W}{\partial\tilde{z}} = -\frac{\beta_0}{2}\frac{\partial S_3}{\partial\tilde{\tau}}. \tag{70}$$

We find from Eqs. (69), (70), and (16) that the fields W and $\tilde{E}$ are related by

$$W(\tilde{\tau},\tilde{z}) = -\frac{\beta_0}{4}|\tilde{E}(\tilde{\tau},\tilde{z})|^2 + W_1(\tilde{\tau}). \tag{71}$$

Here, $W_1(\tilde{\tau})$ is determined by boundary conditions. Without loss of generality, we choose $W_1 \equiv 0$. Using equality (71), we reduce the Bloch equations (15) and (16) to

$$\frac{\partial R}{\partial\tilde{\tau}} = i\left(2\nu_0 - \frac{\beta_0^2}{4}|\tilde{E}|^2\right)R + i\tilde{E}S_3, \tag{72}$$

$$\frac{\partial S_3}{\partial\tilde{\tau}} = \frac{i}{2}(\tilde{E}^* S_\perp - \tilde{E}S_\perp^*). \tag{73}$$

As a result, we obtain the system of equations (69), (72), and (73), which is formally identical to our integrable system suggested previously [14]. This system was used to describe the generation and evolution of ultrashort electromagnetic light pulses in two-level optical media in the quasi-monochromatic approximation. In [23], we found soliton and periodic solutions for this model. The Lax representations for the integrable system (69), (72), and (73) are

$$\partial_{\tilde{\tau}}\Phi = \begin{pmatrix} -i\lambda^2 - i\frac{\beta_1}{4}|\tilde{E}^2| & \gamma\tilde{E} \\ \tilde{\gamma}\tilde{E}^* & i\lambda^2 + \frac{\beta_1}{4}|\tilde{E}^2| \end{pmatrix}\Phi \equiv \mathrm{L}_2\Phi, \tag{74}$$

$$\partial_{\tilde{z}}\Phi = \frac{1}{4(\lambda^2+\nu_0)}\begin{pmatrix} -iS_3 & 2\gamma R \\ 2\tilde{\gamma}R^* & iS_3 \end{pmatrix}\Phi \equiv \mathrm{A}_2\Phi, \tag{75}$$

where

$$\beta_1 = \frac{\beta_0^2}{2}, \quad \gamma = \frac{1}{2}(\beta_1\lambda + i\sqrt{1+\beta_1^2\nu_0}),$$

$$\tilde{\gamma} = -\frac{1}{2}(\beta_1\lambda - i\sqrt{1+\beta_1^2\nu_0}).$$

Some information on the field dynamics can be obtained by analyzing the structure of this Lax pair and by comparing it with a similar Lax pair in [23]. The contribution of the longitudinal acoustic field shows up in the presence of terms with the coefficient β_1 in the matrices L_2 and A_2. As a result, for soliton and other

solutions, inclusion of the longitudinal field manifests itself in a change of the pulse shape and in the appearance of a nonlinear phase addition on the order of $i\beta_1\int_0^{\tilde{\tau}}|\tilde{E}|^2 d\tilde{\tau}$. Since the contribution of the longitudinal field for a small ratio $G_\parallel/G_\perp \sim \varepsilon$ is on the order of ε^2, its contribution to the dynamics of the transverse field in such media can be disregarded, which is attributable to the quasi-monochromatic approximation used above. In the case of ultrashort pulses considered above, i.e., for pulses $\tau_p \sim \omega_B^{-1}$ in duration, the contribution of the longitudinal field is the same in order of magnitude as that of the transverse field. This difference stems from the fact that Eqs. (69) and (71) describe the long-wave–short-wave resonance [7], which is much less effective than the short-wave resonance in the case considered in preceding sections of this paper. We can conclude that the effects related to the longitudinal field are much more pronounced for ultrashort pulses than those in the quasi-monochromatic limit.

In deriving system (69), (72), and (73), we assumed the transverse field to produce rapid oscillations between Zeeman levels. In this case, the nonlinear effects are mainly attributable to the interaction of the transverse field with a two-level medium. Consider the other extreme case, where there are virtually no transitions between levels; i.e., the change in S_3 can be ignored. Applying the approximations used above, we obtain the following additional reduction of Eqs. (69), (72), and (73):

$$\frac{\partial\tilde{E}}{\partial\tilde{z}} = iR, \tag{76}$$

$$\frac{\partial R}{\partial\tilde{\tau}} = i\left(2\nu_0 - \frac{\beta_1}{2}|\tilde{E}|^2\right)R + i\tilde{E}. \tag{77}$$

This integrable system of equations can be reduced to the Thirring model by a simple gauge transformation. It also has stable soliton and other coherent solutions and can be analyzed in detail in terms of the ISTM (see, e.g., [15]). In this case, the existence of soliton and other coherent structures is attributable to the nonlinear phase modulation produced by the longitudinal field. The transverse field manifests itself in the establishment of a coherent coupling between the field and the two-level medium in the linear limit. This example shows that the physical conditions corresponding to a small change in Zeeman level population should be used to observe the coherent nonlinear effects produced by the longitudinal field alone.

5. DISCUSSION

We studied the dynamics of ultrashort acoustic pulses in terms of integrable reductions of the evolution equations that describe the dynamics of a longitudinal–transverse wave propagating along the magnetic field in a medium of impurity ions with an effective spin of 1/2. Let us estimate the parameters of the medium and the intensity of the soliton acoustic pulses that can be generated in such a medium. As an example, we choose a MgO crystal with Fe^{2+} paramagnetic impurities at a liquid helium temperature of $T = 4$ K. The effect of the thermal Zeeman level population decreases at low temperatures. Let the magnetic field strength be such that the Zeeman splitting is $\omega_B = 10^{12}$ s^{-1}. The ratio of thermal energy to Zeeman energy is $k_BT/\hbar\omega_B = 0.5$, where k_B is the Boltzmann constant. In this case, the condition $\tau_p \sim \omega^{-1}$ holds for pulse durations on the order of 100 ps. Under these conditions, the slow-envelope approximation is inapplicable, which corresponds to the case considered here except for the preceding section. When passing to longer pulses, for example, with $\tau_p \sim 1$ ns, the slow-envelope approximation can be applied in terms of the models considered in the preceding section. Note that the condition $\omega_B = 10^{12}$ s^{-1} corresponds to magnetic field strengths that are quite attainable in laboratory conditions. For the coefficients of the medium, we have [8, 16] $G_\perp$, $G_\parallel \sim 10^{-13}$ erg; $n \sim 10^{19}$ cm^{-3}; $n_0 \sim$ 3–4 g/cm^3; $v_\perp$, $v_\parallel \approx (5\text{–}10) \times 10^5$ cm/s; and λ_{11}, $\lambda_{44} \approx (5\text{–}10) \times 10^{11}$ dyn/cm^2. Under these conditions, the intensity of the soliton signal can be $I \sim 10^6$ W/cm^2 [7].

Here, we used only the condition of equal group velocities for the longitudinal and transverse waves and the approximation of unidirectional acoustic wave propagation to study the dynamics of ultrashort pulses. Compared to the results of [6, 7], this allowed us to find a soliton solution corresponding to a more general physical situation and a nonsoliton solution corresponding to the decay of an unstable state. On the other hand, to observe the behavior of the field described here requires milder physical conditions. For example, the condition $\tau_p \ll \omega_B^{-1}$, which was used in [6], requires using ultralow temperatures on the order of 0.1 K, as was pointed out in the same paper. When increasing the temperature by several degrees, the magnetic field strength must be increased because of the thermal level population. This necessitates using acoustic pulses with duration on the order of 1 ps, which, in turn, necessitates an allowance for dispersion. The condition $\tau_p \ll \omega_B^{-1}$ is not required to deduce the integrable model (20) found here.

One-soliton solutions were found in most of the theoretical studies devoted to the dynamics of an acoustic wave in similar problems. For one-phase solutions, to which one-soliton and one-zone periodic solutions belong, the condition of equal group velocities for the transverse and longitudinal waves is unimportant. It is easy to show that, in this case, a reduction equivalent to the corresponding self-similar reduction of Eqs. (16)–(19), to within the renormalization of the coefficients, arises. If the solution for the functions in Eqs. (16)–(19) depends on the self-similar variable $\tau - z/v_s$, then the

one-soliton solution obtained above does not change except for the changes

$$\tau \longrightarrow \tau\sqrt{\frac{v_s - v_\parallel}{v_s - v_\perp}}, \quad \beta_0 \longrightarrow \beta_0\sqrt{\frac{v_s - v_\parallel}{v_s - v_\perp}}.$$

It thus follows, in particular, that the group velocity difference ($v_\parallel \neq v_\perp$) gives rise to a nonlinear phase modulation of the soliton in a symmetric crystal with $G_\parallel = G_\perp$ as well. Strictly speaking, for the model under study, this result is valid for a low value of $(v_\parallel - v_\perp)/v_\perp$, because we used the approximation of unidirectional wave propagation. After simple changes of variables and parameters, similar one-soliton solutions of the models presented in Section 4 can be used to describe the ASIT for unequal group velocities of the transverse and longitudinal waves.

In pure form, the ASIT-associated soliton dynamics requires producing a sufficiently intense pulse with a nearly soliton shape at the boundary of the medium for its observation. At the same time, if the spin system was initially in an unstable state, then the solution that describes the transition to a stable state is a nonsoliton one. In Section 3.4, we derived an expression for the leading edge of the packets of nonlinear oscillations generated in this case. Asymptotically, the dependence of the shape of these oscillations on variables is mainly determined by the variable $\sqrt{\beta_0\chi\tau}$. At $\beta = 0$, i.e., for $G_\perp = G_\parallel$, the corresponding matrix $\hat{L}_1$ in Eq. (22) is symmetric and the gauge transformation can be applied to system (22) and (23). The latter reduces the solution of the problem to studying the Zhakharov–Shabat spectral problem. This leads to the solution of the simpler system of Marchenko equations. It can be shown for the initial and boundary conditions (51) that finding the asymptotics reduces to solving a first-order equation. The right-hand side of this equation is proportional to the self-similar solution of the sine-Gordon equation or the equations equivalent to the Maxwell–Bloch equations for a two-level optical medium at an exact resonance. This solution is completely determined by the real continuum spectrum of the problem (22) and is identical to the solution of the Painleve III equation [24].

Here, we disregarded the nonlinear effects related to the anharmonicity of the crystal lattice. At the same time, the relative crystal deformation in the region of a picosecond soliton can reach $P/n_0 v \sim 10^{-3}$, where the pressure $P \sim 1$ kbar [25]. Under these conditions, the nonlinearities related to the anharmonicity of atomic vibrations at lattice sites significantly affect the dynamics of the acoustic soliton. Including these effects can qualitatively change the dynamics of acoustic waves in the physical system considered above. Although we ignored this kind of effects, they can be included in the integrable models found above as additional perturbing terms.

As another possible physical application of our results, consider the recently predicted effect of anomalous stimulated Brillouin scattering associated with the slowdown of light [26]. Recent progress in coherent nonlinear optics is attributable to the development of a theory for electromagnetically induced transparency [27]. This phenomenon, which was observed in multilevel optical media, causes the group velocity of light in the medium to significantly decrease to values comparable to or even lower than the speed of sound in this medium in the presence of highly variable linear dispersion. Such a decrease in the group velocity of light was experimentally found in cold and hot atomic gases and in crystals doped by rare-earth ions. The decrease in resonant absorption along with strong linear dispersion opens up new mechanisms of resonant interaction between light and acoustic waves. Harris [28] showed that the longitudinal gradient forces acting on two-level atoms increase when an ultraslow optical pulse traveling in a coherent medium is spatially compressed. This increase in force leads to the formation of an atomic bullet and opens up new mechanisms of local ponderomotive light scattering. Matsko *et al.* [29] argue that, because of the strong linear dispersion associated with electromagnetically induced transparency, phase matching can be achieved between the electromagnetic and acoustic waves in a dielectric optical fiber doped by three-level ions. To describe the interaction of an electromagnetic wave with acoustic phonons, Matsko *et al.* [26] used a Hamiltonian similar to that used here to describe the coupling between spin states and acoustic phonons. A resonant optical medium can be simulated by an effective two-level medium. The ponderomotive force associated with local density variations in the optical medium gives a contribution to the acoustic field dynamics similar to the contribution described by the right-hand sides of Eqs. (14) and (15). Apart from the transverse electromagnetic field, the longitudinal acoustic field, whose contribution to the dynamics of the optical medium manifests itself in the emergence of a nonlinear phase modulation, should be taken into account in this scheme. Note that the quasi-monochromatic approximation can be applied to the optical field in this model. This approximation, together with the condition of a low doped ion density, can lead to the reduced integrable model derived in the preceding section. The Brillouin scattering for picosecond acoustic pulses can be studied in terms of this model. The phase-matching condition allows the interaction of these pulses with quasi-monochromatic optical pulses of similar duration to be investigated. The Brillouin scattering mechanism for such ultraslow light can be used to effectively amplify the pulses propagating in ion-doped optical fibers.

REFERENCES

1. A. I. Maimistov and A. M. Basharov, *Nonlinear Optical Waves* (Kluwer, Dordrecht, 1999).

2. A. I. Maĭmistov, Kvantovaya Élektron. (Moscow) **30**, 287 (2000).
3. 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).
4. H.-Y. Hao and H. J. Maris, Phys. Rev. B **64**, 064302 (2001).
5. G. A. Denisenko, Zh. Éksp. Teor. Fiz. **60**, 2269 (1971) [Sov. Phys. JETP **33**, 1220 (1971)].
6. S. V. Voronkov and S. V. Sazonov, Fiz. Tverd. Tela (St. Petersburg) **43**, 1969 (2001) [Phys. Solid State **43**, 2051 (2001)].
7. S. V. Voronkov and S. V. Sazonov, Zh. Éksp. Teor. Fiz. **120**, 269 (2001) [JETP **93**, 236 (2001)].
8. V. A. Golenishchev-Kutuzov, V. V. Samartsev, N. K. Solovarov, and B. M. Khabibulin, *Magnetic Quantum Acoustics* (Nauka, Moscow, 1997).
9. N. S. Shiren, Phys. Rev. B **2**, 2471 (1970).
10. V. V. Samartsev, B. P. Smolyakov, and R. Z. Sharipov, Pis'ma Zh. Éksp. Teor. Fiz. **20**, 644 (1974) [JETP Lett. **20**, 296 (1974)].
11. C. Kittel, *Introduction to Solid State Physics*, 5th ed. (Wiley, New York, 1976; Nauka, Moscow, 1974).
12. A. A. Zabolotskii, Phys. Lett. A **127**, 83 (1988).
13. A. A. Zabolotskii, Physica D (Amsterdam) **40**, 283 (1989).
14. A. A. Zabolotskii, Phys. Lett. A **124**, 500 (1987).
15. A. A. Zabolotskii, Phys. Rev. A **57**, 2958 (1998).
16. J. W. Tucker and V. W. Rampton, *Microwave Ultrasonics in Solid State Physics* (North-Holland, Amsterdam, 1972; Mir, Moscow, 1975).
17. L. D. Landau and E. M. Lifshitz, *Course of Theoretical Physics*, Vol. 7: *Theory of Elasticity*, 4th ed. (Nauka, Moscow, 1987; Pergamon, New York, 1986).
18. S. A. Al'tshuller and B. M. Kozyrev, *Electron Paramagnetic Resonance in Compounds of Transition Elements* (Nauka, Moscow, 1981; Halsted, New York, 1975).
19. Y. R. Shen, *The Principles of Nonlinear Optics* (Wiley, New York, 1984; Nauka, Moscow, 1989).
20. J. D. Gibbon, P. J. Coudrey, J. K. Eilbeck, and R. K. Bullough, J. Phys. A **6**, 1237 (1973).
21. A. E. Borovik and S. I. Kulinich, Pis'ma Zh. Éksp. Teor. Fiz. **39**, 320 (1984) [JETP Lett. **39**, 384 (1984)].
22. L. A. Takhtadzhyan and L. A. Faddeev, *Hamiltonian Methods in the Theory of Solitons* (Nauka, Moscow, 1986; Springer, Berlin, 1987).
23. A. A. Zabolotskii, Zh. Éksp. Teor. Fiz. **107**, 1100 (1995) [JETP **80**, 614 (1995)].
24. A. A. Zabolotskii, Zh. Éksp. Teor. Fiz. **115**, 1158 (1999) [JETP **88**, 642 (1999)].
25. S. A. Akhmanov and V. É. Gusev, Usp. Fiz. Nauk **162**, 3 (1992) [Sov. Phys. Usp. **35**, 153 (1992)].
26. A. B. Matsko, V. V. Rostovstev, M. Flieschhauer, and M. O. Scully, Phys. Rev. Lett. **86**, 2006 (2001); V. Kovalev, Phys. Rev. Lett. **88**, 239301-1 (2002).
27. S. E. Harris, J. E. Field, and A. Imamoglu, Phys. Rev. Lett. **64**, 1107 (1990).
28. S. E. Harris, Phys. Rev. Lett. **85**, 4032 (2000).
29. A. B. Matsko, Yu. V. Rostovtsev, H. Z. Cummins, and M. O. Scully, Phys. Rev. Lett. **84**, 5752 (2000).

Translated by V. Astakhov

Journal of Experimental and Theoretical Physics, Vol. 96, No. 3, 2003, pp. 510–522.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 123, No. 3, 2003, pp. 575–589.

SOLIDS
Structure

Antiferromagnetic-Resonance Spectrum in Charge-Ordered $R_{0.5}Ca_{0.5}MnO_3$ Manganites (R = La, Pr, Tb): The Effect of Orbital and Charge Structures

L. É. Gonchar'* and A. E. Nikiforov
Ural State University, Yekaterinburg, 620083 Russia
**e-mail: lyudmila.gonchar@usu.ru*
Received August 21, 2002

Abstract—A theoretical study is made into the effect of the crystal, orbital, and charge structures on the magnetic structure and spin-wave spectra and on the antiferromagnetic resonance (AFMR) for $R_{0.5}Ca_{0.5}MnO_3$ crystals of monoclinic structure. The model assumes fixed crystal, charge, and orbital structures and enables one to determine the orbitally dependent exchange interaction and single-ion anisotropy for R = La, Pr, Tb. A 16-sublattice weakly noncollinear magnetic *CE*-structure without a ferromagnetic component is obtained. The behavior of magnetic structure in an external magnetic field is simulated, and the values of fields of spin-flop-transition for different Rs are obtained. The law of spin-wave dispersion and the field dependence of the antiferromagnetic-resonance spectrum are calculated. © *2003 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

Attention given to manganites at present is not confined to the well-known effect of colossal magnetoresistance and is directed at a large number of unusual properties. The dielectric phases of manganites likewise are becoming the subject of thorough theoretical and experimental studies. Certain difficulties involved in studying these compounds are associated with the description of all crystal subsystems, namely, the crystal lattice and the charge, orbital and spin subsystems, with regard for their correlation. This paper deals with the magnetic properties of dielectric $R_{0.5}Ca_{0.5}MnO_3$ manganites (R = La, Pr, Tb) of monoclinic structure. The objective of our study is to demonstrate the effect of the crystal, charge, and orbital structures on magnetic ordering and spectra of magnons.

The investigation of the correlation between the orbital and magnetic structures of manganites was begun quite some time ago. Starting with the studies by Goodenough [1] and Wollan and Koehler [2], the presence of such a correlation was determined qualitatively. The quantitative characteristics of interaction of all four subsystems are still being discussed at present. Several models exist, which are used for the description of manganites. The effect of lattice distortions on the magnetic structure is recognized today by most researchers [3]; however, the inclusion of Jahn–Teller lattice distortions, which play the main part in the formation of unusual properties of manganites, is done differently by different authors. In the charge-ordered phases of manganites, the ordering of localized charge carriers formed due to nonisovalent doping is assumed.

Two models are largely used at present for explaining the correlation between the crystal, orbital, and magnetic structures in manganites. The first of these models, referred to as the Kugel'–Khomskii model [4], implies the formation of the orbital structure owing to the orbitally dependent exchange interaction eventually causing a crystal lattice distortion. Various modifications of this model are used in numerous studies [5–9]. This model is used quite frequently for prediction and investigation of orbital excitations (orbitons) [6, 8–10]. However, the temperature of decomposition of the orbitally ordered phase is much higher than the Neél temperature [11, 12], which proves the assumption that the orbital structure is formed as a result of stronger-than-exchange interactions. The second model, referred to as the Kanamori model [13], presumes the presence of the cooperative Jahn–Teller effect. This model likewise presumes the presence of the orbitally dependent exchange interaction; however, the main part is played by the electron-vibrational interaction. The development of the Kanamori model for manganites is described in [14–17].

The description of superexchange interaction in many-electron systems is a fairly complicated problem. In some compounds, the signs of exchange parameters may be defined by the Goodenough–Kanamori rules [1]. These rules, however, do not define the values of exchange integrals. The exchange parameters were originally calculated [5] for pure manganite; however, they turned out to be overestimated compared to experiment [18, 19]. The general approach to the description of pure and doped manganites is based on the Hubbard model with regard for Hund interaction, Coulomb interaction (in-site and intersite), and Jahn–Teller inter-

1063-7761/03/9603-0510$24.00

action in different models. As a rule, the double-exchange contribution is added to the Hamiltonian [6, 9, 15, 16, 20–22] in order to describe additional charge carriers emerging when the rare-earth sublattice is doped with alkali-earth ions. This approach is characterized by a simplified model of the orbital structure ($d_{3z^2-r^2}$ or $d_{x^2-y^2}$-orbitals, with the *x*, *y*, and *z* axes assumed to be dependent on the position of manganese ion). In some studies, the t_{2g} shell is disregarded. The magnetic structure (including pure and charge-ordered manganites) is described in the double-exchange model [5, 7, 16]. Some models [21, 23] are incapable of describing the experimentally observed [2, 24–26] magnetic structure of charge-ordered manganites of the *CE* type. Therefore, it is not always possible to describe the experimental data and more exact calculations need to be performed. The large number of experimental studies into the crystal, charge, orbital, and magnetic structures makes it possible to apply the semiempirical approach to the description of the mechanism of superexchange interaction in dielectric manganites, which was developed in previous papers [17, 27].

The magnetic structure of half-doped charge-ordered manganites was studied in sufficient detail for different compositions of the rare-earth/alkali-earth sublattice [2, 24–26, 28–30]. These compounds are characterized by the presence or absence of charge ordering, and this feature is of main importance from the standpoint of formation of some or other magnetic structure. In the charge-ordered phase, the Mn^{3+} and Mn^{4+} ions are present in equal amounts and are ordered in space, the orbital structure is characteristic of the sublattice of Mn^{3+} ions, the magnetic structure corresponds to the *CE* type, and the easy magnetic axis is directed mainly along the *c* axis (in *Pnma* notation) [2, 24–26, 28, 29, 31]. A fairly accurate symmetry classification was performed for pure manganites [32]; however, there exists the problem of performing a classification of the *CE*-structure proper of the charge-ordered phase and experimental determination of noncollinear components [33]. The charge, orbital, and *CE*-magnetic structures may be destroyed by temperature or magnetic field [3].

We used the model of the orbital and magnetic structures [17, 27, 34], based on fixed crystal and charge structures, strong electron-vibrational interaction, and orbitally dependent superexchange. The transition to the charge-ordered phase leads to localization of holes on manganese ions with the formation of an ordered structure. In compounds of half-doped orthomanganites, the Mn^{3+}/Mn^{4+} ions alternate in the *ac* plane and do not alternate along the *b* axis (*C* type of charge structure). In so doing, the Mn^{3+} ions are Jahn–Teller ions and the Mn^{4+} ions do not exhibit a degeneracy of the ground state in an ideal octahedral coordination. In view of these features, it is possible to use our model [17] to quantitatively estimate the parameters of superexchange interaction and Neél temperature in dielectric manganites. In this study, we calculated the characteristic features of a multisublattice magnetic structure of monoclinic charge-ordered manganites, the spin-wave spectra, and the field dependence of antiferromagnetic resonance under conditions of the external magnetic field directed along the easy magnetic axis. The suggested model is semiempirical; however, it enables one to obtain the magnetic structure and spectra of magnons proceeding from the crystal structure. Our results may be used for interpreting experimental data in case an exact magnetic structure is required.

2. CRYSTAL AND ORBITAL STRUCTURES

At low temperatures, the $R_{0.5}Ca_{0.5}MnO_3$ crystals being treated exhibit a distorted perovskite structure of symmetry $P2_1/m$ and charge ordering [2, 24–26]. This group is a *Pnma* subgroup which describes the space symmetry of a crystal of pure manganite. Because of the presence of manganese ions in different charge states, which are arranged in a staggered order in the basal plane (see Figs. 1 and 2), additional distortions arise in the crystal. The sublattice of manganese ions is divided into three positions (see Table 1). The ions in positions *a* and *b* exhibit *e*-type distortions of oxygen coordination, and the ions in position *f* do not. The orthorhombic distortion of an ideal perovskite lattice [35] varies in accordance with low-symmetry distortion due to the charge nonequivalence of the ions.

1. An *R*-type distortion is a rotation about the pseudocubic $[110]_p$ axis with cell doubling on all three axes ($\{k_{13}\}\tau_9(C_1C_10)$ in the notation of Kovalev [36] or ($\varphi\varphi0$) in the notation of [37]). Three values of the angle φ, namely, φ_1 (*a*), φ_2 (*b*), and φ_3 (*f*), are distinguished in the monoclinic structure, which correspond to this distortion in sign and are different in magnitude.

2. An *M*-type distortion is a rotation of oxygen octahedrons about the $[001]_p$ axis with cell doubling on two axes ($\{k_{11}\}\tau_3(00C_2)$ in the notation of [36] or (00ψ) in the notation of [37]). Three values of the angle ψ, namely, ψ_1 (*a*), ψ_2 (*b*), and ψ_3 (*f*), are distinguished in the monoclinic structure, which correspond to this distortion in sign and are different in magnitude.

3. A Q_ε-type distortion describes the deformation of *e*-type oxygen octahedrons (see Fig. 2) with cell doubling on two axes ($\{k_{11}\}\tau_5$ in the notation of [36] and with the choice of ray $[1/2\ 1/2\ 0]_p$). In the monoclinic structure, this distortion is characteristic only of the coordination of trivalent manganese ions. The doubling occurs on a single $[110]_p$ axis.

4. The coordination of tetravalent manganese is characterized by an *a*-type distortion (uniform compression) because of the smaller radius of the Mn^{4+} ion compared to Mn^{3+}.

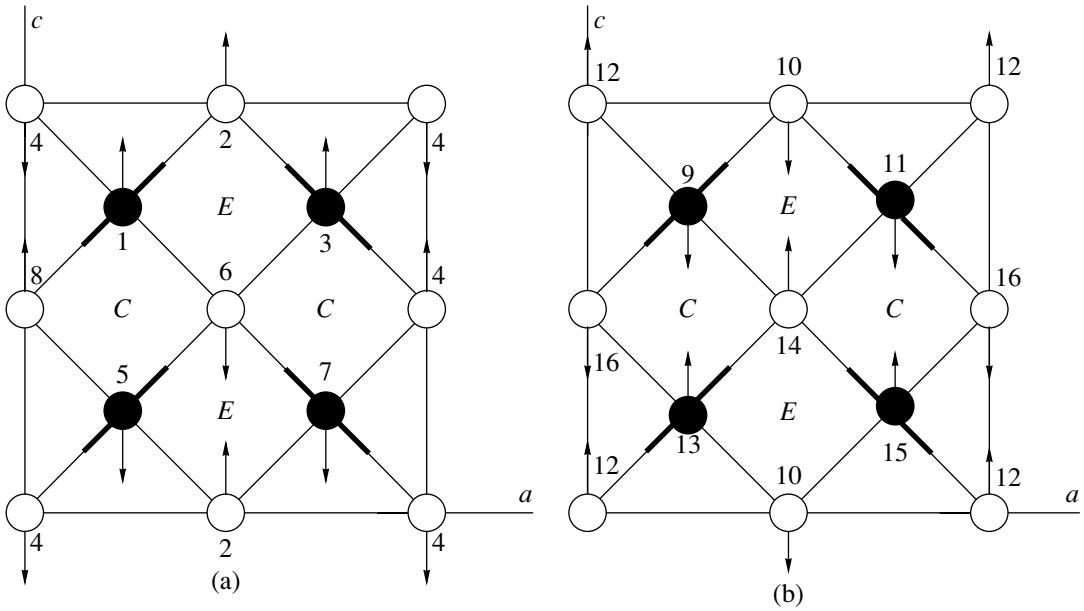

Fig. 1. The orbital and magnetic structures of the charge-ordered phase of half-doped manganite in two neighboring planes along the Y axis (parts (a) and (b)). The ions of oxygen and rare-earth/alkali-earth sublattice are omitted. The hollow symbols indicate Mn^{4+} ions, and the solid symbols indicate Mn^{3+} ions. The arrows indicate the main directions of the magnetic sublattices. The bold lines indicate the orientations of the e_g orbitals. The numbers label the magnetic sublattices.

The corresponding basal distortions are accompanied by adjustment of the lattice; as a result, a Q_θ-type distortion appears and *e*-type distortions of two positions of the Mn^{3+} ion become different in magnitude. Because of the interaction between the *e*-type distortions around Mn^{3+} ions and *a*-type distortions around Mn^{4+} ions, the Mn^{4+} ions shift along the $[\bar{1}10]_p$ axis. This shift defines the wave vector of the orbital structure [17].

Owing to the cooperative Jahn–Teller effect, the orbitally degenerate ground state ^5E of trivalent manganese ions splits. An orbital state with the wave function

$$\psi_n = \varphi_{n\theta}\sin\frac{\phi_n}{2} + \varphi_{n\varepsilon}\cos\frac{\phi_n}{2} \tag{1}$$

is reached on each trivalent manganese ion, where $(\varphi_{n\theta}, \varphi_{n\varepsilon})$ are *E*-level basis functions and the angles ϕ_n obey the relation (see Fig. 1) [27]

$$\begin{aligned}\phi_a \approx -\phi_b, \quad \phi_1 = \phi_5 = \phi_9 = \phi_{13} = \phi_a,\\ \phi_3 = \phi_7 = \phi_{11} = \phi_{15} = \phi_b.\end{aligned} \tag{2}$$

The magnitudes of these angles are expressed in terms of distortions in the crystal,

$$\sin\phi_n = \frac{Q_{\varepsilon n}}{\sqrt{Q_{\theta n}^2 + Q_{\varepsilon n}^2}}, \quad \cos\phi_n = \frac{Q_{\theta n}}{\sqrt{Q_{\theta n}^2 + Q_{\varepsilon n}^2}}. \tag{3}$$

The values of angles of the orbital structure for the compounds being treated are given in Table 2. In the $P2_1/m$ group, *e*-type distortions depend on the shifts of ions (see Table 1) as

$$\begin{aligned}Q_{\theta a} &= \frac{1}{\sqrt{12}}\left[b - \frac{1}{\sqrt{2}}\left(\frac{a}{2} + c\right)\right]\\ &+ \frac{1}{\sqrt{6}}[-(v_{X1} + v_{X3})a + (v_{Z1} - v_{Z3})b],\\ Q_{\varepsilon a} &= \frac{1}{\sqrt{2}}[(v_{X1} - v_{X3})a - (v_{Z1} + v_{Z3})b],\\ Q_{\theta b} &= \frac{1}{\sqrt{12}}\left[b - \frac{1}{\sqrt{2}}\left(\frac{a}{2} + c\right)\right]\\ &+ \frac{1}{\sqrt{6}}[-(v_{X1} + v_{X3})a + (v_{Z2} - v_{Z4})b],\\ Q_{\varepsilon b} &= \frac{1}{\sqrt{2}}[(v_{X1} - v_{X3})a - (v_{Z2} + v_{Z4})b].\end{aligned} \tag{4}$$

The orbital structure of charge-ordered manganite is given in Fig. 1.

The "zigzag" space distribution of Jahn–Teller distortions in a crystal may be qualitatively explained in view of the shift of an Mn^{4+} ion along the *Z* axis caused by the adjustment of the lattice to the breathing distortion of the coordination of Mn^{4+} ions and to the Jahn–

Teller distortion of the coordination of Mn^{3+} ions. This ion "carries along" the neighboring bonds, as a result of which two mutually perpendicular bonds with oxygen become elongated and two bonds shorten. Therefore, the line of shift of tetravalent manganese is the dividing line between trivalent manganese ions with positive and negative Q_ε-type distortions, and the sign of these distortions does not vary along the direction of the shift. In view of the fact that the sign of the shift of an Mn^{4+} ion varies along the X axis, it is obvious that the primitive cell is doubled in this direction. The part played by the shift of tetravalent manganese in the formation of the orbital structure of $La_{0.5}Ca_{0.5}MnO_3$ was previously discussed by Goodenough [1]; however, he suggested shifts in two mutually perpendicular pseudocubic directions. In [24–26], the shift of this ion along precisely the Z axis is assumed.

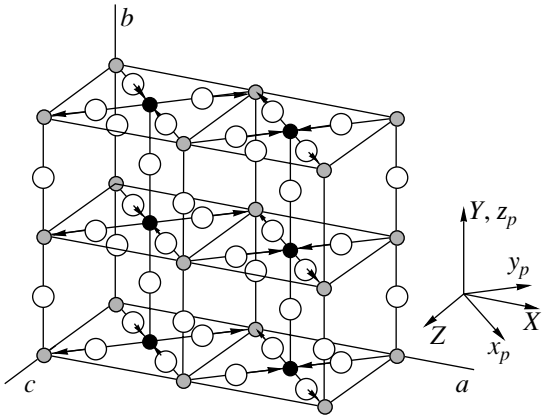

Fig. 2. A fragment of an $R_{0.5}Ca_{0.5}MnO_3$ cell with an ε-type distortion. The R^{3+} and Ca^{2+} ions are omitted; X, Y, Z denote the system of coordinates of the *Pnma* and $P2_1/m$ groups; and x_p, y_p, z_p denote a quasi-cubic system of coordinates.

3. ORBITAL DEPENDENCE OF MAGNETIC INTERACTIONS

Isotropic Exchange

The dependences of the parameters of isotropic exchange of manganese ions in a crystal of charge-ordered manganite on the orbital state of interacting ions of manganese in the crystal were obtained in [17] in view of the difference between the metal-ligand distances r_b (Mn^{3+}–O), r_b' (Mn^{4+}–O), and r_{ac} and the angles of superexchange bond Θ_b, Θ_b', and Θ_{ac} for

Table 1. The coordinates of atoms in a primitive crystal cell in the charge-ordered phase (group of symmetry $P2_1/m$; compared to the reference book [36], the axes are turned for comparison with the designations of *Pnma* axes: *xyz* [36] = *ZXY*)

Ion (position)	$X(a)$	$Y(b)$	$Z(c)$
R, A(1)-2*e*	$1/4 + V_{XR1}$	0	V_{ZR1}
R, A(2)-2*e*	$3/4 + V_{XR1}$	0	V_{ZR2}
R, A(3)-2*e*	V_{XR3}	0	V_{ZR3}
R, A(3)-2*e*	$1/2 + V_{XR3}$	0	V_{ZR3}
Mn^{3+}(1)-2*a*	0	1/4	0
Mn^{3+}(2)-2*b*	1/2	–1/4	0
Mn^{4+}(3)-4*f*	1/4	1/4	$1/2 - V_{ZMn}$
O(1)-2*e*	u_{X1}	0	u_{Z1}
O(2)-2*e*	$1/2 + u_{X1}$	0	u_{Z1}
O(3)-2*e*	$1/4 + u_{X1}$	0	$1/2 + u_{Z2}$
O(4)-2*e*	$3/4 + u_{X1}$	0	$1/2 + u_{Z3}$
O(5)-4*f*	$+ v_{X1}$	$-1/4 + v_Y$	$1/4 + v_{Z1}$
O(6)-4*f*	$5/8 + v_{X1}$	$-1/4 + v_Y$	$1/4 + v_{Z2}$
O(7)-4*f*	$1/8 + v_{X3}$	$-1/4 - v_Y$	$3/4 + v_{Z3}$
O(8)-4*f*	$5/8 + v_{X3}$	$-1/4 - v_Y$	$3/4 + v_{Z4}$

Table 2. The characteristic parameters of distortions of the crystal structure on the charge-ordered phases of crystals of $R_{0.5}Ca_{0.5}MnO_3$ manganites (given along with the name of the compound is the reference to the experimental data on which the calculation was based)

	La [24]	Pr [25]	Tb [26]
ϕ_a, deg	–126.0°	–123.6°	–111.1°
ϕ_b, deg	126.3°	120.6°	112.5°
φ_1, deg	–9.1°	–9.9°	11.6°
φ_2, deg	–9.1°	–9.9°	11.6°
φ_3, deg	–9.3°	–10.1°	11.7°
ψ_1, deg	5.4°	8.0°	–8.8°
ψ_2, deg	3.5°	4.6°	–11.6°
ψ_3, deg	4.7°	6.5°	10.2°
Θ_b, deg	158.3°	156.3°	152.5°
Θ_b', deg	158.3°	156.3°	152.9°
Θ_{ac}, deg	162.3°	157.9°	152.2°
r_b, Å	1.92	1.91	1.90
r_b', Å	1.92	1.91	1.90
r_{ac}, Å	1.96	1.94	1.97

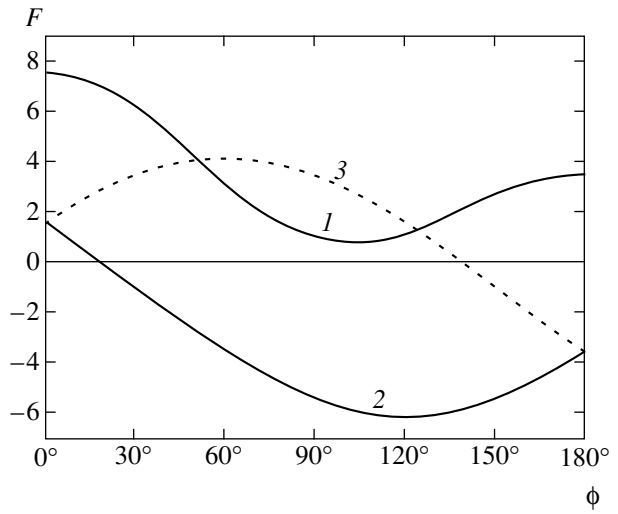

Fig. 3. The dependence of the angular part of exchange integrals (6) along the *b* axis (*1*—F_1) and in the basal plane (*2*—F_2, *3*—F_3) on the Jahn–Teller distortion angle (angle of the orbital structure) ϕ.

pairs of interacting ions along the z_p axis and in the basal plane,

$$J^i = \frac{J_0^i \cos^2\Theta_i}{r_i^{10}} F_i(\phi), \tag{5}$$

where

$$\begin{aligned} F_1(\phi) &= 1 + 2\alpha\cos\phi + \beta\cos^2\phi, \\ \Theta_1 &= \Theta_b, \quad r_1 = r_b, \\ F_2(\phi) &= 1 - \frac{\alpha'}{2}(\cos\phi - \sqrt{3}\sin\phi), \\ \Theta_2 &= \Theta_{ac}, \quad r_2 = r_{ac}, \\ F_3(\phi) &= 1 - \frac{\alpha'}{2}(\cos\phi + \sqrt{3}\sin\phi), \\ \Theta_3 &= \Theta_{ac}, \quad r_3 = r_{ac}, \\ F_4(\phi) &= 1, \quad \Theta_4 = \Theta_b', \quad r_4 = r_b', \end{aligned} \tag{6}$$

and J_0^i, α, α', and β are parameters which are dependent on the sort of magnetic ions in a pair and on the sort of intermediate ion (they may be determined from the experimental data), and ϕ is the angle of Jahn–Teller distortion in a pair of ions (for an Mn^{3+}–Mn^{3+} pair) or on a single ion in a pair (Mn^{3+} for an Mn^{3+}–Mn^{4+} pair). In [17], the parameters of the dependence given by Eqs. (6) were determined, namely, $J_0^1 = 1.24 \times 10^3$ meV Å^{10}, $\alpha = 1.0$, $\beta = 4.5$; $J_0^{2,3} = -4.22 \times 10^2$ meV Å^{10}, $\alpha' = 5.1$, $J_0^4 = 0.97 \times 10^2$ meV Å^{10}. The dependence of the exchange on the Jahn–Teller angle is given in Fig. 3. The structural parameters for $R_{0.5}Ca_{0.5}MnO_3$ are borrowed from [24–26] and are given in Table 2, where r_b is the distance between Mn^{3+} and oxygen along the *b* axis, r_b' is the distance between Mn^{4+} and oxygen along the *b* axis, r_{ac} is the mean distance from Mn^{3+} and Mn^{4+} to oxygen in the *ac* plane, Θ_b is the angle of Mn^{3+}–O–Mn^{3+} superexchange coupling along the *b* axis, Θ_b' is the angle of Mn^{4+}–O–Mn^{4+} superexchange coupling along the *b* axis, and Θ_{ac} is the mean angle of superexchange coupling in the *ac* plane.

The exchange interaction within the basal plane is a result of the presence of the orbital and charge structures. The charge structure defines the ferromagnetic pattern of interaction between Mn^{3+}–Mn^{4+} ions, because the orbitally independent part of this interaction is negative and the presence of an orbitally dependent part of interaction makes it possible to vary the sign of interaction, so that a weak antiferromagnetic exchange arises. The special features of the crystal and charge structures (such as the staggered ordering of manganese ions of different valencies and the shift of Mn^{4+} along the *c* axis) form the orbital structure of the Mn^{3+} sublattice [17]. The presence of this set of factors causes ferromagnetic zigzags related by antiferromagnetic exchange: weak exchange within a plane and strong exchange between the planes.

A characteristic feature of the dependence given by Eqs. (6), as in the case of pure manganite [17, 34], is the possible change of magnitude and sign of the exchange parameter in the *ac* plane (the minus sign indicates ferromagnetic interaction), which will cause a change in the magnetic structure with unchanged crystal structure ($P2_1/m$). This distinguishes the Jahn–Teller compounds from other magnetic dielectrics. A possible manifestation of this feature may be the presence of magnetic ordering of the *A* type in bilayered manganite [20, 27].

Single-Ion Anisotropy and Zeeman Interaction

The single-ion anisotropy emerges in the second order of perturbation theory with respect to spin-orbital interaction. The expression for single-ion anisotropy of a trivalent manganese sublattice in local axes of oxygen octahedrons, which depends on the angle of orbital ordering, was derived in [17, 34],

$$H_{an}^{(n)} = D_n S_{nz_l}^2 + E_n (S_{nx_l}^2 - S_{ny_l}^2), \tag{7}$$

$$D_n = 3P\cos\phi_n, \quad E_n = \sqrt{3}P\sin\phi_n. \tag{8}$$

The values of D_n turn out to be positive ($\cos\phi_n < 0$, $P < 0$) and the same for all magnetic ions in a cell, and E_n has different signs and magnitudes for different positions (*a* and *b*) of Mn^{3+} ions.

The quantity P was determined in [34]: $P \approx -0.1$ meV. Also determined in [34] was the orbital dependence of the *g*-tensors of Zeeman interaction of

Table 3. The parameters of superexchange interaction for the charge-ordered phases of $R_{0.5}Ca_{0.5}MnO_3$ manganites (given along with the name of the compound is the reference to the experimental data on which the calculation was based). The designations correspond to formula (14)

R	J_I, meV	J_{II}, meV	J_{III}, meV	J_{IV}, meV	J_V, meV	J_{VI}, meV	J_{VII}, meV
La [24]	–2.33	–2.41	0.58	0.59	2.67	2.59	1.29
Pr [25]	–2.16	–2.50	0.61	0.85	2.16	2.16	1.21
Tb [26]	–2.07	–1.81	1.03	0.86	1.55	1.64	1.21

Mn^{3+} ions in local axes of oxygen coordination. Because, in the case of field dependences of the spectrum, the effect of nonequivalence and anisotropy of the *g*-tensors is insignificant, we assume the *g*-tensors of manganese ions to be isotropic and equal to two. In this case, the Zeeman interaction is described by the Hamiltonian

$$\hat{H}_{Zeem} = 2\mu_B \sum_n (\mathbf{H} \cdot \mathbf{S}_n). \quad (9)$$

4. RESULTS AND DISCUSSION

The Hamiltonian

$$\hat{H} = \sum_{m>n} J_{mn}(\mathbf{S}_m \cdot \mathbf{S}_n) + \sum_{n=2k-1} \hat{H}_{an}^{(n)} + \hat{H}_{Zeem} \quad (10)$$

was used to calculate the magnetic properties of the charge-ordered phase of manganites. We use the angular dependence of superexchange coupling both in the exchange parameter according to Eqs. (6) and in the remaining terms for taking into account the existence of rotations of oxygen octahedrons. For spin variables in the terms of single-ion anisotropy, the transition from the local system of coordinates related to octahedral axes to the generalized system is accomplished using the matrices

$$M(\varphi, \psi) = \begin{pmatrix} \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & 0 & -\frac{1}{\sqrt{2}} \\ 0 & 1 & 0 \end{pmatrix} \times \begin{pmatrix} \cos\psi & 0 & -\sin\psi \\ 0 & 1 & 0 \\ \sin\psi & 0 & \cos\psi \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\varphi & -\sin\varphi \\ 0 & \sin\varphi & \cos\varphi \end{pmatrix}. \quad (11)$$

The correspondence between the angles of rotation and the orbital structure in a primitive magnetic cell is preassigned as follows:

$$\begin{gathered} Mn_1(\varphi_1, \psi_1, \phi_a), Mn_5(\varphi_1, \psi_1, \phi_a), \\ Mn_9(-\varphi_1, \psi_1, \phi_a), Mn_{11}(-\varphi_1, \psi_1, \phi_a), \\ Mn_3(\varphi_2, \psi_2, \phi_b), Mn_7(\varphi_2, \psi_2, \phi_b), \\ Mn_{13}(-\varphi_2, \psi_2, \phi_b), Mn_{15}(-\varphi_2, \psi_2, \phi_b), \\ Mn_2(\varphi_3, \psi_3), Mn_4(\varphi_3, \psi_3), \\ Mn_6(\varphi_3, \psi_3), Mn_8(\varphi_3, \psi_3), \\ Mn_{10}(-\varphi_3, \psi_3), Mn_{12}(-\varphi_3, \psi_3), \\ Mn_{14}(-\varphi_3, \psi_3), Mn_{16}(-\varphi_3, \psi_3). \end{gathered} \quad (12)$$

The values of the angles of rotation of octahedrons are given in Table 2. In accordance with the experimentally observed magnetic structure of the *CE* type defined by the signs of exchange parameters (see Table 3) in these compounds, the magnetic cell must be doubled as compared to the crystal structure along the *c* axis. So, the magnetic cell contains 16 manganese ions, namely, four ions in position *a*, four ions in position *b*, and eight ions in position *f*. After appropriate transformations, Hamiltonian (10) may be rewritten as the magnetic cell energy in terms of the basis vectors of the magnetic structure for 16 sublattices (in the nearest neighbor approximation),

$$\begin{aligned} E = \frac{1}{16} \times \{ & 4J_1[(\mathbf{a}_1 \cdot \mathbf{A}_1) + (\mathbf{c}_1 \cdot \mathbf{C}_1) + (\mathbf{f}_1 \cdot \mathbf{F}) + (\mathbf{g}_1 \cdot \mathbf{G})] \\ & + 4J_{II}[(\mathbf{a}_2 \cdot \mathbf{A}_1) + (\mathbf{c}_2 \cdot \mathbf{C}_1) + (\mathbf{f}_2 \cdot \mathbf{F}) + (\mathbf{g}_2 \cdot \mathbf{G})] \\ & + 4J_{III}[(\mathbf{a}_1 \cdot \mathbf{A}_1) - (\mathbf{c}_1 \cdot \mathbf{C}_1) + (\mathbf{f}_1 \cdot \mathbf{F}) - (\mathbf{g}_1 \cdot \mathbf{G})] \\ & + 4J_{IV}[(\mathbf{a}_2 \cdot \mathbf{A}_1) - (\mathbf{c}_2 \cdot \mathbf{C}_1) + (\mathbf{f}_2 \cdot \mathbf{F}) - (\mathbf{g}_2 \cdot \mathbf{G})] \\ & + 2J_V[-\mathbf{a}_1^2 + \mathbf{c}_1^2 + \mathbf{f}_1^2 - \mathbf{g}_1^2] \\ & + 2J_{VI}[-\mathbf{a}_2^2 + \mathbf{c}_2^2 + \mathbf{f}_2^2 - \mathbf{g}_2^2] \\ & + J_{VII}[-\mathbf{A}_1^2 + \mathbf{A}_2^2 + \mathbf{A}_3^2 + \mathbf{C}_1^2 - \mathbf{C}_2^2 - \mathbf{C}_3^2 + \mathbf{F}^2 - \mathbf{G}^2] \} \end{aligned} \quad (13)$$

$$+\sum_{i=1,2}\{\lambda_i^1(a_{iX}^2+c_{iX}^2+f_{iX}^2+g_{iX}^2)$$
$$+\lambda_i^2(a_{iY}^2+c_{iY}^2+f_{iY}^2+g_{iY}^2)$$
$$+\lambda_i^3(a_{iZ}^2+c_{iZ}^2+f_{iZ}^2+g_{iZ}^2)$$
$$+\lambda_i^4(a_{iX}a_{iZ}+c_{iX}c_{iZ}+f_{iX}f_{iZ}+g_{iX}g_{iZ})$$
$$+\lambda_i^5(a_{iX}f_{iY}+a_{iY}f_{iX}+c_{iX}g_{iY}+c_{iY}g_{iX})$$
$$+\lambda_i^6(a_{iY}f_{iZ}+a_{iZ}f_{iY}+c_{iY}g_{iZ}+c_{iZ}g_{iY})\}$$
$$+2\mu_B(\mathbf{H}\cdot\{\mathbf{f}_1+\mathbf{f}_2+\mathbf{F}\}),$$

where $\mathbf{f}_1 = \mathbf{S}_1 + \mathbf{S}_5 + \mathbf{S}_9 + \mathbf{S}_{11}$, $\mathbf{g}_1 = \mathbf{S}_1 - \mathbf{S}_5 - \mathbf{S}_9 + \mathbf{S}_{11}$, $\mathbf{a}_1 = \mathbf{S}_1 + \mathbf{S}_5 - \mathbf{S}_9 - \mathbf{S}_{11}$, and $\mathbf{c}_1 = \mathbf{S}_1 - \mathbf{S}_5 + \mathbf{S}_9 - \mathbf{S}_{11}$ are the basis vectors of the magnetic structure for position *a*; $\mathbf{f}_2 = \mathbf{S}_3 + \mathbf{S}_7 + \mathbf{S}_{13} + \mathbf{S}_{15}$, $\mathbf{g}_2 = \mathbf{S}_3 - \mathbf{S}_7 - \mathbf{S}_{13} + \mathbf{S}_{15}$, $\mathbf{a}_2 = \mathbf{S}_3 + \mathbf{S}_7 - \mathbf{S}_{13} - \mathbf{S}_{15}$, and $\mathbf{c}_2 = \mathbf{S}_3 - \mathbf{S}_7 + \mathbf{S}_{13} - \mathbf{S}_{15}$ are the basis vectors of the magnetic structure for position *b*; $\mathbf{F} = \mathbf{S}_2 + \mathbf{S}_4 + \mathbf{S}_6 + \mathbf{S}_8 + \mathbf{S}_{10} + \mathbf{S}_{12} + \mathbf{S}_{14} + \mathbf{S}_{16}$, $\mathbf{G} = \mathbf{S}_2 - \mathbf{S}_4 - \mathbf{S}_6 + \mathbf{S}_8 - \mathbf{S}_{10} + \mathbf{S}_{12} + \mathbf{S}_{14} - \mathbf{S}_{16}$, $\mathbf{A}_1 = \mathbf{S}_2 + \mathbf{S}_4 + \mathbf{S}_6 + \mathbf{S}_8 - \mathbf{S}_{10} - \mathbf{S}_{12} - \mathbf{S}_{14} - \mathbf{S}_{16}$, $\mathbf{A}_2 = \mathbf{S}_2 + \mathbf{S}_4 - \mathbf{S}_6 - \mathbf{S}_8 + \mathbf{S}_{10} + \mathbf{S}_{12} - \mathbf{S}_{14} - \mathbf{S}_{16}$, $\mathbf{A}_3 = \mathbf{S}_2 - \mathbf{S}_4 + \mathbf{S}_6 - \mathbf{S}_8 + \mathbf{S}_{10} - \mathbf{S}_{12} + \mathbf{S}_{14} - \mathbf{S}_{16}$, $\mathbf{C}_1 = \mathbf{S}_2 - \mathbf{S}_4 - \mathbf{S}_6 + \mathbf{S}_8 + \mathbf{S}_{10} - \mathbf{S}_{12} - \mathbf{S}_{14} + \mathbf{S}_{16}$, $\mathbf{C}_2 = \mathbf{S}_2 + \mathbf{S}_4 - \mathbf{S}_6 - \mathbf{S}_8 - \mathbf{S}_{10} - \mathbf{S}_{12} + \mathbf{S}_{14} + \mathbf{S}_{16}$, and $\mathbf{C}_3 = \mathbf{S}_2 - \mathbf{S}_4 + \mathbf{S}_6 - \mathbf{S}_8 - \mathbf{S}_{10} + \mathbf{S}_{12} - \mathbf{S}_{14} + \mathbf{S}_{16}$ are the basis vectors of the magnetic structure for position *f*; and the exchange interaction between magnetic sublattices is given by

$$J_{\mathrm{I}} = J_{1\text{–}2} = J_{1\text{–}8} = J_{4\text{–}5} = J_{5\text{–}6} = J_{9\text{–}10} = J_{9\text{–}16} = J_{12\text{–}13} = J_{13\text{–}14},$$

$$J_{\mathrm{II}} = J_{2\text{–}3} = J_{3\text{–}8} = J_{4\text{–}7} = J_{6\text{–}7} = J_{10\text{–}11} = J_{11\text{–}16} = J_{12\text{–}15} = J_{14\text{–}15},$$

$$J_{\mathrm{III}} = J_{1\text{–}4} = J_{1\text{–}6} = J_{2\text{–}5} = J_{5\text{–}8} = J_{9\text{–}12} = J_{9\text{–}14} = J_{10\text{–}13} = J_{13\text{–}16}, \quad (14)$$

$$J_{\mathrm{IV}} = J_{3\text{–}4} = J_{3\text{–}6} = J_{2\text{–}7} = J_{7\text{–}8} = J_{11\text{–}12} = J_{11\text{–}14} = J_{10\text{–}15} = J_{15\text{–}16},$$

$$J_{\mathrm{V}} = J_{1\text{–}9} = J_{5\text{–}13}, \quad J_{\mathrm{VI}} = J_{3\text{–}11} = J_{7\text{–}15},$$

$$J_{\mathrm{VII}} = J_{2\text{–}10} = J_{4\text{–}12} = J_{6\text{–}14} = J_{8\text{–}16},$$

$$\lambda_i^1 = \frac{\sqrt{3}}{4}P\sin\phi_i\sin2\psi_i,$$

$$\lambda_i^2 = \frac{1}{4}P(3\cos\phi_i\cos^2\varphi_i - \sqrt{3}\sin\phi_i\sin^2\varphi_i\sin2\psi_i),$$

$$\lambda_i^3 = \frac{1}{4}P(3\cos\phi_i\cos^2\varphi_i - \sqrt{3}\sin\phi_i\cos^2\varphi_i\sin2\psi_i),$$

$$\lambda_i^4 = \frac{\sqrt{3}}{2}P\sin\phi_i\cos\varphi_i\cos2\psi_i, \quad (15)$$

$$\lambda_i^5 = \frac{\sqrt{3}}{2}P\sin\phi_i\sin\varphi_i\cos2\psi_i,$$

$$\lambda_i^6 = -\frac{1}{4}P\sin2\varphi_i(3\cos\phi_i + \sqrt{3}\sin\phi_i\sin2\psi_i),$$

$$i = 1, 2.$$

For a quantitative description of the properties of $R_{0.5}Ca_{0.5}MnO_3$, we used the values of the above-identified parameters assembled in Tables 3 and 4.

Magnetic Structure in External Magnetic Field

The minimization of the magnetic energy in the model given by Eq. (13) leads to a multisublattice magnetic structure which has a noncollinear form and is classified for all three positions of a manganese ion as (g_{1X}, 0, g_{1Z}), (g_{2X}, 0, g_{2Z}), (G_X, 0, G_Z) (see Figs. 4a and 4c) disregarding rotational distortions and as (g_{1X}, c_{1Y}, g_{1Z}), (g_{2X}, c_{2Y}, g_{2Z}), (G_X, C_{1Y}, G_Z) (see Figs. 4a and 4d) with regard for rotations, which corresponds to the representation Γ_1 $\mathbf{k}_{12}$ (*a*, *b*) and Γ_2 $\mathbf{k}_{12}$ (*f*) of the $P2_1/m$ group. The approximate direction of the easy magnetic axis of the antiferromagnetic structure is slightly deflected from the *c* axis of the crystal. In contrast to pure manganite, this structure does not exhibit weak ferromagnetism, which makes experimental determination of its details difficult. The structure of the entire magnetic cell approximately corresponds to the CE_Z type, which is confirmed by the experimental results of [24, 25, 31].

Expression (13) enables one to assume the basic components of the magnetic structure. On comparing the coefficients of the squares of components of the basis vectors of the magnetic structure, one can see that the most negative coefficients are observed for the squares of components a_{1X}^2, g_{1X}^2 ($-J_{\mathrm{V}}/8 + \lambda_1^1$) and a_{1Z}^2, g_{1Z}^2 ($-J_{\mathrm{V}}/8 + \lambda_1^3$), a_{2X}^2, g_{2X}^2 ($-J_{\mathrm{VI}}/8 + \lambda_2^1$) and a_{2Z}^2, g_{2Z}^2 ($-J_{\mathrm{VI}}/8 + \lambda_2^3$) of trivalent manganese sublattices. In addition, a_{1X} and a_{2X}, g_{1X} and g_{2X}, a_{1Z} and a_{2Z}, g_{1Z} and g_{2Z} are interrelated via exchange interaction with Mn^{4+} ions, and a_{1X} and a_{1Z}, g_{1X} and g_{1Z}, a_{2X} and a_{2Z}, g_{2X} and g_{2Z} are related in the anisotropy terms with the coefficients $\lambda_{1,2}^1$. It follows from Table 4 and formula (13)

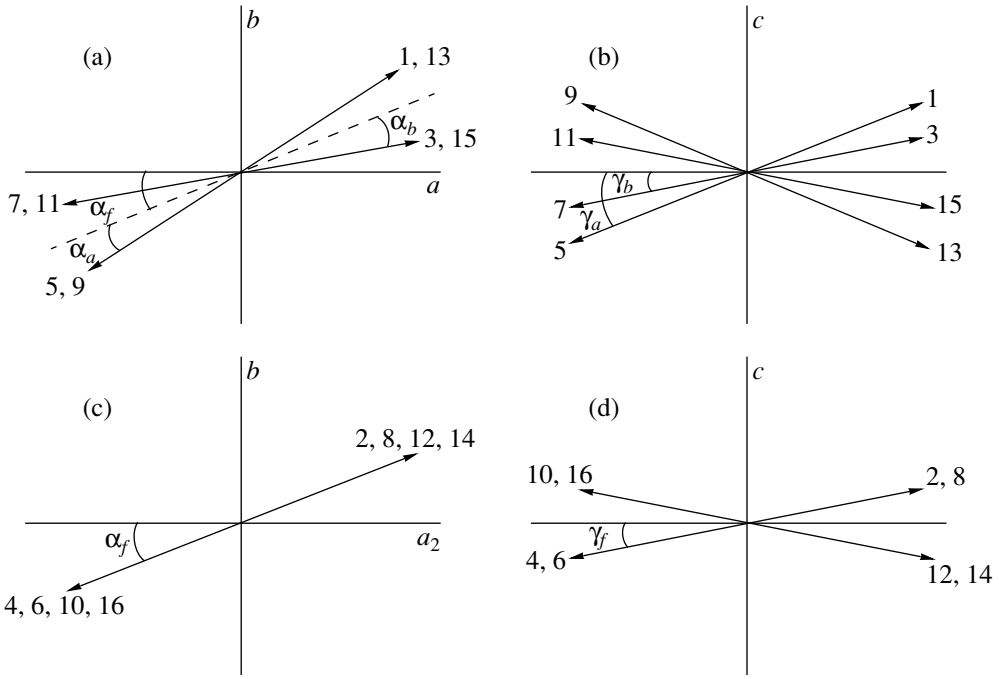

Fig. 4. The magnetic structure of a monoclinic $R_{0.5}Ca_{0.5}MnO_3$ crystal: (a) Mn_a^{3+} and Mn_b^{3+} sublattices, components g_{1X}, g_{1Z}, g_{2X}, g_{2Z}; (b) Mn_a^{3+} and Mn_b^{3+} sublattices, components c_{1Y}, c_{2Y}; (c) Mn_f^{4+} sublattice, components G_X, G_Z; (d) Mn_f^{4+} sublattice, component C_{1Y}.

that the predominance of any of these structures is defined by the exchange interactions with the Mn^{4+} sublattice. Because the negative components of positions *a* and *b*, which are of greatest magnitude, single out different projections of basis vectors (*X* and *Z*, respectively) and $|\lambda_1^1| < |\lambda_2^3|$, one can assume that the *Z* axis will approximately coincide with the easy magnetic axis of the magnetic structure as a whole owing to exchange interaction. The same exchange (contributions to Eq. (13) with the coefficients J_I, $J_{II} < 0$ and J_{III}, $J_{IV} > 0$) defines the basic components of the magnetic structure g_{1Z}, g_{2Z}, G_Z.

One can readily see that, in the absence of rotational distortions in a crystal, the coefficients in expression (13) are $\lambda_{1,2}^1 = \lambda_{1,2}^3 = \lambda_{1,2}^5 = \lambda_{1,2}^6 = 0$, from which it follows that, in this case, structures with the components g_{1X}, g_{2X}, G_X and g_{1Z}, g_{2Z}, G_Z become equivalent.

Table 4. The parameters of single-ion anisotropy for the charge-ordered phases of $R_{0.5}Ca_{0.5}MnO_3$ manganites (given along with the name of the compound is the reference to the experimental data on which the calculation was based). The designations correspond to formula (15)

R	La [24]		Pr [25]		Tb [26]	
i	1	2	1	2	1	2
λ_i^1, meV	0.006	−0.004	0.009	−0.006	−0.012	0.016
λ_i^2, meV	0.042	0.043	0.040	0.037	0.026	0.027
λ_i^3, meV	−0.005	0.005	−0.009	0.007	0.013	−0.014
λ_i^4, meV	0.067	−0.068	0.085	−0.072	0.075	−0.071
λ_i^5, meV	−0.011	0.011	−0.012	0.013	0.016	−0.015
λ_i^6, meV	0.015	0.012	0.017	0.011	−0.006	−0.017

Table 5. The parameters of the magnetic structure of $R_{0.5}Ca_{0.5}MnO_3$ (given along with the name of the compound is the reference to the experimental data on which the calculation was based)

R	α_a, deg	α_b, deg	α_f, deg	γ_a, deg	γ_b, deg	γ_f, deg	α^{exp}, deg
La [24]	–3.57°	3.42°	–4.62°	–0.28°	–0.25°	–0.16°	20°
Pr [25]	–3.87°	3.27°	–9.91°	0.32°	0.29°	0.19°	35°
Tb [26]	–3.73°	4.45°	12.05°	–0.29°	–0.21°	–0.4°	–

Owing to the presence of the orbital structure, the easy magnetic axis of the crystal is singled out and noncollinear components of the magnetic structure emerge. The presence of the $c_{1,2Y}$- and C_{1Y}-components is likewise caused by the inclusion of rotational distortions of the crystal. Disregarding the rotations, one can write for the respective components of the magnetic structure

$$\alpha_a = \frac{S_1\lambda_1^4\sin(2\alpha_f)}{S_2(J_I - J_{III}) + S_1\lambda_1^4\cos(2\alpha_f)},$$
$$\alpha_b = \frac{S_1\lambda_2^4\sin(2\alpha_f)}{S_2(J_{II} - J_{IV}) + S_1\lambda_2^4\cos(2\alpha_f)}, \qquad (16)$$
$$\cos(2\alpha_f) = \frac{S_1(\lambda_1^4 + \lambda_2^4)(J_I - J_{III})(J_{II} - J_{IV})}{S_2\lambda_1^4\lambda_2^4(J_I - J_{III} + J_{II} - J_{IV})},$$

where $S_1 = 2$ is the Mn^{3+} spin and $S_2 = 3/2$ is the Mn^{4+} spin.

When the rotations are taken into account, the calculation is appreciably complicated; therefore, it was performed in the numerical form. The values of the respective noncollinearity angles of the magnetic moments of manganese ions are given in Table 5. According to the experimental data of [24, 25], the magnetic structure in monoclinic charge-ordered manganites has an easy magnetic axis deflected from the *c* axis, with the direction of the latter axis being attributed to the presence of a minor disproportion of the crystal structure that is disregarded in our model. In our model, such a deflection is qualitatively explicable. Similarly to the case of pure and charge-ordered orthorhombic manganite [17], the direction of the easy magnetic axis along the pseudoperovskite diagonal (*a* or *c* axis) is defined by the presence of magnetic anisotropy with the varying sign of the coefficient E_n. In the case of monoclinic structure, the local easy magnetic axis for position *a* is the y_p axis, and for position *b*, the x_p axis. However, the coefficients E_n for these positions have different signs and differ in magnitude as well. As a result, the direction of the easy magnetic axis slightly differs from diagonal. The value of the *X* component of the magnetic structure in our model is on the average less than the experimentally observed value (cf., α_f and α^{exp} in Table 5). Therefore, the orbital structure only partly defines this feature. Depending on the signs of exchange interaction, a zigzag orbital structure in combination with rotational distortions may cause the presence or absence of weak ferromagnetism in a crystal. The $CE_{X,Z}$ structure ($\Gamma_1\,\mathbf{k}_{12}\,(a, b)\,\Gamma_2\,\mathbf{k}_{12}\,(f)$ of the $P2_1/m$ group) corresponds to a fully antiferromagnetic structure, and the $A_{X,Z}$ structure ($\Gamma_1\mathbf{k}_7\,(a, b, f)$ of the $P2_1/m$ group) exhibits a weak ferromagnetic moment similarly to pure manganite. The transition between these structures may be accomplished by varying the orbital angle ϕ to a value exceeding 140°. Such a transition may be accomplished, for example, with the aid of external pressure. Therefore, there exists a possibility of producing weak ferromagnetism in a crystal with the aid of nonmagnetic stimulation.

A characteristic feature in the formation of the magnetic structure of charge-ordered manganites is the presence of an Mn^{4+} sublattice whose magnetic anisotropy is low and is disregarded in our model [17]. Nevertheless, owing to the strong single-ion anisotropy of the neighboring Mn^{3+} ions and to the exchange interaction, the sublattice of tetravalent manganese ions also has the easy magnetic axis directed approximately to the middle between the easy magnetic axes of Mn^{3+} of position *a* and Mn^{3+} of position *b*.

The behavior of the magnetic structure of an antiferromagnet with the easy magnetic axis in an external magnetic field is usually treated in accordance with a simple model with one anisotropic contribution describing the easy magnetic axis (in our case, this is $D(g_{1Z}^2 + g_{2Z}^2 + G_Z^2)$). In this case, with the field directed along the easy magnetic axis, the magnetic structure remains unchanged up to some value of the field (H_{c1}), after which the magnetic sublattices "spin flop" (spin-flop transition) normally to the external field direction. After that, when the external field is strengthened, the direction of magnetic moments gradually approaches the field direction until saturation occurs (spin-flip transition at H_{c2}). With this approach, the behavior of the noncollinear components is not treated.

In our model, the behavior of the magnetic structure is more complicated. This is associated, first of all, with the inclusion of the noncollinear components of the magnetic structure. When the magnetic field increases along the *Z* axis, the components (f_{1X}, a_{1Y}, f_{1Z}), (f_{2X}, a_{2Y}, f_{2Z}),

and $(F_X, A_{1Y}, F_Z) - \Gamma_3\,\mathbf{k}_7$ (a, b, f) interacting with f_{1Z}, f_{2Z}, and F_Z are added to the available components of the magnetic structure; this causes complication of the behavior of the magnetic subsystem in the external magnetic field. One can readily see that any direction of the magnetic field within the *ac* plane leads to the emergence of the same additional components of the magnetic structure. From the standpoint of symmetry, the spin-flop transition, which is observed in these compounds as well [31], fails to lead to (in distinction to pure manganites [34]) a qualitative variation of the magnetic structure: after the critical value of external magnetic field is reached, the ferromagnetic component of the structure emerges along with the change of the approximate easy magnetic axis from *Z* to *X*. A further increase in the field is accompanied by an increase in the $\Gamma_3\,\mathbf{k}_7$-component of the structure and a decrease in the respective components of the initial structure $(\Gamma_1\,\mathbf{k}_{12}$ (a, b) $\Gamma_2\,\mathbf{k}_{12}$ $(f))$. Therefore, it is obvious that, similarly to the case of pure manganite [34], no complete saturation may be attained in this system.

The abundance of additional components of the magnetic structure, which are important from the standpoint of understanding the effect of external magnetic field, significantly hampers analytical calculations. Therefore, we will give below the results of numerical calculations (see Fig. 5). The values of the critical field H_{c1} amounted to 4.2 T for $La_{0.5}Ca_{0.5}MnO_3$, 6.4 T for $Pr_{0.5}Ca_{0.5}MnO_3$, and 5.8 T for $Tb_{0.5}Ca_{0.5}MnO_3$. The value of the field H_{c1} for $Pr_{0.5}Ca_{0.5}MnO_{3-\delta}$ agrees with the value experimentally found for $Pr_{0.5}Ca_{0.5}MnO_{3-\delta}$ in [31] and equal to approximately 2–5 T. No other critical fields were measured. The field of spin-flop transition for charge-ordered manganites has a value which is an order of magnitude less than that for pure manganite (21 T [38], experiment; 19 T [34], theory).

Of great interest, in addition to the behavior of the magnetic structure in an external magnetic field, is the possibility of destruction of charge ordering by an other-than-*CE* magnetic order [20] or by an external magnetic field [3, 31, 39–42]. Different values are cited for the external magnetic fields at which the charge-ordered phase is destroyed. For example, for $Pr_{0.5}Ca_{0.5}MnO_{3-\delta}$, this value is approximately 5–8 T (at $T \sim 80$ K) [31]; for $Pr_{0.5}Sr_{0.5}MnO_3$, 5 or 6 T (at $T \approx 0$) [39]; for various compositions of $Pr_{0.5}Ca_{0.5-x}Sr_xMnO_3$ [40], the transition is assumed to occur at fields of approximately 20 to 30 T ($T < T_N$) for $x = 0$, and no transition was experimentally observed with fields of up to 5 T even in samples with a high strontium content; Respaud *et al.* [41], who investigated $R_{0.5}Ca_{0.5}MnO_3$ compounds (R = La, Pr, Nd, Sm), observed the destruction of the charge-ordered phase by fields of 0–15 T for $La_{0.5}Ca_{0.5}MnO_3$, of 16–25 T for $Pr_{0.5}Ca_{0.5}MnO_3$, of 5–23 T for $Nd_{0.5}Ca_{0.5}MnO_3$, and of 23–26 T for $Sm_{0.5}Ca_{0.5}MnO_3$ at low temperatures (the lower value corresponds to transition with decreasing field, and the higher value, to that with increasing field); Lope *et al.* [42] report, for $La_{0.5}Ca_{0.5}MnO_3$, fields of 0.4–3 T for which a gradual destruction of the antiferromagnetic phase begins at different temperatures. In spite of the different values of the fields at which the charge-ordered phase is destroyed, note that these fields are close, by the order of magnitude, to fields of spin-flop transition. In this way, is it possible to talk about the field dependences of the magnetic structure and of the spin-flop transition proper while treating the crystal as being charge-ordered? Such an approximation is possible for several reasons. First, we assume the presence of an ideal single crystal without twins and domains. In [40–42], polycrystals were investigated. Second, as is reported in [25, 31], at fairly low temperatures the charge-ordered phase is stable; therefore, we can restrict ourselves to a low-temperature approximation. One can assume that no destruction of the charge-ordered phase will occur in the absence of the ferromagnetic component of the magnetic structure and that the mechanism of double exchange will be suppressed. In the presence of even minor ferromagnetism, double exchange causes the emergence in the crystal of regions with delocalized charge carriers, i.e., with deteriorating charge order. As was stated before, the formation of the *A* structure with unchanged symmetry of the crystal lattice is accompanied by the emergence of a weak ferromagnetic component initiating the destruction of the charge-ordered phase. When an external magnetic field is applied, an appreciable ferromagnetic component of the structure may emerge right away (**H** is directed normally to the easy magnetic axis) or at $H > H_{c1}$ (**H** is directed parallel to the easy magnetic axis, see Fig. 5). Other directions represent an intermediate case of transition. In a polycrystal sample, the field of spin-flop transition is the maximal field at which the ferromagnetic moment arises in all grains of the sample. Therefore, one can assume that the spin-flop transition may start the destruction of, or fully destroy, the charge ordering; therefore, the behavior of a charge-ordered

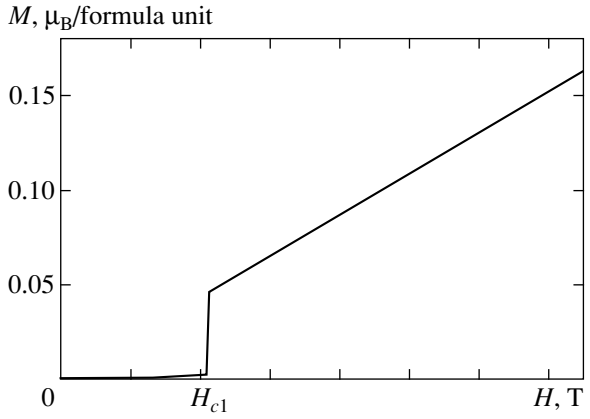

Fig. 5. The total magnetization (per formula unit) *M* as a function of the external magnetic field *H*.

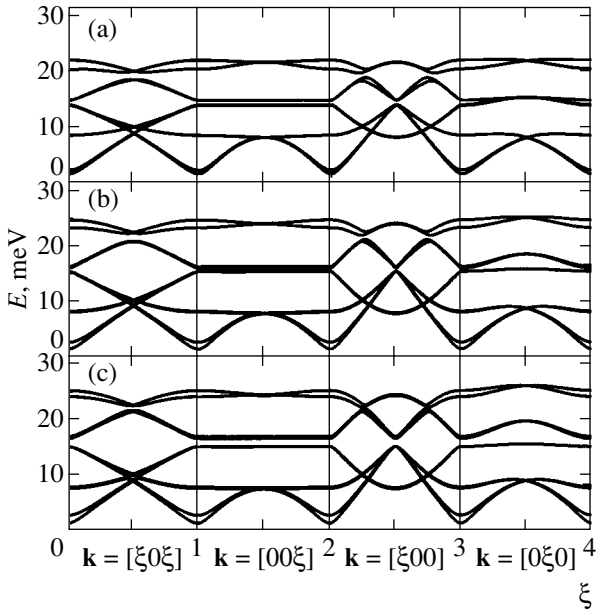

Fig. 6. The dispersion dependences of the energy E of $R_{0.5}Ca_{0.5}MnO_3$ magnons for different directions of the magnetic Brillouin zone on the wave vector ξ: (a) corresponds to $Tb_{0.5}Ca_{0.5}MnO_3$, (b) to $Pr_{0.5}Ca_{0.5}MnO_3$, and (c) to $La_{0.5}Ca_{0.5}MnO_3$.

system in an external magnetic field may be predicted if the field is directed along the easy magnetic axis and little exceeds the field of spin-flop transition.

Spin Waves and Antiferromagnetic Resonance

The dispersion dependences of magnons (Fig. 6) were investigated for the energy given by Eqs. (12) in a linear approximation of spin waves. Because of the 16-sublattice model of the magnetic structure, the spin-wave spectrum has sixteen branches. The spectrum is divided into six zones with several adjacent branches in each zone (two in the first, second, fifth, and sixth zones, and four in the third and fourth zones). A characteristic feature of the spectral dispersion is the marked difference between the curves in the [ξ00] and [00ξ] directions, i.e., along the orbital zigzag and in the perpendicular direction within the *ac* plane. Note that this difference is due both to the strong ferromagnetic interaction within the zigzag and weak antiferromagnetic coupling between zigzags (i.e., magnetic subsystem) and to the relative position of the orbital, charge, and magnetic sublattices. A reduction in the number of sublattices to six (by the scheme of 1' = 1, 13; 2' = 2, 8, 12, 14; 3' = 3, 15; 4' = 4, 6, 10, 16; 5' = 5, 9; 6' = 7, 11; 1'–6' are "new sublattices", and 1–16 are sublattices in the designations of Fig. 1) describes the dispersion dependences in the [ξξ0] and [0ξ0] directions; however, in two other directions being investigated, this model leads to fully identical dispersion dependences; i.e., the effect of the orbital structure on the spin-wave spectrum is lost. In addition, the six-sublattice model presumes disregard of rotational distortions in the single-ion anisotropy, which causes the singling out of the easy magnetic axis of the crystal; i.e., it is incapable of describing a spin-flop transition.

At $k = 0$, the energy spectrum of magnetic excitations may be observed with the aid of AFMR (see Fig. 7). In the absence of an external magnetic field, the lower branches of the spectrum are split and separated by an energy gap. The formation of such a gap is characteristic of an easy-axis antiferromagnet. The removal of degeneracy of the lower spectral branches is caused, as in the case of pure manganite [34], by the presence of orbitally dependent single-ion anisotropy, which is characterized by the alteration of the easy magnetic axis within the *ac* plane. The increase in splitting between degenerate branches is a result of the presence of the sublattice of Mn^{4+} ions which, in our model, do not exhibit magnetic anisotropy.

According to our calculation results, the values of the gap ΔE and splitting δE amounted to ΔE = 1.98 meV and δE = 1.43 meV for $La_{0.5}Ca_{0.5}MnO_3$, to ΔE = 1.81 meV and δE = 0.68 meV for $Pr_{0.5}Ca_{0.5}MnO_3$, and to ΔE = 2.00 meV and δE = 1.18 meV for $Tb_{0.5}Ca_{0.5}MnO_3$. The values for $Pr_{0.5}Ca_{0.5}MnO_3$ qualitatively agree with those experimentally found in [31]: ΔE = 0.64 meV and δE = 0.12 meV. The disagreement between our calculation results and experimental data may be attributed to the presence of oxygen vacancies in the investigated sample; this, according to Hirota *et al.* [38], is the reason for the presence of a minor amount of free charge carriers. The charge carriers bring about double exchange, which, even in the case of low concentration, may cause marked changes in AFMR spectra, for example, as in the case of strontium-doped manganite [43].

In our opinion, the main criterion for qualitative comparison with the given experiment may be provided by the form of field dependence of the spectrum in the direction of the easy magnetization axis (~*c*). The behavior of the field dependence of the frequencies is characteristic of an easy-axis antiferromagnet. When the field increases in the direction of easy magnetization axis, the two lower branches of the spectrum diverge still further. At $H = H_{c1}$, a small jump is observed associated with spin-flop transition. In regular antiferromagnets, given this field, the lower branch has zero energy. In our model, only a small jump of the resonance frequency is expected in the region of fields of spin-flop transition. The subsequent behavior of the field dependence is similar to that of the dependence in the case of the field **H** directed normally to the easy magnetic axis (~*a*), which is attributed to the weakening of the effect of this axis on the spectrum due to the interaction between the effects of the orbital and charge structures, namely, due to the presence of alternation of the sublattices of trivalent manganese ions, which exhibit a strong orbitally dependent single-ion anisot-

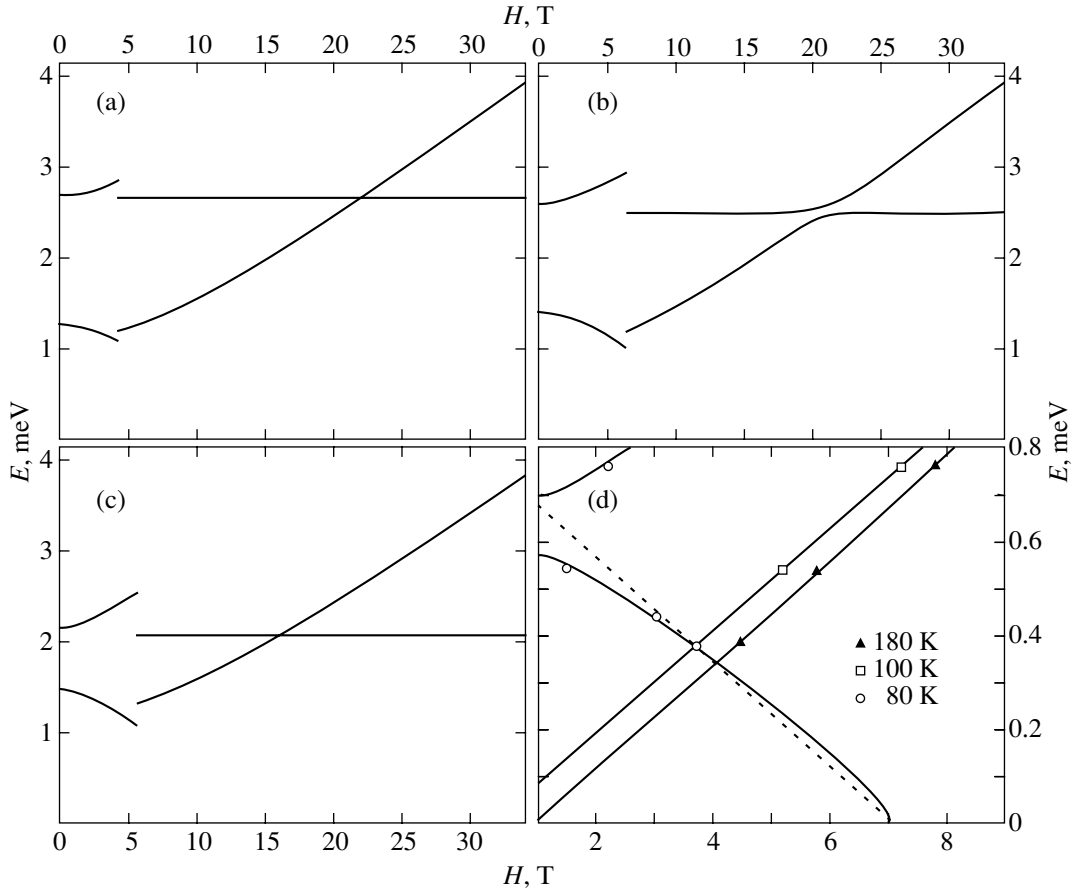

Fig. 7. The AFMR energy E as a function of the external magnetic field H directed parallel to the easy magnetization axis for $R_{0.5}Ca_{0.5}MnO_3$: (a) $La_{0.5}Ca_{0.5}MnO_3$; (b) $Pr_{0.5}Ca_{0.5}MnO_3$; (c) $Tb_{0.5}Ca_{0.5}MnO_3$; and (d) experimental data for $Pr_{0.5}Ca_{0.5}MnO_{3-\delta}$ [31].

ropy, and the sublattices of tetravalent manganese ions, whose magnetic anisotropy may be ignored and the distortion of whose oxygen coordination has no effect on the formation of the easy magnetic axis of the crystal. The magnetic structures with the approximate easy axis in the vicinity of the *a* axis and in the vicinity of the *c* axis differ with respect to energy much less than in the case of pure manganite [34]; therefore, the AFMR frequency jump becomes weaker. A similar behavior was experimentally observed [31] within the range of measurements; however, the field dependence above the field of spin-flop transition is extrapolated by a straight line tending to zero in the low-field region.

5. CONCLUSION

This paper deals with a simple model describing the dispersion and field dependences [31] of spin-wave energy, as well as the field dependence of total magnetization. The model includes orbitally dependent interactions, namely: the exchange and single-ion anisotropy; Zeeman interaction; and orbital and charge structures that are assumed to be fixed and independent of magnetic interactions. No anisotropic exchange is present in the model, and all rotational distortions are taken into account in other interactions.

Because of the simplicity of the model, our results clearly demonstrate the dependence of the magnetic structure and spectra of magnons on the orbital, charge, and crystalline ordering. For the same reason, one can qualitatively explain both the formation of the magnetic structure and its behavior in an external magnetic field.

ACKNOWLEDGMENTS

This study was supported in part by the CRDF program (REC-005), by the Ministry of Education of the Russian Federation (grant no. E00-3.4-277), and by the Russian Foundation for Basic Research (project 02-02-96412). We are grateful to Z. Jirak (Czechia) for placing at our disposal exact experimental data reported [25] at the ICNS2001 Conference (September 13–16, 2001, Munich, Germany).

REFERENCES

1. J. B. Goodenough, Phys. Rev. **100**, 564 (1955).
2. E. O. Wollan and W. C. Koehler, Phys. Rev. **100**, 545 (1955).

3. M. Imada, A. Fujimori, and Y. Tokura, Rev. Mod. Phys. **70**, 1039 (1998); E. Dagotto, T. Hotta, and A. Moreo, Phys. Rep. **344**, 1 (2001).
4. K. I. Kugel' and D. I. Khomskii, Usp. Fiz. Nauk **136**, 621 (1982) [Sov. Phys. Usp. **25**, 231 (1982)].
5. I. Solovyev, N. Hamada, and K. Terakura, Phys. Rev. Lett. **76**, 4825 (1996).
6. G. Khaliullin and V. Oudovenko, Phys. Rev. B **56**, R14243 (1997).
7. I. V. Solovyev and K. Terakura, Phys. Rev. Lett. **83**, 2825 (1999).
8. J. Van den Brink, P. Horsch, F. Mack, and A. M. Oles, Phys. Rev. B **59**, 6795 (1999).
9. J. Van den Brink, G. Khaliullin, and D. Khomskii, Phys. Rev. Lett. **83**, 5118 (1999).
10. E. Saitoh *et al.*, Nature **410**, 180 (2001).
11. D. L. Huber, G. Alejandro, A. Caneiro, *et al.*, Phys. Rev. B **60**, 12155 (1999).
12. M. V. Zimmermann, J. P. Hill, D. Gibbs, *et al.*, Phys. Rev. Lett. **83**, 4872 (1999); H. Kawata, T. Arima, Y. Tokura, *et al.*, Phys. Rev. Lett. **81**, 582 (1999).
13. J. Kanamori, J. Appl. Phys. (Suppl.) **31**, 14S (1960).
14. A. J. Millis, Phys. Rev. B **53**, 8434 (1996).
15. L. F. Feiner and A. M. Oles, Phys. Rev. B **59**, 3295 (1999); S. Fratini, M. Capone, M. Grilli, and D. Feinberg, in *Physics in Local Lattice Distortions*, Ed. by H. Oyanagi and A. Bianconi (Am. Inst. Phys., Melville, New York, 2001); AIP Conf. Proc. **554**, 371 (2000).
16. T. Hotta, S. Yunoki, M. Mayr, and E. Dagotto, Phys. Rev. B **60**, R15009 (1999); T. Hotta, Y. Takada, H. Koizumi, and E. Dagotto, Phys. Rev. Lett. **84**, 2477 (2000); S. Yunoki, T. Hotta, and E. Dagotto, Phys. Rev. Lett. **84**, 3714 (2000).
17. L. E. Gontchar, A. E. Nikiforov, and S. E. Popov, J. Magn. Magn. Mater. **223**, 175 (2001).
18. F. Moussa, M. Hennion, J. Rodriguez-Carvajal, *et al.*, Phys. Rev. B **54**, 15149 (1996).
19. K. Hirota, N. Kaneko, A. Nishizawa, and Y. Endoh, J. Phys. Soc. Jpn. **65**, 3736 (1996).
20. J. Q. Li, Y. Matsui, T. Kimura, and Y. Tokura, Phys. Rev. B **57**, R3205 (1998).
21. R. Maezono, S. Ishihara, and N. Nagaosa, Phys. Rev. B **58**, 11583 (1998); R. Maezono and N. Nagaosa, Phys. Rev. B **62**, 11576 (2000).
22. T. Mizokawa, D. I. Khomskii, and G. A. Sawatzki, Phys. Rev. B **61**, R3776 (2000); T. Mizokawa, D. I. Khomskii, and G. A. Sawatzki, Phys. Rev. B **63**, 024403 (2001).
23. M. Van Veendaal and A. J. Fedro, Phys. Rev. B **59**, 1285 (1999).
24. P. G. Radaelli, D. E. Cox, M. Marezio, and S.-W. Cheong, Phys. Rev. B **55**, 3015 (1997).
25. Z. Jirák, C. Martin, M. Hervieu, and J. Hejtmanek, in *Reported at International Conference on Neutron Scattering ICNS2001, Munich* (see abstract in *Conference Programme and Abstracts ICNS2001*, Ed. by J. Neuhaus and A. Meyer (2001), p. 146).
26. J. Blasco, J. Garcia, J. M. de Teresa, *et al.*, J. Phys.: Condens. Matter **9**, 10321 (1997).
27. L. E. Gontchar and A. E. Nikiforov, Czech. J. Phys. (Suppl. A) **52**, A245 (2002).
28. Z. Jirák, S. Krupička, Z. Šimša, *et al.*, J. Magn. Magn. Mater. **53**, 153 (1985).
29. F. Damay, Z. Jirák, M. Hervieu, *et al.*, J. Magn. Magn. Mater. **190**, 221 (1998).
30. A. Machida, Y. Moritomo, K. Ohoyama, *et al.*, Phys. Rev. B **62**, 80 (2000).
31. S. Kawamata, S. Noguchi, K. Okuda, *et al.*, J. Magn. Magn. Mater. **226–230**, 854 (2001).
32. V. E. Naĭsh, Fiz. Met. Metalloved. **92**, 5 (2001).
33. V. E. Naĭsh, Fiz. Met. Metalloved. **92**, 16 (2001).
34. L. É. Gonchar', A. E. Nikiforov, and S. É. Popov, Zh. Éksp. Teor. Fiz. **118**, 1411 (2000) [JETP **91**, 1221 (2000)].
35. A. E. Nikiforov, S. É. Popov, and S. Yu. Shashkin, Fiz. Met. Metalloved. **87**, 16 (1999).
36. O. V. Kovalev, *Representations of the Crystallographic Space Groups: Irreducible Representations, Induced Representations, and Corepresentations* (Nauka, Moscow, 1986; Gordon and Breach, Yverdon, Switzerland, 1993).
37. K. S. Aleksandrov, A. T. Anistratov, B. V. Beznosikov, and N. V. Fedoseeva, *Phase Transitions in Crystals of Halogenide Compounds* (Nauka, Novosibirsk, 1981).
38. K. Hirota, A. Nishizawa, and Y. Endoh, J. Magn. Magn. Mater. **177–181**, 864 (1998).
39. Y. Tomioka, A. Asamitsu, Y. Morimoto, *et al.*, Phys. Rev. Lett. **74**, 5108 (1995).
40. S. Krupička, M. Maryško, Z. Jirák, and J. Hejtmanek, J. Magn. Magn. Mater. **206**, 45 (1999).
41. M. Respaud, J. M. Broto, H. Rakoto, *et al.*, J. Magn. Magn. Mater. **211**, 128 (2000).
42. J. Lope, P. N. Lisboa-Filho, W. A. C. Passos, *et al.*, J. Magn. Magn. Mater. **226–230**, 500 (2001).
43. A. A. Mukhin, V. Yu. Ivanov, V. D. Travkin, *et al.*, Physica B (Amsterdam) **284–288**, 1414 (2000).

Translated by H. Bronstein

Journal of Experimental and Theoretical Physics, Vol. 96, No. 3, 2003, pp. 523–530.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 123, No. 3, 2003, pp. 590–598.

SOLIDS
Structure

Behavior of Dislocations in Silicon in the Presence of Mechanical and Magnetic Perturbations

A. M. Orlov*, A. A. Skvortsov, and A. A. Solov'ev**
Ul'yanovsk State University, Ul'yanovsk, 432700 Russia
**e-mail: orlovam@ulsu.ru*
***e-mail: scvor@sv.uven.ru*
Received September 6, 2002

Abstract—Various types of dislocation stoppers are identified and their basic parameters are determined. Using dislocation loops as an example, the effect of internal stresses on the motion of linear defects in *n*- and *p*-Si in the field of external elastic forces is estimated. It is found that preliminary magnetic treatment of silicon plates activates the dislocation transport. In the absence of external mechanical loads, displacement of dislocation half-loops (30–50 μm) in the nonuniform field of internal stresses in a silicon crystal with a scratch (stress concentrator) is detected experimentally during isothermal annealing for 0.5–3 h at a temperature of 600–700°C. Dislocation transport is described taking into account the intrinsic (lattice) potential barrier of the crystal and two types of stoppers on the basis of magnetosensitive point defects (dopant) and "forest" dislocations. A kinetic model is proposed for describing the magnetostimulated variation of the mobility of linear defects associated with the formation of long-lived complexes with a paramagnetic impurity. It is found experimentally that the velocity of dislocations in *n*- and *p*-Si increases by a factor of 2 and 3, respectively, upon treatment of the semiconductor in a magnetic field $B = 1$ T for 5–45 min. The "magnetic memory" effect in silicon containing dislocations is detected and kinetic aspects of the effect under natural conditions of sample storage after the removal of the magnetic field are considered. Partial velocities of dislocations and their delay times at various types of stoppers are calculated from the matching of experiment with theory. © *2003 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

It is well known that a weak magnetic field considerably affects the state of linear defects in ionic crystals, metals, and semiconductors [1–13]. It is generally assumed that these effects are mainly due to magnetostimulated evolution of electron spins of paramagnetic impurity centers [12], which changes the mobility of linear defects in the field of internal stresses as well as in the field of external forces.

Available information on dislocation transport in elementary semiconductors [8–10] does not cover a number of important aspects of this phenomenon. For example, there is no information on the division of the role of various types of stoppers interacting with a moving dislocation loop in a doped semiconductor. The field and kinetic dependences of dislocation path lengths in silicon subjected to magnetic field treatment have not been described. The data on the kinetics of dislocation path kinetics in crystals with a nonuniform spatial distribution of linear defects in the presence of internal stresses alone and in the field of external forces are scarce. This study is devoted to analysis of such problems.

We performed experiments with silicon plates doped with phosphorus or boron ($\rho = 2$ or 0.5 Ω cm) with the [111] and [100] crystallographic orientations (Table 1). Dislocation dynamics in regions with a nonuniform dislocation density was studied by scribing the surface of the semiconductor with a diamond prism under a fixed load $P = 0.8$ N according to the technique described

Table 1. Parameters of silicon samples

Conductivity type; impurity	Crystallographic orientation	ρ, Ω cm	Size (mm) and directions of scribing	N_d, cm^{-2}	Deformation around crystallographic axis	σ_{ext}, MPa
n, phosphorus	[111]	2	$30 \times 10 \times 0.4$ [$\bar{1}$ 00], [11$\bar{2}$], [111]	10^4–10^6	[$\bar{1}$ 10]	30
p, boron	[100]	0.5	$25 \times 10 \times 0.6$ [100], [010], [001]	10^4–10^6	[001]	50

1063-7761/03/9603-0523$24.00

in [9, 10]. Displacement of dislocations was caused by internal or external tensile stresses (20–100 MPa). The latter version was realized through four-point bending [3–6] at temperatures of 723–973 K. The paths of dislocation segments were analyzed using the selective etching method [9, 10].

The magnetic field treatment was carried out between the poles of an electromagnet ($B \le 1$ T); the magnetic induction vector was in the plane of the plate and was always perpendicular to inscribed scratches. The magnetic treatment time t_B varied in the range 30–2800 s for a fixed time shift $t^* = 180$ s between the termination of magnetic action and the beginning of high-temperature plastic deformation. The fixation of time t^* was dictated by relaxation processes involving magnetosensitive stoppers after the magnetic treatment of silicon samples.

2. THERMOMECHANICAL SIMULATION OF DISLOCATION

After scratching, a crystal always acquires local stresses σ_{int}, whose high-temperature relaxation in the absence of external mechanical stresses σ_{ext} can lead to generation and subsequent migration of a linear defect. The behavior of such defects must be completely controlled by the stress σ_{int} that redistributes dislocations depending on the spatial distribution of mechanical perturbations in the field of a stress concentrator.

In order to construct the profile of stress σ_{int}, we assume that internal stresses are caused by linear dislocations of the same type, which are arranged in parallel in the vicinity of a scratch. Let us assume that the origin of coordinates is at the point of intersection of the slip plane and the edge of the scratch, the z axis is parallel to dislocations, and the xz plane coincides with the slip plane. For the sake of definiteness, we consider that the Burgers vectors of the dislocations are directed along the x axis. Then the force acting per unit length of a dislocation in the slip plane is equal to $b\sigma_{xy}$, where σ_{xy} is the stress at the point of location of the dislocation.

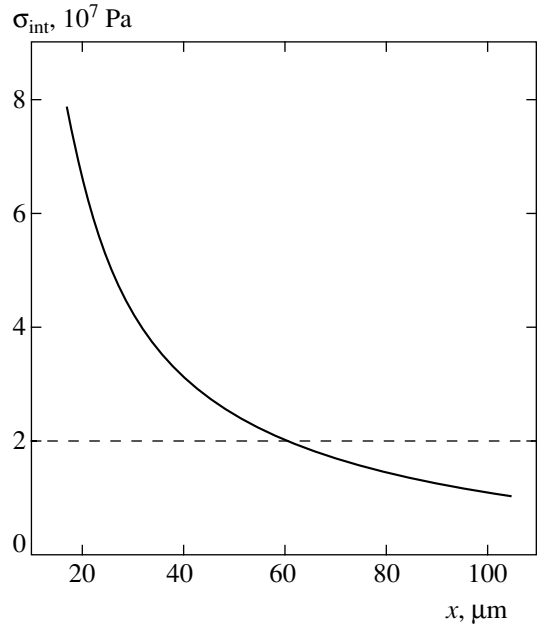

Fig. 1. Profile of internal stresses σ_{int} acting on a dislocation for $\Theta_d^0 = 1$ and $a = 80$ nm. The dashed line indicates the value of Peierls stress for the given crystal.

The stresses created by a linear dislocation decrease in inverse proportion to distance $x - x'$ [14]:

$$\sigma_{xy} = \frac{bG}{2\pi(1-\nu)(x-x')}, \tag{1}$$

where $\nu = 0.3$ is the Poisson ratio, $G = 7.5 \times 10^{10}$ Pa is the shear modulus [15], and x' is the coordinate of the dislocation, measured from the edge of the scratch.

From a large number of dislocations emerging in the vicinity of a scratch ($N_d \approx 2 \times 10^{12}$ m^{-2}), only an insignificant part can move over the crystal. Consequently, elastic stresses created by dislocations are determined by their main array, which is spatially localized in a narrow region of length $a \ll x_{max}$ (x_{max} is the maximum path length of "actuated" dislocations):

$$\sigma_{xy}(x) = -\frac{G}{2\pi(1-\nu)}\int_0^a \frac{\Theta_d(\xi)d\xi}{\xi - x}. \tag{2}$$

Here, $\Theta_d(\xi)$ is the number of dislocations distributed on segment $[0, a]$.

Equation 2 can be simplified considerably by assuming that the dislocation distribution in the "passive" region $\pm a$ is constant:

$$\Theta_d(\xi) = \Theta_d^0 = \text{const}. \tag{3}$$

In this case, integration of Eq. (2), taking into account relation (3) in the indicated limits, gives the stress distribution associated with the scratch:

$$\sigma_{int}(x) = \sigma_{xy}(x) = -\frac{G\Theta_d^0}{2\pi(1-\nu)}\ln\left(1 - \frac{a}{x}\right). \tag{4}$$

This equation is illustrated graphically in Fig. 1. It can be seen that action σ_{int} on dislocations is the smaller, the larger x. A comparison of theoretical values of σ_{int} for silicon with the Peierls stress

$$\sigma_P = \frac{2G}{1-\nu}\exp\left(-\frac{2\pi d}{b(1-\nu)}\right) = 2 \times 10^7 \text{ Pa}$$

makes it possible to estimate the range of action of internal stresses (Fig. 1; $x_{max} \approx 60$ μm). Here, b =

0.38 nm is the modulus of the Burgers vector and d = 0.134 nm is the distance between slip planes [16].

In order to experimentally verify the role of internal stresses, we carried out the first series of experiments without applying an external load to the samples ($\sigma_{ext} = 0$). Typical results of these experiments are represented by curve *1* in Fig. 2.

This series of experiments is distinguished by a clearly observed delay of transport processes at the initial stages of isothermal annealing. For example, no appreciable change in the motion of dislocations could be detected during the first 10 min, while the path lengths of head dislocations during the next 5 min attained a value of 15 μm. A starting delay in dislocation transport (in silicon and germanium) under external loading was also observed earlier by other authors [17–19] and was attributed to the conditions of drawing linear defects to the starting position, preliminary thermal treatment regimes, and experimental conditions.

An analysis of the mobility of linear defects for $\sigma_{ext} = 0$ proved that the path lengths of dislocations rapidly increase in the course of isothermal annealing, although the velocity of motion determined by the elastic stress profile decreases gradually and practically vanishes after 100 min of isothermal holding.

The obtained results enabled us to describe the time evolution of the transport of linear defects in the field of internal stresses in the form of the empirical dependence

$$x = \theta x_{max}\left(1 - \frac{t^0}{t_{\sigma_{ext}=0}}\right), \quad (5)$$

where t^0 and $t_{\sigma_{ext}=0}$ are the starting dislocation delay time and the time of sample annealing for $\sigma_{ext} = 0$, θ being the step function

$$\theta = \frac{f + |f|}{2f} = \begin{cases} 1, & t_{\sigma_{ext}=0} > t^0, \\ 0, & t^0 > t_{\sigma_{ext}=0} > 0, \end{cases}$$

where $f = (t_{\sigma_{ext}=0} - t^0)/t_{\sigma_{ext}=0}$.

Approximation of experimental data allowed us to calculate the values of the delay time (t_0 = 657 s) of dislocation motion and the maximum range of dislocations (equal to 62.44 μm for $t \longrightarrow \infty$), which is in good agreement with the results of calculations (see Fig. 1). Consequently, for large distances $x > x_{max}$, relaxation of internal stresses cannot be realized by motion of dislocations from a scratch.

The application of an external deforming stress σ_{ext} to a sample with scratches radically changes the dynamics of motion of linear defects since the effective

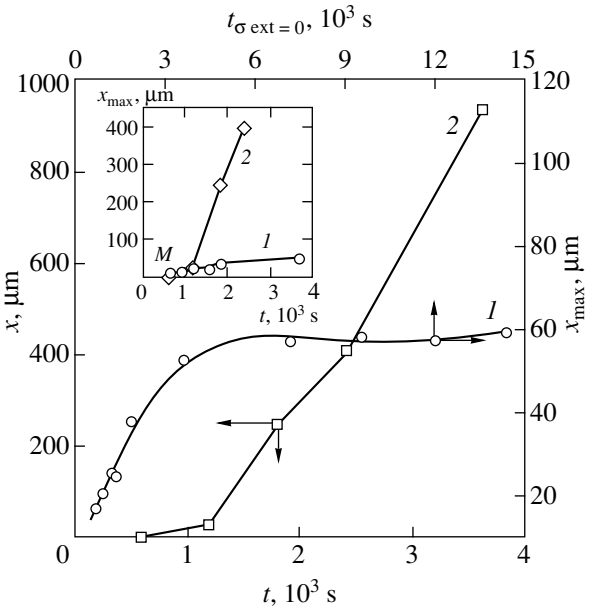

Fig. 2. Time dependence of the range of head dislocations from a scratch inscribed on the (100) surface of *p*-Si under a load of 8.4 N (ρ = 1 Ω cm, T = 950 K): annealing in the absence of external load (*1*) and under an external load of 60 MPa (*2*). The inset shows the initial segment of the curves.

stress σ^* acting on a dislocation at any point is defined as [20, 21]

$$\sigma^* = \sigma_{int} + \sigma_{ext}. \quad (6)$$

Indeed, only the starting delay (about 11 min) and the initial stage of dislocation transport (about 20 min) exhibit nearly complete coincidence of the experimental results represented by curves *1* and *2* in Fig. 2 (see the inset to this figure). It can be seen that the region of responsibility of internal stresses for curve *2* is bounded by point *M* beyond which the motion of dislocations is controlled by external stresses determining the constant rate of their transport (in contrast to curve *1* in Fig. 2). Consequently, the difference in the behavior of the curves observed beyond point *M* is associated with redistribution of stresses in favor of external perturbing forces, while the experimental results for σ_{ext} obtained in the region of influence of internal stresses require the inclusion of σ_{int}. For this reason, theoretical treatment of experimental data for external perturbing factors (including magnetic field) was carried out only for regions beyond the zone of influence of σ_{int}.

3. THEORETICAL ANALYSIS

Let us suppose that the motion of dislocations in the field of external forces is determined by three main types of stoppers: (i) those associated with the intrinsic barrier of the lattice (subscript i = 1), (ii) magnetosen-

sitive point defects ($i = 2$), and (iii) intersecting linear defects including "forest" dislocations ($i = 3$).

Depending on physical conditions, local barriers pinning a dislocation can be overcome in parallel or in series. The dominating manner of overcoming obstacles depends on the ratio of the Peierls barriers σ_P and acting external stresses σ^* [22, 23]. If the applied stress is much lower than the Peierls stress, the motion of dislocation at velocity v' is accompanied by the formation of paired kinks [22, 23]. Otherwise, the motion of dislocations is controlled by consecutive overcoming of stoppers along a route parallel to that in question in the direction of dislocation motion [24]. When the contributions from quantities σ_P and σ^* are commensurate, these two independent routes become equivalent. Let us consider this situation in greater detail.

Let τ' be the mean time of transition of a dislocation ($\sigma_P \gg \sigma^*$) from one metastable state to another. If the number of such transitions is

$$n' = x\sum_{i=1}^{m} C_i,$$

where C_i is the concentration of stoppers per unit path length of the dislocation, the mean velocity of its motion on a segment of length x is given by

$$v' = \frac{x}{n'\tau'}. \tag{7}$$

The mean time

$$\tau' = \frac{1}{2\rho_k}\sum_{i=1}^{m} C_i'\tau_i$$

is determined by transport processes involving paired kinks with the linear density ρ_k, the concentration C_i' of stoppers per unit length along the dislocation line, and the times τ_i' of overcoming the kinks of the relevant types of stoppers. Taking this into account and assuming that the dislocation length $L > 1/\rho_k$, we can write the resultant velocity of the dislocation in the form

$$v' = \frac{2n_k}{\left(\sum_{i=1}^{m} C_i\right)\left(\sum_{i=1}^{m} C_i'\tau_i'\right)} = 2n_k l_k v_k, \tag{8}$$

which matches the equation derived in [22]. Here, $l_k = (\sum_{i=1}^{m} C_i)^{-1}$ is the distance between the nearest stoppers and $v_k = (\sum_{i=1}^{m} C_i'\tau_i')^{-1}$ is the velocity of a kink provided that it overcomes m stoppers in succession.[1]

[1] A similar result for the kink velocity v_k was obtained in [23], where the dynamics of dislocations in crystals with a high Peierls barrier in the presence of point defects ($m = 2$) was studied.

For $\sigma^* > \sigma_P$, velocity v'' of dislocations is limited by point and linear defects which form a discrete spectrum of obstacles [24]. In this case, the resultant velocity v of dislocations and their mechanical mobility $\mu_m = v/F$ are determined by the contribution of each type of stoppers under investigation:

$$\mu_m = \frac{1}{F}(\tau_1 C_1 + \tau_2 C_2 + \tau_3 C_3)^{-1}, \tag{9}$$

where F is the force acting on a defect, $\tau_i = \tau_{0i}\exp(E_i/k_BT)$ are the delay times of dislocations at the corresponding types of stoppers, and E_i is the Peierls barrier ($i = 1$) and the pinning energies at the corresponding types of stoppers ($i = 2, 3$).

Indeed, a moving dislocation encounters different types of stoppers on its path, being held at them for characteristic times. Depending on the number $n_i = xC_i$ of encountered barriers, the dislocation traverses a segment of length x during the time

$$\tau = \sum_{i=1}^{3}\tau_i n_i = x\sum_{i=1}^{3}\tau_i C_i. \tag{10}$$

Then the resultant velocity of dislocations in the field of elastic stresses can be represented in the form

$$v = \frac{x}{\tau} = \left(\sum_{i=1}^{3}\tau_i C_i\right)^{-1}, \tag{11}$$

which confirms the validity of formula (9). Equation (11) shows that the resultant velocity v is controlled by the maximum value of $C_i\tau_i$.

For $\sigma_P \approx \sigma^*$, we must take into account both contributions to the resultant velocity of a dislocation:

$$v = v' + v'' = 2n_k\left[\left(\sum_{i=1}^{m} C_i\right)\left(\sum_{i=1}^{m} C_i'\tau_i'\right)\right]^{-1} + \left(\sum_{i=1}^{3}\tau_i C_i\right)^{-1}.$$

However, in our investigations (see curve *1* in Fig. 2, corresponding to experiment with zero external load), this version of stress distribution is realized only in a narrow spatial domain localized in the vicinity of the scratched region of the crystal and should not be taken into account specially. In the framework of this approach, let us analyze components $C_i\tau_i$.

The first type of stoppers is associated with intrinsic relief of the crystal lattice (Peierls relief), which controls the dislocation transport in a defect-free crystal. This enables us to represent the first term, $C_1\tau_1$, in the form of a constant independent of coordinates and of the time of dislocation travel. Consequently, the partial

velocity of the dislocation must also be constant:

$$v_1 = \frac{1}{C_1\tau_1} = \text{const.} \quad (12)$$

The situation with the second (magnetosensitive) type of stoppers with concentration C_2, which we identified with boron or phosphorus in silicon, is more complicated. Any of these paramagnetic impurities is redistributed between impurity centers with concentrations C_{2a} and C_{2b} with different orientations of electrons spins and the corresponding delay times τ_{2a} and τ_{2b}.

Following the logic of Eqs. (9)–(11) and the material balance equation

$$C_2 = C_{2a} + C_{2b}, \quad (13)$$

we can easily show that, in Eq. (9), we have

$$\tau_2 C_2 = \tau_{2a}C_{2a} + \tau_{2b}C_{2b} = C_2\tau_{2b} - C_{2a}(\tau_{2b} - \tau_{2a}). \quad (14)$$

In this case, the partial velocity v_2 associated with magnetosensitive stoppers must obey the law

$$v_2 = (\tau_{2a}C_{2a} + \tau_{2b}C_{2b})^{-1}, \quad (15)$$

or

$$v_2 = [C_2\tau_{2b} - C_{2a}(\tau_{2b} - \tau_{2a})]^{-1}. \quad (15')$$

However, components C_2, in contrast to C_1, are not constant in time t since the number of stoppers with shorter delay times increases actively in a magnetic field, and relaxation processes occurring after the cessation of magnetic treatment ($t > t_B$) are characterized by the return of the magnetostimulated subsystem of structural defects to the initial unperturbed state. This requires the inclusion of the evolutionary change in components C_2 at various stages of stopper redistribution.

If relaxation processes following magnetic treatment (e.g., for C_{2a}) occur at a rate of

$$\frac{dC_{2a}}{dt} = -kC_{2a}, \quad (16)$$

the concentration of stoppers accelerating the motion of dislocations due to smaller values of τ_{2a} must vary as

$$C_{2a} = C_{2a}^0\exp(-kt^*), \quad (17)$$

where k is the rate constant of recovery of the electronic subsystem, $t^* = t - t_B$ is the relaxation time for magnetostimulated changes, and C_{2a}^0 is the initial value of the concentration of stoppers with the corresponding orientation of electron spins.

The initial value of concentration C_{2a}^0 is determined by the magnetic perturbation level. For a fixed field, the rate of redistribution of magnetosensitive stoppers in favor of C_{2a} can obviously be represented by equations

$$\frac{dC_{2a}^0}{dt_B} = k_{2b}C_{2b}^0 - k_{2a}C_{2a}^0 \quad (18)$$

or

$$\frac{dC_{2a}^0}{dt_B} = k_{2b}(C_2 - C_{2a}^0) - k_{2a}C_{2a}^0, \quad (18')$$

where k_{2a} and k_{2b} are the rate constants of formation of stoppers with different spin orientations. The solution to this equation has the form

$$C_{2a}^0 = \frac{k_{2b}}{k_{2a} + k_{2b}}C_2 - \left(\frac{k_{2b}}{k_{2a} + k_{2b}}C_2 - C_{2a}^{00}\right)\exp[-(k_{2a} + k_{2b})t_B], \quad (19)$$

where C_{2a}^{00} is the equilibrium value of concentration for $t_B = 0$.

Thus, the dislocation velocity controlled by magnetosensitive stoppers must obey Eq. (15) taking into account expressions (17) and (19):

$$v_2 = \left\{C_2\tau_{2b} - \left[\frac{k_{2b}}{k_{2a} + k_{2b}}C_2 - \left(\frac{k_{2b}}{k_{2a} + k_{2b}}C_2 - C_{2a}^{00}\right)\exp[-(k_{2a} + k_{2b})t_B]\right] \times (\tau_{2b} - \tau_{2a})\exp[-k(t - t_B)]\right\}^{-1}. \quad (20)$$

As regards the third type of stoppers with junctions of linear defects [15], their effect on the values of quantities v and μ_m are manifested via the dislocation density N_d. The concentration of such stoppers can be defined as

$$C_3 = \gamma\sqrt{N_d}\cos\alpha\cos\beta, \quad (21)$$

where γ is the number of slip planes and α and β are the angles between the slip planes and the Burgers vectors of interacting dislocations, respectively.

In this case, the dislocation path length in the range of large values of N_d must obey the law

$$x_3 = \frac{t}{C_3\tau_3} = \frac{t}{\tau_3\gamma\sqrt{N_d}\cos\alpha\cos\beta} \quad (22)$$

and must increase indefinitely for small values of $C_3\tau_3$, giving way to transport control by other stoppers.

4. EXPERIMENTAL RESULTS AND DISCUSSION

We verified experimentally the basic equations considered here using the technique described above. The main experimental results are represented in Figs. 3–7.

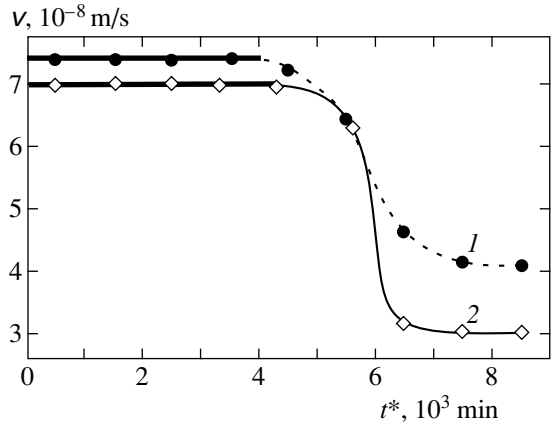

Fig. 3. Evolution of the dislocation velocity during sample storage under natural conditions after treatment in a magnetic field $B = 1$ T for 20 min. Symbols correspond to experiment; bold segments of the curves are the results of calculations based on Eq. (20) ($v = 3.7 \times 10^{-8}$ and 2.4×10^{-8} m/s for $t_B = 0$ for *p*-Si (*1*) and *n*-Si (*2*), respectively).

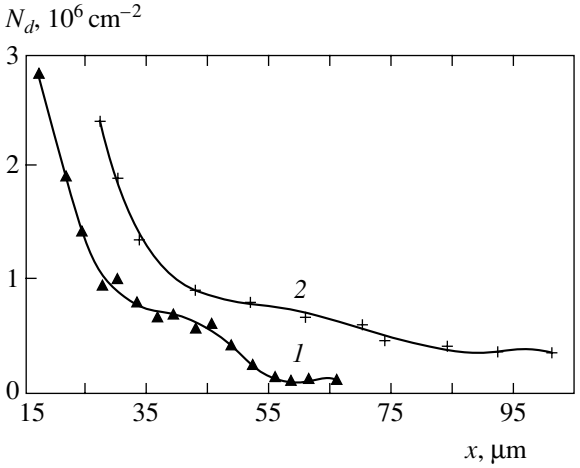

Fig. 4. Spatial distribution of dislocations from a crack in the course of plastic deformation of silicon for 40 min at $T = 948$ K, $t_B = 0$: *n*-Si (*1*) and *p*-Si (*2*).

Preliminary treatment of samples in a magnetic field actuates dislocations, accelerating their motion in the crystal. The effectiveness of this action, which is identified as "magnetic memory", is preserved for at least three days, after which the mobility of dislocations and, hence, their velocity abruptly decrease, tending to values for the initial samples that have not experienced magnetic treatment (Fig. 3). The observed changes in the $v(t^*)$ dependence upon a monotonic attenuation of magnetic memory over 73 h are in good agreement with Eq. (19) for fixed values of t_B. The rate constants k of recovery of the electronic subsystem determined from this matching are given in Table 2.

Figure 4 shows that the expulsion of linear defects to the field of elastic stresses strongly depends on the dislocation concentration profile,[2] increasing, for example, for *n*-Si by a factor of 3 upon a decrease in N_d from 3×10^6 to 0.3×10^6 cm^{-2}. The strongest variations of velocity v occur in spatial regions with large values of N_d, while the velocities of "fast" dislocations in the vicinity of small values of N_d are practically constant or change insignificantly. The depth of penetration of individual dislocations in the temperature range under investigation is a linear function of the time of isothermal annealing of the crystal (0.5–16 h). This indicates the dominating role of "dislocation" stoppers with concentration C_3 only in defective regions adjacent to scratches, where their effect on the resultant velocity v of expulsion is manifested most clearly. This allows us to eliminate the effect of the type of stoppers in question for the most remote dislocations, thus tracing the effect of the magnetic perturbation factor alone.

[2] The term "concentration profile" is applied to the spatial distribution of etch pits in the vicinity of a stress concentrator.

The temporal effect of magnetic field ($B = 1$ T) on remote dislocations is ambiguous both for *n*- and for *p*-Si (Fig. 5). In spite of quantitative difference, we can trace a clear tendency to an increase in velocity and, hence, the concentration C_{2a} of stoppers upon an increase in the time of magnetic treatment of the plates. This points to a decisive role of magnetosensitive stoppers with concentration C_{2a} and its components in the transport of dislocation segments in the range of small N_d.

Good agreement between our experimental results and Eq. (19) enabled us to obtain numerical estimates for the main parameters of dislocation transport (see Table 2) for a fixed degradation time t^*. For example,

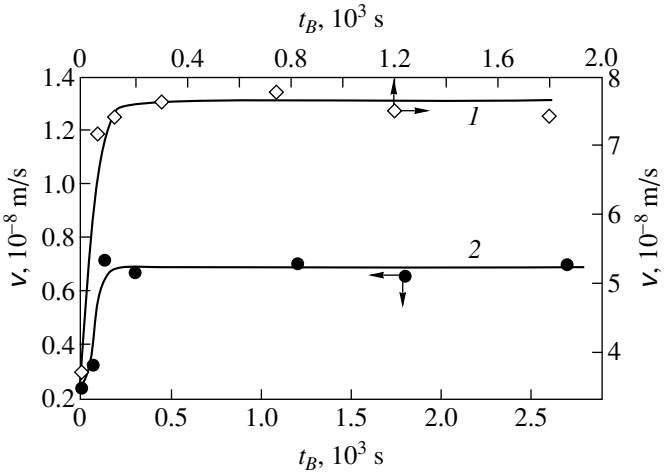

Fig. 5. Dependence of the maximum velocity of expulsion of dislocations in silicon on the time of magnetic treatment of the sample in field $B = 1$ T. Symbols correspond to experiment; the curves are the results of calculations based on Eq. (20).

Table 2. Parameters of magnetostimulated dislocation transport after isothermal annealing for 40 min; B = 1 T, t^* = 180 s

Crystal type	ρ, Ω cm	Impurity concentration, 10^{22} m^{-3}	C_2, 10^7 m^{-1}	C_{2a}^{00}, 10^7 m^{-1}	τ_{2b}, s	τ_{2a}, s	k_{2b}, 10^{-2} s^{-1}	k_{2a}, 10^{-2} s^{-1}	k, 10^{-6} s^{-1}
p	0.5	4.0	3.4	3.3	0.64	0.02	1.3	2.0	1
n	2	0.5	1.8	1.4	1.17	0.06	0.34	0.76	5

the curves in Fig. 6 characterize the redistribution of stoppers with concentration C_{2i} and partial velocities of dislocations $v_{2i} = (C_{2i}\tau_{2i})^{-1}$ with different delay times after controllable stages of magnetic treatment. Large delay times for dislocations at magnetosensitive stoppers (see Table 2), which are approximately 10^{10} times longer than the characteristic times of spin conversion (10^{-11}–10^{-10} s [1–4, 12]) are also worth noting. Consequently, the delayed response of velocity v_2 to the magnetic perturbation factor (see Fig. 4) is connected with delayed processes of formation of complexes on the basis of impurity centers with differently oriented spins rather than with spin conversion:

$$C_{2a} \overset{k_{2a}}{\equiv} C_{2a}^{+} + \xi_{2a},$$
$$C_{2b} \overset{k_{2b}}{\equiv} C_{2b}^{-} + \xi_{2b}. \tag{23}$$

Here, superscripts "+" and "–" determine the configuration of electron spins of impurity centers and ξ_{2i} are structural component of the C_{2i} complexes.

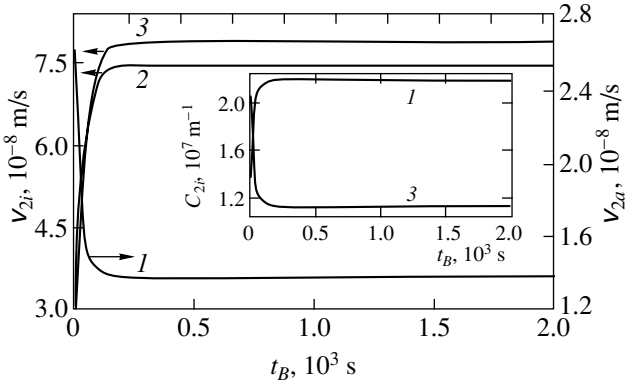

Fig. 6. Effect of magnetic treatment time for *n*-Si on the concentration redistribution of magnetosensitive stoppers and partial velocities of dislocations, calculated by formula (19): C_{2a}, v_{2a} (*1*), resultant velocity (*2*), and C_{2b}, v_{2b} (*3*).

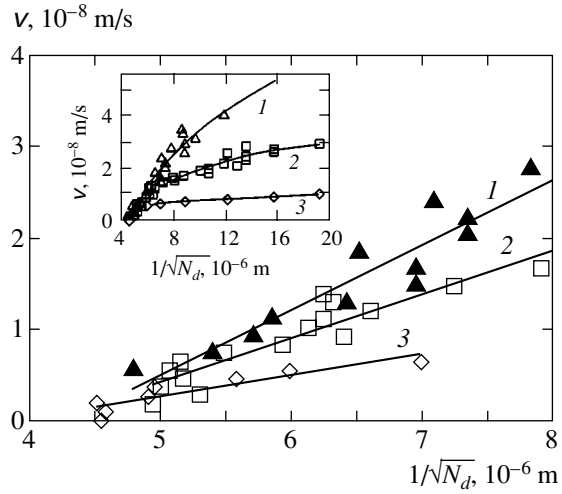

Fig. 7. Dependence of the velocity of expulsion of dislocations in spatial regions controlled by stoppers with concentration C_3 (solid lines) on the quantity $1/\sqrt{N_d}$ determining the distance between dislocations at 903 (*1*), 873 (*2*), and 823 K (*3*). The inset shows the total $v(N_d)$ dependence.

The quantitative information obtained above concerned individual dislocations located at large distances from scratched regions of the crystal. At the same time, regions in the silicon crystal with large values of N_d, i.e., with the dominating effect of "dislocation" stoppers C_3, also carry important information, which can be extracted on the basis of matching the experimental (Fig. 7) and analytic (see Eq. (21)) values of v_3.

Indeed, the velocity of expulsion of dislocations in spatial regions controlled by the type of stoppers in question (see Fig. 7) is correctly described by formula (21) for all temperature regimes under investigation. This enables us to determine the average time of dislocation pinning at the corresponding stoppers, τ_3 = 50 s, as well as the energy barrier E = 0.9 eV for overcoming these stoppers. The activation barriers of dislocation depinning from magnetosensitive stoppers, which have been determined from the values corresponding to individual dislocations, are equal to 2.1 eV. The obtained values are in accord with the available data from the literature [15, 21].

5. CONCLUSIONS

Thus, we have analyzed the dislocation dynamics in silicon in the field of internal and external stresses in monocrystalline silicon plates. We fixed experimentally the motion of dislocation segments in the field of internal stresses over a distance of 30–50 μm from a crack during isothermal annealing of the silicon plates for 0.5–3 h in the temperature range 600–700°C. The

motion observed is associated with the interaction between dislocations in the vicinity of a stress concentrator.

We studied the mechanism of motion of dislocations in the presence of three types of stoppers. It was found that preliminary treatment of silicon plates by a weak magnetic field increases the velocity of expulsion of dislocation segments by a factor of 2–3. We determined the dependence of the maximal path length of dislocations of the time of exposure of the samples to a magnetic field and described the kinetics of magnetostimulated changes in the framework of the model of spin-dependent response of structural defects. Quantitative characteristics of dislocation transport under perturbing action of a magnetic field were determined, and the energy parameters of depinning of linear defects from "dislocations" and magnetosensitive stoppers were estimated.

REFERENCES

1. V. I. Al'shits, E. V. Darinskaya, T. M. Perekalina, *et al.*, Fiz. Tverd. Tela (Leningrad) **29**, 467 (1987) [Sov. Phys. Solid State **29**, 265 (1987)].
2. V. I. Al'shits, E. V. Darinskaya, and E. A. Petrzhik, Fiz. Tverd. Tela (Leningrad) **33**, 3001 (1991) [Sov. Phys. Solid State **33**, 1694 (1991)].
3. Yu. I. Golovin, R. B. Morgunov, and S. E. Zhulikov, Fiz. Tverd. Tela (St. Petersburg) **39**, 495 (1997) [Phys. Solid State **39**, 430 (1997)].
4. Yu. I. Golovin, R. B. Morgunov, V. E. Ivanov, *et al.*, Pis'ma Zh. Éksp. Teor. Fiz. **68**, 400 (1998) [JETP Lett. **68**, 426 (1998)].
5. O. I. Datsko and V. I. Alekseenko, Fiz. Tverd. Tela (St. Petersburg) **39**, 1234 (1997) [Phys. Solid State **39**, 1094 (1997)].
6. E. V. Darinskaya, E. A. Petrzhik, S. A. Erofeeva, *et al.*, Pis'ma Zh. Éksp. Teor. Fiz. **70**, 298 (1999) [JETP Lett. **70**, 309 (1999)].
7. V. I. Al'shits, E. V. Darinskaya, and E. A. Petrzhik, Fiz. Tverd. Tela (St. Petersburg) **24**, 155 (1992) [Sov. Phys. Solid State **34**, 81 (1992)].
8. V. A. Makara, L. P. Steblenko, N. Ya. Gorid'ko, *et al.*, Fiz. Tverd. Tela (St. Petersburg) **43**, 462 (2001) [Phys. Solid State **43**, 480 (2001)].
9. A. A. Skvortsov, A. M. Orlov, V. A. Frolov, *et al.*, Fiz. Tverd. Tela (St. Petersburg) **42**, 1814 (2000) [Phys. Solid State **42**, 1861 (2000)].
10. A. A. Skvortsov, A. M. Orlov, and L. I. Gonchar, Zh. Éksp. Teor. Fiz. **120**, 134 (2001) [JETP **93**, 117 (2001)].
11. M. I. Molotskii, Fiz. Tverd. Tela (Leningrad) **33**, 3112 (1991) [Sov. Phys. Solid State **33**, 1760 (1991)].
12. A. L. Buchachenko, R. Z. Sagdeev, and E. M. Salikhov, *Magnetic and Spin Effects in Chemical Reactions* (Nauka, Novosibirsk, 1978).
13. M. I. Molotskii, R. E. Kris, and V. Fleurov, Phys. Rev. B **51**, 12531 (1995).
14. L. D. Landau and E. M. Lifshitz, *Course of Theoretical Physics*, Vol. 7: *Theory of Elasticity*, 3rd ed. (Nauka, Moscow, 1982; Pergamon, New York, 1986).
15. N. N. Novikov, *Structure and Structure-Sensitive Properties of Real Crystals* (Vishcha Shkola, Kiev, 1983).
16. M. P. Shaskol'skaya, *Crystallography* (Vysshaya Shkola, Moscow, 1984).
17. J. R. Patel, Phys. Rev. Lett. **33**, 1436 (1966).
18. N. Ya. Gorid'ko, V. A. Makara, N. N. Novikov, and L. P. Steblenko, Fiz. Tverd. Tela (Leningrad) **31** (5), 31 (1989) [Sov. Phys. Solid State **31**, 738 (1989)].
19. V. A. Makara, L. P. Steblenko, V. V. Obukhovskiĭ, *et al.*, Fiz. Tverd. Tela (St. Petersburg) **42**, 854 (2000) [Phys. Solid State **42**, 877 (2000)].
20. G. A. Malygin, Fiz. Tverd. Tela (St. Petersburg) **42**, 69 (2000) [Phys. Solid State **42**, 72 (2000)].
21. T. Suzuki, H. Yosinaga, and S. Takeucti, *Dislocation Dynamics and Plasticity* (Syokabo, Tokyo, 1986; Mir, Moscow, 1989).
22. A. N. Orlov, *Introduction to the Defect Theory in Crystals* (Vysshaya Shkola, Moscow, 1983).
23. B. V. Petukhov, Izv. Akad. Nauk SSSR, Ser. Fiz. **51**, 708 (1987).
24. J. Diehl, G. P. Seidel, and L. Niemann, Phys. Status Solidi **12**, 405 (1965).

Translated by N. Wadhwa

Journal of Experimental and Theoretical Physics, Vol. 96, No. 3, 2003, pp. 531–537.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 123, No. 3, 2003, pp. 599–606.

SOLIDS
Structure

Heat Capacity Study of Relaxor $PbMg_{1/3}Nb_{2/3}O_3$ in a Wide Temperature Range

M. V. Gorev[a,*], I. N. Flerov[a], V. S. Bondarev[a], and Ph. Sciau[b]

[a] *L. V. Kirensky Institute of Physics, Siberian Division, Russian Academy of Sciences, Akademgorodok, Krasnoyarsk, 660036 Russia*

**e-mail: gorev@iph.krasn.ru*

[b] *CEMES-CNRS, 31055 Toulouse, France*

Received September 17, 2002

Abstract—Heat capacity of the $PbMg_{1/3}Nb_{2/3}O_3$ compound is measured using the methods of adiabatic and differential scanning calorimetry in the temperature range 80–750 K. Two blurred anomalies on the $C_p(T)$ dependence are observed in wide temperature intervals of 200–400 K and 500–700 K. The results of studies are discussed together with data on the structure and phonon spectrum in the framework of spherical random bond–random field model. © *2003 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

Perovskite-like compounds attract considerable interest owing to rich diversity of their physical properties and possible applications in various technological schemes. Magnetic, dielectric, optical, and electric properties of perovskite-like materials have been studied extensively. New phenomena observed in doped (mixed) perovskites are of special importance.

The perovskite structure is characterized by relative simplicity of the crystal lattice structure, on the one hand, and by striking flexibility, on the other hand, which makes it possible to vary the set of ions forming the lattice over a wide range, thus obtaining the required combination of properties of the material. The substitution of a combination of two cations B and B' of different valence for cation B in the ABX_3 perovskite, which gives $AB_xB'_{1-x}X_3$, is one of the methods modifying the perovskite properties. Among mixed oxygen-based perovskites, compounds with $x = 1/2$ and $x = 1/3$ have been studied most extensively. Depending on the size and valence of the cations, both these types of the compound can be obtained either in an ordered or a disordered state.

The main type of ordering in the cation sublattice in compounds with the $AB_{1/2}B'_{1/2}X_3$ composition is alteration of layers of B and B' ions perpendicular to the triad axis [111] of the simple cubic cell. This gives rise to a structure of the so-called ordered perovskite, or elpasolite, with the $Fm\bar{3}m$ symmetry; the formula of this compound can be written as $A_2BB'X_6$. Such an ordering emerges when the sizes and charges of the B and B' ions differ significantly.

In the case of $AB_xB'_{1-x}X_3$ perovskites with $x = 1/3$, ordered structures with alternating B, B', and B' layers are formed for A = Ba and B = Zn, Mg; B' = Nb, Ta [1]. On the other hand, compounds with A = Pb demonstrate ordering of B' and B" = (1/3)B' + (2/3)B layers alternating along the [111] direction. This type of ordering with a long-range order is observed in $PbMg_{1/3}Ta_{2/3}O_3$. In other Pb-containing compounds, no long-range order is observed. According to the results of structural analysis, ordered regions have a size on the order of 20–800 Å [1, 2].

Compounds with such a small correlation length of ordering are characterized by relaxor-type behavior. Relaxors possess three main features of dielectric response. The phase transition has a clearly manifested diffusion nature, and the permittivity and the temperature corresponding to its maximal value depend significantly of the measuring field frequency. The response in weak fields does not obey the Curie–Weiss law. The room-mean-square polarization exists in a temperature range 200–300 K higher than T_{max} (the temperature corresponding to the permittivity peak), but the average polarization P differs from zero at a temperature much lower than T_{max}.

The $PbMg_{1/3}Nb_{2/3}O_3$ (PMN) compound belongs to the family of an oxygen-containing mixed perovskite-like compound and is a classical representative of materials with relaxor behavior [3]. This compound has a broad peak of ε in the vicinity of $T_{max} \approx (250–270)$ K and displays a strong low-frequency permittivity dispersion. On the other hand, spontaneous polarization is observed at low temperatures (below 250 K) only in polarized samples in electric fields of strength higher than a certain critical value [4] or in those with $PbTiO_3$ admixtures [5]. Such dielectric behavior was initially

explained by the diffusion nature of the ferroelectric phase transition associated with fluctuations of the concentration of Mg^{2+} and Nb^{5+} ions of different valence in the B sites of the $Pm3m$-cubic structure of the ABO_3 perovskite.

Over several decades, PMN has been studied using various methods, but a large number of questions concerning the nature of phase transitions and anomalous behavior of physical properties of this crystal in a wide temperature range remain unanswered. Moreover, the crystal structure and its variation in a wide temperature range are still objects of discussion.

The main attention in the study of relaxors is paid to dielectric properties of these materials, their structure, and spectral parameters. Thermodynamic properties, especially heat capacity and its behavior in a wide temperature range, were studied casually. Nevertheless, thermal parameters of the system may be important in the developing and refining models of the phenomena in question. Calorimetric studies have indisputable advantages since this method makes it possible to detect any type of heat capacity anomalies, which are associated with the electric as well as with the elastic subsystem.

The temperature dependence of heat capacity of PMN was studied by many authors [6–8]. In [6], heat capacity was measured by the differential scanning calorimeter method in the temperature range 140–790 K. The spread in experimental points exceeded 15%, and analysis of heat capacity was carried out only after averaging, which did not allowed the authors of this work to draw unambiguous conclusions on the type of the $C_p(T)$ behavior. In [7, 8], measurements were confined to a temperature of 300 K, which is insufficient for tracing the variation of heat capacity in the region of the permittivity maximum (about 250–350 K). The authors of these publications were mainly interested in the behavior of heat capacity at low temperatures and in the form of the vibrational spectrum of the compound.

Here, we report on the results of heat capacity studies in PMN in a wide temperature range of 80–750 K and analyze its behavior in temperature regions in which anomalies in the electrical and structural properties were observed by many authors.

2. SAMPLE AND MEASURING TECHNIQUE

The experimental sample was in the form of a finely dispersed powder and was prepared at the Center of Material Development and Structural Studies (CEMES) in Toulouse (France). The compound was manufactured using solid phase synthesis from a mixture of corresponding oxides [9].

Heat capacity was measured using two methods. In the temperature range 80–350 K, measurements were made using the adiabatic calorimetry method enabling us to obtain absolute values of the total heat capacity to a high degree of accuracy. A powder sample was placed into an indium container, which was sealed in a helium atmosphere. Helium was used as a heat-exchange gas for leveling out the temperature over the sample. The sample mass was 3.17 g and the container mass was 8.5 g. Measurements were made using the traditional method of discrete heating and in an automated regime of continuous heating [10] at a temperature variation rate of $dT/dt \approx 10^{-1}$ K min^{-1}. In our experiments, we measured the total heat capacity of the sample and the container. The heat capacity of the container was measured separately. The error of heat capacity measurements depends on the method of heating and amounts to (0.1–1.0)%.

In a high-temperature range of 360–750 K, measurements were made on an updated and automated DSM-2M differential scanning calorimeter. In this case, the sample mass was 0.51 g. The error of heat capacity measurements with such a calorimeter is slightly larger than that with an adiabatic calorimeter and amounts to 2–3%.

3. RESULTS AND DISCUSSION

The results of heat capacity measurements in PMN are presented in Fig. 1. The spread in experimental points relative to the smoothing curve does not exceed 1.5%. The error of measurements slightly exceeding the conventional value is due to the fact that the heat capacity of the sample amounts to only a small part of the total heat capacity of the container–sample system. The obtained results are in good agreement with the data obtained in [7, 8] in the temperature range 80–270 K and differ from the results obtained in [6]. The curve describing the temperature dependence of specific heat does not exhibit clearly manifested anomalies typical of traditional phase transitions. However, in temperature regions near 300 and 650 K, two broad, blurred anomalies in heat capacity are observed.

3.1. Separation of the Lattice Component of Heat Capacity

Since we are dealing here with experimental results obtained in a wide temperature range, we must take into account in our analysis the anharmonic contributions to the lattice heat capacity at high temperatures, which may lead to a considerable difference between the isobaric and isochoric heat capacities. Thermal expansion of PMN has been studied repeatedly [9, 11]. The volume expansion coefficient in the temperature range from 300 to 700 K varied from 4×10^{-6} to 30×10^{-6} K^{-1}. To our knowledge, data on compressibility of PMN are not available. We estimated the value of this quantity using the result of publication [12], devoted to the study of the effect of pressure on the unit cell parameters of the $PbZr_{0.5}Ti_{0.5}O_3$ compound allied with PMN. The bulk compressibility modulus was estimated as 2.3×10^{-11} Pa^{-1}.

Using the above estimates, we calculated the anharmonic contribution to the heat capacity of PMN, which amounts to approximately 1 J/mole K at 700 K. In view of the smallness of this value, which is within the experimental error, the difference between C_p and C_v was disregarded in the subsequent analysis. The small value of the anharmonic contribution is due to a small value of the thermal expansion coefficient.

The expression for the lattice heat capacity C_L can be derived from the general relation

$$C_L = \frac{\partial E}{\partial T} = \frac{\partial}{\partial T}\left(\int d\omega G(\omega) n(\omega) \hbar\omega\right), \qquad (1)$$

where E is the internal energy of the system, $n(\omega)$ is the Bose–Einstein distribution function, and $G(\omega)$ is the density of vibrational states.

The density of state function $G(\omega)$ for crystals with a complex composition is unknown as a rule in the entire frequency range. For this reason, we must resort to a simplified approach to obtain a quantitative description of the temperature dependence of heat capacity. In most cases, the analysis of the temperature dependence of heat capacity and the separation of the lattice and anomalous contributions are carried out using a simple model describing the lattice heat capacity of the compound by a combination of the Debye and Einstein functions.

In the case of PMN, such an approach was used in [6], but the treatment of the temperature dependence of heat capacity in a wide temperature range (140–790 K) in the framework of the Debye model with Θ_D = 422.1 K obviously does not fit the real situation. The results of analysis of the data obtained in this approximation are represented by the dashed curve in Fig. 1 and lead to a different value of Θ_D = 520 K. It can clearly be seen that the temperature dependence of heat capacity in the low-temperature range has a non-Debye form [7, 8]: the heat capacity of PMN decreases with temperature almost linearly down to the minimal values attained in experiment. Such deviations from the $C \propto T^3$ dependence are attributed in [8] to the excitation of Einstein oscillators, are typical of a large number of crystals, and are associated with anomalies of real frequency distribution functions in the low-frequency spectral region (below 100 cm^{-1}). The results of lattice heat capacity treatment by a combination of the Debye and Einstein functions are shown by the dashed curve in Fig. 1. In the temperature range 30–200 K, the latter model describes experimental results much better.

Another version of the lattice capacity approximation was implemented in the framework of the fracton model proposed on the basis of analysis of the lattice vibration spectrum of PMN by the inelastic neutron scattering method [13]. Gvasaliya *et al.* [7, 13] believe that the phonon regime is realized at frequencies of

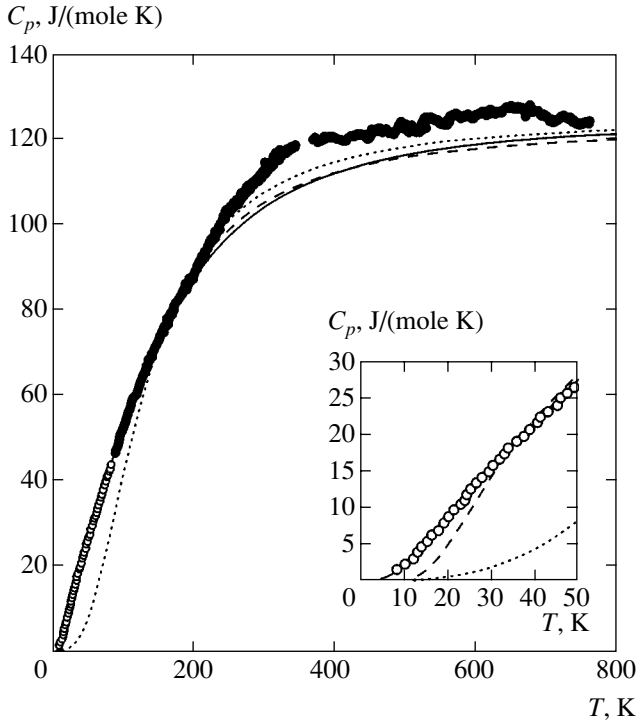

Fig. 1. Temperature dependence of the heat capacity of $PbMg_{1/3}Nb_{2/3}O_3$. Data for the temperature range 7–80 K are borrowed from [8]. Lattice heat capacity is approximated in the framework of the Debye model (dotted curve), a combination of the Debye and Einstein functions (dashed curve), and in the framework of the model developed in [7, 13] (solid curve). The inset shows heat capacity in the temperature range 0–50 K.

$\omega \le \omega_1$, while the fracton regime with spectral dimension d is observed for $\omega_1 \le \omega \le \omega_2$:

$$G(\omega) \propto \begin{cases} \omega^2, & \omega \le \omega_1, \\ \omega^{d-1}, & \omega_1 \le \omega \le \omega_2. \end{cases} \qquad (2)$$

Three parameters (d, ω_1, and ω_2) were determined by us from analysis of experimental data: $\hbar\omega_1/k_B$ = 43.7 K, $\hbar\omega_2/k_B$ = 800 K, and $d \approx 0.8$. A rather nonrealistic value of fracton dimension d is obtained, in all probability, since the dependence $G(\omega) \propto \omega^{d-1}$ was assumed to be valid for $\omega \ge \omega_1$, although it follows from the results obtained in [13] that such a dependence holds only in a narrow frequency range.

Thus, the $G(\omega)$ function used here is rather a certain approximate description of the real frequency spectrum and cannot serve as an argument supporting or rejecting the existence of fractons in the lattice vibration spectrum [7, 13]. Nevertheless, such an approximation makes it possible to describe the lattice heat capacity correctly in the entire temperature range under investigation (see Fig. 1).

The anomalous component $\Delta C_p = C_p - C_L$ of heat capacity, separated by all of the above three methods, is

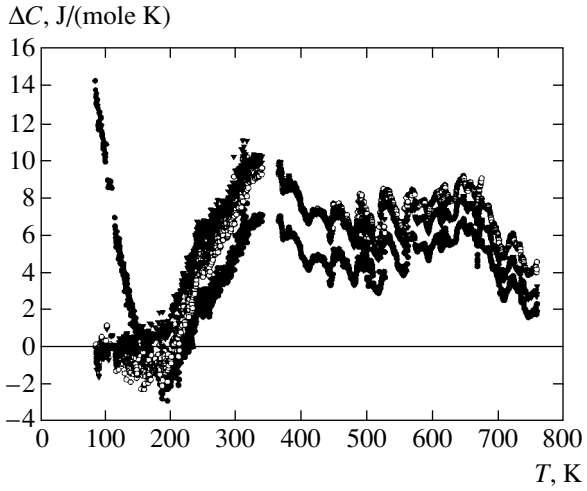

Fig. 2. Anomalous component of the heat capacity of the $PbMg_{1/3}Nb_{2/3}O_3$ compound determined by interpolating the lattice heat capacity in the framework of the Debye model (•), Debye–Einstein model (○), and the model proposed in [7, 13] (▾).

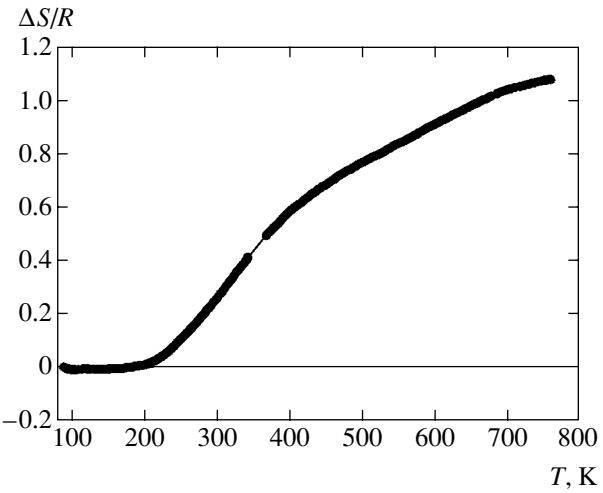

Fig. 3. Temperature dependence of the anomalous entropy of the $PbMg_{1/3}Nb_{2/3}O_3$ compound.

depicted in Fig. 2. In the temperature range $T > 100$ K, the heat capacity has a low sensitivity to fine details of the vibrational spectrum, and the difference in the values of ΔC_p obtained by approximating the lattice heat capacity by a combination of the Debye and Einstein functions and on the basis of a simple description of the lattice vibration spectrum is small and lies within the error of measurements. It should also be noted that the value of $C_p - C_v$ estimated above does not exceed 10–15% of the anomalous heat capacity.

The $\Delta C_p(T)$ dependence clearly displays two regions of anomalous behavior of heat capacity. The temperature range 200–450 K coincides with the region of anomalous behavior of permittivity [3], while deviations from the regular behavior of the refractive index [14], lattice parameters, and thermal expansion coefficient [9] were observed earlier in the temperature range 550–700 K.

The change in entropy associated with the anomalous behavior of heat capacity and calculated by the formula

$$\Delta S(T) = \int_{100\ \mathrm{K}}^{T} \frac{\Delta C_p}{T} dT$$

is depicted in Fig. 3.

In order to find reasons for the emergence of anomalous behavior of heat capacity of PMN, we consider the structural features of this material and the available models of phenomena occurring in it.

3.2. Anomalous Behavior at High Temperatures

The concept of polar nanoregions is one of the most important concepts associated with microscopic properties of relaxors. The first experimental proof of the existence of such regions was obtained by Burns, who studied the refractive index $n(T)$ of single crystals of some disordered ferroelectrics and relaxors including PMN [14]. The refractive index behavior deviates from the linear dependence at temperature T_d (600–650 K for PMN), which is much higher than $T_{\max}$. It was supposed that such an unexpected behavior is due to randomly oriented local polarized regions, which are formed in a nonpolar crystal structure below T_d. This temperature is often referred to as the Burns temperature.

Later, high-resolution electron microscopy studies of PMN [15] at $T < T_d$ revealed the presence of chemically ordered (1 : 1) clusters in which Mg^{2+} and Nb^{5+} ions, which are in the position B, alternate along the axes of the perovskite cubic lattice. The size of these clusters is on the order of 10 Å. In [16, 17], a two-phase model of the PMN structure was proposed on the basis of structure analysis by the neutron elastic scattering method at low temperatures, viz., a cubic structure with a long-range order and nanoregions with a rhombohedral $R3m$ structure and with a correlation length on the order to 100 Å. In these polar nanoregions, Pb and Mg/Nb ions are displaced relative to oxygen ions in the $[111]_c$ directions.

Neutron inelastic scattering experiments [18, 19] with relaxors above the Burns temperature revealed a TO vibration branch for all wave vectors and a TO mode at the center of the Brillouin zone. The value of $(\hbar\omega_0)^2$, where ω_0 is the frequency of this mode, decreases linearly upon cooling down to T_d. Such a behavior matches the behavior of the soft mode in $PbTiO_3$ and other displacement-type ferroelectrics. At the Burns temperature, a phase transition occurs, leading to a rhombohedral ($R3m$) distortion of Nb-rich clus-

ters and to the emergence of polarization in these clusters. However, the correlation length of such clusters is quite small and macroscopic polarization does not appear in the sample.

In the region of soft mode condensation at the Burns temperature, anomalous behavior of heat capacity of PMN can be expected, which was indeed observed in our experiments in the temperature range 600–700 K. The anomaly blurring is probably due to the fact that individual clusters have different phase transition temperatures.

The approximate change in entropy during the high-temperature transition is on the order of (0.4–0.5)R (R is the universal gas constant); this means that processes of the order–disorder type, which are associated with ordering of lead ions in interoctahedral cavities below the Burns temperature, play a significant role in the formation of polar nanoregions.

Taking into account the results of structural studies and the fact that polarization may have eight equivalent directions of the $[111]_c$ type, we could expect that the entropy variation as a result of ordering in the entire crystal is $\Delta S = R\ln 8$. The obtained value of entropy variation amounts to only 20–25% of this value and is in accordance with the volume fraction of polar nanoregions obtained from an analysis of structural data [17].

3.3. Analysis of the Behavior of Heat Capacity in the Framework of the Random Bond–Random Field Theory

Since the discovery of relaxors more than 30 years ago, several models have been proposed for explaining the existence of the ε peak and other peculiar dielectric properties. These are the models of diffusion phase transitions, dipole glasses, reorienting polar clusters, etc. In recent years, the spherical random bond–random field (SRBRF) model has become most popular [20, 21].

The PMN compound below the Burns temperature is treated in the model [20] as a heterogeneous material consisting of Nb-rich regions (or polar clusters) implanted in a quasi-regular array of chemically ordered (1 : 1) regions (chemical clusters). Polar clusters have a typical size of several nanometers and can be reoriented. It is these clusters that are responsible for the observed dielectric behavior. On the contrary, chemical clusters are static and are regarded as sources of random electric fields.

The model Hamiltonian of the system of interacting polar clusters can be formally written in the form

$$H = -\frac{1}{2}\sum_{i,j} J_{ij}\mathbf{S}_i\cdot\mathbf{S}_j - \sum_i \mathbf{h}_i\cdot\mathbf{S}_i - g\sum_i \mathbf{E}_i\cdot\mathbf{S}_i, \quad (3)$$

where $\mathbf{S}_i$ is the dimensionless order parameter, proportional to the dipole moment of a cluster; J_{ij} are random interactions or bonds; $\mathbf{h}_i$ are random local electric fields; and $\mathbf{E}_i$ is the external electric field [20]. As in the theory of spins and dipole glasses, it is assumed that random bonds have a Gaussian distribution and

$$[J_{ij}]^c_{av} = J_0/N, \quad [(J_{ij})^2]^c_{av} = J^2/N. \quad (4)$$

Random fields also obey the Gaussian distribution, and

$$[h_{i\mu}]^c_{av} = 0, \quad [h_{i\mu}h_{j\nu}]^c_{av} = \Delta\delta_{ij}\delta_{\mu\nu}, \quad \mu = x, y, z. \quad (5)$$

Equilibrium values of polarization,

$$P_\mu = \frac{1}{N}\sum_i \langle S_{i\mu}\rangle,$$

and of the order parameter of dipole glass (or the Edwards–Anderson parameter),

$$q_\mu = \frac{1}{N}\sum_i \langle S_{i\mu}\rangle^2,$$

can be determined from the conditions of the free energy minimum:

$$-\frac{2}{3}\beta f = \beta J_0 P^2 - \frac{1}{2}\beta^2 J^2 q^2 - 2z + \ln(2z+\beta^2 J^2 q) - \frac{\beta^2}{2}\frac{J^2 q + \Delta + (J_0 P + gE)^2}{2z+\beta^2 J^2 q}, \quad (6)$$

$$\partial f/\partial P = \partial f/\partial q = \partial f/\partial z = 0, \quad (7)$$

where z is a Lagrangian multiplier introduced to enhance the spherical conditions, simplifying the solution [20]. Equilibrium values of P, q, and z are defined by the following expressions:

$$P = \beta(1-q)(J_0 P + gE),$$
$$q = \beta^2(1-q)^2(J^2 q + \Delta) + P^2, \quad (8)$$
$$2z + \beta^2 J^2 q = 1/(1-q).$$

In zero external field ($E = 0$), the system has two sets of solutions. One of them ($P = 0$, $q \neq 0$) corresponds to the phase without a long-range order, or dipole glass (SG), while the other ($P \neq 0$, $q \neq 0$) describes the phase with a long-range order, or ferroelectric (FE) phase. The ferroelectric phase ($P \neq 0$) can exist only if the parameter J_0 of interaction between clusters exceeds a certain critical value, $J_0 > J_{0c} = \sqrt{J^2 + \Delta}$.

The temperature of transition to the ferroelectric state is defined as

$$T_c = J_0\left(1 - \frac{\Delta}{J_0^2 - J^2}\right)\theta\left(\frac{J_0}{J} - 1\right). \quad (9)$$

The phase diagram of the model is shown in Fig. 4. In the case of PMN, the interaction parameter J_0 is

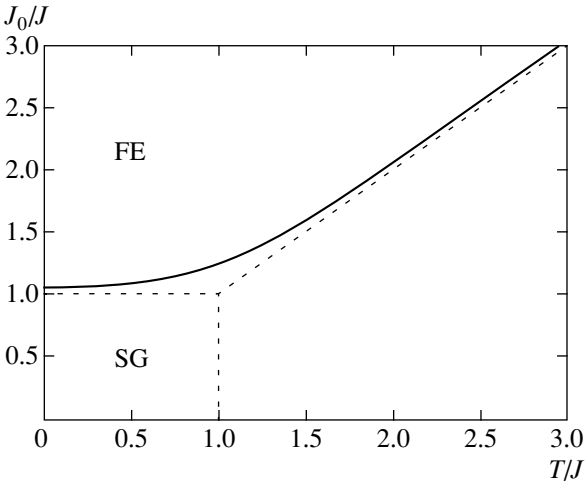

Fig. 4. Phase diagram of the random bond–random field model. Dotted lines correspond to $\Delta = 0$.

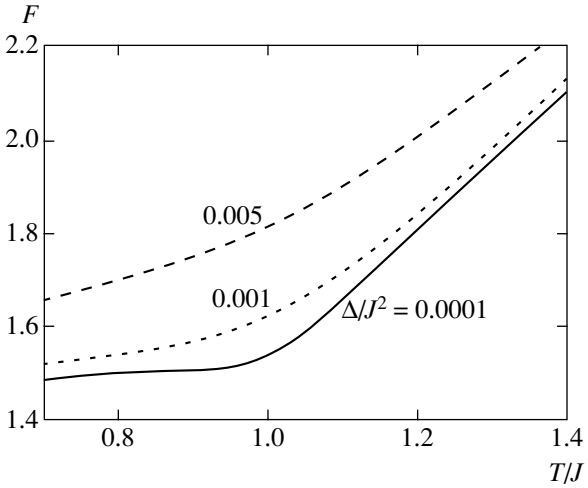

Fig. 6. Temperature dependence of the free energy in the SRBRF model for $E = 0$, $J_0 = 0$, and $\Delta/J^2 = 0.0001$, 0.001, and 0.005.

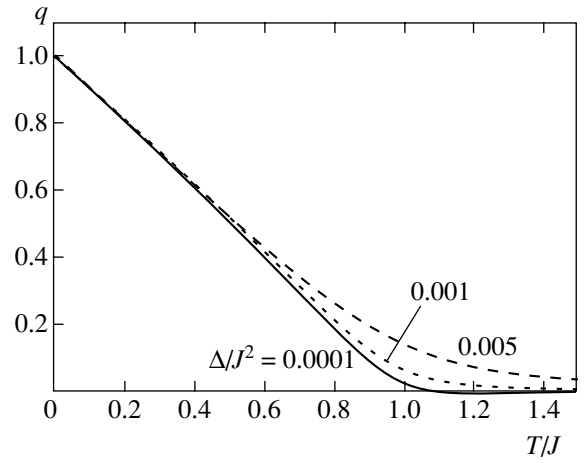

Fig. 5. Temperature dependence of the order parameter q for $E = 0$, $J_0 = 0$, and $\Delta/J^2 = 0.0001$, 0.001, and 0.005.

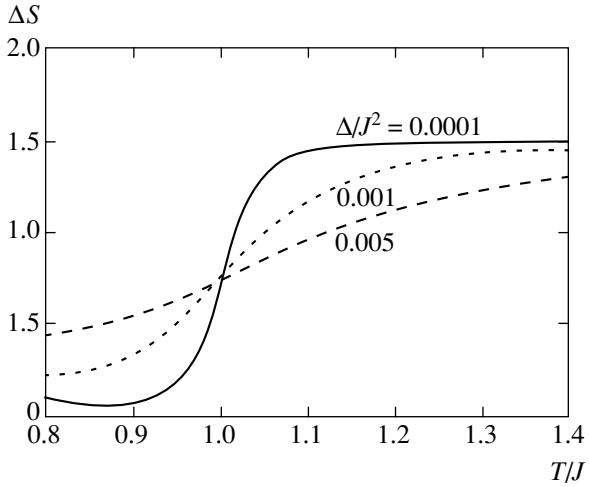

Fig. 7. Temperature dependence of entropy in the SRBRF model for $E = 0$, $J_0 = 0$, and $\Delta/J^2 = 0.0001$, 0.001, and 0.005.

apparently smaller than the critical value, and a transition to the ferroelectric phase with the emergence of a macroscopic order parameter (polarization) is not observed. Accordingly, classical heat capacity anomalies are also absent. However, the existence of parameter q and its temperature variation leads to the emergence of an additional contribution to free energy and heat capacity.

The form of temperature dependences of these contributions was determined as a result of our numerical calculations for various values of the model parameters. The results of calculation of the temperature dependences of order parameter, free energy, entropy, and heat capacity for $\Delta/J^2 = 0.0001$, 0.001, and 0.005 and for $E = 0$, $J_0 = 0$ are presented in Figs. 5–8. It can be seen that, even in the absence of a phase transition to the ferroelectric phase, the heat capacity has a broad anomaly with a peak at $T = J$. For PMN, the value of J

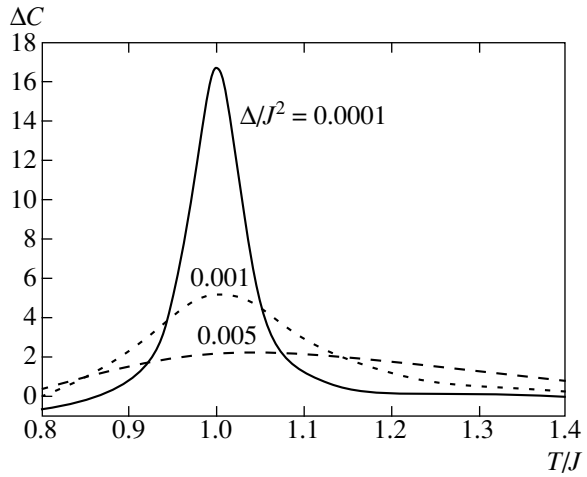

Fig. 8. Temperature dependence of heat capacity in the SRBRF model for $E = 0$, $J_0 = 0$, and $\Delta/J^2 = 0.0001$, 0.001, and 0.005.

is on the order of 300 K, and the observed heat capacity anomaly in the temperature range from 150 to 450 K is in qualitative agreement with the SRBRF model [18, 19].

4. CONCLUSIONS

Our results and the data from the literature lead to the following pattern of the processes occurring in PMN. Above the Burns temperature (≈700 K), PMN is a heterogeneous material containing ordered (1 : 1) regions as well as regions rich in niobium. The sizes of these regions are on the order of tens of nanometers. As the temperature decreases, a transition to the polar *R*3*m* phase takes place at the Burns temperature in regions with an excess of niobium; this transition is accompanied by soft mode condensation and appearance of the blurred heat capacity anomaly. The polarization directions in polar nanoclusters are distributed at random, and macroscopic polarization does not emerge in the system. Further cooling does not lead to any structural changes. Anomalous behavior of dielectric properties and heat capacity in the temperature range 200–400 K is associated with changes occurring in the system of reorientable polar clusters and a transition to the glass-like state.

The results obtained are in good agreement with the results of structural studies and with conclusions of the spherical random bond–random field (SRBRF) model.

ACKNOWLEDGMENTS

This study was supported financially by the Russian Foundation for Basic Research (project no. 00-15-96790).

REFERENCES

1. B. P. Burton and E. Cockayne, Phys. Rev. B **60**, 12542 (1999).
2. E. L. Cross, Ferroelectrics **151**, 305 (1994).
3. G. A. Smolenskiĭ, V. A. Isupov, A. I. Agranovskaya, and S. N. Popov, Fiz. Tverd. Tela (Leningrad) **11**, 2906 (1960) [Sov. Phys. Solid State **11**, 2354 (1960)].
4. G. Calvarin, E. Husson, and Z. G. Ye, Ferroelectrics **165**, 349 (1995).
5. B. Dkhil, J.-M. Kiat, G. Calvarin, *et al.*, Phys. Rev. B **65**, 024104 (2001).
6. A. Fouskova, V. Kohl, N. N. Krainik, and I. E. Mylnikova, Ferroelectrics **34**, 119 (1981).
7. S. N. Gvasaliya, S. G. Lushnikov, Ĭ. Moriya, *et al.*, Kristallografiya **46**, 1110 (2001) [Crystallogr. Rep. **46**, 1025 (2001)].
8. B. A. Strukov, K. A. Minaeva, T. L. Skomorokhova, and V. A. Isupov, Fiz. Tverd. Tela (Leningrad) **8**, 972 (1966) [Sov. Phys. Solid State **8**, 781 (1966)].
9. P. Bonneau, P. Garnier, G. Calvarin, *et al.*, J. Solid State Chem. **91**, 350 (1991).
10. M. V. Gorev, P. I. Gekk, I. M. Iskornev, *et al.*, Izmer. Tekh., No. 8, 33 (1988).
11. L. A. Shebanov, P. P. Kapostins, and J. A. Zvirgzds, Ferroelectrics **56**, 53 (1984).
12. J. Rouquette, J. Haines, V. Bornand, *et al.*, Phys. Rev. B **65**, 214102 (2002).
13. S. N. Gvasaliya, S. G. Lushnikov, I. L. Sashin, and I. G. Siniĭ, Kristallografiya **44**, 284 (1999) [Crystallogr. Rep. **44**, 250 (1999)].
14. G. Burns and F. H. Dacol, Phys. Rev. B **28**, 2527 (1983).
15. A. D. Hilton, D. J. Barber, C. A. Randall, and T. R. Shrout, J. Mater. Sci. **25**, 3461 (1990).
16. N. Mathan, E. Husson, G. Calvarin, and A. Morell, Mater. Res. Bull. **26**, 1167 (1991).
17. N. Mathan, E. Husson, G. Calvarin, *et al.*, J. Phys.: Condens. Matter **3**, 8159 (1991).
18. P. M. Gehring, S. Wakimoto, Z.-G. Ye, *et al.*, Phys. Rev. Lett. **87**, 277601 (2001).
19. K. Hirota, Z.-G. Ye, S. Wakimoto, *et al.*, Phys. Rev. B **65**, 104105 (2002).
20. R. Pirc and R. Blinc, Phys. Rev. B **60**, 13470 (1999).
21. R. Blinc, V. Bobnar, and R. Pirc, Phys. Rev. B **64**, 132103 (2001).

Translated by N. Wadhwa

Journal of Experimental and Theoretical Physics, Vol. 96, No. 3, 2003, pp. 538–542.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 123, No. 3, 2003, pp. 607–611.

SOLIDS
Electronic Properties

Glass Behavior of a Relaxor $Cd_2Nb_2O_7$ in a Weak DC Field

N. N. Kolpakova

Ioffe Physicotechnical Institute, Russian Academy of Sciences,
Politekhnicheskaya ul. 26, St. Petersburg, 194021 Russia
e-mail: kolpakova@mail.ioffe.ru

Received September 10, 2002

Abstract—The temperature behavior of the dielectric response $\varepsilon_{his}(T)$ at a frequency of 1 kHz (110 K < T < 300 K), $\delta\varepsilon(T) = [\varepsilon_{his}(T) - \varepsilon_{ZFC}(T)]$, and of the order parameter $q(T)$, characterizing $\varepsilon(T)$ in a relaxor, is investigated for the relaxor $Cd_2Nb_2O_7$ on samples with different thermal and electric histories in a dc field much weaker than the polarization saturation field. In a weak field (E_{dc} = 0.95 kV/cm), the behavior of $\delta\varepsilon(T) \propto \chi_{nl}$ changes in the region T_f = 184 K and the $\varepsilon_{his}(T)$ curves diverge monotonically at lower temperatures, indicating the development of the glassy state in the system. Analysis of the behavior of $q(T)$ in the framework of the model of spin and dipole glasses shows that random interactions of polar microscopic regions in the presence of random fields play a dominating role in the phase formed below T_f in the ZFC and FH/ZFC regimes. The relative contribution of random fields increases in the FC regime in a weak field and is manifested in the ZFH/FC regime both below and above T_f. *© 2003 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

In recent decades, various models have been proposed for explaining peculiar dielectric behavior of relaxors (relaxation dispersion of $\varepsilon^*(T)$ in a wide frequency range; slowing down of the dynamics of polar microscopic regions, accompanied by relaxation time freezing for $T \longrightarrow T_f^+$; and nonergodic behavior at low temperatures). These models include a system of reoriented polar clusters [1, 2]; a superparaelectric [3, 4] or frustrated (inhomogeneous) ferroelectric split into nanodomains under the action of random fields [5, 6]; and dipole glass with randomly interacting polar microscopic regions in the presence of random fields [7–10]. In spite of intense studies, the following questions remain unclear: (i) Is a relaxor system in zero external electric fields a dipole glass or a frustrated ferroelectric? (ii) What state (glassy or ferroelectric) develops in the system below T_f? (iii) What is the role of dc field E_{dc} in this case? In order to clarify the situation, the temperature dependences of (a) static dielectric nonlinearity $\alpha_3 = \varepsilon_3/\varepsilon_1^4$ in field E_{dc} and in zero field (ε_1 and ε_3 are the first and third harmonics of dielectric response) [9, 10] and (b) quasistatic nonlinear susceptibility $\chi_{nl} = [\varepsilon(E_{dc}) - \varepsilon(0)]/3E_{dc}^2\varepsilon_s^4$ in the frequency range from 100 Hz to 1 MHz, which can be determined from the dielectric response of the system as a function of the field for fixed values of temperature (ε is the permittivity in field E_{dc} and in zero field, while ε_s is the static permittivity) [6, 11] were investigated in PMN ($PbMg_{1/3}Nb_{2/3}O_3$) and PLZT ($Pb_{1-x}La_x(Zr_yT_{1-y})O_3$). It was found that a transition from the ergodic to a nonergodic (glassy) relaxor state occurs in the vicinity of T_f in a weak field E_{dc}, while the ferroelectric phase is formed in a strong field. However, in the frequency range corresponding to relaxation dispersion (1 Hz–1 MHz), which is most interesting for understanding the nature of the relaxor state developing below T_f, the dielectric nonlinearity of the system decreases rapidly [10]. The nonlinear susceptibility $\chi_{nl}(T)$, which can be determined in the course of field scanning, exhibits an anomaly not at T_f, as predicted by the theory of phase transition to the glassy state [12], but at $T > T_f$ [6, 11]. In the case when $\chi_{nl}(T) \propto [\varepsilon(E_{dc}, T) - \varepsilon(0, T)]$ was determined from the measurements of the dielectric response as a function of temperature for fixed values of the field, an anomaly was observed at T_f, but the measurements were made only in a strong field inducing the ferroelectric state [13]. Analogously to many glass systems [14], the permittivity in the case of field cooling $[\varepsilon_{FC}(T) \equiv \varepsilon(E_{dc}, T)]$ of PMN and PLZT in the phase formed below T_f does not coincide with the permittivity for zero field cooling $[\varepsilon_{ZFC}(T) \equiv \varepsilon(0, T)]$ in view of the violation of ergodicity of the system [4, 8, 13, 15–17]. However, the results of such measurements do not give an idea of the behavior of relaxor in zero field as well as in weak and strong fields E_{dc}.

In order to clarify the nature of the relaxor state developing below T_f and the effect of a weak dc field on the relaxation behavior of the system at low tempera-

tures, we analyze here the behavior of (i) permittivity $\varepsilon_{his}(T)$ at a frequency of 1 kHz in a weak field E_{dc} for samples with different thermal and electric histories both in traditional ZFC and FC regimes and in the FH/ZFC (heating in a field following zero field cooling) and ZFH/FC (zero field heating following cooling in the field) regimes and (ii) difference $\delta\varepsilon(T) = [\varepsilon_{his}(T) - \varepsilon_{ZFC}(T)]$ and order parameter $q(T)$ characterizing the deviation of the temperature dependence $\varepsilon(T)$ for relaxors from the Curie–Weiss law [7, 18] using the relaxor $Cd_2Nb_2O_7$ as an example. In previous studies, the behavior of $q(T)$ for relaxor samples with different thermal and electric histories has not been investigated at all. Since the value of $q(T)$ in the general case of glass systems is determined by random interactions, as well as random local electric fields and external field E_{dc} [4, 9, 14, 18], such investigations are of prime importance for understanding the origin of freezing of an ergodic relaxor state below T_f.

The choice of the relaxor $Cd_2Nb_2O_7$ with the pyrochlore-type structure for our experiments was dictated by (i) scarce information on the relaxor state formed in it below $T_f \approx 184$ K [19–21] as compared to perovskites PMN and PLZT; (ii) high symmetry ($Fd3m - O_h^7$) of the paraelectric phase, as in perovskites ($Pm3m - O_h^1$); and (iii) a new type of structural disorder in relaxor systems (orientation disorder of dipoles [22] in contrast to composition disorder in perovskites [1, 3]). Since pyrochlores, as well as perovskites, belong to the centrosymmetric cubic system, familiar models of spin and dipole glasses can be used for analyzing their peculiar dielectric behavior [9–11, 14, 18]. In contrast to perovskites, the relaxor state develops in $Cd_2Nb_2O_7$ simultaneously with ferroelastic (below T_s = 205 K) and ferroelectric (below T_C = 196 K) states, $\varepsilon_{max}(T)$ at frequencies below 13 MHz being observed for $T_{max} < T_C$ [19–22].

2. EXPERIMENTAL TECHNIQUE

Experiments were carried out on ceramic samples possessing a high density (94–96% of the theoretical value) and prepared in accordance with standard technology [20]. Analysis of the samples by X-ray powder diffractometry at room temperature confirmed their pyrochlore-type structure, and no other phases were detected.

Measurements were made on samples having the shape of plates 10 mm in diameter and a thickness of about 2 mm; silver paste electrodes were deposited on the parallel surfaces of the plates. The permittivity $\varepsilon_{his}(T)$ of the samples was measured at a frequency of 1 kHz (in the relaxation dispersion range) with the help of a P5079 capacitance bridge in a weak ac field of $E_{ac} \approx 12$ V/cm and in a dc field of $E_{dc} \approx 0.95$ kV/cm. The dielectric hysteresis loops for $Cd_2Nb_2O_7$, as well as for other relaxors, are extended; as a result, the difference between the coercive field and the polarization saturation field is significant (for a given sample, at 120 K, we have $E_{coer} \sim 1$ kV/cm and $E_{cryt} \sim 2.5$ kV/cm). Consequently, field $E_{dc} \approx 0.95$ kV/cm can change neither the diameter of domains in the frustrated ferroelectric nor the permittivity in the region of T_C [21, 23–25]; it cannot induce a frustrated ferroelectric state in the region of T_f [10, 13], but may affect the interaction between polar microscopic regions [11, 15]. Our measurements were made in the regimes of slow cooling and heating the sample at a rate of 1 K/min; the temperature was stabilized to within ±0.1 K. Sample cooling (ZFC and FC) was always started from 300 K (i.e., at a temperature for which dispersion of $\varepsilon(T)$ is absent [19, 20] and high enough to "erase" all effects produced by previous measurements). At 110 K (i.e., away from T_f), field E_{dc} was applied or removed and the sample was heated (FH/ZFC or ZFH/FC) after holding for about 5 min at this temperature. Such fundamental measurements aimed at testing the glass nature of the phase below T_f [14] have not been made earlier on $Cd_2Nb_2O_7$ in spite of the large number of publications devoted to analysis of $\varepsilon(T)$ in a dc field [21, 23–25].

3. RESULTS AND DISCUSSION

Monotonically diverging $\varepsilon_{his}(T)$ curves and splitting between $\varepsilon_{ZFC}(T)$ and $\varepsilon_{FC}(T)$, as well as between $\varepsilon_{ZFC}(T)$ and $\varepsilon_{FH/ZFC}(T)$, $\varepsilon_{ZFH/FC}(T)$ observed in the temperature

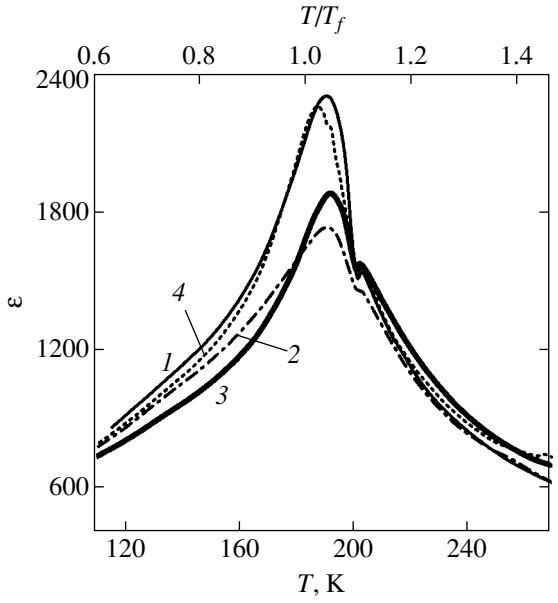

Fig. 1. Temperature dependences of the real part of permittivity $\varepsilon_{his}(T)$ of the $Cd_2Nb_2O_7$ ceramic at a frequency of 1 kHz in a weak electric field E_{dc} = 0.95 kV/cm for samples with different thermal and electrical histories: ZFC (*1*), FH/ZFC (*2*), FC (*3*), and ZFH/FC (*4*).

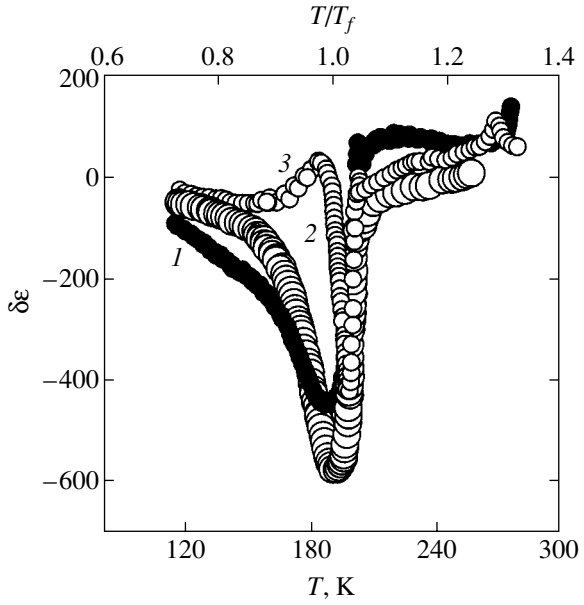

Fig. 2. Temperature dependences of the difference $\delta\varepsilon(T)$ between $\varepsilon_{his}(T)$ and $\varepsilon_{ZFC}(T)$ shown in Fig. 1: FC (*1*), FH/ZFC (*2*), and ZFH/FC (*3*).

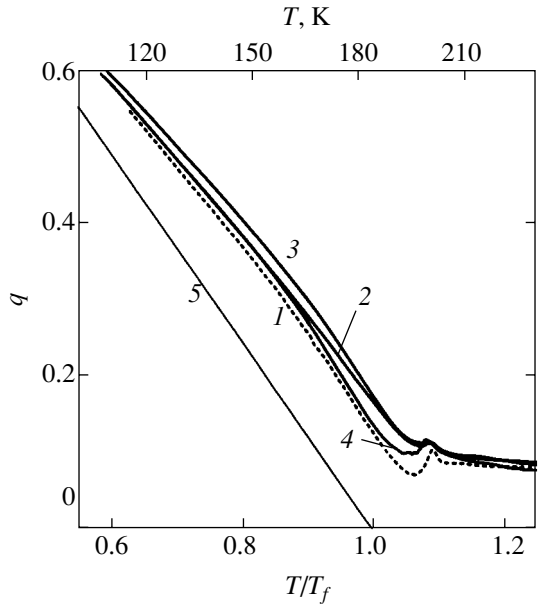

Fig. 3. Temperature dependences of the order parameter $q(T)$ calculated using formula $\varepsilon(T) = C[1 - q(T)]/[T - \Theta(1 - q(T))]$ [18] for $\varepsilon_{his}(T)$ shown in Fig. 1. The values of effective parameters for the samples are as follows: Curie constant $C = 0.67 \times 10^5$ K, the Weiss–Curie temperature Θ = 181 K; ZFC (*1*), FH/ZFC (*2*), FC (*3*), ZFH/FC (*4*); *5* is the solution to the system of equations (2) for $\Delta = 0$, $P = 0$, $E_{dc} = 0$ (spin glasses).

interval 110 K < T < T_f in a weak field (Fig. 1), clearly indicate that the state of the compound is similar to spin and dipole glass states [14]. As $T \longrightarrow T_{max}^{-}$ ($T_{max}^{-} \approx$ 191 K), the divergence between the $\varepsilon_{his}(T)$ curves remains significant and depends on the thermal and electrical histories of the samples. In the region $T > T_{max}$, the divergence becomes gradually smaller and disappears near 270 K.

Writing the relation between field E and polarization $P(E, T)$ induced by it in the form of a power series $P = \varepsilon_0[\varepsilon(0, T)E - \chi_{nl}(T)E^3 + \chi_4(T)E^5 + \ldots]$ (for a cubic structure with a center of symmetry) and considering that the value of $\varepsilon(0, T) \approx \varepsilon_s$ is measured in a weak ac field E_{ac}, while nonlinear effects in the region of T_f are relatively small as compared to $\varepsilon(0, T)$ (for example, $\varepsilon(0, T) \sim 2200$ at T_f; see Fig. 1), the change in the dielectric response ($\varepsilon = 1 + \partial P/\partial E$) induced by field E_{dc} can be represented as a function of nonlinear susceptibility:

$$\begin{aligned}\delta\varepsilon(T) &= \varepsilon_{FC}(E_{dc}, T) - \varepsilon_{ZFC}(0, T)\\ &= -3\chi_{nl}(T)E_{dc}^2,\end{aligned} \quad (1)$$

where ε_0 is the permittivity of vacuum and $\chi_{nl}(T)$ and $\chi_4(T)$ are nonlinearities of the second and fourth order, respectively. It is well known that the nonlinear susceptibility $\chi_{nl}(T)$ acquires an anomaly in the region of T_f upon a transition of the system to the dipole glass state [8, 11], while the normalized nonlinear susceptibility tends to infinity: $\chi_{nl}(T)\varepsilon_s^{-1} \longrightarrow \infty$ [6, 12, 26]. Figure 2 shows for comparison the experimental curves $\delta\varepsilon(T) = [\varepsilon_{his}(T) - \varepsilon_{ZFC}(T)]$ in the case of $\varepsilon_{his} \equiv \varepsilon_{FC}$, which is described by the phenomenological relation (1), as well as for $\varepsilon_{ZFH/FC}$ and $\varepsilon_{FH/ZFC}$. In the interval 110 < T < 205 K, in the FC and FH/ZFC regimes, the value of $\delta\varepsilon(T)$ is negative and passes through a minimum whose position coincides with T_f (curves *1* and *2*). In the ZFH/FC regime, the type of anomaly differs sharply: $\delta\varepsilon(T)$ passes through a minimum at a temperature of 195.8 K = T_f + 11.8 K, reverses its sign at T_f, and passes through a maximum at a temperature of 181.4 K = T_f – 2.6 K (curve *3*).

Formula

$$\varepsilon(T) = C[1 - q(T)]/[T - \Theta(1 - q(T))],$$

describing the dielectric response of relaxors with a glass behavior [7, 18], was used for calculating dependences $q(T)$ for $\varepsilon_{his}(T)$ observed upon variation of experimental conditions (Fig. 3). In contrast to magnetic spin glasses, the order parameter $q(T)$ of $Cd_2Nb_2O_7$ does not vanish above T_f in all cases. At the same time, in contrast to traditional dipole glasses, the $q(T)$ dependence in the region of T_f varies not monotonically, but has a kink. Below 160 K, the order parameter $q(T) \propto T^{-1}$; above 190 K, it is relatively small and exhibits a weak dependence on temperature; the anomaly is absent in the region of transition to the glass state (anomalies in the vicinity of T_s and T_C are associated with a phase transition [22]). Such a behavior of $q(T)$ is typical of dipole glasses and relaxors in the presence of

random interactions and, simultaneously, random fields [9, 14]. Analogously to $\varepsilon_{his}(T)$, a divergence between q_{ZFC} and q_{FC}, $q_{FH/ZFC}$, $q_{ZFH/FC}$ is observed in the temperature range from 110 to 270 K. Below 160 K, the $q_{FH/ZFC}$ and $q_{ZFH/FC}$ curves virtually coincide, while above 160 K, these curves diverge monotonically. In the region of T_f, we have $q_{ZFH/FC} \approx q_{ZFC}$, while $q_{FH/ZFC} \approx q_{FC}$. These features of the behavior of $q_{his}(T)$ indicate that the dynamics of the glass state in a weak dc field depends on the thermal and electrical histories of the samples.

With regard to glass behavior, the free energy of relaxors is determined by two order parameters,

$$P(T) \equiv \langle\langle P_i \rangle\rangle_J, \quad q(T) \equiv \langle\langle P_i P_j \rangle\rangle_J$$

(where P_i and P_j are the local polarizations of clusters); q and P satisfy the system of two equations for an isotropic relaxation system [9, 18],

$$q = \beta^2 J^2 (q + \Delta/J^2)(1 - q)^2 + P^2, \quad (2a)$$

$$P = \beta(1 - q)(J_0 P + E_{dc}), \quad (2b)$$

where J^2 and Δ are the dispersions of random interactions between polar microscopic regions and of random fields, respectively; $\beta = 1/kT$, k being the Boltzmann constant; and $kT_f = (J^2 + \Delta)^{1/2}$. The solutions to Eqs. (2) for $q(T)$ in the case of formation of the glass phase ($J_0 \ll (J^2 + \Delta)^{1/2}$; i.e., $J_0 = 0$, $\Delta \neq 0$, but $\Delta/J^2 \ll 1$, $P = 0$, and $E_{dc} = 0$) show that an increase in parameter Δ/J^2 leads to an increase in $q(T)$ [9, 14, 18]. It was found experimentally that $E_{dc} \neq 0$ also leads to an increase in $q(T)$ [7]. Taking into account these results—as well as the fact that the value of $q_{FH/ZFC}$ for relaxors in the glass phase increases upon the application of a weak field but remains smaller than q_{FC} up to T_f, while the value of $q_{ZFH/FC}$ decreases after the field removal but remains greater than q_{ZFC} (see Fig. 3)—we can assume that the application or removal of field E_{dc} at a low temperature causes a change not only in the value of parameter Δ/J^2, but also in the relative contribution J^2 and Δ of random interactions and random fields. As a result, in the ZFC and FH/ZFC regimes, the behavior of $q(T)$ is mainly determined by random interactions between polar microscopic regions, which limit an increase in $q(T)$ in the presence of random fields Δ and a weak external dc field E_{dc}. In the FC regime, the relative contribution of random fields increases, elevating the value of $q(T)$ for the ZFH/FC regime relative to the ZFC regime both above and below T_f.

4. CONCLUSIONS

The temperature behavior of the dielectric response of the relaxor $Cd_2Nb_2O_7$ at a frequency of 1 kHz (belonging to the relaxation dispersion frequency range) in a weak dc field is investigated for the first time for samples with different thermal and electrical histories with the aim of clarifying the origin and evolution of the relaxor state upon a transition to a phase developing below T_f. Detailed analysis of the temperature dependences $\varepsilon(E)$, $\delta\varepsilon(T) = [\varepsilon_{his}(T) - \varepsilon_{ZFC}(T)]$, and $q(T)$ has been carried out taking into account modern concepts of peculiarities of spin and dipole glasses as well as relaxor with the glass behavior. It is shown that the behavior of a relaxor in the phase developing below T_f is determined by random interactions (contribution of which predominates) as well as random fields (the relative contribution of the latter increases in a weak dc field). The result obtained is of fundamental importance since it forms the basis for understanding the behavior of a relaxor in the temperature range below T_f in a strong dc field and the reasons for the formation of the state of a (frustrated or normal) ferroelectric in a structurally disordered system.

ACKNOWLEDGMENTS

The author is grateful to P.P. Syrnikov, who synthesized the $Cd_2Nb_2O_7$ ceramic, and to A.O. Lebedev, who carried out X-ray structural analysis of the samples.

This study was supported financially by the Russian Foundation for Basic Research (project no. 00-02-16900).

REFERENCES

1. G. A. Smolensky, J. Phys. Soc. Jpn. **28**, 26 (1970).
2. B. E. Vugmeister and H. Rabitz, Phys. Rev. B **57**, 7581 (1998).
3. L. E. Cross, Ferroelectrics **76**, 241 (1987); Ferroelectrics **151**, 305 (1994).
4. H. Gui, B. Gu, and X. Zhang, Phys. Rev. B **52**, 3135 (1995).
5. V. Westphal, W. Kleemann, and D. Glinchuk, Phys. Rev. Lett. **68**, 847 (1992).
6. A. K. Tagantsev and A. E. Glazounov, Phys. Rev. B **57**, 18 (1998); Appl. Phys. Lett. **74**, 1910 (1999).
7. D. Viehland, S. J. Jang, L. E. Cross, and M. Wuttig, Phys. Rev. B **46**, 8003 (1992); J. Appl. Phys. **68**, 2916 (1990).
8. E. V. Colla, E. Yu. Koroleva, N. M. Okuneva, and S. B. Vakhrushev, J. Phys.: Condens. Matter **4**, 3671 (1992); Ferroelectrics **184**, 209 (1996).
9. R. Pirc and R. Blinc, Phys. Rev. B **60**, 13470 (1999).
10. V. Bobnar, Z. Kutnjak, R. Pirc, *et al.*, Phys. Rev. Lett. **84**, 5892 (2000).
11. D. Viehland, S. J. Jang, and L. E. Cross, J. Appl. Phys. **69**, 414 (1991).
12. E. I. Fisher and J. A. Hertz, *Spin Glasses* (Cambridge Univ. Press, Cambridge, 1991).
13. R. Sommer, N. K. Yushin, and J. J. Van der Klink, Phys. Rev. B **48**, 13230 (1993).
14. K. Binder and A. P. Young, Rev. Mod. Phys. **58**, 801 (1986).

15. D. Viehland, J. F. Li, S. J. Jang, *et al.*, Phys. Rev. B **46**, 8013 (1992).

16. Z. Kutnjak, C. Filipic, and A. Levstik, J. Eur. Ceram. Soc. **21**, 1313 (2001).

17. A. Levstik, Z. Kutnjak, C. Filipic, and R. Pirc, Phys. Rev. B **57**, 11204 (1998).

18. D. Sherrington and S. Kirkpatrick, Phys. Rev. Lett. **35**, 1792 (1975).

19. N. N. Kolpakova, P. Czarnecki, W. Nawrocik, *et al.*, Zh. Éksp. Teor. Fiz. **121**, 462 (2002) [JETP **94**, 395 (2002)].

20. N. N. Kolpakova, P. P. Syrnikov, A. O. Lebedev, *et al.*, J. Appl. Phys. **90**, 6332 (2001).

21. Ch. Ang, L. E. Cross, R. Guo, and A. S. Bhalla, Appl. Phys. Lett. **77**, 732 (2000).

22. N. N. Kolpakova, M. Wiesner, G. Kugel, and P. Bourson, Ferroelectrics **201**, 107 (1997); Ferroelectrics **190**, 179 (1997).

23. V. A. Isupov and G. I. Tarasova, Fiz. Tverd. Tela (Leningrad) **25**, 1013 (1983) [Sov. Phys. Solid State **25**, 584 (1983)]; Fiz. Tverd. Tela (Leningrad) **25**, 1018 (1983) [Sov. Phys. Solid State **25**, 587 (1983)].

24. Z. G. Ye, N. N. Kolpakova, J.-P. Rivera, and H. Schmid, Ferroelectrics **124**, 275 (1991).

25. L. S. Kamzina, F. M. Salaev, N. N. Krainik, *et al.*, Fiz. Tverd. Tela (Leningrad) **25**, 2846 (1983) [Sov. Phys. Solid State **25**, 1645 (1983)].

26. J. Hemberger, H. Ries, A. Loidl, and R. Böhmer, Phys. Rev. Lett. **76**, 2330 (1996).

Translated by N. Wadhwa

Journal of Experimental and Theoretical Physics, Vol. 96, No. 3, 2003, pp. 543–554.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 123, No. 3, 2003, pp. 612–624.

SOLIDS
Electronic Properties

Cascade Avalanche Production of Electron–Hole Pairs in Type II Quantum Wells

E. Yu. Perlin*, A. V. Ivanov, and R. S. Levitskiĭ
Vavilov State Optical Institute, Birzhevaya liniya 12, St. Petersburg, 199034 Russia
**e-mail: smperlin@online.ru*
Received June 10, 2002

Abstract—It is shown that the photon avalanche mechanism can be used for producing nonequilibrium electron–hole pairs by low-intensity IR light with a photon energy smaller than the energy gap of a semiconductor by a factor of 3–5. A type II heterostructure with deep quantum wells is proposed to be employed for this purpose. In the model under investigation, the photon avalanche effect is due to a combination of a cascade of one- and two-photon transitions and Auger-type transitions. *© 2003 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

The processes of nonequilibrium electron–hole pair (EHP) production in semiconductors and insulators by intense low-frequency light have been studied in detail (see, for example, [1]). Such processes include interband tunneling in a strong electromagnetic field, multiphoton interband transitions, cascade transitions through local levels in the energy gap, and avalanche pair production. In the case of the avalanche mechanism, free charge carriers (electrons or holes) must possess a kinetic energy sufficient for generating a new EHP. Charge carriers can acquire such an energy only in the field of a high-intensity electromagnetic wave as a result of a cascade of indirect intraband transitions whose probabilities decrease rapidly with the light frequency ω. Multiphoton Auger-type processes [2, 3] occupy an intermediate position between the "ordinary" multiphoton and avalanche EHP production. In this case, the major part of energy required for the EHP production comes from the absorption of several photons, and only the remaining part is due to the kinetic energy of a free electron or a hole heated by light, which is smaller than the energy gap width E_g. For all the processes listed above, the production of an appreciable number of EHPs with an energy exceeding $3\hbar\omega$ is possible only for a light intensity $j \gtrsim 10^9$ W/cm^2. This also applies to multiphoton EHP production in systems with quantum wells (QW), although the production rate in such systems decreases upon an increase in the number of photons involved in the elementary act of transition more slowly than in bulk materials [4]. Here, we disregard radiation in the far IR range, for which the probabilities of transitions involving a large number of photons are high in some cases [5–7], and assume that the light frequency is high as compared to the vibrational excitation energies of the crystal.

In this study, we propose a new efficient mechanism for producing a large number of nonequilibrium EHPs with an excitation energy of $(3–5)\hbar\omega$ for moderate intensities $j \sim 10^4$–10^6 W/cm^2 of long-wave light. Such a mechanism can be realized using a type II semiconducting heterostructure with deep QWs. The model considered below includes a cascade of one- or two-photon transitions combined with Auger-type transitions, which lead to the photon avalanche effect.

Starting from the work by Chivian *et al.* [8], the photon avalanche has become one of the most effective methods of exciting short-wave luminescence by long-wave pumping. The photon avalanche effect has been studied so far for systems of rare earth impurity ions (see, for example, [8–14]). We will describe the photon avalanche mechanism using a simple three-level system as an example. We denote by 1, 2, and 3 the ground state and two excited states of rare earth ions. State 2 is usually regarded as a metastable state. In the absence of pumping, only state 1 is filled. Let the pumping frequency ω be close to the frequency ω_{32} of transition between states 2 and 3, but far from the resonance at the frequency of transition between states 1 and 2 ($\omega_{21} < \omega_{32}$). For low light intensities j, nothing happens in the system since only optical $2 \longrightarrow 3$ transitions could occur between the unfilled states. The situation changes when intensity j becomes higher. Let us suppose that one of the rare earth ions passes to state 2 in some way or other. Having absorbed a quantum $\hbar\omega$, the ion is transformed into state 3. Due to the cross-relaxation mechanism, the ion returns from state 3 to state 2, but one of neighboring rare earth ions is excited in this case from the ground state 1 to state 2. Thus, two ions are in state 2 now. In turn, each of these ions can participate in similar processes, and so on. As a result, a large number of electrons belonging to different ions are accumulated in state 2. This leads to a strong absorption at $2 \longrightarrow 3$ transitions and, hence, to a high occupancy of level 3. In this case, luminescence at a frequency $\Omega > \omega$ is possible on the $3 \longrightarrow 1$ transition. The key role in the

photon avalanche effect is played by the competition between relaxation loss of electrons in state 2 and accumulation of electrons in this state owing to light absorption and cross-relaxation transfer of excitation. The photon avalanche effect is characterized by (a) a clearly manifested threshold nature of the phenomenon (the populations of excited states and light absorption at frequency ω increase jumpwise for $j \approx j_{th}$) and (b) a sharp increase in the setting time τ_{eq} for a quasi-equilibrium electron distribution in the range of threshold light intensities j_{th}. The values of j_{th} decrease upon an increase in the concentration of the rare earth ions. The threshold nature of the photon avalanche effect has made it possible to analyze this effect in terms of the Landau theory of second-order phase transitions [11].

Various schemes of avalanche up-conversion in impurity systems were proposed and implemented during the last decade (see [11, 12] for detailed references). In particular, an effective avalanche cascade scheme of up-conversion in the eight-level model of Tm^{3+} ions in YLF treated in [14] makes it possible to generate radiation at a wavelength of $\lambda \sim 0.29$ μm for pumping with $\lambda = 1.11$ or 0.649 μm.

The typical quasi-equilibrium times τ_{eq} in the electron system of rare earth ions for the photon avalanche effect range from 1 to 100 ms for the threshold pumping energies $E_{sw} \sim 0.1$–10 μJ/μm^2 required for switching the avalanche mechanism. Such a slow course of the photon avalanche effect in systems of rare earth ions naturally limits the possibilities of practical application of this phenomenon in optoelectronics. In this connection, a semiconducting scheme with doped QWs was proposed in [15, 16]. In this scheme, switching to the avalanche up-conversion regime can be executed over a much shorter time with smaller energy expenditures. The advantages of such a scheme are due to large values of oscillator strength for transitions between size quantization subbands, on the one hand, and short relaxation times in the electron system of QWs on the other hand. The role of interionic cross-relaxation transitions in a doped QW is played by intersubband Auger-type transitions in the well. In the system considered in [15, 16], peculiar features of the photon avalanche effect, such as the threshold nature of the process and a strong increase in times τ_{eq} for pumping intensities close to the threshold values, are preserved. The photoinduced dynamics in a doped QW is described in [15, 16] by a system of nonlinear balance equations for their populations of the three lower subbands 1, 2, and 3 (resonant optical transitions occur between subbands 2 and 3). Analysis of these equations proved that the system has two stationary points, viz., a stable node and an unstable node. For the bifurcation value of light intensity $j \approx j_{th}$, these two stationary points degenerate into one point. If the probability W_A of an Auger process in which, as a result of the collision of an electron from subband 3 with an electron from subband 1, both electrons pass to subband 2 can be represented in the form $W_A \approx \gamma_A n_1 n_3$, then we have [16]

$$j_{th} \approx \frac{W_{21}(n_0\gamma_A + W_{31} + W_{32})}{\sigma_{23}(n_0\gamma_A - W_{21} - W_{32})}, \qquad (1)$$

where n_0 is the total (two-dimensional) concentration of charge carriers in a QW, W_{ij} are the rates of relaxation transitions between the ith and jth subbands, and σ_{ij} are the cross sections of light absorption in a transition between the ith and jth subbands. Formula (1) is valid for $n_0\gamma_A > W_{ij}$, $\sigma_{12} = 0$. For finite (but small, as compared to σ_{23}) values of σ_{12}, the (formal) bifurcation intensity j_{th} is complex-valued: $\tilde{j}_{th} = j'_{th} + ij''_{th}$. However, $j''_{th} \ll j'_{th}$ for values σ_{ij}, W_{ij}, and n_0 of current interest. Consequently, if the value of j increases, then the quasi-equilibrium values of n_2 and n_3 sharply increase, while the value of n_1 decreases in the vicinity of j_{th}. With increasing σ_{12}, the change in the population densities n_i near the threshold becomes smoother. The following expression for the time τ_{eq} was derived in [16] for $j = j_{th}$:

$$\tau_{eq} \sim \frac{1}{2W_{21}}\sqrt{\frac{n_0\gamma_A\sigma_{23}}{\sigma_{12}(W_{21} + 2W_{31} + W_{32})}}. \qquad (2)$$

It can be seen that the value of τ_{eq} increases upon a decrease in the rate of arrival of electrons to state 2, in which their number may increase in the avalanche manner.

The scheme of the photon avalanche effect in a system with type II QWs proposed (see Section 2) differs substantially from that considered in [15, 16] and makes it possible to attain luminescence at a wavelength exceeding that of the exciting light by a factor of 3–5 for moderate pumping intensities. The switching energy E_{sw} in this case is on the order of 1–10 pJ/μm^2; i.e., it is four to five orders of magnitude lower than in the system of rare earth ions. It should be noted that a rigorous treatment of the proposed scheme involves considerable difficulties. Nevertheless, we can expect that the extremely simplified approach adopted here will enable us to obtain a qualitative pattern of the process, whose main features will remain unchanged in a more detailed analysis.

2. SCHEME OF CASCADE AVALANCHE UP-CONVERSION IN TYPE II QWs

Let us consider a heterostructure consisting of components A and B and characterized by a type II band diagram (Fig. 1). A region of width $2a$ occupied by component A of the heterostructure (region A) is a rectangular well for electrons of depth U_c and a rectangular barrier for holes of height U_v. We assume that the well depth U_c is quite large (about 1.5–2 eV). Component B

is located outside this region (region *B*). We assume that the well for electrons contains three size-quantization subbands labeled 1, 2, and 3 in increasing order in energy. The energy gaps $\hbar\omega_{ij}$ between subbands are regarded as larger than temperature *T*. We also assume that $\omega_{32} > \omega_{21}$, and the frequency of incident light is $\omega \approx \omega_{32}$ ($\hbar\omega \sim 0.5$–0.8 eV). In contrast to [15, 16], we assume that, under equilibrium conditions, the electron states in the QW are empty (we will also consider the case when subband 1 contains a small number of electrons; see Section 7). If light intensities *j* are low, only a small number of nonequilibrium electrons appear in subband 1 due to one- or two-photon indirect transitions (in the **r** space) from states in the valence band *v* in region *B*. Since the frequency of light is higher than the frequency ω_{21} of direct transitions between subbands 1 and 2, only weak indirect transitions may occur (in the $\mathbf{k}_\perp$ space) between these subbands. Rapid resonant photoexcited transitions could occur between subbands 2 and 3, but these subbands remain virtually empty for small values of *j*. The situation changes dramatically for high values of *j*. As a result of the two-step cascade of weak transitions $v \longrightarrow 1$ and $1 \longrightarrow 2$, a certain number of electrons are still in states with $\mathbf{k}_\perp \neq 0$ of subband 2. These electrons rapidly relax to the bottom of the subband. After this, they may either "drop" to subband 1 due to relaxation involving phonons (see Section 5), or absorb a photon $\hbar\omega$ and pass to subband 3. The oscillator strength for allowed resonant transitions $2 \longrightarrow 3$ is very high. Electrons may drop from subband 3 to subbands 2 and 1 owing to relaxation with the participation of phonons, but the Auger relaxation mechanism also proves to be quite efficient [15]: due to the Coulomb interaction between an electron in subband 3 and an electron in subband 1, both electrons pass to subband 2. Each of these electrons may lead to the emergence of two electrons in subband 2 in the same way, and so on. For high intensities *j* of light, the number of electrons arriving in subband 2 via this mechanism is larger than the number of electrons leaving to subband 1. In this case, the populations in subbands 2 and 3, as well as in the states of the continuous spectrum of the conduction band *c*, increase in an avalanche manner. It should be emphasized that the core of the above-described mechanism of photon avalanche triggering (mechanism I) is strong absorption of light during transitions between excited states of the system, which is combined with the $31 \longrightarrow 22$ Auger transition, leading to multiplication of electrons in subband 2.

In contrast to the model of the photon avalanche effect in a doped QW [15, 16], another Auger process emerges and plays a significant role in the photon avalanche effect in a type II structure: an electron from subband 3 interacts with an electron from the valence band in region *B*, and both electrons enter subband 1 (see Section 4). In this case, the total number of electrons in the QW increases, which lowers the threshold light intensity j_{th}. This Auger process combined with mechanism I ensuring, among other things, a transition of electrons from subband 1 to subband 3 forms the core of another avalanche mechanism (mechanism II).

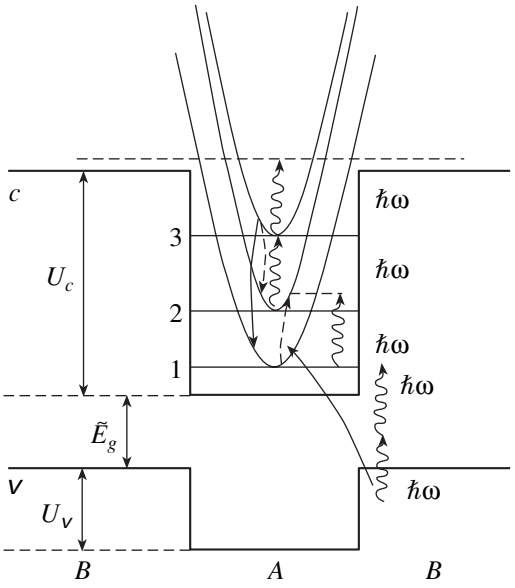

Fig. 1. Diagram of transitions in a type II QW in the case of cascade avalanche up-conversion. Vertical undulated lines with arrows mark optical transitions, solid lines with arrows indicate Auger transitions $3v \longrightarrow 11$, and dashed lines with arrows show Auger transitions $31 \longrightarrow 22$.

Our calculations show that, in the case of single-photon pumping of initiating electrons to subband 1 for cascade avalanche pair production, mechanism I alone would be sufficient in general. However, mechanism II plays an important part in this case, lowering the threshold value of intensity to a considerable extent. In the case of two-photon pumping (in a reasonable range of parameters), mechanism I should be necessarily supplemented with mechanism II for initiating an avalanche. Mechanism II alone (in the case of pure cascade excitation in the $1 \longrightarrow 3$ channel) cannot trigger the photon avalanche effect.

In this study, we also consider explicitly the photo-induced transitions from subband 3 to the states of the continuous spectrum of the conduction band, electron trapping from the continuum to the energy levels in the QW, and recombination of nonequilibrium photoexcited electrons and holes.

3. TRANSITIONS BETWEEN ELECTRON STATES IN A QW AND THE STATES OF CONTINUOUS SPECTRUM IN THE VALENCE BAND AND CONDUCTION BAND

In order to roughly estimate the probabilities of optical transitions between the states of the continuum of the valence band and conduction band and the states of

size-quantization subbands in a well for electrons in region *A*, we use the simplest model with one-band wave functions:

$$\psi_n(\mathbf{r}) = \sum_{\boldsymbol{\mu}, \mathbf{R}_i} e^{i\mathbf{k}_\perp \mathbf{R}_i} \beta_n(R_{i\parallel}) e^{i\boldsymbol{\mu}(\mathbf{r}-\mathbf{R}_i)} u_{n,\boldsymbol{\mu}}(\mathbf{r}), \tag{3}$$

where $u_{n,\boldsymbol{\mu}}(\mathbf{r})$ are the Bloch amplitudes for the nth band. The dimensionless coordinates and wave vectors appearing in the formulas given below are measured in units of a and a^{-1}, respectively, where a is the half-width of the QW. The enveloping even (+) and odd (–) wave functions in the valence band, where region *A* is a barrier for holes having a height U_v, and the enveloping wave functions for electron states in the QW of depth U_c have the form

$$\beta_v^{(\pm)}(z) = \frac{1}{2L^{1/2}} \times \begin{cases} e^{-ik_{v\parallel}(z+1)} + e^{ik_{v\parallel}(z+1)}\gamma^{(\pm)}, & z < -1, \\ \varphi^{(\pm)}/\gamma_+^{(\pm)}, & -1 \le z \le 1, \\ e^{ik_{v\parallel}(z-1)} + e^{-ik_{v\parallel}(z-1)}\gamma^{(\pm)}, & z > 1, \end{cases}$$

$$\beta_{k_i}^{(\pm)}(z) = \sqrt{\frac{\eta_i}{a(1+\eta_i)}} \begin{cases} e^{\eta_i(z+1)} \begin{cases} \cos k_i \\ -\sin k_i, \end{cases} & z < -1, \\ \begin{cases} \cos k_i z \\ \sin k_i z, \end{cases} & -1 \le z \le 1, \\ e^{-\eta_i(z-1)} \begin{cases} \cos k_i \\ \sin k_i, \end{cases} & z > 1, \end{cases} \tag{4}$$

where

$$\varphi^{(\pm)}(z) = \begin{cases} -\sinh \kappa z \\ \cosh \kappa z, \end{cases} \quad \gamma^{(\pm)} = \frac{\gamma_-^{(\pm)}}{\gamma_+^{(\pm)}},$$

$$\gamma_\pm^{(\pm)} = \begin{cases} ik_{v\parallel}\cosh\kappa \pm \kappa \sinh\kappa, \\ ik_{v\parallel}\sinh\kappa \pm \kappa\cosh\kappa, \end{cases}$$

$$k_{v\parallel} = \frac{a}{\hbar}\sqrt{2m_v E_{v\parallel}}, \quad \kappa = \frac{a}{\hbar}\sqrt{2m_v(U_v - E_{v\parallel})}, \tag{5}$$

$$\eta_i = \sqrt{k_0^2 - k_i^2}, \quad k_0 = \hbar^{-1}\sqrt{2m_c a^2 U_c}.$$

For even and odd states in the well, quantities k_i are defined as the roots of transcendental equations

$$\tan\kappa_i = \eta_i/k_i, \quad \tan k_i = -k_i/\eta_i. \tag{6}$$

For overlap integrals of the wave functions of the valence band and the wave functions of the first and second levels in the well, we have, respectively,

$$I_1^{(+)} = \nu_1^{(+)}[\xi_1 \cos k_1(\eta_1 \cosh\kappa + \kappa\sinh\kappa) + \zeta_1(k_1\cosh\kappa\sin k_1 + \kappa\sinh\kappa\cos k_1)], \tag{7}$$

$$I_2^{(-)} = -\nu_2^{(-)}[\xi_2 \sin k_2(\kappa\cosh\kappa + \eta_2\sinh\kappa) + \zeta_2(\kappa\cosh\kappa\sin k_2 - k_2\cos k_2\sinh\kappa)], \tag{8}$$

where

$$\nu_i^{(\pm)} = 2i\sqrt{\frac{a\eta_i}{L(1+\eta_i)}}\frac{k_{v\parallel}}{\xi_i\zeta_i\gamma_-^{(\pm)}}, \quad \xi_i = k_i^2 + \kappa^2, \quad \zeta_i = k_{v\parallel}^2 + \eta_i^2, \tag{9}$$

and L is the normalization length of the sample along the axis of the nanostructure growth.

Expressions for the probabilities of one- and two-photon transitions between the valence band and subband 1 have the form

$$\frac{W_{v1}^{(1,2)}}{S} = \frac{1}{2\pi^2\hbar}\int d^2k_\perp \int dk_{v\parallel} \left|\tilde{M}_1^{(1,2)}(k_{v\parallel})\right|^2 \times \delta\left(\frac{\hbar^2 k_\perp^2}{2m_r} + \frac{\hbar^2 k_{v\parallel}^2}{2m_v} - a^2\Delta_v^{(1,2)}\right), \tag{10}$$

$$\Delta_v^{(p)} = p\hbar\omega - E_{01} - \tilde{E}_g,$$

$$\tilde{M}_1^{(1,2)}(k_{v\parallel}) = \sqrt{\frac{L}{a}}M_1^{(1,2)}(k_{v\parallel}), \tag{11}$$

$$M_1^{(1)}(k_{v\parallel}) = i\frac{eF_\omega}{m\omega}p_{vc}I_1^{(+)},$$

$$M_1^{(2)}(k_{v\parallel}) = \left(\frac{eF_\omega}{m\omega}\right)^2\frac{p_{vc}}{a}\left\{\frac{mk_{v\parallel}}{m_v\omega}I_1^{(+)} + \frac{2i\hbar\cos k_1\sin k_2}{E_{01}-E_{02}-\hbar\omega}\sqrt{\frac{\eta_1\eta_2}{(1+\eta_1)(1+\eta_2)}} \times \left[\frac{\eta_1}{\eta_1+\eta_2} + \frac{k_1(k_1\tan k_1\cot k_2 - k_1)}{k_1^2 - k_2^2}\right]I_2^{(-)}\right\}, \tag{12}$$

where $m_r^{-1} = m_c^{-1} + m_v^{-1}$, $E_{0i} = \hbar^2 k_i^2/2m_c$ is the energy of the ith subband for $\mathbf{k}_\perp \neq 0$; S is the sample cross section, perpendicular to the growth axis; m is the mass of a free electron; p_{vc} is the conventional interband momentum operator; $\tilde{E}_g$ is the energy gap between the top of the valence band in region *B* and the bottom of

the QW in region A; and F_ω is the amplitude of the field of an electromagnetic wave whose polarization vector is directed along the growth axis. Here and below, the energy is measured from the bottom of the QW in region A.

Integrating with respect to $d^2k_\perp$ with the help of the δ function, we obtain

$$\frac{W_{v1}^{(1,2)}}{S} = \frac{2m_r}{\pi^2\hbar^3}\int\limits_0^{\frac{a}{\hbar}\sqrt{2m_v\Delta_v}} dk_{v\parallel}\left|\tilde{M}_1(k_{v\parallel})\right|^2. \qquad (13)$$

Figure 2 shows the dependence of $\sigma_{v1}^{(p)} = W_{v1}^{(p)}/(Sj^p)$ on Δ_v. Here and below, we use the following values of parameters in our calculations: $U_c = 1.74$ eV, $U_v = 0.1$ eV, $a = 3$ nm, $m_v = 0.5m$, and $m_c = 0.04m$. In calculating $\sigma_{v1}^{(2)}$, we used $\tilde{E}_g = 0.9$ eV. For a number n_w of QWs per unit length along the growth axis of a nanostructure approximately equal to 5×10^5 cm^{-1}, the single-photon absorption coefficient at the $v \longrightarrow 1$ transitions amounts to 10^2–10^3 cm^{-1}.

Let us now write the expression for the probability of an optical transition between a state at the bottom of subband 3 in the QW and a state with energy $E_c = \hbar^2k_c^2/2m_ca^2$ in the continuum. Omitting computational details, we can write

$$W_{3c} = \frac{16\pi m_c e^2 a^4}{\hbar^3 c\varepsilon k_c} j\left|A_3A_k\left[j_1(k_c-k_3)+j_1(k_c+k_3)\right.\right. \qquad (14)$$
$$\left.\left. -2\cos k_3\cos(\gamma_c-\varphi-\psi)\frac{\sqrt{\gamma_c^2+(1+\eta_3)^2}}{\gamma_c^2+\eta_3^2}\right]\right|^2,$$

where c is the velocity of light in vacuum; ε is the permittivity;

$$A_3 = [(k_3+\sin k_3\cos k_3)/k_3+(\cos^2 k_3)/\eta_3]^{-1/2},$$

$$A_k = [\sin^2 k_c + k_c^2/(k_c^2-k_0^2\cos^2 k_c)]^{-1/2},$$

$$\gamma_c = \sqrt{k_c^2-k_0^2} \equiv \sqrt{2m_c\Delta_c}/\hbar, \quad \Delta_c \equiv E_c - U_c, \qquad (15)$$

$$\varphi = \arctan[k_c\tan\gamma_c - \gamma_c\tan k_c)/(k_c+\gamma_c\tan\gamma_c\tan k_c)]$$

$$= \arctan\{\gamma_c^{-1}[-1+\eta_3(1+2(\gamma_c^2+2\eta_3+\eta_3^2)^{-1})]\},$$

$$\psi = \arctan\{\gamma_c^{-1}[-1+\eta_3(1+2(\gamma_c^2+2\eta_3+\eta_3^2)^{-1})]\};$$

and $j_1(z) = (\sin z)/z^2 - (\cos z)/z$ is a spherical Bessel function of the first kind. The dispersion of probability W_{3c} is given in Fig. 3. Generally speaking, the energy dependence of W_{3c} described by formulas (14) and (15) is nonmonotonic and is characterized by sharp peaks associated with virtual states in the continuum. However, these peaks are exhibited at energies much higher than the values of our interest.

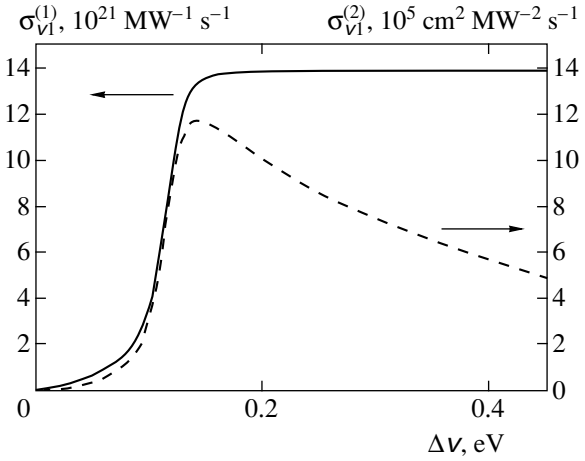

Fig. 2. Cross sections of one- and two-photon absorption via transitions between the valence band in region B and the lower subband in the QW.

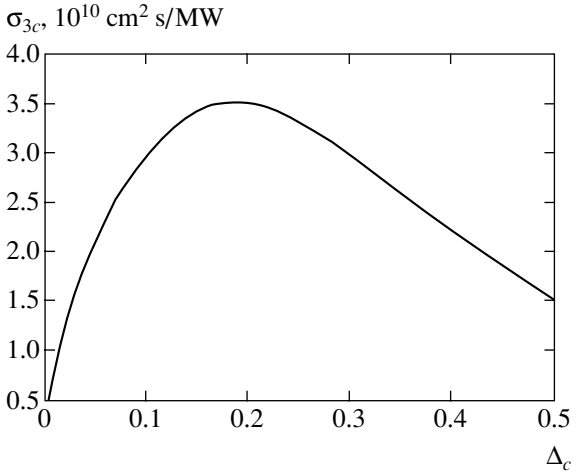

Fig. 3. Dependence of the light absorption cross section σ_{3c} for the transition $3 \longrightarrow c$ on $\Delta_c = E_{3c} + \hbar\omega - U_c$.

4. PROBABILITIES OF AUGER TRANSITIONS

Let us consider Auger transitions of the following type: an electron with a 2D wave vector $\mathbf{k}_{3\perp}$ from subband 3 in the QW interacts with an electron having wave vector $\mathbf{k}_v = (\mathbf{k}_{v\perp}, k_{v\parallel})$ in the valence band in region B. As a result, both electrons pass to subband 1 of the QW to occupy states with wave vectors $\mathbf{k}_{11\perp}$ and $\mathbf{k}_{12\perp}$. The momentum transferred in this case in the plane perpendicular to the growth axis of the nanostructure is given by

$$\mathbf{q}_0 = \mathbf{k}_{3\perp} - \mathbf{k}_{11\perp} = \mathbf{k}_{12\perp} - \mathbf{k}_{v\perp}. \qquad (16)$$

Representing the Coulomb interaction in the form of an expansion into a 3D Fourier series and carrying out standard transformations, we obtain the following expression for a direct matrix element constructed on functions (3)–(6):

$$M^{(d)}_{\mathbf{k}_{3\perp},\mathbf{k}_{v\perp},\mathbf{k}_{3\perp}-\mathbf{q}_0,\mathbf{k}_{v\perp}+\mathbf{q}_0} = \frac{4\pi e^2}{a\varepsilon} \times \int dq_\parallel \iint dz_1 dz_2 \frac{\exp\{iq_\parallel(z_1-z_2)\}}{q_0^2+q_\parallel^2} \times \beta_3^*(z_1)\beta_1(z_1)\beta_v^*(z_2)\beta_1(z_2)\Theta^{(cc)}_{-\mathbf{q}}\Theta^{(vc)}_{\mathbf{q}}, \quad (17)$$

where

$$\Theta^{(mn)}_{\mathbf{q}_0} = \int d\mathbf{r} u^*_{m\mathbf{K}}(\mathbf{r}) u_{n\mathbf{K}+\mathbf{q}_0}(\mathbf{r})$$

are the overlap integrals of Bloch amplitudes for the *m*th and *n*th bands, which are assumed, for the sake of simplicity, to be functions of only $\mathbf{q}_0$ and not of $\mathbf{K}$. In turn, we have

$$\Theta^{(cc)}_{\mathbf{q}} \approx 1, \quad \Theta^{(vc)}_{\mathbf{q}} \approx \boldsymbol{\chi}\mathbf{q}_0, \quad \boldsymbol{\chi} \sim \frac{\hbar}{m}\frac{\mathbf{p}_{cv}}{E_g}, \quad (18)$$

where

$$\boldsymbol{\pi} = \mathbf{p} + \frac{\hbar}{4mc^2}[\boldsymbol{\sigma}\nabla V],$$

$\boldsymbol{\sigma}$ is a Pauli spin matrix, V is the periodic lattice potential, and $\mathbf{p}$ is the momentum operator. Integrating in Eq. (17) with respect to $dq_\parallel$, we obtain

$$M^{(d)}_{\mathbf{k}_{3\perp},\mathbf{k}_{v\perp},\mathbf{k}_{3\perp}-\mathbf{q}_0,\mathbf{k}_{v\perp}+\mathbf{q}_0}(k_{v\parallel}) = \frac{4\pi^2 e^2}{\varepsilon a^2}\chi_\parallel(i\tilde{I}_1 + \sqrt{2}\cos\theta_{\mathbf{q}_0\boldsymbol{\pi}_\perp}\tilde{I}_2) \equiv \frac{4\pi^2 e^2\chi}{S\varepsilon\sqrt{L}}\tilde{M}^{(d)}_{\mathbf{k}_{3\perp},\mathbf{k}_{v\perp},\mathbf{k}_{3\perp}-\mathbf{q}_0,\mathbf{k}_{v\perp}+\mathbf{q}_0}(k_{v\parallel}), \quad (19)$$

where

$$\tilde{I}_1 = \iint dz_1 dz_2 \exp(-q_0|z_1-z_2|)\,\mathrm{sgn}(z_1-z_2) \times \beta_3^*(z_1)\beta_1(z_1)\beta^{(+)*}_{vk_{v\parallel}}(z_2)\beta_1(z_2), \quad (20)$$

$$\tilde{I}_2 = \iint dz_1 dz_2 \exp(-q_0|z_1-z_2|) \times \beta_3^*(z_1)\beta_1(z_1)\beta^{(-)*}_{vk_{v\parallel}}(z_2)\beta_1(z_2). \quad (21)$$

The expression for the transition probability can be written in the form

$$W^{\mathrm{aug}}_{3v,11} = \frac{e^4\chi^2}{(2\pi)^2\hbar\varepsilon^2 a^7}\int d^3k_v \int f(k_{3\perp})d^2k_{3\perp}\int d^2q_0 \times |\tilde{M}^{(d)}_{\mathbf{k}_{3\perp},\mathbf{k}_{v\perp},\mathbf{k}_{3\perp}-\mathbf{q}_0,\mathbf{k}_{v\perp}+\mathbf{q}_0}(k_{v\parallel}) + \tilde{M}^{(e)}_{\mathbf{k}_{3\perp},\mathbf{k}_{v\perp},\mathbf{k}_{3\perp}-\mathbf{q}_0,\mathbf{k}_{v\perp}+\mathbf{q}_0}(k_{v\parallel})|^2 \times \delta\Big[\Delta_v a^{-2} - \frac{\hbar^2 k_{v\parallel}^2}{2m_v} + \frac{\hbar^2}{m_c}(\mathbf{k}_{3\perp}-\mathbf{k}_{v\perp})\cdot\mathbf{q}_0 - \frac{\hbar^2 q_0^2}{m_c} - \frac{\hbar^2 k_{v\perp}^2}{2m_r}\Big], \quad (22)$$

where $\Delta_v = (E_{03} - 2E_{01} - \tilde{E}_g)$ and $f(k_{3\perp})$ is the electron distribution function for subband 3. Superscripts *d* and *e* mark the direct and exchange matrix elements of the transition. In order to simplify computations of the sevenfold multiple integral in the right-hand side of Eq. (22), we can use the following considerations. Under the assumption that the rate of the intrasubband relaxation of electrons in subband 3 is higher than that of other relaxation processes, we assume that all electrons in subband 3 are near its bottom and we can set $\mathbf{k}_{3\perp} = 0$ in the integrand, so that $\int f(k_{3\perp})d^2k_{3\perp} = (2\pi)^2 n_3 a^2$.

The integration over the angles between $\mathbf{k}_{3\perp}$ and $\mathbf{q}_0$ is eliminated with the help of the δ function. Omitting intermediate calculations, we can write

$$\frac{1}{\tau_{\mathrm{aug}}} = \frac{W^{\mathrm{aug}}_{3v,11}}{Sn_3} = \frac{e^4 m_c \chi_\parallel^2}{2\hbar^3 a^3 \varepsilon^2}\int_0^{k^{(m)}_{v\perp}} dk_{v\perp} \times \int_0^{k^{(m)}_{v\parallel}} dk_{v\parallel}\int_{\vartheta_-}^{\vartheta_+}\frac{dq_0^2\left|\tilde{M}^{(d)}_{q_0}(k_{v\parallel}) + \tilde{M}^{(e)}_{q_0}(k_{v\parallel})\right|^2}{\sqrt{(q_0^2-\vartheta_-)(\vartheta_+-q_0^2)}}, \quad (23)$$

$$\vartheta_\pm = \frac{1}{2}(k_{v\perp}\pm\rho),$$

$$\rho = \sqrt{k_{v\perp}^2 + \frac{4m_c a^2}{\hbar^2}\Big(\Delta_v - \frac{\hbar^2 k_{v\parallel}^2}{2m_v} - \frac{\hbar^2 k_{v\perp}^2}{2m_r}\Big)}, \quad (24)$$

$$k^{(m)}_{v\parallel} = \sqrt{\frac{m_v}{2m_c}\Big[\frac{4m_c a^2}{\hbar^2}\Delta_v + k_{v\perp}^2\Big(1+2\frac{m_c}{m_v}\Big)\Big]}, \qquad k^{(m)}_{v\perp} = \frac{1}{\hbar}\sqrt{\frac{4m_v m_c}{2m_c+m_v}\Delta_v}. \quad (25)$$

The characteristic dependence $\tau^{-1}_{\mathrm{aug}}(\Delta_v)$ is depicted in Fig. 4.

An approximate expression for the probability of Auger transitions of the 31 ⟶ 22 type has the form [16]

$$\frac{W_{31,22}}{S} = \frac{32\pi^2 m_c e^4}{\hbar^2 a^2 \varepsilon^2} I(\tilde{n}_1, \tilde{n}_2, \tilde{\delta}), \tag{26}$$

where $\alpha \sim 1$ is a coefficient describing weak interference of the direct and exchange contributions to the probability of the Auger process, $\tilde{n}_i \equiv 4a^2 n_i$, $\tilde{\delta} \equiv 4a^2 m_c \delta/\hbar^2$, and $\delta = E_{03} - 2E_{02} + E_{01}$. The dimensionless function $I(\tilde{n}_1, \tilde{n}_2, \tilde{\delta})$ in the right-hand side of Eq. (26) is successfully approximated by the expression

$$I(\tilde{n}_1, \tilde{n}_3, \tilde{\delta}) \approx a_0 \tilde{n}_1 \tilde{n}_3 [1 + c_0 \tilde{\delta}^3 \exp(-d_0 \tilde{\delta})]^{-1} \times \{\tilde{n}_1 + \tilde{n}_3 + b_0 [1 + (\tilde{\delta}/\tilde{\delta}_0)^8]\}^{-1/2}, \tag{27}$$

where $a_0 \approx 0.1255$, $b_0 \approx 2.44$, $c_0 \approx 0.1$, $d_0 \approx 0.38$, and $\tilde{\delta}_0 \approx 7.37$. For $\tilde{\delta} \lesssim 1$, the quantity $I(\tilde{n}_1, \tilde{n}_2, \tilde{\delta})$ is virtually independent of $\tilde{\delta}$. For very large $\tilde{\delta}$, we have $I^{(d)}(\tilde{n}_1, \tilde{n}_3, \tilde{\delta}) \propto \delta^{-4}$.

Let us consider the role of charge carrier heating due to intrasubband absorption of light. In the case of conventional interband Auger recombination, the inclusion of heating (to be more precise, the quasi-equilibrium non-Fermi component of the distribution function; see, for example, [2]) is of fundamental importance since only charge carriers with a kinetic energy exceeding a certain threshold value can take part in an Auger process. The situation is different for Auger-type processes considered by us here: electrons in subband 3 with zero kinetic energy can participate in such a process. The role of heating is played to a certain extent by the intersubband photoexcitation of carriers, which is taken into account explicitly in our model. The only requirement is that the values of δ and Δ_v be positive, which is presumed by us here. In the case of deep QWs and high values of $\hbar\omega$ we are dealing with, heating of electrons by light is weak in view of the smallness of the intrasubband light absorption coefficient (which is two to three orders of magnitude smaller than the coefficient of intersubband absorption). For example, in the most interesting range of intensities $j \sim j_{th}$, the change in the electron temperature can hardly exceed 10 K according to rough estimates. Naturally, this cannot lead to appreciable consequences for the effects treated here. At the same time, electron heating (in a broad sense) must be taken into account in the high-intensity range along with other factors complicating analysis in this case (see the end of Section 6). However, the intrasubband photoexcitation could play, in principle, an important role indeed only for small values of $\delta, \Delta_v \lesssim 30$–$50$ meV. It should be noted that a possible increase in the probability of Auger-type processes might only increase the effectiveness of the mechanism of cascade avalanche pair generation considered here, lowering the threshold value j_{th} of light intensity.

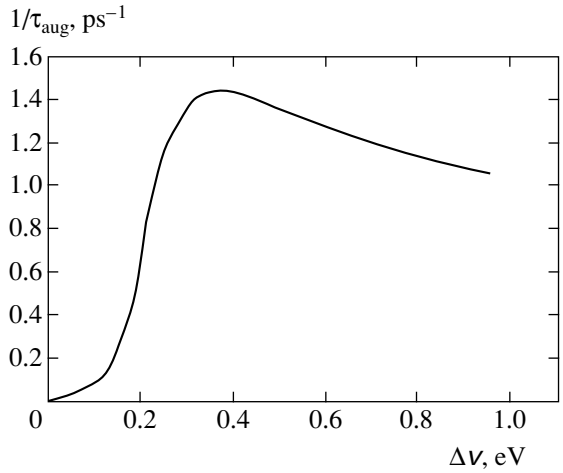

Fig. 4. Dependence of the rate of Auger transitions $3v \longrightarrow 11$ on $\Delta_v = E_{03} - 2E_{01} - E_v$.

5. INTERSUBBAND RELAXATION RATE

Intersubband relaxation in QWs is primarily due to the interaction between electrons and optical phonons. The problems of electron–phonon interaction in systems with QWs have been considered by many authors (see, for example, [17–28]). The complexity of the problem is due to the fact that confinement strongly modifies the vibrational spectrum of the system as well as the electron-vibrational interaction. In particular, the interaction of electrons both with confinement modes and with interface modes must be taken into account correctly. Intersubband scattering was treated in [22, 28]. In the case of interest to us, when the optical phonon energy $\hbar\omega L$ is smaller than the energy E_{01} of the lower subband in the well, the expression for the rate of transitions between the first and second subbands derived in [22] takes the form

$$W_{21}^{(R)} \approx \frac{e^2 m_c \omega_L a}{2\pi^2 \hbar^2} \left(\frac{1}{\varepsilon_\infty} - \frac{1}{\varepsilon_0}\right), \tag{28}$$

where ε_∞ and ε_0 are the high-frequency and static permittivities. In [28], the following expression was obtained instead of (28):

$$W_{21}^{(G)} \approx 4\pi W_{21}^{(R)} \sqrt{\frac{\varepsilon_0}{\varepsilon_\infty}} \Lambda; \tag{29}$$

here, Λ is the function of kinetic energy of an electron, which assumes values approximately equal to $(1$–$2) \times 10^{-1}$ in the energy range of interest. Both the formulas lead to an estimate of $W_{21} \sim 10^{11}$ s^{-1}.

6. BALANCE EQUATION FOR POPULATIONS

While constructing balance equations, it should be borne in mind that, for high pumping intensities j, the electron concentrations in subbands 2 and 3 become comparable and photoinduced transitions involving both absorption and emission of a photon must be taken into account. This does not apply to transitions $v \longrightarrow 1$, $1 \longrightarrow 2$, and $3 \longrightarrow c$, since electrons or holes in final states for these transitions possess a fairly high kinetic energy; losing this energy due to rapid intrasubband or intersubband relaxation, the carriers leave the resonance energy range.

The system of equations for concentrations p, $n_{1,2,3}$, and n_c of nonequilibrium holes in the valence band, electrons in the three subbands, and electrons in the continuum of the conduction band, respectively, has the form

$$\begin{aligned}
\dot{n}_c &= \sigma_{3c} j n_3 - W_{c3} n_c - d_c n_c (p_0 + p), \\
\dot{n}_3 &= W_{c3} n_c - \sigma_{3c} j n_3 - (W_{31} + W_{32} + \tau_{\text{aug}}^{-1}) n_3 \\
&\quad + \sigma_{23} j (n_2 - n_3) - W_{31,22}(n_1, n_3), \\
\dot{n}_2 &= -W_{21} n_2 - \sigma_{23} j (n_2 - n_3) + W_{32} n_3 \\
&\quad + 2 W_{31,22}(n_1, n_3) + \sigma_{12} j n_1, \\
\dot{n}_1 &= -\sigma_{12} j n_1 + W_{21} n_2 + W_{31} n_3 - W_{31,22}(n_1, n_3) \\
&\quad + 2\tau_{\text{aug}}^{-1} n_3 + \sigma_{v1}^{(q)} j^q - d_1 n_1 (p_0 + p), \\
p &= n_1 + n_2 + n_3 + n_c,
\end{aligned} \tag{30}$$

the initial conditions being $n_2, n_3, n_c, p = 0$, $n_1 = n_{10}$ for $t = 0$. Superscript q assumes values 1 or 2 for one- and two-photon transitions between the valence band and subband 1 in the QW. Equations (30) contain "two-dimensional" concentrations p, p_0, and n_c of particles from the continuous band spectrum, which differ from the conventional concentrations $\tilde{p}$, $\tilde{p}_0$, and $\tilde{n}_c$ in factor n_w (see Section 3): $\tilde{n}_c = n_c n_w$, etc. Terms $-d_c n_c (p_0 + p)$ and $-d_1 n_1 (p_0 + p)$ describe conventional bimolecular recombination of electrons in the conduction band and subband 1 with holes from the valence band in region B. The quantity p_0 stands for the equilibrium concentration of holes. For the sake of simplicity, we assume that the equilibrium concentration of electrons in the conduction band is equal to zero. In formulas (30), W_{c3} is the rate of electron capture from the continuous spectrum of the conduction band to subband 3 of the QW. Cross sections σ_{ij} of optical transitions between the ith and jth subbands are estimated taking into account the fact that typical bandwidths of intersubband absorption are equal to 10–30 meV (see, for example, review [29]). In this case, for resonant transitions between neighboring subbands, we obtain $\sigma_{i,i+1} \sim 1$ cm^2/(ps MW) $\times$ (a/5 nm)^2. The value of σ_{12} for a nonresonant indirect transition decreases upon an increase in the resonance detuning. For $\delta \sim 100$ meV, the value of σ_{12} is 3–4 orders of magnitude smaller than σ_{23}.

In accordance with the energy and momentum conservation laws, the states of two electrons in subband 2, into which these electrons enter as a result of the Auger transition $31 \longrightarrow 22$, are characterized by relatively large values of wave vectors $\mathbf{k}_{21\perp}$ and $\mathbf{k}_{22\perp}$ for not very small values of resonance detuning $\omega - \omega_{21}$. Numerical calculations show, however, that for light intensities j of interest to us, concentration n_2 may prove to be sufficient for filling these states. In this case, saturation of Auger transition probabilities takes place, which can be described approximately by a factor of $[1 + (n_2/n_f)^r]^{-1}$. The results of numerical calculations given below were obtained with $n_f \approx 3 \times 10^{12}$ cm^{-2} and $r = 4$. A variation of the values of n_f and r in reasonable limits or of probability W_{c3} of free electron capture by the QW does not change the results qualitatively. This circumstance as well as analysis of the dependence of solutions to Eqs. (30) on other parameters (see discussion in Section 7) suggests that multiparticle effects will not change the results qualitatively. Owing to these effects, the positions of peaks and the intensity of intersubband optical transitions are generally determined by the charge carrier concentration in the QW (see review [30]). A comparison of theory and experiment shows (see, for example, [31]) that the contribution to the short-wave shift of intrasubband IR absorption peaks comes from depolarization effects (plasmon shift) and exciton-type effects (interaction of an excited electron with a hole in the ground state) as well as the effects of direct and exchange Coulomb interactions. A change in the electron populations in subbands 1, 2, and 3 can generally affect the values of $\sigma_{i,i+1}$ to a certain extent. Nevertheless, this effect can hardly be significant in the case of deep QWs under investigation, when plasma frequencies ω_0 of electrons in a well are smaller than the gaps ω_{ij} between size quantization subbands for all values of concentration n_i of interest.

Figure 5 shows the dependence of populations $n_{1,2,3}$ and n_c on the time elapsed from the beginning of pumping. The quasi-equilibrium time τ_{eq} as a function of pumping intensity j is shown in Fig. 6 for one- and two-photon versions of the photon avalanche effect. Figure 7 depicts the dependences of quasi-equilibrium populations $n_{1,2,3}$ and n_c on the pumping intensity j. In our calculations, we used the following values of the parameters appearing in the right-hand sides of Eqs. (30): $a = 3 \times 10^{-7}$ cm, $\alpha = 1.2$, $W_{31} = 0.02$ ps^{-1}, $W_{c3} = 0.01$ ps^{-1}, $W_{32} = 0.07$ ps^{-1}, $W_{21} = 0.1$ ps^{-1}, $\tau_{\text{aug}} = 2$ ps, $n_{10} = 0$, $\delta = 0.05$ eV, $\sigma_{12} = 0.003$ cm^2/ps MW, $\sigma_{23} = 2$ cm^2/ps MW, $\sigma_{3c} = 0.025$ cm^2/ps MW, $\sigma_{v1}^{(1)} = 1.5 \times 10^{10}$ MW^{-1} ps^{-1}, $\sigma_{v1}^{(2)} = 10^4$ cm^{-2} MW^{-2} ps^{-1}, $p_0 = 3 \times$

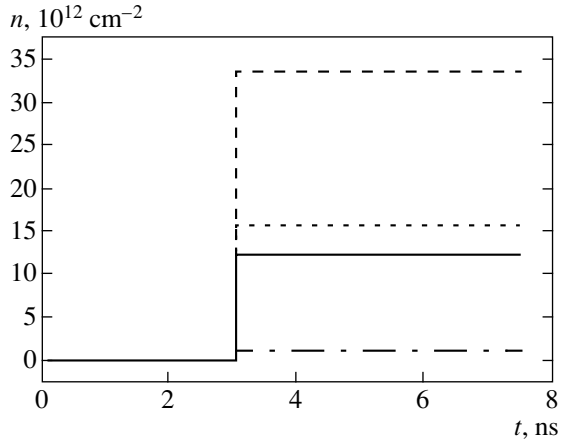

Fig. 5. Dependence of the populations of size quantization subbands and the conduction band on the time elapsed from the beginning of a pumping pulse; $j = 2$ MW/cm^2. Here and in Fig. 7 the populations n_1, n_2, n_3, and n_c are shown by solid, dashed, dotted, and dot-and-dash lines, respectively.

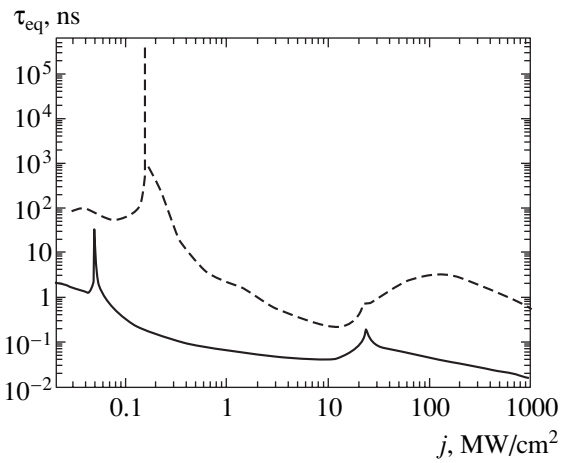

Fig. 6. Dependence of the quasi-equilibrium time τ_{eq} for populations on light intensity j for one-photon (solid curve) and two-photon (dashed curve) transitions between the valence band and the lower size quantization subband.

10^{10} cm^{-2}, $d_1 = 0.01$ cm^2 s^{-1}, and $d_c = 0.003$ cm^2 s^{-1} (see also Section 3). It should be borne in mind that system of equations (30) is obviously inapplicable for $j \gg 10$ MW/cm^2. In this case, the Rabi frequency ω_R for transitions between subbands 2 and 3 is much greater than 10^{13} s^{-1}, i.e., $\omega_R \gg \tau^{-1}$, where τ is the characteristic time of electron momentum relaxation in subband 2 or 3. In such a situation, where we are dealing with the optical Stark effect, a completely different approach is required (see, for example, [32, 33]). In addition, for $j > 10$ MW/cm^2, we must also take into account electron heating in the field of a light wave and a number of nonlinear processes. Nevertheless, we present the $n_i(j)$ dependences for large j also in order to demonstrate the form of solutions to system of equations (30).

7. DISCUSSION

The system of balance equations (30) obviously provides only a very rough description of the kinetics of photoinduced transitions in a system with type II QWs; strictly speaking, a much more detailed approach is required in this case. At the same time, it will be clear from the discussion that the qualitative form of solutions to Eqs. (3) is not very sensitive to the description of elementary processes controlling together the photon avalanche effect. It can hence be expected that a more rigorous solution of the problem would not lead to a qualitative revision of the results obtained here. It should be borne in mind that such a solution is extremely complicated.

The qualitative features of the photon avalanche effect in a type II QW are practically the same for one- and two-photon transitions between the valence band in region *B* and the lower size quantization subband in region *A*. However, since (i) less stringent constraints are imposed on the energy spectrum parameters of the heterostructure, (ii) a higher ratio of the energy of an EHP produced to a pumping light quantum is attained, and (iii) the threshold nature of the effect is manifested most clearly in the two-photon version of the photon avalanche effect, we will mainly consider the results pertaining exactly to this case.

The complexity of the "simplified" system of nonlinear equations (30) hampers its qualitative analysis of the type carried out in [11, 12, 16] and the derivation of simple formulas like Eqs. (1) and (2). For this reason, we will use the results of numerical calculations. It can be seen from Fig. 7b that there exists a threshold light intensity $j = j_{th}$ for which the populations n_i ($i = 1, 2, 3$) and n_c change jumpwise by several orders of magnitude[1] (n_1 and n_2 change by five and ten orders of magnitude, respectively). For the values of parameters used in our computations, we have $j_{th} \approx 149.3$ kW/cm^2. In a wide range of $j_{th} < j < j_{th}^{(1)}$ ($j_{th}^{(1)} \approx 21.5$ MW/cm^2), the populations vary smoothly: n_1 and n_2 attain their peak values and start decreasing, while n_3 and n_c increase monotonically. The populations of subbands 2 and 3 are leveled out for values of j in the vicinity of $j_{th}^{(1)}$. In this case, the main light absorption channel (resonant interband transitions $2 \longrightarrow 3$) is blocked. In addition, in view of a high electron concentration in subband 2, the states in this subband which are final for the main "avalanche-forming" Auger-type process $31 \longrightarrow 22$ turn out to be occupied. Finally, the rate of bimolecular recombination of electrons and holes increases. Taken together, this leads to a sharp decrease in the light energy being absorbed and to a rapid decrease in the

[1] Naturally, while comparing theory and experiment, we must take into account the real intensity distribution over the light beam profile.

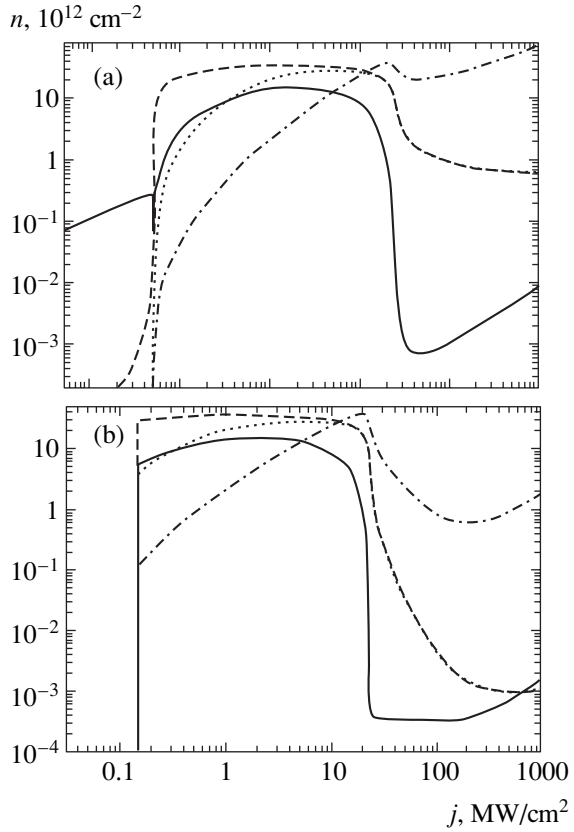

Fig. 7. Dependence of the equilibrium populations on light intensity for (a) one- and (b) two-photon transitions $v \longrightarrow 1$.

populations with increasing j in the region of $j \approx j_{\text{th}}^{(1)}$, although the change is not as strong as near j_{th}. For $j > j_{\text{th}}^{(1)}$, the populations vary smoothly with the light intensity. In this region, populations n_2 and n_3 are almost identical and $n_c \gg n_{2,3} \gg n_1$.

Since the model of cascade avalanche production of pairs under investigation includes a large number of various processes, the stability of the obtained pattern of nonlinear EHP production relative to changes in the parameters appearing in Eqs. (30) should be investigated. Let us first consider the effect of variation of the parameters on the threshold intensity value j_{th}. The dependence of j_{th} on the initial occupancy n_{01} of the lower subband is correctly described by the formula

$$j_{\text{th}}(n_{01}) \approx j_{\text{th}}(0)(1 + \vartheta_1 n_{01})/(1 + \vartheta_2 n_{01}), \qquad (31)$$

where $\vartheta_1 \approx 1.225 \times 10^{-9}$ and $\vartheta_2 \approx 3.65 \times 10^{-9}$ for the chosen values of the remaining parameters. It can be seen that the threshold intensity decreases by a factor of 2.5 as compared to $j_{\text{th}}(0)$ even for $n_{01} \sim 10^9$ cm^{-2}. This fact can be used, in principle, for controlling light by light. Indeed, the presence (absence) of an additional weak light pulse inducing the $v \longrightarrow 1$ transitions creates a situation in which a fixed pumping intensity j is higher (lower) than the threshold value.

The dependence of j_{th} on the two-photon absorption cross section $\sigma_{v1}^{(2)}$ can be approximated by the formula

$$j_{\text{th}}(\tilde{\sigma}_{v1}^{(2)}) \approx j_0 + \chi_0/[1 + \chi_1 \exp(\tilde{\sigma}_{v1}^{(2)}/\chi_2)], \qquad (32)$$

where $\tilde{\sigma}_{v1}^{(2)} = \sigma_{v1}^{(2)}/(10^4\ \text{cm}^2\ \text{MW}^{-2}\ \text{ps}^{-1})$. For the chosen values of other parameters, we have $j_0 \approx 0.05965$ MW/cm^2, $\chi_0 \approx 477$ MW/cm^2, $\chi_1 \approx 5098$, and $\chi_2 \approx 145.616$. If $\tilde{\sigma}_{v1}^{(2)}$ varies in a wide range from 0.05 to 10^5 for which approximation (32) is valid; the value of j_{th} decreases just to one-third of its initial value.

The dependence of j_{th} on the rate of the Auger transitions $3v \longrightarrow 11$ can be described by a relation of the same type as (31):

$$j_{\text{th}}(\tau_{\text{aug}}) \approx j_1(1 + \pi_1 \tau_{\text{aug}}^{-1})/(1 + \pi_2 \tau_{\text{aug}}^{-1}), \qquad (33)$$

where τ_{aug} is measured in picoseconds and $j_1 \approx 0.801$ MW/cm^2, $\pi_1 \approx 13.92$, and $\pi_2 \approx 82.54$ for the adopted values of the other parameters. As the value of τ_{aug}^{-1} increases from 0.005 to 50 ps^{-1}, the value of j_{th} decreases by a factor of less than 5.

If we simultaneously change the values of bimolecular recombination rates d_1 and d_c by a factor of ϕ, we obtain the following expression for j_{th}:

$$j_{\text{th}}(\phi) \approx j^{(0)} + \nu\phi^{\lambda}; \qquad (34)$$

here, $j^{(0)} \approx 0.05$ MW/cm^2, $\nu \approx 0.09448$ MW/cm^2, and $\lambda \approx 0.96913$. For example, if ϕ increases from 0.01 to 100, the values of j_{th} increase from 0.05 to 6.6 MW/cm^2. However, there are no grounds to assume that the recombination rates may correspond to values of $\phi \gg 1$. However, in the region of realistic values of $\phi \lesssim 1$, threshold intensities depend on the recombination rate only slightly. At the same time, the second threshold intensity $j_{\text{th}}^{(1)}$ (see Fig. 7) increases with decreasing ϕ. In this case, the decrease in populations $n_{1,2,3}$ and n_c becomes smoother and loses its "criticality." It should be recalled that the results obtained for $j \gtrsim 10$ MW/cm^2 should be treated with care since the applicability of the balance equations (30) in the range of such high pumping intensities is dubious.

In the model of the photon avalanche effect in a doped QW considered in [15, 16], where the total number of electrons in the QW is fixed, the threshold pumping intensity is independent of the cross section of the photoinduced transition $1 \longrightarrow 2$ (see formula (1) in Section 1). The photon avalanche effect in impurity systems possesses a similar property [8–14]. The situa-

tion is different in the model considered by us here. The dependence of j_{th} on σ_{12} is approximated by the expression

$$j_{th}(\sigma_{12}) \approx j_{12}(1 + f\sigma_{12})/(1 + g_1\sigma_{12} + g_2\sigma_{12}^2), \quad (35)$$

where $j_{12} \approx 1.924$ MW/cm^2, $f \approx 1720$, $g_1 \approx 9181$, $g_2 \approx 5.287 \times 10^6$, and σ_{12} is measured in cm^2 ps^{-1} MW^{-1}. As the value of σ_{12} increases from 3×10^{-4} to 1.5×10^{-2} cm^2 ps^{-1} MW^{-1}, the value of j_{th} decreases from 1.6 to 0.05 MW/cm^2. For $\sigma_{12} \gtrsim 10^{-2}$ cm^2 ps^{-1} MW^{-1}, i.e., in the case when absorption via transitions between the first and second subbands is close to resonant absorption, the system of equations (30) requires an appropriate modification. The reason for the discrepancy between the results obtained in [15, 16] and our results is that the absorption via the $1 \longrightarrow 2$ transition in [15, 16] occurs from the ground state and the absorption in our case takes place from the excited state, while absorption from the ground state occurs only during the $v \longrightarrow 1$ transitions. The value of j_{th} indeed exhibits a weak dependence on cross sections $\sigma_{v1}^{(1,2)}$ (see above).

At the same time, the quasi-equilibrium times τ_{eq} increase with decreasing $\sigma_{v1}^{(1,2)}$ in the same way as they increase with decreasing σ_{12} in the situation considered in [15, 16] (cf. formula (1)). Figure 6 shows that, in the case of two-photon transitions $v \longrightarrow 1$, the values of τ_{eq} are one to two or even more (for $j \approx j_{th}$) orders of magnitude higher than in the case of single-photon transitions. This is due to the inequality $\sigma_{v1}^{(1)} \gg \sigma_{v1}^{(2)} j_{th}$.

Let us consider for comparison a purely cascade scheme of EHP excitation. To this end, we eliminate the Auger transitions $31 \longrightarrow 22$ and $3v \longrightarrow 11$, but assume that transitions between subbands 1 and 2 are of the resonant type; i.e., we increase cross sections σ_{12} by three orders of magnitude, leaving the values of the other parameters unchanged. In this case, for $j > j_{th}$, the populations $n_{1,2,3}$ and n_c turn out to be many orders of magnitude (seven for $j = 200$ kW/cm^2) smaller than in the cascade avalanche scheme.

Systems $In_{0.53}Ga_{0.47}As$ (region *A*)/$AlAs_{0.56}Sb_{0.54}$ [34, 35] (the band parameters of this system are close to those used in our calculations) or $In_{0.3}Ga_{0.7}As$ (region *A*)/AlAs (region *B*) [36] may serve as examples of type II heterostructures with deep QWs.

8. CONCLUSIONS

The analysis carried out here proves that the photon avalanche effect emerges when a type II heterostructure with deep QWs is exposed to IR light with a frequency resonant to the transition between the second and third size-quantization subbands. The effect shows a threshold behavior, with relatively low (~10–100 kW/cm^2) threshold intensities of the IR light. For intensities above the threshold, the number of electron–hole pairs generated is considerable, which gives rise to photoconduction in the direction of the nanostructure growth and to recombination luminescence at a wavelength smaller than that of exciting light by a factor of 3–5. The times for the establishment of quasi-equilibrium distributions in the electron system may be in the nano- or picosecond range depending on pumping conditions. These times increase sharply for pumping intensities close to the threshold values.

ACKNOWLEDGMENTS

This study was supported by the Russian Foundation for Basic Research (project nos. 01-02-17060, 02-02-17311, and 00-02-016337) and by the program Physics of Solid-State Nanostructures-5.

REFERENCES

1. B. G. Gorshkov, A. S. Epifanov, A. A. Manenkov, *et al.*, Tr. Inst. Obshch. Fiz., Akad. Nauk SSSR **4**, 99 (1986).
2. E. Yu. Perlin, A. V. Fedorov, and M. B. Kashevnik, Zh. Éksp. Teor. Fiz. **85**, 1357 (1983) [Sov. Phys. JETP **58**, 787 (1983)].
3. A. M. Danishevskiĭ, E. Yu. Perlin, and A. V. Fedorov, Zh. Éksp. Teor. Fiz. **93**, 1319 (1987) [Sov. Phys. JETP **66**, 747 (1987)].
4. E. Yu. Perlin, Opt. Spektrosk. **82**, 259 (1997) [Opt. Spectrosc. **82**, 235 (1997)].
5. S. D. Ganichev, E. L. Ivchenko, E. Yu. Perlin, *et al.*, Zh. Éksp. Teor. Fiz. **91**, 1233 (1986) [Sov. Phys. JETP **64**, 729 (1986)].
6. H. Minasian and S. Avetisyan, Phys. Rev. B **34**, 963 (1986).
7. E. Yu. Perlin and A. V. Ivanov, Opt. Spektrosk. **87**, 42 (1999) [Opt. Spectrosc. **87**, 36 (1999)].
8. J. S. Chivian, W. E. Case, and D. D. Eden, Appl. Phys. Lett. **35**, 124 (1979).
9. A. W. Kueny, W. E. Case, and M. E. Koch, J. Opt. Soc. Am. B **6**, 639 (1989).
10. A. W. Kueny, W. E. Case, and M. E. Koch, J. Opt. Soc. Am. B **10**, 1834 (1993).
11. S. Guy, M.-F. Joubert, and B. Jacquier, Phys. Rev. B **55**, 8240 (1997).
12. M.-F. Joubert, Opt. Mater. **11**, 181 (1999).
13. M. P. Hehlen, A. Kuditcher, A. L. Lenef, *et al.*, Phys. Rev. B **61**, 1116 (2000).
14. E. Yu. Perlin, A. M. Tkachuk, M.-F. Joubert, *et al.*, Opt. Spektrosk. **90**, 772 (2001) [Opt. Spectrosc. **90**, 691 (2001)].
15. E. Yu. Perlin, Opt. Spektrosk. **5**, 777 (2001) [Opt. Spectrosc. **91**, 729 (2001)].
16. E. Yu. Perlin, J. Lumin. **94–95**, 249 (2001).
17. K. Huang and B. Zhu, Phys. Rev. B **38**, 13377 (1988).
18. M. H. Degani and O. Hipolito, Phys. Rev. B **35**, 7717 (1987).
19. R. Lassnig, Phys. Rev. B **30**, 7132 (1984).

20. T. Tsuchia and T. Ando, Phys. Rev. B **47**, 7240 (1993).
21. B. K. Ridley, Phys. Rev. B **37**, 4583 (1988).
22. B. K. Ridley, Phys. Rev. B **39**, 5282 (1989).
23. B. K. Ridley, Phys. Rev. B **47**, 4592 (1993).
24. H. Rücker, E. Molinari, and P. Lugli, Phys. Rev. B **44**, 3463 (1991).
25. K. J. Nash, Phys. Rev. B **46**, 7723 (1992).
26. C. Trallero-Giner and F. Comas, Phys. Rev. B **37**, 4583 (1988).
27. C. Trallero-Giner, F. Comas, and F. Garsia-Moliner, Phys. Rev. B **50**, 1755 (1994).
28. J. L. Gondar, F. Comas, and F. Castro, Physica B (Amsterdam) **292**, 354 (2000).
29. B.-F. Levine, J. Appl. Phys. **74**, R1 (1993).
30. T. Ando, A. B. Fowler, and F. Stern, Rev. Mod. Phys. **54**, 437 (1982).
31. M. O. Manafresh, F. Szmulowicz, T. Vaughan, *et al.*, Phys. Rev. B **43**, 9996 (1991).
32. S. Schmitt-Rink, D. S. Shemla, and D. A. B. Miller, Adv. Phys. **38**, 89 (1989).
33. E. Yu. Perlin, Zh. Éksp. Teor. Fiz. **105**, 186 (1994) [JETP **78**, 98 (1994)].
34. A. Neogi, T. Mozume, H. Yoshida, *et al.*, IEEE Photon. Technol. Lett. **11**, 632 (1999).
35. A. Neogi, H. Yoshida, T. Mozume, *et al.*, Physica E (Amsterdam) **7**, 183 (2000).
36. C. P. Garcia, A. De Nardis, V. Pellegrini, *et al.*, Appl. Phys. Lett. **77**, 3767 (2000).

Translated by N. Wadhwa

Journal of Experimental and Theoretical Physics, Vol. 96, No. 3, 2003, pp. 555–563.
From Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 123, No. 3, 2003, pp. 625–634.

SOLIDS
Electronic Properties

Pseudogaps in Incommensurate Charge Density Waves and One-Dimensional Semiconductors¶

S. A. Brazovskii[a,b] and S. I. Matveenko[b,*]

[a]*Laboratoire de Physique Théorique et des Modèle Statistiques CNRS, Bât. 100, Université Paris-Sud, 91405 Orsay Cedex, France*
[b]*Landau Institute for Theoretical Physics, Moscow, 117940 Russia*
**e-mail: matveen@landau.ac.ru*

Received September 25, 2002

Abstract—We consider pseudogap effects for electrons interacting with gapless modes. We study generic 1D semiconductors with acoustic phonons and incommensurate charge density waves. We calculate the subgap absorption as it can be observed by means of photoelectron or tunneling spectroscopy. Within the formalism of functional integration and adiabatic approximation, the probabilities are described by nonlinear configurations of an instanton type. Particularities of both cases are determined by the topological nature of stationary excited states (acoustic polarons or amplitude solitons) and by the presence of gapless phonons that change the usual dynamics to the quantum dissipation regime. Below the free-particle edge, the pseudogap starts with an exponential (stretched exponential for gapful phonons) decrease of the transition rates. Deeply within the pseudogap, they are dominated by a power law, in contrast to a nearly exponential law for gapful modes. *© 2003 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION: PSEUDOGAPS IN 1D SYSTEMS

This paper is devoted to the theory of pseudogaps in electronic spectra in application to photoelectron spectroscopy (PES). We study the influence of quantum lattice fluctuations on electronic transitions in the subgap region for one-dimensional (1D) systems with gapless phonons. Low-symmetry systems with gapful spectra were recently addressed by the authors [1], and we refer to this paper for a more comprehensive review and references. Here, we show that sound branches of phonon spectra drastically change the transition rates, making them much more pronounced deeply within the pseudogap. We consider two types of systems: generic 1D semiconductors with acoustic electron–photon (*e–ph*) coupling (conducting polymers, quantum wires, and nanotubes) and incommensurate charge density waves (CDWs) [2], which possess a gapless collective phase mode.

The pseudogap concept [3] refers to various systems where the gap in their bare electronic spectra is partly filled and subgap tails occur. Even for pure systems and at temperature $T = 0$, there can be a rather smeared edge E_g^0, while the spectrum extends deeply inward the gap until some absolute edge E_g, which can be even zero (no true gap at all). The most general reason is that stationary excitations (eigenstates of the total *e–ph* system) are self-trapped states, polarons or solitons, whose energies W_p and W_s are below the free electron ones, thus forming the absolute edge at $E_g < E_g^0$. Nonstationary states filling the pseudogap range $E_g^0 > E > E_g$ can be observed only via instantaneous measurements such as optics, PES, or tunneling. Particularly near E_g^0, the states resemble free electrons in the field of uncorrelated quantum fluctuations of the lattice [4]; here, the self-trapping does not have enough time to develop. However, approaching the exact threshold E_g, the excitations evolve towards eigenstates, which are self-trapped *e–ph* complexes. The pseudogaps must be common in 1D semiconductors precisely because of favorable conditions for self-trapping [5]. The pseudogap is especially pronounced when the bare gap is opened spontaneously as a symmetry breaking effect. In quasi-1D conductors, this symmetry breaking is known as the Peierls–Fröhlich instability leading to CDW formation. Here, the picture of the pseudogap was first suggested theoretically [3] (we also recall [6] and another model [7]) in relation to the absence of a long-range order in 1D CDWs at a finite temperature. In this approach, the smearing of the mean-field electronic gap $2\Delta_0$ corresponds to the disappearance of the true Peierls–Fröhlich transition in favor of a smooth crossover. The pseudogap shape was related to, and derived from, the temperature-dependent finite correlation length ξ. An alternative picture was suggested in [4] and further developed in [8]. It concentrates on effects that persist even at zero temperature and are due to a strong interaction between bare electronic excitations and pertur-

¶This article was submitted by the authors in English.

bations (the amplitude and phase phonons) of the CDW ground state. Here, the pseudogap in instantaneous electronic spectra is related to the transformation of electrons into solitons.

Experimentally, pseudogaps in incommensurate CDWs were first addressed by optic [9–11] and more recently by the PES and ARPES (momentum-resolved PES) methods [12]. The earlier experiments were theoretically interpreted in [13] by compilation of the approaches in [3, 4, 7]. Detailed theories of the subgap absorption in optics have been developed for systems with low symmetries (nondegenerate, such as semiconductors with gapful phonons, or discretely degenerate such as the dimerized Peierls state). They first addressed the general type of polaronic semiconductors [14] with emphasis on long-range Coulomb effects, and then the 1D Peierls system, emphasizing solitonic processes (see [15] and references therein). The authors recently [1] extended the theory of pseudogaps to single electronic spectra in application to PES, and particularly intriguing, to ARPES probes. However, properties of incommensurate CDWs are further complicated by the appearance of a gapless collective mode, resulting in drastic changes. The case of acoustic polarons in a 1D semiconductor belongs to the same class, although this is not usually noticed.

A specific property of 1D systems with continuous degeneracy (with respect to the phase for incommensurate CDWs and to displacements for usual crystals) is that even a single electronic process can create topologically nontrivial excitations—solitons. For incommensurate CDWs, a single electron or hole with the energy near the gap edges $\pm\Delta_0$ spontaneously evolves to a nearly amplitude soliton, while the original particle is trapped at the local level near the gap center. The energy near $0.3\Delta_0$ is released, at first sight, within the time $\omega_{ph}^{-1} \sim 10^{-12}$ s. We see in what follows that there actually also exists a long-scale adaptation process that determines shapes of transition probabilities. Similarly, the usual acoustic polaron in 1D semiconductors is characterized by an electronic density of $\rho \sim \partial\varphi/\partial x$ self-localized within the potential well, and hence, there is a finite increment $\varphi(+\infty) - \varphi(-\infty) \sim \int \rho dx$ of the lattice displacements φ over the length x, which is the signature of topologically nontrivial solitons. These systems with continuous degeneracy form a special class that shows particular properties and must be studied differently than in [1]. They are addressed in this paper.

2. FUNCTIONAL INTEGRALS AND INSTANTONS FOR PES

As a function of the frequency Ω and momentum P, the absorption rate $I(\Omega, P)$ for ARPES can be expressed in terms of the spectral density of the one-electron retarded Green function $G(t, t'; x, x')$ as

$$I(P, \Omega) \propto \mathrm{Im}\int dX e^{-iPX}\int_0^\infty dT e^{i\Omega T} G(X, T, 0, 0). \quad (1)$$

We here address the simple PES, not resolved in momenta, which measures the integrated absorption intensity:

$$I(\Omega) = \frac{1}{2\pi}\int I(P, \Omega) dP.$$

(From now on, we omit all constant factors and set the Planck constant $\hbar = 1$; Ω is then measured with respect to a convenient level, the band edge for semiconductors or the middle of the gap for CDWs.)

We use the adiabatic approximation, which is valid when changes of electronic energies are much larger than the relevant phonon frequencies. Electrons move in a slowly varying phonon potential, e.g., $\mathrm{Re}[\Delta(x, t)\exp(2ik_F x)]$ for an incommensurate CDW, and at any instant t their energies $E(t)$ and wave functions $\psi(x, t)$ are therefore defined as eigenstates for the instantaneous lattice configuration and depend on time only parametrically. In what follows, we work in the Euclidean space $it \longrightarrow t$, which is adequate for studies of classically forbidden processes [14, 16, 17]. The integrated absorption intensity is then given by a functional integral over lattice configurations,

$$I(\Omega) \propto \int_0^\infty dT \int D[\Delta(x, t)]\psi_0(0, T)\psi_0^+(0, 0)e^{-S}, \quad (2)$$

where ψ_0 is the wave function of the particle (which is actually a hole for PES) added and extracted at moments 0 and T. Only the lowest singly filled localized state is relevant for calculations of subgap processes. The energy E_0 of this state is split inside the gap. The action

$$S = S[\Delta(x, t), T] = \left(\int_{-\infty}^0 + \int_T^\infty\right) dt L_0 + \int_0^T dt (L_1 - \Omega), \quad (3)$$
$$L_1 - L_0 = E_0,$$

is expressed through the Lagrangians $L_j[\Delta]$, where the subscripts $j = 0, 1$ correspond to ground states for $2M$ (the bare number) and $2M \pm 1$ electrons in the potential $\Delta(x, t)$. The main contribution comes from saddle points of S, instantons, which are extrema with respect to both the function $\Delta(x, t)$ and the time T. There are also special cases [1], particularly important for ARPES, where the extremum must be taken for the entire integrand in (2), with the wave functions in the prefactor taken into account. Otherwise, the stationary point is determined by $dS/dT = 0$, that is, $E_0(0) = E_0(T) = \Omega$, which determines $T(\Omega)$.

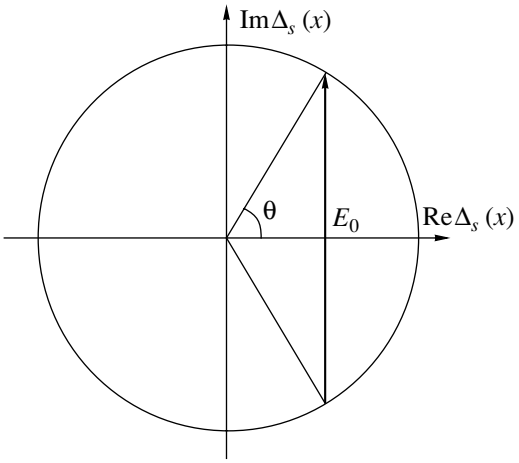

Fig. 1. Trajectory of the chord soliton with phase tails in the complex plane Δ.

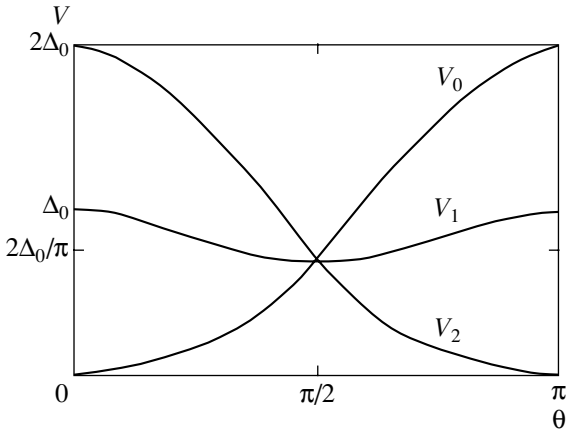

Fig. 2. Self-trapping terms V_ν for chord solitons as functions of the chiral angle 2θ for various fillings ν.

In what follows, we concentrate on most principal features, setting aside calculations of prefactors and the problem of the momentum dependence necessary for ARPES. For a simpler case of nondegenerate systems, they have been studied in [1].

3. CREATION OF AMPLITUDE SOLITONS IN INCOMMENSURATE CDWS

We first consider the subgap electronic spectra for the incommensurate CDW described by the Peierls–Fröhlich model. The incommensurate CDW order parameter is the complex field $\Delta = |\Delta(x, t)|\exp[i\varphi(x, t)]$ acting on electrons by mixing states near the Fermi momenta points $\pm k_F$. The Lagrangians L_j consist of the bare kinetic and potential lattice energies and of the sum over the filled electron levels, in the jth state,

$$L_j = \int dx \frac{2|\partial_t \Delta|^2}{\pi v_F \omega_0^2} + V_j[\Delta(x, t)],$$

where v_F is the Fermi velocity in the metallic state and ω_0 is the amplitude mode frequency ($\omega_0 \ll \Delta_0$ is the condition for the adiabatic approximation).

The important fact is that the stationary state of the system with an odd number of particles, the minimum of V_1, is an amplitude soliton, with the midgap state $E_0 = 0$ occupied by a single electron. Evolution of the free electron with the initial energy $E_0 = \Delta_0$ to the amplitude soliton with $W_s = 2\Delta_0/\pi < \Delta_0$ can be described by the known exact solution for intermediate configurations characterized by the single intragap $E_0 = \Delta_0\cos\theta$ with $0 \le \theta \le \pi$, whence $-\Delta_0 \le E_0 \le \Delta_0$. It was found [8] (see also reviews [18, 19]) to be the chord soliton with 2θ as the total chiral angle, $\Delta(+\infty)/\Delta(-\infty) = \exp(2i\theta)$ (see Fig. 1 and the Appendix for details). The filling numbers $\nu = 0, 1$ of the intragap state correspond to labels $j = 0, 1$. The term $V_0(\theta)$ monotonically increases from $V_0(0) = 0$ for the $2M$ ground state to $V_0(\pi) = 2\Delta_0$ for the $2M + 2$ ground state with two free holes. The term $V_1(\theta) = V_1(\pi - \theta)$ is symmetric and describes both a particle on the $2M$ ground state and a hole on the $2M + 2$ ground state. Obviously, $V_1(0) = V_1(\pi) = \Delta_0$, while the minimum is reached at $\theta = \pi/2$, that is, for a purely amplitude solution: $\min V(\theta) = V_1(\pi/2) = W_s < \Delta_0$, where $W_s = 2\Delta_0/\pi$ is the amplitude solution energy (see Fig. 2). Therefore, to create a nearly amplitude soliton with $\theta = 90°$, the light with $\Omega \approx W_s$ is absorbed by the quantum fluctuation with $E_0(\theta) = W_s$, which is close to the chord soliton with the angle $\theta \approx 50°$.

We note that the amplitude soliton, being an uncharged spin carrier with the topological charge one, is a quasiclassical realization of a spinon in systems with nonretarded attraction of electrons (that is, with high, rather than low, phonon frequencies). Therefore, our analysis is also qualitatively applied to arbitrary nonadiabatic electronic systems provided they are found in the spin-gap regime. (See also the next section.)

It is tempting to use the static solution, with some free parameter, as an ansatz for the time-dependent process; this proved to be successful in gapful cases [1, 15]. Here, however, setting $\theta \longrightarrow \theta(t)$, we would arrive at $\partial_t \Delta \neq 0$ for all x, and the action would therefore be infinite, S being proportional to the system length. The vanishing probability simply reflects the fact that a globally finite perturbation, characteristic of topologically nontrivial solitons, cannot spread over the whole length in a finite time. More generally, as a topologically nontrivial object, the amplitude soliton cannot be created in a pure form: adaptational deformations must appear to compensate the topological charge. These deformations develop over long space–time scales and can be described in terms of the gapless mode, the phase φ, alone. Allowing the time evolution of the chiral angle $\theta \longrightarrow \theta(t)$ within the core, we must therefore also

unhinder the field $\varphi \longrightarrow \varphi(x, t)$ for all x and t. The resulting trajectory is shown in Fig. 1 for an instant of time. Starting from $x \longrightarrow -\infty$ and returning to $x \longrightarrow \infty$, the configuration closely follows the circle $|\Delta| = \Delta_0$, changing almost entirely by phase. Approaching the soliton core, the phase approximately matches the angles $\pm\theta$ that delimit the chord part of the trajectory. The entire trajectory is closed, which leads to a finite action.

Except for a short time scale $T < \xi_0/u$ (see Section 4.2) characterized by small θ and large lengths $\xi = \xi_0/\sin\theta$, the configuration $\Delta(x, t)$ can be divided into the inner part, the core at $|x| \sim \xi$, and the outer part $|x| \gg \xi$, where only perturbations of the phase $\varphi(x, t)$ are important. The inner part can be described by the chord soliton $\Delta_{\mathrm{ChS}}(x, t)$. The chord angle $2\theta(t)$ evolves in time from $\theta(\pm\infty) = 0$ to θ_m in the middle the T interval. As $T \longrightarrow \infty$, that is, near the stationary state of the amplitude soliton, $\theta_m \longrightarrow \pi/2$. This value is actually preserved during most of the T interval, and the changes between $\theta = 0$ and $\theta = \pi/2$ are therefore concentrated within finite ranges $\tau_0 \sim \xi_0/u \ll T$ near the termination points. At large scales, we can see only a jump $\varphi(x, t) \approx \theta(t)\,\mathrm{sgn}\,x$ with $\theta(t) \approx \theta_m\Theta(t)\Theta(T - t)$, where Θ is the standard step function. Because the configuration stays close to the amplitude soliton during the time T, the main core contribution to the action is

$$S_{\mathrm{core}} = (W_s - \Omega)T + \delta S_{\mathrm{core}}, \tag{4}$$

where the first correction $\delta S^0_{\mathrm{core}} = \mathrm{const}$ comes form regions around the instants 0 and T independently. The significant T-dependent contribution $\delta S(T)$ comes from interference of regions 0 and T. Their interaction via gapful excitations, such as the amplitude mode, decays exponentially as $\delta S_{\mathrm{gap}} \propto \exp(-\omega_0 T)$. There are no other contributions for low-symmetry systems, but for an incommensurate CDW, there are sound modes providing the main effect, to be addressed below.

Matching the inner and outer regions is not well defined unless we consider the full microscopic time-dependent model, which is impossible. Fortunately, the long-range effects can be treated easily if we generalize the scheme suggested earlier for static problems of solitons in the presence of interchain interactions [8, 20]. The outer region is described by the action for the soundlike phase mode,

$$S_{\mathrm{snd}}[\varphi(x, t), \theta(t)] = \frac{v_F}{4\pi}\int\!\!\int dx dt\left[\left(\frac{\partial_t\varphi}{u}\right)^2 + (\partial_x\varphi)^2\right], \tag{5}$$
$$\varphi(t, x_s \pm 0) = \mp\theta(t),$$

where u is the phase velocity. The conditions on φ at $x_s \pm 0$ are due to the source provided by the chord soliton that is formed around x_s and enforces the discontinuity 2θ. Integrating $\exp[-S_{\mathrm{snd}}(\varphi, \theta)]$ over $\varphi(x, t)$ with this condition, we arrive at the action for $\theta(t)$,

$$S_{\mathrm{snd}}[\theta] \approx \frac{v_F}{2\pi^2 u}\int\!\!\int dt_{1,2}\dot\theta(t_1)\ln|(t_1 - t_2)|\dot\theta(t_2)$$
$$= \frac{v_F}{2\pi^2 u}\int\!\!\int dt_{1,2}\left(\frac{\theta(t_1) - \theta(t_2)}{t_1 - t_2}\right)^2. \tag{6}$$

The last form of this action is typical of the quantum dissipation problem [21], where $S \sim \sum|\omega||\theta_\omega|^2$. In our case, this dissipation arises from the emission of phase phonons forming a long-range tail in the course of the chord soliton development. Together with V_j, this action can be used to prove the above statements on the time evolution of the chord soliton core.

We now recall that $\dot\theta = \partial_t\theta$ is peaked within narrow regions on the order of ξ_0/u around the time instants $t = 0$ and $T = 0$ and is close to zero elsewhere. Then

$$S_{\mathrm{snd}} \approx \frac{v_F}{4u}\ln\frac{uT}{\xi_0}. \tag{7}$$

There is an even more phenomenological standpoint (see [22] for more details and examples of combined topological defects). The amplitude soliton creates the π-discontinuity along its world line $(0 < t < T, 0)$. To be topologically allowed, that is, to have a finite action, the line must terminate with half-integer vortices located at $(0, 0)$ and $(0, T)$, whose circulation must provide the compensating jump $\delta\varphi = \pi$ along the interval ($\Delta \longrightarrow -\Delta$ combined with $\varphi \longrightarrow \varphi + \pi$ leaves the order parameter $\Delta\exp(i\varphi)$ invariant). The standard energy of vortices for (5) then leads to action (7). Contrary to the usual 2π-vortices, the line connecting the half-integer ones is a physical singularity whose tension gives (4).

Minimizing $S_{\mathrm{tot}} = S_{\mathrm{core}} + S_{\mathrm{snd}}$ with respect to T, we obtain the power law near the amplitude soliton edge $\Omega \geq W_s$,

$$I(\Omega) \propto \left(\frac{\Omega - W_s}{W_s}\right)^\beta, \quad \beta = \frac{v_F}{4u}, \tag{8}$$

which is much more pronounced than the exponential law for gapful cases (see (15) below).

Our derivation suggests a literal long-range order at large (x, t) distances and neglects all fluctuations of the phase except perturbations enforced by the instanton. However, the mean fluctuations of the phase diverge and the order parameter decays in accordance with a power law. These long-range fluctuations are not related to the instanton and can be taken into account a posteriori. This can easily be done by noticing that the eigenfunctions in the prefactor in (2) transform as $\Psi_0 \longrightarrow \Psi_0\exp[i\varphi(x, t)/2]$, and being averaged, contribute the action term

$$\delta S_\varphi = \frac{1}{8}\langle[\varphi(0, 0) - \varphi(0, T)]^2\rangle \approx \frac{u}{4v_F}\ln\frac{uT}{\xi_0}.$$

Therefore, the effect of phase fluctuations, as well as the major role of the form factor, is simply a correction to the value of the index in (8), $\beta \longrightarrow \beta^* = v_F/4u + u/4v_F$. Within our adiabatic approximation $u/v_F \ll 1$, the correction is small, but it builds a bridge to quantum nonadiabatic models where exactly β^* appears as the index of the single-particle Green function with $\gamma_\rho = u/v_F$ identified as the charge channel exponent. The link is completed by noting that the amplitude soliton is a realization of the spinon and that the phase discontinuity in (5) is equivalent, together with fluctuations, to applying the operator

$$\exp\left\{\frac{i}{2}\varphi(x, t) + i\frac{\pi}{2}\frac{\delta}{\delta\varphi}\operatorname{sgn} x\,\Theta(t)\Theta(T-t)\right\},$$

which is our limit for bosonization.

4. ACOUSTIC POLARON AND THE FREE EDGE

4.1. One-Dimensional Semiconductors with Acoustic and Optical Polarons

Behavior near the free edge $\Omega \approx \Delta_0$ is dominated by small fluctuations η in the gap amplitude, $|\Delta| = \Delta_0 + \eta$, and at the Fermi level, $\delta E_F = \varphi' v_F/2$, via the phase gradient $\varphi' = \partial_x\varphi$. We consider it in the framework of the general problem of a combined (gapful and acoustic) polaron. The simpler, compared to the CDW, single-particle formulation bears similar qualitative features but allows a more detailed analysis. We consider electron (hole) states in a 1D dielectric near the edge of a conducting (valence) band. We take into account the gapful mode η with the coupling g_0 and the sound mode (for which we keep the "phase" notation φ) with the velocity u and the coupling g_s. In generic semiconductors, the sound mode is always present as the usual acoustic phonon, while the gapful mode can be present as an additional degree of freedom. In all CDWs, the gapful mode is always present as the amplitude fluctuation $|\Delta| = \Delta_0 + \eta$ while the sound mode appears in incommensurate CDWs as the phase $\Delta = |\Delta|\exp(i\varphi)$.

Within the adiabatic approximation for the electron wave function Ψ, the action S (at imaginary time) is given by

$$S = \int dx \int_0^T dt \left\{ \left(\frac{1}{2m}|\partial_x\Psi|^2 - \Omega|\Psi|^2 \right) + (g_s\partial_x\varphi + g_0\eta)\Psi^\dagger\Psi \right\} \tag{9}$$

$$+ \int dx \int_{-\infty}^{\infty} dt \left\{ \frac{K_s}{2}\left[\left(\frac{\partial_t\varphi}{u}\right)^2 + (\partial_x\varphi)^2\right] + \frac{K_0}{2}\left[\left(\frac{\partial_t\eta}{\omega_0}\right)^2 + \eta^2\right] \right\}.$$

For the incommensurate CDW case, we therefore have $m = \Delta_0/v_F^2$, $g_0 = 1$, $g_s = v_F/2$, $K_s = v_F/2\pi$, $K_0 = 4v_F/\pi$, $2^{3/2}u/v_F = \omega_0/\Delta_0$, and Ω is counted with respect to the edge Δ_0 rather than to the middle of the gap as in the previous section.

It is well known [5] that the stationary state, i.e., the time-independent extremum of (9), corresponds to the self-trapped complex, the polaron. Here, it is composed equally of η and φ', which contribute additively to the static coupling (while the dynamics is completely different):

$$\lambda = \lambda_s + \lambda_0 = \frac{g_s^2}{K_s} + \frac{g_0^2}{K_0}.$$

The polaronic length scale l for $\eta \sim \varphi' \sim |\Psi|^2 \equiv \rho_p(x)$ is $l = 2\lambda/m$ and the total energy is $W_p = -m\lambda^2/24$. The conditions $|W_p| \gg \omega_0$ and $\lambda \gg u$ define the adiabatic, Born–Oppenheimer, approximation. For the CDW case, $\lambda_s = v_F\pi/2$ and $\lambda_0 = v_F\pi/4$, and therefore, $\lambda \sim v_F$; and we arrive at $|W_p| \sim \Delta_0$ and $l \sim \xi_0 = v_F/\Delta_0$, which are the microscopic scales where the single electronic model can be used only qualitatively. The full-scale approach for nearly stationary states was considered in Section 3, but the upper pseudogap region near the free edge Δ_0 is described by model (9) even quantitatively and the most efficiently.

We can integrate over the fields φ and η at all (x, t) to obtain the action in terms of ψ alone, which is now defined only on the interval $(0, T)$ for t,

$$\begin{aligned} S\{\Psi; T\} &= \int dx dt \left(\frac{1}{2m}|\partial_x\Psi|^2 - \Omega|\Psi|^2 \right) \\ &- \frac{1}{2}\iint dt_{1,2} \iint dx_{1,2} \\ &\times \{ U_0(x_1 - x_2, t_1 - t_2)\rho(x_1, t_1)\rho(x_2, t_2) \\ &+ U_s(x_1 - x_2, t_1 - t_2)\partial_x\rho(x_1, t_1)\partial_x\rho(x_2, t_2) \}. \end{aligned} \tag{10}$$

Here, the retarded self-attraction potentials are

$$\begin{aligned} U_s &= \frac{\lambda_s u}{2\pi}\ln\sqrt{x^2 + t^2u^2}, \\ U_0 &= \frac{1}{2}\lambda_0\omega_0\exp[-\omega_0|t|]\delta(x). \end{aligned} \tag{11}$$

An equivalent form, suitable at large T, is obtained via integrating by parts,

$$\begin{aligned} S\{\Psi; T\} &= \int dx \int_0^T dt \left[\frac{1}{2m}|\partial_x\Psi|^2 - \Omega\rho - \frac{\lambda}{2}\rho^2 \right] \\ &+ \frac{1}{2}\iint dt_{1,2} \iint dx_{1,2} \partial_t\rho(x_1, t_1)\partial_t\rho(x_2, t_1)U(x_1 - x_2, t_1 - t_2), \end{aligned} \tag{12}$$

where $U(x, t) = u^{-2}U_s + \omega_0^{-2}U_0$.

The absorption near the absolute edge $\Omega \approx W_p$ is determined by long-time processes when the lattice configuration is almost statically self-consistent. The first term in (12) is nothing but the action S_{st} of the static polaron whose extremum at a given T is

$$S_{st} \approx -T\delta\Omega, \quad \delta\Omega = \Omega - W_p.$$

The second term in (12), S_{tr}, collects contributions only from short transient processes near the impact moments $t = 0, T$, which are seen by the long-length part as $\partial_t \rho(x, t) \approx \rho_p(x)[\delta(t) - \delta(t - T)]$, where ρ_p is the density for the static polaron solution. We obtain

$$S_{tr} \approx \iint dx_{1,2}\rho_p(x_1)\rho_p(x_2)U(x_1 - x_2, T)$$
$$= \frac{\lambda_s}{2\pi u}\ln\frac{uT}{l} + C_0\frac{\lambda_0/l}{\omega_0}\exp(-\omega_0 T) + \text{const}$$

with $C_0 \sim 1$. We see the dominant contribution of the sound mode that grows logarithmically in T, while the part of the gapful mode decays exponentially. If the sound mode is present, the extremum over T is

$$T \approx \frac{\lambda_s}{2\pi u}\frac{1}{\delta\Omega}, \quad S \approx \frac{\lambda_s}{2\pi u}\ln\frac{C_s|W_p|}{\delta\Omega}, \quad C_s \approx 0.9. \quad (13)$$

We find that near the absolute edge $\Omega \approx W_p$, the absorption is given by a power law with the index α that must be large within our adiabatic assumption, $\alpha \gg 1$,

$$I \sim \left(\frac{\delta\Omega}{|W_p|}\right)^{\alpha}, \quad \alpha = \frac{\lambda_s}{2\pi u}. \quad (14)$$

For incommensurate CDW parameters, we obtain $\alpha = v_F/4u$, in full accordance with the exact treatment (8).

Only in the absence of sound modes, $\lambda_s = 0$, the gapful contribution can determine the absolute edge. Minimization of $S = S_{core} + \delta S_{gap}$ over T then leads qualitatively to the result in [1],

$$T \sim \omega_0^{-1}\ln\left|\frac{W_p}{W_p - \Omega}\right|,$$
$$I \propto \exp\left(-\text{const}\frac{|W_p|}{\omega_0} + \frac{\Omega - W_p}{\omega_0}\ln\left|\frac{W_p}{\Omega - W_p}\right|\right) \quad (15)$$

for $\Omega \approx W_p$.

4.2. Free-Electron Edge Vicinity

We now consider the opposite regime near the free edge $\Omega \approx 0$ ($\Omega \longrightarrow \Omega - \Delta_0$ for the incommensurate CDW). Here, entering the pseudogap at $\Omega < 0$, the absorption is determined by fast processes of quantum fluctuations: their characteristic time $T = T(\Omega)$ is short compared to the relevant phonon frequency, $T \ll \omega_0, u/L$, where $L = L(\Omega)$ is the characteristic localization length for the fluctuational electron level at $E_0 = \Omega$. Because T is small, we can neglect all variations in time within $(0, T)$. We then estimate action (10), term by term, as

$$S \approx \frac{C_1 T}{mL^2} - \Omega T - C_2\lambda_s u\left(\frac{T}{L}\right)^2 - C_3\lambda_0\omega_0\frac{T^2}{L}, \quad (16)$$

where $C_i \sim 1$. The condition for its extremum with respect to both L and T yields

$$S \sim \frac{|\Omega|^{3/2}/m^{1/2}}{\max\{|m\Omega|^{1/2}u\lambda_s;\ \omega_0\lambda_0\}},$$

which provides a reasonable interpolation for the absorption in the closest and the more distant vicinities of the free-electron edge. For the purely acoustic case $\lambda_0 = 0$, a variational estimation of the numerical coefficient as $C_1 \approx 1/6$, $C_2 \approx 0.06$ gives

$$I \propto \exp(-\text{const}|\Omega|/mu\lambda_s), \quad \text{const} \approx 2.8. \quad (17)$$

The validity condition $uT/L \sim \sqrt{-\Omega/W_p} \ll 1$ is satisfied by definition of the edge region. This condition is compatible with the low boundary for the frequency, $S \gg 1$, and hence, $-\Omega/W_p \gg u/\lambda_s$, which is small as our basic adiabatic parameter.

For gapful phonons alone, $\lambda_s = 0$ and we arrive at the known result

$$I \propto \exp[-\text{const}|\Omega|^{3/2}/\omega_0], \quad \Omega < 0$$

(see [1] and references therein). However, it was not quite predictable that among the laws $S \propto |\Omega|^{3/2}$ and $S \propto |\Omega|$, it is the smallest contribution to S that wins, proportional to $|\Omega|^{3/2}$ at the lowest $|\Omega|$ and to $|\Omega|$ for larger $|\Omega|$. For the incommensurate CDW, in particular, we have $\lambda_0/\lambda_s \sim 1$ and $u/\omega_0 \sim \xi_0$, and there is no space for the intermediate asymptotic regime $\ln I \propto \Omega$ at $|\Omega| \ll \Delta_0$: beyond the region with $S \propto |\Omega|^{3/2}$, the amplitude fluctuations dominate, the phase-only description is invalid, and the particular nature of amplitude solitons must be taken into account. This regime was considered in Section 3.

The difference between the laws $\ln I \propto -|\Omega|/u$ and $\ln I \propto -|\Omega|^{3/2}/\omega_0$ can be interpreted easily. Indeed, for gapful phonons, we expect the frequency scale to be $\omega_0 \longrightarrow \omega_k = uk \sim u/L \sim u|\Omega|^{1/2}$, where $k \sim 1/L$ is a characteristic wave number and L is the localization length of the fluctuation providing the bound state at $-\Omega$. Then $|\Omega|^{3/2}/\omega_0 \longrightarrow |\Omega|^{3/2}/\omega_k \sim |\Omega|/u$.

While law (17) appears to be the simplest one, it is actually quite uncommon and its derivation is problematic in all systems (cf. [14]). In our case, we note that only at $\lambda_0 \neq 0$, action (16) has the usual saddle point, a minimum over L and a maximum over T. However, for the purely acoustic case $\lambda_0 = 0$, the minimum over L appears only along the extremal line over T. Contrarily, at a given T, the action collapses to either $L \longrightarrow 0$ or $L \longrightarrow \infty$ depending on the value of T with respect to the

threshold $T^* \sim (mu\lambda_s)^{-1}$, which is precisely the inverse width in (17). The paradox can be resolved by inspecting the generic real time formulation (2), but the necessary insight is obtained more easily by another treatment, presented in the next section.

4.3. Quantum Fluctuations as an Instantaneous Disorder with Long-Range Space Correlations

It has already been noticed that in a 1D system, the optical absorption near the band edge can be viewed as for a quenched disorder emulated by instantaneous quantum fluctuations. This asymptotically exact reduction to the time-independent model can be performed as follows. After neglecting the retardation at $T \ll \omega_0, u/L$, the self-interaction term in (10) can be decoupled by the Hubbard–Stratonovich transformation via a time-independent field ζ with the correlator $D(x) = U_0(x, 0) + \partial_x^2 U_s(x, 0)$,

$$S\{\Psi, \zeta; T\} = T\int dx\left(\frac{1}{2m}|\partial_x\Psi|^2 + \zeta(x)\rho(x)\right) + \frac{1}{2}\iint dx_{1,2}\zeta(x_1)D^{-1}(x_1 - x_2)\zeta(x_2). \quad (18)$$

After integration over Ψ and rotation to the real time, it finally becomes the expression for the density of states:

$$\int D[\zeta(x)]\delta(E[\zeta(x)] - \Omega) \times \exp\left[-\frac{1}{2}\iint dx_1 dx_2 \zeta(x_1)D^{-1}(x_1 - x_2)\zeta(x_2)\right],$$

where $E[\zeta(x)]$ is the eigenfunction in the random field ζ,

$$-\frac{\partial_x^2}{2m}\Psi + \zeta\Psi = E\Psi.$$

For the dispersionless phonon alone, e.g., the amplitude mode in the CDW, $D(x) = U_0(x, 0) \sim \delta(x)$, and the known exact results for the uncorrelated disorder [23] provide us with the asymptotic pseudogap formula

$$I(\Omega) \propto \exp\left[-\frac{8}{3^{3/2}}\frac{|W_p|}{\omega_0}\left|\frac{\Omega}{W_p}\right|^{3/2}\right]. \quad (19)$$

For the CDW parameters, it becomes

$$I(\Omega) \propto \exp\left[-\frac{16}{3\pi}\left(2\frac{\Delta_0 - \Omega}{(\Delta_0\omega_0^2)^{1/3}}\right)^{3/2}\right]. \quad (20)$$

Below, we concentrate only on a more problematic case of the sound mode. The correlator $D(x)$ of the "disordered potential" ζ is precisely the mean square of quantum fluctuations of the phonon potential $\zeta = v_F/2\varphi'(x, t)$ at coinciding times: in the Fourier representation, we have

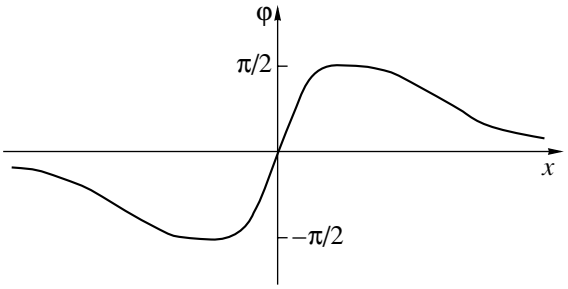

Fig. 3. The acoustic polaron field $\varphi(x, t)$ as a function of x at some moment t.

$$D_k = \int\frac{d\omega}{2\pi}\frac{\lambda_s k^2}{(\omega/u)^2 + k^2} = \frac{1}{2}\lambda_s u|k|.$$

The probability distribution for the Fourier components ζ_k is

$$P[\zeta_k] \propto \exp\left(-\int\frac{dk}{2\pi}\frac{|\zeta_k|^2}{\lambda_s u|k|}\right), \quad (21)$$

which implies that the component with $k = 0$ is excluded, $P(\zeta_0) = 0$. The constraint

$$\zeta_0 = \int_{-\infty}^{\infty}\zeta(x)\,dx = 0 \quad (22)$$

agrees with the properties of the potential proportional to φ' in the time-dependent picture of the previous section, which satisfies condition (22) at any finite t (see Fig. 3). Contrary to the usual expectations of the method of optimal fluctuations, the potential well creating the level E must here be accompanied by compensating barriers. Condition (22) is linked to the paradox in the previous section, i.e., the absence of a finite minimum over the length scale at a given T. Indeed, we can no longer rely on the existence of a bound state at an arbitrarily shallow potential, $E_0 \sim -m(\int dx\zeta(x))^2$, which is zero under condition (22).

While the divergence at small k (large x) is physical, the one at large k in (21) must be regularized to apply in the real space. We proceed by introducing an auxiliary field $\mu(x)$ such that $\zeta = d\mu/dx = \mu'$. We finally arrive at the model of the "nonlocal acoustic disorder,"

$$I(\Omega) \propto \int D[\mu(x)]\delta(E[\partial_x\mu(x)] - \Omega) \times \exp\left[-\frac{\lambda_s}{2u}\iint dx_1 dx_2\frac{(\mu(x_1) - \mu(x_2))^2}{|x_1 - x_2|^2}\right]. \quad (23)$$

Here, the integral in the exponent is already regular at small x. The divergence at large x maintains constraint (22), otherwise

$$\mu(+\infty) - \mu(-\infty) = \int_{-\infty}^{\infty} \zeta(x)dx \neq 0$$

and the integral in (23) would diverge logarithmically, leading to zero probability.

Unfortunately, we are unaware of exact studies for disordered systems with such long-range correlations. Usual scaling estimations [24] for characteristics μ and l give $|\Omega| \sim 1/ml^2 \sim |\mu|/l$, then $|\mu| \sim |\Omega/m|^{1/2}$, and therefore, $\ln I \sim -\mu^2\lambda_s/u \sim -|\Omega|\lambda_s/u$, in accordance with direct estimations and result (17) for the general time-dependent model.

5. DISCUSSION AND CONCLUSIONS

We summarize the obtained results as follows.

The pseudogap starts below the free edge by (stretched) exponential dependences

$$I \propto \exp[-\text{const}(-|\Omega|)^{\gamma}] \qquad (24)$$

with different powers $\gamma = 3/2$ for gapful phonons and $\gamma = 1$ for sound photons. If both modes are present, then the smallest one, with $\gamma = 3/2$, dominates at small Ω. This regime corresponds to free electronic states smeared by instantaneous uncorrelated quantum fluctuations of the lattice.

Deeply within the pseudogap, approaching the absolute threshold W_s or W_p, the exponential law changes for the power law $I(\Omega) \propto (\Omega - W_s)^{\beta}$ with a large exponent β. This contribution dominates over the smooth one from gapful modes $I \propto \exp(\text{const} \times \delta\Omega \ln \delta\Omega)$. The power-law regime corresponds to creation of nearly amplitude solitons surmounted be compensating phase tails. Its description provides a semiclassical interpretation for processes in fully quantum systems of correlated electrons in the spin-gap regime, with the amplitude soliton being a version of the spinon.

These results are different from anything used earlier in either theoretical discussions or interpretation of experimental data [13]. They can vaguely explain unusually wide pseudogaps observed in experiments even at low temperatures for well-formed incommensurate CDWs.

Our results have been derived for single electronic transitions, PES and tunneling. They can also be applied to intergap (particle–hole) optical transitions as long as semiconductors are concerned. For incommensurate CDWs, the results are applied to a vicinity of the free edge. However, the edge at $2E_s$ disappears in favor of the optically active gapless phase mode.

We emphasize in this respect that there cannot be a common pseudogap for processes characterized by different time scales. We must distinguish [8] between short-living states observed in optical, PES (and maybe tunneling) experiments, and long-living states (amplitude solitons and phase solitons) contributing to the spin susceptibility, NMR relaxation, heat capacitance, conductivity, etc. States forming the optical pseudogap are created instantaneously; particularly near the free edge, they are tested over times shorter than the inverse phonon frequencies $\tau_{\text{opt}} \sim \hbar/E_g < \omega_{\text{ph}}^{-1}$ and many orders of magnitude beyond the lifetimes required for current carriers, and even much longer times for thermodynamic contributions. It then follows that analysis of different groups of experimental data [13] within the same picture must be reconsidered. The lack of discriminating different time scales also concerns typical discussions of pseudogaps in high-T_c superconductors.

We conclude that the subgap absorption in systems with gapless phonons is dominated by formation of long space–time tails of relaxation. It applies to both acoustic polarons in 1D semiconductors and solitons in CDWs. Near the free edge, a simple exponential, Urbach-type law appears, competing with stretched exponential laws typical of tails from optimal fluctuations. A deeper part of the pseudogap is dominated by a power-law singularity near the absolute edge.

ACKNOWLEDGMENTS

One of the authors (S.M.) acknowledges hospitality of the Laboratoire de Physique Théorique et des Modèles Statistiques, Orsay and support of the CNRS and the ENS–Landau foundation. This work was partly performed within INTAS grant no. 2212.

APPENDIX

Electronic energies in a complex field Δ are determined by the Dirac Hamiltonian

$$\begin{vmatrix} -iv_F\partial_x & \Delta \\ \Delta^* & iv_F\partial_x \end{vmatrix}, \quad \Delta = |\Delta|e^{i\varphi}.$$

In the ground state, $|\Delta(x, t)| = \Delta_0$, we have $\varphi = \text{const}$, and the electronic spectrum is $E^2 = v_F^2k^2 + \Delta_0^2$, where v_F is the Fermi velocity. However, these free states are not proper excitations. The evolution of added electrons or holes with the initial energy $E_0 \geq \Delta_0$ to the amplitude soliton with $W_s = 2\Delta_0/\pi < \Delta_0$ can be described by an exact solution for intermediate configurations characterized by the singly occupied arbitrary positioned intragap state $E_0 = \Delta_0\cos\theta$ with $0 \leq \theta \leq \pi$, whence $-\Delta_0 < E_0 < \Delta_0$. It was found [8] to be the chord soliton

with 2θ being the total chiral angle, $\Delta(\pm\infty) = \exp(\pm i\theta)$ (see Fig. 1). Namely,

$$\Delta_{\mathrm{ChS}}(x, \theta) = \Delta_0[\cos\theta + i\sin\theta\tanh(k_0 x)]\exp(i\varphi_0), \quad k_0 = \Delta_0\sin\theta, \tag{25}$$

with an arbitrary $\varphi_0 = \text{const}$. The potentials V_ν are known [8] to be given by (see Fig. 2)

$$V_\nu(\theta) = \Delta_0\left[\left(\nu - \frac{2}{\pi}\theta\right)\cos\theta + \frac{2}{\pi}\sin\theta\right],$$

where ν is the filling number of the intragap state, that is, $\nu = 0, 1$ for $j = 0, 1$ while $\nu = 2$ is equivalent to $j = 0$ for the ground state extended by the two particles, $N = 2M + 2$. The term $V_0(\theta)$ monotonically increases from $V_0(0) = 0$ for the $2M$ ground state to $V_0(\pi) = 2\Delta_0$ for the $2M + 2$ ground state with two free holes. Obviously, there is an opposite dependence for $V_2(\theta) = V_0(\pi - \theta)$. Therefore, the total phases slip $2\theta = 0 \longrightarrow 2\theta = 2\pi$ realizes the spectral flow across the gap, also accompanied by the flow of particles for $\nu = 2$ that makes it favorable. The term $V_1(\theta) = V_1(\pi - \theta)$ is symmetric and describes both the particle on the $2M$ ground state and the hole on the $2M + 2$ ground state. Apparently, $V_1(0) = V_1(\pi) = \Delta_0$ (the degenerate ground states are the $2M$ one with an additional free electron for $\theta = 0$ and the $2M + 2$ one with an additional free hole for $\theta = \pi$), while the minimum is $V_1(\pi/2) = W_s < \Delta_0$, where $W_s = 2\Delta_0/\pi$ is the amplitude soliton energy. Therefore, the stationary state of the system with an odd number of particles, the minimum of V_1, is the amplitude soliton with the midgap state $E_0 = 0$ occupied by a single electron.

We note that being an uncharged spin carrier with the topological charge equal to unity, the amplitude soliton is a semiclassical realization of a spinon in systems with nonretarded attraction of electrons (that is, with high, rather than low, phonon frequencies).

REFERENCES

1. S. I. Matveenko and S. A. Brazovskii, Phys. Rev. B **65**, 245108 (2002).
2. G. Grüner, Rev. Mod. Phys. **60**, 1129 (1988).
3. P. A. Lee, T. M. Rice, and P. W. Anderson, Phys. Rev. Lett. **31**, 462 (1973).
4. S. Brazovskii and I. Dzyaloshinskiĭ, Zh. Éksp. Teor. Fiz. **71**, 2338 (1976) [Sov. Phys. JETP **44**, 1233 (1976)].
5. E. I. Rashba, in *Excitons*, Ed. by E. I. Rashba and M. D. Sturge (North-Holland, Amsterdam, 1982), p. 543.
6. A. Bjelis and S. Barisic, J. Phys. Lett. **36**, L169 (1975).
7. M. V. Sadovskii, Zh. Éksp. Teor. Fiz. **66**, 1720 (1974) [Sov. Phys. JETP **39**, 845 (1974)]; Zh. Éksp. Teor. Fiz. **77**, 989 (1979) [Sov. Phys. JETP **50**, 989 (1979)].
8. S. Brazovskii, Pis'ma Zh. Éksp. Teor. Fiz. **28**, 656 (1978) [JETP Lett. **28**, 606 (1978)]; Zh. Éksp. Teor. Fiz. **76**, 677 (1979) [Sov. Phys. JETP **51**, 342 (1980)].
9. H. P. Geserich *et al.*, Physica B (Amsterdam) **143**, 198 (1986); L. Degiorgi, G. Grüner, K. Kim, *et al.*, Phys. Rev. B **49**, 14754 (1994); Phys. Rev. B **52**, 5603 (1995).
10. F. Nad, in *Charge Density Waves in Solids*, Ed. by L. Gor'kov and G. Grüner (Elsevier, Amsterdam, 1989), p. 191.
11. G. Minton and J. W. Brill, Phys. Rev. B **36**, 4254 (1987).
12. G.-H. Gweon, J. W. Allen, R. Claessen, *et al.*, J. Phys.: Condens. Matter **8**, 9923 (1996); A. Terrasi, M. Marsi, H. Berger, *et al.*, Phys. Rev. B **52**, 5592 (1995); Phys. Rev. B **56**, 12647 (1997); L. Perfetti, H. Berger, A. Reginelli, *et al.*, Phys. Rev. Lett. **87**, 216404 (2001).
13. R. H. McKenzie and J. W. Wilkins, Phys. Rev. Lett. **69**, 1085 (1992); Phys. Rev. B **52**, 16428 (1995); K. Kim, R. H. McKenzie, and J. W. Wilkins, Phys. Rev. Lett. **71**, 4015 (1993).
14. A. S. Ioselevich, Zh. Éksp. Teor. Fiz. **81**, 1508 (1981) [Sov. Phys. JETP **54**, 800 (1981)]; Zh. Éksp. Teor. Fiz. **82**, 601 (1982) [Sov. Phys. JETP **55**, 358 (1982)]; Zh. Éksp. Teor. Fiz. **83**, 743 (1982) [Sov. Phys. JETP **56**, 415 (1982)].
15. A. Auerbach and S. Kivelson, Phys. Rev. B **33**, 8171 (1986).
16. R. Rajaraman, *Solitons and Instantons* (North-Holland, Amsterdam, 1989).
17. S. V. Iordanskiĭ and E. I. Rashba, Zh. Éksp. Teor. Fiz. **74**, 1872 (1978) [Sov. Phys. JETP **47**, 975 (1978)].
18. S. Brazovskii and N. Kirova, Sov. Sci. Rev., Sect. A **5**, 99 (1984).
19. S. Brazovskii, in *Charge Density Waves in Solids*, Ed. by L. Gor'kov and G. Grüner (Elsevier, Amsterdam, 1989), p. 425.
20. S. A. Brazovskii and S. I. Matveenko, Zh. Éksp. Teor. Fiz. **99**, 887 (1991) [Sov. Phys. JETP **72**, 492 (1991)].
21. A. O. Caldeira and A. J. Leggett, Phys. Rev. Lett. **46**, 211 (1981).
22. S. Brazovskii, E-print archives, cond-mat/0204147, cond-mat/0006355.
23. B. Halperin, Phys. Rev. A **139**, 104 (1965).
24. I. M. Lifshits, S. A. Gredeskul, and L. A. Pastur, *Introduction to the Theory of Disordered Systems* (Nauka, Moscow, 1982; Wiley, New York, 1988).

Journal of Experimental and Theoretical Physics, Vol. 96, No. 3, 2003, pp. 564–571.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 123, No. 3, 2003, pp. 635–642.

SOLIDS
Electronic Properties

Motion of a 3D Exciton in a Magnetic Field: Exciton–Magnetoexciton "Phase" Transition

Yu. E. Lozovik* and S. Yu. Volkov

Institute of Spectroscopy, Russian Academy of Sciences, Troitsk, Moscow oblast, 142190 Russia

**e-mail: lozovik@isan.troitsk.su*

Received October 9, 2002

Abstract—A rearrangement of the ground state of a Wannier–Mott exciton upon an increase in its momentum is considered. The phase diagram of the electron and the hole experiencing the Coulomb interaction on the magnetic momentum–external magnetic field plane is investigated. A jumplike exciton–magnetoexciton "phase" transition is observed upon an increase in the momentum in fields B weaker than a certain value $B < B_{tr1}$. As momentum P increases above a certain critical value $P_{tr}(B)$, the ground state of the system changes from the hydrogen-like state polarized by the Lorentz force to the magnetoexciton state in which the average distance $\langle \mathbf{r} \rangle$ between the electron and the hole increases jumpwise in the transverse direction relative to the field. As the exciton momentum increases, its wave function is extended along the magnetic field, acquiring the shape of a strongly prolate ellipsoid. It is interesting that the momentum of the transition tends to a finite value $P_0 > 0$ even for $B \longrightarrow 0$. At the point of transition, the exciton energy–momentum relation changes jumpwise from a quadratic law to a relation virtually independent of the momentum. For $B < B_{tr1}$, the exciton–magnetoexciton transition becomes blurred. © *2003 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

It is well known that the ground state of a Wannier–Mott exciton changes significantly in strong magnetic fields in which the characteristic interaction with the field (separation $\hbar\omega$ between the Landau levels) becomes much larger than the exciton binding energy Ry* [1–6].

Since the Schrödinger equation is invariant to simultaneous translation of the electron and the hole and to the corresponding gauge transformation (see [6]), there exists a three-dimensional vector, viz., "magnetic momentum" **P**, which is transformed into the conventional momentum in zero magnetic field [3]. The exciton energy–momentum relation $E(\mathbf{P})$ in very strong magnetic fields changes significantly; in particular, the "magnetic mass" $m_\perp(\mathbf{B}) = (\partial^2 E/\partial \mathbf{P}_\perp^2)^{-1}$ determining the exciton dispersion for small values of **P** in directions transverse to the field noticeably increases [4–6]. In addition, the curve describing the exciton energy as a function of $\mathbf{P}_\perp$ in a strong magnetic field has a point of inflection, and the energy–momentum relation acquires the form $E(\mathbf{P}_\perp) \propto 1/P_\perp$ instead of the quadratic form. This determines the characteristic of a magnetoexciton, viz., an exciton in strong magnetic fields.

Is the rearrangement of an exciton into a magnetoexciton just a blurred transition (crossover)? We will study here the energy–momentum relation of an exciton in arbitrary magnetic fields and show that a jumplike transition occurs in weak magnetic fields, while crossover takes place in magnetic fields exceeding a certain critical value.

We analyze the behavior of a 3D exciton with a nonzero momentum in the entire range of magnetic fields, including relatively weak fields. In our calculations, we use the method of numerical solution of the Schrödinger equation in imaginary time. In fields smaller than a certain critical value, there are two different modes for the ground state of the exciton depending on its momentum. In the first mode (weak magnetic field and small momentum), the ground state of the exciton corresponds to a weakly polarized hydrogen-like function since the Coulomb interaction between the electron and the hole plays the major role for the effective potential describing the exciton internal structure. In the second mode (momenta exceed the critical values), the main contribution to the effective potential comes from the terms describing the interaction of the electron and the hole with the external magnetic field (magnetoexciton mode). In weak magnetic fields, an increase in the exciton momentum turns out to cause a jumplike transition between these two modes, and the wave function of the exciton "jumps" from the Coulomb minimum to the magnetic minimum of the effective potential. As a result of this transition, the width of the wave function increases sharply from the characteristic size of a hydrogen-like state, viz., effective Bohr radius $a^* = \hbar^2\epsilon/2\mu e^2$, to the characteristic size of a magnetoexciton, viz., the magnetic length $l = \sqrt{\hbar/\mu\omega_c}$ (it

should be recalled that we are speaking of weak fields in which $l \gg a^*$). Here, $\omega_c = eB/2\mu c$ is the cyclotron frequency, $\mu = m_e m_h/(m_e + m_h)$ is the reduced mass (m_e and m_h are the masses of the electron and the hole), and ϵ is the permittivity of the material. In this direction of the magnetic field, the exciton wave function is extended significantly. This is due to the fact that the effective potential of the problem has only one (Coulomb) minimum in the direction of the magnetic field, but the wave function in this case is localized far from the characteristic region in which the Coulomb interaction is strong (far from the region of zero spacing between the electron and the hole); for this reason, the state is weakly connected in the direction of the field.

The dependence of the mean distance between the electron and the hole on exciton momentum, which is connected to polarization of a moving exciton, also reflects this structural transformation of the wave function upon an increase in its momentum. According to calculations, for an exciton momentum smaller than a certain critical value P_{tr}, its polarization is a linear function of the momentum and the coefficient in this dependence is associated with the exciton polarizability. For a momentum exceeding P_{tr}, the polarization is also a linear function, but with a different coefficient. This is due to the fact that the distance between the effective potential minima for $P > P_{tr}$ is also a linear function of the momentum, but with a considerably larger coefficient, $\langle r \rangle = cP/eB_\perp$ (which increases upon a decrease in the magnetic field) as in the case of an exciton in very strong magnetic fields. In a small region of the momentum in the vicinity of the transition (near P_{tr}), this dependence is nonlinear. This is due to the fact that, in the vicinity of the transition region, the wave functions concentrated in one of the minima are strongly extended towards the other minimum. For small momenta in arbitrary magnetic fields, the exciton energy–momentum relation is quadratic in the magnetic momentum with a coefficient depending on the field. We have determined the magnetic field dependence of the exciton effective mass from the calculated dispersion.

2. MODEL

Let us consider an electron and a hole in a magnetic field. The initial Hamiltonian has the form

$$H = \frac{1}{2m_e}\left[i\hbar\frac{\partial}{\partial \mathbf{r}_e} - \frac{e}{c}\mathbf{A}(\mathbf{r}_e)\right]^2 + \frac{1}{2m_h}\left[i\hbar\frac{\partial}{\partial \mathbf{r}_h} + \frac{e}{c}\mathbf{A}(\mathbf{r}_h)\right]^2 - \frac{1}{\epsilon\sqrt{(\mathbf{r}_e - \mathbf{r}_h)^2}}.$$

It was mentioned above that the Schrödinger equation is invariant to simultaneous translation of the electron and the hole and to the corresponding gauge transformation [6]. The infinitesimal operator of this transformation commutes with the Hamiltonian, which corresponds to the magnetic momentum conservation law. Let the magnetic field be directed along the z axis. In the cylindrical gauge $\mathbf{A} = [\mathbf{B} \times \mathbf{r}]/2$, the magnetic momentum has the form [3, 4]

$$\mathbf{P}_{2D} = -i\hbar\frac{\partial}{\partial \mathbf{R}_{2D}} + \frac{e}{2c}[\mathbf{B} \times \mathbf{r}], \quad P_z = -i\hbar\frac{\partial}{\partial z}.$$

Here, we have made the substitution

$$\mathbf{r} = \mathbf{r}_h - \mathbf{r}_e, \quad \mathbf{R} = m_e\mathbf{r}_e + m_h\mathbf{r}_h.$$

Subscript $2D$ on the vectors indicates that we are using only the x and y components of a vector, which are transverse to the magnetic field.

We are seeking steady-state solutions, which are the eigenfunctions of the magnetic momentum:

$$\psi'(\mathbf{R}, \mathbf{r}) = \exp\left(-i\frac{e}{2c}[\mathbf{B} \times \mathbf{R}] \cdot \mathbf{r}\right)\psi(\mathbf{R}, \mathbf{r}).$$

Ultimately, the Hamiltonian of the system for function ψ assumes the form

$$H = \frac{1}{2M}\left(\frac{\mu}{\eta}\right)^2\mathbf{P}_{2D}^2 - \frac{\hbar}{2\mu}\Delta + i\frac{eB\hbar}{2c\eta}L_z + \frac{1}{2\mu}\left(\frac{eB}{2c}\mathbf{r}_{2D} - \frac{2\mu}{M}\mathbf{P}_{2D}\right)^2 - \frac{1}{\epsilon|\mathbf{r}|},$$

where

$$\eta^{-1} = m_e^{-1} - m_h^{-1}, \quad L_z = x\frac{\partial}{\partial y} - y\frac{\partial}{\partial x}, \quad M = m_e + m_h.$$

We choose dimensionless units for the coordinate, energy, and momentum:

$$r_0 = \frac{\epsilon\hbar^2}{2\mu e^2}, \quad E_0 = \frac{e^2}{\epsilon r_0}, \quad P_0 = \frac{Me^2}{\epsilon\hbar}$$

(all results that are given below, including the figures, are expressed in these units). For a dimensionless controlling parameter, we take the ratio of the characteristic Coulomb and magnetic energies: $b = \hbar\omega c/E_0$. The other controlling parameter of the equation is the dimensionless momentum.

In dimensionless variables, the Hamiltonian has the form

$$H = \frac{1}{4}\left(\frac{\mu}{\eta}\right)^2 \frac{M}{\mu}\mathbf{P}_{2D}^2 - \Delta + ib\frac{\mu}{\eta}L_z + \left(\frac{b}{2}\mathbf{r}_{2D} - \mathbf{P}_{2D}\right)^2 - \frac{1}{|\mathbf{r}|}. \quad (1)$$

3. COMPUTATIONAL METHOD

We must find the wave function of the ground state and its eigenvalue. In the time-dependent Schrödinger equation

$$i\hbar\frac{\partial}{\partial t}\psi = H\psi,$$

we formally pass to the imaginary time $t' = it$. In this case, the equation assumes the form

$$\hbar\frac{\partial}{\partial t}\psi = -H\psi. \quad (2)$$

Obviously, for $t \longrightarrow \infty$ and for any initial condition, the solution to this equation,

$$\psi(\mathbf{r}, t) = \sum_n C_n\psi_n(\mathbf{r})\exp(-E_n t),$$

tends to the wave function $\psi_0(\mathbf{r})$ of the ground state. Here, $\psi_n(\mathbf{r})$ and E_n are eigenfunctions and eigenvalues of operator H. Consequently, we can solve Eq. (2) by the following method of stabilization (we are seeking a steady-state solution). We will use the explicit algorithm

$$\psi^{n+1} = \psi^n - dt, H\psi^n.$$

We proceed further as follows. First, we set a reasonable approximation ψ^0 of the wave function as the initial state. It can be, for example, a function of the type $\exp(-x^2)$ since the solution to our equation for the ground state has a similar form. This is done to reduce the computation time. Then we carry out an iteration in time and normalize the next approximation (ψ^1) of the wave function:

$$\int \psi^{1*}\psi^1 d\mathbf{r} = 1;$$

this procedure is repeated until the wave function stabilizes (does not vary significantly).

The energy can be determined from the obtained wave function as the mean value of the Hamiltonian in this function:

$$E_0 = \int \psi_0^* H\psi_0 d\mathbf{r}.$$

4. DISCUSSION

We carried out detailed simulation of a 3D exciton for various values of magnetic field and momentum in order to study its "phase" diagram. For concreteness, we took the parameters of the problem for GaAs:

$$\epsilon = 12.3, \quad m_e^* = 0.07m_e, \quad m_h^* = 0.15m_e.$$

As regards the qualitative analysis that will be described below, it is more general in nature.

The effective potential of the exciton (the last two terms in dimensionless Hamiltonian (1)) is determined by two competing interactions: the Coulomb electron–hole potential and the interaction of the electron and the hole with the external magnetic field. Qualitative analysis of the behavior of this effective potential of the exciton upon an increase in its momentum (Fig. 1) helps to grasp the behavior of the wave function, thus explaining the structural transformations experienced by the exciton upon an increase in momentum P. It is convenient to depict schematically these changes in the exciton internal structure upon a change in the magnetic field and momentum in the form of a "phase" diagram of an exciton in a magnetic field in the B–P plane (Fig. 2); we will consider this diagram in detail.

Let us first consider the case of weak magnetic fields ($B_{tr1} > B > 0$) in Fig. 2. For large values of momentum, the effective potential acquires two minima (cf. Figs. 1a, 1b and Figs. 1c, 1d): the first minimum is associated with the Coulomb interaction between the electron and the hole, while the second is due to the parabolic magnetic potential. Each of these minima is associated with a certain set of stationary states. For a large distance between the minima (i.e., in weak magnetic fields), the overlap of the wave functions of the lower states is negligibly small. For momenta $P < P_{tr}$, where

$$P_{tr} \approx \sqrt{2M\left(E_b + \frac{\hbar\omega_c}{2}\right)},$$

the Coulomb energy level E_C lies below the magnetic level E_m (Fig. 3). Thus, for small values of momentum, the hydrogen-like state is the ground state of the exciton, while the state at the magnetic minimum is an excited state. With increasing momentum, the arrangement of energy levels at the minima changes and the lower level at the magnetic minimum becomes the ground state for $P > P_{tr}$; this corresponds to a magnetoexciton. In view of weak overlapping of the wave functions, the repulsion of the corresponding levels is small even in the resonance case, i.e., at the point of quasi-intersection of the quadratic dependence $E(P)$, corresponding to the Coulomb minimum, with the magnetoexciton energy $E \approx \hbar\omega_c$, corresponding to the bound state at the magnetic minimum (at zero Landau level). Point P_{tr} corresponds to the intersection of the

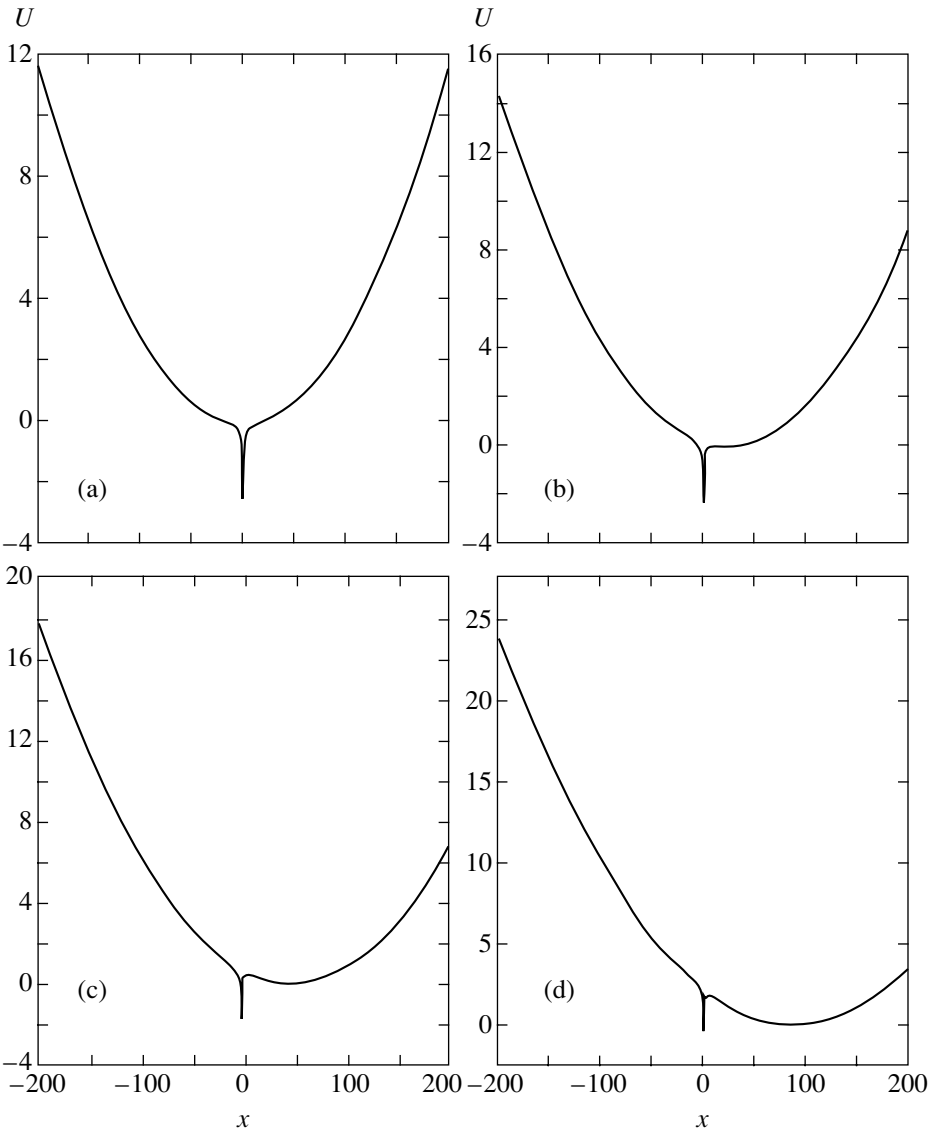

Fig. 1. Transformation of the effective potential upon an increase in momentum in field $B = 0.5$ T for $P = 0.1$ (a), 0.4 (b), 0.8 (c), and 1.5 (d). For better visualization, the $U(x, y, z)$ curves are plotted for $y = 0$ and $z = 0$. For small values of P, the effective potential has only a "Coulomb" minimum at point $x = 0$. As the momentum increases, a new ("magnetic") minimum also appears.

curves describing the energies of the ground states at the local (Coulomb and magnetic) minima of the effective potential (see Fig. 3). After the transition, the state at the Coulomb minimum becomes excited. The following interesting fact is worth noting: the exciton momentum at the point of transition differs from zero even when the magnetic field tends to zero (cf. expression for P_{tr}). A sharp transition in momentum is present up to the critical value of the magnetic field B_{tr1}, at which each local minimum is associated with its own set of bound states. In fields stronger than B_{tr1}, there are no bound states at the individual minima of the effective potential and the wave function of the exciton ground state covers both minima. In this case, the transition is blurred into a crossover from the region with the Coulomb mode to the magnetoexciton mode. The width of the crossover region increases with the field. Finally, for $B = B_{tr2}$, the magnetoexciton is the ground state even for small values of momentum.[1]

Figure 4 shows the results of calculation of the exciton ground state energy as a function of momentum for different values of the magnetic field. For small values of momentum, the energy in arbitrary fields increases in accordance with a quadratic law. As the momentum increases further, the curve describing the dependence of energy on momentum acquires a point of inflection

[1] It should be noted that, in contrast to the two-dimensional indirect exciton considered in [5], the effective potential for a 3D exciton has two minima and not one for large momenta even in strong fields.

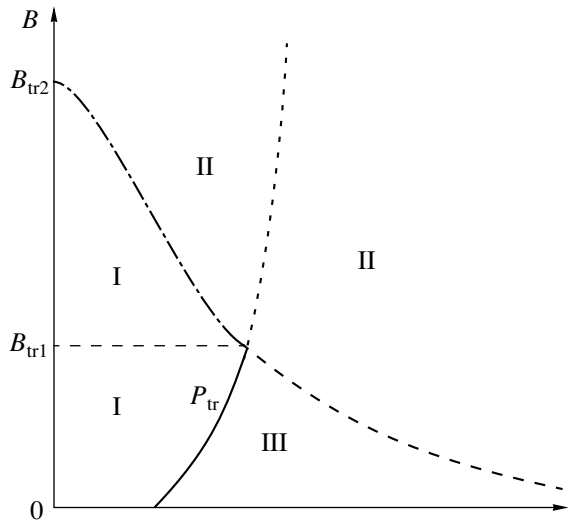

Fig. 2. "Phase" diagram of an exciton. Roman numerals indicate the domains of existence of the "Coulomb" exciton (I), magnetoexciton (II), and magnetoexciton (ground state) for a metastable "Coulomb" exciton (III).

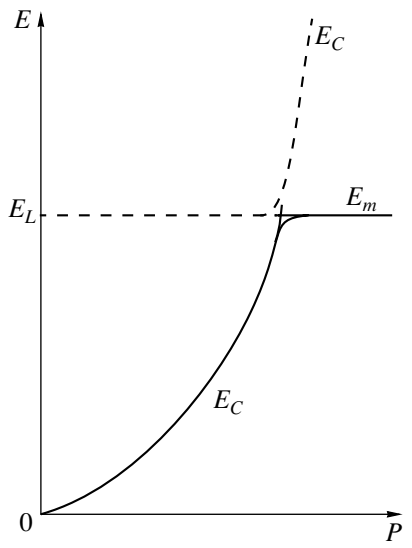

Fig. 3. Momentum dependence of energy at the local minima of the effective potential: transition from the Coulomb regime to a magnetoexciton. Similar quasi-intersections are also observed for excited levels (E_L is the Landau level).

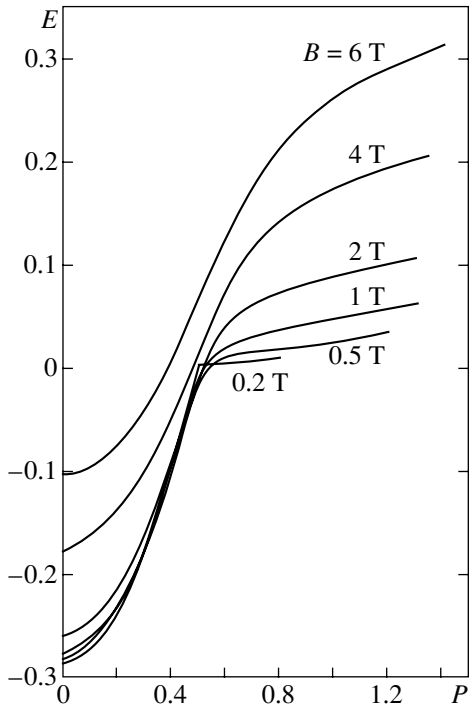

Fig. 4. Exciton energy as a function of momentum for different values of magnetic field.

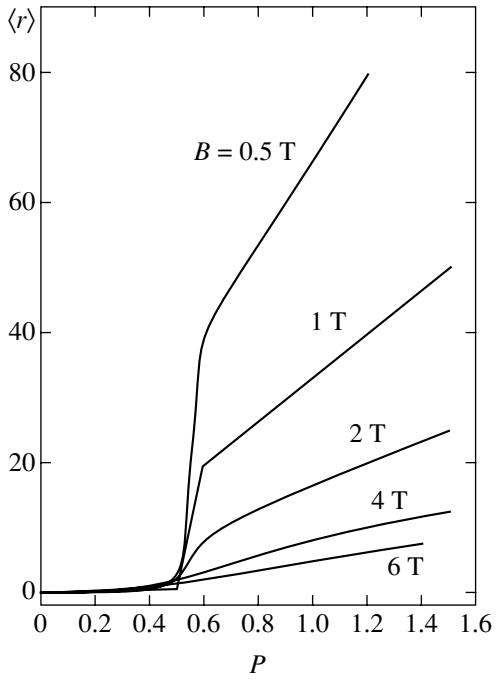

Fig. 5. Average distance $\langle r \rangle$ between the electron and the hole as a function of momentum in different magnetic fields.

as in strong magnetic fields [4, 6], after which the energy remains virtually unchanged. The sharpest variation is observed for $B < B_{tr1}$. For large values of the field, this transition is blurred into a crossover, and the $E(P)$ dependence changes smoothly from a quadratic law to a constant value corresponding to the Landau level. Such a behavior of the exciton energy is associated with a change in the structure of the effective potential (see above).

It is interesting to analyze the behavior of the exciton wave function. For values of the exciton momentum from a small neighborhood of P_{tr}, the wave function, which is at one of the minima, begins to be deformed with increasing P so that it is extended in the direction

of the other minimum of the effective potential. In this transition, the distance between the electron and the hole increases sharply to a value of $\langle r \rangle = cP/eB$ characterizing the regime of a strong magnetic field (see Figs. 5 and 6) [4, 6], and the wave function shifts from the Coulomb minimum to the magnetic minimum. As a result of the transition, the exciton wave function is strongly extended in the direction of the magnetic field. As a matter of fact, the effective potential has only one (Coulomb) minimum in the magnetic field direction; in this case, the wave function is localized far from the characteristic region in which the Coulomb interaction potential is high. Consequently, the state in the direction of the field is weakly connected (see Fig. 7). As the momentum increases, the wave function resembles a strongly prolate ellipsoid. We can estimate the ratio of the widths of the wave function after the transition (characteristic regions of exciton localization) in planes *xz* (along the field) and *xy* (across the field). Only one (Coulomb) minimum exists along the *z* axis (in the direction of the magnetic field); in this case, the characteristic size of localization of the exciton wave function in the direction of the field is determined by the potential

$$U(z) = -\frac{e^2}{\sqrt{z^2 + \rho_0^2}} \approx -\frac{e^2}{\rho_0}\left(1 - \frac{z^2}{2\rho_0^2}\right),$$

where $\rho_0 = 2P/B$. Obviously, the characteristic localization region in this potential is given by

$$R_z \approx \rho_0 \sim P/B.$$

The characteristic size R_{xy} of the wave function in the *xy* plane is determined by the magnetic minimum since the wave function is concentrated just at this minimum after the transition:

$$R_{xy} \approx \sqrt{\frac{h}{m\omega_c}} \propto \sqrt{\frac{1}{B}}.$$

The ratio of these sizes is

$$\frac{R_z}{R_{xy}} \propto \frac{P}{\sqrt{B}}.$$

The rearrangement of the exciton internal structure upon a change in the magnetic field and the exciton momentum (sharp transition in weak fields and crossover in strong fields) also affects the behavior of the exciton polarization. In connection with Fig. 5, which describes the change in the position of the maximum of the exciton wave function upon an increase in momentum *P*, we can make the following remarks. For $P \neq 0$, polarization of the exciton takes place. Indeed, when

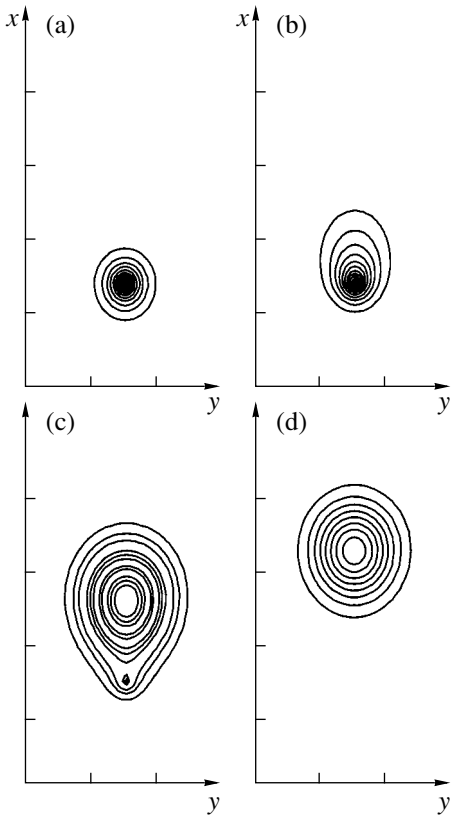

Fig. 6. Transformation of the wave function of the ground state of an exciton in the *xy* plane upon an increase in the momentum depending on the magnetic field for $P = 0$ (a), 0.5 (b), 0.8 (c), and 1.0 (d); $B = 1$ T.

the exciton moves in the frame of reference associated with the center of mass, the electric field

$$\mathbf{E} = \frac{1}{mc}[\mathbf{P} \times \mathbf{B}]$$

is induced. For small values of *P*, field *E* is weak and the response to the electric field is linear in this region:

$$\mathbf{d} = e\langle \mathbf{r} \rangle = \alpha(\mathbf{E}) = \frac{\alpha}{mc}[\mathbf{P} \times \mathbf{B}],$$

where α is the polarizability of the hydrogen-like state. The linear dependence of the induced dipole moment on *P* corresponds to the linear dependence of the position of the maximum of the exciton wave function (in relative coordinates) on *P*. The region of this linear dependence extends to values of *P* for which the Coulomb interaction suppressed due to the shift of the effective potential minimum becomes equal (in order of magnitude) to the interaction with the magnetic field. In this region, the effective potential for the exciton becomes noticeably distorted as compared to the initial Coulomb potential well. For values of momentum exceeding $P = P_{tr}$, the polarization remains linear, but

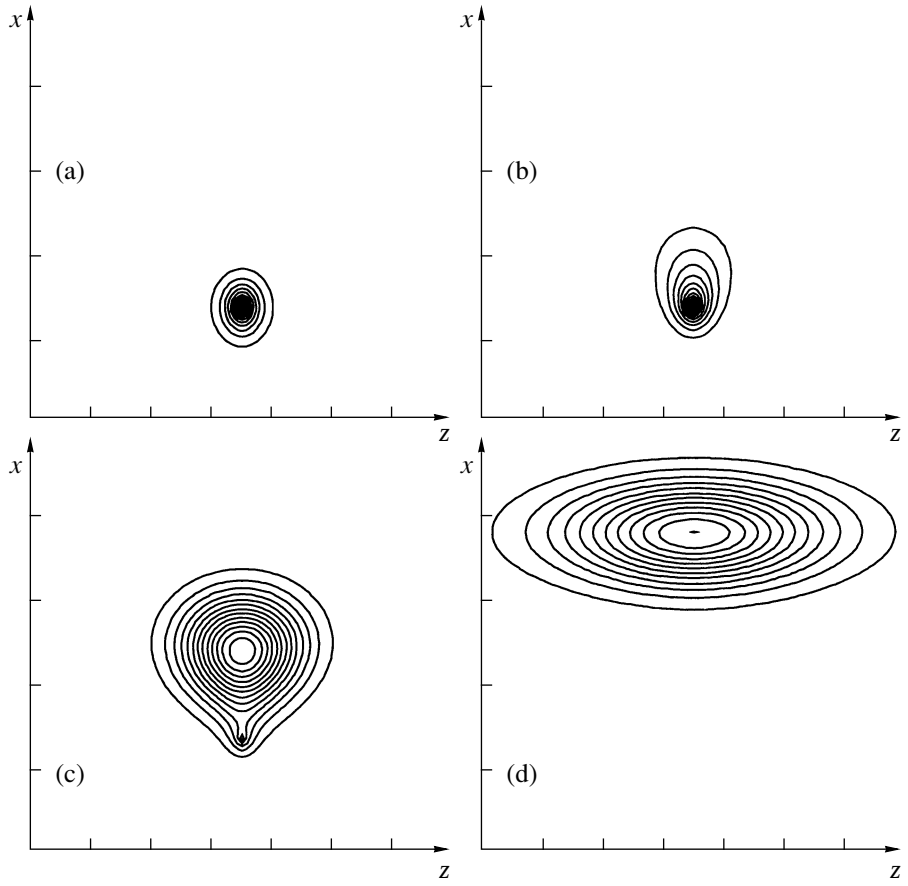

Fig. 7. Transformation of the wave function of the ground state of an exciton in the *xz* plane upon an increase in the momentum depending on the magnetic field for $P = 0$ (a), 0.5 (b), 0.8 (c), and 1.4 (d); $B = 2$ T.

the coefficient of this dependence changes; i.e., the average distance $\langle r \rangle$ increases sharply. In the region of crossover, i.e., in magnetic fields $B > B_{tr1}$ in the vicinity of the momentum P_{tr}, the law of polarization differs significantly from a linear dependence. This is due to the fact that the exciton wave functions concentrated at one of the minima in this region of the phase diagram are strongly deformed so that they are extended in the direction of the other minimum.

In the range of small momenta, the energy–momentum relation for excitons can be described by a parabola (see above) whose parameter is the effective mass of an exciton in a magnetic field, $m_{eff} = 1/2\beta$, where β is the slope of this parabola. Figure 8 shows the dependence of the exciton effective mass on the magnetic field. In zero magnetic field, the effective mass is equal to the initial mass $M = m_e^* + m_h^*$, as expected. As the magnetic field becomes stronger, the exciton effective mass increases monotonically (see Fig. 8).

It should be noted that a similar sharp rearrangement of the structure in magnetic fields weaker than atomic fields $B_0 = m^2ce^3/\hbar^3 \approx 2 \times 10^5$ T is also possible for an atom moving at a very high velocity.

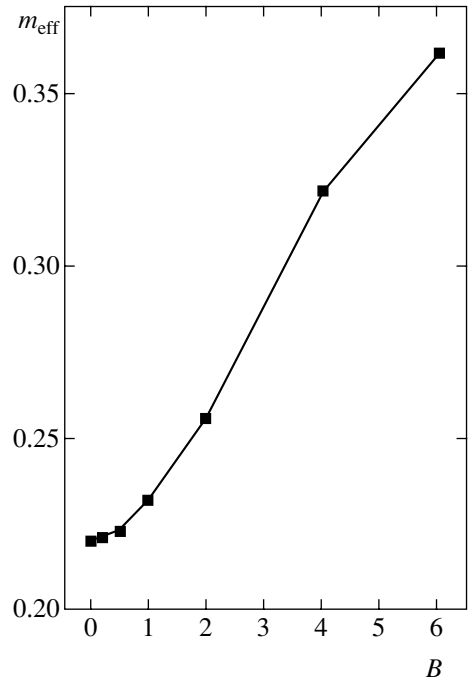

Fig. 8. Effective mass m_{eff} of an exciton as a function of magnetic field *B*.

ACKNOWLEDGMENTS

The authors thank the Russian Foundation for Basic Research and INTAS for the support of this study.

REFERENCES

1. R. J. Elliott and R. Loudon, J. Phys. Chem. Solids **8**, 382 (1959); J. Phys. Chem. Solids **15**, 196 (1960).
2. H. Hasegawa and R. E. Howard, J. Phys. Chem. Solids **21**, 179 (1961).
3. L. P. Gor'kov and I. E. Dzyaloshinskiĭ, Zh. Éksp. Teor. Fiz. **53**, 717 (1967) [Sov. Phys. JETP **26**, 449 (1968)].
4. I. V. Lerner and Yu. E. Lozovik, Zh. Éksp. Teor. Fiz. **78**, 1167 (1980) [Sov. Phys. JETP **51**, 588 (1980)].
5. L. V. Butov, C. W. Lai, D. S. Chemla, *et al.*, Phys. Rev. Lett. **87**, 216804 (2001); L. V. Butov, A. V. Mintsev, Yu. E. Lozovik, *et al.*, Phys. Rev. B **62**, 1548 (2000); Yu. E. Lozovik, I. V. Ovchinnikov, S. Yu. Volkov, *et al.*, Phys. Rev. B **65**, 235304 (2002).
6. Yu. E. Lozovik and A. M. Ruvinsky, Zh. Éksp. Teor. Fiz. **112**, 1791 (1997) [JETP **85**, 979 (1997)]; Phys. Lett. A **227**, 271 (1997).

Translated by N. Wadhwa

Journal of Experimental and Theoretical Physics, Vol. 96, No. 3, 2003, pp. 572–580.
From Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 123, No. 3, 2003, pp. 643–652.

REVIEWS

Stochastic Processes of Demarkovization and Markovization in Chaotic Signals of Human Brain Electric Activity from EEGs during Epilepsy¶

R. M. Yulmetyev[a,*], P. Hänggi[b], and F. M. Gafarov[a]

[a]*Kazan State Pedagogical University, Kazan, 420021 Russia*

**e-mail: rmy@kasan-spu.ru*

[b]*University of Augsburg, D-86135 Augsburg, Germany*

Received July 19, 2002

Abstract—We study the stochastic processes of markovization and demarkovization in chaotic signals of human electroencephalograms (EEGs) during epilepsy using various measures of demarkovization and markovization, namely, the statistical spectrum of a non-Markovity parameter, power spectra of the time correlation function and memory functions of junior orders, and local relaxation and kinetic parameters. The results demonstrate the superiority of the new measures in comparison to the traditional nonlinear measures. We conclude that the applied measures are more appropriate for the quantification of markovization and demarkovization in EEG data and the prediction of epilepsy seizure. © *2003 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

We develop a new approach that could provide us with a powerful means of discrete time series analysis and processing. The subject of our study is human electroencephalogram (EEG) records, because we address our work to those who are interested in signal processing in live complex systems. In studying natural complex systems, very little is usually known about their internal structure and the relationship between their components. The time series describing the dynamics of one or several parameters are typically used for obtaining diagnostic information. The received information is inadequate for describing all the degrees of freedom of this system. Quantitative and qualitative methods proposed recently allow constructing the framework for the description of natural complex systems. It allows diagnosing diseases without going into detail of the internal structure underlying natural complex systems. A similar approach can be used to describe and investigate diversified complex systems as they are related only to the concepts of this framework. Here, we present the results of applying a new framework involving ideas of discrete non-Markovian stochastic processes to the analysis of electric potentials of brain. It turns out that discussing the results in terms of demarkovization and markovization is the best way to uncover the features of seizure dynamics.

Brain cells communicate by producing tiny electrical impulses. In an EEG, electrodes are placed on the scalp over multiple areas of the brain to detect and record the electrical pulses within the brain. The EEG is used to help diagnose the presence and type of seizure disorders, confusion, head injuries, brain tumors, infections, degenerative diseases, and metabolic disturbances that affect the brain.

It is well known that epilepsy is one of the most serious diseases of the human brain [1, 2]. The dynamics of the electric signals accompanying it belongs to a class of nonlinear, nonstationary, and nonergodic processes of complex systems of a live nature [3, 4]. The discrete and non-Markovian properties of time variation of the signals and the sudden alternation of the behavior regimes must be taken into account in analyzing the electrical activity of brain potentials. Together with the fast change of chaotic and regular modes in the behavior of the system, this creates serious problems for the diagnosis and treatment of patients with epilepsy seizure. This is why traditional methods of nonlinear dynamics, such as the Lyapunov exponent, Kolmogorov–Sinai entropy, and correlation and fractal dimensions, are not sufficiently sensitive for the purpose of distinction between different chaotic regimes in epilepsy.

2. BASIC THEORY

Our approach is based on the recent theory for stationary [5] and nonstationary cases [6] of discrete stochastic processes in complex systems. We analyze the stochastic process on the basis of the chain of the cou-

¶This article was submitted by the authors in English.

pled non-Markovian discrete equations for the initial discrete time correlation function (TCF) $a(t)$ ($t = m\tau$),

$$\begin{aligned}\frac{\Delta a(t)}{\Delta t} &= \lambda_1 a(t) - \tau\Lambda_1 \sum_{j=0}^{m-1} M_1(j\tau)a(t-j\tau),\\ \frac{\Delta M_1(t)}{\Delta t} &= \lambda_1 M_1(t) - \tau\Lambda_2 \sum_{j=0}^{m-1} M_2(j\tau)M_1(t-j\tau),\\ \frac{\Delta M_2(t)}{\Delta t} &= \lambda_3 M_2(t) - \tau\Lambda_3 \sum_{j=0}^{m-1} M_3(j\tau)M_2(t-j\tau),\end{aligned} \quad (1)$$

where λ_n is the eigenvalue spectrum of the Liouville operator $i\hat{L}$ and Λ_n are the general relaxation parameters,

$$\lambda_n = i\frac{\langle \mathbf{W}_n^* L \mathbf{W}_n\rangle}{\langle|\mathbf{W}_n|^2\rangle}, \quad \Lambda_n = -\frac{\langle \mathbf{W}_{n-1}(i\hat{L} - \lambda_{n+1})\mathbf{W}_n\rangle}{\langle|\mathbf{W}_{n-1}|^2\rangle}.$$

The kinetic nonlinear finite-difference Eqs. (1) are analogous to the well-known chain of kinetic equations of the Zwanzig–Mori (ZM) type. These ZM equations play a fundamental role in the modern statistical mechanics of nonequilibrium phenomena with continuous time. Kinetic Eqs. (1) can be considered a discrete-difference analogy of hydrodynamic equations for physical phenomena with discrete time. By analogy with [5–7], we define the generalized nonlinear non-Markovity parameter in the frequency-dependent case as

$$\epsilon_i(\omega) = \left\{\frac{\mu_{i-1}(\omega)}{\mu_i(\omega)}\right\}^{1/2}, \quad (2)$$

where $i = 1, 2, \ldots$ and $\mu_i(\omega)$ is the power spectrum of the ith memory function. It is convenient to use this parameter for quantitative description of long-range memory effects in the system considered together with memory functions defined above. The values of $\epsilon_i(\omega)$ allow us to obtain a quantitative estimate of non-Markovity effects and the statistical collective memory in the chaotic changes of the experimentally measured EEG data. The parameters $\epsilon_i(\omega)$ allow classification of all the observed processes into three important types [5]. A Markov process corresponds to the situation where the non-Markovity parameter takes an indefinitely large value $\epsilon_i(\omega) \longrightarrow \infty$, and the quasi-Markov processes correspond to the case where $\epsilon_i(\omega) > 1$. The limit case $\epsilon_i(\omega) \approx 1$ describes non-Markovian processes. In this case, the time scale of memory processes and the correlation dynamics (or the nearest junior and senior memory function) coincide with each other.

3. EXPERIMENTAL DATA

We quantitatively demonstrate the stochastic description of the time–frequency peculiarities of epilepsy. We use experimental data [8] on human EEGs. These files show tonic–clonic seizures of two subjects recorded with a scalp-right-central (C4) electrode (linked earlobes reference). It contains a total of 3 min with about 1 min preseizure, the seizure, and some postseizure activity. The sampling rate is 102.4 Hz (see the papers cited in [8] for more details).

4. NUMERICAL CALCULATIONS

We consider a discrete time series of the electric activity as a one-point stochastic process

$$\begin{aligned}X = \{&x(T), x(T+\tau), x(T+2\tau), \ldots,\\ &x(T+k\tau), \ldots, x(T+\tau N-\tau)\}.\end{aligned} \quad (3)$$

It is convenient to introduce the normalized time correlation function for the quantitative description of time series,

$$\begin{aligned}a(t) &= \frac{1}{(N-m)\sigma^2}\\ &\times \sum_{j=0}^{N-1-m} \delta x(T+j\tau)\delta x(T+(j+m\tau)),\end{aligned} \quad (4)$$

where σ^2 is the variance, N is the number of measurements, and τ is a finite discretization time. The key element of the theory consists in transition from continuum values, variables, and equations to discrete ones. We then obtain a Liouville-like equation of motion for multidimensional state vectors. We can use the method of projection operators in a finite-dimensional vector space. This allows splitting the Liouville-like discrete equation of motion into two mutually orthogonal subspaces, one of which is relevant and the other is irrelevant to discrete time correlations. We have also developed the method for obtaining the set of dynamic orthogonal variables by the Gram–Schmidt orthogonalization procedure.

Dynamical orthogonal variables were calculated from initial time series (3) by the formulas (see [5, 6])

$$\begin{aligned}\hat{\mathbf{W}}_0 &= \mathbf{A}_k^0, \quad \hat{\mathbf{W}}_1 = \left(\frac{\Delta}{\Delta t} - \lambda_1\right)\mathbf{A}_k^0,\\ \hat{\mathbf{W}}_2 &= \left(\frac{\Delta}{\Delta t} - \lambda_2\right)\mathbf{W}_1 + \Lambda_1\mathbf{A}_k^0\\ &= \left[\left(\frac{\Delta}{\Delta t}\right)^2 - \frac{\Delta}{\Delta t}(\lambda_1+\lambda_2) + \lambda_1\lambda_2 + \Lambda_1\right]\mathbf{A}_k^0,\\ \hat{\mathbf{W}}_3 &= \left(\frac{\Delta}{\Delta t} - \lambda_3\right)\mathbf{W}_2 + \Lambda_2\left(\frac{\Delta}{\Delta t} - \lambda_1\right)\mathbf{A}_k^0,\end{aligned} \quad (5)$$

where the parameters λ_i and Λ_i were calculated using (2). Simple, but cumbersome calculations show that the first short-memory function $m_n(t)$ represents a normalized TCF of the first dynamic variable W_n,

$$m_n(t) = \frac{\langle \mathbf{W}_n(0)\mathbf{W}_n(t)\rangle}{\langle \mathbf{W}_n(0)\rangle^2}. \quad (6)$$

We then obtain a chain of finite-difference discrete non-Markovian kinetic equations for the initial time correlation function and memory functions of various orders. We note that all the involved kinetic and relaxation parameters, the time correlation function, and the memory functions can easily be found and calculated directly from the experimental time series. The spectra of memory functions were calculated using the fast Fourier transform.

5. NON-MARKOVIAN PROPERTIES OF EEGs

We have analyzed the time and frequency evolution of the signals during tonic–clonic seizure by means of the time-window technique. We find that the memory function spectra and the statistical spectrum of the non-Markovity parameter are valuable for quantitative and qualitative analysis of epileptic seizures. Numerical parameters based on the theory of discrete non-Markovian processes provide quantitative information about the state of the brain before, during, and after the seizure.

Non-Markovian properties are known to play an essential role in the time dynamics of complex systems. On the basis of our theory [5, 6], we can calculate memory functions $M_i(t)$, $i = 0, 1, 2, 3$ directly from experimental data by Eqs. (2.41)–(2.46) in [6]. We analyze the properties of memory functions by calculating their power frequency spectra. For a quantitative estimation

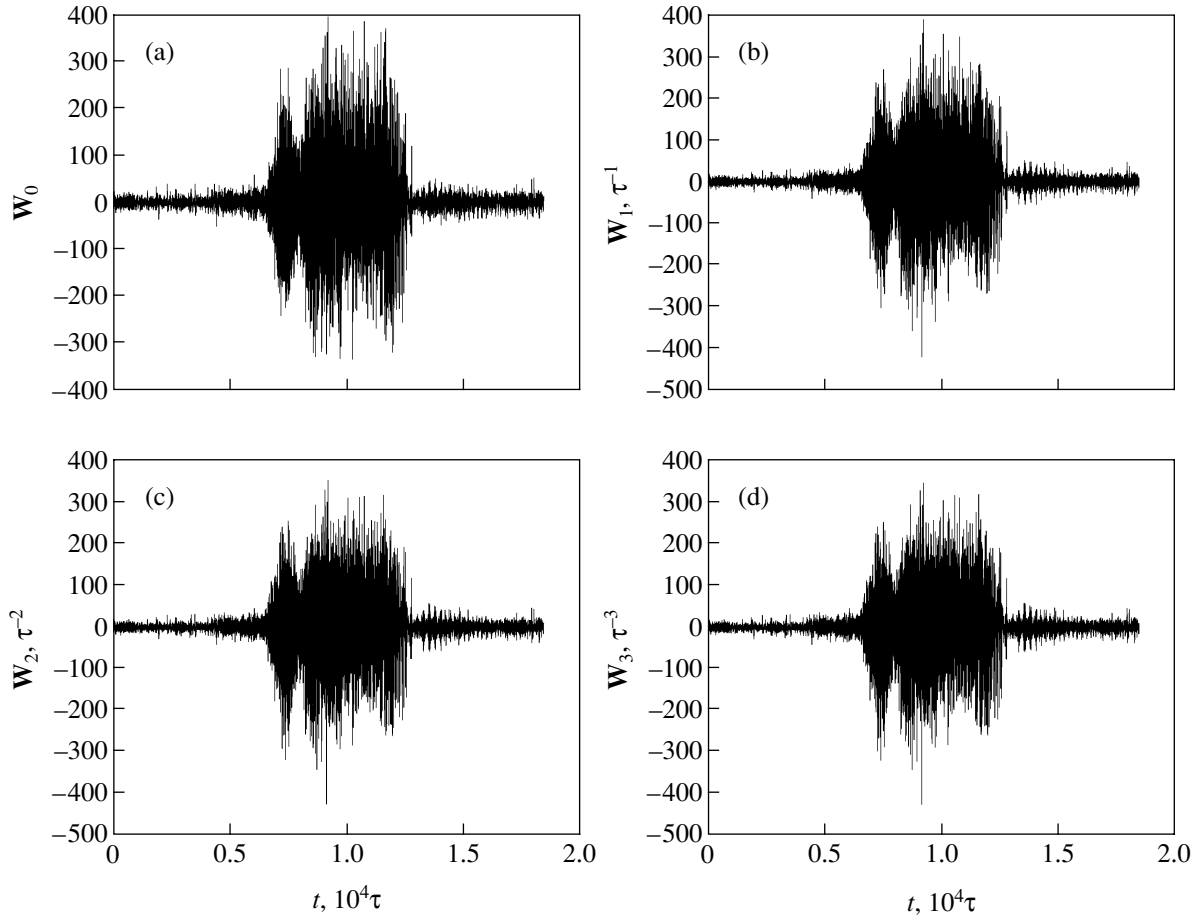

Fig. 1. Time record of the first four orthogonal variables W_0 (a), W_1 (b), W_2 (c), and W_3 (d) of the sampling of electric activity during the tonic–clonic seizure under study. The difference in the dimensions of the four variables must be taken into account in analyzing the scales. The general form of all signals has definite similarity. Simultaneously, some differences in time behavior, $W_i(t)$, are made evident, especially for the states before and after the seizure. We emphasize that the whimsical entanglement of regular and chaotic components is omnipresent in the time-recording window of all the signals. We also note that the difference between the raw EEG data before, during, and after the seizure is sufficiently dramatic. However, simple registration of this fact does not allow us to reveal such subtle features of EEG spectra as the presence or absence of the chaotic or regular components in the signal.

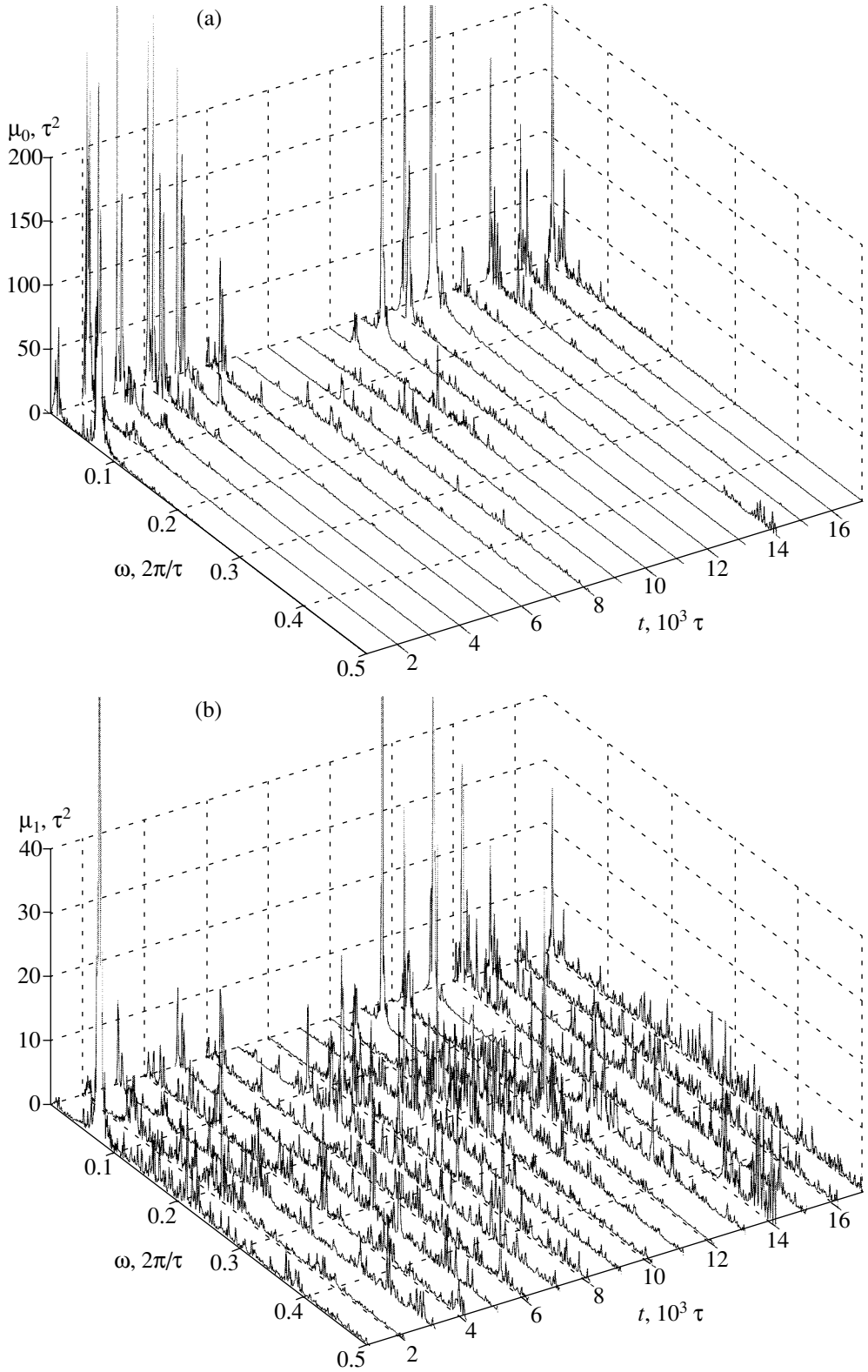

Fig. 2. The window-time behavior of the power spectra $\mu_i(\omega)$, $i = 0$ (a), 1 (b), 2 (c), and 3 (d), for the considered sampling with the tonic–clonic seizure from the short-time window dynamics of the human brain electric activity. The sharp reduction (by almost one order) of intensity of the low-frequency components of the spectra (in the region of δ and ϑ rhythms) attracted our attention at the transition from μ_0 to μ_1, μ_2, and μ_3. The spectra $\mu_i(\omega)$, $i = 1, 2, 3$, contain rather strong noises distributed at regular intervals in the entire frequency region. The intensity in the region of δ and ϑ rhythms sharply decreases in the first half of the seizure (7th, 8th, 9th and, in part, 10th windows) in all μ_i, $i = 0, 1, 2, 3$. The sharp increase in the intensity in the low-frequency region of the spectrum by almost 100 times (in the regions of δ and ϑ rhythms) is observed in the second half of the seizure (11th, 12th, and 13th windows).

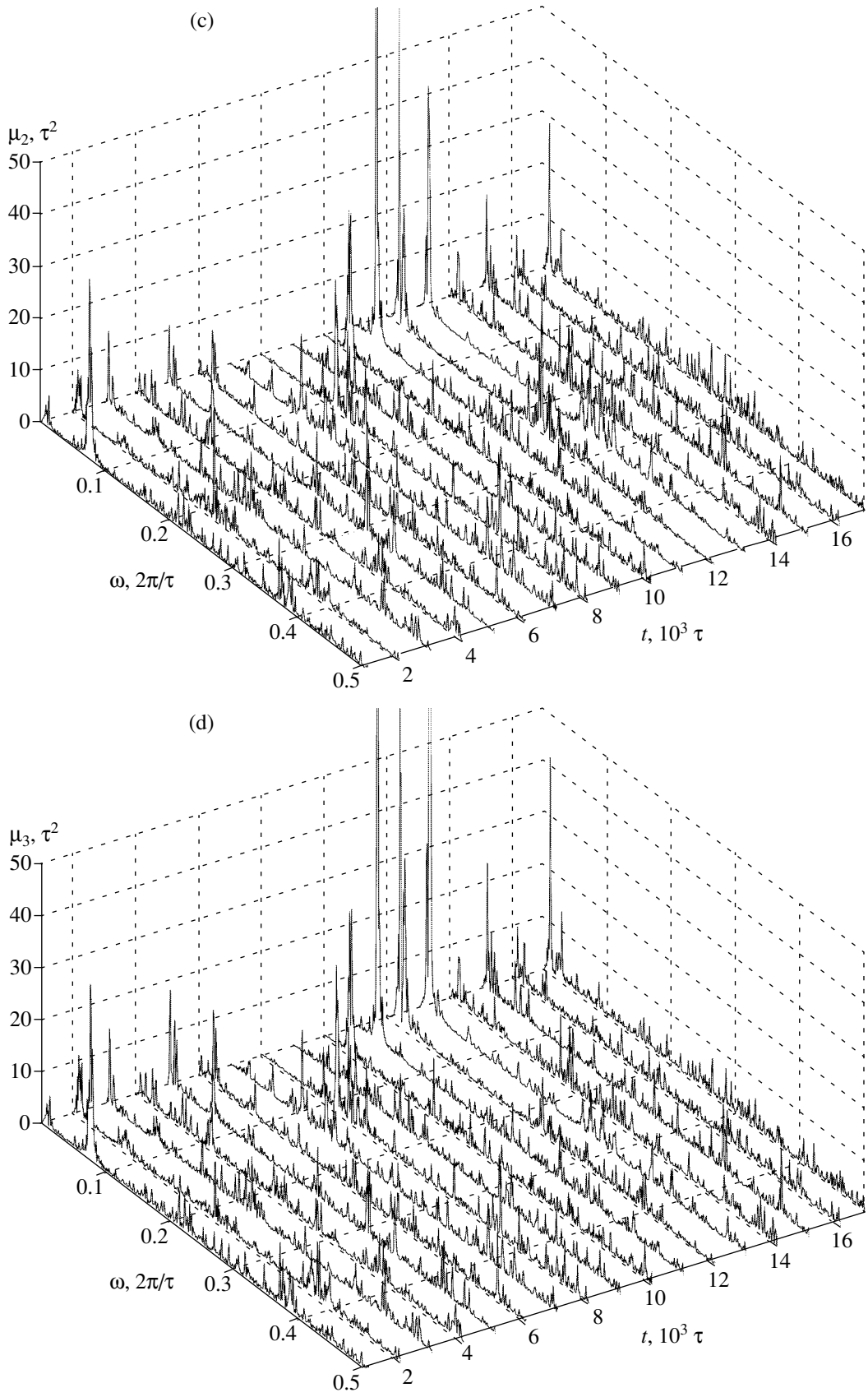
(c)
μ_2, τ^2
50
40
30
20
10
0
0.1
0.2
0.3
0.4
0.5
ω, 2π/τ
2
4
6
8
10
12
14
16
$t, 10^3\ \tau$
(d)
μ_3, τ^2
50
40
30
20
10
0
0.1
0.2
0.3
0.4
0.5
ω, 2π/τ
2
4
6
8
10
12
14
16
$t, 10^3\ \tau$

Fig. 2. (Contd.)

of the non- degree of Markovity, we use the frequency-dependent generalized non-Markovity parameter $\varepsilon_i(\omega)$ introduced by us previously [5]. From the theory in [5, 6], we can also calculate the quantitative values of the kinetic and relaxation parameters λ_1, λ_2, λ_3, Λ_1, and Λ_2 that give additional information on the properties of the complex system under study.

For the observed EEG spectra, we divide the entire time evolution data into nonoverlapping epochs of 1024 data points each. The dynamics of the first four dynamical orthogonal variables W_0, W_1, W_2, and W_3 of the entire data set is presented in Fig. 1. For each epoch, we have calculated the power spectra of the first four memory functions $\mu_0(\omega)$, $\mu_1(\omega)$, $\mu_2(\omega)$, $\mu_3(\omega)$ and the first three points of statistical spectra of the non-Markovity parameter $\varepsilon_1(\omega)$, $\varepsilon_2(\omega)$, and $\varepsilon_3(\omega)$ [5]. The time evolution of the spectra is shown in three-dimensional diagrams (Figs. 2 and 3). The time evolution of the numerical parameters λ_1, λ_2, λ_3, Λ_1, Λ_2 is presented in Fig. 4.

We emphasize that strong demarkovization of the stochastic changes of brain electrical potentials with decreasing numerical values ε_i to the point of unity is exhibited during a tonic–clonic seizure. The chaotic regime of the system is then replaced by the steady non-Markovian-state regime.

It can be seen from Figs. 1a–1d that the time evolution of the dynamic orthogonal variables W_i, $i = 0, 1, 2,$

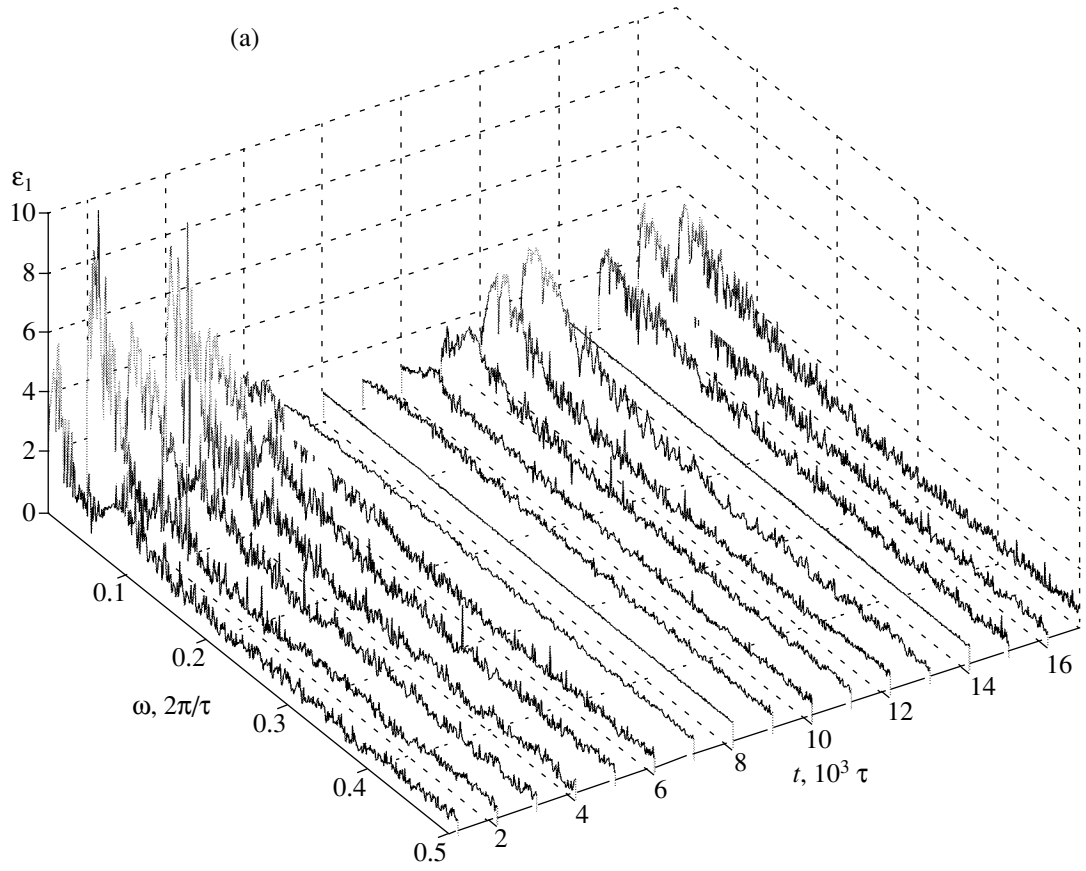

Fig. 3. The window-time behavior of the first three points of the non-Markovity parameter $\varepsilon_i(\omega)$, $i = 1$ (a), 2 (b), and 3 (c), for the long sampling including the tonic–clonic seizure during epilepsy. For the state before the seizure, the quasi-Markovian behavior ($\varepsilon_1 \sim 10$) of the first point $\varepsilon_1(\omega)$ in the low-frequency region (with δ and α rhythms) is obvious. The beginning of the seizure (7th, 8th, 9th, and 10th windows) exhibits a strong non-Markovity ($\varepsilon_1 \sim 1$) on all frequencies of the full spectrum. A weak non-Markovity in the region of δ and ϑ rhythms ($\varepsilon \longrightarrow 4$) is found during the seizure. A strong non-Markovity on all frequencies is established immediately after termination of the seizure (14th window). The frequency behavior of $\varepsilon_3(\omega)$ is characterized by steady non-Markovity ($\varepsilon_3 \longrightarrow 1$) in all the windows and in the entire frequency region. A weak quasi-Markovian noise (in the region of α and β rhythms) appears before the seizure (2nd and 5th windows) and at the end of the seizure (12th, 13th, and 14th windows). The behavior of parameter $\varepsilon_2(\omega)$ is rather peculiar. A strong non-Markovity ($\varepsilon_2 \sim 1$) appears long before the seizure (3rd, 4th, 5th, and 6th windows). Further development of the seizure is accompanied by a slight noise in $\varepsilon_2(\omega)$ in the region of α and β rhythms. The termination of the seizure results in a strong non-Markovity ($\varepsilon_2 \sim 1$) in the 13th window. Noisiness in the entire frequency range of the 14th window then occurs. The steady non-Markovity ($\varepsilon_2 \sim 1$) is appreciable in the 15th, 16th, and 17th windows, appearing after termination of the seizure. The low-frequency (in the region of δ rhythms with $\varepsilon_2 \sim 1$) and high-frequency (in the top border of the γ-spectrum with $\varepsilon_2 \sim 3$) sites of the spectrum are intensely noisy.

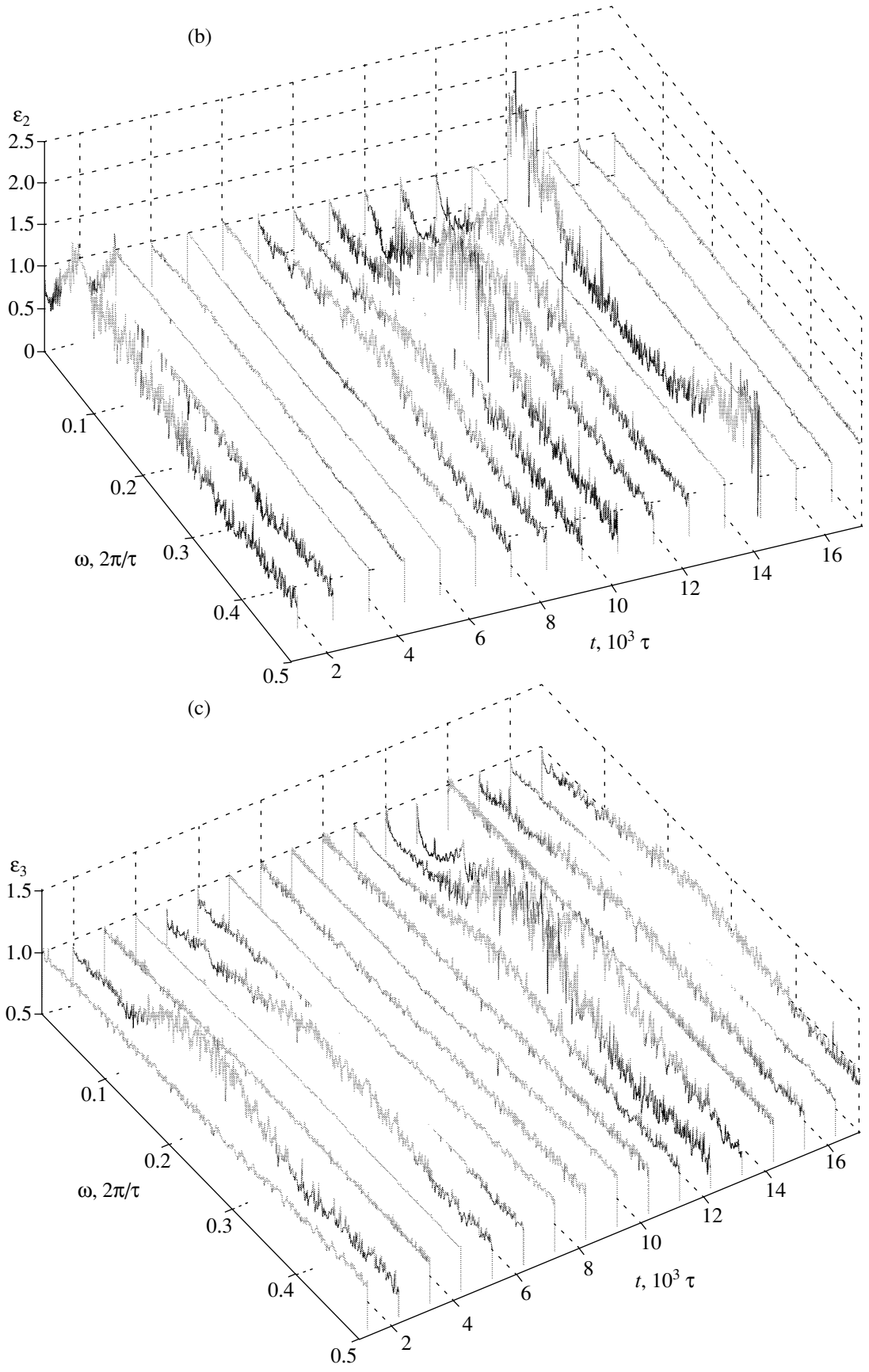
(b)
ε2
2.5
2.0
1.5
1.0
0.5
0
0.1
0.2
0.3
0.4
0.5
ω, 2π/τ
2
4
6
8
10
12
14
16
t, 10^3 τ
(c)
ε3
1.5
1.0
0.5
0.1
0.2
0.3
0.4
0.5
ω, 2π/τ
2
4
6
8
10
12
14
16
t, 10^3 τ

Fig. 3. (Contd.)

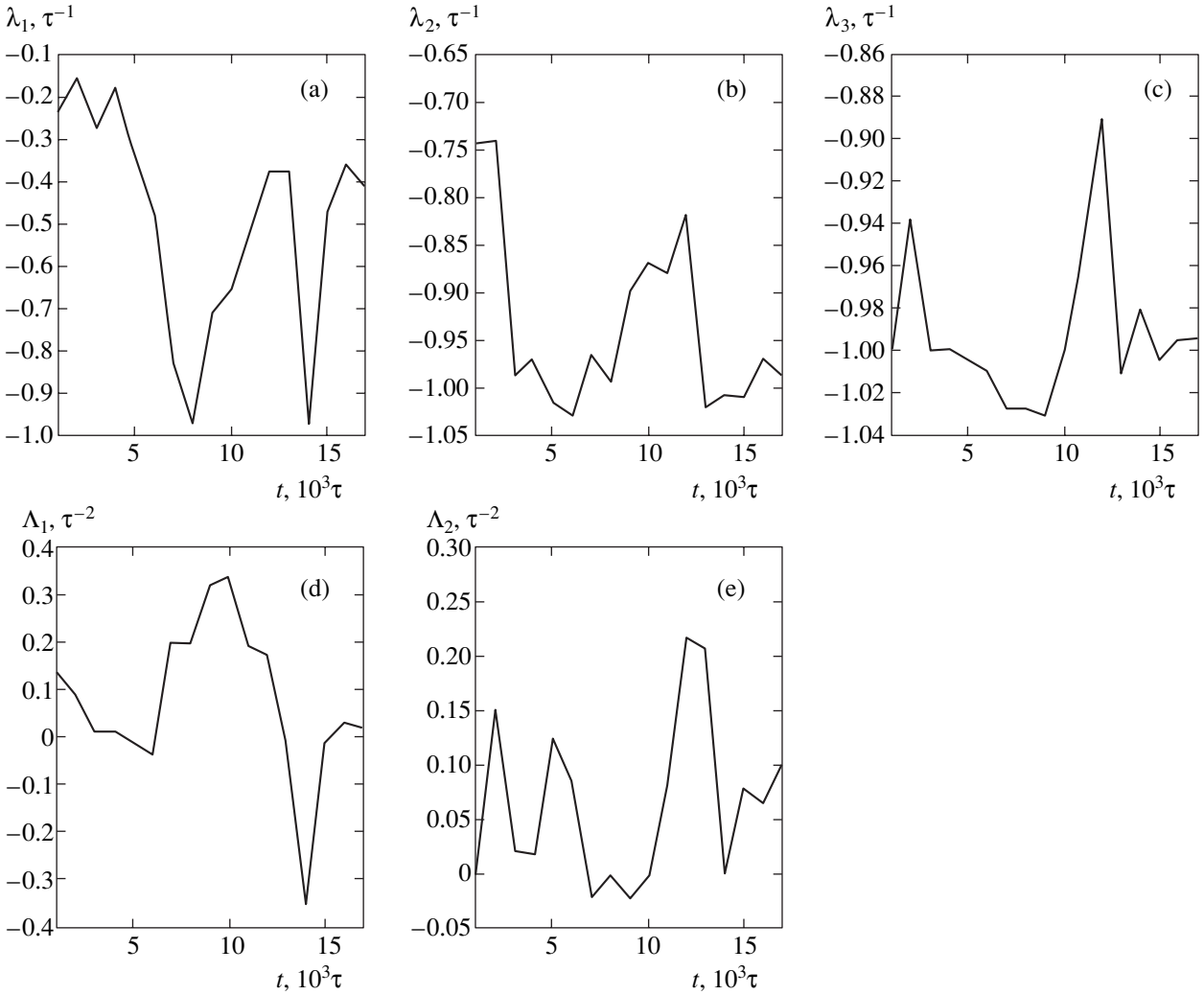

Fig. 4. The window-time behavior of the kinetic (λ_1 (a), λ_2 (b), and λ_3 (c)) and relaxation (Λ_1 (d) and Λ_2 (e)) parameters for the time sampling at epilepsy with the tonic–clonic seizure. The kinetic parameters λ_1, λ_2, and λ_3 are always negative and increase with seizure. The relaxation parameters Λ_1 and Λ_2 change sharply with the sign change from the beginning of the seizure. The most dramatic changes in the behavior of Λ_1 and Λ_2 occur during the seizure in the opposite directions.

3, can be smoothed. Therefore, the scales of these variables before and during the seizure are practically identical. The beginning of the seizure (see Figs. 2a–2d) is characterized by a sharp recession of low-frequency peaks in the spectrum $\mu_0(\omega)$ (7th–10th windows); these peaks in $\mu_0(\omega)$ rise sharply at the end of the seizure and immediately after the seizure. The spectra of $\mu_j(\omega)$, j = 1, 2, 3, differ by white noise and low-frequency bursts on the tail of the seizure. These bursts are most appreciable in the behavior of the spectra $\mu_2(\omega)$ and $\mu_3(\omega)$.

The behavior of the first three points in the statistical spectrum of the non-Markovity parameter $\varepsilon_i(\omega)$, i = 1, 2, 3 (see Fig. 3), turn out to be most indicative and demonstrative. The state before the seizure can be considered quasi-Markovian in the 1st–6th windows for the first level in the low-frequency region (here, $\varepsilon_1(\omega)$ reaches a value of 10) and in the 1st and 2nd windows for the second level ($\varepsilon_2(\omega) \sim 1.5$). The beginning of the seizure (the 7th and 8th windows) is accompanied by the strong non-Markovity of the first level ($\varepsilon_i(\omega) \approx 1$). The increase in low-frequency non-Markovity on the first ($\varepsilon_1 \sim 3.8$), second ($\varepsilon_2 \sim 1.5$), and the third ($\varepsilon_3 \sim 1.5$) relaxation levels is visible at the end of the seizure (10th–13th windows). The behavior on the third level with a value of $\varepsilon_3 \approx 1$ can be considered non-Markovian.

Non-Markovian relaxation behavior on the second level is noteworthy (see Fig. 3b). The strong non-Markovity ($\varepsilon_2 \approx 1$) in the entire frequency region appears long before the seizure in the range from the third to the sixth windows. The weak noise at the mean frequencies is appreciable during the seizure (10th–12th windows). The ending of the seizure coincides with the non-Markovian 13th and quasi-Markovian 14th windows. The

appearance of a strongly pronounced non-Markovian state on the second level with a value of $\varepsilon_2 \approx 1$ is therefore a clear precursor to the seizure. It is significant that a similar precursor is absent in other non-Markovian markers.

The relaxation (λ_1, λ_2, and λ_3) and kinetic (Λ_1 and Λ_2) parameters calculated with the formulas of the theory (see Fig. 4) are very sensitive to approaching the seizure. All the parameters λ_i, $i = 1, 2, 3$, always remain negative and change within wide limits: ($-0.97 \le \lambda_1 \le -0.15$, $-1.03 \le \lambda_2 \le -0.74$, and $-1.03 \le \lambda_3 \le -0.89$) in units of τ^{-1}. Parameters Λ_1 and Λ_2 change sign at the time of the seizure. This corresponds to alternation in the type of solution to the discrete nonlinear kinetic stochastic equation (see Eqs. (2.56)–(2.58) in [6]). All of the above parameters are sensitive to approaching the seizure. A sharp decrease in the values and the sign alternation of λ_i and Λ_i can also be considered a quantitative precursor to a seizure.

Therefore, the increase in parameters $\varepsilon_i(\omega)$ can be considered a markovization of the stochastic process. It may signify an increase in the chaotic components of EEG signals. Simultaneously, a decrease in $\varepsilon_i(\omega)$ to unity is related to demarkovization of the process under study and to an increase in the regular components of the signals. It is obvious from Figs. 1–4 that the specific alternations, fast and sudden changes of chaotic and regular regimes, are inherent features of the stochastic variation of electric potentials during epileptic seizure.

6. CONCLUSIONS

We have clearly demonstrated that the set of kinetic, relaxation, dynamic, and spectral parameters and characteristics of a discrete stochastic process are valuable for quantification of stochastic processes of markovization and demarkovization in EEG data and for prediction of and precursor to epileptic seizure. Because a similar situation is typical of the majority of the phenomena in live systems, our findings are most relevant for life sciences.

ACKNOWLEDGMENTS

This work was supported in part (P.H. and R.Y.) by the Graduiertenkolleg 283: Nonlinear Problems in Analysis, Geometry and Physics, of the Deutsche Forschungsgemeinschaft (DFG), RHSF (grant no. 00-06-00005a), and the Russian Foundation for Basic Research (project no. 02-02-16146).

REFERENCES

1. R. Q. Quiroga, J. Arnhold, K. Lehnertz, and P. Grassberger, Phys. Rev. E **62**, 8380 (2000); R. Q. Quiroga, A. Kraskov, T. Kreuz, and Grassberger, Phys. Rev. E **65**, 041903 (2002).
2. P. A. Robinson, C. J. Rennie, J. J. Wright, *et al.*, Phys. Rev. E **63**, 021903 (2001).
3. L. M. Hively, V. A. Protopopescu, and P. C. Gailey, Chaos **10**, 861 (2000).
4. R. G. Andrzejak, K. Lehnertz, F. Mormann, *et al.*, Phys. Rev. E **64**, 061907 (2001).
5. R. Yulmetyev, P. Hänggi, and F. Gafarov, Phys. Rev. E **62**, 6178 (2000); R. Yulmetyev, F. Gafarov, P. Hänggi, *et al.*, Phys. Rev. E **64**, 066132 (2001).
6. R. Yulmetyev, P. Hänggi, and F. Gafarov, Phys. Rev. E **65**, 046107 (2002).
7. V. Yu. Shurygin and R. M. Yulmetyev, Phys. Lett. A **148**, 199 (1990); Zh. Éksp. Teor. Fiz. **99**, 144 (1991) [Sov. Phys. JETP **72**, 80 (1991)].
8. R. Q. Quiroga, S. Blanco, O. Rosso, *et al.*, Electroencephalogr. Clin. Neurophysiol. **103**, 434 (1997); http://www.vis.caltech.edu/~rodri/data/.