Journal of Experimental and Theoretical Physics, Vol. 98, No. 3, 2004, pp. 395–402.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 125, No. 3, 2004, pp. 453–461.

NUCLEI, PARTICLES, AND THEIR INTERACTION

Polarization of an Electron–Positron Vacuum by a Strong Magnetic Field with an Allowance Made for the Anomalous Magnetic Moments of Particles

V. N. Rodionov

Moscow State Geological Prospecting University, Moscow, 118873 Russia
e-mail: physics@msgpa.ru

Received June 10, 2003

Abstract—Given the anomalous magnetic moments of electrons and positrons in the one-loop approximation, we calculate the exact Lagrangian of an intense constant magnetic field that replaces the Heisenberg–Euler Lagrangian in traditional quantum electrodynamics (QED). We have established that the derived generalization of the Lagrangian is real for arbitrary magnetic fields. In a weak field, the calculated Lagrangian matches the standard Heisenberg–Euler formula. In extremely strong fields, the field dependence of the Lagrangian completely disappears and the Lagrangian tends to a constant determined by the anomalous magnetic moments of the particles. © *2004 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

The quantum corrections to the Maxwellian Lagrangian of a constant electromagnetic field were first calculated by Heisenberg and Euler [1] in 1936. Radiative corrections corresponding to polarization of an electron–positron vacuum by external electromagnetic fields with diagrams containing different numbers of electron loops are still the focus of attention [2–4]. Estimates suggest that the quantum (radiative) corrections could reach the Maxwellian energy density of the electromagnetic field only in exponentially strong electromagnetic fields[1] ($F_c \sim \exp(3\pi/\alpha)H_c$) [5]. The calculations by Heisenberg and Euler are known to contain no approximations in the intensity of external electromagnetic fields, and their results have been repeatedly confirmed by calculations performed in terms of different approaches. On this basis, several authors have identified the field intensity F_c with the validity boundary of universally accepted QED. However, it is clear that, although such quantities were greatly overestimated because the corresponding scale lengths are many orders of magnitude smaller than not only the scale on which weak interactions manifest themselves, but also the Planck length, determining the validity range of traditional QED is currently of fundamental importance. While on the subject of the physics of extremely small distances, we cannot but say that there is a strong analogy[2] between the phenomena that arise for large momentum transfers and the processes in intense electromagnetic fields [2–17]. In fact, the overlapping of seemingly distinctly different areas of physics is not accidental and is suggested by simple dimension considerations.

Allowance for the electromagnetic field intensity based on the exact integrability of the equations of motion is known to play an important role in studying the quantum effects of the interaction between charged particles and the electromagnetic field. In particular, the standard Schwinger correction to the Bohr magneton

$$\mu_0 = \frac{e}{2m},$$

which is called the anomalous magnetic moment of a particle,

$$\Delta\mu = \mu_0\frac{\alpha}{2\pi},$$

manifests itself only in the nonrelativistic limit for weak quasi-static fields [7]. Indeed, when the influence of an intense external field is accurately taken into account, the anomalous magnetic moment of a particle calculated in QED as a one-loop radiative correction decreases with increasing field intensity and increasing energy of the moving particles from the Schwinger value to zero. In particular, for magnetic fields $H \sim H_c$, the anomalous magnetic moment of an electron is described by the asymptotic formula [7, 10]

$$\Delta\mu(H) = \mu_0\frac{\alpha}{2\pi}\frac{H_c}{H}\ln\frac{2H}{H_c}. \quad (1)$$

It follows from (1) that $\Delta\mu(H)$ becomes zero only at one point while decreasing with increasing field. A similar

[1] Here, we use a system of units with $\hbar = c = 1$, $H_c = m^2/|e| = 4.41 \times 10^{13}$ G is the characteristic scale of the electromagnetic field intensity in QED, e and m are the electron charge and mass, and $\alpha = e^2 = 1/137$ is the fine-structure constant.

[2] This analogy was first pointed out by Migdal [6] and Ritus [2].

expression for the anomalous magnetic moment of an electron in an intense constant crossed field $\mathbf{E} \perp \mathbf{H}$ ($E = H$) at $Hp_\perp \gg mH_c$, where $p_\perp$ is the electron momentum component perpendicular to $\mathbf{E} \times \mathbf{H}$, can be represented as [11]

$$\Delta\mu(E) = \mu_0 \frac{\alpha\Gamma(1/3)}{9\sqrt{3}} \left(\frac{3p_\perp H}{mH_c}\right)^{-2/3}. \quad (2)$$

Note that $\Delta\mu(E) \neq 0$ in the entire range of parameters.

Numerous calculations of the Lagrangian for an electromagnetic field (see, e.g., [1–3, 5–7, 11]) have been performed by assuming that the magnetic moment of electrons is exactly equal to the Bohr magneton,[3] i.e., at $\Delta\mu = 0$. However, the following question is of considerable importance in elucidating the internal closeness of QED: What effects will allowance for the anomalous magnetic moments of electrons and positrons produce when calculating the polarization of an electron–positron vacuum by intense electromagnetic fields?

Thus, it is of interest to take the radiative corrections to the Maxwellian Lagrangian of a constant field calculated by the traditional method and compare them with the results that can be obtained from similar calculations by taking into account the nonzero anomalous magnetic moments of particles. The fact that the Lagrangian replacing the Heisenberg–Euler Lagrangian with nonzero anomalous magnetic moments can be calculated by retaining the method of exact solutions of the Dirac equation in arbitrarily intense electromagnetic fields also deserves serious attention. In the approach under development, the suggested theoretical generalization initially contains no constraints on the electromagnetic field intensity.

It should be noted that nonzero anomalous magnetic moments also appear in some of the modified quantum field theories (QFTs) that also describe the electromagnetic interactions. In particular, this is true for a generalization of the traditional QFT known as the theory with "fundamental mass" (see, [12, 13] and references therein). The starting point of this theory is the condition that the mass spectrum of elementary particles is limited. This condition can be represented as

$$m \leq M, \quad (3)$$

where the new universal parameter M is called the fundamental mass. Relation (3) is used as an additional fundamental physical principle that underlies the new QFT. A significant deviation from traditional calculations is the fact that the charged leptons in QED with fundamental mass have magnetic moments that are not equal to the Bohr magneton. This is because, apart from the traditional "minimal" term, the new Lagrangian of the electromagnetic interaction includes "nonminimal" terms. Thus, an electron in modified QED has an anomalous magnetic moment from the outset:

$$\Delta\mu = \mu - \mu_0 = \mu_0 \left(\sqrt{1 + \frac{m^2}{M^2}} - 1\right). \quad (4)$$

An important aspect of the problem under consideration is that the current state of the art in the development of laser physics [14] allows one to carry out a number of optical experiments to directly measure the contributions from the nonlinear vacuum effects predicted by various generalizations of Maxwellian electrodynamics [15]. Therefore, it should be emphasized that experimental verification of the nonlinear vacuum effects with a high accuracy in the presence of relatively weak electromagnetic fields can also provide valuable information about the validity of QED predictions at small distances [16, 17]. Note in passing that precision measurements of various quantities (e.g., the anomalous magnetic moments of an electron and a muon) at nonrelativistic energies, together with studies of the particle interaction at high energies, are of current interest in the same sense.

2. CORRECTION TO THE LAGRANGIAN OF AN ELECTROMAGNETIC FIELD WITH ALLOWANCE FOR ANOMALOUS MAGNETIC MOMENTS OF PARTICLES

Let us consider the correction to the Lagrangian of an electromagnetic field attributable to the polarization of an electron–positron vacuum in the presence of an arbitrarily strong constant magnetic field by taking into account the nonzero anomalous magnetic moments of the particles. To solve this problem, it is convenient, as in the standard approach [5], to represent the electron–positron vacuum as a system of electrons that fill "negative" energy levels. For a constant uniform magnetic field, the Dirac–Pauli equation containing the interaction of a charged lepton with the field (including the anomalous magnetic moment of the particle) has an exact solution [18]. In this case, the energy eigenvalues explicitly depend on the spin orientation with respect to the axis of symmetry specified by the magnetic field direction. Thus, the energy spectrum of an electron that moves in an arbitrarily intense constant uniform magnetic field is

$$E_n(p, H, \xi) = m\left[\frac{p^2}{m^2} + \left(\sqrt{\left|\frac{H}{H_c}\right|(1 + 2n + \xi) + 1} + \xi\frac{\mu - \mu_0}{2\mu_0}\frac{H}{H_c}\right)^2\right]^{1/2}, \quad (5)$$

where p is the electron momentum component along the external field $\mathbf{H}$; $n = 0, 1, 2, \ldots$ is the quantum number of the Landau levels; and $\xi = \pm 1$ characterizes the electron spin component along the magnetic field.

[3] The authors of [4] took into account the anomalous magnetic moments of particles when analyzing the equilibrium processes in a degenerate electron–nucleon gas in a strong magnetic field.

Noting that the radiative correction to the classical density of the Lagrangian is equal, to within the sign, to the total energy density of the electron–positron vacuum in the presence of an external field [5],

$$\mathcal{L}' = -W^H,$$

let us calculate W^H in a constant magnetic field by taking into account the anomalous magnetic moment of the electron. Without dwelling on the details of standard calculations, we represent W^H as

$$W^H = -\frac{|eH|}{(2\pi)^2} \times \int_{-\infty}^{\infty} dp\left[-\varepsilon_0^+(p) + \sum_{n=0}^{\infty}[\varepsilon_n^-(p) + \varepsilon_n^+(p)]\right], \quad (6)$$

where

$$\varepsilon_n^{\pm} = \sqrt{p^2 + m^2\left(\sqrt{1 + 2\frac{H}{H_c}n} \pm \frac{H}{4H_c^*}\right)^2}. \quad (7)$$

Using the Laplace and Fourier integral transforms for the functions that define (6) and performing summation over Landau levels, we can obtain the following formula for $\mathcal{L}'$:

$$\mathcal{L}' = -\frac{m^4\gamma b_1}{8\pi^2}\int_0^{\infty}\frac{d\eta}{\eta^2}e^{-\eta} \times\left[\sinh b + \frac{1}{\pi}\int_{-\infty}^{\infty}\frac{e^{-ix}dx}{x}\cot\left(-\frac{i\gamma\eta}{b_1} + x\gamma\right)\right. \quad (8)$$
$$\left. + {}_1F_2\left(\{1\}, \left\{\frac{1}{4}, \frac{3}{4}\right\}, \frac{b^4}{64x^2}\right)\right],$$

where we use the notation

$$a_1 = \frac{\eta H}{2H_c^*}, \quad b_1 = 1 + \left(\frac{H}{4H_c^*}\right)^2, \quad b = \frac{a_1}{b_1}, \quad \gamma = \frac{H}{H_c};$$

${}_1F_2(z)$ is the generalized hypergeometric function. Formula (8) is an exact expression for the Lagrangian with allowance for the anomalous magnetic moment calculated in the one-loop approximation in an arbitrarily intense magnetic field. An important deviation from the Heisenberg–Euler Lagrangian is that expression (8) contains the additional field scale

$$H_c^* = \frac{m}{4\Delta\mu}. \quad (9)$$

In the theory with fundamental mass, it would be natural to call the quantity

$$H_c^* = \frac{M^2}{e} = \frac{M^2}{m^2}H_c \quad (10)$$

a fundamental field.

Passing to the limits of integration over x from zero to ∞ in (8) and using the evenness of the function ${}_1F_2(z)$, we obtain

$$\mathcal{L}' = -\frac{m^4\gamma b_1}{8\pi^2}\int_0^{\infty}\frac{d\eta}{\eta^2}e^{-\eta} \times\left[\sinh b + \frac{2}{\pi}\int_0^{\infty}\frac{dx}{x}\frac{\sin(2\gamma x)\cos x + \sin x\sinh y}{\cosh(2y) - \cos(2\gamma x)}{}_1F_2(z)\right], \quad (11)$$

where $y = \eta\gamma/b_1$ and $z = -b^4/64x^2$.

In particular, it immediately follows from (11) that

$$\mathrm{Im}\,\mathcal{L}' = 0.$$

The fact that the Lagrangian $\mathcal{L}'$ is real for all possible field intensities suggests the absence of unstable modes; i.e., the vacuum in a constant uniform magnetic field in the case under consideration, as in traditional QED, is stable against the spontaneous production of electron–positron pairs.

Next, let us separate out the integral over x in expression (11). After several obvious substitutions, it reduces to

$$I = \int_0^{\infty}\frac{du[2a_2\sin u\cos(b_2u) + (1 - a_2^2)\sin(b_2u)]}{u[1 + a_2^2 - 2a_2\cos u]} \times {}_1F_2(z_1), \quad (12)$$

where

$$a_2 = e^{-2y}, \quad b_2 = \frac{1}{2\gamma}, \quad z_1 = -\frac{b^4(2\gamma)^2}{64u^2}.$$

Since the expansions

$$\frac{\sin u}{1 + a_2^2 - 2a_2\cos u} = \sin u + a_2\sin(2u) + a_2^2\sin(3u) + \ldots,$$

$$\frac{1 - a_2^2}{1 + a_2^2 - 2a_2\cos u} = 1 + 2a_2\cos u + 2a_2^2\cos(2u) + \ldots$$

are valid, we obtain for (12)

$$I = \int_0^{\infty}\frac{du}{u}\left[-\sin(b_2u) + 2\sum_{k=0}^{\infty}a_2^k\sin[u(k + b_2)]\right] \times {}_1F_2(z_1). \quad (13)$$

It is easy to see that the following expansion of the function ${}_1F_2(z_1)$ at zero may be used in a field that is weak compared to the fundamental field H_c^*:

$$ {}_1F_2(z_1) = 1 + \frac{16}{3}z_1 + \frac{256}{105}z_1^2 + \ldots. \quad (14)$$

Hence, we obtain for (13)

$$ I = \frac{\pi}{2}\frac{1+a_2}{1-a_2} = \frac{\pi}{2}\coth y, \quad (15)$$

where $y = \eta\gamma$.

Substituting (15) into (11) and performing standard regularization of the derived integral [5] yields

$$ \mathcal{L}' = -\frac{m^4}{8\pi^2}\int_0^\infty \frac{e^{-\eta}}{\eta^3}\left[\eta\gamma\coth(\eta\gamma) - 1 - \frac{\eta^2\gamma^2}{3}\right]d\eta. \quad (16)$$

Thus, it follows from (16) that in the limit of a weak field, formula (11) matches the Heisenberg–Euler Lagrangian [1] for an arbitrarily intense constant uniform magnetic field.

Next, let us consider $H > 4H_c^*$. It is easy to verify that in the limit of extremely strong fields, $H \gg 16H_c^{*2}/H_c$, we may again use expansion (14) and can obtain the following expression for integral (13):

$$ I = \frac{\pi}{2}\coth y, \quad (17)$$

where $y = 16\eta H_c^{*2}/(H_c H)$. Results (15) and (17) have a simple meaning: for a sufficiently wide energy gap that separates the electron and positron states, the terms with large numbers k make the largest contribution to integral (13). However, for magnetic fields close to the fundamental field, $H \sim 4H_c^*$, i.e., when the gap width is close to zero, the term with $k = 0$ makes the largest contribution to the sum in the integrand of (13). In this case, integral (13) can be calculated exactly. Our calculations yield

$$ I = \frac{\pi}{2}\cosh\left[\eta\frac{H}{2H_c^*}\frac{1}{1+(H/4H_c^*)^2}\right]. \quad (18)$$

The estimates of integral (12) in the three ranges of magnetic fields ($H \ll H_c^*$, $H \sim 4H_c^*$, and $H \gg H_c^*$) can be represented as a single formula:

$$ I = \frac{\pi}{2}\cosh\left[\eta\frac{H}{2H_c^*}\frac{1}{1+(H/4H_c^*)^2}\right] \times \coth\left[\frac{\eta\gamma}{1+(H/4H_c^*)^2}\right]. \quad (19)$$

Substituting (19) into (11) and regularizing the remaining diverging integral,[4] we obtain

$$ \mathcal{L}' = -\frac{m^4}{8\pi^2}\int_0^\infty \frac{d\eta}{\eta^3}e^{-\eta}\left[\eta b_1\gamma\frac{\cosh[(1+a)\eta\gamma/b_1]}{\sinh(\eta\gamma/b_1)} - b_1^2 - \frac{\eta^2}{6}\gamma^2(2+6a+3a^2)\right], \quad (20)$$

where $a = H_c/2H_c^*$.

3. ASYMPTOTIC RESULTS

In weak fields ($H \ll H_c$) and in the limit of very strong magnetic fields ($H \gg 16H_c^{*2}/H_c$), the integrand in (20) admits an expansion into a series. In the first approximation,

$$ \mathcal{L}' = \frac{m^4\gamma^4}{2880\pi^2 b_1^2}[8 - 15a^2(2+a)^2]\int_0^\infty d\eta\,\eta e^{-\eta}, $$

whence it follows that

$$ \mathcal{L}' = \frac{m^4\gamma^4}{2880\pi^2 b_1^2}[8 - 15a^2(2+a)^2]. \quad (21)$$

Thus, the quantum correction to the Maxwellian Lagrangian in the limit of a weak field ($H \ll H_c$) can be represented as

$$ \mathcal{L}' = \frac{m^4}{360\pi^2}\frac{H^4}{H_c^4} \times\left[1 - \frac{15}{2}a^2 - \frac{15}{2}a^3 - \frac{15}{8}a^4 + \mathcal{O}(\gamma^2; \gamma^2a^2)\right], \quad (22)$$

where the first term matches the standard Heisenberg–Euler formula. The first correction to it attributable to the anomalous magnetic moments of the particles is negative and quadratic in a.

In an extremely strong field ($H \gg 16H_c^{*2}/H_c$), we can also obtain from (21)

$$ \mathcal{L}' = \frac{m^4}{180\pi^2 a^4}(8 - 60a^2 - 60a^3 - 15a^4) \times\left(1 - \frac{8}{\gamma^2a^2}\right). \quad (23)$$

[4] First, as usual [5], the part of the integral that does not contain the magnetic field intensity and that represents the energy of the free vacuum electrons should be discarded. Second, it is necessary to subtract the contribution proportional to H^2 that has already been included in the unperturbed field energy. Discarding this term is related to renormalizing the field intensity and, hence, the charge. Finally, subtracting a contribution on the order of H^4/H^{*4} basically corresponds to renormalizing an additional parameter of the theory—the anomalous magnetic moment of the particle.

According to (23), in the limit of extremely strong fields, the Lagrangian $\mathcal{L}'$ ceases to depend on the field; i.e., as the field grows, the quantum correction to the density of the Lagrangian in the case under consideration asymptotically approaches the constant

$$\mathcal{L}'_\infty = \frac{\alpha^2}{360\pi^2(\Delta\mu)^4}. \quad (24)$$

In a certain sense, the result obtained may be compared with the situation observed in the standard model, where the cross sections for several processes cease to increase with energy if, apart from a photon, vector $W^\pm$- and Z^0 bosons, an additional diagram with a Higgs H boson is included in the analysis. Allowance made for this diagram reduces the increasing terms in amplitude and leads to a behavior of the cross sections consistent with the unitary limit. Since the standard model does not predict the mass of the H-boson, it may well be that this particle is much heavier than the t quark, the heaviest known elementary particle. Thus, $M_H \sim 1$ TeV may prove to be the critical mass that limits the mass spectrum of elementary particles, i.e., acting as the fundamental mass (see (3)).[5]

By comparing the correction $\mathcal{L}'$ with the Lagrangian of the Maxwellian field, we can determine the field intensity

$$F_c^* = \sqrt{\frac{256\alpha}{45\pi}}\frac{H_c^{*2}}{H_c}, \quad (25)$$

at which $\mathcal{L}_0$ becomes equal to (24). For $H = F_c^*$, the quantum correction $\mathcal{L}'$ does not yet reach its asymptotic value of $\mathcal{L}'_\infty$. A comparison of $\mathcal{L}_0$ and $\mathcal{L}'$ in other field ranges clearly shows that the corrections $\mathcal{L}'$ are always small compared to the Lagrangian $\mathcal{L}_0$. The relative corrections $\mathcal{L}'/m^4$ derived from (20) for the anomalous magnetic moments of particles $\Delta\mu_1/\mu_0 = 10^{-3}$, $\Delta\mu_2/\mu_0 = 10^{-3.05}$, and $\Delta\mu_3/\mu_0 = 10^{-3.1}$ are plotted against magnetic field intensity $\gamma = H/H_c$ in Fig. 1.

We estimate the Lagrangian for strong magnetic fields by using formula (20), which we will represent after the substitution $\gamma\eta/b_1 \longrightarrow x$ as

$$\mathcal{L}' = -\frac{m^4\gamma^2}{8\pi^2}\int_0^\infty \frac{\exp(-b_1 x/\gamma)}{x^3} \times\left[x\frac{\cosh[x(1+a)]}{\sinh(x)} - 1 - \frac{x^2}{3}\left(1 + 3a + \frac{3}{2}a^2\right)\right]dx. \quad (26)$$

[5] Note in this connection that the chief goal in the research at the Large Hadron Collider (LHC) at CERN is the search for Higgs bosons in a mass range up to 1 TeV.

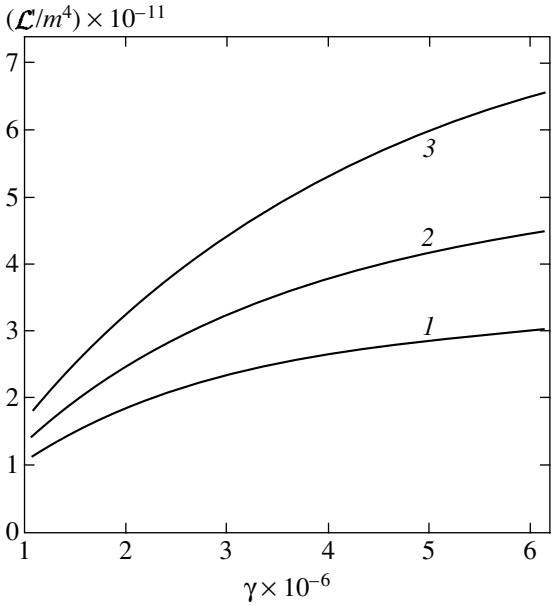

Fig. 1. Normalized Lagrangian $\mathcal{L}'/m^4$ versus magnetic field intensity $\gamma = H/H_c$ for various anomalous magnetic moments of particles: *1*—$\Delta\mu_1/\mu_0 = 10^{-3}$, *2*—$\Delta\mu_2/\mu_0 = 10^{-3.05}$, and *3*—$\Delta\mu_3/\mu_0 = 10^{-3.1}$.

For $H_c \ll H \ll 16H_c^{*2}/H_c$, the range $1 \ll x \ll 16H_c^{*2}/H_c^2$ is important in integral (26). In this case, the hyperbolic functions may be substituted with exponentials and the integrand in (26) becomes

$$\frac{\exp[f_1(\alpha,\gamma,x)]}{x^2} - \frac{\exp[f_2(a,\gamma,x)]}{x^3} - \frac{\exp[f_2(a,\gamma,x)]}{3x}\left(1 + 3a + \frac{3}{2}a^2\right), \quad (27)$$

where

$$f_1(a,\gamma,x) = -\frac{1}{4\gamma}(2-a\gamma)^2 x,$$

$$f_2(a,\gamma,x) = -\frac{1}{\gamma}\left(1 + \frac{a^2}{4}\gamma^2\right)x.$$

For $H_c \ll H < 4H_c^*$ $(1 \ll \gamma < 2/a)$, $f_1 \sim f_2 = -x/\gamma$ and we find from (26) with logarithmic accuracy that

$$\mathcal{L}' = \frac{m^4\gamma^2}{24\pi^2}\left(1 + 3a + \frac{3}{2}a^2\right)\ln\gamma. \quad (28)$$

For $a \longrightarrow 0$, this formula matches the Heisenberg–Euler Lagrangian in the limit of strong magnetic fields, $H \gg H_c$ [5].

If $4H_c^* < H \ll 16H_c^{*2}/H_c$ or $2/a < \gamma \ll 4/a^2$, then $f_1 \sim f_2 = -\gamma a^2 x/4$. The range $1 \ll x \ll 4/(a^2\gamma)$ gives the

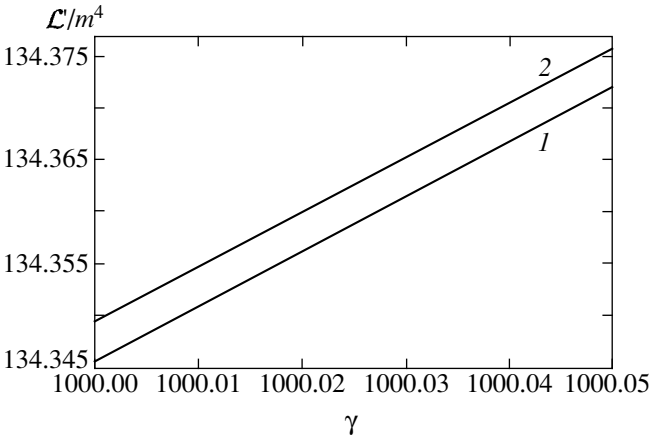

Fig. 2. Lagrangian versus magnetic field intensity $\gamma = H/H_c$ with (curve *1*) and without (curve *2*) an allowance for the anomalous magnetic moment of an electron.

largest contribution to the integral. In this case, we find from (26) that

$$\mathcal{L}' = \frac{m^4\gamma^2}{24\pi^2}\left(1 + 3a + \frac{3}{2}a^2\right)\left(2\ln\frac{2}{a} - \ln\gamma\right). \qquad (29)$$

For $H \gg 16H_c^{*2}/H_c$, we return to the cases considered above (see (23)) where the range $x \ll 1$ gives the largest contribution to integral (26).

If $a\gamma = 2$, i.e., for $H = 4H_c^*$, the exponent (f_1) in one of the exponentials in (27) becomes zero. It is easy to verify that this term is attributable to the contribution from the ground energy state ε_0 (see formula (7)) in which the dependence on particle mass completely drops out at this field intensity. There is no such state with a "dropping" mass in the structure of the Heisenberg–Euler Lagrangian for a fixed field. However, if we consider the passage to the limit $m^2 \longrightarrow 0$, then the Heisenberg–Euler Lagrangian can simulate such an effect. It is easy to see that the total contribution of the ground state is small compared to the contribution of the last term in (27), which owes its origin to the field renormalization in expression (11). A similar conclusion can also be reached by considering integral (16).

The following should be emphasized when commenting on the analogy between the limits $m^2 \longrightarrow 0$ in the Heisenberg–Euler Lagrangian (see formula (16)) and $H \longrightarrow 4H_c^*$ in (26). As we showed above, for $H = 4H_c^*$ in the modified Lagrangian, just as in the Heisenberg–Euler Lagrangian for $m^2 = 0$, the exponent in the terms whose contributions are vanishingly small against the background of the contributions from the renormalization procedure becomes zero. In other words, in both cases, the ground states in the structure of the integrand are equally preferential, but their contribution to the integral is not dominant.

Neglecting the first and the second terms in (27), we find from (26) that

$$\mathcal{L}' = \frac{m^4}{6\pi^2 a^2}\left(1 + 3a + \frac{3a^2}{2}\right)\ln\frac{2}{a}.$$

This result agrees with formulas (28) and (29) from which it can be obtained by the substitution $\gamma = 2/a$. Thus, these functions are continuously joined at $H = 4H_c^*$.

Finally, let us estimate the Lagrangian $\mathcal{L}'$ in terms of traditional QED, i.e., by taking into account the nonzero anomalous magnetic moments of particles in intense electromagnetic fields attributable to radiative effects. Substituting $\Delta\mu$ from (1) into the expression $b_1 = 1 + (\Delta\mu H/m)^2$ yields an estimate of $\mathcal{L}'$ (26) in the limit of extremely strong fields. For a constant magnetic field, $\gamma \gg \mu_0/\Delta\mu$,

$$b_1 \sim \alpha_2 \ln^2(2\gamma),$$

where

$$\alpha_2 = \alpha^2/16\pi^2.$$

In this case, the exponent in (26) is

$$f_3 = -x\frac{b_1}{\gamma} = -\frac{\alpha_2 \ln^2(2\gamma)}{\gamma}x.$$

If $\gamma \gg \alpha_2\ln^2(2\gamma)$, then the range $1 \ll x \ll \gamma/\alpha_2\ln^2(2\gamma)$ is important in (26). Thus, we can find from (26) that

$$\mathcal{L}' = \frac{m^4\gamma^2}{24\pi^2}[\ln\gamma - \ln\alpha_2 - 2\ln(\ln 2\gamma)]. \qquad (30)$$

The first term in (30) is identical to its estimate in the Heisenberg–Euler theory [5]. The relative effective Lagrangian $\mathcal{L}'/m^4$ derived from the integral representation (26) (with $\alpha_2 \approx 3.4 \times 10^{-7}$) is plotted against the magnetic field intensity in Fig. 2 (curve *1*). For comparison, the same figure also shows a plot for $\Delta\mu = 0$ (curve *2*).

Note that allowance for the anomalous magnetic moments of vacuum particles in terms of universally accepted QED leads to a decrease in the radiative correction to the field energy density. Recall that we reached a similar conclusion by considering the static anomalous magnetic moment that arises, in particular, in the modified field theory. Thus, irrespective of the nature of the anomalous magnetic moment attributable to the dynamic or static types of interaction, we obtain consistent results. Our conclusions are also important in studying the anomalous magnetic moment as the

most accurately calculable and measurable (in numerous precision experiments) characteristic of particles.

4. CONCLUSIONS

Our results can be of considerable importance in constructing astrophysical models, in particular, in studying extremely magnetized neutron stars—magnetars; interest in the existence of the latter objects has increased appreciably in recent years (see, e.g., [4] and references therein). According to models for the macroscopic magnetization of bodies composed mostly of neutrons, the intensity of the magnetic fields frozen into them increases from the surface to the central regions and can reach 10^{15}–10^{17} G [19].

Note also that the radiative effects can be enhanced by external intense electromagnetic fields not only in Abelian, but also in non-Abelian quantum field theories. For example, allowance for the influence of an external field on such parameters as the lepton mass and magnetic moment in terms of the standard model leads to nontrivial results. In this case, apart from the electrodynamic contribution, the one-loop mass operator of a charged lepton also contains the contributions from the interaction of $W^{\pm}$-, Z^0-, and H bosons with a vacuum. It is easy to see that, in the absence of an external field, the contribution from weak interactions to the radiative shift of the lepton mass m is suppressed by a factor of $(m/M_i)^2$ ($i = W, Z, H$) compared to the electrodynamic contribution. However, the contributions of weak currents in the ultrarelativistic limit can dominate in intense external fields, as was first noted in [20] (see also [21]).

In close analogy with the quantum corrections to the particle masses, the anomalous magnetic moments of charged leptons in the standard model are attributable to the vacuum radiative effects of electromagnetic and weak interactions and contain the contribution from the hadron polarization of the vacuum. For example, for the anomalous magnetic moment of a muon,

$$a_\mu^{SM} = a_\mu^{QED} + a_\mu^{\text{weak}} + a_\mu^{\text{had}}.$$

According to recent theoretical estimates made in the standard model [22], the contributions from electromagnetic and weak interactions can be written as

$$a_\mu^{QED} = 11\,658\,470.57(0.29)\times 10^{-10},$$

$$a_\mu^{\text{weak}} = 15.1(0.4)\times 10^{-10}.$$

Although the calculations of the contributions from the hadron polarization of a vacuum to a_μ^{SM} have a history that spans almost forty years, a_μ^{had} is currently understood with the largest uncertainty (see, e.g., [23–28]). One of the most reliable estimates for the contributions of the lowest-order hadron polarization of a vacuum that generalizes the data on hadron τ-decay and e^+e^- annihilation appears as follows [23, 24]:[6]

$$a_\mu^{\text{had}} = 692(6)\times 10^{-10}.$$

The theoretical anomalous magnetic moment of a muon in the standard model takes the form [28]

$$a_\mu^{SM} = 11\,659\,177(7)\times 10^{-10}.$$

The results of one of the most recent ($g-2$) experiments aimed at measuring the anomalous magnetic moments of positive polarized muons carried out on a storage ring with superconducting magnets at Brookhaven National Laboratory (BNL) can be represented as

$$a_\mu^{\text{exp}} = 11659204(7)(5)\times 10^{-10}. \quad (31)$$

(Both statistical and systematic errors are included here.) The data obtained yield the difference

$$\Delta_\mu = a_\mu^{\text{exp}} - a_\mu^{SM} = 27\times 10^{-10}, \quad (32)$$

which exceeds the total measurement errors and the uncertainties of the theoretical estimates. According to the most recent reports from the BNL muon ($g-2$) collaboration [29], the relative value of this excess is 2.6. A twofold increase in this accuracy is expected in the immediate future. Clearly, the solution of the muon ($g-2$) problem may lead to the appearance of a new theory beyond the scope of the standard theory.

Recall in this connection that the anomalous magnetic moment of a muon in the modified theory contains a contribution attributable to the new universal parameter M from the outset. According to (4),

$$a_\mu(M) = \frac{m_\mu^2}{2M^2}, \quad (33)$$

where m_μ is the muon mass. It is easy to see that $a_\mu(M)$ is equal in order of magnitude to (32) at $M \sim 1$ TeV.

The principal conclusion drawn from a comparison of the above estimates is that we cannot rule out the possibility that the observed difference between the theoretical and experimental values for Δ_μ is equal to $a_\mu(M)$. As was pointed out above, the parameter M in the new theory may be related to the Higgs boson mass M_H. In this case, the difference between a_μ^{exp} and a_μ^{SM} can provide valuable information about a particle

[6] See, however, [25], where the contribution of the highest orders of hadron polarization of a vacuum were calculated, and the recent papers [26, 27], in which the contribution from the third-order diagrams to a_μ^{had} attributable to photon–photon scattering was taken into account.

whose mass has not been determined in the standard model. Substituting m_μ and the anomalous magnetic moment of a muon into (33), we can easily impose the following constraints on the H-boson mass:

$$1.2\ \text{TeV} \le M_H \le 1.8\ \text{TeV}.$$

The standard model with the Higgs boson mass $M_H \ge 1$ TeV entails several additional features, in particular, the impossibility of describing the weak interactions in the sector of H-, W-, and Z particles in terms of perturbation theory [30]. Naturally, the need for constructing a new nonperturbative theory arises in this case. Apart from the condition for the mass spectrum being limited, $m \le M_H$ (see (2)), the Higgs mechanisms of mass formation and compensation for the discrepancies can become integral elements of one of the most promising versions of the modified theory—the standard model with fundamental mass.

ACKNOWLEDGMENTS

I am grateful to V.G. Kadyshevsky for helpful discussions and valuable remarks and to V.R. Khalilov and A.E. Lobanov for attention to the work and fruitful discussions. This study was supported by the Russian Foundation for Basic Research (project no. 02-02-16784).

REFERENCES

1. W. Heisenberg and H. Euler, Z. Phys. **98**, 714 (1936).
2. V. I. Ritus, Zh. Éksp. Teor. Fiz. **69**, 1517 (1975) [Sov. Phys. JETP **42**, 774 (1975)]; Zh. Éksp. Teor. Fiz. **73**, 807 (1977) [Sov. Phys. JETP **46**, 423 (1977)].
3. V. I. Ritus, hep-th/9812124.
4. V. R. Khalilov, Teor. Mat. Fiz. **133**, 103 (2002).
5. V. B. Berestetskiĭ, E. M. Lifshitz, and L. P. Pitaevskiĭ, *Quantum Electrodynamics*, 3rd ed. (Nauka, Moscow, 1989; Pergamon Press, Oxford, 1982).
6. A. B. Migdal, Zh. Éksp. Teor. Fiz. **62**, 1621 (1972) [Sov. Phys. JETP **35**, 845 (1972)].
7. I. M. Ternov, V. R. Khalilov, and V. N. Rodionov, *Interaction of Charged Particles with Strong Electromagnetic Field* (Mosk. Gos. Univ., Moscow, 1982).
8. V. N. Rodionov, Zh. Éksp. Teor. Fiz. **111**, 3 (1997) [JETP **84**, 1 (1997)].
9. V. N. Rodionov, Zh. Éksp. Teor. Fiz. **113**, 21 (1998) [JETP **86**, 11 (1998)].
10. V. N. Baĭer, V. M. Katkov, and V. M. Strakhovenko, Yad. Fiz. **24**, 379 (1976) [Sov. J. Nucl. Phys. **24**, 197 (1976)].
11. V. I. Ritus, Zh. Éksp. Teor. Fiz. **57**, 2176 (1969) [Sov. Phys. JETP **30**, 1181 (1969)].
12. V. G. Kadyshevskiĭ, Fiz. Élem. Chastits At. Yadra **11**, 5 (1980) [Sov. J. Part. Nucl. **11**, 1 (1980)].
13. V. G. Kadyshevskiĭ, Fiz. Élem. Chastits At. Yadra **29**, 563 (1998) [Phys. Part. Nucl. **29**, 227 (1998)].
14. G. E. Stedman *et al.*, Phys. Rev. A **51**, 4944 (1995).
15. V. I. Denisov and I. P. Denisova, Teor. Mat. Fiz. **129**, 131 (2001); Dokl. Akad. Nauk **378**, 4 (2001) [Dokl. Phys. **46**, 377 (2001)].
16. V. G. Kadyshevskiĭ and V. N. Rodionov, in *Proceedings of Seminar on Symmetry and Integrated Systems*, Ed. by A. N. Sisakyan (Ob. Inst. Yad. Issled., Dubna, 1999), p. 103.
17. V. G. Kadyshevskiĭ and V. N. Rodionov, Teor. Mat. Fiz. **125**, 432 (2000).
18. I. M. Ternov, V. G. Bagrov, and V. Ch. Zhukovskiĭ, Vestn. Mosk. Univ., Ser. 3: Fiz., Astron., No. 1, 30 (1966).
19. R. C. Duncan, astro-ph/0002442.
20. I. M. Ternov, V. N. Rodionov, and A. I. Studenikin, in *Abstracts of All-Union Meeting on Quantum Metrology and Fundamental Physical Constants* (Leningrad, 1982), p. 47.
21. I. M. Ternov, V. N. Rodionov, and A. I. Studenikin, Yad. Fiz. **37**, 1270 (1983) [Sov. J. Nucl. Phys. **37**, 755 (1983)].
22. A. Czarnecki and W. Marciano, Phys. Rev. D **64**, 013014 (2001).
23. M. Davier and A. Hocker, Phys. Lett. B **435**, 427 (1998).
24. M. Davier, S. Eidelman, A. Hocker, and Z. Zhang, hep-ph/0208177.
25. B. Krause, Phys. Lett. B **390**, 392 (1997).
26. E. Bartos, A. Z. Dubnichkova, S. Dubnichka, *et al.*, Nucl. Phys. B **632**, 330 (2002).
27. E. Bartos, S. Dubnichka, A. Z. Dubnichkova, *et al.*, hep-ph/0305051.
28. C. S. Ozben, G. W. Bennett, B. Bousquet, *et al.*, hep-ex/0211044.
29. G. W. Bennett *et al.* (Muon g-2 Collaboration), Phys. Rev. Lett. **89**, 101804 (2002).
30. L. B. Okun', *Leptons and Quarks*, 2nd ed. (Nauka, Moscow, 1981; North-Holland, Amsterdam, 1984).

Translated by V. Astakhov

Journal of Experimental and Theoretical Physics, Vol. 98, No. 3, 2004, pp. 403–416.
From Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 125, No. 3, 2004, pp. 462–475.

NUCLEI, PARTICLES, AND THEIR INTERACTION

QED Correction to Asymmetry for Polarized *ep* Scattering from the Method of Electron Structure Functions¶

A. V. Afanasev[a,b], I. Akushevich[a,b], and N. P. Merenkov[c]

[a]*North Carolina Central University, Durham, NC 27707, USA*
[b]*TJNAF, Newport News, VA 23606, USA*
[c]*National Scientific Center Kharkov Institute of Physics and Technology, Kharkov, 61108 Ukraine*
e-mail: merenkov@kipt.kharkov.ua

Received August 18, 2003

Abstract—The electron structure function method is applied to calculate model-independent QED radiative corrections to the asymmetry of electron–proton scattering. Representations for both spin-independent and spin-dependent parts of the cross section are derived. Master formulas include the leading corrections in all orders and the main contribution of the second order and provide accuracy of the QED corrections at the level of one per mill. Numerical calculations illustrate our analytic results for both elastic and deep inelastic events.

1. INTRODUCTION

Precise polarization measurements in both inclusive [1, 2] and elastic [3, 4] scattering are crucial for understanding the structure and fundamental properties of a nucleon.

One important component of the precise data analysis is radiative effects, which always accompany the processes of electron scattering. The first calculation of radiative corrections to polarized deep inelastic scattering was done by Kukhto and Shumeiko [5], who applied the covariant method of extraction of the infrared divergence [6, 7] to this process. The polarization states were described by 4-vectors, which were kept in their general forms during the calculation. This required a tedious procedure of tensor integration over photonic phase space and, as a result, led to a very complicated structure of the final formulas for the radiative corrections. The next step was taken in [8], where additional covariant expansion of polarization 4-vectors over a certain basis allowed one to simplify the calculation and final results. It resulted in producing the Fortran code POLRAD [9] and Monte Carlo generator RADGEN [10]. These tools are widely used in all current experiments in polarized deep inelastic. Later, the calculation was applied to collider experiments on deep inelastic scattering [11, 12]. We also applied this method to the elastic processes in [13, 14].

However, the method of covariant extraction of the infrared divergence is essentially restricted by the lowest-order radiative corrections. All attempts to go beyond the lowest order lead to very unwieldy formulas, which are difficult to cross check, or to a simple leading log approach [15]. The recent developments are reviewed in [16].

The resolution can be found in applying the formalism of electron structure functions (ESFs). Within this approach, such processes as the electron–positron annihilation into hadrons and the deep inelastic electron–proton scattering in the one-photon exchange approximation can be considered as the Drell–Yan process [17] in the annihilation or scattering channel, respectively. Therefore, the QED radiative corrections to the corresponding cross sections can be written as a contraction of two electron structure functions and the hard part of the cross section (see [18, 19]). Traditionally, these radiative corrections include effects caused by loop corrections and soft and hard collinear radiation of photons and e^+e^- pairs. But it was shown in [19] that this method can be improved by also including effects due to radiation of one noncollinear photon. The corresponding procedure results in a modification of the hard part of the cross section, which takes the lowest-order correction into account exactly and allows going beyond the leading approximation. We applied this approach to the recoil proton polarization in elastic electron scattering in [20]. In the present paper, we calculate radiative corrections to polarized deep inelastic and elastic scattering following [20].

Section 2 gives a short introduction to the structure function method. There, we present two known forms of the electron structure functions, iterative and analytical, which resume singular infrared terms in all orders into the exponent. In this section, we also obtain master formulas for observed cross sections. Leading log results are presented in Section 3. These results are valid for both deep inelastic and elastic cases. We also use the iterative form of electron structure functions to

¶This article was submitted by the authors in English.

extract the lowest-order correction, which can provide a crosscheck via comparison with the known results. In Sections 4 and 5, we describe the procedure of generalizing the results to the next-to-leading order in the deep inelastic and elastic cases. Numerical analysis is presented in Section 6. We consider kinematical conditions of current polarization experiments at fixed targets and collider kinematics. Some conclusions are given in Section 7.

2. ELECTRON STRUCTURE FUNCTIONS

A straightforward calculation based on the quasi-real electron method [21] can be used to write the invariant cross section of the deep inelastic scattering process

$$e^-(k_1) + P(p_1) \longrightarrow e^-(k_2) + X(p_x) \tag{1}$$

as

$$\frac{d\sigma(k_1, k_2)}{dQ^2 dy} = \int\limits_{z_{1m}}^{1} dz_1 \int\limits_{z_{2m}}^{1} dz_2 D(z_1, L) \times \frac{1}{z_2^2} D(z_2, L) \frac{d^2\sigma_{\text{hard}}(\tilde{k}_1, \tilde{k}_2)}{d\tilde{Q}^2 d\tilde{y}}, \quad L = \ln\frac{Q^2}{m^2}, \tag{2}$$

where m is the electron mass and

$$Q^2 = -(k_1 - k_2)^2, \quad y = \frac{2p_1(k_1 - k_2)}{V}, \quad V = 2p_1k_1.$$

The reduced variables that define the hard cross section in the integrand are

$$\tilde{k}_1 = z_1 k_1, \quad \tilde{k}_2 = \frac{k_2}{z_2}, \qquad \tilde{Q}^2 = \frac{z_1}{z_2} Q^2, \quad \tilde{y} = 1 - \frac{1-y}{z_1 z_2}. \tag{3}$$

The electron structure function $D(z, L)$ includes contributions due to photon emission and pair production,

$$D = D^\gamma + D_N^{e^+e^-} + D_S^{e^+e^-}, \tag{4}$$

where D^γ is responsible for the photon radiation and $D_N^{e^+e^-}$ and $D_S^{e^+e^-}$ describe pair production in nonsinglet (by single photon mechanism) and singlet (by double photon mechanism) channels, respectively.

The structure functions on the right-hand side of Eq. (4) satisfy the DGLAP equations [22] (see also [18]). The respective functions $D(z_1, L)$ and $D(z_2, L)$ are responsible for radiation of the initial and final electrons.

There exist different representations for the photon contribution to the structure function [18, 23, 24], but we here use the form given in [18] for D^γ, $D_N^{e^+e^-}$, and $D_S^{e^+e^-}$,

$$D^\gamma(z, Q^2) = \frac{1}{2}\beta(1-z)^{\beta/2-1} \times \left[1 + \frac{3}{8}\beta - \frac{\beta^2}{48}\left(\frac{1}{3}L + \pi^2 - \frac{47}{8}\right)\right] - \frac{\beta}{4}(1+z) + \frac{\beta^2}{32}\left[-4(1+z)\ln(1-z) - \frac{1+3z^2}{1-z}\ln z - 5 - z\right], \tag{5}$$

$$\beta = \frac{2\alpha}{\pi}(L-1),$$

$$D_N^{e^+e^-}(z, Q^2) = \frac{\alpha^2}{\pi^2}\left[\frac{1}{12(1-z)}\left(1 - z - \frac{2m}{\varepsilon}\right)^{\beta/2}\left(L_1 - \frac{5}{3}\right)^2 \times \left(1 + z^2 + \frac{\beta}{6}\left(L_1 - \frac{5}{3}\right)\right)\right]\theta\left(1 - z - \frac{2m}{\varepsilon}\right), \tag{6}$$

$$D_S^{e^+e^-} = \frac{\alpha^2}{4\pi^2}L^2\left[\frac{2(1-z^3)}{3z} + \frac{1}{2}(1-z) + (1+z)\ln z\right]\theta\left(1 - z - \frac{2m}{\varepsilon}\right), \tag{7}$$

where ε is the energy of the parent electron and

$$L_1 = L + 2\ln(1-z).$$

We note that the above form of the structure function $D_N^{e^+e^-}$ includes effects due to real pair production only. The correction caused by the virtual pair is included in D^γ. Terms containing a contribution of the order α^2L^3 are canceled in the sum $D^\gamma + D_N^{e^+e^-}$.

Instead of the photon structure function given by Eqs. (5)–(7), one can use their iterative form [23]

$$D^\gamma(z, L) = \delta(1-z) + \sum_{k=1}^{\infty}\frac{1}{k!}\left(\frac{\alpha L}{2\pi}\right)^k P_1(z)^{\otimes k}, \qquad \underbrace{P_1(z) \otimes \ldots \otimes P_1(z)}_{k} = P_1(z)^{\otimes k},$$

$$P_1(z) \otimes P_1(z) = \int\limits_z^1 P_1(t) P_1\left(\frac{z}{t}\right)\frac{dt}{t}, \tag{8}$$

$$P_1(z) = \frac{1+z^2}{1-z}\theta(1-z-\Delta) + \delta(1-z)\left(2\ln\Delta + \frac{3}{2}\right), \qquad \Delta \ll 1.$$

The iterative form (8) of D^γ does not include any effects caused by pair production. The corresponding nonsinglet part of the structure due to real and virtual pair production can be included into the iterative form of $D^\gamma(z, L)$ by replacing $\alpha L/2\pi$ on the right-hand side of Eq. (8) with the effective electromagnetic coupling

$$\frac{\alpha L}{2\pi} \longrightarrow \frac{\alpha_{\text{eff}}}{2\pi} = -\frac{3}{2}\ln\left(1-\frac{\alpha L}{3\pi}\right), \tag{9}$$

which (within the leading accuracy) is the integral of the running electromagnetic constant.

The lower limits of integration with respect to z_1 and z_2 in the master equation (2) can be obtained from the condition for the existence of inelastic hadronic events,

$$(p_1+\tilde{q})^2 > M_{\text{th}}^2, \quad \tilde{q} = \tilde{k}_1 - \tilde{k}_2, \quad M_{\text{th}} = M + m_\pi, \tag{10}$$

where m_π is the pion mass. This constraint can be rewritten in terms of dimensionless variables as

$$\begin{aligned} &z_1 z_2 + y - 1 - xyz_1 \geq z_2 z_{\text{th}}, \\ &x = \frac{Q^2}{2p_1(k_1-k_2)}, \quad z_{\text{th}} = \frac{M_{\text{th}}^2 - M^2}{V}, \end{aligned} \tag{11}$$

which leads to

$$z_{2m} = \frac{1-y+xyz_1}{z_1 - z_{\text{th}}}, \quad z_{1m} = \frac{1+z_{\text{th}}-y}{1-xy}.$$

The squared matrix element of the considered process in the one-photon exchange approximation is proportional to the contraction of the leptonic and hadronic tensors. Representation (2) reflects the properties of the leptonic tensor. Therefore, it has a universal nature (because of the universality of the leptonic tensor) and can be applied to processes with different final hadronic states. In particular, we can use the electron structure function method to compute radiative corrections to the elastic and deep inelastic (inclusive and semi-inclusive) electron–proton scattering cross sections.

On the other hand, straightforward calculations in the first order in α [5, 8, 21] and the recent calculations of the leptonic current tensor in the second order [25–28] for the longitudinally polarized initial electron demonstrate that, in the leading approximation, the spin-dependent and spin-independent parts of this tensor are the same for the nonsinglet channel contribution. The latter corresponds to photon radiation and e^+e^--pair production through the single-photon mechanism. The difference appears in the second order due to the possibility of pair production in the singlet channel by the double-photon mechanism [28]. Therefore, representation (2), being slightly modified, can be used for the calculation of radiative corrections to cross sections of different processes with a longitudinally polarized electron beam.

In our recent paper [20], we applied the electron structure function method to compute radiative corrections to the ratio of the recoil proton polarizations measured at CEBAF by Jefferson Lab Hall A Collaboration [3]. The aim of this high-precision experiment is the measurement of the proton electric form factor G_E. In the present work, we use this method for calculation of the model-independent part of the radiative corrections to the asymmetry in the scattering of longitudinally polarized electrons on polarized protons at the level of per mill accuracy for elastic and deep inelastic hadronic events.

The cross section of the scattering of the longitudinally polarized electron by the proton with given longitudinal ($\parallel$) or transverse ($\perp$) polarizations for both elastic and deep inelastic events can be written as a sum of the spin-independent and spin-dependent parts,

$$\frac{d\sigma(k_1,k_2,S)}{dQ^2dy} = \frac{d\sigma(k_1,k_2)}{dQ^2dy} + \eta\frac{d\sigma^{\parallel,\perp}(k_1,k_2,S)}{dQ^2dy}, \tag{12}$$

where S is the 4-vector of the target proton polarization and η is the product of the electron and proton polarization degrees. Hereafter, we assume $\eta = 1$.

Master equation (2) describes the radiative corrections to the spin-independent part of the cross section on the right-hand side of Eq. (12), and the corresponding equation for the spin-dependent part is given by

$$\begin{aligned} \frac{d\sigma^{\parallel,\perp}(k_1,k_2,S)}{dQ^2dy} &= \int_{z_{1m}}^1 dz_1 \int_{z_{2m}}^1 dz_2 D^{(p)}(z_1,L) \\ &\times \frac{1}{z_2^2} D(z_2,L)\frac{d^2\sigma_{\text{hard}}^{\parallel,\perp}(\tilde{k}_1,\tilde{k}_2,S)}{d\tilde{Q}^2 d\tilde{y}}, \end{aligned} \tag{13}$$

where

$$D^{(p)} = D^\gamma + D_N^{e^+e^-} + D_S^{e^+e^-(p)},$$

and [28]

$$\begin{aligned} D_S^{e^+e^-(p)} &= \frac{\alpha^2}{4\pi^2}L^2\left(\frac{5(1-z)}{2} + (1+z)\ln z\right) \\ &\times \theta\left(1-z-\frac{2m}{\varepsilon}\right) \end{aligned} \tag{14}$$

describes the radiation of the initial polarized electron.

This representation is valid if radiation of collinear particles does not change the polarizations $S^{\parallel}$ and $S^{\perp}$. Such stabilized 4-vectors of the proton polarization can be written as [8]

$$S_{\mu}^{\parallel} = \frac{2M^2 k_{1\mu} - V p_{1\mu}}{MV},$$
$$S_{\mu}^{\perp} = \frac{u p_{1\mu} + V k_{2\mu} - [2u\tau + V(1-y)]k_{1\mu}}{\sqrt{-uV^2(1-y) - u^2 M^2}}, \quad (15)$$

where

$$u = -Q^2, \quad \tau = M^2/V.$$

It can be verified that, in the laboratory system, the 4-vector $S^{\parallel}$ has the components $(0, \mathbf{n})$, where the 3-vector $\mathbf{n}$ has the orientation of the initial electron 3-momentum $\mathbf{k}_1$. It can also be verified that $S^{\perp}S^{\parallel} = 0$ and that in the laboratory system,

$$S^{\perp} = (0, \mathbf{n}_{\perp}), \quad \mathbf{n}_{\perp}^2 = 1, \quad \mathbf{n} \cdot \mathbf{n}_{\perp} = 0,$$

where the 3-vector $\mathbf{n}_{\perp}$ belongs to the plane $(\mathbf{k}_1, \mathbf{k}_2)$.

If the longitudinal direction $\mathbf{L}$ is chosen along the 3-momentum $\mathbf{k}_1 - \mathbf{k}_2$ in the laboratory system, which coincides with the direction of the 3-vector $\mathbf{q}$ for nonradiative process, and we choose the transverse direction $\mathbf{T}$ in the plane $(\mathbf{k}_1, \mathbf{k}_2)$, then we have the relations

$$\frac{d\sigma^L}{dQ^2 dy} = \cos\theta \frac{d\sigma^{\parallel}}{dQ^2 dy} + \sin\theta \frac{d\sigma^{\perp}}{dQ^2 dy},$$
$$\frac{d\sigma^T}{dQ^2 dy} = -\sin\theta \frac{d\sigma^{\parallel}}{dQ^2 dy} + \cos\theta \frac{d\sigma^{\perp}}{dQ^2 dy},$$
$$\cos\theta = \frac{y + 2xy\tau}{\sqrt{y^2 + 4xy\tau}}, \quad \sin\theta = -2\sqrt{\frac{xy\tau(1-y-xy\tau)}{y^2 + 4xy\tau}},$$

and the master formula (13) for $d\sigma^{\parallel}$ and $d\sigma^{\perp}$.

The asymmetry in elastic scattering and deep inelastic scattering processes is defined as the ratio

$$A^{\parallel,\perp} = \frac{d\sigma^{\parallel,\perp}(k_1, k_2, S)}{d\sigma(k_1, k_2)}, \quad (16)$$

and therefore, calculating the radiative corrections to the asymmetry requires knowing radiative corrections to both spin-independent and spin-dependent parts of the cross section.

Radiative corrections to the spin-independent part were calculated (within the electron structure function approach) in [19]. In the present work, we compute the radiative corrections to the spin-dependent parts for longitudinal and transverse polarizations of the target proton and longitudinally polarized electron beam. For completeness, we briefly recall the result for the unpolarized case.

3. THE LEADING APPROXIMATION

Within the leading accuracy (with the terms of the order $(\alpha L)^n$ taken into account), the electron structure function can be computed, in principle, in all orders of the perturbation theory. In this approximation, we have to take the Born cross section as a hard part on the right-hand sides of Eqs. (2) and (13).

We express the Born cross section in terms of leptonic and hadronic tensors as

$$\frac{d\sigma}{dQ^2 dy} = \frac{4\pi\alpha^2(Q^2)}{VQ^4} L_{\mu\nu}^B H_{\mu\nu}, \quad (17)$$

where $\alpha(Q^2)$ is the running electromagnetic constant, which accounts for the effects of vacuum polarization, and

$$H_{\mu\nu} = -F_1 \tilde{g}_{\mu\nu} + \frac{F_2}{p_1 q} \tilde{p}_{1\mu} \tilde{p}_{1\nu}$$
$$- i\frac{M\epsilon_{\mu\nu\lambda\rho} q_{\lambda}}{p_1 q}\left[(g_1 + g_2)S_{\rho} - g_2 \frac{Sq}{p_1 q} p_{1\rho}\right],$$
$$L_{\mu\nu}^B = -\frac{Q^2}{2} g_{\mu\nu} + k_{1\mu}k_{2\nu} + k_{1\nu}k_{2\mu} + i\varepsilon_{\mu\nu\lambda\rho} q_{\lambda} k_{1\rho}, \quad (18)$$
$$\tilde{g}_{\mu\nu} = g_{\mu\nu} - \frac{q_{\mu}q_{\nu}}{q^2}, \quad \tilde{p}_{1\mu} = p_{1\mu} - \frac{p_1 q}{q^2} q_{\mu}.$$

In Eq. (18), we assume the proton and electron polarization degrees equal to 1. The spin-independent (F_1, F_2) and spin-dependent (g_1, g_2) proton structure functions depend on the two variables

$$x' = \frac{-q^2}{2p_1 q}, \quad q^2 = (p_x - p_1)^2.$$

In the Born approximation, $x' = x$, but these variables differ in the general case, when radiation of photons and electron–positron pairs is allowed.

Because the normalization is chosen, the elastic limit ($p_x^2 = M^2$) can be obtained by simply substituting

$$F_1(x', q^2) \longrightarrow \frac{1}{2}\delta(1-x')G_M^2(q^2),$$

$$F_2(x', q^2) \longrightarrow \delta(1-x')\frac{G_E^2(q^2) + \lambda G_M^2(q^2)}{1+\lambda},$$

$$g_1(x', q^2) \longrightarrow \frac{1}{2}\delta(1-x')\Bigg\{G_M(q^2)G_E(q^2) + \frac{\lambda}{1+\lambda}[G_M(q^2) - G_E(q^2)]G_M(q^2)\Bigg\}, \quad (19)$$

$$g_2(x, q^2) \longrightarrow -\frac{1}{2}\delta(1-x')\frac{\lambda}{1+\lambda} \times [G_M(q^2) - G_E(q^2)]G_M(q^2),$$

$$\lambda = -\frac{q^2}{4M^2}$$

in the hadronic tensor, where G_M and G_E are the magnetic and electric proton form factors.

A simple calculation gives the spin-independent and spin-dependent parts of the well-known Born cross section in the form

$$\frac{d\sigma^B}{dQ^2dy} = \frac{4\pi\alpha^2(Q^2)}{Q^4 y} \times [(1-y-xy\tau)F_2(x, Q^2) + xy^2F_1(x, Q^2)], \quad (20)$$

$$\frac{d\sigma_\parallel^B}{dQ^2dy} = \frac{8\pi\alpha^2(Q^2)}{V^2 y} \times \left[\left(\tau - \frac{2-y}{2xy}\right)g_1(x, Q^2) + \frac{2\tau}{y}g_2(x, Q^2)\right], \quad (21)$$

$$\frac{d\sigma_\perp^B}{dQ^2dy} = -\frac{8\pi\alpha^2(Q^2)}{V^2 y}\sqrt{\frac{M^2}{Q^2}(1-y-xy\tau)} \times \left[g_1(x, Q^2) + \frac{2}{y}g_2(x, Q^2)\right]. \quad (22)$$

Thus, within the leading accuracy, the radiatively corrected cross section of process (1) is defined by Eq. (2) (for its spin-independent part) with (20) as the hard part of the cross section and by Eq. (13) (for its spin-dependent part) with (21) or (22) as the hard part.

It is useful to extract the first-order correction to the Born approximation, as defined by master equation (2). For this purpose, we can use the iterative form of the photon structure function D^γ with $L \longrightarrow L-1$ and

$$\Delta \longrightarrow \Delta_1 = \frac{2(\Delta\varepsilon)}{\sqrt{V}(1-xy)}\sqrt{\tau + z_+},$$

$$z_+ = y(1-x), \quad \frac{2(\Delta\varepsilon)}{\sqrt{V}} \ll 1$$

for $D(z_1, L)$ and

$$\Delta \longrightarrow \Delta_2 = \frac{2(\Delta\varepsilon)}{\sqrt{V}(1-z_+)}\sqrt{\tau + z_+}$$

for $D(z_2, L)$, where $(\Delta\varepsilon)$ is the minimal energy of a hard collinear photon in the special system ($\mathbf{k}_1 - \mathbf{k}_2 + \mathbf{p}_1 = 0$). Straightforward calculations yield the expression

$$\frac{d\sigma^{(1)}(k_1, k_2)}{dQ^2dy} = \frac{\alpha(L-1)}{2\pi} \times \Bigg\{\frac{d\sigma^{(B)}(k_1, k_2)}{dQ^2dy}\left[3 + 2\ln\frac{4(\Delta\varepsilon)^2(z_+ + \tau)}{V(1-z_+)(1-xy)}\right] + \int_{z_{\text{th}}}^{z_+ - \rho} dz\left[\frac{1+z_1^2}{(1-xy)(1-z_1)}\frac{d\sigma^{(B)}(z_1k_1, k_2)}{dQ_t^2dy_t} + \frac{1+z_2^2}{(1-z_+)(1-z_2)}\frac{d\sigma^{(B)}(k_1, k_2/z_2)}{dQ_s^2dy_s}\right]\Bigg\}, \quad (23)$$

where

$$z = \frac{M_x^2 - M^2}{V}, \quad z_1 = \frac{1-y+z}{1-xy},$$

$$z_2 = \frac{1-z_+}{1-z}, \quad \rho = \frac{2(\Delta\varepsilon)}{\sqrt{V}}\sqrt{\tau + z_+},$$

$$Q_t^2 = -q_t^2 = z_1Q^2, \quad Q_s^2 = -q_s^2 = \frac{Q^2}{z_2},$$

$$y_{t,s} = 1 - \frac{1-y}{z_{1,2}}.$$

Similar equations can be derived for the first-order correction to the spin-dependent part of the cross section for both longitudinal and transverse polarizations of the target proton.

4. DEEP INELASTIC SCATTERING CROSS SECTION BEYOND THE LEADING ACCURACY

To go beyond the leading accuracy, we have to improve the expressions for hard parts of the cross sections in master equations (2) and (13) in order to include effects caused by radiation of a hard noncollinear photon. In principle, we can also improve the expression for the D function in order to take collinear next-to-leading effects in the second order of perturbation theory into account. The essential part of these effects is included in our D functions due to the replacement $L \longrightarrow L - 1$. The rest can be written using the results of the corresponding calculations for the double photon emission [27, 30], pair production [28, 31, 32], one-loop corrected Compton tensor [25, 26, 33], and virtual correction [34]. But we restrict ourselves here to the D functions given above in Eqs. (5), (6), (7), and (14).

To compute the improved hard cross section, we must find the full first-order radiative corrections to the cross section of process (1) and subtract from it (to avoid double counting) its leading part defined by Eq. (23) (for the unpolarized case). Therefore, the improved hard part can be written as

$$\frac{d\sigma_{\rm hard}}{dQ^2dy} = \frac{d\sigma^B}{dQ^2dy} + \frac{d\sigma^{(S+V)}}{dQ^2dy} + \frac{d\sigma^H}{dQ^2dy} - \frac{d\sigma^{(1)}}{dQ^2dy}, \quad (24)$$

where $d\sigma^{(S+V)}$ is the correction to the cross section of process (1) due to virtual and soft photon emission and $d\sigma^H$ is the cross section of the radiative process

$$e^-(k_1) + P(p_1) \longrightarrow e^-(k_2) + \gamma(k) + X(p_x). \quad (25)$$

The virtual and soft corrections are factored in a similar way for both polarized and unpolarized cases [19] and can be written as

$$\frac{d\sigma^B}{dQ^2dy} + \frac{d\sigma^{(S+V)}}{dQ^2dy} = \frac{d\sigma^B}{dQ^2dy}\left[1 + \frac{\alpha}{2\pi}\right.$$
$$\left.\times\left(\delta + (L-1)\left(3 + 2\ln\frac{\rho^2}{(1-xy)(1-z_+)}\right)\right)\right],$$
$$\delta = -1 - \frac{\pi^2}{3} - 2f\frac{1-y-xy\tau}{(1-xy)(1-z_+)} - \ln^2\frac{1-xy}{1-z_+}, \quad (26)$$
$$f(x) = \int_0^x \frac{dt}{t}\ln(1-t).$$

To calculate the cross section of radiative process (25), we use the corresponding leptonic tensor in the form

$$L^\gamma_{\mu\nu} = \frac{\alpha}{4\pi^2}(L^{H(un)}_{\mu\nu} + L^H_{\mu\nu})\frac{d^3k}{\omega},$$
$$L^H_{\mu\nu} = 2i\varepsilon_{\mu\nu\lambda\rho}q_\lambda(k_{1\rho}R_t + k_{2\rho}R_s),$$
$$R_t = \frac{u+t}{st} - 2m^2\left(\frac{1}{s^2} + \frac{1}{t^2}\right), \quad (27)$$
$$R_s = \frac{u+s}{st} - 2m^2\frac{s_t}{ut^2}, \quad s_t = \frac{-u(u+Vy-Vz)}{u+V},$$

where ω is the energy of the radiated photon; $L^{H(un)}_{\mu\nu}$ is the leptonic tensor for unpolarized particles (see [33]); and we use the notation

$$s = 2kk_2, \quad t = -2kk_1, \quad q^2 = u + s + t$$

for kinematic invariants. The result for the unpolarized case was derived in [19], and here we rewrite it using standard notation as

$$\frac{d\sigma_{\rm hard}}{dQ^2dy} = \frac{d\sigma^B}{dQ^2dy}\left(1 + \frac{\alpha}{2\pi}\delta\right) + \frac{\alpha}{VQ^2}$$
$$\times\int_{z_{\rm th}}^{z_+} dz\left\{\frac{1-r_1}{1-xy}\hat{P}_tN - \frac{1-r_2}{1-z_+}\hat{P}_sN\right.$$
$$\left. + \int_{r_-}^{r_+} dr\frac{2W}{\sqrt{y^2+4xy\tau}}\right. \quad (28)$$
$$+ P\int_{r_-}^{r_+}\frac{dr}{1-r}\left[\frac{1-\hat{P}_t}{|r-r_1|}\left(\frac{(1+r^2)N}{1-xy} + (r_1-r)T_t\right)\right.$$
$$\left.\left. - \frac{1-\hat{P}_s}{|r-r_2|}\left(\frac{(1+r^2)N}{1-z_+} + (r_2-r)T_s\right)\right]\right\}\frac{\alpha^2(rQ^2)}{r^2},$$

where $r = -q^2/Q^2$ and the limits of integration with respect to r are

$$r_\pm(z) = \frac{1}{2xy(\tau+z_+)}$$
$$\times[2xy(\tau+z) + (z_+-z)(y \pm \sqrt{y^2+4xy\tau})].$$

Here, we used the notation

$$N = 2F_1(x', r) + \frac{2x'}{rxy}\left(\frac{1-y}{xy} - \tau\right)F_2(x', r),$$

$$W = 2F_1(x', r) - \frac{2x'\tau}{rxy}F_2(x', r),$$

$$T_t = -\frac{2x'[1-r(1-y)]}{x^2y^2r}F_2(x', r), \qquad (29)$$

$$T_s = -\frac{2x'(1-y-r)}{x^2y^2r}F_2(x', r),$$

$$r_1 = z_1, \quad r_2 = \frac{1}{z_2}, \quad x' = \frac{xyr}{xyr+z}.$$

The action of the operators $\hat{P}_t$ and $\hat{P}_s$ is defined as

$$\hat{P}_t f(r, x') = f(r_1, x_t), \quad \hat{P}_s f(r, x') = f(r_2, x_s),$$

$$x_t = \frac{xyr_1}{xyr_1+z}, \quad x_s = \frac{xyr_2}{xyr_2+z}.$$

The hard cross section (29) has neither collinear nor infrared singularities. The different terms on the right-hand side of Eq. (29) have singularities at $r = r_1$, $r = r_2$, and $r = 1$. Singularities at the first two points are collinear, and the one at the third point, nonphysical, arising at integration. Collinear singularities vanish because of the action of the operators $\hat{P}_t$ and $\hat{P}_s$ on the terms containing N. The nonphysical singularity cancels because, in the limiting case $r \longrightarrow 1$, we have

$$\frac{r_2-r}{|r_2-r|} = 1, \quad \frac{r_1-r}{|r_1-r|} = -1, \quad T_t + T_s = 0.$$

To derive the hard cross section for the polarized case, we have to use the analogue of Eq. (24) for $d\sigma^{\parallel}$ and $d\sigma^{\perp}$. Taking into account that $d\sigma^{V+S}$ and $d\sigma^{(1)}$ are the same in the polarized and unpolarized cases and using expression (27) for the antisymmetric part of the leptonic tensor to compute $d\sigma^{H}$ in the polarized case, we arrive at

$$\frac{d\sigma^{\parallel,\perp}_{\rm hard}}{dQ^2dy} = \frac{d\sigma^{B}_{\parallel,\perp}}{dQ^2dy}\left(1 + \frac{\alpha}{2\pi}\delta\right) + \frac{\alpha}{Q^4}U^{\parallel,\perp}$$

$$\times\int_{z_{\rm th}}^{z_+} dz\left\{\frac{1-r_1}{1-xy}\hat{P}_t N_t^{\parallel,\perp} + \frac{1-r_2}{1-z_+}\hat{P}_s N_s^{\parallel,\perp}\right.$$

$$+ P\int_{r_-}^{r_+}\frac{dr}{1-r}\left[\frac{1-\hat{P}_s}{|r-r_2|(1-z_+)}\right.$$

$$\times\left((1+r^2)N_s^{\parallel,\perp} + \frac{2(r_2-r)}{r_2}T_s^{\parallel,\perp}\right) - \frac{1-\hat{P}_t}{|r-r_1|}$$

$$\left.\left.\times\left(\frac{(1+r^2)N_t^{\parallel,\perp}}{1-xy} + 2r(r_1-r)T_t^{\parallel,\perp}\right)\right]\right.$$

$$\left.+\int_{r_-}^{r_+} dr\frac{2W^{\parallel,\perp}}{\sqrt{y^2+4xy\tau}}\right\}\frac{x'\alpha^2(Q^2r)}{r^3}, \qquad (30)$$

where

$$U^{\parallel} = 1, \quad U^{\perp} = \sqrt{\frac{M^2}{Q^2}(1-y-xy\tau)^{-1}},$$

$$W^{\parallel} = 4y\tau W, \quad W^{\perp} = 2y^2(1+2x\tau)W,$$

$$W = (1+r)xg_1 + x'g_2,$$

$$N_t^{\parallel} = 2[2r-z-xy(r+2\tau)]g_1 - 8x'\tau g_2,$$

$$N_s^{\parallel} = 2[2-z-xyr(1+2\tau)]g_1 - 8x'\tau g_2,$$

$$N_t^{\perp} = 2[1-y-z+r-xy(r+2\tau)](xyg_1 + 2x'g_2),$$

$$N_s^{\perp} = 2\left[1-y+\frac{1-z}{r}-xy(1+2\tau)\right](xyrg_1 + 2x'g_2),$$

$$T_t^{\parallel} = 2rg_1 - 4x'\tau g_2, \quad T_s^{\parallel} = 2(z-1)(g_1 - 2x'\tau g_2),$$

$$T_t^{\perp} = 2xyrg_1 + 2x'(1-y+r-2xy\tau)g_2,$$

$$T_s^{\perp} = 2(z-1)[xyg_1 + x'(1-y+1/r-2xy\tau)g_2].$$

The polarized hard cross section defined by Eq. (30) is also free from collinear singularities due to the action of the operators $1-\hat{P}_t$ and $1-\hat{P}_s$. The nonphysical singularity at $r = 1$ on the right-hand side of Eq. (30) cancels because

$$T_t^{\parallel,\perp} = \frac{1}{z-1}T_s^{\parallel,\perp}$$

in this limit. We note that radiation of a photon at large angles by the initial and final electrons increases the range of r in (28) and (30), because $r_1 < r < r_2$ for collinear radiation, and now $r_- < r_1$ and $r_+ > r_2$. This may be important if the hadron structure functions are large in these additional regions.

5. HARD CROSS SECTION FOR ELASTIC HADRONIC EVENTS

To describe the hard cross section for elastic hadronic events, we use the replacement defined by (19) in Eqs. (28) and (30). We refer to Eqs. (21)–(23) for the Born cross sections that enter these equations. The function $\delta(1-x')$ is used to integrate with respect to the inelasticity z,

$$\int dz\delta(1-x') = xyr. \tag{31}$$

The final result for the unpolarized case is given by (we do not introduce special notation for the elastic cross section)

$$\frac{d\sigma_{\rm hard}}{dQ^2dy} = \frac{d\sigma^B}{dQ^2dy}\left(1+\frac{\alpha}{2\pi}\delta\right)+\frac{\alpha}{V^2}$$

$$\times\left\{\frac{1-r_1}{1-xy}\hat{P}_tN-\frac{1-r_2}{1-z_+}\hat{P}_sN+\int\limits_{r_-}^{r_+}dr\frac{2W}{\sqrt{y^2+4xy\tau}}\right.$$

$$+P\int\limits_{r_-}^{r_+}\frac{dr}{1-r}\left[\frac{1-\hat{P}_t}{|r-r_1|}\left(\frac{1+r^2}{1-xy}N+(r_1-r)T_t\right)\right.$$

$$\left.\left.-\frac{1-\hat{P}_s}{|r-r_2|}\left(\frac{1+r^2}{1-z_+}N+(r_2-r)T_s\right)\right]\right\}\frac{\alpha^2(Q^2r)}{r}, \tag{32}$$

where

$$N = G_M^2+\frac{2}{xyr}\left(\frac{1-y}{xy}-\tau\right)\frac{G_E^2+\lambda G_M^2}{1+\lambda},$$

$$W = G_M^2 = \frac{2\tau}{xyr}\frac{G_E^2+\lambda G_M^2}{1+\lambda},$$

$$T_t = -\frac{2}{x^2y^2r}[1-r(1-y)]\frac{G_E^2+\lambda G_M^2}{1+\lambda},$$

$$T_s = -\frac{2}{x^2y^2r}(1-r-y)\frac{G_E^2+\lambda G_M^2}{1+\lambda}.$$

The Born cross section on the right-hand side of Eq. (32) is defined as

$$\frac{d\sigma^B}{dQ^2dy} = \frac{4\pi\alpha^2(Q^2)}{V^2}$$

$$\times\left[\frac{1}{2}G_M^2+[1-y(1+\tau)]\frac{G_E^2+\lambda G_M^2}{y^2(1+\lambda)}\right]\delta\left(y-\frac{Q^2}{V}\right). \tag{33}$$

In writing this last equation, we took into account that

$$\delta(1-x) = y\delta\left(y-\frac{Q^2}{V}\right).$$

The spin-dependent hard cross section for elastic hadronic events can be written in a form very similar to (32),

$$\frac{d\sigma_{\rm hard}^{\parallel,\perp}}{dQ^2dy} = \frac{d\sigma_{\parallel,\perp}^B}{dQ^2dy}\left(1+\frac{\alpha}{2\pi}\delta\right)+\frac{\alpha}{V^2}U^{\parallel,\perp}$$

$$\times\left\{\frac{1-r_1}{1-xy}\hat{P}_tN_t^{\parallel,\perp}+\frac{1-r_2}{1-z_+}\hat{P}_sN_s^{\parallel,\perp}\right.$$

$$+\int\limits_{r_-}^{r_+}dr\frac{W^{\parallel,\perp}}{\sqrt{y^2+4xy\tau}}+P\int\limits_{r_-}^{r_+}\frac{dr}{1-r}\left[-\frac{1-\hat{P}_t}{|r-r_1|}\right.$$

$$\times\left(\frac{1+r^2}{1-xy}N_t^{\parallel,\perp}+2r(r_1-r)T_t^{\parallel,\perp}\right)+\frac{1-\hat{P}_s}{|r-r_2|(1-z_+)}$$

$$\left.\left.\times\left((1+r^2)N_s^{\parallel,\perp}+\frac{2(r_2-r)}{r_2}T_s^{\parallel,\perp}\right)\right]\right\}\frac{\alpha^2(Q^2r)}{(4M^2+Q^2r)r^2}, \tag{34}$$

where

$$W^{\parallel} = 4y\tau W,\quad W^{\perp} = 2y^2(1+2x\tau)W,$$

$$W = r[x(1+r)-1]G_M^2+\left[r+\frac{4\tau}{y}(1+r)\right]G_MG_E,$$

$$N_t^{\parallel} = r(2\tau+r)(2-xy)G_M^2+8\tau\left[r\left(\frac{1}{xy}-1\right)-\tau\right]G_MG_E,$$

$$N_s^{\parallel} = r(2\tau+1)(2-xyr)G_M^2+8\tau\left[\frac{1}{xy}-r(1+\tau)\right]G_MG_E,$$

$$N_t^{\perp} = [1-y+r-xy(r+2\tau)]$$

$$\times[-r(2-xy)G_M^2+2(r+2\tau)G_MG_E],$$

$$N_s^{\perp} = \left[1-y+\frac{1}{r}-xy(1+2\tau)\right]$$

$$\times[-r(2-xyr)G_M^2+2r(1+2\tau)G_MG_E],$$

$$T_t^{\parallel} = r\left[(r+2\tau)G_M^2 + 2\tau\left(\frac{2}{xy}-1\right)G_M G_E\right],$$

$$T_s^{\parallel} = -r(1+2\tau)G_M^2 - 2\tau\left(\frac{2}{xy}-r\right)G_M G_E,$$

$$T_t^{\perp} = r\{-[r(1-xy)+1-y-2xy\tau]G_M^2 + [1-y-2xy\tau+r+4\tau]G_M G_E\},$$

$$T_s^{\perp} = r\left[\frac{1}{r}-xy(1+2\tau)+1-y\right]G_M^2 - [2\tau-(2-xyr)+1+r(1-y)]G_M G_E.$$

We note that the argument of the electromagnetic form factors in Eqs. (32) and (34) is $-Q^2r$.

The Born cross sections on the right-hand side of Eq. (34) are given by

$$\frac{d\sigma_{\parallel}^B}{dQ^2dy} = \frac{4\pi\alpha^2(Q^2)}{V(4M^2+Q^2)}\left[4\tau\left(1+\tau-\frac{1}{y}\right)G_M G_E - (1+2\tau)\left(1-\frac{y}{2}\right)G_M^2\right]\delta\left(y-\frac{Q^2}{V}\right), \quad (35)$$

for the longitudinal polarization of the target proton and by

$$\frac{d\sigma_{\perp}^B}{dQ^2dy} = \frac{8\pi\alpha^2(Q^2)}{V(4M^2+Q^2)}\sqrt{\frac{M^2}{Q^2}[1-y(1+\tau)]} \times\left[\left(1-\frac{y}{2}\right)G_M^2 - (1+2\tau)G_M G_E\right]\delta\left(y-\frac{Q^2}{V}\right), \quad (36)$$

for the transverse one. The argument of the form factors in (35) and (36) is $-Q^2$.

The results in this section can be generalized to elastic electron–deuteron scattering in both polarized and unpolarized cases in a very simple way, because the respective deuteron tensors $H_{\mu\nu}^d$ are connected with the proton ones $H_{\mu\nu}^p$ by the relations

$$H_{\mu\nu}^{d(un)} = \frac{4\tau+xyr}{4\tau}H_{\mu\nu}^{p(un)} \times\left(G_M^2 \longrightarrow \frac{2}{3}G_M^{(d)^2}, G_E^2 \longrightarrow G_C^2 + \frac{8x^2y^2r^2}{9(4\tau)^2}G_Q^2\right),$$

$$H_{\mu\nu}^{d(\parallel,\perp)} = -\frac{4\tau+xyr}{8\tau}H_{\mu\nu}^{p(\parallel,\perp)} \times\left(G_M \longrightarrow G_M^{(d)}, G_E \longrightarrow 2G_C + \frac{xyr}{6\tau}G_Q\right),$$

where $G_M^{(d)}$, G_C, and G_Q are the magnetic, charged, and quadrupole deuteron form factors, respectively.

6. NUMERICAL ESTIMATIONS

The formulas obtained in the last section include some operators that emphasize the physical meaning of the transformations performed, but they are not convenient in numerical analysis. Here, we present a unified version of the formulas without any operators. For example, the symbol P is explicitly treated as

$$P\int_{r_-}^{r_+}\frac{dr}{1-r}F(r) = \int_{r_-}^{r_+}\frac{dr}{1-r}(F(r)-F(1)) + F(1)\ln\frac{1-r_-}{r_+-1}.$$

Therefore, all cross sections given by Eqs. (28), (30), (32), and (34) can be written by means of the unified formula

$$\frac{d\sigma_{\text{hard}}^i}{dQ^2dy} = \frac{d\sigma_i^B}{dQ^2dy}\left(1+\frac{\alpha}{2\pi}\delta\right) + \alpha U_i \times\int_{z_{\text{th}}}^{z_+}dz\left(\{L_1^i N_i(r_1) + L_2^i N_i(r_2)\} + \int_{r_-}^{r_+}dr\left\{W_i + T_i + \frac{1}{1-r}\left[N_i(r_1)-N_i(r_2) + \frac{1-r_1}{|r-r_1|}[N_i(r)-N_i(r_1)] + \frac{1-r_2}{|r-r_2|}[N_i(r)-N_i(r_2)]\right]\right\}\right), \quad (37)$$

where

$$L_{1,2}^i = \mp b_i\frac{(1-r_{1,2})^2}{1+r_{1,2}^2} \mp \ln\frac{1-r_-}{r_+-1},$$

$$b_u = -1, \quad b_{l,t} = 1.$$

The index i runs over all polarization states ($i = u, l, t$). The functions $N_i(r)$ and T_i are given by

$$N_i(r) = \frac{1+r^2}{z_+-z}N_i\frac{x'\alpha^2}{r^3},$$

$$T_i = \begin{cases}\pm\dfrac{T_{i1}}{1-r}\dfrac{x'\alpha^2}{r^3}, & r<r_1, \quad r>r_2\\[2ex] T_{i2}\dfrac{x'\alpha^2}{r^3}, & r_1<r<r_2.\end{cases}$$

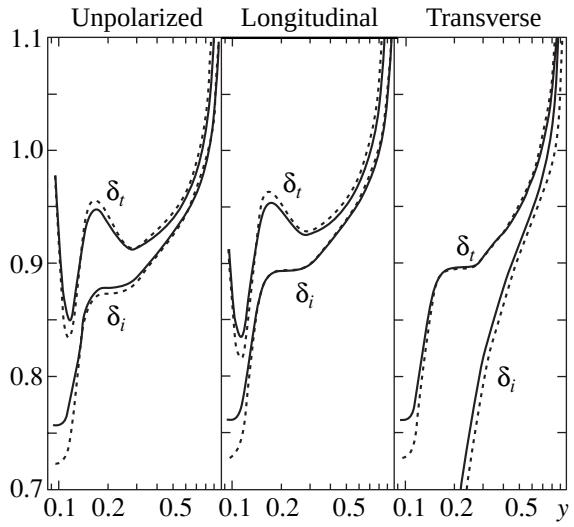

Fig. 1. Radiative correction to unpolarized and polarized (both longitudinal and transverse) parts of the cross section for kinematics close to JLab experiments, $V = 10$ GeV^2, $x = 0.5$.

The pole at $r = 1$ can be reached only in the region $r_1 < r < r_2$, and hence there is no singularity in the terms involving T_{i1}. For T_{i2}, this pole cancels explicitly:

$$T_{u2} = \frac{2(2-y)F_2}{x^2y^2}, \quad T_{l2} = -4(1+r)g_1 + 8x'\tau g_2,$$

$$T_{l2} = -4(1+r)xyg_1 - 4x'\left(r + \frac{1}{r} + 2 - y - 2xy\tau\right)g_2.$$

In the unpolarized case, $N_u = rN/x'$ with N from (28). In other cases, they are

$$N_l = 2\left[-1 - r + \frac{y(1+2x\tau)[1-z+r(1-xy)]}{2-y}\right]g_1 + 8x'\tau g_2,$$

$$N_t = -\frac{4[1-z+r(1-xy)]}{r(2-y)}[xyr(1-y-xy\tau)g_1 + x'(1-y+z+r(1-y+xy))g_2] + 4x'y(1+2x\tau)g_2,$$

$$T_{u1} = -\frac{2(1+r)F_2}{x^2y},$$

$$T_{l1} = 4\frac{y(1+r^2)(1+2x\tau)}{2-y}g_1 + 8x'(1+r)\tau g_2,$$

$$T_{t1} = 4\left\{\frac{1+r^2}{2-y}\left[-2xy(1-y-xy\tau)g_1 + (y-2z+yr(1-2x))\frac{x'}{r}g_2\right] + x'y(1+2x\tau)(1+r)g_2\right\},$$

and

$$W_u = \frac{2W}{\sqrt{y^2+4xy\tau}}\frac{\alpha^2}{r^2}, \quad W_{l,t} = \frac{2W^{\|,\perp}}{\sqrt{y^2+4xy\tau}}\frac{x\alpha^2}{r^3},$$

$$U_u = \frac{1}{VQ^2} \quad U_{l,t} = \frac{\tilde{U}^{\|,\perp}}{Q^4}.$$

The same formulas can be used in the elastic case. Only Eqs. (19) and (31) are needed here. In the elastic case, we must therefore substitute

$$\int dz \longrightarrow xyr,$$

set $x' = 1$ and $z = 0$, and replace the proton structure functions in accordance with (19).

It is believed that the formulas obtained within the presented formalism are not convenient for numerical analysis. There are two reasons for such an opinion. First, the electron structure function in form (5), (6) has a very sharp peak as z tends to unity. Second, because absolute values appear in denominators, the integrand cannot be a continuous function of the integration variables. This produces obstacles for numerical analysis if it is carried out in the traditional style based on adaptive methods of numerical integration, which is used in such programs as TERAD/HECTOR [36] or POLRAD [9]. But it is possible to perform numerical analysis if Monte Carlo integration is used instead of adaptive integration and the regions with sharp peaks are extracted into separate integration subregions. Based on these ideas, we developed the Fortran code ESFRAD,[1] which allows one to perform the numerical analysis without any serious difficulties.

We considered two radiative processes. In the first case, the continuum of hadrons is produced, and in the second case, the proton remains in the ground state. Both of the effects considered contribute to the experimentally observed cross section[2] of deep inelastic scattering. They are usually called the radiative tails from the continuous spectrum and the elastic peak, or simply the inelastic and elastic radiative tails. Below, we study the contributions of the tails numerically within kine-

[1] Electron Structure Function method for RADiative corrections.

[2] Here and below, we mean double differential cross section $\sigma = d\sigma/dydQ^2$.

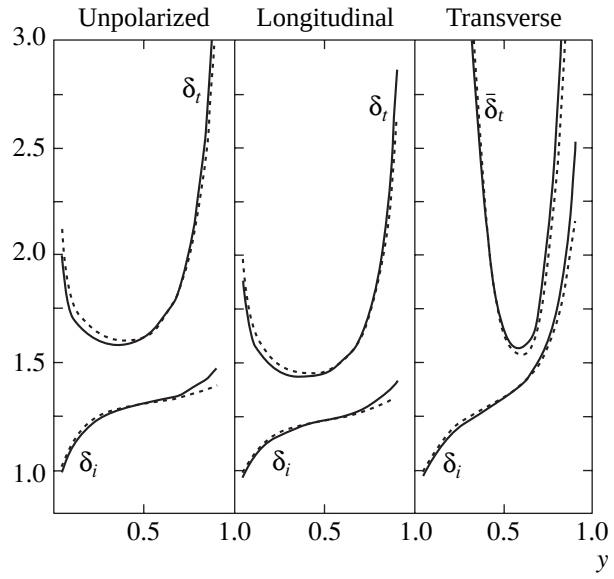

Fig. 2. Radiative correction to unpolarized and polarized (both longitudinal and transverse) parts of the cross section for kinematics close to HERMES experiments, $V = 50\ \mathrm{GeV}^2$, $x = 0.1$, $\bar{\delta}_t = -\delta_t$.

matical conditions of the current experiments on deep inelastic scattering.

We take three typical values of V equal to 10, 50, and $10^5\ \mathrm{GeV}^2$. They correspond to JLab, HERMES, and HERA measurements. Figures 1–3 give the radiative correction factor for all polarization states (unpolarized, longitudinal, or transverse)

$$\delta_{i,t} = \frac{\sigma^{\mathrm{obs}}}{\sigma^B}. \tag{38}$$

The observed double differential cross section is given by master formulas (2) and (13), and the Born cross section is calculated via (20), (21), and (22). Both elastic and inelastic contributions must be taken for σ_{hard}. In this case, we obtain the total radiative corrections factor (δ_t). The subscripts i and t correspond to the cases where the elastic radiative tail is included in the total correction (δ_t) or the inelastic radiative tail contributes only (δ_i). The elastic radiative tail may optionally not be included, because there sometimes exist experimental methods to separate this contribution. We note that, for the HERA kinematics, we do not include it because it is usually separated experimentally. Also, we can extract a one-loop contribution in order to study the effect of higher-order corrections. The observed cross section in this case is given by the sum of the cross sections in Eqs. (23) and (37). We note that this can provide an additional cross check by comparison with POLRAD.

We use rather simple models for spin-averaged and spin-dependent structure functions. It allows us not to mix the pure radiative effects, which are of interest, with the effects due to hadron structure functions. Specifically, we use the so-called D8 model for the spin-average structure function [35] (see also discussion

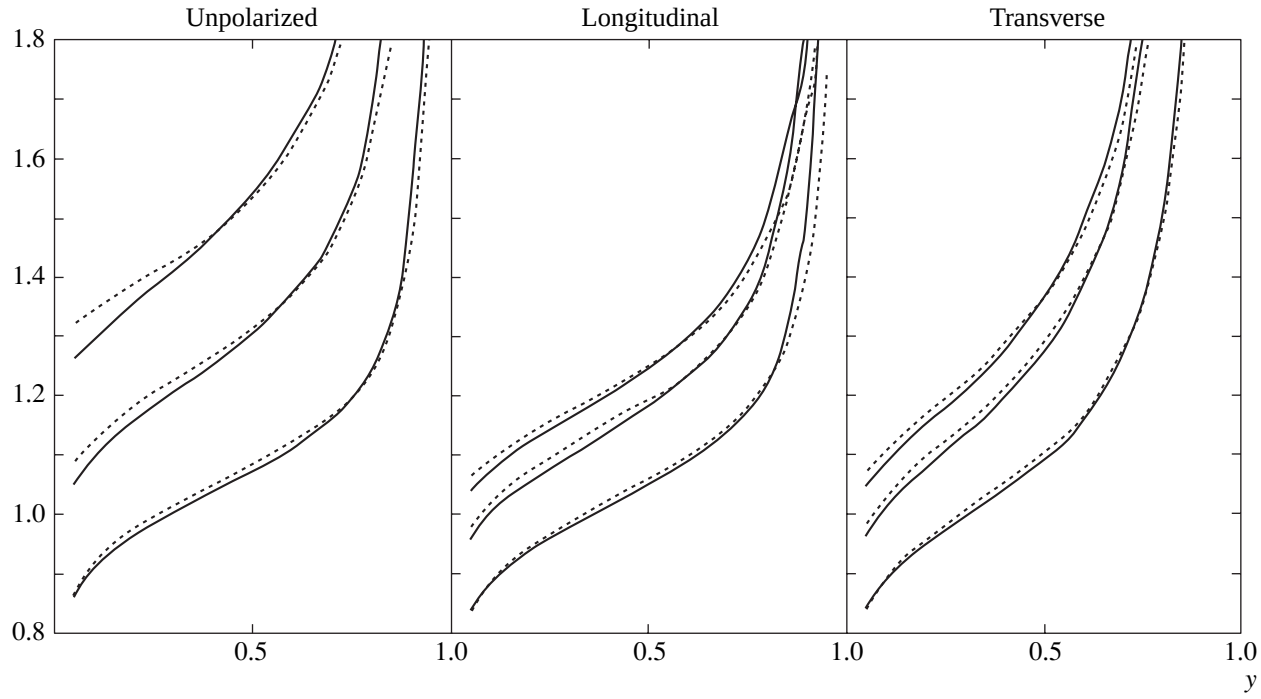

Fig. 3. One-loop and total radiative corrections (dashed and solid lines) for collider kinematics (HERA); $V = 10^5\ \mathrm{GeV}^2$. Lines from top to bottom correspond to different values of $x = 0.001$, 0.01, and 0.1.

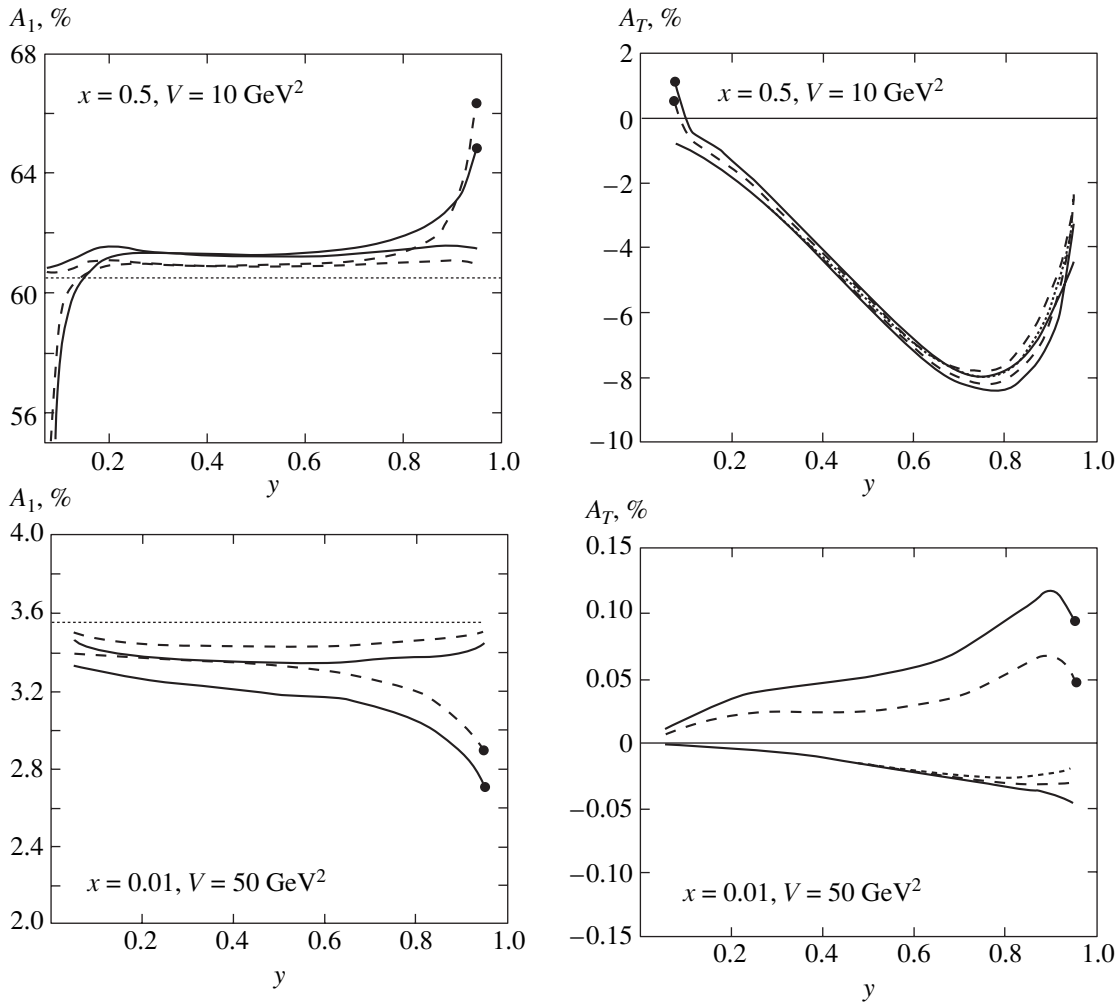

Fig. 4. Radiative correction to asymmetries for the HERMES (lower plots) and JLab (upper plots) kinematics. The dotted line shows the Born asymmetry. The full and dashed lines correspond to the total and one-loop contributions. Asymmetries with the elastic contribution taken into account are marked by dots at the end.

in [9]), and $A_1(x) = x^{0.725}$, suggested in [37]; we set $g_2 = 0$ (the definition of $A_1(x)$ is given below).

From these plots, we can see that the total radiative correction is basically determined by the one-loop correction with some important effect around kinematical boundaries. The sign and value of the higher-order effects are in agreement with the leading log estimations and calculations of the correction to the elastic radiative tail in [38, 39]. Two regions require special consideration: the region of higher y for the HERMES and JLab kinematics and the region near the pion threshold at JLab.

We define the polarization asymmetries in the standard way,

$$A_L = \frac{\sigma_\parallel}{\sigma}, \qquad A_T = \frac{\sigma_\perp}{\sigma}. \tag{39}$$

We can also define the spin asymmetry A_1 as $A_L = DA_1$ (for the chosen model where $g_2 = 0$), where D is the kinematical depolarization factor depending on the ratio R of the longitudinal and transverse photoabsorption cross sections,

$$D = \frac{y(2-y)(1+\gamma^2 y/2)}{y^2(1+\gamma^2) + 2(1-y-\gamma^2 y^2/4)(1+R)},$$

$$R = \frac{\sigma_L}{\sigma_T} = \frac{M(Q^2+\nu^2)}{Q^2\nu}\frac{F_2}{F_1} - 1,$$

where $\nu = yV/2M$ and $\gamma^2 = Q^2/\nu^2$. For fixed x, A_1 is a constant within our model, and it is therefore very convenient for graphical presentation and analysis of different radiative effects. Figure 4 gives the asymmetries A_1 and A_T for the kinematics of HERMES and JLab up to $y = 0.95$. The influence of higher-order and elastic radiative effects can be seen. Figure 5 gives the total correc-

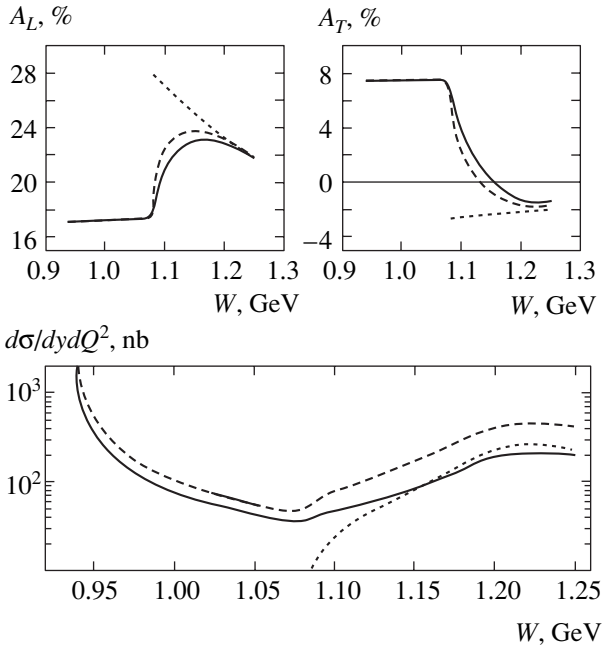

Fig. 5. The cross section (lower plot) and polarization asymmetries (both longitudinal and transverse) for the JLab kinematics ($Q^2 = 1$ GeV^2) near the pion threshold. The dotted line shows the Born cross section and asymmetry. Full and dashed lines correspond to the total and one-loop contributions.

tions to the cross sections and asymmetries for the threshold region of JLab.

7. CONCLUSIONS

We have considered model-independent QED radiative correction to the polarized deep inelastic and elastic electron–proton scattering. Together with the analytic expression for the radiative corrections, we give its numerical values for different experimental situations.

Our analytic calculations are based on the electron structure function method, which allows us to write both the spin-independent and spin-dependent parts of the cross section with the radiative corrections to the leptonic part of interaction taken into account in the form of the well-known Drell–Yan representation. The corresponding radiative corrections explicitly include the first-order correction as well as the leading-log contribution in all orders of the perturbation theory and the main part of the second-order next-to-leading-log contribution. Moreover, any model-dependent radiative correction to the hadronic part of the interaction can be included in our analytic result by inserting it as an additive part of the hard cross section in the integrand sign in master formulas (2) and (13).

To derive the radiative corrections, we take into account the radiation of photons and e^+e^- pairs in collinear kinematics, which produces a large logarithm L in the radiation probability (in D functions), and the radiation of one noncollinear photon, which enlarges the range of the hadron structure function arguments. It may be important that these functions are sufficiently sharp. In this case, the loss in the radiation probability (the loss of L) can be compensated by the increase in the value of the hard cross section.

We note that we extracted the explicit formulas for the first-order contribution at both leading and next-to-leading order levels. We found analytic agreement between these results for the one-loop correction with the previous results in [8], which provides the most important test of the total correction.

On the basis of the analytic results, we constructed the Fortran code ESFRAD.[3] Because of several known reasons discussed in Section 6, results obtained by the electron structure function method are usually not very convenient for precise numerical analysis. But we believe that our numerical procedure based on Monte Carlo integration allows us to overcome the obstacles.

Using the ESFRAD code, we performed numerical analysis for kinematical conditions of the current and future polarization experiments. We found two kinematical regions where the higher-order radiative correction can be important. These are the traditional region of high y and the region around the pion threshold. We gave a detailed analysis of the effects within these regions and presented numerical results within one of the simplest possibilities for modeling the deep inelastic scattering structure functions. Model dependence of the result is certainly an important issue requiring separate investigation for specific applications within the experimental data analysis.

ACKNOWLEDGMENTS

We thank our colleagues at the Jefferson Lab for useful discussions. We thank the US Department of Energy for support under contract no. DE-AC05-84ER40150. The work of N.M. was in addition supported by Rutgers University through the NSF grant no. PHY-9803860 and by Ukrainian DFFD. A.A. acknowledges additional support through the NSF grant no. PHY-0098642.

REFERENCES

1. B. W. Filippone and X. Ji, hep-ph/0101224.
2. M. Anselmino, A. Efremov, and E. Leader, Phys. Rep. **261**, 1 (1995); Erratum: Phys. Rep. **281**, 399 (1997).
3. M. K. Jones *et al.* (Jefferson Lab Hall A Collaboration), Phys. Rev. Lett. **84**, 1398 (2000).
4. L. D. van Buuren (97-01 Collaboration), Nucl. Phys. A **684**, 324 (2001).
5. T. V. Kukhto and N. M. Shumeiko, Nucl. Phys. B **219**, 412 (1983).
6. D. Y. Bardin and N. M. Shumeiko, Nucl. Phys. B **127**, 242 (1977).

[3] Fortran code ESFRAD is available at http://www.jlab.org/~aku/RC.

7. A. A. Akhundov, D. Y. Bardin, L. Kalinovskaya, and T. Riemann, Fortschr. Phys. **44**, 373 (1996).
8. I. V. Akushevich and N. M. Shumeiko, J. Phys. G **20**, 513 (1994).
9. I. Akushevich, A. Ilichev, N. Shumeiko, *et al.*, Comput. Phys. Commun. **104**, 201 (1997).
10. I. Akushevich, H. Bottcher, and D. Ryckbosch, hep-ph/9906408.
11. D. Y. Bardin, J. Bluemlein, P. Christova, and L. Kalinovskaya, Nucl. Phys. B **506**, 295 (1997).
12. I. V. Akushevich, A. N. Ilyichev, and N. M. Shumeiko, J. Phys. G **24**, 1995 (1998).
13. A. Afanasev, I. Akushevich, and N. Merenkov, Phys. Rev. D **64**, 113009 (2001).
14. A. V. Afanasev, I. Akushevich, A. Ilyichev, and N. P. Merenkov, Phys. Lett. B **514**, 269 (2001).
15. I. V. Akushevich, A. N. Ilichev, and N. M. Shumeiko, Phys. At. Nucl. **61**, 2154 (1998).
16. I. Akushevich, E. Kuraev, and B. Shaikhatdenov, Phys. Part. Nucl. **32**, 257 (2001).
17. S. D. Drell and T. Yan, Phys. Rev. Lett. **25**, 316 (1970); Erratum: Phys. Rev. Lett. **25**, 902 (1970).
18. E. A. Kuraev and V. S. Fadin, Yad. Fiz. **41**, 733 (1985) [Sov. J. Nucl. Phys. **41**, 466 (1985)].
19. E. A. Kuraev, N. P. Merenkov, and V. S. Fadin, Yad. Fiz. **47**, 1593 (1988) [Sov. J. Nucl. Phys. **47**, 1009 (1988)].
20. A. Afanasev, I. Akushevich, and N. Merenkov, hep-ph/0009273; Phys. Rev. D **65**, 013006 (2002).
21. V. N. Baier, V. S. Fadin, and V. A. Khoze, Nucl. Phys. B **65**, 381 (1973).
22. V. N. Gribov and L. N. Lipatov, Yad. Fiz. **15**, 781 (1972) [Sov. J. Nucl. Phys. **15**, 438 (1972)]; L. N. Lipatov, Yad. Fiz. **20**, 181 (1974) [Sov. J. Nucl. Phys. **20**, 94 (1975)]; Y. L. Dokshitzer, Zh. Éksp. Teor. Fiz. **73**, 1216 (1977) [Sov. Phys. JETP **46**, 641 (1977)]; G. Altarelli and G. Parisi, Nucl. Phys. B **126**, 298 (1977).
23. S. Jadach, M. Skrzypek, and B. F. Ward, Phys. Rev. D **47**, 3733 (1993).
24. M. Skrzypek, Acta Phys. Pol. B **23**, 135 (1992).
25. I. Akushevich, A. Arbuzov, and E. Kuraev, Phys. Lett. B **432**, 222 (1998).
26. G. I. Gakh, M. I. Konchatnij, and N. P. Merenkov, Pis'ma Zh. Éksp. Teor. Fiz. **71**, 328 (2000) [JETP Lett. **71**, 227 (2000)].
27. M. I. Konchatnij and N. P. Merenkov, Pis'ma Zh. Éksp. Teor. Fiz. **69**, 845 (1999) [JETP Lett. **69**, 893 (1999)].
28. M. I. Konchatnij, N. P. Merenkov, and O. N. Shekhovzova, Zh. Éksp. Teor. Fiz. **118**, 5 (2000) [JETP **91**, 1 (2000)].
29. A. V. Afanasev, I. V. Akushevich, G. I. Gakh, and N. P. Merenkov, Zh. Éksp. Teor. Fiz. **120**, 515 (2001) [JETP **93**, 449 (2001)].
30. N. P. Merenkov, Yad. Fiz. **48**, 1782 (1988) [Sov. J. Nucl. Phys. **48**, 1073 (1988)].
31. N. P. Merenkov, Yad. Fiz. **50**, 750 (1989) [Sov. J. Nucl. Phys. **50**, 469 (1989)].
32. A. B. Arbuzov, E. A. Kuraev, N. P. Merenkov, and L. Trentadue, Zh. Éksp. Teor. Fiz. **108**, 1164 (1995) [JETP **81**, 638 (1995)].
33. E. A. Kuraev, N. P. Merenkov, and V. S. Fadin, Yad. Fiz. **45**, 782 (1987) [Sov. J. Nucl. Phys. **45**, 486 (1987)].
34. R. Barbieri, J. A. Mignaco, and E. Remiddi, Nuovo Cimento A **11**, 824 (1972).
35. P. Amaudruz *et al.* (New Muon Collaboration), Nucl. Phys. B **371**, 3 (1992).
36. A. Arbuzov, D. Y. Bardin, J. Bluemlein, *et al.*, Comput. Phys. Commun. **94**, 128 (1996).
37. A. P. Nagaitsev, V. G. Krivokhijine, I. A. Savin, and G. I. Smirnov, JINR Rapid Commun. (Dubna), No. 3[71]-95, 59 (1995).
38. A. A. Akhundov, D. Y. Bardin, and N. M. Shumeiko, Yad. Fiz. **44**, 1517 (1986) [Sov. J. Nucl. Phys. **44**, 988 (1986)].
39. I. Akushevich, E. A. Kuraev, and B. G. Shaikhatdenov, Phys. Rev. D **62**, 053016 (2000).

Journal of Experimental and Theoretical Physics, Vol. 98, No. 3, 2004, pp. 417–426.
From Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 125, No. 3, 2004, pp. 476–485.

NUCLEI, PARTICLES, AND THEIR INTERACTION

Neutron–Antineutron Oscillations in a Trap Revisited[¶]

B. O. Kerbikov, A. E. Kudryavtsev, and V. A. Lensky

State Research Center Institute of Theoretical and Experimental Physics, Moscow, 117218 Russia
e-mail: borisk@heron.itep.ru, kudryavt@heron.itep.ru, lensky@itep.ru

Received September 9, 2003

Abstract—We reexamine the problem of n–$\bar{n}$ oscillations for ultracold neutrons confined within a trap. We show that, for up to 10^3 collisions with the walls, the process can be described in terms of wave packets. The $\bar{n}$ component grows linearly with time, with the enhancement factor depending on the reflection properties of the walls. *© 2004 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

For quite a long time, physics beyond the Standard Model has continued to be an intriguing subject. Several reactions that may serve as signatures for the new physics have been discussed. One of the most elegant proposals is to look for n–$\bar{n}$ oscillations [1] (see also [2]). There are three possible experimental settings aimed at observation of this process. The first is to establish a limit on nuclear instability because $\bar{n}$ produced inside a nucleus will blow it up. The second is to use a neutron beam from a reactor. This beam propagates a long distance to the target in which the possible $\bar{n}$ component would annihilate and, thus, is detected. The third option, which we discuss in the present paper, is to use ultracold neutrons (UCNs) confined in a trap. The main question is to what extent generation of the $\bar{n}$ component is reduced by the interaction with the trap walls. This subject was addressed by several authors [3–8]. In our opinion, a thorough investigation of the problem is still lacking.

First of all, a clear formulation of the problem of n–$\bar{n}$ oscillations in a cavity has been hitherto missing. Two different approaches were used without presenting sound arguments in favor of their applicability and without tracing connections between them.

In the first approach [4, 5], n–$\bar{n}$ oscillations are considered in the basis of the discrete eigenstates of the trap potential, with the splitting between n and $\bar{n}$ levels and $\bar{n}$ annihilation taken into account. The density of the trap eigenstates, which is proportional to the macroscopic trap volume, is huge and the states cluster together extremely thickly. But these arguments do not suffice to discard the discrete-state approach, because the n–$\bar{n}$ mixing parameter is much smaller than the distance between adjacent levels (see below). The true reason due to which the above treatment is of little physical relevance is as follows. The spectrum of the neutrons provided to the trap by the source is continuous and a certain time is needed for rearrangement of the initial wave function into standing waves corresponding to the trap eigenstates. As is shown below, this time interval appears to be of the order of the β-decay time, and, therefore, the standing wave regime, being interesting by itself, can hardly be reached in the real physical situation.

The second approach [3, 6, 7] treats the neutrons and antineutrons inside a trap as freely moving particles that undergo reflections from the trap walls. Collisions with the walls result in a reduction of the $\bar{n}$ component compared to the case of the free-space evolution. This suppression is due to two factors. The first is the annihilation inside the walls. The second is the phase decoherence of the n and $\bar{n}$ components induced by the difference of the wall potentials acting on n and $\bar{n}$. Reflections of antineutrons from the trap walls were, for the first time, considered in [3]. The purpose of that paper was to investigate the principal possibility of observing n–$\bar{n}$ oscillations in a trap, and the authors estimated the reflection coefficient for antineutrons without paying attention to the decoherence phenomena. Only a single collision with the trap wall was considered in [3]. A comprehensive study of n–$\bar{n}$ oscillations in a trap was presented in [6, 7]. Decoherence and multiple reflections and the influence of gravitational and magnetic fields were included. The approximate equation for the annihilation probability after N collisions obtained in [7, Eq. (3.8)] coincides with the exact formula (59) in the present paper when $N \gg 1$. As we show below, the N-independent asymptotic regime settles at $N \gtrsim 10$.

Derivation of the exact equation for the annihilation probability with an arbitrary number of collisions is not

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

the only purpose of the present work. We already mentioned the problem of the relation between the eigenvalue and the wave-packet approaches. Within the wave-packet approach, some basic notions such as the time between successive collisions and the collision time itself can be defined in a clear and rigorous way. Another question within the wave-packet formalism is the independence of the reflection coefficient from the width of the wave packet and the applicability of the stationary formalism to calculate reflections from the trap walls. These and some other principal points are, for the first time, considered in detail in the present paper.

We also mention that an alternative approach to the evaluation of the reflection coefficients for n and $\bar{n}$ was outlined in [8]. It is based on the time-dependent Hamilton formalism for the interaction of n and $\bar{n}$ with the trap walls. This subject remains outside the scope of the present paper.

The paper is organized as follows. In Section 2, we recall the basic equations describing n–$\bar{n}$ oscillations in free space. Section 3 is devoted to the optical potential approach to the interaction of n and $\bar{n}$ with the trap walls. In Section 4, we analyze the two formalisms proposed to treat n–$\bar{n}$ oscillations in the cavity, namely, box eigenstates and wave packets. In Section 5, reflection from the trap walls is considered. Section 6 contains the main result in this work, the time dependence of the $\bar{n}$ component production probability. In Section 7, conclusions are formulated and problems to be solved outlined.

2. OSCILLATIONS IN FREE SPACE

We start by recalling the basic equations describing n–$\bar{n}$ oscillations in free space. The phenomenological Hamiltonian is a 2×2 matrix in the basis of the two-component n–$\bar{n}$ wave function (we set $\hbar = 1$),

$$H_{jl} = \left(H_j - i\frac{\Gamma_\beta}{2}\right)\delta_{jl} + \epsilon(\sigma_x)_{jl}, \tag{1}$$

where $j, l = n, \bar{n}$, $H_j = k^2/2m - \mu_j B$, μ_j is the magnetic moment, B is the external (e.g., the Earth's) magnetic field, Γ_β is the β-decay width, ϵ is the n–$\bar{n}$ mixing parameter (see below), and σ_x is the Pauli matrix. Assuming the n and $\bar{n}$ wave functions to be plane waves, we write the two-component wave function of the n–$\bar{n}$ system as

$$\hat{\Psi}(x, t) = \begin{pmatrix} \psi_n(t) \\ \psi_{\bar{n}}(t) \end{pmatrix} e^{ikx}. \tag{2}$$

Evolution of the time-dependent part of $\hat{\Psi}(x, t)$ is then described by the equation

$$i\frac{\partial}{\partial t}\begin{pmatrix} \psi_n(t) \\ \psi_{\bar{n}}(t) \end{pmatrix} = \begin{pmatrix} E_n - i\frac{\Gamma_\beta}{2} & \epsilon \\ \epsilon & E_{\bar{n}} - i\frac{\Gamma_\beta}{2} \end{pmatrix}\begin{pmatrix} \psi_n(t) \\ \psi_{\bar{n}}(t) \end{pmatrix}. \tag{3}$$

The difference between E_n and $E_{\bar{n}}$ due to the Earth's magnetic field is

$$\omega = E_{\bar{n}} - E_n = 2|\mu_n|B \approx 6 \times 10^{-12}\ \text{eV}. \tag{4}$$

Diagonalizing the matrix in (3), we find $\psi_n(t)$ and $\psi_{\bar{n}}(t)$ in terms of their values at $t = 0$,

$$\psi_n(t) = \left(\psi_n(0)\left(\cos\nu t + \frac{i\omega}{2\nu}\sin\nu t\right) - \psi_{\bar{n}}(0)\frac{i\epsilon}{\nu}\sin\nu t\right) \times \exp\left[-\frac{1}{2}(i\Omega + \Gamma_\beta)t\right], \tag{5}$$

$$\psi_{\bar{n}}(t) = \left(-\psi_n(0)\frac{i\epsilon}{\nu}\sin\nu t + \psi_{\bar{n}}(0)\left(\cos\nu t - \frac{i\omega}{2\nu}\sin\nu t\right)\right) \times \exp\left[-\frac{1}{2}(i\Omega + \Gamma_\beta)t\right], \tag{6}$$

where $\Omega = E_n + E_{\bar{n}}$, $\nu = (\omega^2/4 + \epsilon^2)^{1/2}$, and $\omega = E_{\bar{n}} - E_n$. In particular, if $\psi_n(0) = 1$ and $\psi_{\bar{n}}(0) = 0$, we have

$$|\psi_{\bar{n}}(t)|^2 = \frac{4\epsilon^2}{\omega^2 + 4\epsilon^2}\exp(-\Gamma_\beta t) \times \sin^2\left(\frac{1}{2}\sqrt{\omega^2 + 4\epsilon^2}\,t\right). \tag{7}$$

The use of this equation to test fundamental symmetries is discussed in [9].

Without the magnetic field, i.e., for $\omega = 0$, and for $t \ll \epsilon^{-1}$, Eq. (7) yields

$$|\psi_{\bar{n}}(t)|^2 \approx \epsilon^2 t^2 \exp(-\Gamma_\beta t). \tag{8}$$

This law (for $t \ll \Gamma_\beta^{-1}$) has been used to establish the lower limit on the oscillation time $\tau = \epsilon^{-1}$. According to the ILL-Grenoble experiment [10],

$$\tau > 0.86 \times 10^8\ \text{s}. \tag{9}$$

The corresponding value of the mixing parameter is $\epsilon \approx 10^{-23}$ eV. This number is used in obtaining numerical results presented below.

The Earth's magnetic field leads to a strong suppression of the n–$\bar{n}$ oscillations. With the value of ω given by (4), Eq. (7) leads to

$$|\psi_{\bar{n}}(t)|^2 \approx \frac{4\epsilon^2}{\omega^2}\exp(-\Gamma_\beta t)\sin^2 t/\tau_B \approx 10^{-23}\sin^2 t/\tau_B, \quad (10)$$

where $\tau_B = (|\mu_n|B)^{-1} \approx 2 \times 10^{-4}$ s. In what follows, we assume that the magnetic field is screened.

For $\omega = 0$ but for arbitrary initial conditions, Eqs. (5) and (6) take the form

$$\psi_n(t) = (\psi_n(0)\cos\epsilon t - i\psi_{\bar{n}}(0)\sin\epsilon t) \times \exp\left[-\left(iE + \frac{\Gamma_\beta}{2}\right)t\right], \quad (11)$$

$$\psi_{\bar{n}}(t) = (-i\psi_n(0)\sin\epsilon t + \psi_{\bar{n}}(0)\cos\epsilon t) \times \exp\left[-\left(iE + \frac{\Gamma_\beta}{2}\right)t\right], \quad (12)$$

where $E = E_n = E_{\bar{n}}$.

3. OPTICAL POTENTIAL MODEL FOR THE TRAP WALL

We remind the reader that neutrons with the energy $E < 10^{-7}$ eV are called ultracold. An excellent review of UCN physics was given in [11] (see also [12]).

A useful relation connecting the neutron velocity v in cm/s and E in eV is given by

$$v[\text{cm/s}] = 10^2(10^9 E[\text{eV}]/5.22)^{1/2}. \quad (13)$$

For $E = 10^{-7}$ eV, the velocity is $v \approx 4.4 \times 10^2$ cm/s.

A less formal definition of UCN involves the notion of the real part of the optical potential corresponding to the trap material (see below). Neutrons with energies less than the height of this potential are called ultracold. The two definitions are essentially equivalent because, as we see in what follows, the real part of the optical potential is of the order 10^{-7} eV for most materials.

Our main interest is in strongly absorptive interaction of the $\bar{n}$ component with the trap walls. We therefore ignore very weak absorption of UCN on the walls [11, 12]. Due to complete reflection from the trap walls, UCN can be stored for about 10^3 s (β-decay time), as was first pointed out in [13].

To be specific, we consider UCN with $E = 0.8 \times 10^{-7}$ eV, which corresponds to $v = 3.9 \times 10^2$ cm/s (see (13)), $k = 12.3$ eV, and de Broglie wavelength $\lambda \approx 10^{-5}$ cm. In the next section, we describe UCN in terms of wave packets, and, hence, the above values must be attributed to the center of the packet.

We treat the interaction of n and $\bar{n}$ with the trap walls in terms of an energy-independent optical potential. The validity of this approach to UCN has been justified in a number of papers (see, e.g., [11, 12, 14]). There is still an open question concerning the discrepancy between theoretical prediction and experimental data on the UCN absorption. Interesting by itself, this problem is outside the scope of our work because, as was already mentioned, absorption of neutrons may be ignored in the n–$\bar{n}$ oscillation process. The low-energy optical potential is given by

$$U_{j\text{A}} = \frac{2\pi}{m}Na_{j\text{A}}, \quad (14)$$

where $j = n, \bar{n}$; m is the neutron mass; N is the number of nuclei in a unit volume; and $a_{j\text{A}}$ is the j–A scattering length, which is real for n and complex for $\bar{n}$. For neutrons, the scattering lengths $a_{n\text{A}}$ are accurately known for various materials [12]. For antineutrons, the situation is different. Experimental data on low-energy $\bar{n}$–A interaction are absent. Only some indirect information may be gained from level shifts in antiprotonic atoms, and, therefore, the values of $a_{\bar{n}\text{A}}$ used in [3, 6, 8, 15] as an input in the n–$\bar{n}$ oscillation problem are similar but not the same. We consider the set of $a_{\bar{n}\text{A}}$ calculated in [16] within the framework of the internuclear cascade model as most reliable. Even this particular model leads to several solutions, and the one that we have chosen for ^{12}C (graphite and diamond) may be called "motivated" by [16]. To estimate the dependence on the material of the walls and to compare our results with those in [3], we also performed calculations for Cu. Scattering lengths for Cu are not given in [16], and we used the solution proposed in [3]. Our calculations were thus performed with the $\bar{n}$–A scattering lengths

$$a_{\bar{n}\text{C}} = (3 - i1)\text{ fm}, \quad a_{\bar{n}\text{Cu}} = (5 - i0.5)\text{ fm}. \quad (15)$$

The scattering lengths for neutrons are [12]

$$a_{n\text{C}} = 6.65\text{ fm}, \quad a_{n\text{Cu}} = 7.6\text{ fm}. \quad (16)$$

The concentrations of atoms N entering (14) are as follows:

$$N_{\text{C(graphite)}} = 1.13 \times 10^{-16}\text{ fm}^{-3},$$

$$N_{\text{C(diamond)}} = 1.63 \times 10^{-16}\text{ fm}^{-3},$$

$$N_{\text{Cu}} = 0.84 \times 10^{-16}\text{ fm}^{-3}.$$

In accordance with (14), the optical potentials are

then given by

$$\begin{aligned} U_{n\mathrm{C(gr)}} &= 1.95\times10^{-7}\ \mathrm{eV},\\ U_{n\mathrm{C(diam)}} &= 2.8\times10^{-7}\ \mathrm{eV},\\ U_{n\mathrm{Cu}} &= 1.66\times10^{-7}\ \mathrm{eV}; \end{aligned} \tag{17}$$

$$\begin{aligned} U_{\bar{n}\mathrm{C(gr)}} &= (0.9-i0.3)\times10^{-7}\ \mathrm{eV},\\ U_{\bar{n}\mathrm{C(diam)}} &= (1.3-i0.4)\times10^{-7}\ \mathrm{eV},\\ U_{\bar{n}\mathrm{Cu}} &= (2-i0.2)\times10^{-7}\ \mathrm{eV}. \end{aligned} \tag{18}$$

In this paper, we consider particles (n and $\bar{n}$) with energies below the potential barrier formed by the real part of the potential. For $\bar{n}$ and ^{12}C, the limiting velocity is $v = 4.15\times10^2$ cm/s.

4. WAVE PACKET VERSUS STANDING WAVES

It is convenient to use the short notation

$$U_j = V_j - iW_j\delta_{j\bar{n}} \tag{19}$$

for optical potentials (17) and (18), where $j = n, \bar{n}$ and the wall material is not indicated explicitly. We consider the following model for the trap in which n–$\bar{n}$ oscillations may possibly be observed. We imagine two walls of type (19) separated by a distance of $L \sim 10^2$ cm, i.e., the one-dimensional potential well of the form

$$U_j(x) = \{\theta(-x-L)+\theta(x)\}\{V_j - iW_j\delta_{j\bar{n}}\}, \tag{20}$$

with $\theta(x)$ being the step function. Our goal is to follow the time evolution of the $\bar{n}$ component in such a trap, assuming that the initial state is a pure n one.

The first question to be answered is how to describe the wave function of the system. Two different approaches seem to be feasible, and both were discussed in the literature [4, 6, 8]. The first is to consider oscillations occurring in the wave packet and to investigate to what extent reflections from the walls distort the picture compared to the free-space regime. The second approach is to consider the eigenvalue problem in potential well (20), to find energy levels for n and $\bar{n}$, and to consider oscillations in this basis. Because of different interactions with the walls, the levels of n and $\bar{n}$ are split and the $\bar{n}$ levels acquire annihilation widths.

At first glance, this approach might seem inadequate because, in a trap with $L \sim 10^2$ cm, the density of states is very high, the characteristic quantum number corresponding to the UCN energy is very large, and the splitting δE between adjacent n levels (or between the levels of the n and $\bar{n}$ spectra) is extremely small. The values of all these quantities are given below, and it follows that $\delta E < 10^{-14}$ eV. However, this approach cannot be discarded without further analysis, because the n–$\bar{n}$ mixing parameter $\epsilon \approx 10^{-23}$ eV is much smaller than δE.

To understand the relation between the two approaches, we note that the initial conditions correspond to a beam of UCN provided by a source. The momentum spectrum of UCN depends on the specific experimental conditions. In order to stay on general grounds and, at the same time, to simplify the problem, we assume that the UCN beam entering the trap has the form of a Gaussian wave packet. We suppose that, at $t = 0$, the center of the wave packet is at $x = x_0$, and, hence,

$$\psi_k(x, t=0) = (\pi a^2)^{-1/4} \times \exp\left(-\frac{(x-x_0)^2}{2a^2} + ikx\right), \tag{21}$$

where a is the width of the wave packet in coordinate space. The normalization of wave function (21) corresponds to one particle in the entire one-dimensional space,

$$\int_{-\infty}^{+\infty} dx|\psi_k(x, t=0)|^2 = 1. \tag{22}$$

For $E = 0.8\times10^{-7}$ eV and the beam resolution $\Delta E/E = 10^{-3}$, we have

$$k = 12.3\ \mathrm{eV}, \quad a = 3.2\times10^{-3}\ \mathrm{cm}. \tag{23}$$

The width of wave packet (21) increases with time according to

$$a' = a\left[1+\left(\frac{t}{ma^2}\right)^2\right]^{1/2} \approx \frac{t}{ma} \tag{24}$$

and becomes comparable with the trap size L for $t \sim 10^3$ s. For the wave hitting the wall and the reflected wave to be clearly resolved, the condition $a'/v \ll \tau_L$ or $a' \ll L$ must be satisfied, where $\tau_L \sim 1$ s is the time between two consecutive collisions with the trap walls. Reflection of the wave packet from the walls is considered in detail in the next section. Here, we show that $t \sim 10^3$ s is the characteristic time needed for the rearrangement of the initial wave packet into stationary states of the trapping box.

We consider the eigenvalue problem for potential well (20). The parameters of potential (20) for neutrons are $V_n \approx 2\times10^{-7}$ eV and $L \approx 10^2$ cm. The number of levels is

$$M \approx \frac{L\sqrt{2mV}}{\pi} \approx \frac{10^8}{\pi}. \tag{25}$$

According to (23), the center of wave packet (21) has the momentum $k = 12.3$ eV, which corresponds to a state with the number of nodes $j \approx 2 \times 10^7$ and $k_j L \approx 6 \times 10^7 \gg 1$. Positions of such highly excited levels in a finite-depth potential are indistinguishable from the spectrum in a potential box with infinite walls. Therefore,

$$\varphi_j(x) \approx \sqrt{\frac{2}{L}} \sin \omega_j x, \quad \omega_j = \frac{\pi j}{L}. \tag{26}$$

Wave functions (26) describe semiclassical states with $j \gg 1$ in a potential well with sharp edges. The "frequency" ω_j is very high compared to the width of the wave packet in momentum space,

$$\omega_j \approx 6 \times 10^5 \text{ cm}^{-1} \gg \nu = \frac{1}{\sqrt{2}a} \approx 2 \times 10^2 \text{ cm}^{-1}.$$

This implies that the wave packet spans over a large number of levels. To determine this number, we note that the distance between adjacent levels around the center of the wave packet is

$$\delta E = E_{j+1} - E_j \sim 10^{-14} \text{ eV}.$$

The highly excited levels within the energy band

$$\Delta E = 10^{-3} E \sim 10^{-10} \text{ eV}$$

corresponding to wave packet (21) are to a high accuracy equidistant, as they should be in the semiclassical regime. The number of states within ΔE is

$$\Delta j = \Delta E / \delta E \sim 10^4$$

and their density in momentum space is

$$\rho(\omega) = a \Delta j \approx L/\pi \sim 10^6 \text{ eV}^{-1}. \tag{27}$$

We can now answer the question formulated at the beginning of this section, namely, whether the n–$\bar{n}$ oscillations in the trap should be described in terms of the wave packet or in terms of the stationary eigenfunctions. At $t = 0$, the wave function has the form of the wave packet (21) provided by the UCN source. Due to collisions with the trap walls, transitions from the initial state (21) into discrete (or quasi-discrete for $\bar{n}$) eigenstates (26) occur.

The time evolution of the initial wave function (21) proceeds according to

$$\psi(x, t) = \int dx' G(x, t; x', 0) \psi_k(x', 0), \tag{28}$$

where $G(x, t; x', 0)$ is the time-dependent Green's function for potential well (20). Using the spectral representation for G, we can write

$$\psi(x, t) = \sum_j e^{-iE_j t} \varphi_j(x) \int dx' \varphi_j^*(x') \varphi_k(x', 0). \tag{29}$$

In the semiclassical approximation, the distance between the adjacent levels is $\delta E = \pi/\tau_L$, and, therefore, one may think that, at $t = \tau_L$, i.e., already at the first collision, the neighboring terms in (29) would cancel each other. But this is not the case. Indeed,

$$\varphi_{j+1}(x) \exp(-iE_{j+1} t) + \varphi_j(x) \exp(-iE_j t)$$

$$= \frac{\exp(-iE_j t)}{i\sqrt{2}L} \left[\exp(i\omega_j x) \left(1 + \exp\left(i \frac{\pi}{L} (x - vt) \right) \right) \right.$$

$$\left. - \exp(-i\omega_j x) \left(1 + \exp\left(-i \frac{\pi}{L} (x + vt) \right) \right) \right].$$

Therefore, there is a constructive interference at $x = \pm vt$ either in the first or in the second term, respectively. This is true with the whole sum of terms in (29) taken into account, and, hence, we can pass from summation to integration in (29). The overlap of the wave functions entering (29) can be easily evaluated, provided the center of the wave packet x_0 is not within the bandwidth distance a' from the trap walls. The overlap is given by the integral

$$\int dx' \varphi_j^*(x') \psi_k(x', 0) \approx \frac{i}{(2\sqrt{\pi} L a)^{1/2}}$$

$$\times \int_{-L}^{0} dx' \exp\left(-\frac{(x' - x_0)^2}{2a^2} + i(k - \omega_j) x' \right)$$

$$= \frac{i}{(2\sqrt{\pi} L a)^{1/2}} \tag{30}$$

$$\times \int_0^L dx'' \exp\left(-\frac{(x'' + x_0)^2}{2a^2} - i(k - \omega_j) x'' \right).$$

At this step, we have omitted the exponential with the high frequency $(k + \omega_j)$. We next take $(x'' + x_0)/(\sqrt{2}a)$ as a new variable and assume that $|x_0| \gg a$, $L - |x_0| \gg a$ (we recall that x_0 is negative because $-L < x < 0$). The result is that

$$\int dx' \varphi_j^*(x') \psi_k(x', 0)$$

$$\approx i \left(\frac{\sqrt{\pi} a}{L} \right)^{1/2} \exp\left(-\frac{a^2}{2} (k - \omega_j)^2 + i(k - \omega_j) x_0 \right). \tag{31}$$

Corrections to (31) are of the order of a/L. We now consider frequency summation in (29). This summation can be replaced by integration over ω, because the density of semiclassical states $\rho(\omega)$ is very high. We thus arrive at

$$\psi(x,t) = \frac{1}{\left[\sqrt{\pi}a\left(1+i\frac{t}{ma^2}\right)\right]^{1/2}} \times\left(\exp\left[-\frac{\alpha(x,t)}{2a^2\left(1+\frac{t^2}{m^2a^4}\right)}\right] + \exp\left[-\frac{\alpha(-x,t)}{2a^2\left(1+\frac{t^2}{m^2a^4}\right)}\right]\right), \tag{32}$$

$$\alpha(x,t) = (x-x_0-v_0t)^2 - it\frac{(x-x_0)^2}{ma^2} - 2ik_0a^2(x-x_0) + i\frac{k_0^2a^2}{m}t + 2ik_0x_0\left(a^2+\frac{t^2}{m^2a^2}\right). \tag{33}$$

The second term in Eq. (32) describes the reflected wave packet (see the next section). According to (21), (28), and (32), all that happens to the wave packet in the trap is broadening and reflections. This is true during some initial period of its life history at least. How long does this period last? The answer to this question may be obtained by estimating the accuracy of performing frequency integration instead of summation over discrete states in (29).

To estimate the time scale for the rearrangement of initial wave packet (21) into trap standing waves (26), it is convenient to introduce the difference

$$\delta\psi(x,t) = \psi_{\text{sum}}(x,t) - \psi_{\text{int}}(x,t)$$

between the "exact" wave function (29) and the approximate integral representation (32). Whenever

$$\delta w(t) = \int dx\left(|\psi_{\text{sum}}|^2 - |\psi_{\text{int}}|^2\right) = 2\int dx\,\Re(\psi_{\text{int}}\delta\psi) \ll 1, \tag{34}$$

we can consider oscillations as proceeding in the wave packet basis. With

$$f(\omega) = \sqrt{\frac{\sqrt{\pi}a}{2L^2}} \times\exp\left(-\frac{a^2}{2}(k_0-\omega)^2 - i\frac{\omega^2}{2m}t + i\omega(x-x_0) + ikx_0\right), \tag{35}$$

we have the estimate

$$\begin{aligned}\delta\psi(x,t) &= \sum_n f(\omega_n) - \int d\omega\rho(\omega)f(\omega)\\ &= -\sum_n\int_{\omega_n}^{\omega_{n+1}} d\omega\rho(\omega)(f(\omega)-f(\omega_n))\\ &\approx -\sum_n\int_{\omega_n}^{\omega_{n+1}} d\omega\rho(\omega)f'(\omega_n)(\omega-\omega_n)\\ &= -\frac{1}{2}\sum_n f'(\omega_n)(\omega_{n+1}-\omega_n).\end{aligned} \tag{36}$$

From (35), we find that

$$\begin{aligned}f'(\omega) &= g(\omega)f(\omega),\\ g(\omega) &= i(x-x_0-vt)-(k_0-\omega)a^2.\end{aligned} \tag{37}$$

Because $f(\omega)$ is a narrow Gaussian peak, we can substitute $g(\omega)$ by $g(k_0)$, and then (36) results in

$$\delta\psi(x,t) \approx \frac{\pi}{2L}(x-x_0-v_0t)\psi_{\text{int}}(x,t). \tag{38}$$

From (34) and (38), we have

$$\delta w \approx \frac{\pi}{2L}\int_{-\infty}^{+\infty} dx|x-x_0-v_0t|\,|\psi_{\text{int}}(x,t)|^2 \propto \frac{a'}{L} \sim \frac{t}{maL} \sim \frac{t}{10^3\ \text{s}}, \tag{39}$$

where a' is given by (23).

Roughly speaking, the time $t \sim 10^3$ s needed for the neutron wave function to rearrange into the trap eigenstate is comparable to the neutron lifetime, and the neutron would rather "die" than adjust to the new boundary conditions. The wave packet formalism is therefore used in what follows. Some additional subtleties arising from the quantization of levels in the trapping box are discussed in Section 7.

5. REFLECTION FROM THE TRAP WALLS

We return to one-dimensional trap (20). Let the particle moving from $x = -\infty$ enter the trap at $t = 0$ through the window at $x = -L$. At $t = \tau_L$, it reaches the wall at $x = 0$, the n component is reflected from the wall, and the $\bar{n}$ component is partly reflected and partly

absorbed. The wave packet describing the interaction with the wall has the form

$$\psi(x,t) = \pi^{-3/4}\sqrt{\frac{a}{2}}\int_{-\infty}^{+\infty} dk\,\psi_j(k,x) \times \exp\left(-\frac{a^2}{2}(k-k_0)^2 + iL(k-k_0) - i\frac{t}{2m}k^2\right), \quad (40)$$

where $j = n, \bar{n}$ and

$$\psi_j(k,x) = e^{ikx} + R(k)e^{-ikx} = e^{ikx} + \rho_j(k)e^{\phi_j(k)}e^{-ikx}. \quad (41)$$

For the n component, $\rho_n(k) = 1$, because we neglect very weak absorption of neutrons at the surface. The integral (40) with the first term in (41) is trivial. To integrate the second term in (41), we note that, due to the Gaussian form factor with $ak_0 \sim 10^3 \gg 1$, the dominant contribution to integral (40) comes from a narrow interval of k around k_0. Expanding $R_j(k)$ at $k - k_0$ and keeping the leading term, we obtain

$$R_j(k) \approx \rho_j(k_0)\exp(i\phi_j(k_0)) \times \left[1 + i\phi_j'(k_0)(k-k_0) + \delta_{j\bar{n}}\frac{\rho_j'(k_0)}{\rho_j(k_0)}(k-k_0)\right] \approx \rho_j(k_0)\exp(i\phi_j(k_0) + i\phi_j'(k_0)(k-k_0)). \quad (42)$$

The validity of the last step for $\bar{n}$ becomes clear from the explicit expressions for $\rho_{\bar{n}}(k)$ and $\phi_{\bar{n}}(k)$ presented below.

Integration in (40) can now be easily performed, with the result [17]

$$\psi_j(x,t) = \frac{1}{\left[\sqrt{\pi}a\left(1 + i\frac{t}{ma^2}\right)\right]^{1/2}} \times \left(\exp\left[-\frac{\alpha_{\rm inc}(x,t)}{2a^2\left(1 + \frac{t^2}{m^2a^4}\right)}\right] + R_j(k_0)\exp\left[-\frac{\alpha_{\rm refl}(x,t)}{2a^2\left(1 + \frac{t^2}{m^2a^4}\right)}\right]\right), \quad (43)$$

$$\alpha_{\rm inc}(x,t) = (x + L - v_0 t)^2 - it\frac{(x+L)^2}{ma^2} - 2ik_0a^2(x+L) + i\frac{k_0^2a^2}{m}t + 2ik_0L\left(a^2 + \frac{t^2}{m^2a^2}\right), \quad (44)$$

$$\alpha_{\rm refl}(x,t) = \alpha_{\rm inc}(-x + \phi', t) + 2ik_0\phi'\left(a^2 + \frac{t^2}{m^2a^2}\right). \quad (45)$$

From (43)–(45), we see that the essence of $R(k)$ in the wave packet formalism is the same as in the time-independent approach. Therefore, imposing standard boundary conditions at $x = 0$, we obtain the reflection coefficients

$$R_j(k) = \rho_j(k)\exp(i\phi_j(k)) = \frac{k - i\kappa_j}{k + i\kappa_j}, \quad (46)$$

$$\kappa_n = [2m(V_n - E)]^{1/2}, \quad \kappa_{\bar{n}} = [2m(V_{\bar{n}} - iW_{\bar{n}} - E)]^{1/2} = \kappa_{\bar{n}}' - i\kappa_{\bar{n}}'', \quad (47)$$

$$\tan\phi_n = \frac{-2k\kappa_n}{k^2 - \kappa_n^2}, \quad \tan\phi_{\bar{n}} = \frac{-2k\kappa_{\bar{n}}'}{k^2 - (\kappa_{\bar{n}}')^2 - (\kappa_{\bar{n}}'')^2}, \quad (48)$$

$$\rho_n = 1, \quad \rho_{\bar{n}}^2 = 1 - \frac{4k\kappa_{\bar{n}}''}{(k + \kappa_{\bar{n}}'')^2 + (\kappa_{\bar{n}}')^2}. \quad (49)$$

For ^{12}C (graphite), in particular,

$$\rho = 0.56, \quad \theta \equiv \phi_{\bar{n}} - \phi_n = 0.72. \quad (50)$$

The first term on the right-hand side of (45) can be written as $[-x + L - v_0(t - \phi'/v_0)]^2$. Hence, the collision time or time delay is [17, 18]

$$\tau_{j,\,\rm coll} = \frac{\phi_j'(k_0)}{v_0} = {\rm Re}\frac{2m}{k\kappa_j}. \quad (51)$$

For neutrons, i.e., for real κ_n, Eq. (51) gives the well-known result

$$\tau_{n,\,\rm coll} = [E(V_n - E)]^{-1/2}.$$

This result is in line with the naive estimate $\tau_{n,\,\rm coll} \sim l/v_0 \sim 10^{-8}$ s [8], where $l \lesssim \lambda$ is the penetration depth.

For ^{12}C (graphite), Eq. (51) yields

$$\tau_{n,\,\rm coll} = 0.7 \times 10^{-8}\ {\rm s}, \quad \tau_{\bar{n},\,\rm coll} = 1.1 \times 10^{-8}\ {\rm s}. \quad (52)$$

Equations (43)–(45) supplemented by the above inequality make it possible to follow the time evolution of the beam inside the trap. We imagine an observer placed at a bandwidth distance from the wall, i.e., at $x = -a$. According to (43)–(45), such an observer concludes that the incident wave (the first term in (43)) dominates at times $t \le \tau_L - \tau_a$, while the reflected wave prevails at $t \ge \tau_L + \tau_a$. With this splitting of the time interval around $N\tau_L$, $N = 1, 2, \ldots$, we use the notation $(N\tau_L-)$ and $(N\tau_L+)$ for the moments before and after the Nth collision. Thus, we can calculate the $\bar{n}$ production

rate, because we have rigorous definitions of the collision time and the time interval between the two subsequent collisions.

6. ANNIHILATION RATE IN A TRAP

We can now inquire into the problem of time-dependence of the $\bar{n}$ production probability. In free space, it is given by $|\psi_{\bar{n}}(t)|^2 = \epsilon^2 t^2$ (see (2)), while in a trap with the complete annihilation or total loss of coherence at each collision, it has a linear time dependence $|\psi_{\bar{n}}(t)|^2 = \epsilon^2 \tau_L t$ [8].

To avoid cumbersome equations and because we consider the time interval $t \ll \Gamma_\beta^{-1}$, we omit $\exp(-\Gamma_\beta t)$ factors. Production of $\bar{n}$ during the collision can also be neglected [8]. The difference in collision times (52) for n and $\bar{n}$ may also be ignored. In the previous section, we have seen that the interaction of the wave packet with the wall is described in terms of reflection coefficients (46).[1]

We assume that, at $t = 0$, a pure-n beam enters the trap at $x = -L$. After crossing the trap, i.e., at $t = (\tau_L-)$, the time-dependent parts of the wave functions are given by (12),[2]

$$\begin{aligned}\psi_n(\tau_L-) &= \cos(\epsilon\tau_L)\exp(-iE\tau_L),\\ \psi_{\bar{n}}(\tau_L-) &= \sin(\epsilon\tau_L)\exp[-i(E\tau_L+\pi/2)].\end{aligned} \tag{53}$$

After the first reflection at $t = (\tau_L+)$, we have

$$\begin{aligned}\psi_{\bar{n}}(\tau_L+) &= \cos(\epsilon\tau_L)\exp[-i(E\tau_L-\phi_n)],\\ \psi_{\bar{n}}(\tau_L+) &= \rho_{\bar{n}}\sin(\epsilon\tau_L)\exp[-i(E\tau_L-\phi_{\bar{n}}+\pi/2)].\end{aligned} \tag{54}$$

Evolution from $t = (\tau_L+)$ to $t = (2\tau_L-)$ again proceeds in accordance with (12),

$$\begin{aligned}\psi_{\bar{n}} &= \frac{1}{2}\sin(2\epsilon\tau_L)(1+\rho e^{i\theta})\\ &\times \exp[-i(2E\tau_L-\phi_{\bar{n}}+\pi/2)]\\ &\approx \epsilon\tau_L(1+\rho e^{i\theta})\exp[-i(2E\tau_L-\phi_{\bar{n}}+\pi/2)],\end{aligned} \tag{55}$$

where $\theta = \phi_{\bar{n}} - \phi_n$ is the decoherence phase and $\rho \equiv \rho_{\bar{n}}$.

[1] An alternative description using time-evolution operators was proposed in [8].

[2] We state this although the Gaussian form factor in (43) also depends on time, the corresponding terms in the time-dependent Schrödinger equation are of the order of $1/ak_0$ compared to the derivative of the exponent $\exp(-iEt)$; we also note that the form factors are the same for n and $\bar{n}$ up to a constant multiplier.

The answer for $\psi(N\tau_L-)$ now seems evident:

$$\begin{aligned}\psi_{\bar{n}}(N\tau_L-) &= \epsilon\tau_L\frac{1-\rho^N e^{iN\theta}}{1-\rho e^{i\theta}}\\ &\times \exp[-i(NE\tau_L-\phi_n+\pi/2)].\end{aligned} \tag{56}$$

This conjecture is easy to verify by induction. For $t = (2\tau_L-)$, the result was derived explicitly in (55). Evolving (56) through one reflection at $t = N\tau_L$ and free propagation from $t = (N\tau_L+)$ to $t = ((N+1)\tau_L-)$, we arrive at (56) with $(N+1)$ instead of N. This completes the proof.

Therefore, the admixture of $\bar{n}$ before the Nth collision, i.e., at $t = N\tau_L-$, is

$$|\psi_{\bar{n}}(N\tau_L-)|^2 = \epsilon^2\tau_L^2\frac{1+\rho^{2N}-2\rho^N\cos N\theta}{1+\rho^2-2\rho\cos\theta}. \tag{57}$$

The annihilation probability at the jth collision is

$$P_a(j) = (1-\rho^2)|\psi_{\bar{n}}(j\tau_L-)|^2. \tag{58}$$

The total annihilation probability after N collisions is therefore given by

$$\begin{aligned}P_a(N) &= (1-\rho^2)\sum_{k=1}^{N}|\psi_{\bar{n}}(k\tau_L)|^2\\ &= \frac{\epsilon^2\tau_L^2(1-\rho^2)}{1+\rho^2-2\rho\cos\theta}\left(N+\frac{\rho^2(1-\rho^{2N})}{1-\rho^2}\right.\\ &\left.-2\rho\frac{\cos\theta-\rho-\rho^N[\cos(N+1)\theta+\rho\cos N\theta]}{1+\rho^2-2\rho\cos\theta}\right).\end{aligned} \tag{59}$$

After several collisions, the terms proportional to ρ^N, ρ^{2N}, and ρ^{N+1} may be dropped, because $\rho \sim 0.5$ (see (50)). Then, (59) takes the form

$$\begin{aligned}P_a(N) &\approx \frac{\epsilon^2\tau_L^2}{1+\rho^2-2\rho\cos\theta}\\ &\times\left(N(1-\rho^2)+1-\frac{(1-\rho^2)^2}{1+\rho^2-2\rho\cos\theta}\right).\end{aligned} \tag{60}$$

Three different regimes may be inferred from (60). For a very strong annihilation, i.e., $\rho \ll 1$,

$$P_a(N) = \epsilon^2\tau_L^2 N = \epsilon^2\tau_L t. \tag{61}$$

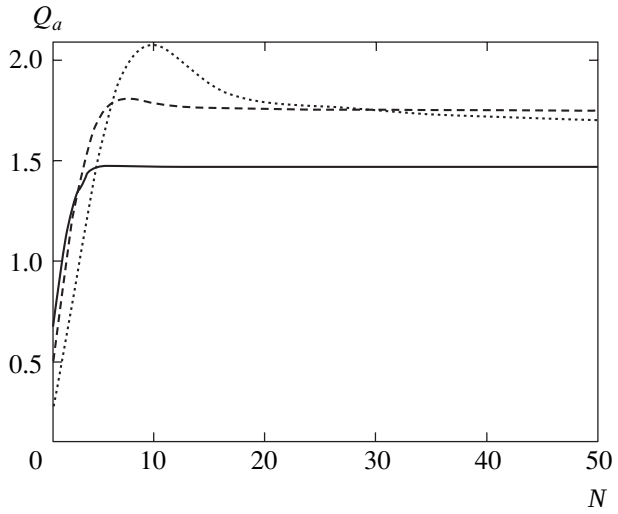

Fig. 1. Plot of the $Q_a(N) = (\epsilon^2\tau_L t)^{-1}P_a(N)$ dependence vs. N. The solid line corresponds to ^{12}C (graphite); the dashed one, to ^{12}C (diamond); and the dotted one, to Cu.

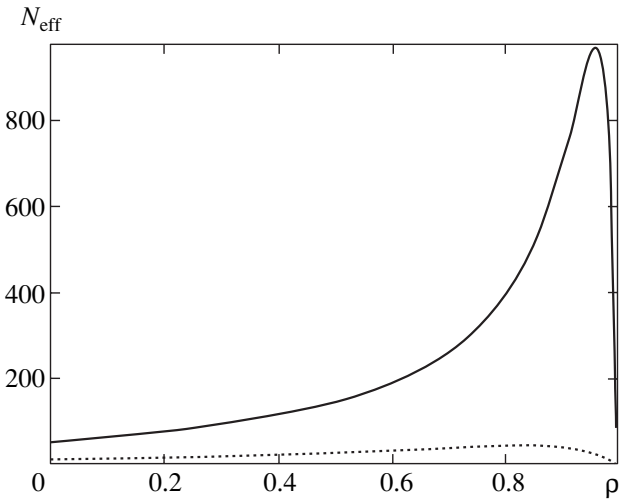

Fig. 2. Plot of the N_{eff} dependence vs. ρ at θ = 0. The solid line is for the number of collisions $N = 50$; the dashed line corresponds to $N = 10$.

For the complete decoherence at each collision, i.e., for $\theta = \pi$,

$$P_a(N) = \epsilon^2\tau_L^2\left(N\frac{1-\rho}{1+\rho} + \frac{\rho(2-\rho)}{(1+\rho)^2}\right) \approx \frac{1-\rho}{1+\rho}\epsilon^2\tau_L t. \quad (62)$$

For the (unrealistic) situation where $\theta = 0$,

$$P_a(N) = \epsilon^2\tau_L^2\left(N\frac{1+\rho}{1-\rho} - \frac{\rho(2+\rho)}{(1-\rho)^2}\right) \approx \frac{1+\rho}{1-\rho}\epsilon^2\tau_L t. \quad (63)$$

For the values of ρ and θ corresponding to optical potentials (17) and (18), the quantity

$$Q_a(N) = (\epsilon^2\tau_L^2 N)^{-1} = (\epsilon^2\tau_L t)^{-1}P_a(N)$$

calculated in accordance with the exact equation (59) is displayed in Fig. 1. This figure shows that the linear time dependence settles after about 10 collisions with the trap walls. The asymptotic value of $Q_a(N)$, which may be called the enhancement factor, is 1.5–2, depending on the wall material.

Proposals have been discussed in the literature [6, 19] to compensate the decoherence phase θ by applying the external magnetic field. Assuming the ideal situation that the regime $\theta = 0$ may be achieved in such a way and also assuming that the reflection coefficient ρ can be varied in the whole range by varying the trap material, we plot the quantity $N_{\text{eff}}(\rho)$ defined as

$$P_a(N) = \epsilon^2\tau_L^2 N_{\text{eff}}(\rho) \quad (64)$$

in Fig. 2. Thus defined, $N_{\text{eff}}(\rho)$ obviously depends also on the number of collisions N; the results for $N = 10$ and $N = 50$ are presented in Fig. 2. This figure shows what can be expected from the trap experiments in the most favorable, although hardly realistic, scenario.

7. CONCLUDING REMARKS

We have reexamined the problem of n–$\bar{n}$ oscillations for UCN in a trap. Our aim was to present a clear formulation of the problem, to calculate the amplitude of the $\bar{n}$ component for an arbitrary observation time and for any given reflection properties of the trap walls. We have shown that, for the physically relevant observation time (i.e., for the time interval less than the β-decay time), the process of n–$\bar{n}$ oscillations is described in terms of wave packets, while the standing-wave regime may settle only at later times. By calculating the difference between the n and $\bar{n}$ collision times, new light has been shed on the decoherence phenomena. For the first time, an exact equation has been derived for the annihilation probability for an arbitrary number of collisions with the trap walls. In line with the conclusions of the previous authors on the subject, this probability grows linearly with time. We have calculated the enhancement factor entering this linear time dependence and found this factor to be 1.5–2, depending on the reflection properties of the wall material.

Despite the extensive investigations reviewed in this article and the results of the present paper, the list of problems for further work is large. The central and most difficult task is to obtain reliable parameters of the optical potential for antineutrons. The beam of $\bar{n}$ with the energy in the range of 10^{-7} eV will hardly be accessible in the near future. Therefore, work has to be continued along the two lines mentioned above: to deduce the parameters of the optical potential the level shifts in antiprotonic atoms and to construct reliable optical

models that can be confronted with the available experimental data on $\bar{n}$–nuclear interaction at higher energies. In a forthcoming publication, we plan to present numerical calculation of the time evolution of a wave packet into standing waves and to discuss some features of n–$\bar{n}$ oscillations in the eigenfunction basis, which were not discussed in [4]. Another task is to perform calculation for the specific geometry of the trap and a realistic spectrum of the neutron beam. This requires an input corresponding to a specific experimental setting.

ACKNOWLEDGMENTS

The authors are grateful to Yu.A. Kamyshkov for interest in this work and discussions. Useful remarks and suggestions from L.N. Bogdanova, F.S. Dzheparov, A.I. Frank, V.D. Mur, V.S. Popov, and G.V. Danilian are gratefully acknowledged.

One of the authors (B.O.K.) expresses his gratitude for financial support to V.A. Novikov, L.B. Okun, INTAS (grant no. 110), and K.A. Ter-Martirosian scientific school (grant no. NSchool-1774).

REFERENCES

1. V. A. Kuz'min, Pis'ma Zh. Éksp. Teor. Fiz. **13**, 335 (1970) [JETP Lett. **13**, 238 (1970)].
2. R. N. Mohapatra and R. E. Marshak, Phys. Rev. Lett. **44**, 1316 (1980).
3. K. G. Chetyrkin, M. V. Kazarnovsky, V. A. Kuzmin, and M. E. Shaposhnikov, Phys. Lett. B **99B**, 358 (1981).
4. S. Marsh and K. W. McVoy, Phys. Rev. D **28**, 2793 (1983).
5. M. Baldo Ceolin, in *Festschrift for Val Telegdi*, Ed. by K. Winter (Elsevier, Amsterdam, 1988), p. 17.
6. R. Golub and H. Yoshiki, Nucl. Phys. A **501**, 869 (1989).
7. H. Yoshiki and R. Golub, Nucl. Phys. A **536**, 648 (1992).
8. B. O. Kerbikov, Phys. At. Nucl. **66**, 2178 (2003).
9. Yu. G. Abov, F. S. Dzheparov, and L. B. Okun, Pis'ma Zh. Éksp. Teor. Fiz. **39**, 493 (1984) [JETP Lett. **39**, 599 (1984)].
10. M. Baldo-Ceolin *et al.*, Z. Phys. C **63**, 409 (1994).
11. I. M. Frank, Usp. Fiz. Nauk **161**, 109 (1991) [Sov. Phys. Usp. **34**, 988 (1991)].
12. V. K. Ignatovich, *The Physics of Ultracold Neutrons* (Nauka, Moscow, 1986; Clarendon Press, Oxford, 1990).
13. Ya. B. Zel'dovich, Zh. Éksp. Teor. Fiz. **36**, 1952 (1959) [Sov. Phys. JETP **9**, 1389 (1959)].
14. M. Lax, Rev. Mod. Phys. **23**, 287 (1951); Phys. Rev. **85**, 621 (1952).
15. J. Hufber and B. Z. Kopelovich, hep-ph/9807210.
16. Ye. S. Golubeva and L. A. Kondratyuk, Nucl. Phys. B (Proc. Suppl.) **56**, 103 (1997).
17. V. M. Galitsky, B. M. Karnakov, and V. I. Kogan, *Problems in Quantum Mechanics* (Nauka, Moscow, 1992).
18. C. J. Goebel and K. W. McVoy, Ann. Phys. (N.Y.) **37**, 62 (1966).
19. V. K. Ignatovich, Phys. Rev. D **67**, 016004 (2002).

Journal of Experimental and Theoretical Physics, Vol. 98, No. 3, 2004, pp. 427–437.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 125, No. 3, 2004, pp. 486–498.

ATOMS, SPECTRA, RADIATION

Energy-Flow Patterns and Bifurcations of Two-Dimensional Cavity Solitons

N. N. Rozanov, S. V. Fedorov, and A. N. Shatsev

Research Institute for Laser Physics, Vavilov State Optical Institute, St. Petersburg, 199034 Russia
e-mail: rosanov@ilph.spb.su

Received May 27, 2003

Abstract—An analysis of two-dimensional spatial optical solitons in a large-aperture class A laser with a saturable absorber is developed. New types of rotating asymmetric solitons are found by computing the governing equation. The existence of weakly and strongly coupled solitons is demonstrated. Essential distinctions between them manifest themselves in the pattern of energy flows. A strongly coupled state evolves with time from an initial superposition of the fields of two solitons via successive bifurcations (topological changes in the energy-flow pattern). © *2004 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

Optical solitons are blobs of light whose linear (diffractive or dispersive) spread is balanced by nonlinear compression. There exist optical solitons of two types. One is the conservative soliton developing and propagating in a transparent medium with negligible losses [1, 2]. The basic characteristics of conservative solitons have continuous spectra. In other words, the corresponding peak intensities or blob widths can vary continuously (within certain limits). The other type is the dissipative soliton (also called *autosoliton*) [3]. Originally, dissipative optical solitons were predicted and analyzed theoretically in large-aperture nonlinear interferometers [4, 5] and lasers with saturable absorbers [6]. By virtue of an additional requirement of energy balance (not imposed in the case of conservative solitons), the basic characteristics of dissipative solitons have discrete spectra. This implies substantial difference in properties between the two types of solitons. Dissipative optical solitons are characterized by an excitation threshold. In conjunction with the discreteness of their spectra, this property enhances their potential utility for data-processing technologies. Progress in experiment and application was stimulated by the introduction of semiconductor microcavities [7, 8]. The current status of theory and experiment in studies of optical solitons was discussed in reviews [9, 10] and a monograph [11].

One-, two-, and three-dimensional dissipative optical solitons can be implemented in certain lasers with and without feedback [11, 12]. In this paper, we consider the spatial solitons that arise in a large-aperture laser with nonlinear (saturable) gain and absorption. Since the longitudinal variation of the electric-field envelope is slow, we can use a paraxial equation averaged over this direction [13], i.e., solitons of this kind are essentially two-dimensional. We restrict our analysis to media with fast optical nonlinearity, whose response is determined by instantaneous values of the electric field envelope. In other words, we neglect the effects due to finite relaxation times (examined in [11, 14–16]). Under this assumption, field dynamics are governed by the complex Ginzburg–Landau equation for the envelope of electric field. The governing equation and its key properties are considered in Section 2. Numerical solution of the governing equation reveals a number of interesting structures that can be interpreted as strongly coupled states of cavity solitons (see [11, 17] and references cited therein). In this paper, we also compute several new structures of this kind. However, the present analysis is focused on the elucidation of the "internal structure" of cavity solitons, which is most obviously manifested in the pattern of energy flow. We apply some well-known methods of the theory of nonlinear oscillations to analyze the pattern of Poynting-vector streamlines (Section 3). The relatively simple patterns corresponding to single solitons with axially symmetric intensity distributions are discussed in Section 4. In Section 5, we consider a pair of interacting solitons and expose qualitative (topological) difference in energy-flow pattern between widely separated (weakly interacting) and closely spaced (strongly interacting) solitons. This analysis indicates that the pattern changes via bifurcations as the distance between the solitons decreases. The main results are summarized in the Conclusions Section.

2. MODEL OF A LASER AND THE BASIC EQUATIONS

Consider a large-aperture laser with saturable gain and absorption in the optical cavity between plane parallel mirrors. In the case of fast optical nonlinearity (for a class A laser), the paraxial equation for the envelope $E(\mathbf{r}_\perp, t)$ averaged over the longitudinal coordinate z (in the mean-field approximation valid when the electric-

field envelope changes weakly over the cavity length) is the generalized complex Ginzburg–Landau equation (see [11])

$$\frac{\partial E}{\partial t} = (i + d)\Delta_\perp E + f(|E|^2)E. \quad (2.1)$$

We use the following dimensionless notation here: t is the time measured in units of field decay time in an empty cavity, d is the effective diffusion coefficient characterizing a weakly dispersive medium ($0 < d \ll 1$), and $\Delta_\perp = \partial^2/\partial x^2 + \partial^2/\partial y^2$ is the transverse Laplace operator. The transverse coordinates x and y are measured in units of Fresnel zone width $w_F = [L_c/2k(1 - |R|)]^{1/2}$, where L_c is the cavity length, k is the wavenumber in a linear medium, and R is the product of the amplitude reflectivities of the cavity mirrors. The function $f(|E|^2)$ characterizes gain and absorption saturation and includes constant (nonresonant) losses. Without specifying this function, we note here that stable localized structures (with field amplitude rapidly decreasing toward periphery) can exist only if

$$\mathrm{Re}\, f_0 = \mathrm{Re}\, f(0) < 0. \quad (2.2)$$

Otherwise, the peripheral structure would be unstable with respect to small perturbations.

Equation (2.1) is invariant under phase shift of the field

$$E \longrightarrow E\exp(i\Phi_0), \quad (2.3)$$

shift in the transverse coordinates

$$E(x, y, t) \longrightarrow E(x + X_0, y + Y_0, t) \quad (2.4)$$

(Φ_0, X_0, and Y_0 are constants), and inversion of either transverse coordinate, such as

$$E(x, y, t) \longrightarrow E(x, -y, t). \quad (2.5)$$

Furthermore, Eq. (2.1) entails the following integral balance energy relation valid for localized structures [12]:

$$\frac{d}{dt}\int |E|^2 d\mathbf{r}_\perp = 2W'(E), \quad (2.6)$$

where the functional $W'(E)$ is defined as

$$W'(E) = \int [|E|^2 \mathrm{Re}\, f(|E|^2) - d|\nabla_\perp E|^2] d\mathbf{r}_\perp. \quad (2.7)$$

For steady structures (including those moving or rotating as a whole), it holds that

$$W'(E) = 0. \quad (2.8)$$

In the case of a negligible frequency detuning, f is a real function ($\mathrm{Im} f = 0$). In the calculations presented below, we used the function f corresponding to a two-level medium with saturable gain and absorption:

$$f(|E|^2) = -1 + \frac{g_0}{1 + |E|^2} - \frac{a_0}{1 + b|E|^2}, \quad (2.9)$$

where g_0 and a_0 are the linear gain and absorption coefficients, respectively; b is the ratio of gain and absorption saturation intensities; and nonresonant losses are represented by the first term on the right-hand side (equal to -1 on the time scale used here).

3. FLOWS OF RADIANT ENERGY

In the paraxial approximation employed here, the Poynting vector **S** averaged over the period of an electromagnetic wave with constant (e.g., almost linear) polarization is related to the envelope E, the real amplitude $A = |E|$, and phase $\Psi = \arg E$ as follows [18, 19]:

$$\mathbf{S} = A^2\nabla\Psi = \mathrm{Im}(E^*\nabla E). \quad (3.1)$$

In accordance with the laser model, we treat the z axis as the predominant wave propagation direction, and governing equation (2.1) involves only the transverse coordinates $\mathbf{r}_\perp = (x, y)$. Therefore, the vector $\mathbf{S}_\perp = (S_x, S_y)$ at a time t can be expressed as

$$\mathbf{S}_\perp = A^2\nabla_\perp\psi = \mathrm{Im}(E^*\nabla_\perp E). \quad (3.2)$$

The corresponding streamlines (curves with tangents parallel to $\mathbf{S}_\perp$ at every point) are conveniently parameterized by equations written in terms of a function τ of the arclength (cf. the ray equation in geometrical optics):

$$\frac{dx}{d\tau} = S_x(x, y), \quad \frac{dy}{d\tau} = S_y(x, y). \quad (3.3)$$

Note that these streamlines would be almost everywhere described by Eq. (3.3) with the Poynting vector replaced by the phase gradient. However, the Poynting vector is more suitable in most cases. Indeed, whereas the phase is not defined at the points of screw wavefront dislocations (where the field amplitude vanishes), the Poynting vector vanishes at these points. Note also that both energy flow and Poynting vector have well-defined physical meaning in the paraxial approximation. The transverse distribution of the Poynting vector combined with the transverse intensity distribution provides an unambiguous description of the field (up to an insignificant constant).

Equations similar to (3.3) have been analyzed in detail in the theory of nonlinear oscillations [20, 21]. The degenerate case corresponds to conservative systems almost everywhere satisfying the relation $\mathrm{div}_\perp \mathbf{S}_\perp = 0$. Conservative systems do not have isolated closed orbits and are not robust. The phase portrait of

such a system consists of "cells" occupied by trajectories of similar type that can be altered by an arbitrarily small change in $S_x(x, y)$ and $S_y(x, y)$ [20, 21]. Examples of Poynting-vector streamlines for conservative optical systems can be found in [22].

Lasers are dissipative systems ($\mathrm{div}_\perp \mathbf{S}_\perp \neq 0$) in which the field exchanges energy with active and absorbing media. To divide a phase portrait into cells of similar behavior of the trajectories described by $x(\tau)$ and $y(\tau)$, one should find the singular (fixed) points (x_0, y_0) at which $S_x(x_0, y_0) = 0$ and $S_y(x_0, y_0) = 0$. According to (3.1), the singular points are of two types, since the Poynting vector vanishes with field amplitude or phase gradient. In the former case, the expansion of the complex field envelope in terms of small deviations of x and y from x_0 and y_0 about a singular point starts from linear terms:

$$E(x, y) = E_{(x)}x + E_{(y)}y + \frac{1}{2}E_{(xx)}x^2 + E_{(yy)}xy + \frac{1}{2}E_{(yy)}y^2 + \ldots, \tag{3.4}$$

where the derivatives, $E_{(x)} = \partial E/\partial x|_{x=y=0}$, etc., are calculated at the singular point. Then,

$$S_x = qy, \quad S_y = -qx, \tag{3.5}$$

where $q = \mathrm{Im}(E_{(x)}E^*_{(y)})$, and Eqs. (3.3) become

$$\frac{dx}{d\tau} = qy, \quad \frac{dy}{d\tau} = -qx. \tag{3.6}$$

Solutions to these equations correspond to trajectories $x^2 + y^2 = R^2$ in the phase plane xy. In this case, the point is a center (i.e., not a robust one) and its type may change when higher order terms of expansion (3.4) are taken into account. The analysis below shows that the point is actually a focus in the case of a cavity soliton with an axially symmetric intensity distribution.

The expansion of the field about a singular point of the other type contains a constant term $E_0 \neq 0$:

$$E(x, y) = E_0 + E_{(x)}x + E_{(y)}y + \frac{1}{2}E_{(xx)}x^2 + E_{(xy)}xy + \frac{1}{2}E_{(yy)}y^2 + \ldots \tag{3.7}$$

Here, the Poynting vector vanishes at the singular point ($x = y = 0$) if $\mathrm{Im}(E_0^* E_{(x)}) = \mathrm{Im}(E_0^* E_{(y)}) = 0$, and the energy-flow streamlines are described by the equations

$$\frac{dx}{d\tau} = \mathrm{Im}(E_0^* E_{(xx)})x + (q + p)y,$$
$$\frac{dy}{d\tau} = (p - q)x + \mathrm{Im}(E_0^* E_{(yy)})y, \tag{3.8}$$

where $p = \mathrm{Im}(E_0^* E_{(xy)})$. The type of a singular point is determined by the roots of the quadratic equation

$$\lambda^2 + \sigma\lambda + \Delta = 0, \tag{3.9}$$

with

$$\sigma = -\mathrm{Im}(E_0^* E_{(xx)}) - \mathrm{Im}(E_0^* E_{(yy)}),$$
$$\Delta = \mathrm{Im}(E_0^* E_{(xx)})\mathrm{Im}(E_0^* E_{(yy)}) - p^2 + q^2. \tag{3.10}$$

When $\sigma \neq 0$ and $\Delta \neq 0$, the singular point is robust (node, focus, or saddle point). In the general case, to divide a phase portrait into cells, one must know not only the singular points and their types, but also nonlocal elements (periodic orbits and saddle separatrices) [20, 21].

Energy-flow streamlines change with system parameters, initial conditions, or time. Of particular interest are bifurcations, i.e., topological changes in the partition of the phase portrait into cells. All types of bifurcations admitted by Eqs. (3.3) have been well studied [20, 21], which facilitates analysis of energy-flow patterns in two-dimensional solitons.

4. SOLITONS WITH AXIALLY SYMMETRIC INTENSITY DISTRIBUTIONS

For a steady soliton, the time dependence of the envelope has the form $\exp(-i\nu t)$, where the frequency shift ν is the eigenvalue of the problem. Solitons with axially symmetric intensity distributions should be considered in polar coordinates r and φ ($x = r\cos\varphi$, $y = r\sin\varphi$). Writing

$$E = F(r)\exp(im\varphi)\exp(-i\nu t), \tag{4.1}$$

where the integer $m = 0, \pm1, \pm2, \ldots$ is a topological index, we obtain an equation for the complex radial function $F(r)$:

$$\frac{d^2F}{dr^2} + \frac{1}{r}\frac{dF}{dr} = \frac{m^2}{r^2}F + \frac{1}{i + d}[i\nu + f(|F|^2)]F = 0. \tag{4.2}$$

The "instantaneous" phase $\Psi = \arg E$ is related to the radial phase $\Psi_0(r) = \arg F$:

$$\Psi(r, \varphi) = \Psi_0(r) + m\varphi. \tag{4.3}$$

It is assumed that time t is fixed and $\Psi_0(0) = 0$ since the constant component of the phase is of no importance. When the field amplitude is small, Eq. (4.2) reduces to a linear equation:

$$\frac{d^2F}{dr^2} + \frac{1}{r}\frac{dF}{dr} - \frac{m^2}{r^2}F - p^2F = 0, \tag{4.4}$$

where

$$p^2 = -\frac{i\nu + f_0}{i + d}. \tag{4.5}$$

Without loss of generality, we can assume that $\mathrm{Re}\,p > 0$. When $m = 0$, we can use (4.5) with $f_0 \longrightarrow f(|F_0|^2)$, where $|F_0|^2$ is the intensity at the center of a fundamental soliton. The solution to (4.4) can be expressed in terms of cylinder functions. However, only their asymptotics can be used at $r \longrightarrow \infty$ (at the periphery of a soliton) and $r \longrightarrow 0$ (when $m \neq 0$, since the field vanishes at the center of the soliton in this case). As $r \longrightarrow \infty$, the complex amplitude F, the real amplitude $A = |F|$, and the radial phase Ψ_0 are

$$\begin{gathered} F \approx F_\infty \exp(-pr), \quad A \approx A_\infty \exp(-p'r), \\ \Psi_0 = \mathrm{const} - p''r, \end{gathered} \tag{4.6}$$

where $p' = \mathrm{Re}\,p > 0$ and $p'' = \mathrm{Im}\,p$.

As $r \longrightarrow 0$, the power series expansion of the solution to (4.4) with $m \neq 0$ or the solution to (4.2) with $m = 0$ has the form

$$F = F_0 r^{|m|}(1 + F_2 r^2 + \ldots). \tag{4.7}$$

Substituting (4.7) into (4.4), we find

$$F_2 = \frac{p^2}{4(|m| + 1)}, \tag{4.8}$$

which yields the lowest order term in the radial phase

$$\begin{aligned} \Psi_0(r) &= \frac{r^2}{4(|m| + 1)}\mathrm{Im}\,p^2 \\ &= -\frac{r^2}{4(|m| + 1)}\mathrm{Im}\frac{f_0 + i\nu}{i + d}. \end{aligned} \tag{4.9}$$

Note that the asymptotic behavior of the radial phase is determined by the same complex quantity p both as $r \longrightarrow \infty$ and as $r \longrightarrow 0$.

In polar coordinates (with basis vectors $\mathbf{e}_r$ and $\mathbf{e}_\varphi$), Poynting vector (3.2) for solitons of this type has the form

$$\mathbf{S}_\perp = A^2(r)\left(\frac{d\Psi_0}{dr}\mathbf{e}_r + \frac{m}{r}\mathbf{e}_\varphi\right). \tag{4.10}$$

Accordingly, Eqs. (3.3) for energy-flow streamlines are rewritten as

$$\frac{dr}{d\tau} = A^2\frac{d\Psi_0}{dr}, \quad \frac{d\varphi}{d\tau} = \frac{m}{r^2}A^2, \tag{4.11}$$

and trajectories are described by the equations

$$\frac{d\varphi}{dr} = \frac{m}{r^2 d\Psi_0/dr}, \quad \varphi - \varphi_0 = m\int_{r_0}^{r}\frac{dr}{r^2 d\Psi_0/dr}. \tag{4.12}$$

The first equation in (4.11) is strictly radial (does not contain φ). As $\tau \longrightarrow \pm\infty$, the radius $r(\tau)$ tends to a constant R_0, where R_0 is a root of the equation

$$\frac{d\Psi_0}{dr}(R_0) = 0,$$

whose roots include $R_0 = 0$. When $m = 0$, the trajectories are radial line segments with endpoints on circles of radius R_0. When $m \neq 0$, these circles are limit cycles, and the state point moves along them with a constant angular velocity $\Omega = mA^2(R_0)/R_0^2$. We should note that stability analysis of energy-flow patterns developed here has nothing to do with time evolution, because a fixed point in time is considered. However, one can consider "stability" with respect to increase in τ (hence the quotation marks). Note that the singular points (nodes or foci) and limit cycles that are "unstable" in the limit of $r \longrightarrow \infty$ become "stable" as the sign of τ reverses. This property can be used in calculations, in particular, to find "unstable" limit cycles.

Since Eq. (4.2) for an axially symmetric localized distribution is a complex one, it is equivalent to a nonlinear system of four first-order ordinary differential equations. The equations for the real amplitude A and the radial phase Ψ_0 entailed by (4.2) are

$$\begin{aligned} &\frac{d^2A}{dr^2} + \frac{1}{r}\frac{dA}{dr} - \frac{m^2}{r^2}A - A\left(\frac{d\Psi_0}{dr}\right)^2 \\ &\quad + \mathrm{Re}\frac{i\nu + f(A^2)}{i + d}A = 0, \\ &\frac{d^2\Psi_0}{dr^2} + 2\frac{1}{A}\frac{dA}{dr}\frac{d\Psi_0}{dr} + \frac{1}{r}\frac{d\Psi_0}{dr} \\ &\quad + \mathrm{Im}\frac{i\nu + f(A^2)}{i + d} = 0. \end{aligned} \tag{4.13}$$

Since Eqs. (4.13) do not contain the phase, the order of the system can be reduced to three. The equivalent sys-

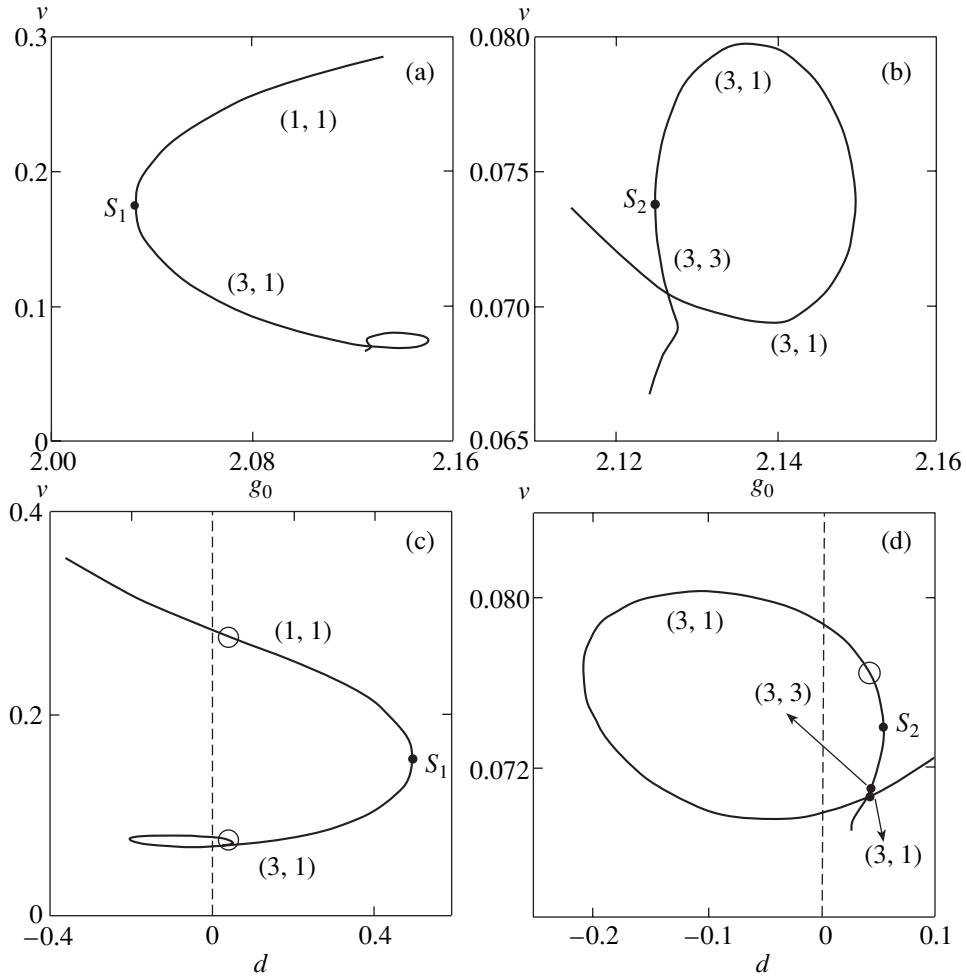

Fig. 1. The eigenvalue ν for solitons with topological charge $m = 1$ versus (a, b) gain g_0 for $d = 0$ and (c, d) diffusion coefficient d for $g_0 = 2.13$. The second loop of the spiral is shown in more detail in panels (b) and (d) as compared to (a) and (c). The numbers at the spiral illustrate the change in (N_Q, N_K) across the points S_N with vertical tangents. Unstable solitons are marked with circles. The dots with arrows near the self-intersection in panel (d) ($d = 0.04$) correspond to the fundamental and excited solitons depicted in Figs. 2c and 2d, respectively.

tem of equations for $A(r)$, $Q(r) = d\Psi_0/dr$, and $K(r) = (1/A)dA/dr$ is

$$\frac{dA}{dr} - KA = 0, \quad \frac{dQ}{dr} + 2KQ + \frac{Q}{r} + \mathrm{Im}\frac{i\nu + f(A^2)}{i + d} = 0, \tag{4.14}$$

$$\frac{dK}{dr} + K^2 - Q^2 + \frac{K - m^2}{r} + \mathrm{Re}\frac{i\nu + f(A^2)}{i + d} = 0.$$

As $r \longrightarrow 0$ and $r \longrightarrow \infty$, the asymptotics of these functions can easily be determined from expressions (4.6) and (4.7). Specifically, as $r \longrightarrow \infty$, we find that $K \longrightarrow -p' < 0$ and $Q \longrightarrow -p'' > 0$. As $r \longrightarrow 0$, both $Q(r)$ and $K(r) - |m|/r$ must vanish simultaneously. These asymptotics can be used to calculate characteristics of the solitons with axially symmetric intensity distributions considered here.

The calculations presented below were performed for $a_0 = 2$ and $b = 10$. When all parameters are held constant, every solution to Eq. (4.14) with appropriate asymptotic behavior corresponds to an eigenvalue ν belonging to a discrete set. The dependence of ν on a control parameter, such as g_0 or d, is graphically represented by a self-intersecting spiral with several loops (see Fig. 1). Different branches of the spiral are characterized by different numbers of zeros N_Q and N_K of the functions $Q(r)$ and $K(r)$ at $r > 0$. The numbers N_Q and N_K change by 1 (when $m = 0$) or 2 (when m is any integer) at the points S_N with vertical tangents in Fig. 1 ($N = 1, 2$).

If $m = 0$, then $N_Q = N_K = 0$ for the outer loop of the spiral (above S_1) and the corresponding localized struc-

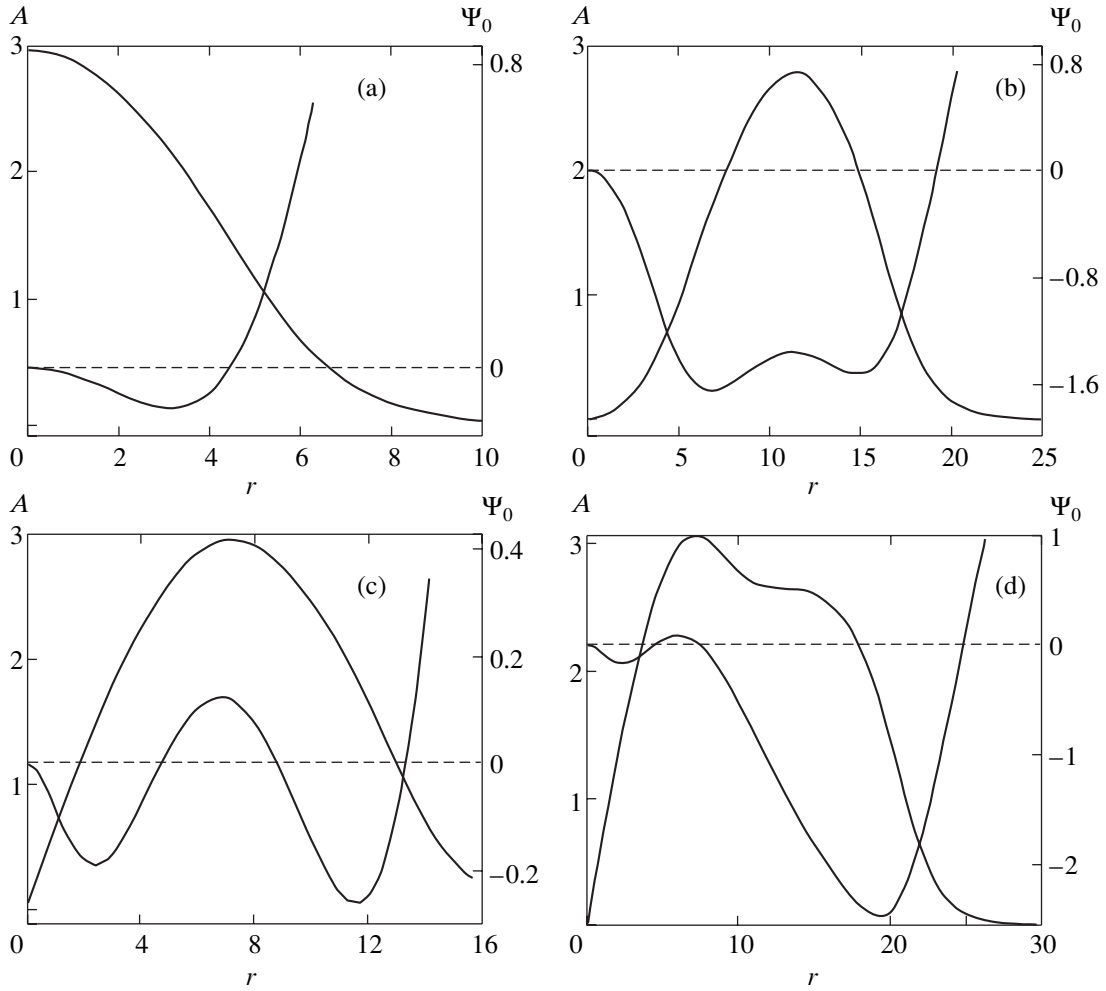

Fig. 2. Radial profiles of amplitude $A(r)$ and "radial" phase $\Psi_0(r)$ for stable fundamental solitons with (a) $m = 0$ ($g_0 = 2.11$, $d = 0.01$, $\nu = 0.12036$) and (c) $m = 1$ ($g_0 = 2.13$, $d = 0.04$, $\nu = 0.070658$) and for excited solitons with (d) $m = 1$ ($g_0 = 2.13$, $d = 0.04$, $\nu = 0.070998$) and (b) $m = 2$ ($g_0 = 2.10$, $d = 0.06$, $\nu = 0.0809$).

tures are unstable. Here, stability (without quotation marks) is interpreted in terms of time evolution and is determined by applying the conventional linear analysis [11]. There exist stable solitons corresponding to the segments of the next loop where $N_Q = 1$ and $N_K = 0$. If $m = 1$ or 2, then there exist no stable solitons corresponding to the outer loop, where $N_Q = N_K = 1$. Stable solitons correspond to the segments of the next loops where $N_Q = 3$ and $N_K = 1$ (fundamental soliton) and $N_Q = 3$ and $N_K = 3$ (excited soliton). The dots with arrows in Fig. 1d represent two close values of ν corresponding to the same value of g_0 and lying on different loops of the spiral near its self-intersection. These points are associated with the fundamental and excited solitons with the radial amplitude and phase distributions illustrated by Figs. 2c and 2d, respectively.

The analysis that follows is focused on stable solitons. When $m = 0$, the field amplitude is a monotonically decreasing function of radius. When $m \neq 0$, the radial profile of the amplitude has a single maximum and the radial phase is an oscillating function of radius with number of oscillations depending on $|m|$ (see Fig. 2). Note that the fundamental and excited states of a stable soliton (which correspond to different loops of spirals similar to those in Fig. 1) differ only quantitatively (cf. Figs. 2b and 2c).

According to the foregoing analysis, the pattern of energy flows in the phase plane is as shown in Fig. 3. The central point, $r = 0$, is a "stable" node when $m = 0$ and a focus when $m \neq 0$. In the latter case, this point is encompassed by an odd number of limit cycles, i.e., circles of radius R_0 defined by the condition

$$Q(R_0) = \frac{d\Psi_0}{dr}(R_0) = 0$$

(see above). The cycles have alternating "stability" properties: the cycle closest to the central point is the "unstable" boundary of the basin of attraction of the

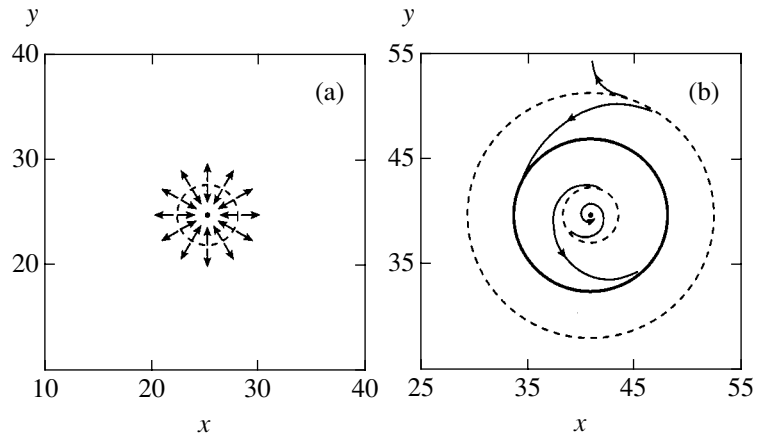

Fig. 3. Phase portraits of single stable solitons with (a) $m = 0$ and (b) $m = 1$. The fixed points at $r = 0$ are node and focus, respectively. Cells occupied by trajectories with different behavior are separated by thick solid ("stable") and dashed ("unstable") limit cycles ($m \neq 0$) or a set of fixed points (in the degenerate case of $m = 0$). The thin curves with arrows are Poynting-vector streamlines. The phase portraits with $m = 1$ and $m = 2$ are topologically equivalent.

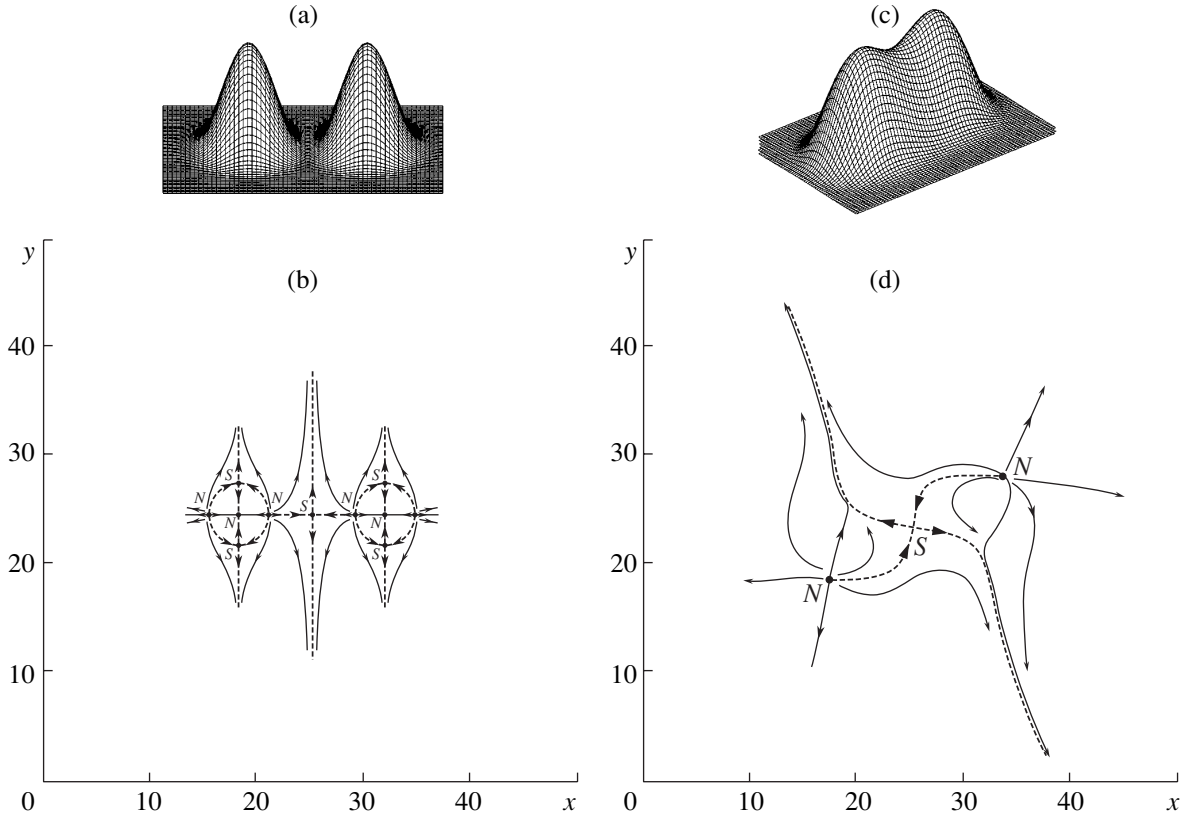

Fig. 4. Bound states of solitons with $m_1 = m_2 = 0$, $g_0 = 2.11$, and $d = 0.06$: (a, c) instantaneous transverse intensity distributions $I(x, y)$ and (b, d) phase portraits for (a, b) weakly and (c, d) strongly coupled solitons. Dashed curves are separatrices associated with saddle points S. Arrows indicate the direction of the Poynting vector.

focus, the next cycle is a "stable" one that attracts trajectories as $\tau \longrightarrow \infty$, and the outermost cycle is "unstable" since the trajectories lying outside it tend to infinity as $r \longrightarrow \infty$. Figures 1–3 illustrate cases when $m \geq 0$. Those corresponding to $m < 0$ can readily be obtained by coordinate inversion, as in (2.5). When $m = 0$, the system is not robust (degenerate). In this case, the circle encompassing the "stable" node is a set of individual fixed points rather than a limit cycle. A small perturbation, such as a distant soliton, gives rise to two pairs of "unstable" nodes N and saddle points S. The four circular arcs that connect them are separatrices incident to

the saddle points (see Fig. 4b). This pattern is robust (not degenerate).

5. WEAK AND STRONG INTERACTION BETWEEN CAVITY SOLITONS

Suppose that the distance L between the centers of two solitons is much greater than the size of a single soliton, which is determined by the radius of the outermost limit cycle. Then, the distortion of their structure caused by the overlap of soliton tails is weak and the instantaneous field can be expressed as

$$\begin{aligned} E &= A(r_1)\exp[i\Psi(r_1) + im_1\varphi_1] \\ &+ A(r_2)\exp[i\Psi_0(r_2) + im_2\varphi_2 + i\vartheta]. \end{aligned} \tag{5.1}$$

For simplicity, we consider solitons whose topological charges are equal in absolute value ($|m_1| = |m_2|$). This implies that the solitons are associated with equal eigenvalues ν and their amplitudes A and radial phases Ψ_0 are similar functions of radius. However, the arguments of these functions are different: r_n and φ_n ($n = 1, 2$) are polar coordinates with origin at the center of the nth soliton:

$$\begin{aligned} r_n^2 &= \left(x \pm \frac{L}{2}\right)^2 + y^2, \quad \cos\varphi_n = \frac{x \pm L/2}{r_n}, \\ \sin\varphi_n &= \frac{y}{r_n}. \end{aligned} \tag{5.2}$$

The constant phase difference ϑ affects the interaction between the solitons [11].

Since the degree of overlap is low, the energy flows near the solitons' centers change insignificantly. Accordingly, the pattern of Poynting-vector streamlines determined by Eqs. (3.3) preserves all closed orbits that encompass the fixed points representing individual solitons (stable and unstable limit cycles, as well as closed orbits consisting of individual separatrices). Additional fixed (typically, saddle) points may arise in the intermediate regions between them. The separatrices emanating from these points partition the plane into cells occupied by trajectories moving away from individual solitons. The intensity distribution and energy-flow pattern corresponding to a steady pair of relatively weakly coupled solitons with zero topological charges are depicted in Figs. 4a and 4b, respectively.

Bound state (5.1) with a constant phase difference ϑ and an arbitrary distance L between the solitons may vary with time. The constant distances $L \gg R_0$ corresponding to steady (stable or unstable) pairs of weakly interacting solitons can be found by substituting (5.1) into (2.8). As the distance between the solitons decreases to $L \leq R_0$ in the course of time, the interaction becomes stronger and the phase portrait becomes qualitatively (topologically) different from that in the case of $L \gg R_0$. Note that this change in internal soliton structure is due to bifurcations associated with variation of time (not control parameter). The corresponding field dynamics are found numerically by solving Eq. (2.1) with function (2.9) by a splitting method with the use of fast Fourier transform.

Figures 4c and 4d show the instantaneous intensity distribution and energy-flow pattern in a steady two-hump structure rotating with a constant angular velocity. This structure can be interpreted as a strongly coupled state of two fundamental cavity solitons (with zero topological charges). Note that the phase portrait in Fig. 4d contains only "unstable" fixed points. In other words, the two nodes that are "stable" in the case of weakly coupled solitons (see Fig. 4b) transform into "unstable" nodes in the case of strongly coupled solitons (see Fig. 4d). Another manifestation of strong coupling is the absence of "individual" closed orbits encompassing the nodes (characteristic of weak coupling), which explains the change in "stability" of the nodes. Almost every trajectory tends to infinity as $\tau \longrightarrow \infty$. The separatrices of the saddle point S lying between the nodes partition the plane into cells occupied by trajectories going to infinity along different directions. The bound solitons rotate as a whole without any dislocation of the wavefront. This demonstrates that field asymmetry is the key condition for rotation [11].

Now, consider the interaction between two cavity solitons with topological charges $m_1 = m_2 = 1$ (see Fig. 5). The corresponding strongly coupled state evolves in time from an initial superposition (5.1) of two independent fields with certain ϑ and L at $t = 0$. In Fig. 5a, the individual solitons that make up the pair are represented by the "stable" foci corresponding to wavefront dislocations. (The field amplitude vanishes at these points and remains finite at other fixed points.) The final steadily rotating structure of a strongly coupled state (see Fig. 5f) preserves only one of the three individual limit cycles associated with each of the two dislocations (the inner "unstable" one, see Fig. 3b). However, the system has two "common" limit cycles encompassing both foci, one of which is "stable" and the other is "unstable." The trajectories lying outside the outer limit cycle go to infinity, whereas those inside it either approach the "stable" common limit cycle or tend to the stable foci (they are partitioned by the separatrices of the saddle point S lying between the foci). The separatrices incident to the saddle point unwind off the individual "unstable" limit cycles, while the outgoing separatrices wind to the "stable" common limit cycle. The overall partition of the phase portrait into cells of steady-state trajectories characterized by different behavior is obvious from Fig. 5f. Actually, transition between the initial and final states involves a time sequence of bifurcations of the energy-flow pattern. Some of these are illustrated by Fig. 5. In particular, the transition between the phase portraits shown in Figs. 5a and 5b involves a bifurcation in which the outer two limit cycles encompassing the left focus coalesce and disappear. The transition from Fig. 5b to Fig. 5c can be associated with a bifurcation at the central saddle point

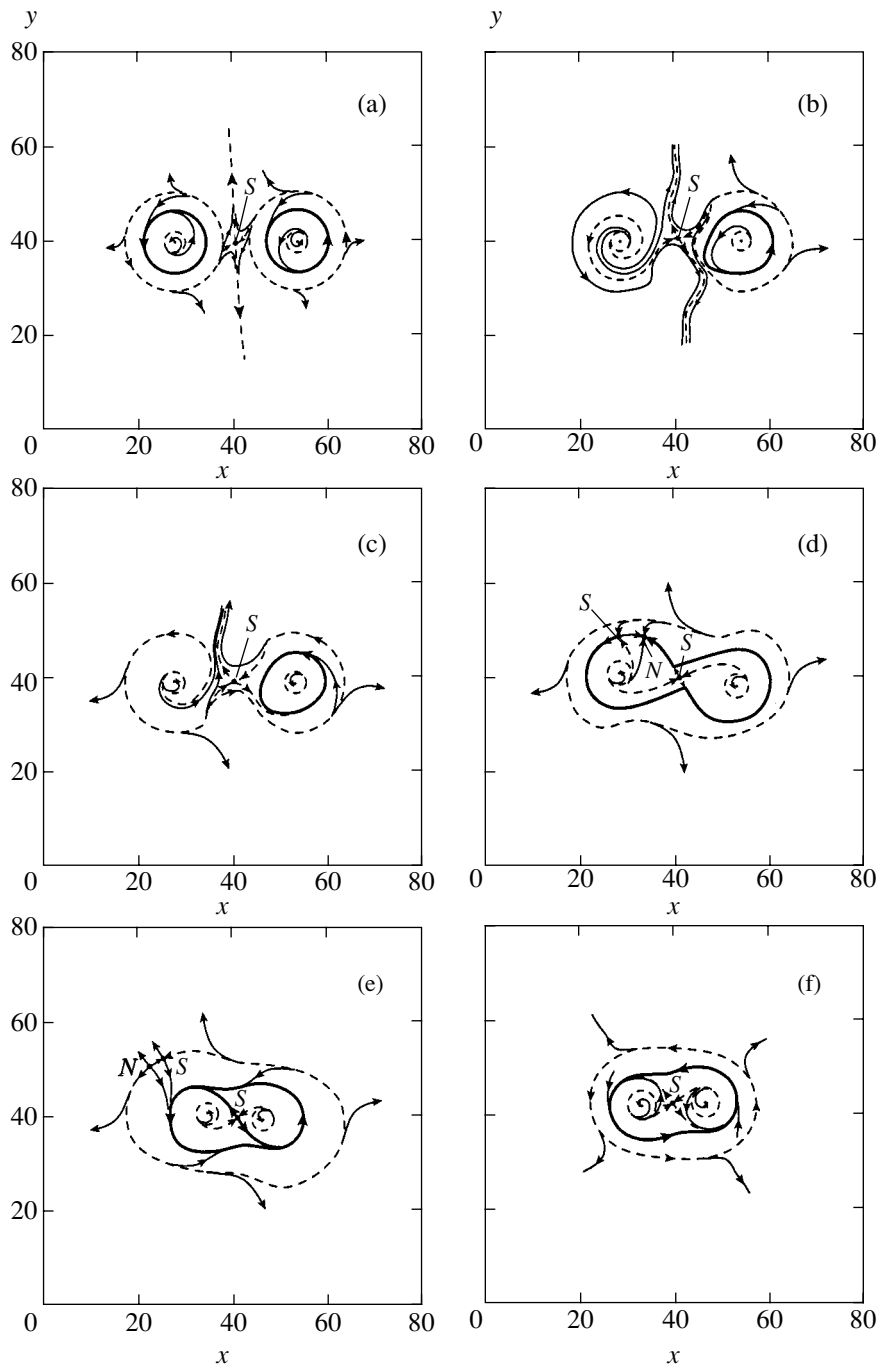

Fig. 5. Instantaneous phase portraits of energy flows for an evolving, strongly coupled, rotating pair of solitons with $m_1 = m_2 = 1$, $g_0 = 2.11$, and $d = 0.06$ at $t = 0$ (a), 5.6 (b), 6 (c), 30 (d), 59 (e), and ∞ (f).

S that gives rise to a homoclinic orbit encompassing the right outer limit cycle. Several bifurcations, including the last one (in transition from Fig. 5e to Fig. 5f), correspond to the appearance and coalescence of two fixed points, a node *N* and a noncentral saddle point *S* (see Figs. 5d and 5e).

The phase portrait of a strongly coupled pair of solitons with opposite topological charges ($m_1 = 1$, $m_2 = -1$) is more complicated. This rotating and slowly moving structure evolves from superposition (5.1) in a relatively long time interval. The wavefront dislocations (centers of individual solitons) are represented by the "stable" foci in Fig. 6. In the final state, both foci are encompassed by single "unstable" limit cycles (with opposite senses of rotation corresponding to opposite charges). The trajectories unwinding off these limit cycles include separatrices of the adjacent saddle

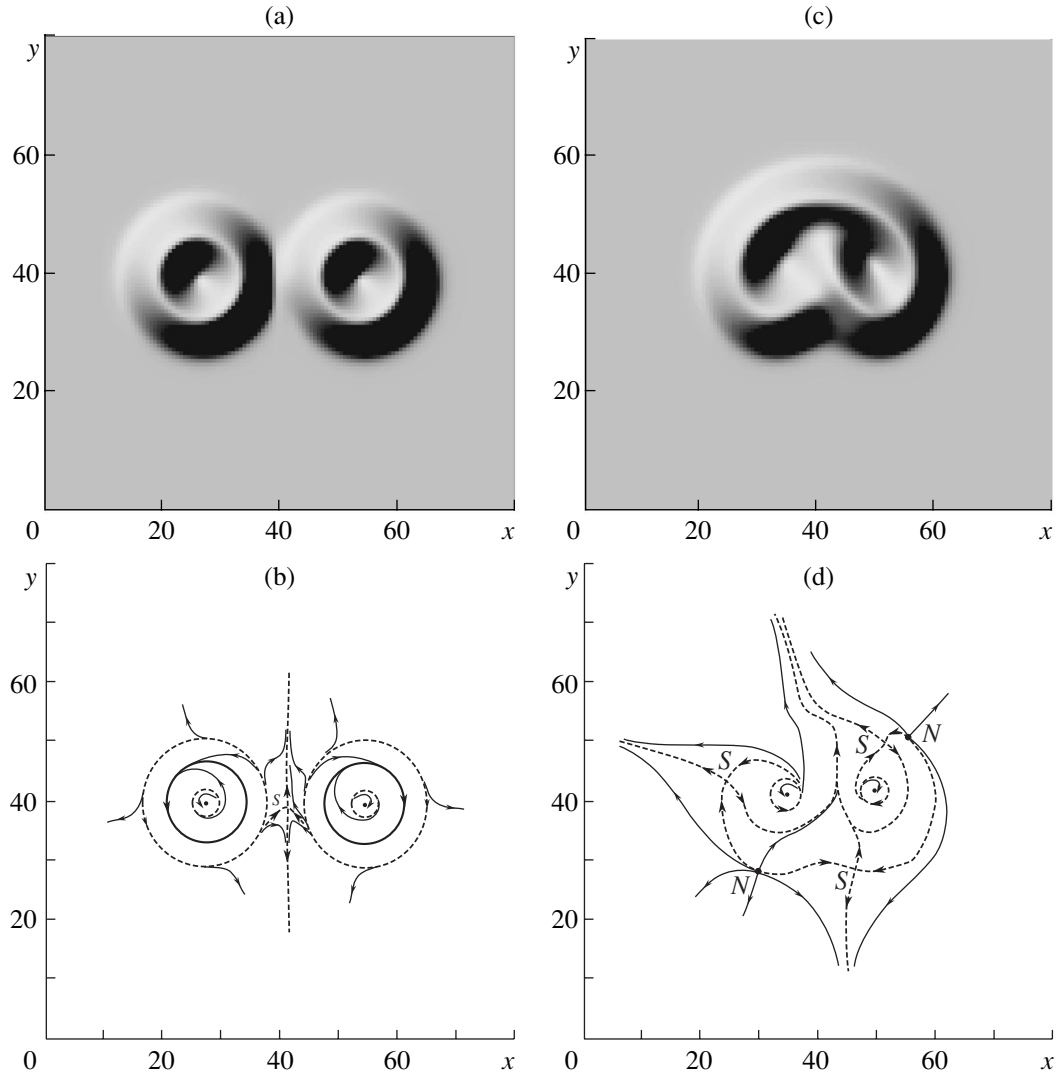

Fig. 6. Instantaneous (a, c) transverse distributions $I(x, y)$ and (b, c) phase portraits for pairs of solitons with $m_1 = 1$ and $m_2 = -1$ at (a, b) $t = 0$ and (c, d) $t = 10000$.

points. Each of the three saddle points emits a separatrix going to infinity. The separatrices emanating from different saddle points asymptotically converge. In addition to those from the two foci and three saddle points, there are two "unstable" nodes in the phase plane. The rays emanating from the nodes include saddle separatrices.

Interactions between cavity solitons can give rise to strongly coupled states of other types. In particular, trains of several solitons with $m = 1$ and two- and three-wave trains including both fundamental and excited solitons were computed by solving Eq. (2.1) (see Figs. 1 and 2 in [23]). All of these states are rotating, and their phase portraits are analogous to those discussed above.

6. CONCLUSIONS

The phase portraits of energy flows (Poynting vector) provide important information about the "internal structure" of cavity solitons. Even in the case of an individual soliton with an axially symmetric intensity distribution, the energy-flow plane is partitioned by concentric circles into several cells occupied by topologically equivalent trajectories (energy-flow streamlines). The radius of each circle corresponds to an extremum in the radial phase distribution. The phase portrait of a pair of weakly interacting solitons retains some "individual" features (fixed points and weakly distorted circular periodic orbits). The additional "collective" features typically include a saddle point with two incident separatrices coming from closed orbits and two separatrices going to infinity. The interaction is weak if the distance between the solitons is much greater than the diameter of the outermost "individual" closed orbit. As the distance between the solitons decreases, a sequence of bifurcations results in a qualitatively (topologically) different phase portrait corresponding to strongly interacting solitons. Individual elements (outer trajectories

in the first place) are replaced by "collective" ones, such as limit cycles encompassing both solitons. Both "stability" properties and number of fixed points may change. Even though the bifurcations alter the distribution of light intensity, the phase portraits of energy flows appear to be more informative.

The variety of strongly coupled soliton structures turns out to be very wide. They include trains and pairs characterized by strongly asymmetric fields. Analysis of the corresponding energy flows is also an important task. Note that analysis of energy diagrams (broadly analogous to Poynting-vector patterns) would also be useful in studies of dissipative solitons of different (not only optical) nature.

ACKNOWLEDGMENTS

We thank A.M. Kokushkin for assistance in preparing the graphs. This work was supported by the Russian Foundation for Basic Research, project no. 02-02-17242.

REFERENCES

1. V. E. Zakharov, S. V. Manakov, S. P. Novikov, and L. P. Pitaevskiĭ, *Theory of Solitons: the Inverse Scattering Method* (Nauka, Moscow, 1980; Consultants Bureau, New York, 1984).
2. E. M. Dianov, P. V. Mamyshev, and A. M. Prokhorov, Kvantovaya Élektron. (Moscow) **15**, 5 (1988).
3. B. S. Kerner and V. V. Osipov, *Autosolitons* (Nauka, Moscow, 1991; Kluwer, Dordrecht, 1994).
4. N. N. Rozanov and G. V. Khodova, Opt. Spektrosk. **65**, 1399 (1988) [Opt. Spectrosc. **65**, 828 (1988)].
5. N. N. Rosanov, A. V. Fedorov, and G. V. Khodova, Phys. Status Solidi B **150**, 545 (1988).
6. N. N. Rozanov and S. V. Fedorov, Opt. Spektrosk. **72**, 1394 (1992) [Opt. Spectrosc. **72**, 782 (1992)].
7. D. Michaelis, U. Peschel, and F. Lederer, Phys. Rev. A **56**, R3366 (1997).
8. M. Brambilla, L. A. Lugiato, F. Prati, *et al.*, Phys. Rev. Lett. **79**, 2042 (1997).
9. V. B. Taranenko and C. O. Weiss, IEEE J. Sel. Top. Quantum Electron. **8**, 488 (2002).
10. IEEE J. Quantum Electron. **39** (2) (2003).
11. N. N. Rosanov, *Spatial Hysteresis and Optical Patterns* (Springer, Berlin, 2002).
12. N. N. Rozanov, *Optical Bistability and Hysteresis in Distributed Nonlinear Systems* (Nauka, Moscow, 1997) [in Russian].
13. A. F. Suchkov, Zh. Éksp. Teor. Fiz. **49**, 1495 (1965) [Sov. Phys. JETP **22**, 1026 (1966)].
14. S. V. Fedorov, A. G. Vladimirov, G. V. Khodova, and N. N. Rosanov, Phys. Rev. E **61**, 5814 (2000).
15. N. N. Rozanov, S. V. Fedorov, and G. V. Khodova, Opt. Spektrosk. **88**, 869 (2000) [Opt. Spectrosc. **88**, 790 (2000)].
16. N. N. Rozanov, S. V. Fedorov, and A. N. Shatsev, Opt. Spektrosk. **91**, 252 (2001) [Opt. Spectrosc. **91**, 232 (2001)].
17. S. V. Fedorov, N. N. Rosanov, A. N. Shatsev, *et al.*, IEEE J. Quantum Electron. **39**, 197 (2003).
18. M. B. Vinogradova, O. V. Rudenko, and A. P. Sukhorukov, *The Theory of Waves*, 2nd ed. (Nauka, Moscow, 1990) [in Russian].
19. S. N. Vlasov and V. I. Talanov, *Wave Self-Focusing* (Inst. Prikl. Fiz. Ross. Akad. Nauk, Nizhni Novgorod, 1997) [in Russian].
20. A. A. Andronov, E. A. Leontovich, I. I. Gordon, and A. G. Maĭer, *Theory of Bifurcation of Dynamic Systems on Surface* (Nauka, Moscow, 1967) [in Russian].
21. N. V. Butenin, Yu. I. Neĭmark, and N. A. Fufaev, *Introduction to Theory of Nonlinear Oscillations* (Nauka, Moscow, 1976) [in Russian].
22. M. S. Soskin and M. V. Vasnetsov, Prog. Opt. **42**, 219 (2001).
23. N. N. Rozanov, S. V. Fedorov, and A. N. Shatsev, Opt. Spektrosk. **95**, 902 (2003) [Opt. Spectrosc. **95**, 843 (2003)].

Translated by A. Betev

Journal of Experimental and Theoretical Physics, Vol. 98, No. 3, 2004, pp. 438–454.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 125, No. 3, 2004, pp. 499–517.

ATOMS, SPECTRA, RADIATION

Kinetics of Atoms in the Field Produced by Elliptically Polarized Waves

O. N. Prudnikov, A. V. Taichenachev, A. M. Tumaĭkin, and V. I. Yudin*

Novosibirsk State University, Institute of Laser Physics, Siberian Division, Russian Academy of Sciences, Novosibirsk, 630090 Russia

**e-mail: ilf@admin.nsu.ru*

Received July 31, 2003

Abstract—The kinetics of atoms with degenerate energy levels in the field produced by elliptically polarized waves is considered in the semiclassical approximation. Analytic expressions for the force acting on an atom and for the diffusion coefficient in the momentum space are derived for the optical transition $J_g = 1/2 \longrightarrow J_e = 1/2$ in the slow atom approximation. These expressions are valid for an arbitrary one-dimensional configuration of the light field and for an arbitrary intensity. The peculiarities of the atomic kinetics are investigated in detail; these peculiarities are associated with ellipticity of light waves and are absent in particular configurations formed by circularly or linearly polarized waves, which were considered earlier. © *2004 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

At the initial stage (up to 1988), the mechanical action of resonance radiation on atoms and, in particular, the motion of atoms in a light field were studied comprehensively in the framework of the simplest model of a two-level atom [1, 2]. This description has made it possible to explain the physical mechanisms and the origin of forces acting on an atom in a light field and to predict the minimal temperature of laser-assisted cooling of atoms (Doppler limit) $k_B T_D \sim \hbar\gamma$, where γ is the natural width of an excited level. Experimental observation of temperatures below T_D [3, 4] stimulated further development of the theory. Detailed analysis proved that sub-Doppler cooling is associated with the degeneracy of the ground atomic state in the projection of the angular momentum. It should be noted that, in spite of a large number of publications concerning the theory of sub-Doppler cooling [5, 6], only a limited class of laser field configurations formed by linearly or circularly polarized waves was studied. Nevertheless, these simple examples have made it possible to single out two main mechanisms of friction leading to sub-Doppler cooling. In the case of counterpropagating linearly polarized waves, we are dealing with the Sisyphean mechanism of friction, associated with the action of induced light pressure, while in the case of counterpropagating circularly polarized waves (σ_+–σ_- configuration), the mechanism is of the orientation type and is associated with the action of spontaneous light pressure [7]. However, the analysis of these particular cases is not exhaustive in view of the nonlinear nature of interaction of atoms with the resonance field and gives no idea about the motion of particles in a field with a more general configuration. Thus, the problem on the kinetics of atoms in the field produced by elliptically polarized waves is interesting in itself. For example, as was proved in our earlier publication [8], ellipticity leads to qualitative differences in the kinetics as compared to the results obtained by using the two-level model even in the simple case of a uniformly polarized field.

Here, the motion of atoms in the field of a one-dimensional configuration formed by waves with arbitrary elliptical polarizations is studied theoretically. In the framework of the semiclassical approach, the kinetics of atoms is described by the Fokker–Planck equation. Analytic expressions are derived for the force acting on an atom as well as for the diffusion and friction coefficients in the slow atom approximation for the optical transition $J_g = 1/2 \longrightarrow J_e = 1/2$ (where J_g and J_e are total angular momenta of the ground (g) state and an excited (e) state). These expressions are analyzed for a number of specific field configurations; in each case, new terms associated with the ellipticity of waves are separated and analyzed. The physical mechanisms leading to these new contributions are interpreted qualitatively.

2. FORMULATION OF THE PROBLEM

We consider the motion of atoms undergoing the optical transition $J_g \longrightarrow J_e$ in a resonance monochromatic field

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

We represent the spatially inhomogeneous vector

amplitude in the form

$$\mathbf{E}(\mathbf{r}) = E(\mathbf{r})\mathbf{e}(\mathbf{r})e^{i\Phi(\mathbf{r})}, \quad (1)$$

where $E(\mathbf{r})$ is the real-valued amplitude and $\mathbf{e}(\mathbf{r})$ is the unit complex polarization vector. The field phase $\Phi(\mathbf{r})$ is defined so that $\mathbf{e}(\mathbf{r}) \cdot \mathbf{e}(\mathbf{r}) = \cos(2\varepsilon(\mathbf{r}))$ is a real-valued quantity determining the local ellipticity of the light field ($\varepsilon(\mathbf{r})$ is the ellipticity parameter and $|\tan\varepsilon|$ is equal to the ratio of the semiminor axis of the polarization ellipse to its semimajor axis). The Hamiltonian of a free atom in a rotating (in the energy pseudospin space) basis has the form

$$\hat{H}_0 = \frac{\hat{\mathbf{p}}^2}{2M} - \hbar\delta\hat{\Pi}_e, \quad (2)$$

where $\delta = \omega - \omega_0$ is the field frequency detuning from the atomic transition frequency ω_0, M is the mass of the atom, and the projection operator

$$\hat{\Pi}_e = \sum_{\mu_e} |J_e, \mu_e\rangle\langle J_e, \mu_e| \quad (3)$$

is constructed from the wave functions of the Zeeman sublevels $|J_e, \mu_e\rangle$ of the excited state. In the dipole approximation, the operator of resonant interaction with field (1) can be written in the form

$$\hat{V}(\mathbf{r}) = \hbar\Omega(\mathbf{r}) \sum_{q=0,\pm1} \hat{D}_q e^q(\mathbf{r}) + \text{H.c.}, \quad (4)$$

where $\Omega = -dE/\hbar$ is the Rabi frequency, d is the reduced matrix element, and $e^q(\mathbf{r})$ are contravariant components of the polarization vector in the cyclic basis

$$\{\mathbf{e}_0 = \mathbf{e}_z, \mathbf{e}_{\pm1} = \mp(\mathbf{e}_x \pm i\mathbf{e}_y)/\sqrt{2}\}.$$

Operator $\hat{D}_q$ can be expressed in terms of the $3jm$ symbols:

$$\hat{D}_q = \sum_{\mu_e, \mu_g} |J_e, \mu_e\rangle(-1)^{J_e - \mu_e} \times \begin{pmatrix} J_e & 1 & J_g \\ -\mu_e & q & \mu_g \end{pmatrix} \langle J_g, \mu_g|. \quad (5)$$

The state of an atomic ensemble is described by the one-particle density matrix $\hat{\rho}$; the quantum-mechanical kinetic equation for this matrix has the form

$$\frac{\partial}{\partial t}\hat{\rho} = -\frac{i}{\hbar}[\hat{H}_0, \hat{\rho}] - \frac{i}{\hbar}[\hat{V}(\hat{\mathbf{r}}), \hat{\rho}] - \hat{\Gamma}\{\hat{\rho}\}, \quad (6)$$

where $\hat{\Gamma}\{\hat{\rho}\}$ is the radiation relaxation operator. Interatomic interactions will be disregarded. Since Eq. (6) describes the evolution of external as well as internal (translational) degrees of freedom of an atom, solving this equation is a complicated problem. However, in the semiclassical approximation, rapid processes of ordering in internal degrees of freedom can be separated from the slow processes associated with translational motion of an atom if the following three basic conditions are satisfied.

1. The dispersion Δp of an atomic momentum considerably exceeds the recoil momentum $\hbar k$ of a photon:

$$\frac{\hbar k}{\Delta p} \ll 1. \quad (7)$$

2. The time of interaction between an atom and the field must exceed the characteristic time of evolution of the internal degrees of freedom of the atom,

$$t \gg \tau = \max\{\gamma^{-1}, (\gamma G)^{-1}\}, \quad (8)$$

where

$$G = \frac{|\Omega|^2}{3(\gamma^2/4 + \delta^2)}$$

is the saturation parameter.

3. The recoil energy $\hbar^2k^2/2M$ is much smaller than $\hbar\tau^{-1}$,

$$\frac{\hbar k^2}{M} \ll \tau^{-1}. \quad (9)$$

It was shown in [1, 2, 9, 10] that, under the above conditions, Eq. (6) can be reduced to the equation $\mathcal{W}(\mathbf{r}, \mathbf{p}, t) = \text{Tr}\{\hat{\rho}(\mathbf{r}, \mathbf{p}, t)\}$ for the atomic distribution function in the phase space (the trace is taken over internal states of the density matrix in the Wigner representation):

$$\left(\frac{\partial}{\partial t} + \sum_i \frac{p_i}{M}\nabla_i\right)\mathcal{W} = -\sum_i \frac{\partial}{\partial p_i}F_i(\mathbf{r}, \mathbf{p})\mathcal{W} + \sum_{ij} \frac{\partial^2}{\partial p_i \partial p_j}D_{ij}(\mathbf{r}, \mathbf{p})\mathcal{W}. \quad (10)$$

The kinetic coefficients $F_i(\mathbf{r}, \mathbf{p})$ and $D_{ij}(\mathbf{r}, \mathbf{p})$ are the components of the force and the diffusion tensor in the Cartesian basis.

3. KINETIC COEFFICIENTS FOR SLOW ATOMS

The kinetic coefficients of the Fokker–Planck equation can be determined from the equations derived by the reduction of Eq. (6) to Eq. (10) and depend on the distribution of atoms over internal states. These coefficients are often determined using the approximation of slow atoms, which are displaced over a distance much

shorter than the wavelength of light during the relaxation time for internal degrees of freedom; i.e., $v\tau \ll \lambda$ or

$$kv \ll \gamma, \gamma G. \tag{11}$$

In this case, to take into account dissipative processes correctly, it is sufficient to confine the analysis to the linear approximation (in velocity $\mathbf{v} = \mathbf{p}/M$) for the force,

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

and to the zeroth approximation for diffusion,

$$D(\mathbf{r}, \mathbf{v})_{ij} \approx D(\mathbf{r})_{ij}. \tag{13}$$

Quantity $\mathbf{F}(\mathbf{r})$ is the force of light pressure acting on a stationary atom. The asymmetric part of tensor ξ_{ij} corresponds to the effective Lorentz force, while its symmetric part defines friction.

The general procedure for deriving the Fokker–Planck equation via the semiclassical expansion in small parameter $\hbar k/\Delta p$ is described in detail in [1, 2, 9, 10]. However, the determination of the explicit form of the kinetic coefficients for atoms with energy levels degenerate in the angular momentum projection in a field of the most general form is a complex mathematical problem even for small values of angular momenta J_g and J_e. For example, in order to determine the friction and diffusion coefficients, it is necessary to calculate and integrate a matrix exponential [9]. The complexity of such calculations increases rapidly with the angular momentum of atomic levels. It is probably for this reason that the atomic kinetics was analyzed only for several particular light field configurations formed by linearly or circularly polarized traveling waves.

An alternative method for determining the kinetic coefficients was proposed (without derivation) in [8]. This method, which is in line with the methods developed earlier, is more convenient in our opinion since the search for the coefficients of friction and diffusion is reduced to solving a single algebraic equation for an auxiliary matrix $\hat{\varphi}$. Here, we will use these results (the derivation is given in Appendix A).

The gradient force acting on a stationary atom is a quantum-mechanical mean of the force operator $\hat{F}$,

$$F_i(\mathbf{r}) = \mathrm{Tr}\{\hat{F}_i(\mathbf{r})\hat{\sigma}(\mathbf{r})\}, \quad \hat{F}_i(\mathbf{r}) = -\nabla_i \hat{V}(\mathbf{r}), \tag{14}$$

where the density matrix $\hat{\sigma}(\mathbf{r})$ is the solution to the optical Bloch equation

$$\left[\left(\frac{\gamma}{2} - i\delta\right)\hat{\Pi}_e\hat{\sigma} + \left(\frac{\gamma}{2} - i\delta\right)\hat{\sigma}\hat{\Pi}_e\right] + \frac{i}{\hbar}[\hat{V}, \hat{\sigma}] - \gamma(2J_e + 1)\sum_q \hat{D}_q^\dagger \hat{\sigma}\hat{D}_q = 0 \tag{15}$$

with the normalization condition $\mathrm{Tr}\{\hat{\sigma}\} = 1$ and describes the steady-state distribution of atoms over magnetic sublevels in the zeroth approximation in the recoil parameter and the velocity of atoms. It should be noted that the explicit analytic form of $\hat{\sigma}$ for all dipole transitions was determined in [11, 12].

The coefficient of friction is proportional to the spatial gradient of $\hat{\sigma}$:

$$\xi_{ij}(\mathbf{r}) = -\hbar k \mathrm{Tr}\{\hat{\varphi}_i \nabla_j \hat{\sigma}\}. \tag{16}$$

The diffusion coefficient can be represented as the sum of two terms,

$$D_{ij}(\mathbf{r}) = D_{ij}^{(s)}(\mathbf{r}) + D_{ij}^{(i)}(\mathbf{r}), \tag{17}$$

the first of which,

$$D_{ij}^{(s)}(\mathbf{r}) = \frac{\gamma}{5}(\hbar k)^2 \times \mathrm{Tr}\left\{\left(\delta_{ij}\hat{\Pi}_e - \frac{2J_e + 1}{4}(\hat{D}_i\hat{D}_j^\dagger + \hat{D}_j\hat{D}_i^\dagger)\right)\hat{\sigma}\right\}, \tag{18}$$

is due to recoil in the case of spontaneous emission, while the second term,

$$D_{ij}^{(i)}(\mathbf{r}) = \frac{\hbar k}{4} \times \mathrm{Tr}\{\hat{\varphi}_i(\hat{\sigma}\widehat{\delta F}_j + \widehat{\delta F}_j\hat{\sigma}) + (\hat{\sigma}\widehat{\delta F}_i + \widehat{\delta F}_i\hat{\sigma})\hat{\varphi}_j\}, \tag{19}$$

is determined by the operator of force fluctuation,

$$\widehat{\delta F}(\mathbf{r}) = \hat{F}(\mathbf{r}) - F(\mathbf{r}). \tag{20}$$

Matrix $\hat{\varphi}$ used in formulas (16) and (19) satisfies the equation (see Appendix A)

$$\left[\left(\frac{\gamma}{2} + i\delta\right)\hat{\Pi}_e\hat{\varphi}_i + \left(\frac{\gamma}{2} - i\delta\right)\hat{\varphi}_i\hat{\Pi}_e\right] - \frac{i}{\hbar}[\hat{V}, \hat{\varphi}_i] - \gamma(2J_e + 1)\sum_q \hat{D}_q\hat{\varphi}_i\hat{D}_q^\dagger = \frac{\widehat{\delta F_i}}{\hbar k} \tag{21}$$

and makes it possible to write the expressions for the friction and diffusion coefficients in a universal form.

4. ANALYTIC EXPRESSIONS FOR KINETIC COEFFICIENTS FOR TRANSITION $J_g = 1/2 \longrightarrow J_e = 1/2$

We will consider here one-dimensional configurations of a light field, when the vector amplitude is a function of only one coordinate *z*. In the general case, such a field is formed by two counterpropagating traveling plane waves with arbitrary intensities and ellipti-

cal polarizations. In this case, the local vector field amplitude (1) can be characterized by four real-valued quantities: the real amplitude, phase, ellipticity, and the angle of orientation of the polarization ellipse. In the coordinate system where the z axis is orthogonal to the polarization plane, vector $\mathbf{e}(z)$ in a cyclic basis can be written in terms of the local ellipticity $\varepsilon(z)$ of the light field and the angle of rotation $\varphi(z)$ of the polarization ellipse:

$$\mathbf{e}(z) = \cos\left(\varepsilon + \frac{\pi}{4}\right)e^{-i\phi}\mathbf{e}_{-1} - \cos\left(\varepsilon - \frac{\pi}{4}\right)e^{-i\phi}\mathbf{e}_{+1}. \quad (22)$$

For simplicity, we will henceforth consider a one-dimensional problem. It should be noted that, in the general case, the kinetic problem can be reduced to a one-dimensional problem only approximately, since recoil processes occurring during spontaneous reemission of photons of the field result in the formation of a coupling between translational degrees of freedom of an atom in all three directions. Coupling in the longitudinal and transverse directions relative to the z axis can be neglected when the width $\Delta p_\perp$ of the momentum distribution in the transverse direction considerably exceeds the width Δp_z of the momentum distribution in the longitudinal direction ($\Delta p_\perp \gg \Delta p_z$) and the recoil effect in the transverse direction does not lead to a noticeable change in the momentum distribution. Such a model is applicable, for example, for a problem on interaction of an atomic beam with a transverse light field. Henceforth, we will assume that the force and the diffusion tensor are functions only of coordinate z and the corresponding velocity component, which will be denoted by v. In the notation adopted here, it is convenient to express kinetic coefficients in terms of the gradients of light field parameters: $\nabla_z\Phi$, $\nabla_z\phi$, and $\nabla_z\varepsilon$ (gradients of the phase, the angle of rotation of the polarization ellipse, and the ellipticity) as well of $\nabla_z\Lambda$ ($\Lambda = \ln E$) (gradient of the logarithm of the real part of the field amplitude).

4.1. Force

Using the explicit form for matrix $\hat{\sigma}$ (15), we obtain the following expression for force in the zeroth order in velocity:

$$F_z = \frac{\hbar G_\varepsilon}{1+G_\varepsilon}\left[\delta(\tan(2\varepsilon)\nabla_z\varepsilon - \nabla_z\Lambda) + \frac{\gamma}{2}\nabla_z\Phi\right]. \quad (23)$$

For the sake of simplicity, we introduced here the effective saturation parameter $G_\varepsilon = G\cos^2(2\varepsilon)$. The first term, proportional to δ, is the force of induced light pressure emerging from the gradients of ellipticity and light field amplitude, while the gradient of phase makes a contribution to the force of spontaneous light pressure, associated with spontaneous rescattering of photons. It should be noted that expression (23) does not contain the contribution from the gradient $\nabla_z\phi$ of orientation of the polarization ellipse. This is due to the specific form of matrix $\hat{\sigma}$ for optical transitions $J_g = J \longrightarrow J_e = J$ (J is a half-integer) [12, 13], for which the magnetic sublevels of the excited state are populated uniformly (isotropically).

4.2. Coefficients of Friction and Diffusion

Coefficients of friction, $\xi = \xi_{zz}$, and induced diffusion, $D^{(i)} = D^{(i)}_{zz}$, can be written in the general form,

$$\xi = \hbar\sum_{\beta\beta'}\chi_{\beta\beta'}\nabla_z\beta\nabla_z\beta', \quad (24)$$

$$D^{(i)} = \hbar^2\sum_{\beta\beta'}\mathcal{D}_{\beta\beta'}\nabla_z\beta\nabla_z\beta', \quad (25)$$

where summation indices β and β' correspond to $\{\Lambda, \varepsilon, \Phi, \phi\}$, and components $\chi_{\beta\beta'}$ and $\mathcal{D}_{\beta\beta'}$ can be expressed in the parameters of the light field. Since in the problem under investigation we are dealing only with gradients along the z axis, the coefficients of friction and induced diffusion contain only symmetric combinations of nondiagonal ($\beta \neq \beta'$) terms: $\chi_{\beta\beta'} + \chi_{\beta'\beta}$ and $\mathcal{D}_{\beta\beta'} + \mathcal{D}_{\beta'\beta}$. It should be noted that expressions $\chi_{\beta\beta'}$ and $\mathcal{D}_{\beta\beta'}$ are also applicable in the 3D formulation of the problem. In this case, it remains for us to calculate the required gradients of the light field parameters to find the tensors of friction and diffusion. Note that, in the most general 3D formulation of the problem, we must take into account, in addition to the available gradients of $\{\Lambda, \varepsilon, \Phi, \phi\}$, the two angles describing the rotation drawing from the plane of the polarization ellipse. However, for 3D problems in which the plane of the polarization ellipse remains unchanged, the above results are sufficient. The expressions for the components of the coefficients of friction $\chi_{\beta\beta'}$ and diffusion $\mathcal{D}_{\beta\beta'}$ are given in Appendix B.

Spontaneous diffusion coefficient $D^{(s)} = D^{(s)}_{zz}$ is proportional to the total population of the excited level and is independent of the external field gradients:

$$D^{(s)} = \frac{\gamma(\hbar k)^2}{12}\frac{G_\varepsilon}{1+G_\varepsilon}. \quad (26)$$

The expressions for the force and the coefficients of friction and diffusion derived here make it possible to describe the kinetics of slow atoms with the optical transition $J_g = 1/2 \longrightarrow J_e = 1/2$, which is a simple model for atoms with energy levels degenerate in the angular momentum projection. It should be noted that simple configurations of light field exist, which are characterized by the gradient of only one parameter. For exam-

ple, a standing uniformly polarized wave is characterized by intensity gradient $\nabla_z\Lambda$, a traveling uniformly polarized wave is characterized by phase gradient $\nabla_z\Phi$, the field of the σ_+–σ_- configuration is characterized by gradient $\nabla_z\phi$ of the orientation of the polarization vector, while the field of the lin $\perp$ lin configuration is characterized by ellipticity gradient $\nabla_z\varepsilon$. For these configurations, the coefficients of friction and diffusion are determined by the corresponding diagonal elements $\chi_{\beta\beta}$ and $\mathcal{D}_{\beta\beta}$.

5. UNIFORMLY POLARIZED FIELD

In a uniformly polarized field, the ellipticity and orientation of the polarization ellipse are independent of the coordinate; i.e., such a field is characterized only by the gradients of phase and intensity. On the other hand, precisely these gradients appear in the description of the atomic kinetics in the nondegenerate two-level model of an atom [1, 2]. For this reason, it would be especially important to consider this type of light fields separately and compare the expressions for the kinetic coefficients (force and coefficients of friction and diffusion) obtained in this way with the available results for the two-level model of an atom. This will enable us to determine the differences associated with the degeneracy of atomic levels in the angular momentum projection and to single out the effects determined by the ellipticity of a light field.

5.1. Elliptically Polarized Traveling Wave

A simple example of a uniformly polarized configuration of a light field is a traveling plane monochromatic wave,

$$\mathbf{E}(z, t) = E_0\mathbf{e}e^{ikz}e^{-i\omega t} + \text{c.c.} \tag{27}$$

Light field amplitude E_0 and polarization vector $\mathbf{e}$ are spatially homogeneous, while phase $\Phi = kz$ is a function of the coordinate. The force acting on an atom in this field is proportional to phase gradient (23) and is a force of spontaneous light pressure by nature. Expressions (B.3) for the friction coefficient, as well as the expressions for the force, coincide in form with the results for a two-level atom with an effective saturation parameter of G_ε. The most significant difference from the model of a two-level atom appears in expression (B.19) for the diffusion coefficient,

$$D^{(i)} = \hbar^2k^2\frac{\gamma}{4}\frac{G_\varepsilon}{(1+G_\varepsilon)^3} \times\left[(1+G_\varepsilon)^2 + 3\sin^2(2\varepsilon) + G_\varepsilon\left(1-\frac{\gamma^2}{\gamma^2/4+\delta^2}\right)\right]; \tag{28}$$

namely, an additional term proportional to $\sin^2(2\varepsilon)$ appears, which is significant in a low-intensity field ($G_\varepsilon \ll 1$). It is well known that induced diffusion of atoms in a traveling wave is associated with fluctuation of the number of photons scattered from an atom [1, 14]. This quantity is determined by the statistics of the number of scattered photons; for a two-level atom in a low-intensity field ($G \ll 1$), the distribution is of the Poisson type. The presence of the additional term in expression (28) indicates a deviation from the Poisson distribution in the statistics of the number of photons scattered in a low-intensity field. These deviations are due to correlation of processes of emission of photons with the σ_+ and σ_- polarization in an elliptically polarized field. Indeed, the scattering probability $\sigma_\pm$ of a circularly polarized quantum of the external field in a small time interval of Δt,

$$p_\pm = \gamma G\frac{1\pm\sin(2\varepsilon)}{2}\pi_\mp\Delta t,$$

depends on the degree of ellipticity and the probability $\pi_\pm$ of finding an atom in the ground state with an angular momentum projection of $\mu_g = \pm 1/2$ ($\pi_+ + \pi_- = 1$). Let us suppose, for example, that a σ_+ photon was scattered in a certain interval; then, $\pi_+ = 1/3$ and $\pi_- = 2/3$ for a subsequent interval in accordance with the relative probabilities of spontaneous decay via the channels $|\mu_e = 1/2\rangle \longrightarrow |\mu_g = 1/2\rangle$ and $|\mu_e = 1/2\rangle \longrightarrow |\mu_g = -1/2\rangle$. On the contrary, after the scattering of a σ_- photon, we have $\pi_+ = 2/3$ and $\pi_- = 1/3$. Consequently, for $\varepsilon \neq 0$, the scattering probability for a photon (of any type),

$$p_+ + p_- = \gamma\frac{G}{2}[1-(\pi_+-\pi_-)\sin(2\varepsilon)]\Delta t,$$

depends on the polarization of the quantum scattered before this event; i.e., the scheme of independent trials is violated. The average number of scattered quanta is determined by the mean values

$$\bar{\pi}_\pm = \frac{1\pm\sin(2\varepsilon)}{2}.$$

Consequently, the statistics of the number of photons is not Poissonian any longer; we can conclude from relation (28) that the statistics is super-Poissonian by nature.

5.2. Uniformly Polarized Standing Wave

In a uniformly polarized standing wave, the real amplitude (1) is a periodic function of coordinate ($E(z) = E_0\cos(kz)$) and polarization vector $\mathbf{e}$ is spatially homogeneous. In such a field, only the gradient of intensity Λ differs from zero; consequently, the force acting on a stationary atom,

$$F(z) = -\frac{\hbar\delta G_\varepsilon\nabla_z\Lambda}{1+G_\varepsilon}, \tag{29}$$

is the force of induced light pressure. The coefficients

of friction and induced diffusion have the form

$$\xi(z) = \hbar\chi_{\Lambda\Lambda}(\nabla_z\Lambda)^2, \quad (30)$$

$$D^{(i)}(z) = \hbar^2\mathcal{D}_{\Lambda\Lambda}(\nabla_z\Lambda)^2. \quad (31)$$

The case of a homogeneously polarized standing wave with an arbitrary analyticity ε was considered in our earlier publication [8]. Henceforth, we will consider only some interesting features in the kinetics which had not been analyzed earlier. In particular, we will find the difference between our results and the results obtained in the nondegenerate model of an atom on the basis of spatially averaged expressions for the coefficients of friction and diffusion.

The gradient force averaged over the spatial period of a light field vanishes ($\langle F(z)\rangle = 0$); the friction coefficient can be represented as the sum of two contributions:

$$\langle\xi\rangle = \langle\xi_1\rangle + \langle\xi_2\rangle. \quad (32)$$

The first term,

$$\langle\xi_1\rangle = \frac{\hbar k^2}{4}\left[\frac{\delta\gamma S_\varepsilon(2+S_\varepsilon)}{(1+S_\varepsilon)^{3/2}(\delta^2+\gamma^2/4)} - \frac{\delta(8+3S_\varepsilon^3+15S_\varepsilon^2+20S_\varepsilon-8(1+S_\varepsilon)^{5/2})}{\gamma(1+S_\varepsilon)^{5/2}}\right] \quad (33)$$

corresponds to the expression in the two-level model of an atom [1] with the effective saturation parameter

$$S_\varepsilon = \frac{|E_0 d|^2\cos^2(2\varepsilon)}{\hbar^2 3(\delta^2+\gamma^2/4)}. \quad (34)$$

The second term,

$$\langle\xi_2\rangle = \hbar k^2\left[\frac{3\delta\sin^2(2\varepsilon)S_\varepsilon(4+3S_\varepsilon)}{4\gamma(1+S_\varepsilon)^{3/2}}\right] \quad (35)$$

is an additional contribution to friction and may be interpreted in terms of the probabilities of transitions between dressed states of the atom (Sisyphus effect) [8, 15]. For small saturation parameters ($S_\varepsilon < 1$), the main contribution to this effect comes from transitions between dressed states, which correspond to the Zeeman sublevels of the ground state; this leads to an additional contribution to friction.

In the low-intensity limit of the light field ($S_\varepsilon \ll 1$), coefficients $\langle\xi_1\rangle$ and $\langle\xi_2\rangle$ have the form

$$\langle\xi_1\rangle = \frac{\hbar k^2}{2}\frac{\delta\gamma S_\varepsilon}{\delta^2+\gamma^2/4}, \quad (36)$$

$$\langle\xi_2\rangle = \hbar k^2 3\sin^2(2\varepsilon)\frac{\delta S_\varepsilon}{\gamma}. \quad (37)$$

Contribution $\langle\xi_1\rangle$ corresponds to the well-known Doppler mechanism of friction [1, 2]. For large detunings ($\delta \gg \gamma$), the relation

$$\frac{\langle\xi_2\rangle}{\langle\xi_1\rangle} \approx 6\sin^2(2\varepsilon)\frac{\delta^2}{\gamma^2}$$

holds, which shows that, in an elliptically polarized field ($\varepsilon \neq 0$), the additional contribution can exceed the result for the two-level model by several orders of magnitude.

In intense fields ($S_\varepsilon \gg 1$), the expressions for the coefficients assume the form

$$\langle\xi_1\rangle = \hbar k^2\frac{\delta}{4\gamma}\frac{(\gamma^2-12\delta^2)\sqrt{S_\varepsilon}}{4\delta^2+\gamma^2}, \quad (38)$$

$$\langle\xi_2\rangle = \hbar k^2\frac{9\sin^2(2\varepsilon)\delta\sqrt{S_\varepsilon}}{4\gamma}. \quad (39)$$

It should be noted that the sign of the friction coefficient in the two-level model of an atom in a high-intensity field is determined not only by the sign of detuning, but also by its magnitude (38). The inclusion of degeneracy of atomic levels leads to the following interesting effect: for field ellipticities $\sin^2(2\varepsilon) > 1/3$, in view of additional contribution (39), the sign of the total friction coefficient (32) is determined only by the sign of detuning: heating for $\delta > 0$ and cooling for $\delta < 0$.

The averaged coefficient of induced diffusion, as well as the friction coefficient, can be presented as the sum of two terms:

$$\langle D^{(i)}\rangle = \langle D_1^{(i)}\rangle + \langle D_2^{(i)}\rangle. \quad (40)$$

The first term represented the familiar result [1] for a two-level atom with a new parameter S_ε:

$$\langle D_1^{(i)}\rangle = \frac{\hbar^2k^2\gamma S_\varepsilon}{8} + \frac{\hbar^2k^2\delta^2}{\gamma}\left[3+\frac{S_\varepsilon}{2} - \frac{15S_\varepsilon^2+40S_\varepsilon+24+\gamma^2S_\varepsilon^2(\delta^2+\gamma^2/4)^{-1}}{8(1+S_\varepsilon)^{3/2}}\right]. \quad (41)$$

The second term,

$$\langle D_2^{(i)}\rangle = \frac{\hbar^2k^2\sin^2(2\varepsilon)}{\gamma} \times\left[\frac{3\delta^2S_\varepsilon(4+3S_\varepsilon)}{8(1+S_\varepsilon)^{2/3}} + \left(\delta^2+\frac{\gamma^2}{4}\right)\frac{1+S_\varepsilon/2-\sqrt{1+S_\varepsilon}}{3\cos^2(2\varepsilon)}\right] \quad (42)$$

is additional as compared to the two-level model of the atom and describes the contributions from the diffusion processes associated with Zeeman degeneracy of energy levels.

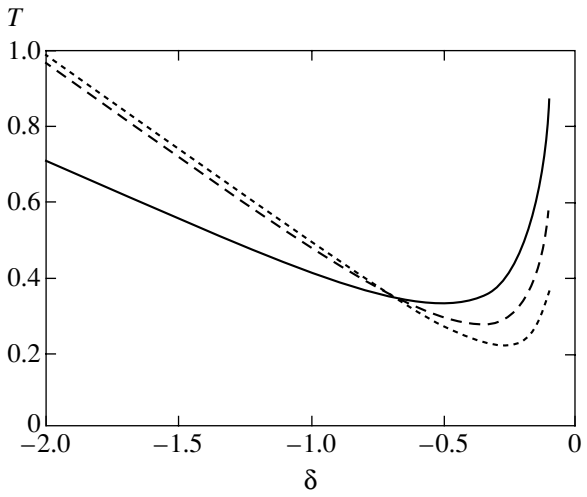

Fig. 1. Temperature in units of $\hbar\gamma/k_B$ as a function of detuning for various ellipticities of the light field: the solid curve corresponds to the linear polarization of the field, the dashed curve corresponds to a field ellipticity of $\sin^2(2\varepsilon) = 1/3$, and the dotted curve is the formal limit for the circular field polarization.

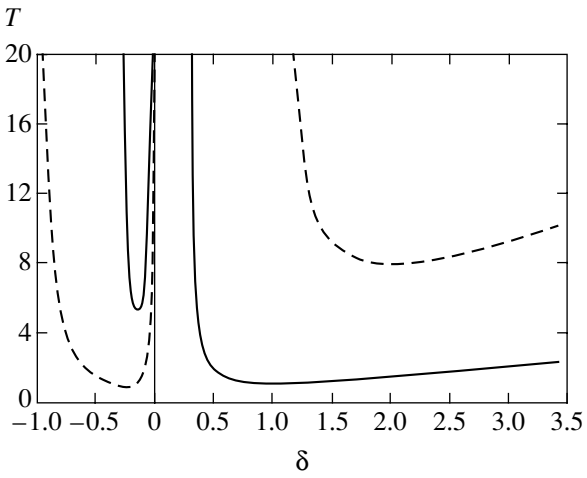

Fig. 2. Temperature in units of $\hbar\gamma S^{1/2}/k_B$ as a function of detuning for various ellipticities of the light field: the solid curve corresponds to the linear polarization of the field and the dashed curve corresponds to a field ellipticity of $\varepsilon = \pi/12$.

According to [1], the expressions written above for the coefficients of friction and diffusion averaged over the space period make it possible to estimate the temperature of the atomic ensemble, $k_BT = -\langle D\rangle/\langle\xi\rangle$, without taking into account localization as a function of the field ellipticity. In a weak field ($S_\varepsilon \ll 1$), we obtain

$$k_BT = -\frac{\hbar}{\delta}\frac{(9\sin^2(2\varepsilon)\delta^2 + \gamma^2)(4\delta^2 + \gamma^2)}{6(12\sin^2(2\varepsilon)\delta^2 + 3\sin^2(2\varepsilon)\gamma^2 + 2\gamma^2)}. \quad (43)$$

The diffusion coefficient is positive; consequently, the steady-state distribution of atoms is possible only for $\langle\xi\rangle < 0$ and the expression for temperature has sense only for $\delta < 0$. For the linear field polarization, the result can be reduced to the known expression for the Doppler temperature of laser cooling in the model of a two-level atom,

$$k_BT(\varepsilon = 0) = -\frac{\hbar(4\delta^2 + \gamma^2)}{12\delta}; \quad (44)$$

the minimum of this temperature is attained for $\delta = -\gamma/2$ and amounts to $k_BT_D = \hbar\gamma/3$. In the general case, the inclusion of the field polarization leads to a slightly different dependence of temperature on detuning (Fig. 1). It can be seen that the temperature can assume values smaller than in the case of the linear polarization of the field. It can be verified that the lowest temperature value is reached for a field ellipticity of $|\varepsilon| \longrightarrow \pi/4$ (i.e., for the circular polarization).

In the other limiting case of a high-intensity light field ($S_\varepsilon \gg 1$), the atomic temperature can be estimated using the following dependence:

$$k_BT = -\frac{\hbar}{\delta} \quad (45)$$
$$\times\frac{(3\cos^2(2\varepsilon) + \sin^2(2\varepsilon))(4\delta^2 + \gamma^2)^2\sqrt{S_\varepsilon}}{6\cos^2(2\varepsilon)(36\sin^2(2\varepsilon)\delta^2 + 9\sin^2(2\varepsilon)\gamma^2 + \gamma^2 - 12\delta^2)}.$$

Here, we can separate two different cases. In the first case, when the light field ellipticity is $\sin^2(2\varepsilon) \geq 1/3$, the direction of the kinetic process is determined only by the sign of δ (cooling for $\delta < 0$). For a field ellipticity of $\sin^2(2\varepsilon) < 1/3$, cooling takes place for detunings belonging to the following two intervals:

$$-\gamma\frac{\sqrt{(3 - 9\sin^2(2\varepsilon))(9\sin^2(2\varepsilon) + 1)}}{6 - 18\sin^2(2\varepsilon)} < \delta < 0,$$
$$\delta > \gamma\frac{\sqrt{(3 - 9\sin^2(2\varepsilon))(9\sin^2(2\varepsilon) + 1)}}{6 - 18\sin^2(2\varepsilon)}. \quad (46)$$

In particular, for the linear field polarization, the expression for temperature assumes the familiar form [1]

$$k_BT(\varepsilon = 0) = \frac{\hbar}{2\delta}\frac{(4\delta^2 + \gamma^2)^2\sqrt{S_\varepsilon}}{12\delta^2 - \gamma^2}. \quad (47)$$

The dependence of temperature on the light field detuning for light field ellipticities of $\sin^2(2\varepsilon) < 1/3$ is shown in Fig. 2.

It should be emphasized once again that the estimate for temperature is the ratio of the friction and diffusion coefficients; i.e., this estimate is based on slow atom approximation (11). It is well known [1], however, that the applicability of this approximation is restricted by

rather severe conditions even in the two-level model of an atom in a standing wave in view of the presence of light field nodes. In the case of a nonzero ellipticity of the light field, this problem requires additional analysis, which will not be carried out here.

5.3. Uniformly Polarized Field of a "Mixed" Type

In the general case, a uniformly polarized field may simultaneously contain the amplitude gradient and the phase gradient. A situation of this kind arises, for example, in the field formed by counterpropagating waves with identical polarization ellipses, but with different amplitudes. The friction and diffusion coefficients in such a field contain, in addition contributions diagonal in gradients, crossed contributions $\chi_{\Lambda\Phi}$ and $\mathcal{D}_{\Lambda\Phi}$. It should be noted that the additional contribution to the friction coefficient exhibits an even dependence on detuning δ, while the additional contribution to the diffusion coefficient is an odd function of δ. Thus, in a "mixed"-type field, an "abnormal" dependence of the friction and diffusion coefficients on the detuning is observed (in particular, the friction coefficient does not vanish in the case of exact resonance, $\delta = 0$, owing to additional contributions $\chi_{\Lambda\Phi}$ and $\chi_{\Phi\Lambda}$; see Appendix B).

6. NONUNIFORMLY POLARIZED FIELDS

It is well known that sub-Doppler cooling of atoms is possible in fields with nonuniform polarization. Simple examples of this kind are fields with an ellipticity gradient (lin $\perp$ lin field configuration) and fields with an orientation gradient (σ_+–σ_- field configuration). In the former case, the field is a superposition of counterpropagating plane waves with linear polarizations oriented at right angles to each other:

$$\mathbf{E}(z, t) = E_0(\mathbf{e}_x e^{ikz} + \mathbf{e}_y e^{-ikz})e^{-i\omega t} + \text{c.c.} \qquad (48)$$

Here, E_0 is the amplitude of each wave, while vectors $\mathbf{e}_x$ and $\mathbf{e}_y$ describe linear polarization of the counterpropagating waves along the **x** and **y** axes, respectively. This expression can be reduced to formula (1) with parameters $E(z) = 2E_0$ and ellipticity $\varepsilon(z) = kz$. The phase and the angle of orientation are independent of coordinate z.

A field of the σ_+–σ_- configuration is a superposition of counterpropagating waves of circular polarizations with opposite directions of rotation:

$$\mathbf{E}(z, t) = E_0(\mathbf{e}_{+1} e^{ikz} + \mathbf{e}_{-1} e^{-ikz})e^{-i\omega t} + \text{c.c.} \qquad (49)$$

As we pass to formula (1), this gives $E(z) = 2E_0$, $\varepsilon(z) = 0$, $\Phi(z) = 0$, and $\phi(z) = kz$. The motion along the z axis changes only the orientation of the polarization vector of the light field.

It should be recalled that the expressions for the kinetic coefficients in these configurations are determined by contributions $(\nabla\varepsilon)^2$ (B.2) and $(\nabla\phi)^2$ (B.18) for the lin $\perp$ lin configuration and $\chi_{\varepsilon\varepsilon}$ (B.4) and $\mathcal{D}_{\varepsilon\varepsilon}$ (B.20) for the σ_+–σ_- configuration, which are diagonal in gradients $\chi_{\phi\phi}$ and $\mathcal{D}_{\phi\phi}$, respectively. "Simple" configurations of a nonuniformly polarized field are undoubtedly convenient for analysis of the atomic kinetics and make it possible to explain the main mechanisms of sub-Doppler cooling in low-intensity fields [7]. However, these mechanisms do not exhaust the atomic kinetics in fields with nonuniform polarization since, as follows from the general results for the friction and diffusion coefficients, the presence in the field of several gradients simultaneously leads to new interesting effects and new mechanisms of sub-Doppler cooling. We will consider these effects in greater detail.

6.1. The ε–θ–$\bar{\varepsilon}$ Field Configuration

A simple example of a field containing all gradients (of intensity, ellipticity, orientation, and phase) is the field formed by counterpropagating plane waves of the same intensity and modulo equal ellipticities, but with opposite directions of rotation:

$$\mathbf{E}(z) = E_0 \mathcal{E}(z) e^{-i\omega t} + \text{c.c.} \qquad (50)$$

Here, E_0 is the amplitude of each of the counterpropagating waves and vector $\mathcal{E}(z) = a_+\mathbf{e}_{+1} + a_-\mathbf{e}_{-1}$ with cyclic components $a_+(z)$ and $a_-(z)$ defines the local polarization ellipse, amplitude, and phase of the field,

$$\begin{aligned} a_+ &= -\cos\left(\varepsilon_0 + \frac{\pi}{4}\right)e^{ikz} - \cos\left(\varepsilon_0 + \frac{\pi}{4}\right)e^{-ikz}e^{-i\theta}, \\ a_- &= \cos\left(\varepsilon_0 + \frac{\pi}{4}\right)e^{ikz} + \cos\left(\varepsilon_0 - \frac{\pi}{4}\right)e^{-ikz}e^{i\theta}, \end{aligned} \qquad (51)$$

where θ is the angle between the principal semiaxes of the polarization ellipses of the counterpropagating waves (Fig. 3) and parameter ε_0 characterizes the degree of ellipticity of the counterpropagating waves. For field configuration (51), we will use the notation ε–θ–$\bar{\varepsilon}$. It should be noted that the familiar configurations lin $\perp$ lin ($\varepsilon_0 = 0$, $\theta = \pi/2$) and σ_+–σ_- ($\varepsilon_0 = \pi/4$) are special cases of the given field configuration.

6.1.1. Gradient force. For low field intensities ($G \ll 1$), the kinetic coefficients assume a simple form. For example, the gradient force splits into the sum of two contributions:

$$\begin{aligned} F(z) &= F^{(i)}(z) + F^{(s)}(z), \\ F^{(i)}(z) &= 2\hbar k \delta S_0 \\ &\times [\cos(2\varepsilon_0)\sin(2kz)(\cos\theta + \cos(2\varepsilon_0)\cos(2kz))] \\ &\times (1 + \cos(2\varepsilon_0)\cos\theta\cos(2kz))^{-1}, \\ F^{(s)}(z) &= \frac{\hbar k \gamma S_0}{2} \frac{\sin(4\varepsilon_0)\sin(2kz)\sin\theta}{1 + \cos(2\varepsilon_0)\cos\theta\cos(2kz)}. \end{aligned} \qquad (52)$$

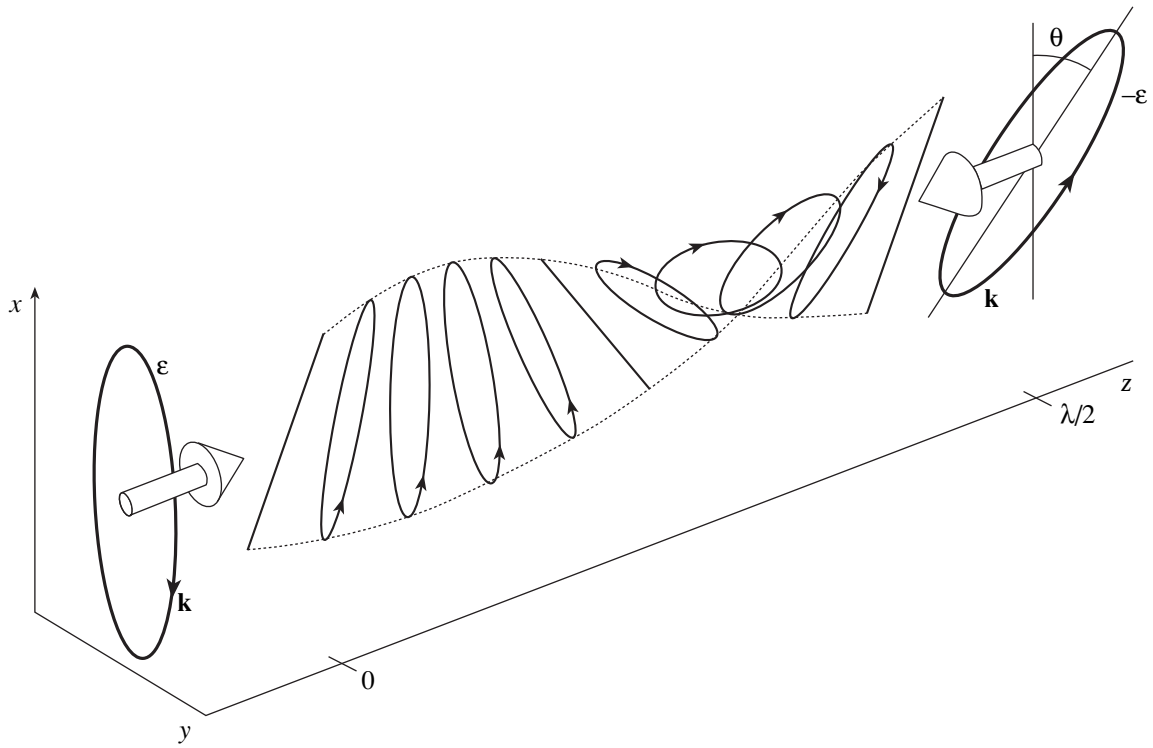

Fig. 3. Spatial configuration ε–θ–ε̄ of a light field. The field is produced by counterpropagating waves of elliptic polarization with ellipticity parameters of ε and –ε.

Here, the saturation parameter is given by

$$S_0 = \frac{|\Omega_0|^2}{3(\delta^2 + \gamma^2/4)}, \tag{53}$$

where $\Omega_0 = -E_0 d/\hbar$, E_0 being the amplitude of the counterpropagating waves. The first term, $F^{(i)}$, is the force of induced light pressure (dipole force), while the second term, $F^{(s)}$, is the force of spontaneous light pressure. It can be seen that the force of spontaneous light pressure induces in this case a periodic optical potential if the ellipticities of the light waves differ from linear or circular ellipticities ($\varepsilon_0 \neq 0, \pi/4$) and also if angle $\theta \neq 0$; in other words, the potential of familiar configurations with linear or circular polarizations is equal to zero. The mechanism of formation of the force of spontaneous light pressure is as follows: spatially inhomogeneous anisotropy of atoms is created by a nonuniformly polarized field; as a result of this anisotropy, atoms rescatter photons of counterpropagating waves with different probabilities, which leads to disbalance of the forces of spontaneous light pressure. It should be emphasized that a spontaneous force in the given field leads to a periodic optical potential that does not vanish in the case of exact resonance ($\delta = 0$). It should also be observed that $F^{(i)}$ is an odd function of detuning δ and an even function of polarization parameters ε_0 and θ, while $F^{(s)}$, on the contrary, is even in δ and odd in ε_0 and θ.

6.1.2. Friction coefficient. Like the gradient force, the friction coefficient slits onto two terms of different origin:

$$\begin{aligned}
\xi(z) &= \xi^{(i)}(z) + \xi^{(s)}(z),\\
\xi^{(i)}(z) &= -\frac{6\hbar k^2\delta}{\gamma}\sin^2\theta\cos^2(2\varepsilon_0)\\
&\times\cos(2kz)[\cos(2kz) + \cos\theta\cos(2\varepsilon_0)]\\
&\times[1 + \cos(2\varepsilon_0)\cos(2kz)\cos\theta]^{-3},\\
\xi^{(s)}(z) &= -\frac{3\hbar k^2}{2}\sin\theta\sin(4\varepsilon_0)\\
&\times[\cos(2kz) + \cos\theta\cos(2\varepsilon_0)]\\
&\times[1 + \cos(2\varepsilon_0)\cos(2kz)\cos\theta]^{-3}.
\end{aligned} \tag{54}$$

After averaging over the spatial period, we obtain

$$\begin{aligned}
\langle\xi^{(i)}\rangle &= -\frac{3\hbar k^2\delta}{\gamma}\frac{\sin^2\theta\cos^2(2\varepsilon_0)}{(1-\cos^2(2\varepsilon_0)\cos^2\theta)^{3/2}},\\
\langle\xi^{(s)}\rangle &= \frac{3\hbar k^2}{8}\frac{\sin(2\theta)\cos(2\varepsilon_0)\sin(4\varepsilon_0)}{(1-\cos^2(2\varepsilon_0)\cos^2\theta)^{3/2}}.
\end{aligned} \tag{55}$$

It is worth noting that average friction coefficients $\langle\xi^{(i)}\rangle$ and $\langle\xi^{(s)}\rangle$ cannot be expressed in analytic form at point $\theta = 0$, $\varepsilon_0 = 0$, in the vicinity of which we have

$$\langle\xi^{(i)}\rangle \sim \frac{\theta^2}{(\theta^2 + 4\varepsilon_0^2)^{3/2}}$$

and

$$\langle\xi^{(s)}\rangle \sim \frac{\theta\varepsilon_0}{(\theta^2 + 4\varepsilon_0^2)^{3/2}}.$$

Using the lin–θ–lin configuration as an example, it was noted in [16] that such a behavior of the friction coefficient is associated with inapplicability of the slow atom approximation in the vicinity of nodes of the field, where, on the one hand, the local saturation parameter is small and, on the other hand, noticeable gradients of the field polarization arise (i.e., the field polarization changes strongly over distances much smaller than wavelength λ). Analysis of situations in these regions requires more exact expressions for the force of friction, which would take into account all orders in the velocity of atoms.

6.1.3. Diffusion coefficient. The spontaneous diffusion coefficient has the form

$$D^{(s)} = \frac{\gamma^2\hbar^2k^2S_0}{6}\left[1 + \cos(2\varepsilon_0)\cos(2kz)\cos\theta - \frac{\cos^2(2\varepsilon_0)\sin^2\theta\sin^2(2kz)}{1 + \cos(2\varepsilon_0)\cos(2kz)\cos\theta}\right]. \quad (56)$$

It is convenient to decompose the induced diffusion coefficients into the terms corresponding to different degrees of detuning, $D^{(i)} = D_0 + D_1 + D_2$, where D_0 is the term independent of detuning,

$$D_0(z) = \frac{\gamma\hbar^2k^2S_0}{2}\frac{1 - \cos^2(2\varepsilon_0)(\cos^2(2kz)\cos^2\theta - \sin^2(2kz)\sin^2\theta) + \sin^2(2\varepsilon_0)}{1 + \cos(2\varepsilon_0)\cos(2kz)\cos\theta} - \frac{3\gamma\hbar^2k^2S_0}{8}\frac{\sin^2(4\varepsilon_0)\sin^2\theta\sin^2(2kz)}{(1 + \cos(2\varepsilon_0)\cos(2kz)\cos\theta)^3}. \quad (57)$$

The term with an abnormal dependence on detuning has the form

$$D_1(z) = \frac{\delta\hbar^2k^2S_0\sin(4\varepsilon_0)\sin\theta\cos(2kz)}{1 + \cos(2\varepsilon_0)\cos(2kz)\cos\theta} \times\left[2 - \frac{3\cos^2(2\varepsilon_0)\sin^2\theta\sin^2(2kz)}{(1 + \cos(2\varepsilon_0)\cos(2kz)\cos\theta)^2}\right], \quad (58)$$

while the part of the diffusion coefficient quadratic in detuning is given by

$$D_2(z) = \frac{6\delta^2\hbar^2k^2S_0\cos^2(2\varepsilon_0)\sin^2\theta\cos^2(2kz)}{\gamma(1 + \cos(2\varepsilon_0)\cos(2kz)\cos\theta)} \times\left[1 - \frac{\cos^2(2\varepsilon_0)\sin^2\theta\sin^2(2kz)}{(1 + \cos(2\varepsilon_0)\cos(2kz)\cos\theta)^2}\right]. \quad (59)$$

After averaging over the spatial period, we obtain the following expressions for the diffusion coefficient. The average spontaneous diffusion coefficient has the form

$$\langle D^{(s)}\rangle = \frac{\hbar^2k\gamma S_0}{6}\frac{Q\sin\theta^2 + \cos(2\theta)}{\cos^2\theta}. \quad (60)$$

The average induced diffusion coefficient is given by $\langle D^{(i)}\rangle = \langle D_0\rangle + \langle D_1\rangle + \langle D_2\rangle$, where

$$\langle D_0\rangle = \frac{\hbar^2k\gamma S_0}{4}\left[\frac{2(1 - Q\sin^2\theta)}{\cos^2\theta} + \frac{\sin^2(2\varepsilon_0)(2\sin^2(2\varepsilon_0) + \cos^2(2\varepsilon_0)\cos^2\theta - 1)}{Q^3}\right],$$
$$\langle D_1\rangle = \frac{\hbar^2k\delta S_0\sin\theta\sin(2\varepsilon_0)(Q-1)[2Q^2(2 + \sin^2\theta) - 3(Q+1)\sin^2\theta]}{Q^3\cos^3\theta}, \quad (61)$$
$$\langle D_2\rangle = \frac{3\hbar^2kS_0\delta^2}{\gamma}\frac{(Q-1)\sin^2\theta}{Q^3\cos^4\theta}(2Q^3\sin^2\theta - 2Q^2(\sin^2\theta + 1) + (Q+1)\sin^2\theta).$$

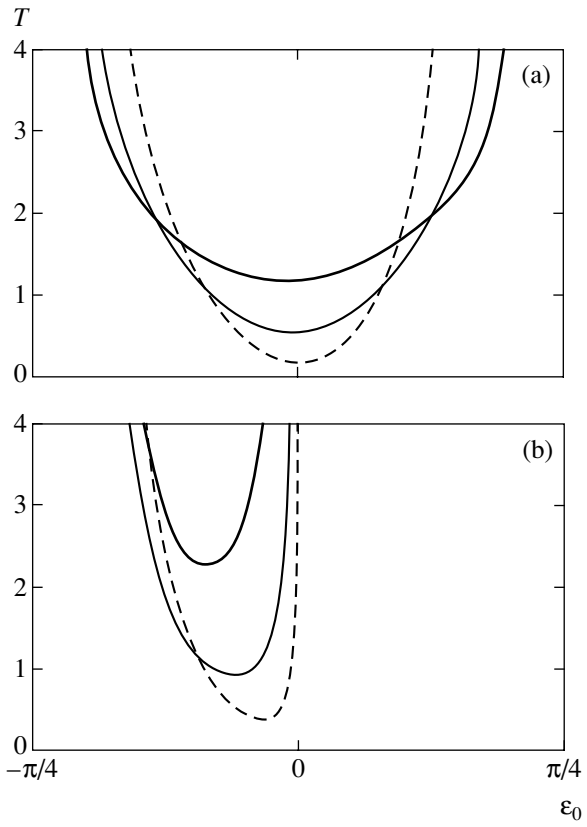

Fig. 4. Temperature in units of $\hbar\gamma S_0/k_B$ as a function of ellipticity ε_0 of light waves for different angles: $\theta = \pi/3$ (bold curves), $\pi/5$ (fine curves), and $\pi/10$ (dashed curves); light field detuning $\delta = 2\gamma$ (a) and 0 (b).

For brevity, we have introduced the notation

$$Q = \sqrt{1 - \cos^2(2\varepsilon_0)\cos^2\theta}.$$

It should be noted that the coefficients of friction and diffusion in the ε–θ–$\bar{\varepsilon}$ field configuration contain contributions $\xi^{(s)}$ and D_1, which are anomalous in detuning; among other things, this may lead to cooling in the case of exact resonance. The direction of the kinetic process (heating or cooling) in this case is determined by the sign of ellipticity ε_0 of counterpropagating light waves and by angle θ determining the mutual orientation of polarization ellipses. The abnormal contribution to friction remains significant for a nonzero detuning as well. Consequently, heating can be replaced by cooling or vice versa by choosing an appropriate ellipticity of the counterpropagating waves and angle θ. The temperature in the cooling region is a function of ellipticity ε_0 of the light waves and angle θ. Figure 4 shows the dependence of the atomic temperature on the ellipticity of light waves for certain angles θ. It can be seen that the shape of the curves is asymmetric in ellipticity parameter ε_0; however, for a large detuning, this effect becomes less pronounced and the temperature minimum is attained for wave ellipticities close to linear. The characteristic increase in the temperature for ellipticities of light waves close to circular is associated with the specific features of the optical transition $J_g = 1/2 \longrightarrow J_e = 1/2$, in which the friction coefficient is equal to zero in the limiting case of the σ_+–σ_- configuration of the light field and, hence, there is no cooling. Note that friction is present for optical transition with larger values of angular moments in the σ_+–σ_- configuration also (e.g., for the transition $J_g = 1 \longrightarrow J_e = 2$) [7].

6.2. The ε–θ–ε Configuration of the Light Field

In this section, we consider an example of the light field configuration formed by elliptically polarized waves with equal ellipticities and the same directions of rotation for the polarization vectors. In contrast to the previous case, a field with the ε–θ–ε configuration contains only the gradients of intensity and ellipticity of the light field (Fig. 5). In particular cases, when angle θ between the principal axes of the ellipses is zero, we return to the case of a uniformly polarized standing wave; for $\theta = \pm\pi/2$ and $\varepsilon_0 = 0$, we have a field of the lin $\perp$ lin configuration, in which only the ellipticity gradient differs from zero.

In the first order in the field intensity, the explicit dependence of the force on polarization parameters ε_0 and θ has the form

$$F(z) = 2\hbar k\delta S_0 \times \frac{\sin(2kz)\cos^2(2\varepsilon_0)(\cos\theta + \cos(2kz))}{1 + \cos(2kz)\cos\theta - \sin(2\varepsilon_0)\sin\theta\sin(2kz)}, \quad (62)$$

where the saturation parameter S_0 is defined in terms of amplitude (53) of the counterpropagating waves. It follows hence that the force averaged over a spatial period differs from zero (effect of rectification of the dipole force):

$$\langle F\rangle = \hbar k\delta S_0 \times \frac{\cos^2(2\varepsilon_0)\sin(2\varepsilon_0)\sin(2\theta)(1 - |\sin\theta|\cos(2\varepsilon_0))^2}{(\cos^2\theta + \sin^2(2\varepsilon_0)\sin^2\theta)^2}. \quad (63)$$

The mean force is an odd function of the light field detuning; in addition, it is an odd function of angle θ and ellipticity ε_0 of the light waves, which is also observed in the general case of an arbitrary field intensity. It can be proved rigorously proceeding from the general symmetry relations for the optical Bloch equation that such a form of the dependence of the force on parameters δ, ε_0, and θ in the ε–θ–ε field configuration takes place for arbitrary optical transitions $J_g \longrightarrow J_e$. Note that the rectified force vanishes in the case of linear and circular polarizations $\varepsilon_0 = 0, \pm\pi/4$ for any θ as

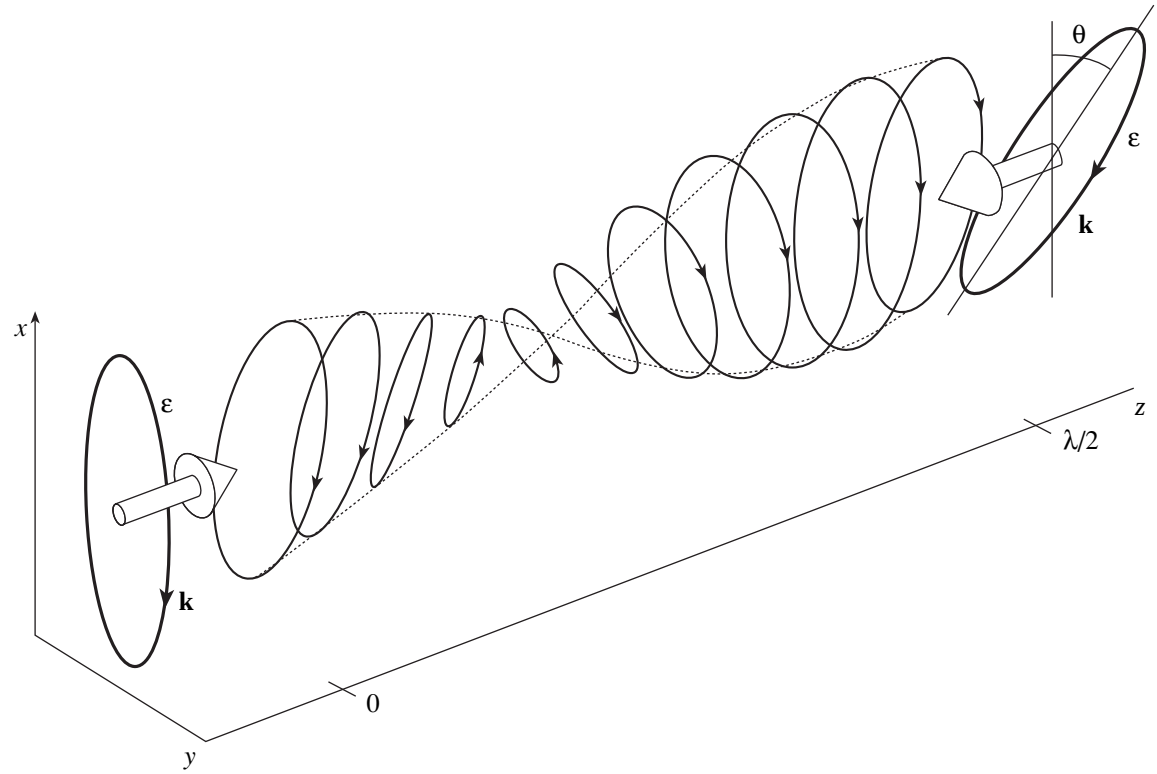

Fig. 5. Spatial configuration ε–θ–ε of a light field. The field is produced by counterpropagating waves of the same ellipticity.

well as for angles $\theta = 0, \pm\pi/2$ in the case of arbitrary ε_0. This becomes obvious if we use the symmetry relations for the force acting on a stationary atom. Reflection in the *xy* plane gives the following relation for the force:

$$\langle F\rangle(\varepsilon_0, \theta) = -\langle F\rangle(\varepsilon_0, -\theta). \quad (64)$$

It can be seen that, for the circular polarization of the waves ($\varepsilon_0 = \pm\pi/4$) or for angles $\theta = 0, \pm\pi/2$, the reflection in the *xy* plane leads to the initial configuration of the light field and, hence, $\langle F\rangle = 0$. Another symmetry relation for the force can be obtained using two consecutive reflections in the *xy* and *xz* planes:

$$\langle F\rangle(\varepsilon_0, \theta) = -\langle F\rangle(-\varepsilon_0, \theta). \quad (65)$$

If the field is produced by waves with the linear polarization ($\varepsilon_0 = 0$), after the spatial reflections mentioned above, we return to the initial configuration of the light field; i.e., $\langle F\rangle(\varepsilon_0 = 0, \theta) = 0$ for any θ.

The effect of rectification of the dipole force and the mechanism for the emergence of this effect was considered in detail in our earlier publication [17]. We only note here that the mechanism for the emergence of the constant component of the dipole force in an ε–θ–ε field is associated with spatially inhomogeneous optical pumping and with the presence of the gradients of the light field intensity and ellipticity.

The linear (in velocity) correction to force (62) can be defined in terms of the friction coefficient; in the case of a low intensity of the light field ($S_0 \ll 1$), the friction coefficient has the form

$$\begin{aligned}\xi(z) = &-\frac{6\hbar k^2\delta}{\gamma}\sin\theta\cos^2(2\varepsilon_0)(\cos(kz)+\cos\theta)\\ &\times(\cos(kz)\sin\theta+\cos\theta\sin(2\varepsilon_0)\sin(kz))\\ &\times[1+\cos\theta\cos(kz)-\sin\theta\sin(2\varepsilon_0)\sin(kz)]^{-3}.\end{aligned} \quad (66)$$

The friction coefficient averaged over the spatial period for $\theta \neq 0, \pi$ and for field polarizations other than circular, $\varepsilon_0 \neq \pm\pi/4$,

$$\langle\xi\rangle = -\frac{3\hbar k^2\delta}{\gamma|\sin\theta\cos(2\varepsilon_0)|}, \quad (67)$$

determines the correction to the force for atoms performing above-barrier motion in the potential created by force (62). Figure 6 shows that the rectified force produces an optical potential in which cold atoms can be localized. It is important to note that, with the appropriate choice of the sign of detuning δ, the force of friction in the vicinity of the minima of potential U_{eff} can lead to cooling of atoms. The latter circum-

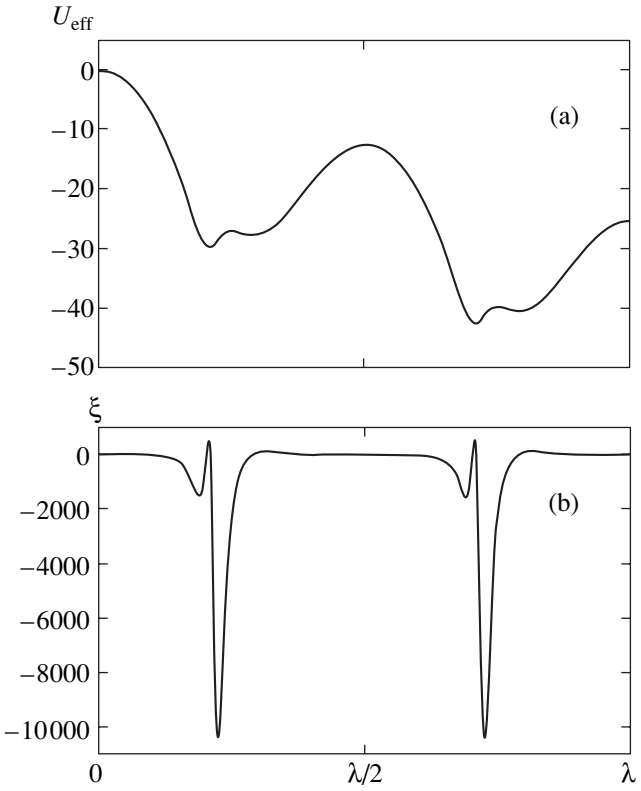

Fig. 6. Spatial dependence of effective potential U_{eff} in units of $\hbar\gamma$ (a) and the friction coefficient in units of $\hbar k^2$ (b) for $\Omega_0 = 100\gamma$ and for parameters optimal for rectification ($\varepsilon_0 = 0.46$, $\theta = 0.51$, $\delta = 62\gamma$).

stance plays an important role for a stable localization of atoms.

6. CONCLUSIONS

We have derived analytic expressions for the kinetic coefficients (the force acting on a stationary atom and coefficients of friction and diffusion) of the Fokker–Planck equation for atoms performing the optical transition $J_g = 1/2 \longrightarrow J_e = 1/2$ in a field of an arbitrary configuration and intensity. The expressions are analyzed for the fields produced by elliptically polarized waves. We have considered two important types of light fields:

(i) fields with a uniform polarization, in which the gradients of ellipticity and spatial orientation are equal to zero and

(ii) nonuniformly polarized fields produced by elliptically polarized waves.

It was shown that, in the first case, in the absence of gradients of the light field polarization, the kinetics of atoms resembles the results obtained in the two-level model of atoms (in particular, our results coincide with the results of this model in the case of the linear ellipticity of the light field). In addition, it was found that our results differ from the predictions of the model of a two-level atom in the case when a uniform polarization of the light field does not coincide with the linear polarization. For light fields of the second type, the effects resulting from the ellipticity of traveling waves that produce the field were demonstrated. The two most interesting configurations were singled out: the ε–θ–$\bar{\varepsilon}$ field created by counterpropagating waves of the opposite ellipticities and the ε–θ–ε field created by counterpropagating waves with the same ellipticity (θ is the angle of mutual orientation of the polarization ellipses for the counterpropagating waves). For example, in the ε–θ–$\bar{\varepsilon}$ field, the ellipticity of traveling waves may lead to sub-Doppler cooling of atoms even in the case of exact resonance ($\delta = 0$); in addition, anomalous (relative to conventional sub-Doppler cooling) regions of heating and cooling are formed. It was shown that, in the field of the ε–θ–ε configuration, the ellipticity of traveling waves leads to rectification of the dipole force. Note that the optical transition $J_g = 1/2 \longrightarrow J_e = 1/2$ chosen for our analysis is a simple example of transition with energy levels degenerate in the angular momentum projection. It should also be observed that the effects described in Section 6 take place for transitions with larger values of angular momenta J_g and J_e as well.

ACKNOWLEDGMENTS

This study was supported by the Russian Foundation for Basic Research (project nos. 01-02-17036, 01-02-17744, 03-02-06047-mas, and 04-02-16488), the Ministry of Education of the Russian Federation (grant no. UR.01.01.060), and the Special Federal Science and Engineering Program "Research and Development Works in Priority Trends of Science and Engineering" for 2002–2006 (state contract no. 01-40-01-06-05).

APPENDIX A

Procedure of Reduction of the Quantum-Mechanical Kinetic Equation for the Density Matrix to the Fokker-Planck Equation for Slow Atoms

The expansion of the kinetic equation for the density matrix in recoil parameter $\hbar k/\Delta p$ has the form

$$\frac{d}{dt}\hat{\rho}(\mathbf{r},\mathbf{p}) = \hat{L}^{(0)}\{\hat{\rho}(\mathbf{r},\mathbf{p})\} + \hbar k\sum_i \frac{\partial}{\partial p_i}\hat{L}_i^{(1)}\{\hat{\rho}(\mathbf{r},\mathbf{p})\} + (\hbar k)^2\sum_{i,j}\frac{\partial^2}{\partial p_i \partial p_j}\hat{L}_{ij}^{(2)}\{\hat{\rho}(\mathbf{r},\mathbf{p})\} + \ldots, \quad \text{(A.1)}$$

where $(\mathbf{r}, \mathbf{p})$ are the coordinates of a point in the phase space. In the zeroth order in recoil effects, the evolution

of the density matrix is determined by the operator

$$\hat{L}^{(0)}\{\hat{\rho}\} = -\left[\left(\frac{\gamma}{2} - i\delta\right)\hat{\Pi}_e\hat{\rho} + \left(\frac{\gamma}{2} + i\delta\right)\hat{\rho}\hat{\Pi}_e\right] - \frac{i}{\hbar}[\hat{V}(\mathbf{r}), \hat{\rho}] + \gamma(2J_e + 1)\sum_q \hat{D}_q^\dagger\hat{\rho}\hat{D}_q, \quad \text{(A.2)}$$

corresponding to the optical Bloch equation (15). The first order terms,

$$\hat{L}_i^{(1)}\{\hat{\rho}\} = -\frac{1}{2\hbar k}(\hat{F}_i(\mathbf{r})\hat{\rho} + \hat{\rho}\hat{F}_i(\mathbf{r})), \quad \text{(A.3)}$$

can be expressed in terms of the force operator (14). Second-order corrections contain both induced terms proportional to the second derivative of operator $\hat{V}(\mathbf{r})$ with respect to the coordinate and the term associated with the recoil effect in spontaneous emission:

$$\hat{L}_{ij}^{(2)}\{\hat{\rho}\} = -\frac{i}{8\hbar k^2}[\nabla_i\hat{F}_j(\mathbf{r}), \hat{\rho}] + \frac{\gamma(2J_e + 1)}{5} \times \sum_{m,n=1,2,3}\left(\delta_{ij}\delta_{nm} - \frac{1}{4}(\delta_{in}\delta_{jm} + \delta_{im}\delta_{jn})\right)\hat{D}_m^\dagger\hat{\rho}\hat{D}_n. \quad \text{(A.4)}$$

In order to derive Eq. (10), we make use of the existence of characteristic time τ. Over time intervals $t \le \tau$, only the change in the internal degrees of freedom of an atom, which strongly depend on the initial conditions, is significant, while time intervals $t \gg \tau$ correspond to the kinetic stage of evolution; accordingly, a change in the internal states of an atom is matched with the change in distribution function $\mathcal{W}$ [2]. Thus, at the kinetic stage of evolution, the density matrix is a linear functional of the distribution function and can be represented in the form

$$\hat{\rho}(\mathbf{r}, \mathbf{p}, t) = \hat{\sigma}(\mathbf{r}, \mathbf{p})\mathcal{W}(\mathbf{r}, \mathbf{p}, t) + \hat{\chi}(\mathbf{r}, \mathbf{p}, t). \quad \text{(A.5)}$$

Here, $\hat{\sigma}(\mathbf{r}, \mathbf{p})$ is the reduced density matrix describing the stationary state of the internal degrees of freedom of an atom moving with velocity $\mathbf{v}$,

$$\frac{d}{dt}\hat{\sigma}(\mathbf{r}, \mathbf{p}) = \hat{L}^{(0)}\{\hat{\sigma}(\mathbf{r}, \mathbf{p})\}, \quad \frac{\partial}{\partial t}\hat{\sigma}(\mathbf{r}, \mathbf{p}) = 0, \quad \text{(A.6)}$$

with the normalization condition $\mathrm{Tr}\{\hat{\sigma}(\mathbf{r}, \mathbf{p})\} = 1$. Closed optical transitions exhibit the property $\mathrm{Tr}\{\hat{L}^{(0)}(\hat{\rho})\} = 0$, indicating the conservation of population. Consequently, in the zeroth order in recoil, we have

$$\frac{d}{dt}\mathrm{Tr}\{\hat{\rho}\} = \frac{d}{dt}\mathcal{W} = \mathrm{Tr}\{\hat{L}^{(0)}\{\hat{\rho}\}\} = 0, \quad \text{(A.7)}$$

i.e., the derivative of the distribution function is at least of the first order of smallness in recoil, and the zeroth order $\hat{\rho}^{(0)}(\mathbf{r}, \mathbf{p}, t)$ can be written in the form

$$\hat{\rho}^{(0)}(\mathbf{r}, \mathbf{p}, t) = \hat{\sigma}(\mathbf{r}, \mathbf{p})\mathcal{W}^{(0)}(\mathbf{r}, \mathbf{p}, t). \quad \text{(A.8)}$$

Thus, $\hat{\sigma}\mathcal{F}$ is the part of the density matrix adiabatically tracing the motion and $\hat{\chi}$ is a small nonadiabatic correction. Indeed, in the zeroth order in the recoil parameter, the equation for the density matrix has the form

$$\frac{d}{dt}\hat{\sigma}\mathcal{W}^{(0)} + \frac{d}{dt}\hat{\chi}^{(0)} = \hat{L}^{(0)}\{\hat{\sigma}\}\mathcal{W} + \hat{L}^{(0)}\{\hat{\chi}^{(0)}\}. \quad \text{(A.9)}$$

In accordance with Eqs. (A.6) and (A.7), we find that, in the zeroth order in recoil, $\hat{\chi}^{(0)}$ satisfies the equation

$$\frac{d}{dt}\hat{\chi}^{(0)}(\mathbf{r}, \mathbf{p}, t) = \hat{L}^{(0)}\{\hat{\chi}^{(0)}(\mathbf{r}, \mathbf{p}, t)\} \quad \text{(A.10)}$$

with the normalization condition $\mathrm{Tr}\{\hat{\chi}^{(0)}(\mathbf{r}, \mathbf{p}, t)\} = 0$. The solution to this equations over time periods t corresponding to the kinetic stage of the evolution is $\hat{\chi}^{(0)} = 0$; thus, nonadiabatic correction $\hat{\chi}$ contains terms starting only from the first order in recoil.

Taking the trace of expression (A.1) and retaining in it only the terms up to the second order in recoil, we obtain

$$\frac{d}{dt}\mathcal{W} = \hbar k\sum_i\frac{\partial}{\partial p_i}\mathrm{Tr}\{\hat{L}_i^{(1)}\{\hat{\sigma}\}\}\mathcal{W} + (\hbar k)^2\sum_{i,j}\frac{\partial^2}{\partial p_i\partial p_j}\mathrm{Tr}\{\hat{L}_j^{(2)}\{\hat{\sigma}\}\}\mathcal{W} + \hbar k\sum_i\frac{\partial}{\partial p_i}\mathrm{Tr}\{\hat{L}_i^{(1)}\{\hat{\chi}\}\}. \quad \text{(A.11)}$$

The last term has a smallness of at least the second order in recoil and contributes to the diffusion coefficient. We will not seek these nonadiabatic corrections to the density matrix as was done, for example, in [15], but will use an alternative approach. To this end, we

supplement Eq. (A.11) with the linear combination of Eq. (A.1),

$$\hbar k\sum_i\frac{\partial}{\partial p_i}\mathrm{Tr}\left\{\hat\varphi_i\frac{d}{dt}\hat\rho\right\} = \hbar k\sum_i\frac{\partial}{\partial p_i}\mathrm{Tr}\{\hat\varphi_i\hat L^{(0)}\{\hat\rho\}\} + (\hbar k)^2\sum_{i,j}\frac{\partial^2}{\partial p_i\partial p_j}\mathrm{Tr}\{\hat\varphi_i\hat L_j^{(1)}\{\hat\rho\}\} + \ldots, \tag{A.12}$$

where $\hat\varphi_i$ is a matrix of dimension $(2J_e + 1) + (2J_g + 1)$. The left-hand side of this equality has the form

$$\hbar k\sum_i\frac{\partial}{\partial p_i}\mathrm{Tr}\left\{\hat\varphi_i\frac{d}{dt}\hat\rho\right\} = \hbar k\sum_i\frac{\partial}{\partial p_i}\mathrm{Tr}\left\{\hat\varphi_i\frac{d}{dt}\hat\sigma\right\}\mathcal{W} + \hbar k\sum_i\frac{\partial}{\partial p_i}\mathrm{Tr}\{\hat\varphi_i\hat\sigma\}\frac{d}{dt}\mathcal{W} + \hbar k\sum_i\frac{\partial}{\partial p_i}\mathrm{Tr}\left\{\hat\varphi_i\frac{d}{dt}\hat\chi\right\}. \tag{A.13}$$

The last term contains the second order in recoil. Over time periods $t \gg \tau$ corresponding to the kinetic stage of evolution, $d\hat\chi/dt = \mathbf{v}\cdot\nabla_{\mathbf{r}}\hat\chi$; consequently, the contribution from this term to the diffusion coefficient can be neglected since only zero-order terms in velocity are retained in the diffusion coefficient for slow atoms. Subtracting expression (A.11) from (A.12) and retaining terms up to the second order in recoil, we ultimately obtain

$$\frac{d}{dt}\mathcal{W} - \hbar k\sum_i\frac{\partial}{\partial p_i}\mathrm{Tr}\{\hat\varphi_i\mathbf{v}\cdot\nabla\hat\sigma\}\mathcal{W} - \hbar k\sum_i\frac{\partial}{\partial p_i}\mathrm{Tr}\{\hat\varphi_i\hat\sigma\}\frac{d}{dt}\mathcal{W} = \hbar k\sum_i\frac{\partial}{\partial p_i}\mathrm{Tr}\{\hat L_i^{(1)}\{\hat\sigma\}\}\mathcal{W} - \hbar k\sum_i\frac{\partial}{\partial p_i}\mathrm{Tr}\{\hat\varphi_i\hat L^{(0)}\{\hat\sigma\}\}\mathcal{W} + (\hbar k)^2\sum_{i,j}\frac{\partial^2}{\partial p_i\partial p_j}\mathrm{Tr}\{\hat L_{ij}^{(2)}\{\hat\sigma\}\}\mathcal{W} - (\hbar k)^2\sum_{i,j}\frac{\partial^2}{\partial p_i\partial p_j}\mathrm{Tr}\{\hat\varphi_i\hat L_j^{(1)}\{\hat\sigma\}\}\mathcal{W} + \hbar k\sum_i\frac{\partial}{\partial p_i}\mathrm{Tr}\{\hat L_i^{(1)}\{\hat\chi\}\} - \hbar k\sum_i\frac{\partial}{\partial p_i}\mathrm{Tr}\{\hat\varphi_i\hat L^{(0)}\{\hat\chi\}\}. \tag{A.14}$$

Note that by choosing matrix $\hat\varphi_i$ appropriately, we can ensure the compensation of the contributions containing nonadiabatic correction $\hat\chi$ so that

$$\mathrm{Tr}\{\hat\varphi_i\hat L^{(0)}\{\hat\chi\}\} = \mathrm{Tr}\{\hat L_i^{(1)}\{\hat\chi\}\}. \tag{A.15}$$

This condition leads to the equation for density matrix $\hat\varphi_i$ (21). Matrix $\hat\varphi$ is a solution to nonhomogeneous linear equation (21) with a source in the form of the fluctuation of the force operator (20). Since the force acting on an atom should be defined in the zeroth and first orders in velocity, while the diffusion coefficient should be defined in the zeroth approximation, it is sufficient to retain in expression (A.14) only the zeroth order (in velocity) of matrix $\hat\sigma$ ($\mathbf{v} = 0$, $\mathbf{r}$) satisfying optical Bloch equation (15) for stationary density matrix $L(0)\{\hat\sigma\} = 0$.

APPENDIX B

Components of Friction and Diffusion Coefficients

The diagonal elements of the friction coefficient components are functions of only odd powers of detuning δ,

$$\chi_{\Lambda\Lambda} = \frac{2\delta}{\gamma}\frac{G_\varepsilon}{(1+G_\varepsilon)^3}\times\left(\frac{\gamma^2}{\gamma^2/4+\delta^2}\frac{1-G_\varepsilon}{2} - G_\varepsilon^2 + 3\sin^2(2\varepsilon)\right), \tag{B.1}$$

$$\chi_{\varepsilon\varepsilon} = -\frac{2\delta}{\gamma}\frac{1}{(1+G_\varepsilon)^3}\left\{3\cos^2(2\varepsilon) + G_\varepsilon\left[10 - 3\cos^2(2\varepsilon) - \frac{\gamma^2}{\gamma^2/4+\delta^2}\tan^2(2\varepsilon)\right] + G_\varepsilon^2\left[\frac{6-\cos^2(2\varepsilon)}{\cos^2(2\varepsilon)} + \frac{\gamma^2}{\gamma^2/4+\delta^2}\tan^2(2\varepsilon)\right] + G_\varepsilon^3\frac{2-\cos^2(2\varepsilon)}{\cos^2(2\varepsilon)}\right\}, \tag{B.2}$$

$$\chi_{\Phi\Phi} = \frac{\delta\gamma}{\gamma^2/4+\delta^2}\frac{G_\varepsilon}{(1+G_\varepsilon)^2}, \tag{B.3}$$

$$\chi_{\phi\phi} = 0. \tag{B.4}$$

We will write the nondiagonal elements associated with correlations of gradients of the amplitude and phase,

$$\chi_{\Lambda\Phi} = \frac{G_\varepsilon}{(1+G_\varepsilon)^2}\left(G_\varepsilon + \frac{\gamma^2/4-\delta^2}{\gamma^2/4+\delta^2}\right), \tag{B.5}$$

$$\chi_{\Phi\Lambda} = \frac{G_\varepsilon}{(1+G_\varepsilon)^3}\times\left(\frac{\delta^2-\gamma^2/4}{\gamma^2/4+\delta^2} + \frac{G_\varepsilon\gamma^2/2}{\gamma^2/4+\delta^2} - 3\sin^2(2\varepsilon)\right), \tag{B.6}$$

ellipticity and phase,

$$\chi_{\varepsilon\Phi} = \frac{G_\varepsilon \tan(2\varepsilon)}{(1+G_\varepsilon)^2}\left(\frac{\delta^2-\gamma^2/4}{\delta^2+\gamma^2/4}-G_\varepsilon\right), \tag{B.7}$$

$$\chi_{\Phi\varepsilon} = \frac{\tan(2\varepsilon)}{(1+G_\varepsilon)^3} \times\left\{1+G_\varepsilon+(1-G_\varepsilon)\left[\frac{\gamma^2 G_\varepsilon}{\gamma^2/4+\delta^2}-3\sin^2(3\varepsilon)\right]\right\}, \tag{B.8}$$

ellipticity and angle,

$$\chi_{\varepsilon\phi} = \frac{G_\varepsilon}{\cos(2\varepsilon)(1+G_\varepsilon)^2} \times\left(2\sin^2(2\varepsilon)\frac{\delta^2}{\gamma^2/4+\delta^2}-1-G_\varepsilon\right), \tag{B.9}$$

$$\chi_{\phi\varepsilon} = 3\frac{\cos(2\varepsilon)}{(1+G_\varepsilon)^2}\left(G_\varepsilon-1-\frac{2G_\varepsilon}{\cos^2(2\varepsilon)}\right), \tag{B.10}$$

ellipticity and amplitude,

$$\chi_{\varepsilon\Lambda} = \frac{2\delta G_\varepsilon \tan(2\varepsilon)}{\gamma(1+G_\varepsilon)^3} \times\left(3\cos^2(2\varepsilon)+G_\varepsilon^2+\frac{3\delta^2 G_\varepsilon+(5G_\varepsilon-2)\gamma^2/4}{\gamma^2/4+\delta^2}\right), \tag{B.11}$$

$$\chi_{\Lambda\varepsilon} = \frac{2\delta\tan(2\varepsilon)}{\gamma(1+G_\varepsilon)^3}\left(G_\varepsilon^2-3\cos^2(2\varepsilon)(1-G_\varepsilon) +\frac{2G_\varepsilon^2(\gamma^2/4-\delta^2)-G_\varepsilon(7\delta^2+11\gamma^2/4)}{\gamma^2/4+\delta^2}\right), \tag{B.12}$$

amplitude and angle,

$$\chi_{\Lambda\phi} = -2\frac{\sin(2\varepsilon)G_\varepsilon}{(1+G_\varepsilon)^2}\frac{\delta^2}{(\gamma^2/4+\delta^2)}, \tag{B.13}$$

$$\chi_{\phi\Lambda} = 3\frac{\sin(2\varepsilon)G_\varepsilon}{(1+G_\varepsilon)^2}, \tag{B.14}$$

and phase and angle,

$$\chi_{\Phi\phi} = \frac{\sin(2\varepsilon)G_\varepsilon}{(1+G_\varepsilon)^2}\frac{\delta\gamma}{\gamma^2/4+\delta^2}, \tag{B.15}$$

$$\chi_{\phi\Phi} = 0. \tag{B.16}$$

The diagonal elements of the diffusion coefficient components $\mathcal{D}_{\beta\beta'}$, which are even functions of δ, have the form

$$\mathcal{D}_{\Lambda\Lambda} = \frac{1}{4\gamma}\frac{G_\varepsilon}{(1+G_\varepsilon)^3}\left\{\gamma^2+12\delta^2\sin^2(2\varepsilon) +G_\varepsilon\left[-\gamma^2+\frac{\gamma^4}{\gamma^2/4+\delta^2}+(4\delta^2+\gamma^2)\tan^2(2\varepsilon)\right] +G_\varepsilon^2[3\gamma^2+2(4\delta^2+\gamma^2)\tan^2(2\varepsilon)]+G_\varepsilon^2\frac{(4\delta^2+\gamma^2)}{\cos^2(2\varepsilon)}\right\}, \tag{B.17}$$

$$\mathcal{D}_{\varepsilon\varepsilon} = \frac{1}{4\gamma}\frac{G_\varepsilon}{(1+G_\varepsilon)^3}\left\{\gamma^2\frac{2-\cos^2(2\varepsilon)}{\cos(2\varepsilon)}+12\delta^2\cos(2\varepsilon) +G_\varepsilon\left[7(4\delta^2+\gamma^2)+\gamma^2\frac{1-5\cos^2(2\varepsilon)}{\cos^2(2\varepsilon)} +\frac{\gamma^4}{\gamma^2/4+\delta^2}\tan^2(2\varepsilon)\right]+G_\varepsilon^2\left[(4\delta^2+\gamma^2)\frac{3+2\cos^2(2\varepsilon)}{\cos^2(2\varepsilon)} +\gamma^2\frac{1-3\cos^2(2\varepsilon)}{\cos^2(2\varepsilon)}\right]+G_\varepsilon^3\frac{4\delta^2+\gamma^2}{\cos^2(2\varepsilon)}\right\}, \tag{B.18}$$

$$\mathcal{D}_{\Phi\Phi} = \frac{\gamma}{4}\frac{G_\varepsilon}{(1+G_\varepsilon)^3} \times\left[(1+G_\varepsilon)^2+3\sin^2(2\varepsilon)+G_\varepsilon\left(1-\frac{\gamma^2}{\gamma^2/4+\delta^2}\right)\right], \tag{B.19}$$

$$\mathcal{D}_{\phi\phi} = \frac{\gamma}{2}\frac{G_\varepsilon}{1+G_\varepsilon}. \tag{B.20}$$

It was noted above that components $\mathcal{D}_{\beta\beta'}$ are symmetric in indices $\beta\beta'$ and, hence, the expression for diffusion coefficient (25) will contain only symmetric sums $\mathcal{D}_{\beta\beta'}+\mathcal{D}_{\beta'\beta}$. These are two normal nondiagonal elements even in detuning,

$$\mathcal{D}_{\varepsilon\Lambda}+\mathcal{D}_{\Lambda\varepsilon} = \frac{\sin(4\varepsilon)}{4\gamma}\frac{G_\varepsilon}{(1+G_\varepsilon)^3} \times\left[12\delta^2-\frac{\gamma^2}{\cos^2(2\varepsilon)}+G_\varepsilon\frac{2}{\cos^2(2\varepsilon)} \times\left(8\delta^2+\gamma^2-\frac{\gamma^4/2}{\gamma^2/4+\delta^2}\right)+G_\varepsilon^2\frac{8\delta^2-\gamma^2}{\cos^2(2\varepsilon)}\right], \tag{B.21}$$

$$\mathcal{D}_{\Phi\phi}+\mathcal{D}_{\phi\Phi} = -\frac{\gamma\sin(2\varepsilon)G_\varepsilon}{(1+G_\varepsilon)^2}. \tag{B.22}$$

Finally, we write four anomalous nondiagonal elements leading to asymmetry in the dependence of the diffu-

sion coefficient on δ:

$$\mathcal{D}_{\Phi\Lambda} + \mathcal{D}_{\Lambda\Phi} = \frac{\delta G_\varepsilon}{(1+G_\varepsilon)^3} \times\left(-3\sin^2(2\varepsilon) + G_\varepsilon\frac{\gamma^2}{\gamma^2/4+\delta^2} + G_\varepsilon^2\right), \quad \text{(B.23)}$$

$$\mathcal{D}_{\phi\Lambda} + \mathcal{D}_{\Lambda\phi} = \frac{2\delta\sin(2\varepsilon)G_\varepsilon}{(1+G_\varepsilon)^2}, \quad \text{(B.24)}$$

$$\mathcal{D}_{\Phi\varepsilon} + \mathcal{D}_{\varepsilon\Phi} = \frac{-\delta\tan(2\varepsilon)G_\varepsilon}{(1+G_\varepsilon)^3} \times\left[3\cos^2(2\varepsilon) + G_\varepsilon\left(3+\frac{\gamma^2}{\gamma^2/4+\delta^2}\right) + G_\varepsilon^2\right], \quad \text{(B.25)}$$

$$\mathcal{D}_{\phi\varepsilon} + \mathcal{D}_{\varepsilon\phi} = \frac{2\delta G_\varepsilon}{(1+G_\varepsilon)^2}\frac{\cos^2(2\varepsilon)+G_\varepsilon}{\cos(2\varepsilon)}. \quad \text{(B.26)}$$

REFERENCES

1. A. P. Kazantsev, G. I. Surdutovich, and V. P. Yakovlev, *The Mechanical Action of Light on Atoms* (Nauka, Moscow, 1991).
2. V. G. Minogin and V. S. Letokhov, *The Pressure of Laser Radiation on Atoms* (Nauka, Moscow, 1986).
3. P. D. Lett, R. N. Watts, S. I. Westbrook, *et al.*, Phys. Rev. Lett. **61**, 169 (1988).
4. A. Aspect, E. Arimondo, R. Kaiser, *et al.*, Phys. Rev. Lett. **61**, 826 (1988).
5. *Laser Cooling and Trapping of Atoms*, Special issue of J. Opt. Soc. Am. B **6**, 11 (1989).
6. *Laser Cooling and Trapping*, Special issue of Laser Phys. **4**, 5 (1994).
7. J. Dalibard and C. Cohen-Tannoudji, J. Opt. Soc. Am. B **6**, 2023 (1989).
8. O. N. Prudnikov, A. V. Taĭchenachev, A. M. Tumaĭkin, and V. I. Yudin, Zh. Éksp. Teor. Fiz. **115**, 791 (1999) [JETP **88**, 433 (1999)].
9. J. Dalibard and C. Cohen-Tannoudji, J. Phys. B **18**, 1661 (1985).
10. J. Javanainen, Phys. Rev. A **44**, 5857 (1991).
11. V. S. Smirnov, A. M. Tumaĭkin, and V. I. Yudin, Zh. Éksp. Teor. Fiz. **96**, 1613 (1989) [Sov. Phys. JETP **69**, 913 (1989)]; A. V. Taĭchenachev, A. M. Tumaĭkin, and V. I. Yudin, Zh. Éksp. Teor. Fiz. **110**, 1727 (1996) [JETP **83**, 949 (1996)]; G. Nienhuis, A. V. Taichenachev, A. M. Tumaikin, and V. I. Yudin, Europhys. Lett. **44**, 20 (1998); A. V. Taĭchenachev, A. M. Tumaĭkin, and V. I. Yudin, Zh. Éksp. Teor. Fiz. **118**, 77 (2000) [JETP **91**, 67 (2000)].
12. A. V. Taĭchenachev, A. M. Tumaĭkin, V. I. Yudin, and G. Nienhuis, Zh. Éksp. Teor. Fiz. **108**, 415 (1995) [JETP **81**, 224 (1995)].
13. A. V. Bezverbnyi, G. Nienhuis, and A. M. Tumaikin, Opt. Commun. **148**, 151 (1998).
14. A. Yu. Pusep, Zh. Éksp. Teor. Fiz. **70**, 851 (1976) [Sov. Phys. JETP **43**, 441 (1976)].
15. J. Dalibard and C. Cohen-Tannoudji, J. Opt. Soc. Am. B **2**, 1707 (1985).
16. V. Finkelstein, P. R. Berman, and J. Guo, Phys. Rev. A **45**, 1829 (1992).
17. O. N. Prudnikov, A. V. Taĭchenachev, A. M. Tumaĭkin, and V. I. Yudin, Zh. Éksp. Teor. Fiz. **120**, 76 (2001) [JETP **93**, 63 (2001)].

Translated by N. Wadhwa

Journal of Experimental and Theoretical Physics, Vol. 98, No. 3, 2004, pp. 455–462.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 125, No. 3, 2004, pp. 518–526.

ATOMS, SPECTRA, RADIATION

Spherical Electron–Ion Systems in Semiclassical Approximation: Atoms and Atomic Clusters

G. V. Shpatakovskaya
Institute of Mathematical Modeling, Russian Academy of Sciences, Moscow, 125047 Russia
e-mail: shpat@imamod.ru
Received August 15, 2003

Abstract—A semiclassical method for calculating the total energy and spatial distribution of electron density in spherically symmetric electron–ion systems is applied to atoms and both solid and hollow atomic clusters. Both exchange–correlation interaction and second-order gradient correction are taken into account. The contribution due to the fourth-order gradient correction is discussed. An expression is proposed for the oscillating correction to the averaged electron density. An expression is obtained for the equilibrium radius of a hollow cluster. The dependence of the equilibrium radius of an endohedral cluster on the valence of the central atom is analyzed. © *2004 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

Spherically symmetric electron–ion systems with fixed distributions of ions can be used to model atoms, ions, neutral and charged atomic clusters, etc. Their characteristics are calculated in the Hartree or Hartree–Fock approximation or by the Kohn–Sham method in density functional theory [1]. However, the complexity of these calculations increases with the number of particles in the system, whereas the simpler and more explicit semiclassical methods become more accurate. In particular, semiclassical approximations have been successfully applied in calculations of thermodynamic properties [2–5] and in analytical treatments of shell effects in the mass spectra of medium-sized and large atomic clusters at zero and finite temperatures [6–8].

In this paper, an improved Thomas–Fermi (ITF) model and corrections to it are used as a basis for a unified algorithm for computing local characteristics (density and potential) and electron energy in a spherically symmetric electron–ion system. The efficiency of the proposed algorithm is demonstrated by applying it to atoms and atomic clusters.

Originally, the semiclassical generalization of the Thomas–Fermi model based on an energy functional allowing for the exchange interaction and the lowest order (gradient) correction was applied to calculate the energy of the electron shell of a free atom in [9]. The energy functional was minimized over the simplest class of trial functions. To develop the method, the authors of [10] proposed a quantum-statistical model in which the electron density was calculated by solving the Euler–Lagrange equation. This model was used to compute an equation of state in the cell approximation for a wide range of density at zero temperature. Its extension to nonzero temperatures was presented in [11].

The semiclassical energy functional on a class of trial density functions has also been applied to describe atomic clusters [12, 13]. In [13], the fourth-order gradient correction was included in the expression for kinetic energy. The authors claimed (referring also to a numerical comparison made in computations of nuclei [14]) that the use of this particular correction made it possible to attain the best accuracy of mean electron density.

The simplest semiclassical functional based on the Thomas–Fermi theory was used in [15] to analyze a hollow cluster as a model of the fullerene C_{60}.

In the present study, a unified semiclassical ITF model is applied to analyze the properties of these and other systems. In Section 2, the general model equation for electron density is written out, including the exchange and correlation interactions and the second-order gradient correction. The model describes average characteristics of the system. However, it can also be used to calculate the oscillations in electron density due to the discrete nature of the electron spectrum. The oscillatory effects are discussed in Section 3, where a corresponding correction to the mean electron density is derived. The results of computations performed for particular atoms, solid and hollow clusters, and fullerenes are presented and discussed in Section 4. Finally, the principal conclusions are summarized.

Some preliminary results of this study were published in [16, 17].

2. EQUATION FOR ELECTRON DENSITY

In the ITF model, the system of N_e electrons occupying a volume V in the field $U_i(\mathbf{r})$ generated by ions at

zero temperature is described by the following energy functional of density $n = n_e(\mathbf{r})$ (in atomic units)

$$E_e[n] = \int d\mathbf{r} \times \left\{ t(n) + n\left[U_i(\mathbf{r}) + \frac{1}{2}U_e(r) \right] + e_{ex}(n) + e_{\text{corr}}(n) \right\}, \quad (1)$$

where $t(n)$ is the kinetic energy density including the second-order correction to the Thomas–Fermi model [2],

$$t(n) = t_{\text{TF}}(n) + \delta_2 t(n) = \frac{3}{10}(3\pi^2)^{2/3} n^{5/3} - \frac{1}{72}\frac{(\nabla n)^2}{n}; \quad (2)$$

$$U_e(\mathbf{r}) = \int d\mathbf{r}' \frac{n_e(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|}$$

is the potential generated by electrons;

$$e_{\text{ex}}(n) = -\frac{3}{4}\left(\frac{3}{\pi}\right)^{1/3} n^{4/3}, \quad U_{\text{ex}} = -\left(\frac{3}{\pi}\right)^{1/3} n^{1/3},$$

$$X = \frac{r_s}{11.4}, \quad r_s = \left(\frac{3}{4\pi n}\right)^{1/3},$$

$$e_{\text{corr}}(n) = -0.033n \quad (3)$$

$$\times \left[(1 + X^3)\ln(1 + X^{-1}) + \frac{X}{2} - X^2 - \frac{1}{3} \right],$$

$$U_{\text{corr}} = -0.033\ln(1 + X^{-1})$$

are the exchange and correlation energy densities [18] and the corresponding potentials.

The extremum condition for the functional subject to conservation of the number of particles $N_e = \int d\mathbf{r} n_e(\mathbf{r})$ entails the Euler–Lagrange equation for density

$$\frac{1}{2}(3\pi^2 n)^{2/3} + \frac{1}{72}\left(\frac{\nabla n}{n}\right)^2 - \frac{1}{36}\frac{\Delta n}{n} + U(\mathbf{r}) - \mu = 0, \quad (4)$$

where the Lagrange multiplier μ is the chemical potential of the system and the effective potential $U(\mathbf{r})$ is the sum of external, electrostatic, exchange, and correlation terms,

$$U(\mathbf{r}) = U_i(\mathbf{r}) + U_e(\mathbf{r}) + U_{\text{ex}}(\mathbf{r}) + U_{\text{corr}}(\mathbf{r}).$$

Using spherical symmetry and changing from the radius to the dimensionless variable $x = r/R$ (R is a reference length) and from the density to the function $\nu(x) = x\sqrt{n}$, one obtains a nonlinear integrodifferential equation for this function and the chemical potential:

$$\frac{1}{18R^2}\frac{d^2\nu}{dx^2} - \frac{(3\pi^2)^{2/3}}{2} x^{-4/3}\nu^{7/3} + (\mu - U)\nu = 0,$$

$$\nu(0) = \nu(L) = 0,$$

$$N_e = 4\pi R^3 \int_0^L dx' \nu^2(x'), \quad (5)$$

$$U_e(x) = \frac{4\pi R^2}{x}\left(\int_0^L dx' \nu^2(x') - \int_x^L dx' \left(1 - \frac{x}{x'}\right)\nu^2(x') \right).$$

It can be solved by Newton's method in a finite-difference approximation for $0 \le x \le L \gg 1$.

The ITF model can be used to evaluate the contributions due to effects ignored in the model (see [7]). In particular, the contribution to energy due to the fourth-order gradient correction is expressed as follows [9]

$$\Delta_4 E = \frac{1}{540(3\pi^2)^{2/3}} \int d\mathbf{r} n^{1/3} \times \left[\left(\frac{\Delta n}{n}\right)^2 - \frac{9}{8}\left(\frac{\nabla n}{n}\right)^2 \frac{\Delta n}{n} + \frac{1}{3}\left(\frac{\nabla n}{n}\right)^4 \right]. \quad (6)$$

3. OSCILLATING ELECTRON DENSITY IN A SHELL

Solving the Schrödinger equation with the potential $U(r)$ obtained in the ITF approximation, one can find the spectrum and wavefunctions of electrons and use the results to calculate the electron density for a discrete spectrum. In what follows, this is done analytically in a semiclassical approximation, and the expression for the oscillating correction to electron density obtained in [19] for an infinite atom is extended to the case of a finite system. The analysis relies on a semiclassical solution to the Schrödinger equation for the radial wavefunction, the Bohr–Sommerfeld quantization condition, and Poisson's summation formula (see [2, 19] for details). The resulting expression for the oscillating correction is

$$n_{\text{osc}}(r) = \frac{1}{2\pi^2 r^2} \int_{U(r)}^{\mu} dE \int_0^{p_E^2(r) r^2} d\lambda^2 \times \sum_{k,s=-\infty}^{\infty} (-1)^{k+s} \frac{\sin[2\sigma_{E\lambda}(r) + 2s\sigma^0_{E\lambda} + 2\pi k\lambda]}{p_{E\lambda}(r)}. \quad (7)$$

Here,

$$p_{E\lambda}(r) = \sqrt{p_E^2(r) - \frac{\lambda^2}{r^2}}, \quad p_E^2(r) = 2(E - U(r)),$$

$$\sigma_{E\lambda}(r) = \int_r^{R_{E\lambda}} dr' p_{E\lambda}(r')$$

are the radial momentum and action, respectively; $\sigma_{E\lambda}^0 = \sigma_{E\lambda}(R'_{E\lambda})$; $R'_{E\lambda}$ and $R_{E\lambda}$ are the left- and right-hand turning points, respectively; and $\lambda = l + 1/2$ (l is the orbital quantum number).

Performing the integral by parts with respect to energy, retaining the integrated term to the leading order in the semiclassical parameter at $E = \mu$, and changing to the variable

$$t = \sqrt{1 - \frac{\lambda^2}{p_\mu^2(r) r^2}} = \frac{p_{\mu\lambda}(r)}{p_\mu(r)}$$

in the integral with respect to λ^2, one obtains

$$n_{\rm osc}(r) \simeq -\frac{p_\mu(r)}{2\pi^2}\int_0^1 dt \sum_{k,s=-\infty}^{\infty} (-1)^{k+s} \times \frac{\cos[2\sigma_{\mu t}(r) + 2s\sigma_{\mu t}^0 + 2\pi k p_\mu(r) r\sqrt{1-t^2}]}{\tau_{\mu t}(r) + s\tau_{\mu t}^0}, \quad (8)$$

where

$$\tau_{\mu t}(r) = \frac{\partial \sigma_{\mu t}(r)}{\partial \mu} = \int_r^{R_{\mu t}} \frac{dr'}{p_{\mu t}(r')}, \quad \tau_{\mu t}^0 = \tau_{\mu t}(R'_{\mu t})$$

is the classical time. Note that a similar expression for the oscillating correction is obtained if $\sigma_{\mu t}(r)$ and $\tau_{\mu t}(r)$ are replaced by the "complementary" quantities

$$\tilde\sigma_{\mu t}(r) = \sigma_{\mu t}^0 - \sigma_{\mu t}(r), \quad \tilde\tau_{\mu t}(r) = \tau_{\mu t}^0 - \tau_{\mu t}(r).$$

These formulas are used in the computations for atoms presented below.

Expression (8) for $n_{\rm osc}(r)$ contains a double sum and an integral of an oscillating function. Their values are mainly determined by the contributions of the integration limits and the regions of slow variation of the argument of an oscillating function. An analysis shows that the argument of the cosine in (8) varies rapidly when $s \neq k - 1$ for attractive potentials of atomic type ($U(r \longrightarrow 0) \longrightarrow -Z/r$, where Z is the charge number) and when $s \neq 2k$ for anharmonic attractive potentials of cluster type ($U(r \longrightarrow 0) \longrightarrow U(0) = \text{const} < 0$). In particular, its derivative with respect to t goes to infinity as $t \longrightarrow 1$ ($\lambda \longrightarrow 0$). Therefore, only the terms with $s = k - 1$ and $s = 2k$ are summed for atoms and clusters, respectively.

Sums of the form

$$\sum_{k=-\infty}^{\infty} (-1)^{ik} \frac{\cos(x + kx^0)}{\alpha + k}$$

can be calculated analytically (see [20]), with $i = 0$ and 1 for atoms and clusters, respectively. Performing the integral by parts and retaining the integrated term to the leading order in the semiclassical parameter at $t = 1$, one obtains an oscillating correction to the electron density in an atom:

$$n_{\rm osc}(r) \approx \frac{\sin[2\tilde\sigma_\mu(r) + \tilde\alpha((2n+1)\pi - 2\sigma_\mu^0)]}{4\pi r^2 p_\mu(r)\tau_\mu^0 \tilde\delta_\mu(r)\sin(\pi\tilde\alpha)}, \quad (9)$$

where

$$\tilde\sigma_\mu(r) = \int_0^r dr' p_\mu(r'), \quad \sigma_\mu^0 = \tilde\sigma_\mu(R_\mu),$$

$$n \le \frac{\sigma_\mu^0}{\pi} \le n+1, \quad \tilde\alpha = \frac{\tilde\tau_\mu(r)}{\tau_\mu^0},$$

$$\tilde\tau_\mu(r) = \int_0^r \frac{dr'}{p_\mu(r')}, \quad \tau_\mu^0 = \tilde\tau_\mu(R_\mu),$$

$$\tilde\delta_\mu(r) = \int_0^r \frac{dr'}{r'^2}\left[\frac{1}{p_\mu(r')} - \frac{1}{\sqrt{2Z/r'}}\right] - \frac{2}{\sqrt{2Zr}}.$$

In the Thomas–Fermi model of a free atom, the radius is infinite, $\tau_\mu^0 = \infty$, and (9) is identical to the expression obtained in [19]. In the case of a cluster, the oscillating correction is

$$n_{\rm osc}(r) \approx -\frac{\sin[2\sigma_\mu(r) + \alpha(n\pi - 2\sigma_\mu^0)]}{8\pi r^2 p_\mu(r)\tau_\mu^0 \delta_\mu(r)\sin(0.5\pi\alpha)}, \quad (10)$$

where

$$n - \frac{1}{2} \le \frac{2\sigma_\mu^0}{\pi} \le n + \frac{1}{2}, \quad \alpha = \frac{\tau_\mu(r)}{\tau_\mu^0},$$

$$\delta_\mu(r) = \int_r^{R_\mu} \frac{dr'}{r'^2 p_\mu(r')},$$

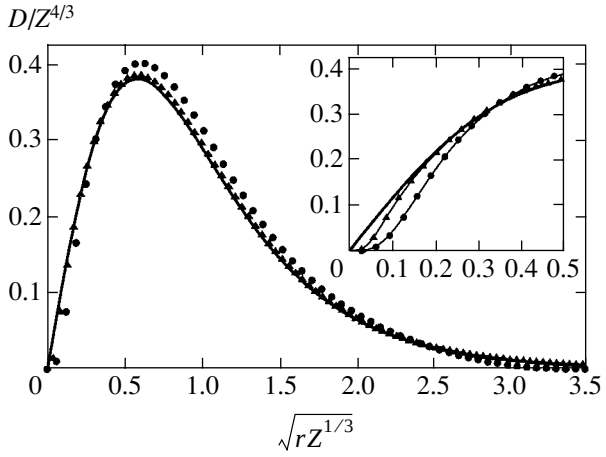

Fig. 1. Radial electron-density distribution $D(r) = 4\pi r^2 n(r)$ in a free atom predicted by the Thomas–Fermi model (solid curves) and by the ITF model with $Z = 10$ (circles) and 80 (triangles).

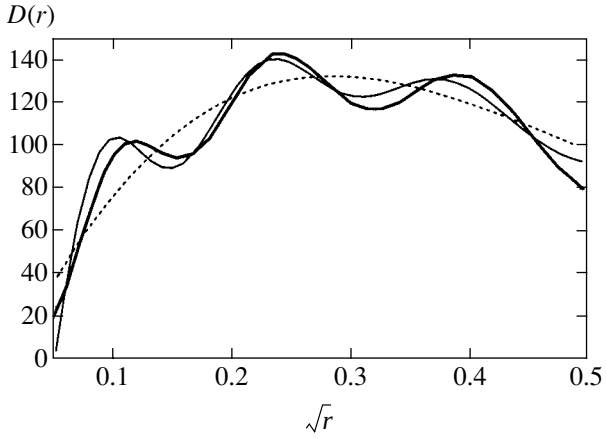

Fig. 2. Radial electron-density distribution in the central region of the mercury atom (without exchange and correlation interactions) predicted by the ITF model (dotted curve), Hartree model (solid curve), and the ITF model with oscillating correction (9) (thin curve).

the integrals in $\sigma_\mu(r)$ and $\tau_\mu(r)$ with respect to r' are taken from r to the right-hand turning point R_μ, $\sigma_\mu^0 = \sigma_\mu(0)$, and $\tau_\mu^0 = \tau_\mu(0)$.

Let us briefly discuss the applicability of the expressions obtained here. Integrating by parts and retaining the integrated term, one obtains an accurate estimate for the integral if the derivative of the argument of the oscillating function (cosine) in (8) with respect to t is sufficiently large. In the present case, this derivative (at $t = 1$) is $2p_\mu^2(r)r^2\tilde{\delta}_\mu(r)$ for an atom and $2p_\mu^2(r)r^2\delta_\mu(r)$ for a cluster. Accordingly, the domains where the estimation procedure used here is not valid are determined by the points where the factors in these products vanish. There exist exactly three such points in the case of an atom, $r = 0$, $r = R_\mu$, and $r = r_0$ ($\tilde{\delta}_\mu(r_0) = 0$), which restricts the applicability of expression (9) to $0 \ll r \ll r_0$. In the case of a cluster, there are two such points, $r = 0$ and $r = R_\mu$, and the interval of applicability is wider.

4. ATOMS, ATOMIC CLUSTERS, HOLLOW CLUSTERS, AND FULLERENES

In this section, the formulas obtained above are applied to specific spherically symmetric systems: medium-sized and large atoms, solid sodium clusters, hollow clusters, and endohedrally doped hollow clusters. The possibility of using a hollow spherical cluster as a model of the fullerene C_{60} is also discussed.

The only difference between the electron–ion systems considered in this section lies in the potential $U_i(\mathbf{r})$ generated by ions.

4.1. Atom

The external field in an atom is generated by a nucleus with charge number Z ($Z = N_e$ in a neutral atom):

$$U_i(\mathbf{r}) = -Z/r.$$

Figure 1 compares the radial electron-density distributions calculated by solving Eq. (5) for neon ($Z = 10$) and mercury ($Z = 80$) with those predicted by the Thomas–Fermi model. It is well known that the model yields a divergent density at the origin, $n_{TF} \propto r^{-3/2}$, and the corresponding radial density distribution behaves as a square root, $D(r \longrightarrow 0) \propto \sqrt{r}$. In the ITF model, as in [10], the effects of second order in the semiclassical parameter are treated in a self-consistent manner, electron density is constant at the origin, and the radial density distribution is a quadratic function of radius ($D(r \longrightarrow 0) \propto r^2$). In the system of units employed here, the Thomas–Fermi model yields the same curve for any Z. The figure demonstrates that the ITF results tend to this universal curve as the charge number Z increases, while their correct behavior near the origin is preserved (see the enlarged inset). This is a good illustration of the well-known validity conditions $Z^{-1/3} \ll 1$ and $r > 1/Z$ for the semiclassical description as applied to an atom.

Figure 2 compares the electron density in the central region of the mercury atom predicted by the ITF model (without exchange and correlation terms) with and without the use of oscillating correction (9) against that calculated in the Hartree approximation [21]. Within its limits of applicability, the oscillating correction accurately describes the oscillation associated with the concentration of electrons in the K, L, and M shells.

The top row in Table 1 shows the results of a self-consistent calculation of the total electron-shell energy for the mercury atom and some of its components, including the exchange energy, the correlation energy,

Table 1. Energy of the electron shell of the mercury atom according to the ITF model

E	E_{ex}	E_{corr}	$\delta_2 E$	$\delta_4 E$	$E - E_{corr} + \delta_4 E$	E_{stat}
-1.9616×10^4	-3.3196×10^2	-1.0086×10^1	1.2772×10^3	7.8599×10^2	-1.8820×10^4	-1.8400×10^4
	-3.2560×10^2	-9.9442×10^0	1.2770×10^3	7.8632×10^2		

and the second-order gradient correction. The bottom row presents the exchange and correlation energies calculated as additive corrections (not self-consistently) by solving Eq. (5) without exchange and correlation terms. It is obvious that the self-consistent and non-self-consistent results are very close. This agrees with the assertion in [7] about the energy correction associated with the small terms ignored in functional (1).

As noted in the Introduction, the calculations performed for nuclei and clusters in [13, 14] took into account the fourth-order gradient correction in a self-consistent manner. This correction can also be estimated for an atom by using expression (6) with the density obtained by solving Eq. (5). The resulting estimates based on the ITF model (with and without exchange and correlation terms) are also presented in Table 1. First, note that the fourth-order correction is smaller than the second-order one by only a factor of 1.5, which implies slow convergence of the expansion. Second, the fourth-order gradient correction cannot be used separately, since it must be combined with quantum-mechanical corrections to the exchange energy, which have yet to be found (see [9, 22]). Third, the self-consistent treatment of the exchange–correlation interaction weakly affects the values of gradient corrections; however, this is not true for clusters (see below).

The last two columns in Table 1 present the values of electron energy obtained by using the ITF model (without the correlation term and with fourth-order gradient correction (6) added) and the well-known semiclassical formula

$$E_{stat} = -0.768745Z^{7/3} + 0.5Z^2 - 0.2699Z^{5/3}. \quad (11)$$

For $Z > 4$, expression (11) corresponds to the dependence of the electron-shell energy on Z in the Hartree–Fock model within fractions of a percent. A comparison of the present results with E_{stat} shows that the error in the calculated energy is reduced from 6.5% to 2.3% by taking into account $\delta_4 E$.

The correspondence between the corrections calculated in this study and the summands in (11) can be explained as follows. The first term in (11) is the energy predicted by the Thomas–Fermi model, the second one is the Scott correction, and the third one is the sum of the exchange energy and a finite part of the second-order gradient correction. Subtracting the exchange energy from the last term in (11) for $Z = 80$, one finds that the finite part of the second-order quantum-mechanical correction is much smaller than the value obtained in the present calculations (fourth column). This is explained by recalling that the Scott correction was shown in [23] to arise from the summation of a divergent series of quantum-mechanical corrections to energy. The divergence is due to the inapplicability of the semiclassical approximation in the central region of an atom. Thus, the large values of $\delta_2 E$ and $\delta_4 E$ given by the ITF model are explained by the contributions of this region that are responsible for the Scott correction.

4.2. Atomic Cluster

In the jellium model of an atomic cluster, N_i ions uniformly distributed over the volume of radius R generate the potential

$$U_i(\mathbf{r}) = \begin{cases} -\dfrac{N_e}{2R}\left[3 - \left(\dfrac{r}{R}\right)^2\right], & r \le R, \\ -\dfrac{N_e}{R}, & r > R, \end{cases}$$

where $N_e = wN_i$ (w is the valence of an atom). The radius of a cluster is related to the number of electrons in it: $R = r_s N_e^{1/3}$, where r_s is the Wigner–Seitz radius.

Figure 3 takes the electron density distribution in the cluster Na_{58} ($r_s = 3.92$, $R(N = 58) = 15.17$) predicted by the ITF model with and without oscillating correction (10) and compares it with the results obtained by the Kohn–Sham method borrowed from [24]. Expression (10) provides a good approximation of the oscillation amplitude and phase outside the neighborhoods of $r = 0$ and $r = R_\mu = 16.5$, in agreement with the analysis presented in Section 3.

Figure 4 compares the total energy per atom

$$\frac{E}{N} = \frac{E_e + E_i}{N}$$

for a sodium cluster with its analog for bulk metal. The energy of ions uniformly distributed over a ball of radius R is $E_i = 0.6N_e^2/R$. The energy calculated in this study agrees (within 1–5%) with the semiclassical results obtained in [13]. The agreement improves with increasing number of particles, i.e., with a decreasing semiclassical parameter (which is proportional to

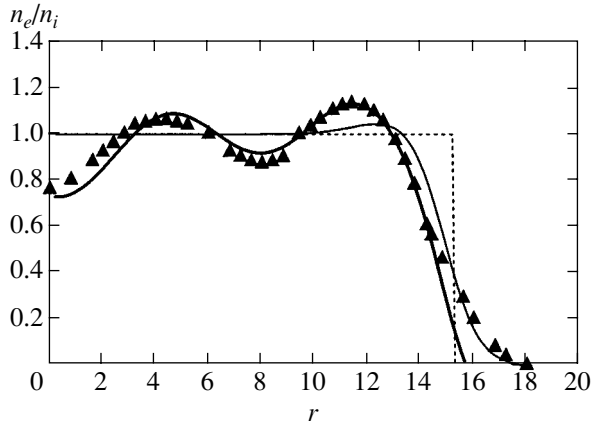

Fig. 3. Relative electron-density distribution in Na_{58} predicted by the ITF models with and without oscillating correction (10) (thick and thin solid curves, respectively) and by the Kohn–Sham method in [24] (symbols). The dotted curve is the distribution of ions.

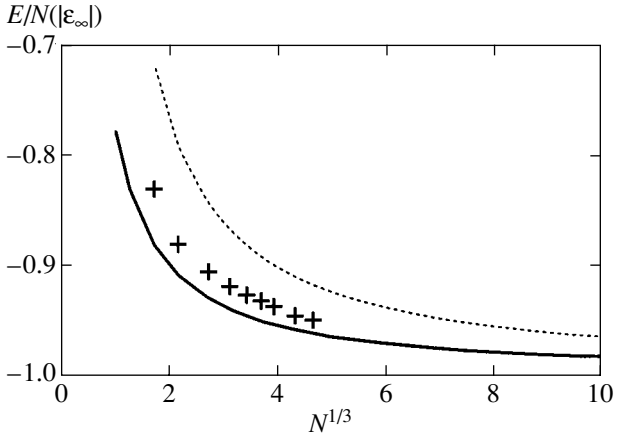

Fig. 4. Total energy per atom in sodium clusters in units of the absolute value of energy per atom in a metal ($|\varepsilon_\infty|$ = 2.252 eV). The dotted and solid curves correspond to the ITF models with and without the fourth-order gradient correction (6), respectively. The symbols were obtained in self-consistent calculations using correction (6) [13].

$N_e^{-1/3}$). Furthermore, Fig. 4 demonstrates that the additive contribution of correction (6) is too large and a better result is obtained without it.

The results presented in Table 2 for the cluster Na_{100} illustrate the contributions to the energy of electrons due to the exchange, correlation, and gradient terms. As in Table 1, the top and bottom rows here show, respectively, the results predicted by the complete ITF model and by the ITF model without exchange–correlation terms. Additive fourth-order correction (6) calculated for clusters according to the complete ITF model is by a factor of 1.5 greater than the contribution of the second-order correction (top row). However, both gradient corrections are reduced and their ratio changes (bottom row) by dropping the exchange–correlation terms in Eq. (5), whereas the exchange and correlation terms calculated self-consistently (top row) and as additive corrections (bottom row) are similar in both cases.

Table 2. Energies of electrons and ions in the Na_{100} cluster

E_e	E_{ex}	E_{corr}	$\delta_2 E$	$\delta_4 E$	E_i
–337.71	–11.254	–3.6959	0.2126	0.36	329.76
	–10.577	–3.5387	0.10552	0.0478	

Thus, exchange and correlation effects are much more important for clusters as compared to free atoms. When they are taken into account, the values of gradient terms are substantially modified and the use of the fourth-order correction leads to unsatisfactory results. Since this correction has no theoretical justification (see Section 4.1), the good results of the self-consistent treatment of the fourth-order correction in [13, 14] should be interpreted as accidental and attributed to the use of a more suitable class of trial functions.

Finally, note that the contribution of the shell-related oscillatory effects to the energy of cluster electrons can be calculated by the method described in [7].

4.3. Hollow Cluster and Fullerene

The calculated characteristics of the hollow cluster proposed in [15] as a model of the fullerene C_{60} will now be presented. The charge of N_i ions in a hollow cluster is "smeared out" over a spherical shell of radius R. This distribution of ions generates the potential

$$U_i(\mathbf{r}) = \begin{cases} -\dfrac{N_e}{R}, & r \le R, \\ -\dfrac{N_e}{r}, & r > R, \end{cases}$$

which acts on mutually interacting valence electrons ($N_e = wN_i$, where $w = 1$ and $N_e = N_i = 60$). The energy of ions uniformly distributed over the sphere is $E_i = 0.5N_e^2/R$. According to [15], the Thomas–Fermi model of a hollow cluster does not admit any finite equilibrium radius R_0. The finite value $R_0 = 7.36$ close to $R_0^{(\exp)} = 6.75$ measured for the fullerene C_{60} was obtained in [15] by using the ion energy

$$E_i = 0.4311\frac{N_e^2}{R}, \qquad (12)$$

which corresponds to the real C_{60} molecule with ions located at the vertices of the truncated icosahedron of radius R.

The calculations performed for a hollow cluster in this study show that the equilibrium radius R_0 predicted

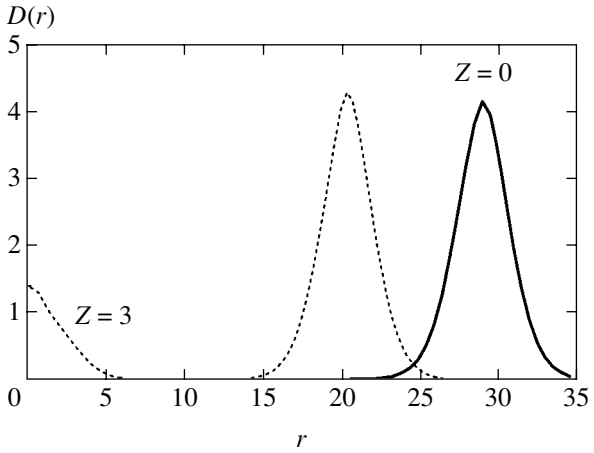

Fig. 5. Radial electron-density distributions in a hollow cluster ($Z = 0$, solid curve) and in a hollow cluster with a trivalent central atom ($Z = 3$, dotted curve) ($N_e - Z = 60$).

by the ITF model is finite. Its value was determined by minimizing the total energy of the cluster, $E = E_e + E_i$, as a function of R. As a result, a reasonable square-root dependence of the equilibrium radius on the number of particles was obtained: $R_0 = \alpha N_e^{1/2}$ with $\alpha = 3.743$. When this formula is applied to the fullerene, the result is too large ($R_0 = 29$), because the energy of ions is substantially overestimated by replacing their actual distribution with a spherical shell. The corresponding radial electron-density distribution is shown in Fig. 5. When the actual (not self-consistent) energy of ions given by (12) is used, the ITF model yields $R_0 = 5$ [16].

Analysis of spherical hollow clusters based on self-consistent modeling is of special interest because it can elucidate, in particular, the dependence of R_0 on the valence Z of the atom located at the center of the cluster. The real system corresponding to this model is a highly symmetric fullerene with a central atom, such as LaC_{60}. In this case, the potential generated by the ions and their energy are expressed as

$$U_i(\mathbf{r}) = -\frac{Z}{r} - \begin{cases} \dfrac{N_e - Z}{R}, & r \le R \\ \dfrac{N_e - Z}{r}, & r > R, \end{cases}$$

$$E_i = \frac{N_e^2 - Z^2}{2R},$$

where Z is the charge number of the central ion and the number N_e of electrons includes the Z valence electrons of the central atom. Calculations were performed for $N_e - Z = 60$ electrons in a spherical shell and $Z = 0$, 1, 2, and 3. The resulting monotonically decreasing function is accurately approximated by the expression

$$R_0(Z) = R_0(0)\exp(-0.04Z^2), \quad R_0(0) = 29.$$

Figure 5 shows the radial electron-density distribution $D(r) = 4\pi r^2 n_e(r)$ calculated for a hollow cluster ($Z = 0$) and an endohedrally doped hollow cluster with a trivalent central atom ($Z = 3$) having the corresponding equilibrium radii.

5. CONCLUSIONS

Based on an improved semiclassical model, a unified efficient algorithm is proposed for calculating the characteristics of a spherically symmetric many-body electron–ion system. The Euler–Lagrange equation for density is derived by minimizing the semiclassical energy functional including the exchange–correlation interaction and second-order gradient terms. This equation is applied to calculate averaged characteristics of various physical systems, such as atoms and solid or hollow clusters.

The proposed model can be used to calculate the total energy of electrons and their local characteristics (density and potential distributions). The radial Schrödinger equation with this potential is solved analytically in the semiclassical approximation. The resulting expression is used to obtain an oscillating correction to density in a finite system and calculate electron-density distributions for atoms and atomic clusters reflecting the discreteness of their electronic spectra. The contributions of the fourth-order corrections based on the averaged electron densities to the energies of these systems are evaluated. The use of these corrections in calculations is shown to be unjustified.

An analysis of the properties of a hollow cluster with ions uniformly "smeared" over a sphere is presented. In particular, the spatial density distribution is calculated for the valence electrons, and the total energy of a hollow cluster is determined as a function of radius for a given number of atoms in the cluster. It is shown that the resulting curve has a minimum at a finite (equilibrium) value of radius. A reasonable square-root dependence of the equilibrium radius on the number of particles in a cluster is obtained. The dependence of the equilibrium radius of an endohedrally doped hollow cluster on the valence of the atom located at its center is analyzed. The possibility of using hollow spherical clusters as models of the fullerenes C_{60} and LaC_{60} is also discussed. It is shown that an adequate quantitative description cannot be developed, because the energy of ions is highly overestimated when their actual distribution in a fullerene is approximated by a spherical shell.

ACKNOWLEDGMENTS

This work was supported in part by the Russian Foundation for Basic Research, project nos. 03-01-00438 and 02-01-00185, and by the Ministry of Industry and Science of the Russian Federation, grant NSh-2060.2003.2.

REFERENCES

1. *Theory of the Inhomogeneous Electron Gas*, Ed. by S. Lundqvist and N. H. March (Plenum, New York, 1983; Mir, Moscow, 1987).
2. D. A. Kirzhnits, Yu. E. Lozovik, and G. V. Shpatakovskaya, Usp. Fiz. Nauk **111**, 3 (1975) [Sov. Phys. Usp. **16**, 587 (1975)].
3. G. V. Shpatakovskaya, Teplofiz. Vys. Temp. **23**, 42 (1985).
4. E. A. Kuz'menkov and G. V. Shpatakovskaya, Int. J. Thermophys. **13**, 315 (1992).
5. D. A. Kirzhnits and G. V. Shpatakovskaya, Preprint No. 33, FI RAN (Physical Inst., Russian Academy of Sciences, Moscow, 1998).
6. G. V. Shpatakovskaya, Pis'ma Zh. Éksp. Teor. Fiz. **70**, 333 (1999) [JETP Lett. **70**, 334 (1999)]; cond-mat/0001116.
7. G. V. Shpatakovskaya, Zh. Éksp. Teor. Fiz. **118**, 87 (2000) [JETP **91**, 76 (2000)].
8. G. V. Shpatakovskaya, Pis'ma Zh. Éksp. Teor. Fiz. **73**, 306 (2001) [JETP Lett. **73**, 268 (2001)].
9. D. A. Kirzhnits, Zh. Éksp. Teor. Fiz. **32**, 115 (1957) [Sov. Phys. JETP **5**, 64 (1957)].
10. N. N. Kalitkin and L. V. Kuz'mina, Fiz. Tverd. Tela (Leningrad) **13**, 2314 (1971) [Sov. Phys. Solid State **13**, 1938 (1971)].
11. F. Perro, Phys. Rev. A **30**, 586 (1979).
12. M. Membrado, A. F. Pacheco, and J. Saudo, Phys. Rev. B **41**, 5643 (1990).
13. C. Yannouleas and U. Landman, Phys. Rev. B **48**, 8376 (1993).
14. M. Brack and P. Quentin, Nucl. Phys. A **361**, 35 (1981).
15. D. P. Clougherty and X. Zhu, Phys. Rev. A **56**, 632 (1997).
16. V. Ya. Karpov and G. V. Shpatakovskaya, Élektron. Zh. Issled. Rossii **191**, 2118 (2002), http://zhurnal.ape.relarn.ru/article/2002/191.pdf.
17. G. V. Shpatakovskaya, in *Physics of Extremal States*, Ed. by V. E. Fortov *et al.* (Inst. Probl. Khim. Fiz. Ross. Akad. Nauk, Chernogolovka, 2003), p. 119.
18. L. Hedin and B. I. Lundqvist, J. Phys. C **4**, 2064 (1971).
19. D. A. Kirzhnits and G. V. Shpatakovskaya, Zh. Éksp. Teor. Fiz. **62**, 2082 (1972) [Sov. Phys. JETP **35**, 1088 (1972)].
20. I. S. Gradshteĭn and I. M. Ryzhik, *Tables of Integrals, Sums, Series, and Products*, 4th ed. (Gostekhizdat, Moscow, 1962; Academic, New York, 1980).
21. D. R. Hartree and W. Hartree, Proc. R. Soc. London **149**, 210 (1935).
22. D. A. Kirzhnits, Zh. Éksp. Teor. Fiz. **34**, 1625 (1958) [Sov. Phys. JETP **7**, 1116 (1958)].
23. D. A. Kirzhnits and G. V. Shpatakovskaya, Zh. Éksp. Teor. Fiz. **108**, 1238 (1995) [JETP **81**, 679 (1995)].
24. M. Brack, Rev. Mod. Phys. **65**, 677 (1993).

Translated by A. Betev

Journal of Experimental and Theoretical Physics, Vol. 98, No. 3, 2004, pp. 463–477.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 125, No. 3, 2004, pp. 527–542.

ATOMS, SPECTRA, RADIATION

Optical Third-Harmonic Generation in One-Dimensional Photonic Crystals and Microcavities

M. G. Martemyanov*, T. V. Dolgova, and A. A. Fedyanin**
Moscow State University, Moscow, 119992 Russia
*e-mail: *misha@shg.ru; **URL:http://www.shg.ru*
Received August 25, 2003

Abstract—The formalism of nonlinear transfer matrices is used to develop a phenomenological model of a cubic nonlinear-optical response of one-dimensional photonic crystals and microcavities. It is shown that third-harmonic generation can be resonantly enhanced by frequency-angular tuning of the fundamental wave to the photonic band-gap edges and the microcavity mode. The positions and amplitudes of third-harmonic resonances at the edges of a photonic band gap strongly depend on the value and sign of the dispersion of refractive indexes of the layers that constitute the photonic crystal. Model calculations elucidate the role played by phase matching and spatial localization of the fundamental and third-harmonic fields inside a photonic crystal as the main mechanisms of enhancement of third-harmonic generation. The experimental spectrum of third-harmonic intensity of a porous silicon microcavity agrees with the calculated results. © *2004 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

Photonic crystals have been extensively studied in recent years because of their unique dispersion properties and the possibility of modulating the spectral density of optical field modes, which manifests itself by the formation of photonic band gaps [1]. Fundamental interest in photonic crystals, in particular, stems from peculiar nonlinear optical effects, such as bistability [2] and optical switching [3] due to modulation of the refractive index of one-dimensional photonic crystal layers in a high-intensity field. This modulation causes a dynamic or quasi-stationary shift of the photonic band gap in a photonic crystal with a cubic nonlinearity. In such crystals, one can observe four-wave mixing and excitation of the waveguide mode at the anti-Stokes frequency which propagates along interfaces [4]. In media with modulation only of nonlinear susceptibility with a period of the order of the wavelength, nonlinear diffraction effects are observed [5].

The use of photonic crystals for effectively generating radiation at the second harmonic frequency was suggested in [6] and experimentally implemented for the first time in [7]. Phase mismatch between the fundamental and second harmonic waves is minimized by adding the reciprocal lattice vector of the periodic medium to the wave vectors of the interacting waves. When one of the frequencies is tuned to the edge of a photonic band gap, the phase matching condition for pumping and second harmonic waves is satisfied, which results in the enhancement of second harmonic generation in photonic crystals [8–11].

Third-harmonic generation (THG) in a photonic crystal can occur either directly due to cubic susceptibility or in a cascade manner as a result of quadratic susceptibility. The first process was considered for an infinite photonic crystal in [12], where it was shown that there were structure parameters at which phase matching conditions were simultaneously satisfied for the fundamental and second-harmonic waves and the fundamental and third-harmonic waves. With these parameters, the time evolution of second- and third-harmonic intensities was studied, and it was shown that the pump energy could not be completely transfered to the second or third harmonic. In cascade THG by a one-dimensional photonic crystal with quadratic susceptibility [13, 14], simultaneous phase matching of the pump with the second and third harmonics can also be achieved by adjusting the optical thicknesses of photonic crystal layers. The calculations reported in [13] were performed for a photonic crystal with an infinite number of layers, and only phase matching effects were therefore studied. In [14], a photonic crystal with a finite number of layers was considered, and effects of the spatial localization of fields related to its finite dimensions were taken into account. Pumping field localization effects can be enhanced by the introduction of a defect into a photonic crystal. In such a microcavity with distributed mirrors, the electromagnetic field resonant to the microcavity mode is effectively localized, which enhances second [15, 16] or third [17] harmonic generation. The enhancement of harmonic wave generation at the photonic band gap edge and in the microcavity mode is a result of the combined effects of phase matching due to the periodicity of Bragg reflectors and field localization caused by the presence of a microcavity spacer and finite photonic-crystal dimensions [11]. A key parameter that determines the enhancement of third harmonic generation is the dispersion of refractive indexes of layers in a photonic crystal or microcavity.

Its compensation is the essence of the phase matching at the photonic band gap edges in a photonic crystal and in the microcavity mode. However, the dependence of the magnitude and spectral position of third harmonic resonances on the magnitude and sign of the dispersion of layers constituting a photonic crystal or microcavity has not been studied yet.

In this work, we study THG in one-dimensional finite photonic crystals and microcavities characterized by a cubic nonlinearity $\hat{\chi}^{(3)}$. The formalism of nonlinear transfer matrices is used to elucidate the role played by each of the mechanisms of the enhancement of third harmonic generation, that is, phase matching and pumping and third harmonic field localization when the pumping wave is in resonance with the microcavity mode or photonic band gap edge. The spectra of third harmonic intensity in the spectral range containing a photonic band gap with a microcavity mode and a passband region are calculated in the approximation of a given pumping field. The dependence of third harmonic generation enhancement on the dispersion of the refractive indexes of the layers that constitute a photonic crystal is studied. This dispersion determines the mutual arrangement of photonic band gaps and microcavity modes at the pump and third-harmonic wavelengths. The calculation results are compared with the experimental third harmonic spectrum generated in a microcavity made from mesoporous silicon.

2. NONLINEAR TRANSFER-MATRIX METHODS FOR CALCULATING THIRD-HARMONIC GENERATION IN PHOTONIC CRYSTALS

2.1. Nonlinear Transfer-Matrix Method

There are several approaches to calculating optical harmonic generation in one-dimensional photonic crystals. One of these is via solving a system of reduced equations obtained in the method of slowly varying amplitudes [12–14, 18]. This approach can be used to take into account energy transfer from the pump to the generated harmonic; analyze simultaneous generation of the second and third harmonics; and examine the time evolution of the fundamental and the second- and third-harmonic waves, which is important for studying harmonic generation by femtosecond pulses. In the approximation of constant fundamental field, THG is described by a single inhomogeneous equation, which can be directly solved using the Green function of a multilayer structure [4, 19]; the solution is constructed based on linearly independent solutions to the homogeneous wave equation, which can be found using the formalism of transfer matrices [20]. Lastly, a convenient method is the extension of the formalism of transfer matrices to harmonic generation. This method is applicable in the approximation of constant fundamental wave and when the fields are stationary, that is, under the condition that the pumping pulse width is much greater than the time of fundamental-wave propagation across the photonic crystal. There are two equivalent approaches. The first one, described in [21], uses the formalism of Green functions suggested by Sipe [22]. The second approach given in [23] is based on the formalism of coupled and free harmonic waves introduced by Bloembergen and Pershan [24]. Both rely on direct solution of an inhomogeneous wave equation with the use of the Green functions of a photonic crystal and provide additional information about the nonlinear optical processes under consideration, such as contributions of separate layers to the resulting third-harmonic wave.

The method of nonlinear transfer matrices suggested in [23] can conveniently be used to calculate nonlinear-optical effects in one-dimensional photonic crystals because of its simplicity and form optimal for numerical calculations. The problem of THG in photonic crystals can be decomposed into three sequential stages. First, the fundamental wave propagation in a multilayer linear structure is described taking into account multiple-reflection interference, and the spatial pumping field distribution within the photonic crystal is calculated. At the second stage, cubic polarization induced in a medium with nonzero $\hat{\chi}^{(3)}$ is determined. Lastly, the linear problem of third harmonic wave propagation in a layered structure is solved taking into account coupled and homogeneous waves, and the intensity of the third harmonic wave that emerges from the photonic crystal is found.

Let the z axis be perpendicular to the surface of the photonic crystal and xz be the plane of pumping wave incidence (Fig. 1a). A monochromatic linearly polarized fundamental wave with frequency ω, wave vector $\mathbf{k}_1^+$, and amplitude $\mathbf{E}_1^+$ propagates in half-space 1 at angle θ_1 to the normal to the surface of the photonic crystal. We assume that the photonic crystal is optically inactive and nonmagnetic; *s*- and *p*-polarized waves will therefore be considered separately. The electromagnetic field in the jth layer is a superposition of two plane forward waves (propagating in the positive direction along the z axis) and backward,

$$\begin{aligned}\mathbf{E}_j^{\omega}(z,t) &= \mathbf{E}_j^{+}\exp[ik_{z,j}^{\omega}(z-d_{ij}) + (ik_{x,j}^{\omega}x - i\omega t)] \\ &+ \mathbf{E}_j^{-}\exp[-ik_{z,j}^{\omega}(z-d_{ij}) + (ik_{x,j}^{\omega}x - i\omega t)],\end{aligned} \quad (1)$$

where

$$k_{z,j}^{\omega} = |\mathbf{k}_j|\cos\theta_j, \quad k_{x,j}^{\omega} = |\mathbf{k}_j|\sin\theta_j,$$

d_{ij} is the z coordinate of the boundary between the ith and jth layers, $\mathbf{k}_j$ is the wave vector, and θ_j is the angle of refraction of the pumping wave in the jth layer. The $\exp(ik_{x,j}^{\omega}x - i\omega t)$ term will further be omitted because of the translational invariance of the problem in the xy

plane and its stationary character. The field at the left boundary of the jth layer is represented by the two-component vector $\mathbf{E}_j \equiv (E_j^+ + E_j^-)$; the first component of this vector is the amplitude of the forward wave and the second one, that of the backward wave. The relation between $\mathbf{E}_j$ and $\mathbf{E}_k$ in the kth layer at its left boundary is given by the two 2×2 matrices

$$\begin{pmatrix} E_k^+ \\ E_k^- \end{pmatrix} = \begin{pmatrix} 1/t_{kj} & r_{kj}/t_{kj} \\ r_{kj}/t_{kj} & 1/t_{kj} \end{pmatrix} \times \begin{pmatrix} \exp(ik_{z,j}^{\omega} d_j) & 0 \\ 0 & \exp(-ik_{z,j}^{\omega} d_j) \end{pmatrix} \begin{pmatrix} E_j^+ \\ E_j^- \end{pmatrix}. \quad (2)$$

Here, the first matrix contains the Fresnel reflectivity r_{kj} and transmissivity t_{kj} for the wave incident from the kth to the jth layer and relates fields to the left and right of the kjth interface. In what follows, this matrix is denoted by $\mathbf{M}_{kj}$. The second $\Phi_j(d_j)$ matrix describes field propagation in the jth layer of thickness d_j from the left to right boundary. The fields in half-spaces 1 and l can therefore be related. Under the assumption that the backward wave is absent in half-space l and that a wave with unit amplitude is incident on the 1–2 boundary, we obtain

$$\begin{pmatrix} t \\ 0 \end{pmatrix} = \begin{pmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{pmatrix} \begin{pmatrix} 1 \\ r \end{pmatrix}. \quad (3)$$

Here, r and t are the reflectivity and transmissivity of the photonic crystal as a whole. They are determined from (3), which gives

$$r = -\frac{T_{21}}{T_{22}}, \quad t = \frac{T_{11}T_{22} - T_{12}T_{21}}{T_{22}},$$

where $T_{\alpha\beta}$ are the elements of the transfer matrix for the photonic crystal as a whole, $\mathbf{T} \equiv \mathbf{M}_{l(l-1)}\mathbf{\Phi}_{(l-1)} \ldots \mathbf{M}_{21}$. The fundamental field distribution within the photonic crystal is given by

$$\mathbf{E}_j^{\omega}(z) = \mathbf{\Phi}_j(z - d_{j(j-1)})\mathbf{M}_{j(j-1)}\mathbf{\Phi}_{(j-1)} \ldots \mathbf{M}_{21} \begin{pmatrix} 1 \\ r \end{pmatrix}. \quad (4)$$

Spatial distribution (4) can be used to calculate the spatial distribution of the cubic polarization wave and inhomogeneous third harmonic can be calculated; this is done in the next Section.

The third stage of solving the problem can be considerably simplified. In the approximation of a given pumping field, cubic polarization in each layer is a third harmonic source independent of the other layers. Thus, we can solve the problem of third-harmonic generation and propagation for a single photonic crystal layer with a cubic susceptibility. The third-harmonic field generated by the photonic crystal can then be obtained by summing such partial third-harmonic outputs of all layers and taking into account their phases.

Let the jth layer be nonlinear. The interference of the coupled $\mathbf{E}^{(s)}$ and free $\mathbf{E}_j$ third harmonic waves is included in the boundary conditions, which, for the ijth and jkth boundaries, have the form

$$\mathbf{\Phi}_i\mathbf{E}_i = \mathbf{M}_{ij}\mathbf{E}_j + \mathbf{M}_i^{(s)}\mathbf{E}^{(s)}, \quad \mathbf{M}_{kj}\mathbf{\Phi}_j\mathbf{E}_j + \mathbf{M}_k^{(s)}\mathbf{\Phi}^{(s)}\mathbf{E}^{(s)} = \mathbf{E}_k. \quad (5)$$

The amplitude of the inhomogeneous third-harmonic wave in (5) is calculated in the jth nonlinear layer at its left boundary, and all homogeneous waves are third-harmonic waves. Matrices $\mathbf{M}$ with index (s) are constructed similarly to the usual transfer matrices, but the Fresnel coefficients in their elements contain refractive indexes for the inhomogeneous third-harmonic wave in the nonlinear layer and the free third harmonic in the layer whose number equals the lower index of the matrix. The $\mathbf{\Phi}^{(s)}$ matrix is similar to Φ_j and is obtained from the latter by replacing the wave vectors of the free third harmonic with the wave vectors of the inhomogeneous wave. System (5) yields

$$\mathbf{E}_k = \mathbf{M}_{kj}\mathbf{\Phi}_j(\mathbf{M}_{ji}\mathbf{\Phi}_i\mathbf{E}_i + \mathbf{S}_j), \quad (6)$$

where the vector

$$\mathbf{S}_j = (\overline{\mathbf{\Phi}}_j\mathbf{M}_j^{(s)}\mathbf{\Phi}^{(s)} - \mathbf{M}_j^{(s)})\mathbf{E}^{(s)} \quad (7)$$

is singled out for convenience. The $\overline{\mathbf{\Phi}}_j$ matrix is inverse to $\mathbf{\Phi}_j$. Equation (6) determines the third-harmonic field in the kth layer as a superposition of the waves transmitted from the ith layer and generated in the jth layer by nonlinear sources. Equation (7) contains the contribution of the nonlinear jth layer to the third-harmonic wave; the term in parentheses takes into account the interference of the homogeneous and inhomogeneous waves. Under the assumption that no external field with the third-harmonic wavelength is incident on the photonic crystal, (6) can be rewritten as

$$\mathbf{R}_{jl}\begin{bmatrix} E_l^+(j) \\ 0 \end{bmatrix} - \mathbf{L}_{j1}\begin{bmatrix} 0 \\ E_1^-(j) \end{bmatrix} = \mathbf{S}_j, \quad (8)$$

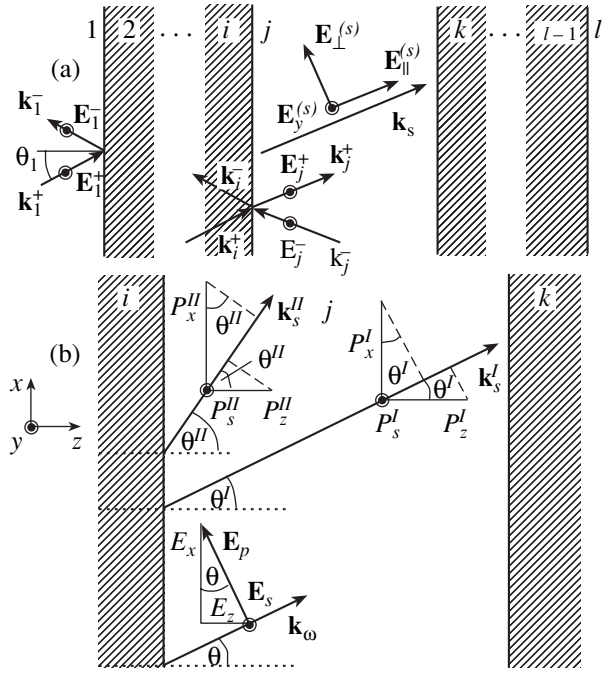

Fig. 1. (a) Scheme of a one-dimensional photonic crystal (layers 2 … $l-1$) (half-spaces 1 and l denote the vacuum and substrate, *s*-polarized fundamental wave is shown) and (b) THG scheme in the *j*th nonlinear layer.

where the matrices

$$\mathbf{R}_{jl} \equiv \overline{\mathbf{\Phi}}_j \mathbf{M}_{jk} \overline{\mathbf{\Phi}}_k \mathbf{M}_{k(k+1)} \ldots \overline{\mathbf{\Phi}}_{(l-1)} \mathbf{M}_{(l-1)l},$$

$$\mathbf{L}_{j1} \equiv \mathbf{M}_{ji} \mathbf{\Phi}_i \mathbf{M}_{i(i-1)} \ldots \mathbf{\Phi}_2 \mathbf{M}_{21}$$

characterize the propagation of homogeneous third-harmonic waves from half-spaces 1 and l to the *j*th layer. It follows that, given $\mathbf{S}_j$, we can find the amplitude and phase of the third-harmonic fields $E_1^-(j)$ and $E_l^+(j)$ generated in the *j*th photonic crystal layer and emerging from the photonic crystal into the vacuum (half-space 1) and substrate (half-space l). The total third-harmonic field in the substrate or vacuum is the sum of $E_l^+(j)$ or $E_1^-(j)$ taken over all layers.

2.2. Inhomogeneous Third-Harmonic Waves in a Photonic Crystal

Let us obtain equations describing inhomogeneous third-harmonic waves for direct THG by means of a cubic susceptibility $\hat{\chi}^{(3)}$. Equations for cascade THG caused by quadratic susceptibility $\hat{\chi}^{(2)}$ can be obtained similarly. The cascade process can be ignored when the nonlinear medium has an inversion center. The generation of the bulk dipole second harmonic is then forbidden, and the second harmonic is only generated by dipole surface and bulk quadrupole sources; conversely, the generation of the bulk dipole third harmonic is allowed. As a result, the cascade process becomes less effective.

A nonlinear photonic crystal layer can be treated as a medium rotationally isotropic in the plane of the layer. Such a medium is characterized by the symmetry groups ∞/mm and $\infty 2$ (cylinders) and ∞m (a cone with a symmetry axis of an infinite order perpendicular to the surface of the layer). Equations for inhomogeneous third-harmonic waves for other symmetry groups can be obtained similarly. The tensor of dipole cubic susceptibility $\hat{\chi}^{(3)}$ invariant with respect to the ∞m, $\infty 2$, and ∞/mm groups and symmetrical with respect to permutations of the last three indices has four independent nonzero components [25],

$$\begin{aligned} \chi_1 \equiv \chi_{xxyy} &= \chi_{yyxx} = \frac{1}{3}\chi_{xxxx} = \frac{1}{3}\chi_{yyyy}, \\ \chi_2 &\equiv \chi_{yyzz} = \chi_{xxzz}, \\ \chi_3 &\equiv \chi_{zzzz}, \\ \chi_4 &\equiv \chi_{zzxx} = \chi_{zzyy}. \end{aligned} \tag{9}$$

Let the *j*th layer with cubic susceptibility be situated between two linear layers i and k (Fig. 1). Let us determine the amplitude of the inhomogeneous third harmonic $\mathbf{E}^{(s)}$ on the *ji*th interface from the fundamental field amplitude $\mathbf{E}_j$ on the same boundary. The angle between the fundamental wave vector

$$k_\omega = n_\omega \frac{\omega}{c},$$

where n_ω is the refractive index at the fundamental frequency, and the z axis is θ. Here and throughout, index j numbering layers is omitted. The dipole cubic polarization is given by the convolution

$$\begin{aligned} \mathbf{P}^{3\omega} &= \hat{\chi}^{(3)} \vdots (\mathbf{E}^+ \exp(ik_z^\omega z) + \mathbf{E}^- \exp(-ik_z^\omega z))^3 \\ &\equiv \mathbf{P}^{I+} \exp(ik_z^{s,I} z) + \mathbf{P}^{I-} \exp(-ik_z^{s,I} z) \\ &+ \mathbf{P}^{II+} \exp(ik_z^{s,II} z) + \mathbf{P}^{II-} \exp(-ik_z^{s,II} z). \end{aligned} \tag{10}$$

Here, $k_z^\omega = k_\omega \cos\theta$ is the projection of the fundamental wave vector onto the normal to the interface. In Eq. (10), all cubic polarization terms are divided into two types. The terms of the first type have the z wave vector component $k_z^{s,I} = 3k_z^\omega$ and are obtained by the convolution of three fundamental waves propagating in the same direction. The other terms have the normal projection of the wave vector $k_z^{s,II} = k_z^\omega$ and are obtained by the convolution of three fundamental

waves one of which has the z component of the wave vector opposite to the projections of the other two waves. The propagation of cubic polarization and inhomogeneous third-harmonic waves is determined by the effective permittivity calculated as

$$k^{s,(I,II)} = \frac{3\omega}{c}\sqrt{\epsilon^{I,II}}.$$

For the waves of type *I*,

$$k^{s,I} = 3k_\omega = \frac{3\omega}{c}\sqrt{\epsilon(\omega)};$$

that is,

$$\epsilon^I = \epsilon(\omega)$$

and the polarization wave propagates in the medium collinearly to the fundamental wave, $\theta^I = \theta$. Similar calculations for the polarization waves of type *II* give

$$\epsilon^{II} = \epsilon(\omega)(1 + 8\sin^2\theta)/9.$$

The angle between axis z and the propagation direction of the polarization waves of type *II* is different from θ; it is given by

$$\theta^{II} = \arctan(3k_x^\omega/k_z^\omega).$$

Taking into account nonzero components (9) of the tensor $\hat{\chi}^{(3)}$, the projections of the two-component vectors $\mathbf{P}^I = (\mathbf{P}^{I+}, \mathbf{P}^{I-})$ and $\mathbf{P}^{II} = (\mathbf{P}^{II+}, \mathbf{P}^{II-})$ onto the coordinate axes expressed in terms of the *s*- and *p*-polarized fundamental wave components can be written as

$$\begin{aligned}\mathbf{P}_x^I &= \chi_1(\mathbf{E}_p^3\cos^3\theta + \mathbf{E}_s^2\mathbf{E}_p\cos\theta)\\ &\quad + \chi_2\mathbf{E}_p^3\cos\theta\sin^2\theta,\\ \mathbf{P}_y^I &= \chi_1(\mathbf{E}_s^3 + \mathbf{E}_s\mathbf{E}_p^2\cos^2\theta) + \chi_2\mathbf{E}_s\mathbf{E}_p^2\sin^2\theta,\\ \mathbf{P}_z^I &= -\frac{1}{3}\chi_3\mathbf{E}_p^3\sin^3\theta\\ &\quad - \chi_4(\mathbf{E}_p\mathbf{E}_s^2\sin\theta + \mathbf{E}_p^3\sin\theta\cos^2\theta)\end{aligned} \tag{11}$$

for $\mathbf{P}^I$ and

$$\begin{aligned}\mathbf{P}_x^{II} &= \chi_1(3\mathbf{E}_p^3\cos^3\theta + 2\mathbf{E}_p\mathbf{E}_s\mathbf{E}_s\cos\theta + \mathbf{E}_s\mathbf{E}_s\mathbf{E}_p\cos\theta)\\ &\quad + 3\chi_2\mathbf{E}_p^3\sin^2\theta\cos\theta,\\ \mathbf{P}_y^{II} &= \chi_1(3\mathbf{E}_s^2 + 2\mathbf{E}_s\mathbf{E}_p\mathbf{E}_p\cos^2\theta\\ &\quad + \mathbf{E}_p\mathbf{E}_p\mathbf{E}_s\cos^2\theta)\\ &\quad + \chi_2(\mathbf{E}_p\mathbf{E}_p\mathbf{E}_s\sin^2\theta + 2\mathbf{E}_s\mathbf{E}_p\mathbf{E}_p\sin^2\theta),\\ \mathbf{P}_z^{II} &= -\chi_3\mathbf{E}_p^3\sin^3\theta\\ &\quad - \chi_4\sin\theta(3\mathbf{E}_p^3\cos^2\theta + \mathbf{E}_s\mathbf{E}_s\mathbf{E}_p + 2\mathbf{E}_p\mathbf{E}_s\mathbf{E}_s)\end{aligned} \tag{12}$$

for $\mathbf{P}^{II}$. Here, we use the equation

$$E_y = E_s, \quad E_x = E_p\cos\theta, \quad E_z = -E_p\sin\theta.$$

The product of three two-component vectors is given by the rule

$$\mathbf{E}_a\mathbf{E}_b\mathbf{E}_c = \begin{pmatrix}\mathbf{E}_a^+\cdot\mathbf{E}_b^+\cdot\mathbf{E}_c^+\\ \mathbf{E}_a^-\cdot\mathbf{E}_b^-\cdot\mathbf{E}_c^-\end{pmatrix} \tag{13}$$

for the waves of type *I* and by the rule

$$\mathbf{E}_a\mathbf{E}_b\mathbf{E}_c = \begin{pmatrix}\mathbf{E}_a^+\cdot\mathbf{E}_b^+\cdot\mathbf{E}_c^-\\ \mathbf{E}_a^-\cdot\mathbf{E}_b^-\cdot\mathbf{E}_c^+\end{pmatrix} \tag{14}$$

for the waves of type *II*. Cubic polarization can conveniently be decomposed into components with polarization directions normal and parallel to the wave vector of the inhomogeneous wave and the *s*-polarized component,

$$\begin{aligned}\mathbf{P}_\parallel^{I,II} &= \mathbf{P}_x^{I,II}\sin(\theta^{I,II}) + \mathbf{P}_z^{I,II}\cos(\theta^{I,II}),\\ \mathbf{P}_\perp^{I,II} &= \mathbf{P}_x^{I,II}\cos(\theta^{I,II}) - \mathbf{P}_z^{I,II}\sin(\theta^{I,II}),\\ &\quad \mathbf{P}_y^{I,II}.\end{aligned} \tag{15}$$

For such components, the transition to third-harmonic inhomogeneous waves is simple [24]:

$$\begin{aligned}\mathbf{E}^{(s)I,II} &= \frac{4\pi}{\epsilon^{I,II} - \epsilon(3\omega)}(\mathbf{P}_y^{I,II} + \mathbf{P}_\perp^{I,II})\\ &\quad - \frac{4\pi}{\epsilon(3\omega)}\mathbf{P}_\parallel^{I,II}.\end{aligned} \tag{16}$$

Equations (16) for third-harmonic inhomogeneous waves are substituted into (7). The reduction of the general problem of THG in a photonic crystal to THG in a photonic crystal containing a single layer with cubic susceptibility allows partial contributions of each photonic crystal layer to the total third-harmonic wave to be calculated. In combination with the feasibility of calculating the fundamental field at each point in the photonic crystal, this is a convenient tool for analyzing the mechanisms of nonlinear-optical phenomena in one-dimensional photonic crystals. Variation of the ini-

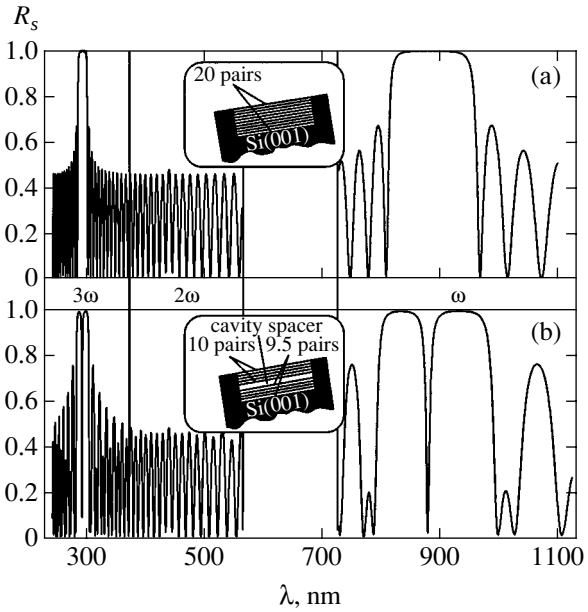

Fig. 2. Dependence of the linear reflection coefficient of the *s*-polarized fundamental wave on its wavelength calculated for (a) a photonic crystal and (b) a microcavity. Photonic crystal and microcavity schemes are shown in the insets.

tial calculation parameters, such as the fundamental-radiation wavelength or the corresponding incidence angle, allows us to obtain the frequency and angular spectra of the linear reflection coefficient and the intensities and phases of third harmonic waves.

Note that (11), (12), and (15) specify possible polarizations of third-harmonic waves. If the fundamental wave is *s*-polarized, only an *s*-polarized third harmonic is generated (the first term of the P_y component), and if the fundamental wave is *p*-polarized, the third harmonic is also *p*-polarized (the first terms of the P_x and P_z components).

3. THIRD-HARMONIC GENERATION IN PHOTONIC CRYSTALS

3.1. The Absence of Dispersion

Calculations of the enhancement of third-harmonic generation at the edge of the photonic band gap are performed for a photonic crystal. For a microcavity, we study THG enhancement effects that appear when the fundamental radiation is tuned across the region in the frequency-angle space that corresponds to the microcavity mode. The influence of the microcavity layer is weak at the edge of the photonic band gap, and third-harmonic enhancement is similar to that characteristic of photonic crystals. The model photonic crystal consists of 20 pairs of alternating layers with refractive indexes $n_\omega^H = 1.93$ and $n_\omega^L = 1.61$ and optical thicknesses $\lambda_0/4$, where $\lambda_0 = 960$ nm (Fig. 2). The refractive indexes selected are close to the real refractive indexes of porous silicon layers (porous silicon was used to fabricate the microcavity studied experimentally in this work). The microcavity is obtained by doubling the thickness of the central twenty-first layer. It is assumed that the photonic crystal and microcavity are placed between two half-spaces, air and a linear medium with a refractive index equal to the refractive index of silicon. THG from the silicon substrate is ignored. Fundamental radiation comes from air. The mechanisms of THG enhancement are studied in the frequency domain, that is, under fundamental wavelength variations, which allows both photonic band gap edges to be observed for the same sample. This cannot be done by means of angular third-harmonic spectroscopy with a fixed fundamental wavelength and normal component of the fundamental wave vector tuned by changing the angle of incidence. All frequency spectra are calculated at the incidence angle $\theta_1 = 45°$. The fundamental wavelength λ_ω is varied from 730 to 1100 nm. The refractive indexes of photonic crystal layers are assumed to be constant in the spectral ranges of the fundamental and third harmonic wavelengths.

The spectra of the linear reflection coefficient of the photonic crystal and microcavity are shown in Fig. 2. The same figure contains the spectra of the linear reflection coefficient in the wavelength ranges of the second (360–560 nm, this spectrum is further denoted by $R_{2\omega}$) and third (240–370 nm, spectrum $R_{3\omega}$) harmonics. The reflection coefficient R_ω is close to unity from 820 to 940 nm, which is a manifestation of the photonic band gap. The $R_{3\omega}$ spectrum also contains a wavelength region with the reflection coefficient close to unity (photonic band gap at third-harmonic wavelengths), whereas there is no photonic band gap in the $R_{2\omega}$ spectrum, because the phase difference between the waves reflected from layers with optical thicknesses $\lambda/4$ or $3\lambda/4$ is πn, where n is an even integer, and the waves are added in phase. If the optical thickness of layers is $\lambda/2$, the phase difference between the waves reflected from such layers is πm, where m is an odd integer, and the waves interfere destructively. The presence of a resonant layer manifests itself in the spectrum of the linear reflection coefficient by a dip within the photonic band gap corresponding to the microcavity mode. The microcavity mode is present in both R_ω and $R_{3\omega}$ spectra.

The spatial distributions of the fundamental field amplitude shown in Fig. 3 were calculated at the wavelengths corresponding to the minima in the spectrum of the linear reflection coefficient that are closest to photonic band gap (Figs. 3b, 3c), in the region within the photonic band gap of the photonic crystal (Fig. 3d), and in the microcavity mode (Fig. 3a). The fundamental field amplitude is $E_\omega = |E_j^+ + E_j^-|$, where E_j^+ and E_j^- are the complex components of the $\mathbf{E}_j^\omega$ two-component pumping field vector given by (4). When λ_ω is tuned across the photonic band gap, the fundamental field

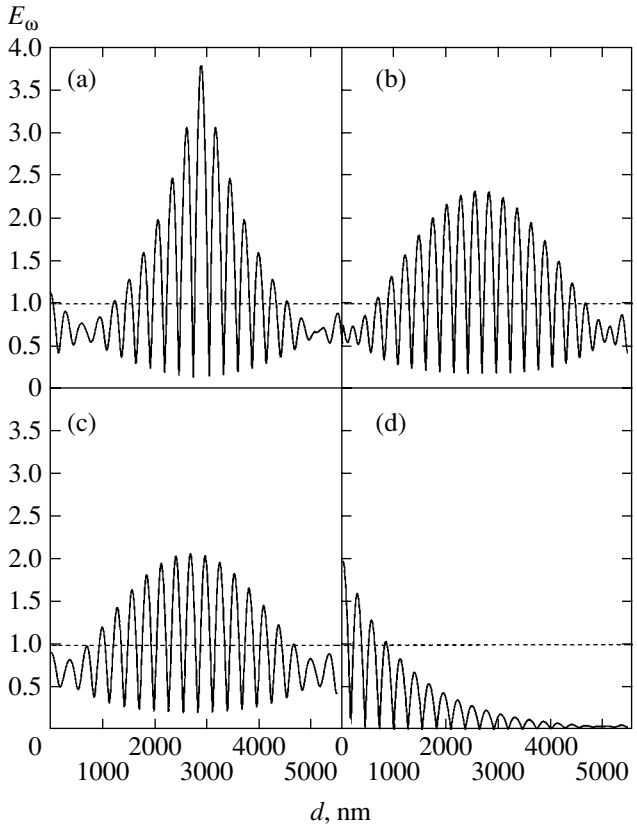

Fig. 3. Fundamental field amplitude distributions inside a photonic crystal and microcavity: (a) for the microcavity mode, fundamental wavelength is $\lambda_\omega = 877$ nm; (b) at the short-wavelength edge of the photonic crystal band gap, $\lambda_\omega = 806$ nm; (c) at the long-wavelength edge of the photonic band gap, $\lambda_\omega = 966$ nm; and (d) within the photonic band gap, $\lambda_\omega = 870$ nm. The dashed line is the incident wave amplitude equal to one.

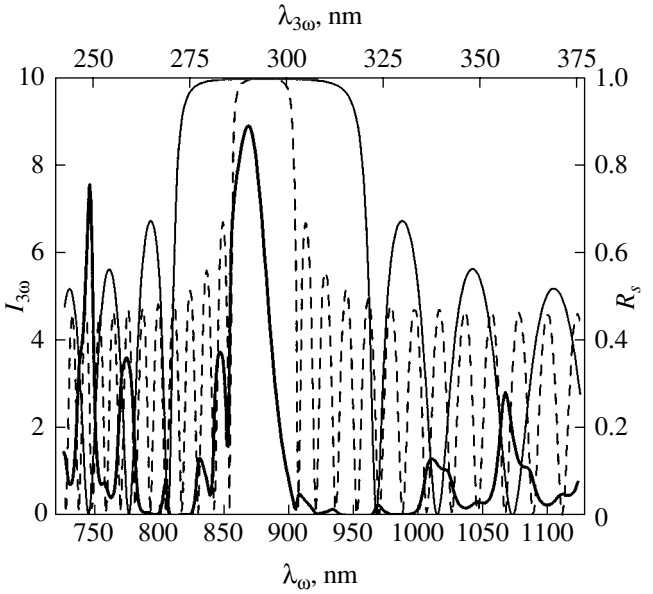

Fig. 4. The third harmonic intensity spectrum of a photonic crystal calculated for *s*-polarized fundamental and third harmonic waves (*SS* geometry) (the thick solid line); the reflection coefficient spectra of an *s*-polarized wave calculated in the regions of fundamental (solid thin line) and third-harmonic (dashed line) wavelengths. The third-harmonic intensity $I_{3\omega}$ is given in arbitrary units.

exponentially decays as the depth of penetration into the photonic crystal increases (Fig. 3d). The fundamental wave which is in resonance with the microcavity mode is strongly localized inside the resonant layer. At the chosen microcavity parameters, E_ω increases approximately fourfold. At the short-wavelength and long-wavelength edges of the photonic band gap, the fundamental field is localized less strongly and is enhanced 2.1–2.3 times. This effect is caused by the finite photonic crystal length; with a larger number of layers, the fundamental field is distributed more evenly. The amplitude of the pumping field resonant to the microcavity mode reaches a maximum in the microcavity layer and sharply decreases in several neighboring layers, whereas in a photonic crystal, the fundamental field tuned to photonic band gap edges is more evenly distributed over the photonic crystal. This means that for effective THG in a photonic crystal, phase matching of the homogeneous third harmonic waves generated by various layers and having amplitudes of the same order is important.

The third-harmonic intensity spectrum $I_{3\omega}$ generated in a photonic crystal in the absence of refractive index dispersion is shown in Fig. 4. Multiple peaks located both within the photonic band gap and near every R_ω spectrum minimum to the left and right outside the photonic band gap are observed in the third-harmonic spectrum. Third-harmonic enhancement is less pronounced at the edge of the photonic band gap, but the amplitude of third-harmonic peaks increases in the next linear reflection coefficient minima.

When the fundamental wavelength is tuned across the photonic band gap, the amplitude of the fundamental field decreases exponentially as the depth of penetration into the photonic crystal increases and the source of third-harmonic generation is several photonic crystal layers near surface. Upon tuning λ_ω to the minimum of the reflection coefficient, the fundamental field effectively penetrates deep into the photonic crystal and all its layers become sources of third-harmonic waves. At the same time, the amplitudes of the peaks of third-harmonic intensity are commensurate no matter whether λ_ω is tuned across the photonic band gap or the minima of the reflection coefficient spectrum. This is evidence of dephasing of third-harmonic waves coming from different photonic crystal layers; these waves destructively interfere with each other. It can be expected that the inclusion of dispersion into calculations (when the refractive indexes of photonic crystal layers have different values at the fundamental and third-harmonic wavelengths) will change the phases

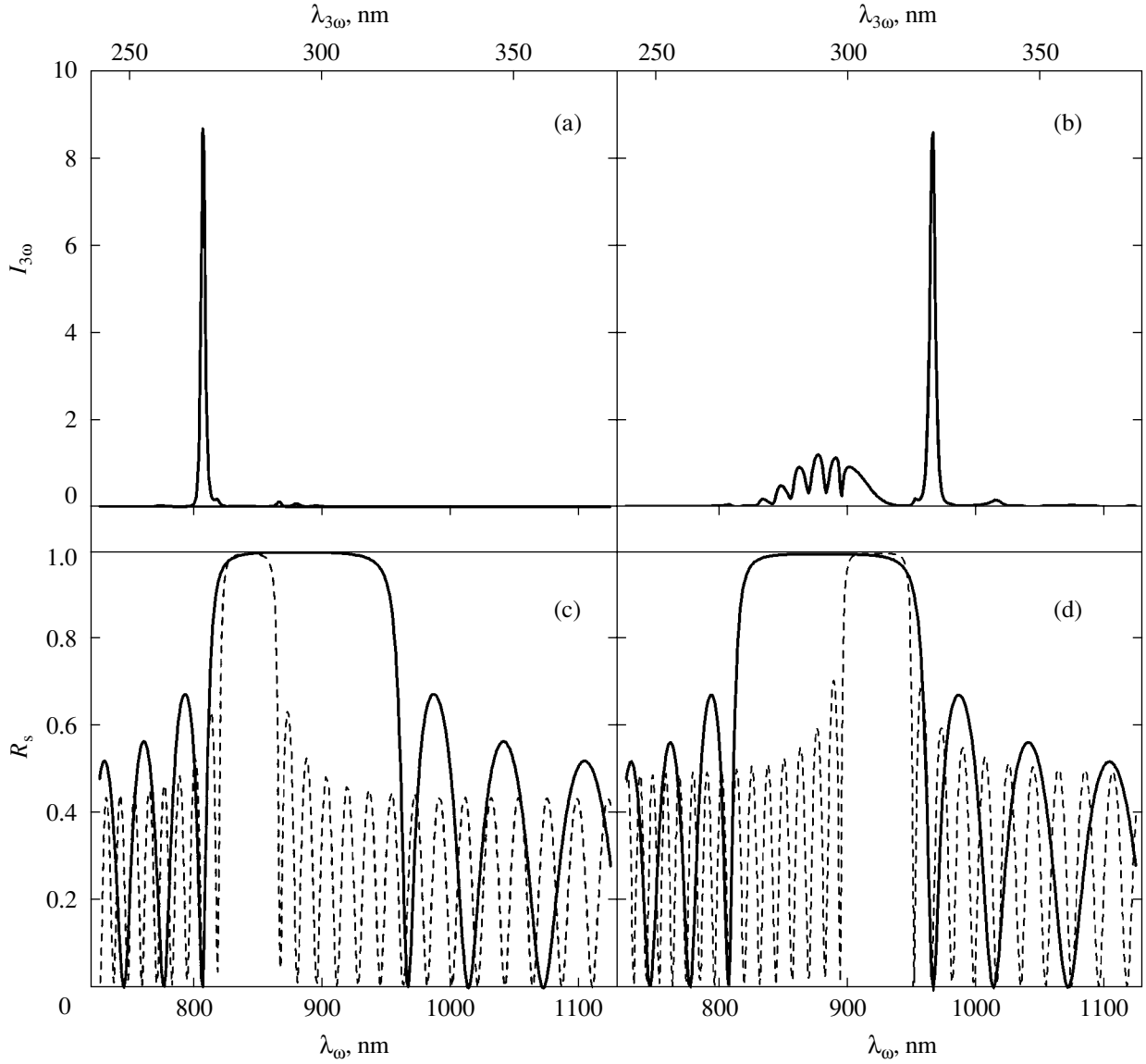

Fig. 5. (a, b) Third-harmonic intensity spectra of a photonic crystal calculated for the *SS* geometry in the presence of dispersion of photonic crystal layers and (c, d) *s*-polarized wave reflection coefficient spectra calculated in the wavelength ranges of fundamental waves (solid line) and the third harmonic (dashed line). Third-harmonic intensity units in Figs. 5a, 5b and 4 are incommensurate.

of partial third-harmonic waves coming from different layers and the peaks corresponding to the minimum of R_ω will gain in third-harmonic intensity.

3.2. Enhancement of Third-Harmonic Generation in the Presence of Dispersion

Characteristic third-harmonic intensity spectra calculated in the presence of dispersion of the refractive indexes of photonic crystal layers are shown in Figs. 5a and 5b. The bottom panels show the linear reflection coefficient spectra R_ω and $R_{3\omega}$. Figures 5a and 5c correspond to the dispersion of the refractive indexes of optically less dense photonic crystal layers $\Delta n_{3\omega} = n_{3\omega}^L - n_\omega^L = -0.045$, and Figs 5b and 5d, to $\Delta n_{3\omega} = n_{3\omega}^L - n_\omega^L = 0.051$. The refractive index at the fundamental wavelength is taken to be $n_\omega^L = 1.610 + i \times 0.00003$; that is, absorption is virtually absent. The dispersion of the refractive indexes of optically denser photonic crystal layers is set equal to $\Delta n_{3\omega} n_\omega^H / n_\omega^L$. The $I_{3\omega}(\lambda_\omega)$ spectra show that the third-harmonic intensity resonantly increases in the spectral region of photonic band gap edges. At $\Delta n_{3\omega} < 0$, the intensity of the third harmonic peak at the first R_ω spectrum minimum to the left of the photonic band gap (the short-wavelength edge of the photonic band gap) increases by no less than two orders of magnitude as compared to the intensity of the peaks

within the photonic band gap. At $\Delta n_{3\omega} > 0$, the third harmonic intensity at the long-wavelength edge of the photonic band gap increases by approximately an order of magnitude.

The correspondence between the sign of dispersion and the spectral position of the third-harmonic intensity peak with respect to the photonic band gap is unambiguous; namely, at $\Delta n_{3\omega} < 0$ and $\Delta n_{3\omega} > 0$, third-harmonic intensity resonances are observed only at the short- and long-wavelength photonic band gap edges, respectively. This is clearly seen in Fig. 6, where the dependence of the amplitude of the third-harmonic intensity peaks $I_{3\omega}^{\max}$ on dispersion $\Delta n_{3\omega}$ are shown at the short- and long-wavelength photonic band gap edges. The main conclusion is that we do not observe simultaneous enhancement of the third harmonic at both photonic band gap edges. The $I_{3\omega}^{\max}(\Delta n_{3\omega})$ dependences are oscillatory in character. Their maxima that appear, for instance, at dispersion $\Delta n_{3\omega}$ values equal to 0.051, –0.114, and –0.045 are reached when λ_ω coincides with the minimum of the spectrum of the reflection coefficient for the fundamental wave R_ω and one of the minima of the $R_{3\omega}$ spectrum is observed at the third harmonic wavelength (Figs. 5c, 5d). The farther from the photonic band gap minimum of the $R_{3\omega}$ spectra that coincides with the photonic band gap edge in the R_ω spectrum corresponds to the weaker the third harmonic resonance. It follows that a key factor of third-harmonic enhancement is the coincidence of the pumping wavelength with the photonic band gap edge and of the third-harmonic wavelength with the minimum of the reflection coefficient $R_{3\omega}$.

Third-harmonic resonances are observed not only at the photonic band gap edge but also at the fundamental wavelengths at which the R_ω spectrum has the second, third, etc., minima. These resonances are enhanced to a lesser degree than the peaks at the edge of the photonic band gap. This is clearly seen from Fig. 7, where the whole set of all third harmonic frequency spectra obtained as $\Delta n_{3\omega}$ changed from –0.21 to 0.19 in steps of 0.002 in wavelength steps of 1 nm is shown. Peaks at the edges of the photonic band gap are quite pronounced. The enhancement at the short-wavelength edge is more intense, and the enhancement at the other R_ω spectrum minima is much weaker.

The resonance enhancement of third-harmonic generation at the edge of the photonic band gap is determined by multiple-reflection interference of both fundamental and third-harmonic waves, because the strongest enhancement of the third harmonic is observed when the third-harmonic wavelength occurs at an $R_{3\omega}$ spectrum minimum. To elucidate the role played by interference effects at the fundamental wavelength we must use (4) to calculate the spatial distribution of the amplitude of the fundamental wave within the photonic

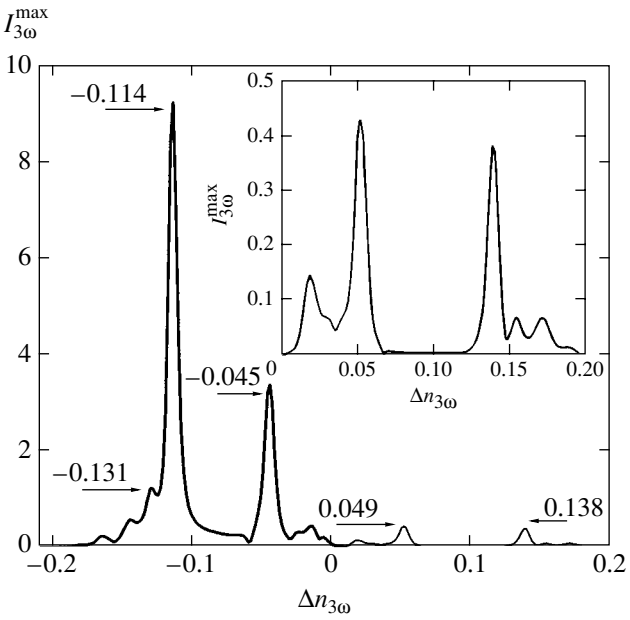

Fig. 6. Dispersion $\Delta n_{3\omega}$ dependences of maximum third-harmonic peak intensity at the right (thin line) and left (thick line) edges of the photonic band gap. Arrows show the dispersion values that correspond to the characteristic third-harmonic intensity maxima. The region with positive $\Delta n_{3\omega}$ values is shown in the inset on an enlarged scale.

crystal (Fig. 3) and the amplitude of inhomogeneous third harmonic waves at the boundary of each layer [Eq. (16)]. The multiple-reflection interference effects at the third-harmonic wavelength are determined by comparing the amplitude and phase of the output homogeneous third-harmonic waves $E_1^-(j)$ generated separately by each jth layer [see Eq. (8)] and the amplitude and phase of inhomogeneous third harmonic waves at the boundary of each layer.

The spatial distribution of the amplitude of the inhomogeneous third-harmonic wave inside the photonic crystal when λ_ω corresponds to the edge of the photonic band gap is shown in Fig. 8a. The shape of the $E^{(s)}(j)$ dependence reproduces that of the distribution of the fundamental wave amplitude. When the dispersion $\Delta n_{3\omega}$ is introduced into calculations, the shape of the $E^{(s)}(j)$ dependence does not change and enhancement in the middle of the photonic crystal caused by fundamental field localization is retained, while the amplitude of the inhomogeneous wave decreases as $\Delta n_{3\omega}$ increases. The phase distribution of the inhomogeneous third-harmonic wave at the boundary of each layer is determined by the fundamental field phase at the same boundary. The phase difference between third-harmonic fields in neighboring layers with equal refractive indexes is close to π. For this reason, each point of the polar diagram in which inhomogeneous third harmonics are shown with their phases (Fig. 8b) has a corre-

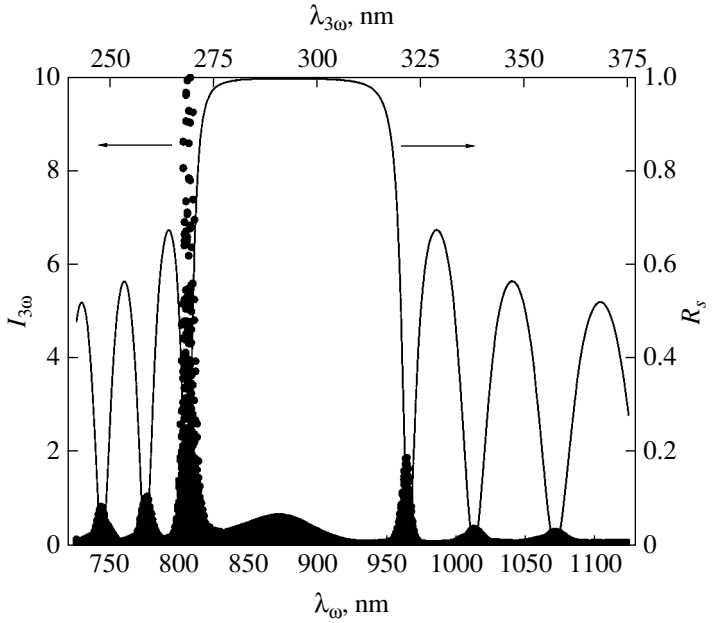

Fig. 7. Solid circles are the whole set of the third-harmonic intensity spectra obtained with $\Delta n_{3\omega}$ varied from –0.21 to 0.19, and the solid line is the linear reflection coefficient spectrum.

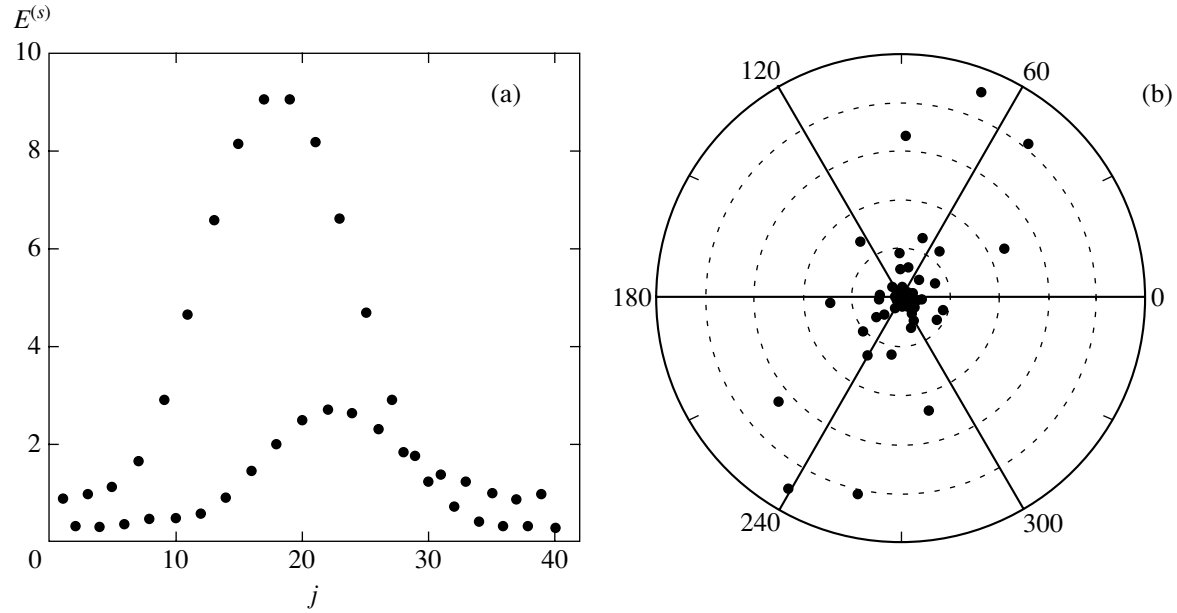

Fig. 8. (a) Dependence of the amplitude of the inhomogeneous third-harmonic wave determined by (16) on the number of the photonic crystal layer and (b) amplitudes and phases of inhomogeneous third-harmonic waves. The calculations were performed for $\Delta n_{3\omega} = 0$.

sponding point with a comparable amplitude and an almost opposite phase. The phase of inhomogeneous waves does not change at $\Delta n_{3\omega} \neq 0$.

The dependences of the amplitudes and phases of the third-harmonic partial waves E_1^- emerging from the photonic crystal on the layer number j are shown in Fig. 9. These dependences were calculated at the maxima of the $I_{3\omega}^{\max}(\Delta n_{3\omega})$ dependence (Figs. 9a–9d) and at its minimum (Figs. 9e, 9f). The amplitudes of inhomogeneous third-harmonic waves at the minimum and maximum of the $I_{3\omega}^{\max}(\Delta n_{3\omega})$ dependence are of the same order of magnitude (Fig. 8), and the amplitude of emerging free third harmonic waves at the minimum of the $I_{3\omega}^{\max}(\Delta n_{3\omega})$ dependence is by a factor no less than 3

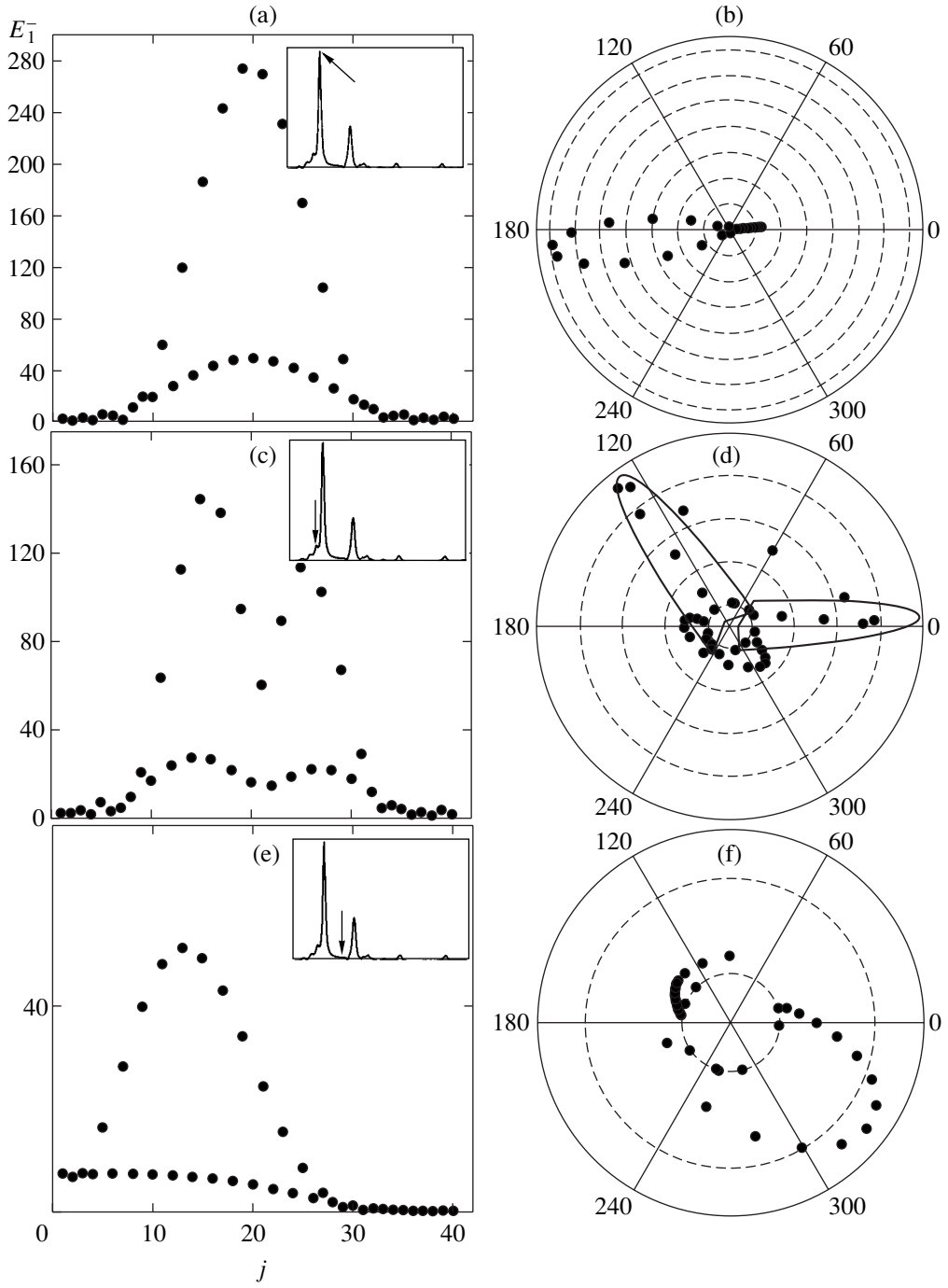

Fig. 9. (a, c, e) Dependences of the partial third-harmonic E_1^- amplitude on the layer number j calculated when the fundamental wavelength is tuned to the edge of the photonic band gap for dispersion $\Delta n_{3\omega}$ values of (a) –0.114, (c) –0.131, and (e) –0.085 and (b, d, f) partial third-harmonic waves with their phases. The insets show schematically the $I_{3\omega}^{\max}(\Delta n_{3\omega})$ dependence shown in Fig. 6. The points at which the calculations were made are marked by arrows.

smaller than at its maximum. This implies additional amplitude enhancement of the emerging third harmonic by constructive multiple-reflection interference of third-harmonic waves; namely, the partial third harmonic-wave generated in the jth nonlinear layer reaches an interference maximum outside the photonic crystal. This is accompanied by phase matching of partial third-harmonic fields outside the photonic crystal, and the phases of the E_1^- waves become localized in a narrow angular interval. The weakest phase matching of partial third-harmonic waves is observed at the min-

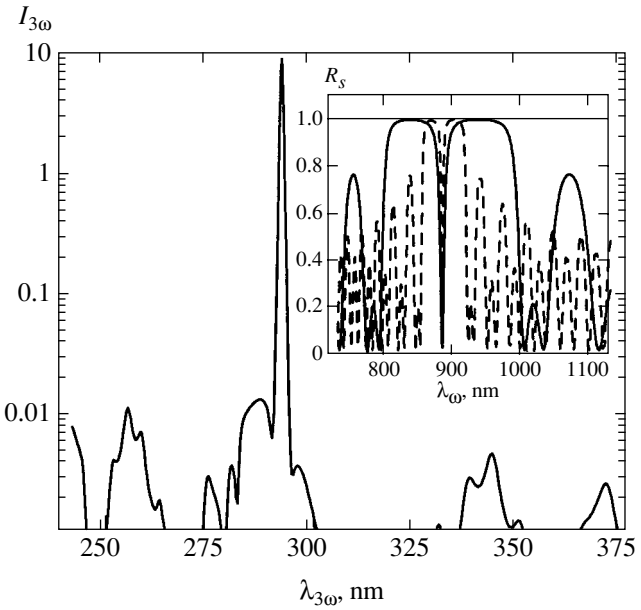

Fig. 10. The intensity spectrum of the third harmonic generated in a microcavity and calculated for the *SS* geometry. Shown in the inset are the reflection coefficient spectra of an *s*-polarized wave calculated in the pumping (solid line) and third harmonic (dashed line) wavelength ranges. The $R_{3\omega}$ spectrum is plotted on a triply enlarged wavelength scale.

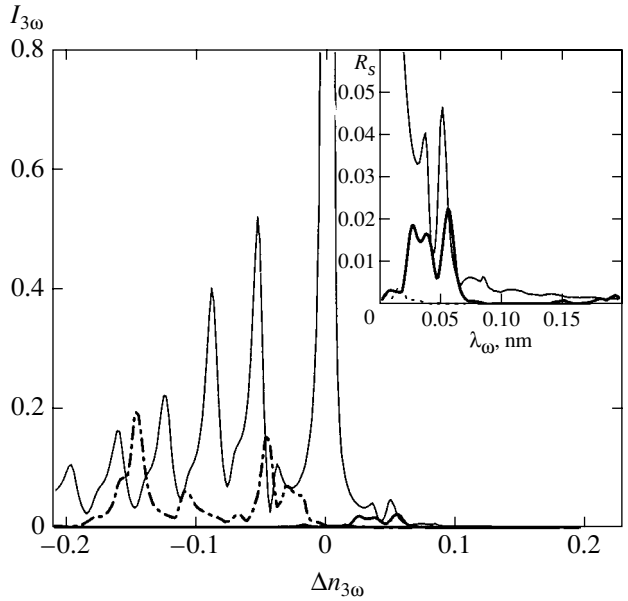

Fig. 11. Maximum intensity of third-harmonic peaks generated in the microcavity mode (thin line) and at the long-wavelength (thick line) and short-wavelength (dot-and-dash line) edges of the photonic band gap as a function of dispersion $\Delta n_{3\omega}$; the dependences in the $\Delta n_{3\omega} > 0$ region are shown in the inset on an enlarged scale.

imum of the $I_{3\omega}^{\max}(\Delta n_{3\omega})$ dependence. The phase jump through $2\pi/3$ in Fig. 9d explains the reason of a maximum third-harmonic enhancement when the photonic band gap edges in the R_ω spectrum and the minima in the $R_{3\omega}$ spectra closest to the photonic band gap coincide; namely, the larger the $\Delta n_{3\omega}$ value, the larger phase changes of partial third-harmonic waves and the stronger their destructive interference. As a result, the highest intensity of the total emerging third harmonic is attained because, first, the amplitude of the emerging partial third-harmonic waves increases and, second, their phases are matched (Figs. 9a, 9b). The smallest $I_{3\omega}^{\max}$ value is observed when the photonic band gap in the R_ω spectrum corresponds to the edge of the photonic band gap in the $R_{3\omega}$ spectrum (Figs. 9e, 9f). The maximum of the $E_1^-(j)$ dependence for the selected photonic crystal parameters is then shifted by six layers toward the photonic crystal surface with respect to the maximum of the spatial distribution of the inhomogeneous third-harmonic wave amplitude. This is explained by the third-harmonic wavelength falling within the photonic band gap in the $R_{3\omega}$ spectrum. The contributions of deeper photonic crystal layers to the emerging third harmonic then exponentially decrease.

The plots in Fig. 9 were constructed for $\Delta n_{3\omega} < 0$, when the third harmonic resonance is observed at the short-wavelength edge of the photonic band gap. The dependences for $\Delta n_{3\omega} > 0$ are similar.

4. THIRD HARMONIC GENERATION IN MICROCAVITIES

The dependence of the intensity of the third harmonic generated in a microcavity on its wavelength is shown in Fig. 10. The third-harmonic spectrum has a peak corresponding to the resonance between the fundamental radiation and the microcavity mode. Its intensity is more than three orders of magnitude higher than the intensity of the third harmonic in other spectral regions. No enhancement is observed at the edge of the photonic band gap. If $\Delta n_{3\omega} \neq 0$, the resonance peak amplitude substantially changes and THG enhancement at the edge of the photonic band gap arises. The $I_{3\omega}^{\max}(\Delta n_{3\omega})$ dependences of the maxima of third-harmonic intensity peaks at the resonance between the fundamental radiation and the mode and between the fundamental radiation and the short- and long-wavelength edges of the photonic band gap are shown in Fig. 11. If $\Delta n_{3\omega} < 0$, the resonance amplitudes in the microcavity mode and at the left edge of the photonic band gap increase. If $\Delta n_{3\omega} > 0$, we observe enhancement in the mode and at the right edge of the photonic band gap. THG enhancement in the microcavity mode is maximum at zero dispersion. If dispersion is nonzero, its amplitude decreases by approximately an order of magnitude and oscillates as a function of $\Delta n_{3\omega}$.

The localization of the fundamental field resonant to the mode results in a very substantial increase in the amplitude of partial third-harmonic waves, which

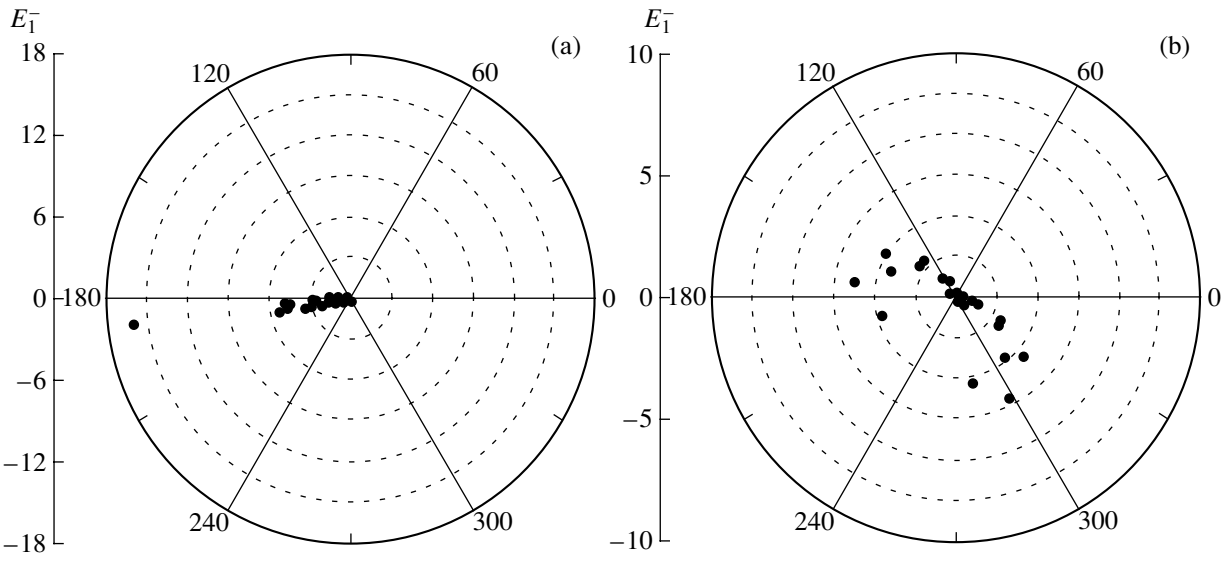

Fig. 12. Polar diagrams of complex partial third harmonic waves $E_1^-(j)$ corresponding to (a) the maximum of the $I_{3\omega}^{\max}(\Delta n_{3\omega})$ dependence, $\Delta n_{3\omega} = 0$ and (b) one of the minima of the $I_{3\omega}^{\max}(\Delta n_{3\omega})$ dependence, $\Delta n_{3\omega} = -0.04$.

depend on the third power of the fundamental field amplitude. Irrespective of the phase difference between these waves, the minimum intensity of the emerging third-harmonic wave is then no less than the maximum amplitude of the third-harmonic wave generated at the edge of the photonic band gap. The partial third-harmonic wave amplitudes in the microcavity mode remain almost invariant as $\Delta n_{3\omega}$ is varied, whereas their phases change substantially. The $E_1^-(j)$ phases at the $I_{3\omega}^{\max}(\Delta n_{3\omega})$ peaks are nearly equal (see Fig. 12a). At the minimum reached at $\Delta n_{3\omega} = -0.04$ (Fig. 12b), the partial waves $E_1^-(j)$ are out of phase, and their interference is destructive.

It follows that, when the fundamental wavelength is tuned to the photonic band gap edges and the microcavity mode, the fundamental wave is localized in the neighborhood of the cavity layer. The fundamental field enhancement is stronger when the fundamental wavelength is in resonance with the microcavity mode. As a result, the amplitude of inhomogeneous partial third-harmonic waves increases. At certain dispersion $\Delta n_{3\omega}$ values, the amplitudes of partial homogeneous third-harmonic waves become maximum because of multiple-reflection interference and their phases are close to each other. This results in a resonant increase in the third-harmonic intensity. The third-harmonic intensity peak in the microcavity mode is less sensitive to phase matching of partial waves because of strong fundamental field amplitude enhancement. THG enhancement in the microcavity mode is observed almost irrespective of the dispersion value, whereas the peaks at the short-wavelength edge of the photonic band gap appear at a negative dispersion, and those at the long-wavelength edge, when dispersion is positive.

5. AN EXPERIMENTAL STUDY OF THIRD HARMONIC GENERATION IN MICROCAVITIES

We experimentally studied THG in microcavities. The sample was a one-dimensional microcavity made from mesoporous silicon by electrochemical etching of a heavily doped *p*-type silicon plate with a (100) crystallographic orientation in a solution of hydrofluoric acid following the procedure described in [11]. The sample was two one-dimensional photonic crystals consisting of five pairs of quarter-wave ($\lambda_0 = 1300$ nm) porous silicon layers separated by a half-wave cavity layer. The λ_0 value is the spectral position of the microcavity mode at the normal fundamental wave incidence. The refractive index and thickness were $n_\omega^H \approx 1.93$ and $d_H \approx 170$ nm for optically denser layers and $n_\omega^L \approx 1.61$ and $d_L \approx 200$ nm for less dense ones. The porous silicon microcavity layer has the refractive index $n^L = 1.61$ and thickness $d_{\text{res}} = 400$ nm. The spectral dependences of the third harmonic intensity were measured using a tunable optical parametric generator. Its output linearly polarized radiation had the following characteristics: pulse width of 4 ns, pulse energy of about 10 mJ at a 800 nm wavelength, and tuning range of 410–690 nm for the signal wave and 735–2200 nm for the idle wave. We used the idler tuning range because the photonic

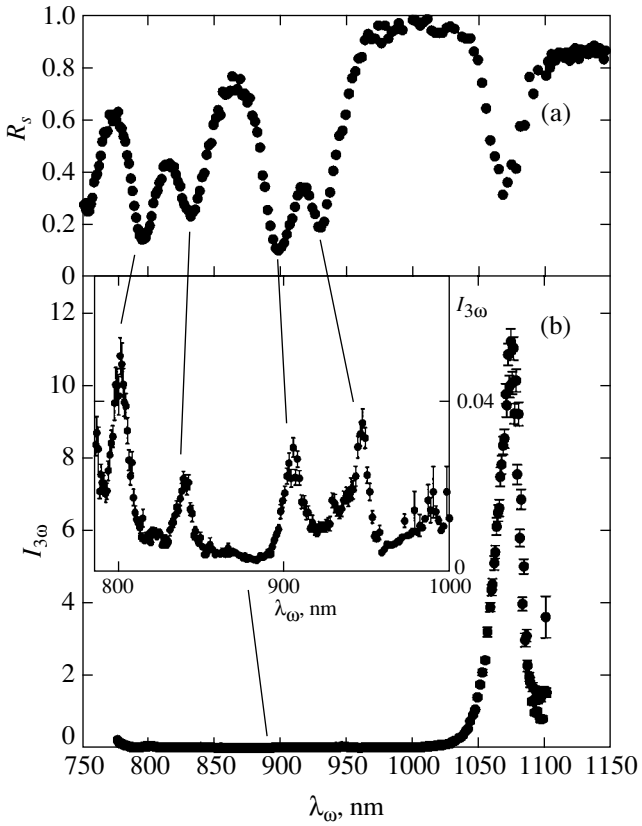

Fig. 13. (a) Spectrum of the reflection coefficient of the *s*-polarized pumping wave and (b) dependence of the intensity of the *s*-polarized third harmonic wave on the wavelength of *s*-polarized pumping. The sample is a microcavity made from porous silicon with $\lambda_0 = 1300$ nm.

band gap and the allowed mode of the microcavity were within it. Third-harmonic radiation reflected from the microcavity sample was separated by UV filters and detected by a photoelectron multiplier. The frequency spectra of third-harmonic intensity were measured for fundamental and third-harmonic waves polarized in the plane of the sample (*SS* geometry) at a 60° angle of fundamental wave incidence.

The third-harmonic intensity spectrum is shown in Fig. 13b, and the spectrum of the linear reflection coefficient for the fundamental wave is shown in Fig. 13a. The $I_{3\omega}(\lambda_\omega)$ spectral dependence has a resonance at a fundamental wavelength of $\lambda_\omega \approx 1075$ nm, which coincides with the microcavity mode wavelength shifted to shorter waves with respect to λ_0 at an oblique angle of incidence. We also observe resonant third-harmonic features when λ_ω is tuned to the spectral regions of the photonic band gap edge and outside the gap (inset to Fig. 13). The third-harmonic intensity in these resonance peaks is three orders of magnitude lower than in the microcavity mode. The positions of third-harmonic intensity resonances correlate with the minima of the spectrum of the linear reflection coefficient for the fundamental wave.

6. CONCLUSIONS

We developed a phenomenological model of optical THG in one-dimensional photonic crystals and microcavities. The model is based on the formalism of transfer matrices. It describes the generation and propagation of third-harmonic waves through a photonic crystal by taking into account multiple reflections of fundamental and third-harmonic wave at interfaces and interference between the homogeneous and inhomogeneous third-harmonic waves. The spectra of third-harmonic intensity in photonic crystals and microcavities were studied for the example of nonlinear optical media with the limiting point symmetry groups $\infty 2$, ∞m, and ∞/mm of layers constituting photonic crystals. The spectra were studied in the spectral range of wavelengths and incidence angles of fundamental waves containing the photonic band gap and regions near its edges. At the resonance between the fundamental wave and the microcavity mode, the third-harmonic intensity increases by more than three orders of magnitude. When the fundamental wavelength coincides with photonic band gap edges, we observe resonance enhancement of third-harmonic intensity, which depends on the magnitude and sign of dispersion. If the refractive indexes of photonic crystal layers at the third-harmonic wavelength $n_{3\omega}$ are smaller than the refractive indexes at the fundamental wavelength, we observe resonant third-harmonic enhancement by a factor exceeding 100 at the short-wavelength edge of the photonic band gap. If $n_{3\omega} > n_\omega$, THG is enhanced at the long-wavelength edge of the photonic band gap.

The main mechanism of resonant THG enhancement in the microcavity mode is fundamental field localization, which manifests itself by a four- to tenfold increase in the third-harmonic amplitude. An additional enhancement factor is phase matching of partial third-harmonic waves from each photonic crystal layer when they leave the microcavity. THG enhancement at the edges of the photonic band gap is caused to a greater extent by an increase in the amplitude of emerging partial third-harmonic waves and their phase matching as a result of multiple-reflection interference of third-harmonic waves and, to a lesser extent, by fundamental wave localization within the photonic crystal.

In agreement with the calculated third-harmonic spectra, the experimental spectrum of the intensity of the third harmonic generated in a one-dimensional microcavity fabricated from porous silicon shows that the intensity of the third harmonic increases approximately 1000 times in the microcavity mode. Additional third-harmonic peaks are observed at the left edge of the photonic band gap and outside the gap. The spectral positions of the resonances coincide with the minima of the linear fundamental wave reflection coefficient to within several nanometers.

ACKNOWLEDGMENTS

The authors are deeply indebted to O.A. Aktsipetrov for the statement of the problem and numerous useful discussions.

REFERENCES

1. K. Sakoda, *Optical Properties of Photonic Crystals* (Springer, Berlin, 2001).
2. W. Chen and D. L. Mills, Phys. Rev. Lett. **58**, 160 (1987).
3. M. Scalora, J. P. Dowling, C. M. Bowden, and M. J. Bloemer, Phys. Rev. Lett. **73**, 1368 (1994).
4. A. V. Andreev, A. V. Balakin, A. B. Kozlov, *et al.*, J. Opt. Soc. Am. B **19**, 1865 (2002).
5. I. Freund, Phys. Rev. Lett. **21**, 1404 (1968).
6. N. Bloembergen and J. Sievers, Appl. Phys. Lett. **17**, 483 (1970).
7. J. P. van der Ziel and M. Ilegems, Appl. Phys. Lett. **28**, 437 (1976).
8. J. Martorell, R. Vilaseca, and R. Corbalan, Appl. Phys. Lett. **70**, 702 (1997).
9. A. V. Balakin, V. A. Bushuev, N. I. Koroteev, *et al.*, Opt. Lett. **24**, 793 (1999).
10. Y. Dumeige, P. Vidakovic, S. Sauvage, *et al.*, Appl. Phys. Lett. **78**, 3021 (2001).
11. T. V. Dolgova, A. I. Maidykovski, M. G. Martemyanov, *et al.*, J. Opt. Soc. Am. B **19**, 2129 (2002).
12. V. V. Konotop and V. Kuzmiak, J. Opt. Soc. Am. B **16**, 1370 (1999).
13. V. V. Konotop and V. Kuzmiak, J. Opt. Soc. Am. B **17**, 1874 (2000).
14. M. Centini, G. D'Aguanno, M. Scalora, *et al.*, Phys. Rev. E **64**, 46606 (2001).
15. T. V. Dolgova, A. I. Maĭdykovskiĭ, M. G. Martem'yanov, *et al.*, Pis'ma Zh. Éksp. Teor. Fiz. **73**, 8 (2001) [JETP Lett. **73**, 6 (2001)].
16. T. V. Dolgova, A. I. Maidykovski, M. G. Martemyanov, *et al.*, Appl. Phys. Lett. **81**, 2725 (2002).
17. T. V. Dolgova, A. I. Maĭdykovskiĭ, M. G. Martem'yanov, *et al.*, Pis'ma Zh. Éksp. Teor. Fiz. **75**, 17 (2002) [JETP Lett. **75**, 15 (2002)].
18. G. D'Aguanno, M. Centini, M. Scalora, *et al.*, J. Opt. Soc. Am. B **19**, 2111 (2002).
19. A. V. Andreev, A. V. Balakin, A. B. Kozlov, *et al.*, J. Opt. Soc. Am. B **19**, 2083 (2002).
20. M. Born and E. Wolf, *Principles of Optics*, 2nd ed. (Pergamon, New York, 1964; Nauka, Moscow, 1970).
21. N. Hashizume, M. Ohashi, T. Kondo, and R. Ito, J. Opt. Soc. Am. B **12**, 1894 (1995).
22. J. E. Sipe, J. Opt. Soc. Am. B **4**, 481 (1987).
23. D. S. Bethune, J. Opt. Soc. Am. B **6**, 910 (1989).
24. N. Bloembergen and P. S. Pershan, Phys. Rev. **128**, 606 (1962).
25. Yu. I. Sirotin and M. P. Shaskolskaya, *Fundamentals of Crystal Physics*, 2nd ed. (Nauka, Moscow, 1979; Mir, Moscow, 1982).

Translated by V. Sipachev

Journal of Experimental and Theoretical Physics, Vol. 98, No. 3, 2004, pp. 478–482.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 125, No. 3, 2004, pp. 543–547.

ATOMS, SPECTRA, RADIATION

Angular Distribution of Transient Radiation from an Ultrarelativistic Particle in a Magnetic Field

M. I. Ryazanov

Moscow Institute of Engineering Physics (Technological University), Kashirskoe sh. 31, Moscow, 115409 Russia

e-mail: ryazanov@theor.mephi.msk.su

Received September 5, 2003

Abstract—It is shown that a magnetic field acting on an ultrarelativistic charged particle escaping from a conductor changes the intensity of transient radiation. The angular and frequency distribution of transient radiation in the magnetic field is determined. The possibility of determining the energy of the ultrarelativistic particle from the change in the azimuthal asymmetry of transient radiation emitted by this particle in the magnetic field is discussed. © *2004 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

It is well known that the energy and momentum conservation laws hold for the emission of a photon by a free ultrarelativistic particle for a small longitudinal (directed along the particle velocity) momentum transfer,

$$\Delta p \sim \frac{h\omega}{c}\left(\frac{mc^2}{E}\right)^2,$$

so that the wave process of radiation formation takes place over a particle path of length

$$\frac{h}{\Delta p} = \frac{c}{\omega}\left(\frac{E}{mc^2}\right)^2.$$

For ultrarelativistic particles, the length of radiation formation (coherence length) may be of a macroscopic size. On such a long path, the competing processes may significantly change the motion of a particle in the region of radiation formation, thus lowering the radiation intensity. An example of this effect is the influence of multiple scattering on bremsstrahlung [1–3] or the effect of polarization of the medium on bremsstrahlung [4].

In the case of an ultrarelativistic particle and the conductor-vacuum interface, transient radiation emerging when the charged particle crosses the interface between the two media [5–7] is also formed in a spatial region having a length of $(c/\omega)(E/mc^2)^2$ and a much smaller transverse size. In view of the large coherence length, the action of an external field on the particle may change the type of motion of the particle in the course of radiation formation. This affects the process of radiation formation since the angular distribution of radiation changes and its intensity decreases.

In an external electric field, for a particle escaping from a conductor along the normal to the interface, this effect leads to the emergence of azimuthal asymmetry in the distribution of transient radiation, which strongly depends on the Lorentz factor γ [8]. The action of a magnetic field on transient radiation differs from the action of an electric field; it would hence be interesting to estimate the effect of a magnetic field on transient radiation emitted by an ultrarelativistic particle.

2. TRANSIENT RADIATION EMITTED BY A NONUNIFORMLY MOVING CHARGE

Let a particle bearing a charge e and having a velocity $v \approx c$ escape at instant $t = 0$ from a conductor ($z < 0$) into vacuum ($z > 0$), where it is acted upon by a constant uniform magnetic field **H** parallel to the conductor surface. We choose the direction of the x axis along the magnetic field and consider transient radiation from the particle escaping along the z axis perpendicular to the conductor surface. The law of motion of the particle can be written in the form

$$\mathbf{r}(t) = \mathbf{R}(t) + \mathbf{s}(t),$$

where $\mathbf{R}(t)$ and $\mathbf{s}(t)$ are the normal and tangential components of the radius vector of the particle relative to the surface, respectively. It is well known that, for such a motion of the charge, the field outside the conductor coincides with the field of two charges escaping from the same point $\mathbf{r} = 0$ at instant $t = 0$, viz., charge e moving according to the law $\mathbf{r}(t)$ and an image charge $-e$ moving according to the law

$$\mathbf{r}(t) = -\mathbf{R}(t) + \mathbf{s}(t)$$

(see [9]).

For such motion of the two charges, the angular and frequency distribution of the radiant energy has the form ($\mathbf{n} = \mathbf{r}/r$, $\mathbf{v}(t) = d\mathbf{R}/dt$, $\mathbf{u}(t) = d\mathbf{s}/dt$)

$$\begin{aligned}\frac{d^2E}{d\omega d\Omega} &= \frac{e^2\omega^2}{4\pi^2c^3}\left|\int_0^\infty dt[\mathbf{n}\{\mathbf{v}(t)+\mathbf{u}(t)\}]\right.\\ &\times\exp\{i\omega t - i\mathbf{k}\cdot\mathbf{R}(t) - i\mathbf{k}\cdot\mathbf{s}(t)\}\\ &+\int_0^\infty dt[\mathbf{n}\{\mathbf{v}(t)-\mathbf{u}(t)\}]\\ &\left.\times\exp\{i(\omega + i\mathbf{k}\cdot\mathbf{R}(t) - i\mathbf{k}\cdot\mathbf{s}(t)\}\right|^2.\end{aligned} \tag{1}$$

The action of the field may be significant when the terms independent of the field are partly cancelled out. For example, in the ultrarelativistic case, the exponent of one of the exponential functions is on the order of $\omega - \mathbf{k}\cdot\mathbf{v}$ and is smaller than the exponent of the other exponential, which is on the order of $\omega + \mathbf{k}\cdot\mathbf{v}$. The integral with the rapidly oscillating exponential is small and can be omitted. The effect of the external field can be disregarded in the first approximation in all cases when the main terms are not mutually cancelled out. This enables us to transform relation (1) to

$$\begin{aligned}\frac{d^2E}{d\omega d\Omega} &= \frac{e^2\omega^2}{4\pi^2c^3}\left|\int_0^\infty dt[\mathbf{n}\{\mathbf{v}(t)+\mathbf{u}(t)\}]\right.\\ &\left.\times\exp\{i(\omega - i\mathbf{k}\cdot\mathbf{R}(t) - i\mathbf{k}\cdot\mathbf{s}(t)\}\right|^2.\end{aligned} \tag{2}$$

3. TRANSIENT RADIATION IN A MAGNETIC FIELD

With the above choice of the coordinate axes, the velocity components of an ultrarelativistic particle in a magnetic field have the form ($\Omega \equiv eH/mc\gamma$)

$$u_x(t) = 0, \quad u_y(t) = v_0\sin\Omega t, \quad v(t) = v_0\cos\Omega t. \tag{3}$$

It follows hence that the law of motion of the particle can be written in the form

$$\begin{aligned}x(t) = 0, \quad y(t) &= (v_0/\Omega)(1-\cos\Omega t),\\ z(t) &= (v_0/\Omega)\sin\Omega t.\end{aligned} \tag{4}$$

We confine our analysis to the case when the change in the particle velocity due to the action of the magnetic field in the region of radiation formation considered here is small as compared to the initial velocity of the particle. This means that the inequality $u_y \ll v_0$ is satisfied; i.e., $\Omega t \ll 1$. In this case, the law of motion of the charge can be represented in the form

$$\begin{aligned}u_x(t) = 0, \quad u_y(t) &\approx v_0\Omega t, \quad v(t) \approx v_0,\\ x(t) = 0, \quad y(t) &= (v_0\Omega)t^2/2, \quad z(t) = v_0 t.\end{aligned} \tag{5}$$

The law of motion of the image charge in this approximation has the form

$$\begin{aligned}u_x(t) = 0, \quad u_y(t) &\approx v_0\Omega t, \quad v(t) \approx -v_0,\\ x(t) = 0, \quad y(t) &= (v_0\Omega)t^2/2, \quad z(t) = -v_0 t.\end{aligned} \tag{6}$$

Considering that $v_0 \gg u$, we can disregard quantity u in the preexponential factor in expression (2); in this case, the angular and frequency distribution of the energy of transient radiation can be written in the form ($Q \equiv k_y v_0\Omega/2$)

$$\begin{aligned}\frac{d^2E}{d\omega d\Omega} &= \frac{e^2\omega^2}{4\pi^2c^3}[\mathbf{n}\cdot\mathbf{v}_0]^2\\ &\times\left|\int_0^\infty dt\exp\{i(\omega t + -i\mathbf{k}\cdot\mathbf{v}_0 t - iQt^2)\}\right|^2.\end{aligned} \tag{7}$$

The integral contained in this formula cannot be reduced to elementary functions and can be expressed in terms of the Fresnel integrals

$$\begin{aligned}S(x) &= (2/\pi)^{1/2}\int_0^x \sin^2 t\,dt,\\ C(x) &= (2/\pi)^{1/2}\int_0^x \cos^2 t\,dt.\end{aligned} \tag{8}$$

For $\Omega < 0$, i.e., for $k_y = k\sin\vartheta\sin\varphi < 0$, integration gives

$$\begin{aligned}&\int_0^\infty dt\exp\{i(\omega - \mathbf{k}\cdot\mathbf{v})t + i|Q|t^2\}\\ &= \left(\frac{\pi}{2|Q|}\right)^{1/2}\exp\left\{-\frac{i(\omega-\mathbf{k}\cdot\mathbf{v})^2}{4|Q|}\right\}\\ &\times\left\{\left[\frac{1}{2} - C\left(\frac{\omega-\mathbf{k}\cdot\mathbf{v}}{2|Q|^{1/2}}\right)\right] + i\left[\frac{1}{2} - S\left(\frac{\omega-\mathbf{k}\cdot\mathbf{v}}{2|Q|^{1/2}}\right)\right]\right\}.\end{aligned} \tag{9}$$

The angular and frequency distribution of transient

radiation for $\sin\vartheta\sin\varphi < 0$ has the form

$$\frac{d^2E}{d\omega d\Omega} = \frac{e^2\omega^2}{\pi c^3}\frac{[\mathbf{n}\times\mathbf{v}]^2}{|Q|} \times\left\{\left[\frac{1}{2} - C\left(\frac{\omega-\mathbf{k}\cdot\mathbf{v}}{2|Q|^{1/2}}\right)\right]^2 + \left[\frac{1}{2} - S\left(\frac{\omega-\mathbf{k}\cdot\mathbf{v}}{2|Q|^{1/2}}\right)\right]^2\right\}. \quad (10)$$

For $\Omega > 0$, i.e., for $k_y = k\sin\vartheta\sin\varphi > 0$, integration yields

$$\int_0^\infty dt\exp\{i(\omega-\mathbf{k}\cdot\mathbf{v})t - i|Q|r^2)\} = \left(\frac{\pi}{2|Q|}\right)^{1/2}\exp\left\{\frac{i(\omega-\mathbf{k}\cdot\mathbf{v})^2}{4|Q|}\right\} \times\left\{\left[\frac{1}{2} + C\left(\frac{\omega-\mathbf{k}\cdot\mathbf{v}}{2|Q|^{1/2}}\right)\right] - i\left[\frac{1}{2} + S\left(\frac{\omega-\mathbf{k}\cdot\mathbf{v}}{2|Q|^{1/2}}\right)\right]\right\}. \quad (11)$$

In contrast to expression (10), the angular and frequency distribution of transient radiation for $\sin\vartheta\sin\varphi > 0$ assumes the form

$$\frac{d^2E}{d\omega d\Omega} = \frac{e^2\omega^2}{\pi c^3}([\mathbf{n}\times\mathbf{v}]^2/|Q|) \times\left\{\left[\frac{1}{2} + C\left(\frac{\omega-\mathbf{k}\cdot\mathbf{v}}{2|Q|^{1/2}}\right)\right]^2 - \left[\frac{1}{2} + S\left(\frac{\omega-\mathbf{k}\cdot\mathbf{v}}{2|Q|^{1/2}}\right)\right]^2\right\}. \quad (12)$$

For large x, Fresnel integrals $C(x)$ and $S(x)$ oscillate in the vicinity of 1/2 with an amplitude decreasing slowly upon an increase in x:

$$S(x) = 1/2 - (1/2\pi)^{1/2}(1/x)\cos x^2, \quad C(x) = 1/2 + (1/2\pi)^{1/2}(1/x)\sin x^2. \quad (13)$$

For small x, the Fresnel integrals increase rapidly from zero for $x = 0$ to values on the order of unity for $x \sim 1$. In the vicinity of zero, we have

$$S(x) = (2/\pi)^{1/2}(x^3/3), \quad C(x) = (2/\pi)^{1/2}x. \quad (14)$$

We introduce auxiliary functions $f(x)$ and $g(x)$, defined by the equations

$$1/2 - S(x) = g(x)\sin(\pi x^2/2) + f(x)\cos(\pi x^2/2), \quad (15)$$

$$1/2 - C(x) = g(x)\cos(\pi x^2/2) - f(x)\sin(\pi x^2/2). \quad (16)$$

For $\sin\vartheta\sin\varphi < 0$, the radiant energy distribution (10) can be written as

$$\frac{d^2E}{d\omega d\Omega} = \frac{e^2\omega^2}{\pi^2c^3}\left(\frac{[\mathbf{n}\cdot\mathbf{v}]^2}{|Q|}\right) \times\left\{g^2\left(\frac{\omega-\mathbf{k}\cdot\mathbf{v}}{2|Q|^{1/2}}\right) + f^2\left(\frac{\omega-\mathbf{k}\cdot\mathbf{v}}{2|Q|^{1/2}}\right)\right\}. \quad (17)$$

In the region $0 \le x < \infty$, auxiliary functions $f(x)$ and $g(x)$ can be approximated with an error smaller than 2×10^{-3} by the expressions [10]

$$f(x) \approx \frac{1 + 0.926x}{2 + 1.792x + 3.104x^2}, \quad g(x) \approx \frac{1}{2 + 4.142x + 3.492x^2 + 6.670x^3}. \quad (18)$$

More accurate approximations can be found in [11].

4. EFFECT OF A MAGNETIC FIELD ON TRANSIENT RADIATION

As the magnetic field tends to zero, the argument $(\omega - \mathbf{k}\cdot\mathbf{v})/2|Q|^{1/2}$ of the Fresnel integrals in expressions (10) and (12) tends to infinity. Consequently, in the range of frequencies and angles for which

$$(\omega-\mathbf{k}\cdot\mathbf{v}) \gg |Q|^{1/2} \sim \left|\frac{k_y v_0 eH}{mc\gamma}\right|^{1/2},$$

the magnetic field practically does not affect transient radiation. The angular and frequency distribution of transient radiation is deformed by the magnetic field when the opposite inequality holds,

$$\left|\frac{v_0 eH\omega\vartheta\sin\varphi}{mc^2\gamma}\right|^{1/2} \gg (\omega-\mathbf{k}\cdot\mathbf{v})^2 \sim \omega^2\left(\vartheta + \frac{1}{\gamma^2}\right)^2. \quad (19)$$

For characteristic angles of $\vartheta \sim 1/\gamma$ of the emission of radiation, this inequality assumes the form

$$\left|\frac{eH\sin\varphi}{mc^2}\right| \gg \frac{\omega}{\gamma^4}. \quad (20)$$

When inequalities (19) and (20) are satisfied, the argument of the Fresnel integrals is small; consequently, we can use approximate expressions (14) for the Fresnel integrals. In region (20), we have

$$C\left(\frac{\omega-\mathbf{k}\cdot\mathbf{v}}{2|Q|^{1/2}}\right) \approx \frac{\omega-\mathbf{k}\cdot\mathbf{v}}{|\pi Q|^{1/2}} \gg S\left(\frac{\omega-\mathbf{k}\cdot\mathbf{v}}{2|Q|^{1/2}}\right), \quad (21)$$

so that distribution (10) of transient radiation for $\sin\varphi < 0$ can be written in the form

$$\frac{d^2E}{d\omega d\Omega} = \frac{e^3\omega^2}{\pi c^3}\frac{[\mathbf{n}\times\mathbf{v}]^2}{|Q|}\left\{\frac{1}{2} - \frac{\omega - \mathbf{k}\cdot\mathbf{v}}{|\pi Q|^{1/2}}\right\}. \quad (22)$$

In the range of angles where $\sin\varphi > 0$, distribution (12) of transient radiation has the form

$$\frac{d^2E}{d\omega d\Omega} = \frac{e^3\omega^2}{\pi c^3}\frac{[\mathbf{n}\times\mathbf{v}]^2}{|Q|}\left\{\frac{1}{2} + \frac{\omega - \mathbf{k}\cdot\mathbf{v}}{|\pi Q|^{1/2}}\right\}. \quad (23)$$

It should be emphasized that, for small values of angle φ, the conditions for applicability of expressions (22) and (23) are violated; consequently, a direct transition from relation (22) to (23) is ruled out.

5. AZIMUTHAL ASYMMETRY OF RADIATION DISTRIBUTION

If a charged particle escapes from a conductor at right angles to its surface, the angular distribution of radiation in zero magnetic field exhibits azimuthal symmetry. The action of a magnetic field breaks the axial symmetry of the angular distribution. Defining the direction of emission of radiation by angles ϑ and φ in a spherical system of coordinates with the axis directed along the initial velocity of the particle and considering that the radiation emitted by the ultrarelativistic particle is concentrated in the region of small angles ϑ, we can represent the argument of Fresnel integrals $C(x)$ and $S(x)$ in expressions (10) and (12) in the form

$$\frac{\omega - \mathbf{k}\cdot\mathbf{v}}{2|Q|^{1/2}} = \frac{(\omega/2)(\vartheta^2 + \gamma^{-2})}{[(\Omega\omega/2)\vartheta\sin\varphi]^{1/2}}. \quad (24)$$

The limiting case in which the value given by formula (24) tends to infinity corresponds to transient radiation in zero magnetic field. Then the angular and frequency distribution (2) is transformed into the conventional distribution for transient radiation,

$$\frac{d^2E}{d\omega d\Omega} = \frac{e^2}{\pi^2 c}\frac{[\mathbf{n}\times\mathbf{v}]^2}{\{1 - (\mathbf{n}\cdot\mathbf{v}/c)^2\}^2}. \quad (25)$$

It can be seen from formula (24) that this limiting case emerges for $\varphi = 0$ as well as for $\vartheta \longrightarrow 0$. However, the region of $\vartheta \ll 1/\gamma$ makes a small contribution to the radiation intensity and can be disregarded. Thus, radiation emitted in the plane passing through the particle velocity and the magnetic field (i.e., for $\varphi = 0$) does not depend on the magnetic field strength on the whole. However, the intensity of radiation propagating in the plane perpendicular to the magnetic field and passing through the initial velocity of the particle (i.e., for $\varphi = \pi/2$) can noticeably decrease depending on the external magnetic field strength. The ratio of the intensity of radiation emitted at angles $\varphi = 0$ and $\varphi = \pi/2$ is obviously close to unity if the argument of functions $C(x)$ and $S(x)$ is much greater than unity. Considerable deviations of this ratio from unity corresponds to small values of quantity (24), i.e., to the region in which the following inequality holds:

$$(\omega/2)(\vartheta^2 + \gamma^{-2}) \ll [(\Omega\omega/2)\vartheta\sin\varphi]^{1/2} \approx [(\omega\vartheta/\gamma)(eH/mc)\sin\varphi]^{1/2}. \quad (26)$$

For characteristic angles $\vartheta \sim 1/\gamma$ and $\varphi \sim \pi/2$, this inequality assumes the form

$$\frac{eH}{mc} \gg \frac{\omega}{\gamma^3}. \quad (27)$$

The intensity ratio for radiation emitted at angles of $\varphi = 0$ and $\pi/2$ can be obtained in the form

$$\frac{d^2E(\omega, \vartheta, \varphi = \pi/2)}{d^2E(\omega, \vartheta, \varphi = 0)} = \frac{\pi\omega^3 mc}{8ceH}\frac{\{1 + (\gamma\vartheta)^2\}^2}{\vartheta\gamma^3}. \quad (28)$$

For very small angles ϑ, inequality (27) is violated, and relation (28) becomes inapplicable. It follows from the above arguments that azimuthal asymmetry in the angular distribution of transient radiation strongly depends on the particle energy. From the presence or absence of azimuthal symmetry, one can judge whether or not the field strength is high enough for affecting transient radiation.

6. DISCUSSION

It follows hence that an external magnetic field affects the process of formation of transient radiation from an ultrarelativistic particle only for a large length of radiation formation and, hence, strongly depends on the radiation frequency and the Lorentz factor of the particle. For example, for millimeter wavelengths and for a Lorentz factor of 10^3, the coherence length is on the order of several meters over such a length. The magnetic field must noticeably change the motion of a particle for the effect in question to be manifested. In this case, for a particle escaping along the normal to the conductor–vacuum interface in an external constant and uniform magnetic field, azimuthal asymmetry in transient radiation is observed if inequality (26) is satisfied.

The magnetic field strength for which noticeable asymmetry takes place is determined by inequality (27) and strongly depends on the Lorentz factor of the particle and the radiation frequency.

Azimuthal asymmetry of transient radiation in a magnetic field can be measured quite easily, while its strong dependence on the Lorentz factor ensures sufficiently high accuracy in energy measurements. Hence,

it can be concluded that the measurement of azimuthal asymmetry in the angular distribution of transient radiation in an external field may become a convenient method for measuring the energy of ultrarelativistic particles.

It is important to emphasize that we have considered only the transient radiation formed near the surface of a conductor over a particle path with a length on the order of the coherence length. The action of the field on a particle also leads to emission of radiation on the subsequent path of the particle; however, this radiation is not connected any longer with the intersection of the interface by the particle and is conventional radiation emitted by a particle moving in a magnetic field. The properties of such radiation are determined by the specific nature of the subsequent motion of the particle in the field and its contribution to the total radiation can be different. In comparison with experiment, the contribution from such radiation should be taken into account, but it is inexpedient to include it in the general analysis. This is due to the fact that, first, such radiation depends on the velocity of the particle and not on its energy; second, the distribution of this radiation strongly depends on the condition of subsequent motion of the particle in the experimental setup.

ACKNOWLEDGMENTS

The author takes a pleasant opportunity to thank B.A. Dolgoshein for interesting remarks.

REFERENCES

1. L. D. Landau and I. Ya. Pomeranchuk, Dokl. Akad. Nauk SSSR **92**, 535 (1953); Dokl. Akad. Nauk SSSR **92**, 735 (1953).
2. A. B. Migdal, Phys. Rev. **103**, 1811 (1956).
3. A. B. Migdal, Zh. Éksp. Teor. Fiz. **32**, 633 (1957) [Sov. Phys. JETP **5**, 527 (1957)].
4. M. L. Ter-Mikaelyan, Dokl. Akad. Nauk SSSR **94**, 1033 (1954).
5. V. L. Ginzburg and I. M. Frank, Zh. Éksp. Teor. Fiz. **16**, 15 (1946).
6. V. E. Pafomov, Zh. Éksp. Teor. Fiz. **33**, 1074 (1957) [Sov. Phys. JETP **6**, 829 (1957)].
7. G. M. Garibyan, Zh. Éksp. Teor. Fiz. **33**, 1403 (1957) [Sov. Phys. JETP **6**, 1079 (1957)].
8. M. I. Ryazanov, Zh. Éksp. Teor. Fiz. **122**, 999 (2002) [JETP **95**, 861 (2002)].
9. V. E. Pafomov, Izv. Vyssh. Uchebn. Zaved., Radiofiz. **10**, 240 (1967).
10. *Handbook of Mathematical Functions*, Ed. by M. Abramowitz and I. A. Stegun, 2nd ed. (Dover, New York, 1971; Nauka, Moscow, 1979).
11. J. Boersma, Math. Comput. **14**, 380 (1960).

Translated by N. Wadhwa

Journal of Experimental and Theoretical Physics, Vol. 98, No. 3, 2004, pp. 483–488.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 125, No. 3, 2004, pp. 548–555.

ATOMS, SPECTRA, RADIATION

Anisotropic Laser-Induced Evaporation of Graphite Films

G. M. Mikheev[a], R. G. Zonov[a], A. N. Obraztsov[b], and A. P. Volkov[b]

[a]*Institute of Applied Mechanics, Ural Division, Russian Academy of Sciences, Izhevsk, 426000 Russia*
[b]*Moscow State University, Moscow, 119992 Russia*
e-mail: gmmikheev@udmnet.ru

Received April 21, 2003

Abstract—An experimental investigation of the effect of linearly polarized high-energy pulsed laser light, normally incident on a carbon thin film, is reported. The material under study consists of platelike graphite crystallites with basal crystallographic planes mostly oriented perpendicular to the substrate surface. An increase is revealed in the fraction of the graphite crystallites oriented perpendicular to the polarization plane. Laser light is found to cause significant anisotropy in diffuse scattering by the film surface. Experimental observations are explained by a model of anisotropic evaporation of graphite-like carbon material due to polarization dependence of the absorption and reflection coefficients for a rough surface. © *2004 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

Currently, much attention is given to analysis of various allotropic forms of carbon having unique physical properties, such as diamond, graphite, carbon nanotubes, and fullerenes. In particular, it was shown in our recent publications that some graphite-like films are characterized by extremely low threshold electric field strengths for electron tunneling from their surface to vacuum [1]. The field-emission properties of these films are analogous to those of carbon nanotubes (e.g., see [2–4]). One distinctive feature of graphite-like materials with low field-emission thresholds is that the basal crystallographic planes of the constituent well-ordered platelike graphite crystallites are mostly oriented perpendicular to the film surface [1, 2]. Structural properties of these graphite films can manifest themselves in other phenomena as well. In this paper, we present the results of a study of the effect of high-energy laser light on the morphology and optical properties of graphite films.

It should be noted that one well-known effect of pulsed laser light on solids is the formation of diverse periodic surface structures (e.g., see [5–12]). Normally, a laser-induced periodic surface structure is characterized by a surface height varying with a period $d \sim \lambda$, where λ is the wavelength of the laser light [6, 7]. Moreover, large-scale roughness (with a length scale varying from 10 to 300 μm) may develop on the surface of an opaque material in the zone irradiated by pulsed laser light with $\lambda \leq 1$ μm [8, 10, 12].

2. EXPERIMENTAL

We studied specimens of carbon films obtained by means of our standard technique of plasma deposition from a methane–hydrogen mixture (e.g., see [1, 13]). The films were deposited on 25×25 mm^2 silicon substrates. A single-mode YAG:Nd^{3+} laser was used as a high-power light source. The laser cavity was designed to produce linearly polarized light at 1064 nm with pulse energy of up to 50 mJ [14]. The half-width of Gaussian laser pulses was about 22 ns, and the beam diameter was 2 mm. The laser beam was directed through a converging lens with a focal length of 10 cm onto the film surface under study. The irradiance on the carbon-film surface was varied gradually by varying the lens-to-specimen distance between 12 and 35 cm. The specimens were irradiated by one or several laser light pulses in air.

The film structure and morphology before and after irradiation was analyzed by means of a Neophot 32 optical microscope with a resolution of at least 0.4 μm, a Solver P47 atomic-force microscope (AFM), and a LEO 1550 scanning electron microscope (SEM). The angular dependence of the efficiency of diffuse scattering by the examined carbon films was analyzed by means of an apparatus based on an LEF-3M ellipsometer (see Fig. 1), with a He–Ne laser used as a 1 mW circularly polarized light source. Diffuse scattered light *b* was observed at a constant angle φ in the plane *yz* perpendicular to the plane *xz* of incidence of beam *a*. The angle of incidence of beam *a* in the *xz* plane was equal to the angle of reflection of beam *b* in the *yz* plane. We studied the intensity I_R of reflected light on the specimen's orientation characterized by the angle α of rotation in the *xy* plane at various constant angles φ. In the starting position of a specimen, the sides of the silicon substrate were aligned with the *x* and *y* axes.

3. RESULTS

In its original surface morphology, a graphite-like film consists of platelike graphite crystallites several tens of nanometers thick, with other dimensions vary-

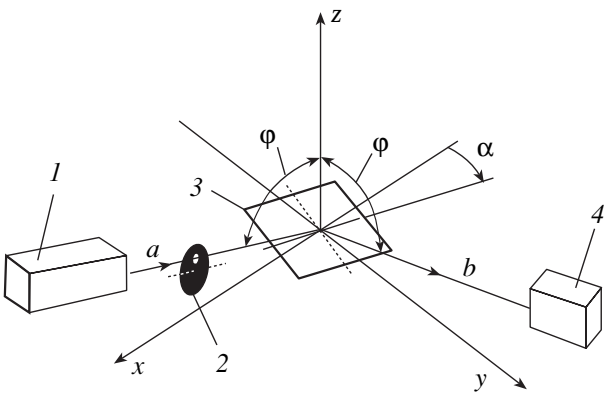

Fig. 1. Optical arrangement for studying diffuse light scattering by carbon film surface: (*1*) He–Ne laser; (*2*) beam chopper; (*3*) carbon film on silicon substrate; (*4*) photomultiplier; *xyz* is a Cartesian coordinate system.

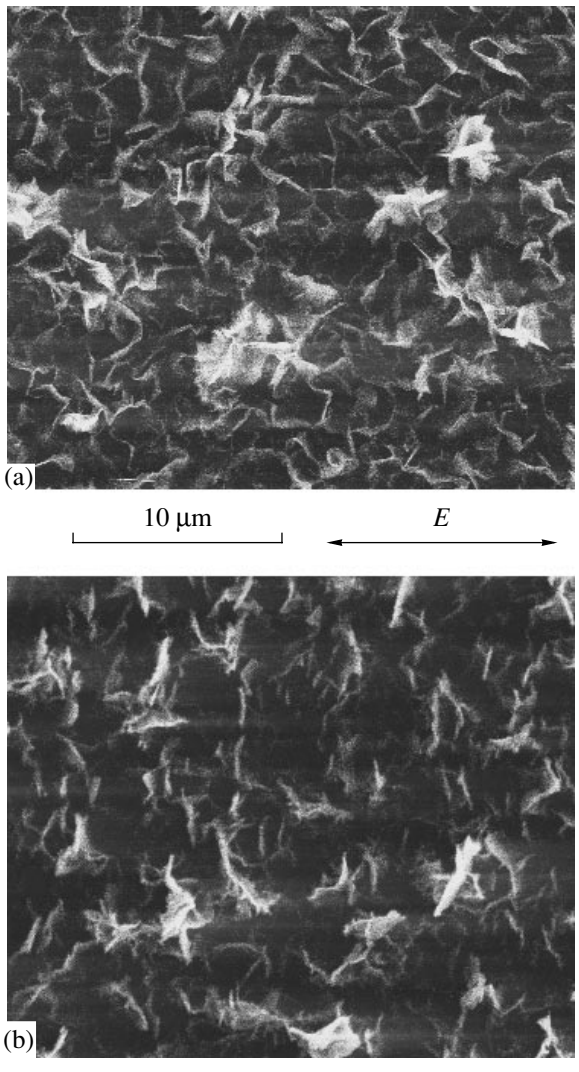

Fig. 2. Electron-microscopic images of carbon surface (a) before and (b) after irradiation by laser light. Arrows indicate the direction of the laser electric field vector *E*.

ing from one to several microns (see Fig. 2a). The predominant orientation of the platelike crystallite planes is parallel to the basal crystallographic plane of graphite (0001) and perpendicular to the substrate. There is no particular orientation in other directions. This orientation of the crystallites explains the changes in diffuse scattering efficiency that can be observed with the naked eye as the angle of incidence or observation is varied. Note that the diffuse reflection efficiency is invariant with respect to rotation of the specimen about an axis perpendicular to its surface. These observations are consistent with previously reported results [1, 2, 13, 15].

Irradiation with high-energy laser light substantially changes the film-surface morphology, as illustrated by images obtained by means of SEM (see Fig. 2b) and AFM (Fig. 3). After irradiation with five to ten pulses (depending on the beam irradiance, which varied between 10 and 30 MW/cm^2), the surface had numerous elements stretched along a certain direction. Apparently, these were platelike graphite crystallites that had existed on the film surface and were not strongly affected by laser light. Comparative electron microscopy of the original and laser-irradiated carbon surfaces (see Figs. 2a and 2b, respectively) leads to the following conclusions.

I. Irradiation by laser light results in a partial disordering of the graphite-like material under study, which manifests itself in changes in its surface morphology and in the efficiency of secondary electron emission. The latter effect is inferred from the observed change in SEM image contrast. Additional evidence of the disordering has been obtained in Raman-scattering studies.

II. There is no indication of any periodic structure on the surface after repeated irradiation by laser pulses [5–7, 11]. The absence of periodic surface structure was also noted in [7], where the effect of irradiation with 40 ns XeCl excimer laser pulses on rough diamond film surfaces was investigated.

III. An increase is revealed in the fraction of the graphite crystallites perpendicular to the polarization plane. This conclusion about predominant orientation of crystallites on the irradiated surface segment is corroborated by atomic-force microscopy of surface morphology (Fig. 3).

This last effect is most clearly manifested in the AFM image of the film surface shown in Fig. 3. Moreover, the figure demonstrates that the width of an individual crystallite at a height of 0.6 μm can be as large as 2 μm in the *A–A* cross section (see Fig. 3b), whereas its width in the perpendicular *B–B* cross section is 0.6 μm (Fig. 3c). Analogous characteristics are exhibited by other morphological elements of the film surface irradiated by laser light. All of them are parallel to the *A–A* cross section. Note that the surface irradiated by five to ten pulses of laser light changes color from the original metallic gray to velvet black. Further increase in the number *N* of irradiating pulses leads to decrease in the height of surface elements. When *N* is

sufficiently large, the film-surface color changes from black back to gray and no morphological elements with predominant orientation are observed in AFM images.

An analysis of experimental results shows that the orientation of morphological elements in the SEM and AFM images of the graphite film surface is perpendicular to the polarization plane of the incident laser light. An angular displacement of the irradiated specimen about the axis of the laser beam (normal to the film plane) results in a corresponding rotation of the predominant orientation. When the polarization plane of laser light is changed by means of two quarter-wave plates without rotating the specimen, the structures created by laser irradiation on the carbon film surface change correspondingly.

The predominant orientation of crystallites along the direction normal to the surface is also confirmed by the aforementioned visual observations of diffuse scattering anisotropy with respect to the angle of incidence in the plane perpendicular to the film surface [2]. Moreover, we found that the efficiency of diffuse scattering by an irradiated surface depends on the angle of rotation of the specimen about an axis perpendicular to the substrate. This effect was analyzed in detail by measuring the intensity of diffuse scattered light at various angles of rotation of the specimen.

Figure 4 shows the intensity I_R of diffuse scattered He–Ne laser light measured as a function of the angle of specimen rotation α. Whereas no orientational dependence of I_R on α is observed (see Fig. 4a), the dependence $I_R(\alpha)$ obtained after irradiation by high-energy laser light with polarization characterized by an angle γ_1 exhibits two pronounced peaks over the interval of complete revolution about the axis (Fig. 4b). For an angle of rotation α measured relative to an arbitrary direction, the peaks in $I_R(\alpha)$ correspond to $\alpha_{11} \approx 83°$ and $\alpha_{12} \approx 264°$; that is, $\alpha_{12} - \alpha_{11} \approx 180°$. By scrutinizing the curve $I_R(\alpha)$ obtained for the film irradiated by laser light with a polarization angle γ_2 differing from γ_1 by $\Delta\gamma$, it was found that the corresponding peaks in $I_R(\alpha)$ were observed at angles differing by $\Delta\gamma$. As an example, Fig. 4c shows the dependence $I_R(\alpha)$ measured for $\Delta\gamma = 45°$, in which case $\alpha_{21} \approx 129°$ and $\alpha_{22} \approx 309°$. The following relations were found to hold within the measurement error: $\alpha_{21} - \alpha_{11} \approx \alpha_{22} - \alpha_{12} \approx \Delta\gamma$.

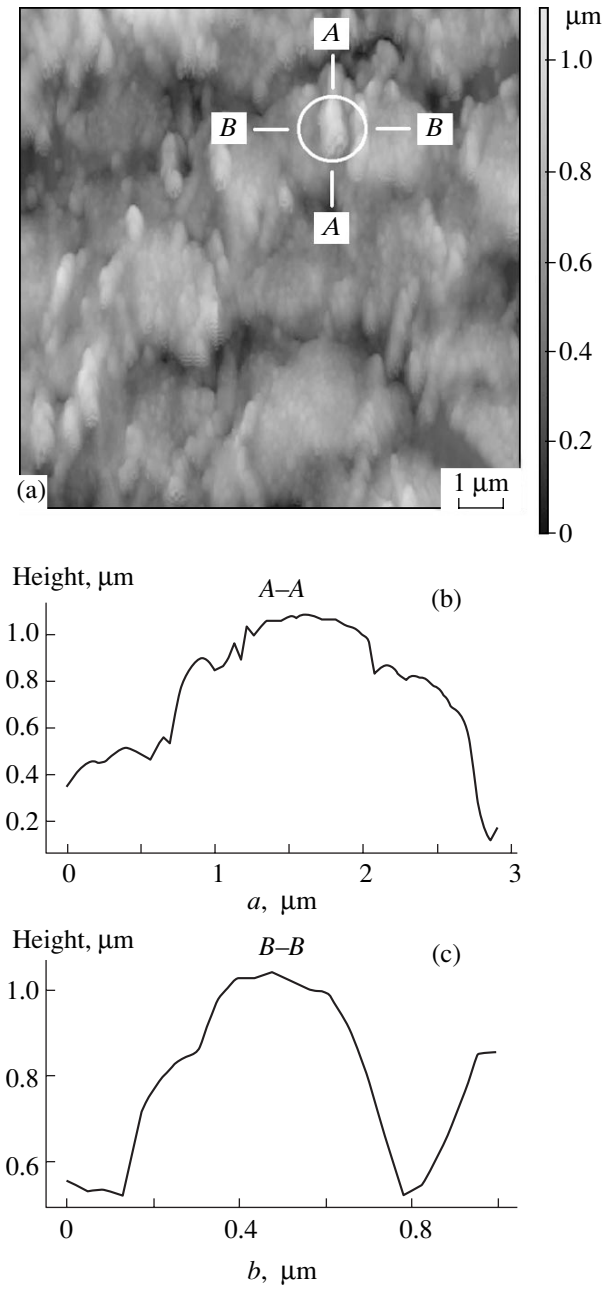

Fig. 3. AFM image of carbon film after laser irradiation (a) and the mutually perpendicular cross-sectional profiles of a structural element on the film surface in planes *A–A* (b) and *B–B* (c).

4. DISCUSSION

According to [6], the formation of periodic surface patterns is caused by diffraction of a spatially coherent incident light wave by a rough surface and interference of diffracted waves with the wave that penetrates the medium and creates a periodic temperature field on the surface. Large-scale periodic pattern formation is associated with instability development at the interface between a melt and an optical-breakdown plasma, melt displacement caused by the vaporizing target material, initial capillary-wave pattern formation, and other effects [12]. As noted above, our experiments give no indication of pattern formation on the irradiated carbon film surface. However, we did observe stretched elements mostly oriented perpendicular to the polarization plane of laser light. It should be noted that these parallel elements are chaotically distributed over the film surface, but their presence is clearly manifested in the angular dependence of diffuse scattering. This can be explained by anisotropic evaporation of platelike graphite crystallites from the original film surface.

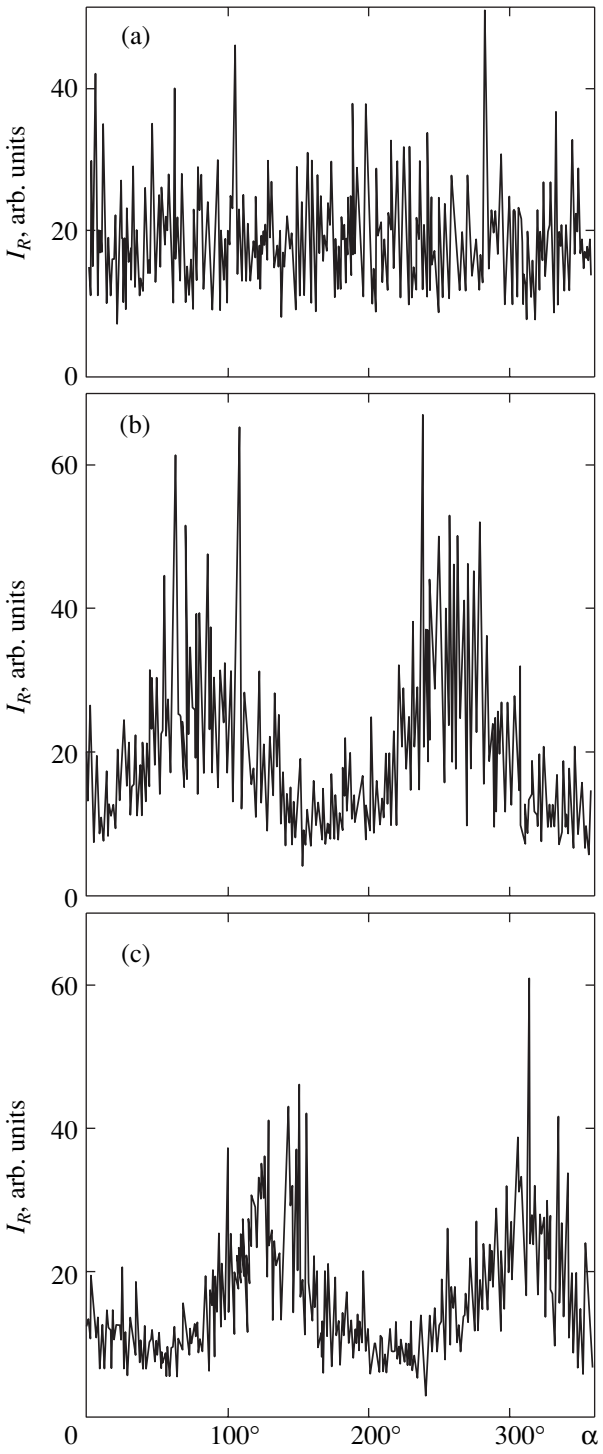

Fig. 4. Intensity I_R of scattered He–Ne laser light versus the angle of rotation α of a specimen, measured at φ = 75° (a) before laser irradiation and after irradiation by pulsed laser light polarized at angles (b) γ_1 and (c) γ_2 differing by 45°.

To elucidate this phenomenon, let us consider the effect of laser light on a rough film surface consisting of hemispherical elements. Suppose that a laser beam is normally incident on the film surface. For the laser

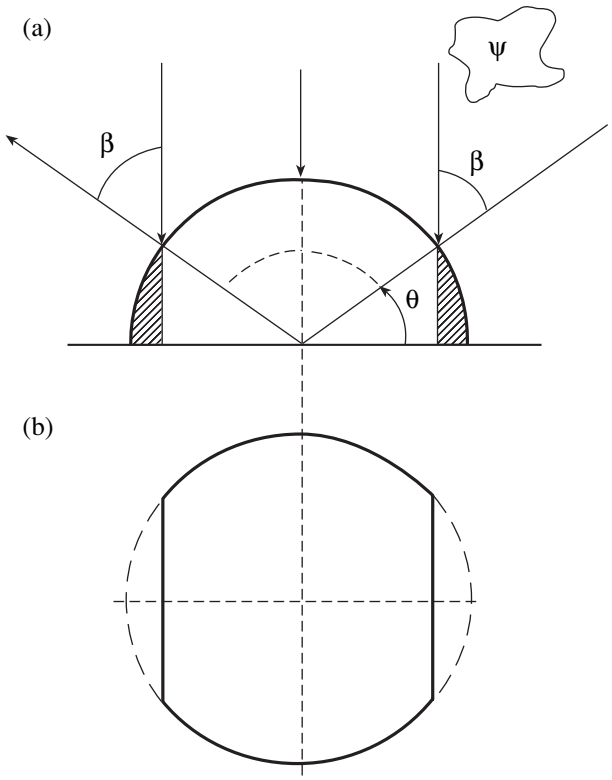

Fig. 5. Schematic illustration of interaction between a laser beam normally incident on a surface and a hemispherical element on the surface: (a) diametrical cross-sectional view of the hemisphere; (b) top view of the hemispherical element after irradiation by a laser beam polarized in the plane ψ.

polarization plane ψ passing through an arbitrary diameter of a hemisphere, the angle of incidence β is determined by the azimuthal angle θ: $\beta = |90° - \theta|$ (see Fig. 5).

The reflection efficiency and absorption coefficient for absorbing media depend on the angle of incidence and polarization of the beam. To evaluate the reflection coefficients R_p and R_s, which correspond, respectively, to polarization in the plane of incidence (*p*-polarization) and in the plane perpendicular to the plane of incidence (*s*-polarization), we can use the following exact formulas for isotropic absorbing media [16, 17]:

$$R_s = \frac{a^2 + b^2 - 2a\cos\beta + \cos^2\beta}{a^2 + b^2 + 2a\cos\beta + \cos^2\beta},$$

$$R_p = R_s\frac{a^2 + b^2 - 2a\sin\beta\tan\beta + \sin^2\beta\tan^2\beta}{a^2 + b^2 + 2a\sin\beta\tan\beta + \sin^2\beta\tan^2\beta},$$

$$\frac{R_p}{R_s} = \frac{a^2 + b^2 - 2a\sin\beta\tan\beta + \sin^2\beta\tan^2\beta}{a^2 + b^2 + 2a\sin\beta\tan\beta + \sin^2\beta\tan^2\beta},$$

where

$$a = \left\{\frac{1}{2}[((n^2-\chi^2-\sin^2\beta)^2+4n^2\chi^2)^{1/2} + (n^2-\chi^2-\sin^2\beta)]\right\}^{1/2},$$

$$b = \left\{\frac{1}{2}[((n^2-\chi^2-\sin^2\beta)^2+4n^2\chi^2)^{1/2} - (n^2-\chi^2-\sin^2\beta)]\right\}^{1/2},$$

with n and χ denoting the real and imaginary parts of the complex refractive index of the absorbing medium. Borrowing the real and imaginary parts of the permittivity of graphite at 1 μm from [18], we obtain $n^2 = 1.3$ and $\chi^2 = 4.9$ for graphite at the neodymium laser wavelength. Then, we use the formulas written out above to plot the absorption coefficients $A_p = 1 - R_p$ and $A_s = 1 - R_s$ (see Fig. 6a). Figure 6b shows the dependence of A_p/A_s on β plotted in a similar manner. According to Fig. 6, $A_p = A_s = 0.48$ for $\beta = 0$ and $A_p = A_s = 0$ for $\beta = 90°$. As β varies from 0 to 90°, the absorption coefficient A_s monotonically decreases. The dependence of A_p on β exhibits different behavior. At $\beta_{max} = 66°$, the absorption coefficient A_p reaches its maximum value 0.67, whereas $A_s = 0.23$, i.e., $A_p/A_s = 2.9$. At higher values of β, the ratio A_p/A_s monotonically increases.

Note that laser-induced fracture and evaporation of the material under study are characterized by threshold conditions depending on the absorption coefficient. Accordingly, evaporation (or fracture) of a graphite-like material induced by laser light with a certain irradiance at points on a spherical surface where $\theta \le \beta_{max}$ and $\theta \ge \pi - \beta_{max}$ (see Fig. 5a) is much more efficient in the case of *p*-polarization as compared to *s*-polarization. Even though the shapes of graphite crystallites in the examined films substantially differ from a hemisphere, the analysis developed above explains the mechanism of their selective ablation with respect to the laser polarization direction. In the case of linearly polarized laser light, both heating and subsequent fracture of a hemispherical element are anisotropic. On the lateral parts of such elements, spherical half-segments of the material oriented perpendicular to the polarization plane evaporate in the first place. As a result, the original hemispherical shape of the element change, and its top view becomes a curvilinear trapezoid whose parallel sides are perpendicular to the polarization plane of the laser beam (see Fig. 5b).

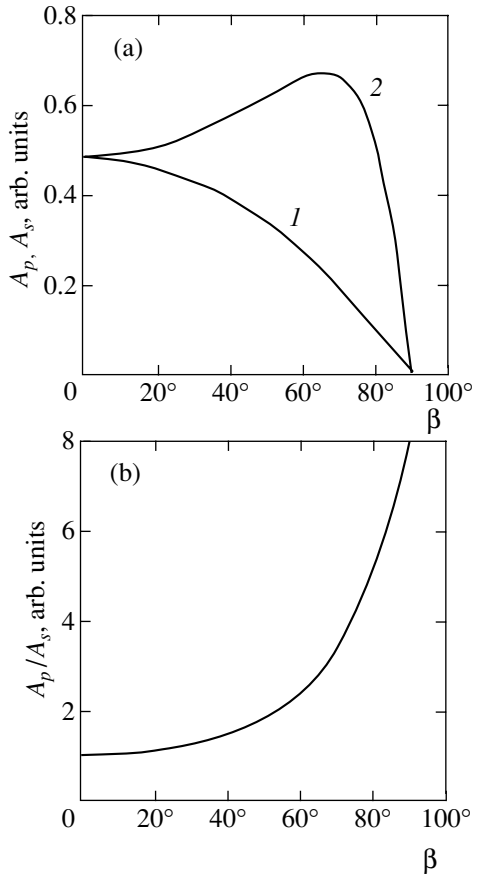

Fig. 6. Calculated dependences on the angle of incidence β for graphite at a wavelength of 1 μm: (a) absorption coefficients A_s (curve *1*) and A_p (curve *2*); (b) A_p/A_s.

Thus, we can explain how irradiation by high-energy laser light can cause a certain preferred direction to appear on an initially isotropic rough film surface. The morphology of a real carbon film surface is inhomogeneous and is substantially different from the model considered above. However, the anisotropic evaporation mechanism outlined above can obviously manifest itself in a similar manner in this case as well.

In experimental studies of film structure by means of a He–Ne laser, the angle between the incident beam and the direction of observation of the scattered light was 90° (see Fig. 1). Therefore, the dependence presented in Fig. 4a can be explained by diffuse scattering by a fractal rough surface [19]. It is clear that the patterns developing on an irradiated surface substantially change the diffuse scattering function, as illustrated by the experimental results shown in Figs. 4b and 4c.

5. CONCLUSIONS

It is demonstrated that irradiation of graphite-like films by linearly polarized high-energy pulsed laser light leads to the development of spatially oriented structures. The structure orientation determined by the

laser polarization plane is explained by anisotropic evaporation of the carbon film.

ACKNOWLEDGMENTS

We are sincerely grateful to A.E. Murav'ev, A.Yu. Popov, and K. Efimovs for technical assistance provided in the course of this study. This work was supported by the International Association for Cooperation with Scientists from the former Soviet Union (INTAS), grant no. 01-0254.

REFERENCES

1. A. N. Obraztsov, A. P. Volkov, A. I. Boronin, and S. V. Koshcheev, Zh. Éksp. Teor. Fiz. **120**, 970 (2001) [JETP **93**, 846 (2001)].
2. A. N. Obraztsov, A. P. Volkov, I. Yu. Pavlovskiĭ, *et al.*, Pis'ma Zh. Éksp. Teor. Fiz. **69**, 381 (1999) [JETP Lett. **69**, 411 (1999)].
3. J.-M. Bonard, H. Kind, Th. Stöckli, and L.-O. Nilson, Solid-State Electron. **45**, 893 (2001).
4. A. V. Eletskiĭ, Usp. Fiz. Nauk **172**, 401 (2002) [Phys. Usp. **45**, 369 (2002)].
5. M. J. Birnbaum, Appl. Phys. **36**, 3688 (1965).
6. S. A. Akhmanov, V. I. Emel'yanov, N. I. Koroteev, and V. N. Seminogov, Usp. Fiz. Nauk **147**, 675 (1985) [Sov. Phys.–Usp. **28**, 1084 (1985)].
7. V. P. Ageev, L. L. Buĭlov, V. I. Konov, *et al.*, Dokl. Akad. Nauk SSSR **303**, 598 (1988) [Sov. Phys. Dokl. **33**, 840 (1988)].
8. V. P. Ageev, L. L. Buĭlov, and V. I. Konov, Kvantovaya Élektron. (Moscow) **16**, 1214 (1989).
9. D. O. Barsukov, G. M. Gusakov, and A. I. Frolov, Kvantovaya Élektron. (Moscow) **18**, 1477 (1991).
10. A. B. Braĭlovskiĭ, I. A. Dorofeev, A. B. Ezerskiĭ, *et al.*, Zh. Tekh. Fiz. **61**, 129 (1991) [Sov. Phys. Tech. Phys. **36**, 324 (1991)].
11. A. F. Banishev, V. S. Golubev, and O. D. Khramova, Laser Phys. **3**, 1198 (1993).
12. V. V. Voronov, S. I. Dolgaev, S. V. Lavrishchev, *et al.*, Kvantovaya Élektron. (Moscow) **30**, 710 (2000).
13. A. N. Obraztsov, A. A. Zolotukhin, A. O. Ustinov, *et al.*, Carbon **41**, 836 (2003).
14. G. M. Mikheev, D. I. Maleev, and T. N. Mogileva, Kvantovaya Élektron. (Moscow) **19**, 45 (1992) [Sov. J. Quantum Electron. **22**, 37 (1992)].
15. A. N. Obraztsov, I. Yu. Pavlovsky, A. P. Volkov, *et al.*, Diamond Relat. Mater. **8**, 814 (1999).
16. A. P. Prishivalko, *Light Reflection from Absorbing Media* (Akad. Nauk BSSR, Minsk, 1963), p. 26.
17. G. M. Mikheev and V. S. Idiatulin, Kvantovaya Élektron. (Moscow) **24**, 1007 (1997) [Quantum Electronics **27** (11), 978 (1997)].
18. *Optical Properties of Semiconductors: Handbook*, Ed. by V. I. Gavrilenko, A. M. Grekhov, D. V. Korbutyak, and V. G. Litovchenko (Naukova Dumka, Kiev, 1987), p. 198.
19. A. A. Potapov, *Fractals in Radiophysics and Radiolocation* (Logos, Moscow, 2002), p. 194.

Translated by A. Betev

Journal of Experimental and Theoretical Physics, Vol. 98, No. 3, 2004, pp. 489–495.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 125, No. 3, 2004, pp. 556–562.

ATOMS, SPECTRA, RADIATION

Two-Scale Angular-Momentum Evolution Induced by Elliptically Polarized Resonance Radiation for a Two-Level Atom with $J_g \longrightarrow J_e = J_g + 1$ Optical Transition in a Constant Magnetic Field

K. A. Nasyrov

Institute of Automatics and Electrometry, Siberian Division, Russian Academy of Sciences, Novosibirsk, 630090 Russia

e-mail: nasyrov@iae.nsk.su

Received August 7, 2003

Abstract—The dynamics of a two-level atom with optical transition $J_g \longrightarrow J_e = J_g + 1$ under the action of elliptically polarized resonance radiation in the presence of a constant field directed along the radiation wave vector is studied in the approximation of a semiclassical description of the angular momentum orientation. It is shown that the atomic distribution over the angular momentum orientation exhibits a two-scale time evolution. At the first stage, after the beginning of irradiation, the angular momenta of atoms get oriented (over comparatively short time intervals) along the magnetic field, as well as in the opposite direction, depending on their initial orientation and the ellipticity of radiation. At the second (longer) stage, the redistribution of atoms takes place, as a result of which they are oriented predominantly in one of the above directions. The duration of the second stage is an exponential function of angular momentum *J*. © *2004 MAIK "Nauka/Interperiodica".*

The scope of problems that involve the interaction of atoms with resonance radiation in the presence of a magnetic field has been expanded in recent years. We are speaking primarily of magnetooptical traps for neutral atoms [1], magnetometers based on optical pumping [2, 3], and so on. A characteristic feature of such problems is that the degeneracy of atomic states in the angular momentum projection must be included in view of the vector nature of the interaction of the electromagnetic field with the atoms. Even if we confine our analysis to the model of an atom with two energy states, the problem becomes essentially a multilevel problem in view of the above-mentioned degeneracy and an analytic solution can be obtained only in a few specific cases [4–6]. The dynamics of an atom can be traced only via numerical solution of the corresponding equations, which can provide answers to specific questions, although it is difficult to analyze the dependences on the parameters of the problem in this case.

However, semiclassical approaches to the description of rotational motion and, in particular, the angular momentum orientation can be useful here [7]. In these approximations, the equations for the density matrix are simplified to such an extent that analytic solutions can be obtained and their parametric analysis can be carried out. A necessary condition for the validity of a semiclassical description of rotational motion is a large angular momentum of the quantum system, $J \gg 1$. However, essentially quantum effects, which are formally manifested in the presence of narrow angular structures with a size smaller than the quantum indeterminacy $\Delta\theta \sim 1/\sqrt{J}$ in the angular momentum orientation in the density matrix, can also be exhibited in such systems [7, 8]. In the present study, we consider only such solutions which will be referred to as semiclassical and in which the characteristic scale of angular strictures exceeds the quantum-mechanical indeterminacy in the angular momentum orientation. Nevertheless, the results obtained when the condition $J \gg 1$ holds remain qualitatively valid for not very large values of $J \sim 3$–5 as well. This remark is especially important since atoms are mainly characterized by not very large values of *J*. For example, the alkali atom Cs exhibits a closed transition $F \longrightarrow F + 1$ from the hyperfine state with a total angular momentum of $F = 4$. Only some atoms (e.g., Fr isotopes) possess an angular momentum sufficient for assuming that the semiclassical approximation is quite applicable for their description.

It was shown in [8], where the classical approximation was used for describing rotational motion, that an atom with optical transition $J_g \longrightarrow J_e = J_g + 1$ in the field of elliptically polarized resonance radiation and in the presence of a magnetic field directed along the wave vector orients its angular momentum either parallel or antiparallel to the magnetic field, depending on its initial orientation and the ellipticity of the resonance field

1063-7761/04/9803-0489$26.00

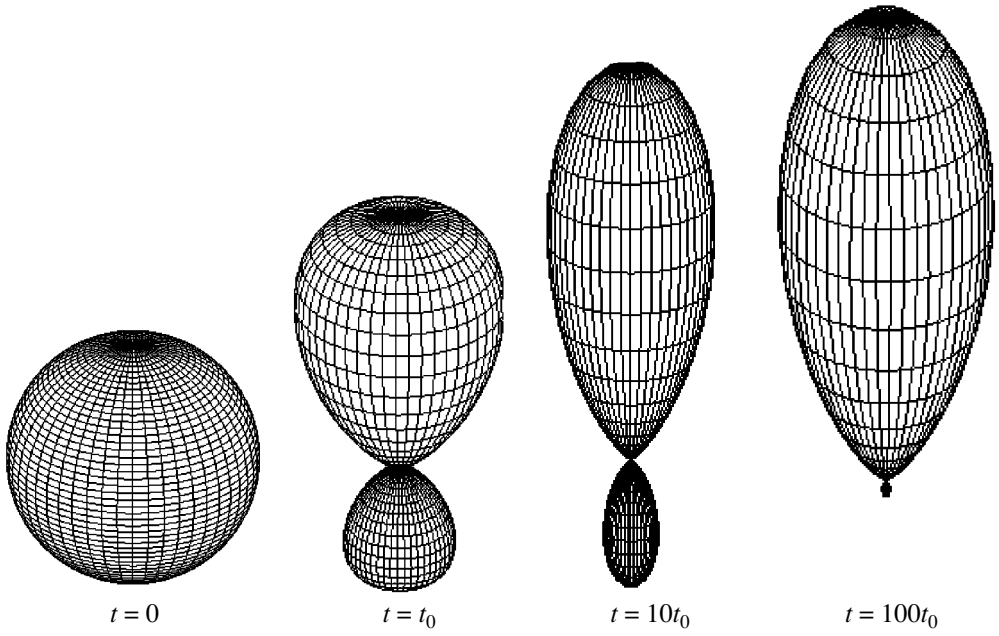

Fig. 1. Evolution of the angular momentum orientation distribution for an atom with $J_g = 4$, $J_e = 5$; the ellipticity of radiation is $x_0 = -0.3$ and $t_0 = J\Gamma/G^2$. The z axis and the magnetic field are directed upwards.

polarization. However, such dynamics of the angular momentum is observed only for a certain time interval,

$$t_0 = J\Gamma/G^2,$$

where J is the angular momentum, G is the Rabi frequency, and Γ is the line half-width. Numerical calculations based on the optical Bloch equations in the *JM* representation for a two-level degenerate system show that, in the case of the elliptic polarization of radiation, one of the directions of the angular momentum is unstable. For long time intervals $t \gg t_0$, the overwhelming majority of particles are concentrated only in the vicinity of one of the possible directions of the angular momentum (Fig. 1). The results of calculations obtained in the *JM* representation are visualized with the help of the method described in [7]. The figure shows that, in the case of population of the lower state with an initially isotropic distribution over the direction of the angular momentum, an anisotropic distribution is formed at the first stage over time intervals $t \sim t_0$ (for the chosen ellipticity of polarization, the angular momenta of part of the atoms are oriented along the magnetic field, while the angular momenta of a slightly smaller part of atoms are oriented in the opposite direction). For longer periods of time, the latter part becomes smaller and smaller until the angular momenta of nearly all particles are oriented along the magnetic field (it should be stipulated that, in view of quantum indeterminacy $\Delta\theta \sim 1/\sqrt{J}$ in the angle of orientation of angular momentum, a certain distribution of angular momentum directions about the magnetic field direction always exists). It is worth noting that the second stage is much longer than the first.

The present study aims to describe this phenomenon on the basis of a semiclassical representation of angular momentum [9]. In particular, the factors determining the duration of the second stage in the redistribution of particles over the direction of angular momentum will be clarified.

We consider a quantum system with two energy levels degenerate in the angular momentum projection, which resonantly interacts with radiation so that radiation induces transitions from the lower (ground) state g to the upper (excited) state e and back. Passing to a semiclassical description of rotational motion, we will use the $\phi\theta\alpha$ representation introduced in [9] for a rotator or a spherical top. Here, angles θ and ϕ characterize the polar and azimuth angles of the direction of the angular momentum, while angle α determines the orientation of the rotator axis in a plane orthogonal to the angular momentum direction. In exact resonance with the exciting field, the equations for the density matrix for the quantum system considered here in the $\phi\theta\alpha$ representation have the form [9]

$$\left(\frac{d}{dt} + \Gamma_e\right)\rho_{ee} = -i\exp\left(\frac{i\hat{w}}{2J}\right)G\rho_{ge}$$

$$+ i\exp\left(\frac{-i\hat{w}}{2J}\right)G^*\rho_{eg},$$

$$\frac{d}{dt}\rho_{gg} + \omega_H \frac{d}{d\phi}\rho_{gg} = i\exp\left(\frac{-i\hat{w}}{2J}\right)G\rho_{ge}$$
$$- i\exp\left(\frac{i\hat{w}}{2J}\right)G^*\rho_{eg} + \hat{\Gamma}_e\rho_{ee}, \quad (1)$$
$$\left(\frac{d}{dt} + \Gamma\right)\rho_{eg} = -i\exp\left(\frac{i\hat{w}}{2J}\right)G\rho_{gg}$$
$$+ i\exp\left(\frac{-i\hat{w}}{2J}\right)G\rho_{ee},$$
$$\rho_{ge} = \rho_{eg}^*, \quad 2J = J_e + J_g + 1.$$

Here, Γ_e is the constant of spontaneous decay of the excited state to the ground state. We assume that there exists only one decay channel for state *e*. Constant Γ is the rate of coherence breaking between levels *e* and *g*. In the absence of other mechanisms of destruction of this coherence besides spontaneous breaking, we have $\Gamma = \Gamma_e/2$.

We consider the case of a weak magnetic field, when the Zeeman frequency ω_H is much smaller than the relaxation constant Γ; consequently, we disregard the magnetic field effects in Eqs. (1) for nondiagonal elements of the density matrix and for the excited state. The effect of rotation of the angular momentum of the atom in the magnetic field is included only in the equation for the diagonal element of the density matrix for the ground state. Equations (1) are written in a coordinate system with the *z* axis directed along the magnetic field.

Term $\hat{\Gamma}_e\rho_{ee}$ on the right-hand side of the equation for the diagonal element ρ_{gg} of the density matrix describes the arrival at the ground state due to spontaneous decay of excited state *e*. The specific form of this term for various types of optical transitions was determined in [8]. In the particular case of transition $J_g \longrightarrow J_e = J_g + 1$, this term has the form

$$\frac{\hat{\Gamma}_e - \Gamma_e}{\Gamma_e} = \frac{J_e - J_g}{J}\left(1 + \frac{J_e - J_g}{2J}\right) + \frac{1}{4J^2}\Delta, \quad (2)$$

where Δ is the angular Laplacian in angular variables θ and ϕ.

In Eqs. (1), the following notation is used:

$$G(\phi, \theta, \alpha) = \sum_\sigma G_\sigma D^1_{\sigma, J_g - J_e}(\phi, \theta, \alpha)$$
$$= \tilde{G}(\phi, \theta)\exp(i(J_e - J_g)\alpha),$$
$$\tilde{G}(\phi, \theta) = \sum_\sigma G_\sigma D^1_{\sigma, J_g - J_e}(\phi, \theta, 0), \quad G_\sigma = \frac{E_\sigma d_{eg}}{\hbar\sqrt{2J+1}}.$$

Here, E_σ are the circular components of the polarization vector of the electric field of the wave ($\sigma = 0, \pm 1$); d_{eg} is the reduced matrix element of the dipole moment for the *e*–*g* transition; and $D^1_{\sigma, J_g - J_e}$ is the Wigner rotation matrix. We will consider the case when the direction of propagation of radiation coincides with the direction of the magnetic field. Operator $\hat{w}$ acts in accordance with the rule [9]

$$\hat{w}PQ = \left(\frac{\partial}{\partial\phi} - \cos\theta\frac{\partial}{\partial\alpha}\right)P\frac{\partial}{\partial\cos\theta}Q$$
$$- \frac{\partial}{\partial\cos\theta}P\left(\frac{\partial}{\partial\phi} - \cos\theta\frac{\partial}{\partial\alpha}\right)Q. \quad (3)$$

We consider the situation when the radiation intensity is not high; i.e., the Rabi frequency $G \ll \Gamma$. In this case, the population of the excited state is always much smaller than that of the ground state and we can disregard ρ_{ee} as compared to ρ_{gg} as well as the time derivatives in the equations for ρ_{ge} and ρ_{eg}. Summing the equations for the diagonal elements of the density matrix in this approximation for the radiation intensity, we arrive at the equation

$$\frac{\partial}{\partial t}\rho_{gg} + \omega_H\frac{\partial}{\partial\phi}\rho_{gg}$$
$$= 2\sin\left(\frac{\hat{w}}{2J}\right)(G\rho_{ge} + G^*\rho_{eg}) + (\hat{\Gamma}_e - \Gamma_e)\rho_{ee}. \quad (4)$$

In the semiclassical approximation, when it is assumed that the angular scales of characteristic variations of the density matrix considerably exceed the quantum indeterminacy $1/\sqrt{J}$ in the angular momentum orientation, we can retain only a few first terms in the expansion of operators of the type $\exp(i\hat{w}/2J)$ into a Taylor series. In particular, for our subsequent analysis, it is sufficient to retain in Eq. (4) for ρ_{gg} only the terms on the order up to $1/J^2$; consequently, in the expansion of the sine in Eq. (4), we retain only the first nonzero term. In addition, ρ_{ee} and nondiagonal elements of the density matrix can be calculated to within $1/J$:

$$\rho_{ee} = -\frac{i}{\Gamma_e}(G\rho_{ge} - G^*\rho_{eg}) + \frac{\hat{w}}{2J\Gamma_e}(G\rho_{ge} + G^*\rho_{eg}),$$
$$\rho_{eg} = -\frac{i}{\Gamma}G\rho_{gg} + \frac{\hat{w}}{2J\Gamma}G\rho_{gg}, \quad (5)$$
$$\rho_{ge} = \frac{i}{\Gamma}G^*\rho_{gg} + \frac{\hat{w}}{2J\Gamma}G^*\rho_{gg}.$$

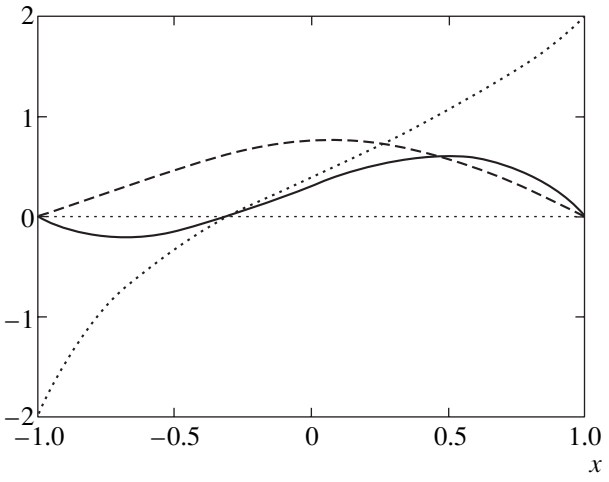

Fig. 2. Normalized drift velocity $uJ\Gamma/\bar{G}^2$ (solid curve), normalized diffusion coefficient $DJ^2\Gamma/\bar{G}^2$ (dashed curve), and the ratio u/DJ of the drift velocity to the diffusion coefficient (dotted curve). The ellipticity of polarization is $x_0 = -0.3$.

We assume further that the Zeeman frequency ω_H is much higher than the velocity $\bar{G}^2/J\Gamma$ of rotation of the angular momentum [8] (here, $\bar{G}^2 = |G_{+1}|^2 + |G_{-1}|^2$). This allows us to assume that population ρ_{gg} has a uniform distribution over angle ϕ. In this approximation, we will study the angular momentum distribution over angle θ,

$$\rho = \frac{1}{2\pi}\int_0^{2\pi}\rho_{gg}d\phi. \tag{6}$$

Proceeding in the same way as described in [8] and retaining only the terns on the order of $1/J$ and $1/J^2$, after certain calculations, we obtain

$$\frac{\partial}{\partial t}\rho + \frac{\partial}{\partial\cos\theta}\left(u\rho - D\frac{\partial}{\partial\cos\theta}\rho\right) = 0, \tag{7}$$

$$u = \frac{\sin^2\theta}{2\Gamma J}[|G_{-1}|^2(1+\cos\theta) - |G_{+1}|^2(1-\cos\theta)],$$

$$D = \frac{\sin^2\theta}{8\Gamma J^2}[|G_{-1}|^2(1+\cos\theta)(3-\cos\theta) + |G_{+1}|^2(1-\cos\theta)(3+\cos\theta)].$$

It can be seen that the dynamics of the angular momentum distribution over its projection on the z axis obeys a diffusion-drift equation, the terms proportional to $1/J$ determining the drift velocity u of the motion of angular momentum in coordinate $\cos\theta$, while the terms proportional to $1/J^2$ are responsible for diffusion.

Thus, if we consider the case when $J \gg 1$, which is necessary for the semiclassical description of the angular momentum to be valid, the diffusion term is of the next order of smallness in parameter $1/J$ as compared to the drift term. This means that the angular momentum dynamics at the initial instant following the application of laser radiation is mainly determined by the drift; at later stages, diffusion begins to play a significant role.

To simplify the subsequent analysis, we introduce the notation $x = \cos\theta$. In accordance with Eqs. (7), velocity u is alternating in the range of admissible values of x ($-1 \le x \le 1$) and assumes zero value for

$$x_0 = \frac{|G_{+1}|^2 - |G_{-1}|^2}{|G_{-1}|^2 + |G_{+1}|^2}. \tag{8}$$

Quantity x_0 also characterizes the ellipticity of the radiation polarization (in particular, $x_0 = 0$ corresponds to the linear polarization and $x_0 = \pm 1$ to the right and left circular polarization). It follows from Eqs. (7) that $u > 0$ for $x > x_0$ and, conversely, $u < 0$ for $x < x_0$ (see Fig. 2). In other words, all particles with initial coordinate $x > x_0$ move to the right and would be accumulated at point $x = 1$ in the absence of diffusion, while particles with initial coordinate $x < x_0$ would gather at point $x = -1$. If we assume that the initial distribution of particles over x is uniform, the number of particles gathered at points $x = 1$ and $x = -1$ would be

$$N^+ = \frac{1-x_0}{2}, \quad N^- = \frac{1+x_0}{2}, \tag{9}$$

respectively. We will use the following normalization for the number of particles:

$$N^+ + N^- = 1. \tag{10}$$

However, the steady-state solution to Eq. (7) has the form

$$\rho_0(x) = \frac{1}{Z}\exp\left(\int_{x_0}^{x}\frac{u}{D}dx_1\right),$$
$$Z = \int_{-1}^{1}\exp\left(\int_{x_0}^{x}\frac{u}{D}dx_1\right)dx. \tag{11}$$

Using this equation to evaluate the total number of par-

ticles for $x > x_0$ and $x < x_0$, we obtain

$$N_0^+ = \frac{1}{Z}\int_{x_0}^{1}\exp\left(\int_{x_0}^{x}\frac{u}{D}dx_1\right)dx,$$
$$N_0^- = \frac{1}{Z}\int_{-1}^{x_0}\exp\left(\int_{x_0}^{x}\frac{u}{D}dx_1\right)dx, \quad (12)$$

which does not coincide in the general case with the initial distribution (9) for N^+ and N^-. It follows hence that the initial redistribution of particles, which takes the characteristic time

$$t_0 = 1/u = J\Gamma/\bar{G}^2$$

is followed by a longer process, during which a true steady-state distribution of particles (11), (12) sets in. The duration of this process is determined by diffusion of particles in a region in the vicinity of point x_0. Integrating Eq. (7) with respect to x from x_0 to 1, we obtain the equation for the time variation of N^+,

$$\frac{d}{dt}N^+ = j_0, \quad (13)$$

where j_0 is the particle flux at point x_0.

It follows from Eqs. (11) and (12) that function u/D (the dependence of this function on x is shown in Fig. 2) plays an important role. This function possesses the property

$$\left.\frac{u}{D}\right|_{x=\pm 1} = \pm 2J, \quad (14)$$

which is useful for subsequent calculations. To determine the value of j_0, we will use the fact that the second stage of the redistribution of particles to the stationary value is a slow process. From the standpoint of mathematics, we are using the presence of small parameter $1/J$ in the problem, which in fact is manifested in that the diffusion velocity is smaller than the drift velocity by a factor of $1/J$. This circumstance suggests that, at large distances from point x_0, in each of the regions $x > x_0$ and $x < x_0$, for long time intervals, we have a quasi-stationary distribution

$$\rho^+(x) = \frac{N^+}{N_0^+}\rho_0(x), \quad x > x_0, \quad (15)$$

$$\rho^-(x) = \frac{N^-}{N_0^-}\rho_0(x), \quad x < x_0. \quad (16)$$

On the other hand, owing to the quasi-stationary condition, the following equation is valid in the vicinity of point $x = x_0$:

$$u\rho - D\frac{d}{dx}\rho = j. \quad (17)$$

The solution to this equation is obvious:

$$\rho(x) = \exp\left(\int_{x_0}^{x}\frac{u}{D}dx_1\right)$$
$$\times\left[\rho(x_0) - \int_{x_0}^{x}\frac{j}{D}\exp\left(-\int_{x_0}^{x_1}\frac{u}{D}dx_2\right)dx_1\right]. \quad (18)$$

In the vicinity of point x_0, we can also use the expansion

$$u(x) = (x - x_0)u', \quad u' = \left.\frac{du}{dx}\right|_{x=x_0}, \quad (19)$$

and calculate explicitly the integral in the exponential function on the right-hand side of Eq. (18):

$$\exp\left(-\int_{x_0}^{x}\frac{u}{D}dx_1\right) \approx \exp\left(-\frac{u'}{2D_0}(x - x_0)^2\right). \quad (20)$$

Since the latter function rapidly attenuates, we can rightfully take the quantity j/D outside the integral in Eq. (18) at point $x = x_0$ and use in the region

$$|x - x_0| > \sqrt{2D_0/u'} \sim 1/\sqrt{J}$$

the following relations instead of (18):

$$\rho^+(x) = \exp\left(\int_{x_0}^{x}\frac{u}{D}dx_1\right)$$
$$\times\left[\rho(x_0) - \frac{j_0}{D_0}\int_{x_0}^{1}\exp\left(-\int_{x_0}^{x_1}\frac{u}{D}dx_2\right)dx_1\right], \quad (21)$$

$$\rho^-(x) = \exp\left(\int_{x_0}^{x}\frac{u}{D}dx_1\right)$$
$$\times\left[\rho(x_0) + \frac{j_0}{D_0}\int_{-1}^{x_0}\exp\left(-\int_{x_0}^{x_1}\frac{u}{D}dx_2\right)dx_1\right]. \quad (22)$$

Comparing expressions (21) and (22) with (15)

and (16), we find that

$$j_0 = \frac{D_0}{Z\int_{-1}^{1}\exp\left(-\int_{x_0}^{x}\frac{u}{D}dx\right)}\left(\frac{N^-}{N_0^-}-\frac{N^+}{N_0^+}\right). \quad (23)$$

We will calculate N_0^+ and N_0^- by using obvious approximations, such as

$$ZN_0^+ = \int_{x_0}^{1}\exp\left(\int_{x_0}^{x}\frac{u}{D}dx_1\right)dx = \exp\left(\int_{x_0}^{1}\frac{u}{D}dx\right) \times\int_{x_0}^{1}\exp\left(-\int_{x}^{1}\frac{u}{D}dx_1\right)dx \approx \frac{\exp\left(\int_{x_0}^{1}\frac{u}{D}dx\right)}{2J}. \quad (24)$$

On segment $x_0 \le x \le 1$, we approximately assume that

$$\frac{u}{D} \approx 2J\frac{x-x_0}{1-x_0}, \quad (25)$$

which finally gives

$$ZN_0^+ \approx \frac{\exp(J(1-x_0))}{2J}. \quad (26)$$

Analogously, for N_0^- we obtain

$$ZN_0^- \approx \frac{\exp(J(1+x_0))}{2J}. \quad (27)$$

Finally, using approximation (19), (20), we obtain the following expression for j_0:

$$j_0 = 2J\sqrt{\frac{D_0u'}{2\pi}}(N^-\exp(-J(1+x_0)) - N^+\exp(-J(1-x_0))) = \sqrt{\frac{3}{2\pi J}}\frac{(1-x_0^2)^{3/2}}{2}\frac{\bar{G}^2}{\Gamma}(N^-\exp(-J(1+x_0)) - N^+\exp(-J(1-x_0))). \quad (28)$$

Substituting this expression for flux j_0 into Eq. (13) and using normalization condition (10), we obtain

$$\frac{d}{dt}N^+ = \sqrt{\frac{3}{2\pi J}}\frac{(1-x_0^2)^{3/2}}{2}\frac{\bar{G}^2}{\Gamma} \times[\exp(-J(1+x_0)) - N^+ \times(\exp(-J(1+x_0)) + \exp(-J(1-x_0)))]. \quad (29)$$

It follows from this equation that the duration of the second stage in the evolution of the angular momentum orientation is given by

$$t_1 = t_0\sqrt{\frac{2\pi}{3J}}\frac{1}{(1-x_0^2)^{3/2}}\frac{e^J}{\cosh(Jx_0)}. \quad (30)$$

Simple estimates show that the second stage duration is $t_1 \approx 40t_0$ for $J = 5$ and for ellipticity $x_0 = -0.3$.

Thus, in the description of the angular momentum based on the semiclassical approximation, the problem of the interaction between a two-level atom and elliptically polarized radiation in the presence of a magnetic field can be formally reduced to the diffusion-drift equation of particle transfer in the coordinate corresponding to the angular momentum projection on the quantization axis. For the optical transition

$$J_g \longrightarrow J_e = J_g + 1,$$

the drift velocity is alternating; as a result, particles are concentrated depending on the angular momentum orientation parallel or antiparallel to the magnetic field. However, when the polarization differs from linear, one of these directions dominates in the sense that, after a long time, the angular momenta of most particles are oriented in this direction. As a result of diffusion, particles are "pumped" from the less advantageous direction of their angular momentum orientation to the opposite direction. However, the rate of such pumping is proportional to the number density of particles in the vicinity of point x_0, at which the drift velocity vanishes. In view of the properties of the solution to the diffusion-drift equation, the number density of particles in this region is found to be exponentially small,

$$\exp\left(-\int_{x_0}^{\pm 1}(u/D)dx\right) \sim e^{-J},$$

as compared to the number density of particles concentrated in the above-mentioned directions. This circumstance determines the slow rate of the second process.

After completion of the second stage, the numbers of particle oriented along and against the magnetic field satisfy the equilibrium relation

$$N_0^- = \exp(2Jx_0)N_0^+,$$

which is quite sensitive to ellipticity x_0 for large values of J.

ACKNOWLEDGMENTS

This study was financed by the Russian Foundation for Basic Research (project no. 01-02-17433).

REFERENCES

1. A. Aspect, E. Arimondo, R. Kaiser, *et al.*, Phys. Rev. Lett. **61**, 826 (1988).
2. R. Wynands and A. Nagel, Appl. Phys. B: Laser Opt. **B68**, 1 (1999).
3. C. Andreeva, S. Cartaleva, Y. Dancheva, *et al.*, Phys. Rev. A **66**, 012502 (2002).
4. E. Arimondo, Prog. Opt. **35**, 257 (1995).
5. V. S. Smirnov, A. M. Tumaĭkin, and V. I. Yudin, Zh. Éksp. Teor. Fiz. **96**, 1613 (1989) [Sov. Phys. JETP **69**, 913 (1989)].
6. A. M. Tumaĭkin and V. I. Yudin, Zh. Éksp. Teor. Fiz. **98**, 81 (1990) [Sov. Phys. JETP **71**, 43 (1990)].
7. K. A. Nasyrov and A. M. Shalagin, Zh. Éksp. Teor. Fiz. **81**, 1649 (1981) [Sov. Phys. JETP **54**, 877 (1981)].
8. K. A. Nasyrov, Phys. Rev. A **63**, 043406 (2001).
9. K. A. Nasyrov, J. Phys. A: Math. Gen. **32**, 6663 (1999).

Translated by N. Wadhwa

Journal of Experimental and Theoretical Physics, Vol. 98, No. 3, 2004, pp. 496–502.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 125, No. 3, 2004, pp. 563–569.

PLASMA, GASES

Vacuum Microarc

A. V. Zharinov
All-Russia Institute of Electrical Engineering, Moscow, 111250 Russia
e-mail: agilim@vei.ru
Received August 5, 2003

Abstract—The physics of stationary vacuum microarc in a wide interelectrode gap with the perveance corresponding to a geometry of the Müller electron projector type and the Langmuir–Blodgett function $\alpha^2 \geq 5$ is considered on a qualitative level. Under these conditions, the electric field at the cathode can exhibit a significant (severalfold) increase due to a positive space charge of microarc, which makes field electron emission possible. The most important features of the continuity equation, Poisson equation, and thermal conductivity equation describing this system are considered. © *2004 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

Electric microarc is a fundamental phenomenon playing the role of a living cell in the complicated "organism" of high-current arc discharge, widely used, albeit still on the basis of empirical knowledge, in both science and technology [1]. As is known, a vacuum arc can be initiated in various ways, in particular, it appears as a result of the electric discharge between two electrodes in vacuum, thus restricting the maximum possible potential difference between electrodes to a level of about 10^5 V per millimeter of discharge gap width. Alternatively, a vacuum arc arises upon breakage of the contact between current-carrying electrodes. Depending on the circuit parameters, the current of a vacuum arc may vary from a fraction of ampere to tens and hundreds of kiloamperes, while the voltage drop varies from the ionization potential to several dozens of volts. A vacuum arc can exist between electrodes made of any metals (from mercury to tungsten) possessing substantially different evaporation rates, thermal conductivities, and other properties.

In most cases, electric arc is studied in a discharge gap with a width below 1 cm, where an important or even decisive role is played by anode evaporation. A high-current arc consists of numerous autonomous arcs, each arising from its own cathode spot emitting electrons. The discharge current via one cathode spot can vary from about 0.4 A for mercury up to 10^2 A for tungsten [2]. The current density measured on a cathode spot may reach up to 10^8 A/cm^2. Therefore, the cathode spot radius can be estimated at about 5×10^{-4} cm, the electron density is on the order of 10^{18} cm^{-3}, and the Debye radius is $5 \times 10^{-7}\sqrt{T_e}$ cm (T_e is the electron temperature expressed in electronvolts).

Thus, according to the results of observations, a vacuum arc with a sufficiently small current is a microscopic self-sustained discharge arising from a single cathode spot on the surface of a massive cathode. This very type of discharge, referred to below as vacuum microarc, is studied in this paper.

It is commonly accepted that the typical vacuum microarc consists of three parts: (i) a cathode spot; (ii) an active near-cathode region called the cathode layer, from which all ions are collected in the cathode spot to provide for continuous reproduction of the necessary flow of atoms from the cathode surface, the working temperature, the electric field strength, and the electron emission current I; and (iii) a passive anode layer playing the role of a conductor between the outer surface of the cathode layer and the anode surface. In the case of a distant anode, the anode layer is characterized by predominating negative space charge and by a very low density of atoms. Ionization of atoms in the anode layer does not play any significant role in the mechanism of self-sustained vacuum microarc. The conditions of existence and the properties of vacuum microarc are fully determined by interaction of the active cathode layer with the cathode spot surface.

The active cathode layer can be represented by a spherical sector of solid angle Ω with a cathode spot with an area of $\pi R^2 \approx \Omega r_0^2$ at the apex. The atomic flux from the cathode spot surface, with an angular distribution described approximately by the cosine law, propagates within a solid angle $\Omega_0 > \Omega$ irrespective of the electric field structure. A part of the atomic flux within the angle $\Delta\Omega = \Omega_0 - \Omega$ bypasses the active cathode layer and does not participate in discharge.

Figure 1 shows a schematic diagram of such an idealized vacuum arc with $\Delta\Omega = 0$. The outer boundary of the active cathode layer has a potential of φ_c and is spaced by $r_0 + a$ from the center (point O) of the spherical sector. Analysis of this model, despite its simplified character, allows the most important parameters to be established and their interrelation to be understood.

1063-7761/04/9803-0496$26.00

In the scheme of Fig. 1, unknown parameters include the potential φ_c, the electron emission current I, the ion current I_i to the cathode, the spot radius R, the solid angle Ω (or the quantity $r_0^2 \approx \pi R^2/\Omega$), and the cathode layer thickness a.

All the cathode characteristics important from the standpoint of the microarc's existence are considered as known. These include the evaporation rate, the rate of cathode sputtering by ion bombardment, and the density of electron emission current as a function of the cathode temperature (T_c) and the electric field strength (E_c) at the cathode, as well as thermal conductivity, electric conductivity, etc.

Based on the known properties of the cathode, vacuum microarc is described by jointly solving the equation of continuity for particle fluxes in the active cathode layer, the Poisson equation, and the thermal balance equation for the cathode spot surface. Rigorous solution of this problem is a very complicated task. For this reason, an analysis of the scheme in Fig. 1 will be performed using justified simplifying assumptions implying that a stationary vacuum microarc exists in a certain limited region of parameters.

2. THE EQUATION OF CONTINUITY AND THE CONDITION OF EXISTENCE OF VACUUM MICROARC

Let us restrict the consideration to a region of parameters in which the ion current fraction is small: $i = I_i/I \ll 1$. As is known, the Langmuir collisionless cathode layer is characterized by $i \geq \sqrt{m/M}$, where m and M are the electron and ion masses, respectively. For example, the latter ratio for tungsten is $\sqrt{m/M} \approx 1.7 \times 10^{-3}$. In the case of $i \geq 2\sqrt{m/M}$, the electric field E_c accounts for greater than 0.7 of the value corresponding to $i \gg 1$ (i.e., to a cathode layer with positive space charge). Therefore, the condition of $i \ll 1$ does not exclude field electron emission, provided that a field strength on the order of $E_c \approx 10^7$–10^8 V/cm is consistent with a joint solution of the Poisson and continuity equations. Thus, the region of parameters corresponding to $i \ll 1$ admits the electron emission via the Richardson–Dushman–Schottky and Fowler–Nordheim mechanisms.

Under the condition of $i \ll 1$, we may assume that I = const in the active cathode layer. Then, the ion current component can be expressed as $I_i = I\sigma_i N_g$, where σ_i is the average ionization cross section and N_g is the "atomic density" in the gas phase of the cathode layer:

$$N_g = \int_{r_0}^{r_0 + a} n_g(r)dr = \frac{i}{\sigma_i}. \quad (1)$$

For obtaining estimates, we will use the well-known approximated formula

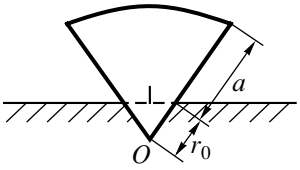

Fig. 1. Schematic diagram of the model of vacuum microarc (see the text for explanations).

$$\sigma_i = \sigma_m \frac{2.7}{W} \ln W,$$

where σ_m is the maximum ionization cross section, $W = \varphi_e/U_i$, φ_e is the kinetic energy of electron, an U_i is the ionization potential. For example, $i = 1.7 \times 10^{-3}$ and $\sigma_i = 10^{-16}$ cm^2 correspond to $N_g = 1.7 \times 10^{13}$ cm^{-2}. For such a low atomic density and a transport cross section of $\sigma_T < 10^{15}$ cm^2, electrons are barely scattered on atoms in the cathode layer. It should be also noted that, for $\sigma_m = 5 \times 10^{-16}$ cm^2 and $U_i = 8$ V, the average cross section $\sigma_i = 10^{-16}$ cm^2 can be obtained for two values of W: $W_1 = 1.083$ ($\varphi_e \approx 8.76$ V) and $W_2 = 54$ ($\varphi_e \approx 430$ V).

For the known values of i and σ_i, formula (1) gives a necessary condition for the existence of a stationary discharge obeying the equation of continuity for the fluxes of atoms, ions, and emitted electrons.

Let us consider the behavior of atoms in more detail, assuming their radial motion at a constant velocity of V_g. Then, atoms will travel the distance r without losing electrons with a probability of

$$\mu(r) = \exp\left(-\frac{r_0}{x_0}\right)\left(1 - \frac{r_0}{r}\right), \quad (2)$$

where

$$x_0 = eV_g/\sigma_i j \quad (3)$$

is the characteristic ionization length in the flow of electrons with the initial current density j (at $r = r_0$).

As a result of ionization, the density of atoms drops faster than according to the quadratic law,

$$n_g(r) = n_0 \frac{r_0^2}{r^2}\mu(r),$$

and the integration yields

$$N_g = \frac{i}{\sigma_i} = n_0 x_0 (1 - \mu), \quad (4)$$

where $\mu = \mu(r_0 + a)$. Note that, with allowance for

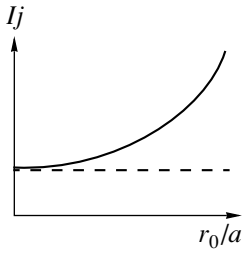

Fig. 2. The behavior of $Ij(r_0/a)$ for vacuum microarc with χ = const.

expression (3), formula (4) is equivalent to the relation

$$ki = (1-\mu), \tag{5}$$

where $k = j/eq_0$ and $q_0 = n_0V_g$ is the density of the atomic flux from the cathode surface at $r = r_0$. The probability $\mu(r_0 + a)$ is naturally called the transparency of the cathode layer, while the reciprocal quantity $\chi = \mu^{-1}$ is called the compression.

Another important process is the resonance recharge of atoms in the counterflow of ions. With allowance for this factor, the quantity k in relation (5) should be replaced by

$$\bar{k} \approx k\left(1 + \frac{\sigma_n}{\sigma_i}\frac{i}{2}\right),$$

where σ_n is the recharge cross section. Then, relation (5) should be rewritten as

$$\bar{k}i = h(1-\mu),$$

where the coefficient h takes into account the reverse flux of recharged ions ($1 < h < 2$). Since $\sigma_n > 10^2\sigma_i$, the recharge process may significantly decrease the transparency and thickness of the cathode layer, while leaving N_g = const. It should be noted that, as the ion current fraction i increases, the gas compression in the active cathode layer takes place predominantly due to the resonance recharge.

The model of vacuum microarc under consideration, characterized by a "zero cone transparency" ($\mathbf{V}_g \parallel \mathbf{r}$; $\Delta\Omega = 0$), is optimum from the standpoint of a minimum necessary rate of reproduction of the atomic flux from the cathode surface. An analogous situation with negligibly small "edge transparency" is actually possible for $R \gg a$.

The transparency of the cathode layer can vary in response to deformation of the spherical sector, which leads to a change in x_0, a, and/or r_0. However, for a real vacuum microarc in a stationary regime, the integral transparency must exactly compensate for reproduction of the atomic flux from the cathode spot surface by means of evaporation and cathode sputtering. If no such reproduction takes place, a vacuum arc can exist only in the case of zero integral transparency or infinite compression.

It is naturally assumed that ions are reflected from the cathode in the neutral state, rather than condensed on the cathode surface, so that the atomic flux density q_0 obeys the relation

$$eq_0 = j_i(1 + \gamma_i) + eq_s,$$

where $\gamma_i(\varphi_c)$ is the cathode sputtering coefficient and $q_s(T_c)$ is the evaporation rate. In a stationary regime,

$$\mu = \frac{\gamma_i j_i + eq_s}{eq_0}. \tag{6}$$

In particular, at a sufficiently low cathode temperature T_c for which cathode sputtering predominates (while still $\gamma_i \ll 1$), we obtain

$$\begin{aligned} \mu(\varphi_c) &= \gamma_i(\varphi_c) \ll 1, \\ \chi(\varphi_c) &= \mu^{-1} \gg 1. \end{aligned} \tag{7}$$

Thus, for the given values of φ_c = const and $eq_s \ll \gamma_i j_i$, the microarc transparency and compression are constant and determined by cathode sputtering. In this case,

$$\frac{a}{x_0} = \left(1 + \frac{a}{r_0}\right)\ln\chi. \tag{8}$$

For $\sigma_i = 10^{-16}$ cm^2 and $V_g = 5 \times 10^4$ cm/s, we obtain an estimate of $x_0 = 80/j$, for which relation (8) can be rewritten in terms of the product of the current I and current density j:

$$Ij \approx 3 \times 10^4\left(1 + \frac{r_0}{a}\right)^2 \ln^2\chi.$$

For $\ln\chi = 5$ ($\chi \approx 148$), we obtain

$$Ij \approx 7.5 \times 10^5\left(1 + \frac{r_0}{a}\right)^2. \tag{8'}$$

In the limit of $r_0/a \ll 1$, this yields $Ij \longrightarrow 7.5 \times 10^5$ A^2/cm^2 = const; in the other limiting case, $r_0/a \gg 1$, we obtain $Ij \longrightarrow 7.5 \times 10^5 r_0^2/a^2$. This behavior of $Ij(r_0/a)$ for χ = const is illustrated in Fig. 2.

Using to the results reported by Daalder for vacuum microarc on copper cathode (see review by Harris [2]),

$$I = 50 \text{ A}, \quad j \approx 2 \times 10^8 \frac{\text{A}}{\text{cm}^2}, \quad Ij = 10^{10} \frac{\text{A}^2}{\text{cm}^2},$$

we obtain an estimate of $R \approx 2.8 \times 10^{-4}$ cm. Taking $Ij \approx 10^{10}$ A^2/cm^2, in relation (8'), we obtain $r_0/a \approx 115$. Rela-

tion (8) for $r_0/a \gg 1$ correctly reflects the dependence of a/x_0 on $\ln\chi$, since the edge transparency can be ignored. Then, applying relation (8) to the experimental data of Daalder, we may conclude that the maximum current of a field-emission microarc, $I \approx 50$ A, corresponds to a "degenerate" one-dimensional geometry of the active cathode layer ($r_0/a \gg 1$).

The scheme of vacuum microarc under consideration implies that the superimposed flows of atoms, ions, and electrons inside the spherical segment proceed in the radial direction and do not intersect the cone surface. This idealization leads to a severalfold decrease in the lower limiting value of the microarc current (called threshold current). The threshold current corresponds to a certain minimum ratio $(r_0/a)_{\min}$ for which the atomic density N_g decreases below a minimum possible level ($N_g < i/\sigma$), leading to the appearance of "vacuum" in the cathode layer volume and to the quenching of arc discharge. In the vicinity of this threshold, the discharge may be conventionally called point microarc. It is obvious that, for $(r_0/a)_{\min} \le 1$, the threshold current coincides in order of magnitude with that for $r_0/a \ll 1$.

In the field emission regime at $j \sim 10^8$ A/cm^2, the threshold current of a point microarc according to formula (8') is approximately 7.5×10^{-3} A. It is interesting to note that a current of the same order of magnitude is observed for field emission from a microscopic point [3]. Therefore, the question naturally arises as to how can the field electron emission from a microscopic point be distinguished from that in the case of microarc.

In this context, it is necessary to emphasize the need for taking into account the ionization and recharge processes in the description of stationary electron emission for $j \sim 10^8$ A/cm^2, when the ionization length is $x_0 \sim 10^{-6}$ cm, that is, much smaller than the point radius. It is not excluded that stationary field emission in the Müller projector can be observed for a flat cathode.

Finally, note that the above question has a physical meaning only provided that both degenerate and point microarcs involve an electric field strength of $E \sim 10^8$ V/cm.

3. ELECTRIC POTENTIAL AND FIELD DISTRIBUTION IN VACUUM MICROARC

In many investigations, it is assumed that the active cathode layer is filled with a quasineutral plasma and the current density obeys the "3/2 law," which is valid for the Langmuir collisionless cathode layer. However, there is some doubt concerning the validity of these assumptions in the case of vacuum microarc. At a low density of atoms ($N_g < 10^{14}$ cm^{-2}) and a monotonic potential distribution in the cathode layer, all electrons travel the distance to the potential virtually without colliding with atoms in this layer. For this reason, the cathode layer is characterized by predominating positive space charge and the Poisson equation in the zero approximation can be written with neglect of the space charge of electrons. This assumption does not seem preposterous in view of a microscopic size of vacuum microarc.

In order to check this, let us obtain estimates for a one-dimensional cathode layer in which

$$n_g(x) = n_0 \exp(-x/x_0),$$

and the ion current density at a low transparency can be expressed as

$$j_i(x) = en_0 V_g \exp(-x/x_0).$$

Moving toward the cathode, ions are retarded as a result of the Coulomb interaction with the intense flow of electrons and the resonance recharge caused by the counterflow of atoms. Therefore, it can be reasonably assumed that the average ion velocity is comparable to that of the neutral atoms, $V_i \sim V_g$. Then, $en_i = en_0 \exp(-x/x_0)$, the electric field at the cathode obeys the condition

$$|E_c| \le 4\pi e n_0 x_0, \tag{9}$$

and the voltage drop is

$$\varphi_c \le 4\pi e n_0 x_0^2. \tag{10}$$

Formulas (9) and (10) reflect equality of the densities of atoms and ions in the cathode layer, $N_g \approx N_i$, and can be rewritten as

$$|E_c| \le 1.8 \times 10^{-6} N_g \ [\text{V/cm}], \tag{9'}$$

$$\varphi_c \le 1.44 \times 10^{-4} \frac{N_g}{j} \ [\text{V}]. \tag{10'}$$

For $i \sim 10^{-2}$ and $\sigma_i = 10^{-16}$ cm^2, we obtain $N_g = 10^{-14}$ cm^2, $|E_c| \le 1.8 \times 10^8$ V/cm, and $\varphi_c \ll 144$ V (for $E_c > 10^8$ V/cm, the field emission current density is $j \gg 10^8$ A/cm^2). These estimates seem to be quite realistic.

It should be also noted that, for $j \sim 10^8$ A/cm^2, the frequency of the Coulomb electron–ion collisions is much greater than that of the electron–atom collisions. Therefore, the electric conductivity in the cathode layer must correspond, in order of magnitude, to the Spitzer formula

$$j \approx 13 T_e^{3/2}(x) \frac{d\varphi}{dx}.$$

For a negligibly small ion current fraction, inelastic electron collisions are insignificant and the electron

temperature $T_e(x) \approx T_c + (2/3)e\varphi$ increases in proportion to the potential φ. In this case, the electric field at the cathode is $|E_c| \approx j/13T_e^{3/2}$. For $j \approx 2 \times 10^8$ A/cm^2 and $T_c \leq 0.3$ eV, this yields $|E_c| \geq 10^8$ V/cm. The potential $\varphi(x)$ can be determined from the formula $jx \approx 2.9\varphi^{5/2}$, which can be written for $eV_g/\sigma_i \approx 10^2$ as $\varphi(x) \approx 4(x/x_0)^{2/5}$. For $a/x_0 = \ln\chi \approx 5$, this yields $\varphi_c \approx 7.6$ V. The electric conductivity according to Spitzer corresponds to electroneutrality, but the order of magnitude is retained even in the case of a two- to threefold decompensation.

The first special feature of vacuum microarc apparently consists in that the ion concentration exponentially increases toward the cathode, rather than decreasing in this direction. The second (and main) feature is the positive space charge in the cathode layer. According to estimates, the electric field strength at the cathode and the current density may reach levels typical of the field emission at low values of φ_c. Moreover, it is quite possible that the maximum field strength E_c at the cathode of a microarc may exceed the values achieved as a result of the field enhancement at a microscopic point.

In concluding this section, it should be emphasized that the initial assumption concerning a passive role of the anode layer in vacuum microarc is valid provided that the microarc current $I(\varphi_c)$ is consistent with the current $I_a(L_a, \varphi_c, \varphi_a)$ limited by the perveance of the anode volume, that is, depends on the anode potential φ_a and the discharge gap width L_a. For $I > I_a$, a nonstationary regime can take place as a result of the field sagging in the region of $x > a$ or even an aperiodic instability of the Pierce type can develop with the formation of a virtual cathode and current breakage. Such processes were considered in detail, with description of original experiments, by Nezlin [4].

Thus, with exponential growth of the gas and ion density toward the cathode, field emission under stationary vacuum microarc conditions seems to be quite possible even from a flat cathode surface.

4. THERMOPHYSICS OF A VACUUM MICROARC CATHODE

The thermophysics of cathode spots is a very interesting and complicated problem of independent basic significance. However, judging by the available literature, this problem was never systematically nor thoroughly studied and was not given proper attention in reviews. For the description of thermal processes in the cathode spot of vacuum microarc, it will be necessary to make simplifying assumptions as it was done in the preceding sections.

Consider a cathode spot in the form of a hemispherical well of radius r_0 on the surface of a massive cathode at a temperature of T_c. The microarc current $I = 2\pi r_0^2 j$ determines the characteristic scale of the Joule power evolved on the cathode,

$$Q_0 = IT_c\sqrt{L}.$$

Here, $L = 2.45 \times 10^{-8}$ is the Lorenz number, which can be considered as a fundamental constant determining the relationship between the thermal conductivity λ, the electric conductivity σ, and the temperature in the Wiedemann–Franz law $\lambda = \sigma TL$.

A heat flux from the cathode layer via the well surface is

$$Q_\lambda = 2\pi r_0^2 \lambda \left|\frac{dT}{dr}\right|_{r_0} > 0.$$

Let us assume that no heat exchange with the gas discharge volume takes place on the flat cathode surface outside the well. Then, the temperature of the well surface is determined by the thermal conductivity equation

$$\mathrm{div}\,\mathrm{grad}T = -\frac{j^2(r)}{\sigma(r)\lambda}$$

with the boundary condition $T|_{r \gg r_0} = \bar{T}$. A solution of this equation can be presented in the following form:

$$t = \cos\alpha - \beta\sin\alpha, \tag{11}$$

where

$$t = \frac{\bar{T}}{T_c}, \quad \beta = \frac{Q_\lambda}{Q_0}, \quad \alpha = \frac{jr_0\sqrt{L}}{\lambda}.$$

The current I can be expressed in terms of α as

$$Ij = 2.56 \times 10^8 \lambda^2 \alpha^2 \frac{\mathrm{A}^2}{\mathrm{cm}^2}. \tag{12}$$

The angle α determines the rate of temperature decrease in the bulk of cathode. In particular, for $\beta = 0$, we obtain

$$\frac{T(r)}{T_c} = \cos\left(\alpha - \frac{\alpha r_0}{r}\right).$$

In the regime of field emission from tungsten ($\lambda \approx 1$) for $j \approx 2.56 \times 10^8$ A/cm^2, we obtain $I \approx \alpha^2$. At a sufficiently small cathode spot radius, the microarc current can be very small (on the order of 10^{-2} A). For $\alpha \ll 1$, formula (11) yields $t \approx 1 - \beta\alpha$, which implies that the spot can be "cold" for $\beta\alpha \leq 1/2$.

Using the condition $\beta\alpha \leq 1/2$, we obtain

$$Q_\lambda \leq \frac{Q_0}{2\alpha} = \frac{\alpha T_c\sqrt{L}}{2} \approx 8 \times 10^{-5}\alpha T_c \ll 1 \text{ W}.$$

Since $\beta \sim 1/2\alpha \gg 1$, the ion-bombardment-induced

heating predominates. However, the heat flux is such that

$$Q_\lambda \le I_i(\varphi_c + U_i) \approx iI(\varphi_c + U_i).$$

For $i \sim 10^{-2}$, $I \sim 10^{-2}$ A, $T_c \approx 600$ K, and $\alpha \sim \sqrt{I} \sim 10^{-1}$, we obtain an estimate of $\varphi_c \ge 40$ V. This implies that a low-current stationary microarc with $I \sim 10^{-2}$ A is possible both on a big massive cathode and on a cathode with dimensions on the order of several micrometers. For the above estimates, the spot diameter is smaller than the characteristic scale of the polycrystalline structure of a massive cathode. In this case, using the values of λ from handbooks is physically meaningless and estimation of the minimum current of a field-emission microarc becomes a difficult problem.

In the opposite limiting case of $\beta = Q_\lambda/Q_0 \ll 1$, expression (11) implies that, for $\alpha \longrightarrow \pi/2$, $t \longrightarrow 0$ and the problem has no stationary solutions. For $\alpha = \pi/2$, this situation takes place when

$$(Ij)_{cr} = 6.3 \times 10^8 \lambda^2 \text{ A}^2/\text{cm}^2. \quad (13)$$

For copper ($\lambda = 3.5$), this yields $(Ij)_{cr} = 7.74 \times 10^9$ A^2/cm^2. This value is in satisfactory agreement with the result obtained by Daalder: $(Ij)_{cr} = 10^{10}$ A^2/cm^2. Thus, the upper limiting microarc current is probably related to a thermal instability of the cathode spot developed at $\beta \longrightarrow 0$ [2].

Returning to the approximation of $t \approx 1 - \alpha\beta$, note that, irrespective of the temperature, the equality $\alpha\beta\rho = 1$ is valid for ρ determined by the temperature gradient on the spot surface via the relation

$$\left|\frac{dT}{dr}\right|_{r_0} = \frac{T_c}{r_0\rho}.$$

Therefore, $\rho \longrightarrow 1$ implies $\alpha\beta \longrightarrow 1$ and, apparently, a thermal instability ($t \longrightarrow 0$) in a point microarc can develop for $\alpha \ll 1$. It is almost impossible to study rather complicated thermophysics of vacuum microarc in experiments on natural microscopic objects without recourse to modeling on specially designed setups.

5. CONCLUSIONS

It is believed that a high-current low-voltage vacuum arc consists of numerous fragments ("group spots") representing complexes of vacuum microarcs [2, 5]. Therefore, investigation into the properties of vacuum microarc is of primary significance. This task can be solved both by studying natural microscopic objects and by modeling separate processes and parts of a microarc. For example, gas compression, potential distribution, and stability of the quasispherical cathode layer can probably be studied in a non-self-sustained discharge with an incandescent cathode of macroscopic dimensions. The thermophysics of the microscopic cathode spot, including the problems of thermal instability, can also be studied in detail using a probe or a cathode filament that has been self-heated in discharge, as was done by Pustogarov [6] for plasmatrons. Modeling seems to offer a quite justified method stimulating fundamental rather than purely technological approach to arc discharge and promising both new important basic knowledge and effective solutions of applied problems.

To summarize, the following conclusions can be emphasized.

1. At a high current density, the cathode layer features densification (described in terms of compression χ) related to a higher probability of ionization and resonance recharge of atoms in the gas phase. The characteristic size of the cathode layer is on the order of the ionization length $x_0 \approx 10^2/j$ (for $j > 10^7$ A/cm^2, this yields a microscopic value below 10^{-5} cm).

2. On the microscopic scale, the entire cathode layer is characterized by a predominating positive space charge. For this reason, the electric field at the cathode E_c and the potential φ_c can be estimated by the order of magnitude as $E_c \approx 1.8 \times 10^{-6} N_g$ [V/cm] and $\varphi_c \approx 1.4 \times 10^{-4} N_g/j$ [V], where $N_g = i/\sigma_i \approx n_0 x_0$ [cm^{-2}]. For $N_g \ge 10^{-14}$ cm^2, this yields $E_c \ge 10^8$ V/cm, $\varphi_c < 10^2$ V, and $j \approx 10^8$ A/cm^2. Therefore, a positively charged active cathode layer of vacuum microarc provides conditions for field electron emission.

3. In a stationary microarc, the compression is inversely proportional to the rate of reproduction of the atomic flux from the cathode spot. In the vicinity of a cathode sputtering threshold, the compression may reach a level of $\chi > 10^4$.

4. For the cathode layer geometry modeled by a spherical segment, the compression depends on the ionization length x_0, the cathode spot radius R, and the cathode layer thickness a. For a preset constant compression, there are certain possible relations between the spatial characteristics x_0, R, and a corresponding to various stationary values of the microarc current I and the current density j:

(i) for $R \gg a$ (degenerate one-dimensional regime),

$$(Ij)_1 \approx 3 \times 10^4 \frac{R^2}{a^2} (\ln\chi)^2;$$

(ii) for $a \gg R$ (point microarc regime),

$$(Ij)_2 \approx 3 \times 10^4 (\ln\chi)^2.$$

For a point microarc with $j \approx 10^8$ A/cm^2, the microarc current is on the same order of magnitude as the current of field emission from a microscopic point.

5. The ascending current–voltage characteristic and self-sustained character of vacuum microarc, together with the possibility of stationary states, admit the formation of complexes (clusters [2]) of a large number of microarcs on a smooth (flat) surface, arranged in a certain order and satisfying the conditions of a stable thermal regime and self-sustained operation. If this hypothesis is valid, it will be possible to create stationary, high-current high-voltage field emission vacuum diodes using smooth massive cathodes.

ACKNOWLEDGMENTS

The author is grateful to Yu.A. Kovalenko and A.N. Ermilov for support and promotion of this study.

REFERENCES

1. G. A. Mesyats, *Ectons* (Nauka, Ekaterinburg, 1993).
2. L. Harris, in *Vacuum Arcs: Theory and Application*, Ed. by J. Lafferty (Wiley, New York, 1980; Mir, Moscow, 1982).
3. V. N. Shrednik, in *Unheated Cathodes*, Ed. by M. I. Elinson (Sovetskoe Radio, Moscow, 1974), p. 216.
4. M. V. Nezlin, *Beam Dynamics in a Plasma* (Énergoatomizdat, Moscow, 1982).
5. V. I. Rakhovskiĭ, *Physical Principles of the Commutation of Electric Current in Vacuum* (Nauka, Moscow, 1970).
6. A. V. Pustogarov, in *Experimental Studies of Plasmatrons* (Nauka, Novosibirsk, 1977), p. 315.

Translated by P. Pozdeev

Journal of Experimental and Theoretical Physics, Vol. 98, No. 3, 2004, pp. 503–507.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 125, No. 3, 2004, pp. 570–575.

PLASMA, GASES

Electroacoustic Oscillations of Aluminum Oxide Particles in the Thermal Plasma

G. S. Dragan
Mechnikov National University, Odessa, 65026 Ukraine
e-mail: pipeaes@te.net.ua
Received August 5, 2003

Abstract—The frequencies of natural electroacoustic oscillations of aluminum oxide particles in a laminar disperse aluminum flame are determined experimentally using the capacitive method. A computational model is proposed for estimating the natural frequency of oscillations of charged particles in the smoky plasma taking into account the Doppler effect. It is shown that, for a natural frequency of oscillations of 51 kHz, two measured maxima at frequencies of 30 and 60 kHz in the oscillation spectrum correspond to the Doppler frequencies. © *2004 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

Combustion products obtained as a result of burning of a metal powder in a gaseous oxidizer contain fine-dispersed particles of metal oxide, whose size may change depending on the combustion regime, from hundredths of a micrometer (volumetric condensation in the gas-phase combustion regime) to several micrometers (heterogeneous combustion and condensation) [1]. When the temperature of combustion products exceeds 2000 K, condensed particles are charged as a result of thermionic emission from the particle surface; consequently, the gas phase contains free electrons. If the gas phase is free of easy-ionized admixture of alkali metal atoms, the ionization of the gas can be disregarded. Such a variety of the combustion product plasma is sometimes referred to as a plasma sol, which in turn is a variety of low-temperature plasma with a condensed disperse phase (CDP) [2]. A plasma containing the condensed phase in the form of smoke grains is called a smoky plasma [3].

A specific feature of plasmas with CDP is the thermodynamic interaction at the interface, as a result of which intrinsic electric fields associated with surface processes, the charge state, and the mobility of charged plasma components (condensed particles, electrons, and ions) are induced in the plasma [2, 4]. In all probability, these processes in plasmas lead to the formation of linear chains of smoke grains [5] as well as ordered 3D structures, which were discovered in [6] and demonstrated in review [7]. In standing strata of gas discharge in dusty plasma, plasma–dust structures have a more perfect shape and were hence called plasma crystals [8, 9].

On the other hand, the interfacial thermodynamic interaction in the plasma with CDP gives rise to new modes and instabilities [10–12]. Acoustic oscillation in the condensed phase of a smoky plasma, which emerge as a result of electric charge fluctuations on the surface of grains, was studied by us for the first time in [10]. Later [11], analogous oscillations were studied in the dusty plasma of a dc glow discharge and were interpreted as a consequence of the plasma–dust current instability. The most comprehensive analysis of oscillation processes and instabilities in dusty plasmas is given in review [12]. Unfortunately, smoky plasmas with their specific features were outside of the author's interest, although some of the results considered in this review are applicable in both cases. Thus, electroacoustic oscillations in smoky plasmas have not been studied in actual practice. In addition, the above-mentioned features of plasmas with CDP (such as the formation of ordered structures and the emergence of vibrational–oscillatory processes) may be interrelated.

This study is devoted to experimental investigation of the spectrum of natural low-frequency electroacoustic oscillations of the condensed disperse phase in the front of a laminar flame of aluminum powder. An attempt is made to interpret the experimental results as a consequence of the formation of an ordered structure of aluminum oxide particles in the plasma flow.

First of all, we note the distinguishing features of two common types of plasma with a CDP:

(i) The dusty plasma is formed when light dust particles are introduced into a gas discharge under a low pressure or emerges in cosmic space in the presence of dust particles [12]; the typical features of such plasmas are the difference between the electron and ion temperatures, the insignificant role of thermionic emission from the surface of particles during their charging, and the collisionless nature of processes in the bulk; as a result, the self-consistent Vlasov equation for electrons can be applied.

1063-7761/04/9803-0503$26.00 © 2004 MAIK "Nauka/Interperiodica"

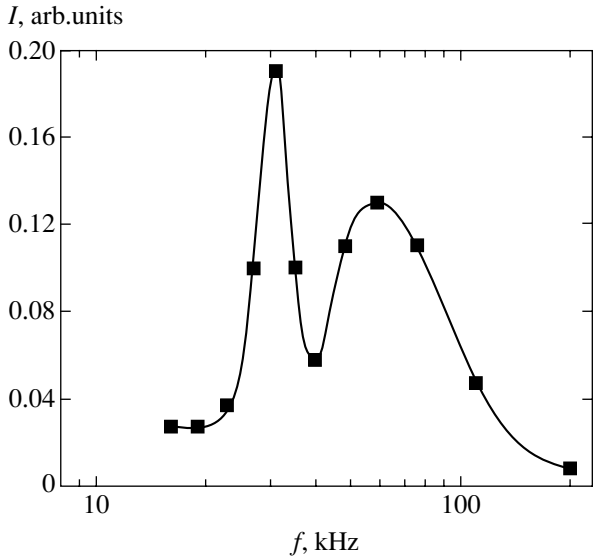

Fig. 1. Spectrum of low-frequency electroacoustic oscillations in the smoky plasma of aluminum oxide.

(ii) The smoky plasma contains smoke grains, i.e., particles formed in the combustion products in the course of volumetric condensation or as residues of the burning-out fuel [2]; free electrons are generated in the gas phase due to the thermionic emission from grains or as a result of ionization of gas atoms; the distribution of the self-consistent electrostatic potential in the vicinity of free charges and charged grains can be described by the Poisson–Boltzmann equation.

Thus, the laminar flame of aluminum particles under investigation is a typical example of a smoky plasma.

2. EXPERIMENTAL TECHNIQUE AND RESULTS

The experiments were performed on the setup described in [1]. Fine-disperse aluminum powder of the grade ASD-4 with an average grain size of 4 μm was carried by the air flow through a metallic pipe into the combustion zone formed above the pipe cross section after the ignition of the air–fuel mixture. The experimental technique made it possible to distribute the powder uniformly in the flow and to organize a laminar diffuse flame. The flame was conical in shape with a height of $L = 0.12$ m and a base diameter of $d = 0.028$ m at the burner throat. Two plane-parallel metallic plates of 0.16 m in height and with a width of 0.11 m were spaced 0.065 m apart and mounted along the flame axis. The plates were arranged so that there was an air gap between the plates and the flame. One of the plates was grounded via a resistor $R = 1$ kΩ. A constant potential of 4 kV relative to the grounded plate was applied to the other plate. As the flame propagated in the electrode gap, a voltage drop was created across resistor R and detected by a storage oscilloscope. The same signal was fed to a low-frequency spectrum analyzer.

Spectral analysis of radiation emitted by the flame in the visible spectral range revealed that the width L of the combustion front was approximately equal to 10^{-3} m and that the combustion products contained the gaseous and condensed phases. The gas phase contained uncontrollable admixture of sodium atoms and molecules with a high ionization potential.

The condensed phase was investigated by sampling of the flame followed by the electron microscopic analysis. It was found that the condensed phase was represented by Al_2O_3 submicrometer spherical smoke grains with a cubic mean diameter of 0.12 μm. The number density of particles and their mean charge were determined using the techniques described in [1]. The mean value of the number density of grains was $(2 \pm 1) \times 10^{17}$ m^{-3} and their mean charge was 30 ± 10 in units of electron charge.

The volume-averaged temperature of grains was determined from the continuous radiation spectrum by the polychromatic method, while the temperature of the gas phase was calculated from the absolute intensity of resonance lines for Na. Experiments proved that the temperature of the gas phase was equal to the temperature of the condensed phase to within the experimental error and amounted to $T = (3150 \pm 100)$ K.

The processing of the experimental data on oscillation processes in a smoky plasma in the electrode gap proved that the signal recorded from resistor R was alternating and bipolar in spite of the fact that a constant voltage was supplied to it. The averaged amplitude value of the signal was $U = 0.8 \pm 0.2$ V and the duration of oscillations was $\tau = 70 \pm 30$ ms. The oscillation process was induced periodically and at random.

The spectral composition of signal I is shown in Fig. 1. It can be seen that the oscillation process occurs in the frequency range from 15 to 200 kHz with two characteristic frequencies in the vicinity of 30 and 60 kHz with clearly manifested peaks. The first peak at a frequency of 30 kHz has a larger amplitude and low dispersion. The half-width of the distribution function amounts approximately to 1 kHz. The second peak with a smaller amplitude has a larger half-width of 6 kHz.

3. DISCUSSION OF EXPERIMENTAL RESULTS

In order to explain the observed oscillations of the voltage drop across standard resistor R, we consider the processes occurring in the electrode gap. The constant voltage applied to the plates obviously does not induce a current in the circuit since there is no contact between the plasma and the plates. The presence of an air gap between the plasma and the plates also ensures the absence of conduction current in the plasma and, hence, its polarization. In this case, current may appear in the circuit only as a result of a change in the permittivity in the electrode gap; i.e., oscillations are capacitive by nature. Considering that the flame under study is sta-

tionary and oscillations of its size with such a frequency are hardly probable, we can assume that oscillations take place within the flame.

Let us determine the main plasma parameters. In the smoky plasma under investigation, the temperature of the electronic component is close to the gas temperature in view of a high collision rate (the pressure of the ambient is 10^5 Pa) and the absence of conduction current in the plasma flow. The gas phase is formed by molecules with a high ionization potential value; consequently, the ionization of these molecules is insignificant. We can assume that free electrons appear in the gaseous phase as a result of the thermionic emission from the surface of Al_2O_3 particles. Then the number density of free electrons can be determined from the quasi-neutrality condition

$$n_e = Zn_p, \tag{1}$$

where n_e is the concentration of free electrons in the gas phase, n_p is the concentration of Al_2O_3 particles, and Z is the mean charge of particles in units of elementary charge. Substituting the experimental values of the charge and number density of particles, we obtain the average value of the number density of free electrons, which is equal to 6×10^{18} m^{-3}.

The Debye screening length in such a plasma medium is defined as the distance from the surface of a particle at which the value of the Debye potential is smaller than the Coulomb potential by a factor of 2.7 and is given by

$$D = \sqrt{\frac{kT}{4\pi e^2 n_e}} = 1.6 \times 10^{-6} \text{ m}.$$

This value is close to the mean distance between particles, $l \approx n_p^{-3} \approx 1.7 \times 10^{-6}$ m, and is much smaller than the characteristic scale of the plasma (this is a necessary condition for the existence of plasmas). The plasma frequency for the electron component is 1.4×10^{10} s^{-1}.

Clearly, electron oscillations cannot induce low-frequency oscillations; hence, we assume that the oscillations are induced by smoke grains. When charged particles are displaced relative to one another, a retrieving force appearing as a result of electrostatic interaction leads to oscillations propagating in the condensed phase of the smoky plasma. It can easily be verified that the potential energy of interaction between particles is comparable to the energy of thermal motion. Taking into account the electric origin of the interaction between the particles and the low-frequency nature of oscillations, we will refer to these oscillations as electroacoustic. It is interesting to note that oscillations of particles may in turn lead to the propagation of ultrasound in the gas phase.

4. THEORETICAL DESCRIPTION

We assume that a smoky plasma exhibits a certain ordering resembling the crystalline structure [8, 9] or its grains form the linear chains described in [5]. Then it becomes possible to consider a one-dimensional linear model, following the approach developed in [13], where the problem of propagation of waves in crystal lattices is described.

We assume that the system is monodisperse. Then a plane wave propagating along a chain of smoke grains can be represented in the form

$$u_j = A\exp[i(kja - \omega t)], \tag{2}$$

where u_j is the displacement of the jth particle relative to its equilibrium position in the chain, A is the amplitude of longitudinal oscillations, k is the wave number, ω is the angular frequency, a is the mean distance between particles, and t is the time.

Let us consider the interaction of the jth particle with its nearest neighbors with numbers $j - 1$ and $j + 1$. Taking into account the screening of the surface charge of a particle by the volume charge of electrons, we can write the expressions for the distribution of potential φ in the vicinity of the particle and strength E of the electric field produced by the particle in the form

$$\varphi = \frac{eZ}{r}\exp\left(\frac{r_p - r}{D}\right),$$

$$E = \frac{eZ}{r^2}\exp\left(\frac{r_p - r}{D}\right) + \frac{ez}{rD}\exp\left(\frac{r_p - r}{D}\right),$$

where e is the electron charge and r_p is the particle radius.

This leads to the following expression for the force of electrostatic interaction between the jth and $(j + 1)$th particles:

$$F_{j,j+1} = \frac{Z^2e^2}{(a + u_{j+1} - u_j)^2}\exp\left(-\frac{a + u_{j+1} - u_j}{D}\right) + \frac{Z^2e^2}{(a + u_{j+1} - u_j)D}\exp\left(-\frac{a + u_{j+1} - u_j}{D}\right). \tag{3}$$

An analogous expression can be derived for force $F_{j,j-1}$.

The resultant force acting on the jth particle is

$$F_j = F_{j,j+1} - F_{j,j-1}. \tag{4}$$

We expand expression (3) and a similar expression for $F_{j,j-1}$ into a Taylor series; assuming that u_{j-1}, u_j,

$u_{j+1} \ll a$, we confine our analysis to the linear terms of the expansion. Then formula (4) is transformed to

$$F_j = \mu(2u_j - u_{j+1} - u_{j-1}), \quad (5)$$

where

$$\mu = \frac{Z^2e^2}{a^3D^2}\exp\left(\frac{r_p - a}{D}\right)(a^2 + 2D^2 + 2aD). \quad (6)$$

The equation of motion of a particle of mass m can be written in the form

$$m\frac{d^2u_j}{dt^2} = -\mu(2u_j - u_{j+1} - u_{j-1}). \quad (7)$$

Substituting expression (2) into (7), we obtain the dispersion equation for longitudinal waves propagating along a linear chain in the condensed phase of a smoky plasma:

$$\omega = \pm\omega_0 \sin\frac{ka}{2}, \quad (8)$$

$$\omega_0 = 2\sqrt{\frac{\mu}{m}}. \quad (9)$$

The plus and minus signs correspond to waves propagating in opposite directions.

In the range of long waves, when $ka \ll 1$, expression (8) assumes the form

$$\omega = v_0 k = ak\left(\frac{\mu}{m}\right)^{1/2}. \quad (10)$$

To estimate characteristic quantities, we will use the plasma parameters obtained from the experiment: $n_p = 2.0 \times 10^{17}$ m^{-3}, $r_p = 0.06$ μm, $T = 3150$ K, $Z = 30$, and the particle density $\rho = 3570$ kg/m^3.

Expressions (6), (9) and (10) then give

$$\mu \approx 8.4 \times 10^{-8} \text{ N/m}, \quad f_0 = \frac{\omega_0}{2\pi} \approx 51 \text{ kHz},$$

$$v_0 \approx 0.33 \text{ m/s}.$$

Pay attention to the rather low velocity of wave propagation (0.33 m/s), which is due to a large mass of particles ($m \approx 10^{-18}$ kg) as compared to the atomic mass.

To analyze the spectral composition of oscillations, we consider the frequency (density of states) distribution of modes. According to [13], the density of states for a linear one-particle chain can be written in the form

$$g(\omega) = \frac{1}{\pi}\frac{dk}{d\omega}. \quad (11)$$

Taking into account expression (8), we obtain

$$g(\omega) = \frac{1}{\pi v_0}\left[1 - \left(\frac{\omega}{\omega_0}\right)^2\right]^{-1/2}. \quad (12)$$

It can be seen that the density of states exhibits an explicit dependence on ω and turns to infinity for $\omega = \omega_0$. If we take into account the polydisperse nature of the system, the charge distribution of particles, and energy dissipation, the density of states will probably not turn to infinity. However, a maximum of the frequency spectrum must be in the vicinity of ω_0. It should be recalled that $f_0 = 51$ kHz. Experiments show the presence of two peaks near 30 and 60 kHz. It should be noted in this connection that, in spite of quantitative agreement between the experimental results and the predictions of such an approximate model, a serious drawback exists since the model predicts only one peak.

To explain the results obtained here, we consider the possibility of formation of two peaks in the oscillation spectrum as a result of the Doppler effect. In the combustion front, the smoky plasma moves at a velocity close to $v_{pl} \approx 0.6$ m/s, which exceeds the wave velocity v_0. For a monodisperse linear chain of particles, the phase (v_{ph}) and the group (v_g) velocities of longitudinal waves are given by

$$v_{ph} = \frac{\omega}{k} = v_0\frac{\sin(ka/2)}{ka/2}, \quad (13)$$

$$v_g = \frac{d\omega}{dk} = v_0\cos\frac{ka}{2}. \quad (14)$$

Taking into account Eq. (8), for $\omega = \omega_0$, we obtain $\sin(ka/2) = 1$ and $\cos(ka/2) = 0$. Then the group velocity is $v_g = 0$. Consequently, a standing wave formed at a frequency of $\omega = \omega_0$ can be represented as the result of summation of two counterpropagating traveling waves. In accordance with expression (13), the velocity of these waves is $v_{ph} = 2v_0/\pi$. One of these waves propagates along the plasma flow, while the other runs in the opposite direction. In a reference frame attached to the setup, we determine the values of frequencies of perceived signals,

$$f_1 = f_0\left(1 + \frac{v_{pl}}{v_{ph}}\cos\alpha\right)^{-1} = 41 \text{ kHz},$$

$$f_2 = f_0\left(1 - \frac{v_{pl}}{v_{ph}}\cos\alpha\right)^{-1} = 68 \text{ kHz},$$

where α is the angle between the direction of motion of the plasma and the wave vector. In our case, this angle is determined by the flame geometry: $\alpha = \arctan(2L/d)$, where $L = 0.12$ m is the flame height

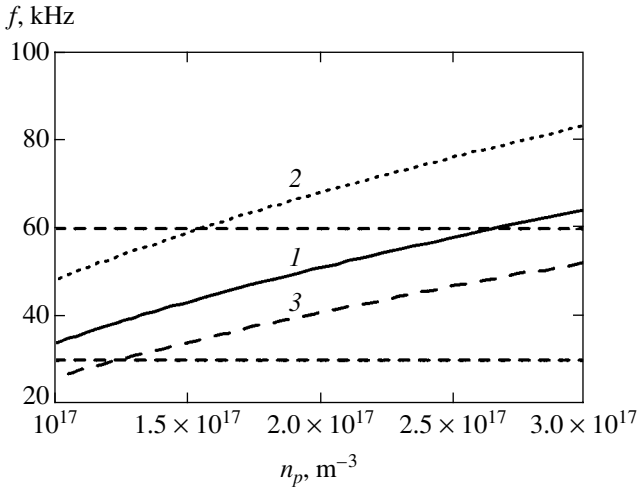

Fig. 2. Dependence of the natural frequency of oscillations (*1*) and Doppler frequencies f_1 (*2*) and f_2 (*3*) on the number density of smoke grains.

and d = 0.028 m is the flame diameter at the burner throat.

Thus, the values of frequencies obtained for oscillations of the condensed phase in a smoky plasma match the experimental values, which leads to the conclusion that the Doppler effect can be observed in the laminar flame of a metallic powder.

Let us consider the dependence of the Doppler frequencies of oscillations on the number density of smoke grains of aluminum oxide, shown in Fig. 2. Curve *1* describes the theoretical dependence of the natural frequency of oscillations in the condensed phase on the number density of particles, while curves *2* and *3* correspond to Doppler frequencies f_1 and f_2, respectively. Note that the oscillation frequency for charged smoke grains depends on their number density. As the value of n_p increases by a factor of three, the oscillation frequency of particles is almost doubled. It can be seen that the natural frequency of oscillations of particles and the frequencies of detected Doppler waves are determined not only by the parameters of the condensed phase, but also by the properties of particles, which determine their charge state (e.g., the work function for electrons escaping from the surface of particles to the plasma).

The horizontal dashed lines in the figure show for comparison the experimental frequency values of 30 and 60 kHz. It can be seen from the graphs that the experimental values of oscillation frequencies virtually coincide to within the error in the measurements of concentration of smoke grains.

Thus, the above analysis leads to the conclusion that an ordered structure of condensed submicrometer particles can be formed in the front of a laminar disperse flame. On the other hand, taking into account the assumption concerning the arrangement of particles in the flow in the form of chains, we can assume that the spatial ordering of particles and their charge state are responsible for the observed effects.

It is interesting to note that oscillations of smoke grains are macroscopic by nature; consequently, these oscillations may induce acoustic oscillations of the gas phase of the flame. The latter oscillations should apparently be detected by acoustic methods.

ACKNOWLEDGMENTS

The author expresses his sincere gratitude to A.V. Florko and N.I. Poletaev for fruitful discussions.

REFERENCES

1. A. N. Zolotko, Ya. I. Vovchuk, N. I. Poletaev, and A. V. Florko, Fiz. Goreniya Vzryva **32**, 24 (1996).
2. V. I. Vishnyakov, G. S. Dragan, and S. V. Margashchuk, *Plasma Chemistry*, Ed. by B. M. Smirnov (Énergoatomizdat, Moscow, 1990), No. 16, p. 98.
3. G. S. Dragan and V. I. Vishnyakov, in *Proceedings of 30th EPS Conference on Controlled Fusion and Plasma Physics* (St. Petersburg, Russia, 2003); http://eps2003.ioffe.ru/publics/pdfs/O-1.3B-pre.pdf.
4. G. S. Dragan, Vestn. Odessk. Gos. Univ., Fiz.–Mat. Nauki **8**, 163 (2003).
5. M. P. Kir'yakova and M. N. Chesnokov, Fiz. Aérodispersnykh Sist. **1**, 126 (1969).
6. G. S. Dragan, A. A. Mal'gota, *et al.*, in *Proceedings of International Scientific and Technical Meeting on Coal MGDÉS* (Alma-Ata, 1982), p. 77.
7. I. T. Yakubov and A. G. Khrapak, Sov. Technol. Rev. B **2**, 269 (1989).
8. J. H. Chu and Lin I, Phys. Rev. Lett. **72**, 4009 (1994).
9. A. P. Nefedov, O. F. Petrov, V. I. Molotkov, and V. E. Fortov, Pis'ma Zh. Éksp. Teor. Fiz. **72**, 313 (2000) [JETP Lett. **72**, 218 (2000)].
10. V. I. Vishnyakov, G. S. Dragan, and S. V. Margashchuk, in *Proceedings of III All-Union Meeting on Physics of Low-Temperature Plasma with CDP* (Odessa, 1988), p. 17.
11. V. I. Molotkov, A. P. Nefedov, V. M. Torchinskiĭ, *et al.*, Zh. Éksp. Teor. Fiz. **116**, 902 (1999) [JETP **89**, 477 (1999)].
12. P. K. Shukla, Phys. Plasmas **8**, 1791 (2001).
13. J. S. Blakemore, *Solid State Physics*, 2nd ed. (Cambridge Univ. Press, Cambridge, 1985; Mir, Moscow, 1988).

Translated by N. Wadhwa

Journal of Experimental and Theoretical Physics, Vol. 98, No. 3, 2004, pp. 508–514.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 125, No. 3, 2004, pp. 576–583.

PLASMA, GASES

Generation of Even Harmonics in a Relativistic Laser Plasma of Atomic Clusters

V. P. Kraĭnov and V. S. Rastunkov

Moscow Institute of Physics and Technology, Dolgoprudnyĭ, Moscow oblast, 141700 Russia
e-mail: krainov@online.ru

Received August 25, 2003

Abstract—It is shown that the irradiation of atomic clusters by a superintense femtosecond laser pulse gives rise to various harmonics of the laser field. They arise as a result of elastic collisions of free electrons with atomic ions inside the clusters in the presence of the laser filed. The yield of even harmonics whose electromagnetic field is transverse is attributed to the relativism of the motion of electrons and the consideration of their drift velocity associated with the internal ionization of atoms and atomic ions of a cluster. These harmonics are emitted in the same direction as odd harmonics. The conductivities and electromagnetic fields of the harmonics are calculated. The generation efficiency of the harmonics slowly decreases as the harmonic number increases. The generation of even harmonics ceases when the drift velocity of electrons becomes equal to zero and only the oscillation velocity of electrons is nonzero. The results can also be applied to the irradiation of solid-state targets inside a skin layer. © *2004 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

Interaction between superintense femtosecond laser pulses and large clusters [1, 2] or solid-state targets (in a skin layer) generates a plasma that consists of relativistic electrons and multicharged atomic ions. The process of multiple field-induced ionization is of tunneling or above-barrier character [3] because the Keldysh parameter γ in a superatomic field is very small:

$$\gamma = \frac{\omega\sqrt{2E_Z}}{F} \ll 1. \qquad (1)$$

Here, F and ω are the amplitude of the electric field strength and the frequency of laser radiation, respectively, and E_Z is the ionization potential of an atomic ion with the charge multiplicity Z. Throughout this paper, we use the atomic system of units $e = m_e = \hbar = 1$. The collision ionization of atomic ions is essential only in weak electromagnetic fields, when the electron velocity is small. Cluster beams have definite advantages over solid-state targets owing to the absence of a thin skin layer and a low reflection of an electromagnetic field from the surface.

In the case of a linearly polarized laser field, electrons leave atomic ions during multiple ionization and have an essentially nonuniform angular distribution with respect to the drift velocities (i.e., with respect to the initial velocities of electrons at the moment of ionization). Indeed, the characteristic values of the initial momenta of electrons along and perpendicular to the polarization of the laser field in the nonrelativistic case are given by [4–6]

$$p_\parallel = \sqrt{\frac{3\omega}{2\gamma^3}}, \quad p_\perp = \sqrt{\frac{F}{2\sqrt{2E_Z}}}, \quad p_\parallel \gg p_\perp. \qquad (2)$$

Of course, electrons may be ejected even with greater drift velocities (see formula (4) below), but with lesser probability. In the field of a titanium–sapphire laser of intensity 10^{19} W/cm^2 and for the ionization potential of a multicharged atomic ion of 500 eV, we have $p_\parallel \approx$ 100 au (c = 137 au); i.e., a typical longitudinal drift momentum is relativistic.

The oscillatory motion of electrons in the filed of a superintense laser pulse is still more relativistic. The relativistic momentum of the oscillatory motion is on the order of

$$p_\mathrm{F} = \frac{F}{\omega}. \qquad (3)$$

In the field of a titanium–sapphire laser of intensity 10^{19} W/cm^2, $p_\mathrm{F} \sim 300$ au; i.e., the oscillatory motion of electrons is essentially relativistic.

Collisions between electrons and multicharged atomic ions in the presence of a laser field give rise to an induced emission of field harmonics due to the nonmonochromatic motion of a free electron in the laser field. The nonrelativistic case for a linearly polarized laser field ($F/\omega \ll c$) has already been considered in detail by Silin [7–9]. In this limit, only odd harmonics (along the polarization vector of the laser field) are emitted. Silin also considered the case of weak relativ-

ism [10], when even harmonics are also emitted. However, the longitudinal field of this radiation is polarized along the wave vector of the external laser field; therefore, this radiation exists only inside the plasma and is not emitted outside. A similar case in the general relativistic statement has recently been considered in [11].

In [11], the drift momenta $p_\parallel$ and $p_\perp$ were small and therefore neglected compared to the oscillation momentum p_F. However, it was shown for the first time in [11] that the consideration of both oscillation and drift momenta leads to the generation of even harmonics, which can be experimentally observed. Indeed, the electric-field vector of these harmonics contains a component directed along the electric field of the external laser field; i.e., the field of even harmonics is transverse and is different from zero in the wave zone outside the plasma region. According to the results of [10, 11], this component vanishes at $p_\parallel = p_\perp = 0$.

In view of inequality (2), we assume that only the longitudinal drift momentum $p_\parallel$ is different from zero. To simplify the problem mathematically, we will not average over the distribution of this momentum at the moment of ionization, as it was done in Silin's works [7–9]; we just fix its value. Indeed, there is not much difference between the dependence of the harmonic yield on a current value of the longitudinal momentum and on the longitudinal temperature defined by formula (2). Under the tunneling ionization, the distribution over longitudinal drift momenta formally coincides with the Maxwell distribution [4, 12]:

$$w \propto \exp\left(-p_\parallel^2 \frac{\gamma^3}{3\omega}\right). \quad (4)$$

2. MOTION OF A RELATIVISTIC ELECTRON IN A SUPERINTENSE LASER FIELD

When solids are irradiated by a superintense laser field, the problem is complicated due to the fact that a larger part of a pulse is reflected by the surface of the skin layer. The electric field inside the thin skin layer is very small compared to the electric field of the incident electromagnetic wave and compared to the magnetic field inside the skin layer. The motion of a free relativistic electron inside the skin layer is essentially different from its motion in vacuum (in the latter case, the electron trajectory looks like figure 8 in the case of linear polarization). In particular, in the case of vacuum, the amplitude of two-dimensional oscillations of a relativistic electron in the plane passing through the polarization vector and the wave vector of the field is on the order of c/ω. This quantity is much greater than the skin depth c/ω_p (under the standard condition of a dense plasma, $\omega_p \gg \omega$), where $\omega_p = \sqrt{4\pi N_e}$ is the plasma frequency (N_e is the concentration of free electrons). Thus, in the case of a solid-state target, the oscillations of an electron are substantially distorted and damped due to the skin layer; the electron trajectory becomes similar to a one-dimensional trajectory along the polarization vector and closer to the surface of a solid.

One does not face such a problem when irradiating atomic clusters (of course, one also does not face such a problem when irradiating atomic gases; however, because of the low density of gases, the harmonic yield is small in the case of gas targets). The cluster radius R (of about tens of angstroms) is less than the skin depth (hundreds of angstroms), so that the external electromagnetic filed easily passes through a whole cluster. However, the oscillation amplitude of a relativistic electron is $c/\omega \gg R$. Therefore, the generation of harmonics occurs only at the moments when a relativistic electron passes through a cluster during its oscillations. In the case of large clusters, the external ionization of clusters is insignificant, so that there is not enough time for a cluster to substantially expand due to the Coulomb explosion during a femtosecond laser pulse.

Accordingly, the intensity of harmonics, calculated for the motion of an electron in a cluster medium, must be multiplied by a small factor $\omega R/c \ll 1$, which represents a fraction of the time that a relativistic electron remains inside a cluster. Bearing this in mind, we consider the motion of a free electron in the field of a superintense laser wave neglecting the effects of laser pulse focusing. The effect of the plasma medium reduces to the fact that the wave vector

$$k = \frac{\sqrt{\omega^2 - \omega_p^2}}{c}$$

of the laser field in the medium is different from the wave vector of a free electron in vacuum.

The Newton equations for the motion of a relativistic electron in the field of a linearly polarized wave can be solved analytically (although in the implicit form) [13]. Choose axis x along the propagation direction of the wave, axis y along its polarization, and axis z along the direction of the magnetic field. The kinematic momentum of an electron along axis y is defined by the relation

$$p_y(t) = p_\parallel + \frac{F}{\omega}\cos\varphi. \quad (5)$$

Here, $p_\parallel$ is the drift momentum along the polarization axis and $\varphi = \omega t - kx$ is the phase of the electromagnetic wave. The kinematic momentum of an electron along axis x (when the transverse drift momentum is neglected) is defined by

$$p_x(t) = \frac{1}{2\kappa}\left(\frac{F}{\omega}\cos\varphi + p_\parallel\right)^2 + \frac{c^2 - \kappa^2}{2\kappa}. \quad (6)$$

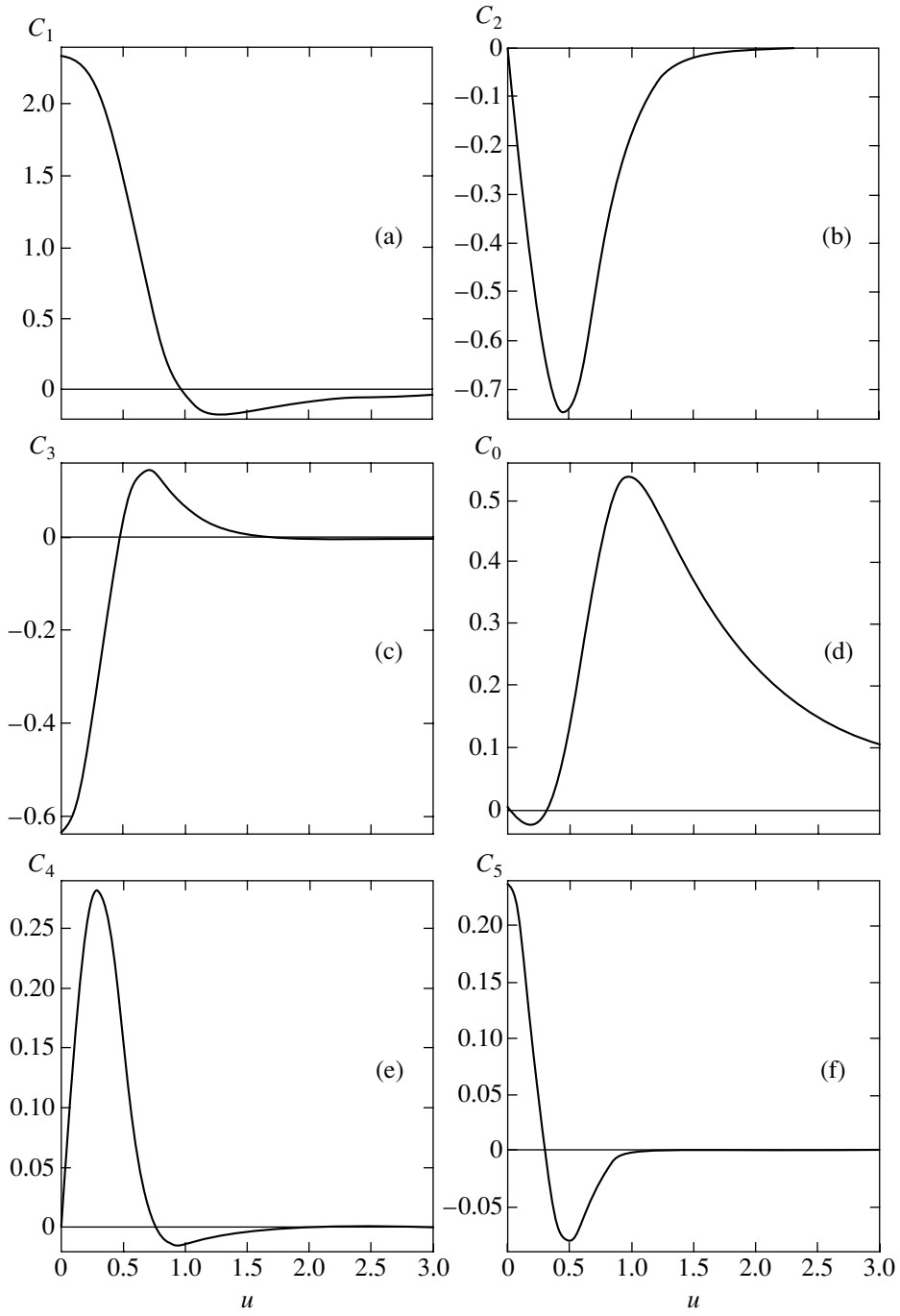

Fig. 1. Coefficients (a) C_1, (b) C_2, (c) C_3, (d) C_0, (e) C_4, and (f) C_5 as functions of the dimensionless drift momentum u.

Here, the constant κ is given by

$$\kappa = \sqrt{c^2 + p_\parallel^2 + \frac{F^2}{2\omega^2}}. \tag{7}$$

Finally, we set $p_z(t) = 0$: there is no motion along magnetic field (again, when one neglects the transverse drift momentum).

The components of the kinematic velocities of electrons along axes y and x are equal to

$$v_y(t) = \frac{2c\kappa p_y(t)}{c^2 + \kappa^2 + p_y^2(t)}, \qquad v_x(t) = \frac{2c\kappa p_x(t)}{c^2 + \kappa^2 + p_y^2(t)}, \tag{8}$$

respectively. Finally, the time differential dt can be

expressed in terms of the differential of the phase of the field, $d\varphi$, by the relation

$$dt = \frac{c^2 + \kappa^2 + p_y^2(t)}{2\omega\kappa^2} d\varphi. \quad (9)$$

The transport cross section of relativistic elastic scattering of an electron by an atomic ion of charge Z at small angles is defined by the Mott formula [14] (in atomic units):

$$\sigma_M = \frac{4\pi Z^2 \Lambda}{p^2(t) v^2(t)}. \quad (10)$$

Here, Λ is the Coulomb logarithm and $p(t)$ and $v(t)$ are the total momentum and velocity of the electron, respectively. In the limit of large velocities, the Coulomb logarithm represents a quantum logarithm [7].

The frequency of elastic electron–ion scattering is given by

$$\nu_{ei} = \sigma_M N_i v = \frac{4\pi Z^2 N_i \Lambda}{p^2(t) v(t)}. \quad (11)$$

Here, N_i is the concentration of atomic ions. Multiplying (11) by the velocity vector $\mathbf{v}$ of the electron, by the electron concentration N_e, and by the time interval dt, we obtain the density of the electric current of electrons:

$$d\mathbf{j} = -N_e \mathbf{v} \nu_{ei} dt. \quad (12)$$

This density has components along axes x and y. Note that this relation is also valid in the relativistic case (the so-called Pauli formula [13]).

The component of current (12) along axis x gives rise to the longitudinal electric field, which, as we mentioned above, does not exist outside the plasma region. Therefore, below we will concentrate only on the component of the electric current density that is directed along axis y. Substituting the expressions for the total velocity and momentum of the electron that were obtained above into (12), we obtain

$$dj_y = -AFf(\varphi) d\varphi. \quad (13)$$

Here, we denote

$$A = \frac{4\pi Z^2 N_e N_i \Lambda \omega}{F^3} \quad (14)$$

and define the function

$$f(\varphi) = \frac{(u + \cos\varphi)[1 + s(u\cos\varphi + (1/4)\cos 2\varphi)]}{[(u + \cos\varphi)^2 + s(u\cos\varphi + (1/4)\cos 2\varphi)^2]^{3/2}}. \quad (15)$$

The dimensionless constants u and s are defined by the relations

$$s = \left(\frac{F}{\omega\kappa}\right)^2, \quad u = \frac{p_\parallel \omega}{F}. \quad (16)$$

Equation (13) implies the following expression for the y component of the tensor of electric conductivity:

$$\sigma_y = \frac{1}{F}\int dj_y = -A\int_0^\varphi f(\varphi) d\varphi. \quad (17)$$

The conductivity σ_y is a nonlinear function of the electric field intensity F.

Expanding the integrand in (17) in terms of Fourier series, we obtain the following set of harmonics:

$$\sigma_y = -A\sum_{n=1}^{\infty} C_n \sin n\varphi - AC_0\varphi. \quad (18)$$

Here, the coefficients C_n of the Fourier series are defined by the following integral:

$$C_n = \frac{1}{\pi n}\int_0^{2\pi} f(\varphi)\cos n\varphi d\varphi, \quad (19)$$

$$C_0 = \frac{1}{2\pi}\int_0^{2\pi} f(\varphi) d\varphi. \quad (20)$$

One can see that both odd and even harmonics of conductivity are different from zero. They are coherent to the field of the incident electromagnetic wave. There also exists a zero-order harmonic, which corresponds to a constant electric field.

3. CALCULATION OF CONDUCTIVITY AT HARMONIC FREQUENCIES

The coefficients C_n in (19) and (20), which determine the conductivity for the harmonics of the external electromagnetic field, were calculated numerically as functions of the dimensionless drift momentum of an electron (see (16))

$$u = p_\parallel \frac{\omega}{F}. \quad (21)$$

We fixed a value of the dimensionless oscillation momentum of the electron

$$w = \frac{F}{\omega c}. \quad (22)$$

Then, the constant s, defined by relation (16), can be expressed in terms of u and w by the relation

$$s = \frac{1}{u^2 + 1/2 + 1/w^2}.$$

Figure 1 represents the calculated values of the coefficients C_n with $n = 0$–5 for a typical relativistic case of $w = F/\omega c = 2$, which corresponds to a peak intensity of 8×10^{18} W/cm^2 of a titanium–sapphire laser.

Figure 1a corresponds to a field generated at the fundamental harmonic. For drift momenta $u < 1$, we have $C_1 > 0$, which corresponds to the normal (positive) conductivity of the electron current (electrons move opposite to the direction of the electric field). When $u > 1$, the conductivity becomes negative (electrons move along the field). As expected, the field of the fundamental harmonic is the greatest among the fields of all the other harmonics. This makes it possible to determine the Joule absorption of electromagnetic energy by an atomic medium [15]. According to Fig. 1a, this absorption is determined by electrons with small drift velocities, which dominate in the expression for the absorption integrated over all drift velocities. The value of $C_1(0)$ coincides with that obtained in [11] for the case of $w = 2$, as expected.

In principle, an electron generated during tunneling or above-barrier ionization by an ac field may have any value of its drift momentum. However, the probability of large values of the drift momentum is suppressed due to the exponentially small probability (4) of generation of such electrons and due to the small value of the coefficient C_1 for large values of $p_\parallel$ (see Fig. 1a).

Figure 1b shows C_2 as a function of the dimensionless drift momentum u. According to the results of [11], the second harmonic along the polarization axis of the field is not generated for $u = 0$. The probability of its generation is maximal when $u \approx 0.5$ and decreases as u increases. The inequality $C_2 < 0$ implies that the conductivity of the second harmonic is negative (electrons move along the electric field vector of the electromagnetic field). Comparing Figs. 1a and 1b, we can conclude that the intensity of the second harmonic is not much smaller than that of the fundamental harmonic. However, substantial generation of the second harmonic occurs only for relativistic values of the drift momentum of an electron.

The coefficient C_3, which represents the amplitude of the third harmonic, is shown in Fig. 1c. The value of $C_3(0)$ also coincides with that obtained in [11] for the case of $w = 2$, as expected. When $u < 0.5$, the conductivity of the third harmonic is negative, whereas, for $u > 0.5$, it becomes positive.

Figure 1d corresponds to the static part of conductivity. It vanishes when $u = 0$ in accordance with the results of [11]. The static conductivity is mainly positive; even though its magnitude is smaller than the conductivity at the fundamental frequency, it is still sufficiently high. It slowly decreases as the drift momentum u increases.

Figure 1e represents the coefficient C_4 for the fourth harmonic. In the large, the intensity of this harmonic decreases as the harmonic number increases. As expected for an even harmonic, we have $C_4(0) = 0$.

Finally, Fig. 1f represents the coefficient C_5 for the conductivity at the fifth harmonic. The value of $C_5(0)$ coincides with that obtained in [11] for the case of $w = 2$. The conductivity at the fifth harmonic is positive for $u < 0.3$ and negative for $u > 0.3$.

The analysis of the results obtained allows us to draw the following conclusion. In a relativistic laser plasma, not only odd but also even harmonics, as well as a constant electric field along the polarization axis of the external linearly polarized electromagnetic field, are efficiently generated.

4. INTENSITY OF RELATIVISTIC HARMONICS

The expressions for the currents obtained above can be used for determining the electromagnetic fields of the generated harmonics according to Silin's approach [16]. According to (18), the Maxwell equation for the projection of the vector potential onto the polarization axis y of the external electromagnetic field (at the frequency of the nth harmonic) is expressed as

$$-\frac{1}{c^2}\frac{\partial^2 A_y^{(n)}}{\partial t^2} + \frac{\partial^2 A_y^{(n)}}{\partial x^2} + \frac{4\pi}{c} j_y'^{(n)} = -\frac{4\pi}{c}\int dj_y^{(n)} = -\frac{4\pi}{c}\sigma_y^{(n)} F = \frac{4\pi}{c} A C_n F \sin n\varphi. \quad (23)$$

Here, $j_y'^{(n)}$ denotes the density of the electron current that is not associated with the collisions between electrons and atomic ions but is attributed to the electromagnetic field of the generated harmonic (see below). The corresponding equation for the electric-field intensity at the harmonic frequency,

$$F_y^{(n)} = -\frac{1}{c}\frac{\partial A_y^{(n)}}{\partial t},$$

is obtained from (23) by differentiating with respect to time:

$$-\frac{\partial^2 F_y^{(n)}}{\partial t^2} + c^2\frac{\partial^2 F_y^{(n)}}{\partial x^2} - 4\pi\frac{\partial j_y'^{(n)}}{\partial t} = -4\pi A n\omega C_n F \cos[n(\omega t - kx)]. \quad (24)$$

Here, $k^2 = (\omega^2 - \omega_p^2)/c^2$ is the square of the wave number for the incident electromagnetic wave.

From Newton's second law of motion (in the nonrelativistic case, from the fact that the current generated by the field of a harmonic is small), we have

$$\frac{\partial j_y'^{(n)}}{\partial t} = N_e F_y^{(n)}.$$

Substituting this equation into (24), we find its solution

$$F_y^{(n)} = \frac{4\pi A\omega n C_n F}{(n^2-1)\omega_p^2}\cos[n(\omega t - kx)]. \quad (25)$$

This solution is also valid in the relativistic case $\omega_p > \omega$ because the field does not decay at a distance equal to a cluster size. Substituting the value of the constant A from (14) into (25), we finally obtain

$$F_y^{(n)} = \frac{Z\omega_p^2\omega^2 n C_n \Lambda}{(n^2-1)F^2}\cos[n(\omega t - kx)]. \quad (26)$$

From (26), we obtain the following expression for the ratio of the harmonic intensity to the intensity of the external electromagnetic field:

$$\eta^{(n)} = \frac{\left|F_y^{(n)}\right|^2}{\left|F\cos\varphi\right|^2} = \left|\frac{ZC_n\omega_p^2\omega^2 n\Lambda}{(n^2-1)F^3}\right|^2. \quad (27)$$

This ratio decreases as the intensity of the incident wave increases and as the harmonic number n increases.

Evaluating $F \sim \omega c$ for the general relativistic case, we obtain the following estimate for the generation efficiency of harmonics:

$$\eta^{(n)} \propto \left(\frac{Ze^2\omega_p^2 C_n n\Lambda}{m_e(n^2-1)c^3\omega}\right)^2. \quad (28)$$

Here, we recovered the charge and mass of an electron, which we set equal to unity above. The efficiency of a harmonic increases as the density of the atomic medium increases (whereby clusters are more efficient than a gaseous medium) and as the laser-field frequency ω decreases. The estimates obtained are also valid when the plasma frequency is greater than the laser frequency.

5. CONCLUSIONS

The generation of harmonics was experimentally observed by the authors of [17] on argon clusters (see also the survey [3]). It was shown that odd harmonics from the third to the ninth are generated on clusters consisting of several thousands of argon atoms; the generation efficiency is greater than that obtained with a gaseous medium of the same average density by a factor of 5. Moreover, in the case of clusters, one observed the generation of higher order harmonics than in the case of a gaseous medium. There was no generation of even harmonics because the intensity of laser radiation was less than 10^{18} W/cm^2 in the experiment. The generation of harmonics associated with the nonlinearity of the Mie oscillations (surface plasma oscillations of the electron cloud in a cluster) was insignificant in view of the small anharmonicity of the Mie oscillations. This conclusion was confirmed by numerical calculations for small metal clusters [18].

The results of this work can also be applied to the irradiation of solid-state targets by superintense laser pulses, where the aforementioned phenomena occur inside the skin layer. Even and odd harmonics of the laser field (from the second to the tenth) were observed by the authors of [19] for intensities higher than 10^{19} W/cm^2. The generation region of the harmonics corresponded to the electron concentration ranging from 10^{21} to 10^{23} cm^{-3}.

The results obtained in the present study show that the generation of even harmonics is determined by the drift velocity of electrons. During above-barrier ionization, an electron may acquire a sufficiently high drift velocity. Of course, in a superintense laser field, this velocity is not given by relation (2) but must be determined from relativistic theory. Preliminary estimates show that, even at intensities on the order of 10^{20} W/cm^2, this velocity is nonrelativistic in contrast to the oscillation velocity of electrons. However, an electron may acquire relativistic energy during a laser pulse when heating a plasma. This heating is attributed to the induced inverse bremsstrahlung of laser energy during collisions between electrons and multicharged atomic ions, reflections from the inner surface of a cluster, elastic scattering by charged clusters, excitation of surface plasma oscillations (Mie oscillations), etc. However, electron heating in a plasma always decreases as the kinetic energy of electrons increases because the collision rate of electrons with other objects decreases. The experimental results of [20] on the irradiation of argon clusters by a superintense femtosecond laser pulse have shown that the typical electron temperature is several keV. The energy spectra of electrons were measured in [21, 22] under the irradiation of xenon clusters by a 150-fs laser pulse with a peak intensity of 2×10^{16} W/cm^2. The mean energy of electrons was no greater than 2 keV. In spite of the fact that the drift velocity of electrons is small, it is this velocity that is responsible for the generation of even harmonics in a relativistic laser plasma.

ACKNOWLEDGMENTS

This work was supported by the Russian Foundation for Basic Research (project nos. 02-02-16678 and 04-02-16499), by the BRHE (project no. MO-011-0), and by the ISTC (project no. 2155).

REFERENCES

1. T. Ditmire, T. Donnelly, A. M. Rubenchik, *et al.*, Phys. Rev. A **53**, 3379 (1996).
2. G. Grillon, Ph. Balcou, J.-P. Chamberlet, *et al.*, Phys. Rev. Lett. **89**, 065005 (2002).
3. V. P. Krainov and M. B. Smirnov, Phys. Rep. **370**, 237 (2002).
4. P. B. Corkum, N. H. Burnett, and F. Brunel, Phys. Rev. Lett. **62**, 1259 (1989).
5. N. B. Delone and V. P. Krainov, J. Opt. Soc. Am. B **8**, 1207 (1991).
6. V. P. Krainov, J. Phys. B: At. Mol. Opt. Phys. **36**, L169 (2003).
7. V. P. Silin, Kvantovaya Élektron. (Moscow) **27**, 283 (1999).
8. V. P. Silin, Zh. Éksp. Teor. Fiz. **114**, 864 (1998) [JETP **87**, 468 (1998)].
9. V. P. Silin, Zh. Éksp. Teor. Fiz. **117**, 926 (2000) [JETP **90**, 805 (2000)].
10. V. P. Silin, Kratk. Soobshch. Fiz. **8**, 32 (1998).
11. V. P. Krainov, Phys. Rev. E **68**, 027401 (2003).
12. N. B. Delone and V. P. Krainov, *Multiphoton Processes in Atoms*, 2nd ed. (Springer, Berlin, 2000).
13. L. D. Landau and E. M. Lifshitz, *The Classical Theory of Fields*, 7th ed. (Nauka, Moscow, 1988; Pergamon Press, Oxford, 1975).
14. V. B. Berestetskiĭ, E. M. Lifshitz, and L. P. Pitaevskiĭ, *Quantum Electrodynamics*, 3rd ed. (Nauka, Moscow, 1989; Pergamon Press, Oxford, 1982).
15. G. Ferrante, M. Zarcone, and S. A. Uryupin, Phys. Plasmas **8**, 4745 (2001).
16. V. P. Silin, Zh. Éksp. Teor. Fiz. **47**, 2254 (1964) [Sov. Phys. JETP **20**, 1510 (1964)].
17. T. D. Donnelly, T. Ditmire, K. Neumann, *et al.*, Phys. Rev. Lett. **76**, 2472 (1996).
18. F. Calvayrac, P.-G. Reinhard, and E. Suraud, J. Phys. B: At. Mol. Opt. Phys. **31**, 1367 (1998).
19. M. Tatarakis, A. Gopal, I. Watts, *et al.*, Phys. Plasmas **9**, 2244 (2002).
20. T. Auguste, P. D'Oliveĭra, S. Hulin, *et al.*, Pis'ma Zh. Éksp. Teor. Fiz. **72**, 54 (2000) [JETP Lett. **72**, 38 (2000)].
21. T. Ditmire, E. Springate, J. W. G. Tisch, *et al.*, Phys. Rev. A **57**, 369 (1998).
22. R. A. Smith, J. W. G. Tisch, T. Ditmire, *et al.*, Phys. Scr. **80**, 35 (1999).

Translated by I. Nikitin

Journal of Experimental and Theoretical Physics, Vol. 98, No. 3, 2004, pp. 515–526.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 125, No. 3, 2004, pp. 584–597.

PLASMA, GASES

Analysis of Pair Correlation Functions for Macroscopic Particles in Dusty Plasmas: Numerical Simulation and Experiment

O. S. Vaulina*, O. F. Petrov, and V. E. Fortov

Institute for High Energy Densities, Joint Institute for High Temperatures, Russian Academy of Sciences, Moscow, 127412 Russia

**e-mail: industpl@redline.ru*

Received August 29, 2003

Abstract—Pair correlation is analyzed for systems of macroscopic particles with various isotropic interaction potentials. Under certain conditions, the behavior of the pair correlation function is determined by an effective order parameter and its decrease toward infinity follows an asymptotic power law. When the effective parameter is smaller than a certain critical value, the decay of pair correlation is much steeper. Experimental results concerning the form of the pair correlation function are presented for liquid-like dust structures localized in the near-electrode plasma sheath of a high-frequency capacitive discharge. An analysis of numerical and experimental results shows that melting dynamics in these systems are analogous to those characteristic of a topological phase transition. © *2004 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

Laboratory dusty plasmas are partially ionized gases containing micrometer-sized dust particles, which can have large negative or positive charges (10^3–$10^5 e$) and combine into quasi-steady liquid- or crystal-like structures [1–5]. These plasma-dust structures provide a unique tool both for analyzing the properties of essentially nonideal plasmas and for gaining deeper understanding of self-organization of matter. Studies of the properties of nonideal dusty plasmas play an important role in developing new phenomenological models of liquid-like systems. These studies are particularly important because strong interaction between particles in liquids makes it impossible to develop an analytical description of their thermodynamic characteristics based on the use of a small parameter, as done in the theory of gases [6–10].

The equilibrium properties of a liquid are comprehensively described by a set of probability density functions $g_s(\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_s)$ for particles at points $\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_s$. In the case of an isotropic binary interaction, the physical properties of a liquid (such as pressure, density, energy density, and compressibility) are determined by a pair correlation function $g(r) = g_2(|\mathbf{r}_1 - \mathbf{r}_2|)$ [6–9]. This function can be expressed as follows [10]:

$$g(r) = \exp(-[U(r)/T] + N(r) + B(r)), \qquad (1)$$

where $U(r)$ is the potential energy of binary interaction, T is the kinetic energy of chaotic (thermal) motion of particles, $N(r)$ is determined by the functions $g_1(\mathbf{r}_1)$ and $g(r)$, and $B(r)$ represents the effects due to higher order correlations and has the form of a complicated integral of $g_s(\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_s)$ if $s > 2$. In the popular hypernetted chain approximation, $B(r) = 0$, and $N(r)$ is determined using the Ornstein–Zernike relation [9, 10]. However, numerical studies have shown that the use of the hypernetted chain approximation leads to unsatisfactory results even for weakly nonideal systems [9–13]. Only allowance for higher order correlations (calculation of $B(r) \neq 0$) ensures agreement with numerical simulations [10]. Thus, determination of $g(r)$ generally requires not only information about the binary interaction potential, but also knowledge of the behavior of $g_s(\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_s)$ for $s > 2$ or use of some approximations of these correlation functions.

Unlike real fluids, laboratory dusty plasmas provide a good model for examining physical properties of nonideal systems, because dust particles are large enough to be imaged, which facilitates application of direct nonintrusive diagnostic methods. Interaction between dust particles in plasmas is commonly described by the Yukawa-type screened Coulomb potential

$$\varphi = \varphi_c \exp(-r/\lambda), \qquad (2)$$

where r is distance, λ is the screening radius, $\varphi_c = eZ_p/r$ is the Coulomb potential, and Z_p is the dust-particle charge. This assumption is consistent both with measurements of forces acting between two dust particles [14] and with a computed structure of a screening cloud [15] only at relatively short distances from the particle ($r < 5\lambda$). The screening weakens with increasing r, and

the asymptotic behavior of φ at $r \gg \lambda_D$ is governed by the power law [16]

$$\varphi \approx eZ_p a_p / r^2.$$

The results reported in [14–16] were obtained for isolated dust particles in plasmas. However, it remains unclear how the potential of interaction between two particles is modified by influence of other particles in a dense dust cloud, ionization of gas both in the cloud and outside it, collisions of electrons or ions with neutrals in the ambient gas, and other factors. Thus, the real potentials of interaction between particles are not known for dust grains in plasmas, and neither are they for many other physical systems in which interparticle interaction forces play an essential role.

Determination of the parameters responsible for the state of a system of interacting particles is an important task in the physics of nonideal dusty plasmas, as well as in other natural sciences. In particular, two dimensionless parameters responsible for mass transfer and phase state in Yukawa dissipative systems (with $\kappa = r_p/\lambda < 6$) were found in [17, 18]: the effective "nonideality" parameter

$$\Gamma^* = \Gamma\{(1 + \kappa + \kappa^2/2)\exp(-\kappa)\}^{1/2}$$

and the scaling parameter

$$\xi = \nu_{fr}^{-1} eZ_p\{(1 + \kappa + \kappa^2/2)\exp(-\kappa)n_p/\pi m_p\}^{1/2},$$

where n_p is the particle concentration, $\Gamma = (Z_p e)^2/Tr_p$ is the Coulomb coupling parameter, ν_{fr} is the friction coefficient for dust particles, and $r_p = n_p^{-1/3}$ is the mean distance between particles. A numerical model was tested against the laboratory experimental conditions in various dusty gas-discharge plasmas in [19–21]. Experimental studies showed that dust-particle dynamics in these plasmas can be described in terms of the parameters Γ^* and ξ. However, the parameters of the potential of interaction between particles can be determined only if additional information about its form is available.

The behavior of the pair correlation function reflects the phase state of a system. For example, liquid-like nonideal systems are characterized by short-range ordering of particles, whereas the functions $g(r)$ used for crystalline lattices describe long-range ordering. Numerical simulations showed that the effective parameter Γ^* of a Yukawa system (with $\kappa < 6$) completely determines the pair correlation function $g(r)$ (describing both long- and short-range ordering) in the interval from $\Gamma^* < 1$ to the point of crystallization into a body-centered cubic (BCC) lattice at $\Gamma^* \longrightarrow 106$. Thus, it was noted that the spatial correlation of particles in three-dimensional Yukawa systems with $\kappa < 6$ depends only on the ratio of the second derivative φ'' of the binary interaction potential $\varphi(r)$ at r_p to the particle temperature T. Moreover, both melting and crystallization processes (at $\Gamma^* \approx 102$–106) and formation of well-ordered clusters of dust grains (at $\Gamma^* \approx 22$–25) occur at nearly constant values of Γ^* [17, 18]. One may reasonably assume that this property holds under certain conditions for potentials of more general form describing binary interactions in many-particle systems. In the present study, we examine this assumption by analyzing pair correlation functions and conditions for phase transitions in systems characterized by various repulsive potentials.

Certain physical systems exhibit topological phase transitions between low- and high-temperature phases [22–30]. Topological phase transitions are more commonly observed in low-dimensional systems. Such a transition can be interpreted as a kind of "melting" that eliminates the positional ordering of the low-temperature phase at $T > T_m$ and preserves its orientational ordering (which breaks down at $T > T_0 > T_m$). Physically, this phase transition is explained by the formation of topological defects (dislocations and disclinations) in crystalline lattices. In a theory of this phenomenon developed for two-dimensional systems, melting is interpreted as transformation of a crystal into an isotropic liquid via an intermediate hexatic phase [22]. The theory was corroborated both by experimental studies of quasi-two-dimensional nonideal systems and by recent numerical simulations of extended two-dimensional many-particle systems with various binary interaction potentials [23–30]. In real monolayers of macroscopic particles, the topological mechanism of melting frequently manifests itself as topological excitations (vortices and antivortices) characteristic of finite two-dimensional systems [22]. Experimental investigation of topological phase transitions is a difficult task, because essential qualitative changes in real systems may be obscured by quasi-two-dimensional effects induced by small perturbations [23–26]. Nevertheless, short-range orientational ordering is observed not only in dust subsystems consisting of several (four to ten) dust layers, but also in simulated three-dimensional liquid-like structures. An analysis of three-particle correlation in systems of macroscopic particles with screened interaction potential (2) has revealed the formation of well-ordered clusters at $\Gamma^* > 25$ [31]. This observation is in good agreement with numerical results [17]. The qualitative changes can be attributed to topological defects of a three-dimensional lattice, and the transition can be associated with the existence of two distinct liquid phases by analogy with the topological phase transitions in two-dimensional systems.

Orientational ordering is commonly analyzed in terms of boundary angular autocorrelation functions $g_6(t)$ and static correlation functions $g_6(r)$ and $g(r)$ (see [23–30]). The correlation functions of two-dimensional systems were found to exhibit universal behavior in topological phase transitions [22]: their decay with

increasing r follows exponential and power laws in the high- and low-temperature phases, respectively. Thus, the existence of two liquid phases must be manifested as a different variation of the correlation with distance between particles. In this paper, we present numerical results concerning the spatial asymptotics of pair correlation of interacting dust grains over a wide range of an "order" parameter and examine the experimental pair correlation functions obtained for liquid-like dust structures in the near-electrode plasma sheath of a high-frequency capacitive discharge.

2. NUMERICAL SIMULATION

2.1. Parameters of the Numerical Analysis

Correct simulation of plasma-dust particle transport must rely on a molecular-dynamics method, in which a system of ordinary differential equations containing a Langevin force F_{br} is solved. This force represents random impacts by molecules of the ambient gas or other random processes that underlie the relaxation of the kinetic temperature T of dust grains to the equilibrium value characterizing the energy of their stochastic motion [32–34]. Microscopic processes in homogeneous extended clouds of interacting macroscopic particles are simulated by setting periodic boundary conditions and taking into account not only the random force F_{br} responsible for thermal motion, but also the forces F_{int} acting between pairs of particles [17–21]:

$$m_p\frac{d^2\mathbf{l}_k}{dt^2} = \sum_j F_{int}(l)\big|_{\mathbf{l}=|\mathbf{l}_k-\mathbf{l}_j|}\frac{\mathbf{l}_k-\mathbf{l}_j}{|\mathbf{l}_k-\mathbf{l}_j|} - m_p\nu_{fr}\frac{d\mathbf{l}_k}{dt} + \mathbf{F}_{br}, \qquad (3)$$

$$F_{int}(l) = -eZ_p\frac{\partial\varphi}{\partial l},$$

where $\mathbf{l} = |\mathbf{l}_k - \mathbf{l}_j|$ is the separation between particles, m_p is the particle mass, ν_{fr} is the friction coefficient associated with collisions between dust particles and ambient-gas neutrals [35, 36], and φ is the potential of interaction between particles (the interaction energy is $U(r) \sim eZ_p\varphi(r)$). Computations were performed for the Yukawa potential with $\kappa = 2.4$ and 4.8. Correct simulation of molecular dynamics was ensured by using discretization cells of size $R \gg \lambda$ [37]. In our computations, $R \approx 5n_p^{-1/3} > (12–24)\lambda$. Additional computations were performed for the following combinations of power and exponential laws frequently used to model repulsion in kinetics of interacting particles [34, 38, 39]:

$$\varphi = \varphi_c b(r_p/r)^n, \qquad (4)$$

$$\varphi = \varphi_c\{a\exp(-\kappa_1 r/r_p) + b\exp(-\kappa_2 r/r_p)\}, \qquad (5)$$

$$\varphi = \varphi_c\{a\exp(-\kappa_1 r/r_p) + b(r_p/r)^n\}, \qquad (6)$$

where a, b, κ_1, and κ_2, and n are parameters. Both potential (2) and models (5) and (6) (with $n = 1$ in the latter) are of special interest in the physics of dusty plasmas. They can be used to allow for weaker screening at relatively large distances between dust grains [14–16]. These models have also been applied to describe repulsion between atoms in covalent metals [38, 39] or polymers [34, 38].

To analyze pair correlations in systems of particles with isotropic interaction potentials (4)–(6), three-dimensional equations of motion (3) were solved for specific values of parameters defined by analogy with those characterizing Yukawa systems: the effective order parameter

$$\Gamma^* = (Z_p^* e)^2/Tr_p, \qquad (7)$$

and the scaling parameter

$$\xi = \omega^*/\nu_{fr}, \qquad (8)$$

where the frequency of collisions between macroscopic particles is calculated as

$$\omega^* = eZ_p^*(n_p/\pi m_p)^{1/2} \qquad (9)$$

and the effective particle charge is

$$Z_p^* e = \{Z_p e\varphi''/2n_p\}^{1/2}. \qquad (10)$$

Note that the effective particle charge does not have any particular physical meaning here. However, the use of (10) makes it possible to retain Γ^*, ξ, and ω^* as universally applicable parameters in models with interaction potentials of any type.

The computations were performed for 125 independent particles in the central cell, while the number of particles taken into account in computing binary interactions reached approximately 3000. The binary interaction potential was cut off at the distance $L_{cut} = 4l_p$. To ensure that numerical results are independent of the number of particles and cut-off distance, we performed additional computations for 512 actual particles for $L_{cut} = 7l_p$ with $\Gamma^* = 1.5$, 17.5, 25, 49, and 92. A detailed description of the numerical procedure can be found in [18, 21]. The value of ξ was varied between 0.04 and 3.6, i.e., within the limits characteristic of the experimental conditions in gas-discharge plasmas. The value of Γ^* was varied between 1 and 110.

Our computations showed that the effective parameter Γ^* completely characterizes the ordering and phase states of the simulated particle systems if the

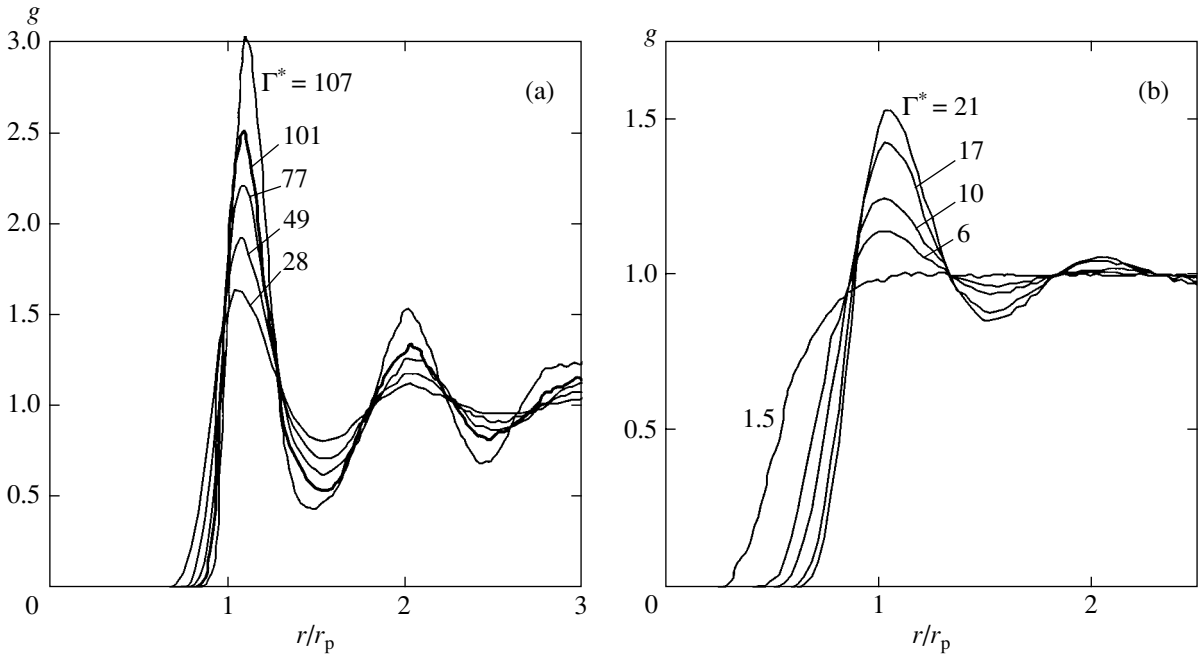

Fig. 1. Pair correlation functions $g(r/r_p)$ for Yukawa systems ($\kappa < 6$) for several values of Γ^* indicated at the curves.

following empirical condition for long-range interaction is satisfied:

$$2\pi|\varphi'(r_p)| > |\varphi''(r_p)|r_p. \quad (11)$$

In the first (linear) approximation, this criterion specifies conditions under which the force acting between two particles separated by the mean interparticle distance is greater than the force typically arising in collisions of macroscopic particles.

2.2. Ordering in Dissipative Dust-Particle Systems with Various Isotropic Repulsive Potentials

Ordering in the simulated systems was analyzed by using the pair correlation function $g(r)$ and the structure factor $S(q)$. Figure 1 shows the pair correlation functions obtained for Yukawa systems in a wide range of Γ^*. Figure 2 compares these functions with the functions $g(r)$ computed using various potentials subject to empirical condition (11) for two values of Γ^* and two values of ξ. Figures 3 and 4 illustrate the dependence of the first maxima g_1 and S_1 of the functions $g(r)$ and $S(q)$ and the corresponding distances $r = d_{g1}$ and $q = d_{S1}$ on Γ^*. Here, vertical bars represent the absolute deviations of these quantities for $\xi = 0.04$–3.6 and for various potentials satisfying (11). To compare the pair correlations computed for dissipative systems ($\nu_{fr} \neq 0$) with solutions to reversible equations of motion for nondissipative Yukawa systems ($\nu_{fr} = 0$) and with results obtained for a one-component plasma model, Fig. 3 also shows the maximum values of $g(r)$ and $S(q)$ found in [10, 40].

Our numerical study shows that spatial correlation of dust grains in the simulated systems is independent of friction (ν_{fr}) and is determined by the value of Γ^* under conditions ranging from a gaseous state ($\Gamma^* \sim 1$) to the point of crystallization into a BCC lattice ($\Gamma^* \approx$

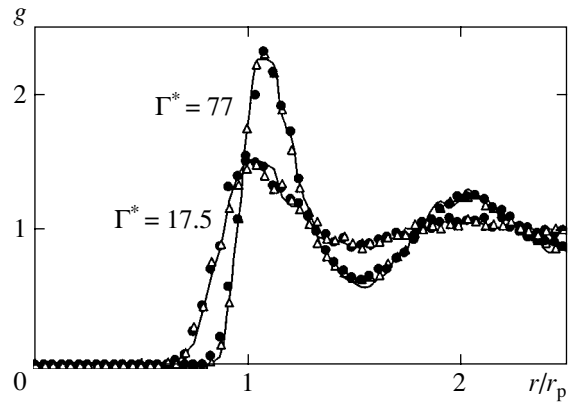

Fig. 2. Comparison of $g(r/r_p)$ for several model potentials and several values of ξ and Γ^*. For $\Gamma^* = 77$, the solid curve, triangles, and circles correspond to $\xi = 0.14$ and $\varphi/\varphi_c = \exp(-4.8r/r_p)$, $\xi = 0.14$ and $\varphi/\varphi_c = 0.1\exp(-2.4r/r_p) + \exp(-4.8r/r_p)$, and $\xi = 1.22$ and $\varphi/\varphi_c = \exp(-4.8r/r_p) + 0.05r_p/r$, respectively. For $\Gamma^* = 17.5$, the solid curve, triangles, and circles correspond to $\xi = 1.22$ and $\varphi/\varphi_c = \exp(-2.4r/r_p)$, $\xi = 1.22$ and $\varphi/\varphi_c = 0.1\exp(-2.4r/r_p) + \exp(-4.8r/r_p)$, and $\xi = 0.14$ and $\varphi/\varphi_c = 0.05(r_p/r)^3$, respectively.

102–104). Indeed, the first maximum points of the functions g and S corresponding to crystalline structure are characteristic of the BCC lattice: $d_{q1} \approx r_1 = (3\sqrt{3}/4n_p)^{1/3}$, $d_{S1} \approx q_1 = 2\pi(\sqrt{2}\,n_p)^{1/3}$, and $k_p = 2\pi(n_p)^{1/3}$ (see Fig. 4). Jumps in the values of the first maxima of $g(r)$ and $S(q)$ from 2.65 to 3.1 are observed as the normalized order parameter Γ* varies from the crystallization point $\Gamma^*_{jump} \approx$ 102–104 to the melting point $\Gamma^*_m \approx$ 106–107 (see Fig. 3). Thus, $\Gamma^*_{jumpm} \approx 104.5(\pm 2)$ can be interpreted as the point of phase transition between a liquid-like state and a BCC lattice.

Since $\Gamma^*_{jumpm} \approx 104.5$ (±2%) is independent of the ambient viscosity, this result is consistent with molecular-dynamics simulations of crystallization in Yukawa systems with zero friction [40–42]. The deviations of their results from $\Gamma^*_{jumpm} \approx 104.5$ vary within ±5% and can be attributed to difference in numerical procedures (number of particles, integration step, etc.) and to choice of Γ* associated with either melting or crystallization point of the system. It should be noted that $\Gamma^*_{jumpm} \approx 104.5(\pm 2\%)$ agrees with the theoretical results obtained in [43], where the value of the order parameter on the phase-transition line in the BCC model was 105(±3%). (The latter value is consistent with numerical results based on various criteria for crystallization [44] and melting [45].)

Note also that the form of a correlation function $g(r)$ satisfying condition (11) is determined by the value of Γ*. Therefore, the methods for determining the potential of interaction between particles from measurements of the structure factor based on the hypernetted chain approximation (using direct relations between $g(r)$, $S(q)$, and $\varphi(r)$ [8–10]) cannot be applied to the systems in question. Furthermore, the result obtained here can explain the widespread use of various phenomenological melting and crystallization criteria specifying the maximum values of correlation functions or the ratios of their maximum and minimum values on phase-transition lines (when $r \neq 0$) irrespective of the interparticle interaction potential. In particular, one can use the similarity of pair correlation functions (see Section 2.3) and the numerical values of their maxima g_1 (see Fig. 3) to obtain the well-known ratio of g_1 to the first minimum of $g(r)$, equal to 5, on the crystallization line.

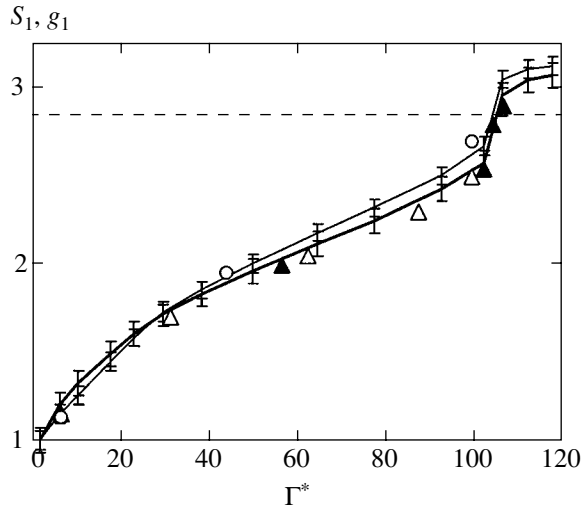

Fig. 3. First maxima of structure factor S_1 (thin curve) and pair correlation function g_1 (thick curve) vs. Γ*: closed triangles represent g_1 in the nondissipative Yukawa model ($\nu_{fr} = 0$) [40]; open triangles, g_1 in the BCC model [10]; open circles, S_1 in the BCC model [10]. Vertical bars are the absolute deviations for $\xi = 0.04$–3.6 and for various potentials satisfying (11).

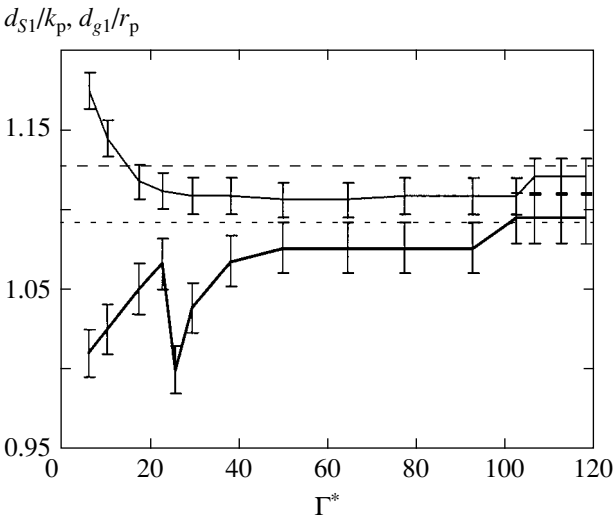

Fig. 4. Relative locations of the maximum of S_1, d_{S1}/k_p (thin curve), and the maximum of g_1, d_{g1}/r_p (thick curve), versus Γ*. Dashed curves represent the maximum points of the correlation function for BCC lattice. Vertical bars are the absolute deviations for $\xi = 0.04$–3.6 and for various potentials satisfying (11).

2.3. Pair Correlation in Liquid-Like Particle Systems and Dust Cluster Formation

To analyze the asymptotic decay of pair correlation with increasing distance between particles, we normalized the correlation function $h(r) = g(r) - 1$ to $h_1 = \max(h(r))$. Figure 5 shows the results obtained for several values of Γ*. An analysis of numerical results shows that $h(r)$ has pronounced maxima when Γ* > 3–5. For Γ* between 28 and 102, the behavior of $h(r)/h_1$ is determined by the value of Γ* (see Section 2.2). Its decay with increasing distance follows an asymptotic power law. At $r > r_p$, it can be approximated by the function (see Fig. 5a)

$$h(r/r_p)/h_1 \approx (\beta r_p/r)^{2.75}\sin(2\pi\beta(r/r_p + \beta - 1)), \quad (12)$$

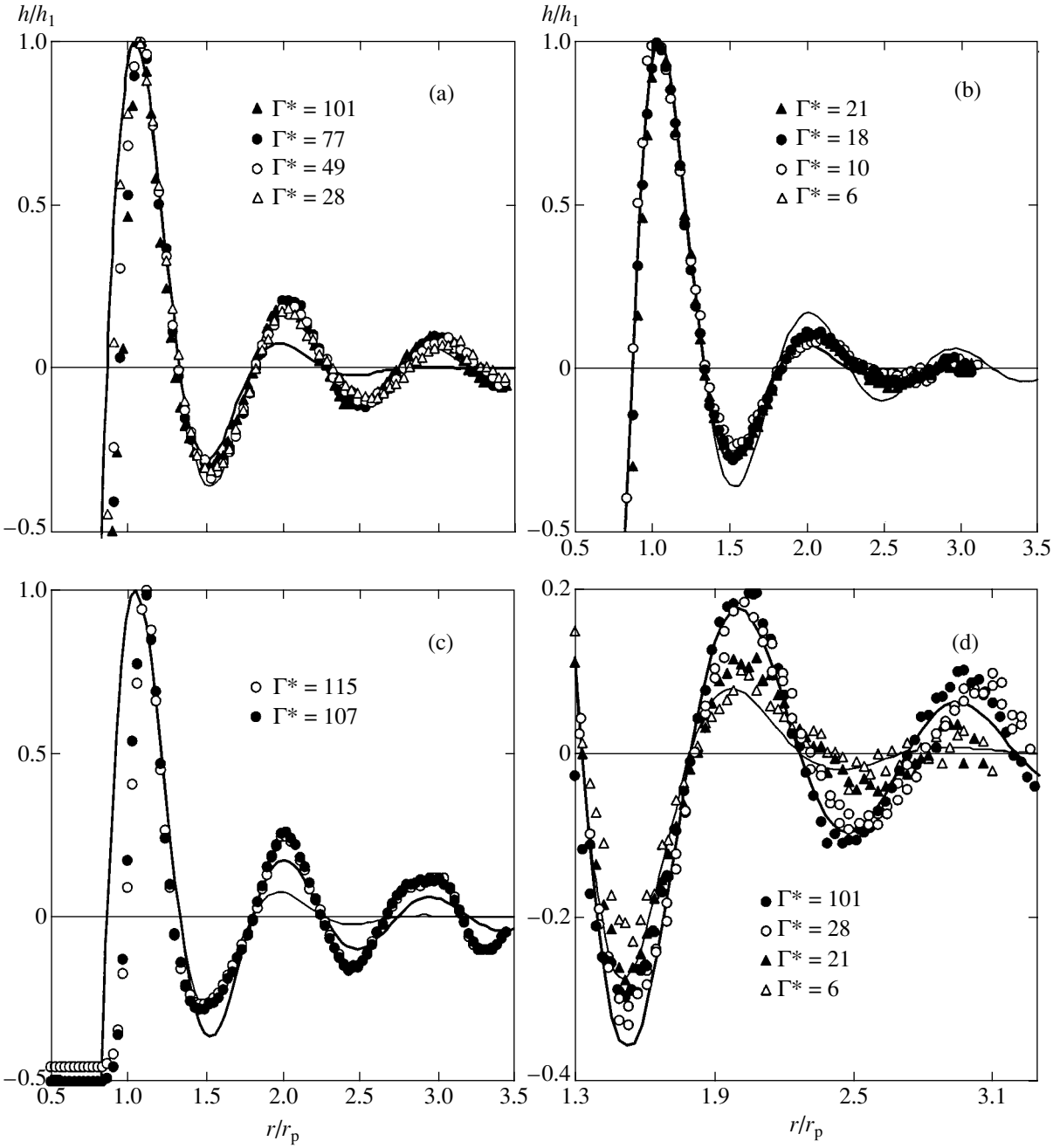

Fig. 5. Normalized pair correlation functions $h(r/r_p)/h_1$ for Yukawa systems with Γ* indicated in the panels, approximations (12) and (13) (thick and thin curves, respectively), and (d) an enlarged fragment comparing the numerical results obtained for several values of Γ*.

where $\beta \approx 1.07$. When the effective order parameter is smaller than $\Gamma^* \sim 21$, the decrease in pair correlation is much steeper and can be approximated by an exponential (see Fig. 5b):

$$h(r/r_p)/h_1 \approx \exp\{2.75(\beta - r/r_p)\} \times \sin(2\pi\beta(r/r_p + \beta - 1)). \quad (13)$$

Note that approximation of the pair correlation function is a difficult task, because the high-frequency thermal fluctuations due to stochastic motion of dust grains affect the measured value of $h(r)/h_1$, particularly when r is large and the fluctuation-induced error is comparable to $h(r)/h_1$ (see Fig. 5d). In the present case, one can hardly use a more suitable function to approximate numerical data for practical applications, because the dust-subsystem correlation functions measured in actual experiments exhibit stochastic variations due both to fluctuations in the ambient plasma and to instru-

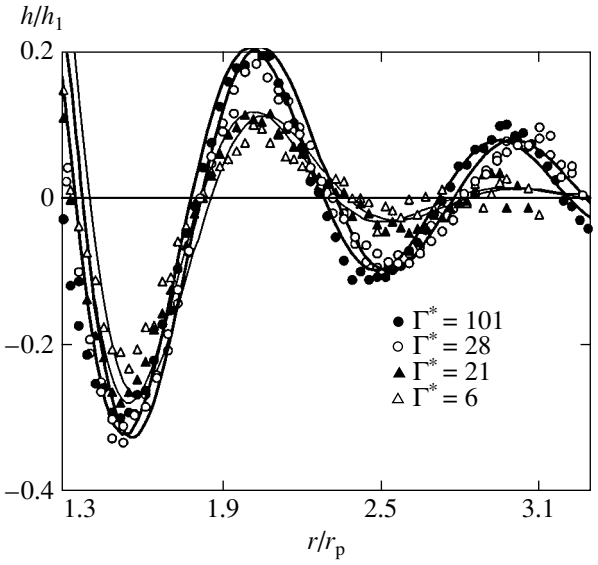

Fig. 6. Functions $h(r/r_p)/h_1$ for Yukawa systems with Γ^* indicated in the panels and approximations (14) and (15) (thick and thin curves, respectively). Shift between curves corresponding to different Γ^* is due to change in the location of the maximum of g_1 (see Fig. 4).

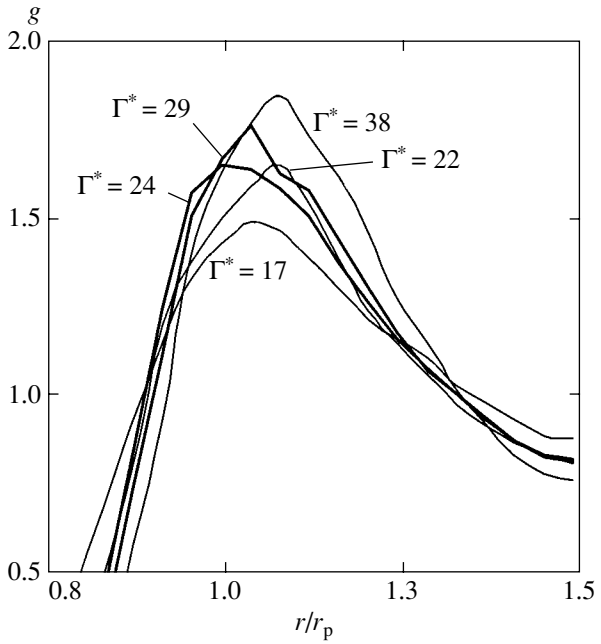

Fig. 7. Maxima of pair correlation functions for $\Gamma^* \approx 17$–40. Thick curves correspond to $\Gamma^* = 24$ and 29.

mentation noise. Nevertheless, substantial qualitative difference in asymptotic behavior of pair correlations is found between weakly ($\Gamma^* \le 21$) and strongly ($\Gamma^* \ge 28$) nonideal systems. When $\Gamma^* \ge 28$, the decay of the pair correlation with increasing distance follows power law (12). When the effective parameter is smaller than a certain critical value $\Gamma^* \sim 21$, the decrease in pair correlation can be approximated by exponential law (13). Note that this result disagrees with experimental observations reported in [30], where $g(r)$ was found to decay exponentially for both hexatic and isotropic liquid phases in a monolayer, and a power-law approximation was obtained only for a crystalline phase. This disagreement can be explained by the essentially two-dimensional structure of the monolayer (distinct from the three-dimensional system analyzed here). The behavior of the correlation function predicted for a crystalline phase by simulating three-dimensional systems of interacting particles is illustrated by Fig. 5c. It is obvious that the decay of pair correlation is much steeper here, as compared to that characteristic of liquid-like systems.

Improvement of the accuracy of approximations of $g(r)$ by numerical fitting would not be physically justified. However, such approximations may facilitate computations of thermodynamic characteristics of liquid-like systems determined by the pair correlation function $g(r)$, such as pressure, energy density, and compressibility (see [9, 10]). A detailed analysis of approximating functions for $g(r)$ curves at $0 \le r \le r_p$ was presented in [10]. The following functions can be sug-

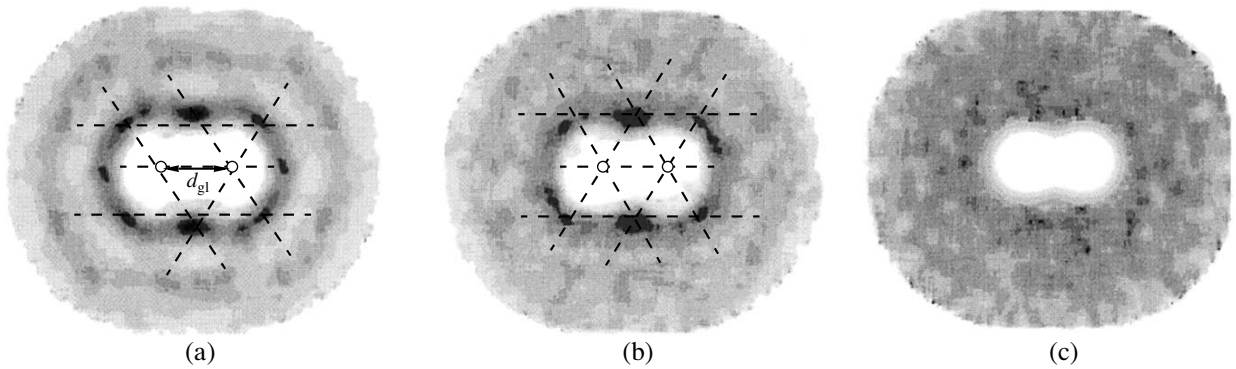

Fig. 8. Slices of three-particle correlation functions g_3 obtained by numerical simulation for $\Gamma^* = 37.5$ (a), 17.5 (b), and 1.5 (c).

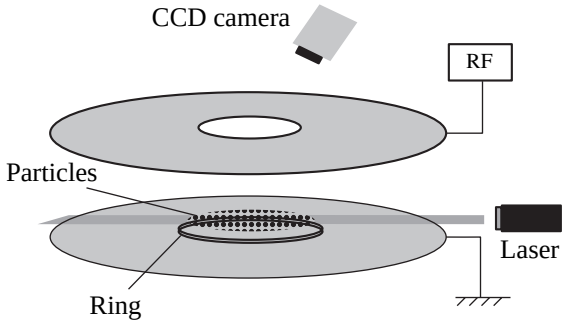

Fig. 9. Experiment in high-frequency capacitive discharge.

gested as approximations of $h(r)$ at $r \geq r_p$ (see Fig. 6):

$$h(r/r_p)/h_1 \approx (\beta r_p/r)^{2.5} \times (\sin(2\pi d_{g1}(r/r_p + \beta - 1)) + 0.1)/g_1, \quad 28 \leq \Gamma^* \leq 102, \tag{14}$$

$$h(r/r_p)/h_1 \approx \exp\{2.25(\beta - r/r_p)\} \times (\sin(2\pi d_{g1}(r/r_p + \beta - 1)) + 0.1)/g_1, \quad 4 \leq \Gamma^* \leq 21. \tag{15}$$

These functions allow for the shift in $g(r)$ associated with change in the location of its maximum g_1 as a function of Γ^* (see Fig. 4).

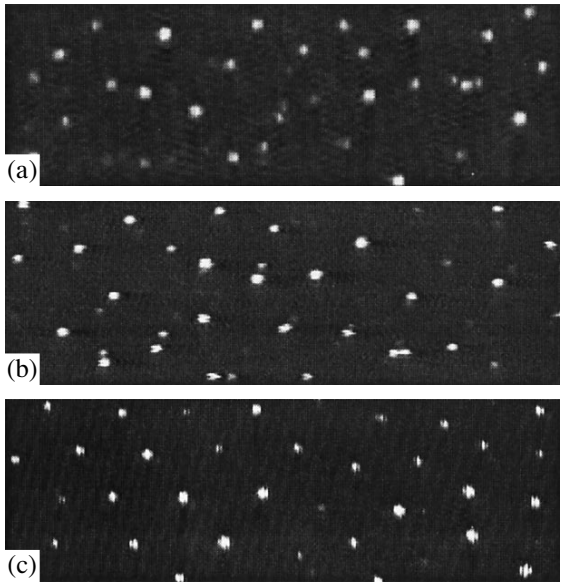

Fig. 10. Images of dust-cloud particles in the near-electrode plasma sheath of discharge for (a) P = 3 Pa and W = 10 W, (b) P = 3 Pa and W = 2 W, and (c) P = 7 Pa and W = 10 W.

The difference in the asymptotics of the pair correlation functions corresponding to different values of the

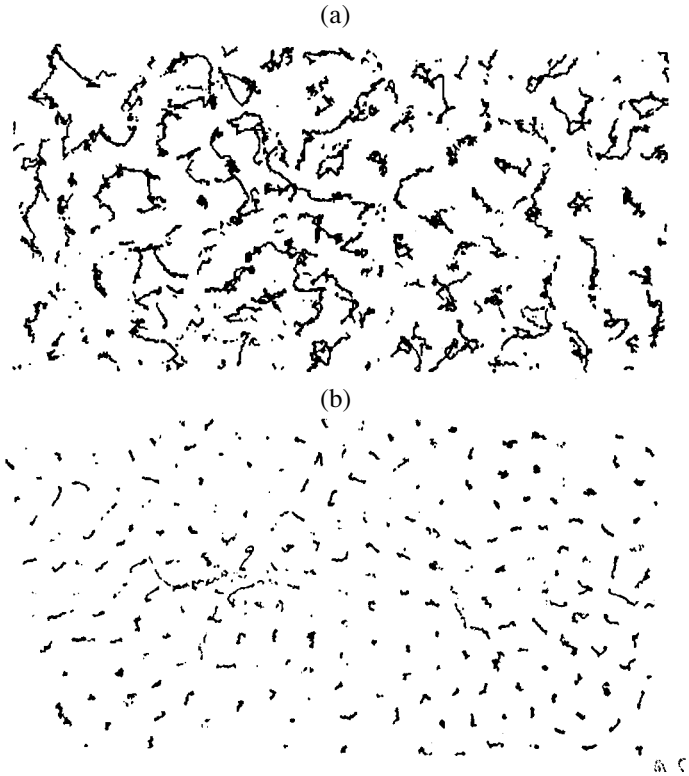

Fig. 11. Trajectories of dust grains over the averaging time interval for pair correlation functions for (a) P = 3 Pa and W = 10 W and (b) P = 7 Pa and W = 10 W.

order parameter can be compared with the results of numerical experiments reported in [17, 31]. In particular, formation of groups (clusters) of macroscopic particles was observed in [17] at $\Gamma^* \approx 22$–24. This phenomenon was accompanied by a sharp decrease in diffusivity D and a shift in the location of the first maximum of $g(r)$ (see Fig. 4), and the system's properties changed qualitatively at this point (at least, those analogous to properties of a solid). Figure 7 illustrates the behavior of the maximum of $g(r)$ at $\Gamma^* \approx 22$–24 for three-dimensional Yukawa systems. The critical value $\Gamma^* \approx 23.5$ corresponds to the condition of one particle per sphere of Wigner–Seitz radius

$$a_{WS} = (4\pi n_p/3)^{-1/3}$$

($g(r) = 0$ at $r < a_{WS}$), and the mean free path

$$r_{p-p} \approx (3T/\omega^{*2} m_p)^{1/2}$$

for particle–particle collisions is close to a_{WS}.

Subsequently, the formation of well-ordered clusters of macroscopic particles at $\Gamma^* > 22$–24 was revealed in numerical simulations and laboratory experiments by analyzing three-particle correlation functions [31]. Figure 8 shows slices of three-particle correlation functions $g_3(r_{12}, r_{23}, r_{31})$ ($r_{ij} = |\mathbf{r}_i - \mathbf{r}_j|$) at $r_{12} = d_{g1}$ computed for several values of Γ^*. To facilitate comparison, the slices are normalized to the maximum of $g_3(r_{12}, r_{23}, r_{31})$ (black and white areas correspond to $g_3 = 1$ and 0, respectively).

The formation of well-ordered dust clusters can be interpreted as the onset of orientational (short-range) ordering in the systems in question with increasing order parameter (see Fig. 8a). The observed qualitative changes can be attributed to the topological defects of three-dimensional crystalline structure responsible for the existence of two distinct (isotropic and orientationally ordered) liquid-like phases (by analogy with topological phase transitions in two-dimensional systems). Therefore, with increasing Γ^*, the simulated system should exhibit a solid-like dynamical behavior that can be described by the jump model developed for molecular liquids [6]. In this model, a liquid-phase molecule resides in an equilibrium position (site) until its energy becomes sufficiently large for the molecule to break free from the potential bonding with adjacent molecules and jump into a new site surrounded by different molecules. Agreement of the dynamics of Yukawa systems with predictions of this model at $\Gamma^* > 40$–50 is supported by numerical results [17]. To elucidate the nature of the observed effects, quantitative characteristics of orientational ordering (mean time of specific orientation of macroscopic particles, topological entropy, rates of variation of autocorrelation functions, etc.) must be determined as functions of the order parameter.

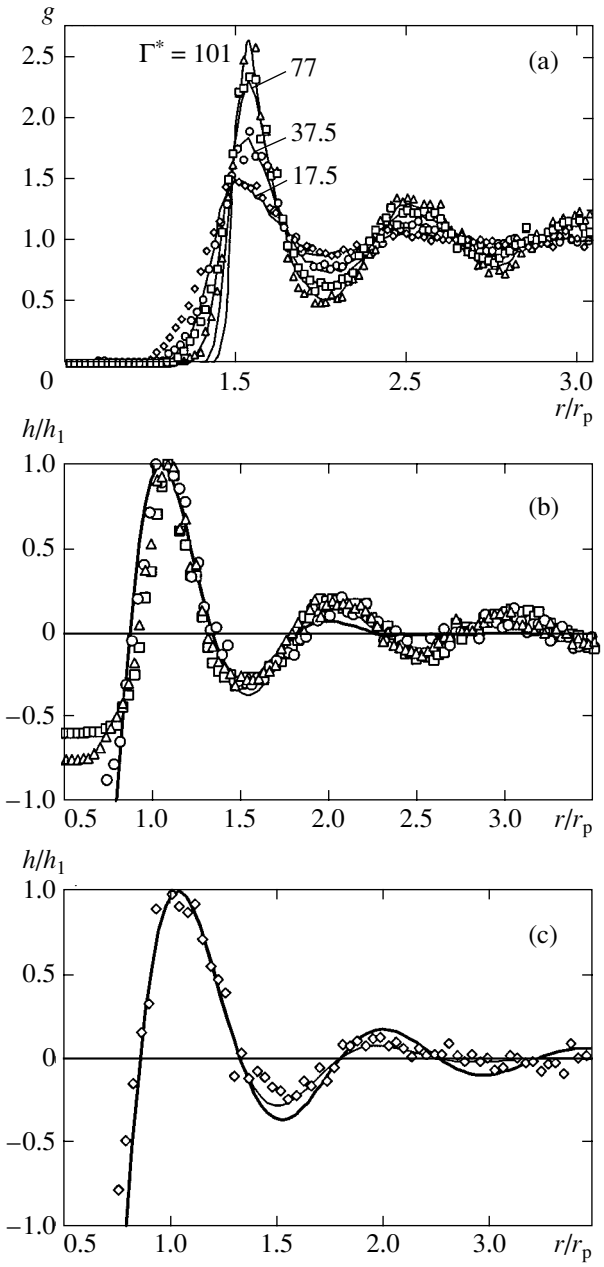

Fig. 12. Measured $g(r/r_p)$ (a) and $h(r/r_p)/h_1$ (b, c) for P = 3 Pa and W = 10 W (diamonds), P = 3 Pa and W = 2 W (circles), and P = 7 Pa and W = 10 W (triangles). Solid curves are (a) $g(r/r_p)$ computed for several Γ^* and (b, c) approximations (12) and (13) (thick and thin curves, respectively).

3. EXPERIMENT

The setup used to study pair correlations of dust particles in the near-electrode plasma sheath of a high-frequency capacitive discharge is schematized in Fig. 9.

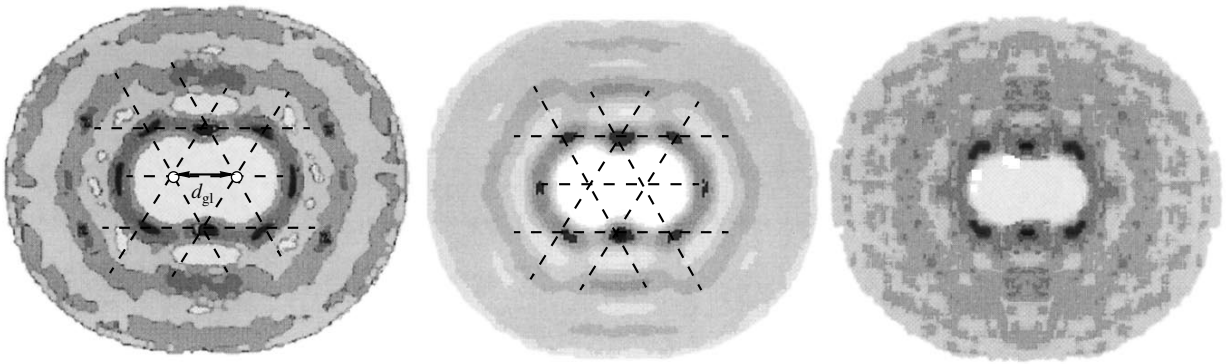

Fig. 13. Slices of three-particle correlation functions g_3 obtained in experiments for (a) P = 7 Pa and W = 10 W, (b) P = 3 Pa and W = 2 W, and (c) P = 3 Pa and W = 10 W.

The experiments were conducted in argon at a pressure of P = 2–10 Pa and discharge power of $W \approx$ 2–10 W. Monodisperse melamine-formaldehyde particles of radius $a_p \approx 1.7$ μm and density $\rho_p \approx 1.5$ g cm^{-3} were used as the dust component. Four to ten layers of dust particles were observed. The dust cloud was sliced with a 200–300 μm thick He–Ne laser sheet and its images were recorded with a CCD camera at a framing speed of 25 fps. Figure 10 shows fragments of dust-cloud images obtained under different experimental conditions.

The images were processed by means of a special computer program that determined the locations and displacements of individual particles in dust structures. In all cases under analysis, we observed quasi-steady liquid-like structures. The mean distance r_p between dust grains in these structures varied from 260 to 350 μm. Since the structures under analysis consisted of several layers of macroscopic particles, no large-scale vortices analogous to those observed in [24, 25] were observed. Figure 11 shows the trajectories of grains recorded during the averaging time interval for pair correlation functions (between 1 and 2 s).

The images were processed to obtain correlation functions $g(r)$ and $g_3(r_{12}, r_{23}, r_{31})$ averaged over 1 to 2 s under constant experimental conditions. Figure 12a shows $g(r/r_p)$ obtained for several values of P and W. Figure 12b compares measured functions $h(r)/h_1$ with approximations (12) and (13) found in numerical experiments. It is clear that the measured pair correlation functions agree with the functions $g(r)$ obtained for the simulated systems with $\Gamma^* \approx 17.5$ and $\Gamma^* > 37$, even though the actual experimental conditions differed from those set in the computed homogeneous problem. Note that the decay of spatial correlation of macroscopic particles follows a power law when $\Gamma^* > 37$ and is approximately exponential when $\Gamma^* \approx 17.5$.

Despite the difference between parameters of the three-dimensional homogeneous problem and the experimental conditions, the measured three-particle correlation functions are also consistent with those obtained by numerical simulation (see Fig. 8). Figure 13 shows slices of measured three-particle correlation functions $g_3(r_{12}, r_{23}, r_{31})$ for $r_{12} = d_{g1}$ (most probable interparticle distance determined as the point of maximum for measured $g(r)$). A comparison of the results presented here for several values of discharge parameters reveals a short-range orientational ordering in the dust structures, which is manifested by the maxima of $g_3(r_{12}, r_{23}, r_{31})$ at the vertices of the hexagonal clusters depicted by dashed lines in Figs. 13a and 13b. As the maximum of the pair correlation function increases, these maxima (separated by a distance r nearly equal to d_{g1}) grow and additional maxima, separated by a distance $r \approx 2d_{g1}$, appear (see Fig. 13a).

Thus, our analysis of the spatial correlation of macroscopic particles suggests that the behavior of pair correlation functions in experimentally observed dust structures should be independent of the binary interaction potential. At $\Gamma^* \approx$ 22–24, the dust subsystem appears to exhibit a transformation into an isotropic liquid analogous to the topological phase transition in two-dimensional systems. This conjecture can be supported by experimental studies of boundary angular autocorrelation and spatial correlation functions ($g_6(t)$, $g_6(r)$) and by measurements of topological entropy.

4. CONCLUSIONS

A numerical analysis of correlation functions is performed for extended three-dimensional many-particle systems with various binary interaction potentials. It is shown that the form of the pair correlation function is determined by the value of Γ^* for systems characterized by various repulsive potentials. This implies that the form and parameters of the potentials cannot be determined by inverting $g(r)$. Our results show that pair correlation of particles with interaction potentials con-

sidered in this study can be described (and, therefore, certain thermodynamic characteristics of liquid-like systems can be determined) without performing the complicated calculations required to find spatial closure approximations. For a system of this kind, the pair correlation function can be approximated by a function depending on two dimensionless parameters (Γ^* and r/r_p). Furthermore, substantially different forms of this function at $r/r_p > 1$ will be obtained only for weakly correlated and strongly nonideal structures (when $\Gamma^* < 22$ and $\Gamma^* > 28$, respectively).

The behavior of $h(r)/h_1 = (g(r) - 1)/(g_1 - 1)$ for liquid-like systems at $\Gamma^* \geq 28$ and $\Gamma^* \leq 22$ weakly depends on the order parameter if $r \geq r_p$, in which case $h(r)/h_1$ can be approximated by the product of a harmonic function with a function describing the decay of spatial correlation. In strongly correlated systems (at $\Gamma^* \geq 28$), the asymptotic decay of pair correlation follows a power law. When the effective order parameter is smaller than a certain critical $\Gamma^* \approx 22$, the decay of the correlation function can be described by an exponential law. The pair correlation of macroscopic particles measured in liquid-like dust structures localized in the near-electrode plasma sheath of a high-frequency capacitive discharge is in good agreement with numerical results. The change in the decay of the pair correlation function observed in both numerical and experimental studies can be attributed to the existence of two distinct liquid phases: an isotropic phase and a phase characterized by short-range orientational ordering. This conjecture is supported by observations of well-ordered clusters of dust grains.

ACKNOWLEDGMENTS

This work was supported, in part, by the Russian Foundation for Basic Research, project no. 01-02-16658, and by INTAS, grant no. 01-0391. We thank A.V. Chernyshev, A.V. Gavrikov, and I.A. Shakhova for providing experimental images.

REFERENCES

1. J. H. Chu and Lin I, Phys. Rev. Lett. **72**, 4009 (1994).
2. H. Thomas, G. Morfill, V. Demmel, *et al.*, Phys. Rev. Lett. **73**, 652 (1994).
3. A. Melzer, T. Trottenberg, and A. Piel, Phys. Lett. A **191**, 301 (1994).
4. V. E. Fortov, A. P. Nefedov, V. M. Torchinskiĭ, *et al.*, Pis'ma Zh. Éksp. Teor. Fiz. **64**, 86 (1996) [JETP Lett. **64**, 92 (1996)].
5. A. M. Lipaev, V. I. Molotkov, A. P. Nefedov, *et al.*, Zh. Éksp. Teor. Fiz. **112**, 2030 (1997) [JETP **85**, 1110 (1997)].
6. Ya. I. Frenkel', *Kinetic Theory of Liquid* (Nauka, Leningrad, 1975; Oxford Univ. Press, Oxford, 1976).
7. E. M. Lifshitz and L. P. Pitaevskiĭ, *Course of Theoretical Physics*, Vol. 5: *Statistical Physics* (Nauka, Moscow, 1979; Pergamon, New York, 1980).
8. R. Balescu, *Equilibrium and Nonequilibrium Statistical Mechanics* (Wiley, New York, 1975; Mir, Moscow, 1978).
9. N. K. Airawadi, Phys. Rep. **57**, 241 (1980).
10. S. Ichimaru, Rev. Mod. Phys. **54**, 1017 (1982).
11. H. J. Raverche and R. D. Mountain, J. Chem. Phys. **57**, 3987 (1972).
12. H. J. Raverche and R. D. Mountain, J. Chem. Phys. **57**, 4999 (1972).
13. S. Wang and J. A. Krumhansl, J. Chem. Phys. **56**, 4287 (1972).
14. U. Konopka, G. E. Morfill, and R. Ratke, Phys. Rev. Lett. **84**, 891 (2000).
15. J. E. Daugherty, R. K. Porteous, M. D. Kilgore, *et al.*, J. Appl. Phys. **72**, 3934 (1992).
16. J. E. Allen, Phys. Scr. **45**, 497 (1992).
17. O. S. Vaulina and S. V. Vladimirov, Phys. Plasmas **9**, 835 (2002).
18. O. S. Vaulina, S. V. Vladimirov, O. F. Petrov, *et al.*, Phys. Rev. Lett. **88**, 245002 (2002).
19. V. E. Fortov, O. S. Vaulina, O. F. Petrov, *et al.*, Phys. Rev. Lett. **90**, 245005 (2003).
20. O. S. Vaulina, O. F. Petrov, V. E. Fortov, *et al.*, Fiz. Plazmy **29**, 698 (2003) [Plasma Phys. Rep. **29**, 642 (2003)].
21. V. E. Fortov, O. S. Vaulina, O. F. Petrov, *et al.*, Zh. Éksp. Teor. Fiz. **123**, 798 (2003) [JETP **96**, 704 (2003)].
22. J. M. Kosterlitz and D. J. Thouless, J. Phys. C **6**, 1181 (1973).
23. H. M. Thomas and G. E. Morfill, Nature **379**, 806 (1996).
24. G. E. Morfill, H. M. Thomas, and M. Zuzic, in *Advances in Dusty Plasma*, Ed. by P. K. Shukla, D. A. Mendis, and T. Desai (World Sci., Singapore, 1997), p. 99.
25. Lin I, Wen-Tan Juan, C. H. Chiang, *et al.*, in *Advances in Dusty Plasma*, Ed. by P. K. Shukla, D. A. Mendis, and T. Desai (World Sci., Singapore, 1997), p. 143.
26. Wen-Tan Juan, M. H. Chen, and Lin I, Phys. Rev. E **64**, 016402 (2001).
27. K. Bagchi, Y. C. Andersen, and W. Swope, Phys. Rev. Lett. **76**, 255 (1996).
28. A. Jaster, Europhys. Lett. **42**, 277 (1998).
29. K. Zahn and G. Maret, Phys. Rev. Lett. **85**, 3656 (2000).
30. A. H. Marcus and S. A. Rice, Phys. Rev. Lett. **77**, 2577 (1996).
31. O. S. Vaulina, O. F. Petrov, V. E. Fortov, *et al.*, Phys. Rev. Lett. (in press).
32. V. V. Zhakhovskiĭ, V. I. Molotkov, A. P. Nefedov, *et al.*, Pis'ma Zh. Éksp. Teor. Fiz. **66**, 392 (1997) [JETP Lett. **66**, 419 (1997)].
33. O. S. Vaulina, A. P. Nefedov, O. F. Petrov, *et al.*, Zh. Éksp. Teor. Fiz. **118**, 1319 (2000) [JETP **91**, 1147 (2000)].

34. A. A. Ovchinnikov, S. F. Timashev, and A. A. Belyĭ, *Kinetics of Diffusely Controlled Processes* (Khimiya, Moscow, 1986).

35. E. M. Lifshitz and L. P. Pitaevskiĭ, *Physical Kinetics* (Nauka, Moscow, 1979; Pergamon Press, Oxford, 1981).

36. N. A. Fuchs, *The Mechanics of Aerosols* (Dover, New York, 1964).

37. R. T. Farouki and S. Hamaguchi, Phys. Lett. **61**, 2973 (1992).

38. *Molecular Interactions from Diatomic Molecules to Biopolymers*, Ed. by B. Pulman (Mir, Moscow, 1981).

39. I. M. Torrens, *Interatomic Potentials* (Academic, New York, 1972).

40. M. O. Robbins, K. Kremer, and G. S. Grest, J. Chem. Phys. **88**, 3286 (1988).

41. E. J. Meijer and D. Frenkel, J. Chem. Phys. **94**, 2269 (1991).

42. S. Hamaguchi, R. T. Farouki, and D. H. E. Dubin, Phys. Rev. E **56**, 4671 (1997).

43. W. L. Slattery, G. D. Doollen, and H. E. DeWitt, Phys. Rev. A **21**, 2087 (1980).

44. E. L. Plollock and J. P. Hansen, Phys. Rev. A **8**, 3110 (1973).

45. H. M. van Horn, Phys. Lett. A **28A**, 707 (1969).

Translated by A. Betev

Journal of Experimental and Theoretical Physics, Vol. 98, No. 3, 2004, pp. 527–537.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 125, No. 3, 2004, pp. 598–608.

PLASMA, GASES

Electromagnetic Radiation at Twice the Plasma Frequency Emitted from the Region of Interaction of Two Short Laser Pulses in a Rarefied Plasma

L. M. Gorbunov[a] and A. A. Frolov[b]

[a]*Lebedev Physical Institute, Russian Academy of Sciences, Leninskiĭ pr. 53, Moscow, 119991 Russia*
[b]*Institute for High Energy Densities of Associated Institute for High Temperatures, Russian Academy of Sciences, Izhorskaya ul. 13/19, Moscow, 125412 Russia*
e-mail: frolov@ihed.ras.ru
Received September 12, 2003

Abstract—We study the electromagnetic radiation at twice the plasma frequency, which emerges because of the interaction of two identical counterpropagating short laser pulses in a rarefied plasma and caused by excitation of small-scale standing plasma waves in the pulse overlap region. The energy, spectral, and angular characteristics of radiation are investigated, and the dependence of these characteristics on the parameters of the laser pulses is analyzed. The possibility of applying this effect for diagnostics of localized plasma oscillations is discussed. *© 2004 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

In our previous publication [1], we considered the possibility of "impact" excitation of localized coherent small-scale plasma oscillations in a rarefied plasma by counterpropagating short laser pulses. These oscillations could become an attractive object for studying both laser pulses and the processes determining the evolution of plasma oscillations (destruction of oscillation coherence, wave breaking, and the development of Langmuir turbulence). Here, we consider a possible method for diagnostics of localized plasma waves from radiation produced by these waves in the vicinity of a twofold plasma frequency $2\omega_p$.

Radiation at a frequency of $2\omega_p$ has been widely discussed starting from 1950s in connection with solar flares [2]. A possible physical mechanism responsible for this radiation, which is now commonly referred to as coalescence of two plasmons, was mentioned for the first time in [3]. Since the end of the 1960s, this effect has attracted the attention of researchers in connection with general problems of strong Langmuir turbulence (see, for example, [4]). Some of the observed singularities in the spectrum of radiation reflected from an inhomogeneous laser plasma in the region of the twofold laser frequency $2\omega_0$ were also attributed to two-plasmon fusion [5].

The effect considered here is also based on elementary nonlinear coalescence of two plasmons accompanied by the generation of photons. The actually observed characteristics of the radiation at twice the plasma frequency (such as intensity, directivity diagram, polarization, and linewidth) are determined by the superposition of the fields of transverse waves (photons) originating from a large number of elementary processes and depend on the properties of plasma waves (their spatial structure, damping, degree of coherence, and method of excitation).

In this paper, we consider the radiation at twice the plasma frequency resulting from localized small-scale standing plasma waves excited in a plasma by the interaction of two short laser pulses. It was shown in [1] that the collision of such pulses in the plasma gives rise to a short-lived standing electromagnetic wave in the region where these pulses overlap, producing ponderomotive forces, which vary periodically in space with a wave number of $2k_0$, where k_0 is the wave number of laser radiation. Under the action of these forces, small-scale coherent plasma perturbations are formed. The time evolution of these perturbations depends on the duration of the pulses. If the pulse duration τ is longer than the reciprocal plasma frequency ω_p^{-1}, these perturbations exist only during the time of interaction of the pulses and disappear as the pulses move apart. If the pulse duration is on the order of or smaller than the plasma period, such plasma oscillations persist even after termination of the interaction between the pulses in the form of a localized coherent plasma wave. It will be shown below that, as a result of such an impact excitation, the electromagnetic radiation at twice the plasma frequency can be emitted from the region of localization of plasma oscillations. The intensity of this radiation and its directivity diagram are essentially determined by the parameters of interacting pulses and carry information both on the pulses propagating in the plasma and on the Langmuir waves remaining in the region of their interaction.

1063-7761/04/9803-0527$26.00 © 2004 MAIK "Nauka/Interperiodica"

It should be noted that the effect of radiation at twice the plasma frequency from a localized standing Langmuir wave considered here is analogous in a certain sense to the effect of generation of second harmonic radiation from counterpropagating surface polaritons, which is known in solid-state physics [6].

2. EXCITATION OF PLASMA OSCILLATIONS DURING THE INTERACTION OF LASER PULSES

In our previous publication [1], we studied the excitation of small-scale plasma oscillations during the interaction of two identical short laser pulses in the hydrodynamic nondissipative approximation. The infinitely long lifetime of oscillations emerging in this model leads to a singularity in the spectral density of the energy radiated. To avoid this difficulty, we must take into account the damping of plasma oscillations, which is possible only in the framework of the kinetic theory. For this reason, in this section, as well as in the Appendix, we will develop a kinetic approach to describing the interaction of short laser pulses.

We consider two identical laser pulses propagating in a plasma towards each other along the z axis. We write the electric field $\mathbf{E}_L$ of the pulses in the form

$$\mathbf{E}_L(\mathbf{r}, t) = \frac{1}{2} \times \{\mathbf{E}_0(\mathbf{r}, t)\exp(-i\omega_0 t) + \mathbf{E}_0^*(\mathbf{r}, t)\exp(i\omega_0 t)\}. \quad (2.1)$$

Here, ω_0 is the laser frequency and $\mathbf{E}_0$ is the complex amplitude varying slowly with time on the ω_0^{-1} scale and exhibiting the coordinate dependence

$$\mathbf{E}_0(\mathbf{r}, t) = \mathbf{E}_+(\mathbf{r}, t)\exp(ik_0 z) + \mathbf{E}_-(\mathbf{r}, t)\exp(-ik_0 z), \quad (2.2)$$

where k_0 is the longitudinal wave number connected with frequency ω_0 via the dispersion relation $k_0 c = \sqrt{\omega_0^2 - \omega_p^2}$, $\omega_p = \sqrt{4\pi e^2 N_{0e}/m}$ being the plasma frequency, which is assumed to be smaller than ω_0; e, m, and N_{0e} are the charge, mass, and concentration of electrons in the plasma; and $\mathbf{E}_\pm(\mathbf{r}, t)$ are the amplitudes of laser pulses propagating from left to right (plus sign) and from right to left (minus sign), which change in space insignificantly over a scale of k_0^{-1}.

In the course of interaction, the pulses generate a standing electromagnetic wave associated with quasi-static small-scale ponderomotive forces. These forces induce periodic electron density perturbations (with a wave number of $2k_0$) and small-scale quasi-static electric fields in the plasma. Substituting formula (2.2) into definition (A.23) of the high-density potential and separating the terms proportional to $\exp(\pm 2ik_0 z)$, we obtain

$$\phi(\mathbf{r}, t) = \frac{e^2}{4m\omega_0^2}\{\exp(2ik_0 z)(\mathbf{E}_+ \cdot \mathbf{E}_-^*) + \exp(-2ik_0 z)(\mathbf{E}_- \cdot \mathbf{E}_-^*)\}. \quad (2.3)$$

We assume that the electric field amplitude of the laser pulses in the region of their interaction has an axially symmetric Gaussian form,

$$\mathbf{E}_+(\mathbf{r}, t) = \mathbf{e}_+ E_{0L}\exp\left(-\frac{\xi^2}{2L^2} - \frac{\rho^2}{2R^2}\right),$$
$$\mathbf{E}_-(\mathbf{r}, t) = \mathbf{e}_- E_{0L}\exp\left(-\frac{\eta^2}{2L^2} - \frac{\rho^2}{2R^2}\right), \quad (2.4)$$

where $\xi = z - V_g t$ and $\eta = z + V_g t$ are the longitudinal coordinates in the systems comoving with the pulses, $V_g = c^2 k_0/\omega_0$ is the group velocity of the pulses, L is the length of a pulse connected with its duration via the relation $\tau = L/V_g$, $\rho = \sqrt{x^2 + y^2}$ is the transverse coordinate, R is the pulse width, E_{0L} is the maximal value of the electric field amplitude, and $\mathbf{e}_+$ and $\mathbf{e}_-$ are the vectors determining the polarization of the laser pulses.

The Fourier component of the electric field of plasma perturbations can be expressed in terms of the corresponding component of the ponderomotive potential with the help of formula (A.27) (see Appendix). In accordance with formula (2.4), the latter component has the form

$$\phi(\omega, \mathbf{k}) = \frac{mV_E^2\pi^2 R^2 L\tau(\mathbf{e}_+ \cdot \mathbf{e}_-)}{4} \times \exp\left(-\frac{\omega^2\tau^2}{4} - \frac{k_\perp^2 R^2}{4}\right) \times \left[\exp\left(-\frac{(k_z - 2k_0)^2 L^2}{4}\right) + \exp\left(-\frac{(k_z + 2k_0)^2 L^2}{4}\right)\right], \quad (2.5)$$

where $V_E = eE_{0L}/m\omega_0$.

Bearing in mind that we are dealing with short laser pulses of duration τ on the order of ω_p^{-1}, we will study plasma oscillations remaining in the overlap region of the pulses after termination of their interaction. Since ponderomotive forces in this case are equal to zero, only free plasma oscillations can exist in the plasma;

the frequency ω and the wave vector $\mathbf{k}$ of these oscillations are connected via the dispersion relation [7]

$$\varepsilon^l(\omega, \mathbf{k}) = 1 + \delta\varepsilon^l(\omega, \mathbf{k}) = 0. \tag{2.6}$$

If the thermal velocity of electrons, $V_T = \sqrt{T/m}$ (T is the electron temperature), is smaller than the phase velocity (ω/k) of plasma waves for the Maxwellian distribution function F, expression (A.28) leads to the following relation [7]:

$$\delta\varepsilon^l(\varepsilon, \mathbf{k}) = -\frac{\omega_p^2}{\omega^2}\left(1 + 3\frac{k^2V_T^2}{\omega^2}\right) + i\sqrt{\frac{\pi}{2}}\frac{\omega\omega_p^2}{k^3V_T^3}\exp\left(-\frac{\omega^2}{2k^2V_T^2}\right). \tag{2.7}$$

Bearing in mind that the ratio of the thermal velocity of electrons to the phase velocity of a plasma wave is small (in our case, this corresponds to the inequality $2V_T/c < \omega_p/\omega_0$), we can represent the Fourier component (A.27) of plasma waves in the form

$$\langle \mathbf{E}_2(\omega, \mathbf{k})\rangle = -\frac{i\mathbf{k}}{2e}\phi(\omega_L, \mathbf{k})\omega_L \times\left(\frac{1}{\omega - \omega_L + i\gamma_L} - \frac{1}{\omega + \omega_L + i\gamma_L}\right), \tag{2.8}$$

where

$$\omega_L(k) = \omega_p\left(1 + \frac{3}{2}k^2r_D^2\right), \quad \gamma_L(k) = \sqrt{\frac{\pi}{8}}\frac{\omega_P}{k^3r_D^3}\exp\left(-\frac{3}{2} - \frac{1}{2k^2r_D^2}\right), \tag{2.9}$$

and $r_D = V_T/\omega_p$ is the Debye radius for electrons.

In contrast to the hydrodynamic result [1], expression (2.8) takes into account the thermal correction in the dispersion relation for Langmuir waves as well as Landau damping. Although both these effects are weak in the given approximation, consistent inclusion of damping plays a fundamental role, limiting the lifetime of excited plasma waves and removing the singularity in the frequency spectrum of the radiation at twice the plasma frequency.

Formula (2.8) combined with relation (2.5) defines the electric field of a small-scale standing plasma wave excited by the pulses in their overlap region. The lifetime of such a wave is characterized by decrement $\gamma_L(2k_0)$, and the range of wave vectors in the transverse and longitudinal directions is characterized, respectively, by the transverse and longitudinal sizes R and L of the pulses. Considering that the laser wavelength is small as compared to the pulse sizes, we will henceforth assume that $k = 2k_0$ in formulas (2.9).

3. ELECTRIC FIELD OF LOW-FREQUENCY RADIATION

The interaction between plasma waves induces a nonlinear current. In the approximation quadratic in the wave fields, the Fourier component of current density $\mathbf{j}_2$ in the general form can be written as [8]

$$j_i(\omega, \mathbf{k}) = -\frac{i\omega}{8\pi}\int\frac{d\omega' d\mathbf{k}'}{(2\pi)^4} \times S_{ijs}(\omega, \mathbf{k}, \omega', \mathbf{k}')E_j(\omega'', \mathbf{k}'')E_s(\omega', \mathbf{k}'), \tag{3.1}$$

where $\omega'' = \omega - \omega'$, $\mathbf{k}'' = \mathbf{k} - \mathbf{k}'$, and $\mathbf{E}(\omega, \mathbf{k}) = \langle \mathbf{E}_2(\omega, \mathbf{k})\rangle$. In the case we are interested in, when the phase velocities of interacting waves are larger than the thermal velocity of electrons ($\omega > kV_T$, $\omega' > k'V_T$, $\omega'' > k''V_T$), tensor $S_{ijs}(\omega, \mathbf{k}, \omega', \mathbf{k}')$ has the form [8]

$$S_{ijs}(\omega, \mathbf{k}, \omega', \mathbf{k}') = -\frac{4\pi ie^3N_{0e}}{m^2\omega\omega'\omega''} \times\left[\frac{k_i}{\omega}\delta_{js} + \frac{k_j''}{\omega''}\delta_{is} + \frac{k_s'}{\omega'}\delta_{ij}\right]. \tag{3.2}$$

Using formula (2.8), we obtain the nonlinear current (3.1) of a standing plasma wave:

$$\mathbf{j}(\omega, \mathbf{k}) = \frac{\omega_p^2}{32\pi em}\int\frac{d\omega' d\mathbf{k}'}{(2\pi)^4}\frac{\omega_L^2\phi(\omega_L, \mathbf{k}')\phi(\omega_L, \mathbf{k}'')}{\omega'\omega''} \times\left[\frac{\mathbf{k}}{\omega}(\mathbf{k}'\cdot\mathbf{k}'') + \frac{\mathbf{k}'}{\omega''}k''^2 + \frac{\mathbf{k}''}{\omega'}k'^2\right] \times\left(\frac{1}{\omega' - \omega_L + i\gamma_L} - \frac{1}{\omega' + \omega_L + i\gamma_L}\right) \times\left(\frac{1}{\omega'' - \omega_L + i\gamma_L} - \frac{1}{\omega'' + \omega_L + i\gamma_L}\right). \tag{3.3}$$

Current (3.3) is proportional to the squared intensity of the plasma wave field or the fourth power of the laser field.

The radiation field is generated only by the vortex part of current (3.3). Separating this part, we obtain from the Maxwell equations the spectral density of the vortex electric field $\mathbf{E}^{tr}(\omega, \mathbf{r})$,

$$\mathbf{E}^{\mathrm{tr}}(\omega, \mathbf{r}) = -4\pi i\omega\,\mathrm{curl}\,\mathrm{curl}\int\frac{d\mathbf{k}}{(2\pi)^3}\frac{\exp(i\mathbf{k}\cdot\mathbf{r})}{k^2T(\omega, k)}\mathbf{j}(\omega, \mathbf{k}), \tag{3.4}$$

where $T(\omega, k) = \omega^2\varepsilon(\omega) - c^2k^2$ and $\varepsilon(\omega) = 1 - \omega_p^2/\omega^2$ is the transverse permittivity of the plasma. Relations (3.3) and (3.4) determine the low-frequency electromagnetic field induced in a rarefied plasma as a result of collision of two short laser pulses.

Substituting relation (3.3) into (3.4), integrating with respect to frequencies ω' and wave vectors $\mathbf{k}'$, and discarding the small terms proportional to $\exp(-k_0^2L^2)$, we obtain

$$\mathbf{E}^{\mathrm{tr}}(\omega, \mathbf{r}) = -\frac{i\omega\omega_p}{em}\frac{\pi^{5/2}m^2V_E^4k_0^2R^2L\tau^2}{16\sqrt{2}}(\mathbf{e}_+\cdot\mathbf{e}_-)^2 \times \exp\left(-\frac{\omega^2\tau^2}{8}\right)\mathrm{curl}\,\mathrm{curl}\,\mathbf{e}_z\frac{\partial}{\partial z} \times\int\frac{d\mathbf{k}}{(2\pi)^3}\frac{1}{k^2T(\omega,k)}\exp\left(i\mathbf{k}\cdot\mathbf{r} - \frac{k_z^2L^2}{8} - \frac{k_\perp^2R^2}{8}\right) \quad (3.5)$$
$$\times\left\{\frac{1+\mathrm{erf}\left(\frac{i(\omega-2\omega_L)\tau}{\sqrt{8}}\right)}{\omega-2\omega_L+2i\gamma_L}\exp\left[-\frac{(\omega-2\omega_L)^2\tau^2}{8}\right] - \frac{1+\mathrm{erf}\left(\frac{i(\omega+2\omega_L)\tau}{\sqrt{8}}\right)}{\omega+2\omega_L+2i\gamma_L}\exp\left[-\frac{(\omega+2\omega_L)^2\tau^2}{8}\right]\right\},$$

where

$$\mathrm{erf}(z) = \frac{2}{\sqrt{\pi}}\int_0^z dt\exp(-t^2)$$

is the probability integral of the complex argument, $\mathbf{e}_z$ is the unit vector in the direction of the z axis, $\omega_L = \omega_p(1 + 6k_0^2r_D^2)$ is the plasma oscillation frequency taking into account spatial dispersion, and

$$\gamma_L = \sqrt{\frac{\pi}{8}}\frac{\omega_p}{8k_0^3r_D^3}\exp\left(-\frac{3}{2} - \frac{1}{8k_0^2r_D^2}\right).$$

We are naturally interested in the radiation field in the wave zone at large distances from the region of localization of plasma oscillations ($r \gg R, L$). In this case, the integral over wave vectors in expression (3.5) can be evaluated by the steepest descent method. Taking into account the contribution from the pole, $k = \sqrt{\varepsilon(\omega)}\omega/c$, we arrive at the following expression for the spectral density of the electric field of low-frequency electromagnetic radiation:

$$\mathbf{E}^{\mathrm{tr}}(\omega, \mathbf{r}) = \frac{\pi^{3/2}\omega^2\tau^2}{8\sqrt{2}}\frac{\omega_p}{\omega_0}\frac{V_E^3}{8c^3}E_{0L} \times\frac{k_0^2R^2L}{r}(\mathbf{e}_+\cdot\mathbf{e}_-)^2\sqrt{\varepsilon(\omega)}\sin\theta\cos\theta\mathbf{e}_\theta \times\exp\left(i\frac{\omega}{c}\sqrt{\varepsilon(\omega)}r - \frac{\omega^2\tau^2}{8} - \frac{\omega^2}{c^2}\varepsilon(\omega)\frac{L^2\cos^2\theta + R^2\sin^2\theta}{8}\right) \quad (3.6)$$
$$\times\left\{\frac{1+\mathrm{erf}\left(\frac{i(\omega-2\omega_L)\tau}{\sqrt{8}}\right)}{\omega-2\omega_L+2i\gamma_L}\exp\left(-\frac{(\omega-2\omega_L)^2\tau^2}{8}\right) - \frac{1+\mathrm{erf}\left(\frac{i(\omega+2\omega_L)\tau}{\sqrt{8}}\right)}{\omega+2\omega_L+2i\gamma_L}\exp\left(-\frac{(\omega+2\omega_L)^2\tau^2}{8}\right)\right\},$$

where θ is the angle between the z axis and the direction of radius vector $\mathbf{r}$ and $\mathbf{e}_\theta$ is the unit vector in the meridional direction.

In order to find the space–time dependence of the low-frequency electric field in the wave zone, we carry out the inverse Fourier transformation in time. Taking into account the contributions from poles $\omega = \pm2\omega_L - 2i\gamma_L$, we obtain from relation (3.6)

$$\mathbf{E}^{\mathrm{tr}}(\mathbf{r}, t) = -\frac{\sqrt{3}\pi^{3/2}\omega_p^2\tau^2}{2\sqrt{2}}\frac{\omega_p}{\omega_0}\frac{V_E^3}{8c^3}E_{0L}\frac{k_0^2R^2L}{r}(\mathbf{e}_+\cdot\mathbf{e}_-)^2 \times\sin\theta\cos\theta\mathbf{e}_\theta\exp\left\{-2\gamma_L\left(t - \frac{2}{3}\frac{r}{c}\right) - \frac{\omega_L^2\tau^2}{2} - \frac{3}{8}k_p^2(L^2\cos^2\theta + R^2\sin^2\theta)\right\} \quad (3.7)$$
$$\times\sin\left\{2\omega_L\left[t - \frac{\sqrt{3}}{2}\frac{r}{c}(1 + 2k_0^2r_D^2)\right]\right\}.$$

This relation describes a diverging spherical electromagnetic wave having a doubled plasma frequency and polarized in the meridional direction. It should be noted that expression (3.7) holds only at large distances from the region of localization of small-scale plasma oscilla-

tions ($r \gg R, L$). In addition, it becomes valid only after some time following the passage of the leading front of radiation through a given point (for $t - r/c \gg \omega_p\tau^2$, $\omega_p R^2/c^2$).

4. ANGULAR, SPECTRAL, AND ENERGY CHARACTERISTICS OF RADIATION

The energy emitted in unit interval of frequencies $d\omega$ into unit solid angle do is connected with spectral density (3.6) of the electric field through the relation [9]

$$\frac{dW}{d\omega do} = \frac{c\sqrt{\varepsilon(\omega)}}{4\pi^2} r^2 \left|\mathbf{E}^{\mathrm{tr}}(\omega, \mathbf{r})\right|^2, \qquad (4.1)$$

where frequency ω is regarded as positive.

Using relation (3.6), we can write the expression for low-frequency radiation energy (4.1) in the form

$$\frac{dW}{d\omega do} = \frac{\sqrt{\pi}\omega_p^2\tau^2\omega^4\tau^3}{64}(\mathbf{e}_+ \cdot \mathbf{e}_-)^4 k_0^2 R^2 \varepsilon^{3/2}(\omega)$$

$$\times \left(\frac{V_E^2}{4c^2}\right)^3 W_L \sin^2\theta\cos^2\theta$$

$$\times \exp\left(-\frac{\omega^2\tau^2}{4} - \frac{\omega^2}{c^2}\varepsilon(\omega)\frac{L^2\cos^2\theta + R^2\sin^2\theta}{4}\right) \qquad (4.2)$$

$$\times \left| \frac{1 + \mathrm{erf}\left(\frac{i(\omega - 2\omega_L)\tau}{\sqrt{8}}\right)}{\omega - 2\omega_L + 2i\gamma_L} \exp\left(-\frac{(\omega - 2\omega_L)^2\tau^2}{8}\right) \right.$$

$$\left. - \frac{1 + \mathrm{erf}\left(\frac{i(\omega + 2\omega_L)\tau}{\sqrt{8}}\right)}{\omega + 2\omega_L + 2i\gamma_L} \exp\left(-\frac{(\omega + 2\omega_L)^2\tau^2}{8}\right) \right|^2,$$

where

$$W_L = \frac{E_{0L}^2 \pi^{3/2} R^2 L}{8\pi}$$

is the energy of a laser pulse. For weakly damped plasma oscillations ($\omega_p \gg \gamma_L$), expression (4.2) has a sharp peak in the vicinity of frequency $\omega = 2\omega_L$, corresponding to the emission of electromagnetic waves at a doubled plasma frequency. In the vicinity of resonance

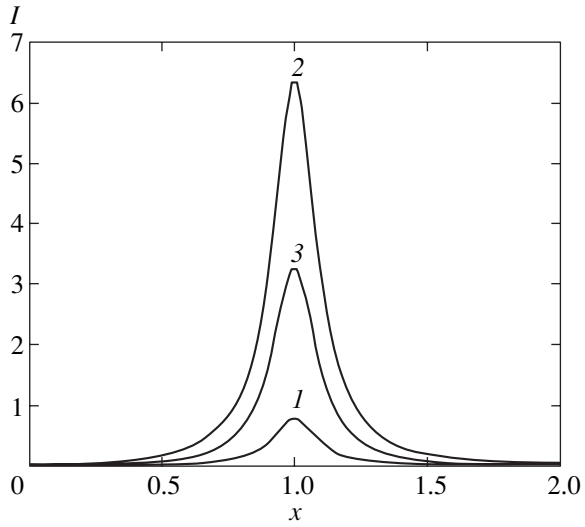

Fig. 1. Dependence of the dimensionless energy spectrum I of radiation emitted at angles of 45° and 135° on the dimensionless frequency $x = \omega/2\omega_L$ for $\gamma_L/\omega_L = 0.1$ and for fixed intensity and radius of laser pulses for three values of parameter $\beta = (\omega_L\tau)^2/2$, characterizing the pulse duration: $\beta = 0.25$ (*1*), 1 (*2*), and 2 (*3*).

($\omega = 2\omega_L$), formula (4.2) can be noticeably simplified and assumes the form

$$\frac{dW}{d\omega do} = \frac{3\sqrt{3\pi}\omega_p^6\tau^5 k_0^2 R^2}{32}(\mathbf{e}_+ \cdot \mathbf{e}_-)^4$$

$$\times \left(\frac{V_E^2}{4c^2}\right)^3 W_L \frac{\sin^2\theta\cos^2\theta}{(\omega - 2\omega_L)^2 + 4\gamma_L^2} \qquad (4.3)$$

$$\times \exp\left\{-\omega_L^2\tau^2 - \frac{3}{4}k_p^2(L^2\cos^2\theta + R^2\sin^2\theta) \right.$$

$$\left. - \frac{(\omega - 2\omega_L)^2\tau^2}{8}\right\}.$$

It follows from this formula that, for a fixed angle, both the radiation intensity and the linewidth depend on the damping rate of plasma oscillations as well as on the pulse duration τ. The reason for such dependence is that the pulse duration determines not only the efficiency of excitation of plasma waves, but also the longitudinal size of the region of their interaction. This size is associated with the spread in the longitudinal components of the wave vectors of excited plasma oscillations and with the possibility that the condition for the merging of two plasma waves with the formation of a transverse electromagnetic wave will be fulfilled.

Figure 1 shows the dependence of the dimensionless radiation energy spectrum,

$$I = \frac{dW}{d\omega do}\frac{512\omega_L^2}{3\sqrt{3}\pi k_0^2 R^2(\mathbf{e}_+\cdot\mathbf{e}_-)^4(V_E^2/4c^2)^3 cE_{0L}^2R^2}$$

$$\times\exp\left(\frac{3}{8}k_p^2R^2\right) = \frac{1}{\left(\frac{\omega}{2\omega_L}-1\right)^2+\frac{\gamma_L^2}{\omega_L^2}}$$

$$\times\left(\frac{\omega_L^2\tau^2}{2}\right)^3\exp\left\{-\frac{\omega_L^2\tau^2}{2}\left[\left(\frac{\omega}{2\omega_L}-1\right)^2+\frac{11}{4}\right]\right\},$$

at angles of $\theta = 45°$ and $135°$ for three values of parameter $\omega_L^2\tau^2/2$ characterizing the duration of pulses for their fixed intensity and radius. The pulse duration mainly affects the efficiency of plasma wave excitation and, hence, the radiation intensity. In addition, a relatively weak influence of the pulse duration on the radiation linewidth is observed. The shorter the pulses, the wider the spread in the longitudinal components of the plasma waves excited by these pulses and the larger the possible number of elementary processes of plasmon coalescence. It should be recalled that the coalescence of two plasmons with opposite wave vectors into one is possible only if these vectors differ in magnitude by a value on the order of ω_L/c.

Integrating expression (4.3) with respect to frequency, we obtain the angular distribution of the radiated energy:

$$\frac{dW}{do} = \frac{(3\pi)^{3/2}\omega_p^5\tau^5k_0^2R^2}{32}(\mathbf{e}_+\cdot\mathbf{e}_-)^4$$

$$\times\left(\frac{V_E^2}{4c^2}\right)^3 W_L\frac{\omega_p}{2\gamma_L}\sin^2\theta\cos^2\theta \qquad (4.4)$$

$$\times\exp\left\{-\omega_L^2\tau^2-\frac{3}{4}k_p^2(L^2\cos^2\theta+R^2\sin^2\theta)\right\}.$$

Let us analyze the dependence of the directivity diagram of the low-frequency electromagnetic radiation on the relation between the longitudinal and transverse sizes of laser pulses. The equation for the optimal angle θ_{max} corresponding to the maximum radiation energy follows from relation (4.4) and has the form

$$\sin^4\theta_{max}-\sin^2\theta_{max}\left[1+\frac{8}{3k_p^2(R^2-L^2)}\right] \qquad (4.5)$$

$$+\frac{4}{3k_p^2(R^2-L^2)} = 0.$$

If the pulses have identical or nearly identical longitudinal and transverse sizes (such that $k_p^2(R^2-L^2)\ll 1$), Eq. (4.5) leads to $\sin^2\theta_{max} = 0.5$, which corresponds to the following values of the angles:

$$\theta_{max} = \frac{\pi}{4}, \frac{3\pi}{4}. \qquad (4.6)$$

Radiation is concentrated in narrow cones with angles of 45° and 135° relative to the axis along which the pulses propagate. This result coincides with that obtained in [10, 11] from analysis of radiation emitted from solar flares in the approximation such that the spectrum of Langmuir noise in the wave vector space is concentrated in a small neighborhood of two opposite and rather long wave vectors.

When pulses with transverse sizes much larger than their longitudinal sizes ($R \gg L$) collide, the angular directivity of low-frequency radiation essentially depends on the parameter k_pR. For narrow pulses ($k_pR \ll 1$), the radiation intensity has the maximal value for angles defined by relation (4.6). If, however, pulses have large transverse sizes, Eq. (4.5) leads to

$$\theta_{max} = \sqrt{\frac{4}{3}\frac{1}{k_pR}}, \quad \pi-\sqrt{\frac{4}{3}\frac{1}{k_pR}}. \qquad (4.7)$$

In this case, radiation is emitted at small angles relative to the direction of pulse propagation. However, the radiation intensity is small in accordance with (4.4).

If the longitudinal size of colliding pulses exceeds their transverse size ($L \gg R$), the radiant energy attains its maximal value for angles (4.6) provided that $\omega_p\tau \ll 1$. If the opposite inequality ($\omega_p\tau > 1$) holds, Eq. (4.5) leads to

$$\theta_{max} = \sqrt{\frac{\pi}{2}} \pm \sqrt{\frac{4}{3}\frac{1}{\omega_p\tau}}. \qquad (4.8)$$

In this case, radiation is emitted in the direction perpendicular to the direction of pulse propagation, but the value of radiant energy is exponentially small.

Figure 2 shows the angular dependences of dimensionless energy of the radiation at twice the plasma frequency for three values of parameter $(R/L)^2$, characterizing the radii of laser pulses, for their fixed energy and duration:

$$g = \frac{dW}{do}\frac{L^2\exp(\omega_L^2\tau^2)}{24\sqrt{3}k_p^5k_0^2(\mathbf{e}_+\cdot\mathbf{e}_-)^4W_L^4}$$

$$\times\left(\frac{4m^2c^2\omega_0^2}{e^2}\right)^3\frac{\gamma_L}{\omega_L} = \left(\frac{L}{R}\right)^4\sin^2\theta\cos^2\theta$$

$$\times\exp\left\{-\frac{3}{4}\omega_L^2\tau^2\left(\cos^2\theta+\frac{R^2}{L^2}\sin^2\theta\right)\right\}.$$

It can be seen that the radiated energy decreases with increasing pulse radius and the directivity diagram becomes narrower. These effects are due to the fact that an increase in the pulse radius narrows the angular spread in the wave vectors of the excited plasma waves and reduces the number of elementary events of their coalescence (cf. [10, 11]). In addition, in accordance with results (4.7), optimal angle θ_{max} decreases with increasing transverse size of a laser pulse.

The total energy of low-frequency electromagnetic radiation can be determined from relation (4.4) by integrating over solid angle *do*. For pulses with identical spatial sizes $L = R$, integration is carried out analytically and the total energy has the form

$$W = \frac{\sqrt{3}(\pi)^{5/2}\omega_p^5\tau^5k_0^2R^2}{20}(\mathbf{e}_+\cdot\mathbf{e}_-)^4\left(\frac{V_E^2}{4c^2}\right)^3 W_L \times\frac{\omega_p}{2\gamma_L}\exp\left(-\frac{7}{4}\omega_p^2\tau^2\right). \quad (4.9)$$

The low-frequency radiant energy (4.9) attains its maximal value

$$W_{max} = \frac{\sqrt{6}k_0^2R^2}{5}(\mathbf{e}_+\cdot\mathbf{e}_-)^4\left(\frac{5\pi}{7e}\right)^{5/2}\left(\frac{V_E^2}{4c^2}\right)^3\frac{\omega_p}{2\gamma_L}W_L, \quad (4.10)$$

in the case of collision of two laser pulses with a duration equal to

$$\omega_p\tau = \sqrt{\frac{10}{7}}. \quad (4.11)$$

In spite of the fact that expression (4.10) for the energy contains the product of two large parameters $k_0^2R^2 \gg 1$ and $\omega_p/\gamma_L \gg 1$, the energy value must be smaller than the energy of a small-scale standing plasma wave [1],

$$W_p = \frac{\pi\omega_p^2\tau^2}{\sqrt{2}}\frac{V_E^2}{4c^2}W_L\exp\left(-\frac{\omega_p^2\tau^2}{2}\right). \quad (4.12)$$

Comparing expression (4.12) with formula (4.10), we find that our treatment is valid provided that

$$k_0^2R^2\left(\frac{V_E^2}{4c^2}\right)^2 \ll \frac{\gamma_L}{\omega_p}. \quad (4.13)$$

Another limitation on the plasma parameters is connected with the condition of smallness of the Landau damping for small-scale plasma oscillations ($\gamma_L \ll \omega_L$). For values of parameter $k_0r_D < 0.2$, ratio γ_L/ω_L does not exceed 0.1 and plasma oscillations are weakly damped.

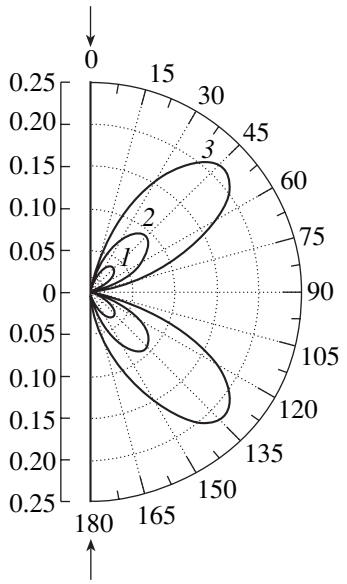

Fig. 2. Angular dependence of the dimensionless energy g of the radiation at twice the plasma frequency for short laser pulses ($\omega_p\tau = \sqrt{4/3}$) for various values of parameter $\alpha = (R/L)^2$ characterizing the pulse width: $\alpha = \sqrt{2}$ (*1*), 1 (*2*), and $1/\sqrt{2}$ (*3*).

For example, for $k_0r_D = 0.2$, the damping decrement amounts to $\gamma_L = 0.1\omega_L$ and the plasma oscillation frequency ω_L is equal to $1.24\omega_p$ on account of spatial dispersion.

Let us estimate the value of energy emitted upon the collision of two short laser pulses of intensity $I_L = 10^{16}$ W/cm^2, wavelength $\lambda_0 = 0.8$ μm ($\omega_0 = 2.4\times10^{15}$ s^{-1}), duration $\tau = 27$ fs, and focal spot diameter $2R = 16$ μm in a rarefied plasma of density $N_{0e} = 1.7\times10^{19}$ cm^{-3} and temperature $T_e = 20$ eV. For such parameters, the energy of the laser pulses amounts to $W_L = 0.96\times10^{-3}$ J, while the energy of a small-scale standing plasma wave is $W_p = 1.4\times10^{-3}W_L$. In this case, small-scale plasma oscillations attenuate weakly with a decrement of $\gamma_L = 0.1\omega_L$, and the energy emitted at a frequency of $\omega = 2\omega_L = 1.7\times10^{14}$ s^{-1}, which is approximately 1/15 of the laser carrier frequency, is directed at angles of 45° and 135° and has a value of $W = 2.8\times10^{-6}W_L = 2\times10^{-3}W_p = 2.7\times10^{-9}$ J. It should be emphasized once again that the frequency ω_L of Langmuir waves, taking into account spatial dispersion, is equal to $1.24\omega_p$ for such values of plasma parameters.

5. CONCLUSIONS

Radiation at twice the plasma frequency may carry information on the evolution of free small-scale localized plasma oscillations (which are not sustained by external sources) under the conditions of laser experi-

ments. The model considered here presumes the linearity of plasma waves and, in particular, their dissipation due to Landau damping. We have considered the energy characteristics of radiation, its spectrum, and directivity diagram precisely for such conditions. The discrepancy between the results of measurements and calculations may indicate the importance and type of nonlinear processes disregarded in our model of evolution of plasma waves. For example, the effects associated with the mobility of ions (induced scattering from ions or decay of a plasma wave into a plasma wave and an acoustic wave) make the plasma wave spectrum and, hence, directivity of radiation at twice the plasma frequency more isotropic. A nonlinear dissipation mechanism affects the shape of the radiation spectrum and the radiation intensity.

In analyzing plasma waves, the linear approximation is limited to moderate intensities of laser pulses of 10^{15}–10^{16} W/cm^2, while modern laser technology makes it possible to obtain pulses with a much higher intensity of 10^{18}–10^{20} W/cm^2; for such intensities, plasma waves being excited may be nonlinear and may contain a large number of spatial harmonics, which are multiple of $2k_0$. It can be expected that the processes of coalescence of higher order Langmuir waves will lead to electromagnetic radiation not only at a doubled plasma frequency, but at higher frequencies, which are multiple of the Langmuir frequency.

Nevertheless, using such high-intensity laser pulses, it is also possible to realize the limiting case of excitation of linear plasma waves considered here. Indeed, it follows from the results that the efficiency of plasma wave excitation depends only on parameter $\mathbf{e}_+ \cdot \mathbf{e}_-$. Plasma waves are excited most effectively for parallel polarization vectors of laser pulses; conversely, the effect of excitation of Langmuir waves disappears for orthogonal polarizations. By changing the angle between the polarization vectors, it is possible to affect the amplitudes of excited plasma waves and, hence, the spectrum and intensity of low-frequency radiation.

It should be noted that the radiation effects analogous to that considered here might also take place in other media [6] as well as in plasma for other types of waves. For example, doubled-frequency radiation emerging as a result of coalescence of two surface waves was considered in [12].

ACKNOWLEDGMENTS

This study was supported by the Russian Foundation for Basic Research (project nos. 01-02-16723 and 02-02-16110).

APPENDIX

Let us consider a quasi-static response of a plasma, quadratic in the rf field, by using the Vlasov kinetic equation [7]. We will seek the electron distribution function $f(\mathbf{r}, \mathbf{v}, t)$ in the form of a power series in the small parameter ε_E, which is equal to the ratio of the electron quiver velocity V_E to the phase velocity of electromagnetic radiation and is proportional to the high-frequency electric field strength $\mathbf{E}_L(\mathbf{r}, t)$,

$$f = F + f_1 + f_2 + \ldots,$$

where F is the electron distribution function in the absence of a laser pulse and f_n is the nth order perturbation of the distribution function in parameter ε_E.

In the linear approximation ($n = 1$), the equation for f_1 has the form

$$\frac{\partial f_1}{\partial t} + \mathbf{v} \cdot \frac{\partial f_1}{\partial \mathbf{r}} + \frac{e}{m}\left(\mathbf{E}_L + \frac{1}{c}\mathbf{v} \times \mathbf{B}_L\right)\frac{\partial F}{\partial \mathbf{v}} = 0, \quad \text{(A.1)}$$

where $\mathbf{B}_L$ is the high-frequency magnetic laser field, connected with the electric field via the equation

$$\mathrm{curl}\mathbf{E}_L = -\frac{1}{c}\frac{\partial \mathbf{B}_L}{\partial t}.$$

Keeping in mind another small parameter, ε_T, which is equal to the actual ratio of the electron thermal velocity V_T to the phase velocity of electromagnetic radiation, we can represent the solution to Eq. (A.1) in the form of a power series in the electron velocity:

$$f_1 = \frac{\partial F}{\partial \nu_j}\left(-A_j + \nu_i\frac{\partial C_i}{\partial x_j} - \nu_i\nu_l\frac{\partial^2 G_i}{\partial x_l \partial x_j} + \ldots\right), \quad \text{(A.2)}$$

where

$$\mathbf{A} = \frac{e}{m}\int_{-\infty}^{t} dt' \mathbf{E}_L(\mathbf{r}, t)$$

is the high-frequency electron velocity,

$$\mathbf{C} = \int_{-\infty}^{t} dt' \mathbf{A}(\mathbf{r}, t')$$

is the electron displacement in the rf field, and

$$\mathbf{G} = \int_{-\infty}^{t} dt' \mathbf{C}(\mathbf{r}, t').$$

While deriving formula (A.2), we assumed that there was no laser pulse at a given point $\mathbf{r}$ for $t \longrightarrow -\infty$. An expression similar to (A.2) was discussed in [13] under the assumption that F is the Maxwell distribution function.

In the approximation quadratic in ε_E, the equation for f_2 has the form

$$\frac{\partial f_2}{\partial t} + \mathbf{v} \cdot \frac{\partial f_2}{\partial \mathbf{r}} + \frac{e}{m}\left(\mathbf{E}_2 + \frac{1}{c}\mathbf{v} \times \mathbf{B}_2\right) \cdot \frac{\partial F}{\partial \mathbf{v}} = -I, \quad \text{(A.3)}$$

where the quantity I appearing on the right-hand side can be expressed in terms of functions defined by the linear approximation:

$$I = \frac{e}{m}\left(\mathbf{E}_L + \frac{1}{c}\mathbf{v}\times\mathbf{B}_L\right)\cdot\frac{\partial f_1}{\partial \mathbf{v}}. \quad (A.4)$$

Using formula (A.2), we can write expression (A.4) in the form

$$I = R_j\frac{\partial F}{\partial \nu_j} + D_{ij}\frac{\partial^2 F}{\partial \nu_i \partial \nu_j}. \quad (A.5)$$

Here, coefficients R_j and D_{ij} are proportional to the square of the rf field and are power functions of the electron velocity vector component:

$$R_j = R_j^0 + \nu_l S_{lj} + \ldots, \quad (A.6)$$

$$D_{ij} = D_{ij}^0 + \nu_l Q_{lij} + \nu_l\nu_m P_{lmij} + \ldots, \quad (A.7)$$

where

$$R_j^0 = \frac{\partial A_i}{\partial t}\frac{\partial C_i}{\partial x_j}, \quad (A.8)$$

$$S_{lj} = \left(\frac{\partial A_i}{\partial x_l} - \frac{\partial A_l}{\partial x_i}\right)\frac{\partial C_i}{\partial x_j} - \frac{\partial A_i}{\partial t}\frac{\partial}{\partial x_j}\left(\frac{\partial G_i}{\partial x_l} + \frac{\partial G_l}{\partial x_i}\right), \quad (A.9)$$

$$D_{ij}^0 = -\frac{1}{2}\frac{\partial}{\partial t}(A_iA_j), \quad (A.10)$$

$$Q_{lij} = \frac{1}{2}\left(A_i\frac{\partial}{\partial x_j} + A_j\frac{\partial}{\partial x_i}\right)A_l - \frac{1}{2}\frac{\partial}{\partial x_l}(A_iA_j) + \frac{1}{2}\left(\frac{\partial A_i}{\partial t}\frac{\partial}{\partial x_j} + \frac{\partial A_j}{\partial t}\frac{\partial}{\partial x_i}\right)C_l, \quad (A.11)$$

$$P_{lmij} = \frac{1}{2}\left(\frac{\partial A_i}{\partial x_l}\frac{\partial}{\partial x_j} + \frac{\partial A_j}{\partial x_l}\frac{\partial}{\partial x_i}\right)C_m - \frac{1}{2}\left(\frac{\partial A_i}{\partial t}\frac{\partial}{\partial x_j} + \frac{\partial A_j}{\partial t}\frac{\partial}{\partial x_i}\right)\frac{\partial G_m}{\partial x_l} - \frac{1}{2}\left(\frac{\partial A_l}{\partial x_i}\frac{\partial}{\partial x_j} + \frac{\partial A_l}{\partial x_j}\frac{\partial}{\partial x_i}\right)C_m. \quad (A.12)$$

In deriving tensor D_{ij}, we took into account its symmetry relative to permutation of indices i and j.

We have so far not used the explicit expression for the laser field. Consequently, formulas (A.5)–(A.12) are valid for an arbitrary dependence of $\mathbf{E}_L$ on coordinates and time.

We define the time dependence of the laser field in the form (2.1), retaining an arbitrary coordinate dependence of the amplitude. Substituting formula (2.1) into the definitions of quantities **A**, **C**, and **G** and integrating them with respect to time, we confine our analysis to the first-order terms in the derivatives of the slowly varying amplitude $\mathbf{E}_0$ everywhere except in (A.10). We will also take into account the second-order terms in small parameter $\varepsilon_t = (1/\omega_0)(\partial/\partial t)$ in the expression for D_{ij}^0. As a result, we find that coefficients R_j and D_{ij}^0 in formula (A.5) contain terms varying slowly with frequency along with the terms varying at the doubled frequency. Accordingly, we represent distribution function f_2 as well as fields $\mathbf{E}_2$ and $\mathbf{B}_2$ in the form of the sum of rapidly and slowly varying terms, denoting the latter by angle brackets $\langle\ldots\rangle$. As a result, we obtain from Eq. (A.3)

$$\left(\frac{\partial}{\partial t} + \mathbf{v}\cdot\frac{\partial}{\partial \mathbf{r}}\right)\langle f_2\rangle + \frac{e}{m}\left(\langle\mathbf{E}_2\rangle + \frac{1}{c}\mathbf{v}\times\langle\mathbf{B}_2\rangle\right)\cdot\frac{\partial F}{\partial \mathbf{v}} = -\langle I\rangle, \quad (A.13)$$

where, in accordance with Eq. (A.5), we have

$$\langle I\rangle = \langle R_j\rangle\frac{\partial F}{\partial \nu_j} + \langle D_{ij}\rangle\frac{\partial^2 F}{\partial \nu_i\partial \nu_j}. \quad (A.14)$$

The coefficients in this formula have the form

$$\langle R_j\rangle = \langle R_j^0\rangle + \nu_l\langle S_{lj}\rangle + \ldots, \quad (A.15)$$

$$\langle R_j^0\rangle = -\frac{e^2}{4m^2\omega_0^2}\left[\frac{\partial|E_0|^2}{\partial x_j} + 2\frac{i}{\omega_0}\left(E_{0i}\frac{\partial^2 E_{0i}^*}{\partial t\partial x_j} - E_{0i}^*\frac{\partial^2 E_{0i}}{\partial t\partial x_j}\right)\right], \quad (A.16)$$

$$\langle S_{lj}\rangle = -\frac{ie^2}{4m^2\omega_0^3}\left\{\frac{\partial}{\partial x_j}\left(E_{0i}\frac{\partial E_{0l}^*}{\partial x_i} - E_{0i}^*\frac{\partial E_{0l}}{\partial x_i}\right) + \frac{\partial}{\partial x_l}\left(E_{0i}\frac{\partial E_{0i}^*}{\partial x_j} - E_{0i}^*\frac{\partial E_{0i}}{\partial x_j}\right) + \frac{i}{\omega_0}\times\left[2\frac{\partial^2 E_{0i}}{\partial t\partial x_j}\left(\frac{\partial E_{0i}^*}{\partial x_l} - \frac{\partial E_{0i}^*}{\partial x_i}\right) - \frac{\partial E_{0i}}{\partial x_j}\frac{\partial}{\partial t}\left(\frac{\partial E_{0i}^*}{\partial x_l} - \frac{\partial E_{0l}^*}{\partial x_i}\right) + 3E_{0i}\frac{\partial^2}{\partial t\partial x_j}\left(\frac{\partial E_{0i}^*}{\partial x_l} + \frac{\partial E_{0l}^*}{\partial x_i}\right) + \text{c.c.}\right]\right\}, \quad (A.17)$$

$$\langle D_{ij}\rangle = \langle D_{ij}^0\rangle + \nu_l\langle Q_{lij}\rangle + \nu_l\nu_m\langle P_{lmij}\rangle, \quad (A.18)$$

$$\langle D_{ij}^0\rangle = \frac{ie^2}{8m^2\omega_0}\left\{\frac{i}{\omega_0}\frac{\partial}{\partial t}(E_{0j}E_{0i}^* + E_{0i}E_{0j}^*) + \frac{1}{\omega_0^2}\frac{\partial}{\partial t}\left(E_{0j}\frac{\partial E_{0i}^*}{\partial t} + E_{0i}\frac{\partial E_{0j}^*}{\partial t} - \text{c.c.}\right)\right\}, \quad (A.19)$$

$$\langle Q_{lij}\rangle = -\frac{e^2}{8m^2\omega_0^2}\left\{\frac{\partial}{\partial x_l}(E_{0i}^* E_{0j} + \text{c.c.})\right.$$
$$+\frac{i}{\omega_0}\frac{\partial}{\partial t}\left(E_{0i}\frac{\partial E_{0l}^*}{\partial x_j} + E_{0j}\frac{\partial E_{0l}^*}{\partial x_i} - \text{c.c.}\right) \quad \text{(A.20)}$$
$$\left.+\frac{i}{\omega_0}\frac{\partial}{\partial x_l}\left(E_{0j}\frac{\partial E_{0i}^*}{\partial t} + E_{0i}\frac{\partial E_{0j}^*}{\partial t} - \text{c.c.}\right)\right\},$$

$$\langle P_{lmij}\rangle = \frac{ie^2}{8m^2\omega_0^3}\left\{\frac{\partial}{\partial x_l}\left(E_{0i}^*\frac{\partial E_{0m}}{\partial x_j} + E_{0j}^*\frac{\partial E_{0m}}{\partial x_i} - \text{c.c.}\right)\right.$$
$$+\frac{i}{\omega_0}\left(-3E_{0i}\frac{\partial^3 E_{0m}^*}{\partial t\partial x_l\partial x_j} - 3E_{0j}\frac{\partial^3 E_{0m}^*}{\partial t\partial x_l\partial x_i} + \frac{\partial^2 E_{0i}}{\partial t\partial x_l}\frac{\partial E_{0m}^*}{\partial x_j}\right.$$
$$+\frac{\partial^2 E_{0j}}{\partial t\partial x_l}\frac{\partial E_{0m}^*}{\partial x_i} + \frac{\partial E_{0l}}{\partial x_i}\frac{\partial^2 E_{0m}^*}{\partial t\partial x_j} + \frac{\partial E_{0l}}{\partial x_j}\frac{\partial^2 E_{0m}^*}{\partial t\partial x_i} \quad \text{(A.21)}$$
$$\left.\left.-2\frac{\partial E_{0i}}{\partial x_l}\frac{\partial^2 E_{0m}^*}{\partial t\partial x_j} - 2\frac{\partial E_{0j}}{\partial x_l}\frac{\partial E_{0m}^*}{\partial t\partial x_i} + \text{c.c.}\right)\right\}.$$

Expressions (A.13)–(A.21) allow us to consider various effects to different degrees of accuracy in small parameters ε_T (smallness of the thermal velocity of particles as compared with the phase velocity of waves) and ε_t (smallness of the amplitude variation over a period of the rf field) in the approximation quadratic in the rf field. In particular, using these relations, it is possible to analyze the problem of so-called nonstationary ponderomotive forces, which had been considered for a long time mainly in the framework of the hydrodynamic description of a plasma (see the review article [14]).

Including the largest terms in parameters ε_t and ε_E in expression (A.14), we can write Eq. (A.13) in the form

$$\left(\frac{\partial}{\partial t} + \mathbf{v}\cdot\frac{\partial}{\partial \mathbf{r}}\right)\langle f_2\rangle$$
$$+\frac{e}{m}\left(\langle \mathbf{E}_2\rangle + \frac{1}{c}\mathbf{v}\times\langle \mathbf{B}_2\rangle\right)\cdot\frac{\partial F}{\partial \mathbf{v}} \quad \text{(A.22)}$$
$$= \frac{1}{m}\frac{\partial\phi}{\partial \mathbf{r}}\cdot\frac{\partial F}{\partial \mathbf{v}} + \frac{1}{m}\left[\left(\frac{\partial}{\partial t} + \mathbf{v}\cdot\frac{\partial}{\partial \mathbf{r}}\right)\phi_{ij}\right]\frac{\partial^2 F}{\partial v_i \partial v_j},$$

where

$$\phi_{ij} = \frac{e^2}{8m\omega_0^2}(E_{0j}E_{0i}^* + E_{0j}^* E_{0i}),$$
$$\phi = \delta_{ij}\phi_{ij} = \frac{e^2}{4m\omega_0^2}|E_0|^2. \quad \text{(A.23)}$$

Quantity ϕ is usually referred to as the rf potential.

Carrying out the Fourier transformation in Eq. (A.22) in coordinates and time, we obtain the following expression for $\langle f_2(\omega, \mathbf{k}, \mathbf{v})\rangle$:

$$\langle f_2(\omega, \mathbf{k}, \mathbf{v})\rangle = \frac{1}{i(\omega - \mathbf{k}\cdot\mathbf{v})}$$
$$\times\left\{-\frac{e}{m}\left(\langle \mathbf{E}_2\rangle + \frac{1}{c}\mathbf{v}\times\langle \mathbf{B}_2\rangle\right) + \frac{1}{m}i\mathbf{k}\phi(\omega, \mathbf{k})\right\} \quad \text{(A.24)}$$
$$\times\frac{\partial F}{\partial \mathbf{v}} + \frac{1}{m}\phi_{ij}(\omega, \mathbf{k})\frac{\partial^2 F}{\partial v_i \partial v_j}.$$

Using this expression, we find the Fourier component of the electron current density,

$$\langle j_i(\omega, \mathbf{k})\rangle = e\int d\mathbf{v} v_i \langle f_2(\omega, \mathbf{k}, \mathbf{v})\rangle$$
$$= -\frac{i\omega}{4\pi}\delta\varepsilon_{ij}\left(\langle E_{2j}\rangle - \frac{i}{e}k_j\phi\right), \quad \text{(A.25)}$$

where we have assumed that distribution function F is isotropic and introduced the standard notation [7]

$$\delta\varepsilon_{ij}(\omega, \mathbf{k}) = \frac{4\pi e^2}{m\omega}\int d\mathbf{v}\frac{v_i}{\omega - \mathbf{k}\cdot\mathbf{v}}\frac{\partial F}{\partial v_j}. \quad \text{(A.26)}$$

It should be noted that the rf field is taken into account in Eq. (A.22) via the two terms on the right-hand side. In the lowest approximation in small parameters ε_t and ε_T, the contribution to current density (A.25) comes only from the term proportional to the ponderomotive potential.

Confining our analysis to potential quasi-static fields and using expression (A.25), we find the Fourier component of the electric field strength,

$$\langle \mathbf{E}_2\rangle = \frac{i\mathbf{k}}{e}\phi(\omega, \mathbf{k})\frac{\delta\varepsilon^l(\omega, \mathbf{k})}{1 + \delta\varepsilon^l(\omega, \mathbf{k})}, \quad \text{(A.27)}$$

where $\delta\varepsilon^l$ is the partial contribution of electrons to the longitudinal permittivity of the plasma,

$$\delta\varepsilon^l(\omega, \mathbf{k}) = \frac{4\pi e^2}{mk^2}\int d\mathbf{v}\frac{1}{\omega - \mathbf{k}\cdot\mathbf{v}}\left(\mathbf{k}\cdot\frac{\partial F}{\partial \mathbf{v}}\right). \quad \text{(A.28)}$$

It should be emphasized that the approach to describing the quasi-static response of the plasma to the rf field developed in the Appendix can be used for analyzing a wider range of problems.

REFERENCES

1. L. M. Gorbunov and A. A. Frolov, Zh. Éksp. Teor. Fiz. **120**, 583 (2001) [JETP **93**, 510 (2001)].
2. V. V. Zheleznyakov, *Radioemission of the Sun and Planets* (Nauka, Moscow, 1964); *Radiation Processes in Plasmas*, Ed. by G. Bekefi (Wiley, New York, 1966; Mir, Moscow, 1971); S. A. Kaplan, S. B. Pikel'ner, and

V. N. Tsytovich, *Physics of the Solar Atmospheric Plasma* (Nauka, Moscow, 1977), p. 191.

3. V. L. Ginzburg and V. V. Zheleznyakov, Zh. Éksp. Teor. Fiz. **35**, 699 (1958).
4. V. N. Tsytovich, *Theory of Turbulent Plasma* (Atomizdat, Moscow, 1971; Plenum, New York, 1974); E. N. Kruchina, R. Z. Sagdeev, and V. D. Shapiro, Pis'ma Zh. Éksp. Teor. Fiz. **32**, 443 (1980) [JETP Lett. **32**, 419 (1980)].
5. M. V. Goldman, D. L. Newman, D. Russel, *et al.*, Phys. Plasmas **2**, 1947 (1995).
6. G. Blau, J. L. Coutaz, and R. Reinisch, Opt. Lett. **18**, 1352 (1993); E. V. Alieva, G. N. Zhizhin, V. A. Sychugov, *et al.*, Pis'ma Zh. Éksp. Teor. Fiz. **62**, 794 (1995) [JETP Lett. **62**, 817 (1995)].
7. V. P. Silin and A. A. Rukhadze, *Electromagnetic Properties of Plasma and Plasmalike Media* (Atomizdat, Moscow, 1961), pp. 81, 89; E. M. Lifshitz and L. P. Pitaevskiĭ, *Physical Kinetics* (Nauka, Moscow, 1979; Pergamon Press, Oxford, 1981).
8. V. V. Pustovalov and V. P. Silin, Tr. Fiz. Inst. im. P.N. Lebedeva, Akad. Nauk SSSR **61**, 109 (1972).
9. L. D. Landau and E. M. Lifshitz, *The Classical Theory of Fields*, 7th ed. (Nauka, Moscow, 1982), p. 229.
10. I. H. Cairns, J. Plasma Phys. **38**, 179 (1987).
11. A. J. Willes, P. A. Robinson, and D. B. Melrose, Phys. Plasmas **3**, 149 (1996).
12. R. Dragila and S. Vukovic, Phys. Rev. A **34**, 407 (1986).
13. P. Mora and R. Pellat, Phys. Fluids **22**, 2408 (1979).
14. G. W. Kentwill and D. A. Jones, Phys. Rep. **145**, 320 (1987).

Translated by N. Wadhwa

Journal of Experimental and Theoretical Physics, Vol. 98, No. 3, 2004, pp. 538–545.
From Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 125, No. 3, 2004, pp. 609–617.

SOLIDS
Structure

Low-Energy Singlet Dynamics of Spin-1/2 Kagomé Heisenberg Antiferromagnets¶

A. V. Syromyatnikov and S. V. Maleyev
St. Petersburg Nuclear Physics Institute, St. Petersburg, 188300 Russia
e-mail: syromyat@gtn.ru
Received June 30, 2003

Abstract—We suggest a new approach for description of the low-energy sector of spin-$\frac{1}{2}$ kagomé Heisenberg antiferromagnets (KAFs). We show that a kagomé lattice can be represented as a set of blocks containing 12 spins, having the form of stars and arranged in a triangular lattice. Each of these stars has two degenerate singlet ground states that can be considered in terms of pseudospin $\frac{1}{2}$. Taking into consideration symmetry, we show that the KAF lower singlet band is created from these degenerate states by the interstar interaction. We demonstrate that this band is described by the effective Hamiltonian of a magnet in the external magnetic field. The general form of this Hamiltonian is established. The Hamiltonian parameters are calculated up to the third order of perturbation theory. The ground-state energy calculated in the model considered is lower than those evaluated numerically in the previous finite clusters studies. A way of experimental verification of this picture using neutron scattering is discussed. It is shown that the approach presented cannot be directly extended to KAFs with larger spin values. © *2004 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

Unusual low-temperature properties of kagomé antiferromagnets (KAFs) have attracted much attention from both theorists and experimenters over the last decade. Apparently, the most intriguing features were observed in specific heat C measurements of SrCrGaO (spin-$\frac{3}{2}$ kagomé material) [1]. A peak at $T \approx 5$ K has been revealed that is practically independent of magnetic field up to 12 T, and C appeared to be proportional to T^2 at $T \lesssim 5$ K.

There is no appropriate theory describing the low-energy KAF sector. Qualitative understanding of low-temperature spin-$\frac{1}{2}$ KAF physics is based mostly on results of numerous finite-cluster investigations [2–6]. They revealed a gap separating the ground state from the upper triplet levels and a band of nonmagnetic singlet excitations with a very small or zero gap inside the spin gap. The number of states in the band increases with the number of sites N as α^N. For samples with up to 36 sites, it was found that $\alpha = 1.15$ and 1.18 for even and odd N, respectively [2, 5]. It is now believed that this wealth of singlets is responsible for a low-T specific heat peak and explains its field independence [1, 7].

The origin of the singlet band and the nature of the ground state are still under debate. Previous exact diagonalization studies of clusters with $N \leq 36$ [4, 8] revealed an exponential decay of the spin–spin and dimer–dimer correlation functions, and therefore the point of view that KAF is a spin liquid is widely accepted [2, 4–12].

A quantum dimer model (QDM) is now considered the best candidate for description of low-energy KAF properties [6, 9, 13–15]. In QDMs proposed for the kagomé problem in some recent papers [6, 13–16], the spin Hilbert space is restricted to the states in which spins are paired into first-neighbor singlets. The main argument to support this restriction is the coincidence of the low-energy spectrum and the number of lower singlet excitations in samples with up to 36 sites with the exact diagonalization results [13, 15, 16]. At the same time, it was noted that further studies are required to analyze this problem. As was recently demonstrated in [15], an effective Hamiltonian describing the low-energy KAF singlet sector can be written in this approach. Unfortunately, it appears to be quite cumbersome and makes it possible to obtain the result under a number of crude approximations only [14, 15].

In our recent paper [17], we suggested another approach for spin-$\frac{1}{2}$ KAF that differs from the QDMs discussed above. We proposed to consider a kagomé lattice as a set of stars with 12 spins arranged in a triangular lattice (see Fig. 1). Numerical diagonalization has shown that a single star has two degenerate singlet ground states separated from the upper triplet levels by

¶This article was submitted by the authors in English.

a gap. These states form a singlet energy band as a result of interstar interaction. It was assumed that this band determines the low-energy KAF singlet sector. We have shown that it is described by the Hamiltonian of a magnet in the external magnetic field where degenerate states of the stars are represented in terms of two projections of pseudospin $\frac{1}{2}$.

This picture possibly reflects only the lowest part of the lower singlet sector because the number of states in the band within our approach is $2^{N/12} \approx 1.06^N$ [17], whereas it is now believed that it should be scaled by the 1.15^N law obtained numerically for clusters with $N \le 36$ [13, 16].

In this, more comprehensive paper, we develop this star concept. Using the symmetry considerations presented in Section 2, we prove that the singlet band arising from the ground states of stars determines the KAF lower singlet sector. This band is studied in Section 3, where the general form of the effective Hamiltonian is established. The Hamiltonian parameters are calculated up to the third order of perturbation theory. The ground-state energy calculated in the model considered is lower than these energies evaluated numerically in previous finite cluster studies. A comparison between our model and the QDM is carried out. We demonstrate that our approach cannot be directly extended to KAFs with larger spin values. A way to verify this picture experimentally using neutron scattering is discussed. We summarize our results in Section 4.

2. SYMMETRY CONSIDERATION

We start with the Hamiltonian of the spin-$\frac{1}{2}$ kagomé Heisenberg antiferromagnet,

$$\mathcal{H}_0 = J_1 \sum_{\langle i,j \rangle} \mathbf{S}_i \cdot \mathbf{S}_j + J_2 \sum_{(i,j)} \mathbf{S}_i \cdot \mathbf{S}_j, \qquad (1)$$

where $\langle i, j \rangle$ and (i, j) denote nearest and next-nearest neighbors on the kagomé lattice, respectively, shown in Fig. 1. The case where $|J_2| \ll J_1$ is considered in this paper. We discuss the possibility of both signs of the next-nearest-neighbor interactions, ferromagnetic and antiferromagnetic. As is shown below, although the second term in Eq. (1) is small, it can be important for the low-energy properties.

A kagomé lattice can be represented as a set of stars arranged in a triangular lattice (see Fig. 1). We first neglect the interaction between stars and set $J_2 = 0$ in Eq. (1). A star is a system of 12 spins. We now consider its properties in detail.

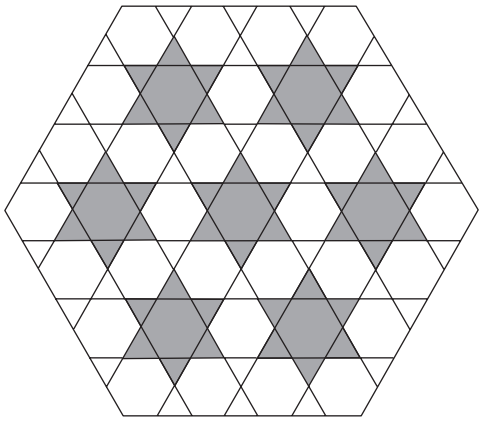

Fig. 1. Kagomé lattice (KL). There is a spin at each lattice site. The KL can be considered as a set of stars arranged in a triangular lattice. Each star contains 12 spins. A unit cell is also presented (dark region). There are four unit cells per star.

Because Hamiltonian (1) commutes with all of the projections of the total spin operator, all of the star levels are classified by the values of S, irreducible representations (IRs) of its symmetry group, and are degenerate with respect to S_z. The star symmetry group C_{6v} contains six rotations and reflections with respect to six lines passing through the center. There are four one-dimensional and two two-dimensional IRs, which are presented in Appendix 1. In their bases, the matrix of the Hamiltonian has a block structure. Each block has been diagonalized numerically.

As a result, it was found that the star has a doubly degenerate singlet ground state separated from the lower triplet level by a gap $\Delta \approx 0.26J_1$. Ground-state wave functions can be represented as

$$\Psi_+ = \frac{1}{\sqrt{2 + 1/16}}(\phi_1 + \phi_2), \qquad (2)$$

$$\Psi_- = \frac{1}{\sqrt{2 - 1/16}}(\phi_1 - \phi_2), \qquad (3)$$

where functions ϕ_1 and ϕ_2 are shown schematically in Fig. 2. The bold line there represents the singlet state of the corresponding two spins, i.e., $(|\uparrow\rangle_i|\downarrow\rangle_j - |\downarrow\rangle_i|\uparrow\rangle_j)/\sqrt{2}$.

It can be shown that ϕ_1 and ϕ_2 are not orthogonal: their scalar product is $(\phi_1\phi_2) = 1/32$. They contain six singlets, each having the energy $-S(S + 1)J_1 = -3/4J_1$. It can be shown that the interaction between singlets does not contribute to the energy of the ground states which is consequently equal to $-4.5J_1$.

Functions ϕ_1 and ϕ_2 are invariant under rotations of the star and transform into each other under reflections. Hence, Ψ_+ is invariant under all symmetry group transformations. In contrast, function Ψ_- is invariant under

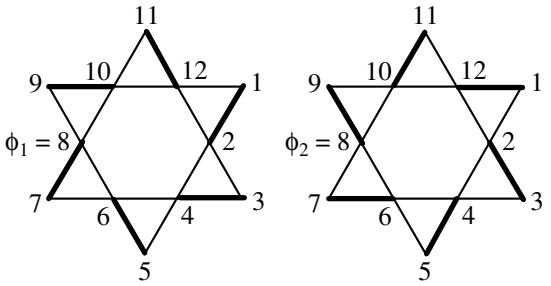

Fig. 2. Schematic representation of the two singlet ground-state wave functions ϕ_1 and ϕ_2 of a star. The bold line shows the singlet state of two neighboring spins, i.e., $(|\uparrow\rangle_i|\downarrow\rangle_j - |\downarrow\rangle_i|\uparrow\rangle_j)/\sqrt{2}$.

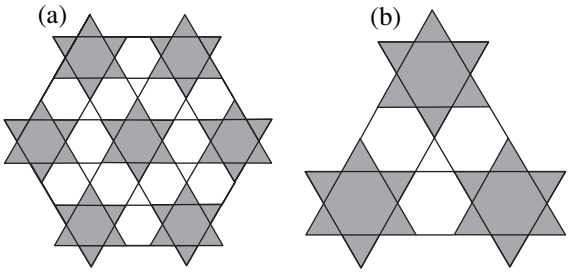

Fig. 3. (a) A cluster where operator of the interaction between stars has the same symmetry group C_{6v} as the whole cluster; (b) the only configuration of three stars giving a nonzero contribution to the third term in the perturbation expansion.

rotations, changes sign under reflections, and is transformed under representation (A.3). Therefore, the ground state has accidental degeneracy. As shown in the next section, the next-nearest-neighbor interaction, which has the same symmetry as the original Hamiltonian, removes this degeneracy.

The KAF containing $\mathcal{N}$ noninteracting stars has an energy spectrum with a large level degeneracy when $\mathcal{N} \gg 1$. For example, the ground-state degeneracy is $2^{\mathcal{N}}$ and that of the lowest triplet level is $3\mathcal{N}2^{(\mathcal{N}-1)}$. Interaction between stars gives rise to an energy band from each such group of levels, and it is a very difficult task to follow their evolution. On the other hand, group theory makes it possible to draw certain conclusions about the KAF low-energy sector. We now show that the singlet band that stemming from the ground state cannot be overlapped by those originating from the upper singlet levels.

We consider a cluster with seven stars, shown in Fig. 3a, and we begin with neglecting the interaction between them. The symmetry group of the cluster is also C_{6v}. The ground state has a degeneracy of $2^7 = 128$. The corresponding wave functions transformed under IRs of C_{6v} are constructed as linear combinations of products containing Ψ_+ or Ψ_- for each star. It is easy to show, by using the standard procedure of constructing bases in irreducible representations ([18, 19] and Appendix 1), that there are at least two ground-state wave functions of the cluster discussed that are transformed under any given IR.

It is important to mention that the operator of the interstar interaction in the cluster has the same symmetry as the intrastar interaction, which is the sum of the star Hamiltonians. Moreover, the interaction between stars commutes with the square of the total spin operator. Therefore, if we increase the interstar interaction from zero, all the levels move in energy, but their classification cannot be changed. Levels can cross each other as the interaction rises from 0 to J_1, but crossing is forbidden for levels of the same symmetry. This is a consequence of a symmetry theorem proved in [18, 19]. Hence, we can conclude that the lower singlet sector of the cluster is formed by states that stem from the original 128 lower levels.

We can obtain the same conclusions considering clusters with the symmetry C_{6v} with a large number of stars. Therefore, we assume in what follows that the KAF low-energy singlet sector is formed by the states originating from those in which each star is in one of the states Ψ_+ or Ψ_-. Because bands with $S \neq 0$ can overlap the singlet bands under discussion, we have to suppose that the KAF low-energy properties are determined by the lowest states in this singlet band.

Because the interaction between stars commutes with the square of the total spin operator, bands of different S can be studied independently. The KAF lower singlet sector is considered in detail in the next section. Investigation of states with $S \neq 0$ is outside the scope of this paper.

3. SINGLET DYNAMICS

In this section, we derive the general form of the effective Hamiltonian describing the lower singlet sector. The interstar interaction is considered a perturbation. Although it is not small compared to the intrastar interaction, there are reasons presented below to use perturbation expansion here.

Two-star coupling. We begin with considering the interaction between two nearest stars, still neglecting the second term in Eq. (1). Initially, there is a fourfold degenerate ground state with the wave functions $\{\Psi^{(1)}_{n_1}\Psi^{(2)}_{n_2}\}$ (where $n_i = +, -$, and the upper index denotes the stars) and the energy

$$E^{(0)}_{n_1 n_2} = E^{(0)}_{n_1} + E^{(0)}_{n_2} = -9J_1.$$

As can be seen from Fig. 4, the interaction has the form

$$V = J_1(\mathbf{S}^{(1)}_1 \cdot \mathbf{S}^{(2)}_1 + \mathbf{S}^{(1)}_3 \cdot \mathbf{S}^{(2)}_3). \quad (4)$$

According to the standard theory [18], the following

conditions should be satisfied to consider V as a perturbation:

$$\left|C_{m_1m_2}^{n_1n_2}\right| = \left|\frac{V_{n_1n_2;\,m_1m_2}}{E_{n_1n_2}^{(0)} - E_{m_1m_2}^{(0)}}\right| \ll 1, \tag{5}$$

where $V_{n_1n_2;\,m_1m_2} = \langle\Psi_{n_1}^{(1)}\Psi_{n_2}^{(2)}|V|\Psi_{m_1}^{(1)}\Psi_{m_2}^{(2)}\rangle$, m_1m_2 denotes excited singlet levels of the two stars, and $n_i = +, -$. We have calculated $C_{m_1m_2}^{n_1n_2}$ for $n_i = +, -$ using wave functions obtained numerically and found that all of these coefficients do not exceed 0.09. Conditions (5) are, therefore, satisfied. Then, the maximum value of the sum $\sum_{m_1m_2}\left|C_{m_1m_2}^{n_1n_2}\right|^2$ is 0.28, which is also sufficiently small. Thus, the interaction between stars is considered a perturbation in what follows.

We proceed with calculations of corrections to the initial ground-state energy of two stars. For this, because the state is fourfold degenerate, we must solve a secular equation [18]. The corresponding matrix elements in the third order of perturbation theory are given by [18]

$$\begin{aligned} H_{n_1n_2;\,k_1k_2} &= V_{n_1n_2;\,k_1k_2} \\ &+ \sum_{m_1,\,m_2}\frac{V_{n_1n_2;\,m_1m_2}V_{m_1m_2;\,k_1k_2}}{E_{n_1n_2}^{(0)} - E_{m_1m_2}^{(0)}} \\ &+ \sum_{m_1,\,m_2}\sum_{q_1,\,q_2}\frac{V_{n_1n_2;\,m_1m_2}V_{m_1m_2;\,q_1q_2}V_{q_1q_2;\,k_1k_2}}{(E_{n_1n_2}^{(0)} - E_{m_1m_2}^{(0)})(E_{n_1n_2}^{(0)} - E_{q_1q_2}^{(0)})}, \end{aligned} \tag{6}$$

where $n_i, k_i = +, -$. Obviously, the first term in Eq. (6) is zero, and the second term can be represented as

$$\begin{aligned} H_{n_1n_2;\,k_1k_2} &= -i\int_0^\infty dt e^{-\delta t + iE_{n_1n_2}^{(0)}t} \\ &\times\langle\Psi_{n_1}^{(1)}\Psi_{n_2}^{(2)}|Ve^{-it(\mathcal{H}_0^{(1)} + \mathcal{H}_0^{(2)})}V|\Psi_{k_1}^{(1)}\Psi_{k_2}^{(2)}\rangle, \end{aligned} \tag{7}$$

where $\mathcal{H}_0^{(i)}$ are Hamiltonians of the corresponding stars. The third term in Eq. (6) is to be considered later. Using the symmetry of the functions Ψ_+ and Ψ_- discussed above and the invariance of the system under reflection with respect to the dotted line in Fig. 4, it can be shown that the only nonzero elements belong to the first and to the second diagonals (i.e., with $n_1 = k_1$, $n_2 = k_2$ and with $n_1 \neq k_1$, $n_2 \neq k_2$). We have calculated them numerically with a very high precision by expansion of the operator exponent up to the power 150.[1] The results can be represented as

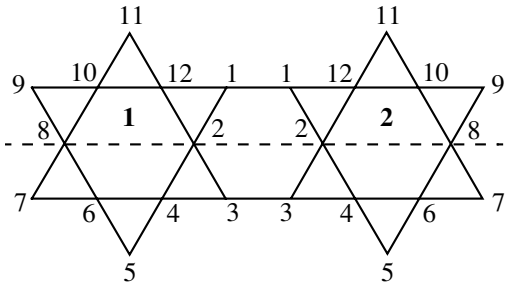

Fig. 4. Interactions between two stars: $V = J_1(\mathbf{S}_1^{(1)} \cdot \mathbf{S}_1^{(2)} + \mathbf{S}_3^{(1)} \cdot \mathbf{S}_3^{(2)})$ and $\tilde{V} = J_2(\mathbf{S}_1^{(1)} \cdot \mathbf{S}_2^{(2)} + \mathbf{S}_2^{(1)} \cdot \mathbf{S}_1^{(2)} + \mathbf{S}_1^{(2)} \cdot \mathbf{S}_3^{(2)} + \mathbf{S}_3^{(1)} \cdot \mathbf{S}_2^{(2)})$, where the superscripts label the stars. The system is symmetric under reflection with respect to the dotted line.

$$H_{++;\,++} = -a_1 + a_2 - a_3, \tag{8}$$

$$H_{+-;\,+-} = -a_1 + a_3, \tag{9}$$

$$H_{-+;\,-+} = -a_1 + a_3, \tag{10}$$

$$H_{--;\,--} = -a_1 - a_2 - a_3, \tag{11}$$

where $a_1 = 0.256J_1$, $a_2 = 0.015J_1$, and $a_3 = 0.0017J_1$. The terms of the second diagonal $H_{++;\,--} = H_{--;\,++} = -H_{+-;\,-+} = -H_{-+;\,+-} = a_4 = -0.0002J_1$ are much smaller than a_1, a_2, and a_3. Therefore, the interaction shifts all the levels by the value $-a_1$ and removes their degeneracy. The constants a_2, a_3, and a_4 determine the level splitting. It is seen that the splitting is very small compared to the shift.

Therefore, the KAF appears to be a set of two-level interacting systems, and the low-energy singlet sector of the Hilbert space can naturally be represented in terms of pseudospins: $|\uparrow\rangle = \Psi_-$ and $|\downarrow\rangle = \Psi_+$. It follows from Eqs. (8)–(11) that, in these terms, the interaction between stars is described by the Hamiltonian of a ferromagnet in the external magnetic field,

$$\mathcal{H} = \sum_{\langle i,\,j\rangle}[\mathcal{J}_z s_i^z s_j^z + \mathcal{J}_y s_i^y s_j^y] + h\sum_i s_i^z + \mathcal{C}, \tag{12}$$

where $\langle i, j\rangle$ now denotes nearest-neighbor pseudospins, arranged in a triangular lattice formed by the stars;

[1] The difference in these results from those obtained by the expansion of the exponent up to a power of 149 is on the order of 10^{-5}%. So, the method gives nearly precise values. The results would be the same if one were to use the more common expression (6) and eigenfunctions of the star obtained numerically.

Contributions to parameters of the effective Hamiltonian (13) from the terms V^1, V^2, and V^3 of the perturbation expansion. Interaction J_2 has been taken into account in V^1 and V^2 terms only. $\mathcal{N}$ is the number of stars in the lattice

	V^1	V^2	V^3*		Totals
			two-stars	three-stars	
$\mathcal{J}_z$	0	$-0.007J_1 + 0.002J_2$	$-0.013J_1$	$0.010J_1$	$-0.010J_1 + 0.002J_2$
$\mathcal{J}_y$	0	$-0.001J_1 + 0.007J_2$	$-0.001J_1$	$0.001J_1$	$-0.001J_1 + 0.007J_2$
$\mathcal{J}_x$	0	0	$0.067J_1$	0	$0.067J_1$
h	$-0.563J_2$	$-0.092J_1 - 0.218J_2$	$-0.161J_1$	$0.080J_1$	$-0.173J_1 - 0.781J_2$
$\Delta\mathcal{C}$**	$-0.009J_2\mathcal{N}$	$-0.768J_1\mathcal{N} + 1.530J_2\mathcal{N}$	$-0.361J_1\mathcal{N}$	$0.304J_1\mathcal{N}$	$-0.825J_1\mathcal{N} + 1.521J_2\mathcal{N}$

* This term implies two-star coupling shown in Fig. 4 and three-star interaction in the configuration presented in Fig. 3b.

** Correction to the value $\mathcal{C}_0 = -4.5J_1\mathcal{N}$ for noninteracting stars.

s is the spin-$\frac{1}{2}$ operator; $\mathcal{C} = -5.268J_1\mathcal{N}$; $\mathcal{J}_z = 4a_3 = -0.007J_1$; $\mathcal{J}_y = 4a_4 = -0.001J_1$; and $h = -6a_2 = -0.092J_1$. Here, $\mathcal{N} = N/12$ is the number of stars in the lattice. The factor 6 appears in the expression for h because each star interacts with six neighbors. We see that the magnetic field in the effective Hamiltonian (12) is much larger than the exchange. In this approximation, stars, therefore, behave almost like free spins in external magnetic field and the ground state of the KAF has a long-range singlet order, which settles on the triangular star lattice and is formed by stars in Ψ_- states.

V_3 corrections. The field remains the largest term in the effective Hamiltonian, and the KAF ground state has the same long-range order if we take the V^3 terms in the perturbation series into account. For the two-star coupling, the V^3 corrections have the form given by Eq. (6). The terms V^3 imply analysis of the three-star interaction as well. Nonzero contributions from them have only been obtained for the configuration presented in Fig. 3b. The secular matrix for three stars is 8×8 in size. We have calculated the V^3 corrections with a very high precision using the integral representation similar to that in Eq. (7) for the second term in Eq. (6). All the operator exponents were expanded up to a power of 150. As a result, the low-energy properties of the KAF are described by the effective Hamiltonian

$$\mathcal{H} = \sum_{\langle i,j \rangle} [\mathcal{J}_z s_i^z s_j^z + \mathcal{J}_x s_i^x s_j^x + \mathcal{J}_y s_i^y s_j^y] + h\sum_i s_i^z + \mathcal{C}, \quad (13)$$

where all parameters are given in the table. It describes two-star coupling. We omit the three pseudospin terms in Eq. (13) that have the form $s_i^z s_j^z s_k^z$ and $s_i^z s_j^y s_k^y$ and describe the three-star interaction. The corresponding coefficients are on the order of $10^{-3}J_1$ and $10^{-4}J_1$, respectively, and are negligible in comparison with those of the retained terms. We stress that, within our precision, the Hamiltonian in Eq. (13) is an exact mapping of the original Heisenberg model to the low-energy sector (the excitation energy $\omega \sim \max\{\mathcal{J}_z, \mathcal{J}_y, \mathcal{J}_x, h\} \ll J_1$).

As follows from the study of the V^3 corrections in the table, the common shift given by them remains much larger than the level splitting in both cases of the two- and three-star coupling. At the same time, the values of the V^3 perturbation terms are approximately two times smaller than those of the V^2 ones. The change in the effective Hamiltonian from the V^3 terms is, therefore, significant, and analysis of the perturbation series cannot be completed at this point for correctly establishing the effective Hamiltonian. Unfortunately, such work requires large computer capacity that is not at our disposal. We must restrict ourselves to the present precision.

One can judge the applicability of the perturbation series from the following values of the ground-state energy of two interacting stars, shown in Fig. 4, obtained numerically and using the first two orders of perturbation theory. The ground-state energy of two noninteracting stars is $-9J_1$. That of two interacting ones calculated numerically by power method [20] is $-9.62J_1$. On the other hand, the ground-state energy obtained using table is $-9.42J_1$ (the respective contributions of the V^2 and V^3 terms are $-0.27J_1$ and $-0.15J_1$).

Effective Hamiltonian structure. Although perturbation theory is unsatisfactory in the star model and many perturbation terms are to be taken into account, we can now prove that Eq. (13) is the most general form of the effective Hamiltonian assuming that n-pseudospin couplings with $n > 2$ are small, as it was in the case of $n = 3$ discussed above. We consider possible terms of the form $s_i^z s_j^+$, $s_i^z s_j^-$, s_i^+, and s_i^-. In these cases, the numbers of functions Ψ_+ and Ψ_- to the right of the

corresponding matrix elements differ from those to the left by unity. As has been pointed out above, a kagomé lattice contains lines of symmetry reflections, and the star Hamiltonian and interstar interaction are invariant under these reflections. Because Ψ_+ are invariant and Ψ_- changes sign under these transformations, the matrix elements are equal to themselves with the opposite sign and must, therefore, be zero. Another possible term

$$s_i^x s_j^y = -\frac{i}{4}(s_i^+ s_j^+ - s_i^- s_j^- + s_i^- s_j^+ - s_i^+ s_j^-)$$

cannot appear in the effective Hamiltonian because the corresponding matrix elements should be imaginary.

Ground state. As is clear from table, $\mathcal{J}_x$ and h are the largest parameters of Hamiltonian (13) in our approximation. Therefore, the KAF behaves as the Ising antiferromagnet in the perpendicular magnetic field. In this case, the classical value of the field at which spin-flip occurs is $h_{s-f} = \mathcal{J}_x$, which is approximately 2.6 times smaller than h. The ground state must, therefore, remain ordered with all the stars in the Ψ_- state.

The ground-state energy and that of the upper edge of the singlet band calculated using table are

$$(-4.5J_1 + \Delta\mathcal{C} + h/2 + 3\mathcal{J}_z/4)\mathcal{N} = -0.452J_1N$$

and

$$(-4.5J_1 + \Delta\mathcal{C} - h/2 + 3\mathcal{J}_z/4)\mathcal{N} = -0.437J_1N,$$

respectively. Corrections from $\mathcal{J}_x$ to these values in the first nonzero order of perturbation theory are given by $(3/16)\mathcal{N}\mathcal{J}_x^2/h$ and are negligible. At the same time, the ground-state energy of the largest cluster with $N = 36$ that has previously been considered numerically is $-0.438J_1N$ [5]. Hence, we believe that clusters used in the previous studies were too small to reflect the Heisenberg KAF low-energy sector at $J_2 = 0$ properly.

Interaction J_2. We now show that, in spite of its smallness, the next-nearest neighbor interaction can play an important role for low-energy properties. We have calculated J_2 corrections to the parameters of effective Hamiltonian (13) for the first and the second terms in Eq. (6) only. There are 12 intrinsic J_2 interactions in each star, which splits the doubly degenerate ground state and, as is seen from table, makes a contribution to magnetic field h and constant $\mathcal{C}$.

As is clear from Fig. 4, the two-star coupling is now given by the operator

$$\tilde{V} = J_2(\mathbf{S}_1^{(1)} \cdot \mathbf{S}_2^{(2)} + \mathbf{S}_2^{(1)} \cdot \mathbf{S}_1^{(2)} + \mathbf{S}_2^{(1)} \cdot \mathbf{S}_3^{(2)} + \mathbf{S}_3^{(1)} \cdot \mathbf{S}_2^{(2)}).$$

Corrections proportional to J_2 were calculated in the same way as above and are also presented in table. It is clear that the contribution of the next-nearest interactions to the magnetic field becomes significant if $|J_2| \sim 0.1J_1$. If $J_2 < 0$ (ferromagnetic interaction), they can even change the sign of h.

The effect of the next-nearest ferromagnetic coupling for KAF properties was previously studied in [2] numerically on finite clusters with $N \le 27$ in a wide range of the values of J_2. It was shown there that, at $|J_2|/J_1 \ll 1$, the ground state has the $\sqrt{3} \times \sqrt{3}$ magnetic structure. At $|J_2|/J_1 \ll 1$, the ground state is disordered and there is a band of singlet excitations inside the triplet gap. As we demonstrated above, this band is a result of the star ground-state degeneracy in our approach.

T^2 specific heat behavior. There has been much speculation on the low-temperature dependence of the KAF specific heat $C \propto T^2$ observed experimentally for $S = 3/2$ (see [7, 21] and references therein). As we found above, low-T properties are described by effective Hamiltonian (13) of a magnet, which has a spectrum of the form $\epsilon_{\mathbf{q}} = \sqrt{(cq)^2 + \Delta'^2}$ at $q \ll 1$ and can be in ordered or disordered phases depending on particular values of the parameters. Small Δ' here implies the proximity to the quantum critical point at which $C \propto T^2$. Such a situation arises in the singlet dynamics of the model of interactive plaquets [21]. We do not present the corresponding analysis here because the parameters of the effective Hamiltonian could be changed in subsequent orders of perturbation theory.

Experimental verification. In both cases of the ordered and disordered ground state, the approach presented in this paper can be checked by inelastic neutron scattering: the corresponding intensity for the singlet–triplet transitions should have a periodicity in the reciprocal space corresponding to the star lattice. This picture is similar to that observed in the dimerized spin-Pairls compound $CuGeO_3$ [22]. In this case, inelastic magnetic scattering has a periodicity that corresponds to the dimerized lattice.

Comparison with QDMs. We point out that states in which all stars are in the Ψ_- or Ψ_+ state can be represented as linear combinations of some first-neighbor dimer states proposed in [15, 16] for QDM. However, our approach to the kagomé problem is not equivalent to the QDM. In particular, we take all intermediate states into account in considering the star interaction via perturbation theory in Eq. (6), whereas the QDM is restricted to the first-neighbor dimer subspace as regards the dimer dynamics.

Unfortunately, we cannot carry out a complete comparison between QDM and the star approach at the present stage. The effective Hamiltonian derived in [15] was analyzed under crude approximations only. At the same time, the model presented here also requires further studies of the applicability of perturbation theory for description of the interstar interaction. We also note that some present-day results obtained within these two

approaches contradict each other. For example, our model gives the ordered ground state, whereas the authors of [15] suggest that it is not ordered.

Cases of $S > 1/2$. We finally note that our consideration of the $S = 1/2$ KAF cannot be extended directly to the cases of larger spins. Although functions presented in Fig. 2, where the bold line shows the singlet state of the corresponding two spins, remain eigenfunctions of the Hamiltonian for $S > 1/2$, we have found numerically that they are not ground states of the star with $S = 1$ and $S = 3/2$. All details of calculations are presented in Appendix 2. Another approach for KAFs with $S > 1/2$ should therefore be proposed.

4. CONCLUSIONS

In this paper, we present a model of the low-energy physics of spin-$\frac{1}{2}$ KAFs. The spin lattice can be represented as a set of stars that are arranged in a triangular lattice and contain 12 spins (see Fig. 1). Each star has two degenerate singlet ground states with different symmetry, which can be described in terms of pseudospin. It is shown that interaction between the stars leads to the band of singlet excitations that determines the low-energy KAF properties. The low-energy dynamics is described by the Hamiltonian of a spin-$\frac{1}{2}$ magnet in the external magnetic field given by Eq. (13). The Hamiltonian parameters are calculated in the first three orders of perturbation theory and are summarized in the table. Within our precision, the KAF has an ordered singlet ground state with all the stars in the state given by Eq. (3). The ground-state energy is lower than that calculated in the previous finite cluster studies. We show that our model cannot be extended directly to KAFs with $S > 1/2$.

The approach discussed in this paper can be verified experimentally under inelastic neutron scattering: the corresponding intensities for singlet–triplet transitions should have periodicity in the reciprocal space corresponding to the star lattice.

ACKNOWLEDGMENTS

We are grateful to D.N. Aristov and A.G. Yashenkin for interesting discussions. This work was supported by the Russian Foundation for Basic Research (project nos. SS-1671.2003.2, 03-02-17340, and 00-15-96814), Goskontrakt (grant no. 40.012.1.1.1149), and Russian State Programs "Quantum Macrophysics," "Collective and Quantum Effects in Condensed Matter," and "Neutron Research of Solids."

APPENDIX 1

Irreducible Representations of the Group C_{6v}

The symmetry group C_{6v} contains six rotations C^k by the angles $2\pi k/6$ ($k = 0, 1, ..., 5$) and six reflections, which can be written as $C^k u_1$, where u_1 is the operator of a reflection. One-dimensional irreducible representations can be represented as follows [18, 19]:

$$C^k \sim 1, \qquad u_1 \sim 1, \tag{A.1}$$

$$C^k \sim (-1)^k, \quad u_1 \sim 1, \tag{A.2}$$

$$C^k \sim 1, \qquad u_1 \sim -1, \tag{A.3}$$

$$C^k \sim (-1)^k, \quad u_1 \sim -1. \tag{A.4}$$

For two-dimensional representations, we have [18, 19]

$$C^k \sim \begin{pmatrix} e^{i\frac{2\pi l}{6}k} & 0 \\ 0 & e^{-i\frac{2\pi l}{6}k} \end{pmatrix}, \qquad u_1 \sim \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \tag{A.5}$$

where two inequivalent representations are given by $l = 1$ and $l = 2$.

APPENDIX 2

Star with S = 1 and S = 3/2

In this appendix, we present the details of numerical calculations showing that functions presented in Fig. 2, where the bold line depicts the singlet state of the corresponding two spins, are not ground states of the star with $S = 1$ and $S = 3/2$, as this was for $S = 1/2$.

A simple numerical method for determining the eigenvalue of a Hermitian operator H of the maximum modulus (power method [20]) was used. It is based on the following statement. We consider a state of the system $f = \sum_i c_i \psi_i$, where the sum may not include all the H eigenfunctions. For a given f, the eigenvalue E_{extr} of the maximum modulus is determined by

$$\lim_{n\to\infty} \frac{\langle f|H^{n+1}|f\rangle}{\langle f|H^n|f\rangle} = E_{\text{extr}}. \tag{B.1}$$

This becomes evident by noting that $\langle f|H^n|f\rangle = \sum_i |c_i|^2 E_i^n$.

Equation (B.1) can be used in numerical calculations in the following way. The corresponding expression is calculated for $n = 1, 2, \ldots, n_{\max}$. Convergence

can therefore be controlled by comparing results with different n. Studying a full set of vectors f and taking n_{max} large enough to match the necessary precision, one can find the eigenvalue of H with the largest modulus.

In the case of the star, the maximum eigenvalue of the Hamiltonian is $E_{max} = 18S^2J_1$ (this energy has the state in which all the spins are along the same direction) and the energy of singlet states shown in Fig. 2 is $E_{ss} = -6S(S + 1)J_1$. Because $E_{max} > |E_{ss}|$ for $S > 1/2$, we have to take $H = \mathcal{H}_0 - WI$ to investigate the lower $\mathcal{H}_0$ levels, where $\mathcal{H}_0$ is the star Hamiltonian given by Eq. (1), I is the unit matrix, and $W = (E_{max} + E_{ss})/2 + J_1$. Eigenvalues of H are therefore shifted down relative to those of $\mathcal{H}_0$ by the same value W such that the H eigenvalue with the largest modulus becomes equal to the $\mathcal{H}_0$ ground-state energy minus W.

We have not found the ground-state energy for the star with $S = 1$ and $S = 3/2$ by this method because the full set of vectors f should be examined for that. This operation requires much computer time. However, studying a number of vectors f, we have obtained that there are states lower than those discussed above, at least, by the energy $1.8J_1$. The method yielded E_{extr} with the prescribed precision to the second decimal position at $n_{max} = 100–300$ depending on f and S.

REFERENCES

1. A. P. Ramirez, B. Hessen, and M. Winklemann, Phys. Rev. Lett. **84**, 2957 (2000).
2. P. Lecheminant, B. Bernu, C. Lhuillier, *et al.*, Phys. Rev. B **56**, 2521 (1997).
3. C. Zeng and V. Elser, Phys. Rev. B **42**, 8436 (1990).
4. P. W. Leung and V. Elser, Phys. Rev. B **47**, 5459 (1993).
5. C. Waldtmann, H.-U. Everts, B. Bernu, *et al.*, Eur. Phys. J. B **2**, 501 (1998).
6. C. Zeng and V. Elser, Phys. Rev. B **51**, 8318 (1995).
7. P. Sindzingre, G. Misguich, C. Lhuillier, *et al.*, Phys. Rev. Lett. **84**, 2953 (2000).
8. J. T. Chalker and J. F. G. Eastmond, Phys. Rev. B **46**, 14 201 (1992).
9. S. Sachdev, Phys. Rev. B **45**, 12377 (1992).
10. K. Yang, L. K. Warman, and S. M. Girvin, Phys. Rev. Lett. **70**, 2641 (1993).
11. V. Kalmeyer and R. B. Laughlin, Phys. Rev. Lett. **59**, 2095 (1987).
12. J. B. Marston and C. Zeng, J. Appl. Phys. **69**, 5962 (1991).
13. F. Mila, Phys. Rev. Lett. **81**, 2356 (1998).
14. G. Misguich, D. Serban, and V. Pasquier, Phys. Rev. Lett. **89**, 137202 (2002).
15. G. Misguich, D. Serban, and V. Pasquier, Phys. Rev. B **67**, 214413 (2003).
16. M. Mambrini and F. Mila, Eur. Phys. J. B **17**, 651 (2000).
17. A. V. Syromyatnikov and S. V. Maleyev, Phys. Rev. B **66**, 132408 (2002).
18. L. D. Landau and E. M. Livshitz, *Course of Theoretical Physics*, Vol. 3: *Quantum Mechanics: Non-Relativistic Theory*, 3rd ed. (Nauka, Moscow, 1974; Pergamon Press, Oxford, 1977).
19. M. I. Petrashen' and E. D. Trifonov, *Applications of Group Theory in Quantum Mechanics* (Nauka, Moscow, 1967; Iliffe, London, 1969).
20. B. N. Parlett, *The Symmetric Eigenvalue Problem* (Prentice Hall, Englewood Cliffs, N.J., 1980; Mir, Moscow, 1983).
21. V. N. Kotov, M. E. Zhitomirsky, and O. P. Sushkov, Phys. Rev. B **63**, 064 412 (2001).
22. L. P. Regnault, M. Ain, B. Hennion, *et al.*, Phys. Rev. B **53**, 5579 (1996).

Journal of Experimental and Theoretical Physics, Vol. 98, No. 3, 2004, pp. 546–555.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 125, No. 3, 2004, pp. 618–629.

SOLIDS
Electronic Properties

Effect of Spin Structures and Nesting on the Shape of the Fermi Surface and Pseudogap Anisotropy in the *t–t'–U* Hubbard Models

A. A. Ovchinnikov† and M. Ya. Ovchinnikova

Institute of Chemical Physics, Russian Academy of Science, Moscow, 119977 Russia
e-mail: movchin@center.chph.ras.ru

Received June 2, 2003

Abstract—The effect of two types of spin structures on the shape of the Fermi surface and on the map of photoemission intensities for the *t–t'–U* Hubbard model is investigated. The stripe phase with a period of 8*a* and the spiral spin structure are calculated in the mean field approximation. It is shown that, in contrast to electron-type doping, hole-doped models are unstable to the formation of such structures. Pseudogap anisotropy is different for *h*- and *e*-doping and is determined by the spin structure. In accordance with ARPES data for $La_{2-x}Sr_xCuO_4$, the stripe phase is characterized by quasi-one-dimensional FS segments in the vicinity of points $M(\pm\pi, 0)$ and by suppression of the spectral density for $k_x = k_y$. It is shown that spiral structures exhibit polarization anisotropy: different segments of the FS correspond to electrons with different spin polarizations.

1. INTRODUCTION

Angle-resolved photoemission spectroscopy (ARPES) is an effective method for studying the electron bands and Fermi surfaces (FS) of cuprates [1, 2]. This method provides an image of the FS projection onto the CuO_2 plane. The results obtained in early studies (see [1, 2] and the literature cited therein) corresponded to hole-type FSs centered at point $Y(\pi, \pi)$ of the two-dimensional Brillouin zone. Other versions of the FS topology were also discussed later. In particular, the existence of electron-type FSs with the center at point $\Gamma(0, 0)$ was predicted for $Bi_2Sr_2CaCu_2O_{8+\delta}$ (BSCCO) [3]. The revision of the problem [4, 5] apparently confirmed the initial assumption. At the same time, a transition from the *h*-type FS to the *e*-type FS was discovered for $La_{2-x}Sr_xCuO_2$ (LSCO) upon a transition from the underdoped to overdoped region of the phase diagram [6, 7]. On the basis of ARPES data, the existence of a *d*-wave superconducting (SC) gap was proved and a pseudogap was discovered in antinodal directions in the underdoped compound BSCCO. Recently, band splitting and the FS were observed in bilayer cuprates, and the phase diagram boundary was discovered beyond which such splitting disappears [8–11]. Analysis of photoemission induced by circularly polarized light revealed a state with time-reversal symmetry breaking (TRSB) in the underdoped compound BSCCO [12].

Wide application of photoemission intensity maps in the k_x, k_y, ω space involves the determination of matrix elements and requires methods for obtaining information on the FS from the ARPES data. The problem of matrix elements was discussed in detail in [13, 14]. However, in strongly correlated systems, the topology, shape, and intensities of individual (main and shadow) segments of the FS also depend on spin and charge structures, which ensure the lowest energy.

This study is devoted to model analysis of the effect of various periodic spin structures on the FS shape and their manifestations in photoemission intensities. The analysis is based on the *t–t'–U* Hubbard model. In contrast to static structures in magnesium-based compounds, cuprates are instead characterized by dynamic structures. The lifetimes of such structures are longer than the corresponding times $t \sim 10^{-6}$–10^{-9} s in μ-SR experiments. Over short time intervals and for processes with an energy resolution of $\delta E > \hbar/t$, local spin structures can be treated in the static approximation. In this case, it is natural to find the structures that correspond to the FS features observed in the ARPES experiments.

Experimental indications of the existence of spin and charge structures in cuprates are incommensurate peaks in the spin susceptibility $\chi''(q, \omega \longrightarrow 0)$ for $q = (\pi \pm \delta q, \pi)$, $(\pi, \pi \pm \delta q)$ in LSCO, which are obtained from inelastic neutron scattering [15]; a linear structure along the CuO bonds with a period of 4*a* (*a* is the lattice constant), which was observed in BSCCO from the Fourier analysis of tunnel spectra [16]; and a periodic staggered $4a \times 4a$ structure around vortices in the mixed state of BSCCO [17].

In this study, we analyze the FS topology and the intensity maps of photoemission spectra for a number

† Deceased.

of possible spin structures in the t–t'–U Hubbard models. We will use the extremely simple language of the mean field (MF) method for interpreting the pseudogap (PG) to give examples of structures with different types of PG anisotropy and to consider some properties of photoemission of cuprates in light of the results obtained for periodic structures. Earlier [18], current states of the type of an orbital ferromagnet were proposes as latent order parameters (OP) responsible for the emergence of the pseudogap. In this work, the range of possible latent OPs is extended to stripe structures and spiral spin structures. Such structures are stabilized due to the removal of degeneracy in the states of "hot spots," i.e., Van Hove singularities (VHS) of spectra or the removal of degeneracy in bands located on parallel FS segments in the case of nesting.

The article has the following structure. In Section 2, the t–t'–U Hubbard model is considered and a simplifies mean-field approximation is used for describing the normal sate. The MF equations are formulated for an arbitrary periodic structure with spin density waves (SDW) and charge density waves (CDW). The methods for visualizing the FS and the intensity map of ARPES spectra are considered in Section 3. Typical examples of FS and ARPES intensity maps are given in Section 4 for homogeneous MF solutions with the AF spin order. On such solutions, the emergence of the pseudogap with various types of anisotropy in h- and e-doped systems is traced. The results obtained for inhomogeneous spin and charge structures are considered. These structures include stripe phases of antiphase AF domains along the y bonds with a domain width of $4a$, spiral spin states, and periodic 1D and 2D structures with charge modulation. In Section 5, various degrees of stability of homogeneous AF solutions relative to the formation of stripe and spiral spin structures for e- and h-doped models are demonstrated. Features of photoemission for these structures, the methods for their testing, and the correspondence of these structures to some ARPES data for cuprates are discussed in Sections 6 and 7.

2. PERIODIC SOLUTIONS OF THE HUBBARD MODEL IN THE FRAMEWORK OF THE MEAN FIELD METHOD

We will study the effect of nesting and the formation of periodic structures on the FS and the energy bands in the normal state of cuprates by applying the MF method to the initial Hamiltonian of the t–t'–U Hubbard model,

$$H = T + \sum_n U n_{n\uparrow} n_{n\downarrow}, \quad T = \sum_{\sigma, k} \epsilon_k c^\dagger_{k\sigma} c_{k\sigma}, \tag{1}$$

$$\epsilon_k = 2t(\cos k_x + \cos k_y) + 4t' \cos k_x \cos k_y. \tag{2}$$

We will assume that $t = 1$ and will measure all energies and parameters U and t' in units of t. Hamiltonian (1) is insufficient for describing the SC order in the MF approximation. In accordance with some approaches, the SC state arises due to attraction between electrons of neighboring sites, which is of the correlation type. The interaction of the type of correlated jumps in the effective Hamiltonian was derived, for example, in [19, 20]. In the empirical version, this can be, for example, Hamitonian (1) supplemented with the interaction between nearest neighbors of the form

$$V = \kappa \sum_{\langle nm \rangle, \sigma} c^\dagger_{n\sigma} c^\dagger_{m,-\sigma} c_{m,-\sigma} c_{n,\sigma} \tag{3}$$

with $\kappa < 0$. In this study, however, we will analyze only the normal state and confine the analysis to the MF treatment of the initial Hamiltonian (1).

Let us consider a periodic structure with the 2D translation vectors

$$E_1 = (E_{1x}, E_{1y}), \quad E_2 = (E_{2x}, E_{2y}). \tag{4}$$

Suppose that a unit cell contains n_c centers with coordinates $j = (j_x, j_y)$ so that an arbitrary site of the 2D lattice,

$$n = n(L, j) = (n_1, n_2) = E_1 L_1 + E_2 L_2 + (j_x, j_y) \tag{5}$$

is described by integers L_1, L_2 (unit cell coordinates) and by numbers $j = (j_x, j_y)$ fixing a lattice site. Two-dimensional vectors B_1, B_2 of the reciprocal lattice satisfy the equations $E_i B_j = 2\pi\delta_{ij}$. (Components E_i and B_j are given in units of constants of the direct and reciprocal lattices.)

We denote by $\tilde{k}$ the quasimomentum within the principal Brillouin zone $\tilde{G}$ of the periodic structure in contrast to quasimomentum k varying within the Brillouin zone G of the initial lattice. The areas of $\tilde{G}$ and G amount to $4\pi^2/n_c$ and $4\pi^2$ and are limited by the conditions

$$\tilde{k} \in \tilde{G} : |\tilde{k} B_i| \le \pi; \quad k \in G : |k_{x(y)}| \le \pi. \tag{6}$$

The order parameters for the periodic MF solutions are the electron densities and the vectors of average spin for each site of the unit cell,

$$r_j = \frac{1}{N_L} \sum_L \langle r_{n(L,j)} \rangle, \quad S_{\mu j} = \frac{1}{N_L} \sum_L \langle S_{\mu, n(L,j)} \rangle. \tag{7}$$

Here, index μ enumerates the spin vector components and $N_L = N/n_c$ is the number of unit cells, n_c being the number of centers in a cell.

In the MF approximation, the mean energy in model (1) is given by

$$\overline{H} = \langle T \rangle + N_L U \sum_j \left(r_j^2 - \sum_\mu S_{j\mu}^2 \right) \tag{8}$$

and the wave function is determined by the population

of the one-electron eigenstates $\chi^{\dagger}_{k\lambda}$ of the linearized Hamiltonian

$$H_{lin} = T + N_L\sum_j\{2Ur_j\hat{r}_j - 2US_{\mu j}\hat{S}_{\mu j}\} = \sum_{\tilde{k}\in\tilde{G}}\hat{h}_{\tilde{k}}. \quad (9)$$

The latter can be split into independent contributions for each value of the reduced quasimomentum $\tilde{k}$. Here, $\hat{r}_j$ and $\hat{S}_{\mu j}$ are the operators corresponding to one-electron mean values (7). In the momentum representation, the eigenstates of Hamiltonian (9) can be expanded in the set of $2n_c$ Fermi operators,

$$\chi^{\dagger}_{\tilde{k}\lambda} = \sum_{m,\sigma} c^{\dagger}_{\tilde{k}+Bm,\sigma}W_{m\sigma,\lambda}(\tilde{k}),$$
$$m = (m_1, m_2); \quad Bm = B_1m_1 + B_2m_2; \quad (10)$$
$$\lambda = 1, \ldots, 2n_c.$$

Here, $\tilde{k}$ varies within $\tilde{G}$ and the set of pairs of integers (m_1, m_2) is such that vectors $\tilde{k} + Bm$ cover the entire phase space G. Matrix $W_{m\sigma,\lambda}$ of eigenvectors and vector $E_{\tilde{k},\lambda}$ of eigenvalues are determined by diagonalozation of matrix $h_{\tilde{k}}$ in the basis $\{c^{\dagger}_{\tilde{k}+Bm,\sigma}\}$:

$$(h_{\tilde{k}})_{m\sigma,m',\sigma'}W_{m'\sigma',\lambda} = W_{m\sigma,\lambda}E_{\tilde{k},\lambda}. \quad (11)$$

Here, we have

$$(h_{\tilde{k}})_{m\sigma,m',\sigma'} = \delta_{mm'}\delta_{\sigma\sigma'}\epsilon_{\tilde{k}+Bm}$$
$$+ U\sum_j\varphi(j, m'-m)[r_j\delta_{\sigma\sigma'} - S_{\mu j}(\sigma_\mu)_{\sigma\sigma'}], \quad (12)$$

$$\varphi(j, m) = \exp[i(j - j_0)Bm],$$
$$Bm = B_1m_1 + B_2m_2. \quad (13)$$

The choice of j_0 (the origin from which the sites in a cell are counted) is arbitrary and affects only the phases. In turn, order parameters (7) themselves can be calculated in terms of the eigenvector matrix W and Fermi functions f:

$$\{r_j, S_{\mu j}\} = \frac{1}{2N}\sum_{\tilde{k}\in\tilde{G}}\sum_{m,s,m',s'}\{\sigma_0, \sigma_\mu\}_{ss'}\varphi(j, m'-m)$$
$$\times W^*_{ms,\lambda}(\tilde{k})W_{m's',\lambda}(\tilde{k})f(E_{\tilde{k}\lambda} - \mu). \quad (14)$$

The Pauli matrices σ_μ and σ_0 in this expression correspond to components $S_{\mu j}$ and r_j, respectively. Equations (11)–(14) determine the self-consistent solutions in the MF approximation.

For a periodic structure with a definite symmetry, the combination of equivalent atoms into groups considerably reduces the number of independent order parameters and often reduces matrix (12) to the real matrix when j_0 is chosen at the symmetry center of the unit cell of the structure.

3. METHODS FOR VISUALIZING THE FERMI SURFACE AND THE GENERALIZED FERMI SURFACE

The Fermi boundaries and the quantity that can be called the generalized Fermi boundary (GFS) can be determined from the ARPES data using several methods. These methods are considered in detail in [13], and each method can be put in correspondence to a method of constructing FS and GFS maps in model calculations.

One of these methods employs the photoemission intensity map $I(k, \omega)$ for electrons with momentum component k in the *ab* plane and with energy $E_e = h\nu - \omega$:

$$I(k, \omega) = |M(k)|^2A(k\omega)f(\omega)\otimes R_{\omega k}, \quad (15)$$

$$A(k\omega) = \frac{1}{Z}\sum_{\alpha,\gamma}|\langle\gamma|c_{k\sigma}|\alpha\rangle|^2$$
$$\times e^{\beta(E_\alpha - \mu N)}\delta(E_\gamma - E_\alpha - \mu - h\omega). \quad (16)$$

Intensity (15) is determined by the product of the squared matrix element $M(k)$, spectral density $A(k\omega)$, and Fermi function f. In order to compare the intensity with the observed ARPES signal, a convolution of the product with the Gaussian function $R_{\omega k}$ [13] with parameters characterizing the energy and momentum resolution is usually carried out in Eq. (15). In Eq. (16), α and γ characterize the states of the entire system, $\beta = 1/kT$, and μ is the chemical potential. The dependence of matrix element M on k and its role were studied in [13, 14]. Here, we assume for simplicity that the matrix element is constant since we are interested in the effect of periodic structures and transition processes on spectral density $A(k\omega)$ and the photoemission intensity.

In the one-electron MF approximation, we have

$$A(k\omega) = \frac{1}{N}\sum_{\tilde{k}\in\tilde{G}}\sum_{m,\sigma,\lambda}\left|W_{m\sigma,\lambda}(\tilde{k})\right|^2$$
$$\times\bar{\delta}(E_{\tilde{k}\lambda} - \mu - \omega)\delta_{k,\tilde{k}+Bm}. \quad (17)$$

Two-dimensional index $m = (m_1, m_2)$ passes through all independent transition vectors $Bm = B_1m_1 + B_2m_2$; $\lambda = 1, \ldots, 2n_c$ enumerates proper Fermi operators of linearized Hamiltonian $h_{\tilde{k}}$ with reduces momentum $\tilde{k}$. We calculate function $A(k\omega)$ using the standard substitution of the δ function in Eq. (17) by a function of a finite width Ω, e.g., of the form $\bar{\delta}(\omega) = \cosh^{-2}(\omega/\Omega)$. Another method for introducing broadening is to use

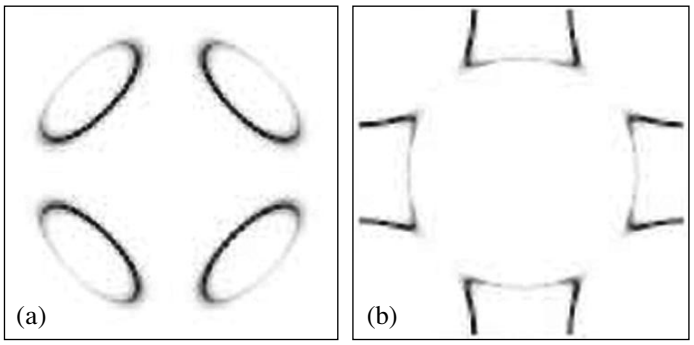

Fig. 1. Fermi surfaces for homogeneous AF solutions in hole- (a) and electron-doped (b) models with parameters $U = 4$, $t' = 0.3$, $\delta = |1 - n| = 0.2$ in the underdoped region on plane $-\pi < k_x, k_y < \pi$. Hole and electron pockets are located around points $(\pm\pi/2, \pm\pi/2)$ or $(\pm\pi, \pm\pi)$.

the self-energy part chosen empirically in [13]. The construction of map $I(k_x, k_y, \omega = 0)$ makes it possible to visualize both the main and shadow segments of the Fermi boundary, which are manifested with a lower or higher intensity.

It should be noted that band energies are periodic in the k space: $E_{\tilde{k}+Bm,\lambda} = E_{\tilde{k},\lambda}$ for any $m = (m_1, m_2)$. However, quantity $|W_{m\sigma,\lambda}(\tilde{k})|^2$ in spectral function (17) and, accordingly, photoemission intensity (15) do not possess such a periodicity. For this reason, different FS segments are manifested with different intensities even for the matrix element in Eq. (15) independent of k in view of the compound nature of band operators (10) in the presence of SDW and CDW structures (i.e., transition processes). In calculating $I(k, \omega)$, the correspondence of a given vector k to quantities $\tilde{k}$ and $m = (m_1, m_2)$ is determined by the equation $k = \tilde{k}(k) + m_1B_1 + m_2B_2$.

Other methods, which are adequate to the processing of ARPES data, were also proposed for visualizing the FS in model calculations. One of such methods employs the construction of the map of gradient $g_k = |\nabla_k \bar{n}_k|$ of smoothed occupancy $\bar{n}_k = n_k \otimes R_k$. It is also possible to construct the maps of intensity averaged over a definite frequency window $2\Delta\omega$:

$$I_{\Delta\omega}(k, \omega) = \int d\omega' I(k, \omega') R\left(\frac{\omega - \omega'}{\Delta\omega}\right). \qquad (18)$$

Here, R is the corresponding Gaussian function of width $\Delta\omega$, which imitates a finite resolution in ω. The construction of such maps involves the normalization of a function to its maximum value. Consequently, the brightness and width of the Fermi boundaries on such maps is determined by the width of smoothing function R_k or frequency window $\Delta\omega$ in Eq. (18). In particular, for a large width $\Delta\omega$, a map of functions $I_{\Delta\omega}$ or g_k shows not only the actual Fermi boundaries with a sharp occupancy step, but also the boundaries with a substantial but smoother variation of $n(k)$. The connection between such segments and the emergence of a dielectric gap and anisotropic pseudogaps will be illustrated below.

4. TOPOLOGY OF THE FERMI SURFACE AND GENERALIZED FERMI SURFACE FOR HOMOGENEOUS ANTIFERROMAGNETIC MEAN-FIELD SOLUTIONS

In this section, we partly repeat the well-known results obtained in [21–28]. Homogeneous AF solutions in the MF approximation are characterized by the average alternating spin $d_0 = (-1)^{n_x + n_y}\langle S_{nz}\rangle$. The corresponding magnetic Brillouin zone $\tilde{G}_{AF}$ is confined by the limits $|k_x \pm k_y| \le \pi$. The known energies of the upper and lower Hubbard bands are given by

$$\epsilon_{\tilde{k}\lambda} = 4t'\cos k_x \cos k_y \pm \sqrt{U^2 d_0^2 + 4t^2(\cos k_x + \cos k_y)^2} + \text{const}. \qquad (19)$$

The Van Hove singularity in the density of states (DOS) of the lower Hubbard band corresponds to hot spots $M = (\pm\pi, 0), (0, \pm\pi)$. The shape of the FS critically depends of parameter t' and on the doping level. For $t' = 0$, the band energy is constant for all values of k along the boundary of the magnetic Brillouin zone. For $t' > 0$, energy $\epsilon_{k,\lambda=1}$ at points M is lower than at diagonal points $S(\pm\pi/2, \pm\pi/2)$. For this reason, in accordance with the well-known pattern, the FS bounds hole pockets centered at point S for a low doping level. Figure 1a is a map of intensity $I(k, \omega = 0)$ calculated by formulas (15) and (17). Since band energy ϵ_M at point M lies below the chemical potential, photoemission of electrons from such segments with $k = k_M$ requires an energy of $\Delta_{PG} = (\mu - \epsilon_M) > 0$, which is equal to the work function. Accordingly, the curve describing the energy distribution of photoelectrons (EDC) will be shifted by $-\Delta_{PG}$. Such a shift indicates the emergence of a pseudogap in the normal state in directions $k \sim k_M$ in the

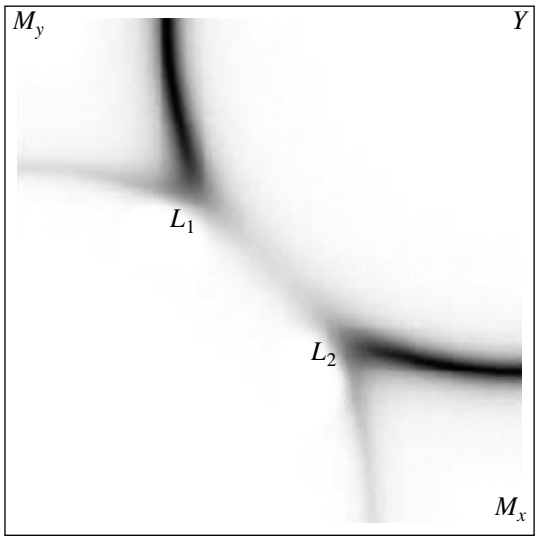

Fig. 2. Map of averaged intensity (18) with $\Delta\omega = 0.05t$ on plane $0 < k_x, k_y < \pi$ for the same electron-doped model as in Fig. 1b. Line L_1–L_2 of the generalized Fermi boundary corresponds to energies below the Fermi level, i.e., to the emergence of a PG with anisotropy differing from anisotropy of the PG in BSCCO.

given homogeneous AF solutions. For small values of $t'/t \sim 0.1$, the pseudogap disappears when the doping level increases to a certain value of δ_{opt}. At this instant, the chemical potential passes through a VHS in the DOS. The peak of the DOS at the Fermi level for $\delta = \delta_{opt}$ ensures the maximal value of T_c for such a doping level. Simultaneously, the FS changes its topology at $\delta = \delta_{opt}$, being transformed into a single large FS with electron-type segments in the vicinity of hot spots. In this case, optimal doping level δ_{opt} increases with t'. In [21–23, 25, 27], such a behavior of phase curves $T_c(\delta)$ is described and the geometrical interpretation of the PG is given for the t–t'–U and t–t'–J models on the basis of more rigorous approaches. In these approaches, the value of δ_{opt} is smaller than in the simple MF approximation.

When the effective attraction between electrons from neighboring sites is included, homogeneous MF solutions lead to an SC order with the d symmetry, which is compatible with a local AF order [22, 23]. For underdoped cuprates, the total gap (the shift of the EDC edge in ARPES), which depends on the PG and SC gaps in accordance with the relation

$$\Delta = \sqrt{\Delta_{PG}^2 + \Delta_{SC}^2}, \qquad (20)$$

explains [26, 27] the nonlinear dependence of gap $\Delta(z)$ on $z = \cos k_x - \cos k_y$ and other features of the observed gap [29–31].

In the case of electron doping, a similar analysis shows that electron pockets appear around points M in underdoped models (Fig. 1b). The main (nonshadow) segments of the FS in this case face point $Y(\pi, \pi)$.

Figure 2 shows the map of averaged intensity (18) with a frequency window width of $\Delta\omega = 0.05t$. As $\Delta\omega \longrightarrow 0$, the maximum $\max I(k, \omega) \longrightarrow \infty$ on the true Fermi boundaries, and segment L_1–L_2 on the normalized intensity map becomes invisible. In addition to the FS around the electron pockets, we can see boundary L_1–L_2 corresponding to the maximum of the occupancy gradient in the diagonal direction. On this segment of the boundary, we have $\epsilon_{k,\lambda=2} - \mu < 0$. This indicates the emergence of a PG on segment L_1–L_2 around points $k_S = (\pm\pi/2, \pm\pi/2)$. The neighborhoods of these points are responsible for the formation of VHS in the DOS of the upper Hubbard band in a model with $t' > 0$. Further e-doping transforms the FS for homogeneous MF solutions into a single large surface around point $Y(\pi, \pi)$. As a result of inclusion of a pairing interaction of type (3) and d superconductivity combined with nonstandard PG anisotropy for the underdoped region, the minimal energy Δ_{min} of Fermi excitations does not vanish even in the nodal d directions of the SC gap:

$$\Delta_{min} = \min\left(\sqrt{\Delta_{PG}^2 + \Delta_{SC}^2}\right) \neq 0. \qquad (21)$$

As regards the behavior of such quantities as heat capacity and $\lambda^{-2}(T)$ for $T \longrightarrow 0$, the finite value of the minimal gap for Fermi excitations is perceived as the generalized s symmetry of the SC order. In [32], the crossover from some properties characteristic of d symmetry to those typical of s symmetry of the SC order was considered for n-type cuprates $X_{2-x}Ce_xCuO_4$, X = Nd and Pr. Homogeneous model solutions predict reverse evolution of these properties from those typical of s-type superconductors to the properties characteristic of d-type superconductors upon an increase in the doping level. In spite of this contradiction, the solution is interesting as an example of the fact that d superconductivity combined with nonstandard PG anisotropy may imitate in some properties the SC order with the generalized s symmetry.

The electron pockets around points M in n-type underdoped cuprates, which were predicted for the first time in [22, 24], were indeed observed on the ARPES intensity maps for underdoped compounds NCCO [33]. The evolution of FSs from nonconnected small FSs to a large FS of the hole type around point $Y(\pi, \pi)$ was traced. The parameters of the t–t'–t''–U model reproducing the observed evolution were selected in [28]. Among other things, the assumption that effective potential $U = U(\delta)$ decreases with increasing doping level had to be made. According to [28], such a screening of U combined with higher harmonics ($\sim t''$) in the band energy leads to the simultaneous formation of hole pockets around point $k = (\pi/2, \pi/2)$ and electron pockets around k_M from the lower and upper Hubbard bands, respectively. For $\delta \sim 0.2$, the Mott gap in compounds NCCO is closed [28]. The emergence of a PG in

diagonal directions in underdoped NCCO compounds is also confirmed by Raman scattering data [34].

Proceeding from the FS shape in underdoped models, general considerations concerning a possible symmetry of the SC order in these compounds can be formulated. Suppose that the SC transition is caused by attraction of type (3) for electrons at the neighboring sites of bonds with the x and y orientations. This interaction could lead to states $\sum_k \varphi^{s(d)}(k) c^\dagger_{k\uparrow} c^\dagger_{k\downarrow}$ of coupled pairs with $\varphi^{s(d)}(k) = \cos k_x \pm \cos k_y$ of generalized s or d symmetry. However, the state of a coupled pair must be orthogonal (or nearly orthogonal) to one-site state $c^\dagger_{n\uparrow} c^\dagger_{n\downarrow}$ of the pair, which is suppressed by one-center repulsion U. In underdoped model with h or e doping, in which pockets of only one (electron or hole) type are present, the d symmetry is the only possibility to achieve orthogonality of the function of a pair to the one-center function in view of orthogonality of the angular parts of these functions. Indeed, the main (nonshadow) segments of the FS in this case are either completely inside the magnetic Brillouin zone boundary (see Fig. 1a) or completely outside this boundary (see Fig. 1b). Consequently, the s function $\varphi^s(k) = \cos k_x + \cos k_y$ of a pair does not change its sign on the main regions of the FS and, hence, cannot be orthogonal to the one-center function of the pair.

The situation changes for $\delta > 0.15$ in the electron-doped models with the parameters selected in [28]. In this case, electron pockets around point M due to filling of the upper Hubbard band coexist with hole pockets around point $S(\pm\pi/2, \pm\pi/2)$ due to partial depletion of the lower Hubbard band. As a result, the main segments of the FS lie partly outside and partly inside the magnetic Brillouin zone boundary. Such segments correspond to different signs of pair function $\varphi^s(k) = \cos k_x + \cos k_y$; i.e., the orthogonality of the pair s-wave function to the one-center function can be ensured due to orthogonality of the "radial" parts of the functions if we conditionally refer to quantity $z = \cos k_x + \cos k_y$ as a radial variable. It was very important to verify experimentally whether such an SC order of s symmetry is realized in NCCO and PCCO compounds for $\delta > 0.15$. Analogous situations and problems may also arise during the formation of periodic SDW and CDW structures since the energy profile along the magnetic Brillouin zone boundary changes not only upon a change in t' and t'', but also as a result of the formation of these structures.

5. INSTABILITY TO THE FORMATION OF SPIN STRUCTURES IN n- AND p-TYPE CUPRATES

The degree of instability of homogeneous AF states to the formation of periodic spin and charge structures in doped models can be estimated on the basis of MF calculations. Figure 3 shows the dependences of mean energy $\bar{H} = \langle H \rangle$ on doping level $\delta = |1 - n|$ for the normal state in hole- and electron-doped models for a number of structures in comparison with the energy of homogeneous paramagnetic (PM) and AF solutions. In addition to the latter solutions, MF solutions were obtained (or sought) for the following structures.

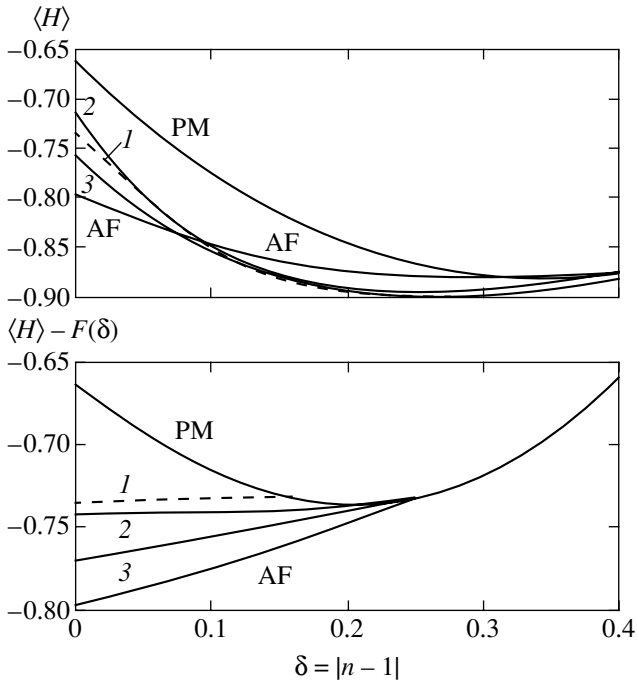

Fig. 3. (top) Dependence of the mean energy (per site) in the hole-doped t–t'–U model with parameters $U = 4.0$ and $t' = 0.3$ for different spin structures: for a PM state, for a homogeneous AF state, for a stripe phase with a period of $8a$ of the structure along the x axis (dashed curve), and for spiral states with $Q = \pi(\eta, 1)$ (curve *2*) and $Q = \pi(\eta, \eta)$ (curve *3*) for $\eta = 0.8$. (bottom) The energies of the same structures in the electron-doped model. For convenience of representation, common function $F(\delta) = U(n^2 - 1)$ is subtracted from all energies.

1. A stripe structure consisting of antiphase AF stripes parallel to the y axis with domain walls at the bonds. The structure is characterized by a unit cell with eight centers and vectors $E_{1,2} = (4a, \pm a)$. For analogous structures with domain walls passing through lattice sites, the mean energy is close to but slightly higher than the energy of the first structure. The Fermi surfaces of these structures are also similar; for this reason, we will consider only the former structure.

2. Spiral spin structures with

$$\langle S_n \rangle = d_0 [\mathbf{e}_x \cos Qn + \mathbf{e}_y \sin Qn], \qquad (22)$$

$Q = Q_x = \pi(\eta, 1)$ or $Q = Q_{xy} = \pi(\eta, \eta)$ with $\eta \sim 0.75$–0.8.

3. Staggered structures with antiphase square $4a \times 4a$ domains. We do not consider the data on these struc-

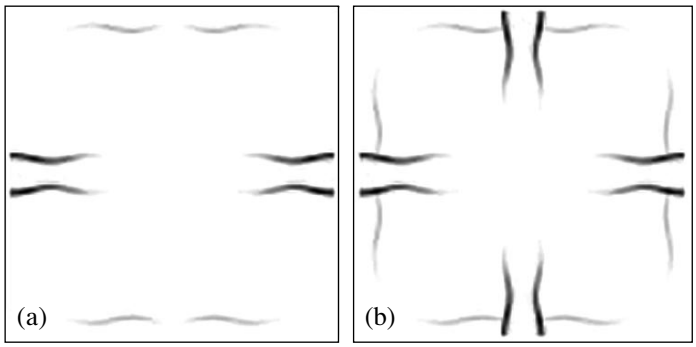

Fig. 4. (a) Photoemission intensity map $I(k, \omega = 0)$ for a stripe structure with AF stripes of width $4a$, parallel to the y axis, and with domain walls centered at bonds. (b) The same averaged over two stripe structures with the x and y orientations. The model parameters are $U = 6$ and $t' = 0.1$.

tures since their energies are higher than the energies of the stripe structures and their FSs have a nonrealistic form.

4. The current states of an orbital antiferromagnet, which were proposed in [18], cannot be realized in model (1) in the MF approximation without using additional interactions with $U/t = 4$–6. In the case of hole-doping, analogous solutions with alternating spin currents in plaquets and with the rotation of the local spin by $\pi/2$ upon a transition between neighboring sites along the perimeter of a plaquet exist only for large values of U ($U/t \geq 5$). The energies of such structures are higher than the energy for stripe and spiral structures and the shape of their FSs is nonrealistic. For this reason, the data on these structures are not given here.

The calculations were made mainly for structures with a fixed (commensurate) period of $8a$, although the optimal size of domains or vector $\mathbf{Q}$ for the spiral state depends on the doping level. Dependences $Q(\delta)$ for the spiral state were calculated by many authors (see, for example, [35]). However, we are interested here in the typical features of FSs and anisotropy of PGs for definite structures.

It can be seen from Fig. 3a that, in the case of h doping, homogeneous AF states are unstable to the formation of stripe structures and spiral spin states. The formation of such structures extends the doping region in which nonzero values of local spin are preserved ($\langle S_n \rangle \neq 0$); vanishing of these values corresponds to merging of energy $\bar{H}(\delta)$ for a given structure with the energy of the PM state. At the same time, in the case of electron doping, the MF energies of all the structures listed above were found to be higher than the energy of the homogeneous AF state. The higher stability of the AF state with e doping corresponds to a broader doping region of the long-range AF order in n-type cuprates. The peak in the spin susceptibility for $Q \sim (\pi, \pi)$ in these cuprates (in contrast to incommensurate values of Q in p-type cuprates) also indicates the absence of stripe phases or spiral spin structures in these compounds.

6. FERMI SURFACES FOR STRIPE STRUCTURES: ANALOGY WITH THE DATA ON COMPOUND LSCO

Figure 4a shows the intensity map $I(k, \omega = 0)$ calculated by formulas (15) and (17) for a stripe structure. The structure consists of AF stripes parallel to the y axis for a model with $U/t = 6$, $t'/t = 0.1$ for $n = 0.8$. In the case of the homogeneous AF solution, this doping is close to that for which the Fermi level passes through a VHS in the DOS of the lower Hubbard band. The stripe structure parallel to the y axis splits the VHS, makes hot spots $M_x = (\pi, 0)$ and $M_y = (0, \pi)$ nonequivalent, and forms the main and shadow horizontal segments of the FS. Figure 4b shows an intensity map symmetrized in structures with domains of the x and y orientations. The FS shape is analogous to that obtained in other calculations of stripe phases [36]. It differs considerably from the FS typical of homogeneous AF solutions in a model with $t' > 0$. The main difference, i.e., the absence of the Fermi boundary in the diagonal direction, indicates the emergence of a PG at $k_x = \pm k_y$. As a consequence, we can expect that, even in the case of the d symmetry of the SC state, the minimal energy of Fermi excitations differs from zero in accordance with an equation similar to Eq. (20) for homogeneous AF solutions with e doping. As the value of t' increases to 0.3, small hole pockets are formed additionally at points $k_S = (\pi/2, \pi/2)$. The major part of the previous zero Fermi boundary, in particular, at point $M_y = (0, \pi)$, becomes a PG and dielectrized Fermi boundary; only quasi-one-dimensional segments of the FS, which are normal to the directions of stripes (y axis), are preserved. Accordingly, a quasi-one-dimensional conductivity of such a structure can also be expected. Invisible PG segments of the generalized Fermi boundary can be visualized while constructing smoothed intensity (18) with a large frequency window $\Delta\omega$.

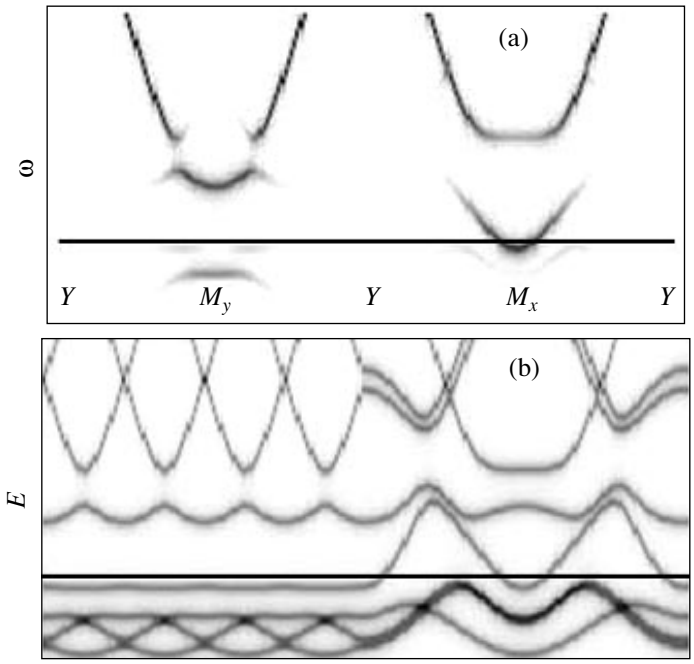

Fig. 5. (a) Self-energies $E_\lambda(k)$ in the MF problem as functions of quasimomentum varying along the contour $Y(-\pi, \pi)$–$M_y(0, \pi)$–$Y(\pi, \pi)$–$M_x(\pi, 0)$–$Y(\pi, -\pi)$ for a stripe structure with a period of $8a$ along the x axis. (b) Map of intensities $I(k, \omega)$ for k varying along the same contour. The map reveals the same energy levels $E_\lambda(k)$, but with different weights determined by the structure of band states. The parameters of the model are the same as in Fig. 4. The Fermi level is marked by the horizontal line.

Figure 5a shows one-electron energies $E_\lambda(k)$ (eigenvalues of the MF problem) as functions of the quasimomentum varying along the contour

$$Y(-\pi, \pi) - M_y(0, \pi) - Y(\pi, \pi) - M_x(\pi, 0) - Y(\pi, -\pi).$$

As usual, band energies $E_\lambda(k)$ are periodic functions with a period of $\pi/4$ on the first (horizontal) segment of contour $Y - M_y - Y$. However, the intensity map reveals only nonshadow energy levels for a given k. Figure 5b represents such a map on plane k, ω for k varying along the same contour. The intersection of energy levels with the Fermi level in the vicinity of M_x corresponds to the quasi-one-dimensional FS shown in Fig. 4a.

The conservation of the FS in the vicinity of M_x and the formation of a pseudogap in the vicinity of M_y and for $k_x \sim k_y$ are due to the action of the spin-dependent mean field; the principal harmonic of this field is $F(n) \propto \cos Q_\eta n$ with $Q_\eta = (\eta\pi, \pi)$ (here, $\eta = 0.75$). This field elevates the zero level $\epsilon(0, \pi)$ at point M_y, repelling it from lower zero levels ϵ_k at point $k = (\pm\eta\pi, 0)$ in the vicinity of M_x. The same field lowers (below the chemical potential) zero level $\epsilon(\pi, 0)$ at point M_x, repelling it from higher zero levels $\epsilon(\pm 0.25\pi, \pi)$ near M_y.

The ARPES data for underdoped LSCO compound [6, 7, 37] are in qualitative agreement with the features of the intensity maps represented in Fig. 4. The presence of two segments of the FS with different properties (in the vicinity of points M and in diagonal directions), systematic suppression of the spectral weight in the vicinity of point $(\pi/2, \pi/2)$ as compared to BSCCO compounds or with overdoped LSCO samples, and straight FS segments near point $(\pi, 0)$ with a width $\sim\pi/2$ were also interpreted as the proof of the existence of inhomogeneous structures in underdoped compound LSCO [6, 7, 37] (in particular, the combination of order–disorder with stripe structures [38]). Another proof of the existence of stripe structures is the observation of neutron scattering peaks for incommensurate momenta $(\pi \pm \delta, \pi)$ and $(\pi, \pi \pm \delta)$ in compound LSCO [15].

Anisotropy of the FS (see Fig. 4a) suggests a revision of the admissible symmetry of SC order compatible with a stripe structure. For structures symmetric relative to transposition of axes $x \longleftrightarrow y$, the d-wave SC order $\langle c^\dagger_{k,\uparrow} c^\dagger_{-k,\downarrow} \rangle \propto (\cos k_x - \cos k_y)$ is expected. It ensures the orthogonality of the pair function to one-center pair function $c^\dagger_{n\uparrow} c^\dagger_{n\downarrow}$ suppressed by one-center repulsion U. However, in the case of the quasi-one-dimension FS depicted in Fig. 4a, the generalized s-wave pair function $\langle c^\dagger_{k,\uparrow} c^\dagger_{-k,\downarrow} \rangle \propto (\cos k_x + \cos k_y)$ is more probable. The latter function may be orthogonal to the one-center pair function due to nodal lines $k_x \pm k_y = \pm\pi$. The role of nodal lines in the angular dependence of the SC gap will now be played by the nodal lines of the "radial" part of the pair function if quantity $z = \cos k_x + \cos k_y$ can be treated as the radial variable.

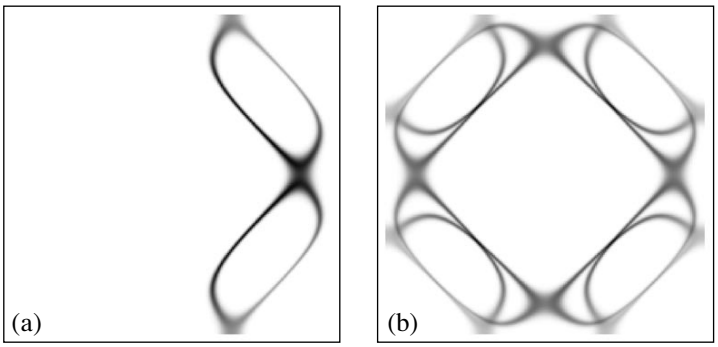

Fig. 6. (a) Photoemission intensity map $I_\uparrow(k, \omega = 0)$ for electrons with spin polarization $\sigma = \uparrow$ for a spiral structure with $Q = \pi(0.8, 1)$. (b) The same averaged over two spin polarizations $\sigma = \uparrow, \downarrow$ and two structures with $Q = (0.8\pi, \pi)$ and $Q = (\pi, 0.8\pi)$ in the model with parameters $U = 6$ and $t' = 0.1$.

Verification of this hypothesis requires appropriate calculations.

7. SPIRAL SPIN STRUCTURES AND SPIN POLARIZATION OF VARIOUS SEGMENTS OF THE FERMI SURFACE

It was found earlier [39] that spiral spin structures (22) exhibit polarization anisotropy of the FS. Different segments of the FS correspond to electrons with different predominant spin polarizations. Indeed, the mean field from a spiral spin structure mixes one-electrons states $\{c^\dagger_{k,\uparrow}, c^\dagger_{k+Q,\downarrow}\}$ and leads to splitting of VHS as in the case of stripe structures. However, in contrast to stripe structures, occupancies $n_{k,\sigma} = \langle c^\dagger_{k\sigma} c_{k\sigma}\rangle$ and intensities $I_\sigma(k\omega)$ of photoemission of electrons with a fixed polarization σ depend on the spin polarization σ. Quantities $I_\sigma(k\omega)$ are defined by formulas (15) and (17), but do not contain summation over σ on the right-hand side of Eq. (17).

Figure 6 shows the intensity map $I_{\sigma=\uparrow}(k, \omega = 0)$ for a spiral state with $Q = (0.8\pi, \pi)$ for spin polarization $\sigma = \uparrow$ on the z axis perpendicular to the plane of rotation of mean spins with a spiral configuration. Thus, FSs revealed in photoemission of electrons with polarization $\sigma = \uparrow$ exhibit anisotropy. For the opposite spin polarization, the FS is the reflection of the FS depicted in Fig. 6a in the plane $x = 0$. Figure 6b shows the Fermi surfaces symmetrized in spins and in two types of structures with $Q = (0.8\pi, \pi)$ and $Q = (\pi, 0.8\pi)$. In the vicinity of point M, the symmetrized FSs have the intersection with M–Y line (typical of hole-type FS) as well as with $M - \Gamma$ line (typical of the electron-type FS). Such double intersections were apparently observed in the ARPES spectra for BSCCO [1, 2, 40]. Direct comparison with experiment is impossible since the two-layer splitting is disregarded in the model.

Polarization anisotropy of the FS directly reflects the presence of spin currents $J_\uparrow = -J_\downarrow$. According to [39], this anisotropy could be responsible for the effect of time-reversal symmetry breaking (TRSB) in the dichroism of the ARPES signal observed for the underdoped compound BSCCO [12]. (In the alternative hypothesis [12, 41], the TRSB effect is explained by a peculiar aligning of circular microcurrents.) A direct observation of polarization anisotropy of the FS requires selective (in spin polarization) measurement of the photoemission intensity. In the measurements of the total photoemission intensity, such a selectivity was achieved in the so-called spin-orbit photoemission [42]. Such measurements are also possible in principle in ARPES. It is important to continue the study of TRSB in the underdoped BSCCO compound; in particular, it is expedient to find out whether this effect and the current associated with it are of volume or surface nature. For the time being, it remains unclear whether the ground state in BSCCO can be presented as a set of quasi-static domains with a spiral structure and with a system of spin currents associated with this set.

8. CONCLUSIONS

The MF analysis of the normal state in the t–t'–U Hubbard model revealed that the FS topology and PG anisotropy in the underdoped region depend on the sign of t', the type of (e- or h-) doping, and the spin structure. In hole-doped models, the homogeneous AF mean-field solution is found to be unstable to the formation of two types of spin structures, viz., the stripe phase and the spiral spin structure. However, in models with electron doping, the homogeneous AF solution remains the lowest among the solutions considered. This corresponds to a large (in doping level) region of existence of the AF order in n-type cuprates (such as NCCO and PCCO) as well as to the commensurate peak observed in neutron scattering for these compounds at $Q = (\pi, \pi)$. For homogeneous AF solutions in underdoped models, a PG arises in the antinodal direction $k \sim (\pi, 0), (0, \pi)$ for hole doping and in diagonal (nodal) directions $k_x = k_y$ for electron doping. The latter solution gives example of a system for which d-type superconductivity is combined with a finite minimal gap of Fermi excitations. In accor-

dance with ARPES data for the LSCO compound, the calculation of the stripe phase showed that it is characterized by quasi-one-dimensional segments of the FS, the emergence of a PG, and the suppression of spectral density in the diagonal direction and in the direction parallel to stripe domains. Such anisotropy of the FS and PG is incompatible with the *d* symmetry of the SC order. For a spiral spin structure, polarization anisotropy of the FS is detected when different FS segments correspond to different spin polarizations of electrons. This property can be used for testing spiral spin structures. Among unsolved problems, we can mention the study of the structures of valence bonds and current states and the transition from analysis of quasi-static structures to dynamic fluctuations.

ACKNOWLEDGMENTS

The author is grateful sincerely to V.Ya. Krivnov for helpful remarks and for his assistance in the research.

This study was supported financially by the Russian Foundation for Basic Research (project no. 03-03-32141).

REFERENCES

1. A. Damascelli, Z.-X. Shen, and Z. Hussain, Rev. Mod. Phys. **75**, 473 (2003).
2. Z. M. Shen and D. S. Dessau, Phys. Rep. **253**, 1 (1995).
3. Y.-D. Chuang, A. D. Gromko, D. S. Dessau, *et al.*, Phys. Rev. Lett. **83**, 3717 (1999).
4. H. M. Fretwell, A. Kaminski, J. Mesot, *et al.*, Phys. Rev. Lett. **84**, 4449 (2000).
5. S. V. Borisenko, M. S. Golden, S. Legner, *et al.*, Phys. Rev. Lett. **84**, 4453 (2000).
6. A. Ino, C. Kim, T. Mizokawa, *et al.*, J. Phys. Soc. Jpn. **68**, 1496 (1999).
7. T. Yoshida, X. J. Zhou, M. Nakamura, *et al.*, cond-mat/0011172.
8. D. L. Feng, N. P. Armitage, D. H. Lu, *et al.*, Phys. Rev. Lett. **86**, 5550 (2001).
9. Y. D. Chuang, A. D. Gromko, A. Fedorov, *et al.*, Phys. Rev. Lett. **87**, 117002 (2001).
10. A. A. Kordyuk, S. V. Borisenko, T. K. Kim, *et al.*, cond-mat/0110379; Phys. Rev. Lett. **89**, 077003 (2002).
11. A. Kaminski, S. Rosenkranz, H. M. Fretwell, *et al.*, cond-mat/0210531.
12. A. Kaminsky, S. Rosenkranz, H. M. Fretwell, *et al.*, cond-mat/0203133.
13. S. V. Borisenko, A. A. Kordyuk, S. Legner, *et al.*, Phys. Rev. B **64**, 094513 (2001).
14. M. Lindroos, S. Sahrakorpi, and A. Bansil, Phys. Rev. B **65**, 054514 (2002).
15. R. J. Birgeneau and G. Shirane, in *Physical Properties of High Temperature Superconductors*, Ed. by D. M. Ginzberg (World Sci., Singapore, 1989; Mir, Moscow, 1991), Vol. 1.
16. C. Howald, H. Eisaki, N. Kaneko, *et al.*, cond-mat/0201546; Phys. Rev. B **67**, 014533 (2003).
17. J. E. Hoffman, E. A. Hudson, K. M. Lang, *et al.*, Science **295**, 466 (2002).
18. S. Chakravarty, R. B. Laughlin, D. K. Morr, and Ch. Nayak, cond-mat/0005443; Phys. Rev. B **63**, 094503 (2001).
19. J. E. Hirsch, Phys. Rev. Lett. **54**, 1317 (1985).
20. A. A. Ovchinnikov and M. Ya. Ovchinnikova, Zh. Éksp. Teor. Fiz. **110**, 342 (1996) [JETP **83**, 184 (1996)]; Zh. Éksp. Teor. Fiz. **112**, 1409 (1997) [JETP **85**, 767 (1997)].
21. N. M. Plakida, V. S. Oudovenko, R. Horsch, and A. J. Liechtenstein, Phys. Rev. B **55**, 11997 (1997).
22. A. A. Ovchinnikov and M. Ya. Ovchinnikova, Phys. Lett. A **249**, 531 (1998).
23. A. A. Ovchinnikov, M. Ya. Ovchinnikova, and E. A. Plekhanov, Pis'ma Zh. Éksp. Teor. Fiz. **67**, 350 (1998) [JETP Lett. **67**, 369 (1998)]; Zh. Éksp. Teor. Fiz. **114**, 985 (1998) [JETP **87**, 534 (1998)]; Zh. Éksp. Teor. Fiz. **115**, 649 (1999) [JETP **88**, 356 (1999)].
24. R. O. Kuzian, R. Hayn, A. F. Barabanov, and L. A. Maksimov, Phys. Rev. B **58**, 6194 (1998).
25. F. Onufrieva, P. Pfeuty, and M. Kisilev, Phys. Rev. Lett. **82**, 2370 (1999).
26. A. A. Ovchinnikov and M. Ya. Ovchinnikova, Zh. Éksp. Teor. Fiz. **118**, 1434 (2000) [JETP **91**, 1242 (2000)].
27. C. Kusko and R. S. Markiewicz, Phys. Rev. Lett. **84**, 963 (2000).
28. C. Kusko, R. S. Markiewicz, M. Lindroos, and A. Bansil, cond-mat/0201117; Phys. Rev. B **66**, 140513 (2002).
29. M. R. Norman, H. Ding, M. Randeria, *et al.*, Nature **392**, 157 (1998).
30. T. Timusk and B. Statt, Rep. Prog. Phys. **62**, 61 (1999).
31. M. R. Norman and C. Pepin, cond-mat/0302336.
32. H. Balci, V. N. Smolyaninova, P. Fournier, *et al.*, Phys. Rev. B **66**, 174510 (2002).
33. N. P. Armitage, D. H. Lu, D. L. Feng, *et al.*, Phys. Rev. Lett. **86**, 1126 (2001).
34. A. Koitzsch, G. Blumberg, A. Gozar, *et al.*, cond-mat/0304175.
35. A. A. Ovchinnikov and M. Ya. Ovchinnikova, Zh. Éksp. Teor. Fiz. **116**, 1058 (1999) [JETP **89**, 564 (1999)].
36. X. J. Zhou, P. Bogdanov, S. A. Kellar, *et al.*, Science **286**, 268 (1999).
37. A. Ino, C. Kim, M. Nakamura, *et al.*, Phys. Rev. B **62**, 4137 (2000).
38. X. J. Zhou, T. Yoshida, S. A. Kellar, *et al.*, Phys. Rev. Lett. **86**, 5578 (2001).
39. M. Ya. Ovchinnikova, Zh. Éksp. Teor. Fiz. **123**, 1082 (2003).
40. D. S. Dessau, Z. X. Shen, D. M. King, *et al.*, Phys. Rev. Lett. **71**, 2781 (1993).
41. C. M. Varma, Phys. Rev. Lett. **83**, 3538 (1999).
42. G. Ghiringhelli, L. H. Tjeng, A. Tanaka, *et al.*, Phys. Rev. B **66**, 75101 (2002).

Translated by N. Wadhwa

Journal of Experimental and Theoretical Physics, Vol. 98, No. 3, 2004, pp. 556–564.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 125, No. 3, 2004, pp. 630–639.

SOLIDS
Electronic Properties

The Band Structure of *n*-Type Cuprate Superconductors with the *T*'(*T*) Structure Taking into Account Strong Electron Correlation

V. A. Gavrichkov and S. G. Ovchinnikov

Kirenskiĭ Institute of Physics, Siberian Division, Russian Academy of Sciences, Akademgorodok, Krasnoyarsk, 660036 Russia

e-mail: gav@iph.Krasn.ru; sgo@iph.Krasn.ru

Received June 6, 2003

Abstract—The spectral density, dispersion relations, and the position of the Fermi level for *n*-doped compositions based on NCO and LCO were calculated within the framework of the generalized tight binding method. As distinguished from LCO, the dielectric gap in NCO is nonlinear in character. We observe a virtual level both at the bottom of the conduction band and at the top of the valence band in both compounds. However, its position corresponds to the extreme bottom of the conduction band in LCO and is 0.1–0.2 eV above the bottom in NCO. This explains why we observe Fermi level pinning in *n*-LCO as the concentration of the doping component grows and reproduce its absence in NCCO at low doping values. We also found both compositions to be unstable in a narrow concentration range with respect to a nonuniform charge density distribution. The relation between the phase diagram for NCCO and the calculated electronic structure is discussed. © *2004 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

Superconducting *n*-type cuprates $Nd_{2-x}Ce_xCuO_4$ (NCCO) and $Pr_{2-x}Ce_xCuO_4$ (PCCO) have certain special features that distinguish them from *p*-type cuprates. Rare-earth metal and oxygen layers are shifted in their crystal structure with respect to CuO_2 in such a way that there is no nearest neighbor (the apical oxygen) for copper along the *c* axis; that is, the structural element of the CuO_2 layer is a plane CuO_4 cluster, whereas it is a CuO_6 octahedron in $La_{2-x}Sr_xCuO_4$ (LSCO). The phase diagram of NCCO is substantially different from the phase diagram of LSCO (Fig. 1). The initial undoped Nd_2CuO_4 (NCO) composition is an antiferromagnet and dielectric. Doping with electrons rather weakly suppresses the antiferromagnetic state [1] because of the diamagnetic dilution mechanism [2]. The superconducting state borders on the antiferromagnetic phase and exists in a narrow concentration range $x_{min} < x < x_{max}$, where $x_{min} = 0.14$ and $x_{max} = 0.17$. In the normal state, the electrical conduction of NCCO is described by a Fermi-liquid quadratic temperature dependence [3], as distinguished from linear dependences for hole high-T_C superconductors [4]. It was shown for NCCO by angle-resolved photoelectron spectroscopy [5] that the dielectric gap in this compound was not rectilinear. The minimum of the conduction band and the maximum of the valence band belong to different Brillouin zone points, $\mathbf{k} = (\pi, 0)$ and $\mathbf{k} = (\pi/2, \pi/2)$, respectively. The dispersion relations in NCCO for the top of the valence band, however, remain virtually identical to those in LSCO. Angle-resolved photoelectron spectra also revealed the appearance of intragap states when either NCCO was doped with electrons or LSCO was doped with holes [5]. As distinguished from LSCO, where chemical potential pinning occurs at low *x*, NCCO shows a more complex concentration dependence of the chemical potential [6].

This is one more important difference between NCCO and LSCO, which is related to the possible role that neodymium *f* electrons can play in the formation of the state of heavy fermions at low temperatures [7]. However, first, no unambiguous experimental substan-

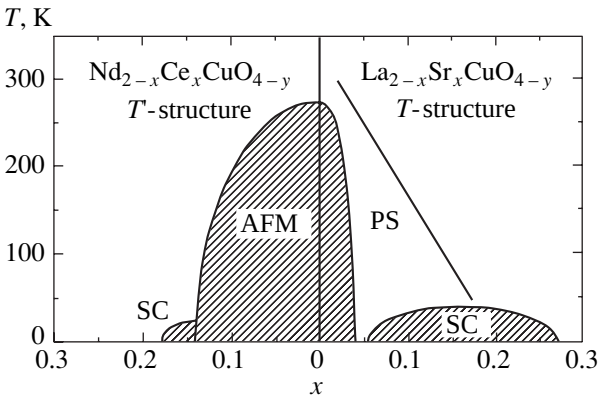

Fig. 1. Phase diagram of LSCO and NCCO. Composition regions: SC, superconducting phase; PS, pseudogap state; and AFM, dielectric phase in the antiferromagnetic state.

tiation [8] of the existence of such a state has been obtained and no explanation of the large coefficient γ value in the linear temperature dependence of heat capacity, $c = \gamma T$, where $\gamma \approx 4J/k^2$ for $x > 0.15$ [9], has been suggested. Secondly, at a low doping level ($0.05 \le x < 0.14$), the γ value is an order of magnitude smaller, $\gamma = 0.3J/k^2$ [9], which leads us to conclude that, even if the state of heavy fermions of the new type does exist in NCCO, these effects manifest themselves at higher concentrations because of the low Nd–Cu spin–spin coupling constant. In this work, we consider low doping levels, and, at $x < 0.15$, the beautiful physics of heavy fermions [7] remains outside the scope of our analysis.

The purpose of this work was to study the electronic structure of NCO and La_2CuO_4 (LCO) undoped and weakly doped with electrons. We use the same calculation methods taking into account strong electron correlation as were earlier used by us to study hole cuprates.

Note that strong electron correlation is of fundamental importance and should be correctly taken into account. Indeed, one-electron band calculations give the ground state of LCO and NCO in the form of a metal with a half-filled band, which is at variance with experiment. At the same time, the simplest strong electron correlation models like the Hubbard model are excessively simplified, and the degree of their applicability to describing the band structure of a particular substance is not known a priori. Our experience in studying the electronic structure of high-T_C cuprates taking into account strong electron correlation shows [10–12] that the multiband *p–d* model [13] is the most suitable at excitation energies up to 3 eV inside the valence and conduction bands. This model takes into account two *d* orbitals of copper, $d_{x^2-y^2} \equiv d_x$ and $d_{3z^2-r^2} \equiv d_z$. We used the generalized tight binding method to calculate the band structure of quasi-particles taking into account strong electron correlation [14]. In this method, the many-electron Hamiltonian within the cell is exactly diagonalized, many-electron molecular orbitals are found, and Hubbard *X*-operators are constructed at the first stage. At the second stage, intercell jumps are included and the band structure of the crystal is calculated. Particular examples of calculations of hole cuprates by this method are given in [11, 12, 15].

In this work, we calculated the spectral density and dispersion for the conduction band in compounds undoped and weakly doped with electrons. The calculations by the generalized tight binding method were performed for NCCO with the *T'* structure and *n*-type La_2CuO_4 with the *T* structure. We show that there is a virtual level typical of systems with strong electron correlation both at the bottom of the conduction band and at the top of the valence band [11] in LCO and NCO. The positions of this level in the two compounds are, however, different. The observed asymmetry also results in different concentration dependences of the Fermi level in NCCO and *n*-LCO. The character of valence band dispersion remains virtually identical in compounds of both types. All results obtained for *n*-LCO have no experimental analogs and are predictive in character, because no *n*-LCO materials with the *T* structure have been prepared as yet. In particular, *n*-type superconducting compositions $La_{2-x}Ce_xCuO_4$ obtained in [16] had the *T'* structure. Nevertheless, there is a possibility of inverting *p*-type LCO by the field effect, as with field-effect transistors, where applying a positive voltage to the gate results in the formation of an inversion layer in a *p*-type semiconductor at the boundary with the gate. A theoretical study of the electronic structure of *n*-LCO with the *T* structure is therefore of interest.

In Section 2, we discuss the most important changes in the generalized tight binding method for *n*-type cuprates and give the initial equations for dispersion relations and spectral density. The dispersion dependences for the conduction band in NCCO in comparison with similar dependences for *n*-LCO and experimental dependences are studied in Section 3. In Section 4, we calculate the partial contributions of various orbitals to the spectral density of the conduction band and study the density of states at the bottom of the conduction band in both compounds. The positions of the Fermi level in NCCO at various doping component concentrations are determined in Section 5, where the instability of the state with a uniform charge density distribution is also studied. The results of our calculations are briefly summarized in Section 6.

2. DISPERSION RELATIONS AND SPECTRAL DENSITY OF QUASI-PARTICLE STATES IN NCCO AND *n*-LCO

In the generalized tight binding method, the Hamiltonian of the CuO_2 layer can be written in the form

$$H = \sum_{i\lambda\sigma} \varepsilon_i^{\lambda} a_{i\lambda\sigma}^{\dagger} a_{i\lambda\sigma} + \frac{1}{2} \sum_{i,j} \sum_{\lambda_1\lambda_2\sigma_1\sigma_2\sigma_3\sigma_4} V_{ij}^{\lambda_1\lambda_2} a_{i\lambda_1\sigma_1}^{\dagger} a_{i\lambda_1\sigma_3} \times a_{i\lambda_2\sigma_2}^{\dagger} a_{i\lambda_2\sigma_4} + \sum_{\langle i,j\rangle} \sum_{\lambda_2\lambda_2\sigma} t_{ij}^{\lambda_1\lambda_2} a_{i\lambda_1\sigma}^{\dagger} a_{j\lambda_2\sigma}. \quad (1)$$

Here, $a_{i\lambda\sigma}$ is the hole annihilation operator in the Wannier state on node *i* (copper or oxygen) for orbital λ and spin σ. Two copper orbitals ($d_{x^2-y^2}$ and d_{z^2}) and two $p_{x/y}$ and p_z orbitals on each oxygen node that form σ bonds with the specified copper orbitals are included. Among the Coulomb matrix elements, we can identify intraatomic Hubbard elements $U_d(U_p)$ for repulsion in one copper (oxygen) orbital between electrons with opposite spins, interorbital Coulomb elements $V_d(V_p)$,

$J_d(J_p)$ exchange matrix elements, and interatomic Coulomb repulsion parameters V_{pd}, which we, for simplicity, consider identical for all orbitals. The last term in (1) describes interatomic copper–oxygen jumps with the parameters $t_{pd}^{x^2-y^2,\,x/y} \equiv t_{pd}$ and $t_{pd}^{z^2,\,x} \equiv t_{pd}/\sqrt{3}$ and oxygen–oxygen jumps with the parameter $t_{pp}^{x,\,y} \equiv t_{pp}$. The charge transfer energy will be denoted by $\Delta_{pd} = \varepsilon_p - \varepsilon_{d_{x^2-y^2}}$, and the energy of splitting of the d level in the uniaxial crystal field component, by $\Delta_d = \varepsilon_{dz^2} - \varepsilon_{d_{x^2-y^2}}$. Of six O^{2-} ions, two apical ones are situated along the c axis in the T structure of the LCO composition. Their effects are controlled by two calculation parameters, t'_{pd} and t'_{pp}, which are the integrals of electron jumps from copper and in-pane oxygen to apical oxygen.

In the generalized tight binding method, the band structure of quasi-particles including strong electron correlation effects is calculated in two stages. At the first stage, the CuO_2 layer lattice is partitioned into many unit cells, and the Hamiltonian within one cell is exactly diagonalized. In addition to selecting the CuO_6 cluster as the unit cell, the problem of constructing the Wannier functions of b_{1g} and a_{1g} symmetry on the initial oxygen orbitals is solved [11, 12]. The many-electron molecular orbitals $|n, p\rangle$ (where $n = 0, 1, 2, \ldots$ is the number of holes in the cell and p denotes the set of the other orbital and spin indices) obtained by diagonalizing the cell Hamiltonian H_0 are used to construct the Hubbard operators of this cell, $X^m = |n+1, p\rangle\langle n, q|$, and one-electron operators, $a_{f\lambda\sigma} = \sum_m \gamma_{\lambda\sigma}(m) X^m_{f\sigma}$. Here, the band index of quasi-particles m numbers one-electron excitations from the initial state $|n, q\rangle$ to the final state $|n+1, p\rangle$ [17].

As distinguished from LCO, oxygen and rare-earth metal elements in NCO are known to form their own separate planes in the environment of the CuO_2 plane, and the plane of the rare-earth metal is closest to the CuO_2 plane. In this situation, both parameters should be close to zero, $t'_{pd} = 0$ and $t'_{pp} = 0$. Additional changes in the other parameters were not introduced beforehand and were taken from the calculations of the electronic structure of p-type cuprates [11]. The initial parameters of our Hamiltonian were:

$$\varepsilon_{d_x} = 0, \quad \varepsilon_{d_z} = 2 \text{ eV},$$
$$\varepsilon_p = 1.6 \text{ eV}, \quad \varepsilon_{p_z} = 0.5 \text{ eV},$$
$$t_{pp} = 0.46 \text{ eV}, \quad t'_{pp} = 0, \quad t'_{pd} = 0, \quad U_d = 9 \text{ eV},$$
$$U_p = 4 \text{ eV}, \quad V_{pd} = 1.5 \text{ eV}, \quad J_d = 1 \text{ eV}.$$

In the generalized tight binding method, the dispersion relations and spectral density can be written in the form [11]

$$\left\| \frac{(E - \Omega^G_m)\delta_{mn}}{F^G_\sigma(m)} - 2\sum_{\lambda\lambda'} \gamma^*_{\lambda\sigma}(m) T^{PG}_{\lambda\lambda'}(\mathbf{k}) \gamma_{\lambda'\sigma}(n) \right\| = 0, \quad (2)$$

$$A_\sigma(\mathbf{k}, E) = \left(-\frac{1}{\pi}\right)\sum_\lambda \mathrm{Im}(G^{\lambda\lambda}_{k\sigma}) = \left(-\frac{1}{\pi}\right) \times \sum_{\lambda mn} \gamma_{\lambda\sigma}(m)\gamma^+_{\lambda\sigma}(n)\mathrm{Im}(D^{mn}_{k\sigma}(AA) + D^{mn}_{k\sigma}(BB)), \quad (3)$$

where

$$G^{\lambda\lambda'}_{k\sigma} = \langle\langle a_{k\lambda\sigma} | a^\dagger_{k\lambda'\sigma}\rangle\rangle_E = \sum_{mn} \gamma_{\lambda\sigma}(m)\gamma^\dagger_{\lambda'\sigma}(n) D^{mn}_{k\sigma}, \quad (4)$$

$$\hat{D}_{k\sigma} = \begin{pmatrix} \hat{D}_{k\sigma}(AA) & \hat{D}_{k\sigma}(AB) \\ \hat{D}_{k\sigma}(BA) & \hat{D}_{k\sigma}(BB) \end{pmatrix}, \quad D^{mn}_{k\sigma}(AB) = \langle\langle X^m_{k\sigma} | Y^n_{k\sigma}\rangle\rangle_E. \quad (5)$$

Here, indices P and G run over the A and B antiferromagnetic state sublattices. Equations (2) and (3) were obtained for the antiferromagnetic phase [11, 12] using the equations of motion for Green function (5) in the Hubbard I approximation for intercell jumps. The elements of the tight binding matrix

$$T^{AA}_{\lambda\lambda'}(k) = T^{BB}_{\lambda\lambda'}(k) = \frac{2}{N}\sum_{R_1} T^{AA}_{\lambda\lambda'}(\mathbf{R}_1) e^{ikR_1},$$

$$T^{AB}_{\lambda\lambda'}(k) = T^{BA}_{\lambda\lambda'}(k) = \frac{2}{N}\sum_{R_2} T^{AB}_{\lambda\lambda'}(\mathbf{R}_2) e^{ikR_2}$$

in the five-orbital d_x, d_z, b, a, p_z basis take the form

$$T_{\lambda\lambda'}(\mathbf{R}) = \begin{pmatrix} 0 & 0 & -2t_{pd}\mu_{ij} & 0 & 0 \\ 0 & 0 & \frac{2t_{pd}\xi_{ij}}{\sqrt{3}} & \frac{2t_{pd}\lambda_{ij}}{\sqrt{3}} & 0 \\ -2t_{pd}\mu_{ij} & \frac{2t_{pd}\xi_{ij}}{\sqrt{3}} & -2t_{pp}\nu_{ij} & 2t_{pp}\chi_{ij} & -2t'_{pp}\xi_{ij} \\ 0 & \frac{2t_{pd}\lambda_{ij}}{\sqrt{3}} & 2t_{pp}\chi_{ij} & 2t_{pp}\nu_{ij} & -2t'_{pp}\lambda_{ij} \\ 0 & 0 & -2t'_{pp}\xi_{ij} & -2t'_{pp}\lambda_{ij} & 0 \end{pmatrix}; \quad (6)$$

the μ_{ij}, ξ_{ij}, λ_{ij}, ν_{ij}, and χ_{ij} coefficients were given

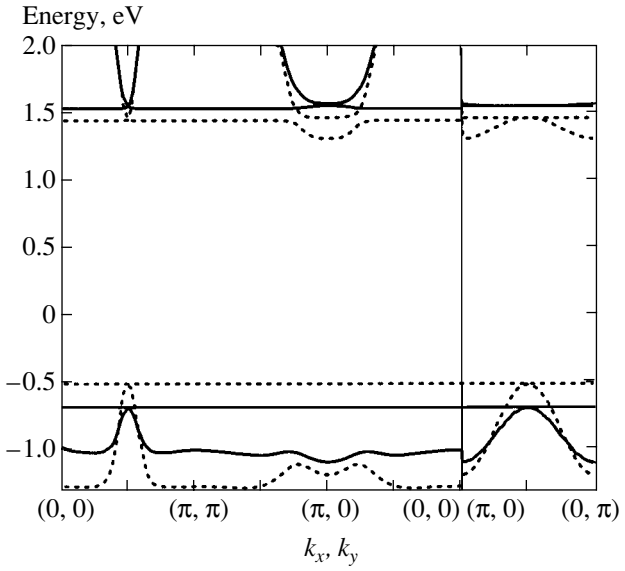

Fig. 2. Dispersion dependences for *n*-LCO (solid line) and NCCO (dotted line). Doping level $x = 0.03$.

in [11]. Equation (2) is an analog of the dispersion equation in the tight binding method and differs from it in two respects. First, quasi-particle energies are calculated in the form $\Omega_m^G = \varepsilon_{2qG} - \varepsilon_{1pG}$, that is, in the form of resonances between many-particle states from different configuration space sectors. Secondly, the occupation number $F_\sigma^G(m) = \langle X_{f_G\sigma}^{pp} \rangle + \langle X_{f_G\sigma}^{qq} \rangle$ leads to concentration dependences of both dispersion and spectral density amplitude (3). Quasi-particle states with different m can overlap and interact, like singlet and triplet two-hole states of *p*-type cuprates do [11, 12].

3. DISPERSION OF BANDS

The band dispersion for *n*-LCO and NCCO at a concentration of the doping *n*-component of $x = 0.03$ in the antiferromagnetic phase is shown in Fig. 2. Calculations only reproduce dispersion in the immediate vicinity of the dielectric gap. This is sufficient for analyzing the spectrum of quasi-particles involved in the superconducting state.

The bottom of the conduction band is formed as a result of the dispersion of the local state with an energy of $\Omega_c = E(1, {}^2b_{1g}) - E(0, a_{1g})$, and the top of the valence band is formed by excitations with the participation of the two-hole singlet $\Omega_v = E(2, {}^1A_{1g}) - E(1, {}^2b_{1g})$ and triplet $\Omega_v^1 = E(2, {}^3B_{1g}) - E(1, {}^2b_{1g})$ (Fig. 3). Both compounds have a virtual level at the bottom of the conduction band. This level is similar in nature to that at the top of the valence band (Fig. 3) [11]. Namely, there are two types of quasi-particles that correspond to possible transitions Ω_c and Ω_v. One of the quasi-particles in the undoped compound corresponds to the transition between empty states, which gives zero contributions to dispersion and spectral density. In the one-hole sector of the configuration space, the empty state is one of the components of the spin doublet in each of the sublattices of the Néel antiferromagnetic state of the CuO_2 layer. The vacuum sector corresponds to the a_{1g} singlet state of the fully occupied p^6d^{10} shell. The existence of two singlet states in the vacuum and two-hole sectors (Fig. 3) is the main reason for the existence of the dispersionless virtual level not only at the top of the valence band but also at the bottom of the conduction band in *n*-LCO and NCCO.

Conduction band dispersion in *n*-LCO was calculated with the parameter values obtained in studying the

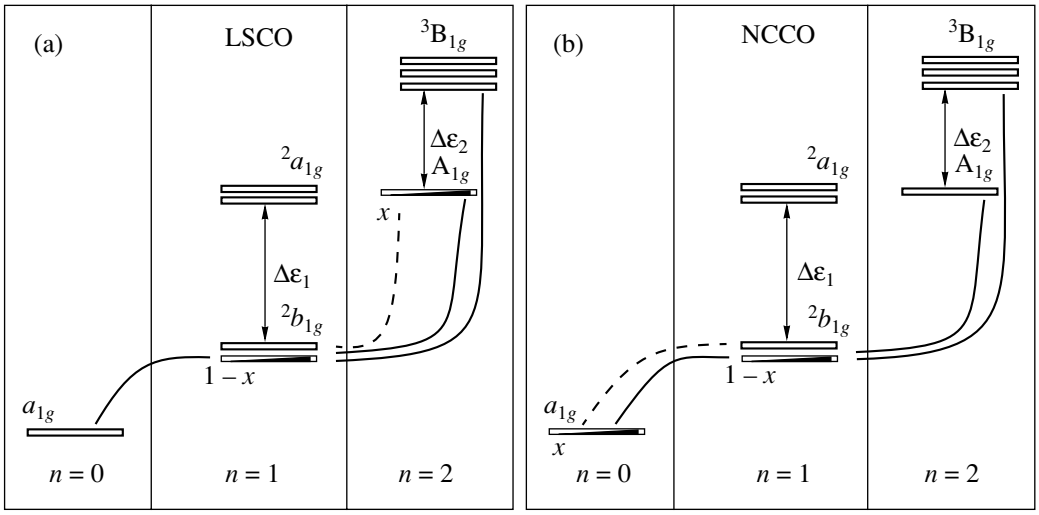

Fig. 3. Configuration space scheme for LSCO and NCCO. The solid lines correspond to quasi-particle transitions that form rigid bands, and the dashed lines, to impurity bands.

electronic structure of *p*-type cuprates. Mere comparison of the spectra for *n*-LCO and NCCO in Fig. 2 shows that the degeneracy of the spectrum of LCO at point $X = (\pi, 0)$ is accidental. Conversely, the intersection of two conduction bands, the broad band and the band of virtual level states, at point $M = (\pi/2, \pi/2)$ is caused by CuO_2 layer symmetry and does not depend on the parameter values used in the calculations.

The broad band and the band of virtual level states behave differently as the doping level increases [11, 15]. The broad bands remain virtually unchanged; they will further be called rigid by analogy with the rigid band model. The spectral density and dispersion for the virtual level state bands increase as the degree of doping grows, and they will be called "impurity" bands. Quotation marks (further omitted) are not meaningless, because these states have no bearing on the true local impurity potential.

The transport of quasi-particles in the valence and conduction bands occurs with different effective transport integrals. We therefore observe different dispersion dependences for different bands. Indeed, calculations give $t'_{\text{c.b.}}/t_{\text{c.b.}} = -0.05$ and $t'_{\text{v.b.}}/t_{\text{v.b.}} = -0.14$ for NCO and $t'_{\text{c.b.}}/t_{\text{c.b.}} = 0.05$ and $t'_{\text{v.b.}}/t_{\text{v.b.}} = -0.085$ for LCO, where $t_{\text{c.b.}}$ ($t'_{\text{c.b.}}$) and $t_{\text{v.b.}}$ ($t'_{\text{v.b.}}$) are the effective transport integrals between the nearest (next-nearest) neighbors for the conduction and valence bands, respectively. The most significant change in passing from LCO to NCO is the formation of a nonlinear dielectric gap because of the formation of a new minimum at the *X* point of the conduction band. Calculations show that the appearance of a rectilinear gap is also accompanied by a change in the sign of the $t'_{\text{c.b.}}/t_{\text{c.b.}}$ ratio. There are moderate changes in the valence band, but they do not lead to qualitative differences in the dispersion dependences for the *n* and *p* materials.

The reproduction of dispersion at the bottom of the conduction band in NCCO requires the initial parameters to be changed as follows:

$$\varepsilon_{d_x} = 0.2 \text{ eV}, \quad \varepsilon_{d_z} = 2 \text{ eV},$$
$$\varepsilon_p = 1.6 \text{ eV}, \quad \varepsilon_{p_z} = 0.5 \text{ eV},$$
$$t_{pp} = 0.56 \text{ eV}, \quad t'_{pp} = 0.1 \text{ eV},$$
$$t'_{pd} = 0, \quad U_d = 9 \text{ eV},$$
$$U_p = 4 \text{ eV}, \quad V_{pd} = 1.5 \text{ eV}, \quad J_d = 1 \text{ eV}.$$

It follows that dispersion calculations in NCCO largely result in changes in the t'_{pd} and t'_{pp} values and, to a lesser extent, in Δ_{pd} and t_{pp}. A smaller Δ_{pd} value corresponds to a smaller dielectric gap in NCO, $E_g = 1.6$ eV. A smaller t_{pp} value in LCO in turn corresponds to the presence of orthorhombic distortions in the system of CuO_6 octahedra. As a consequence, a small increase in t_{pp} in NCCO is responsible for the conduction band minimum at point *X* at low electron concentrations. In combination, changes in precisely these parameters qualitatively modify the dielectric gap and make it nonlinear in NCO.

Among the known materials based on *n*-LCO and having the *T* structure, the calculated dispersion might be observed in $La_2Cu_{1-x}Zn_xO_4$ compositions [1], because Zn^{2+} has a completely filled d^{10} shell, which formally corresponds to filling the vacuum sector, as in *n*-type materials such as NCCO. However in reality, the Zn impurity leads to a strongly bound carrier and the violation of translational invariance over the spin lattice. It appears that, because of strong impurity effects, photoemission measurements for $La_2Cu_{1-x}Zn_xO_4$, similar to those performed for NCCO [5], cannot be made.

4. PARTIAL CONTRIBUTIONS TO SPECTRAL DENSITY

We calculated the spectral density $A(\mathbf{k}, E)$ for the rigid and impurity bands in NCCO (Fig. 4) at a low concentration of the doping component $x = 0.03$. The spectral density is characterized by two peaks corresponding to the rigid and impurity bands. The dependences of the peak amplitudes for (a) the rigid and (b) impurity bands along the symmetrical Brillouin zone directions are plotted in Fig. 4. Figure 5b shows how the virtual level with zero spectral weight at $x = 0$ transforms into an impurity band with spectral weight *x*. The overshooting of triplet states into the conduction band is insignificant, and this is one more source of asymmetry of the states of *p*- and *n*-type carriers. Similar dependences for the conduction band in *n*-LCO are shown in Fig. 5. These results cannot however be compared with experimental data because of the absence of *n*-type compounds based on LCO with the *T* structure.

As follows from calculations of the density of states, there is a region with a reduced density between the impurity and rigid bands (Fig. 6). This pseudogap vanishes at point $\overline{M} = (\pi/2, \pi/2)$ (Fig. 2). For this reason, the passage of the Fermi level from the rigid to the impurity band may be accompanied by a decrease in the density of states on this level. The pseudogap itself is magnetic in nature, as follows from its absence in the paramagnetic phase. Because of the special features inherent in the spectrum, the pseudogap is more pronounced for the density of states of NCCO.

5. THE CONCENTRATION DEPENDENCE OF THE FERMI LEVEL

Calculations of the dependence of the Fermi level position on doping in NCCO for the antiferromagnetic phase exhibit considerable differences from the dependence obtained for *n*-LCO. Indeed, at low concentrations, the Fermi level in NCCO goes deep into rigid

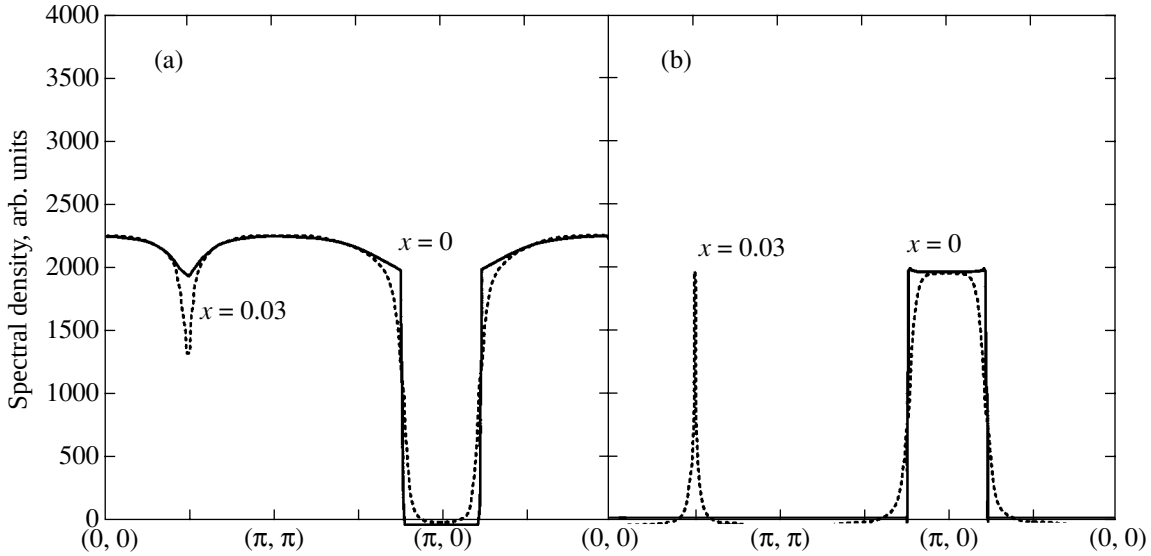

Fig. 4. Spectral density of quasi-particle states along symmetrical Brillouin zone directions in NCO (a) for the rigid conduction band and (b) for the band of virtual level states.

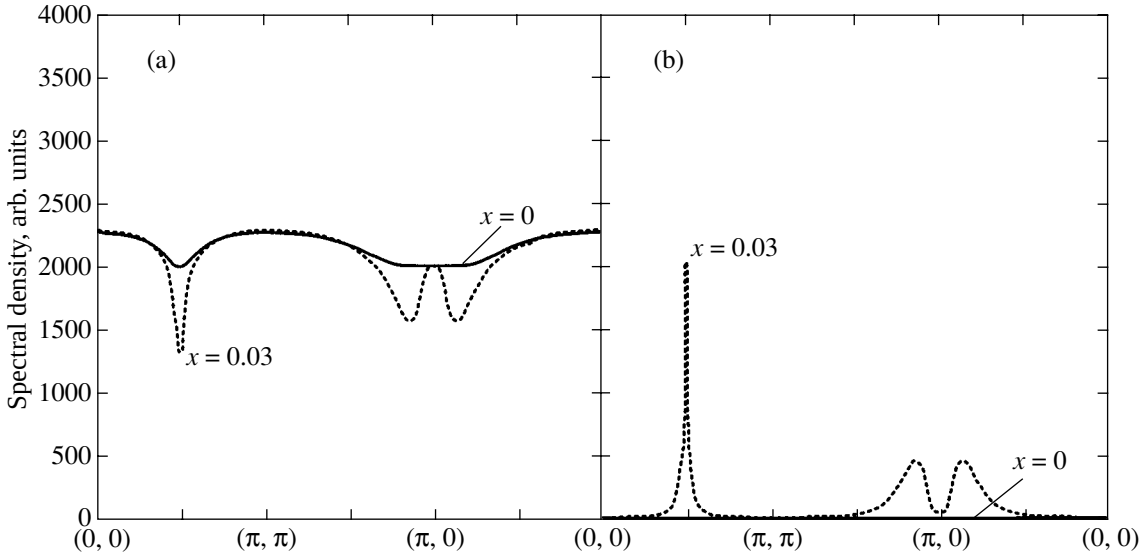

Fig. 5. Spectral density of quasi-particle states along symmetrical Brillouin zone directions in *n*-LCO (a) for the rigid conduction band and (b) for the impurity band.

conduction band states and only then, at $x = x_1$, into impurity band states. This corresponds to $x_1 = 0.08$–0.1 in Fig. 7. The Fermi level reenters rigid band states at $x_2 = 0.18$–0.2. A similar dependence of the Fermi level in *n*-LCO shows pinning only at low concentrations. Indeed, already at very low concentrations, the Fermi level occurs in the zone of impurity band states that are being formed. Because the number of states on the Fermi level begins to grow more slowly than x (Fig. 8b), pinning ends, and the Fermi level enters the rigid conduction band. This occurs at $x_2 = 0.11$–0.12.

The calculated concentration dependence of the Fermi level contains a Δx interval where the rate of growth of the number of states in the impurity band exceeds the rate of increasing the number of carriers $\partial\mu/\partial x = (\partial^2\Phi/\partial x^2)_{T,P} < 0$, which is evidence of possible phase stratification at the given doping level. Such compositions cannot be stably homogeneous, because the sought distribution corresponds to the thermodynamic potential Φ maximum. For instance, the dependence of $\partial\mu/\partial x$ on x for NCCO (Fig. 9) shows that $\Phi(x)$ has an instability region $\Delta x \approx 0.03$ wide. This region separates

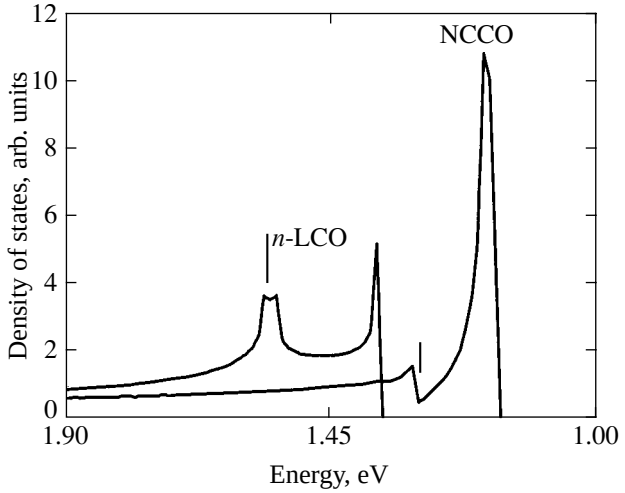

Fig. 6. Density of states at the bottom of the conduction band in NCCO and *n*-LCO. The vertical line indicates the region with a reduced density of states (pseudogap).

two regions with thermodynamic potential minima. Even at small dopant concentrations, *n*-LCO occurs in the $\Delta x \approx 0.05$ region with a maximum thermodynamic potential. Although the materials under consideration are systems with fixed numbers of carriers, both systems can either be divided into macroscopic regions capable of exchanging particles or experience the transition to the superconducting state, where the number of particles is no longer conserved. The origin of the instability of the homogeneous normal state is related to the presence of antiferromagnetic order, because the impurity band disappears in the paramagnetic phase. It follows that we observe an unusual relation between the nonuniform charge density distribution and the presence of antiferromagnetic ordering. The inclusion of zero spin fluctuations causes instability region Δx narrowing, but does not negate its existence [15].

We also observe a qualitative correlation between Fermi level movement to the antiferromagnetic state and the phase diagram of NCCO. Indeed, the concentration region of Fermi level residence in the impurity band, or, which is the same, in the pseudogap region correlates with the superconducting region in the phase diagram. In NCCO and *n*-LCO, the Fermi level reaches the pseudogap at different dopant concentrations $x_1(\text{NCCO}) > x_1(n\text{-LO}) = 0$. In NCCO, the Fermi level enters impurity band states with an already well-developed spectral density. This is seen from Fig. 8a, where the spectral density in the impurity band is nonzero already at $x = x_1$. It follows that the presence of a finite spectral density of impurity band states at the Fermi level corresponds to the superconducting region in the phase diagram of NCCO. The immediate consequence of a correlation of this type would be the beginning of the superconducting region in the phase diagram of NCCO at T_C higher than T_C for *n*-LCO, this region being narrower on the concentration scale. Accordingly, we also have $x_{max} \approx x_2$. Such an equality was observed for PCCO in [18], whose authors were able to

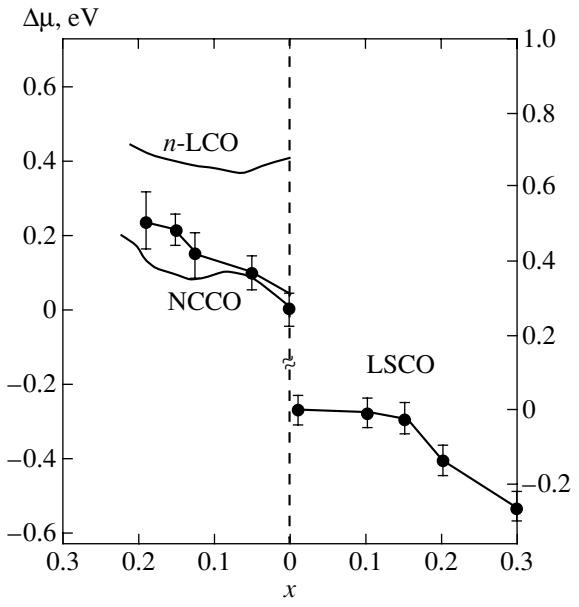

Fig. 7. Dependence of the chemical potential shift $\Delta\mu(x)$ on the concentration of the doping *n*-component in NCCO and *n*-LCO. The experimental data on NCCO and LSCO were taken from [6].

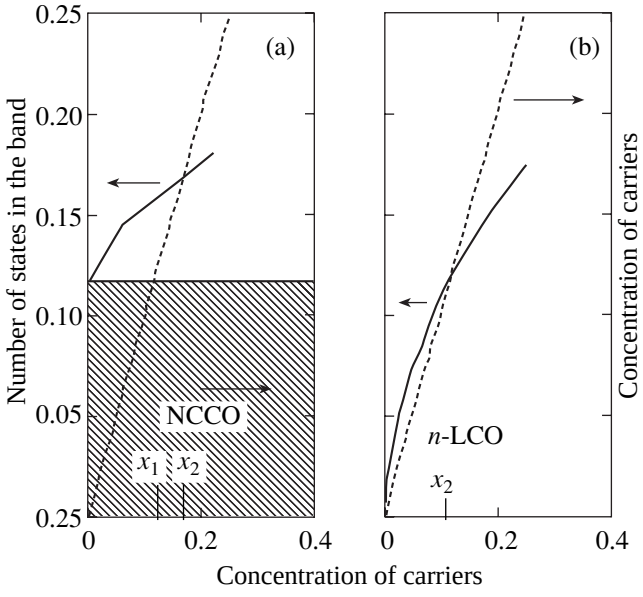

Fig. 8. Dependence of the total number of states on the concentration of the doping component for the impurity bands in NCCO and *n*-LCO; x_1 and x_2 correspond to the entrance to and exit from the band of virtual level states, respectively. The solid line shows the number of states, and the dashed line, the concentration of electrons *x*. The shaded region in (a) corresponds to the contribution of the rigid band.

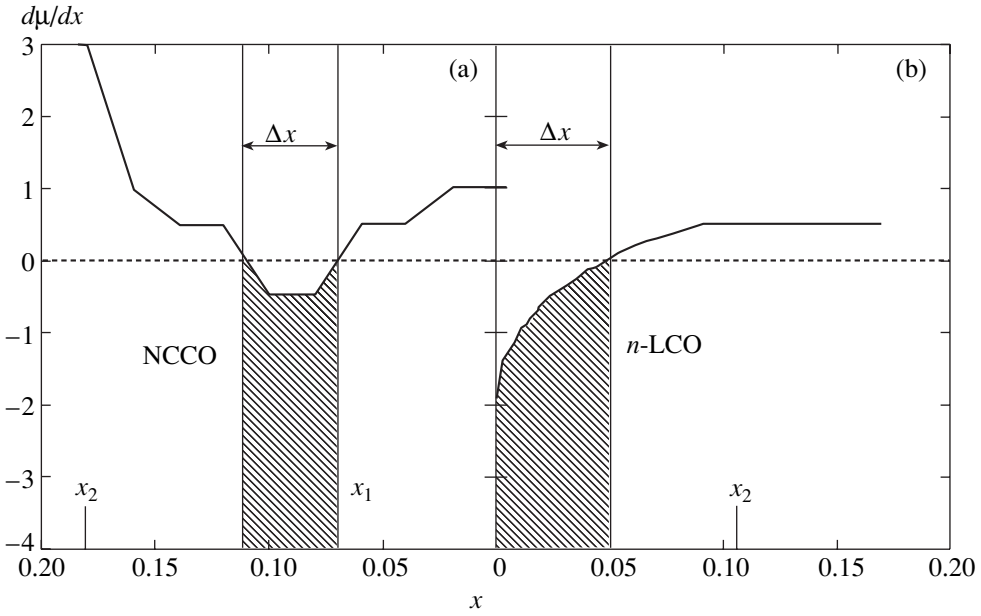

Fig. 9. Dependence of the second thermodynamic potential derivative $\partial^2\Phi/\partial x^2$ on the concentration of the doping component in NCCO and *n*-LCO; x_2 corresponds to the exit of the Fermi level from the impurity band.

study the $T^*(x)$ dependence of the characteristic temperature for the pseudogap state in the superconducting phase of PCCO in magnetic fields higher than H_{c2}. In our calculations, the pseudogap state is an attribute of the impurity band rather than the precursor of the superconducting state. It can be identified as a special feature of the electronic structure of materials with strong electron correlation in the antiferromagnetic phase and with a singlet ground state in one of the configuration space sectors of the unit cell of the material under study.

6. CONCLUSIONS

The results of our calculations can be summarized as follows:

(1) Common to the dispersion dependences for NCCO and *n*-LCO is the presence of a virtual level at the bottom of the conduction band and at the top of the valence band in the antiferromagnetic phase. The reason for its existence is the presence of singlet states in the vacuum (a_{1g} is a closed shell) and two-particle (A_{1g}) configuration space sectors of both compounds. The rigid conduction band in NCCO has a minimum at point *X* of the Brillouin zone at low doping levels. Because of accidental degeneracy of the rigid band and virtual level at point *X* in *n*-LCO, its dielectric gap is rectilinear, whereas the gap in NCCO is not. The last conclusion is in agreement with the angle-resolved photoelectron spectroscopy data on weakly doped NCCO compounds [5].

(2) The concentration dependences of the Fermi level for NCCO with the *T*' structure and *n*-type LCO with the *T* structure are not symmetrical. We explain this asymmetry by the different positions of the virtual level with respect to the bottom of the rigid conduction band in these compositions. As a consequence, we observe pinning of the Fermi level by states developed on the virtual level in *n*-type LCO at low dopant concentrations. In NCCO, the Fermi level is immediately immersed into rigid conduction band states and only enters impurity band states when the degree of doping increases further. The probability of pinning the Fermi level by them, however, actually decreases as the doping level grows.

(3) We observed that, in our calculations, the regions of Fermi level pinning by the impurity band were virtually Δx regions with an instability of the form $(\partial^2\Phi/\partial x^2)_{T,P} < 0$, which could be the reason for a nonuniform charge density distribution in the compositions under consideration.

(4) A qualitative correspondence exists between the phase diagram of NCCO and the concentration dependence of the Fermi level: namely, the concentration region of Fermi level residence in the impurity band correlates with the concentration region of the superconducting state in these compounds. There is no such correspondence for *n*-LCO because of the absence of the corresponding materials with the *T* structure. Our results, however, show that the hypothetical phase diagram for LCO of the *p*/*n* type with the *T* structure should be more symmetrical than the diagram of NCCO/LSCO.

ACKNOWLEDGMENTS

This work was financially supported by the Russian Foundation for Basic Research (project no. 03-02-16124), RFFI-KKFN "Eniseĭ" (project no. 02-02-97705),

INTAS (project no. 01-0654), integration program of URO and Siberian Division, Russian Academy of Sciences (project no. 9), and the "Quantum Macrophysics" program of the Russian Academy of Sciences.

REFERENCES

1. B. Keimer, N. Belk, R. J. Birgeneau, *et al.*, Phys. Rev. B **46**, 14034 (1992).
2. S. G. Ovchinnikov, Physica C (Amsterdam) **228**, 81 (1994).
3. S. J. Hagen, J. I. Peng, Z. Y. Li, and R. I. Greene, Phys. Rev. B **43**, 13606 (1991).
4. H. Takagi, B. Batlog, H. L. Kao, *et al.*, Phys. Rev. Lett. **69**, 2975 (1992).
5. N. P. Armitage, F. Ronning, D. H. Lu, *et al.*, Phys. Rev. Lett. **88**, 257001 (2002).
6. N. Harima, J. Matsuno, A. Fujimori, *et al.*, Phys. Rev. B **64**, 220507(R) (2001).
7. P. Fulde, Physica B (Amsterdam) **230**–**232**, 1 (1997).
8. W. Henggeler, B. Roessli, A. Furrer, *et al.*, Phys. Rev. Lett. **80**, 1300 (1998).
9. E. Maiser, W. Mexner, R. Schafer, *et al.*, Phys. Rev. B **56**, 12961 (1997).
10. S. G. Ovchinnikov, Usp. Fiz. Nauk **167**, 1043 (1997) [Phys. Usp. **40**, 993 (1997)].
11. V. A. Gavrichkov, S. G. Ovchinnikov, A. A. Borisov, and E. G. Goryachev, Zh. Éksp. Teor. Fiz. **118**, 422 (2000) [JETP **91**, 369 (2000)].
12. V. A. Gavrichkov, A. A. Borisov, and S. G. Ovchinnikov, Phys. Rev. B **64**, 235124 (2001).
13. Ya. B. Gaididei and V. M. Loktev, Phys. Status Solidi B **147**, 307 (1988).
14. S. G. Ovchinnikov and I. S. Sandalov, Physica C (Amsterdam) **161**, 607 (1989).
15. A. A. Borisov, V. A. Gavrichkov, and S. G. Ovchinnikov, Zh. Éksp. Teor. Fiz. **124**, 862 (2003) [JETP **97**, 773 (2003)].
16. M. Naito and M. Hepp, Jpn. J. Appl. Phys. **39**, L485 (2000).
17. R. O. Zaĭtsev, Zh. Éksp. Teor. Fiz. **68**, 207 (1975) [JETP **41**, 100 (1975)].
18. L. Alff, Y. Krockenberger, B. Welter, *et al.*, Nature **422**, 698 (2003).

Translated by V. Sipachev

Journal of Experimental and Theoretical Physics, Vol. 98, No. 3, 2004, pp. 565–571.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 125, No. 3, 2004, pp. 640–647.

SOLIDS
Electronic Properties

On the Nature of Different Temperature Dependences of the Size of Antiphase Domains in Commensurate Long-Period Structures

O. I. Velikokhatnyĭ[a], S. V. Eremeev[a,b], I. I. Naumov[a], and A. I. Potekaev[b,*]

[a] *Institute of Strength Physics and Materials Science, Siberian Division, Russian Academy of Sciences, Akademicheskiĭ pr. 2/1, Tomsk, 634021 Russia*

[b] *Kuznetsov Physicotechnical Institute, pl. Novosobornaya 1, Tomsk, 634021 Russia*

**e-mail: kanc@spti.tsu.ru, sergey@eremeev.tomsk.ru*

Received August 5, 2003

Abstract—First-principle calculations of the electronic structure and electronic susceptibility were performed to study the relation between the nesting properties of the Fermi surface and the character of the temperature dependence of long-period structures of two types exemplified by Ag_3Mg and Al_3Ti alloys. The observed temperature dependence of the long period length $2M$ in the Al_3Ti alloy was analyzed. It was shown that the temperature dependence of the size of the antiphase domain in long-period commensurate structures was determined by the special features of the electronic structure of the system, in particular, by the geometry of Fermi surface nesting regions. © *2004 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

Ordered alloys with long-period structures are one of the interesting and promising classes of metallic alloys. They differ from the usual ordered systems with simple superstructures by the presence of antiphase boundaries, which periodically or quasi-periodically disturb the ordered arrangement of atoms. Antiphase boundaries, which are usually energetically unfavorable in ordered alloys, are equilibrium structure elements in systems with long-period structures. Ordered alloys with long-period structures have quite definite stability regions in temperature–composition phase diagrams [1].

Studies of the mechanical properties of such alloys showed [2–4] that strengthening against the decay of the supersaturated solid solution could effectively be combined with strengthening against atomic ordering. This allows unusual disperse stable decay structures to be formed. Their alloys possess high mechanical properties, which are stable over the whole temperature range in which the ordered state of the matrix is retained [5, 6].

According to their character, long-period structures can be divided into two groups, namely, incommensurate and commensurate structures. Incommensurate structures are, for instance, formed in the CuAu, Cu_3Au, Au_3Cu, Cu_3Pd, and Cu_3Pt alloys [7]. They are characterized by random distances M between antiphase boundaries (antiphase domains of different lengths are stochastically distributed along the ⟨001⟩ direction). When the composition and temperature are varied, the half-period $\overline{M}$ averaged over the chaotic ensemble changes continuously and takes on various values, including irrational.

In our preceding works [8, 9] we studied the special features of the electronic structure of the Cu–Au, Cu–Pd, and Cu–Pt alloys and answered the question of why incommensurate long-period structures characterized by irrational periods and smeared antiphase boundaries are formed in them. The reasons for the existence of such long-period structures only in a narrow temperature range were elucidated. We also explained the observed dependence of the period length on the degree of long-range ordering η in the Cu–Au alloy [8] and the reasons for the formation of two-dimensional long-period structures in the Au_3Cu and Cu_3Pd alloys [9].

Consider the special features of alloys with commensurate long-period structures, such as Ag_3Mg, Cu_3Al, Au_3Cd, Al_3Ti, Pt_3V, etc.

It was found in studying Ag_3Mg alloys with various magnesium contents that the antiphase half-period was constant, $M = 2$, in a certain concentration range that corresponded to the $D0_{23}$ superstructure [10, 11] or $\overline{M}$ slightly decreased as the concentration of magnesium increased [12]. Importantly, the antiphase boundaries were sharply defined, and $\overline{M}$ remained virtually constant as temperature varied. This type of alloys also includes Au_3Zn, Au_3Cd, and Au_3Mn [13].

Somewhat different results were obtained for Al_3Ti alloys [14–17], for which high- and low-temperature phases with long-period structures were observed [14]. Several comparatively simple commensurate structures

were formed at low temperatures, whereas a series of fairly complex configurations were identified at high temperatures. For instance, the ⟨211⟩ configuration (this notation corresponds to the 2–1–1 sequence of domains, where 2 and 1 stand for domains with $M = 2$ and $M = 1$, respectively) was observed at $T = 700°C$, which corresponded to $\overline{M} \approx 1.33$. At $T = 900°C$, the configuration was ⟨21⟩, and, accordingly, $\overline{M} \approx 1.5$; at $T = 1150°C$, it was ⟨221⟩ with $\overline{M} \approx 1.68$, and, at $T = 1200°C$, the ⟨22221⟩ configuration with $\overline{M} \approx 1.76$ formed. Note that $\overline{M}$ obviously tends to increase as temperature grows, whereas the low-temperature phase is formed as a mixture of antiphase domains with $M = 2$ and $M = 1$ with obvious predominance of $D0_{22}$ superstructure elements ($M = 1$). The simplest long-period structure $D0_{22}$ was observed in the alloy of the stoichiometric composition [15]. At a 60–73 at. % Al, structures with $4/3 \leq \overline{M} \leq 2$ were observed and the $\overline{M}$ value was temperature-dependent. The dependence of $\overline{M}$ on the composition and temperature of annealing was also reported in [16, 17]. In [18], Pt–V alloys with long-period structures at compositions close to Pt_3V were studied. During annealing at 930°C, the $D0_{22}$ ordered structure (of the Al_3Ti type) transformed into a structure with $L1_2$-type ordering, and, at 1036°C, a disordered state was formed. It is directly stated in [19] that the high-temperature (above 1000°C) phase has the $L1_2$ structure, and the low-temperature one (below 900°C), the $D0_{22}$ structure. An increase in $\overline{M}$ with temperature was observed.

To summarize, the most characteristic features of commensurate structures are as follows:

(1) the low-temperature phase is formed as a commensurate superstructure;

(2) the high-temperature state is a mixture of commensurate elements;

(3) the antiphase boundaries are sharply defined planes near which there is no essential structural peculiarities;

(4) the "mean" antiphase domain size runs over rational values as the composition of alloys changes.

Nevertheless, we can distinguish between two groups of commensurate long-period structures. The first group includes such alloys as Ag_3Mg, Cu_3Al, Au_3Cd, Au_3Zn, and Au_3Mn. The mean half-period remains virtually constant in these alloys as temperature varies. The second group includes the Al_3Ti, Au_4Zn, and Pt_3V alloys. These alloys sometimes show a very substantial temperature dependence of the mean domain size.

In this work, we study the differences in the temperature dependences of domain dimensions in two types of commensurate long-period structures exemplified by the Ag_3Mg and Al_3Ti alloys.

2. CALCULATION DETAILS

The principal value calculated in our approach is the generalized susceptibility of noninteracting electrons $\chi(\mathbf{q})$,

$$\chi(\mathbf{q}) = \frac{2\Omega}{(2\pi)^3} \times \int d^3\mathbf{k} \sum_{\lambda,\lambda'} \frac{f(\varepsilon_\lambda(\mathbf{k}))[1 - f(\varepsilon_{\lambda'}(\mathbf{k}+\mathbf{q}))]}{\varepsilon_{\lambda'}(\mathbf{k}+\mathbf{q}) - \varepsilon_\lambda(\mathbf{k})}.$$

This value is calculated from the electron energy spectrum of the alloy. If the system contains plane Fermi surface regions or Fermi surface regions whose shapes coincide and these regions are separated by the nesting vector $2\mathbf{k}_F$, the $\chi(\mathbf{q})$ function has some singularity at the same vector. Depending on the "quality" of nesting (that is, on the degree of similarity of the coinciding regions), the electronic susceptibility singularity may be a kink, a step, or even a peak. The more pronounced the singularity of $\chi(\mathbf{q})$, the larger the energy gain of the formation of a long-period structure.

The electron energy spectrum was calculated by the full-potential linear muffin-tin orbital (LMTO) method in the local electron density approximation [20]. The exchange-correlation potential was taken according to Barth and Hedin [21], the integration over occupied states was performed by the tetrahedron method [22], 165 (Ag_3Mg) and 126 (Al_3Ti) reference points were used in self-consistent calculations of the $\varepsilon_\lambda(\mathbf{k})$ spectrum, and 1771 (Ag_3Mg) and 4851 (Al_3Ti) points in the irreducible part of the Brillouin zone were used in electronic susceptibility $\chi(\mathbf{q})$ calculations. The generalized susceptibility of noninteracting electrons $\chi(\mathbf{q})$ was calculated only taking into account the energy bands that intersected the Fermi level and determined the behavior of this value. The lattice parameter a was set equal to 7.766 au for Ag_3Mg and 7.264 au for Al_3Ti with $c/a = 2.23$.

3. RESULTS AND DISCUSSION

The electron energy spectrum $\varepsilon(\mathbf{k})$ calculated for a hypothetical Ag_3Mg alloy with $L1_2$-type ordering is shown in Fig. 1. Its characteristic feature, like that of all noble metal-based alloys, is a bright Ag d-band localized below the Fermi level in such a way that the geometry of the Fermi surface is determined by the 17, 18, and 19 bands, which are virtually fully s and p states of Ag and Mg.

The susceptibility $\chi(\mathbf{q})$ calculated along the Γ–X Brillouin zone direction for the Ag_3Mg alloy with

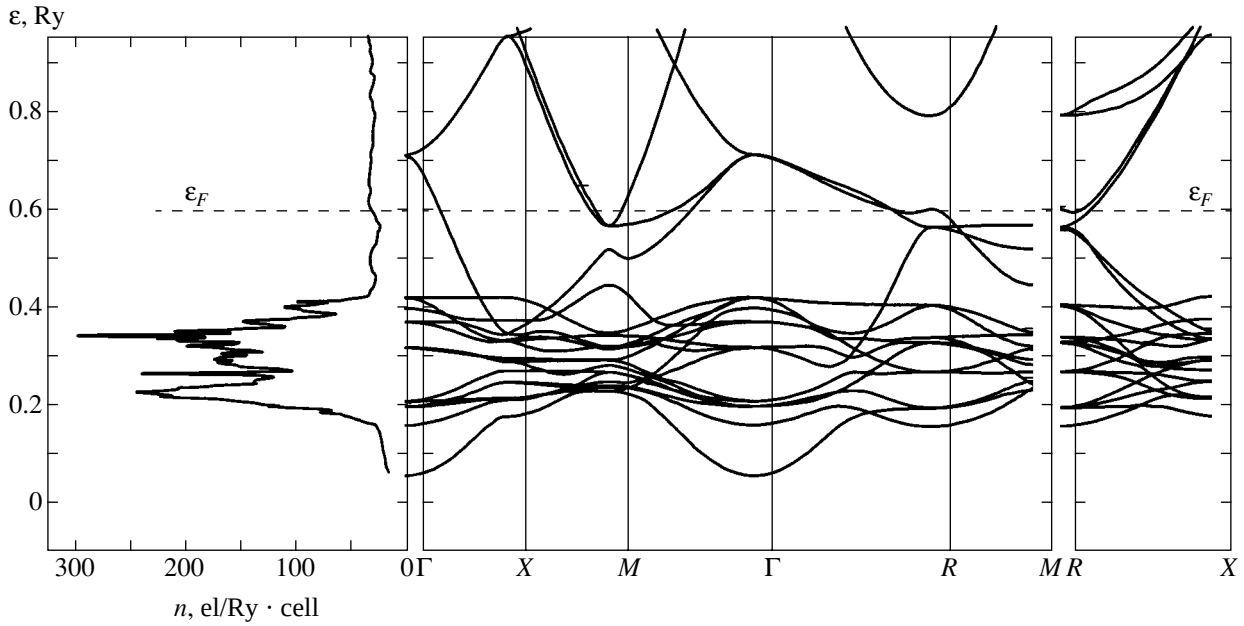

Fig. 1. Electron energy spectrum $\varepsilon(\mathbf{k})$ and the density of electronic states $n(\varepsilon)$ of the Ag_3Mg alloy with the $L1_2$ structure.

$L1_2$-type ordering is shown in Fig. 2. It has a sharp maximum at the wave vector $\mathbf{q}_n = 2\pi/a$ [0.284, 0, 0], which is evidence of the instability of the hypothetical $L1_2$ phase with respect to the formation of a long-period structure with half-period $\overline{M} = 1.76$. This corresponds with experimental observations. Indeed, the Ag_3Mg system has never been observed to have the $L1_2$ structure, it undergoes the transition immediately from the disordered to the long-period state characterized by a mixture of domains with the $D0_{22}$ ($M = 1$) and $D0_{23}$ ($M = 2$) ordering types. The elements of the $D0_{23}$ superstructures are more numerous, and the mean half-period of the antiphase domain $\overline{M}$ is 1.75 [10], which is in excellent agreement with the calculated value specified above (1.76).

An analysis of the partial contributions to the total susceptibility $\chi(\mathbf{q})$ showed that the susceptibility maximum appeared because of the interband electron transitions 18–19 and 19–18 (Fig. 2). Ultimately, this maximum is caused by the geometric features of the Fermi surface shown in Fig. 3. In the vicinity of the Brillouin zone point M, there are two vast electronic regions of the 18th and 19th Fermi surface sheets which virtually coincide in shape and are separated by the $\mathbf{q}_n = 2\pi/a$ [0.284, 0, 0] vector mentioned above. This high degree of similarity is responsible for the well-defined $\chi(\mathbf{q})$ maximum. It can therefore be suggested that the specified special features of the geometry of the Fermi surface and, ultimately, the electronic structure inherent in the hypothetical Ag_3Mg phase with $L1_2$-type ordering contribute to the experimentally observed mixture of domains with the $D0_{22}$ and $D0_{23}$ superstructures.

The electron energy spectrum $\varepsilon(\mathbf{k})$ of the Al_3Ti alloy with the $L1_2$ structure is shown in Fig. 4. The $\chi(\mathbf{q})$ susceptibility along the ⟨100⟩ Brillouin zone direction cal-

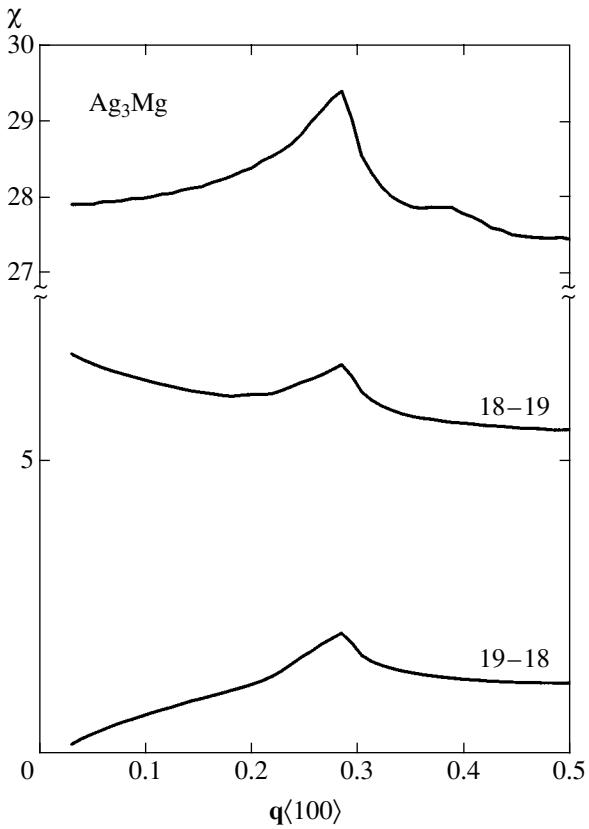

Fig. 2. Electronic susceptibility $\chi(\mathbf{q})$ (the upper curve) and its partial contributions calculated for Ag_3Mg in the ⟨100⟩ direction.

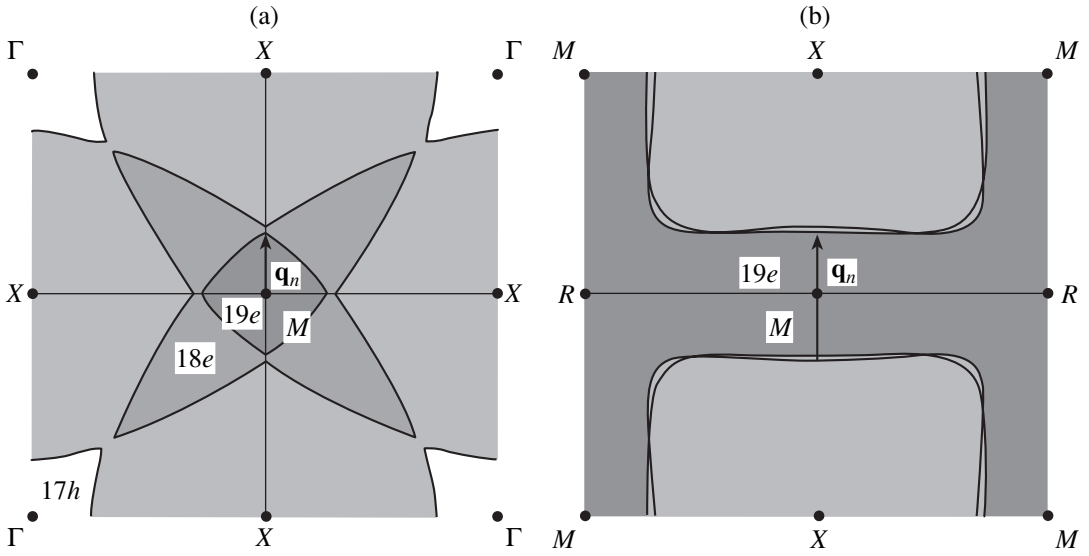

Fig. 3. Fragments of Fermi surface sections for Ag_3Mg: (a) in the $z = 0$ plane and (b) in the $z = 0.5(2\pi)/a$ plane. The nesting vector is shown by the arrow.

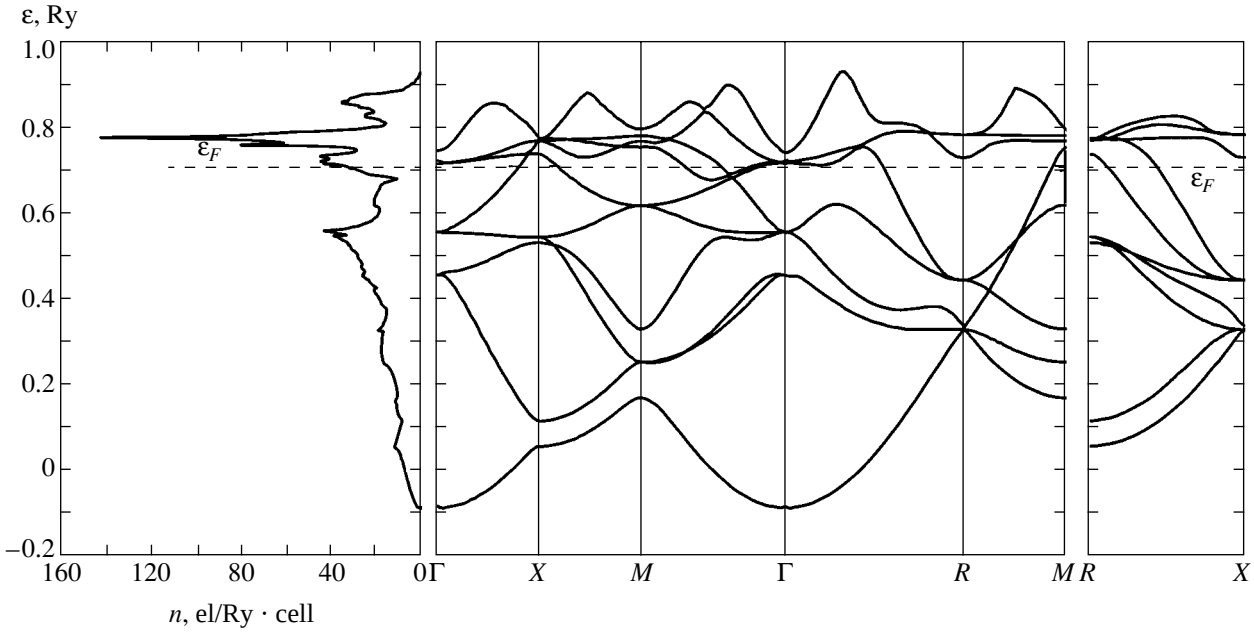

Fig. 4. Electron energy spectrum $\varepsilon(\mathbf{k})$ and the density of electronic states $n(\varepsilon)$ of the Al_3Ti alloy with the $L1_2$ structure.

culated from the electron energy spectrum of Al_3Ti is shown in Fig. 5. Of interest is the local susceptibility maximum at $\mathbf{q}_n = 0.35$; it characterizes the instability of this system with respect to long period formation with $\bar{M} = 1.47$. An analysis of the partial contributions to the total $\chi(\mathbf{q})$ susceptibility shows that the contribution of 7–7 intraband transitions shown in Fig. 5 is fully responsible for this local susceptibility maximum. One can see that, at $\mathbf{q}_n = 0.35$, this contribution has a pronounced maximum, which is evidence of the presence of coinciding Fermi surface regions separated by the vector $\mathbf{q}_n = 0.35(2\pi/a)\langle 001\rangle$ (Fig. 6).

Note that the observed low-temperature $D0_{22}$ structure has a fairly high degree of tetragonality $c/a = 2.23$ [23]. It would therefore be reasonable to calculate the electron energy spectrum and susceptibility for the $L1_2$ structure with a tetragonal distortion equal to that in the $D0_{22}$ structure, that is, $c/a = 1.115$. One point should be mentioned. It can be suggested from general considerations that the larger the long period length, the smaller should

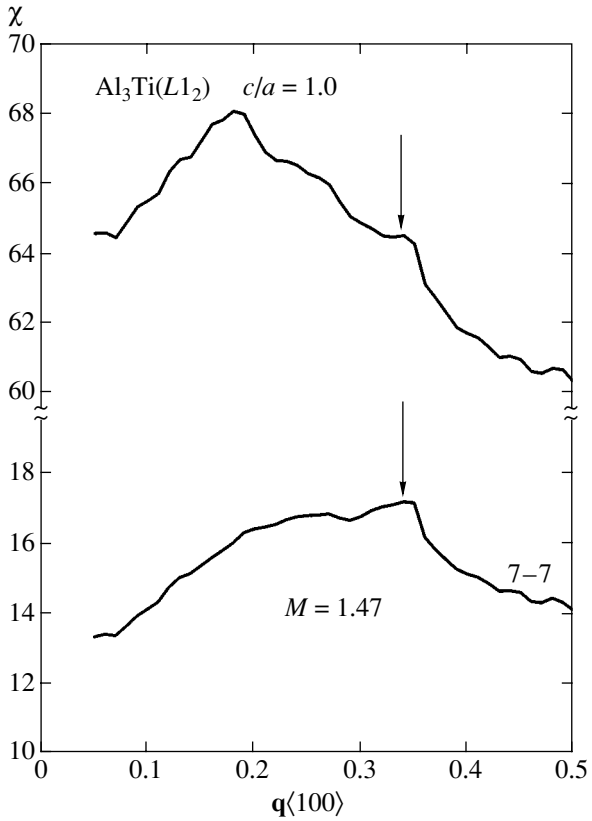

Fig. 5. Electronic susceptibility $\chi(\mathbf{q})$ of the Al_3Ti alloy calculated for the $L1_2$ structure in the $\langle 001\rangle$ direction. The lower curve is the partial contribution of intraband 7–7 electronic transitions. The arrow indicates the local susceptibility maximum.

be the degree of tetragonality of the base cells that constitute the given structure. In the limit of an infinitely long period, this long-period structure transforms into the usual cubic $L1_2$ structure. On the other hand, the greatest degree of tetragonality should be characteristic of the simplest long-period structure of Al_3Ti, that is, $D0_{22}$. It follows that, for completeness, we must also study structures with intermediate tetragonal distortion degrees.

We analyzed the dependence of $\chi(\mathbf{q})$ on the c/a parameter in the range of parameter values from 1 to 1.115. The singularity observed in the cubic $L1_2$ structure at $\mathbf{q}_n = 0.35(2\pi/a)\langle 001\rangle$ shifts to the right as c/a increases and gradually vanishes (see Fig. 7). A new singularity at $\mathbf{q}_n = 0.42(2\pi/a)$, however, appears at $c/a = 1.10$. This singularity not only shifts to the right but also fairly sharply increases as the degree of tetragonality becomes still larger. At the experimental degree of tetragonality, this singularity corresponds to the vector $\mathbf{q}_n = 0.46(2\pi/a)$, which in turn corresponds to the half-period of length $\overline{M} = 1.1$. There exist coinciding Fermi surface regions for all the vectors specified above.

The above results can be represented in the form of the dependence of the mean half-period length $\overline{M}$ on the tetragonality parameter c/a of the base $L1_2$ cell. At the experimentally observed degree of tetragonality, calculations give $\overline{M} = 1.1$, which is very close to the true low-temperature $D0_{22}$ structure ($M = 1$). The antiphase domain size increases as the degree of tetragonality decreases, and $\overline{M}$ approximates 1.5 in the cubic structure. Here, $(c/a)^{-1}$ plays the role of something like temperature. The lower the degree of tetragonality, the higher the temperature and the longer the long period of the long-period structure. These results can be put in qualitative correspondence with the experimental temperature dependence of $\overline{M}$ [14].

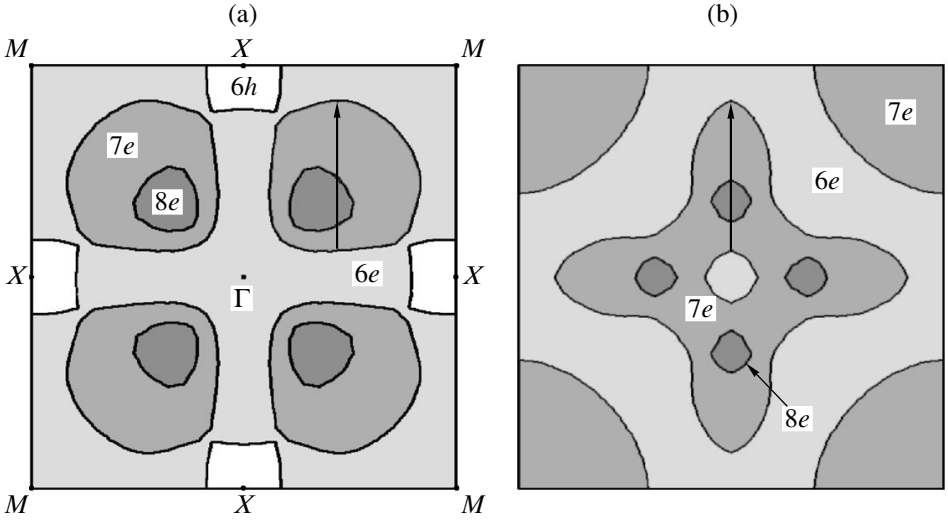

Fig. 6. Fragments of Fermi surface sections for the Al_3Ti alloy in two mutually orthogonal planes: (a) $z = 0$ and (b) $y = 0.23(2\pi/a)$. The nesting vector is shown by the arrow.

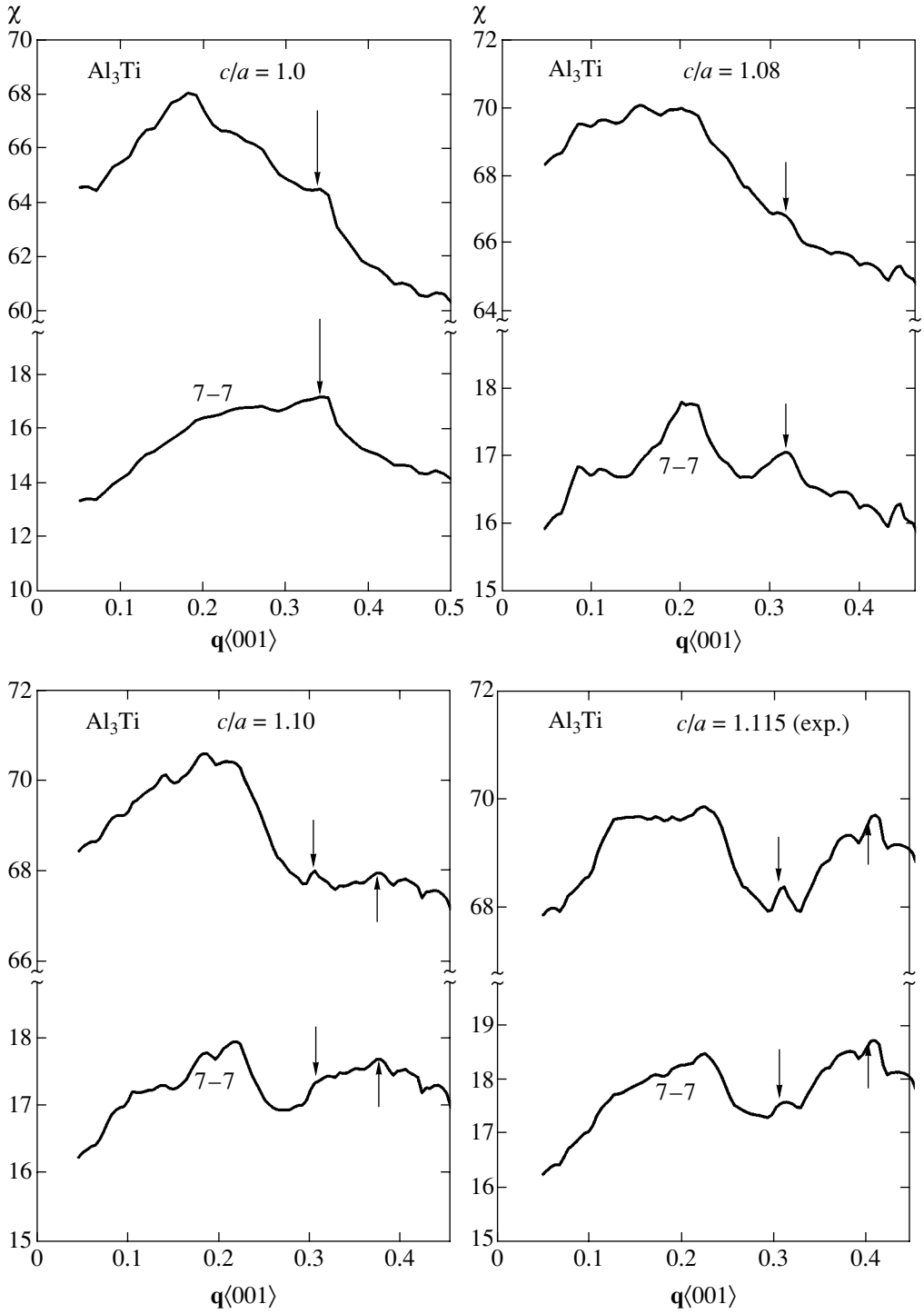

Fig. 7. Electronic susceptibility $\chi(\mathbf{q})$ of the Al_3Ti alloy in the cubic $L1_2$ structure at various degrees of tetragonality (the upper curves) and the partial contributions of the 7–7 intraband transitions (the lower curves). Local susceptibility maxima are indicated by the arrows.

To summarize, the temperature dependence of the size of antiphase domains in commensurate long-period structures is determined by the special features of the electronic structure of the system, in particular, by the geometry of the nesting Fermi surface regions. The degree of coincidence of the flattened Fermi surface sheets is high in the Ag_3Mg alloy, and a fairly sharp $\chi(\mathbf{q})$ maximum is formed. The nesting quality is lower in Al_3Ti, and the susceptibility singularity does not have a pronounced peak. The shape of the Fermi surface and the nesting vector length change as temperature decreases (the degree of tetragonality increases). In

contrast, no nesting changes occur in Ag_3Mg, and long-period structures remain stable even at low temperatures. In our view, this difference explains the difference in stability of commensurate long-period structures under temperature variations.

ACKNOWLEDGMENTS

This work was supported by CRDF and by the Ministry of Education of the Russian Federation under the program "Basic Research and Higher Education," grant V1-P-16-09.

REFERENCES

1. A. I. Potekaev, I. I. Naumov, V. V. Kulagina, V. N. Udodov, O. I. Velikokhatnyĭ, and S. V. Eremeev, *Natural Long-Period Nanostructures* (NTL, Tomsk, 2002).
2. O. D. Shashkov, L. N. Buinova, V. I. Syutkina, *et al.*, Fiz. Met. Metalloved. **28**, 1029 (1969).
3. L. N. Buĭnova, V. I. Syutkina, O. D. Shashkov, and É. S. Yakovleva, Fiz. Met. Metalloved. **29**, 1221 (1970).
4. L. N. Buĭnova, V. I. Syutkina, O. D. Shashkov, and É. S. Yakovleva, Fiz. Met. Metalloved. **33**, 1195 (1972).
5. L. N. Buĭnova, V. I. Syutkina, O. D. Shashkov, and É. S. Yakovleva, Fiz. Met. Metalloved. **34**, 561 (1972).
6. V. I. Syutkina and É. S. Yakovleva, Fiz. Tverd. Tela (Leningrad) **6**, 2688 (1966) [Sov. Phys. Solid State **6**, 2141 (1966)].
7. M. J. Marsinkowski and L. Zwell, Acta Metall. **11**, 373 (1963).
8. O. I. Velikokhatnyĭ, S. V. Eremeev, I. I. Naumov, and A. I. Potekaev, Zh. Éksp. Teor. Fiz. **117**, 548 (2000) [JETP **90**, 479 (2000)].
9. O. I. Velikokhatnyĭ, S. V. Eremeev, I. I. Naumov, and A. I. Potekaev, Pis'ma Zh. Éksp. Teor. Fiz. **69**, 548 (1999) [JETP Lett. **69**, 589 (1999)].
10. M. Guymont and D. Gratias, Acta Crystallogr. A **35**, 181 (1979).
11. A. Gangulee and B. Bever, Trans. Metall. Soc. AIME **242i**, 278 (1968).
12. K. Hanhi, J. Maki, and P. Paalassalo, Acta Metall. **19**, 15 (1971).
13. A. I. Potekaev, Izv. Vyssh. Uchebn. Zaved., Fiz., No. 6, 3 (1995).
14. A. Loiseau, G. van Tendeloo, R. Portier, *et al.*, J. Phys. (Paris) **46**, 595 (1985).
15. A. Loiseau, J. Electron Microsc. **35** (Suppl. 1), 843 (1986).
16. R. Milida, Jpn. J. Appl. Phys. **25**, 1815 (1986).
17. A. Loiseau, J. Planes, and F. Ducastelle, in *Alloy Phase Stability: Proceedings of NATO Advanced Study Institute, Maleme, 1987*, Ed. by G. M. Stocks and A. Gonis (Kluwer Academic, Dordrecht, 1989), p. 101.
18. J. Planes, A. Loiseau, F. Ducastelle, and G. van Tendeloo, in *EMAG 87: Analytical Electron Microscopy*, Ed. by C. W. Lorimer (Inst. of Metals, London, 1988), p. 261.
19. D. Schryvers and S. Amelinckx, MRS Bull. **20**, 367 (1983).
20. S. Savrasov and D. Savrasov, Phys. Rev. B **46**, 12181 (1992).
21. U. Barth and L. Hedin, J. Phys. C **5**, 1629 (1972).
22. J. Rath and A. J. Freeman, Phys. Rev. B **11**, 2109 (1975).
23. W. B. Pearson, *A Handbook of Lattice Spacings and Structures of Metals and Alloys* (Pergamon, New York, 1958).

Translated by V. Sipachev

Journal of Experimental and Theoretical Physics, Vol. 98, No. 3, 2004, pp. 572–581.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 125, No. 3, 2004, pp. 648–658.

SOLIDS
Electronic Properties

Characteristics of the Phase-Separated State in Manganites: Relationship with Transport and Magnetic Properties

K. I. Kugel[a], A. L. Rakhmanov[a], A. O. Sboychakov[a], M. Yu. Kagan[b], I. V. Brodsky[b], and A. V. Klaptsov[b]

[a]*Institute of Theoretical and Applied Electrodynamics, Russian Academy of Sciences, Izhorskaya ul. 13/19, Moscow, 125412 Russia*

[b]*Kapitza Institute for Physical Problems, Russian Academy of Sciences, ul. Kosygina 2, Moscow, 119334 Russia*

e-mail: kugel@orc.ru

Received August 19, 2003

Abstract—The temperature and magnetic field dependence of the resistivity, magnetoresistance, and magnetic susceptibility of phase-separated manganites in the temperature range corresponding to nonmetallic behavior are considered within the framework of a model of inhomogeneous state with allowance for the existence of ferromagnetically correlated regions even in the absence of long-range magnetic order. The main attention is given to the interval of high temperatures and weak fields. The main characteristics of the phase-separated state of manganites are evaluated from a comparison of the theoretical results with available experimental data.

1. INTRODUCTION

Unusual properties and a rich phase diagram of manganites inspired a large number of papers devoted to various aspects of the physics of these compounds. The special interest in manganites in recent years is related to the possible existence of various inhomogeneous charge and spin states such as lattice and magnetic polarons, droplet and stripe structures, etc. [1–3]. Analogous phenomena are inherent in many strongly correlated systems where the electron–electron interaction energy is greater than the kinetic energy. One of the most spectacular manifestations of such a behavior is the formation of ferromagnetic (FM) droplets (ferrons) in slightly doped antiferromagnetic (AFM) semiconductors [4]. Another example is the formation of a string (linear trace of frustrated spins) during the motion of a hole in an AFM insulator [5].

The above examples refer to the so-called electron phase separation, whereby a single charge carrier changes its local electronic environment. In addition to such a small-scale phase separation, manganites can also feature a large-scale phase separation corresponding to the coexistence of different phases characteristic of first-order phase transitions (e.g., of the transition between AFM and FM states). An example of large-scale phase separation is the formation of relatively large FM droplets in an AFM matrix. Such droplets with linear dimensions on the order of 100–1000 Å were observed, in particular, by the method of neutron diffraction [6]. Note also that the attraction between one-electron FM droplets (mediated by either elastic or magnetic dipole interactions) can result in merging of the ferrons and the formation of intermediate- to large-scale inhomogeneities [7]. There are clear experimental indications suggesting that the phase separation is inherent in both magnetically ordered phases and the paramagnetic (PM) state [1–3, 8]. Therefore, the formation of inhomogeneous states has proved to be a phenomenon typical of manganites in various parts of their phase diagrams. Moreover, the phase separation must strongly influence the magnetic and transport properties of manganites.

The role of phase separation is most often considered in the region of the AFM state and especially in the vicinity of the transition between AFM and FM states. However, as mentioned above, manganites can be inhomogeneous even in the PM state, at temperatures exceeding the corresponding phase transition temperature. An analysis of the available experimental data reveals a substantial similarity in the high-temperature behavior of the resistivity, magnetoresistance, and magnetic susceptibility of various manganites with different low-temperature states [9–12]. In addition, the magnetoresistance turns out to be rather large far from the FM–AFM transition and even in the PM region. Furthermore, the magnetic susceptibility of manganites is significantly higher than that for typical antiferromagnets. These experimental data clearly indicate the existence of significant FM correlations in the high-temperature range.

We will proceed from the assumption that ferromagnetically correlated regions exist in manganites above

the temperatures characterizing the onset of long-range magnetic (FM or AFM) ordering. This assumption allows us to describe the characteristic features of the resistivity, magnetoresistance, and magnetic susceptibility of manganites in a nonmetallic state within the framework of a common model. The consideration below is based on the model of conductivity in phase-separated manganites developed in [9, 13–15] and makes use of the experimental data for manganites of various compositions reported in [9–12]. However, this paper is not restricted to considering only one-electron magnetic droplets (ferrons). A part of the previously obtained results will be generalized to the case of an arbitrary number of electrons in ferromagnetically correlated regions.

In Section 2, the temperature dependence of the resistivity is analyzed for an inhomogeneous state with the density of the ferromagnetically correlated regions far from the percolation threshold. In Sections 3 and 4, we discuss within the same assumptions the magnetoresistance of manganites and their magnetic susceptibility, respectively. It will be shown that the proposed model of the inhomogeneous state provides for an adequate description of the high-temperature behavior of manganites. A comparison of the theoretical results and experimental data allows us to determine the main characteristics of ferromagnetically correlated regions in manganites.

2. RESISTIVITY

The temperature dependence of the resistivity of manganites will be analyzed based on the notions developed previously [13].

This physical picture is essentially as follows. We consider a system comprising small ferromagnetic droplets dispersed in a nonferromagnetic insulating matrix, in which charge is transferred via tunneling of the charge carriers from one droplet to another. In the general case, the tunneling probability depends on an external magnetic field. We assume that the droplets do not overlap and the whole system is far from the percolation threshold. Each droplet may contain k charge carriers. Every new charge carrier tunneling to a given droplet experiences Coulomb repulsion from the carriers already occurring in this droplet. The repulsion energy A is assumed to be relatively large ($A > k_BT$), so that the main contribution to the conductivity is related to the processes involving droplets containing k, $k + 1$, and $k - 1$ carriers.

Under these conditions, an expression for the resistivity $\rho(T)$ has the following form:

$$\rho = \frac{k_BT\exp(A/2k_BT)}{128\pi e^2\omega_0 l^5 k n^2}, \qquad (1)$$

where e is the electron charge, ω_0 is a frequency corresponding to the characteristic energy of electrons in a droplet, l is the characteristic tunneling length, and n is the concentration of FM droplets. Formula (1) can be readily derived as described in [13] (see Appendix). This expression is a straightforward generalization of the formula obtained in [13] for the conductivity of one-electron droplets. The resistivity (1) exhibits a thermoactivation behavior, whereby the activation energy is equal to a half of the Coulomb repulsion energy (for more detail, see [13]).

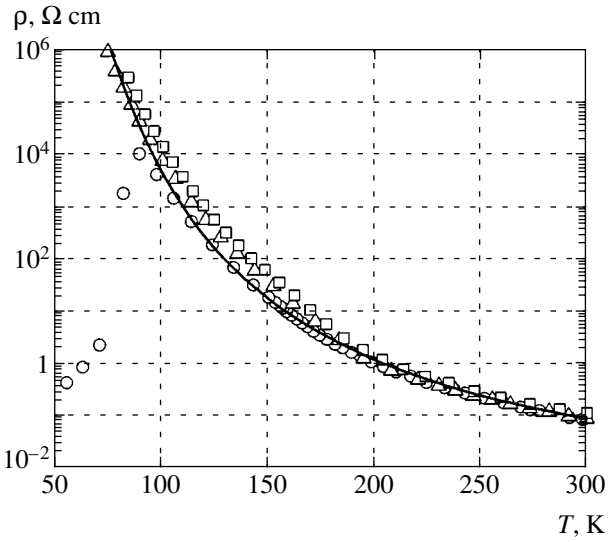

Fig. 1. Temperature dependence of the resistivity of $(La_{1-y}Pr_y)_{0.7}Ca_{0.3}MnO_3$ [9]: (squares) $y = 1$, with $^{16}O \longrightarrow {}^{18}O$ isotope substitution; (triangles) $y = 0.75$, with $^{16}O \longrightarrow {}^{18}O$ isotope substitution; (circles) $y = 0.75$, with ^{16}O isotope. The solid curve shows the results of calculations using formula (1).

Expression (1) provides for a fairly good description of the temperature dependence of the resistivity for various manganites. This is illustrated in Figs. 1–4, showing the $\rho(T)$ curves for six manganites, constructed using the experimental results reported in [9–12] (detailed numerical data were kindly provided by the authors of these papers). As can be seen from these data, the samples differ in chemical composition, the type of crystal structure, the magnitude of the resistivity (at a fixed temperature, the latter varies for different samples by more than two orders of magnitude), and the low-temperature behavior (some of the samples behave as metals and the other, as insulators). In the high-temperature range (above the point of the FM phase transition), $\rho(T)$ exhibits identical behavior for all samples and is well described by the universal relation

$$\rho(T) \propto T\exp(A/2k_BT)$$

represented by solid curves in Figs. 1–4.

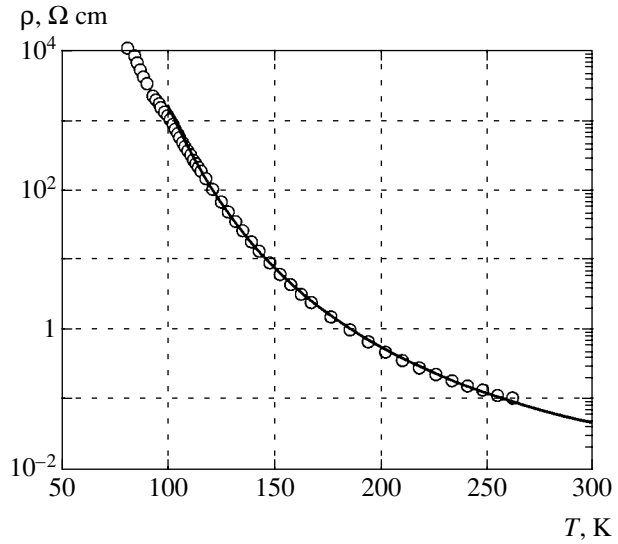

Fig. 2. Temperature dependence of the resistivity of $Pr_{0.71}Ca_{0.29}MnO_3$: (circles) experimental data [10]; (solid curve) calculation using formula (1).

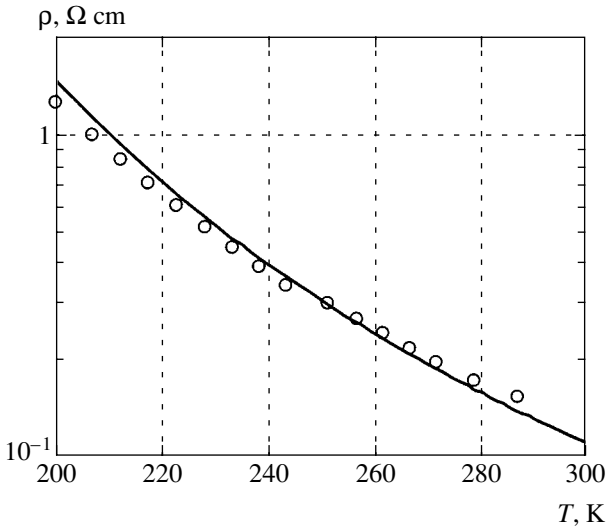

Fig. 3. Temperature dependence of the resistivity of a layered manganite with the composition $(La_{0.4}Pr_{0.6})_{1.2}Sr_{1.8}Mn_2O_7$: (circles) experimental data [11]; (solid curve) calculation using formula (1).

Using Eq. (1) and experimental data, it is possible to determine some quantitative characteristics of the phase-separated state. In particular, an analysis carried out by Zhao *et al.* [12, 16] demonstrated that interpretation of the experimental data in terms of this relation gives an accurate estimate of the Coulomb energy *A*. For example, the data presented in Figs. 1–4 allow the Coulomb barrier *A* to be determined with an accuracy of 2–3%; for the compounds under consideration, this value falls within a rather narrow range from 3500 to 3700 K (see Table 1). As for the characteristic frequency ω_0 entering into formula (1), it was pointed out in [12, 13, 16] that this quantity varies within a very restricted interval of 10^{13}–10^{14} Hz. This estimate can be obtained, for example, from the uncertainty principle: $\hbar\omega_0 \sim \hbar^2/2ma^2$, where *a* is the characteristic size of the droplet and *m* is the electron mass. Indeed, assuming that $a \approx 1$–2 nm, we arrive at the latter interval. These values of the droplet size provide for a correct (to within an order of magnitude) estimate of the Coulomb barrier energy *A*: taking into account that this energy is on the order of $e^2/\varepsilon a$ and substituting permittivity $\varepsilon \sim 10$ we obtain the value of *A* consistent with the experimental data.

It is rather difficult to estimate the tunneling length *l*. However, we can ascertain that, in the domain of applicability of relation (1), this length cannot be much smaller than the distance between droplets [13]. Otherwise, the behavior of the resistivity would be different. In the quasiclassical approximation, the tunneling length is on the order of the characteristic size of the wave function, provided that the barrier height is comparable to the depth of the potential well. In our case, the size of the electron wave function is on the order of the ferron size and the height of the barrier virtually coincides with the depth of the potential well. The latter naturally follows from the adopted model of ferron formation [2]. Therefore, it is reasonable to suggest that the tunneling length is on the order of the ferron size (several nanometers), although in the general case it can substantially differ from *a*.

Another nontrivial task is to estimate the concentration *n* of ferrons. On the one hand, following Zhao *et al.* [12, 16], the concentration *n* could be determined via the dopant concentration *x* as $n \approx x/d^3$. However, this

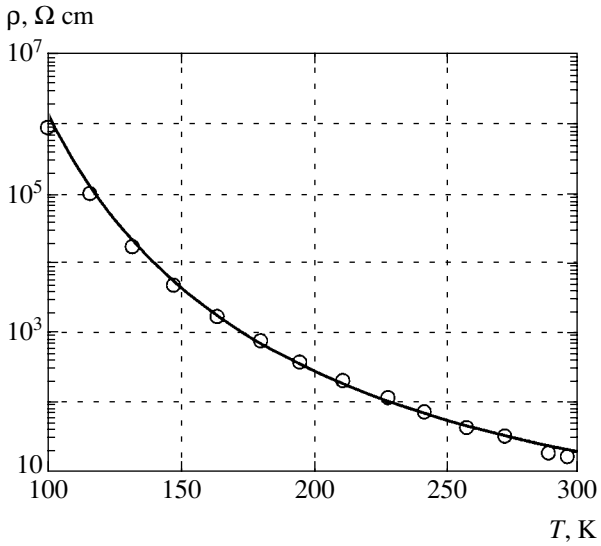

Fig. 4. Temperature dependence of the resistivity of $La_{0.8}Mg_{0.2}MnO_3$: (circles) experimental data [12]; (solid curve) calculation using formula (1).

Table 1. The Coulomb energy A, the resistivity ρ (at 200 K), and the product l^5n^2k for some manganites, estimated using formula (1) and the experimental data presented in Figs. 1–4

Manganite composition	A, K	ρ(200 K), Ω cm	l^5n^2k, cm^{-1}	Ref.
$(La_{1-y}Pr_y)_{0.7}Ca_{0.3}MnO_3$	3650	1.25	2×10^5	[9] (see Fig. 1)
$Pr_{0.71}Ca_{0.29}MnO_3$	3500	0.57	3×10^5	[10] (see Fig. 2)
$(La_{0.4}Pr_{0.6})_{1.2}Sr_{1.8}Mn_2O_7^*$	3600	1.5	1.5×10^5	[11] (see Fig. 3)
$La_{0.8}Mg_{0.2}MnO_3$	3700	283	1×10^3	[12] (see Fig. 4)

* The general chemical formula of this system is $(La_{0.4}Pr_{0.6})_{2-2x}Sr_{1+2x}Mn_2O_7$.

approach leads to at least two discrepancies. First, even for a moderate concentration of a divalent element, x = 0.1–0.2, the droplets would overlap to form a continuous FM metallic cluster. However, the material can be insulating even at higher dopant concentrations of x = 0.5–0.6, at least in a high-temperature range. Second, as can be seen from the experimental data, a relation between the dopant concentration and the conductivity of manganites is rather complicated. For some compounds, a twofold variation of x can change the resistivity by two orders of magnitude [12, 16], while in some other cases, $\rho(x)$ exhibits a nonmonotonic behavior within a certain concentration range.

It should be noted that these difficulties are inherent not only in our model of phase separation, but also arise in other models attempting to describe the properties of manganites (e.g., in polaronic models [17, 18]). Unfortunately, this circumstance was not given proper attention in [12, 16] in the interpretation of experimental data in terms of the existing models of the conductivity in manganites. The natural conclusion is that the number of carriers involved in the charge transfer process does not coincide with the concentration x of a divalent dopant. This is especially obvious in the case of charge ordering, when part of the charge carriers introduced by doping are localized and form a periodic structure.

Therefore, using expression (1) and experimental data, it is possible to evaluate the combination l^5n^2k. Table 1 summarizes the values of the Coulomb energy A, the resistivity ρ (at 200 K), and the product l^5n^2k estimated using formula (1) and the experimental data presented in Figs. 1–4. Note that, while the estimate for A is accurate to within ±50 K, the combination l^5n^2k can be estimated only by the order of magnitude (at least, such is the uncertainty of the value of ω_0).

3. MAGNETORESISTANCE

Previously [9, 14, 15], it was demonstrated that the model of phase separation adopted here leads to a rather unusual dependence of the magnetoresistance MR(T, H) on the temperature and magnetic field. At relatively high temperatures and not very strong magnetic fields, the magnetoresistance is described by the formula (see Appendix)

$$\mathrm{MR} \approx 5\times10^{-3}\frac{\mu_B^3 S^5 N_{\mathrm{eff}}^3 Z^2 g^3 J^2 H_a}{(k_B T)^5}H^2. \qquad (2)$$

Here, μ_B is the Bohr magneton, S is the average spin of a manganese ion, N_{eff} is the number of manganese atoms in a droplet, Z is the number of nearest neighbors of a manganese ion, g is the Landé factor, J is the exchange integral of the FM interaction, and H_a is the effective magnetic anisotropy field of a droplet. The law MR $\propto H^2/T^5$ was experimentally observed for a number of manganites in the region of their nonmetallic behavior [9, 10]. The same high-temperature behavior of the magnetoresistance can be obtained by processing the experimental data reported in [11, 12] (see Figs. 5–8].

The value of S depends on the relative content of the trivalent and tetravalent manganese ions and varies from 3/2 to 2. For purposes of estimation, we assume that S = 2. The value of Z is, in fact, the number of manganese ions interacting with a conduction electron occurring in a droplet. It is reasonable to assume that Z is of the same order of magnitude as the number of nearest neighbors for a manganese ion, that is, $Z \approx 6$. The Landé factor g is determined from experimental data. For manganese, g is usually assumed to be close to that in the pure spin case (g = 2). The exchange integral J characterizes the magnetic interaction between a conduction electron and a molecular field generated by ferromagnetically correlated spins in a droplet. This molecular field accounts for the low-temperature ferromagnetism. Therefore, we can evaluate J using the well-known relationship of the molecular field theory, $S(S+1)ZJ/3 \sim k_B T_C$, where T_C is the Curie temperature. The latter parameter is determined from experiment (e.g., neutron diffraction or magnetization measurements). For example, T_C of manganites of the La–Pr–Ca system is about 100–120 K [6].

The magnetic anisotropy of manganites related to the crystal structure of these compounds is usually not

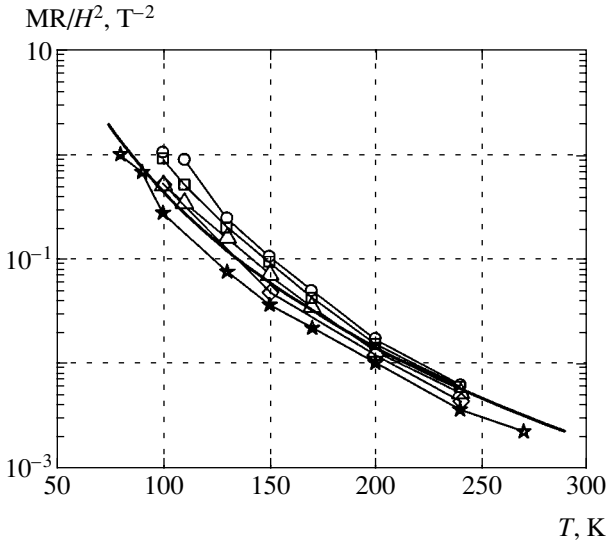

Fig. 5. Temperature dependence of the MR/H^2 ratio for $(La_{1-y}Pr_y)_{0.7}Ca_{0.3}MnO_3$ [9]: (circles) $y = 0.75$; (squares) $y = 0.75$, with 30% of $^{16}O \longrightarrow {}^{18}O$ isotope substitution; (triangles) $y = 0.75$, with $^{16}O \longrightarrow {}^{18}O$ isotope substitution; (diamonds) $y = 1$; (asterisks) $y = 1$, with $^{16}O \longrightarrow {}^{18}O$ isotope substitution. The solid curve shows the results of calculations using relation (2): MR $\propto 1/T^5$.

too high. This implies that the main contribution to the effective magnetic anisotropy field H_a is related to the droplet shape anisotropy and can be evaluated as $H_a = \pi(1 - 3\tilde{N})M_s$, where $\tilde{N}$ is the demagnetization factor of the droplet (along the main axis) and M_s is the magnetic moment per unit volume of the droplet. Below, we assume that the droplet is rather elongated ($\tilde{N} \ll 1$) and $M_s = Sg\mu_B\pi/d^3$, so that $H_a \approx 2$ kOe.

The value of N_{eff} is naturally determined by the size of a droplet and, in principle, it could be found from the neutron diffraction experiments. However, to the best of our knowledge, no such measurements were performed in a wide temperature range for the systems under consideration. For this reason, N_{eff} is treated here as a fitting parameter. Using Eq. (2) and the above estimates, we can determine the N_{eff} value from experimental data on the magnetoresistance (in the range of parameters corresponding to MR $\sim H^2/T^5$). The results of such data processing are summarized in Table 2. In Figs. 5–8, solid curves show the data used in the fitting procedure. The value of T_C was taken to be equal to 120 K.

Thus, the size of the ferromagnetically correlated regions for all manganite compositions under consideration turns out to be nearly the same at temperatures of 200–300 K. The volume of these regions is approximately equal to that of a ball with a diameter of 7–8 lattice periods. It can naturally be assumed that, within a droplet volume, the number of charge carriers involved in the tunneling process equals the number of dopant atoms. Hence, we can write that $k = N_{eff}x$, where x is the atomic fraction of the dopant. The values of x and k are presented in Table 2.

4. MAGNETIC SUSCEPTIBILITY

The concentration of droplets can be evaluated based on the magnetic susceptibility data, by assuming that the main contribution to the susceptibility comes from the ferromagnetically correlated regions. At high

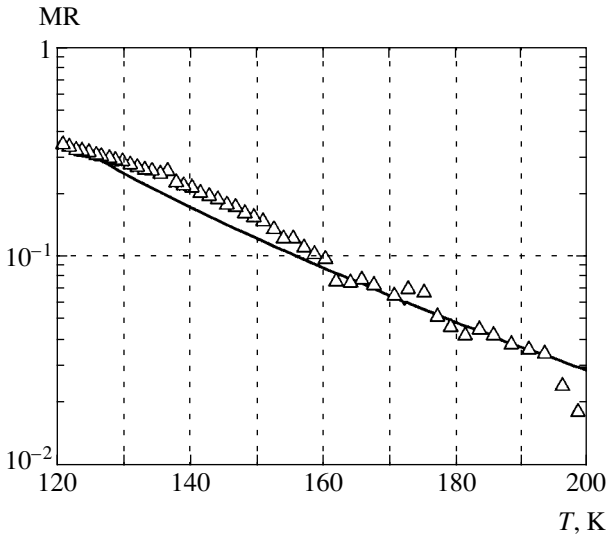

Fig. 6. Temperature dependence of the magnetoresistance of $Pr_{0.71}Ca_{0.29}MnO_3$ for $H = 2$ T (triangles) experimental data [10]; (solid curve) calculation using formula (2).

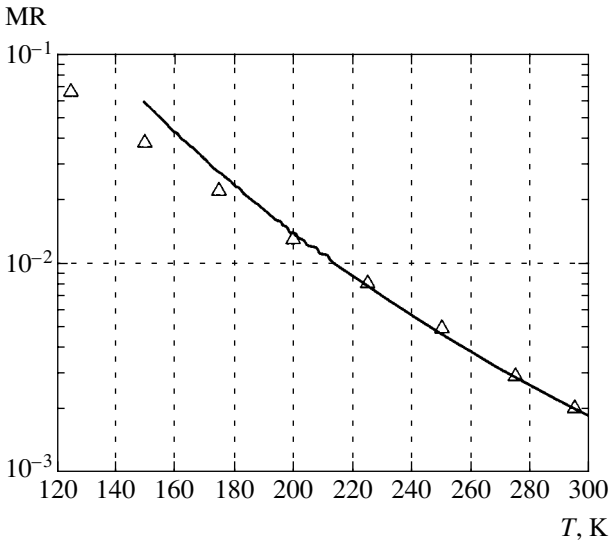

Fig. 7. Temperature dependence of the magnetoresistance of $(La_{0.4}Pr_{0.6})_{1.2}Sr_{1.8}Mn_2O_7$ for $H = 1$ T (triangles) experimental data [11]; (solid curve) calculation using formula (2).

temperatures ($k_BT \gg \mu_B g S N_{eff} H$, $\mu_B g S N_{eff} H_a$), the susceptibility $\chi(T)$ can be written as

$$\chi(T) = \frac{n(\mu_B g S N_{eff})^2}{3k_B(T-\Theta)}, \quad (3)$$

where Θ is the Curie–Weiss constant.

The results of the experimental data processing in terms of relation (3) are presented in Table 3. The experimental temperature dependences of the magnetic susceptibility of manganites are shown in Figs. 9–12, where solid curves illustrate the fitting procedure. Using these results, we can also estimate the concentration of the FM phase as $p = nN_{eff}d^3$. For all samples, the value of lattice parameter d was taken to be equal to 3.9 Å. Based on the data in Tables 1–3, it is also possible to estimate the carrier tunneling length l.

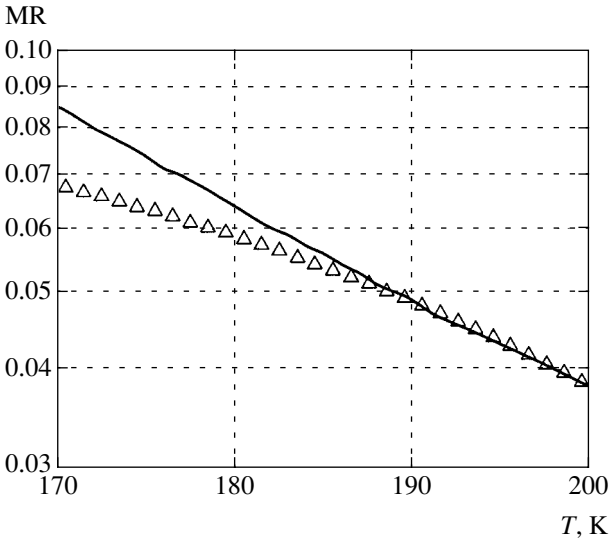

Fig. 8. Temperature dependence of the magnetoresistance of $La_{0.8}Mg_{0.2}MnO_3$ for H = 1.5 T (triangles) experimental data [12]; (solid curve) calculation using formula (2).

5. DISCUSSION

The analysis performed in the previous sections demonstrates that a simple model of electron tunneling between the ferromagnetically correlated regions (FM droplets) provides a description of the conductivity and the magnetoresistance data for a wide class of manganites. A comparison of the theoretical predictions with the experimental data on the temperature dependence of the resistivity, magnetoresistance, and magnetic susceptibility allowed us to determine various characteristics of the phase-separated state, such as the size of FM droplets, their concentration, and the number of electrons in a droplet, and to estimate the characteristic tunneling length of the charge carriers. The obtained values of parameters seem to be quite reasonable. Indeed, the characteristic tunneling length is on the order of the FM droplet size, while the concentration of the FM phase in the high-temperature range is substantially smaller than the percolation threshold and varies from about 1 to 7%. Note also that the droplets contain 50–100 charge carriers and the parameter A evaluated from the experimental data agrees in order of magnitude with the Coulomb energy of an extra electron in a metal ball

Table 2. The effective number N_{eff} of manganese atoms, the number of electrons k in a droplet, and the dopant fraction x for some manganites, estimated using formula (2) and the experimental data presented in Figs. 5–8

Manganite composition	N_{eff}	x	k	Ref.
$(La_{1-y}Pr_y)_{0.7}Ca_{0.3}MnO_3$	250	0.3	75	[9] (see Fig. 5)
$Pr_{0.71}Ca_{0.29}MnO_3$	200	0.29	58	[10] (see Fig. 6)
$(La_{0.4}Pr_{0.6})_{1.2}Sr_{1.8}Mn_2O_7$	250	0.4	100	[11] (see Fig. 7)
$La_{0.8}Mg_{0.2}MnO_3$	265	0.2	53	[12] (see Fig. 8)

Table 3. The Curie–Weiss constant Θ, the density of ferrons n, the FM phase fraction p, and the effective tunneling length l for some manganites, estimated using formula (3) and the experimental data presented in Figs. 9–12

Manganite composition	Θ, K	n, cm^{-3}	p	l, Å	Ref.
$(La_{1-y}Pr_y)_{0.7}Ca_{0.3}MnO_3$	55	1.8×10^{18}	0.03	24	[9] (see Fig. 9)
$Pr_{0.71}Ca_{0.29}MnO_3$	105	6×10^{18}	0.07	17	[10] (see Fig. 10)
$(La_{0.4}Pr_{0.6})_{1.2}Sr_{1.8}Mn_2O_7$	255	2.5×10^{18}	0.04	19	[11] (see Fig. 11)
$La_{0.8}Mg_{0.2}MnO_3$	150	0.6×10^{18}	0.01	14	[12] (see Fig. 12)

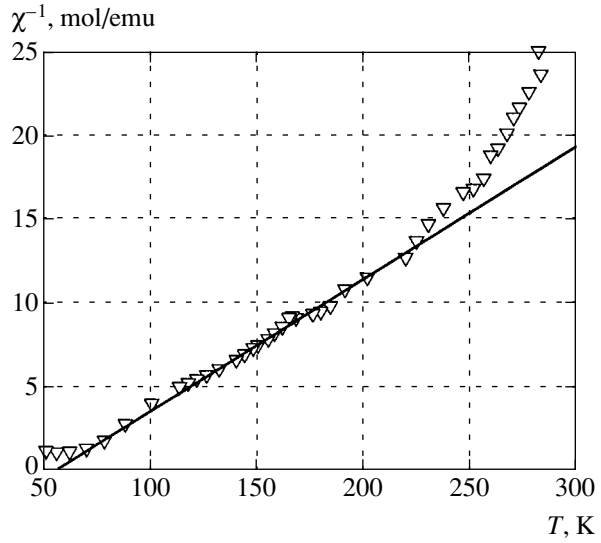

Fig. 9. Temperature dependence of the inverse magnetic susceptibility of $(La_{1-y}Pr_y)_{0.7}Ca_{0.3}MnO_3$ with $y = 1$: (triangles) experimental data [9]; (solid curve) calculations using relation (3). The other manganites of this group exhibit analogous behavior of $\chi(T)$ in the high-temperature range (see [9]).

(7–8)*d* in diameter. The obtained estimates of the droplet parameters (characteristic tunneling barrier, size, and tunneling length) are close for manganites with strongly different transport properties.

Another characteristic feature of the phase-separated manganites is a large magnitude of the $1/f$ noise in the temperature range corresponding to the dielectric state [19, 20]. In the framework of the phase separation model adopted here, the following expression was derived for the Hooge constant [13, 14],

$$\alpha_H = \frac{\langle \delta U^2 \rangle_\omega V_s \omega}{U_{dc}^2} = 2\pi^2 l^3 \ln^2\left(\frac{\tilde{\omega}_0}{\omega}\right), \qquad (4)$$

where $\langle \delta U^2 \rangle_\omega$ is the spectral density of the voltage fluctuations, V_s is the sample volume, U_{dc} is the applied voltage, and $\tilde{\omega}_0 = \omega_0 \exp(A/2k_BT)$. Substituting estimated values of the parameters presented in the tables and in the text, we obtain $\alpha_H \approx 10^{-16}$ cm^3 at a temperature of 100–200 K and frequencies within 1–1000 s^{-1}. This value of α_H is 3–5 orders of magnitude higher than the corresponding values for semiconductors.

Thus, we have a rather consistent scheme describing the transport properties of manganites under conditions when the ferromagnetically correlated regions do not form a percolation cluster. Moreover, the proposed approach proves to be valid for a fairly wide range of dopant concentrations. However, as mentioned above, a relationship between the concentration of FM droplets and the doping level is still incompletely clear. If the above picture of the phase separation is valid, it becomes evident that not all electrons or holes introduced by doping participate in the transport processes. Now we will try to find some qualitative arguments illustrating the possible differences in the effective concentrations of charge carriers below and above the transition from PM to magnetically ordered state.

The phase diagram of a typical manganite contains a region of the AFM state with FM phase inclusions in the range of low temperatures and low doping levels.

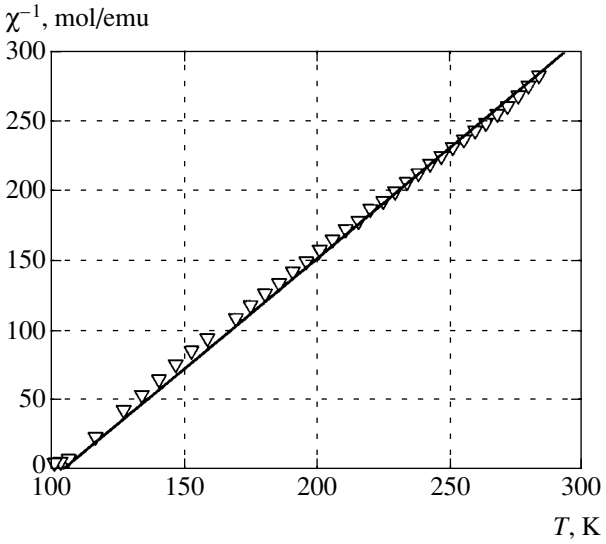

Fig. 10. Temperature dependence of the inverse magnetic susceptibility of $Pr_{0.71}Ca_{0.29}MnO_3$: (triangles) experimental data [10]; (solid curve) calculation using formula (3). The density of a porous sample was assumed to be 0.8 of the theoretical value.

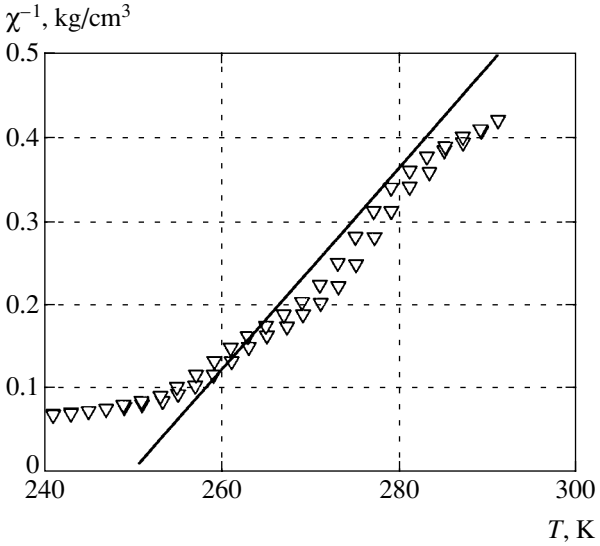

Fig. 11. Temperature dependence of the inverse magnetic susceptibility of $(La_{0.4}Pr_{0.6})_{1.2}Sr_{1.8}Mn_2O_7$: (triangles) experimental data [11]; (solid curve) calculation using formula (3).

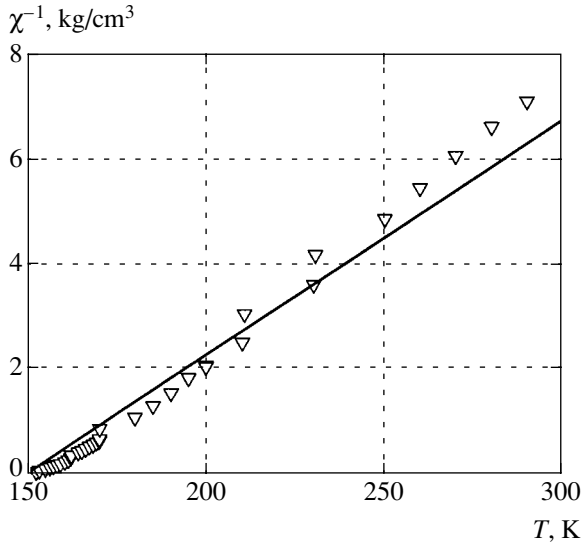

Fig. 12. Temperature dependence of the magnetoresistance of $La_{0.8}Mg_{0.2}MnO_3$: (triangles) experimental data [12]; (solid curve) calculation using formula (3).

When the doping level is increased, the transition from AFM to FM phase is observed. At high temperatures, manganites are in the PM state. As the temperature decreases, the transition from PM to AFM or FM state takes place depending on the doping level. Let us consider the behavior of such a system in the vicinity of a triple point. In the AFM phase, the radius R of a region converted by one electron into the FM state can be estimated as [3]

$$R = d(\pi t/4J_{ff}S^2Z)^{1/5},$$

where J_{ff} is the AFM interaction constant. For the high-temperature PM phase, the radius R_T of a region converted by one electron into the FM state corresponds to the size of the so-called temperature ferron and equals to $R_T = d(\pi t/4k_BT\ln(2S+1))^{1/5}$ [3].

The critical concentration $x_c \approx 0.15$ corresponding to overlap of the low-temperature ferrons can be estimated as $x_c \sim (3/4\pi)(d/R)^3$. For the high-temperature ferrons, the corresponding estimate is $\delta_c \sim (3/4\pi)(d/R_T)^3$. Substituting the expressions for the radii of high- and low-temperature ferrons, we can estimate the ratio x_c/δ_c in the vicinity of the triple point corresponding to the coexistence of FM, AFM, and PM phases:

$$\frac{x_c}{\delta_c} \sim \left[\frac{T\ln(2S+1)}{zJ_{ff}S^2}\right]^{3/5} \sim \left[\frac{T_C\ln(2S+1)}{T_N}\right]^{3/5}, \quad (5)$$

where T_C and T_N are the Curie and the Néel temperatures, respectively. For the manganites under consideration, we have $T_C \sim T_N \sim 120$–150 K and $\ln(2S+1) \sim 1.6$ for $S = 2$, hence $\delta_c \le x_c$. This is consistent with experimental data showing that $\delta \sim 1$–7%. Thus, although formula (5) correctly reflects the observed tendency, the nature of the charge disbalance in the PM region is incompletely clear. Probably, for $x > x_c$ (in real manganites, the concentration x can reach a level of about 50%), the residual charge is localized in the PM matrix outside the temperature ferrons. A detailed study of this problem will be presented elsewhere.

ACKNOWLEDGMENTS

The authors are grateful to V.A. Aksenov, N.A. Babushkina, S.W. Cheong, I. Gordon, L.M. Fisher, D.I. Khomskiĭ, F.V. Kusmartsev, V.V. Moshchalkov, A.N. Taldenkov, I.F. Voloshin, G. Williams and X.Z. Zhou for fruitful discussions and kindly provided experimental data.

This study was supported by the Russian Foundation for Basic Research (project nos. 02-02-16708, 03-02-06320, and NSh-1694.2003.2), INTAS (grant no. 01-2008), and CRDF (grant no. RP2-2355-MO-02).

APPENDIX

We present a short derivation of formulas for the resistivity (1) and the magnetoresistance (2). Consider a system of small ferromagnetic droplets dispersed in a dielectric matrix exposed to an electric field **E**, in which charge is transferred via tunneling of the charge carriers from one droplet to another. In the ground state, each droplet contains k charge carriers. Every new charge carrier tunneling to a given droplet experiences the Coulomb repulsion from the carriers already occurring in this droplet. The repulsion energy A is assumed to be relatively large ($A > k_BT$), so that the main contribution to the conductivity is related to the processes involving droplets containing k, $k+1$, and $k-1$ carriers.

Let N be the total number of droplets in the system and N_1, N_2, and N_3 be the numbers of droplets containing k, $k+1$, and $k-1$ carriers, respectively. Assuming the total number of droplets in the system to be constant, so that $N_2 = N_3$ and $N_1 + 2N_2 = N$, we can write the partition function of the system [13],

$$Z = \sum_{m=0}^{N/2} C_N^m C_{N-m}^m \exp(-m\beta), \quad \beta = \frac{A}{k_BT}, \quad (6)$$

where C_N^m are the binomial coefficients. Using the Stirling formula, replacing summation by integration, and calculating the corresponding integral by the steep-

est descent method, we obtain the following expressions for the average vales of N_1, N_2, and N_3:

$$\begin{aligned} \bar{N}_2 &= \bar{N}_3 = N\exp\left(-\frac{A}{2k_BT}\right), \\ \bar{N}_1 &= N - 2\bar{N}_2 = N\left[1 - 2\exp\left(-\frac{A}{2k_BT}\right)\right]. \end{aligned} \tag{7}$$

According to our model, electron tunneling implies one of the four possible events: (i) two droplets with k electrons are converted into droplets with $k + 1$ and $k - 1$ electrons; (ii) process inverse to that in the previous case; (iii) droplets with k and $k + 1$ electrons exchange sites with each other; (iv) the same for droplets with k and $k - 1$ electrons. The total current density is given by the sum of the contributions from all these processes: $j = j_1 + j_2 + j_3 + j_4$. As was demonstrated in [13], all the four processes equally contribute to the total current.

For example, consider the third process. The density of electrons involved in this exchange is $(k + 1)n_2$, where n_2 is the concentration of droplets with $k + 1$ electrons. Assuming that $k \gg 1$, we obtain

$$j_3 \approx ekn_2\left\langle \sum_i v^i \right\rangle = ekn_2\left\langle \sum_i \frac{r^i\cos\theta^i}{\tau(r^i, \theta^i)} \right\rangle, \tag{8}$$

where $\langle\ldots\rangle$ denotes the statistical average with the sum taken over all droplets with $k + 1$ electrons, $\langle v^i \rangle$ is the average velocity of electrons along the electric field **E**, r^i is the distance between droplets involved in this process, θ^i is the angle between vector **E** and the direction of electron motion, and $\tau(r^i, \theta^i)$ is the characteristic tunneling time. The standard expression for τ is as follows [21]:

$$\tau(r, \theta) = \omega_0^{-1}\exp\left(\frac{r}{l} - \frac{eEr\cos\theta}{k_BT}\right), \tag{9}$$

where l is the tunneling length and ω_0 is the characteristic frequency. Far from the percolation threshold, the average in relation (8) is the spatial average of the velocity v^i multiplied by the number N_1 of droplets accessible for the jump.

Assuming the electric field strength to be small ($eEl/k_BT \ll 1$) and the repulsion energy A to be large ($A/k_BT \gg 1$), we obtain in the first order with respect to the field E:

$$\left\langle \sum_i v^i \right\rangle = \frac{eE\omega_0}{k_BT}Ne^{-A/2k_BT}\langle r^2\cos^2\theta e^{-r/l}\rangle_V, \tag{10}$$

where $\langle\ldots\rangle_V$ denotes the average over the sample volume. Substituting Eq. (10) into relation (8) and integrating over space, we determine the corresponding current component. Taking into account the obvious relation $\rho = j_3/4E$, we arrive at formula (1). As can be seen, the resistivity decreases with the temperature according to the law

$$\rho(T) \propto T\exp\left(\frac{A}{2k_BT}\right),$$

which is characteristic of the systems with tunneling conductivity [21].

Now let us proceed to the calculation of magnetoresistance MR. The probability of tunneling depends, in particular, on the mutual orientation of the electron spin and the magnetic moment of a droplet. Orientation of the ferromagnetically correlated regions in the magnetic field H leads to an increase in the transition probability and, hence, to a decrease in the resistance with increasing field strength—in agreement with experiment. The conductivity of the system can be represented as $\sigma(H) = \sigma_0\langle\Sigma(H)\rangle$, where $\Sigma(H)$ is the "spin" contribution to the probability of electron tunneling. For this definition, MR = $\langle\Sigma(H)\rangle - 1$.

Denoting the effective magnetic moment of a droplet by $M = \mu_B g N_{\text{eff}} S$ and assuming the interaction between droplets to be negligibly small, we write the free energy of a droplet in the magnetic field in the following form:

$$U(H) = U(0) - M(H\cos\theta + H_a\cos^2\psi), \tag{11}$$

where θ is the angle between the applied field **H** and the magnetic moment **M**, H_a is the anisotropy field, and ψ is the angle between the anisotropy axis and the direction of the magnetic moment (for the sake of simplicity, we consider the case of uniaxial anisotropy). Let **H** be parallel to the z axis, and let the anisotropy axis lie in the (x, z) plane and make the angle β with vector **H**. In this configuration,

$$\cos\psi = \sin\theta\sin\beta\cos\varphi + \cos\theta\cos\beta,$$

where φ is the angle between the x axis and the projection of **M** onto the (x, y) plane.

In the classical limit, a given orientation of vector **M** corresponds to the probability

$$\begin{aligned} &P(H, \theta, \varphi) \\ &= A(H)\exp\left[\frac{M(H\cos\theta + H_a\cos^2\psi(\theta, \varphi))}{k_BT}\right], \end{aligned} \tag{12}$$

where $A(H)$ is the normalization factor. The eigenstates of an electron correspond to conservation of the spin projection $s = \pm 1/2$ onto the effective field direction in

a ferromagnetically correlated region. Let an electron interact with Z magnetic moments in the droplet. The energy of this interaction is $E_s = -JSZs$. Since the product JSZ is on the order of the Curie temperature, E_s is much greater than the energy of interaction between the electron spin and the magnetic field, provided that $H \ll 100$ T. In this case, the effective field direction coincides with the direction of vector **M** and the probability for the electron spin projection to be s can be written as

$$P_s = \frac{\exp(-E_s/k_BT)}{2\cosh(E_s/k_BT)}. \quad (13)$$

Upon transfer from droplet 1 to droplet 2, an electron occurs in an effective field making an angle ν with that in the initial state, for which

$$\cos\nu = \cos\theta_1\cos\theta_2 + \sin\theta_1\sin\theta_2\cos(\varphi_1 - \varphi_2)$$

(indices 1 and 2 refer to the droplet number). Then, the work performed for the electron transfer from droplet 1 to droplet 2 is $\Delta E_s = E_s(1 - \cos\nu)$. Accordingly, the probability of this transfer is proportional to $\exp(-\Delta E_s/k_BT)$. Taking into account all the probability factors introduced above, the final expression can be written as

$$\langle\Sigma(H)\rangle = \int_0^{2\pi} d\varphi_1 \int_0^{2\pi} d\varphi_2 \int_0^{\pi} \sin\theta_1 d\theta_1 \times \int_0^{\pi} \sin\theta_2 d\theta_2 P(\theta_1, \varphi_1) P(\theta_2, \varphi_2) \times \sum_{s = \pm 1/2} P_s \exp\left(-\frac{\Delta E_s}{k_BT}\right). \quad (14)$$

In the high-temperature range, where k_BT is much greater compared to the Zeeman energy $\mu_B g S N_{eff} H$ and the magnetic anisotropy energy $\mu_B g S N_{eff} H_a$, relations (12)–(14) yield formula (2).

The limits of applicability of the above expressions for the resistivity and magnetoresistance of manganites are considered in more detail elsewhere [13–15].

REFERENCES

1. E. Dagotto, T. Hotta, and A. Moreo, Phys. Rep. **344**, 1 (2001).
2. E. L. Nagaev, Phys. Rep. **346**, 387 (2001).
3. M. Yu. Kagan and K. I. Kugel', Usp. Fiz. Nauk **171**, 577 (2001) [Phys. Usp. **44**, 553 (2001)].
4. É. L. Nagaev, Pis'ma Zh. Éksp. Teor. Fiz. **6**, 484 (1967) [JETP Lett. **6**, 18 (1967)].
5. L. N. Bulaevskiĭ, É. L. Nagaev, and D. I. Khomskiĭ, Zh. Éksp. Teor. Fiz. **54**, 1562 (1968) [Sov. Phys. JETP **27**, 836 (1968)].
6. A. M. Balagurov, V. Yu. Pomjakushin, D. V. Sheptyakov, *et al.*, Phys. Rev. B **64**, 024420 (2001).
7. J. Lorenzana, C. Castellani, and C. Di Castro, Phys. Rev. B **64**, 235127 (2001); Phys. Rev. B **64**, 235128 (2001).
8. N. I. Solin, V. V. Mashkautsan, A. V. Korolev, *et al.*, Pis'ma Zh. Éksp. Teor. Fiz. **77**, 275 (2003) [JETP Lett. **77**, 230 (2003)].
9. N. A. Babushkina, E. A. Chistotina, K. I. Kugel', *et al.*, Fiz. Tverd. Tela (St. Petersburg) **45**, 480 (2003) [Phys. Solid State **45**, 508 (2003)].
10. L. M. Fisher, A. V. Kalinov, I. F. Voloshin, *et al.*, Phys. Rev. B **68**, 174403 (2003).
11. P. Wagner, I. Gordon, V. V. Moshchalkov, *et al.*, Europhys. Lett. **58**, 285 (2002).
12. J. H. Zhao, H. P. Kunkel, X. Z. Zhou, and G. Williams, J. Phys.: Condens. Matter **13**, 9349 (2001).
13. A. L. Rakhmanov, K. I. Kugel, Ya. M. Blanter, and M. Yu. Kagan, Phys. Rev. B **63**, 174424 (2001).
14. A. O. Sboychakov, A. L. Rakhmanov, K. I. Kugel', *et al.*, Zh. Éksp. Teor. Fiz. **122**, 869 (2002) [JETP **95**, 753 (2002)].
15. A. O. Sboychakov, A. L. Rakhmanov, K. I. Kugel, *et al.*, J. Phys.: Condens. Matter **15**, 1705 (2003).
16. J. H. Zhao, H. P. Kunkel, X. Z. Zhou, and G. Williams, Phys. Rev. B **66**, 184428 (2002).
17. M. Ziese and C. Srinitiwarawong, Phys. Rev. B **58**, 11519 (1998).
18. G. Jakob, W. Westerburg, F. Martin, and H. Adrian, Phys. Rev. B **58**, 14966 (1998).
19. V. Podzorov, M. Uehara, M. E. Gershenson, *et al.*, Phys. Rev. B **61**, R3784 (2000).
20. V. Podzorov, M. E. Gershenson, M. Uehara, and S.-W. Cheong, Phys. Rev. B **64**, 115113 (2001).
21. N. F. Mott and E. A. Davis, *Electronic Processes in Non-Crystalline Materials*, 2nd ed. (Clarendon Press, Oxford, 1979; Mir, Moscow, 1982).

Translated by P. Pozdeev

Journal of Experimental and Theoretical Physics, Vol. 98, No. 3, 2004, pp. 582–593.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 125, No. 3, 2004, pp. 659–672.

SOLIDS
Electronic Properties

Optical Properties of a Coherent Phase in Electron–Hole Systems: Stimulated Light Scattering and Multibeam Processes

Yu. E. Lozovik, I. V. Ovchinnikov, and V. A. Sharapov

Institute of Spectroscopy, Russian Academy of Sciences, Troitsk, Moscow oblast, 142190 Russia
e-mail: lozovik@isan.troitsk.ru

Received September 5, 2003

Abstract—A theoretical study is reported of stimulated light scattering, including wave-vector reversal and anomalous transmission, by a coherent phase in electron–hole (e–h) systems of low and high charge-carrier density. For these two cases the coherent phase is taken to be a Bose–Einstein condensate of excitons or a BCS-like state of e–h pairs, respectively. The scattering mechanism is laser-induced coherent recombination of two excitons or two coherent e–h pairs, respectively. The e–h system is assumed to exist within a GaAs/AlGaAs double quantum well or bulk GaAs. The emission rate of two photons with antiparallel momenta is estimated. Multiphoton emission due to multiexciton coherent recombination is covered. Methods for detecting the effects predicted are proposed. © *2004 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

Theory tells us that macroscopic quantum coherence will arise in the electron–hole (e–h) system of a semiconductor or semimetal if its temperature is decreased to below a certain point [1–8]. Considerable interest has been shown in the detection of the phenomenon in both three-dimensional (3D) and two-dimensional (2D) systems. Recent years have seen significant experimental advances in this field, especially for the quasi-2D system of crossed excitons in a double quantum well (DQW) [9–12]. The predictions that had been made about their behavior include anomalous transport, superfluidity, drag, and quasi-Josephson phenomena [13–31].

Let n be the charge-carrier density. Two extreme cases are generally considered. The high-density case is $na_B^3 \gg 1$ [1] or $na_B^2 \gg 1$ [13] for 3D or 2D systems, respectively. The low-density case is $na_B^3 \ll 1$ [3, 4] or $na_B^2 \ll 1$ [14] for 3D or 2D systems, respectively. Here, a_B is the Bohr radius.

In the low-density case, e–h systems can contain excitons that should pass to a coherent state if cooled to a sufficiently low temperature. For 3D systems a coherent phase may be formed by Bose–Einstein (BE) condensation [3, 4]. A 2D system first displays a localized BE condensate, which can be characterized by an order parameter of fluctuating phase; this state then changes into the low-temperature Berezinski–Kosterlitz–Thouless phase, which shows quasi-long-range order (see [15] and references therein).

In the high-density case, coherent pairing should arise between an electron and a hole in a manner resembling the Bardeen–Cooper–Schrieffer (BCS) mechanism,[1] the size of the e–h pairs being much greater than the mean distance between them [1, 2, 6]. The BCS-like state of quasi-2D e–h systems was studied in [13, 15]. In particular, the case where electrons are spatially separated from holes was covered [13]. Later on, detailed research was conducted into phenomena possible in dense e–h systems, such as superfluidity, phase equilibrium, tunneling, quasi-Josephson phenomena, and drag [13–30].

For intermediate charge-carrier densities, a liquid exciton phase was investigated in quasi-2D [17] and 3D [35] systems. The phase diagram and properties of a quasi-2D system below the transition point were examined in [13, 15].

The formation of a coherent phase should confer new optical properties on the e–h system [36–40]. One

[1] Consider a quasi-equilibrium dense e–h system, which may be found in a semimetal or created by laser irradiation. It can display a coherent phase only if the respective Fermi surfaces of electrons and holes fulfill the nesting condition $v_F\delta p < \Delta$, where v_F is the Fermi velocity, δp is the Fermi-surface separation for the pairing particles, and Δ is the energy gap due to pairing. The transition point falls fairly slowly with decreasing size of the nesting Fermi-surface parts [7]. Over a narrow range of δp a state with a nonuniform gap may also exist [32, 33]. A competing type of ground state is a liquid metallic phase in the form of e–h droplets, which are more stable in multivalley semiconductors such as Ge and Si [34].

of them is a strong, narrow luminescence line related to exciton recombination in the BE condensate; its intensity is proportional to the density n_0 of the BE condensate. However, with strong exciton–exciton interaction, the ground state is characterized by a fairly large density of uncondensed excitons, whose radiative recombination makes it very difficult to detect the condensate line. It is therefore desirable that a better indicator be found of a coherent phase in e–h systems.

This paper theoretically treats two-photon emission due to coherent recombination of two excitons or two coherent e–h pairs in the coherent phase of an e–h system in the low- and the high-density case, respectively. The emission should provide a clear indication of a coherent phase being present in the e–h system because the recombination can be represented in terms of the anomalous Green functions.

In the low-density case the coherent phase is taken to be a BE condensate of excitons [40]. Although the two-exciton recombination is a second-order process with respect to exciton–photon interaction, its weakness is compensated for by its rate varying as n_0^2 rather than n_0 as with one-exciton recombination.

In 3D systems the total momentum of the emitted photons must be zero, because both excitons have zero momentum. The photon momenta are therefore antiparallel. With 2D systems, the argument applies only to photon momentum components parallel to the plane of the system. This feature could be detected from the time-dependent angular correlation of photocounts measured with appropriately arranged photon counters [37, 41], as in the Brown–Twiss experiment.

If the two-exciton coherent recombination in a 3D system is induced by laser photons of momentum **k**, it must produce two photons with the respective momenta **k** and –**k**. The same is true of in-plane photon momentum components in 2D systems. This phenomenon might be seen as the reversal of the wave vector by photoinduced coherent recombination of excitons. In 2D systems, in-plane reversal is responsible for two anomalous beams: reversed and transmitted (Fig. 5). Furthermore, BE-condensed excitons should exhibit multiphoton effects linked to multiexciton coherent recombination [40]. They are considered in what follows.

Three main topics are addressed in this study. First, we examine stimulated light scattering by a 2D coherent phase in a GaAs/AlGaAs DQW (Fig. 1) for the two extreme cases. Second, we investigate optical wave-vector reversal for bulk GaAs in the high-density case. Third, we explore three- and four-exciton coherent recombination in a BE condensate of excitons in Cu_2O. This line of research is taken up in light of experimental efforts to bring about the BE condensation of excitons in Cu_2O [7, 8, 39, 42–49].

2. COHERENT OPTICAL PROCESSES AND THE ANOMALOUS GREEN FUNCTIONS

It is well known that laser irradiation of a semiconductor can generate e–h pairs with a density dependent on the radiation intensity. If the lifetime of such excitations is far longer than their energy relaxation time, the e–h system can be in a quasi-equilibrium state and display a number of equilibrium phases. The recombination of an e–h pair in a direct-gap semiconductor, such as GaAs, involves emitting a photon. If the temperature is low enough, a coherent phase should arise in the e–h system, which would be manifested in nonzero means that relate one-particle states having antiparallel momenta. It is this anomaly that makes it possible to produce two photons with antiparallel momenta (in 3D systems) by annihilation of two excitons or two coherent e–h pairs. This process can also be induced by resonant laser irradiation.

The presence of a coherent phase is indicated by a nonvanishing anomalous Green function. In the low-density case this Green function is

$$\hat{G}(\omega, \mathbf{p}) = -i\int dt e^{i\omega t}\langle T a_{\mathbf{p}}(t) a_{-\mathbf{p}}(0)\rangle$$
$$= -\frac{\beta}{[\omega - (\varepsilon(\mathbf{p}) - i\gamma)][\omega + (\varepsilon(\mathbf{p}) - i\gamma)]},$$

where $a_{\mathbf{p}}$ is the annihilation operator for an exciton of momentum **p**; T is the Wick time-ordering operator; $\varepsilon(\mathbf{p}) \equiv \varepsilon_{\mathbf{p}}$ and γ are the dispersion relation and the inverse lifetime of elementary excitations in the e–h system, respectively; and $\beta = \rho_{\text{cond}} V_0$, with ρ_{cond} and V_0 denoting the spatial density of the exciton BE condensate and the zeroth Fourier coefficient of exciton–exciton interaction, respectively. From here on, we set $\hbar = 1$ by taking suitable units. For simplicity, we assume that γ does not depend on the energy or momentum of an elementary excitation.

For GaAs the dispersion relation is

$$\varepsilon_{\mathbf{p}} = \sqrt{\frac{v^2 k^2}{M} + \left(\frac{k^2}{2M}\right)^2}, \quad (1)$$

where $v = \sqrt{\beta/M}$ is the speed of sound and M is the exciton mass.

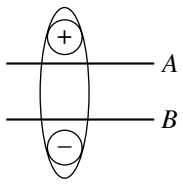

Fig. 1. GaAs/AlGaAs double quantum well, with layers *A* and *B* separated by a tunnel-thin barrier.

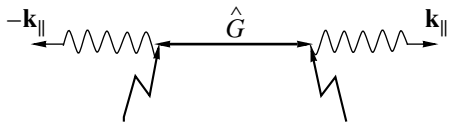

Fig. 2. Two-exciton recombination in GaAs.

In the high-density case the anomalous Green function is

$$F(\omega, \mathbf{p}) = -i\int dt e^{i\omega t}\langle T e_{\mathbf{p}}(t) h_{-\mathbf{p}}(0)\rangle$$
$$= \Delta\prod_{\pm}\left[\omega - \frac{1}{2}[\varepsilon_1(\mathbf{p}) + \varepsilon_2(\mathbf{p})] \pm (\xi(\mathbf{p}) + i\eta)\right]^{-1}, \quad (2)$$

where $e_{\mathbf{p}}$ and $h_{\mathbf{p}}$ are the respective annihilation operators for an electron and a hole and $\xi(\mathbf{p}) \equiv \xi_{\mathbf{p}}$ is the dispersion relation of an elementary excitation in the e–h system. As with the BCS theory, we take

$$\xi_{\mathbf{p}} = \sqrt{\Delta^2 + (\varepsilon_1(\mathbf{p}) - \varepsilon_2(\mathbf{p}))^2},$$

where Δ is the gap and $\varepsilon_1(\mathbf{p})$ and $\varepsilon_2(\mathbf{p})$ are the respective electron dispersion relations for the electron band of the layer *A* and the hole band of the layer *B*, assuming zero interaction. From here on the electron and the hole band will be referred to as band 1 and band 2, respectively. The electron dispersion relations are taken as

$$\varepsilon_\alpha(\mathbf{p}) = (-1)^\alpha\left(-\frac{p^2}{2m_\alpha} + \frac{p_F^2}{2m_\alpha}\right),$$

where m_1 and m_2 are the effective electron masses for the bands indicated and p_F is the Fermi momentum; this is determined by the density of excited electrons.

In quasi-2D systems, only an in-plane momentum component is conserved, so that the normal component of photon momentum is subject to energy conservation only.

Let us consider the emission of two photons with respective in-plane momentum components $\mathbf{k}_1$ and $\mathbf{k}_2$ such that $\mathbf{k}_1 = -\mathbf{k}_2 = \mathbf{k}_\parallel$. The indices *a* and *b* will mean that the corresponding quantity refers to the low- or the high-density case, respectively.

The emission rate is given by

$$\mathcal{W}_{(a,b),\mathbf{k}_\parallel} = \int 2\pi\delta(\omega_1 + \omega_2)\left|\mathcal{M}_{(a,b)}(\omega_1, \mathbf{k}_\parallel)\right|^2 \times \prod_{i=1,2}\frac{dk_{\perp,i}}{2\pi}, \quad (3)$$

where $\mathcal{M}_{(a,b)}(\omega, \mathbf{k}_\parallel)$ is the transition matrix element, $k_{\perp,i}$ is the vertical coordinate of the *i*th photon momentum, and ω_i is the *i*th photon energy as reckoned from the chemical potential μ of the e–h system:

$$\omega_i(\mathbf{k}_\parallel, k_{\perp,i}) = c\sqrt{\mathbf{k}_\parallel^2 + k_{\perp,i}^2} - \mu.$$

In the two extreme cases, $\mu = \mu_a = E_g - E_b$ and $\mu = \mu_b = E_g + \mu_e + \mu_h$, where E_g is the semiconductor energy gap; E_b is the exciton binding energy; μ_e and μ_h are the chemical potentials of charge carriers relative to the conduction and the valence band, respectively. Since E_g exceeds E_b, μ_e, and μ_h by several orders of magnitude, we have $\mu_a \approx \mu_b \approx E_g$.

3. STIMULATED LIGHT SCATTERING BY A 2D COHERENT PHASE: THE LOW-DENSITY CASE

Let us investigate two-exciton coherent recombination for a GaAs/AlGaAs DQW in the low-density case. The Hamiltonian of exciton–photon interaction is

$$H_{\text{int}}^{(a)} = \sum_{\mathbf{k}}\frac{g_{\mathbf{k}}}{\sqrt{L}}a_{\mathbf{k}}^\dagger c_{\mathbf{k}} + \text{H.c.}, \quad (4)$$

where $g_{\mathbf{k}}$ is the coupling constant and *L* is the thickness of a constituent quantum well. Since *L* will not appear in the final formula, we conveniently set $L = 1$.

Consider a system of excitons that interact with each other, $N_{\text{exc}}\rangle$ being the total number of excitons before the recombination. This process (Fig. 2) is the transition from the state

$$|\Phi_0\rangle = |N_{\text{exc}}\rangle|\mathbf{0}\rangle$$

to the state

$$|\Phi_f\rangle = |(N-2)_{\text{exc}}\rangle|1_{\mathbf{k}_1}, 1_{\mathbf{k}_2}\rangle,$$

where $|N_{\text{exc}}\rangle$ is the exciton ground state; $|\mathbf{0}\rangle$ is the photon ground state; and $|1_{\mathbf{k}_1}, 1_{\mathbf{k}_2}\rangle = c_{\mathbf{k}_1}^\dagger c_{\mathbf{k}_2}^\dagger|\mathbf{0}\rangle$, with $c_{\mathbf{k}_1}^\dagger$ denoting the creation operator for a photon of momentum $(\mathbf{k}_1, k_{\perp,1})$. To the lowest order the matrix element of the transition is

$$S_{0\to f}^{(a)} = \frac{1}{2}\int\langle\Phi_f|H_{\text{int}}^{(a)}(t_1)H_{\text{int}}^{(a)}(t_2)|\Phi_0\rangle dt_1 dt_2$$
$$= \mathcal{M}_a(\omega, \mathbf{k}_\parallel)\delta(\omega_1 + \omega_2) = g^2\hat{G}(\omega, \mathbf{k}_\parallel)\delta(\omega_1 + \omega_2). \quad (5)$$

Notice that the matrix element is nonzero only if $\mathbf{k}_1 = -\mathbf{k}_2 = \mathbf{k}_\parallel$ for the in-plane photon momentum compo-

nents. If $\gamma \ll \varepsilon_{\mathbf{k}_\parallel}$, we can use the approximation

$$|\mathcal{M}_a(\omega, \mathbf{k}_\parallel)|^2 \approx \frac{\pi g^4 \beta^2}{\gamma(4\varepsilon_{\mathbf{k}_\parallel}^2 + \gamma^2)} \sum_{\pm} \delta(\omega \pm \varepsilon_{\mathbf{k}_\parallel}). \quad (6)$$

As a result, Eq. (3) becomes

$$\mathcal{W}_{a, \mathbf{k}_\parallel} = \frac{\beta^2}{\gamma(4\varepsilon_{\mathbf{k}_\parallel}^2 + \gamma^2)} \times \prod_{\pm} \left[g^2 \int 2\pi \delta(\omega \pm \varepsilon_{\mathbf{k}_\parallel}) \frac{dk_\perp}{2\pi} \right], \quad (7)$$

where ω must be regarded as a function of $\mathbf{k}_\parallel$ and $k_\perp$. The quantity in brackets is the inverse lifetime of an exciton with the momentum $\mathbf{k}$. Approximation (6) implies that the two photons differ in energy by an amount on the order of $\varepsilon_{\mathbf{k}_\parallel}$, which is small compared with μ. It follows that the magnitudes of the photon momenta are very close to $k_0 \approx E_g/c$. Thus, the photon momenta make almost the same angles with the plane of the DQW.

The emission rate for one photon with the in-plane momentum component $\mathbf{k}$ is given by $W_{a, \mathbf{k}} = n_{\mathbf{k}} \tau_{\mathbf{k}}^{-1}$, where $n_{\mathbf{k}} = \beta/2\varepsilon_{\mathbf{k}}$ is the total number of excitons with the momentum $\mathbf{k}$.

The above formulas enable us to quantitatively compare the respective rates of the one- and the two-photon emission:

$$\alpha_a = \frac{\mathcal{W}_{a, \mathbf{k}}}{W_{a, \mathbf{k}}} = \frac{2\varepsilon_{\mathbf{k}} \beta}{\gamma(4\varepsilon_{\mathbf{k}}^2 + \gamma^2)\tau_{\mathbf{k}}}.$$

Let us take $\beta = 0.8 \times 10^{13}\ \mathrm{s}^{-1} \equiv 0.5 \times 10^{-3}$ eV and $\tau_{\mathbf{k}} \approx 10^{-8}$ s. Also, the effective exciton mass in GaAs is 0.22 times the free-electron mass. The speed of sound v in BE-condensed excitons is then close to 2×10^5 cm/s. Assuming that $\gamma \sim v/l \approx 10^8$–$10^9\ \mathrm{s}^{-1}$, where l is the exciton mean free path, we find that γ is between 10^8 and $10^9\ \mathrm{s}^{-1}$. Thus, α_a varies from 0 to 10, depending on $\mathbf{k}$.

4. STIMULATED LIGHT SCATTERING BY A 2D COHERENT PHASE: THE HIGH-DENSITY CASE

We now proceed to the coherent recombination of two e–h pairs for the same heterostructure in the high-density case, $na_B^2 \gg 1$. Consider a system of coherent e–h pairs with the volume V, which initially consists of N such pairs. The Hamiltonian of electron–hole–photon interaction is

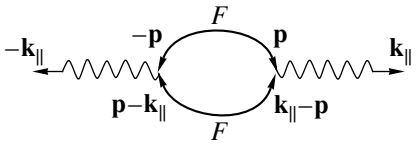

Fig. 3. Coherent recombination of two e–h pairs in GaAs.

$$H_{\mathrm{int}}^{(b)} = \sum_{\mathbf{q}, \mathbf{k}} \frac{f_{\mathbf{q}, \mathbf{k}}}{\sqrt{V}} e_{\mathbf{q}}^\dagger h_{\mathbf{k}-\mathbf{q}}^\dagger c_{\mathbf{k}} + \mathrm{H.c.} \quad (8)$$

Since $f_{\mathbf{q}, \mathbf{k}}$ is almost independent of $\mathbf{q}$, we shall neglect the dependence and simply write $f_{\mathbf{k}}$. Further, we set $V = 1$.

The recombination is the transition from the state

$$|\Phi_0\rangle = |\psi\rangle_{(N)}|0\rangle$$

to the state

$$|\Phi_{\mathrm{f}}\rangle = |\psi\rangle_{(N-2)}|1_{\mathbf{k}_1}\rangle|1_{\mathbf{k}_2}\rangle,$$

where $|\psi\rangle_{(N)}$ is the ground-state wave function, an analog of the one in the BCS theory [6]. This process involves emission of two photons (Fig. 3). The corresponding matrix element is written as

$$S_{0\to f}^{(b)} = \frac{1}{2}\int \langle \Phi_{\mathrm{f}}| H_{\mathrm{int}}^{(b)}(t_1) H_{\mathrm{int}}^{(b)}(t_2)|\Phi_0\rangle dt_1 dt_2 = \mathcal{M}_b(\omega_1, \mathbf{k}_1)\delta(\omega_1 + \omega_2), \quad (9)$$

$$\mathcal{M}_b(\omega, \mathbf{k}) = f_{\mathbf{k}}^2 \int F(\omega_{\mathbf{p}}, \mathbf{p})\, F(\omega - \omega_{\mathbf{p}}, \mathbf{k} - \mathbf{p}) \frac{d\mathbf{p}}{(2\pi)^3} d\omega_{\mathbf{p}}. \quad (10)$$

Integration with respect to $\omega_{\mathbf{p}}$ yields

$$\mathcal{M}_b(\omega, \mathbf{k}) = i\pi\Delta^2 f_{\mathbf{k}}^2 \int \frac{[\alpha_{\mathbf{k}, \mathbf{p}}(\xi_{\mathbf{p}} + \xi_{\mathbf{k}-\mathbf{p}}) + (\xi_{\mathbf{p}} + \xi_{\mathbf{k}-\mathbf{p}} + 2i\eta)(\xi_{\mathbf{p}} - \xi_{\mathbf{k}-\mathbf{p}} - \omega)]}{(\xi_{\mathbf{p}} + i\eta)(\xi_{\mathbf{k}-\mathbf{p}} + i\eta)[(\omega - \alpha_{\mathbf{k}, \mathbf{p}})^2 - (\xi_{\mathbf{p}} + \xi_{\mathbf{k}-\mathbf{p}} + 2i\eta)^2](2\pi)^3} d\mathbf{p}, \quad (11)$$

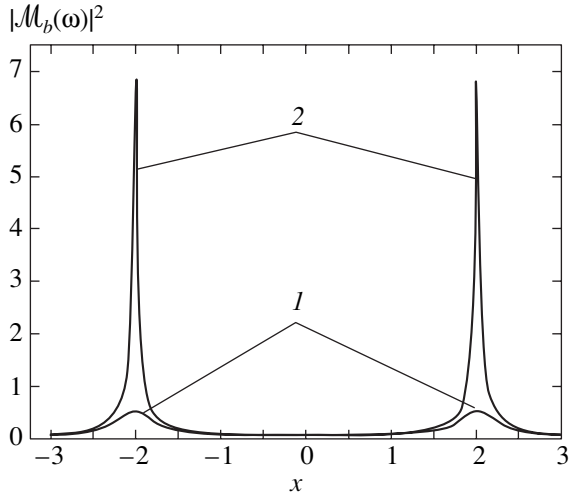

Fig. 4. Graphs of $|\mathcal{M}_b(\omega)|^2$ vs. $x \equiv \omega/\Delta$ for η/Δ equal to (*1*) 0.1 and (*2*) 0.01.

where

$$\alpha_{\mathbf{k},\mathbf{p}} = \frac{\varepsilon_1(\mathbf{p}) + \varepsilon_2(\mathbf{p})}{2} - \frac{\varepsilon_1(\mathbf{k}-\mathbf{p}) + \varepsilon_2(\mathbf{k}-\mathbf{p})}{2}.$$

As Eq. (11) includes a cumbersome integrand, we focus on the case $\mathbf{k} = 0$, in which the emitted photon momenta are perpendicular to the plane of the DQW. In qualitative terms the result should also work for reasonable $|\mathbf{k}|$.

With $\mathbf{k} = 0$, Eq. (11) becomes

$$\mathcal{M}_b(\omega, 0) \equiv \mathcal{M}_b(\omega) = i\pi\Delta^2 f_0^2 \int \frac{1}{(\xi_\mathbf{p} + i\eta)(\omega'^2 - 4(\xi_\mathbf{p} + i\eta)^2)} \frac{d\mathbf{p}}{(2\pi)^3}.$$

The integrand being actually a function of $|\mathbf{p}| \equiv p$ only, we change to the variable $\varepsilon = p^2/2M_1$, where $M_1 = 2m_1m_2/(m_1 + m_2)$. The resultant single integral can be calculated in analytical form. For $\Delta \ll p_F^2/2M_1$, we thus obtain

$$\mathcal{M}_b(\omega) = \frac{M_1}{2\pi}\Delta^2 f_0^2 \left(-\frac{A^- \ln(\alpha^-\beta^+/\alpha^+\beta^-) + A^+ D}{2\omega r} + i\frac{A^+ \ln(\alpha^-\beta^+/\alpha^+\beta^-) - A^- D}{2\omega r} \right),$$

where

$$r = \sqrt{(\omega'^2 - 4\Delta^2 - 4\eta^2)^2 + 16\omega'^2\eta^2},$$

$$A^\pm = \frac{1}{\sqrt{2}}\sqrt{r \pm (\omega'^2 - 4\Delta^2 - 4\eta^2)},$$

$$\alpha^\pm = \sqrt{\left(\Delta \pm \frac{\omega}{2} - \frac{A^+}{2}\right)^2 + \left(\eta + \frac{A^-}{2}\right)^2},$$

$$\beta^\pm = \sqrt{\left(\Delta \pm \frac{\omega}{2} + \frac{A^+}{2}\right)^2 + \left(\eta - \frac{A^-}{2}\right)^2},$$

$$D = \arcsin\left(\frac{\Delta + \omega - A^+/2}{\alpha^+}\right) - \arcsin\left(\frac{\Delta - \omega - A^+/2}{\alpha^-}\right) + \arccos\left(\frac{\Delta - \omega + A^+/2}{\beta^-}\right) - \arccos\left(\frac{\Delta + \omega + A^+/2}{\beta^+}\right).$$

Let us examine the function $|\mathcal{M}_b(\omega)|^2$. Figure 4 shows its graphs for different η/Δ. Notice that $|\mathcal{M}_b(\omega)|^2$ is nonzero only near the points $\omega = \pm 2\Delta$. Owing to the fact that these regions disappear as $\eta/\Delta \longrightarrow 0$, we take the approximation

$$|\mathcal{M}_b(\omega)|^2 = \left(\frac{M_1}{2\pi}\right)^2 \Delta f_0^4 S_\eta [\delta(\omega - 2\Delta) + \delta(\omega + 2\Delta)], \quad (12)$$

where $S_\eta \approx 0.3\ln(\eta/\Delta)$. Approximation (12) enables us to estimate the rate of two-photon emission in the normal direction. As with Eq. (7), we calculate integral (3) to obtain

$$\mathcal{W}_b = \left(\frac{M_1}{2\pi}\right)^2 \frac{\Delta S_\eta}{n^2} W_b^2, \quad (13)$$

where W_b is the rate of one-photon emission. Let us set $n \sim 10^{12}$ cm^{-2}, $\Delta \sim 0.001$ eV, $\eta/\Delta \sim 0.001$, $W_b \sim 10^8$ s^{-1}, and $M_1 = 0.1m_e$. Then

$$\alpha_b = \mathcal{W}_b/W_b \sim 10^{-6}–10^{-5}. \quad (14)$$

If the recombination is induced by resonant laser radiation with N_0 photons per mode, the rate of emission in the opposite direction to the incident beam is given by

$$W_{opp} = (N_0 + 1)\mathcal{W}_b. \quad (15)$$

If $N_0 \sim 10^5$–10^6, the emission rate W_{opp} will be comparable with the rate of one-photon emission, so that the former effect is, in principle, detectable.

Also note that in both extreme cases the 2D nature of the e–h system implies that a laser mode will induce the emission of photons with the wave vectors $(\mathbf{k}_\parallel, \pm k_{\perp,1})$ and $(-\mathbf{k}_\parallel, \pm k_{\perp,2})$, where $k_{\perp,1} \approx k_{\perp,2}$. It follows that aside from backscattering the laser-induced recombination can produce photons with a reversed in-plane momentum component that will cross the DQW (see beam *4* of Fig. 5). This phenomenon might be seen as anomalous transmission.

5. OPTICAL WAVE-VECTOR REVERSAL BY A 3D COHERENT PHASE IN THE HIGH-DENSITY CASE

Let us consider stimulated light scattering by bulk GaAs in the high-density case ($na_B^3 \gg 1$). The Hamiltonian of electron–hole–photon interaction and the transition matrix element are again given by Eqs. (8) and (9), respectively; however, **k** and **p** are now 3D vectors in (10). Note that both the magnitudes and the directions of the emitted photon momenta are subject to momentum and energy conservation, so that the two photon momenta must be antiparallel and equal in magnitude.

Let us calculate $\mathcal{M}_b(\omega, \mathbf{k}) \equiv \mathcal{M}_b(\mathbf{k})$. As in the 2D case, integration with respect to $\omega_\mathbf{p}$ yields Eq. (11). The z axis being aligned with **k**, we recast Eq. (11) to

$$\mathcal{M}_b(\mathbf{k}) = (2M_1)^{3/2}\sqrt{\Delta}\, f_\mathbf{k}^2 I(p_i), \quad i = 1, \ldots, 6. \quad (16)$$

Here, $I(p_i)$ is an appropriate integral with the dimensionless parameters

$$p_1 = \frac{k^2}{2M_1\Delta}, \quad p_2 = \frac{k^2}{2M_2\Delta}, \quad p_3 = \frac{M_1 k^2}{2M_2^2\Delta},$$

$$p_4 = \frac{p_F^2}{2M_1\Delta}, \quad p_5 = \frac{\omega}{\Delta} \approx \frac{E_g}{\Delta}, \quad p_6 = \frac{\eta}{\Delta},$$

where $M_2 = 2m_1m_2/(m_2 - m_1)$ with $m_1 \neq m_2$. It can be calculated numerically.

With bulk GaAs, we have $E_g \approx 1.5$ eV, $n = 10^{21}$ cm^{-3}, $p_F = 3 \times 10^7$ cm^{-1}, $M_2 \approx 0.2m_e$, $\Delta = 0.2p_F^2/2M_1$, and $|\mathbf{k}| \approx E_g/c$. Calculating p_i, we obtain $|I(p_i)|^2 \approx 250$. The rate of two-photon emission is thus estimated as ${}^c\mathcal{W}_{b,\mathbf{k}} \approx 5 \times 10^{21}$ s^{-1}.

6. MULTIEXCITON COHERENT RECOMBINATION IN A BOSE–EINSTEIN CONDENSATE: THE CASE OF Cu_2O

This section is concerned with the coherent recombination of three or four BE-condensed excitons in bulk Cu_2O. The processes will be investigated by generalizing the above results concerning two-exciton recombination. The rate of direct e–h recombination being very low in Cu_2O, an exciton decays mostly by emission of a photon and an optical phonon. The Hamiltonian of exciton–photon–phonon interaction is

$$V^{(1)} = \sum_{\mathbf{p}-\mathbf{k}-\mathbf{q}=0} \frac{\lambda_{\mathbf{k},\mathbf{q}}}{\sqrt{V}} c_{\mathbf{p}+\mathbf{q}} Q_\mathbf{p} b_\mathbf{q} + \text{H.c.}, \quad (17)$$

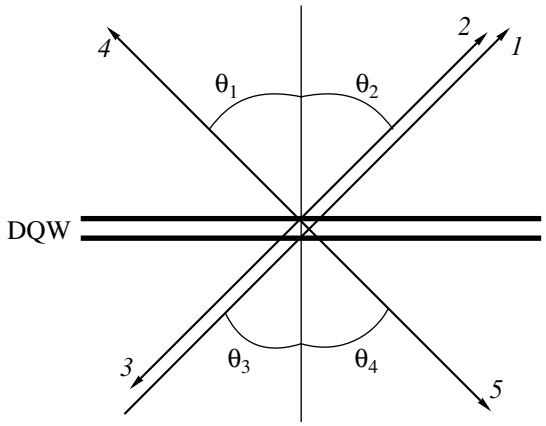

Fig. 5. Stimulated light scattering from a 2D system involving two-photon emission: (*1*) incident beam; (*2*) ordinary transmitted beam, ($\mathbf{k}_\parallel$, $k_\perp$); (*3*) reversed beam, $-(\mathbf{k}_\parallel, k_\perp)$; (*4*) anomalous transmitted beam, $(-\mathbf{k}_\parallel, k_\perp)$; and (*5*) ordinary reflected beam, $(\mathbf{k}_\parallel, -k_\perp)$. The angles are $\theta_1 = \theta_2 = \theta_3 = \theta_4$.

where $b_\mathbf{q}$ is the annihilation operator for a phonon of momentum **q** and λ is the effective coupling constant.

If the BE condensate experiences N-exciton recombination creating N photons, it makes the transition

$$\langle i| \xrightarrow{V^{(1)}} c_{\mathbf{k}_1} \ldots c_{\mathbf{k}_N} a_0^N \langle i| = \langle f|.$$

As with two-exciton recombination, the processes could be detected by Brown–Twiss measurements using a coincidence circuit with N photon counters. However, we here consider an alternative approach based on multibeam laser-induced recombination.

Since BE-condensed excitons have zero momentum, so must be the total momentum of the N photons created:

$$\sum_{i=1}^{N} \mathbf{k}_i = 0. \quad (18)$$

Recall that we reckon photon energy from the exciton chemical potential μ, so that the photon dispersion relation is

$$\omega_\mathbf{k} = ck - \mu \equiv c(k - k_0).$$

BE-condensed excitons are thus assigned zero energy ($\mu = 0$). Due to energy conservation the photon energies ω_i must obey the constraint

$$\sum_{i=1}^{N} \omega_i = 0, \quad \omega_i \equiv \omega_{\mathbf{k}_i}. \quad (19)$$

If N-exciton recombination is induced by $N - 1$ laser beams with respective wave vectors $\mathbf{k}_i$ ($i = 1, \ldots, N-1$), it should produce a plane outgoing wave whose wave

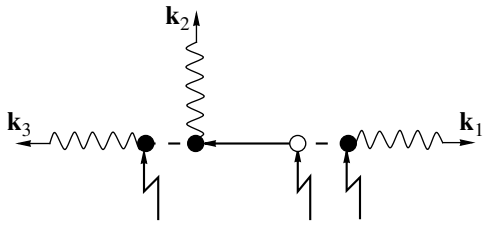

Fig. 6. Three-exciton recombination in a BE condensate.

vector $\mathbf{k}_N$ and photon energy ω_N are subject to constraints (18) and (19), respectively:

$$\mathbf{k}_N = -\sum_{i=1}^{N-1} \mathbf{k}_i, \tag{20}$$

$$\omega_N = -\sum_{i=1}^{N-1} \omega_i. \tag{21}$$

As compared with concurrent spontaneous emission, the intensity of the stimulated emission is higher by a factor of

$$\prod_{i=1}^{N-1} (N_i + 1), \tag{22}$$

where N_i is the mean number of photons per mode for the ith incident beam.

Let us determine conditions for an outgoing wave to be detectable. As with two-exciton recombination, one-exciton stimulated recombination will not contribute to the photoluminescence in the direction of the outgoing wave if the incident wave vectors are oriented appropriately. Accordingly, the background emission in the outgoing-wave direction is due to one-exciton spontaneous recombination only. The rate W_N of N-exciton stimulated recombination should therefore be compared with the rate W_1 of one-exciton spontaneous recombination. Assume that each incident beam has 10^3 photons per mode. The detection is possible if

$$10^{3(N-1)} W_N > W_1, \tag{23}$$

where

$$W_1 = \rho_{\text{cond}}/\tau, \tag{24}$$

with $\tau = \pi/\lambda^2 k_0^2 \approx 10^{-5}$ s being the exciton lifetime in Cu_2O; hence $\lambda \approx 2.5 \times 10^2$ s^{-1} cm$^{3/2}$.

Three-exciton recombination is also connected with phonon–exciton interaction (see below). The corresponding Hamiltonian is

$$V^{(2)} = \sum_{\mathbf{p}-\mathbf{k}-\mathbf{q}=0} \frac{g_{\mathbf{k},\mathbf{q}}}{\sqrt{V}} a_{\mathbf{p}+\mathbf{q}} a_{\mathbf{p}} b_{\mathbf{q}} + \text{H.c.} \tag{25}$$

Assume that the coupling constant is $g = 10^2\lambda$. This implies that the rate of phonon exchange between excitons is higher than that of photon emission by four orders of magnitude (so that the characteristic time is about 10^{-10} s).

Using the Fermi golden rule, we find that photons are created at the rate

$$W_N = N\int 2\pi\delta\left(\sum_{i=1}^{N} \omega_i\right) \times |\mathcal{M}_N(\mathbf{k}_1 \ldots \mathbf{k}_N)|^2 V^N \prod_{i=1}^{N-1} \frac{d\mathbf{k}_i}{(2\pi)^3}, \tag{26}$$

where $\mathcal{M}_N$ is the matrix element of the process and the prefactor N is due to N photons being created in every elementary process. Note that $\mathcal{M}_N$ can be expanded in terms of N factors of the form $V^{-1/2}$, so that W_N is actually independent of V. This allows us to set $V = 1$.

6.1. Three-Exciton Recombination

The matrix element $\mathcal{M}_3(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3)$ of three-exciton recombination is the sum of matrix elements obtained from the one from Fig. 6 by all possible permutations of the photon-vertex arguments:

$$\mathcal{M}_3(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3) = g\lambda^3 \rho_{\text{cond}}^{3/2} \sum_{i \neq j} \mathcal{G}(\omega_i)\mathcal{G}(\omega_j) G_{-\mathbf{k}_j}(\omega_j), \quad i, j = 1, 2, 3, \tag{27}$$

where $G(\omega, \mathbf{k})$ is the exciton Green function,

$$G(\omega, \mathbf{k}) = \frac{\omega + \sqrt{\varepsilon_{\mathbf{k}}^2 + \beta^2}}{[\omega - (\varepsilon_{\mathbf{k}} - i\gamma)][\omega + (\varepsilon_{\mathbf{k}} - i\gamma)]}, \tag{28}$$

and $\mathcal{G}(\omega)$ is the phonon Green function,

$$\mathcal{G}(\omega) = \frac{2\Omega}{[\omega - (\Omega - i\delta)][\omega + (\Omega - i\delta)]}. \tag{29}$$

The latter is assumed to be independent of $\mathbf{k}$ because the magnitude of any wave vector involved is of the order of k_0 (see Section 3), which is small in terms of

the Brillouin zone; with such wave vectors, the optical-phonon energy can be regarded as a constant.

The rate of the process is mainly determined by the photon-energy regions in which virtual particles are closest to their mass surfaces, at which resonance occurs. The photon energies (with the dispersion relation selected) are of the order of the elementary-excitation energy $\varepsilon_{\mathbf{k}}$ in the BE condensate or the optical-phonon energy Ω (see (31)). The orders of magnitude of $\varepsilon_{\mathbf{k}}$ and Ω are much less than that of the exciton chemical potential $\mu \approx E_g$ as reckoned from the valence-band maximum. We therefore take k_0 as an approximation to the magnitudes of photon wave vectors (see Appendix for details). Accordingly, $\mathcal{M}_3(\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3)$ in (27) becomes a function of photon energies only. Assuming that $\delta \ll \Omega$ and $\gamma \ll \varepsilon_{k_0}$, where ε_{k_0} is the elementary-excitation energy, we apply the pole expansion to $|\mathcal{M}_3|^2$ and obtain the sum of six equal resonant terms:

$$|\mathcal{M}_3(\omega_1, \omega_2, \omega_3)|^2 \approx 6g^2\lambda^6\rho_{\text{cond}}^3 \left|\mathcal{G}(\omega_2)\mathcal{G}(\omega_1)G_{k_0}(\omega_1)\right|^2.$$

Owing to the above approximation the rate W_3 of three-exciton stimulated recombination is expressed as

$$W_3 = 18\frac{g^2\lambda^6\rho_{\text{cond}}^3 k_0^3}{\pi c^3}\int \left|\mathcal{G}(\omega_1)G_{k_0}(\omega_1)\right|^2 \frac{\delta\omega_1}{2\pi} \times \int \left|\mathcal{G}(\omega_2)\right|^2 \frac{d\omega_2}{2\pi}.$$

Since $\Omega \approx 10^{-2}$ eV $\gg \varepsilon_{k_0} \approx 10^{-4}$ eV, we arrive at

$$W_3 = 18\frac{g^2\lambda^6\rho_{\text{cond}}^3 k_0^3}{\pi c^3}\frac{1}{\Omega^2\delta}\left[\frac{1}{\delta} + \frac{1}{\gamma}\left(1 + \frac{\beta^2}{2\varepsilon_{k_0}^2}\right)\right] = 18\frac{g^2\lambda^4\rho_{\text{cond}}^2 k_0}{c^2\Omega^2\delta}\left[\frac{1}{\delta} + \frac{1}{\gamma}\left(1 + \frac{\beta^2}{2\varepsilon_{k_0}^2}\right)\right]W_1. \quad (30)$$

To estimate W_3, we take the exciton mass as $2.7m_e$ and the other parameters as $\beta \approx 0.5$ meV (0.5×10^{12} s^{-1}), $\rho_{\text{cond}} \approx 10^{19}$ cm^{-3}, $\gamma \approx 0.1\beta$, and $\delta \approx 10^9$ s^{-1}. The relative permittivity and energy gap of Cu_2O are $\varepsilon \approx 9$ and $E_g \approx 2$ eV, so that $c = c_0/\sqrt{\varepsilon} \approx 10^{10}$ cm/s and $k_0 = 3 \times 10^5$ cm^{-1}. Thus,

$$W_3 \approx 10^{-2}W_1;$$

that is, one out of a hundred excitons decays by three-exciton stimulated recombination. Noting that the rate is high enough to meet requirement (23), we conclude that the process is in principle detectable.

Three-exciton stimulated recombination could be produced and detected by illuminating the BE condensate with two laser beams making an angle of $2\pi/3$ with one another (see Appendix); the wave vector of the outgoing beam will be coplanar with the incoming wave vectors and will make the stated angle with each of them.

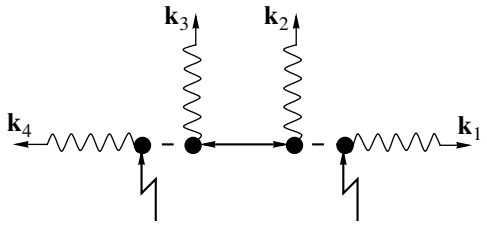

Fig. 7. Four-exciton recombination in a BE condensate.

As already noted, the recombination rate is mainly determined by the photon-energy regions in which virtual particles are closest to their mass surfaces. Specifically, the corresponding energies of the three emitted photons are close to any of the following sets

$$\begin{aligned}&(\mu + \Omega, \mu - \Omega, \mu),\\ &(\mu + \Omega, \mu \pm \varepsilon_{k_0}, \mu - \Omega \mp \varepsilon_{k_0}),\\ &(\mu - \Omega, \mu \mp \varepsilon_{k_0}, \mu + \Omega \pm \varepsilon_{k_0}),\end{aligned} \quad (31)$$

where μ is the exciton chemical potential. The first set corresponds to the first term in brackets appearing in (30), so that the resonant condition is fulfilled by a virtual optical phonon. The other sets correspond to the second term, implying that an elementary excitation of the BE condensate will be at resonance.

Thus, the recombination rate can be increased by setting the respective photon energies of the laser beams to any two members of any energy set in (31), the outgoing photon energy being equal to the remaining member.

6.2. Four-Exciton Recombination

To the lowest order with respect to exciton–phonon–photon interaction, the matrix element of four-exciton stimulated recombination is given by

$$\mathcal{M}_4(\mathbf{k}_1 \ldots \mathbf{k}_4) = \rho_{\text{cond}}\lambda^4 \times \sum_{\substack{l \neq m\\ m \neq n\\ n \neq l}} F_{-(\mathbf{k}_1 + \mathbf{k}_m)}(-(\omega_l + \omega_m))\mathcal{G}(-\omega_l)\mathcal{G}(-\omega_n),$$

$$l, m, n = 1, \ldots, 4.$$

It corresponds to 12 diagrams constructed from that of Fig. 7 by photon-vertex permutation. The two exciton lines that are not shown in Fig. 7 are implicitly included

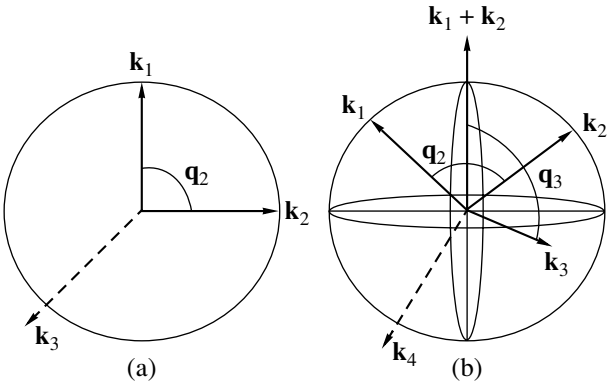

Fig. 8. The magnitude of an outgoing wave vector $\mathbf{k}_N$ relative to the mutual orientation of the incoming ones with a fixed magnitude, for (a) $N = 3$ and (b) $N = 4$.

in the anomalous Green function of BE-condensed excitons.

Applying the pole expansion to $|\mathcal{M}_4|^2$, we obtain the sum of 12 equal resonant terms:

$$|\mathcal{M}_4|^2 = 12\rho_{\text{cond}}^2\lambda^8\left|F_{|\mathbf{k}_1+\mathbf{k}_2|}(\omega_1+\omega_2)\mathcal{G}(\omega_1)\mathcal{G}(\omega_3)\right|^2. \quad (32)$$

As with three-exciton recombination, we take the magnitudes of photon wave vectors to be k_0 (see Appendix for details), so that the recombination rate W_4 is expressed as

$$W_4 = 48\rho_{\text{cond}}^2\lambda^8\frac{k_0^4}{2\pi^2c^4}\int\frac{du}{2\pi}\int_0^{2k_0}dk|F_k(u)|^2 \times\int\frac{d\omega_1}{2\pi}\left|\mathcal{G}(\omega_1)\right|^2\int\frac{d\omega_3}{2\pi}\left|\mathcal{G}(\omega_3)\right|^2,$$

where $u = \omega_1 + \omega_2$ and $k = |\mathbf{k}_1 + \mathbf{k}_2|$. Assuming that γ is independent of $\mathbf{k}$ and is much less than ε_{k_0}, we see that

$$\int\frac{du}{2\pi}\int_0^{2k_0}dk|F_k(u)|^2 = \frac{\pi\beta^2}{8\gamma^2 v}.$$

With $v \approx 0.5\times10^6$ cm/s, we arrive at

$$W_4 = \frac{3\rho_{\text{cond}}^2\lambda^8\beta^2k_0^4}{\pi c_v c^4\delta^2\gamma^2} = \frac{3\rho_{\text{cond}}\lambda^6\beta^2k_0^2}{c_v c^2\delta^2\gamma^2}W_1. \quad (33)$$

Calculation yields

$$W_4 = 5\times10^{-3}W_1;$$

that is, one out of two hundred excitons decays by four-exciton stimulated recombination. Noting that the rate is high enough to meet requirement (23), we conclude that the process is, in principle, detectable.

The wave vectors of the photons created are oriented as in Fig. 8b of Appendix. Due to condition (18) the total momentum of any two photons is in the opposite direction to that of the others. Also, the angle between any two wave vectors is the same as that between the other two. However, the wave vectors are not generally coplanar.

Four-exciton stimulated recombination could be produced and detected by illuminating the BE condensate with three laser beams. The wave vectors of the incoming beams should be oriented as described above, the outgoing wave vector being given by Eq. (20).

As with three-exciton recombination, the rate of four-exciton recombination is highest if the photon energies are equal to the members of the set

$$(\mu+\Omega, \mu-\Omega\pm\varepsilon_{k_0}, \mu-\Omega, \mu+\Omega\mp\varepsilon_{k_0}).$$

Accordingly, the recombination rate can be increased by setting the respective photon energies of the laser beams to any three members of the set, the outgoing photon energy being equal to the remaining member.

7. CONCLUSIONS

We investigated stimulated light scattering, including wave-vector reversal, from an e–h coherent phase in a GaAs/AlGaAs DQW and bulk GaAs in the low- and the high-density case, the scattering mechanism being the coherent recombination of two excitons or two e–h pairs, respectively. The estimated rates of two-photon emission indicate that the scattering is detectable in both of the extreme cases. If the incident laser radiation provides 10^2–10^5 photons per mode, the rate of two-photon emission will be comparable to that of one-photon emission. We also considered multiexciton coherent recombination, for bulk Cu_2O. Similarly, the coherent recombination of multiple correlated e–h pairs is possible in the high-density case.

Four-wave mixing is another candidate way of studying BE-condensed excitons [50–52]. However, it is important to note that this method creates a coherent exciton system (in the form of exciton density waves) before it is examined, whereas our approach deals with an existing one.

It should be emphasized that stimulated light scattering by an e–h system is possible only if the system contains a coherent phase; therefore, this optical phenomenon could serve as an indicator of the presence of such phase. Another potential test for a coherent phase is the detection of a pair of correlated photons with antiparallel momenta that have been produced by coherent recombination of two excitons or two e–h pairs in the coherent phase. This could be done by Brown–Twiss

measurements with two appropriately positioned detectors. A similar method should work for multiexciton coherent recombination.

ACKNOWLEDGMENTS

Yu.E.L. is grateful to Yu.M. Kagan, L.V. Keldysh, L.A. Maksimov, and V.G. Lysenko for helpful discussions of the results.

This work was supported by grants from INTAS and the Russian Foundation for Basic Research.

APPENDIX

In general, the integral in Eq. (26) is to be taken with respect to $N-1$ wave vectors in a 3D space. It is possible to reduce the total number of variables of integration if we can set $k_i = k_0$ for $i = 1, \ldots, N-1$, where k_i is the magnitude of the ith wave vector. Specifically, we introduce spherical coordinates in the 3D space and so bring the integral to the form

$$\int \ldots 2\pi\delta\left(\sum_{i=1}^{N} c(k_i - k_0)\right)\prod_{i=1}^{N-1}\frac{k_i^2 d\omega_i d\cos\theta_i d\phi_i}{c(2\pi)^3}, \quad (34)$$

the wave vectors being represented as

$$\mathbf{k}_i = k_i(\cos\theta_i;\ \sin\theta_i\cos\phi_i;\ \sin\theta_i\sin\phi_i). \quad (35)$$

With $k_i = k_0$ for each i, the product terms become

$$\frac{k_0^2 d\omega_i d\cos\theta_i d\phi_i}{c(2\pi)^3},$$

and the delta-function term is recast to

$$\frac{1}{c}\delta\left(\left|\sum_{i=1}^{N-1}\mathbf{k}_i\right| - k_0\right). \quad (36)$$

With this approximation, energy conservation dictates that the incoming wave vectors be oriented so that the outgoing wave vectors as given by Eq. (20) have the magnitude k_0.

Three-Photon Case

Without loss of generality, we introduce spherical coordinates according to Fig. 8a. Consequently,

$$|\mathbf{k}_1 + \mathbf{k}_2| = k_0\sqrt{2(1+\cos\theta_2)}$$

and expression (36) becomes

$$\frac{1}{ck_0}\delta(\sqrt{2(1+\cos\theta_2)} - 1) = \frac{1}{ck_0}\delta\left(\cos\theta_2 + \frac{1}{2}\right).$$

Integration with respect to the angles brings expression (34) to a simpler form:

$$\frac{k_0^3}{\pi c^3}\int \ldots \prod_{i=1}^{2}\frac{d\omega_i}{2\pi}. \quad (37)$$

The fact that $\cos\theta_2 = -1/2$ implies that the angle between the wave vectors is $2\pi/3$.

Four-Photon Case

In the four-photon case the matrix element depends on the photon energies and $|\mathbf{k}_1 + \mathbf{k}_2|$. To reduce the number of integrations in (26), we represent the orientation of $\mathbf{k}_3$ in terms of its angle with $\mathbf{k}_1 + \mathbf{k}_2$ (Fig. 8b), so that

$$\left|\sum_{i=1}^{N-1}\mathbf{k}_i\right| = \sqrt{|\mathbf{k}_1 + \mathbf{k}_2|^2 + k_0^2 + 2k_0|\mathbf{k}_1 + \mathbf{k}_2|\cos\theta_3}.$$

Expression (36) thus becomes

$$c^{-1}\delta(\sqrt{|\mathbf{k}_1 + \mathbf{k}_2|^2 + k_0^2 + 2k_0|\mathbf{k}_1 + \mathbf{k}_2|\cos\theta_3} - k_0) = (c|\mathbf{k}_1 + \mathbf{k}_2|)^{-1}\delta(\cos\theta_3 - X),$$

where $X = -(2k_0)^{-1}|\mathbf{k}_1 + \mathbf{k}_2| \in [-1, 0]$. Integration with respect to θ_3 transforms (34) into

$$\int \ldots \frac{k_0^2}{c^2|\mathbf{k}_1 + \mathbf{k}_2|}\frac{d\omega_3}{(2\pi)}\prod_{i=1}^{2}\frac{k_0^2 d\omega_i}{c(2\pi)^3}d\cos\theta_i d\phi_i.$$

Next, we represent the orientation of $\mathbf{k}_2$ in terms of its angle with $\mathbf{k}_1$ (Fig. 8b), so that

$$|\mathbf{k}_1 + \mathbf{k}_2| = k_0\sqrt{2(1+\cos\theta_2)}.$$

Changing from $\cos\theta_2$ to $|\mathbf{k}_1 + \mathbf{k}_2|$ as a variable of integration,

$$\frac{d\cos\theta_2}{|\mathbf{k}_1 + \mathbf{k}_2|} = \frac{d(|\mathbf{k}_1 + \mathbf{k}_2|)}{k_0^2}, \quad |\mathbf{k}_1 + \mathbf{k}_2| \in (0, 2k_0),$$

brings (34) to the form

$$\frac{k_0^4}{2\pi^2 c^4}\int \ldots d|\mathbf{k}_1 + \mathbf{k}_2|\prod_{i=1}^{3}\frac{d\omega_i}{2\pi}.$$

REFERENCES

1. L. V. Keldysh and Yu. V. Kopaev, Fiz. Tverd. Tela (Leningrad) **6**, 2791 (1964) [Sov. Phys. Solid State **6**, 2219 (1964)].
2. A. N. Kozlov and L. A. Maksimov, Zh. Éksp. Teor. Fiz. **48**, 1184 (1965) [Sov. Phys. JETP **21**, 790 (1965)].

3. S. A. Moskalenko, Fiz. Tverd. Tela (Leningrad) **4**, 276 (1962) [Sov. Phys. Solid State **4**, 199 (1962)].
4. L. V. Keldysh and A. N. Kozlov, Zh. Éksp. Teor. Fiz. **54**, 978 (1968) [Sov. Phys. JETP **27**, 521 (1968)].
5. I. M. Blatt, K. W. Boer, and W. Brandt, Phys. Rev. **126**, 1691 (1962).
6. D. Jerome, T. M. Rice, and W. Kohn, Phys. Rev. **158**, 462 (1967).
7. Yu. E. Lozovik and V. I. Yudson, Fiz. Tverd. Tela (Leningrad) **17**, 1613 (1975) [Sov. Phys. Solid State **17**, 1054 (1975)].
8. A. L. Ivanov, H. Haug, and L. V. Keldysh, Phys. Rep. **296**, 237 (1998); B. I. Halperin and T. M. Rice, Solid State Phys. **21**, 115 (1968); S. A. Moskalenko and D. W. Snoke, *Bose–Einstein Condensation of Excitons and Biexcitons and Coherent Nonlinear Optics with Excitons* (Cambridge Univ. Press, New York, 2000).
9. A. A. Dremin, V. B. Timofeev, A. V. Larionov, *et al.*, Pis'ma Zh. Éksp. Teor. Fiz. **76**, 526 (2002) [JETP Lett. **76**, 450 (2002)]; A. V. Larionov, V. B. Timofeev, P. A. Ni, *et al.*, Pis'ma Zh. Éksp. Teor. Fiz. **75**, 689 (2002) [JETP Lett. **75**, 570 (2002)]; A. V. Larionov, V. B. Timofeev, D. Hvam, *et al.*, Pis'ma Zh. Éksp. Teor. Fiz. **75**, 233 (2002) [JETP Lett. **75**, 200 (2002)].
10. L. V. Butov, A. C. Gossard, and D. S. Chemla, Nature **418**, 751 (2002); L. V. Butov, C. W. Lai, A. L. Ivanov, *et al.*, Nature **417**, 47 (2002).
11. S. A. Moskalenko, M. A. Liberman, D. W. Snoke, *et al.*, Phys. Rev. B **66**, 245316 (2002).
12. V. V. Krivolapchuk, E. S. Moskalenko, and A. L. Zhmodikov, Phys. Rev. B **64**, 045313 (2001); J.-P. Cheng, J. Kono, B. D. McCombe, *et al.*, Phys. Rev. Lett. **74**, 450 (1995); L. V. Butov, A. Zrenner, G. Abstreiter, *et al.*, Phys. Rev. Lett. **73**, 304 (1994); U. Sivan, P. M. Solomon, and H. Shtrikman, Phys. Rev. Lett. **68**, 1196 (1992); T. Fukuzawa, E. E. Mendez, and J. M. Hong, Phys. Rev. Lett. **64**, 3066 (1990); J. A. Kash, M. Zachau, E. E. Mendez, *et al.*, Phys. Rev. Lett. **66**, 2247 (1991); J. B. Stark, W. H. Knox, D. S. Chemla, *et al.*, Phys. Rev. Lett. **65**, 3033 (1990).
13. Yu. E. Lozovik and V. I. Yudson, Pis'ma Zh. Éksp. Teor. Fiz. **22**, 556 (1975) [JETP Lett. **22**, 274 (1975)]; Zh. Éksp. Teor. Fiz. **71**, 738 (1976) [Sov. Phys. JETP **44**, 389 (1976)].
14. Yu. E. Lozovik and A. V. Poushnov, Phys. Lett. A **228**, 399 (1997); I. V. Lerner and Yu. E. Lozovik, Zh. Éksp. Teor. Fiz. **80**, 1488 (1981) [Sov. Phys. JETP **53**, 763 (1981)]; A. V. Klyuchnik and Yu. E. Lozovik, J. Low Temp. Phys. **38**, 761 (1980); Yu. E. Lozovik and M. Willander, Appl. Phys. A **71**, 379 (2000).
15. Yu. E. Lozovik and O. L. Berman, Pis'ma Zh. Éksp. Teor. Fiz. **64**, 526 (1996) [JETP Lett. **64**, 573 (1996)].
16. Yu. E. Lozovik, O. L. Berman, and V. G. Tsvetus, Phys. Rev. B **59**, 5627 (1999).
17. Yu. E. Lozovik, O. L. Berman, and M. Willander, J. Phys. C **14**, 12457 (2002); A. V. Klyuchnik and Yu. E. Lozovik, Fiz. Tverd. Tela (Leningrad) **20**, 625 (1978) [Sov. Phys. Solid State **20**, 364 (1978)].
18. D. Yoshioka and A. H. MacDonald, J. Phys. Soc. Jpn. **59**, 4211 (1990).
19. X. M. Chen and J. J. Quinn, Phys. Rev. Lett. **67**, 895 (1991).
20. G. E. W. Bauer, Phys. Rev. Lett. **64**, 60 (1990); Phys. Scr. **45**, 154 (1992).
21. X. Zhu, P. B. Littlewood, M. S. Hybertsen, and T. M. Rice, Phys. Rev. Lett. **74**, 1633 (1995).
22. Y. Naveh and B. Laikhtman, Phys. Rev. Lett. **77**, 900 (1996).
23. Yu. E. Lozovik and A. M. Ruvinskiĭ, Zh. Éksp. Teor. Fiz. **112**, 1791 (1997) [JETP **85**, 979 (1997)].
24. A. Imamoglu, Phys. Rev. B **57**, R4195 (1998).
25. S. Conti, G. Vignale, and A. H. MacDonald, Phys. Rev. B **57**, R6846 (1998).
26. A. V. Korolev and M. A. Liberman, Phys. Rev. Lett. **72**, 270 (1994); S. V. Korepov and M. A. Liberman, Phys. Lett. A **233**, 121 (1997); S. A. Moskalenko and M. A. Liberman, Phys. Rev. B **65**, 064303 (2002).
27. S. I. Shevchenko, Phys. Rev. Lett. **72**, 3242 (1994).
28. R. Ferreira and G. Bastard, Rep. Prog. Phys. **60**, 345 (1997).
29. A. Parlangeli, P. C. M. Christianen, J. C. Maan, *et al.*, Phys. Rev. B **62**, 15323 (2000); A. Parlangeli, P. C. M. Christianen, A. K. Geim, *et al.*, Phys. Status Solidi **164**, 587 (1997).
30. H. C. Tso, P. Vasilopoulos, and F. M. Peeters, Phys. Rev. Lett. **70**, 2146 (1993).
31. J. S. Thakur, D. Neilson, and M. P. Das, Phys. Rev. B **57**, 1801 (1998); L. Swierkowski, J. Szymanski, and D. Neilson, Acta Phys. Pol. A **84**, 445 (1993).
32. P. Fulde and R. A. Ferrel, Phys. Rev. **135**, A550 (1964).
33. A. I. Larkin and Yu. N. Ovchinnikov, Zh. Éksp. Teor. Fiz. **47**, 1136 (1964) [Sov. Phys. JETP **20**, 762 (1964)].
34. T. M. Rice, J. C. Hensel, T. G. Phillips, and G. A. Thomas, *The Electron Hole Liquid in Doped Semiconductors: Theoretical Aspects, Experimental Aspects* (Academic, New York, 1977; Mir, Moscow, 1980).
35. L. V. Keldysh and A. P. Silin, Kratk. Soobshch. Fiz. **8**, 33 (1975).
36. V. A. Gergel', R. F. Kazarinov, and R. A. Suris, Zh. Éksp. Teor. Fiz. **53**, 544 (1967) [Sov. Phys. JETP **26**, 354 (1967)].
37. B. Laikhtman, Europhys. Lett. **43**, 53 (1998).
38. H. Shi, G. Verechaka, and A. Griffin, Phys. Rev. B **50**, 1119 (1994).
39. Yu. E. Lozovik and A. V. Poushnov, Phys. Rev. B **58**, 6608 (1998).
40. Yu. E. Lozovik and I. V. Ovchinnikov, Pis'ma Zh. Éksp. Teor. Fiz. **74**, 318 (2001) [JETP Lett. **74**, 288 (2001)]; Phys. Rev. B **66**, 075124 (2002).
41. A. Olaya-Castro, F. J. Rodriguez, L. Quiroga, and C. Tejedor, Phys. Rev. Lett. **87**, 246403 (2001).
42. *Bose–Einstein Condensation*, Ed. by A. Griffin, D. W. Snoke, and S. Stringari (Cambridge Univ. Press, Cambridge, 1995).

43. S. Denev and D. W. Snoke, Phys. Rev. B **65**, 085211 (2002).

44. A. Schmitt, L. Banyai, and H. Haug, Phys. Rev. B **63**, 205113 (2001).

45. G. M. Kavoulakis and A. Mysyrowicz, Phys. Rev. B **61**, 16619 (2000); G. M. Kavoulakis, Y. C. Chang, and G. Baym, Phys. Rev. B **55**, 7593 (1997).

46. M. Y. Shen, T. Yokouchi, S. Koyama, and T. Goto, Phys. Rev. B **56**, 13066 (1997).

47. E. Benson, E. Fortin, and A. Mysyrowicz, Solid State Commun. **101**, 313 (1997); Phys. Status Solidi B **191**, 345 (1995); E. Fortin, S. Fafard, and A. Mysyrowicz, Phys. Rev. Lett. **70**, 3951 (1993); D. W. Snoke, J. P. Wolfe, and A. Mysyrowicz, Phys. Rev. Lett. **64**, 2543 (1990).

48. J. L. Lin and J. P. Wolfe, Phys. Rev. Lett. **71**, 1222 (1993); D. W. Snoke and J. P. Wolfe, Phys. Rev. B **42**, 7876 (1990).

49. H. Shi, G. Verechaka, and A. Griffin, Phys. Rev. B **50**, 1119 (1994); M. Inoue and E. Hanamura, J. Phys. Soc. Jpn. **41**, 1273 (1976).

50. P. Kner, S. Bar-Ad, M. V. Marquezini, *et al.*, Phys. Rev. B **60**, 4731 (1999); Phys. Status Solidi **164**, 579 (1997).

51. S. R. Bolton, U. Neukirch, L. J. Sham, *et al.*, Phys. Rev. Lett. **85**, 2002 (2000); Phys. Rev. B **63**, 115303 (2001).

52. K. El Sayed, D. Birkedal, V. G. Lyssenko, and J. M. Hvam, Phys. Rev. B **55**, 2456 (1997).

Translated by A. Sharshakov

Journal of Experimental and Theoretical Physics, Vol. 98, No. 3, 2004, pp. 594–604.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 125, No. 3, 2004, pp. 673–683.

SOLIDS
Electronic Properties

Properties of a Fast Josephson Vortex

A. S. Malishevskiĭ, V. P. Silin, and S. A. Uryupin

Lebedev Institute of Physics, Russian Academy of Sciences,
Leninskiĭ pr. 53, Moscow, 119991 Russia
e-mail: malish@sci.lebedev.ru

Received May 22, 2003

Abstract—A model is formulated for the analytic description of vortices in a system consisting of a long Josephson junction (JJ) and a waveguide that is magnetically coupled to this junction. The application of this model made it possible to determine the allowed range of velocities of a vortex. It is established that a free vortex can move with a velocity much greater than the Swihart velocity of a Josephson junction. Such a vortex is called fast. The effect of the waveguide on the forced motion of vortices is studied. It is shown that fast vortices can be generated at relatively small values of current. *© 2004 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

Model equations for the difference between the phases of a condensate wave function on different sides of a Josephson junction (JJ) (see, for example, [1–6]) are widely used for the description of the properties of Josephson vortices. The main simplification of the theory in model approaches lies in replacing a sine nonlinearity by relatively simple piecewise linear functions. The efficiency of the application of model approaches has been demonstrated both while describing vortices that carry a single quantum of a magnetic flux [7] and while describing more complicated multivortex structures [8, 9]. Up to now, the model description has been used for studying vortices either in an isolated JJ or in two magnetically coupled JJs [10]. Since the application of model approaches in the theory of simple Josephson structures proved to be efficient, it seems expedient to use a model approach in the theory of vortices in a structure consisting of a JJ coupled to a waveguide. The statement of such a problem was given in [11]. According to this statement, in the present study, we consider vortices in a JJ coupled to a planar waveguide. To obtain analytic characteristics, we apply the Sakai–Tateno–Pedersen model in which the sine of the phase difference is simulated by a sawtooth function of form (2.12) (see below). The application of this model, on the one hand, allows one to correctly describe the allowed and forbidden velocity bands for a freely moving vortex (these bands were determined in [12] in a theory that does not take into account dissipation and does not use an approximate representation for the sine of the phase difference) and, on the other hand, allows one to construct an analytic solution to a system of coupled equations for the phase differences in the JJ and in the waveguide even when dissipation is significant. The latter fact is of interest for the problem of the motion of a vortex under a transport current. A remarkable property of the Sakai–Tateno–Pedersen model is the possibility of describing the role of the Cherenkov effect in a JJ.

In the present paper, we set up a problem of the velocity of a Josephson vortex in a JJ to which a magnetically coupled waveguide is connected. An analytic examination of such a problem is possible owing to the model consideration of the coupling between the JJ and the waveguide. We show that the coupling to a waveguide characterized by the large Swihart velocity allows one to obtain a Josephson vortex moving with a velocity that may be much greater than that of an isolated JJ. This opens up the possibility of implementing fast Josephson vortices.

The results of the model approach are described as follows. In Section 2, relying on the equations that describe a JJ coupled to a waveguide, we describe a free motion of an elementary vortex in such a system. We indicate the range of velocities in which a fast Josephson vortex may exist. The existence of this range of velocities is associated with the effect of the waveguide on the JJ. In Section 3, we investigate the forced motion of this vortex due to a transport current. For sufficiently small losses, we determine the phase differences on the JJ and on the waveguide walls within the Sakai–Tateno–Pedersen model. The use of the model approach has allowed us to determine the contributions, to the phase differences, of the terms that vary over a relatively large scale determined by dissipation. We show how this scale manifests itself in the structure of a vortex and determine the contributions of dissipation in the JJ and the waveguide to the relation between the vortex velocity and current. We show that the forced motion of a fast Josephson vortex may occur under relatively low densities of the transport current. In Section 4, we describe the effect of Cherenkov losses in the JJ coupled to a waveguide on the transport current. In Section 5, we discuss the final results.

1063-7761/04/9803-0594$26.00 © 2004 MAIK "Nauka/Interperiodica"

2. A FREE VORTEX

Consider a double-sandwich-type layered system that consists of three superconducting layers S_1, S_2, and S_3 and two nonsuperconducting layers I and W. The superconducting layers occupy the domains $x < -d$, $d < x < d + L$, and $x > d + L = 2d_w$ and have the London depths λ_1, λ_2, and λ_3, respectively. The substance in the layer I of thickness $2d$, which is sandwiched between the layers S_1 and S_2, has dielectric permittivity ϵ and conductivity σ. The layer W of thickness $2d_w$, which separates the superconductors S_2 and S_3, has dielectric permittivity ϵ_w and conductivity σ_w. Assume that the layer I is thin enough that Cooper pairs may tunnel through it, thus creating a Josephson current of critical density j_c. The thickness of the layer W is assumed to be so large that one can neglect the Josephson current through W compared to the displacement and conductivity currents. This fact allows us to treat the system considered as a JJ magnetically coupled to a waveguide.

Let φ be the phase difference of the condensate wave functions of the superconductors S_1 and S_2 on their boundaries $x = -d$ and $x = d$ with the layer I and φ_w be the phase difference of the condensate wave functions of the superconductors S_2 and S_3 on their boundaries $x = d + L$ and $x = d + L + 2d_w$ with the waveguide W. Assume that the characteristic space scales of variations of φ and φ_w are large compared to the London lengths. Then, following [13, 14], we obtain the following system of equations for φ and φ_w (cf. [11, 12]):

$$\omega_j^2 F[\varphi(z,t)] + \frac{\partial^2\varphi(z,t)}{\partial t^2} = V_s^2\frac{\partial^2\varphi(z,t)}{\partial z^2} + SV_s^2\frac{\partial^2\varphi_w(z,t)}{\partial z^2}, \quad (2.1)$$

$$\frac{\partial^2\varphi_w(z,t)}{\partial t^2} = V_{sw}^2\frac{\partial^2\varphi_w(z,t)}{\partial z^2} + S_wV_{sw}^2\frac{\partial^2\varphi(z,t)}{\partial z^2}. \quad (2.2)$$

Here,

$$V_s^2 \equiv c^2\frac{2d}{\epsilon}\frac{\lambda_3 + 2d_w + \lambda_2\coth(L/\lambda_2)}{\Delta} \quad (2.3)$$

and

$$V_{sw}^2 \equiv c^2\frac{2d_w}{\epsilon_w}\frac{\lambda_1 + 2d + \lambda_2\coth(L/\lambda_2)}{\Delta} \quad (2.4)$$

are the Swihart velocities in the JJ and the waveguide, respectively, where

$$\Delta \equiv \left(\lambda_1 + 2d + \lambda_2\coth\frac{L}{\lambda_2}\right)\times\left(\lambda_3 + 2d_w + \lambda_2\coth\frac{L}{\lambda_2}\right) - \lambda_2^2\operatorname{cosech}^2\frac{L}{\lambda_2} > 0,$$

the coupling constants between the JJ and the waveguide that appear on the right-hand sides of (2.1) and (2.2) are given by

$$S \equiv \frac{\lambda_2\operatorname{cosech}(L/\lambda_2)}{\lambda_3 + 2d_w + \lambda_2\coth(L/\lambda_2)}, \quad (2.5)$$

and

$$S_w \equiv \frac{\lambda_2\operatorname{cosech}(L/\lambda_2)}{\lambda_1 + 2d + \lambda_2\coth(L/\lambda_2)}, \quad (2.6)$$

respectively; ω_j is the Josephson frequency of the JJ; and $F[\varphi]$ is the density of the Josephson current normalized by j_c. We emphasize that the expressions for V_s and V_{sw} contain the thickness L of the superconducting layer S_2 through which the JJ is coupled to the waveguide. This means that V_s and V_{sw} determine the Swihart velocities with regard to the above coupling. In the limit as $L \longrightarrow \infty$, (2.3) and (2.4) reduce to

$$V_s \longrightarrow v_s \equiv c\sqrt{\frac{1}{\epsilon}\frac{2d}{\lambda_1 + \lambda_2 + 2d}},$$

and

$$V_{sw} \longrightarrow v_{sw} \equiv c\sqrt{\frac{1}{\epsilon_w}\frac{2d_w}{\lambda_2 + \lambda_3 + 2d_w}},$$

respectively. The quantities v_s and v_{sw} are the Swihart velocities of noninteracting isolated JJ and waveguide, respectively.

For vortex structures propagating with constant velocity v when $\varphi(z,t) = \psi(\zeta)$ and $\varphi_w(z,t) = \psi_w(\zeta)$, $\zeta \equiv z - vt$, Eqs. (2.1) and (2.2) yield

$$\omega_j^2 F[\psi(\zeta)] - (V_s^2 - v^2)\psi''(\zeta) = SV_s^2\psi_w''(\zeta), \quad (2.7)$$

$$-(V_{sw}^2 - v^2)\psi_w''(\zeta) = S_wV_{sw}^2\psi''(\zeta). \quad (2.8)$$

Setting $\psi(-\infty) = \psi_w(-\infty) = 0$ and $\psi'(-\infty) = \psi_w'(-\infty) = 0$, from (2.8) we obtain

$$\psi_w(\zeta) = -S_w\frac{V_{sw}^2}{V_{sw}^2 - v^2}\psi(\zeta). \quad (2.9)$$

Substituting (2.9) into (2.7), we obtain the following equation for the phase difference ψ on the JJ:

$$\omega_j^2 F[\psi(\zeta)] = \frac{(v_1^2 - v^2)(v_2^2 - v^2)}{V_{sw}^2 - v^2}\psi''(\zeta), \quad (2.10)$$

where the velocities v_1 and v_2 are defined by

$$v_m \equiv \sqrt{\frac{V_s^2 + V_{sw}^2}{2} + (-1)^m \sqrt{\frac{(V_s^2 - V_{sw}^2)^2}{4} + SS_w V_s^2 V_{sw}^2}} > 0, \quad m = 1, 2. \quad (2.11)$$

Below, in the main body of the paper, we will apply the Sakai–Tateno–Pedersen model [1–3] with the following sawtooth function $F[\psi]$:

$$F[\psi] = F_{STP}[\psi] = \begin{cases} (2/\pi)\psi, & -\pi/2 < \psi < \pi/2, \\ (2/\pi)(\pi - \psi), & \pi/2 < \psi < 3\pi/2, \\ (2/\pi)(\psi - 2\pi), & 3\pi/2 < \psi < 5\pi/2. \end{cases} \quad (2.12)$$

According to (2.12) and Eqs. (2.7) and (2.8), the wavenumbers of small perturbations of the phase difference that are characterized by the coordinate dependence $\exp(ik\zeta)$ are determined by the equation

$$\pm\frac{2}{\pi}\omega_j^2 + \frac{(v_1^2 - v^2)(v_2^2 - v^2)}{V_{sw}^2 - v^2}k^2 = 0, \quad (2.13)$$

where the "+" corresponds to the perturbations of ψ near $\psi = 0$ or 2π and the "–" corresponds to the perturbations of ψ near $\psi = \pi$. Equation (2.13) can be rewritten in the following equivalent form:

$$\left[\pm\frac{2}{\pi}\omega_j^2 + (V_s^2 - v^2)k^2\right](V_{sw}^2 - v^2) = SS_w V_s^2 V_{sw}^2 k^2. \quad (2.14)$$

This equation corresponds to the coupling between the ordinary (for the "+") and extraordinary (for "–") Swihart waves of the JJ and an electromagnetic wave in the waveguide.

In the case of "+," we obtain $k = \pm ik_j(v)$ from (2.13), and, in the case of "–," we obtain $k = \pm k_j(v)$, where

$$k_j(v) \equiv \sqrt{\frac{2}{\pi}}\omega_j \sqrt{\frac{V_{sw}^2 - v^2}{(v_1^2 - v^2)(v_2^2 - v^2)}}. \quad (2.15)$$

Using $k_j(v)$, we rewrite Eq. (2.10) for $\psi(\zeta)$ as

$$F_{STP}[\psi(\zeta)] = \frac{2}{\pi}k_j^{-2}(v)\psi''(\zeta). \quad (2.16)$$

The formal similarity between this equation and the equation corresponding to the isolated JJ [7] makes it possible to write out the following solution to Eq. (2.16) that describes an elementary vortex (a 2π kink):

$$\psi(\zeta) = \frac{\pi}{2}\exp\left[k_j(v)\zeta + \frac{\pi}{4}\right], \quad \zeta < -\frac{\pi}{4k_j(v)}, \quad (2.17)$$

$$\psi(\zeta) = \pi + \frac{\pi}{\sqrt{2}}\sin[k_j(v)\zeta], \quad -\frac{\pi}{4k_j(v)} < \zeta < \frac{\pi}{4k_j(v)}, \quad (2.18)$$

$$\psi(\zeta) = 2\pi - \frac{\pi}{2}\exp\left[-k_j(v)\zeta + \frac{\pi}{4}\right], \quad \zeta > \frac{\pi}{4k_j(v)}. \quad (2.19)$$

The phase difference ψ_w on the waveguide walls that corresponds to this vortex is determined from (2.9) and is expressed as

$$\psi_w(\zeta) = -\frac{\pi}{2}\frac{S_w V_{sw}^2}{V_{sw}^2 - v^2}\exp\left[k_j(v)\zeta + \frac{\pi}{4}\right], \quad \zeta < -\frac{\pi}{4k_j(v)},$$

$$\psi_w(\zeta) = -\pi\frac{S_w V_{sw}^2}{V_{sw}^2 - v^2} - \frac{\pi}{\sqrt{2}}\frac{S_w V_{sw}^2}{V_{sw}^2 - v^2}\sin[k_j(v)\zeta], \quad -\frac{\pi}{4k_j(v)} < \zeta < \frac{\pi}{4k_j(v)},$$

$$\psi_w(\zeta) = -2\pi\frac{S_w V_{sw}^2}{V_{sw}^2 - v^2} + \frac{\pi}{2}\frac{S_w V_{sw}^2}{V_{sw}^2 - v^2}\exp\left[-k_j(v)\zeta + \frac{\pi}{4}\right], \quad \zeta > \frac{\pi}{4k_j(v)}.$$

For the solution (2.17)–(2.19) to describe a solitary vortex, i.e., for the equalities $\psi(-\infty) = 0$ and $\psi(\infty) = 2\pi$ to hold, it is necessary that $k_j(v)$ be real and given by (2.15). This requirement implies that the values of

the vortex velocity may only belong to the two allowed bands

$$0 < v < v_1, \tag{2.20}$$

$$V_{sw} < v < v_2. \tag{2.21}$$

On the contrary, in the interval (v_1, V_{sw}) and for $v > v_2$, $k_j(v)$ is pure imaginary, which does not correspond to the motion of a solitary vortex. Recall that, in an isolated JJ, a vortex may move with arbitrary velocities below the Swihart velocity v_s. The splitting of the domain of admissible velocities of a vortex and the appearance of the forbidden band $[v_1, V_{sw}]$ is associated with the effect of the waveguide on the JJ.

The expressions for the velocities v_1 and v_2 that determine the right boundaries of the allowed bands have an especially simple form in the case of weak coupling between the JJ and the waveguide, when the coupling constants S and S_w are small. When $V_{sw} < V_s$, we have

$$v_1 \approx \left(1 - \frac{1}{2}SS_w\frac{V_s^2}{V_s^2 - V_{sw}^2}\right)V_{sw}, \tag{2.22}$$

$$v_2 \approx \left(1 + \frac{1}{2}SS_w\frac{V_{sw}^2}{V_s^2 - V_{sw}^2}\right)V_s. \tag{2.23}$$

In particular, when $V_{sw} \ll V_s$, this implies that the allowed band (2.20) extends from zero to nearly the Swihart velocity of the waveguide; this band is followed by a relatively narrow forbidden band ($V_{sw} - v_1 \approx SS_wV_{sw}/2$), which, in turn, is followed by the broad second allowed band (2.21) extending from the Swihart velocity of the waveguide to a velocity slightly greater than the Swihart velocity of the JJ. In other words, when the Swihart velocity of the waveguide is small compared with the Swihart velocity of the JJ, the effect of the waveguide manifests itself in a slight increase in the limit velocity of a vortex compared with V_s and in the emergence of a narrow (compared with both V_s and V_{sw}) forbidden band for velocities.

When $V_{sw} = V_s$ and the coupling between the JJ and the waveguide is weak, we have

$$v_1 \approx \left(1 - \frac{1}{2}\sqrt{SS_w}\right)V_s,$$

$$v_2 \approx \left(1 + \frac{1}{2}\sqrt{SS_w}\right)V_s.$$

These formulas show that the width of the first allowed band is close to V_s, while the forbidden band and the second allowed band are relatively narrow: $V_{sw} - v_1 \approx v_2 - V_{sw} \approx \sqrt{SS_w}\,V_s/2$. This means that, just as in the previous case, the effect of the waveguide manifests itself, first, in a slight increase in the limit velocity of a vortex compared with V_s and, second, in the emergence of a narrow, compared with V_s, forbidden band for velocities.

Finally, when $V_{sw} > V_s$ and the coupling between the JJ and the waveguide is still weak, from (2.11) we obtain

$$v_1 \approx \left(1 - \frac{1}{2}SS_w\frac{V_{sw}^2}{V_{sw}^2 - V_s^2}\right)V_s, \tag{2.24}$$

$$v_2 \approx \left(1 + \frac{1}{2}SS_w\frac{V_s^2}{V_{sw}^2 - V_s^2}\right)V_{sw}. \tag{2.25}$$

In particular, in the most interesting case when $V_{sw} \gg V_s$, the first allowed band extends from zero to nearly the Swihart velocity of the JJ, which is followed by a relatively broad forbidden band $(1 - SS_w/2)V_s \le v \le V_{sw}$, which, in turn, is followed by the second allowed band

$$V_{sw} < v < v_2 \approx \left(1 + \frac{1}{2}SS_w\frac{V_s^2}{V_{sw}^2}\right)V_{sw}. \tag{2.26}$$

The width of this band is small compared with both V_{sw} and V_s. Since this region lies near the Swihart velocity of the waveguide, which was assumed to be large compared with the Swihart velocity of the JJ, it can be called the existence region of a fast Josephson vortex.

As already mentioned, in an isolated JJ, a vortex may move with a velocity less than the Swihart velocity v_s in the JJ. When the JJ is connected to a waveguide, a new phenomenon arises, a fast Josephson vortex. This vortex moves with velocity much greater than the Swihart velocity of the JJ. Such an increase in the vortex velocity is attractive since the Swihart velocity of the waveguide may be close to the velocity of light $c/\sqrt{\varepsilon_w}$ in the substance that fills the waveguide. Thus, formula (2.26) points to the following interesting fact: a Josephson vortex may move with very high velocity that is close to the velocity of light in the dielectric that fills the waveguide.

Thus, when the Swihart velocity of the waveguide is large compared with the Swihart velocity of the JJ, the effect of the waveguide is manifested in the fact that, first, the right boundary of the first allowed band slightly decreases compared with V_s, second, the forbidden band, which extends nearly from V_s to V_{sw}, is broader than both allowed bands, and, third, there appears a region of a fast vortex that moves with a velocity close to the Swihart velocity of the waveguide.

We emphasize that the splitting of the range of velocities in which a Josephson 2π kink may exist and the emergence of a forbidden velocity band are not related to the choice of the Sakai–Tateno–Pedersen nonlinearity but are attributed to the effect of the waveguide on the JJ and also occur, for example, in the conventional model with sine nonlinearity (see Appendix 1).

3. FORCED MOTION OF A VORTEX

In this section, we consider a forced motion of a vortex under the action of a dc transport current of density j that flows through the JJ. As in the previous section, we consider a uniformly moving vortex when the accelerating effect of the transport current is compensated by dissipative losses.

Using $\beta \equiv 4\pi\sigma/\epsilon$ and $\beta_w \equiv 4\pi\sigma_w/\epsilon_w$ instead of (2.1) and (2.2) to describe the losses in the JJ and the waveguide, respectively, we obtain (cf. [11])

$$\omega_j^2 F[\varphi(z,t)] + \frac{\partial^2\varphi(z,t)}{\partial t^2} + \omega_j^2\frac{j}{j_c} + \beta\frac{\partial\varphi(z,t)}{\partial t} = V_s^2\frac{\partial^2\varphi(z,t)}{\partial z^2} + SV_s^2\frac{\partial^2\varphi_w(z,t)}{\partial z^2}, \quad (3.1)$$

$$\frac{\partial^2\varphi_w(z,t)}{\partial t^2} + \beta_w\frac{\partial\varphi_w(z,t)}{\partial t} = V_{sw}^2\frac{\partial^2\varphi_w(z,t)}{\partial z^2} + S_wV_{sw}^2\frac{\partial^2\varphi(z,t)}{\partial z^2}. \quad (3.2)$$

In the case of a steady motion at a constant velocity v, from (3.1) and (3.2), we obtain

$$\omega_j^2 F[\psi(\zeta)] - (V_s^2 - v^2)\psi''(\zeta) + \omega_j^2\frac{j}{j_c} - \beta v\psi'(\zeta) = SV_s^2 f'(\zeta), \quad (3.3)$$

$$-(V_{sw}^2 - v^2)f'(\zeta) - \beta_w v f(\zeta) = S_wV_{sw}^2\psi''(\zeta), \quad (3.4)$$

where $f(\zeta) \equiv \psi_w'(\zeta)$.

When $F[\psi]$ is used in the form (2.12) corresponding to the Sakai–Tateno–Pedersen model, the wave vectors of small perturbations of ψ and f are determined from the equation

$$\left[\pm\frac{2}{\pi}\omega_j^2 + (V_s^2 - v^2)k^2 - i\beta vk\right] \times [(V_{sw}^2 - v^2)k - i\beta_w v] = SS_wV_s^2V_{sw}^2k^3, \quad (3.5)$$

in which the "+" and "–" have the same meaning as in Eqs. (2.13) and (2.14). In contrast to the dissipationless case, Eq. (3.5) has three solutions. Let us write out these solutions on the assumption that dissipation is small.

When the left-hand side of Eq. (3.5) has a "+," two roots $k = \pm ik_j(v) + i\alpha k_j(v)$ of this equation, where

$$\alpha \equiv \frac{\pi}{4}\frac{vk_j(v)\beta(v)}{\omega_j^2} \ll 1, \quad (3.6)$$

$$\beta(v) \equiv \beta + SS_w\frac{V_s^2V_{sw}^2}{(V_{sw}^2 - v^2)^2}\beta_w, \quad (3.7)$$

differ from the roots $k = \pm ik_j(v)$ of the dissipationless equation (2.14) by a small dissipative term $i\alpha k_j(v)$. The third root of Eq. (3.5) with a "+" on the left-hand side, which is associated with dissipation, is $k = i(1 - \alpha_w)k_w$, where

$$k_w \equiv \frac{v\beta_w}{V_{sw}^2 - v^2}, \quad (3.8)$$

$$\alpha_w \equiv \frac{\pi}{2}SS_w\frac{V_s^2V_{sw}^2}{V_{sw}^2 - v^2}\frac{k_w^2}{\omega_j^2} \ll 1. \quad (3.9)$$

When the left-hand side of Eq. (3.5) has a "–," in the low-loss limit, this equation has the following solutions: $k = \pm k_j(v) + i\alpha k_j(v)$, $i(1 + \alpha_w)k_w$.

Conditions (3.6) and (3.9) for the smallness of α and α_w agree with our assumption that dissipation is small.

Let us write a solution to the system of Eqs. (3.3) and (3.4), which describes a Josephson 2π kink and the associated field in the waveguide in the first allowed velocity band (2.20), for $k_w > 0$. Assume that $\psi(\zeta)$ takes the values $\pi/2$ and $3\pi/2$ at points $\zeta = -\zeta_0$ and $\zeta = \zeta_0$, respectively.

We will seek a solution to systems (3.3), (3.4) as a superposition of constants and terms of the form $\exp(ik\zeta)$, where k is a solution to Eq. (3.5). Then, requiring that the solution do not contain terms that grow exponentially as $|\zeta| \longrightarrow \infty$, we obtain

$$\psi(\zeta) = -\frac{\pi}{2}\frac{j}{j_c} + \frac{\pi}{2}\left(1 + \frac{j}{j_c}\right) \times \exp[(1-\alpha)k_j(v)(\zeta + \zeta_0)], \quad (3.10)$$

$$\psi_w'(\zeta) = A_w\exp[(1-\alpha)k_j(v)(\zeta + \zeta_0)] \quad (3.11)$$

in the tail of the vortex ($\zeta < -\zeta_0$), where $-\pi j/2j_c < \psi < \pi/2$;

$$\psi(\zeta) = \pi + \frac{\pi}{2}\frac{j}{j_c} + \{B\sin[k_j(v)\zeta] + C\cos[k_j(v)\zeta]\} \times \exp[-\alpha k_j(v)\zeta] + D\exp[-(1 + \alpha_w)k_w\zeta], \quad (3.12)$$

$$\psi'_w(\zeta) = \{B_w \sin[k_j(v)\zeta] + C_w \cos[k_j(v)\zeta]\} \times \exp[-\alpha k_j(v)\zeta] + D_w \exp[-(1+\alpha_w)k_w\zeta] \quad (3.13)$$

in the middle part of the vortex ($-\zeta_0 < \zeta < \zeta_0$), where $-\pi/2 < \psi < 3\pi/2$; and

$$\psi(\zeta) = 2\pi - \frac{\pi}{2}\frac{j}{j_c} - \left(1 - \frac{j}{j_c}\right) \times E\exp[-(1+\alpha)k_j(v)(\zeta - \zeta_0)] + \left(1 - \frac{j}{j_c}\right) \times \left(E - \frac{\pi}{2}\right)\exp[-(1-\alpha_w)k_w(\zeta - \zeta_0)], \quad (3.14)$$

$$\psi'_w(\zeta) = E_w \exp[-(1+\alpha)k_j(v)(\zeta - \zeta_0)] + F_w \exp[-(1-\alpha_w)k_w(\zeta - \zeta_0)] \quad (3.15)$$

in the head of the vortex ($\zeta > \zeta_0$), where $3\pi/2 < \psi < 2\pi - (\pi j/2j_c)$.

When writing Eqs. (3.10) and (3.14), we took into account that $\psi(-\zeta_0 - 0) = \pi/2$ and $\psi(\zeta_0 + 0) = 3\pi/2$.

From the continuity condition for the functions $\psi(\zeta)$, $\psi'(\zeta)$, and $\psi'_w(\zeta)$ at $\zeta = \pm\zeta_0$ and from the relations between the quantities B, C, D, E, A_w, B_w, C_w, D_w, E_w, and F_w that arise under the substitution of (3.10)–(3.15) into any of the equations of the system (3.3), (3.4), we obtain a system of 12 equations from which we can determine the size of the middle region $2\zeta_0$, the relation between the current density j and the velocity of a vortex structure, and the coefficients of the exponential and trigonometric functions in (3.10)–(3.15). From this system of equations, in the linear approximation in dissipation, we obtain

$$\frac{j(v)}{j_c} = \frac{\pi}{4}\left(1 + \frac{\pi}{4}\right)\frac{v k_j(v)\beta(v)}{\omega_j^2} = \frac{\sqrt{\pi}}{2\sqrt{2}}\left(1 + \frac{\pi}{4}\right)\frac{v}{\omega_j}\sqrt{\frac{V_{sw}^2 - v^2}{(v_1^2 - v^2)(v_2^2 - v^2)}} \times \left[\beta + SS_w\frac{V_s^2 V_{sw}^2}{(V_{sw}^2 - v^2)^2}\beta_w\right], \quad (3.16)$$

$$\zeta_0 = \frac{\pi}{4k_j(v)} + O(\beta^2, \beta_w^2, \beta\beta_w). \quad (3.17)$$

Again, under the assumption that α and α_w are small, we obtain the expressions, given in Appendix 2, for the coefficients that define the functions $\psi(\zeta)$ and $\psi'_w(\zeta)$.

Substituting (3.17) and the corresponding expressions from Appendix 2 into (3.10)–(3.15), we obtain the following expressions for a Josephson 2π kink and the associated field in the waveguide in the allowed velocity band (2.20) to the lowest order in small dissipation:

$$\psi(\zeta) \approx -\frac{\pi}{2}\frac{j}{j_c} + \frac{\pi}{2}\left(1 + \frac{j}{j_c}\right) \times \exp\left\{(1-\alpha)k_j(v)\left[\zeta + \frac{\pi}{4k_j(v)}\right]\right\}, \quad (3.18)$$

$$\psi'_w(\zeta) \approx -\frac{\pi}{2}S_w\frac{V_{sw}^2}{V_{sw}^2 - v^2}k_j(v)\left[1 + \frac{\pi\alpha}{4} - \frac{k_w}{k_j(v)}\right] \times \exp\left\{(1-\alpha)k_j(v)\left[\zeta + \frac{\pi}{4k_j(v)}\right]\right\} \quad (3.19)$$

in the tail of the vortex ($\zeta < -\pi/4k_j(v)$), where $-\pi j/2j_c < \psi < \pi/2$;

$$\psi(\zeta) \approx \pi + \frac{\pi}{2}\frac{j}{j_c} + \frac{\pi}{\sqrt{2}}\{\sin[k_j(v)\zeta] - \alpha\cos[k_j(v)\zeta]\}\exp[-\alpha k_j(v)\zeta] + \frac{\pi^2}{2}\frac{SS_w V_s^2 V_{sw}^2}{\omega_j^2(V_{sw}^2 - v^2)}k_w^2 \exp[-(1+\alpha_w)k_w\zeta], \quad (3.20)$$

$$\psi'_w(\zeta) \approx \frac{\pi}{\sqrt{2}}S_w\frac{V_{sw}^2}{V_{sw}^2 - v^2} \times \{k_w \sin[k_j(v)\zeta] - k_j(v)\cos[k_j(v)\zeta]\} \times \exp[-\alpha k_j(v)\zeta] + \pi S_w\frac{V_{sw}^2}{V_{sw}^2 - v^2}k_w \times \exp[-(1+\alpha_w)k_w\zeta] \quad (3.21)$$

in the middle part of the vortex ($-\pi/4k_j(v) < \zeta < \pi/4k_j(v)$), where $-\pi/2 < \psi < 3\pi/2$; and

$$\psi(\zeta) \approx 2\pi - \frac{\pi}{2}\frac{j}{j_c} - \frac{\pi}{2}\left(1 - \frac{j}{j_c}\right) \times \exp\left\{-(1+\alpha)k_j(v)\left[\zeta - \frac{\pi}{4k_j(v)}\right]\right\} - \pi^2 SS_w\frac{V_s^2 V_{sw}^2}{\omega_j^2(V_{sw}^2 - v^2)}k_w^2 \times \exp\left\{-(1-\alpha_w)k_w\left[\zeta - \frac{\pi}{4k_j(v)}\right]\right\}, \quad (3.22)$$

$$\psi'_w(\zeta) \approx -\frac{\pi}{2}S_w\frac{V_{sw}^2}{V_{sw}^2 - v^2}k_j(v)\left[1 - \frac{\pi\alpha}{4} + \frac{k_w}{k_j(v)}\right]$$

$$\times \exp\left\{-(1+\alpha)k_j(v)\left[\zeta - \frac{\pi}{4k_j(v)}\right]\right\} \quad (3.23)$$

$$+ 2\pi S_w\frac{V_{sw}^2}{V_{sw}^2 - v^2}k_w \exp\left\{-(1-\alpha_w)k_w\left[\zeta - \frac{\pi}{4k_j(v)}\right]\right\}$$

in the head of the vortex ($\zeta > \pi/4k_j(v)$), where $3\pi/2 < \psi < 2\pi - (\pi j/2j_c)$.

In the second allowed velocity band (2.21) of a vortex, when $k_w < 0$, a solution to the system of Eqs. (3.3), (3.4) is sought in a form similar to (3.10)–(3.15) with the difference that the contributions (associated with the dissipation) of the third roots $i(1 \pm \alpha_w)k_w$ of Eqs. (3.5) arise in the middle part and the tail of a vortex rather than in the middle part and the head. This solution has a form similar to (3.18)–(3.23) (see Appendix 3). Here, the relation between the transport current and the vortex velocity in a linear approximation in β and β_w is given by (3.16) as before.

To conclude this section, we note that, to obtain formula (3.16), we can apply a method that involves the approximate solution of Eqs. (3.3) and (3.4) (see Appendix 4). This method allows one to write the function $j(v)$, which differs from (3.16) only by a numerical factor, in the model with sine nonlinearity as well (see Appendix 4):

$$\frac{j(v)}{j_c} = \frac{4\sqrt{2}}{\pi^{3/2}}\frac{vk_j(v)\beta(v)}{\omega_j^2}. \quad (3.24)$$

4. THE EFFECT OF CHERENKOV LOSSES ON THE FORCED MOTION OF A VORTEX

A remarkable property of the Sakai–Tateno–Pedersen approach is the fact that it makes it possible to consider the effect of Cherenkov losses on the motion of a vortex. The results of this analysis are presented in this section as applied to the conditions under which $V_s \ll V_{sw}$ and there exists a fast vortex. When $V_s \ll V_{sw}$, there are two velocity ranges in each of which vortices exist.

First, consider a range of small velocities when one can speak of the motion of a relatively slow vortex with a velocity of $v < v_1 \approx (1 - SS_w)^{1/2}V_s$. To describe the Cherenkov losses of a slow vortex, it is sufficient to take into account the spatial dispersion of the JJ. Let us restrict the analysis to the limit of relatively weak spatial dispersion of the JJ, which is possible when the vortex velocity satisfies the condition

$$1 - (v/v_1)^2 \ll 1. \quad (4.1)$$

In addition, when considering the Cherenkov losses, we will neglect dissipation in the JJ and the waveguide. This approach is justified for low dissipation and small Cherenkov losses, when their effect on the motion of a vortex is additive. Under these conditions, the forced motion of a slow vortex is described by Eqs. (3.1) and (3.2) with $\beta = \beta_w = 0$ when a term of the form

$$\frac{1}{2}\omega_j^2\lambda_j^2\lambda_J^2\frac{\partial^4\varphi(z,t)}{\partial z^4}, \quad (4.2)$$

where $\lambda_j \equiv v_s/\omega_j$ and $\lambda_J^2 \equiv (\lambda_1^3 + \lambda_2^3)(\lambda_1 + \lambda_2 + 2d)^{-1}$, is added to the right-hand side of (3.1) to take into account the weak spatial dispersion of the JJ. In writing the small term (4.2), we neglect the interaction between the JJ and the waveguide. Taking into account the above changes in Eqs. (3.1) and (3.2), we use the following equation to describe the forced motion of a slow vortex:

$$F[\psi(\zeta)] = \frac{2}{\pi}k_j^{-2}(v)\psi''(\zeta) + \frac{1}{2}\lambda_j^2\lambda_J^2\psi^{IV}(\zeta) - \frac{j}{j_c}, \quad (4.3)$$

where $k_j^2(v) \approx (2/\pi)\omega_j^2(v_1^2 - v^2)^{-1}$. Equation (4.3) formally differs from the equation considered in [7] in that the coefficients λ^2 are replaced by λ_J^2 and $(v_s^2 - v^2)\omega_j^{-2}$ is replaced by $(2/\pi)k_j^{-2}(v)$. This allows one to make use of the mathematical result of [7]. For instance, following [7], for the velocities of a slow vortex that satisfy the condition

$$v_1^2 \gg v_1^2 - v^2 \gg \frac{2}{\sqrt{\pi}}\frac{\lambda_J}{\lambda_j}v_s^2 \approx \frac{2}{\sqrt{\pi}}\frac{\lambda_J}{\lambda_j}v_1^2, \quad (4.4)$$

we obtain the following relation between the current and the vortex velocity:

$$\frac{j}{j_c} = \frac{\varepsilon^4}{8}\left(\sin\frac{\pi}{2\varepsilon} - \frac{\varepsilon}{2}\cos\frac{\pi}{2\varepsilon}\right)^2, \quad (4.5)$$

where $\varepsilon \equiv \sqrt{\pi}\lambda_j\lambda_J k_j^2(v) \ll 1$. Formula (4.5) gives an oscillatory relation between the current and the vortex velocity, which is established when there is balance between the effects of current and Cherenkov losses due to the irradiation of waves by the vortex. In this case, the minima of the function (4.5) correspond to a discrete set of eigenvelocities v_n of a free motion of a Josephson vortex; this set is associated with the internal structure, of the vortex, created by extraordinary Swihart waves Cherenkov-trapped by the vortex (see [7]). In the case of a slow vortex, the discrete set of velocities v_n is given by

$$v_n \approx \left(1 - \frac{2}{\sqrt{\pi}}\frac{\lambda_J}{\lambda_j}n\right)v_1, \quad (4.6)$$

where, according to inequalities (4.4), positive integers n lie in the interval $1 \ll n \ll \lambda_j/\lambda_J$.

Let us consider the range of velocities of a fast vortex when $V_{sw} < v < v_2 \approx V_{sw} + (SS_w/2)(V_s^2/V_{sw})$. The properties of a fast vortex are primarily determined by the waveguide. This means that, when considering the Cherenkov losses of a fast vortex, it suffices to take into account the spatial dispersion of the waveguide. As in the case of a slow vortex, we will neglect the small dissipation in the JJ and the waveguide. Then, to describe the Cherenkov losses of a fast vortex, we have Eqs. (3.1) and (3.2) with $\beta = \beta_w = 0$ when a term of the form

$$\frac{1}{2}v_{sw}^2\lambda_w^2\frac{\partial^4\varphi_w(z,t)}{\partial z^4}, \tag{4.7}$$

where $\lambda_w^2 \equiv (\lambda_2^3 + \lambda_3^3)(\lambda_2 + \lambda_3 + 2d_w)^{-1}$ is added to the right-hand side of Eq. (3.2). The small term (4.7) does not take into account weak interaction between the JJ and the waveguide. When the velocity of a fast vortex satisfies the condition

$$v_2^2 - v^2 \ll SS_wV_s^2, \tag{4.8}$$

modified equation (3.2) enables one to express ψ_w in terms of ψ and write the following equation for the phase difference of the fast vortex:

$$F[\psi(\zeta)] = \frac{2}{\pi}k_j^{-2}(v)\psi''(\zeta) + \frac{1}{2}\lambda_j^2\lambda_{eff}^2\psi^{IV}(\zeta) - \frac{j}{j_c}, \tag{4.9}$$

where the effective length λ_{eff} depends on the vortex velocity:

$$\begin{aligned}\lambda_{eff} &\equiv \lambda_w\frac{\sqrt{SS_w}}{v^2 - V_{sw}^2}v_{sw}V_{sw}\frac{V_s}{\lambda_j\omega_j}\\ &\approx \sqrt{SS_w}\frac{V_{sw}^2}{v^2 - V_{sw}^2}\lambda_w.\end{aligned} \tag{4.10}$$

Since $v_2^2 \approx V_{sw}^2 + SS_wV_s^2$, inequality (4.8) implies that $v^2 - v_{sw}^2 \approx SS_wV_s^2$. For such velocities of a vortex, the effective length λ_{eff} remains finite. However, owing to the condition $V_{sw} \gg V_s$, $\lambda_{eff} \approx \lambda_w V_{sw}^2/V_s^2\sqrt{SS_w}$ may substantially exceed λ_J; this makes ir possible to neglect the dispersion of the JJ when considering the Cherenkov losses of a fast vortex. A formal similarity between Eqs. (4.3) and (4.9) implies that the relation between the current and the velocity of a fast vortex is also given by (4.5), which now contains a new small parameter

$$\begin{aligned}\varepsilon \equiv \sqrt{\pi}\lambda_j\lambda_{eff}k_j^2(v) &= \frac{2}{\sqrt{\pi}}\frac{v_{sw}}{V_{sw}}\frac{\lambda_w}{\lambda_j}\\ &\times \sqrt{SS_w}\frac{v_sV_s}{v_2^2 - v^2} \ll 1.\end{aligned} \tag{4.11}$$

Setting $v_{sw} \approx V_{sw}$ and $v_s \approx V_s$, we can see that conditions (4.8) and (4.11) are consistent if

$$SS_w \gg \frac{4}{\pi}\left(\frac{\lambda_w}{\lambda_j}\right)^2. \tag{4.12}$$

In the case of a fast vortex, the minima of the oscillation function (4.5) are attained at a velocity close to the eigenvelocities of freely moving fast vortices, which are approximately given by

$$v_n \approx \left(1 - \frac{2}{\sqrt{\pi}}\sqrt{SS_w}\frac{\lambda_w}{\lambda_j}\frac{V_s^2}{v_2^2}n\right)v_2. \tag{4.13}$$

Here, according to inequalities (4.8) and (4.11), positive integers n lie in the interval $1 \ll n \ll \sqrt{SS_w}(\lambda_j/\lambda_w)$.

5. DISCUSSION

Note that Eq. (3.16) differs from the function $j(v)$ for an isolated JJ [7],

$$\frac{j(v)}{j_c} = \frac{\sqrt{\pi}}{2\sqrt{2}}\left(1 + \frac{\pi}{4}\right)\frac{v}{\sqrt{v_s^2 - v^2}}\frac{\beta}{\omega_j},$$

by the velocity dependence of the term associated with losses in the JJ and by the presence of a term determined by the losses in the waveguide. We emphasize that the terms in (3.16) that contain β and β_w, which characterize the dissipation in the JJ and the waveguide, respectively, depend differently on velocity.

For the most interesting case of a fast Josephson vortex, whose velocity is given by (2.26), the function (3.16) can approximately be represented as

$$\begin{aligned}\frac{j(v)}{j_c} &\approx \frac{\sqrt{\pi}}{2\sqrt{2}}\left(1 + \frac{\pi}{4}\right)\frac{1}{\omega_j}\\ &\times\left[\beta + \frac{1}{4}SS_w\frac{V_s^2}{(v - V_{sw})^2}\beta_w\right]\sqrt{\frac{v - V_{sw}}{v_2 - v}},\end{aligned}$$

$$V_{sw} < v < \left(1 + \frac{1}{2}SS_w\frac{V_s^2}{V_{sw}^2}\right)V_{sw}.$$

Under the condition

$$\beta < \frac{V_{sw}^2}{SS_w V_s^2}\beta_w,$$

the dissipation is mainly determined by losses in the waveguide. In this case, we have

$$\frac{j(v)}{j_c} \approx \frac{\sqrt{\pi}}{8\sqrt{2}}\left(1+\frac{\pi}{4}\right)\frac{\beta_w}{\omega_j}SS_w\frac{V_s^2}{\sqrt{(v_2-v)(v-V_{sw})^3}}.$$

For $v = (V_{sw} + 3v_2)/4 \approx [1 + (3/8)SS_w(V_s^2/V_{sw}^2)]V_{sw}$, the function on the right-hand side attains its minimum, equal to

$$\frac{j_{min}}{j_c} \approx \frac{4\sqrt{2\pi}}{3\sqrt{3}}\left(1+\frac{\pi}{4}\right)\frac{1}{SS_w}\frac{V_{sw}^2}{V_s^2}\frac{\beta_w}{\omega_j}.$$

Thus, the motion of a fast Josephson vortex may be realized when the transport current is greater than j_{min}. This means that the forced motion of a fast vortex, which occurs due to the coupling between the JJ and the waveguide, does not require large values of current.

The expression for $\psi'_w(\zeta)$ and the terms in $\psi(\zeta)$ that do not contain the small quantity k_w in the arguments of the exponential functions are expressed to the first order in β and β_w. The coefficients of $\exp[(1 + \alpha_w)k_w\zeta]$ in the expressions for $\psi(\zeta)$ are proportional to k_w^2. Taking into account these small contributions to $\psi(\zeta)$ does overlap the accuracy limit because precisely these contributions make it possible to write correct expressions for $\psi'_w(\zeta)$. In (3.22), the last small dissipative term ($\propto k_w^2$) is localized within the scale $\sim k_w^{-1}$, which is much greater than the scale $\sim k_j^{-1}(v)$, in which the third term is localized. We emphasize that, despite the fact that the terms in the phase differences that correspond to the small third roots of Eqs. (3.5) are localized in a large spatial region, they give small corrections to the law (3.16). To illustrate this fact, we write the energy of the system (per unit length of axis y) as

$$H = \frac{\phi_0^2}{32\pi^3c^2}\int_{-\infty}^{\infty}d\zeta\left\{\frac{\epsilon\omega_j^2}{d}\int_0^{\psi(\zeta)}d\psi F_{STP}[\psi]\right.$$

$$+\frac{\epsilon(v^2+V_s^2)}{2d}[\psi'(\zeta)]^2+\frac{\epsilon_w(v^2+V_{sw}^2)}{2d_w}[\psi'_w(\zeta)]^2$$

$$\left.+\frac{1}{2}\left(S\frac{\epsilon V_s^2}{d}+S_w\frac{\epsilon_w V_{sw}^2}{d_w}\right)\psi(\zeta)\psi'_w(\zeta)\right\}.$$

Since a vortex moves uniformly, its energy is conserved. Therefore, taking into account (3.3) and (3.4), calculating the derivative dH/dt, and equating it to zero, we obtain the following relation:

$$\frac{\phi_0 j(v)v}{c} = \frac{\phi_0^2 v^2}{32\pi^3c^2}$$
$$\times\int_{-\infty}^{\infty}d\zeta\left\{\frac{\epsilon}{d}\beta[\psi'(\zeta)]^2+\frac{\epsilon_w}{d_w}\beta_w[\psi'_w(\zeta)]^2\right\}. \quad (5.1)$$

The left-hand side of Eq. (5.1) represents the power of the Lorenz force associated with the effect of the transport current on the vortex. The right-hand side of (5.1) contains the power of friction forces associated with the bulk ohmic losses in the JJ and in the waveguide. Substituting the terms of $\psi(\zeta)$ and $\psi'_w(\zeta)$ that contain $\exp[(1 \pm \alpha_w)k_w\zeta]$ into (5.1), we obtain

$$\frac{\phi_0^2}{16\pi c^2}\frac{\epsilon_w}{d_w}\frac{S_w^2V_{sw}^4}{(V_{sw}^2-v^2)}v^3\beta_w^2.$$

This expression shows that the energy losses of a vortex per unit time that are associated with the excitation of large-scale perturbations (whose scale $\sim k_w^{-1} \propto \beta_w^{-1}$ is determined by dissipation) are proportional to β_w^2. In other words, in a linear approximation in dissipation, dissipative contributions, to the phase difference, localized on large scales are inessential for the dependence $j(v)$. The possibility of drawing such a conclusion is one of advantages of our analysis, which is due to the use of the model description.

If we ignore the effect of dissipation on the coordinate dependence of the phase differences, we can obtain formula (3.16) by substituting dissipationless expressions (2.9) and (2.17)–(2.19) into (5.1). The dependence $j(v)$ in a linear approximation in the case of sine nonlinearity can be obtained similarly: substituting (2.9) and (A1.2) into (5.1), which is independent of the form of the nonlinearity $F[\psi]$, we obtain expression (3.24).

Finally, we established the influence of the Cherenkov losses on the transport current in the Sakai–Tateno–Pedersen model, which manifests itself both for vortex velocities less than the Swihart velocity of the JJ and for velocities of a fast vortex. In either case, the function $j(v)$ exhibits oscillatory behavior, which was earlier established for a simple case of an isolated JJ. The minima of $j(v)$ correspond to the velocities of a free motion of both slow (cf. [7]) and fast Josephson vortices. Under the conditions of small dissipation and low Cherenkov losses, the oscillating part of $j(v)$ is added to the mono-

tonic part of $j(v)$ associated with dissipation in the JJ and the waveguide.

6. CONCLUSIONS

Thus, we have formulated a theory for the effect of a magnetically coupled waveguide on the vortex in a Josephson junction. We have established the properties in a fast Josephson vortex with a velocity much greater than the Swihart velocity of the JJ.

ACKNOWLEDGMENTS

This work was supported in part by a grant from the President of the Russian Federation for supporting young Russian scientists (project no. MK-1809-2003.02), a grant for supporting leading scientific schools of the Russian Federation (project no. NSh-1385-2003.2), and by the Ministry of Science and Industry of the Russian Federation (project no. 40.012.11.1357).

APPENDIX 1

The analysis presented in Section 2 can be extended to the case of a standard sine nonlinearity when the density of the Josephson current depends on φ by the law $j_c \sin\varphi(z, t)$; then, instead of Eq. (2.16), we have

$$\sin\psi(\zeta) = \frac{2}{\pi}k_j^{-2}(v)\psi''(\zeta). \qquad \text{(A.1.1)}$$

This equation differs from the well-known sine-Gordon equation in that the coefficient of the second derivative depends on velocity. A solution to (A1.1) corresponding to a 2π kink is given by

$$\psi(\zeta) = 4\arctan\left\{\exp\left[\sqrt{\frac{\pi}{2}}k_j(v)\zeta\right]\right\}. \qquad \text{(A.1.2)}$$

For solution (A1.2) to describe a vortex, the quantity $k_j(v)$ should be real, or, which is equivalent, the vortex velocity should take values only within two allowed bands (2.20) and (2.21). The latter means that, in the model with sine nonlinearity, the range of allowed velocities of a 2π kink in the JJ coupled to a waveguide also splits into two regions separated by a forbidden band. Thus, the emergence of the forbidden velocity band is not associated with the Sakai–Tateno–Pedersen nonlinearity that we chose in the main body of the text but is determined by the coupling between an eigen electromagnetic mode of the waveguide and the Swihart wave of the JJ.

APPENDIX 2

The coefficients in the expressions (3.11)–(3.15) for the phase differences are given by

$$A_w \approx -\frac{\pi}{2}S_w(v)\left[1 + \frac{\pi\alpha}{4} - \frac{k_w}{k_j(v)}\right]k_j(v),$$

$$B = \frac{\pi}{\sqrt{2}} + O(\beta^2, \beta_w^2, \beta\beta_w),$$

$$C \approx -\frac{\pi}{\sqrt{2}}\alpha, \quad D \approx \frac{\pi^2}{2}\frac{SS_w(v)V_s^2}{\omega_j^2}k_w^2,$$

$$B_w \approx \frac{\pi}{\sqrt{2}}S_w(v)k_w,$$

$$C_w = -\frac{\pi}{\sqrt{2}}S_w(v)k_j(v) + O(\beta^2, \beta_w^2, \beta\beta_w),$$

$$D_w \approx \pi S_w(v)k_w, \quad E \approx \frac{\pi}{2} - \pi^2\frac{SS_w(v)V_s^2}{\omega_j^2}k_w^2,$$

$$E_w \approx -\frac{\pi}{2}S_w(v)k_j(v)\left[1 - \frac{\pi\alpha}{4} + \frac{k_w}{k_j(v)}\right],$$

$$F_w \approx 2\pi S_w(v)k_w,$$

where $S_w(v) \equiv S_w V_{sw}^2/(V_{sw}^2 - v^2)$.

APPENDIX 3

For low dissipation, a solution to Eqs. (3.3) and (3.4) corresponding to the motion of an elementary vortex in a JJ with velocities $V_{ws} < v < v_2$ has the form

$$\psi(\zeta) \approx -\frac{\pi}{2}\frac{j}{j_c} + \frac{\pi}{2}\left(1 + \frac{j}{j_c}\right)$$

$$\times \exp\left\{(1-\alpha)k_j(v)\left[\zeta + \frac{\pi}{4k_j(v)}\right]\right\}$$

$$+ \pi^2 SS_w\frac{V_s^2V_{sw}^2}{\omega_j^2(V_{sw}^2 - v^2)}k_w^2$$

$$\times \exp\left\{-(1-\alpha_w)k_w\left[\zeta + \frac{\pi}{4k_j(v)}\right]\right\},$$

$$\psi_w'(\zeta) \approx -\frac{\pi}{2}S_w\frac{V_{sw}^2}{V_{sw}^2 - v^2}k_j(v)\left[1 + \frac{\pi\alpha}{4} - \frac{k_w}{k_j(v)}\right]$$

$$\times \exp\left\{(1-\alpha)k_j(v)\left[\zeta + \frac{\pi}{4k_j(v)}\right]\right\}$$

$$-2\pi S_w \frac{V_{sw}^2}{V_{sw}^2 - v^2} k_w \exp\left\{-(1-\alpha_w)k_w\left[\zeta + \frac{\pi}{4k_j(v)}\right]\right\}$$

in the tail of the vortex ($\zeta < -\pi/4k_j(v)$) and

$$\psi(\zeta) \approx 2\pi - \frac{\pi}{2}\frac{j}{j_c} - \frac{\pi}{2}\left(1 - \frac{j}{j_c}\right)$$

$$\times \exp\left\{-(1+\alpha)k_j(v)\left[\zeta - \frac{\pi}{4k_j(v)}\right]\right\},$$

$$\psi_w'(\zeta) \approx -\frac{\pi}{2} S_w \frac{V_{sw}^2}{V_{sw}^2 - v^2} k_j(v)\left[1 - \frac{\pi\alpha}{4} + \frac{k_w}{k_j(v)}\right]$$

$$\times \exp\left\{-(1+\alpha)k_j(v)\left[\zeta - \frac{\pi}{4k_j(v)}\right]\right\}$$

in the head of the vortex ($\zeta > \pi/4k_j(v)$); in the middle part of the vortex, a solution in the second allowed velocity band is obtained from (3.20) and (3.21) by changing the signs of the last terms.

APPENDIX 4

If dissipation is sufficiently small, one can obtain the following approximate expression for $f(\zeta)$ from (3.4):

$$f(\zeta) \approx -S_w \frac{V_{sw}^2}{V_{sw}^2 - v^2}[\psi'(\zeta) - k_w \psi(\zeta)]. \qquad \text{(A.4.1)}$$

Substituting (A4.1) into (3.3), by analogy with the dissipationless case, we obtain the following equation for $\psi(\zeta)$:

$$F[\psi] - \frac{2}{\pi} k_j^{-2}(v)\psi''(\zeta) = \frac{j}{j_c} + \frac{v\beta(v)}{\omega_j^2}\psi'(\zeta). \qquad \text{(A.4.2)}$$

When $F[\psi]$ corresponds to the nonlinearity in the Sakai–Tateno–Pedersen model, Eq. (A4.2) differs from that considered in [7] only in that the coefficients $k_j(v)$ and $\beta(v)$ depend on velocity. The similarity between Eq. (A4.2) and Eq. (2.1) from [7] allows one to use the results of [7] and write the dependence (3.16) $j(v)$ to a linear approximation in dissipation.

If $F[\psi] = \sin\psi$, then the formal similarity between Eq. (A4.2) and Eq. (2.5) from [15] allows one to use the results of [15] and write out the dependence (3.24).

REFERENCES

1. S. Sakai and H. Tateno, Jpn. J. Appl. Phys. **22**, 1374 (1983).
2. S. Sakai and N. F. Pedersen, Phys. Rev. B **34**, 3506 (1986).
3. S. Sakai, Phys. Rev. B **36**, 812 (1987).
4. A. F. Volkov, Physica C (Amsterdam) **183**, 177 (1991).
5. A. F. Volkov, Physica C (Amsterdam) **192**, 306 (1992).
6. V. M. Eleonskiĭ and N. E. Kulagin, Zh. Éksp. Teor. Fiz. **119**, 971 (2001) [JETP **92**, 844 (2001)].
7. V. P. Silin and A. V. Studenov, Zh. Éksp. Teor. Fiz. **117**, 1230 (2000) [JETP **90**, 1071 (2000)].
8. A. S. Malishevskii, V. P. Silin, and S. A. Uryupin, Phys. Lett. A **253**, 333 (1999).
9. A. S. Malishevskiĭ, V. P. Silin, and S. A. Uryupin, Zh. Éksp. Teor. Fiz. **117**, 771 (2000) [JETP **90**, 671 (2000)].
10. A. S. Malishevskiĭ, V. P. Silin, and A. V. Studenov, Kratk. Soobshch. Fiz., No. 12, 3 (2000).
11. V. V. Kurin and A. V. Yulin, Phys. Rev. B **55**, 11659 (1997).
12. A. S. Malishevskii, V. P. Silin, and S. A. Uryupin, Phys. Lett. A **306**, 153 (2002).
13. A. A. Abrikosov, *Fundamentals of the Theory of Metals* (Nauka, Moscow, 1987; North-Holland, Amsterdam, 1988).
14. V. P. Silin and S. A. Uryupin, Zh. Éksp. Teor. Fiz. **108**, 2163 (1995) [JETP **81**, 1179 (1995)].
15. D. W. McLaughlin and A. C. Scott, Phys. Rev. B **18**, 1652 (1978).

Translated by I. Nikitin

Journal of Experimental and Theoretical Physics, Vol. 98, No. 3, 2004, pp. 605–611.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 125, No. 3, 2004, pp. 684–692.

SOLIDS
Electronic Properties

Dynamic Control of Electronic States in a Double Quantum Dot under Weak Dissipation Conditions

V. A. Burdov* and D. S. Solenov
Lobachevskiĭ State University, Nizhni Novgorod, 603950 Russia
**e-mail: burdov@phys.unn.ru*
Received August 13, 2003

Abstract—The problem of the time evolution of an electron wave packet in a symmetric double quantum dot under the action of a strong alternating electric field and a slowly varying bias voltage is solved theoretically under the conditions when the electron subsystem can transfer its energy to a single resonator mode. It is shown that the possibility of energy exchange between the electron subsystem and the resonator does not hamper the formation of stable electronic states localized in the left or right quantum dot (i.e., polarized states possessing a positive or negative dipole moment). An adiabatic change in the bias voltage may alter the direction of the dipole moment of the given state (which corresponds to an electron transition from one quantum dot to the other). © *2004 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

Considerable advances in nanotechnologies necessitate the development of new devices operating in the single-electron mode. In such devices, it is essential to control the wave function of an individual electron using, for example, alternating external electromagnetic fields. The control of charge or spin dynamics of individual electrons by external alternating fields can be referred to as the dynamic control of electronic states.

Suitable objects for investigations in this field include structures consisting of several tunnel-coupled quantum dots. In particular, the possibility of creating electronic states localized within a quantum dot and subsequent controllable transfer of such a localized state to a neighboring quantum dot or well was considered earlier as applied to double quantum dots [1] and wells [2–4] as well as lattices of quantum wells or dots [5].

Generally speaking, the studies in this field began with the works on the dynamic localization effect in lattices of quantum wells and in double quantum wells [6–9]. It was shown in these publications that in order to "lock" the electron density in a quantum well, a strong alternating field with a definite relation between the amplitude and frequency is required as a rule. For this reason, a strong alternating field with a slowly varying amplitude [2, 3] was proposed in [1–5] for controlling an electron wave packet for its transfer from one site to a neighboring site; alternatively, in addition to a strong harmonic field with a constant amplitude, it was proposed that an adiabatically varying bias voltage be applied [1, 4, 5].

It should be emphasized that the results obtained in the above-mentioned publications corresponded to a completely coherent mode of evolution of a wave packet. Dissipative processes associated with the interaction of the electron subsystem with radiation fields were completely disregarded. This study aims at theoretical analysis of the effect of weak dissipation on the possibility of a dynamic control of electronic state in a double quantum dot. For this purpose, we will consider a model in which an electron located in a double quantum dot interacts with the resonator and can emit energy into one of the resonator modes (in actual practice, this can correspond, for example, to the phonon mode of an impurity center in a quantum dot).

2. BASIC EQUATIONS OF THE PROBLEM

Let us consider a symmetric double quantum dot in a slowly and monotonically varying electric field $E(t)$ and in an alternating electric field $F\cos\omega t$ with a constant amplitude. We assume that, in zero field, the splitting energy $\hbar\Delta$ for the ground level of size quantization in the double quantum dot (splitting is due to weak tunnel coupling between the dots) is much lower than the size quantization energy. The wave functions $\chi_{0,1}(\mathbf{r})$ of the stationary states of two lower levels with energies $\pm\hbar\Delta/2$ are symmetric and antisymmetric, respectively, to the sign reversal of coordinate z (we assume that the centers of the quantum dots lie on the z axis at a distance of $\pm L$ from the origin).

We assume that an electron in the quantum dot can exchange energy with the resonator. For simplicity, we will consider energy radiation to a single resonator mode, which will be simulated here (following [10]) by

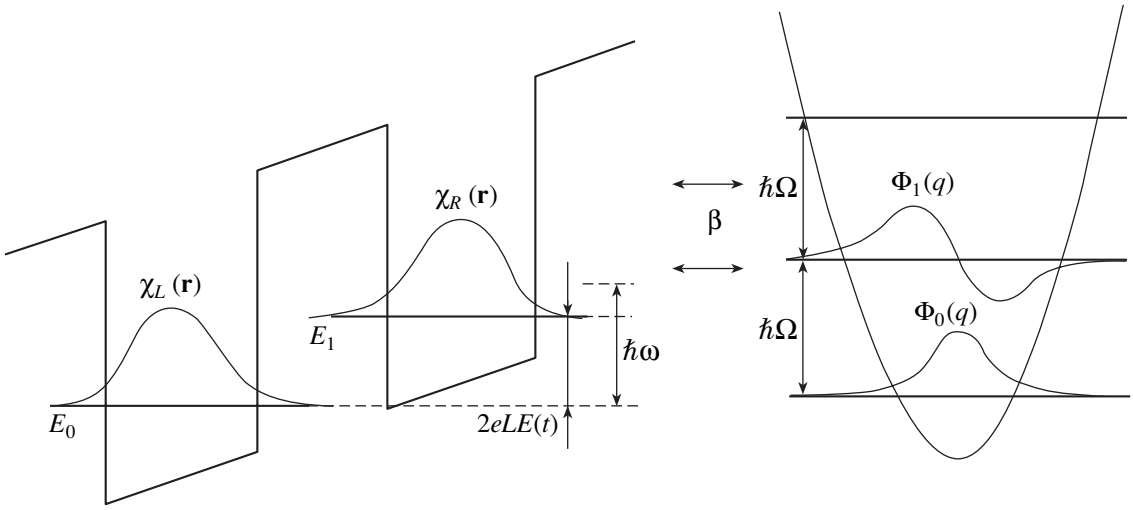

Fig. 1. Energy diagram of a two-level double quantum dot interacting with an oscillator of frequency Ω.

a harmonic oscillator of mass m and frequency Ω (see Fig. 1). The operator of the interaction is linear in displacement q of the oscillator from the equilibrium position and is proportional to the electron dipole moment $D(z)$ in the double quantum dot (i.e., this operator is an antisymmetric function of coordinate z).

The total Hamiltonian of the system has the form

$$\hat{H} = \hat{H}_e(\mathbf{r}) + \hat{H}_v(q) + ez(E(t) + F\cos\omega t) + CqD(z), \quad (1)$$

where $\hat{H}_e(\mathbf{r})$ is the electron Hamiltonian, $\hat{H}_v(q)$ is the Hamiltonian of the harmonic oscillator of frequency Ω, C is the constant of interaction of the electron with the oscillator field, and $-e$ is the electron charge. We assume that fields $E(t)$ and F are quite strong in the sense that the characteristic value of the electron potential energy in the electric field is much higher than the splitting energy (eLF, $eLE(t) \gg \hbar\Delta$). At the same time, we assume that the values of eLF and $eLE(t)$ are much smaller than the characteristic energy of size quantization in the quantum dot, which is of the same order of magnitude as the potential barrier height. The latter condition enables us to disregard the probability of system excitation to the upper levels. As a result, the electron quantum dynamics essentially affects only two lower levels with energies $\pm\hbar\Delta/2$. These assumptions enable us to use the two-level approximation for the electron Hamiltonian.

In order to determine the wave function $\Psi(q, \mathbf{r}, t)$ of the system, we must solve the Schrödinger equation

$$i\hbar\frac{\partial\Psi}{\partial t} = \hat{H}\Psi \quad (2)$$

with Hamiltonian (1). We will seek the wave function of the system in the form of the following expansion:

$$\Psi(q, \mathbf{r}, t)$$
$$= \chi_L(\mathbf{r})\exp\left\{i\frac{EL}{\hbar}\int E(t)dt + i\frac{eLF}{\hbar\omega}\sin\omega t\right\}$$
$$\times\sum_{n=0}^{\infty} l_n(t)\Phi_n(q)\exp\left\{-i\Omega\left(n+\frac{1}{2}\right)t\right\} \quad (3)$$
$$+\chi_R(\mathbf{r})\exp\left\{-i\frac{eL}{\hbar}\int E(t)dt - i\frac{eLF}{\hbar\omega}\sin\omega t\right\}$$
$$\times\sum_{n=0}^{\infty} r_n(t)\Phi_n(q)\exp\left\{-i\Omega\left(n+\frac{1}{2}\right)t\right\}.$$

Here, $l_n(t)$ and $r_n(t)$ are unknown expansion coefficients, $\Phi_n(q)$ are the stationary eigenfunctions of the oscillator Hamiltonian, and $\chi_{L,R}(\mathbf{r})$ are orthonormal functions connected with $\chi_{0,1}(\mathbf{r})$ via the relations

$$\chi_L(\mathbf{r}) = \frac{\chi_0(\mathbf{r}) - \chi_1(\mathbf{r})}{\sqrt{2}}, \quad \chi_R(\mathbf{r}) = \frac{\chi_0(\mathbf{r}) + \chi_1(\mathbf{r})}{\sqrt{2}}.$$

The expansion in basis $\chi_{L,R}(\mathbf{r})$ is more convenient for us than the expansion in basis $\chi_{0,1}(\mathbf{r})$ of stationary states since functions $\chi_{L,R}(\mathbf{r})$ are almost completely localized in the left and right quantum dots, respectively. Consequently, the squares of the moduli of coefficients l_n and r_n are the joint probabilities of an electron being in the left or right quantum dot and of the oscillator being in the nth state.

The substitution of wave function (3) into Schrödinger equation (2) leads to equations of the form

$$i\frac{dl_n}{dt} = -\frac{\Delta}{2}r_n\exp\left\{-i\frac{2eL}{\hbar}\int E(t)dt - i\frac{2eLF}{\hbar\omega}\sin\omega t\right\}$$

$$-\beta\omega(\sqrt{n}l_{n-1}\exp\{i\Omega t\} + \sqrt{n+1}l_{n+1}\exp\{-i\Omega t\}),$$

$$i\frac{dr_n}{dt} = -\frac{\Delta}{2}l_n\exp\left\{i\frac{2eL}{\hbar}\int E(t)dt + i\frac{2eLF}{\hbar\omega}\sin\omega t\right\} \quad (4)$$

$$+\beta\omega(\sqrt{n}r_{n-1}\exp\{i\Omega t\} + \sqrt{n+1}r_{n+1}\exp\{-i\Omega t\}),$$

where the dimensionless coupling parameter

$$\beta = C\sqrt{\frac{\hbar}{2m\Omega}}\frac{\langle\chi_R|D(z)|\chi_R\rangle}{\hbar\omega}$$

is introduced. In all subsequent calculations, we will assume that $\beta \ll 1$, which indicates a weak coupling of the electron subsystem with the resonator.

Let us suppose that field $E(t)$ varies with time according to the law

$$E(t) = E_0(1 + \mu\omega t), \quad (5)$$

where the dimensionless parameter μ characterizes the rate of variation of field $E(t)$ and is small. The $E(t)$ dependence in form (5) is naturally not the single possible dependence. We choose this dependence as the simplest from an infinite set of various monotonic dependences.

Note that in the case when an electron experiences the action of a constant field alone (i.e., for $\mu = F = 0$), the two-well potential becomes asymmetric, the energy levels in the double quantum dot diverge by $\sqrt{\hbar^2\Delta^2 + 4e^2L^2E_0^2}$, and the wave functions of stationary states are localized in each potential well separately. In the limit of a strong field, the quantum transition energy approaches a value of $2eLE_0$ and the wave functions corresponding to each of these energy levels become equal to $\chi_{L,R}(\mathbf{r})$ as shown in Fig. 1. If we now "switch on" an alternating field of frequency ω, the quantum system can attain resonance when

$$2eLE_0 = \hbar\omega s, \quad (6)$$

where s is an arbitrary integer.

For $\mu \neq 0$, with an appropriate choice of the time origin, we can always make the constant component of field $E(t)$ exactly satisfy the resonance condition, e.g., with $s = 1$. Assuming that instant $t = 0$ is chosen precisely in this way, we will consider, without loss of generality, the fact that E_0 satisfies the resonance condition with $s = 1$. Then, Eqs. (4) can be written in the following form, where the resonance terms with different values of s are explicitly separated:

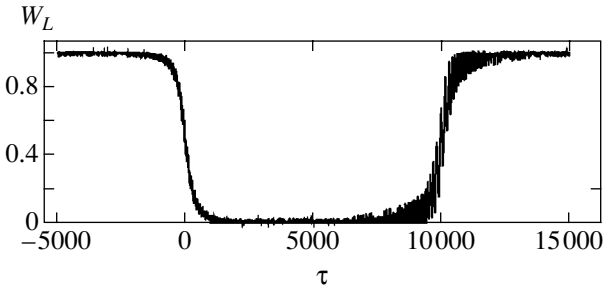

Fig. 2. Probability of population of the left quantum dot for $\beta = 0$, $\lambda = 7.7$, $\mu = 10^{-4}$, and $\Delta/\omega = 0.2$.

$$i\frac{dl_n}{d\tau} = -r_n\exp\left\{-\frac{i\mu\tau^2}{2}\right\}\sum_{s=-\infty}^{\infty}\delta_s\exp\{-i(s+1)\tau\}$$

$$-\beta(\sqrt{n}l_{n-1}\exp\{i(1+\eta)\tau\}$$

$$+\sqrt{n+1}l_{n+1}\exp\{-i(1+\eta)\tau\}),$$

$$i\frac{dr_n}{d\tau} = -l_n\exp\left\{\frac{i\mu\tau^2}{2}\right\}\sum_{s=-\infty}^{\infty}\delta_s\exp\{i(s+1)\tau\} \quad (7)$$

$$+\beta(\sqrt{n}r_{n-1}\exp\{i(1+\eta)\tau\}$$

$$+\sqrt{n+1}r_{n+1}\exp\{-i(1+\eta)\tau\}).$$

Here, we introduced the dimensionless time $\tau = \omega t$; parameter $\eta = \Omega/\omega - 1$ characterizes the closeness of the oscillator frequency to the resonance value, $\delta_s = J_s(\lambda)\Delta/2\omega \ll 1$, and $\lambda = 2eLF/\hbar\omega$.

In the subsequent numerical calculations (whose results are presented in Figs. 2–6), we assume everywhere that $\lambda = 7.7$, $\mu = 10^{-4}$, and $\Delta/\omega = 0.2$. Choosing energy $\hbar\Omega = 10^{-2}$ eV (which corresponds to the characteristic values of phonon frequencies) and distance $L = 10$ nm and taking into account the fact that frequency ω is close to oscillator frequency Ω, we obtain a value of the alternating field amplitude F approximately equal to 40 kV/cm. The resonance value of field E_0 in this case is close to 5 kV/cm, which corresponds to bias voltages in the structure equal to 0.1–0.01 V. The splitting energy $\hbar\Delta$ is found to be 2 meV; i.e., it is the smallest energy scale of the problem, which was presumed earlier. It should be noted that the chosen value of parameter μ corresponds to the increase in field $E(t)$ from the first resonance value ($s = 1$) to the second one ($s = 2$) over a time approximately equal to 10^{-9} s. As regards the coupling parameter β, its value is determined by a specific type of interaction and may vary over a wide interval (from nearly indefinitely small values to several units).

For $\mu = 0$, the two-level electron system can be in exact resonance with the alternating electric field for an infinitely long time. For $\mu \neq 0$, for each instant, we can introduce the concept of instantaneous energy of level splitting, which is approximately equal to $2eLE(t)$ in a strong field. In this case, the quantum system is effectively close to resonance (when the action of the external field is the strongest) for a certain finite time interval determined by expression (5) and the relation between parameters μ and δ. After this time, the two-level system deviates from resonance and the type of the action of the alternating electric field on this system changes.

We will now consider the solution of system (7). We assume that, at the initial instant (here, this instant should be a certain time $\tau = -\tau_0$ exceeding all characteristic time scales of the problem, but not large enough for the field $E(-\tau_0)$ to be negative), the oscillator is not excited and the left quantum dot is populated, while the right dot is absolutely free; i.e., we assume that

$$l_0(-\tau_0) = 1, \quad l_n(-\tau_0) = 0, \quad n \geq 1, \qquad r_n(-\tau_0) = 0, \quad n \geq 0. \tag{8}$$

Obviously, initial condition (8) corresponds to the minimum energy of the electron + oscillator system.

3. DYNAMICS OF POPULATION OF QUANTUM STATES OF THE SYSTEM

In the simplest case, when there is no coupling between the double quantum dot and the resonator ($\beta = 0$), system (7) splits into independent pairs of equations for amplitudes $l_n(\tau)$ and $r_n(\tau)$ at each level of the oscillator. To find the solution in the vicinity of resonance, it is sufficient to retain only one term with $s = -1$ in each sum over s in Eqs. (7). All these pairs of equations are completely identical and can be reduced to a single second-order equation, say, for coefficient $l_n(\tau)$,

$$\frac{d^2 l_n}{d\tau^2} + i\mu\tau\frac{dl_n}{d\tau} + \delta^2 l_n = 0 \tag{9}$$

(we have omitted index –1 on δ).

An approximate solution to Eq. (9) was obtained earlier in [4] by using the WKB method; here, we give only the final expression for the total probability of filling for the left quantum dot:

$$W_L(\tau) = \frac{\sqrt{4\delta^2 + \mu^2\tau^2} - \mu\tau}{2\sqrt{4\delta^2 + \mu^2\tau^2}}. \tag{10}$$

While deriving this relation, we chose the value of τ_0 equal to infinity since Eq. (9) was derived in the single-resonance approximation. Expression (10) describes the transfer of the electron density from the left to the right quantum dot during an adiabatic transition of the two-level electron system through a resonance.

Thus, for $\tau < 0$, the separation between two lower levels of the double quantum dot (the instantaneous value of this quantity is $2eLE(t)$) is found to be smaller than energy $\hbar\omega$ of an alternating field quantum. The larger the value of $|\tau|$, the farther the two-level system from resonance and the higher the probability (10) of occupation of the left quantum dot. With increasing time, the separation between the energy levels of the double quantum dot increases and, for $\tau = 0$, this separation coincides with $\hbar\omega$, which means that the two-level system is in strict resonance (6). In this case, probability W_L is equal to 1/2. A further increase in τ leads to further divergence of the energy levels and to a gradual withdrawal of the system from resonance. This process is accompanied by a decrease in probability W_L to zero and virtually complete relocation of the electron density to the right quantum dot. The characteristic transition time is determined by ratio δ/μ, which amounts to a value on the order of 10^{-10} s in dimensional units.

Note that the exact solution to Eq. (9) represented in Fig. 2 slightly differs from the approximate solution described by expression (10). This is due to the following two circumstances. First, the numerical calculations were made for a wider range of time τ embracing two neighboring resonances with $s = 1$ and $s = 2$, while solution (10) is valid only in the vicinity of one resonance. Second, as can be seen from Fig. 2, probability $W_L(\tau)$ rapidly oscillates, although the amplitude of these oscillations is small. Obviously, expression (10) describing a smooth decreasing function is the result of averaging of the exact solution over these rapid oscillations in the region of the first resonance.

The oscillations observed in the exact solution are in fact analogous to Rabi oscillations, which now occur in a strong alternating field. If field $E(t)$ were strictly constant, the solution to Eq. (9) would have the form $l_0(\tau) = \cos\delta\tau$. In accordance with this solution, the electron wave packet would oscillate between the two quantum dots with constant frequency δ, which has the meaning of the Rabi frequency in a strong field (since the field is strong, frequency δ is not linear in field amplitude F). However, in view of variability of field $E(t)$, the cosinusoidal dependence is observed only for times $|\tau| \lesssim \delta^2/\mu$; for chosen values of δ and μ (see the caption to Fig. 2), the value of this time is on the order of unity. For $\tau > \delta^2/\mu$, the second term that "quenches" the amplitude of oscillations dominates in Eq. (9). Since the value of δ^2/μ if found to be much smaller than the period $2\pi/\delta$ of Rabi oscillations, not a single oscillation with complete transfer of charge from the left to the right quantum dot is observed.

Let us now consider a more interesting situation when the electron subsystem interacts with the resonator field; i.e., $\beta \neq 0$. The interaction with the resonator can obviously break the dynamic control in the system since it becomes possible for the electron located, say, in the right quantum dot in a state disadvantageous from the

energy point of view to transfer its excess energy to the resonator. As a result, the resonator is excited and leaves its ground state, while the electronic state that initially possessed a certain polarization can be destroyed.

In order to find out whether it is possible to preserve the dynamic control of the electron distribution in a double quantum dot under the conditions of energy exchange with the resonator, we must solve system (7) for $\beta \neq 0$. In this case, system (7) obviously cannot be represented in a closed form and can be solved only numerically. At a certain step, it is necessary to truncate Eqs. (7), which corresponds to the replacement of the oscillator by an N-level system with an equidistant spectrum.

Calculations show that the number of retained energy levels of the oscillator and the solutions to Eqs. (7) themselves differ significantly for different values of parameters η and β. Further, we consider two basically different cases, i.e., the nonresonance case, when the frequency of the oscillator significantly differs from the frequency of the external alternating field (i.e., $\eta \neq 0$), and the resonance case, when $\eta = 0$.

The main problem is to evaluate the total (i.e., summed over all states of the oscillator) probability of population of any of quantum dots (e.g., left) as a function of time, which is now defined as

$$W_L(\tau) = \sum_{n=0}^{\infty} |l_n(\tau)|^2. \quad (11)$$

Further, we calculate and discuss the behavior of individual probabilities $|l_n(\tau)|^2$ and $|r_n(\tau)|^2$, which are essential for determining the sum in Eq. (11).

Let us begin with the nonresonance case. Figure 3 shows the probability of exciting the oscillator to the first level for $\eta = 0.1$. It can be seen from the figure that it is as though probabilities $|l_1(\tau)|^2$ and $|r_1(\tau)|^2$ alternate on the τ axis: when one probability decreases, the other starts to increase, and vice versa. As regards the oscillator, it is excited most strongly when the two-level electron system is in the vicinity of the second resonance for $\tau \sim 10^4$. Outside this neighborhood, probabilities $|l_n(\tau)|^2$ and $|r_n(\tau)|^2$ decrease by one or two orders of magnitude. However, the excitation probability does not exceed a few thousandths in the vicinity of the resonance as well. Calculations show that the probabilities of excitation of the oscillator to higher levels have values rapidly decreasing with increasing level number. For example, the probabilities of excitation of the oscillator to the second level turn out to be lower than the probabilities represented in Fig. 3 by 3–4 orders of magnitude. Consequently, the oscillator is almost always unexcited and, hence, hardly affects the electron distribution in quantum dots.

Calculating the probabilities in the nonresonance case ($\eta = 0.1$), we replaced the oscillator by a four-level system; i.e., $0 \le n \le 3$. The addition of the fifth level leads to corrections on the order of 10^{-10} to the values of probabilities, which is beyond the required accuracy limits.

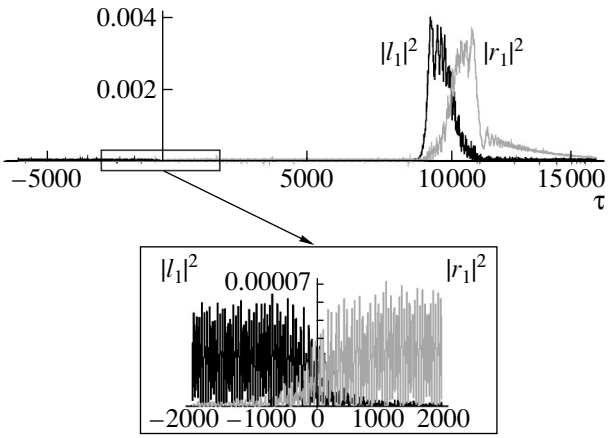

Fig. 3. Probability of excitation of the oscillator to the first level for $\eta = 0.1$, $\beta = 0.01$, $\lambda = 7.7$, $\mu = 10^{-4}$, and $\Delta/\omega = 0.2$.

We also calculated the total probability (11) of populating the left quantum dot with an electron; in the nonresonance case, this probability coincides with $|l_0(\tau)|^2$ to within fractions of percent (see Fig. 3). Calculations proved that the difference of the probabilities

$$\Delta W_L(\tau) = W_L(\tau, \beta = 0.01) - W_L(\tau, \beta = 0)$$

remains small everywhere and amounts to $\sim 10^{-4}$. It can be proved analytically that this difference in the nonresonance case must be determined by the second power of coupling parameter β, which is confirmed by the results of calculations.

Consequently, the dynamics of the electron density remains the same as in the absence of coupling with the resonator. Upon each passage of the two-level system through the resonance, the double quantum dot reverses the direction of polarization since the electron density performs transition from one quantum dot to the other.

Let us now consider the resonance situation, when $\eta = 0$. Figure 4 shows the joint probabilities of populating one of the quantum dots by an electron and of the oscillator being in one of its states for $\beta = 0.01$ and $\lambda = 7.7$. It can be seen from Fig. 4 that, after the transition from the first resonance $s = 1$, probabilities $|l_n(\tau)|^2$ decrease to values close to zero until the two-level system finds itself in the vicinity of the next resonance with $s = 2$. Conversely, probabilities $|r_n(\tau)|^2$ increase in this range of τ from almost zero values to a few tenths for several first energy levels of the oscillator.

After the passage through the second resonance, the probabilities of populating the right quantum dot decrease to zero for all levels of the oscillator, while probabilities $|l_n(\tau)|^2$ increase. Upon the passage through each next resonance, an increase or a decrease in the probabilities obviously changes to a decrease or an increase, respectively. The following tendency can be

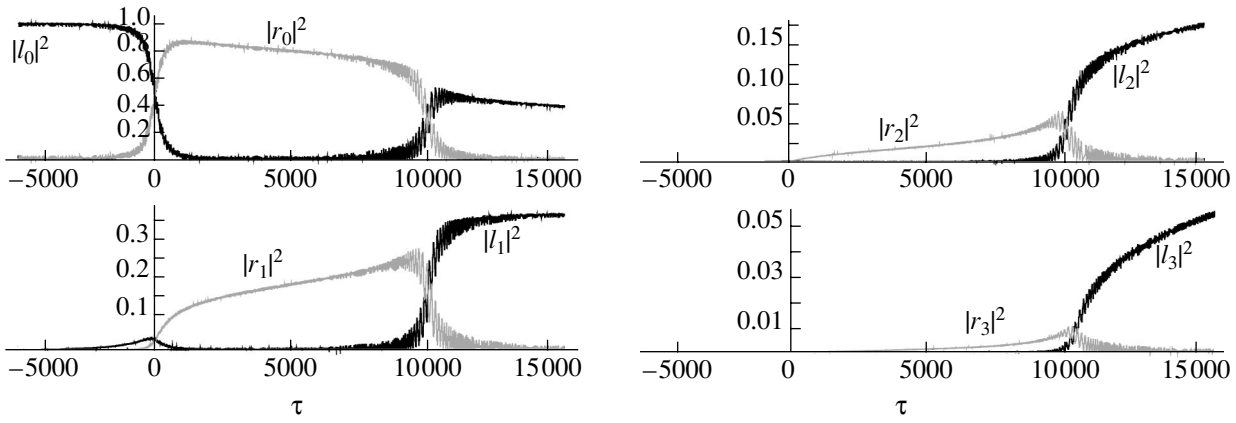

Fig. 4. Dynamics of excitation of the oscillator in the resonance case ($\eta = 0$) for $\beta = 0.01$, $\lambda = 7.7$, $\mu = 10^{-4}$, and $\Delta/\omega = 0.2$.

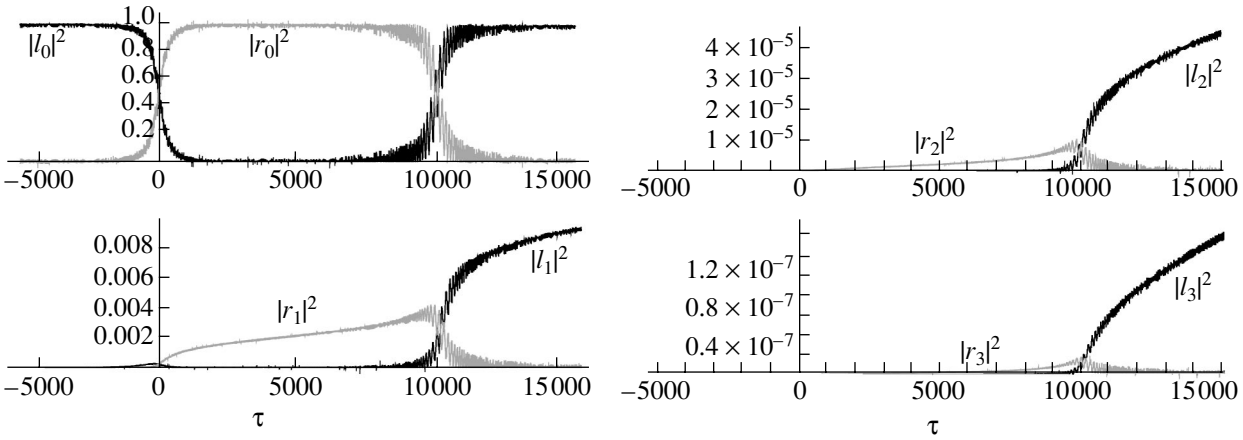

Fig. 5. Dynamics of excitation of the oscillator for $\beta = 0.001$. The values of the remaining parameters are the same as in Fig. 3.

traced: the longer the time of interaction between the electron subsystem and the resonator, the higher the degree of resonator excitation; i.e., higher and higher energy levels of the oscillator become involved in the process of energy exchange between the resonator and the external alternating field. For this reason, the number N of oscillator levels that should be retained in the calculations obviously depends on the time interval during which we are going to observe the system. For instance, for the time interval $-5000 < \tau < 15000$ over which our calculations were made (see Fig. 4), the number of energy levels of the oscillator was $N = 8$.

Figure 5 illustrates the case of a weaker coupling ($\beta = 10^{-3}$ for $\lambda = 7.7$). It can be seen that the general form of the dependence of probabilities $|l_n|^2$ and $|r_n|^2$ on time τ does not change as compared to the case when $\beta = 0.01$. As before, the oscillator is gradually excited with time to its higher levels, but the process of its excitation occurs much more slowly in view of a weaker coupling with the electron subsystem. For this reason, in the interval of time variation shown in Fig. 5, coefficients $l_0(\tau)$ and $r_0(\tau)$ play the major role, while all remaining amplitudes still have no time to increase appreciably. In the region $-5000 < \tau < 15000$, the values of probabilities $|l_n(\tau)|^2$ and $|r_n(\tau)|^2$ for $\beta = 0.001$ differ insignificantly from their values in the case of complete absence of coupling. However, with increasing τ, these differences gradually accumulate and the behavior of probabilities for $\beta = 0.01$ qualitatively repeats the dependences shown in Fig. 4 for $\beta = 0.01$.

As in the nonresonance case, we calculate the total probability $W_L(\tau)$ by using formula (11). The results of calculation for $\beta = 0.01$ are shown in Fig. 6, where the probability difference $\Delta W_L(\tau)$ is depicted. It can be seen that the difference is extremely small in region $\tau < 0$ (up to resonance region), constituting a small fraction of a percent of the probability $W_L(\tau)$ proper, which is shown in Fig. 2. In the vicinity of the resonance, the difference noticeably increases and reaches several percent. Further, in region $\tau > 0$, as we move away from the resonance, the value of $W_L(\tau)$ itself is insignificant, while difference $\Delta W_L(\tau)$, whose absolute value slightly decreases in this region, nevertheless becomes of the order of probability $W_L(\tau)$ itself.

In the region of the second resonance, a noticeable "spike" of $\Delta W_L(\tau)$ is observed (although the probability

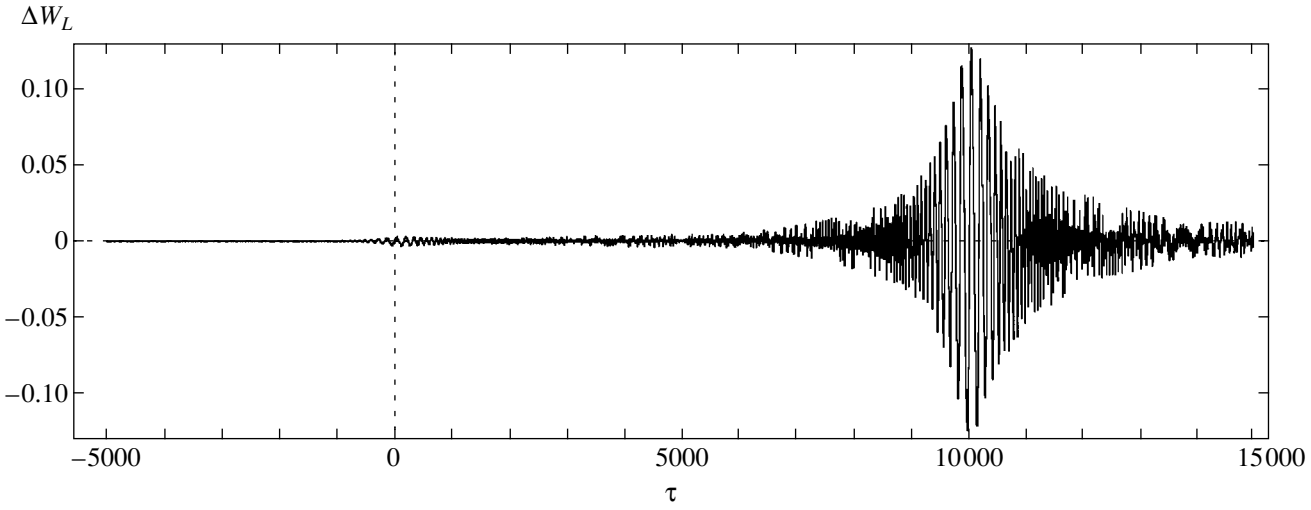

Fig. 6. Variation of population probability for the left quantum dot, associated with the interaction between the electron subsystem and the resonator, in the resonance case for $\lambda = 7.7$, $\mu = 10^{-4}$, and $\Delta/\omega = 0.2$.

$W_L(\tau)$ itself sharply increases in this region), after which the difference decreases again. On the whole, as follows from Fig. 6, the deviation of probability $W_L(\tau)$ from its value calculated for $\beta = 0$ does not exceed 0.1 in absolute value. It should be noted that, in contrast to the nonresonance case, the difference is two orders of magnitude larger and is apparently determined by the first power of coupling parameter β.

It should be noted that the value of difference $\Delta W_L(\tau)$ for $\beta = 0.001$ turns out to be one or two orders of magnitude smaller than for $\beta = 0.01$. For this reason, dependence $W_L(\tau)$ repeats the dependence shown in Fig. 2 even to a higher degree of accuracy.

Thus, probability $W_L(\tau)$ of populating the left quantum dot for various values of parameter β remains virtually the same as in the absence of coupling both in the nonresonance ($\eta = 0$) and in the resonance ($\eta = 0.1$) cases. Such a behavior indicates that the electron density, as before, performs transitions (as in the case when $\beta = 0$) from one quantum dot to the other upon an adiabatic passage of the two-level electron system through resonance. Even in the case when an electron wave packet is localized in the right quantum dot, such an "energetically disadvantageous" electronic state polarized against the external field $E(t)$ is found to be stable. Its stability is due to the "locking" effect of the strong alternating field, which confines the electron density in the right quantum dot despite the possible energy exchange between the electron subsystem and the resonator.

On the whole, it can be concluded that the interaction of the electron subsystem with the resonator does not disturb the dynamic control over electron states. This is mainly due to the fact that the external alternating field in our model possesses an infinitely large energy, whose finite part is transferred via the two-level electron system to the resonator. The state of the "transmitter" itself does not change in this case since any energy losses in it are compensated by the strong alternating field.

ACKNOWLEDGMENTS

The authors thank Yu.A. Romanov and A.M. Satanin for their interest in this research.

This study was financed by the Ministry of Education of the Russian Federation (grant nos. E02-3.1.-336 and UR.01.01.057) and the Russian Foundation for Basic Research (project no. 02-02-17495).

REFERENCES

1. V. A. Burdov, Pis'ma Zh. Éksp. Teor. Fiz. **76**, 25 (2002) [JETP Lett. **76**, 21 (2002)].
2. R. Bavli and H. Metiu, Phys. Rev. A **47**, 3299 (1993).
3. A. A. Gorbatsevich, V. V. Kapaev, and Yu. V. Kopaev, Zh. Éksp. Teor. Fiz. **107**, 1320 (1995) [JETP **80**, 734 (1995)].
4. V. A. Burdov, Zh. Éksp. Teor. Fiz. **116**, 217 (1999) [JETP **89**, 119 (1999)].
5. V. A. Burdov and D. S. Solenov, Phys. Lett. A **305**, 427 (2002).
6. D. H. Dunlap and V. M. Kenkre, Phys. Rev. B **34**, 3625 (1986).
7. S. Raghavan, V. M. Kenkre, D. H. Dunlap, *et al.*, Phys. Rev. A **54**, R1781 (1996).
8. F. Grossmann, T. Dittrich, P. Jung, and P. Hanggi, Phys. Rev. Lett. **67**, 516 (1991); F. Grossmann and P. Hanggi, Europhys. Lett. **18**, 571 (1992).
9. Y. Dakhnovskii and R. Bavli, Phys. Rev. B **48**, 11010 (1993).
10. A. B. Pippard, *Physics of Vibration* (Cambridge Univ. Press, Cambridge, 1978; Vysshaya Shkola, Moscow, 1989).

Translated by N. Wadhwa

Journal of Experimental and Theoretical Physics, Vol. 98, No. 3, 2004, pp. 612–618.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 125, No. 3, 2004, pp. 693–700.

SOLIDS
Electronic Properties

Experimental Evidence of a Two-Phase Magnetic State in Thin Epitaxial Films of $Re_{0.6}Ba_{0.4}MnO_3$ (Re = La, Pr, Nd, Gd)

O. Yu. Gorbenko[a], R. V. Demin[a], A. R. Kaul'[a], L. I. Koroleva[a], R. Szymczak[b], H. Szymczak[b], and M. Baran[b]

[a] *Moscow State University, Moscow, 119992 Russia*
[b] *Institute of Physics, Polish Academy of Sciences, 02668 Warsaw, Poland*
e-mail: koroleva@ofef343.phys.msu.su

Received October 9, 2003

Abstract—Thin epitaxial films of $Re_{0.6}Ba_{0.4}MnO_3$ (Re = La, Pr, Nd, Gd) on (001)-oriented single crystal $SrTiO_3$ and $ZrO_2(Y_2O_3)$ substrates have been prepared and studied. All films possess a cubic perovskite structure, except for the film with Re = La, which exhibited a rhombohedral distortion of the perovskite lattice. The results show evidence for the presence of two magnetic phases, ferromagnetic (FM) and antiferromagnetic (AFM), in the films studied: (i) the magnetization isotherm $M(H)$ appears as a superposition of a linear component (characteristic of antiferromagnets) and a small spontaneous magnetization component; (ii) the magnetic moment per formula unit is significantly reduced as compared to the value expected for the complete FM or ferrimagnetic ordering; (iii) there is a difference between magnetizations of the samples cooled with and without an applied magnetic field, which is preserved in the entire range of magnetic fields studied (50 kOe); (iv) the temperature dependence of the magnetization $M(T)$ in strong magnetic fields is close to linear (for the composition with Re = Gd, $M(T)$ is described by a Langevin function for superparamagnets with a cluster moment of $22\mu_B$); and (v) the magnetization hysteresis loops of the field-cooled samples are shifted along the field axis. The exchange integral (characterizing the Mn–O–Mn coupling via the FM–AFM phase boundary) estimated from the latter shift is $|J| = 10^{-6}$ eV. This value is two orders of magnitude lower than the negative exchange integral between the FM layers in $ReMnO_3$, which makes the presence of a transition layer at the FM–AFM phase boundary unlikely. The temperature dependences of electrical resistance and magnetoresistance exhibit maxima at the Curie temperature (T_C), where the magnetoresistance reaches a colossal value. This behavior indicates that the two-phase magnetic state is caused by a strong *s–d* exchange. © *2004 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

In recent years, much attention has been devoted to manganites of the $Re_{1-x}A_xMnO_3$ system, where Re is a rare earth ion and A = Ca, Sr, or Ba. This interest is related to the phenomenon of colossal magnetoresistance observed in some of these compounds at room temperature. The compounds with A = Sr and Ca were studied most thoroughly, while $Re_{1-x}Ba_xMnO_3$ compositions were characterized to a much lower extent, especially in the case of thin films. Only thin films of the $La_{1-x}Ba_xMnO_3$ system with $x = 0.2$ and 0.33 were studied [1–3], in which the room-temperature magnetoresistance R_0/R_H reached about 50% at $H = 0.8$ and 5 T for the former and latter composition, respectively. It was also reported [4] that the magnetoresistance of the compound with $x = 0.33$ strongly depends on the degree of nonstoichiometry with respect to oxygen.

Compounds of the $La_{1-x}Ba_xMnO_3$ system are of considerable interest because of very high Curie temperatures: $T_C = 362$ was observed for $x = 0.3$ [5–7]. Such a high T_C value is due to a relatively large average radius $\langle r_A \rangle$ of A cations, since it is known that the Curie temperature of manganites of the ABO_3 type increases with $\langle r_A \rangle$ value [8, 9]. At the same time, there is an opposite trend (so-called misfit effect) related to a difference between the radii of different A cations (Re^{3+} and Ba^{2+}), which decreases the Curie temperature [10].

At present, there is no commonly accepted opinion about the crystal structure of $Re_{1-x}Ba_xMnO_3$ compounds. In $La_{1-x}Ba_xMnO_3$ compositions with $0.2 \le x \le 0.4$, Radaelli *et al.* observed a hexagonal structure (space group, $R\bar{3}c$), while Barnabe *et al.* [7] reported on a more complicated crystal structure (in agreement with the neutron diffraction data [11]) in stoichiometric compositions of the $Re_{1-x}Ba_xMnO_3$ systems (Re = La, Pr) with $x = 0.4$. At the same time, Raman spectroscopy and X-ray data [12] for polycrystalline samples of the $La_{1-x}Ba_xMnO_3$ system with $x \ge 0.35$ showed evidence of a phase separation into cubic $La_{0.65}Ba_{0.35}MnO_3$ and hexagonal $BaMnO_3$ phases. It was suggested [12] that the structure of compounds with $x \ge 0.35$ cannot accommodate Ba^{2+} ions because of their large size.

We have studied the crystal structure and the magnetic and electrical properties of thin epitaxial films of $Re_{1-x}Ba_xMnO_3$ compounds (Re = La, Pr, Nd, Gd). The observed features have been interpreted within the framework of the modern theory of magnetic semiconductors. As was noted above, only films of the $La_{1-x}Ba_xMnO_3$ system were reported previously [1–3]. The films of $Re_{1-x}Ba_xMnO_3$ with other rare earth elements have been prepared and characterized for the first time.

2. EXPERIMENTAL METHODS

All films were grown by metalorganic chemical vapor deposition (MOCVD) using aerosols of volatile organometallic compounds. The aerosols of diglyme solutions of the initial compounds with a total concentration of 0.02 mol/l were produced by an ultrasonic source. The initial compounds for MOCVD were $R(thd)_3$ (R = La, Pr, Nd, Gd), $Mn(thd)_3$, and $Ba(thd)_2(Phen)_2$, where thd = 2,2,6,6-tetramethylheptane-3,5-dionate and Phen = *o*-phenanthroline. The samples were prepared in a reactor with an induction heated substrate holder at a substrate temperature of 800°C, an oxygen partial pressure of 3 mbar, and a total pressure of 6 mbar. The deposition rate was 1 μm/h and the film thicknesses ranged within 300–400 nm. The films were deposited onto (001)-oriented single crystal $SrTiO_3$ and $ZrO_2(Y_2O_3)$ substrates.

The films were studied by scanning electron microscopy (in combination with electron probe microanalysis) and by X-ray diffraction. The magnetization of thin films was measured with a SQUID magnetometer and the electric resistance was determined by the conventional four-point-probe technique.

3. EXPERIMENTAL RESULTS AND DISCUSSION

3.1. Structural Characteristics of Thin Films

According to the X-ray diffraction data, the films obtained on (001)-oriented $SrTiO_3$ substrates represented single-phase perovskites epitaxially grown in the "cube over cube" mode. The pseudocubic lattice parameter of the perovskite phase monotonically decreased with the ion radius of Re^{3+}. Only the X-ray diffraction pattern of the film of $La_{0.6}Ba_{0.4}MnO_3$ exhibited weak superstructural reflections and splitting of the pseudocubic reflections expected for a rhombohedral distortion of the perovskite lattice (space group, $R\bar{3}c$), in accordance with the behavior of this material in the ceramic state [7]. In all other $Re_{0.6}Ba_{0.4}MnO_3$ films, we observed neither peak splitting nor superstructural reflections characteristic of the rhombohedral, tetragonal, or orthorhombic distortions encountered in the perovskite structures of rare earth manganites described in the literature. The increase in symmetry up to the cubic type can be explained by an increase in the parameter of disorder with decreasing radius of the rare earth ion (and, accordingly, by an increase in the difference of ion radii of the rare earth element and barium statistically occupying the A sublattice sites in the perovskite structure). This effect should be most pronounced in compositions with the level of barium doping approaching 0.5.

The films grown on (001)-oriented $ZrO_2(Y_2O_3)$ substrates exhibited simultaneous growth in two directions, (001) and (110). In the films of perovskite manganites studied previously, we usually observed only one of these orientations, namely, (110) [13, 14]. The appearance of another type we explain by an increase in the lattice parameter of barium-doped perovskite, which leads to a change in the lattice mismatch between film and substrate.

3.2. Magnetic Properties

Figure 1 shows the temperature dependence of the magnetization $M(T)$ measured in various magnetic fields for films of the PrBaMnO and GdBaMnO systems with $x = 0.4$ on $SrTiO_3$ substrates (below, this subscript will be omitted). The $M(T)$ curves of the LaBaMnO and NdBaMnO films on the $SrTiO_3$ substrates, as well as of the NdBaMnO film on the $ZrO_2(Y_2O_3)$ substrate, were very much like those depicted in Fig. 1a for PrBaMnO films on the $SrTiO_3$ substrates. These magnetization measurements were performed in two modes. The upper curves for each field strength H were obtained by initially cooling a sample from $T = 300$ to 5 K in the given field, after which the temperature variation of the magnetization was measured in the course of heating of this field-cooled (FC) sample. The lower curves were obtained for a sample cooled in the absence of a magnetic field, after which the field was applied and the $M(T)$ curve of the zero-field-cooled (ZFC) sample was measured in the heating mode.

The data in Fig. 1 reveal differences between the temperature dependences of the magnetization for the FC and ZFC samples. The magnetization of FC samples is higher than that of the otherwise identical ZFC samples. This difference increases with decreasing temperature and is more pronounced in lower magnetic fields. The $M(T)$ curves of the ZFC samples studied in small fields exhibit maxima at a certain temperature T_f; at $T > T_f$, the curves of the FC and ZFC samples coincide. In sufficiently large fields, the maximum in the $M(T)$ curves of the ZFC samples is not observed, but a difference between the cures of FC and ZFC samples is retained in the entire range of magnetic fields studied (up to 50 kOe). The only exception was the GdBaMnO film, for which the difference disappeared at $H < 6$ kOe (see the inset to Fig. 1b).

Figure 2 shows the magnetization isotherms measured at various temperatures for the FC samples of

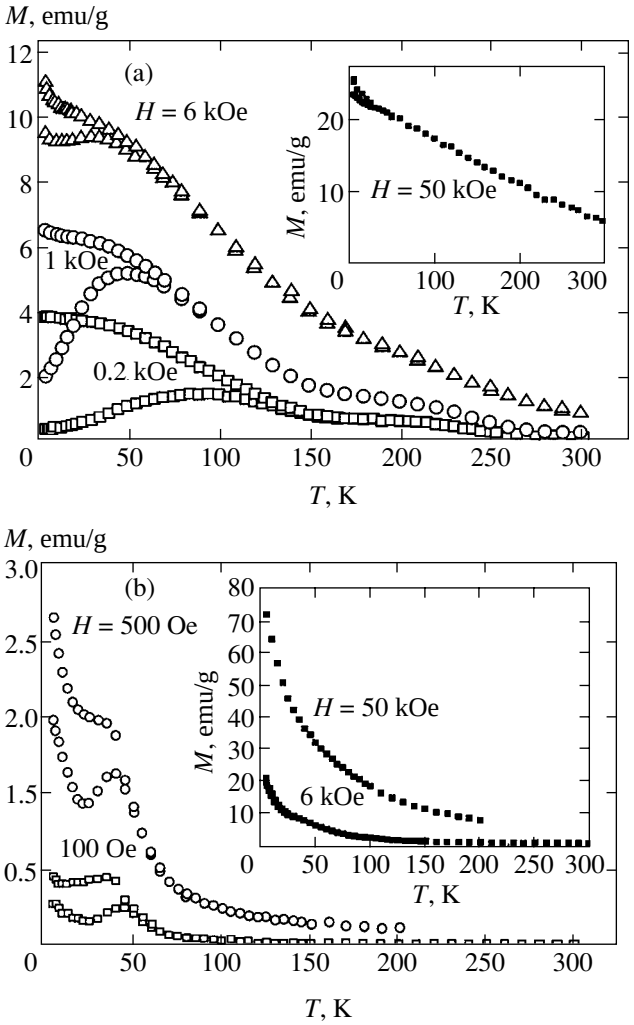

Fig. 1. Temperature dependence of the magnetization $M(T)$ of thin films of $Re_{0.6}Ba_{0.4}MnO_3$ with Re = Pr (a) and Gd (b) on $SrTiO_3$ substrates measured in various magnetic fields. The upper and lower curves for each field strength H refer to the FC and ZFC samples, respectively (see the text).

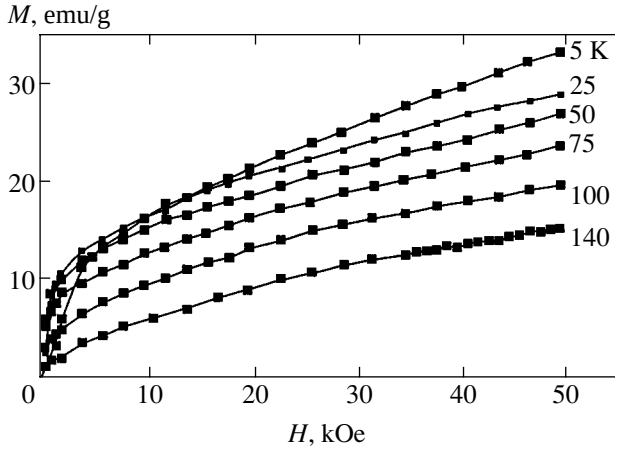

Fig. 2. Magnetization isotherms of $Nd_{0.6}Ba_{0.4}MnO_3$ films on $SrTiO_3$ substrates measured at various temperatures.

NdBaMnO films on $SrTiO_3$ substrates. The magnetization isotherms of the other films studied in our experiments were analogous to those presented in Fig. 2. As can be seen, the $M(H)$ curves appear as superpositions of a small spontaneous magnetization component and a linear component characteristic of antiferromagnets. By extrapolating the linear parts of the $M(H)$ curves to intersection with the M axis, we determined the spontaneous magnetization component and calculated the experimental magnetic moments μ_{exp} (expressed in Bohr magnetons per formula unit, μ_B/FU). The values of μ_{exp} at $T = 5$ K for the films of all studied compositions are given in the table. For comparison, we also present the values of theoretical magnetic moments μ_{th} (μ_B/FU) calculated for three variants of spin ordering for Mn^{3+}, Mn^{4+}, and Re^{3+} ions: the first value refers to FM ordering only between Mn^{3+} and Mn^{4+}; the second, to FM ordering of Mn^{3+}, Mn^{4+}, and Re^{3+}; and the third, to FM ordering of Mn^{3+} and Mn^{4+}, and AFM ordering of Re^{3+}. These values were calculated using the following pure spin moments of the rare earth ions: $2\mu_B$ (Pr^{3+}), $3\mu_B$ (Nd^{3+}), and $7\mu_B$ (Gd^{3+}). As can be seen, the values of μ_{exp} are much lower than μ_{th} in all cases. This experimental fact also provides evidence for the presence of two magnetic phases in the films under consideration.

In the films with Re = La, Pr, and Nd, the behavior of $M(T)$ in strong magnetic fields is close to linear (see Fig. 1a for the PrBaMnO films), which is untypical of ferromagnets. As is known, the temperature dependence of the magnetization in ferromagnets is described by the Brillouin function [15]. Obviously, the exact value of the Curie temperature can be determined only in experiments performed in the absence of external magnetic fields, because such a field suppresses and smears the phase transition. In practice, however, the Curie temperature is frequently determined by extrapolating the steepest (upper) part of the $M(T)$ curve to the temperature axis. In the general case, this yields a certain characteristic temperature T_C' that is close to the Curie temperature. The characteristic temperatures determined in this way for the films studied in our experiments strongly depend on the field applied during the $M(T)$ measurements. The values of T_C' obtained by this method are listed in the table. As can be seen, the T_C' value significantly increases with increasing H. For example, the T_C' values for the LaBaMnO film in a field of 100 Oe is 283 K and that in a field of 50 kOe is 90 K higher, amounting to 373 K. In magnetically homogeneous materials possessing spontaneous magnetization (for example, in ferromagnets) this difference does not exceed 10 K.

As can be seen from Fig. 1b, the film of GdBaMnO exhibits a difference between magnetizations of the FC and ZFC samples only in sufficiently small fields: no such difference is observed for $H = 6$ kOe. The shape of

Magnetic properties of thin epitaxial films of $Re_{0.6}Ba_{0.4}MnO_3$ (Re = La, Pr, Nd, Gd) on (001)-oriented $SrTiO_3$ and $ZrO_2(Y_2O_3)$ substrates

Re	La	Pr	Nd	Nd	Gd
Substrate	$SrTiO_3$	$SrTiO_3$	$SrTiO_3$	$ZrO_2(Y_2O_3)$	$SrTiO_3$
μ_{exp}, μ_B/FU	1.93	1.15	1.52	1.35	3.28
μ_{th}, μ_B/FU	3.6	3.6, 4.8, 2.4	3.6, 5.4, 1.8	3.6, 5.4, 1.8	3.6, 7.8, 0.6
ΔH, Oe	57	400	230	300	380
K_u, erg/cm^3	10^4	1.9×10^4	1.2×10^4	3.2×10^4	2.7×10^4
T_C', K (H = 100 Oe)	283	160	142	145	75
T_C', K (H = 6 kOe)	308	213	170	210	26
T_C', K (H = 50 kOe)	373	369	260		50
T_f, K (H = 100 Oe)	100	95	73	75	
T_f, K (H = 6 kOe)	50	50	41	38	
T_f, K (H = 50 kOe)	30	33		33	
$T_{\rho(max)}$, K	284	116			

the $M(T)$ curves for this film also differs from that of the curves in Fig. 1a (the latter are also typical of the films of other compounds studied). In weak fields, the $M(T)$ curves of the ZFC samples of GdBaMnO (Fig. 1b) pass through a maximum and then exhibit a minimum followed by a sharp increase. The curves of the FC samples exhibit only an inflection point (instead of the maximum and minimum observed for the ZFC samples) followed by a sharp growth with decreasing temperature. In a field of $H > 6$ kOe, the difference between $M(T)$ curves of the FC and ZFC samples disappears and the magnetization in both cases monotonically decreases with increasing temperature, showing no singularities observed in lower fields. The magnetic moment of the film measured at 5 K in a field of 50 kOe is $3.28\mu_B$/FU.

At the same time, the $M(T)$ curves measured in the fields $H < 6$ kOe resemble those of a ferrimagnet with the compensation point. In this case, the theoretical low-temperature magnetic moment μ_{th} per formula unit at low temperatures must be equal to the difference between the magnetic moments of Gd^{3+} and manganese ions, that is, to $0.6\mu_B$/FU. The experimentally observed magnetic moment at $H = 500$ Oe, where the compensation point is still observed, does not exceed $0.08\mu_B$/FU, which is lower than μ_{th} by a factor of 7.5. This suggests that only a part of the sample (about 13%) is ferrimagnetic, while the remaining part is antiferromagnetic. It should be noted that the ferrimagnetic state with a compensation point was previously observed in the related compound $Gd_{0.67}Ca_{0.33}MnO_3$ [16]. The presence of an AFM phase in the GdBaMnO film is also evidenced, as was noted above, by the magnetization curves. For $H \geq 6$ kOe, the compensation point in the $M(T)$ curves is no longer observed and the magnetization strongly increases, so that the magnetic moment reaches $3.28\mu_B$/FU at 5 K in a field of 50 kOe. This value is significantly greater than that in a sample with complete ferrimagnetic ordering ($0.6\mu_B$/FU), but still markedly smaller that the magnetic moment in the case of complete FM ordering ($7.8\mu_B$/FU. Apparently, the moments of Gd^{3+} and manganese ions exhibit FM ordering, but the FM phase occupies only a part of the sample volume.

The above considerations suggest that the phase with spontaneous magnetization in a GdBaMnO film occurring in a two-phase magnetic state exhibits a magnetic field-induced transition from ferrimagnetic to FM ordering. The experimental $M(T)$ curve observed in a field of 50 kOe is well described by the Langevin function for an ensemble of superparamagnetic clusters with a ferromagnetic moment of $\mu = 22\mu_B$ and a true magnetization of $M_0 = 73.6$ emu/g,

$$M/M_0 = \coth(\mu_H/kT) - kT/\mu H, \quad (1)$$

where M is the magnetization at a given temperature and M_0 is the true magnetization. This is illustrated in Fig. 3, showing a good fit of experimental data (symbols) to a theoretical curve constructed according to relation (1). Assuming that the spins of Gd^{3+} and manganese ions in a cluster are ferromagnetically ordered, the cluster includes approximately three formula units. These results also provide evidence that a two-phase (FM–AFM) magnetic state exists in the GdBaMnO

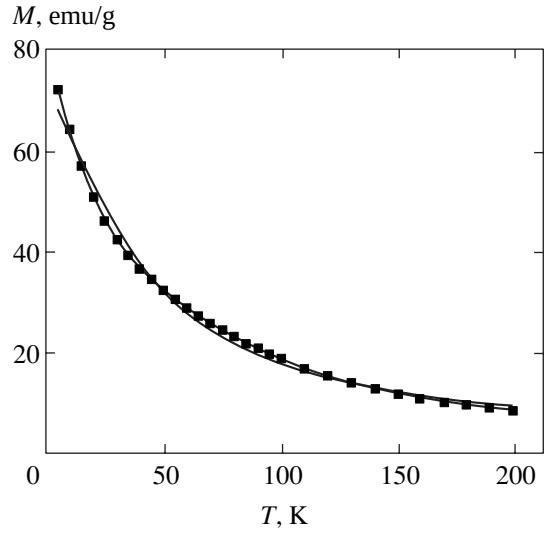

Fig. 3. Temperature dependence of the magnetization $M(T)$ of a thin film of $Gd_{0.6}Ba_{0.4}MnO_3$ on a $SrTiO_3$ substrates measured in a magnetic field of $H = 50$ kOe (black squares). Solid curve shows the Langevin function calculated for an ensemble of superparamagnetic clusters with the magnetic moment $\mu = 22\mu_B$ and the true magnetization $M_0 = 73.6$ emu/g.

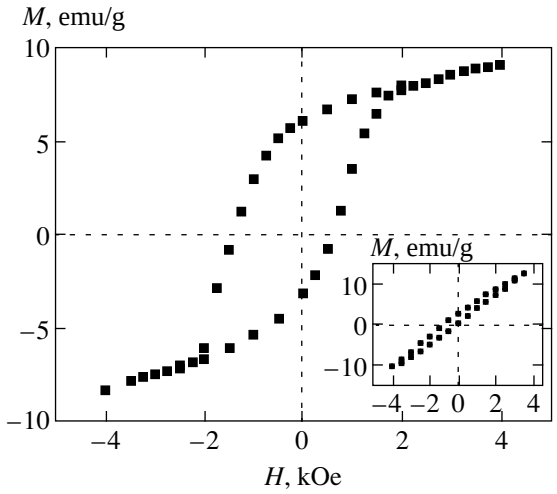

Fig. 4. Magnetization hysteresis loop of a thin film of $Pr_{0.6}Ba_{0.4}MnO_3$ on a $SrTiO_3$ substrates measured at 5 K upon cooling the sample in a magnetic field of 4 kOe. The inset shows the same for a thin film of $Gd_{0.6}Ba_{0.4}MnO_3$ on a $SrTiO_3$ substrate.

film, with the magnetic moment of the FM cluster amounting to approximately $22\mu_B$ in a magnetic field of 50 kOe.

The existence of a two-phase magnetic state in the films under consideration is also confirmed by the fact that the magnetization hysteresis loops of FC samples are shifted along the field axis. This is illustrated in Fig. 4, showing such shifted loops for PrBaMnO and GdBaMnO films. An analogous shift of the hysteresis loop was originally observed in partly oxidized cobalt [17], where it was attributed to the exchange interaction between FM Co particles and their AFM shells of CoO. This phenomenon was called exchange anisotropy. Later, the shift of hysteresis loops in the samples cooled in a weak magnetic field was considered evidence for a spin glass state. Nevertheless, this phenomenon has been explained only for cluster spin glasses and must be absent in a true spin glass containing only randomly oriented spins. Kouvel [18] explained the shift of the magnetization hysteresis loops observed in CuMn and AgMn spin glasses by an inhomogeneous distribution of manganese ions: the regions depleted of Mn were assumed to be ferromagnetic, while the Mn-rich regions featuring exchange interaction were considered antiferromagnetic.

The shift of the magnetization hysteresis loops observed in our samples unambiguously point to the existence of a two-phase (FM–AFM) magnetic state with exchange interact ion between the FM and AFM regions of the films. This shift can be expressed as

$$\Delta H = K_u/M_s, \qquad (2)$$

where K_u is the exchange anisotropy constant and M_s is the saturation magnetization. The exchange anisotropy constants K_u calculated for all films have proved to be on the order of 10^4 erg/cm^3 (see table). Using these values, it is possible to determine exchange integral J characterizing the Mn–O–Mn coupling via the FM–AFM phase boundary, provided that the area of this interface is known. Unfortunately, no such data are available at present.

3.3. Electrical Properties

The resistivity ρ and magnetoresistance $\Delta\rho/\rho = (\rho_H - \rho_{H=0})/\rho_{H=0}$ of LaBaMnO, PrBaMnO, and NdBaMnO films on $SrTiO_3$ substrates were studied at $T > 78$ K and $H \le 8.2$ kOe. The resistances of GdBaMnO films on $SrTiO_3$ and NdBaMnO films on $ZrO_2(Y_2O_3)$ were so high (because of very small film thicknesses) that we failed to measure the ρ values below T_C by the four-point-probe technique. Figure 5 shows the $\rho(T)$ curves of the films studied. Figure 6 presents data on the behavior of $\{\Delta\rho/\rho\}(T)$ for PrBaMnO and LaBaMnO. The magnetoresistance is negative and the $\rho(T)$ and $|\{\Delta\rho/\rho\}|(T)$ curves exhibit maxima. The temperatures of maxima for the latter curves are lower than those for the former ones, which is typical of magnetic semiconductors. The resistivities at maxima were as follows: $\sim 10^{-3}$ Ω cm for PrBaMnO and NdBaMnO, and $\sim 10^{-5}$ Ω cm for LaBaMnO. The magnetoresistance at maximum is very large, reaching 43% for PrBaMnO.

The presence of peaks in the temperature dependences of ρ and $\Delta\rho/\rho$, together with the colossal magnetoresistance, are indicative of the existence of a two-phase magnetic state related to a strong *s–d* exchange in the films studied. Evidently, these manganites are

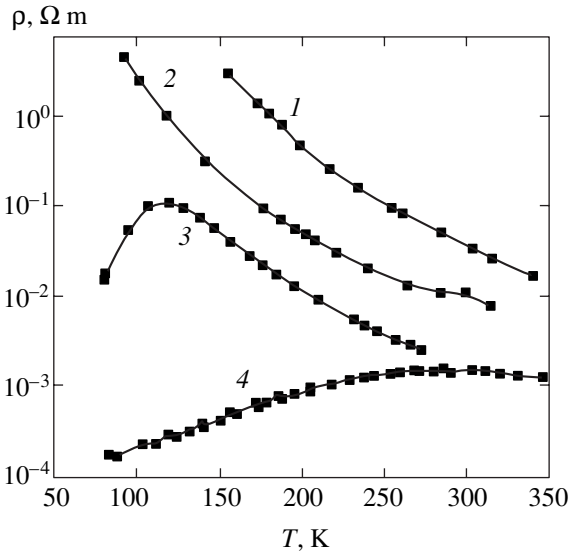

Fig. 5. Temperature dependence of the resistivity ρ of thin films of $Re_{0.6}Ba_{0.4}MnO_3$ with Re = Gd (*1*), Nd (*2*), Pr (*3*), and La (*4*) on $SrTiO_3$ substrates.

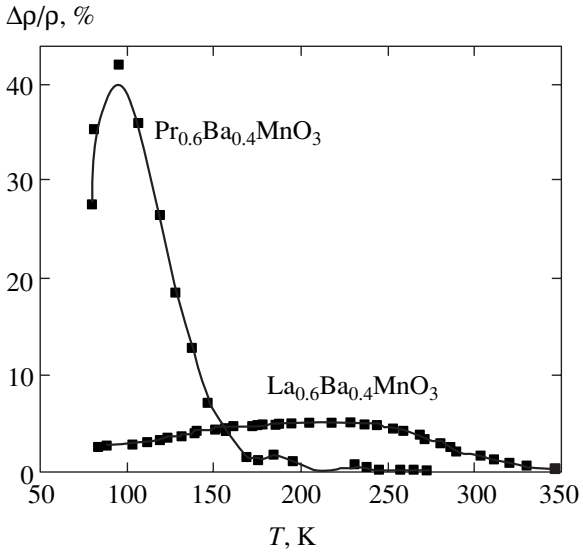

Fig. 6. Temperature dependence of the magnetoresistance of thin $Pr_{0.6}Ba_{0.4}MnO_3$ and $La_{0.6}Ba_{0.4}MnO_3$ films on $SrTiO_3$ substrates measured in a magnetic field of 8.5 kOe.

essentially the AFM semiconductors $LaMnO_3$, $PrMnO_3$, $NdMnO_3$, and $GdMnO_3$ doped with Ba^{2+} ions. Judging by the resistivity ρ, the films of LaBaMnO, PrBaMnO, and NdBaMnO on $SrTiO_3$ substrates occur in a conducting two-phase magnetic state with AFM clusters deprived of charge carriers (holes) are dispersed in a conducting FM matrix. This conducting two-phase magnetic state related to a strong *s–d* exchange has been described in review [19]. This state is characterized by a sharp increase in resistivity in the vicinity of the Curie point. There are two mechanisms by which the impurity magnetism influences the resistance: the scattering of charge carriers (decreasing their mobility) and the formation of a tail (representing localized states) in the conduction band. In the vicinity of the Curie point, charge carriers sharply lose the mobility and exhibit partial localization in the band tail, which explains the appearance of a maximum in $\rho(T)$ near T_C. Under the action of a magnetic field, these charge carriers are delocalized from the band tail and their mobility increases that leads to a colossal magnetoresistance.

As can be seen from Fig. 5, the film of GdBaMnO possesses the maximum resistivity among the samples studied, exceeding 1 Ω m at $T = 150$ K (i.e., significantly above T_C). This suggests the presence of an insulating two-phase magnetic state comprising FM drops, where charge carriers (holes) are concentrated due to an *s–d* exchange energy gain [19], dispersed in an insulating AFM matrix.

4. CONCLUSIONS

The magnetic properties of the films described above resemble those of cluster spin glasses. Indeed, there are differences between magnetizations of the FC and ZFC samples (Fig. 1), the magnetic moment per formula unit at 5 K is strongly reduced (see table), and the $M(T)$ curve shape differs from that of the Brillouin function. On the other hand, there are significant distinctions as well. First, the M values of the FC samples of spin glasses are independent of the temperature at $T < T_f$, provided that the cluster size does not vary with the temperature (this condition is usually valid for spin glasses). In contrast, the magnetizations of our FC samples increase with decreasing temperature. Second, in spin glasses, the behavior of the magnetization in FC and ZFC samples differs only in small fields not exceeding several kOe, whereas our films (except GdBaMnO) exhibit this difference in the entire range of magnetic fields studied (up to 50 kOe). These facts can be explained assuming that a decrease in the temperature leads to an increase in the volume of the FM phase in the two-phase magnetic state. The same factor can account for the difference between the $M(T)$ curve and the Brillouin function. Third, the magnetization isotherms of spin glasses are substantially nonlinear (see Fig. 2), while those of our films are superpositions of a small spontaneous magnetization component and the linear component characteristic of antiferromagnets.

The shift ΔH of the magnetization hysteresis loops of the FC samples along the H axis is unambiguous evidence in favor of the two-phase magnetic state, although this state is also observed in spin glasses (where it is entirely due to the presence of FM and AFM regions and the exchange interaction between them, as was pointed out by Kouvel [18]). Using the field shift ΔH, we have estimated the exchange integral J characterizing the Mn–O–Mn coupling via the FM–AFM phase boundary in some manganites occurring in an insulating two-phase (FM–AFM) magnetic state [20].

It was found that $|J| \sim 10^{-6}$ eV. This value is two orders of magnitude lower than a negative exchange integral between the FM layers in $LaMnO_3$ ($|J_1| = 5.8 \times 10^{-4}$ eV) determined from the neutron scattering data [21]. This result indicates that the presence of a transition layer with tilted spins at the FM–AFM phase boundary is unlikely.

In the material occurring in a two-phase magnetic state, the charge carriers are concentrated in the FM phase and are absent from the AFM phase. For this reason, the topology of the two-phase magnetic state is determined by the Coulomb forces and the interfacial energy. As can be seen from data presented in the table, the FM phase volume in LaBaMnO, PrBaMnO, and NdBaMnO on $SrTiO_3$ substrates occur in a conducting two-phase magnetic state with the FM phase filling the space between insulating AFM spheres. Since the K_u values are on the same order of magnitude for all the films studied (a part of which occur in the conducting two-phase magnetic state) and for the manganites reported in [20] (occurring in an insulating two-phase magnetic state), we may suggest that the area of the FM–AFM phase boundary in the two cases is also comparable. Therefore, the conclusions made in [20] can be expanded to include the films considered above, so that the presence of a transition layer with tilted spins at the FM–AFM phase boundary is unlikely. The GdBaMnO films apparently occur in an insulating two-phase magnetic state. For this material, the magnetic moment of the FM clusters estimated using the Langevin function (1) for $H = 50$ kOe is 22 μ_B, which corresponds to three formula units.

ACKNOWLEDGMENTS

This study was supported by the Russian Foundation for Basic Research, project no. 03-02-16100.

REFERENCES

1. R. von Helmolt, J. Wecker, B. Holzapfel, *et al.*, Phys. Rev. Lett. **71**, 2331 (1993).
2. G. C. Xiong, Q. Li, H. L. Ju, *et al.*, Appl. Phys. Lett. **66**, 1427 (1995).
3. T. Kanki, H. Tanaka, and T. Kawai, Phys. Rev. B **64**, 224418 (2001).
4. H. L. J. Ju, J. Gopalakrishnan, J. L. Peng, *et al.*, Phys. Rev. B **51**, 6143 (1995).
5. G. H. Jonker and J. H. van Santen, Physica (Amsterdam) **16**, 337 (1950); G. H. Jonker, Physica (Amsterdam) **22**, 702 (1956).
6. P. G. Radaelli, M. Marezio, H. Y. Hwang, and C. W. J. Cheong, Solid State Chem. **122**, 444 (1996).
7. A. Barnabe, F. Millange, A. Maignan, *et al.*, Chem. Mater. **10**, 252 (1998).
8. F. Millange, A. Maignan, V. Caignaert, *et al.*, Z. Phys. B **101**, 169 (1996).
9. A. Maignan, Ch. Simon, V. Caignaert, *et al.*, Z. Phys. B **99**, 305 (1996).
10. L. M. Rodriguez-Martinez and J. P. Attfied, Phys. Rev. B **54**, R15622 (1996).
11. Z. Jirak, E. Pollet, A. F. Andersen, *et al.*, Eur. J. Solid State Inorg. Chem. **27**, 421 (1990).
12. C. Roy and R. C. Budhani, J. Appl. Phys. **85**, 3124 (1999).
13. O. Yu. Gorbenko, R. V. Demin, A. R. Kaul', *et al.*, Fiz. Tverd. Tela (St. Petersburg) **40**, 290 (1998) [Phys. Solid State **40**, 263 (1998)].
14. A. I. Abramovich, L. I. Koroleva, A. V. Michurin, *et al.*, Zh. Éksp. Teor. Fiz. **118**, 455 (2000) [JETP **91**, 399 (2000)].
15. S. V. Vonsovskiĭ, *Magnetism: Magnetic Properties of Dia-, Para-, Ferro-, Antiferro-, and Ferrimagnets* (Nauka, Moscow, 1971; Wiley, New York, 1974).
16. G. J. Snyder, C. H. Booth, F. Bridges, *et al.*, Phys. Rev. B **55**, 6453 (1997).
17. W. H. Meiklejohn and C. P. Bean, Phys. Rev. **105**, 904 (1957).
18. J. S. Kouvel, J. Phys. Chem. Solids **21**, 57 (1961); J. Phys. Chem. Solids **24**, 795 (1963).
19. É. L. Nagaev, Usp. Fiz. Nauk **166**, 833 (1996) [Phys. Usp. **39**, 781 (1996)]; Phys. Rep. **346**, 381 (2001).
20. 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)].
21. F. Moussa, M. Hennion, and J. Rodriguez-Carvajal, Phys. Rev. B **54**, 15149 (1996).

Translated by P. Pozdeev