Journal of Experimental and Theoretical Physics, Vol. 95, No. 5, 2002, pp. 789–799.
From Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 122, No. 5, 2002, pp. 915–927.

GRAVITATION, ASTROPHYSICS

Electromagnetic Activity of a Pulsating Paramagnetic Neutron Star¶

S. I. Bastrukov[a, b, c, *], D. V. Podgainy[a], J. Yang[c, d], and F. Weber[e]

[a] *Joint Institute for Nuclear Research, Dubna, Moscow oblast, 141890 Russia*
[b] *Center for High-Energy Astrophysics, Ewha Womans University, Seoul 120-750; Asia Pacific Center for Theoretical Physics, Pohang 790-784, South Korea*
[c] *Center for High-Energy Astrophysics, Ewha Womans University, Seoul 120-750, South Korea*
[d] *Center for High-Energy Physics, Kyungpook National University, Daegu 702-701, South Korea*
[e] *Department of Physics, University of Notre Dame, Notre Dame, IN 46556-5670, USA*
**e-mail: bast@thsunl.jinr.ru*

Received June 10, 2002

Abstract—The fact that neutron star matter possesses the capability of maintaining a highly intense magnetic field has been and still is among the most debatable issues in pulsar astrophysics. Over the years, there were several independent suggestions that the dominant source of pulsar magnetism is either the field-induced or the spontaneous magnetic polarization of the baryon material. The Pauli paramagnetism of degenerate neutron matter is one of the plausible and comprehensive mechanisms of the magnetic ordering of neutron magnetic moments, promoted by a seed magnetic field inherited by the neutron star from a massive progenitor and amplified by its implosive contraction due to the magnetic flux conservation. Adhering to this attitude and based on the equations of magnetoelastic dynamics underlying continuum mechanics of single-axis magnetic insulators, we investigate electrodynamics of a paramagnetic neutron star undergoing nonradial pulsations. We show that the suggested approach regains a recent finding of Akhiezer *et al.* [1] that the spin-polarized neutron matter can transmit perturbations by low-frequency transverse magnetoelastic waves. We found that nonradial torsional magnetoelastic pulsations of a paramagnetic neutron star can serve as a powerful generator of a highly intense electric field producing the magnetospheric polarization charge whose acceleration along the open magnetic field lines leads to the synchrotron and curvature radiation. Analytic and numerical estimates for periods of nonradial torsional magnetoelastic modes are presented and are followed by a discussion of their possible manifestation in currently monitored activity of pulsars and magnetars. © *2002 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

Recent years have seen a resurgence of interest in the magnetic properties of neutron star matter [1–4] and of the early advanced hypothesis that a considerable contribution to the ultrastrong magnetic field of these compact objects can be attributed to spin polarization of stellar material [5–8]. This development calls into question our understanding of the laws governing continuum mechanics and macroscopic electrodynamics of magnetically ordered nuclear matter. To the best of our knowledge, the first significant step in this direction was made in [1], cited in the abstract and hereafter referred to as the Akhiezer–Laskin–Peletminskii (ALP) model, advocating ferromagnetism of neutron stars. Using equations of the magnetohydrodynamic type adopted from macroscopic electrodynamics of ferromagnetic dielectrics [9], it was shown that magnetically ordered neutron matter can transmit perturbations by low-frequency magnetoelastic waves along with the well-known high-frequency spin waves typical of ferromagnetic solids [9–11]. The observation of these oscillatory motions in currently monitoring neutron stars is crucial, in our opinion, for unambiguous identification of the permanent magnetism of stellar material. This attitude motivates our present work, continuing investigations begun in [1], aimed at searching characteristic features of electromagnetic activity of neutron stars owing its origin to nonradial magnetoelastic pulsations of paramagnetic neutron stars.

The fingerprints of the Pauli mechanism of the field-induced (nonspontaneous) spin polarization of neutron star matter can be traced in the existing scenario of the pulsar birth in a supernova event [12, 13]. The catastrophic collapse of the massive main sequence star exhausting its nuclear fuel implies that implosive contraction of a weakly magnetized massive star is accompanied by intensive neutronization of stellar material due to the inverse β process

$$e + p \longrightarrow n + \nu_e$$

responsible for fast cooling of pulsars [12]. Because this urca process is controlled by weak, parity-violating interaction, it is expected that the magnetic anisotropy caused by the presence of a seed magnetic field intro-

¶This article was submitted by the authors in English.

duces in the final product of the collapse a tremendous difference between the number of neutrons with spin magnetic moments directed along the seed magnetic field and those with oppositely directed spins, such that the main body of the newly born neutron star mass develops a permanent magnetization of the paramagnetic type. The amplification of the magnetic field in this process is attributed to implosive contraction that proceeds with the preserved magnetic flux.

Following this line of argument, we consider the homogeneous model of a paramagnetic neutron star undergoing nonradial pulsations triggered either by the implosive effect of a supernova event or by gamma-bursting starquakes. In doing this, we utilize a somewhat different, as compared to the ALP model, form of the macroscopic equations governing the motions of magnetically polarized neutron matter, adopted from the macroscopic electrodynamics of single-axis magnetoelastic insulators [14]. One of the purposes is to show that the proposed approach is interesting in its own right because the continuum mechanics of magnetically polarized stellar matter is less studied in astrophysics compared to magnetohydrodynamics underlying our understanding of the motions of highly conductive stellar matter threaded by a magnetic field. Different aspects of this project have been reported in proceedings of several recent conferences [15–17], and our goal here is to bring them together in an extended fashion.

The paper is organized as follows. In Section 2, the macroscopic equations of the magnetoelastic dynamics of spin-polarized nuclear matter are introduced and the dispersion equation for the wave transport of magnetization is derived. Section 3 presents a variational calculation of the periods of nonradial torsional pulsations of paramagnetic neutron stars with emphasis on the generation of the magnetospheric polarization charge responsible for the radiation from the star; the obtained analytic estimates are quantified using parameters that are typical of radio pulsars and magnetars. The last section provides a brief summary of the results obtained.

2. GOVERNING EQUATIONS FOR MAGNETOELASTIC DYNAMICS

In what follows, we assume, as in most of the works cited above, that permanently magnetized baryon matter of a neutron star possesses properties of a degenerate Fermi gas of neutrons condensed by self-gravity to the normal nuclear density $\rho = 2.8 \times 10^{14}$ g cm^{-3}. To describe the equilibrium state of spin-polarized neutron star matter, we use a linear constitutive equation in the form given in [18],

$$\mathbf{M} = \chi \mathbf{B}. \tag{2.1}$$

Here, $\chi > 0$ stands for the average paramagnetic susceptibility of homogeneous neutron star matter, which is estimated to be $\chi \approx 2\chi_F$ at the normal nuclear density [5–8], where χ_F is the Pauli paramagnetic susceptibility of zero-temperature, degenerate, neutron Fermi gas compressed to the nuclear density,

$$\chi_F = \frac{3}{2} n \frac{\mu_n^2}{\epsilon_F} \approx 1.3 \times 10^{-4},$$

and **B** denotes the fossil magnetic field frozen in the neutron star core. The macroscopic description of motions of neutron star matter in terms of the theory of continuous media implies that the space scale of material displacements is much larger than the spacing between baryons. The basic suggestion underlying continuum models of neutron star material is to identify the behavior of many-component spin-polarized baryon matter with that the spin-polarized neutron degenerate Fermi gas of the equivalent density ρ subjected to the standard continuity equation

$$\frac{d\rho}{dt} + \rho \frac{\partial v_k}{\partial x_k} = 0. \tag{2.2}$$

Hereafter,

$$d/dt = \partial/\partial t + \mathbf{v} \cdot \nabla$$

stands for the convective derivative. The second suggestion of particular interest is to consider the magnetization field $\mathbf{m}(\mathbf{r}, t)$ (magnetic moment per unit volume) as an independent dynamical variable of motion, on equal footing with the bulk density $\rho(\mathbf{r}, t)$ and the elastic displacement velocity $\mathbf{v}(\mathbf{r}, t)$. According to [14], the distinguishing feature of mechanical behavior of magnetoelastic insulators is that the dynamics of their intrinsic deformations is controlled by a driving force originating from antisymmetric magnetic stresses τ_{ij} (see also [19]). The dynamical equation of magnetoelasticity is given by

$$\rho \frac{dv_i}{dt} = \frac{\partial \tau_{ik}}{\partial x_k}, \quad \tau_{ik} = \frac{1}{2}[m_i B_k - m_k B_i]. \tag{2.3}$$

Thus, the antisymmetric form of the magnetic stress tensor τ_{ij} exhibits a substantially non-Hookean character[1] of magnetoelasticity, which comes into play only when the direction of the local magnetization **m** deviates from the direction of the equilibrium magnetiza-

[1] The linear elastodynamics of material displacements u_i in an isotropic solid under pure shear deformations that are not accompanied by density fluctuations is described by the Lamé equation [20]

$$\rho \frac{\partial^2 u_i}{\partial t^2} = \frac{\partial \sigma_{ik}}{\partial x_k}, \quad \frac{\partial u_k}{\partial x_k} = 0, \tag{2.4}$$

$$\sigma_{ik} = 2\mu u_{ik}, \quad u_{ik} = \frac{1}{2}\left(\frac{\partial u_k}{\partial x_i} + \frac{\partial u_i}{\partial x_k}\right),$$

where σ_{ik} is the symmetric tensor of elastic stresses, μ is the shear modulus, and u_{ik} is the strain tensor.

tion **M**. The constitutive equation for the evolution of **m** is given by

$$\frac{dm_i}{dt} = \omega_{ik}m_k, \quad \omega_{ik} = \frac{1}{2}\left(\frac{\partial v_k}{\partial x_i} - \frac{\partial v_i}{\partial x_k}\right), \tag{2.5}$$

where ω_{ik} is interpreted as the antisymmetric rate-of-deformation tensor [14].

The above equations of dissipation-free magnetoelastic dynamics can be represented in the following equivalent vector form:

$$\frac{d\rho(\mathbf{r}, t)}{dt} + \rho(\mathbf{r}, t)\nabla\cdot\mathbf{v}(\mathbf{r}, t) = 0, \tag{2.6}$$

$$\rho\frac{d\mathbf{v}(\mathbf{r}, t)}{dt} = \frac{1}{2\chi}\nabla\times[\mathbf{m}(\mathbf{r}, t)\times\mathbf{M}], \quad \mathbf{M} = \chi\mathbf{B}, \tag{2.7}$$

$$\begin{aligned}\frac{d\mathbf{m}(\mathbf{r}, t)}{dt} &= [\boldsymbol{\omega}(\mathbf{r}, t)\times\mathbf{M}],\\ \boldsymbol{\omega}(\mathbf{r}, t) &= \frac{1}{2}[\nabla\times\mathbf{v}(\mathbf{r}, t)].\end{aligned} \tag{2.8}$$

This form accentuates the fact that the magnetoelastic driving force

$$\mathbf{f}(\mathbf{r}, t) = \nabla\times\boldsymbol{\tau}(\mathbf{r}, t)$$

in Eq. (2.7) is inextricably related to the magnetic torque density

$$\boldsymbol{\tau}(\mathbf{r}, t) = \frac{1}{2}[\mathbf{m}(\mathbf{r}, t)\times\mathbf{B}];$$

we again see that magnetoelastic effects manifest themselves when the magnetization field **m** deviates from the direction of the saturated magnetization $\mathbf{M} = \chi\mathbf{B}$. Equation (2.8) describing differential rotation of the magnetization about the magnetic anisotropy axis is the standard equation of precession under which the direction of **m** changes but the magnitude does not. It is noteworthy that similar equations have recently been used in the study of the large-scale motions of a poorly conducting interstellar medium possessing properties of gas-based ferrocolloidal soft matter consisting of tiny ferromagnetic solid grains suspended in a dense magnetically passive and electrically neutral fluid [22].

2.1. Wave Transport of Magnetization in Paramagnetic Neutron Star Matter

Applying the standard linearization procedure to Eqs. (2.6)–(2.8),

$$\mathbf{v} \longrightarrow \mathbf{v}_0 + \delta\mathbf{v}(\mathbf{r}, t), \quad \mathbf{m} \longrightarrow \mathbf{m}_0 + \delta\mathbf{m}(\mathbf{r}, t),$$

where

$$\mathbf{v}_0 = 0, \quad \mathbf{m}_0 = \mathbf{M} = \chi\mathbf{B},$$

we obtain,

$$\nabla\cdot\delta\mathbf{v}(\mathbf{r}, t) = 0, \quad \nabla\cdot\delta\mathbf{m}(\mathbf{r}, t) = 0, \tag{2.9}$$

$$\rho\frac{\partial\delta\mathbf{v}(\mathbf{r}, t)}{\partial t} = \frac{1}{2\chi}\nabla\times[\delta\mathbf{m}(\mathbf{r}, t)\times\mathbf{M}], \tag{2.10}$$

$$\frac{\partial\delta\mathbf{m}(\mathbf{r}, t)}{\partial t} = \frac{1}{2}[[\nabla\times\delta\mathbf{v}(\mathbf{r}, t)]\times\mathbf{M}]. \tag{2.11}$$

This set of coupled equations describes transmission of linear fluctuations in incompressible spin-polarized baryon matter that are not accompanied by the appearance of the density of magnetic poles (the right-hand sides of Eqs. (2.9)). Substitution of the plane-wave form of the fluctuating variables

$$\begin{aligned}\delta\mathbf{v} &\propto \exp(i\omega t - i\mathbf{k}\cdot\mathbf{r}),\\ \delta\mathbf{m} &\propto \exp(i\omega t - i\mathbf{k}\cdot\mathbf{r})\end{aligned} \tag{2.12}$$

into (2.11) leads to the transversality conditions

$$\mathbf{k}\cdot\delta\mathbf{v} = 0, \quad \mathbf{k}\cdot\delta\mathbf{m} = 0.$$

Inserting (2.12) into (2.10) yields

$$\omega\rho\delta\mathbf{v} = -\frac{1}{2\chi}(\mathbf{k}\cdot\mathbf{M})\delta\mathbf{m}.$$

After substitution of (2.12) into (2.11), we obtain

$$\omega\delta\mathbf{m} = -\frac{1}{2}[(\mathbf{k}\cdot\mathbf{M})\delta\mathbf{v} - \mathbf{k}(\delta\mathbf{v}\cdot\mathbf{M})].$$

Taking the scalar product of the last equation with $\mathbf{k} \neq 0$ and considering the above transversality conditions, we obtain

$$\delta\mathbf{v}\cdot\mathbf{M} = 0.$$

Given this, the link between the frequency and the wave vector in the magnetoelastic wave is defined by the coupled equations

$$\begin{aligned}\omega\rho\delta\mathbf{v} + \frac{1}{2\chi}(\mathbf{k}\cdot\mathbf{M})\delta\mathbf{m} &= 0,\\ \omega\delta\mathbf{m} + \frac{1}{2}(\mathbf{k}\cdot\mathbf{M})\delta\mathbf{v} &= 0.\end{aligned} \tag{2.13}$$

Eliminating $(\mathbf{k}\cdot\mathbf{M})$, we find that magnetoelastic oscillatory motions satisfy the energy equipartition principle

$$\frac{\rho\delta\mathbf{v}^2}{2} = \frac{\delta\mathbf{m}^2}{2\chi}, \tag{2.14}$$

which states that, in the magnetoelastic wave, the kinetic energy of fluctuating elastic displacements equals the mean potential energy of fluctuating magnetization. The compatibility of Eqs. (2.13) leads to the dispersion relation of the magnetoelastic wave,

$$\omega^2 = \frac{(\mathbf{k}\cdot\mathbf{M})^2}{4\chi\rho} = \frac{\chi}{4\rho}(\mathbf{k}\cdot\mathbf{B})^2 = V_M^2k^2\cos^2\theta, \tag{2.15}$$

where θ is the angle between **k** and **M**. It is remarkable that the speed of the wave transport of magnetization

$$V_M = \sqrt{\frac{MB}{4\rho}} \tag{2.16}$$

in paramagnetic neutron matter is proportional to the intensity of the fossil magnetic field **B**; in ferromagnetic neutron matter, this speed is proportional to the intensity of the spontaneous magnetization **M**. This is noteworthy because the magnetoelastic wave transport of magnetization is characterized by a dispersion-free law, $\omega \propto k$ in contrast to spin waves that have quadratic dispersion in k, $\omega \propto k^2$. It is therefore expected that, under the cooling of paramagnetic neutron star, the temperature variation of the equilibrium magnetization $M(T)$ follows the Curie law

$$\frac{M(T)}{B} = \chi(t) \propto T^{-1},$$

which is due to the dispersion-free nature of magnetophonons, instead of the Bloch law

$$M(0) - M(T)/M(0) \propto T^{3/2}$$

for ferromagnetic dielectrics, which is due to quadratic dispersion of magnons.

Deserving special comment is the case of the homogeneous spherical mass of paramagnetic matter, which is obviously of particular relevance for neutron stars. In the case of the homogeneous spherical mass of (nonferromagnetic) magnetics, the internal magnetic field is uniform and is expressed by the equations

$$\mathbf{B} + 2\mathbf{H} = 0$$

and

$$\mathbf{B} = \mathbf{H} + 4\pi\mathbf{M},$$

which imply that

$$\mathbf{B} = \frac{8\pi}{3}\mathbf{M};$$

see, for instance, [21, §76, problem 2], where it is emphasized that the latter equations hold for solely nonferromagnetic materials. Substituting this latter value of **B** into (2.16), we find

$$V_M = \sqrt{\frac{2\pi}{3}\frac{M^2}{\rho}}.$$

This form of the speed of the magnetoelastic wave is very similar to that found in [1]. On this ground, we can conclude that magnetoelastic waves is a feature generic to the permanent magnetization of neutron star matter of both ferromagnetic and paramagnetic types. For condensed media possessing a highly pronounced property of magnetic polarizability, the considered magnetoelastic dynamic wave has the same physical significance as the Alfvén magnetohydrodynamic wave does for incompressible magnetoactive plasma.

Quantitatively, the speed of a magnetoelastic wave in paramagnetic neutron matter compressed to the normal nuclear density with the magnetic field strength $B \sim 10^{12}$–10^{14} G, typical of pulsars and magnetars, falls into the interval $10^5 < V_M < 10^6$ cm/s; for comparison, the speed of the zero-temperature longitudinal sound wave is

$$c_s = \sqrt{\frac{v_F^2}{3}} \approx 10^9 \text{ cm/s}.$$

The transverse magnetoelastic wave is therefore a slowly propagating exitation in spin-polarized neutron star matter possessing properties of the degenerate paramagnetic Fermi gas of neutrons.

3. NONRADIAL MAGNETOELASTIC PULSATIONS OF A PERMANENTLY MAGNETIZED NEUTRON STAR

The purpose of the remainder of this paper is to elucidate the character of mechanical distortions of a neutron star caused by strong coupling between fluctuations of the local magnetization and material displacement and their effect on electromagnetic activity of a paramagnetic neutron star. In doing this, we focus on nonradial magnetoelastic pulsations, which are of particular interest in pulsar astrophysics [23–25]. Circumstantial evidence for the neutron star pulsations is the coherence of millisecond micropulses inferred in [26].

The eigenfrequencies of nonradial magnetoelastic pulsations can be computed on the basis of the energy variational principle. The starting point of this method is the energy balance equation

$$\frac{\partial}{\partial t}\int\frac{\rho\delta\mathbf{v}^2}{2}dV = \int[\delta\mathbf{m}\times\mathbf{B}]\cdot\delta\boldsymbol{\omega}\,dV,$$
$$\delta\boldsymbol{\omega} = \frac{1}{2}[\nabla\times\delta\mathbf{v}(\mathbf{r},t)], \tag{3.1}$$

which is obtained by taking the scalar product of (2.10) with $\delta\mathbf{v}$ and integrating by parts over the star volume; the surface integral is then dropped because the crustal material of a neutron star possesses properties of a magnetoactive solid-state plasma in which the magnetic ordering effects are heavily suppressed. The left-hand side of (3.1) exhibits a substantially rotational character of motions accompanying magnetoelastic pulsations of a permanently magnetized neutron star. The next step is to use the factorized representation of the velocity and vorticity fields

$$\delta\mathbf{v}(\mathbf{r},t) = \mathbf{a}(\mathbf{r})\dot{\alpha}(t), \quad \delta\boldsymbol{\omega}(\mathbf{r},t) = \boldsymbol{\phi}(\mathbf{r})\dot{\alpha}(t),$$
$$\boldsymbol{\phi}(\mathbf{r}) = \frac{1}{2}[\nabla\times\mathbf{a}(\mathbf{r})], \tag{3.2}$$

where $\mathbf{a}(\mathbf{r})$ is the field of instantaneous displacements and $\alpha(t)$ defines the temporal evolution of fluctuations.

Inserting (3.2) into (2.11) and eliminating time derivatives, we obtain

$$\delta\mathbf{m}(\mathbf{r}, t) = \boldsymbol{\mu}(\mathbf{r})\alpha(t),$$
$$\boldsymbol{\mu}(\mathbf{r}) = [\boldsymbol{\phi}(\mathbf{r})\times\mathbf{M}] = \frac{1}{2}[[\nabla\times\mathbf{a}(\mathbf{r})]\times\mathbf{M}]. \quad (3.3)$$

Substitution of (3.2) and (3.3) into (3.1) leads to

$$\frac{dH}{dt} = 0, \quad H = \frac{M\dot{\alpha}^2}{2} + \frac{K\alpha^2}{2} \longrightarrow \ddot{\alpha} + \omega^2\alpha = 0, \quad (3.4)$$
$$\omega^2 = \frac{K}{M},$$

where the inertia M and the stiffness K of magnetoelastic vibrations are given by

$$M = \int\rho\mathbf{a}^2 dV,$$
$$K = \chi^{-1}\int\boldsymbol{\mu}^2 dV = \frac{1}{4\chi}\int[[\nabla\times\mathbf{a}]\times\mathbf{M}]^2 dV. \quad (3.5)$$

Thus, computing the frequency of the magnetoelastic mode requires specifying the field $\mathbf{a}$ of instantaneous displacements that have a differentially rotational character, as follows from the expression for the coefficient K of the restoring force of magnetoelastic pulsations.

3.1. Comments on Nonradial Elastic Pulsations of a Solid Star

The eigenmodes of neutron stars associated with deformation properties of incompressible baryon material, highly robust to mechanical distortions, can be specified, as was first suggested in [23], by spheroidal and torsional modes of shear elastic vibrations of a solid sphere. This terminology is due to Lamb [27], who first tackled the latter problem and gave its solution for substantially radial spheroidal and torsional elastic vibrations of a solid sphere (see, e.g., [28]). In the meantime, the case of nonradial pulsations, which is of particular interest in the astrophysics of compact stars, has not been considered in the literature on elasticity and therefore deserves a special analysis. Essentially, the problem is as follows. From classical equation of elastodynamies (2.4), it follows that the field of material displacements

$$\mathbf{u}(\mathbf{r}, t) = \mathbf{a}(\mathbf{r})\alpha(t)$$

corresponding to standing elastic waves of pure shear a solution of the Helmholtz equation

$$\nabla^2\mathbf{u} + k^2\mathbf{u} = 0.$$

Clearly, this equation holds for the solenoidal field of instantaneous displacements,

$$\nabla^2\mathbf{a} + k^2\mathbf{a} = 0,$$

where

$$k^2 = \omega^2/c_t^2$$

and

$$c_t^2 = \mu/\rho$$

is the speed of elastic shear waves in solid bulk. The poloidal solution

$$\mathbf{a}_p = A_p(L)\nabla\times\nabla\times[\mathbf{r}j_L(kr)P_L(z)]$$

describes even-parity spheroidal modes. The toroidal solution

$$\mathbf{a}_t = A_t(L)\nabla\times[\mathbf{r}j_L(kr)P_L(z)]$$

describes odd-parity torsional modes; hereafter, $j_L(kr)$ is the spherical Bessel function and $P_L(z)(z = \cos\theta)$ is the Legendre polynomial of the multipole degree L. General properties of solenoidal vector fields, both the toroidal and the poloidal ones, can be found in [29]. The arbitrary constants and the frequencies of these modes are customarily found from the boundary condition of a stress-free surface,

$$n_k\sigma_{ik}\big|_{r=R} = 0$$

(where n_i are components of the unit vector normal to the surface), which leads to a transcendental dispersion equation whose roots are determined by the nodal structure of Bessel functions. In the case of low-frequency nonradial substantially long wavelengths, $\lambda \longrightarrow \infty$, with

$$k = \omega/c_t = 2\pi/\lambda \longrightarrow 0,$$

the Helmholtz equation of standing shear waves is reduced to the vector Laplace equation for the solenoidal field of elastic displacements,

$$\nabla^2\mathbf{a} = 0, \quad \nabla\cdot\mathbf{a} = 0. \quad (3.6)$$

The poloidal and toroidal solutions of (3.6) are given by [24]

$$\mathbf{a}_p = N_p(L)\nabla\times\nabla\times[\mathbf{r}r^L P_L(z)]$$
$$= N_p(L+1)\nabla r^L P_L(z) \quad (3.7)$$

$$\mathbf{a}_t = N_t(L)\nabla\times[\mathbf{r}r^L P_L(z)]. \quad (3.8)$$

From the standpoint of Lamb's solutions for the fields of displacements, the spherical Bessel function $j_L(kr)$ determining the radial dependence of $\mathbf{a}(r, \theta)$ asymptotically tends in the long-wavelength limit to the function r^L that has no nodes in the interval $0 < r < R$; from this, the term nonradial vibrations is derived. The frequencies of nonradial shear modes can be computed from the above-expanded energy variational principle. Taking the scalar product of Lamé equation (2.4) with

$$u_i(\mathbf{r}, t) = a_i(\mathbf{r})\alpha(t)$$

and integrating over the volume, we obtain

$$M\ddot{\alpha} + K\alpha = 0, \quad M = \int \rho a_i a_i dV,$$
$$K = \frac{\mu}{2}\int\left(\frac{\partial a_i}{\partial x_j} + \frac{\partial a_j}{\partial x_i}\right)^2 dV. \tag{3.9}$$

Substituting into (3.9) the poloidal and the toroidal displacement fields in Eqs. (3.7) and (3.8), respectively, allows us to analytically express the respective frequencies of nonradial spheroidal and torsional shear modes $\omega_s(L)$ and $\omega_t(L)$ of a spherical mass of an elastic solid through the multipole degree L as

$$\omega_s(L) = \omega_E[2(2L+1)(L-1)]^{1/2},$$
$$\omega_t(L) = \omega_E[(2L+3)(L-1)]^{1/2}, \quad \omega_E = \frac{c_t}{R}, \tag{3.10}$$

where $\omega_E = [\mu/(\rho R^2)]^{1/2}$ is the natural unit of frequency of elastic shear vibrations. Equations (3.10) were obtained in recent works [30] in a somewhat different context. The goal of this short comment was to demonstrate the efficiency of the energy variational principle in the study of nonradial vibrations, which allows computing the frequency of both the even-parity *s* mode and the odd-parity *t* mode of the solid sphere on an equal footing. It is also noteworthy that the problem of inertia! waves in a uniformly rotating solid, which, in our opinion, is of particular interest in the study of pulsations of rotating neutron stars, was only recently considered and solved in [31].

The fact that spin-polarized neutron matter can transmit perturbations by transverse waves indicates that the magnetic field penetrating into the body of the star imparts to stellar material a supplementary portion of elasticity generic to solids. In computing periods of nonradial magnetoelastic pulsations of a permanently magnetized neutron star, it therefore seems natural to use the fields of instantaneous displacements corresponding to nonradial spheroidal and torsional shear vibrations of a solid sphere. In doing this, we note that the poloidal vector field associated with spheroidal nonradial pulsations is irrotational,

$$\nabla \times \mathbf{a}_p = 0.$$

This implies that a paramagnetic neutron star does not support nonradial spheroidal pulsations (because the coefficient of the restoring force K in Eq. (3.5) vanishes), but solely supports nonradial torsional shear pulsations coupled with fluctuations in magnetization.

3.2. Periods of Torsional Magnetoelastic Pulsations

Under the global nonradial differentially rotational vibrations of a neutron star, the velocity field of torsional material displacements is described by [24, 25]

$$\delta\mathbf{v}(\mathbf{r}, t) = \dot{\mathbf{u}}(\mathbf{r}, t) = \mathbf{a}_t(\mathbf{r})\dot{\alpha}(t)$$
$$= \frac{1}{2}[\delta\boldsymbol{\omega}(\mathbf{r}, t)\times\mathbf{r}], \tag{3.11}$$
$$\delta\boldsymbol{\omega}(\mathbf{r}, t) = N_t\nabla r^L P_L(z)\dot{\alpha}(t).$$

The constant N_t is eliminated from the boundary condition

$$\delta\mathbf{v}|_{r=R} = \frac{1}{2}[\boldsymbol{\Omega}(t)\times\mathbf{R}], \quad \boldsymbol{\Omega}(t) = \dot{\alpha}(t)\nabla P_L(z),$$

as

$$N_t = \frac{1}{R^{L-1}}. \tag{3.12}$$

The dipole field, with $L = 1$, corresponds to the rigid-body rotation of the star, because the angular velocity becomes a homogeneous vector. The differentially rotational deformations of the star corresponding to quadrupole, $L = 2$, and octupole, $L = 3$, overtones of nonradial torsional pulsations are illustrated in Fig. 1. In spherical polar coordinates, the components of the toroidal field of instantaneous displacements $\mathbf{a}_t(\mathbf{r})$ in the star undergoing torsional nonradial pulsations about the polar axis are given by

$$a_r = 0, \quad a_\theta = 0,$$
$$a_\phi = N_t r^L(1-z^2)^{1/2}\frac{dP_L(z)}{d\mu}. \tag{3.13}$$

Computed with this field, the parameter of inertia as a function of the multipole degree of vibration is given by

$$M(L) = \int \rho \mathbf{a}_t^2 dV = 4\pi\rho R^5\frac{L(L+1)}{(2L+1)(2L+3)}. \tag{3.14}$$

It is easy to see that at $L = 1$, this parameter equals the moment of inertia of rigid sphere,

$$M(L=1) = \frac{2}{5}\mathcal{M}R^2,$$

where

$$\mathcal{M} = \frac{4\pi}{3}\rho R^3$$

is the star mass.

In the general case, the direction of the equilibrium magnetic anisotropy $\mathbf{M}$ can be tilted to the polar axis about which the torsional pulsations of the star occur,

$$M_r = M[(1-z^2)^{1/2}\cos\phi\sin\beta + z\cos\beta],$$
$$z = \cos\theta,$$
$$M_\theta = M[z\cos\phi\sin\beta - (1-z^2)^{1/2}\cos\beta], \tag{3.15}$$
$$M_\phi = -M\sin\phi\sin\beta,$$

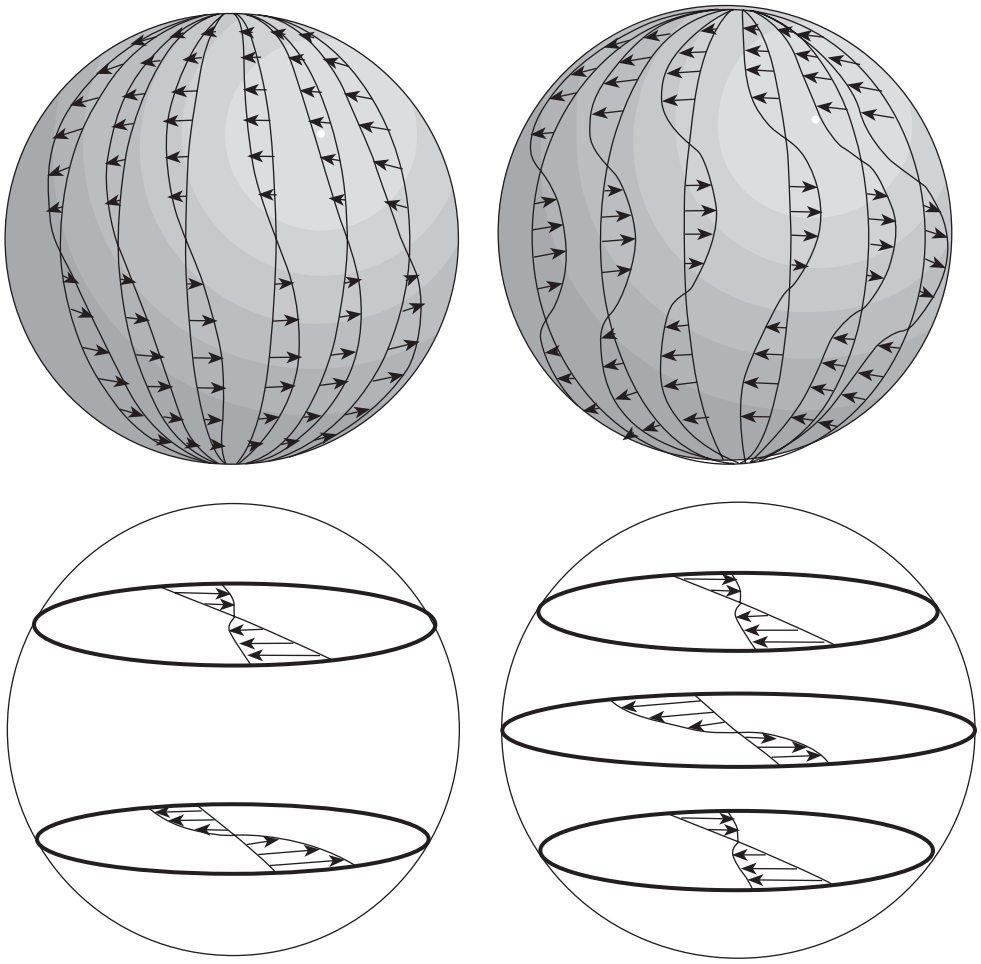


Fig. 1. Geometrical illustration of torsional deformations of a neutron star undergoing quadrupole (left) and octupole (right) nonradial pulsations.

where β is the inclination angle between the polar axis z and the vector $\mathbf{M}$. After a simple, but fairly tedious, calculation of integrals, we obtain the following analytic form of the stiffness:

$$K(L) = \frac{1}{4\chi}\int [[\nabla \times \mathbf{a}_t] \times \mathbf{M}]^2 dV$$

$$= \pi MBR^3 \frac{L(L^2-1)(L+1)}{4L^2-1} \times \cos\beta\left[1 + \frac{3L-1}{2(L-1)}\tan\beta\right]. \quad (3.16)$$

The frequency of a nonradial torsional magnetoelastic mode is given by

$$\omega^2(L) = \omega_M^2(L^2-1)\frac{2L+3}{2L-1}\cos\beta \times\left[1 + \frac{3L-1}{2(L-1)}\tan\beta\right], \quad \omega_M^2 = \frac{V_M^2}{R^2}, \quad (3.17)$$

where ω_M is the natural unit of frequency and $V_M = [MB/4\rho]^{1/2}$ is the speed of the magnetoelastic wave in bulk. This mode can be considered as a magnetoelastic counterpart of Walker's mode for spherical homogeneous mass of a ferromagnetic solid [11]. For the adopted constitutive equation of paramagnetic matter $B = \chi^{-1}M$, this frequency is given by

$$\omega_M^2 = \frac{M^2}{4\chi\rho R^2}. \quad (3.18)$$

For an ideal homogeneous magnetic sphere, with $B = (8\pi/3)M$, this frequency is given by

$$\omega_M^2 = \frac{2\pi}{3}\frac{M^2}{\rho R^2}. \quad (3.19)$$

The corresponding period is $P_M = (2\pi\omega_M)^{-1}$. This mode, which is said to be the magnetotorsional or m/t mode in what follows, is unique to the permanent magnetization of neutron star matter and is an axial or abnormal parity mode. In the case where $\beta = 0$ (the model of the aligned

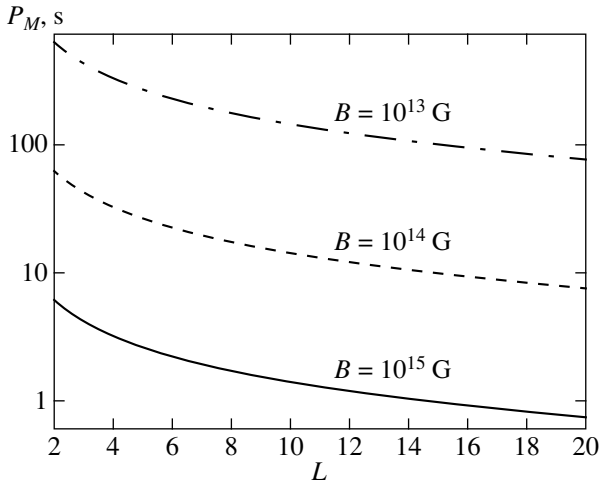


Fig. 2. The period P_M of nonradial torsional magnetoelastic pulsations (in seconds) of a neutron star as a function of the multipole degree L of vibration

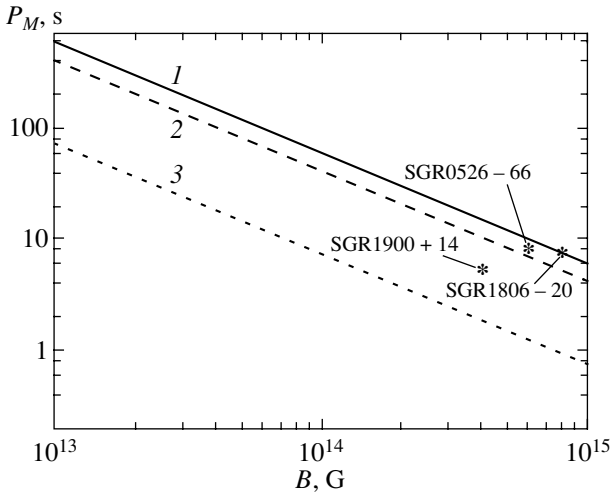


Fig. 3. The period of nonradial torsional magnetoelastic pulsations (in seconds) versus the magnetic field intensity B and data on periods of pulsed gamma emission of soft gamma repeaters taken from [38]; $M = 1.4M_{sun}$, $R = 12$ km; $L = 2$ (*1*), 3 (*2*), 20 (*3*).

magnetic torsator), the frequency of the *m*/*t* mode is given by

$$\omega(L) = \omega_M[(L^2 - 1)(2L + 3)/(2L - 1)]^{1/2}$$

(see [16]). This equation implies that the asymptotic shortening of the period $P(L) = (2\pi\omega(L))^{-1}$ as $L \longrightarrow \infty$ is inversely proportional to the multipole degree of vibrations, $P(L) \propto 1/L$. On the other hand, this indicates the lengthening of periods as the multipole degree of vibration L decreases. It seems quite plausible that, under the implosive effect of a supernova event or starquake, the permanently magnetized core of the nascent neutron star can show a highly restless oscillatory behavior characterized by rather large values of L, whereas a mature object becomes quieter and its transition to lower overtones of magnetoelastic pulsations is accompanied by lengthening of periods.

3.3. Application to Pulsars and Magnetars

To estimate the timing of magnetoelastic pulsations, we evaluate here periods of the *m*/*t* mode for a homogeneous model of a paramagnetic neutron star with the standard parameters, the mass $M = 1.4M_\odot$ and the radius $R = 12$ km, and with the magnetic susceptibility taken from the model of the degenerate paramagnetic Fermi gas of neutrons condensed to the normal nuclear density, which corresponds to the homogeneous neutron star model with the above parameters. In Fig. 2, we plot the period $P(L)$ as a function of the multipole degree of vibration L, computed in the model of the aligned magnetic torsator for the magnetic field intensity typical of both radio pulsars, $B \sim 10^{11}$–10^{13}, and supermagnetic anomalous *X*-ray pulsars and soft gamma repeaters $B \sim 10^{14}$–10^{15}, dubbed magnetars [34]. For a neutron star with the magnetic field of the Crab pulsar, the expected period of the *m*/*t* mode is $P \sim 3$–5 min. It is remarkable that the computed periods are close to those for pulsed gamma emission of currently monitored soft gamma repeaters (see, e.g., [35]). One of the salient features of the soft gamma repeater radiation activity is that they do not display radiation in the radio region. The pulsed gamma emission of soft gamma repeaters becomes well discernible just after highly energetic gamma bursts [36], which are presumably associated with irregular starquakes [37]. In Fig. 3, the period of torsional magnetoelastic pulsations of a paramagnetic neutron star is pictured in juxtaposition with data on the periodic pulsed radiation of soft gamma repeaters. Bearing in mind that the computed periods fall into the realm of pulsed gamma emission of magnetars, we conjecture that the detected 5–10 s periodicity of their pulsed gamma activity is powered by nonradial torsional magnetoelastic vibrations exhibiting permanent magnetization of this class of neutron stars.

3.4. Magnetosphere of a Permanently Magnetized Neutron Star

One more remarkable inference of the model under consideration is that a paramagnetic neutron star undergoing nonradial torsional magnetoelastic pulsations is capable of generating a periodically oscillating electric field inducing the magnetospheric effect that has many features in common with the Goldreich–Julian effect [32, 33]. This can be readily seen from the Minkowski equation describing the dielectric polarizability **D** in

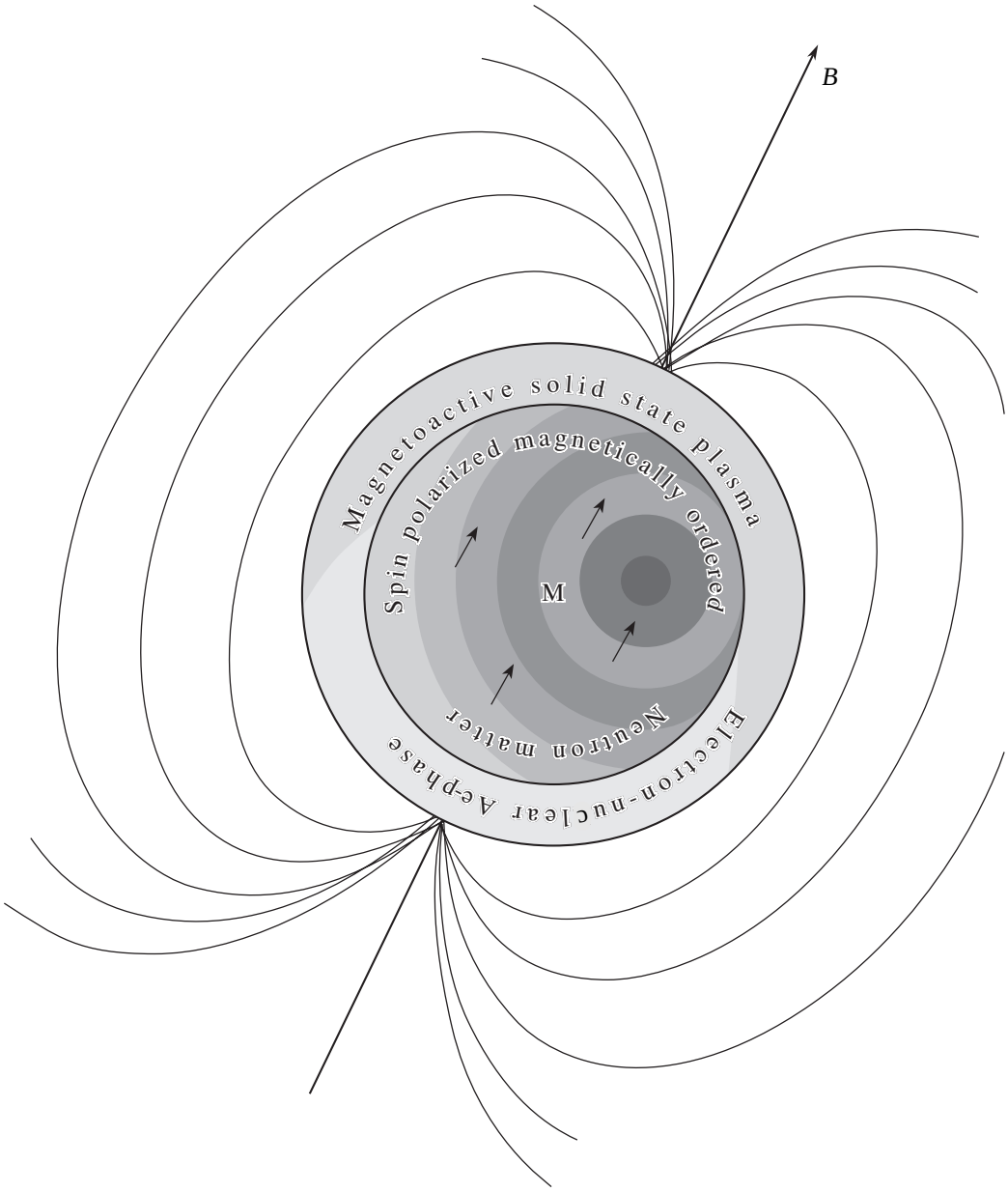


Fig. 4. Cross section of the paramagnetic neutron star structure according to expected electromagnetic properties of stellar matter.

moving permanently magnetized matter of nonferromagnetic type (see, e.g. [21]),

$$\mathbf{D} = \epsilon\mathbf{E} + \frac{4\pi(2\epsilon+1)}{3c}[\mathbf{v}\times\mathbf{M}], \qquad (3.20)$$

under the assumption that the dielectric permeability of spin-polarized baryon matter is infinitely large, $\epsilon \longrightarrow \infty$, as in metallic solids. For a linear, small-amplitude, differentially rotational fluctuations of such matter around the equilibrium state with $\mathbf{v}_0 = 0$, Eq. (3.20) is reduced to

$$\delta\mathbf{D} = \epsilon\delta\mathbf{E} + \frac{4\pi(2\epsilon+1)}{3c}[\delta\mathbf{v}\times\mathbf{M}], \qquad (3.21)$$

and as $\epsilon \longrightarrow \infty$ the last equation becomes[2]

$$\delta\mathbf{E} = -\frac{8\pi}{3c}[\delta\mathbf{v}\times\mathbf{M}], \quad \delta\mathbf{v} = \frac{1}{2}[\delta\boldsymbol{\omega}\times\mathbf{r}]. \qquad (3.22)$$

Identifying the angular velocity magnitude with the frequency of magnetotorsional pulsations, we find the inten-

[2] Equation (3.22) has the same physical meaning as the equation

$$\delta\mathbf{E} = -\frac{1}{c}[\delta\mathbf{v}\times\mathbf{B}], \quad \mathbf{u} = \frac{1}{2}[\delta\boldsymbol{\omega}\times\mathbf{r}],$$

in the Goldreich–Julian theory [32] of pulsar magnetosphere resulting from the perfect conductivity condition $\sigma \longrightarrow \infty$ in Ohm's law:

$$\mathbf{j} = \sigma(\delta\mathbf{E} + \frac{1}{c}[\delta\mathbf{v}\times\mathbf{B}]).$$

sity of equatorial electric field $E \sim 10^{10}$–$10^{12}\, P^{-1}$ V/cm. This field pulls off the charged particles from the star surface and accelerates them along the open magnetic field lines frozen into the neutron star; the electric force $F_e \sim eE$ is much greater than the Newtonian force of gravitational attraction $F_g \sim mg$: the ratio $F_e/F_g \sim 10^8$–10^{10}. The density of the resultant magnetospheric polarization charge is given by

$$\delta\rho = \frac{1}{4\pi}\nabla \cdot \delta\mathbf{E} = -\frac{4}{3c}(\delta\boldsymbol{\omega} \cdot \mathbf{M}). \qquad (3.23)$$

Numerically, the particle density of the polarization charge $\delta n_c = |\delta\rho/e|$ is of the order $10^{-2} B P_M^{-1}$ cm^{-3}. We can expect that magnetoelastic pulsations causing periodic fluctuations of the open magnetic field lines frozen into the star core should affect the electromagnetic (synchrotron and/or curvature [39]) radiation by periodic deviations of the beam direction. For neutron stars with a magnetic field intensity typical of radio pulsars, the above periodicity manifests itself as a long periodic modulation of the main pulse train. In searching for this effect, satellite-based telescopes seem to be more promising, because proper rotation of the Earth highly limits the monitoring time of radio pulsars by stationary Earth-based telescopes. Understandably, this discussion is suggestive rather than conclusive.

4. SUMMARY AND CONCLUSION

While the magnetic flux conservation in the process of contraction of the main sequence star, predicted in [40], serves as a sufficiently reliable guide in estimating the surface magnetic field for both pulsars and magnetars, the electrodynamics of neutron star matter responsible for the long-term stability of such highly intense fields remains one of the challenges in astrophysics of compact stars (e.g., [41]). One of the plausible explanations is that the fossil magnetic field of a collapsed massive star, amplified by processes of catastrophic implosion, resides in the star interior by causing strong spin polarization of baryon matter in the neutron star core such that the main body of the neutron star mass comes into gravitational equilibrium in the state of permanent magnetization promoted by Pauli paramagnetism. The resultant structure of the paramagnetic neutron star relevant to this scenario, pictured in Fig. 4, is thought of as a dense magnetic core (composed of spin-polarized baryon matter) covered by a magnetoactive solid-state plasma (composed of highly mobile electrons and the crystallized structure of immobilized protons and nuclei). It is noteworthy that the presence of a magnetic core provides a natural justification of the magnetoplasma processes in the neutron star crust like Alvfén waves [42] and helicons [43].

To explore characteristic features of electromagnetic activity of a neutron star owing its origin to the permanent magnetization of stellar material, we have considered a highly idealized model of a homogeneous paramagnetic star undergoing global nonradial magnetoelastic pulsations. Highlighted are magnetoelastic dynamics equations adopted from the macroscopic theory of poorly conducting magnetics; it was shown that this theory can be efficiently utilized in the study of motions of permanently magnetized stars associated with large-scale transport of magnetization in an incompressible magnetically ordered stellar matter. What is newly disclosed here is that a permanently magnetized neutron star can support torsional nonradial magnetoelastic pulsations generating the electric field responsible for the neutron star magnetosphere. The net outcome of this paper is that the paramagnetic magnetization of neutron star matter is not inconsistent with the available data on electromagnetic activity of both pulsars and magnetars.

ACKNOWLEDGMENTS

The authors are grateful to Profs. M.G. Park, M. Van Putten, and C.H. Lee for discussion of some problems considered in this paper and to Prof. S.A. Hayward for critical reading of the manuscript. This project is partly supported by the Korea Research Foundation (grant no. 2001-041-D00052) and the Russian Foundation for Basic Research (project no. 00-02-26112).

REFERENCES

1. A. I. Akhiezer, N. V. Laskin, and S. V. Peletminskii, Zh. Éksp. Teor. Fiz. **109**, 1981 (1996) [JETP **82**, 1066 (1996)].
2. M. Kutschera and W. Wójick, Acta Phys. Pol. B **27**, 2277 (1996).
3. P. Haensel and S. Bonazzola, Mon. Not. R. Astron. Soc. **314**, 1017 (1996).
4. M. Kutschera, Mon. Not. R. Astron. Soc. **307**, 784 (1999); T. Tatsumi, Phys. Lett. B **489**, 280 (2000).
5. D. H. Brownell and J. Callaway, Nuovo Cimento B **60**, 169 (1969); S. D. Silverstein, Phys. Rev. Lett. **23**, 139 (1969); M. J. Rice, Phys. Lett. A **29**, 637 (1969).
6. J. W. Clark and N.-C. Chao, Nuovo Cimento Lett. **2**, 185 (1969); J. W. Clark, Phys. Rev. Lett. **23**, 1463 (1969); E. Ostgaard, Nucl. Phys. A **154**, 202 (1970); V. R. Pandharipanda, V. K. Garde, and J. K. Srivastava, Phys. Lett. B **38**, 485 (1970).
7. R. F. O'Connel and K. M. Rousell, Astron. Asprohys. **18**, 198 (1972); J. Pfarr, Z. Phys. A **251**, 152 (1972); J. Schmidt-Burgk, Astron. Astrophys. **26**, 335 (1973); P. Haensel, Phys. Rev. C **11**, 1822 (1975); M. Modarres and J. M. Irvine, J. Phys. G **5**, 1333 (1977).
8. D. N. Sedrakyan, K. M. Shakhbazyan, and A. G. Movsesyan, Astrofizika **21**, 547 (1984); R. F. Sawyer, Phys. Rev. C **40**, 865 (1989); S. Marcos, R. Niembro, M. L. Quelle, and J. Navarro, Phys. Lett. B **271**, 373 (1991); M. Kutschera and W. Wójick, Acta Phys. Pol. B **23**, 947 (1992).

9. A. I. Akhiezer, V. G. Baryakhtar, and S. V. Peletminskii, *Spin Waves* (Nauka, Moscow, 1967; Interscience, New York, 1968).
10. E. M. Lifshitz and L. P. Pitaevski, *Statistical Physics* (Butterworth-Heinenann, Oxford, 1998), Part 2.
11. A. G. Gurevich and G. A. Melkov, *Magnetic Oscillations and Waves* (Nauka, Moscow, 1994).
12. F. Weber, *Pulsars as Astrophysical Laboratories for Nuclear and Particle Physics* (Institute of Physics, Bristol, 1999).
13. G. S. Bisnovaty-Kogan, *Stellar Physics* (Springer-Verlag, Berlin, 2002).
14. H. F. Tiersten, J. Math. Phys. **5**, 1298 (1964).
15. S. Bastrukov, J. Yang, D. Podgainy, and O. Streltsova, in *Hot Points in Astrophysics*, Ed. by V. B. Belyaev and G. V. Efimov (Joint Institute for Nuclear Researches, Dubna, 2000), p. 206.
16. S. Bastrukov, J. Yang, D. Podgainy, and D. F. Weber, in *Explosive Phenomena in Astrophysical Compact Objects*, Ed. by C.-H. Lee, M. Rho, I. Yi, and H. K. Lee; AIP Conf. Proc. **556**, 197 (2001).
17. S. Bastrukov, J. Yang, D. Podgainy, and F. Weber, in *Soft Gamma Repeaters*, Ed. by M. Ferroti and S. Mareghetti (2002).
18. C. Kittel, *Introduction to Solid State Physics* (Wiley, New York, 1996).
19. W. F. Brown, *Magnetoelastic Principles in Ferromagnetism* (North-Holland, Amsterdam, 1962).
20. L. D. Landau, E. M. Lifshitz, A. M. Kosevich, and L. P. Pitaevski, *Theory of Elasticity* (Butterworth-Heinenann, Oxford, 1995).
21. L. D. Landau, E. M. Lifshitz, and L. P. Pitaevski, *Electrodynamics of Continuous Media* (Butterworth-Heinenann, Oxford, 1995).
22. J. Yang and S. I. Bastrukov, Pis'ma Zh. Éksp. Teor. Fiz. **71**, 577 (2000) [JETP Lett. **71**, 395 (2000)]; A. A. Mamun and P. K. Shukla, Pis'ma Zh. Éksp. Teor. Fiz. **75**, 259 (2002) [JETP Lett. **75**, 213 (2002)]; S. I. Bastrukov and J. Yang, Phys. Scr. **65**, 340 (2002).
23. H. M. van Horn, Astrophys. J. **236**, 899 (1980); P. M. McDermott, H. M. van Horn, and C. J. Hansen, Astrophys. J. **325**, 725 (1988); T. E. Strohmayer, Astrophys. J. **372**, 573 (1991).
24. S. I. Bastrukov, F. Weber, and D. V. Podgainy, J. Phys. G **25**, 107 (1999).
25. S. I. Bastrukov, I. V. Molodtsova, D. V. Podgainy, *et al.*, Fiz. Élem. Chastits At. Yadra **30**, 992 (1999) [Phys. Part. Nucl. **30**, 436 (1999)].
26. T. E. Strohmayer, J. M. Cordes, and H. M. van Horn, Astrophys. J. **389**, 685 (1992).
27. H. Lamb, *Hydrodynamics* (Cambridge Univ. Press, Cambridge, 1932; Gostekhizdat, Moscow, 1947).
28. A. E. H. Love, *A Treatise on the Mathematical Theory of Elasticity* (Cambridge Univ. Press, Cambridge, 1927; ONTI, Moscow, 1935); P. M. Morse and H. Feshbach, *Methods of Theoretical Physics* (McGraw-Hill, New York, 1953; Inostrannaya Literatura, Moscow, 1958).
29. S. Chandrasekhar, *Hydrodynamic and Hydromagnetic Stability* (Clarendon, Oxford, 1961).
30. S. I. Bastrukov, Phys. Rev. E **49**, 3166 (1994); S. I. Bastrukov and D. V. Podgainy, Physica A (Amsterdam) **250**, 345 (1998).
31. I. E. Dzyaloshinskii, Phys. Lett. A **202**, 403 (1995).
32. P. Goldreich and W. H. Julian, Astrophys. J. **157**, 869 (1969).
33. V. S. Beskin, A. V. Gurevich, and Ya. N. Istomin, *Physics of the Pulsar Magnetosphere* (Cambridge Univ. Press, Cambridge, 1992); F. C. Michel and H. Li, Phys. Rep. **318**, 227 (1999).
34. R. C. Duncan and C. Thompson, Astrophys. J. **392**, L9 (1992); B. Paczyński, Acta Astron. **42**, 145 (1992).
35. K. Hurley, C. Kouveliotou, P. Woods, *et al.*, Astrophys. J. Lett. **519**, L143 (1999).
36. O. Blaes, R. Blandford, P. Goldreich, and P. Madau, Astrophys. J. **343**, 839 (1989); L. M. Franco, B. Link, and R. I. Epstein, Astrophys. J. **543**, 987 (2000).
37. B. Cheng, R. I. Epstein, R. A. Guyer, and C. Young, Nature **382**, 518 (1996); R. Duncan, Astrophys. J. **498**, L45 (1998).
38. C. Y. Cardall, M. Prakash, and J. Lattimer, Astrophys. J. **554**, 322 (2001).
39. K. Hirotani, Astrophys. J. **549**, 495 (2000).
40. V. L. Ginzburg, Dokl. Akad. Nauk SSSR **156**, 43 (1964) [Sov. Phys. Dokl. **9**, 329 (1964)]; L. Woltjer, Astrophys. J. **140**, 1309 (1964).
41. G. Chanmugam, Annu. Rev. Astron. Astrophys. **30**, 143 (1992); V. Trimble, Beam Line (Stanford) **25**, 41 (1995); S. Konar and D. Bhattacharya, Mon. Not. R. Astron. Soc. **308**, 795 (1999).
42. S. I. Bastrukov, V. V. Papoyan, and D. V. Podgainy, Pis'ma Zh. Éksp. Teor. Fiz. **64**, 593 (1996) [JETP Lett. **64**, 637 (1996)]; Astrophysics **40**, 46 (1997).
43. P. Goldreich and A. Reisenegger, Astrophys. J. **395**, 250 (1992); S. I. Vainshtein, S. M. Chitre, and A. V. Olinto, Phys. Rev. E **61**, 4422 (2000).

Journal of Experimental and Theoretical Physics, Vol. 95, No. 5, 2002, pp. 800–804.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 122, No. 5, 2002, pp. 928–932.

NUCLEI, PARTICLES, AND THEIR INTERACTION

Capture of Ultracold Neutrons in Thin Absorbing Films

V. V. Lomonosov* and A. I. Gurevich

Russian Research Centre Kurchatov Institute, pl. Kurchatova 1, Moscow, 123182 Russia

**e-mail: lom@alpha.ru*

Received May 18, 2002

Abstract—The dependence of the ultracold neutron capture cross section in targets with a thickness smaller than the neutron wavelength is calculated in the time-dependent quantum theory. It is shown that, for low velocities of neutrons, their capture cross section $\sigma_c \sim v$, i.e., tends to zero as the neutron velocity v tends to zero.

1. INTRODUCTION

This study was initiated by the experimental results [1] on transmission of ultracold neutrons through samples containing natural gadolinium and enriched in the Gd^{157} isotopes (see also [2]). Two isotopes, Gd^{155} and Gd^{157}, have resonances in the radiative capture cross section, which lie in the vicinity of the thermal region. For this reason, natural gadolinium is characterized by a very large capture cross section in the thermal region. If we assume that the dependence of the capture cross section on the neutron velocity v is determined by the $1/v$ law [3], the value of the capture cross section for ultracold neutrons becomes abnormally large (about 15 Mb). The measurement of the dependence of the capture cross sections for neutrons on their velocity in [1] proved that it is described by the $1/v$ law down to velocities equal approximately to 9 m/s, while considerable deviations from this law are observed for lower velocities, and the capture cross section starts decreasing. Thus, the $1/v$ law in this neutron velocity region becomes inapplicable. Rauch *et al.* [1] explain this effect by fluctuations of the number of scattering nuclei in the interaction volume, which are associated with specific samples. Such an explanation is possible in principle, but gives rise to a number of other questions. First, does the $1/v$ law hold for perfectly fabricated targets? Second, do physical limitations on the magnitude of the capture cross section exist when the neutron velocity tends to zero? And third, the dependence of the cross section of capture at nuclei constituting dense matter may differ considerably from this law since the $1/v$ law is valid only for the interaction of a neutron with a free nucleus. We will consider the latter circumstance in greater detail. With decreasing velocity, the neutron wavelength λ increases together with the volume of the substance with which the neutron interacts. For ultracold neutrons, the number of nuclei in the interaction volume is $N \approx n\lambda^3 \gg 1$, where n is the number density of nuclei in the substance. Since the elastic scattering of a neutron from this ensemble occurs coherently, the elastic scattering cross section is a nonlinear function of N, while the cross sections of inelastic processes depend on N linearly in the first approximation. Consequently, the relations between elastic and inelastic processes for an ensemble of nuclei and for an individual nucleus, for which the conditions of coherent scattering hold, differ significantly. This effect of suppression of inelastic channels [4] for Bragg scattering of thermal neutrons by ideal crystal is well known. A similar effect is also observed in the case of Bragg scattering of X rays [5]. Coherent scattering is possible for thermal neutrons and X rays whose wavelength is comparable with the separation between scattering centers only if Bragg's condition is satisfied. Coherent scattering of ultracold neutrons, whose wavelength is much larger than the separation between nuclei, by an ensemble of target nuclei occurs irrespective of the fulfillment of Bragg's conditions. Hence the suppression of inelastic channels must be observed under certain physical conditions. In order to simplify calculations, we assume that the wavelength of ultracold neutrons $\lambda \le d$, where d is the thickness of the target. We will use the single-resonance approximation for considering the neutron capture process. For isotopes Gd^{155} (the energy of the first resonance is $\varepsilon_0 \approx 0.0268$ eV and the energy of the second resonance is $\varepsilon_1 \approx 2$ eV) and Gd^{157} ($\varepsilon_0 \approx 0.0314$ eV and $\varepsilon_1 \approx 2.825$ eV), the total width of the resonances is much smaller than the distance between them; consequently, the condition for the single-resonance approximation is satisfied.

2. QUANTUM-MECHANICAL DESCRIPTION OF PASSAGE OF ULTRACOLD NEUTRONS THROUGH A SUBSTANCE

Time evolution of the quantum-mechanical system neutrons + target nuclei is determined by the total Hamiltonian. We divide the total Hamiltonian into two components: H_0, which determines the state of target

nuclei and free motion of the incident neutron beam, and the Hamiltonian V, which is responsible for the interaction between projectile neutrons and target nuclei. In the resonance approximation, we expand the total wave function of the system neutrons + target nuclei into a series in the eigenfunctions of operator H_0:

$$\Psi(t) = A_{n_\mathbf{p}}(t)\phi_\mathbf{p}\Phi_N + \sum_{i=1}^{N} C^i_\mathbf{p}(t)\Phi^*_N + \sum_{\mathbf{p}'} B_{\mathbf{p},\mathbf{p}'}(t)\phi_{\mathbf{p}'}\Phi_N + \sum_{\mathbf{k}} B_{\mathbf{p},\mathbf{k}}(t)\Phi_N|1\mathbf{k}\rangle. \quad (1)$$

Here, $A_{n_\mathbf{p}}(t)$ is the amplitude of the state in which the target nuclei are in the ground state and there exists an initial distribution of the incident neutron flow; $C^i_\mathbf{p}(t)$ is the amplitude of the state in which a target nucleus at point $\mathbf{r}_i$ is excited and one neutron from the initial distribution with momentum $\mathbf{p}$ is absorbed; $B_{\mathbf{p},\mathbf{p}'}(t)$ is the amplitude of the state in which target nuclei are in the ground state, a neutron with momentum $\mathbf{p}$ is absorbed from the initial state, and a neutron with momentum $\mathbf{p}'$ is emitted; and $B_{\mathbf{p},\mathbf{k}}(t)$ is the amplitude of the state in which a neutron with momentum $\mathbf{p}$ is absorbed from the initial distribution and a γ quantum with momentum $\mathbf{k}$ is emitted; one of target nuclei has increased its mass number by unity.

Using the Heitler method [6], we can write the following system of equations for Fourier transforms of amplitudes, which is required for the solution of the time-dependent Schrödinger equation:

$$\begin{aligned}
(\omega + i\epsilon)A_{n_\mathbf{p}}(\omega) &= 1 + \sum_{i=1}^{N}\sum_{\mathbf{p}} V^*_\mathbf{p} e^{-i\mathbf{p}\cdot\mathbf{r}_i} C^i_\mathbf{p}(\omega),\\
(\omega - \varepsilon_0 + \varepsilon_p + i\epsilon)C^i_\mathbf{p}(\omega) &= V_\mathbf{p} e^{i\mathbf{p}\cdot\mathbf{r}_i} A_{n_\mathbf{p}}(\omega)\\
&+ \sum_{\mathbf{p}'} V_{\mathbf{p}'} e^{i\mathbf{p}'\cdot\mathbf{r}_i} B_{\mathbf{p},\mathbf{p}'}(\omega) + \sum_{\mathbf{k}} V_\mathbf{k} e^{i\mathbf{k}\cdot\mathbf{r}_i} B_{\mathbf{p},\mathbf{k}}(\omega),\\
(\omega - \varepsilon_{p'} + \varepsilon_p + i\epsilon)B_{\mathbf{p},\mathbf{p}'}(\omega) &= \sum_{i=1}^{N} V^*_{\mathbf{p}'} e^{-i\mathbf{p}'\cdot\mathbf{r}_i} C^i_\mathbf{p}(\omega),\\
(\omega - \omega_k + \varepsilon_p + \Delta M + i\epsilon)B_{\mathbf{p},\mathbf{k}}(\omega) &= \sum_{i=1}^{N} V^*_\mathbf{k} e^{-i\mathbf{k}\cdot\mathbf{r}_i} C^i_\mathbf{p}(\omega).
\end{aligned} \quad (2)$$

Here, $V_\mathbf{p}$ and $V_{\mathbf{p}'}$ are the matrix elements corresponding to the absorption of a neutron with momenta $\mathbf{p}$ and $\mathbf{p}'$, respectively, by a target nucleus; $V_\mathbf{k}$ is the matrix element corresponding to the absorption of a γ quantum by a compound nucleus; ω_k, ε_p, and $\varepsilon_{p'}$ are the energies of a γ quantum, a neutron with momentum $\mathbf{p}$, and a neutron with momentum $\mathbf{p}'$, respectively; $\Delta M = E_{A+1} - E_A$ is the binding energy for a neutron in a compound nucleus with $A + 1$ nucleons, and ϵ is a quantity tending to zero. While writing the system of equations (2), we assumed, for simplicity of calculations, that a compound nucleus emits only one γ quantum, and there is no cascade emission of quantum. Substituting the expressions for $B_{\mathbf{p},\mathbf{p}'}(\omega)$ and $B_{\mathbf{p},\mathbf{k}}(\omega)$ into the equation for $C^i_\mathbf{p}(\omega)$ and using the equality

$$\xi(x) = \frac{1}{x + i\epsilon} = P\left(\frac{1}{x}\right) - i\pi\delta(x),$$

we obtain

$$\begin{aligned}
&\left(\omega - \varepsilon_0 + \varepsilon_p + i\frac{\Gamma_n}{2} + i\frac{\Gamma_\gamma}{2}\right)C^i_\mathbf{p}(\omega)\\
&\quad = V_\mathbf{p} e^{i\mathbf{p}\cdot\mathbf{r}_i} A_{n_\mathbf{p}}(\omega)\\
&\quad - \left(\frac{\Gamma_n}{2}\sum_{i\neq j}^{N}\frac{\exp(ip|\mathbf{r}_i - \mathbf{r}_j|)}{p|\mathbf{r}_i - \mathbf{r}_j|} + \frac{\Gamma_\gamma}{2}\right.\\
&\quad \left.\times \sum_{i\neq j}^{N}\frac{\exp(ik|\mathbf{r}_i - \mathbf{r}_j|)}{k|\mathbf{r}_i - \mathbf{r}_j|}\right)C^j_\mathbf{p}(\omega),
\end{aligned} \quad (3)$$

where

$$\Gamma_n = 2\pi\sum_{\mathbf{p}'}|V_{\mathbf{p}'}|^2\xi(\omega - \varepsilon_{p'} + \varepsilon_p),$$

$$\Gamma_\gamma = 2\pi\sum_{\mathbf{k}}|V_\mathbf{k}|^2\xi(\omega - \omega_k + \varepsilon_p + \Delta M),$$

and we have also taken into account the fact that $|\mathbf{p}| \approx |\mathbf{p}'|$ in the case of elastic scattering of neutrons. For ultracold neutrons, the condition $\Gamma_\gamma p/\Gamma_n k \ll 1$ usually holds ($\Gamma_\gamma \sim 10^3\Gamma_n$, $p \sim 10^5$ cm^{-1}, $k \sim 10^{10}$ cm^{-1} for an energy of a quantum of the order of 1 MeV, and $\Gamma_\gamma p/\Gamma_n k \sim 10^{-2}$). If we introduce the number density $n(\mathbf{r})$ of target nuclei in this approximation and denote $\Gamma = \Gamma_n + \Gamma_\gamma$, we can write expression (3) in the form

$$\begin{aligned}
&\left(\omega - \varepsilon_0 + \varepsilon_p + i\frac{\Gamma}{2}\right)C_\mathbf{p}(\mathbf{r},\omega) = V_\mathbf{p}\exp(i\mathbf{pr})A_{n_\mathbf{p}}(\omega)\\
&\quad - \frac{\Gamma_n}{2}\int d\mathbf{r}' n(\mathbf{r}')C_\mathbf{p}(\mathbf{r}',\omega)\frac{\exp(ip|\mathbf{r} - \mathbf{r}'|)}{p|\mathbf{r} - \mathbf{r}'|}.
\end{aligned} \quad (4)$$

For the sake of definiteness, we denote the incident neutron flow along the normal to the target surface. In this case, we have $n(\mathbf{r}) = n$ for all x, y for $0 \le z \le d$ and

$n(\mathbf{r}) = 0$ at all other points in the space. It is obvious from symmetry considerations that $C_{\mathbf{p}}(\mathbf{r}, \omega)$ cannot depend on coordinates x and y. Consequently, we can integrate Eq. (4) with respect to these coordinates:

$$\left(\omega - \varepsilon_0 + \varepsilon_p + i\frac{\Gamma}{2}\right)C_{\mathbf{p}}(z, \omega) = V_{\mathbf{p}}\exp(ipz)A_{n_{\mathbf{p}}}(\omega) - i\frac{\pi\Gamma_n n}{p^2}\int_0^d dz' C_{\mathbf{p}}(z', \omega)\exp(ip|z - z'|). \quad (5)$$

Differentiating Eq. (5) twice with respect to z, we obtain the differential equation

$$C''_{\mathbf{p}}(z, \omega) + \alpha^2 C_{\mathbf{p}}(z, \omega) = 0, \quad (6)$$

where

$$\alpha^2 = p^2\left(1 - \frac{2\pi\Gamma_n}{p^3}\frac{1}{\omega - \varepsilon_0 + \varepsilon_p + i\Gamma/2}\right). \quad (7)$$

The solution to Eq. (6) has the form

$$C_{\mathbf{p}}(z, \omega) = Ae^{i\alpha z} + Be^{-i\alpha z}. \quad (8)$$

Coefficients A and B can be determined from the integral equation (5) by substituting solution (8) at points $z = 0$ and $z = d$. The general form of coefficients A and B is rather cumbersome; if, however, condition $pd \ll 1$ holds, the expressions for coefficients A and B can be simplified considerably and have the form

$$A = \frac{V_{\mathbf{p}}A_{n_{\mathbf{p}}}(\omega)}{\omega - \varepsilon_0 + \varepsilon_p + i(\Gamma/2)f(\alpha d)}\frac{1 - e^{-i\alpha d}}{e^{i\alpha d} - e^{-i\alpha d}},$$
$$B = -\frac{V_{\mathbf{p}}A_{n_{\mathbf{p}}}(\omega)}{\omega - \varepsilon_0 + \varepsilon_p + i(\Gamma/2)f(\alpha d)}\frac{1 - e^{i\alpha d}}{e^{i\alpha d} - e^{-i\alpha d}}, \quad (9)$$

where

$$f(\alpha d) = 1 + i\frac{2\pi\Gamma_n}{\Gamma}\frac{n}{p^2\alpha}\frac{2 - e^{-i\alpha d} - e^{i\alpha d}}{e^{i\alpha d} - e^{-i\alpha d}}. \quad (10)$$

Substituting expression (9) into the expression for the general solution (8), we obtain

$$C_{\mathbf{p}}(z, \omega) = \frac{V_{\mathbf{p}}A_{n_{\mathbf{p}}}(\omega)}{\omega - \varepsilon_0 + \varepsilon_p + i(\Gamma/2)f(\alpha d)} \times \frac{e^{i\alpha z}(1 - e^{-i\alpha d}) - e^{-i\alpha z}(1 - e^{i\alpha d})}{e^{i\alpha d} - e^{-i\alpha d}}. \quad (11)$$

Using this relation and the system of equations (2), we can find all the amplitudes of states and calculate the corresponding cross sections. In order to calculate the capture cross section, we use the technique developed by Heitler [6]. We present amplitude $B_{\mathbf{p},\mathbf{k}}(\omega)$ in the form

$$B_{\mathbf{p},\mathbf{k}}(\omega) = U_{\mathbf{p},\mathbf{k}}(\omega)\xi(\omega - \omega_k + \varepsilon_p + \Delta M)A_{n_{\mathbf{p}}}(\omega). \quad (12)$$

According to [6], the capture cross section can be written in the form

$$\sigma_c = 2\pi\sum_{\mathbf{k}}|U_{\mathbf{p},\mathbf{k}}(\omega = 0)|^2\frac{\delta(E_{in} - E_{fin})}{j_{\mathbf{p}}}, \quad (13)$$

where $j_{\mathbf{p}}$ is the current density of projectile neutrons, and E_{in} and E_{fin} are the initial and final energies of the target + neutron system. The expression for $U_{\mathbf{p},\mathbf{k}}(\omega)$ has the form

$$U_{\mathbf{p},\mathbf{k}}(\omega) = n\int\frac{dxdydzV_{\mathbf{p}}V^*_{\mathbf{k}}\exp(-i\mathbf{k}\cdot\mathbf{r})}{\omega - \varepsilon_0 + \varepsilon_p + i(\Gamma/2)f(\alpha d)} \times \frac{e^{i\alpha z}(1 - e^{-i\alpha d}) - e^{-i\alpha z}(1 - e^{i\alpha d})}{e^{i\alpha d} - e^{-i\alpha d}}. \quad (14)$$

For subsequent calculations, we require an explicit expression for coefficient α in terms of the resonance parameters. For $\omega = 0$, we obtain from Eq. (7)

$$\alpha^2 = p^2\left(1 - \frac{4\pi\Gamma_n^0}{\Gamma}\frac{n}{p_0^3}\frac{\varepsilon_0}{\varepsilon_p}\frac{\Delta - i}{\Delta^2 + 1}\right), \quad (15)$$

where Γ_n^0 is the neutron width of the nuclear zero level p_0 is the neutron momentum for the resonance energy ε_0, and $\Delta = 2(\varepsilon_p - \varepsilon_0)/\Gamma$. The energy of projectile ultracold neutrons satisfies the inequality

$$\varepsilon_p \le \frac{4\pi\Gamma_n^0}{\Gamma}\frac{n}{p_0^3}\varepsilon_0, \quad (16)$$

and the quantity α can be written in the form

$$\alpha \approx \left(\frac{4\pi\Gamma_n^0}{\Gamma}\frac{n}{p_0}\right)^{1/2}\left(1 + \frac{i}{2|\Delta|}\right). \quad (17)$$

Let us consider two limiting cases. Let us suppose that $\alpha d \ll 1$; it can easily be demonstrated that the capture cross section is defined as

$$\sigma_c = \frac{N\sigma_{n\gamma}^0\sqrt{\frac{\varepsilon_0}{\varepsilon_p}}}{\Delta^2 + \left(1 + \delta_1\sqrt{\frac{\varepsilon_0}{\varepsilon_p}}\right)^2}. \quad (18)$$

If $\alpha d \gg 1$, the expression for the capture cross section has the form

$$\sigma_c = \frac{N\sigma_{n\gamma}^0\sqrt{\frac{\varepsilon_0}{\varepsilon_p}}}{\left(\Delta + \delta_2\sqrt{\frac{\varepsilon_0}{\varepsilon_p}}\right)^2 + \left(1 + \delta_2|\Delta|\sqrt{\frac{\varepsilon_0}{\varepsilon_p}}\right)^2}. \quad (19)$$

Here, $\sigma_{n\gamma}^0$ is the capture cross section at exact resonance, N is the number of nuclei in the target, and

$$\delta_1 = \frac{\pi\Gamma_n^0}{\Gamma}\frac{nd}{p_0^2},$$

$$\delta_2 = \sqrt{\frac{\pi\Gamma_n^0 n}{\Gamma p_0^3}}\frac{1}{\Delta^2 + 1}.$$

It can be seen from the expressions (18) and (19) for the capture cross section that, if the inequality $\delta_1\sqrt{\varepsilon_0/\varepsilon_p} \ll 1$ is satisfied in the case (18) and the inequality $\delta_2\sqrt{\varepsilon_0/\varepsilon_p} \ll 1$ is satisfied in the case (19), the quantity

$$\sigma_c \propto \sqrt{\frac{\varepsilon_0}{\varepsilon_p}} \propto \frac{1}{v},$$

i.e., the $1/v$ law holds in this neutron energy range. If, however, $\delta_1\sqrt{\varepsilon_0/\varepsilon_p} \gg 1$ or $\delta_2\sqrt{\varepsilon_0/\varepsilon_p} \gg 1$, the capture cross section

$$\sigma_c \propto \sqrt{\frac{\varepsilon_0}{\varepsilon_p}} \propto v$$

and the value of the capture cross section tends to zero upon a further decrease in the neutron energy. At the same time, it is shown in [7] that the value of the elastic scattering cross section increases upon a decrease in the neutron energy and attains its limiting value of the order of the geometrical size of the target in the limit when the neutron energy tends to zero.

3. DISCUSSION OF RESULTS

Let us compare the obtained results with experimental data on the dependence of the cross section of capture of ultracold neutrons on their velocity (energy) for a metallic Gd^{157} film. For the Gd^{157} isotope, the resonance parameters are $g\Gamma_n^0 \approx 0.56$ MeV, $\Gamma \approx 106$ MeV, and $\varepsilon_0 \approx 0.03$ eV. In the experiments [1], a target of thickness $d \approx 114$ Å was used. On the basis of these values, we calculate the neutron velocity for which the capture cross section starts decreasing. It can easily be shown that this value of Gd^{157} can be determined from the inequality $\delta_1\sqrt{\varepsilon_0/\varepsilon_p} > 1$, under which $\sigma_c \propto v$. Substituting the resonance parameters into this expression, we obtain $v < 4$–5 m/s, which is in good agreement with the results described in [4].

The expressions (18) and (19) derived above for the capture cross section make it possible to determine the maximum value of the cross section. We differentiate expressions (18) and (19) with respect to neutron energy and find the value of energy for which these expressions attain their maximum values. For expression (18), we obtain

$$(\varepsilon_p)_{\max} = \varepsilon_0\frac{\delta_1^2}{\Delta^2 + 1},$$

while the corresponding expression for Eq. (19) is

$$(\varepsilon_p)_{\max} = \varepsilon_0\delta_2^2.$$

Accordingly, the maximum possible values of capture cross sections are given by

$$(\sigma_c)_{\max} = \frac{N\sigma_{n\gamma}^0}{2\delta_1}\frac{1}{\sqrt{\Delta^2 + 1} + 1} \quad (20)$$

for $\alpha d \ll 1$ and

$$(\sigma_c)_{\max} = \frac{N\sigma_{n\gamma}^0}{2\delta_2}\frac{1}{\Delta^2 + 1} \quad (21)$$

for $\alpha d \gg 1$. It should be noted that, if the $1/v$ law held, the value of the neutron capture cross section for $\varepsilon_p = (\varepsilon_p)_{\max}$ under the condition $\alpha d \gg 1$ would be twice as large as the value of the neutron capture cross section given by relation (21).

It should be noted in conclusion that the expression (18) derived above for the capture cross section provides an answer to the academic question concerning the dependence of the neutron capture cross section for an "isolated nucleus" for a neutron velocity tending to zero. Indeed, it follows from relation (18) that, for any nonzero value of parameter δ_1, there exists a neutron energy (or velocity) for which $\delta_1(\varepsilon_0/\varepsilon_p)_{\max}^{1/2} > 1$, and the increase in the value of the capture cross section ceases for energies $\varepsilon_p < (\varepsilon_p)_{\max}$; in the limit of still lower energies, the capture cross section vanishes. The value of δ_1 is determined by the resonance parameters of the nucleus and by the value of the number density of nuclei in the volume $V \sim \lambda^3$ of interaction of a neutron with nuclei. As the neutron velocity tends to zero, the neutron wavelength tends to infinity, and the interaction volume becomes infinitely large; consequently, there always exist nuclei identical to the target nucleus even in an ideal experiment on neutron

scattering by a single isolated nucleus in an infinitely large volume of interaction. Thus, the parameter δ_1 in this case becomes nonzero and, hence, the capture cross section is bounded.

REFERENCES

1. H. Rauch, M. Zawisky, and Ch. Stellmach, Phys. Rev. Lett. **83**, 4955 (1999).
2. J. Felber, R. Gähler, and R. Golub, Phys. Rev. Lett. **85**, 5667 (2000); H. Rauch, M. Zawiski, Ch. Stellmach, and P. Geltenbort, Phys. Rev. Lett. **85**, 5668 (2000).
3. I. I. Gurevich and L. V. Tarasov, *Low-Energy Neutron Physics* (Nauka, Moscow, 1965; North-Holland, Amsterdam, 1968).
4. A. M. Afanas'ev and Yu. Kagan, Zh. Éksp. Teor. Fiz. **48**, 327 (1965) [Sov. Phys. JETP **21**, 215 (1965)].
5. G. Bormann, Z. Phys. B **127**, 297 (1950).
6. W. Heitler, *Quantum Theory of Radiation* (Oxford Univ. Press, New York, 1954).
7. A. I. Gurevich and V. V. Lomonosov, Zh. Éksp. Teor. Fiz. **109**, 916 (1996) [JETP **82**, 493 (1996)].

Translated by N. Wadhwa

Journal of Experimental and Theoretical Physics, Vol. 95, No. 5, 2002, pp. 805–819.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 122, No. 5, 2002, pp. 933–949.

ATOMS, SPECTRA, RADIATION

Single-Atom Laser: Coherent and Nonclassical Effects in the Regime of a Strong Atom–Field Correlation

S. Ya. Kilin* and T. B. Karlovich

Stepanov Institute of Physics, National Academy of Sciences of Belarus, Minsk, 220072 Belarus
**e-mail: kilin@dragon.bas-net.by*

Received November 13, 2001

Abstract—Based on the approximation of strong correlations between an atom and an intracavity field, which implies the equal probabilities of finding the atom in the ground state and n photons in the field and of finding the atom in an excited state and $n-1$ photons in the field, it is shown that the conditional states of a field generated by a single-atom laser are described by the diagonal part of the generalized coherent Mittag–Leffler state. The quasi-distributions P and Q of the intracavity-field probability amplitude are found, and the boundedness of the Glauber function on a segment is shown. The possibility of inversionless lasing is demonstrated, and the absence of a lasing threshold is found for some region of parameters. The regimes of generation of the amplitude-squeezed states of the field are studied and the parameters of the system providing the maximum squeezing are determined. It is shown that the atom–field states are entangled at weak pump intensities. © *2002 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

A single-atom laser attracts attention not only as a limiting case of a device capable of generating coherent radiation but also as a simplest quantum-optical system that can be completely theoretically described and studied experimentally. The modern experimental methods allow one to isolate an atom in an optical cavity whose Q factor is so high that a single photon is capable of saturating a resonance transition in the atom [1].

Historically, the study of the lasing properties of a single-atom laser has been initiated by the analysis of three- and four-level schemes [2–4], which are commonly used in macroscopic lasers. However, it was found that a simplest two-level scheme provides the observation of all the new effects that were predicted for three- and four-level schemes. Among these effects are the sub-Poisson statistics of light [2, 6], the multipeak fluorescence spectrum [3], and inversionless lasing [6]. It was shown that an incoherently pumped two-level single-atom laser operates similarly to a usual laser, featuring the subthreshold and superthreshold regimes and self-quenching [4, 7].

However, a number of questions and even formulations of the problems have remained outside the scope of these studies, possibly, because of the cumbersome equations used in them. One of them is the question about a state of the field produced in the cavity upon incoherent pumping of the atom. To which class of states does this state belong? What is the reason for the appearance of its nonclassical properties? What is the reason for the inversionless oscillation in a single two-level atom laser? Is the common state of the atom and field entangled?

In this paper, we found a stationary density matrix in a comparatively simple form using the approximation of strong correlations between the atom and field, which corresponds to a large interaction constant between the atom and field. We showed that conditional density matrices of the generated field are described by the diagonal part of the density matrix of the generalized Mittag–Leffler states. These states are the eigenfunctions of the generalized annihilation operator and generalize naturally usual coherent states. We found the quasi-distributions P and Q of the intracavity-field probability amplitude for these states and demonstrated the boundedness of the Glauber function on a segment. This property is inherent in the Glauber function only. Neither the Q nor the W function possesses this property. Due to a strong atom–field correlation, a single-atom laser is also capable of generating generalized coherent states without inversion and even without a threshold in some region of parameters. When the decay rate of the excited state of the atom is lower than the decay rate of the field in the cavity, the fluctuations of the number of photons are lower than the shot noise level. We found a maximum value of the amplitude squeezing for a single-atom laser. The analysis of mixing of the atomic and field subsystems showed that quantum correlations in this system are possible only at relatively weak pump intensities.

The content of the paper is presented in the following way. In Section 2, we derive the master equation for the density matrix of our model. Section 3, which is devoted to the properties of the quasi-probability functions P, Q, and W, is divided into two subsections. In Section 3.1, we present the general and stationary equations for the Glauber P function and the corresponding

equations for the distribution of the number of photons and normally ordered field moments. In this section, we also give the physical interpretation of stationary equations as the balance equations for the level populations of the combined atom–field quantum system. In Section 3.2, we formulate the basic approximation of the paper and present analytic expressions for the P and Q functions. We also discuss in this section the physical reason for the boundedness of the Glauber function on a segment for radiation from a single-atom laser. In Section 4, we consider the distribution of photons and show that a single-atom laser emits radiation corresponding to the generalized coherent Mittag–Leffler states (more exactly, to their diagonal part). In Section 5, we describe the properties of inversionless oscillation in a single-atom laser. Section 6 contains the description of the sub-Poisson statistics of this laser. The classical and quantum correlation properties of the atom and field are discussed in detail in Section 7. The methods of regularization of P functions for the phase-independent states of the field are considered in the Appendix.

2. THE MODEL

A single-atom laser is considered within the framework of a model system consisting of a two-level atom with the ground state $|1\rangle$ and the excited state $|2\rangle$, whose interaction with a resonance mode of the field is described by the Janes–Cummings Hamiltonian with the interaction constant g. The atom is pumped incoherently with the mean rate R_{12}. In addition, the decay of the resonance mode of the field with the rate k and the decay of the atom with the rate R_{21} are taken into account. The master equation for the density matrix, reduced over the states of the surroundings, in the interaction representation has the form

$$\begin{aligned}\dot{\hat{\rho}} &= g[a^+\sigma_- - a\sigma_+, \hat{\rho}] - R_{12}L(\sigma_+, \sigma_-)\hat{\rho} \\ &\quad - R_{21}L(\sigma_-, \sigma_+)\hat{\rho} - 2kL(a, a^+)\hat{\rho}.\end{aligned} \tag{2.1}$$

Here, the transition operators (σ_+, σ_-) of the atom and (a^+, a) of the field satisfy the commutation relations $[\sigma_+, \sigma_-] = 2\sigma_z$, $[a, a^+] = 1$ for Fermi and Bose particles, respectively, and relaxation is described by the Lindblood operator L:

$$2L(x, y)\hat{\rho} = [x, \hat{\rho}y] + [x\hat{\rho}, y].$$

We assume that the frequency of the atomic transition coincides with the frequency of the fundamental mode of the cavity. We consider only optical frequencies in the model of a single-atom laser and neglect the contribution from thermal photons.

3. DISTRIBUTION FUNCTIONS OF THE INTRACAVITY-FIELD PROBABILITY AMPLITUDE

3.1 General Relations

To study the s-ordered quasi-probability functions $R^{(s)}(\alpha)$ of the field (see Appendix), it is convenient to pass from the equation for the density matrix (2.1) to the equation for the normally ordered characteristic operator $F_N \equiv F$:

$$\begin{aligned}\frac{\partial F}{\partial t} &= -g\lambda^*\sigma_- F + g\lambda F\sigma_+ + g\left[\sigma_-, \frac{\partial F}{\partial \lambda}\right] + g\left[\sigma_+, \frac{\partial F}{\partial \lambda^*}\right] \\ &\quad - \frac{R_{21}L(\sigma_+, \sigma_-)F}{2} - \frac{R_{12}L(\sigma_-, \sigma_+)F}{2} \\ &\quad - k_0\left(\lambda^*\frac{\partial F}{\partial \lambda^*} + \lambda\frac{\partial F}{\partial \lambda}\right).\end{aligned} \tag{3.1}$$

This equation can be transformed with the help of the integral transformation (A.1) to a system of linear differential equations in partial derivatives for any of the $R^{(s)}$ functions in the basis of atomic states. The equations are the simplest for the Glauber distribution function $R_{ij}^{(0)} \equiv P_{ij}$. Assuming that the initial state of the field is phase-independent, i.e., the diagonal elements P_{11} and P_{22} depend only on the square of the modulus $|\alpha|^2$ of the field amplitude, while the nondiagonal elements P_{12} and P_{21} are zero, the equations for the elements P_{ij} can be represented as the equation for the elements of the column vector $\mathbf{P}$: $P_1 = P_{11}$, $P_2 = P_{22}$, $P_3 = P_{12}/\alpha^*$, and $P_4 = P_{21}/\alpha$, which depend only on $|\alpha|^2 \equiv p$:

$$\frac{\partial \mathbf{P}}{\partial t} = A\mathbf{P} + \frac{\partial}{\partial p}(B\mathbf{P}). \tag{3.2}$$

Here, matrices A and B are determined by the expressions

$$\begin{aligned}A &= \begin{pmatrix} -R_{12} & R_{21} & gp & gp \\ R_{12} & -R_{21} & -gp & -gp \\ -g & g & -\gamma_{12} & 0 \\ -g & g & 0 & -\gamma_{21} \end{pmatrix}, \\ B &= \begin{pmatrix} 2kp & 0 & -gp & -gp \\ 0 & 2kp & 0 & 0 \\ 0 & -g & 2kp & 0 \\ 0 & -g & 0 & 2kp \end{pmatrix},\end{aligned} \tag{3.3}$$

and the decay rate of the nondiagonal elements is

$$\gamma_{12} = \gamma_{21} = \frac{R_{21} + R_{12} - 2k}{2}.$$

By solving Eqs. (3.2) and using the relation between the s-ordered functions (A.3), we can determine all the

three quasi-probability functions. Note that the system of four linear equations in partial derivatives (3.2) is an initial system for considering statistical properties of the field. By using some approximations, for example, the adiabatic approximation, we can exclude the functions P_{12} and P_{21} from consideration by expressing them explicitly in terms of P_1 and P_2. A system of two equations in second-order partial derivatives obtained by neglecting the terms responsible for spontaneous emission can be reduced to one equation for the unconditional function of the field distribution $P = P_{11} + P_{22}$, which, in turn, using some additional assumptions, can be reduced to the Fokker–Planck equation. In this paper, we will not use the adiabatic approximation and some other approximations which reduce system (3.2) to the Fokker–Planck equation for the function P. In addition, we will not consider here a nonstationary solution of Eqs. (3.2) and will analyze only stationary distribution functions Q_i and P_i and the distributions of the number of photons $\rho_i(n) \equiv \langle n|\langle i|\hat{\rho}|i\rangle|n\rangle$ and of normally ordered moments $m_i(n) \equiv \mathrm{Sp}_f(\langle i|a^{+n}a^n\hat{\rho}|i\rangle)$, which are related to Q_i and P_i [see (A.4). (A.7), (A.10), (A.15)]. According to (3.2), the continuous functions P_1 and P_2 and these two sets of discrete values satisfy in the stationary regime the following system of equations and two recurrent sequences, respectively:

$$\begin{cases} \left[2\nu + p\left(1 - \dfrac{\partial}{\partial p}\right)\right]P_2 + pP_1 = 2\underline{a}^2 P_1, & (3.3a) \\ p\left(1 - \dfrac{\partial}{\partial p}\right)P_2 - pP_1 = \dfrac{2}{\eta}\left(\mu - p\dfrac{\partial}{\partial p}\right)pP, & (3.3b) \end{cases}$$

$$\begin{cases} (2\nu + n + 1)\rho_2(n) + (n+1)\rho_1(n+1) = 2\underline{a}^2\rho_1(n), & (3.4a) \\ \rho_2(n) - \rho_1(n+1) = = \dfrac{2}{\eta}\{(\mu + n + 1)\rho(n+1) - (n+2)\rho(n+2)\}, & (3.4b) \end{cases}$$

$$\begin{cases} (2\nu + n + 1)m_2(n) + m(n+1) = 2\underline{a}^2 m_1(n), & (3.5a) \\ (n+1)m_2(n) + m_2(n+1) - m_1(n+1) = \dfrac{2}{\eta}(m + n + 1)m(n+1). & (3.5b) \end{cases}$$

Here, $P = P_1 + P_2$; $\rho = \rho_1 + \rho_2$; $m = m_1 + m_2$; the parameter $2\nu + 1 = R_{21}/k$ is equal to the normalized decay rate of the excited level $|2\rangle$; the parameter $2\underline{a}^2 = R_{12}/k$ is equal to the excitation rate of incoherently pumped states $|n\rangle$; the quantity $\eta = g^2/k^2$ determines the normalized rate of spontaneous decay to a resonance mode; and the parameter $\mu = \underline{a}^2 + \nu$ is equal to the total dephasing rate caused by incoherent perturbations of R_{12} and R_{21} after subtraction of the half decay rate $k/2$ of the field.

Equations (3.4a) have a simple physical meaning. They describe the balance of the transitions between adjacent groups of states having the same number of excitations (the number of photons if an atom is in the ground state or the number of photons plus unity if an atom is in the excited state). On the left-hand side of Eq. (3.4a), an average number of transitions occurring from the states $|n+1\rangle|1\rangle$ and $|n\rangle|2\rangle$ to the states $|n\rangle|1\rangle$ and $|n-1\rangle|2\rangle$ appears (see Fig. 1), while, on the right-hand side of this equation, the number of reverse transitions appears. Equation (3.4b) can be interpreted as a balance equation for the number of transitions between the states having the same number of photons in the cavity. To do this, it is necessary to introduce the probabilities of spontaneous (w_n) and stimulated (nw_n) transitions induced in an atom by the intracavity field. Then, the number of transitions occurring from the states $|n+1\rangle|2\rangle$ and $|n+1\rangle|1\rangle$ to the states $|n\rangle|2\rangle$ and $|n\rangle|1\rangle$ (see Fig. 1) should be equal to the number of reverse transitions:

$$\begin{aligned}(n+1)\rho_2(n+1) + (n+1)\rho_1(n+1) \\ + (n+1)w_n\rho_1(n+1) = (n+1)w_n\rho_2(n).\end{aligned} \quad (3.6)$$

By equating (3.6) and (3.4b), we find that the normalized density of spontaneous transitions induced by the intracavity field is

$$w_n = \frac{\eta}{2}\left\{\mu + (n+1)\left[1 - \frac{(n+2)\rho(n+2)}{(n+1)\rho(n+1)}\right]\right\}^{-1}. \quad (3.7)$$

It follows from Eq. (3.7) that (i) the probability w_n depends on the pump $\underline{a}^2$, the value of w_n decreasing (self-quenching) with increasing $\underline{a}^2$ (in the general case, when $\underline{a}^2$ exceeds a certain value), and (ii) w_n

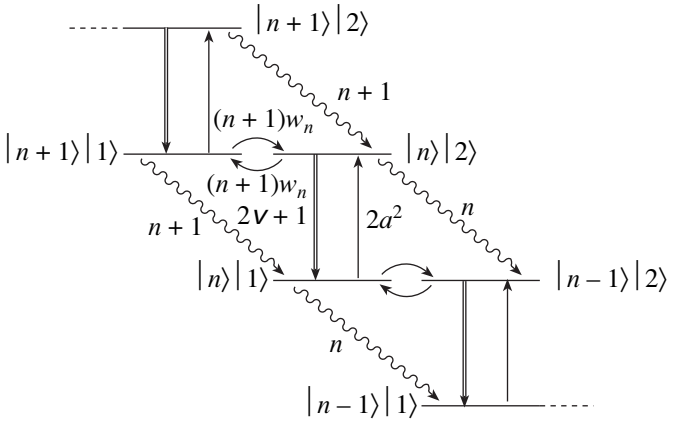


Fig. 1. Energy level diagram and transitions of an atom in a cavity. The splitting of levels having the same number of excitations is not shown. The wavy arrows show transitions involving photons escaping from the cavity (transition rates normalized to the decay rate k are determined by the number of photons in the cavity). The double arrows indicate transitions caused by the decay of the atom (rates are $2\nu + 1$); the arrows directed upward indicate transitions caused by incoherent pumping (rates are $2\underline{a}^2$). The arclike arrows indicate spontaneous and stimulated transitions in the atom, which are induced by the intracavity field (w_n is the spontaneous transition rate in the presence of n photons in the cavity).

depends on the number of photons in the cavity due to a partial destruction of the interference between intracavity photons, which become distinguishable because of the interaction with the atom: photons "belonging" to the atom in the ground state differ from photons corresponding to the excited atomic state. The dependence of w_n on the number of photons is not a priori,[1] but can be determined only after finding the distribution $\rho(n)$ of the number of photons. However, in the limiting case of strong interactions, when the parameter η is so large that the right-hand sides of Eqs. (3.3b) and (3.4b) can be neglected, the dependence of w_n on n is insignificant.

3.2. Strong Interactions ($\eta \gg 1$). The Boundedness of the Glauber Function on a Segment

In the case of strong interactions, a strong coupling appears between population of the levels having the same numbers of excitation. According to (3.4b), in this case,

$$\rho_2(n) = \rho_1(n+1), \tag{3.8}$$

which reflects the leveling of populations of the states $|n\rangle|2\rangle$ and $|n+1\rangle|1\rangle$ before transitions from these states to other states. Such a population leveling results, in combination with the balance of the number of transitions between adjacent groups of states with the same number of excitations, in a simple algebraic relation between the conditional densities of quasi-probabilities P_1 and P_2:

$$\nu P_2 = (a^2 - p)P_1. \tag{3.9}$$

This relation is obtained by subtracting Eq. (3.3b) from (3.3a) under the condition $\eta \gg 1$. Algebraic relation (3.9), together with differential equation (3.3a), allows us to find the functions P_1 and P_2 in the explicit form:

$$P_1(p) = c\theta(\underline{a}^2 - p)\frac{(\underline{a}^2 - p)^{\nu-1}}{\Gamma(\nu)}e^p, \tag{3.10a}$$

$$P_2(p) = c\theta(\underline{a}^2 - p)\frac{(\underline{a}^2 - p)^{\nu}}{\Gamma(\nu+1)}e^p. \tag{3.10b}$$

Here, $\theta(x)$ is the Heaviside step function, and the normalization constant

$$c = \frac{1}{\pi}\Gamma(\nu+1)\underline{a}^{-2\nu}\left[1 + 2\frac{\underline{a}^2}{\nu+1}N_{\nu+1}(\underline{a}^2)\right]^{-1} \tag{3.11}$$

is expressed in terms of the incomplete gamma function $\gamma(\nu, x)$ or the Mittag–Leffler function [8]

$$E_{\alpha,\beta}(x) = \sum_n \frac{x^n}{\Gamma(\alpha n + \beta)}$$

as

$$N_\nu(x) = \nu x^{-\nu}e^x\gamma(\nu, x) \equiv \Gamma(1+\nu)E_{1,1+\nu}(x). \tag{3.12}$$

According to (3.10), the Glauber functions P_1 and P_2 are expressed in terms of the kernel of the convolution

$$I^\beta * f(x) = \frac{\theta(x)}{\Gamma(\beta)}x^{\beta-1} * f(x)$$

$$= \frac{1}{\Gamma(\beta)}\int_0^x (x-t)^{\beta-1}f(t)dt,$$

which is called the fractional Riemann–Liouville integration of order β [9]. The order $\nu \equiv R_{21}/2k - 1/2$ corresponds to the function $P_1(p)$, and the order $\nu + 1$ corresponds to the function $P_2(p)$. In the case of the negative order of the fractional integration, this operation is called the fractional differentiation because, when n is an integer, we have

$$I^{-n} * f(x) = \left(\frac{\partial}{\partial x}\right)^n f(x).$$

It is obvious that in this case the transformation kernel is a generalized function. According to (3.10), the Glauber function $P_2(p)$ is always a positively defined distribution function, whereas the function $P_1(p)$ for

[1] Note that the dominator in Eq. (3.7), which is proportional, according to (3.4b), to the difference $\rho_2(n) - \rho_1(n+1)$, does not vanish at any values of the parameters, i.e., the function w_n is always a positively defined quantity, which can be proved by the contradiction method using Eq. (3.4a).

$-0.5 \le \nu \le 0$ is no longer a distribution function and becomes a generalized function. The nonclassical properties of light for such a state are considered in Section 6.

The characteristic feature of functions P_1 and P_2 is their boundedness on the segment $[0, \underline{a}^2]$, which means that the field amplitudes exceeding $\underline{a}$ cannot be observed. The presence of the boundary value of the field amplitude is explicitly manifested in the interval $0 \le \nu \le 1$, when $1 < R_{21}/2k < 3$ (see Fig. 2). To demonstrate the fact that the boundedness of the P functions is not related to the strong interaction approximation, which we used in deriving expression (3.10), we calculated numerically the P function by solving exact algebraic Eqs. (3.4) and (3.5) and using regularized expansions (A.10) and (A.14) of the P functions in the Laguerre polynomials. One can see from Fig. 2 (curve *2*) that the maximum of the distribution probability also corresponds to the boundary value of $\underline{a}^2$, but outside the maximum the P function exhibits nonclassical properties, oscillating near zero.

It is obvious that, because the features of the P function are smoothed when the field amplitude is described in terms of the Q or W functions (A.3), the existence of the upper bound for the field amplitude does not restrict the Q and W functions on a segment. Thus, the Q_i functions are determined, in the limit under study, according to (3.9), (3.10), and (A.3), by the expressions

$$Q_1(p) = c\Gamma(\nu+1)e^{-p}\left(\frac{\underline{a}}{\sqrt{p}}\right)^{\nu} I_{\nu}(2a\sqrt{p}), \qquad (3.13a)$$

$$\begin{aligned} Q_2(p) &= e^{-p}\frac{\partial}{\partial p}(Q_1 e^{p}) \\ &= c\Gamma(\nu+1)e^{-p}\left(\frac{\underline{a}}{\sqrt{p}}\right)^{\nu+1} I_{\nu+1}(2\underline{a}\sqrt{p}), \end{aligned} \qquad (3.13b)$$

where $I_\nu(x)$ is the modified Bessel function of order ν. It follows from these expressions that the functions $Q_1(p)$ and $Q_2(p)$ do not vanish at any value of the field amplitude (see Fig. 2).

The boundary value of the field amplitude $|\alpha_b|^2 = \underline{a}^2 = R_{12}/2k$ corresponds to the statistical equality of the number $R_{12}\Delta t$ of excitations of the atom during the interval Δt upon incoherent pumping (i.e., the number of the $|1\rangle|n\rangle \longrightarrow |2\rangle|n\rangle$ transitions followed by the $|2\rangle|n\rangle \longrightarrow |1\rangle|n+1\rangle$ transition, which increases the number of intracavity photons by unity and occurs, according to (3.8), with the probability 1/2) and the number $2k|\alpha_b|^2\Delta t$ of transitions from the states $|1\rangle|n\rangle$ ($n \approx |\alpha_b|^2$) to the states $|1\rangle|n-1\rangle$ (i.e., transitions that reduce the number of intracavity photons by unity). Such a statistical equilibrium, along with the multiphoton interference corresponding to the representation of the Glauber function in the form of superposition contributions from different n-photon states [see (A.10) and (A.14)], causes the boundedness of the functions P_1 and P_2 on a segment. Note that, along with the fact that the functions $Q_i(p)$ and $W_i(p)$ do not possess the boundedness on a segment, the distribution functions of the number of photons $\rho_1(n)$ and $\rho_2(n)$ change only weakly above the boundary value $n_b \approx \underline{a}^2$. Let us present, for example, the expressions for the functions P_1 and P_2 for the case when the decay rates of the atom (R_{21}) and field (k) ($\nu = 0$) are the same:

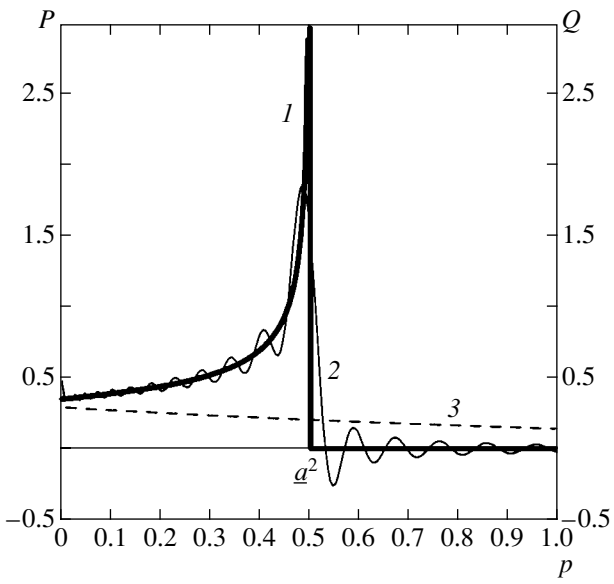


Fig. 2. (*1, 2*) P and (*3*) Q distribution functions calculated for the parameters $\sqrt{\eta} = 200$, $\underline{a}^2 = 1/2$, and $\nu = 1/2$. Curve *1* corresponds to the strong correlation approximation; curve *2* is calculated by solving numerically exact algebraic equations (3.4) and (3.5) and using regularized expansions (A.10) and (A.14) for the number of terms in the series $N = 3500$.

$$\begin{aligned} P_1 &= \frac{\delta(\underline{a}^2 - p)}{\pi[2 - \exp(-\underline{a}^2)]}, \\ P_2 &= \frac{\exp[-(\underline{a}^2 - p)]\theta(\underline{a}^2 - p)}{\pi[2 - \exp(-\underline{a}^2)]}. \end{aligned} \qquad (3.14)$$

Therefore, in this limit, the state of the field, when the atom is in the ground state, is a coherent state averaged over a random phase. The distribution of the number of photons in this state is Poissonian, with the possible detection of photons corresponding to the values of $\underline{a}^2$ exceeding the boundary value. According to (3.8), the distribution of photons for the field corresponding to the atom in the excited state is a displaced Poissonian. For this field, the numbers of photons exceeding $\underline{a}^2$ are also accessible. Note that the Poisson distribution of photons is not caused by intense pumping but appears due to the coherent atom–field interaction. For the

three-level Λ scheme, this possibility was pointed out in paper [4].

4. PHOTON DISTRIBUTION: THE GENERALIZED COHERENT MITTAG–LEFFLER STATES

It follows from Eqs. (3.4a) and (3.8) that, in the limit of strong interactions, the stationary density matrices ρ_1 and ρ_2 can be found by solving the eigenvalue problem

$$\left(1 + \frac{\beta}{aa^+}\right)a\rho_i a^+ = \underline{a}^2\rho_i. \quad (4.1)$$

The parameter $\beta = \nu$ for the density matrix ρ_1 and $\beta = \nu + 1$ for the density matrix ρ_2. Indeed, the calculation of the diagonal elements in the Fock basis from Eq. (4.1) taking into account the known equalities

$$a|n\rangle = \sqrt{n}|n-1\rangle, \quad a^+|n\rangle = \sqrt{n+1}|n+1\rangle$$

yields Eqs. (3.4a). The problem (4.1) can be reformulated as the eigenvalue problem

$$A_\beta|\underline{a}, \beta\rangle = \underline{a}|\underline{a}\beta\rangle \quad (4.2)$$

for the generalized annihilation operator

$$A_\beta = \sqrt{1 + \frac{\beta}{aa^+}}a. \quad (4.3)$$

Then, the solution of (4.1) is described as the diagonal part of the density matrix of the generalized coherent state $|\underline{a}, \beta\rangle$:

$$\rho_1 = n_1\mathrm{diag}\{|\underline{a}, \nu\rangle\langle\underline{a}, \nu|\}, \quad (4.4a)$$

$$\rho_2 = n_2\mathrm{diag}\{|\underline{a}, \nu+1\rangle\langle\underline{a}, \nu+1|\}. \quad (4.4b)$$

Here, n_1 and n_2 are the populations of the ground and excited states, respectively:

$$n_2 = 1 - n_1 = \frac{b}{1+2b}, \quad b = \frac{\underline{a}^2}{\nu+1}N_{\nu+1}(\underline{a}^2);$$

and the normalized generalized coherent states $|\underline{a}, \nu\rangle$ are determined by the superposition of the Fock states:

$$|\underline{a}, \beta\rangle = N_\beta^{-1/2}\sum_{n=0}^{\infty}\underline{a}^n\sqrt{\frac{\Gamma(1+\beta)}{\Gamma(1+\beta+n)}}|n\rangle. \quad (4.5)$$

The normalization constant $N_\beta(\underline{a}^2)$ is expressed in terms of the Mittag–Leffler function $E_{1,1+\beta}(\underline{a}^2)$ (3.12). The authors of paper [8] proposed to call this class of generalized coherent states the coherent Mittag–Leffler states because they are connected with these special functions. It is obvious that, for $\beta = 0$, these states coincide with usual coherent states. For $\beta = 0$, states (4.5) generalize coherent states in the obvious way: in the expansion of the coherent state $|\underline{a}\rangle$ in Fock states, the quantity $n! \equiv \Gamma(1+n)$ is replaced by $\Gamma(1+\beta+n)$, while the exponential normalization constant $\exp(|\alpha|^2)$ is replaced by the function $\Gamma(1+\beta)E_{1,1+\beta}(\underline{a}^2)$. Taking into account that the Fock state $|n\rangle$ is generated from vacuum by the operator $A_\beta^{+n}/\sqrt{\Gamma(1+\beta+n)}$,

$$|n\rangle = A_\beta^{+n}/\sqrt{\Gamma(1+\beta+n)}|0\rangle,$$

we obtain another representation for the states (4.5):

$$|\underline{a}, \beta\rangle = (E_{1,1+\beta}(\underline{a}^2))^{-1/2}E_{1,1+\beta}(\underline{a}A_\beta^+)|0\rangle, \quad (4.6)$$

which is analogous to the generator of coherent states from vacuum.

The generalized deformed annihilation A_β and creation A_β^+ operators satisfy the commutation relations

$$[A_\beta, A_\beta^+] = 1 + \beta|0\rangle\langle 0|, \quad (4.7)$$

which determine the uncertainty relations for the operators A_β and A_β^+ and their combinations. For the states (4.5) and (4.6), these relations represent equalities. Relations (4.7) also show the noncanonical nature of transformations (deformations) of the operators a, a^+ to the operators A_β, A_β^+.

The photon distribution for generalized coherent Mittag–Leffler states and, hence, for the states described by density matrices (4.4) is determined by the expression

$$|\langle n|\underline{a}, \beta\rangle|^2 = (E_{1,1+\beta}(\underline{a}^2))^{-1}\frac{\underline{a}^{2n}}{\Gamma(1+\beta+n)}. \quad (4.8)$$

The properties of the photon distribution are usually related to the lasing properties of the system (the presence of a maximum of the distribution function) and to the stability of the radiation produced (the width of the distribution function compared to that of the Poisson light). Depending on the parameter β, the distribution (4.8) can be both sub-Poissonian ($\beta < 0$) and super-Poissonian $\beta > 0$, retaining its lasing properties.

Before discussing the generation possibilities of a single-atom laser and the stability of the photon distribution, we present another method for the description of systems, which is related to the deformed operators A_β and A_β^+ [10]. By using the relation

$$|n\rangle_\beta = \frac{A_\beta^{+n}}{\sqrt{n!}}|0\rangle,$$

we form a set of basis states which differ from the Fock states by the absence of normalization. It is obvious that, in this deformed basis, the relations

$$A_\beta^+|n\rangle_\beta = \sqrt{n+1}|n\rangle_\beta, \quad A_\beta|n\rangle_\beta = \sqrt{n}|n-1\rangle_\beta,$$

are fulfilled for the operators A_β and A_β^+, which are similar to relations for operators a and a^+ in the usual Fock

space, and hence, the commutation relations $[A_\beta, A_\beta^+] = 1$ are fulfilled. By using this circumstance, we can assume that the Hamiltonian operator of a deformed oscillator is

$$H_\beta = \hbar\omega\frac{A_\beta^+A_\beta + A_\beta A_\beta^+}{2}.$$

In the Fock representation, the eigenvalues of this Hamiltonian are

$$E_\beta(n) = \hbar\omega\left[(1-\delta_{n0})\left(n+\beta+\frac{1}{2}\right)+\delta_{n0}\frac{1+\beta}{2}\right];$$

i.e., the structure of the spectrum of the deformed oscillator coincides with that of a usual oscillator, the exception being the ground state, whose energy differs from that of the $n = 1$ state by the value $\hbar(1 + \beta/2)$, which is greater or smaller than a quantum $\hbar\omega$, depending on the sign of the parameter β. In this approach, the state (4.4a) is related to the deformed oscillator in which the frequency of the $|1\rangle$–$|0\rangle$ transition can be lower than $\hbar\omega$ ($\nu < 0$). For the (4.4b) state, the frequency of a similar transition is always greater than $\hbar\omega$ by $\hbar\omega(1 + (\nu + 1)/2)$.

5. INVERSIONLESS LASING

It follows from photon distributions (4.4) that the population n_2 of the upper laser level is always lower than the population n_1 of the lower level. The question arises of how in this case such a device as a single atom in a high-Q cavity, which is called a single-atom laser [3, 4, 11], can generate emission that is close to coherent emission and whether this emission is produced by overcoming a certain threshold or it appears without a threshold. The question of how to define the lasing threshold in the case when a fraction of spontaneous emission in the cavity mode relative to a total spontaneous emission (in the cavity and noncavity modes) approaches unity ($\beta_c \longrightarrow 1$) has been discussed in detail in the literature [11, 12]. It is obvious that the reasoning applied to a usual laser ($\beta_c \longrightarrow 0$), which is based on the balance rate equations and on a comparison of the average rates of stimulated and spontaneous transitions, as well as the values of the loss rates, gives information on the average emission intensity (an average number of photons) and its increase with increasing the incoherent pump power. However, the nature of emission produced by such a source and its coherent properties remain unclear. To answer this question, it is necessary to study the distribution functions of this emission. Based on the study of these functions, some possible criteria were proposed for a passage from the regime of incoherent emission, which is characterized by an exponential photon distribution, to coherent emission, which is characterized by a Poisson photon distribution with a maximum located not at zero. One of such obvious criteria is the appearance of a maximum of the distribution function $\rho(n)$, which can be determined from the relation $\rho(0) < \rho(1)$ [11, 13]. This criterion is approximate because it cannot distinguish coherent emission from incoherent emission in the region of small amplitudes $|\alpha|^2 < 1$. The application of this criterion for usual lasers coincides with the threshold condition according to which the average rate of stimulated emission – the average rate of absorption = the average rate of the photon loss. However, in the region $\beta_c \longrightarrow 1$ (and we consider namely this operation regime of a single-atom laser), this usual threshold condition, which was obtained from rate equations, is never fulfilled, although, as shown above, emission of a single-atom laser can be coherent. Indeed, it follows from the balance equation (3.6) that

$$k\langle nw_n\rangle_2 - k\langle nw_n\rangle_1 = k\langle n\rangle - k\langle w_n\rangle_2,$$

i.e., the difference of the average rates of stimulated emission

$$k\langle nw_n\rangle_2 \equiv k\sum_n nw_n\rho_2(n)$$

and absorption

$$k\langle nw_n\rangle_1 \equiv k\sum_n nw_n\rho_1(n)$$

is always smaller than the loss rate

$$k\langle n\rangle \equiv \sum_n n(\rho_1(n) + \rho_2(n))$$

by the value of the rate

$$k\langle w_n\rangle_2 = k\sum_n w_n\rho_2(n)$$

of spontaneous emission to a cavity mode of the field, which should be expected for a two-level atom under stationary conditions. This means that photons emitted spontaneously to the cavity mode of the field do not distort completely the coherent properties of emission being produced, and under certain conditions, the device under study can generate highly coherent emission. This is explained by the fact that photons belonging to the same cavity mode cannot be distinguished from each other in principle: there is no difference between photons that appeared in the cavity mode due to stimulated or spontaneous emission processes. In this connection, we should emphasize that, although the role of stimulated processes in the production of coherent emission in a single-atom laser decreases, no coherent emission can be produced only due to spontaneous transitions. In addition, the term “laser” as applied to a device producing coherent emission should be perceived with stipulations as a historical term corresponding to a device that is obtained from a usual laser as a passage to the limit.

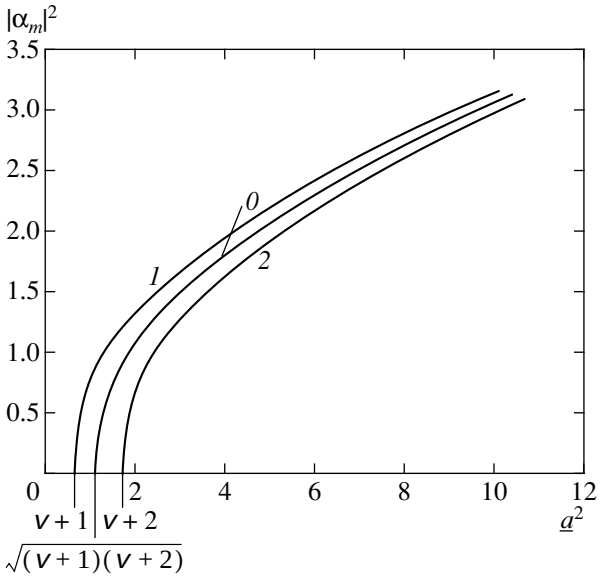


Fig. 3. Dependence of the maximum $|\alpha_m|^2$ of the unconditional (curve *0*) and conditional Q_1 and Q_2 (curves *1* and *2*) functions on the pump intensity $\underline{a}^2$. The pump thresholds $\underline{a}_1^2 = \sqrt{(\nu + 1)(\nu + 2)}$, $\underline{a}_{1(1)}^2 = \nu + 1$, and $\underline{a}_{1(2)}^2 = \nu + 2$ corresponding to the maxima of functions Q, Q_1, and Q_2 are indicated by arrows; the curves are plotted for $\nu = -1/2$.

Taking the above considerations into account, we can show that, when the coherent pump exceeds the value

$$\underline{a}_1^2 = \sqrt{(\nu + 1)(\nu + 2)}, \tag{5.1}$$

the probability $\rho(1)$ of the presence of one photon exceeds the probability of the absence of photons, i.e., the maximum of the distribution $\rho(n)$ is displaced from zero, which can be treated as a lasing threshold (threshold in the photon distribution). As the incoherent pump is further increased, the maximum of the discrete distribution $\rho(n)$ shifts to the region of greater numbers of photons, taking successively the increasing values of m when the pump achieves the m-photon threshold

$$\underline{a}_m^2 = \sqrt{(\nu + m)(\nu + m + 1)}. \tag{5.2}$$

Conditions (5.1) for the appearance of lasing and the dependence (5.2) of the amplitude maximum on the pump can also be obtained by considering the conditional Q functions (3.13). Thus, the maximum of the distributions Q_1 and Q_2 appear when the relations

$$2\underline{a}^2 = x\frac{I_\nu(x)}{I_{\nu+1}(x)}, \quad 2\underline{a}^2 = x\frac{I_{\nu+1}(x)}{I_{\nu+2}(x)},$$

are satisfied, respectively. Here, $x = 2\sqrt{p}\underline{a}$. The solution to these equations for $p = 0$ gives two lasing thresholds for the conditional states of the field, $\underline{a}_{1(1)}^2 = \nu + 1$ and $\underline{a}_{1(2)}^2 = \nu + 2$, while the consideration of the unconditional Q function leads to the threshold equal to (5.1) (see Fig. 3). Note that the inversionless lasing studied here differs from the types of inversionless lasing described in [5] for three- and four-level schemes. The authors of [5] assumed that the inversionless lasing appeared due to a strong interference between adjacent transitions. We assume that the inversionless lasing in our system is caused by a strong atom–field correlation, which is absent in the traditional consideration of spontaneous and stimulated transitions in two-level systems. It is because of such a correlation that relation (3.8) is satisfied, which obviously provides the absence of population inversion in the atomic subsystem: even a significant excess of the pump $2\underline{a}^2$ over losses $2\nu + 1$ is eliminated by an instantaneous leveling of the populations of the states $|n\rangle|2\rangle$ and $|n + 1\rangle|1\rangle$, which corresponds to the reconstruction of the inversionless distribution of populations over atomic states and the increase in the number of photons by unity.

Therefore, in the case of a strong correlation of the atom–field system, an incoherent pumping of the atomic subsystem can excite the photon subsystem to the highly intense states, which are close to coherent states. In this sense, this excitation regime is close to coherent excitation of a field oscillator by an amplitude-stabilized classical current. The latter is especially clearly manifested in the region $0 < \nu < 1$, where, according (3.10), the state of the field is classical (functions P_i are positively defined) and the unconditional Glauber distribution function has a maximum for any arbitrarily low pump intensities, demonstrating excitation of the field states, which are similar to coherent states, without any threshold [4, 7]. A field oscillator excited by a classical current behaves similarly.

The region of applicability of the strong correlation approximation (3.8) is determined by the boundedness of the rate w_n of spontaneous transitions (3.7) induced by the intracavity field. The region of applicability can be estimated from the inequality

$$w_n > \frac{\eta/2}{\mu + n + 1} \gg 1,$$

which shows that, along with the inequalities $g \gg k$, R_{12}, R_{21}, ν, the inequality $g \gg kn$ should also be satisfied, which obviously cannot be valid for arbitrary n for a fixed value of g. Therefore, an increase in the pump $\underline{a}^2$, resulting in the population of excited states of the field, will obviously limit the approximation used. Our numerical calculations show that the strong correlation approximation can be applied until the population of the excited atomic state approaches 1/2, i.e., the parameter b determining saturation is limited [see (4.4)].

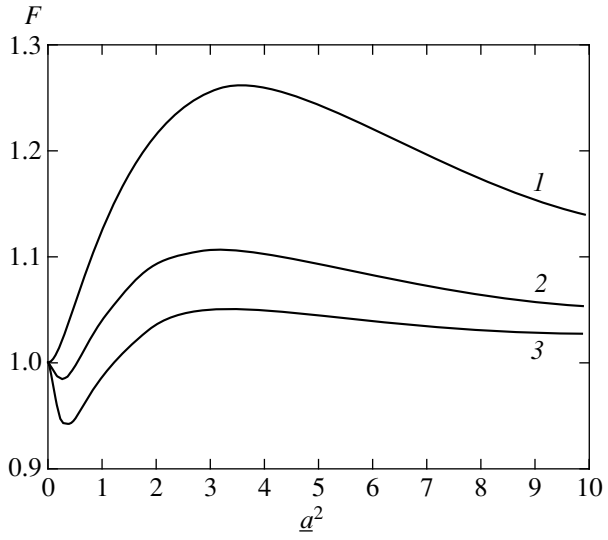


Fig. 4. Fano factor of the intracavity field in the strong correlation approximation as a function of the incoherent pump intensity $\underline{a}^2$ for $\sqrt{\eta}$ = 200 and ν = 1/2 (*1*), –0.25 (*2*), and –0.495 (*3*).

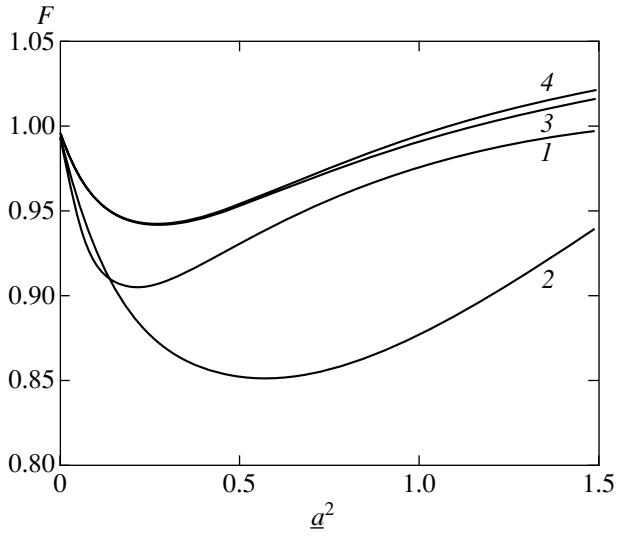


Fig. 5. Fano factor of the intracavity field calculated numerically from exact recurrent relations (3.5) as a function of the incoherent pump intensity for ν = –0.495 and $\sqrt{\eta}$ = 0.6 (*1*), 1.5 (*2*), 20 (*3*), and 200 (*4*).

6. SUB-POISSON STATISTICS OF LIGHT (ν < 0)

As mentioned above, the Glauber function (3.10a) in the region ν < 0 becomes a generalized function of higher order than the usual δ function, which is reflected in the nonclassical properties of radiation emitted by a single-mode laser. The distribution of the number of photons becomes sub-Poissonian. According to (3.10) or (4.4) and (4.5), the normally ordered moments $m_1(n)$ and $m_2(n)$ in the strong correlation approximation are expressed in terms of the moments

$$m_{\underline{a},\beta}(n) = \langle \underline{a}, \beta | a^{+n} a^n | \underline{a}, \beta \rangle = \frac{\underline{a}^{2n}}{N_\beta(\underline{a}^2)} \left(\frac{d}{d\underline{a}^2} \right)^n N_\beta(\underline{a}^2) \quad (6.1)$$

of the generalized Mittag–Leffler states (4.5):

$$m_1(n) = n_1 m_{\underline{a}^2, \nu}(n), \quad m_2(n) = n_2 m_{\underline{a}^2, \nu+1}(n). \quad (6.2)$$

Figure 4 shows the values of the Fano factor

$$F = \frac{\langle n^2 \rangle - \langle n \rangle^2}{\langle n \rangle} = 1 - m(1) + \frac{m(2)}{m(1)}$$

for the unconditional distribution of the number of phonons in the cavity as a function of the incoherent pump $\underline{a}^2$ for different parameters ν. One can see from Fig. 4 that, for ν < 0, the Fano factor becomes smaller than unity irrespective of other parameters of the system, and as ν decreases down to the limiting value of –1/2, the radiation squeezing increases and reaches 5.5%. The minimum value of the Fano factor is achieved for the pump intensity approximately equal to 1/2 and shifts to zero by the value ν + 1/2 with increasing decay rate of the atom.

To study the possibility of achieving the maximum amplitude squeezing in a single-atom laser, we calculated the dependences of the Fano factor on the pump intensity for different interaction constants (Fig. 5). The dependences were calculated using exact recurrent relations (3.5) without using the strong correlation approximation. It follows from Fig. 5 that the squeezing increases with η, achieves the maximum value, and then decreases down a certain constant value determined by expressions (6.1) and (6.2). The numerical calculations of the Fano factor show that the maximum squeezing reaches 15% when $2\nu + 1 \ll 1$, $\eta = (3/2)^2$, and $\underline{a}^2 = 1$. This result gives a more exact region of parameters at which the maximum squeezing is achieved compared to [6].

7. CLASSICAL AND QUANTUM CORRELATIONS: THE ENTANGLED STATES OF AN ATOM AND A FIELD

As mentioned above, the interaction between the atom and field upon incoherent pumping of the atom produces a correlated atom–field state. According to (3.2) and (3.4), the combined stationary density matrix corresponding to this state has the form

$$\hat{\rho}_{st} = |1\rangle\langle 1|\rho_1 + |2\rangle\langle 2|\rho_2 + \frac{1}{\sqrt{\eta}}(\sigma_+ a\rho + \rho a^+ \sigma_-), \quad (7.1)$$

where ρ_i is the conditional field density matrices and $\rho = \rho_1 + \rho_2$ is the unconditional density matrix of the field.

In the strong interaction limit ($\eta \gg 1$), which corresponds to the neglect of the last term in (7.1), the produced mixed state exhibits strong atom–field correlations: the ground state of the atom is coupled with the generalized coherent Mittag–Leffler state with the parameter ν (4.4a), while the excited state of the atom is coupled with the same field state but with the parameter $\nu + 1$(4.4b). However, this type of correlations has a classical nature because the density matrix of a correlated atom–field system in this case is a mixture of the factorized states of the atom and field ($|1\rangle|n\rangle$ and $|2\rangle|n\rangle$) with the corresponding probabilities $\rho_1(n)$ and $\rho_2(n)$. This type of correlations does not require for its description the use of a superposition of such factorized atom–field states. This type of correlations can be explained within the framework of the theory of correlation of classical random quantities.

In this section, we consider the question about the possibility of producing nonclassical correlations in the system under study, i.e., correlations that manifest the so-called entangled states. In our case, these are the states that include the superpositions of factorized atom-field states. Unfortunately, the general methods for analyzing properties of entangled systems in mixed states have not been developed so far. For this reason, we consider here only particular cases for which such methods have been developed.

One of these methods—an analytic calculation of the degree of entanglement of the states, the entanglement formation entropy E_F—was proposed by Wouters as applied to the mixed states of two two-level systems. According to this method,

$$E_F = -x\log_2 x - (1-x)\log_2(1-x). \tag{7.2}$$

Here, the quantity

$$x = \frac{1+\sqrt{1-C^2}}{2}$$

is expressed in terms of the so-called concurrency parameter

$$C(\rho) = \max\{0, \lambda_1 - \lambda_2 - \lambda_3 - \lambda_4\}, \tag{7.3}$$

where λ_i are the eigenvalues of the matrix $\rho\tilde{\rho}$ arranged in descending order. The adjoint matrix $\tilde{\rho}$ is calculated by the rule

$$\tilde{\rho} = (\sigma_y \otimes \sigma_y)\rho^*(\sigma_y \otimes \sigma_y),$$

where σ_y are spin Pauli operators for particles with the spin 1/2.

The quantum system studied here can be represented as a system of two spins in the case when one can consider the number of photons that does not exceed unity. In this case, one of the two-level quantum systems is an atom and another system is an intracavity field with the states $|0\rangle$ and $|1\rangle$.

In the strong correlation approximation, retaining the last term in (7.1) and using expression (4.4) for the calculation of the density matrix (7.1), we obtain in the one-photon limit (weak incoherent pumping) the eigenvalues of the $\rho\tilde{\rho}$ matrix:

$$\lambda_1 = \frac{s+s_1}{1+2s}, \quad \lambda_2 = \frac{s-s_1}{1+2s}, \quad \lambda_3 = \lambda_4 = \frac{\sqrt{s_2}}{1+2s}, \tag{7.4}$$

where

$$s = \frac{\underline{a}^2}{\nu+1}, \quad s_1 = \frac{s}{\sqrt{\eta}}, \quad s_2 = \frac{\underline{a}^2 s}{\nu+2}.$$

The constants s_1 and s_2 are much smaller than the constant s, so that the concurrency parameter is always negative. Therefore, the quantity E_F for this system is zero, and the atom–field states are not entangled.

Not using the strong correlation approximation, in the case of excitation of only one-photon state, we can estimate the entanglement of the atomic and field subsystems from the expression

$$C = 2(z_3 - \sqrt{z_4})(1 + z_1 + z_2 + z_3 + z_4) \tag{7.5}$$

for the concurrency parameter, where the parameters

$$z_1 = \frac{\rho_1(1)}{\rho_1(0)} = \frac{\eta\underline{a}^2}{2(\nu+1)(\eta+2\nu+1)},$$

$$z_2 = \frac{\rho_2(0)}{\rho_1(0)} = \frac{(\eta+2\nu+2)z_1}{\eta},$$

$$z_3 = \frac{\rho_2(1)}{\rho_1(0)} = \frac{(\eta+2\nu+4)\underline{a}^2}{2(\nu+2)(\eta+2\nu+2)},$$

$$z_4 = \mathrm{Sp}(a\rho) = \frac{z_1}{\sqrt{\eta}(z_1+z_2+z_3+z_4)}$$

were calculated from recurrent expressions (3.7), taking into account (7.1). According to (7.5), the atomic and field systems will be in an entangled state if

$$\frac{\eta z_3}{z_1} > 0. \tag{7.6}$$

The dependence of C on the strength of the atom–field interaction and the parameter ν within the framework of this approximation is presented in Fig. 6, where condition (7.6) is also shown. One can see from this figure that the concurrency parameter increases with increasing interaction η and achieves its maximum value at $\sqrt{\eta} = 0.4$ and $\nu = -1/2$. Figure 7 shows the dependences of the concurrency parameter C and the degree

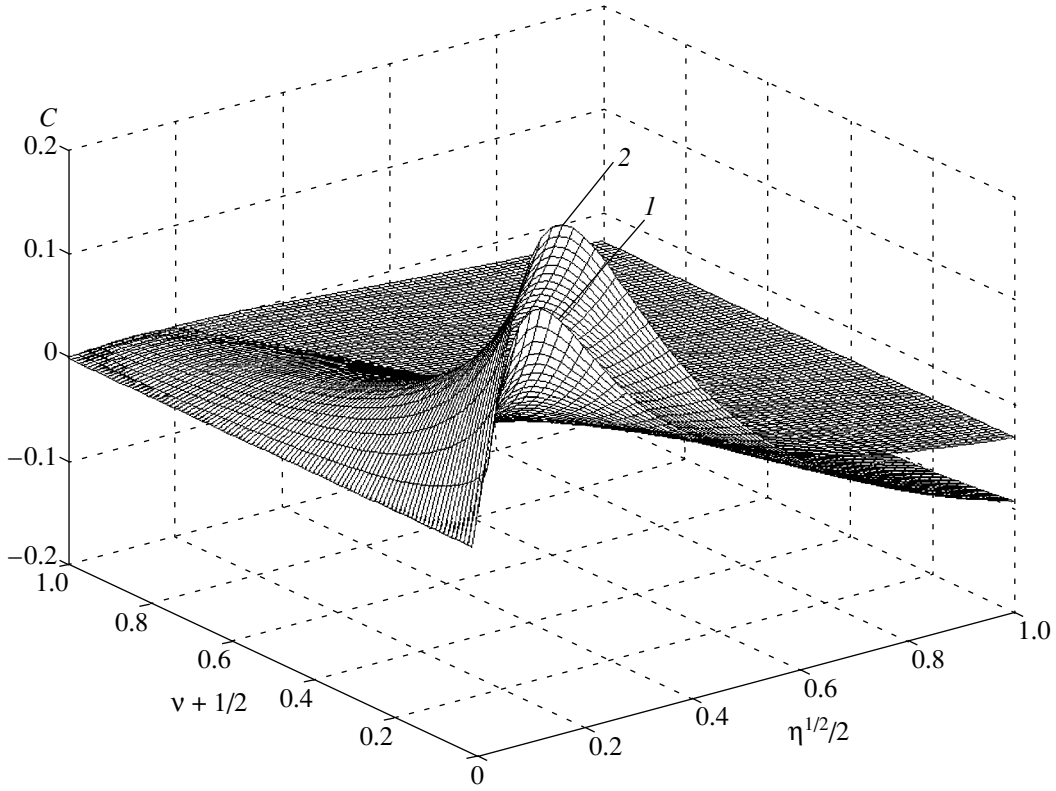


Fig. 6. Dependence of the concurrency parameter C on the parameter ν and the strength of the atom–field interaction η for $\underline{a}^2 = 0.05$ (*1*) and 0.25 (*2*).

of entanglement E_F of the atom–field states on the incoherent pump intensity for optimal values of the parameters $\sqrt{\eta} = 0.4$ and $\nu = -1/2$. One can see that both these dependences achieve the maximum simultaneously at $\underline{a}^2 = 0.33$.

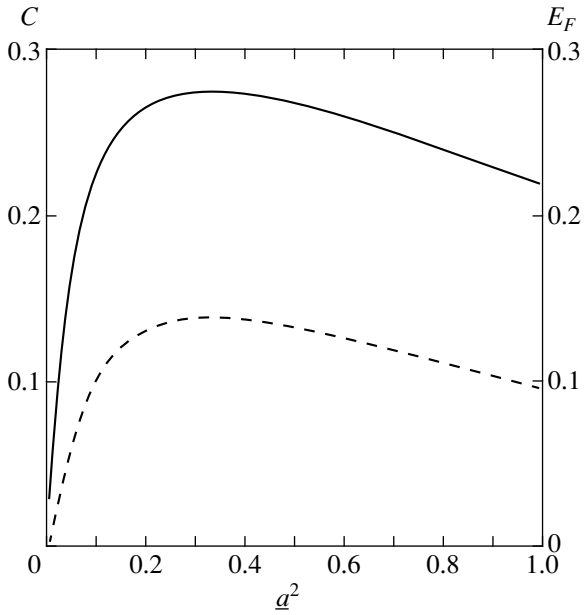


Fig. 7. Dependence of the concurrency parameter C (solid curve) and the degree E_F of entanglement (dashed curve) of the atom–field states on the incoherent pump intensity.

8. CONCLUSIONS

We have studied in detail the state of a field produced by a single-atom laser in the case of a strong coupling between the atom and field of both quantum (related to the entangled atom–field states) and classical (in the absence of these states) nature. Unlike a usual macroscopic laser, where such a coupling is absent and the atomic and field states are independent, the field states in the single-atom laser depend on the level at which the atom is found. In the stationary state, the probability of finding n photons and the atom in the excited state is equal to the probability of finding $n + 1$ photons and the atom in the ground state. We have shown that, in the stationary state, the field states belonging to the ground or excited atomic levels are the generalized Mittag–Leffler states averaged over phase fluctuations and having loss balance indices ν differing by unity. We have found that the conditional Glauber P functions corresponding to these states are expressed in terms of the kernel of the operation of fractional Riemann–Liouville integration (differentiation) of the order equal to the loss index or exceeding it by unity. In the case when this index is zero (atomic losses are equal to field losses), the generalized coherent states belong-

ing to the ground atomic state are transformed to a usual coherent state averaged over phase fluctuations. This demonstrates that emission produced by a single-atom laser upon incoherent pumping of the atom is close to emission produced by a classical current with a stabilized amplitude and a random phase.

The presence of a rigid classical correlation between the atomic and field states also results in inversionless lasing, which differs from inversionless lasing in three-level macroscopic systems. When the loss index is negative (the rate of field losses exceeds the rate of atomic losses), the laser emission is described by a nonclassical sub-Poisson statistics. Another nonclassical feature of the emission produced by a single-atom laser is a boundedness of the *P* function on a segment.

Although the atom–field correlations are classical for most values of the parameters, at low pump powers the entangled atom–field states can be generated with the maximum entanglement formation entropy of about 12%.

Note in conclusion that a strong correlation between an atom and a cavity field mode considered in this paper has been achieved in modern experiments with single atoms in high-Q optical cavities [14, 15]. For example, in experiments [15] with preliminarily cooled Cs atoms (the resonance transition corresponds to one of the components of the 852.4-nm D_2 line; the dipole decay rate constant of the $6S_{1/2}$, $F = 4$, $m_F = 4 \longrightarrow 6P_{3/2}$, $F = 5$, $m_F = 5$ transition is $\gamma_\perp = 2\pi \times 2.6$ MHz) captured by a dipole trap tuned far away from the resonance with the capture time $\tau = 28$ ms in a cavity of length $l = 44.6 \times 10^{-6}$ m and the decay constant $k = 2\pi \times 4$ MHz, the interaction constant was $g = 2\pi \times 32$ MHz, corresponding to a strong correlation regime with the parameter $\eta = 64$. In this case, the frequency of the cavity field mode almost exactly coincided with the atomic-transition frequency and the frequency of the probe low-intensity laser radiation with the average number of photons $\bar{n} \approx 0.1$.

APPENDIX

Properties of Phase-Independent Distribution Functions

The statistical properties of a single-mode field averaged over the atomic states can be conveniently studied using the *s*-ordered quasi-probability functions

$$R^{(s)}(\alpha) = \frac{1}{\pi^2}\int d^2\lambda F_N(\lambda) \times \exp(-\lambda\alpha^* + \lambda^*\alpha - s|\lambda|^2/2), \quad \text{(A.1)}$$

which are defined as a two-dimensional Fourier transform of the normally ordered field characteristic function [16]

$$F_N = \mathrm{Sp}_f(\exp(\lambda a^+)\exp(-\lambda^* a)\rho).$$

Here, the superscript $s = 0$ corresponds to the *P* function of the Glauber distribution, $s = 1$ corresponds to the Wigner function *W*, and $s = 2$ corresponds to the distribution function *Q*. For an incoherently pumped single-atom laser, both the characteristic function and the distribution function prove to be phase-independent functions because of the invariance of the master equation (2.1) with respect to the simultaneous phase shift of the operators related to the atom and field: $a \longrightarrow a\exp(i\phi)$, $\sigma_- \longrightarrow \sigma_-\exp(i\phi)$ [17]. In this case, the distribution functions $R_s(|\alpha|^2)$ are calculated as a one-dimensional Bessel transformation of the characteristic function:

$$R^{(s)}(|\alpha|^2) = \frac{2}{\pi}\int d|\lambda||\lambda| J_0(2|\lambda||\alpha|)F_N(|\lambda|^2)\exp(-s|\lambda|^2/2). \quad \text{(A.2)}$$

In addition, the Wigner function and *Q* function are related to the *P* function by the integral expression

$$R^{(s)}(|\alpha|) = \frac{4}{s}\exp\left(-\frac{2}{s}|\alpha|^2\right) \times \int_0^\infty d|\beta||\beta|\exp\left(-\frac{2}{s}|\beta|^2\right)I_0\left(\frac{4}{s}|\alpha||\beta|\right)P(|\beta|), \quad \text{(A.3)}$$

where $I_0(x)$ is a modified Bessel function.

By using the expansion of the phase-independent characteristic function into a power series over normally ordered moments

$$F_N(|\lambda|^2) = \sum_{n=0}^{\infty}(-1)^n\frac{|\lambda|^{2n}}{(n!)^2}\langle a^{+n}a^n\rangle, \quad \text{(A.4)}$$

we can represent the Wigner and *Q* functions as a sum of Laguerre polynomials:

$$R^{(s)}(|\alpha|^2) = \frac{1}{\pi}\sum_{n=0}^{\infty}\frac{(-1)^n}{s^{n+1}n!}\langle a^{+n}a\rangle \times \exp\left(-\frac{|\alpha|^2}{s}\right)L_n\left(\frac{|\alpha|^2}{s}\right), \quad \text{(A.5)}$$

where Laguerre polynomials are the known orthogonal polynomials defined by the series [18]

$$L_n(x) = \frac{e^x}{n!}\frac{d^n}{dx^n}(x^n e^{-x}) = \sum_{l=0}^{n}(-1)^l\frac{n!}{l!(n-l)!}\frac{x^l}{l!}. \quad \text{(A.6)}$$

For the *P* function ($s = 0$), expansion (A.5) is not defined. The regularized expansion of the *P* function in a series over the Laguerre polynomials was

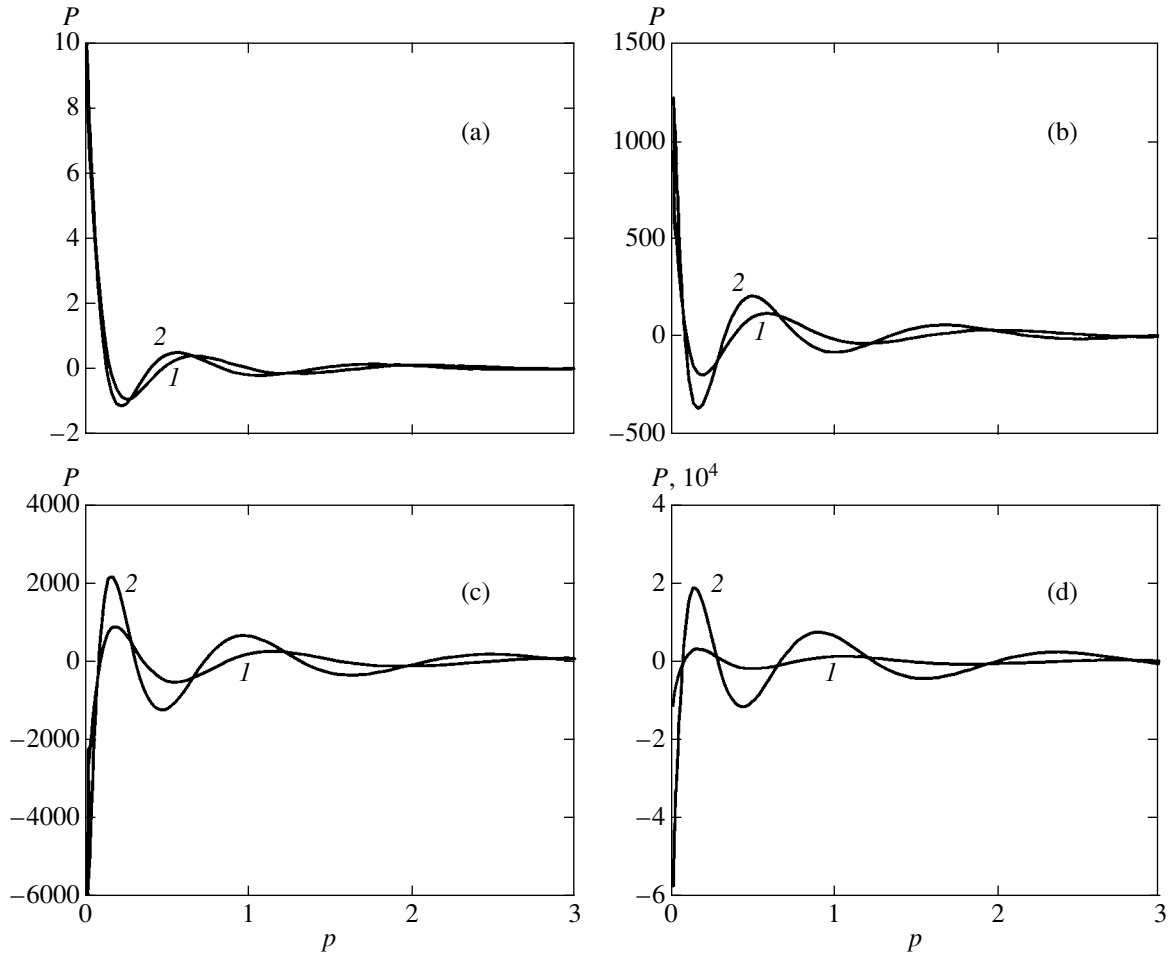


Fig. 8. Approximate distributions of the *P* function for even and odd Fock states. The curves were calculated using Laguerre polynomials L_N with $N = 25$ (*1*) and 30 (*2*); $n_0 = 0$ (a), 2 (b), 3 (c), 5 (d).

obtained in [19]. For the phase-independent distribution, it has the form

$$P(|\alpha|^2) = \frac{1}{\pi}\exp[-(\xi-1)|\alpha|^2] \times \sum_{j=0}^{\infty} L_j(\xi|\alpha|^2)\sum_{s=0}^{\infty}\frac{(-\xi)^s}{s!(j-s)!}\rho(s), \quad \text{(A.7)}$$

where the regularization parameter $\xi \geq 1$ and $\rho(s)$ are the diagonal elements of the density matrix in the Fock basis.

Another method for regularization of the *P* function also exists. Let us make a formal substitution $1 = \exp(-q|\alpha|^2)\exp(q|\alpha|^2)$ in the integrand in expression (A.2) for $s = 0$ and write the expansion

$$e^{q|\lambda|^2}F_N(|\lambda|^2) = \sum_{n=0}^{\infty}\sum_{m=0}^{n}\frac{(-1)^m\langle a^{+m}a^m\rangle}{m!^2(n-m)!}q^{n-m}|\lambda|^{2n} = \sum_{n=0}^{\infty}L_n\left(\frac{\partial^2}{\partial\lambda\partial(-\lambda^*)}\right)F_N\left(\frac{|\lambda|^2}{q}\right)\Big|_{\lambda=0}q^n|\lambda|^{2n}. \quad \text{(A.8)}$$

Assuming that the series (A.8) converges and using the relation

$$\int_0^{\infty} d|\lambda||\lambda|^{2n+1}J_0(2|\lambda||\alpha|)\exp(-q|\lambda|^2) = \frac{n!}{2q^{n+1}}\exp\left(-\frac{|\alpha|^2}{q}\right)L_n\left(\frac{|\alpha|^2}{q}\right), \quad \text{(A.9)}$$

we obtain the required expansion of the *P* function in a series over the Laguerre polynomials and normally ordered moments $\langle a^{+n}a^n\rangle$:

$$P(|\alpha|^2) = \frac{1}{\pi q}\sum_{n=0}^{\infty}\exp\left(-\frac{|\alpha|^2}{q}\right)L_n\left(\frac{|\alpha|^2}{q}\right) \times L_n\left(\frac{\partial^2}{\partial\lambda\partial(-\lambda^*)}\right)F_N\left(\frac{|\alpha|^2}{q}\right)\Big|_{\lambda=0}. \quad \text{(A.10)}$$

An example of a phase-independent nonclassical state of the field is a purely Fock state $\rho_0 = |n_0\rangle\langle n_0|$. The normally ordered moments for this state are

$$\langle a^{+k}a^k\rangle = \frac{\theta(n_0-k)n_0!}{(n_0-k)!}, \tag{A.11}$$

where $\theta(x)$ is the Heaviside function. By substituting (A.11) into expansion (A.10), we obtain the regularized expression for the Glauber function of the Fock state:

$$P_{n_0}(|\alpha|^2) = \frac{1}{\pi q}\sum_{n=0}^{\infty}\exp\left(-\frac{|\alpha|^2}{q}\right)L_n\left(\frac{|\alpha|^2}{q}\right) \times \sum_{k=0}^{n}(-q)^k C_n^k C_{n_0}^k. \tag{A.12}$$

Expression (A.12) for $n_0 = 0$ gives, in particular, the representation of the delta function for the positive semiaxis of the argument values in the form of a series over the Laguerre polynomials:

$$\delta(|\alpha|^2) = \sum_{n=0}^{\infty}\exp(-|\alpha|^2)L_n(|\alpha|^2). \tag{A.13}$$

Figure 8 shows the plots of approximate distributions for even and odd Fock states. For even states, the central peak of the P function is located in the region of positive values. As the order n of the Laguerre polynomials appearing in expansion (A.12) increases, the height of the central peak increases, as well as the heights of other negative and positive minima of the distribution function. As the number n_0 of photons increases, the distribution function becomes more singular, the number and intensity of oscillations increase and their period decreases, and the central peak becomes more intense than for $n_0 = 0$. For odd Fock states, the central peak of the distribution is located in the region of negative values. Its intensity increases with increasing number n_0 of photons, remaining in the region of negative values of the P function. In other aspects, the behavior of the quasi-probability function for an odd number of photons is similar to that for an even number of photons.

Taking into account the expansion of the density matrix in the Fock basis

$$\rho = \sum\rho(n)|n\rangle\langle n|$$

and using expression (A.12), we obtain the relation between the P function and the photon distribution $\rho(n)$:

$$P(|\alpha|^2) = \frac{1}{\pi q}\sum_{n=0}^{\infty}\rho(n)\sum_{m=0}^{\infty}\exp\left(-\frac{|\alpha|^2}{q}\right)L_m\left(\frac{|\alpha|^2}{q}\right)S_m^n, \tag{A.14}$$

$$S_m^n = \sum_{k=0}^{m}(-q)^k C_n^k C_m^k.$$

By varying the parameter q in expression (A.14), we can control the convergence of individual parts of the series. In addition, by summing the series (A.14) over other variables, we can represent it in the form (A.7).

Comparison of (A.10) and (A.14) yields the known relation between moments and the number of photons:

$$\langle a^{+m}a^m\rangle = \sum_{k=m}^{\infty}\frac{\rho(k)k!}{(k-m)!}.$$

The inverse relation

$$\rho(n) = (-1)^n\sum_{m=n}^{\infty}(-1)^m\frac{\langle a^{+m}a^m\rangle}{n!(m-n)!}$$

can be obtained by representing the Q function for the phase-independent states in the form (A.4) as

$$Q(|\alpha|^2) = \frac{1}{\pi}\langle\alpha|\rho|\alpha\rangle = \frac{1}{\pi}\exp(-|\alpha|^2)\sum_{n=0}^{\infty}\frac{\rho(n)}{n!}|\alpha|^{2n}. \tag{A.15}$$

Comparison of the direct and inverse relations between the moments and numbers of photons shows that a series of coefficients $\rho(k)$ always converges, whereas a series of moments $\langle a^{+m}a^m\rangle$ converges when a series of terms $k!\rho(k)$ converges.

ACKNOWLEDGMENTS

This work was supported by a joint grant of the Russian and Belarussian Foundations for Basic Research.

REFERENCES

1. H. J. Kimble, in *Cavity Quantum Electrodynamics*, Ed. by P. R. Berman (Academic, New York, 1994), p. 203; H. Walther, Opt. Spektrosk. **91**, 359 (2001) [Opt. Spectrosc. **91**, 327 (2001)].
2. Yu. Mu and C. M. Savage, Phys. Rev. A **46**, 5944 (1992).
3. M. Löffler, G. M. Meyer, and H. Walther, Phys. Rev. A **55**, 3923 (1997).
4. M. Löffler, G. M. Meyer, and H. Walther, Europhys. Lett. **37**, 317 (1997).
5. A. V. Kozlovskiĭ and A. N. Oraevskiĭ, Zh. Éksp. Teor. Fiz. **115**, 1210 (1999) [JETP **88**, 666 (1999)].
6. H. J. Carmichael, Phys. Rev. A **56**, 5065 (1997); O. A. Kocharovskaya and Ya. I. Khanin, Pis'ma Zh. Éksp. Teor. Fiz. **48**, 581 (1988) [JETP Lett. **48**, 630 (1988)].
7. T. B. Karlovich and S. Ya. Kilin, Opt. Spektrosk. **91**, 375 (2001) [Opt. Spectrosc. **91**, 343 (2001)].
8. J.-M. Sixdeniers, K. A. Penson, and A. I. Solomon, J. Phys. A **32**, 7543 (1999).
9. Yu. A. Brychkov and A. P. Prudnikov, *Integral Transforms of Generalized Functions* (Nauka, Moscow, 1977; Gordon and Breach, New York, 1989).
10. V. I. Man'ko, G. Marmo, E. C. G. Sudarshan, and F. Zaccaria, Phys. Scr. **55**, 128 (1997).

11. P. R. Rice and H. J. Carmichael, Phys. Rev. A **50**, 4318 (1994).
12. F. DeMartini and G. R. Jacobovitz, Phys. Rev. Lett. **60**, 1711 (1988); F. DeMartini, F. Cairo, P. Mataloni, and F. Verezegnassi, Phys. Rev. A **46**, 4220 (1992); G. Björk and Y. Yamamoto, IEEE J. Quantum Electron. **QE-27**, 2386 (1991); Y. Yamamoto, S. Machida, and G. Björk, Opt. Quantum Electron. **24**, S215 (1992); Y. Yamamoto and R. E. Slusher, Phys. Today **46** (6), 66 (1993); X. Wang, R. A. Linke, G. Devlin, and H. Yokoyama, Phys. Rev. A **47**, R2488 (1993).
13. M. Lax and M. Zwanzinger, Phys. Rev. A **7**, 750 (1973); M. O. Scully and M. S. Zubariy, *Quantum Optics* (Cambridge Univ. Press, Cambridge, 1999), p. 336.
14. C. J. Hood, T. W. Lynn, A. C. Doherty, *et al.*, Science **287**, 1447 (2000); P. W. H. Pinkse, T. Fischer, P. Maunz, *et al.*, J. Mod. Opt. **47**, 2769 (2000).
15. J. Ye, D. W. Vernooy, and H. J. Kimble, quant-ph/9908007.
16. S. Ya. Kilin, *Quantum Optics: Fields and Their Detection* (Nauka i Tekhnika, Minsk, 1990), p. 31; V. Bužek and P. L. Knight, Prog. Opt. **34**, 11 (1995).
17. B. Daeubler, H. Risken, and L. Schoendoff, Phys. Rev. A **48**, 3955 (1993).
18. *Handbook of Mathematical Functions*, Ed. by M. Abramowitz and I. A. Stegun (National Bureau of Standards, Washington, 1964; Nauka, Moscow, 1979).
19. J. Perina and L. Mista, Phys. Lett. A **27**, 217 (1968).

Translated by M. Sapozhnikov

Journal of Experimental and Theoretical Physics, Vol. 95, No. 5, 2002, pp. 820–832.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 122, No. 5, 2002, pp. 950–964.

ATOMS, SPECTRA, RADIATION

Self-Diffraction of Light Waves by a Nonlocal Photorefractive Grating in a Crystal with the $\bar{4}3m$ Symmetry

R. V. Litvinov

Tomsk State University of Control Systems and Radioelectronics, Tomsk, 634050 Russia
e-mail: litvinov@ed.rk.tusur.ru

Received February 5, 2002

Abstract—A frequency-degenerate steady-state two-wave interaction on a dynamic transmitting phase grating formed in a cubic crystal of the $\bar{4}3m$ symmetry group with a nonlocal photorefractive response is considered in the paraxial approximation. The conservation laws for the nonlinear system of equations of coupled waves, derived for an arbitrary orientation of interaction relative to the crystallographic axes and the polarization of incident light waves, indicate that the contribution to energy exchange between the interacting waves may come from fluxes in different directions. The possibility of nonunidirectional energy pumping from one wave to the other upon a change in their polarization state due to the interaction is demonstrated. For the transverse configuration of the interaction and linear polarization of incident waves, explicit analytic expressions for the scalar amplitudes of the orthogonal components of the light field are derived in the linear approximation in the coefficient of modulation of the interference pattern of light. The possibility of rotation of the polarization planes of light waves without a change in their intensity is demonstrated. For three particular configurations, the dependence of the efficiency of interaction of linearly polarized waves on the reduced length, orientation of the polarization vectors of the incident light waves, and the ratio of their intensities are analyzed. © *2002 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

Dynamic scattering of light by phase inhomogeneities formed due to local cubic nonlinearity for a high power of radiation (of the order of a kilowatt per square centimeter) is observed in many media [1–4]. For low radiation powers (of the order of a milliwatt per square centimeter), self-action of light in photorefractive noncentrosymmetric crystals is clearly manifested. In the general case, however, the photorefractive response in such crystals cannot be described by a simple cubic nonlinearity. Perturbations of dielectric properties of the medium are induced in this case through the linear electro-optical effect by the photoinduced field of the space charge. The relation between this field and the light intensity can be described by a system of nonlinear constitutive equations whose structure is determined by the collection of processes leading to the spatial separation of the electric charge [4–11]. In contrast to media with local nonlinearity mechanisms, photorefractive crystals can exhibit a nonlocal response in the case of strict frequency degeneracy of two light beams exchanging energy in this case [4–6, 9–27].

The two-wave interaction on a dynamic phase grating was considered in [1, 2, 4–6, 9–11] in the approximation with a preset polarization of light waves. This can be done for anisotropic photorefractive crystals if the polarization of the waves incident on a crystal coincides with the polarization of one of the intrinsic optical modes of the medium. In this case, the distance Δk between the wave surfaces of these modes in the $\mathbf{k}$ space must be sufficient for disregarding the intermode interaction. For example, in the ferroelectric crystals $LiNbO_3$, $LiTaO_3$, and $BaTiO_3$, the value of Δk is on the order of 10^4 cm^{-1}, which is considerably larger than the two-wave amplification coefficients $\Gamma \sim 10^2$ cm^{-1} attained in such crystal. Hence, the polarization of light waves in a crystal does not change as a result of interaction if the incident waves are polarized along the ordinary and extraordinary axes. The theory of frequency-degenerate scalar two-wave interaction implies that the steady-state energy exchange on a nonlocal phase grating is unidirectional and that the steady-state energy exchange on a local phase grating is prohibited; this was also pointed out in an earlier publication [28].

An analysis of two-wave interaction in cubic photorefractive crystals $Bi_{12}GeO_{20}$, $Bi_{12}SiO_{20}$, and $Bi_{12}TiO_{20}$ proved that the change in the polarization state of light waves due to gyrotropy and linear birefringence induced by an external field influences considerably the effectiveness of energy exchange [12–16, 18–20]. Moreover, the additional transformation of the polarization structure of the light field directly due to self-diffraction may change the direction of heat exchange upon an increase in the interaction length or the dependence of this direction on the intensity ratio of the incident waves [21–27]. Such a transformation removes the prohibition on the energy exchange for two-wave interaction on a local photorefractive grating. The reasons of change in the polarization of light waves directly due to self-diffraction are the different efficiencies of two

Components H_{EE}, $H_{EM} = H_{ME}$, and H_{MM} of the coupling matrix; its eigenvalues $p_{y,z}$; and angle φ between the intrinsic y' axis and the coordinate axis y for various configurations of the two-wave interaction in cubic crystals of symmetry groups $\bar{4}3m$ and 23

Configuration	H_{EE}	$H_{EM} = H_{ME}$	H_{MM}	Eigenvalues	φ
longitudinal $\mathbf{x}^o \uparrow\uparrow [110]$, $\mathbf{z}^o \uparrow\uparrow [001]$	–1	0	0	$p_y = -1$, $p_z = 0$	0
transverse $\mathbf{x}^o \uparrow\uparrow [110]$, $\mathbf{z}^o \uparrow\uparrow [1\bar{1}0]$	0	–1	0	$p_{y,z} = \pm 1$	–45°
diagonal $\mathbf{x}^o \uparrow\uparrow [110]$, $\mathbf{z}^o \uparrow\uparrow [\bar{1}11]$	$\frac{1}{\sqrt{3}}$	0	$-\frac{2}{\sqrt{3}}$	$p_y = \frac{1}{\sqrt{3}}$, $p_z = -2p_y$	0
$\mathbf{x}^o \uparrow\uparrow [110]$, $\mathbf{z}^o \uparrow\uparrow [\bar{1}1\bar{2}]$	0	$-\frac{1}{\sqrt{3}}$	$\sqrt{\frac{2}{3}}$	$p_{y,z} = \frac{1}{\sqrt{6}} \pm \frac{1}{\sqrt{2}}$	–62.6°

Note: Arrows indicate the orientation of the coordinate unit vector along the positive direction of the corresponding crystallographic axis.

intramode processes and the strong intermode interaction. The latter is due to a small value of Δk which does not exceed 10 cm^{-1}. The decrease in the effectiveness of intermode interaction upon an increase in the interaction length due to Bragg's detuning equal to Δk is compensated by an increase in the coupling constant γ attaining values of the order of Δk and higher due to the application of nonstationary mechanisms of formation of photorefractive holograms [29, 30].

The natural circular birefringence and the linear birefringence induced by an external electric field complicate to a considerable extent the qualitative pattern of the vectorial two-wave interaction in crystals of the 23 symmetry, especially for comparable intensities of light waves, for which the approximation of a preset field of a high-intensity wave is inapplicable. Self-diffraction in nongyrotropic cubic crystals of the $\bar{4}3m$ symmetry (such as GaAs, InP, or CdTe) in zero external electric field is more obvious; in this case, the polarization state of light waves may change only due to the interaction. In the particular case of the longitudinal configuration of the interaction in crystals with such a symmetry group (see table), the exact solution of the problem of vectorial self-diffraction of plane monochromatic light waves by a local photorefractive grating with an amplitude proportional to the modulation coefficient of the light interference pattern was obtained in [22].

Here, we consider the steady-state vectorial interaction of two monochromatic light waves by a transmitting nonlocal photorefractive grating in a cubic crystal of the $\bar{4}3m$ symmetry of an arbitrary orientation. Main attention is paid to an analysis of the features of energy exchange between the waves, which are associated with their polarization due to self-diffraction.

2. MODEL

Figure 1 shows a schematic diagram of symmetric interaction of two cocurrent light waves, viz., a signal wave

$$\tilde{\mathbf{S}} = \mathbf{S}\exp[i(\omega t - \mathbf{k}_S \cdot \mathbf{r})]$$

and a reference wave

$$\tilde{\mathbf{R}} = \mathbf{R}\exp[i(\omega t - \mathbf{k}_R \cdot \mathbf{r})],$$

in a cubic nonmagnetic photorefractive crystal of symmetry group $\bar{4}3m$. Permittivity perturbations $\Delta\varepsilon_{ij}$ responsible for the interaction are induced in the crystal by the field $\mathbf{E}_{sc}$ of the space charge through the linear electro-optical effect. In the case under investigation, field $\mathbf{E}_{sc}$ is formed due to spatial separation of charge carriers excited by the light field of intensity

$$I = I_0\left\{1 + \left[\frac{m}{2}\exp(iKz) + \text{c.c.}\right]\right\} \propto \left|\tilde{\mathbf{S}} + \tilde{\mathbf{R}}\right|^2.$$

This field exhibits a rapidly oscillating dependence on the transverse coordinate z along the vector

$$\mathbf{K} = \mathbf{k}_R - \mathbf{k}_S \quad (|\mathbf{K}| = K)$$

and a slowly varying dependence on the longitudinal coordinate x (interaction length) along the normal to the input face of the crystal. This enables us to set $\mathbf{E}_{sc} = E_{sc}\mathbf{z}^o$, where $\mathbf{z}^o = \mathbf{K}/K$, and represent the permittivity perturbations in the form

$$\Delta\varepsilon_{ij} = n^4 r_{41} E_{sc} |\delta_{ijk}| z_k^0 = n^4 r_{41} E_{sc} g_{ij}, \qquad (1)$$

where n is the refractive index, $r_{41} = r_{52} = r_{63}$ is the only nonzero independent electro-optical coefficient for cubic noncentrosymmetric crystals with symmetry groups $\bar{4}3m$ and 23, δ_{ijk} is the Levi-Civita symbol, and z_k^o are the components of the unit vector $\mathbf{z}^o$ in the crystal-physics system of coordinates.

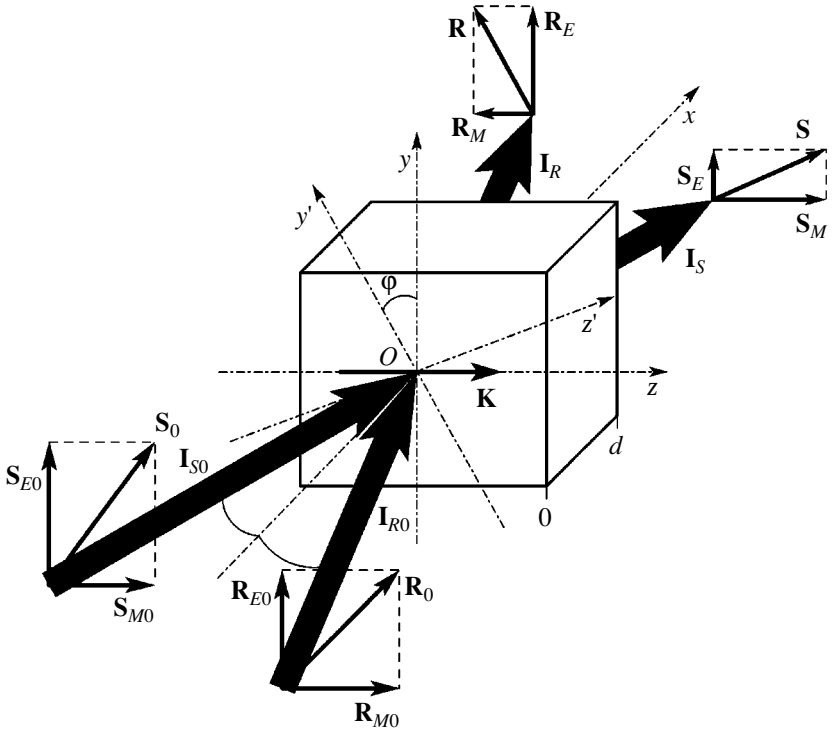


Fig. 1. Symmetric two-wave interaction in a photorefractive crystal.

In the framework of the band theory of solids, the field E_{sc} of a space charge is connected with the luminous intensity I through a system of nonlinear differential (constitutive) equations whose structure is determined by the type and amount of photoactive impurities and traps in the crystal as well as by the mechanisms of transfer of photoexcited charge carriers [4–11]. For large modulation coefficients $m \approx 1$, the contribution of higher harmonics to the spatial dependence [17, 31–39]

$$E_{sc} = \sum_{s=-\infty}^{\infty} \frac{E_s}{2} \exp(isKz)$$

becomes significant. However, the effect of higher diffraction orders on the effectiveness of the interaction between the main beams in thick crystals is weak [1, 34, 38]. In this case, the nonlinearity of processes leading to spatial separation of electric charges affects the nonlinear dependence of the amplitude of the fundamental harmonic on the modulation coefficient, which can be represented in the steady-state conditions in the form

$$E_1 = E_{\text{eff}} \sum_{n=1}^{\infty} a_n m^n = E_{\text{eff}} f(m).$$

Here, the effective amplitude E_{eff} and function $f(m)$ are determined by the band model parameters of the crystal and external conditions [17, 31, 32, 35, 39]. In the general case, amplitude $E_1 = E_1' - iE_1''$ is a complex quantity, and the photorefractive grating contains the local (proportional to E_1') as well as a nonlocal (E_1'') component. A local grating is formed during charge separation due to the drift in a strong constant electric field or due to the linear photogalvanic effect [4–11, 31–33]. A nonlocal grating is formed due to the diffusion mechanism of charge separation or due to the drift in an external periodic electric field with period T satisfying the condition $\tau_R \ll T \ll \tau_{di}$, where τ_R and τ_{di} are the recombination and dielectric relaxation times, respectively [7, 9, 29, 35–39]. A nonlocal photorefractive response can also be realized due to the circular photogalvanic effect [10]. The amplitudes of the local and nonlocal components of the fundamental harmonic of the space-charge field may attain several tens of kilovolts per centimeter for large modulation coefficients $m \approx 1$. In the case of a nonlocal response of the crystal, the two-wave amplification coefficient in the intensity of a weak light wave has a value of the order of 10 cm^{-1} and higher [4, 9, 15–25, 27, 29, 37, 38].

We confine our analysis to the steady-state two-wave interaction in the case of a nonlocal photorefractive response of the crystal, when the effective amplitude is an imaginary quantity, $E_{\text{eff}} = -iE''$, and function $f(m)$ has real values for a real-valued argument. In this case, the permittivity perturbation tensor $\Delta\boldsymbol{\varepsilon}$ can be represented in the form

$$\Delta\boldsymbol{\varepsilon} = \left[-\frac{in^4 r_{41} E''}{2} f(m)\exp(iKz) + \text{c.c.}\right]\mathbf{g}, \qquad (2)$$

where the components of tensor $\mathbf{g}$ are defined by the relation $g_{ij} = |\delta_{ijk}| z_k^o$ (see formula (1)). We will use the paraxial approximation, which is valid for small angles between the wave vectors $\mathbf{k}_{S,R}$ and the x axis, assuming that the vector amplitudes of the light waves have only the y and z components (TE and TM, respectively):

$$\mathbf{S} = S_E \mathbf{y}^o + S_M \mathbf{z}^o, \quad \mathbf{R} = R_E \mathbf{y}^o + R_M \mathbf{z}^o$$

(see Fig. 1). In these approximations, the equations for coupled waves describing slow variations of the TE and TM components of the light field over the interaction length can easily be derived from the wave equation for the light field $\tilde{\mathbf{E}} = \tilde{\mathbf{S}} + \tilde{\mathbf{R}}$ in the form

$$\frac{dS_E}{dx} = \frac{\gamma}{4} f(m)(H_{EE}R_E + H_{EM}R_M), \qquad (3)$$

$$\frac{dS_M}{dx} = \frac{\gamma}{4} f(m)(H_{ME}R_E + H_{MM}R_M), \qquad (4)$$

$$\frac{dR_E}{dx} = -\frac{\gamma}{4} f^*(m)(H_{EE}S_E + H_{EM}S_M), \qquad (5)$$

$$\frac{dR_M}{dx} = -\frac{\gamma}{4} f^*(m)(H_{ME}S_E + H_{MM}S_M). \qquad (6)$$

Here,

$$m = 2\frac{S_E R_E^* + S_M R_M^*}{I_0} \qquad (7)$$

is the modulation coefficient of the interference pattern, which is expressed in terms of the amplitudes of the TE and TM components of the light waves;

$$\gamma = \frac{2\pi n^3 r_{41} E''}{\lambda}$$

is the coupling constant; and

$$H_{MM} = \mathbf{z}^o \cdot \mathbf{g} \cdot \mathbf{z}^o, \quad H_{EE} = \mathbf{y}^o \cdot \mathbf{g} \cdot \mathbf{y}^o,$$

$$H_{EM} = H_{ME} = \mathbf{y}^o \cdot \mathbf{g} \cdot \mathbf{z}^o;$$

$$I_0 = |S_E|^2 + |S_M|^2 + |R_E|^2 + |R_M|^2.$$

In Eqs. (3)–(6), we have omitted the terms proportional to the linear absorption coefficient α, which does not affect the polarization state of interacting waves. The inclusion of absorption only leads to an identical (equal to $\exp(-\alpha x/2)$) decrease in the TE and TM components of the light field.

Equations (3)–(7) describe the variation of the amplitudes of the TE and TM components of the light field for a steady-state two-wave interaction in cubic crystals of symmetry $\bar{4}3m$ with an arbitrary orientation in zero external field. In the linear approximation in the modulation coefficient ($f(m) \equiv m$), these equations coincide with the equations for coupled waves used in [18, 19, 22, 24] in an analysis of the interaction in gyrotropic crystals of symmetry 23, if we omit in the latter equations the terms describing the natural circular birefringence as well as the linear birefringence induced by the external field.

3. FIRST INTEGRALS OF EQUATIONS FOR COUPLED WAVES

The system of equations (3)–(6) for coupled waves obviously has the integral

$$I_0 = |S_E|^2 + |S_M|^2 + |R_E|^2 + |R_M|^2 = \text{const},$$

corresponding to the energy conservation law for the light field in the case of its redistribution between two interacting waves in a nonabsorbing photorefractive crystal.

We can derive a conservation law that makes it possible to find first integrals that will be used below for constructing the general solution to the nonlinear system of equations (3)–(7). This system can be used to express the spatial derivative of the modulation coefficient in the form

$$\frac{dm}{dx} = f(m)\frac{\gamma}{2I_0}[H_{EE}(|R_E|^2 - |S_E|^2) + H_{MM}(|R_M|^2 - |S_M|^2) + 2H_{EM}\text{Re}(R_E^* R_M - S_E^* S_M)]. \qquad (8)$$

The light interference pattern formed at the boundary $x = 0$ by incident light waves with different elliptic polarizations is displaced relative to the interference pattern formed by linearly polarized incident waves ($S_{E0,M0} = S_{E0,M0}^*$ and $R_{E0,M0} = R_{E0,M0}^*$). In the case when the origin on the z axis coincides with one of the maxima of the interference pattern formed in the case of linear polarization of incident waves, the magnitude of this displacement is $\Delta z = -\arg(m_0)/K$. Assuming that the problem is unbounded along the transverse coordinate z, we can consider that the origin on the z axis always coincides with one of the maxima of the interference pattern at the boundary $x = 0$ and treat the quantity m_0 as real-valued ($\text{Im}(m_0) = 0$) for an arbitrary polarization of incident light waves. Then, in view of real-valuedness of the function $f(m_0)$ and coefficients H_{MM}, H_{EE}, and H_{EM}, Eq. (8) leads to the conservation law

$$\text{Im}m = \text{Im}(S_E R_E^* + S_M R_M^*) = 0, \qquad (9)$$

indicating the absence of bending for the interference fringes along the interaction length.

For an arbitrary orientation of the interaction relative to the crystallophysical axes, all four elements of the so-called coupling matrix [18, 19, 22]

$$\mathbf{H} = \begin{pmatrix} H_{MM} & H_{ME} \\ H_{ME} & H_{EE} \end{pmatrix} \tag{10}$$

may differ from zero. These elements are given in the table for four particular cases. It follows from Eqs. (3)–(6) that the orientation of the polarization vectors of light waves in the crystal does not change in the case of longitudinal configuration and the TE polarization of incident waves ($S_E|_{x=0} = S_{E0} \neq 0$, $R_E|_{x=0} = R_{E0} \neq 0$, $S_{M0} = R_{M0} = 0$). In the approximation linear in the modulation coefficient ($f(m) \equiv m$), the energy exchange between light waves in this configuration is described by the well-known relations in the scalar theory [1, 6, 11]. The solution to Eqs. (3) and (5) for the same configuration, but for the linear polarization of incident waves, which is oriented arbitrarily in the *yz* plane, was derived in the same approximation in [22]. The solution to the complete system of equations (3)–(7) for waves can also be obtained in the case of an arbitrary orientation of the interaction relative to crystallophysical axes. Taking into account the conservation law (9), we can show that the equations for four characteristics of the equation in partial derivatives, which determines first integrals of Eqs. (3)–(6), coincide in structure with these equations if we omit the factor $\gamma f(m)/4 = \gamma f(m)^*/4$. Relations for these characteristics are linear functions of four arbitrary initial conditions and can be solved relative to any three initial conditions. Inverse relations expressing the initial conditions as functions of the characteristics can be used as the first three independent integrals of the dynamic system (3)–(7) (see, for example, [40]). The general method of determining its independent integrals is inconvenient in view of the cumbersome analytic calculations used. However, a preliminary analysis based on this method predicts the integrability of such a dynamic system in the general case.

In order to find a compact form of the first integrals, we pass from the initially introduced coordinate axes *y* and *z* to new axes; it is convenient to use for such axes the intrinsic axes *y*' and *z*' of matrix **H** (see Fig. 1). The angle φ between the axes *y*' and *y* is given in the table for particular cases of the interaction. In the new system of coordinates, the equations for the *y* and *z* components of the vector amplitudes

$$\mathbf{S} = S_y \mathbf{y}^{o\prime} + S_z \mathbf{z}^{o\prime}, \quad \mathbf{R} = R_y \mathbf{y}^{o\prime} + R_z \mathbf{z}^{o\prime}$$

of light waves interacting on a nonlocal photorefractive grating form the following system of equations:

$$\frac{dS_{y,z}}{dx} = \frac{\gamma}{4} f\left(2\frac{S_y R_y^* + S_z R_z^*}{I_0}\right) p_{y,z} R_{y,z}, \tag{11}$$

$$\frac{dR_{y,z}}{dx} = -\frac{\gamma}{4} f\left(2\frac{S_y^* R_y + S_z^* R_z}{I_0}\right) p_{y,z} S_{y,z}. \tag{12}$$

Here, we have taken into account the fact that the modulation coefficient can be expressed in terms of the new variables as

$$m = \frac{2(S_y R_y^* + S_z R_z^*)}{I_0}; \quad p_{y,z} = H_\Sigma \pm D$$

are the eigenvalues of matrix **H** whose values are given in the table for particular cases of the interaction; and

$$H_\Sigma = \frac{H_{MM} + H_{EE}}{2}, \quad D = (H_\Delta^2 + H_{ME}^2)^{1/2},$$

$$H_\Delta = \frac{H_{MM} - H_{EE}}{2}.$$

Considering that the conservation law (9) in the new system of coordinates assumes the form

$$\mathrm{Im} m = \mathrm{Im}(S_y R_y^* + S_z R_z^*) = 0,$$

we can derive from Eqs. (11) and (12) for the *y* and *z* components the following integrals:

$$\begin{aligned} I_y' &= S_y'^2 + R_y'^2, \quad I_z' = S_z'^2 + R_z'^2, \\ I_y'' &= S_y''^2 + R_y''^2, \quad I_z'' = S_z'^2 + R_z''^2. \end{aligned} \tag{13}$$

Here, the superscripts denote the real and imaginary components of the complex amplitudes of the TE and TM components:

$$S_{y,z} = S_{y,z}' + iS_{y,z}'', \quad R_{y,z} = R_{y,z}' + iR_{y,z}''.$$

Integrals I_y', I_z' and I_y'', I_z'' describe the conservation of the overall energy of the light field,

$$I_y' + I_z' + I_y'' + I_z'' = I_0,$$

during its redistribution between the light waves in the course of self-diffraction, as well as the conservation of parts of this energy,

$$I_y = I_y' + I_y'' = |S_y|^2 + |R_y|^2,$$

$$I_z = I_z' + I_z'' = |S_z|^2 + |R_z|^2,$$

concentrated in the orthogonal polarization components whose orientation coincides with the orientation of the eigenvectors of the coupling matrix. In addition, in the case of interaction of elliptically polarized waves, the energy of components turned through angle π/2 in phase relative to each other is also conserved:

$$I' = I_y' + I_z', \quad I'' = I_y'' + I_z''.$$

The existence of the laws of conservation of I_y and I_z as well as I' and I'' may modify significantly the qualitative

pattern of energy exchange between waves in the case of vectorial self-diffraction as compared to the case of scalar self-diffraction with the conservation law

$$|S|^2 + |R|^2 = I_0$$

[1, 4–6, 9–11]. In the latter case, the interaction leads only to unidirectional pumping of energy from one wave to the other. On the other hand, in the case of vectorial self-diffraction, the contribution to energy exchange may come from two fluxes in different directions. For example, the energy exchange between the z components of the light field corresponding to integral I_z may amplify the signal wave and suppress the reference wave, while the energy exchange between the y components corresponding to integral I_y may suppress the signal wave and amplify the reference wave. Such a situation may be realized for the orientation of interaction relative to the axes of the crystal, which corresponds to the eigenvalues p_y and p_z of the coupling matrix with opposite signs (see table). This follows (e.g., for the boundary condition $S_{y0} = S_{z0}$ and $R_{y0} = R_{z0}$) from the additional integrals $\Im'_{S,R}$ and $\Im''_{S,R}$ that can be obtained from Eqs. (11) and (12), respectively, in the form

$$\Im'_S = p_z \operatorname{sgn} R'_{y0} \arcsin\frac{S'_y}{\sqrt{I'_y}} - p_y \operatorname{sgn} R'_{z0} \arcsin\frac{S'_z}{\sqrt{I'_z}}, \quad (14)$$

$$\Im'_R = p_z \operatorname{sgn} S'_{y0} \arcsin\frac{R'_y}{\sqrt{I'_y}} - p_y \operatorname{sgn} S'_{z0} \arcsin\frac{R'_z}{\sqrt{I'_z}}. \quad (15)$$

If we replace the prime in the last two expressions by a double prime, we obtain expressions for integrals $\Im''_{S,R}$.

It should be noted that the latter integrals are directly connected with transformation of the polarization state of light waves upon self-diffraction. Another supplementary set of integrals $\Im_{Sy,Sz}$ and $\Im_{Ry,Rz}$, which is also associated with such a transformation, can be obtained in the form

$$\Im_{Sy} = \operatorname{sgn} R'_{y0} \arcsin\frac{S'_y}{\sqrt{I'_y}} - \operatorname{sgn} R''_{y0} \arcsin\frac{S''_y}{\sqrt{I''_y}}, \quad (16)$$

$$\Im_{Sz} = \operatorname{sgn} R'_{z0} \arcsin\frac{S'_z}{\sqrt{I'_z}} - \operatorname{sgn} R''_{z0} \arcsin\frac{S''_z}{\sqrt{I''_z}}. \quad (17)$$

If we carry out the simultaneous substitution $S \longrightarrow R$ and $R \longrightarrow S$ in these expressions, we obtain expressions for integrals $\Im_{Ry,Rz}$.

The real-valuedness of coefficients H_{MM}, H_{EE}, and H_{EM} indicates that the polarization of light waves in the crystal remains linear ($S_{E,M} = S^*_{E,M}$, $R_{E,M} = R^*_{E,M}$) in the case of their linear polarization at the boundary $x = 0$, changing only its orientation in the general case. This feature is also characteristic of the two-wave interaction in cubic gyrotropic crystals of the 23 symmetry in zero external field [24, 26]. The transformation of the linear polarization of incident light waves into the elliptical polarization in photorefractive cubic crystals becomes possible in the presence of an external electric field, when linear birefringence is induced in a crystal [18–21, 23, 25, 27].

It should be noted that the set of integrals (13)–(17) is not independent, but any seven integrals from this set are independent. This makes it possible to uncouple Eqs. (11) and (12) and obtain a solution for the real and imaginary parts of each scalar amplitude in quadratures, which can be reduced, in some particular cases, either to transcendental equations or to explicit analytic expressions for these components. For example, the solutions in quadratures for the y and z components of the amplitudes of light waves interacting in a crystal in the approximation linear in the modulation coefficient,

$$f\left[\frac{2(S_y R^*_y + S_z R^*_z)}{I_0}\right] \equiv \frac{2(S_y R^*_y + S_z R^*_z)}{I_0},$$

in the case of incidence of linearly polarized waves ($S''_{y,z} = R''_{y,z} = 0$) can be obtained in the form

$$\int_{S'_{y0,z0}}^{S'_{y,z}} \left\{ \xi(I'_{y,z} - \xi^2) + \frac{I'_{z,y}}{2}\sqrt{I'_{y,z} - \xi^2} \times \sin\left[2\left(\frac{p_{z,y}}{p_{y,z}} \arcsin\frac{\xi}{\sqrt{I'_{y,z}}} \mp \operatorname{sgn}(R'_{y0,z0}) \frac{\Im'_S}{p_{y,z}} \right) \right] \right\}^{-1} d\xi = \frac{p_{y,z}\gamma x}{2I_0}, \quad (18)$$

$$\int_{R'_{y0,z0}}^{R'_{y,z}} \left\{ \xi(I'_{y,z} - \xi^2) + \frac{I'_{z,y}}{2}\sqrt{I'_{y,z} - \xi^2} \times \sin\left[2\left(\frac{p_{z,y}}{p_{y,z}} \arcsin\frac{\xi}{\sqrt{I'_{y,z}}} \mp \operatorname{sgn}(S'_{y0,z0}) \frac{\Im'_R}{p_{y,z}} \right) \right] \right\}^{-1} d\xi = -\frac{p_{y,z}\gamma x}{2I_0}. \quad (19)$$

The solution to Eqs. (11) and (12) in this case can be written in a different form if we introduce the auxiliary functions $\phi_{y,z}$ [22] satisfying the expression $\Im_\phi = p_y\phi_z - p_z\phi_y$ using the relations

$$S'_{y,z} = (I'_{y,z})^{1/2}\cos\phi_{y,z}, \quad R'_{y,z} = (I'_{y,z})^{1/2}\sin\phi_{y,z}.$$

In some particular cases of two-wave interaction, we can rationalize the integrands in formulas (18) and (19) and obtain an explicit solution for the scalar amplitude. The simplest of such cases is the linear configuration, when $p_z = 0$ and $p_y = -1$ [22]. We will derive below the expressions for amplitudes S'_y, S'_z, R'_y, and R'_z in the transverse configuration. In view of relation $p_z = -2p_y$, the integrands can be rationalized in the diagonal configuration also, and their integration in the case of an arbitrary orientation of the polarization vectors of incident waves and an arbitrary relation between their intensities leads to transcendental equations in amplitudes S'_y, S'_z, R'_y, and R'_z. In the general case, including the last configuration from the table, numerical integration must be used for obtaining the distributions of these amplitudes over the interaction length or their dependence at the output of the crystal on the parameters of the waves incident on it (e.g., the orientation of the polarization vectors or the ratio of their intensities).

4. SELF-DIFFRACTION IN THE CASE OF TRANSVERSE CONFIGURATION AND LINEAR POLARIZATION OF INTERACTING WAVES

In this configuration, in view of relation $p_{y,z} = \pm 1$, integral $\Im'_S$ (or $\Im'_R$) can be transformed as

$$I_3 = S'_y R'_z + S'_z R'_y.$$

Another supplementary integral, which is convenient to use instead of integral I'_y or I'_z, has the form

$$I_4 = R'_y R'_z - S'_y S'_z.$$

Rationalization of the integrands in formulas (18) and (19) makes it possible to obtain the amplitudes S'_y, S'_z, R'_y, and R'_z in the form

$$\begin{aligned} S'_{y,z} &= \operatorname{sgn}(S'_{y0,z0})\sqrt{I'_{y,z}}\frac{1}{\sqrt{1+\beta^2_{y,z}(x)}}, \\ R'_{y,z} &= \operatorname{sgn}(S'_{y0,z0})\sqrt{I'_{y,z}}\frac{\beta_{y,z}(x)}{\sqrt{1+\beta^2_{y,z}(x)}}. \end{aligned} \tag{20}$$

Here, the functions $\beta_y(x)$ and $\beta_z(x)$ for $I_3 I_4 \neq 0$ are connected through the relation

$$\frac{\beta_y + \beta_z}{\beta_y \beta_z - 1} = \frac{I_3}{I_4}$$

and have the form

$$\begin{aligned} \beta_{y,z}(x) = \tanh\Bigg[&\pm\frac{\gamma x}{2}\frac{I_3 I_4}{I_0 I'_{y,z}}\sqrt{1+\frac{b^2_{y,z}}{4}} \\ &+ \operatorname{arctanh}\frac{\beta_{y0,z0} + b_{y,z}/2}{\sqrt{1+b^2_{y,z}/4}}\Bigg]\sqrt{1+\frac{b^2_{y,z}}{4}} - \frac{b_{y,z}}{2}, \end{aligned} \tag{21}$$

where

$$b_{y,z} = \frac{I'^2_{y,z} + I_3^2 - I_4^2}{I_3 I_4}, \quad \beta_{y0,z0} = \frac{R'_{y0,z0}}{S'_{y0,z0}}.$$

Formulas (20) and (21) imply that self-diffraction of light waves by a photorefractive grating in the transverse configuration attains saturation upon an increase in the reduced length γx when the amplitudes S'_y, S'_z, R'_y, and R'_z become independent of γx, in contrast to the longitudinal configuration, for which distributions of the TE and TM components oscillating in γx are also possible [22]. In the general case, the distributions of amplitudes S'_y, S'_z, and R'_y, R'_z and the intensities of light waves $I_S = S'^2_y + S'^2_z$ and $I_R = R'^2_y + R'^2_z$ over γx are nonmonotonic. The intensity and polarization of light waves in saturation and the reduced length typical of this regime are determined by the relation between the intensities I_{R0} and I_{S0} of the incident waves and by the orientation of their polarization vectors.

It follows from the arguments put forth in the previous section that the difference in the signs of the eigenvalues $p_{y,z} = \pm 1$ of the coupling matrix indicates the possibility of the existence of two fluxes in different directions in the total energy exchange between the light waves. As for photorefractive crystals of the 23 symmetry, the existence of energy fluxes in different directions for the two-wave interaction in a crystal of symmetry $\bar{4}3m$ makes it possible to control the direction of energy pumping not only by rotating the polarization plane of the waves incident on the crystal [4, 9, 10, 13–20], but also by changing the ratio of their intensities [21–27].

The intensities of light waves do not change upon self-diffraction if energy fluxes in different directions compensate each other. This condition holds in the case of incidence on a crystal of light waves with the TE or TM polarization and with the same intensity, when the equalities $I'_y = I'_z = |I_3|$ and $I_4 = 0$ hold, and the functions

$$\beta_y = \exp\left(-\frac{\gamma x}{2}\right), \quad \beta_z = \frac{1}{\beta_y} = \exp\left(\frac{\gamma x}{2}\right)$$

are ordinary exponential functions with exponents of opposite signs. Consequently, the intensities of the interacting light waves do not change ($I_S = I_R = I_0/2$) for any magnitude of the reduced length γx. However, their polarization vectors rotate monotonically upon an increase in γx, so that

$$|S_y'| = |R_z'| \longrightarrow \sqrt{I_0/2}, \quad |S_z'| = |R_y'| \longrightarrow 0$$

for $\gamma x \gg 1$ ($|S_z'|/|S_y'| = |R_y'|/|R_z'| \approx 0.08$ for $\gamma x = 5$). Thus, under the saturation conditions, the polarization vectors of interacting waves are oriented along different intrinsic axes of matrix **H** (are mutually orthogonal), and the amplitude of the photorefractive grating, which is proportional to the coefficient of modulation of the interference pattern, is equal to zero.

When light waves with the TE and TM polarization incident on the crystal have different intensities, energy fluxes between the interacting waves have different directions and compensate each other only in the regime of saturation, when $I_S = I_R = I_0/2$ again. In this case, the flux directed towards the light wave with a lower intensity is larger than the opposite flux. Consequently, the wave with a lower intensity is amplified upon an increase in γx, while the wave with a higher intensity attenuates. If we turn the sample crystal through 180° about the x axis (see Fig. 1) or change the initial ratio $\beta_{\text{first}} = I_{R0}/I_{S0}$ of the intensities of incident waves by the reverse ratio $\beta_{\text{second}} = \beta_{\text{first}}^{-1}$ without rotating the crystal, the weak wave will be amplified all the same. This is typical of a nonunidirectional energy exchange in the two-wave interaction on a local or nonlocal photorefractive grating in cubic gyrotropic crystals [21, 23–27]. Figure 2 shows the dependences of the amplitudes S_y', S_z', and R_y', R_z' and the intensities I_S and I_R of the light waves on γx for $\beta = I_{R0}/I_{S0} = 0.5$. In contrast to the previous case, neither of the amplitudes is equal to zero in saturation; i.e., the polarization vectors of the light waves do not coincide with the intrinsic axes of the coupling matrix. The polarization vector of the signal wave, which rotates upon an increase in γx, does not reach the position coinciding with the y axis (the angle between this vector and the y axis is $\theta_S = -35.2°$ for $\gamma x = 15$). The polarization vector of the reference wave passes though the position of the z' axis ($\theta_R = 54.7°$ for $\gamma x = 15$). However, the angle between the polarization vectors $\Delta\theta \longrightarrow 90°$ for $\gamma x \gg 1$ as in the previous case. In saturation, we have $S_y' = R_z'$ and $S_z' = -R_y'$. Typically, the intensities of the light waves approach monotonically the value $I_0/2$ upon an increase in γx. Figure 3 shows the dependences of the amplitudes S_y', S_z' and R_y', R_z' and the intensities I_S and I_R of the light waves on γx for the ratio $\beta = I_{R0}/I_{S0} = 2$, the polarizations of the signal wave along the intrinsic axis

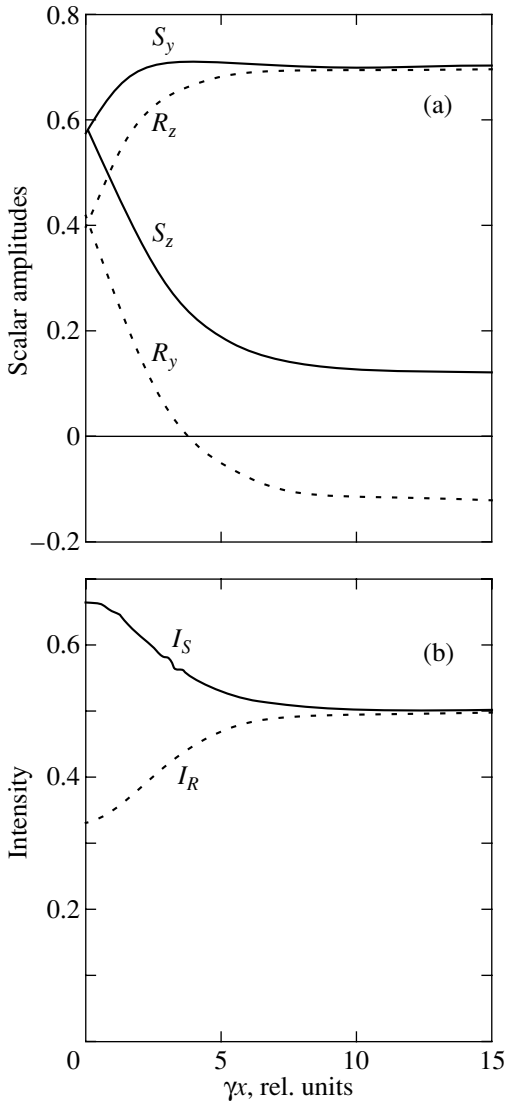


Fig. 2. Distribution of (a) scalar amplitudes and (b) intensities of light waves over the interaction length in the transverse configuration for the TE (or TM) polarization of incident waves.

z' of the coupling matrix, and the TE polarizations of the reference wave. In this case, as before, the angle between the polarization vectors of the light waves $\Delta\theta \longrightarrow 90°$ for $\gamma x \gg 1$, but this saturation value is attained upon an increase in γx slightly more slowly ($\theta_S = -65.5°$ and $\theta_R = 23.9°$ for $\gamma x = 15$). In contrast to the previous case, the polarization vector of the signal wave, which rotates upon an increase in γx, passes through the position of the y' axis, while the polarization vector of the reference wave does not attain the position of the z' axis. The intensity distribution for light waves over the interaction length is not monotonic. For $\gamma x < 3$, the reference wave is amplified, while the signal wave attenuates; for $\gamma x = 3$, we have $I_R = 0.87$ and $I_S = 0.13$. For $\gamma x > 3$, the intensity of the reference

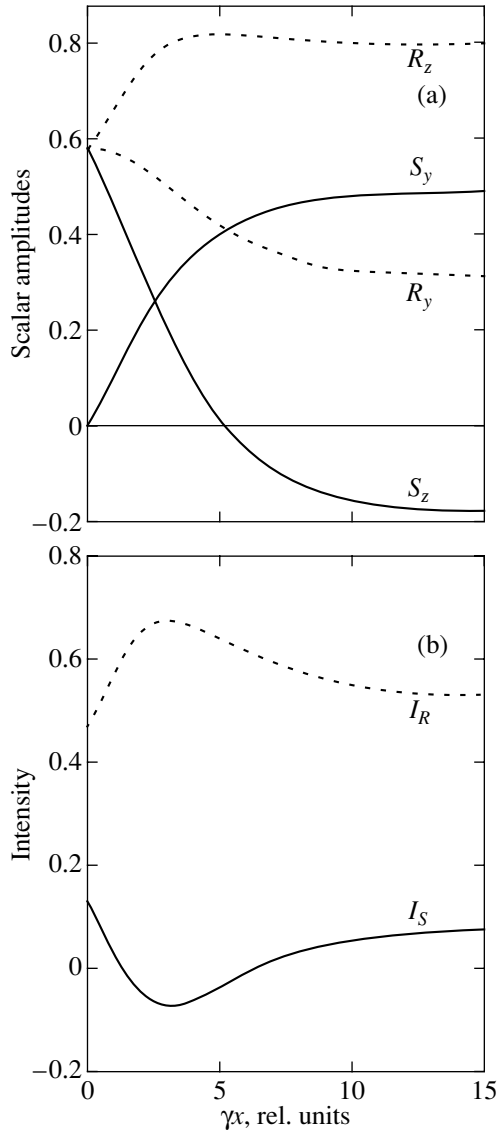


Fig. 3. Distribution of (a) scalar amplitudes and (b) intensities of light waves over the interaction length in the transverse configuration for the TE polarization of the incident reference wave and for the polarization of the signal wave along the intrinsic axis z' of the coupling matrix.

wave decreases, while the intensity of the signal wave increases to saturation values ($I_{R\infty} = 0.77$ and $I_{S\infty} = 0.23$).

If the waves incident on the crystal are polarized along the y' axis, we have $I_3 = I_4 = I'_z = 0$ and $I'_y = I_0$. If, however, the waves are polarized along the z' axis, we have $I_3 = I_4 = I'_y = 0$ and $I'_z = I_0$. In both cases, the orientation of the polarization vectors does not change, and the intensity distribution of light waves over γx is described by the conventional relations from the scalar theory [1]:

$$I_S = \frac{I_0}{1 + \beta \exp(-\gamma x)}, \quad I_R = \frac{I_0}{1 + \beta^{-1} \exp(\gamma x)}.$$

5. ENERGY EXCHANGE FOR DIFFERENT CONFIGURATIONS OF INTERACTION AND FOR THE SAME LINEAR POLARIZATION OF INCIDENT WAVES

It is convenient to use the two-wave amplification coefficient

$$\Gamma = \frac{1}{x} \ln \frac{I_S I_{R0}}{I_{S0} I_R}$$

as a quantity characterizing the effectiveness of the overall energy exchange in the case of the two-wave interaction on a photorefractive grating. In the case under investigation, this coefficient can be represented in the form

$$\Gamma = \frac{1}{x} \ln \frac{S_y'^2 + S_z'^2}{R_y'^2 + R_z'^2} \beta. \qquad (22)$$

In the case of the scalar interaction, coefficient Γ is independent of the ratio $\beta = I_{R0}/I_{S0}$ of the intensities of incident waves and of the interaction length x [1, 4, 6, 9–11].

In the general case, the polarization of light waves changes as a result of self-diffraction, leading to a change in the direction of energy exchange upon an increase in the reduced length γx for an invariable intensity ratio β of the incident waves [21–27] or dependence $\Gamma(\beta)$ for $\gamma x \equiv$ const. If one of the inequalities $\Gamma(\beta) > 0$ or $\Gamma(\beta) < 0$ is valid in the interval $\beta \in (0, \infty)$, the direction of the overall energy pumping does not change, and the energy exchange is unidirectional. If, however, the values of $\Gamma(\beta)$ reverse sign upon a variation of β in the given interval, the energy exchange is not unidirectional. In the same configuration and under the condition $\gamma x \gg 1$, the form of the dependence $\Gamma(\beta)$ is determined to a considerable extent by the orientation of the polarization vectors of the light waves incident on the crystal. This is illustrated by the dependences $\Gamma(\beta)$ shown in Figs. 4a, 4c, and 4e for the transverse and diagonal configurations of the interaction as well as the last configuration from those given in the table, respectively, for $\gamma x = 10$ and various angles θ'_0 between the polarization vectors of the incident waves and the intrinsic axis y' of matrix **H**. In turn, the form of the dependence of the coefficient Γ on the orientation of the polarization vectors is determined by the value of β. The dependences of the coefficient $\Gamma(\theta'_0)$ for different values of β in the cases listed above are shown in Figs. 4b, 4d, and 4f. It should be noted that, in the transverse configuration, the dependences $\Gamma(\beta)$ for a certain angle θ'_0 and $\Gamma(\beta)$ for the angle $90° - \theta'_0$ are symmetric about the origin of coordinates, while the dependences $\Gamma(\theta'_0)$ for a certain intensity ratio β_{first} of the incident

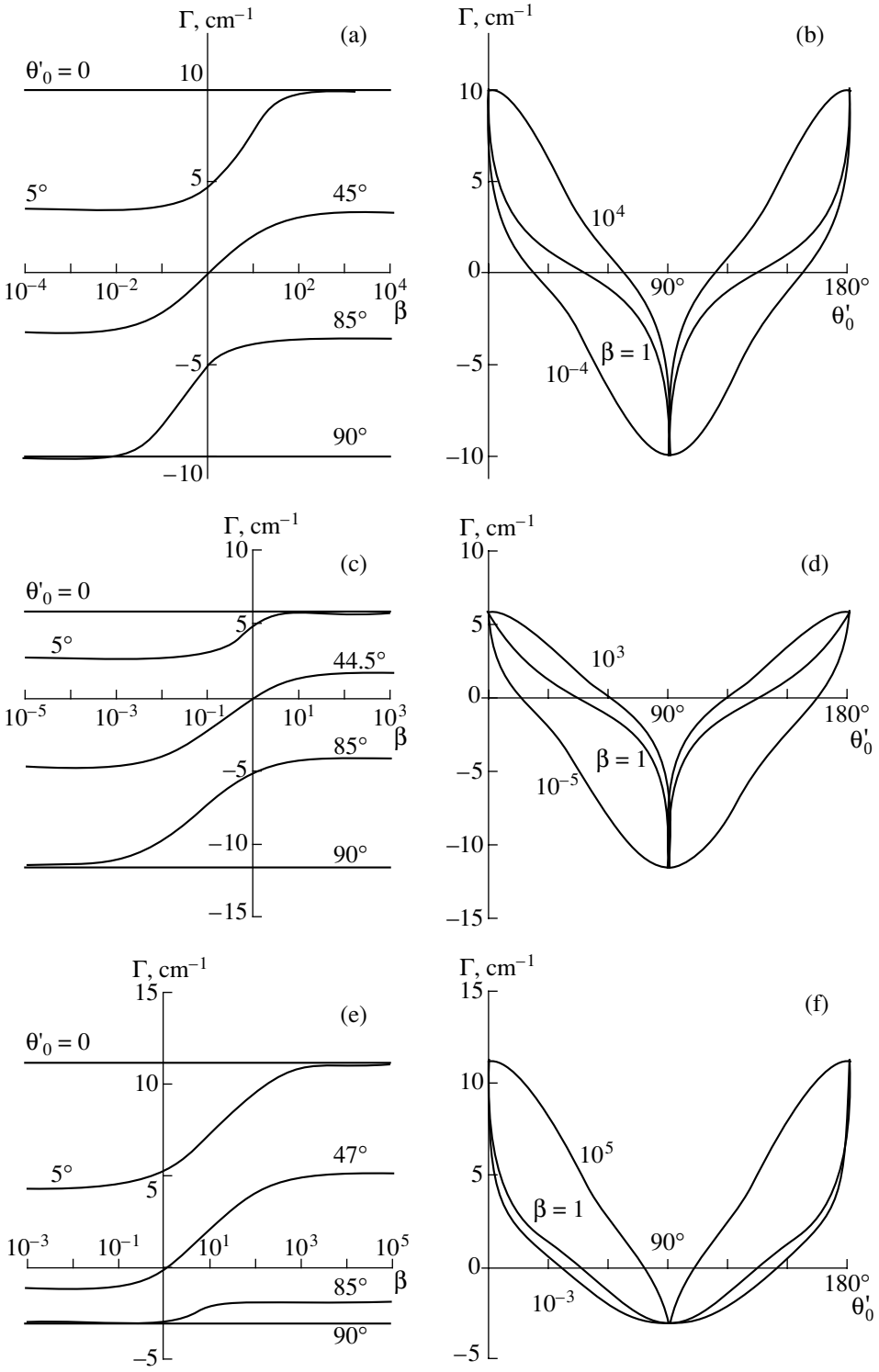


Fig. 4. Dependences of the two-wave amplification coefficient on the intensity ratio β of incident waves and on their polarization angle θ'_0 for (a, b) transverse, (c, d) diagonal, and (e, f) the last configuration from the table.

waves and $\Gamma(90° - \theta'_0)$ for $\beta_{second} = \beta_{first}^{-1}$ are symmetric about the abscissa axis. This is due to the fact that the coefficients H_{EE} and H_{MM} determining the effectiveness of energy exchange between identical polarization components of the light waves are equal to zero: $H_{EE} = H_{MM} = 0$ ($H_\Sigma = 0$). For the other two configurations of the interaction, $H_\Sigma \neq 0$, and hence the dependences $\Gamma(\beta)$ and $\Gamma(\theta'_0)$ do not possess the above symmetries.

Coefficient Γ is independent of β if the polarization vectors of the incident light waves are oriented either along the intrinsic axis y' of matrix $\mathbf{H}$, when $\Gamma = \gamma p_y$, or along the intrinsic axis z', when $\Gamma = \gamma p_z$. Since the nonunidirectional contribution in this case is equal to zero, the effectiveness of the overall energy exchange has the maximal value. For the longitudinal configuration, $p_{y,z} = \pm 1$; consequently, the values of the amplification coefficient for $\theta_0' = 0$ and $\theta_0' = 90°$ differ only in sign: $\Gamma = \gamma$ and $\Gamma = -\gamma$. In the diagonal configuration, the eigenvalues $p_{y,z}$ differ in absolute value also; consequently, the change in the angle of polarization of incident waves from $\theta_0' = 0$, when $\Gamma = \gamma/\sqrt{3}$, to $\theta_0' = 90°$, when $\Gamma = -2\gamma/\sqrt{3}$, not only changes the direction of overall energy pumping, but also doubles its effectiveness. In the case of the latter configuration, the difference between the coefficient $\Gamma \approx 1.1\gamma$ for $\theta_0' = 0$ and $\Gamma \approx -0.3\gamma$ for $\theta_0' = 90°$ is even stronger. A typical feature of a scalar interaction is that the rotation of the crystal through 180° about the x axis results in the coincidence of the values of Γ for $\theta_0' = 0$ and $\theta_0' = 90°$ with the values of this coefficient for $\theta_0' = 90°$ and $\theta_0' = 0$, respectively, before this rotation.

When the polarization vectors deviate from the direction of the intrinsic axes of the coupling matrix, the coefficient Γ becomes a monotonic function of β with two asymptotes parallel to the abscissa axis, whose position on the plots strongly depends on angle θ_0'. The asymptotic values of this coefficient $\Gamma(\beta \longrightarrow 0) = \Gamma_R$ and $\Gamma(\beta \longrightarrow \infty) = \Gamma_S$ for identical linear polarizations of incident light waves can be obtained with the help of relations (18) and (19), respectively, in the form

$$\Gamma_{S,R} = \gamma[H_\Sigma + \cos(2\theta_0')D] \pm \frac{1}{x}\ln\left\{1 + \left\{\frac{\sin(2\theta_0')D}{H_\Sigma + \cos(2\theta_0')D} \times \left[1 - \exp\left[\mp\frac{\gamma x}{2}(H_\Sigma + \cos(2\theta_0')D)\right]\right]\right\}^2\right\}. \quad (23)$$

It should be noted that the coefficient Γ obtained on the basis of the solution to Eq. (11) in the approximation of a given light field of an intense reference wave ($R_y = R_{y0}$, $R_z = R_{z0}$) coincides with the coefficient $\Gamma_S = \ln(I_S/I_{S0})$. In turn, the coefficient of two-wave amplification obtained on the basis of the solution to Eq. (12) in the approximation of a given light field of an intense signal wave ($S_y = S_{y0}$, $S_z = S_{z0}$) is equal to coefficient Γ_R with the opposite sign ($\ln(I_R/I_{R0})/x = -\Gamma_R$). Under the condition

$$\frac{\gamma x[H_\Sigma + \cos(2\theta_0')D]}{2} \gg 1,$$

which is satisfied in the cases corresponding to the monotonic curves in Fig. 4, the asymptotic values satisfy the following approximate relations:

$$\Gamma_S \approx \gamma[H_\Sigma + \cos(2\theta_0')D],$$
$$\Gamma_R \approx 2\ln\left|\frac{H_\Sigma + \cos(2\theta_0')D}{\sin(2\theta_0')D}\right|\frac{1}{x}.$$

It is worth noting that, if this condition is satisfied, the coefficient Γ_R is independent of the coupling constant γ.

The inequality $\Gamma_S \neq \Gamma_R$ indicates nonreciprocity of the amplification of a weak signal wave for $\beta_{\text{first}} \longrightarrow \infty$ and attenuation of a weak reference wave for $\beta_{\text{second}} = (\beta_{\text{first}})^{-1} \longrightarrow 0$. This is due to the fact that the contribution to the intensity of the weak light field, whose vector amplitude can be obtained in the form

$$\mathbf{A} = \mathbf{A}_0\exp\left[\pm\frac{\gamma x}{2}(H_\Sigma + \cos(2\theta_0')D)\right] - \mathbf{A}_\perp\frac{\sin(2\theta_0')D}{H_\Sigma + \cos(2\theta_0')D} \times\left\{1 - \exp\left[\pm\frac{\gamma x}{2}(H_\Sigma + \cos(2\theta_0')D)\right]\right\} \quad (24)$$

in the approximation of a preset field of the strong wave, comes not only from the component proportional to the vector amplitude $\mathbf{A}_0$ of the incident wave ($\mathbf{A}_0 \cdot \mathbf{A}_0 = I_{A0}$), but also from the orthogonal component proportional to $\mathbf{A}_\perp$ ($\mathbf{A}_\perp \cdot \mathbf{A}_\perp = I_{A0}$, $\mathbf{A}_0 \cdot \mathbf{A}_\perp = 0$). The plus sign in front of the coupling constant γ in formula (24) corresponds to amplification of a weak signal wave ($\mathbf{A} \equiv \mathbf{S}$), while the minus sign corresponds to attenuation of a weak reference wave ($\mathbf{A} \equiv \mathbf{R}$). It should be noted that relation (24) is similar in structure to the relations derived in [21, 23–27] for the vector amplitude of a weak wave interacting with a strong pumping wave in a gyrotropic photorefractive crystal of symmetry 23. The polarization vector of the strong wave is equal to $\mathbf{A}_0/|\mathbf{A}_0|$; consequently, the orthogonal component of the weak light field does not affect the interference pattern with the modulation coefficient

$$m = m_0\exp\left[\pm\gamma x\frac{H_\Sigma + \cos(2\theta_0')D}{2}\right].$$

Consequently, it does not produce a reverse effect on the photorefractive grating whose amplitude is proportional to m. This component appears as a result of diffraction of a strong wave by this grating (formed by the

light field components proportional to $\mathbf{A}_0/|\mathbf{A}_0|$), which is accompanied by the transformation of the polarization state to the orthogonal state. Such a diffraction always amplifies a weak light wave. On the other hand, the interaction between the light field components proportional to $\mathbf{A}_0/|\mathbf{A}_0|$ may either amplify or suppress this component of the weak wave. In the case of amplification, the energy fluxes associated with diffraction and interaction have the same direction, while the directions of these fluxes in the case of attenuation are different. When the intensities of incident light waves are leveled out ($\beta \longrightarrow 1$), both waves acquire orthogonal polarization components. In this case, the interference pattern is formed both due to the interference of the components with the initial polarization proportional to $\mathbf{A}_0$ and due to the interference of the components proportional to $\mathbf{A}_\perp$, which also produce the reverse effect on the photorefractive grating in this case.

For the polarization angle $\theta_0' \neq 0$ and $\theta_0' \neq 90°$, two fluxes in different directions, which determine the overall effectiveness of energy exchange between the interacting light waves, compensate each other ($\Gamma = 0$) for a certain value of θ_0', which depends on the intensity ratio β of incident waves, the reduced length γx, and the orientation of interaction relative to the crystal axes (see Figs. 4b, 4d, and 4f). For a small value of $\gamma x \ll 1$, the dependences of the two-wave amplification coefficient on θ_0' remain virtually unchanged upon the variation of β:

$$\Gamma(\theta_0') \approx \gamma[H_\Sigma + \cos(2\theta_0')D].$$

In this case, $\Gamma = 0$ for $\theta_0' = 45°$ and $\theta_0' = 135°$ in the case of the transverse configuration, $\theta_0' = 35.3°$ and $\theta_0' = 144.7°$ in the case of the diagonal configuration, and $\theta_0' = 62.6°$ and $\theta_0' = 117.4°$ in the case of the last configuration from the table. In the general case, an increase in γx leads to an increase in one of the nonunidirectional fluxes and a decrease in the other flux. These fluxes are now compensated for a different angle of polarization. The only exception is the case of transverse configuration of interaction considered in the previous section for $\theta_0' = \pm 45°$ and $\beta = 1$, when $\Gamma = 0$ for any value of γx.

The polarization dependences $\Gamma(\theta_0')$ indicate that, for $\beta \gg 1$, the attenuation of a weak light wave as a result of self-diffraction in all the above cases has a strong selectivity in angles θ_0' in the vicinity of $\theta_0' = 0$. For the transverse and diagonal configurations, such a selectivity also exists for $\beta \ll 1$ in the vicinity of $\theta_0' = 90°$. This effect is associated with the existence of a nonunidirectional energy flux, which always amplifies a weak wave and increases upon a deviation of the polarization of incident waves from the directions of intrinsic axes of matrix $\mathbf{H}$. For $\beta \longrightarrow 1$, this flux increases for the above deviation at a lower rate, and selectivity is deteriorated. However, for $\beta = 1$, in contrast to the cases $\beta \ll 1$ and $\beta \gg 1$, the dependence of the effectiveness of energy exchange on the angle θ_0' may have two selective regions. For $\gamma x \ll 1$, when the contribution of nonunidirectional energy exchange is insignificant, there are no selective regions in the polarization dependence of coefficient Γ.

6. CONCLUSIONS

Thus, we have analyzed the energy exchange and variations of the polarization state of light waves interacting on a transmitting nonlocal photorefractive grating in a crystal of symmetry $\bar{4}3m$ of an arbitrary orientation. It is shown that contribution to energy exchange may come from nonunidirectional fluxes emerging as a result of transformation of the polarization state of light waves. Exact analytic solutions and the results of numerical analysis were obtained in the approximation for the grating amplitude linear in the modulation coefficient m for the case when linearly polarized light waves are incident on the crystal. It should be noted that the first integrals of the equations for coupled waves obtained by us are independent of corrections nonlinear in m. For this reason, the main qualitative features of the interaction, which are associated with the mutual effect of the energy exchange between the light waves and the change in their polarization state, remain unchanged after the inclusion of these corrections.

In particular cases of interaction when the polarization of incident waves coincides with intrinsic axes of the coupling matrix, the polarization of light waves in the crystal does not change, and the energy exchange between waves is unidirectional. In the general case, the light field in the crystal contains polarization components orthogonal to the polarization of incident light waves. These components make a nonunidirectional contribution to the overall energy exchange, which always amplifies a weak light wave. Typically, for a large difference in the intensities of incident waves, when the light field of a strong wave can be regarded as preset, the orthogonal polarization component of a weak wave appears due to anisotropic diffraction by an inhomogeneous photorefractive grating and does not produce a reciprocal effect on it.

Depending on the orientation of the polarization vectors of incident light waves and on their intensity ratio, a unidirectional energy exchange without a change in the orientation of polarization vectors of the interacting waves or a nonunidirectional energy exchange with a rotation of polarization vectors can be realized. In the particular case of transverse configuration of the two-wave interaction with the same intensity of the incident waves with the TE or TM polarization, energy fluxes propagating in different directions compensate each other completely, and the intensities of the waves in the crystal remain unchanged. Nevertheless,

their polarization vectors change their orientation upon an increase in the reduced interaction length γx so that, for $\gamma x \gg 1$, these vectors become orthogonal and oriented along different intrinsic axes of the coupling matrix. Under this condition, the polarization vectors for all the configurations of interaction involving a change in the polarization state of light waves are orthogonal for an arbitrary intensity ratio of the waves incident on the crystal. However, in the general case, the orientation of polarization vectors for $\gamma x \gg 1$ does not coincide with the orientation of the intrinsic axes of the coupling matrix.

ACKNOWLEDGMENTS

The author is grateful to S.M. Shandarov and B.N. Poizner for fruitful discussions.

REFERENCES

1. V. L. Vinetskiĭ, N. V. Kukhtarev, S. G. Odulov, *et al.*, Usp. Fiz. Nauk **129**, 113 (1979) [Sov. Phys. Usp. **22**, 742 (1979)].
2. B. Ya. Zel'dovich, N. F. Pilipetskiĭ, and V. V. Shkunov, *Wave Front Reversal* (Nauka, Moscow, 1985).
3. Y. R. Shen, *The Principles of Nonlinear Optics* (Wiley, New York, 1984; Nauka, Moscow, 1989).
4. S. G. Odulov, M. S. Soskin, and A. I. Khizhnyak, *Dynamic Lattice Lasers* (Nauka, Moscow, 1990).
5. N. Kukhtarev, V. Markov, and S. Odulov, Opt. Commun. **23**, 338 (1977).
6. N. V. Kukhtarev, V. B. Markov, S. G. Odulov, *et al.*, Ferroelectrics **22**, 949 (1979); **22**, 961 (1979).
7. S. I. Stepanov and G. S. Trofimov, Zh. Tekh. Fiz. **55**, 559 (1985) [Sov. Phys. Tech. Phys. **30**, 331 (1985)].
8. P. Tayebati and D. Mahgerefteh, J. Opt. Soc. Am. B **8**, 1053 (1991).
9. M. P. Petrov, S. I. Stepanov, and A. V. Khomenko, *Photorefractive Crystals in Coherent Optics* (Nauka, St. Petersburg, 1992).
10. B. I. Sturman and V. M. Fridkin, *Photovoltaic Effect in Media without Center of Symmetry and Related Phenomena* (Nauka, Moscow, 1992).
11. B. I. Sturman, Zh. Tekh. Fiz. **48**, 1010 (1978) [Sov. Phys. Tech. Phys. **23**, 589 (1978)].
12. M. P. Petrov, S. V. Miridonov, S. I. Stepanov, *et al.*, Opt. Commun. **31**, 301 (1979).
13. N. V. Kukhtarev, B. D. Pavlik, V. V. Sorokin, *et al.*, Kvantovaya Élektron. (Moscow) **13**, 326 (1986).
14. N. V. Kukhtarev, M. S. Brodin, and V. I. Volkov, Fiz. Tverd. Tela (Leningrad) **30**, 2757 (1988) [Sov. Phys. Solid State **30**, 1588 (1988)].
15. G. Pauliat, C. Besson, and G. Roosen, IEEE J. Quantum Electron. **23**, 1736 (1989).
16. C. Stace, A. K. Powell, K. Walsh, *et al.*, Opt. Commun. **70**, 509 (1989).
17. G. A. Brost, J. Opt. Soc. Am. B **9**, 1454 (1992).
18. H. C. Pedersen and P. M. Johansen, J. Opt. Soc. Am. B **12**, 592 (1995).
19. H. Tuovinen, A. A. Kamshilin, and T. Jaaskelainen, J. Opt. Soc. Am. B **14**, 3383 (1997).
20. B. I. Sturman, E. V. Podivilov, K. H. Ringhofer, *et al.*, Phys. Rev. E **60**, 3332 (1999).
21. R. V. Litvinov and S. M. Shandarov, Opt. Spektrosk. **83**, 334 (1997) [Opt. Spectrosc. **83**, 313 (1997)].
22. Yi Hu, K. H. Ringhofer, and B. I. Sturman, J. Appl. Phys. **68**, 931 (1999).
23. V. Yu. Krasnoperov, R. V. Litvinov, and S. M. Shandarov, Fiz. Tverd. Tela (St. Petersburg) **41**, 632 (1999) [Phys. Solid State **41**, 568 (1999)].
24. R. V. Litvinov, S. M. Shandarov, and S. G. Chistyakov, Fiz. Tverd. Tela (St. Petersburg) **42**, 1397 (2000) [Phys. Solid State **42**, 1435 (2000)].
25. R. V. Litvinov and S. M. Shandarov, Kvantovaya Élektron. (Moscow) **31**, 973 (2001).
26. A. G. Mart'yanov, S. M. Shandarov, and R. V. Litvinov, Fiz. Tverd. Tela (St. Petersburg) **44**, 1006 (2002) [Phys. Solid State **44**, 1050 (2002)].
27. R. V. Litvinov, Kvantovaya Élektron. (Moscow) **32**, 535 (2002).
28. A. A. Chaban, Zh. Éksp. Teor. Fiz. **57**, 1387 (1969) [Sov. Phys. JETP **30**, 751 (1969)].
29. S. I. Stepanov and M. P. Petrov, Opt. Commum. **53**, 292 (1985).
30. Ph. Refregier, L. Solymar, H. Rajbenbach, *et al.*, J. Appl. Phys. **58**, 45 (1985).
31. K. Blotekjaer, J. Appl. Phys. **48**, 2495 (1977).
32. M. G. Moharam, T. K. Gaylord, R. Magnusson, *et al.*, J. Appl. Phys. **50**, 5642 (1979).
33. R. Saxena and T. Y. Chang, J. Opt. Soc. Am. B **9**, 1467 (1992).
34. A. V. Dugin, B. Ya. Zel'dovich, P. N. Il'inykh, *et al.*, Kvantovaya Élektron. (Moscow) **19**, 1129 (1992).
35. S. M. Shandarov, N. I. Nazhestkina, O. V. Kobozev, *et al.*, J. Appl. Phys. **68**, 1007 (1999).
36. G. V. Calvo, B. Sturman, F. Agulló-Lopez, *et al.*, Phys. Rev. Lett. **84**, 3839 (2000).
37. O. V. Kobozev, A. E. Mandel', S. M. Shandarov, *et al.*, Kvantovaya Élektron. (Moscow) **30**, 514 (2000).
38. S. M. Shandarov, V. Yu. Krasnoperov, V. A. Kartashov, *et al.*, Neorg. Mater. **37**, 728 (2001).
39. R. V. Litvinov, Izv. Vyssh. Uchebn. Zaved., Fiz., No. 10, 20 (2001).
40. A. P. Kartashov and B. L. Rozhdestvenskiĭ, *Ordinary Differential Equations and Foundations of Calculus of Variations* (Nauka, Moscow, 1986).

Translated by N. Wadhwa

Journal of Experimental and Theoretical Physics, Vol. 95, No. 5, 2002, pp. 833–843.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 122, No. 5, 2002, pp. 965–977.

ATOMS, SPECTRA, RADIATION

Model of a Periodic Sequence (Trajectory) of the Results of Measurements of Atomic States at a Micromaser Output

G. P. Miroshnichenko

St. Petersburg State Institute of Fine Mechanics and Optics (Technical University), St. Petersburg, 197101 Russia
e-mail: gpmirosh@yahoo.com; mirosh@mkk.ifmo.ru

Received March 28, 2002

Abstract—The method of stochastic recurrent relation is used for simulating a random sequence (trajectory) of the counts of a detector of atomic states at the output of a single-atom micromaser. A random sequence of the relative frequencies of counts of a detector during a fixed time interval is calculated. The frequencies of counts vary randomly near the average time-independent level. It is assumed that these average levels are reproducible observables for each random trajectory. A micromaser can jumpwise transfer from one average level of the relative frequencies of counts to another. It is assumed that a certain subensemble of the states of a field mode corresponds to each measured average level of the frequencies of counts. A method is proposed for calculating possible average relative frequencies of counts and corresponding (mean) reduced density matrices of the mode ρ_{st}. The matrix ρ_{st} characterizes a micromaser during its development along a specified periodic trajectory. It is found by solving the eigenvalue problem for the evolution operator on the period. An analytic method for solving this problem is developed. The matrix ρ_{st} is the solution to the inverse problem of the reconstruction of the statistics of a field mode from the statistics of a random trajectory. The procedure of selecting the parameters of the evolution operator on the period is discussed through a numerical example. © *2002 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

At present, many fundamental concepts of quantum mechanics can be directly verified and find practical applications due to a high level of experimental methods in atomic physics and quantum optics. Thus, the concept of entangled states of quantum-mechanical systems is used in rapidly developing scientific fields such as quantum data processing and quantum calculations and communication. The operation of quantum information and communication systems is based on quantum measurements in information elements during which the reduction of the state occurs. Individual atoms or ions captured in traps, as well as single photons, play the role of cubits. Experiments in which measurements are performed in a specially prepared individual quantum system solve the general theoretical problems of the interpretation of fundamentals of quantum mechanics [1].

In this respect, the so-called single-atom micromaser offers interesting possibilities [2]. In such a micromaser, an individual quantum system—a separated mode of the microwave cavity—interacts during each period (and, hence, becomes entangled) with the second quantum system—a Rydberg atom, which is excited to a maser energy level before entering the cavity. At the instant the atom leaves the cavity (and the systems no longer interact with each other but remain in the entangled state), a quantum measurement of the energy of the atom is performed. This measurement yields indirect information on the state of a quantized mode at the measurement instant. The measurement process is repeated during each flight of the atom, resulting, due to the so-called reverse action of a measuring instrument on an object being measured, in the unusual dynamics of a field mode. In the literature devoted to these problems, it is pointed out that such indirect measurements serve as a main source of information on the state of a quantized microwave mode of the cavity.

The problem concerning the relation between the statistical properties of a mode being generated with the statistics of successive counts of detectors, which are selective with respect to the states of escaping atoms, has been considered in many papers. The authors of paper [3] performed numerical simulation of the dynamics of a field mode of a micromaser subjected to quantum measurements of atomic states. They used the recurrent Filipowicz relation [2] for calculating the reduced density matrix (RDM) of the field after each measurement event. Such a calculation yields the realization of a random process, which represents a sequence of detector counts and the related RDM sequence. We will call this simulation technique the method of stochastic recurrent relation (SRR). The scheme of calculations was ultimately generalized, the atoms were incident on a detector at equal time intervals, and the detector had 100% efficiency. The only random variable used in calculations simulated the detector operation and showed in which state, the upper or lower, the atom was detected. The authors of paper

1063-7761/02/9505-0833$22.00

[3] have obtained important results, which were qualitatively confirmed in experimental paper [4]. It was shown in these papers that, in the steady state at a certain relation between parameters in the sequence of detector counts, several (in these papers, two) quasistationary states, which transformed jumpwise to each other at random moments, were observed. Such a behavior of a micromaser is explained by a hypothesis of quantum jumps performed by the micromaser from one minimum of the effective Filipowicz potential [2] to another, with close depth. More complicated SRR calculations were carried out in papers [5, 6], where random variables were introduced which take into account the Poisson scatter of the expectation time between the flights of subsequent atoms and the imperfect parameters of detectors. The authors of paper [5] have obtained and verified experimentally the expression relating the Mandel parameter Q for the number of photons of a field mode with a similar parameter for atoms detected in the ground state. SRRs were used in papers [6] for simulating measuring schemes including a classical microwave field and allowing the study of the dynamics of a phase of a quantum mode during measurements of atomic subsystems.

In a series of papers [7–10], the theory of a micromaser taking into account the detection of the atomic subsystem was developed based on the time "coarse-grain" differential master equation obtained in paper [7]. In paper [8], a nonlinear master equation was found which describes the dynamics of a field mode in the time interval between successive detector counts. In paper [9], a linear master equation was obtained for a not normalized conditional RDM, and in paper [10] a propagator was derived for the corresponding subensemble. In papers [8–10], basic concepts of a new field—the statistics of atomic counts—were formulated, which allow one to relate the statistical properties of the sequence of counts of a detector selective with respect to atomic states with the statistical properties of a field mode. In these papers, the expressions were obtained for the joint probability of detecting k atoms in the ground state and m atoms in the excited state during a fixed observation interval. By using these expressions, the authors found the Mandel parameter Q for atoms in a certain energy state, expressions for correlation and cross-correlation functions for detector counts separated by a fixed time interval (irrespective of the states of atoms detected during the chosen time interval), and expressions for the probability density of the expectation time between two successive counts. These papers confirm the results obtained in paper [5] and present general expressions describing the relation between two-atomic correlation functions with two-time correlation functions of the intensity of the micromaser field (for equal time intervals).

The concepts of papers [8–10] were generalized in subsequent papers [11–15]. Thus, the authors of papers [11, 12] studied the effect of the non-Poisson pumping of a micromaser on the state-selective statistics of atoms leaving a detector. The consideration of the non-Poisson statistics required the use of the theory of stochastic point processes for the description of the statistics of the arrival times of the pump atoms. The authors of paper [13], where studies [6] were developed, obtained expressions relating the statistics of atomic counts with the temporal evolution of the intracavity field phase. By using a linear master equation, the authors of paper [14] obtained new expressions for the statistical parameters of atomic counts (in particular, they determined the probability of n detector counts for a specified sequence of the counts). The validity of the expressions was verified by solving numerically the master equation. In paper [15], the first attempt was made to determine the spectrum of a micromaser in the state belonging to a subensemble, when the RDM was developed between the subsequent detector counts. The calculation was performed assuming that the detector efficiency was low.

We developed in this paper an alternative method for measuring statistical parameters of the micromaser field using the statistical processing of the counts of a detector of atomic states. The method is based on the Von Neumann quantum theory of measurements [16]. A quantum measurement event results in the selection of a certain quantum-mechanical subensemble of the field-mode states. The separation into subensembles is determined by a detector and depends on its parameters. Information obtained in the measurement event can be used for a more detailed study of a field-mode state at different instants of time. Thus, according to papers [3, 4], one can determine in which of the minima of the Filipowicz potential [2] a quantum mode is located at the observation moment and study the characteristics of a quantum jump of the mode to a competing minimum. Our method is based on the fact that during the residing time of the mode in a certain potential minimum, the average statistical characteristics of the sequence of detector counts have time to be established. We can separate a subensemble of the field-mode states and find the average stationary RDM of the subensemble by measuring the average relative frequency of counts of the selective detector. We assume that, to find the average RDM of the subensemble, we can use, instead of a random sequence of (trajectory) of detector counts, a periodic trajectory providing the measured average relative frequency of counts. Such a substitution, as shown in this paper, is possible because of a weak dependence of the SRR solutions on the order of the sequence of detector counts. In other words, random successive counts can be interchanged within some limits to replace a random trajectory by a periodic trajectory. We developed in Section 4 an analytic method for calculating a stationary RDM ρ_{st} for a periodic sequence. In fact, the method solves an inverse problem of reconstructing the RDM ρ_{st} of a subensemble of field states from statistical characteristics of the sequence of counts. By using ρ_{st}, we can calculate various average statistical characteristics of the field mode

and atomic beam acting during a specified time interval. In the papers mentioned above, this problem was solved with the help of a steady field RDM $\rho^{(ss)}$, which was obtained without indirect measurements ($\rho^{(ss)}$ is exactly defined in Section 3). However, the statistical characteristics obtained with the help of $\rho^{(ss)}$ can be correct if they are observed during large time intervals, when the micromaser has been in all its quasi-stationary states many times. The method we developed allows us to study the statistical properties of a field mode in more detail. If the micromaser has one stationary state, whereas other states are unlikely, then, as shown in this paper, the calculations performed with the help of these two matrices yield the same results.

2. MODEL OF A SINGLE-ATOM MICROMASER

The model of a single-atom micromaser is based on the Jaynes–Cummings Hamiltonian H [17] and was developed in paper [2]. The Hamiltonian has the form

$$H = \omega\hat{a}^+\hat{a} + \omega_0\hat{S}_3 + g(\hat{a}^+\hat{S}_- + \hat{a}\hat{S}_+).$$

Here, $\hat{S}_+$, $\hat{S}_-$, and $\hat{S}_3$ are the operators of the group SU(2)

$$S_3 = \frac{|1\rangle\langle 1| - |0\rangle\langle 0|}{2}, \quad S_- = |0\rangle\langle 1|, \quad S_+ = |1\rangle\langle 0|,$$

satisfying the common commutation relations; $\hat{a}^+$ and $\hat{a}$ are the Bose operators of creation and annihilation of the field-mode photons; ω is the frequency of the quantum mode of the cavity; ω_0 is the atomic transition frequency; g is the parameter of the atom–field interaction; $|0\rangle$ and $|1\rangle$ are the ground and excited atomic states of the maser levels; and $\hbar = 1$. Below, we consider the case of $\omega = \omega_0$.

According to the Filipowicz theory [2], a micromaser operates cyclically, its operation period consisting of the time τ of interaction of an atom with a quantized mode of the cavity and of the time T ($T \gg \tau$) of field relaxation to the Planck distribution. Let us describe analytically the development of the RDM of a quantum mode during a period [2]. We assume that all the atoms flying into a cavity are prepared in the excited state. Let σ_{at} be the initial density matrix of an atom. Then, before the escape of an atom from the cavity after the termination of the interaction of the mode, the RDM $\rho(t)$ is described by the expression

$$\rho(\tau) = \mathrm{Sp}_{at}[\exp(-iH\tau)\rho(0) \otimes \sigma_{at}\exp(iH\tau)].$$

Here, Sp_{at} is the trace of the matrix over atomic states. For the initial conditions chosen, as follows from paper [18], each RDM diagonal develops independently. Consider the dynamics of the principal diagonal of the field RDM—the probability distribution law of the number n of photons in the Fock basis. This relation has the following matrix form:

$$\rho_n(\tau) = \cos^2(g\tau\sqrt{n+1})\rho_n(0) + \sin^2(g\tau\sqrt{n})\rho_{n-1}(0),$$

$$\rho(t) = \sum_n \rho_n(t)|n\rangle\langle n|.$$

After the escape of an atom from the cavity, the principal diagonal of the field RDM develops as

$$\frac{d\rho_n(t)}{dt} = \gamma(n_b+1)[(n+1)\rho_{n+1}(t) - n\rho_n(t)] + \gamma n_b[n\rho_{n-1}(t) - (n+1)\rho_n(t)].$$

Here, γ is the relaxation rate of the field; n_b is the average Planck number of photons in the cavity. We retain the notation $\hat{a}^+$ and $\hat{a}$ for creation and annihilation operators, respectively, which now act on the Fock projector $|n\rangle\langle n|$ according to the rules

$$a^+|n\rangle\langle n| = \sqrt{n+1}|n+1\rangle\langle n+1|,$$

$$a|n\rangle\langle n| = \sqrt{n}|n-1\rangle\langle n-1|.$$

Then, the principal diagonal of the field RDM after the termination of the cycle of duration T can be represented with the help of the evolution operator (by neglecting the field relaxation during a short interval τ of interaction) in the form

$$\rho(T) = W(T)(Q0 + Q1)\rho(0).$$

Here,

$$W(T) = \exp(T\gamma L),$$

$$L = -(2n_b+1)\hat{a}^+\hat{a} - n_b + (n_b+1)\hat{a}\sqrt{\hat{a}^+\hat{a}} + n_b\hat{a}^+\sqrt{\hat{a}\hat{a}^+}, \quad (1)$$

$$Q0 = \hat{a}^+\frac{\sin(g\tau\sqrt{\hat{a}^+\hat{a}+1})}{\sqrt{\hat{a}^+\hat{a}+1}}, \quad (2)$$

$$Q1 = \cos^2(g\tau\sqrt{\hat{a}^+\hat{a}+1}). \quad (3)$$

The above equation describes the principal diagonal of the field RDM for a full ensemble of quantum states of the mode, without performing an indirect quantum measurement.

To simulate the process of measurements of the states of escaping atoms, we divide the full ensemble into three subensembles and introduce a random variable ξ, which satisfies the condition $\xi = 0$ if the detector found an atom in the lower state, $\xi = 1$ if an atom was found in the upper state, and $\xi = 2$ if the detector did not produce any count because its efficiency is less than unity. Let us neglect the duration of the measuring procedure and assume that the measurement is performed at the moment of the atom escape from the cavity. We

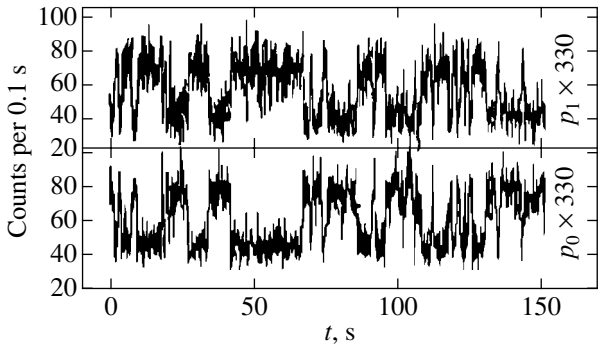


Fig. 1. Experimental results from paper [4].

denote by $\rho(l)$ the vector of the principal diagonal of the RDM at the beginning of the lth cycle, which satisfies the normalization condition

$$\mathrm{Sp}\rho(l) = \sum_{n=0} \rho_n(l) = 1. \tag{4}$$

Here, Sp is the trace over the field states. Let us denote by $S(\xi, T)$ the evolution operator of the principal diagonal of the RDM $\rho(l)$ for a subensemble corresponding to the value of the variable ξ found at the lth cycle an to the cycle duration T. We have

$$\begin{aligned} S(\xi, T) &= W(T)D(\xi), \\ D(0) &= \varepsilon_0 Q0, \quad D(1) = \varepsilon_1 Q1, \\ D(2) &= (1-\varepsilon_0)Q0 + (1-\varepsilon_1)Q1. \end{aligned} \tag{5}$$

Here, ε_1 and ε_0 are the efficiencies of detecting atoms in the upper (lower) states, respectively. It is obvious that $S(\xi, T)$ satisfies the condition (division of the total ensemble into three subensembles)

$$S(0, T) + S(1, T) + S(2, T) = W(T)(Q0 + Q1).$$

The vector of the principal RDM diagonal, found in the previous cycle, is an initial vector for the subsequent cycle. After each measurement event, this vector should be renormalized. Taking the above considerations into account, we obtain a nonlinear SRR describing the dynamics of the principal diagonal of the RDM during an indirect quantum measurement:

$$\rho(l+1) = \frac{S(\xi_l, T)\rho(l)}{\mathrm{Sp}S(\xi_l, T)\rho(l)}. \tag{6}$$

Here, the symbol ξ_l denotes the value of a random variable ξ detected in the lth cycle. Below, we will call the dependence of the random variable ξ on the number ξ_l of the operating cycle of the micromaser a trajectory. Let us denote the probability of finding atom in the lower state in the lth cycle by a detector as

$$a_0 = \mathrm{Sp}D(0)\rho(l) \tag{7}$$

and, similarly, for the upper state as

$$a_1 = \mathrm{Sp}D(1)\rho(l). \tag{8}$$

These probabilities are equal to the product of the corresponding efficiencies ε_0 and ε_1 of the detector and the quantum-mechanical probabilities of finding an atom in the upper or lower states, which are calculated from $\mathrm{Sp}Q0\rho(l)$ and $\mathrm{Sp}Q1\rho(l)$, respectively. The probabilities a_0 and a_1 are used to generate a random variable ξ in each cycle.

3. BASIC PROPERTIES OF THE STOCHASTIC EVOLUTION OF THE FIELD RDM DURING MEASUREMENT OF AN ATOMIC STATE AT THE MICROMASER OUTPUT

In experiments [4], a random sequence of detector counts was measured which showed in which state an atom was found. The randomness of this sequence is caused by a number of reasons. An atomic beam has a Poisson statistics of the expectation time T of atoms [2]:

$$P(T) = R\exp(-RT). \tag{9}$$

Here, R is the injection rate of atoms and $P(T)$ is the probability density of the expectation time. A detector which does not measure some atoms due to its low efficiency also has random properties. The results of measurements—atoms escaping from the cavity in the upper or lower state—are distributed randomly. The authors of paper [4] studied the statistics of counts selective with respect to atomic states and presented the plots of the relative frequencies of counts of a detector finding the atom in the upper (p_1) or lower (p_0) states during the observation interval $\Delta t_{av} \approx 0.1$ s. Figure 1 shows the corresponding results obtained in paper [4]. These quantities are

$$p_0 = \frac{k}{\Delta t_{av} R}, \tag{10}$$

$$p_1 = \frac{m}{\Delta t_{av} R}. \tag{11}$$

Here, k is the number of atoms found in the lower state, m is the number of atoms found in the upper state from the total number $\Delta t_{av}R$ of atoms that have flown during the time interval Δt_{av}. The frequencies of counts were determined by sampling approximately 300 flying atoms by directly counting the number of favorable events. The experiment was performed for $g\tau \approx 0.92$, $N_{ex} = R/\gamma \approx 200$, and $R \approx 3300$ s^{-1}. The data were obtained with a detector having the efficiency $\varepsilon_1 = \varepsilon_0 \approx 0.35$. The dependence was measured after the establishment of a field in the cavity, and therefore it is typical for a stationary state of a micromaser subjected to measurements. A characteristic feature of a random realization in a micromaser under stationary conditions is the existence of several (two, as in paper [4]) average relative frequencies of counts of the detector finding atoms

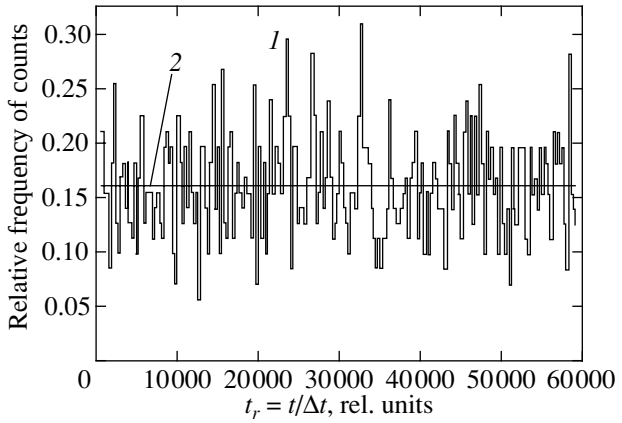


Fig. 2. The relative frequency of counts of a detector for atoms found in the ground states. The expectation times are random and are described by a Poisson distribution; (*1*) relative frequency of counts p_0 for atoms in the ground state obtained in the interval $\Delta t_{av} = 284\Delta t$; (*2*) solid straight horizontal line is the average level $\bar{p}_0$ of the frequency of counts. On the abscissa, the time $t_r = t/\Delta t$ is plotted in units of the elementary interval $\Delta t = 1/4R$. The parameters of calculation are presented in the text.

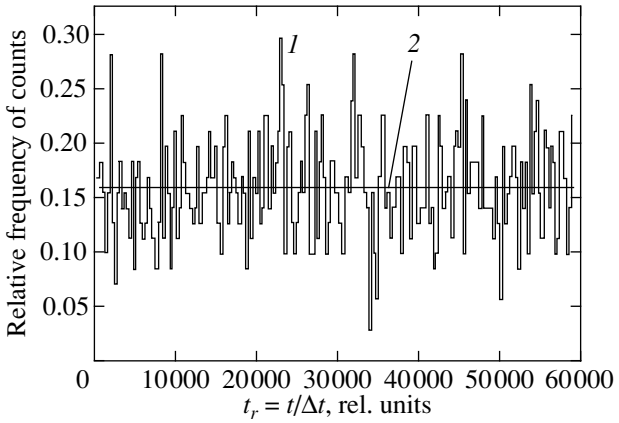


Fig. 3. The relative frequency of counts of a detector for atoms found in the ground state. Simulation is performed with the help of the evolution operator W_F averaged over the Poisson distribution of the expectation times. The average time interval between adjacent atoms is chosen to be $1/R = 4\Delta t$. The notation and calculation parameters are as in Fig. 2.

in the upper ($\bar{p}_1$) and lower ($\bar{p}_0$) states. The average relative frequencies of counts are retained during noticeable time intervals (for tens of seconds, in Fig. 1). For brevity, we will call each such interval the time t_{st} of residence in a certain stationary state. The relative frequencies of counts p_0 (10) and p_1 (11) are random functions, and they vary randomly near their average values $\bar{p}_0$ and $\bar{p}_1$ during the residence time. Sharp (quantum, according to the title of paper [4]) jumps can occur between average levels. The average number of photons in the mode drastically changes during the jump (in paper [4], approximately from 70 to 149 and vice versa). The moment of a jump itself is unpredictable, and its nature and dynamics have not been studied.

To simulate the RDM dynamics in the presence of measurements during each operation cycle of a micromaser, it is necessary to generate random variables simulating the Poisson nature of an incoming beam (variable T) and random results of the detector operation (variable ξ). The vector $\rho(l+1)$ can be calculated from expression (6), where T is the expectation time found for the lth cycle. The results of simulation are presented in Fig. 2, where on the abscissa the time $t_r = t/\Delta t$ is plotted in relative units, scaled to the elementary time interval Δt. The number of intervals Δt between adjacent atoms is random and is generated with the help of the distribution

$$\bar{P}(s) = R\Delta t(1 - R\Delta t)^{s-1}, \quad s = 1, 2, \ldots, \qquad (12)$$

which is close to a Poisson distribution. Here, $\bar{P}(s)$ is the probability of expectation of s elementary intervals between adjacent atoms or of time $T = s\Delta t$. The interval Δt is a parameter of the distribution, its choice being restricted, according to [9], by the condition

$$\tau \ll \Delta t$$

providing single-atom processes.

The plots in Figs. 2 and 3 were calculated for $\Delta t = 1/4R$. For the experiment [4] $\Delta t \approx 10^{-4}$ s, therefore, the plots presented in Figs. 2 and 3 correspond approximately to micromaser operation for 5 s. It is obvious that

$$\frac{\bar{P}(s)}{\Delta t} \xrightarrow[\Delta t \to 0]{} P(T),$$

where $P(T)$ is defined by expression (9). Curve *1* in Fig. 2 represents the dependence of the relative frequency of counts p_0 (10), obtained during the time intervals $\Delta t_{av} = N_{ex}/R = \gamma^{-1}$, on the time t_r. The interval Δt_{av}, as in paper [4], is determined by the lifetime γ^{-1} of the field mode. Solid straight line *2* in Fig. 2 shows the average steady level of the frequency of counts $\bar{p}_0$ obtained by averaging curve *1* in time. The following parameters were used in calculations:

$$N_{ex} = 71, \quad \Delta t_{av} = 284\Delta t, \quad n_b = 0.1,$$
$$g\tau = 0.92, \quad \varepsilon_0 = \varepsilon_1 = \frac{1}{3}, \quad \Delta t = \frac{1}{4R}. \qquad (13)$$

The plots presented in Fig. 2 were obtained using expression (6), when the stochasticity was produced by random variables T and ξ. Curve *1* is a theoretical model [with parameters (13)] of an experimental curve. Let us show that a random dependence with the same statistical properties can be obtained by a simpler simulation. For this purpose, we average SRR (6) over the ensemble of expectation times (12), using the method

proposed in paper [2]. The averaged expressions have the form

$$W = \frac{R\Delta t \exp(R\Delta t L/N_{ex})}{I-(1-R\Delta t)\exp(R\Delta t L/N_{ex})}, \quad (14)$$

$$S(\xi) = WD(\xi), \quad \xi = 0, 1, 2, \quad (15)$$

$$\rho(l+1) = \frac{S(\xi_l)\rho(l)}{\mathrm{Sp}S(\xi_l)\rho(l)}. \quad (16)$$

Here, I is the unit operator. In the limit

$$W \xrightarrow[\Delta t \to 0]{} W_F = \left(I - \frac{L}{N_{ex}}\right)^{-1}, \quad (17)$$

we obtain the result of Filipowicz [2] for the averaged evolution operator W_F in the relaxation process. Figure 3 shows the results of calculation performed with the same parameters (13) but using expressions (14)–(16). The SRR (16) describes the dynamics of a micromaser under the action of a periodic atomic beam with the time interval between the atoms equal to 1/R. Comparison of Figs. 2 and 3 shows that the random dependences of p_0 on time (curves *1*) are close in both these figures, i.e., their average levels $\bar{p}_0$ coincide and the average values of their dispersions are also close to each other (which was verified by direct calculations). This is explained by a sufficiently large averaging interval Δt_{av}, resulting in the smoothing of curve *1* in Fig. 2. This comparison allows one to study the model by using expressions (14)–(16) without generation of the Poisson variable T, which simplifies the analysis. The steady RDM $\rho^{(ss)}$ mentioned in the Introduction, which is obtained without quantum measurements, satisfies, in our notation, the equation

$$\rho^{(ss)} = W_F(Q0 + Q1)\rho^{(ss)}. \quad (18)$$

4. PROPERTIES OF EVOLUTION OPERATORS $S(\xi)$ AND THEIR PRODUCTS: THE PERIODIC TRAJECTORIES

As follows from experiments and numerical simulations, the average relative frequencies of counts $\bar{p}_1$ and $\bar{p}_0$ of an atomic-state-selective detector are maintained in a micromaser during the residence time t_{st}. The values of these rates are determined by the law of alternation of the counting sequence. It is obvious that the same average relative frequency of counts can be obtained at many trajectories, the periodic trajectories being the simplest among them. We denote by $\{\xi_j, 1 \le j \le L\}$ a set of values taken by a random variable ξ over the period L. Consider the features of the micromaser dynamics by choosing the period operator

$$SL = \prod_{\{\xi_j,\, 1 \le j \le L\}} S(\xi_j)$$

as an evolution operator. The vector of the principal diagonal of the RDM is determined at the end of each period. A nonlinear dynamic recurrent relation describing the evolution of the micromaser over a periodic trajectory has, similarly to (16), the form

$$\rho(p+1) = \frac{SL\rho(p)}{\mathrm{Sp}SL\rho(p)}. \quad (19)$$

Here, p is the period number. Unlike (16), Eq. (19) describes the evolution along a specified trajectory. All the operators SL have eigenvalues that are smaller than unity, so that the trace is not conserved [in contrast to the operator in the recurrent relation (18)]. Nevertheless, due to its nonlinearity, relation (19) has a stationary solution. Let us obtain the general solution to (19) by the Fourier method. For this purpose, we expand the initial vector $\rho(0)$ in the eigenvectors of the operator SL:

$$\rho(0) = \sum_\lambda c_\lambda \rho_\lambda. \quad (20)$$

Here, c_λ are the expansion coefficients, and ρ_λ and λ are the eigenvector and the eigenvalue of the operator SL:

$$SL\rho_\lambda = \lambda\rho_\lambda. \quad (21)$$

Let us denote the maximum eigenvalue in expansion (20) by $\tilde{\lambda}$ and obtain the solution of (19) on the period p in the form

$$\rho(p) = \left[\rho_{\tilde{\lambda}} + \sum_{\lambda \ne \tilde{\lambda}} \frac{c_\lambda}{c_{\tilde{\lambda}}}\left(\frac{\lambda}{\tilde{\lambda}}\right)^p \rho_\lambda\right]\left[1 + \sum_{\lambda \ne \tilde{\lambda}} \frac{c_\lambda}{c_{\tilde{\lambda}}}\left(\frac{\lambda}{\tilde{\lambda}}\right)^p\right]^{-1}. \quad (22)$$

In the limit, we have

$$\rho(p) \xrightarrow[p \to \infty]{} \rho_{\tilde{\lambda}}.$$

It follows from (22) that any eigenvector entering the expansion on the initial state and having the largest eigenvalue can be a stationary vector of Eq. (19). As a rule, the initial vector (20) has a nonzero projection on the eigenvector with the maximum possible eigenvalue $\lambda_{\max}$ of the operator SL. It is reasonable to treat this vector as the ground stationary state ρ_{st} for a periodic trajectory (this state was mentioned in the Introduction)

$$\rho_{st} = \rho_{\lambda_{\max}}. \quad (23)$$

Let us elucidate the meaning of the eigenvalue of the period operator SL (21). For this purpose, we rewrite Eq. (19) for any stationary vector ρ in the form

$$\rho = \frac{SL\rho}{\mathrm{Sp}SL\rho} \quad (24)$$

and compare it with (21). One can see that the eigenvalue of the eigenvector normalized according to (4) is

$$\lambda = \mathrm{Sp}SL\rho. \quad (25)$$

But the quantity $\mathrm{Sp}SL\rho$ is equal to the conditional probability of finding L atoms by a detector in the states where a random variable ξ takes values from the set $\{\xi_j, 1 \le j \le L\}$ (under the condition that the initial state of the field was ρ). Let us call for brevity the quantity $\lambda = \mathrm{Sp}SL\rho$ the probability of a periodic trajectory of period L. Therefore, the eigenvalue problem (21) for the product of operators $S(\xi)$ becomes especially important because its solutions give discrete (quantized) probability levels $\mathrm{Sp}SL\rho = \lambda$ and possible stationary states of the field $\rho = \rho_\lambda$ for a periodic trajectory. The stationary states ρ_{st} (23) determine the probabilities of detecting an atom in the lower (7) or upper (8) states at the end of each period

$$\bar{a}_0 = \mathrm{Sp}D(0)\rho_{st}, \tag{26}$$

$$\bar{a}_1 = \mathrm{Sp}D(1)\rho_{st}. \tag{27}$$

These probabilities satisfy the exact relation

$$\frac{\bar{a}_0}{\varepsilon_0} + \frac{\bar{a}_1}{\varepsilon_1} = 1. \tag{28}$$

The average relative frequencies of counts $\bar{p}_1$ and $\bar{p}_0$ of a detector for a periodic trajectory can be calculated by averaging over the trajectory period $L = k + m + r$:

$$\bar{p}_1 = \frac{m}{L}, \tag{29}$$

$$\bar{p}_0 = \frac{k}{L}. \tag{30}$$

Here, k is the number of atoms in the lower state and m is the number of atoms in the upper state detected over the period L. The detector did not find r atoms over the period.

In paper [8], an important property of ergodicity of a random sequence of detector counts was proved (formula (2.36) in [8]). In our notation, this relation has the form

$$\bar{\bar{p}}_0 = \mathrm{Sp}D(0)\rho^{(ss)}. \tag{31}$$

Here, $\bar{\bar{p}}_0$ is the average relative frequency of counts for the lower state of atoms found for an infinite averaging interval. This property is inherent in the steady RDM $\rho^{(ss)}$ (18). Let us show that the ground stationary state ρ_{st} (23), which is typical for a micromaser evolving along a specified periodic trajectory, gives close values for the probabilities (26), (27) and frequencies (29), (30). For this purpose, the parameters k, m, r, ε_1, and ε_0 of the problem should be related by the expression

$$\frac{\bar{p}_0}{\varepsilon_0} + \frac{\bar{p}_1}{\varepsilon_1} = \frac{k}{k+m+r\varepsilon_0}\frac{1}{\varepsilon_0} + \frac{m}{k+m+r\varepsilon_0}\frac{1}{\varepsilon_1} = 1, \tag{32}$$

which is similar to equality (28).

Let us solve the eigenvalue problem (21) for the period operator

$$SL = S^k(0)S^m(1)S^r(2) \tag{33}$$

and find the eigenvectors ρ_λ and the ground stationary state ρ_{st} (23). The operators $S(\xi)$ are defined in (15); analysis is performed for the limiting case $\Delta t \longrightarrow 0$ (17). To solve this problem, we will use the linearization method. Let us perform the orthogonal transformation of displacement of operators $\hat{a}^+$ and $\hat{a}$ in expressions (1)–(3), (15), and (17),

$$\hat{a}^+ \longrightarrow \hat{a}^+ + \mu^+, \quad \hat{a} \longrightarrow \hat{a} + \mu,$$

where μ^+ and μ are numbers, and linearize the operator SL (33), by retaining in it the terms that are linear and quadratic in $\hat{a}^+$, $\hat{a}$. The region of applicability of the linearization method is limited by the condition

$$g\tau < 1; \tag{34}$$

the Filipowicz parameter $\Theta = g\tau\sqrt{N_{ex}}$ can be arbitrary, but $N_{ex} > 1$. As follows from numerical calculations, the formulas of the linearization method can also be used at larger values of $g\tau$ than (34). Therefore, the estimate (34) is rather rough. Let us choose numbers μ^+ and μ so that the terms in SL that are linear over $\hat{a}^+$ and $\hat{a}$ vanish. Let us introduce the notation

$$\nu = \frac{\mu^+\mu}{N_{ex}}$$

and obtain the expressions

$$\chi = \frac{\bar{p}_0}{\varepsilon_0}, \tag{35}$$

$$\begin{aligned} F(\nu) &= (\sin^2(\theta\sqrt{\nu}) - \chi)^2[\nu(\varepsilon_0 - \varepsilon_1) + 1 - \varepsilon_1] \\ &+ (\sin^2(\Theta\sqrt{\nu}) - \chi)\{\nu[\chi(\varepsilon_0 - \varepsilon_1) - 1 + \varepsilon_1\varepsilon_0] \\ &+ (1 - \varepsilon_1)[\chi(2 - \varepsilon_0) - 1]\} \\ &+ (1 - \varepsilon_0)(1 - \varepsilon_1)(1 - \chi)(\nu - \chi), \end{aligned} \tag{36}$$

$$\begin{aligned} G(\nu) &= \left\{\frac{\nu(1 + \sin^2(\Theta\sqrt{\nu}))^2}{(1+\nu)^2}\varepsilon_0\right\}^k \left\{\frac{1 - \sin^4(\Theta\sqrt{\nu})}{1+\nu}\varepsilon_1\right\}^m \\ &\times \left\{\frac{\sin^2(\Theta\sqrt{\nu})[\nu(\varepsilon_1 - \varepsilon_0) + \varepsilon_1 - 1] + \nu(2 - \varepsilon_1 - \varepsilon_0) + 1 - \varepsilon_1}{(1+\nu)^2}(1 + \sin^2(\Theta\sqrt{\nu}))\right\}^r. \end{aligned} \tag{37}$$

Expressions (36) and (37) have been obtained by performing simple but cumbersome algebraic transformations, so that their derivation is not presented here. We present simplified expressions for the functions $F(\nu)$ and $G(\nu)$, which are valid for small $n_b \ll 1$ and large $\Theta \gg 1$.

The zeroes of the function $F(\nu)$ determine the value of the parameter ν at which the terms in *SL* that are linear over $\hat{a}^+$ and $\hat{a}$ vanish. The function $F(\nu)$ has a set of zeroes. Each of the zeroes corresponds to a series of almost equidistant eigenvalues, the distance between them being dependent on the series number. Each of the series has the ground-state eigenvector. This vector is localized in the Fock state, the position of its maximum is determined by the position of the corresponding zero of the function $F(\nu)$, and the corresponding eigenvalue is determined by the value of the function $G(\nu)$ at the zero chosen. The functions $F(\nu)$ and $G(\nu)$ are characterized by an important relation: positions of the maxima of the function $G(\nu)$ coincide with zeroes of the function $F(\nu)$. Analysis of quadratic terms shows that the vectors located in the so-called capture zones [18] prove to be stable. The positions of these zones on the axis of numbers ν is specified by the double inequality

$$\left(\frac{\pi}{2}\frac{1+2q}{\Theta}\right)^2 \le \nu \le \left(\frac{\pi}{2}\frac{2+2q}{\Theta}\right)^2, \quad q = 0, 1, \ldots. \quad (38)$$

Here, the integer q is the number of the capture zone. Note that the functions $F(\nu)$ and $G(\nu)$ are independent of the order of sequence of factors in the period operator *SL* (33) and determine the properties of a beam consisting of $(r+m+k)!/r!m!k!$ periodic trajectories of the same period L. The coefficients of the quadratic form over operators $\hat{a}^+$ and $\hat{a}$ depend on the interchange of the factors. Of most interest is a stable state ρ_{st} (23) corresponding to the maximum eigenvalue. When the parameters of the operator *SL* (33) are related by expression (32), the ground state is localized in the capture zone (38) in the interval $0 \le \nu \le 1$. An important property of capture zones is that their position (38) is independent of the numbers k, m, and r and the detector efficiency.

Consider a simplified case of coinciding efficiencies of the detector:

$$\varepsilon_1 = \varepsilon_0 = \varepsilon.$$

Expressions (32), (35)–(37) are simplified and take the form

$$\varepsilon = \frac{k+m}{k+m+r}, \quad (39)$$

$$\chi = \frac{\bar{p}_0}{\varepsilon} = \frac{k}{k+m}, \quad (40)$$

$$F(\nu) = (\sin^2(\Theta\sqrt{\nu}) - \chi)^2 + (\sin^2(\Theta\sqrt{\nu}) - \chi) \times [\chi(2-\varepsilon) - 1 - \nu(\varepsilon+1)] + (1-\varepsilon)(1-\chi)(\nu-\chi), \quad (41)$$

$$G(\nu) = \left[\frac{\nu(1+\sin^2(\Theta\sqrt{\nu}))^2}{(1+\nu)^2}\right]^k \times \left[\frac{1-\sin^4(\Theta\sqrt{\nu})}{1+\nu}\right]^m \times \left[\frac{2\nu+\cos^2(\Theta\sqrt{\nu})}{(1+\nu)^2}(1+\sin^2\Theta\sqrt{\nu})\right]^r \varepsilon^{k+m}(1-\varepsilon)^r. \quad (42)$$

Let us derive the known result of Filipowicz from expression (41). To do this, we set $k = 0$, $m = 0$, $\chi \ne 0$ in (41). Then, we obtain from the condition $F(\nu) = 0$ the equation

$$\sin^2(\Theta\sqrt{\nu}) = \nu, \quad (43)$$

which determines the positions of minima of the effective Filipowicz potential. This result becomes clear if we note that, for $k = 0$, $m = 0$, and $\chi \ne 0$, the evolution operator *SL* (33) becomes a multiple of the operator $W_F(Q0 + Q1)$ (18). The vector $\rho^{(ss)}$ is the eigenvector of this operator, with the eigenvalue equal to 1. It is for this operator that the result (43) was obtained in paper [2].

To justify relation (32) chosen above, note that in this case the equation $F(\nu) = 0$ always has the root ν satisfying two approximate conditions

$$\sin^2(\Theta\sqrt{\nu}) \approx \chi, \quad \nu \approx \chi. \quad (44)$$

It is the required root, which determines the stable state ρ_{st} corresponding to the maximum eigenvalue of the operator *SL*. The meaning of approximate equalities (44) is as follows. Because the ground state represents a strongly localized vector, then, according to expressions (26), (2), and (5), we have in the linearization approximation

$$\bar{a}_0 = \mathrm{Sp}D(0)\rho_{st} \approx \varepsilon\sin^2(\Theta\sqrt{\nu}). \quad (45)$$

Then, according to (40), we obtain the relation between the probabilities $\bar{a}_0$ (26), $\bar{a}_1$ (27) of the detector count (at the end of each period) and the average relative frequencies of counts $\bar{p}_0$ (29), $\bar{p}_1$ (30) of the detector [analogue of equality (31)]

$$\bar{p}_0 \approx \bar{a}_0, \quad \bar{p}_1 \approx \bar{a}_1. \quad (46)$$

The second approximate equality (44) determines the position of the maximum probability of photon numbers for the ground state ρ_{st} as a function of the Fock number n:

$$n = \nu N_{ex} \approx \frac{\bar{a}_0}{\varepsilon} N_{ex}. \quad (47)$$

The ratio $\bar{a}_0/\varepsilon$ is the quantum-mechanical probability of finding an atom in the lower state at the end of each period of a periodic trajectory.

Let us find the values of the parameter Θ for which approximate equalities (44) are replaced by exact ones. For this purpose, we rewrite the equation $F(\nu) = 0$ in the form that is similar to the Filipowicz equation (43):

$$\sin^2(\Theta\sqrt{\nu}) = SN(\nu),$$

$$SN(\nu) = \frac{1}{2}\{\nu(\varepsilon+1) + 1 + \chi\varepsilon - \sqrt{[\nu(\varepsilon+1)+1+\chi\varepsilon]^2 - 4[\nu(1-\varepsilon+2\varepsilon\chi)+\varepsilon\chi]}\}.$$

By substituting $SN(\nu)$ into expression (42), we obtain the equation for the envelope $V(\nu)$ of the maxima of the function $G(\nu)$. We do not present here the expression for the function $V(\nu)$ because it is too cumbersome. The plot of this function is shown in Fig. 4. One can see from Fig. 4 that the function $V(\nu)$ has one maximum at the value of ν for which two exact equalities

$$\nu = \chi, \quad SN(\nu) = \chi \tag{48}$$

are satisfied. For the values of the parameter

$$\Theta = -\frac{\arcsin\sqrt{\chi} + j\pi}{\sqrt{\chi}}, \quad j = 1, 2, \ldots, \tag{49}$$

expressions (48) give an exact value of the argument ν at which the maximum of the envelope coincides with the main maximum of the function $G(\nu)$. By substituting (48) and (49) into $G(\nu)$ (42), we obtain the maximum value of the function $G(\nu)$—the maximum eigenvalue of the period operator (33) or the maximum probability of a periodic trajectory with period $L = k + m + r$:

$$\lambda_{\max} \approx G_{\max} = \bar{p}_0^k \bar{p}_1^m (1 - \bar{p}_0 - \bar{p}_1)^{L-k-m}. \tag{50}$$

Here, $\bar{p}_0$ and $\bar{p}_1$ are determined by expressions (29) and (30). By multiplying (50) by the statistical weight, we obtain the expression for the total probability of the beam of trajectories with period L:

$$\mathrm{Prob}(k, m, L) = \bar{p}_0^k \bar{p}_1^m (1 - \bar{p}_0 - \bar{p}_1)^{L-k-m} \times \frac{L!}{k!m!(L-k-m)!}, \tag{51}$$

which represents a two-dimensional binomial distribution. This expression generalizes [if approximate equality (46) is used] expression (3) from paper [5] to the case of an imperfect detector. In paper [5], an expression similar to (51) was obtained under the condition that the field mode is found in the state $\rho^{(ss)}$. Here, expression (51) is obtained for a periodic trajectory when the field mode is in the state ρ_{st} (23).

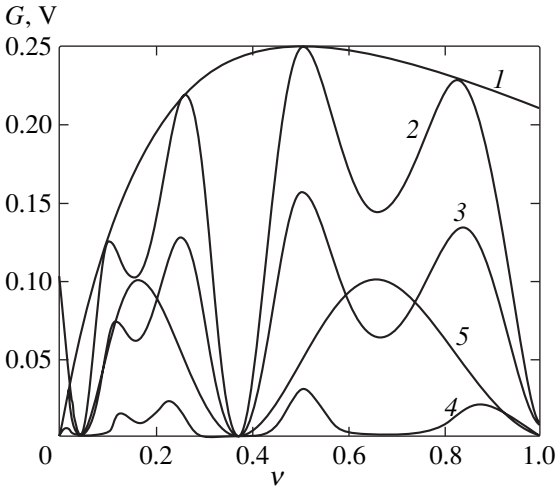


Fig. 4. Function $G(\nu)$ and the envelope $V(\nu)$ for parameters $k = 1$, $m = 1$, $\chi = 0.5$: (*1*) $V(\nu)$ for $r = 0$, $\varepsilon = 1$; (*2*) $G(\nu)$ for $r = 0$, $\varepsilon = 1$; (*3*) $G(\nu) \times 10$ for $r = 2$, $\varepsilon = 0.5$; (*4*) $G(\nu) \times 10^2$ for $r = 20$, $\varepsilon = 0.091$; (*5*) $\cos^2(\Theta\sqrt{\nu})/10$ is plotted to show capture zones. Other calculation parameters are as in Fig. 2.

5. DISCUSSION

As follows from experiments [4], a sequence of detector counts—the trajectory ξ_l—is a random function of the number l of the flown atom. By using statistical processing, one can obtain the relative frequencies of finding an atom in the upper (lower) state, as well as the relative frequency of transmission of an atom by an operating detector at a specified interval of time (counting statistics). In the experimental paper [4], these quantities were calculated in each interval Δt_{av} along a random trajectory. It was found that random realizations had a characteristic property: under stationary conditions, the frequencies of counts $\bar{p}_0$ and $\bar{p}_1$ of the detector averaged over the residence time t_{st}, which indicate the detection of the upper or lower atomic state, no longer depended on time.

The aim of this work was to solve the inverse problem, namely, using the average relative frequencies of counts $\bar{p}_0$ and $\bar{p}_1$ measured in experiments in the specified time interval to derive and solve the dynamic (rather than stochastic) equation for the field RDM. It is obvious that this evolution equation will be dynamic if the trajectory of detector counts is specified (fixed). For this purpose, we used a periodic trajectory in this paper. We showed in Section 3 how the randomness related to the Poisson statistics of an atomic beam can be avoided. To do this, we used the method proposed in paper [2] and replaced SRR (6) by SRR (16) averaged over the distribution (12) of expectation times. The recurrent relation (16) corresponds to a periodic atomic beam with the same time intervals between the atoms, which are equal to 1/R. It was noted in Section 3 that, upon passage from SRR (6) to SRR (16), the statistical prop-

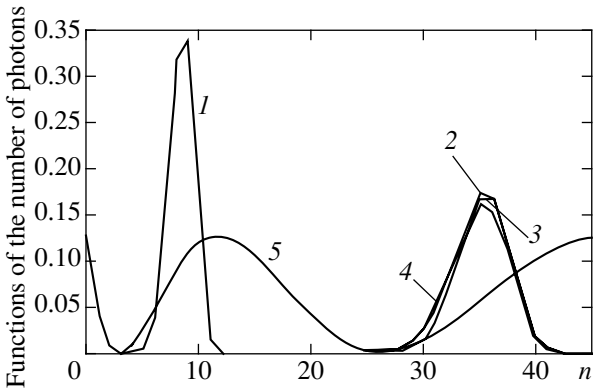


Fig. 5. Eigenvectors for the maximum eigenvalue of the period operator $SL = S^k(0)S^m(1)S^r(2)$ as functions of the number n of photons: (*1*) ρ_{st} for $k = 1$, $m = 4$, $r = 0$; (*2*) ρ_{st} for $k = 1$, $m = 1$, $r = 0$; (*3*) ρ_{st} for $k = 1$, $m = 1$, $r = 20$; (*4*) $\rho^{(ss)}$; (*5*) $\cos^2(g\tau\sqrt{n})/8$ is plotted to show capture zones. Calculation parameters are as in Fig. 2.

erties of random frequencies of counts p_0 and p_1 were conserved. To derive the dynamic recurrent relation (19) for determining the stationary vector of the principal diagonal of the RDM ρ_{st}, it is necessary to choose an evolution matrix for a period. The parameters of the period matrix SL (33)—the integers k, m, and L—are chosen by using the experimental values of average relative frequencies of counts $\bar{p}_0$ and $\bar{p}_1$ and expressions (29) and (30). Below, we present a numerical example of application of this procedure. The statistical processing of a random sequence, which was used to plot the functions in Fig. 2, gives the results

$$\bar{p}_0 = \frac{k}{k+m+r} \approx 0.162, \quad \bar{p}_1 = \frac{m}{k+m+r} \approx 0.163.$$

This result agrees with expression (32):

$$\bar{p}_0 + \bar{p}_1 \approx \varepsilon = \frac{1}{3}.$$

To obtain the minimum period, we can choose the values

$$k = 1, \quad m = 1, \quad \chi = \frac{k}{k+m} = \frac{1}{2}.$$

Then, the period is

$$L = \frac{k}{\bar{p}_0} \approx 6.$$

The parameters of a periodic trajectory are

$$k = 1, \quad m = 1, \quad r = 4.$$

We should substitute these parameters into Eq. (33), solve the eigenvalue problem (21), and find the vector of the principal diagonal of the RDM ρ_{st} for the maximum eigenvalue λ_{max}. The vector ρ_{st} (23) is the required solution to the inverse problem. It can be used for calculating the statistical characteristics of a field mode of a micromaser subjected to quantum measurements of the states of the atomic subsystem under stationary conditions. Figure 5 shows the plots of the principal-diagonal vector of the RDM ρ_{st} for $k = 1$, $m = 1, 4$, and $r = 0, 20$. The curves are plotted for the value of Θ satisfying expression (49) for $j = 2$. The vector ρ_{st} represents a single peak localized in the capture zone (38) with the number $q = 1$.

The obtained vector ρ_{st} describes a subensemble of the states of a field mode, which is determined by the average relative frequency of counts $\bar{p}_0$ ($\bar{p}_1$) detected during the residence time t_{st}. The localization region of this vector is the capture zone with the number $q = 1$. The total ensemble of the field states under stationary conditions is described by the vector $\rho^{(ss)}$ (18). The plot of the vector $\rho^{(ss)}$ for the same value of Θ is also shown in Fig. 5. According to the Filipowicz theory [2], in this case, the vector $\rho^{(ss)}$ represents a single peak localized in the zone with $q = 1$. In other words, the total ensemble of the field states under these conditions is determined by one average value of the relative frequency of counts $\bar{p}_0$ ($\bar{p}_1$). In this case, the plots of $\rho^{(ss)}$ and ρ_{st} should virtually coincide, which is demonstrated in Fig. 5. It follows from Fig. 4 that it does not matter which of the detectors, with low or high efficiency, was used for determining the parameters of a periodic trajectory in experiments. Indeed, the position of the maxima of the function $G(\nu)$ in Fig. 4 weakly depends on the detector efficiency. The same property is demonstrated in Fig. 5, where the plots of ρ_{st} are shown for several values of r [different values of ε (39)]. By using a highly efficient detector, we can detect trajectories with short residence times t_{st} in the stationary state.

Note in conclusion that the results obtained in this paper are valid in the linearization approximation. This approximation can be used when capture zones (38) can be considered separately from each other. The eigenvectors (21) of the problem for the operator (33) localized in different zones are weakly overlapped. This is the case when the probability of the trajectory segment weakly depends on the sequence of counts and is determined by the number of counts. Small variations in the sequence of the counts, conserving the statistical characteristics of the realization ξ_j, only weakly affect the principal-diagonal vector of the field RDM. In this case, the use of the corresponding periodic trajectory for solving the inverse problem is justified.

A more complicated dynamics of a micromaser during quantum measurements is observed when the conditions of applicability of the method are violated. In this case, capture zones (especially, with smaller numbers) lose their individuality. Under stationary conditions, two (or several) competing average relative frequencies of counts of the detector can exist. The vector $\rho^{(ss)}$ can have several peaks located in different capture

zones (38). The trajectory becomes unstable, and sharp jumps can occur between the average levels of the relative frequencies of counts. These jumps occur during a time interval that is much shorter than the residence time t_{st}. This instability can be caused by an unlikely event—a critical fluctuation, when a random variable determining the dynamics of the micromaser (the Poisson variable T or a variable ξ simulating the order of the detector operation) takes unlikely values in succession. This case requires a separate study.

REFERENCES

1. J. M. Raimond, M. Brune, and S. Haroche, Rev. Mod. Phys. **73**, 565 (2001).
2. P. Filipowicz, J. Javanainen, and P. Meystre, Phys. Rev. A **34**, 3077 (1986).
3. P. Meystre and E. M. Wright, Phys. Rev. A **37**, 2524 (1988).
4. O. Benson, G. Raithel, and H. Walther, Phys. Rev. Lett. **72**, 3506 (1994).
5. G. Rempe and H. Walter, Phys. Rev. A **42**, 1650 (1990).
6. C. Wagner, R. J. Brecha, A. Schenzle, and H. Walter, Phys. Rev. A **46**, R5350 (1992); **47**, 5068 (1993).
7. L. A. Lugiato, M. O. Scully, and H. Walther, Phys. Rev. A **36**, 740 (1987).
8. H.-J. Briegel, B.-G. Englert, N. Sterpi, and H. Walter, Phys. Rev. A **49**, 2962 (1994).
9. U. Herzog, Phys. Rev. A **50**, 783 (1994).
10. C. Wagner, A. Schenzle, and H. Walter, Opt. Commun. **107**, 318 (1994).
11. U. Herzog, Phys. Rev. A **52**, 602 (1995).
12. U. Herzog, Phys. Rev. A **61**, 047801 (2000).
13. B.-G. Englert, T. Gantsog, A. Schenzle, *et al.*, Phys. Rev. A **53**, 4386 (1996).
14. D. B. Johnson and W. C. Schieve, Phys. Rev. A **63**, 033808 (2001).
15. R. R. McGowan and W. C. Schieve, Phys. Rev. A **59**, 778 (1999).
16. J. von Neumann, *Mathematische Grudlagen der Quantenmechanik* (Springer-Verlag, Berlin, 1932).
17. E. T. Jaynes and F. W. Cummings, Proc. IEEE **51**, 89 (1963).
18. I. P. Vadeĭko, G. P. Miroshnichenko, A. V. Rybin, and J. Timonen, Opt. Spektrosk. **89**, 328 (2000) [Opt. Spectrosc. **89**, 300 (2000)]; G. P. Miroshnichenko, I. P. Vadeĭko, A. V. Rybin, and J. Timonen, Pis'ma Zh. Éksp. Teor. Fiz. **72**, 647 (2000) [JETP Lett. **72**, 449 (2000)].

Translated by M. Sapozhnikov

Journal of Experimental and Theoretical Physics, Vol. 95, No. 5, 2002, pp. 844–850.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 122, No. 5, 2002, pp. 978–985.

ATOMS, SPECTRA, RADIATION

The Ionization of Atoms in an Intense Nonclassical Electromagnetic Field

A. M. Popov* and O. V. Tikhonova

Skobel'tsyn Research Institute of Nuclear Physics, Moscow State University, Vorob'evy gory, Moscow, 119992 Russia
**e-mail: popov@mics.msu.su*

Received April 12, 2002

Abstract—The ionization of atoms in a strong nonclassical unimodal electromagnetic field was considered. It was shown that the probability of ionization could substantially change depending on the quantum state of the field even at a constant mean number of quanta in the radiation mode. The difference of ionization rates was especially large for multiphoton ionization processes. It was, in particular, shown that a nonclassical field could be much more effective from the point of view of the ionization of atoms than a classical field of the same intensity. The characteristics of the decay of a bound atomic system state in a strong nonclassical field were studied without invoking perturbation theory. *© 2002 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

The introduction of the theory of "squeezed" electromagnetic field states [1] followed by their experimental observation [2, 3] actually opened a new chapter in the physics of the interaction of high-power electromagnetic radiation with matter [4]. In practice, we deal with a new macroscopic quantum effect, that is, a purely quantum (nonclassical) electromagnetic field state, nevertheless containing a large (huge) number of quanta [5]. Although the degrees of laser radiation squeezing currently attainable are not large, it is likely that there are no physical limitations that can, in the future, prevent the generation of high-intensity fields whose properties are very far from those of classical electromagnetic fields [5]. This means that the semiclassical theory (quantum medium and classical electromagnetic field) that has earlier been successfully used to describe interactions between high-power electromagnetic radiation and matter (for instance, in the dynamics of laser generation [6, 7], in various nonlinear optics problems [7, 8], and in studying elementary events in intense light fields [9, 10]) should be replaced by a consistent quantum theory, in which both the atomic (molecular) subsystem and electromagnetic field are described quantum mechanically. The new theory should, as a particular case, include the old semiclassical theory, when the state of an electromagnetic field as a quantum object can be described in the classical limit. Such a consistent quantum approach can naturally be expected to reveal new effects inexplicable in terms of the traditional semiclassical approach.

In this work, we consider the dynamics of multiphoton ionization of an atom in a strong nonclassical unimodal electromagnetic field. Namely, we study the action on an atom of an electromagnetic field in the Fock and coherent states and in the squeezed vacuum state. Depending on the quantum state of the field, the probabilities of ionization can be different even when the number of quanta in the radiation mode remains unchanged. The difference of ionization rates is especially large for multiphoton ionization. In particular, a nonclassical field can be much more effective from the point of view of the ionization of atoms than a classical field of the same intensity. The special features of the decay of a bound atomic system state in a strong nonclassical field were studied outside the framework of perturbation theory.

2. QUANTUM ELECTROMAGNETIC FIELD

Further, we only consider one field mode with frequency ω. This field can conveniently be described by a pair of conjugate values, namely, a (the vector potential) and $\varepsilon = -(da/dt)/c$ (the electric field strength). Both these values satisfy the harmonic oscillator equation.

In quantum theory, Hamilton operator $H_f(\varepsilon)$ can be put in correspondence to a field mode. This is the Hamiltonian of the one-dimensional harmonic oscillator. Clearly, the field mode is characterized by a set of stationary states with the energies

$$E_k = \hbar\omega(k + 1/2), \tag{1}$$

where k has the meaning of the number of field quanta (photons) in the stationary (Fock) state $|k\rangle$. The ground field state, or electromagnetic vacuum, is the state with a zero number of quanta, $k = 0$. The probability density distribution for observing field strength ε in measurements is then given by the equation

$$\rho_0(\varepsilon) = \frac{1}{\varepsilon_0\sqrt{\pi}}\exp\left(-\frac{\varepsilon^2}{\varepsilon_0^2}\right), \tag{2}$$

1063-7761/02/9505-0844$22.00

where $\varepsilon_0 = \sqrt{4\hbar\omega^4/\pi^2 c^3}$ determines electric field uncertainty in the vacuum state.[1] The probability density distribution in an arbitrary stationary state $|k\rangle$ is written as

$$\rho_k(\varepsilon) = \frac{1}{2^k k!\sqrt{\pi}} H_k^2(\varepsilon)\exp(-\varepsilon^2) \quad (3)$$

(here, H_k is the Hermitian polynomial).

Note that an electromagnetic field in an arbitrary stationary state is a purely quantum object. Although field energy in a stationary state can be very high, the mean field strength ε averaged over the quantum state is zero. Therefore, the mean force acting on an electric charge is also zero. From the point of view of quantum theory, a classical electromagnetic field is the coherent state of a field oscillator with a large mean number of quanta [11]. This state can be written in the energy representation as

$$\psi_c(\varepsilon) = \sum_k \alpha_k |k\rangle, \quad (4)$$

where the amplitudes in the expansion over stationary field states are given by

$$\alpha_k = \frac{z^k}{\sqrt{k!}}\exp\left(-\frac{|z|^2}{2}\right), \quad (5)$$

where z is a complex number. Clearly, the squares of the moduli of the amplitudes $|\alpha_k|^2$ are determined by the Poisson distribution, and the Poisson distribution parameter $\langle k\rangle = |z|^2$ has the meaning of the mean number of field quanta in state (4). It is easy to show that the mean electric field strength in state (4) varies with time by the harmonic law[2]

$$\langle\varepsilon\rangle = \sqrt{2\langle k\rangle}\cos(\omega t),$$

and the quantum field strength uncertainty is

$$\Delta\varepsilon = \sqrt{D_\varepsilon} = 1/\sqrt{2}$$

(here, D_ε is the electric field variance). If $\langle k\rangle \gg 1$, the quantum uncertainty in ε can be ignored because $\Delta\varepsilon \ll \langle\varepsilon\rangle$, and we arrive at the conclusion that changes in the electric field as a function of time are determined by the solution to the classical equation of motion. The field energy and radiation intensity are then proportional to the mean number of quanta in the mode $\langle k\rangle$ and are related to field strength $\langle\varepsilon\rangle$ by classical electrodynamics equations.

Clearly, the classical field state is but a particular case of a strong electromagnetic field (field with the number of quanta $\langle k\rangle \gg 1$). The number of strong electromagnetic field realizations that cannot be described classically is infinite. In recent years, so-called squeezed states have usually been considered. In the coordinate representation, these states are described by wave functions of the form [12]

$$\psi(\varepsilon) = \frac{1}{\sqrt{\beta\sqrt{\pi}}}\exp(ia\varepsilon)\exp\left(-\frac{\varepsilon^2}{2\beta^2}\right), \quad (6)$$

which depend on two parameters, α and β. Clearly, if $\alpha = 0$ and $\beta = 1$, we have an electromagnetic vacuum, and if $a \neq 0$ and $\beta = 1$, the field is in the coherent state [12]. The special feature of the state described by wave function (6) is as follows: if $\beta \ll 1$, electric field variance $D_\varepsilon = \beta^2/2$ is substantially smaller than in the vacuum and coherent states. Conversely, for the coordinate conjugate to the electric field, we have

$$D_a = \frac{1}{2\beta^2} \gg \frac{1}{2}.$$

The uncertainty relation

$$D_\varepsilon D_a = 1/4$$

is, of course, satisfied. During the time evolution of state (6), the variances of the a and ε coordinates oscillate at a twice the electric field frequency. As a result, the state with

$$D_\varepsilon = 1/2\beta^2, \quad D_a = \beta^2/2$$

is formed in half the optical cycle. In this sense, squeezed states with the β and $1/\beta$ parameters are physically equivalent. Further, the $K = \beta$ (if $\beta > 1$) or $K = 1/\beta$ (if $\beta < 1$) parameter will be called the squeezing parameter of electromagnetic field states.

Among the squeezed electromagnetic field states, of special interest is the squeezed vacuum state. This state is described by (6) with $\alpha = 0$ and $\beta \neq 1$; that is, it only differs from the electromagnetic vacuum state in the width of the distribution of ε. Although the mean field strength in this state is also zero at an arbitrary instant of time, the electromagnetic field energy in the squeezed vacuum state proves to be high if $\beta \ll 1$ or $\beta \gg 1$. The expansion of the squeezed vacuum state $\psi_{sq}(\varepsilon)$ in stationary field states gives

$$\psi_{sq}(\varepsilon) = \sum_k \alpha_{2k}|2k\rangle, \quad (7)$$

where the expansion coefficients α_{2k} are determined by the equation (see Appendix)

$$\alpha_{2k} = (-1)^k\sqrt{\frac{2\beta}{1+\beta^2}}\frac{\sqrt{(2k)!}}{2^k k!}\left(\frac{1-\beta^2}{1+\beta^2}\right)^k, \quad (8)$$

and all odd coefficients are zero.

Let us use wave function (6) to calculate the mean field energy over a quantum state,

$$\langle E\rangle = \frac{\hbar\omega}{2}\left(\frac{\beta^2}{2} + \frac{1}{2\beta^2} + a^2\right), \quad (9)$$

[1] We further use the dimensionless field coordinate $\varepsilon/\varepsilon_0$.

[2] This expression corresponds to a real parameter z value.

and introduce the mean number of field quanta in an arbitrary field state by the equation

$$\langle E\rangle = \hbar\omega(\langle k\rangle + 1/2). \quad (10)$$

Equations (9) and (10) determine the relation between the squeezing parameter and the mean number of quanta,

$$\langle k\rangle = \frac{1}{4}\left(\beta - \frac{1}{\beta}\right)^2 + \frac{a^2}{2}. \quad (11)$$

In particular, for the squeezed vacuum state, we have

$$\langle k\rangle = \frac{1}{4}\left(\beta - \frac{1}{\beta}\right)^2 \longrightarrow \frac{1}{4\beta^2} \quad (12)$$

at $\beta \ll 1$.

In practice, precisely the squeezed vacuum state is of the greatest interest, on the one hand, because the variance of the number of quanta (at a given mean number of quanta) is then maximum. We will show later on that this circumstance has an important influence on the consequences of the action of such a field on an atomic system. On the other hand, modern experiments on the generation of nonclassical electromagnetic fields in parametric amplification processes [13] are capable of producing states close to precisely the squeezed vacuum state. Although the degree of squeezing currently attainable in such experiments is comparatively small, $K = 1/\beta \approx 10$, we see no reason why squeezing cannot be increased to $K \sim 10^8$–10^9.

Below, we consider the result of the action of an electromagnetic field in the stationary, coherent, and squeezed vacuum states on an atomic system under the conditions when the mean numbers of quanta in the field mode are equal in these states.

3. ATOM AND ITS INTERACTION WITH ELECTROMAGNETIC FIELDS

Let an atom be characterized by a set of stationary states of a discrete spectrum ($|n\rangle$, $n = 1, 2, 3, \ldots$) and continuum $|E\rangle$, and let these states satisfy the equations

$$H_0|n\rangle = E_n|n\rangle, \quad H_0|E\rangle = E|E\rangle,$$

where H_0 is the atomic Hamiltonian. We assume that, at the initial instant of time, the atom is in the ground state $|1\rangle$ and has the ionization potential $I = -E_1$.

The complete Hamiltonian of the "atom + electromagnetic field" system can be written as

$$H(\mathbf{r}, \varepsilon) = H_0(\mathbf{r}) + H_f(\varepsilon) + V(\mathbf{r}, \varepsilon), \quad (13)$$

where $\mathbf{r}$ is the set of atomic subsystem coordinates and $V(\mathbf{r}, \varepsilon)$ is the interaction operator between the atomic subsystem and the electromagnetic field. In the dipole approximation and in the $d\varepsilon$ gauge, this operator is written in the form

$$V = -ez\varepsilon. \quad (14)$$

Here, it is assumed that the z axis is directed along the ε vector. It is important for the further analysis that the interaction between an atom and an electromagnetic field is weak. Precisely this circumstance allows us to admit the independent existence of atomic and field subsystems as a zeroth approximation and to take into account their interaction by means of perturbation theory. Physically, this means that the evolution of the "atom + field" system should be interpreted in the basis of stationary states that describe subsystems without taking into account their interaction. We will therefore seek the total wave function of the system $\Psi(\mathbf{r}, \varepsilon, t)$ in the form

$$\Psi(\mathbf{r}, \varepsilon, t) = \sum_{n,k} C_{n,k}(t)|n\rangle|k\rangle\exp\left(-\frac{i}{\hbar}(E_n + E_k)t\right) + \sum_k \int dE\, C_{E,k}(t)|E\rangle|k\rangle\exp\left(-\frac{i}{\hbar}(E + E_k)t\right), \quad (15)$$

where E_k is given by (1), and the $C_{n,k}$ and $C_{E,k}$ expansion coefficients determine the amplitudes of the probabilities of finding atoms in the $|n\rangle(|E\rangle)$ state and the electromagnetic field in the $|k\rangle$ state. Substituting expansion (15) into the Schrödinger equation yields the system of equations for the $C_{n,k}$ and $C_{E,k}$ amplitudes

$$i\hbar\frac{dC_{n,k}}{dt} = -\frac{\varepsilon_0}{\sqrt{2}}\sum_{n'} d_{nn'}\exp\left(\frac{i}{\hbar}(E_n - E_{n'})t\right) \times \{C_{n',k+1}\sqrt{k+1}\exp(-i\omega t) + C_{n',k-1}\sqrt{k}\exp(i\omega t)\} - \frac{\varepsilon_0}{\sqrt{2}}\int dE' d_{nE'}\exp\left(\frac{i}{\hbar}(E_n - E')t\right) \times \{C_{E',k+1}\sqrt{k+1}\exp(-i\omega t) + C_{E',k-1}\sqrt{k}\exp(i\omega t)\}, \quad (16)$$

where $d_{nn'} = \langle n|ez|n'\rangle$ and $d_{nE'} = \langle n|ez|E'\rangle$ are the matrix elements of the dipole operator, which acts in the space of atomic states. The equation for the $C_{E,k}$ amplitude is obtained from (16) by replacing state $|n\rangle$ with $|E\rangle$. Equation (16) was derived taking into account that the matrix element of the field operator is only nonzero if $k = k \pm 1$ and that

$$\langle k+1|\varepsilon|k\rangle = \frac{1}{\sqrt{2}}\sqrt{k+1}.$$

Clearly, the photoionization of an atom is the transition from the initial $\Phi_i(\mathbf{r}, \varepsilon) = |1\rangle|\psi_i\rangle$ state to the final $\Phi_f = |E\rangle|\psi_f\rangle$ state under excitation $V(z, \varepsilon)$. Here, $|\psi_i\rangle$ and $|\psi_f\rangle$ are the initial and final electromagnetic field states.

4. SINGLE-PHOTON IONIZATION OF AN ATOM IN A NONCLASSICAL ELECTROMAGNETIC FIELD

Let the initial atomic state $|1\rangle$ be related to the continuum by a single-photon transition. The probability of the photoionization of the atom can then be calculated in first-order perturbation theory. Solving (16) then yields the equation for the probability of ionization in unit time

$$\frac{d\dot{W}_i^{(1)}}{dE} = \frac{2\pi}{\hbar}\sum_k |\alpha_k|^2 k |d_{EI}|^2 \frac{\varepsilon_0^2}{2}\delta(E + I - \hbar\omega), \quad (17)$$

where α_k are the amplitudes of the expansion of the initial field state ψ_i in the basis of stationary states, $\psi_i = \sum_k \alpha_k |k\rangle$, and the δ function determines the energy of the final atomic state.

As

$$\sum_k |\alpha_k|^2 k = \langle k\rangle \quad (18)$$

is the mean number of quanta in field state $|\psi_i\rangle$, we find from (17) that the probability of ionization is independent of the specific form of the distribution of $|\alpha_k|^2$ and is only determined by the mean number of quanta in the field mode. This means that the single-photon ionization of an atomic system in a nonclassical electromagnetic field has no special features within the limits of first-order perturbation theory applicability.

5. MULTIPHOTON IONIZATION

If the ionization of an atom can occur as a multiphoton process, solving (16) at the level of Nth-order perturbation theory yields

$$\frac{d\dot{W}_i^{(N)}}{dE} = \frac{2\pi}{\hbar}\left|d_{EI}^{(N)}\right|^2\left(\frac{\varepsilon_0^2}{2}\right)^N \times \sum_{k=0}^{\infty} |\alpha_k|^2 k(k-1)\ldots(k-N+1)\delta(E + I - N\hbar\omega), \quad (19)$$

where $d_{EI}^{(N)}$ is the multiphoton matrix element. In particular, for $N = 2$, we have

$$d_{EI}^{(2)} = \sum_{n'}\frac{d_{En'}d_{n'1}}{E_{n'} - E_1 - \hbar\omega} + \int\frac{d_{EE'}d_{E'1}}{E' - E_1 - \hbar\omega}dE'.$$

Equation (19) shows that the probability of photoionization is determined by the specific electromagnetic field quantum state. For instance, if the field is in the coherent state, the summation in (19) yields

$$\frac{d\dot{W}_i^{(N)}}{dE} = \frac{2\pi}{\hbar}\left|d_{EI}^{(N)}\right|^2 \times\left(\frac{\varepsilon_0^2}{2}\right)^N \langle k\rangle^N \delta(E + I - N\hbar\omega), \quad (20)$$

where $\langle k\rangle$ is the mean number of field quanta determined by (18). As radiation intensity $P \sim \varepsilon_0^2\langle k\rangle$, (20) can be obtained in the semiclassical approach by applying Nth-order perturbation theory.[3]

The situation is different for a non-Poisson distribution of $|\alpha_k|^2$. For instance, we find from (19) for the stationary state of a field with k_0 quanta that photoionization is only possible if $k_0 \geq N$ and that, at $k_0 = N$, the ionization leads to the vacuum field state. In practice, of greatest interest is the situation with $k_0 \gg N$. Equation (19) then yields

$$\frac{d\dot{W}_i^{(N)}}{dE} \approx \frac{2\pi}{\hbar}\left|d_{EI}^{(N)}\right|^2\left(\frac{\varepsilon_0^2}{2}\right)^N \times k_0^N\left(1 - \frac{N(N-1)}{k_0}\right)\delta(E + I - N\hbar\omega); \quad (21)$$

that is, the probability of photoionization is then somewhat lower than for the coherent state with $\langle k\rangle = k_0$.

Of special interest is electromagnetic field in the squeezed vacuum state. Indeed, number-of-quanta distribution (8) is then characterized by a large width (variance), which increases as the degree of squeezing $K = 1/\beta$ grows. For this reason, the probability of observing a number of quanta substantially larger than the mean value $\langle k\rangle$ is many times larger in a strongly squeezed than in the coherent state. In this situation, the probability of nonlinear N-photon ionization in a nonclassical squeezed field should be much higher than in a classical field, and this difference should increase with increasing multiphoton order of the process. For instance, for $N = 2$, (19) gives

$$\frac{d\dot{W}_i^{(2)}}{dE} \approx \frac{2\pi}{\hbar}\left|d_{EI}^{(2)}\right|^2\left(\frac{\varepsilon_0^2}{2}\right)^2 \sum_{k=0}^{\infty}|\alpha_k|^2 k(k-1) \times \delta(E + I - 2\hbar\omega) \sim \langle k^2\rangle - \langle k\rangle. \quad (22)$$

Here,

$$\langle k^2\rangle = \sum k^2 |\alpha_k|^2$$

is the mean square of the number of field quanta. As

$$\langle k^2\rangle = \langle k\rangle^2 + D_k$$

[3] Note that an arbitrary field state characterized by the Poisson distribution of $|\alpha_k|^2$ is not necessarily coherent. A certain phase relation between the α_k amplitudes is also required. In calculating the probability of N-photon ionization in the lowest order of perturbation theory, this phase relation is, however, inessential; that is, all field states with equal $|\alpha_k|^2$ values are then equivalent.

(D_k is the variance of the number-of-quanta distribution), we find from (22) that

$$\frac{d\dot{W}_i^{(2)}}{dE} \approx \frac{2\pi}{\hbar}\left|d_{EI}^{(2)}\right|^2\left(\frac{\varepsilon_0^2}{2}\right)^2$$
$$\times(\langle k\rangle^2 + D_k - \langle k\rangle)\delta(E + I - 2\hbar\omega).$$

As mentioned, in real laser beams with $\langle k\rangle \gg 1$, the difference of atom ionization rates observed in stationary versus coherent electromagnetic fields is negligibly small. A different situation arises if the electromagnetic field is in the squeezed vacuum state. We then have

$$D_k = 2(\langle k\rangle^2 + \langle k\rangle),$$

which gives

$$\frac{d\dot{W}_i^{(2)}}{dE} \approx \frac{2\pi}{\hbar}\left|d_{EI}^{(2)}\right|^2\left(\frac{\varepsilon_0^2}{2}\right)^2$$
$$\times(3\langle k\rangle^2 + \langle k\rangle)\delta(E + I - 2\hbar\omega);$$

that is, if $\langle k\rangle \gg 1$, the probability of ionization is three times higher than in semiclassical theory for equal mean numbers of photons in the radiation mode.

Note once more that, among all squeezed electromagnetic field states described by (6) which have equal mean numbers of quanta in the field mode, precisely the squeezed vacuum state is characterized by the largest variance of the number of quanta and is therefore most effective in the multiphoton ionization of atoms. Indeed, variance D_k calculations for state (6) with $1/\beta \gg a \gg 1$ (that is, for a strongly squeezed state with a certain degree of coherence) give

$$D_k \approx 2\langle k\rangle^2\left(1 - \frac{a^2}{\langle k\rangle}\right),$$

which is smaller than the D_k value in the squeezed vacuum state.

Generally, the $d\dot{W}_i^{(N)}/dE$ value for nonlinear N-photon ionization is determined by the Nth moment

$$\langle k^N\rangle = \sum k^N|\alpha_k|^2,$$

and we can expect a still larger difference in atomic ionization rates at the same number of mode photons depending on the quantum state of the field. For the squeezed vacuum state with a large mean number of quanta, we have [14, 15]

$$\langle k^N\rangle \sim (2N-1)!!\langle k\rangle^N, \tag{23}$$

and, with $N \approx 10$, which is characteristic of the ionization of atoms by radiation in the optical frequency range, the difference in the rates of ionization by squeezed and classical fields can amount to several orders of magnitude.

For squeezed field states described by (6), a similar formula for $\langle k^N\rangle$ at $1/\beta \gg a \gg 1$ has the form

$$\langle k^N\rangle \sim (2N-1)!!\langle k\rangle^N\left(1 - N\frac{a^2}{2\langle k\rangle}\right). \tag{24}$$

This means that, among all field states of form (6) with equal $\langle k\rangle$ values, precisely the squeezed vacuum state most effectively acts on an atom.

To conclude this section, note that (19) is the general expression for a transition with the absorption of N field quanta. If the transition to the continuum can occur as an M-photon process ($M = [I/\hbar\omega] + 1$) and $M < N$, (19) describes above-threshold ionization with the absorption of $N - M$ above-threshold quanta. It can therefore be stated that the energy spectrum of photoelectrons formed under the action of a squeezed electromagnetic field on an atom should be substantially different from the spectrum obtained in the interaction of an atom with a classical field, namely, the intensity of peaks corresponding to the absorption of a large number of above-threshold quanta should be anomalously high in the squeezed field.

6. GOING BEYOND THE SCOPE OF PERTURBATION THEORY: ONE DISCRETE LEVEL AND CONTINUUM

In this section, we consider the ionization of an atom having a single discrete level in more detail. We will show, without invoking perturbation theory, that the time evolution of an atomic state substantially depends on the quantum electromagnetic field state even when single-photon ionization occurs.

Ignoring free–free transitions in the equation for the $C_{E,k}$ amplitude in the continuum allows system (16) to be written in the form

$$\begin{aligned} i\hbar\frac{dC_{1,k}}{dt} &= -\frac{\varepsilon_0}{\sqrt{2}} \\ &\times\int dE' d_{IE'}\exp\left(\frac{i}{\hbar}(E_1 - E')t\right) \\ &\times\{C_{E',k+1}\sqrt{k+1}\exp(-i\omega t) \\ &\quad + C_{E',k-1}\sqrt{k}\exp(i\omega t)\}, \\ i\hbar\frac{dC_{E,k}}{dt} &= -\frac{\varepsilon_0}{\sqrt{2}}d_{E1}\exp\left(\frac{i}{\hbar}(E - E_1)t\right) \\ &\times\{C_{1,k+1}\sqrt{k+1}\exp(-i\omega t) \\ &\quad + C_{1,k-1}\sqrt{k}\exp(i\omega t)\}. \end{aligned} \tag{25}$$

Applying the method of adiabatic continuum elimination [9] makes it possible to easily obtain the following

equation for the $C_{1,k}$ amplitude:

$$\frac{dC_{1,k}}{dt} + \frac{\pi}{2\hbar}\varepsilon_0^2|d_{E1}|^2 k C_{1,k} = 0, \tag{26}$$

where $E = \hbar\omega - I$.

The integration of (26) yields the equation for the probability of ionization of an atom in time t,

$$\begin{aligned} W_i(t) &= 1 - \sum_k |C_{1,k}(t)|^2 \\ &= 1 - \sum_k |\alpha_k|^2 \exp\left(-\frac{\pi}{\hbar}\varepsilon_0^2|d_{EI}|^2 kt\right). \end{aligned} \tag{27}$$

Here, the α_k values determine the initial electromagnetic field state. Expanding the exponent in Eq. (27) into a series up to first-order terms yields Eq. (17), which corresponds to first-order perturbation theory. Generally, a nonexponential atomic state decay is expected in a nonclassical field for all field states except the stationary state. For a coherent state with a large mean number of quanta ($\langle k\rangle \gg 1$), the k value under the summation sign in (27) can, however, be replaced by $\langle k\rangle$, because the uncertainty in the number of quanta $\Delta k \sim \sqrt{\langle k\rangle} \ll \langle k\rangle$. This gives

$$W_i(t) = 1 - \sum_k |C_{1,k}(t)|^2 = 1 - \exp\left(-\frac{\pi}{\hbar}\varepsilon_0^2|d_{EI}|^2\langle k\rangle t\right),$$

which corresponds to the solution to the problem in the semiclassical approximation [9].

In the squeezed vacuum state, the number-of-quanta distribution is characterized by a large variance; that is, the partial contributions of states with large ($k \gg \langle k\rangle$) and small ($k \ll \langle k\rangle$) numbers of quanta are large. For this reason, the decay rate of an atomic state in a field that is in the squeezed vacuum state is initially substantially higher than when the field oscillator is in a stationary or coherent state and then gradually decreases. There is a nonzero residual probability of nonionization determined by the probability that the field is in the vacuum state,

$$1 - W_i(t \longrightarrow \infty) = |\alpha_{k=0}|^2.$$

7. CONCLUSION

To summarize, we showed in this work that the probability of ionization of an atomic system could substantially change depending on the quantum electromagnetic field state even at a fixed mean number of quanta in the radiation mode. The difference in ionization rates was especially large for multiphoton ionization processes. It was, in particular, shown that a nonclassical field could be much more effective from the point of view of the ionization of atoms than an equally intense classical field.

We can naturally expect that the effectiveness of the action of a nonclassical electromagnetic field on atom-molecular systems (compared with the action of a classical electromagnetic field) should also manifest itself in some other physical phenomena. It can, in particular, be expected that nonlinear polarization in a medium under the action of the squeezed electromagnetic field state can be anomalously large; therefore, squeezed fields can be very effective in the generation of optical radiation harmonics [5]. We can also expect a substantial increase in the probability of above-barrier absorption of electromagnetic field quanta in multiphoton ionization, in the probability of multiquantum absorption in scattering of electrons by atoms in the presence of nonclassical electromagnetic fields, etc.

Note also that the most important difference in the dynamics of interaction between an atomic system and an electromagnetic field arises if the mean number of quanta in the field mode is comparatively small, $\langle k\rangle \geq N$. Such a situation is, for instance, characteristic of the interaction of an atom in a microcavity with a field mode containing few quanta [16].

ACKNOWLEDGMENTS

This work was financially supported by the Russian Foundation for Basic Research (projects nos. 00-02-16046 and 00-15-96554) and CRDF (project RP1-2259).

APPENDIX

Clearly, the amplitudes of the squeezed vacuum state expansion in stationary field states are determined by the equation

$$\begin{aligned} \alpha_{2k} &= \frac{1}{\sqrt{2^{2k}(2k)!\pi\beta}} \\ &\times \int_{-\infty}^{\infty} H_{2k}(\xi)\exp\left(-\frac{\xi^2}{2}\left(1 + \frac{1}{\beta^2}\right)\right)d\xi. \end{aligned} \tag{A.1}$$

Let us express the Hermitian polynomial via the degenerate hypergeometric function,

$$H_{2k}(\xi) = (-1)^k\frac{(2k)!}{k!}F\left(-k, \frac{1}{2}, \xi^2\right), \tag{A.2}$$

to rewrite (A.1) in the form

$$\begin{aligned} \alpha_{2k} &= \frac{(-1)^k}{\sqrt{2^{2k}(2k)!\pi\beta}}\frac{(2k)!}{k!} \\ &\times \int_0^{\infty} F\left(-k, \frac{1}{2}, z\right)\exp(-\lambda z)z^{-1/2}dz, \end{aligned} \tag{A.3}$$

where

$$\lambda = \frac{1}{2}\left(1 + \frac{1}{\beta^2}\right).$$

The integration in (A.3) gives (see [17])

$$\alpha_{2k} = \frac{(-1)^k}{\sqrt{\lambda\beta}}\frac{\sqrt{(2k)!}}{2^k k!}F\left(-k, \frac{1}{2}, \frac{1}{2}, \frac{1}{\lambda}\right). \quad \text{(A.4)}$$

Taking into account that

$$F\left(-k, \frac{1}{2}, \frac{1}{2}, \frac{1}{\lambda}\right) = \left(1 - \frac{1}{\lambda}\right)^k,$$

we can rewrite (A.4) as

$$\alpha_{2k} = \frac{(-1)^k}{\sqrt{\lambda\beta}}\frac{\sqrt{(2k)!}}{2^k k!}\left(1 - \frac{1}{\lambda}\right)^k. \quad \text{(A.5)}$$

This corresponds to Eq. (8) in the main body of the paper.

REFERENCES

1. D. Stoler, Phys. Rev. D **1,** 3217 (1970); **4**, 1925 (1971).
2. R. E. Slusher, L. W. Holleberg, B. Yurke, and J. C. Mertz, Phys. Rev. Lett. **55**, 2409 (1985).
3. Wu Ling-An, H. J. Kimble, J. L. Hall, and Wu Huifa, Phys. Rev. Lett. **57**, 2520 (1986).
4. D. F. Walls, Nature **306**, 141 (1983).
5. V. P. Bykov, Usp. Fiz. Nauk **161** (10), 145 (1991) [Sov. Phys. Usp. **34**, 910 (1991)].
6. O. Svelto, *Principles of Lasers* (Plenum, New York, 1976; Mir, Moscow, 1979).
7. S. A. Akhmanov and S. Yu. Nikitin, *Physical Optics* (Mosk. Gos. Univ., Moscow, 1998).
8. D. N. Klyshko, *Physical Principles of Quantum Electronics* (Nauka, Moscow, 1986).
9. M. V. Fedorov, *Atomic and Free Electrons in a Strong Light Field* (World Sci., Singapore, 1998).
10. N. B. Delone and V. P. Kraĭnov, *Nonlinear Ionization of Atoms by Laser Emission* (Fizmatlit, Moscow, 2001).
11. R. J. Glauber, Phys. Rev. **131**, 2766 (1963).
12. I. Sh. Averbukh and N. F. Perel'man, Usp. Fiz. Nauk **161** (7), 41 (1991) [Sov. Phys. Usp. **34**, 572 (1991)].
13. L. Mandel and E. Wolf, *Optical Coherence and Quantum Optics* (Cambridge Univ. Press, Cambridge, 1995; Fizmatlit, Moscow, 2000).
14. J. Janszky and Y. Yushin, Phys. Lett. A **137**, 451 (1989).
15. A. V. Masalov, Opt. Spektrosk. **70**, 648 (1991) [Opt. Spectrosc. **70**, 382 (1991)].
16. D. Merschede and H. Walter, Phys. Rev. Lett. **54**, 551 (1985).
17. L. D. Landau and E. M. Lifshitz, *Course of Theoretical Physics*, Vol. 3: *Quantum Mechanics: Non-Relativistic Theory* (Nauka, Moscow, 1974; Pergamon, New York, 1977).

Translated by V. Sipachev

Journal of Experimental and Theoretical Physics, Vol. 95, No. 5, 2002, pp. 851–860.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 122, No. 5, 2002, pp. 986–998.

ATOMS, SPECTRA, RADIATION

The Mode Structure and Spectral Properties of Supercontinuum Emission from Microstructure Fibers

A. B. Fedotov[a], Ping Zhou[b], Yu. N. Kondrat'ev[c], S. N. Bagayev[d], V. S. Shevandin[c], K. V. Dukel'skiĭ[c], V. B. Smirnov[e], A. P. Tarasevitch[b], D. von der Linde[b], and A. M. Zheltikov[a, *]

[a]*Physics Faculty, International Laser Center, Moscow State University, Vorob'evy gory, Moscow, 119992 Russia*

[b]*Institut für Laser- und Plasmaphysik, Universität Essen, D-45117 Essen, Germany*

[c]*Vavilov State Optical Institute, All-Russia Research Center, Mendeleevskaya liniya 1, St. Petersburg, 199034 Russia*

[d]*Institute of Laser Physics, Siberian Division, Russian Academy of Sciences, pr. Lavrent'eva 13/3, Novosibirsk, 630090 Russia*

[e]*Russian Center of Laser Physics Research Institute, St. Petersburg State University, Petrodvorets, St. Petersburg, 198504 Russia*

**e-mail: zheltikov@top.phys.msu.su*

Received May 15, 2002

Abstract—The mode structure and spectral properties of supercontinuum emission generated by femtosecond pulses of Ti:sapphire laser radiation in microstructure fibers are studied. The long-wavelength (720–900 nm) and visible (400–600 nm) parts of supercontinuum emission are shown to be spatially separated in microstructure-fiber modes, which can be isolated with an appropriate spectral filtering. The spatial modes thus isolated in spectrally sliced supercontinuum emission possess a spatial quality sufficient for further efficient frequency conversion. The possibility of achieving a high spectral quality of supercontinuum emission is also demonstrated. We explore the ways to control the spectrum of supercontinuum emission by matching parameters of the pump pulse with the parameters of a microstructure fiber and by tuning the initial chirp of the pump pulse. The results of our studies show that supercontinua produced in microstructure fibers offer new approaches to designing a new generation of optical parametric amplifiers and broadband radiation sources for spectroscopic, metrological, and biomedical applications. *© 2002 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

Supercontinuum (SC) generation [1] is a nonlinear-optical phenomenon involving spectral superbroadening of a light pulse resulting from the joint action of the whole set of nonlinear-optical effects, such as self- and cross-phase modulation, four-wave mixing (FWM), and stimulated Raman scattering (SRS), often accompanied by soliton formation and propagation and modulation instabilities. As demonstrated recently, supercontinuum generation can be radically enhanced with the use of fibers of a new type—microstructure (MS) fibers [2–8]. Fibers of this type may provide a high degree of light confinement in the fiber core [9, 10], simultaneously offering broad opportunities for engineering the dispersion of guided modes [11]. Due to their remarkable properties, microstructure fibers make nonlinear optics accessible even to unamplified femtosecond pulses [12]. Enhancement of the broad class of nonlinear-optical phenomena [9, 10, 13], accompanying the propagation of femtosecond pulses in microstructure fibers, makes it possible to generate a supercontinuum emission [6, 14–17] starting with nano- and even subnanojoule energies of laser radiation.

Investigations of supercontinuum generation in microstructure fibers have revealed several interesting physical phenomena related to the nonlinear-optical interactions of ultrashort light pulses, providing a deeper insight into the scenarios of spectral superbroadening of such pulses in different regimes of waveguiding in microstructure fibers [18–20]. The range of applications of supercontinuum generation in microstructure fibers is rapidly expanding, leading to revolutionary changes in optical metrology [21–24], opening new horizons in optical coherence tomography [25], and suggesting new solutions for the creation of compact and practical sources of broadband emission based on supercontinuum generation in microstructure fibers.

It would be very important now, in view of numerous spectroscopic, metrological, and tomographic applications of supercontinua generated in microstructure fibers, not only to improve the reproducibility and stability of temporal characteristics and the spectral content of supercontinuum emission, but also to ensure a high spatial mode quality of this emission. In many practically important cases, highly efficient supercon-

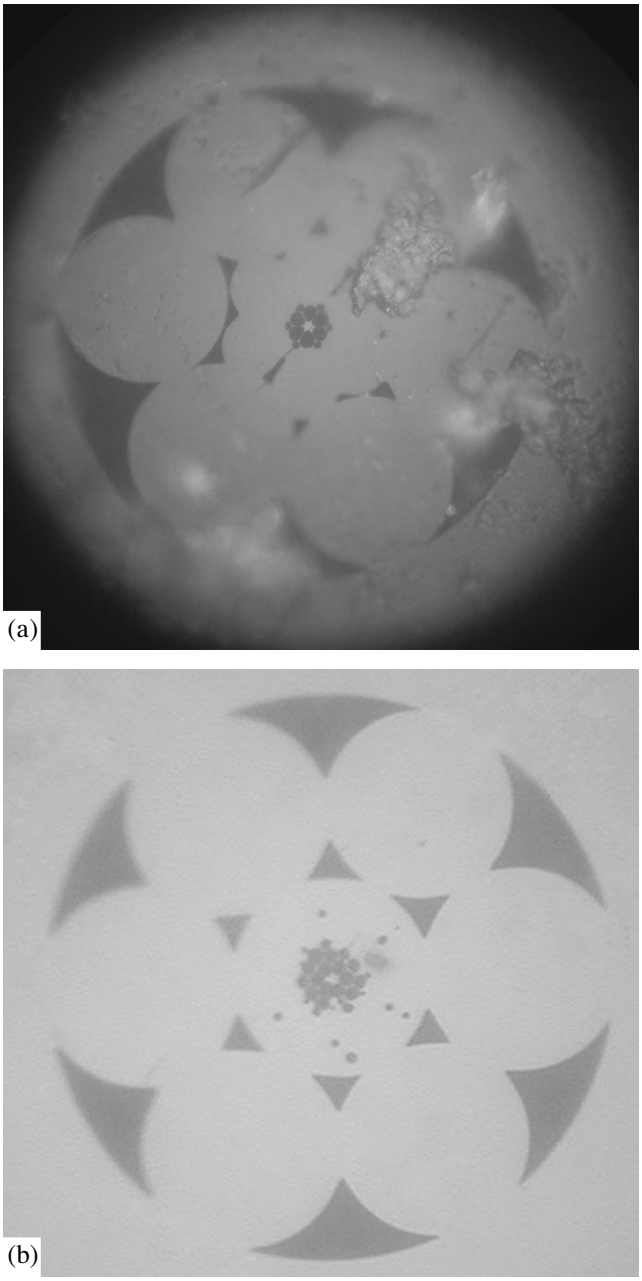


Fig. 1. Microscope cross-sectional images of microstructure fibers: (a) a fiber with a single hexagonal cycle of air holes around the fiber core with a diameter of 2 μm and (b) a fiber with two hexagonal cycles of air holes around the fiber core with a diameter of 2 μm.

tinuum generation in microstructure fibers involves multimode-phase-matched four-wave mixing, which leads to the spectral superbroadening. Supercontinua are also emitted in the multimode regime under these conditions.

In this paper, we propose and experimentally implement a method of spatial filtering and spectral slicing of supercontinuum emission generated in a microstructure fiber with a small core diameter, where multimode-phase-matched four-wave mixing results in a preferable generation of new spectral components emitted as a part of the supercontinuum in a certain (possibly high-order) guided mode. Such fibers have been fabricated and investigated as a part of the work presented in this paper. The results of our experimental studies demonstrate the possibility of separating isolated spatial modes in supercontinuum emission produced in such fibers within different spectral ranges. The proposed method of spatial mode filtering provides a high spatial quality of supercontinuum emission, which is sufficient to allow efficient further frequency conversion of spectrally sliced supercontinuum emission. This frequency convertibility of SC spatial modes was demonstrated by our experiments where spectrally sliced supercontinuum emission was mixed with the fundamental radiation of a Ti:sapphire laser in a nonlinear crystal to produce a sum-frequency signal. We will also demonstrate the possibility of achieving a high spectral quality of supercontinua produced in microstructure fibers and explore the ways to control the spectrum of supercontinuum emission by matching parameters of the pump pulse with the parameters of a microstructure fiber and by varying the initial chirp of the pump pulse. The results of our studies show that supercontinua produced in microstructure fibers offer new approaches to designing a new generation of optical parametric amplifiers and broadband radiation sources for spectroscopic, metrological, and biomedical applications.

2. EXPERIMENTAL

Experiments on supercontinuum generation were performed with a family of microstructure optical fibers where a cladding consists of one, two, or more hexagonal cycles of air holes. A system of smaller auxiliary air holes in the cladding of these fibers improves the confinement of light field in the fiber core (Figs. 1a, 1b). The technology employed to fabricate MS fibers used in our experiments was similar to a standard procedure described, e.g., in [2, 26]. However, instead of using a preform consisting of a set of identical capillaries, we started with a preform consisting of fused silica fibers with different diameters. In particular, to fabricate the basis fiber of the created family—a fiber with a single hexagonal cycle of air holes around the fiber core (Fig. 1a)—we employed a preform whose central part included a fiber with the minimum diameter surrounded by six capillaries. The fabrication of a such a fiber with a minimally microstructured cladding was reported earlier in [27, 28]. The possibility of changing the spatial sizes of the structure was built in the process of fiber fabrication. The minimum diameter of the core in the created MS fibers was equal to 1 μm. The air-filling fraction of the microstructure part of the cladding in the created fibers, as can be seen from Fig. 1, is very high, providing a high refractive index step between the core and the cladding in the fiber. A system of smaller auxiliary air holes in the cladding of these fibers,

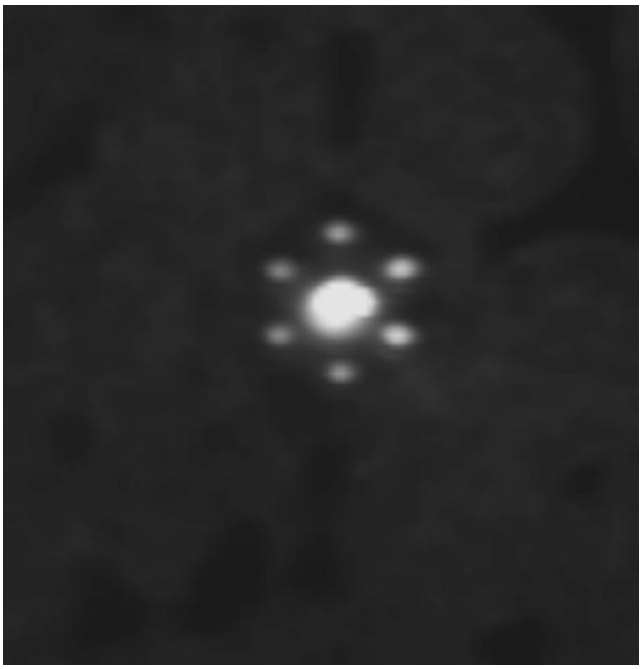


Fig. 2. Transverse intensity distribution of Ti:sapphire-laser radiation at the output end of a microstructure fiber with a single hexagonal cycle of air holes around the fiber core with a diameter of 2 μm in the low-power-pump no-supercontinuum-emission regime.

increasing the refractive-index step between the fiber core and the fiber cladding and preventing the guided modes from leaking into the cladding, adds to the confinement of light field in the fiber core and reduces optical losses of fiber modes (as can be seen from Fig. 2, which displays an image of a Ti:sapphire laser radiation intensity distribution measured at the output end of the fiber). The fiber with such a geometry is a basis fiber for the created family of fibers since the cladding of this fiber has a minimal number of holes (a single cycle of holes, plus auxiliary holes, improving light-field confinement in the fiber core).

To fabricate fibers with a more complicated structure, we modified the preform. In a microstructure fiber shown in Fig. 1b, the fused silica core is surrounded by two hexagonal cycles of air holes and a system of smaller auxiliary holes, improving light-field confinement in the fiber core. The increase in the number of cycles of air holes around the fiber core reduces the magnitude of fiber losses. Optical losses have been determined for microstructure fibers of this type from the results of measurements [28] performed on ~100-m MS-fiber segments. The magnitude of optical losses was estimated to be 2–3 dB/m for fibers with a single hexagonal cycle of air holes in the cladding and 0.4–0.5 dB/m for fibers with two cycles of air holes.

Spectral broadening and supercontinuum generation in the created MS fibers were studied in our experiments with the use of femtosecond pulses produced by a Ti:sapphire laser system. This laser system included a Ti:sapphire master oscillator and a regenerative amplifier and was capable of generating 40-fs pulses of 800-nm radiation with an energy up to 0.2 mJ per pulse and a repetition rate of 1 kHz. The energy of laser pulses coupled into the fiber ranged from 0.1 up to 50 nJ. Experiments were performed with fiber samples with a length of 4–200 cm. The laser beam was focused onto the input end of a fiber sample, placed on a three-dimensional translation stage, with a microobjective (Fig. 3). Radiation coming out of the fiber was collimated with an identical microobjective and was split

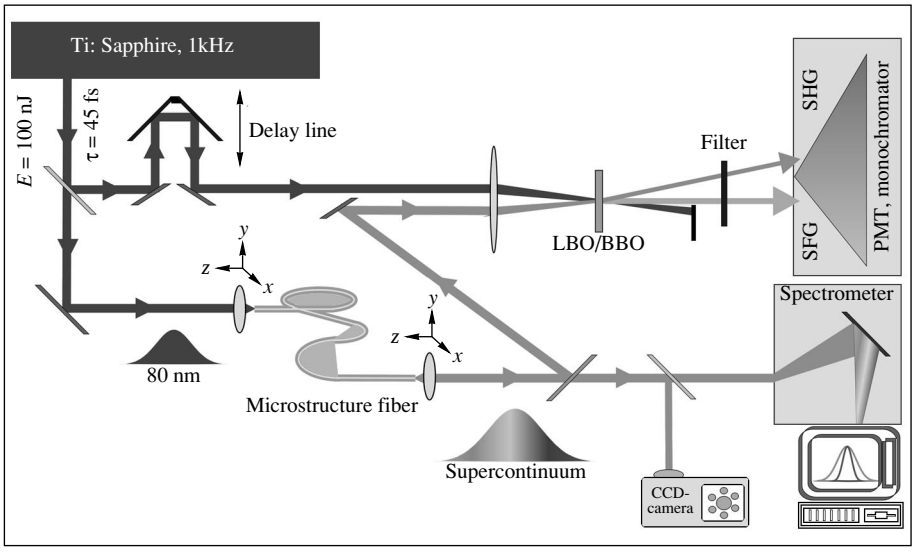


Fig. 3. Diagram of the experimental setup for studying supercontinuum generation in microstructure fibers: SFG, the signal resulting from sum-frequency generation in the nonlinear crystal; SHG, the signal of second-harmonic generation in the nonlinear crystal; PMT, photoelectric multiplier.

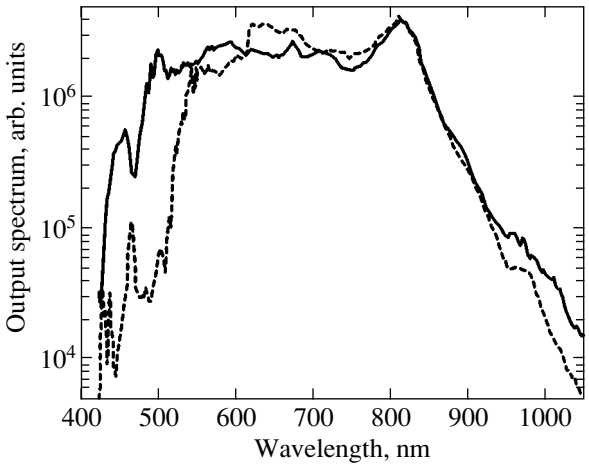


Fig. 4. Spectra of SC emission generated by 40-fs pulses with an energy of 2 nJ (dashed curve) and 3 nJ (solid curve) in a 1.5-m MS fiber with a single ring of air holes around the fiber core (Fig. 1a) and a core diameter of 3 μm.

into two beams. One of these beams was delivered to a spectrograph, while the other one was used to visualize the transverse intensity distribution in the emission coming out of the MS fiber by imaging the output end of the fiber onto a CCD camera.

3. THE MODE STRUCTURE OF SUPERCONTINUUM EMISSION

Propagation of femtosecond laser pulses through an MS fiber was accompanied by a considerable spectral broadening of these pulses. With only a few nanojoules of Ti:sapphire laser radiation coupled into an MS fiber sample with a length of several centimeters, we observed the generation of SC emission with a spectral bandwidth exceeding an octave. Figure 4 shows typical spectra of an SC generated by 40-fs pulses with an energy of 2 and 3 nJ in a 1.5-m MS fiber with a single ring of air holes around the fiber core and a core diameter of 3 μm.

Figure 2 displays the transverse intensity distribution of pump radiation at the output end of the fiber in the regime of low-power pump, when no SC is generated. The transverse structure of the guided mode of pump radiation, as can be seen from this image, features a sixfold rotational symmetry, which is also characteristic of the fiber structure. Supercontinuum emission was generally produced in the multimode regime in our experiments. However, we were able to filter isolated spatial modes for different spectral ranges of supercontinuum emission using a set of color-glass filters. Figures 5a–5c present typical results of such experiments performed for an SC generated in a 1.5-m-long MS fiber with a single ring of air holes around the fiber core with a core diameter of 3 μm. The transverse intensity distribution of SC emission measured with a filter providing maximum transmission within the range of 720–900 nm (Fig. 5a) has a bell-like shape, displaying a single maximum on the beam axis. The visible part of SC emission (400–600 nm), on the other hand, has a doughnut-like spatial mode structure (Fig. 5b) under the same experimental conditions.

With a slight variation in the initial conditions of mode excitation at the input end of the MS fiber, the doughnut mode of the visible part of SC emission (Fig. 5b) was transformed into a more complicated, two-lobe pattern shown in Fig. 5c. Both the doughnut-like mode of Fig. 5b and the two-lobe mode of Fig. 5c remained reproducible and stable and were observed for MS fibers with lengths ranging from several centimeters up to 2 m. Apparently, because of the poorer spatial overlapping between the pump beam and the two-lobe mode, the short-wavelength part of SC emission in the case of the two-lobe mode (the dashed line in Fig. 6) was much less intense than in the case of the doughnut mode (the solid line in Fig. 6).

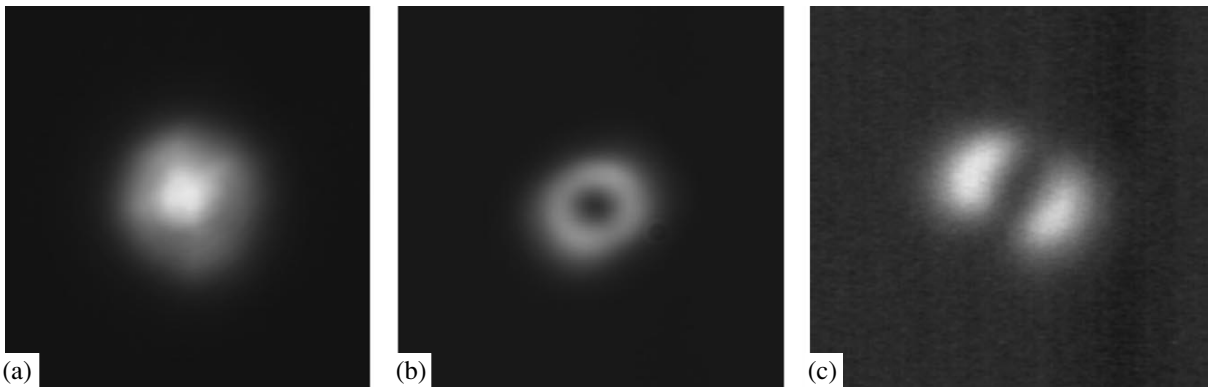


Fig. 5. Transverse intensity distributions of SC emission generated in a 1.5-m-long MS fiber with a single ring of air holes around the fiber core with a core diameter of 3 μm measured within the spectral range of (a) 720–900 and (b, c) 400–600 nm. With a variation in the initial conditions at the input end of the fiber, a doughnut mode of the visible part of SC emission (b) tends to transform into a two-lobe mode (c).

Physically, our idea of using microstructure fibers with small fiber cores for the generation of supercontinuum emission that could be spectrally sliced in separated spatial modes is based on the fact that the difference between propagation constants for adjacent fiber modes supported by a fiber at a certain frequency increases with a decrease in the fiber core radius. Within the framework of the elementary theory of optical fibers [29], the relation between the difference of propagation constants $\Delta\beta$ characterizing two adjacent guided modes in a fiber and the fiber core radius a is given by

$$\Delta\beta \approx \frac{\pi^2 c f_n}{4a^2 \omega n_c},$$

where c is the speed of light, ω is the radiation frequency, f_n is a function of the mode index, and n_c is the refractive index of the fiber core. This elementary relation is very instructive, however, as it explains in a very simple way why MS fibers with small core diameters may generate mode-separable supercontinuum emission in the regime when FWM processes are phase-matched only for a certain spatial mode of the nonlinear signal generated through FWM for each SC slice (or at least for some of the SC slices).

Transformation of the spatial distribution of SC emission observed in our experiments (Fig. 5) indicates changes in multimode phase matching for FWM processes contributing to spectral superbroadening within different spectral ranges. Our measurements performed on MS fibers with a core diameter of 3 μm also show that phase matching is achieved only for a certain spatial mode of the emitted signal within each of the studied spectral ranges. This circumstance allows isolated spatial modes to be separated by spectrally slicing SC emission.

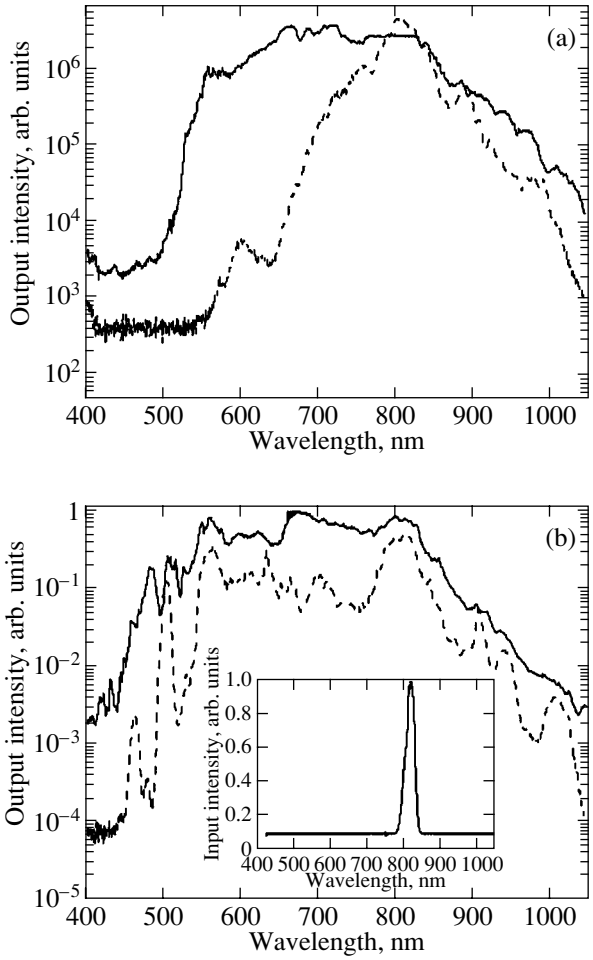


Fig. 6. Spectra of supercontinuum emission generated by 40-fs 5-nJ pulses of 800-nm Ti:sapphire laser radiation in (a) 7-cm and (b) 1-m MS fibers with a single ring of air holes around the fiber core with a core diameter of 3 μm in the regime of (solid line) doughnut-like and (dashed line) two-lobe mode of the visible part of the supercontinuum. The spectrum of the Ti:sapphire pulse coupled into the fiber is shown in the inset in Fig. 6b.

4. NONLINEAR-OPTICAL FREQUENCY CONVERSION OF SPECTRALLY SLICED SUPERCONTINUUM EMISSION AND CROSS-CORRELATION MEASUREMENTS

Frequency convertibility of spectrally sliced supercontinuum is an important criterion of the quality of spatial modes of SC emission. Based on this criterion, we may also judge whether SC emission generated in MS fibers and spectrally sliced with the use of the above-described technique can be employed in practice for spectroscopic studies and pump–supercontinuum probe measurements.

With these circumstances in mind, we experimentally assessed the efficiency of nonlinear-optical frequency conversion for spectrally sliced spatial modes of SC emission produced in an MS fiber. The sum-frequency signal was produced in our experiments by mixing different parts of SC emission with the fundamental radiation of the above-described Ti:sapphire laser in a 100-μm-thick LBO crystal. Figure 7 presents the results of these measurements performed with the use of the long-wavelength part (~720–900 nm) of SC emission (the spectrum of this radiation is shown by curve *1* in panel (a) of Fig. 7) mixed with the fundamental radiation of the Ti:sapphire laser in the LBO crystal in the noncollinear geometry of sum-frequency generation (SFG). The broadband sum-frequency signal was produced within the spectral range from 370 up to 430 nm in the direction determined by phase-matching conditions (see diagram *1* in Fig. 7). This geometry of sum-frequency generation allowed the efficiency of frequency conversion of about 0.1% to be achieved. We also observed collinear second-harmonic generation (SHG) using the long-wavelength part of SC emission as a pump (see diagram *2* in Fig. 7). The efficiency of this second-harmonic generation process under our experimental conditions was more than an order of magnitude lower than the efficiency of noncollinear

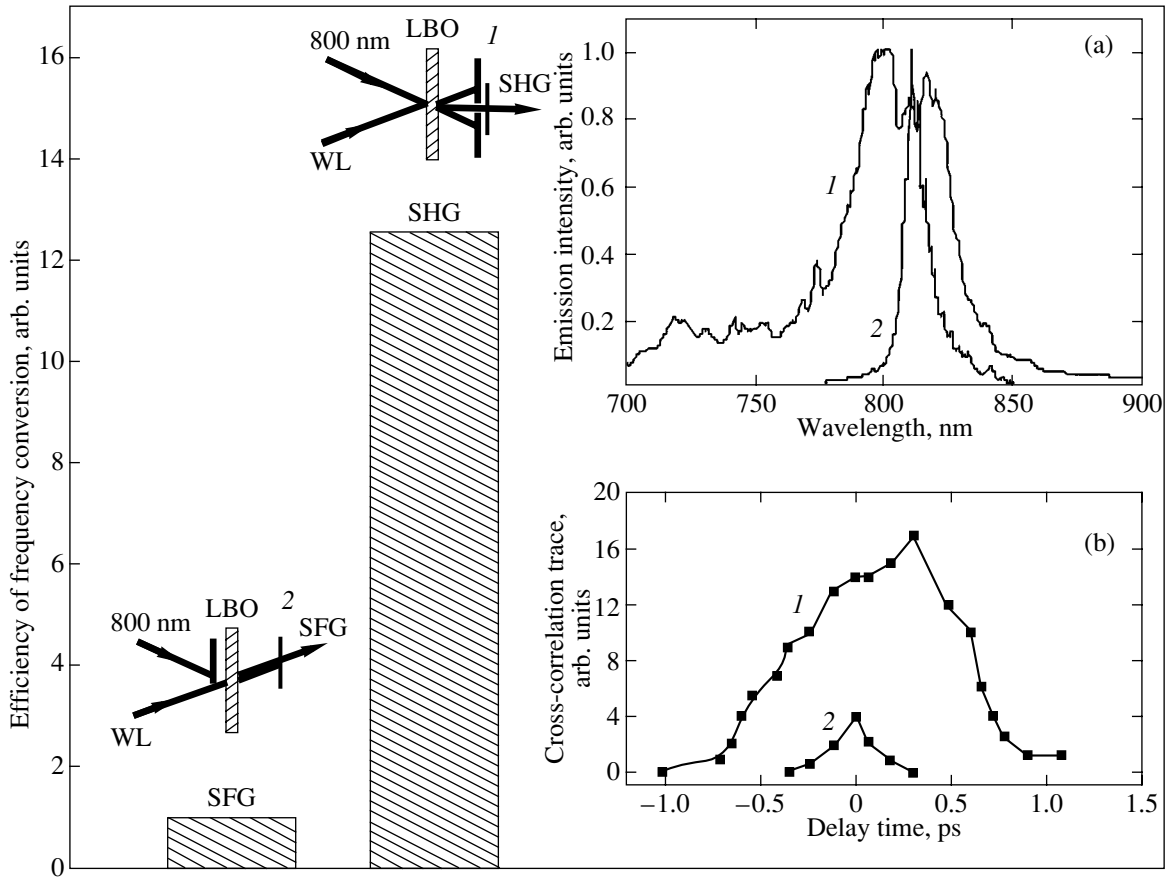


Fig. 7. Nonlinear-optical frequency conversion of spectrally sliced white light (WL) and cross-correlation measurements. (a) The spectrum of supercontinuum emission produced within the range of wavelengths from 720 up to 900 nm in a microstructure fiber with a single ring of holes around the fiber core, a core diameter of 3 μm, and the length of 1.5 m (curve *1*) and the spectrum of low-intensity Ti:sapphire laser pulses transmitted through the same microstructure fiber with virtually no or very weak spectral broadening (curve *2*). (b) Cross-correlation traces for a pulse of the 720–900 nm spectral slice of supercontinuum (curve *1*) and weakly broadened low-intensity Ti:sapphire laser pulses transmitted through the same microstructure fiber with virtually no or very weak spectral broadening (curve *2*). Beam diagrams of (*1*) noncollinear sum-frequency generation (SFG) and (*2*) collinear second-harmonic generation (SHG) are also shown. The heights of the vertical bars represent the relative efficiencies of the SFG and SHG processes.

sum-frequency generation (the relative efficiencies of SFG and SHG processes are shown on the left-hand side of Fig. 7). The results of these experiments show that the efficiency of nonlinear-optical frequency conversion of spectrally sliced SC emission can be made high enough to allow nonlinear-optical spectroscopic and time-resolved measurements and to use SC generated in MS fibers as seed radiation for optical parametric amplification.

Sum-frequency and second-harmonic generation experiments performed with spectrally sliced SC emission also allow the characteristic pulse duration to be estimated for different parts of SC emission. Panel (b) of Fig. 7 presents the results of such cross-correlation experiments, where the intensity of the sum-frequency signal was measured as a function of the delay time between the fundamental radiation pulse of the Ti:sapphire laser and the broadband emission pulse coming out of the fiber and passing through a set of optical filters. Cross-correlation traces measured with the use of this approach were compared with the results of cross-correlation measurements performed in the same geometry for low-intensity Ti:sapphire laser pulses transmitted through the MS fiber with virtually no or very weak spectral broadening (the spectrum of this signal is shown by line *2* in panel (a) of Fig. 7). The cross-correlation trace measured for the broadband signal of spectrally sliced supercontinuum (curve *1* in panel (b) of Fig. 7) was much broader than the cross-correlation trace measured for the signal with virtually no or very weak spectral broadening (curve *2* in panel (b) of Fig. 7). This comparison shows that different spectral components emitted as a part of the supercontinuum are characterized by different delay times. This effect is associated with the dispersion properties of MS-fiber-guided modes and can be employed to temporally and spatially resolve different frequency components in pump–supercontinuum probe experiments [30–32].

5. CHIRP CONTROL OF SUPERCONTINUUM GENERATION IN MICROSTRUCTURE FIBERS

In this section, we will present the results of experimental studies demonstrating the possibility of controlling the process of supercontinuum generation in a microstructure fiber by changing the initial chirp of the pump pulse. Optimizing the initial chirp of the input pulse, one can improve the efficiency of SC generation in an MS fiber and increase the spectral width of SC emission. The initial chirp of the pump pulse coupled into an MS fiber has an influence on the efficiency of SC generation and the shape of the SC spectrum through several physical mechanisms. Dispersion spreading of a short pump pulse propagating through an MS fiber is one of the most important among these factors. With an appropriate choice of the initial chirp, one can radically change the evolution of the waveform of a light pulse propagating through the fiber. When nonlinear-optical processes and high-order dispersion effects are negligible, the evolution of the duration of a Gaussian pulse with a chirp parameter α and the initial pulse duration τ_0 as a function of the propagation coordinate x is given by [33]

$$\tau(x) = \tau_0\sqrt{(1-\alpha k_2 x)^2 + (x/L_d)^2},$$

where k_2 is the group-velocity dispersion and $L_d = \tau_0^2/|k_2|$ is the dispersion length. The minimum pulse duration under these conditions is achieved at the distance

$$l_c = \alpha\tau_0^2 L_d[1 + (\alpha\tau_0^2)^2]^{-1}.$$

Thus, by choosing the sign and the absolute value of the initial chirp, one can precompensate for a dispersion spreading of the pump pulse within a certain section of an MS fiber (with a length on the order of l_c). An initially chirped pulse then first experiences compression while propagating through an MS fiber. This phase of compression is then, of course, followed by the normal dispersion spreading of the pulse.

Group-delay effects [34] and the sensitivity of phase matching to the temporal profile of the phase in the pulse [35] also limit the generation of broadband emission in an MS fiber. In particular, the frequency dependence of the group velocity of light pulses leads to the spatial walk-off of the generated frequency components with respect to the spectral components of the pump pulse. The characteristic distance l_g where the group delay becomes comparable with the pump pulse duration τ can be estimated as

$$l_g = \frac{\tau}{\left|u_p^{-1} - u_s^{-1}\right|},$$

where u_p and u_s are the group velocities of the pump pulse and the signal pulse corresponding to a certain group of spectral components. In view of large spectral widths characteristic of the SC generation process, group-delay effects may have a considerable influence on the width and the structure of SC spectra.

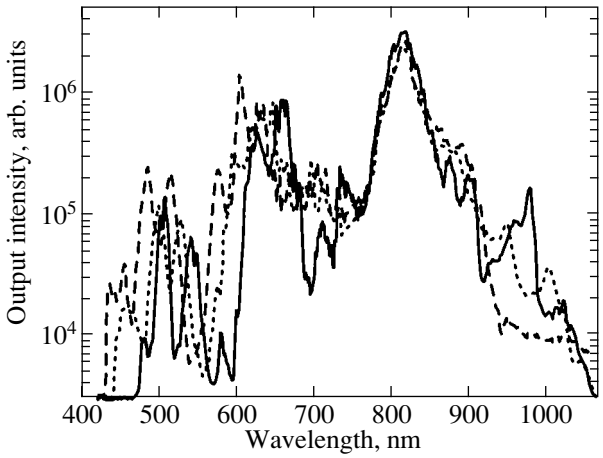


Fig. 8. The influence of the initial chirp of pump pulses on the spectrum of supercontinuum emission from a microstructure fiber with a single ring of holes around the fiber core (Fig. 1a), a core diameter of 3 μm, and the length of about 10 cm. The chirp parameter α and the initial pump pulse duration τ_0 are equal to (solid line) $\alpha = 0$, $\tau_0 = 62$ fs; (dotted line) $\alpha = -7 \times 10^{-5}$ fs^{-2}, $\tau_0 = 65$ fs; and (dashed line) $\alpha = -10^{-4}$ fs^{-2}, $\tau_0 = 70$ fs.

The goal of optimization of the initial chirp of the pump pulse coupled into the fiber is to reduce the dispersion of the pump pulse in the fiber, improve phase matching with the inclusion of nonlinear phase shifts, and increase the effective length of interaction between the spectral components of the pump pulse and frequency-separated components arising in the process of SC generation. Figure 8 displays the results of experiments on supercontinuum generation in a 10-cm MS fiber with a single cycle of air holes around the fiber core (Fig. 1a) and a core diameter of 3 μm. As can be seen from the presented results, the variation in the initial chirp of the pump pulse noticeably changes the spectra of SC emission. Unfortunately, a detailed quantitative analysis of the physical information encoded in chirp-sensitive spectra of supercontinuum emission is impeded by the fact that a variation in the initial chirp of the pump pulse under conditions of our experiments simultaneously slightly changes the duration of this pulse, which, in turn, leads to a variation in the intensity of pump radiation. It is important to mention, however, that a negative chirping of the pump pulse improves the efficiency of generation of high-frequency components in SC emission (as can be seen from Fig. 8) even in spite of a slight increase in the pulse duration, which is equal to 62, 65, and 70 fs for the spectra shown by the solid, dotted, and dashed lines in Fig. 8, respectively.

In view of the large number of physical factors involved in SC generation by ultrashort laser pulses in an MS fiber, a linear chirp is generally insufficient to optimize this process. Methods of optimal phase con-

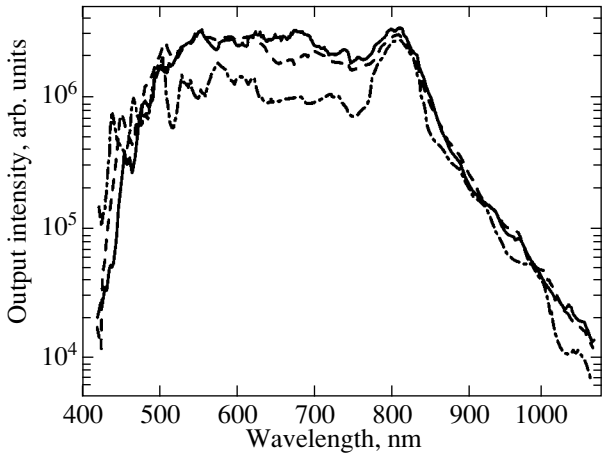


Fig. 9. Spectra of supercontinuum emission generated by a 50-fs Ti:sapphire laser pulse with an energy of (dash-dotted line) 8, (dashed line) 12, and (solid line) 16 nJ in a 1-m microstructure fiber with a single ring of holes around the fiber core (Fig. 1a) and a core diameter of 4 μm.

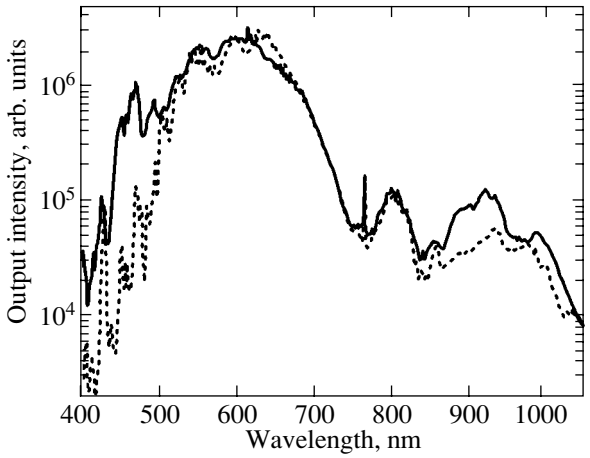


Fig. 10. Pump depletion in supercontinuum generation: spectra of supercontinuum emission generated by a 60-fs Ti:sapphire laser pulse with an energy of 8 nJ (dashed line) and 12 nJ (solid line) in a 1-m microstructure fiber with a single ring of holes around the fiber core (Fig. 1a) and a core diameter of 2.5 μm.

trol [36–39] with the use of spatial light modulators [40–43], allowing light pulses with arbitrary amplitude and phase profiles to be produced, seem to offer much promise for optimizing the initial chirp of the input pulse for the efficient generation of supercontinuum emission with controllable spectral and temporal parameters.

6. SPECTRAL PROPERTIES OF SUPERCONTINUUM EMISSION

The spectral quality is the key property of supercontinuum for numerous spectroscopic and metrological applications, as well as for ultrashort-pulse synthesis and for using supercontinuum emission to seed optical parametric amplification. In this section, we will show that MS fibers allow the generation of supercontinuum with a high spectral quality using femtosecond pulses of moderate powers.

The results of experimental studies presented in Fig. 9 demonstrate that the parameters of input laser pulses can be matched with the characteristics of an MS fiber in such a way as to allow the generation of SC emission with flat spectra spanning over nearly an octave. Supercontinuum emission with such a spectrum can be employed for spectroscopic purposes, as well as seed radiation for optical parametric amplification.

The possibility of achieving highly efficient conversion of pump pulse energy into the visible range is an important property of SC emission generated in MS fibers. This possibility is illustrated by Fig. 10, which displays the spectra of SC emission generated by 60-fs pulses of Ti:sapphire laser radiation with an energy on the order of 10 nJ coupled into a 1-m MS fiber with a single cycle of air holes around the fiber core. The pump field is depleted under these conditions, resulting in a highly efficient generation of a broadband visible light.

It should be emphasized that the spectral quality of SC emission attainable with MS fibers (Fig. 9) is much higher than the quality of white-light emission conventionally employed for optical parametric amplification. This circumstance, complemented by the tunability and controllability of SC spectra (Figs. 4, 6, 9, 10) by matching parameters of the pump pulse with the parameters of an MS fiber, including the possibility of achieving the pump-depletion regime (Fig. 10), suggests ways of using MS fibers for creating a new generation of optical parametric amplifiers and broadband sources for spectroscopic studies.

7. CONCLUSION

Thus, based on the investigation of the mode structure and spectral properties of supercontinuum emission produced in microstructure fibers, we proposed and experimentally implemented the method of spatial filtering and spectral slicing of this supercontinuum emission. The key physical effect behind our method of spatial and spectral filtering of SC emission is associated with the increase in the separation between the propagation constants corresponding to adjacent guided modes with a decrease in the fiber core diameter. Microstructure fibers provide sufficiently small core diameters for the realization of this approach. Supercontinua in such fibers can be generated in the regime when multimode-phase-matched four-wave mixing results in a preferable generation of new spectral components emitted as a part of the supercontinuum in a certain (perhaps, high-order) guided mode. The results of our experimental studies presented in this paper demonstrate the possibility of separating isolated

spatial modes in supercontinuum emission produced in such fibers within different spectral ranges. The proposed method of spatial mode filtering provides a high spatial quality of supercontinuum emission, which is sufficient to allow efficient further frequency conversion of spectrally sliced supercontinuum emission. This frequency convertibility of SC spatial modes was demonstrated by our experiments where spectrally sliced supercontinuum emission was mixed with the fundamental radiation of a Ti:sapphire laser in a nonlinear crystal to produce a sum-frequency signal. The method of spatial and spectral filtering of supercontinuum emission from microstructure fibers developed in this paper opens the way to conveniently and efficiently employ MS-fiber-generated supercontinuum emission for spectroscopic applications, time-resolved measurements, optical metrology, and coherence tomography, offering, at the same time, new solutions for synthesizing ultrashort light pulses.

Our studies have also demonstrated the possibility of achieving a high spectral quality of supercontinua produced in microstructure fibers. We have explored the ways to control the spectrum of supercontinuum emission by matching parameters of the pump pulse with the parameters of a microstructure fiber and by varying the initial chirp of the pump pulse. The results of our investigations show that supercontinua produced in microstructure fibers offer new approaches to designing a new generation of optical parametric amplifiers and broadband radiation sources for spectroscopic, metrological, and biomedical applications.

ACKNOWLEDGMENTS

This study was supported in part by the President of Russian Federation Grant no. 00-15-99304, the Russian Foundation for Basic Research (project no. 00-02-17567), the Volkswagen Foundation (project no. I/76 869), CRDF Award no. RP2-2266, and the “Fundamental Metrology” Federal Science and Technology Program of the Russian Federation.

REFERENCES

1. *The Supercontinuum Laser Source*, Ed. by R. Alfano (Springer-Verlag, Berlin, 1989).
2. J. C. Knight, T. A. Birks, P. St. J. Russell, and D. M. Atkin, Opt. Lett. **21**, 1547 (1996).
3. J. C. Knight, J. Broeng, T. A. Birks, and P. St. J. Russell, Science **282**, 1476 (1998).
4. R. F. Cregan, B. J. Mangan, J. C. Knight, *et al.*, Science **285**, 1537 (1999).
5. P. J. Bennet, T. M. Monro, and D. J. Richardson, Opt. Lett. **24**, 1203 (1999).
6. J. K. Ranka, R. S. Windeler, and A. J. Stentz, Opt. Lett. **25**, 25 (2000).
7. A. B. Fedotov, A. M. Zheltikov, L. A. Mel’nikov, *et al.*, Pis’ma Zh. Éksp. Teor. Fiz. **71**, 407 (2000) [JETP Lett. **71**, 281 (2000)].
8. A. M. Zheltikov, Usp. Fiz. Nauk **170**, 1203 (2000).
9. N. G. R. Broderick, T. M. Monro, P. J. Bennett, and D. J. Richardson, Opt. Lett. **24**, 1395 (1999).
10. A. B. Fedotov, A. M. Zheltikov, A. P. Tarasevitch, and D. von der Linde, Appl. Phys. B **73**, 181 (2001).
11. J. C. Knight, J. Arriaga, T. A. Birks, *et al.*, IEEE Photonics Technol. Lett. **12**, 807 (2000).
12. D. A. Akimov, A. A. Ivanov, M. V. Alfimov, *et al.*, Appl. Phys. B **74**, 307 (2002).
13. A. N. Naumov, A. B. Fedotov, A. M. Zheltikov, *et al.*, J. Opt. Soc. Am. B **19** (9), 2183 (2002).
14. St. Coen, A. H. L. Chau, R. Leonhardt, *et al.*, Opt. Lett. **26**, 1356 (2001).
15. T. A. Birks, W. J. Wadsworth, and P. St. J. Russell, Opt. Lett. **25**, 1415 (2000).
16. R. Holzwarth, M. Zimmermann, Th. Udem, *et al.*, Opt. Lett. **26**, 1376 (2001).
17. D. A. Akimov, A. B. Fedotov, A. A. Podshivalov, *et al.*, Pis’ma Zh. Éksp. Teor. Fiz. **74**, 515 (2001) [JETP Lett. **74**, 460 (2001)].
18. J. Herrmann, U. Griebner, N. Zhavoronkov, *et al.*, Phys. Rev. Lett. **88**, 173901 (2002).
19. W. J. Wadsworth, A. Ortigosa-Blanch, J. C. Knight, *et al.*, J. Opt. Soc. Am. B **19** (9), 2148 (2002).
20. S. Coen, A. H. L. Chau, R. Leonhardt, *et al.*, J. Opt. Soc. Am. B **19**, 753 (2002).
21. S. A. Diddams, D. J. Jones, Jun Ye, *et al.*, Phys. Rev. Lett. **84**, 5102 (2000).
22. D. J. Jones, S. A. Diddams, J. K. Ranka, *et al.*, Science **288**, 635 (2000).
23. R. Holzwarth, T. Udem, T. W. Hänsch, *et al.*, Phys. Rev. Lett. **85**, 2264 (2000).
24. S. N. Bagayev, A. K. Dmitriyev, S. V. Chepurov, *et al.*, Laser Phys. **11**, 1270 (2001).
25. I. Hartl, X. D. Li, C. Chudoba, *et al.*, Opt. Lett. **26**, 608 (2001).
26. A. B. Fedotov, M. V. Alfimov, A. A. Ivanov, *et al.*, Laser Phys. **11**, 138 (2001).
27. A. M. Zheltikov, Ping Zhou, V. V. Temnov, *et al.*, Quantum Electron. **32** (6), 542 (2002).
28. A. B. Fedotov, S. O. Konorov, Yu. N. Kondrat’ev, *et al.*, Laser Phys. **12** (10) (2002).
29. A. W. Snyder and J. D. Love, *Optical Waveguide Theory* (Chapman and Hall, London, 1983; Radio i Svyaz’, Moscow, 1987).
30. S. A. Kovalenko, A. L. Dobryakov, J. Ruthmann, and N. P. Ernsting, Phys. Rev. A **59**, 2369 (1999).
31. S. A. Kovalenko, R. Schanz, V. M. Farztdinov, *et al.*, Chem. Phys. Lett. **323**, 312 (2000).
32. A. L. Dobryakov, S. A. Kovalenko, V. M. Farztdinov, *et al.*, Solid State Commun. **116**, 437 (2000).

33. S. A. Akhmanov, V. A. Vysloukh, and A. S. Chirkin, *The Optics of Femtosecond Pulses* (Nauka, Moscow, 1988).
34. A. B. Fedotov, A. N. Naumov, A. M. Zheltikov, *et al.*, J. Opt. Soc. Am. B **19** (9), 2156 (2002).
35. A. N. Naumov and A. M. Zheltikov, Opt. Express **10**, 122 (2002).
36. P. Brumer and M. Shapiro, Chem. Phys. Lett. **126**, 541 (1986).
37. D. J. Tannor, R. Kosloff, and S. A. Rice, J. Chem. Phys. **85**, 5805 (1986).
38. R. S. Judson and H. Rabitz, Phys. Rev. Lett. **68**, 1500 (1992).
39. D. Zeidler, S. Frey, K.-L. Kompa, and M. Motzkus, Phys. Rev. A **64**, 023420 (2001).
40. A. M. Weiner, D. E. Leaird, G. P. Wiederrecht, and K. A. Nelson, J. Opt. Soc. Am. B **8**, 1264 (1991).
41. C. W. Hillegas, J. X. Tull, D. Goswami, *et al.*, Opt. Lett. **19**, 737 (1994).
42. A. M. Weiner, Prog. Quantum Electron. **19**, 161 (1995).
43. M. M. Wefers and K. A. Nelson, Science **262**, 1381 (1993).

Translated by A. Zheltikov

Journal of Experimental and Theoretical Physics, Vol. 95, No. 5, 2002, pp. 861–863.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 122, No. 5, 2002, pp. 999–1002.

ATOMS, SPECTRA, RADIATION

The External Field Effect on Transient Radiation of an Ultrarelativistic Particle

M. I. Ryazanov

Moscow State Institute of Engineering Physics (Technical University), Kashirskoe sh. 31, Moscow, 115409 Russia
e-mail: ryazanov@theor.mephi.msk.su

Received June 24, 2002

Abstract—An external field determining the law of particle motion was shown to change the intensity of transient radiation in the ultrarelativistic case. The angular and frequency distribution of transient radiation in an external field was obtained. The possibility of determining the energy of an ultrarelativistic particle from the measured azimuthal asymmetry of particle transient radiation in an external field was discussed. © *2002 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

The formation of transient radiation when a fast particle crosses the boundary surface between media [1–5] occurs in a finite region of space at the interface. When a photon of frequency ω is emitted by an ultrarelativistic particle having the energy

$$E \equiv \gamma mc^2 \gg mc^2,$$

the path length of the particle in the region of radiation formation (the coherence length) is $l \sim (c/\omega)\gamma^2$, and the cross section of this region is small compared with l. Transient radiation arises without external field participation. However, if a particle deflected by a field escapes the region of radiation formation, this disturbs the radiation formation process, and its intensity decreases. At high energies and low frequencies, coherence length l can be large to the extent that an external field can withdraw the particle from the formation region, thereby decreasing the intensity of transient radiation.

It would be interesting to estimate the external field effect on transient radiation and find the region of frequencies and energies in which this effect can be substantial.

2. TRANSIENT RADIATION OF A CHARGE MOVING WITH ACCELERATION

Let a particle with charge e fly out at a velocity $v \approx c$ from a conductor ($z < 0$) into a vacuum ($z > 0$), where an external field parallel to the interface $z = 0$ acts. We assume that particle velocity $\mathbf{v}(t)$ changes insignificantly during radiation formation time $1/(\omega - \mathbf{k}\cdot\mathbf{v})$. This allows us to write the law of particle motion as

$$\mathbf{r}(t) = \mathbf{v}t + \mathbf{u}t + \mathbf{w}t^2/2$$

($\mathbf{v}$ is the perpendicular and $\mathbf{u}$ is the normal velocity component, $\mathbf{v} \perp \mathbf{w}$). If a constant uniform electric field $\mathbf{E}$ acts on the particle and $\mathbf{E} \perp \mathbf{u}$, then

$$\mathbf{w} = \frac{e\mathbf{E}}{m\gamma}.$$

The boundary conditions at the surface of the conductor $z = 0$ can be satisfied by the introduction of a fictitious charge-image $-e$ moving by the law

$$\mathbf{r}_1(t) = -\mathbf{v}t + \mathbf{u}t + \mathbf{w}t^2/2.$$

The field outside the conductor then coincides with the field of these two charges when they move in the vacuum. The angular and frequency distribution of radiated energy has the form

$$\frac{d^2E}{d\omega d\Omega} = \frac{e^2\omega^2}{\pi^2c^3} \times \left| \int_0^\infty dt \exp\{i(\omega - \mathbf{k}\cdot\mathbf{u})t - i\mathbf{k}\cdot\mathbf{w}t^2/2\} \times \{[\mathbf{n}\times\mathbf{v}]\cos(\mathbf{k}\cdot\mathbf{v}t)\} - \{it[\mathbf{n}\cdot(\mathbf{u} + \mathbf{w}t)]\sin(\mathbf{k}\cdot\mathbf{v}t)\} \right|^2. \quad (1)$$

The action of the field can be substantial if field-independent terms in part cancel each other. For instance, in the ultrarelativistic case, the difference

$$\omega - \mathbf{k}\cdot\mathbf{v} - \mathbf{k}\cdot\mathbf{u} \sim \omega/\gamma^2 \ll \omega$$

is small in the exponential functions $\exp\{i(\omega - \mathbf{k}\cdot\mathbf{v} - \mathbf{k}\cdot\mathbf{u})t - i\mathbf{k}\cdot\mathbf{w}t^2/2\}$. If the main terms do not cancel each other, external field effects can always be ignored in a first approximation. This allows (1) to be rewritten as

$$\frac{d^2E}{d\omega d\Omega} = \frac{e^2\omega^2}{4\pi^2c^3}\left|[\mathbf{n}\cdot(\mathbf{v}-\mathbf{u})]\right.$$
$$\left.\times\int_0^\infty dt\exp\{i(\omega-\mathbf{k}\cdot\mathbf{v}-\mathbf{k}\cdot\mathbf{u})t - i\mathbf{k}\cdot\mathbf{w}t^2/2\}\right|^2. \quad (2)$$

The integral in (2) does not reduce to elementary functions and can be expressed via the Fresnel integrals

$$S(x) = \sqrt{\frac{2}{\pi}}\int_0^x \sin t^2 dt,$$
$$C(x) = \sqrt{\frac{2}{\pi}}\int_0^x \cos t^2 dt. \quad (3)$$

The integration gives

$$\int_0^\infty dt\exp\{(\omega-\mathbf{k}\cdot\mathbf{v}-\mathbf{k}\cdot\mathbf{u})t - i\mathbf{k}\cdot\mathbf{w}t^2/2\}$$
$$= \sqrt{\frac{\pi}{|\mathbf{k}\cdot\mathbf{w}|}}\exp\left\{-\frac{i(\omega-\mathbf{k}\cdot\mathbf{v}-\mathbf{k}\cdot\mathbf{u})^2}{2\mathbf{k}\cdot\mathbf{w}}\right\}$$
$$\times\left\{\left[\frac{1}{2}+\mathrm{sgn}(\mathbf{k}\cdot\mathbf{w})\frac{C(\omega-\mathbf{k}\cdot\mathbf{v}-\mathbf{k}\cdot\mathbf{u})}{\sqrt{|2\mathbf{k}\mathbf{w}|}}\right]\right. \quad (4)$$
$$\left. - i\,\mathrm{sgn}(\mathbf{k}\cdot\mathbf{w})\left[\frac{1}{2}+\mathrm{sgn}(\mathbf{k}\cdot\mathbf{w})S\left(\frac{\omega-\mathbf{k}\cdot\mathbf{v}-\mathbf{k}\cdot\mathbf{u}}{\sqrt{|2\mathbf{k}\mathbf{w}|}}\right)\right]\right\}.$$

The angular and frequency distribution of transient radiation has the form

$$\frac{d^2E}{d\omega d\Omega} = \frac{e^2\omega^2}{\pi c^3}\frac{[\mathbf{n}\times\mathbf{v}]^2}{|\mathbf{k}\cdot\mathbf{w}|}$$
$$\times\left\{\left[\frac{1}{2}+\mathrm{sgn}(\mathbf{k}\cdot\mathbf{w})\frac{C(\omega-\mathbf{k}\cdot\mathbf{v}-\mathbf{k}\cdot\mathbf{u})}{\sqrt{|2\mathbf{k}\mathbf{w}|}}\right]^2\right. \quad (5)$$
$$\left. + \left[\frac{1}{2}+\mathrm{sgn}(\mathbf{k}\cdot\mathbf{w})S\left(\frac{\omega-\mathbf{k}\cdot\mathbf{v}-\mathbf{k}\cdot\mathbf{u}}{\sqrt{|2\mathbf{k}\mathbf{w}|}}\right)\right]^2\right\}.$$

At large x, the $C(x)$ and $S(x)$ Fresnel integrals oscillate about 1/2 with an amplitude slowly decreasing as x increases,

$$S(x) = \frac{1}{2} - \sqrt{\frac{1}{2\pi}}\frac{1}{x}\cos x^2,$$
$$C(x) = \frac{1}{2} + \sqrt{\frac{1}{2\pi}}\frac{1}{x}\sin x^2. \quad (6)$$

At small x, the Fresnel integrals rapidly increase from zero at $x = 0$ to values of the order of one at $x \sim 1$. In the vicinity of zero, we have

$$S(x) = \sqrt{\frac{2}{\pi}}\frac{x^3}{3}, \quad C(x) = \sqrt{\frac{2}{\pi}}x. \quad (7)$$

Let us introduce auxiliary functions $f(x)$ and $g(x)$ determined by the equations

$$\frac{1}{2} - S(x) = g(x)\sin(\pi x^2/2) + f(x)\cos(\pi x^2/2), \quad (8)$$

$$\frac{1}{2} - C(x) = g(x)\cos(\pi x^2/2) - f(x)\sin(\pi x^2/2). \quad (9)$$

The distribution of radiated energy [Eq. (5)] for $\mathbf{k}\cdot\mathbf{w} < 0$ takes the form

$$\frac{d^2E}{d\omega d\Omega} = \frac{e^2\omega^2}{\pi^2c^3}\frac{[\mathbf{n}\times\mathbf{v}]^2}{|\mathbf{k}\cdot\mathbf{w}|}$$
$$\times\left\{g^2\left(\frac{\omega-\mathbf{k}\cdot\mathbf{v}-\mathbf{k}\cdot\mathbf{u}}{\sqrt{|2\mathbf{k}\cdot\mathbf{w}|}}\right)\right. \quad (10)$$
$$\left. + f^2\left(\frac{\omega-\mathbf{k}\cdot\mathbf{v}-\mathbf{k}\cdot\mathbf{u}}{\sqrt{|2\mathbf{k}\cdot\mathbf{w}|}}\right)\right\}.$$

In the $0 \le x < \infty$ region, the $f(x)$ and $g(x)$ auxiliary functions can be approximated with an error smaller than 2×10^{-3} by the equations [6]

$$f(x) \approx \frac{1+0.926x}{2+1.792x+3.104x^2},$$
$$g(x) \approx \frac{1}{2+4.142x+3.492x^2+6.670x^3}. \quad (11)$$

More accurate approximations can be found in [7].

3. AZIMUTHAL RADIATION ASYMMETRY AT NORMAL INCIDENCE

If a charged particle crosses the surface of a conductor along the normal to this surface, the distribution of radiation has azimuthal symmetry in the absence of a field. An external field breaks this symmetry. The direction of radiation emission will be specified by the ϑ and φ angles of the spherical coordinate system with the axis along z. Let particle acceleration in the field be directed along the x axis. The argument of the $f(x)$ and $g(x)$ functions in (10) at normal particle incidence ($\mathbf{u} = 0$) can then be written as

$$\frac{\omega-\mathbf{k}\cdot\mathbf{v}}{\sqrt{|2\mathbf{k}\mathbf{w}|}} = \omega\frac{1-\cos\vartheta+1/2\gamma^2}{\sqrt{2w(\omega/c)\sin\vartheta\cos\varphi}}. \quad (12)$$

The limiting case in which the value of (12) tends to infinity corresponds to transient radiation in the absence of an external field. The angular and frequency

distribution of radiation [Eq. (1)] then transforms into the distribution typical of usual transient radiation,

$$\frac{d^2E}{d\omega d\Omega} = \frac{e^2}{\pi^2 c}\frac{[\mathbf{n}\times\mathbf{v}]^2}{\{1-(\mathbf{n}\cdot\mathbf{v}/c)^2\}^2}. \quad (13)$$

It follows from (12) that this limiting situation corresponds to $\varphi = \pm\pi/2$ and to $\vartheta \longrightarrow 0$. The $\vartheta \ll 1/\gamma$ region, however, makes a small contribution to the intensity of radiation and can be ignored. It follows that radiation emitted in the plane which is normal to the external field and passes through particle velocity (that is, at $\varphi = \pm\pi/2$) is fully independent of the external field strength. The intensity of radiation propagating in the plane that passes through particle velocity and field (that is, at $\varphi = 0$) can, however, substantially decrease depending on external field. The ratio between radiation intensities corresponding to $\varphi = 0$ and $\varphi = \pm\pi/2$ can be obtained in the form

$$\frac{d^2E(\varphi=0)}{d^2E(\varphi=\pi/2)} = \frac{\pi\omega^2\{1-(\mathbf{n}\cdot\mathbf{v}/c)^2\}^2}{|\mathbf{k}\cdot\mathbf{w}|} \times\left\{g^2\left(\frac{\omega-\mathbf{k}\cdot\mathbf{v}}{\sqrt{|2\mathbf{k}\cdot\mathbf{w}|}}\right)+f^2\left(\frac{\omega-\mathbf{k}\cdot\mathbf{v}}{\sqrt{|2\mathbf{k}\cdot\mathbf{w}|}}\right)\right\}. \quad (14)$$

This ratio substantially decreases as particle energy increases and the angle of radiation emission decreases. For instance, when the argument of Fresnel integrals (12) is close to one, $g^2 + f^2 \approx 1/3$, and (14) becomes

$$\frac{d^2E(\varphi=0)}{d^2E(\varphi=\pi/2)} \approx \frac{\pi\omega c}{6|\mathbf{n}\cdot\mathbf{w}|}\left(\vartheta^2+\frac{1}{\gamma^2}\right)^2, \quad (15)$$

where it is taken into account that ultrarelativistic particle radiation is concentrated at $\vartheta \ll 1$. It follows that azimuthal asymmetry of the angular distribution of transient radiation strongly depends on the energy of the particle. The presence or absence of azimuthal asymmetry shows whether or not field strength is sufficient for influencing transient radiation.

4. DISCUSSION

Equation (15) shows that, at normal incidence, azimuthal radiation asymmetry is substantial when the value of (12) is greater than unity. In a uniform external electric field **E** at angles $\vartheta \sim 1/\gamma$, this condition is met if

$$e|\mathbf{E}| \sim (\omega/\gamma^2)mc = mc^2/\lambda\gamma^2. \quad (16)$$

Transient optical range radiation at $\gamma \sim 10^4$ substantially changes in a field of about 500 V/cm. Transient radiation azimuthal anisotropy in the centimeter radio frequency band at $\gamma \sim 10^3$ appears in a field of the order of or stronger than 1 V/cm. These field values correspond to maximum azimuthal anisotropy of transient radiation. In lower fields, the azimuthal anisotropy is small.

Azimuthal anisotropy of transient radiation in an external field is fairly easy to measure, its dependence on the Lorentz factor is quadratic, and increasing energy therefore improves the conditions for measuring anisotropy in a fairly strong field. The conclusion can be drawn that measurements of azimuthal anisotropy of the angular distribution of transient radiation in an external field can conveniently be used to determine the energy of ultrarelativistic particles.

As condition (16) of maximum azimuthal anisotropy actually depends on the ratio between the external field strength and radiation frequency ($|\mathbf{E}|/\omega$), convenient external field conditions can be attained by varying the frequency of radiation.

Note that we only considered transient radiation formed close to the surface of a conductor along particle path of the order of the coherence length. The action of a field on a particle also causes radiation along further particle trajectory. This radiation is, however, not related to particle crossing of the interface; this is usual radiation caused by particle movement in an external field. Such radiation is determined by the specific character of further particle movement in an external field, and its contributions can be various. In comparison with experiment, this radiation should be taken into account, but including it in a general consideration is inexpedient. Indeed, such radiation depends on the velocity of a particle rather than its energy, and its distribution strongly depends on the conditions of further particle movement in the measuring device.

REFERENCES

1. V. L. Ginzburg and I. M. Frank, Zh. Éksp. Teor. Fiz. **16**, 15 (1946).
2. M. L. Ter-Mikaelian, *High Energy Electromagnetic Processes in Condensed Media* (Akad. Nauk Arm. SSR, Yerevan, 1969; Wiley, New York, 1972).
3. G. M. Garibyan and Yan Shi, *X-ray Transient Radiation* (Akad. Nauk Arm. SSR, Yerevan, 1983).
4. V. L. Ginzburg and V. N. Tsytovich, *Transient Radiation and Transient Scattering* (Nauka, Moscow, 1984).
5. B. Dolgoshein, Nucl. Instrum. Methods Phys. Res. A **326**, 434 (1993).
6. *Handbook of Mathematical Functions*, Ed. by M. Abramowitz and I. A. Stegun (National Bureau of Standards, Washington, 1964; Nauka, Moscow, 1979).
7. J. Boersma, Math. Comput. **14**, 380 (1960).

Translated by V. Sipachev

Journal of Experimental and Theoretical Physics, Vol. 95, No. 5, 2002, pp. 864–877.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 122, No. 5, 2002, pp. 1003–1018.

PLASMA, GASES

Interaction of Charged Dust Particles in Clouds of Thermodynamically Equilibrium Charges

V. A. Gundienkov and S. I. Yakovlenko*
Institute of General Physics, Russian Academy of Sciences, ul. Vavilova 38, Moscow, 119991 Russia
**e-mail: syakov@kapella.gpi.ru*
Received March 22, 2002

Abstract—The solution of the Poisson–Boltzmann equation for a cloud of charges surrounding two charged dust particles treated as Debye atoms forming a Debye molecule is investigated numerically using Cassini coordinates. The electric force exerted on a dust particle by the other dust particle was determined by integrating the electrostatic pressure on the surface of the dust particle. It is shown that attractive forces appear when the following two conditions are satisfied. First, the Debye radius (corresponding to the electron density at half the mean distance between the dust particles) must be approximately equal to half the mean distance between the dust particles. Attraction between the dust particles emerges at a distance equal approximately to half the mean distance between the dust particles. Second, attraction takes place when like charges are concentrated predominantly on the dust particles. If the particles carry a small fraction of charge of the same polarity, repulsion between the particles takes place at all distances. © *2002 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

The study of a plasma in which charged particles of micrometer size play a significant role (so-called dust plasma) is interesting from the fundamental and applied points of view (see [1, 2] and the literature cited therein). It is especially interesting because of the collective effects observed in a dust plasma due to its nonideality [3–6]. The properties of a nonideal plasma are often considered in the so-called one-component approximation. In this case, one of the charges is sort of smeared uniformly over the space, and polarization effects are taken into account in the form of corrections in some cases.

Apparently, the physics of processes occurring in a dust plasma is basically different. The major object of investigation must be dust particles surrounded by clouds of charged particles with masses much smaller than the dust particle mass. A charged dust particle surrounded by a cloud of charges of the opposite sign is an analogue of an atom in gas kinetics. Generally speaking, the charge shell of such a "dust atom" may not be in thermodynamic equilibrium. However, we will consider the situation when the charge distribution in the shell is of the Boltzmann type. Such a dust atom can naturally be referred to as a Debye atom [7] in contrast to a Thomas–Fermi atom in which the charge shell is formed by a degenerate electron gas. Similarly, we can introduce the concept of a Debye molecule [8, 9] and a Debye crystal. The properties of such Debye systems are determined analytically by the Boltzmann distribution and the Poisson equation, i.e., by the Poisson–Boltzmann equation.

According to the results of a number of experiments (see, for example, [3–6]), dust particles of a micrometer size in thermionic plasmas, gas-discharge plasmas, and nuclear-excited plasmas may form 3D structures. Consequently, it is natural to assume the presence of attractive forces emerging due to polarization of charge shells of Debye atoms. However, convincing theoretical results demonstrating the attraction of Debye atoms have not been obtained so far. The exact solution of the Poisson–Boltzmann equation shows that charged planes in an electron cloud as well as in a plasma always exhibit repulsion rather than attraction [10, 11]. Numerical calculations of the forces of interaction between Debye atoms [8, 9], like the results of approximate analysis [12, 13], were not quite reliable until recently.

The problem of interaction of particles in a dust plasma is close to the problem of interaction of colloidal particles in electrolytes. The very concept of Debye radius for plasmas was borrowed from the theory of electrolytes. However, in spite of the long history of investigation into the physics of colloidal particles in electrolytes [10], the origin of attractive forces in electrolytes has not been clarified completely (at least, for the case when the diameter of a colloidal particle is smaller than the Debye radius; see, for example, [14–18]).

Below, we make an attempt to demonstrate reliably the existence of polarization forces of attraction between Debye atoms and to determine the conditions under which attraction appears. This work differs methodically from other publications devoted to an analysis of interaction of charged dust particles in plasmas and in electrolytes.

1063-7761/02/9505-0864$22.00

First, in contrast to a number of publications (see, for example, [10, 12–18]), we consider the situation when the total charge of dust particles is not negligibly small as compared to the total charge of the plasma particles (of the same polarity) surrounding dust particles. Moreover, it will be shown (see [19] for preliminary results) that attraction is most significant in the opposite limiting case, i.e., when almost the entire charge of one polarity is concentrated in dust particles and the clouds, accordingly, contain charges of the same (opposite) sign.

Second, in an analysis of the properties of a Debye molecule, we rely mainly on the fact that a Debye atom has a definite structure. In particular, the charge of a dust particle cannot be considered approximately as a delta function as a rule even if its radius is much smaller than the Debye radius.

Third, we calculate directly the resultant force exerted on a dust particle by another dust particle and by the charge shell rather than the potential energy of the system. The dependence of the energy of interaction between dust particles on their separation is determined by the integration of this force. The Poisson–Boltzmann equation is solved in our case in an infrequently used system of coordinate based on Cassinian ovals [20, 21]. This enables us to calculate the field strength in the vicinity of the surface of a small dust particle to a high degree of accuracy and to determine reliably the force acting on the dust particle.

2. FORMULATION OF THE PROBLEM

2.1. Poisson–Boltzmann Equation

For the sake of definiteness, we will consider below a thermal emission plasma and speak of positively charged dust particles and of the electron shell of charges. However, the main results are also valid for an electric-discharge dust plasma as well as a plasma ionized by an external source of hard radiation, when dust particles bear a negative charge and charge shells consist predominantly of positive ions.

Let the electron gas surrounding charged particles be formed due to emission of electrons from dust particles having a rather high temperature T. In addition, dust particles are in partly ionized gas. In order to find the spatial distribution of potential ϕ, the field strength $-\nabla\phi$, and the charge density $\rho = e(N_i - N_e)$, we must solve the Poisson equation

$$\nabla(-\nabla\phi) = 4\pi\rho.$$

In this equation, the densities N_i and N_e of ions and electrons are determined by the Boltzmann distribution

$$N_i = N_{i0}\exp(-e\phi/T), \quad N_e = N_{e0}\exp(e\phi/T),$$

where N_{i0} and N_{e0} are the densities of ions and electrons at the points where the potential is equal to zero, and ∇ is the Hamilton vector operator.

Thus, the Poisson–Boltzmann equation has the form

$$\Delta\phi = 4\pi e(N_{e0}\exp(e\phi/T) - N_{i0}\exp(-e\phi/T)), \quad (1)$$

where $\Delta = \nabla^2$ is the Laplace operator; the temperature of particles and plasma is assumed to be the same.

2.2. Dimensionless Quantities

We will measure the length in units of the Debye radius $r_D = (T/4\pi e^2 N_{e0})^{1/2}$ corresponding to the electron density at the points where the potential is equal to zero. We introduce dimensionless quantities (potential φ, electric field strength $\mathbf{E}$, and the electron density n_e) with the help of relations

$$\varphi = \frac{\phi e}{T}, \quad \mathbf{E} = -\frac{\nabla\phi e r_D}{T}, \qquad n_e = r_D^3 N_e = n_D\exp\varphi, \quad (2)$$

where $n_D = r_D^3 N_{e0}$.

In our estimates below, we will proceed, as a rule, from the experimental conditions [3], in which $N_{e0} = 2.5 \times 10^{10}$ cm^{-3} and $T = 0.146$ eV = 1700 K; for the characteristic quantities, we have $r_D = 18$ μm, $T/e = 0.146$ V, and $T/er_D = 80$ V/cm. For the mean radius of a dust particle $r_p = 0.4$ μm ($r_0 = r_p/r_D = 2.23 \times 10^{-2}$) and its charge $Z_p e = 500e$, we have the field strength on the particle surface $Z_p e/r_0^2 = 4.5 \times 10^4$ V/cm ($E_0 = E(r_0) = 550$).

For the dimensionless quantities, Eq. (1) can be reduced to the following equations for the dimensionless potential φ:

$$\Delta\varphi = \exp\varphi - \delta\exp(-\varphi), \qquad \nabla\mathbf{E} = -(\exp\varphi - \delta\exp(-\varphi)), \quad \mathbf{E} = -\nabla\varphi. \quad (3)$$

Here, $\delta = N_{i0}/N_{e0}$ is the parameter characterizing additional ionization of the gas. In view of the quasi-neutrality of the plasma, $0 \le \delta \le 1$.

2.3. Boundary Conditions

Following [7], we will use the term "Debye atom" for a single charged dust particle surrounded by a cloud of lighter charges in thermodynamic equilibrium; two or more dust particles will be referred to as a Debye molecule [8, 9]. Formally, the analyses of a Debye atom and a Debye molecule differ only in the geometry of the problem. While analyzing a Debye atom, we can get by with the solution of the one-dimensional Poisson equation, assuming that the electron cloud is spherically symmetric. In an analysis of a diatomic Debye molecule, we can assume that the problem is symmetric about the x axis connecting the nuclei (dust particles). In this case, it is sufficient to consider the two-dimensional equation (3) in the coordinate plane xy. As we go over to an analysis of a Debye molecule, the problem is

complicated considerably due to the choice of the boundary conditions.

In the real physical problem, the charge $Z_p e$ of dust particles and their radius r_p are specified (the formation of the dust particle charge is described in [22–24]). Consequently, one of the boundary conditions is that imposed on the field strength on the surface S of dust particles:

$$\mathbf{E}_0 = -\nabla\varphi|_S. \tag{4}$$

The particle charge in this case is defined as

$$Z_p = \frac{-r_D^2}{4\pi e}\int_S \nabla\phi d\mathbf{s}, \quad z_p = \frac{1}{4\pi}\int_S \mathbf{E} d\mathbf{s}. \tag{5}$$

Here, z_p is the dimensionless charge of a particle, which is connected with the charge Z_p of the particle in units of electron charge through the relation

$$Z_p = 4\pi z_p n_D;$$

the area of the surface is measured in squares of the Debye radius.

In the second boundary condition, we must define the surface S' on which the field is equal to zero:

$$\nabla\varphi|_{S'} = 0. \tag{6}$$

The zero value of the electric field at the boundary follows from the quasi-neutrality of the system of charges under investigation. Surface S' determines the boundary of the Debye system under investigation.

In an analysis of a Debye molecule, we mainly aim at determining the resultant electrostatic force acting on the dust particles as a function of the distance d between them. In this case, it is more convenient to proceed from other boundary conditions [8, 9]. We specify not the field, but the constant potential on the surface of dust particles,

$$\varphi|_S = \varphi_0 = \text{const}.$$

The field strength $\mathbf{E}_0$ on the surface of a dust particle is determined from the solution of the Poisson–Boltzmann equation. The resultant force is determined with the help of the integral of electrostatic pressure over the surface of a dust particle. In order to obtain the required value of charge z_p (5), we must change appropriately the value of potential φ_0.

In the case under investigation, the force of interaction between dust particles is directed along the z axis and is defined as

$$F = \frac{1}{8\pi}\int_S (\nabla\phi)^2|_S ds_z, \quad f = \int_S E_0^2 ds_z. \tag{7}$$

Here, ds_z is the component of the surface area element $d\mathbf{s}$ along the z axis, the force F is connected with the dimensionless force f through the relation

$$F = (T^2/8\pi e^2) f,$$

and the electric pressure is directed along the outward normal to the surface of dust particles.

3. SOME PROPERTIES OF DEBYE ATOMS

3.1. Debye Atom

3.1.1. Poisson–Boltzmann equation. The properties of a Debye molecule are mainly determined by the properties of the Debye atoms constituting this molecule. In particular, for a large distance between dust particles, a Debye molecule must decay into two independent Debye atoms. We will use this fact below. For this reason, before calculating the force of interaction, we consider some properties of Debye atoms (see also [25]).

In the one-dimensional (i.e., in a planar, cylindrically, or spherically symmetric) case, Eq. (3) assumes the form

$$\frac{1}{r^k}\frac{d}{dr}\left(r^k\frac{d\varphi}{dr}\right) = \exp\varphi - \delta\exp(-\varphi),$$
$$\varphi|_{r=a_0} = 0, \tag{8}$$
$$E(r)|_{r=a_0} \equiv -\left.\frac{d\varphi}{dr}\right|_{r=a_0} = 0.$$

Here, $k = 0, 1, 2$ for the planar, cylindrically, and spherically symmetric cases, respectively; depending on the geometry, point $r = 0$ corresponds to the beginning of a planar layer, the center of the cylinder, or the center of the sphere. One of the boundary conditions specifies the boundary $r = a_0$ of the Debye atom, at which the field is equal to zero.

We will consider below the spherically symmetric case (when $k = 2$), which simulates a Debye atom, and the planar case enabling us to consider the variation of the potential near the dust particle surface [6–8]. In the spherically symmetric case, a convenient characteristic of a Debye atom is the dimensionless charge contained in a sphere of radius r, which is defined as

$$z(r) = r^2 E(r).$$

3.1.2. Debye atom in a cloud of like charges. The case $\delta = 0$, when charge shells are formed by particles with charges of the same polarity, corresponds, for example, to a thermionic plasma [3] or a gas ionization such that the charge of one polarity is completely concentrated on dust particles (see, for example, [24]). We choose the value of a_0 equal to half the mean distance between dust particles:

$$a_0 = a_p \equiv (N_p^{-1/3}/2r_D),$$

where N_p is the dust particle density (see Fig. 1). The value of $N_p^{-1/3}/2$ is smaller than the Wigner–Seitz radius $r_{WS} \equiv (4\pi N_p/3)^{-1/3}$ by 24%.

We will consider the most interesting situation when the dust particle radius r_p is much smaller than the separation $r_0 = r_p/r_D \ll a_0$ between dust particles. In experiments [3], $N_p = 5 \times 10^7$ cm^{-3}; accordingly, $a_p = 0.755$, and the conditions of the smallness of the dust particle radius are satisfied:

$$a_p r_D/r_p = a_0/r_0 = 34.$$

The results of analysis of Eq. (8) for the spherically symmetric case ($k = 2$) show [25, 26] that, for a moderate charge

$$z_p \equiv Z_p e^2/r_D T < a_0^3/3$$

of a small particle, $r_0 \ll a_0$, the distributions of charge, field, and potential around a dust particle are defined by the relations

$$z(r) = \frac{a_0^3}{3}\left[1 - \left(\frac{r}{a_0}\right)^3\right], \quad E(r) = \frac{z(r)}{r^2},$$
$$\varphi(r) = \frac{a_0^3}{3}\left[\frac{a_0}{r} - \frac{3}{2} + \left(\frac{r}{2a_0}\right)^2\right]. \qquad (9)$$

For a large charge $z_p > a_0^3/3$ far away from the dust particle surface, for

$$r_0 - 3r_0^2/a_0^3 > r > a_0,$$

expressions (9) are valid as before. A change in the dependences is observed near the surface ($r < r_0 - 3r_0^2/a_0^3$), where a sharp decrease in $z(r)$, $E(r)$, and $\varphi(r)$ is observed (see Fig. 2 and [25]). In other words, for a large charge of a dust particle, the Debye radius of the atom has a certain core formed by the charge shell in the vicinity of the surface of the dust particle. The charge of the particle together with the core is $z_{cor} \equiv a_0^3/3$. The screening of this "remaining" charge occurs at a large distance r close to a_0.

The condition of the large particle charge

$$z_p \equiv Z_p e^2/r_D T > z_{cor}$$

can be written for the particle charge measured in units of electron charge:

$$Z_p > Z_{cor} \equiv \frac{\pi N_{e0}}{6N_p}.$$

According to the results of measurements [3], the charge of dust particles is quite large:

$$Z_p = 500 > Z_{cor} = 262, \quad z_p = 0.273 > z_{cor} = 0.143.$$

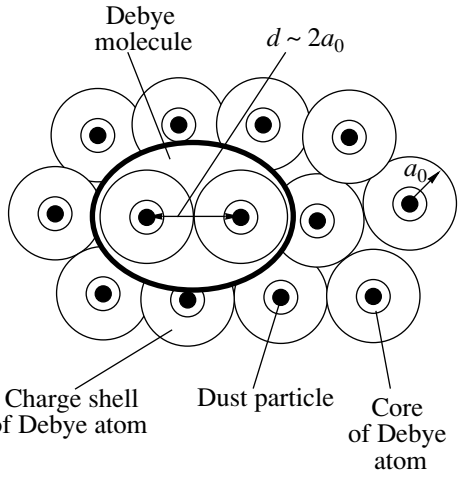


Fig. 1. Schematic diagram of an aggregate of Debye atoms and a Debye molecule.

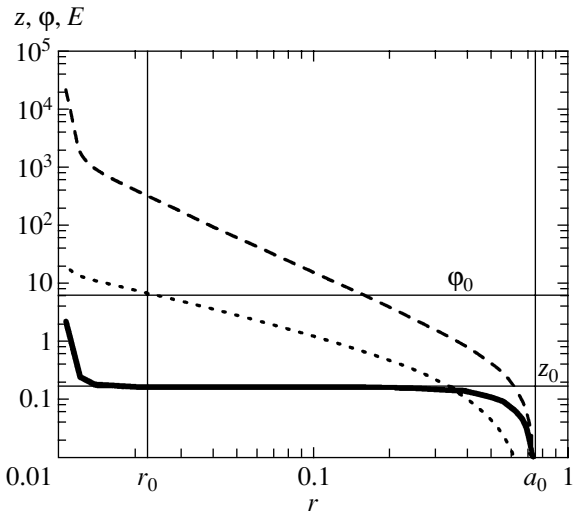


Fig. 2. Charge z (solid curve), field strength E (dotted curve), and potential φ (dashed curve) as a function of the distance r to the particle center, measured in units of the Debye radius ($\delta = 0$). The radius of a Debye atom $a_0 = a_p \equiv N_p^{-1/3}/2r_D = 0.755$ is chosen for the experimental conditions [3]: $T = 1700$ K, $N_p = 5 \times 10^7$ cm^{-3}, $N_{e0} = 2.5 \times 10^{10}$ cm^{-3}, $r_D = 18$ μm, $r_p = 0.4$ μm, $r_0 = r_p/r_D = 2.23 \times 10^{-2}$.

Calculations show, however (see Fig. 2), that for values of electron density and temperature measured in [3], the charge of dust particles of the given radius in the thermal equilibrium state must be $Z_p = 286$ ($z_p = 0.156$), which is smaller than the measured value $Z_p \approx 500$. Consequently, either the measured values of plasma parameters are considerably inaccurate, or the dust particle charge is nonequilibrium under the experimental conditions [3] (see also [24]).

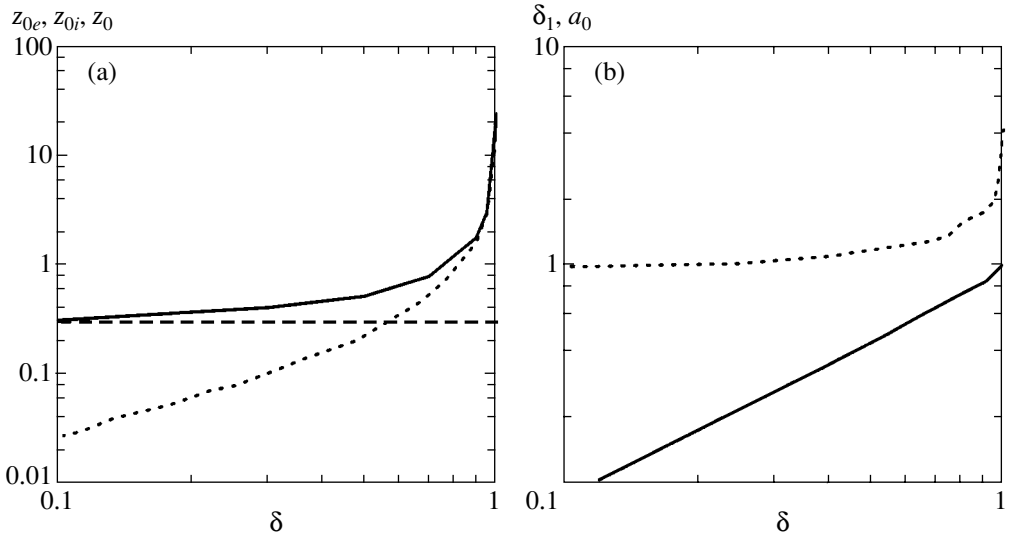


Fig. 3. Relation between the parameters characterizing a dust plasma and the parameters of a Debye atom: (a) dependence of the dimensionless charge z_{0e} of the electron shell (solid curve) and z_{0i} of the ion shell (dotted curve), as well as of the uncompensated charge of the shell $z_0 = z_{0e} - z_{0i}$ (dashed curve) on parameter δ; (b) dependence of the parameter δ_1 characterizing the ratio of the total number of ions in a Debye atom to the number of electrons (solid curve) and of the radius a_0 of a Debye atom (dotted curve) on the parameter δ characterizing the ratio of the number of ions to the number of electrons at the boundary of a Debye atom. The parameters of a dust particle are $r_0 = 0.1$, $\varphi(r_0) = 2.4$, $z(r) = 0.28$.

3.1.3. Debye atom in a plasma. In the case $\delta \neq 0$, when charge clouds consist of particles of both polarities, the radius of a Debye atom is defined, as before, as the distance $r = a_0$ at which the charge of a dust particle is completely compensated by free charges of the plasma ($E(a_0) = 0$). As in the case $\delta = 0$, the radius of a Debye atom is equal to half the mean distance between dust particles. For $\delta = 1$, we can consider an isolated dust particle in an infinitely large volume of the plasma. For $\delta \longrightarrow 1$, the radius of a Debye atom tends to infinity ($a_0 \longrightarrow \infty$). As a matter of fact, the finite charge z_0 of a particle may be compensated completely by a quasi-neutral plasma only if the plasma size is infinitely large. If $\delta < 1$, the radius of a Debye atom is finite.

The dimensionless charges of electrons and ions contained in the charge shell are defined as

$$z_{0e} = \int_{r_0}^{a_0} \exp(\varphi(r)) r^2 dr,$$
$$z_{0i} = \delta \int_{r_0}^{a_0} \exp(-\varphi(r)) r^2 dr, \quad z_0 = z_{0e} - z_{0i}. \tag{10}$$

The quantity $\delta_1 \equiv z_{0i}/z_{0e}$ gives the ratio of the free charge of ions in a Debye atom to the electron charge. Generally speaking, the value of δ_1 must be a complex function of parameters δ, a_0, and φ_0. However, in the cases when the main contribution to integration in expressions (10) comes from the region of small values of potential $\varphi(r) \ll 1$, we can assume approximately that $\delta_1 \approx \delta$.

Figure 3 illustrates the dependences of z_{0e}, z_{0i}, and δ_1 on δ. In the results presented in Fig. 3, the value of a_0 for different values of δ was chosen as large as possible for the radius of a dust particle corresponding to the experiments [3]:

$$r_0 = r_p/r_D = 2.23 \times 10^{-2}.$$

This was carried out by "test firing": when the value of a_0 was chosen greater than that in Fig. 3, the particle charge becomes infinitely large ($z(r_0) \longrightarrow \infty$). The obtained dependences $z(r)$ and $\varphi(r)$ were used for determining $z_0 = z(r_0)$ and $\varphi_0 = \varphi(r_0)$ for $r_0 = 0.1$ (see also Fig. 4).

It can be seen that, as the value of δ increases, the number of both positive and negative charges in the shell of a Debye atom increases due to an increase in its volume (see Fig. 3). At the same time, the number of uncompensated charges $z_0 = z_{0e} - z_{0i}$ remains virtually unchanged upon a change in δ. In the range of parameters under investigation, $\delta_1 \approx \delta$.

As in the case of $\delta = 0$, for a given value of r_0, the value of a_0 cannot be infinitely large for an indefinitely large value of the particle charge z_0. For a large value of $E_0 = z_0/r_0^2$, at a distance $r - r_0 \sim 1/E_0$ from the particle surface, a sharp decrease in the dependences $z(r)$, $E(r)$, and $\varphi(r)$ is observed due to the screening by charges of the opposite polarity (see Fig. 4). In this case, the value of a_0 is bounded by a certain limiting value $a_{0\max} \equiv a_0$

($E_0 \longrightarrow \infty$). This limiting value increases logarithmically for $\delta \longrightarrow 1$:

$$a_{0\max} \approx \frac{1}{2}\ln\frac{1}{1-\delta} + \frac{1}{2}$$

when $0.9 \le \delta \le 0.999$ (see [25]).

Since a Debye atom has a core screening the charge of the dust particle, we cannot ascribe the unscreened value of the charge to the dust particle while considering the interaction of Debye atoms.

3.2. On the Interaction of Dust Particles

3.2.1. Interaction of nonpolarized particles. If we imagine a situation when the shells of Debye atoms separated by distance d do not interact with one another, attractive forces cannot emerge between dust particles. Only repulsive forces will act between them. Indeed, the force of interaction between nonpolarized shells can be expressed in the form

$$f(d) = z_{\text{eff}}(d)z_p/d^2.$$

Here, $z_{\text{eff}}(d) = E(d)d^2$ is the total charge within the sphere of radius d surrounding a dust particle (uncompensated part of the particle charge). In view of quasineutrality of a Debye atom, $z_{\text{eff}}(r) \ge 0$ for $r > r_0$. The charges of the same sign repel one another:

$$z_{\text{eff}}(d)z_p \ge 0.$$

For the emergence of attractive forces, a rearrangement (polarization) of charge shells is required. In this case, the number of charges attracting dust particles to the center of a Debye molecule must increase on its axis.

3.2.2. Interaction of charged planes. The Poisson–Boltzmann equation (4) in the planar case ($k = 0$) has a solution in quadratures. This enables us to consider the force of interaction between planes and to determine the requirements to the accuracy of the solution to this equation in the vicinity of the surface of a dust particle.

An analysis shows that the electrostatic interaction between the planes surrounded by a cloud of like charges as well as between the planes located in a plasma leads to the repulsion between the planes [10, 11]. By way of an illustration, we consider the case of $\delta = 0$, which allows us to obtain simple analytic expressions required for estimating the requirements to the accuracy of calculation of the field and potential near the surface of a dust particle.

We consider the electrostatic pressure exerted on a charged conducting plane located between two conducting (left and right) planes with the same charge density. If necessary, one of the planes can be removed to an infinitely large distance.

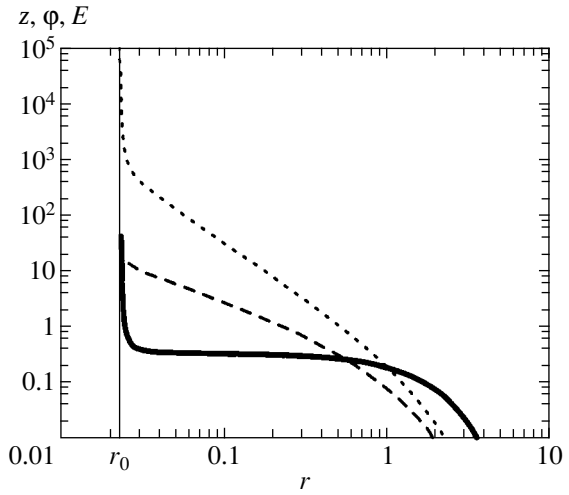


Fig. 4. Dependence of charge z (solid curve), electric field strength E (dotted curve), and potential φ (dashed curve) on the distance r to the center of a particle, measured in units of the Debye radius ($\delta = 0.999$). The radius of a microscopic particle $r_0 = 2.23 \times 10^{-2}$ is chosen for the experimental conditions [2].

Integration of the Poisson–Boltzmann equation for the planar case gives [11]

$$\varphi(x) = \ln(E^2 + E_1^2), \quad E(x) = E_1\tan[(a_0 - x)E_1/2].$$

The quantities $E_1 \equiv \exp(\varphi_1/2)$ and φ_1 are connected with a_0 through the relation

$$a_0 = \frac{2}{E_1}\arctan\frac{E_0}{E_1}.$$

Here, x is the distance to the plane under investigation, which is assumed for simplicity to be infinitely thin, and φ_1 is the value of the potential at point a_0, where the field strength is equal to zero. For identical charge densities on the planes, the value of a_0 is equal to half the distance between the planes.

Although the potential on the left and right of the conducting plane is the same ($\varphi(-0) = \varphi(0) = \varphi_0$), the field strengths on the surface under investigation on the left ($E(-0) = E_{01}$) and right ($E(0) = E_{02}$) sides are different. This gives rise to the electrostatic pressure p on the plane:

$$p = E_{02}^2 - E_{01}^2.$$

The quantity a_0 is a monotonically decreasing function of E_1. Accordingly, $E_{01} > E_{02}$ and $p < 0$ if, for example, the distance to the left plane $2a_{01}$ is larger than the distance $2a_{02}$ to the right plane. In other words, the resultant force of pressure is directed towards the more remote plane. In particular, if one of the planes is removed to an infinitely large distance, the remaining two planes repel each other.

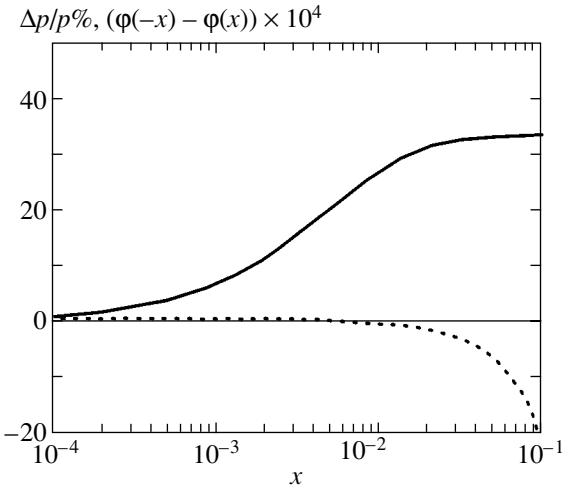


Fig. 5. Dependence of the error in determining pressure (see relation (11)) exerted on a conducting plane and of the difference in potentials on the left and right of the plane on the distance x to the plane. The plane under investigation is between two other charged planes; all the planes are under the potential $\varphi_0 = 10$. Half the distance to the left plane is $a_0 = 6.27$, and half the distance to the right plane is $a_0 = 2.08$; $p = 2$ in this case.

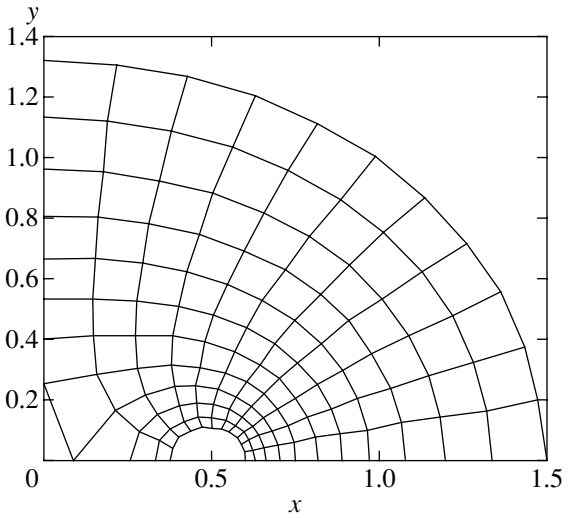


Fig. 6. Cassini coordinates for the distance $d = 1$ between the foci, which corresponds approximately to the transition from attraction to repulsion at $a_0 \approx 1$.

Thus, attraction between dust particles can emerge only in a nonplanar geometry.

3.2.3. On the accuracy of computation of potential near the surface. In numerical integration of the Poisson–Boltzmann equation, the value of the field strength is determined at points of the mesh on which the difference scheme is constructed. The value of E_0 determined approximately corresponds to the value of the field at a distance of the order of mesh spacing from the dust particle surface. Let us find the error in the pressure being determined resulting from inaccuracy in determining the position corresponding to E_0. The relative error in pressure determined from points separated from the plane by distances x and $-x$ in the planar geometry is given by

$$\frac{\Delta p}{p} = \frac{\left|p - (E^2(-x) - E^2(x))\right|}{p}. \tag{11}$$

It can be seen from Fig. 5 that, if the potential of the plane is significant ($\varphi_0 \gg 1$), the value of $\Delta p/p$ amounts to tens of percent even at small distances $x \sim 0.01$, while the difference $\varphi(-x) - \varphi(x)$ between potentials on the left and right is virtually equal to zero. In other words, numerical integration requires a very high accuracy in determining the derivative of the potential near the particle surface, which dictates a very small mesh size in the vicinity of the surface.

At the same time, in order to determine the magnitude of the force acting on a particle, the method of solving the Poisson–Boltzmann equation must ensure the maximum accuracy just in the region near the surface of dust particles. We are mainly interested in distances between dust particles exceeding considerably their diameter. In ordinary systems of coordinates, it is difficult to attain a sufficient accuracy in calculation of the force acting on small dust particles.

4. METHOD OF SOLUTION OF TWO-CENTER PROBLEM

4.1. Cassini Coordinates

We used orthogonal coordinates constructed on the basis of the well-known Cassinian oval [20, 21] for a certain particular case.

The relation between variables u and v defining a point on the Cassinian oval with Cartesian coordinates in the quadrant $x > 0$, $y > 0$ is determined by the following expressions:

$$x(u, v) = \frac{d}{2\sqrt{2}} \times \sqrt{\sqrt{\exp 2u + 2\exp u\cos v + 1} + \exp u\cos v + 1}, \tag{12a}$$

$$y(u, v) = \frac{d}{2\sqrt{2}} \times \sqrt{\sqrt{\exp 2u + 2\exp u\cos v + 1} - \exp u\cos v - 1}. \tag{12b}$$

For the entire zy plane, the coordinate mesh is obtained by the mirror reflection relative to the z and y axes (d is the distance between the foci of the ovals located at points $(-d/2, 0)$ and $(d/2, 0)$). Variable $\infty > u > -\infty$ is a certain analogue of the radial variable. For $u < 0$, the curve has the shape of two independent ovals; for $u = 0$, the coordinate line is a Bernoulli lemniscate, i.e., an oval with an infinitely slim "waist." For $0.65 > u > 0$, we have an oval with a waist, while for $u > 0.65$, the oval

has an ellipsoidal shape. Variable $\pi > v > 0$ is analogous to an angle in polar coordinates. For $v = 0$, the point lies on the ray $(d/2, \infty)$ along the abscissa axis; for $v = \pi$, the locus approaches the angle formed by the segment $(0, d/2)$ on the abscissa axis and the ray $(0, \infty)$ along the ordinate axis. The shape of coordinate lines is illustrated in Fig. 6. The application of coordinates (12) provides the following important advantages. First, the family of Cassinian ovals corresponds qualitatively to the pattern of equipotential surfaces for two like-charged particles located at the foci of the oval. Second, the domain of the solution of the equation in these coordinates becomes rectangular. Third, the density of ovals increases exponentially towards the surface of dust particles. This allows us to use a homogeneous mesh even for large distances between small particles.

4.2. On the Method of Solution

Without going into details, we describe the main feature of the method of solution. It is convenient to use Cassini coordinates in the situation when the radius r_0 of dust particles is much smaller than both the Debye radius ($r_0 \ll 1$) and the radius of a Debye atom ($r_0 \ll a_0$). Considering the range of large distances between particles ($d > 5r_0$), we can treat small dust particles as Cassinian ovals close to circles. On a small oval, it is convenient to define the value of potential φ_0. At the same time, the cloud of charges enveloping dust particles can be described by an ellipsoidal oval for $d < 5a_0$. It is convenient to define the zero value of the field on this oval.

The surface of a dust particle and the outer surface corresponding to the boundary of the Debye molecule (on which the field vanishes) are described in coordinates (8) by the constants

$$u_{\min} = \ln\left(\frac{4r_0}{d^2}(d + r_0)\right),$$
$$u_{\max} = \ln\left(\frac{4a_0}{d^2}(d + a_0)\right). \quad (13)$$

The boundary conditions (2) in this case have the form

$$\varphi|_{u = u_{\min}} = \varphi_0, \quad \left.\frac{\partial\varphi}{\partial u}\right|_{u = u_{\max}} = 0. \quad (14)$$

The Poisson–Boltzmann equation (3) with the boundary conditions (13), (14) was solved by the Gauss–Newton method of iterations using the program packet MATLAB.

The variation of the potential in Cassini coordinates and in Cartesian coordinates is illustrated in Fig. 7.

In order to calculate the charge and the force of interaction of dust particles, we considered 3D coordinates formed by rotating 2D coordinates (8) about the x axis. In these coordinates, we calculated the force of interaction between dust particles by formula (7). The energy of interaction of dust particles was calculated by the formula

$$U(d) = \int_d^\infty f(x)dx + \text{const}. \quad (15)$$

The constant was usually chosen so that the energy $U(d)$ at the minimum was equal to zero,

$$\min(U(d)) = 0.$$

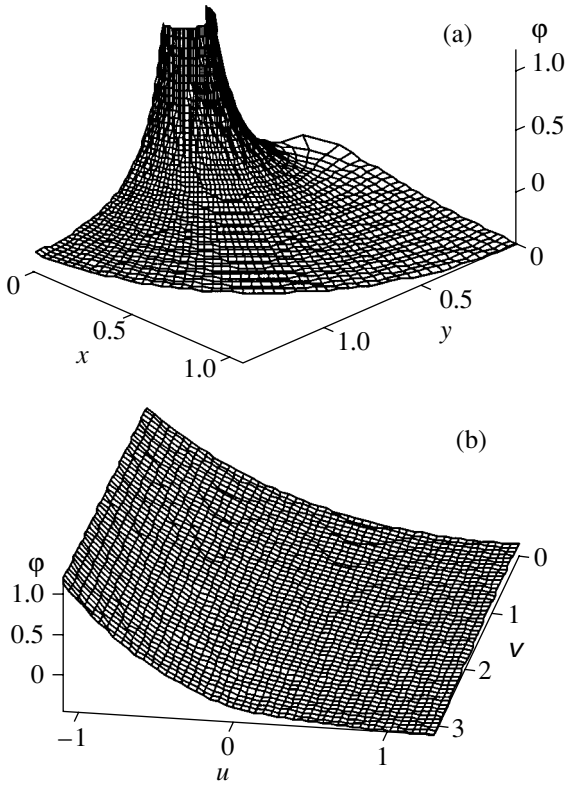


Fig. 7. Potential surfaces (a) $\varphi(x, y)$ in Cartesian coordinates and (b) $\varphi(u, v)$ in Cassini coordinates. The solution is obtained for $\delta = 0$, $d = 1$, $r_0 = 0.1$, $a_0 = 0.755$, and $\varphi_0 = 1.16$.

5. RESULTS OF CALCULATIONS

5.1. Debye Molecule in a Cloud of Like Charges ($\delta = 0$)

5.1.1. Choice of parameters for calculations. Calculations were made for parameters φ_0, r_0, and a_0, which correspond to a single Debye atom for $d \geq a_0$. For this purpose, we first solved the spherically symmetric problem, in which the values of the field and potential at point $r = a_0$ were assumed to be equal to zero. The solution to this problem was used to determine the potential φ_0 on a particle with the given radius r_0. Then we solved the two-center problem with these values of φ_0, r_0, and a_0 for $d = 10a_0$. The results of the solutions for the spherically symmetric and two-center problems

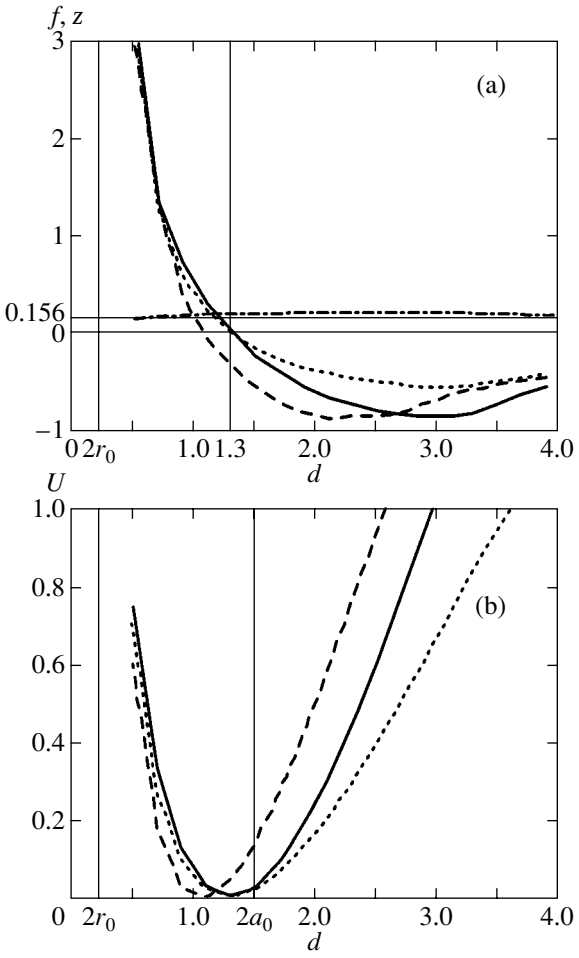


Fig. 8. Dependence of (a) the force component along the x axis and (b) the potential energy of interaction between dust particles on their separation d for $\delta = 0$. A positive value of the force corresponds to repulsion between dust particles, while a negative value corresponds to their attraction. The normalization of the potential energy is chosen so that the energy $U(d)$ at the minimum is equal to zero. Solid curves correspond to a constant potential $\varphi_0 = 1.16$ on the surface of dust particles; dotted curves correspond to a constant charge $z_0 = 0.156$ of a dust particle, which is ensured by the selection of $\varphi_0(d)$; and dashed curves correspond to a constant charge ensured by the choice of $a_0(d)$ for $\varphi_0 = 1.16$. The dot-and-dash curve in (a) gives the dependence of the dust particle charge on distance d for the case of constant potential $\varphi_0 = 1.16$.

coincided to a high degree of accuracy. In the subsequent series of calculations, we passed to smaller values of d.

In the series of calculations with the results presented in Fig. 8, we used the plasma parameters from [3] and set $a_0 = 0.755$. Calculations showed that the region of comparatively large distances $d \sim 2a_0$ is most interesting. Considering that the electron cloud near the surface of dust particles for $d \sim 2r_0$ is polarized weakly, we assumed for convenience of calculations that the particle radius $r_0 = 0.1$ is five times as large as the experimental value. Accordingly, the potential $\varphi_0 = 1.16$ borrowed from the solution to the one-center problem for $r_0 = 0.1$ has a value much smaller than the potential on the surface of a small dust particle ($\varphi_0 = 6.5$ for $r_0 = 0.0223$). In other words, a small charged conducting sphere was replaced by a larger sphere with a charge partly compensated by the charges of the electron cloud. According to calculations, deep electron shells are indeed polarized only slightly, and such a replacement is justified (see below).

5.1.2. Dependence of the force of interaction between dust particles on their separation. In order to determine the dependence of the force of interaction between dust particles on the separation d, we carried out a series of computations with given values of φ_0, r_0, and a_0. The particle charge z_0 in this case is also a function of d. For this reason, we made additional computations with values of φ_0 or a_0 changed in such a way that the particle charge z_0 did not depend on d.

Calculations proved that, for small distances between particles ($d \sim r_0$), repulsion takes place. This does not match the results of numerical calculations [8, 9], in which attraction was observed for $d \sim r_0$. Apparently, the above-mentioned error in calculation of the derivative of potential near the surface of a dust particle was large in the computations [8, 9]. The resultant force is very sensitive to this error. In actual practice, repulsion of particles at small distances prevails over polarization-induced attraction due to the fact that charge envelopes close to the surface of a dust particle are polarized weakly.

The equilibrium distance $d = d_0$ between dust particles, at which the force component reverses its sign, is of considerable interest. In the calculations presented in Fig. 8, we have $d_0 \approx 1.3$, which is slightly smaller than the mean distance $2a_0 = 1.5$ between dust particles. The position of point d_0 weakly depends on the quantities (φ_0, a_0 or z_0, a_0) preserved in calculations upon the variation of d. The variation of a_0 (for constant z_0 and φ_0) affects the value of d_0 more strongly. Apparently, confinement of the constant charge $z_0 = \text{const}$ by varying the particle potential $\varphi_0 = \varphi_0(d)$ is more closely related to the physics of interaction of charged dust particles.

Since the problem cannot be treated as a binary problem for $d \gg a_0$, we describe the results of calculations only for comparatively small values of $d < 4a_0$. For the separation between particles $d > 2a_0$, the repulsion from other particles surrounding the two particles under consideration becomes significant (see Fig. 1).

Knowing the force of attraction $F(2a_0)$ between dust particles separated by the mean interparticle spacing $2a_0$, we can estimate the electrostatic pressure compressing the dust particle gas,

$$P_E \approx F(2a_0)N_p^{2/3} = (T^2/8\pi e^2)N_p^{2/3} f(2a_0), \qquad (16)$$

and the surface tension of the "dust liquid,"

$$\sigma_E \approx F N_p^{1/3} = (N_p^{1/3}T^2/8\pi e^2) f(2a_0).$$

Comparing the electrostatic pressure exerted on dust particles with the gas-kinetic pressure of dust particles and the gas-kinetic pressure of free electrons, we obtain

$$P_E/N_pT = (T/8\pi e^2 N_p^{1/3}) f(2a_0),$$

$$P_E/N_eT = (T/8\pi e^2 N_e) N_p^{2/3} f(2a_0).$$

Under the conditions of experiments [3], we have

$$|f(2a_0)| \approx 0.2,$$

$$P_E = 9.7\times 10^{-7}|f(2a_0)| \text{ torr} \approx 2\times 10^{-7} \text{ torr},$$

$$\sigma_E = 3.5\times 10^{-9}|f(2a_0)| \text{ N/m} \approx 7\times 10^{-10} \text{ N/m},$$

$$P_E/N_pT \approx 20, \qquad P_E/N_eT \approx 0.04.$$

It should be noted, however, that a comparison of the electrostatic pressure exerted on dust particles with the gas-dynamic electron pressure does not allow us to draw any important conclusions since electrons are not free, but are in the electric field of dust particles. At the same time, a gas of Debye atoms in the mixture with a neutral gas must exhibit a tendency to compression under the given experimental conditions [3]. Such a situation was considered in [27–29]. An analysis of the effect of interaction between Debye atoms on gasdynamic properties of a dust plasma is beyond the scope of this study.

5.1.3. Dependence on the size of the Debye atom. We carried out a series of calculations for various values of a_0 (see Fig. 9). The calculations show that attraction takes place only for $a_0 \le 1$. Even for $a_0 > 1.12$, the sign-reversal point of the force is situated at a large distance $d_0 > 4a_0$.

The condition $a_0 = a_p/r_D < 1$ can be written for dimensional quantities in the form

$$N_{e0} > N_{ecr} \equiv \frac{T}{\pi e^2} N_p^{3/2}. \qquad (17)$$

Electrostatic compressive forces vanish when $d_0 = 2a_0$, i.e., for $a_0 = 1$. Accordingly, the condition $a_0 = 1$ or $N_e = N_{ecr}$ is the equilibrium condition for a gas of Debye atoms.

In this case, the condition of a large charge on a particle, $z_p > 1/3$, can be written for the charge of the dust particle in units of electron charge:

$$Z_p > Z_{ecr} \equiv \frac{\pi}{6}\frac{N_{ecr}}{N_p}.$$

Under our experimental conditions, we have $N_{ecr} = 4.4\times 10^{10}$ cm^{-3} and $Z_{ecr} = 460$. These quantities are of the same order of magnitude as those measured in the experiments [3]: $N_{e0} \approx 2.5\times 10^{10}$ cm^{-3} and $Z_p \approx 500$.

In view of the fact that the depth of the potential well cannot be determined from an analysis of two Debye atoms without taking into account the action of other

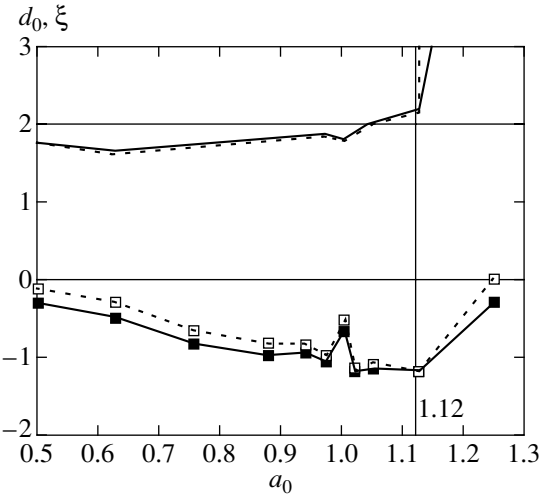


Fig. 9. Dependence of the coordinate d_0 of the sign-reversal point (bold curves) and of the steepness ξ of the force at point d_0 (fine curves with squares) on the size a_0 of the Debye atom. Solid curves correspond to a potential on the surface of dust particles corresponding to a_0 and independent of d; dashed curves correspond to a constant charge of a dust particle, ensured by the choice of $\varphi_0(d)$.

particles, we will characterize the force of interaction of Debye atoms by the steepness at the point of intersection with the abscissa axis:

$$\xi = f'(d)\big|_{d=d_0} = U''(d)\big|_{d=d_0}.$$

The frequency of oscillation of dust particles about the equilibrium position can be expressed in terms of the quantity ξ,

$$\omega = |\xi|^{1/2}\omega_0,$$

where $\omega_0 = v_T/a_p$, $v_T = \sqrt{2T/m_p}$ is the thermal velocity of dust particles, and m_p is their mass. Under the experimental conditions [3], we have $m_p \approx 2\times 10^{-12}$ g, $v_T \sim 0.5$ cm/s, and $a_p \approx 1.4\times 10^{-3}$ cm. It follows hence that the frequency of oscillations is $\omega_0 = 357$ s^{-1} and the period of oscillations is $2\pi/\omega_0 = 18$ ms. It can be seen from Fig. 9 that the strongest coupling takes place for $0.5 < a_0 < 1$. Under these conditions, the gas of Debye atoms must have a tendency to compression (cf. [29]).

For small-radius dust particles, the size of the Debye atom is also smaller. For example, the maximum value of the radius of the Debye atom $a_{0\max} \equiv a_0(z_0 \longrightarrow \infty)$ as a function of r_0 can be approximated by the expression $a_{0\max} = 3r_0^{0.3}$ or $r_0 = (a_{0\max}/3)^{10/3}$ for $r_0 < 0.02$ [25]. Consequently, the radius of a dust particle must not be too large or too small. For $0.5 < a_{0\max} < 1$, we have the con-

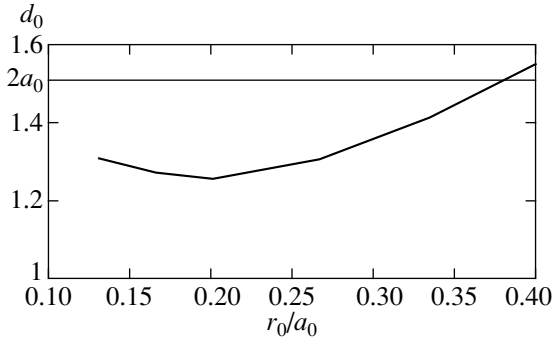


Fig. 10. Dependence of the coordinate d_0 of the sign-reversal point for the force on r_0/a_0. Here, r_0 can be treated as the radius of the domain in which the polarization of the charge cloud is disregarded. Potential φ_0 for $r = r_0$ was determined for $a_0 = 0.755$.

dition $2.5 \times 10^{-3} < r_0 < 2.6 \times 10^{-2}$. In the experiments [3], $r_0 = 2.23 \times 10^{-2}$, and this condition is satisfied.

5.1.4. On the effect of the dust particle size. It was mentioned above that, in our calculations, we replaced a small charged sphere by a conducting sphere of a larger size with a charge partially compensated by free charges in the shell of the Debye atom. A natural question arises concerning the correctness of such a replacement. In order to verify this, we carried out several series of calculations with different values of r_0 and, accordingly, φ_0. The difference in the results is insignificant for dust particle radii smaller than the radius a_0 of the Debye atom. For instance, for $a_0 = 0.755$ (see Fig. 10), the difference in the position of the point $d_0 = 1.28$ at which attraction is replaced by repulsion upon a change in the dust particle radius in the range $r_0 = 0.1$–0.2 (and for the choice of the values of φ_0 corresponding to the given value of r_0) has a spread less than 2%, which is within the error limits of the calculations.

The effect of the dust particle size becomes significant for $r_0 > 0.3a_0$. For $r_0 > 0.4$, the polarization-induced attraction decreases to such an extent that the coordinate of the force sign-reversal point becomes larger than the mean distance between particles ($d_0 > 2a_0$). Consequently, it can be concluded that a considerable contribution to the polarization forces comes not only from the periphery of the Debye atom ($r \approx a_0$), but also from the charges separated by a distance $r \approx 0.3a_0$ from the center of a dust particle. In other words, the force of attraction is formed due to polarization of the majority of electrons in the charge shell. Consequently, the attractive forces can hardly be calculated using some approximate methods.

5.2. Debye Molecule in a Cloud of Unlike Charges ($\delta \neq 0$)

5.2.1. Dependence of the force of interaction between dust particles on their separation. In the case of two Debye atoms in a nearly quasi-neutral plasma ($1 - \delta \ll 1$), the separation between these atoms can be chosen as large as desired (see Fig. 1). However, in order to assume that the interaction is binary, the condition $N_p^{-1/3} \gg 2a_0 r_D$ must be satisfied. It should be recalled that the radius of a Debye atom for $1 - \delta \ll 1$ may considerably exceed the Debye radius, $a_0 > 1$ (see above and [25]).

As in the case when $\delta = 0$, we carried out series of calculations with preset values of φ_0, r_0, and a_0 for a Debye atom to determine the dependence of the force of interaction between dust particles on distance d. Additional computations were made with values of φ_0 or a_0 modified so that the particle charge z_0 was independent of d. As in the case of $\delta = 0$, we chose the value of r_0 greater than the radius of the atomic core, thus simulating a dust particle by a conducting sphere of a larger size with a charge partially compensated by the free charges from the shell of the Debye atom. Thus, the polarization of the core was disregarded.

In the results presented in Fig. 11, the value of a_0 for different values of δ (see table) was chosen so that it corresponds to the extremely large charge for the dust particle radius corresponding to the experimental conditions [3]:

$$r_0 = r_p/r_D = 2.23 \times 10^{-2}.$$

This was done by test firing: when the value of a_0 is greater than that given in table, the particle charge becomes infinitely large ($z(r_0) \longrightarrow \infty$). The obtained

Parameters of calculations for various values of δ in Fig. 11

Type of curve and figure	δ	a_0	φ_0	z_0
Dotted curves in (a) and (b)	0.999	4.1	2.455	0.282
Dashed curves in (a) and (b)	0.9	1.71	2.426	0.283
Dot-and-dash curves in (c) and (d)	0.7	1.288	2.413	0.272
Dotted curves in (c) and (d)	0.5	1.12	2.378	0.286
Dashed curves in (c) and (d)	0.3	1.018	2.364	0.286
Solid bold curves in (c) and (d)	0.1	0.94	2.292	0.279

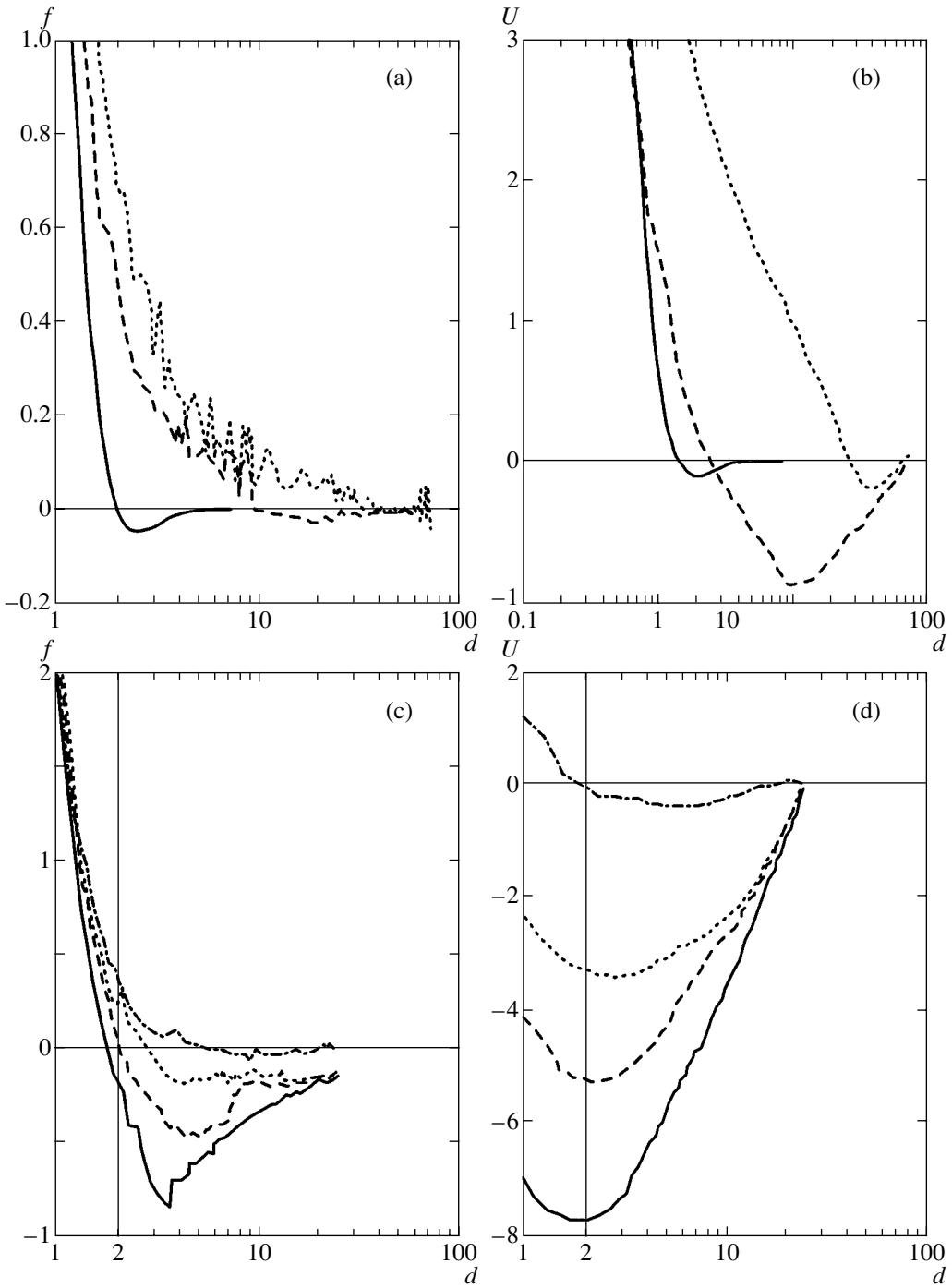


Fig. 11. Dependences (a, c) of the *x* component of force and (b, d) of the potential energy *U* of interaction between dust particles on their separation *d* for different values of $\delta \neq 0$ (see table). In all cases, $r_0 = 0.1$. Solid curves correspond to the analytic expressions from [7]: $f(d) = \text{const}(1/d)(1 + d - d^2/2)\exp(-d)$ and $U(d) = \text{const}(1/d^2)(1 - d/2)\exp(-d)$.

dependences $z(r)$ and $\varphi(r)$ were used for determining $z_0 = z(r_0)$ and $\varphi_0 = \varphi(r_0)$ for $r_0 = 0.1$.

For $1 - \delta \ll 1$, attraction of a particle could not be detected in the range of parameters under investigation. Attraction appears only when a noticeable fraction of the positive charge of the plasma is carried by dust particles (for $\delta < 0.7$; see Fig. 11). The smaller the fraction of the plasma charge, the stronger the maximum force of attraction and the larger the potential well depth.

The attenuation of attractive forces upon an increase in δ has a simple explanation. The results of the above calculations show that, for $\delta = 0$, attractive forces

emerge due to the fact that electrons are accumulated near the x axis between the centers of particles and ensure the attraction to the center of the Debye molecule. This attraction exceeds the repulsion of dust particle charges screened by inner layers of the electron shells of Debye atoms. For $1 - \delta \ll 1$, the effect of screening of the dust particle charge by the inner layers of electrons shells remains unchanged. However, the attraction to the center of the molecule attenuates considerably due to the fact that not only electrons but also the positive charges repelling dust particles are concentrated in this region.

In the case of small values of plasma charge ($\delta \ll 1$), the potential well depth is large enough (of the order of several values of temperature). It should be borne in mind, however, that the binary treatment is valid for distances of the order of the diameter $2a_0$ of the Debye atom.

5.2.2. On analytical approaches. The above conclusion concerning the absence of attraction for $1 - \delta \ll 1$ contradicts the results of recent approximate analyses [12, 13] (see Fig. 11). It follows from the results of these calculations that attraction between dust particles takes place in an analysis of the linearized Poisson–Boltzmann equation for $\delta = 1$ in the region

$$r > \frac{\sqrt{3}+1}{\sqrt{2}} = 1.93.$$

This result is surprising. As a matter of fact, the action of point charges (located at points $\mathbf{r}_1$ and $\mathbf{r}_2$) in a plasma in the linear approximation [12, 13] is assumed to be independent, so that the value of potential at point $\mathbf{r}$ is determined by the sum of the screened potentials of point charges:

$$\varphi(\mathbf{r}) = \Phi(|\mathbf{r} - \mathbf{r}_1|) + \Phi(|\mathbf{r} - \mathbf{r}_2|),$$

$$\Phi(x) = (z_p\sqrt{2}/x)\exp(-x/\sqrt{2}), \quad |\mathbf{r}_2 - \mathbf{r}_1| = d.$$

In accordance with the simple considerations presented above in Subsection 3.2, attractive forces are ruled out in the absence of rearrangement of the charge shell of a dust particle under the action of another dust particle for identically charged dust particles. In the linear approximation [12, 13], the following inequality must hold:

$$f(d) = -z_p \left.\frac{\partial \Phi(x)}{\partial x}\right|_{x=d} = z_p^2 \frac{\sqrt{2}+d}{d^2}\exp\left(-\frac{d}{\sqrt{2}}\right) > 0; \tag{18}$$

this corresponds to repulsion.

Inaccuracy of the results [12, 13] is apparently associated with the following circumstance. The force (18) acting directly on a dust particle was supplemented in [12, 13] by the attractive force exerted on the electron shell of a dust particle by another dust particle. Such an approximation would be justified if the charge shells of dust particles would be rigidly connected with the dust particle charges through some other forces. However, there are no extraneous rigid forces in the problem under consideration. The presence of the force of attraction of the electron shell of one charge to another charge, which was calculated in [12, 13], only indicates that the given configuration of the charge shell is not equilibrium, and the force of attraction to another charge must lead to polarization of the charge shell, which was disregarded in [12, 13]. There are no grounds to add this polarizing force to the force acting directly on the dust particle.

In this case, we can draw an analogy with the polarization-induced attractive forces between atoms, which are considerable at large distances between the nuclei (as compared to the size of electron shells). It is well known that no polarization-induced interaction of spherically symmetric atoms takes place in the first order of perturbation theory. It appears only in the second order of perturbation theory, when the polarization of the electron shell of an atom by the charges of another atom is taken into account. An ordinary atom differs from a Debye atom only in that the electrons of the shell move in it according to quantum-mechanical and not classical laws. The origin of polarization-induced forces is the same for an ordinary and a Debye atom.

6. CONCLUSIONS

Let us summarize the results of the above analysis.

1. A Debye atom has the following structure: a core surrounded by an electron shell. For large charges of the dust particle, the high-density region (core) of the electron cloud screens considerably the large charge of the dust particle near its surface. In this connection, while considering the interaction of Debye atoms, we cannot ascribe the unscreened value of charge to a dust particle. The dust particle charge screened by the core has a universal value determined by the distance between dust particles. It is screened by the electron shell of the Debye atom.

2. Attractive forces are associated with polarization of charge shells of Debye atoms. In the absence of polarization, attraction is also absent. The force of attraction is formed due to polarization of the majority of electrons of the charge shell. The polarization of the core is insignificant.

3. Forces of attraction between dust particles emerge at a comparatively large distance approximately equal to the mean separation between dust particles. In this case, the Debye radius must be equal approximately to half the mean distance between dust particles.

4. Attraction takes place if like charges are concentrated predominantly at dust particles. If dust particles

carry a small fraction of the charge of some polarity, repulsion is observed at any distance.

5. For a certain relation between the electron density and the density of dust particles, the Debye "liquid" is in equilibrium: the electrostatic forces of interaction between dust particles vanish.

Since attractive forces appear at large distances, the problem of the formation of dust liquids and crystals can be solved correctly only if many-particle interaction is taken into account. However, we can draw the following two conclusions concerning the criteria for the emergence of collective phenomena on the basis of the results presented by us here:

(a) In the case of a thermionic plasma, the electron density must be such that the Debye radius is equal approximately to half the mean value between dust particles.

(b) For a gas-discharge or a nuclear-excited plasma, the properties of the ionization source and the density of dust particles must be matched so that the main (usually negative) charge is carried by dust particles.

ACKNOWLEDGMENTS

The authors are grateful to A.N. Tkachev for fruitful discussions of the results of this work and [12, 13] and to Yu.I. Syts'ko for the discussion of computational aspects of the problem.

REFERENCES

1. V. E. Fortov and I. T. Yakubov, *Nonideal Plasma* (Énergoatomizdat, Moscow, 1994).
2. N. N. Tsytovich, Usp. Fiz. Nauk **167**, 57 (1997) [Phys. Usp. **40**, 53 (1997)].
3. V. E. Fortov, A. P. Nefedov, O. F. Petrov, *et al.*, Zh. Éksp. Teor. Fiz. **111**, 467 (1997) [JETP **84**, 256 (1997)].
4. V. E. Fortov, V. S. Filinov, A. P. Nefedov, *et al.*, Zh. Éksp. Teor. Fiz. **111**, 889 (1997) [JETP **84**, 489 (1997)].
5. V. E. Fortov, V. I. Vladimirov, L. V. Deputatova, *et al.*, Dokl. Akad. Nauk **366**, 184 (1999) [Dokl. Phys. **44**, 279 (1999)].
6. A. P. Nefedov, O. F. Petrov, and V. E. Fortov, Usp. Fiz. Nauk **167**, 1215 (1997) [Phys. Usp. **40**, 1163 (1997)].
7. A. N. Tkachev and S. I. Yakovlenko, Zh. Tekh. Fiz. **69** (1), 53 (1999) [Tech. Phys. **44**, 48 (1999)].
8. S. I. Yakovlenko, Pis'ma Zh. Tekh. Fiz. **25** (16), 83 (1999) [Tech. Phys. Lett. **25**, 670 (1999)].
9. S. I. Yakovlenko, Kratk. Soobshch. Fiz., No. 9, 3 (1999).
10. B. Derjagin and L. Landau, Acta Physicochim. URSS **14**, 633 (1941).
11. S. I. Yakovlenko, Pis'ma Zh. Tekh. Fiz. **27** (9), 83 (2001) [Tech. Phys. Lett. **27**, 389 (2001)].
12. D. N. Gerasimov and O. A. Sinkevich, Teplofiz. Vys. Temp. **37**, 853 (1999).
13. A. S. Ivanov, Phys. Lett. A **290**, 304 (2001).
14. A. M. Larsen and D. G. Grier, Nature (London) **385**, 230 (1997).
15. D. G. Grier, Nature (London) **393**, 621 (1998).
16. W. R. Bowen and A. O. Sharif, Nature (London) **393**, 663 (1998).
17. J. C. New, Phys. Rev. Lett. **82**, 1072 (1999).
18. M. Tokuyama, Phys. Rev. E **59**, R2550 (1999).
19. V. A. Gundienkov and S. I. Yakovlenko, Kratk. Soobshch. Fiz., No. 12, 3 (2001); Pis'ma Zh. Tekh. Fiz. **28** (10), 46 (2001) [Tech. Phys. Lett. **28**, 422 (2001)].
20. *Mathematical Encyclopedia*, Ed. by Yu. V. Prokhorov (Bol'shaya Rossiĭskaya Éntsiklopediya, Moscow, 1995).
21. I. N. Bronshteĭn and K. A. Semendyaev, *Handbook on Mathematics* (Nauka, Moscow, 1964).
22. A. N. Tkachev and S. I. Yakovlenko, Pis'ma Zh. Tekh. Fiz. **25** (1), 52 (1999) [Tech. Phys. Lett. **25**, 21 (1999)].
23. S. I. Yakovlenko, Pis'ma Zh. Tekh. Fiz. **26** (16), 47 (2000) [Tech. Phys. Lett. **26**, 337 (2000)].
24. S. I. Yakovlenko, Pis'ma Zh. Tekh. Fiz. **26** (26), 38 (2000) [Tech. Phys. Lett. **26**, 1045 (2000)].
25. S. I. Yakovlenko, Kratk. Soobshch. Fiz., No. 1, 3 (2002).
26. E. G. Gibson, Phys. Rev. **9**, 2389 (1996).
27. S. A. Maĭorov, A. N. Tkachev, and S. I. Yakovlenko, Usp. Fiz. Nauk **164**, 298 (1994) [Phys. Usp. **37**, 279 (1994)].
28. S. A. Mayorov, A. N. Tkachev, and S. I. Yakovlenko, Phys. Scr. **51**, 498 (1995).
29. S. I. Yakovlenko, Izv. Vyssh. Uchebn. Zaved., Fiz. **38**, 3 (1995).

Translated by N. Wadhwa

Journal of Experimental and Theoretical Physics, Vol. 95, No. 5, 2002, pp. 878–880.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 122, No. 5, 2002, pp. 1019–1021.

FLUIDS

Convective Instability Mechanism for a Binary Mixture Heated from Above

L. Kh. Ingel

Scientific Production Association Typhoon, Obninsk, Kaluga oblast, 249038 Russia
e-mail: lingel@obninsk.com

Received February 8, 2002

Abstract—It is shown that the intensification of a stabilizing buoyancy flux directed from the surface to the bulk of a binary mixture (e.g., salt water) may give rise to convective instability. *© 2002 MAIK "Nauka/Interperiodica".*

Some physical mechanisms are known to lead to a paradoxical (at first glance) result: contrary to intuitive ideas, convective instability emerges in a liquid when it is heated not from below, but from above, i.e., in situations when the density of the liquid increases with the depth [1, 2]. Another possibility of this kind is considered in this communication.

We consider a semi-infinite layer of a liquid medium stratified both in temperature and in the admixture concentration (see the figure below). For the sake of definiteness, we can speak, for example, of salt water in which the contribution to the density stratification comes from the vertical distributions $T(z)$ and $s(z)$ of temperature and salt concentration, respectively (the z axis is directed vertically downwards from the horizontal surface of the medium). We assume that the medium is heated from above so that a constant vertical temperature gradient

$$\gamma_T^{(1)} = \frac{\partial T^{(1)}}{\partial z} < 0$$

is formed in it. At the same time, the medium is also characterized by the admixture stratification:

$$\gamma_s^{(1)} = \frac{\partial s^{(1)}}{\partial z} < 0.$$

In the framework of the conventional approximation [1, 3, 4], we assume that the density ρ is a linear function of temperature and admixture concentration:

$$\rho = \rho_0[1 - \alpha(T - T_0) + \beta(s - s_0)]. \quad (1)$$

Here, α is the thermal expansion coefficient of the medium, and $\beta > 0$ is the corresponding coefficient for the admixture concentration (in oceanology, this coefficient is referred to as the salinity contraction coefficient). The subscript "0" marks the constant (reference) values of the corresponding quantities. For constant values of the gradients $\gamma_T^{(1)}$ and $\gamma_s^{(1)}$, the density of the medium is a linear function of the depth z:

$$\rho^{(1)}(z) = \rho_0[1 + (-\alpha\gamma_T^{(1)} + \beta\gamma_s^{(1)})z] \quad (2)$$

(the superscript (1) corresponds to the initial "background" state).

The above-mentioned gradients $\gamma_T^{(1)}$ and $\gamma_s^{(1)}$ correspond to certain heat and admixture fluxes on the horizontal surface of the medium:

$$Q_T^{(1)} = -c_p\rho_0\kappa\gamma_T^{(1)} > 0, \quad Q_s^{(1)} = -\rho_0\chi\gamma_s^{(1)} > 0, \quad (3)$$

where c_p is the heat capacity of the medium, and κ and χ are the transfer coefficients.

Although the admixture flux directed to the bulk of the medium makes a destabilizing contribution to the density stratification (for $\gamma_s^{(1)} < 0$, the corresponding term in Eq. (2) decreases with increasing depth), heating from above is assumed to be strong enough so that the density stratification is quite stable on the whole. It should be recalled that the condition

$$\frac{\partial\rho}{\partial z} > 0,$$

which is equivalent to the condition

$$-\beta\gamma_s < -\alpha\gamma_T$$

or

$$\gamma_s > \frac{\alpha}{\beta}\gamma_T,$$

is insufficient for hydrodynamic stability of such a system. If the values of the transfer coefficient κ and χ are different ($\chi \ll \kappa$ for salt water), the instability condition may be much more stringent in view of the possible specific instability mechanism associated with "double diffusion" [1, 5]:

$$-\gamma_s < -\frac{\chi}{\kappa}\frac{\alpha}{\beta}\gamma_T \text{ or } \gamma_s > \frac{\chi}{\kappa}\frac{\alpha}{\beta}\gamma_T. \quad (4)$$

1063-7761/02/9505-0878$22.00

We assume that, in the background state considered above, this condition is satisfied. As applied to the fluxes, this corresponds to the fulfillment of the inequality

$$Q_T^{(1)} > c_p\left(\frac{\kappa}{\chi}\right)^2\frac{\beta}{\alpha}Q_s^{(1)}. \quad (5)$$

Let us suppose that the fluxes at the boundary $z = 0$ change at a certain instant $t = 0$. For simplicity, we assume that the fluxes instantaneously assume new constant values $Q_T^{(2)}$ and $Q_s^{(2)}$, these new values corresponding to an increase in the downstream buoyancy flux due to a decrease in the admixture flux:

$$Q_s^{(2)} < Q_s^{(1)}.$$

The sign of the admixture flux may even change: $Q_s^{(2)} < 0$ (downstream freshening instead of salinization). We assume that heating of the medium from above continues, although it becomes weaker:

$$0 < Q_T^{(2)} < Q_T^{(1)}.$$

The jumpwise variation of the fluxes at the boundary obviously leads to evolution of the temperature and admixture concentration fields. The analysis of this evolution is reduced to the solution of 1D thermal conductivity and diffusion equations,

$$\frac{\partial T}{\partial t} = \kappa\frac{\partial^2 T}{\partial z^2}, \quad \frac{\partial s}{\partial t} = \chi\frac{\partial^2 s}{\partial z^2}, \quad (6)$$

with the initial conditions

$$T = T^{(1)} = T_0 + \gamma_T^{(1)}z,$$

$$s = s^{(1)} = s_0 + \gamma_s^{(1)}z \text{ at } t = 0$$

and with the boundary conditions

$$T \longrightarrow T^{(1)}, \quad s \longrightarrow s^{(1)} \text{ for } z \longrightarrow \infty,$$

$$\frac{\partial T}{\partial z} = \gamma_T^{(2)} = -\frac{Q_T^{(2)}}{c_p\rho_0\kappa},$$

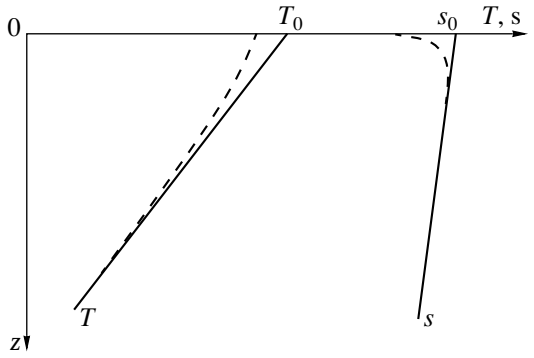
$$\frac{\partial s}{\partial z} = \gamma_s^{(2)} = -\frac{Q_s^{(2)}}{\rho_0\chi} \text{ for } z = 0.$$

The solution, which can easily be found using the superposition method, has the form

$$T = T_0 + \gamma_T^{(1)}z - \frac{2(Q_T^{(1)} - Q_T^{(2)})}{c_p\rho_0\kappa}\sqrt{\kappa t}\,\mathrm{i}\,\mathrm{erf}\frac{z}{\sqrt{4\kappa t}},$$

$$s = s_0 + \gamma_s^{(1)}z - \frac{2(Q_s^{(1)} - Q_s^{(2)})}{\rho_0\chi}\sqrt{\chi t}\,\mathrm{i}\,\mathrm{erf}\frac{z}{\sqrt{4\chi t}},$$

where i erf is the symbol of the multiple error integral [6]. The evolution of the fields is presented schematically in the figure. Perturbations penetrate the medium from the surface according to diffusion laws (the penetration depth is proportional to $\sqrt{t}$). However, the proportionality factors for temperature and admixture concentration may differ considerably. For example, for salt water, $\kappa \gg \chi$, so that the rate of penetration of the temperature perturbation into the medium is an order of magnitude higher. The perturbation amplitude also increases in proportion to $\sqrt{t}$. For the vertical gradient, we obtain



Schematic diagram of vertical profiles of temperature T and admixture concentration s. Solid lines correspond to the initial state and dashed curves describe evolution after a change of fluxes at the surface.

$$\frac{\partial T}{\partial z} = \gamma_T^{(1)} - (\gamma_T^{(1)} - \gamma_T^{(2)})\mathrm{erfc}\frac{z}{\sqrt{4\kappa t}},$$

$$\frac{\partial s}{\partial z} = \gamma_s^{(1)} - (\gamma_s^{(1)} - \gamma_s^{(2)})\mathrm{erfc}\frac{z}{\sqrt{4\chi t}}.$$

The asymptotic forms of solutions for small,

$$z \ll \min\{\sqrt{4\kappa t}, \sqrt{4\chi t}\},$$

and large,

$$z \gg \max\{\sqrt{4\kappa t}, \sqrt{4\chi t}\},$$

depths are especially simple. At large depths, we have

$$\frac{\partial T}{\partial z} \approx \gamma_T^{(1)}, \quad \frac{\partial s}{\partial z} \approx \gamma_s^{(1)}.$$

In other words, the background state is preserved at such depths: the perturbation associated with a change in the fluxes at the surface has no time to spread to such distances. In this region, the medium is stable, as would also be the case for $t < 0$. At small depths, we have

$$\frac{\partial T}{\partial z} \approx \gamma_T^{(2)}, \quad \frac{\partial s}{\partial z} \approx \gamma_s^{(2)};$$

in other words, the stratification in this region has time to adjust to new values of the fluxes at the surface. Since the downstream buoyancy flux has only increased for $t \geq 0$, it is natural to expect that the medium in the

region in which perturbations from the surface have arrived has become even more stable to convection. This is true if the new values of the fluxes satisfy the following condition similar to relation (5):

$$Q_T^{(2)} > c_p\left(\frac{\kappa}{\chi}\right)^2\frac{\beta}{\alpha}Q_s^{(2)}. \quad (7)$$

The fulfillment of this condition is especially obvious in the case when the medium is heated from above ($Q_T^{(2)} > 0$) as before, but the flux of admixture at the surface vanishes or becomes negative.

However, apart from the asymptotic forms considered above, an intermediate region also exists. In particular, if the value of κ is greater than χ by two orders of magnitude, there obviously exists (and continuously expands) a region

$$\sqrt{4\chi t} \leqslant z \ll \sqrt{4\kappa t} \quad (8)$$

in which

$$\frac{\partial T}{\partial z} \approx \gamma_T^{(2)}, \quad \frac{\partial s}{\partial z} \approx \gamma_s^{(1)}.$$

We are speaking of the region to which thermal perturbations from the surface have time to penetrate, while salinity perturbations have no time to penetrate into this region. Here, condition (4) may be violated since the value of γ_s remains unchanged (background), while the value of γ_T has increased (manages to adjust itself to new conditions at the boundary).

Thus, if

$$\gamma_T^{(2)} > \frac{\kappa}{\chi}\frac{\beta}{\alpha}\gamma_s^{(1)} \text{ or } Q_T^{(2)} < c_p\left(\frac{\kappa}{\chi}\right)^2\frac{\beta}{\alpha}Q_s^{(1)},$$

the stability conditions in the intermediate region (8) are violated, and convection associated with double (differential) diffusion must appear.

It should be emphasized that this result is nontrivial. The buoyancy fluxes set at the boundary are of a stabilizing nature: at any instant, the steady-state stratification corresponding to them is stable, which rules out the emergence of convective instability. However, owing to the unsteady nature of these fluxes and the difference in the velocities of propagation of temperature and concentration perturbations, a region in which the stability conditions are violated may be formed in the bulk of the medium.

It can easily be seen that the above assumption concerning the semi-infiniteness of the medium is not of fundamental importance. A similar effect is also possible when nonsteady stabilizing fluxes are set at the lower boundary of the medium.

ACKNOWLEDGMENTS

This study was supported financially by the Russian Foundation for Basic Research (project no. 02-05-64203) and the International Center of Science and Technology (project no. G-553).

REFERENCES

1. J. S. Turner, *Buoyancy Effects in Fluids* (Cambridge Univ. Press, Cambridge, 1973; Mir, Moscow, 1977).
2. L. Kh. Ingel', Usp. Fiz. Nauk **167**, 779 (1997) [Phys. Usp. **40**, 741 (1997)].
3. L. D. Landau and E. M. Lifshitz, *Course of Theoretical Physics*, Vol. 6: *Fluid Mechanics* (Nauka, Moscow, 1986; Pergamon, New York, 1987).
4. G. Z. Gershuni and E. M. Zhukhovitskiĭ, *Convective Stability of Incompressible Fluid* (Nauka, Moscow, 1972).
5. G. Walin, Tellus **16**, 389 (1964).
6. *Handbook of Mathematical Functions*, Ed. by M. Abramowitz and I. A. Stegun (National Bureau of Standards, Washington, 1964; Nauka, Moscow, 1979).

Translated by N. Wadhwa

Journal of Experimental and Theoretical Physics, Vol. 95, No. 5, 2002, pp. 881–885.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 122, No. 5, 2002, pp. 1022–1026.

SOLIDS
Structure

Decay of High-Frequency Acoustic Excitations in Chain-type Lattices and Weak Localization

E. P. Chulkin

Physicotechnical Institute, Ural Division, Russian Academy of Sciences, ul. Kirova 132, Izhevsk, 426001 Russia
e-mail: chulkin@otf.fti.udmurtia.su

Received March 5, 2002

Abstract—The attenuation of high-frequency sound in disordered quasi-one-dimensional semiconducting and dielectric crystals, which is associated with three-phonon decay and elastic scattering by structure defects is considered theoretically. It is shown that specific interference processes occurring in the regime of weak localization of acoustic vibrational excitations considerably affect the propagation of sound. This mechanism of sound attenuation can be observed experimentally from the anomalies in the frequency dependence of the reciprocal attenuation length of sound. *© 2002 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

Attenuation of sound due to decay processes in standard weakly anisotropic three-dimensional defect-free crystal lattices was considered for the first time in [1]. It was shown that anisotropy does not introduce any considerable changes in the lattice attenuation coefficient of sound. The mechanism of propagation of high-frequency acoustic excitations during combined action of elastic scattering by defects and three-phonon decay processes was analyzed analytically [2, 3] and investigated experimentally [4–6]. It was found that, in the helium temperature range, three-phonon decay processes modify significantly the diffusion mode of propagation of acoustic excitations when elastic scattering plays a dominating role.

The vibrational spectrum of chain-type lattices exhibits quasi-one-dimensional properties in the entire range of the frequency spectrum except at singular points and spectrum boundaries in contrast to standard three-dimensional compounds due to a strong anisotropy of interaction between atoms. As a result, localization effects become significant in the propagation of acoustic excitations in nonideal low-dimensional lattices [7, 8]. It should be emphasized that well-defined quasi-local modes do not appear in such lattices under the conditions of diagonal disorder [9, 10]. Consequently, the effect of weak localization of vibrational modes can be analyzed without taking into account the renormalization of the vibrational spectrum [11] and the delay effect [12–14], which are found to be insignificant in this case.

This study aims at an analysis of the weak localization effect on the attenuation of high-frequency sound in nonideal chain-type crystals in the situation in which only decay is significant among all three-phonon processes.

We consider the localization of acoustic vibrational modes with displacement vectors parallel and perpendicular to weakly linked chains. Vibrational modes of the first kind are longitudinally polarized excitations (*l* modes). The modes of the second type are so-called bending excitations (*b* modes) [15–17]. Such modes were observed in experiments on inelastic neutron scattering in the quasi-one-dimensional compound $(Ta_{1-x}Nb_xSe_4)_2I$ [18] and manifested indirectly in the anomalous behavior of the low-temperature lattice heat capacity of the recently discovered new phase of carbon (carbolite) [19].

2. FORMULATION OF THE PROBLEM

We consider a crystal with isolated impurity atoms. Its dynamic properties will be described by the standard Hamiltonian taking into account cubic anharmonism,

$$H = H_0 + H_{\text{imp}} + H_{\text{int}} = H' + H_{\text{int}},$$

where

$$H_0 = \frac{1}{2M_0}\sum_{n,\alpha}(p_n^\alpha)^2 + \frac{1}{2}\sum_{\substack{n,n'\\ \alpha,\beta}}\Phi^{(0)\alpha\beta}_{n,n'}u_n^\alpha u_{n'}^\beta,$$

$$H_{\text{imp}} = \frac{1}{2}\left(\frac{1}{M}-\frac{1}{M_0}\right)\sum_{n,\alpha}c_n(p_n^\alpha)^2$$

$$+\frac{1}{2}\sum_{\substack{n,n'\\ \alpha,\beta}}\Delta\Phi^{\alpha\beta}_{nn'}u_n^\alpha u_{n'}^\beta, \qquad (1)$$

$$H_{\text{int}} = \frac{1}{6}\sum_{\substack{n,n',n''\\ \alpha,\beta,\gamma}} \Phi^{\alpha\beta\gamma}_{n,n',n''} u^{\alpha}_{n} u^{\beta}_{n'} u^{\gamma}_{n''},$$

$$\Delta\Phi^{\alpha\beta}_{nn'} = \Phi^{\alpha\beta}_{nn'} - \Phi^{(0)\alpha\beta}_{nn'}.$$

Here, H_0 is the Hamiltonian of the unperturbed harmonic atomic lattice; H_{imp} is the perturbation associated with impurities in this system; H' is the Hamiltonian of the harmonic nonideal lattice; H_{int} describes the dynamic anharmonic interaction between ions; the quantities u^{α}_{n} and p^{α}_{n} are the Cartesian components of the displacement and momentum operators of the nth atom; M and M_0 are the masses of an impurity atom and an atom of an ideal lattice (we assume that the impurity is heavy, i.e., $M \gg M_0$); and $\Phi_{nn'}$ and $\Phi^{0}_{nn'n''}$ are the matrix elements of the second- and third-order force parameters. The subscript "0" marks the parameters of the regular system. Factor c_n is equal to zero if a lattice atom is at the nth site and to unity if a point defect occupies this site. The configuration average $\langle c_n\rangle_n$ is equal to the impurity concentration c. For the sake of simplicity, we will henceforth assume that the force parameter matrices are diagonal in the Cartesian indices. In order to simplify notation, we will denote by n the combination of the site (n) and Cartesian (α) indices. In carrying out specific calculations, we assume that the disorder is of the diagonal type; i.e., impurities are treated as isotopic defects. Here, we do not distinguish between $\Phi_{nn'n''}$ and $\Phi^{(0)}_{nn'n''}$. Thus, we consider only the matrix anharmonism, which is assumed to be weak. The obtained results can be generalized to the case of a nondiagonal disorder, when $\Delta\Phi_{nn'} \neq 0$.

3. MODEL OF A CHAIN CRYSTAL

For the sake of definiteness, we assume that the lattice of a quasi-one-dimensional crystal is tetragonal with the unit cell parameters a and b. The effective interaction between atoms in the basal plane xy ($\|$) is considered to be much weaker than the interaction along the chain axis z ($\perp$). In this case, we have three characteristic force parameters, which satisfy the inequality

$$\left|\Phi^{(0)s_\perp}_{xx}\right| \gg \left|\Phi^{(0)s_\|}_{xx}\right| \gg \left|\Phi^{(0)s_\|}_{zz}\right|. \tag{2}$$

The reduced force parameters correspond to three characteristic frequencies: $\omega^2_3 \gg \omega^2_2 \gg \omega^2_1$. The dispersion relations for acoustic longitudinal and bending vibrational modes are defined as

$$\omega^2_l(\mathbf{k}) = \frac{(\omega_\perp b k_\perp)^2}{2} + \omega^2_\|\left(\sin^2\frac{k_x a}{2} + \sin^2\frac{k_y a}{2}\right), \tag{3}$$

$$\omega^2_b(\mathbf{k}) = (\omega_1 a k_\perp)^2 + \frac{(\omega_3 b^2 k^2_\perp)^2}{\pi} + \omega^2_2\left(\sin^2\frac{k_x a}{2} + \sin^2\frac{k_y a}{2}\right), \tag{4}$$

where $\omega_3 \approx \omega_\perp$ and $\omega_1 \approx \omega_\|$. The frequencies $\omega_\perp$ and $\omega_\|$ can be expressed in terms of parameters $\Phi^{(0)s_\perp}_{zz}$ and $\Phi^{(0)s_\|}_{zz}$ ($|\Phi^{(0)s_\perp}_{zz}| \gg |\Phi^{(0)s_\|}_{zz}|$) [7]. It is important that the inequality $\omega_\perp \gg \omega_\|$ holds.

We introduce a one-particle retarded Green's function G^+ [20] defined on the operators of dynamic atomic displacements u_n:

$$G^+_{nn'}(t-t') = -i\theta(t-t')\langle[u_n(t), u_{n'}(t')]\rangle. \tag{5}$$

The angle brackets indicate statistical averaging with Hamiltonian H'. In the momentum representation, the Green's function of the jth polarization mode averaged over impurity configurations is defined as

$$(\tilde{G}^+_j)^{-1}(\mathbf{k},\omega) = (\bar{G}^+_j)^{-1}(\mathbf{k},\omega) - \Pi^j(\mathbf{k},\omega). \tag{6}$$

Here, $\bar{G}^+_j(\mathbf{k},\omega)$ is the configuration-averaged retarded one-particle Green's function corresponding to the overall harmonic Hamiltonian H' and $\Pi^j(\mathbf{k},\omega)$ is the polarization operator. Here,

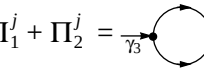
$$\bar{G}^+_j(\mathbf{k},\omega) = \left[\omega^2 - \omega^2_j(\mathbf{k}) - i\frac{\omega}{\tau^j_i(\omega)}\right]^{-1}, \tag{7}$$

where the lifetime for elastic processes is given by

$$\tau^j_i(\omega) = \left[\frac{\pi}{2}c\epsilon^2\omega^2 g_j(\omega)\right]^{-1}, \tag{8}$$

$g_j(\omega)$ is the spectral partial function of the density of states of vibrational modes, and $\epsilon = (M - M_0)/M_0$. We assume that the temperature is comparatively low, and the Rayleigh mechanism of quasiparticle decay prevails over the anharmonic mechanism (see below). As regards the polarization operator Π^j, it has the following form in the cubic anharmonism approximation:

$$\Pi^j = \Pi^j_1 + \Pi^j_2 =$$



. (9)

The graphical relation Π^j_1 describes the standard three-phonon process of spontaneous decay of an acoustic quantum in the presence of defects and Π^j_2 describes the interaction between decay acoustic phonons and phonon density fluctuations in the vicinity of defects [21, 22]. The lines with arrows in relation (9) correspond to Green's function $\bar{G}^{+(-)}$, and vertex U emerges as a result of summation of fan diagrams and characterizes the processes of coherent backward scattering of phonons by defects. It is well known that they deter-

mine the weak localization regime. The explicit expression for vertex U^j is given in [7]. In our calculations, we assume that the contributions from the *l* and *b* modes to the vertex U^j are independent. We specify the anharmonic interaction parameter Φ_j^3 in the standard approximation:

$$\Phi_j^3(\mathbf{k}, \mathbf{k}_1, \mathbf{k}_2) = -i\tilde{\gamma}_3\omega_j(\mathbf{k})\omega_j(\mathbf{k}_1)\omega_j(\mathbf{k}_2),$$
$$\tilde{\gamma}_3 = \frac{\gamma_3}{(\gamma_\parallel^2\gamma_\perp)^{1/2}}. \tag{10}$$

Here, $\gamma_\parallel$, $\gamma_\perp$ ($\gamma_\perp \gg \gamma_\parallel$), and γ_3 are effective harmonic and anharmonic force constants. It should be noted that these constants satisfy the approximate relation (accurate to within an order of magnitude)

$$\frac{\gamma_3^2\omega_{j\max}}{\gamma_\parallel^2\gamma_\perp} = \tilde{\gamma}_3^2\omega_{j\max} \approx 10\frac{\langle u^2\rangle}{a^2}$$
$$= 10\delta_A^j = 10\frac{(\gamma_G^j)^2 h}{M\omega_{j\max}a^2}, \tag{11}$$

where $\langle u^2\rangle$ is the mean square of atomic displacements, $\omega_{j\max} \approx \omega_{3(\perp)}$ is the maximum frequency in the acoustic spectrum, *a* is the characteristic atomic spacing, δ_A is the anharmonicity parameter, γ_G^j is the partial Grüneisen factor for the *j*th vibrational mode, *M* is the mass of an atom, and *h* is Planck's constant. The value of δ_A can be of the order of 10^{-2}–10^{-1}, and not 10^{-3} (see, for example, [23]). The individual terms of the polarization operator appearing in relation (9) in view of relations (10) are defined as

$$\Pi_1^j(\mathbf{k}, \omega) = i\tilde{\gamma}_3^2\omega_j^2(\mathbf{k})\int_0^\omega \frac{d\omega_1}{2\pi}\sum_{\mathbf{k}_1}\omega_j^2(\mathbf{k}_1)\omega_j^2(\mathbf{k}-\mathbf{k}_1)$$
$$\times \overline{G}_j^+(\mathbf{k}_1, \omega_1)\overline{G}_j^-(\mathbf{k}-\mathbf{k}_1, \omega-\omega_1) \tag{12}$$
$$\times\left[1-\exp\left(\frac{\omega_1-\omega}{T}\right)\right]^{-1},$$

$$\Pi_2^j(\mathbf{k}, \omega) = i\tilde{\gamma}_3^2\omega_j^2(\mathbf{k})\int_0^\omega \frac{d\omega_1}{2\pi}\sum_{\mathbf{k}_1,\mathbf{q}}\omega_j^2(\mathbf{k}_1)\omega_j^2(\mathbf{k}-\mathbf{k}_1)$$
$$\times \overline{G}_j^+(\mathbf{k}_1, \omega_1)\overline{G}_j^-(\mathbf{k}-\mathbf{k}_1, \omega-\omega_1)U^j(\mathbf{q}; \omega, \omega_1)$$
$$\times\left[1-\exp\left(\frac{\omega_1-\omega}{T}\right)\right]^{-1} \tag{13}$$
$$\times \overline{G}_j^+(\mathbf{k}_1-\mathbf{q}, \omega_1)\overline{G}_j^-(\mathbf{k}+\mathbf{q}-\mathbf{k}_1, \omega-\omega_1).$$

The upper integration limit in these relations is determined by the energy conservation law.

4. ATTENUATION OF HIGH-FREQUENCY SOUND

Let us now consider the attenuation of high-frequency sound in a chain-type compound with a diagonal disorder. The propagation of sound depends on the elasticity of the crystal lattice. Its attenuation $\Gamma^j(\omega)$ is determined by the imaginary part of the polarization operator of the one-particle lattice Green's function:

$$\Gamma^j(\omega) = \Gamma_1^j(\omega) + \Gamma_2^j(\omega)$$
$$= \frac{\mathrm{Im}\Pi_1^j(\omega_j(\mathbf{k})) + \mathrm{Im}\Pi_2^j(\omega_j(\mathbf{k}))}{2\omega_j(\mathbf{k})}. \tag{14}$$

We assume that $\omega \approx \omega_j(\mathbf{k})$. In order to obtain the required order-of-magnitude estimates of $\mathrm{Im}\Pi_1^j$ and $\mathrm{Im}\Pi_2^j$, we confine ourselves to three-phonon processes of a phonon decay into two phonons with close energies. In addition, we will consider the situation when the standard anharmonic interaction of thermal phonons can be disregarded in a nonideal crystal, i.e., when the condition $\tau_i^{-1}(\omega) \gg \tau_N^{-1}(\omega)$ holds, where τ_N is the time of relaxation due to normal anharmonic decay processes. The latter inequality is equivalent to the condition

$$c\epsilon^2 \gg \delta_A^j\frac{\omega}{\omega_{j\max}}F\left(\frac{\omega}{2T}\right),$$
$$F\left(\frac{\omega}{2T}\right) = \left[1-\exp\left(-\frac{\omega}{2T}\right)\right]^{-1}, \quad T < \omega. \tag{15}$$

Here, $\omega_{j\max} \approx \omega_\perp$ for $j = l$ and $\omega_{j\max} \approx \omega_3$ for $j = b$. Taking into account what has been said above, we determine the damping factor Γ^j in the high-frequency limit $\omega\tau_i^j(\omega) \gg 1$ ($\omega > T$). Substituting relation (7) into formulas (12) and (13) and taking into account relation (14), we obtain

$$\Gamma_1^j(\omega) = \frac{\delta_A^j}{c\epsilon^2}\frac{\omega}{\omega_{j\max}}\omega_j(\mathbf{k})F\left(\frac{\omega}{2T}\right), \tag{16}$$

$$\Gamma_2^j(\omega) = \frac{\delta_A^{l(b)}}{c\epsilon^2}\frac{\omega_j(\mathbf{k})}{\tau_i^{l(b)}(\omega)\omega_{\perp(3)}}\left(\frac{\omega}{\omega_{\parallel(2)}}\right)^2 F\left(\frac{\omega}{2T}\right). \tag{17}$$

For comparison, we write the attenuation of an acoustic quantum of frequency ω due to spontaneous decay in the absence of defects:

$$\Gamma_0^l(\omega) = \frac{1}{3}\delta_A^l\omega\left(\frac{\omega}{\omega_\perp}\right)^2 F\left(\frac{\omega}{2T}\right), \tag{18}$$

$$\Gamma_0^b(\omega) = \frac{1}{5}\delta_A^b\omega\left(\frac{\omega}{\omega_3}\right)^{3/2} F\left(\frac{\omega}{2T}\right). \tag{19}$$

The mechanism of sound damping described by relations (16) and (17) is significant in the low-temperature

region ($T < \omega$), where the scattering of decay phonons is sensitive to defects. The expression for Γ_2^j in the form (17) is valid in the frequency range $\omega_{\perp(3)}^2 > \omega^2 \gg 2\omega_{\parallel(2)}^2$ in which the dispersion relations (3) and (4) exhibit a quasi-one-dimensional behavior. While deriving relation (17), we also assumed that the chain "locking" parameter $\omega_{\parallel(2)}^2 \tau_i^{l(b)}(\omega)/\omega \ll 1$. In order to estimate the relative contribution of the weak localization effect to sound damping, we compare Γ_1^j and Γ_2^j. Relations (16) and (17) lead to (see also [7])

$$\frac{\Gamma_2^l}{\Gamma_1^l} = c\epsilon^2 \frac{\omega}{\omega_\perp}\left(\frac{\omega}{\omega_\parallel}\right)^2,$$
$$\frac{\Gamma_2^b}{\Gamma_1^b} = c\epsilon^2 \left(\frac{\omega}{\omega_3}\right)^{1/2}\left(\frac{\omega}{\omega_2}\right)^2. \quad (20)$$

It should be noted that a comparison of Γ_2^j (17) and Γ_0^j (18), (19) for $\omega_{\parallel(2)} < \omega$ shows that $\Gamma_2^j > \Gamma_0^j$.

Let us consider relations (20). It was found that, if the measure of crystal "defectiveness" $c\epsilon^2 \leq 1$ and the values of $\omega_{\parallel(2)}$ characterizing the intensity of interaction between the chains are smaller than ω, a situation when $\Gamma_2^j \leq \Gamma_1^j$ is possible in the low-temperature region ($T < \omega$), where the anharmonic decay of phonons is strong and their fusion is insignificant. In other words, the effect of backward coherent scattering on sound damping may be significant due to a large phase volume for the quasi-one-dimensional dynamic behavior of a chain-type crystal. It was noted above that $\omega_3 \approx \omega_\perp$, $\omega_1 \approx \omega_\parallel$, and $\omega_1 \ll \omega_2$. Consequently, in accordance with relations (20), the effect of the weak localization on the longitudinal vibrational mode damping for a fixed parameter $c\epsilon^2$ is stronger than for the bending vibrational mode. We will use the experimental results obtained in [18] for the compound $(Ta_{1-x}Nb_xSe_4)_2I$: $\omega_\perp \approx 1$ THz, $\omega_\parallel \approx 0.15$ THz, and $\omega_2 \approx 0.5$ THz. For record-high frequencies $\omega \approx 0.1$ THz of sound waves used in acoustic measurements and for a disorder parameter $c\epsilon^2 \approx 1$, we obtain the following numerical estimate of relations (20): $\Gamma_2^l/\Gamma_1^l \approx 0.1$ and $\Gamma_2^b/\Gamma_1^b \approx 0.01$.

The reciprocal sound attenuation length (the imaginary part of the wave number) corresponding to the contribution of the weak localization effect [21, 24] is defined by the relation

$$\frac{1}{l_{\perp(\parallel)}^j(\omega)} = \frac{\Gamma_2^j(\omega)}{v_{\perp(\parallel)}^j(\omega)}, \quad (21)$$

where $v_{\perp(\parallel)}^j$ are the mean group velocities of the jth vibrational mode along and across the chains, respectively [7]. Using relations (17) and (2.17) and (2.18) from [7], we obtain

$$\frac{1}{l_\perp^l} = \sqrt{2}\frac{\delta_A^l}{b}\left(\frac{\omega}{\omega_\perp}\right)^3\left(\frac{\omega}{\omega_\parallel}\right)^2,$$
$$\frac{1}{l_\parallel^l} = 2\frac{\delta_A^l}{a}\left(\frac{\omega}{\omega_\perp}\right)^2\left(\frac{\omega}{\omega_\parallel}\right)^4; \quad (22)$$

$$\frac{1}{l_\perp^b} = \frac{\sqrt{\pi}}{2}\frac{\delta_A^b}{b}\left(\frac{\omega}{\omega_3}\right)^2\left(\frac{\omega}{\omega_2}\right)^2,$$
$$\frac{1}{l_\parallel^b} = 4\frac{\delta_A^b}{a}\left(\frac{\omega}{\omega_3}\right)^{3/2}\left(\frac{\omega}{\omega_2}\right)^4. \quad (23)$$

Pay attention to the fact that all reciprocal values of length defined by relations (22) and (23) exhibit different frequency dependences. This allows us to identify this mechanism of sound damping from the frequency dependence. We obtain numerical estimates of the quantities $1/l_{\parallel(\perp)}^j$ using the values of the unit cell parameters $a = 9.5$ Å and $c = 12.8$ Å for the system $(Ta_{1-x}Nb_xSe_4)_2I$ [18]. If we assume that the value of anharmonicity parameter δ_A is of the order of 10^{-2}, we obtain $1/l_\perp^l \approx 50$ cm^{-1}, $1/l_\parallel^l \approx 400$ cm^{-1}, $1/l_\perp^b \approx 60$ cm^{-1}, and $1/l_\parallel^b \approx 2600$ cm^{-1}.

5. CONCLUSIONS

We have analyzed the effect of weak localization of acoustic vibrational modes on the damping factor of high-frequency sound in a nonideal strongly anisotropic chain-type anharmonic crystal lattice. We considered longitudinally polarized excitations as well as excitations resembling bending waves in noninteracting chains. It is shown that, in the case when elastic scattering by defects dominates over three-phonon decay processes, specific interference scattering processes lead to noticeable renormalizations of the sound damping factor. The anharmonic interference mechanism can in principle prevail over the Rayleigh and standard anharmonic mechanisms. Unfortunately, we are not aware of experimental data on the propagation of acoustic excitations in substantially disordered quasi-one-dimensional crystals. The renormalization of the damping factor of high-frequency sound predicted in this work can be observed experimentally in compounds of the carbolite type [19] using the scheme of thermal pulses [4] or acoustic measurements in the microwave region.

ACKNOWLEDGMENTS

The author thanks the reviewer for valuable remarks.

REFERENCES

1. V. L. Pokrovskiĭ and A. M. Dykhne, Zh. Éksp. Teor. Fiz. **39**, 720 (1960) [Sov. Phys. JETP **12**, 503 (1960)].
2. D. V. Kazakovtsev and I. B. Levinson, Pis'ma Zh. Éksp. Teor. Fiz. **27**, 194 (1978) [JETP Lett. **27**, 181 (1978)].
3. D. V. Kazakovtsev and Y. B. Levinson, Phys. Status Solidi B **96**, 117 (1979).
4. S. N. Ivanov, A. V. Taranov, and E. N. Khazanov, Zh. Éksp. Teor. Fiz. **99**, 1311 (1991) [Sov. Phys. JETP **72**, 731 (1991)].
5. S. N. Ivanov, E. N. Khazanov, P. Paszkiewicz, *et al.*, Z. Phys. B **99**, 535 (1996).
6. B. A. Danil'chenko, S. N. Ivanov, D. V. Poplavskiĭ, *et al.*, Zh. Éksp. Teor. Fiz. **112**, 325 (1997) [JETP **85**, 179 (1997)].
7. A. P. Zhernov and E. P. Chulkin, Zh. Éksp. Teor. Fiz. **117**, 350 (2000) [JETP **90**, 308 (2000)].
8. E. P. Chulkin, A. P. Zhernov, and T. N. Kulagina, Fiz. Nizk. Temp. **26**, 173 (2000) [Low Temp. Phys. **26**, 128 (2000)].
9. M. A. Ivanov and Yu. V. Skripnik, Fiz. Tverd. Tela (Leningrad) **32**, 2965 (1990) [Sov. Phys. Solid State **32**, 1722 (1990)].
10. M. A. Ivanov, A. M. Kosevich, E. S. Syrkin, *et al.*, Fiz. Nizk. Temp. **19**, 434 (1993) [Low Temp. Phys. **19**, 305 (1993)].
11. A. P. Zhernov, E. I. Salamatov, and E. P. Chulkin, Phys. Status Solidi B **165**, 355 (1991).
12. A. L. Burin, L. A. Maksimov, and I. Y. Polishchuk, Physica B (Amsterdam) **210**, 49 (1995).
13. I. Y. Polishchuk, L. A. Maksimov, and A. L. Burin, Phys. Rep. **288**, 205 (1997).
14. A. P. Zhernov and E. P. Chulkin, Zh. Éksp. Teor. Fiz. **113**, 930 (1998) [JETP **86**, 507 (1998)].
15. I. M. Lifshits, Zh. Éksp. Teor. Fiz. **22**, 475 (1952).
16. A. M. Kosevich, *Theory of Crystal Lattice* (Vishcha Shkola, Khar'kov, 1988).
17. E. G. Brovman, Yu. Kagan, and A. Kholas, Zh. Éksp. Teor. Fiz. **61**, 2429 (1971) [Sov. Phys. JETP **34**, 1300 (1971)].
18. J. E. Lorenzo, R. Currat, P. Monceau, *et al.*, J. Phys.: Condens. Matter **10**, 5039 (1998).
19. A. F. Gurov, V. N. Kopylov, K. Kusano, *et al.*, Phys. Rev. B **56**, 11629 (1997).
20. Yu. M. Kagan, in *Proceedings of School on Theory of Defects in Crystals, Tbilisi, 1969*, Vol. 2, p. 93.
21. A. P. Zhernov and E. P. Chulkin, Zh. Éksp. Teor. Fiz. **109**, 602 (1996) [JETP **82**, 321 (1996)].
22. A. P. Zhernov and E. P. Chulkin, Phys. Status Solidi B **193**, 67 (1996).
23. M. Saint-Paul, S. Holtmeier, R. Britel, *et al.*, J. Phys.: Condens. Matter **8**, 2021 (1996).
24. E. P. Chulkin, A. P. Zhernov, and T. N. Kulagina, Fiz. Nizk. Temp. **25**, 1218 (1999) [Low Temp. Phys. **25**, 912 (1999)].

Translated by N. Wadhwa

Journal of Experimental and Theoretical Physics, Vol. 95, No. 5, 2002, pp. 886–900.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 122, No. 5, 2002, pp. 1027–1043.

SOLIDS
Structure

Lattices of Nonspherical Ferromagnetic Grains with Magnetic Dipole Interaction: Theory and Experimental Examples

E. Z. Meĭlikhov* and R. M. Farzetdinova

Institute of Molecular Physics, Russian Research Centre Kurchatov Institute, pl. Kurchatova 1, Moscow, 123182 Russia

**e-mail: meilikhov@imp.kiae.ru*

Received May 16, 2002

Abstract—The magnetic properties (ground state, magnetic phase diagram, and phase transitions in a magnetic field) of two- and three-dimensional lattices of ferromagnetic grains with the intergrain dipole interaction are studied. The main attention is paid to the lattices formed by nonspherical grains (prolate and oblate ellipsoids of revolution) and their extreme forms (rodlike and disc-shaped grains). An analysis shows that the conclusions of the theory are in good agreement with the results of experiments. © *2002 MAIK "Nauka/Interperiodica"*.

1. INTRODUCTION

As a result of rapid development of technology, magnetic nanostructures have been transformed from the object of academic studies to the object of experimental investigation and have become elements of real microelectronic (nanoelectronic) devices. The term "nanostructure" will be applied here only to a definite type of such systems, namely, regular three- and two-dimensional (3D and 2D) lattices formed by identical one-domain ferromagnetic nanometer-size grains. We assume that the exchange interaction only establishes a ferromagnetic ordering of elementary magnetic moments in grains, while the interaction between the grains is of the pure dipole type (this means that the separation between the grains is larger than the characteristic range of the exchange interaction). In addition, we confine our analysis to grains in the form of (prolate and oblate) ellipsoids of revolution or their extreme forms: pointlike (spherical), rodlike and disk-shaped grains. In all cases (except for grains in the form of oblate ellipsoids whose magnetic moment lies in the equatorial plane), we assume that the magnetic moment of a grain is directed along its easy-magnetization axis (easy axis) determined either by the geometry (rods and prolate ellipsoids) or by crystal anisotropy (pointlike or spherical grains).

Real nanostructures interesting for practical applications are constructed so that the directions of easy axes or planes of the grains are parallel. In the former case, this makes unnecessary the analysis of different types of magnetic vortex states typical of a system of dipoles with a freely rotating moment [1]. The magnetic properties of some simple structures of this type were studied in detail earlier (simple lattices of pointlike grains were mainly studied). For example, it was shown in [2, 3] that the ground state of 2D square and 3D cubic lattices formed by such grains with parallel (antiparallel) magnetic moments is always antiferromagnetic. This conclusion also remains valid for systems consisting of uniformly magnetized (one-domain) spherical grains since the field of such a grain coincides with the field of an equivalent pointlike dipole placed at the center of the grain, and the energy of dipole interaction of such grains coincides with the energy of interaction between two equivalent pointlike dipoles [4]. It was found, however, that lowering of the symmetry of any element of the system (lattice type or the shape of grains) may facilitate its transition to a magnetic state more advantageous from the energy point of view.

In particular, this is due to the fact that, when the shape of a grain differs considerably from a sphere, its field at small distances from the surface differs from the field of an equivalent dipole placed at its center. For a prolate ellipsoid, its field at points close to the equatorial plane of the grain is much weaker than the field of an equivalent dipole; the opposite situation is observed for a grain in the form of an oblate ellipsoid.[1] Consequently, a 2D lattice of nonspherical grains whose major axes lie in the lattice plane has a larger tendency to ferromagnetism (than for a spherical shape of grains) in the first case and to antiferromagnetism in the second case [5].

In view of various fundamental, technological, and technical circumstances, real systems are virtually never symmetric (in the above sense), and their asym-

[1] The magnitude of the field in the equatorial (relative to the direction of the magnetic moment) plane constitutes 50% of the field at its axis for a spherical grain, 30% for a grain in the form of a strongly prolate ellipsoid of revolution, and 85% for a grain in the form of a strongly oblate ellipsoid of revolution.

1063-7761/02/9505-0886$22.00

metry is characterized by a variety of parameters. There are numerous examples of such systems, including planar periodic structures of nonspherical magnetic grains, which are being extensively investigated at present and are treated as the media with considerable potential for elevating the magnetic recording density [6]. Typical structures of this kind have the form of 2D rectangular lattices of one-domain extended magnetic grains with uniaxial geometric anisotropy. The shape of such grains resembles an ellipsoid of revolution with an axes ratio of 3–5; the period of such lattices is comparable with the grain size. Another example related to the problem under investigation is magnetic dielectric nanocomposites. The electrical conduction of such a (3D) system is due to tunnel electron transitions between the grains [7], whose probability is determined by the mutual orientation of magnetic moments of adjacent grains. Consequently, it is clear that the resistance of such a medium depends directly on its magnetic state. The same refers to the (giant) magnetoresistance of such a system. Finally, it is appropriate to mention a slightly unexpected object in the context of our discussion, namely, ultrathin films of ferromagnetic metals on monocrystalline substrates. It was found that, for a certain effective thickness of such films, a long-range ferromagnetic order is established in them [8]. It was shown in a recent publication [9] that, under certain conditions, such films grow through natural lithography, i.e., via the formation of nuclei which are distributed more or less uniformly over the substrate area and which are transformed, during their subsequent growth, into ellipsoidal grains whose major axes are parallel to one another.

Thus, systems of the type under investigation are quite diverse. This work aims at an analysis of their magnetic properties such as the ground state, the magnetic phase diagram, and magnetic phase transitions. In contrast to known publications (see, for example, [1–3]) in which similar systems were investigated, we consider more general oblique lattices formed by nonspherical (or, which is the same, non-pointlike) dipoles. It is shown that the inclusion of these new circumstances modifies the properties of such systems considerably.

In the final section of this paper, the obtained results are used to describe the properties of a number of real systems.

2. MAGNETIC STATE OF A SYSTEM OF FERROMAGNETIC GRAINS WITH DIPOLE INTERACTION

Let us consider a system of identical one-domain ferromagnetic grains whose centers are at the sites of a 3D tetragonal lattice with periods l_x, l_y, and l_z, and whose magnetic moments $\boldsymbol{\mu}$ are identical and parallel to one another. Such a magnetic anisotropy may be due to crystal anisotropy of grains or anisotropy in their shape, or is manifested (in the absence of the former anisotropies) in a weak external magnetic field (and become decisive at a low temperature). Further, we assume that these moments form angle θ (latitude) with the x axis and are turned through an angle φ (longitude) about the x axis relative to the xz plane. On account of the long-range dipole–dipole interaction, the magnetic state of such a system is determined by the magnetic field $H_d = \sum'_{m,l,n} H_{mln}$ created in the volume of an individual grain (located, for definiteness, at the origin of coordinates) by all the remaining grains. Here, H_{mln} is the field component created at the origin by a grain with the center at point (ml_x, ll_y, nl_z), where m, l, and n are integers, and parallel to magnetic moments of the grains; the sum does not contain the term with $m = l = n = 0$.

The energy of interaction between the chosen grain and the magnetic field H_d is given by

$$W_d = -\frac{\mu}{V}\int_V H_d(r)\,dV, \tag{1}$$

where integration is carried out over the volume of the central grain. The ground state of the system corresponds to the configuration of the magnetic moments $\boldsymbol{\mu}_{mln}$ of the grains, for which the energy W_d attains its minimum value. In our case, there exists only one ferromagnetic configuration (the magnetic moments of grains are directed along their major axes and are parallel to one another); for the antiferromagnetic state of the system, we confine the analysis to the situation when the magnetic moments of grains form two identical magnetic sublattices of antiparallel magnetic moments. In this case,

$$\boldsymbol{\mu}_{mln} = \boldsymbol{\mu}\exp[i\pi(\phi_x m + \phi_y l + \phi_z n)],$$

where phases ϕ_x, ϕ_y, ϕ_z may assume the values 0 or 1; for the ferromagnetic state, $\phi_x = \phi_y = \phi_z = 0$ or $\phi_x = \phi_y = \phi_z = 1$, while six different antiferromagnetic states correspond to different combinations of these phases, in which at least one phase is equal to zero and at least one phase is equal to unity.[2]

Let i_1 be the absolute value of magnetization of each of the sublattices (the total magnetization of the system is $i = 2i$ in the ferromagnetic state and $i = 0$ in the antiferromagnetic state). In accordance with the mean-field theory, the energy of the dipole–dipole interaction is proportional to magnetization i_1: $W_d = -\lambda i_1$. From all the states under investigation, the one that is most advantageous from the energy point of view corresponds to the highest value of the dipole interaction parameter λ. Thus, in order to determine the type of the ground state of the system under investigation, it is sufficient to find and compare the values of these parameter corresponding to the ferromagnetic and various anti-

[2] Any other set of integral phases ϕ_x, ϕ_y, ϕ_z is identical to one of the eight sets enumerated above.

ferromagnetic configurations of the magnetic moments of the grains.

As regards the magnetization i_1 of the sublattices, it is defined by the self-consistent equation of the mean-field theory,

$$i_1 = i_0 \tanh[\lambda i_1/kT], \tag{2}$$

where $i_0 = (1/2)\mu N_g$, $N_g = 1/(l_x l_y l_z)$ being the grain concentration. This equation determines, as usual, the Curie temperature $T_C = \lambda i_0/k$ of the system. It should be noted, however, that it is applicable only in the case when the difference in the energies W_d of the ferromagnetic and antiferromagnetic states of the system is larger than kT. Otherwise, we must take into account thermal fluctuations of the magnetic order.

We will consider below the following three situations: (i) pointlike (zero-dimensional) but anisotropic grains, (ii) rodlike (one-dimensional) grains magnetized along their axes, and (iii) three-dimensional grains in the form of prolate and oblate ellipsoids of revolution (including the extreme case of an oblate ellipsoid, viz., two-dimensional disks). The first two situations are interesting since, on the one hand, they permit an exact solution and, on the other hand, are limiting cases for the third, much more real, situation, which can be described only numerically.

While calculating the dipole interaction energy, the following circumstance should be borne in mind. It was noted in [2] that the energy W_F of the ferromagnetic state strongly depends on the sample shape:

$$-W_F \propto \lambda_F \propto (N - 4\pi/3),$$

where N is the demagnetization factor in the direction of the magnetic moment (the energy of any antiferromagnetic state is independent of N). With increasing N (corresponding to the extension of the sample in this direction), parameter λ_F increases, and the ferromagnetic state could become advantageous. Calculations show, however, that this does not take place even in very extended samples with a simple cubic lattice of pointlike dipoles, and their ground state remains antiferromagnetic. Nevertheless, the sample shape must be fixed in a certain way in calculations. All the results described below correspond to spherical (actually, cubic) samples for which $N = 4\pi/3$ irrespective of parameters β_y and β_z of lattice extension along the y and z axes.

2.1. 3D Lattice of Pointlike Grains

In this case (which is realized when the separation between the grains is much larger than their size), the magnetic field component for each grain at the origin under investigation is given by

$$H_{mln} = \mu\left[-\frac{1}{R_{mln}^3} + \frac{3(\mathbf{e}_0 \cdot \mathbf{R}_{mln})^2}{R_{mln}^5}\right] \times \exp(i\pi(\phi_x m + \phi_y l + \phi_z n)), \tag{3}$$

where $\mathbf{R}_{mln} = (ml_x, ll_y, nl_z)$ is the radius vector of the grain (m, l, n) and $\mathbf{e}_0 = (\cos\theta, \sin\theta\sin\varphi, \sin\theta\cos\varphi)$ is the unit vector parallel to the magnetic moments of the grains. In this case, the magnetic interaction energy has the form

$$W_d = -\mu H_d,$$

$$H_d = \sum_{m,l,n}{}' H_{mln} = \mu \sum_{m,l,n}{}'\left[-\frac{1}{R_{mln}^3} + \frac{3(\mathbf{e}_0 \cdot \mathbf{R}_{mln})^2}{R_{mln}^5}\right] \times \exp(i\pi(\phi_x m + \phi_y l + \phi_z n)). \tag{4}$$

Using this relation, we obtain

$$\lambda = 2\mu\beta_y\beta_z \sum_{m,l,n}{}'\left[-\frac{1}{(m^2 + \beta_y l^2 + \beta_z^2 n^2)^{3/2}} + \frac{3[m\cos\theta + (\beta_y l\sin\varphi + \beta_z n\cos\varphi)\sin\theta]^2}{(m^2 + \beta_y^2 l^2 + \beta_z^2 n^2)^{5/2}}\right] \times \exp(i\pi(\phi_x m + \phi_y l + \phi_z n)), \tag{5}$$

where $\beta_y = l_y/l_x$ and $\beta_z = l_z/l_x$ are the lattice extension coefficients along the y and z axis, respectively. Taking into account the fact that all the sums whose terms are odd relative to the summation variables m, l, n are equal to zero for a tetragonal lattice, we can represent relation (5) in the form

$$\lambda = 2\mu\beta_y\beta_z \times [\lambda_0\cos^2\theta + \lambda_{\pi/2}\sin^2\theta + 3\lambda_\varphi\sin^2\theta\sin^2\varphi], \tag{6}$$

where

$$\lambda_0 = \sum_{m,l,n}{}' \frac{(2m^2 - \beta_y^2 l^2 - \beta_z^2 n^2)}{(m^2 + \beta_y^2 l^2 + \beta_z^2 n^2)^{5/2}} \times \exp(i\pi(\phi_x m + \phi_y l + \phi_z n)), \tag{7}$$

$$\lambda_{\pi/2} = \sum_{m,l,n}{}' \frac{(-m^2 - \beta_y^2 l^2 + 2\beta_z^2 n^2)}{(m^2 + \beta_y^2 l^2 + \beta_z^2 n^2)^{5/2}} \times \exp(i\pi(\phi_x m + \phi_y l + \phi_z n)), \tag{8}$$

$$\lambda_\varphi = \sum_{m,l,n}{}' \frac{(\beta_y^2 l^2 - \beta_z^2 n^2)}{(m^2 + \beta_y^2 l^2 + \beta_z^2 n^2)^{5/2}} \times \exp(i\pi(\phi_x m + \phi_y l + \phi_z n)). \quad (9)$$

The transformation used in [10] considerably improves the convergence of the dipole sums (7)–(9), which can be represented in the form of the sums of exponentially decreasing terms:

$$\lambda_0 = \left(\frac{4\pi}{3\beta_y\beta_z}\right) \sum_{m,n=-\infty}^{\infty} (2X_{mn} - Y_{mn} - Z_{mn}), \quad (10)$$

$$\lambda_{\pi/2} = \left(\frac{4\pi}{3\beta_y\beta_z}\right) \sum_{m,n=-\infty}^{\infty} (2Z_{mn} - Y_{mn} - X_{mn}), \quad (11)$$

$$\lambda_\varphi = \left(\frac{4\pi}{3\beta_y\beta_z}\right) \sum_{m,n=-\infty}^{\infty} (Y_{mn} - Z_{mn}), \quad (12)$$

where

$$X_{mn} = \sum_{l=1}^{\infty}\left(\frac{1}{l} + \alpha_{mn}^{(x)}\right)\cos(l\pi\phi_x)\exp(-\alpha_{mn}^{(x)}l),$$

$$\alpha_{mn}^{(x)} = 2\pi\left[\left(\frac{m+\phi_y/2}{\beta_y}\right)^2 + \left(\frac{n+\phi_z/2}{\beta_z}\right)^2\right]^{1/2},$$

$$Y_{mn} = \sum_{l=1}^{\infty}\left(\frac{1}{l} + \alpha_{mn}^{(y)}\right)\cos(l\pi\phi_y)\exp(-\alpha_{mn}^{(y)}l),$$

$$\alpha_{mn}^{(y)} = 2\pi\left[\left(m+\frac{\phi_x}{2}\right)^2 + \left(\frac{n+\phi_z/2}{\beta_z}\right)^2\right]^{1/2},$$

$$Z_{mn} = \sum_{l=1}^{\infty}\left(\frac{1}{l} + \alpha_{mn}^{(z)}\right)\cos(l\pi\phi_z)\exp(-\alpha_{mn}^{(z)}l),$$

$$\alpha_{mn}^{(z)} = 2\pi\left[(m+\phi_x/2)^2 + \left(\frac{n+\phi_y/2}{\beta_y}\right)^2\right]^{1/2}.$$

The evaluation and comparison of the values of the dipole interaction parameter λ for different magnetic configurations makes it possible to construct the magnetic phase diagram of the system. In the case under investigation, when there exist six simple two-sublattice antiferromagnetic configurations along with the ferromagnetic configuration, such a diagram may be quite complicated. Its form is determined by four parameters corresponding to lattice extension (β_y, β_z) and the direction of the magnetic moment of the dipoles (θ, ϕ).

Figures 1a and 1b show, by way of an example, the angular dependences of the dipole interaction parameters for two different magnetic configurations, one of which corresponds to a cubic lattice of dipoles with magnetic moments perpendicular to the y axis ($\phi = 0$), while the other corresponds to a tetragonal lattice with dipoles oriented so that $\phi \neq 0$. It can be seen that the ground state of the simple cubic lattice of pointlike dipoles is antiferromagnetic, although the specific form of the corresponding magnetic configuration is determined by the values of the above parameters. A more detailed analysis confirms the general nature of this conclusion for any tetragonal lattice of pointlike dipoles.

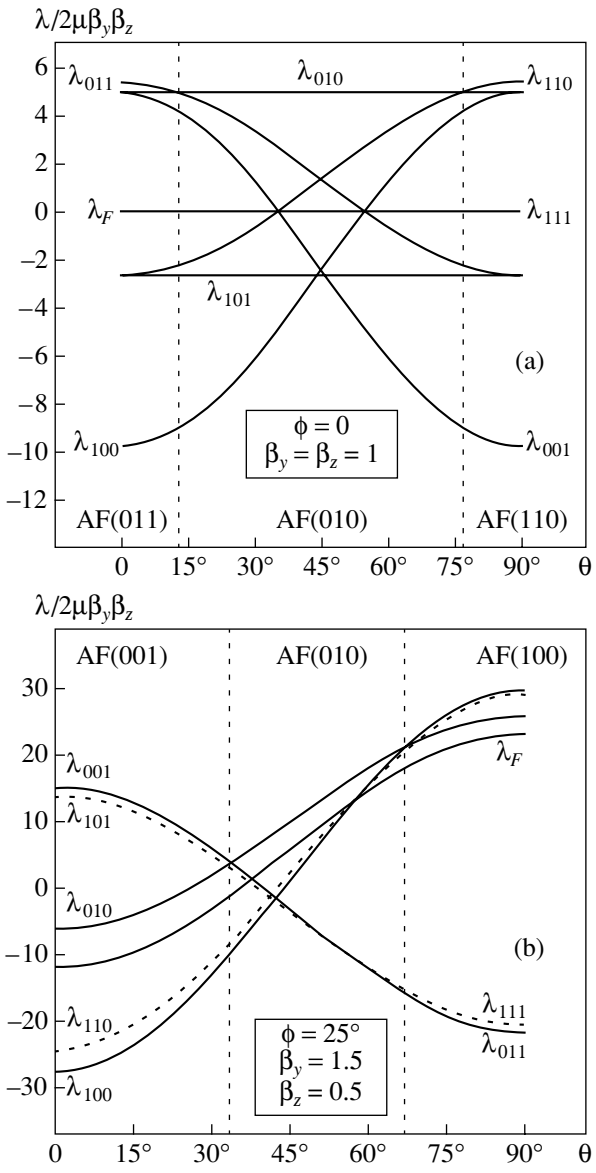


Fig. 1. Angular dependences $\lambda(\theta)$ of the dipole interaction parameters for a (a) cubic and (b) tetragonal lattice of pointlike dipoles. The domains of ferromagnetic (F) and antiferromagnetic (AF) states are indicated. Subscript F correspond to the ferromagnetic configuration, and subscripts 100, 110, etc. correspond to the values of phases ϕ_x, ϕ_y, and ϕ_z.

2.2. 2D Lattice of Pointlike Grains

For obvious reasons, it is more convenient to analyze the properties of the system under investigation for a 2D lattice of anisotropic pointlike magnetic dipoles. In order to obtain the required relations, it is sufficient to set $l = 0$ in all the above formulas. In this case, we obtain the following expression for the dipole interaction parameter:[3]

$$\lambda = 2\left(\frac{\mu\beta}{l_x}\right) \times \{(3\cos^2\theta - 1)\Sigma_m + \beta^2(3\sin^2\theta - 1)\Sigma_n\}, \quad (13)$$

where $\beta \equiv \beta_z$ and

$$\Sigma_m = \sum_{m,n}{}' \frac{m^2 e^{i\pi(\phi_x m + \phi_z n)}}{(m^2 + \beta^2 n^2)^{5/2}}, \qquad \Sigma_n = \sum_{m,n}{}' \frac{n^2 e^{i\pi(\phi_x m + \phi_z n)}}{(m^2 + \beta^2 n^2)^{5/2}}. \quad (14)$$

The transformation used in [10] considerably improves the convergence of the dipole sums (14), which can be represented after this in the form

$$\Sigma_m = \frac{16\pi^2}{3\beta^3}\sum_{m=1}^{\infty}\sum_{k=-\infty}^{\infty}\cos(\pi m\phi_x)\left(k + \frac{\phi_z}{2}\right)^2 \times K_2\left[(2\pi m/\beta)\left|k + \frac{\phi_z}{2}\right|\right], \quad (15)$$

$$\Sigma_n = \frac{16\pi^2}{3\beta^2}\sum_{n=1}^{\infty}\sum_{k=-\infty}^{\infty}\cos(\pi n\phi_z)\left(k + \frac{\phi_x}{2}\right)^2 \times K_2\left[2\pi n\beta\left|k + \frac{\phi_x}{2}\right|\right], \quad (16)$$

where K_2 is the Macdonald function. The latter rapidly decreases exponentially upon an increase in the argument; consequently, it is sufficient to retain only those terms in sums (15) and (16) which correspond to the minimal (in absolute value) arguments of this function. Simple approximate expressions for these sums, which correspond to different magnetic configurations, have the following form:

$$\Sigma_m = \frac{4\pi^2}{9\beta}\left\{1 + \frac{24}{\beta^2} \times \left[K_2\left(\frac{2\pi}{\beta}\right) + 5K_2\left(\frac{4\pi}{\beta}\right) + 10K_2\left(\frac{6\pi}{\beta}\right)\right]\right\}, \quad (17)$$

$$\Sigma_n = \frac{4\pi^2}{9\beta^4}\{1 + 24\beta^2 \times [K_2(2\pi\beta) + 5K_2(4\pi\beta) + 10K_2(6\pi\beta)]\} \quad (18)$$

for the ferromagnetic configuration ($\phi_x = 0$, $\phi_z = 0$);

$$\Sigma_m = \frac{8\pi^2}{3\beta^3} \times \left[K_2\left(\frac{\pi}{\beta}\right) + K_2\left(\frac{2\pi}{\beta}\right) + 10K_2\left(\frac{3\pi}{\beta}\right)\right], \quad (19)$$

$$\Sigma_n = -\frac{2\pi^2}{9\beta^4}\{1 + 48\beta^2 \times [K_2(2\pi\beta) + 3K_2(4\pi\beta) + 10K_2(6\pi\beta)]\} \quad (20)$$

for the configuration S_{01} ($\phi_x = 0$, $\phi_z = 1$);

$$\Sigma_m = -\frac{2\pi^2}{9\beta}\left\{1 + \frac{48}{\beta^2} \times \left[K_2\left(\frac{2\pi}{\beta}\right) + 3K_2\left(\frac{4\pi}{\beta}\right) + 10K_2\left(\frac{6\pi}{\beta}\right)\right]\right\}, \quad (21)$$

$$\Sigma_n = \frac{8\pi^2}{3\beta^2}[K_2(\pi\beta) + K_2(2\pi\beta) + 10K_2(3\pi\beta)] \quad (22)$$

for the configuration S_{10} ($\phi_x = 1$, $\phi_z = 0$); and

$$\Sigma_m = -\frac{8\pi^2}{3\beta^3} \times \left[K_2\left(\frac{\pi}{\beta}\right) - K_2\left(\frac{2\pi}{\beta}\right) + 10K_2\left(\frac{3\pi}{\beta}\right)\right], \quad (23)$$

$$\Sigma_n = -\frac{8\pi^2}{3\beta^2}[K_2(\pi\beta) - K_2(2\pi\beta) + 10K_2(3\pi\beta)] \quad (24)$$

for the configuration S_{11} ($\phi_x = 1$, $\phi_z = 1$).

The angular dependences $\lambda(\theta)$ of the dipole interaction parameter calculated using formula (13) using the exact formulas (14) and approximate relations (17)–(24) (ferromagnetic and antiferromagnetic AF(10), AF(01), AF(11) magnetic configurations, respectively) are presented in Figs. 2a and 2b. For moderately deformed rectangular lattices ($0.5 < \beta < 2$), relations (17)–(24) provide a result that practically does not dif-

[3] Relation (13) can be written in the symmetric form $\lambda = \lambda_0\cos^2\theta + \lambda_{\pi/2}\sin^2\theta$, where $\lambda_0 = 2(\mu\beta/l_x)(2\Sigma_m - \beta^2\Sigma_n)$ and $\lambda_{\pi/2} = 2(\mu\beta/l_x)(2\beta^2\Sigma_n - \Sigma_m)$ are the values of λ corresponding to angles $\theta = 0$ and $\theta = \pi/2$, respectively.

fer from the exact result; however, the error increases with the strain (the error becomes equal to 4% for $\beta = 2.5$ or $\beta = 0.4$).

The evaluation and comparison of the values of the dipole interaction parameter λ for different magnetic configurations makes it possible to construct the magnetic phase diagram of the system (see above). Figure 3 shows such a diagram for a rectangular lattice of pointlike (but anisotropic) dipoles. It can be seen that for a slight deviation of the direction of such dipoles from the sides of the lattice, its ground state is always antiferromagnetic (irrespective of the ratio of the lattice periods). On the contrary, for large angles of inclination of the dipoles, the ground state of the lattice becomes ferromagnetic.

2.3. 3D Lattice of Rodlike Grains

An analysis of a lattice formed by rodlike grains (linear dipoles of a finite length) is interesting since the parameters of the magnetic phase diagram of such a lattice can be determined to a high degree of accuracy on the basis of the well-known simple analytic expression for the energy of interaction of such dipoles and can be used for estimating the errors in a rough model of lattices formed by 3D ellipsoidal grains (see below).

If the center of one of such dipoles is at the origin of the coordinate system and the position of the center of another dipole is determined by the radius vector $\mathbf{R}_{mln}$, and if such dipoles are parallel, the energy of their interaction can be written in the form [5]

$$w_d = \frac{\mu^2}{a^2} \times \left(\frac{2}{|\mathbf{R}_{mln}|} - \frac{1}{|\mathbf{R}_{mln} + 2a\mathbf{e}_0|} - \frac{1}{|\mathbf{R}_{mln} - 2a\mathbf{e}_0|} \right), \tag{25}$$

where μ and $2a$ are the magnetic moment and the length of each dipole, respectively. It follows hence that

$$\lambda = 2\mu\beta_y\beta_z\gamma^2 \sum_{m,l,n}{}' \exp[i\pi(\phi_x m + \phi_y l + \phi_z n)] \left[\frac{2}{(m^2 + \beta_y^2 l^2 + \beta_z^2 n^2)^{1/2}} - \frac{1}{[(m + 2\gamma^{-1}\cos\theta)^2 + (\beta_y l + 2\gamma^{-1}\sin\theta\sin\varphi)^2 + (\beta_z n + 2\gamma^{-1}\sin\theta\cos\varphi)^2]^{1/2}} - \frac{1}{[(m - 2\gamma^{-1}\cos\theta)^2 + (\beta_y l - 2\gamma^{-1}\sin\theta\sin\varphi)^2 + (\beta_z n - 2\gamma^{-1}\sin\theta\cos\varphi)^2]^{1/2}} \right], \tag{26}$$

where $\gamma = l_x/a$.

2.4. 2D Lattice of Rodlike Grains

In this case, the general formula (26) is simplified as follows:

$$\lambda = 2\left(\frac{\mu\beta}{l_x}\right) \sum_{m,n}{}' \exp(i\pi(\phi_x m + \phi_z n)) \left\{ \left[\frac{1}{(m^2 + \beta^2 n^2)^{1/2}} - \frac{1}{[(m + 2\gamma^{-1}\cos\theta)^2 + (\beta n + 2\gamma^{-1}\sin\theta)^2]^{1/2}} \right] + \left[\frac{1}{(m^2 + \beta^2 n^2)^{1/2}} - \frac{1}{[(m - 2\gamma^{-1}\cos\theta)^2 + (\beta n - 2\gamma^{-1}\sin\theta)^2]^{1/2}} \right] \right\}. \tag{27}$$

2.5. Ellipsoidal Grains

The method of approximate calculation of the energy W_d of intergrain dipole interaction for a lattice of 3D ellipsoidal grains was proposed in [11]. This method is applicable for strongly prolate or strongly oblate uniformly magnetized ellipsoids of revolution and makes it possible to represent the expression for the energy w_{mln} of their pair interaction in the form of a series in the derivatives of potential ψ of the magnetic field created by such a grain.

Let the center of one of the grains be at the origin and the center of another grain be at the point with coordinates x_0, y_0, z_0. If the magnetic moments $\boldsymbol{\mu}$ of the grains are directed along the x axis, $w_{mln} = -\mu\langle H\rangle_{mln}$, where

$$\langle H\rangle_{mln} = -\frac{1}{V}\iiint \left[\frac{\partial\psi(x, y, z)}{\partial x}\right] dx\,dy\,dz = -\frac{1}{V}\iiint \left[\frac{\partial\psi(x', y', z')}{\partial x'}\right] dx'\,dy'\,dz', \tag{28}$$

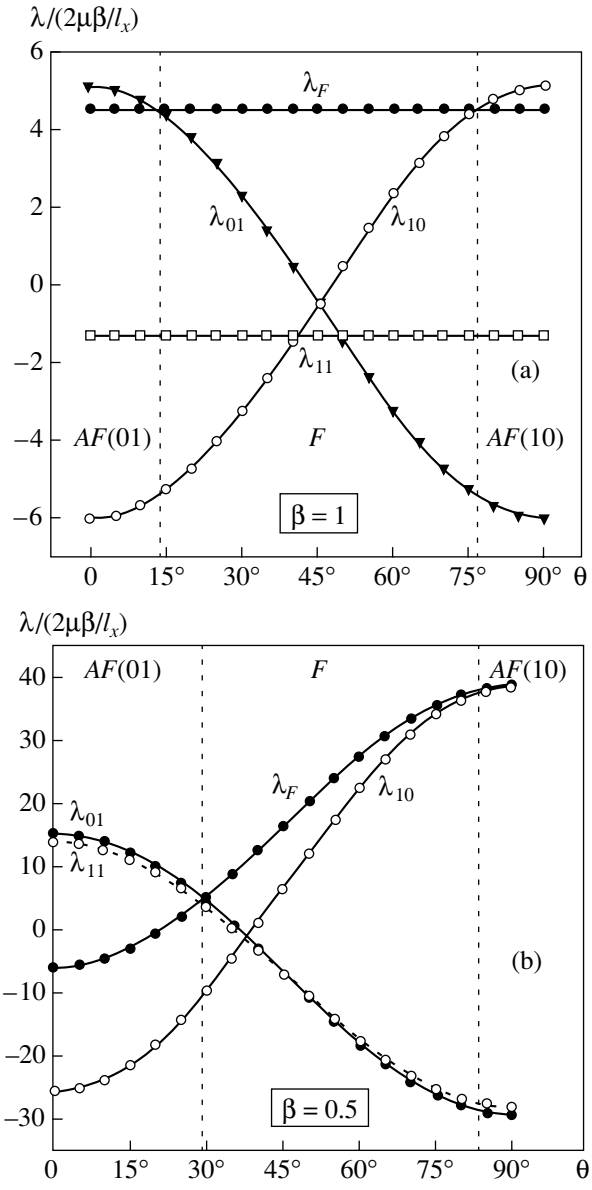


Fig. 2. Angular dependences $\lambda_F(\theta)$, $\lambda_{10}(\theta)$, $\lambda_{01}(\theta)$, and $\lambda_{11}(\theta)$ of the dipole interaction parameters for a (a) quadratic and (b) rectangular lattice of pointlike dipoles. Solid curves are calculated using exact formulas (14), and symbols correspond to approximate relations (17)–(24). The domains of ferromagnetic (F) and antiferromagnetic (AF) states are indicated.

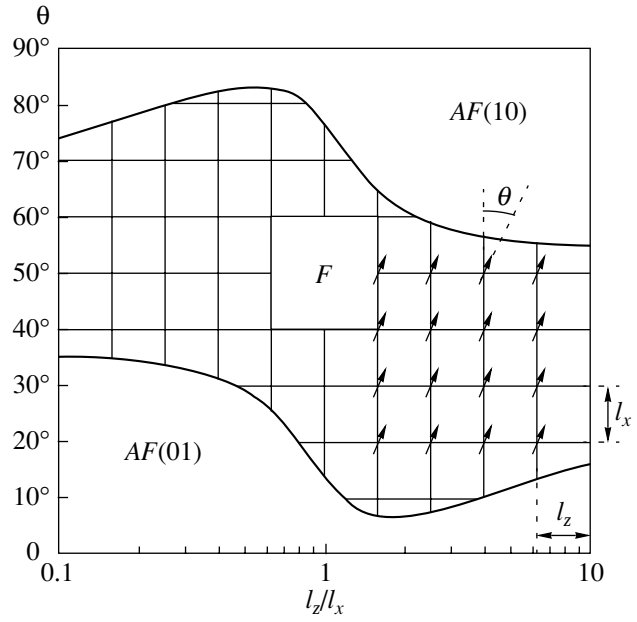


Fig. 3. Magnetic phase diagram of a rectangular lattice of pointlike dipoles parallel to one another. The ground state of the system corresponds to the antiferromagnetic configuration S_{01} for small values of the angle of inclination θ and to the antiferromagnetic configuration S_{10} for large angles. The hatched region corresponds to the ferromagnetic state.

V is the grain volume, and integration is carried out over the volume of the second grain. The coordinates $x' = x - x_0$, $y' = y - y_0$, $z' = z - z_0$ correspond to the coordinate system obtained by parallel translation of the initial system with the origin at the center of the second grain ($x_0 = ml_x$, $y_0 = ll_y \equiv \beta_y l_x$, $z_0 = nl_z \equiv \beta n l_x$).

In view of the geometrical anisotropy, the magnetic moments of the grains are always directed along their major axis, which will be assumed to be parallel to the x axis (it is the only major axis for a prolate ellipsoid of revolution and any of the major axes for an oblate ellipsoid). In this case, integral (28) can be written in the form

$$\iiint\left[\frac{\partial\psi(x', y', z')}{\partial x'}\right]dx'dy'dz' = \int_{-a}^{a} S(x')dx',$$
$$S(x') = \iint_{C_r}\left[\frac{\partial\psi(x', y', z')}{\partial x'}\right]dy'dz', \tag{29}$$

where the integration domain C_r is a circle with the center at the point (x', 0, 0) and radius $r = r(x') = b(1 - x'^2/a^2)^{1/2}$ for a prolate ellipsoid and $r = r(x') = a(1 - x'^2/a^2)^{1/2}$ for an oblate ellipsoid; a, b, and c are the ellipsoid semiaxes ($a > b = c$ for a prolate ellipsoid and $a = b > c$ for an oblate ellipsoid whose minor axis is directed along the z axis).

Expanding potential $\psi(x', y', z')$ into the Taylor series

$$\frac{\partial\psi(x', y', z')}{\partial x'} = \psi'_{0xx} + [\psi''_{0xx}x' + \psi''_{0xy}y' + \psi''_{0xz}z'] + \frac{1}{2}[\psi'''_{0xxx}x'^2 + \psi'''_{0xyy}y'^2 + \psi'''_{0xzz}z'^2] + [\psi'''_{0xxy}x'y' + \psi'''_{0xyz}y'z' + \psi'''_{0xxz}x'z'] + \ldots, \tag{30}$$

where

$$\psi'_{0x} = \partial\psi/\partial x'|_{x'=y'=z'=0},\ \psi''_{0xx} = \partial^2\psi/\partial x'^2|_{x'=y'=z'=0},$$
$$\text{and } \psi'''_{0xxx} = \partial^3\psi/\partial x'^3|_{x'=y'=z'=0}$$

(similar notation is used for the derivatives y' and z'), and substituting this expansion into relation (29), we can see that only those terms in expansion (28), which do not contain terms with odd powers of x', y', z' survive after integration. For a prolate ellipsoid, taking into account relation $\psi''_{0yy} + \psi''_{0zz} = -\psi''_{0xx}$, we finally obtain

$$\langle H_x\rangle_{\text{mln}} = -\left[\frac{3}{2}\left(\frac{\psi_0(a,0,0)-\psi_0(-a,0,0)}{2a}\right)-\frac{1}{2}\psi'_{0x}\right] + \frac{3}{20}a^2\left(1+\frac{2b^2}{3a^2}\right)\psi'''_{0xxx}+\left(\frac{b}{a}\right)^4 O(\psi^{\text{V}}), \quad (31)$$

where $O(\psi^{\text{V}})$ is the sum of terms proportional to the fifth-order derivatives of the magnetic potential.

Similarly, we find that the "central" oblate ellipsoid with the equatorial plane xy creates in this plane a magnetic field $\langle H\rangle_{\text{mln}}$ whose components are defined by the relations

$$\langle H_x\rangle_{\text{mln}} = -\left[\frac{3}{2}\left(\frac{\psi_0(a,0,0)-\psi_0(-a,0,0)}{2a}\right)-\frac{1}{2}\psi'_{0x}\right] + \frac{3}{20}a^2\left[\left(\psi'''_{0xxx}-\frac{2}{3}\psi'''_{0xyy}\right)+\frac{2c^2}{3a^2}(\psi'''_{0xxx}+\psi'''_{0xyy})\right] + O(\psi^{\text{V}}), \quad (32)$$

$$\langle H_y\rangle_{\text{mln}} = -\left[\frac{3}{2}\left(\frac{\psi_0(0,a,0)-\psi_0(0,-a,0)}{2a}\right)-\frac{1}{2}\psi'_{0y}\right] + \frac{3}{20}a^2\left[\left(\psi'''_{0yyy}-\frac{2}{3}\psi'''_{0xxy}\right)+\frac{2c^2}{3a^2}(\psi'''_{0yyy}+\psi'''_{0yxx})\right] + O(\psi^{\text{V}}). \quad (33)$$

Numerical calculations show that the contribution of the term $O(\psi^{\text{V}})$ is negligibly small (less than 1%) in all cases of practical importance. Thus, the application of expressions (31)–(33) taking into account explicitly written terms ensures an accuracy not worse than 1% in the calculation of magnetic energy.

As regards the potential of the magnetic field created by the central grain in the volume of the second grain, the corresponding relations are well known [12]: for a prolate ellipsoid, we have

$$\psi = \left(\frac{3\mu x}{e^3a^3}\right)(\operatorname{arctanh} t - t),\quad t = \frac{e}{\sqrt{1+\xi}}, \quad (34)$$

where $e = (1 - b^2/a^2)^{1/2}$ is the eccentricity and ξ is the larger root of the equation $[(y^2 + z^2)/a^2]/(1 - e^2 + \xi) + (x/a)^2/(1 + \xi) = 1$, while, for an oblate ellipsoid, we have

$$\psi = \left(\frac{3\mu x}{e^3c^3}\right)\left[\arctan t - \frac{t}{1+t^2}\right],\quad t = \frac{e}{\sqrt{1+\xi}}, \quad (35)$$

where $e = [(a^2/c^2 - 1)^{1/2}]$ and ξ is the larger root of the equation $[(x^2 + y^2)/c^2]/(1 + e^2 + \xi) + (z/c)^2/(1 + \xi) = 1$.

2.6. 3D Lattice of Ellipsoidal Grains

Relations (31)–(33), which determine the mean magnetic field $\langle H\rangle_{\text{mln}}$ of the central grain, are written in the coordinate system in which the x axis coincides with the direction of magnetic moments ($\theta = 0$). Using a linear coordinate transformation, we can easily generalize these equations to the case when the direction of these moments is characterized by angles $\theta \neq 0$ and $\phi \neq 0$. However, for a 3D lattice, we confine our analysis to a simpler situation with $\theta = 0$. For a tetragonal lattice of ellipsoidal grains with magnetic moments parallel to the lattice side directed along the x axis, parameter λ determining the type of magnetic order is given by

$$\lambda = 2\mu\beta_y\beta_z\sum_{m,l,n}{}'\left[\frac{\langle H_x\rangle_{mln}}{H_0}\right]\exp(i(\phi_x m+\phi_y l+\phi_z n)), \quad (36)$$
$$H_0 = \frac{\mu}{l_x^3}.$$

2.7. 2D Lattice of Prolate Ellipsoidal Grains

For prolate ellipsoidal grains whose major axes lie in the plane of a 2D rectangular lattice (xz plane), in the general case ($\theta \neq 0$), we obtain

$$\lambda = 2\left(\frac{\mu\beta}{l_x}\right)\sum_{m,n}{}'\left[\frac{\langle\tilde{H}\rangle_{mn}}{H_0}\right]\exp(i(\phi_x m+\phi_z n)), \quad (37)$$

where the mean field $\langle\tilde{H}\rangle_{mn}$ should be calculated using formulas (31) in the coordinate system $\tilde{x} = x\cos\theta + z\sin\theta$, $\tilde{z} = z\cos\theta - x\sin\theta$.

This expression for the dipole interaction parameter λ was used for determining the limits of applicability of our approximate model for prolate ellipsoidal grains. A comparison of the results corresponding to a rectangular lattice and obtained on the basis of the exact formula (27) for linear dipoles and the approximate relation (37) for similar strongly prolate ellipsoids with the eccentricity $e = 0.9999$ (the ratios of the axes $a/b \approx 70$) is illustrated in Fig. 4. It can be seen that the approximate model correctly reflects all features of the complex phase diagram of the system and leads to qualitatively correct results in the entire range of parameters except in the situation close to contact of grains occurring due to their finite transverse dimensions.

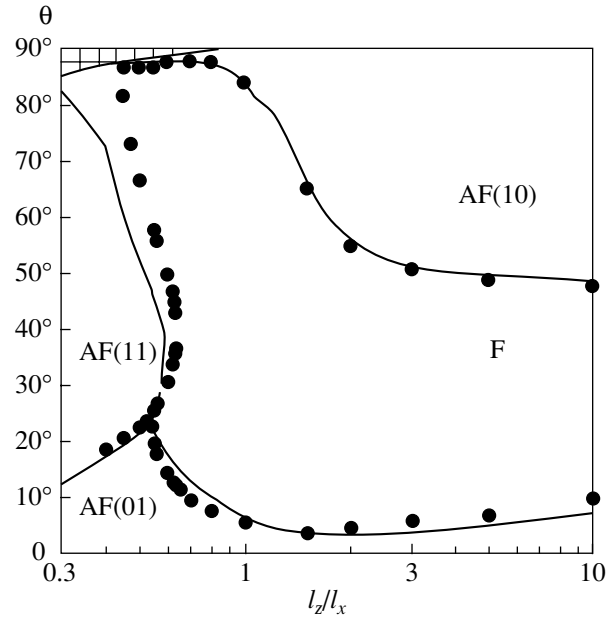


Fig. 4. Comparison of magnetic phase diagrams of a rectangular lattice of linear dipoles (solid curves) and strongly prolate (e = 0.9999) ellipsoids (points) of the same length $2a = (2/3)l_x$. The hatched region corresponds to contacting ellipsoidal grains.

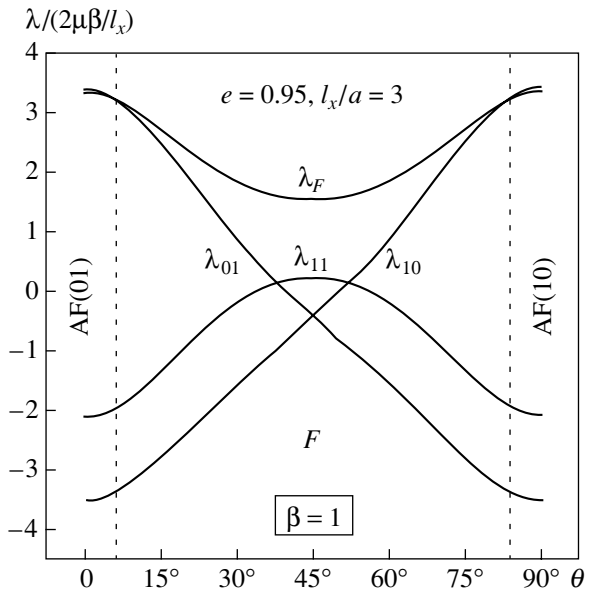


Fig. 5. Dependences of the dipole interaction parameters for a square lattice of ellipsoidal grains on the angle of inclination of their magnetic moments (cf. Fig. 2a). Ellipsoidal grains: e = 0.95, l_x/a = 3.

As before, the evaluation and comparison of the values of the dipole interaction parameter λ for different magnetic configurations of magnetic moments of grains makes it possible to construct the magnetic phase diagram of the system. The dependences of these parameters for a square lattice of ellipsoidal grains on the angle of inclination of their magnetic moments are presented in Fig. 5. Examples of magnetic phase diagrams for two lattices of ellipsoidal grains with different scaling ratios l_x/a are shown in Figs. 6a and 6b. For a small value of this ratio, there exist lattices whose ground state is ferromagnetic even for θ = 0 (e.g., all lattices with β < 1 are of this type for l_x/a = 2.5). In addition, a comparison of Figs. 6a and 6b shows that, upon a decrease in this ratio, the antiferromagnetic phase vanishes in the range of small angles θ, indicating the possibility for a transition of the system from the antiferromagnetic to the ferromagnetic state. It will be shown below that magnetic phase transitions in thin films of magnetic metals may be associated precisely with this feature of the phase diagrams of lattices formed by ellipsoidal grains.

Unfortunately, the approximate nature of the model does not permit the exact determination of individual boundaries on the phase diagrams in all cases. For example, in the cases illustrated in Figs. 6a and 6b, the difference in the energies W_F and W_{01} of the corresponding states in the range of parameters θ < 10° and β > 1 amounts to less than 0.1%. Consequently, it is impossible to determine the actual type of magnetic ordering in this region in the framework of our model. We can only state that the boundary between these phases probably lies in the hatched rectangular regions.

2.8. 2D Lattice with Oblate Ellipsoidal Grains

Let us suppose that the equatorial planes of oblate ellipsoidal grains coincide with a plane of a 2D rectangular lattice (*xy* plane). In the absence of crystal anisotropy, their magnetic moments also lie in this plane, although they are not "attached" to a certain direction. It would be unjustified from the physical point of view to assume that the ground magnetic state in this case corresponds to one of the collinear antiferromagnetic configurations, say, of the type S_{01} studied in the previous analysis, in which the magnetic moments of the grains are parallel to one another. Among candidates for the ground state, we can consider, for example, the fan configuration of magnetic moments, where the latter are turned through angles ±χ relative to one of the lattice sides (see the inset to Fig. 7).

Such states for a 2D square lattice of pointlike magnetic dipoles were studied in [1], where it was proved that the energies of all fan configurations are identical irrespective of angle χ. It was found, however, that such a degeneracy is typical only of pointlike dipoles and is removed as we pass to oblate ellipsoidal grains. In the latter case, the energy of a fan configuration can be calculated by the formula

$$\lambda(\chi) = 2\left(\frac{\mu\beta}{l_x}\right)\Lambda(\chi),$$

$$\Lambda(\chi) = \sum_{m,l}{}' e^{i\pi l}\left[\frac{\langle H_x\rangle_{ml}}{H_0}\cos(\chi-\chi_{ml}) - \frac{\langle H_y\rangle_{ml}}{H_0}\sin(\chi-\chi_{ml})\right], \qquad (38)$$

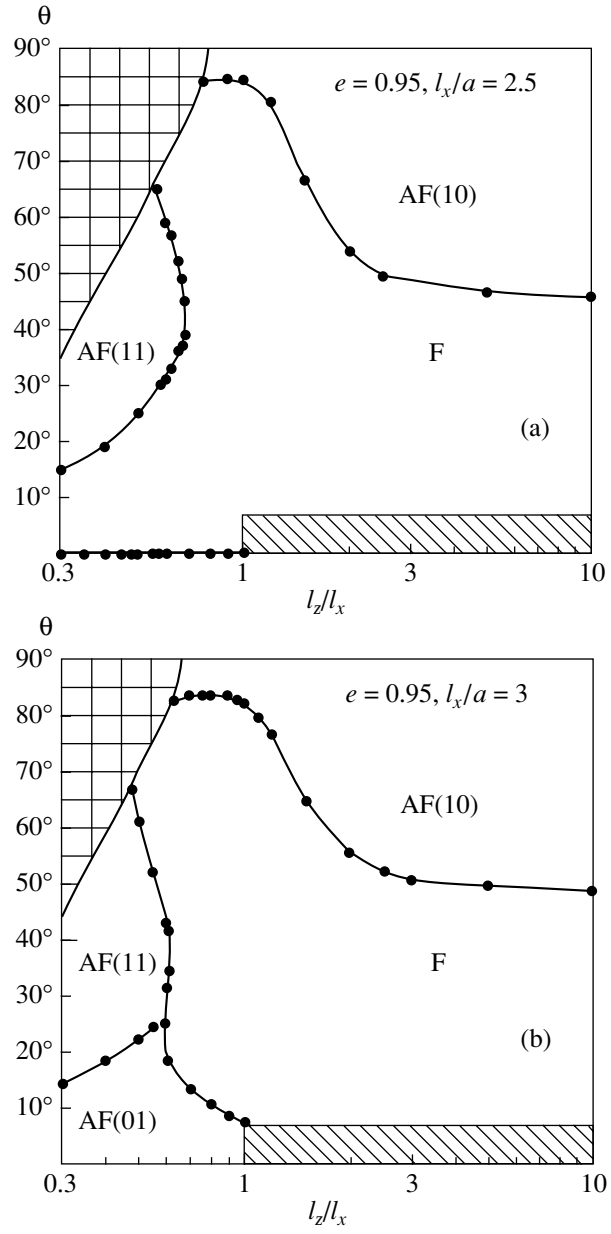


Fig. 6. Magnetic phase diagrams of rectangular lattices of ellipsoidal grains with eccentricity $e = 0.95$ and $l_x/a = 2.5$ (a) and 3 (b). The notation is the same as in Fig. 4. The boundaries of the upper hatched regions forbidden by the geometry of the problem correspond to contacting grains. Phase boundaries for $\theta \ll 10°$, $\beta > 1$ lie within the lower hatched regions.

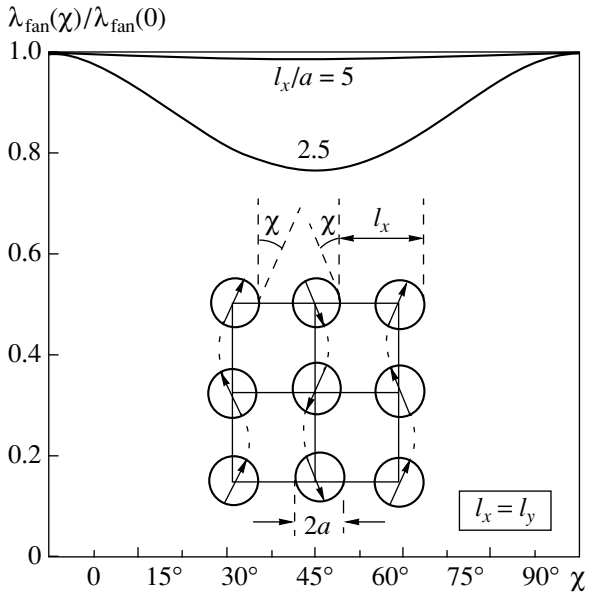


Fig. 7. Dependences of the energy of fan antiferromagnetic configurations for a square lattice of disk-shaped grains on the angle χ of rotation of their magnetic moments (see inset) for various distances between the grains. The inset represents a fan configuration in the form of wavy magnetic field lines.

where $\chi_{ml} = \chi e^{i\pi(m+l)}$; the fields $\langle H_x \rangle_{ml}$ and $\langle H_y \rangle_{ml}$ are defined by relations (32) and (33) in the coordinate system $\tilde{x} = x\cos\chi + y\sin\chi$, $\tilde{y} = y\cos\chi - x\sin\chi$; and the potential of the field created by the central grain is defined by relation (35), in which we must set $\xi = (e^2 + 1)[(x^2 + y^2)/a^2 - 1]$.

Obviously, a noticeable difference from the lattice of pointlike dipoles can be expected only in the case when the shape of grains differs considerably from spherical and their size is comparable with the lattice period ($l \sim a$). Figure 7 shows an example of the dependence of the energy of a fanlike antiferromagnetic configuration on angle χ, which is obtained in this way for very flat ($c/a = 10^{-4}$) and closely spaced grains (disks) in a square lattice. It follows that any vortex configuration in this case is inferior in energy relative to the collinear configuration S_{01}, which represents the ground state. Calculations show that this conclusion remains in force for any rectangular lattice of disk-shaped grains. Similarly, the wavy configurations for which $\chi_{ml} = \chi e^{i\pi l}$ is also less advantageous from the energy point of view (see Fig. 8).

2.9. 2D Lattice of Oblate Ellipsoidal Grains in a Magnetic Field

The energy disadvantage of the ferromagnetic configuration of the ground state can be suppressed by applying an external magnetic field, which inevitably leads to a transition to the ferromagnetic state (if the field is strong enough). Depending on whether this magnetic field is directed along or across the magnetic moments of the sublattices of the initial antiferromagnetic state, such a transition follows different scenarios. In one case, the phase transition occurs through the formation, growth, and coalescence of nuclei of the other (ferromagnetic) phase; this is the process that should be considered in the general theory of kinetics of phase transitions.

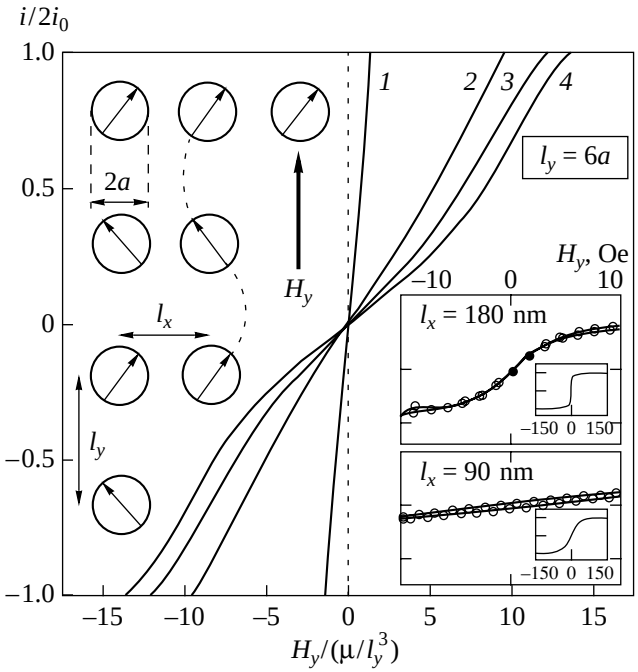


Fig. 8. Field dependences of magnetization for various rectangular lattices of disk-shaped grains of diameter 2*a*. The magnetic field is directed along the longer lattice period $l_y = 6a$. The smaller lattice period l_x is equal to $6a$ (curve *1*), $4a$ (*2*), $3.33a$ (*3*), and $3a$ (*4*). The left inset shows a wavy configuration of the magnetic moments of grains in the field parallel to the longer side of the lattice. The right inset presents the experimental field dependences of magnetization of lattices of flat grains [14].

In another case, the magnetic moment of grains rotate freely and coherently in their plane (under the assumption that there is no magnetic anisotropy in the crystal, as, for example, in the case of grains of a soft magnetic material). The moments of different sublattices have a tendency to align themselves in the direction of the magnetic field by rotating in the opposite directions. For the ground antiferromagnetic state S_{01}, such a transition occurs through a wavy configuration (which is disadvantageous in zero field). When the magnetic moments of the grains rotate through angle χ, the energy of their dipole interaction increases by

$$\Delta W_d(\chi) = -\Delta\lambda(\chi) i_0 \equiv [\lambda(0) - \lambda(\chi)] i_0,$$

where the dipole interaction parameter λ is calculated by formula (38) for $\chi_{ml} = \chi e^{i\pi l}$. This energy increase for each grain is compensated by the decrease in its Zeeman energy,

$$\Delta W_d(\chi) = \mu H \sin\chi,$$

which gives

$$H = \frac{\mu \Delta\Lambda(\chi)}{l_x^3 \sin\chi},$$

where $\Delta\Lambda(\chi) = \Lambda(0) - \Lambda(\chi)$. At the same time, the lattice magnetization is given by

$$i = 2i_0 \sin\chi = \frac{\mu\beta}{l_x^2}\sin\chi.$$

The last two relations determine the parametric relation between the magnetic field H and the equilibrium magnetization i in this field. Such a magnetization could be observed upon an infinitely slow variation of the field in the absence of magnetic anisotropy of any kind. A series of such field dependences for 2D lattices of disk-shaped grains is presented in Fig. 8. As the small lattice period (along which the moments of the grains in the initial antiferromagnetic state are directed) decreases, the magnetization saturation field of the system increases.

3. EXPERIMENTAL EXAMPLES

3.1. Magnetism of Ultrathin Films

Experiments with ultrathin films of iron, cobalt, and nickel revealed that, for a certain effective thickness, a long-range ferromagnetic order is established in such films [8], but the origin and mechanisms of this phenomenon remain not quite clear. The control of the topology of such films and, in particular, an analysis of the geometry and magnetic properties of islets consisting of Co (Fe, Ni) atoms, the fraction of the substrate covered by these atoms, the structure of the film itself, and also the evolution of relevant parameters in the course of the film growth are very important for the development of physical ideas concerning the mechanism of establishment of the magnetic order in such films.

In this connection, we can mention a recent publication [9] in which it is shown that, for a high (room) temperature of monocrystalline (110)-oriented Cu substrate, the Co film deposited on it grows through natural lithography, i.e., through the formation of nuclei (distributed more or less uniformly over the substrate area), which are transformed into ellipsoidal grains with major axes oriented (with a small spread of 5°–10°) along the [001] axes of the substrate. It was found using scanning tunnel microscopy (STM) that the size of these grains increases in the course of film growth, but the distance between their centers and the shape of the grains (i.e., their eccentricity) remain unchanged.

Experiments show that such films become ferromagnetic only if their thickness is large enough. (The effective thickness d of an inhomogeneous film consisting of grains is equal to the thickness of a homogeneous film containing the same number of atoms and is measured by the number of effective monatomic layers.) Gu *et al.* [9] believe that the critical thickness d_c of the film is the thickness for which its hysteresis loop exhibits a nonzero coercivity for the first time. According to their measurements, $d_c \approx 4.6$ monolayers, although noticeable nonlinearity of the magnetic-field dependence of

the film magnetization, which is a consequence of intergrain interaction, appears even for $d \approx 4$ monolayers.

We assume that the described process of transition to the ferromagnetic state can be attributed to the change in the magnetic state of a 2D system of ellipsoidal Co grains with the long-range dipole interaction in accordance with the magnetic phase diagram whose examples were given above. In the framework of this model, the film growth is reduced only to a change in the scaling factor whose role in the problem is played by the size a of the semimajor axis of an ellipsoidal grain.

In order to pass to a quantitative description of the process, we must connect the parameters characterizing the size and shape of the grains (the semiaxis length a and eccentricity e) with the effective thickness d of the film. For this purpose, we can use the experimentally determined (for the same system) relation $\epsilon = \epsilon(d)$ between the fraction ϵ of the open surface of the substrate (which is not occupied by Co grains) and the effective thickness of the film [13]. Obviously, the fraction of the surface covered by grains is $1 - \epsilon = \pi ab/(l_x l_z)$; this leads to the required expression for the parameter l_x/a determining the nature of the magnetic phase diagram:

$$l_x/a = [l_x^{(0)}/a^{(0)}][(1 - \epsilon(d^{(0)}))/(1 - \epsilon(d))],$$

where the parameters labeled by the superscript (0) correspond to the thinnest film (in experiments [9], $d^{(0)} = 0.9$ monolayers). Figure 9 shows such a dependence plotted on the basis of the experimental data [13] for a Co film on a (110) oriented Cu substrate.

According to Gu *et al.* [9], the grain shape is close to ellipsoidal with the eccentricity $e \approx 0.95$, and the ratio of the average distances between the grains (lattice periods) is $\beta = l_z/l_x = 0.5$–1. As regards the angles θ of grain orientation, it follows from the STM images presented in [9] that these angles are distributed in a narrow interval near $\theta = 0$. With increasing effective thickness of the film, the grain size increases, while lattice periods remain practically unchanged (right up to the contact between the grains). In this case, the ratio l_x/a, which determines the magnetic state of the system in accordance with Fig. 6, decreases monotonically, which leads to the magnetic phase transition from the initially antiferromagnetic state to the ferromagnetic state. The critical thickness d_c of the film at which this transition occurs and subsequent dependence of the Curie temperature T_C of the emerging ferromagnetic state are determined (see above) by the value of the dipole interaction parameter λ for various magnetic states of the system.

Proceeding from the STM images of Co films presented in [9], we assumed in our calculations that (i) the axes of all grains are parallel to one another ($\theta = 0$), (ii) $\beta \equiv l_z/l_x = 0.5$, and (iii) $l_x^{(0)}/a^{(0)} = 4$. Figure 10 shows the dependence $T_C(d) = \lambda(d) i_0/k$ of the Curie temperature of the films on their thickness, determined for the above values of the parameters. For the chosen set of parameters, a transition of the system of grains from the antiferromagnetic to the ferromagnetic state occurs at a critical thickness $d_c \approx 4.3$ monolayers, which is close to the value indicated in [9]. After the attainment of such a thickness, the Curie temperature of the formed ferromagnetic state immediately assumes a finite value and then increases with the film thickness. The scale of the initial Curie temperature is determined by the value of μi_0 for $\mu = 300\mu_B$, $i_0 = 10^{-4}$ G cm (which corresponds to the volume of grains of 40 nm^3 and their concentration $N = 10^{-14}$ cm^{-2}) and amounts to $T_C \sim 300$ K.

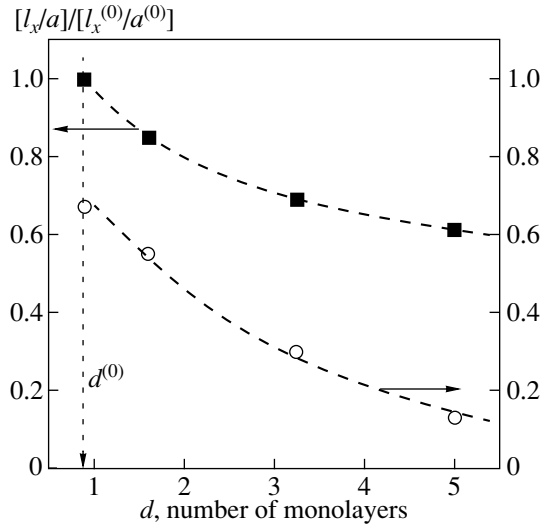


Fig. 9. Experimental dependence $\epsilon = \epsilon(d)$ of the fraction of open (not occupied by Co grains) (100)-oriented surface of the Cu substrate on the effective thickness of the Co film [13] (circles) and corresponding theoretical values of parameter l_x/a (squares). The parameters marked by the superscript (0) correspond to the thinnest film of thickness $d_0 \approx 0.9$ monolayers. The experimental results are approximated by the exponential dependence $\epsilon = \exp(-d/2.56)$ (lower dashed curve), while the results of calculations are approximated by the hyperbolic dependence $[l_x/a]/[l_x^{(0)}/a^{(0)}] = [1 - \epsilon(d^{(0)})]/[1 - \epsilon(d)] = 0.971/d^{0.289}$ (upper dashed curve).

3.2. 2D Lattices of Disk-Shaped Grains in a Magnetic Field

Modern electron lithography makes it possible to create artificial periodic magnetic structures with a size of individual elements (grains) up to tens of nanometers. On the one hand, the sizes of such grains are large enough for their ferromagnetic properties to be virtually identical to the properties of large objects made of the same material, but on the other hand, these are so small that the grains themselves are one-domain ferromagnets. These structures with the dipole interaction

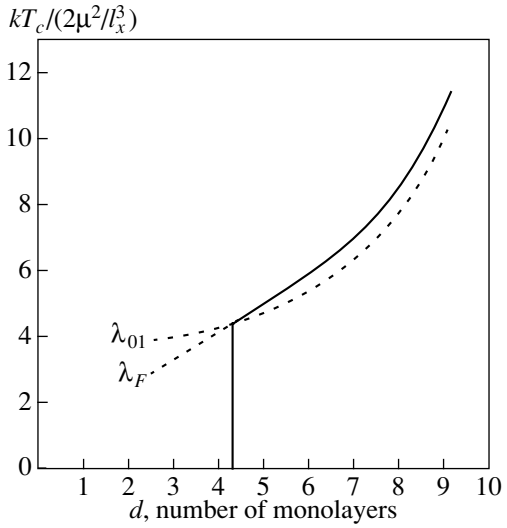


Fig. 10. Dependence of the Curie temperature of the film on its effective thickness. Major axes of grains are parallel to one another ($\theta = 0$), eccentricity $e = 0.95$; grain lattice parameters: $\beta = 0.5$, $l_x^{(0)}/a^{(0)} = 4$.

between grains are just the objects studied by us here. In order to analyze the experimental manifestation of the rearrangement of the magnetic configuration for this type of structures, we consider the results obtained in [14], where 2D rectangular lattices of circular plane grains made of a magnetically soft material (Supermalloy $Ni_{80}Fe_{14}Mo_5$) were studied. The grain size was as follows: diameter $2a = 60$ nm and thickness $h = 7$ nm, which corresponds to the grain magnetic moment $\mu = (\pi a^2 h)I_s = 1.6 \times 10^{-14}$ G cm^3, where $I_s = 800 \pm 60$ G is the saturation magnetization of the grain material. One of the lattice periods ($l_y = 180$ nm) remained unchanged, while the other (l_x) varied in the range 90–180 nm. Since $l_x \le l_y$, the ground state of the system corresponded to the magnetic configuration S_{01}. The magnetic field applied along the y axis must transform the system to the ferromagnetic state via an intermediate wavy magnetic configuration (see above). The characteristic scale of the field required for this purpose is defined by the quantity μ/l_y^3, which is equal to 2.7 Oe in our case. In accordance with the dependences presented in Fig. 8, the saturation magnetization of the system must take place in the fields $H_y = 5.5$ and 35 Oe, respectively, for lattices $l_x = 180$ and 90 nm. The experimentally measured values of these fields [14] $H_y \approx 6$ and 40 Oe are in satisfactory agreement with the results of computations (see the inset to Fig. 8).

3.3. Magnetic Recording Density

The principle of magnetic recording is that the magnetic state (magnitude and direction of magnetization) of a small region of a magnetic medium deposited on the surface of a disk (tape) is memorized. Usually, the medium consists of small grains (of size 100–1000 Å) of a magnetic material (e.g., Fe–Co alloy). A special device known as the head can be positioned over any region of the disk and change the magnetic state of this region (information recording) or determine the state of this region (information readout). The latest advancement in this field was the demonstration of a disk with the recording density of about 5 Gbit/cm^2. A bit of information on the disk is recorded on a region containing approximately 100 grains.

The extremely high density of information storage can obviously be attained by recording one bit on a single magnetic grain. For this purpose, it is necessary to create a special magnetic carrier containing individual periodically arranged magnetic nanoparticles of the same size, shape, and orientation. The periodicity of the arrangement is required for information recording/readout, while the uniformity in the above geometrical parameters is required for storing the recorded information for a long time (of the order of 10 years).

The main mechanism leading to a reversal of the magnetic moment (and, hence, to a loss of information) in small grains is a thermal excitation of such a process. The time of information storage is determined by the characteristic time τ of rotation of the magnetic moment, which is defined as

$$\tau = \tau_0 \exp(\Delta_0/kT), \tag{39}$$

where $\tau_0 = 10^{-10}$–10^{-9} s and Δ_0 is the height of the energy barrier separating two stable states of the grain magnetization. The time τ of information storage exhibits a very strong (exponential) dependence on Δ_0, and in order to obtain $\tau \ge 10$ years, it is necessary to satisfy the condition

$$\Delta_0/kT \ge 40. \tag{40}$$

It is well known [15] that

$$\Delta_0 = \frac{1}{2}\nu V_0 I_s^2, \quad \nu = N_b - N_a, \tag{41}$$

where N_a and N_b are the demagnetization coefficients for a grain along the a and b axes, respectively; V_0 is the grain volume; and I_s is the magnetization of its material.

In our subsequent estimates, we consider Fe grains for which $I_s = 1700$ G and assume that their temperature is $T = 300$ K (room temperature). Substituting relations (41) into (40), we obtain

$$V_0 \ge V_{\min}, \quad V_{\min}[\text{nm}^3] = 1000/\nu, \tag{42}$$

$V_{\min}$ being the minimal volume of Fe grains for which the magnetic moment preserves its direction for 10 years. It should be emphasized that the obtained estimate corresponds to isolated (noninteracting) magnetic grains. Relations (42) imply that, in order to reduce $V_{\min}$, we must take grains with large values of parameter ν, i.e., grains in the form of a strongly prolate

ellipsoid (rodlike grains). It is sufficient to confine our analysis to the ratio of the grain axes $a/b = 4$–10 ensuring the value of $\nu = 5$–6. A further increase in this ratio (i.e., transition to rodlike grains) does not increase parameter ν appreciably. Assuming that $\nu = 5.5$, we obtain the final estimate

$$V_0 \geq V_{\min}, \quad V_{\min} = 180 \text{ nm}^3. \tag{43}$$

The maximum attainable value of information storage density corresponds to close packing of such grains. For the grains of the shape under investigation (ellipsoid of revolution with the axes ratio $a/b = 6$), the recording density is about 1000 Gbit/cm^2.

However, the magnetic interaction of closely spaced grains, which reduces the information storage time like thermal excitation, may obviously reduce the obtained estimate significantly. In order to demonstrate the strength of the latter effect, we use the relation [16]

$$\frac{\Delta_H}{\Delta_0} = (1-\eta)^2, \quad \eta = \frac{H}{\nu I_s} = \frac{HV_0}{\nu\mu}, \tag{44}$$

which describes the lowering of the height Δ_H of the energy barrier separating two stable states of a grain due to the magnetic field H created by all the remaining grains in the region of its location. This relation is based on the assumption that the magnetic moment of the chosen grain and the external magnetic field are antiparallel.

Let the major axes of all grains be parallel to the x axis ($\theta = 0$). Then, in the approximation of rodlike grains, the x component of the resultant field of all grains at the origin (the site of the chosen grain) is determined by a relation similar to (27):

$$H/H_0 = \frac{1}{\gamma^3}\sum_{m,n}{}' e^{i\pi\phi_{mn}}\left[\frac{2}{(m^2+\beta^2n^2)^{1/2}} - \frac{1}{[(m+2\gamma^{-1})^2+\beta^2n^2]^{1/2}} - \frac{1}{[(m-2\gamma^{-1})^2+\beta^2n^2]^{1/2}}\right], \tag{45}$$

where $H_0 = \mu/a^3$, $\phi_{mn} = 1$ if the direction of the magnetic moment of the grain located at point (ml_x, nl_z) coincides with the positive direction of the x axis; otherwise, $\phi_{mn} = -1$ (it should be recalled that $\gamma = l_x/a$, and $\beta = l_z/l_x$). It is clear from geometrical considerations that $l_x > 2a$ and $l_z > 2b$.

Obviously, the magnitude of the field H depends on the sign distribution of magnetic moments of surrounding grains, i.e., on the information recorded in the vicinity of the grain under investigation. While calculating the information storage time, we must proceed from the most unfavorable distribution of these moments (ensuring the maximum possible magnetic field). Such (the worst) configuration can easily be established: if the magnetic moment of the chosen grain (located at the origin) is negative (opposite to the positive direction of the x axis), the signs of the magnetic moments of surrounding grains in such a configuration depend on whether or not they fall into a cone with the angle $2\arccos(1/\sqrt{3})$ and the axis parallel to the x axis. In the former case, the magnetic moment is positive, and in the latter case, it is negative. In particular, the magnetic moments of the grains located on the x and z axes are positive and negative, respectively. The dependences of field H on the geometrical parameters β and γ of the lattice formed by the grains calculated for this case are shown in Fig. 11.

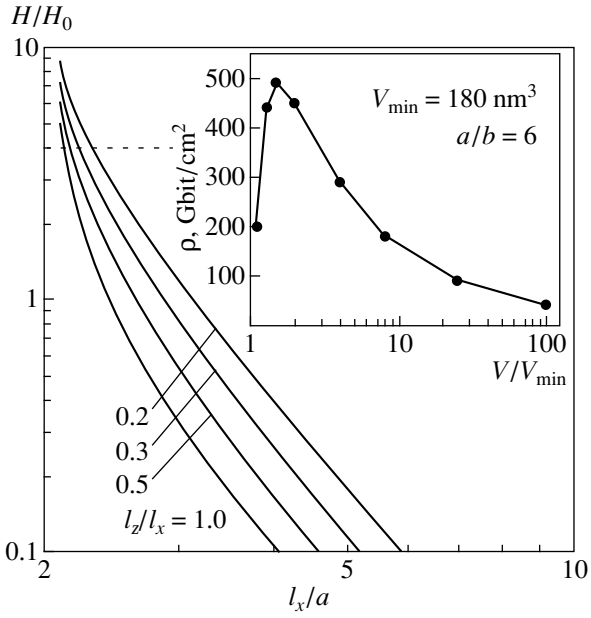


Fig. 11. Dependences of "magnetization reversal" field H on the geometrical parameters $\beta = l_z/l_x$ and $\gamma = l_x/a$ of a lattice of rodlike grains of length $2a$ with magnetic moment μ ($H_0 = \mu/a^3$). The inset shows the dependence of attainable density ρ of magnetic recording on the volume V_0 of Fe grains in the form of ellipsoids of revolution with the axes ratio $a/b = 6$ ($V_{\min} = 180$ nm^3). The dashed line corresponds to condition (46) for $X = 4$.

In order to compensate the effect of magnetic interaction between grains, their volume must be increased as compared to the minimum volume $V_{\min}$. It follows from relations (40)–(42) that the required volume $X = V_0/V_{\min}$ must satisfy the condition $X \geq 1/(1-\eta)^2$, where $\eta = 0.125(H/H_0)$ for $a/b = 6$ and $\nu = 5.5$. Thus, the admissible values of periods l_x and l_z of the lattice formed by grains must satisfy the relation[4]

$$H(l_x, l_z)/H_0 \leq 8\left(1 - \frac{1}{\sqrt{X}}\right). \tag{46}$$

[4] Condition (46) indicates, in particular, that H must be equal to zero for $V = V_{\min}$ ($X = 1$), which corresponds to grains separated by infinitely long distances.

Among these values, we must select those corresponding to the maximum surface density of grains, equal to $1/(l_x l_z)$. The horizontal straight line $H/H_0 = \text{const}$ drawn in Fig. 11 corresponds to a certain value of X. It can be seen that the optimum values of lattice periods are always those corresponding to the point of intersection of this line with the curve $H(l_x, l_z)$ for the minimum ratio of the periods $\beta = l_z/l_x$ dictated by geometrical considerations (in our case, $\beta = b/a \approx 0.2$). The dependence of the maximum attainable density of magnetic recording determined in this way on the volume of grains is shown in the inset to Fig. 11. It can be seen that, although the dipole interaction between grains considerably reduces this density (in our case, to half the rated value), it can still be high enough and attain values of approximately 500 Gbit/cm^2, which is two orders of magnitude higher than the density of recording in the best modern magnetic disks.

The approximation of rodlike dipoles used in the previous analysis is completely applicable only for grains separated by large distances (as compared to the grain size) from the origin (the site of the selected molecule). Obviously, the largest errors appear in this approximation when we calculate the field created by the nearest grains. However, special calculations made by us proved that the resultant action of remote but numerous grains exceeds the action of a few nearest grains (this is a consequence of the long-range nature of the field of a magnetic dipole). Thus, the approximation of rodlike grains correctly describes the effect of magnetic interaction between the grains.

4. CONCLUSIONS

We have described the methods for analyzing the magnetic properties of 2D and 3D lattices of nonspherical ferromagnetic grains with the intergrain dipole interaction: the ground state, the magnetic phase diagram, and the change in the magnetic state under the action of an external magnetic field. The obtained results can be used for describing the properties of a number of real systems, including 2D periodic structures of magnetic grains suitable for creating magnetic memory systems with a high recording density, ultrathin films of ferromagnetic metals on monocrystalline substrates, and rectangular lattices of disk-shaped magnetically soft grains. The methods developed for describing the properties of such systems are in good agreement with the results of relevant experiments.

ACKNOWLEDGMENTS

This study was supported by the Russian Foundation for Basic Research (project nos. 00-02-17191, 01-02-16420, and 02-02-16974) and the PICS-RFBR program (project no. 01-02-22004).

REFERENCES

1. P. I. Belobrov, R. S. Gekht, and V. I. Ignatchenko, Zh. Éksp. Teor. Fiz. **84**, 1097 (1983) [Sov. Phys. JETP **57**, 636 (1983)].
2. J. A. Sauer, Phys. Rev. **57**, 142 (1940).
3. J. M. Luttinger and L. Tisza, Phys. Rev. **70**, 954 (1946).
4. J. D. Jackson, *Classical Electrodynamics* (Wiley, New York, 1999).
5. M. D. Costa and Yu. G. Pogorelov, Phys. Status Solidi A **189**, 923 (2001).
6. R. L. White, J. Magn. Magn. Mater. **209**, 1 (2000).
7. J. S. Moodera and G. Mathon, J. Magn. Magn. Mater. **209**, 248 (2000).
8. W. Dürr, M. Taborelli, O. Paul, *et al.*, Phys. Rev. Lett. **62**, 206 (1989); C. M. Schneider, P. Dressler, P. Schuster, *et al.*, Phys. Rev. Lett. **64**, 1059 (1990); J. de la Figuera, J. E. Pietro, C. Okal, and R. Miranda, Phys. Rev. B **47**, 13043 (1993); A. K. Schmid, D. Atlan, H. Itoh, *et al.*, Phys. Rev. B **48**, 2855 (1993); F. Huang, M. T. Kief, G. J. Mankey, and R. F. Willis, Phys. Rev. B **49**, 3962 (1994); F. O. Shumann, M. E. Buckley, and J. A. C. Bland, Phys. Rev. B **50**, 16424 (1994); H. J. Elmers, J. Haushild, H. Höche, *et al.*, Phys. Rev. Lett. **73**, 898 (1994).
9. E. Gu, S. Hope, M. Tselepi, and J. A. C. Bland, Phys. Rev. B **60**, 4092 (1999).
10. H. Benson and D. L. Mills, Phys. Rev. **178**, 839 (1969).
11. E. Z. Meĭlikhov and R. M. Farzetdinova, Zh. Éksp. Teor. Fiz. **121**, 875 (2002) [JETP **94**, 751 (2002)].
12. L. D. Landau and E. M. Lifshitz, *Course of Theoretical Physics*, Vol. 8: *Electrodynamics of Continuous Media* (Fizmatgiz, Moscow, 1959; Pergamon, New York, 1960).
13. M. T. Kief and W. F. Egelhoff, Phys. Rev. B **47**, 10785 (1993).
14. R. P. Cowburn, A. O. Adeyeye, and M. E. Welland, New J. Phys. **1**, 16.1 (1999).
15. *Introduction to Magnetic Recording*, Ed. by R. M. White (IEEE Press, New York, 1985).
16. E. Z. Meĭlikhov, Zh. Éksp. Teor. Fiz. **116**, 2182 (1999) [JETP **89**, 1184 (1999)].

Translated by N. Wadhwa

Journal of Experimental and Theoretical Physics, Vol. 95, No. 5, 2002, pp. 901–911.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 122, No. 5, 2002, pp. 1044–1056.

SOLIDS
Structure

Noncollinear Magnetic States in Iron–Chromium-Type Multilayers

V. N. Men'shov* and V. V. Tugushev

Russian Research Centre Kurchatov Institute, pl. Kurchatova 1, Moscow, 123182 Russia
**e-mail: vvtugushev@mail.ru, sasha@mics.msu.su*

Received May 21, 2002

Abstract—A model of noncollinear magnetic ordering in Fe/Cr-type multistructures was suggested. The model was based on the idea of charge (and, as a consequence, spin) density redistribution near a metal–metal interface. A peculiar state of the whole structure characterized by strong short-range antiferromagnetic ordering in the interlayer and a pronounced dependence of magnetic characteristics on the properties of the boundary between iron and chromium layers was shown to be formed in a certain temperature range. Inhomogeneous antiferromagnetic structures with a vector order parameter were found, and the effective exchange coupling between neighboring iron layer moments was calculated using the Ginzburg–Landau expansion of the thermodynamic potential. The results were used to discuss the experimental data on Fe/Cr superlattices obtained in neutron scattering and magnetization measurements. *© 2002 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

Interest in multilayers (sandwiches and superlattices) of the FM/AFM type, where FM and AFM are ferromagnetic and antiferromagnetic materials, respectively, stems from increasing requirements on the parameters of layered nanostructures used in modern magnetic recording technologies. Such multilayers are as a rule prepared by molecular-beam epitaxy with the use of layer-by-layer deposition techniques. They are unique objects of fundamental studies and have quite a number of unusual transport and magnetic properties (for instance, giant magnetoresistance).

Transition-metal-based multilayers, for instance, multilayers containing iron and cobalt as ferromagnets and chromium and manganese as antiferromagnets, rank among the most popular FM/AFM nanostructures. In particular, at least three recent reviews concentrate on Fe/Cr-type structures, which will be the subject matter of our study [1–3]. Indeed, Fe/Cr multilayers were the first example of magnetic structures with exchange potential $J_1(L)$ between FM layers separated by an AFM interlayer in which the sign of this potential oscillated as a function of interlayer thickness L. The $J_1(L)$ potential has a complex dependence on the geometry of the structure (thickness and growth direction), the topology of the Fermi surface in the chromium interlayer, and the quality of the Fe/Cr interface. The $J_1(L)$ dependence contains "fast" and "slow" components oscillating as functions of L. These components are related to different types of electronic and hole regions of the chromium Fermi surface; we also observe exchange potential periodicity to experience disturbance ("phase slip") caused by a peculiar magnetic order type, namely, by a spin density wave in chromium. Along with parallel and antiparallel orientations of the magnetic moments of neighboring iron layers, orientations with $\psi(L)$ angles (angles between iron moments) that differed from 0 or π were observed in various experiments in the thermodynamically equilibrium state and in the absence of an external magnetic field. Further, we use the terms "collinear state" and "noncollinear state" for orientations with $\psi = (0, \pi)$ and $\psi \neq (0, \pi)$, respectively; these terms have no deeper meaning.

All the existing theoretical schemes (see [3]) give more or less plausible descriptions of the dependence of potential $J_1(L)$ on interlayer thickness in structures with ideally smooth Fe/Cr interfaces and predict the collinear state of such structures to be more favorable than noncollinear. The appearance of noncollinear states is usually related to defects of the Fe/Cr interface, namely, to interface roughness [1, 2]. A consideration of effective exchange in a system with a rough, that is, not ideally smooth, Fe/Cr interface is, however, an exceedingly complex task involving many unknown parameters. The point is that, at the microscopic level, roughness is related to the presence of monoatomic steps, interdiffusion of the components, and other technological deviations from uniformity of real interfaces formed during nanostructure growth. This in turn causes frustrations of exchange coupling between Fe and Cr and the formation of magnetic domain walls on both sides of the Fe/Cr surface. These deviations from ideal surface smoothness were found to initiate very serious changes in the macroscopic magnetic properties of multilayers; in particular, they could cause the formation of noncollinear states [4–6].

1063-7761/02/9505-0901$22.00

Several interesting phenomenological models have been suggested to describe this effect. The most effective one is the biquadratic exchange scheme [7], which is, so far as we know, closest to the correct interpretation of the experimental data on Fe/Cr multilayers. This scheme directly relates the formation of noncollinear ordering of iron moments to the presence of alternating steps and terraces on Fe/Cr interfaces. On the one hand, these steps and terraces substantially suppress the short-wave component of usual (bilinear) exchange, that is, the $J_1(L)$ potential, and, on the other, they form a peculiar additional (biquadratic) exchange potential $J_2(L)$ caused by partial disturbance of strong coupling between Fe and Cr spins near monosteps (see the discussion in review [3]).

No microscopic theory underlying the biquadratic exchange scheme has been suggested, first and foremost, because of the absence of a correct description of the complex influence of the Fe/Cr interface on antiferromagnetic ordering with a spin density wave in the chromium interlayer, through which iron layers interact with each other. Non-self-consistent approaches of type [8] are incapable of even qualitatively explaining the magnetic phase diagram of Fe/Cr multilayers, not to mention fine details of spin density distributions within the interlayer. Using these approaches to solve the delicate problem of the influence of interface imperfections on favorableness or unfavorableness of collinear compared with noncollinear ordering is therefore meaningless.

Recently, we have suggested a model for self-consistently describing spin density distributions in Fe/Cr multilayers [9, 10]. This model has been used to analyze collinear states and has allowed us to explain the key features of the phase diagram of Fe/Cr multilayers in the temperature–interlayer thickness variables in terms of short-range order and local phase transitions into the state with a spin density wave. In this work, we generalize this model to noncollinear states. The microscopic reason why the noncollinear rather than collinear state is energetically favorable in the presence of monosteps and terraces on the Fe/Cr interface will be elucidated; this reason is related to the rearrangement of the spin density wave in the chromium surface layer and the formation of a peculiar domain wall (rotation of the spin density wave phase) near a monostep on the Fe/Cr interface. The biquadratic exchange model will be substantiated at the microscopic level, and limitations imposed on its applicability by interlayer thickness, temperature, and interface roughness will be analyzed.

2. MODEL AND PROBLEM STATEMENT

As in [9, 10], we consider a model of a three-layer nanostructure comprising two ferromagnetic plates (Fe) separated by an antiferromagnetic layer (Cr). Let the technological Fe/Cr interface lie in the ($\mathbf{n}_y$, $\mathbf{n}_z$) plane, which is orthogonal to the direction of structure growth $\mathbf{n}_x$ along one of the [100] cubic axes (here and throughout, $\mathbf{n}_x$, $\mathbf{n}_y$, and $\mathbf{n}_z$ are basis unit vectors). Consider the temperature T range corresponding to short-range antiferromagnetic order in the chromium layer; that is, $T_N < T \ll T_C$, where T_N is the Néel temperature in the chromium layer and T_C is the Curie temperature in iron plates. Ferromagnetic layers are assumed to be sufficiently thick for magnetization $\mathbf{S}$ within ferromagnetic plates at $T \ll T_C$ to be considered uniform and independent of T. At the same time, antiferromagnetic layer thickness L changes within fairly wide limits ($L > 2\xi_0$, where ξ_0 is the coherence length of the order of ten chromium monolayers; in the temperature range that we are considering, sublattice magnetization $\boldsymbol{\sigma}$ within the antiferromagnetic layer can be fairly nonuniform and can strongly depend on T).

Layer dimensions in the $\mathbf{n}_y$ and $\mathbf{n}_z$ directions are assumed to be much larger than thickness L. We can then use the simplified one-dimensional model in which $\boldsymbol{\sigma}(x)$ depends on a single spatial coordinate for ideally smooth interfaces. In the presence of interface extensive defects (such as monosteps separating planar terraces), $\boldsymbol{\sigma}$ also depends on the y and z coordinates, and the one-dimensional model can then be used to approximately describe a nanostructure fragment situated far from a monostep. A stricter criterion of the applicability of the one-dimensional model and the role played by $\boldsymbol{\sigma}(x, y, z)$ distribution nonuniformity within the interlayer will be discussed in Section 5. For now, this model will be used as a basis for the further analysis. We restrict our consideration to transversely polarized spin density waves, when $\boldsymbol{\sigma}(x) \perp \mathbf{n}_x$ (precisely this situation is discussed most frequently, but a detailed analysis of the reasons for that would lead us far beyond the scope of this work).

Let us introduce order parameter $\boldsymbol{\Delta}(x) = U\boldsymbol{\sigma}(x)$, where U is the effective potential, whose explicit form is not discussed here (e.g., see review [11]). This parameter has the dimension of energy and describes the envelope of the spin density wave. For a transversely polarized one-dimensional spin density wave, we have

$$\boldsymbol{\Delta}(x) = \mathbf{n}_y\Delta_y(x) + \mathbf{n}_z\Delta_z(x), \quad \boldsymbol{\Delta}(x) \perp \mathbf{n}_x, \tag{1}$$

where $|x| \le l$ (l is the interlayer half-width, $L = 2l$). We assume that $\boldsymbol{\Delta}(x)$ is small ($|\boldsymbol{\Delta}| \ll \pi T$) and slowly varies in space ($|\partial\boldsymbol{\Delta}/\partial x| \ll \pi T/\xi_0$). On this assumption, let us write the thermodynamic potential $F[\boldsymbol{\Delta}]$ of the antiferromagnetic layer under consideration in the form of the Ginzburg–Landau expansion in powers of $\boldsymbol{\Delta}(x)$ and $\partial\boldsymbol{\Delta}(x)/\partial x$. A detailed justification of this approach to the problem can be found in [9, 10]. We therefore give the expression for $F[\boldsymbol{\Delta}]$ without much comment:

$$F = F_V + F_S, \tag{2}$$

$$F_V = \frac{1}{2}\int_{-l}^{l} f_V dx, \tag{3}$$

$$f_V = c_1\mathbf{\Delta}^2 + c_2 v_F^2\mathbf{\Delta}'^2 + c_2\mathbf{\Delta}^4, \tag{4}$$

$$F_S = \frac{\nu}{4}(\mathbf{\Delta}^2(l) + \mathbf{\Delta}^2(-l)) + \frac{1}{2}(\mathbf{A}(l)\mathbf{\Delta}(l) + (-1)^{N+1}\mathbf{A}(-l)\mathbf{\Delta}(-l)). \tag{5}$$

The F_V and F_S values are the volume and surface contributions to the total thermodynamic potential, respectively. The c_1, c_2, ν, and A coefficients are given in [9, 10] (e.g., see [11] for their calculation); v_F is the Fermi velocity of the electrons of plane chromium Fermi surface regions responsible for the formation of spin density waves; and N is the number of chromium monolayers in the interlayer. Everywhere below, $c_1, c_2 > 0$, $\nu < 0$, and the $\mathbf{A}(\pm l)$ vector is proportional to magnetization $\mathbf{S}(\pm l)$ in the corresponding ferromagnetic plate.

Expansion (3), (4) is strictly valid almost in the whole $|x| < l$ region except antiferromagnetic layer regions of size of about ξ_0 near the interface, where local approximation (2) to functional F is incorrect. Fine details of charge and spin redistribution on such scales are of no interest to us; they can be considered to be taken into account by the ν and $\mathbf{A}$ coefficients. The term in (5) linear in $\mathbf{\Delta}$ (the exchange term) is directly related to exchange interaction between ferromagnetic plate and antiferromagnetic layer spins. The term in (5) quadratic in $\mathbf{\Delta}$ (the Coulomb term) is determined by charge flow between layers of different metals and by the resulting appearance of a contact potential drop between ferromagnetic and antiferromagnetic layers. According to [9, 10], the Coulomb term predominates over the exchange term in a wide temperature range, $T > T_N$, and determines the characteristic temperature T_0 of the formation of short-range antiferromagnetic order; a spin density wave is induced by an increase in the electronic polarizability in the interlayer close to the FM/AFM interface. At the same time, the exchange term determines the details of the spatial dependence of the spin density wave and its orientation with respect to magnetization $\mathbf{S}(\pm l)$ in the ferromagnetic plates. The role played by this term can become very important in the immediate vicinity of the T_0 point, when Δx sharply decreases and the spin density wave is largely induced by spin density redistribution in the interlayer under the influence of the exchange field of the ferromagnetic plate. The critical temperatures in the Fe/Cr system are $T_0 \approx 550$–600 K, $T_N \approx 311$ K, and $T_C \approx 1040$ K ($T_N < T_0 < T_C$). We assume that interlayer thickness L is larger than 25–30 Å.

Formally, the problem can be formulated as follows. In [9, 10], thermodynamically equilibrium states of functional (2) were studied only under the condition $\mathbf{A}(l) = \pm\mathbf{A}(-l)$, that is, for the collinear orientation of iron moments. Only one $\mathbf{\Delta}(x)$ vector function component (Δ_y or Δ_z) was assumed to be nonzero, and the energetically optimal configuration of the spin density wave at a given (parallel or antiparallel) mutual orientation of $\mathbf{A}(\pm l)$ vectors and at a fixed number of monolayers N (even or odd) in the interlayer was calculated. Below, we analyze a more general (and, as it turns out, much more complex for calculations) situation when the $\mathbf{A}(\pm l)$ vectors have mutual orientations characterized by arbitrary angles ψ ($0 < \psi < \pi$), that is, are noncollinear. At the first stage, the ψ angle and the number of monolayers N play the role of given external parameters, and the optimization of functional $F[\mathbf{\Delta}]$ [Eq. (2)] is performed by conditionally varying it with respect to the two-component ($\Delta_y, \Delta_z \neq 0$) vector order parameter [Eq. (1)]. At the second stage, the obtained $\mathbf{\Delta}(x)$ solution is used to calculate $F(\psi, N)$, and the most energetically favorable configuration for magnetizations $\mathbf{S}(\pm l)$ in iron layers is determined. Lastly, the limits of the applicability of the developed approach are analyzed. It turns out that this approach can conveniently be used to fairly approximately model the influence of interface roughness on the formation of spin density waves if monosteps are spaced well apart, terraces are broad, and the interlayer is not too thin.

3. THE STRUCTURE OF SHORT-RANGE ANTIFERROMAGNETIC ORDER WITH A NONCOLLINEAR SPIN DENSITY WAVE

Varying functional $F[\mathbf{\Delta}]$ [Eq. (2)] with respect to order parameter $\mathbf{\Delta}(x)$ yields the self-consistency equation

$$v_F^2 c_2\mathbf{\Delta}'' - c_1\mathbf{\Delta} - 2c_2\mathbf{\Delta}^3 = 0 \tag{6}$$

with the boundary conditions

$$\mathbf{\Delta}'(\pm l) \mp \frac{1}{D}\mathbf{\Delta}(\pm l) + (\pm 1)^N\frac{\mathbf{A}(\pm l)}{2c_2 v_F^2} = 0. \tag{7}$$

The $D = 2c_2 v_F^2/|\nu|$ parameter (the spatial scale related to charge density redistribution near the interface known as "interpolation length") weakly depends on temperature in the model under consideration. The $\mathbf{A}(\pm l)$ vector is given in the form

$$A_x(\pm l) = 0, \quad A_y(\pm l) = \pm A\sin(\psi/2), \quad A_z(\mp l) = A\cos(\psi/2), \tag{8}$$

where $A > 0$, which corresponds to antiferromagnetic exchange at the Fe/Cr interface, and ψ is the angle between the magnetization directions in the ferromagnetic plates. Earlier, only the situations with $\psi = 0, \pi$ were considered, when only one $\mathbf{\Delta}(x)$ vector component (Δ_y or Δ_z) was nonzero, and it proved possible to obtain exact solutions to problem (6), (7) [9, 10]. These were classified as one symmetrical $\Delta_+(x)$ and two antisym-

metric $\Delta_{-}^{(1,2)}(x)$ solutions with respect to the replacement $x \longrightarrow -x$. Each solution had its own definite existence region in the (T, L) phase diagram.

Consider vector solutions to problem (6), (7) given in form (1). If the number of monolayers N is odd (if N is even, it is easy to see from (5), (7), and (8) that all calculations will be identical except for the substitutions $\psi \longrightarrow \psi - \pi$, $\Delta_z \longrightarrow \Delta_y$, and $\Delta_y \longrightarrow -\Delta_z$, $\mathbf{\Delta}(x)$ will be sought in the quasi-helicoidal form

$$\begin{aligned} \Delta_z(x) &= \Delta(x)\cos\Theta(x), \\ \Delta_y(x) &= \Delta(x)\sin\Theta(x), \end{aligned} \tag{9}$$

where $\Delta(x) > 0$ is the modulus and $\Theta(x)$ is the order parameter phase; this phase takes on the values $\Theta(\pm l) = \pi \pm \varphi/2$ at interlayer boundaries in the selected reference system for counting angles. The $\varphi = \Theta(+l) - \Theta(-l)$ angle characterizes antiferromagnetic structure sloping in the interlayer ($0 < \varphi < \pi$). We therefore have $\Theta(\pm l) = \pi \pm \psi/2$ at $\varphi = \psi$, which corresponds to the strictly antiparallel mutual orientation of the $\boldsymbol{\sigma}$ and $\mathbf{S}$ vectors at the corresponding Fe/Cr interfaces and to a minimum of the F_S surface contribution to thermodynamic potential at a fixed $\Delta(\pm l)$ amplitude. The problem of the ratio between the φ and ψ angles optimal from the point of view of the total energy F of the system has no obvious solution. The point is that changes in $\Delta(x)$ and $\Theta(x)$ are stringently related to each other. This relation is determined by a system of two nonlinear equations obtained by substituting (9) into (6),

$$v_F \frac{d\Theta}{dx} = \frac{B}{\Delta^2}, \tag{10}$$

$$\left(v_F \Delta \frac{d\Delta}{dx}\right)^2 = \Delta^6 + \frac{c_1}{c_2}\Delta^4 + C\Delta^2 - B^2, \tag{11}$$

where C and B are real constants of integration, which determine the type of the solution. The equality $B = 0$ corresponds to a scalar antiferromagnetic structure with $\Theta(x) = \text{const}$, and the $\Delta_{+}(x)$ and $\Delta_{-}^{(1,2)}(x)$ functions mentioned above then satisfy (11) (see [9, 10]). If $B \neq 0$, which corresponds to a vector antiferromagnetic structure, (11) with $c_1 > 0$ has two different solutions, which are bounded in magnitude in the $|x| \leq l$ segment. These solutions can be written in terms of the Jacobi elliptic functions [12]. An analysis shows that, in the temperature range of interest to us, only one solution exists. This solution is characterized by the following spatial amplitude and phase distributions:

$$\Delta(x) = \frac{v_F k'}{\xi}\left[\left(\frac{\lambda\xi}{l}\right)^2 \mathrm{nc}^2\left(\frac{\lambda x}{l}, k\right) - \beta\right]^{1/2}, \tag{12}$$

$$\lambda = \frac{1}{\xi}\left[\frac{1 - 3\beta k'^2}{1 - 2k'^2}\right]^{1/2},$$

$$\Theta(x) = \frac{B}{v_F}\int_0^x \frac{dx}{\Delta^2(x)} + \pi,$$

$$B^2 = \left(\frac{v_F k'}{\xi}\right)^6 \frac{(1 + k^2)(2 - k^2)}{(2k^2 - 1)^2}\beta(\beta - \beta_+)(\beta - \beta_-), \tag{13}$$

$$\beta_+ = \frac{1}{2 - k^2}, \quad \beta_- = \frac{k^2}{k'^2}\frac{1}{1 + k^2}.$$

Here, k is the modulus of the elliptic function and $k' = \sqrt{1 - k^2}$ is an additional modulus, $0 \leq k \leq 1$. The region of allowed parameter β values is limited by the conditions $\beta \geq \beta_+$ for $k^2 \leq 1/2$ or $\beta_+ \geq \beta \geq 0$ for $k^2 \geq 1/2$. These conditions are related to the characteristic lengths of the system and to the $\psi - \varphi$ difference angle by boundary conditions (7). It follows from (7) that $B \propto A\sin[(\psi - \varphi)/2]$; that is, only the state with a linearly polarized spin density wave is possible in the absence of exchange coupling at the interface ($A = 0$). If $A \neq 0$, the formation of a state with a vector order parameter of type (9) should necessarily be accompanied by the frustration of exchange couplings at the interfaces; in other words, $\psi - \varphi \neq 0$.

Constraints (7) for solution (12), (13) can be written in the form

$$P = \frac{B}{v_F}\cot\left(\frac{\psi - \varphi}{2}\right), \tag{14}$$

$$P^2 + \left(\frac{B}{v_F}\right)^2 = \left(\frac{A\lambda k'}{2c_2 v_F l}\right)^2\left[\mathrm{nc}^2(\lambda, k) - \beta\left(\frac{1}{\lambda\xi}\right)^2\right], \tag{15}$$

where

$$P = \left(\frac{v_F \lambda k'}{l}\right)^2\left(\frac{\lambda}{l}\frac{\mathrm{sn}(\lambda, k)\mathrm{dn}(\lambda, k)}{\mathrm{cn}^3(\lambda, k)} - \frac{1}{D}\left[\mathrm{nc}^2(\lambda, k) - \beta\left(\frac{1}{\lambda\xi}\right)^2\right]\right). \tag{16}$$

These equations together with (13) determine the parameters (k, β, and φ) of the $\mathbf{\Delta}(x)$ vector structure at $x = \pm l$. Unfortunately, (14)–(16) can only be analyzed numerically in the whole range of parameter values. In several important limiting cases, such an analysis can, however, be performed analytically. For instance, near the $T_0(l)$ temperature, which determines the upper boundary of the existence of short-range order with a spin density wave induced by a charge flow, we have $k'^2 \ll 1$ [9, 10].

Assuming k' to be small to the extent that $(k'\exp(l/\xi))^2 \ll 1$, we can approximate the elliptic func-

tions in (12) and (13) by hyperbolic ones [12]. This substantially simplifies (12)–(15), and we can write

$$\Delta(x) = \frac{v_F k'}{\xi}\left[\cosh^2\left(\frac{x}{\xi}\right) - \beta\right]^{1/2}, \quad (17)$$

$$\tan\Theta(x) = \sqrt{\frac{\beta}{1-\beta}}\tanh\left(\frac{x}{\xi}\right). \quad (18)$$

Setting $x = \pm l$ in (18), we obtain the following relation between β and angle φ:

$$\frac{1}{\beta} = 1 + \tanh^2\left(\frac{l}{\xi}\right)\cot^2\left(\frac{\varphi}{2}\right). \quad (19)$$

In the $k'^2 \ll 1$ limit, (14) determines the relation between the φ and ψ angles,

$$\tan\left(\frac{\varphi}{2}\right) = \tan\left(\frac{\psi}{2}\right)\frac{\tanh(l/\xi) - \xi/D}{\coth(l/\xi) - \xi/D}. \quad (20)$$

Note that the condition $\xi/D = \tanh(l/\xi)$, which reduces the right-hand side of (20) to zero, determines temperature $T_0(l)$. Above this temperature, only a solution with a small $\Delta(x)$ amplitude induced by a weak exchange field $\mathbf{A}(\pm l)$ at the interface exists. By simplifying (15) at $T > T_0$, we obtain the k' parameter in the explicit form

$$k'^2 = \frac{\left(\frac{A\xi^2}{2c_2 v_F^3}\right)^3\left[\cosh^2\left(\frac{l}{\xi}\right) - \beta\right]}{\left(\tanh\left(\frac{l}{\xi}\right)\cosh^2\left(\frac{l}{\xi}\right) - \frac{\xi}{D}\left[\cosh^2\left(\frac{l}{\xi}\right) - \beta\right]\right)^2 + \beta(1-\beta)}. \quad (21)$$

It is interesting to trace the transition to collinear solutions in (19)–(21). If $\psi \longrightarrow 0$, then $\varphi \longrightarrow 0$, $\beta \longrightarrow 0$, and $\Delta(x) \propto \cosh(x/\xi)$; if $\psi \longrightarrow \pi$, then $\varphi \longrightarrow \pi$, $\beta \longrightarrow 1$, and $\Delta(x) \propto \sinh(x/\xi)$. The $\Delta(l)$ amplitudes for $\psi = 0$ and $\psi = \pi$ coincide with exponential accuracy if $l \gg \xi$,

$$\frac{\Delta(l, \psi = \pi)}{\Delta(l, \psi = 0)} - 1 \propto \exp\left(-\frac{2l}{\xi}\right), \quad (22)$$

and are strongly different if $l \ll \xi$,

$$\frac{\Delta(l, \psi = \pi)}{\Delta(l, \psi = 0)} \propto \left(\frac{l}{\xi}\right)^2. \quad (23)$$

The region of the applicability of the small-amplitude approximation to spin density waves in solving the problem at temperatures above $T_0(l)$ can be estimated as

$$\left(\frac{A\xi^2}{2c_2 v_F^3}\right)^2 \ll \left(\tanh\left(\frac{l}{\xi}\right) - \frac{\xi}{D}\right)^2. \quad (24)$$

A closer approach to the $T_0(l)$ point requires including terms of higher order in k'^2 in (12)–(16). This leads to very cumbersome calculations, which cannot be performed within the scope of this work.

Consider the $T < T_0(l)$ temperature region, within which there exists a solution with the $\Delta(x)$ spin density wave amplitude induced by a charge flow at the interface even in the absence of exchange coupling, that is, at $A = 0$. The $\mathbf{\Delta}_0(x) = \mathbf{n}_z\Delta_+(x)$ solution linearly polarized along $\mathbf{n}_z$ and symmetrical with respect to x, which is most favorable energetically in the whole range of model parameter values at $A = 0$ [9, 10], can naturally be used as a zeroth approximation with respect to the exchange potential. This solution is easy to obtain from (12), (14), and (15) at $A = \beta = \varphi = 0$. The only unknown parameter that remains, k', is determined by the nontrivial solution to (16) at $P = \beta = 0$. Let us denote this parameter by $k' = k'_0$ ($k = k_0$) and find the first nonvanishing correction to $\mathbf{\Delta}_0(x)$, which appears and increases as A deviates from zero. According to (12)–(15), the φ angle is related to the β parameter as follows:

$$\varphi = \frac{2\sqrt{\beta}}{k_0}\frac{l}{\xi}\left(\frac{E(\lambda_0, k_0)}{\lambda_0} - k'^2_0\right), \quad (25)$$

$$\sqrt{\beta} = \frac{A\xi^2}{2c_2 v_F^3}\frac{\sqrt{1-2k'^2_0}}{k_0 k'_0}\mathrm{nc}(\lambda_0, k_0)\sin\left(\frac{\psi}{2}\right), \quad (26)$$

where $E(\lambda, k)$ is the incomplete elliptic integral of the second kind and the argument $\lambda_0 = \lambda$ ($\beta = 0$, $k = k_0$). Changes in k are given by

$$\frac{d}{d(k^2)}(\lambda\,\mathrm{sn}(\lambda, k)\mathrm{dc}(\lambda, k))|_{k = k_0, \beta = 0}(k^2 - k_0^2) = \frac{A\xi l}{2c_2 v_F^3}\frac{\sqrt{1-2k_0^2}}{k'_0}\mathrm{cn}(\lambda_0, k_0)\cos\frac{\psi}{2}. \quad (27)$$

These equations are noticeably simplified for a thin interlayer ($l \ll D$, $\lambda_0^2 \ll 1$),

$$\varphi = 2k_0 l\sqrt{\beta}/\xi, \quad (28)$$

$$\sqrt{\beta} = \frac{A\xi^2}{2c_2 v_F^3}\frac{\sqrt{1-2k'^2_0}}{k_0 k'_0}\sin\left(\frac{\psi}{2}\right), \quad (29)$$

$$-k_0'^2 = \frac{AlD^2}{4c_2 v_F^3 \xi}\frac{\sqrt{1-2k_0'^2}}{k_0'}\cos\left(\frac{\psi}{2}\right), \quad (30)$$

where

$$k_0'^2 = \frac{1}{2}\left(1-\frac{lD}{\xi^2}\right).$$

Hence it follows that the region of the applicability of our approach at $T < T_0(l)$ and $l \ll D$ is limited by the inequality

$$\left(\frac{A\xi^2}{c^2 v_F^3}\right)^2 \ll 1-\frac{lD}{\xi^2}. \quad (31)$$

The corresponding equations for a thick interlayer are obtained by applying a more complex procedure involving the expansion in the small parameter $\exp(-2l/\xi)$. This yields

$$\varphi = 2\sqrt{\beta}, \quad (32)$$

$$\sqrt{\beta} = \frac{A\xi^2}{2c_2 v_F^3}\frac{\sqrt{\xi^2-D^2}}{Dk_0'^2}\sin\left(\frac{\psi}{2}\right), \quad (33)$$

$$k'^2 - k_0'^2 = \frac{AD^2}{2c_2 v_F^3}\left[1+\frac{\xi}{4D}\left(\frac{3l}{\xi}+1\right)\right.$$
$$\left.\times\left(1-\left(\frac{D}{\xi}\right)^2\right)-\frac{1}{4}\frac{1+(D/\xi)^2}{1-(D/\xi)^2}\right]^{-1}\frac{\cos(\psi/2)}{\sqrt{1-(D/\xi)^2}}, \quad (34)$$

where

$$k_0'^2 = 16\exp\left(-\frac{2l}{\xi}\right)\frac{\xi-D}{\xi+D}.$$

Equations (32)–(34) are valid as long as

$$\frac{A\xi^2}{c_2 v_F^3} \ll \exp\left(-\frac{2l}{\xi}\right).$$

To summarize, we showed how a weak deformation $\boldsymbol{\delta}(x) = \boldsymbol{\Delta}(x) - \boldsymbol{\Delta}_0(x)$ of the order parameter, which appears because of interaction between a spin density wave and plate moments at the interface, could be described via several parameters $(\varphi, \sqrt{\beta}, k^2 - k_0^2) \sim A$.

4. THE ENERGY OF ANTIFERROMAGNETIC ORDERING WITH A NONCOLLINEAR SPIN DENSITY WAVE

Substituting the $\boldsymbol{\Delta}(x)$ function that corresponds to a thermodynamic potential extremum into (2)–(5) yields the energy of noncollinear state (11)–(13) in the form

$$F(\psi) = F_V(\psi) + F_S(\psi), \quad (35)$$

$$F_V(\psi) = \frac{c_2 v_F^4 l}{3\xi^4(1-2k'^2)^2}\left[k'^2(1-3k'^2)+3\beta k'^4\right.$$
$$\times(2k'^2-\beta(1-k'^2+k'^4))+\frac{2(1-3\beta k'^2)}{\lambda}$$
$$\times\left((2k'^2-1)\left(E(\lambda,k)-\frac{\mathrm{sn}(\lambda,k)\mathrm{dn}(\lambda,k)}{\mathrm{cn}(\lambda,k)}\right)\right.$$
$$\left.\left.+k'^2(1-3\beta k'^2)\frac{\mathrm{sn}(\lambda,k)\mathrm{dn}(\lambda,k)}{\mathrm{cn}^3(\lambda,k)}\right)\right], \quad (36)$$

$$F_S(\psi) = -\frac{c_2 v_F^4 \lambda^2 k'^2}{Dl^2}\left(\mathrm{nc}^2(\lambda,k)-\beta\left(\frac{l}{\lambda\xi}\right)^2\right)$$
$$-\frac{A\lambda v_F k'}{l}\left(\mathrm{nc}^2(\lambda,k)-\beta\left(\frac{l}{\lambda\xi}\right)^2\right)^{1/2}\cos\left(\frac{\psi-\varphi}{2}\right). \quad (37)$$

These equations are valid for an odd number N of chromium monolayers. The (k, β, φ) parameters are related by (12) and (14)–(16) to the (l, D, ξ, A) nanostructure characteristics and to the ψ angle. The $F(\psi)$ dependence, which is of primary interest to us, cannot be written out explicitly at arbitrary (l, D, ψ, A) values. We will therefore concentrate on several limiting situations.

At a high temperature $T > T_0(l)$, in the region where the small-amplitude spin density wave approximation (24) is applicable, we can expand (36) and (37) in powers of the small parameter $(k'\exp(l/\xi))^2 \ll 1$ and use relations (19)–(21). For odd N values, cumbersome calculations give

$$F(\psi) = F_0 + F_1\cos\psi + F_2\cos^2\psi. \quad (38)$$

Here, in the approximations of the lowest order in parameter A, which we are interested in, we have $(F_0, F_1) \sim A^2$ and $F_2 \sim A^4$, where $(F_0, F_1) < 0$ and $F_2 > 0$. The necessity of retaining terms of the order of A^4 will be justified below. The F_0, F_1, and F_2 coefficients are complex functions of the lengths (l, D, ξ); these functions will not be written in the general form. For thin interlayers with $l \ll \xi$, we obtain

$$F_0 = F_1 = -\frac{A^2\xi^2}{8c_2 v_F^2 l}\frac{1}{1-\xi^2/lD},$$
$$F_2 = \frac{A^4\xi^8}{64l^3 c_2^3 v_F^8(1-\xi^2/lD)^4}, \quad (39)$$

and, for thick layers with $l \gg \xi$, we have

$$F_0 = -\frac{A^2\xi}{4c_2 v_F^2}\frac{1}{1-\xi/D},$$
$$F_1 = \frac{2F_0\exp(-2l/\xi)}{1-\xi/D}, \quad (40)$$

$$F_2 = \frac{A^4\xi^5}{16c_2^3v_F^8}\frac{(1+\xi/D)(5-3\xi/D)}{(1-\xi/D)^6}\exp\left(-\frac{4l}{\xi}\right).$$

It follows from (38)–(40) that the $J(\psi) = F(\psi) - F_0$ effective potential of magnetic moment interactions at $T > T_0(l)$ is a value of the second order in A [because $\Delta(x) \sim k' \sim A$—see (17) and (21)], which rapidly increases in the $T \longrightarrow T_0(l)$ limit. Note also that the term containing $\cos^2\psi$ in (38), which is of fundamental importance, as will be made clear below, is determined by the A^4 value and is always positive. Clearly, when using (39) and (40), we must bear in mind the conditions of the applicability of the approximations used to obtain these equations [see (24)], that is,

$$\frac{A\xi^3}{2c_2v_F^3l} \ll 1 - \frac{\xi^2}{lD} \tag{41}$$

for thin layers and

$$\frac{A\xi^2}{2c_2v_F^3l} \ll 1 - \frac{\xi}{lD} \tag{42}$$

for thick ones. It follows that (38) has the form of the effective energy in the biquadratic exchange model (see Introduction), in which the ratio between the bilinear and biquadratic components is much larger than one, $|F_1/F_2| \gg 1$.

Next, consider the temperature region below the $T_0(l)$ point, where an explicit expression for $F(\psi)$ can also be obtained in certain instances. For instance, if (31) is satisfied, Eqs. (28)–(30) can be used to obtain $F(\psi)$ for odd N in the form

$$F(\psi) = F_0 + F_{1/2}\cos\left(\frac{\psi}{2}\right), \tag{43}$$

in the approximation of the lowest order in A, which we are interested in. We have $(F_0, F_{1/2}) < 0$, and F_0 is independent of A, whereas $F_{1/2} \sim A$. For a thin interlayer ($l \ll \xi$), we can write

$$F_0 = -\frac{c_2v_F^4l}{4\xi^4}\left(1 - \frac{\xi^2}{lD}\right)^2,$$
$$F_{1/2} = -Av_F\left(\frac{1 - lD/\xi^2}{2lD}\right)^{1/2}, \tag{44}$$

whereas, for a thick layer ($l \gg \xi$), we have

$$F_0 = -\frac{c_2v_F^4}{D^3}, \quad F_{1/2} = -\frac{Av_F}{D}\left(1 - \left(\frac{D}{\xi}\right)^2\right)^{1/2}. \tag{45}$$

Equation (43) does not contain terms of the order of A^2 or of higher orders in A [note that these are, in particular, $F_1\cos\psi$- and $F_2\cos^2\psi$-type terms, such as are present in Eq. (38) for the $T > T_0(l)$ region, where $F_1 \sim A^2$ and $F_2 \sim A^4$], because calculating them is a very laborious task, and, below the $T_0(l)$ point, the expansion (43) terms proportional to F_0 and $F_{1/2}$ predominate. Equation (43) fundamentally differs from (38) by the presence of the term linear in A (generally, by the presence of terms with odd powers of A). As a result, the angular dependences $F(\psi)$ in (38) and (43) have substantially different characters. What is more, unlike (38), dependence (43) does not at all reduce to the formula for the effective energy in the biquadratic exchange model.

Consider the physical meaning of this difference. If $T > T_0(l)$, short-range antiferromagnetic order in the chromium layer only arises to the extent to which it is induced by exchange interactions with ferromagnetic iron moments, and the spin density wave amplitude $\Delta(x) \sim A$. It follows that the exchange contribution to energy begins with terms quadratic in A. If $T < T_0(l)$, short-range antiferromagnetic order in the chromium interlayer is induced by charge density redistribution close to the Fe/Cr interface even in the zero order with respect to A, and the spin density wave formed in such a way is oriented in the exchange field created by the ferromagnetic moments of iron plates. The exchange contribution to the energy $F(\psi)$ therefore begins with terms first-order in A. Equations (38) and (43) therefore correspond to different aspects of exchange interaction at the Fe/Cr interface, namely, the induction and the orientation of spin density waves above and below the $T_0(l)$ temperature of short-range antiferromagnetic order formation, respectively.

Note in conclusion that, for an even number N of monolayers in the chromium interlayer, the $\psi \longrightarrow \psi - \pi$ substitution should be made in (38) and (43). Clearly, we then have $\cos\psi \longrightarrow -\cos\psi$ in (38) and $\cos(\psi/2) \longrightarrow \sin(\psi/2)$ in (43). It follows that energy $F(\psi)$ sharply changes as a result of a small (by one monolayer) change in the thickness of the chromium interlayer, which is an important property of the system under consideration.

Equations (38) and (43) allow us to draw the unambiguous conclusion that the absolute minimum of $F(\psi)$ corresponds to $\psi = 0$ for odd N and $\psi = \pi$ for even N. If the ψ angle is determined self-consistently rather than by some external factor such as the action of a magnetic field, only collinear states corresponding to either ferromagnetic or antiferromagnetic orientations of Fe moments in the neighboring plates exist under thermodynamic equilibrium conditions at ideally smooth interfaces, when all interlayer cross sections contain equal numbers of monolayers N. The situation with nonideal interfaces, however, is unlikely to be equally simple.

5. THE INFLUENCE OF INTERLAYER THICKNESS FLUCTUATIONS ON THE FORMATION OF NONCOLLINEAR STATES

As mentioned in the Introduction, the problem of the interaction of a spin density wave with a nonideal (rough) interface in Fe/Cr-type multilayers is exceedingly complex because of the necessity of taking into account many factors that influence this interaction. On the one hand, fluctuations (both large- and small-scale) of charge and exchange interaction potentials between Fe and Cr are observed close to the interface. These fluctuations are related to interdiffusion, frustration of interatomic bonds, and other defects formed during structure growth. On the other hand, spin density waves experience deformations both in the direction of structure growth ($\mathbf{n}_x$) and in the plane of the interface ($\mathbf{n}_y$, $\mathbf{n}_z$). It follows that, formally, initial model (2)–(5) should be complicated at least in two respects: first, the $\nu(\mathbf{r})$ and $A(\mathbf{r})$ parameters should be considered random functions of the $\mathbf{r} = (x, y, z)$ spatial coordinate, and, secondly, we must abandon the one-dimensional approximation in calculating $\mathbf{\Delta}(\mathbf{r})$. Solving a three-dimensional self-consistency equation for order parameter $\mathbf{\Delta}(\mathbf{r})$ at a given configuration of sources (ν, A) followed by averaging over these configurations is exceedingly complex if not impossible. Averaging over small-scale (of the order of interatomic distances) configurations can in principle be performed within the framework of the standard model of the interaction of spin density waves with point impurities [11], but the influence of large-scale fluctuations poses a much more serious problem, which we will not discuss in this work. In addition to the two factors specified above, there is a "geometric" factor that influences the interaction of spin density waves with a nonideally smooth interface. This factor is related to changes (fluctuations) of interlayer thickness in the presence of inhomogeneities of various types on the surface that separates chromium and iron layers. Indeed, the surface contribution F_S to the total energy of the structure sharply changes when the thickness of the interlayer changes by as little as one monolayer, which gives us a hint that the role played by the geometric factor under consideration may be fairly important. The strong dependence of energy $F(\psi, N)$ on the number (odd or even) of monolayers in the interlayer leads us to the idea that this effect can qualitatively be taken into account by a comparatively simple model in analyzing the formation of noncollinear spin density waves in structures with almost ideally smooth interfaces.

Consider a surface with monosteps spaced far apart, that is, with linear defects which jumpwise change the number of chromium monolayers in the interlayer by one. These monosteps separate vast ideally smooth terraces, where the number of monolayers in the interlayer does not change. Let us divide the whole structure into fragments containing either an odd or an even number of monolayers in the interlayer in the cross section normal to the interface. Let Λ be the fraction of fragments with odd N and $(1 - \Lambda)$ be the fraction of fragments with even N. We assume that all fragments are characterized by the same ψ angle between iron moments in the iron plates. We can then, in a very crude approximation, write the total energy of the system as the partial sum of the energies of separate fragments,

$$\langle F(\psi)\rangle = \Lambda F(\psi) + (1 - \Lambda)F(\pi - \psi), \qquad (46)$$

where $F(\psi)$ is given by (38) or (43). Clearly, (46) completely ignores the contribution of spin density wave deformations in interlayer regions adjoining the interface in the vicinity of monosteps. Indeed, it is natural to suggest that the value and form of the $\mathbf{\Delta}(\mathbf{r})$ order parameter at a large distance from monosteps should insignificantly differ from the corresponding characteristic in an interlayer with ideally smooth boundaries and with the same N and ψ values. Of course, spin density waves experience rearrangements close to monosteps (for instance, domain walls are formed), and these disturbances extend over the $\xi_\perp$ characteristic length over the interface in the $(\mathbf{n}_y, \mathbf{n}_z)$ direction. The simplest estimates show that $\xi_\perp \ll \xi$ in the linear defect model in a system with a spin density wave (e.g., see [11]), and the domain wall energy is small compared with the contribution of terraces included in (46) if $\xi_\perp \ll l_\perp$, where $l_\perp$ is the characteristic terrace length in the $\mathbf{n}_y$ or $\mathbf{n}_z$ direction. Generally, calculations of the spin density wave deformation energy go beyond the one-dimensional approximation and require a special study, which is outside the scope of this work.

Let us use (46) to estimate the influence of interlayer thickness fluctuations on the formation of a thermodynamically equilibrium noncollinear state. At $T > T_0(l)$, (46) becomes

$$\langle F(\psi)\rangle = F(\pi/2) + J_1\cos\psi + J_2\cos^2\psi, \qquad (47)$$

where $J_1 = (2\Lambda - 1)F_1$, $J_2 = F_2 > 0$, and $F(\pi/2) = F_0 < 0$, in agreement with (38)–(40). The effective interaction energy between the ferromagnetic plates $E = \langle F(\psi)\rangle - F(\pi/2)$ has the form characteristic of the biquadratic exchange model (see Introduction), and the J_1 bilinear potential is large compared with the J_2 biquadratic potential ($|J_1| > 2J_2$) at almost all Λ in the $0 \le \Lambda \le 1$ interval except a narrow $|\Lambda - 1/2| \ll 1$ region, where the relation $J_1 \le 2J_2$ may hold. We obtained these estimates using the $|F_1| \gg F_2$ inequality, which follows from (38)–(40). If $|J_1| > 2J_2$, the $\langle F(\psi)\rangle$ function given by (47) reaches a minimum either at $\psi = 0$ or at $\psi = \pi$. However, if the condition $|J_1| \le 2J_2$ is met, the $\psi_0 = \arccos(-J_1/2J_2)$ angle corresponds to the minimum of $\langle F(\psi)\rangle$; that is, a noncollinear state is formed. At $\Lambda = 1/2$, we have $\psi_0 = \pi/2$; that is, iron moments are oriented mutually orthogonally in the neighboring plates.

In the $T < T_0(l)$ temperature region, (46) becomes

$$\langle F(\psi)\rangle = F\left(\frac{\pi}{2}\right) + J_+\cos\left(\frac{\psi}{2}\right) + J_-\sin\left(\frac{\psi}{2}\right), \qquad (48)$$

where $J_+ = \Lambda F_{1/2}$, $J_- = (1 - \Lambda)F_{1/2}$, $F(\pi/2) = F_0 < 0$, and $F_{1/2} < 0$, in agreement with (43)–(45). Note that, like (43) above, (48) does not contain $J_1\cos\psi$- and $J_2\cos^2\psi$-type terms, which are always small in the approach that we use. The angular dependence of the $E = \langle F(\psi)\rangle - F(\pi/2)$ effective exchange energy is fundamentally different from its analogue in the biquadratic exchange model. A minimum is attained at the $\psi_0 = 2\arccos(\Lambda/\sqrt{\Lambda^2 + (1 - \Lambda^2)})$ angle; that is, $\psi_0 \neq (0, \pi)$ at all Λ different from 0 and 1. If $\Lambda \longrightarrow 1/2$, ψ_0 tends to $\pi/2$, as in the biquadratic exchange model. Note that ψ_0 only depends on the Λ parameter, which characterizes geometric perfection of the Fe/Cr interface, but is independent of the set of (l, ξ, D, A) values. In this respect, model (48) at temperatures $T < T_0(l)$ obviously differs from model (47) at $T > T_0(l)$.

To summarize, even qualitative estimates by (47) and (48) show that interlayer thickness fluctuations caused by Fe/Cr interface nonideality play an important role in the formation of noncollinear states. The region of Λ parameters where such states exist is narrow at $T > T_0(l)$, when spin density waves are induced by exchange interactions at the Fe/Cr interface, and sharply broadens at $T < T_0(l)$, when spin density waves are formed as a result of charge density redistribution near the interface. Such a difference in the behavior of the system above and below $T_0(l)$ is caused by strong short-range antiferromagnetic ordering in the interlayer at $T < T_0(l)$, which results in an "orientation" mechanism of interactions between spin density waves and magnetic moments in the ferromagnetic plates. This mechanism is in principle absent when interlayers are paramagnetic or when they are antiferromagnetic at temperatures above $T_0(l)$; it cannot be reproduced within the framework of the standard indirect exchange scheme of the RKKY type.

6. CONCLUSION

The mechanism of formation and the conditions of the existence of noncollinear states in the three-level model described by (2)–(5) were considered in the preceding sections in much detail. The problem of using these results to interpret the experimental data on Fe/Cr-type multilayers, however, requires some refinements to be made, primarily concerning the limits of the applicability of model (2)–(5) itself.

A natural limitation on interlayer thickness L from below is the inequality $L > 2\xi_0$, which allows the "surface" and "volume" contributions to thermodynamic potential (2) to be separated. Equally obvious is the lower temperature limit $T > T_N$, where T_N is the temperature of long-range ordering in the antiferromagnetic interlayer. The last inequality is, however, only a necessary but not sufficient condition of the applicability of our model. In reality, temperatures are limited by the condition $T > T_2 \geq T_N$, where T_2 is the temperature below which the use of the density of the volume contribution to the thermodynamic potential in form (4) is inadmissible, and taking into account terms of higher order in $\boldsymbol{\Delta}(x)$ and $\boldsymbol{\Delta}'(x)$ in the Ginzburg–Landau expansion is necessary (more details can be found in [11]). It can be stated with much confidence that both applicability conditions specified above are fulfilled at $T >$ 300–350 K and $L >$ 30–40 Å.

One more restriction is related to the requirement of uniform ferromagnetic layer magnetization; this requirement is not explicitly present in (47) and (48) but is implied and corresponds to the absence of ferromagnetic layer separation into magnetic domains. In all probability, this condition is satisfied in fairly thick (40–50 Å or more) ferromagnetic layers, but can be violated in thinner layers, especially if the Fe/Cr interface contains monosteps. In the presence of monosteps, the formation of domain walls in Fe plates in the immediate vicinity of monosteps, which then separate regions with opposite magnetization polarizations in the ferromagnetic layer, can be more favorable energetically than the formation of domain walls in the Cr interlayer, contrary to what has been suggested above. A scalar spin density wave with a $\Delta_+(x)$ amplitude then forms in both antiferromagnetic interlayer fragments, one with odd and the other with even N, separated by a monostep. It is likely that precisely this situation arises in three-layer Fe/Cr/Fe structures with a wedge-shaped interlayer grown on a thick whisker and coated by a thin (about 20 Å) iron film. The domain structure of a thin iron plate observed by scanning electron microscopy with polarization analysis [1] is direct evidence of oscillations of interlayer exchange coupling caused by monolayer changes in the thickness of the chromium interlayer, whereas the "noncollinear state" proper is only formed in a narrow domain wall region near a monostep at the Fe/Cr interface. This situation, which corresponds to a magnetic structure geometry other than that considered in this work, requires a special study.

Taking into account the restrictions specified above (thick ferromagnetic and antiferromagnetic layers and high temperatures), consider some experimental data which provide direct or indirect evidence of the existence of noncollinear states in Fe/Cr-type multilayers and which can be interpreted in terms of the model that we use. Neutron diffraction, transport, and magnetooptical data have made it possible to roughly construct the magnetic phase diagram (T, L) of the system under consideration [1–3]. This diagram contains the high-temperature ($T > T_0$), intermediate ($T_N < T < T_0$), and low-temperature ($T < T_N$) regions, which have sharply different properties. In the $T > T_0$ region, the state of the interlayer is paramagnetic, the effective coupling between neighboring plate moments is very weak [2], and the structure virtually breaks up into independent ferromagnetic layers separated by a nonmagnetic inter-

layer. Below the T_0 temperature (which substantially depends on thickness L, especially when the interlayer is comparatively thin: $T_0 \approx 500$ K at $L \geq 55$ Å, and $T_0 \approx 600$ K at $L \approx 15$ Å), the effective coupling between neighboring plate moments sharply increases [2], and, in the intermediate temperature range $T_N < T < T_0$, a quite peculiar magnetic state of the structure as a whole is formed. According to [4, 5], this state is characterized by strong short-range antiferromagnetic order, a nonuniform spin density distribution across the interlayer, and a substantial dependence of magnetic characteristics on the properties of the interface between the magnetic and antiferromagnetic layers. Lastly, at $T < T_N$, the effective magnetic coupling between neighboring plate moments again sharply decreases [2]; long-range antiferromagnetic order of the type of an incommensurate spin density wave, similar to that existing in bulk chromium, arises in the interlayer; and the system represents a set of alternating ferromagnetic and antiferromagnetic layers with very weak interlayer interactions.

It follows that the $T_N < T < T_0$ temperature range is the most interesting to study, but information about the magnetic structure of the system at such temperatures is very contradictory. In already mentioned work [5] on the diffraction of neutrons on epitaxial Fe/Cr multilayers, the transition between the intermediate and low-temperature regions was smeared to a large extent (almost over 100 K). This transition was interpreted by the authors as a smooth change of phases with commensurate and incommensurate spin density waves in the chromium interlayer. On the other hand, Fe/Cr multilayers grown differently than in [5] were studied by similar methods in [13]. In [13], no noticeable difference in the magnetic structure of the interlayer was observed between the low-temperature and intermediate regions. According to the authors, the structure experienced only the transition at T_N from a phase with an incommensurate spin density wave directly into the paramagnetic state. Such discrepancies are usually related to the strong influence of the quality of the interface on the formation of spin density waves in chromium interlayers; in our view, this influence is largely a consequence of electron scattering by small-scale surface potential fluctuations. From the point of view of the possibility of formation of noncollinear states, precisely the $T_N < T < T_0$ intermediate region with short-range antiferromagnetic order in the chromium interlayer is most promising, because strong magnetic coupling between neighboring ferromagnetic Fe layers is observed precisely at these temperatures. The results of our analysis (see Section 5) predict a substantial increase in interlayer coupling at $T_N < T < T_0$ compared with the $T > T_0$ temperature region. These results are in a sense substantiated by experiments [5] on large-angle neutron scattering in Fe/Cr superlattices with interlayers $L = 42$ Å thick. Such experiments are very important for directly determining the type of magnetic structure in the temperature range of interest to us, and their results clearly indicate the noncollinear character of ordering of ferromagnetic layer magnetic moments and the presence of helicoidal spin density wave components in antiferromagnetic interlayers. The neutron diffraction pattern with a multicomponent intensity obtained in [5] could not be used to unambiguously reproduce the shape of the spin density wave. The authors therefore resorted to the so-called torsional phenomenological model [7] for the purpose of data processing in the spirit of helicoidal magnetic order with a spin density wave. A preliminary analysis showed that the microscopic noncollinear state model suggested above is also in satisfactory agreement with the results obtained in [5] and explains the observed "three-peak" dependence of the intensity of neutron scattering on wave vector k in the direction of structure growth.

Indirect evidence of the existence of a noncollinear state at $T_N < T < T_0$ can be found in works on Brillouin scattering and magnetometry [1–4, 14] and on magnetization and magnetoresistance [1–3, 15]. Residual magnetization $\mathbf{M}(\mathbf{H} \longrightarrow 0)$ in external field $\mathbf{H}$ measured in these works was close to half the saturation magnetization. This result was interpreted by the authors as a consequence of an almost orthogonal mutual orientation of neighboring ferromagnetic layers, and rapid saturation of the $\mathbf{M}(\mathbf{H})$ dependence was in conformity with the biquadratic exchange scheme. We have not studied our model in an external magnetic field yet because this problem is not so trivial as may seem at first sight. The type of spin density waves can change from symmetrical to antisymmetric in comparatively thick interlayers and a fairly strong field $\mathbf{H}$, and the corresponding $\Delta_-(x)$ solution to the self-consistency equation (see [10]) can be more energetically favorable than the $\Delta_+(x)$ solution obtained in this work in the absence of an external magnetic field. As a result, a structure with an amplitude domain wall of the spin density wave within the interlayer or even at the interface can form. Clearly, calculations of all these variants are far beyond the scope of this work. Such calculations are, nevertheless, necessary in prospect.

Work [6] holds a special position among other works on the problem of the existence of noncollinear states that we are considering. This work substantially broadens the class of layered structures with antiferromagnetic ordering of the type of spin density waves in interlayers between ferromagnetic plates. In [6], the magnetic and transport characteristics of $Fe/Cr_{1-x}Fe_x$ superlattices with $x = 0.06$ were studied. We cannot go into details of this interesting investigation and will only mention the main points. It was found that comparatively thick ($L > 36$ Å) interlayers were characterized by two antiferromagnetic ordering temperatures, T_N and T_0. The low-temperature ($T < T_N$) antiferromagnetic phase unambiguously corresponded to the state with a commensurate spin density wave, which was observed in bulk $Cr_{1-x}Fe_x$ alloys with $x > 0.024$. The

high-temperature ($T > T_0$) phase was paramagnetic, and, in the intermediate $T_N < T < T_0$ temperature range approximately 150 K wide, a complex nonuniform antiferromagnetic state with an unclear structure was formed. As in Fe/Cr superlattices, effective coupling between neighboring iron layers through the interlayer was exceedingly weak at $T > T_0$ and $T < T_N$ but sharply increased in the $T_N < T < T_0$ temperature range. The T_N and T_0 temperatures were interpreted in [6] as transition points between two antiferromagnetic states and between antiferromagnetic and paramagnetic states, respectively. What is more, the special features of the hysteresis of magnetization and magnetoresistance led the authors of [6] to suggest that the magnetic state in the $T_N < T < T_0$ temperature range was noncollinear, and they made at attempt at describing it in terms of the biquadratic exchange scheme. In our view, a thorough study of the results obtained in [6] sheds light on several interesting issues. Incidentally, one of these is quite obvious: the T_N temperature is not at all the transition point between antiferromagnetic states with commensurate and incommensurate spin density waves, as is claimed in the overwhelming majority of works (including reviews [1–3]). Rather, this is, in agreement with our model, the transition point between antiferromagnetic short- and long-range order states, which have origins of different natures and different spin density wave structures.

Lastly, we would like to mention work [16] on ferromagnetic resonance in Fe/Cr superlattices. The results obtained in that work are evidence of the formation of a noncollinear state and are interpreted in terms of the phenomenological biquadratic exchange scheme. Unfortunately, small antiferromagnetic layer thicknesses (from 7.6 to 10 Å) used in the systems studied in [16] interfere with directly applying our model to analyze these experimental results. We, nevertheless, hope that, basically, we reached a correct understanding of the mechanism of formation of noncollinear states, and the consistent theoretical approach suggested above gives a qualitatively correct description of the experimental situation in Fe/Cr-type multilayers even outside the limits of its formal applicability.

ACKNOWLEDGMENTS

One of the authors (V. V. T.) is deeply indebted to N.M. Kreĭnes for discussions of the problem of noncollinear states in Fe/Cr multilayers and for providing information about experiments on ferromagnetic resonance including unpublished data.

This work was financially supported by the Russian Foundation for Basic Research (project no. 01-02-16175).

REFERENCES

1. D. T. Pierce, J. Unguris, R. J. Celotta, and M. D. Stiles, J. Magn. Magn. Mater. **200**, 290 (1999).
2. H. Zabel, J. Phys.: Condens. Matter **11**, 9303 (1999).
3. R. S. Fishman, J. Phys.: Condens. Matter **13**, R235 (2001).
4. A. Schreyer, J. F. Anker, Th. Zeidler, *et al.*, Phys. Rev. B **52**, 16066 (1995).
5. A. Schreyer, C. F. Majkrzak, Th. Zeidler, *et al.*, Phys. Rev. Lett. **79**, 4914 (1997).
6. E. E. Fullerton, C. H. Sowers, and S. D. Bader, Phys. Rev. B **56**, 5468 (1997).
7. J. C. Slonczewski, J. Magn. Magn. Mater. **150**, 13 (1995).
8. R. S. Fishman and Z. P. Shi, Phys. Rev. B **59**, 13849 (1999); R. S. Fishman, Phys. Rev. Lett. **78**, 1351 (1997).
9. M. Avignon, V. Men'shov, and V. Tugushev, Europhys. Lett. **56**, 132 (2001).
10. V. N. Men'shov and V. V. Tugushev, Zh. Éksp. Teor. Fiz. **120**, 899 (2001) [JETP **93**, 786 (2001)].
11. V. V. Tugushev, in *Electronic Phase Transitions*, Ed. by W. Hanke and Yu. V. Kopaev (North-Holland, Amsterdam, 1992), Modern Problems in Condensed Matter Sciences, Vol. 32, p. 239.
12. *Handbook of Mathematical Functions*, Ed. by M. Abramowitz and I. A. Stegun (National Bureau of Standards, Washington, 1964; Nauka, Moscow, 1979).
13. E. E. Fullerton, S. D. Bader, and J. L. Robertson, Phys. Rev. Lett. **77**, 1382 (1996).
14. P. Vavassori, M. Grimsditch, and E. Fullerton, J. Magn. Magn. Mater. **223**, 284 (2001).
15. E. E. Fullerton, K. T. Riggs, C. H. Sowers, and S. D. Bader, Phys. Rev. Lett. **75**, 330 (1995).
16. A. B. Drovosekov, O. V. Zhotikova, N. M. Kreĭnes, *et al.*, Zh. Éksp. Teor. Fiz. **116**, 1817 (1999) [JETP **89**, 986 (1999)].

Translated by V. Sipachev

Journal of Experimental and Theoretical Physics, Vol. 95, No. 5, 2002, pp. 912–916.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 122, No. 5, 2002, pp. 1057–1062.

SOLIDS
Electronic Properties

Tunnel Magnetotransport in Heterostructures with Quantum Rings

V. M. Kovalev and A. V. Chaplik*

Institute of Semiconductor Physics, Siberian Division, Russian Academy of Sciences, pr. Akademika Lavrent'eva 13, Novosibirsk, 630090 Russia

**e-mail: chaplik@isp.nsc.ru*

Received April 1, 2002

Abstract—A tunnel current through a heterostructure whose barrier contains quantum rings is calculated. The plane of the rings is parallel to the barrier interface. In a magnetic field perpendicular to this plane, a tunnel current at a fixed bias experiences Aharonov–Bohm oscillations under the variation of magnetic flux through a ring; however, these oscillations are not strictly periodic. *© 2002 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

The problem of resonant tunneling of electrons has a long history that actually dates back to the Ramsauer–Tausend phenomenon (an abnormally small scattering cross section of electrons by atoms at certain energies of an impinging particle). The effect of impurities in the barrier on the tunnel current was repeatedly discussed as applied to solid state physics. In Schmidlin's work [1], the role of impurities was reduced to increasing the tunneling transparency due to the local reduction of the barrier height caused by the fluctuation potential of the impurities. Kane [2] analyzed a doped *p*–*n* junction and showed that a tunnel current can be increased (as compared with a "pure" structure) due to transitions between the "tails" of the density of states near the band edges.

As far as we are concerned, Parker and Mead [3] were the first who showed (by an example of the Schottky barrier) that the presence of impurity levels in the barrier makes the tunneling a two-step process, thus considerably increasing the partial tunnel current at a resonance energy as compared with a direct tunnel current. The authors of [3] applied a kinetic method; i.e., it was required that the balance condition had to be satisfied for particles arriving at the impurity level from the semiconductor and outgoing from the impurity to the metal electrode. Such an approach is classical and can be applied under ordinary classical conditions that allow one to neglect interference phenomena when calculating the probability of a two-step transfer.

In [4], Chaplik and Entin developed a quantum theory of resonant tunneling through a short-range impurity center in a barrier (1D barrier + 3D impurity) and showed that, in addition to the energy resonance, there also is a geometric resonance in the tunnel current: the peak current attains its maximum for a certain definite position of the impurity in the barrier. Naturally, this is the symmetry plane for a symmetric barrier; however, when a voltage is applied to the tunnel structure (which is always the case in the experiment), the point of geometric maximum is displaced, and the impurity centers situated in the barrier successively "enter" or "leave" the resonance. This results in a rather complicated shape of the current–voltage characteristic, which has a negative-slope region under certain conditions.

Larkin and Matveev [5] generalized the Bardeen method of tunnel Hamiltonian to the case of resonant tunneling through one or two short-range impurities and determined the conductivity and the correlation function of conductivity fluctuations in mesoscopic semiconductor contacts.

All the aforementioned studies involved the impurity centers with a single discrete energy level. For a short-range impurity center, the resonance component of the partial tunnel current is described by the Breit–Wigner-type formula

$$j_{\text{res}}(E) \propto \frac{\Gamma^2}{\Gamma^2 + (E - E_0)^2}, \qquad (1)$$

where E_0 is the particle energy on the impurity center, E is the energy of the impinging particle, and Γ is the tunneling width of the level.

If a barrier contains an impurity with several bound states, or a quantum point, the problem can easily be solved in the limit $\Gamma \ll \Delta E$, where ΔE is on the order of the distance between the impurity levels (quantum-point levels). In this case, the net current is a sum of resonance contributions and, represented as a function of applied voltage V (i.e., difference between the Fermi levels at the contacts), it has a staircase form.[1] When a magnetic field **B** is applied to the system perpendicular to the heteroboundary, the tunnel current (for fixed V)

[1] Until (under very large biases) the bottom of the conduction band of the emitter becomes higher than the highest impurity level in the barrier.

becomes a steplike and monotonic function of the field because the number of resonance levels below the Fermi level varies monotonically with the field.

All the aforesaid applies to the situation when the tunneling occurs from a 3D emitter and, under an applied magnetic field, the spectrum of the particles impinging on the barrier remains continuous. In the case of a 2D emitter, the electron spectrum becomes discrete both in and "before" the barrier, and the current–voltage characteristic represents a series of spikes corresponding to resonance transitions between two systems of discrete levels.

The aim of the present study is to consider the tunnel transport through a heterostructure containing quantum rings in its plane. From general considerations, one should expect that the tunnel current will exhibit the Aharonov–Bohm effect under the application of a perpendicular magnetic field to the system. Indeed, we will show below that the tunnel current oscillates under the variation of the magnetic field (flux), although, generally, these oscillations are neither monochromatic nor strictly periodic. Recall that progress in technology has led to the development of ensembles of quantum rings with the radius on the order of the Bohr radius in a relevant material [6]. Experiments on tunnel magnetotransport may prove to be a useful means for investigating such structures.

2. SPECTRUM AND WAVE FUNCTIONS OF A QUANTUM RING

We assume (according to the results of [6]) that a ring consists of a narrowband material (for example, InAs) and is immersed into a barrier, which is a wideband semiconductor (GaAs). The energy diagram of the system, corresponding to the distance from the ring center equal to its radius, is shown in Fig. 1; the coordinate z is measured along a normal to the heterostructure (the tunnel current flows in this direction). Thus, the interior and exterior regions of the ring are occupied with the barrier material, and the resonant tunneling corresponds to the particle trajectories that pass through the points on the ring.

Obviously, the model of one-dimensional ring $|\Psi|^2 \propto \delta(r-a)\delta(z-z_0)$ is insufficient for solving the problem posed because the resonant tunneling is determined by the overlap of the wave function of a particle impinging on the barrier and that of a particle bound in the ring. We will represent the potential of the ring in cylindrical coordinates by the expression

$$U(\mathbf{R}) = -U_0\delta(r-a)u(z-z_0), \tag{2}$$

where $U_0 > 0$, a is the Bohr radius, z_0 is the position of the ring in the barrier plane, and $u(z)$ is a dimensionless function that is everywhere positive and has a sharp maximum at $z = z_0$ (see Fig. 1). Assume that the potential well is sufficiently narrow along z so that the wave

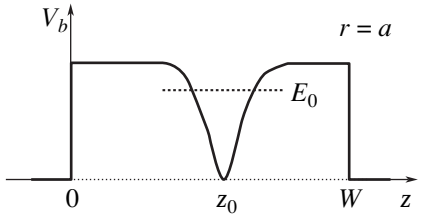


Fig. 1. The energy diagram of the structure with quantum rings.

function of the bound state cannot appreciably change over a distance on the order of the width of the function $u(z)$. Let $u(q)$ be the Fourier transform of $u(z)$. Then, after separating the variable φ, the Schrödinger equation in the q representation in z is expressed as

$$\nabla_r^2\psi(r,q) - \left(k^2 + q^2 + \frac{m^2}{r^2}\right)\psi(r,q) + \frac{2MU_0}{\hbar^2}\delta(r-a)\psi(r,z_0)u(q) = 0. \tag{3}$$

Here, $E = -\hbar^2k^2/2M$; M is the electron mass in the barrier, where the electron mainly resides according to model (2); $m = 0, \pm1, \pm2, \ldots$ is the azimuthal quantum number; and

$$\psi(r,z_0) = \frac{1}{2\pi}\int\psi(r,q)e^{iqz_0}dq. \tag{4}$$

Applying a conventional method to solving the radial equation (3) with the potential in the form of a δ function and determining $\psi(r, z_0)$ in a self-consistent manner, we obtain from (4) the following equation for the energy eigenvalues in the ring:

$$1 = \lambda a\int_{-\infty}^{\infty} K_{|m|}(a\sqrt{k^2+q^2}) \times I_{|m|}(a\sqrt{k^2+q^2})u(q)dq, \tag{5}$$

where $I_{|m|}$ and $K_{|m|}$ are Bessel functions of imaginary argument of the first and third kind, respectively. As is clear from (5), one cannot replace $u(z)$ by a δ function because all the energy levels E_m in the ring will tend to minus infinity in this case. However, as will be clear from the analysis below, the main, resonant, part of the rotational spectrum and the tunneling conductivity depends on the form of the potential through a single parameter E_0. Hence, the ring model (2) provides a fairly reasonable approximation. For moderately small radii of the ring, the rotational quantum $\hbar^2/2Ma^2$ is much less than the depth of the ground state $|E_0|$; i.e., $ka \gg 1$. In this limit, one can easily derive the following

asymptotic expression from (5) (the argument of the Bessel functions is greater than their index):

$$E_m = E_0 + \frac{\hbar^2 m^2}{2Ma^2} + \frac{\hbar^2 m^2 (m^2 + 5/2)}{2Ma^2 (k_0 a)^2}. \quad (6)$$

Here, the second term corresponds to the spectrum of a one-dimensional ring (a plane rotator), while the third term represents a correction due to the finite width of the electron wave function spread out near the circle $r = a$.

In what follows, we need the wave functions of the bound states corresponding to energy E_m. These states are solutions to Eq. (3) that are finite at zero and decrease as $r \longrightarrow \infty$; they are represented as follows:

$$\begin{aligned}\psi_m(r < a, z) &= c_m \int_{-\infty}^{\infty} K_{|m|}(a\sqrt{k_m^2 + q^2}) \\ &\times I_{|m|}(r\sqrt{k_m^2 + q^2}) \exp[q(z - z_0)]dq, \\ \psi_m(r > a, z) &= c_m \int_{-\infty}^{\infty} I_{|m|}(a\sqrt{k_m^2 + q^2}) \\ &\times K_{|m|}(r\sqrt{k_m^2 + q^2}) \exp[q(z - z_0)]dq,\end{aligned} \quad (7)$$

where c_m is a normalizing factor.

One can see that the wave functions depend on the form of $u(z)$ only through k_m, i.e., ultimately through the parameter E_0.

3. TUNNEL CURRENT IN THE MODEL OF A δ-SHAPED SOLENOID

Suppose that an infinitely thin solenoid with magnetic flux Φ passes through a quantum ring situated inside the barrier of a tunnel structure. The flow of tunneling electrons is parallel to the solenoid. The energy levels in the one-dimensional ring are classified according to the momentum $m = 0, \pm 1, \pm 2$. Accordingly, a plane wave incident to the system is expanded in terms of cylindrical harmonics:

$$\Psi(\mathbf{r}) = \frac{1}{\sqrt{2\pi\Omega}} \sum J_\nu(k_\parallel r) \exp(im\varphi),$$

where Ω is a normalizing volume, J is a Bessel function, $\nu = |m + \Phi/\Phi_0|$, r and φ are cylindrical coordinates in the plane of the structure, and $k_\parallel$ is the projection of the wave vector of the incident wave onto this plane; the number m is preserved during tunneling.

The wave functions of the bound states in the ring and the spectrum of appropriate energies are given by formulas (5)–(7) in which the index m of the Bessel functions should be replaced by ν.

To determine the tunnel current, we apply the Bardeen method and calculate the tunnel width Γ of the bound state and the transition amplitudes T_p and $T_{p'}$ from the state in the ring to the left and right contacts, respectively. As a result, we obtain

$$\begin{aligned}\Gamma &= \Gamma_l + \Gamma_r, \\ \Gamma_{l(r)} &= \pi \sum_{p(p')} |T_{p(p')}(E_{p(p')})|^2 \delta(E_{p(p')} - E_m),\end{aligned} \quad (8)$$

where E_p and $E_{p'}$ are, respectively, the energies of the electrons impinging on and transmitted through the barrier and $T_{p(p')} = E_m \langle \Psi_{p(p')} | \Psi_m \rangle$ are the overlapping integrals of the functions of the bound and free states, respectively. The tunneling probability from state p (on the left) to state p' (on the right of the barrier) through the level E_m is given by

$$\omega_{pp'} = \frac{2\pi}{\hbar} \frac{|T_p|^2 |T_{p'}|^2}{\Gamma^2 + (E_p - E_m)^2} \delta(E_p - E_{p'}). \quad (9)$$

For the tunnel current, we have

$$I = e \sum_{p,\, p'} \omega_{pp'} \{ f(E_p - \mu) - f(E_{p'} - \mu + eV) \},$$

where f is the Fermi–Dirac distribution function, μ is the chemical potential, and V is the voltage applied to the barrier.

For the conductivity of the system at $T = 0$ K, we have

$$\begin{aligned}g(\Phi) &= g_0 \sum_m E_m^2 \\ &\times \frac{I_{|m+\Phi|}(k_m^2 a^2 \theta_1^2/2) I_{|m+\Phi|}(k_m^2 a^2 \theta_2^2/2)}{\left(1 + \frac{\pi}{6} \frac{(m + \Phi/\Phi_0)^2 - 1/4}{k_m a}\right)(\Gamma^2 + (\mu - E_m)^2)},\end{aligned} \quad (10)$$

where

$$\theta_1^2 = \frac{V_b \sqrt{1 - E_m/V_b}}{k_b z_0 E_m}, \quad \theta_2^2 = \frac{V_b \sqrt{1 - E_m/V_b}}{k_b (W - z_0) E_m},$$

V_b is the barrier height, W is its width, and $k_b = \sqrt{2MV_b/\hbar^2}$.

Recall that, here, the energies E_m are functions of the combination $|m + \Phi/\Phi_0|$; therefore, the tunneling conductivity $g(\Phi)$, being a sum over m from $-\infty$ to $+\infty$, is a periodic function of the magnetic flux with period Φ_0. The numerically calculated values of $g(\Phi)$ are shown in Fig. 2. Because the model with a δ-shaped solenoid has a purely illustrative character, we performed the calculation for the case when electrons inside and outside the barrier have identical effective masses.

4. UNIFORM MAGNETIC FIELD

We consider the tunneling from a 3D emitter to a 3D collector. Outside the barrier, tunneling electrons are described by the Fock–Darwin wave functions (when the moment with respect to the normal to the system and the momentum along the field are specified). The general case involves very laborious calculations; therefore, we restrict the analysis to the limit case $B < \hbar\omega_c$, where $\hbar\omega_c$ is the Landau quantum. Obviously, the above inequality is equivalent to the condition $4\Phi \gg \Phi_0$ (the number of flux quanta passing through the ring is large). Suppose that the Fermi level lies between the lowest and the first Landau levels of the impinging electron; i.e., let us restrict ourselves to the range of $\delta\Phi$ that contains several quanta Φ_0 but is significantly less than the value of Φ such that $\hbar\omega_c \sim \mu$. Then, the functions $\Psi_{0,m}(\mathbf{r})$, which correspond to the lowest Landau level and $m \in (-\infty, 0)$, serve as the wave functions of the initial state (on the left of the barrier). Hence, one can see that the tunneling conductivity is not a periodic function of the magnetic flux since it is expressed in terms of a sum over a semi-infinite interval of m.

The tunneling widths are determined by the expressions

$$\Gamma_l = \frac{4}{3}\left(\frac{M_1}{2\hbar^2}\right)^{1/2} k_m^2 a^2 \frac{E_m^2 G^2(\nu)}{1 + \frac{\pi}{6}\frac{\nu^2 - 1/4}{ka}} \times \frac{\sqrt{E_m - \hbar\omega_c/2}}{\left(1 - \frac{M_1}{M_2}\right)\left(E_m - \frac{\hbar\omega_c}{2}\right) + \frac{M_1}{M_2}V_b} \times \exp\left(-2k_b z_0 \sqrt{1 - \frac{E_m - \hbar\omega_c/2}{V_b}}\right). \quad (11)$$

Here, one should make the substitution $z_0 \longrightarrow W - z_0$; M_1 and M_2 are the electron effective masses outside and inside the barrier, respectively.

In Eq. (11),

$$G(\nu) = \int_0^\infty \frac{J_\nu(sa)}{s^2 + k^2} R_\nu(s) s\,ds,$$

$$R_\nu(s) = \frac{2^{1+|m|/2}}{\sqrt{|m|!}} a_H^{1+\nu} \frac{\Gamma\left(1 + \frac{\nu + |m|}{2}\right)}{\Gamma(1+\nu)} \times s^\nu F\left(1 + \frac{\nu + |m|}{2};\ 1 + \nu;\ -a_H^2 s^2\right) \quad (12)$$

is the Fourier–Bessel transform of the wave functions of the state outside the barrier for the lowest Landau level with a moment $m \le 0$.

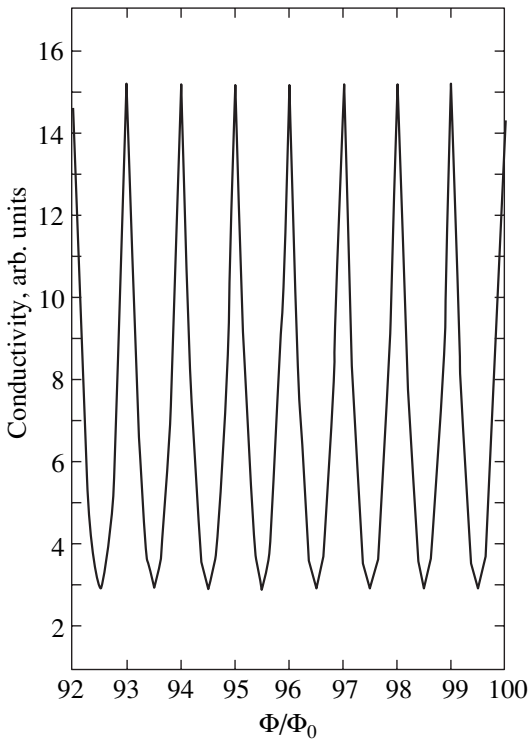


Fig. 2. Tunneling conductivity versus magnetic flux (thin solenoid).

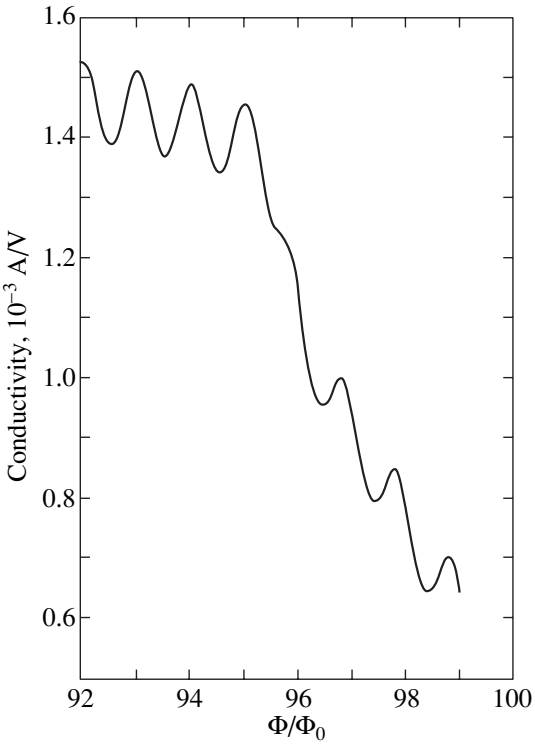


Fig. 3. Tunneling conductivity versus magnetic flux (uniform field).

The conductivity ($T = 0$) is given by

$$g(\Phi) = \frac{16}{9}\pi\frac{e^2}{\hbar}\sum_{m=-\infty}^{0}\frac{M_1 k_m^4 a^4}{\hbar^2 a^2} \times \frac{\mu - \hbar\omega_c/2}{\left(\left(1-\frac{M_1}{M_2}\right)\left(E_m - \frac{\hbar\omega_c}{2}\right) + \frac{M_1}{M_2}V_b\right)^2} \times \exp\left(-2k_b W\sqrt{1-\frac{\mu-\hbar\omega_c/2}{V_b}}\right) \times \frac{E_m^4 G^4(\nu)}{\left(1+\frac{\pi}{6}\frac{(m+\Phi/\Phi_0)^2 - 1/4}{k_m a}\right)(\Gamma^2 + (\mu - E_m)^2)}. \quad (13)$$

The numerical calculation was performed for the following parameters: $\mu = 1.03|E_0|$, $M_1 = 0.025m_0$ (InAs), $M_2 = 0.07m_0$ (GaAs), $a = 300$ Å, $V_b = 0.55$ eV, $W = 100$ Å, and $z_0 = W/2$.

The results of calculations (Fig. 3) show that the oscillating character of $g(\Phi)$ is preserved. The characteristic interval between the spikes is equal to Φ_0, the modulation depth of conductivity is about 15%, and $\Gamma \sim 0.002$–0.01 eV for a magnetic field strength of $H \sim 4.04$–4.4 T.

A decrease in the conductivity as the magnetic field increases is attributed to the fact that, for a fixed total energy of a tunneling electron, its "longitudinal" energy (the continuous part of the spectrum) decreases with the field. Therefore, the interval between this energy and the barrier ridge increases, and the decay length of the wave function along z decreases. As a result, the overlapping integral T_p decreases.

Thus, we have shown that a tunnel current through a heterostructure with quantum rings exhibits a specific Aharonov–Bohm effect. The tunnel current as a function of magnetic field for a given voltage across the structure has the form of modulated oscillations with a characteristic period Φ_0 in the flux. In the magnetic field scale, this interval is much less than the scale related to the Landau quantization for rings with radii on the order of 10–100 nm in fields on the order of 1–10 T.

ACKNOWLEDGMENTS

This work was supported by the Russian Foundation for Basic Research (project no. 02-02-16377) and by the programs "Physics of Solid State Nanostructures" "Quantum Macrophysics" and "Mesoscopic Systems" of the Ministry of Education of the Russian Federation.

REFERENCES

1. F. W. Schmidlin, J. Appl. Phys. **37**, 2823 (1968).
2. E. O. Kane, Phys. Rev. **131**, 79 (1963).
3. G. H. Parker and C. A. Mead, Appl. Phys. Lett. **14**, 21 (1969); Phys. Rev. **184**, 780 (1969).
4. A. V. Chaplik and M. V. Éntin, Zh. Éksp. Teor. Fiz. **67**, 208 (1974) [Sov. Phys. JETP **40**, 106 (1974)].
5. A. I. Larkin and K. A. Matveev, Zh. Éksp. Teor. Fiz. **93**, 1030 (1987) [Sov. Phys. JETP **66**, 580 (1987)].
6. R. J. Warburton *et al.*, Nature **405**, 926 (2000).

Translated by I. Nikitin

Journal of Experimental and Theoretical Physics, Vol. 95, No. 5, 2002, pp. 917–926.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 122, No. 5, 2002, pp. 1063–1073.

SOLIDS
Electronic Properties

Peculiarities of the Magnetic, Galvanomagnetic, Elastic, and Magnetoelastic Properties of $Sm_{1-x}Sr_xMnO_3$ Manganites

A. I. Abramovich*, L. I. Koroleva, and A. V. Michurin**

Moscow State University, Moscow, 119992 Russia
**e-mail: avramovich@ofef343.phys.msu.ru*
***e-mail: koroleva@ofef343.phys.msu.ru*

Received February 5, 2002

Abstract—Manganites of the $Sm_{1-x}Sr_xMnO_3$ system (x = 0.33, 0.4, and 0.45) possess giant negative values of the magnetoresistance $\Delta\rho/\rho$ and the volume magnetostriction ω near the Curie temperature T_C. In the compound with x = 0.33, the isotherms of $\Delta\rho/\rho$, ω, and magnetization σ exhibit smooth variation and do not reach saturation up to maximum magnetic field strengths (120 kOe) studied (according to the neutron diffraction data, this substance comprises a ferromagnetic (FM) matrix with distributed clusters of a layered antiferromagnetic (AFM) structure of the *A* type). In the compounds with x = 0.4 and 0.45 containing, besides the FM matrix and *A*-type AFM phase, a charge-ordered AFM phase of the *CE* type (thermally stable to higher temperatures as compared to the *A*-type AFM and the FM phases), the same isotherms measured at $T \geq T_C$ show a jumplike increase in the interval of field strengths between H_{c1} and H_{c2} and then reach saturation. In the interval $H_{c1} < H < H_{c2}$, the σ, ω, and $\Delta\rho/\rho$ values exhibit a metastable behavior. At temperatures above T_C, the anisotropic magnetostriction changes sign, which is indicative of rearrangements in the crystal structure. The giant values of ω and $\Delta\rho/\rho$ observed at $T \geq T_C$ for all compounds, together with excess (relative to the linear) thermal expansion and a maximum on the $\rho(T)$ curve, are explained by the phenomenon of electron phase separation caused by a strong *s–d* exchange. The giant values of magnetoresistance and volume magnetostriction (with ω reaching ~10^{-3}) are attributed to an increase in the volume of the FM phase induced by the applied magnetic field. In the compound with x = 0.33, this increase proceeds smoothly as the FM phase grows through the FM layers in the *A*-type AFM phase. In the compounds with x = 0.4 and 0.45, the FM phase volume increases at the expense of the charge-ordered *CE*-type AFM structure (in which spins of the neighboring manganese ions possess an AFM order). The jumps observed on the $\sigma(H)$ curves, whereby the magnetization σ reaches ~70% of the value at T = 1.5 K, are indicative of a threshold character of the charge-ordered phase transition to the FM state. Thus, the giant values of ω and $\Delta\rho/\rho$ are inherent in the FM state, appearing as a result of the magnetic-field-induced transition of the charge-ordered phase to the FM state, rather than being caused by melting of this phase. © *2002 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

The interest in manganites is related primarily to the phenomenon of colossal magnetoresistance observed in some compounds at room temperature, which can be employed in various sensor devices. Various properties of manganites were extensively studied and the results of numerous investigations have been summarized in reviews [1–12]. The colossal magnetoresistance is usually observed in the region of temperatures corresponding to breakage of a magnetically ordered state characterized by nonzero spontaneous magnetization.

Previously [13–18], we established that the colossal magnetoresistance in $La_{1-x}Sr_xMnO_3$ ($0.1 \leq x \leq 0.3$) and $Sm_{0.55}Sr_{0.45}MnO_3$ compounds is accompanied by a large negative volume magnetostriction (on the order of 10^{-4}–10^{-3}), which makes possible the application of these manganites in various magnetomechanical devices. In the former system, we observed a correlation between magnetoresistance and volume magnetostriction, which was manifested by a decrease in both values with increasing x and by the absence of saturation in their isotherms up to maximum magnetic field strengths (~120 kOe) studied, while the isotherms of magnetization already exhibited saturation in fields $H <$ 10 kOe. This behavior was explained in terms of the coexistence of ferromagnetic (FM) and antiferromagnetic (AFM) phases in the crystalline compounds, with the charge carriers (in this case, holes) concentrated in the FM phase. This type of two-phase magnetic state related to a strong *s–d* exchange was described in [19] and reviewed in [1, 2]. It should be noted that a notion about the Curie temperature of a sample occurring in such a state is rather conditional.

A compound of the $Sm_{1-x}Sr_xMnO_3$ system with x = 0.5 exhibits a charge-ordered state. Investigations of the neutron diffraction at 1.5 K $\leq T \leq$ 300 K showed that the

compound $^{154}Sm_{0.6}Sr_{0.4}MnO_3$ contains the clusters of three types—FM, *A*-type AFM, and charge-ordered (*CE*-type) AFM [20]—and that an AFM order of the *A* type disappears above 120 K, while the charge-ordered *CE* state still exists at 150 K but is absent at 180 K. A spontaneous magnetic moment appears at $T_C \approx 110$ K; at the same temperature, the resistivity ρ exhibits a maximum. At $T < T_C$, the compound possesses conductivity of the metallic type. The temperature dependence of the conductivity activation energy exhibits a jump at 180 K; in compounds possessing a charge-ordered state, this is usually interpreted as a thermal breakage of this order. The above values of T_C, T_N, and the temperature of the charge order breakage (T_{CO}) are close to those determined [21] from the electron and neutron diffraction data and the temperature dependence of magnetization. Runov *et al.* [20, 22, 23] showed that the samples of $Sm_{1-x}Sr_xMnO_3$ compounds with $x = 0.25$ and 0.4 at low temperatures contain FM clusters with 180–250 Å dimensions coexisting with large clusters up to several thousand angstroms in size. At the same time, no charge-ordered phase clusters were observed in the compound with $x = 0.25$ [20]. Luzyanin *et al.* [24] studied the second harmonic of the magnetic susceptibility and suggested that the paramagnetic phase features AFM correlations with a weak FM component coexisting with charge-ordered AFM domains up to several hundred angstroms in size. The latter clusters exhibit thermal breakage at temperatures significantly above the Curie temperature.

Thus, the magnetic structure of $Sm_{1-x}Sr_xMnO_3$ compounds is more complicated as compared to that of the $La_{1-x}Sr_xMnO_3$ system. Compounds of the former system with *x* close to 0.5, in contrast to $La_{1-x}Sr_xMnO_3$ compounds, contain charge-ordered clusters. This order exhibits thermal breakage at higher temperatures than does the magnetic order in the FM and *A*-type AFM phases of the sample. On the other hand, the compounds of both systems showed the colossal magnetoresistance and large volume magnetostriction near T_C. First, we studied magnetostriction in the $Sm_{1-x}Sr_xMnO_3$ system only for a compound with $x = 0.45$ [16–18]. Previously [13–18], we suggested that the colossal magnetoresistance and large volume magnetostriction are explained by the same reason: a growth of the FM part of a crystal, in which the charger carriers are concentrated. However, the mechanism of this growth remained unclear. As will be shown in Section 3 below, compounds of the $Sm_{1-x}Sr_xMnO_3$ system with $x = 0.4$ and 0.45 exhibit the colossal magnetoresistance and large volume magnetostriction even in a high-temperature range, where the magnetic order in both FM and *A*-type AFM parts is broken but a charge-ordered state is still retained in a certain part of the sample.

The aim of this paper is as follows. Based on the results of a complex study of the behavior of magnetization, paramagnetic susceptibility, thermal expansion, magnetostriction, electric resistance, and magnetoresistance in the $Sm_{1-x}Sr_xMnO_3$ compounds with $x = 0.33$, 0.4 and 0.45, we will elucidate some peculiarities in these characteristics related to the presence of charge-ordered phase clusters. It should be noted that the $Sm_{1-x}Sr_xMnO_3$ system offers a convenient object for such investigation, since the compounds with $x = 0.4$ and 0.45 contain charge-ordered clusters. Moreover, only these clusters still retain a magnetic order at $T \geq T_N$, while that in the FM and *A*-type AFM parts is broken as a result of heating. The compound with $x = 0.33$ differs from those with $x = 0.4$ and 0.45 in that the former does not contain charge-ordered clusters [20]. For all the three compounds, $T_C < T_N$. Thus, we will elucidate the character (jumplike versus smooth) of the FM phase growth at the expense of the charge-ordered and *A*-type AFM phases considered separately. The results are treated based on the recent theoretical works of Nagaev [2] and Dagotto *et al.* [3].

2. EXPERIMENTAL METHODS

The samples of compounds of the $Sm_{1-x}Sr_xMnO_3$ system were synthesized and analyzed by O.Yu. Gorbenko and A.R. Kaul'. The synthesis was conducted according to a standard ceramic technology. The phase composition and crystal lattice parameters were monitored by X-ray diffraction measured on a Siemens D5000 diffractometer. According to these data, the samples represent single-phase perovskites with an orthorhombic structure (*Pnma* space group) and the lattice parameters ($T = 300$ K) $a = 5.424(1)$ Å, $b = 7.678(2)$ Å, $c = 5.434(2)$ Å for $x = 0.45$ and $a = 5.436(2)$ Å, $b = 7.679(4)$ Å, $c = 5.459(2)$ Å for $x = 0.33$. The orthorhombicity parameter of 0.2% (calculated using the lattice parameters) indicates that the structure is close to cubic. A single-phase state of the synthesized samples was confirmed by the Raman spectroscopy measurements performed on a Jobin–Yvon T64000 spectrometer with triple monochromator. The Raman spectra showed the presence of the phonon modes characteristic of the orthorhombic manganites with *Pnma* symmetry.

The magnetization was studied using a vibrating-sample magnetometer in a range of temperatures from 1.5 to 300 K and magnetic field strengths up to 130 kOe. The initial magnetic susceptibility in an alternating magnetic field with the frequency varied from 0.8 to 8 kHz was measured on an F-5063 ferrometer, while the paramagnetic susceptibility was studied by a weighing technique with electromagnetic compensation. The electric resistance was determined by the four-point-probe technique. The magnetostriction and thermal expansion in the temperature range from 4.2 to 300 K were measured using strain gauges with a resistance of 92.30 ± 0.01 Ω and a tension sensitivity coefficient of 2.26. One strain gauge was glued to a sample, and another, to a quartz crystal, with the same orienta-

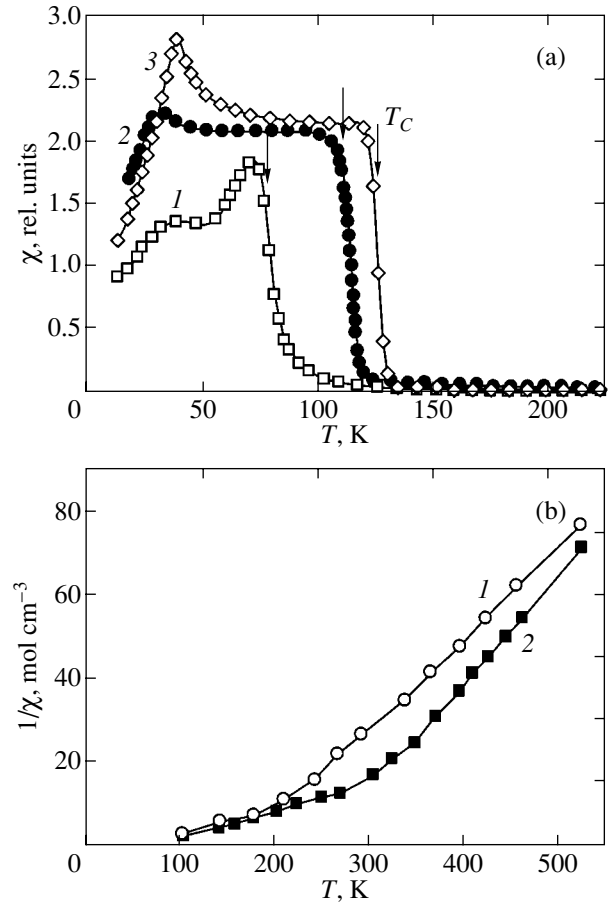


Fig. 1. Temperature dependences of (a) the initial magnetic susceptibility χ measured in an alternating magnetic field of 1 Oe with a frequency of 8 kHz for $Sm_{1-x}Sr_xMnO_3$ compounds with x = 0.33 (*1*), 0.4 (*2*), and 0.45 (*3*) and (b) the paramagnetic susceptibility of the compounds with x = 0.33 (*1*) and 0.45 (*2*).

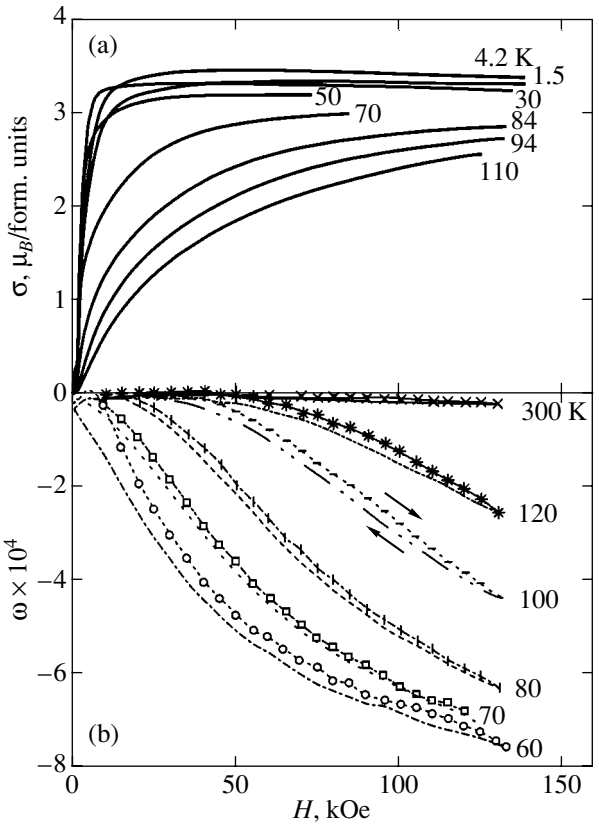


Fig. 2. The isotherms of (a) magnetization σ and (b) volume magnetostriction ω for $Sm_{1-x}Sr_xMnO_3$ with x = 0.33 measured with increasing and decreasing the applied field strength as indicated by arrows.

tion relative to the applied magnetic field. The magnetostriction parallel ($\lambda_\parallel$) and perpendicular ($\lambda_\perp$) to the applied magnetic field were measured, after which the volume ($\omega = \lambda_\parallel + 2\lambda_\perp$) and anisotropic ($\lambda_t = \lambda_\parallel - \lambda_\perp$) magnetostriction were calculated. The measurements of thermal expansion and magnetostriction in strong pulsed magnetic fields were performed in the laboratory of R. Ibarra (Saragossa University, Spain).

3. EXPERIMENTAL RESULTS

3.1. Magnetic Properties

For all compounds of the system studied, the temperature dependence of the initial magnetic susceptibility $\chi(T)$ measured in an alternating magnetic field with a frequency of 8 kHz exhibited a maximum in the low-temperature region (T < 50 K) and a sharp drop at higher temperatures (Fig. 1a). In the compounds studied, the Curie point was determined as the temperature of a minimum in the $(d\chi/dT)(T)$ curve. For the $Sm_{1-x}Sr_xMnO_3$ compounds with x = 0.33, 0.4, and 0.45, this yields 79, 112, and 126 K, respectively. These values are close to the published data (see [20, 21] and references therein). It should be noted that the low-temperature peak position is virtually independent of x; this maximum is probably related to an increase in the magnetic anisotropy of FM and AFM clusters in the two-phase magnetic state. The paramagnetic susceptibility of all compositions obeys the Curie–Weiss law at temperatures above ~$2T_C$ (Fig. 1b), with the paramagnetic Curie points at θ = 175, 194, and 250 K for x = 0.33, 0.4, and 0.45, respectively. In the temperature interval $T_C < T < 2T_C$, the experimental curves deviate from the Curie–Weiss law.

Figures 2a and 3a show the isotherms of the magnetization $\sigma(H)$ for the compounds with x = 0.33 and 0.45, respectively (the curves for x = 0.4 are much like those in Fig. 3a). As can be seen, the $\sigma(H)$ curves display no peculiarities below T_C and reach saturation at $H \le 30$ kOe. The spontaneous magnetic moment, determined by extrapolating $\sigma(H)$ curves to a zero field at 1.5 K, is close to the values expected in the case of complete FM ordering of the magnetic moments of Mn^{3+} and Mn^{4+} ions in the compositions studied. The magnetization

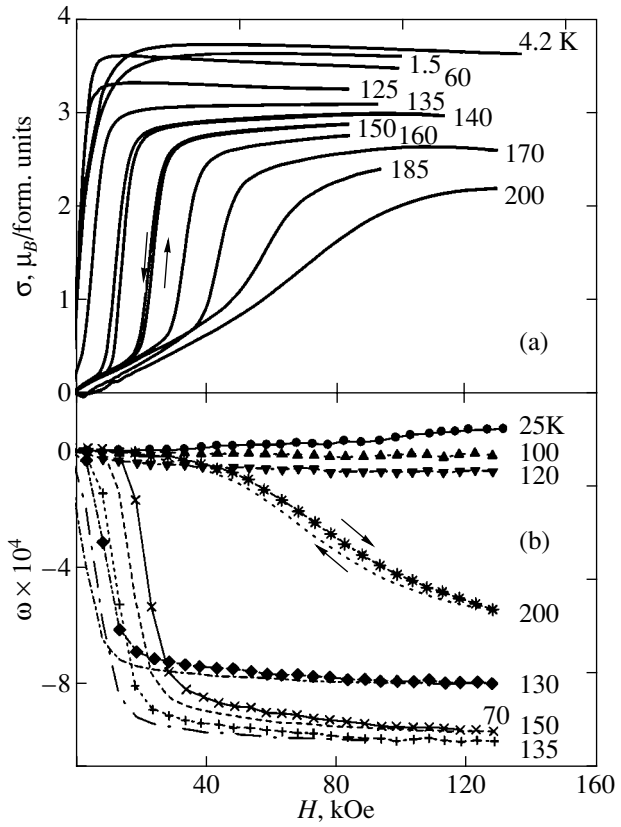


Fig. 3. The isotherms of (a) magnetization σ and (b) volume magnetostriction ω for $Sm_{1-x}Sr_xMnO_3$ with $x = 0.45$ measured with increasing and decreasing the applied field strength as indicated by arrows.

isotherms of various compositions are significantly different at $T \geq T_C$. Indeed, the samples with $x = 0.4$ and 0.45 exhibit a jumplike increase in the magnetization in a certain interval of field strength $H_{c1} < H < H_{c2}$ (Fig. 3a). Figure 4 shows the temperature variation of the H_{c1} and H_{c2} values. As can be seen, both characteristic fields increase with the temperature by a linear law. In the region of a jump, the $\sigma(H)$ curves exhibit a hysteresis whose width decreases with an increase in the temperature. As can be seen from Fig. 2a, no such jump is observed on the $\sigma(H)$ curves of a sample with $x = 0.33$.

Figures 5a and 5c show the temperature variation of magnetization of the same compositions in various magnetic fields. For a compound with $x = 0.45$, a decrease in the temperature leads to a sharp increase in the magnetization in the vicinity of T_C for the samples measured in magnetic fields below 20 kOe. In stronger fields, the transition exhibits smearing. The $\sigma(T)$ curves for the compound with $x = 0.4$ are very much like those presented in Fig. 5a. For the sample with $x = 0.33$, the transition at T_C is smeared in the entire range of magnetic fields studied (Fig. 5c).

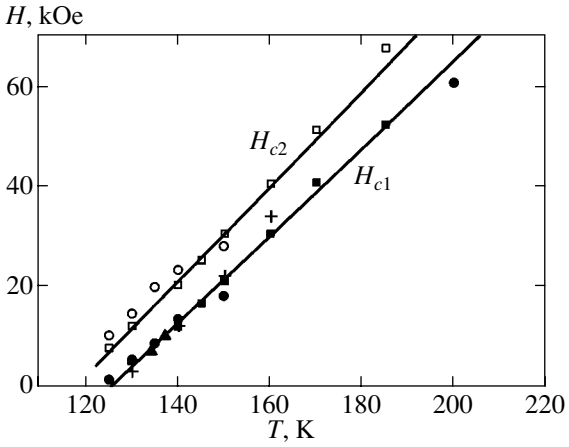


Fig. 4. Plots of the critical field strengths H_{c1} and H_{c2} versus temperature for $Sm_{1-x}Sr_xMnO_3$ with $x = 0.45$. Symbols shows the experimental data obtained from (□, ■) magnetization, (○, ●) magnetostriction, and (▲, +) magnetoresistance measurements ((+) data from [27, Fig. 4]).

3.2. Elastic and Magnetoelastic Properties

Figure 6a shows the temperature variation of the linear thermal expansion $\Delta L/L(T)$ for all three compositions studied. As can be seen the $\Delta L/L(T)$ exhibits a sharp change in the vicinity of T_C for the compounds with $x = 0.4$ and 0.45, and a smoother variation for the sample with $x = 0.33$. For the first two compositions, a relative change in the sample volume is very large: $\Delta V/V = 3\Delta L/L \approx 0.1\%$. The application of a magnetic field leads to suppression of the jump in $\Delta L/L(T)$ and in the negative magnetostriction. The samples of all compositions exhibit a considerable temperature hysteresis of $\Delta L/L(T)$ in the region of T_C.

Figures 2b and 3b show the isotherms of the volume magnetostriction $\omega(H)$ for the compounds with $x = 0.33$ and 0.45, respectively. Figures 5b, 5d and 6b, 6c present the temperature dependences of the volume ($\omega(T)$) and anisotropic ($\lambda_t(T)$) magnetostriction of the same compositions. The curves of $\omega(H)$, $\omega(T)$, and $\lambda_t(T)$ for the compound with $x = 0.4$ are very much like those depicted in Figs. 3b, 5b, and 6c for $x = 0.45$. As can be seen from Fig. 3b, the sample with $x = 0.45$ exhibits a change in the behavior of $\omega(H)$ in the vicinity of T_C, whereby a jump appears in the curves. These jumps are observed at approximately the same values of the magnetic field strength (for close temperature) as the jumps in the curves of $\sigma(H)$ (Fig. 3a). As can be seen from Fig. 4, showing the temperature dependence of the critical fields H_{c1} and H_{c2} determined from the magnetization and magnetostriction measurements, the experimental points are satisfactorily fitted to straight lines. In the region of $T = T_C$, the $\omega(H)$ curves exhibit saturation at a field about 40 kOe, while an increase and decrease in the magnetic field strength reveals a hysteresis in the volume magnetostriction.

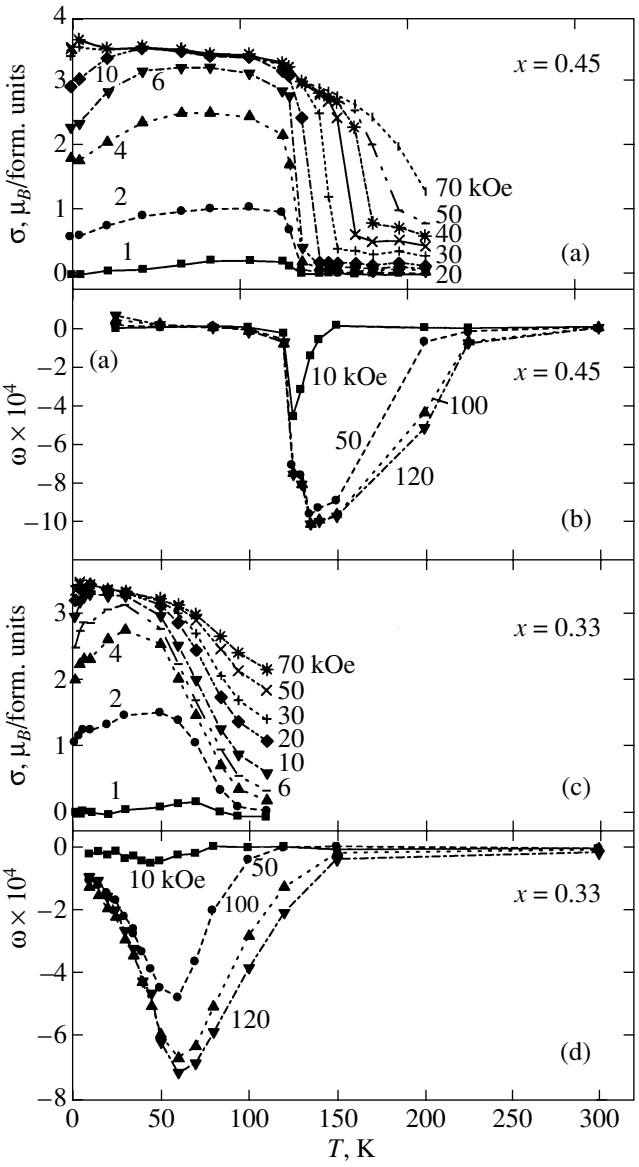


Fig. 5. Temperature dependences of (a, c) the magnetization σ and (b, d) the volume magnetostriction ω for $Sm_{1-x}Sr_xMnO_3$ compounds with $x = 0.33$ and 0.45 measured in various magnetic fields.

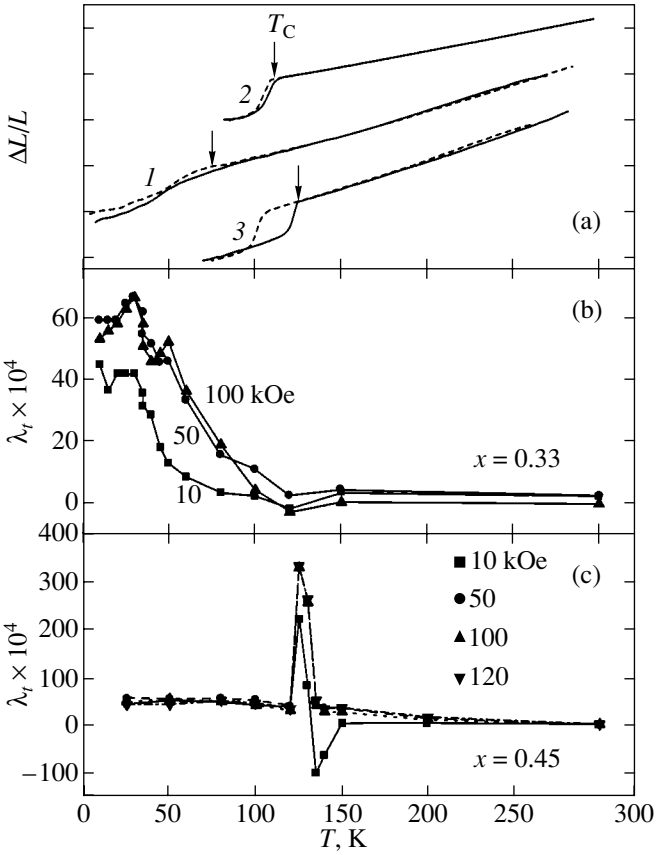


Fig. 6. Temperature dependences of (a) the linear thermal expansion $\Delta L/L(T)$ (in units of 1×10^{-5}) for $Sm_{1-x}Sr_xMnO_3$ with $x = 0.33$ (*1*), 0.4 (*2*), and 0.45 (*3*) (solid and dashed curves measured in the heating and cooling modes, respectively) and (b, c) the anisotropic magnetostriction of the compounds with $x = 0.33$ and 0.45 measured in various magnetic fields.

For the compounds with $x = 0.4$ and 0.45, the volume magnetostriction is negative in the entire temperature range studied and possesses a very small absolute value everywhere, except for a very narrow interval in the vicinity of T_C where the $\omega(T)$ curve passes through a minimum (Fig. 5b). At this minimum, $|\omega|$ acquires giant values: 5×10^{-4} in a field of $H = 8.4$ kOe and 10^{-3} in a field of $H = 50$ kOe for $x = 0.45$ (and approximately the same values for $x = 0.4$).

As can be seen from Fig. 6c, the anisotropic magnetostriction λ_t for the compounds with $x = 0.4$ and 0.45 acquires small absolute values, being almost constant in the entire temperature range studied, except within a narrow interval in the vicinity of T_C. In this interval, the λ_t value increases severalfold and changes sign in the interval of field strengths $0 < H < 2$ kOe, being positive below T_C and negative above this temperature. For the fields $H > 20$ kOe, the $\lambda_t(T)$ exhibits a maximum at $T = T_C$. A different behavior of $\omega(H)$, $\omega(T)$, and $\lambda_t(T)$ is observed for the compound with $x = 0.33$ (Figs. 2b, 5d, and 6b). Indeed, the isotherms of volume magnetostriction are far from saturation in the entire temperature range studied, while the magnetization isotherms already exhibit saturation at $H \leq 30$ kOe. Negative volume magnetostriction in this compound is observed in the entire temperature range below T_C. The curves of $\omega(T)$ pass through a minimum somewhat below T_C, while a maximum absolute value of $|\omega|$ (on the order of 7×10^{-4}) is attained at a field strength of 120 kOe (Fig. 5d). The λ_t value is positive in the entire temperature range studied, exhibiting a maximum in the vicinity of T_C and then quite rapidly decreasing to zero (Fig. 6b).

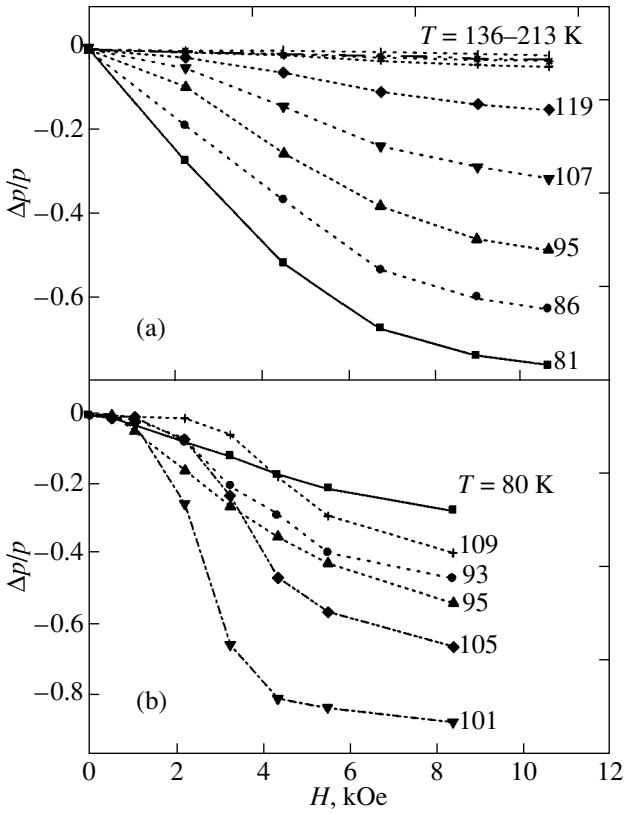


Fig. 7. The isotherms of magnetoresistance for $Sm_{1-x}Sr_xMnO_3$ with $x = 0.33$ (a) and 0.4 (b).

3.3. Electrical and Galvanomagnetic Properties

In the vicinity of T_C, all the compounds studied are characterized by a $\rho(T)$ curves with a maximum, with the ρ values at the peak increased by several orders of magnitude as compared to those at low temperatures. In the low-temperature region, the $\rho(T)$ curve has a shape characteristic of metals. The magnetoresistance $\Delta\rho/\rho = [\rho(H) - \rho(H = 0)]/\rho(H = 0)$ is negative for all compound in the entire temperature range studied (from $T = 4.2$ K to $T > T_C$), with the absolute value reaching a maximum near T_C.

Figure 7 presents the plots of $\Delta\rho/\rho$ versus H for the compounds with $x = 0.33$ and 0.4. As can be seen from Fig. 7b, the latter compound exhibits a sharp growth of $|\Delta\rho/\rho|$ when the field strength reaches H_{c1} (the same behavior was observed for the compound with $x = 0.45$). Figure 7a shows that the $\Delta\rho/\rho$ value in the compound with $x = 0.33$ smoothly increases with the field strength and the curves exhibit no saturation. It should be noted that the magnetoresistance of samarium-containing manganites reaches giant values in relatively weak magnetic fields as compared to those for the manganites of other rare-earth metals (e.g., lanthanum, praseodimium). Indeed, we observed the magnetoresistance reaching 83, 72, and 44% in the samples of $Sm_{1-x}Sr_xMnO_3$ with $x = 0.33$, 0.4, and 0.45, respectively, in a magnetic field of 8.4 kOe.

3.4. Instability of Magnetoelastic Properties

The compounds with $x = 0.4$ and 0.45 exhibit instability in their magnetic properties in the interval of magnetic field strengths $H_{c1} < H < H_{c2}$ at temperatures in the vicinity of T_C. This is manifested by an increase in magnetization, magnetostriction, and magnetoresistance of the samples exposed to the magnetic field with the strength increased from H_{c1} to H_{c2}. If the applied magnetic field is decreased in the same interval, the characteristics vary in the reverse direction. It was found that the time variation of magnetostriction follows an exponential law.

The experimental results for H = const (fast field buildup to $H = 8.4$ kOe at various temperatures in the region of T_C) are well described by the activation relationship $\tau = \tau_0 \exp(E_0/kT)$ with $\tau_0 = 1.4 \times 10^{-7}$ s and $E_0/k = 0.17$ eV. The experiments at T = const in the same range of field strengths are described by the activation law

$$\tau = \tau_0 \exp[(\mu_0 mH/kT)],$$

while the results involving both temperature and field variations are well described by the equation

$$\tau = \tau_0 \exp[(\mu_0 mH + E_0)/kT],$$

where $m = 230\mu_B$ and $E_0 = 0.17$ eV.

4. DISCUSSION OF RESULTS

The results of investigation of the magnetic, galvanomagnetic, elastic, and magnetoelastic properties of $Sm_{1-x}Sr_xMnO_3$ manganites (with $x = 0.33$, 0.4, and 0.45) in a range of temperatures from 1.5 to 300 K and magnetic field strength up to 130 kOe revealed the following features in behavior of the system under consideration. In the vicinity of T_C, the samples exhibit excess (relative to the linear) thermal expansion reaching up to 0.03% (Fig. 6a), giant negative magnetoresistance (Fig. 7), and large negative volume magnetostriction (Figs. 2b, 3b, 5b, and 5d). The latter two temperature dependences exhibit maxima. The absolute value of magnetoresistance in the compounds with $x = 0.33$, 0.4, and 0.45 in a magnetic field of 8.4 kOe reaches 83, 72, and 44%, respectively. Record values were observed for the volume magnetostriction of compounds: for the compounds with $x = 0.4$ and 0.45, we observed $|\omega| \sim 5 \times 10^{-4}$ in a field of 8.4 kOe and $\sim 10^{-3}$ in a field of $H = 50$ kOe. The measurements of paramagnetic susceptibility showed (Fig. 1b) that all compositions obeys the Curie–Weiss law at temperatures significantly above T_C, typically beginning at $T \approx 2T_C$, which is indicative of the presence of magnetic clusters in the system above T_C.

The behavior of magnetization, thermal expansion, volume magnetostriction, and magnetoresistance of the compound with $x = 0.33$ is very much like that previously observed by some of the authors in the $La_{1-x}Sr_xMnO_3$ system [13–15]: the $\Delta\rho/\rho(H)$ and $\omega(H)$ curves do not exhibit saturation up to maximum field strengths studied ($H = 130$ kOe), while the magnetization $\sigma(H)$ already saturates at $H \sim 20$ kOe (Figs. 2a and 7a). The above features can be explained in terms of the electron phase separation caused by a strong *s–d* exchange. The compound with $x = 0.33$ comprises a conducting FM matrix with distributed insulating clusters of the *A*-type AFM phase. This structure features a conducting two-phase magnetic state caused by a strong *s–d* exchange as described in reviews [1, 2].

The two-phase magnetic state is characterized by a sharp increase in resistivity in the vicinity of the Curie point. There are two mechanisms through which the magnetic-impurity interaction influences the resistivity: (i) the scattering of charge carriers, leading to a decrease in their mobility, and (ii) the formation of a tail of localized states in the conduction band. In the vicinity of the Curie point, there is a sharp decrease in the mobility of charge carriers and they are partly localized in the band tail, which explains the appearance of a maximum at T_C in the $\rho(T)$ curve. The applied magnetic field produces delocalization of the charge carriers localized in the band tail, which gives rise to the colossal magnetoresistance. Yanase and Kasuya [19] showed that the FM part of a sample occurring in a two-phase magnetic state is characterized by reduced lattice parameters. On heating above T_C, the sample exhibits anomalous expansion. Application of a magnetic field at $T \geq T_C$ increases the FM order near the magnetic impurity (in this case, Sr ions) to an extent greater than average over the crystal because the effect is enhanced by the *s–d* exchange (magnetic field restores the FM phase broken by heating, after which the sample exhibits a lattice contraction corresponding to this phase). It should be noted that, for a sample occurring in a two-phase magnetic state, the Curie point is essentially the temperature of thermal breakage of the FM phase.

It should be emphasized that the above explanation of the phenomenon of giant negative magnetoresistance, large negative volume magnetostriction, maximum in the resistivity ρ, and an excess thermal expansion near T_C is related to a change in the FM phase volume under the action of an applied magnetic field and/or the temperature. In $Sm_{1-x}Sr_xMnO_3$ with $x = 0.33$, this change proceeds smoothly, probably, because the FM phase volume increases at the expense of AFM clusters of the *A* type. These clusters possess a layered AFM structure comprising FM ordered sublattices with AFM-ordered magnetic moments. The sample FM part volume increases as the FM phase grows through the FM layers in the *A*-type AFM phase.

Behavior of the isotherms of magnetization, volume magnetostriction, and magnetoresistance in the samples of $Sm_{1-x}Sr_xMnO_3$ with $x = 0.4$ and 0.45 is different from that observed for the compound with $x = 0.33$. As was pointed out in the Introduction, the former two compositions contain, in contrast to that with $x = 0.33$, charge-ordered clusters which exhibit thermal breakage at higher temperatures than does the conducting FM phase and the *A*-type AFM clusters deprived of charge carriers. As can be seen from Figs. 3 and 7b, these samples exhibit a jumplike increase in the magnetization, volume magnetostriction, and magnetoresistance at $T \geq T_C$ in the region of critical fields $H_{c1} < H < H_{c2}$, where the H_{c1} and H_{c2} values linearly increase with the temperature (Fig. 4). Within the critical field interval, the above magnetic characteristics exhibit temporal relaxation. In the region of a jump, the $\sigma(H)$ curve exhibits a hysteresis whose width decreases with increasing temperature. At $H \geq H_{c2}$, the aforementioned isotherms exhibit saturation.

If the Curie temperature is determined by extrapolating the steepest part of the $\sigma(T)$ curve to intersection with the temperature axis, the T_C' values obtained in this way depend on the magnetic field strength at which the measurements of $\sigma(T)$ were performed. For the compounds with $x = 0.4$ and 0.45, the T_C' values vary in a nonuniform manner, slowly in the region of weak fields and much faster in stronger fields, so that eventually the T_C' value exhibits a twofold increase at 70 kOe as compared to T_C' determined from the temperature dependence of the initial magnetic susceptibility (Fig. 5a). For the compound with $x = 0.33$, the increase in T_C' with the field strength is still more pronounced and proceeds more uniformly, which makes determination of the T_C' values in strong fields impossible (Fig. 5c). Such a strong dependence of T_C' on the field H is not typical of ferromagnets. For example, the T_C' values determined by the same method for Gd and $CdCr_2Se_4$ in the same interval of field strengths increase by only a few Kelvins [25]. All this suggests that the concept of Curie temperature in magnetically heterogeneous (nonsingle-phase) materials is rather conditional.

The results of numerical modeling performed by Dagotto *et al.* [3] showed that the concentration transition at $x = 0.5$ from FM to charge-ordered state in manganites is a first-order phase transition and that the compositions close to $x = 0.5$ contain magnetic clusters of various types, representing FM, *A*-type AFM, and charge-ordered states. Also considered in [3] was the possibility that there exists a conducting FM phase containing both the *A*-type AFM and charge-ordered clusters. Apparently, this possibility is realized in the $Sm_{1-x}Sr_xMnO_3$ compositions with $x = 0.4$ and 0.45, as evidenced by the metallic type of conductivity at $T < T_C$ in combination with a high level of Jahn–Teller distortions of the manganese–oxygen octahedra. According to the data reported in [20], a high level of these distor-

tions in the compound with $x = 0.4$ is untypical of a strongly diluted system of Jahn–Teller ions (60% Mn^{3+}): this level is approximately as high as in $LaMnO_3$ (100% Mn^{3+}) [26]. Such distortions must lead to the localization of charge carriers, but $Sm_{0.6}Sr_{0.4}MnO_3$ exhibited conductivity of the metallic type below T_C. This is possible only provided that the conductivity possesses a percolation character and is mediated by tie links surrounding the insulating clusters (these links occupy a much smaller part of the sample as compared to that occupied by the insulating clusters). Then, the jumplike growth of $\sigma(H)$, $|\omega|(H)$, and $|\Delta\rho/\rho|(H)$, as well as the nonuniform variation of $T_C'(H)$ in the compounds with $x = 0.4$ and 0.45, can be related to a transition of the charge-ordered clusters to the FM state under the action of an applied magnetic field.

Thus, the large volume magnetostriction and the colossal magnetoresistance in the compounds with $x = 0.4$ and 0.45 (as well as in that with $x = 0.33$) are related to the FM phase appearing as a result of the magnetic-field-induced transition of the charge-ordered clusters to the FM state. In the charge-ordered phase, spins of the neighboring Mn ions are oriented in the opposite directions. Therefore, conversion of this phase into the FM state requires a considerably greater amount of energy than the analogous conversion of the *A*-type AFM phase. As can be seen from Figs. 3, 5a, 5b, and 7b, this transition proceeds in a jumplike manner when the applied magnetic field strength reaches the first threshold value H_{c1}. The magnetic-field-induced transition of the charge-ordered clusters to the FM state must be accompanied by modification of the crystal structure and, hence, manifested by a change in the anisotropic magnetostriction. Indeed, the data in Fig. 6c show a change in the sign of λ_t in this temperature region, which is accompanied by the appearance of certain features in the $\lambda_t(T)$ curve: a maximum in the region of positive λ_t values and a minimum in the region of negative λ_t values.

A strong field dependence of the T_C' values determined as described above can be explained as follows. In the compounds with $x = 0.4$ and 0.45, T_C' in weak fields exhibits a smooth increase related to the magnetic-field-induced restoration of the FM phase (broken by heating) and, probably, to an increase in the volume of this phase at the expense of the *A*-type AFM clusters. When the applied field strength reaches the H_{c1} threshold, the charge-ordered phase exhibits a jumplike transition to the FM state. This conversion is analogous to a first-order phase transition, which is confirmed by the fact that the system exhibits instability of the magnetic, elastic, magnetoelastic, electrical, and galvanomagnetic properties in the interval of field strengths between H_{c1} and H_{c2} (see Sections 3.1–3.4). Above this interval, thermal breakage of the FM phase slows down because the energy is gained in the *s–d* exchange. This explains a smooth increase in T_C' with H in the range of $H > H_{c2}$. In the compound with $x = 0.33$, the smooth growth of T_C' with H is related both to the magnetic-field-induced restoration of the FM phase and to the FM phase growth through the FM sublattices of the *A*-type AFM clusters.

Thus, the colossal magnetoresistance, large negative volume magnetostriction, and the excess thermal expansion near T_C observed in samples of the $Sm_{1-x}Sr_xMnO_3$ system with $x = 0.33$, 0.4, and 0.45 are explained by the phenomenon of electron phase separation caused by a strong *s–d* exchange, whereby the charge carriers are concentrated in the FM phase. The presence of charge-ordered clusters in the compounds with $x = 0.4$ and 0.45 brings certain special features in behavior of the magnetoresistance and volume magnetostriction, which is manifested by jumps in the corresponding isotherms related to the magnetic-field-induced transition of the charge-ordered clusters to the FM state.

5. CONCLUSION

We have studied a model $Sm_{1-x}Sr_xMnO_3$ system (with $x = 0.33$, 0.4, and 0.45) to elucidate how the charge ordering can influence the magnetic, elastic, magnetoelastic, and galvanomagnetic properties of manganites. This system is a convenient object for such investigation, since the compounds with $x = 0.4$ and 0.45 comprise a ferromagnetic phase with distributed *A*-type AFM clusters and charge-ordered clusters such that $T_C < T_N < T_{CO}$. The latter relation implies that the behavior observed at $T > T_N$ is related to the magnetic-field-induced modification of only the charge-ordered clusters. The compound with $x = 0.33$, in contrast to those with $x = 0.4$ and 0.45, contains only the FM and *A*-type AFM phases.

For these reasons, a comparison of the properties of compounds with $x = 0.4$ and 0.45 to those of the compound with $x = 0.33$ allowed us to establish the nature of the colossal magnetoresistance and large volume magnetostriction observed in the system under consideration and to reveal the peculiarities of these effects related to the charge ordering. In the compounds with $x = 0.4$ and 0.45, the isotherms of magnetization, volume magnetostriction, and magnetoresistance measured at T immediately above T_C exhibit a slow initial growth with the field strength, followed by a significant jumplike increase in the interval between the critical fields H_{c1} and H_{c2}, the values of which linearly increase with the temperature (Figs. 3, 4, and 7b). As the field grows further (above H_{c2}), the aforementioned isotherms exhibit saturation. In contrast, the same isotherms of the composition with $x = 0.33$ exhibit a smooth character and show no evidence of saturation up to maximum field strengths studied (Figs. 2 and 7a).

In the compounds with $x = 0.4$ and 0.45, the magnetization exhibits a jump and acquires a large value. For example (see Fig. 3a) a sample of the composition with $x = 0.45$ at $T = 200$ K ($T_C = 126$ K) has $\sigma = 2\mu_B$/form. unit, while the magnetization at $T = 1.5$ K amounts to $3.5\mu_B$/form. unit (which corresponds to a complete FM ordering of the moments of Mn^{3+} and Mn^{4+} ions at $T = 1.5$ K). From this, we conclude that the compounds with $x = 0.4$ and 0.45 feature the transformation of charge-ordered clusters into the FM state at the threshold field strength H_{c1}. This conversion is analogous to a first-order phase transition, since the process has a jumplike character and the system exhibits instability of the magnetization, magnetostriction, and magnetoresistance, as manifested by the temporal variation of their isotherms and by a difference of the isotherms measured with increasing and decreasing the field strength (Figs. 3 and 7b).

In the compound with $x = 0.33$, the structure of which contains no charge-ordered clusters, the values of magnetization, magnetostriction, and magnetoresistance at T_C increase with the field in a smooth manner. In this material, the FM phase volume increases at the expense of a decreasing fraction of the *A*-type AFM phase. The AFM ordering of the *A*-type corresponds to a layered structure in which moments of the neighboring FM layers are oriented in the opposite directions. In the compound with $x = 0.33$, the FM part increases with the field smoothly, as the FM phase grows through the FM layers in the *A*-type AFM phase.

In the charge-ordered phase, spins of the neighboring Mn ions are oriented in the opposite directions. The results of our experiments showed that conversion of the charge-ordered clusters into the FM state requires applying threshold fields. As the threshold field strength is reached, such clusters completely transform into the FM state. This transition must be accompanied by modification of the crystal structure and, hence, manifested by a change in the anisotropic magnetostriction. Indeed, the data in Fig. 6c show a change in the sign of t in this temperature region, which is accompanied by the appearance of certain features in the $\lambda_t(T)$ curve: a maximum in the region of positive λ_t values and a minimum in the region of negative λ_t values. In contrast, the λ_t value in the compound with $x = 0.33$ is positive in the entire temperature range studied, including the region of T_C, and rapidly drops to zero for fields above T_C (Fig. 6b).

As was shown by Nagaev [1, 2] and by Yanase and Kasuya [19], the colossal magnetoresistance and large volume magnetostriction are explained by the magnetic-field-induced increase in volume of the FM phase (in which the charge carriers are concentrated) in a magnetically inhomogeneous sample and the resulting gain in the *s–d* exchange energy. According to this, the colossal magnetoresistance, large negative volume magnetostriction, and the excess thermal expansion near T_C observed in samples of the $Sm_{1-x}Sr_xMnO_3$ system with $x = 0.33$, 0.4, and 0.45 are explained by the phenomenon of electron phase separation caused by a strong *s–d* exchange, whereby the charge carriers are concentrated in the FM phase. The presence of charge-ordered clusters in the compounds with $x = 0.4$ and 0.45 brings certain special features in behavior of the magnetoresistance and volume magnetostriction, which is manifested by jumps in the corresponding isotherms related to the magnetic-field-induced transition of the charge-ordered clusters to the FM state.

It was experimentally established that the samples of $Sm_{1-x}Sr_xMnO_3$ with $x = 0.33$, 0.4, and 0.45 at $T > T_C$ exhibit, besides a colossal magnetoresistance of 83, 72, and 42%, respectively, a record level of volume magnetostriction reaching $\omega \sim 5 \times 10^{-4}$ in a field of 8.4 kOe and 10^{-3} in a field of $H = 50$ kOe. This is the second system of manganites after $La_{1-x}Sr_xMnO_3$ in which we observed a large volume magnetostriction accompanied by the colossal magnetoresistance in the same temperature region. These results open new possibilities for the application of manganites in magnetomechanical devices and sensors.

ACKNOWLEDGMENTS

The authors are grateful to M.R. Ibarra and C. Marquina for their help, attention, and hospitality during their stay at the University of Saragossa. Special thanks go to O.Yu. Gorbenko and A.R. Kaul' for the preparation and analysis of samples.

This study was supported by the Russian Foundation for Basic Research (project nos. 00-15-96695 and 00-02-17810), the INTAS Foundation (grant 97-open-30253), and NATO (grant HTECH LG 972942).

REFERENCES

1. É. L. Nagaev, Usp. Fiz. Nauk **166**, 833 (1996) [Phys. Usp. **39**, 781 (1996)].
2. E. L. Nagaev, Phys. Rep. **346**, 381 (2001).
3. E. Dagotto, T. Hotta, and A. Moreo, Phys. Rep. **344**, 1 (2001).
4. C. N. R. Rao, A. K. Cheetham, and R. Mahesh, Chem. Mater. **8**, 2421 (1996).
5. A. P. Ramirez, J. Phys.: Condens. Matter **9**, 8171 (1997).
6. A. Raveau, A. Maignan, C. Martin, and M. Hervieu, Chem. Mater. **10**, 2641 (1998).
7. L. P. Gor'kov, Usp. Fiz. Nauk **168**, 665 (1998) [Phys. Usp. **41**, 589 (1998)].
8. Y. Tokura and Y. Tamioka, J. Magn. Magn. Mater. **200**, 1 (1999).
9. J. M. D. Coey, M. Viret, and M. von Molnar, Adv. Phys. **48**, 167 (1999).
10. V. M. Loktev and Yu. G. Pogorelov, Fiz. Nizk. Temp. **26**, 231 (2000) [Low Temp. Phys. **26**, 171 (2000)].
11. Yu. A. Izyumov and Yu. N. Skryabin, Usp. Fiz. Nauk **171**, 121 (2001).

12. M. Yu. Kagan and K. I. Kugel', Usp. Fiz. Nauk **171**, 577 (2001).
13. L. I. Koroleva, R. V. Demin, and A. M. Balbashov, Pis'ma Zh. Éksp. Teor. Fiz. **65**, 449 (1997) [JETP Lett. **65**, 474 (1997)].
14. R. V. Demin, L. I. Koroleva, and A. M. Balbashov, Phys. Lett. A **231**, 279 (1997).
15. R. V. Demin, L. I. Koroleva, R. Szymszak, and H. Szymszak, Pis'ma Zh. Éksp. Teor. Fiz. **75**, 402 (2002) [JETP Lett. **75**, 331 (2002)].
16. A. Abramovich, L. Koroleva, A. Michurin, *et al.*, Physica B (Amsterdam) **293**, 38 (2000).
17. A. I. Abramovich, L. I. Koroleva, A. V. Michurin, *et al.*, Fiz. Tverd. Tela (St. Petersburg) **42**, 1451 (2000) [Phys. Solid State **42**, 1494 (2000)].
18. C. Margina, R. Ibarra, A. I. Abramovich, *et al.*, J. Magn. Magn. Mater. **226–230**, 999 (2001).
19. A. Yanase and T. Kasuya, J. Phys. Soc. Jpn. **25**, 1025 (1968).
20. V. V. Runov, D. Yu. Chernyshov, A. I. Kurbakov, *et al.*, Zh. Éksp. Teor. Fiz. **118**, 1174 (2000) [JETP **91**, 1017 (2000)].
21. C. Martin, A. Maignan, M. Hervieu, and B. Raveau, Phys. Rev. B **60**, 12191 (1999).
22. V. V. Runov, H. Glattli, G. V. Kopitsa, *et al.*, Pis'ma Zh. Éksp. Teor. Fiz. **69**, 323 (1999) [JETP Lett. **69**, 353 (1999)].
23. V. Runov, H. Glattli, G. Kopitsa, *et al.*, Physica B (Amsterdam) **276–278**, 795 (2000).
24. I. D. Luzyanin, V. A. Ryzhov, D. Yu. Chernyshov, *et al.*, Phys. Rev. B **64**, 094432 (2001).
25. K. P. Belov, L. I. Koroleva, M. A. Shalimova, *et al.*, Zh. Éksp. Teor. Fiz. **72**, 1994 (1977) [Sov. Phys. JETP **45**, 1047 (1977)].
26. J. Rodriguez-Carvajal, M. Hennion, F. Moussa, *et al.*, Phys. Rev. B **57**, R3189 (1998).
27. Y. Tomioka, H. Kuwahara, A. Asamitsu, *et al.*, Appl. Phys. Lett. **70**, 3609 (1997).

Translated by P. Pozdeev

Journal of Experimental and Theoretical Physics, Vol. 95, No. 5, 2002, pp. 927–939.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 122, No. 5, 2002, pp. 1074–1088.

SOLIDS
Electronic Properties

Intersubband Collective Excitations in a Quasi-Two-Dimensional Electron System in External Magnetic Field

L. V. Kulik[a, b, *], I. V. Kukushkin[a, b], V. E. Kirpichev[a, b], S. V. Tovstonog[a, b], V. E. Bisti[a], K. von Klitzing[b], and K. Eberl[b]

[a]*Institute of Solid State Physics, Russian Academy of Sciences, p/o Chernogolovka, Moscow oblast, 142432 Russia*
[b]*Max-Planck-Institut für Festkörperforschung, Stuttgart 70569, Germany*
**e-mail: kulik@issp.ac.ru*

Received May 21, 2002

Abstract—The spectrum of collective excitations in a quasi-two-dimensional electron system was studied by the method of Raman scattering spectroscopy. In an applied magnetic field, such systems exhibit collective excitations related to the electron transitions between dimensionally quantized subbands with a change in the Landau level index (intersubband Bernstein modes). It is shown that these modes interact with the fundamental intersubband excitations of the charge and spin densities, the interaction energy being determined by the excitation quasimomentum. Interaction of the intersubband Bernstein modes and the fundamental intersubband excitations with quasi-two-dimensional LO phonons was studied. Behavior of the new branches of collective excitations in a quasi-two-dimensional electron system possessing more than one occupied Landau level was studied and the nature of these branches was determined. *© 2002 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

Neutral excitations in low-dimensional electron systems have been extensively studied in the past decades by both theoretical and experimental methods [1, 2]. Special attention was devoted to quasi-two-dimensional electron systems (2DES), in which new basic phenomena were discovered including integer and fractional quantum Hall effects. Combining the properties of 2D and 3D systems, 2DESs possess complicated spectra of neutral excitations. Since the electron motion in 2DESs is confined (quantized) in one of the spatial directions, the excitation spectrum consists of intra- and intersubband branches. The intrasubband branches represent the usual 2D excitations related to the electron transitions within the same quantum confinement subband. We will concentrate on the intersubband excitations related to the electron transitions between different quantum confinement subbands.

The intersubband excitations have direct analogs in neither 2D nor 3D electron systems. The number of quantum confinement subbands and, hence, of various intersubband excitations in real 2DESs is large. For this reason, we will restrict our consideration to the fundamental branches of intersubband excitations related to the electron transitions from the ground to the first quantum confinement subband. Description of the other branches of intersubband excitations can be constructed by an analogous scheme. It will be assumed that the quantum electrical limit is realized, whereby the Fermi level of electrons measured from the ground subband energy level lies below the first excitation band bottom. In this case, the spectrum of neutral intersubband excitations comprises two collective modes, including the fundamental collective excitations of charge and spin densities, and the continuum of single-particle excitations (SPEs) [3–6]. The charge density excitations (CDEs) are related to the electron transitions with spin conservation, while the spin density excitations (SDEs) involve changes in the spin of an excited electron.

The fundamental CDEs and SDEs can be considered as the singlet and triplet states of the exciton formed by an electron in the excited subband and a hole under the Fermi level of electrons of the ground subband (Fig. 1). In contrast to the Mott exciton, the CDE and SDE energies are nondegenerate because the CDE energy includes the energy of macroscopic polarization of the electron system (depolarization shift). Since the polarization vector oscillates in the direction perpendicular to that of the CDE quasimomentum, the CDEs can be considered as an electron analog of TA phonons. Via the electric field of the macroscopic polarization, the fundamental CDEs can interact with quasi-two-dimensional LO phonons to form the bound CDE–LO modes [4].

In an external magnetic field perpendicular to the 2DES plane, the energy spectrum of electrons is completely quantized (to become quasi-zero-dimensional), which leads to a basic rearrangement of the spectrum of

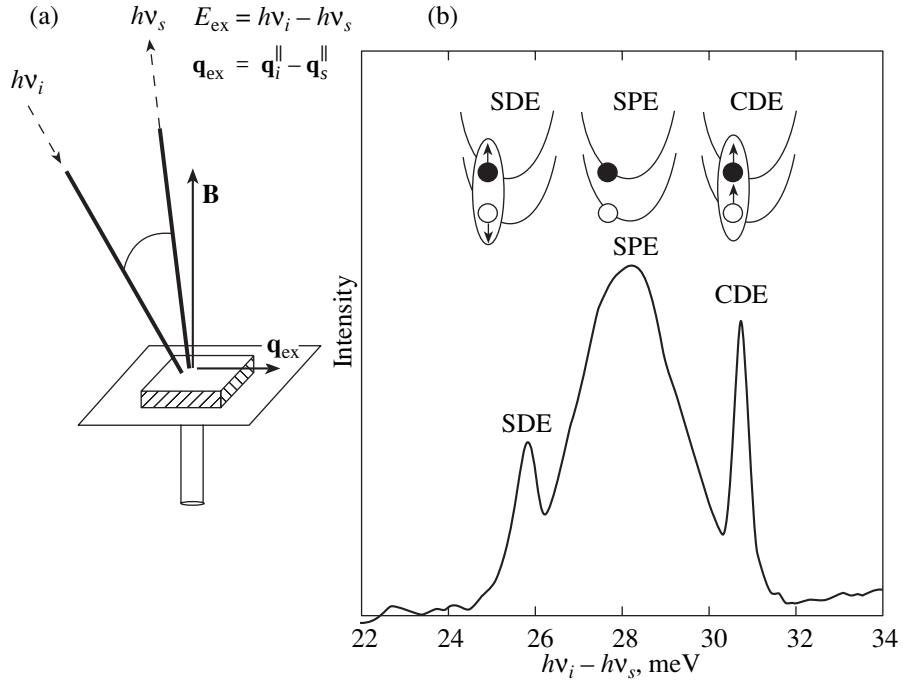


Fig. 1. Raman scattering in 2DES: (a) schematic diagram of the experimental geometry. The energy difference $h\nu_i - h\nu_s$ and the difference of projections of the incident (excitation) and scattered photon momenta onto the 2DES plane, $\mathbf{q}_i^{\parallel} - \mathbf{q}_s^{\parallel}$, are transferred to a 2DES excitation; (b) typical Raman spectrum of intersubband excitations in a 250-Å-wide QW with an electron density of 6.8×10^{11} cm^{-2}, measured in a zero magnetic field for $q = 0.4 \times 10^5$ cm^{-1}. Diagram at the top shows the possible types of intersubband excitations, including the charge density (CDE), spin density (SDE), and single-particle (SPE) excitations.

neutral excitations. All electron excitations in the magnetic field possess a collective character. As a result, a spectrum of collective excitations—intersubband Bernstein modes (ISBMs)—is formed instead of the continuum of single-particle excitations. Besides ISBM, the spectrum retains the fundamental excitations of charge and spin densities, now related to the intersubband transitions with conservation of the Landau level index.

The fundamental CDEs and SDEs in a magnetic field were observed by methods of IR absorption and Raman scattering spectroscopy [1]. However, direct observation of the ISBMs encountered experimental problems for a long time because the optical transitions related to ISBM excitation are dipole-forbidden both in the IR absorption and in the Raman scattering process (the Landau level index is not retained). Ando [7] showed that the problem of conservation of the Landau level index can be solved by applying an external magnetic field parallel to the quantum well plane, thus mixing the transverse (perpendicular to the 2DES plane) and longitudinal (parallel to the 2DES plane) motions of electrons [8–10]. It should be noted that such mixing significantly modifies the spectrum of intersubband magnetic excitations. Only ISBMs with a zero quasimomentum q in a tilted magnetic field admit a rather simple description [7], while nonzero quasimomenta make the description quite difficult [11].

This paper presents the results of a presently most complete experimental investigation of the energy spectrum and dispersion of intersubband excitations in a perpendicular magnetic field, summarizing the results obtained by the authors within the last five-year period. The presentation proceeds as follows. Section 2 gives a description of the original experimental approach developed by the authors for measuring the Raman spectra. This method allows the Raman spectra to be measured under conditions of extremely low temperatures and superstrong magnetic fields and ensures long-term stability of the informative signal in the magnetic field, not achievable with the other techniques. Section 3 considers the spectrum of intersubband excitations in the region of small quasimomenta ($q \longrightarrow 0$) and presents experimental evidence of the basic energy relations between ISBMs (intersubband analog of the Kohn theorem) [1, 12]. Section 4 addresses the interaction of ISBMs with the fundamental intersubband CDEs and SDEs [12]. Section 5 is devoted to the interaction of the intersubband excitations of a 2DES with the phonon subsystem of a semiconductor on which the 2DES is based. It will be shown that the fundamental CDEs in polar semiconductors interact with LO phonons, while ISBMs interact with the CDE–LO hybrid modes [13]. Section 6 discusses the new branches of intersubband

collective excitations considered as antiphase oscillations of the electron subsystems of 2DES on various Landau levels [14].

2. EXPERIMENTAL METHOD

The experiments were performed with a series of high-quality MBE-grown heterostructures representing asymmetric, selectively doped separate 250-Å-wide $Al_xGa_{1-x}As/GaAs$ quantum wells (QWs). The electron density n_s in the samples varied within (0.5–6.8) $\times$ 10^{11} cm^{-2} and the electron mobility μ, within (1–7) $\times$ 10^6 cm^2/(V s). The Raman spectral line width was 0.15 meV.

The measurements were conducted at 1.5 K in a cryostat with superconducting solenoid, with the magnetic field strengths varied from 0 to 11 T in the Faraday geometry. The optical excitation of the electron system was effected by radiation of a tunable titanium–sapphire laser at a photon energy exceeding the GaAs band gap width (E_0) and an excitation power density of 0.1–1 W/cm^2. Experiments employed an original double-fiber technique free of significant disadvantages inherent in the standard setups with optical windows, such as contamination of the optical tract and detuning of the optical system by a magnetic field sweep. The new technique ensured long-term stability of the signal of Raman scattering measured in the presence of a magnetic field. The first optic fiber was used to excite the electron system, and the second fiber detected the Raman scattering signal (Fig. 1a). The detecting optic fiber served as an effective *in situ* premonochromator, filtering off a large fraction of the laser radiation reflected from the sample surface and all of the signal of the intrinsic inelastic scattering in the exciting fiber. The quasimomentum transfer to 2DES in the course of the Raman scattering was determined by the arrangement of fibers relative to the sample surface, the maximum momentum transfer reaching 1.2×10^5 cm^{-1}.

In order to distinguish between the resonances related to CDEs and SDEs, the light scattering spectra were measured for both parallel and perpendicular configurations of the linear polarization vectors of the exciting and scattered photons. In the parallel configuration, the incident and scattered photons had the same polarization vectors and, hence, the spectrum displayed the lines due to excitations without spin rotation. In the perpendicular configuration, the polarization vectors of the incident and scattered photons were perpendicular and, hence, the spectrum contained the lines of excitations with spin rotation [15]. The polarization was analyzed by means of linear polarizers placed in liquid helium between the sample and fiber edges. The Raman signal passed through a U-1000 double monochromator and was detected by a CCD camera. The spectral resolution of the measuring system was 0.03 meV. Simultaneous measurement of the Raman and luminescence spectra allowed the electron density to be monitored under quasi-continuous photoexcitation conditions [16].

3. INTERSUBBAND BERNSTEIN MODES

Figure 1b presents a typical Raman spectrum scattering in a zero magnetic field for a QW with an electron density of $n_s = 6.8 \times 10^{11}$ cm^{-2}, measured for $q = 0.4 \times 10^5$ cm^{-1} and $E_L = 1.587$ eV. The spectrum consists of two narrow lines denoted by SDE and CDE, peaked at the energies of 25.7 and 30.7 meV, and a broad band (denoted SPE) situated in between the former two. The CDE and SDE lines are observed only in the parallel and perpendicular polarization configurations of the exciting and scattered photons, respectively. Thus, the CDE line corresponds to the fundamental excitation of the charge density, while the SDE line corresponds to that of the spin density. The broad SPE band is related to the continuum of single-particle excitations [4]. This band is present in the spectra obtained in both parallel and perpendicular polarization configurations (single-particle excitations can either involve or not the spin rotation). In the long-wave limit, the SPE band shifts away from the SDE peak to higher energies by a value equal to the intersubband exciton energy, while the CDE peak binding energy increases by an amount equal to the depolarization shift.

In a magnetic field perpendicular to the QW plane, the SPE band exhibits splitting into individual components. These spectra are presented in Fig. 2, where the SPE band components are denoted by $ISBM_{+1}$, $ISBM_{-1}$, $ISBM_{-2}$, $ISBM_{-3}$, and L_0. As the magnetic field strength is increased, the $ISBM_{+n}$ ($ISBM_{-n}$) components shift upward (downward) relative to the SPE band maximum. The energy resonance of the $ISBM_{+1}$ ($ISBM_{-1}$) and CDE (SDE) lines is manifested by the interaction (anticrossing) of levels (Fig. 2). In a strong magnetic field, the lines of the fundamental intersubband excitations (SDE and CDE) and the intermediate line L_0 become dominating in the Raman spectrum.

Using the experimental spectra, we determined dependences of the line energies on the applied magnetic field strength (Fig. 3). As can be seen, the CDE and SDE energies are independent of the magnetic field because these excitations are related to the intersubband electron transitions between the Landau levels with the same indices. With neglect of the nonparabolic shape of the conduction band of GaAs in the energy interval of the intersubband level splitting, the energies of all such transitions are equal and independent of the magnetic field strength. On the contrary, the energies of $ISBM_{\pm n}$ lines form the negative and positive "fans" of the Landau levels, originating from the energy of intersubband quantization. The mass determined from the slopes of these Landau levels is $0.071m_0$, where m_0 is

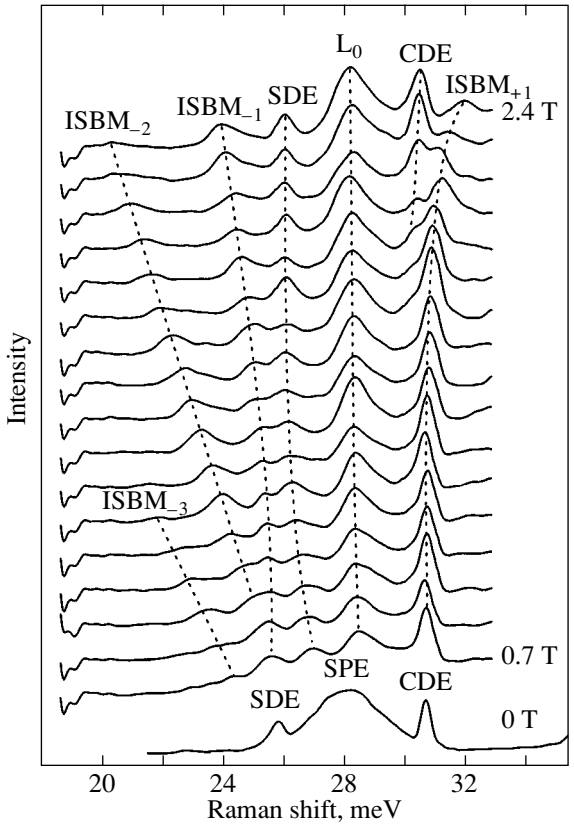


Fig. 2. Raman spectra of a 250-Å-wide QW with an electron density of 6.8×10^{11} cm^{-2}, measured for $q = 0.4 \times 10^5$ cm^{-1} and $E_L = 1.587$ eV in a range of magnetic fields from 0.7 to 2.4 T with an 0.1-T step, in comparison to the spectrum measured in a zero magnetic field.

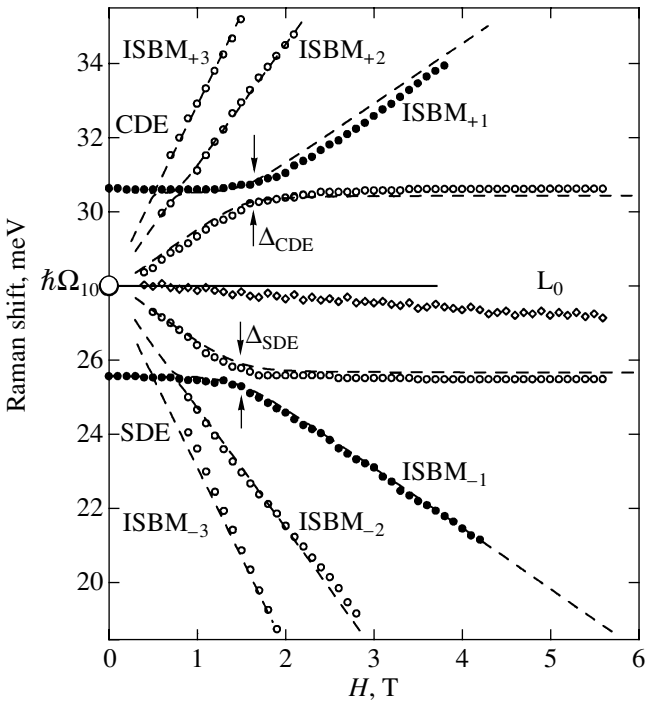


Fig. 3. The energies of intersubband magnetic excitations in a 250-Å-wide QW with an electron density of 6.8×10^{11} cm^{-2}. Points represent the experimental data; dashed curves show the results of calculations. For the sake of clarity, only the theoretical curves for CDEs are shown above the $\hbar\Omega_{10}$ level and only the curves for SDEs are depicted below this level. The SPE continuum is indicated by a large open circle.

the free electron mass. This estimate virtually coincides with the effective electron mass for GaAs ($m^* = 0.067m_0$). Thus, we may conclude that the $ISBM_{+1}$ and $ISBM_{-1}$ lines are related to the intersubband Bernstein modes (i.e., the intersubband electron transitions between the Landau levels with different numbers). The experimental energies of the intersubband Bernstein modes for $q \longrightarrow 0$ are well described by the relation

$$E_{B\pm n} = |\hbar\Omega_{10} \pm n\hbar\omega_c|, \quad n \neq 0, \tag{1}$$

where Ω_{10} and $\omega_c = eB/m^*c$ are the electron intersubband and cyclotron frequencies, respectively. Formula (1) for 2DES is an analog of the Kohn theorem for the intersubband excitations, according to which the energies of the intersubband Bernstein modes for zero quasimomentum transfer are independent of the electron–electron interaction [1]. A remarkable fact is that expression (1) contains no terms reflecting the features (shape, height, and width) of a potential bounding the 2DES, the intersubband energy $\hbar\Omega_{10}$ being the only quantity characterizing the transverse motion of electrons.

4. INTERACTION OF INTERSUBBAND BERNSTEIN MODES WITH FUNDAMENTAL CDE AND SDE MODES

Equation (1) describes the experimental results corresponding to small quasimomentum transfer ($q \longrightarrow 0$). However, as the quasimomentum increases, the $ISBM_{+n}$ ($ISBM_{-n}$) and CDE (SDE) mode energies deviate from linear relations in the region of mode resonances and formula (1) becomes inapplicable to description of the ISBM energies (Fig. 4). This situation is analogous to that with the intersubband excitations: while the energies of intersubband excitations and intersubband Bernstein modes for $q = 0$ are determined only by the effective mass of free electrons (Kohn theorem and the intersubband analog), the dispersion relations are determined by the many-body Coulomb interactions. Therefore, in order to describe the ISBMs in the case of nonzero quasimomenta, it is necessary to modify Eq. (1) so as to take into account the interparticle Coulomb interactions.

Theoretical calculations of the intersubband excitations with an allowance for the electron–electron interaction were performed within the framework of a local

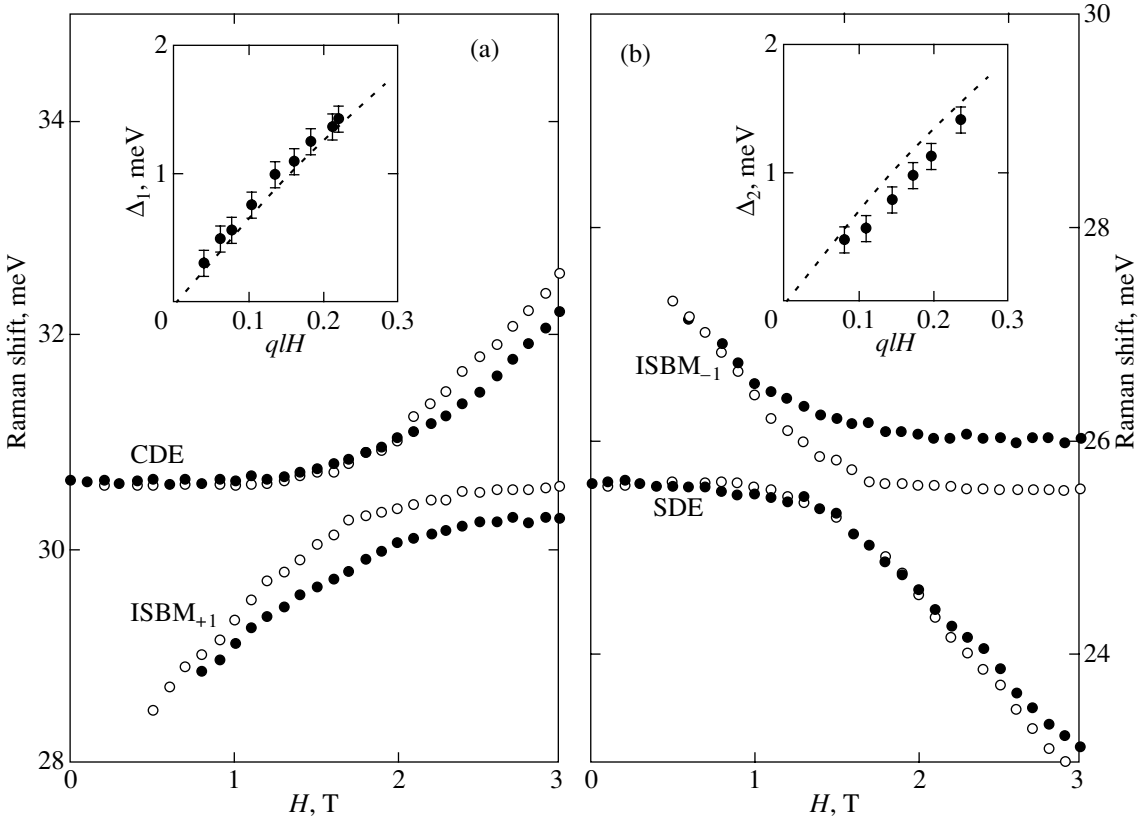


Fig. 4. Hybrid mode energies in the region of energy resonance of (a) ISBM$_{+1}$ and CDE and (b) ISBM$_{-1}$ and SDE modes for the quasimomentum values 0.4×10^5 and 0.8×10^5 cm^{-1}. The insets show the experimental (points) and theoretical (dashed lines) dispersion of the energy gaps Δ_1 and Δ_2.

density approximation (TDLDA model) [3, 7, 17]. According to this approximation, the energies of collective excitations correspond to the poles of a total polarization function (below, the indices refer to excitations of the charge (CD) and spin (SD) densities). For a perpendicular magnetic field, it is possible to neglect the interactions between intra- and intersubband excitations [18]. In this case, the energies of collective intersubband excitations correspond to the poles of the intersubband part of the total polarization function

$$\tilde{\chi}_i^{10}(q, \omega) = \frac{\chi^{10}(q, \omega)}{1 - \gamma_i(q)\chi^{10}(q, \omega)}, \qquad (2)$$

where $\chi^{10}(q, \omega)$ is the intersubband polarization function of the 2DES in the magnetic field without interaction [18]:

$$\chi^{10}(q, \omega) = \frac{2}{\hbar\pi l_B}\sum_{N=0}^{\infty}\left(\sum_{S=0}^{\infty} n_F^N |J_{N+S,N}(q)|^2 \times \frac{\Omega_{10} + S\omega_c}{(\omega + i\delta)^2 - (\Omega_{10} + S\omega_c)^2} + \sum_{S=1}^{\infty} n_F^{N+S} |J_{N+S,N}(q)|^2 \frac{\Omega_{10} - S\omega_c}{(\omega + i\delta)^2 - (\Omega_{10} - S\omega_c)^2}\right), \qquad (3)$$

$$n_F^N = \begin{cases} 1, & N \le \tilde{N}, \\ 0, & N > \tilde{N}, \end{cases}$$

and $\tilde{N} = \pi n_s \hbar/m^*\omega_c - 1$ is the index of the highest occupied Landau level. The square of the matrix element is expressed as

$$|J_{N,N'}|^2 = \frac{N_2!}{N_1!}\left[\frac{(l_B q)^2}{2}\right]^{N_1 - N_2} \times \exp\left[-\frac{(l_B q)^2}{2}\right]\left[L_{N_2}^{N_1 - N_2}\left(\frac{(l_B q)^2}{2}\right)\right]^2,$$

where $N_1 = \max(N, N')$, $N_2 = \min(N, N')$, $L_N^{N'}(x)$ are the associated Laguerre polynomials, and $l_B = \sqrt{\hbar c/eB}$ is the magnetic length. The parameter γ_i is determined by the depolarization shift and the exchange correlation energy β_i:

$$\gamma_{CD}(q) = V(q)/\epsilon(\omega) - \beta_{CD},$$

$$\gamma_{SD}(q) = -\beta_{SD}.$$

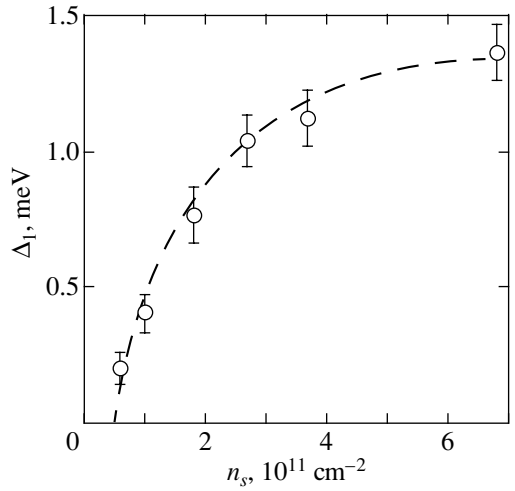


Fig. 5. The plot of energy gap Δ_1 versus electron density n_s in a 250-Å-wide QW, measured for a fixed quasimomentum of 1.1×10^5 cm^{-1}. The CDE and ISBM$_{+1}$ modes do not intersect at $n_s \le 5 \times 10^{10}$ cm^{-2}. Points represent the experimental data; the dashed line shows the results of theoretical calculations.

The matrix element of the Coulomb interaction entering into an expression for the depolarization shift is

$$V(q) = \frac{2\pi e^2}{\epsilon_\infty q}\int dz_1 \int dz_2 \psi_0(z_1)\psi_1(z_1) \times \exp(-q|z_1 - z_2|)\psi_0(z_2)\psi_1(z_2), \quad (4)$$

while the dynamic screening of the Coulomb interaction by the optical LO phonons is included in the frequency-dependent dielectric function

$$\epsilon(\omega) = \frac{\omega^2 - \omega_{\mathrm{LO}}^2}{\omega^2 - \omega_{\mathrm{TO}}^2}, \quad (5)$$

where ω_{LO} and ω_{TO} are the corresponding phonon frequencies and $\epsilon_\infty \approx 12.86$ is the permittivity of GaAs.

The exchange correlation energy in the local density approximation is

$$\beta_i = \int dz \psi_0^2(z) U_i(z) \psi_1^2(z),$$
$$U_i(z) = \left.\frac{\partial V_i}{\partial \rho_i}\right|_{\rho_{CD} = n_s\psi_0^2(z),\, \rho_{SD} = 0}, \quad (6)$$

where V_i is the exchange correlation potential and ρ_i is either the 3D electron density $\rho_{\mathrm{CD}} = \rho_\uparrow + \rho_\downarrow$ or the spin density $\rho_{\mathrm{SD}} = \rho_\uparrow - \rho_\downarrow$ ($\rho_\uparrow$ and $\rho_\downarrow$ are the densities of the spin subsystems of the 2DES). Assuming that the temperature is much higher than the Zeeman energy and both spin subsystems are equally occupied, and selecting V_i as in [19], we arrive at the following expressions:

$$U_{CD}(z) = -1.706 a_B^3 r_s^2 \left(1 + \frac{0.6213 r_s}{11.4 + r_s}\right)\mathrm{Ry}^{\mathrm{GaAs}},$$
$$U_{SD}(z) = -1.706 a_B^3 r_s^2 \times \left(1 - 0.036 r_s + \frac{1.36 r_s}{1 + 10 r_s}\right)\mathrm{Ry}^{\mathrm{GaAs}}, \quad (7)$$

where

$$r_s(z) = (4\pi a_B^3 n(z)/3)^{-1/3}, \quad n(z) = n_s\psi_0^2(z),$$

$a_B \approx 100$ Å is the Bohr electron radius for GaAs, and $\mathrm{Ry}^{\mathrm{GaAs}} \approx 5.71$ meV. The first like terms in the expressions for $U_i(z)$ reflect the exchange interaction, while the other terms describe the electron correlations.

The electron wave functions in the fundamental and excited quantum confinement subbands, $\psi_0(z)$ and $\psi_1(z)$, which are necessary for determining the depolarization shift and the exchange correlation energy, were obtained as self-consistent solutions to the system of one-dimensional Kohn–Sham and Poisson equations [20]. Restricting the calculation to terms on the order of $(ql_B)^6$ (that is, ignoring the electron transitions with $|\Delta n| > 3$), we obtain the curves plotted in Fig. 3. The experimental differences of the ISBM$_{+n}$ (ISBM$_{-n}$) and CDE (SDE) energies in the region of mode resonances are in a reasonable agreement with the results of numerical calculations (Fig. 3). Thus, we may conclude that the observed anticrossing of ISBM$_{+n}$ (ISBM$_{-n}$) and CDE (SDE) modes are manifestations of the many-body Coulomb interaction, which mixes the collective excitations of the 2DES related to the intersubband electron transitions with and without a change of the Landau level index.

By varying the excitation quasimomentum and the electron density, we studied dependence of the energies of hybrid modes on the magnitude of the Coulomb interaction (both Coulomb terms, representing the depolarization shift and the exchange correlation energy, are monotonic functions of q and n_s) [6]. Figure 4 shows plots of the Raman shift of the hybrid modes versus magnetic field strength in the region of mode resonances for a fixed electron density $n_s = 6.8 \times 10^{11}$ cm^{-2} and two values of the quasimomentum. As can be seen, the energy gap Δ_1 (Δ_2) between the hybrid modes increases when q doubles from 0.4×10^5 to 0.8×10^5 cm^{-1}; the Δ_1 (Δ_2) energy is a linear function of q in the range of experimentally accessible quasimomenta ($ql_B \ll 1$—see Fig. 4).

Figure 5 shows dependence of the Δ_1 value on the electron density for a fixed quasimomentum $q = 1.1 \times 10^5$ cm^{-1}. In the region of small n_s ($< 2 \times 10^{11}$ cm^{-2}), the gap Δ_1 is described by a linear function of the electron

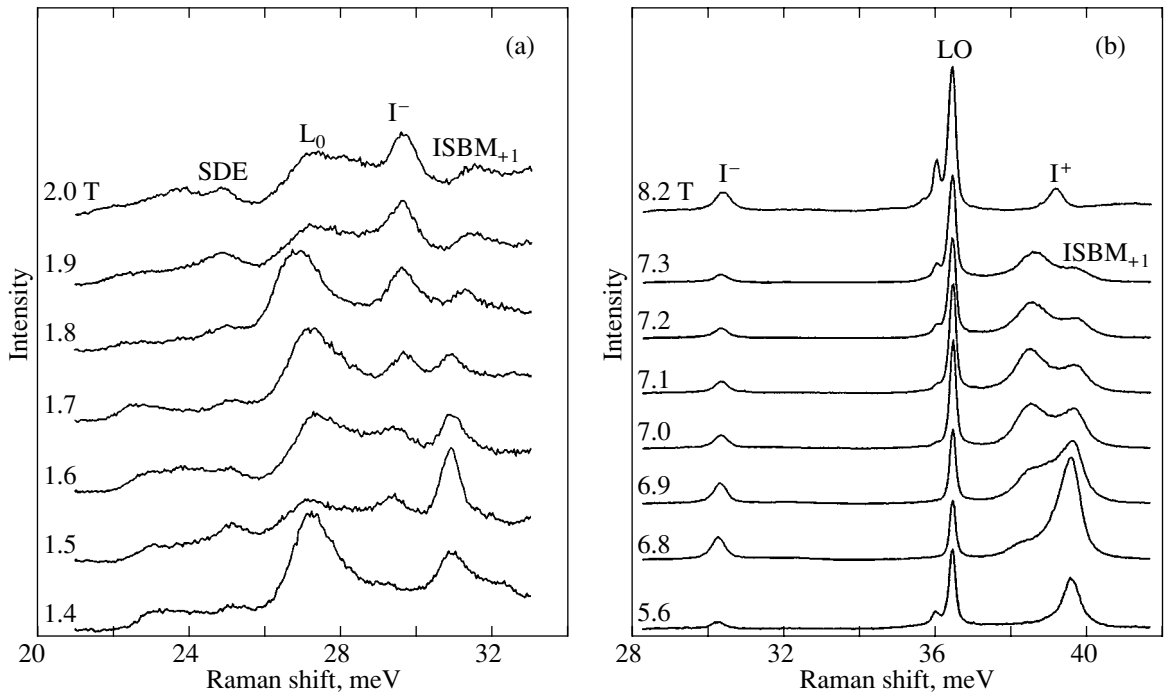


Fig. 6. Raman spectra of a 250-Å-wide QW with an electron density of 6.8×10^{11} cm^{-2}, measured in the region of energy resonances of (a) $ISBM_{+1}$ and I^- and (b) $ISBM_{+1}$ and I^+ modes for $q = 1.1 \times 10^5$ cm^{-1}, $E_L = 1.59$ eV, and various magnetic field strengths (indicated at the curves). LO is the line of a bulk LO phonon in GaAs.

density intersecting the abscissa axis at a certain finite critical value $n_s^c \approx 4 \times 10^{10}$ cm^{-2}, rather than at $n_s \longrightarrow 0$. The critical electron density is determined by equality of the depolarization shift and the exchange correlation energy (these quantities depend differently on the electron density [6]). At an electron density below the critical value, the energy of the CDE mode is lower than the intersubband quantization energy $\hbar\Omega_{10}$; that is, the energy balance between $ISBM_{+1}$ and CDE modes becomes impossible. It should be also noted that the Δ_1 value exhibits saturation at $n_s \geq 2 \times 10^{11}$ cm^{-2}, which is related to a mixing of the $ISBM_{+1}$ mode with the CDE–LO mode. This effect is considered in detail in the following section. If the interaction with the LO phonons is ignored, dependence of the Δ_1 value on the electron density must be close to linear.

Thus, we have studied the interaction between intersubband Bernstein modes, involving a change in the Landau level index by unity, and the fundamental CDEs and SDEs, but did not consider the behavior of ISBMs with large indices. The energies of such modes are described by Eq. (1), and their interaction with the fundamental modes in the long-wave limit is small (see Fig. 3). For example, the interaction of CDE and $ISBM_{+2}$ modes becomes detectable only at $q = 1.2 \times 10^5$ cm^{-1}, which is close to the upper limit of the experimentally accessible interval of quasimomenta. The calculation shows that the energy of the interaction between CDE (SDE) and $ISBM_{+n}$ ($ISBM_{-n}$) is proportional to $(ql_B)^n$ (for $ql_B \ll 1$).

5. INTERACTION OF FUNDAMENTAL AND BERNSTEIN MODES WITH LO PHONONS

The effect of LO phonons on the spectrum of intersubband electron excitations was studied using a series of samples in which the intersubband quantization energy $\hbar\Omega_{10}$ was close to the energy of LO phonons in GaAs. Figure 6 shows the Raman spectra for a single QW with $n_s = 6.8 \times 10^{11}$ cm^{-2}, measured for $q = 1.1 \times 10^5$ cm^{-1} in two spectral intervals: below (a) and above (b) the bulk LO phonon energy in GaAs. The QW width and the electron density were selected so as to ensure that the CDE energy almost coincided with LO phonon energy in GaAs. In this case, the macroscopic polarization field of CDE, oscillating at the LO phonon frequency, mixes the CDE mode with the LO phonon mode and the spectrum displays two CDE–LO hybrid modes denoted by I^- and I^+.

Application of a perpendicular magnetic field leads to the appearance of ISBMs in the Raman spectrum. Figure 7b shows plots of the Raman shifts for the $ISBM_{+1}$, two hybrid CDE–LO, and the bulk LO phonon (reference) modes versus magnetic field strength for GaAs. For the sake of clarity, the other spectral features (ISBM modes with $n \neq +1$ and SDE modes) are omitted. In the region of energy resonances of the $ISBM_{+1}$

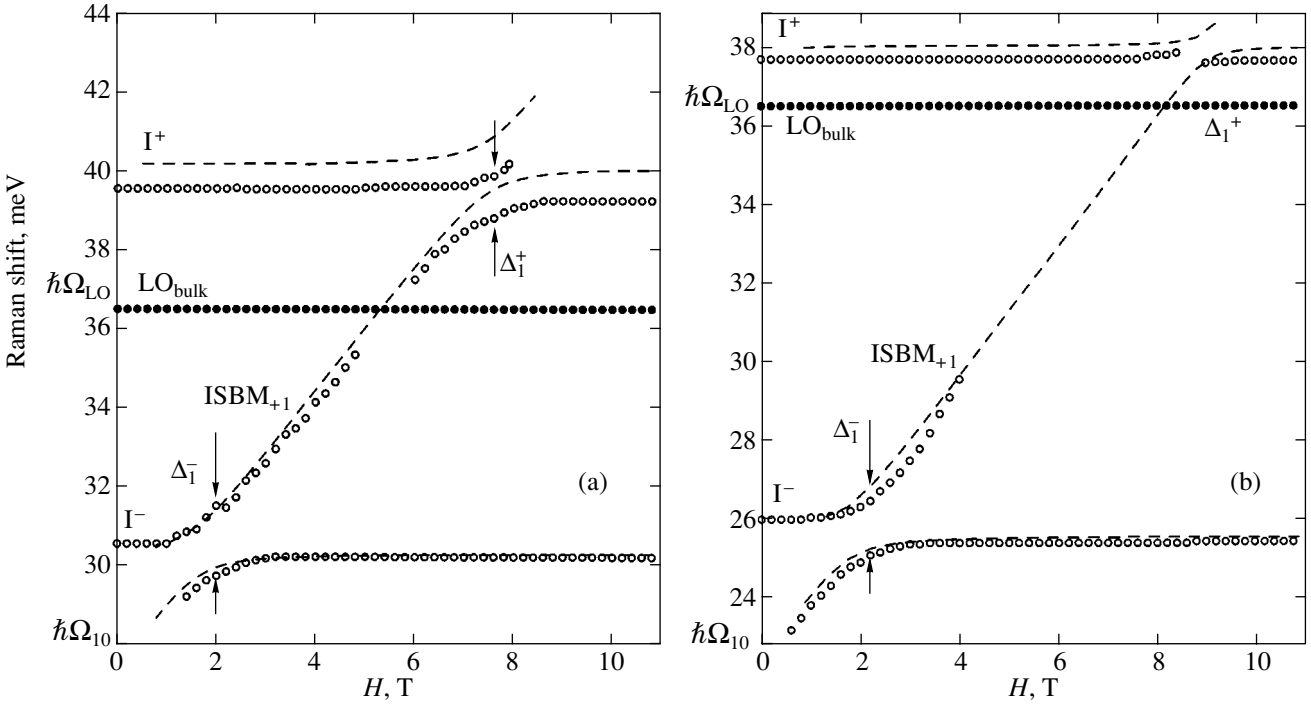


Fig. 7. The energies of intersubband magnetic excitations in a 250-Å-wide QW with an electron density of (a) 6.8×10^{11} and (b) 3.8×10^{11} cm^{-2}. Open circles represent the experimental data; dashed curves show the results of theoretical calculations; black circles show the reference energy level corresponding to a bulk LO phonon.

mode with the I$^-$ and I$^+$ modes, the spectrum displays anticrossing, which is evidence of the interaction between the ISBM$_{+1}$ mode with each of the hybrid CDE–LO modes. Thus, we may conclude that the 2DES under consideration features the formation of triple ISBM–CDE–LO modes. However, it is possible to show that, in the region of accessible quasimomenta ($ql_B \ll 1$), the ISBM$_{+1}$ mode interacts only with the CDE component of the CDE–LO hybrid mode, rather than with the LO phonons.

We have studied the energy of interaction between the ISBM$_{+1}$ mode and the I$^-$ and I$^+$ modes by varying the degree of the CDE mode mixing with the LO phonon mode (Fig. 7). This was achieved by decreasing the electron density in the QW (and, hence, the intersubband energy $\hbar\Omega_{10}$) and by driving the CDE mode out of the resonance with the LO phonon. Figure 7 shows plots of the Raman shift versus the magnetic field strength for two samples with the same QW widths but different electron densities (6.8×10^{11} cm^{-2} versus 3.8×10^{11} cm^{-2}). As can be seen from these data, the gap Δ_1^+ between ISBM$_{+1}$ and LO-like modes (I$^+$) decreases with the electron density (i.e., with the LO-like mode energy approaching the bulk LO phonon energy in GaAs). After the further decrease in the charge carrier density (from 3.8×10^{11} to 2.7×10^{11} cm^{-2}), the Δ_1^+ value becomes zero. At the same time, the energy gap Δ_1^- between the ISBM$_{+1}$ and CDE-like mode (I$^-$) remains almost unchanged. This behavior of Δ_1^+ and Δ_1^- indicates that the ISBM$_{+1}$ mode interacts only with the CDE component of the CDE–LO hybrid mode, while interaction with the LO phonon component in the accessible interval of quasimomenta is negligibly small.

The experimental spectra were compared to the results of numerical calculations performed within the framework of the local density approximation (see above) in the long-wave limit ($\chi^{10}(q,\omega) \sim (ql_B)^2$) with neglect of the transitions with $|n| > 1$. As can be seen from Fig. 7, the theoretical spectrum is in quite good agreement with experiment in the entire range of magnetic fields.

As was demonstrated in Section 4, the energy gap between the CDE–ISBM$_{+1}$ hybrid modes is a linear function of the quasimomentum for $ql_B \ll 1$. We have also measured the dispersion curves for Δ_1^- and Δ_1^+ (Fig. 8). The plots of $\Delta_1^-(q)$ and $\Delta_1^+(q)$ are close to linear, which is consistent with the theoretical calculation using the local density approximation. An important result is that both energy gaps, Δ_1^- and Δ_1^+, vanish when $q \longrightarrow 0$, which implies that the ISBM energy at $q = 0$ is independent of the degree of the electron–electron interaction screening by LO phonons. This experimen-

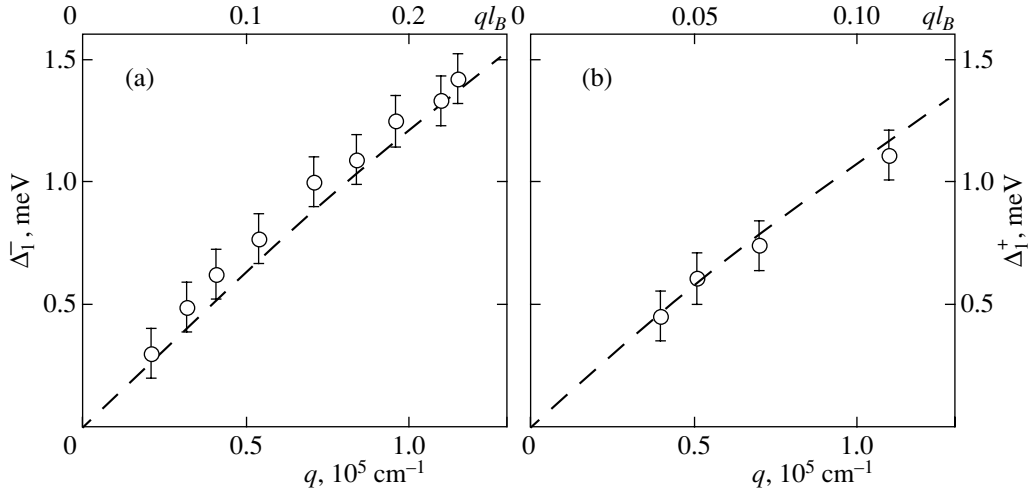


Fig. 8. Dispersion of the energy gaps (a) Δ_1^- and (b) Δ_1^+ in a 250-Å-wide QW with an electron density of 6.8×10^{11} cm^{-2}. Open circles present the experimental data; dashed curves show the results of theoretical calculations.

tal observation indicates that the basic relation (1) remains valid in polar semiconductors.

6. ANTIPHASE INTERSUBBAND MODES

As was shown above, the Raman spectra exhibit, besides resonances related to the ISBM modes and fundamental intersubband excitations, an additional resonance manifested by the L_0 line (see Fig. 3). When $B \longrightarrow 0$, the L_0 line merges with the SPE band reflecting continuum of the single-particle excitations. Based on this experimental observation, Brozak *et al.* [10] suggested that the L_0 resonance is also related to the single-particle excitations. Another argument in favor of the single-particle nature of this line was that no collective modes occurring between the fundamental intersubband excitations (CDE and SDE) can take place according to the local density approximation.

However, we presented arguments against assignment of the L_0 line to the single-particle excitations [12]. Strictly speaking, no single-particle excitations can exist in a translation-invariant 2DES occurring in a perpendicular magnetic field: an excited electron and a hole under the Fermi level always form a bound state (magnetoexciton) [21, 22]. For large quasimomenta, the binding energies of magnetoexcitons are small and they can be destroyed by any small perturbation. Magnetoexcitons with large quasimomenta can be considered as analogous to single-particle excitations. However, in order to detect such excitations at small quasimomenta such as those accessible in the standard experiments on the Raman scattering ($ql_B \ll 1$), it is required that the excitations be strongly scattered from impurities and roughnesses of the 2DES heteroboundaries. This implies that the Raman scattering cross section for the "single-particle" excitations must depend on the sample quality. However, no such dependence was observed [12]. The absence of collective modes corresponding to the L_0 resonance in the local density approximation can be explained by the fact that this model only roughly takes into account the exchange interaction (the exchange energy is assumed to depend only on the total local electron density in the QW). In order to elucidate the nature of the L_0 peak, we studied the Raman spectra of a series of high-quality heterostructures with single QWs. The results of these measurements showed that the L_0 resonance is related to the collective intersubband excitations in the 2DES, representing the antiphase or optical branches of CDEs and SDEs [14].

Figure 9 shows the Raman spectra of three QWs possessing different electron configurations. As can be seen, all spectra display three dominating lines: L_0 and two others corresponding to the fundamental intersubband (SDE and CDE) modes. In weak magnetic fields, the L_0 mode energy is close to the intersubband energy $\hbar\Omega_{10}$; as the magnetic field strength grows, the L_0 mode energy drops. Under the mode resonance conditions, the L_0 and SDE lines exhibit repulsion (Fig. 10). This behavior of lines is indicative of the interaction between the SDE mode and the excitation related to the L_0 resonance. Since modes of different symmetries cannot interact, we may conclude that the L_0 line is due to a collective 2DES excitation possessing the same quasimomentum as that of SDE. Therefore, the L_0 line cannot be related to a process in which the quasimomentum is not conserved.

Interpretation of the L_0 line in terms of the single-particle excitations is also inconsistent with the dependence of the mode energy on the magnetic field strength. As can be seen from Fig. 10, the L_0 mode energy significantly decreases when the field strength grows. However, it is known that the intersubband

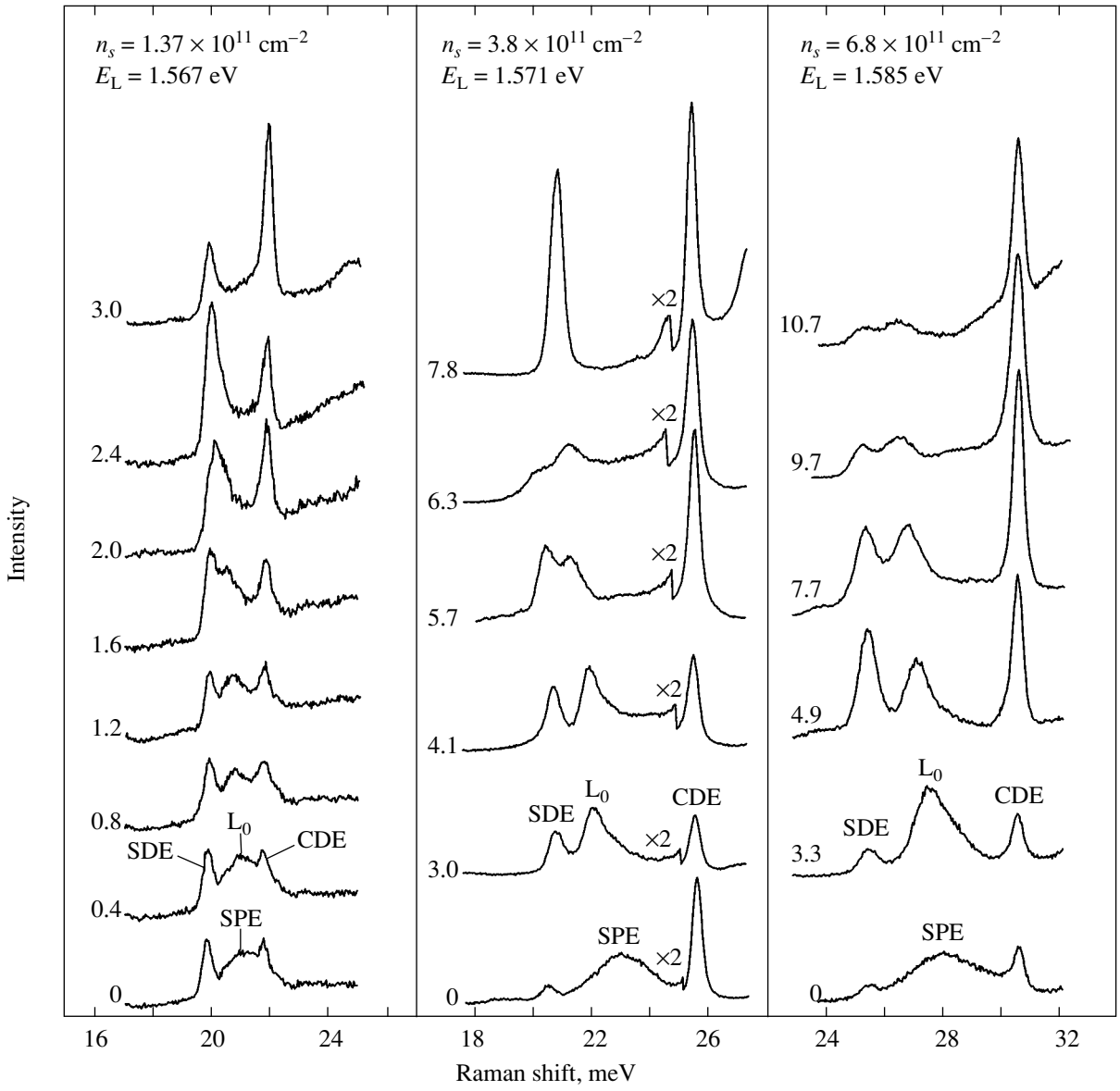


Fig. 9. Raman spectra of a 250-Å-wide QW with different electron densities, measured using various laser excitation energies (indicated at the top) in various magnetic fields (indicated at the curves).

energy and, hence, the single-particle excitation energy are independent of the field. Nor can we relate a shift of the L_0 mode energy in the magnetic field to the cyclotron energy, since the tangent slope α of the relation between the L_0 mode energy and the magnetic field strength depends on the electron density of the 2DES (Fig. 11): the value of α decreases by more than half when the electron density grows from 1.37×10^{11} to 6.8×10^{11} cm^{-2}. Nevertheless, there exists an invariant quantity related to the L_0 line shift that is insensitive to parameters of the sample studied. Such a quantity, invariant with respect to the electron density, is the critical electron occupancy factor ν_{cr} for which the L_0 mode energy becomes equal to the SDE mode energy (Fig. 11).

A theoretical description of the collective excitation modes corresponding to the L_0 line was developed within the framework of the Hartree–Fock approximation in which the exchange interaction depends not only on the total electron density in the QW (as in the local density approximation), but on the distribution of electrons over the Landau levels as well [23]. Calculations performed with neglect of the electron transitions changing the Landau level index predict that there are $2N$ collective intersubband modes (N is the index of the highest occupied Landau level) combined into pairs, each pair including excitations of the charge and spin densities. The energy of a pair of modes is independent of the magnetic field strength. These are the equiphase or acoustic modes corresponding to the fundamental

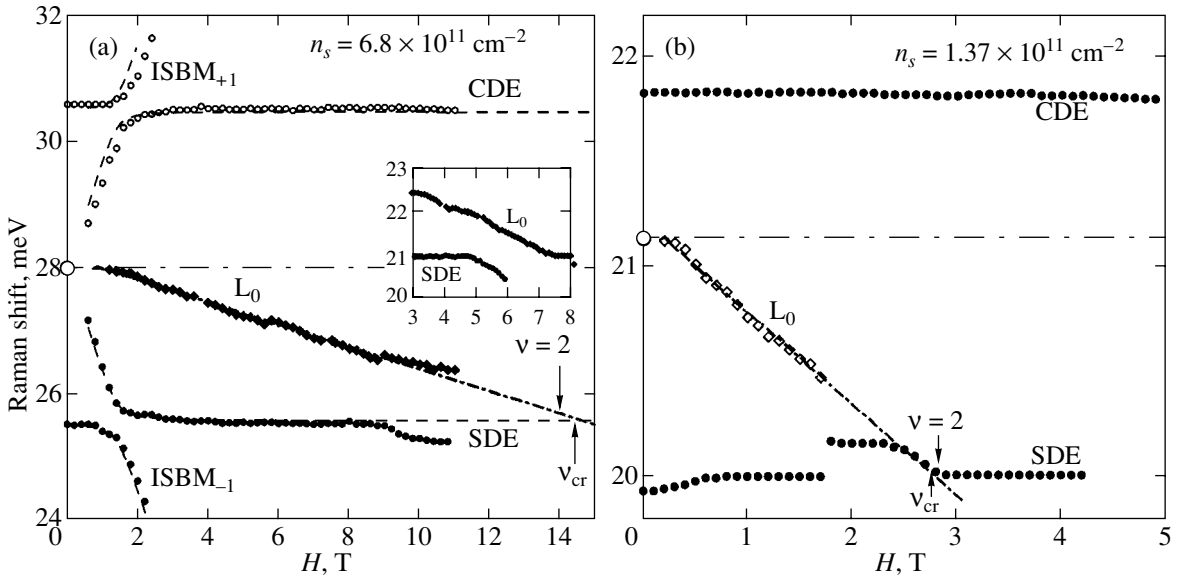


Fig. 10. The energies of intersubband magnetic excitations in a 250-Å-wide QW with an electron density of (a) 6.8×10^{11} and (b) 1.37×10^{11} cm^{-2}, measured for $q = 0.4 \times 10^5$ cm^{-1}. Dash-and-dot lines show the intersubband splitting and L_0 resonance energies in a linear approximation. Dashed lines represent the calculation in the local density approximation ($\chi^{10}(q, \omega) \sim (ql_B)^2$). The inset illustrates repulsion between the L_0 and SDE lines.

CDEs and SDEs. The energies of all other (antiphase or optical) modes are close to the intersubband energy at $B \longrightarrow 0$ and decrease with increasing magnetic field strength. The acoustic CDE and SDE modes are split by a value equal to the depolarization shift, while the energies of optical excitations coincide in each pair and differ but little between pairs [23].

The above excitations are, in a certain sense, analogous to the phonons in crystals. In the case when a crystal unit cell contains a single atom, the phonon spectrum of the crystal represents an acoustic branch. Each extra atom introduced into the cell adds an optical phonon branch. In our case, such a unit cell is the magnetic flux quantum (see the bottom diagram in Fig. 12). An analog of the acoustic phonon branch is offered by the fundamental CDEs and SDEs, the energies of which are determined by the total electron density of the 2DES and are independent of the magnetic field strength. All the other antiphase modes are analogs of the optical branches. The number of such modes corresponding to the excitation of each particular type (CDE and SDE) is $N - 1$. Naturally, the above analogy between phonons and the collective magnetic excitations in the 2DES is rather conditional. For example, the electron occupancy factor is a continuous function of the magnetic field strength (the number of electrons per magnetic flux quantum varies continuously), whereas the number of atoms changes discretely. Accordingly, the energies of optical modes also continuously vary with the electron occupancy factor.

Obviously, optical excitations cannot be observed for electron occupancy factors $\nu < 2$, that is, when only one (the lowest) Landau level is occupied (one electron with each spin per magnetic flux quantum). Indeed, only the acoustic CDE and SDE modes are present in the Raman spectra for $\nu < 2$ (Fig. 10). An assumption that the L_0 resonance must probably have a multicomponent structure for occupancy factors $\nu > 4$, when more than one optical branch is present in the spectra of magnetic excitations, is also confirmed by experiment. For still greater occupancies ($\nu \gg 4$), the L_0-resonance

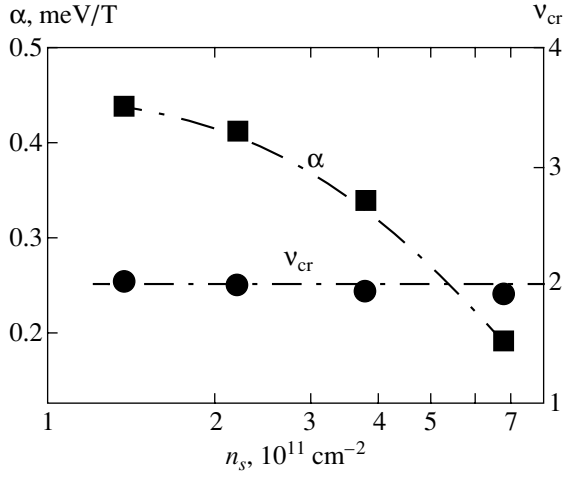


Fig. 11. Plots of the tangent slope α of the relation between the L_0 mode energy and the magnetic field strength and the critical electron occupancy factor ν_{cr} for QWs with various electron densities n_s. Dash-and-dot lines are drawn for convenience.

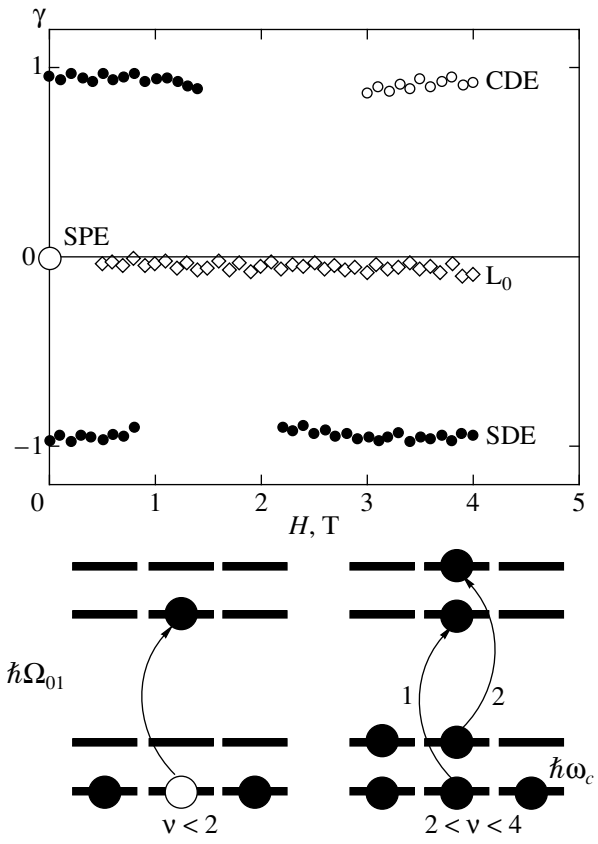


Fig. 12. The degree of polarization γ of the main peaks in the Raman spectrum of a 250-Å-wide QW with an electron density of 6.8×10^{11} cm^{-2}. The bottom diagram illustrates the appearance of an additional degree of freedom in the case of antiphase excitations: for $2 < \nu < 4$, two different CDEs (SDEs) can be constructed for the two electron transitions. For $\nu < 2$, there is a single electron transition and one excitation of each type.

half-width becomes large, which indicates that the L_0 line is multicomponent. As the occupancy factor decreases, the resonance narrows and (for $2 < \nu < 4$) the L_0 line half-width becomes equal to that of the SDE line. On the other hand, the half-widths of the acoustic intersubband SDE and CDE modes are independent of the magnetic field strength (Fig. 10).

In order to demonstrate that the L_0 resonance corresponds to the excitation of both charge and spin densities, we measured the Raman spectra using two different polarization configurations. Figure 12 shows the degree of polarization $\gamma = (I_- - I_+)/(I_- + I_+)$ of the observed lines (I_- and I_+ are the Raman signal intensities in the parallel and perpendicular polarization configurations). As can be seen, the fundamental CDE and SDE modes are observed only in the co- and cross-polarized configurations of the incident and scattered photons, respectively, while the L_0 resonance is observed with equal intensity in both configurations. Thus, the L_0 resonance consists of two components, CDE and SDE, in agreement with the theory [23].

It should be noted that the above antiphase excitations exist not only in the case of fundamental intersubband CDEs and SDEs, but are present in all other intersubband excitations as well. Recently, the antiphase ISBMs were theoretically studied in [24]. In experiment, it is rather difficult to separate the Raman signal from the equiphase and antiphase Bernstein modes, since the energies of the former and latter signals in large magnetic fields almost coincide [24].

7. CONCLUSION

We have considered the spectrum of intersubband collective excitations in a perpendicular magnetic field. The intersubband excitations of a quasi-two-dimensional electron system in a magnetic field were studied by both experimental and theoretical methods. It was demonstrated that, at a quasimomentum close to zero, the energies of the intersubband Bernstein modes are well described by the basic equation (1). At a nonzero quasimomentum, the intersubband Bernstein modes interact with the fundamental charge and spin density excitations, with the interaction energy depending on the quasimomentum and the electron density. The intersubband Bernstein modes do not interact with the LO phonons at small quasimomenta ($ql_B \ll 1$). However, in the case when the LO phonon mode is mixed with the fundamental charge density excitation (CDE), the intersubband Bernstein modes interact with the CDE–LO hybrid modes. New branches of charge and spin density excitations were considered, which are related to the antiphase oscillations of the electron subsystems on different Landau levels.

ACKNOWLEDGMENTS

This study was supported by the Russian Foundation for Basic Research.

REFERENCES

1. T. Ando, A. B. Fowler, and F. Stern, Rev. Mod. Phys. **54**, 437 (1982).
2. *Perspectives in Quantum Hall Effect*, Ed. by S. Das Sarma and A. Pinczuk (Wiley, New York, 1997).
3. A. C. Tselis and J. J. Quinn, Phys. Rev. B **29**, 3318 (1984).
4. A. Pinczuk, S. Schmitt-Rink, G. Danan, *et al.*, Phys. Rev. Lett. **63**, 1633 (1989).
5. D. Gammon, B. V. Shanabrook, J. C. Ryan, *et al.*, Phys. Rev. Lett. **68**, 1884 (1992).
6. S. Ernst, A. R. Goñi, K. Syassen, and K. Eberl, Phys. Rev. Lett. **72**, 4029 (1994).
7. T. Ando, Phys. Rev. B **19**, 2106 (1979).
8. W. Beinvogl and J. F. Koch, Phys. Rev. Lett. **40**, 1736 (1978).

9. E. Batke, G. Weimann, and W. Schlapp, Phys. Rev. B **43**, 6812 (1991).
10. G. Brozak, B. V. Shanabrook, D. Gammon, and D. S. Karzer, Phys. Rev. B **47**, 9981 (1993).
11. L. V. Kulik, I. V. Kukushkin, V. E. Kirpichev, *et al.*, Phys. Rev. B **61**, 1712 (2000); L. V. Kulik, I. V. Kukushkin, V. E. Kirpichev, *et al.*, submitted to Phys. Rev. B.
12. V. E. Kirpichev, L. V. Kulik, I. V. Kukushkin, *et al.*, Phys. Rev. B **59**, R12751 (1999).
13. L. V. Kulik, I. V. Kukushkin, V. E. Kirpichev, *et al.*, Phys. Rev. B **61**, 12717 (2000).
14. L. V. Kulik, I. V. Kukushkin, V. E. Kirpichev, *et al.*, Phys. Rev. Lett. **86**, 1837 (2001).
15. G. Abstreiter, M. Cardona, and A. Pinczuk, in *Light Scattering in Solid IV*, Ed. by M. Cardona and G. Guntherodt (Springer-Verlag, Berlin, 1984), p. 5.
16. I. V. Kukushkin and V. B. Timofeev, Adv. Phys. **45**, 147 (1996).
17. I. K. Marmorkos and S. Das Sarma, Phys. Rev. B **48**, 1544 (1993).
18. L. Wendler and R. Pechstedt, J. Phys.: Condens. Matter **2**, 8881 (1990).
19. O. Gunnarson and B. I. Lundqvist, Phys. Rev. B **13**, 4274 (1976).
20. P. Hohenberg and W. Kohn, Phys. Rev. **140**, B864 (1964); W. Kohn and L. J. Sham, Phys. Rev. **140**, A1333 (1965).
21. I. V. Lerner and Yu. E. Lozovik, Zh. Éksp. Teor. Fiz. **78**, 1167 (1980) [Sov. Phys. JETP **51**, 588 (1980)].
22. C. Kallin and B. I. Halperin, Phys. Rev. B **30**, 5655 (1984).
23. V. E. Bisti, Pis'ma Zh. Éksp. Teor. Fiz. **69**, 543 (1999) [JETP Lett. **69**, 584 (1999)].
24. V. E. Bisti, Pis'ma Zh. Éksp. Teor. Fiz. **73**, 17 (2001) [JETP Lett. **73**, 21 (2001)].

Translated by P. Pozdeev

Journal of Experimental and Theoretical Physics, Vol. 95, No. 5, 2002, pp. 940–943.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 122, No. 5, 2002, pp. 1089–1092.

SOLIDS
Electronic Properties

Zeros of the Order Parameter in the Superconducting Phase of the UGe_2 Band Ferromagnet

I. A. Fomin

Kapitsa Institute for Physical Problems, Russian Academy of Sciences, Moscow, 117334 Russia
e-mail: fomin@kapitza.ras.ru

Received July 1, 2002

Abstract—Under the assumption of a strong spin–orbital interaction, two forms of the order parameter are obtained for two superconducting phases of the ferromagnetic UGe_2 that are allowed by the crystal symmetry. For each of the two phases, symmetry nodes in the gap of Fermi excitations are found, and the consequences of the existence of nodes, which can be used for experimental phase identification, are discussed. © *2002 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

The UGe_2 band ferromagnet becomes a superconductor in the interval of pressures $11 < P_c < 16$ kbar and temperatures $T_s < 0.8$ K [1, 2]. This temperature is small compared with the Curie temperature T_C at the same pressure. The separation of Fermi surfaces for electrons with opposite spin projections, which is known from experiments, is 2 or 3 orders of magnitude greater than the superconducting gap. Such a wide separation makes the singlet Cooper pairing impossible; however, it does not inhibit the formation of Cooper pairs with parallel spins, i.e., pairs in the triplet state. Under the triplet Cooper pairing, the order parameter is a complex vector function $\mathbf{d}(\mathbf{k})$. In magnets, the symmetry with respect to time reversal is broken. As a result, the superconducting phases of a ferromagnetic metal are, generally, nonunitary; i.e., $\mathbf{d}$ and $\mathbf{d}^*$ are not proportional to each other, or, in other words, $\mathbf{d} \times \mathbf{d}^* \neq 0$. If the symmetry group of the normal phase is known, one can classify the superconducting phases that can arise from a given normal phase [3]. An attempt at such a classification for UGe_2 was made in [4]. In particular, it was shown that the symmetry of the normal phase is described by the $D_2(C_2)$ group, which is isomorphic to D_2. Four possible forms of the order parameter were obtained, which are transformed according to the four representations of D_2: A, B_1, B_2, and B_3. In [5], it was shown that, if the composition rules for antiunitary symmetry elements are taken into account, then the basis functions corresponding to the representations A and B_1 are equivalent to each other; i.e., they are transformed to the same corepresentation of the magnetic group. Two other functions are also equivalent to each other. Thus, in this case, we have only two different symmetry types of the superconducting order parameter. The necessity to use corepresentations for the classification of superconducting ferromagnetic phases was first noted in [6] in relation to the investigation of $ZrZn_2$.

In order to experimentally identify the type of the order parameter that is realized in UGe_2, it is important to investigate the existence and arrangement of nodes in the gap of the Fermi excitation spectrum for each of the possible forms of the order parameter. This investigation is the purpose of the present paper. The symmetry related part of the discussion could also be applied to the URhGe superconducting ferromagnet [7].

2. BASIS FUNCTIONS

First, we obtain the general form of the functions $\mathbf{\Psi}_A$ and $\mathbf{\Psi}_B$, which are transformed by two different corepresentations A and B of the magnetic group $D_2(C_2)$. This group consists of four operators. Two of them—the identity (E) and the rotation by π about the axis z (C_2^z)—are unitary. The other two—RC_2^x and RC_2^y—include the operation R of time reversal and, therefore, are antiunitary. The corepresentations are formed by the matrices G_1 and G_z corresponding to the unitary operators, and F_x and F_y, which correspond to the antiunitary operators. For one-dimensional corepresentations, they are complex numbers. According to the general rules for multiplying matrices that form corepresentations (see [8]), they satisfy the following equations: $G_z^2 = G_1$, $F_x \cdot F_x^* = G_1$, $F_y \cdot F_y^* = G_1$, and $F_x \cdot F_y^* = G_z$. These equations have two solutions, which generate two different corepresentations. One of them (say, A) has the form

$$G_1 = 1; \quad G_z = 1; \quad F_x = e^{2i\phi}; \quad F_y = e^{2i\phi}. \quad (1)$$

The other one, B, is

$$G_1 = 1; \quad G_z = -1; \quad F_x = e^{2i\phi}; \quad F_y = -e^{2i\phi}, \quad (2)$$

where ϕ is a real scalar. The factor 2 in the exponents is introduced to make the form of the basis functions more convenient.

Now, we write Ψ_A in the form

$$\Psi_A = \mathbf{x} f_x(\mathbf{k}) + \mathbf{y} f_y(\mathbf{k}) + \mathbf{z} f_z(\mathbf{k}), \quad (3)$$

where **x**, **y**, and **z** are the unit vectors along the second-order axes **b**, **c**, and **a**, respectively. The axis **a** is the direction of easy magnetization in the ferromagnetic phase. All functions $f_{x,y,z}(\mathbf{k})$ are odd; i.e., $f_x(-\mathbf{k}) = -f_x(\mathbf{k})$, and so on. When the operators E, C_2^z, RC_2^x, and RC_2^y act on Ψ_A, this function must be multiplied by numbers defined by Eqs. (1). This yields conditions for the functions $f_x(\mathbf{k})$, $f_y(\mathbf{k})$, and $f_z(\mathbf{k})$. Since the operators under considerations are linear (antilinear), the conditions are imposed on each function f_x, f_y, and f_z individually. For example, for $f_x(\mathbf{k})$, we have

$$f_x(-k_x, -k_y, k_z) = -f_x(k_x, k_y, k_z);$$
$$f_x^*(k_x, -k_y, -k_z) = e^{2i\phi} f_x(k_x, k_y, k_z);$$
$$f_x^*(-k_x, k_y, -k_z) = e^{2i\phi} f_x(k_x, k_y, k_z).$$

Conditions for $f_y(\mathbf{k})$ are obtained from the above ones by replacing all subscripts x with y and conversely. For $f_z(\mathbf{k})$, we have

$$f_z(-k_x, -k_y, k_z) = f_z(k_x, k_y, k_z);$$
$$f_z^*(k_x, -k_y, -k_z) = -e^{2i\phi} f_z(k_x, k_y, k_z);$$
$$f_z^*(-k_x, k_y, -k_z) = -e^{2i\phi} f_z(k_x, k_y, k_z).$$

The function

$$\Psi_A = e^{-i\phi_A}\{\hat{\mathbf{x}} k_x(a_{11} + ik_xk_ya_{10}) + \hat{\mathbf{y}} k_y(a_{22} + ik_xk_ya_{20}) + \hat{\mathbf{z}} k_z(a_{33} + ik_xk_ya_{30})\} \quad (4)$$

satisfies all the conditions above. Here, ϕ_A, a_{11}, … are real functions of k_x^2, k_y^2, and k_z^2. The function Ψ_A determined from Eq. (4) differs from that proposed in [4] by the factor $e^{-i\phi_A}$. Setting $\phi_A = \pi/2$ and using the aforementioned result in [5], we can redefine arbitrary functions so as to reduce Eq. (4) to the form Ψ_{B_1} obtained in [4].

For the corepresentation B, a similar reasoning yields the conditions

$$f_x(-k_x, -k_y, k_z) = -f_x(k_x, k_y, k_z),$$
$$f_x^*(k_x, -k_y, -k_z) = e^{2i\phi} f_x(k_x, k_y, k_z),$$
$$f_x^*(-k_x, k_y, -k_z) = e^{2i\phi} f_x(k_x, k_y, k_z)$$

for $f_x(\mathbf{k})$ and the conditions

$$f_z(-k_x, -k_y, k_z) = f_z(k_x, k_y, k_z),$$
$$f_z^*(k_x, -k_y, -k_z) = -e^{2i\phi} f_z(k_x, k_y, k_z),$$
$$f_z^*(-k_x, k_y, -k_z) = -e^{2i\phi} f_z(k_x, k_y, k_z)$$

for $f_z(\mathbf{k})$. The general form of the basis function for B is

$$\Psi_B = e^{-i\phi_B}\{\hat{\mathbf{x}} k_z(b_{13} + ik_xk_yb_{10}) + \hat{\mathbf{y}} k_z(ib_{23} + k_xk_yb_{20}) + \hat{\mathbf{z}} k_x(b_{31} + ik_xk_yb_{30})\}. \quad (5)$$

Under an appropriate choice of the phase factor $e^{-i\phi_B}$, it goes to the function Ψ_{B_3} or Ψ_{B_4} (see [4]).

3. NODES

Under the triplet Cooper pairing, the gap in the spectrum of Fermi excitations is determined (see [9]) by the eigenvalues of the matrix

$$(\Delta_k\Delta_k^\dagger)_{\alpha\beta} = \mathbf{d}(\mathbf{k})\cdot\mathbf{d}^*(\mathbf{k})\delta_{\alpha\beta} + i[\mathbf{d}(\mathbf{k})\times\mathbf{d}^*(\mathbf{k})]\sigma_{\alpha\beta}. \quad (6)$$

For nonunitary phases, this matrix has two different eigenvalues. Each of them equals the square of the gap magnitude for one of the spin orientations. Separating the real and imaginary parts of the vector $\mathbf{d}(\mathbf{k})$ by setting $\mathbf{d}(\mathbf{k}) = \mathbf{d}_1(\mathbf{k}) + i\mathbf{d}_2(\mathbf{k})$, we obtain, for the eigenvalues $(\Delta_k\Delta_k^\dagger)_{\alpha\beta}$,

$$|\Delta_{1,2}|^2 = \mathbf{d}_1^2(\mathbf{k}) + \mathbf{d}_2^2(\mathbf{k}) \pm 2|\mathbf{d}(\mathbf{k})\times\mathbf{d}^*(\mathbf{k})|.$$

The gap vanishes under the following conditions:

(1) $\mathbf{d}_1(\mathbf{k}) = 0$ and $\mathbf{d}_2(\mathbf{k}) = 0$; then, the gap vanishes for both spin orientations.

(2) $|\mathbf{d}_1(\mathbf{k})| = |\mathbf{d}_2(\mathbf{k})|$ and $|\mathbf{d}_1(\mathbf{k})| \perp |\mathbf{d}_2(\mathbf{k})|$; then, only one of the gaps vanishes.

A direct verification shows that if no special assumptions are made about the unknown functions a_{11}, a_{10}, …, b_{13}, b_{10}, …, which appear in Eqs. (4) and (5), then, for both types of the order parameter Ψ_A and Ψ_B, there are no nodes in the gap. This assertion is based solely on the properties of the order parameter symmetry. The fact that the magnetic polarization in UGe_2 and in URhGe is strong was not used.

It is known that the separation of Fermi surfaces with different spin projections suppresses the pairing amplitude for quasiparticles with opposite spins. For singlet pairing, the superconducting state is already broken at the separation $2I > \sqrt{2}\,\Delta_0$, where Δ_0 is the gap magnitude at zero temperature in the absence of polarization [10]; moreover, the state changes very rapidly as the polarization increases.

The formal cause of the pairing suppression is the fact that, under the polarization, the singularity of the scattering amplitude of quasiparticles with opposite momenta changes. Without the polarization, the scattering amplitude has the singularity $\ln(\omega_D/\Delta_0)$ in the

second order in the interaction. When the Fermi surfaces for two spin projections are separated by $2I$, the singularity becomes $\ln(\omega_D/I)$. At $I \gg \Delta_0$, this contribution can be included in the regular part of the scattering amplitude. The transition from the state $\Delta_{\uparrow\downarrow} \neq 0$ to the state $\Delta_{\uparrow\downarrow} = 0$ must occur at $I \sim \Delta_0 \sim T_s$. Taking into account that the condition $I \gg T_s$ is certainly true for UGe_2 and URhGe, we will assume that $\Delta_{\uparrow\downarrow} = 0$. In vector notation, this is equivalent to the condition $d_z(\mathbf{k}) = 0$. Under this condition, we obtain, for the two types of the order parameter,

$$\begin{aligned}\mathbf{\Psi}_A &= e^{-i\phi_A}\{\hat{\mathbf{x}}k_x(a_{11} + ik_xk_ya_{10}) \\ &+ \hat{\mathbf{y}}k_y(a_{22} + ik_xk_ya_{20})\},\end{aligned} \tag{7}$$

$$\begin{aligned}\mathbf{\Psi}_B &= e^{-i\phi_B}\{\hat{\mathbf{x}}k_z(b_{13} + ik_xk_yb_{10}) \\ &+ \hat{\mathbf{y}}k_z(ib_{23} + k_xk_yb_{20})\}.\end{aligned} \tag{8}$$

The order parameter $\mathbf{\Psi}_A$, as well as the gap on both Fermi surfaces, vanishes at the points $k_x = 0$ and $k_y = 0$. These are symmetry nodes. To verify this fact, consider the function $\mathbf{\Psi}_A(0, 0, k_z)$ and apply the operator C_2^z to it. By Eqs. (1), we have $C_2^z\mathbf{\Psi}_A(0, 0, k_z) = \mathbf{\Psi}_A(0, 0, k_z)$. On the other hand, the definition of C_2^z yields

$$\begin{aligned}C_2^z\mathbf{\Psi}_A(0, 0, k_z) &= -\mathbf{x}f_x(0, 0, k_z) \\ -\mathbf{y}f_y(0, 0, k_z) &= -\mathbf{\Psi}_A(0, 0, k_z).\end{aligned} \tag{9}$$

Therefore, $\mathbf{\Psi}_A(0, 0, k_z) = 0$.

Similarly, Eq. (8) shows that $\mathbf{\Psi}_B$ vanishes at the line $k_z = 0$. These are also symmetry nodes, since

$$\begin{aligned}C_2^z\mathbf{\Psi}_B(k_x, k_y, 0) &= -\mathbf{x}f_x(-k_x, -k_y, 0) \\ -\mathbf{y}f_y(-k_x, -k_y, 0) &= \mathbf{\Psi}_B(k_x, k_y, 0).\end{aligned} \tag{10}$$

On the other hand, it follows from Eqs. (2) that

$$C_2^z\mathbf{\Psi}_B(k_x, k_y, 0) = -\mathbf{\Psi}_B(k_x, k_y, 0);$$

therefore, $\mathbf{\Psi}_B(k_x, k_y, 0) = 0$.

4. DISCUSSION OF THE RESULTS

We see that two feasible superconducting phases of UGe_2 differ in the type and arrangement of nodes in the gap of the excitation spectrum. For the *A*-type phase (see formula (7)), the nodes are isolated and are at the intersection points of the Fermi surfaces with the direction of the easy magnetization axis. For the *B*-type phase (8), they form lines on the equators of the Fermi surfaces that are perpendicular to this axis. The existence of nodes leads to a power law for thermodynamic quantities as functions of the temperature at $T \ll T_s$, where the exponent depends on the type of nodes. The analysis of power laws for thermodynamic quantities is a conventional method for the identification of unconventional superconducting phases (see, e.g., [11]). Here, we discuss only the specific features of the low-temperature behavior caused by the magnetic polarization of UGe_2. The width of the gap generally depends on the spin projection. For example, it may be that $\Delta_\downarrow \ll \Delta_\uparrow$ if the separation of Fermi surfaces is large. Then, there exists an interval of temperatures $\Delta_\downarrow \ll T < T_s$ for which the lesser gap has almost no effect on the temperature dependences of thermodynamic quantities and the contribution of Fermi particles with the spin "down" direction to these quantities is the same as for the normal phase. Another specific feature of superconducting ferromagnets is the existence of the magnetization field $\mathbf{H}_M = 4\pi\mathbf{M}$. For UGe_2, we have $H_M \sim 1$ kOe. This is significantly greater than the field H_{C1} estimated on the basis of the temperature; i.e., the superconducting UGe_2 is in a mixed state. The combination of vortices with the line of nodes in the gap oriented perpendicularly to the vortex axes leads, according to the prediction in [12], to the appearance of a finite state density on the Fermi level, which, in turn, results in a linear (in temperature) contribution to the heat capacity c_s; moreover, $c_s \sim c_n\sqrt{H_M/H_{c2}}$. Because of the square root dependence on the field, this contribution to the heat capacity is greater than the contribution of electrons localized on vortices if the field is small compared to H_{c2}. The contribution to the heat capacity discussed here is expected to appear in the *B*-type phase and should be absent in the *A*-type phase. Thus, the expected difference of low-temperature properties of *A* and *B* phases presumably will make it possible to identify the superconducting phases observed in UGe_2 and URhGe.

ACKNOWLEDGMENTS

The work was supported by the Russian Foundation for Basic Research (project no. 01-02-16714). A part of this study was conducted in the Grenoble Research Center of the French Commissariat for Atomic Energy. I am grateful to J. Flouké for his hospitality and stimulating discussions. I am also grateful to Fourier University for financial support and to V.P. Mineev and A. Huxley for helpful remarks.

REFERENCES

1. S. S. Saxena, P. Agarwal, K. Ahilan, *et al.*, Nature **406**, 587 (2000).
2. A. Huxley, I. Sheikin, E. Ressouche, *et al.*, Phys. Rev. B **63**, 144519 (2001).
3. G. E. Volovik and L. P. Gor'kov, Zh. Éksp. Teor. Fiz. **88**, 1412 (1985) [Sov. Phys. JETP **61**, 843 (1985)].
4. I. A. Fomin, Pis'ma Zh. Éksp. Teor. Fiz. **74**, 116 (2001) [JETP Lett. **74**, 111 (2001)].

5. V. P. Mineev, cond-mat/0204263.

6. K. V. Samokhin and M. B. Walker, cond-mat/0203309.

7. D. Aoki, A. Huxley, E. Ressouche, *et al.*, Nature **413**, 613 (2001).

8. E. P. Wigner, *Group Theory and Its Application to the Quantum Mechanics of Atomic Spectra* (Academic, New York, 1959; Inostrannaya Literatura, Moscow, 1961), Chap. 26.

9. D. Vollhardt and P. Wölfle, *The Superfluid Phases of Helium 3* (Taylor and Francis, London, 1990), p. 71.

10. A. A. Abrikosov, *Fundamentals of the Theory of Metals* (Nauka, Moscow, 1987; North-Holland, Amsterdam, 1988).

11. V. P. Mineev and K. V. Samokhin, *Introduction to the Theory of Unusual Superconductivity* (Mosk. Fiz. Tekh. Inst., Moscow, 1998).

12. G. E. Volovik, Pis'ma Zh. Éksp. Teor. Fiz. **58**, 457 (1993) [JETP Lett. **58**, 469 (1993)].

Translated by A. Klimontovich

Journal of Experimental and Theoretical Physics, Vol. 95, No. 5, 2002, pp. 944–956.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 122, No. 5, 2002, pp. 1093–1106.

NONLINEAR PHYSICS

Cusp Catastrophe in Slowly Varying Equilibriums

B. I. Suleĭmanov

Institute of Mathematics and Computer Center, Ufa Research Center, Russian Academy of Sciences, Ufa, 650077 Russia
e-mail: bis@imat.rb.ru

Received November 27, 2001

Abstract—Solutions of the equations $v_x + v^3 - tv + x = 0$ and $v_{xx} = v^3 - tv + x$, which describe the nucleation of domain walls occurring in the neighborhood of cusps of slowly varying equilibriums, are analyzed. Examples related to the diffusion in smoothly inhomogeneous media are considered. © *2002 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

1.1. Partial differential equations (PDEs)

$$L(S, D_S)u(S) = \sum_{i+j=1}^{2} A_{ij}(S)\frac{\partial^{i+j}u}{\partial s_1^i \partial s_2^j} G(S, u) \quad (S = (s_1, s_2)) \tag{1.1}$$

describe various phenomena in inhomogeneous media. Generally, their solutions defy analytical analysis.

Even the simplest first-order equation

$$p\frac{\partial u}{\partial s_1} + q\frac{\partial u}{\partial s_2} = G(S, u) \quad (p > 0, q > 0 \text{ are constants}) \tag{1.2}$$

for the local parameter of a time-independent plane medium in a given velocity field (p, q) [1, pp. 150, 158] cannot be solved explicitly.

The investigation of solutions of second-order equations, such as the time-independent part

$$p_1\frac{\partial u}{\partial s_1} - \frac{\partial^2 u}{\partial s_1^2} - \frac{\partial^2 u}{\partial s_2^2} = G(S, u) \quad (p_1 \text{ is a constant}) \tag{1.3}$$

of the two-dimensional nonlinear diffusion equation

$$\frac{\partial u}{\partial \zeta} + p_1\frac{\partial u}{\partial s_1} - \frac{\partial^2 u}{\partial s_1^2} - \frac{\partial^2 u}{\partial s_2^2} = G(S, u), \tag{1.4}$$

is even more complicated. Here, in addition to the random walk of particles, their drift with a constant velocity is taken into account [2, p. 44] (in the absence of the drift, the time-independent diffusion equation

$$-\frac{\partial^2 u}{\partial s_1^2} - \frac{\partial^2 u}{\partial s_2^2} = G(S, u) \tag{1.5}$$

does not contain the first derivatives of $u(S)$). Another example is provided by the time-dependent one-dimensional diffusion equation

$$\frac{\partial u}{\partial s_1} - \frac{\partial^2 u}{\partial s_2^2} = G(S, u). \tag{1.6}$$

However, for smooth [3, 4] inhomogeneities

$$A_{ij}(S) = h_{ij}(\varepsilon S), \quad G(S, u) = f(\varepsilon S, u), \quad \varepsilon \ll 1, \tag{1.7}$$

considered in this paper, the situation is different: effective methods for the description of solutions to Eq. (1.1), (1.7) and equivalent PDEs

$$L(X, \varepsilon D_X)u(X) = f(X, u), \quad X = \varepsilon S, \tag{1.8}$$

are available (see [4–15] and references therein).

1.2. Often, solutions to Eq. (1.8) have asymptotic expansions

$$u = u_0(X) + \varepsilon u_1(X) + \varepsilon^2 u_2(X) + \ldots, \tag{1.9}$$

in which $u = u_0(X)$ is a slowly varying equilibrium of Eq. (1.1), (1.7):

$$f(X, u_0) = 0, \quad X = (x_1, x_2). \tag{1.10}$$

As an example, we cite some initial value problems solved with the help of such asymptotic expansions.

(1) Let the local medium parameter, which is a solution to Eq. (1.2), (1.7), be given, at $s_1 = 0$, by a slowly varying function $g(\varepsilon s_2)$. Then, the dilatations $x_1 = \varepsilon s_1$ and $x_2 = \varepsilon s_2$ reduce the description of the medium to solving the problem

$$\varepsilon\left(p\frac{\partial u}{\partial x_1} + q\frac{\partial u}{\partial x_2}\right) = f(x_1, x_2, u), \quad u|_{x_1 = 0} = g(x_2). \tag{1.11}$$

If, on the interval $0 \le x_1 \le K$, the solution $u = u_0(X)$ to the limit equation (1.10) to Eq. (1.11) is unique and satisfies the stability condition

$$f_u(x_1, x_2, u_0(x_1, x_2)) < 0, \tag{1.12}$$

then the solution to the initial value problem (1.11) has an asymptotic expansion (1.9) on this interval outside an infinitely small (as $\varepsilon \longrightarrow 0$) neighborhood of the line $x_1 = 0$ (see [10]).

(2) Under similar restrictions on the functions $u(0, s_2)$ and $G(S, u)$, the solution to the initial value problem for Eq. (1.6), (1.7) is expanded into series (1.9) at an infinitely small distance from the line $x_1 = 0$ (as $\varepsilon \longrightarrow 0$) (see Section 2.3).

(3) If the sink function has form (1.7) and the initial value is $u|_{\zeta = 0} = g(\varepsilon S)$, the initial value problem for the diffusion equation (1.4) is reduced to the problem

$$\frac{\partial u}{\partial \zeta} + \varepsilon p_1 \frac{\partial u}{\partial x_1} - \varepsilon^2 \left(\frac{\partial^2 u}{\partial x_1^2} + \frac{\partial^2 u}{\partial x_2^2} \right) = f(x_1, x_2, u), \qquad u|_{\zeta = 0} = g(X) \tag{1.13}$$

by the change of variables $X = \varepsilon S$. Under the same assumptions on $f(X, u)$, the solution to this problem at $\zeta \gg 1$ is represented by the asymptotic solution (1.9) to the equation

$$\varepsilon p_1 \frac{\partial u}{\partial x_1} - \varepsilon^2 \left(\frac{\partial^2 u}{\partial x_1^2} + \frac{\partial^2 u}{\partial x_2^2} \right) = f(x_1, x_2, u). \tag{1.14}$$

The asymptotic expansion of the solution of the initial value problem (1.13) is a sum of the right-hand side of (1.9) and the series

$$\Pi_0(\zeta, X) + \varepsilon \Pi_1(\zeta, X) + \varepsilon^2 \Pi_2(\zeta, X) + \ldots$$

consisting of exponentially small (at $\zeta \gg 1$) terms (see the beginning of Section 4.2).

Note that the solution to (1.13) obviously tends to the root $u_0(X)$ as $\zeta \longrightarrow \infty$ and $\varepsilon \longrightarrow 0$. Indeed, according to (1.12), this is an asymptotically stable equilibrium of the limiting ordinary differential equation ODE (1.13)

$$u_\zeta = f(X, u),$$

which differs from other solutions by exponentially small quantities as $\zeta \longrightarrow \infty$ (see [16, p. 289]). By the same reason, solutions to problem (1.11) and to the initial value problem considered in Section 4 tend to $u_0(X)$.

Generally, the fact that solutions to PDEs (1.8) can often be represented by series (1.9) depends on how often roots to Eq. (1.10) satisfy conditions (1.12): since, for the majority of smooth functions $f(X, u)$, $f_u(X, u_0(X))$ can vanish only on isolated curves of the plane X [17, 18], about half of all smooth functions $f(X, u)$ satisfy condition (1.12) in domains that contain no such lines.

We give one more example of a boundary value problem that shows that, if this condition is satisfied, then asymptotic expansions (1.9) are characteristic of boundary value problems as well. Assume that, in a closed domain G with a smooth boundary dG, Eq. (1.10) has three roots, and the minimal $u_0^1(X)$ and maximal $u_0^3(X)$ of them satisfy condition (1.12). We also assume that, in the domain G, $u(X, \varepsilon)$ satisfies the equation

$$-\varepsilon^2 \left(\frac{\partial^2 u}{\partial x_1^2} + \frac{\partial^2 u}{\partial x_2^2} \right) = f(x_1, x_2, u), \tag{1.15}$$

which is equivalent to Eq. (1.5), (1.7). Furthermore, we assume that it satisfies one of the two boundary conditions

$$u|_{dG} = a(X) \quad \text{or} \quad \left. \frac{\partial u}{\partial n} \right|_{dG} = b(X), \tag{1.16}$$

where n is the normal to dG. Under certain additional conditions, solutions to problems (1.15), (1.16) have two different asymptotic expansions of the form (1.9) on different sides of the shock curve defined by the equation

$$\int_{u_0^1(x_1, x_2)}^{u_0^3(x_1, x_2)} f(x_1, x_2, u) du = 0. \tag{1.17}$$

In one of them, the principal term is $u_0^1(X)$; and in the other one, it is $u_0^3(X)$ [11–13].[1]

1.3. In studies devoted to physical problems (e.g., [3–5]), simple states described by series (1.9) are mentioned only in passing, and the typical nature of zero curves of $f_u(X, u_0(X))$ on the plane X is ignored (see [17, 18]) (these curves consist of smooth parts formed by fold points and joined at cusps). Therefore, the fact that asymptotic expansions (1.9) become inappropriate at singular points $u_0(X)$ is also ignored.

However, for ODE (1.8), the case of a fold in the solution to the limit equation (1.10) has been thoroughly analyzed in [19–23]. It is clear that a similar scheme can be used for the analysis of this type of singularities in PDEs (an investigation of this sort can be found in Section 4). In particular, it is obvious that the behavior of the corresponding solutions to PDE (1.8) in the neighborhood of the fold points of the roots to Eq. (1.10) is also described by solutions of the Riccati equation

$$\Gamma_\xi = \xi - \Gamma^2 \tag{1.18}$$

or, when Eq. (1.8) does not include the first derivatives, by the Painlevé equation

$$\Gamma_{\xi\xi} = \xi - \Gamma^2.$$

[1] A smooth change from one asymptotic expansion to the other occurs in the neighborhood of curve (1.17), which has a finite extent in terms of the variables S. In terms of these variables, such asymptotic solutions to Eq. (1.8) (of the type of smooth shock waves) are interpreted as domain walls [3].

The case of a cusp, which is specific to PDEs, has not yet been considered. However, this analysis is essential for understanding the behavior of asymptotic solutions (1.9) "beyond" the cusp. Indeed, even if "before" this point the solution to Eq. (1.10) was unique, it is three-valued "beyond" this point, i.e., in the domain where the roots of (1.10) overlap (see Section 2.2). For example, the assumption that in this case a domain wall is formed with a front localized in the vicinity of a curve, which is determined from the condition of equality of areas (1.17), is most often false. Indeed, an analysis of neighborhoods of root cusps (1.10) (see Sections 2, 3, and 5) shows that the formation of such a domain wall is characteristic only of solutions to the PDE (1.8) that do not include the first-order derivatives of $u(X)$. The main conclusion of this analysis is that the formation of domain walls with fronts localized in a vanishingly narrow (as $\varepsilon \longrightarrow 0$) neighborhood of one of the boundaries of the domain where the roots of (1.10) overlap is typical for the majority of PDEs (1.8).

In Section 4 of this study, an example is analyzed that explains how the terms of (1.8) with second-order derivatives affect structures formed beyond the cusp.[2] However, a great part of the reasoning used in that section and that employed earlier is identical, and Section 4 can be omitted on first reading.

2. REFERENCE EQUATIONS

2.1. In this section, we show that solutions to the PDE (1.8) that have asymptotic expansions (1.9) are described, in the neighborhood of the cusps for $u_0(X)$, by solutions of the Abel ODE

$$P(v, x) = v_x + v^3 - tv + x = 0 \tag{2.1}$$

or, if $h_{10}(X) = h_{01}(X) = 0$ in Eq. (1.8), by solutions to one of the equations

$$v_{xx} = v^3 - tv + x, \quad v_{xx} + v^3 - tv + x = 0. \tag{2.2}$$

Moreover, the leading term $g(x, t)$ of the asymptotic expansions (at $x^2 + t^2 \longrightarrow \infty$) of solutions to Eqs. (2.1) or (2.2) in the matching domains with the asymptotic expansions (1.9) (see [15]) is a root of the equation

$$g^3 - tg + x = 0. \tag{2.3}$$

2.2. We begin with the presentation of certain conclusions drawn in [18].

(1) On the plane X, cusps X^* are typical for which the first three coefficients of the Taylor series of the smooth functions $f(X^*, u)$

$$\sum_{j=0}^{\infty} \frac{1}{j!}\frac{\partial^j f}{\partial u^j}(X^*, u^*)(u - u^*)^j$$

at the points u^*—the roots to the equations $f(X^*, u) = 0$—are zero. Due to x_1^* and x_2^*, no more than two constraints can be imposed on the expansions of $f(X, u)$ at the points $(X = X^*, u = u^*)$. Therefore, in the corresponding Taylor series at these points

$$\begin{aligned} f(X, u) &= a(x_1 - x_1^*) + b(x_2 - x_2^*) \\ &+ (u - u^*)[c(x_1 - x_1^*) + d(x_2 - x_2^*)] + e(u - u^*)^3 \\ &+ \sum_{i+j>1} c_{ij0}(x_1 - x_1^*)^i (x_2 - x_2^*)^j \\ &+ (u - u^*)\sum_{i+j>1} c_{ij1}(x_1 - x_1^*)^i (x_2 - x_2^*)^j + \sum_{k>3} c_{00k}(u - u^*)^k \\ &+ \sum_{k>1} (u - u^*)^k \sum_{i+j>0} c_{ijk}(x_1 - x_1^*)^i (x_2 - x_2^*)^j, \end{aligned} \tag{2.4}$$

the constants a, b, c, and d, along with the constant e, are nonzero.

(2) There exist constants c_{ij}^k [18, pp. 45, 46, 52] such that the change of variables

$$\begin{aligned} u - u^* &= \sum_{i+j=1}^{\infty} c_{ij}^0 Y^i Z^j + U\left[1 + \sum_{i+j=1}^{\infty} c_{ij}^1 Y^i Z^j\right] \\ &+ \sum_{k=2}^{\infty} U^k \sum_{i+j=0}^{\infty} c_{ij}^k Y^i Z^j, \\ Y &= a(x_1 - x_1^*) + b(x_2 - x_2^*), \\ Z &= c(x_1 - x_1^*) + d(x_2 - x_2^*) \end{aligned} \tag{2.5}$$

take Eq. (1.10), which is determined by series (2.4), to the equation

$$\delta(Y, Z) + \sigma(Y, Z)U + eU^3 = 0, \tag{2.6}$$

$$\begin{aligned} \delta(Y, Z) &= Z + \sum_{i+j>1} \delta_{ij} Y^i Z^j, \\ \sigma(Y, Z) &= Y + \sum_{i=j>1} \sigma_{ij} Y^i Z^j. \end{aligned} \tag{2.7}$$

[2] The relationship of domain walls, which are solutions to the time-independent PDE (1.14), to solutions of the time-dependent diffusion equation (1.4), (1.7) is discussed in the Conclusions.

(3) At $e\sigma \geq 0$, Eq. (2.6) has a unique root, and, for $e\sigma < 0$, it is unique only outside the overlapping domain

$$|\delta| < \frac{(-4\sigma^3)^{1/2}}{(27e)^{1/2}};$$

inside this domain, the solution to (2.6) is three-valued. Series (2.5) is also three-valued in this case.

2.3. In the neighborhood of the cusp, we pass to dilated variables. Since the leading term of series (2.5), being a solution to Eq. (2.6), depends both on δ and on σ, all three terms on the left-hand side of this equation should be balanced: if the order of smallness of U in the new variables is ε^k, then $\delta(Y, Z)$ and $\sigma(Y, Z)$ must be of order ε^{3k} and ε^{2k}, respectively. Due to this consideration and the form of series (2.5) and (2.7), it is clear that the dilatations should be as follows:

$$Z = \varepsilon^{3k}z, \quad Y = \varepsilon^{2k}y, \quad u - u^* = \varepsilon^k V, \qquad (2.8)$$

where $k > 0$ is a constant. These dilatations reduce series (2.4) to the form

$$f(X, u) = \varepsilon^{3k}\left(z + yV + eV^3 + \sum_{j>1}\varepsilon^{jk}P_j(z, y, V)\right). \qquad (2.9)$$

Therefore, according to the general idea of matching asymptotic expansions [15], k must be chosen so as to make the left-hand side of Eq. (1.8) of order ε^{3k} in variables (2.8).

In the new variables, the operations of differentiation with respect to x_1 and x_2 in the principal order with respect to the parameter ε are actually the differentiations with respect to z:

$$\frac{\partial}{\partial x_1} = \frac{a}{\varepsilon^{3k}}\frac{\partial}{\partial z} + \frac{c}{\varepsilon^{2k}}\frac{\partial}{\partial y}, \quad \frac{\partial}{\partial x_2} = \frac{b}{\varepsilon^{3k}}\frac{\partial}{\partial z} + \frac{d}{\varepsilon^{2k}}\frac{\partial}{\partial y}.$$

Therefore, the change of variables (2.8) in the principal order reduces the operator $L(X, \varepsilon D_X)$ (see (1.1) and (1.7)) on the left-hand side of Eq. (1.8) to the differentiation with respect to z:

$$L(X, \varepsilon D_X) = \varepsilon\frac{M}{\varepsilon^{3k}}\frac{\partial}{\partial z} + \varepsilon^2\frac{N}{\varepsilon^{6k}}\frac{\partial^2}{\partial z^2} + \ldots. \qquad (2.10)$$

(In the case of the general position, the constants

$$M = ah_{10}(X^*) + bh_{01}(X^*)$$

and

$$N = a^2h_{20}(X^*) + abh_{11}(X^*) + b^2h_{02}(X^*)$$

are nonzero: see Subsection 2.2.) Thus, k is found from the equation

$$\min(1 - 2k, 2 - 5k) = 3k.$$

It yields $k = 1/5$, and change (2.8) reduces Eq. (1.8) to the equation

$$M\frac{\partial V}{\partial z} - z - yV - eV^3 = O(\varepsilon^{1/5});$$

the limit equation to this one is a first-order ODE. This is also true for $N = 0$; it is only important that the constant M is nonzero (for the PDEs (1.2), (1.3), and (1.6), it is equal to $pa + qb$, p_1a, and a, respectively).

However, for the PDE (1.8) that does not include the first-order derivatives, M is equal to zero. In this case, it follows from the right-hand sides of Eqs. (2.9) and (2.10) that $k = 1/4$ and the change of variables (2.8) reduces this PDE (1.8) to the equation

$$N\frac{\partial^2 V}{\partial z^2} - z - yV - eV^2 = O(\varepsilon^{1/4})$$

(for the time-independent diffusion equation (1.15), $N = -a^2 - b^2$).

2.4. If the constant k is chosen as specified above, the substitution of the series

$$V = V_0(z, y) + \sum_{n=1}^{\infty}\varepsilon^{nk}V_n(z, y) \qquad (2.11)$$

into the equations obtained from the PDE (1.8) by the change of variables (2.8) and then by equating the coefficients of equal powers of ε yields ODEs for the coefficients of series (2.11). For their leading terms, the ODEs have the form

$$M\frac{\partial V_0}{\partial z} - z - yV_0 - eV_0^3 = 0 \qquad (2.12)$$

if Eq. (1.8) includes the first derivatives of $u(X)$; otherwise, if no first derivatives appear in this equation, the ODE has the form

$$N\frac{\partial^2 V_0}{\partial z^2} - z - yV_0 - eV_0^3 = 0. \qquad (2.13)$$

The ODE (2.12) is transformed into Eq. (2.1) by the change of variables

$$z = -\left(\frac{M^3}{e}\right)^{1/5}x, \quad y = -(M^2e)^{1/5}t, \quad V_0 = -\left(\frac{M}{e^2}\right)^{1/5}v, \qquad (2.14)$$

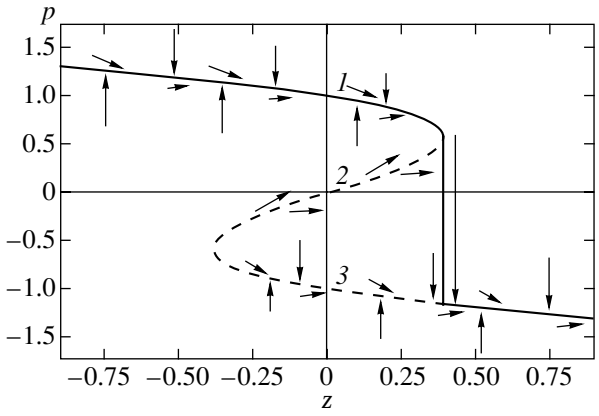


Fig. 1. Directions of the phase velocity of system (3.3) (arrows); the limit of the solution to system (3.3) at $t = \infty$ (solid curve); and branches *1*, *2*, and *3* of the curve $z = p - p^3$ (dashed curves in the places where they differ from the solid curve).

and the ODE (2.13) is transformed into Eq. (2.2) by the change of variables

$$z = \operatorname{sgn}(e)\left(\frac{N^3}{e}\right)^{1/8} x,$$
$$y = -\operatorname{sgn}(e)(Ne)^{1/4} t, \qquad (2.15)$$
$$V_0 = -\left(\frac{N}{e^3}\right)^{1/8} v.$$

The requirement that the behavior of the asymptotic expansions (1.9) at $X \longrightarrow X^*$ is matched with the behavior of series (2.11) at $y^2 + z^2 \longrightarrow \infty$ imposes certain conditions on the asymptotics $V_n(z, y)$. In particular, it is seen from the right-hand side of expansion (2.9) that the leading term of the asymptotics $V_0(z, y)$ at $y^2 + z^2 \longrightarrow \infty$ in the matching domain of expansions (1.9) and (2.11) is a root of the equation obtained from Eqs. (2.12), (2.13) by replacing the terms containing the derivatives with zero. Each change of variables (2.14) and (2.15) transforms this root into the root of Eq. (2.3).

3. A SPECIAL SOLUTION TO EQUATION (2.1)

3.1. For the analysis of solutions of the PDE (1.8) such that the leading terms $u_0(X)$ of their asymptotic expansions (1.9) have singularities of cusp, it is of prime importance to find out the asymptotics $U(X, \varepsilon)$ "beyond" those points. For this reason, the most interesting thing concerning the corresponding solution to Eq. (2.1) is its behavior at $t \longrightarrow \infty$.

The change of variables

$$s = \frac{x}{|t|^{3/2}}, \quad v(x, t) = |t|^{1/2} r(s, t) \qquad (3.1)$$

reduces the solutions to the reference equation (2.1) to the solutions

$$p(s, t) = r(s, t), \quad z = s \qquad (3.2)$$

of the slow–fast autonomous system

$$|t|^{-5/2} p_s = \operatorname{sgn}(t) p - z - p^3, \quad z_s = 1. \qquad (3.3)$$

This allows us to use the following reasoning to analyze the behavior of solutions.[3]

An analysis of Fig. 1, where the directions of the phase velocity of system (3.3) ($t^{5/2}(p - p^3 - z)$, 1) are depicted, shows the following (note that, in contrast to branch *2* of the curve $z(p) = p - p^3$, the stability condition $(p - p^3 - z)'_p = 1 - 3p^2 < 0$ is satisfied on its branches *1* and *3*). As *s* increases from $-\infty$ to the value corresponding to the instance

$$s = s_0 = \frac{2}{3^{3/2}}, \qquad (3.4)$$

when branch *1* vanishes, the corresponding solution (3.2) to system (3.3) moves in the vicinity of this branch and then, having jumped into a small neighborhood of the point ($p = -2/3^{1/2}$, $z = s_0$) of branch *3*, remains in the vicinity of this branch as *s* continues to increase.

A similar reasoning suggests that the unique root of Eq. (2.3) is the leading term of the asymptotics $v(x, t)$ for all *x* as $t \longrightarrow -\infty$. Indeed, the change of variables (3.1) reduces Eq. (2.1) to the equation

$$|t|^{-5/2} r_s = \operatorname{sgn}(t) r - r^3 - s, \qquad (3.5)$$

which turns into the cubic equation

$$\operatorname{sgn}(t) r - r^3 - s = 0 \qquad (3.6)$$

when $t \longrightarrow -\infty$. The latter equation has the unique root $r = r_0(s)$ for $t < 0$. Since

$$(\operatorname{sgn}(t) r - r^3 - s)'_r\big|_{r = r_0(s)} = \operatorname{sgn}(t) - 3r_0(s)^2 < 0, \qquad (3.7)$$

it is evident that as *s* increases from $-\infty$ to ∞, the discussed solution to Eq. (3.5) moves along this root of Eq. (3.6).

3.2. The reasoning above suggests two conclusions.

(1) The leading term of the asymptotics $v(x, t)$ at $x^2 + t^2 \longrightarrow \infty$ outside the curve where the roots of Eq. (2.3) $x = s_0 t^{3/2}$ merge coincides with the smoothly varying root of Eq. (2.3).

(2) The neighborhood of this merge curve is a shock layer; when this layer is crossed from left to right along the axis *x*, the values of $v(x, t)$ sharply decrease from $\sqrt{t/3}\,(1 + o(1))$ down to $-2\sqrt{t/3}\,(1 + o(1))$.

The validity of these conclusions was verified by numerical simulation of the behavior of the solution to

[3] It is reduced to the application of the standard qualitative analysis technique [24, pp. 20–22] developed for general-type slow–fast systems.

Eq. (2.1), which was assumed to approach the root of Eq. (2.3) at any t both when $x \longrightarrow \infty$ and $x \longrightarrow -\infty$. The simulation, which was performed on large intervals $-L < x < L$, used the simple iterative procedure

$$v_{\text{new}}(k) = v_{\text{old}}(k) + 0.05hP[v_{\text{old}}(k), k],$$
$$k = -N+1, \ldots, -1, 0, 1, \ldots, N-1.$$

In this procedure, the differential operator $P(v, x)$ on the left-hand side of (2.1) was approximated by the finite-difference operator

$$P[v, k] = \frac{v_{k+1} - v_k}{h} + v_k^3 - tv_k + kh, \quad h = \frac{L}{N};$$

the function $g(kh, t)$ determined by the maximal root of Eq. (2.3) for negative x (and by its minimal root for nonnegative x) was used as the zero approximation $v_{\text{st}}(k)$ for $k > 0$ ($k \le 0$).

The results of calculations by this rapidly converging procedure are partially presented in Fig. 2; they completely confirm both conclusions above.

3.3. It follows from the above reasoning and the results obtained in [24] that the following series give complete asymptotic expansions $v(x, t)$ (at $t \longrightarrow \pm\infty$) on both sides of the curve $x = s_0 t^{3/2}$:

$$\sum_{j=0}^{\infty} r_j(s)|t|^{(1-5j)/2}.$$

Here, $r_0(s)$ are the roots of the cubic equation (3.6) satisfying condition (3.7) (at $t > 0$, they are different for $s > s_0$ and $s < s_0$), and $r_j(s)$ are recurrently expressed in terms of $r_0(s)$ upon substituting the series

$$\sum_{j=0}^{\infty} r_j(s)|t|^{-5j/2} \tag{3.8}$$

into Eq. (3.5) and equating the coefficients of equal powers of $|t|$.

Asymptotic expansions at $|x| \longrightarrow \infty$ that are suitable in the neighborhood of the line $t = 0$ are given by the series

$$v(x, t) = x^{1/3}\sum_{k=0}^{\infty} g_k(\rho)x^{-5k/3}, \quad \rho = t/x^{2/3}, \tag{3.9}$$

where $g_0(\rho)$ is the unique root of the equation

$$1 - \rho g_0 + g_0^3 = 0.$$

This series is obtained from the asymptotic expansions $v(x, t)$ at $t \longrightarrow \pm\infty$ by passing from the variable t to x and from s to ρ. Indeed, the coefficients $r_n(s)$ of expansion (3.8) have the asymptotic expansions

$$r_n(s) = \sum_{k=0}^{\infty} r_{nk}s^{(1-5k-2n)/3}$$

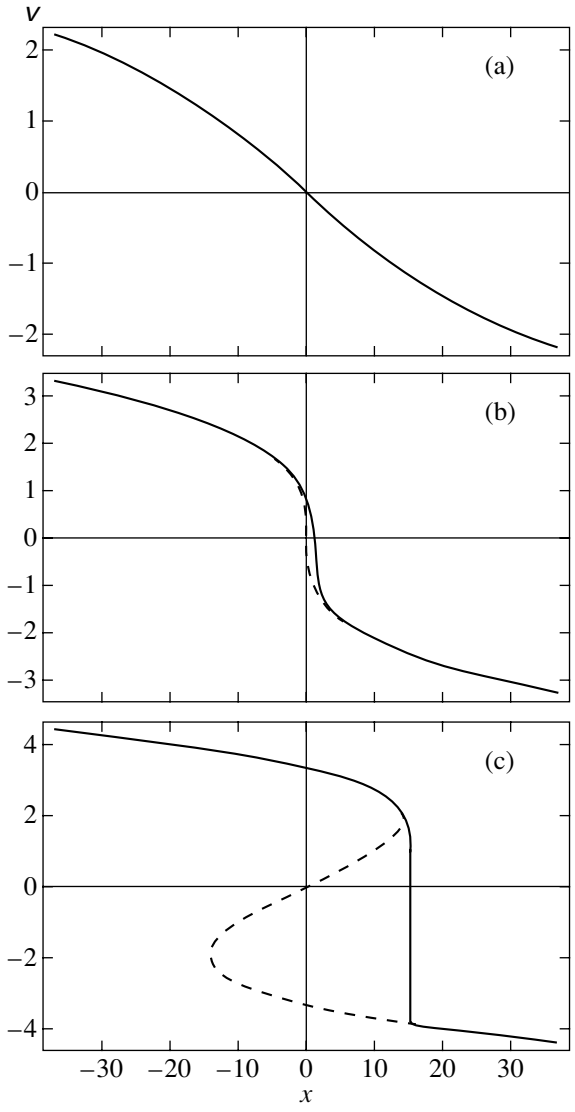


Fig. 2. The solution v to Eq. (2.1) as a function of x at various t (solid curves) and the plots of the roots of the cubic equation (2.3) where they differ from the plots of v (dashed curve) at the instances of time $t = -11$ (a), 0 (b), and 11 (c).

when $s \longrightarrow \pm\infty$. Multiplying series (3.8) by $|t|^{1/2}$ and substituting the above expressions for the coefficients into these products, we obtain

$$\sum_{k=0}^{\infty} |t|^{1/2-5k/2}s^{1/3-5k/3}\sum_{n=0}^{\infty} r_{nk}s^{-2n/3}$$
$$= \sum_{k=0}^{\infty} x^{1/3-5k/3}\sum_{n=0}^{\infty} r_{nk}s^{-2n/3}.$$

Substituting $\text{sgn}(t)\rho$ for $s^{-2/3}$ in this series, we obtain (3.9).

3.4. For $t \gg 1$, the system of ODEs (3.3) is similar to the system

$$\mu p_s = p - p^3 - z, \quad z_s = p, \quad \mu \ll 1. \tag{3.10}$$

The latter system describes the Van der Pol oscillator [14, 24, 25]. Both systems belong to the class of slow–fast systems of ODEs; therefore, the solution to Eq. (2.1) can be completely described in the neighborhood of the curve $x = s_0 t^{3/2}$ using the technique presented, e.g., in [24]. Unfortunately, the computations are very tedious, and we only describe the shock layer of the solution to Eq. (2.1) approximately as was done for Eq. (3.10) in [14].

In the first approximation, the jump of the corresponding solution $r(s, t)$ to Eq. (3.5) when s goes beyond the point (3.4) is given by the separatrix solution of the ODE

$$R_\sigma + \left(R - \frac{1}{3^{1/2}}\right)^2\left(R + \frac{2}{3^{1/2}}\right) = 0, \qquad \sigma = t^{5/2}\left(s - \frac{2}{3^{3/2}}\right) - T(t), \tag{3.11}$$

which exponentially decreases down to the value $-2 \times 3^{-1/2}$ when $\sigma \longrightarrow \infty$. At $\sigma \longrightarrow -\infty$, the following asymptotic expression is true:

$$R = \frac{1}{3^{1/2}} + \frac{1}{3^{1/2}\sigma} - \frac{\ln|\sigma|}{3^{3/2}\sigma^2} + O\left(\frac{\ln^2|\sigma|}{\sigma^3}\right). \tag{3.12}$$

Asymptotics (3.11) and (3.12) determine $R(\sigma)$ up to the function $T(t)$. Its asymptotics as $t \longrightarrow \infty$ can be found only when higher order approximations of $r(s, t)$ are constructed. Moreover, it turns out that the complete asymptotic expansion

$$r_{\text{int}}(\sigma, t) = R(\sigma) + \ldots,$$

in which the separatrix solution of Eq. (3.11) is the leading term, is compatible only with the part of expansion (3.8) that is valid on the right of the point s_0 (the corresponding matching is performed for $s \longrightarrow s_0 + 0$ and $\sigma \longrightarrow \infty$). At $\sigma \longrightarrow -\infty$ and $s \longrightarrow s_0 - 0$, the asymptotic expansion $r_{\text{int}}(\sigma, t)$ and the part of expansion (3.8) that approximates $r(s, t)$ on the left of s_0 are matched with the help of the intermediate expansion

$$r_{\text{inm}}(\nu, t) = \frac{1}{3^{1/2}} + \frac{w_1(\nu)}{t^{5/6}} + \frac{w_2(\nu)}{t^{5/3}} + \ldots, \tag{3.13}$$

which depends on the scaling variable

$$\nu = t^{5/3}(s - s_0).$$

Its coefficients $w_j(\nu)$ are uniquely found from the recurrent sequence of ODEs

$$\frac{dw_1}{d\nu} = -\nu - \sqrt{3}w_1^2, \tag{3.14}$$

$$\frac{dw_2}{d\nu} = -2\sqrt{3}w_1 w_2 - w_1^3, \tag{3.15}$$

$$\ldots$$

and the conditions imposed on the behavior of $w_j(\nu)$ when $\nu \longrightarrow \infty$. These conditions are obtained by matching expansion (3.13) with the part of series (3.8) that approximates $r(s, t)$ at $s < s_0$. The dilatations

$$\nu = -\frac{q}{3^{1/6}}, \quad w_1(\nu) = -\frac{\Gamma(q)}{3^{1/3}} \tag{3.16}$$

reduce Eq. (3.14) to Eq. (1.18) with the independent variable q. Its general solution

$$(\ln(c_1\text{Ai}(q) + c_2\text{Bi}(q)))'_q \tag{3.17}$$

is the logarithmic derivative of a combination of the Airy functions $\text{Ai}(q)$ and $\text{Bi}(q)$. It follows (see [20]) from the well-known asymptotics of $\text{Ai}(q)$ and $\text{Bi}(q)$ at $q \longrightarrow \infty$ (see [26]) and the matching conditions for the asymptotic expansions (3.8) and (3.13) that $c_2 = 0$ in solution (3.17). Therefore,

$$w_1(\nu) = -\frac{1}{3^{1/3}}\frac{\text{Ai}'(q)}{\text{Ai}(q)}. \tag{3.18}$$

This solution is smooth on the interval (ξ_0, ∞), where ξ_0 is the root of $\text{Ai}(q)$ that is nearest to $q = 0$. When $q \longrightarrow \xi_0$, we have

$$w_1(\nu) = -\frac{(1 + O((q - \xi_0)^2))}{3^{1/3}(q - \xi_0)} = \frac{(1 + O((\nu + 3^{-1/6}\xi_0)^2))}{3^{1/2}(\nu + 3^{-1/6}\xi_0)}. \tag{3.19}$$

Relations (3.18), (3.19), and the power growth of the asymptotics for $w_2(\nu)$ at $\nu \longrightarrow \infty$, which follows from the matching condition for (3.18) and (3.13), uniquely determine the desired solution to Eq. (3.15):

$$w_2(\nu) = \frac{1}{3^{7/6}\text{Ai}(q)^2}\left[\int_0^\infty \frac{\text{Ai}'(k)^3}{\text{Ai}(k)}dk - \int_0^q\left(\frac{\text{Ai}'(k)^3}{\text{Ai}(k)} - \frac{\text{Ai}'(\xi_0)^2}{k - \xi_0}\right)dk - \text{Ai}'(\xi_0)^2\ln\left|\frac{q - \xi_0}{\xi_0}\right|\right].$$

When $\nu \longrightarrow -3^{-1/6}\xi_0$, we have

$$w_2(\nu) = \frac{1}{3^{3/2}(\nu + 3^{-1/6}\xi_0)^2} \times \left[-\ln\left|\frac{\nu + 3^{-1/6}\xi_0}{3^{-1/6}\xi_0}\right| + \int_0^\infty \frac{\text{Ai}'(k)^3}{\text{Ai}'(\xi_0)^2\text{Ai}(k)}dk + \int_{\xi_0}^0\left(\frac{\text{Ai}'(k)^3}{\text{Ai}'(\xi_0)^2\text{Ai}(k)} - \frac{1}{k - \xi_0}\right)dk\right] \times (1 + O((\nu + 3^{-1/6}\xi_0)^2)). \tag{3.20}$$

Now, we derive from estimates (3.19) and (3.20) that the asymptotics

$$T(t) = -\frac{\xi_0}{3^{1/6}}t^{5/6} + \frac{5}{18}\ln t + \frac{1}{3}\left[\ln\left|\frac{\xi_0}{3^{1/6}}\right| + \int_0^\infty \frac{\mathrm{Ai}'(k)^3}{\mathrm{Ai}'(\xi_0)^2\mathrm{Ai}(k)}dk + \int_{\xi_0}^0\left(\frac{\mathrm{Ai}'(k)^3}{\mathrm{Ai}'(\xi_0)^2\mathrm{Ai}(k)} - \frac{1}{k-\xi_0}\right)dk\right] + o(1) \quad (3.21)$$

satisfies the matching condition of the behavior of asymptotic expansion (3.13) at $\nu \longrightarrow -3^{-1/6}\xi_0$ with that of the expansion $r_{\mathrm{int}}(\sigma, t)$ at $\sigma \longrightarrow -\infty$ in the leading order with respect to t.

Indeed, the expansions $r_{\mathrm{int}}(\sigma, t)$ and $r_{\mathrm{inm}}(\nu, t)$ are matched for the values of σ of an order greater than $O(\ln t)$. Hence, the form of the variable σ defined in (3.11) and estimate (3.21) suggest that, in the process of matching, at $\sigma \longrightarrow -\infty$, the quantity

$$\sigma + T(t) + \frac{t^{5/6}\xi_0}{3^{1/6}} = t^{5/2}\left(s - \frac{2}{3^{1/2}}\right) + \frac{t^{5/6}\xi_0}{3^{1/6}} = t^{5/6}\left(\nu + \frac{\xi_0}{3^{1/6}}\right)$$

also tends to $-\infty$; therefore, (3.19) and (3.20) yield the expansion

$$r_{\mathrm{inm}}(\nu, t) = \frac{1}{3^{1/2}} + \frac{1}{3^{1/2}\sigma} - \frac{\ln|\sigma|}{3^{3/2}\sigma^2} + \frac{1}{3^{3/2}\sigma^2}\left[-3\left(T(t) + \frac{t^{5/6}\xi_0}{3^{1/6}}\right) + \frac{5}{6}\ln t + \ln\left|3^{-1/6}\xi_0\right| + \int_0^\infty \frac{\mathrm{Ai}'(k)^3}{\mathrm{Ai}'(\xi_0)^2\mathrm{Ai}(k)}dk + \int_{\xi_0}^0\left(\frac{\mathrm{Ai}'(k)^3}{\mathrm{Ai}'(\xi_0)^2\mathrm{Ai}(k)} - \frac{1}{k-\xi_0}\right)dk\right] + \ldots.$$

In this series, we substitute the right-hand side of (3.21) for $T(t)$ to obtain the following representation of the intermediate asymptotic expansion $r_{\mathrm{inm}}(\nu, t)$:

$$r_{\mathrm{inm}}(\nu, t) = \frac{1}{3^{1/2}} + \frac{1}{3^{1/2}\sigma} - \frac{\ln|\sigma|}{3^{3/2}\sigma^2} + \ldots.$$

Up to the terms that are not written out explicitly, this representation coincides with the sum of the first three terms of asymptotics (3.12) for the leading term of the internal asymptotic expansion $r_{\mathrm{int}}(\sigma, t)$.

Asymptotics (3.21) completes the description of the behavior of the solution to Eq. (3.5) in the shock layer in the principal order with respect to t. This gives a description of the behavior of the universal special solution of Eq. (2.1) accurate to the leading term of its asymptotic expansion at $x^2 + t^2 \longrightarrow \infty$.

4. AN EXAMPLE: THE FORMATION OF A DOMAIN WALL FROM THE SOLUTION TO THE DIFFUSION EQUATIONS (1.6), (1.7)

4.1. Consider the solution to the PDE equivalent to Eq. (1.6), (1.7)[4]

$$\varepsilon u_\tau = \varepsilon^2 u_{\chi\chi} + f(\tau, \chi, u), \quad \varepsilon \ll 1, \quad (4.1)$$

with the initial value

$$u|_{\tau=0} = a(\chi). \quad (4.2)$$

Assume that ($\tau = \tau^* > 0$, $\chi = \chi^*$) is the general position cusp for the roots of the equation

$$f(\tau, \chi, u) = 0 \quad (4.3)$$

and that this equation has a unique root $u = u_0(\tau, \chi)$ for $\tau \le \tau^*$ for which the conventional stability condition

$$f_u(\tau, \chi, u_0(\tau, \chi)) < 0 \quad (4.4)$$

is fulfilled. Then, within any interval $\delta < \tau < \tau^* - \delta$ with the boundaries determined by a constant δ ($0 < \delta < \tau^*/2$) independent of ε, the solution to the boundary value problem under consideration is expanded in a series of type (1.9):

$$u = u_0(\tau, \chi) + \varepsilon u_1(\tau, \chi) + \varepsilon^2 u_2(\tau, \chi) + \ldots. \quad (4.5)$$

4.2. Series (4.5) does not satisfy the initial condition (4.2). However, if the stability condition (4.4) is fulfilled, the residual can be corrected in a standard way (see [27, Chapter 3]) by adding the series

$$\Pi u = \Pi_0(\zeta, \chi) + \varepsilon\Pi_1(\zeta, \chi) + \varepsilon^2\Pi_2(\zeta, \chi) + \ldots, \quad \zeta = \frac{\tau}{\varepsilon},$$

to the right-hand side of (4.5); this series consists of exponentially small (at $\zeta \longrightarrow \infty$) terms.

The reference to [27], which is devoted to slow–fast systems of ODEs rather than to PDEs, is correct. Indeed, as in the situation considered in [27, Chapter 3], the coefficients $\Pi_k(\zeta, \chi)$ of the correction series Πu are solutions to the sequence of initial value problems for the ODEs

$$\frac{\partial\Pi_0(\zeta, \chi)}{\partial\zeta} = f(0, \chi, u_0(0, \chi) + \Pi_0(\zeta, \chi)),$$

[4] It is quite natural to consider the inhomogeneity such that its rates of change in time and in the spatial variable have the same order of smallness (see [4] and the end of [5], where Eq. (4.1) is mentioned in connection with a problem in combustion theory).

$$\Pi_0 u(0, \chi) = a(\chi) - u_0(0, \chi),$$

$$\frac{\partial \Pi_k(\zeta, \chi)}{\partial \zeta} = f_u(0, \chi, u_0(0, \chi) + \Pi_0(\xi, \chi))\Pi_k(\zeta, \chi) + G_k(\zeta, \chi),$$

$$\Pi_k(0, \chi) = -u_k(0, \chi), \quad k \geq 1,$$

where $G_k(\zeta, \chi)$ depend on the functions $\Pi_j(\zeta, \chi)$ ($j = 0, 1, \ldots, k-1$). The reasoning in [27] proves that the solutions to this sequence of problems are exponentially small when $\zeta \longrightarrow \infty$. The fact that $G_k(\zeta, \chi)$ includes the term $\Pi_{k-2}(\zeta, \chi)''_{\chi\chi}$ does not change this property at $\zeta \longrightarrow \infty$, since all derivatives of Π are exponentially small when $\zeta \longrightarrow \infty$; indeed, they are solutions to the Cauchy problem for the ODE

$$\frac{\partial g}{\partial \zeta} - f_u(0, \chi, u_0(0, \chi) + \Pi_0(\zeta, \chi))g = H(\zeta, \chi)$$

with exponentially small (at $\zeta \longrightarrow \infty$) right-hand sides (see [27, p. 63]).

4.3. The behavior of solutions to the initial value problem (4.1), (4.2) in the neighborhood of the cusp (τ^*, χ^*) has been analyzed in Sections 2 and 3. "Beyond" this neighborhood, it forms a domain wall. Its asymptotics at $\varepsilon \longrightarrow 0$ is much like the asymptotics (at $t \longrightarrow \infty$) of the special solution to Eq. (2.1) described in Section 3. This is for the following reasons:

(1) The front of the domain wall is localized in the vanishingly small neighborhood of one of the two fold lines originating at (τ^*, χ^*) (the fold lines are the boundaries of the domain where the roots of Eq. (4.3) overlap and in which Eq. (4.3) has three solutions).

(2) "Plateaus" of the domain wall on different sides of the small neighborhood of the shock fold line $\chi = \varphi(\tau)$ are described by two asymptotic expansions (4.5); their leading terms are the minimal, $u_0^1(\tau, \chi)$, and the maximal, $u_0^3(\tau, \chi)$, of the three roots of Eq. (4.3). Being smooth extensions of the unique (at $\tau < \tau^*$) root of this equation through the rays ($\tau = \tau^*$, $\chi > \chi^*$) and ($\tau = \tau^*$, $\chi < \chi^*$), both these roots satisfy the stability condition (4.4) within their smoothness domains.

(3) The asymptotics of the domain wall with respect to the parameter ε in the neighborhood of the shock fold line $\chi = \varphi(\tau)$ is also very similar to the asymptotics (at $t \longrightarrow \infty$) of the solution to Eq. (2.1) in the neighborhood of the curve $x = s_0 t^{3/2}$.

4.4. It should be noted that the structures of these asymptotics have a quite substantial difference. Namely, instead of the solution of Eq. (3.11), which is integrable by quadratures, the behavior of the formed domain wall in the neighborhood of its front in the principal order in ε is described by a monotonic solution to the boundary value problem for a nonintegrable second-order differential equation.

This problem appears upon the change of variable

$$\eta = \frac{\chi - \varphi(\tau)}{\varepsilon} - \mu(\tau, \varepsilon), \tag{4.6}$$

which is similar to (3.11). Changing the variable in Eq. (4.1), we see that the leading term $U(\eta, \tau)$ of the internal asymptotic expansion

$$u_{\text{int}}(\eta, \tau, \varepsilon) = U_0(\eta, \tau) + \ldots$$

of the solution to Eq. (4.1) satisfies the second-order ODE

$$-\varphi_\tau U_\eta - U_{\eta\eta} = f(\tau, \varphi(\tau), U), \tag{4.7}$$

which has exactly two equilibriums

$$U = U_0(\tau), \quad U = U_1(\tau).$$

They are the multiple and, respectively, simple roots of the equation

$$f(\tau, \varphi(\tau), U) = 0, \tag{4.8}$$
$$f_U(\tau, \varphi(\tau), U_0(\tau)) = 0,$$
$$f_U(\tau, \varphi(\tau), U_1(\tau)) < 0.$$

However, the difference described above is not so large, since the monotonic solution $U(\eta, \tau)$ to Eq. (4.7) we are interested in is an analog and some kind of "extension" of the monotonic separatrix solution $R(\sigma)$ to Eq. (3.11); therefore, it is clear that its convergence to the limit value $U_1(\tau)$ is exponential:

$$U - U_1(\tau) = O(\exp(-c(\tau)|\eta|)), \quad c(\tau) < 0. \tag{4.9}$$

Moreover, the asymptotics

$$U = U_0(\tau) + \frac{2\varphi'(\tau)}{f_2(\tau)\eta} - \frac{4}{f_2(\tau)}\left(1 + \frac{f_3(\tau)\varphi'(\tau)^2}{3 f_2^2(\tau)}\right) \times \frac{\ln|\eta|}{\eta^2} + \frac{\Delta(\tau)}{\eta^2} + O\left(\frac{\ln^2|\eta|}{|\eta|^3}\right) \tag{4.10}$$

describes its convergence to the limit value $U_0(\tau)$. Here, $f_2(\tau)$ and $f_3(\tau)$ are the constants appearing in the Taylor series

$$f(\tau, \varphi(\tau), U) = \frac{f_2(\tau)}{2}(U - U_0(\tau))^2 + \frac{f_3(\tau)}{6}(U - U_0(\tau))^3 + \ldots$$

for the right-hand side of Eq. (4.7) at the point $U = U_0(\tau)$, and $\Delta(\tau)$ is an arbitrary function which can be set to zero without loss of generality due to the choice of the yet undetermined function $\mu(\tau, \varepsilon)$ in (4.6).

4.5. Relations (4.9) and (4.10) determine the desired solution to Eq. (4.7) only up to the phase shift $\mu(\tau, \varepsilon)$. Its asymptotics at $\varepsilon \longrightarrow 0$ is constructed in the same

way as asymptotics (3.21). It is based on the existence of the intermediate asymptotic expansion

$$u_{\text{inm}}(\tau, \lambda) = U_0(\tau) + \varepsilon^{1/3}\Upsilon_1(\tau, \lambda) + \varepsilon^{2/3}\Upsilon_2(\tau, \lambda) + \ldots, \tag{4.11}$$

which depends on the scaling variable

$$\lambda = \frac{\chi - \varphi(\tau)}{\varepsilon^{2/3}}. \tag{4.12}$$

Here, $U_0(\tau)$ is the multiple root of Eq. (4.8), and the other coefficients are recurrently determined from the sequence of ODEs that is obtained from Eq. (4.1) by the change of variables (4.12), the substitution of series (4.11) into the resulting equation, and the subsequent equating of the coefficients of equal powers of ε:

$$-\varphi_\tau\frac{\partial\Upsilon_1}{\partial\lambda} = \frac{f_{uu}}{2}\Upsilon_1^2 + f_\chi\lambda, \tag{4.13}$$

$$-\varphi_\tau\frac{\partial\Upsilon_2}{\partial\lambda} = f_{uu}\Upsilon_1\Upsilon_2 + f_{\chi u}\lambda\Upsilon_1 + \frac{\partial^2\Upsilon_1}{\partial\lambda_2} + \frac{f_{uuu}}{6}\Upsilon_1^3 - U_0'(\tau), \tag{4.14}$$

$$\ldots.$$

Here, f_χ, $f_{\chi u}$, f_{uu}, and f_{uuu} are the derivatives of $f(\tau, \chi, u)$ at $\chi = \varphi(\tau)$, $u = U_0(\tau)$.

Using the dilatations

$$\Upsilon_1(\lambda) = -\left(\frac{4f_\chi\varphi_\tau}{f_{uu}^2}\right)^{1/3}\Gamma(\xi),$$
$$\lambda = r(\tau)\xi = -\left(\frac{2\varphi_\tau^2}{f_\chi f_{uu}}\right)^{1/3}\xi, \tag{4.15}$$

we reduce Eq. (4.13) to Eq. (1.18), which has the solution

$$\Gamma(\xi) = (\ln|\text{Ai}(\xi)|)_\xi'. \tag{4.16}$$

This solution is smooth on the interval (ξ_0, ∞), where the left boundary is the first zero of the function Ai(ξ). According to (4.15) and (4.16), we have, at $\lambda \longrightarrow r(\tau)\xi_0$,

$$\Upsilon_1(\xi) = \frac{2\varphi_\tau}{f_{uu}(\lambda - r(\tau)\xi_0)}(1 + O((\lambda - r(\tau)\xi_0)^2)). \tag{4.17}$$

The desired solution to Eq. (4.14) can now be obtained by a procedure similar to that used for solving Eq. (3.15):

$$\Upsilon_2(\xi) = \frac{1}{\text{Ai}(\xi)^2}\left[\int_0^\xi\left(\text{Ai}(\kappa)^2H(\tau, \kappa) - \frac{\Omega(\tau)}{\kappa - \xi_0}\right)d\kappa + \Omega(\tau)\ln\left|\frac{\xi - \xi_0}{\xi_0}\right| - \int_0^\infty\text{Ai}(\kappa)^2H(\tau, \kappa)d\kappa\right]. \tag{4.18}$$

Here,

$$H(\tau, \kappa) = \left(\frac{2\varphi_\tau^2}{f_\chi f_{uu}}\right)^{1/3}\left[\frac{2f_{\chi u}}{f_{uu}}\kappa\frac{\text{Ai}'(\kappa)}{\text{Ai}(\kappa)} - \frac{f_\chi}{\varphi_\tau^2}\left(\frac{\text{Ai}'(\kappa)}{\text{Ai}(\kappa)}\right)'' - \frac{2f_\chi f_{uuu}}{3f_{uu}^2}\left(\frac{\text{Ai}'(\kappa)}{\text{Ai}(\kappa)}\right)^3 - U_0'(\tau)\right],$$

and $\Omega(\tau)$ is the residue of the function $\text{Ai}(\xi)^2H(\tau, \xi)$ at the point ξ_0:

$$\Omega(\tau) = -2\text{Ai}'(\xi_0)^2\left(\frac{2f_\chi^2}{\varphi_\tau^4 f_{uu}}\right)^{1/3}\left(1 + \frac{\varphi_\tau^2 f_{uuu}}{3f_{uu}^2}\right).$$

Using representation (4.18) for Υ_2, we see that, at $\lambda \longrightarrow r(\tau)\xi_0$,

$$\Upsilon_2(\xi) = \left[-\frac{4}{f_{uu}}\left(1 + \frac{\varphi_\tau^2 f_{uuu}}{3f_{uu}^2}\right)\frac{\ln|\lambda - r(\tau)\xi_0|}{(\lambda - r(\tau)\xi_0)^2} + \frac{h(\tau)}{(\lambda - r(\tau)\xi_0)^2}\right](1 + O((\lambda - r(\tau)\xi_0)^2)), \tag{4.19}$$

where

$$h(\tau) = \frac{4\ln|r(\tau)\xi_0|}{f_{uu}}\left[1 + \frac{\varphi_\tau^2 f_{uuu}}{3f_{uu}^2}\right] + \frac{1}{\text{Ai}'(\xi_0)^2}\left(\frac{2\varphi_\tau^2}{f_\chi f_{uu}}\right)^{2/3} \times\left[\int_0^{\xi_0}\left(\text{Ai}(\kappa)^2H(\tau, \kappa) - \frac{\Omega(\tau)}{\kappa - \xi_0}\right)d\kappa - \int_0^\infty\text{Ai}(\kappa)^2H(\tau, \kappa)d\kappa\right].$$

The terms of the intermediate asymptotic expansion (4.11) have an increasing singularity at $r(\tau)\xi_0$, where this expansion is invalid: in the neighborhood of $\lambda = r(\tau)\xi_0$, the expansion $u_{\text{int}}(\nu, \tau, \varepsilon)$ is the correct approximation. The matching requirement for expansions $u_{\text{int}}(\nu, \tau, \varepsilon)$ and $u_{\text{inm}}(\lambda, \tau, \varepsilon)$ and relations (4.10), (4.17), and (4.19) allows us to find the following asymptotics of $\mu(\tau, \varepsilon)$ at $\varepsilon \longrightarrow 0$:

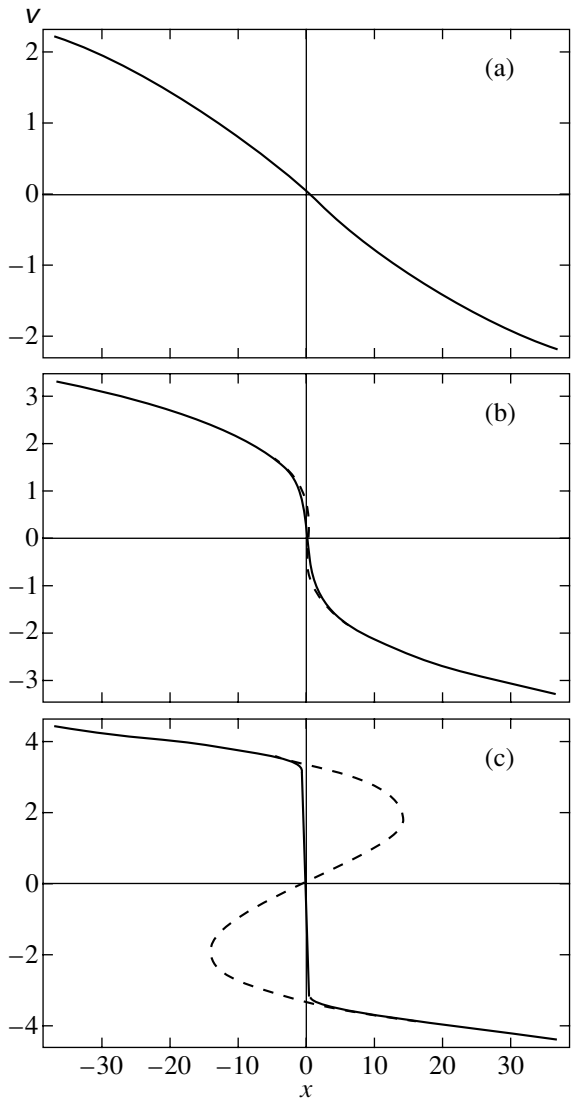


Fig. 3. The solution v to the first equation in (2.2) as a function of x at various t (solid curves) and the plots of the roots of the cubic equation (2.3) where they differ from the plots of v (dashed curve) at the instances of time $t = -11$ (a), 0 (b), and 11 (c).

$$\mu(\tau, \varepsilon) = \frac{r(\tau)\xi_0}{\varepsilon^{1/3}}$$

$$-\frac{2}{3\varphi_\tau}\left(1 + \frac{\varphi_\tau^2 f_{uuu}}{3 f_{uu}^2}\right)\ln\varepsilon + \frac{f_{uu}h(\tau)}{2\varphi_\tau} + o(1).$$

Thus, we have obtained a description of the formed domain wall in the principal order.

5. SPECIAL SOLUTIONS TO THE ODE (2.2)

5.1. Figure 3 presents a part of the numerical computation results which were performed by an iterative procedure similar to that described in Subsection 3.2. These computations simulate the behavior of the special solution to the first ODE in (2.2) at various t. It is seen from these computations that, in addition to the monotonicity and oddness of this solution, its asymptotics at $x^2 + t^2 \longrightarrow \infty$ outside the unique jump curve is also specified by the smoothly changing root of Eq. (2.3). However, in the case under consideration, the jump in the asymptotics $v(x, t)$ at $t \longrightarrow \infty$ occurs in the neighborhood of the ray ($x = 0$, $t \geq 0$), which forms the so-called Maxwell set [17, p. 306] for the cusp catastrophe described by Eq. (2.3). Hence,[5] if the PDE (1.1) does not include the first-order derivatives of $u(S)$, one can expect that the corresponding solutions to Eq. (1.8) behind the condensation point of the roots of Eq. (1.10) take the form of domain walls with fronts localized near the curves defined by the condition of equality of areas (1.17).

The change of variables (3.1) reduces the first equation in (2.2) to the ODE

$$t^{-4}r_{ss} = r^3 - \mathrm{sgn}(t)r + s. \qquad (5.1)$$

The leading terms of the asymptotic expansions of the corresponding solutions to this equation at $t \longrightarrow \infty$,

$$r_{\mathrm{out}}(s, t) = r_0(s) + \sum_{j=1}^{\infty} t^{-4j} r_j(s), \qquad (5.2)$$

coincide with the maximal, $r_{\max}(s)$ (for $s < 0$), and the minimal, $r_{\min}(s)$ (for $s > 0$), roots of the cubic equation (3.7). The standard reasoning used in the matching method makes it is easy to show that, in the small neighborhood of $s = 0$, these two parts of expansion (5.2) are joined by the internal expansion

$$r_{\mathrm{int}}(y, t) = p_0(y) + \sum_{j=1}^{\infty} t^{-2j} p_j(y), \quad y = st^2. \qquad (5.3)$$

Its coefficients satisfy the recurrent sequence of ODEs obtained by the substitution of (5.3) into (5.1) and equating the coefficients of equal powers of the small parameter t^{-2}. These equations are supplemented by the matching conditions of the behavior of their solutions at $y \longrightarrow \pm\infty$ with the behavior of the external asymptotic expansions (5.2) at $s \longrightarrow 0$. In particular, for the odd (in y) leading term of expansion (5.3), the ODE

$$(p_0)''_{yy} = p_0^3 - p_0$$

is complemented by the condition

$$\lim_{y \to \infty} p_0(y) = -1.$$

[5] The Maxwell set is distinguished by the fact that the integral $\int_{g_1(x,t)}^{g_3(x,t)} (x - tg + g^3)dg$, with the limits defined by two different roots of Eq. (2.3), vanishes at this set.

Hence, we derive that

$$p_0(y) = -\tan\frac{y}{\sqrt{2}}.$$

When $t \longrightarrow -\infty$, the solution to Eq. (2.2) also has an asymptotic expansion defined by the change of variables (3.1) and series (5.2). However, in this case, the leading term $r_0(s)$ of series (5.2), for all s, is given by the unique (at $t < 0$) root of Eq. (2.7). The following series give its asymptotic expansions at $|x| \longrightarrow \infty$, which are suitable in the neighborhood of the line $t = 0$:

$$v(x, t) = x^{1/3}\sum_{j=0}^{\infty} x^{-8j/3} g_j(\rho), \quad \rho = \frac{t}{x^{2/3}}. \qquad (5.4)$$

Here, $g_0(\rho)$ is the unique root of the equation

$$1 - \rho g + g^3 = 0.$$

The function $v(x, t)$ is an analog of the Il'in solution to the Burgers equation [15, p. 287], which describes the influence of small dissipation on the nucleation of shock waves [15, 28]. Our solution to Eq. (2.2) gives a similar description of the nucleation of domain walls formed by solutions to the boundary value problems (1.15), (1.16).

5.2. We can formulate no general position boundary value problems that lead to the corresponding special solutions to the second equation in (2.2). Also, no description of their uniform asymptotics at $x^2 + t^2 \longrightarrow \infty$ is available. Numerical simulation suggests that solutions to the second equation in (2.2) at $t \gg 1$ are characterized by high-frequency oscillation domains, which can be analyzed using the Kuzmak–Whitham averaging [29, 30].

6. CONCLUSIONS

The main conclusion formulated at the end of the Introduction is as yet purely theoretical. However, the fact that our considerations are rather general gives reason to expect that corresponding experimental phenomena will be discovered. Indeed, the formation of structures predicted in this paper is as natural as, for example, the spontaneous formation of shock waves in gas dynamics.

In this connection, the most interesting equation of form (1.8) discussed in this paper is the PDE (1.14). As has already been noted, it is equivalent to the time-independent part of the diffusion equation (1.4), (1.7). True enough, it is difficult to completely analyze the relationship of the main conclusion of this paper about the solutions to Eq. (1.14) to solutions to the time-dependent Eq. (1.4), (1.7) (see the related paper [13]). However, certain sufficient conditions that ensure that solutions to, say, initial value time-dependent problem (1.13) approach the corresponding asymptotic solutions to the time-independent Eq. (1.14) can be easily formulated just now.

Assume that there is a unique cusp of solutions to Eq. (1.10) on the plane X. Moreover, we assume that there exists a unique root of Eq. (1.10) outside the overlapping domain and it satisfies the stability condition (1.12). It is clear that this root is the attractor for all solutions to the ODE

$$u_\zeta = f(X, u).$$

which is the limit equation to Eq. (1.14). Assume that, within the overlapping domain (1.10), the initial function $g(X)$ in problem (1.13) is such that the solution to the limit initial value problem

$$u_\zeta = f(X, u), \quad u|_{\zeta = 0} = g(X),$$

is attracted, within this domain, to the root of Eq. (1.10) that determines one of the "plateaus" of the domain wall (the solution to Eq. (1.14)). It is clear that, for certain (sufficiently large) ζ, this solution in the form of series (1.9) is an asymptotics at $\varepsilon \longrightarrow 0$ for the solution to problem (1.13). In order for the corresponding asymptotic solution to Eq. (1.14) at such ζ to be a solution to the time-dependent problem (1.13) at every X, it is sufficient to take a smooth initial function $g(X)$ that is either less or greater (depending on the situation) than any root of Eq. (1.10).

ACKNOWLEDGMENTS

This work was supported by the Russian Foundation for Basic Research (project no. 00-01-00663) and by the Major Scientific Schools Foundation (project no. 00-15-96038). I am grateful to V.E. Adler for his help in computations.

REFERENCES

1. Ya. B. Zel'dovich and A. D. Myshkis, *Elements of Mathematical Physics* (Nauka, Moscow, 1973).
2. Yu. M. Svirezhev, *Nonlinear Waves, Dissipative Structures and Catastrophe in Ecology* (Nauka, Moscow, 1987).
3. A. Vl. Gurevich and R. G. Mints, Usp. Fiz. Nauk **142**, 61 (1984) [Sov. Phys. Usp. **27**, 19 (1984)].
4. L. A. Ostrovskiĭ, Izv. Vyssh. Uchebn. Zaved., Radiofiz. **17**, 395 (1973).
5. V. S. Berman, Prikl. Mat. Mekh. **42**, 450 (1978).
6. I. A. Molotkov and S. A. Vakulenko, *Concentration of Nonlinear Waves* (Leningr. Gos. Univ., Leningrad, 1988).
7. I. A. Molotkov, S. A. Vakulenko, and M. A. Bisyarin, *Nonlinear Localizated Wave Processes* (Yanus-K, Moscow, 1999).
8. N. N. Rozanov, Zh. Éksp. Teor. Fiz. **80**, 96 (1981) [Sov. Phys. JETP **53**, 47 (1981)].
9. V. P. Maslov, V. G. Danilov, and K. A. Volosov, *Mathematical Simulation of Heat-and-Mass Transfer. Evolution of Dissipative Structures* (Nauka, Moscow, 1987).
10. T. N. Sysoeva, Vestn. Mosk. Univ., Ser. 1: Mat., Mekh., No. 1, 32 (1983).

11. P. Fife and U. Grinli, Usp. Mat. Nauk **29**, 103 (1974).
12. H. H. Nefedov, Diff. Ur. **31**, 1142 (1995).
13. V. F. Butuzov and I. V. Nedel'ko, Diff. Ur. **38**, 222 (2001).
14. J. Colle, *Perturbation Methods in Applied Mathematics* (Blaisdell Publishing Company, Waltham, 1968; Mir, Moscow, 1972).
15. A. M. Il'in, *Matching of Asymptotic Expansions of Solutions of Boundary-Value Problems* (Nauka, Moscow, 1987).
16. M. V. Fedoryuk, *Ordinary Differential Equations* (Nauka, Moscow, 1985).
17. T. Poston and M. Stewart, *Catastrophe Theory and Its Applications* (Pitman, London, 1978; Mir, Moscow, 1980).
18. R. Gilmore, *Catastrophe Theory for Scientists and Engineers* (Wiley, New York, 1981; Mir, Moscow, 1984), Part 1.
19. R. Haberman, Stud. Appl. Math. **57**, 247 (1977).
20. R. Haberman, SIAM J. Appl. Math. **37**, 69 (1979).
21. D. C. Diminie and R. Haberman, J. Nonlinear Sci. **10**, 198 (2000).
22. O. M. Kiselev, J. Nonlinear Math. Phys. **8**, 65 (2001).
23. O. M. Kiselev and S. G. Glebov, Russ. J. Math. Phys. **9**, 21 (2002).
24. *Differential Equations with a Small Parameter and Relaxation Oscillations*, Ed. by E. F. Mishchenko and N. Kh. Rozov (Nauka, Moscow, 1975; Plenum, New York, 1980).
25. A. A. Dorodnitsyn, Prikl. Mat. Mekh. **11**, 313 (1947).
26. *Handbook of Mathematical Functions*, Ed. by M. Abramowitz and I. A. Stegun (National Bureau of Standards, Washington, 1964; Nauka, Moscow, 1979).
27. A. B. Vasil'eva and V. F. Butuzov, *Asymptotic Expansions of Solutions of Singular Perturbed Equations* (Nauka, Moscow, 1973).
28. V. R. Kudashev and B. I. Suleĭmanov, Prikl. Mat. Mekh. **65**, 458 (2001).
29. G. E. Kuzmak, Prikl. Mat. Mekh. **23**, 216 (1959).
30. M. V. Fedoryuk, Zh. Vychisl. Mat. Mat. Fiz. **26**, 198 (1986).

Translated by A. Klimontovich

Journal of Experimental and Theoretical Physics, Vol. 95, No. 5, 2002, pp. 957–964.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 122, No. 5, 2002, pp. 1107–1115.

NONLINEAR PHYSICS

Calculation of the Diffusion Coefficient in Acoustic Turbulence

N. A. Silant'ev

Instituto Nacional de Astrofísica, Optica y Electrónica, C. P. 72000, Pue. Mexico
Pulkovo Astronomical Observatory, Russian Academy of Sciences, Pulkovskoe sh. 65, St. Petersburg, 196140 Russia
e-mail: silant@inaoep.mx

Received June 5, 2002

Abstract—The first (Born) approximation commonly used to calculate the diffusion coefficient D_T of a passive scalar in acoustic turbulence is shown to be insufficient. Even for a small main parameter—the Mach number, $M \ll 1$—the next approximation gives a larger contribution to D_T than does the first approximation, but negative in sign. We present a procedure for correctly calculating D_T based on the solution of a nonlinear DIA (direct interaction approximation) equation for the mean Green's function of the problem. We include an additional term in the general formula for D_T that directly describes the compressibility of acoustic turbulence. This term has not been known previously and has been disregarded even in the Born approximation. A positive value was obtained for $D_T = CM^3 u_0/p_0$. The spectrum $E(x)$ was assumed to be smooth at distances $\Delta x \sim M^2 \ll 1$. © *2002 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

The diffusion of passive scalar particles and fields in a turbulent medium is a major problem in the theory of turbulence, which is important in practical terms. For the model of turbulence in an unbounded medium, there are exact formulas for the turbulent diffusion coefficient D_T in the Lagrangian [1, 2] and Eulerian [3, 4] representations. From a practical point of view, of particular importance are calculations in the Eulerian representation, which is used below. In this representation, an exact calculation of D_T involves determining the stochastic Green's function of the problem, $G(\mathbf{r}, t; \mathbf{r}', t')$, and its subsequent statistical averaging with the components of velocity field $\mathbf{u}(\mathbf{r}, t)$ (see [4] for more details).

By acoustic turbulence, we mean a medium with chaotically propagating acoustic waves. The parameters of such a medium—the gas-velocity correlators

$$\langle u_i(\mathbf{r}, t) u_j(\mathbf{r}', t') \rangle \equiv B_{ij}(\mathbf{R}, \tau), \quad \mathbf{R} = \mathbf{r} - \mathbf{r}', \quad \tau = t - t' \tag{1}$$

—are determined by the stochastic boundary conditions and by the sources that maintain continuous wave formation. Such turbulence is rarely encountered in pure form (e.g., in the solar corona). However, it is commonly used in theoretical works on magnetic dynamo (see [5–7] and references therein) as an example of simple turbulence in which analytic calculations can be performed to completion with a small main parameter—the Mach number, $M = u_0/c$ (here, $u_0^2 = B_{ii}(0, 0)$ is the mean square of the oscillation velocity amplitude and c is the speed of sound).

The first (Born) approximation is used in all these works to calculate the turbulent diffusion coefficient D_T, because the next approximation contains an additional small factor, $M^2 \ll 1$. The standard Green's function

$$G_m(R, \tau) = (4\pi D_m \tau)^{-3/2} \exp\left(-\frac{R^2}{4D_m\tau}\right),$$

which describes molecular diffusion with coefficient D_m in a gas at rest, is used as the Green's function $G(R, \tau)$. Since D_m is small, the limiting form $G_m(R, \tau) \longrightarrow \delta(\mathbf{R})$ is considered.

In addition, the authors used an incomplete formula to calculate D_T that contained no contribution from the compressibility proper (the correlator $\langle u_i \mathrm{div}\,\mathbf{u} \rangle$). In [4], we showed that including this additional term significantly increases D_T even in the Born approximation (if the damping coefficient $k(p)$ of an acoustic wave is of the order of D_m or D_T).

In this paper, we pointed out that, to assess whether it is correct to use the Born approximation in calculating D_T, it is necessary to determine the Green's function more accurately and to estimate the contribution from the next approximation containing the fourth-order velocity correlators to D_T. Here, we carry out this program.

First, we show that the contribution from the fourth-order correlators to D_T, which we denote by $D_T^{(1)}$ (see formula (19) in [4]), is slightly larger than the contribution $D_T^{(0)}$ from the pair correlator (1). Moreover, $D_T^{(1)}$ is negative and, hence, the total turbulent diffusion

coefficient $D_T^{(0)} + D_T^{(1)}$ is negative. We emphasize that this remains valid even if the function $G_m(R, \tau)$ with the diffusion coefficient D_T is used. It is only important that the damping coefficient k be much larger than D_T, which probably always holds. Thus, if purely diffusion functions are used as the Green's functions, then the total turbulent diffusion coefficient is negative. In this case, allowance for compressibility is unimportant if $k \gg D_T$. Note that the estimate $D_T^{(1)} \approx M^2 D_T^{(0)} \ll D_T^{(0)}$ from [6] is incorrect, because it disregards the resonant nature of the contribution from acoustic harmonics in the expression for $D_T^{(1)}$.

Why is the derived turbulent diffusion coefficient negative in the first place? It is negative apparently because the Green's function was chosen wrongly. When acoustic turbulence is considered, the Green's function must reflect the main property of passive scalar transport—the mainly oscillatory motion of passive scalar particles. Even the form of the velocity correlator (1) in the space of wave numbers (see [7]),

$$B_{nm}(\mathbf{R}, \tau) = \frac{1}{(2\pi)^3}\int d\mathbf{p} \exp(i\mathbf{p}\cdot\mathbf{R})\tilde{B}_{nm}(\mathbf{p}, \tau), \quad (2)$$

$$\tilde{B}_{nm}(\mathbf{p}, \tau) = 2\pi^2 p_n p_m \frac{E(p)}{p^4}\cos(cp\tau) \times \exp[-k(p)p^2\tau], \quad (3)$$

which contains the oscillating factor $\cos(cp\tau)$, shows that the Fourier transform of the mean Green's function $\tilde{G}(p, \tau)$ must also contain oscillating terms. Here, $k(p)$ is the damping coefficient of the p wave and $E(p)$ is the acoustic turbulence spectrum determined from the expression

$$u_0^2 \equiv \langle u^2(\mathbf{r}, t)\rangle = \int_0^\infty dp E(p). \quad (4)$$

Recall that the Fourier transform of the diffusion Green's function $\tilde{G}(p, \tau) = \exp(-Dp^2\tau)$ contains no oscillating terms.

The damping of an isolated acoustic wave is attributable to viscosity in the medium and $k(p)$ is small, of the order of the molecular diffusion coefficient D_m (see [8]). In an ensemble of interacting waves, the situation is different. Zakharov and Sagdeev [9] developed a model of acoustic turbulence with nonlinear wave interaction. They derived the following expression for $k(p)$: $k(p) = E(p)/c$, where $E(p) \approx \text{const}\, u_0^2 \sqrt{p_0}\, p^{-3/2}$ for $p > p_0$ and $\lambda_0 = 1/p_0$ is the characteristic wavelength. When $p \longrightarrow 0$, $E(p)$ tends to zero. Using the dimensionless wave number $x = p/p_0$, we introduce a dimensionless turbulence spectrum by means of the relation $E(p) \equiv u_0^2/p_0 E(x)$. In this notation, the result of [9] can be written as

$$E(x) = \text{const}\, x^{-3/2}, \quad k(p) = ME(x)u_0\lambda_0.$$

The damping $k(p)$ for a small Mach number, $M = u_0/c \ll 1$, is weak, being smaller than $u_0\lambda_0$. However, as we will see below, the turbulent diffusion coefficient is even smaller, of the order of $M^3 u_0\lambda_0$.

Expression (3) is the pair gas-velocity correlator for a homogeneous, isotropic, and stationary ensemble of acoustic harmonic waves. Substituting (3) into (2) yields an expression for the scalar product of the velocities at the same place but at different times:

$$\langle \mathbf{u}(\mathbf{r}, t)\cdot\mathbf{u}(\mathbf{r}, t')\rangle = \int_0^\infty dp E(p)\cos(cp\tau)\exp[-k(p)p^2\tau]. \quad (5)$$

Here, $\tau = t - t'$. We see from (5) that such a statistical ensemble describes a nonperiodic, chaotic gas motion even if wave damping is ignored ($k = 0$); i.e., the turbulent diffusion of matter takes place simply because of the superposition of chaotic, incoherent acoustic harmonic waves. It is only important that the wave spectrum be continuous.

If damping is ignored, an ensemble of waves of the same wavelength ($E(p) = u_0^2\delta(p - p_0)$) produces a periodic overall motion, which may be very intricate in space. In this case, the diffusion mixing of particles is weak and attributable to viscous (dynamical) damping: here, there is no kinematic diffusion similar to a macroscopic random walk. Below, we consider the diffusion of passive scalar particles only in a turbulent medium with a continuous wave spectrum, which is clearly most natural.

The fact that $D_T \sim M^3 u_0\lambda_0 \ll u_0\lambda_0$ reflects an important property of acoustic turbulence: the motion of a passive scalar particle is mainly oscillatory and does not lead to diffusion. If the particle passed to a different trajectory corresponding to a different wave during each oscillation period, then, clearly, the diffusion coefficient would be much larger, of the order of $u_0\lambda_0$. Thus, the actual Green's function $\tilde{G}(p, \tau)$ must describe these oscillations, and the diffusion functions $\tilde{G}(p, \tau) = \exp(-Dp^2\tau)$ are unsuitable for calculating D_T. For nonacoustic turbulence, the so-called self-consistent method [10] is efficient in calculating D_T. In this method, the diffusion function

$$\tilde{G}(p, \tau) = \exp(-D_s p^2\tau)$$

with the unknown diffusion coefficient D_s is substituted for the unknown Green's function in the exact formula for D_T. Solving the resulting algebraic equation yields D_s that closely matches the D_T values calculated by

more accurate methods. The success of the self-consistent method stems from the fact that the diffusion Green's function is close to the exact Green's function. For acoustic turbulence, the self-consistent method yields large negative D_T, which is clearly contrary to fact. This also suggests that the actual Green's function for the time scales that determine the transport mechanism of passive scalar particles is far from being a purely diffusion one.

Below, we use the solution to a nonlinear integral equation for the Green's function to calculate D_T. This equation actually has oscillating terms.

2. BASIC EQUATIONS

We take the number density of passive scalar particles, $n(\mathbf{r}, t)$, as the passive scalar field. The continuity equation for $n(\mathbf{r}, t)$ in the random velocity field $\mathbf{u}(\mathbf{r}, t)$ of the main gas is a stochastic equation to determine $n(\mathbf{r}, t)$:

$$\left(\frac{\partial}{\partial t} - D_m\nabla^2\right)n(\mathbf{r}, t) = -\mathrm{div}[\mathbf{u}(\mathbf{r}, t)n(\mathbf{r}, t)]. \quad (6)$$

Assume that the statistical ensemble of the field $\mathbf{u}(\mathbf{r}, t)$ is homogeneous, isotropic, and stationary and that it is described by correlator (3). The mean $\langle\mathbf{u}\rangle = 0$.

The Green's function $G(1; 2) \equiv G(\mathbf{r}, t; \mathbf{r}', t')$ of Eq. (6) satisfies the linear equation

$$G(1; 2) = G_m(1-2) - \int d3\, G_m(1-3)\nabla_i^{(3)}[u_i(3)G(3; 2)]. \quad (7)$$

Below, we use the following convenient notation: $f(1) = f(\mathbf{r}_1, t_1)$, $f(1-2) = f(\mathbf{r}_1 - \mathbf{r}_2, t_1 - t_2) \equiv f(\mathbf{R}, \tau)$, $dn = d\mathbf{r}_n dt_n$, etc.

For large times and scales, averaging of Eq. (6) is known (see [3, 4]) to result in the diffusion equation [the left-hand side of (6)] for the mean number density $\langle n(\mathbf{r}, t)\rangle$ with the diffusion coefficient $D_m + D_T$. The exact formula for the stationary turbulent diffusion coefficient is

$$D_T = \frac{1}{3}\int d\mathbf{R}\int_0^\infty d\tau[\langle u_i(1)G(1; 2)u_i(2)\rangle - \mathbf{R}\cdot\langle\mathbf{u}(1)G(1; 2)\mathrm{div}\mathbf{u}(2)\rangle]. \quad (8)$$

For incompressible turbulence ($\mathrm{div}\mathbf{u} = 0$), the second term in (8) vanishes. It is this significant term that was disregarded in [6, 7] for compressible acoustic turbulence. The authors of these papers restricted themselves to substituting the free term $G_m(1-2)$ of Eq. (7) for $G(1; 2)$ into the first term of expression (8). As we show below, substituting several iterations of Eq. (7) into (8) yields the relation $D_T^{(0)} + D_T^{(1)} < 0$, i.e., a definitely absurd result. The reason why substituting the iterations of Eq. (7) when calculating D_T is inefficient was explained in detail in the Introduction.

In [3, 11], we derived a new, renormalized equation for the Green's function $G(1; 2)$ instead of Eq. (7), which includes the mean Green's function $\langle G(1; 2)\rangle \equiv G(1-2)$ as a free term. Clearly, $G(1-2)$ correctly describes the convective-oscillatory transport of passive scalar particles in acoustic turbulence; it is hoped that substituting iterations of the renormalized equation into (8) will yield the correct value for D_T. Thus, the problem reduces to writing and solving the equation for the mean Green's function $G(1-2)$.

It follows from Eq. (7) that $G(1-2)$ depends on the fluctuational part $G'(1; 2)$ of the Green's function and vice versa. Therefore, an attempt to write a separate equation only for $G(1-2)$ leads to an hierarchy of nonlinear equations for $G(1-2)$ (the situation is similar to the closure problem in the theory of turbulence). The simplest equation of this hierarchy, with a quadratic nonlinearity, is called the DIA (direct interaction approximation) equation. It was first written and analyzed in [12]:

$$G(1-2) = G_m(1-2) + \int d3\int d4\, G_m(1-3)\times\nabla_i^{(3)}G(3-4)\nabla_j^{(4)}B_{ij}(3-4)G(4-2). \quad (9)$$

Subsequently, it emerged that this equation was efficient in calculating the turbulent diffusion coefficients D_T for various models of incompressible turbulence (see [10]). Our results confirm that it is also efficient for acoustic turbulence.

In calculations, it is convenient to use the function $\tilde{g}(p, s)$—the Fourier transform in $\mathbf{R}$ and the Laplace transform in τ of $G(R, \tau)$. In addition, we make use of the dimensionless variables $x = p/p_0$ and $t = cp_0\tau$. In these variables, the DIA equation for $\tilde{g}(x, s)$ takes the form

$$\tilde{g}(x, s) = \left[s + \frac{x^2}{\gamma} + \frac{M^2}{2}\int_0^\infty dy E(y)\times\int_{-1}^1 d\mu\,\mu x(\mu x - y)\times\int_0^\infty dt e^{-(s+\eta y^2)t}\cos(yt)\tilde{G}(|\mathbf{x}-\mathbf{y}|, t)\right]^{-1}. \quad (10)$$

Here, we introduced the dimensionless quantities $\gamma = c/p_0D_m$ and $\eta(y) = k(y)p_0/c$; μ is the cosine of the angle between vectors $\mathbf{x}$ and $\mathbf{y}$. Recall the definition of the dimensionless turbulence spectrum: $E(p) = u_0^2E(x)/p_0$. Since the molecular diffusion is weak, $\gamma \approx \lambda_0/l \gg 1$ (l is the mean free path of the gas molecules; $D_m \sim cl$). This is the largest parameter of the problem. Even the gas-

dynamical equations themselves suggest that $\lambda_0 \gg l$. If we use the damping coefficient $k(p)$ from [9], then $\eta(y) = M^2E(y)$; i.e., this is a small quantity but much larger than $1/\gamma$ if, of course, the Mach number is moderately large, which we assume below. The case of a small Mach number, $\gamma M^2 \ll 1$, describes a fluid that is virtually at rest. In this limit, the integral term in (10) can be discarded to give

$$\tilde{g}(x, s) \approx \frac{1}{s + x^2/\gamma};$$

i.e., the Green's function actually matches G_m.

It is easy to see that for $x \ll 1$ and $s \ll 1$, the integral term in (10) transforms into the diffusion expression $(D_T^{(0)}/D_m)(x^2/\gamma)$ with the turbulent diffusion coefficient

$$D_T^{(0)} = \frac{u_0 M}{3p_0}\int_0^\infty dx E(x)\int_0^\infty dt \cos(yt)e^{-\eta(x)x^2 t} \times\left[\tilde{G}(x, t) + x\frac{\partial\tilde{G}(x, t)}{\partial x}\right]. \quad (11)$$

Recall that $\tilde{G}(x, t)$ is the Fourier transform of the Green's function $G(\mathbf{R}, \tau)$ in $\mathbf{R}$ in dimensionless variables x and t. Expression (11) can also be derived from the general formula (8) if we take the DIA expression for $G(\mathbf{R}, \tau)$ as $G(1; 2)$ and pass to the Fourier representation.

The condition $s \ll 1$ implies that $\tau \gg 1/cp_0 = T_0$, where T_0 is the characteristic oscillation period of acoustic waves. In this case, the diffusion propagation of a passive scalar is established after many gas oscillations rather than after the characteristic velocity-correlation damping time, $\tau_{\text{damp}} \sim 1/k(p_0)p_0^2$, which is much longer than $T_0(\tau_{\text{damp}} \sim T_0/M^2)$ and tends to infinity as $M^2 \longrightarrow 0$. As we mentioned above, the diffusion is mainly produced by a nonperiodic gas motion, which results from the superposition of chaotic harmonic acoustic waves with a broad turbulence spectrum $E(p)$. Therefore, the time it takes for diffusion to be established is unrelated to the velocity-correlation damping time.

3. CALCULATING THE DIFFUSION COEFFICIENT

To calculate $D_T^{(0)}$ and the contribution of the fourth-order correlators $D_T^{(1)}$ (see formula (19) from [4]), it is convenient to introduce the following auxiliary functions:

$$\tilde{g}_c(x, p, q) = \int_0^\infty dt \exp(-pt)\cos(qt)\tilde{G}(x, t) \equiv \mathrm{Re}\tilde{g}(x, p - iq),$$
$$\tilde{g}_c(x, p, q) = \int_0^\infty dt \exp(-pt)\sin(qt)\tilde{G}(x, t) \equiv \mathrm{Im}\tilde{g}(x, p - iq). \quad (12)$$

Expression (11) for $D_T^{(0)}$ can then be written as

$$D_T^{(0)} = \frac{u_0 M}{3p_0}\int_0^\infty dx E(x) \times\left[\tilde{g}_c(x, p, q) + x\frac{\partial\tilde{g}_c(x, p, q)}{\partial x}\right]_{p = \eta(x)x^2,\, q = x}. \quad (13)$$

Assuming that $s = p - iq$ in (10) and separating out the real and imaginary parts yields

$$\tilde{g}_c(x, p, q) = \frac{\alpha(x, p, q)}{\alpha^2(x, p, q) + \beta^2(x, p, q)},$$
$$\tilde{g}_s(x, p, q) = \frac{\beta(x, p, q)}{\alpha^2(x, p, q) + \beta^2(x, p, q)}, \quad (14)$$

$$\alpha(x, p, q) = p + \frac{x^2}{\gamma} + \frac{M^2}{4}\int_0^\infty dy E(y)\int_{-1}^1 d\mu\mu x(\mu x - y) \times[\tilde{g}_c(|\mathbf{x} - \mathbf{y}|, p + \eta(y)y^2, q + y) + \tilde{g}_c(|\mathbf{x} - \mathbf{y}|, p + \eta(y)y^2, q - y)],$$
$$\beta(x, p, q) = q - \frac{M^2}{4}\int_0^\infty dy E(y)\int_{-1}^1 d\mu\mu x(\mu x - y) \times[\tilde{g}_s(|\mathbf{x} - \mathbf{y}|, p + \eta(y)y^2, q + y) + \tilde{g}_s(|\mathbf{x} - \mathbf{y}|, p + \eta(y)y^2, q - y)]. \quad (15)$$

The system of equations (14) and (15) is equivalent to the DIA equation (10). Its advantage is that it allows the diffusion coefficients $D_T^{(0)}$ and $D_T^{(1)}$ to be calculated directly, without calculating the Green's function $\tilde{G}(x, t)$ itself. In this case, we need not know the functions $\tilde{g}_c(x, p, q)$ and $\tilde{g}_s(x, p, q)$ for all possible values of the variable p—expression (13) includes only

$$p = \eta(x)x^2 = M^2E(x)x^2 \ll 1.$$

It is easy to verify that $\alpha(x, p, q)$ is small, of the order of $M^2 \ll 1$, for small p. In contrast, $\beta(x, p, q) \sim q$ contains no small parameter in the first approximation. For

small p, the asymptotic solution of system (14) and (15) is

$$\alpha(x, p, q) = p + \frac{\pi}{6}M^2x^2E(q) + O(M^4),$$
$$\beta(x, p, q) = q - \frac{M^2}{6}x^2\int_0^\infty dy\frac{E(y)}{q+y} + O(M^4). \quad (16)$$

Here, we discarded the extremely small term x^2/γ. Since α is small compared to β, we can assume that $\tilde{g}_c(x, p, q) \approx \pi\delta(q)$ in our intermediate calculations for sufficiently smooth spectra $E(x)$. However, $\tilde{g}_c$ is no longer a δ-shaped function in the main formula (13), where $\alpha \sim x^2$ and $\beta \sim x$. The spiky form of $\tilde{g}_c$ suggests that the diffusion in acoustic turbulence is resonant in nature—waves close in frequency give the largest contribution to the diffusion. Therefore, we find from (15) that α is proportional to $E(q)$ and that the diffusion coefficient itself [see (18)] is proportional to the square of the spectrum.

The first terms in (16), i.e., $\alpha = p$ and $\beta = q$, correspond to the Born approximation used in [6, 7]. Substituting these values into (13) yields

$$D_T^{(\mathrm{Born})} = \frac{u_0M}{3p_0}\int_0^\infty dxE(x)\eta(x) \longrightarrow \frac{u_0M^3}{3p_0}\int_0^\infty dxE^2(x). \quad (17)$$

The expression after the arrow corresponds to the model from [9], in which $k(p) = E(p)/c$. Note that retaining the discarded term x^2/γ gives an addition M^2D_m, with 2/3 of this addition arising from the compressibility of acoustic turbulence.

Using the DIA expression (16), we obtain

$$D_T^{(0)} = \frac{u_0M}{3p_0}\int_0^\infty dxE(x)\left[\eta(x) + \frac{\pi}{2}M^2E(x)\right] \longrightarrow \frac{u_0M^3}{3p_0}\left(1 + \frac{\pi}{2}\right)\int_0^\infty dxE^2(x). \quad (18)$$

Comparison of (17) and (18) indicates that $D_T^{(0)}$ is approximately a factor of 2.5 larger than $D_T^{(\mathrm{Born})}$, with 2/3 of the additional contribution to $D_T^{(\mathrm{Born})}$ (the term with $\pi/2$) being attributable to compressibility. If we assume that $\eta \ll M^2$, then all of the diffusion is defined by the remaining expression. This expression describes the walk of a passive scalar particle in the field of a mixture of incoherent acoustic harmonic waves in the absence of any damping. Previously, we pointed out the resonant nature of this diffusion mechanism [$D_T \sim E^2(x)$]. Now, we see that this diffusion mechanism is more efficient than the diffusion due to damping. In the absence of damping, $p = 0$ and $\tilde{g}_c(x, 0, q)$ is simply the cosine transformation of the Green's function $\tilde{G}(x, t)$ in time t and it determines the diffusion coefficient. If we take $\tilde{G}(x, t) \equiv 1$, as in [6, 7], then we obtain $\tilde{g}_c(x, 0, q) = \delta(q)$. According to (13), this leads to the relation $D_T^{(\mathrm{Born})} \sim E(0) \equiv 0$, i.e., to the absence of diffusion. The actual Green's function $\tilde{G}(x, t)$ is close to the solution of the DIA equation (10). This function significantly depends on time and not only in the form of the diffusion exponential $\exp(-D_Tp^2\tau)$; it also contains harmonically varying terms like $\cos(\omega(x)t)$ and $\sin(\omega(x)t)$. These oscillating terms describe the diffusion of a passive scalar particle due to the passage from one wave to another and due to a gradual recession from the initial position when a continuous wave spectrum exists. Let us now show that these oscillating terms in $\tilde{G}(x, t)$ actually exist.

The asymptotic solutions (16) hold even after the first iteration of system (15). This implies that, to calculate D_T, it will suffice to use a linearized DIA equation when we substitute the corresponding molecular Green's function or unity for $\tilde{G}(|\mathbf{x} - \mathbf{y}|, t)$ in the kernel of Eq. (10) by ignoring small (of the order of $1/\gamma \ll 1$) damping. Thus, when analyzing diffusion, it will suffice to consider the following explicit expression for $\tilde{g}(x, s)$:

$$\tilde{g}(x, s) = \left[s + \frac{M^2x^2}{3}\int_0^\infty dyE(y)\frac{s + \eta(y)y^2}{[s + \eta(y)y^2]^2 + y^2}\right]^{-1}. \quad (19)$$

The sufficiency of using the linearized DIA equation for small Struchal numbers, $u_0\tau_0/R_0 \le 1$, for incompressible turbulence was shown in [3] (here, τ_0 and R_0 are the characteristic lifetime and velocity correlation scale, respectively). In our case of acoustic turbulence, the Mach number, $M \ll 1$, acts as the Struchal number. The inverse Laplace transform of expression (19) can be determined in analytic form only for the spectrum $E(x) = \delta(x - 1)$, although we do not consider this spectrum:

$$\tilde{G}(x, t) = \frac{\exp(-\eta\delta t/(1+\delta))}{1+\delta} \times\left\{1 + \delta\left[\cos(\omega(x)t) + \frac{\eta(4+\delta)}{2\omega^3(x)}\sin(\omega(x)t)\right]\right. \quad (20)$$
$$\left.\times\exp\left(-\frac{\eta(2-\delta)}{2(1+\delta)}t\right)\right\} + O(\eta^2).$$

Here,

$$\eta \equiv \eta(x) = \frac{p_0 k(p)}{c}, \quad \delta = \frac{M^2 x^2}{3}, \quad \omega^2(x) = 1 + \delta.$$

In the limit $\eta = 0$, the Green's function (20) takes the form

$$\tilde{G}(x,t) = \frac{1}{1 + M^2 x^2/3} \times \left[1 + \frac{M^2 x^2}{3}\cos\left(\sqrt{1 + \frac{M^2 x^2}{3}}\,t\right)\right]. \tag{21}$$

We see from formulas (20) and (21) that the contribution of the oscillating terms is large. Interestingly, the dispersion law $\omega^2(x) = 1 + M^2x^2/3$ differs from the dispersion law for acoustic waves [$\omega(x) = x$)], and $\omega(x)$ does not depend on the damping parameter $\eta(x)$. The frequency $\omega(x) \approx Mx/\sqrt{3}$ for large x. This limiting relation holds for an arbitrary spectrum, as is easy to see from Eq. (19) by assuming that $s \gg y$ and $s \gg \eta$. Note also that the group velocity corresponding to $\omega(x)$ is much lower than the speed of sound c:

$$V_{\text{group}} = c\frac{\partial\omega(x)}{\partial x} = \frac{2M^2 x}{3\sqrt{1 + M^2x^2/3}}c \ll c. \tag{22}$$

Although, for times much longer than the time it takes for diffusion to be established, the spot of passive scalar particles increases its size $R(\tau)$ as $R^2(\tau) \approx 6D_T\tau$, the group velocity appears to qualitatively characterize the smearing of this spot during the initial diffusion period.

Determining the frequencies $\omega(x)$ by ignoring the correlation damping, $\eta = 0$, reduces to finding purely imaginary [$s = \pm i\omega(x)$] roots of the equation

$$1 + \frac{M^2 x^2}{3}\int_0^\infty dy \frac{E(y)}{y^2 - \omega^2} = 0. \tag{23}$$

A simple example of the spectrum $E(x) = 1/(b-1)$ for x inside the interval $(1, b)$ and $E(x) = 0$ outside this interval shows that such roots exist. In this case, even two frequencies exist: one lies inside the interval $(1, b)$ and the other lies outside this interval, $\omega > b$.

The DIA equation describes the contribution from all the pair velocity correlators, including some of the fourth-order correlators, to the passive scalar transport. The contribution from the so-called irreducible fourth-order correlators (see [13] for more details), when the velocities at four space-time points are averaged crosswise, $\langle u_i(1)u_j(3)\rangle\langle u_n(2)u_m(4)\rangle$ for $t_1 \geq t_2 \geq t_3 \geq t_4$, is not included in this equation. When considering the fourth or higher order correlators, we assume, as is usually done, that the velocity-field ensemble is Gaussian; i.e., these correlators can be expressed as the product of all possible pair correlators. This assumption appears to be justified for estimates of the contribution from high-order correlators. An explicit formula for the contribution of the irreducible fourth-order correlators $D_T^{(1)}$ is given in [4] [see formula (19)]. After simple trigonometric transformations and after discarding terms of the order of $M^2 \ll 1$ in the integrand, we obtain

$$D_T^{(1)} = -\frac{u_0 M^3}{12 p_0}\int_0^\infty dx \int_0^\infty dy \int_{-1}^1 d\mu\mu E(x)E(y) \times [2xy + (x^2 + y^2)\mu + xy\mu^2] \times \tilde{g}_s(y, \eta(y)y^2, y)\tilde{g}_s(x, \eta(x)x^2, x) \times \tilde{g}_c(|\mathbf{x} + \mathbf{y}|, \eta(x)x^2 + \eta(y)y^2, x - y) + O\left(\frac{u_0 M^5}{p_0}\right). \tag{24}$$

Here, μ is the cosine of the angle between vectors $\mathbf{x}$ and $\mathbf{y}$. At $x \approx y$, $\tilde{g}_c$ is a δ-shaped function with a peak width of the order of $M^2 \ll 1$. Assuming the spectrum $E(x)$ to be smooth in the interval $\Delta x \sim M^2 \ll 1$ and discarding terms of the order of $M^2 \ll 1$ in the integrand finally yields

$$D_T^{(1)} = -\frac{\pi u_0 M^3}{9 p_0}\int_0^\infty dx E^2(x) + O\left(\frac{u_0 M^5}{p_0}\right). \tag{25}$$

Note that the contribution $D_T^{(1)}$ is attributable precisely to the damping $\eta(x)$. If the free terms in (16) are discarded, then the integrand in (24) becomes of the order of M^2 and the expression for $D_T^{(1)}$ transforms into a minor addition of the order of u_0M^5/p_0. Of course, this is a purely mental operation; damping always exists, if only because of viscosity, and the contribution (25) is real. This contribution was overlooked in [6], which resulted in an incorrect estimate, $D_T^{(1)} \sim M^2 D_T^{(0)} \ll D_T^{(0)}$. In [7], the contribution $D_T^{(1)}$ was not estimated at all.

If we retain only the free terms in expressions (16), which corresponds to the Born approximation used in [6, 7], then we will obtain a negative total turbulent diffusion coefficient:

$$D_T^{(\text{Born})} + D_T^{(1)} = -\frac{u_0 M}{3 p_0}\int_0^\infty dx E(x) \times \left[\frac{\pi}{3}M^2 E(x) - \eta(x)\right] \longrightarrow -\frac{(\pi - 3)u_0 M^3}{9 p_0}\int_0^\infty dx E^2(x). \tag{26}$$

Allowance for compressibility does not change this result, because the diffusion coefficient D_m (or even $D_T \sim u_0 M^3/p_0$) is much smaller than the damping coefficient $k \sim u_0 M/p_0$. The expression after the arrow corresponds to the model from [9], for which $\eta(x) = M^2 E(x)$.

Only using an oscillating Green's function $\tilde{G}(x, t)$ like (21), i.e., including the second term in expression (16) for $\alpha(x, p, q)$, and allowing for compressibility lead to the correct positive total turbulent diffusion coefficient $D_T = D_T^{(0)} + D_T^{(1)}$:

$$D_T = \frac{u_0 M}{3p_0}\int_0^\infty dx E(x)\left[\eta(x) + \frac{\pi}{6}M^2 E(x)\right] \longrightarrow \frac{(6+\pi)u_0 M^3}{18 p_0}\int_0^\infty dx E^2(x) + O\left(\frac{u_0 M^5}{p_0}\right). \quad (27)$$

The sixth-order correlators give a contribution of the order of $u_0 M^5/p_0$ to D_T, i.e., much smaller than the main expression (27). Since (26), owing to velocity-correlation damping, accounts for about 3% of (27), we conclude that random walks in the field of an incoherent superposition of harmonic acoustic waves with a continuous spectrum are the main diffusion mechanism for passive scalar particles in acoustic turbulence.

4. CONCLUSION

Below, we present our main results. We have shown for the first time that it is incorrect to use only the first (Born) approximation to calculate the turbulent diffusion coefficient D_T in acoustic turbulence, because the contribution from the next approximation is negative and slightly exceeds in magnitude the contribution from the first approximation even for very small Mach numbers, $M \ll 1$. We presented a procedure for correctly calculating D_T based on the asymptotic solution of a nonlinear DIA equation for the mean Green's function $\tilde{G}(p, \tau)$ in the form of damped (with time) oscillations. The correspondence of this Green's function to the transport physics of passive scalar particles was shown to be a crucial factor.

The diffusion of a passive scalar particle in acoustic turbulence takes place for the following two reasons. First, the damping of correlated gas motions due to viscosity and, more importantly, due to nonlinear wave interaction prevents the return of the passive scalar particle to its original position. Second, as we showed here for the first time, an incoherent superposition of waves results in a random walk of the particle in the medium even if this damping is ignored. Our calculations indicate that the second mechanism is much stronger than the first mechanism and that it leads to the final formula (27) for the diffusion coefficient D_T. As a result, it emerged that an efficient method of calculating D_T involves using the formula of the first approximation with the velocity-correlator damping ignored and using solution (19) to the linearized DIA equation as the Green's function. In this method of calculating D_T, damping must be disregarded in order not to calculate the large negative correction from the second approximation. This correction results from the contribution of damping and the fourth-order velocity correlators and almost completely offsets the contribution of damping to D_T from the first approximation. In general, the positive contribution to D_T from the damping mechanism in the first approximation is equal in order of magnitude to the contribution from the second mechanism. The contribution of incompressibility [see the second term in (8)], which was previously disregarded, accounts for about 2/3 of the total diffusion coefficient. This additional term in the general formula (8) for D_T was first obtained in [3].

Our results apply to acoustic turbulence with a continuous wave spectrum. For single-mode turbulence [$E(p) = \delta(p - p_0)$], the diffusion is governed only by the first mechanism, i.e., by damping. This is immediately seen from expression (20) for the Green's function for this case.

We gave a purely mathematical solution to the problem of correctly calculating the diffusion coefficient for passive scalar particles in acoustic turbulence with a continuous wave spectrum. For a more complete physical understanding of the diffusion of passive scalar particles in acoustic turbulence, it would undoubtedly be useful to derive our formula for D_T purely qualitatively. The dependence $D_T \propto u_0 \lambda_0 M^3$ can apparently be justified from dimension considerations as follows. Disregarding the small contribution from the first diffusion mechanism, we can assume that D_T depends on c, u_0, and $\lambda_0 \approx 1/p_0$. An elementary diffusion event, the jump of a particle from one wave to another, in a statistically isotropic medium can depend on the velocity of the first wave squared and on the velocity of the second wave squared, i.e., $D_T \propto u_0^4 \propto M^4$. The dimensional proportionality coefficient can be only $c\lambda_0$. As a result, we obtain

$$D_T = \text{const}\, c\lambda_0 M^4 = \text{const}\, u_0 \lambda_0 M^3,$$

which matches formula (27).

REFERENCES

1. G. I. Taylor, Proc. London Math. Soc. A **20**, 196 (1921).
2. H. K. Moffatt, J. Fluid Mech. **65**, 1 (1974).
3. A. Z. Dolginov and N. A. Silant'ev, Geophys. Astrophys. Fluid Dyn. **63**, 139 (1992).
4. N. A. Silant'ev, Zh. Éksp. Teor. Fiz. **114**, 930 (1998) [JETP **87**, 505 (1998)].
5. S. I. Vaĭnshteĭn, Dokl. Akad. Nauk SSSR **195**, 793 (1970) [Sov. Phys. Dokl. **15**, 1090 (1971)].

6. S. I. Vaĭnshteĭn and Ya. B. Zel'dovich, Usp. Fiz. Nauk **106**, 431 (1972) [Sov. Phys. Usp. **15**, 159 (1972)].
7. A. P. Kazantsev, A. A. Ruzmaĭkin, and D. V. Sokolov, Zh. Éksp. Teor. Fiz. **88**, 487 (1985) [Sov. Phys. JETP **61**, 285 (1985)].
8. L. D. Landau and E. M. Lifshitz, *Course of Theoretical Physics*, Vol. 6: *Fluid Mechanics* (Nauka, Moscow, 1986; Pergamon, New York, 1987).
9. V. E. Zakharov and R. Z. Sagdeev, Dokl. Akad. Nauk SSSR **192**, 297 (1970) [Sov. Phys. Dokl. **15**, 439 (1970)].
10. N. A. Silant'ev, Zh. Éksp. Teor. Fiz. **111**, 871 (1997) [JETP **84**, 479 (1997)].
11. N. A. Silant'ev, Zh. Éksp. Teor. Fiz. **101**, 1216 (1992) [Sov. Phys. JETP **74**, 650 (1992)].
12. R. H. Kraichnan, J. Fluid Mech. **5**, 497 (1959).
13. V. I. Tatarskii, *Wave Propagation in a Turbulent Medium* (Nauka, Moscow, 1967; McGraw-Hill, New York, 1961).

Translated by V. Astakhov

Journal of Experimental and Theoretical Physics, Vol. 95, No. 5, 2002, pp. 965–972.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 122, No. 5, 2002, pp. 1116–1124.

NONLINEAR PHYSICS

Nonlinear Localization of Excitations and the Dynamics of Solitons in Self-Modulated Systems

A. S. Kovalev[a, *] and I. V. Gerasimchuk[b, **]

[a]*Verkin Institute for Low Temperature Physics and Engineering, National Academy of Sciences of Ukraine, Kharkov, 61103 Ukraine*

[b]*Institute for Theoretical Physics, National Scientific Center Kharkov Institute of Physics and Technology, Kharkov, 61108 Ukraine*

**e-mail: kovalev@ilt.kharkov.ua*

***e-mail: igera@ukr.net*

Received June 18, 2002

Abstract—The nonlinear dynamics of systems with a spatially periodic ground state was studied. The dynamics of kinks against the background of a periodic soliton structure was considered for the example of the sine-Klein–Gordon model that described a fluxon lattice in a long Josephson contact in an external magnetic field and an incommensurate structure of a surface atomic layer or adatom chains on the surface of a crystal. The velocity of moving kinks was shown to be bounded from above and from below if the ground state was spatially periodic. © *2002 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

In recent years, interest of researchers working on the theory of nonlinear waves and solitons has shifted to studying the nonlinear dynamics of real physical systems with their discrete character, internal structure, and spatial nonuniformity. Of special interest are layered substances in which material parameters are periodically modulated in space ("modulated systems"). Examples of such media are layered crystals with polyatomic unit cells (in particular, high-T_c superconductors), multilayer magnets (which offer much promise for technological applications), corrugated optical fiber waveguides, etc. In all these media, spatial periodicity results in a band structure of the spectrum of linear waves and the formation of gaps ("forbidden bands") in this spectrum. The existence of peculiar localized excitations, so-called "gap solitons," in forbidden spectrum regions becomes possible if medium nonlinearity is taken into account [1–3]. Unfortunately, gap solitons can only be studied by approximate methods in the small-amplitude limit within the framework of simplified models.

In this work, we call attention to the possibility of the existence of gap solitons of a different origin, which can be studied by exact analytic methods with the use of so-called integrable models. Consider systems in which all material medium parameters are spatially uniform, but the ground state is spatially periodic. Examples of such systems are numerous and well known. The ground state of a long Josephson contact in a magnetic field higher than critical is a periodic fluxon lattice [4, 5]. Another example is an incommensurate periodic structure that can be formed by surface atoms as, for instance, on the [111] surface of gold [6, 7], because of different electronic states and, therefore, different elastic properties in the bulk and in the surface layer.

Similar but one-dimensional incommensurate structures can be formed by adsorbed atom chains on striated [110] surfaces of crystals [8]. Examples of such systems (hereafter called "self-modulated") are antiferromagnets with nonuniform exchange-relativistic Dzyaloshinski interaction and with the ground state in the form of a periodic helicoidal structure [9, 10] and the domain structure of ferromagnets with magnetic dipole interaction taken into account. In these examples, the spectrum of linear excitations also contains gaps, but soliton excitations whose frequencies lie within these spectral gaps are in many respects different from those in usual modulated media.

As a first step, we studied the dynamics of the simplest one-parameter topological solitons of the "kink" type [11], which propagate in a self-modulated medium, that is, through a periodic lattice of similar kinks.

2. FORMULATION OF THE MODEL: A PERIODIC GROUND STATE AND THE SPECTRUM OF LINEAR EXCITATIONS ABOVE IT

Consider a simple example of a self-modulated system, namely, an incommensurate structure of a surface atomic layer or of a chain of adsorbed atoms on the surface of a crystal. Interaction between surface atoms will be taken into account in the harmonic approximation. It is assumed that the equilibrium interatomic distance in

the absence of a surface field is b and is different from the interatomic distance in the bulk, which equals a. The effect of bulk (substrate) on the adatoms can be modeled by a periodic potential relief with the period a. For simplicity, this relief is selected in the form of a trigonometric function. (The influence of adsorbed atoms on the substrate is ignored; that is, the substrate is considered absolutely rigid.) The potential energy of this system has the form

$$U = \sum_n U_0\left\{1 - \cos\frac{2\pi y_n}{a}\right\} + \sum_n \frac{\alpha}{2}(y_n - y_{n-1} - b)^2, \quad (1)$$

where y_n is the coordinate of the nth atom and α is the elastic interaction constant along the chain. The dynamic equation for atomic displacements

$$\nu_n = y_n - an$$

in this model (the Frenkel–Kontorova model [12]) has the form

$$m\nu_{\tau\tau} + \frac{2\pi U_0}{a}\sin\frac{2\pi\nu_n}{a} + \alpha(2\nu_n - \nu_{n+1} - \nu_{n-1}) = 0. \quad (2)$$

Using the long-wave approximation and the dimensionless variables

$$u = \frac{2\pi\nu}{a}, \quad x = 2n\pi\sqrt{\frac{U_0}{\alpha a^2}}, \quad t = 2\tau\pi\sqrt{\frac{U_0}{ma^2}},$$

we obtain the well-known sine-Gordon equation [11]

$$u_{tt} - u_{xx} + \sin u = 0. \quad (3)$$

In the same approximation and in the same variables, total energy (1) takes the form

$$U = E_0\int dx\left\{\frac{u_t^2}{2} + \frac{u_x^2}{2} + (1 - \cos u) + \xi u_x\right\}, \quad (4)$$

where

$$E_0 = \frac{a}{2\pi}\sqrt{U_0\alpha}.$$

The incommensurability of the chain of adatoms and the substrate is characterized by the dimensionless parameter

$$\xi = \sqrt{\frac{\alpha a^2}{U_0}}\frac{a - b}{a}.$$

For a long Josephson contact, field $u(x)$ describes the phase difference of the wave functions of superconductors at the contact, and the ξ parameter is proportional to the external magnetic field applied in the contact plane [13].

The last term ξu_x in Eq. (4) for the energy has a divergent form and does not influence dynamic equation (3). Changes in the ξ parameter, however, change the potential energy of the system and can alter its ground state. At $b = a$, the ground state corresponds to the trivial solution of (3) with $u \equiv 0$ and energy $E = 0$. At $b \neq a$, the situation is more complex.

For definiteness, set $b > a$ ($\xi < 0$). In addition to the $u \equiv 0$ solution, (3) then has a nontrivial static solution [4, 5]

$$u_0(x) = \pi + 2\mathrm{am}\left\{\frac{x}{k}, k\right\}, \quad (5)$$

where $\mathrm{am}(z, k)$ is the elliptic amplitude and k is the modulus of this elliptic function. Solution (5) describes a "stretched" system with a periodic chain of 2π kinks (a chain of "one-dimensional" dislocations in the surface layer or a fluxon lattice in a Josephson contact) with the distance $L = 2kK(k)$ between them, where $K(k)$ is the complete elliptic integral of the first kind. The natural kink width is a value of the order of one. In the initial dimensional variables, this width is given by the equation

$$\Lambda = \frac{a}{2\pi}\sqrt{\frac{\alpha a^2}{U_0}}.$$

The energy density of such a periodic structure (per period) is

$$\varepsilon = \frac{U}{L} = E_0\left\{\frac{2}{k^2}\left[\frac{2E(k)}{K(k)} - 1\right] - \frac{\xi\pi}{kK(k)} + 2\right\}. \quad (6)$$

Here, $E(k)$ is the complete elliptic integral of the second kind. Energy density ε depends on the ξ parameter of incommensurability of interatomic distances in the chain and in the substrate. At small ξ parameter values, the ground state of the system is uniform, and periodic solution (5) with a high energy can only exist if pressure is applied to the chain at infinity. However, if parameter ξ exceeds the critical value $\xi_c = -4/\pi$, at which

$$b_c = a + \frac{4}{\pi}\sqrt{\frac{U_0}{\alpha}}, \quad (7)$$

the energy minimum corresponds to periodic state (5) with the modulus of the elliptic function determined by the equation

$$\frac{E(k)}{k} = \frac{\xi}{\xi_c}.$$

Small-amplitude excitations against the background of the nontrivial ground state with $w = u - u_0 \ll 1$ are described by the linear equation

$$w_{tt} - w_{xx} + \left(1 - 2\mathrm{cn}^2\left(\frac{x}{k}, k\right)\right)w = 0. \quad (8)$$

Periodic solutions of this Lamé equation for linear waves above the ground state were discussed in [14]

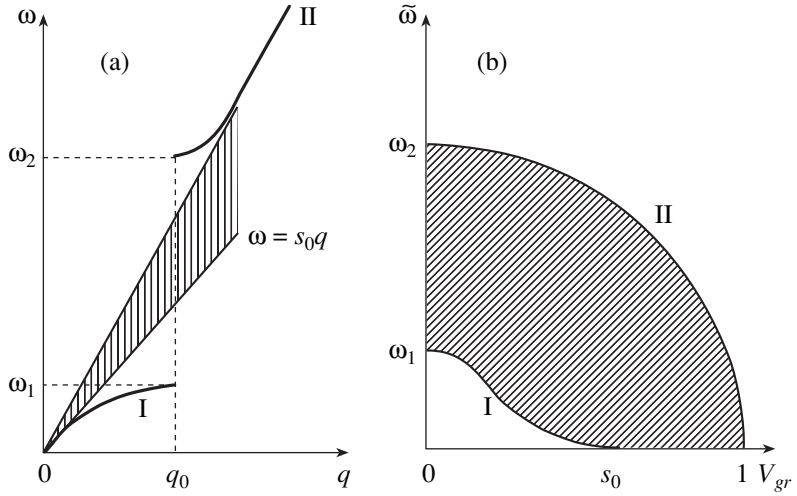


Fig. 1. (a) Dispersion law for linear waves in a self-modulated system with a periodic ground state and (b) the same dispersion law in the frame of reference moving at the group velocity of linear waves.

(also see [15]). They are expressed via elliptic eta and theta functions,

$$w = \frac{H(x/k-\eta)}{\Theta(x/k)}\exp[i(qx-\omega t)], \qquad (9)$$

and the dispersion law of linear waves $\omega = \omega(q)$, where q is the wave number, is implicitly written as

$$\omega = \frac{1}{k}\mathrm{dn}(\eta, k), \quad q = \frac{i}{k}Z(\eta, k), \qquad (10)$$

where the $Z(\eta, k)$ zeta function is purely imaginary in the problem under consideration. As distinguished from the dispersion law for linear waves above a uniform ground state $\omega^2 = 1 + q^2$, the spectrum of linear waves above periodic ground state (5) consists of two branches separated by a gap at $q = q_0 = \pi/L$ (see Fig. 1a). The first Goldstone branch (I) describes natural oscillations of the kink lattice, which is treated as an effective chain of anharmonically coupled quasi-particles. In the long-wave limit ($\omega, q \longrightarrow 0$), the dispersion law has the "sound" form $\omega = s_0 q$, and "the velocity of sound" in the kink lattice is

$$s_0 = \frac{k'K(k)}{E(k)} < 1,$$

where $k' = \sqrt{1-k^2}$. If the density of kinks in the incommensurate structure is low and $L \gg 1$ ($k' \ll 1$), it is easy to find the dependence of s_0 on L,

$$s_0 = 2L\exp(-L/2).$$

In the initial dimensional variables, this dependence takes the form

$$s_0 = \sqrt{\frac{\alpha a^2}{m}}\frac{2L}{\Lambda}\exp\left\{-\frac{L}{2\Lambda}\right\}.$$

In the opposite limit, when $q \longrightarrow q_0 = \pi/L$ and $\omega \longrightarrow \omega_1 = k'/k$, the linear wave represents antiphase kink oscillations and is described by the function

$$w = \mathrm{cn}\left(\frac{x}{k}, k\right)\sin(\omega_1 t).$$

The second (activation) spectrum branch (II) is bounded from below by the frequency $\omega_2 = 1/k$ (at $q = q_0$), and the corresponding solution

$$w = \mathrm{sn}\left(\frac{x}{k}, k\right)\sin(\omega_2 t)$$

describes antiphase oscillations of regions between kinks at fixed positions of solitons themselves. At large wave numbers $q \gg q_0$, the spectrum again assumes the sound form, but the velocity of sound is then larger than s_0 (in our variables, larger than one). This velocity corresponds to the limiting velocity of linear waves in a uniform medium without solitons. In the $L \longrightarrow \infty$ ($k \longrightarrow 1$) limit, the lower branch of linear waves disappears, and the upper branch takes the form

$$\omega = \sqrt{1+q^2}, \quad \omega_2 \longrightarrow 1.$$

It is expedient to rewrite the dispersion law in terms of the group velocity of linear waves

$$V_{gr} = \partial\omega/\partial q$$

and to write wave frequencies in the frame of reference that moves at the group velocity,

$$\tilde{\omega} = \omega - qV_{gr}.$$

The resulting dispersion law is shown in Fig. 1b, where curves I and II correspond to the lower and upper spectrum branches in Fig. 1a. The hatched region between them refers to two-parameter dynamic solitons of the

envelope, which will not be considered. Below, we restrict our analysis to studying the dynamics of kinks (one-parameter topological solitons) in an incommensurate structure (a periodic lattice of kinks of the same sign). Such soliton solutions correspond to the $s_0 < V_{gr} < 1$ region of the velocity axis in Fig. 1b and to the hatched sector of velocities in Fig. 1a.

3. MOTION OF A KINK THROUGH A SOLITON LATTICE

Consider the propagation of a one-parameter soliton (additional kink) when the ground state is given by periodic function (5), that is, the propagation through a periodic lattice of kinks of the same sign. The exact solution to this problem is simplest to find using the Darbu transformation, which allows more complex solutions to be obtained if a "seed" solution is known [here, solution (5) for the ground state]. The Darbu transformation is especially simple to use if the seed solution depends on a single variable, as in the problem under consideration. Indeed, we have $u = u_0(x)$, and u is independent of time. (More complex solutions against the background of a moving periodic structure are easily obtained using the Lorentz transformation.) A problem similar to ours was considered in [16, 17] for solitons of the sine-Gordon equation that propagated against the background of the standing monochromatic wave $u_0 = u_0(\omega t)$ and in [18] for solitons of the hyperbolic sine-Gordon equation against the background of a monochromatic wave.

The Darbu transformation for sine-Gordon equation (3) that we are using is well known [19]. To write this transformation in a compact form, it is convenient to pass from the initial field variable $u(x, t)$ to new variables V and W related to the initial field u as

$$V = i(u_x + u_t), \quad W = \exp(iu). \tag{11}$$

For these variables, (3) is written in the Lax form,

$$V_x - V_t = \frac{1}{2}\left(W - \frac{1}{W}\right), \tag{12}$$

$$W_x + W_t = VW. \tag{13}$$

There is an auxiliary linear problem for two complex-valued functions $\psi_1(x, t)$ and $\psi_2(x, t)$ that corresponds to system (12), (13). Let us introduce the column function

$$\psi = \begin{Bmatrix} \psi_1 \\ \psi_2 \end{Bmatrix}.$$

An arbitrary solution $u(x, t)$ [therefore, also $V(x, t)$ and $W(x, t)$] can then be put in correspondence with the overdetermined system of equations [19]

$$4\psi_x = \begin{vmatrix} V & \lambda + \dfrac{W}{\lambda} \\ \lambda + \dfrac{1}{W\lambda} & -V \end{vmatrix} \psi, \tag{14}$$

$$4\psi_t = \begin{vmatrix} V & \lambda - \dfrac{W}{\lambda} \\ \lambda - \dfrac{1}{W\lambda} & -V \end{vmatrix} \psi. \tag{15}$$

Generally, the complex parameter λ in (14) and (15) is the Darbu transformation parameter. The initial sine-Gordon equation [Eq. (3)] is the condition of system (14), (15) consistency. Solving (14), (15) with a given seed solution $u_0(x, t)$ (consequently, with given $V_0(x, t)$ and $W_0(x, t)$ functions) and an arbitrary λ parameter allows a new $u(x, t)$ [or $V(x, t)$ and $W(x, t)$] solution to be constructed. Naturally, the λ parameter should be selected such that the new $u(x, t)$ solution will be real. Let us introduce the notation

$$\sigma_1 = \lambda\frac{\psi_2}{\psi_1}, \quad \sigma_2 = \lambda\frac{\psi_1}{\psi_2}. \tag{16}$$

The relation between the new and seed solutions is then given by the formulas

$$V = -V_0 + \sigma_2 - \sigma_1, \quad W = W_0\sigma_1/\sigma_2. \tag{17}$$

Using (11), we can easily find the final form of the equation that relates the new and old solutions in the initial field variables,

$$u(x, t) = u_0(x, t) - 2i\ln\frac{\psi_2(u_0, \lambda)}{\psi_1(u_0, \lambda)}. \tag{18}$$

It follows that the central problem is to solve system (14), (15) of linear equations with variable coefficients. In our case, the problem is simplified, because the initial solution for the ground state $u = u_0(x)$ is independent of time. Therefore, (15) becomes a differential equation with constant coefficients and includes coordinate x as a parameter. For the periodic ground state under consideration [Eq. (5)], the V_0 and W_0 functions are written as

$$V_0 = \frac{2i}{k}\mathrm{dn}\left(\frac{x}{k}, k\right),$$
$$W_0 = \left\{\mathrm{sn}\left(\frac{x}{k}, k\right) - i\,\mathrm{cn}\left(\frac{x}{k}, k\right)\right\}^2, \tag{19}$$

and system (15) of linear equations is easily solved. The key problem is to correctly select the λ transformation

parameter at which the new $u(x, t)$ solution is real. We will see below that selecting arbitrary real numbers for λ leads to real $u(x, t)$ solutions, and the λ parameter itself characterizes the velocity of soliton propagation.

The solution to (15) has the form

$$\begin{aligned}\psi_1 &= a(x)\exp(\mu t) + b(x)\exp(-\mu t),\\ \psi_2 &= a(x)\frac{\lambda(4\mu - V_0)}{\lambda^2 - W_0}\exp(\mu t)\\ &\quad - b(x)\frac{\lambda(4\mu + V_0)}{\lambda^2 - W_0}\exp(-\mu t),\end{aligned} \tag{20}$$

where

$$\mu = \frac{1}{4}\sqrt{(\lambda + 1/\lambda)^2 - 4/k^2}.$$

Substituting (20) into the first pair of equations (14) and equating the coefficients of $\exp(\pm\mu t)$ to zero yields equations for $a(x)$ and $b(x)$ solvable in quadratures. The corresponding expressions have the form

$$\begin{aligned}a(x) &= \exp\left\{\mu\int_0^x \frac{\lambda^2 + W_0}{\lambda^2 - W_0}dx\right.\\ &\quad \left. - \frac{1}{2}\int_0^x \frac{V_0 W_0}{\lambda^2 - W_0}dx + C_1\right\},\\ b(x) &= \exp\left\{-\mu\int_0^x \frac{\lambda^2 + W_0}{\lambda^2 - W_0}dx\right.\\ &\quad \left. - \frac{1}{2}\int_0^x \frac{V_0 W_0}{\lambda^2 - W_0}dx + C_2\right\}.\end{aligned} \tag{21}$$

It follows from (16), (17) and the unimodular character of the W function [see (11)] that the ψ_2/ψ_1 function is unimodular. It is easy to check that this requirement is met if the C_1 and C_2 constants of integration satisfy the condition

$$C_1 - C_2 = i\pi/2.$$

Let us separate out the real and imaginary parts in the first integral in (21) and use the relation

$$W_0 - W_0^* = 2V_{0x}.$$

The equation for the ψ_2/ψ_1 ratio, which determines new solution (18) for $u(x, t)$, then becomes

$$\frac{\psi_2}{\psi_1} = \exp(i\rho)\frac{\exp(i\varphi) - i\exp(\vartheta)}{1 + i\exp(i\varphi + \vartheta)}, \tag{22}$$

where $\vartheta = 2\mu(t + f(x))$ and

$$f(x) = \frac{k^2}{4}\left(\lambda^2 - \frac{1}{\lambda^2}\right)\int_0^x \frac{dx}{\mathrm{dn}^2(x/k, k) + (2k\mu)^2}, \tag{23}$$

$$\begin{aligned}\varphi &= \arctan\frac{\mathrm{dn}(x/k, k)}{2k\mu},\\ \rho &= \arctan\frac{2\mathrm{sn}(x/k, k)\mathrm{cn}(x/k, k)}{2\mathrm{sn}^2(x/k, k) - (1 + \lambda^2)}.\end{aligned} \tag{24}$$

The differentiation of the $(\varphi \pm \rho)$ function with respect to the coordinate followed by simple transformations involving elliptic Jacobi functions and by integration transforms these equations into

$$\varphi \pm \rho = \mp\,\mathrm{am}\left(\frac{x}{k}, k\right) \pm \mathrm{am}\left(\frac{x \pm \Delta}{k}, k\right), \tag{25}$$

where the Δ phase shift of the solution implicitly depends on the k and λ parameters as follows:

$$\mathrm{sn}\left(\frac{\Delta}{k}, k\right) = \frac{2\lambda}{(1 + \lambda^2)k}. \tag{26}$$

Note that all the formulas given above are based on the assumption that $\lambda > 1$. This corresponds to positive f values; that is, the soliton moves in the negative direction.

The $f(x)$ function in (22) can be represented by the sum of two terms,

$$f(x) = \frac{x}{v} + \chi(x),$$

where the mean value of the $\chi(x)$ periodic function is zero. Of key importance is the $f(x)$ function component that linearly grows as the coordinate increases. This function determines the mean velocity of soliton propagation through the incommensurate structure,

$$v = \frac{4K(k)\lambda^2}{(\lambda^4 - 1)k^2}\left\{\int_0^{K(k)} \frac{dz}{\mathrm{dn}^2(z, k) + (2k\mu)^2}\right\}^{-1}. \tag{27}$$

It follows that, in (22), the phase

$$\vartheta = \frac{2\mu}{v}(x + vt) + 2\mu\chi(x)$$

describes soliton movement in the negative direction at mean velocity v. This movement is accompanied by periodic oscillations when the soliton passes through the points at which soliton lattice kinks are situated. Substituting (22) and (25) into (18) yields the final

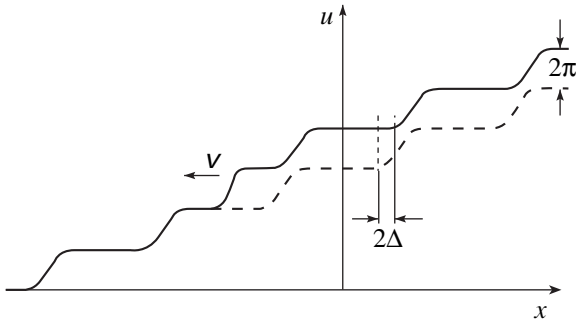


Fig. 2. Movement of an additional kink through a lattice of solitons. The dashed line corresponds to the field distribution $u_0(x)$ in the initial periodic incommensurate structure.

solution for the motion of the additional soliton through the soliton lattice,

$$u(x,t) = \pi + 2\mathrm{am}\left(\frac{x}{k}, k\right) - 2i\ln\left\{\frac{\exp(i\kappa_+) - i\exp(\vartheta + i\kappa)}{\exp(-i\kappa_+) + i\exp(\vartheta - i\kappa)}\right\}. \quad (28)$$

Here, the notation

$$\kappa_\pm = \frac{1}{2}\left\{\mathrm{am}\left(\frac{x \pm \Delta}{k}, k\right) - \mathrm{am}\left(\frac{x}{k}, k\right)\right\} \quad (29)$$

is used for convenience.

In spite of the awkwardness of solution (26)–(29), it has a simple physical interpretation, and an analysis of its asymptotic behaviors poses no difficulty.

4. DISCUSSION

First, consider the asymptotic behaviors of the solution. If $x \longrightarrow \pm\infty$ ($\vartheta \longrightarrow \pm\infty$), it follows from (28) that

$$u(x \longrightarrow -\infty, t) \approx \pi + 2\mathrm{am}\left(\frac{x+\Delta}{k}, k\right),$$
$$u(x \longrightarrow +\infty, t) \approx \pi + 2\mathrm{am}\left(\frac{x-\Delta}{k}, k\right) + 2\pi. \quad (30)$$

The 2π term in (30) in the expression for $u(x \longrightarrow +\infty)$ corresponds to the appearance of an additional 2π kink in the initial lattice of solitons. Because of the presence of this additional soliton, the initial lattice is as a whole stretched by 2Δ. The Δ value depends on the velocity at which this soliton moves.

The described dynamics is schematically shown in Fig. 2. The dashed line corresponds to field distribution $u_0(x)$ in the initial periodic incommensurate structure. The additional kink, which moves through the lattice, has a steeper front in the figure (which corresponds to velocities close to one). This kink displaces the whole structure behind it by $2\Delta(v)$. Consider the limiting cases of this movement.

In the $\lambda \longrightarrow \infty$ limit (then $\mu \longrightarrow \infty$), the kink velocity tends to its limiting value $v \longrightarrow 1$ [the limiting velocity in Lorentz-invariant model (3)]. The width of the kink $l = v/2\mu$ then tends to zero, and the kink transforms into a singular Heaviside theta function. The 2Δ phase shift of the solution vanishes at infinity: the singular kink moves through an undeformed initial periodic structure at the limiting velocity, and

$$u(+\infty) - u_0(+\infty) = 2\pi.$$

The dynamics of the kink moving in a periodic structure at minimum velocity s_0 is of interest. In this limit, $\mu \longrightarrow 0$, and λ tends to its minimum value

$$\lambda_{\min} = \frac{1+k'}{k}.$$

The width of the moving kink then tends to infinity, and the kink "spreads." Its limiting velocity s_0 coincides with the sound velocity of long-wave oscillations of the initial lattice of kinks, and the 2Δ phase shift tends to $2\Delta = L$, that is, to the period of the initial periodic structure. Regular alternation of kinks in the self-modulated structure is then restored, and the $u(+\infty) - u_0(+\infty)$ shift vanishes.

The limiting case of additional soliton movement at velocities close to s_0 can be given a clear physical interpretation if the incommensurate structure has a large period, $L \gg 1$ (when $k' \ll 1$). In this limit, separate kinks can be treated as weakly interacting quasi-particles that form a one-dimensional chain. The solution to (3) for an isolated kink is well known [11, 12],

$$u = 4\arctan\left[\exp\left(\frac{x - vt}{\sqrt{1-v^2}}\right)\right]. \quad (31)$$

Substituting this solution into expression (4) for the total energy yields the energy of the isolated kink in the form

$$E = E_0\left\{\frac{8}{\sqrt{1-v^2}} + 2\pi\xi\right\}. \quad (32)$$

Obviously, velocity v in dimensionless variables is related to velocity V_k in the initial variables as

$$v = \frac{V_k\sqrt{m}}{\sqrt{\alpha a^2}}.$$

It follows from (32) that, for $v \ll 1$, the kinetic energy of the kink is $E = MV_k^2/2$ and its effective mass is

$$M = \frac{4}{\pi}\sqrt{\frac{U_0}{\alpha a^2}}m. \quad (33)$$

Further, it follows from (32) that the energy of the isolated static soliton becomes negative at $\xi < -4/\pi$. Above, we obtained this critical incommensurability parameter from the exact periodic solution given by (5). At large incommensurability parameter values $|\xi| > |\xi_c|$, the formation of kinks becomes energetically favorable. Their density, however, remains finite because of mutual repulsion. The repulsion energy between two kinks of the same sign was calculated in [20] for the sine-Gordon equation. At a distance between kinks far exceeding their size, this energy has the form

$$U(\Delta) \approx 32E_0 \exp(-\Delta/\Lambda), \tag{34}$$

where Δ is the distance between solitons and Λ is the characteristic kink width introduced above.

The coordinate of the Nth kink in the soliton lattice can be written in the form

$$y_N = LN + \zeta_N,$$

where L is the distance between kinks in the ground state and ζ_N is small kink displacements. The total energy of the system (above the ground-state energy) is then written as

$$E = \sum_N \left\{ \frac{1}{2} M \left(\frac{d\zeta_N}{d\tau} \right)^2 + 32E_0 \exp\left(-\frac{L}{\Lambda} \right) \times \exp\left(-\frac{\zeta_N - \zeta_{N-1}}{\Lambda} \right) \right\}. \tag{35}$$

This corresponds to the well-known and completely integrable Toda model [21, 22]. We will only consider the long-wave approximation, when relative displacements of neighboring kinks are substantially smaller than their width,

$$\zeta_N - \zeta_{N-1} \ll \Lambda.$$

Passing to the continuous coordinate of kink centers $NL \longrightarrow z$, we then easily obtain the well-known Boussinesq equation [11]

$$M\zeta_{\tau\tau} - \frac{32E_0L^2}{\Lambda^2} \exp\left(-\frac{L}{\Lambda} \right) \times \left\{ \zeta_{zz} + \frac{L^2}{12} \zeta_{zzzz} - \frac{L}{\Lambda} \zeta_z \zeta_{zz} \right\} = 0 \tag{36}$$

in the main nonlinear approximation, that is, only taking into account terms quadratic in $\zeta(z, \tau)$.

It is easy to see that, if the dispersion ($\sim\zeta_{zzzz}$) and nonlinear ($\sim\zeta_z\zeta_{zz}$) terms are ignored, the obtained wave equation describes kink lattice oscillation waves propagating at the velocity

$$s_0 = \sqrt{\frac{\alpha a^2}{m}} \frac{2L}{\Lambda} \exp\left\{ -\frac{L}{2\Lambda} \right\}.$$

Above, we obtained the same expression for s_0 from the exact solution for linear waves in an incommensurate structure.

The soliton solution to (36), which describes a "superkink" against the background of a lattice of kinks, has the form [11]

$$\zeta = -\sqrt{3}\Lambda \sqrt{\frac{V_k^2}{s_0^2} - 1} \times \left\{ 1 + \tanh\left[\sqrt{3} \sqrt{\frac{V_k^2}{s_0^2} - 1} \frac{z - V_k\tau}{L} \right] \right\}. \tag{37}$$

In Fig. 1b, a $V_k > s_0$ velocity axis region in the vicinity of the $V_k = s_0$ point corresponds to this solution. In the $V_k \longrightarrow s_0$ limit, we have

$$\zeta(+\infty) - \zeta(-\infty) = 0.$$

Above, this result was obtained from the exact solution.

5. CONCLUSION

In this work, we considered the dynamics of a kink moving through an incommensurate structure, through a lattice of kinks (the dynamics of a fluxon moving through a fluxon lattice in a long Josephson contact). The exact analytic solution was found and studied to describe this dynamics in terms of the sine-Gordon equation. In the limit of a large lattice period and a low velocity of the additional soliton, a qualitative physical description of the dynamics of the soliton was suggested. The results can be used to theoretically describe changes in the density of fluxons in a long Josephson contact under applied magnetic field variations.

ACKNOWLEDGMENTS

This work was partly supported by INTAS (grant INTAS-99 no. 0167).

REFERENCES

1. D. L. Mills and J. Trullinger, Phys. Rev. B **36**, 947 (1987).
2. O. A. Chubykalo, A. S. Kovalev, and O. V. Usatenko, Phys. Rev. B **47**, 3153 (1993).
3. A. B. Aceves and S. Wabnitz, Phys. Lett. A **141**, 37 (1989).
4. R. A. Ferell and R. E. Prange, Phys. Rev. Lett. **40**, 479 (1963).
5. D. J. Scalapino, in *Tunneling Phenomena in Solids*, Ed. by E. Burstein and S. Lundqvist (Plenum, New York, 1969), p. 477.

6. U. Harten, A. M. Lahee, J. P. Toennies, and Ch. Woll, Phys. Rev. Lett. **54**, 2619 (1985).
7. M. Mansfield and R. J. Needs, J. Phys.: Condens. Matter **2**, 2361 (1990).
8. O. Braun and Yu. S. Kivshar, Phys. Rep. **306**, 4 (1998).
9. V. G. Bar'yakhtar, T. K. Soboleva, and A. L. Sukstanskiĭ, Fiz. Tverd. Tela (Leningrad) **27**, 2428 (1985) [Sov. Phys. Solid State **27**, 1454 (1985)].
10. E. P. Stefanovskiĭ and A. L. Sukstanskiĭ, Zh. Éksp. Teor. Fiz. **104**, 3434 (1993) [JETP **77**, 628 (1993)].
11. A. M. Kosevich and A. S. Kovalev, *Introduction to the Nonlinear Physical Mechanics* (Naukova Dumka, Kiev, 1989).
12. A. M. Kosevich, *Theory of Crystal Lattice* (Vishcha Shkola, Khar'kov, 1988).
13. I. O. Kulik and I. K. Yanson, in *Josephson Effect in Superconducting Tunnel Structures* (Nauka, Moscow, 1970).
14. P. Lebwohl and M. Stephen, Phys. Rev. **163**, 376 (1967).
15. E. T. Whittaker and G. N. Watson, *Modern Analysis* (Cambridge Univ. Press, New York, 1940), Chap. XXIII.
16. V. M. Eleonskiĭ, N. E. Kulagin, and N. S. Novozhilova, Zh. Éksp. Teor. Fiz. **89**, 2174 (1985) [Sov. Phys. JETP **62**, 1255 (1985)].
17. V. M. Eleonskiĭ and N. E. Kulagin, Fiz. Met. Metalloved. **65**, 868 (1988).
18. A. S. Kovalev and N. E. Kulagin, Preprint No. 14-87, FTINT AN USSR (Khar'kov, 1987).
19. M. A. Sall', Teor. Mat. Fiz. **52**, 227 (1982).
20. A. S. Kovalev, A. D. Kondratyuk, and A. I. Landau, Preprint No. 26-89, FTINT AN USSR (Khar'kov, 1989).
21. H. Toda, Prog. Theor. Phys. Suppl. **45**, 174 (1970).
22. H. Toda, Prog. Theor. Phys. Suppl. **59**, 1 (1976).

Translated by V. Sipachev

Journal of Experimental and Theoretical Physics, Vol. 95, No. 5, 2002, pp. 973–979.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 122, No. 5, 2002, pp. 1125–1132.

MISCELLANEOUS

Autowaves in a Near-Threshold Medium

A. Yu. Dovzhenko[a], S. V. Maklakov[b], I. E. Rumanov[b], and E. N. Rumanov[b, *]

[a]*Landau Institute for Theoretical Physics, Russian Academy of Sciences, Chernogolovka, Moscow oblast, 142432 Russia*

[b]*Institute of Structural Macrokinetics and Material Science Problems, Russian Academy of Sciences, Chernogolovka, Moscow oblast, 142432 Russia*

**e-mail: ed@ism.ac.ru*

Received April 12, 2002

Abstract—Approximate and numerical methods are used to study the behavior of autowaves for parameters close to the propagation threshold. Under these conditions, the variations in wave velocity and amplitude are slow. A quasi-steady-state equation is derived for the velocity. This equation describes the relaxation to a steady state (uniform motion) in the above-threshold region and the initial damping stage that determines the time scale of this process in the below-threshold region. As the threshold is approached, the time scales indefinitely increase in the above- and below-threshold regions of parameters. Small random inhomogeneities of the active medium and other "noise" sources produce intense velocity pulsations. These pulsations are comparable in scale to the mean velocity (as in the case of strong turbulence) and resemble the critical fluctuations in order parameter near the point of a continuous phase transition in their statistical properties. The pulsation spectrum exhibits a sharp peak at zero frequency. In contrast to flicker noise, this peak disappears as one recedes from the threshold. The solutions to the quasi-steady-state equation and the results of numerical simulations agree as long as the fluctuations are small—as in the theory of continuous transitions, beyond the fluctuation region. © *2002 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

Traveling pulses are a variety of autowaves [1]. The pulse structure can be clearly imagined by considering a reaction wave moving along a tube with porous walls. Because the reaction products behind the wave filter through the walls, they are gradually replaced with the original material from the space surrounding the tube. Since the mixture composition that was before the wave passage is restored, a new pulse can travel in the wake of the first pulse. Thus, a periodic autowave is a sequence of pulses. In a two-dimensional reactor (e.g., a Petri dish), a traveling pulse can turn into a rotating spiral. Indeed, if the tube is closed into a ring, then the pulse will rotate with a period that, of course, must exceed the restoration time of the original mixture composition (this requires that the ring radius be large enough). A flat layer of the active medium can be represented as a set of such concentric rings. Since the pulse velocity, clearly, does not depend on the ring radius, the pulses in the outer rings will lag behind to form a spiral bent in the direction of rotation. The circular waves and other complex structure diverging from the leading center can be "constructed" from traveling pulses.

The reaction wave that forms the leading edge of the pulse propagates due to the diffusion of active particles into the original mixture. In order of magnitude, its velocity u and width b are

$$u \sim \sqrt{\chi/\tau_R}, \quad b \sim \sqrt{\chi\tau_R}, \tag{1.1}$$

where χ is the activator diffusion coefficient and τ_R is the reaction time scale. Estimate (1.1) suggests that this time is short compared to the restoration time. In this case, the restoration stage does not affect appreciably the pulse itself (its velocity and amplitude). It is important only for the formation of periodic waves, spirals, etc. At the same time, activator losses (e.g., through the tube walls) slow down the reaction. The velocity u decreases and b increases, which contributes to the losses. The existence of a propagation threshold, which was detected in various autowaves (see the review [2]), is the result of such a feedback. From a mathematical point of view, the problem of wave propagation with losses has two solutions; the fast wave is stable, while the slow wave is unstable. At the bifurcation point, the two solutions merge and disappear, so this point determines the threshold value for a control parameter (e.g., the ratio of loss time to τ_R). The propagation threshold was first found in [3] for combustion waves. In this case, heat acts as the activator. Because of the activation dependence of the reaction rate on temperature, the region where the reaction takes place is narrow compared to the wavelength. This allowed Zel'dovich [3] to derive approximate formulas for the velocity of steady-state (uniformly moving) waves and a condition for their coalescence.

We consider unsteady-state waves propagating under near-threshold conditions. Under these conditions, the relaxation to uniform motion (for above-threshold parameters) or the wave damping (in the

1063-7761/02/9505-0973$22.00

below-threshold region) is slow, which makes it possible to derive a quasi-steady-state equation that describes the variations in wave velocity with time. Although a model of an activated exothermic reaction is used to derive the equation, its form near the threshold does not depend on the specific model and is valid for any reaction waves. It thus follows that this conclusion can be extended to traveling pulses and, eventually, to all of the autowaves constructed from them. The difference between the stable and unstable solutions is small near the threshold. This gives rise to a situation of an almost indifferent equilibrium similar to that arising near the point of a continuous phase transition. Deviations from a steady state (a uniformly moving wave) arise easily and relax slowly. These deviations (critical fluctuations) are thermal for continuous phase transitions, while, in our case, they are attributable to inhomogeneities of the active medium and to other sources of "noise." In general, the latter significantly exceeds the thermal pumping. As regards the statistical properties of the velocity fluctuations (the frequency spectrum, etc.), they do not depend on noise and are inherent in the wave itself. Indeed, the correlation time of the fluctuations increases indefinitely as the threshold is approached, whereas, for noise, this time clearly remains finite. Therefore, any noise near the threshold can be treated as δ-correlated (white) noise.

The formula for the velocity correlator derived from the quasi-steady-state equation in the linear approximation is invalid in the near-threshold region, where the pulsations cease to be small. The boundaries of this "fluctuation" region can be determined by comparing, by analogy with phase transitions, the rms deviation of the velocity and the deviation of the mean velocity from its threshold value. For continuous phase transitions, the extent of the fluctuation region is a parameter of the material that undergoes a transition. In an active system, this extent together with the fluctuation intensity depends on the noise level. In particular, this allows the fluctuation region to be extended to facilitate its study by introducing controllable noise.

2. THE QUASI-STEADY-STATE EQUATION

According to (1.1), the velocity of a traveling pulse does not depend on the restoration stage. If the restoration time is large compared the activator loss time τ_c, then the threshold effect does not depend on this stage either. Therefore, we disregard the filtering through the tube walls and consider the wave of activated transformation of the original material into a product instead of the traveling pulse. For an exothermic reaction, heat acts as the activator. The wave moves because of heat transport to the original material, heating causes an intense reaction, heat is released, etc. In this case, thermal diffusivity is the transport coefficient in (1.1). If the activation energy E is high enough, the heat release is concentrated near the thermal peak—in the reaction zone. Ahead, in the heating zone, the reaction can be ignored, because the temperature is low; behind, in the cooling zone, the reaction does not proceed, because all of the material was used up. The three-zone wave structure is characterized by three parameters: the Lewis number L, the ratio of diffusion coefficient to thermal diffusivity; the Semenov number S, the ratio of cooling time τ_c to reaction time $\tau_R(T_b)$,

$$T_b = T_0 + \frac{Q}{c}, \tag{2.1}$$

where T_0 is the initial temperature, Q is the reaction heat, and c is the heat capacity; and the Zel'dovich number

$$Z = \frac{E}{T_b^2}(T_b - T_0). \tag{2.2}$$

For $Z \gg 1$, the reaction-zone width b is small compared to the wave width. Therefore, we can roughly consider this zone as a surface on which the heat and diffusion fluxes change abruptly [3]. Thus, we have the equations for the wave

$$\frac{\partial n}{\partial t} + u(t)\frac{\partial n}{\partial x} - L\frac{\partial^2 n}{\partial x^2} = -\varphi\delta(x), \tag{2.3}$$

$$\frac{\partial T}{\partial t} + u(t)\frac{\partial T}{\partial x} - \frac{\partial^2 T}{\partial x^2} = \varphi\delta(x) - \frac{T}{S}, \tag{2.4}$$

written in a coordinate system in which the reaction zone is at rest at $x = 0$. The boundary conditions are

$$\begin{aligned} x &\longrightarrow -\infty, \quad n = 1, \quad T = 0; \\ x &\longrightarrow \infty, \quad n = 0, \quad T = 0. \end{aligned} \tag{2.5}$$

We use $\tau_R(T_b)$ as the time scale and choose the length and velocity scales according to (1.1). Here, n is the concentration of the original material, T is the temperature measured from T_0 on the scale $T_b - T_0$, $u(t)$ is the instantaneous wave velocity,

$$\varphi^2 = u_0^2 \exp[-Z(1 - T_m)], \tag{2.6}$$

u_0 is the velocity of uniform wave motion for $S \longrightarrow \infty$, and $T_m = T(x = 0)$. Problem (2.3)–(2.6) for

$$S > S_{\text{th}} \approx 2eZu_0^{-2}$$

has two steady-state solutions [3]

$$u_1(S) > u_{\text{th}} \approx u_0/\sqrt{e}, \quad u_2(S) < u_{\text{th}},$$

with the lower branch being unstable. For $S < S_{\text{th}}$, there are no solutions in the form of steady-state waves. At the threshold $S = S_{\text{th}}$, the difference $1 - T_m \approx Z^{-1} \ll 1$, which justifies the approximate representation of the activation law (2.6).

Consider the unsteady-state solutions to this problem that describe the relaxation to a uniformly moving wave for $S > S_{\text{th}}$ and its damping for $S < S_{\text{th}}$. These pro-

cesses are slow near S_{th}, which, as we will see below, allows an approximate equation to be derived for the velocity $u(t)$. For $t \gg u_{th}^{-2}$, the effect of the initial conditions must vanish. According to (2.3)–(2.5), we can then write for the point $x = 0$

$$\frac{1}{2\sqrt{\pi L}}\int_0^t \frac{\varphi(t')}{\sqrt{t-t'}}\exp\left[-\frac{l^2(t', t)}{4L(t-t')}\right]dt' \approx 1, \qquad (2.7)$$

$$\frac{1}{2\sqrt{\pi}}\int_0^t \frac{\varphi(t')}{\sqrt{t-t'}}\exp\left[-\frac{l^2(t', t)}{4(t-t')} - \frac{t-t'}{S}\right]dt' \approx T_m(t), \quad (2.8)$$

where

$$l(t', t) = \int_{t'}^{t} u(t'')dt''. \qquad (2.9)$$

For (2.7), the total transformation of the material is assumed, so $n = 0$ behind the reaction zone. The integrands in (2.7) and (2.8) are nonzero only in a time interval of the order of u_0^{-2} near the upper limit.

Equalities (2.6)–(2.8) allow φ, u, and T_m to be determined. When the wave moves uniformly, $u = \varphi$. Near the threshold, the velocity changes slowly and u differs from φ only slightly. In the integrands of (2.7) and (2.8), we use the following expansions everywhere except the exponentials:

$$\varphi(t') \approx \varphi(t) - (t-t')\frac{du}{dt},$$

$$u(t') \approx u(t) - (t-t')\frac{du}{dt}, \qquad (2.10)$$

$$\langle u\rangle \approx u(t) - \left(\frac{1}{2}\right)(t-t')\frac{du}{dt},$$

where $\langle u\rangle$ is defined by the equality

$$l(t', t) = \langle u\rangle(t-t').$$

If the condition

$$\left|\frac{du}{dt}\right| \ll u_{th}^3 \qquad (2.11)$$

is satisfied, then expansions (2.10) are valid as long as the difference $(t-t') \le u_{th}^{-2}$. As we said above, only this t' interval gives a significant contribution when integrals (2.7) and (2.8) are calculated. For their calculations, we introduce a new integration variable in (2.7), which is related to t' by

$$\xi_1^2 = l^2(t', t)[4L(t-t')]^{-1}.$$

The integration over ξ_1 should be carried out in the interval from 0 to $l(0, t)/(2\sqrt{t})$ Since the integral rapidly converges, the upper limit may be set to infinity. The left-hand side of (2.7) then takes the form

$$\frac{2}{\sqrt{\pi}}\int_0^\infty e^{-\xi_1^2}\varphi(\xi_1)[2u(\xi_1) - \langle u\rangle]^{-1}d\xi_1. \qquad (2.12)$$

Let us simplify the factor at the exponential in (2.12) by retaining the terms with the derivative du/dt to a power not higher than the first power:

$$1 + u_{th}^{-1}\left[\varphi(t) - u(t) + 2L\xi_1^2 u_{th}^{-2}\frac{du}{dt}\right]. \qquad (2.13)$$

The instantaneous and mean velocities in (2.13) were substituted with the threshold value u_{th} everywhere except the small difference $\varphi - u$. Substituting (2.12) and (2.13) into (2.7) yields

$$u - \varphi = Lu_{th}^{-2}\frac{du}{dt}. \qquad (2.14)$$

Let us perform similar transformations for the integrand in (2.8). After changing the integration variable $t' \longrightarrow \xi_2$,

$$\xi_2^2 = l^2(t', t)[4(t-t')]^{-1} + (t-t')S^{-1}; \qquad (2.15)$$

using expansions (2.10); and linearizing in du/dt, we obtain

$$T_m(t) - \frac{\varphi}{\sqrt{\varphi^2 + 4/S}} = -u_{th}^{-3}K\frac{du}{dt},$$
$$K = L - 1 + \frac{2}{Z}(L+2). \qquad (2.16)$$

A zero K would imply that we cannot restrict ourselves to the first order in the expansion in du/dt used. Therefore, (2.16) is valid as long as K is not too small. Curiously, the $K = 0$ line for large Z is close to the boundary of oscillatory instability for the wave under consideration [4]. Having eliminated the temperature T_m from (2.6) and (2.16), we substitute u for φ in the derived equality. Indeed, according to (2.14), the difference $\varphi - u$ is proportional to du/dt. However, the above substitution results in corrections $\sim(du/dt)(u - u_{th})$, which may be disregarded near the threshold. Introducing the quantities

$$s = SS_{th}^{-1} - 1, \quad v = uu_{th}^{-1} - 1 \qquad (2.17)$$

and retaining the highest terms in s and v, we obtain the equation

$$\frac{dv}{dt} = \frac{s - 2v^2}{t_*}, \quad t_* = \frac{KZ}{u_{th}^2}. \qquad (2.18)$$

For $S > S_{th}$, Eq. (2.18) has two solutions that correspond to a constant velocity, with the slow branch $u(S)$ being unstable. For $S < S_{th}$, the velocity decreases with time and the wave is damped. For $|s| \ll 1$, the slowdown

becomes small when the velocity approaches u_{th}.This stage of slow variations is typical of the growth of explosive instability near the limit (see [5]) and of chaotic regimes with intermittency [6]. If two above-threshold regions are separated by a below-threshold segment about $u_{th}t_*|s|^{-1/2}$ in length, then the damped wave can traverse this segment, whereupon its velocity and amplitude are restored in the above-threshold region [7]. This "tunneling" significantly affects the behavior of autowaves under near-threshold conditions. Since any actual medium is inhomogeneous, it consists of randomly arranged above- and below-threshold segments when its mean parameters are close to their threshold values. In this case, the wave percolation condition does not require the existence of an above-threshold cluster threading the entire medium. The appearance of small but closely spaced clusters will suffice.

3. VELOCITY FLUCTUATIONS

Consider a stable wave that uniformly moves at constant $S > S_{th}$ with velocity u_{st}, with the temperature in the reaction zone being

$$T(x = 0) = T_{st} \approx 1 - 2(Su_{st}^2)^{-1}. \tag{3.1}$$

Now, let the parameter S include a small addition that is a random function of time:

$$\begin{aligned} S^{-1} &= S_0^{-1}[1 + y(t)], \\ \langle y(t_1)y(t_2)\rangle &= \langle y^2\rangle\delta(t_1 - t_2), \quad \langle y^2\rangle \ll 1. \end{aligned} \tag{3.2}$$

The following term will then be added to the left-hand side of (2.8):

$$\begin{aligned} &-\frac{T_{st}}{2S_0\sqrt{\pi}}\int_0^t \frac{y(t')dt'}{\sqrt{t-t'}} \\ &\times\left\{\int_{-\infty}^0 d\xi \exp[k_+\xi - A(t, t', \xi)]\right. \\ &\left.+\int_0^\infty d\xi \exp[k_-\xi - A(t, t', \xi)]\right\}, \end{aligned} \tag{3.3}$$

where

$$k_\pm = \frac{u_{st}}{2} \pm \left(\frac{u_{st}^2}{4} + S_0^{-1}\right)^{1/2},$$

$$A(t, t', \xi) = \frac{t-t'}{S_0} + \frac{[u_{st}(t-t') - \xi]^2}{4(t-t')}.$$

Accordingly, Eq. (2.18) linearized near u_{st} takes the form

$$\frac{dw}{dt} + \gamma w = -f(t), \tag{3.4}$$

where

$$w = v - v_{st}, \quad v_{st} = \sqrt{\frac{s}{2}}, \quad \gamma = 2\sqrt{2s}\frac{1}{t_*},$$

$$\begin{aligned} f(t) &= \frac{8}{KS_0\sqrt{\pi}}\int_0^{(u_{st}/2)\sqrt{t}} d\zeta y(t - 4\zeta^2 u_{st}^{-2})2\zeta \\ &\times\int_0^\infty \exp[-(\nu + \zeta)^2]d\nu. \end{aligned} \tag{3.5}$$

Given (3.2) and in the asymptotic limit of large t and t', the correlation function

$$\begin{aligned} \langle w(t)w(t')\rangle &= \exp[-\gamma(t + t')] \\ &\times\int_0^t\int_0^{t'} \langle f(t_1)f(t_2)\rangle \exp[\gamma(t_1 + t_2)]dt_1dt_2 \end{aligned} \tag{3.6}$$

reduces to

$$\langle w(t)w(t')\rangle \approx \frac{1}{2}\left(\frac{u_{th}}{KS_{th}}\right)^2\frac{\langle y^2\rangle}{\gamma}\exp(-\gamma|t - t'|). \tag{3.7}$$

In writing the coefficients on the right-hand side of (3.7), we substituted $u_{st} \longrightarrow u_{th}$ and $S_0 \longrightarrow S_{th}$. For thermal fluctuations, the "pumping" $\langle y^2\rangle$ is chosen to provide an equilibrium value for the mean square of the deviations for the fluctuating quantity (see [8]). In this case, $\langle y^2\rangle \propto \gamma$. For the autowave under consideration, $\langle y^2\rangle$ is determined by inhomogeneities of the medium and by other noise sources; clearly, it does not depend on γ. The growth of fluctuations near the threshold ($\gamma \longrightarrow 0$) results from a low stability of the steady state (a uniformly moving wave) under these conditions. This effect is similar to the critical behavior near the point of a continuous phase transition. According to (3.7), the velocity fluctuations have a Lorentz spectrum with a peak at $\omega = 0$; the peak width is $\Delta\omega = \gamma$. Significantly, the properties of the fluctuations (except their intensity) do not depend on noise. The correlation time increases as the threshold is approached, whereas, for noise, this time clearly remains limited. Therefore, any noise near the threshold can be assumed to be a δ-correlated one. Note also that, in contrast to flicker noise [9], the low-frequency peak disappears as one recedes from the threshold. We took into account the pattern obtained by using the quasi-steady-state equation when analyzing the results of our numerical simulations of autowaves.

4. NUMERICAL SIMULATIONS

We studied a wave with a narrow (but finite) reaction zone. We chose $L = 1$ to avoid complications related to diffusion instability. Such a wave can be simulated by using the equations

$$\frac{\partial \eta}{\partial t} + u_{tr}\frac{\partial \eta}{\partial x} - \frac{\partial^2 \eta}{\partial x^2} = \Phi(\eta, \theta),$$
$$\Phi(\eta, \theta) = (1-\eta)\exp\left(\frac{\theta}{1+\beta\theta}\right), \tag{4.1}$$

$$\frac{\partial \theta}{\partial t} + u_{tr}\frac{\partial \theta}{\partial x} - \frac{\partial^2 \theta}{\partial x^2} = Z\Phi(\eta, \theta) - \frac{\theta + Z}{S} \tag{4.2}$$

with the boundary conditions

$$x = 0, \quad \eta = 0, \quad \theta = -Z;$$
$$x = x_0, \quad \eta = 1, \quad \frac{\partial \theta}{\partial x} = 0. \tag{4.3}$$

In (4.1)–(4.3), the concentration of the reaction product is $\eta = 1 - n$ and the dimensionless temperature θ is measured from T_b on the scale $T_b^2 E^{-1}$, so $\theta = -Z$ corresponds to T_0; $\beta = T_b/E$. Since, in contrast to (2.5), we have to take a finite interval, $0 < x < x_0$, for our numerical simulations, the boundary conditions must soften the effect of the ends, if possible. At a sufficiently large x_0, we can restrict ourselves to substituting the condition $T(x \longrightarrow \infty) = 0$ for $\partial\theta/\partial x = 0$ at $x = x_0$, as in (4.3). The following simplified form of a steady-state wave was used as the initial conditions:

$$0 < x < \frac{x_0}{2}, \quad \eta = \exp\left[u_{tr}\left(x - \frac{x_0}{2}\right)\right],$$
$$\theta + Z = Z\exp\left[k_+\left(x - \frac{x_0}{2}\right)\right],$$
$$\frac{x_0}{2} < x < x_0, \quad \eta = 1, \quad \theta + Z = Z\exp\left[k_-\left(x - \frac{x_0}{2}\right)\right],$$
$$k_\pm = \frac{u_{tr}}{2} \pm \left[\left(\frac{u_{tr}}{2}\right)^2 + S^{-1}\right]^{1/2}.$$

This choice allows us to speed up the attainment of a steady (for $S > S_{th}$) or quasi-steady (for $S < S_{th}$) state. The instantaneous wave velocity was determined from the motion of the maximum of the function Φ. In the above-threshold region $S > S_{th}$, the "trail" velocity u_{tr} was chosen to reduce the displacement of maxΦ from the point $x_0/2$, if possible. In this case, the effect of the end points $x = 0$ and $x = x_0$ on the wave parameters was negligible (for $x_0 = 2000$). Thus, we found the dependence $u_{st}(S)$ and the characteristic relaxation time to a steady state t_r (this time was determined from the time at which the difference between the instantaneous velocity and u_{st} decreased to $0.01u_{st}$). Subsequently, we introduced noise in the form (3.2) at $u_{tr} = u_{st}$; the product of 0.1 and random numbers from the interval (–1, +1), which changed at each step, was used as the function $y(t)$. A built-in subroutine of the GNU Fortran-77 compiler served as the random number generator.

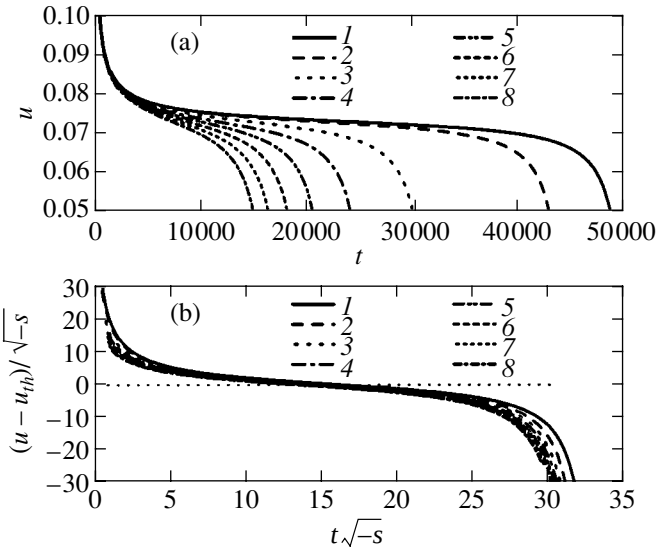


Fig. 1. Damped-autowave velocity versus time. (a) Numerical simulations: $10^4 S = 2.469$ (*1*), 2.470 (*2*), 2.475 (*3*), 2.480 (*4*), 2.485 (*5*), 2.490 (*6*), 2.495 (*7*), and 2.500 (*8*). (b) The same curves after scaling the axes.

In all cases, we used $Z = 10$ and $\beta = 0.09$. For our system of parabolic equations, we employed an implicit scheme that ensured stability and convergence of the computational procedure [10]; the steps in space and in time were 0.1 and 0.2, respectively. We also computed two cases with a spatial step of 0.01 for $10^4/S$ equal to 2.3 and 2.4. Decreasing the step did not cause significant changes in the computational results: the mean source power Φ in a steady state was conserved with a relative error $<10^{-4}$. For the above parameters, the width of the reaction zone in the waves under consideration is $b \approx 1$ and the width of the heating zone is $Zb \approx 10$; the corresponding (Michelson) time scale is $\tau \approx 100$ (because the velocity scale is about 0.1). The procedure described above yielded $u_{th} \approx 0.073 \pm 0.001$ and $S_{th} \approx 4057 \pm 1$. Figure 1 shows $u(t)$ for several values of S ($<S_{th}$). We see how the wave damping slows down as the threshold is approached. This slowdown and the curve shape are in agreement with the solution to the quasi-steady-state equation (2.18) for $s < 0$:

$$v(t) = \sqrt{-\frac{s}{2}}\cot\left(\frac{t}{t_*}\sqrt{-2s} + \alpha_i\right),$$
$$\alpha_i = \operatorname{arccot}\left(v_i\sqrt{-\frac{2}{s}}\right), \tag{4.4}$$

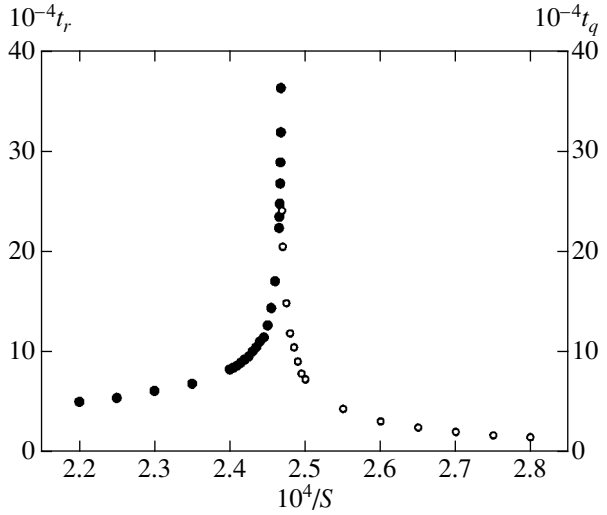


Fig. 2. Time scale for autowaves versus distance from the threshold. The time t_r is determined from the time when $u - u_{st}$ decreases to $0.01u_{st}$ and t_q is determined from the position of the point of inflection in the curves of Fig. 1a (and similar curves for other S).

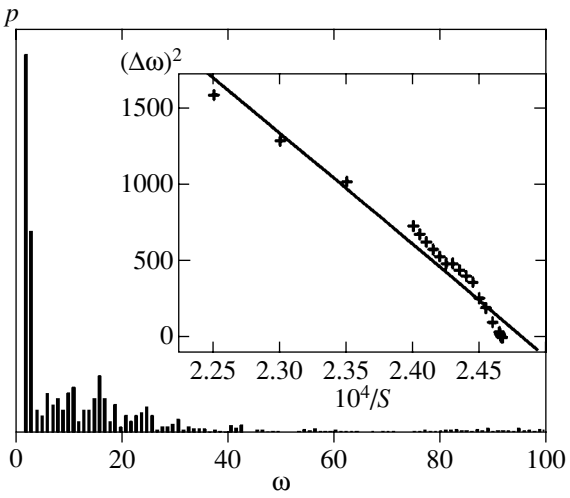


Fig. 3. A typical spectrum of the $u(t) - u_{st}$ velocity pulsations triggered by white noise. The frequency ω is in units of $2\pi/t_0$; the computational time is $t_0 = 60\,000$. The spectral-peak width Δω is plotted against distance from the threshold in the inset.

where $v_i = v(t = 0)$. According to (4.4), the damping time scale $t_q \propto (-s)^{1/2}$. As was noted in Section 2, this behavior is typical of the growth of explosive instability and of regimes with intermittency. After scaling, the curves approached but did not coalesce because of the small inaccuracies in determining S_{th} and u_{th}. (When a value of S close to S_{th} is specified in computations, the relative error in the difference between these values, i.e., s, increases. A similar remark is true for $u - u_{th}$.) In general, the agreement between the simulation results and (4.4) is satisfactory.

The time scale is plotted against S in Fig. 2. The plot resembles the temperature dependence of susceptibility, scattering cross sections, etc., near the point of a continuous phase transition. However, the singularity at $S = S_{th}$ is different in nature from that for continuous transitions. In the latter case, a new phase can be obtained from the old phase by means of a small deformation. The wave solution disappears at the threshold (explosive instability); the difference between the old and new regimes is finite. The left and right parts of the plot in Fig. 2 have a different meaning. This is the relaxation time to a steady state (uniform wave motion) for $S > S_{th}$ and the time of transition to a new regime (in our case, the wave damping time) for $S < S_{th}$. If we continue our comparison with phase transitions, then a finite phase difference (e.g., in density) corresponds to a first-order phase transition. A phase equilibrium is possible in such a transition. The position of the phase boundary is stable, for example, for fixed system temperature and volume. When the volume is varied, the boundary is displaced and one of the phases grows at the expense of the other. Without allowance for the boundary energy, the transition appears smooth until the complete disappearance of this phase. Actually, for a sufficiently small amount of phase, it is unstable and disappears abruptly—there is a threshold in this case as well. Let, for example, a liquid drop be in equilibrium with vapor. During an isothermal expansion, its radius decreases because of its evaporation. When the critical radius corresponding to unstable (Gibbs) equilibrium is reached, the drop abruptly disappears and the pressure increases. For first-order phase transitions, the near-threshold region is more difficult to observe than that for autowaves.

For the cases with the inclusion of noise in the form (3.2), the steady state ($u = u_{st}$) found without any noise addition played the role of initial conditions. The value of $\langle y^2 \rangle$ calculated for checking purposes is close to its theoretical value (1/300). At such a noise level and for L and Z given above, $s \le 10^{-6}$ must correspond to the fluctuation region. This estimate is obtained if the fluctuation region is specified by the condition $\langle w^2 \rangle \ge v_{st}^2$, by analogy with the theory of continuous transitions. The variance for the velocity pulsations and its increase as the threshold is approached obtained during our computational experiments agree with (3.7). It should be noted that, without noise, pulsations related to rounding-off in our computations are observed. These pulsations also grow as $S \longrightarrow S_{th}$, remaining much smaller than the pulsations due to the introduced noise for all S (the variance ratio is no less than 10). A typical pulsation spectrum is shown in Fig. 3. The "line" frequencies are $\omega_n = 2\pi n/t_0$, where n is an integer and t_0 is the computational time. The line intensity distribution corresponds to a continuous spectrum. No resonances were observed at $\omega \neq 0$. The peak at zero frequency nar-

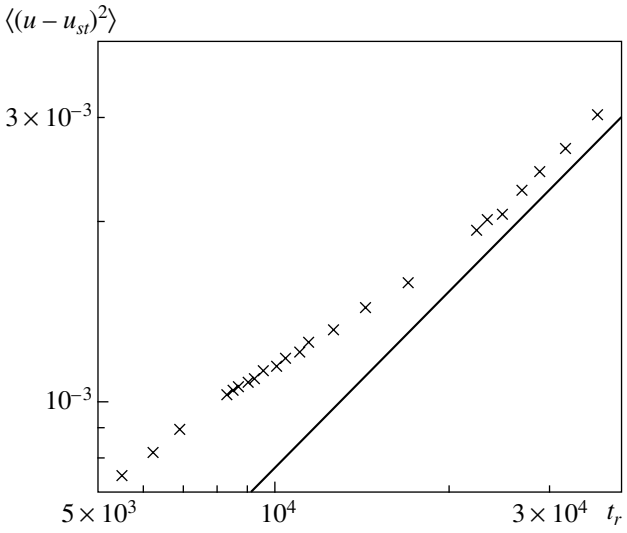


Fig. 4. Relation between the dispersion of velocity pulsations, $\langle (u - u_{st})^2 \rangle$, and the relaxation time to a uniformly moving wave ($u = u_{st}$). The data points correspond to S = 4444, 4348, 4255, 4167, 4149, 4132, 4115, 4098, 4082, 4065, and 4057. For convenience, a straight line with a slope angle of $\pi/4$ is drawn in the figure.

rows as the threshold is approached; this narrowing is shown in the insert to Fig. 3. The time t_0 near the threshold was increased to 120 000 to satisfy the condition $t_0\Delta\omega \gg 1$. Since the line intensity is a random function of frequency, we used the following procedure to determine the peak width $\Delta\omega$. The sum of intensities was computed for the first N lines. The number N was increased until this sum ceased to increase (with a relative error $<10^{-6}$). Let the maximum sum computed in this way be equal to I. Then, $\Delta\omega$ is the frequency of the line with number N_1 at which the sum of intensities for the first N_1 lines is equal to $I/2$. There is no saturation for white noise, $N_1 \approx N/2$; $\Delta\omega = \gamma \propto \sqrt{s}$ for a Lorentz spectrum. As we see from Fig. 3, the dependence of $(\Delta\omega)^2$ on s is actually almost linear: near the threshold, $S^{-1} \approx (1 - s)/S_{th}$.

For each $S > S_{th}$, the velocity variance determined in our computation with noise can be correlated with the relaxation time shown in Fig. 2 (the latter was, of course, determined from our computations without noise). In approximation (3.7), these two quantities are proportional to $1/\gamma$. If we show the plot on a log–log scale, as in Fig. 4, then the data points must lie near the straight line with a slope angle of $\pi/4$. Using the least-squares method, we obtained a slope that can be called a critical index equal to 0.764.

5. CONCLUSION

A detailed analysis of our numerical simulations indicates that the behavior of the wave velocity and its fluctuations can be qualitatively described by a quasi-steady-state equation in the range of parameters where the fluctuations are small.

Based on the equality $2v^2 = s$, which defines the velocity of steady-state waves near the threshold [3], we can assume that the small deviations of v from v_{st} obey the quasi-steady-state equation (2.18). In this case, however, the parameter t_* remains indefinite. The derivation of Eq. (2.18) in terms of the problem (2.3)–(2.6) allows t_* to be determined, with condition (2.11) being satisfied. The "parabolic" form of the right-hand side of Eq. (2.18) corresponds to any case of explosive instability (coalescence of the two branches of the solution). The steady state (e.g., an autowave) near the point of coalescence is formally stable against infinitely small perturbations. At the same time, the basin of attraction of this regime in state space is small, which, as we see, leads to an increase in the intensity of low-frequency fluctuations (soft modes) and in their correlation time. In this respect, the hard transition made by the system because of the growth of explosive instability does not differ from its soft transition. The parameters of the new pattern formed by the hard transition cannot be determined using the quasi-steady-state equation. However, near the bifurcation point, the transition includes a slow stage, which allows its time scale to be estimated.

ACKNOWLEDGMENTS

We thank A.G. Merzhanov and L.N. Shchur for helpful remarks. This study was supported by the Russian Foundation for Basic Research (project no. 01-02-17487).

REFERENCES

1. V. A. Vasil'ev, Yu. M. Romanovskiĭ, and V. G. Yakhno, Usp. Fiz. Nauk **128**, 625 (1979) [Sov. Phys. Usp. **22**, 615 (1979)].
2. A. G. Merzhanov and E. N. Rumanov, Rev. Mod. Phys. **71**, 1173 (1999).
3. Ya. B. Zel'dovich, Zh. Éksp. Teor. Fiz. **11**, 159 (1941).
4. A. P. Aldushin and S. G. Kasparyan, Dokl. Akad. Nauk SSSR **244**, 67 (1979) [Sov. Phys. Dokl. **24**, 29 (1979)].
5. A. G. Merzhanov, V. V. Barzykin, and V. G. Abramov, Khim. Fiz. **15**, 3 (1996) [Chem. Phys. Repts. **15**, 793 (1996)].
6. P. Manneville and Y. Pomeau, Physica D (Amsterdam) **1**, 219 (1980).
7. A. Yu. Dovzhenko, S. V. Maklakov, and É. N. Rumanov, Dokl. Akad. Nauk **374**, 321 (2000) [Dokl. Phys. **45**, 446 (2000)].
8. L. D. Landau and E. M. Lifshitz, *Course of Theoretical Physics*, Vol. 5: *Statistical Physics* (Nauka, Moscow, 1995; Pergamon, Oxford, 1980), Part 1.
9. M. B. Weissman, Rev. Mod. Phys. **60**, 537 (1988).
10. A. A. Samarskiĭ, *The Theory of Difference Schemes* (Nauka, Moscow, 1989), p. 279.

Translated by V. Astakhov